



(12) 发明专利

(10) 授权公告号 CN 101261277 B

(45) 授权公告日 2014. 03. 12

(21) 申请号 200710199696. 5

WO 199108291 A1, 1991. 06. 13, 全文 .

(22) 申请日 2001. 07. 26

US 5616561 A, 1997. 04. 01, 全文 .

(30) 优先权数据

09/628, 112 2000. 07. 27 US

Daniel B. Constam ET AL. Regulation of Bone Morphogenetic Protein Activity by Pro. 《J. CELL. BIOL. 》. 1999, 第 144 卷 (第 5 期), 第 139-149 页 .

(62) 分案原申请数据

01814907. 3 2001. 07. 26

审查员 姚进孝

(73) 专利权人 约翰斯·霍普金斯大学医学院

地址 美国马里兰州

(72) 发明人 李瑟瑾 A·C·麦克弗伦

(74) 专利代理机构 北京纪凯知识产权代理有限公司

公司 11245

代理人 赵蓉民 陆惠中

(51) Int. Cl.

G01N 33/68 (2006. 01)

(56) 对比文件

WO 199421681 A1, 1994. 09. 29, 全文 .

权利要求书2页 说明书111页 附图3页

(54) 发明名称

原肌生长抑制素肽及其使用方法

(57) 摘要

本发明提供了一种基本上纯化的原肌生长抑制素多肽的肽部分。例如,本发明提供了原肌生长抑制素多肽如一种肌生长抑制素功能前区或一种成熟肌生长抑制素肽的蛋白水解片断,以及肌生长抑制素功能前区或肌生长抑制素的功能性肽部分。也提供了一种突变体原肌生长抑制素多肽,它可抵抗蛋白水解的切割,例如切割成为一种肌生长抑制素功能前区和一种成熟的肌生长抑制素肽。本发明也提供了一种编码原肌生长抑制素多肽的肽部分的一种多核苷酸。另外,也提供了与原肌生长抑制素多肽的肽部分特异结合的抗体。本发明进一步提供了鉴定原肌生长抑制素多肽的功能性肽部分的方法。

1. 包含肌生长抑制素功能前区的原肌生长抑制素多肽的基本上纯化的肽部分在制备用于减轻体重、恶病质、肥胖病和 II 型糖尿病的药物中的应用。

2. 根据权利要求 1 所述的应用,其中所述的原肌生长抑制素多肽是脊椎动物原肌生长抑制素多肽。

3. 根据权利要求 1 所述的应用,其中所述的原肌生长抑制素多肽选自:

一种是如 SEQ ID NO :2 中列出的氨基酸序列的人原肌生长抑制素多肽 ;和

一种是如 SEQ ID NO :4 中列出的氨基酸序列的鼠科动物原肌生长抑制素多肽。

4. 根据权利要求 1 所述的应用,其中所述的原肌生长抑制素多肽选自:

一种是如 SEQ ID NO :8 中列出的氨基酸序列的鸡原肌生长抑制素多肽 ;和

一种是如 SEQ ID NO :6 中列出的氨基酸序列的大鼠原肌生长抑制素多肽。

5. 根据权利要求 1 所述的应用,其中所述的原肌生长抑制素多肽选自:

一种是如 SEQ ID NO :10 中列出的氨基酸序列的狒狒原肌生长抑制素多肽 ;

一种是如 SEQ ID NO :12 中列出的氨基酸序列的牛原肌生长抑制素多肽 ;

一种是如 SEQ ID NO :18 中列出的氨基酸序列的火鸡原肌生长抑制素多肽 ;

一种是如 SEQ ID NO :14 中列出的氨基酸序列的猪原肌生长抑制素多肽 ;和

一种是如 SEQ ID NO :16 中列出的氨基酸序列的绵羊原肌生长抑制素多肽。

6. 根据权利要求 1 所述的应用,其中所述的原肌生长抑制素多肽是为如 SEQ ID NO :20 中列出的氨基酸序列的斑马鱼原肌生长抑制素多肽。

7. 根据权利要求 1 所述的应用,其中所述的原肌生长抑制素多肽选自:

一种是如 SEQ ID NO :27 中列出的氨基酸序列的鲑鱼等位基因 1 原肌生长抑制素多肽 ;

和

一种是如 SEQ ID NO :29 中列出的氨基酸序列的鲑鱼等位基因 2 原肌生长抑制素多肽。

8. 根据权利要求 1 所述的应用,其中所述的肽是包含肌生长抑制素功能前区的原肌生长抑制素多肽的蛋白水解片段,所述的肽具有选自以下的一种活性:

a) 肌生长抑制素信号转导激发活性 ;

b) 肌生长抑制素信号转导抑制活性 ;

c) 肌生长抑制素结合活性 ;

d) 原肌生长抑制素结合活性 ;

e) 细胞定位活性 ;和

f) a)、b)、c)、d) 和 e) 的任何组合。

9. 根据权利要求 8 所述的应用,其中所述的肽是由原肌生长抑制素多肽裂解产生的,在其蛋白水解切割位点是如下氨基酸序列:

Arg-Xaa-Xaa-Arg (SEQ ID NO :21)

10. 根据权利要求 9 所述的应用,其中所述的蛋白水解切割位点相当于 SEQ IN NO :2 中列出的 263 至 266 氨基酸残基。

11. 根据权利要求 8 所述的应用,其是为原肌生长抑制素多肽 1 至 262 位氨基酸残基的肌生长抑制素功能前区,所述原肌生长抑制素多肽选自 :SEQ ID NO :2、4、6、8、10、12、14、16、18、20、27 和 29。

12. 一种突变体原肌生长抑制素多肽,包括在蛋白水解切割位点上破坏蛋白水解裂解

作用的突变,其蛋白水解切割位点是如下氨基酸序列:

Arg-Xaa-Xaa-Arg (SEQ ID NO :21),

其中所述的突变包括 SEQ ID NO :21 的 Arg 残基的突变。

13. 一种基本纯化的多核苷酸,其中所述的多核苷酸编码原肌生长抑制素多肽的功能前区。

14. 根据权利要求 13 所述的多核苷酸,其中所述的功能前区进一步包括原肌生长抑制素信号肽。

15. 一种编码权利要求 12 所述的突变体原肌生长抑制素多肽的多核苷酸。

16. 产生含权利要求 13 所述的多核苷酸的转基因非人生物体的方法。

17. 产生含权利要求 15 所述的多核苷酸的转基因非人生物体的方法。

原肌生长抑制素肽及其使用方法

[0001] 本申请为分案申请,原申请的申请号是 01814907.3(PCT/US01/23510),申请日是 2001 年 7 月 26 日,发明名称为“原肌生长抑制素肽及其使用方法”。

技术领域

[0002] 本发明一般地涉及了原肌生长抑制素 (promyostatin) (生长分化因子 -8 ;GDF-8) 的肽部分,和更特别地涉及在细胞中影响肌生长抑制素 (myostatin) 信号转导作用的组合物,和采用这些组合物调节肌生长抑制素在细胞中信号转导的方法。

背景技术

[0003] 在美国,决定要减肥的个人每年花费的时间、精力和费用是惊人的。对于这些人中的许多人,其目的不仅仅是要更好看些,更重要的是要避免与超重相关的一些医学问题。在美国超过半数以上的成人群体被认为是超重的。而且,美国成年男人的 20 至 30%、成年妇女的 30 至 40% 被认为是肥胖症,在贫穷和少数人种中比例是最高的。肥胖症,其定义是超过肥胖平均水平至少大约 20% 以上,其流行在过去几十年中大幅增加,并已经在儿童人群中成为一种主要的问题。所有儿童的 20% 现在被认为是超重的,这个数字意味着在过去 5 年中增加了一倍。

[0004] 肥胖症和与其直接相关的医学问题是全世界发病率和死亡率的主要原因。肥胖症是形成许多病理状况的一个主要危险因素,包括动脉粥样硬化、高血压、心脏病发作、II 型糖尿病、胆囊疾病、和某种癌症,并引起过早死亡。心脏疾病在美国是死亡率的最主要因素,II 型糖尿病在美国影响了超过 1600 万人并是疾病引起死亡的最主要原因之一。

[0005] 80% 以上的 II 型糖尿病发生在肥胖人中。尽管 II 型糖尿病影响所有人种,但在本土美国人、非洲美国人和美籍西班牙人中特别普遍。很明显,II 型糖尿病,过去几乎只发生在年龄超过 40 岁的成人中,现在发生在儿童,在过去 5 年中报道的病例几乎增至三倍。II 型糖尿病也称为非胰岛素依赖型糖尿病,特征是胰岛素对葡萄糖的反应性分泌降低,以及机体对胰岛素作用的抵抗,即便是血循环中的胰岛素水平一般是正常的或升高的。II 型糖尿病可影响许多不同组织和器官的功能,可能引起血管疾病、肾衰、视网膜病和神经病。

[0006] 与肥胖症相关的医学问题相反,严重的体重减轻一般发生在患有某些慢性疾病的患者中,也是医学干预的一种挑战。这种体重减轻的分子基础,就恶病质来说,不是很清楚。但是,很明显,恶病质使这些疾病的处理变得复杂起来,并与患者的不佳预后有关。恶病质的影响在癌症和 AIDS 患者中发生的消耗综合征中表现得很明显。

[0007] 尽管已经做了大量的努力以尝试阐明涉及体重调节中的生物学过程,所提供的结果较实际应用价值更多的是浮夸。例如,苗条蛋白 (leptin) 的发现曾经被欢呼为了解人类脂肪聚积分子基础的重大突破,并许诺使用它治疗肥胖症。在动物中的研究表明苗条蛋白涉及传输调节食欲的内部信号,并认为苗条蛋白可能被用于治疗患有肥胖症的人。但是,使用苗条蛋白治疗肥胖症的进程是缓慢的,迄今苗条蛋白没有达到最初的期望。

[0008] 病态肥胖的治疗目前限制于手术去除部分小肠,从而减少食物 (和卡路里) 吸收

的量。对于中等肥胖,仅有的“治疗方法”是摄入健康饮食和规律地锻炼,这种方法已经证明至多一定程度地有效。因此,需要确定涉及体重调节,包括肌肉发育和脂肪聚积的生物学因子,由此可以开发治疗如肥胖症和恶病质的疾病的方法。本发明满足了这种需求,并提供了其它的优势。

[0009] 发明概述

[0010] 本发明涉及一种基本上纯化的原肌生长抑制素多肽的肽部分。在此原肌生长抑制素多肽的实例是具有如下氨基酸序列的哺乳动物原肌生长抑制素,如人原肌生长抑制素 (SEQ ID NO :2);小鼠原肌生长抑制素 (SEQ ID NO :4);大鼠原肌生长抑制素 (SEQ ID NO :6);狒狒原肌生长抑制素 (SEQ ID NO :10);牛原肌生长抑制素 (SEQ ID NO :12);猪原肌生长抑制素 (SEQ ID NO :14);羊原肌生长抑制素 (SEQ ID NO :16);鸟类原肌生长抑制素,具有如鸡原肌生长抑制素 (SEQ ID NO :8)和火鸡原肌生长抑制素 (SEQ ID NO :18)的氨基酸序列;和鱼类原肌生长抑制素如具有如 SEQ ID NO :20 中列出的氨基酸序列的斑马鱼原肌生长抑制素多肽。另外,原肌生长抑制素多肽的实施例是含有在此公开的鲑鱼等位基因 1 (SEQ ID NO :27) 的部分或鲑鱼等位基因 2 (SEQ ID NO :29) 的部分的多肽。

[0011] 本发明也涉及了生长分化因子 (GDF) 多肽 (原-GDF) 的蛋白水解片段或其功能性肽部分。这样的蛋白水解片段在此的实例是原肌生长抑制素多肽和原 GDF-11 多肽的蛋白水解片段,和这种蛋白水解片段的功能性肽部分。原 GDF 多肽的蛋白水解片段或其功能性肽部分的特性,部分是具有或影响与 GDF 信号转导相关的活性。如此,原肌生长抑制素多肽或其功能性肽部分可具有,例如肌生长抑制素受体结合活性、肌生长抑制素信号转导激发或抑制活性、肌生长抑制素结合活性、或原肌生长抑制素结合活性。在一个实施方案中,蛋白水解片段是原肌生长抑制素多肽在蛋白水解切割识别位点被切割产生的片段,该识别位点具有氨基酸序列 Arg-Xaa-Xaa-Arg (SEQ ID NO :21)。这样的蛋白水解识别位点的实例是 Arg-Ser-Arg-Arg (SEQ ID NO :22) 序列,在 SEQ ID NO :1 (原肌生长抑制素) 中显示为氨基酸残基 263 至 266,或在 SEQ ID NO :25 (人原 GDF-11) 中的氨基酸残基 295 至 298, Arg-Ile-Arg-Arg (SEQ ID NO :23) 序列在 SEQ ID NO :20 中显示为氨基酸残基 263 至 266。

[0012] 在另一个实施方案中,蛋白水解片段是一种 GDF 功能前区,例如一种肌生长抑制素功能前区,它包含了原肌生长抑制素多肽的氨基酸残基 1 至 262,或其功能性肽部分,或一种 GDF-11 功能前区,它包含了原-GDF-11 多肽的氨基酸残基 1 至大约 295,或其功能性肽部分。肌生长抑制素功能前区的实施例是在 SEQ ID NO :4 和 SEQ ID NO :6 中列出的氨基酸残基 1 至大约 263;以及 SEQ ID NO :2、SEQ ID NO :10、SEQ ID NO :12、SEQ ID NO :8、SEQ ID NO :18、SEQ ID NO :14、SEQ ID NO :16、SEQ ID NO :20 中列出的氨基酸残基 1 至大约 262。肌生长抑制素功能前区的功能性肽部分的实施例是可与肌生长抑制素或原肌生长抑制素特异地相互作用的肌生长抑制素功能前区的肽部分。GDF-11 功能前区的实例是 SEQ ID NO :25 的氨基酸残基 1 至大约 295,GDF-11 功能前区的功能性肽部分的实施例是可与成熟 GDF-11 或原 GDF-11 多肽特异地相互作用的 GDF-11 功能前区的肽部分。优选的是,GDF 功能前区的功能性肽部分可减少或抑制相对应的 GDF 或相关的 GDF 激发信号转导的能力。

[0013] 仍然在另一个实施方案中,原 GDF 多肽的蛋白水解片段是一个成熟的 GDF 肽,或其功能性肽部分。因此,蛋白水解片段可能是一个成熟的肌生长抑制素肽 C- 末端,它包含了原肌生长抑制素肽大约 268 至 374 氨基酸,或是一个成熟的 GDF-11 肽 C- 末端,它包含了

原 GDF-11 多肽的大约 299 至 407 氨基酸。成熟的肌生长抑制素肽的实例是在 SEQ ID NO : 4 和 SEQ ID NO :6 中列出的大约 268 至 375 氨基酸残基 ;在 SEQ ID NO :2、SEQ ID NO :10、SEQ ID NO :12、SEQ ID NO :8、SEQ ID NO :18、SEQ ID NO :14、SEQ ID NO :16、SEQ ID NO :20 中列出的大约 267 至 374 氨基酸残基 ;和 SEQ ID NO :27 和 SEQ ID NO :29 的相应氨基酸残基。成熟肌生长抑制素的功能性肽部分的实例是具有肌生长抑制素信号转导激发或抑制活性的肌生长抑制素的肽部分。成熟 GDF-11 肽的实例是 SEQ ID NO :25 的氨基酸残基 299 至 407,成熟 GDF-11 肽的功能性肽部分的实例是具有 GDF-11 信号转导激发或抑制活性的成熟 GDF-11 的肽部分。成熟 GDF 肽的功能性肽部分的活性可以是,例如与其受体特异性相互作用的能力,减少或抑制成熟 GDF 肽与其受体特异性相互作用能力的活性,或任何其他激发或抑制 GDF 信号转导活性的活性。

[0014] 本发明进一步涉及了一种突变体原 GDF 多肽,例如一种突变体原肌生长抑制素多肽或突变体原 GDF-11 多肽,后者含有可在蛋白水解切割位点破坏蛋白水解裂解作用的一个氨基酸突变,该蛋白水解切割位点具有氨基酸序列 Arg-Xaa-Xaa-Arg (SEQ ID NO :21)。因此,突变体原肌生长抑制素或原 GDF-11 多肽具有 SEQ ID NO :21 的 Arg 残基突变,这样突变体多肽不能被切割成为功能前区和成熟的 GDF 肽。优选地,突变体 GDF 多肽与相应的或相关的野生型 GDF 多肽相比,具有显性失活活性。例如,突变体原肌生长抑制素多肽或突变体原 GDF-11 多肽相对肌生长抑制素或 GDF-11 分别具有显性失活活性。

[0015] 本发明也涉及了一种编码原肌生长抑制素或原 GDF-11 多肽的肽部分的多核苷酸。例如,该多核苷酸可编码肌生长抑制素或 GDF-11 功能前区,或其功能性肽部分。本发明进一步涉及了可与原肌生长抑制素多肽的肽部分,例如,与肌生长抑制素功能前区或其功能性肽部分特异结合的抗体,以及可特异地与原 GDF-11 多肽的肽部分结合的抗体。也提供了一种试剂盒,其中含有原肌生长抑制素或原 GDF-11 多肽的肽部分,或突变体原肌生长抑制素或原 GDF-11 多肽,或编码原肌生长抑制素或原 GDF-11 多肽的肽部分的多核苷酸,或可特异结合如此一种肽部分或的抗体,或者其混合物。

[0016] 另外,本发明涉及了一种可鉴定肌生长抑制素功能前区或 GDF-11 功能前区的功能性肽部分的方法,这些肽部分可特异地与肌生长抑制素肽、GDF-11 肽或两者相互作用。本发明的一种方法可通过,例如检测肌生长抑制素功能前区或 GDF-11 功能前区的肽部分与肌生长抑制素特异性相互作用的能力来实施 ;并检测肽部分与肌生长抑制素的特异性相互作用。在一个实施方案中,本发明的一种方法可采用计算机系统来实施,其中要检测肌生长抑制素功能前区或 GDF-11 功能前区的虚拟肽部分与虚拟的肌生长抑制素肽特异相互作用的能力。在另一个实施方案中,该方法通过将功能前区的肽部分与肌生长抑制素肽在适合功能前区与肌生长抑制素肽特异地相互作用的条件下进行接触而实施。

[0017] 附图简述

[0018] 图 1 显示了小鼠原肌生长抑制素 (SEQ ID NO :4) ;大鼠原肌生长抑制素 (SEQ ID NO :6) ;人原肌生长抑制素 (SEQ ID NO :2) ;狓狓原肌生长抑制素 (SEQ ID NO :10) ;牛原肌生长抑制素 (SEQ ID NO :12) ;猪原肌生长抑制素 (SEQ ID NO :14) ;羊原肌生长抑制素 (SEQ ID NO :16) ;鸡原肌生长抑制素 (SEQ ID NO :8) ;火鸡原肌生长抑制素 (SEQ ID NO :18) ;和斑马鱼原肌生长抑制素 (SEQ ID NO :20) 的氨基酸序列。氨基酸相对人原肌生长抑制素 (SEQ ID NO :2) 进行编号。虚线表示引导最大同源性的间隙。序列中相同的残基被划

上阴影。

[0019] 图2显示了小鼠原肌生长抑制素 (SEQ ID NO :4) 和斑马鱼原肌生长抑制素 (SEQ ID NO :20) 的氨基酸序列, 鲑鱼等位基因 1 原肌生长抑制素 (SEQ ID NO :27 ;“鲑鱼 1”) 和鲑鱼等位基因 2 原肌生长抑制素 (SEQ ID NO :29 ;“鲑鱼 2”) 的部分氨基酸序列。相应的人原肌生长抑制素的氨基酸部位标记在每排的左侧 (比较图 1 ;鲑鱼 1 的第一个氨基酸对应于人原肌生长抑制素 218 ;鲑鱼 2 的第一个氨基酸对应于人原肌生长抑制素 239)。虚线表示引导最大同源性的间隙。相对的氨基酸位置, 包括间隙, 沿着每排的上方标记。序列中相同的残基被划上阴影。

[0020] 发明详述

[0021] 本发明提供了一种基本上纯的原肌生长抑制素 (promyostatin) 多肽的肽部分。原肌生长抑制素, 以前被称为生长分化因子 -8 (GDF-8), 包括一个氨基末端的功能前区 (prodomain) 和一个 C-末端的成熟的肌生长抑制素 (myostatin) 肽 (见美国专利 5,827,733 号)。在从原肌生长抑制素中分割后, 肌生长抑制素的活性由成熟的肌生长抑制素肽实现。因此, 原肌生长抑制素是一个前体多肽, 可被蛋白水解分割产生活性的肌生长抑制素。如在此所公开的, 肌生长抑制素的功能前区可抑制肌生长抑制素的活性, GDF-11 的活性, 或两者的活性。

[0022] 本发明也提供了一个原 -GDF-11 多肽的基本上纯的肽部分。原 -GDF-11, 以前一般被称为 GDF-11, 包括一个氨基酸末端的功能前区和一个 C-末端的成熟 GDF-11 肽 (见国际公开第 WO 98/35019 号, 在此引入作为参考文献)。在从原 -GDF-11 中分割后, GDF-11 的活性由成熟的 GDF-11 肽实现。因此, 原 -GDF-11 像原肌生长抑制素一样, 是一个前体多肽, 可被蛋白水解分割产生活性的 GDF-11。如在此所公开的, GDF-11 的功能前区可抑制 GDF-11 活性, 肌生长抑制素活性, 或抑制两者的活性。

[0023] 原肌生长抑制素和原 -GDF-11 是转化生长因子 - θ (TGF- θ) 超家族的成员, 该超家族由在许多细胞类型中控制增殖、分化和其它功能的多功能多肽组成。TGF- θ 超家族, 包括一组可在胚胎发育过程中对分化过程影响广泛的结构相关蛋白, 包括例如, 缪氏抑制物质 (MIS), 它是正常雄性发育过程中是必需的 (Behringer 等人, Nature 345 :167, 1990), 果蝇 decapentaplegic (DPP) 基因产物, 它是背腹轴形成和成虫盘形态发生所必需的 (Padgett 等人, Nature 325 :81-84, 1987), 爪蟾 Vg-1 基因产物, 它位于卵的植物极上 (Weeks 等人, Cell 51 :861-867, 1987), 激活素 (activin) (Mason 等人, Biochem. Biophys. Res. Comm. 135 :957-964, 1986), 它可诱导爪蟾胚的中胚层和前端结构的形成 (Thomsen 等人, Cell 63 :485, 1990), 和骨形态发生蛋白 (BMP, 成骨素, OP-1), 它可重新诱导软骨和骨的形成 (Sampath 等人, J. Biol. Chem. 265 :13198, 1990)。TGF- θ 家族成员可影响许多分化过程, 包括脂肪形成、肌形成、软骨发生、血细胞生成、和上皮细胞分化 (Massague, Cell 49 :437, 1987 ;Massague, Ann. Rev. Biochem. 67 :753-791, 1998 ;每一篇在此都引入作为参考文献)。

[0024] 许多 TGF- θ 家族成员对其它肽生长因子具有调节作用 (正调节作用或负调节作用)。特别的是, TGF- θ 超家族的某些成员具有与神经系统功能相关的表达方式或活性。例如, 抑制素 (inhibitin) 和激活素 (activin) 都在脑中表达 (Meunier 等人, Proc. Natl. Acad. Sci., USA 85 :247, 1988 ;Sawchenko 等人, Nature 344 :615, 1988), 激活素可

作为一个神经细胞生存分子而起作用 (Schubert 等人, Nature 344 :868, 1990)。另一个家族成员, 生长分化因子 -1 (GDF-1), 在其表达方式上是神经系统特异性的 (Lee, Proc. Natl. Acad. Sci., USA 88 :4250, 1991), 和其它家族成员如 Vgr-1 (Lyons 等人, Proc. Natl. Acad. Sci., USA 86 :4554, 1989 ; Jones 等人, Development 111 :531, 1991)、OP-1 (Ozkaynak 等人, J. Biol. Chem. 267 :25220, 1992)、和 BMP-4 (Jones 等人, Development 111 :531, 1991), 也在神经系统中表达。因为骨骼肌可产生一种因子或多种因子, 促进运动神经元的存活 (Brown, Trends Neurosci. 7 :10, 1984), 肌生长抑制素 (GDF-8) 和 GDF-11 在肌肉中的表达表明肌生长抑制素和 GDF-11 是神经元的营养因子。像这样, 调节肌生长抑制素, GDF-11 或两者活性的方法可用于治疗神经变性疾病, 如肌萎缩侧索硬化症或肌营养不良, 或在移植前在培养物中保存细胞或组织。

[0025] TGF- θ 家族的蛋白可合成为大的前体蛋白, 它随后经历蛋白水解, 从 C-末端切除一簇大约 110 至 140 个氨基酸碱性残基, 形成一个功能前区肽和一个 C-末端的成熟肽。这个家族成员蛋白的 C-末端成熟肽在结构上是相关的, 不同家族成员根据其同源性可分成不同的亚组。尽管在特殊的亚组内的同源性范围为 70% 至 90% 的氨基酸序列同一性, 在亚组之间的同源性明显更低, 一般在 20% 至 50%。在每个成员中, 有活性的类型似乎是一个二硫键联接的 C-末端肽片断的二聚体。

[0026] 原肌生长抑制素和原 -GDF-11 多肽在哺乳动物、鸟类和鱼类中已经被鉴定, 肌生长抑制素在许多其它的物种, 包括脊椎动物和无脊椎动物中是有活性的。在胚胎发育过程中和在成体动物中, 例如, 肌生长抑制素特异的表达在肌源性谱系的细胞中 (McPherron 等人, Nature 387 :83-90, 1997, 在此引入作为参考文献)。在早期胚胎发生过程中, 肌生长抑制素表达在发育体节的生肌节腔隙中。在晚期胚胎期和成体动物中, 肌生长抑制素广泛的表达在骨骼肌组织中, 尽管在肌肉之间的表达水平变化非常大。肌生长抑制素的表达也可在脂肪组织中检测到, 尽管比在肌肉中的水平更低。类似地, GDF-11 在骨骼肌和脂肪组织, 以及成人胸腺、脾脏和子宫中表达, 也在不同发育时期的脑内表达。

[0027] 来自多个物种的原肌生长抑制素多肽具有实质上相同的序列, 人、小鼠、大鼠和鸡的成熟肌生长抑制素 C-末端序列的氨基酸序列 100% 相同 (见图 1)。在此原肌生长抑制素多肽的实例 (见图 1) 为人的原肌生长抑制素 (SEQ ID NO :2) ; 鼠肌生长抑制素 (SEQ ID NO :4) ; 大鼠肌生长抑制素 (SEQ ID NO :6) ; 狒狒原肌生长抑制素 (SEQ ID NO :10) ; 牛原肌生长抑制素 (SEQ ID NO :12) ; 猪原肌生长抑制素 (SEQ ID NO :14) ; 羊原肌生长抑制素 (SEQ ID NO :16) ; 鸡原肌生长抑制素 (SEQ ID NO :8) ; 火鸡原肌生长抑制素 (SEQ ID NO :18) ; 和斑马鱼原肌生长抑制素 (SEQ ID NO :20) 的氨基酸序列。在此原肌生长抑制素多肽的实例也可以包括部分的鲑鱼等位基因 1 (SEQ ID NO :27 ; “鲑鱼 1”) 和鲑鱼等位基因 2 (SEQ ID NO :29 ; “鲑鱼 2”, 见图 2) 的一种多肽。在此公开的编码这些原肌生长抑制素多肽的核酸分子分别是 SEQ ID NO :1, 3, 5, 9, 11, 13, 15, 7, 17, 19, 26 和 28 (也见, McPherron 和 Lee, Proc. Natl. Acad. Sci., USA 94 :12457, 1997, 在此引入作为参考文献)。在此一个原 -GDF-11 多肽的实例是人的原 -GDF-11 (SEQ ID NO :25), 它由 SEQ ID NO :24 编码。

[0028] 考虑到在原肌生长抑制素多肽中的广泛保守性, 特别是在如同具有像人和鱼之间的差别性的物种中, 从任何物种中获得编码肌生长抑制素的多核苷酸和在任何物种中识别原肌生长抑制素或肌生长抑制素的表达将是很普通的, 其中这些多核苷酸包括鲑鱼 1 和鲑

鱼 2 序列的剩余部分。特别的是,成熟的肌生长抑制素序列与 TGF- θ 超家族的其它成员具有明显的同源性,肌生长抑制素含有大多数其它家族成员内和其它物种间高度保守的残基。而且,像 TGF- θ 和抑制素 θ 一样,除了实际上存在于所有其它家族成员中的 7 个半胱氨酸残基以外,肌生长抑制素含有额外的一对半胱氨酸残基。肌生长抑制素与 Vgr-1 (45% 序列同一性) 是最同源的。像 TGF- θ 超家族的其它成员,肌生长抑制素可被合成为一个大的前体原肌生长抑制素多肽,它可被蛋白水解切割成一个活性的肌生长抑制素肽。

[0029] 编码多种生物体的原肌生长抑制素多肽的多核苷酸可用熟知的程序和算法基于与公开序列的同一性(或同源性)来识别。同源性或同一性经常采用序列分析软件来测定,这些软件如遗传学计算机集团(威斯康星大学生物技术学中心,大学路 1710 号, Madison, WI53705)的序列分析软件包。这样的软件可通过对多种删除,替代和其它修饰作用赋予不同程度的同源性值,从而匹配相似的序列。术语“同源性”和“同一性”,当在此两个或更多个核酸或多肽序列的情况下使用时,是指当被比较和被排列以最大的对应于比较窗或指定区域时,采用任何数量的序列比较算法测量或通过人工排列和目测时,两个或更多个相同的序列或子序列,或具有特殊比例的氨基酸残基或相同的核苷酸。

[0030] 为了比较序列,一般一个序列可作为一个参考序列,用它可以检测其它序列。当采用一种序列比较算法时,检测和参考序列输入计算机,如果需要,指定子序列的坐标,并指定序列算法程序的参数。采用缺省的程序参数或指定可选的参数。然后序列比较算法基于程序参数对检测序列与参考序列计算等同的序列百分比。

[0031] 术语“比较窗”在此广泛的使用以涵盖一定数目连续位点的任何一个的一个片段,该数目例如,大约 20 至 600 个位点,例如,氨基酸或核苷酸位点普遍是大约 50 至大约 200 个位点,更普遍的是大约 100 至大约 150 个位点,其中在两个序列被最佳的排列之后,一个序列可与相同数目连续位点的一个参考序列进行比较。进行比较的序列排列的方法在本领域中是为人熟知的。进行比较的最佳序列排列可通过,例如 Smith 和 Waterman 局部同源性算法 (Adv. Appl. Math. 2 :482,1981), Needleman 和 Wunsch 的同源性排列算法 (J. Mol. Biol. 48 :443,1970), Person 和 Lipman 的相似法检索 (Proc. Natl. Acad. Sci., USA 85 :2444,1988),上述每种方法在此均引入作为参考文献;这些算法的计算机化应用(威斯康星遗传学软件包中的 GAP, BESTFIT, FASTA, 和 TFASTA, 遗传学计算机集团,575 Science Dr., Madison, WI);或人工排列和目测。确定同源性或同一性的其它算法包括,例如,除 BLAST 程序以外(国立生物学信息中心的基本局部排列检索工具),有 ALIGN, AMAS(多次排列序列分析), AMPS(蛋白质多次序列排列), ASSET(排列片段统计估算工具), BANDS, BESTSCOR, BIOSCAN(生物学序列比较分析结点), BLIMPS(Blocks IMPROVED 检索器), FASTA, Intervals&Points, BMB, CLUSTAL V, CLUSTAL W, CONSENSUS, LCONSENSUS, WCONSENSUS, Smith-Waterman 算法, DARWIN, Las Vegas 算法, FNAT(强制核苷酸排列工具), 框架排列, 框架检索, DYNAMIC, FILTER, FSAP(Fristensky 序列分析包), GAP(球形排列程序), GENAL, GIBBS, GenQuest, ISSC(敏感性序列比较), LALIGN(局部序列排列), LCP(局部容量程序), MACAW(多次排列构建和分析工作台), MAP(多次排列程序), MBLKP, MBLKN, PIMA(图形引导的多次序列排列), SAGA(遗传算法的序列排列)和 WHAT-IF。这样的排列程序也可用于筛选基因组数据库以识别具有实质上相同序列的多核苷酸序列。

[0032] 许多基因组数据库可用于比较,包括,例如作为人类基因组测序项目(J. Roach,

[http://weber.u.Washington.edu/~roach/human genome progress 2.html](http://weber.u.Washington.edu/~roach/human%20genome%20progress%20.html)) 一部分的人类基因组的主要部分。另外,至少有 21 个基因组已经完全被测序,包括,例如,生殖器支原体 (*M. genitalium*), 甲烷球菌 (*M. jannaschii*), 流感嗜血杆菌 (*H. influenzae*), 大肠杆菌, 酵母 (酿酒酵母), 和果蝇 (*D. melanogaster*)。在如鼠,秀丽新小杆线虫 (*C. elegans*) 和拟南芥 (*Arabidopsis sp.*) 等生物体模型基因组的测序中已经获得了明显的进展。几个含有一些功能性信息说明的基因组信息数据库由不同的机构维护,也经互联网上获得,例如, <http://www.tigr.org/tdb>; <http://www.genetics.wisc.edu>; <http://genome-www.stanford.edu/~ball>; <http://hiv-web.lanl.gov>; <http://www.ncbi.nlm.nih.gov>; <http://www.ebi.ac.uk>; <http://Pasteur.fr/other/biology>; 和 <http://www.genome.wi.mit.edu>。

[0033] 有用的算法的一个实施方案是 BLAST 和 BLAST2.0 算法,由 Altschul 等人所描述 (Nucleic Acids Res. 25 :3389-3402, 1977; J. Mol. Biol. 215 :403-410, 1990, 每一篇文献在此都引入作为参考文献)。执行 BLAST 分析的软件可公开地从国家生物技术学信息中心获得 (<http://www.ncbi.nlm.nih.gov>)。这种算法首先通过在查询序列中识别短的长度为 W 的字符串,识别出高分值序列配对 (HSP), 当在一个数据库序列中与相同长度的字符串排列时,该序列配对匹配或满足一些阳性意义的阈值评分 T 。 T 被称为是邻近的字符串分值的阈值 (Altschul 等人, 见前, 1977, 1990)。这些最初的邻近字符串点击可作为启动检索寻找含有它们的更长的 HSP 的开端。字符串点击可沿着每个序列两个方向延伸,只要累计的排列分值增加。对核苷酸序列的累计分值的计算可以采用参数 M (对一对相匹配的残基的奖励分值; 一般 > 0)。对于氨基酸序列,采用一种评分矩阵来计算累计的分值。字符串点击在每个方向的延伸当出现以下情况时停止: 累计排列分值从达到的最大值下降至数量 X 时; 由于一个或多个阴性的残基排列分值的积累,导致累计的分值到达 0 或之下时; 或达到每个序列的末端时。BLAST 算法参数 W , T 和 X 决定了排列的敏感性和速度。BLASTN 程序 (对于核苷酸序列) 采用的缺省字符串长度 (W) 为 11, 期望值 (E) 为 10, $M = 5$, $N = 4$, 并比较两条链。对于氨基酸序列, BLASTP 程序采用的缺省字符串长度为 3, 期望值 (E) 为 10, BLOSUM62 评分矩阵 (见 Henikoff 和 Henikoff, Proc. Natl. Acad. Sci., USA 89 :10915, 1989) 排列 (B) 为 50, 期望值 (E) 为 10, $M = 5$, $N = -4$, 并比较两条链。

[0034] BLAST 算法也在两个序列之间进行相似性的统计学分析 (见, 例如, Karlin 和 Altschul, Proc. Natl. Acad. Sci. USA 90 :5873, 1993, 在此引入作为参考)。由 BLAST 算法提供了一种相似性的测量方法是最小和数概率 ($P(N)$), 它提供了一种概率的指示法, 根据它在两个核苷酸或氨基酸序列之间的匹配将偶然发生。例如, 如果在比较检测核酸与参考核酸之间的最小和数概率少于大约 0.2, 更优选少于大约 0.01, 最优选少于大约 0.001 时, 该核酸被认为与参考序列是相似的。

[0035] 在一个实施方案中, 蛋白质和核酸序列同源性是采用基本局部排列检索工具 (“BLAST”) 进行评价的。尤其是, 5 个特殊的 BLAST 程序用来进行下列的工作:

[0036] (1) BLASTP 和 BLAST3 将一种氨基酸查询序列与一个蛋白质序列数据库进行比较;

[0037] (2) BLASTN 将一种核苷酸查询序列与一个核苷酸序列数据库进行比较;

[0038] (3) BLASTX 将一个查询核苷酸序列 (两条链) 的 6 框的概念翻译产物与一个蛋白质序列数据库进行比较;

[0039] (4) TBLASTN 将一个查询蛋白质序列与一个所有六个可读框（两条链）翻译的核苷酸序列数据库进行比较；和

[0040] (5) TBLASTX 将一个核苷酸查询序列的 6 框翻译作用与一个核苷酸序列数据库的 6 框翻译作用进行比较。

[0041] BLAST 程序通过识别相似的片段来鉴定同源序列，在此相似的片段是指在查询氨基酸或核酸序列和一种检测序列之间的“高分值的片段对”，该检测序列优选从一种蛋白质或核酸序列数据库中获得。高分值的片段对优选采用一种评分矩阵的方法进行鉴定（即排列），在本领域中许多这样的矩阵是已知的。优选地，使用的评分矩阵是 BLOSUM62 矩阵（Gonnet 等人，*Science* 256:1443-1445, 1992；Henikoff 和 Henikoff，*Proteins* 17:49-61, 1993，两文在此均引入作为参考）。不太优选的是，也可使用 PAM 或 PAM250 矩阵（Schwartz 和 Dayhoff，主编，“检测距离关系的矩阵：蛋白质序列和结构图集”（*Matrices for Detecting Distance Relationships: Atlas of Protein Sequence and Structure*）（Washington，国家生物医学研究基金会 1978））。BLAST 程序可通过美国国家医学图书馆访问，例如，在 www.ncbi.nlm.nih.gov。

[0042] 在上述算法中使用的参数可根据研究的序列长度和同源性程度进行调整。在一些实施方案中，没有用户指令时，参数可以是算法所提供的缺省参数。

[0043] 编码一个原肌生长抑制素的多核苷酸可来自任何生物体，包括例如，小鼠、大鼠、牛、猪、人、鸡、火鸡、斑马鱼、鲑鱼、长须鲸、其它水生生物和其它物种。水生生物的实例包括那些属于鱼塘鱼类的生物，如鳟鱼、红点鲑、香鱼、鲤鱼、鲫鱼、金鱼、昏白鱼、银鱼、鳊鱼、海鳗、沙丁鱼、飞鱼、巴西刺鲈、鲷、鸚鵡鲈鱼、新西兰真鲷、鲭、白腹鲭、鲔鱼、金枪鱼、鲣鱼、鲱鱼、岩鱼、平鱼（fluke）、鳎鱼、比目鱼、河豚、鲑鱼；那些属于头足类的生物，如鱿鱼、墨鱼、章鱼；那些属于斧足类的生物，如蛤（如，硬壳蛤、蛤仔、圆蛤、海蛤蜊、软壳蛤）；乌蛤、蛤贝、海螺；扇贝（如，海扇贝、海湾扇贝、calloo）；海螺、蜗牛、海参、赤贝；牡蛎（如贞洁巨牡蛎（*C. virginica*），海湾牡蛎，新西兰牡蛎，太平洋牡蛎）；那些属于腹足类的生物，如螺旋贝、鲍鱼（如绿鲍鱼、粉红鲍鱼、红鲍鱼）；和那些属于甲壳类的生物，如龙虾，包括但不限于刺蜥、岩蜥和美洲蜥；对虾；河虾，包括但不限于罗氏对虾（*M. rosenbergii*），*P. styllrolls*，印度对虾（*P. indicus*），日本对虾（*P. japonicus*），斑节对虾（*P. monodon*），*P. vannemel*，*M. ensis*，*S. melantho*，*N. norvegicus*，冷水河虾；蟹，包括但不限于蓝蟹、白嘴蟹、石蟹、王蟹、雪花蟹、雪蟹、褐蟹、邓杰内斯蟹、约拿蟹、红树蟹、软壳蟹；虾蛄、磷虾、河螯虾（*langostinos*）；小龙虾 / 龙虾，包括但不限于蓝色的、栗色的、红爪的、红沼泽的、软壳的、白色的；环节动物门；脊索动物门，包括但不限于爬虫类如美洲鳄鱼和龟；两栖类生物，包括蛙；和棘皮类，包括但不限于海胆。

[0044] 本发明提供一个原肌生长抑制素多肽基本上纯的肽部分，和一个原 -GDF-11 多肽的基本上纯的肽部分。如在此所使用的，提到“原 -GDF”，例如原肌生长抑制素或原 -GDF-11，它意味着全长的多肽，包括氨基末端的功能前区和羧基末端的生物学活性 GDF 肽。另外，功能前区包括一个信号肽（引导序列），它由功能前区氨基末端的开始 15 至 30 个氨基酸组成。信号肽可从全长的原 -GDF 多肽中切割出来，后者可进一步在一个 Arg-Xaa-Xaa-Arg (SEQ ID NO:21) 蛋白水解切割位点上被切断。

[0045] 在此所谈氨基酸残基的参照物是根据图 1 和 2（也见序列列表）中显示的全长

原 -GDF 多肽而言。也应该认识到在此所谈特定多肽是开始或者终止在“大约”一个特殊的氨基酸残基处。在此文中使用术语“大约”是因为认识到在蛋白水解切割识别位点处或邻近部位,一个特殊的蛋白酶可切断原 -GDF 多肽,或从识别位点切下一个或几个氨基酸。如此,例如,谈到一个具有 SEQ ID NO :4 的大约 1 至 263 氨基酸残基的序列的肌生长抑制素功能前区时,将包括原肌生长抑制素的一个氨基末端肽部分,包括信号肽,并具有在氨基酸残基 257 至氨基酸残基 269,优选在氨基酸残基 260 至氨基酸残基 266 终止的一个羧基末端。

[0046] 类似地,信号肽可从一个原 -GDF 多肽的大约 15 至 30 氨基酸残基的任何位置上被切断,例如,在残基 15, 20, 25 或 30 上切断,不影响剩余功能前区的功能。因此,为方便起见,在此一般谈到一个原 -GDF 多肽的肽部分时,信号肽已经从中被切下来,其起始在大约氨基酸残基 20 处。但是,将要认识到的是信号肽的切割可以在一个原 -GDF 多肽的大约开始 15 至 30 个氨基酸之中的任意氨基酸位置处。如此,例如,谈到一个具有 SEQ ID NO :4 的大约 20 至 263 氨基酸残基的序列的肌生长抑制素功能前区时,将包括原肌生长抑制素的一个肽部分,它缺少原肌生长抑制素的大约开始 15 至 30 个氨基酸,包含信号肽,并具有在氨基酸残基 257 至氨基酸残基 269,优选在氨基酸残基 260 至氨基酸残基 266 终止的一个羧基末端。

[0047] 一般来说,在此提到一个原 -GDF 多肽或一个 GDF 功能前区时,其开始在大约氨基酸 1 处。但考虑到上述的公开内容,将认识到这样的缺少信号肽的原 GDF 多肽或 GDF 功能前区也包括在本发明中。在此方面进一步应认识到在本发明的一个肽中存在或缺少一个信号肽可影响,例如细胞的区室,一种肽,例如原肌生长抑制素功能前区将穿过它,该肽将最终定位在此,包括是否该肽可从细胞中被分泌出来。因此,本发明进一步提供了一个原 -GDF 多肽的基本上纯的信号肽部分。如在此所公开的,这样一种信号肽可用来靶向一种药物,特别是一种肽药物,至相同的细胞区室中,如同该信号肽所起源的天然 GDF 一样。

[0048] 术语“肽”或“肽部分”在此广泛用来指通过肽键连接的两个或更多个的氨基酸。术语“片段”或“蛋白水解片段”在此也用来指一种可通过在一个多肽上进行蛋白水解反应产生的产物,即一种通过切断多肽中的肽键所产生的肽。尽管术语“蛋白水解片段”一般在此用来指一种可通过蛋白水解反应产生的一种多肽,应该认识到该片段不仅不必必须通过蛋白水解反应来产生,而且也可采用化学合成或重组 DNA 技术的方法来产生,如在下面所详述的,这些方法可产生一种合成肽,与蛋白水解片段是相同的。考虑到原肌生长抑制素和其它 TGF- θ 超家族成员之间已公开的同源性,将要认识到本发明所述肽的特征部分地在于它不存在于以前公开的该超家族的成员之中。采用上述的计算机算法可很容易地确定,是否一个原肌生长抑制素或原 -GDF-11 多肽的肽部分存在于以前所公开的 TGF- θ 超家族成员中。

[0049] 一般来说,本发明所述肽含有至少 6 个氨基酸,一般含有大约 10 个氨基酸,可含有 15 个或更多的氨基酸,特别是 20 个或更多的氨基酸。应该认识到术语“肽”在此并不用来说明组成分子的氨基酸的特殊大小或数目,本发明所述肽可含有可达几个氨基酸残基或更多的残基数目。例如,全长的成熟 C-末端肌生长抑制素肽含有超过 100 个氨基酸,一个全长的功能前区肽可含有超过 260 个氨基酸。

[0050] 如在此所使用的,术语“基本上纯化的”或“基本上纯的”或“分离的”是指所提到的分子,例如一种肽或一种多核苷酸,其形式相对地不含有蛋白质、核酸、脂类、碳水化合物

或其它与其天然相关的物质。一般来说,基本上纯净的肽、多核苷酸、或样品中其它分子组成至少占 20%,一般至少占样品的大约 50%,通常至少占样品的大约 80%,特殊的占样品的大约 90%或 95%或更多。判断本发明所述肽或所述多核苷酸基本上是纯净的,可以采用熟知的方法来进行,例如通过进行电泳,将相对分离的条带鉴定为特殊的分子。一种基本上纯的多核苷酸,例如可通过克隆多核苷酸或化学或酶合成的方法获得。一种基本上纯净的肽,例如可通过一种化学合成的方法,或采用蛋白质纯化的方法来获得,然后进行蛋白质水解,如果需要,通过色谱或电泳的方法来进行进一步纯化。

[0051] 本发明的一种肽可分别通过与原肌生长抑制素或原 -GDF-11 序列的比较,确定该肽的氨基酸序列包含在原肌生长抑制素或原 -GDF-11 多肽序列中来鉴定。但应该认识到,本发明的一种肽不需要与原肌生长抑制素或原 -GDF-11 的相应氨基酸序列相等。因此,例如,本发明的一种肽可对应原肌生长抑制素多肽的氨基酸序列,但可与天然发生序列有所不同,例如在对应的 L-氨基酸的位置含有一个或多个 D-氨基酸;或含有一个或多个氨基酸类似物,例如,已经发生衍生化作用或者是在其活性侧链上具有被修饰的氨基酸。类似地,在肽的一个或多个肽键上可以被修饰。另外,在氨基末端或羧基末端或两者上的活性基团可以被修饰。这样的肽被修饰后,例如,可改善对蛋白酶、氧化剂或肽在生物学环境中可能遇到的其它活性物质的稳定性,并因此在本发明所述方法的实施过程中特别有用。当然肽可以被修饰使其在生物学环境中降低稳定性,以便减少该肽在环境中具有活性的时间。

