

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公开说明书

[21] 申请号 200510092639.8

[51] Int. Cl.

C12N 15/12 (2006.01)
C07K 14/435 (2006.01)
C12N 15/63 (2006.01)
C12N 15/81 (2006.01)
C12N 15/85 (2006.01)
C12P 21/02 (2006.01)

[43] 公开日 2006年8月30日

[11] 公开号 CN 1824777A

[51] Int. Cl. (续)

C07K 16/18 (2006.01)
C07K 16/28 (2006.01)
A61K 38/17 (2006.01)
A61P 25/00 (2006.01)
A61P 9/00 (2006.01)
A61P 1/00 (2006.01)
G01N 33/53 (2006.01)

[22] 申请日 2000.9.25

[21] 申请号 200510092639.8

分案原申请号 00812265.2

[30] 优先权

[32] 1999. 9. 24 [33] NL [31] 1013140

[32] 1999. 9. 24 [33] EP [31] 99203140.1

[32] 2000. 7. 28 [33] EP [31] 00202683.9

[32] 2000. 7. 31 [33] US [31] 60/222,047

[71] 申请人 索尔瓦药物有限公司

地址 荷兰韦斯普

[72] 发明人 W·德里尔斯尼德尔 C·比戈尔

C·劳肯 G·尼斯 J·维尼玛

[74] 专利代理机构 中国国际贸易促进委员会专利商
标事务所

代理人 刘晓东

权利要求书7页 说明书88页 附图9页

[54] 发明名称

新的人 G 蛋白偶联受体

[57] 摘要

本发明的多核苷酸和多肽涉及 G 蛋白偶联受体家族,称为 IGS4 家族。本发明还涉及抑制或激活这些多核苷酸和多肽的作用,包含所述多核苷酸的载体、包含这种载体的宿主细胞、和 IGS4 基因过度表达、错误表达、表达不足、或受遏制(敲除动物)的转基因动物。本发明还涉及用于筛选能够作为 IGS4 家族的激动剂或拮抗剂的化合物的方法,以及这些 IGS4 多肽、多核苷酸、激动剂、或拮抗剂的用途。本发明的优选用途涉及神经系统、胃肠系统、心血管系统、骨骼肌、和/或甲状腺的紊乱,和/或肺病、免疫学疾病、和泌尿生殖系统紊乱。本发明还涉及 IGS4 多肽的相关配体 - 称为神经介肽 U 的神经肽 - 的鉴定。

1. 包含选自下组的核苷酸序列的分离多核苷酸：
 - (a) 编码SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的IGS4多肽的核苷酸序列；
 - (b) 编码由荷兰Baarn的真菌菌种保藏中心保藏物编号CBS102221或保藏物编号CBS102222所含DNA插入片段编码的多肽的核苷酸序列，特别是包含SEQ ID NO: 1、SEQ ID NO: 3、SEQ ID NO: 5、或SEQ ID NO: 7的核苷酸序列；
 - (c) 在其全长上与核苷酸序列(a)或(b)具有至少80%（优选至少90%）序列同一性的核苷酸序列；
 - (d) 与核苷酸序列(a)或(b)或(c)互补的核苷酸序列。
2. 权利要求1的多核苷酸，其中所述多核苷酸包含SEQ ID NO: 1中所含的编码SEQ ID NO: 2的IGS4多肽的核苷酸序列、SEQ ID NO: 3中所含的编码SEQ ID NO: 4的IGS4多肽的核苷酸序列、SEQ ID NO: 5中所含的编码SEQ ID NO: 6的IGS4多肽的核苷酸序列、或SEQ ID NO: 7中所含的编码SEQ ID NO: 8的IGS4多肽的核苷酸序列。
3. 权利要求1的多核苷酸，其中所述多核苷酸包含在其全长上与SEQ ID NO: 1、SEQ ID NO: 3、SEQ ID NO: 5、或SEQ ID NO: 7的核苷酸序列，或者与荷兰Baarn的真菌菌种保藏中心保藏物编号CBS102221或保藏物编号CBS102222所含DNA插入片段的序列至少80%同一的核苷酸序列。
4. 权利要求3的多核苷酸，其是SEQ ID NO: 1、SEQ ID NO: 3、SEQ ID NO: 5、或SEQ ID NO: 7的多核苷酸，或是荷兰Baarn的真菌菌种保藏中心保藏物编号CBS102221或保藏物编号CBS102222所含DNA插入片段的多核苷酸。
5. 权利要求1-4的多核苷酸，其是DNA或RNA。
6. 编码IGS4神经介肽受体蛋白、优选哺乳动物神经介肽受体蛋白的分离核苷酸序列，所述蛋白质展示以高亲和力结合神经介肽U，优

选神经介肽U-8、神经介肽U-23、和/或神经介肽U-25。

7. 权利要求6中编码IGS4神经介肽受体蛋白的分离核苷酸序列，所述蛋白质展示在脑、骨骼肌、小脑、睾丸、胼胝体、脊髓、黑质、髓质、丘脑、尾状核、脑桥、伏核、胎脑、胃、心、甲状腺、肺、胸腺、前列腺、和/或气管中表达。

8. 编码IGS4神经介肽受体蛋白、优选哺乳动物神经介肽受体蛋白的分离核苷酸序列，所述蛋白质展示以高亲和力结合神经介肽U，优选神经介肽U-8、神经介肽U-23、和/或神经介肽U-25，所述蛋白质展示在脑、骨骼肌、小脑、睾丸、胼胝体、脊髓、黑质、髓质、丘脑、尾状核、脑桥、伏核、胎脑、胃、心、甲状腺、肺、胸腺、前列腺、和/或气管中表达，且所述核苷酸序列选自权利要求1-5中定义的核苷酸序列。

9. 包含表达体系的DNA或RNA分子，其中当所述表达体系存在于相容宿主细胞中时能够生成包含与SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的多肽，或者与由荷兰Baarn的真菌菌种保藏中心保藏物编号CBS102221或保藏物编号CBS102222所含DNA插入片段编码的多肽具有至少80%同一性的氨基酸序列的IGS4多肽。

10. 包含表达体系的分离DNA或RNA分子，其中所述表达体系能够生成包含神经介肽受体蛋白、优选哺乳动物神经介肽受体蛋白氨基酸序列的IGS4多肽，所述蛋白质展示以高亲和力结合神经介肽U，优选神经介肽U-8、神经介肽U-23、和/或神经介肽U-25，且展示在脑、骨骼肌、小脑、睾丸、胼胝体、脊髓、黑质、髓质、丘脑、尾状核、脑桥、伏核、胎脑、胃、心、甲状腺、肺、胸腺、前列腺、和/或气管中表达。

11. 包含权利要求9或10的表达体系的宿主细胞。

12. 权利要求11的宿主细胞，其是酵母细胞。

13. 权利要求11的宿主细胞，其是动物细胞。

14. 由权利要求11-13的细胞衍生的IGS4受体膜制剂。

15. 用于生成IGS4多肽的方法，包括在足以生成所述多肽的条件下培养权利要求11-13的宿主，并由培养物回收该多肽。

16. 用于产生生成IGS4多肽的细胞的方法，包括用权利要求9或10的表达体系转化或转染宿主细胞，使得宿主细胞在适当的培养条件下生成IGS4多肽。

17. 包含在其全长上与SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的氨基酸序列，或者与由荷兰Baarn的真菌菌种保藏中心保藏物编号CBS102221或保藏物编号CBS102222所含DNA插入片段编码的多肽至少80%同一的氨基酸序列的IGS4多肽。

18. 权利要求17的多肽，其包含SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的氨基酸序列，或者由荷兰Baarn的真菌菌种保藏中心保藏物编号CBS102221或保藏物编号CBS102222所含DNA插入片段编码的氨基酸序列。

19. 包含神经介肽受体蛋白、优选哺乳动物神经介肽受体蛋白的氨基酸序列的分离IGS4多肽，所述蛋白质展示以高亲和力结合神经介肽U，优选神经介肽U-8、神经介肽U-23、和/或神经介肽U-25。

20. 权利要求19的分离IGS4多肽，其包含神经介肽受体蛋白的氨基酸序列，所述蛋白质展示在脑、骨骼肌、小脑、睾丸、胼胝体、脊髓、黑质、髓质、丘脑、尾状核、脑桥、伏核、胎脑、胃、心、甲状腺、肺、胸腺、前列腺、和/或气管中表达。

21. 包含神经介肽受体蛋白、优选哺乳动物神经介肽受体蛋白的氨基酸序列的分离IGS4多肽，所述蛋白质展示以高亲和力结合神经介肽U，优选神经介肽U-8、神经介肽U-23、和/或神经介肽U-25，所述蛋白质展示在脑、骨骼肌、小脑、睾丸、胼胝体、脊髓、黑质、髓质、丘脑、尾状核、脑桥、伏核、胎脑、胃、心、甲状腺、肺、胸腺、前列腺、和/或气管中表达，且所述氨基酸序列选自权利要求17-18中定义的氨基酸序列。

22. 对权利要求17-21的IGS4多肽具有免疫特异性的抗体。

23. 用于治疗需要增强权利要求17-21的IGS4多肽的活性或表达的对象的方法，包括：

(a) 对该对象施用治疗有效量的所述受体的激动剂；和/或

(b) 对该对象提供包含在其全长上与编码SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的IGS4多肽，或者由荷兰Baarn的真菌菌种保藏中心保藏物编号CBS102221或保藏物编号CBS102222所含DNA插入片段编码的多肽的核苷酸序列具有至少80%同一性的核苷酸序列；或者以某种形式与上述核苷酸序列之一互补从而在体内引起所述受体活性生成的核苷酸序列的分离多核苷酸；和/或

(c) 对该对象提供包含编码包含IGS4神经介肽受体蛋白、优选哺乳动物IGS4神经介肽受体蛋白的核苷酸序列的分离多核苷酸，所述蛋白质展示以高亲和力结合神经介肽U，优选神经介肽U-8、神经介肽U-23、和/或神经介肽U-25。

24. 用于治疗需要抑制权利要求17-21的IGS4多肽活性或表达的对象的方法，包括：

(a) 对该对象施用治疗有效量的所述受体的拮抗剂；和/或

(b) 对该对象施用抑制编码所述受体的核苷酸序列表达的核酸分子；和/或

(c) 对该对象施用治疗有效量的与所述受体竞争配体的多肽。

25. 用于诊断对象体内与权利要求17-21的IGS4多肽的表达或活性有关的疾病或对这种疾病的易感性的方法，包括：

(a) 测定所述对象基因组中编码所述IGS4多肽的核苷酸序列中是否存在突变；和/或

(b) 在由所述对象衍生的样品中分析所述IGS4多肽表达的存在与否或其表达量。

26. 用于鉴定权利要求17-21的IGS4多肽的激动剂的方法，包括：

(a) 使生成IGS4多肽的细胞接触待测化合物；并

(b) 测定所测化合物是否引起由IGS4多肽的激活产生的信号。

27. 由权利要求26的方法鉴定的激动剂。

28. 用于鉴定IGS4神经介肽受体蛋白、优选哺乳动物IGS4神经介肽受体蛋白的激动剂的方法，所述蛋白质展示以高亲和力结合神经介肽U，优选神经介肽U-8、神经介肽U-23、和/或神经介肽U-25，包括：

(a)使生成IGS4神经介肽受体蛋白的细胞接触待测化合物；并
(b)测定所测化合物是否引起由IGS4神经介肽受体蛋白的激活产生的信号。

29. 权利要求28所述用于鉴定IGS4神经介肽受体蛋白的激动剂的方法，其中所述激动剂对神经系统包括中枢神经系统（CNS）和周围神经系统（PNS）、胃肠系统、心血管系统、骨骼肌、和/或甲状腺的紊乱有效，和/或对肺病、免疫学疾病、和泌尿生殖系统紊乱也有效。

30. 由权利要求28或29的方法鉴定的激动剂，优选对神经系统包括中枢神经系统（CNS）和周围神经系统（PNS）、胃肠系统、心血管系统、骨骼肌、和/或甲状腺的紊乱有效，和/或对肺病、免疫学疾病、和泌尿生殖系统紊乱也有效的激动剂。

31. 用于鉴定权利要求17-21的IGS4多肽的拮抗剂的方法，包括：
(a)使生成IGS4多肽的细胞接触激动剂；并
(b)测定在存在候选化合物时由所述激动剂产生的信号是否减小。

32. 由权利要求31的方法鉴定的拮抗剂。

33. 用于鉴定IGS4神经介肽受体蛋白、优选哺乳动物IGS4神经介肽受体蛋白的拮抗剂的方法，所述蛋白质展示以高亲和力结合神经介肽U，优选神经介肽U-8、神经介肽U-23、和/或神经介肽U-25，包括：

(a)使生成IGS4神经介肽受体蛋白的细胞接触激动剂；并
(b)测定在存在候选化合物时由所述激动剂产生的信号是否减小。

34. 权利要求33中用于鉴定IGS4神经介肽受体蛋白的拮抗剂的方法，其中所述拮抗剂对神经系统包括中枢神经系统（CNS）和周围神经系统（PNS）、胃肠系统、心血管系统、骨骼肌、和/或甲状腺的紊乱有效，和/或对肺病、免疫学疾病、和泌尿生殖系统紊乱也有效。

35. 由权利要求33或34的方法鉴定的拮抗剂，优选对神经系统包括中枢神经系统（CNS）和周围神经系统（PNS）、胃肠系统、心血管系统、骨骼肌、和/或甲状腺的紊乱有效，和/或对肺病、免疫学疾病、

和泌尿生殖系统紊乱也有效的拮抗剂。

36. 由权利要求16的方法产生的重组宿主细胞或其表达IGS4多肽的膜。

37. 产生遗传修饰的非人动物的方法，包括：

(a) 将核酸分子的编码部分与在所述动物中能够驱动高水平基因表达或者能够在该基因通常不表达的细胞类型中进行表达的调控序列进行连接，其中所述编码部分基本上由编码具有SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的氨基酸序列或者由荷兰Baarn的真菌菌种保藏中心保藏物编号CBS102221或保藏物编号CBS102222所含DNA插入片段编码的氨基酸序列的蛋白质的核酸序列或者所述序列之一的生物活性部分组成；或者

(b) 分离并改造核酸分子的编码部分，所述编码部分主要由编码具有SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的氨基酸序列或者由荷兰Baarn的真菌菌种保藏中心保藏物编号CBS102221或保藏物编号CBS102222所含DNA插入片段编码的氨基酸序列的蛋白质的核酸序列或者所述序列之一的生物活性部分组成，并将所述序列重新导入所述动物的基因组，使得编码具有SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的氨基酸序列或者由荷兰Baarn的真菌菌种保藏中心保藏物编号CBS102221或保藏物编号CBS102222所含DNA插入片段编码的氨基酸序列的蛋白质的内源等位基因完全或部分失活。

38. 用于测定某种物质是否是IGS4受体的潜在配体的方法，包括：

(a) 在存在和不存在所述物质的情况中，使表达权利要求17-21之一或者SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8之一的受体的细胞，或者使包含权利要求17-21之一或者SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8之一的所述受体之一的受体膜制剂接触经标记的神经介肽U；并

(b) 测量神经介肽U与所述IGS4的结合。

39. 权利要求17-21任一项的多肽，其特征为所述多肽可结合神

经介肽U, 优选神经介肽U-8、神经介肽U-23、和/或神经介肽U-25, 且显示亲和力为至少大约 $\log EC_{50} = -6$ 。

40. 权利要求17-21任一项的多肽, 其特征为所述多肽可结合神经介肽U, 优选神经介肽U-8、神经介肽U-23、和/或神经介肽U-25, 且显示亲和力为至少大约 $\log EC_{50} = -9$ 。

新的人G蛋白偶联受体

描述

本发明涉及已经鉴定的新多核苷酸、由它们编码的多肽、以及这些多核苷酸和多肽的用途及其生产。更具体的说,本发明的多核苷酸和多肽涉及G蛋白偶联受体(G-protein coupled receptor, GPCR),下文称为IGS4。IGS4存在两种多态性形式,下文称为IGS4A和IGS4B。本发明还涉及抑制或激活这些多核苷酸和多肽的作用,包含所述多核苷酸的载体、包含这种载体的宿主细胞、和IGS4基因过度表达、错误表达、表达不足、或受遏制(敲除动物)的转基因动物。本发明还涉及用于筛选能够作为所述G蛋白偶联受体IGS4的激动剂或拮抗剂的化合物的方法,以及IGS4的相关配体。

发明背景

已经充分证明,许多具有医学重要性的生物学过程是由参与涉及G蛋白和/或第二信使(如cAMP)的信号转导途径的蛋白质介导的(Lefkowitz, 自然(Nature) 51: 353 - 354, 1991)。在本文中,这些蛋白质称为参与G蛋白途径的蛋白质。这些蛋白质的范例包括GPC受体,诸如肾上腺素能试剂和多巴胺的受体(B. K. Kobilka等人,美国国家科学院进展(Proc. Natl. Acad. Sci. USA) 84: 46 - 50, 1987; B. K. Kobilka等人,科学(Science) 238: 650 - 656, 1987; J. R. Bunzow等人,自然(Nature) 336: 783 - 787, 1988); G蛋白自身; 效应蛋白,如磷脂酶C、腺苷酸环化酶、和磷酸二酯酶; 和促动蛋白(actuator protein),如蛋白激酶A和蛋白激酶C(M. I. Simon等人,科学(Science) 252: 802 - 808, 1991)。

例如,在信号转导的一种形式中,在激素结合GPCR后,受体与异源三聚体的G蛋白相互作用,并诱导GDP由鸟嘌呤核苷酸结合位点解离

下来。在鸟嘌呤核苷酸的正常细胞浓度，GTP立即填充该位点。GTP与G蛋白 α 亚基的结合引起G蛋白与受体的解离且G蛋白解离成 α 与 $\beta\gamma$ 亚基。然后，携带GTP的形式结合激活后的腺苷酸环化酶。由G蛋白自身（ α 亚基具有内在的GTP酶活性）催化GTP水解变成GDP，使G蛋白恢复其基础的无活性形式。 α 亚基的GTP酶活性本质上是控制开/关转换的内在时钟。 α 亚基结合GDP的形式对 $\beta\gamma$ 具有高亲和力，随后 α GDP与 $\beta\gamma$ 的再次结合使系统恢复基础状态。如此，G蛋白起到双重作用：既作为中间物，将信号由受体传递至效应物（在这个范例中是腺苷酸环化酶）；又作为时钟，控制信号的持续时间。

已经鉴定，G蛋白偶联受体的膜结合型超家族具有7个推定的跨膜结构域。认为这些结构域是由胞外或胞质环连在一起的跨膜 α -螺旋。G蛋白偶联受体包括广泛的生物学活性受体，诸如激素、病毒、生长因子、和神经受体。

G蛋白偶联受体家族包括多巴胺受体，它们结合用于治疗CNS紊乱的神经安定药。该家族成员的其它范例包括但不限于降钙素、肾上腺素能、神经肽Y、促生长素抑制素、神经降压素、神经激肽、辣椒素、VIP、CGRP、CRF、CCK、缓激肽、促生长激素神经肽（galanin）、促胃动素、nociceptin、内皮缩血管肽、cAMP、腺苷、毒蝇碱、乙酰胆碱、5-羟色胺、组胺、凝血酶、激肽、促卵泡激素、视蛋白、内皮分化基因-1、视红紫质、添味剂（odorant）、和巨细胞病毒的受体。

大多数G蛋白偶联受体在前两个胞外环中都具有单个保守半胱氨酸残基，认为它们形成的二硫键稳定了功能蛋白的结构。7个跨膜区命名为TM1、TM2、TM3、TM4、TM5、TM6、和TM7。连接TM5与TM6的胞质环可能是G蛋白结合结构域的主要成分。

大多数G蛋白偶联受体在第3个胞质环和/或羧基末端包含潜在的磷酸化位点。对于几种G蛋白偶联受体，诸如 β -肾上腺素受体，由蛋白激酶A和/或特异受体激酶进行的磷酸化介导受体的脱敏。

最近发现，某些GPCR，如降钙素受体样受体，可能与称为受体活性修饰蛋白（RAMP）的小型单个穿膜蛋白相互作用。GPCR与某种RAMP

的这种相互作用决定了哪种天然配体对GPCR-RAMP联合具有相关亲和力并调控复合物的功能信号活性(L. M. McLathie等人, 自然(Nature) 393: 333 - 339, 1998)。

对于有些受体, 认为G蛋白偶联受体的配体结合位点包含由G蛋白偶联受体的几个跨膜结构域形成的亲水槽, 该槽由G蛋白偶联受体的疏水残基围绕。推测G蛋白偶联受体的每个跨膜螺旋的亲水侧面向内侧, 并形成有极性的配体结合位点。在几种G蛋白偶联受体中, TM3涉及配体结合位点, 诸如TM3天冬氨酸残基。TM5的丝氨酸、TM6天冬酰胺、和TM6或TM7的苯丙氨酸或酪氨酸也涉及配体结合。

G蛋白偶联受体可在细胞内通过异源三聚体的G蛋白偶联多种胞内酶、离子通道、和转运蛋白(参阅Johnson等人, 内分泌回顾(Endoc. Rev.) 10: 317 - 331, 1989)。不同的G蛋白 α 亚基优先刺激特定的效应物, 从而调控细胞中的多种生物学功能。已经鉴定, G蛋白偶联受体胞质残基的磷酸化是调控G蛋白偶联某些G蛋白偶联受体的重要机制。在哺乳动物宿主体内的许多部位发现了G蛋白偶联受体。

在已知药物中, 超过半数是由受体产生的, 主要是GPCR类(Drews, 自然生物技术(Nature Biotechnology) 14: 1516, 1996)。这说明这些受体作为治疗靶具有经证明、经检验的历史记录。显然, 需要鉴定和表征可用于预防、改善、或校正功能紊乱或疾病的其它受体, 包括但不限于PNS紊乱、精神病、和CNS紊乱, 包括精神分裂症, 发作性阵发性焦虑(episodic paroxysmal anxiety, EPA)症诸如强迫症(OCD)、创伤后精神紧张性障碍(PTSD)、恐怖症、和恐慌, 重型抑郁症, 双相性精神障碍, 帕金森症, 广泛性焦虑症, 孤独症, 谵妄, 多发性硬化症, 阿尔茨海默症/痴呆, 和其它神经变性疾病, 严重精神发育迟缓, 运动障碍, 亨廷顿病, 图雷特综合症, 抽搐, 震颤, 张力障碍, 痉挛, 厌食症, 贪食症, 中风, 上瘾/依赖/渴望, 睡眠障碍, 癫痫, 偏头痛; 注意缺陷/多动症(ADHD); 心血管疾病, 包括心力衰竭, 心绞痛, 心律失常, 心肌梗死, 心脏肥大, 低血压, 高血压(如原发性高血压、肾性高血压、或肺动脉高血压), 血栓形成, 动脉硬

化, 脑血管痉挛, 蛛网膜下出血, 脑缺血, 脑梗死, 周围血管疾病, 雷诺病, 肾病(如肾衰竭); 异常脂血症(dyslipidemias); 肥胖症; 呕吐; 胃肠紊乱, 包括肠易激惹综合症(IBS), 炎症性肠病(IBD), 胃食管反流病(GERD), 活动能力障碍(motility disorder), 和胃排空延迟状况, 诸如手术后或糖尿病性胃轻瘫, 和糖尿病, 溃疡(如胃溃疡); 腹泻; 其它疾病, 包括骨质疏松; 炎症; 感染, 诸如细菌、真菌、原生动物、和病毒感染, 特别是由HIV-1或HIV-2引起的感染; 疼痛; 癌症; 由化疗诱导的损伤; 肿瘤侵入; 免疫紊乱; 尿潴留; 哮喘; 变态反应; 关节炎; 良性前列腺肥大; 内毒素休克; 脓毒病; 糖尿病并发症; 和妇科疾病。