[0052] 与原肌生长抑制素或原 -GDF-11 多肽中的相应序列相比较,本发明的一种肽的序列也可通过插入一种保守的氨基酸替代肽中的一个或几个氨基酸而进行修饰。保守的氨基酸替代作用包括用具有相对相同化学特性的另一个氨基酸残基替代一个氨基酸残基,例如一个疏水残基如异亮氨酸、缬氨酸、亮氨酸或蛋氨酸替代另一个,或一个极性残基替代另一个,例如精氨酸替代赖氨酸;或谷氨酸替代天冬氨酸;或谷氨酸替代天冬酰胺;或类似的情况。细查图 1,原肌生长抑制素多肽可被修饰位置的实例是很明显的,其中显示了实质上不影响原肌生长抑制素或肌生长抑制素活性的肌生长抑制素功能前区和成熟肌生长抑制素肽中的多个氨基酸差异。

[0053] 本发明也提供了一种生长分化因子 (GDF) 多肽 (原 -GDF 多肽) 或其功能性肽部分的一种基本上纯的蛋白水解片段。原 -GDF 多肽的蛋白水解片段在此的实例是原肌生长抑制素多肽的蛋白水解片段和原 -GDF-11 多肽的蛋白水解片段。如在此所公开的,与原 -GDF 蛋白水解片段相同的原 -GDF 多肽的肽部分可通过一种化学的方法或一种重组 DNA 的方法来产生。考虑到此处披露内容,其它 GDF 多肽的蛋白水解片段可很容易的进行制备和使用。

[0054] 总体来说,对应原 -GDF 多肽蛋白水解片段的肽的实例是一个羧基末端 (C-末端) 成熟 GDF 片段,它可特异地与一个 GDF 受体相互作用,并影响 GDF 的信号转导,以及一个氨基末端功能前区片段,它可包括一个信号肽,如在此所公开的,它可特异地与原 -GDF 多肽或成熟的 GDF 肽相互作用,并影响其引发 GDF 信号转导作用的能力。例如,原肌生长抑制素多肽的蛋白水解片段包括一个 C-末端成熟肌生长抑制素肽,它可特异地与一个肌生长抑制素受体相互作用,并诱导肌生长抑制素信号转导作用;以及一个氨基末端功能前区片段,它可特异地与肌生长抑制素相互作用,因此减少或抑制肌生长抑制素诱发肌生长抑制素信号转导作用的能力。

[0055] 原 -GDF 多肽或其功能性肽部分的一个蛋白水解片段的特性部分地在于具有或影

响与 GDF 信号转导作用刺激或抑制作用相关的活性。例如,一个原肌生长抑制素或其功能性肽部分具有肌生长抑制素受体结合活性,肌生长抑制素信号转导作用刺激或抑制活性,肌生长抑制素结合活性,原肌生长抑制素结合活性,或其联合作用。因此,术语“功能性肽部分”,当在此用来说明原 -GDF 多肽时,是指可与其受体特异地相互作用,并刺激或抑制 GDF 信号转导作用的原 -GDF 多肽的一个肽部分;它可特异的与一个成熟的 GDF 或原 -GDF 相互作用;或可发挥细胞定位活性,即信号肽的活性。应该认识到,例如,全长成熟的肌生长抑制素肽的功能性肽部分,不需要具有与成熟肌生长抑制素相同的活性,包括刺激肌生长抑制素信号转导作用的能力,因为成熟肽的功能性肽部分,例如,具有特异地与肌生长抑制素受体相互作用的能力,而不需要具有激活信号转导作用通路的能力。鉴定这样一种可用作肌生长抑制素拮抗剂的原 -GDF 多肽的功能性肽部分的方法,在此被公开或在本领域是已知的。因此,在一个实施方案中,原肌生长抑制素多肽的一种功能性肽部分可特异地与一种肌生长抑制素受体相互作用,并可作为激发肌生长抑制素信号转导作用的一种激动剂或减少或抑制肌生长抑制素信号转导作用的一种拮抗剂。

[0056] 在另一个实施方案中,原肌生长抑制素多肽的一种功能性肽部分可特异地与一种原肌生长抑制素多肽或一种成熟的肌生长抑制素肽相互作用,因此可阻断肌生长抑制素信号转导作用。这样的一种原肌生长抑制素的功能性肽部分的作用有,例如,可阻止原肌生长抑制素多肽断裂成为成熟的肌生长抑制素;与成熟的肌生长抑制素肽形成一个复合体;或有其它一些机制。在肽 - 肌生长抑制素复合体形成的部位,复合体可阻断肌生长抑制素的信号转导,例如通过减少或抑制肌生长抑制素与其受体特异地相互作用的能力,或通过与受体结合,该形式缺乏诱发肌生长抑制素信号转导作用的能力。

[0057] 原 -GDF 多肽的蛋白水解片段可通过在蛋白水解切割位点切割多肽来产生,该位点具有共有氨基酸序列 Arg-Xaa-Xaa-Arg (SEQ ID NO:21)。这样的蛋白水解识别位点的实例是 Arg-Ser-Arg-Arg (SEQ ID NO:22) 序列,在 SEQ ID NO:1 (原肌生长抑制素) 中显示为氨基酸残基 263 至 266,或 SEQ ID NO:25 (人原 -GDF-11;也见,图 2 的相对位置 267 至 270) 中的氨基酸残基 295 至 298,在 SEQ ID NO:20 显示为氨基酸残基 263 至 266 的 Arg-Ile-Arg-Arg (SEQ ID NO:23) 序列。

[0058] 除了信号肽的蛋白水解位点以外,例如原肌生长抑制素多肽,含有其它两个潜在的蛋白酶解加工位点 (Lys-Arg 和 Arg-Arg)。在后一个蛋白酶解加工位点或其附近切割原肌生长抑制素多肽,可产生一个生物学活性的 C-末端成熟人肌生长抑制素片段,其中该位点包含在共有 Arg-Xaa-Xaa-Arg (SEQ ID NO:21) 蛋白水解切割识别位点之中 (见,例如,SEQ ID NO:2 的氨基酸残基 263 至 266)。全长的成熟肌生长抑制素肽的实例含有大约 103 至大约 109 个氨基酸,其预测分子量为大约 12,400 道尔顿 (Da)。另外,肌生长抑制素可形成二聚体,后者的预期分子量大约为 23 至 30 千道尔顿 (kDa)。二聚体可以是肌生长抑制素同型二聚体,或是杂二聚体,例如,与 GDF-11 或另一个 GDF 或 TGF- θ 家族成员一起。

[0059] 本发明的一个蛋白水解片段的实例是 GDF 功能前区,例如一个肌生长抑制素功能前区,它包括大约原肌生长抑制素多肽的氨基酸残基 20 至 262,或其功能性肽部分,或 GDF-11 的功能前区,它包括原 -GDF-11 多肽的大约氨基酸残基 20 至 295,或其功能性肽部分,两者都进一步含有信号肽,该信号肽由各自的原 -GDF 多肽的大约氨基酸 1 至 20 组成。肌生长抑制素功能前区的进一步实例是,在 SEQ ID NO:4 和 SEQ ID NO:6 中列出的氨基酸

残基大约 20 至 263 ;以及在 SEQ ID NO :2, SEQ ID NO :10, SEQ ID NO :12, SEQ ID NO :8, SEQ ID NO :18, SEQ ID NO :14, SEQ ID NO :16, SEQ ID NO :20 中列出的氨基酸残基大约 20 至 262,它们可通过相应的原肌生长抑制素多肽的蛋白水解切割,通过化学合成,或可从编码蛋白水解片段的重组多核苷酸的表达而产生。肌生长抑制素功能前区的功能性肽部分的实例是可特异与肌生长抑制素或原肌生长抑制素相互作用的原肌生长抑制素功能前区的肽部分。一个 GDF-11 功能前区的实例是 SEQ ID NO :25 的大约氨基酸残基 20 至 295,它可进一步包括信号肽,该信号肽由 SEQ ID NO :25 的大约氨基酸残基 1 至 20 组成,GDF-11 功能前区的功能性肽部分的实例是可特异与成熟 GDF-11 或原 -GDF-11 多肽相互作用的 GDF-11 功能前区的肽部分。优选的是,GDF 功能前区的功能性肽部分可抑制相应 GDF 或相关 GDF 刺激信号转导作用的能力,例如通过减少或抑制 GDF 特异地与其受体相互作用的能力,或通过受体结合成为一个无活性复合物。在一个实施方案中,本发明提供了一个原 -GDF 多肽的功能性片段,特别是一个 GDF 功能前区的功能性片段,可有效地连接于一个 GDF 信号肽,优选一个肌生长抑制素信号肽或 GDF-11 信号肽,分别由原肌生长抑制素或原 -GDF-11 的开始大约 15 至 30 个末端氨基酸组成。

[0060] 如在此所公开的,肌生长抑制素功能前区或 GDF-11 功能前区可与成熟的肌生长抑制素, GDF-11 或两者相互作用,因此可减少或抑制成熟 GDF 与其受体特异相互作用的能力(见实施例 7 和 8)。因此,获得肌生长抑制素功能前区的一个功能性肽部分的方法,例如,可采用在此提供的方法,通过检测肌生长抑制素功能前区的肽部分,并鉴定可特异地与肌生长抑制素或与原肌生长抑制素相互作用,减少或抑制肌生长抑制素与肌生长抑制素受体的相互作用或肌生长抑制素刺激肌生长抑制素信号转导作用的能力的功能前区功能性肽部分。

[0061] 可特异地与肌生长抑制素相互作用的原肌生长抑制素功能前区的一个功能性肽部分,或另一个 GDF 功能前区的一个功能性肽部分,也可采用任何已知用于识别特异的蛋白质-蛋白质相互作用的测定方法来鉴定。这样的测定方法包括,例如,凝胶电泳法,亲和层析法,Fields 和 Song 的双杂种系统法(the two hybridsystem) (*Nature* 340 :245-246, 1989 ;也见美国专利 5,283,173 ;Fearon 等人, *Proc. Natl. Acad. Sci., USA* 89 :7958-7962, 1992 ;Chien 等人, *Proc. Natl. Acad. Sci., USA* 88 :9578-9582, 1991 ;Young, *Biol. Reprod.* 58 :302-311 (1998),每一篇文献都在此引入作为参考文献),逆向双杂种测定方法 (Leanna 和 Hannink, *Nucl. Acid Res.* 24 :3341-3347, 1996,在此引入作为参考文献),阻遏型反式激活系统(美国专利 5,885,779,在此引入作为参考),噬菌体展示系统 (Lowman, *Ann. Rev. Biophys. Biomol. Struct.* 26 :401-424, 1997,在此引入作为参考),GST/HIS 破坏测定,突变体操纵子 (WO 98/01879,在此引入作为参考文献),蛋白补充系统(美国专利 5,776,689,在此引入作为参考文献),以及类似方法(见,例如, Mathis, *Clin. Chem.*, 41 :139-147, 1995 Lam, *Anticancer Drug Res.* 12 :145-167, 1997 ;Phizicky 等人, *Microbiol. Rev.* 59 :94-123, 1995 ;每一篇文献在此都引入作为参考文献)。

[0062] GDF 功能前区的一个功能性肽部分也可采用分子建模的方法进行鉴定。例如,一个成熟的肌生长抑制素肽的一个氨基酸序列可输入具有合适的模拟软件的计算机系统中,可产生肌生长抑制素的三维表象(“虚拟的肌生长抑制素”)。原肌生长抑制素的氨基酸序列也可被输入至计算机系统中,这样模拟软件可模拟原肌生长抑制素序列的部分,例如,功能

前区的部分,可鉴定出能与虚拟肌生长抑制素特异相互作用的功能前区的这些肽部分。特异相互作用的基线可通过模拟虚拟的肌生长抑制素和全长的原肌生长抑制素功能前区而预先确定,鉴定在虚拟肌生长抑制素中被功能前区“接触”的氨基酸残基,因为这样的一个相互作用已知可抑制肌生长抑制素的活性。

[0063] 应该认识到,这样的方法,包括双杂种体系测定法和分子建模方法,也可用来鉴定包括在本发明中的其它特异性相互作用的分子。因此诸如双杂种体系测定法可用来鉴定 GDF 受体如肌生长抑制素受体,例如可采用能特异地与一个 ActRIIA 或 Act RIIB 受体相互作用的肌生长抑制素肽或其肽部分作为测定中的一个结合成分,并鉴定能与肌生长抑制素肽特异相互作用的 GDF 受体。类似地,分子建模的方法可用来鉴定能与一个成熟 GDF 肽如成熟的肌生长抑制素,或与一个 GDF 受体,特异性相互作用的药物,因此该药物可用作 GDF 或 GDF 受体介导的信号转导作用的激动剂或拮抗剂。这样一种药物可以是,例如肌生长抑制素功能前区或 GDF-11 功能前区的一个功能性肽部分,或一种可模拟 GDF 功能前区作用的化学药物。

[0064] 用于在此公开的目的的模拟系统是基于通过例如结晶学分析或核磁共振分析得到的结构信息,或根据一级结构序列信息(见,例如, Dunbrack 等人,“会议综述:蛋白质结构预测技术关键性评价第二次会议(Meeting review:the Secondmeeting on the Critical Assessment of Techniques for Protein StructurePrediction)(CASP2)(Asilomar,加利福尼亚,1996年12月13-16日)”Fold Des. 2(2):R27-42,(1997); Fischer 和 Eisenberg,Protein Sci. 5:947-55,1996;(也见美国专利 5,436,850); Havel, Prog. Biophys. Mol. Biol. 56:43-78,1991; Lichtarge 等人, J. Mol. Biol. 274:325-37,1997; Matsumoto 等人, J. Biol. Chem. 270:19524-31,1995; Sali 等人, J. Biol. Chem. 268:9023-34,1993; Sali, Molec. Med. Today 1:270-7,1995a; Sali, Curr. Opin. Biotechnol. 6:437-51,1995b; Sali 等人, Proteins 23:318-26,1995c; Sali, Nature Struct. Biol. 5:1029-1032,1998; 美国专利 5,933,819; 美国专利 5,265,030,每一篇在此都引入作为参考文献)。

[0065] 原肌生长抑制素多肽或 GDF 受体的晶体结构坐标可用来设计可与该蛋白质结合的化合物,并可用多种方法改变其物理或生理学特性。蛋白质的结构坐标也可用来在小分子数据库中计算机筛选可与该多肽结合的药物,以便开发调节或结合的药物,它可作为 GDF 信号转导作用的激动剂或拮抗剂发挥作用。这样的药物可通过计算机采用标准的方程匹配动力学数据来鉴定(见,例如, Segel,“酶动力学”(J. Wiley&sons 1975),在此引入作为参考文献)。

[0066] 采用晶体结构数据来设计抑制剂或结合剂的方法在本领域是已知的。例如,GDF 受体坐标可添加到其它可获得的相似受体的坐标上,包括具有一个结合抑制物的受体,可为抑制物与受体的相互作用提供最接近的方式。在合理化药物设计实践中,使用的计算机程序也可用来鉴定一些化合物,这些化合物可复制发现的相互作用特性,例如类似于在成熟肌生长抑制素和共结晶的肌生长抑制素功能前区之间发现的相互作用特性。详细了解特异性相互作用的性质有助于修饰化合物以改变或改善其可溶性,药物动力学和类似特性,而不影响结合活性。

[0067] 执行晶体结构信息设计药物所需功能的计算机程序是为人熟知的。这样程序的实

例包括, Catalyst Databases™—一个可访问化学数据库如 BioByte Master File, Derwent WDI 和 ACD 的信息检索程序; Catalyst/HYPO™—产生化合物的模型, 并可用药物侯选物的结构推测解释活性的变化; Ludi™—通过鉴定和匹配互补极性和疏水基团, 将分子插入到蛋白质的活性位点中; 以及 Leapfrog™—采用遗传学算法在用户控制的参数下“生成”新的配体。

[0068] 许多普通用途的机器均可使用这种程序, 或更方便的是组建更专业化的仪器来执行这种操作。一般来说, 实例的实施可在一个或多个在可编程的系统中运行的计算机程序中执行, 每个系统至少包含一个处理器, 至少一个数据存储系统(包括易失性存储器和非易失性储器和/或存储元件), 至少一个输入装置, 至少一个输出装置。程序在处理器上运行以之行施在此所述的各种功能。

[0069] 每个这样的程序都可以以任何所需的计算机语言, 包括例如, 机器语言, 汇编语言, 高等程序语言, 或面向目标的程序语言, 来实现与计算机系统的交流。在任何情况下, 语言都可以是一个已编译或解释语言。计算机程序典型地将被储存在存储介质或装置中, 例如, ROM, CD-ROM, 磁性或光学介质, 或类似物, 当存储介质或装置被计算机读取执行在此所述的步骤时, 这些介质可以被一般或特殊用途程序计算机所读取以配置和运行计算机。该系统也可被认为是当作一个计算机可读取的存储介质, 用计算机程序配置, 其中这样配置的存储介质可使计算机以一种特殊的和预定的方式运行以执行在此所述的各种功能。

[0070] 本发明的实例包括一些系统, 例如, 基于互联网的系统, 特别是存储和处理坐标信息的计算机系统, 如在此所述, 这些信息来自结晶学或 NMR 分析, 或氨基酸或核苷酸序列信息。如在此所使用的, 术语“计算机系统”是指用来分析在此出现的坐标或序列的硬件部件, 软件部件和数据存储部件。典型地计算机系统包括一个处理、访问和操纵序列数据的处理器。处理器可以是任何已知类型的中央处理器, 例如 Intel 公司的奔腾 II 或奔腾 III 处理器, 或太阳, 摩托罗拉, 康柏, 先进微装置 (AMD) 或国际商用机器的类似处理器。

[0071] 典型地计算机系统是一个一般用途的系统, 包括处理器和一个或多个存储数据的内部数据存储部件, 以及一个或多个在数据存储部件中检索数据的数据检索装置。专业技术人员可很容易的理解现在可见到的任何一个计算机系统都是适合的。

[0072] 在一个实施方案中, 计算机系统包括连接于系统总线的一个处理器, 该总线与主存储器相连接, 优选用 RAM 来实现, 和一个或多个内部数据存储装置, 如一个硬盘驱动器或其它记录数据的计算机可读介质。在一些实施方案中, 计算机系统进一步包括一个或多个可在内部数据存储装置上读取存储数据的数据检索装置。

[0073] 数据检索装置可能代表, 例如, 一个软盘驱动器, 一个光盘驱动器, 一个磁带驱动器, 或一个能与远程数据存储系统连接的调制解调器(如, 通过互联网)。在一些实施方案中, 内部的数据存储装置是一种可移动的计算机可读介质, 如一张软盘, 一张光盘, 一张磁带等等, 上面包含已记录的控制逻辑和/或数据。计算机系统可能有利地包含或编有适当的软件, 以便一旦插入数据检索装置, 就从数据存储部件中读取控制逻辑和/或数据。

[0074] 计算机系统一般包含一个显示器, 它用来对计算机用户进行显示输出数据。也应该指出的是计算机系统可在一个网络或广域网络中连接到其它的计算机系统上以便提供对该计算机系统的集中式访问。

[0075] 只要需要鉴定一种可特异地与肌生长抑制素或与 GDF 受体相互作用的化学实体

时,可以使用任何能筛选其具有与该分子的特异相互作用能力的化学实体或片段的方法。这种过程的开始可通过,例如,在计算机屏幕上观察肌生长抑制素和肌生长抑制素功能前区。然后选择的功能前区的肽部分或可作为模拟物的化学实体,可被定位在多种方向,或插入至肌生长抑制素的单个结合位点中。可采用软件如 Quanta 和 Sybyl 完成插入过程,然后进行能量最小化,并用标准的分子力学力场,如 CHARMM 和 AMBER 完成分子动力学。

[0076] 可特别采用专门的计算机程序来选择功能前区的肽部分,或有用的化学实体,例如 GDF 受体激动剂或拮抗剂。例如,这样的程序包括 GRID (Goodford, J. Med. Chem., 28 : 849-857, 1985 ;可从牛津大学获得,牛津,英国) ;MCSS (Miranker 和 Karplus, Proteins : Structure. Function and Genetics 11 :29-34, 1991, 可从 Molecular Simulation 获得, Burlington MA) ;AUTODOCK (Goodsell 和 Olsen, Proteins : Structure. Function and Genetics 8 :195-202, 1990, 可从 Scripps Research Institute, LaJolla CA) ;DOCK (Kuntz 等人, J. Mol. Biol. 161 :269-288, 1982, 可从加利福尼亚大学获得,旧金山 CA), 每一篇文献在此都引入作为参考文献。

[0077] 已经被选择的合适的肽或药物可组装成一个单一的化合物或结合药物。装配可通过观察显示在计算机屏幕的三维影像上片段之间的相互关系来进行,然后采用软件如 Quanta 或 Sybyl 进行人工模型构建。能辅助本领域专业技术人员连接单个化学实体或片段的有用的程序,包括,例如, CAVEAT (Bartlett 等人, Special Pub., Royal Chem. Soc. 78 : 182-196, 1989, 可从加利福尼亚大学获得,伯克利 CA) ;3D 数据库系统如 MACCS-3D (MDL 信息系统, 圣莱昂纳多 CA ;综述,见 Martin, J. Med. Chem. 35 :2145-2154, 1992) ;HOOK (可从 Molecular Simulations 获得, Burlington, Mass.), 每一篇在此均引入作为参考文献。

[0078] 除了采用逐步剔除的方式,如上所述每次一个片段或一个化学实体,构建或鉴定这样特异相互作用的药物的方法以外,药物可设计成一个整体或重新采用一个空闲的活性位点,或任意的包含一个已知的可特异相互作用的药物的一部分,例如一个全长的肌生长抑制素功能前区,其与肌生长抑制素特异地相互作用。这样的方法包括,例如, LUDI (Bohm, J. Comp. Aid. Molec. Design 6 :61-78, 1992, 可从 Biosym Technologies 获得, 圣地亚哥 CA) ;LEGEND (Nishibata 和 Itai, Tetrahedron 47 :8985, 1991, 可从 Molecular Simulations, Burlington MA) ;LeapFrog (可从 Tripos Associates 获得, 圣路易 MO), 以及那些由 Cohen 等人 (J. Med. Chem. 33 :883-894, 1990) 和 Navia 和 Murcko (Curr. Opin. Struct. Biol. 2 :202-210, 1992) 所描述的,每一篇在此均引入作为参考文献。

[0079] 在本领域中可得到专门的计算机软件来评价化合物的形变能和静电相互作用。设计用于这些目的的程序实例包括 Gaussian 92, 修正版 C (Frisch, Gaussian 公司, Pittsburgh PA, 1992) ;AMBER, 4.0 版 (Kollman, 加利福尼亚大学旧金山分校, 1994) ;QUANTA/CHARMM (Molecular Simulations 公司, Burlington MA, 1994) 和 Insight II/ Discover (Biosym Technologies 公司, 圣地亚哥 CA, 1994)。这些程序可采用,例如 Silicon 图形工作站, IRIS 4D/35 或 IBM RISC/6000 工作站 550 型来执行。其它的硬件系统和软件包对本领域专业技术人员是已知的,其速度和容量可不断的更改。

[0080] 鉴定一种可与目标分子,例如与一个成熟的 GDF 肽如成熟的肌生长抑制素,或一个 GDF 受体特异相互作用的药物的分子建模过程,可如在此所公开的那样进行。第一步,执行一个靶分子,例如肌生长抑制素的虚拟表现形式。因此,在一个实施方案中,本发明可

提供一个靶分子的虚拟表现形式,其中靶分子选自原-GDF多肽,例如,原肌生长抑制素;原-GDF多肽的肽部分;GDF受体;以及GDF受体的一个相关功能区,例如,GDF结合功能区。靶分子的虚拟表现形式可被显示出来或保存在计算机系统存储器中。该过程以起始状态启动,含有虚拟的靶分子,然后转换至一种状态,其中含有一个或多个虚拟受检分子的数据库存储至计算机系统的一个存储器中。如上所述,存储器可以是任何类型的存储器,包括RAM或其它内部的存储装置。

[0081] 然后该过程转换至一种状态,其中第一个虚拟的受检分子对虚拟的靶分子特异相互作用的能力被确定,其中含有虚拟受检分子的数据库被开放以分析虚拟靶分子和虚拟受检分子的相互作用,并进行分析,其中受检分子可以是一组受检分子中的一个。确定特异性相互作用的根据是维持在计算机系统软件所进行的计算,或通过与一个预先确定的特异相互作用进行比较,后者可存储在计算机系统的存储器内,在适当时可被访问。

[0082] 然后该过程转换至一种状态,其中特异的相互作用在此被检测到,虚拟的受检分子被显示,或被存储在计算机上的第二个数据库中。如果合适,在需要时,重复该过程检测虚拟的靶分子和第二个虚拟的受检分子,第三个虚拟的受检分子,等等。

[0083] 如果确定一个虚拟的受检分子能特异地与虚拟的靶分子相互作用,所鉴定的虚拟受检分子从数据库中移出,并显示给用户。这种状态通知用户,在被输入的范围内显示名称或结构的分子可与靶分子特异的相互作用。一旦所鉴定的受检分子的名称显示给用户,该过程转换至判定状态,其中可得出结论是否在数据库中是否存在更多的虚拟受检分子或要被检测。如果在数据库中不存在更多的分子,该过程终止在结束状态。但是,如果在数据库中存在更多的受检分子,该过程转换至一种状态,其中指示器指向数据库中的下一个受检分子以便检测其特异的结合活性。以这种方式,检测新分子与虚拟的靶分子特异相互作用的能力。

[0084] 如上所述,这样的方法可用在包括在发明权利要求范围内的多个方面。因此,这些方法可用来鉴定一个原肌生长抑制素功能前区的肽部分,该功能区可与肌生长抑制素特异地相互作用,可减少或抑制肌生长抑制素与其受体相互作用的能力或影响肌生长抑制素引发信号转导作用的能力。类似地,这些方法可用来鉴定小的有机分子,它们可模拟GDF功能前区的作用,因此可减少或抑制肌生长抑制素或GDF-11的信号转导作用。这些方法也可用来鉴定可与GDF受体特异相互作用的药物,例如,Act RIIA, Act RIIB或其它GDF受体,这样的药物可用作GDF受体激动剂或拮抗剂,它们可在细胞中调节GDF信号转导作用。另外,这些方法提供了一种手段来鉴定以前未知的原-GDF多肽或GDF受体,例如,通过鉴定特殊多肽的保守结构特征。

[0085] 类似TGF- θ 超家族的其它成员,活性的GDF肽可作为前体多肽表达,它可断裂成为成熟的、生物学活性的形式。相应地,仍然在另一个实施方案中,原-GDF多肽的蛋白水解片段是一个成熟的GDF肽,或一个成熟GDF肽的功能性肽部分,其中如上所述,功能性肽部分可具有GDF激动剂或拮抗剂的活性。蛋白水解片段可以是一个成熟的C-末端肌生长抑制素肽,它大约包括原肌生长抑制素多肽的氨基酸残基268至374(见图1;也见图2),或一个成熟的C-末端GDF-11多肽,它包括原-GDF-11多肽的大约氨基酸残基299至407。全长的成熟肌生长抑制素肽的实例是在SEQ ID NO:4和SEQ ID NO:6中列出的氨基酸残基大约268至375;在SEQ ID NO:2,SEQ ID NO:10,SEQ ID NO:12,SEQ ID NO:8,SEQ ID NO:18,

SEQ ID NO:14, SEQ ID NO:16 和 SEQ ID NO:20 中列出的氨基酸残基大约 267 至 374, 以及 SEQ ID NO:27 中的氨基酸残基大约 49 至 157 和 SEQ ID NO:29 中的氨基酸残基大约 28 至 136。全长的成熟 GDF-11 肽的实例是 SEQ ID NO:25 的氨基酸残基大约 299 至 407。成熟 GDF 肽的功能性肽部分的实例是成熟肌生长抑制素或成熟 GDF-11 的肽部分, 与成熟的 GDF 肽相比它们具有激动剂或拮抗剂活性。优选的, 成熟的 GDF 肽活性具有与其受体特异相互作用的能力。

[0086] 如在此所公开的, 成熟的肌生长抑制素肽 (在此一般是指“肌生长抑制素”) 可通过与表达在细胞表面上的肌生长抑制素受体的特异相互作用 (见实施例 7) 而诱发肌生长抑制素的信号转导活性。因此肌生长抑制素的功能性肽部分可通过采用在此描述的一种方法 (实施例 7) 或本领域已知的方法, 检测成熟肌生长抑制素肽的肽部分而获得, 并鉴定能特异与一个肌生长抑制素受体, 例如, 表达在细胞上的激活素 IIA 型受体 (Act RIIA) 或 Act RIIB 受体相互作用的肌生长抑制素的功能性肽部分。

[0087] 肌生长抑制素功能前区可减少或抑制肌生长抑制素信号转导活性。在一个实施方案中, 肌生长抑制素功能前区可与肌生长抑制素特异的相互作用, 因此减少或抑制了肌生长抑制素肽与其受体特异相互作用的能力。如在此所公开的, 一个前体原肌生长抑制素也缺乏与肌生长抑制素受体特异相互作用的能力, 因此可减少或抑制原肌生长抑制素断裂成为成熟肌生长抑制素的原肌生长抑制素突变体提供了一种减少或抑制肌生长抑制素信号转导作用的方法。相应地, 在另一个实施方案中, 本发明提供了一个突变体原 -GDF 多肽, 它含有一个或多个氨基酸突变, 这些突变可阻止突变体原 -GDF 蛋白水解成为活性成熟的 GDF 肽。

[0088] 本发明的突变体原 -GDF 多肽发生的突变, 可影响蛋白水解位点如共有的蛋白水解切割识别位点 Arg-Xaa-Xaa-Arg (SEQ ID NO:21) 处的断裂, 该位点存在于原 -GDF 多肽中。因此, 突变可以是 SEQ ID NO:21 的一个 Arg 残基的突变, 这样以致突变体原肌生长抑制素, 例如, 不能断裂成为肌生长抑制素功能前区和成熟的肌生长抑制素肽。但是, 突变也可发生在不是蛋白水解切割位点的位点上, 可改变蛋白酶结合原 -GDF 多肽的能力, 影响在切割位点上的蛋白水解作用。本发明的突变体原 -GDF 多肽, 例如, 一个突变体原肌生长抑制素或突变体原 -GDF-11 相对肌生长抑制素或 GDF-11 具有显性失活活性, 因此可用来减少或抑制肌生长抑制素或 GDF-11 在细胞中的信号转导作用。

[0089] 本发明也提供了一个基本上纯化的多核苷酸, 如上所述, 它编码一个原肌生长抑制素多肽或一个突变体原肌生长抑制素的肽部分, 或一个原 -GDF-11 多肽或突变体原 -GDF-11 的肽部分。如在下面更详细公开的, 本发明也提供了可用作调节细胞上肌生长抑制素效应的药物的多核苷酸, 并进一步提供了一个编码 GDF 受体的多核苷酸, 或其功能性肽部分。在下面的公开书中也提供了这样多核苷酸的实例。同样地, 应该认识到下面的公开内容与在此公开的本发明的许多实施方案是相关的。

[0090] 术语“多核苷酸”在此广泛用于表示两个或更多的脱氧核糖核苷酸或核糖核苷酸组成的序列, 它们通过一个磷酸二酯键联系在一起。像这样, 术语“多核苷酸”包括 RNA 和 DNA, 可以是一条基因或其一部分, cDNA, 合成的多聚脱氧核糖核酸序列, 或类似物, 可以是单链或双链, 也可以是 DNA/RNA 杂交链。而且, 在此所使用的术语“多核苷酸”包括天然存在的核酸分子, 可从细胞中分离, 也可以是合成的分子, 例如, 可通过化学合成或酶学的方法

法如通过聚合酶链式反应 (PCR) 来制备。在许多实施方案中,本发明的一种多核苷酸含有核苷类似物或核苷酸类似物,或主链键而不是磷酸二酯键(见上)。

[0091] 一般来说,包含多核苷酸的核苷酸是天然存在的脱氧核糖核苷酸,如与 2'-脱氧核糖连接的腺嘌呤、胞嘧啶、鸟嘌呤或胸腺嘧啶,或核糖核苷酸如与核糖连接的腺嘌呤、胞嘧啶、鸟嘌呤或尿嘧啶。但是,多核苷酸也可含有核苷酸类似物,包括非天然存在的合成核苷酸或修饰的天然核苷酸。这样的核苷酸类似物在本领域内是为人熟知的,并在市面上可购得,含有这样的核苷酸类似物的多核苷酸也是一样(Lin 等人, *Nucl. Acids Res.* 22 : 5220-5234(1994); Jellinek 等人, *Biochemistry* 34 : 11363-11372(1995); Pagratis 等人, *Nature Biotechnol.* 15 : 68-73(1997), 每一篇文献在此均引入作为参考文献)。

[0092] 连接多核苷酸的核苷酸的共价键一般是磷酸二酯键。但是,共价键也可以是许多其它键中的任意一种,包括硫二酯键、磷酸硫代硫酸键、肽样键或任何其它本领域已知可用于连接核苷酸产生合成多核苷酸的键(见,例如, Tam 等人, *Nucl. Acids Res.* 22 : 977-986(1994); Ecker 和 Crooke, *BioTechnology* 13 : 351360(1995), 每一篇在此均引入作为参考文献)。掺入非天然的核苷酸类似物或连接核苷酸或类似物的化学键,在多核苷酸暴露于含有核酸降解活性的环境中时特别有用,这种环境包括,例如一种组织培养基或在给一个活体给药时,因为修饰的多核苷酸不太容易被降解。

[0093] 由天然存在的核苷酸和磷酸二酯键组成的多核苷酸可被化学合成来产生,或采用重组 DNA 的方法,使用一个合适的多核苷酸作为模板来产生。相对比的是,由核苷酸类似物或不是磷酸二酯键的共价键组成的多核苷酸一般采用化学合成,尽管如 T7 聚合酶的酶可以将某种类型的核苷酸类似物掺入到多核苷酸中,因此可用来从一个合适的模板重组产生一个这样的多核苷酸(Jellinek 等人,见前,1995)。

[0094] 当一个多核苷酸编码一种肽,例如,原肌生长抑制素的肽部分或一个肽药物,编码的序列一般包含在一个载体中,并可有效地与合适的调控元件相连接,后者包括,如果需要,一个组织特异性启动子或增强子。编码肽可进一步被操作与,例如,肽标记如 His-6 标记或类似物连接,后者便于在靶细胞上鉴定药物的表达。多聚组氨酸标记肽如 His-6 可采用二价阳离子如镍离子,钴离子或类似物来进行检测。其它的肽标记包括,例如,FLAG 抗原表位,它可采用抗 FLAG 抗体来检测(见,例如, Hopp 等人, *BioTechnology* 6 : 1204(1988); 美国专利 5,011,912, 每一篇均在此引入作为参考文献); c-myc 抗原表位,它可采用该抗原表位的特异抗体来检测;生物素,它可采用链霉亲和素或亲和素来检测;和谷胱甘肽 S- 转移酶,它可采用谷胱甘肽来检测。这样的标记可提供其它的优势,即它们可便于分离有效连接的肽或肽药物,例如,在需要获得一个对应于肌生长抑制素多肽的蛋白水解片段的基本上纯化的肽。

[0095] 如在此所使用的,术语“有效连接”或“有效联系”的意思是两个或更多的分子根据相互的位置进行定位,这样它们可作为一个单一的单位,并可产生一种属于一个或两个分子或其联合体的功能。例如,编码本发明一种肽的一个多核苷酸序列可有效的与一个调控元件连接,其中调控元件可将其调节效应赋予该多核苷酸,这种方法类似于该调控元件作用于在细胞中与该调控元件相联系的一种多核苷酸序列。第一个多核苷酸编码序列也可有效的与第二个(或更多的)编码序列相连接,这样嵌合的多肽可从有效连接的编码序列中表达。嵌合的多肽可以是一个融合多肽,其中两个(或更多的)编码肽被翻译成一个

单一多肽,即通过一个肽键共价结合;或可被翻译成两个单独的多肽,在翻译中,可有效的互相连接形成一个稳定的复合体。

[0096] 一个嵌合多肽一般可表现其肽组分的一些或所有特性。像这样,如在此所公开的,嵌合多肽在实施本发明的方法中是特别有用的。例如,在一个实施方案中,本发明的一种方法可在细胞中调节肌生长抑制素的信号转导。因此,当一个嵌合多肽的一个肽组分可编码细胞区室的定位功能区,第二个肽组分编码显性失活 Smad 多肽时,功能性嵌合多肽可被细胞区室定位功能区转运至细胞的区室中,并具有 Smad 多肽的显性失活活性,因此可调节细胞中的肌生长抑制素信号转导作用。

[0097] 细胞区室化功能区是为人所熟知的,包括,例如质膜定位功能区,核定位信号,线粒体膜定位信号和内质网定位信号,或类似物(见,例如, Hancock 等人, EMBO J. 10: 4033-4039, 1991; Buss 等人, Mol. Cell. Biol. 8: 3960-3963, 1988; 美国专利 5, 776, 689, 每一篇文章在此均加入作为参考)。这样一种功能区可用来将一种药物靶向至细胞中的一个特殊区室中,或引导药物从细胞中分泌出来。例如,肌生长抑制素受体的激酶功能区如 Act RIIB 一般与质膜的内表面是相联系的。因此,包含显性失活肌生长抑制素受体激酶功能区的嵌合多肽,例如缺乏激酶活性的显性失活 Act RIIB 受体,可进一步包含质膜定位功能区,因此可将显性失活 Act RIIB 激酶功能区定位在细胞膜的内表面。

[0098] 如在此所公开的,原 -GDF 信号肽具有细胞定位活性。如在此所使用的,术语“细胞定位活性”是指信号肽指引有效连接的肽转运至一个或多个特异的细胞内区室中,或引导分子从细胞中分泌出来。像这样,原 -GDF 信号肽可特殊地用于指引肽或其它与信号肽有效连接的药物转运至相同的细胞内区室中,这与具有实质上相同信号肽的天然表达的 GDF 一样。而且,信号肽,例如一个原肌生长抑制素信号肽包含原肌生长抑制素的大约开始 15 至 30 个氨基酸,可引导一个有效连接的药物通过与含有信号肽的天然原 -GDF 相同的通路从细胞中分泌出来。因此,实施本发明的一种方法的特别有用的药物包括一个 GDF 功能前区,或其与 GDF 信号肽有效连接的功能性肽部分,优选原肌生长抑制素或原 -GDF-11 信号肽。

[0099] 本发明的一种多核苷酸,包括用于实施本发明方法中的一种多核苷酸,可直接与一种靶细胞接触。例如,用作反义分子的寡核苷酸,核酶,或三联药物 (triplexing agent) 可直接与一种靶细胞接触,因此进入细胞并影响它们的功能。一个多核苷酸药物也可特异地与一种多肽相互作用,例如,肌生长抑制素受体(或肌生长抑制素),因此可改变肌生长抑制素与受体特异性相互作用的能力。这样的多核苷酸,以及制备和鉴定这种多核苷酸的方法在此被公开或在本领域中是已知的(见,例如, O = Connell 等人, Proc. Natl. Acad. Sci., USA 93: 5883-5887, 1996; Tuerk 和 Gold, Science 249: 505-510, 1990; Gold 等人, Ann. Rev. Biochem. 64: 763-797, 1995; 每一篇在此均引入作为参考文献)。

[0100] 本发明的一个多核苷酸,其可编码原 -GDF 多肽如原肌生长抑制素的肽部分,或编码突变体原肌生长抑制素多肽,或编码 GDF 受体或其功能性肽部分,或是可用在实施本发明方法中的多核苷酸药物,可被包含在一个载体中,这样可便于多核苷酸的操作,包括将多核苷酸引入一个靶细胞中。载体可以是一个克隆载体,可用于保留多核苷酸,或可以是一个表达载体,它除了含有多核苷酸以外,还含有用于表达多核苷酸的调控元件,当多核苷酸编码一种肽时,用以在一种特殊的细胞中表达所编码的肽。一种表达载体可以包含表达元件,例如这些元件是,维持编码的多核苷酸的转录所必须的,或在被克隆进载体前,调控元件可

有效地与多核苷酸连接。

[0101] 一个表达载体(或多核苷酸)一般含有或编码一个启动子序列,它能提供组成型,或如果需要,诱导性或组织特异性或发育期特异性的编码多核苷酸的表达,一个聚腺苷酸(poly-A)识别序列,以及一个核糖体识别位点或内部核糖体进入位点,或其它的调控元件如一个增强子,它可以是组织特异性的。如果需要,载体也可含有原核或真核宿主系统复制所需的元件,或两者都有。这样的载体,包括质粒载体和病毒载体如细菌噬菌体,杆状病毒,逆转录病毒,慢病毒,腺病毒,痘苗病毒,塞姆利基森林病毒和腺伴随病毒载体,都是已知的,并可从商业渠道中获得(Promega, Madison WI; Stratagene, La Jolla CA; GIBCO/BRL, Gaithersburg MD)或由本领域的专业技术人员来构建(见,例如, Meth. Enzymol. 185卷, Goeddel, 主编(Academic Press, Inc., 1990); Jolly, Canc. Gene Ther. 1:51-64, 1994; Flotte, J. Bioenerg. Biomemb. 25:37-42, 1993; Kirshenbaum 等人, J. Clin. Invest. 92:381-387, 1993; 每一篇在此均引入作为参考文献)。

[0102] 四环素(tet)诱导型启动子可特别用于驱动本发明的多核苷酸的表达,例如,一个编码肌生长抑制素的显性失活形式的多核苷酸,其中蛋白酶解加工位点已经突变,或编码肌生长抑制素功能前区的多核苷酸,它可与一个成熟的肌生长抑制素肽形成复合体,或编码GDF受体的显性失活形式的多核苷酸。在将四环素,或四环素类似物给与一个含有与tet诱导型启动子有效连接的多核苷酸的受试者的过程中,可诱导编码肽的表达,因此肽可以影响其活性,例如,因此一种肽药物可减少或抑制肌生长抑制素的信号转导作用。可采用这样一种方法,例如,来诱导成体生物体的肌肉过度生长。

[0103] 多核苷酸也可有效地与组织特异性调控元件相连接,例如,一个肌肉细胞特异性调控元件,这样编码肽的表达被限制在个体的肌肉细胞中,或培养细胞中,例如器官培养的混合细胞群体中的肌肉细胞中。肌肉细胞特异调控元件包括,例如,肌肉肌酸肌酶启动子(Sternberg 等人, Mol. Cell. Biol. 8:2896-2909, 1988, 在此引入作为参考文献)和肌球蛋白轻链增强子/启动子(Donoghue 等人, Proc. Natl. Acad. Sci., USA 88:5847-5851, 1991, 在此引入作为参考文献),在本领域均是为人熟知的。

[0104] 病毒表达载体可特别用来将多核苷酸引入细胞中,特别是受试者的细胞中。病毒载体提供了这样的优势,即它们可以以相对高的效率感染宿主细胞,并可感染特殊的细胞类型。例如,编码原肌生长抑制素功能前区或其功能性肽部分的多核苷酸可被克隆进一个杆状病毒载体中,然后它被用于感染一个昆虫宿主细胞,因此提供了一种产生大量编码功能前区的方法。病毒载体也可来自一种感染目标生物体细胞的病毒,该生物体,如脊椎动物如哺乳动物、鸟类或鱼类宿主细胞。病毒载体可特别地用于将一个用在实施本发明方法中的多核苷酸引入一个靶细胞中。已经开发出了用于特殊宿主系统的病毒载体,特别是哺乳动物系统,包括,例如逆转录病毒载体,其它的慢病毒如那些以人免疫缺陷型病毒(HIV)为基础的病毒载体,腺病毒载体,腺伴随病毒载体,疱疹病毒载体,痘苗病毒载体,和类似载体(见 Miller 和 Rosman, BioTechniques 7:980-990, 1992; Anderson 等人, Nature 392:25-30 Suppl., 1998; Verma 和 Somia, Nature 389:239-242, 1997; Wilson, New Engl. J. Med. 334:1185-1187(1996), 每一篇文献在此均引入作为参考文献)。

[0105] 当逆转录病毒,例如,用于基因转移时,理论上是可以开发出复制感受态逆转录病毒,因为在用来产生逆转录病毒载体的包装细胞系中重组逆转录载体和病毒基因序列。包

装细胞系,其中已经减少或消除了重组产生复制感受态病毒的可能性,它可用来将复制感受态逆转录病毒产生的可能性降至最低。所有用来感染细胞的逆转录病毒载体上清液均被采用标准的测定方法如 PCR 和逆转录酶测定来筛选复制感受态病毒。逆转录病毒载体可允许异源基因整合进入宿主细胞基因组中,这可允许基因在细胞分裂后传递给子细胞。

[0106] 包含在一个载体中的一种多核苷酸,可通过任何本领域已知的各种方法引入到细胞中 (Sambrook 等人, 分子克隆:实验室手册 (Molecular Cloning: A Laboratory Manual) (Cold Spring Harbor Laboratory Press 1989); Ausubel 等人, 现代分子生物学方法 (Current Protocols in Molecular Biology), John Wiley 和 Sons, Baltimore, MD (1987, 和 1995 年增刊), 每一篇文献在此均加入作为参考)。这样的方法包括,例如,转染、脂质转染法、微注射、电穿孔和用病毒载体感染;可能包括使用脂质体、微乳剂或类似物,它们可促进多核苷酸进入细胞中,并可保护多核苷酸在进入细胞前免遭降解。特殊方法的选择将依赖于,例如多核苷酸要引入的细胞,以及细胞是否从培养物中分离,或存在于培养的组织或器官或原位中。

[0107] 通过用病毒载体感染将多核苷酸引入细胞中是特别有益处的,因为它可有效的将核酸分子在离体或体内引入细胞内(见,例如,美国专利 5,399,346,在此引入作为参考)。而且,病毒是非常专一性的,其选择的基础是在一种或几种特殊类型细胞中感染和增殖的能力。因此,它们的天然特异性可用来将包含在载体中的核酸分子靶向至特殊类型的细胞中。像这样,基于 HIV 的载体可用来感染 T 细胞,基于腺病毒的载体可用来,例如,感染呼吸上皮细胞,基于疱疹病毒的载体可用来感染神经元细胞,和类似情况。其它载体,如腺伴随病毒具有更大的宿主细胞范围,因此可用来感染多种细胞类型,尽管病毒或非病毒载体也可用特殊的受体或配体修饰通过受体介导的形式改变靶向的特异性。

[0108] 本发明也提供了一些抗体,它们可特异地与原肌生长抑制素多肽或突变体原肌生长抑制素多肽的肽部分结合。本发明中特别有用的抗体包括可特异与肌生长抑制素功能前区,或其功能性肽部分结合的抗体,和可与原肌生长抑制素多肽结合并减少或抑制原肌生长抑制素蛋白水解断裂为成熟肌生长抑制素肽的抗体。另外,本发明的一种抗体可以是与 GDF 受体特异结合的抗体,或其功能性肽部分,如下面所描述的。制备和分离本发明的抗体的方法在下面更详细的进行描述,其公开内容在此引入作为参考文献。

[0109] 肌生长抑制素对于骨骼肌质量的正常调节是很重要的。相对于野生型的小鼠,敲除肌生长抑制素的小鼠,缺乏肌生长抑制素,因为增生和肥大的联合作用,其肌肉的量是前者的两至三倍。如在此所公开的,敲除肌生长抑制素的小鼠的脂肪聚积大幅度减少,至少部分是因为全身骨骼肌组织合成代谢增加。相反的是,在裸鼠中肌生长抑制素的过度表达诱导了消耗综合征,类似于在人类慢性疾病如癌症或 AIDS 的患者中所观察到的恶病质状态。如在此进一步公开的,肌生长抑制素活性可通过一种具有 Smad 信号转导通路特性的信号转导作用进行介导。相应地,本发明提供了调节细胞肌生长抑制素效应的方法,其方法是将细胞与一种可影响细胞内肌生长抑制素信号转导的药物接触。

[0110] 如在此所使用的,术语“调节”,当用来指肌生长抑制素在细胞上的作用时,其意思是细胞中的肌生长抑制素信号转导作用是增加的或减少或抑制的。术语“增加”和“减少或抑制”用来指相对于肌生长抑制素信号转导活性的基线水平,它是在缺乏肌生长抑制素时的信号转导通路的水平,或在存在肌生长抑制素时是正常细胞的活性水平。例如,肌生长

抑制素信号转导通路可在与肌生长抑制素接触的肌肉细胞中表现特殊的活性,在肌肉细胞与肌生长抑制素功能前区的进一步的接触中,肌生长抑制素信号转导活性被减少或抑制。像这样,肌生长抑制素功能前区是一种可用来减少或抑制肌生长抑制素信号转导作用的药物。类似地,另一个 GDF 家族成员的功能前区如 GDF-11 功能前区,或另一个 TGF- β 家族成员如激活素功能前区, MIS 功能前区,或类似物,可用来降低肌生长抑制素的信号转导作用。术语“减少或抑制”在此一起使用,是因为认识到,在一些情况下,肌生长抑制素信号转导作用的水平可降低至用一种特殊检测方法可检测到的水平之下。像这样,采用这种测定方法来确定是否保持低水平的肌生长抑制素信号转导作用,或信号转导作用是否完全被抑制是不可能的。

[0111] 如在此所使用的,术语“肌生长抑制素信号转导作用”是指一系列事件,一般是发生在细胞中的一系列蛋白-蛋白的相互作用,其原因是肌生长抑制素与表达在细胞表面上的肌生长抑制素受体之间的特异相互作用。像这样,肌生长抑制素信号转导作用可被检测,例如,可通过检测肌生长抑制素与细胞上其受体的特异相互作用,通过检测在细胞中肌生长抑制素信号转导作用中涉及的一个或多个多肽的磷酸化,通过检测一个或多个由于肌生长抑制素信号转导所特定诱导的基因的表达,或通过检测对肌生长抑制素信号转导作用反应而发生的表型改变(见实施例)。如在此所公开的,一种用在本发明方法中的药物可作为刺激肌生长抑制素信号转导的激动剂或作为减少或抑制肌生长抑制素信号转导的拮抗剂。

[0112] 本发明的方法在此一般以关于肌生长抑制素为例。但应该认识到,本发明的方法可更广泛地包括通过将细胞与影响细胞内由 GDF 产生的信号转导的药物接触,调节其它 GDF 肽,例如 GDF-11,对细胞的作用。实行本发明所有范围的方法因为本公开内容而很容易地理解,包括例如,识别 GDF 受体的方法,鉴定药物的方法,该药物调节由于 GDF 与其受体特异性相互作用而产生的信号转导,和类似方法。

[0113] 肌生长抑制素信号转导通路在此的实例是 Smad 通路,它是在肌生长抑制素特异地与一个激活素 II 型受体的细胞外功能区作用过程中启动的,通过细胞内多肽的相互作用而进行传导,多肽包括细胞中的 Smad 蛋白。一般来说,肌生长抑制素信号转导与特殊的细胞内多肽如 Smad 多肽的磷酸化或去磷酸化有关。因此,在存在肌生长抑制素时,细胞中的肌生长抑制素信号转导可通过与缺少肌生长抑制素时多肽磷酸化水平比较,检测一个或多个 Smad 多肽磷酸化增加的水平来进行检测。本发明的一种方法提供了增加或减少肌生长抑制素信号转导作用的手段,因此涉及肌生长抑制素信号转导通路中的 Smad 多肽的磷酸化水平将分别增加至正常的水平之上或降低至期望的水平之下。

[0114] 本发明的方法的施行,例如,可通过在合适的条件下,将靶细胞和一种影响细胞肌生长抑制素信号转导的药物相接触。合适的条件可通过将细胞置于合适的培养基中,该细胞可以是一种分离的细胞或是一种组织或器官的成分,或与生物体中的原位细胞接触。例如,含有细胞的培养基可与影响肌生长抑制素与表达在细胞上的肌生长抑制素受体特异相互作用的能力的一种药物接触,或者与影响细胞内肌生长抑制素信号转导通路的一种药物接触。一般来说,细胞可以是受试者一种组织或器官的成分,在这种情况下,接触细胞可包括将药物给与受试者。但是,细胞也可在培养液中处理,然后保存在培养液中,给与一个受试者,或用来产生转基因的非人动物。

[0115] 用在本发明方法中的一种药物可以是任何类型的分子,例如,一种多核苷酸,一种

肽,一种肽模拟体 (peptidomimetic),类肽 (peptoids) 如 vinylogous 类肽,一种小的有机分子,或类似物,可以以任何方式作用影响肌生长抑制素的信号转导。该药物可通过结合肌生长抑制素或肌生长抑制素受体如激活素受体在细胞外作用,因此可改变肌生长抑制素与其受体特异相互作用的能力,或可在细胞内作用改变细胞内的肌生长抑制素信号转导。另外,该药物可以是一种激动剂,可模仿或增强细胞上肌生长抑制素的作用,例如,肌生长抑制素与其受体特异相互作用的能力,因此增加了细胞中的肌生长抑制素信号转导作用;或可以是一个拮抗剂,它可减少或抑制肌生长抑制素在细胞上的作用,因此减少或抑制细胞中的肌生长抑制素的信号转导。

[0116] 如在此所使用的,术语“特异的相互作用”或“特异的结合”或类似的情况意味着两种分子形成一种复合体,在生理环境下该复合体是相对稳定的。在此所使用的术语涉及多种相互作用,包括,例如,肌生长抑制素与肌生长抑制素受体的相互作用,肌生长抑制素信号转导通路细胞内组分的相互作用,抗体与其抗原的相互作用,以及肌生长抑制素功能前区与肌生长抑制素的相互作用。一种特异的相互作用的特征是,解离常数至少为大约 $1 \times 10^{-6} \text{M}$,一般至少为大约 $1 \times 10^{-7} \text{M}$,通常至少大约 $1 \times 10^{-8} \text{M}$,特殊的是至少大约 $1 \times 10^{-9} \text{M}$ 或 $1 \times 10^{-10} \text{M}$ 或更大。一般一种特异的相互作用在生理条件下是稳定的,包括,例如,在活个体如人类或其它脊椎动物或无脊椎动物中发生的条件,以及发生在细胞培养中的条件,细胞培养如用来保留哺乳动物细胞或来自另一个脊椎动物机体或无脊椎动物机体的细胞。另外,一种特异的相互作用如肌生长抑制素功能前区和肌生长抑制素的细胞外相互作用一般在一定条件下是稳定的,这些条件是那些用来水产养殖具有商业价值的海生生物的条件。确定两个目的分子是否特异的相互作用的方法是为人所熟知的,包括,例如平衡透析,表面胞质团共振 (surface plasmon resonance) 和类似方法。

[0117] 一种可改变肌生长抑制素与其受体特异相互作用的药物可通过,例如,与肌生长抑制素的结合来发挥作用,这样肌生长抑制素不能特异地与其细胞受体相互作用,通过与肌生长抑制素竞争结合其受体,或通过绕过肌生长抑制素与其受体的特异相互作用的需求以便诱发肌生长抑制素的信号转导。截短型肌生长抑制素受体如肌生长抑制素受体的一种可溶性细胞外功能区是可与肌生长抑制素结合的药物一个实施方案,因此可隐蔽肌生长抑制素,并减少或抑制其与细胞表面肌生长抑制素受体特异相互作用的能力。肌生长抑制素功能前区或其功能性肽部分是可结合肌生长抑制素的药物的另一个实施方案,因此可减少或抑制肌生长抑制素与细胞表面肌生长抑制素受体的特异相互作用。这样的肌生长抑制素拮抗剂可用于本发明方法的实践中,特别是减少或抑制细胞中的肌生长抑制素信号转导作用。

[0118] 卵泡素抑制素 (Follistatin) 是可与肌生长抑制素结合的药物另一个实施方案,因此可减少或抑制肌生长抑制素与其受体特异相互作用的能力。卵泡素抑制素可结合并抑制多种 TGF- β 家族成员,包括肌生长抑制素 (GDF-8; 美国专利 6,004,937) 和 GDF-11 (Gamer 等人, *Devel. Biol.* 208:222-232, 1999) 和,因此可用来实施如所公开的该方法。尽管以前已经有人描述使用卵泡素抑制素来调节肌生长抑制素的作用 (美国专利 6,004,937),但在本公开书之前并不了解卵泡素抑制素可减少或抑制肌生长抑制素与肌生长抑制素受体如 Act RIIB 的特异相互作用的能力。

[0119] 一种可用在本发明的方法中的药物也可与一种细胞肌生长抑制素受体相互作用,

因此可与肌生长抑制素竞争受体。这样的药物可以是,例如,一种可特异结合细胞表面肌生长抑制素受体的一种抗体,包括肌生长抑制素结合区的所有或一部分,因此可阻止肌生长抑制素与受体的特异相互作用。这样一种抗肌生长抑制素受体抗体的选择是以其特异结合受体而不激活肌生长抑制素信号转导作用的能力,因此它可用作减少或抑制肌生长抑制素信号转导作用的一种肌生长抑制素拮抗剂;或以其特异结合受体并激活肌生长抑制素信号转导作用的能力作为肌生长抑制素的一种激动剂。抗体可以采用一种肌生长抑制素受体,或受体的细胞外功能区作为免疫原来产生,或是一种抗独特型抗体,可针对抗肌生长抑制素抗体,模拟肌生长抑制素。抗 GDF 受体抗体在下面更为详细的讨论。

[0120] 在本发明方法中使用的一种药物也可以是一种减少或抑制原 -GDF 多肽蛋白水解断裂成为一个活性的成熟 GDF 肽的药物,因此可减少或抑制 GDF 信号转导作用。这样一种药物可以是一种蛋白酶抑制剂,特别是一个能抑制可识别并切断 Arg-Xaa-Xaa-Arg (SEQ ID NO :21) 蛋白水解识别位点的蛋白酶活性的抑制剂。当原 -GDF 是原肌生长抑制素,可减少或抑制蛋白酶与肌生长抑制素中 Arg-Xaa-Xaa-Arg (SEQ ID NO :21) 蛋白水解断裂位点特异结合的抗肌生长抑制素抗体也可用来减少或抑制原肌生长抑制素的蛋白水解作用,因此可降低所产生的成熟肌生长抑制素的量。这样一种抗体可结合蛋白水解断裂位点,或可结合一些原 -GDF 多肽上的其它位点,这样蛋白酶的结合和切割被减少或抑制。

[0121] 另外,在本发明方法中使用的一种药物可以是一种突变体肌生长抑制素受体,例如,它对肌生长抑制素结合的反应缺乏肌生长抑制素信号转导活性,或具有组成型肌生长抑制素信号转导活性。例如,突变体肌生长抑制素受体在其激酶功能区具有一个点突变,即缺失,或类似情况,这样该受体缺乏激酶活性。这样一种显性失活突变体肌生长抑制素受体缺乏传递肌生长抑制素信号转导的能力,尽管它可具有与肌生长抑制素特异结合这样的事实。

[0122] 在本发明方法中使用的一种药物也可了解涉及肌生长抑制素信号转导通路中的一种细胞内多肽的水平或活性。如在此所公开的,肌生长抑制素对肌肉生长的调节涉及一种信号转导通路的成分,它可被激活素 II 型受体激活(见实施例 7 和 9;也见实施例 14)。肌生长抑制素可特异的与表达在培养 OS 细胞上(实施例 7)的激活素 II 型受体(Act RIIB)相互作用。低亲和性结合表明肌生长抑制素在体内与 ActRIIB 的结合可能涉及其它因子,类似于 TGF- β ,后者当 I 型受体也存在时,它与 II 型受体具有更高的亲和性(Attisano 等人, *Cell* 75 :671-680, 1993),或类似其它的系统,需要其它的分子来将配体递呈给信号受体(Massague, 见前, 1998 ;Wang 等人, *Cell* 67 :795-805, 1991)。

[0123] 肌生长抑制素与 Act RIIB 的特异相互作用表明肌生长抑制素信号转导涉及 Smad 信号转导通路的组分。因此,Smad 信号转导通路提供了在细胞上调节肌生长抑制素作用的靶标,影响 Smad 通路的药物可用来调节细胞中的肌生长抑制素信号转导。

[0124] 用来调节 GDF 信号转导的细胞内多肽组分的水平或活性的药物包括激动剂,它可增加信号转导活性,和拮抗剂,它可减少或抑制信号转导活性。对于肌生长抑制素,例如,可增加肌生长抑制素信号转导活性的药物的实例是磷酸酶抑制剂,它可减少或抑制 Smad 多肽的去磷酸化,因此可延长 Smad 的信号转导活性。显性失活 Smad 6 或 Smad 7 多肽,它消除 Smad 6 和 Smad 7 对肌生长抑制素信号转导的抑制效应,是可通过增加 Smad 信号转导而增加肌生长抑制素信号转导活性的药物的另外的实例。

[0125] 可减少或抑制肌生长抑制素信号转导活性的拮抗剂的实例是显性失活 Smad 多肽如显性失活 Smad 2, Smad 3 或 Smad 4, 其中 C 末端磷酸化位点已经突变。抑制性的 Smad 多肽如 Smad 6 和 Smad 7, 可抑制 Smad 2 和 Smad 3 的激活; 以及 c-ski 多肽, 可结合 Smad 多肽并抑制信号转导作用, 是用来通过减少 Smad 信号转导作用来减少或抑制肌生长抑制素信号转导作用的拮抗剂的另外的实例。

[0126] 当在细胞内作用的药物是一种肽, 它可与细胞直接接触, 或一种编码该肽 (或多肽) 的多核苷酸可引入细胞中, 该肽可在细胞中表达。可认识到的是一些用在本发明方法中的肽相对是较大的, 因此可能不容易穿过细胞膜。但是, 已知多种方法可将肽引入细胞内。选择将这样的肽引入细胞中的方法部分要依靠所提供多肽要进入的靶细胞的特性。例如, 当靶细胞, 或包括靶细胞的几种类型细胞, 表达一种受体, 在结合一种特殊配体过程中被内化至细胞中时, 肽药物可有效地与配体相连接。一旦与受体的结合, 肽可通过受体介导的内吞作用转运至细胞中。肽药物也可被封装在一个脂质体或配制在一种液体复合体中, 它可便于肽进入细胞中, 可进一步被修饰以如上所述表达一种受体 (或配体)。肽药物也可通过肽的改造而被引入至细胞中, 这种改造使其含有蛋白转导功能区如人免疫缺陷型病毒 TAT 蛋白转导功能区, 这样便于肽转运进入细胞中 (见 Schwarze 等人, Science 285 :1569-1572 (1999), 在此引入作为参考; 也见 Derossi 等人, J. Biol. Chem. 271 :18188 (1996))。

[0127] 靶细胞也可与编码肽药物的多核苷酸接触, 后者可在细胞中表达。表达的肽药物可以是一个突变体 GDF 受体或其肽部分。突变体 GDF 受体的实例包括肌生长抑制素受体的激酶缺陷形式, 如一种显性失活 Act RIIA 或 Act RIIB, 它们可以但不需要具有与配体 (如, 肌生长抑制素) 特异结合的能力; 一种截短型肌生长抑制素或其它 GDF 受体, 如一种肌生长抑制素受体的可溶形式, 可与肌生长抑制素结合, 因此可将其与细胞肌生长抑制素受体的特异相互作用中隔离开; Smad 多肽的显性失活形式如显性失活 Smad 3, 其中 C 末端磷酸化位点已经突变 (Liu 等人, Proc. Natl. Acad. Sci., USA 94 :10669-10674, 1997); Smad 7 多肽, 可抑制 Smad 2 和 Smad3 的激活作用 (Heldin 等人, Nature 390 :465-471, 1997); 或 c-ski 多肽, 可结合 Smad 多肽并抑制 Smad 的信号转导作用 (Sutrave 等人, Genes Devel. 4 :1462-1472, 1990)。

[0128] C-ski 肽药物在细胞中的表达可特别用来调节肌生长抑制素的信号转导。缺乏 c-ski 的小鼠显示骨骼肌群严重减少 (Berk 等人, Genes Devel. 11 :2029-2039, 1997), 而在肌肉中过度表达 c-ski 的转基因小鼠显示大量的肌肉肥大 (Sutrave 等人, 见前, 1990)。C-ski 作用并阻断一些 Smad 蛋白的活性, 包括 Smad 2, Smad 3 和 Smad 4, 它们可介导 TGF- β 和激活素 II 型受体的信号 (Luo 等人, Genes Devel. 13 :2196-1106, 1999; Stroschein 等人, Science 286 :771-774, 1999; Sun 等人, Mol. Cell 4 :499-509, 1999a; Sun 等人, Proc. Natl. Acad. Sci., USA 96 :112442-12447, 1999b; Akiyoshi 等人, J. Biol. Chem. 274 :35269, 1999)。因此根据本发明的公开书, 肌生长抑制素活性可通过 Act RIIB 的结合而被介导, 将要认识到使用 Smad 通路的肌生长抑制素或任何 GDF 的活性, 可通过在靶细胞中增加或降低 c-ski 的表达而被调节。

[0129] 用于本发明方法中的药物可以是一个多核苷酸, 它可与上述的细胞接触或被引入其中。一般来说, 但不是必须的, 多核苷酸被引入细胞中, 在此它可直接影响其功能, 或在转

录或翻译后,或这些情况都存在。例如,如上所述,多核苷酸可编码一种肽药物,后者在细胞中表达,可调节肌生长抑制素活性。这样一种表达的肽,例如,可以是一个突变体原肌生长抑制素多肽,它不能被切断形成活性的肌生长抑制素;或可以是一个突变体肌生长抑制素受体,例如,一个截短型肌生长抑制素受体细胞外功能区;一个有效地与膜锚着功能区连接的肌生长抑制素受体细胞外功能区;或一个突变体的缺乏蛋白激酶活性的肌生长抑制素受体。将一个多核苷酸引入细胞的方法的实例见下或在本领域中已知。

[0130] 一个用在本发明方法中的多核苷酸药物也可以是,或编码,一个反义分子,一个核酶或三联药物。例如,该多核苷酸可以是(或编码)一个反义核苷酸序列,如一个反义 c-ski 核苷酸序列,它可作为一个激动剂增加细胞中肌生长抑制素的信号转导;或一个反义 Smad 核苷酸序列,根据特殊的 Smad 反义核苷酸序列,它可作为一个激动剂增加肌生长抑制素的信号转导或作为一个拮抗剂减少或抑制肌生长抑制素信号转导。这样的多核苷酸可直接与靶细胞接触,一旦被细胞摄取,可影响它们的反义、核酶或三联活性;或可被一个引入细胞中的多核苷酸编码,因此一旦表达多核苷酸即可以产生,例如,一个反义 RNA 分子或核酶,其可以影响它的活性。

[0131] 一个反义多核苷酸,核酶或三联药物与靶序列互补,后者可以是一个 DNA 或 RNA 序列,例如,信使 RNA,可以是一个编码序列,一个包含内含子-外显子连接的核苷酸序列,一个调控序列如一个 SD 序列(Shine-Delgrno 序列),或类似物。互补的程度为多核苷酸,例如,一个反义多核苷酸可与细胞内的靶序列特异性相互作用。根据反义或其它多核苷酸的总长度,对于靶序列一个或几个错配是可被耐受的,不丢失其靶序列多核苷酸的特异性。因此,如果很少有含有,例如 20 个核苷酸的反义分子可耐受任何错配,那么几个不匹配将不影响反义分子的杂交效率,该反义分子,例如,可与编码一个细胞多肽的靶 mRNA 的全长互补。可耐受的错配数目是可被估计的,例如,采用已知的公式确定杂交动力学(见 Sambrook 等人,见前,1989)或采用如在此所公开的方法或本领域已知的方法来经验性的确定,特别是细胞中确定反义多核苷酸,核酶或三联药物的存在可降低靶序列的水平或细胞中靶序列编码的多肽的表达。

[0132] 用作反义分子,核酶或三联药物的多核苷酸,可抑制翻译或切断核酸分子,因此可在细胞中调节肌生长抑制素的信号转导。一种反义分子,例如,可与一个 mRNA 结合形成一个不能在细胞中翻译的双链分子。优选至少含有大约 15 至 25 个核苷酸的反义寡核苷酸,因为它们可很容易地被合成,并可特异地与靶序列杂交,尽管从引入靶细胞的多核苷酸也可表达更长的反义分子。用作反义分子的特异核苷酸序列可采用已知的方法来鉴定,例如,基因步移法(见,例如,Seimiya 等人, *J. Biol. Chem.* 272 :4361-4636(1997),在此引入作为参考文献)。当反义分子直接与靶细胞接触,它可有效地与化学活性基团如铁联 EDTA 相连接,这样可在杂交位点上切断靶 RNA。作为比较,一个三联药物可使转录停止(Maher 等人, *Antisense Res. Devel.* 1 :227(1991); Helene, *Anticancer Drug Design* 6 :569(1991))。因此,三联药物可被设计成,例如可识别 Smad 基因调控元件的序列,因此可降低或减少细胞中 Smad 多肽的表达,可调节靶细胞中肌生长抑制素的信号转导。

[0133] 本发明也提供了一种方法,它可鉴定能改变 GDF 如肌生长抑制素在细胞中作用的药物,特别是可改变 GDF 与其细胞受体特异相互作用能力的药物。这样的药物可通过增加或降低 GDF 与其受体特异相互作用的能力来起作用,可分别用于增加或降低 GDF 信号转导

作用。在此本发明的一种筛选方法的实例是采用一种肌生长抑制素受体,例如,一种激活素 II 型受体如 ActRIIA 或 ActRIIB。

[0134] 本发明的一种筛选方法可,例如,通过在合适的条件下肌生长抑制素,或其功能性肽部分,与一种肌生长抑制素受体如 Act RIIA 或 Act RIIB,和要检测的药物接触。肌生长抑制素,受体和药物可以以任何所需要的顺序接触。像这样,筛选的方法可用来鉴定竞争性或非竞争性抑制肌生长抑制素与受体结合的药物,可介导或增强肌生长抑制素与受体结合的药物,可诱导特异结合的肌生长抑制素从受体上解离下来的药物,和影响肌生长抑制素诱发信号转导能力的药物,这样的药物具有激动剂或拮抗剂活性。进行适当的反应控制以证实药物的作用是对肌生长抑制素或其它 GDF 受体特异的。

[0135] 实施本发明筛选方法的合适条件可以是任何可允许肌生长抑制素特异与其受体相互作用的条件,包括在此公开的方法(见实施例 7 和 9)或本领域已知的方法。因此,实施筛选测定的合适条件可以是,例如,采用基本上纯的肌生长抑制素受体的体外条件;细胞培养条件,使用正常表达肌生长抑制素受体的细胞,例如一种脂肪细胞或一种肌肉细胞,或一种已经被基因工程修饰在其表面表达功能性肌生长抑制素受体的细胞;或在生物体存在的原位条件下。

[0136] 本发明的一种筛选方法也可采用如上所述的分子建模的方法来实现。采用分子建模的方法提供了一种便利,经济有效的方法来在大量的群体如一个潜在药物的组合文库中筛选那些药物,这些潜在药物大多数可能特异的与 GDF 受体相互作用,因此降低了需要采用生物学测定筛选的潜在药物数量。采用分子建模的方法鉴定可与 GDF 受体如 Act RIIB 特异相互作用的药物过程中,采用在此公开的方法检测所选择的药物调节 GDF 如细胞上的肌生长抑制素的作用的能力。

[0137] 被检测药物调节肌生长抑制素作用的能力可采用在此公开的方法(见实施例 7 和 9)或本领域已知的方法来检测。术语“被检测药物”或“受检分子”在此广泛用来指许多在本发明方法中被进行激动剂或拮抗剂活性检测的任何药物。尽管该方法一般用作筛选测定法,鉴定以前未知的分子,这些分子如在此所述可作为激动剂或拮抗剂药物,该方法也可用来证实实际上具有特殊活性的已知药物具有这样的活性,例如使药物活性标准化。

[0138] 本发明的方法的实施可以,例如通过将肌生长抑制素与一种细胞接触,该细胞已经被基因工程修饰以表达 Act RIIB 受体,并通过检测参与肌生长抑制素信号转导通路中 Smad 多肽磷酸化来确定药物的作用,例如一种显性失活 Act RIIB 的作用。如果需要,该细胞可进一步被基因工程修饰包含一种报告子核苷酸序列,其表达依赖于肌生长抑制素信号转导通路,例如,依赖于 Smad 通路的激活,被检测药物的作用可通过比较存在和不存在药物,肌生长抑制素或两者的情况下报告子核苷酸序列的表达来确定。报告子核苷酸序列的表达的检测可以,例如通过检测报告子核苷酸序列的 RNA 转录物,或通过检测报告子核苷酸序列编码的多肽。一种多肽报告子可以是,例如,一种 θ -内酰胺酶,氯霉素乙酰基转移酶,腺苷脱氨酶,氨基糖甙磷酸转移酶,二氢叶酸还原酶,潮霉素-B 磷酸转移酶,胸苷激酶, θ -半乳糖苷酶,荧光素酶或黄嘌呤鸟嘌呤磷酸核苷转移酶多肽或类似物,其检测可以通过,例如检测放射活性,发光,化学发光,荧光,酶活性,或由于报告子多肽引起的特异结合。

[0139] 本发明的一种筛选方法可提供这样的优势,即它可适应高通量分析,因此可用于筛选检测药物的组合文库以鉴定那些可调节肌生长抑制素对细胞作用的药物,包括那些

可改变肌生长抑制素与肌生长抑制素受体的特异相互作用的药物。制备可检测所需活性的分子组合文库的方法在本领域是为人熟知的,并包括,例如,制备噬菌体展示肽库的方法,该肽可以是限制性的肽(见,例如,美国专利 5,622,699;美国专利 5,206,347;Scott 和 Smith, *Scienc* 献在此均引入作为参考文献);一个肽文库(美国专利 5,264,563,在此引入作为参考 e 249:386-390,1992;Markland 等人, *Gene* 109:13-19,1991;每一篇文文献);一个肽模拟文库(Blondelle 等人, *Trends Anal. Chem.* 14:83-92,1995);一个核酸文库(0 = Connell 等人,见前,1996;Tuerk 和 Gold,见前,1990;Gold 等人,见前,1995;每一篇在此都引入作为参考文献);一个寡糖文库(York 等人, *Carb. Res.*, 285:99-128, 1996;Liang 等人, *Science*, 274:1520-1522,1996;Ding 等人, *Adv. Expt. Med. Biol.*, 376:261-269,1995;每一篇在此都引入作为参考文献);一个脂蛋白文库(de Kruif 等人, *FEBS Lett.* 399:232-236,1996,在此引入作为参考文献);一个糖蛋白或糖脂文库(Karaoglu 等人, *J. Cell. Biol.*, 130:567-577,1995,在此引入作为参考文献);或一个含有,例如药物或其它药剂的化学文库(Gordon 等人, *J. Med. Chem.* 37:1385-1401,1994;Ecker 和 Crooke, *Bio/Technology*, 13:351-360,1995;每一篇在此都引入作为参考文献)。多核苷酸可特别用作可调节肌生长抑制素与其受体特异相互作用的药物,因为核酸分子具有与细胞靶位的结合特异性,该靶位包括天然存在的细胞多肽,并且因为具有这样特异性的合成分子可很容易的制备和被鉴定(见,例如,美国专利 5,750,342,在此引入作为参考文献)。

[0140] 考虑到本公开书,将要认识到多种动物模型系统可用作研究工具以鉴定可用于实践本发明方法的药物。例如,转基因小鼠或其它实验动物可采用在此公开的多种肌生长抑制素抑制剂构建物来制备,转基因非人生物体可直接被检测生物体中表达一种特殊药物的不同水平的来确定所产生的作用。另外,转基因生物体,例如,一种转基因小鼠,可与其它小鼠,例如,与 ob/ob, db/db, 或棕灰色的黄色致死突变体小鼠杂交,以确定用于治疗或防止如肥胖症、II 型糖尿病或类似的疾病的一种肌生长抑制素抑制剂表达的最佳水平。像这样,本发明提供了转基因的非人生物体,特别是含有编码肌生长抑制素功能前区的一种多核苷酸的转基因生物体,该多核苷酸包含肌生长抑制素信号肽,或编码突变体原肌生长抑制素多肽的一种多核苷酸。

[0141] 已知有多种方法来产生转基因动物。在一种方法中,将前核期的胚胎(“单细胞胚”)从雌性生物中取出,转基因被微注射进该胚胎中,其中转基因经染色体整合入所得成熟动物体的生殖细胞和体细胞中。在另一种方法中,分离胚胎干细胞,转基因通过电穿孔法、质粒转染或微注射的方法掺入至干细胞中;干细胞然后再导入到胚胎中,在此它们定居并形成生殖细胞系。将多核苷酸微注射入哺乳动物物种中的方法在,例如美国专利 4,873,191 中进行了描述,该文在此引入作为参考文献。仍然在另一个方法中,胚胎细胞用含有转基因的逆转录病毒进行感染,其中胚胎的生殖细胞含有染色体上整合的转基因。

[0142] 当转基因的动物是鸟类时,向受精卵的前核进行微注射是有问题的,因为鸟的受精卵一般在输卵管中的第一个 24 小时进行细胞分裂,因此前核是不易接近的。因此,逆转录病毒感染的方法优选制备转基因鸟类(见美国专利 5,162,215,在此引入作为参考文献)。但如果要在鸟类进行微注射,胚胎要在前次产卵后的大约 2.5 小时从处死的雌鸟中获得,转基因被微注射进入胚盘的细胞浆中,胚胎在宿主的壳中培育直至成熟(Love 等, *Biotechnology* 12,1994)。当转基因的动物是牛或猪时,微注射可因为卵不透明而受到妨

碍,这样使核很难被传统的相差显微镜所鉴别。为了克服这种问题,卵首先被离心分离出前核以便更好的观察。

[0143] 本发明的非人转基因动物可以是牛,猪,羊,鸟或其它生物。转基因可在不同的发育时期被引入胚胎靶细胞中,根据胚胎靶细胞的发育时期可选择不同的方法。受精卵是最佳的微注射的靶标。使用受精卵作为基因转移的靶标有一个主要的优势是注射的 DNA 可在第一次分裂前整合入宿主的基因中 (Brinster 等人, Proc. Natl. Acad. Sci., USA 82 : 4438-4442, 1985)。结果,所有转基因非人动物的细胞可携带掺入的转基因,因此可有效的将转基因传递给转基因者的后代,因为 50% 的生殖细胞可携带转基因。

[0144] 一种转基因动物可通过杂交两种嵌合动物而产生,每一种动物在用于繁殖的细胞中都包含外源基因物质。得到的后代 25% 将是具有纯合的外源基因物质的转基因动物,获得的动物的 50% 是杂合的,剩余的 25% 缺少外源基因物质,具有野生型的表现型。

[0145] 在微注射方法中,转基因例如通过凝胶电泳的方法被消化和纯化而不含任何载体 DNA。转基因可含有一个有效相连的启动子,它可与转录中涉及的细胞蛋白相互作用,并提供了组成型表达、组织特异性表达、发育期特异性表达、或类似情况。这样的启动子包括那些来自巨细胞病毒 (CMV)、莫罗尼白血病毒 (MLV) 和疱疹病毒,以及来自编码金属硫蛋白、骨骼激动蛋白、磷酸苯丙酮酸羧化酶 (PEPCK)、磷酸甘油酸 (PGK)、二氢叶酸酶 (DHFR) 和胸苷激酶 (TK) 的基因。也可使用来自病毒长末端重复顺序 (LTR) 的启动子如罗斯肉瘤病毒 LTR。当要转基因的动物是鸟时,优选的启动子包括那些来自鸡 θ - 球蛋白基因、鸡溶菌酶基因和鸟造白细胞组织增生病毒。用于胚胎干细胞质粒转染的构建物将使用其它的调控元件,包括例如,刺激转录的增强子元件、剪接受体、终止和多腺苷酸化信号、允许转录的核糖体结合位点、和类似物。

[0146] 在逆转录病毒感染方法中,发育中的非人胚胎可在体外培养至囊胚期。在此期间,分裂球可以是逆转录病毒的靶标 (Jaenich, Proc. Natl. Acad. Sci. USA 73 :1260-1264, 1976)。可通过酶学处理去除透明带来有效感染分裂球 (Hogan 等人, Manipulating the Mouse Embryo (Cold Spring Harbor Laboratory Press, 1986))。用来引入转基因的病毒载体系统典型的是一个装载转基因的复制缺陷型逆转录病毒 (Jahner 等人, Proc. Natl. Acad. Sci., USA 82 :6927-6931, 1985; Van der Putten 等人, Proc. Natl. Acad. Sci., USA 82 :6148-6152, 1985)。通过在一个产生病毒的细胞单层上培养分裂球可容易地和有效地获得转染 (Van der Putten 等人, 见前, 1985; Stewart 等人, EMBO J. 6 :383-388, 1987)。可选择的是,可在后面的时期中进行感染。病毒或产生病毒的细胞可注射进囊胚腔中 (Jahner 等人, Nature 298 :623-628, 1982)。大多数原代转基因细胞是转基因的嵌合体,因为掺入作用仅发生在形成转基因的非人动物的一个细胞亚群中。进一步,原代转基因细胞可在基因组的的不同位置上含有多种转基因的逆转录插入物,它们通常在后代中发生分离。另外,还可能通过在子宫内对中期妊娠的胚胎进行逆转录病毒感染,将转基因引入生殖细胞系中,尽管效率低 (Jahner 等人, 见前, 1982)。

[0147] 胚胎干细胞 (ES) 也可作为引入靶基因的靶标。ES 细胞可从体外培养的植入前胚胎中获得,并可与胚胎融合 (Evans 等人, Nature 292 :154-156, 1981; Bradley 等人, Nature 309 :255-258, 1984; Gossler 等人, Proc. Natl. Acad. Sci., USA 83 :9065-9069, 1986; Robertson 等人, Nature 322 :455-448, 1986)。转基因可通过 DNA 转染或通过逆转

录病毒介导的转导作用被有效地导入 ES 细胞中。这样转化的 ES 细胞然后可与来自非人动物的囊胚结合在一起。然后 ES 细胞定居到胚胎中,形成所得嵌合体动物的生殖细胞系(见 Jaenisch, *Science* 240 :1468-1474, 1988)。

[0148] 如在此所公开的,肌生长抑制素可至少部分地通过 Smad 信号转导通路发挥其活性,肌生长抑制素的表达与多种病理状况是相关的。像这样,本发明提供了治疗多种与肌生长抑制素相关的病理状况的新靶位,这些病理状况包括代谢性状况如肥胖症和 II 型糖尿病。相应地,本发明提供了通过调节受试者肌肉细胞或脂肪组织细胞中肌生长抑制素信号转导作用来缓解一个受试者病理状况严重性的方法,其中病理状况的特征至少部分地是肌肉或脂肪组织异常的重量、发育或代谢活性。

[0149] 肌生长抑制素可作为肌肉生长的负性调节物 (McPherron 等人,见前,1997)。肌生长抑制素敲除小鼠的重量大约超出野生型同窝出生仔 25% 至 30%,这种检测到的小鼠体重增加完全是由于骨骼肌组织重量的大幅增加。在缺乏肌生长抑制素的小鼠中,骨骼肌的重量是相应野生型同窝出生仔的大约 2 至 3 倍。在纯合子基因敲除小鼠中,这种肌肉质量的增加是增生和肥大联合作用的结果。

[0150] 如在此所公开的,杂合肌生长抑制素敲除小鼠也有骨骼肌重量增加,尽管在程度上少于纯合子突变体小鼠所观察到的程度,这样可证明肌生长抑制素在体内是剂量依赖性的(见实施例 1)。而且,肌生长抑制素在动物中的过度表达对肌肉生长具有相反的效应。例如,荷载表达肌生长抑制素的肿瘤的裸鼠发展成消耗综合征,其特征是肌肉和脂肪重量的大幅丢失(见实施例 8)。这种在裸鼠中的综合征类似于患有慢性疾病如癌症或 AIDS 的患者中发生的恶病质状态。

[0151] 肌生长抑制素免疫反应物质的血清水平在肌肉消耗方面与患者的状态是相关的 (Gonzalez-Kadavid 等人, *Proc. Natl. Acad. Sci., USA* 95 :14938-14943, 1998, 在此引入作为参考文献)。因此,患有 AIDS 的患者,也显示恶病质的征象,测量只有全身体重的减轻,其肌生长抑制素免疫反应物质的血清水平较没有 AIDS 的正常男性或没有体重减轻的 AIDS 患者有轻微的升高。但是,这些结果的解释是复杂的,因为在血清样品中检测到的肌生长抑制素免疫反应物质在 SDS 胶上没有所预期的经可靠性加工的肌生长抑制素的迁移率。

[0152] 如在此所公开的,肌生长抑制素不仅影响肌肉的重量,而且影响生物体的总体代谢。例如,肌生长抑制素在脂肪组织中表达,当动物年老时,肌生长抑制素缺陷小鼠的脂肪聚积有大幅降低(见实施例 II 和 III)。尽管在此没有提议肌生长抑制素作用的机制,但肌生长抑制素的作用可以是肌生长抑制素对脂肪组织的直接作用,或可以是由于缺少肌生长抑制素活性对骨骼肌组织引起的间接作用。不管何种机制,因肌生长抑制素活性降低导致的总体合成代谢对肌肉组织的影响,可改变生物体的总体代谢,并影响能量以脂肪的形式储备,如通过在肥胖小鼠株(棕灰色的黄色致死(A^y)小鼠)中导入肌生长抑制素突变所显示的,它的脂肪聚积可被抑制 5 倍(见实施例 5)。在含有肌生长抑制素突变的刺鼠中,异常的葡萄糖代谢也被部分抑制。这些结果表明抑制肌生长抑制素的方法可用于治疗或防止如肥胖症和 II 型糖尿病的代谢疾病。

[0153] 例如,本发明的方法可用于减轻多种病理状况的严重性,包括例如,与如癌症(见 Norton 等人, *Crit. Rev. Oncol. Hematol.* 7 :289-327, 1987) 的慢性疾病相关的恶病质,以及如 II 型糖尿病、肥胖症和其它代谢性疾病的病态。如在此使用的,术语“病理状况”是指一

种疾病,其特征在于至少部分是肌肉或脂肪组织异常的重量、发育或代谢活性。这样的病理状况,包括例如,肥胖症;与肥胖相关的状况例如,动脉粥样硬化、高血压、和心肌梗塞;肌肉消耗性疾病如肌营养不良、神经肌肉疾病、恶病质和厌食症;和代谢性疾病如 II 型糖尿病,它通常但并不一定是与肥胖症相关,这些病理状况对采用本发明方法的治疗是特别有反应的。

[0154] 如在此使用的,术语“异常”,当用于指肌肉或脂肪组织的数量、发育或代谢活性时,是与专业临床医生或其它相关的技术人员会认为正常的或理想的数量、发育或代谢活性相比较的相对意义而使用的。这种正常或理想的数值对于临床医生是已知的,是基于一般在相应人群中的健康个体中观察到或需要的平均数值。例如,临床医生知道,一个具有特定高度和体型的人,肥胖症与体重大约超过“理想”体重范围 20%以上相关。但临床医生应当认识到,仅仅是体重超过相应人群中同样高度和体型者期望体重 20%或以上的健身者并不一定是肥胖的。类似地,技术人员会知道,通过例如让患者接受多种力量检测并与相应人群中平均的健康个体的期望结果进行比较,可确定一个表现出肌肉活动异常地减少的患者具有异常的肌肉发育。

[0155] 本发明的一种方法可缓解一种病理状况的严重性,该状况的特征至少部分是肌肉或脂肪组织异常的重量、发育或代谢活性,该法通过在与该状况病因学相关的肌肉或脂肪细胞中调节肌生长抑制素信号转导作用。如在此使用的,术语“减轻”,当用在关于病理状况的严重性时,是指与该状况相关的体征或症状被减轻。要监测的体征或症状是一种特定病理状况的特征,并为专业临床医生所熟知,监测该体征或症状的方法也一样。例如,当病理状况是 II 型糖尿病时,专业临床医生可检测受试者的葡萄糖水平、葡萄糖清除率以及类似指标。当病理状况是肥胖症或恶病质时,临床医生可简单的监测受试者的体重。

[0156] 要给与受试者的药物是在促进药物与靶细胞接触的条件下用药的,如果合适,可进入细胞中。例如,可通过将多核苷酸加入一个可感染细胞的病毒载体中推动多核苷酸药物进入细胞。如果不能得到对细胞类型特异的病毒载体,载体可被修饰以表达一种对靶细胞上表达的配体(或受体)特异的受体(或配体),或可封装在一个脂质体中,它也可被修饰以包含这样一种配体(或受体)。一种肽药物可通过许多方法被引入细胞中,包括例如,通过将肽工程改造以含有一个蛋白转导功能区如人免疫缺陷病毒 TAT 蛋白转导功能区,它可加速肽转运至细胞中(见 Schwarze 等人,见前,1999;Derossi 等人,见前,1996)。

[0157] 药物在靶细胞中的存在可被直接鉴定,例如通过有效的将一个可检测的标记物连接到药物上,通过采用对药物、特别是一种肽药物有特异性的抗体,或通过检测由于药物造成的下游效应,例如细胞中 Smad 多肽磷酸化的减少。一种药物可被标记以便采用本领域已知的方法检测到(Hermanson,“生物连接技术”(Bioconjugate Techniques)(Academic Press 1996),在此引入作为参考;也见,Harlow 和 Lane,见前,1988)。例如,肽或多核苷酸药物可用多种可检测到的部分标记,包括放射性标记、酶如碱性磷酸酶、生物素、荧光色素和类似物。当药物包含在一个试剂盒中时,标记药物的试剂也可包含在该试剂盒中,或可单独从商业渠道购买到试剂。

[0158] 用在本发明方法中的一种药物可被应用到病理状况的位置,或用任何给靶细胞提供多核苷酸或肽的方法。如在此使用的,术语“靶细胞”是指要接触药物的肌肉细胞或脂肪细胞。为了给一个活的受试者用药,药物一般被制成适合给与受试者的药制剂。因此,发

明提供了一种含有药物的药学组合物，该药物在药学可接受的载体中被用来调节细胞内肌生长抑制素的信号转导作用。像这样，此药物可用作治疗患有如在此定义的病理状况的受试者的药剂。

[0159] 药学可接受的载体是本领域中熟知的，包括例如，水溶液如水或生理缓冲盐水或其它溶剂或媒介物如乙二醇、甘油、油如橄榄油、或可注射的有机酯。一种药学可接受的载体可含有生理学可接受的化合物，其作用是例如稳定或增加缀合物的吸收。这样的生理学可接受的化合物包括例如，碳水化合物如葡萄糖、蔗糖或葡聚糖，抗氧化剂如抗坏血酸或谷胱甘肽，螯合剂，低分子量蛋白或其它的稳定剂或赋形剂。本领域的专业技术人员会了解，包括生理学可接受化合物的药理学可接受载体的选择，依赖于例如治疗药物的生理-化学特性和组合物给药的途径，该途径可以是，例如口服或胃肠外如静脉内，通过注射、插管、或其它本领域已知的方法。药学组合物也可包含第二种试剂如诊断试剂、营养物质、毒素、或治疗药物如一种癌症化疗药物。

[0160] 药物可被并入封装材料中，成为如水包油的乳剂、微乳剂、微胶粒、混合微胶粒、脂质体、微球体或其它聚合物基质（见，例如，Gregoriadis, Liposome Technology, 第1卷 (CRC 出版社, Boca Raton, FL 1984) ; Fraley 等人, Trends Biochem. Sci., 6 :77 (1981), 每一篇在此均引入作为参考文献)。例如由磷脂或其它脂类组成的脂质体是无毒的、生理学可接受的和可代谢的载体，相对易于制备和应用。“隐形”脂质体（见，例如，美国专利 5, 882, 679 ; 5, 395, 619 ; 和 5, 225, 212, 每一篇都在此引入作为参考文献）是这种特别用来制备药学组合物的封装材料的实施例，该组合物可用来实行本发明的方法，可使用其它类似的“掩蔽”脂质体，这种脂质体延长了治疗药物保留在循环中的时间。例如阳离子脂质体也可被特异的受体或配体修饰 (Morishita 等人, J. Clin. Invest., 91 :2580-2585 (1993), 在此引入作为参考文献)。另外，一种多核苷酸药物可采用例如腺病毒-聚赖氨酸 DNA 复合体被引入细胞中（见，例如，Michael 等人, J. Biol. Chem. 268 :6866-6869 (1993), 在此引入作为参考文献）。

[0161] 含有改变肌生长抑制素信号转导作用的药物的药学组合物，其给药途径部分地依赖于该分子的化学结构。例如多肽和多核苷酸，当口服给药时不是特别地有用，因为它们可在消化道中被降解。但是，例如化学修饰多肽的方法是熟知的，使它们对内源性蛋白酶的降解不太敏感，或在经过消化道时更易被吸收（见，例如，Blondelle 等人，见前，1995 ; Ecker 和 Crook，见前，1995）。另外，肽药物可采用 D-氨基酸来制备，或可含有一种或多种基于肽模拟体的功能区，肽模拟体是模拟了肽功能区结构的有机分子；或基于一种类肽如 vinylogous 类肽。

[0162] 如在此所公开的药学组合物可通过多种途径分别给一个个体用药，包括例如，经口服或胃肠外如静脉内、肌肉内、皮下、眼窝内、囊内、腹膜内、直肠内、池内，或采用例如皮肤贴片或经皮电离子透入疗法，经过皮肤被动的或易化吸收。此外，药学组合物可通过注射、插管、口服或局部给药，后者可以是被动的，例如通过直接涂敷一种药膏，或例如采用一种鼻喷雾剂或吸入剂而起效，在这种情况下，组合物中的一个组分是一种合适的抛射剂。药学组合物也可被用在病理状况的部位，例如通过静脉或动脉内进入供应肿瘤的血管中。