发明概述

本发明的一个方面涉及IGS4多肽(包括IGS4A和IGS4B多肽多态型)、多核苷酸、和重组材料及其生产方法。本发明的另一个方面涉及这些IGS4多肽、多核苷酸、和重组材料的使用方法。这些用途包括但不限于作为治疗靶, 用于治疗PNS紊乱、精神病、和CNS紊乱, 包括精神分裂症, 发作性阵发性焦虑(episodic paroxysmal anxiety, EPA) 症诸如强迫症(OCD)、创伤后精神紧张性障碍(PTSD)、恐怖症、和恐慌, 重型抑郁症, 双相性精神障碍, 帕金森症, 广泛性焦虑症, 孤独症, 谵妄, 多发性硬化症, 阿尔茨海默症/痴呆, 和其它神经变性疾病, 严重精神发育迟缓, 运动障碍, 亨廷顿病, 图雷特综合症, 抽搐, 震颤, 张力障碍, 痉挛, 厌食症, 贪食症, 中风, 上瘾/依赖/渴望, 睡眠障碍, 癫痫, 偏头痛; 注意缺陷/多动症(ADHD); 心血管疾病, 包括心力衰竭, 心绞痛, 心律失常, 心肌梗死, 心脏肥大, 低血压, 高血压(如原发性高血压、肾性高血压、或肺动脉高血压), 血栓形成, 动脉硬化, 脑血管痉挛, 蛛网膜下出血, 脑缺血, 脑梗死, 周围血管疾病, 雷诺病, 肾病(如肾衰竭); 异常脂血症(dyslipidemias); 肥胖症; 呕吐; 胃肠紊乱, 包括肠易激惹综合症(IBS), 炎症性肠病(IBD), 胃食管反流病(GERD), 活动能力障碍(motility disorder),

和胃排空延迟状况，诸如手术后或糖尿病性胃轻瘫，和糖尿病，溃疡（如胃溃疡）；腹泻；其它疾病，包括骨质疏松；炎症；感染，诸如细菌、真菌、原生动物、和病毒感染，特别是由HIV-1或HIV-2引起的感染；疼痛；癌症；由化疗诱导的损伤；肿瘤侵入；免疫紊乱；尿潴留；哮喘；变态反应；关节炎；良性前列腺肥大；内毒素休克；脓毒病；糖尿病并发症；和妇科疾病，等等。本发明的优选用途涉及神经系统（包括中枢神经系统（CNS）和周围神经系统（PNS））、胃肠系统、心血管系统、骨骼肌、和/或甲状腺的紊乱，和/或肺病、免疫学疾病、和泌尿生殖系统紊乱。

本发明的另一个方面涉及使用本发明提供的材料来鉴定激动剂和拮抗剂的方法，以及使用经鉴定的化合物来治疗与IGS4不平衡有关的状况。本发明的另一个方面涉及用于检测与IGS4活性或水平异常有关的疾病的诊断实验。本发明的另一个方面涉及由IGS4表达或活性异常所引起紊乱的动物模型系统。参照本发明鉴定的优选激动剂或拮抗剂适用于治疗神经系统（包括中枢神经系统（CNS）和周围神经系统（PNS））、胃肠系统、心血管系统、骨骼肌、和/或甲状腺的紊乱，和/或肺病、免疫学疾病、和泌尿生殖系统紊乱。

本发明还涉及本发明IGS4多肽的同类配体的鉴定。发现称为神经介肽U的神经肽结合所述IGS4多肽有高亲和力。

表1: SEQ ID NO: 1和SEQ ID NO: 3的IGS4A-DNA

```

5' -
GGCTCAGCTTGAAACAGAGCCTCGTACCAGGGGAGGCTCAGGCCTTGGATTTTAATGTCA
GGGATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAACTAGAAGATCCATTC
CAGAAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCCTCTGCGGACCTCGGCGCAGC
CACTTCTTCCCTCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCAATTGGC
AATGTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTAC
TACCTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGAATGCCCCTGGAG
GTCTATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCCGTGGGCTGCTACTTCAAG
ACGGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTG
GAGCGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGG
GCCCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCAACACC
AGCATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCCGGCC
ACCTGTACGGTCATCAAGCCCATGTGGATCTACAATTTTCATCATCCAGGTCACCTCCTTC
CTATTCTACCTCCTCCCCATGACTGTTCATCAGTGTCTCTACTACCTCATGGCACTCAGA
CTAAAGAAAGACAAATCTCTTGAGGCAGATGAAGGGAATGCAAATATTCAAAGACCCTGC
AGAAAAATCAGTCAACAAGATGCTGTTTGTCTTGGTCTTAGTGTGTTGCTATCTGTTGGGCC
CCGTTCCACATTGACCGACTCTTCTTCAGCTTTGTGGAGGAGTGGAGTGAATCCCTGGCT
GCTGTGTTCAACCTCGTCCATGTGGTGTGTCAGGTGTCTTCTTCTACCTGAGCTCAGCTGTC
AACCCCATTTATCTATAACCTACTGTCTCGCCGCTTCCAGGCAGCATTCAGAAATGTGATC
TCTTCTTTCCACAAAACAGTGGCACTCCCAGCATGACCCACAGTTGCCACCTGCCAGCGG
AACATCTTCCCTGACAGAAATGCCACTTTGTGGAGCTGACCGAAGATATAGGTCCCCAATTC
CCATGTCAGTCATCCATGCACAACCTCTCACCTCCCAACAGCCCTCTCTAGTGAACAGATG
TCAAGAACAAACTATCAAAGCTTCCACTTTAAACAAAACCTGAATTTCTTTCAGAGCTGACT
CTCCTCTATGCCTCAAAACTTCAGAGAGGAACATCCCATAAATGTATGCCTTCTCATATGA
TATTAGAGAGGTAGAATGGCTCTTACAACCTCATGTACCCATTGCTAGTTTTTTTTTTTAA
ATAAACGTGAAAACCTGAGAGTTAGATCTGGTTTTCAAAACCCAAAGACTGCCTGATTTTTAG
TTATCTTTCCACTATCCTAACTGCCTCATGCCCTTCACTAGTTTCATGCCAAGAACGTGA
CTGGAAAGGCATGGCACCTATACCTTGATTAATTTCCATTAATGGAAATGGTTCGTCCCTG
AGTCATCTACGTTCCGAGTCAGGCTGTCACTCCTACTA - 3'

```

表2: SEQ ID NO: 5和SEQ ID NO: 7的IGS4B-DNA

5' -

GGCTCAGCTTGAAACAGAGCCTCGTACCAGGGGAGGCTCAGGCCTTGGATTTTAATGTCA
GGGATGGAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAACTAGAAGATCCATTC
CAGAAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGC
CACTTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTGTGGTGGGGTCAATTGGC
AATGTCTTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCACCAACTAC
TACCTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCTGCTCCTTGGAAATGCCCTGGAG
GTCTATGAGATGTGGCGCAACTACCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAG
ACGGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTG
GAGCGCTACGTGGCCATCCTACACCCGTTCCGCGCAAACCTGCAGAGCACCCGGCGCCGG
GCCCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCAACACC
AGCATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCCGGCC
ACCTGTACGGTCATCAAGCCCATGTGGATCTACAATTTTCATCATCCAGGTCACCTCCTTC
CTATTCTACCTCCTCCCCATGACTGTCATCAGTGTCTCTACTACCTCATGGCACTCAGA
CTAAAGAAAGACAAATCTCTTGAGGCAGATGAAGGAATGCAAATATTCAAAGACCCTGC
AGAAAATCAGTCAACAAGATGCTGTTTGTCTTGGTCTTAGTGTGGTCTATCTGTTGGGCC
CCGTTCCACATGACCGACTCTTCTTCAGCTTTGTGGAGGAGTGGACTGAATCCCTGGCT
GCTGTGTTCAACCTCGTCCATGTGGTGTGAGGTGTCTTATTCTACCTGAGCTCAGCTGTC
AACCCATTATCTATAACCTACTGTCTCGCCGCTTCCAGGCAGCATTCCAGAATGTGATC
TCTTCTTTCCACAAACAGTGGCACTCCAGCATGACCCACAGTTGCCACCTGCCAGCGG
AACATCTTCTGACAGAATGCCACTTTGTGGAGCTGACCGAAGATATAGGTCCCCAATTC
CTATGTCAGTCATCCGTGCACAACTCTCACCTCCCAACAGCCCTCTCTAGTGAACAGATG
TCAAGAACAAACTATCAAAGCTTCCACTTTAACAAAACCTGAATTCTTTCAGAGCTGACT
CTCCTCTATGCCTCAAACTTCAGAGAGGAACATCCATAATGTATGCCTTCTCATATGA
AATTAGAGAGGTAGAATGGCTCTTACAACCTCATGTACCCATTGCTAGTFTTTTTTTTAA
ATAAACGTGAAAACCTGAGAGTTAGATCTGGTTTCAAACCCAAGACTGCCTGATTTTTAG
TTATCTTTCCACTATCCTAACTGCCTCATGCCCCCTTCACTAGTTCATGCCAAGAACGTGA
CTGGAAAGGCATGGCACCTATACCTTGATTAATTTCCATTAATGGAAATGGTTCGTCTTG
AGTCATCTACGTTCCGAGTCAGGCTGTCACTCCTACTA - 3'

表3: SEQ ID NO: 9和SEQ ID NO: 11的IGS4A-64-DNA

```

5' -
GGCTCAGCTTGAAACAGAGCCTCGTACCAGGGGAGGCTCAGGCCTTGGATTTTAATGTCA
GGGATGGAAAACCTTCAGAAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTC
CAGAAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGC
CACTTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGC
AATGTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTAC
TACCTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCTCTGCTCCTTGAATGCCCTGGAG
GTCTATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCCGTGGGCTGCTACTTCAAG
ACGGCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTG
GAGCGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGG
GCCCTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCAACACC
AGCATCCATGGCATCAAGTTCCTACTTCCCCAATGGGTCCCTGGTCCCAGGTTTCGGCC
ACCTGTACGGTCATCAAGCCCATGTGGATCTACAATTTTCATCATCCAGGTCACCTCCTTC
CTATTCTACCTCCTCCCCATGACTGTCATCAGTGTCTCTACTACCTCATGGCACTCAGA
CTAAAGAAAGACAAATCTCTTGAGGCAGATGAAGGGAATGCAAATATTCAAAGACCCTGC
AGAAAATCAGTCAACAAGATGCTGTCTTTGTGGAGGAGTGGAGTGAATCCCTGGCTGCTG
TGTTCAACCTCGTCCATGTGGTGTGAGGTGTCTTCTTCTACCTGAGCTCAGCTGTCAACC
CCATTATCTATAACCTACTGTCTCGCCGCTTCCAGGCAGCATTCCAGAATGTGATCTCTT
CTTTCACAAAACAGTGGCACTCCCAGCATGACCCACAGTTGCCACCTGCCAGCGGAACA
TCTTCCTGACAGAATGCCACTTGTGGAGCTGACCGAAGATATAGGTCCCAATTCCCAT
GTCAGTCATCCATGCACAACTCTCACCTCCCAACAGCCCTCTCTAGTGAACAGATGTCAA
GAACAACTATCAAAGCTTCCACTTTAACAAAACCTGAATTCTTTCAGAGCTGACTCTCC
TCTATGCCTCAAACTTCAGAGAGGAACATCCCATAATGTATGCCTTCTCATATGATATT
AGAGAGGTAGAATGGCTCTTACAACCTCATGTACCCATTGCTAGTTTTTTTTTTTAATAA
ACGTGAAAACCTGAGAGTTAGATCTGGTTTCAAAACCCAAGACTGCCTGATTTTTAGTTAT
CTTTCACCTATCCTAACTGCCTCATGCCCTTCACTAGTTCATGCCAAGAACGTGACTGG
AAAGGCATGGCACCTATACCTTGATTAATTTCCATTAATGGAAATGGTTCGTCTGAGTC
ATCTACGTTCCGAGTCAGGCTGTCACTCCTACTA - 3'

```

表4: SEQ ID NO: 2和SEQ ID NO: 4的IGS4A-蛋白质
(不含括号内的3个氨基酸)

```
(MSG) MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGV
IGNVLVCLVILQHQAAMKTPTNYLFLSLAVSDLLVLLLGMPLLEVYEMWRNYPFLFGPVGCV
FKTALFETVCFASILSITTVSVERYVAILHHPFRAKLQSTRRRALRILGIVWGFSVLFSLP
NTSIHGKIFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMA
LRLKKDKSLEADEGNANIQRPCRKSVNKMLFVLVLFVAICWAPFHIDRLFFSFVEEWSES
LAAVFNLVHVSVGVFFYLSSAVNPPIIYNLLSRRFQAAFQNVISSFHKQWHSQHDPQLPPA
QRNIFLTECHFVELTEDIGPQFPCQSSMHNHSLPTALSSEQMSRTNYQSFHFNKT
```

表5: SEQ ID NO: 6和SEQ ID NO: 8的IGS4B-蛋白质
(不含括号内的3个氨基酸)

```
(MSG) MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGV
IGNVLVCLVILQHQAAMKTPTNYLFLSLAVSDLLVLLLGMPLLEVYEMWRNYPFLFGPVGCV
FKTALFETVCFASILSITTVSVERYVAILHHPFRAKLQSTRRRALRILGIVWGFSVLFSLP
NTSIHGKIFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMA
LRLKKDKSLEADEGNANIQRPCRKSVNKMLFVLVLFVAICWAPFHIDRLFFSFVEEWTES
LAAVFNLVHVSVGVLFYLSSAVNPPIIYNLLSRRFQAAFQNVISSFHKQWHSQHDPQLPPA
QRNIFLTECHFVELTEDIGPQFLCQSSVHNSHLPTALSSEQMSRTNYQSFHFNKT
```

表6: SEQ ID NO: 10和SEQ ID NO: 12的IGS4A-64-蛋白质
(不含括号内的3个氨基酸)

```
(MSG) MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGV
IGNVLVCLVILQHQAAMKTPTNYLFLSLAVSDLLVLLLGMPLLEVYEMWRNYPFLFGPVGCV
FKTALFETVCFASILSITTVSVERYVAILHHPFRAKLQSTRRRALRILGIVWGFSVLFSLP
NTSIHGKIFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTSFLFYLLPMTVISVLYYLMA
LRLKKDKSLEADEGNANIQRPCRKSVNKMLSLWRSGVNPWLLCSTSSMWCQVSSST
```

发明详述

定义

提供下列定义以便于理解本文中频繁使用的某些术语。

“IGS4”指包含SEQ ID NO: 2或SEQ ID NO: 4 (IGS4A)、SEQ ID NO: 6或SEQ ID NO: 8 (IGS4B) 中列出的氨基酸序列的多肽或其变体等等。特别优选IGS4B多肽。

“受体活性”或“受体的生物学活性”指所述IGS4的代谢或生理学功能，包括相似的活性或改进的活性或非期望副作用降低的这些活性。还包括所述IGS4的抗原性和免疫原性。

“IGS4基因”指包含SEQ ID NO: 1、SEQ ID NO: 3、SEQ ID NO: 5、或SEQ ID NO: 7中列出的核苷酸序列的多核苷酸或其变体和/或其互补物。

在用于本文时，“抗体”包括多克隆和单克隆抗体；嵌合的、单链的、和人源化的抗体；以及Fab片段，包括Fab或其它免疫球蛋白表达文库的产物。

“分离的”指“人为的”由天然状态改变和/或与天然环境分开。因此，若自然界中存在的组合物或物质成为“分离的”，则它发生了变化或与其最初环境分开，或者二者兼之。例如，存活动物中天然存在的多核苷酸或多肽不是“分离的”，但是与天然状态的共存物质分开的相同多核苷酸或多肽是“分离的”，正如用于本文的该术语。

“多核苷酸”一般指任何多核糖核苷酸或多脱氧核糖核苷酸，可以是未修饰的RNA或DNA，或者是经修饰的RNA或DNA。“多核苷酸”包括但不限于单链和双链DNA、单链与双链区混合的DNA、单链和双链RNA、单链与双链区混合的RNA、包含DNA与RNA的杂合分子（其中DNA和RNA可以是单链，或更通常地是双链，或者是单链与双链区的混合）。另外，“多核苷酸”还可包括包含RNA或DNA或二者的三链区。术语“多核苷酸”还包括包含一种或多种经修饰碱基的DNA或RNA，以及为了稳定或其它原因而修饰了主链的DNA或RNA。“经修饰”碱基包括例如三

苯甲基化碱基和非常见碱基（诸如肌昔）。已经对DNA和RNA进行了多种修饰；因此，“多核苷酸”涵盖自然界中常常发现的多核苷酸的化学、酶促、或代谢修饰形式，以及病毒和细胞特征性的DNA和RNA化学形式。“多核苷酸”还涵盖相对较短的多核苷酸，通常称为寡核苷酸。

“多肽”指包含通过肽键或经修饰肽键（即peptide isosteres）彼此相连的2个或多个氨基酸的任何肽或蛋白质。“多肽”包括短链，常常称为肽、寡肽、或寡聚物；也包括长链，一般称为蛋白质；和/或其联合。多肽可包含20种基因编码氨基酸以外的氨基酸。“多肽”包括通过天然加工（诸如翻译后加工）或通过本领域众所周知的化学修饰技术进行了修饰的氨基酸序列。这些修饰详细描述于基础教科书、更详细的专论、以及庞大的研究文献中。修饰可发生于多肽中的任何位置，包括肽主链、氨基酸侧链、和氨基或羧基末端。应当理解，相同类型的修饰可以相同或不同程度存在于指定多肽中的几个位点。同样，指定多肽可包含许多类型的修饰。作为遍在蛋白化作用的结果，多肽可以是分支的，也可以是具有或不具有分支的环状。环状的、分支的、和分支环状的多肽可来自翻译后天然加工，或者由合成方法产生。修饰包括乙酰化、酰化、ADP-核糖基化、酰胺化、黄素的共价吸附、血红素部分的共价吸附、核苷酸或核苷酸衍生物的共价吸附、脂类或脂类衍生物的共价吸附、磷脂酰肌醇的共价吸附、交联、环化、二硫键的形成、脱甲基、共价交联的形成、胱氨酸的形成、焦谷氨酸的形成、甲酰化、 γ -羧化、糖基化、GPI锚的形成、羟化、碘化、甲基化、肉豆蔻酸化、氧化、蛋白水解加工、磷酸化、异戊二烯化、外消旋、硒化、硫化、由tRNA介导的向蛋白质添加氨基酸（诸如精氨酰化）、和遍在蛋白化作用。参阅，例如《蛋白质 - 结构和分子特性》（PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES），第2版，T.E.Creighton, W.H.Freeman and Company, 纽约, 1993; F.Wold, “翻译后蛋白质修饰：观点和展望”（Posttranslational Protein Modifications: Perspectives and Prospects），《蛋白质的翻译后共价修饰》（POSTTRANSLATIONAL COVALENT MODIFICATION OF

PROTEINS)第1-12页, B. C. Johnson编, Academic出版社, 纽约, 1983; Seifter等人, “蛋白质修饰和非蛋白质辅助因子的分析”(Analysis for protein modifications and nonprotein cofactors), 酶学方法(Meth. Enzymol.) 182: 626-646, 1990; 和Rattan等人, “蛋白质合成: 翻译后修饰和衰老”(Protein Synthesis: Posttranslational Modifications and Aging), 纽约科学院年鉴(Ann. NY Acad. Sci.) 663: 48-62, 1992.

在用于本文时, “变体”指分别与参考多核苷酸或多肽不同、但保留了本质特性(诸如本质的生物学、结构、调控、或生化特性)的多核苷酸或多肽。典型的多核苷酸变体与另一种参考多核苷酸具有不同的核苷酸序列。变体核苷酸序列中的变化有可能改变或不改变参考多核苷酸编码的多肽的氨基酸序列。正如下文所述, 核苷酸变化可在参考多核苷酸编码的多肽中导致氨基酸替代、添加、删除、融合、和截短。典型的多肽变体与另一种参考多肽具有不同的氨基酸序列。一般而言, 差异是有限的, 因此参考多肽与变体的序列在整体上密切相关, 在许多区域相同。变体与参考多肽的氨基酸序列可能因一处或多处替代、添加、和删除及其任意组合而不同。替代或插入的氨基酸残基可以是或不是由遗传密码子编码的。多核苷酸或多肽的变体可以是天然存在的, 诸如等位基因变体, 或者非天然存在的。可通过诱变技术或直接合成来产生多核苷酸和多肽的非天然存在变体。

“同一性”是核苷酸序列或氨基酸序列一致性的度量。一般而言, 排列序列以获得最高水平匹配。“同一性”本质上具有本领域认可的含义, 可使用发表的技术进行计算。参阅, 例如《计算分子生物学》(COMPUTATIONAL MOLECULAR BIOLOGY), A. M. Lesk编, 牛津大学出版社, 纽约, 1988; 《生物计算: 信息学和基因组计划》(BIOCOMPUTING: INFORMATICS AND GENOME PROJECTS), D. W. Smith编, Academic出版社, 纽约, 1993; 《序列信息的计算机分析》(COMPUTER ANALYSIS OF SEQUENCE DATA), 第1部分, A. M. Griffin和H. G. Griffin编, Humana出版社, 新泽西, 1994; 《分子生物学中的序列分析》(SEQUENCE

ANALYSIS IN MOLECULAR BIOLOGY), G. von Heinje, Academic出版社, 1987; 和《序列分析引物》(SEQUENCE ANALYSIS PRIMER), M. Gribskov和J. Devereux编, M Stockton出版社, 纽约, 1991。有许多方法可用于测量两种多核苷酸或多肽序列之间的同一性, 术语“同一性”对于熟练技术人员是众所周知的(H. Carillo和D. Lipton, SIAM应用数学杂志(SIAM J. Applied Math.) 48: 1073, 1988)。常用于测定两种序列之间的同一性或相似性的方法包括但不限于《大型计算机指南》中公开的方法, J. Bishop编, Academic出版社, 圣地亚哥, 1994; 和H. Carillo和D. Lipton, SIAM应用数学杂志(SIAM J. Applied Math.) 48: 1073, 1988。用于测定同一性和相似性的方法已编译成计算机程序。用于测定两种序列之间的同一性和相似性的优选计算机程序方法包括但不限于GCG程序软件包(J. Devereux等人, 核酸研究(Nucleic Acids Research) 12(1): 387, 1984)、BLASTP、BLASTN、FASTA(S. F. Atschul等人, 分子生物学杂志(J. Molec. Biol.) 215: 403, 1990)。词语“同源性”可替代“同一性”。

作为例示, 例如, 具有与参考核苷酸序列SEQ ID NO: 1具有至少95%“同一性”的核苷酸序列的多核苷酸意指, 多核苷酸的核苷酸序列与参考序列大致相同, 只是在参考核苷酸序列SEQ ID NO: 1的每100个核苷酸中多核苷酸序列可包含多达5个核苷酸差异。换言之, 为了获得具有与参考核苷酸序列至少95%同一的核苷酸序列的多核苷酸, 可在参考序列中删除或用其它核苷酸替代多达5%的核苷酸, 或者可在参考序列中插入多达参考序列核苷酸总数5%的任何数目核苷酸, 或者多达参考序列核苷酸总数5%的任何数目核苷酸可以是删除、插入、或替代的组合。参考序列的这些突变可发生于参考核苷酸序列的5'或3'末端位置或者这些末端位置之间的任何位置, 它们或是以单个核苷酸散布于参考序列中, 或是以一个或多个连续组散布于参考序列中。

相似的, 例如, 具有与参考氨基酸序列SEQ ID NO: 2具有至少95%“同一性”的氨基酸序列的多肽意指, 多肽的氨基酸序列与参考序列大致相同, 只是在参考氨基酸序列SEQ ID NO: 2的每100个氨基酸中

多肽序列可包含多达5个氨基酸改变。换言之，为了获得具有与参考氨基酸序列至少95%同一的氨基酸序列的多肽，可在参考序列中删除或用其它氨基酸替代多达5%的氨基酸，或者可在参考序列中插入多达参考序列氨基酸总数5%的任何氨基酸。参考序列的这些改变可发生于参考氨基酸序列的氨基或羧基末端位置或者这些末端位置之间的任何位置，它们或是以单个残基散布于参考序列中，或是以一个或多个连续组散布于参考序列中。

本发明的多肽

在一个方面，本发明涉及IGS4多肽（或IGS4蛋白质）。IGS4多肽包括SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、和SEQ ID NO: 8的多肽，以及具有由1999年9月24日保藏于荷兰Baarn的真菌菌种保藏中心（Centraalbureau voor Schimmelcultures）的保藏物编号CBS102221或保藏物编号CBS102222所含DNA插入片段编码的氨基酸序列的多肽；包含SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的氨基酸序列的多肽，以及包含由荷兰Baarn的真菌菌种保藏中心的保藏物编号CBS102221或保藏物编号CBS102222所含DNA插入片段编码的氨基酸序列的多肽；包含在其全长上与SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的氨基酸序列，和/或与由荷兰Baarn的真菌菌种保藏中心保藏物编号CBS102221或保藏物编号CBS102222所含DNA插入片段编码的氨基酸序列具有至少80%同一性的氨基酸序列的多肽，仍然更优选与所述氨基酸序列至少90%同一性，甚至仍然更优选至少95%同一性。此外，高度优选有至少97%，特别是至少99%同一性。IGS4多肽还包括在其全长上与具有SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的氨基酸序列的多肽或者与具有由荷兰Baarn的真菌菌种保藏中心保藏物编号CBS102221或保藏物编号CBS102222所含DNA插入片段编码的氨基酸序列的多肽具有至少80%同一性的氨基酸序列的多肽，仍然更优选与SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8有至少90%同一性，