[0163] 在实行本发明方法中要给与的药物总量可以单次剂量给与受试者，或者通过大丸剂或者在相对较短的时间内输注，或可以采用分段治疗方案给药，其中可在较长的时间内

多次用药。本领域的技术人员会知道,在一个受试者中治疗一种病理状况所需的药学组合物的量依赖于许多因素,包括受试者的年龄和一般健康状况,以及给药途径和要给药的治疗次数。考虑到这些因素,本专业的技术人员会在必要时调整特定的剂量。一般来说,药学组合物的剂型、给药途径和频率,最初是采用 I 期和 II 期临床试验来确定的。

[0164] 药学组合物可被制成口服剂型,如片剂、或溶液或悬液形式;或可与一种有机或无机的载体、或适合肠或胃肠外应用的赋形剂组成混合物,并可例如与通常的无毒性的、药学可接受的载体复合成片剂、丸剂、胶囊剂、栓剂、溶液、乳剂、悬液、或其它适合应用的形式。除了以上公开的以外,载体还包括葡萄糖、乳糖、甘露糖、阿拉伯树胶、明胶、甘露醇、淀粉糊、三硅酸镁、云母、玉米淀粉、角质素、硅胶、马铃薯淀粉、尿素、中链甘油三酯、葡聚糖,和其它适合用于制造制剂的载体,其形式是固体、半固体或液体。另外的辅剂可使用稳定剂、增稠剂或着色剂和香水,例如一种稳定干燥剂如 triulose(见,例如,美国专利 5,314,695)。

[0165] 本发明也提供了一种在受试者体内调节肌肉组织或脂肪组织生长的方法。如在此所公开的,GDF 受体如 Act RIIA 和 Act RIIB 参与了介导 GDF 的效应,GDF 如肌生长抑制素,它参与了肌肉组织和脂肪组织的形成。因此,在一个实施方案中,调节肌肉组织或脂肪组织生长的方法包括影响来自 GDF 受体如一个激活素受体,例如 Act RIIA 或 Act RIIB 的信号转导作用。这样一种方法的实施可通过用突变体 GDF 受体接触组织中的细胞,或在细胞中表达突变体 GDF 受体,它具有显性失活活性、组成型活性、或类似活性。

[0166] 在另一个实施方案中,在一个生物体中调节肌肉组织或脂肪组织生长的方法可通过给与该生物体一种药物来实施,该药物影响肌生长抑制素的信号转导。优选地,该药物是一个或编码一个肌生长抑制素功能前区或一个突变体原肌生长抑制素多肽,它们每一个都可包含肌生长抑制素的信号肽。如在此使用的,术语“生长”是以相对含义使用的,是指已经接受本发明方法的生物体中肌肉组织重量或脂肪组织重量与相应的未接受本发明方法的生物体相比较。因此,当本发明方法的实施使肌生长抑制素的信号转导被减少或抑制时,将要认识到与相应的肌生长抑制素信号转导作用没有受影响的生物体(或生物体群体)的肌肉质量相比,该生物体中肌肉组织的生长会导致肌肉质量的增加。

[0167] 本发明的一种方法可用来增加肌肉的重量或减少生物体的脂肪含量或两者均有。例如,当这样一种方法在一种可用作食物来源的生物体上实施时,食物的蛋白含量可被增加,胆固醇水平可被降低,食品的品质可被改善。本发明的一种方法也可被用来减少一种生物体中肌肉组织的生长,例如一种对环境有害的生物体,使该生物体在环境中几乎不能进行竞争。因此,本发明的方法可在任何表达肌生长抑制素的真核生物上实施,包括脊椎动物生物例如哺乳动物、鸟类或鱼类,或可以是无脊椎动物例如软体动物、棘皮动物、腹足动物或头足类动物。

[0168] 药物可以是如在此所公开的任何可改变肌生长抑制素信号转导的药物,并可以任何方便的方式给与生物体。例如当要处理的生物体是在水中养殖的鱼、虾、扇贝、或类似物时,药物可加到供养生物体的水中,或掺进它们的食物中,特别是当药物是一种可溶性的肽或是小的有机分子时。

[0169] 当在本发明的方法中使用的药物是编码一种肽药物、一种反义药物或类似物的多核苷酸时,可选择含多核苷酸的非人生物体的生殖细胞,并可产生表达药物的转基因生物

体。优选地,多核苷酸是在一个可诱导的调控元件的控制下,这样被多核苷酸编码的药物可在所需的时间和时间段内表达。相应地,本发明提供了转基因的非人生物体,以及由这些生物体产生的食品。这样的食品因为肌肉组织增加而具有增加的营养价值。转基因的非人动物可以是在此公开的任何物种,包括脊椎动物生物如牛、猪、羊、鸡、火鸡和鱼,以及无脊椎动物类如虾、龙虾、蟹、乌贼、牡蛎和鲍鱼。

[0170] TGF- θ 家族成员及其与细胞表面受体特异性相互作用的调节将要被开始被阐明。因此 TGF- θ 家族一个成员的功能前区和 TGF- θ 家族另一个成员的成熟区的共同表达与细胞内二聚作用相关,并发生生物学活性的同型二聚体的分泌 (Gray 等人, *Science* 247:1328,1990)。例如, BMP-2(骨形态发生蛋白-2) 功能前区和 BMP-4(骨形态发生蛋白-4) 成熟功能区的一起使用可引起 BMP-4 成熟区的表达大幅度增加 (Hammonds 等人, *Mol.Endocrinol.* 5:149,1991)。对于已经研究的大多数家族成员来说,同型二聚体类是有生物学活性的,而对于其它家族成员如抑制素 (Ling 等人, *Nature* 321:779,1986) 和 TGF- θ (Cheifetz 等人, *Cell* 48:409,1987), 杂二聚体也已经被检测到且似乎与相应的同型二聚体具有不同的生物学特性。

[0171] 受体-配体相互作用的研究已经揭示了大量关于细胞如何对外界刺激发生反应的信息,并已导致治疗性重要化合物的开发,如红细胞生成素、集落刺激因子和血小板源生长因子 (PDGF)。因此在鉴定介导 TGF- θ 家族成员作用的受体方面已经进行了不断的努力。如在此所公开的,肌生长抑制素与一个激活素 II 型受体特异地相互作用。对这种相互作用的证实为检定用于农业和人类治疗目的的拮抗剂和激动剂提供了靶标,例如,治疗多种病理状况如肥胖症、II 型糖尿病和恶病质。这种特异的相互作用的证实也为检定其它的肌生长抑制素受体、以及其它生长分化因子的特异性受体提供了一种方法。相应地,本发明提供了 GDF 受体,它可与一种 GDF 或 GDF 的混合物,例如与肌生长抑制素、GDF-11、或两者的混合物,发生特异地相互作用。

[0172] 本发明的一个 GDF 受体在此以肌生长抑制素受体为例,特别是激活素 II 型受体,它特异地与肌生长抑制素和 GDF-11 相互作用。但是,包括在本发明中的还有,与肌生长抑制素特异地相互作用、但不与 GDF-11 相互作用的肌生长抑制素受体,以及特异地与 GDF-11 而不是肌生长抑制素相互作用的 GDF-11 受体,等。为了便于讨论,本发明的受体在此一般地以“GDF 受体”被提及,并以肌生长抑制素受体为例,它是至少与肌生长抑制素特异性相互作用的受体。像这样,当文献一般地提到肌生长抑制素与肌生长抑制素受体的特异性相互作用时,要认识到本公开书更广泛的包括任何 GDF 受体,包括至少与 GDF-11 特异地相互作用的 GDF-11 受体。

[0173] 也提供了一个表达 GDF 受体多肽的重组细胞系,以及特异地与受体结合的抗体,编码受体的基本上纯化的多核苷酸,和基本上纯化的 GDF 受体多肽。也提供了 GDF 受体的肽部分,包括例如,GDF 受体如肌生长抑制素受体的可溶性细胞外功能区,如在此所公开的,它可改变肌生长抑制素与细胞肌生长抑制素受体的特异性相互作用;GDF 受体的组成型活性细胞内激酶功能区,它可在细胞内诱导、刺激,或者换句话说维持 GDF 信号转导作用;或 GDF 受体的其它截短型部分,它具有调节肌生长抑制素或其它 GDF 信号转导作用的能力。

[0174] 本发明还提供了鉴定 GDF 受体多肽的方法,包括:采用核苷酸探针或抗体探针筛选基因组或 cDNA 文库的方法,该文库是表达文库;应用例如 GDF,如肌生长抑制素或其功能

性肽部分筛选对其有反应的细胞,从而表达该受体的方法;双杂种体系试验,如上所述,采用例如 GDF 肽作为一个杂交的组成成分,从表达 GDF 受体的细胞中制备的 cDNA 文库表达的肽作为第二个杂交的组成成分,和类似的情况。

[0175] 如上所述,与 GDF 受体,例如肌生长抑制素受体如 Act RIIB,特异性相互作用的物质可以采用该受体鉴定,以筛选这种物质。相反地,已经被确定具有与肌生长抑制素受体如 Act RIIB 受体特异性相互作用的物质,可以被用于筛选其他的肌生长抑制素受体或其它 GDF 受体。这样一种方法可包括,将诸如该物质(或肌生长抑制素或其它 GDF)的成分与表达 GDF 受体的细胞,在足以使该物质(或 GDF)与受体特异相互作用的条件下孵育,GDF 受体可以是截短型膜结合受体或一个可溶性受体;测量结合到受体上的该物质(或 GDF);并分离受体。如上所述的分子建模方法也可以被用作筛选方法以鉴定一个 GDF 受体或其功能性肽部分。

[0176] 还提供了非人转基因动物,具有以表达 GDF 受体为特征的表现型,其表现型是通过动物体细胞和生殖细胞中含有的转基因赋予的。该转基因含有编码 GDF 受体如肌生长抑制素受体多肽的多核苷酸。产生这种转基因动物的方法在这里公开,或否则为本领域已知。

[0177] 本发明提供了一个基本纯化的、编码 GDF 受体全部或一个肽部分的多核苷酸。尽管 GDF 受体在这里用一个激活素 II 型受体作为例证,但编码激活素 II 型受体的多核苷酸以前曾经被描述过(美国专利 5,885,794)。因此,应当认识到,这种激活素 II 型受体并不包含在本发明中(Massague,见前,1998;Heldin 等人,见前,1997)。类似地,激活素 I 型受体,包括 Act RIB;TGF- θ 受体,包括 TGF- θ RJ 和 TGF- θ RII;和 BMP 受体,包括 BMP RIA、BMP RIB 和 BMP RII 受体已经被描述过并为本领域熟知(Massague,见前,1998;Heldin 等人,见前,1997),因此不包含在本发明的 GDF 受体中。

[0178] 本发明的一个多核苷酸可以编码一个具有肌生长抑制素受体活性的多肽,例如肌生长抑制素结合活性,或可以编码一个突变体肌生长抑制素受体,例如在激酶功能结构区中有突变的突变体肌生长抑制素受体,这样使该突变体充当一个显性失活的肌生长抑制素受体(见上)。因此,本发明的多核苷酸可以是天然存在的、合成的或有意操作的多核苷酸。例如,mRNA 序列部分可以因 RNA 剪接方式的替换或替换 RNA 转录启动子而被改变。作为另一个实施方案,多核苷酸可以接受定点诱变。多核苷酸也可以是反义核苷酸序列。本发明的 GDF 受体多核苷酸包括由于遗传密码而被简并的序列。存在 20 种天然氨基酸,其中大多数由不止一个密码子编码。因此,假如由多核苷酸编码的 GDF 受体多肽氨基酸序列在功能上没有改变,那么所有简并的核苷酸序列都包括在本发明中。编码肌生长抑制素受体多肽的核苷酸序列也包括在内。

[0179] 编码本发明 GDF 受体的多核苷酸的寡核苷酸部分也包括在本发明中。这种寡核苷酸通常至少大约 15 个碱基长,它足以使寡核苷酸选择性地杂交到一个编码受体的多核苷酸上,而且长度可以是至少大约 18 个核苷酸或 21 个核苷酸或以上。如这里所用的,术语“选择性杂交作用”或“选择性杂交”是指在中度严格或高度严格生理条件下的杂交,它可以区分相关的核苷酸序列和非相关的核苷酸序列。

[0180] 在核酸杂交反应中,用于获得特定严格水平的条件会因要杂交的核酸的特性而异。例如,长度、互补性程度、核苷酸序列组成(例如,相对 GC:AT 含量)和核酸类型,即寡核苷酸或靶核酸序列是 DNA 或是 RNA,都可以在选择杂交条件中考虑。一个另外的考虑是核

酸之一是否是固定化的,例如在过滤器上。选择合适严格条件的方法可以根据经验确定的或是用各种公式估计的,都是本领域熟知的(见,例如 Sambrook 等人,见前,1989)。

[0181] 一个逐步增高的严格条件的实例如下:2X SSC/0.1% SDS 在大约室温下(杂交条件);0.2X SSC/0.1% SDS 在大约室温下(低度严格条件);0.2X SSC/0.1% SDS 在大约 42°C 下(中度严格条件);和 0.1X SSC 在大约 68°C 下(高度严格条件)。可以仅用这些条件中的一个进行冲洗,例如高度严格条件,或可以使用每一个条件,例如每个 10 至 15 分钟,按上面所列的顺序,重复所列的任一或全部步骤。

[0182] 本发明编码 GDF 受体的多核苷酸可以通过几种方法中的任一种来获得。例如多核苷酸可以用杂交或以计算机为基础的技术分离,如本领域所熟知的。这些方法包括但不限于:1) 用探针与基因组或 cDNA 文库杂交以检测同源的核苷酸序列;2) 表达文库的抗体筛选以检测具有共用结构特征的克隆 DNA 片段;3) 采用能够退火到目的 DNA 序列上的引物对基因组 DNA 或 cDNA 进行聚合酶链式反应(PCR);4) 序列数据库的计算机搜索以找到相似的序列(见上);5) 扣除 DNA 文库差示筛选;和 6) 双杂种体系试验,采用例如一个成熟 GDF 肽作为杂交物之一。

[0183] 根据本公开书,即一个激活素受体与肌生长抑制素特异性相互作用,寡核苷酸探针可以根据编码激活素受体的序列而设计,例如编码与肌生长抑制素结合的细胞外功能区的序列,并用于筛选从细胞中制备的文库,如肌肉细胞或脂肪细胞,它们对肌生长抑制素起反应,从而加快编码肌生长抑制素受体的多核苷酸的鉴定。选择出的克隆可以被进一步筛选,例如通过将插入物亚克隆到一个表达载体中,并在克隆序列表达后用肌生长抑制素筛选表达的多肽。

[0184] 本发明的多核苷酸,例如一个编码肌生长抑制素受体的多核苷酸,可以来自脊椎动物种属,包括哺乳动物、鸟类或鱼类,或来自无脊椎动物。如果可以获得适合的探针,依赖核酸杂交的筛选步骤可以从任何生物体中分离任何基因序列。相对应于编码所研究蛋白的序列一部分的寡核苷酸探针可以被化学合成。这需要知道氨基酸序列的短寡肽链。考虑到遗传密码的简并性,编码受体的多核苷酸序列可以从遗传密码的一段序列中推导出来。于是,当序列被简并时,可以进行混合的加成反应。这包括变性双链 DNA 的异种混合物。对于这种筛选,杂交作用优选地在单链 DNA 或变性的双链 DNA 上进行。当有关目的多肽的 mRNA 序列存在极低量时,杂交作用在探测此来源的 cDNA 克隆中是特别有用的。因此,通过采用严格的杂交条件,以避免非特异结合,在与靶核酸完全互补的混合物中将靶 DNA 杂交到一个寡核苷酸探针上,使放射自显影成像可被用于识别特异的 cDNA 克隆(Wallace 等人, *Nucl. Acid Res.*, 9:879, 1981, 在此引入作为参考文献)。可选择地,一个扣除文库因而可以被用来排除非特异的 cDNA 克隆。

[0185] 当不知道所需多肽的整个氨基酸序列时, DNA 序列的直接合成是不可能的,精选的方法是合成 cDNA 序列。分离目的 cDNA 序列的标准步骤之一是形成在质粒或噬菌体中制备的 cDNA 文库,其中文库来自 mRNA 的逆转录,后者在具有高水平基因表达的供体细胞中是丰富的。当与聚合酶链式反应技术联合使用时,甚至极少表达的产物也可以被克隆。当已知多肽氨基酸序列的重要部分时,在已经变性成单链形式的 cDNA 克隆拷贝上进行的杂交过程可以使用标记的单链或双链 DNA 或 RNA 探针序列,它们复制了推定在靶 cDNA 中存在的序列(Jay 等人, *Nucl. Acid Res.*, 11:2325, 1983, 在此引入作为参考文献)。

[0186] 可以用一个 GDF 受体特异的抗体,如抗-Act RIIB 抗体,筛选 cDNA 表达文库,如 λ gt11 文库寻找 GDF 受体肽。抗体可以是多克隆或单克隆的,并可被用于检测指示 GDF 受体 cDNA 存在的表达产物。也可以用一个 GDF 肽筛选这样一种表达文库,例如用肌生长抑制素或其一个功能性肽部分鉴定一个克隆,后者编码至少肌生长抑制素受体的肌生长抑制素结合功能区的一部分。

[0187] 编码突变体 GDF 受体和突变体 GDF 受体多肽的多核苷酸也包括在本发明中。编码 GDF 受体的多核苷酸的改变可以是基因内突变,如点突变、无义 (STOP) 突变、错义突变、剪接位点突变或移码,或可以是杂合的或纯合的缺失,可以是自发突变或可以是例如用重组 DNA 方法人工改造的。这种改变可以用本领域专业技术人员已知的标准方法检测,包括但不限于核苷酸序列分析, Southern 印迹分析, PCR 为基础的分析如多重 PCR 或序列标记位点 (STS) 分析、或原位杂交分析。GDF 受体多肽可以用标准的 SDS-PAGE、免疫沉淀反应分析、western 印迹分析或类似方法来分析。突变体 GDF 受体的实例是截短型 GDF 受体,包括:可溶性细胞外功能区,它可具有特异地结合其同源物 GDF 的能力但缺少激酶功能区;细胞内 GDF 受体激酶功能区,它可具有组成型激酶活性;以及含点突变体 GDF 受体,它破坏了受体的激酶活性或受体的配体结合能力;以及类似物。这种 GDF 受体突变体用于调节 GDF 信号转导,并因而用于实施本发明的各种方法。

[0188] 编码 GDF 受体的多核苷酸可以通过将该多核苷酸导入适合的宿主细胞而在体外表达。“宿主细胞”可以是任何细胞,特定的载体可以在其中增殖,且在适当的时候载体中包含的多核苷酸可以在其中表达。术语“宿主细胞”包括一个原始宿主细胞的任何子代。应当理解,由于例如在复制过程中发生的突变,宿主细胞的所有子代与亲代细胞可能是不相同的。然而,当使用“宿主细胞”时,这种子代也包括在其中。获得宿主细胞的方法是本领域熟知的,宿主细胞短暂地或稳定地含有导入的本发明多核苷酸。

[0189] 本发明的 GDF 受体多核苷酸可被插入一个载体中,它可以是克隆载体或重组表达载体。术语“重组表达载体”是指质粒、病毒或本领域已知的其它载体,它们已经通过插入或合并一个多核苷酸而被操作,特别是对于本发明,是编码 GDF 受体的全部或一个肽部分的多核苷酸。这种表达载体含有一个启动子序列,它加速该宿主的插入基因序列的有效转录。表达载体通常含有一个复制起点,一个启动子,以及可以对转化细胞进行表现型筛选的特殊基因。适合用于本发明的载体包括但不限于,在细菌中表达的以 T7 为基础的表达载体 (Rosenberg 等人, 基因 (Gene) 56 :125, 1987)、在哺乳动物细胞中表达的 pMSXND 表达载体 (Lee 和 Nathans, J. Biol. Chem. 263 :3521, 1988) 和在昆虫细胞中表达的杆状病毒来源的载体。DNA 片段可以存在于载体中,可有效地连接到调控元件,例如一个启动子,它可以是 T7 启动子、金属硫蛋白 I 启动子、多角体蛋白启动子、或其它所需的启动子,特别是组织特异性启动子或诱导性启动子。

[0190] 编码 GDF 受体的多核苷酸序列可以在原核细胞或真核细胞中表达。宿主可以包括微生物、酵母、昆虫和哺乳动物生物体。在原核细胞中表达具有真核细胞或病毒序列的多核苷酸的方法是本领域熟知的,同样的是能够在宿主中表达和复制的有生物学功能的病毒和质粒 DNA 载体。含有本发明多核苷酸的表达载体的构建方法是熟知的,同样的是在选择转录或翻译控制信号中要考虑的因素,包括例如,要多核苷酸是否优选在特定的细胞类型中或在特定的条件下表达 (见,例如, Sambrook 等人, 见前, 1989)。

[0191] 各种宿主细胞 / 表达载体系统可以被用来表达 GDF 受体编码序列,包括但不限于,微生物如用重组细菌噬菌体 DNA、质粒 DNA 或粘粒 DNA 表达载体转化的细菌;用重组酵母表达载体转化的酵母细胞;用重组病毒表达载体如花椰菜花叶病毒或烟草花叶病毒感染,或用重组质粒表达载体如 Ti 质粒转化的植物细胞系统;用重组病毒表达载体如杆状病毒感染的昆虫细胞;用重组病毒表达载体如逆转录病毒、腺病毒或痘苗病毒载体感染的动物细胞系统;和为稳定表达而基因工程改造的转化动物细胞系统。当表达的 GDF 受体是翻译后修饰时,例如通过糖基化作用,选择一个能够影响所需修饰作用的宿主细胞 / 表达载体系统,例如哺乳动物宿主细胞 / 表达载体系统,可以是特别有益的。

[0192] 依据所用的宿主细胞 / 载体系统,可以在表达载体中使用许多适合的转录和翻译元件中的任何一个,包括组成型和诱导型启动子、转录增强元件、转录终止子,以及类似物 (Bitter 等人, *Meth. Enzymol.* 153 :516-544, 1987)。例如,当在细菌系统中克隆时,可以使用诱导型启动子如细菌噬菌体 Σ 的 pL, plac, ptrp, ptac (ptrp-lac 杂交启动子), 以及类似启动子。当在哺乳动物细胞系统中克隆时,可以使用来自哺乳动物细胞基因组的启动子如人或鼠金属硫蛋白启动子,或来自哺乳动物病毒的启动子如逆转录病毒长末端重复顺序、腺病毒延迟启动子或痘苗病毒 7.5K 启动子。由重组 DNA 或合成技术产生的启动子也可以被用来提供插入 GDF 受体编码序列的转录。

[0193] 在酵母细胞中,可以使用许多含组成型或诱导型启动子的载体 (见 Ausubel 等人, 见前, 1987, 见 13 章; Grant 等人, *Meth. Enzymol.* 153 :516-544, 1987; Glover, *DNA 克隆* (DNA Cloning) 第 II 卷 (IRL Press, 1986), 见第 3 章; Bitter, *Meth. Enzymol.* 152 :673-684, 1987; 也见, *酵母菌属酵母菌的分子生物学* (The Molecular Biology of the Yeast *Saccharomyces*) (Strathern 等人主编, Cold Spring Harbor Laboratory 出版社, 1982), 第 I 和 II 卷)。可以使用组成型酵母启动子如 ADH 或 LEU2, 或诱导型启动子如 GAL (Rothstein, *DNA 克隆* 第 II 卷 (见前, 1986), 第 3 章)。可选择地, 可以使用促进外源 DNA 序列整合进酵母染色体的载体。

[0194] 真核细胞系统, 特别是哺乳动物表达系统, 使表达的哺乳动物蛋白得到适当的翻译后修饰。具有正确加工初级转录物、糖基化、磷酸化和基因产物方便地插入质膜的细胞机制的真核细胞可被用作表达 GDF 受体多肽或其功能性肽部分的宿主细胞。

[0195] 使用重组病毒或病毒元件以指导表达的哺乳动物细胞系统可被改造。例如, 当应用腺病毒表达载体时, GDF 受体编码序列可被连接在一个腺病毒转录 / 翻译控制复合体上, 如延迟启动子和三联前导序列。可选择地, 可以使用痘苗病毒 7.5K 启动子 (Mackett 等人, *Proc. Natl. Acad. Sci., USA* 79 :7415-7419, 1982; Mackett 等人, *J. Virol.* 49 :857-864, 1984; Panicali 等人, *Proc. Natl. Acad. Sci., USA* 79 :4927-4931, 1982)。特别有用的是牛乳头状瘤病毒载体, 它可以作为染色体外遗传因子复制 (Sarver 等人, *Mol. Cell. Biol.* 1 :486, 1981)。此 DNA 进入小鼠细胞后不久, 质粒复制到大约每个细胞 100 至 200 个拷贝。插入的 cDNA 的转录不需该质粒整合进宿主细胞染色体, 因而产生了高水平的表达。通过在质粒中包含一个选择性标记, 如新基因, 这些载体可被用于稳定表达。可选择地, 逆转录病毒基因组可被修饰以用作能够导入和指导 GDF 受体基因在宿主细胞中表达的载体 (Cone 和 Mulligan, *Proc. Natl. Acad. Sci., USA* 81 :6349-6353, 1984)。也可以用诱导型启动子获得高水平表达, 包括但不限于金属硫蛋白 IIA 启动子和热休克启动子。

[0196] 为长期、高产量生产重组蛋白,优选稳定的表达。宿主细胞可以用 GDF 受体 cDNA 转化,而不是使用含病毒复制起点的表达载体,GDF 受体 cDNA 由合适的表达控制元件调控,如启动子、增强子序列、转录终止子和聚腺苷酸化位点,以及一个选择性标记。在重组质粒中的选择性标记可带来对这种选择的抗性,使细胞稳定地将质粒整合进其染色体中并生长形成集落,集落反过来可以被克隆并发展成细胞系。例如,在导入外源 DNA 后,可以让改造的细胞在富集培养基中生长 1 至 2 天,然后转移到选择性培养基中。可以使用许多选择系统,包括但不限于,单纯疱疹病毒胸腺嘧啶核苷激酶 (Wigler 等人, 细胞 (Cell) 11 :223,1977)、次黄嘌呤 - 鸟嘌呤磷酸核糖基转移酶 (Szybalska 和 Szybalski, Proc. Natl. Acad. Sci., USA 48 :2026,1982)、和腺嘌呤磷酸核糖基转移酶 (Lowy 等人, 细胞 (Cell) 22 :817,1980)。这些基因可分别被用在 tk^- 、 $hgpRT^-$ 或 $aprt^-$ 细胞中。此外,抗代谢物抗性可被用作选择下列基因的基础:dhfr 赋予对氨基喋呤的抗性 (Wigler 等人, Proc. Natl. Acad. Sci., USA 77 :3567,1980 ;O' Hare 等人, Proc. Natl. Acad. Sci., USA 78 :1527,1981) ;gpt 赋予对霉酚酸的抗性 (Mulligan 和 Berg, Proc. Natl. Acad. Sci., USA 78 :2072,1981) ;neo 赋予对氨基糖苷类 G-418 的抗性 (Colberre-Garapin 等人, J. Mol. Biol. 150 :1,1981) ;和 hygro 赋予对潮霉素的抗性 (Santerre 等人, 基因 (Gene) 30 :147,1984)。还有对另外的选择性基因的描述,包括:trpB 使细胞应用吲哚代替色氨酸;hisD 使细胞应用 histinol 代替组氨酸 (Hartman 和 Mulligan, Proc. Natl. Acad. Sci., USA 85 :8047,1988) ;和 ODC (鸟氨酸脱羧酶) 赋予对鸟氨酸脱羧酶抑制剂,2-(氟甲基)-DL-鸟氨酸,DFMO 的抗性 (McConlogue, Curr. Comm. Mol. Biol. (Cold Spring Harbor Laboratory 出版社,1987))。

[0197] 当宿主是真核细胞时,可以使用这些 DNA 转染方法,如钙磷共沉淀法,常规的机械方法如微注射、电穿孔、插入包封在脂质体中的质粒,或病毒载体。真核细胞也可以用编码本发明 GDF 受体的 DNA 序列和第二个外源 DNA 分子共转化,后者编码选择性表型如单纯疱疹胸腺嘧啶核苷激酶基因。另一个方法是使用真核细胞病毒载体,如猿病毒 40 (SV40) 或牛乳头状瘤病毒,以暂时地感染或转化真核细胞并表达蛋白 (Gluzman, 真核病毒载体 (Eukaryotic Viral Vectors) (ColdSpring Harbor Laboratory 出版社,1982))。

[0198] 本发明还提供了稳定的重组细胞系,该细胞系的细胞表达 GDF 受体多肽并含有编码 GDF 受体的 DNA。适合的细胞类型包括但不限于,NIH 3T3 细胞 (小鼠)、C2C12 细胞、L6 细胞和 P19 细胞。C2C12 和 L6 成肌细胞在培养基中自行分化并依赖特定的生长条件形成肌管 (Yaffe 和 Saxel, Nature 270 :725-727,1977 ;Yaffe, Proc. Natl. Acad. Sci., USA 61 :477-483,1968)。P19 是一个胚胎癌细胞系。这些细胞被描述在例如美国模式菌种收藏中心 (American Type Culture Collection (ATCC)) 的细胞系目录中。这些细胞可以用已知的方法稳定地转化 (见,例如, Ausubel 等人,见前,1995,见 9.5.1-9.5.6 节)。

[0199] GDF 受体可以用诱导型或组成型调节元件从本发明的重组多核苷酸表达,如在此所述。所需的蛋白编码序列和一个有效连接的启动子可以作为非复制型 DNA (或 RNA) 分子被导入受体细胞,它可以是线性分子或共价闭合环状分子。所需分子的表达可以是由于导入序列的临时性表达,或者是由于稳定地保存在细胞中的多核苷酸,例如通过整合进宿主细胞染色体,因而引起永久表达。因此,细胞可以是稳定地或暂时地转化的 (转染的) 细胞。

[0200] 一个可用的载体的实例是能够整合所需的基因序列进入宿主细胞染色体的载体。已稳定地在其染色体中整合了导入 DNA 的细胞也可以通过导入一个或多个标记来筛选,这

些标记可以筛选含有表达载体的宿主细胞。标记可以补充宿主的营养缺陷如 leu2 或 ura3, 它们是常见的酵母营养缺陷型标记; 可以赋予杀生物剂的抗性, 例如对抗生素或重金属离子如铜或类似物的抗性。选择性标记基因可以被直接连接到要表达的 DNA 基因序列上, 或通过共转染被导入同一个细胞。

[0201] 导入的序列可被合并进能够在受体宿主内自主复制的质粒或病毒载体中。各种载体中的任何一个可被用于此目的。选择特定质粒或病毒载体中的重要因素包括: 含载体的受体细胞可以容易被识别并从不含载体的细胞中挑选出来; 在特定宿主细胞中所需载体的拷贝数量; 和是否是需要在不同种属的宿主细胞间“穿梭”。

[0202] 对于哺乳动物宿主, 用于表达的载体系统有几种。一类载体应用 DNA 元件, 提供了在自主复制的染色体外质粒, 质粒来自动物病毒如牛乳头状瘤病毒、多瘤病毒、腺病毒、或 SV40 病毒。第二类载体包括痘苗病毒表达载体。第三类载体依靠所需的基因序列整合进宿主染色体。已稳定地在其染色体中整合了导入 DNA 的细胞也可以通过导入一个或多个标记基因 (如上所述) 来选择, 这些标记基因可以选择含有表达载体的宿主细胞。选择性标记基因可以被直接连接到要表达的 DNA 序列上, 或通过共转染被导入同一个细胞。可以包括另外的元件以提供编码的 mRNA 或肽的最佳合成, 包括例如剪接信号、转录启动子或增强子, 和转录或翻译终止信号。插入了适当的调控元件的 cDNA 表达载体是本领域熟知的 (见, 例如, Okayama, *Mol. Cell. Biol.* 3:280, 1983)。

[0203] 一旦制备了用于表达的含载体或 DNA 序列的构建物, DNA 构建物就可以被导入合适的宿主。可以用各种方法将多核苷酸导入细胞, 包括例如转染或转化的方法, 如原生质体融合、钙磷沉淀、和电穿孔或其它常规的技术, 例如当载体是病毒载体时的感染。

[0204] 本发明还提供了具有表达重组 GDF 受体的细胞的转基因动物。这种转基因动物可以被选择为具有脂肪含量下降或肌肉重量增加, 或两者都有, 因而这些动物可以被用作高肌肉和蛋白含量、低脂肪和胆固醇含量的食品的来源。它们在生殖细胞和体细胞中已经在染色体上发生了改变, 以致 GDF 特别是肌生长抑制素的产生保持在“正常”水平, 但肌生长抑制素受体的产生数量减少或完全破坏, 使得动物细胞结合肌生长抑制素的能力下降, 结果肌肉组织超过正常水平, 优选地脂肪或胆固醇水平没有升高。因此, 本发明也包括由动物提供的食品。这种食品因肌肉组织增加而提高了营养价值。本发明的转基因非人动物包括牛、猪、绵羊和禽类动物, 以及其它脊椎动物, 进一步包括转基因无脊椎动物。

[0205] 本发明也提供了一种产生具有高肌肉含量的动物食品的方法。这种方法可包括修饰动物原核胚胎生殖细胞的遗传结构, 将胚胎植入假妊娠雌性的输卵管中, 因而使胚胎成熟为正常分娩子代, 检测子代存在的转基因以鉴定转基因阳性的子代, 杂交转基因阳性的子代以获得进一步转基因阳性的子代, 加工子代以获得食品。生殖细胞的修饰包括改变遗传组成, 以减少或抑制天然存在的、编码产生肌生长抑制素受体蛋白的基因表达。例如, 转基因可以包括一个对编码肌生长抑制素受体的多核苷酸特异的反义分子; 可以包括一个无功能序列, 替换或插入内源性肌生长抑制素受体基因或转基因中; 或可以编码一个肌生长抑制素受体拮抗剂, 例如, 显性失活肌生长抑制素受体如显性失活 Act RIIB。

[0206] 如这里所用的, 术语“动物”是指除人之外任何的鸟、鱼或哺乳动物, 并包括任何发育阶段, 包括胚胎和胎儿阶段。农场动物如猪、山羊、绵羊、母牛、马、兔或类似动物; 啮齿类如小鼠; 和家养宠物如猫和狗包括在术语“动物”的意思内。此外, 在这里使用的术语“生物

体”包括如上所述的动物,以及其它真核细胞,包括例如,其它脊椎动物如爬行类和两栖类,以及如上所述的无脊椎动物。

[0207] 如这里所用的,术语“转基因”,当用来指动物或生物体时,是指动物或生物体的细胞已经被遗传学操作,以获得一个被细胞稳定地保留的外源多核苷酸序列。操作可以是例如,微注射一个多核苷酸,或用含多核苷酸的重组病毒感染。因此,这里所用的术语“转基因”是指其中一个或多个细胞获得重组多核苷酸的动物(生物体),重组多核苷酸可以在细胞内被整合进染色体,或可以保留为染色体外复制多核苷酸,这样可能被改造成为酵母人工染色体。术语“转基因动物”也包括“生殖细胞系”转基因动物。生殖细胞系转基因动物是遗传信息已经被摄取并合并进生殖系细胞的转基因动物,因而具有将信息传递给子代的能力。如果这种子代事实上具有一些或全部该信息,那么该子代也被认为是转基因动物。本发明进一步包括转基因生物体。

[0208] 转基因生物体可以是任何生物体,其基因组已经通过对早期阶段胚胎或受精卵的体外操作而被改变,或通过任何转基因技术导致特异基因敲除。术语“基因敲除”是指细胞内或体内基因的定向破坏,导致其功能完全丧失。可以通过在要使之无功能的基因中的插入而使转基因动物的靶基因无功能,例如通过同源重组,或通过任何其它在细胞内破坏基因功能的方法。

[0209] 可被用于实施本发明所述的转基因可以是一个含修饰的 GDF 受体编码序列的 DNA 序列。优选地,修饰的 GDF 受体基因是通过在胚胎干细胞中的同源定向破坏的基因。例如,可以删除 GDF 受体基因的整个成熟 C-末端区(见实施例 13)。任选地,可以用另一个多核苷酸如无功能的 GDF 受体序列,插入或替换而完成破坏(或删除)。还可以在生物体的一个细胞中导入或表达反义 GDF 受体多核苷酸、或通过在细胞中表达抗体或显性失活 GDF 受体,而使之成为“敲除”表现型。在适当时,可以在此使用编码具有 GDF 受体活性的蛋白,但在核苷酸序列上由于基因编码的简并性而不同于天然存在的 GDF 基因序列的多核苷酸,同样可使用截短的形式、等位基因突变体和种间同源物的多核苷酸。

[0210] 本发明还提供了特异地结合 GDF 受体并从而阻断 GDF 与受体结合的抗体。这种抗体可被用于,例如缓解诸如与肌肉组织有关的细胞增殖性异常的病理状况。

[0211] 与 GDF 受体,特别是肌生长抑制素受体特异结合的单克隆抗体可以增加骨骼肌的发育。在所述要求保护的方法的优选实例中,GDF 受体单克隆抗体、多肽、或多核苷酸被给予有病理状况的患者,如肌肉消耗性疾病、神经肌肉病、肌肉萎缩、衰老、或类似情况。GDF 受体抗体,特别是抗肌生长抑制素受体抗体,也可以给予有病理状况的患者,如肌肉营养不良、脊髓损伤、外伤性损伤、充血性阻塞性肺病(COPD)、AIDS 或恶病质。

[0212] 在一个优选的实施方案中,抗肌生长抑制素受体抗体经静脉内、肌肉内或皮下注射给予有肌肉消耗性疾病或异常的患者;优选地,单克隆抗体以大约 0.1Tg/kg 至大约 100mg/kg 的剂量范围给药;更优选地在大约 1Tg/kg 至 75mg/kg 之间;最优选地从大约 10mg/kg 至 50mg/kg。抗体可以通过例如使用大药丸或通过缓慢输注给药。缓慢输注优选在 30 分钟至 2 小时之内。抗肌生长抑制素受体抗体,或其它抗 GDF 受体抗体,可以被制成适于给予患者的制剂。这种制剂是本领域已知的。

[0213] 给药方案要由主治医生考虑各种因素来确定,这些可改变肌生长抑制素受体蛋白的作用,例如需要形成的组织量、组织损伤部位、损伤组织的状况、伤口大小、损伤组织的类

型、患者的年龄、性别、和饮食、任何感染的严重性、给药时间和其它临床因素。剂量可以因重制中使用的基质和要在组合物中使用的药剂类型而不同,例如药剂是抗肌生长抑制素受体抗体。通常全身性或可注射的用药,如静脉内、肌肉内或皮下注射。用药通常以最小有效剂量开始,在预先选择的时间内增加剂量直至观察到阳性效果。然后,剂量的逐步增加要限制在这种逐步递增可产生相应的效果增加的水平上,同时顾及可能出现的任何不利影响。在最终组合物中添加可以帮助增加肌肉质量的其它已知生长因子,如 IGF I (胰岛素样生长因子 I)、人、牛或鸡生长激素,也可能影响剂量。在给予抗肌生长抑制素受体抗体的实例中,抗体通常在大约 0.1Tg/kg 至大约 100mg/kg 的剂量范围给药;更优选地在大约 10mg/kg 至 50mg/kg 之间。

[0214] 如这里所用的,术语“抗体”以其广义使用,包括多克隆和单克隆抗体,以及这些抗体的抗原结合片段。用于本发明方法的抗体或其抗原结合片段具有如下特征,例如具有对 GDF 受体如肌生长抑制素受体的抗原决定簇的特异性结合活性。此外,如上讨论的,本发明的抗体可以是与原肌生长抑制素多肽的一个肽部分,特别是肌生长抑制素功能前区或其功能性肽部分,特异地结合的抗体。应当认识到,下面为 GDF 受体抗体的制备和定性而举例说明的方法进一步可应用于本发明中其他抗体的制备和定性,包括特异地结合肌生长抑制素功能前区的抗体、特异地结合原肌生长抑制素多肽并减少或抑制原肌生长抑制素蛋白水解地切割为肌生长抑制素的抗体,和类似情况。

[0215] 当术语“特异地结合”或“特异性结合活性”用于抗体时,指抗体和特定抗原决定簇间相互作用的解离常数至少大约 1×10^{-6} ,一般地至少大约 1×10^{-7} ,通常至少大约 1×10^{-8} ,特别是至少大约 1×10^{-9} 或 1×10^{-10} 或以下。如此,保留对 GDF 受体抗原决定簇的特异性结合活性的抗体片段 Fab、F(ab')₂、Fd 和 Fv,包括在抗体的定义中。对本发明来说,如果同 TGF- θ 受体或 BMP 受体相比,抗体对肌生长抑制素受体的结合亲和力至少大 2 倍,一般至少大 5 倍,特别地至少大 10 倍,那么该特异地与肌生长抑制素受体抗原决定簇起反应的抗体被认为基本上不与 TGF- θ 受体或 BMP 受体反应。

[0216] 如这里所用的术语“抗体”包括天然存在的抗体以及非天然存在的抗体,包括例如单链抗体、嵌合抗体、双功能抗体和人源化抗体,及其抗原结合片段。这些非天然存在的抗体可以用固相肽合成法构建,可以重组地产生,或通过例如筛选由重链可变区和轻链可变区组成的组合文库而获得(见 Huse 等人, *Science* 246:1275-1281(1989),在此引入作为参考文献)。制造例如嵌合抗体、人源化抗体、CDR-移植抗体、单链抗体和双功能抗体的这些和其它方法是本领域专业技术人员熟知的(Winter 和 Harris, *Immunol. Today* 14:243-246,1993;Ward 等人, *Nature* 341:544-546,1989;Harlow 和 Lane, 抗体:实验室指南 (Antibodies:A laboratory manual)(ColdSpring Harbor Laboratory 出版社,1988);Hilyard 等人, 蛋白工程:实用方法 (ProteinEngineering:A practical approach)(IRL 出版社 1992);Borrabeck, 抗体工程 (AntibodyEngineering),第 2 版,(牛津大学出版社 1995);每一篇文章在此均引入作为参考文献)。

[0217] 与 GDF 受体特异地结合的抗体可以应用该受体作为免疫原产生,并通过使用激活素受体如 Act RIB、Act RIIA 或 Act RIIB、或 BMP 受体如 BMP RII、BMP RIA 和 BMP RIB 去除与受体例如 TGF- θ I 型或 II 型受体发生交叉反应的抗体(见 Messague,见前,1998)。本发明的抗体可很方便地用肌生长抑制素受体的肽部分产生,该部分在 TGF- θ 、激活素或 BMP

受体中不存在。类似地,特异地结合肌生长抑制素功能前区的抗体可以用功能前区或其功能性肽部分作为免疫原而产生。当这种肽无免疫原性时,可以通过下面的方法使之具有免疫原性,将此半抗原偶联到一个载体分子上如牛血清白蛋白(BSA)或钥孔血蓝素(KLH),或以融合蛋白的方式表达肽部分。偶联半抗原到载体分子上的各种其它载体分子和方法是本领域熟知的(见,例如,Harlow和Lane,见前,1988)。

[0218] 如果需要,可以制备用于本发明方法的含有抗体或其它试剂的试剂盒。这种除药物外,试剂盒可以含有一个药物组合物,其中药物可以经重制以给受试者用药。试剂盒也可以含有,例如检测抗体或检测抗体与GDF受体特异结合的试剂。这些用于标记的试剂或另外识别抗体在这里被描述,并为本领域已知。

[0219] 在例如兔、山羊、小鼠或其它哺乳动物中产生多克隆抗体的方法是本领域熟知的(见,例如,Green等人,“多克隆抗血清的产生”(Production of Polyclonal Antisera),在免疫化学规程(Immunochemical Protocols)中(Manson主编,Humana出版社1992),1-5页;Coligan等人,“在兔、大鼠、小鼠和仓鼠中生产多克隆抗血清”(Production of Polyclonal Antisera in Rabbits, Rats, Mice and Hamsters),在Curr. Protocols Immunol.(1992)中,2.4.1节;每一篇文章均在此引入作为参考文献)。此外,单克隆抗体可以用本领域熟知和常规的方法获得(Harlow和Lane,见前,1988)。例如,用肌生长抑制素受体或其抗原决定簇片段免疫小鼠,将脾细胞与一个合适的骨髓瘤细胞系如SP/02骨髓瘤细胞融合,以产生杂交瘤细胞。可以用标记抗原筛选克隆的杂交瘤细胞系,以识别分泌具有合适特异性的单克隆抗体的克隆,表达具有所需特异性和亲合力的抗体的杂交瘤可以被分离并用作抗体的持续来源。可以进一步筛选不能与肌生长抑制素受体特异地结合的抗体。例如,这种抗体用于制备临床应用的标准化试剂盒。表达例如抗肌生长抑制素受体单链抗体的重组噬菌体,也提供了可用于制备标准化试剂盒的抗体。

[0220] 制备单克隆抗体的方法是熟知的(见,例如,Kohler和Milstein,Nature256:495,1975,在此引入作为参考文献;又见,Coligan等人,见前,1992,见2.5.1-2.6.7节;Harlow和Lane,见前,1988)。简要地,单克隆抗体可以通过以下步骤获得:用含抗原的组合物给小鼠注射,通过取出血清样本证实抗体产生的存在,取出脾脏以获得B淋巴细胞,用骨髓瘤细胞融合B淋巴细胞以产生杂交瘤,克隆杂交瘤,选择对抗原产生抗体的阳性克隆,并从杂交瘤培养物中分离抗体。

[0221] 单克隆抗体可以通过各种广为接受的技术从杂交瘤培养物中分离和纯化,例如包括,蛋白-A琼脂糖凝胶亲和层析、尺寸排除层析和离子交换层析(Coligan等人,见前,1992,见2.7.1-2.7.12节和2.9.1-2.9.3节;又见,Barnes等人,“免疫球蛋白G(IgG)的纯化”,在Meth. Molec. Biol.10:79-104(Humana出版社1992)中,在此引入作为参考文献)。体外和体内操作单克隆抗体的方法为本领域专业技术人员熟知。体外操作可以在适合的培养基中进行,如Dulbecco改进的Eagle培养基或RPMI1640培养基,任选地补充一种哺乳动物血清如胎牛血清或微量元素和生长支持添加剂(growth sustaining supplement)如正常小鼠腹腔渗出细胞、脾细胞、骨髓巨噬细胞。体外生产提供了相对纯的抗体制剂,可以大规模生产大量的所需抗体。大规模杂交瘤的培养可以通过均相悬浮培养在气升式反应器、连续搅拌反应器、或固定化或细胞截留培养器中进行。体内操作可以通过将细胞克隆注射进与亲代细胞组织相容的哺乳动物中,例如同系基因小鼠中进行,以引起产抗体肿瘤的生

长。任选地,动物在注射前用烃类,特别是油,例如降植烷(四甲基十五碳烷)进行预致敏。1至3周后,从动物体液中回收所需的单克隆抗体。

[0222] 这里公开的抗体的治疗应用也是本发明的一部分。例如,本发明的抗体也可以来自类人的灵长类抗体。在狒狒中产生治疗用抗体的一般技术可发现于,例如 Goldenberg 等人,国际专利出版物 WO 91/11465(1991);和 Losman 等人, Int. J. Cancer46 :310,1990 中,每一篇文章在此均引入作为参考文献。

[0223] 治疗用抗 GDF 受体抗体也可以来自“人源化”单克隆抗体。通过将来自小鼠免疫球蛋白的重链可变区和轻链可变区的小鼠互补性决定区移入人的可变区中,然后在小鼠对应物的框架区内替代以人的残基,产生了人源化单克隆抗体。使用来自人源化单克隆抗体的抗体成分消除了小鼠恒定区免疫原性带来的潜在问题。克隆小鼠免疫球蛋白可变区的一般技术是已知的(见,例如,Orlandi 等人, Proc. Natl. Acad. Sci., USA 86 :3833,1989,在此整体引入作为参考文献)。生产人源化单克隆抗体的技术也是已知的(见,例如, Jones 等人, Nature 321 :522,1986 ;Riechmann 等人, Nature 332 :323,1988 ;Verhoeyen 等人, Science 239 :1534,1988 ;Carter 等人, Proc. Natl. Acad. Sci., USA 89 :4285,1992 ;Sandhu, Crit. Rev. Biotechnol. 12 :437,1992 ;和 Singer 等人, J. Immunol. 150 :2844,1993 ;每一篇文章均在此引入作为参考文献)。

[0224] 本发明的抗体也可以来自从组合免疫球蛋白文库中分离的人抗体片段(见,例如, Barbas 等人, 方法:免疫学方法手册 (METHODS :A Companion to Methods in Immunology) 2 :119,1991 ;Winter 等人, Ann. Rev. Immunol. 12 :433,1994 ;每一篇文章在此引入作为参考文献)。用于产生人免疫球蛋白噬菌体文库的克隆和表达载体可以从例如 STRATAGENE 克隆系统 (La Jolla, CA) 中获得。

[0225] 本发明的抗体也可以来自人单克隆抗体。这种抗体从转基因小鼠中获得,该小鼠已经被“改造”以对抗原激发反应产生特异的人抗体。在此技术中,人的重链基因座和轻链基因座元件被导入来自胚胎干细胞系的小鼠株中,后者含有靶向破坏的内源性重链基因座和轻链链基因座。转基因小鼠可以合成对人抗原特异的人抗体,且小鼠可被用于产生分泌人抗体的杂交瘤。从转基因小鼠中获得人抗体的方法已被描述,例如, Green 等人, Nature Genet. 7 :13,1994 ;Lonberg 等人, Nature368 :856,1994 ;和 Taylor 等人, Int. Immunol. 6 :579,1994 ;每一篇文章均在此引入作为参考文献。

[0226] 本发明的抗体片段可以通过抗体的蛋白水解或通过编码片段的 DNA 在大肠杆菌中的表达来制备。抗体片段可以用传统的方法,经胃蛋白酶或木瓜蛋白酶消化整个抗体来获得。例如,抗体片段可以用胃蛋白酶通过抗体的酶裂解产生,以提供一个表示为 $F(ab')_2$ 的 5S 片段。此片段可以进一步用巯基还原剂裂解,和任选具有一种对由二硫键断裂所产生的巯基基团具有封阻作用的基团,裂解以产生 3.5SFab' 单价片段。可选择地,用胃蛋白酶的酶裂解直接地产生了 2 个单价 Fab' 片段和一个 Fc 片段(见,例如, Goldenberg, 美国专利 4,036,945 和美国专利 4,331,647,每一篇文章在此引入作为参考文献,以及其中包含的参考文献 ;Nisonhoff 等人, Arch. Biochem. Biophys. 89 :230. 1960 ;Porter, Biochem. J. 73 :119,1959 ;Edelman 等人, Meth. Enzymol. ,1 :422(Academic Press 1967),每一篇文章在此引入作为参考文献 ;又见, Coligan 等人,见前,1992,见 2.8.1-2.8.10 和 2.10.1-2.10.4 节)。

[0227] 假如片段特异地与被完整抗体识别的抗原结合,也可以使用其它裂解抗体的方法,如分离重链以形成单价轻/重链片段,进一步裂解片段,或其它酶学的、化学的或遗传学的技术。例如,Fv片段包括V_H和V_L链的联合。此联合可以是非共价的(Inbar等人,Proc.Natl.Acad.Sci.,USA 69:2659,1972)。可选择地,可变区链可以被一个分子间二硫键连接,或被一个化学药品如戊二醛交联(Sandhu,见前,1992)。优选地,包括V_H和V_L链的Fv片段由肽接头连接。这些单链抗原结合蛋白(sFv)通过构建一个结构基因而制备,结构基因包括编码V_H和V_L功能区的DNA序列,通过一个寡核苷酸连接。结构基因被插入一个表达载体中,随后被导入一个宿主细胞如大肠杆菌中。重组宿主细胞合成单一的多肽链,有接头肽桥接两个V功能区。产生sFvs的方法被描述于,例如Whitlow等人,方法:酶学方法手册(Methods:A Companion to Methods in Enzymology)2:97,1991;Bird等人,Science242:423-426,1988;Ladner等人,美国专利4,946,778;Pack等人,Bio/Technology11:1271-1277,1993;每一篇文章均在此引入作为参考文献;又见,Sandhu,见前,1992。

[0228] 抗体片段的另一个形式是编码单一互补性决定区(CDR)的肽。CDR肽(“最小识别单位”)可以通过构建编码目的抗体CDR的基因来获得。这种基因通过采用例如聚合酶链式反应来制备,以从产抗体细胞的RNA合成可变区(见,例如,Larrick等人,方法酶学方法手册(Methods:A Companion to Methods in Enzymology)2:106,1991,在此引入作为参考文献)。

[0229] 本发明还提供了一个鉴定GDF受体多肽的方法。这样一种方法可以通过下面的步骤进行,例如,将含GDF多肽的成分,与表达全长受体或截短受体的细胞,在足以令GDF结合到受体上的条件下孵育;检测GDF多肽与受体的结合;并分离受体。GDF可以是任何已知的GDF(如GDF-1-16),优选地是GDF-8(肌生长抑制素)或GDF-11。分离受体的方法在下面的实例章节中被更详细地描述。因此,本发明也提供了一个基本纯化的GDF受体,以及比天然存在的GDF受体有较少氨基酸残基的GDF受体肽类和肽衍生物。这种肽类和肽衍生物可在研究肌肉消耗性疾病并开发更有效的治疗方法中用于作为研究和诊断工具。

[0230] 本发明进一步提供了GDF受体变异体。如这里所用的,术语“GDF受体变异体”是指模仿GDF受体至少部分结构的分子。GDF受体变异体可用于减少或抑制GDF结合,从而如这里公开的,缓解一种病理状况。GDF受体变异体的实例包括但不限于,截短型GDF受体如GDF受体的可溶性细胞外功能区;显性失活的GDF受体如显性失活的ActRIIB受体,它缺少激酶活性;或其它截短型或突变体GDF受体。

[0231] 本发明不仅涉及天然存在的GDF受体的肽类和肽衍生物,而且涉及GDF变异体,包括突变体GDF受体,和化学合成的、特异地结合GDF的GDF受体衍生物,如肌生长抑制素。例如,本发明设想改变GDF受体的氨基酸序列。GDF受体可以通过改变编码蛋白的DNA而改变。优选地,采用具有相同或相似特性的氨基酸,仅进行保守氨基酸的改变。氨基酸取代的实例包括:丙氨酸改变为丝氨酸;精氨酸改变为赖氨酸;天门冬酰胺改变为谷氨酰胺或组氨酸;天门冬氨酸改变为谷氨酸;半胱氨酸改变为丝氨酸;谷氨酰胺改变为天门冬酰胺;谷氨酸改变为天门冬氨酸;甘氨酸改变为脯氨酸;组氨酸改变为天门冬酰胺或谷氨酰胺;异亮氨酸改变为亮氨酸或缬氨酸;亮氨酸改变为缬氨酸或异亮氨酸;赖氨酸改变为精氨酸、谷氨酰胺或谷氨酸;蛋氨酸改变为亮氨酸或异亮氨酸;苯丙氨酸改变为酪氨酸、亮氨酸或

蛋氨酸 ; 丝氨酸改变为苏氨酸 ; 苏氨酸改变为丝氨酸 ; 色氨酸改变为酪氨酸 ; 酪氨酸改变为色氨酸或苯丙氨酸 ; 缬氨酸改变为异亮氨酸或亮氨酸。

[0232] 用于本发明的变异体包括 GDF 受体的类似体、同系物、突变蛋白质和模拟体, 它们保留了与其各自的 GDF 特异地结合的能力。在另一个实施方案中, 还设想具有显性失活活性的变异体 GDF 受体, 不管此变异体是否也与其 GDF 特异地起作用。GDF 受体肽类指具有这些活性的 GDF 受体的氨基酸序列部分。变异体可以通过化学修饰、蛋白水解酶消化、或两者联合, 从 GDF 受体本身直接产生。此外, 可以使用遗传工程技术, 以及直接从氨基酸残基合成多肽的方法。

[0233] 肽类可以用这种常用的方法合成, 如 α -氨基的 t-BOC 或 Fmoc 保护。两种方法包含通过从肽的 C 末端开始每一步加入一个氨基酸而逐步合成 (Coligan 等人, 现代免疫学手册 (Current Protocols in Immunology) (Wiley Interscience, 1991), 第 9 单元, 在此引入作为参考文献)。本发明的肽类也可以采用共聚 (苯乙烯-二乙烯基苯), 其中每克聚合物含 0.1-1.0mMol 胺, 通过熟知的固相肽合成方法合成 (Merrifield, J. Am. Chem. Soc., 85 : 2149, 1962 ; Stewart 和 Young, 固相肽合成 (Solid Phase Peptides Synthesis) (Freeman, 旧金山, 1969), 见 27-62 页, 每一篇文章在此均引入作为参考文献)。在化学合成完成时, 该肽类可以解保护, 并通过用液态 HF-10% 苯甲醚在 0°C 处理大约 1/4-1 小时而从聚合物上裂解。在试剂蒸发后, 用 1% 醋酸溶液从聚合物中提取肽类, 然后冻干产生粗原料。通常地, 这样可以在 Sephadex G-15 上, 用 5% 醋酸作为溶剂, 通过如凝胶过滤技术而纯化。柱中的适当组分冻干将产生均一的肽或肽衍生物, 它们然后通过标准的技术而被定性, 如氨基酸分析、薄层层析、高效液相色谱、紫外吸收光谱法、摩尔旋光、溶解度, 并通过固相埃德曼降解来定量。

[0234] 模拟 GDF 受体的结合和功能的非肽化合物 (“模拟体 (mimetics)”) 可以用 Saragovi 等人概括的方法产生 (Science 253 : 792-95, 1991, 在此引入作为参考文献)。模拟体是模拟蛋白二级结构成分分子 (Johnson 等人, “肽旋转模拟体” (Peptide Turn Mimetics), 在 生物技术和药剂学 (Biotechnology and Pharmacy) (Pezzuto 等人主编 ; Chapman 和 Hall, 纽约 1993) 中, 在此引入作为参考文献)。应用肽模拟体的基本原理是, 蛋白的肽骨骼主要地以使氨基酸侧链促进分子间相互作用的方式存在。对于本发明来说, 一个合适的模拟体可以认为相当于一个 GDF 受体的等同物。

[0235] 较长的肽类可以通过“天然化学”连接技术将肽类连接在一起而产生 (Dawson 等人, Science 266 : 776, 1994, 在此引入作为参考文献)。变异体可以通过应用基因组或 cDNA 克隆方法的重组技术而产生。可以应用位点特异和区域定向诱变技术 (Ausuble 等人, 见前, 1989 和 1990 至 1993 增刊), 见第 1 卷, 第 8 章 ; 蛋白工程 (Protein Engineering) (Oxender 和 Fox 主编, A. Liss, Inc., 1987))。此外, 可以应用接头扫描和 PCR 介导技术来诱变 (Erlich, PCR 技术 (Stockton 出版社 1989) ; Ausuble 等人, 见前, 1989 至 1993)。应用上面任何技术的蛋白测序、结构和建模方法公开在上面引用的参考文献中。

[0236] 本发明还提供了阻断 GDF 特异结合到其受体上的 GDF 受体结合剂。例如, 这种结合剂在研究上面描述的肌肉消耗性疾病中, 作为研究和诊断的工具以及用于有效的治疗方法中, 并可以用这里公开的方法鉴定, 例如分子建模方法。此外, 含 GDF 受体结合剂的药组合物可以代表有效的疗法。在本发明的上下文中, 词组“GDF 受体结合剂”表示 : GDF 受体

的天然存在的配体,例如 GDF-1 至 GDF-16 ;GDF 受体的合成配体,或天然或合成配体的适当衍生物。配体的确定和分离是本领域熟知的 (Lemer, Trends Neurosci. 17 :142-146, 1994, 在此引入作为参考文献)。

[0237] 在另一个实施方案中,本发明涉及阻碍 GDF 受体和 GDF 间结合的 GDF 受体结合剂。这种结合剂可以通过竞争性抑制、非竞争性抑制或无竞争性抑制进行阻碍。GDF 受体和一个或多个 GDF 间正常结合的阻碍可以产生有用的药理学效果。

[0238] 本发明还提供了一个鉴定与 GDF 受体结合的组合物方法。此方法包括,将含该组合物的成分与 GDF 受体在足以使组成成分特异地相互作用的条件下孵育,并检测组合物与 GDF 受体的结合。如上所述,结合到 GDF 受体上的组合物包括肽类、肽模拟体、多肽、化学化合物和生物试剂。孵育包括将反应物置于使所测组合物与 GDF 受体接触的条件下,并如会在体内发生的那样,提供适于特异性相互作用的条件。接触可以在溶液中或在固相内进行。如上所述,所测的配体 / 组合物可以任选地是一个组合文库,以筛选大量组合物。在本发明方法中鉴定的组合物,在溶液中或在结合到一个固体支持物上之后,通过一般用于检测特定 DNA 序列的任一方法如 PCR、寡聚物限制性 (Saiki 等人, Bio/Technology 3 :1008-1012, 1985, 在此引入作为参考文献)、等位基因特异的寡核苷酸 (ASO) 探针分析 (Conner 等人, Proc. Natl. Acad. Sci., USA 80 :278, 1983, 在此引入作为参考文献)、寡核苷酸连接测定 (OLA) (Landegren 等人, Science 241 :1077, 1988, 在此引入作为参考文献)、和类似方法 (见 Landegren 等人, Science 242 :229-237, 1988. 在此引入作为参考文献), 被进一步评价、检测、克隆、测序等。

[0239] 为确定一个组合物是否可以功能性地与受体蛋白复合,可以通过监测外源基因编码蛋白的蛋白水平的改变,或通过这里公开的任何其它方法,监测外源基因的诱导。当鉴定一个组合物可以诱导外源基因的转录时,可以得出的结论是,此组合物可以特异地结合到受体蛋白上,该受体蛋白由编码原始样本测定组合物的核酸编码。

[0240] 外源基因的表达可以通过例如功能试验或蛋白产物分析来监测。因此,外源基因是一个提供可分析 / 可测定的表达产物的基因,以使外源基因的表达可以检测。这种外源基因包括但不限于,报告基因如氯霉素乙酰基转移酶基因、碱性磷酸酶基因、 θ -半乳糖激酶、荧光素酶基因、绿色荧光蛋白基因、鸟嘌呤黄嘌呤磷酸核糖基转移酶、碱性磷酸酶,和抗生素抗性基因如新霉素磷酸转移酶 (见上)。

[0241] 外源基因的表达表示组合物和 GDF 受体特异性的相互作用 ;因此,结合性或阻断性组合物可以被鉴定和分离。本发明的组合物可以用已知的、常用的蛋白纯化技术,如提取、沉淀、离子交换层析、亲和层析、凝胶过滤和类似方法,从培养基或细胞中提取和纯化。组合物可以用结合到柱基质上的、修饰的受体蛋白细胞外功能区,通过亲和层析,或通过肝素层析而分离。

[0242] 如上所述,组合化学方法也包括在本发明的筛选方法中,该法可鉴定结合到 GDF 受体上的化合物。因此,筛选方法也具有鉴定变异体、结合剂或阻断剂等用途,它们如果不是物理地 (如,空间地),就是功能性地作为所需要的拮抗剂或激动剂起作用。

[0243] 下面的实施例试图举例说明,但并不限制本发明。

[0244] 实施例 1

[0245] 肌生长抑制素以剂量依赖的方式起作用

[0246] 此实施例表明,肌生长抑制素抑制肌肉生长的活性依赖于肌生长抑制素的体内表达水平。

[0247] 肌生长抑制素是骨骼肌质量的负性调节物 (McPherron 等人, 见前, 1997; McPherron 和 Lee, 见前, 1997)。属肌生长抑制素基因缺失纯合子的肌生长抑制素敲除小鼠, 总体重增加 25-30%。对纯合子敲除小鼠的检查显示, 肌肉质量的增加是由于全身骨骼肌大约增加 100-200%。

[0248] 肌生长抑制素突变杂合子的小鼠, 总体重也增加。然而, 杂合子增加的量小于纯合子增加的量, 且在许多检测中仅有一个年龄和性别组有统计学显著性。为确定杂合子小鼠是否在野生型小鼠和纯合子小鼠之间具有中间的表现型, 对肌肉重量的分析扩展到杂合子小鼠。杂合子小鼠的单个肌肉标本比野生型小鼠重大约 25-50%。这些结果显示, 肌生长抑制素基因缺失的杂合小鼠在野生型小鼠和纯合的肌生长抑制素敲除小鼠之间具有一个中间的表现型, 表明肌生长抑制素在体内产生一个剂量依赖的效应。

[0249] 这些结果提示, 肌生长抑制素活性的调节可以用于治疗肌肉消耗性疾病和其它伴随肌生长抑制素活动的代谢性异常。而且, 肌生长抑制素的剂量依赖作用提示, 可以在肌生长抑制素活性没有达到完全抑制的情况下获得治疗效果, 因此, 假如例如某种活性水平产生了受试者不需要的的作用, 可以调节肌生长抑制素的活性。

[0250] 实施例 2

[0251] 肌生长抑制素的作用随敲除小鼠的年龄而下降

[0252] 此实施例表明, 伴随着突变体小鼠肌肉重量的下降, 野生型小鼠和纯合的肌生长抑制素敲除小鼠间体重的差异减小。

[0253] 肌生长抑制素敲除小鼠在 5 月龄时比野生型小鼠重大约 25-30% (McPherron 等人, 见前, 1997)。然而, 当动物变老时, 这个总体重的差异显著地变小或完全地消失。为了确定此效果是否是由于敲除小鼠体重的相对丢失, 因为例如肌肉衰退, 或因为野生型小鼠获得的体重相对较大, 将肌肉重量作为年龄的函数对其进行了详细分析。

[0254] 在从 2 个月到 17 个月检测的所有年龄中, 纯合突变的小鼠胸肌重量显著高于同窝出生的野生型小鼠。在 5 个月时观察到最显著的差异, 此时突变体小鼠的胸肌重量大约多 200%。尽管在较大的年龄, 胸肌重量略有下降, 但突变体小鼠此肌肉的重量仍然比野生型小鼠大 2 倍。在检测的所有其它肌肉中观察到同样的基本倾向, 包括肱三头肌、四头肌、腓肠肌和跖肌、和胫骨前肌。在雄性和雌性小鼠中都观察到相似的倾向。这些结果表明, 突变体和野生型小鼠间随年龄观察到的总体重差别的下降, 是由于突变体小鼠肌肉重量的轻度下降。

[0255] 实施例 3

[0256] 肌生长抑制素以剂量依赖的方式影响脂肪聚积

[0257] 此实施例表明, 肌生长抑制素敲除小鼠不能聚积脂肪, 且脂肪聚积的减少与肌生长抑制素体内表达的水平相关。

[0258] 如实施例 2 所示, 由于肌生长抑制素突变体肌肉重量的下降没有完全解释所观察到的, 野生型动物最终重量与突变体小鼠大致相同, 于是检测了野生型和突变体小鼠脂肪聚积的量。检测雄性小鼠腹股沟、附睾和腹膜后的脂肪垫。在 2 月龄时, 这些脂肪垫中任何一个的重量在野生型和突变体小鼠间没有差异。到 5 至 6 月龄时, 野生型和杂合的敲除小

鼠都具有较大范围的脂肪垫重量,平均,当动物达到 9 至 10 月龄时,脂肪垫重量增加大约 3 至 5 倍。由于在这些动物中观察到的大范围脂肪垫重量,一些动物显示比其它动物更大的增加(高达 10 倍)。

[0259] 与野生型和杂合的敲除小鼠相反,肌生长抑制素纯合突变体小鼠的脂肪垫重量在相对窄的范围内,在 2 月龄小鼠与 9 至 10 月龄小鼠中实际上是相同的。因此,野生型小鼠中随着年龄而增加的脂肪聚积在纯合的肌生长抑制素敲除小鼠中观察不到。在纯合突变体小鼠中,此脂肪聚积的差异与肌肉重量轻度下降一起,作为年龄的函数,完全解释了观察到的野生型动物最终与突变体动物具有相同的总体重。

[0260] 在 9 至 10 月龄,杂合敲除小鼠的平均脂肪垫重量介于野生型小鼠和纯合突变体小鼠的中间。尽管由于在这些小鼠和野生型小鼠中脂肪垫的重量范围宽,使这个差异没有统计学显著性,但这些结果仍然提示肌生长抑制素对脂肪聚积具有剂量依赖的作用,与其对肌肉生长的作用相似。

[0261] 实施例 4

[0262] 肌生长抑制素对代谢的影响

[0263] 此实施例表明,血清胰岛素和葡萄糖水平、以及代谢活动受肌生长抑制素表达水平的影响。

[0264] 为确定肌生长抑制素突变体小鼠中的骨骼肌肥大和缺少脂肪聚积是否是由于对总体代谢的影响,检测了突变体小鼠的代谢概况。同野生型对照小鼠相比(表 1),肌生长抑制素突变体小鼠的血清甘油三酯和胆固醇水平显著地降低。肌生长抑制素突变体小鼠的血清胰岛素水平也显示较低。但在纯合突变体小鼠和野生型小鼠之间,喂食和空腹葡萄糖水平都没有区别(表 1),且两组小鼠在糖耐量试验中都有正常的反应。结果显示,尽管纯合肌生长抑制素敲除小鼠的血清胰岛素水平低于野生型动物,但它能够保持血清葡萄糖的正常水平。

[0265] 表 1. 血清参数

[0266]

	+/+	-/-	
甘油三酯 (mg/dl)	131.5±16.5	66.8±11.4	p = 0.012
胆固醇 (mg/dl)	138.3±8.1	94.5±6.8	p = 0.0034
喂食葡萄糖 (mg/dl)	114.0±4.8	119.3±5.2	n. s. (p = 0.43)
空腹葡萄糖 (mg/dl)	86.5±3.8	103.3±9.3	n. s. (p = 0.13)

[0267] +/+ 代表野生型小鼠;-/- 代表纯合敲除小鼠

[0268] 为确定代谢率的差异是否可以解释突变体小鼠缺少脂肪聚积,用热量计比较了野生型和突变体小鼠的耗氧率。突变体小鼠的基础代谢率和总代谢率比野生型对照小鼠低。这些结果提示,突变体小鼠缺少脂肪聚积不是由于代谢活动率较高。

[0269] 实施例 5

[0270] 肌生长抑制素在遗传肥胖的小鼠中影响脂肪聚积

[0271] 此实施例表明,在肥胖遗传学模型小鼠中,肌生长抑制素表达缺乏抑制脂肪的聚积。

[0272] 为确定肌生长抑制素活性缺乏是否不仅在正常小鼠,而且在肥胖小鼠中抑制脂肪聚积,检测了代表肥胖遗传学模型的棕灰色的黄色致死 (A^y) 小鼠 (Yen 等人, FASEB J. 8 : 479-488, 1994) 中肌生长抑制素突变的影响。产生致死性黄色和肌生长抑制素突变的双重杂合小鼠,并检测这些双重杂合小鼠杂交的子代小鼠。

[0273] 与 A^y/a , 肌生长抑制素 $+/+$ 小鼠相比, A^y/a , 肌生长抑制素 $-/-$ 双突变小鼠的总体重令人注目地下降 (大约 9 克)。考虑到 A^y/a , 肌生长抑制素 $-/-$ 双突变体的骨骼肌比 A^y/a , 肌生长抑制素 $+/+$ 小鼠多大约 2 至 3 倍,此总体重的下降更突出。双突变体的肌肉比 A^y/a , 肌生长抑制素 $+/+$ 小鼠多大约 10 克,因此,其余组织的总重量减少大约是 19 克。

[0274] 总体重下降由总体脂肪含量减少引起。如表 2 所示,与 A^y/a , 肌生长抑制素 $+/+$ 小鼠相比, A^y/a , 肌生长抑制素 $-/-$ 双突变体的子宫旁和腹膜后脂肪垫重量减少 5 倍至 6 倍。这些结果提示,肌生长抑制素突变的存在显著地抑制肥胖症的脂肪聚积。

[0275] 肌生长抑制素突变的存在也显著地影响葡萄糖代谢。缺少肌生长抑制素突变的野灰色致死性小鼠有很不正常的糖耐量试验结果,其血糖水平常常达到 450 至 600mg/dl 并在 4 个小时时间内仅仅慢慢恢复到基线水平。如前所述,对雌性野灰色致死性小鼠的影响小于雄性小鼠,一些雌性在此试验中的反应几乎正常 (见, Yen 等人, 见前, 1994)。相反,尽管 A^y/a , 肌生长抑制素 $-/-$ 小鼠的糖耐量试验轻度异常,但在 A^y/a , 肌生长抑制素 $+/+$ 小鼠中没有一只观察到很大的异常。

[0276] 这些结果提示,肌生长抑制素突变在野灰色致死性小鼠中至少部分地抑制异常糖代谢的形成。显著地,肌生长抑制素突变杂合子的小鼠同肌生长抑制素 $+/+$ 和肌生长抑制素 $-/-$ 小鼠相比,具有中间的反应,因此证实了肌生长抑制素的剂量依赖作用。

[0277] 实施例 6

[0278] 重组肌生长抑制素的纯化

[0279] 此实施例提供了一个制备和分离重组肌生长抑制素的方法。

[0280] 为阐明肌生长抑制素的生物学活性,纯化了大量肌生长抑制素蛋白以进行生物试验。稳定的产生高水平肌生长抑制素蛋白的中国仓鼠卵巢 (CHO) 细胞系,采用甲氨喋呤选择方案,通过共扩增一个含二氢叶酸还原酶盒的肌生长抑制素表达盒来产生 (McPherron 等人, 见前, 1997)。通过在羟磷灰石、刀豆植物血凝素琼脂糖凝胶、DEAE 琼脂糖和肝素琼脂糖凝胶上的连续分级分离,从最高生产细胞系的条件培养基中纯化肌生长抑制素。银染色分析显示,在这 4 个柱层析步骤后获得的纯化蛋白 (称作“肝素洗脱物”),由分子量大约 35 千道尔顿 (kDa) 和 12kDa 的两种成分组成。

[0281] 纯化的蛋白制剂用各种标准测定,表明是两个肌生长抑制素功能前区肽的复合体,和一个二硫键连接的肌生长抑制素肽成熟 C 末端的二聚体。首先,western 印迹分析,用针对原肌生长抑制素序列特异部分而产生的抗体鉴定作为功能前区的 35kDa 带和作为成熟 C 末端肽的 12kDa 带。其次,在非还原条件下,与直接抗成熟 C 末端肽的抗体起作用的种类,具有与二硫键连接的二聚体一致的电泳迁移率。第三,功能前区与成熟 C 末端肽的摩尔比率大约为 1 : 1。第四,功能前区和成熟 C 末端肽也能经过 4 柱层析步骤共纯化。最后,即使 C 末端区不含公认的 N-连接的糖基化信号,成熟的 C 末端肽结合到刀豆植物血凝素柱上,提示成熟的 C 末端肽由于其与功能前区肽相互作用而结合到该柱上,功能前区含有潜在的 N-连接的糖基化位点。

[0282] 这些结果提示,由遗传学修饰的 CHO 细胞产生的肌生长抑制素,以蛋白酶水解加工的形式分泌,得到的功能前区和成熟 C 末端区非共价地联合,以形成一个复合体,它含有两个功能前区肽和一个二硫键连接的 C 末端蛋白水解片段二聚体,与 TGF- β 的描述相似。在 TGF- θ 复合体中,C 末端二聚体以无活性的潜在形式存在 (Miyazono 等人, J. Biol. Chem. 263 :6407-6415,1988),通过用酸、离液剂、活性氧类、或纤溶酶处理,或通过与其它蛋白相互作用包括血小板反应蛋白和整连蛋白 Iv β 6,可以从此潜在复合体中释放活性种类 (Lawrence 等人, Biochem. Biophys. Res. Comm. 133 :1026-1034,1985 ;Lyons 等人, J. Cell. Biol. 106 :1659-1665,1988 ;Schultz-Cherry 和 Murphy-Ullrich, J. Cell. Biol. 122 :923-932,1993 ;Barcellos-Hoff 和 Dix, Mol. Endocrinol. 10 :1077-1083,1996 ;Munger 等人, 细胞 (Cell) 96 :319-328,1999)。进一步,将纯化的功能前区肽 (还已知为潜伏期伴随肽或 LAP) 加到 TGF- θ 复合体中抑制了纯化的 C 末端二聚体在体外和体内的生物学活性 (Gentry 和 Nash, 生物化学 29 :6851-6857,1990 ;Bottinger 等人, Proc. Natl. Acad. Sci. USA 93 :5877-5882,1996)。

[0283] 由功能前区和成熟 C 末端肽组成的肝素洗脱物用 HPLC C4 反相柱进一步纯化。C 末端二聚体早于功能前区从 HPLC 柱上洗脱,因此,可以分离不含功能前区的 C 末端二聚体。也可以获得主要含功能前区的级分,尽管这些级分含有少量的 C 末端二聚体。这些蛋白中的一些也表现为高分子量复合体。高分子量复合体的性质尚不清楚,但根据有或无还原剂的 western 印迹分析,这些复合体可至少含有一个功能前区肽和一个 C 末端成熟肌生长抑制素肽,两者由一个或更多二硫键连接。事实上;在富含前肽的 HPLC 级分 (HPLC 级分 35-37) 中出现的成熟 C 末端肽,大多数存在于这些高分子量复合体中。这些较高分子量的复合体似乎代表由遗传学修饰的 CHO 细胞分泌的不正确折叠的蛋白。

[0284] 实施例 7

[0285] 肌生长抑制素与激活素受体特异地相互作用

[0286] 此实施例表明,肌生长抑制素特异地结合一个在培养的细胞上表达的激活素 II 型受体,且此特异性结合被肌生长抑制素功能前区抑制。

[0287] 已经鉴定了一些 TGF- β 家族成员的受体,大多数是单一的跨膜丝氨酸 / 苏氨酸激酶 (Massague 和 Weis-Garcia, Cancer Surveys 27 :41-64,1996)。例如,已知激活素 II 型受体 (Act RIIA 和 / 或 Act RIIB) 结合 TGF- β 超家族的成员。缺少 Act RIIB 受体小鼠的表现型显示前 / 后轴型缺陷和肾脏异常,与 GDF-11 敲除小鼠中观察到的非常相似 (McPherron 等人, Nat. Genet. 22 :260-264,1999 ;Oh 和 Li, Genes Devel. 11 :1812-1826,1997)。因为在成熟 C 末端区, GDF-11 和肌生长抑制素 (GDF-8) 的氨基酸序列有 90% 同一性,所以检测了肌生长抑制素特异地与激活素 II 型受体相互作用的能力。

[0288] 肌生长抑制素通过放射性碘化作用标记,用转染了 Act RIIB 表达构建物的 COS 细胞进行结合性研究。肌生长抑制素特异地与转染的 COS 细胞相互作用。肌生长抑制素的结合性被过量未标记的肌生长抑制素以剂量依赖的方式竞争,且在用空载体转染的对照 COS 细胞中显著地降低。对用 BMP RII 或 TGF- β RII 表达构建物转染的细胞没有发生有效的结合。肌生长抑制素对 Act RIIB 转染细胞的结合是可饱和的,通过斯卡查德分析测定的结合亲和力为大约 5nM。

[0289] 受体结合试验也被用于检测肌生长抑制素功能前区抑制成熟 C 末端二聚体与此

系统中 Act RIIB 特异地相互作用的能力。添加纯化的功能前区肽以剂量依赖的方式阻断了 C 末端二聚体结合 Act RIIB 转染的 COS 细胞的能力。这些结果提示,肌生长抑制素功能前区是肌生长抑制素的天然抑制剂。

[0290] 实施例 8

[0291] 增高的肌生长抑制素水平导致重量减轻

[0292] 此实施例表明,体内肌生长抑制素水平升高可以引起相当大的重量减轻。

[0293] 在一组实验中,表达肌生长抑制素的 CHO 细胞被注射给裸鼠。有表达肌生长抑制素的 CHO 细胞瘤的裸鼠,在细胞注射后大约 12 至 16 天的过程中显示出严重的消瘦。此消瘦症状在注射了任何各种对照 CHO 细胞系的裸鼠中没有观察到,这些对照细胞系经过了相似的选择处理但不表达肌生长抑制素。此外,用于转染 CHO 细胞的构建物中的肌生长抑制素编码序列是受金属硫蛋白启动子控制的,当给带肌生长抑制素表达瘤的小鼠供应含硫酸锌的水时,消瘦症状加重。Western 印迹分析显示,带有表达肌生长抑制素的 CHO 细胞的裸鼠,血清中有高水平的肌生长抑制素蛋白。这些结果提示,裸鼠的消瘦症状是对肌生长抑制素水平升高而引起的反应,如下面讨论的,用纯化的肌生长抑制素注射小鼠观察到的相似作用证实了此结果。

[0294] 在带有表达肌生长抑制素的 CHO 细胞的裸鼠中观察到的令人注目的体重减轻,主要地是由于脂肪和肌肉重量都不成比例的减轻。同带对照 CHO 细胞瘤的小鼠相比,白脂肪垫重量(肩胛内白、子宫和腹膜后脂肪)减少超过 90%。肌肉重量也严重地下降,在 16 天时,表达肌生长抑制素小鼠的单个肌肉重量大约是对照小鼠的一半。此肌肉重量的减轻反映在纤维体积和蛋白含量的相应下降。

[0295] 带有表达肌生长抑制素的 CHO 细胞瘤的小鼠还形成了严重的低血糖症。然而,体重减轻和低血糖症不是由于进食的差异,因为在 16 天研究过程中的每一个检测时间段内,所有小鼠消耗相同量的食物。这些结果提示,肌生长抑制素过度表达引起显著的体重减轻,这与患有慢性疾病,如癌症或 AIDS,的病人中发生的恶病质消瘦症状是相似的。

[0296] 随着用低剂量肌生长抑制素的更慢性的给药,观察到了脂肪重量的改变。例如,每天 2 次注射 1 μ g 肌生长抑制素蛋白 7 天,导致许多不同白脂肪垫(肩胛内白、子宫和腹膜后脂肪垫)的重量下降大约 50%,而对褐色脂肪(肩胛内褐)没有显著影响。这些结果证实,肌生长抑制素可以引起体重减轻,并在极端的情况下,引起活体消瘦症状。

[0297] 实施例 9

[0298] 结合到激活素受体的肌生长抑制素的定性

[0299] 此实施例描述了在体内鉴定肌生长抑制素结合激活素受体与肌生长抑制素产生的生物学作用之间关系的方法。

[0300] Act RIIA 或 Act RIIB 敲除小鼠可被用于在体内证实 Act RIIA 或 Act RIIB 是肌生长抑制素的受体。对这些小鼠详细的肌肉分析可以确定激活素受体敲除是否与肌肉纤维数量或大小的改变有关。由于 Act RIIA/Act RIIB 双纯合突变体在胚胎形成过程中早期死亡(Song 等人, *Devel. Biol.* 213 :157-169, 1999),因此只能检测各种纯合/杂合的组合。但是,可以产生组织特异的或条件性敲除小鼠,使得两种基因可以仅在肌肉中都“删除”,因而可以在出生后检测双纯合的敲除小鼠。

[0301] 可以随小鼠的年龄检测对脂肪组织的作用,以确定在敲除小鼠中脂肪细胞数量或

这些脂肪细胞中的脂质聚积是否改变。用从胶原酶处理的组织中制备的细胞悬液测定脂肪细胞数量和大小 (Rodbell, *J. Biol. Chem.* 239 :375-380, 1964 ;Hirsch 和 Gallian, *J. Lipid Res.* 9 :110-119, 1968)。通过测量干尸重量以及提取脂质后残余的干尸重量来确定动物的总脂质含量 (Folch 等人, *J. Biol. Chem.* 226 :497-509, 1957)。

[0302] 还可以检测各种血清参数,包括喂食和空腹葡萄糖和胰岛素、甘油三酯、胆固醇和苗条蛋白。如上面公开的,血清甘油三酯和血清胰岛素在肌生长抑制素突变体动物中下降。还可以用糖耐量试验检测激活素受体敲除小鼠对外源葡萄糖负荷的反应能力。如上面公开的,5月龄野生型和肌生长抑制素突变体小鼠对葡萄糖负荷的反应基本上是相同的。通过测量小鼠年老时的这些参数可以扩展此观察。在糖耐量试验过程中的不同时间还可以测量血清胰岛素水平。

[0303] 也可以用一个热量计 (Columbus Instruments) 监测基础代谢率。如上面公开的,3月龄时,肌生长抑制素突变体小鼠的代谢率较其对应的野生型低。此分析可以扩展到年老小鼠,而且还可以测量这些动物的呼吸商。通过测量基础体温以及当置于4°C时它们保持体温的能力,可以确定保持正常生热作用的能力。也可以检测褐色脂肪的重量,及 UCP1、UCP2 和 UCP3 在褐色脂肪、白脂肪、肌肉和其它组织中的表达水平 (Schrauwen 等人, 1999)。

[0304] 可以监测相对于体重增加的食物摄取量,并可以计算喂养效率。此外,可以监测在高脂膳食下动物的体重增加。保持高脂膳食的野生型小鼠迅速地聚积脂肪,然而这里公开的结果提示,激活素受体突变动物将相对地保持瘦态。

[0305] 这些研究的结果可以提供肌生长抑制素对小鼠影响的更完全的概观,特别是有关其总体代谢状态,从而理解关于肌生长抑制素敲除小鼠抑制脂肪聚积的能力是否是肌肉中肌生长抑制素突变的合成代谢作用,它引起能量应用的变换,使得很少能量可被用来以脂肪的形式储存。例如,脂肪聚积的减少可以是由于生热作用速度增加。这些结果也会提供一个基础,以比较在肥胖症和 II 型糖尿病的不同遗传学模型范围内,肌生长抑制素活性的作用。

[0306] 实施例 10

[0307] 在肥胖症和 II 型糖尿病遗传学模型中,肌生长抑制素作用的特性

[0308] 此实施例描述了确定肌生长抑制素在治疗肥胖症或 II 型糖尿病中的作用的方法。

[0309] 同野生型小鼠相比,肌生长抑制素突变体小鼠的总体脂肪聚积令人注目地减少,提示可以改变肌生长抑制素活性以治疗或预防肥胖症或 II 型糖尿病。可以在这些代谢性疾病的几个特性清楚的小鼠模型背景中检测肌生长抑制素突变的作用,这些模型包括,例如“肥胖”小鼠 (ob/ob)、“糖尿病”小鼠 (db/db)、和野灰色黄色致死性 (A^y) 突变品系。上面描述的每个参数和试验在所有这些品系实际上是异常的 (见,例如, Yen 等人,见前, 1994, Friedman 和 Halaas, *Nature* 395 :763-770, 1998)。通过构建双重突变体,然后对双重突变体动物与合适的仅携带 ob/ob、db/db、或野灰色黄色致死性突变的对照同窝出生仔一起进行各种试验,可以检测肌生长抑制素突变在携带这些其它突变的小鼠中减缓或抑制这些异常形成的能力。

[0310] 如上面公开的, A^y 小鼠中肌生长抑制素突变与肌生长抑制素突变 A^y 小鼠大约 5 倍的脂肪聚积抑制相关,并如糖耐量试验所评定的,与部分抑制糖代谢异常的形成有关。这些

结果可以扩展到包括另外各种年龄的动物,且可以用 ob/ob 和 db/db 突变体进行相似的研究。由于这些突变都是隐性的,可以产生肌生长抑制素突变和 ob 或 db 突变双重纯合的小鼠。为检测肌生长抑制素功能部分缺失在这些遗传学模型系统中的作用,也检测了 ob 或 db 突变纯合和肌生长抑制素突变杂合的小鼠。已经产生了肌生长抑制素和 ob 突变双重杂合的小鼠,可以检测来自这些双重杂合小鼠交配的子代,特别是关于脂肪聚积和糖代谢。在肥胖突变体中,这些异常之一或两者都部分抑制,可能提示肌生长抑制素是治疗肥胖症和 II 型糖尿病的靶标。

[0311] 实施例 11

[0312] 表达能够影响肌生长抑制素活性的显性失活多肽的转基因小鼠的特性

[0313] 此实施例描述了通过表达显性失活多肽,在出生后鉴定肌生长抑制素作用的方法,该显性失活多肽可以阻断肌生长抑制素表达或肌生长抑制素信号转导。

[0314] 肌生长抑制素抑制剂

[0315] 在出生后调节肌生长抑制素活性,可被用来确定肌生长抑制素对肌肉纤维数量(增生)和肌肉纤维大小(肥大)的作用。条件性肌生长抑制素敲除小鼠,其肌生长抑制素基因在动物生命的限定时间被删除,可被用于这些研究。tet 调节物和 cre 重组酶一起提供了产生这种小鼠的系统。在此系统中,cre 的表达是由给予强力霉素诱导的。

[0316] 也可以产生一个从可诱导的启动子表达肌生长抑制素抑制剂的转基因小鼠,这样可以在动物生命过程中的限定时间减弱或抑制肌生长抑制素的活性。四环素调节物被用于产生这种转基因小鼠,其中肌生长抑制素的表达由强力霉素诱导。

[0317] 一个 tet 系统的修饰,应用杂交逆向 tet- 反式激活剂(VP16 活化功能区与突变体逆向 tet 阻遏物的融合蛋白)和杂交逆向 tet- 反式阻遏物(哺乳动物 Kox1 的 KRAB 阻遏物功能区与天然 tet 阻遏物的融合蛋白)的共表达,可以特别地用来产生转基因小鼠(Rossi 等人, Nat. Genet. 20 :389-393,1998 ;Forster 等人, Nucl. Acids Res. 27 :708-710,1999)。在此系统中,杂交逆向 tet- 反式阻遏物结合 tet 操纵基因序列,并仅在四环素存在时激活转录;杂交逆向 tet- 反式阻遏物结合 tet 操纵基因序列并仅在没有四环素时抑制转录。通过这两个融合蛋白的共表达,靶启动子的基础活性在没有四环素时被 tet- 反式阻遏物沉默,并在给予四环素时被逆向 tet- 反式激活剂激活。

[0318] 可以产生两型转基因系。在第一型,在肌肉特异的启动子,例如肌肉肌酸激酶启动子(Sternberg 等人,见前,1988)或肌球蛋白轻链增强子/启动子(Donoghue 等人,见前,1991)的控制下,转基因编码肌生长抑制素抑制剂多肽。筛选在骨骼肌中 tet 调节物特异性表达的单个转基因系,为两个启动子中的每一个选择几个独立的系并检测,以证实观察到的任何作用不是由于例如整合位点特异的作用。已经构建了一个含 2 个在肌球蛋白轻链启动子/增强子控制下的 tet 调节物的构建物,并可以用于前核注射。在第二型系中,转基因含有在最小 CMV 启动子控制下的肌生长抑制素抑制剂多肽,它进一步含有 tet 操纵基因序列。

[0319] 肌生长抑制素抑制剂可以是显性失活形式的肌生长抑制素,或肌生长抑制素功能前区,如这里公开的,后者可以抑制肌生长抑制素活性。TGF- β 家族成员的显性失活形式已经被描述(见,例如, Lopez 等人, Mol. Cell Biol. 12 :1674-14679,1992 ;Wittbrodt 和 Rosa, Genes Devel. 8 :1448-1462,1994),并含有例如突变的蛋白水解切割位点,从而防止

蛋白被加工成生物活性形式。当在细胞内与内源性野生型基因共表达时,突变体蛋白与野生型蛋白一起形成无功能杂合二聚体,因此作为显性失活起作用。已经构建了一个在原肌生长抑制素切割位点内含有突变的突变体肌生长抑制素多肽,并在 293 细胞中,通过以不同的比率与野生型肌生长抑制素共表达,可以检测显性失活的作用。可以用 western 印迹分析检测暂时转染了构建物的 293 细胞的条件培养基,并可以检测突变体阻断成熟 C 末端二聚体形成的能力。

[0320] 还可以使用仅编码肌生长抑制素功能前区的表达构建物。如上面公开的,功能前区与成熟 C 末端二聚体形成一个紧密的复合体,并阻断成熟 C 末端肌生长抑制素二聚体在表达受体的培养细胞中结合 Act RIIB 的能力。通过与 TGF- β 类比,肌生长抑制素功能前区也能够以无活性潜在复合体形式在体内保留成熟 C 末端二聚体。

[0321] 这些转基因动物可以与表达 tet 调节物的动物一起培育,产生含 tet 调节物和抑制剂靶标的构建物的双重转基因系。这些双重转基因系可以被筛选,以找到其中所有不同的成分都合适地表达的系。在饮水中给予强力霉素前和后,采用获取自各系中有代表性小鼠的各种肌肉和对照组织的 RNA,进行 Northern 印迹分析,可以用来鉴定这种转基因系。在没有强力霉素的情况下不表达任何转基因,而有强力霉素时仅在肌肉中表达转基因的转基因系将被选择。

[0322] 对选择的转基因动物给予强力霉素并检测其对肌肉质量的影响。给怀孕的母体使用强力霉素以在胚胎形成过程中诱导抑制剂的表达。在转基因动物发育过程中,阻断肌生长抑制素活性的效果可以与肌生长抑制素敲除小鼠中观察到的效果相比。由于同肌生长抑制素最初表达的时间相比,驱动 tet 调节物表达的启动子可以在发育过程的较晚时间被诱导,因此对转基因小鼠肌肉质量的影响可以与肌生长抑制素敲除小鼠中发生的效果相比。

[0323] 通过在出生后不同的时间给双重转基因小鼠使用强力霉素,可以检测出生后抑制肌生长抑制素活性的效果。强力霉素处理可以在例如 3 周龄时开始,并可以在 5 月龄时分析动物,此时,肌生长抑制素敲除小鼠与野生型小鼠肌肉重量的差异最大。检测了抑制剂对动物肌肉质量的影响。也可以对肌肉进行组织学检测以确定对纤维数量和纤维大小的影响。此外,可以进行转基因动物各种肌肉的纤维类型分析,以确定对 I 型和 II 型纤维是否具有选择性作用。