甚至仍然更优选至少95%同一性。此外，高度优选有至少97%，特别是至少99%。优选的IGS4多肽展示受体的至少一种生物学活性。

在本发明的另一个实施方案中，IGS4多肽可以是较大蛋白质（诸如融合蛋白）的一部分。包含额外的氨基酸序列常常是有利的，这些额外氨基酸序列包括分泌序列或前导序列、原序列、诸如多个组氨酸残基等有助于纯化的序列、诸如抗原性肽标签等有助于检测的序列（诸如血细胞凝集素HA标签）、或在重组生产过程中有助于稳定的额外序列。

本发明还包括IGS4多肽的片段。片段指具有与上述IGS4多肽氨基酸序列的部分而非全部相同的氨基酸序列的多肽。与IGS4多肽一样，片段可以是“独立的”，或者是包含于较大多肽中而形成一部分或区域，最优选作为单个连续区域。本发明的多肽片段的代表性范例包括例如IGS4多肽氨基酸编号大约第1-20位、第21-40位、第41-60位、第61-80位、第81-100位、和第101位-末端的片段。在本文中，“大约”包括在一端或在两端比具体叙述的范围大或小几个、5、4、3、2、或1个氨基酸。

优选片段包括例如具有IGS4多肽的氨基酸序列、但删除了包含氨基末端的一系列连续残基或包含羧基末端的一系列连续残基或删除了两段连续残基（一段包含氨基末端，另一段包含羧基末端）的截短型多肽。参照本发明的截短型多肽的范例是分别由多核苷酸SEQ ID NO: 9和SEQ ID NO: 11编码的SEQ ID NO: 10和SEQ ID NO: 12的多肽。同样优选由结构或功能属性表征的片段，诸如包含 α -螺旋和 α -螺旋形成区、 β -折叠和 β -折叠形成区、转角和转角形成区、卷曲和卷曲形成区、亲水区、疏水区、 α 两亲区、 β 两亲区、柔性区、表面形成区、底物结合区、和高抗原指数区的片段。其它优选的片段是生物学活性片段。生物学活性片段是那些介导受体活性的片段，包括那些具有相似活性或改进活性，或者非期望活性降低的片段。还包括那些在动物（尤其是人）中具有抗原性或免疫原性的片段。

因此，本发明的多肽包括具有与SEQ ID NO: 2、SEQ ID NO: 4、

SEQ ID NO: 6、或SEQ ID NO: 8的氨基酸序列和/或与具有由荷兰Baarn的真菌菌种保藏中心的保藏物编号CBS102221或保藏物编号CBS102222所含DNA插入片段编码的氨基酸序列的多肽至少80%同一的氨基酸序列的多肽，或者与相应片段具有至少80%同一性的其片段。优选的是，所有这些多肽片段保留了受体的生物学活性，包括抗原性。确定序列和片段的变体也构成本发明的一部分。优选变体是那些因保守氨基酸替代（即用具有相似特征的残基进行替代）而与参照物不同的变体。典型的这类替代是Ala、Val、Leu、与Ile之间；Ser与Thr之间；酸性残基Asp与Glu之间；Asn与Gln之间；和碱性残基Lys与Arg之间；或芳香族残基Phe与Tyr之间的替代。特别优选的是其中替代、删除、或添加及其任意组合几个、5-10、1-5、或1-2个氨基酸的变体。

关于本发明的多肽，还发现它们显示对结合神经介肽U有高亲和力，特别是神经介肽U8（8个氨基酸的寡肽）、神经介肽U23（23个氨基酸的寡肽）、和/或神经介肽U25（25个氨基酸的寡肽）。在本发明的内容中，术语“高亲和力”理解为描述显示log EC₅₀值至少低于-6.00（大约660nM），优选log EC₅₀值低于-7.00（大约55nM），更优选log EC₅₀值低于-9.00（大约500pM-1.2nM），最优选log EC₅₀值低于-10.00（大约50-100pM）的配体结合。

在文献中，将神经肽神经介肽U的两种形式，即神经介肽U8和神经介肽U25，描述成最初由猪脊髓分离，可刺激子宫并具高血压性的肽（Minamino等人，生物化学和生物物理学研究通讯(Biochem. Biophys. Res. Commun.) 130: 1078-1085, 1985）。对于神经介肽U23（23个氨基酸的寡肽），参阅例如Okimura等人，肽化学(Pept. Chem.) 32: 321-324, 1995, vol. date 1994; Salmon等人，生物化学杂志(J. Biol. Chem.) 275(7): 4549-4554, 2000。随后由许多物种分离了神经介肽U(NMU)，如大鼠(NMU23)、人(NMU25)、青蛙(NMU25)、狗(NMU8和NMU25)、兔(NMU25)、和鸡(NMU25)。由此，Domin等人描述了使用特异放射性免疫实验对大鼠、猪、豚鼠、和人组织提取物中的神经介肽U免疫反应性进行鉴定(生物化学和生物物理学研究通

讯 (Biochem. Biophys. Res. Commun.) 140: 1127 - 1134, 1986)。Conlon等人 (神经化学 (Neurochem.) 51: 988 - 991, 1988) 建立了来自大鼠回肠的神经介肽U23的一级结构。Minamino等人 (生物化学和生物物理学研究通讯 (Biochem. Biophys. Res. Commun.) 156: 355 - 360, 1988) 主要使用用于猪神经介肽U8的免疫亲和层析和放射性免疫实验由小肠分离了大鼠神经介肽U, 通过微量序列分析测定了大鼠神经介肽U的氨基酸序列, 并通过合成确认了结构。虽然猪神经介肽U的C末端七肽酰胺结构在大鼠神经介肽U中是完全保守的, 但是该肽其余部分与猪神经介肽U25相比显示9个氨基酸的取代和2个氨基酸的删除。Domin等人 (生物化学杂志 (J. Biol. Chem.) 264: 20881 - 20885, 1989) 还测定了青蛙 *Rana temporaria* 中神经介肽U的分布、一级结构、和相对生物学活性, 发现整个序列是显示与猪和大鼠神经介肽U有显著序列相似性的25肽。在进一步的研究中, Domin等人 (调节肽 (Regul. Pept.) 41: 1 - 8, 1992) 由鸡纯化了神经介肽U的鸟类同系物。微量序列分析将肽鉴定为长度为25个氨基酸残基, 并且鸡神经介肽U在其具有生物活性的C端区域显示与猪肽有显著序列相似性。O'Harte等人 (肽 (Peptides) 12: 11 - 15, 1991) 描述了狗神经介肽U25的分离、结构鉴定、和药理学活性。此外, 对于兔神经介肽U25, 发现它缺乏翻译后加工位点的保守性 (Kage等人, 调节肽 (Regul. Pept.) 33: 191 - 198, 1991); 因此, 在兔神经介肽U中, 在猪和狗神经介肽U25中发现的Arg16-Arg17双碱性残基加工位点被Arg-Gly取代, 但是这个潜在的单碱性加工位点不被肠中的切割酶利用。

在所研究的物种中, 发现位于肽C末端的5个氨基酸几乎完全保守, 说明该区具有主要的重要性。因此, 哺乳动物神经介肽共享生物学活性似乎必需的C端序列 “Phe-Leu-Phe-Arg-Pro-Arg-Asn-酰胺”。NMU分布于胃肠道和中枢神经系统 (CNS) 二者。在大鼠中, 在回肠中发现神经介肽 (NMU) 浓度最高, 然后是垂体、下丘脑、脊髓、甲状腺、和泌尿生殖道。免疫组织化学研究显示, 只在神经纤维 (主要是肠肌和粘膜下神经丛) 中和在除了胃以外的所有区域的粘膜中发现了肠中

的NMU免疫反应性，而在内分泌细胞中没有发现NMU免疫反应性。在大鼠脑中，在广泛分布于除了小脑以外的全脑的纤维中发现有NMU免疫反应性。已经分离了编码NMU前体的人和大鼠基因。二者在C端编码NMU，中部编码其它潜在肽产物（Lo等人，分子内分泌学杂志（J. Mol. Endocrinol.）6: 1538 - 1544, 1992; Austin等人，分子内分泌学杂志（J. Mol. Endocrinol.）14: 157 - 169, 1995）。在大鼠子宫中鉴定了高亲和力NMU结合，并显示对GTP--S敏感（Nandha等人，内分泌学（Endocrinology）133: 482 - 486, 1993），说明NMU的受体应该是G蛋白偶联受体。然而，NMU的生理学作用在很大程度上仍然是未知的。神经介肽U可引起平滑肌的有力收缩、升高动脉血压、改变肠离子转运、和在低剂量刺激肾上腺皮质的功能和生长。NMU还显示可减小肠上动脉和门静脉中的血流，同时略为加大胰组织中的血流。

此外，根据国际专利申请WO 90/01330所述，神经介肽U8和U25适用于治疗胃肠道紊乱，如可用于在胃肠出血和餐后低血压的治疗中选择性降低流向胃肠道的血流。

已经鉴定了本发明的IGS4多肽是应答神经介肽U或与其足够相似的配体的G蛋白偶联受体。因此，应答神经介肽U的IGS4受体，特别是IGS4B受体，将极大有助于对神经介肽U和与其足够相似的配体的生理学机制的理解，以及对相关疾病的理解。

表5-8显示了本发明多肽的组织分布和表达水平，由此熟练技术人员可估计表达的定位和相关性。例如，关于本发明多肽的组织分布，通过如MTE（多种组织表达）分析、Northern印迹分析、和定量RT-PCR表达分析发现了本发明的IGS4多肽具体在如脑、骨骼肌、小脑、胸腺、髓质、甲状腺、气管、丘脑、黑质、胼胝体、尾状核、脑桥、伏核（nucleus accumbens）、胎脑、和胃中以中等水平表达（分别相对于在MTE印迹中作为100%的睾丸中或在定量RT-PCR中作为100%的脊髓中的表达）；而且如在心、肺、和前列腺中以相关水平表达（若通过定量RT-PCR可检测到）。例如，若表达水平为至今发现的睾丸或脊髓中最高表达的表达值（设为100%）的至少20%，则认为它们是中等的。例如，若至

少通过定量RT-PCR分析可检测到表达水平,则认为表达水平是相关的。应当领会,为任何器官指出的表达水平是构成器官的特定组织和细胞类型中表达水平的平均值。因此,若发现关于器官的表达水平只是相关的,这并不排除特定区域内(如器官的特定组织和/或细胞类型)具有局部的中等甚至高度表达。

这些结果说明IGS4多肽优选作用于神经系统(包括中枢神经系统(CNS)和周围神经系统(PNS))、胃肠系统、心血管系统、骨骼肌、和/或甲状腺,和/或肺病、免疫学疾病、和泌尿生殖系统紊乱。

因此,在另一个实施方案中,本发明涉及包含神经介肽受体蛋白的氨基酸序列的分离IGS4多肽,优选哺乳动物神经介肽受体蛋白,所述蛋白质展示以高亲和力结合神经介肽U,优选神经介肽U-8、神经介肽U-23、和/或神经介肽U-25。具体而言,包含神经介肽受体蛋白的氨基酸序列的分离IGS4多肽是展示在脑、骨骼肌、小脑、睾丸、胼胝体、脊髓、黑质、髓质、丘脑、尾状核、脑桥、伏核、胎脑、胃、心、甲状腺、肺、胸腺、前列腺、和/或气管中表达(至少可通过Northern和/或MTE和/或定量RT-PCR分析检测到)的蛋白质。在该实施方案的变体中,本发明涉及包含神经介肽受体蛋白、优选哺乳动物神经介肽受体蛋白的氨基酸序列的IGS4多肽,所述蛋白质展示以高亲和力结合神经介肽U,优选神经介肽U-8、神经介肽U-23、和/或神经介肽U-25,所述蛋白质展示在脑、骨骼肌、小脑、睾丸、胼胝体、脊髓、黑质、髓质、丘脑、尾状核、脑桥、伏核、胎脑、胃、心、甲状腺、肺、胸腺、前列腺、和/或气管中表达(至少可通过Northern和/或MTE和/或定量RT-PCR分析检测到),而且所述氨基酸序列选自上文定义的氨基酸序列。

可以任何合适方式制备本发明的IGS4多肽。这些多肽包括分离的天然存在的多肽、重组产生的多肽、合成产生的多肽、或由这些方法联合产生的多肽。用于制备这些多肽的方法在本领域是众所周知的。

本发明的多核苷酸

本发明的另一个方面涉及IGS4多核苷酸。IGS4多核苷酸包括编码IGS4多肽（包括IGS4A和IGS4B）及其片段的分离多核苷酸，以及与其密切相关的多核苷酸。更具体的说，本发明的IGS4多核苷酸包括包含分别编码SEQ ID NO: 2或SEQ ID NO: 4的IGS4A多肽或者SEQ ID NO: 6或SEQ ID NO: 8的IGS4B多肽的SEQ ID NO: 1、SEQ ID NO: 3、SEQ ID NO: 5、或SEQ ID NO: 7中所含核苷酸序列的多核苷酸，具有SEQ ID NO: 1、SEQ ID NO: 3、SEQ ID NO: 5、或SEQ ID NO: 7的特定序列的多核苷酸，和基本上相当于荷兰Baarn的真菌菌种保藏中心的保藏物编号CBS102221或保藏物编号CBS102222所含DNA插入片段的多核苷酸。

IGS4多核苷酸还包括包含在其全长上与编码SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的IGS4多肽的核苷酸序列具有至少80%同一性的核苷酸序列的多核苷酸，包含在其全长上与SEQ ID NO: 1、SEQ ID NO: 3、SEQ ID NO: 5、或SEQ ID NO: 7的核苷酸序列至少80%同一的核苷酸序列的多核苷酸，和基本上相当于荷兰Baarn的真菌菌种保藏中心的保藏物编号CBS102221或保藏物编号CBS102222所含DNA插入片段的多核苷酸。

在这方面，特别优选具有至少90%同一性的多核苷酸，尤其优选具有至少95%同一性的多核苷酸。此外，高度优选具有至少97%同一性的多核苷酸，最优选具有至少98-99%同一性的多核苷酸，尤其是具有至少99%同一性的多核苷酸。IGS4多核苷酸还包括与SEQ ID NO: 1、SEQ ID NO: 3、SEQ ID NO: 5、或SEQ ID NO: 7中所含核苷酸序列或者与荷兰Baarn的真菌菌种保藏中心的保藏物编号CBS102221或保藏物编号CBS102222所含DNA插入片段具有足够同一性，从而在可用于扩增或者作为探针或标记的条件下发生杂交的核苷酸序列。本发明还提供了与这些IGS4多核苷酸互补的多核苷酸。

如公开数据库的BLSAT搜索结果所示，本发明的IGS4与G蛋白偶联受体家族的其它蛋白质在结构上相关。表1的氨基酸序列（SEQ ID NO: 2）在其大部分长度上（316个氨基酸残基）与人孤儿G蛋白偶联受体（编号043664，Tan等人，基因组学（Genomics）52（2）：223-229，1998）

具有大约46%同一性（使用BLAST, S.F. Altschul等人, 核酸研究 (Nucleic Acids Res.) 25: 3389 - 3402, 1997）。与大鼠神经降压素1受体（编号P20789, K. Tanaka等人, 神经元 (Neuron) 4: 847 - 854, 1990）存在27%同源性（第61 - 349位氨基酸残基）。表1的核苷酸序列（SEQ ID NO: 1）在第120 - 864位核苷酸残基上与孤儿G蛋白偶联受体63%同一（编号AF044600, 对应于蛋白质序列043664）。此外, 在第94 - 1137位残基上与人生长激素促分泌素受体存在53%同一性（A. D. Howard等人, 科学 (Science): 273: 974 - 977, 1996）。因此, 预计本发明的IGS4多肽和多核苷酸与其同源多肽和多核苷酸具有相似生物学功能/特性, 而且它们的效用对本领域技术人员是显而易见的。

可由天然来源诸如基因组DNA获得本发明的多核苷酸。具体而言, 可设计编码特定GPCR基因亚族内保守区的简并PCR引物。使用简并引物对基因组DNA或cDNA进行的PCR扩增反应将引起考虑中的基因家族几个成员（已知的和新的）的扩增（当使用基因组模板时, 简并引物必须位于相同外显子内）（Libert等人, 科学 (Science) 244: 569 - 572, 1989）。可使用众所周知的商业性技术（如F. M. Ausubel等人, 《分子生物学现行方案》（Current Protocols in Molecular Biology）, 2000）来合成本发明的多核苷酸。

编码SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的IGS4多肽的核苷酸序列可能分别与SEQ ID NO: 1（第55 - 1299位核苷酸）或SEQ ID NO: 3（第64 - 1299位核苷酸）或SEQ ID NO: 5（第55 - 1299位核苷酸）或SEQ ID NO: 7（第64 - 1299位核苷酸）中所含多肽编码序列相同, 或者它可能是不同的核苷酸序列, 而是因遗传密码冗余性（简并性）, 分别与SEQ ID NO: 1、SEQ ID NO: 3、SEQ ID NO: 5、或SEQ ID NO: 7中所含多肽编码序列相比显示改变, 但是也编码SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的多肽。

在另一个实施方案中, 本发明涉及编码IGS4神经介肽受体蛋白优选编码哺乳动物神经介肽受体蛋白、的分离核苷酸序列, 所述蛋白质

展示以高亲和力结合神经介肽U，优选神经介肽U-8、神经介肽U-23、和/或神经介肽U-25。具体而言，分离核苷酸序列编码展示在脑、骨骼肌、小脑、睾丸、胼胝体、脊髓、黑质、髓质、丘脑、尾状核、脑桥、伏核、胎脑、胃、心、甲状腺、肺、胸腺、前列腺、和/或气管中表达（至少可通过Northern和/或MTE和/或定量RT-PCR分析检测到）的IGS4神经介肽受体蛋白。在该实施方案的变体中，本发明涉及编码IGS4神经介肽受体蛋白优选编码哺乳动物神经介肽受体蛋白、的分离核苷酸序列，所述蛋白质展示以高亲和力结合神经介肽U，优选神经介肽U-8、神经介肽U-23、和/或神经介肽U-25，所述蛋白质展示在脑、骨骼肌、小脑、睾丸、胼胝体、脊髓、黑质、髓质、丘脑、尾状核、脑桥、伏核、胎脑、胃、心、甲状腺、肺、胸腺、前列腺、和/或气管中表达（至少可通过Northern和/或MTE和/或定量RT-PCR分析检测到），而且所述核苷酸序列选自上文定义的核苷酸序列。

在将本发明的多核苷酸用于IGS4多肽的重组生产时，多核苷酸自身可包含成熟多肽或其片段的编码序列；成熟多肽或其片段的编码序列与其它编码序列（诸如编码前导序列或分泌序列、蛋白原序列、前蛋白序列、前蛋白原序列、或其它融合肽部分）处于读码框中。例如，可编码有助于纯化融合多肽的标记序列。在本发明这方面的某些优选实施方案中，标记序列是六组氨酸肽，如pQE载体（Qiagen公司）中提供的，且描述于Gentz等人，美国国家科学院进展（Proc. Natl. Acad. Sci. USA）86：821 - 824，1989），或HA标签。多核苷酸还可包含5'和3'非编码序列，诸如转录但不翻译的序列、剪接和聚腺苷酸化信号、核糖体结合位点、和稳定mRNA的序列。

另一个优选的实施方案是编码包含SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的IGS4多肽氨基酸序列但其中几个、5 - 10、1 - 5、1 - 3、1 - 2、或1个氨基酸残基被替代、删除、或添加、或其任意联合的IGS4变体的多核苷酸。

可使用本领域普遍知道的方法来改造本发明的多核苷酸，从而为了多种目的，包括但不限于修饰基因产物的克隆、加工、和/或表达，

而改变IGS4编码序列。通过随机片段化和基因片段与合成寡核苷酸的PCR重装配进行的DNA改组，可用于改造核苷酸序列。例如，由寡核苷酸介导的定点诱变可用于导入突变，从而产生氨基酸替代、产生新的限制性位点、改变修饰（如糖基化或磷酸化）模式、改变密码子偏爱性、产生剪接变体、等等。

本发明还涉及与本文上述序列发生杂交的多核苷酸。在这方面，本发明尤其涉及在严谨条件下与上述多核苷酸发生杂交的多核苷酸。在用于本文时，术语“严谨条件”指只有序列间存在至少80%（优选至少90%，更优选至少95%，甚至更优选至少97%，特别是至少99%）同一性时才发生杂交。

与SEQ ID NO: 1、SEQ ID NO: 3、SEQ ID NO: 5、或SEQ ID NO: 7中所含核苷酸序列或其片段同一或足够同一的本发明多核苷酸，可作为cDNA和基因组DNA的杂交探针，用于分离编码IGS4的全长cDNA和基因组克隆，和用于分离与IGS4基因具有高度序列相似性的其它基因（包括编码来自非人物种的同系物和直向进化同源物的基因）的cDNA和基因组克隆。本领域熟练技术人员完全了解这些杂交技术。通常，这些核苷酸序列与参考序列80%同一，优选90%同一，更优选95%同一。探针通常将包含至少5个核苷酸，优选至少8个核苷酸，更优选至少10个核苷酸，甚至更优选至少12个核苷酸，特别是至少15个核苷酸。最优选的是，这些探针将具有至少30个核苷酸，而且可以具有至少50个核苷酸。特别优选的探针的范围是30 - 50个核苷酸。

为了获得编码IGS4多肽（包括来自非人物种的同系物和直向进化同源物）的多核苷酸的一个实施方案包括下列步骤：在严谨杂交条件下用具有SEQ ID NO: 1、SEQ ID NO: 3、SEQ ID NO: 5、或SEQ ID NO: 7或其片段的经标记探针筛选合适文库，并分离包含所述多核苷酸序列的全长cDNA和基因组克隆。这些杂交技术对本领域技术人员是众所周知的。严谨杂交条件如上文定义的，或者是在含50% 甲酰胺、5x SSC（150mM NaCl、15mM 柠檬酸三钠）、50mM 磷酸钠（pH7.6）、5x Denhardt氏液、10% 硫酸葡聚糖（w/v）、和20 μg/ml 变性剪切鲑鱼精DNA的

溶液中于42℃保温过夜，随后用0.1x SSC于大约65℃清洗滤膜。

本发明的多核苷酸和多肽可作为研究试剂和材料用于开发人和动物疾病的治疗和诊断。

载体、宿主细胞、表达

本发明还涉及包含本发明多核苷酸的载体，和用本发明载体基因工程改造的宿主细胞，还涉及通过重组技术生产本发明的多肽。使用由本发明DNA构建物衍生的RNA，无细胞翻译系统也可用于生产这些蛋白质。

为了重组生产，可基因工程改造宿主细胞以掺入本发明多核苷酸的表达系统或其部分。可通过许多标准实验室指南（诸如Davis等人，《分子生物学基本方法》（BASIC METHODS IN MOLECULAR BIOLOGY），1986；和Sambrook等人，《分子克隆：实验室指南》（MOLECULAR CLONING: A LABORATORY MANUAL），第2版，冷泉港实验室出版社，冷泉港，纽约，1989）中描述的方法，诸如磷酸钙转染、由DEAE-葡聚糖介导的转染、转位（transvection）、显微注射、由阳离子脂类介导的转染、电穿孔、转导、刮擦装载（scrape loading）、轰击导入、或感染，将多核苷酸导入宿主细胞。

合适宿主的代表性范例包括细菌细胞，诸如链球菌、葡萄球菌、大肠杆菌、链霉菌、和枯草芽孢杆菌细胞；真菌细胞，诸如酵母细胞和曲霉细胞；昆虫细胞，诸如果蝇S2和Spodoptera Sf9细胞；动物细胞，诸如CHO、COS、HeLa、C127、3T3、BHK、HEK293、和Bowes黑素瘤细胞；和植物细胞。

可使用很多种表达系统。这些系统包括由染色体、游离体、和病毒衍生的系统，如由细菌质粒，细菌噬菌体，转座子，酵母游离体，插入元件，酵母染色体元件，病毒诸如杆状病毒、乳多空病毒（诸如SV40）、痘苗病毒、腺病毒、禽痘病毒、假狂犬病病毒、和逆转录病毒衍生的载体，以及由上述联合衍生的载体，诸如由质粒和噬菌体遗传元件衍生的载体，诸如粘粒和噬菌粒。表达系统可包含调控以及引

起表达的控制区。一般而言，可使用适用于在宿主中维持、增殖、或表达多核苷酸产生多肽的任何系统或载体。可通过多种众所周知的常规技术，诸如Sambrook等人《分子克隆：实验室指南》（见上文）中所述，将合适的核苷酸序列插入表达系统。