[0324] 强力霉素可以用不同的剂量并在不同的时间用药,以鉴定肌生长抑制素抑制剂的作用。双重转基因小鼠也可以被慢性地维持使用强力霉素,然后如上所述检测对脂肪垫和其它相关代谢参数的影响。这些研究的结果可以证实,出生后调节肌生长抑制素活性可以增加肌肉质量或减少脂肪聚积,因而提示,靶向肌生长抑制素可以用于临床治疗各种肌肉消耗性和代谢性疾病。

[0325] 肌生长抑制素

[0326] 也可以检测含肌生长抑制素转基因的转基因小鼠,并可以将肌生长抑制素表达产生的影响与含表达肌生长抑制素的 CHO 细胞的裸鼠中的观察效果进行比较。与上面描述的相似,肌生长抑制素可被置于条件性 (tet) 和组织特异性调控元件的控制之下,可以检测转基因小鼠中肌生长抑制素的表达以确定发生的消瘦症状是否与裸鼠中的观察相似。肌生长抑制素转基因可以包括例如来自 SV40 的加工信号,使转基因能够与内源性肌生长抑制素基因相区别。

[0327] 在给予强力霉素后的各时间点,可以从肌生长抑制素转基因小鼠中分离血清样本,并可以测定血清中肌生长抑制素转基因产物的水平。随着时间监测动物的总体重以确定动物是否表现显著的重量减轻。此外,分离和称重了单个肌肉和脂肪垫,并在选择的肌肉样本中测定肌肉纤维的数量、大小和类型。

[0328] 肌生长抑制素转基因表达的水平可以因给予动物的强力霉素的剂量不同而变化。转基因表达可以被监测,采用例如,肌肉中转基因 RNA 水平的 northern 印迹分析,或血清中肌生长抑制素蛋白的水平。测定肌生长抑制素转基因表达的特定水平可以确定与肌生长抑制素引起的消瘦程度的相关性。转基因系也可以与肌生长抑制素敲除小鼠杂交,以产生转基因为唯一肌生长抑制素表达源的小鼠。可以监测发育过程中各时间点的肌生长抑制素表达,并确定肌生长抑制素对纤维数量、纤维大小、和纤维类型的影响。获得可精确和迅速控制肌生长抑制素表达的小鼠提供了一个有力的工具,以进一步鉴定肌生长抑制素信号转导通路,并检测各种潜在地可能用于调节肌生长抑制素信号转导的试剂的效果。

[0329] 肌生长抑制素信号转导的效应物

[0330] 可以产生含肌生长抑制素信号转导通路任一显性失活形式的转基因小鼠,该信号转导通路可以包括 TGF- θ 信号转导通路的成分,它特异地在骨骼肌中表达。如这里公开的,通过由激活素 II 型受体诱导的通路介导信号转导的 Smad 蛋白,可能涉及肌生长抑制素信号转导。

[0331] Act RIIB 可以结合与肌生长抑制素高度相关的 GDF-11 (McPherron 等人,见前,1997;Gamer 等人,见前,1999;Nakashima 等人, *Mech. Devel.* 80 :185-189,1999),且能够结合抑制 Smad2、Smad3、和 Smad4 的 c-ski 的表达令人注目地影响肌肉生长 (Sutrave 等人,见前,1990;Berk 等人,见前,1997;又见, Luo 等人,见前,1999;Stroschein 等人,见前,1999;Sun 等人,见前,1999a 和 b;Akiyoshi 等人,见前,1999)。如这里公开的,肌生长抑制素特异地与 Act RIIB 相互作用,因此能够发挥其生物学作用,至少部分通过体内结合激活素 II 型受体并激活 Smad 信号通路。

[0332] 采用在 Act RIIB/Smad 信号转导通路的特定点上阻断或能够被阻断的转基因小鼠系,可以检测 Smad 信号转导通路在调节肌肉生长中的作用。肌肉肌酸激酶启动子或肌球蛋白轻链增强子 / 启动子可被用于驱动 Smad 信号转导通路各种抑制剂的表达。

[0333] 用于此系统中的抑制剂可以包括例如,卵泡素抑制素;显性失活的 Act RIIB 受体;显性失活的 Smad 多肽如 Smad3;c-ski;或抑制性 Smad 多肽如 Smad7。卵泡素抑制素可以结合并抑制某种 TGF- β 家族成员的活性,包括 GDF-11 (Gamer 等人,见前,1999)。例如通过表达细胞外功能区,特别是 Act RIIB 细胞外功能区的可溶性形式,或通过表达缺少激酶功能区或含有使突变体受体缺少激酶活性的突变的截短型 Act RIIB 受体,可以获得显性失活形式的激活素 II 型受体。Smad7 作为一个抑制性 Smad 起作用,它可以阻断由激活素、TGF- β 、和 BMP 诱导的信号转导通路。例如,通过突变 Smad3 的 C 末端磷酸化位点可以构建 Smad3 的显性失活形式,从而阻断 Smad3 的功能 (Liu 等人,见前,1997)。在转基因小鼠中,c-ski 过度表达已经发现与肌肉肥大有关联 (Sutrave 等人,见前,1990)。

[0334] 可以制备转基因小鼠并检测每一个起始系,以使转基因适当的、肌肉特异性的表达。检测被选择小鼠的总体重、个体肌肉重量、及肌肉纤维大小、数量和类型。如上所述,那些显示对肌肉质量有明显影响的系可以进一步检测脂肪聚积和其它相关的代谢参数。应

用这些不同的试剂寻靶激活素受体 /Smad 信号转导通路中的特定步骤,是特别有信息价值的,因为不同试剂的信号转导通路在不同的步骤中有重叠。例如,卵泡素抑制素结合并抑制激活素和 GDF-11 活性但不影响 TGF- β ,而显性失活的 Smad3 可以通过激活素和 TGF- β 受体阻断信号传导。Smad7 可能是更多向性的,因为它还阻断通过 BMP 受体的信号传导。本研究可使调节肌生长抑制素活性的特定靶位得以确定,从而提供了开发药物和其它试剂的各种策略,这些药物和试剂调节肌生长抑制素的信号转导并进而调控肌生长抑制素的活性。

[0335] 特别地,这里描述的转基因系可以用于确定出生后阻断肌生长抑制素功能或 Smad 信号转导通路对形成肥胖症或 II 型糖尿病的作用。例如,抑制性转基因可以杂交进 ob/ob、db/db 和 A^y 突变体小鼠中。在没有强力霉素时,抑制剂转基因不表达,因此动物与每个亲代突变体小鼠没有差别。在强力霉素存在的情况下,抑制剂被表达并可以阻断肌生长抑制素活性。可以检测在这些突变体动物中阻断肌生长抑制素活性对形成代谢性异常的影响。

[0336] 抑制剂的表达可以在较早年龄时诱导,例如在 3 周龄,以使效果最大化。此外,可以在代谢性异常严重到不可逆转之前阻断肌生长抑制素活性。动物可以维持给予强力霉素,并在不同的年龄用上面描述的试验测评,包括涉及脂肪聚积和葡萄糖代谢的试验。在 ob/ob、db/db 和 A^y 突变体动物中,可以发现在一个或多个试验结果出现异常的年龄组中存在延迟。可以用已经形成一些肥胖症或 II 型糖尿病迹象的老年动物进行相似的研究,并测定肌生长抑制素活性阻断对各种参数的影响,包括脂肪重量和葡萄糖代谢。这些研究结果可以进一步确定特异性靶标,它们在预防或治疗肥胖症或 II 型糖尿病的努力中可以被操作控制。

[0337] 实施例 12

[0338] 肌生长抑制素特性对诱导恶病质的影响

[0339] 此实施例描述了确定肌生长抑制素信号转导在恶病质形成和进展中的作用的的方法。

[0340] 在正常个体中,激活素受体和 Smad 通路可以构成涉及介导肌生长抑制素活性的至少部分信号转导通路,因此,可以涉及介导个体中因肌生长抑制素过高水平而发生的影响。如这里公开的,例如恶病质,至少可以部分地通过异常的高水平肌生长抑制素介导。如此,通过 Smad 通路调控信号转导的方法可以提供一个开发药物的新策略,以开发治疗全身肌肉消耗、特别是恶病质的药物。

[0341] Smad 信号转导通路在恶病质中的作用,可以通过检测上面描述的各种转基因系对恶病质的易感性而确定,恶病质可以由以下因子诱导,如白介素 -6 (IL-6 ;Black 等人, 内分泌学 128 :2657-2659, 1991, 在此引入作为参考文献)、肿瘤坏死因子 -I (TNF-I ;Oliff 等人, 细胞 50 :555-563, 1987, 在此引入作为参考文献)、或某种肿瘤细胞。就 IL-6 和 TNF-I 来说,抑制剂转基因可被杂交进裸鼠背景中,然后用产生 IL-6 或 TNF-I 的 CHO 细胞激发动物,当 IL-6 或 TNF-I 以此方式过度表达时,导致裸鼠消瘦。可以采用上面描述的、用于制备过度产生肌生长抑制素的细胞的方法,制备过度产生 IL-6 或 TNF-I 的 CHO 细胞。例如, TNF-I cDNA 可被克隆进 pMSXND 表达载体 (Lee 和 Nathans, J. Biol. Chem. 263 :3521-3527, 1988), 然后可以在逐渐增加甲氨喋呤浓度的情况下,逐步选择携带表达构建物扩增拷贝的细胞。

[0342] 肿瘤细胞如 Lewis 肺癌细胞 (Matthys 等人, Eur. J. Cancer 27 :182-187, 1991, 在此引入作为参考文献) 或结肠 26 腺癌细胞 (Tanaka 等人, J. Cancer Res. 50 :2290-2295,

1990,在此引入作为参考文献),可以导致小鼠恶病质,也可被用于这些研究。当这些细胞系在小鼠中长成肿瘤时,引起严重的消瘦。于是,可以在这里描述的各种转基因小鼠中检测这些肿瘤的影响。应当认识到,各种肿瘤细胞将仅在一定的遗传背景下生长。例如,Lewis 肺癌细胞常规地在 C57BL/6 小鼠中生长,而结肠 26 癌细胞常规地在 BALB/c 小鼠中生长。因此,转基因可被回交进这些或其它遗传学背景中,以使肿瘤细胞生长。

[0343] 可以监测各种参数包括总体重、单个肌肉重量、肌纤维大小和数量、食物摄取量和血清参数包括葡萄糖水平。此外,可以检测血清肌生长抑制素水平和肌肉中肌生长抑制素 RNA 的水平,以证实增高的肌生长抑制素表达与恶病质相关。这些研究的结果可以证实,在这些实验模型中,肌生长抑制素的作用是恶病质诱导剂的下游。结果还可证实,在这些模型中,Smad 信号转导通路对于恶病质的形成是必须的,并可表明通过调节 Smad 信号传送可以在恶病质的治疗中获得治疗益处。

[0344] 实施例 13

[0345] 生长分化因子-8(GDF-8)和 GDF-11 受体的鉴定和特性描述

[0346] 此实施例描述了 GDF-8(肌生长抑制素)和 GDF-11 的细胞表面受体鉴定和特性描述的方法。

[0347] 纯化的 GDF-8 和 GDF-11 蛋白将主要用于评定生物学活性。为鉴定 GDF-8 和 GDF-11 作用的潜在靶细胞,要寻找表达其受体的细胞。为此,纯化的蛋白将用氯胺 T 方法进行放射性碘化,此方法已成功地用于标记此超家族的其它成员如 TGF- β (Cheifetz 等人,见前,1987)、激活素 (Sugino 等人, *J. Biol. Chem.* 263 :15249-15252,1988) 和骨形态发生蛋白 (Paralkar 等人, *Proc. Natl. Acad. Sci., USA* 88 :3397-3401,1991),以进行受体结合研究。每个成熟加工形式的 GDF-8 和 GDF-11 都含有多个酪氨酸残基。将采用两个不同的方法鉴定这些蛋白的受体。

[0348] 一个方法将确定受体的数量、亲合力和分布。在细胞培养中生长的全细胞、胚胎或成年组织的冷冻切片、或从组织或培养的细胞中制备的全膜组分将与标记的蛋白孵育,并将测定结合蛋白的数量或分布。对于涉及细胞系或细胞膜的试验,结合量的测定将通过在数次冲洗后测量结合到培养皿上的细胞上的放射活性,或就细胞膜而言,在离心或用滤器保留膜后,测量与膜一起沉积的放射活性。对于涉及原代培养的试验,细胞数量可以更有限,将通过用照相乳剂覆盖而直接地显示结合部位。对于涉及冷冻切片的试验,将通过暴露这些切片于高分辨 Beta-max 超级胶片而显示配体结合部位;如果需要更好的定位,将把切片浸泡于照相乳剂中。对于所有这些试验,将通过添加过量未标记蛋白作为竞争者来确定特异性结合(例如,见 Lee 和 Nathans,见前,1988)。

[0349] 第二个方法将用生物化学的方法描述受体的特性。细胞膜制剂或在细胞培养中生长的潜在靶细胞将与标记的配体一起孵育,受体/配体复合物将用双琥珀酰亚胺辛二酸酯共价地交联,双琥珀酰亚胺辛二酸酯已经被普遍地用于鉴定各种配体的受体,包括 TGF- β 超家族成员的受体 (Massague 和 Like, *J. Biol. Chem.* 260 :2636-2645,1985)。交联的复合物通过 SDS 聚丙烯酰胺凝胶电泳被分离,以寻找过量未标记蛋白的条带,在没有过量未标记蛋白存在时不会有该标记条带;而在过量未标记蛋白存在时会有该标记条带。推定受体的分子量将通过减去配体的分子量而估算。这些试验将解决的一个重要问题是,GDF-8 和 GDF-11 是否像其它 TGF- β 超家族成员一样,通过 I 型和 II 型受体传递信号 (Massague 和

Weis-Garcia, 见前, 1996)。

[0350] 一旦检测这些分子的受体的一个方法已经完成, 将进行更详细的分析以确定结合亲和性和特异性。将使用斯卡查德分析确定结合部位的数量和解离常数。通过进行 GDF-8 和 GDF-11 之间的交叉竞争分析, 将有可能确定它们是否能够结合到同一个受体上及它们的相对亲和性。这些研究将给出关于分子是否经过相同或不同的受体发送信号的提示。将用其它 TGF- β 家族成员进行竞争试验以确定特异性。这些配体中的某一些从商业途径获得, 其它一些从 Genetics Institute, Inc. 获得。

[0351] 对于这些试验, 将检测各种胚胎和成年组织以及细胞系。根据 GDF-8 特异地表达在骨骼肌上和 GDF-8 敲除小鼠的表现型, 最初的研究集中在胚胎和成年的肌肉组织以制备细胞膜, 并用冷冻切片研究受体。此外, 如所述, 将在妊娠期的不同天数从胚胎中分离和培养成肌细胞, 或从成年肌肉中分离和培养卫星细胞 (Vivarelli 和 Cossu, *Devel. Biol.* 117 : 319-325, 1986 ; Cossu 等人, *Cell Diff.* 9 : 357-368, 1980)。在细胞培养后各天, 将在这些原代细胞上进行结合研究, 通过放射自显影定位结合部位, 使得结合部位可以与各种肌原性的标记物如肌肉肌球蛋白共同被定位 (Vivarelli 等人, *J. Cell. Biol.* 107 : 2191-2197, 1988), 并将与细胞的分化状态如多核肌小管的形成与结合相关联。除应用原代细胞以外, 将应用细胞系寻找受体。特别是, 最初集中在 3 个细胞系上, C2C12、L6 和 P19。C2C12 和 L6 成肌细胞在细胞培养中自发分化并依赖特定的生长条件形成肌小管 (Yaffe 和 Saxel, 见前, 1977 ; Yaffe, 见前, 1968)。P19 胚胎瘤细胞可被诱导分化成各种细胞类型, 包括在有 DMSO 的情况下分化成骨骼肌细胞 (Rudnicki 和 McBurney, 畸胎癌和胚胎干细胞: 一个实用方法 (Teratocarcinomas and Embryonic Stem Cell: A practical approach) (E. J. Robertson, IRL 出版社, 剑桥 1987)。将在各种生长条件下及各种分化阶段对这些细胞系进行受体结合研究。尽管最初的研究集中在肌肉细胞上, 但也将检测其它组织和细胞类型的 GDF-8 和 GDF-11 受体的存在情况。

[0352] 重组人 GDF-8 (rhGDF-8) 同型二聚体将被用在这些结合研究中。RhGDF-8 用 CHO 细胞表达, 并纯化到大约 90% 的纯度。rhGDF-8 具有期望的 25kDa 至 27kDa 的分子量, 一旦还原, 还原成 12kDa 的单体。在受体-配体结合试验中采用 I-125 标记的 GDF-8, 两个成肌细胞系, L6 和 G-8, 与 GDF-8 结合。结合是特异性的, 因为非标记的 GDF-8 有效地竞争标记配体的结合。解离常数 (Kd) 是 370pM, 且 L6 成肌细胞具有较高数量 (5,000 受体 / 细胞) 的细胞表面结合蛋白。GDF-11 (BMP-11) 与 GDF-8 是高度同源的 (> 90%)。受体结合研究显示, GDF-8 和 GDF-11 结合到 L6 成肌细胞上的相同结合蛋白上。重要的是要确立 GDF-8 是否与已知的 TGF- θ 受体结合。TGF- θ 不竞争 GDF-8 的结合, 提示 GDF-8 受体与 TGF- θ 受体不同。GDF-8 受体不在所有的成肌细胞系上表达, 包括 4 个成肌细胞系, C2C12、G7、MLB13MYCc14 和 BC3H1, 它们不结合 GDF-8。

[0353] 可以获得编码 GDF-8 和 GDF-11 受体的基因。作为理解 GDF-8 和 GDF-11 发挥其生物学作用的机制的第一步, 重要的是克隆编码其受体的基因。从上面的实验中, 对 GDF-8 和 GDF-11 是否结合相同的或不同的受体将会更清楚。还会有关于这些受体的组织和细胞类型分布的重要信息。应用此信息, 将采取两个不同的方法克隆受体基因。

[0354] 第一个方法将使用表达克隆策略。事实上, 这是最初由 Mathews 和 Vale (细胞 65 : 973-982, 1991) 及 Lin 等人 (细胞 68 : 775-785, 1992) 用来克隆第一个激活素和 TGF- β 受

体的策略。从表达最高相对数目的高亲和力结合部位的组织或细胞类型中将获得 Poly-A 选择的 RNA, 并用于在哺乳动物表达载体 pcDNA-1 中制备一个 cDNA 文库, pcDNA-1 含有一个 CMV 启动子和 SV40 复制起点。文库将被铺平板, 来自每个平板的细胞将汇入肉汤并冷冻。来自每个库的等分试样将生长以制备 DNA。每一个库将在小室玻片中被暂时地转染进 COS 细胞, 转染的细胞将与碘化的 GDF-8 或 GDF-11 一起孵育。在冲走未结合的蛋白后, 配体结合部位将通过放射自显影显示。一旦识别一个阳性库, 来自该库的细胞将以较低的密度重新铺平板, 并将重复此步骤。然后将阳性库铺平板, 将每个克隆精选进并如所述再分析 (Wong 等人, *Science* 228 :810-815, 1985)。

[0355] 最初, 将使用大小为 1500 个菌落的库进行筛选。为确信在这样复杂的混合物中鉴定出一个阳性克隆, 将进行采用 TGF- β 和克隆的 II 型受体的对照实验。编码 TGF- β II 型受体的编码序列将被克隆进 pcDNA-1 载体中, 用此构建物转化的细菌将与来自我们文库的细菌以各种比例混合, 包括 1 : 1500。然后, 由此混合物制备的 DNA 将转染进 COS 细胞, 与碘化的 TGF- β 一起孵育, 并通过放射自显影显示。如果在 1 : 1500 比例下观察到阳性信号, 那么将筛选含有 1500 个克隆的库。否则, 将使用与相应比率相对应的较小的库, 在这种情况下, 在对照实验中, 该操作过程足以敏感到鉴定到一个阳性信号。

[0356] 大多数 TGF- β 超家族成员的受体已经被鉴定属于跨膜丝氨酸 / 苏氨酸激酶家族 (Massague 和 Weis-Garcia, 见前, 1996), 利用此事实还将使用第二个相似的尝试克隆 GDF-8 和 GDF-11 受体的策略。由于这些受体的胞浆功能区在序列上相关, 将使用简并 PCR 探针以克隆此受体家族的成员, 它们在含有 GDF-8 和 GDF-11 结合部位的组织中表达。事实上, 这是已经被用于鉴定此受体家族大多数成员的方法。一般的策略将是, 设计与已知受体保守区相应的简并引物, 将这些引物用于对制备自合适的 RNA 样本 (最可能来自骨骼肌) 的 cDNA 进行 PCR, 亚克隆 PCR 产物, 最后对单个亚克隆测序。当序列被鉴定时, 它们将被用作杂交探针以从进一步分析中消除重复的克隆。然后, 将测试被鉴定的受体结合纯化的 GDF-8 和 GDF-11 的能力。由于此筛选将仅产生小的 PCR 产物, 因此将从制备自合适组织的 cDNA 文库中获得每个受体的全长 cDNA 克隆, 插入 pcDNA-1 载体中, 转染进 COS 细胞, 并将测试转染的细胞结合碘化 GDF-8 或 GDF-11 的能力。理想地, 在此筛选中鉴定的每一个受体将检测与这些配体结合的能力。然而, 被鉴定受体的数量可以很大, 分离所有全长 cDNA, 检测它们需要相当大的努力。几乎肯定地说, 一些被鉴定受体会相当于已知的受体, 对于它们, 或者从其它研究者那里获得全长 cDNA 克隆, 或者根据公开发表的序列通过 PCR 扩增编码序列, 应当是简单的。对于新的序列, 将通过 northern 印迹分析确定组织分布, 最优先考虑的事是指向那些表达方式与上面确定的 GDF-8 和 / 或 GDF-11 结合部位的分布最紧密相似的受体。

[0357] 特别地, 已知这些受体分成两类, I 型和 II 型, 它们可以根据序列区分并均需完全的活性。某些配体在没有 II 型受体时不能结合 I 型受体, 而其它则能够与两型受体都结合 (Massague 和 Weis-Garcia, 见前, 1996)。上面概述的交联实验应当给出有关 I 型和 II 型受体是否也都涉及 GDF-8 和 GDF-11 信号发送的一些指示。如果如此, 为全面理解 GDF-8 和 GDF-11 如何传输其信号, 克隆这两个受体亚型将是重要的。由于不能预测在没有 II 型受体时, I 型受体是否能够与 GDF-8 和 GDF-11 相互作用, 所以将首先克隆 II 型受体。只有当这些配体的至少一个 II 型受体已经被鉴定后, 才会进行鉴定 GDF-8 和 GDF-11 I 型受体的尝

试。一般的策略将是, II 型受体与 PCR 筛选中鉴定的每一个 I 型受体共转染, 然后通过交联检测转染的细胞。如果 I 型受体是 GDF-8 或 GDF-11 受体复合体的一部分, 将在转染的细胞中应当探测到两个交联的受体种类, 一个相当于 I 型受体, 另一个相当于 II 型受体。

[0358] 由于至少一个称作 GDNF 的 TGF- β 超家族成员能够通过一个完全不同类型的、涉及 GPI- 联成分(糖基磷脂酰肌醇联成分)(GDNFR- α) 和受体酪氨酸激酶(c-ret; Trupp 等人, *Nature* 381:785-789, 1996; Durbec 等人, *Nature* 381:789-793, 1996; Treanor 等人, *Nature* 382:80-83, 1996; Jing 等人, *细胞* 85:1113-1124, 1996) 的受体复合体发送信号, 对 GDF-8 和 GDF-11 受体的寻找进一步复杂化。尽管 GDNF 是 TGF- β 超家族成员在相关性上最遥远的, 其它 TGF- β 家族成员也能通过类似的受体系统发送信号肯定是可能的。如果 GDF-8 和 GDF-11 确实通过相似的受体复合体发送信号, 表达筛选方法应当至少能够鉴定出此复合体中 GPI- 联成分(实际上, GDNFR- α 采用表达筛选方法来鉴定)。就 GDNF 来说, GDNF- 和 c-ret- 缺陷型小鼠的相似表现型提示 c-ret 是 GDNF 的潜在受体。

[0359] 实施例 14

[0360] GDF-11 敲除小鼠的制备和特性

[0361] GDF-11 敲除小鼠的表现型在许多方面与携带一种对某些 TGF- θ 超家族成员受体的受体的存在缺失的小鼠表现型相似, 包括激活素 IIB 型受体(Act RIIB)。为确定 GDF-11 的生物学功能, GDF-11 基因胚胎干细胞被同源定位破坏。

[0362] 根据 Stratagene(La Jolla, CA) 提供的产品说明书, 在 λ FIXII 中制备小鼠 129SvJ 基因组文库。GDF-11 基因的结构从限制性图谱和从文库中分离的噬菌体克隆的部分测序中推导。制备靶向构建物的载体由 Philip Soriano 和 Kirk Thomsa 惠赠。为保证得到的小鼠不具有 GDF-11 功能, 去除了完整的成熟 C 末端区并替换成一个 neo 基因盒。用靶向构建物转染 R1ES 细胞, 用 9-[1,3-二羟-2-丙氧甲基]鸟嘌呤(2TM) 和 G418(250Tg/ml) 选择, 并通过 Southern 印迹分析法分析。

[0363] GDF-11 基因的同源定位在 8/115 9-[1,3-二羟-2-丙氧甲基]鸟嘌呤/G418 双重抗性 ES 细胞克隆中观察到。在将几种靶向克隆注射进 C57BL/6J 胚胎后, 从一个 ES 克隆中获得了嵌合体, 当与雌性 C57BL/6J 和 129/SvJ 杂交时, 产生杂合仔。C57BL/6J/129/SvJ 杂种 F1 杂合体的杂交产生 49 个野生型(34%)、94 个杂合子(66%), 没有纯合突变体成年子代。相似地, 在 129/SvJ 背景中没有见到成年纯合无效动物(32 个野生型(36%) 和 56 个杂合子突变体(64%) 动物)。

[0364] 为确定纯合突变体死亡的年龄, 在不同妊娠年龄, 从曾经与杂合雄性交配的杂合雌性中分离成窝胚胎, 测定基因型。在检测的所有胚胎阶段, 纯合突变体胚胎以接近预期的频率 25% 出现。在杂种新生小鼠中, 不同基因型也以预期的孟德尔比率出现, 1 : 2 : 1(34+/(28%), 61+/(50%) 和 28-/(23%))。纯合突变体小鼠活着出生并能够呼吸和喂奶。但所有纯合突变体在出生后第一个 24 小时内死亡。准确的死亡原因不清楚, 但死亡率可能与纯合突变体的肾脏严重的发育不良或完全缺失有关。

[0365] 纯合突变体动物可以通过其极短或缺失的尾巴而被容易地识别。为进一步鉴定这些纯合突变体动物的尾巴缺陷, 检测了它们的骨骼以确定尾椎骨破坏的程度。但是, 对胚胎晚期和新生小鼠的野生型和突变体骨骼标本的比较显示, 差异不仅存在于动物的尾部区域, 而且也存在于许多其它区域。几乎在记录了差异的每一例中, 异常看起来表现为脊柱节

段一致的变形,其中特定的节段看起来具有更前面节段的形态特征。这些变形在从颈部区域到尾部区域的整个轴向骨骼中是明显的。除轴向骨骼中见到的缺陷以外,骨骼的其余部分,如头盖骨和四肢骨看起来是正常的。

[0366] 突变体新生动物中的脊柱前变形最容易表现在胸区,那里有显著的胸椎(T)节段数量的增加。所有检测的野生型动物显示典型的13个胸椎,每一个伴有一对肋骨。相反,纯合突变体小鼠显示胸椎数量惊人的增加。所有检测的纯合突变体有另外4至5对肋骨,总共17至18对,尽管在这些动物当中,超过1/3当中,第18肋显示为不发育。因此,正常应当对应于腰椎(L)的L1至L4或L5节段在突变体动物中看起来已经变形为胸椎节段。

[0367] 此外,胸部区域内的变形也是明显的,其中一个胸椎具有另一个胸椎的形态学特征。例如,在野生型小鼠中,第7对肋骨与胸骨附着,剩下的6对不附着或游离。在纯合突变体中,附着和游离的肋骨对数量分别增加至10-11和7-8。因此,在野生型动物中都有游离肋骨的胸椎节段T8、T9、T10,在某些情况下甚至有T11,在突变体动物中都变形为具有更前面的胸椎的典型特征,即出现肋骨附着到胸骨上。与此发现一致,正常在野生型动物的T10上发现的过渡性棘突和过渡性关节突,在纯合突变体中发现在T13上。在某些突变体动物中还记录了胸部区域内另外的变形。例如,在野生型小鼠中,来自T1的肋骨一般地接触到胸骨的顶端。但是,在检测的2/23杂种和2/3129/SvJ纯合突变体小鼠中,T2看起来已经变形为具有与T1相似的形态;即,在这些动物中,来自T2的肋骨延长到与胸骨顶端接触。在这些情况下,来自T1的肋骨看起来与第二对肋骨融合。最后,在82%的纯合突变体中,一般出现在T2上的长棘突移到T3位置上。在某些其它纯合突变体中,见到在其它的胸椎水平有脊柱胸骨肋骨对的不对称融合。

[0368] 前变形不限于胸部区域。我们观察到的前部的最大变形在第6颈椎(C6)水平。在野生型小鼠中,通过腹侧存在的两个前结节可以容易地识别C6。在几个纯合突变体小鼠中,尽管这两个前结节中的一个存在于C6上,但另一个却出现在C7的位置。因此,在这些小鼠中,C7看起来已经部分变形为具有与C6相似的形态。另外一个纯合突变体在C7上有2个前结节,但在C6上保留了一个前结节,成为一个完全的C7到C6变形,但部分的C6到C5变形。

[0369] 轴向骨骼的变形也延伸到腰部区域。野生型动物一般地仅有6个腰椎,而纯合突变体有8至9个。由于上面描述的数据提示,4至5个腰椎节段变形成胸椎节段,所以在突变体中,至少有6个腰椎一定来自通常由骶椎和尾椎产生的节段。因此,纯合突变体小鼠具有总共33-34个骶前椎骨,相比在野生型小鼠中一般存在26个骶前椎骨。最常见的骶前椎骨类型是,突变体小鼠C7/T18/L8和C7/T18/L9,相比野生型小鼠是C7/T13/L6。由于后肢相对于前肢的位置后移了7至8个节段,因此甚至在没有详细的检查骨骼的情况下,就观察到在突变体动物中存在额外的骶前椎骨。

[0370] 在纯合突变体小鼠中,尽管骶椎和尾椎也被影响,但每个变形的确切性质不是容易识别的。在野生型小鼠中,尾部节段S1和S2同S3和S4相比,具有典型的宽横突。在突变体中,没有看起来是可识别的S1或S2椎骨。取而代之,突变体动物有几个看起来具有与S3形态相似的椎骨。此外,所有4个骶椎的横突通常相互融合在一起,尽管在新生时经常仅见到前3个椎骨的融合。然而,在纯合突变体中,骶椎的横突通常是不融合的。在最尾的区域,所有突变体动物还有严重的椎骨畸形及广泛的软骨融合。尽管融合的严重性使计数尾

部区域椎骨的数量很困难,但在几个动物中数出高达 15 个横突。不能确定这些是否代表了突变体的骶或尾椎,因为区分 S4 和尾椎的形态学标准即使在野生型新生动物中也没有建立。不管其身份,此区域椎骨的总数量显著地少于大约 30 的正常数量。因此,尽管突变体较野生型小鼠有显著多的胸椎和腰椎,但由于尾巴的截短,突变体节段的总数量是减少的。

[0371] 杂合小鼠在轴向骨骼中也显示异常,尽管其表现型比纯合小鼠更像野生型。杂合小鼠中最明显的异常是存在一个额外的带有一对肋骨的胸椎节段。此变形存在于受检的每一个杂合动物中,且在每一例中额外的肋骨对都附着在胸骨上。因此, T8, 其伴随的肋骨一般不与胸骨接触,看起来已变形为更前胸椎的形态学特征,且 L1 看起来已变形为后胸椎的形态学特征。在杂合小鼠中也见到各种程度的表现出前变形的其它异常。这些包括, T2 特征性的长棘突移动一个节段到 T3, 关节突和棘突从 T10 移到 T11, C6 上的前结节移到 C7, T2 变形到 T1, T2 伴随的肋骨与胸骨顶端接触。

[0372] 为理解 GDF-11 突变体小鼠中见到的轴向型异常的基础,检测了在发育各阶段分离的突变体胚胎,并与野生型胚胎比较。通过大体形态学检查,直到妊娠期第 9.5 天分离的纯合突变体胚胎不易与相应的野生型胚胎区分。特别地,在任何特定的发育年龄,存在的体节数目在突变体和野生型胚胎间是一样的,提示体节形成的速率在突变体中是没有改变的。在妊娠期后第 10.5-11.5 天,通过后肢后移 7-8 体节,可以容易地区分突变体胚胎和野生型胚胎。在此阶段,尾部发育异常也容易显现。合起来看,这些数据提示,在突变体骨骼中观察到的异常代表了节段本身的真正变形,而不是通过例如增加体节生成速率插入额外的节段。

[0373] 已知含同源异形框的基因表达改变引起果蝇和脊椎动物的变形。为了解 Hox 基因(脊椎动物含同源异形框的基因)的表达模式在 GDF-11 零突变体中是否改变,在妊娠期后第 12.5 天的野生型、杂合子和纯合子突变体胚胎中,通过全样原位杂交,测定了 3 个代表性 Hox 基因, Hoxc-6、Hoxc-8 和 Hoxc-11 的表达模式。Hoxc-6 的表达模式在野生型胚胎中横跨原脊椎 8-15,对应于胸部节段 T1-T8。但在纯合突变体中, Hoxc-6 表达模式向后迁移并延伸至原脊椎 9-18(T2-T11)。用 Hoxc-8 探针见到相似的迁移。在野生型胚胎中, Hoxc-8 在原脊椎 13-18(T6-T11)中表达,但在纯合突变体胚胎中, Hoxc-8 在原脊椎 14-22(T7-T15)中表达。最后, Hoxc-11 表达也向后迁移,因为表达的前界从野生型胚胎中的原脊椎 28 改变到突变体胚胎中的原脊椎 36(注意到,由于后肢的位置在突变体胚胎中也向后迁移, Hoxc-11 的表达模式在野生型和突变体中相对于后肢看起来是相似的)。这些数据提供了进一步的证据,即突变体动物中见到的骨骼异常代表着相同的改变。

[0374] GDF-11 小鼠的表现型提示, GDF-11 在胚胎形成的早期作为轴向结构的整体调节物起作用。为开始调查 GDF-11 发挥其作用的机制,通过全样原位杂交在小鼠早期胚胎中检测了 GDF-11 的表达模式。在这些阶段, GDF-11 表达的原始部位与已知的中胚层细胞产生的位置准确地相关。GDF-11 的表达首先在妊娠期后第 8.25-8.5 天(8-10 体节)的原条区探测到,这是内移细胞形成发育胚胎的中胚层的部位。在 8.75 天时,表达保持在原条中,但在妊娠期后第 9.5 天时,尾芽代替原条成为新的中胚层细胞来源, GDF-11 的表达迁移到尾芽。因此,在这些早期阶段, GDF-11 看起来在发育中的胚胎区域内合成,在此新的中胚层细胞产生并可能获得其位置的认同。

[0375] GDF-11 敲除小鼠的表现型在几个方面与 TGF- θ 超家族某些成员的受体,如激活

素 IIB 型受体 (Act RIIB), 缺失小鼠的表现型相似。与 GDF-11 敲除小鼠的情况一样, Act RIIB 敲除小鼠具有额外的肋骨对, 且肾脏缺陷的范围从肾脏发育不良到完全缺失。这些小鼠表现型的相似性提出了 Act RIIB 可以是 GDF-11 受体的可能性。但是, Act RIIB 可能不是 GDF-11 的唯一受体, 因为 GDF-11 敲除小鼠的表现型比 Act RIIB 小鼠的表现型更严重。例如, 尽管 GDF-11 敲除动物有 4-5 对额外的肋骨并在整个轴向骨骼上显示相同的变形, 但 Act RIIB 敲除动物仅有 3 对额外的肋骨且不在其它轴水平显示变形。此外, 数据提示, GDF-11 敲除小鼠的肾脏缺陷也比 Act RIIB 敲除小鼠更严重。Act RIIB 敲除小鼠在左 / 右轴形成中显示缺陷, 如肺异构和心脏范围的缺陷, 这些我们还没有在 GDF-11 敲除小鼠中观察到。尽管产生的这些配体的敲除小鼠中没有显示左 / 右轴形成的缺陷, 但 Act RIIB 可以结合激活素和某些骨形态发生蛋白 (BMP)。

[0376] 如果 GDF-11 确实直接作用于中胚层细胞以建立位置认同, 那么这里给出的数据将与 GDF-11 作用的短范围或形态发生素模式一致。即, 当中胚层细胞在 GDF-11 表达的位点产生时, GDF-11 可以作用在中胚层前体上以建立 Hox 基因表达的模式, 或可选择地, 在胚胎后端产生的 GDF-11 可以扩散以形成形态发生素梯度。无论 GDF-11 的作用机制是什么, 在 GDF-11 敲除动物中仍然出现大体的前 / 后结构图式发育, 提示 GDF-11 可能不是前 / 后结构特化的唯一调节物。不过, GDF-11 作为轴向结构图式发育的整体调节物而发挥重要作用是明确的, 对此分子的进一步研究将导致关于脊椎动物胚胎中如何沿前 / 后轴建立位置认同的重要的新认识。

[0377] 预期在 GDF-8 敲除动物中有相似的表现型。例如, 当与野生型比较时, 预期 GDF-8 敲除动物有肋骨数目增加、肾脏缺陷和解剖差异。

[0378] 尽管本发明根据上面的实施例中已经得到描述, 应当理解, 修饰和变化都包含在本发明的精神和范围内。因此, 本发明仅受下面的权利要求限制。

[0379] 序列表

[0380]

- <110> 约翰斯·霍普金斯大学医学院
- <120> 原肌生长抑制素肽及其使用方法
- <130> JHU1120-11CN
- <150> PCT/US 01/23510
<151> 2001-07-26
- <150> 09/628, 112
<151> 2000-07-27
- <160> 29
- <170> FastSEQ for Windows Version 4.0
- <210> 1
<211> 2743
<212> DNA
<213> 人类
- <220>
<221> CDS
<222> (59)... (1183)
- <400> 1
- | | |
|--|-----|
| aagaaaagta aaaggaagaa acaagaacaa gaaaaaagat tatattgatt ttaaaatc | 58 |
| atg caa aaa ctg caa ctc tgt gtt tat att tac ctg ttt atg ctg att | 106 |
| Met Gln Lys Leu Gln Leu Cys Val Tyr Ile Tyr Leu Phe Met Leu Ile | |
| 1 5 10 15 | |
| gtt gct ggt cca gtg gat cta aat gag aac agt gag caa aaa gaa aat | 154 |
| Val Ala Gly Pro Val Asp Leu Asn Glu Asn Ser Glu Gln Lys Glu Asn | |
| 20 25 30 | |
| gtg gaa aaa gag ggg ctg tgt aat gca tgt act tgg aga caa aac act | 202 |
| Val Glu Lys Glu Gly Leu Cys Asn Ala Cys Thr Trp Arg Gln Asn Thr | |
| 35 40 45 | |
| aaa tct tca aga ata gaa gcc att aag ata caa atc ctc agt aaa ctt | 250 |
| Lys Ser Ser Arg Ile Glu Ala Ile Lys Ile Gln Ile Leu Ser Lys Leu | |

[0381]

50	55	60	
cgt ctg gaa aca gct cct aac atc agc aaa gat gtt ata aga caa ctt			298
Arg Leu Glu Thr Ala Pro Asn Ile Ser Lys Asp Val Ile Arg Gln Leu			
65	70	75	80
tta ccc aaa gct cct cca ctc cgg gaa ctg att gat cag tat gat gtc			346
Leu Pro Lys Ala Pro Pro Leu Arg Glu Leu Ile Asp Gln Tyr Asp Val			
85	90	95	
cag agg gat gac agc agc gat ggc tct ttg gaa gat gac gat tat cac			394
Gln Arg Asp Asp Ser Ser Asp Gly Ser Leu Glu Asp Asp Asp Tyr His			
100	105	110	
gct aca acg gaa aca atc att acc atg cct aca gag tct gat ttt cta			442
Ala Thr Thr Glu Thr Ile Ile Thr Met Pro Thr Glu Ser Asp Phe Leu			
115	120	125	
atg caa gtg gat gga aaa ccc aaa tgt tgc ttc ttt aaa ttt agc tct			490
Met Gln Val Asp Gly Lys Pro Lys Cys Cys Phe Phe Lys Phe Ser Ser			
130	135	140	
aaa ata caa tac aat aaa gta gta aag gcc caa cta tgg ata tat ttg			538
Lys Ile Gln Tyr Asn Lys Val Val Lys Ala Gln Leu Trp Ile Tyr Leu			
145	150	155	160
aga ccc gtc gag act cct aca aca gtg ttt gtg caa atc ctg aga ctc			586
Arg Pro Val Glu Thr Pro Thr Thr Val Phe Val Gln Ile Leu Arg Leu			
165	170	175	
atc aaa cct atg aaa gac ggt aca agg tat act gga atc cga tct ctg			634
Ile Lys Pro Met Lys Asp Gly Thr Arg Tyr Thr Gly Ile Arg Ser Leu			
180	185	190	
aaa ctt gac atg aac cca ggc act ggt att tgg cag agc att gat gtg			682
Lys Leu Asp Met Asn Pro Gly Thr Gly Ile Trp Gln Ser Ile Asp Val			
195	200	205	
aag aca gtg ttg caa aat tgg ctc aaa caa cct gaa tcc aac tta ggc			730
Lys Thr Val Leu Gln Asn Trp Leu Lys Gln Pro Glu Ser Asn Leu Gly			
210	215	220	
att gaa ata aaa gct tta gat gag aat ggt cat gat ctt gct gta acc			778
Ile Glu Ile Lys Ala Leu Asp Glu Asn Gly His Asp Leu Ala Val Thr			

[0382]

225	230	235	240	
ttc cca gga cca gga gaa gat ggg ctg aat ccg ttt tta gag gtc aag				826
Phe Pro Gly Pro Gly Glu Asp Gly Leu Asn Pro Phe Leu Glu Val Lys				
	245	250	255	
gta aca gac aca cca aaa aga tcc aga agg gat ttt ggt ctt gac tgt				874
Val Thr Asp Thr Pro Lys Arg Ser Arg Arg Asp Phe Gly Leu Asp Cys				
	260	265	270	
gat gag cac tca aca gaa tca cga tgc tgt cgt tac cct cta act gtg				922
Asp Glu His Ser Thr Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr Val				
	275	280	285	
gat ttt gaa gct ttt gga tgg gat tgg att atc gct cct aaa aga tat				970
Asp Phe Glu Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg Tyr				
	290	295	300	
aag gcc aat tac tgc tct gga gag tgt gaa ttt gta ttt tta caa aaa				1018
Lys Ala Asn Tyr Cys Ser Gly Glu Cys Glu Phe Val Phe Leu Gln Lys				
	305	310	315	320
tat cct cat act cat ctg gta cac caa gca aac ccc aga ggt tca gca				1066
Tyr Pro His Thr His Leu Val His Gln Ala Asn Pro Arg Gly Ser Ala				
	325	330	335	
ggc cct tgc tgt act ccc aca aag atg tct cca att aat atg cta tat				1114
Gly Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr				
	340	345	350	
ttt aat ggc aaa gaa caa ata ata tat ggg aaa att cca gcg atg gta				1162
Phe Asn Gly Lys Glu Gln Ile Ile Tyr Gly Lys Ile Pro Ala Met Val				
	355	360	365	
gta gac cgc tgt ggg tgc tca tgagatttat attaagcgtt cataacttcc				1213
Val Asp Arg Cys Gly Cys Ser				
	370	375		
taaaacatgg aaggttttcc cctcaacaat tttgaagctg tgaaattaag taccacaggc				1273
tataggccta gtagtatgcta cagtcactta agcataagct acagtatgta aactaaaagg				1333
gggaatatat gcaatggttg gcatttaacc atccaaacaa atcatacaag aaagttttat				1393
gatttccaga gtttttgagc tagaaggaga tcaaattaca tttatgttcc tatatattac				1453
aacatcggcg aggaaatgaa agcgattctc cttgagttct gatgaattaa aggagtatgc				1513
tttaaagtct atttctttaa agttttgttt aatatttaca gaaaaatcca catacagtat				1573

[0383]

```

tggtaaaatg caggattggt atataccatc attcgaatca tccttaaaca cttgaattta 1633
tattgtatgg tagtatactt ggtaagataa aattccacaa aaatagggat ggtgcagcat 1693
atgcaatttc cattcctatt ataattgaca cagtacatta acaatccatg ccaacgggtgc 1753
taatacgata ggctgaatgt ctgaggctac caggtttatac acataaaaaa cattcagtaa 1813
aatagtaagt ttctcttttc ttcaggtgca ttttctaca cctccaaatg aggaatggat 1873
tttctttaat gtaagaagaa tcatttttct agaggttggc tttcaattct gtagcatact 1933
tggagaaact gcattatctt aaaaggcagt caaatgggtg ttgtttttat caaaatgca 1993
aaataacata cttggagaag tatgtaattt tgtcttttga aaattacaac actgccttgg 2053
caacactgca gtttttatgg taaaataata gaaatgatcg actctatcaa tattgtataa 2113
aaagactgaa acaatgcatt tatataatat gtatacaata ttgttttgta aataagtgc 2173
tcctttttta tttactttgg tatattttta cactaaggac atttcaaatt aagtactaag 2233
gcacaaagac atgtcatgca tcacagaaaa gcaactactt atatttcaga gcaaattagc 2293
agattaaata gtggtcttaa aactccatat gttaatgatt agatggttat attacaatca 2353
ttttatattt ttttcatgga ttaacattca cttatggatt catgatggct gtataaagtg 2413
aatttgaat ttcaatgggt tactgtcatt gtgtttaaat ctcaacgttc cattatttta 2473
atacttgcga aacattact aagtatacca aaataattga ctctattatc tgaatgaag 2533
aataaactga tgctatctca acaataactg ttacttttat tttataattt gataatgaat 2593
atatttctgc atttatttac ttctgttttg taaattggga tttgttaat caaatttatt 2653
gtactatgac taaatgaaat tatttcttac atctaatttg tagaaacagt ataagttata 2713
ttaaagtgtt ttcacatttt tttgaaagac 2743

```

<210> 2

<211> 375

<212> PRT

<213> 人类

<400> 2

```

Met Gln Lys Leu Gln Leu Cys Val Tyr Ile Tyr Leu Phe Met Leu Ile
 1           5           10           15
Val Ala Gly Pro Val Asp Leu Asn Glu Asn Ser Glu Gln Lys Glu Asn
           20           25           30
Val Glu Lys Glu Gly Leu Cys Asn Ala Cys Thr Trp Arg Gln Asn Thr
           35           40           45
Lys Ser Ser Arg Ile Glu Ala Ile Lys Ile Gln Ile Leu Ser Lys Leu
           50           55           60
Arg Leu Glu Thr Ala Pro Asn Ile Ser Lys Asp Val Ile Arg Gln Leu
           65           70           75           80
Leu Pro Lys Ala Pro Pro Leu Arg Glu Leu Ile Asp Gln Tyr Asp Val
           85           90           95
Gln Arg Asp Asp Ser Ser Asp Gly Ser Leu Glu Asp Asp Asp Tyr His
           100          105          110
Ala Thr Thr Glu Thr Ile Ile Thr Met Pro Thr Glu Ser Asp Phe Leu
           115          120          125
Met Gln Val Asp Gly Lys Pro Lys Cys Cys Phe Phe Lys Phe Ser Ser

```

[0384]

130	135	140
Lys Ile Gln Tyr Asn Lys Val Val Lys Ala Gln Leu Trp Ile Tyr Leu		
145	150	155
Arg Pro Val Glu Thr Pro Thr Thr Val Phe Val Gln Ile Leu Arg Leu		
	165	170
Ile Lys Pro Met Lys Asp Gly Thr Arg Tyr Thr Gly Ile Arg Ser Leu		
	180	185
Lys Leu Asp Met Asn Pro Gly Thr Gly Ile Trp Gln Ser Ile Asp Val		
	195	200
Lys Thr Val Leu Gln Asn Trp Leu Lys Gln Pro Glu Ser Asn Leu Gly		
	210	215
Ile Glu Ile Lys Ala Leu Asp Glu Asn Gly His Asp Leu Ala Val Thr		
225	230	235
Phe Pro Gly Pro Gly Glu Asp Gly Leu Asn Pro Phe Leu Glu Val Lys		
	245	250
Val Thr Asp Thr Pro Lys Arg Ser Arg Arg Asp Phe Gly Leu Asp Cys		
	260	265
Asp Glu His Ser Thr Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr Val		
	275	280
Asp Phe Glu Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg Tyr		
	290	295
Lys Ala Asn Tyr Cys Ser Gly Glu Cys Glu Phe Val Phe Leu Gln Lys		
305	310	315
Tyr Pro His Thr His Leu Val His Gln Ala Asn Pro Arg Gly Ser Ala		
	325	330
Gly Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr		
	340	345
Phe Asn Gly Lys Glu Gln Ile Ile Tyr Gly Lys Ile Pro Ala Met Val		
	355	360
Val Asp Arg Cys Gly Cys Ser		
	370	375

<210> 3

<211> 2676

<212> DNA

<213> 日本小鼠

<220>

<221> CDS

<222> (104)... (1231)

<400> 3

```
gtctctcgga cggatcatgc actaatatatt cacttggcat tactcaaaag caaaaagaag 60
aaataagaac aagggaaaaa aaaagattgt gctgattttt aaa atg atg caa aaa 115
```

[0385]

	Met Met Gln Lys	
	1	
ctg caa atg tat gtt tat att tac ctg ttc atg ctg att gct gct ggc		163
Leu Gln Met Tyr Val Tyr Ile Tyr Leu Phe Met Leu Ile Ala Ala Gly		
5	10	15
		20
cca gtg gat cta aat gag ggc agt gag aga gaa gaa aat gtg gaa aaa		211
Pro Val Asp Leu Asn Glu Gly Ser Glu Arg Glu Glu Asn Val Glu Lys		
	25	30
		35
gag ggg ctg tgt aat gca tgt gcg tgg aga caa aac acg agg tac tcc		259
Glu Gly Leu Cys Asn Ala Cys Ala Trp Arg Gln Asn Thr Arg Tyr Ser		
	40	45
		50
aga ata gaa gcc ata aaa att caa atc ctc agt aag ctg cgc ctg gaa		307
Arg Ile Glu Ala Ile Lys Ile Gln Ile Leu Ser Lys Leu Arg Leu Glu		
	55	60
		65
aca gct cct aac atc agc aaa gat gct ata aga caa ctt ctg cca aga		355
Thr Ala Pro Asn Ile Ser Lys Asp Ala Ile Arg Gln Leu Leu Pro Arg		
	70	75
		80
gcg cct cca ctc cgg gaa ctg atc gat cag tac gac gtc cag agg gat		403
Ala Pro Pro Leu Arg Glu Leu Ile Asp Gln Tyr Asp Val Gln Arg Asp		
	85	90
		95
		100
gac agc agt gat ggc tct ttg gaa gat gac gat tat cac gct acc acg		451
Asp Ser Ser Asp Gly Ser Leu Glu Asp Asp Asp Tyr His Ala Thr Thr		
	105	110
		115
gaa aca atc att acc atg cct aca gag tct gac ttt cta atg caa gcg		499
Glu Thr Ile Ile Thr Met Pro Thr Glu Ser Asp Phe Leu Met Gln Ala		
	120	125
		130
gat ggc aag ccc aaa tgt tgc ttt ttt aaa ttt agc tct aaa ata cag		547
Asp Gly Lys Pro Lys Cys Cys Phe Phe Lys Phe Ser Ser Lys Ile Gln		
	135	140
		145
tac aac aaa gta gta aaa gcc caa ctg tgg ata tat ctc aga ccc gtc		595
Tyr Asn Lys Val Val Lys Ala Gln Leu Trp Ile Tyr Leu Arg Pro Val		
	150	155
		160
aag act cct aca aca gtg ttt gtg caa atc ctg aga ctc atc aaa ccc		643

[0386]

Lys Thr Pro Thr Thr Val Phe Val Gln Ile Leu Arg Leu Ile Lys Pro	
165	170 175 180
atg aaa gac ggt aca agg tat act gga atc cga tct ctg aaa ctt gac	691
Met Lys Asp Gly Thr Arg Tyr Thr Gly Ile Arg Ser Leu Lys Leu Asp	
	185 190 195
atg agc cca ggc act ggt att tgg cag agt att gat gtg aag aca gtg	739
Met Ser Pro Gly Thr Gly Ile Trp Gln Ser Ile Asp Val Lys Thr Val	
	200 205 210
ttg caa aat tgg ctc aaa cag cct gaa tcc aac tta ggc att gaa atc	787
Leu Gln Asn Trp Leu Lys Gln Pro Glu Ser Asn Leu Gly Ile Glu Ile	
	215 220 225
aaa gct ttg gat gag aat ggc cat gat ctt gct gta acc ttc cca gga	835
Lys Ala Leu Asp Glu Asn Gly His Asp Leu Ala Val Thr Phe Pro Gly	
	230 235 240
cca gga gaa gat ggg ctg aat ccc ttt tta gaa gtc aag gtg aca gac	883
Pro Gly Glu Asp Gly Leu Asn Pro Phe Leu Glu Val Lys Val Thr Asp	
	245 250 255 260
aca ccc aag agg tcc cgg aga gac ttt ggg ctt gac tgc gat gag cac	931
Thr Pro Lys Arg Ser Arg Arg Asp Phe Gly Leu Asp Cys Asp Glu His	
	265 270 275
tcc acg gaa tcc cgg tgc tgc cgc tac ccc ctc acg gtc gat ttt gaa	979
Ser Thr Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr Val Asp Phe Glu	
	280 285 290
gcc ttt gga tgg gac tgg att atc gca ccc aaa aga tat aag gcc aat	1027
Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg Tyr Lys Ala Asn	
	295 300 305
tac tgc tca gga gag tgt gaa ttt gtg ttt tta caa aaa tat ccg cat	1075
Tyr Cys Ser Gly Glu Cys Glu Phe Val Phe Leu Gln Lys Tyr Pro His	
	310 315 320
act cat ctt gtg cac caa gca aac ccc aga ggc tca gca ggc cct tgc	1123
Thr His Leu Val His Gln Ala Asn Pro Arg Gly Ser Ala Gly Pro Cys	
	325 330 335 340
tgc act ccg aca aaa atg tct ccc att aat atg cta tat ttt aat ggc	1171

[0387]

Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr Phe Asn Gly
 345 350 355

aaa gaa caa ata ata tat ggg aaa att cca gcc atg gta gta gac cgc 1219
 Lys Glu Gln Ile Ile Tyr Gly Lys Ile Pro Ala Met Val Val Asp Arg
 360 365 370

tgt ggg tgc tca tgagctttgc attaggttag aaacttccca agtcatggaa 1271
 Cys Gly Cys Ser
 375

ggctttcccc tcaatttcga aactgtgaat tcaagcacca caggctgtag gccttgagta 1331
 tgctctagta acgtaagcac aagctacagt gtatgaaacta aaagagagaa tagatgcaat 1391
 ggttggcatt caaccaccaa aataaacat actataggat gttgtatgat ttccagagtt 1451
 tttgaaatag atggagatca aattacatt atgtccatat atgtatatta caactacaat 1511
 ctaggcaagg aagtgagagc acatcttggtg gtctgctgag ttaggagggt atgattaaaa 1571
 ggtaaagtct tatttcctaa cagtttact taatatttac agaagaatct atatgtagcc 1631
 tttgtaaagt gtaggattgt tatcatttaa aaacatcatg tacacttata tttgtattgt 1691
 atacttggta agataaaatt ccacaaagta ggaatggggc ctcacataca cattgccatt 1751
 cctattataa ttggacaatc caccacgggt ctaatgcagt gctgaatggc tcctactgga 1811
 cctctcgata gaacactcta caaagtacga gtctctctct cccttcagg tgcacttcca 1871
 cacacacagc actaagtgtt caatgcattt tctttaagga aagaagaatc ttttttcta 1931
 gaggtcaact ttcagtcaac tctagcacag cgggagtgac tgctgcatct taaaaggcag 1991
 ccaaacagta ttcatttttt aatctaaatt tcaaaatcac tgtctgcctt tatcacatgg 2051
 caattttgtg gtaaaataat ggaaatgact ggttctatca atattgtata aaagactctg 2111
 aaacaattac atttatataa tatgtataca atattgtttt gtaaataagt gtctcctttt 2171
 atatttactt tggatattt ttacactaat gaaattcaa atcattaag tacaagaca 2231
 tgtcatgtat cacaaaaaag gtgactgctt ctatttcaga gtgaattagc agattcaata 2291
 gtggtcttaa aactctgtat gtttaagatta gaaggttata ttacaatcaa tttatgtatt 2351
 ttttacatta tcaacttatg gtttcatggt ggctgtatct atgaatgtgg ctcccagtca 2411
 aatttcaatg cccaccatt ttaaaaatta caagcattac taaacatacc aacatgtatc 2471
 taaagaata caaatatggt atctcaataa cagctacttt tttattttat aatttgacaa 2531
 tgaatacatt tcttttattt acttcagttt tataaattgg aactttgttt atcaaatgta 2591
 ttgtactcat agctaaatga aattatttct tacataaaaa tgtgtagaaa ctataaatta 2651
 aagtgttttc acatttttga aaggc 2676

- <210> 4
- <211> 376
- <212> PRT
- <213> 日本小鼠

<400> 4
 Met Met Gln Lys Leu Gln Met Tyr Val Tyr Ile Tyr Leu Phe Met Leu
 1 5 10 15

[0388]

Ile Ala Ala Gly Pro Val Asp Leu Asn Glu Gly Ser Glu Arg Glu Glu
 20 25 30
 Asn Val Glu Lys Glu Gly Leu Cys Asn Ala Cys Ala Trp Arg Gln Asn
 35 40 45
 Thr Arg Tyr Ser Arg Ile Glu Ala Ile Lys Ile Gln Ile Leu Ser Lys
 50 55 60
 Leu Arg Leu Glu Thr Ala Pro Asn Ile Ser Lys Asp Ala Ile Arg Gln
 65 70 75 80
 Leu Leu Pro Arg Ala Pro Pro Leu Arg Glu Leu Ile Asp Gln Tyr Asp
 85 90 95
 Val Gln Arg Asp Asp Ser Ser Asp Gly Ser Leu Glu Asp Asp Asp Tyr
 100 105 110
 His Ala Thr Thr Glu Thr Ile Ile Thr Met Pro Thr Glu Ser Asp Phe
 115 120 125
 Leu Met Gln Ala Asp Gly Lys Pro Lys Cys Cys Phe Phe Lys Phe Ser
 130 135 140
 Ser Lys Ile Gln Tyr Asn Lys Val Val Lys Ala Gln Leu Trp Ile Tyr
 145 150 155 160
 Leu Arg Pro Val Lys Thr Pro Thr Thr Val Phe Val Gln Ile Leu Arg
 165 170 175
 Leu Ile Lys Pro Met Lys Asp Gly Thr Arg Tyr Thr Gly Ile Arg Ser
 180 185 190
 Leu Lys Leu Asp Met Ser Pro Gly Thr Gly Ile Trp Gln Ser Ile Asp
 195 200 205
 Val Lys Thr Val Leu Gln Asn Trp Leu Lys Gln Pro Glu Ser Asn Leu
 210 215 220
 Gly Ile Glu Ile Lys Ala Leu Asp Glu Asn Gly His Asp Leu Ala Val
 225 230 235 240
 Thr Phe Pro Gly Pro Gly Glu Asp Gly Leu Asn Pro Phe Leu Glu Val
 245 250 255
 Lys Val Thr Asp Thr Pro Lys Arg Ser Arg Arg Asp Phe Gly Leu Asp
 260 265 270
 Cys Asp Glu His Ser Thr Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr
 275 280 285
 Val Asp Phe Glu Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg
 290 295 300
 Tyr Lys Ala Asn Tyr Cys Ser Gly Glu Cys Glu Phe Val Phe Leu Gln
 305 310 315 320
 Lys Tyr Pro His Thr His Leu Val His Gln Ala Asn Pro Arg Gly Ser
 325 330 335
 Ala Gly Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu
 340 345 350
 Tyr Phe Asn Gly Lys Glu Gln Ile Ile Tyr Gly Lys Ile Pro Ala Met
 355 360 365

[0389]

Val Val Asp Arg Cys Gly Cys Ser
370 375

<210> 5
<211> 1131
<212> DNA
<213> 褐家鼠

<220>
<221> CDS
<222> (1)... (1128)

<400> 5
atg att caa aaa ccg caa atg tat gtt tat att tac ctg ttt gtg ctg 48
Met Ile Gln Lys Pro Gln Met Tyr Val Tyr Ile Tyr Leu Phe Val Leu
1 5 10 15
att gct gct ggc cca gtg gat cta aat gag gac agt gag aga gag gcg 96
Ile Ala Ala Gly Pro Val Asp Leu Asn Glu Asp Ser Glu Arg Glu Ala
20 25 30
aat gtg gaa aaa gag ggg ctg tgt aat gcg tgt gcg tgg aga caa aac 144
Asn Val Glu Lys Glu Gly Leu Cys Asn Ala Cys Ala Trp Arg Gln Asn
35 40 45
aca agg tac tcc aga ata gaa gcc ata aaa att caa atc ctc agt aaa 192
Thr Arg Tyr Ser Arg Ile Glu Ala Ile Lys Ile Gln Ile Leu Ser Lys
50 55 60
ctc cgc ctg gaa aca gcg cct aac atc agc aaa gat gct ata aga caa 240
Leu Arg Leu Glu Thr Ala Pro Asn Ile Ser Lys Asp Ala Ile Arg Gln
65 70 75 80
ctt ctg ccc aga gcg cct cca ctc cgg gaa ctg atc gat cag tac gac 288
Leu Leu Pro Arg Ala Pro Pro Leu Arg Glu Leu Ile Asp Gln Tyr Asp
85 90 95
gtc cag agg gat gac agc agt gac ggc tct ttg gaa gat gac gat tat 336
Val Gln Arg Asp Asp Ser Ser Asp Gly Ser Leu Glu Asp Asp Asp Tyr
100 105 110
cac gct acc acg gaa aca atc att acc atg cct acc gag tct gac ttt 384
His Ala Thr Thr Glu Thr Ile Ile Thr Met Pro Thr Glu Ser Asp Phe
115 120 125

[0390]

cta atg caa gcg gat gga aag ccc aaa tgt tgc ttt ttt aaa ttt agc	432
Leu Met Gln Ala Asp Gly Lys Pro Lys Cys Cys Phe Phe Lys Phe Ser	
130 135 140	
tct aaa ata cag tac aac aaa gtg gta aag gcc cag ctg tgg ata tat	480
Ser Lys Ile Gln Tyr Asn Lys Val Val Lys Ala Gln Leu Trp Ile Tyr	
145 150 155 160	
ctg aga gcc gtc aag act cct aca aca gtg ttt gtg caa atc ctg aga	528
Leu Arg Ala Val Lys Thr Pro Thr Thr Val Phe Val Gln Ile Leu Arg	
165 170 175	
ctc atc aaa ccc atg aaa gac ggt aca agg tat acc gga atc cga tct	576
Leu Ile Lys Pro Met Lys Asp Gly Thr Arg Tyr Thr Gly Ile Arg Ser	
180 185 190	
ctg aaa ctt gac atg agc cca ggc act ggt att tgg cag agt att gat	624
Leu Lys Leu Asp Met Ser Pro Gly Thr Gly Ile Trp Gln Ser Ile Asp	
195 200 205	
gtg aag aca gtg ttg caa aat tgg ctc aaa cag cct gaa tcc aac tta	672
Val Lys Thr Val Leu Gln Asn Trp Leu Lys Gln Pro Glu Ser Asn Leu	
210 215 220	
ggc att gaa atc aaa gct ttg gat gag aat ggg cat gat ctt gct gta	720
Gly Ile Glu Ile Lys Ala Leu Asp Glu Asn Gly His Asp Leu Ala Val	
225 230 235 240	
acc ttc cca gga cca gga gaa gat ggg ctg aat ccc ttt tta gaa gtc	768
Thr Phe Pro Gly Pro Gly Glu Asp Gly Leu Asn Pro Phe Leu Glu Val	
245 250 255	
aaa gta aca gac aca ccc aag agg tcc cgg aga gac ttt ggg ctt gac	816
Lys Val Thr Asp Thr Pro Lys Arg Ser Arg Arg Asp Phe Gly Leu Asp	
260 265 270	
tgc gat gaa cac tcc acg gaa tcg cgg tgc tgt cgc tac ccc ctc acg	864
Cys Asp Glu His Ser Thr Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr	
275 280 285	
gtc gat ttc gaa gcc ttt gga tgg gac tgg att att gca ccc aaa aga	912
Val Asp Phe Glu Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg	
290 295 300	

[0391]

Leu Met Gln Ala Asp Gly Lys Pro Lys Cys Cys Phe Phe Lys Phe Ser
 130 135 140
 Ser Lys Ile Gln Tyr Asn Lys Val Val Lys Ala Gln Leu Trp Ile Tyr
 145 150 155 160
 Leu Arg Ala Val Lys Thr Pro Thr Thr Val Phe Val Gln Ile Leu Arg
 165 170 175
 Leu Ile Lys Pro Met Lys Asp Gly Thr Arg Tyr Thr Gly Ile Arg Ser
 180 185 190
 Leu Lys Leu Asp Met Ser Pro Gly Thr Gly Ile Trp Gln Ser Ile Asp
 195 200 205
 Val Lys Thr Val Leu Gln Asn Trp Leu Lys Gln Pro Glu Ser Asn Leu
 210 215 220
 Gly Ile Glu Ile Lys Ala Leu Asp Glu Asn Gly His Asp Leu Ala Val
 225 230 235 240
 Thr Phe Pro Gly Pro Gly Glu Asp Gly Leu Asn Pro Phe Leu Glu Val
 245 250 255
 Lys Val Thr Asp Thr Pro Lys Arg Ser Arg Arg Asp Phe Gly Leu Asp
 260 265 270
 Cys Asp Glu His Ser Thr Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr
 275 280 285
 Val Asp Phe Glu Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg
 290 295 300
 Tyr Lys Ala Asn Tyr Cys Ser Gly Glu Cys Glu Phe Val Phe Leu Gln
 305 310 315 320
 Lys Tyr Pro His Thr His Leu Val His Gln Ala Asn Pro Arg Gly Ser
 325 330 335
 Ala Gly Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu
 340 345 350
 Tyr Phe Asn Gly Lys Glu Gln Ile Ile Tyr Gly Lys Ile Pro Ala Met
 355 360 365
 Val Val Asp Arg Cys Gly Cys Ser
 370 375

<210> 7

<211> 1128

<212> DNA

<213> 红原鸡

<220>

<221> CDS

<222> (1)... (1125)

<400> 7

atg caa aag ctg gca gtc tat gtt tat att tac ctg ttc atg cag atc

48

[0393]

Ile Lys Pro Met Lys Asp Gly Thr Arg Tyr Thr Gly Ile Arg Ser Leu	
180	185
190	
aaa ctt gac atg aac cca ggc act ggt atc tgg cag agt att gat gtg	624
Lys Leu Asp Met Asn Pro Gly Thr Gly Ile Trp Gln Ser Ile Asp Val	
195	200
205	
aag aca gtg ctg caa aat tgg ctc aaa cag cct gaa tcc aat tta ggc	672
Lys Thr Val Leu Gln Asn Trp Leu Lys Gln Pro Glu Ser Asn Leu Gly	
210	215
220	
atc gaa ata aaa gct ttt gat gag act gga cga gat ctt gct gtc aca	720
Ile Glu Ile Lys Ala Phe Asp Glu Thr Gly Arg Asp Leu Ala Val Thr	
225	230
235	240
ttc cca gga cca gga gaa gat gga ttg aac cca ttt tta gag gtc aga	768
Phe Pro Gly Pro Gly Glu Asp Gly Leu Asn Pro Phe Leu Glu Val Arg	
245	250
255	
gtt aca gac aca ccg aaa cgg tcc cgc aga gat ttt ggc ctt gac tgt	816
Val Thr Asp Thr Pro Lys Arg Ser Arg Arg Asp Phe Gly Leu Asp Cys	
260	265
270	
gat gag cac tca acg gaa tcc cga tgt tgt cgc tac ccg ctg aca gtg	864
Asp Glu His Ser Thr Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr Val	
275	280
285	
gat ttc gaa gct ttt gga tgg gac tgg att ata gca cct aaa aga tac	912
Asp Phe Glu Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg Tyr	
290	295
300	
aaa gcc aat tac tgc tcc gga gaa tgc gaa ttt gtg ttt cta cag aaa	960
Lys Ala Asn Tyr Cys Ser Gly Glu Cys Glu Phe Val Phe Leu Gln Lys	
305	310
315	320
tac ccg cac act cac ctg gta cac caa gca aat ccc aga ggc tca gca	1008
Tyr Pro His Thr His Leu Val His Gln Ala Asn Pro Arg Gly Ser Ala	
325	330
335	
ggc cct tgc tgc aca ccc acc aag atg tcc cct ata aac atg ctg tat	1056
Gly Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr	
340	345
350	
ttc aat gga aaa gaa caa ata ata tat gga aag ata cca gcc atg gtt	1104

[0395]

Phe Asn Gly Lys Glu Gln Ile Ile Tyr Gly Lys Ile Pro Ala Met Val
 355 360 365

gta gat cgt tgc ggg tgc tca tga 1128
 Val Asp Arg Cys Gly Cys Ser
 370 375

<210> 8

<211> 374

<212> PRT

<213> 红原鸡

<400> 8

Gln Lys Leu Ala Val Tyr Val Tyr Ile Tyr Leu Phe Met Gln Ile Ala
 1 5 10 15
 Val Asp Pro Val Ala Leu Asp Gly Ser Ser Gln Pro Thr Glu Asn Ala
 20 25 30
 Glu Lys Asp Gly Leu Cys Asn Ala Cys Thr Trp Arg Gln Asn Thr Lys
 35 40 45
 Ser Ser Arg Ile Glu Ala Ile Lys Ile Gln Ile Leu Ser Lys Leu Arg
 50 55 60
 Leu Glu Gln Ala Pro Asn Ile Ser Arg Asp Val Ile Lys Gln Leu Leu
 65 70 75 80
 Pro Lys Ala Pro Pro Leu Gln Glu Leu Ile Asp Gln Tyr Asp Val Gln
 85 90 95
 Arg Asp Asp Ser Ser Asp Gly Ser Leu Glu Asp Asp Asp Tyr His Ala
 100 105 110
 Thr Thr Glu Thr Ile Ile Thr Met Pro Thr Glu Ser Asp Phe Leu Val
 115 120 125
 Gln Met Glu Gly Lys Pro Lys Cys Cys Phe Phe Lys Phe Ser Ser Lys
 130 135 140
 Ile Gln Tyr Asn Lys Val Val Lys Ala Gln Leu Trp Ile Tyr Leu Arg
 145 150 155 160
 Gln Val Gln Lys Pro Thr Thr Val Phe Val Gln Ile Leu Arg Leu Ile
 165 170 175
 Lys Pro Met Lys Asp Gly Thr Arg Tyr Thr Gly Ile Arg Ser Leu Lys
 180 185 190
 Leu Asp Met Asn Pro Gly Thr Gly Ile Trp Gln Ser Ile Asp Val Lys
 195 200 205
 Thr Val Leu Gln Asn Trp Leu Lys Gln Pro Glu Ser Asn Leu Gly Ile
 210 215 220
 Glu Ile Lys Ala Phe Asp Glu Thr Gly Arg Asp Leu Ala Val Thr Phe
 225 230 235 240

[0396]

Pro Gly Pro Gly Glu Asp Gly Leu Asn Pro Phe Leu Glu Val Arg Val
 245 250 255
 Thr Asp Thr Pro Lys Arg Ser Arg Arg Asp Phe Gly Leu Asp Cys Asp
 260 265 270
 Glu His Ser Thr Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr Val Asp
 275 280 285
 Phe Glu Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg Tyr Lys
 290 295 300
 Ala Asn Tyr Cys Ser Gly Glu Cys Glu Phe Val Phe Leu Gln Lys Tyr
 305 310 315 320
 Pro His Thr His Leu Val His Gln Ala Asn Pro Arg Gly Ser Ala Gly
 325 330 335
 Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr Phe
 340 345 350
 Asn Gly Lys Glu Gln Ile Ile Tyr Gly Lys Ile Pro Ala Met Val Val
 355 360 365
 Asp Arg Cys Gly Cys Ser
 370

<210> 9

<211> 1128

<212> DNA

<213> 狒狒

<220>

<221> CDS

<222> (1)... (1125)

<400> 9

atg caa aaa ctg caa ctc tgt gtt tat att tac ctg ttt atg ctg att 48
 Met Gln Lys Leu Gln Leu Cys Val Tyr Ile Tyr Leu Phe Met Leu Ile
 1 5 10 15

gtt gct ggt cca gtg gat cta aat gag aac agt gag caa aaa gaa aat 96
 Val Ala Gly Pro Val Asp Leu Asn Glu Asn Ser Glu Gln Lys Glu Asn
 20 25 30

gtg gaa aaa gag ggg ctg tgt aat gca tgt act tgg aga caa aac act 144
 Val Glu Lys Glu Gly Leu Cys Asn Ala Cys Thr Trp Arg Gln Asn Thr
 35 40 45

aaa tct tca aga ata gaa gcc att aaa ata caa atc ctc agt aaa ctt 192
 Lys Ser Ser Arg Ile Glu Ala Ile Lys Ile Gln Ile Leu Ser Lys Leu
 50 55 60

[0397]

ttc cca gga cca gga gaa gat ggg ctg aat ccc ttt tta gag gtc aag	768
Phe Pro Gly Pro Gly Glu Asp Gly Leu Asn Pro Phe Leu Glu Val Lys	
245 250 255	
gta aca gac aca ccc aaa aga tcc aga agg gat ttt ggt ctt gac tgt	816
Val Thr Asp Thr Pro Lys Arg Ser Arg Arg Asp Phe Gly Leu Asp Cys	
260 265 270	
gat gag cac tca aca gaa tcg cga tgc tgt cgt tac cct cta act gtg	864
Asp Glu His Ser Thr Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr Val	
275 280 285	
gat ttt gaa gct ctt gga tgg gat tgg att atc gct cct aaa aga tat	912
Asp Phe Glu Ala Leu Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg Tyr	
290 295 300	
aag gcc aat tac tgc tct gga gag tgt gaa ttt gta ttt tta caa aaa	960
Lys Ala Asn Tyr Cys Ser Gly Glu Cys Glu Phe Val Phe Leu Gln Lys	
305 310 315 320	
tat cct cat act cat ctg gta cac caa gca aac ccc aga ggt tca gca	1008
Tyr Pro His Thr His Leu Val His Gln Ala Asn Pro Arg Gly Ser Ala	
325 330 335	
ggc cct tgc tgt act ccc aca aag atg tct cca att aat atg cta tat	1056
Gly Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr	
340 345 350	
ttt aat ggc aaa gaa caa ata ata tat ggg aaa att cca gcc atg gta	1104
Phe Asn Gly Lys Glu Gln Ile Ile Tyr Gly Lys Ile Pro Ala Met Val	
355 360 365	
gta gac cgc tgc ggg tgc tca tga	1128
Val Asp Arg Cys Gly Cys Ser	
370 375	

<210> 10

<211> 375

<212> PRT

<213> 狒狒

<400> 10

[0399]

```

Met Gln Lys Leu Gln Leu Cys Val Tyr Ile Tyr Leu Phe Met Leu Ile
 1           5           10           15
Val Ala Gly Pro Val Asp Leu Asn Glu Asn Ser Glu Gln Lys Glu Asn
          20           25           30
Val Glu Lys Glu Gly Leu Cys Asn Ala Cys Thr Trp Arg Gln Asn Thr
          35           40           45
Lys Ser Ser Arg Ile Glu Ala Ile Lys Ile Gln Ile Leu Ser Lys Leu
 50           55           60
Arg Leu Glu Thr Ala Pro Asn Ile Ser Lys Asp Ala Ile Arg Gln Leu
 65           70           75           80
Leu Pro Lys Ala Pro Pro Leu Arg Glu Leu Ile Asp Gln Tyr Asp Val
          85           90           95
Gln Arg Asp Asp Ser Ser Asp Gly Ser Leu Glu Asp Asp Asp Tyr His
          100          105          110
Ala Thr Thr Glu Thr Ile Ile Thr Met Pro Thr Glu Ser Asp Phe Leu
          115          120          125
Met Gln Val Asp Gly Lys Pro Lys Cys Cys Phe Phe Lys Phe Ser Ser
          130          135          140
Lys Ile Gln Tyr Asn Lys Val Val Lys Ala Gln Leu Trp Ile Tyr Leu
 145          150          155          160
Arg Pro Val Glu Thr Pro Thr Thr Val Phe Val Gln Ile Leu Arg Leu
          165          170          175
Ile Lys Pro Met Lys Asp Gly Thr Arg Tyr Thr Gly Ile Arg Ser Leu
          180          185          190
Lys Leu Asp Met Asn Pro Gly Thr Gly Ile Trp Gln Ser Ile Asp Val
          195          200          205
Lys Thr Val Leu Gln Asn Trp Leu Lys Gln Pro Glu Ser Asn Leu Gly
          210          215          220
Ile Glu Ile Lys Ala Leu Asp Glu Asn Gly His Asp Leu Ala Val Thr
 225          230          235          240
Phe Pro Gly Pro Gly Glu Asp Gly Leu Asn Pro Phe Leu Glu Val Lys
          245          250          255
Val Thr Asp Thr Pro Lys Arg Ser Arg Arg Asp Phe Gly Leu Asp Cys
          260          265          270
Asp Glu His Ser Thr Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr Val
          275          280          285
Asp Phe Glu Ala Leu Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg Tyr
          290          295          300
Lys Ala Asn Tyr Cys Ser Gly Glu Cys Glu Phe Val Phe Leu Gln Lys
 305          310          315          320
Tyr Pro His Thr His Leu Val His Gln Ala Asn Pro Arg Gly Ser Ala
          325          330          335
Gly Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr
          340          345          350

```

[0400]

Phe Asn Gly Lys Glu Gln Ile Ile Tyr Gly Lys Ile Pro Ala Met Val	
355	360
Val Asp Arg Cys Gly Cys Ser	
370	375
<210> 11	
<211> 1128	
<212> DNA	
<213> 牛	
<220>	
<221> CDS	
<222> (1)... (1125)	
<400> 11	
atg caa aaa ctg caa atc tct gtt tat att tac cta ttt atg ctg att	48
Met Gln Lys Leu Gln Ile Ser Val Tyr Ile Tyr Leu Phe Met Leu Ile	
1 5 10 15	
gtt gct ggc cca gtg gat ctg aat gag aac agc gag cag aag gaa aat	96
Val Ala Gly Pro Val Asp Leu Asn Glu Asn Ser Glu Gln Lys Glu Asn	
20 25 30	
gtg gaa aaa gag ggg ctg tgt aat gca tgt ttg tgg agg gaa aac act	144
Val Glu Lys Glu Gly Leu Cys Asn Ala Cys Leu Trp Arg Glu Asn Thr	
35 40 45	
aca tcg tca aga cta gaa gcc ata aaa atc caa atc ctc agt aaa ctt	192
Thr Ser Ser Arg Leu Glu Ala Ile Lys Ile Gln Ile Leu Ser Lys Leu	
50 55 60	
cgc ctg gaa aca gct cct aac atc agc aaa gat gct atc aga caa ctt	240
Arg Leu Glu Thr Ala Pro Asn Ile Ser Lys Asp Ala Ile Arg Gln Leu	
65 70 75 80	
ttg ccc aag gct cct cca ctc ctg gaa ctg att gat cag ttc gat gtc	288
Leu Pro Lys Ala Pro Pro Leu Leu Glu Leu Ile Asp Gln Phe Asp Val	
85 90 95	
cag aga gat gcc agc agt gac ggc tcc ttg gaa gac gat gac tac cac	336
Gln Arg Asp Ala Ser Ser Asp Gly Ser Leu Glu Asp Asp Asp Tyr His	
100 105 110	
gcc agg acg gaa acg gtc att acc atg ccc acg gag tct gat ctt cta	384

[0401]

Ala Arg Thr Glu Thr Val Ile Thr Met Pro Thr Glu Ser Asp Leu Leu 115	120	125	
acg caa gtg gaa gga aaa ccc aaa tgt tgc ttc ttt aaa ttt agc tct Thr Gln Val Glu Gly Lys Pro Lys Cys Cys Phe Phe Lys Phe Ser Ser 130	135	140	432
aag ata caa tac aat aaa cta gta aag gcc caa ctg tgg ata tat ctg Lys Ile Gln Tyr Asn Lys Leu Val Lys Ala Gln Leu Trp Ile Tyr Leu 145	150	155	160
agg cct gtc aag act cct gcg aca gtg ttt gtg caa atc ctg aga ctc Arg Pro Val Lys Thr Pro Ala Thr Val Phe Val Gln Ile Leu Arg Leu 165	170	175	528
atc aaa ccc atg aaa gac ggt aca agg tat act gga atc cga tct ctg Ile Lys Pro Met Lys Asp Gly Thr Arg Tyr Thr Gly Ile Arg Ser Leu 180	185	190	576
aaa ctt gac atg aac cca ggc act ggt att tgg cag agc att gat gtg Lys Leu Asp Met Asn Pro Gly Thr Gly Ile Trp Gln Ser Ile Asp Val 195	200	205	624
aag aca gtg ttg cag aac tgg ctc aaa caa cct gaa tcc aac tta ggc Lys Thr Val Leu Gln Asn Trp Leu Lys Gln Pro Glu Ser Asn Leu Gly 210	215	220	672
att gaa atc aaa gct tta gat gag aat ggc cat gat ctt gct gta acc Ile Glu Ile Lys Ala Leu Asp Glu Asn Gly His Asp Leu Ala Val Thr 225	230	235	240
ttc cca gaa cca gga gaa gat gga ctg act ccc ttt tta gaa gtc aag Phe Pro Glu Pro Gly Glu Asp Gly Leu Thr Pro Phe Leu Glu Val Lys 245	250	255	768
gta aca gac aca cca aaa aga tct agg aga gat ttt ggg ctt gat tgt Val Thr Asp Thr Pro Lys Arg Ser Arg Arg Asp Phe Gly Leu Asp Cys 260	265	270	816
gat gaa cac tcc aca gaa tct cga tgc tgt cgt tac cct cta act gtg Asp Glu His Ser Thr Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr Val 275	280	285	864
gat ttt gaa gct ttt gga tgg gat tgg att att gca cct aaa aga tat			912

[0402]

Asp Phe Glu Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg Tyr
 290 295 300

 aag gcc aat tac tgc tct gga gaa tgt gaa ttt gta ttt ttg caa aag 960
 Lys Ala Asn Tyr Cys Ser Gly Glu Cys Glu Phe Val Phe Leu Gln Lys
 305 310 315 320

 tat cct cat acc cat ctt gtg cac caa gca aac ccc aga ggt tca gcc 1008
 Tyr Pro His Thr His Leu Val His Gln Ala Asn Pro Arg Gly Ser Ala
 325 330 335

 ggc ccc tgc tgt act cct aca aag atg tct cca att aat atg cta tat 1056
 Gly Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr
 340 345 350

 ttt aat ggc gaa gga caa ata ata tac ggg aag att cca gcc atg gta 1104
 Phe Asn Gly Glu Gly Gln Ile Ile Tyr Gly Lys Ile Pro Ala Met Val
 355 360 365

 gta gat cgc tgt ggg tgt tca tga 1128
 Val Asp Arg Cys Gly Cys Ser
 370 375

<210> 12

<211> 375

<212> PRT

<213> 牛

<400> 12

Met Gln Lys Leu Gln Ile Ser Val Tyr Ile Tyr Leu Phe Met Leu Ile
 1 5 10 15
 Val Ala Gly Pro Val Asp Leu Asn Glu Asn Ser Glu Gln Lys Glu Asn
 20 25 30
 Val Glu Lys Glu Gly Leu Cys Asn Ala Cys Leu Trp Arg Glu Asn Thr
 35 40 45
 Thr Ser Ser Arg Leu Glu Ala Ile Lys Ile Gln Ile Leu Ser Lys Leu
 50 55 60
 Arg Leu Glu Thr Ala Pro Asn Ile Ser Lys Asp Ala Ile Arg Gln Leu
 65 70 75 80
 Leu Pro Lys Ala Pro Pro Leu Leu Glu Leu Ile Asp Gln Phe Asp Val
 85 90 95
 Gln Arg Asp Ala Ser Ser Asp Gly Ser Leu Glu Asp Asp Asp Tyr His
 100 105 110

[0403]

Ala Arg Thr Glu Thr Val Ile Thr Met Pro Thr Glu Ser Asp Leu Leu
 115 120 125
 Thr Gln Val Glu Gly Lys Pro Lys Cys Cys Phe Phe Lys Phe Ser Ser
 130 135 140
 Lys Ile Gln Tyr Asn Lys Leu Val Lys Ala Gln Leu Trp Ile Tyr Leu
 145 150 155 160
 Arg Pro Val Lys Thr Pro Ala Thr Val Phe Val Gln Ile Leu Arg Leu
 165 170 175
 Ile Lys Pro Met Lys Asp Gly Thr Arg Tyr Thr Gly Ile Arg Ser Leu
 180 185 190
 Lys Leu Asp Met Asn Pro Gly Thr Gly Ile Trp Gln Ser Ile Asp Val
 195 200 205
 Lys Thr Val Leu Gln Asn Trp Leu Lys Gln Pro Glu Ser Asn Leu Gly
 210 215 220
 Ile Glu Ile Lys Ala Leu Asp Glu Asn Gly His Asp Leu Ala Val Thr
 225 230 235 240
 Phe Pro Glu Pro Gly Glu Asp Gly Leu Thr Pro Phe Leu Glu Val Lys
 245 250 255
 Val Thr Asp Thr Pro Lys Arg Ser Arg Arg Asp Phe Gly Leu Asp Cys
 260 265 270
 Asp Glu His Ser Thr Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr Val
 275 280 285
 Asp Phe Glu Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg Tyr
 290 295 300
 Lys Ala Asn Tyr Cys Ser Gly Glu Cys Glu Phe Val Phe Leu Gln Lys
 305 310 315 320
 Tyr Pro His Thr His Leu Val His Gln Ala Asn Pro Arg Gly Ser Ala
 325 330 335
 Gly Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr
 340 345 350
 Phe Asn Gly Glu Gly Gln Ile Ile Tyr Gly Lys Ile Pro Ala Met Val
 355 360 365
 Val Asp Arg Cys Gly Cys Ser
 370 375

<210> 13

<211> 1128

<212> DNA

<213> 猪

<220>

<221> CDS

<222> (1)... (1125)

[0404]

<400> 13			
atg caa aaa ctg caa atc tat gtt tat att tac ctg ttt atg ctg att			48
Met Gln Lys Leu Gln Ile Tyr Val Tyr Ile Tyr Leu Phe Met Leu Ile			
1	5	10	15
gtt gct ggt ccc gtg gat ctg aat gag aac agc gag caa aag gaa aat			96
Val Ala Gly Pro Val Asp Leu Asn Glu Asn Ser Glu Gln Lys Glu Asn			
20	25	30	
gtg gaa aaa gag ggg ctg tgt aat gca tgt atg tgg aga caa aac act			144
Val Glu Lys Glu Gly Leu Cys Asn Ala Cys Met Trp Arg Gln Asn Thr			
35	40	45	
aaa tct tca aga cta gaa gcc ata aaa att caa atc ctc agt aaa ctt			192
Lys Ser Ser Arg Leu Glu Ala Ile Lys Ile Gln Ile Leu Ser Lys Leu			
50	55	60	
cgc ctg gaa aca gct cct aac att agc aaa gat gct ata aga caa ctt			240
Arg Leu Glu Thr Ala Pro Asn Ile Ser Lys Asp Ala Ile Arg Gln Leu			
65	70	75	80
ttg ccc aaa gct cct cca ctc cgg gaa ctg att gat cag tac gat gtc			288
Leu Pro Lys Ala Pro Pro Leu Arg Glu Leu Ile Asp Gln Tyr Asp Val			
85	90	95	
cag aga gat gac agc agt gat ggc tcc ttg gaa gat gat gat tat cac			336
Gln Arg Asp Asp Ser Ser Asp Gly Ser Leu Glu Asp Asp Asp Tyr His			
100	105	110	
gct acg acg gaa acg atc att acc atg cct aca gag tct gat ctt cta			384
Ala Thr Thr Glu Thr Ile Ile Thr Met Pro Thr Glu Ser Asp Leu Leu			
115	120	125	
atg caa gtg gaa gga aaa ccc aaa tgc tgc ttc ttt aaa ttt agc tct			432
Met Gln Val Glu Gly Lys Pro Lys Cys Cys Phe Phe Lys Phe Ser Ser			
130	135	140	
aaa ata caa tac aat aaa gta gta aag gcc caa ctg tgg ata tat ctg			480
Lys Ile Gln Tyr Asn Lys Val Val Lys Ala Gln Leu Trp Ile Tyr Leu			
145	150	155	160
aga ccc gtc aag act cct aca aca gtg ttt gtg caa atc ctg aga ctc			528
Arg Pro Val Lys Thr Pro Thr Thr Val Phe Val Gln Ile Leu Arg Leu			
165	170	175	

[0405]

atc aaa ccc atg aaa gac ggt aca agg tat act gga atc cga tct ctg Ile Lys Pro Met Lys Asp Gly Thr Arg Tyr Thr Gly Ile Arg Ser Leu 180 185 190	576
aaa ctt gac atg aac cca ggc act ggt att tgg cag agc att gat gtg Lys Leu Asp Met Asn Pro Gly Thr Gly Ile Trp Gln Ser Ile Asp Val 195 200 205	624
aag aca gtg ttg caa aat tgg ctc aaa caa cct gaa tcc aac tta ggc Lys Thr Val Leu Gln Asn Trp Leu Lys Gln Pro Glu Ser Asn Leu Gly 210 215 220	672
att gaa atc aaa gct tta gat gag aat ggt cat gat ctt gct gta acc Ile Glu Ile Lys Ala Leu Asp Glu Asn Gly His Asp Leu Ala Val Thr 225 230 235 240	720
ttc cca gga cca gga gaa gat ggg ctg aat ccc ttt tta gaa gtc aag Phe Pro Gly Pro Gly Glu Asp Gly Leu Asn Pro Phe Leu Glu Val Lys 245 250 255	768
gta aca gac aca cca aaa aga tcc agg aga gat ttt gga ctc gac tgt Val Thr Asp Thr Pro Lys Arg Ser Arg Arg Asp Phe Gly Leu Asp Cys 260 265 270	816
gat gag cac tca aca gaa tct cga tgc tgt cgt tac cct cta act gtg Asp Glu His Ser Thr Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr Val 275 280 285	864
gat ttt gaa gct ttt gga tgg gac tgg att att gca ccc aaa aga tat Asp Phe Glu Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg Tyr 290 295 300	912
aag gcc aat tac tgc tct gga gag tgt gaa ttt gta ttt tta caa aaa Lys Ala Asn Tyr Cys Ser Gly Glu Cys Glu Phe Val Phe Leu Gln Lys 305 310 315 320	960
tac cct cac act cat ctt gtg cac caa gca aac ccc aga ggt tca gca Tyr Pro His Thr His Leu Val His Gln Ala Asn Pro Arg Gly Ser Ala 325 330 335	1008
ggc ccc tgc tgt act ccc aca aag atg tct cca atc aat atg cta tat Gly Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr 340 345 350	1056

[0406]

ttt aat ggc aaa gaa caa ata ata tat ggg aaa att cca gcc atg gta 1104
 Phe Asn Gly Lys Glu Gln Ile Ile Tyr Gly Lys Ile Pro Ala Met Val
 355 360 365

gta gat cgc tgt ggg tgc tca tga 1128
 Val Asp Arg Cys Gly Cys Ser
 370 375

<210> 14

<211> 375

<212> PRT

<213> 猪

<400> 14

Met Gln Lys Leu Gln Ile Tyr Val Tyr Ile Tyr Leu Phe Met Leu Ile
 1 5 10 15
 Val Ala Gly Pro Val Asp Leu Asn Glu Asn Ser Glu Gln Lys Glu Asn
 20 25 30
 Val Glu Lys Glu Gly Leu Cys Asn Ala Cys Met Trp Arg Gln Asn Thr
 35 40 45
 Lys Ser Ser Arg Leu Glu Ala Ile Lys Ile Gln Ile Leu Ser Lys Leu
 50 55 60
 Arg Leu Glu Thr Ala Pro Asn Ile Ser Lys Asp Ala Ile Arg Gln Leu
 65 70 75 80
 Leu Pro Lys Ala Pro Pro Leu Arg Glu Leu Ile Asp Gln Tyr Asp Val
 85 90 95
 Gln Arg Asp Asp Ser Ser Asp Gly Ser Leu Glu Asp Asp Asp Tyr His
 100 105 110
 Ala Thr Thr Glu Thr Ile Ile Thr Met Pro Thr Glu Ser Asp Leu Leu
 115 120 125
 Met Gln Val Glu Gly Lys Pro Lys Cys Cys Phe Phe Lys Phe Ser Ser
 130 135 140
 Lys Ile Gln Tyr Asn Lys Val Val Lys Ala Gln Leu Trp Ile Tyr Leu
 145 150 155 160
 Arg Pro Val Lys Thr Pro Thr Thr Val Phe Val Gln Ile Leu Arg Leu
 165 170 175
 Ile Lys Pro Met Lys Asp Gly Thr Arg Tyr Thr Gly Ile Arg Ser Leu
 180 185 190
 Lys Leu Asp Met Asn Pro Gly Thr Gly Ile Trp Gln Ser Ile Asp Val
 195 200 205
 Lys Thr Val Leu Gln Asn Trp Leu Lys Gln Pro Glu Ser Asn Leu Gly
 210 215 220