为了使翻译产生的蛋白质分泌进入内质网内腔、周质空间、或胞外环境，可将合适的分泌信号掺入目的多肽。这些信号可以是内源的，或者是外源的，即衍生自不同物种。

一般而言，若要表达IGS4多肽用于筛选实验，优选在细胞表面产生多肽。此时，可在用于筛选实验前收获细胞。在IGS4多肽的亲合力或功能活性受到受体活性修饰蛋白（RAMP）修饰的情况中，最有可能优选在细胞表面共表达相关RAMP，而且这常常是需要的。此时同样需要在用于筛选实验前收获表达IGS4多肽和相关RAMP的细胞。若IGS4多肽分泌进入培养基，则可回收培养基以回收并纯化多肽；若生成于胞内，则在回收多肽前首先必须裂解细胞。可通过本领域技术人员众所周知的方法来回收表达IGS4多肽的膜。一般而言，这些方法包括收获表达IGS4多肽的细胞，并通过诸如但不限于pottering的方法将细胞匀浆。可通过将悬浮液清洗一次或几次来回收膜。

可通过众所周知的方法由重组细胞培养物回收并纯化IGS4多肽，包括硫酸铵或乙醇沉淀、酸提取、阴离子或阳离子交换层析、磷酸纤维素层析、疏水相互作用层析、亲和层析、羟基磷灰石层析、和凝集素层析。最优选的是，采用高效液相层析来进行纯化。当多肽在分离和/或纯化过程中发生变性时，可采用众所周知用于蛋白质重新折叠的技术来重新产生有活性的构象。

诊断实验

本发明还涉及IGS4多核苷酸作为诊断剂的用途。与功能障碍有关的突变型IGS4基因的检测将提供诊断工具，从而增加或明确由IGS4的表达不足、过度表达、或表达变化引起的疾病或其易感性的诊断。同样，此时可能需要相关受体活性修饰蛋白的共表达，以获得预期质量

的诊断实验。可通过多种技术，在DNA水平上检测IGS4基因内携有突变的个体。

可由受试者的细胞，诸如由血液、尿液、唾液、活组织样品、或尸解材料，获得用于诊断的核酸。基因组DNA可直接用于检测，或者可在分析前通过PCR或其它扩增技术进行酶促扩增。RNA或cDNA也可以相似方式使用。可通过扩增产物与正常基因型相比的大小变化来检测删除和插入。可通过将扩增后的DNA与经标记IGS4核苷酸序列进行杂交来鉴定点突变。可通过RNA酶消化或熔解温度的差异来区分完美匹配的序列与错误匹配的双链体。还可通过DNA片段在含或不含变性剂的凝胶中的电泳迁移率的变化，或者通过直接的DNA测序来检测DNA序列差异(参阅如Myers等人，科学 (Science) 230: 1242, 1985)。还可通过核酸酶保护实验(诸如RNA酶和S1保护)或化学切割法来揭示特定位置的序列变化(参阅Cotton等人，美国国家科学院进展 (Proc. Natl. Acad. Sci. USA) 85: 4397 - 4401, 1985)。在另一个实施方案中，可构建包含IGS4核苷酸序列或其片段的寡核苷酸探针阵列来进行如遗传突变的高效筛选。阵列技术的方法是众所周知的，而且具有普遍适用性，可用于在分子遗传学中解决多种问题，包括基因表达、遗传连锁、和遗传可变性(参阅例如M. Chee等人，科学 (Science) 274: 610 - 613, 1996)。

诊断实验提供了通过上述方法检测IGS4基因的突变来诊断或确定对下列疾病易感性的方法：PNS紊乱、精神病、和CNS紊乱，包括精神分裂症，发作性阵发性焦虑 (episodic paroxysmal anxiety, EPA) 症诸如强迫症 (OCD)、创伤后精神紧张性障碍 (PTSD)、恐怖症、和恐慌，重型抑郁症，双相性精神障碍，帕金森症，广泛性焦虑症，孤独症，谵妄，多发性硬化症，阿尔茨海默症/痴呆，和其它神经变性疾病，严重精神发育迟缓，运动障碍，亨廷顿病，图雷特综合症，抽搐，震颤，张力障碍，痉挛，厌食症，贪食症，中风，上瘾/依赖/渴望，睡眠障碍，癫痫，偏头痛；注意缺陷/多动症 (ADHD)；心血管疾病，包括心力衰竭，心绞痛，心律失常，心肌梗死，心脏肥大，低血压，

高血压（如原发性高血压、肾性高血压、或肺动脉高血压），血栓形成，动脉硬化，脑血管痉挛，蛛网膜下出血，脑缺血，脑梗死，周围血管疾病，雷诺病，肾病（如肾衰竭）；异常脂血症（dyslipidemias）；肥胖症；呕吐；胃肠紊乱，包括肠易激惹综合症（IBS），炎症性肠病（IBD），胃食管反流病（GERD），活动能力障碍（motility disorder），和胃排空延迟状况，诸如手术后或糖尿病性胃轻瘫，和糖尿病，溃疡（如胃溃疡）；腹泻；其它疾病，包括骨质疏松；炎症；感染，诸如细菌、真菌、原生动物、和病毒感染，特别是由HIV-1或HIV-2引起的感染；疼痛；癌症；由化疗诱导的损伤；肿瘤侵入；免疫紊乱；尿潴留；哮喘；变态反应；关节炎；良性前列腺肥大；内毒素休克；脓毒症；糖尿病并发症；和妇科疾病。具体而言，本发明的诊断实验提供了用于诊断或确定对下列疾病易感性的方法：神经系统（包括中枢神经系统（CNS）和周围神经系统（PNS））、胃肠系统、心血管系统、骨骼肌、和/或甲状腺的紊乱，和/或肺病、免疫学疾病、和泌尿生殖系统紊乱。

另外，可通过包括由衍生自受试者的样品测定IGS4多肽或IGS4 mRNA水平的异常降低或升高的方法来诊断PNS紊乱、精神病、和CNS紊乱，包括精神分裂症，发作性阵发性焦虑（episodic paroxysmal anxiety, EPA）症诸如强迫症（OCD）、创伤后精神紧张性障碍（PTSD）、恐怖症、和恐慌，重型抑郁症，双相性精神障碍，帕金森症，广泛性焦虑症，孤独症，谵妄，多发性硬化症，阿尔茨海默症/痴呆，和其它神经变性疾病，严重精神发育迟缓，运动障碍，亨廷顿病，图雷特综合症，抽搐，震颤，张力障碍，痉挛，厌食症，贪食症，中风，上瘾/依赖/渴望，睡眠障碍，癫痫，偏头痛；注意缺陷/多动症（ADHD）；心血管疾病，包括心力衰竭，心绞痛，心律失常，心肌梗死，心脏肥大，低血压，高血压（如原发性高血压、肾性高血压、或肺动脉高血压），血栓形成，动脉硬化，脑血管痉挛，蛛网膜下出血，脑缺血，脑梗死，周围血管疾病，雷诺病，肾病（如肾衰竭）；异常脂血症（dyslipidemias）；肥胖症；呕吐；胃肠紊乱，包括肠易激惹综合症

(IBS), 炎症性肠病 (IBD), 胃食管反流病 (GERD), 活动能力障碍 (motility disorder), 和胃排空延迟状况, 诸如手术后或糖尿病性胃轻瘫, 和糖尿病, 溃疡 (如胃溃疡); 腹泻; 其它疾病, 包括骨质疏松; 炎症; 感染, 诸如细菌、真菌、原生动物、和病毒感染, 特别是由HIV-1或HIV-2引起的感染; 疼痛; 癌症; 由化疗诱导的损伤; 肿瘤侵入; 免疫紊乱; 尿潴留; 哮喘; 变态反应; 关节炎; 良性前列腺肥大; 内毒素休克; 脓毒病; 糖尿病并发症; 和妇科疾病。具体而言, 可通过包括由衍生自对象的样品测定IGS4多肽或IGS4 mRNA水平的异常降低或升高的方法来诊断神经系统 (包括中枢神经系统 (CNS) 和周围神经系统 (PNS))、胃肠系统、心血管系统、骨骼肌、和/或甲状腺的紊乱, 和/或肺病、免疫学疾病、和泌尿生殖系统紊乱。可使用任何本领域众所周知的用于多核苷酸定量的技术在RNA水平测量表达的降低或升高, 诸如PCR、RT-PCR、RNA酶保护、Northern印迹、和其它杂交方法。可用于测定衍生自宿主的样品中蛋白质 (诸如IGS4) 水平的实验技术对本领域技术人员是众所周知的。这些实验方法包括放射性免疫实验、竞争性结合实验、Western印迹分析、和ELISA实验。

另一方面, 本发明涉及用于诊断疾病或其易感性的诊断试剂盒, 特别是PNS紊乱、精神病、和CNS紊乱, 包括精神分裂症, 发作性阵发性焦虑 (episodic paroxysmal anxiety, EPA) 症诸如强迫症 (OCD)、创伤后精神紧张性障碍 (PTSD)、恐怖症、和恐慌, 重型抑郁症, 双相性精神障碍, 帕金森症, 广泛性焦虑症, 孤独症, 谵妄, 多发性硬化症, 阿尔茨海默症/痴呆, 和其它神经变性疾病, 严重精神发育迟缓, 运动障碍, 亨廷顿病, 图雷特综合症, 抽搐, 震颤, 张力障碍, 痉挛, 厌食症, 贪食症, 中风, 上瘾/依赖/渴望, 睡眠障碍, 癫痫, 偏头痛; 注意缺陷/多动症 (ADHD); 心血管疾病, 包括心力衰竭, 心绞痛, 心律失常, 心肌梗死, 心脏肥大, 低血压, 高血压 (如原发性高血压、肾性高血压、或肺动脉高血压), 血栓形成, 动脉硬化, 脑血管痉挛, 蛛网膜下出血, 脑缺血, 脑梗死, 周围血管疾病, 雷诺病, 肾病 (如肾衰竭); 异常脂血症 (dyslipidemias); 肥胖症; 呕吐; 胃肠紊乱,

包括肠易激惹综合症(IBS), 炎症性肠病(IBD), 胃食管反流病(GERD), 活动能力障碍 (motility disorder), 和胃排空延迟状况, 诸如手术后或糖尿病性胃轻瘫, 和糖尿病, 溃疡 (如胃溃疡); 腹泻; 其它疾病, 包括骨质疏松; 炎症; 感染, 诸如细菌、真菌、原生动物、和病毒感染, 特别是由HIV-1或HIV-2引起的感染; 疼痛; 癌症; 由化疗诱导的损伤; 肿瘤侵入; 免疫紊乱; 尿潴留; 哮喘; 变态反应; 关节炎; 良性前列腺肥大; 内毒素休克; 脓毒病; 糖尿病并发症; 和妇科疾病, 其包含:

(a) IGS4多核苷酸, 优选SEQ ID NO: 1、SEQ ID NO: 3、SEQ ID NO: 5、或SEQ ID NO: 7的核苷酸序列或其片段; 和/或

(b) 与(a)互补的核苷酸序列; 和/或

(c) IGS4多肽, 优选SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的多肽或其片段; 和/或

(d) 针对IGS4多肽的抗体, 优选针对SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的多肽的抗体; 和/或

(e) IGS4多肽的相关生物学或抗原性特性需要的RAMP多肽。

应当领会, 在任何这些试剂盒中, (a)、(b)、(c)、(d)、或(e)可包含实质性成份。优选的是, 本发明涉及用于诊断或确定下列疾病或其易感性的诊断试剂盒: 神经系统 (包括中枢神经系统 (CNS) 和周围神经系统 (PNS))、胃肠系统、心血管系统、骨骼肌、和/或甲状腺的疾病, 和/或肺病、免疫学疾病、和泌尿生殖系统紊乱。

染色体实验

本发明的核苷酸序列还可用于染色体鉴定。将序列特异靶向单个人染色体上的特定位置, 并能够与其发生杂交。参照本发明将相关序列对染色体进行作图是将那些序列与基因相关疾病联系起来的重要的第一步。一旦将序列定位于精确的染色体位置, 可使序列在染色体上的物理位置与遗传图谱资料相关联。这些资料可在例如V. McKusick的“人类孟德尔遗传” (可通过网络由Johns Hopkins大学Welch医学图

书馆获得)中找到。然后通过连锁分析(物理上相邻基因的共遗传)鉴定基因与已定位于相同染色体区域的疾病之间的关系。

还可测定受影响与不受影响的个体之间的cDNA或基因组序列差异。若在有些或所有受影响的个体中观察到突变,而在任何正常个体中没有观察到该突变,则它有可能是疾病的病因。

抗体

本发明的多肽或其片段或其类似物,或者表达它们(如果需要的话,还与相关RAMP一起表达)的细胞,还可作为免疫原用于产生针对IGS4多肽的免疫特异性抗体。术语“免疫特异性”指抗体对本发明多肽的亲合力显著高于对现有技术中其它相关多肽的亲合力。

使用常规方案,通过对动物(优选非人动物)施用多肽或者携带表位的片段、类似物、或细胞,可获得针对IGS4多肽产生的抗体。对于单克隆抗体的制备,可使用提供由连续细胞系培养物产生的抗体的任何技术。范例包括杂交瘤技术(G.Kohler和C.Milstein,自然(Nature)256:495-597,1975)、trioma技术、人B细胞杂交瘤技术(Kozbor等人,今日免疫学(Immunology Today)4:72,1983)、和EBV-杂交瘤技术(Cole等人,《单克隆抗体和癌症治疗》(MONOCLONAL ANTIBODIES AND CANCER THERAPY),第77-96页,Alan R. Liss公司,1985)。

上述抗体可用于分离或鉴定表达多肽的克隆,或用于通过亲和层析纯化多肽。

针对这些IGS4多肽或针对IGS4多肽-RAMP复合物的抗体还可用于治疗PNS紊乱、精神病、和CNS紊乱,包括精神分裂症,发作性阵发性焦虑(episodic paroxysmal anxiety,EPA)症诸如强迫症(OCD)、创伤后精神紧张性障碍(PTSD)、恐怖症、和恐慌,重型抑郁症,双相性精神障碍,帕金森症,广泛性焦虑症,孤独症,谵妄,多发性硬化症,阿尔茨海默症/痴呆,和其它神经变性疾病,严重精神发育迟缓,运动障碍,亨廷顿病,图雷特综合症,抽搐,震颤,张力障碍,痉挛,

厌食症，贪食症，中风，上瘾/依赖/渴望，睡眠障碍，癫痫，偏头痛；注意缺陷/多动症（ADHD）；心血管疾病，包括心力衰竭，心绞痛，心律失常，心肌梗死，心脏肥大，低血压，高血压（如原发性高血压、肾性高血压、或肺动脉高血压），血栓形成，动脉硬化，脑血管痉挛，蛛网膜下出血，脑缺血，脑梗死，周围血管疾病，雷诺病，肾病（如肾衰竭）；异常脂血症（dyslipidemias）；肥胖症；呕吐；胃肠紊乱，包括肠易激惹综合症（IBS），炎症性肠病（IBD），胃食管反流病（GERD），活动能力障碍（motility disorder），和胃排空延迟状况，诸如手术后或糖尿病性胃轻瘫，和糖尿病，溃疡（如胃溃疡）；腹泻；其它疾病，包括骨质疏松；炎症；感染，诸如细菌、真菌、原生动物、和病毒感染，特别是由HIV-1或HIV-2引起的感染；疼痛；癌症；由化疗诱导的损伤；肿瘤侵入；免疫紊乱；尿潴留；哮喘；变态反应；关节炎；良性前列腺肥大；内毒素休克；脓毒症；糖尿病并发症；和妇科疾病等等。优选的是，本发明抗体可用于治疗神经系统（包括中枢神经系统（CNS）和周围神经系统（PNS））、胃肠系统、心血管系统、骨骼肌、和/或甲状腺的紊乱，和/或肺病、免疫学疾病、和泌尿生殖系统紊乱。

动物

本发明的另一个方面涉及作为由IGS4异常表达或活性所引起紊乱的模型的非人动物系统。非人动物模型系统也可用于进一步鉴定IGS4基因的活性。这些系统可作为筛选策略的一部分，所述策略设计用于鉴定能够治疗下列基于IGS4的紊乱的化合物，诸如PNS紊乱、精神病、和CNS紊乱，包括精神分裂症，发作性阵发性焦虑（episodic paroxysmal anxiety, EPA）症诸如强迫症（OCD）、创伤后精神紧张性障碍（PTSD）、恐怖症、和恐慌，重型抑郁症，双相性精神障碍，帕金森症，广泛性焦虑症，孤独症，谵妄，多发性硬化症，阿尔茨海默症/痴呆，和其它神经变性疾病，严重精神发育迟缓，运动障碍，亨廷顿病，图雷特综合症，抽搐，震颤，张力障碍，痉挛，厌食症，贪

食症，中风，上瘾/依赖/渴望，睡眠障碍，癫痫，偏头痛；注意缺陷/多动症（ADHD）；心血管疾病，包括心力衰竭，心绞痛，心律失常，心肌梗死，心脏肥大，低血压，高血压（如原发性高血压、肾性高血压、或肺动脉高血压），血栓形成，动脉硬化，脑血管痉挛，蛛网膜下出血，脑缺血，脑梗死，周围血管疾病，雷诺病，肾病（如肾衰竭）；异常脂血症（dyslipidemias）；肥胖症；呕吐；胃肠紊乱，包括肠易激惹综合症（IBS），炎症性肠病（IBD），胃食管反流病（GERD），活动能力障碍（motility disorder），和胃排空延迟状况，诸如手术后或糖尿病性胃轻瘫，和糖尿病，溃疡（如胃溃疡）；腹泻；其它疾病，包括骨质疏松；炎症；感染，诸如细菌、真菌、原生动物、和病毒感染，特别是由HIV-1或HIV-2引起的感染；疼痛；癌症；由化疗诱导的损伤；肿瘤侵入；免疫紊乱；尿潴留；哮喘；变态反应；关节炎；良性前列腺肥大；内毒素休克；脓毒症；糖尿病并发症；和妇科疾病。具体而言，这些系统作为筛选策略的一部分，所述策略设计用于鉴定能够治疗基于IGS4的神经系统（包括中枢神经系统（CNS）和周围神经系统（PNS））、胃肠系统、心血管系统、骨骼肌、和/或甲状腺的紊乱，和/或肺病、免疫学疾病、和泌尿生殖系统紊乱的化合物。如此，动物模型可用于鉴定可有效治疗由IGS4表达或活性异常引起的紊乱的制药学化合物、疗法、和干预法。另外，这些动物模型可用于测定动物对象中的LD₅₀和ED₅₀。这些数据可用于测定IGS4紊乱潜在治疗的体内功效。

基于IGS4表达或活性异常、基于IGS4的紊乱的动物模型系统可包括非重组动物以及经重组改造的转基因动物。

IGS4疾病的动物模型可包括例如遗传模型。可利用例如IGS4序列（诸如上文所述），联合本领域熟练技术人员众所周知的用于产生转基因动物的技术，改造展示基于IGS4紊乱样症状的动物模型。例如，可将IGS4序列导入目的动物的基因组，并在其中过度表达或错误表达，或者，若存在内源IGS4序列，则它们可过度表达、错误表达、或者遭到破坏从而表达不足或灭活IGS4基因表达。

为了过度表达或错误表达IGS4基因序列，可将IGS4基因序列的编码部分与在目的动物类型中能够驱动高水平基因表达或能够在该基因通常不表达的细胞类型中进行表达的调控序列进行连接。这些调控序列对本领域熟练技术人员是众所周知的，而且它们的使用不需过多实验。

为使内源IGS4基因序列表达不足，可分离这种序列并进行改造，使得在重新导入目的动物的基因组后，内源IGS4基因的等位基因将失活或“敲除”。优选的是，通过基因打靶导入经改造的IGS4基因序列，使得在经改造IGS4基因序列整合到动物基因组中后，内源IGS4序列遭到破坏。这个部分的下文将讨论基因打靶。

任何物种的动物，包括但不限于小鼠、大鼠、兔、松鼠、豚鼠、猪、微型猪、山羊、和非人灵长类（如狒狒、猴、和黑猩猩），可用于产生IGS4相关紊乱的动物模型。

可使用任何本领域已知技术将IGS4转基因导入动物，以产生转基因动物的起始系（founder line）。这些技术包括但不限于原核显微注射（P. C. Hoppe和T. E. Wagner, 1989, 美国专利号4,873,191）；由逆转录病毒介导的基因转移进入种系（van der Putten等人，美国国家科学院进展（Proc. Natl. Acad. Sci. USA）82: 6148 - 6152, 1985）；胚胎干细胞中的基因打靶（Thompson等人，细胞（Cell）56: 313 - 321, 1989）；胚胎的电穿孔（Lo, 分子细胞生物学（Mol. Cell. Biol.）3: 1803 - 1814, 1983）；和由精子介导的基因转移（Lavitrano等人，细胞（Cell）57: 717 - 723, 1989）；等。这些技术的回顾参阅Gordon, “转基因动物”（Transgenic Animals），国际细胞学回顾（Intl. Rev. Cytol.）115: 171 - 229, 1989（完整收入本文作为参考）。

本发明提供了在其所有细胞中携带IGS4转基因的转基因动物，以及在有些而非所有细胞中携带转基因的动物，即嵌合动物（参阅例如Jakobovits, 现行生物学（Curr. Biol.）4: 761 - 763, 1994描述的技术）。转基因可以是以单个转基因整合，或者是串联体，如头-头串联或头-尾串联。还可通过例如Lasko等人的传授（M. Lasko等人，

美国国家科学院进展 (Proc. Natl. Acad. Sci. USA) 89: 6232 - 6236, 1992), 将转基因选择性导入特定细胞类型并在其中被激活。这种细胞类型特异激活需要的调控序列将取决于具体的目的细胞类型, 而且对于本领域技术人员是显而易见的。

当需要将IGS4转基因整合到内源IGS4基因的染色体位点中时, 优选基因打靶。简而言之, 当要使用这种技术时, 设计包含一些与目的内源IGS4基因同源的核苷酸序列 (如小鼠IGS4基因的核苷酸序列) 的载体, 用于通过与染色体序列的同源重组而整合并破坏内源IGS4基因或其等位基因的核苷酸序列的功能。还可通过例如Gu等人的传授 (H. Gu等人, 科学 (Science) 265: 103 - 106, 1994), 将转基因选择性导入特定细胞类型, 从而只在该细胞类型中灭活目的内源基因。这种细胞类型特异灭活需要的调控序列将取决于具体的目的细胞类型, 而且对于本领域技术人员是显而易见的。

一旦产生了转基因动物, 可利用标准技术测定重组IGS4基因和蛋白质的表达。可通过Southern印迹分析或PCR技术对动物组织分析是否发生了转基因的整合, 如此进行初步筛选。还可使用包括但不限于对由动物获得的组织样品的Northern印迹分析、原位杂交技术、和RT-PCR等技术, 评价转基因动物的组织中IGS4转基因的mRNA表达水平。还可使用针对目的靶基因转基因产物的特异性抗体在免疫细胞化学上评价表达靶基因的组织样品。然后可进一步评价在易于检测的水平表达IGS4基因mRNA或IGS4转基因肽 (使用针对靶基因产物表位的抗体在免疫细胞化学上检测) 的IGS4转基因动物, 以鉴定那些显示基于IGS4紊乱的特征性症状的动物。

一旦产生了IGS4转基因起始动物 (即那些在目的细胞或组织中表达IGS4蛋白质、且优选展示基于IGS4紊乱的症状的动物), 可将它们培育、近交、远交、或杂交以产生该具体动物的群体。这些培育策略的范例包括但不限于: 将含超过一个整合位点的起始动物远交以建立不同的系; 不同系的近交以产生因各IGS4转基因加成表达的结果而以更高水平表达目的IGS4转基因的复合型IGS4转基因系; 将杂合转基因

动物杂交以产生指定整合位点的纯合子动物，从而既提高了表达又消除了对通过DNA分析来筛选动物的可能需要；将不同纯合子系杂交以产生复合型杂合子或纯合子系；将动物与不同的近交遗传背景进行育种以检验修饰等位基因对IGS4转基因表达和IGS4样症状发展的影响。这样的一种方法是将IGS4转基因起始动物与野生型品系杂交以产生展示上述IGS4相关紊乱样症状的F1代。然后，若发现纯合子靶基因转基因动物是可生育的，则可将F1代近交以建立纯合子种系。

疫苗

本发明的另一个方面涉及用于在哺乳动物中诱导免疫学应答的方法，包括对哺乳动物施用（例如通过接种）IGS4多肽或其片段，如果需要则与RAMP多肽一起施用，施用量足够产生抗体和/或T细胞免疫应答以保护动物免于PNS紊乱、精神病、和CNS紊乱，包括精神分裂症，发作性阵发性焦虑（episodic paroxysmal anxiety, EPA）症诸如强迫症（OCD）、创伤后精神紧张性障碍（PTSD）、恐怖症、和恐慌，重型抑郁症，双相性精神障碍，帕金森症，广泛性焦虑症，孤独症，谵妄，多发性硬化症，阿尔茨海默症/痴呆，和其它神经变性疾病，严重精神发育迟缓，运动障碍，亨廷顿病，图雷特综合症，抽搐，震颤，张力障碍，痉挛，厌食症，贪食症，中风，上瘾/依赖/渴望，睡眠障碍，癫痫，偏头痛；注意缺陷/多动症（ADHD）；心血管疾病，包括心力衰竭，心绞痛，心律失常，心肌梗死，心脏肥大，低血压，高血压（如原发性高血压、肾性高血压、或肺动脉高血压），血栓形成，动脉硬化，脑血管痉挛，蛛网膜下出血，脑缺血，脑梗死，周围血管疾病，雷诺病，肾病（如肾衰竭）；异常脂血症（dyslipidemias）；肥胖症；呕吐；胃肠紊乱，包括肠易激惹综合症（IBS），炎症性肠病（IBD），胃食管反流病（GERD），活动能力障碍（motility disorder），和胃排空延迟状况，诸如手术后或糖尿病性胃轻瘫，和糖尿病，溃疡（如胃溃疡）；腹泻；其它疾病，包括骨质疏松；炎症；感染，诸如细菌、真菌、原生动物、和病毒感染，特别是由HIV-1或HIV-2引起的感染；

疼痛；癌症；由化疗诱导的损伤；肿瘤侵入；免疫紊乱；尿潴留；哮喘；变态反应；关节炎；良性前列腺肥大；内毒素休克；脓毒病；糖尿病并发症；和妇科疾病等等。本发明的还有一个方面涉及用于在哺乳动物中诱导免疫学应答的方法，包括借助在体内指导IGS4多核苷酸表达的载体投递IGS4多肽，从而诱导这种免疫学应答以产生抗体，保护所述动物免于疾病。本发明具体涉及用于在哺乳动物中诱导免疫学应答的方法，包括给哺乳动物接种IGS4多肽或其片段，如果需要还与RAMP多肽一起接种，接种量足够产生抗体和/或T细胞免疫应答以保护所述动物免于神经系统（包括中枢神经系统（CNS）和周围神经系统（PNS））、胃肠系统、心血管系统、骨骼肌、和/或甲状腺的紊乱，和/或肺病、免疫学疾病、和泌尿生殖系统紊乱。

本发明的还有一个方面涉及免疫学/疫苗制剂（组合物），在导入哺乳动物宿主后，将在该哺乳动物中诱导针对IGS4多肽的免疫学应答，其中组合物包含IGS4多肽或IGS4基因。这些免疫学/疫苗制剂（组合物）可以是治疗性的免疫学/疫苗制剂，或者是预防性的免疫学/疫苗制剂。疫苗制剂还可包含合适的载体。由于IGS4多肽可能在胃中降解，因此优选胃肠外施用（包括皮下、肌肉内、静脉内、皮内、等注射）。适用于胃肠外施用的制剂包括水性和非水性无菌注射液，可包含抗氧化剂、缓冲液、抑菌剂、和使得制剂与接受者血液等渗的溶质；以及水性和非水性无菌悬浮液，可包含悬浮剂或增稠剂。制剂可存在于单位剂量或多份剂量的容器中，例如密封的安瓿瓶或药瓶，可以冻干状态储存，只需要在临使用前加入无菌液体载体。疫苗制剂还可包含用于增强制剂免疫原性的佐剂系统，诸如水包油系统和本领域知道的其它系统。剂量将取决于疫苗的比活性，而且易于通过常规实验测定。

筛选实验

本发明的IGS4多肽可用于筛选结合受体并激活（激动剂）或抑制激活（拮抗剂）本发明受体多肽的化合物的方法。因此，本发明的多

肽还可用于评价例如细胞、非细胞制剂、化学文库、和天然产物混合物中小分子底物和配体的结合情况。这些底物和配体可以是天然底物和配体，或者是结构或功能模拟物。

IGS4多肽负责生物学功能，包括病理学。因此，希望找到刺激IGS4或可抑制IGS4功能的化合物和药物。一般而言，激动剂可用于下列状况的治疗和预防目的，诸如PNS紊乱、精神病、和CNS紊乱，包括精神分裂症，发作性阵发性焦虑（episodic paroxysmal anxiety, EPA）症诸如强迫症（OCD）、创伤后精神紧张性障碍（PTSD）、恐怖症、和恐慌，重型抑郁症，双相性精神障碍，帕金森症，广泛性焦虑症，孤独症，谵妄，多发性硬化症，阿尔茨海默症/痴呆，和其它神经变性疾病，严重精神发育迟缓，运动障碍，亨廷顿病，图雷特综合症，抽搐，震颤，张力障碍，痉挛，厌食症，贪食症，中风，上瘾/依赖/渴望，睡眠障碍，癫痫，偏头痛；注意缺陷/多动症（ADHD）；心血管疾病，包括心力衰竭，心绞痛，心律失常，心肌梗死，心脏肥大，低血压，高血压（如原发性高血压、肾性高血压、或肺动脉高血压），血栓形成，动脉硬化，脑血管痉挛，蛛网膜下出血，脑缺血，脑梗死，周围血管疾病，雷诺病，肾病（如肾衰竭）；异常脂血症（dyslipidemias）；肥胖症；呕吐；胃肠紊乱，包括肠易激惹综合症（IBS），炎症性肠病（IBD），胃食管反流病（GERD），活动能力障碍（motility disorder），和胃排空延迟状况，诸如手术后或糖尿病性胃轻瘫，和糖尿病，溃疡（如胃溃疡）；腹泻；其它疾病，包括骨质疏松；炎症；感染，诸如细菌、真菌、原生动物、和病毒感染，特别是由HIV-1或HIV-2引起的感染；疼痛；癌症；由化疗诱导的损伤；肿瘤侵入；免疫紊乱；尿潴留；哮喘；变态反应；关节炎；良性前列腺肥大；内毒素休克；脓毒症；糖尿病并发症；和妇科疾病。拮抗剂可用于多种状况的治疗和预防目的，诸如PNS紊乱、精神病、和CNS紊乱，包括精神分裂症，发作性阵发性焦虑（episodic paroxysmal anxiety, EPA）症诸如强迫症（OCD）、创伤后精神紧张性障碍（PTSD）、恐怖症、和恐慌，重型抑郁症，双相性精神障碍，帕金森症，广泛性焦虑症，孤独症，谵妄，

多发性硬化症, 阿尔茨海默症/痴呆, 和其它神经变性疾病, 严重精神发育迟缓, 运动障碍, 亨廷顿病, 图雷特综合症, 抽搐, 震颤, 张力障碍, 痉挛, 厌食症, 贪食症, 中风, 上瘾/依赖/渴望, 睡眠障碍, 癫痫, 偏头痛; 注意缺陷/多动症 (ADHD); 心血管疾病, 包括心力衰竭, 心绞痛, 心律失常, 心肌梗死, 心脏肥大, 低血压, 高血压 (如原发性高血压、肾性高血压、或肺动脉高血压), 血栓形成, 动脉硬化, 脑血管痉挛, 蛛网膜下出血, 脑缺血, 脑梗死, 周围血管疾病, 雷诺病, 肾病 (如肾衰竭); 异常脂血症 (dyslipidemias); 肥胖症; 呕吐; 胃肠紊乱, 包括肠易激惹综合症 (IBS), 炎症性肠病 (IBD), 胃食管反流病 (GERD), 活动能力障碍 (motility disorder), 和胃排空延迟状况, 诸如手术后或糖尿病性胃轻瘫, 和糖尿病, 溃疡 (如胃溃疡); 腹泻; 其它疾病, 包括骨质疏松; 炎症; 感染, 诸如细菌、真菌、原生动物、和病毒感染, 特别是由HIV-1或HIV-2引起的感染; 疼痛; 癌症; 由化疗诱导的损伤; 肿瘤侵入; 免疫紊乱; 尿潴留; 哮喘; 变态反应; 关节炎; 良性前列腺肥大; 内毒素休克; 脓毒症; 糖尿病并发症; 和妇科疾病。具体而言, 本发明可用于筛选结合受体并激活 (激动剂) 或抑制激活 (拮抗剂) IGS4神经介肽受体蛋白的化合物的方法, 优选哺乳动物IGS4神经介肽受体蛋白, 所述蛋白质展示以高亲和力结合神经介肽U, 优选神经介肽U-8、神经介肽U-23、和/或神经介肽U-25。这些筛选实验特别适用于筛选对神经系统 (包括中枢神经系统 (CNS) 和周围神经系统 (PNS))、胃肠系统、心血管系统、骨骼肌、和/或甲状腺的紊乱, 和/或肺病、免疫学疾病、和泌尿生殖系统紊乱有效的化合物。

一般而言, 这些筛选流程包括产生在其表面表达本发明受体多肽且如果需要则还在其表面共表达RAMP的适当细胞。这些细胞包括来自哺乳动物、酵母、果蝇、或大肠杆菌的细胞。然后使表达受体 (或细胞膜包含表达的受体) 的细胞接触待测化合物以观察结合情况, 或者对功能应答的刺激或抑制情况。

一种筛选技术包括在测量由受体激活引起的胞外pH、胞内pH、或

胞内钙变化的系统中使用表达本发明受体的细胞（例如经转染的CHO细胞）。在这种技术中，可使化合物接触表达本发明受体多肽的细胞。然后测量第二信使应答（如信号转导、pH变化、或钙水平变化），以确定潜在化合物是否激活或抑制受体。

另一种方法包括通过测定由受体介导的信号（诸如cAMP积累和/或腺苷酸环化酶活性）的调节来筛选受体抑制剂。这种方法包括用本发明受体转染真核细胞，从而在细胞表面表达该受体。然后在存在潜在拮抗剂的情况中使细胞暴露于本发明受体的激动剂。若潜在拮抗剂结合受体，并由此抑制受体结合，则由激动剂介导的信号将受到调节。

用于检测本发明受体的激动剂或拮抗剂的另一种方法是美国专利号5,482,835（收入本文作为参考）中描述的基于酵母的技术。

这些实验可仅仅测试候选化合物的结合情况，其中通过与候选化合物直接或间接相连的标记物，或者在涉及经标记竞争剂的竞争的实验中，检测对携带受体的细胞的粘附。此外，这些实验可使用适用于表面携带受体的细胞的检测系统来测试候选化合物是否导致由受体激活产生的信号。通常在存在已知激动剂的情况中鉴定激活的抑制剂，并在存在候选化合物的情况中观察激动剂对激活的影响。

由此，可筛选出显示与本发明IGS4受体发生配体结合的候选化合物。在本发明的内容中，“配体结合”理解为描述化合物对IGS4受体的亲和力显示 $\log EC_{50}$ 值至少低于-6.00（大约660nM），优选 $\log EC_{50}$ 值低于-7.00（大约55nM），更优选 $\log EC_{50}$ 值低于-9.00（大约500pM - 1.2nM），最优选 $\log EC_{50}$ 值低于-10.00（大约50 - 100pM）。

因此，在本发明的一个方面，涉及用于确定某种物质是否是IGS4受体的潜在配体的方法，包括：

(a) 在存在和不存在所述物质的情况中，使表达上文定义的IGS4神经介肽受体之一或者SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、和SEQ ID NO: 8的受体之一的细胞，或者使包含上文定义的IGS4神经介肽受体之一或者SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、和SEQ ID NO: 8的受体之一的受体膜制剂接触经标记的神经介肽U；并

(b) 测量神经介肽U与IGS4的结合。

此外，这些实验可仅仅包括下列步骤：混合候选化合物与含IGS4多肽的溶液以形成混合物，测量混合物中的IGS4活性，并将混合物的IGS4活性与标准进行比较。

IGS4 cDNA、蛋白质、和针对该蛋白质的抗体还可用于建立用于检测添加化合物对细胞中IGS4 mRNA和蛋白质生成的影响的实验。例如，可建立ELISA，通过本领域知道的标准方法，使用单克隆和多克隆抗体测量IGS4蛋白的分泌水平或细胞结合水平。这可用于发现对适当操作的细胞或组织抑制或增强IGS4生成的试剂（也分别称为拮抗剂或激动剂）。用于进行筛选实验的标准方法在本领域是众所周知的。

潜在IGS4拮抗剂的范例包括抗体或（在有些情况中）与IGS4的配体密切相关的寡核苷酸或蛋白质，如配体的片段，或结合受体但不引发应答从而可预防受体活性的小分子。

因此，在另一个方面，本发明涉及用于鉴定IGS4多肽的激动剂、拮抗剂、配体、受体、底物、酶、等；或者降低、升高、和/或增强IGS4多肽生成的化合物的筛选试剂盒，其包含：

(a) IGS4多肽，优选SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的多肽；

(b) 表达IGS4多肽，优选SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的多肽的重组细胞；

(c) 表达IGS4多肽的细胞膜，优选表达SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的IGS4多肽的细胞膜；或者

(d) 针对IGS4多肽的抗体，优选针对SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、或SEQ ID NO: 8的IGS4多肽的抗体。

应当领会，在任何这些试剂盒中，(a)、(b)、(c)、或(d)可包含实质性成份。

预防和治疗方法

本发明提供了用于治疗涉及IGS4活性量过度或不足的异常状况

的方法。

若IGS4的活性过度，则可使用几种方法。一种方法是与制药学可接受载体一起，以有效抑制激活的量，对对象施用上文所述的抑制剂化合物（拮抗剂），通过阻断配体与IGS4的结合，或者通过抑制与RAMP多肽或第二信号的相互作用，由此减轻异常状况。

另一种方法是施用能够与内源IGS4竞争性结合配体的IGS4多肽的可溶性形式。这种竞争剂的典型实施方案包括IGS4多肽的片段。

还有一种方法是使用表达阻断技术来抑制编码内源IGS4的基因的表达。已知这种技术包括反义序列的使用，或是内部生成或是分开施用。参阅，例如O'Connor，神经化学杂志（*J Neurochem*）56: 560, 1991；《寡脱氧核苷酸作为基因表达的反义抑制剂》（*Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression*），CRC出版社，Boca Raton，佛罗里达，美国，1988。或者，可使用与基因形成三链体的寡核苷酸。参阅，例如Lee等人，核酸研究（*Nucleic Acids Res*）6: 3073, 1979；Cooney等人，科学（*Science*）241: 456, 1988；Dervan等人，科学（*Science*）251: 1360, 1991）。可施用这些寡聚物本身，或者可在体内表达相关寡聚物。合成的反义或三链体寡核苷酸可包含经修饰碱基或经修饰主链。后者的范例包括甲基磷酸酯、硫代磷酸酯、或肽核酸主链。为了免受核酸酶的降解作用而提供保护，将这些主链掺入反义或三链体寡核苷酸，这在本领域是众所周知的。用这些或其它经修饰主链合成的反义和三链体分子也构成本发明的一个部分。

另外，可使用对IGS1 mRNA序列特异的核酶来预防IGS4多肽的表达。核酶是具有催化活性的天然或合成的RNA（参阅例如N. Usman等人，结构生物学现行观点（*Curr. Opin. Struct. Biol.*）6（4）：527-533, 1996）。合成核酶可设计成在选定位置特异切割IGS1 mRNA，从而预防IGS1 mRNA翻译成功能多肽。可用天然磷酸核糖主链和天然碱基合成核酶，正如RNA分子中常见的。或者，为了针对核糖核酸酶的降解作用提供保护，可用非天然主链（例如2'-O-甲基RNA）合成核酶，而

且可包含经修饰碱基。

对于涉及IGS4表达和活性不足的异常状况的治疗，也可使用几种方法。一种方法是与制药学可接受载体一起，对对象施用治疗有效量的激活IGS4的化合物（即上文所述激动剂），由此减轻异常状况。或者，可采用基因疗法来实现对象体内相关细胞内源生成IGS4。例如，如上所述，可改造本发明的多核苷酸，从而在复制缺陷型逆转录载体中进行表达。然后可分离逆转录病毒表达构建物，并导入用包含编码本发明多肽的RNA的逆转录质粒载体转导的包装细胞，使得包装细胞产生包含目的基因的侵染性病毒颗粒。可对对象施用这些生产者细胞，从而在体内改造细胞和在体内表达多肽。基因疗法的概述，参阅《人类分子遗传学》（Human Molecular Genetics），第20章“基因疗法和其它基于分子遗传学的治疗方法”（及其引用的参考文献），T. Strachan和A. P. Read，BIOS Scientific Publishers有限公司，1996。

可对需要这种疗法的任何对象应用任何上述治疗方法，包括例如哺乳动物，诸如狗、猫、牛、马、兔、猴，最优选为人。

配方和施用

肽（诸如IGS4多肽的可溶性形式）以及作为激动剂和拮抗剂的肽或小分子可以与合适的制药学载体进行联合配制。这些配方包含治疗有效量的多肽或化合物，以及制药学可接受的载体或赋形剂。配方应当适合施用模式，且完全属于本领域技术范围之内。本发明还涉及制药学包装和试剂盒，其中包含装有一种或多种本发明上述组合物成份的一种或多种容器。

本发明的多肽和其它化合物可单独使用，或者与其它化合物（诸如治疗性化合物）联合。

制药学组合物的全身施用的优选形式包括注射，通常是静脉内注射。也可使用其它注射路径，诸如皮下、肌肉内、或腹膜内。全身施用的候选方法包括使用渗透剂（诸如胆盐或梭链孢酸或其它去污剂）

的穿粘膜和穿皮肤施用。另外，若正确配制成肠配方或胶囊配方，则口服施用也是可能的。

需要的剂量范围取决于选择的肽或化合物、施用的路径、配方的本质、对象状况的本质、和主治医生的判断。合适的剂量范围是0.1 - 100 $\mu\text{g}/\text{kg}$ 对象体重。然而，考虑到可利用的化合物的多样性和不同施用路径的效率不同，预计所需剂量的变化很大。例如，预计口服施用要求的剂量比静脉内注射要大。正如本领域完全领会的，可使用标准经验性常规程序来调整这些剂量水平的变化以进行优化。

在治疗中使用的多肽也可在对象体内内源生成，这就是上文所述的常常称为“基因疗法”的治疗形式。因此，例如可使用多核苷酸（诸如DNA或RNA），例如通过使用逆转录病毒质粒载体，改造来自对象的细胞，从而离体编码多肽。然后，将细胞导入对象体内。

实施例

下列实施例只是为了进一步更详细的例示本发明，因此不应当认为这些实施例以任何方式限制本发明的范围。

实施例1：编码新的G蛋白偶联受体的cDNA的克隆

实施例1a：编码新的G蛋白偶联受体（GPCR）的基因组片段的同源性PCR克隆

使用基于PCR的同源性克隆策略来分离编码新的G蛋白偶联受体的部分基因组DNA序列。根据神经降压素受体基因家族（N. Vita等人，FEBS通讯（FEBS Lett.）317：139 - 142，1993；N. Vita等人，欧洲药理学杂志（Eur. J. Pharmacol.）360：265 - 272，1998）的保守区，即跨膜结构域1（TM1）和3（TM3）的内部及TM3与胞内环2（I2）之间的边界，设计了正向（F22）和反向（R44和R46）简并PCR引物：

F22 (TM1):

5'-CTCATCTTCGCGGTGGGC(A or G)C(A,C,G or T)G(C or T)(A,C,G or T)GG-3' (SEQ ID NO: 13)

R44 (TM3/12):

5'-GGCCAGGCAGCGCTCCGCGCT(C or 肌昔)A(A or G)(A,C,G or T)C(C or T)(A,C,G or T)GC(A,G or T)-3' (SEQ ID NO: 14)

R46 (TM3):

5'-GAA(A or G)TA(A or G)TAGCC(A or G)CG(A or G)CAGCC(A or T)-3' (SEQ ID NO: 15)

为了遏制神经降压素受体家族已知成员的扩增，选择的引物R44的3'末端核苷酸位置与NTR1和NTR2二者cDNA的对应位置不互补。PCR反应在60 μ l体积中进行，包含100ng 人基因组DNA (Clontech)、6 μ l GeneAmp™ 10x PCR缓冲液II (100mM Tris-HCl pH8.3、500mM KCl, Perkin Elmer)、3.6 μ l 25mM MgCl₂、0.36 μ l dNTP混合物 (含每种dNTP 25mM)、1.5U AmpliTaq-Gold™聚合酶 (Perkin Elmer)、和30pmol 每种简并正向 (F22) 及反向 (R44) 引物。将反应管于95℃加热10min，然后进行35个循环的变性 (95℃, 1min)、退火 (55℃, 2min)、和延伸 (72℃, 3min)。最后将反应管于72℃加热10min。

对于半嵌套PCR反应，使用1 μ l 1/50稀释的初级PCR反应液作为模板，分别使用F22和R46作为简并正向和反向引物。在与初级PCR反应相同的条件下进行半嵌套PCR反应。

将半嵌套PCR反应产物在琼脂糖凝胶上进行大小分离，并用溴化乙锭进行染色。虽然预期为 \pm 220bp的片段，但是只看见 \pm 120bp的片段。使用Qiaex-IF™纯化试剂盒 (Qiagen) 由凝胶纯化该片段，并使用pGEM-T试剂盒 (Promega) 参照供应商推荐的流程连接到pGEM-T质粒中。将由此产生的重组质粒用于转化感受态大肠杆菌 SURE™ 2 细菌 (Stratagene)。将经转化细胞涂布在含100 μ g/ml 氨苄青霉素、0.5mM IPTG、和50 μ g/ml X-gal的LB琼脂平板上。使用Qiagen-tip 20微量制备试剂盒 (Qiagen)，由个别菌落的微量培养物纯化质粒DNA。使用ABI Prism™ BigDye™终止物循环测序简易反应试剂盒 (PE-ABI)，使用插入片段侧翼引物，对经纯化质粒DNA进行DNA测序反应。

表7: 所用寡聚物引物的概述

SEQ ID NO: 13	F22: 5'-CTCATCTTCGCGGTGGGC(A or G)C(A,C,G or T)G(C or T)(A,C,G or T)GG-3'
SEQ ID NO: 14	R44: 5'-GGCCAGGCAGCGCTCCGCGCT(C or 肌昔)A(A or G)(A,C,G or T)C(C or T)(A,C,G or T)GC(A,G or T)-3'
SEQ ID NO: 15	R46: 5'-GAA(A or G)TA(A or G)TAGCC(A or G)CG(A or G)CAGCC(A or T)-3'
SEQ ID NO: 16	AP1: 5'-CCATCCTAATACGACTCACTATAGGGC-3'
SEQ ID NO: 17	AP2: 5'-ACTCACTATAGGGCTCGAGCGGC-3'
SEQ ID NO: 18	IGS4R1: 5'GGATCCCAAATAAGAAAGGGTAGTTGC-3'
SEQ ID NO: 19	IGS4R2: 5'AAAGGGTAGTTGCGCCACATCTCATAGAC-3'
SEQ ID NO: 20	IGS4F5: 5'AGGTCTATGAGATGTGGCGCAACTACCCT-3'
SEQ ID NO: 21	IGS4F6: 5'ATGTGGCGCAACTACCCTTCTTATTTGGG-3'
SEQ ID NO: 22	R74: 5'-CGGAAGTTGGCGGACACG(A or G)(A,C or G)(A or G)TT(A or G)TA-3'
SEQ ID NO: 23	IGS4F7: 5'-GCTCAGCTTGAAACAGAGCCTCGTACC-3'
SEQ ID NO: 24	IGS4F8: 5'-CCATGTGGATCTACAATTCATCATCC-3'
SEQ ID NO: 25	IGS4F9: 5'-AAGACAAATCTCTTGAGGCAGATGAAGGG-3'
SEQ ID NO: 26	IGS4F10: 5'-GATGCTGTTTGTCTTGGTCTTAGTGTTTGC-3'
SEQ ID NO: 27	IGS4R5: 5'-GGATGATGAAATTGTAGATCCACATGGGC-3'
SEQ ID NO: 28	IGS4R6: 5'-TGTGGAGAAGTCTCTCAAAGTGTGG-3'
SEQ ID NO: 29	IGS4R7: 5'-TAGTAGGAGTGACAGCCTGACTCGGAACG-3'
SEQ ID NO: 30	IGS4R8: 5'-AACGTAGATGACTCAGGACGAACCATTTCC-3'
SEQ ID NO: 31	IGS4F11: 5'-TCGTACCAGGGGAGGCTCAGGC-3'

通过EtOH/NaOAc沉淀来纯化测序反应产物，并在ABI 377自动化测序仪上进行分析。

克隆HNT1552的插入片段的序列分析显示，它有可能编码GPCR家族新成员的一部分。我们将这种新的GPCR序列称为IGS4。

实施例1b: 包含完整IGS4编码序列的cDNA片段的克隆

通过RACE (cDNA末端的快速扩增)分析和RT-PCR扩增二者获得了IGS4 cDNA的完整编码序列。使用与Marathon™ cDNA扩增试剂盒

(Clontech, 目录编号K1802-1) 一起提供的衔接头引物1和2 (AP1: SEQ ID NO: 16; AP2: SEQ ID NO: 17) 与IGS4特异引物, 在来自脑或睾丸的Marathon-Ready™ cDNA (Clontech, 目录编号分别为7400-1和7414-1) 上进行5'和3' RACE PCR反应。参照由Clontech提供的Marathon-Ready™ cDNA用户手册的指示, 进行PCR RACE反应。将RACE产物在琼脂糖凝胶上分离, 用溴化乙锭染色, 并转移到Hybond N⁺膜上。将印迹在经修改的Church缓冲液(0.5M 磷酸盐、7% SDS、10mMEDTA) 中于65℃预杂交2小时, 然后在含 2×10^6 cpm/ml ³²P标记的IGS4 cDNA探针的相同缓冲液中于65℃杂交过夜。使用Prime-It II试剂盒™ (Stratagene), 参照供应商提供的指示, 通过随机引发掺入[α -³²P]dCTP将IGS4 cDNA探针放射性标记至比活 $>10^9$ cpm/ μ g。在高严谨度(2x SSC/0.1% SDS, 室温, 2x30min; 随后0.1x SSC/0.1% SDS, 65℃, 2x40min) 清洗经杂交印迹, 并放射自显影过夜。由制备性凝胶纯化杂交片段, 克隆到pGEM-T载体中, 并如上所述测序。

使用IGS4特异引物IGS4R1 (SEQ ID NO: 18) 和IGS4R2 (SEQ ID NO: 19) (根据克隆HNT1552的DNA序列而设计) 对人脑cDNA的第一轮半嵌套5' RACE分析产生了克隆HNT1886和HNT1887 (图1)。这些克隆向IGS4 cDNA序列的上游延伸, 并越过推定的翻译起始密码子。同样, 使用IGS4特异引物IGS4F5 (SEQ ID NO: 20) 和IGS4F6 (SEQ ID NO: 21) 对人脑cDNA的第一轮3' RACE分析产生了克隆HNT1874 - 1878和HNT1902 - 1903 (图1)。这些克隆延伸了已知IGS4 cDNA的3'端。

将此刻获得的所有序列装配成包含长开放读码框的单一毗连序列群, 它编码与人孤儿受体FM-3 (Tan等人, 基因组学 (Genomics) 52: 223 - 229, 1998; GenBank编号AF044600和AF044601) 最相似的新蛋白质的一部分。为了分析IGS4的RNA表达分布型, 在供应商推荐的条件下, 将包含来自不同人组织的RNA的Master Blot™膜 (Clontech, 目录编号7770-1) 与³²P标记的克隆HNT1903插入片段进行杂交。由睾丸RNA获得最强杂交, 而由前列腺、胃、脊髓、海马、延髓、甲状腺、胸腺、肺、和气管获得的信号要弱得多。

由于毗连序列群的序列尚不包含完整的IGS4编码序列，因此我们在人全脑RNA上进行了RT-PCR同源性克隆实验，使用IGS4特异引物IGS4F6 (SEQ ID NO: 21) 和简并引物R74 (SEQ ID NO: 22)，后一引物是根据包括神经降压素受体1和2、生长激素促分泌素受体 (A. D. Howard等人, 科学 (Science) 273: 974 - 977, 1996)、及孤儿GPCR FM-3和GPR38 (K. K. McKee等人, 基因组学 (Genomics) 46: 426 - 434, 1997) 在内的GPCR亚族的保守区 (TM7/C端胞内部分) 而设计的。RT-PCR反应在50 μ l 体积中在500ng来自人脑的总RNA上进行，参照供应商的推荐使用Titar™单管RT-PCR系统 (Boehringer, 目录编号 1, 888, 382)。RT-PCR条件如下：于55℃ 逆转录45min；于94℃ 变性2min，随后20个循环的降温PCR反应 (于94℃ 变性30sec，于60℃ 退火30sec[-0.25℃/循环]，和于68℃ 延伸2min)，和另一轮30个PCR循环 (于94℃ 变性30sec，于55℃ 退火30sec，和于68℃ 延伸3min[+5sec/循环])。最后以68℃ 7min的额外延伸步骤作为结束。使用放射性标记的克隆HNT1903的插入片段，通过Southern印迹分析反应产物。使用QiaexIP™试剂盒 (Qiagen)，由凝胶纯化与探针发生杂交的 \pm 690bp 片段，并克隆到pGEM-T载体中，产生克隆HNT2210 - 2212。这些克隆的序列分析能够在3'方向上延伸现有的IGS4 cDNA毗连序列群。

由于延伸后的IGS4 cDNA毗连序列群仍然不含翻译终止密码子，因此设计了另外的IGS4特异3' RACE引物 (IGS4F7 - 10, SEQ ID NO: 23 - 26)。在来自人睾丸的Marathon Ready™ cDNA (Clontech, 目录编号7414-1) 上进行嵌套或半嵌套3' RACE反应。将IGS4特异条带 (使用IGS4特异探针通过Southern印迹评价) 克隆到pGEM-T中，产生克隆HNT2289 - 2290 (AP1/IGS4F5→AP2/IGS4F9)、HNT2293 - 2295 (AP1/IGS4F6→AP2/IGS4F9)、HNT2296 - 2297 (AP1/IGS4F7→AP2/IGS4F9)、HNT2308 - 2310 (AP1/IGS4F8→AP2/IGS4F10) 和 HNT2253 (AP1/IGS4F7→AP1/IGS4F5)。在睾丸Marathon Ready™ cDNA上进行的额外5' RACE PCR反应产生克隆HNT2279 - 2281

(AP2/IGS4R6→AP2/IGS4R5)。(注意:如AP1/IGS4F5→AP2/IGS4F9指克隆是由在使用引物对AP1/IGS4F5的初级RACE PCR反应并用引物对AP2/IGS4F9嵌套后获得的IGS4特异片段产生的)。

这些克隆的序列分析能够在3'方向上进一步延伸现有的IGS4 cDNA毗连序列群,但是仍未达到IGS4编码序列的末端。IGS4毗连序列群DNA序列对表达序列标签(EST)数据库(dbest)的计算机辅助同源性搜索(Blastn; S. F. Altschul等人,核酸研究(Nucleic Acids Res.) 25: 3389 - 3402, 1997)显示,存在与IGS4毗连序列群3'端交叠的EST序列(编号N45474)(交叠区的同一性接近100%)。EST N45474在3'方向上进一步延伸IGS4 DNA毗连序列群至翻译终止密码子和3'非翻译区(3'-UTR)。另外,鉴定到都覆盖IGS4 mRNA 3'-UTR的另一组EST(图2)。根据这些EST的3'-UTR设计了另外的IGS4特异引物(IGS4R7 - 8, SEQ ID NO: 29 - 30)。使用引物IGS4F7 (SEQ ID NO: 23)、IGS4F11 (SEQ ID NO: 31)、和IGS4R7 - 8 (SEQ ID NO: 29 - 30)的多种组合,在来自人睾丸的Marathon Ready™ cDNA上进行初级PCR反应。将PCR管于95℃加热2min,然后进行35个循环的变性(95℃, 30sec)、退火(65℃, 30sec)、和延伸(72℃, 90sec)。最后将反应管于72℃加热10min。同样在相同条件下进行嵌套PCR反应。由凝胶纯化±1630bp的DNA片段,并克隆到pGEM-T载体中,获得下列克隆:HNT2311、HNT2312、和HNT2317(IGS4F7/IGS4R7→IGS4F11/IGS4R8); HNT2313、HNT2324、HNT2326、和HNT2328(IGS4F11→IGS4R8); HNT2314、HNT2315、和HNT2322(IGS4F11→R7)。由纯化的1630bp PCR片段获得克隆HNT2363,该片段是由人睾丸Marathon Ready™ cDNA使用IGS4F11/R7引物对在如下略为改动的条件下扩增的:首先于94℃变性2min后,将PCR管进行15个循环的变性(94℃, 15sec)、退火(65℃, 30sec)、和延伸(72℃, 2min),随后另外20个循环的变性(94℃, 15sec)、退火(65℃, 30sec)、和延伸(72℃, 2min; +10sec/循环),最后是72℃ 7min的延伸步骤。这些克隆的序列分析能够装配IGS4 cDNA共有序列(图1)。对所有克隆的仔细检查显示,它们事实上属于2种序

列类型，有5个核苷酸位置是不同的。这些变体序列对应于人群中的多态性。我们将这些不同的cDNA类型称为IGS4ADNA (SEQ ID NO: 1和SEQ ID NO: 3) 和IGS4BDNA (SEQ ID NO: 5和SEQ ID NO: 7)。共有序列包含长的开放读码框，其中包含处于同一读码框的2个起始密码子 (IGS4ADNA和IGS4BDNA中的第55 - 57位 (SEQ ID NO: 1和SEQ ID NO: 5) 和第64 - 66位 (SEQ ID NO: 3和SEQ ID NO: 7))，预测显示与PGCR蛋白质具有较好同源性的415 (SEQ ID NO: 2和SEQ ID NO: 6) 或412 (SEQ ID NO: 4和SEQ ID NO: 8) 个氨基酸的蛋白质。蛋白质的亲水性分析 (J. Kyte等人, 分子生物学杂志 (J. Mol. Biol.) 157: 105 - 132, 1982; P. Klein等人, 生物化学和生物物理学学报 (Biochim. Biophys. Acta) 815: 468 - 476, 1985) 也指示存在7个跨膜结构域。由于第一个ATG起始密码子处于弱Kozak翻译起始序列内，而第二个处于强Kozak序列内，因此IGS4A/B蛋白质有可能由第二个甲硫氨酸起始，长度为412个氨基酸 (M. Kozak, 基因 (Gene) 234: 187 - 208, 1999)。但是，也不能排除有些 (甚至完全) 由第一个ATG起始。在5个多态性核苷酸中，4个 (IGS4A/BDNA中的第947、999、1202、和1216位) 导致编码的氨基酸残基的转变，而第5个 (IGS4A/BDNA中的第1381位) 位于3'-UTR内。各自预测的蛋白质序列称为IGS4APROT (SEQ ID NO: 2和SEQ ID NO: 4) 和IGS4BPROT (SEQ ID NO: 6和SEQ ID NO: 8)。(注意1: 在此文件中，IGS4APROT和IGS4BPROT序列代表最长的可能 (415个氨基酸) 序列，但是应当理解，实际的蛋白质可能在氨基末端短少3个氨基酸; 因此，表4和表5中IGS4APROT和IGS4BPROT的前3个氨基酸用括号括起来了) (注意2: 在此文件中，IGS4指一般性IGS4序列，而不考虑具体的等位基因类型)。IGS4蛋白质序列对公开的结构域蛋白质数据库的同源性搜索显示与人孤儿GPCR FM-3 (编号043664; C. P. Tan等人, 基因组学 (Genomics) 52: 223 - 229, 1998) 具有最佳同源性 (在IGS4A第26 - 342位氨基酸中具有46% 的同一性)。

IGS4 cDNA序列对DNA数据库的同源性搜索产生了许多条目，它们也是由IGS4基因座衍生的 (概述见图2)：

- 发现了都由IGS4 cDNA的3'-UTR衍生的10个EST序列条目（编号W61169、AI432384、W61131、AI023570、F01358、F03770、Z38158、R40869、R37725、和H11333）、2个STS（序列标签位点）（编号G20615和G05725）、和1个基因组序列（编号AQ078563）。

- EST编号N45474编码IGS4编码序列的3'端和3'-UTR的一部分（见上文）。

- 发现了由42个未排序的毗连序列群以任意顺序装配而成的“工作草稿”高通量基因组序列（编号AC008571，版本AC008571.1，1999年8月3日保藏），其中我们在4个分离的区域检测到完整的IGS4 cDNA序列。这些区域最有可能对应于不同的IGS4外显子，因为它们的侧翼就是规范的剪接供体和受体序列。根据这种分析，IGS4ADNA（或IGS4BDNA）序列中不同外显子的位置可定义如下：外显子1（第1-780位）、外显子2（第781-865位）、外显子3（第866-991位）、和外显子4（第992-1658位）。AC008571基因组序列属于IGS4A等位基因型。

- 发现了6个交叠EST条目（编号H11359、R13890、R13353、F07531、F05108、和F05107），其中装配的DNA序列在其3'端与IGS4外显子2和外显子3的前端交叠。然而外显子2上游的DNA序列与IGS4外显子1完全不同。可能这6个EST是由源自其它启动子的转录本衍生的。

- 最后发现了编码外显子2的2个基因组序列条目（编号AQ019411和AQ015065）。

我们在上述不同实验中分离的许多IGS4 cDNA克隆中，除了由未剪接（或部分剪接）的转录本衍生的许多克隆外，我们还发现了包含64bp（IGS4ADNA中的第866-929位）删除的许多克隆。至今，我们只发现了多态性A型的截短转录本。我们将这种剪接变体cDNA序列称为IGS4A-64DNA（SEQ ID NO: 9和SEQ ID NO: 11）。由于这种删除准确的发生于外显子2/外显子3边界，而经删除片段的最后2个核苷酸是“AG”，因此可能这种删除代表另外的剪接事件，其中外显子3中的“AG”作为剪接受体。由这种剪接变体编码的IGS4A读码框在删除点后发生读码框移位。编码的（截短的）296个氨基酸的蛋白质称为IGS4A-64PROT

(SEQ ID NO: 10和SEQ ID NO: 12)。IGS4A-64PROT序列的亲水性分析显示这种蛋白质只包含5个跨膜结构域(对应于IGS4APROT的TM结构域1-5)。这种截短的受体可能具有生理学相关性。

在涂布于含100 μ g/ml氨苄青霉素的LB琼脂平板上之后,将包含质粒HNT2322(包含IGS4ADNA插入片段)的细菌菌株再次克隆,并保藏于Innogenetics N.V.菌株表(ICCG4320)和荷兰Baarn的Centraalbureau voor Schimmelculturen(CBS)(编号CBS102221)。由再次克隆的分离物制备质粒DNA,并对插入片段再次测序,发现与IGS4ADNA序列相同。

在涂布于含100 μ g/ml氨苄青霉素的LB琼脂平板上之后,将包含质粒HNT2363(包含IGS4BDNA插入片段)的细菌菌株再次克隆,并保藏于Innogenetics N.V.菌株表(ICCG4340)和荷兰Baarn的Centraalbureau voor Schimmelculturen(CBS)(编号CBS102222)。由再次克隆的分离物制备质粒DNA,并对插入片段再次测序,发现与IGS4BDNA序列相同。

实施例2: 由神经介肽U在CHO α 16-IGS4细胞中诱导的胞内钙浓度特异性变化

实施例2a: 实验流程: 方法和材料

A. 用于IGS4转染CHO α 16细胞的方法和材料

实验中使用了下列材料: 包含IGS4 DNA序列的载体(IGS4-pcDNA3.1); SuperFect转染试剂(Qiagen); 含10% FCS的Nut-Mix F12(Gibco); 0.028mg/ml 庆大霉素(Gibco); 0.22mg/ml 潮霉素(Gibco)。

用于克隆选择的材料: 含10% FCS的Nut-Mix F12; 0.028mg/ml 庆大霉素; 0.22mg/ml 潮霉素; 和0.55mg/ml 遗传霉素(Gibco)。

● 应用了下列方法: 如制造商(Qiagen)所述用SuperFect转染试剂进行转染。将细胞铺在24孔板中至50%汇合。每个孔加入0.6 μ g/ μ l 质粒DNA及1 μ l SuperFect转染试剂。24小时后,更换培养基,并

通过含遗传霉素的选择培养基来选择经转染细胞克隆。通过RT-PCR和Northern印迹来鉴定表达IGS4的细胞克隆。

B. 用于FLIPR实验的方法和材料

细胞的制备:

对于细胞制备, 采用了下列材料: 平板: 清澈、平底、黑孔的96孔板 (Costar); 培养基: 生长培养基: 添加10% 胎牛血清 (Gibco) 的含Glutamax (Gibco) 的Nut-Mix F12 (HAM); 培养箱: 5% CO₂, 37°C (Nuair)。

方法进行如下: 在实验前24或48小时, 将细胞接种到黑壁微量培养板中。48小时培养的细胞密度是 0.8×10^4 细胞/孔, 24小时培养的细胞密度是 2.2×10^4 细胞/孔。所有步骤都是在无菌条件下进行的。

染料装载:

为了观察胞内钙水平的变化, 细胞必须“装载”对钙敏感的荧光染料。这种染料, 称为FLUO-4 (Molecular Probes), 只有在与钙形成复合物时, 才会在488nm受到激发, 并发出500 - 560nm范围的光。该染料在4 μ M终浓度下使用。加入Pluronic acid来提高染料的溶解度和细胞对染料的摄取。向染料介质中加入Probenicid (一种阴离子交换蛋白质抑制剂) 以提高染料在细胞中的保持。

使用了下列材料:

- 2mM染料原液: 将1mg Fluo-4 (Molecular Probes) 溶于443 μ l低水DMSO (Sigma)。分装保存于-20°C。
- 20% pluronic acid溶液: 于37°C将400mg pluronic acid (Sigma) 溶于2ml低水DMSO (Sigma)。保存于室温。
- 染料/pluronic acid混合物: 临使用前, 将等体积的染料原液与20% pluronic acid混合。染料与pluronic acid的终浓度分别为1mM与10%。
- 250mM Probenicid原液: 将710mg Probenicid (Sigma) 溶于5ml 1N NaOH, 并与5ml添加20mM HEPES的不含酚红的Hank氏BSS (Gibco) 混合。

- 装载缓冲液: 10.5ml添加20mM HEPES的不含酚红的Hank氏BSS (Gibco)、105 μ l Probenicid、和210 μ l 1M HEPES。

- 清洗缓冲液: 添加20mM HEPES (Gibco) 和2.5mM Probenicid的不含酚红的Hank氏BSS (Gibco)。

方法进行如下: 临加到装载缓冲液中之前, 将2mM染料原液与20% (w/v) Pluronic acid等体积混合。由孔中吸出生长培养基, 而不搅动汇合的细胞层。使用Multidrop (Labsystems) 将100 μ l 装载介质分配到每个孔中。将细胞在5% CO₂、37℃的培养箱中保温30分钟。为了计算背景荧光, 有些孔不装载染料。这些孔的背景荧光来自细胞的自身荧光。装载染料后, 用清洗缓冲液 (自动化Denley细胞清洗器) 清洗细胞3次, 将基础荧光降低至背景以上计数20,000 - 25,000。加入100 μ l 缓冲液, 并将细胞于37℃保温, 直至开始实验。

C. 化合物平板的制备

制备3 μ M (终浓度的3x) 肽用于初步筛选。对于浓度应答曲线, 制备浓度范围30 μ M - 100nM的肽溶液。所有肽都是在含0.1% BSA (Sigma) 的缓冲液中稀释的。

使用了下列材料: 肽: 猪神经介肽U25、大鼠神经介肽U23、猪神经介肽U8 (Bachem); 稀释缓冲液: 添加20mM HEPES (Gibco) 和0.1% BSA (Sigma) 的不含酚红的Hank氏BSS (Gibco); 平板: 清澈、平底96孔板 (Costar)。

D. 实验

FLIPR设置参数设置成0.4sec曝光时长, 滤器1, 50 μ l 流体添加, 移液器高度125 μ l, 分配速度40 μ l/sec, 不混合。

实施例2b: 结果

为了鉴定孤儿G蛋白偶联受体 (GPCR) IGS4的内源配体, 将IGS4 (IGS4A和IGS4B两种形式) 稳定转染到中国仓鼠卵巢 (CHO) 细胞中。由于不知道IGS4的G蛋白偶联机制, 故使用特定的CHO细胞株。这些CHO细胞稳定表达G蛋白G α 16 (CHO G α 16, Molecular Devices), 它是GPCR

的“通用衔接物”(G. Milligan, F. Marshall, 和S. Rees, 《G α 16作为通用的G蛋白衔接物: 激动剂筛选策略的暗示》(G α 16 as a universal G protein adapter: implications for agonist screening strategies.), TIPS 17: 235 - 237, 1996)。

对产生的CHO G α 16-IGS4细胞在荧光成像平板读数器(FLIPR)上进行功能筛选, 测量应答推定肽配体的胞内钙动员。在神经介肽U23的浓度10nM, 诱导了较大的、瞬时的、且强烈的钙应答。相反, CHO G α 16细胞和表达另一种无关孤儿GPCR的CHO G α 16细胞不应答神经介肽U23。使用IGS4B的这些实验结果显示于图4。

此外, 调查了IGS4激活对猪和大鼠神经介肽U异构体的浓度依赖性(对于IGS4A和IGS4B二者)。在范围 10^{-6} - 10^{-12} M内, 猪神经介肽U25、大鼠神经介肽U23、猪神经介肽U8在FLIPR实验中诱导特异的由IGS4介导的钙动员。测试的所有3种神经介肽U异构体都引起IGS4B的相同最大激活, LogEC₅₀值为 -10.09 ± 0.08 (神经介肽U8, n = 4; 80pM)、 -10.61 ± 0.08 (神经介肽U23, n = 10; 50pM)、和 -9.14 ± 0.09 (神经介肽U25, n = 3; 1.2nM)。因此, 所有3种肽都引起特别是IGS4B的有力激活, 说明神经介肽U是该受体的天然激动剂。使用IGS4B的这些实验结果显示于图3a(神经介肽U8)、图3b(神经介肽U23)、和图3c(神经介肽U25)。

对于IGS4A受体, 发现了稍微低一点的亲和力, 但是仍然显示神经介肽U肽是一般IGS4受体的较好配体。IGS4A的LogEC₅₀值如下: 神经介肽U8的LogEC₅₀值 = -9.3 ± 0.09 (n = 1; 485pM); 神经介肽U23的LogEC₅₀值 = -7.27 ± 0.16 (n = 6; 53nM); 和神经介肽U25的LogEC₅₀值 = -6.18 ± 0.14 (n = 3; 658nM)。

神经介肽U激活IGS4后看见的钙动员应答说明该受体偶联Gq/11亚族的G蛋白。另外, CHO G α 16-IGS4细胞的胞内cAMP基础水平不受猪神经介肽U8(1和10 μ M)的调控, 说明该受体不与Gs亚族的G蛋白偶联(资料未显示)。

实施例3: 人多组织表达阵列(MTETM)上的IGS4杂交

使用Prime-It II试剂盒™(Stratagene), 通过随机引发掺入[α - 32 P]dCTP将人IGS4A DNA(来自pGEMT-hIGS4A[ICCG#4320]的 ± 730 bp BamHI/HindIII插入片段)放射性标记至比活 $>10^9$ cpm/ μ g。通过Sephadex G-50层析将经标记探针与游离标记物纯化开来, 于95℃变性5min, 并加入ExpressHyb杂交液至终浓度1-1.5 $\times 10^6$ cpm/ml。参照供应商的推荐, 将人多组织表达(MTE™)阵列(Clontech, 目录编号7775-1)在ExpressHyb液中于65℃预杂交30min和杂交过夜。

将已杂交的MTE™阵列用2x SSC/1% SDS于65℃清洗20min, 5次; 然后用0.1x SSC/0.5% SDS于55℃清洗20min, 2次。清洗后, 通过磷光成像(Cyclone Storage Phosphor System, Packard)使阵列放射自显影(图5)。使用OptiQuant图象分析软件(Packard)定量分析MTE™阵列的杂交数据。对包含RNA的不同斑点位置的信号强度校正由空白位置获得的平均背景信号。由包含大肠杆菌DNA的斑点获得的信号强度被视作代表不展示IGS4表达的样品。信号强度低于大肠杆菌DNA的样品被视作是阴性的。

通过由每个值减去由大肠杆菌DNA观察到的杂交信号(认为它是背景信号), 重新计算RNA阵列上不同组织的杂交信号。显示较低杂交信号的所有组织被视作是低于背景的, 且是IGS4阴性的。将相对于睾丸(100%)中的表达水平绘图, 提供于图7。

实施例4: IGS4组织分布的Northern印迹分析

使用Prime-It II试剂盒™(Stratagene), 通过随机引发掺入[α - 32 P]dCTP将人IGS4A DNA(来自pGEMT-hIGS4A[ICCG#4320]的 ± 730 bp BamHI/HindIII插入片段)放射性标记至比活 $>10^9$ cpm/ μ g。通过Sephadex G-50层析将经标记探针与游离标记物纯化开来, 于95℃变性5min, 并加入ExpressHyb杂交液至终浓度1-1.5 $\times 10^6$ cpm/ml。参照供应商的推荐, 将人Northern印迹(Clontech, 目录编号7760-1、7759-1、7767-1、7755-1、和7769-1)在ExpressHyb液中于65℃预杂交30min和杂交过夜。

杂交后, 将Northern印迹用2x SSC/0.05% SDS于室温清洗10min, 4次; 然后用0.1x SSC/0.1% SDS于50℃清洗40min, 2次。清洗后, 使用磷光体贮存平板(Cyclone Storage Phosphor System, Packard)和X射线胶片使Northern印迹放射自显影。Northern印迹的结果显示于图6。

Northern印迹分析的结果似乎与阵列杂交(实施例3)在很大程度上是一致的。在睾丸中发现了至今最强的信号(2.4kb转录本)。在胸腺、脊髓、髓质、甲状腺、丘脑、和黑质中发现了较弱的2.4kb条带, 在胼胝体、尾状核、和胃中发现了很弱的条带。对于有些组织, 在Northern上看不到2.4kb条带, 而在MTE阵列上观察到强烈至中等的杂交信号(如全脑、大脑皮层、肺、颞叶和额叶、扁桃体、小脑、肾、及海马)。

实施例5: 定量RT-PCR分析

还使用LightCycler™仪器(Roche Diagnostics)和IGS4特异TaqMan™探针, 通过实时定量RT-PCR(Q-PCR), 测定了不同人组织中的IGS4表达水平。

实施例5a: 实验流程

A. cDNA合成

在逆转录前, 将来自人总RNA组I至V(Clontech, 目录编号K4000-1至K4004-1)的3μg总RNA用3U DNase I(Life Technologies, 目录编号18068-015)在30μl反应体积(20mM Tris pH8.3、50mM KCl、2mM KCl)中于室温处理15分钟, 以破坏可能污染的基因组DNA。加入3μl 25mM EDTA并于65℃加热10分钟来终止反应。将2.6μg经DNA酶处理的RNA与1.3μg oligo(dT)(Life Technologies, 目录编号18418-012)退火, 并使用Omniscript逆转录酶(Qiagen, 目录编号205111)参照酶供应商推荐的方案在52μl反应体积中于37℃保温1小时进行逆转录。于93℃加热5分钟灭活Omniscript逆转录酶。

B. Q-PCR

定量PCR反应在20 μ l反应混合物中进行, 其中包含1x TaqMan™通用PCR Mastermix (PE Applied Biosystems, 目录编号4304437)、0.12mg/ml BSA、900nM IGS4特异正向和反向引物 (IP14,963和IP14,964)、250nM IGS4特异TaqMan™探针 (IP14,962)、和1.6 μ l (IGS4) 或0.16 μ l (GAPDH: 甘油醛-3-磷酸脱氢酶) cDNA合成反应液作为模板。为了建立IGS4标准曲线, 使用了系列稀释 ($10^7 - 10^1$ 拷贝/反应) 的IGS4质粒ICCG4320; 而对于GAPDH标准曲线, 使用系列稀释的人脑cDNA合成反应液 (0.16 μ l、0.016 μ l、和0.0016 μ l) 作为模板。1x TaqMan™通用PCR Mastermix包含AmpliTaq Gold™ DNA聚合酶、AmpErase™ UNG(尿嘧啶-N-糖基化酶)、含dUTP的dNTP混合物、passive reference、和优化的缓冲液成份。IGS4特异引物和TaqMan探针是使用Primer Express™软件 (PE Applied Biosystems) 设计的。在与IGS4所述相同的条件下进行人GADPH的定量PCR反应, 不同之处是使用来自TaqMan™ GAPDH对照试剂盒 (PE Applied Biosystems, 目录编号402869; 序列信息不能由PE Applied Biosystems获得) 的GAPDH特异引物和TaqMan™探针。

PCR反应在LightCycler仪器™的玻璃毛细管中进行。首先于50℃保温2分钟以进行AmpErase™ UNG反应(从而启动并激活AmpliTaq Gold DNA聚合酶(95℃ 10min))后, 将反应混合物进行40个循环的变性(95℃, 15sec)和退火/延伸(60℃ [GAPDH]或68℃ [IGS4], 1min)。使用LightCycler软件3.0版进行实验样品的定量。在IGS4质粒拷贝范围 $10 - 10^7$ 内, 在IGS4质粒的量与报导染料的释放之间获得较好的线性相关性。我们还用GAPDH TaqMan™探针, 使用系列稀释的脑cDNA, 获得了线性标准曲线。相对GAPDH表达水平处于骨骼肌的0.4 - 10.2%范围内, 而骨骼肌在所有测试组织中具有最高的GAPDH表达水平。相对IGS4表达水平表述成相对于脊髓中检测水平的比例, 而脊髓在所有测试组织中具有最高的IGS4表达(图8)。我们还在对GAPDH持家基因的表达进行标准化后, 将相对IGS4表达水平绘图(图8)。

实施例5b: 结果

使用IGS4特异TaqMan探针的Q-PCR显示, 在脊髓中发现了最高表达水平(未对GAPDH标准化)。脊髓中的IGS4表达水平总计11,467拷贝mRNA/ng pA RNA(假定cDNA合成反应的效率为100%, 且假定pA RNA占总RNA的2%)。还在脑(脊髓水平的41%)、骨骼肌(37%)、小脑(31%)、睾丸(19%)、肺(12%)、和心(11%)中发现了高IGS4表达水平。在胎脑(5%)、气管(4%)、前列腺(2%)、和甲状腺(1.4%)中发现有较低水平。对GAPDH表达标准化后, 相对IGS4表达模式在很大程度上保持不变, 除了骨骼肌是例外, 其中相对表达水平跌至脊髓的2%。由于不清楚对GADPH的标准化是否是正确的程序(GAPDH表达水平在不同细胞/组织类型中预期变化较大或较小), 我们宁愿关注未标准化的相对表达水平。

意识到睾丸、脊髓、和脑似乎是最显著的表达位点之一, 这些Q-PCR数据似乎符合来自RNA阵列(实施例3)和Northern印迹(实施例4)杂交实验的表达数据。然而, Q-PCR分析还显示许多其它组织中有重要表达, 诸如骨骼肌、小脑、肺、和心。

在不知道特异受体的情况下, 已经有人提出, 配体神经介肽U是神经肽或神经调质(J. Domin, M. A. Ghatei, P. Chohan, 和S. R. Bloom, 肽(Peptides) 8: 779 - 784, 1987)。我们的调查显示, IGS4是神经介肽U受体家族的新成员, 在CNS和PNS区域、胃肠系统、免疫系统、泌尿生殖系统、心血管系统、骨骼肌、甲状腺、和肺中表达。

表8: 用于IGS4 Q-PCR反应的PCR寡核苷酸引物和TaqMan探针的概述

SEQ ID NO: 32	IP 14,963	5'-CCTCTTCAGCCTGGCGGTCTCTG-3'
SEQ ID NO: 33	IP 14,964	5'-GGAGGCGAAGCACACGGTCTCA-3'
SEQ ID NO: 34	IP 14,962	5'(FAM)-AGATGTGGCGCAACTACCCTTTCTTGTTCCGGGCC-(TAMRA)3'

此说明书中引用的所有发表物, 包括但不限于专利和专利申请,

完整收入本文作为参考，就像特定的且个别的指定将单独的发表物完整收入本文作为参考如完整列出一样。

图的简述

图1: 分离用于产生共有IGS4 cDNA序列的不同cDNA克隆的相对位置的示意图。还指出了所用5'和3' RACE引物(分别为IGS4R#和IGS4F#), 以及EST编号N45474的位置。引物IGS4R6位于内含子1内。有些克隆(如HNT2311、HNT2312、和HNT2253)只是部分测序(只指出了测序的部分)。共有序列A和共有序列B分别表示IGS4等位基因类型A和B的共有序列。对每一个克隆指出(阴影盒)了在4个多态性位置每一处鉴定的核苷酸。“S”表示克隆HNT2211和HNT2212中的序列差异, 指“C”或“T”。IGS4A和IGS4B共有序列的编码区用“**”指示。由于共有序列的5'端存在一些残余序列差异, IGS4ADNA和IGS4BDNA序列只是由第86位至末端。

图2: 不同DNA数据库条目相对于IGS4 cDNA序列的相对位置的示意图。IGS4 cDNA序列用盒指示(IGS4编码序列的位置用实心盒指示)。在IGS4 cDNA序列上面指出(“=”)了IGS4外显子1-4的相对位置。基因组序列AC008571中编码外显子1->4的部分分别用AC008571a->d指示。AC008571序列内的这些片段的位置是: AC008571a(AC008571反向互补的第13129-13908位)、AC008571b(AC008571的第51676-51760位)、AC008571c(AC008571反向互补的第79978-80103位)、和AC008571d(AC008571反向互补的第83060-83728位)。G05725和G20615是STS(序列标识的序列)条目, 而F05107、F05108、F07531、R13353、R13890、H11359、N45474、W61169、AI432384、W61131、AI023570、F01358、F03770、Z38158、R40869、R37725、和H11333是EST条目。基因组克隆AQ019411和AQ015065中包含IGS4外显子2的部分用“=”指示。与IGS4 cDNA序列完全不同的EST序列F05107、F05108、F07531、R13353、R13890、和H11359的5'部分用“*”指示。AQ078563是基因组克隆。

图3: 不同神经介肽U异构体对IGS4受体的激活。将CHO α 16-IGS4B

细胞在96孔板中培养过夜，并装载Fluo-4AM。用FLIPR (Molecular Devices) 测量由受体介导的Ca²⁺变化。将由CCD照相机检测到的荧光最大变化标准化成1，并表述成计数。

图3a: 神经介肽U8的结果;

图3b: 神经介肽U23的结果;

图3c: 神经介肽U25的结果。

图4: 神经介肽U23在表达IGS4B的CHO α 16细胞中诱导胞内Ca²⁺动员。对细胞系CHO α 16-IGS4、CHO α 16、和用另一种孤儿GPCR转染的CHO α 16应用10nM 神经介肽U23。在96孔板中培养细胞过夜，并装载Fluo-4AM。用FLIPR (Molecular Devices) 测量由受体介导的胞内Ca²⁺变化，表述成由CCD照相机检测到的计数。

图5: 使用人IGS4探针的人多组织表达阵列。

图6: 使用IGS4探针的Northern印迹分析。

图7: IGS4表达分析 (MTE印迹)。

图8: IGS4 mRNA相对于在脊髓中观察到的表达的相对表达水平。显示了未标准化和GAPDH标准化的两种表达水平。

序列表

<110> SOLVAY PHARMACEUTICALS B.V.

<120>新的人G蛋白偶联受体

<130> SPW99.06/HA 00.19

<140>

<141>

<160> 34

<170> PatentIn Ver. 2.1

<210> 1

<211> 1658

<212> DNA

<213> 人类 (Homo sapiens)

<220>

<221> CDS

<222> (55)..(1299)

<223> IGS4A 长型

<400> 1

```

ggctcagctt gaaacagagc ctcgtaccag gggaggctca ggccttggat ttta atg      57
                                                    Met
                                                    1

tca ggg atg gaa aaa ctt cag aat gct tcc tgg atc tac cag cag aaa      105
Ser Gly Met Glu Lys Leu Gln Asn Ala Ser Trp Ile Tyr Gln Gln Lys
           5                10                15

cta gaa gat cca ttc cag aaa cac ctg aac agc acc gag gag tat ctg      153
Leu Glu Asp Pro Phe Gln Lys His Leu Asn Ser Thr Glu Glu Tyr Leu
           20                25                30

gcc ttc ctc tgc gga cct cgg cgc agc cac ttc ttc ctc ccc gtg tct      201
Ala Phe Leu Cys Gly Pro Arg Arg Ser His Phe Phe Leu Pro Val Ser
           35                40                45

gtg gtg tat gtg cca att ttt gtg gtg ggg gtc att ggc aat gtc ctg      249
Val Val Tyr Val Pro Ile Phe Val Val Gly Val Ile Gly Asn Val Leu
           50                55                60                65

gtg tgc ctg gtg att ctg cag cac cag gct atg aag acg ccc acc aac      297
Val Cys Leu Val Ile Leu Gln His Gln Ala Met Lys Thr Pro Thr Asn
           70                75                80

tac tac ctc ttc agc ctg gcg gtc tct gac ctc ctg gtc ctg ctc ctt      345
Tyr Tyr Leu Phe Ser Leu Ala Val Ser Asp Leu Leu Val Leu Leu Leu
           85                90                95

gga atg ccc ctg gag gtc tat gag atg tgg cgc aac tac cct ttc ttg      393
Gly Met Pro Leu Glu Val Tyr Glu Met Trp Arg Asn Tyr Pro Phe Leu

```

100	105	110	
ttc ggg ccc gtg ggc tgc tac ttc aag acg gcc ctc ttt gag acc gtg Phe Gly Pro Val Gly Cys Tyr Phe Lys Thr Ala Leu Phe Glu Thr Val 115	120	125	441
tgc ttc gcc tcc atc ctc agc atc acc acc gtc agc gtg gag cgc tac Cys Phe Ala Ser Ile Leu Ser Ile Thr Thr Val Ser Val Glu Arg Tyr 130	135	140	489
gtg gcc atc cta cac ccg ttc cgc gcc aaa ctg cag agc acc cgg cgc Val Ala Ile Leu His Pro Phe Arg Ala Lys Leu Gln Ser Thr Arg Arg 150	155	160	537
cgg gcc ctc agg atc ctc ggc atc gtc tgg ggc ttc tcc gtg ctc ttc Arg Ala Leu Arg Ile Leu Gly Ile Val Trp Gly Phe Ser Val Leu Phe 165	170	175	585
tcc ctg ccc aac acc agc atc cat ggc atc aag ttc cac tac ttc ccc Ser Leu Pro Asn Thr Ser Ile His Gly Ile Lys Phe His Tyr Phe Pro 180	185	190	633
aat ggg tcc ctg gtc cca ggt tgc gcc acc tgt acg gtc atc aag ccc Asn Gly Ser Leu Val Pro Gly Ser Ala Thr Cys Thr Val Ile Lys Pro 195	200	205	681
atg tgg atc tac aat ttc atc atc cag gtc acc tcc ttc cta ttc tac Met Trp Ile Tyr Asn Phe Ile Ile Gln Val Thr Ser Phe Leu Phe Tyr 210	215	220	729
ctc ctc ccc atg act gtc atc agt gtc ctc tac tac ctc atg gca ctc Leu Leu Pro Met Thr Val Ile Ser Val Leu Tyr Tyr Leu Met Ala Leu 230	235	240	777
aga cta aag aaa gac aaa tct ctt gag gca gat gaa ggg aat gca aat Arg Leu Lys Lys Asp Lys Ser Leu Glu Ala Asp Glu Gly Asn Ala Asn 245	250	255	825
att caa aga ccc tgc aga aaa tca gtc aac aag atg ctg ttt gtc ttg Ile Gln Arg Pro Cys Arg Lys Ser Val Asn Lys Met Leu Phe Val Leu 260	265	270	873
gtc tta gtg ttt gct atc tgt tgg gcc ccg ttc cac att gac cga ctc Val Leu Val Phe Ala Ile Cys Trp Ala Pro Phe His Ile Asp Arg Leu 275	280	285	921
ttc ttc agc ttt gtg gag gag tgg agt gaa tcc ctg gct gct gtg ttc Phe Phe Ser Phe Val Glu Glu Trp Ser Glu Ser Leu Ala Ala Val Phe 290	295	300	969
aac ctc gtc cat gtg gtg tca ggt gtc ttc ttc tac ctg agc tca gct Asn Leu Val His Val Val Ser Gly Val Phe Phe Tyr Leu Ser Ser Ala 310	315	320	1017
gtc aac ccc att atc tat aac cta ctg tct cgc cgc ttc cag gca gca Val Asn Pro Ile Ile Tyr Asn Leu Leu Ser Arg Arg Phe Gln Ala Ala			1065

325	330	335	
ttc cag aat gtg atc tct tct ttc cac aaa cag tgg cac tcc cag cat			1113
Phe Gln Asn Val Ile Ser Ser Phe His Lys Gln Trp His Ser Gln His			
340	345	350	
gac cca cag ttg cca cct gcc cag cgg aac atc ttc ctg aca gaa tgc			1161
Asp Pro Gln Leu Pro Pro Ala Gln Arg Asn Ile Phe Leu Thr Glu Cys			
355	360	365	
cac ttt gtg gag ctg acc gaa gat ata ggt ccc caa ttc cca tgt cag			1209
His Phe Val Glu Leu Thr Glu Asp Ile Gly Pro Gln Phe Pro Cys Gln			
370	375	380	385
tca tcc atg cac aac tct cac ctc cca aca gcc ctc tct agt gaa cag			1257
Ser Ser Met His Asn Ser His Leu Pro Thr Ala Leu Ser Ser Glu Gln			
390	395	400	
atg tca aga aca aac tat caa agc ttc cac ttt aac aaa acc			1299
Met Ser Arg Thr Asn Tyr Gln Ser Phe His Phe Asn Lys Thr			
405	410	415	
tgaattcttt cagagctgac tctcctctat gcctcaaaac ttcagagagg aacatcccat			1359
aatgtatgcc ttctcatatg atattagaga ggtagaatgg ctcttacaac tcatgtaccc			1419
attgctagtt tttttttttt aataaacgtg aaaactgaga gttagatctg gtttcaaaac			1479
ccaagactgc ctgattttta gttatctttc cactatccta actgcctcat gccccttcac			1539
tagttcatgc caagaacgtg actggaaaagg catggcacct ataccttgat taatttccat			1599
taatggaaat ggttegtcct gagtcatcta cgttccgagt caggctgtca ctctacta			1658
<210> 2			
<211> 415			
<212> PRT			
<213> 人类			
<400> 2			
Met Ser Gly Met Glu Lys Leu Gln Asn Ala Ser Trp Ile Tyr Gln Gln			
1	5	10	15
Lys Leu Glu Asp Pro Phe Gln Lys His Leu Asn Ser Thr Glu Glu Tyr			
20	25	30	
Leu Ala Phe Leu Cys Gly Pro Arg Arg Ser His Phe Phe Leu Pro Val			
35	40	45	
Ser Val Val Tyr Val Pro Ile Phe Val Val Gly Val Ile Gly Asn Val			
50	55	60	
Leu Val Cys Leu Val Ile Leu Gln His Gln Ala Met Lys Thr Pro Thr			
65	70	75	80

Gln Ser Ser Met His Asn Ser His Leu Pro Thr Ala Leu Ser Ser Glu
385 390 395 400

Gln Met Ser Arg Thr Asn Tyr Gln Ser Phe His Phe Asn Lys Thr
405 410 415

<210> 3

<211> 1658

<212> DNA

<213> 人类

<220>

<221> CDS

<222> (64)..(1299)

<223> IGS4A 短型

<400> 3

ggctcagctt gaaacagagc ctcgtaccag gggaggctca ggccttggat tttaatgtca 60

ggg atg gaa aaa ctt cag aat gct tcc tgg atc tac cag cag aaa cta 108
Met Glu Lys Leu Gln Asn Ala Ser Trp Ile Tyr Gln Gln Lys Leu
1 5 10 15

gaa gat cca ttc cag aaa cac ctg aac agc acc gag gag tat ctg gcc 156
Glu Asp Pro Phe Gln Lys His Leu Asn Ser Thr Glu Glu Tyr Leu Ala
20 25 30

ttc ctc tgc gga cct cgg cgc agc cac ttc ttc ctc ccc gtg tct gtg 204
Phe Leu Cys Gly Pro Arg Arg Ser His Phe Phe Leu Pro Val Ser Val
35 40 45

gtg tat gtg cca att ttt gtg gtg ggg gtc att ggc aat gtc ctg gtg 252
Val Tyr Val Pro Ile Phe Val Val Gly Val Ile Gly Asn Val Leu Val
50 55 60

tgc ctg gtg att ctg cag cac cag gct atg aag acg ccc acc aac tac 300
Cys Leu Val Ile Leu Gln His Gln Ala Met Lys Thr Pro Thr Asn Tyr
65 70 75

tac ctc ttc agc ctg gcg gtc tct gac ctc ctg gtc ctg ctc ctt gga 348
Tyr Leu Phe Ser Leu Ala Val Ser Asp Leu Leu Val Leu Leu Leu Gly
80 85 90 95

atg ccc ctg gag gtc tat gag atg tgg cgc aac tac cct ttc ttg ttc 396
Met Pro Leu Glu Val Tyr Glu Met Trp Arg Asn Tyr Pro Phe Leu Phe
100 105 110

ggg ccc gtg ggc tgc tac ttc aag acg gcc ctc ttt gag acc gtg tgc 444
Gly Pro Val Gly Cys Tyr Phe Lys Thr Ala Leu Phe Glu Thr Val Cys
115 120 125

ttc gcc tcc atc ctc agc atc acc acc gtc agc gtg gag cgc tac gtg 492
Phe Ala Ser Ile Leu Ser Ile Thr Thr Val Ser Val Glu Arg Tyr Val

130	135	140	
gcc atc cta cac ccg ttc cgc gcc aaa ctg cag agc acc cgg cgc cgg			540
Ala Ile Leu His Pro Phe Arg Ala Lys Leu Gln Ser Thr Arg Arg Arg			
145	150	155	
gcc ctc agg atc ctc ggc atc gtc tgg ggc ttc tcc gtg ctc ttc tcc			588
Ala Leu Arg Ile Leu Gly Ile Val Trp Gly Phe Ser Val Leu Phe Ser			
160	165	170	175
ctg ccc aac acc agc atc cat ggc atc aag ttc cac tac ttc ccc aat			636
Leu Pro Asn Thr Ser Ile His Gly Ile Lys Phe His Tyr Phe Pro Asn			
180	185	190	
ggg tcc ctg gtc cca ggt tcg gcc acc tgt acg gtc atc aag ccc atg			684
Gly Ser Leu Val Pro Gly Ser Ala Thr Cys Thr Val Ile Lys Pro Met			
195	200	205	
tgg atc tac aat ttc atc atc cag gtc acc tcc ttc cta ttc tac ctc			732
Trp Ile Tyr Asn Phe Ile Ile Gln Val Thr Ser Phe Leu Phe Tyr Leu			
210	215	220	
ctc ccc atg act gtc atc agt gtc ctc tac tac ctc atg gca ctc aga			780
Leu Pro Met Thr Val Ile Ser Val Leu Tyr Tyr Leu Met Ala Leu Arg			
225	230	235	
cta aag aaa gac aaa tct ctt gag gca gat gaa ggg aat gca aat att			828
Leu Lys Lys Asp Lys Ser Leu Glu Ala Asp Glu Gly Asn Ala Asn Ile			
240	245	250	255
caa aga ccc tgc aga aaa tca gtc aac aag atg ctg ttt gtc ttg gtc			876
Gln Arg Pro Cys Arg Lys Ser Val Asn Lys Met Leu Phe Val Leu Val			
260	265	270	
tta gtg ttt gct atc tgt tgg gcc ccg ttc cac att gac cga ctc ttc			924
Leu Val Phe Ala Ile Cys Trp Ala Pro Phe His Ile Asp Arg Leu Phe			
275	280	285	
ttc agc ttt gtg gag gag tgg agt gaa tcc ctg gct gct gtg ttc aac			972
Phe Ser Phe Val Glu Glu Trp Ser Glu Ser Leu Ala Ala Val Phe Asn			
290	295	300	
ctc gtc cat gtg gtg tca ggt gtc ttc ttc tac ctg agc tca gct gtc			1020
Leu Val His Val Val Ser Gly Val Phe Phe Tyr Leu Ser Ser Ala Val			
305	310	315	
aac ccc att atc tat aac cta ctg tct cgc cgc ttc cag gca gca ttc			1068
Asn Pro Ile Ile Tyr Asn Leu Leu Ser Arg Arg Phe Gln Ala Ala Phe			
320	325	330	335
cag aat gtg atc tct tct ttc cac aaa cag tgg cac tcc cag cat gac			1116
Gln Asn Val Ile Ser Ser Phe His Lys Gln Trp His Ser Gln His Asp			
340	345	350	
cca cag ttg cca cct gcc cag cgg aac atc ttc ctg aca gaa tgc cac			1164
Pro Gln Leu Pro Pro Ala Gln Arg Asn Ile Phe Leu Thr Glu Cys His			

```

          355                360                365
ttt gtg gag ctg acc gaa gat ata ggt ccc caa ttc cca tgt cag tca 1212
Phe Val Glu Leu Thr Glu Asp Ile Gly Pro Gln Phe Pro Cys Gln Ser
          370                375                380

tcc atg cac aac tct cac ctc cca aca gcc ctc tct agt gaa cag atg 1260
Ser Met His Asn Ser His Leu Pro Thr Ala Leu Ser Ser Glu Gln Met
          385                390                395

tca aga aca aac tat caa agc ttc cac ttt aac aaa acc tgaattcttt 1309
Ser Arg Thr Asn Tyr Gln Ser Phe His Phe Asn Lys Thr
          400                405                410

cagagctgac tctcctctat gcctcaaaac ttcagagagg aacatcccat aatgtatgcc 1369

ttctcatatg atattagaga ggtagaatgg ctcttacaac tcatgtaccc attgctagtt 1429

tttttttttt aataaacgtg aaaactgaga gttagatctg gtttcaaaac ccaagactgc 1489

ctgatttttta gttatctttc cactatccta actgcctcat gcccttcac tagttcatgc 1549

caagaacgtg actggaaagg catggcacct ataccttgat taatttccat taatggaaat 1609

ggttcgtcct gagtcatcta cgttccgagt caggctgtca ctctacta 1658

<210> 4
<211> 412
<212> PRT
<213> 人类

<400> 4
Met Glu Lys Leu Gln Asn Ala Ser Trp Ile Tyr Gln Gln Lys Leu Glu
  1                5                10                15

Asp Pro Phe Gln Lys His Leu Asn Ser Thr Glu Glu Tyr Leu Ala Phe
          20                25                30

Leu Cys Gly Pro Arg Arg Ser His Phe Phe Leu Pro Val Ser Val Val
          35                40                45

Tyr Val Pro Ile Phe Val Val Gly Val Ile Gly Asn Val Leu Val Cys
          50                55                60

Leu Val Ile Leu Gln His Gln Ala Met Lys Thr Pro Thr Asn Tyr Tyr
          65                70                75                80

Leu Phe Ser Leu Ala Val Ser Asp Leu Leu Val Leu Leu Leu Gly Met
          85                90                95

Pro Leu Glu Val Tyr Glu Met Trp Arg Asn Tyr Pro Phe Leu Phe Gly
          100                105                110

Pro Val Gly Cys Tyr Phe Lys Thr Ala Leu Phe Glu Thr Val Cys Phe
          115                120                125

```

Ala Ser Ile Leu Ser Ile Thr Thr Val Ser Val Glu Arg Tyr Val Ala
 130 135 140

Ile Leu His Pro Phe Arg Ala Lys Leu Gln Ser Thr Arg Arg Arg Ala
 145 150 155 160

Leu Arg Ile Leu Gly Ile Val Trp Gly Phe Ser Val Leu Phe Ser Leu
 165 170 175

Pro Asn Thr Ser Ile His Gly Ile Lys Phe His Tyr Phe Pro Asn Gly
 180 185 190

Ser Leu Val Pro Gly Ser Ala Thr Cys Thr Val Ile Lys Pro Met Trp
 195 200 205

Ile Tyr Asn Phe Ile Ile Gln Val Thr Ser Phe Leu Phe Tyr Leu Leu
 210 215 220

Pro Met Thr Val Ile Ser Val Leu Tyr Tyr Leu Met Ala Leu Arg Leu
 225 230 235 240

Lys Lys Asp Lys Ser Leu Glu Ala Asp Glu Gly Asn Ala Asn Ile Gln
 245 250 255

Arg Pro Cys Arg Lys Ser Val Asn Lys Met Leu Phe Val Leu Val Leu
 260 265 270

Val Phe Ala Ile Cys Trp Ala Pro Phe His Ile Asp Arg Leu Phe Phe
 275 280 285

Ser Phe Val Glu Glu Trp Ser Glu Ser Leu Ala Ala Val Phe Asn Leu
 290 295 300

Val His Val Val Ser Gly Val Phe Phe Tyr Leu Ser Ser Ala Val Asn
 305 310 315 320

Pro Ile Ile Tyr Asn Leu Leu Ser Arg Arg Phe Gln Ala Ala Phe Gln
 325 330 335

Asn Val Ile Ser Ser Phe His Lys Gln Trp His Ser Gln His Asp Pro
 340 345 350

Gln Leu Pro Pro Ala Gln Arg Asn Ile Phe Leu Thr Glu Cys His Phe
 355 360 365

Val Glu Leu Thr Glu Asp Ile Gly Pro Gln Phe Pro Cys Gln Ser Ser
 370 375 380

Met His Asn Ser His Leu Pro Thr Ala Leu Ser Ser Glu Gln Met Ser
 385 390 395 400

Arg Thr Asn Tyr Gln Ser Phe His Phe Asn Lys Thr
 405 410

```

<210> 5
<211> 1658
<212> DNA
<213> 人类

<220>
<221> CDS
<222> (55)..(1299)
<223> IGS4B 长型

<400> 5
ggctcagctt gaaacagagc ctcgtaccag gggaggctca ggccttggat ttta atg 57
                                                Met
                                                1

tca ggg atg gaa aaa ctt cag aat gct tcc tgg atc tac cag cag aaa 105
Ser Gly Met Glu Lys Leu Gln Asn Ala Ser Trp Ile Tyr Gln Gln Lys
                    5                10                15

cta gaa gat cca ttc cag aaa cac ctg aac agc acc gag gag tat ctg 153
Leu Glu Asp Pro Phe Gln Lys His Leu Asn Ser Thr Glu Glu Tyr Leu
                    20                25                30

gcc ttc ctc tgc gga cct cgg cgc agc cac ttc ttc ctc ccc gtg tct 201
Ala Phe Leu Cys Gly Pro Arg Arg Ser His Phe Phe Leu Pro Val Ser
                    35                40                45

gtg gtg tat gtg cca att ttt gtg gtg ggg gtc att ggc aat gtc ctg 249
Val Val Tyr Val Pro Ile Phe Val Val Gly Val Ile Gly Asn Val Leu
                    50                55                60                65

gtg tgc ctg gtg att ctg cag cac cag gct atg aag acg ccc acc aac 297
Val Cys Leu Val Ile Leu Gln His Gln Ala Met Lys Thr Pro Thr Asn
                    70                75                80

tac tac ctc ttc agc ctg gcg gtc tct gac ctc ctg gtc ctg ctc ctt 345
Tyr Tyr Leu Phe Ser Leu Ala Val Ser Asp Leu Leu Val Leu Leu Leu
                    85                90                95

gga atg ccc ctg gag gtc tat gag atg tgg cgc aac tac cct ttc ttg 393
Gly Met Pro Leu Glu Val Tyr Glu Met Trp Arg Asn Tyr Pro Phe Leu
                    100                105                110

ttc ggg ccc gtg ggc tgc tac ttc aag acg gcc ctc ttt gag acc gtg 441
Phe Gly Pro Val Gly Cys Tyr Phe Lys Thr Ala Leu Phe Glu Thr Val
                    115                120                125

tgc ttc gcc tcc atc ctc agc atc acc acc gtc agc gtg gag cgc tac 489
Cys Phe Ala Ser Ile Leu Ser Ile Thr Thr Val Ser Val Glu Arg Tyr
                    130                135                140                145

gtg gcc atc cta cac ccg ttc cgc gcc aaa ctg cag agc acc cgg cgc 537
Val Ala Ile Leu His Pro Phe Arg Ala Lys Leu Gln Ser Thr Arg Arg
                    150                155                160

```

cgg gcc ctc agg atc ctc ggc atc gtc tgg ggc ttc tcc gtg ctc ttc	585
Arg Ala Leu Arg Ile Leu Gly Ile Val Trp Gly Phe Ser Val Leu Phe	
165 170 175	
tcc ctg ccc aac acc agc atc cat ggc atc aag ttc cac tac ttc ccc	633
Ser Leu Pro Asn Thr Ser Ile His Gly Ile Lys Phe His Tyr Phe Pro	
180 185 190	
aat ggg tcc ctg gtc cca ggt tcg gcc acc tgt acg gtc atc aag ccc	681
Asn Gly Ser Leu Val Pro Gly Ser Ala Thr Cys Thr Val Ile Lys Pro	
195 200 205	
atg tgg atc tac aat ttc atc atc cag gtc acc tcc ttc cta ttc tac	729
Met Trp Ile Tyr Asn Phe Ile Ile Gln Val Thr Ser Phe Leu Phe Tyr	
210 215 220 225	
ctc ctc ccc atg act gtc atc agt gtc ctc tac tac ctc atg gca ctc	777
Leu Leu Pro Met Thr Val Ile Ser Val Leu Tyr Tyr Leu Met Ala Leu	
230 235 240	
aga cta aag aaa gac aaa tct ctt gag gca gat gaa ggg aat gca aat	825
Arg Leu Lys Lys Asp Lys Ser Leu Glu Ala Asp Glu Gly Asn Ala Asn	
245 250 255	
att caa aga ccc tgc aga aaa tca gtc aac aag atg ctg ttt gtc ttg	873
Ile Gln Arg Pro Cys Arg Lys Ser Val Asn Lys Met Leu Phe Val Leu	
260 265 270	
gtc tta gtg ttt gct atc tgt tgg gcc ccg ttc cac att gac cga ctc	921
Val Leu Val Phe Ala Ile Cys Trp Ala Pro Phe His Ile Asp Arg Leu	
275 280 285	
ttc ttc agc ttt gtg gag gag tgg act gaa tcc ctg gct gct gtg ttc	969
Phe Phe Ser Phe Val Glu Glu Trp Thr Glu Ser Leu Ala Ala Val Phe	
290 295 300 305	
aac ctc gtc cat gtg gtg tca ggt gtc tta ttc tac ctg agc tca gct	1017
Asn Leu Val His Val Val Ser Gly Val Leu Phe Tyr Leu Ser Ser Ala	
310 315 320	
gtc aac ccc att atc tat aac cta ctg tct cgc cgc ttc cag gca gca	1065
Val Asn Pro Ile Ile Tyr Asn Leu Leu Ser Arg Arg Phe Gln Ala Ala	
325 330 335	
ttc cag aat gtg atc tct tct ttc cac aaa cag tgg cac tcc cag cat	1113
Phe Gln Asn Val Ile Ser Ser Phe His Lys Gln Trp His Ser Gln His	
340 345 350	
gac cca cag ttg cca cct gcc cag cgg aac atc ttc ctg aca gaa tgc	1161
Asp Pro Gln Leu Pro Pro Ala Gln Arg Asn Ile Phe Leu Thr Glu Cys	
355 360 365	
cac ttt gtg gag ctg acc gaa gat ata ggt ccc caa ttc cta tgt cag	1209
His Phe Val Glu Leu Thr Glu Asp Ile Gly Pro Gln Phe Leu Cys Gln	
370 375 380 385	

tca tcc gtg cac aac tct cac ctc cca aca gcc ctc tct agt gaa cag 1257
 Ser Ser Val His Asn Ser His Leu Pro Thr Ala Leu Ser Ser Glu Gln
 390 395 400

atg tca aga aca aac tat caa agc ttc cac ttt aac aaa acc 1299
 Met Ser Arg Thr Asn Tyr Gln Ser Phe His Phe Asn Lys Thr
 405 410 415

tgaattcttt cagagctgac tctcctctat gcctcaaac ttcagagagg aacatcccat 1359

aatgtatgcc ttctcatatg aaattagaga ggtagaatgg ctcttacaac tcatgtaccc 1419

attgctagtt tttttttttt aataaacgtg aaaactgaga gttagatctg gtttcaaac 1479

ccaagactgc ctgattttta gttatctttc cactatccta actgcctcat gccccttcac 1539

tagttcatgc caagaacgtg actggaaagg catggcacct ataccttgat taatttccat 1599

taatggaaat ggttcgtcct gagtcatcta cgttccgagt caggctgtca ctcctacta 1658

<210> 6

<211> 415

<212> PRT

<213> 人类

<400> 6

Met Ser Gly Met Glu Lys Leu Gln Asn Ala Ser Trp Ile Tyr Gln Gln
 1 5 10 15

Lys Leu Glu Asp Pro Phe Gln Lys His Leu Asn Ser Thr Glu Glu Tyr
 20 25 30

Leu Ala Phe Leu Cys Gly Pro Arg Arg Ser His Phe Phe Leu Pro Val
 35 40 45

Ser Val Val Tyr Val Pro Ile Phe Val Val Gly Val Ile Gly Asn Val
 50 55 60

Leu Val Cys Leu Val Ile Leu Gln His Gln Ala Met Lys Thr Pro Thr
 65 70 75 80

Asn Tyr Tyr Leu Phe Ser Leu Ala Val Ser Asp Leu Leu Val Leu Leu
 85 90 95

Leu Gly Met Pro Leu Glu Val Tyr Glu Met Trp Arg Asn Tyr Pro Phe
 100 105 110

Leu Phe Gly Pro Val Gly Cys Tyr Phe Lys Thr Ala Leu Phe Glu Thr
 115 120 125

Val Cys Phe Ala Ser Ile Leu Ser Ile Thr Thr Val Ser Val Glu Arg
 130 135 140

Tyr Val Ala Ile Leu His Pro Phe Arg Ala Lys Leu Gln Ser Thr Arg
 145 150 155 160

Arg Arg Ala Leu Arg Ile Leu Gly Ile Val Trp Gly Phe Ser Val Leu
 165 170 175
 Phe Ser Leu Pro Asn Thr Ser Ile His Gly Ile Lys Phe His Tyr Phe
 180 185 190
 Pro Asn Gly Ser Leu Val Pro Gly Ser Ala Thr Cys Thr Val Ile Lys
 195 200 205
 Pro Met Trp Ile Tyr Asn Phe Ile Ile Gln Val Thr Ser Phe Leu Phe
 210 215 220
 Tyr Leu Leu Pro Met Thr Val Ile Ser Val Leu Tyr Tyr Leu Met Ala
 225 230 235 240
 Leu Arg Leu Lys Lys Asp Lys Ser Leu Glu Ala Asp Glu Gly Asn Ala
 245 250 255
 Asn Ile Gln Arg Pro Cys Arg Lys Ser Val Asn Lys Met Leu Phe Val
 260 265 270
 Leu Val Leu Val Phe Ala Ile Cys Trp Ala Pro Phe His Ile Asp Arg
 275 280 285
 Leu Phe Phe Ser Phe Val Glu Glu Trp Thr Glu Ser Leu Ala Ala Val
 290 295 300
 Phe Asn Leu Val His Val Val Ser Gly Val Leu Phe Tyr Leu Ser Ser
 305 310 315 320
 Ala Val Asn Pro Ile Ile Tyr Asn Leu Leu Ser Arg Arg Phe Gln Ala
 325 330 335
 Ala Phe Gln Asn Val Ile Ser Ser Phe His Lys Gln Trp His Ser Gln
 340 345 350
 His Asp Pro Gln Leu Pro Pro Ala Gln Arg Asn Ile Phe Leu Thr Glu
 355 360 365
 Cys His Phe Val Glu Leu Thr Glu Asp Ile Gly Pro Gln Phe Leu Cys
 370 375 380
 Gln Ser Ser Val His Asn Ser His Leu Pro Thr Ala Leu Ser Ser Glu
 385 390 395 400
 Gln Met Ser Arg Thr Asn Tyr Gln Ser Phe His Phe Asn Lys Thr
 405 410 415

<210> 7
 <211> 1658
 <212> DNA
 <213> 人类

<220>

<221> CDS

<222> (64)..(1299)

<223> IGS4B 短型

<400> 7

ggctcagctt gaaacagagc ctcgtaccag gggaggctca ggccttggat tttaatgtca 60

ggg atg gaa aaa ctt cag aat gct tcc tgg atc tac cag cag aaa cta 108
 Met Glu Lys Leu Gln Asn Ala Ser Trp Ile Tyr Gln Gln Lys Leu
 1 5 10 15

gaa gat cca ttc cag aaa cac ctg aac agc acc gag gag tat ctg gcc 156
 Glu Asp Pro Phe Gln Lys His Leu Asn Ser Thr Glu Glu Tyr Leu Ala
 20 25 30

ttc ctc tgc gga cct cgg cgc agc cac ttc ttc ctc ccc gtg tct gtg 204
 Phe Leu Cys Gly Pro Arg Arg Ser His Phe Phe Leu Pro Val Ser Val
 35 40 45

gtg tat gtg cca att ttt gtg gtg ggg gtc att ggc aat gtc ctg gtg 252
 Val Tyr Val Pro Ile Phe Val Val Gly Val Ile Gly Asn Val Leu Val
 50 55 60

tgc ctg gtg att ctg cag cac cag gct atg aag acg ccc acc aac tac 300
 Cys Leu Val Ile Leu Gln His Gln Ala Met Lys Thr Pro Thr Asn Tyr
 65 70 75

tac ctc ttc agc ctg gcg gtc tct gac ctc ctg gtc ctg ctc ctt gga 348
 Tyr Leu Phe Ser Leu Ala Val Ser Asp Leu Leu Val Leu Leu Leu Gly
 80 85 90 95

atg ccc ctg gag gtc tat gag atg tgg cgc aac tac cct ttc ttg ttc 396
 Met Pro Leu Glu Val Tyr Glu Met Trp Arg Asn Tyr Pro Phe Leu Phe
 100 105 110

ggg ccc gtg ggc tgc tac ttc aag acg gcc ctc ttt gag acc gtg tgc 444
 Gly Pro Val Gly Cys Tyr Phe Lys Thr Ala Leu Phe Glu Thr Val Cys
 115 120 125

ttc gcc tcc atc ctc agc atc acc acc gtc agc gtg gag cgc tac gtg 492
 Phe Ala Ser Ile Leu Ser Ile Thr Thr Val Ser Val Glu Arg Tyr Val
 130 135 140

gcc atc cta cac ccg ttc cgc gcc aaa ctg cag agc acc cgg cgc cgg 540
 Ala Ile Leu His Pro Phe Arg Ala Lys Leu Gln Ser Thr Arg Arg Arg
 145 150 155

gcc ctc agg atc ctc ggc atc gtc tgg ggc ttc tcc gtg ctc ttc tcc 588
 Ala Leu Arg Ile Leu Gly Ile Val Trp Gly Phe Ser Val Leu Phe Ser
 160 165 170 175

ctg ccc aac acc agc atc cat ggc atc aag ttc cac tac ttc ccc aat 636
 Leu Pro Asn Thr Ser Ile His Gly Ile Lys Phe His Tyr Phe Pro Asn
 180 185 190

ggg tcc ctg gtc cca ggt tgg gcc acc tgt acg gtc atc aag ccc atg 684
 Gly Ser Leu Val Pro Gly Ser Ala Thr Cys Thr Val Ile Lys Pro Met
 195 200 205

tgg atc tac aat ttc atc atc cag gtc acc tcc ttc cta ttc tac ctc 732
 Trp Ile Tyr Asn Phe Ile Ile Gln Val Thr Ser Phe Leu Phe Tyr Leu
 210 215 220

ctc ccc atg act gtc atc agt gtc ctc tac tac ctc atg gca ctc aga 780
 Leu Pro Met Thr Val Ile Ser Val Leu Tyr Tyr Leu Met Ala Leu Arg
 225 230 235

cta aag aaa gac aaa tct ctt gag gca gat gaa ggg aat gca aat att 828
 Leu Lys Lys Asp Lys Ser Leu Glu Ala Asp Glu Gly Asn Ala Asn Ile
 240 245 250 255

caa aga ccc tgc aga aaa tca gtc aac aag atg ctg ttt gtc ttg gtc 876
 Gln Arg Pro Cys Arg Lys Ser Val Asn Lys Met Leu Phe Val Leu Val
 260 265 270

tta gtg ttt gct atc tgt tgg gcc ccg ttc cac att gac cga ctc ttc 924
 Leu Val Phe Ala Ile Cys Trp Ala Pro Phe His Ile Asp Arg Leu Phe
 275 280 285

ttc agc ttt gtg gag gag tgg act gaa tcc ctg gct gct gtg ttc aac 972
 Phe Ser Phe Val Glu Glu Trp Thr Glu Ser Leu Ala Ala Val Phe Asn
 290 295 300

ctc gtc cat gtg gtg tca ggt gtc tta ttc tac ctg agc tca gct gtc 1020
 Leu Val His Val Val Ser Gly Val Leu Phe Tyr Leu Ser Ser Ala Val
 305 310 315

aac ccc att atc tat aac cta ctg tct cgc cgc ttc cag gca gca ttc 1068
 Asn Pro Ile Ile Tyr Asn Leu Leu Ser Arg Arg Phe Gln Ala Ala Phe
 320 325 330 335

cag aat gtg atc tct tct ttc cac aaa cag tgg cac tcc cag cat gac 1116
 Gln Asn Val Ile Ser Ser Phe His Lys Gln Trp His Ser Gln His Asp
 340 345 350

cca cag ttg cca cct gcc cag cgg aac atc ttc ctg aca gaa tgc cac 1164
 Pro Gln Leu Pro Pro Ala Gln Arg Asn Ile Phe Leu Thr Glu Cys His
 355 360 365

ttt gtg gag ctg acc gaa gat ata ggt ccc caa ttc cta tgt cag tca 1212
 Phe Val Glu Leu Thr Glu Asp Ile Gly Pro Gln Phe Leu Cys Gln Ser
 370 375 380

tcc gtg cac aac tct cac ctc cca aca gcc ctc tct agt gaa cag atg 1260
 Ser Val His Asn Ser His Leu Pro Thr Ala Leu Ser Ser Glu Gln Met
 385 390 395

tca aga aca aac tat caa agc ttc cac ttt aac aaa acc tgaattcttt 1309
 Ser Arg Thr Asn Tyr Gln Ser Phe His Phe Asn Lys Thr
 400 405 410

cagagctgac tctcctctat gcctcaaaac ttcagagagg aacatcccat aatgtatgcc 1369
 ttctcatatg aaattagaga ggtagaatgg ctcttacaac tcatgtaccc attgctagtt 1429
 tttttttttt aataaacgtg aaaactgaga gttagatctg gtttcaaaac ccaagactgc 1489
 ctgattttta gttatctttc cactatccta actgcctcat gcccttcac tagttcatgc 1549
 caagaacgtg actggaaagg catggcacct atacctgat taatttccat taatggaaat 1609
 ggttcgtcct gagtcatcta cgttccgagt caggctgtca ctcctacta 1658

<210> 8
 <211> 412
 <212> PRT
 <213> 人类

<400> 8
 Met Glu Lys Leu Gln Asn Ala Ser Trp Ile Tyr Gln Gln Lys Leu Glu
 1 5 10 15
 Asp Pro Phe Gln Lys His Leu Asn Ser Thr Glu Glu Tyr Leu Ala Phe
 20 25 30
 Leu Cys Gly Pro Arg Arg Ser His Phe Phe Leu Pro Val Ser Val Val
 35 40 45
 Tyr Val Pro Ile Phe Val Val Gly Val Ile Gly Asn Val Leu Val Cys
 50 55 60
 Leu Val Ile Leu Gln His Gln Ala Met Lys Thr Pro Thr Asn Tyr Tyr
 65 70 75 80
 Leu Phe Ser Leu Ala Val Ser Asp Leu Leu Val Leu Leu Leu Gly Met
 85 90 95
 Pro Leu Glu Val Tyr Glu Met Trp Arg Asn Tyr Pro Phe Leu Phe Gly
 100 105 110
 Pro Val Gly Cys Tyr Phe Lys Thr Ala Leu Phe Glu Thr Val Cys Phe
 115 120 125
 Ala Ser Ile Leu Ser Ile Thr Thr Val Ser Val Glu Arg Tyr Val Ala
 130 135 140
 Ile Leu His Pro Phe Arg Ala Lys Leu Gln Ser Thr Arg Arg Arg Ala
 145 150 155 160
 Leu Arg Ile Leu Gly Ile Val Trp Gly Phe Ser Val Leu Phe Ser Leu
 165 170 175
 Pro Asn Thr Ser Ile His Gly Ile Lys Phe His Tyr Phe Pro Asn Gly
 180 185 190
 Ser Leu Val Pro Gly Ser Ala Thr Cys Thr Val Ile Lys Pro Met Trp

195	200	205
Ile Tyr Asn Phe Ile Ile Gln Val Thr Ser Phe Leu Phe Tyr Leu Leu 210	215	220
Pro Met Thr Val Ile Ser Val Leu Tyr Tyr Leu Met Ala Leu Arg Leu 225	230	235 240
Lys Lys Asp Lys Ser Leu Glu Ala Asp Glu Gly Asn Ala Asn Ile Gln 245	250	255
Arg Pro Cys Arg Lys Ser Val Asn Lys Met Leu Phe Val Leu Val Leu 260	265	270
Val Phe Ala