[0407]

Lys Ser Ser Arg Leu Glu Ala Ile Lys Ile Gln Ile Leu Ser Lys Leu	
50	55
60	
cgc ctg gaa aca gct cct aac atc agc aaa gat gct ata aga caa ctt	240
Arg Leu Glu Thr Ala Pro Asn Ile Ser Lys Asp Ala Ile Arg Gln Leu	
65	70
75	80
ttg ccc aag gct cct cca ctc cgg gaa ctg att gat cag tac gat gtc	288
Leu Pro Lys Ala Pro Pro Leu Arg Glu Leu Ile Asp Gln Tyr Asp Val	
85	90
95	
cag aga gat gac agc agc gac ggc tcc ttg gaa gac gat gac tac cac	336
Gln Arg Asp Asp Ser Ser Asp Gly Ser Leu Glu Asp Asp Asp Tyr His	
100	105
110	
gtt acg acg gaa acg gtc att acc atg ccc acg gag tct gat ctt cta	384
Val Thr Thr Glu Thr Val Ile Thr Met Pro Thr Glu Ser Asp Leu Leu	
115	120
125	
gca gaa gtg caa gaa aaa ccc aaa tgt tgc ttc ttt aaa ttt agc tct	432
Ala Glu Val Gln Glu Lys Pro Lys Cys Cys Phe Phe Lys Phe Ser Ser	
130	135
140	
aag ata caa cac aat aaa gta gta aag gcc caa ctg tgg ata tat ctg	480
Lys Ile Gln His Asn Lys Val Val Lys Ala Gln Leu Trp Ile Tyr Leu	
145	150
155	160
aga cct gtc aag act cct aca aca gtg ttt gtg caa atc ctg aga ctc	528
Arg Pro Val Lys Thr Pro Thr Thr Val Phe Val Gln Ile Leu Arg Leu	
165	170
175	
atc aaa ccc atg aaa gac ggt aca agg tat act gga atc cga tct ctg	576
Ile Lys Pro Met Lys Asp Gly Thr Arg Tyr Thr Gly Ile Arg Ser Leu	
180	185
190	
aaa ctt gac atg aac cca ggc act ggt att tgg cag agc att gat gtg	624
Lys Leu Asp Met Asn Pro Gly Thr Gly Ile Trp Gln Ser Ile Asp Val	
195	200
205	
aag aca gtg ttg caa aac tgg ctc aaa caa cct gaa tcc aac tta ggc	672
Lys Thr Val Leu Gln Asn Trp Leu Lys Gln Pro Glu Ser Asn Leu Gly	
210	215
220	
att gaa atc aaa gct tta gat gag aat ggt cat gat ctt gct gta acc	720

[0409]

Ile Glu Ile Lys Ala Leu Asp Glu Asn Gly His Asp Leu Ala Val Thr 225	230	235	240	
ttc cca gaa cca gga gaa gaa gga ctg aat cct ttt tta gaa gtc aag Phe Pro Glu Pro Gly Glu Glu Gly Leu Asn Pro Phe Leu Glu Val Lys	245	250	255	768
gta aca gac aca cca aaa aga tct agg aga gat ttt ggg ctt gat tgt Val Thr Asp Thr Pro Lys Arg Ser Arg Arg Asp Phe Gly Leu Asp Cys	260	265	270	816
gat gag cac tcc aca gaa tct cga tgc tgt cgt tac cct cta act gtg Asp Glu His Ser Thr Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr Val	275	280	285	864
gat ttt gaa gct ttt gga tgg gat tgg att att gca cct aaa aga tat Asp Phe Glu Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg Tyr	290	295	300	912
aag gcc aat tac tgc tct gga gaa tgt gaa ttt tta ttt ttg caa aag Lys Ala Asn Tyr Cys Ser Gly Glu Cys Glu Phe Leu Phe Leu Gln Lys	305	310	315	960
tat cct cat acc cat ctt gtg cac caa gca aac ccc aaa ggt tca gcc Tyr Pro His Thr His Leu Val His Gln Ala Asn Pro Lys Gly Ser Ala	325	330	335	1008
ggc cct tgc tgt act cct aca aag atg tct cca att aat atg cta tat Gly Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr	340	345	350	1056
ttt aat ggc aaa gaa caa ata ata tat ggg aag att cca ggc atg gta Phe Asn Gly Lys Glu Gln Ile Ile Tyr Gly Lys Ile Pro Gly Met Val	355	360	365	1104
gta gat cgc tgt ggg tgc tca tga Val Asp Arg Cys Gly Cys Ser	370	375		1128

<210> 16

<211> 375

<212> PRT

<213> 羊

[0410]

<400> 16
 Met Gln Lys Leu Gln Ile Phe Val Tyr Ile Tyr Leu Phe Met Leu Leu
 1 5 10 15
 Val Ala Gly Pro Val Asp Leu Asn Glu Asn Ser Glu Gln Lys Glu Asn
 20 25 30
 Val Glu Lys Lys Gly Leu Cys Asn Ala Cys Leu Trp Arg Gln Asn Asn
 35 40 45
 Lys Ser Ser Arg Leu Glu Ala Ile Lys Ile Gln Ile Leu Ser Lys Leu
 50 55 60
 Arg Leu Glu Thr Ala Pro Asn Ile Ser Lys Asp Ala Ile Arg Gln Leu
 65 70 75 80
 Leu Pro Lys Ala Pro Pro Leu Arg Glu Leu Ile Asp Gln Tyr Asp Val
 85 90 95
 Gln Arg Asp Asp Ser Ser Asp Gly Ser Leu Glu Asp Asp Asp Tyr His
 100 105 110
 Val Thr Thr Glu Thr Val Ile Thr Met Pro Thr Glu Ser Asp Leu Leu
 115 120 125
 Ala Glu Val Gln Glu Lys Pro Lys Cys Cys Phe Phe Lys Phe Ser Ser
 130 135 140
 Lys Ile Gln His Asn Lys Val Val Lys Ala Gln Leu Trp Ile Tyr Leu
 145 150 155 160
 Arg Pro Val Lys Thr Pro Thr Thr Val Phe Val Gln Ile Leu Arg Leu
 165 170 175
 Ile Lys Pro Met Lys Asp Gly Thr Arg Tyr Thr Gly Ile Arg Ser Leu
 180 185 190
 Lys Leu Asp Met Asn Pro Gly Thr Gly Ile Trp Gln Ser Ile Asp Val
 195 200 205
 Lys Thr Val Leu Gln Asn Trp Leu Lys Gln Pro Glu Ser Asn Leu Gly
 210 215 220
 Ile Glu Ile Lys Ala Leu Asp Glu Asn Gly His Asp Leu Ala Val Thr
 225 230 235 240
 Phe Pro Glu Pro Gly Glu Glu Gly Leu Asn Pro Phe Leu Glu Val Lys
 245 250 255
 Val Thr Asp Thr Pro Lys Arg Ser Arg Arg Asp Phe Gly Leu Asp Cys
 260 265 270
 Asp Glu His Ser Thr Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr Val
 275 280 285
 Asp Phe Glu Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg Tyr
 290 295 300
 Lys Ala Asn Tyr Cys Ser Gly Glu Cys Glu Phe Leu Phe Leu Gln Lys
 305 310 315 320
 Tyr Pro His Thr His Leu Val His Gln Ala Asn Pro Lys Gly Ser Ala
 325 330 335

[0411]

gcc aca acc gaa acg att atc aca atg cct acg gag tct gat ttt ctt Ala Thr Thr Glu Thr Ile Ile Thr Met Pro Thr Glu Ser Asp Phe Leu 115 120 125	384
gta caa atg gag gga aaa cca aaa tgt tgc ttc ttt aag ttt agc tct Val Gln Met Glu Gly Lys Pro Lys Cys Cys Phe Phe Lys Phe Ser Ser 130 135 140	432
aaa ata caa tat aac aaa gta gta aag gca caa tta tgg ata tac ttg Lys Ile Gln Tyr Asn Lys Val Val Lys Ala Gln Leu Trp Ile Tyr Leu 145 150 155 160	480
agg caa gtc caa aaa cct aca acg gtg ttt gtg cag atc ctg aga ctc Arg Gln Val Gln Lys Pro Thr Thr Val Phe Val Gln Ile Leu Arg Leu 165 170 175	528
att aaa ccc atg aaa gac ggt aca aga tat act gga att cga tct ttg Ile Lys Pro Met Lys Asp Gly Thr Arg Tyr Thr Gly Ile Arg Ser Leu 180 185 190	576
aaa ctt gac atg aac cca ggc act ggt atc tgg cag agt att gat gtg Lys Leu Asp Met Asn Pro Gly Thr Gly Ile Trp Gln Ser Ile Asp Val 195 200 205	624
aag aca gtg ttg caa aat tgg ctc aaa cag cct gaa tcc aat tta ggc Lys Thr Val Leu Gln Asn Trp Leu Lys Gln Pro Glu Ser Asn Leu Gly 210 215 220	672
atc gaa ata aaa gct ttt gat gag aat gga cga gat ctt gct gta aca Ile Glu Ile Lys Ala Phe Asp Glu Asn Gly Arg Asp Leu Ala Val Thr 225 230 235 240	720
ttc cca gga cca ggt gaa gat gga ctg aac cca ttt tta gag gtc aga Phe Pro Gly Pro Gly Glu Asp Gly Leu Asn Pro Phe Leu Glu Val Arg 245 250 255	768
gtt aca gac aca cca aaa cgg tcc cgc aga gat ttt ggc ctt gac tgc Val Thr Asp Thr Pro Lys Arg Ser Arg Arg Asp Phe Gly Leu Asp Cys 260 265 270	816
gac gag cac tca acg gaa tct cga tgt tgt cgc tac ccg ctg aca gtg Asp Glu His Ser Thr Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr Val 275 280 285	864

[0413]

gat ttt gaa gct ttt gga tgg gac tgg att ata gca cct aaa aga tac 912
 Asp Phe Glu Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg Tyr
 290 295 300

aaa gcc aat tac tgc tct gga gaa tgt gaa ttc gta ttt cta cag aaa 960
 Lys Ala Asn Tyr Cys Ser Gly Glu Cys Glu Phe Val Phe Leu Gln Lys
 305 310 315 320

tac ccg cac act cac ctg gta cac caa gca aat cca aga ggc tca gca 1008
 Tyr Pro His Thr His Leu Val His Gln Ala Asn Pro Arg Gly Ser Ala
 325 330 335

ggc cct tgc tgc aca ccc acc aag atg tcc cct ata aac atg ctg tat 1056
 Gly Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr
 340 345 350

ttc aat gga aaa gaa caa ata ata tat gga aag ata cca gcc atg gtt 1104
 Phe Asn Gly Lys Glu Gln Ile Ile Tyr Gly Lys Ile Pro Ala Met Val
 355 360 365

gta gat cgt tgc ggg tgc tca tga 1128
 Val Asp Arg Cys Gly Cys Ser
 370 375

<210> 18

<211> 375

<212> PRT

<213> 吐绶鸡

<400> 18

Met Gln Lys Leu Ala Val Tyr Val Tyr Ile Tyr Leu Phe Met Gln Ile
 1 5 10 15
 Leu Val His Pro Val Ala Leu Asp Gly Ser Ser Gln Pro Thr Glu Asn
 20 25 30
 Ala Glu Lys Asp Gly Leu Cys Asn Ala Cys Thr Trp Arg Gln Asn Thr
 35 40 45
 Lys Ser Ser Arg Ile Glu Ala Ile Lys Ile Gln Ile Leu Ser Lys Leu
 50 55 60
 Arg Leu Glu Gln Ala Pro Asn Ile Ser Arg Asp Val Ile Lys Gln Leu
 65 70 75 80
 Leu Pro Lys Ala Pro Pro Leu Gln Glu Leu Ile Asp Gln Tyr Asp Val
 85 90 95

[0414]

Gln Arg Asp Asp Ser Ser Asp Gly Ser Leu Glu Asp Asp Asp Tyr His
 100 105 110
 Ala Thr Thr Glu Thr Ile Ile Thr Met Pro Thr Glu Ser Asp Phe Leu
 115 120 125
 Val Gln Met Glu Gly Lys Pro Lys Cys Cys Phe Phe Lys Phe Ser Ser
 130 135 140
 Lys Ile Gln Tyr Asn Lys Val Val Lys Ala Gln Leu Trp Ile Tyr Leu
 145 150 155 160
 Arg Gln Val Gln Lys Pro Thr Thr Val Phe Val Gln Ile Leu Arg Leu
 165 170 175
 Ile Lys Pro Met Lys Asp Gly Thr Arg Tyr Thr Gly Ile Arg Ser Leu
 180 185 190
 Lys Leu Asp Met Asn Pro Gly Thr Gly Ile Trp Gln Ser Ile Asp Val
 195 200 205
 Lys Thr Val Leu Gln Asn Trp Leu Lys Gln Pro Glu Ser Asn Leu Gly
 210 215 220
 Ile Glu Ile Lys Ala Phe Asp Glu Asn Gly Arg Asp Leu Ala Val Thr
 225 230 235 240
 Phe Pro Gly Pro Gly Glu Asp Gly Leu Asn Pro Phe Leu Glu Val Arg
 245 250 255
 Val Thr Asp Thr Pro Lys Arg Ser Arg Arg Asp Phe Gly Leu Asp Cys
 260 265 270
 Asp Glu His Ser Thr Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr Val
 275 280 285
 Asp Phe Glu Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg Tyr
 290 295 300
 Lys Ala Asn Tyr Cys Ser Gly Glu Cys Glu Phe Val Phe Leu Gln Lys
 305 310 315 320
 Tyr Pro His Thr His Leu Val His Gln Ala Asn Pro Arg Gly Ser Ala
 325 330 335
 Gly Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr
 340 345 350
 Phe Asn Gly Lys Glu Gln Ile Ile Tyr Gly Lys Ile Pro Ala Met Val
 355 360 365
 Val Asp Arg Cys Gly Cys Ser
 370 375

<210> 19

<211> 1125

<212> DNA

<213> 斑马鱼

<220>

<221> CDS

[0415]

<222> (1)... (1122)

<400> 19

atg cat ttt aca cag gtt tta att tct cta agt gta tta att gca tgt	48
Met His Phe Thr Gln Val Leu Ile Ser Leu Ser Val Leu Ile Ala Cys	
1 5 10 15	
ggt cca gtg ggt tat gga gat ata acg gcg cac cag cag cct tcc aca	96
Gly Pro Val Gly Tyr Gly Asp Ile Thr Ala His Gln Gln Pro Ser Thr	
20 25 30	
gcc acg gag gaa agc gag ctg tgt tcc aca tgt gag ttc aga caa cac	144
Ala Thr Glu Glu Ser Glu Leu Cys Ser Thr Cys Glu Phe Arg Gln His	
35 40 45	
agc aag ctg atg aga ctg cat gcc atc aag tcc caa att ctt agc aaa	192
Ser Lys Leu Met Arg Leu His Ala Ile Lys Ser Gln Ile Leu Ser Lys	
50 55 60	
ctc cga ctc aag cag gct cca aac atc agc cgg gac gtg gtc aag cag	240
Leu Arg Leu Lys Gln Ala Pro Asn Ile Ser Arg Asp Val Val Lys Gln	
65 70 75 80	
ctg tta ccc aaa gca ccg cct ttg caa caa ctt ctg gat cag tac gat	288
Leu Leu Pro Lys Ala Pro Pro Leu Gln Gln Leu Leu Asp Gln Tyr Asp	
85 90 95	
gtt tta gga gat gac agt aag gat gga gct gtg gaa gag gac gat gaa	336
Val Leu Gly Asp Asp Ser Lys Asp Gly Ala Val Glu Glu Asp Asp Glu	
100 105 110	
cat gcc acc aca gag acc atc atg acc atg gcc aca gaa cct gac ccc	384
His Ala Thr Thr Glu Thr Ile Met Thr Met Ala Thr Glu Pro Asp Pro	
115 120 125	
att gtt caa gta gat cgg aaa ccg aag tgt tgc ttt ttc tcc ttc agt	432
Ile Val Gln Val Asp Arg Lys Pro Lys Cys Cys Phe Phe Ser Phe Ser	
130 135 140	
ccg aag atc caa gcg aac ccg atc gta aga gcg cag ctc tgg gtt cat	480
Pro Lys Ile Gln Ala Asn Arg Ile Val Arg Ala Gln Leu Trp Val His	
145 150 155 160	
ctg aga ccg gcg gag gag gcg acc acc gtc ttc tta cag ata tct cgg	528

[0416]

Leu Arg Pro Ala Glu Glu Ala Thr Thr Val Phe Leu Gln Ile Ser Arg	
165	170
ctg atg ccc gtt aag gac gga gga aga cac cga ata cga tcc ctg aaa	576
Leu Met Pro Val Lys Asp Gly Gly Arg His Arg Ile Arg Ser Leu Lys	
180	185
atc gac gtg aac gca gga gtc acg tct tgg cag agt ata gac gta aag	624
Ile Asp Val Asn Ala Gly Val Thr Ser Trp Gln Ser Ile Asp Val Lys	
195	200
cag gtg ctc acg gtg tgg tta aaa caa ccg gag acc aac cga ggc atc	672
Gln Val Leu Thr Val Trp Leu Lys Gln Pro Glu Thr Asn Arg Gly Ile	
210	215
gag att aac gca tat gac gcg aag gga aac gac ttg gcc gtc act tca	720
Glu Ile Asn Ala Tyr Asp Ala Lys Gly Asn Asp Leu Ala Val Thr Ser	
225	230
acc gag act ggg gag gat gga ctg ctc ccc ttt atg gag gtg aaa ata	768
Thr Glu Thr Gly Glu Asp Gly Leu Leu Pro Phe Met Glu Val Lys Ile	
245	250
tca gag ggc cca aaa cga atc cgg agg gac tcc gga ctg gac tgc gat	816
Ser Glu Gly Pro Lys Arg Ile Arg Arg Asp Ser Gly Leu Asp Cys Asp	
260	265
gag aat tcc tca gag tct cgc tgc tgc agg tac cct ctc act gtg gac	864
Glu Asn Ser Ser Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr Val Asp	
275	280
ttc gag gac ttt ggc tgg gac tgg att att gct cca aaa cgc tat aag	912
Phe Glu Asp Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg Tyr Lys	
290	295
gcg aat tac tgt tca gga gaa tgc gac tac atg tac ctg cag aag tat	960
Ala Asn Tyr Cys Ser Gly Glu Cys Asp Tyr Met Tyr Leu Gln Lys Tyr	
305	310
ccc cac acc cat ctg gtg aac aag gcc agt ccg aga gga acg gct ggg	1008
Pro His Thr His Leu Val Asn Lys Ala Ser Pro Arg Gly Thr Ala Gly	
325	330
ccc tgc tgc act ccc acc aag atg tct ccc atc aac atg ctt tac ttt	1056

[0417]

Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr Phe
 340 345 350

aac ggc aaa gag cag atc atc tac ggc aag atc cct tcg atg gta gta 1104
 Asn Gly Lys Glu Gln Ile Ile Tyr Gly Lys Ile Pro Ser Met Val Val
 355 360 365

gac cgc tgt ggc tgc tca tga 1125
 Asp Arg Cys Gly Cys Ser
 370

<210> 20

<211> 374

<212> PRT

<213> 斑马鱼

<400> 20

Met His Phe Thr Gln Val Leu Ile Ser Leu Ser Val Leu Ile Ala Cys
 1 5 10 15
 Gly Pro Val Gly Tyr Gly Asp Ile Thr Ala His Gln Gln Pro Ser Thr
 20 25 30
 Ala Thr Glu Glu Ser Glu Leu Cys Ser Thr Cys Glu Phe Arg Gln His
 35 40 45
 Ser Lys Leu Met Arg Leu His Ala Ile Lys Ser Gln Ile Leu Ser Lys
 50 55 60
 Leu Arg Leu Lys Gln Ala Pro Asn Ile Ser Arg Asp Val Val Lys Gln
 65 70 75 80
 Leu Leu Pro Lys Ala Pro Pro Leu Gln Gln Leu Leu Asp Gln Tyr Asp
 85 90 95
 Val Leu Gly Asp Asp Ser Lys Asp Gly Ala Val Glu Glu Asp Asp Glu
 100 105 110
 His Ala Thr Thr Glu Thr Ile Met Thr Met Ala Thr Glu Pro Asp Pro
 115 120 125
 Ile Val Gln Val Asp Arg Lys Pro Lys Cys Cys Phe Phe Ser Phe Ser
 130 135 140
 Pro Lys Ile Gln Ala Asn Arg Ile Val Arg Ala Gln Leu Trp Val His
 145 150 155 160
 Leu Arg Pro Ala Glu Glu Ala Thr Thr Val Phe Leu Gln Ile Ser Arg
 165 170 175
 Leu Met Pro Val Lys Asp Gly Gly Arg His Arg Ile Arg Ser Leu Lys
 180 185 190
 Ile Asp Val Asn Ala Gly Val Thr Ser Trp Gln Ser Ile Asp Val Lys
 195 200 205

[0418]

Gln Val Leu Thr Val Trp Leu Lys Gln Pro Glu Thr Asn Arg Gly Ile
 210 215 220
 Glu Ile Asn Ala Tyr Asp Ala Lys Gly Asn Asp Leu Ala Val Thr Ser
 225 230 235 240
 Thr Glu Thr Gly Glu Asp Gly Leu Leu Pro Phe Met Glu Val Lys Ile
 245 250 255
 Ser Glu Gly Pro Lys Arg Ile Arg Arg Asp Ser Gly Leu Asp Cys Asp
 260 265 270
 Glu Asn Ser Ser Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr Val Asp
 275 280 285
 Phe Glu Asp Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg Tyr Lys
 290 295 300
 Ala Asn Tyr Cys Ser Gly Glu Cys Asp Tyr Met Tyr Leu Gln Lys Tyr
 305 310 315 320
 Pro His Thr His Leu Val Asn Lys Ala Ser Pro Arg Gly Thr Ala Gly
 325 330 335
 Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr Phe
 340 345 350
 Asn Gly Lys Glu Gln Ile Ile Tyr Gly Lys Ile Pro Ser Met Val Val
 355 360 365
 Asp Arg Cys Gly Cys Ser
 370

<210> 21

<211> 4

<212> PRT

<213> 人工序列

<220>

<223> 蛋白水解裂解位点

<221> 变体

<222> (0)... (0)

<223> Xaa = 任何氨基酸

<400> 21

Arg Xaa Xaa Arg

1

<210> 22

<211> 4

<212> PRT

<213> 真核细胞

[0419]

<220>
 <221> SITE
 <222> (0)... (0)
 <223> 蛋白水解加工位点

<400> 22
 Arg Ser Arg Arg
 1

<210> 23
 <211> 4
 <212> PRT
 <213> 真核细胞

<220>
 <221> SITE
 <222> (0)... (0)
 <223> 蛋白水解加工位点

<400> 23
 Arg Ile Arg Arg
 1

<210> 24
 <211> 1393
 <212> DNA
 <213> 人类

<220>
 <221> CDS
 <222> (54)... (1274)
 <223> GDF-11

<400> 24
 ccgcgggact cggcgtecc cgccccccag tctccctcc cctcccctcc agc atg 56
 Met
 1

gtg ctc gcg gcc ccg ctg ctg ctg ggc ttc ctg ctc ctc gcc ctg gag 104
 Val Leu Ala Ala Pro Leu Leu Leu Gly Phe Leu Leu Leu Ala Leu Glu
 5 10 15

ctg cgg ccc cgg ggg gag gcg gcc gag ggc ccc gcg gcg gcg gcg gcg 152
 Leu Arg Pro Arg Gly Glu Ala Ala Glu Gly Pro Ala Ala Ala Ala Ala

[0420]

20	25	30	
gcg gcg gcg gcg gcg gca gcg gcg ggg gtc ggg ggg gag cgc tcc agc Ala Ala Ala Ala Ala Ala Ala Ala Gly Val Gly Gly Glu Arg Ser Ser 35 40 45			200
cgg cca gcc ccg tcc gtg gcg ccc gag ccg gac ggc tgc ccc gtg tgc Arg Pro Ala Pro Ser Val Ala Pro Glu Pro Asp Gly Cys Pro Val Cys 50 55 60 65			248
gtt tgg cgg cag cac agc cgc gag ctg cgc cta gag agc atc aag tcc Val Trp Arg Gln His Ser Arg Glu Leu Arg Leu Glu Ser Ile Lys Ser 70 75 80			296
cag atc ttg agc aaa ctg ccg ctc aag gag gcg ccc aac atc agc cgc Gln Ile Leu Ser Lys Leu Arg Leu Lys Glu Ala Pro Asn Ile Ser Arg 85 90 95			344
gag gtg gtg aag cag ctg ctg ccc aag gcg ccg ccg ctg cag cag atc Glu Val Val Lys Gln Leu Leu Pro Lys Ala Pro Pro Leu Gln Gln Ile 100 105 110			392
ctg gac cta cac gac ttc cag ggc gac gcg ctg cag ccc gag gac ttc Leu Asp Leu His Asp Phe Gln Gly Asp Ala Leu Gln Pro Glu Asp Phe 115 120 125			440
ctg gag gag gac gag tac cac gcc acc acc gag acc gtc att agc atg Leu Glu Glu Asp Glu Tyr His Ala Thr Thr Glu Thr Val Ile Ser Met 130 135 140 145			488
gcc cag gag acg gac cca gca gta cag aca gat ggc agc cct ctc tgc Ala Gln Glu Thr Asp Pro Ala Val Gln Thr Asp Gly Ser Pro Leu Cys 150 155 160			536
tgc cat ttt cac ttc agc ccc aag gtg atg ttc aca aag gta ctg aag Cys His Phe His Phe Ser Pro Lys Val Met Phe Thr Lys Val Leu Lys 165 170 175			584
gcc cag ctg tgg gtg tac cta cgg cct gta ccc cgc cca gcc aca gtc Ala Gln Leu Trp Val Tyr Leu Arg Pro Val Pro Arg Pro Ala Thr Val 180 185 190			632
tac ctg cag atc ttg cga cta aaa ccc cta act ggg gaa ggg acc gca Tyr Leu Gln Ile Leu Arg Leu Lys Pro Leu Thr Gly Glu Gly Thr Ala			680

[0421]

195	200	205	
ggg gga ggg ggc gga ggc cgg cgt cac atc cgt atc cgc tca ctg aag			728
Gly Gly Gly Gly Gly Gly Arg Arg His Ile Arg Ile Arg Ser Leu Lys			
210	215	220	225
att gag ctg cac tca cgc tca ggc cat tgg cag agc atc gac ttc aag			776
Ile Glu Leu His Ser Arg Ser Gly His Trp Gln Ser Ile Asp Phe Lys			
	230	235	240
caa gtg cta cac agc tgg ttc cgc cag cca cag agc aac tgg ggc atc			824
Gln Val Leu His Ser Trp Phe Arg Gln Pro Gln Ser Asn Trp Gly Ile			
	245	250	255
gag atc aac gcc ttt gat ccc agt ggc aca gac ctg gct gtc acc tcc			872
Glu Ile Asn Ala Phe Asp Pro Ser Gly Thr Asp Leu Ala Val Thr Ser			
	260	265	270
ctg ggg cgg gga gcc gag ggg ctg cat cca ttc atg gag ctt cga gtc			920
Leu Gly Pro Gly Ala Glu Gly Leu His Pro Phe Met Glu Leu Arg Val			
	275	280	285
cta gag aac aca aaa cgt tcc cgg cgg aac ctg ggt ctg gac tgc gac			968
Leu Glu Asn Thr Lys Arg Ser Arg Arg Asn Leu Gly Leu Asp Cys Asp			
	290	295	300
gag cac tca agc gag tcc cgc tgc tgc cga tat ccc ctc aca gtg gac			1016
Glu His Ser Ser Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr Val Asp			
	310	315	320
ttt gag gct ttc ggc tgg gac tgg atc atc gca cct aag cgc tac aag			1064
Phe Glu Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg Tyr Lys			
	325	330	335
gcc aac tac tgc tcc ggc cag tgc gag tac atg ttc atg caa aaa tat			1112
Ala Asn Tyr Cys Ser Gly Gln Cys Glu Tyr Met Phe Met Gln Lys Tyr			
	340	345	350
cgg cat acc cat ttg gtg cag cag gcc aat cca aga ggc tct gct ggg			1160
Pro His Thr His Leu Val Gln Gln Ala Asn Pro Arg Gly Ser Ala Gly			
	355	360	365
ccc tgt tgt acc ccc acc aag atg tcc cca atc aac atg ctc tac ttc			1208
Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr Phe			

[0422]

370	375	380	385	
aat gac aag cag cag att atc tac ggc aag atc cct ggc atg gtg gtg				1256
Asn Asp Lys Gln Gln Ile Ile Tyr Gly Lys Ile Pro Gly Met Val Val				
	390	395	400	
gat cgc tgt ggc tgc tct taagtgggtc actacaagct gctggagcaa				1304
Asp Arg Cys Gly Cys Ser				
	405			
agacttgggtg ggtgggtaac ttaacctctt cacagaggat aaaaaatgct tgtgagiatg				1364
acagaaggga ataaacaggc ttaaagggt				1393
<210> 25				
<211> 407				
<212> PRT				
<213> 人类				
<400> 25				
Met Val Leu Ala Ala Pro Leu Leu Leu Gly Phe Leu Leu Leu Ala Leu				
1	5	10	15	
Glu Leu Arg Pro Arg Gly Glu Ala Ala Glu Gly Pro Ala Ala Ala Ala				
	20	25	30	
Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Val Gly Gly Glu Arg Ser				
	35	40	45	
Ser Arg Pro Ala Pro Ser Val Ala Pro Glu Pro Asp Gly Cys Pro Val				
	50	55	60	
Cys Val Trp Arg Gln His Ser Arg Glu Leu Arg Leu Glu Ser Ile Lys				
65	70	75	80	
Ser Gln Ile Leu Ser Lys Leu Arg Leu Lys Glu Ala Pro Asn Ile Ser				
	85	90	95	
Arg Glu Val Val Lys Gln Leu Leu Pro Lys Ala Pro Pro Leu Gln Gln				
	100	105	110	
Ile Leu Asp Leu His Asp Phe Gln Gly Asp Ala Leu Gln Pro Glu Asp				
	115	120	125	
Phe Leu Glu Glu Asp Glu Tyr His Ala Thr Thr Glu Thr Val Ile Ser				
	130	135	140	
Met Ala Gln Glu Thr Asp Pro Ala Val Gln Thr Asp Gly Ser Pro Leu				
145	150	155	160	
Cys Cys His Phe His Phe Ser Pro Lys Val Met Phe Thr Lys Val Leu				
	165	170	175	
Lys Ala Gln Leu Trp Val Tyr Leu Arg Pro Val Pro Arg Pro Ala Thr				
	180	185	190	
Val Tyr Leu Gln Ile Leu Arg Leu Lys Pro Leu Thr Gly Glu Gly Thr				

[0423]

195	200	205
Ala Gly Gly Gly Gly Gly Gly Arg Arg His Ile Arg Ile Arg Ser Leu		
210	215	220
Lys Ile Glu Leu His Ser Arg Ser Gly His Trp Gln Ser Ile Asp Phe		
225	230	235
Lys Gln Val Leu His Ser Trp Phe Arg Gln Pro Gln Ser Asn Trp Gly		
245	250	255
Ile Glu Ile Asn Ala Phe Asp Pro Ser Gly Thr Asp Leu Ala Val Thr		
260	265	270
Ser Leu Gly Pro Gly Ala Glu Gly Leu His Pro Phe Met Glu Leu Arg		
275	280	285
Val Leu Glu Asn Thr Lys Arg Ser Arg Arg Asn Leu Gly Leu Asp Cys		
290	295	300
Asp Glu His Ser Ser Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr Val		
305	310	315
Asp Phe Glu Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg Tyr		
325	330	335
Lys Ala Asn Tyr Cys Ser Gly Gln Cys Glu Tyr Met Phe Met Gln Lys		
340	345	350
Tyr Pro His Thr His Leu Val Gln Gln Ala Asn Pro Arg Gly Ser Ala		
355	360	365
Gly Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr		
370	375	380
Phe Asn Asp Lys Gln Gln Ile Ile Tyr Gly Lys Ile Pro Gly Met Val		
385	390	395
Val Asp Arg Cys Gly Cys Ser		400
405		

<210> 26

<211> 476

<212> DNA

<213> 鲑鱼-1

<220>

<221> CDS

<222> (3)... (473)

<400> 26

gg cag ccg gag acg aat tgg ggg atc gag att aat gcg ttc gac tcg	47
Gln Pro Glu Thr Asn Trp Gly Ile Glu Ile Asn Ala Phe Asp Ser	
1 5 10 15	

aag gga aat gat ctg gcc gtt acc tca gca gaa gcg gga gaa gga ctg	95
Lys Gly Asn Asp Leu Ala Val Thr Ser Ala Glu Ala Gly Glu Gly Leu	

[0424]

	20	25	30	
caa ccc ttc atg gag gtg acg att tca gag ggc ccg aag cgc tcc agg				143
Gln Pro Phe Met Glu Val Thr Ile Ser Glu Gly Pro Lys Arg Ser Arg				
	35	40	45	
aga gac tcg ggc ctg gac tgt gac gag aac tcc ccc gag tcc cgc tgt				191
Arg Asp Ser Gly Leu Asp Cys Asp Glu Asn Ser Pro Glu Ser Arg Cys				
	50	55	60	
tgc cgc tac ccc ctc acg gta gac ttt gaa gac ttt ggc tgg gac tgg				239
Cys Arg Tyr Pro Leu Thr Val Asp Phe Glu Asp Phe Gly Trp Asp Trp				
	65	70	75	
att att gcc ccc aag cgc tac aag gcc aac tac tgc tct ggt gag tgt				287
Ile Ile Ala Pro Lys Arg Tyr Lys Ala Asn Tyr Cys Ser Gly Glu Cys				
	80	85	90	95
gag tac atg cac ctg cag aag tac ccc cac acc cac ctg gtg aac aag				335
Glu Tyr Met His Leu Gln Lys Tyr Pro His Thr His Leu Val Asn Lys				
	100	105	110	
gct aac cct cgc ggc acc gca ggg ccc tgc tgc acc ccc acc aag atg				383
Ala Asn Pro Arg Gly Thr Ala Gly Pro Cys Cys Thr Pro Thr Lys Met				
	115	120	125	
tcc ccc atc aac atg ctc tac ttc aac cgc aaa gag cag atc atc tac				431
Ser Pro Ile Asn Met Leu Tyr Phe Asn Arg Lys Glu Gln Ile Ile Tyr				
	130	135	140	
ggc aag atc ccc tcc atg gtg gtg gac cgt tgc gga tgc tcg				473
Gly Lys Ile Pro Ser Met Val Val Asp Arg Cys Gly Cys Ser				
	145	150	155	
tga				476
<210> 27				
<211> 157				
<212> PRT				
<213> 鲑鱼-1				
<400> 27				
Gln Pro Glu Thr Asn Trp Gly Ile Glu Ile Asn Ala Phe Asp Ser Lys				
	1	5	10	15

[0425]

Gly Asn Asp Leu Ala Val Thr Ser Ala Glu Ala Gly Glu Gly Leu Gln
 20 25 30
 Pro Phe Met Glu Val Thr Ile Ser Glu Gly Pro Lys Arg Ser Arg Arg
 35 40 45
 Asp Ser Gly Leu Asp Cys Asp Glu Asn Ser Pro Glu Ser Arg Cys Cys
 50 55 60
 Arg Tyr Pro Leu Thr Val Asp Phe Glu Asp Phe Gly Trp Asp Trp Ile
 65 70 75 80
 Ile Ala Pro Lys Arg Tyr Lys Ala Asn Tyr Cys Ser Gly Glu Cys Glu
 85 90 95
 Tyr Met His Leu Gln Lys Tyr Pro His Thr His Leu Val Asn Lys Ala
 100 105 110
 Asn Pro Arg Gly Thr Ala Gly Pro Cys Cys Thr Pro Thr Lys Met Ser
 115 120 125
 Pro Ile Asn Met Leu Tyr Phe Asn Arg Lys Glu Gln Ile Ile Tyr Gly
 130 135 140
 Lys Ile Pro Ser Met Val Val Asp Arg Cys Gly Cys Ser
 145 150 155

<210> 28

<211> 412

<212> DNA

<213> 鲑鱼-2

<220>

<221> CDS

<222> (2)... (409)

<400> 28

g gtt acc tca act gaa gcc gga gaa gga ctg caa ccc ttc atg gag gtg 49
 Val Thr Ser Thr Glu Ala Gly Glu Gly Leu Gln Pro Phe Met Glu Val
 1 5 10 15

aag att tcg gag ggc ccg aag cgc tcc agg aga gat tcg ggc ctg gac 97
 Lys Ile Ser Glu Gly Pro Lys Arg Ser Arg Arg Asp Ser Gly Leu Asp
 20 25 30

tgt gat gag aac tcc ccc gag tcc cgc tgc tgc egg tac ccc ctc acg 145
 Cys Asp Glu Asn Ser Pro Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr
 35 40 45

gtg gac ttt gaa gac ttt ggc tgg gac tgg att att gcc ccc aag cgc 193
 Val Asp Phe Glu Asp Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg
 50 55 60

[0426]

tac aag gcc aac tac tgc tct ggt gag tgc gag tac atg cac ctg cag 241
 Tyr Lys Ala Asn Tyr Cys Ser Gly Glu Cys Glu Tyr Met His Leu Gln
 65 70 75 80

aag tac ccc cac acc cac ctg gtg aac aag gct aac cct cgc ggc acc 289
 Lys Tyr Pro His Thr His Leu Val Asn Lys Ala Asn Pro Arg Gly Thr
 85 90 95

gcg ggg ccc tgc tgc acc ccc acc aag atg tcc ccc atc aac atg ctc 337
 Ala Gly Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu
 100 105 110

tac ttc aac cgc aaa gag cag atc atc tac ggc aag atc ccc tcc atg 385
 Tyr Phe Asn Arg Lys Glu Gln Ile Ile Tyr Gly Lys Ile Pro Ser Met
 115 120 125

gtg gtg gac cgc tgc ggc tgc tcg tga 412
 Val Val Asp Arg Cys Gly Cys Ser
 130 135

<210> 29
 <211> 136
 <212> PRT
 <213> 鲑鱼-2

<400> 29
 Val Thr Ser Thr Glu Ala Gly Glu Gly Leu Gln Pro Phe Met Glu Val
 1 5 10 15
 Lys Ile Ser Glu Gly Pro Lys Arg Ser Arg Arg Asp Ser Gly Leu Asp
 20 25 30
 Cys Asp Glu Asn Ser Pro Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr
 35 40 45
 Val Asp Phe Glu Asp Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg
 50 55 60
 Tyr Lys Ala Asn Tyr Cys Ser Gly Glu Cys Glu Tyr Met His Leu Gln
 65 70 75 80
 Lys Tyr Pro His Thr His Leu Val Asn Lys Ala Asn Pro Arg Gly Thr
 85 90 95
 Ala Gly Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu
 100 105 110
 Tyr Phe Asn Arg Lys Glu Gln Ile Ile Tyr Gly Lys Ile Pro Ser Met
 115 120 125
 Val Val Asp Arg Cys Gly Cys Ser
 130 135

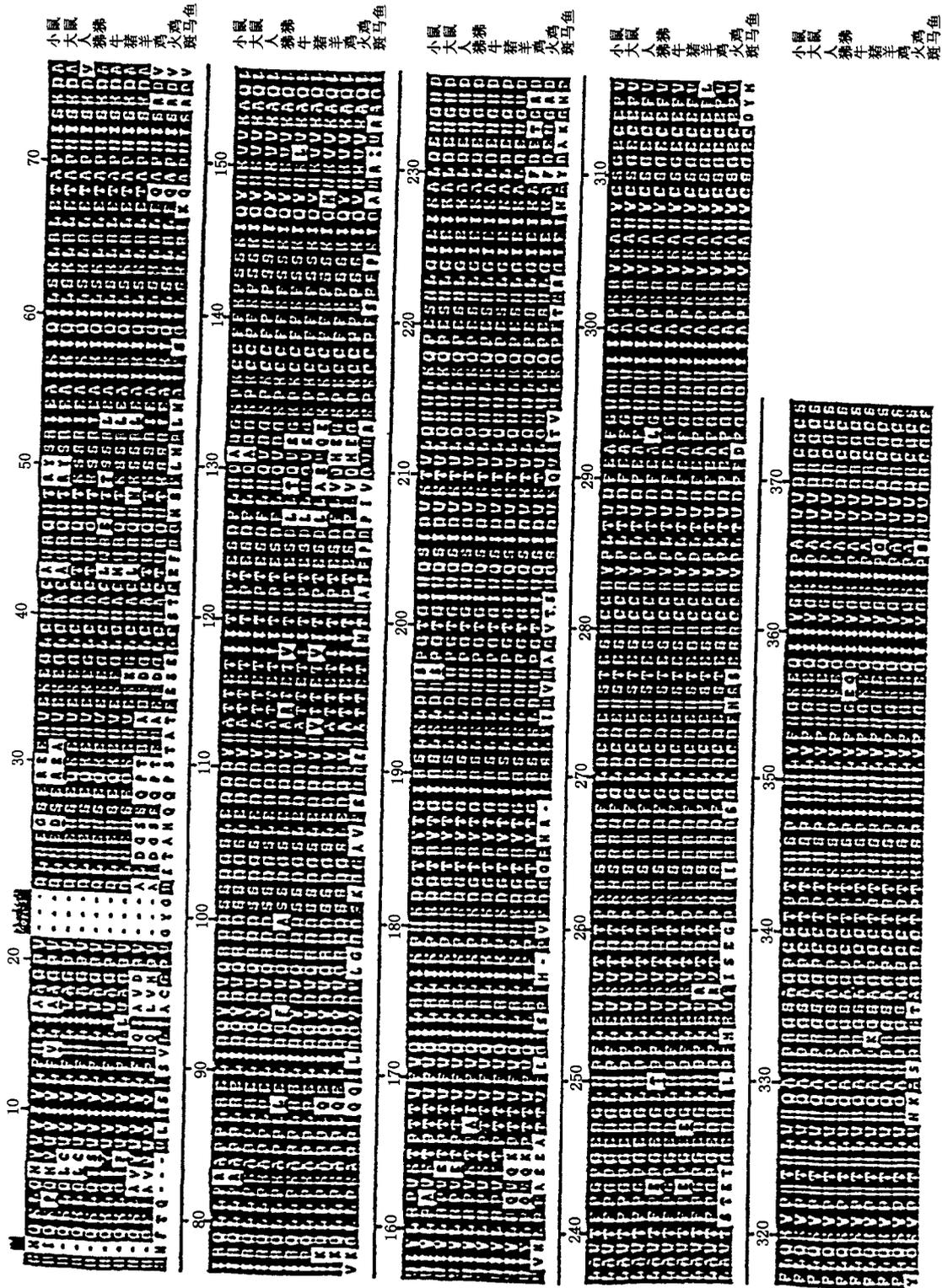


图 1

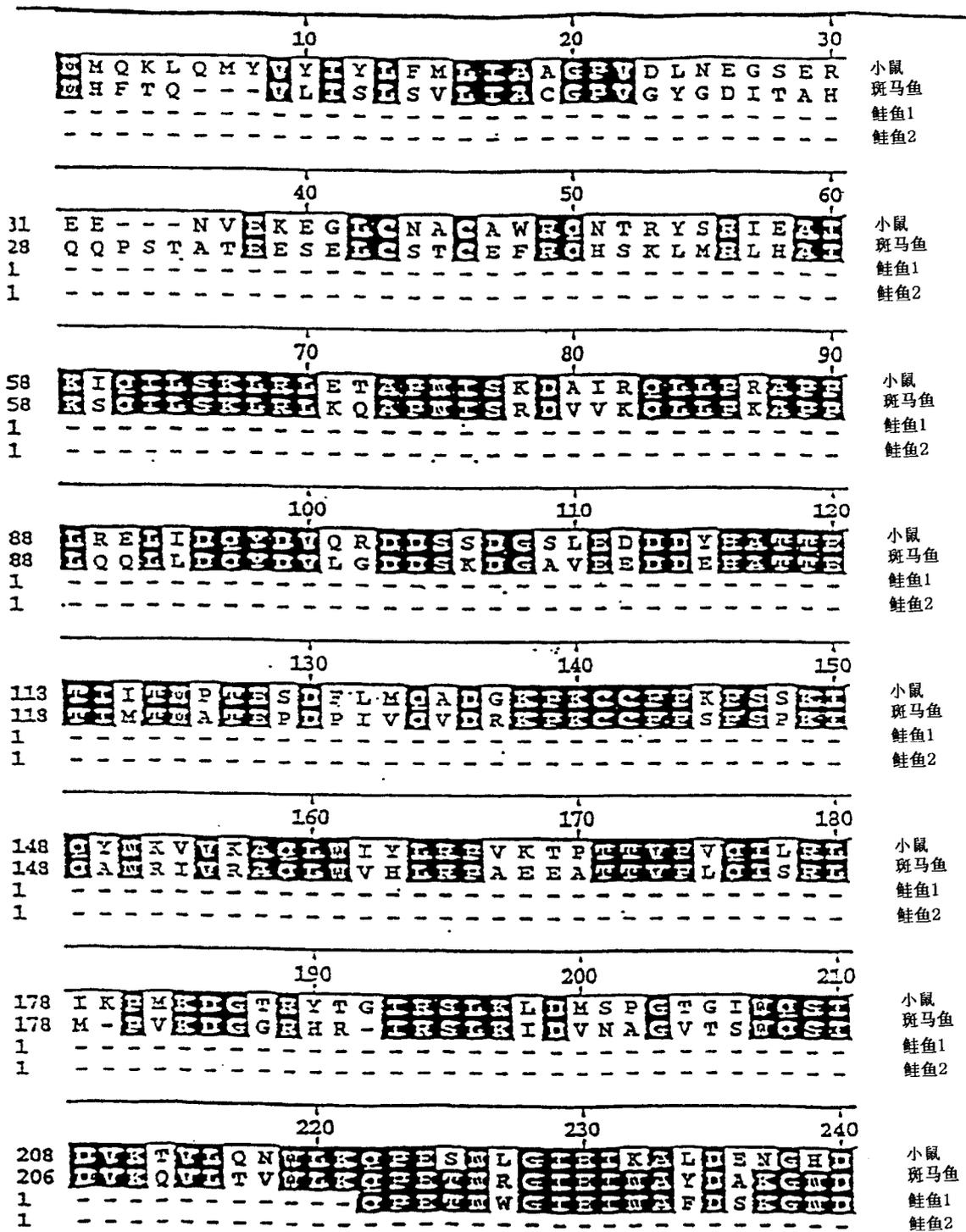


图 2

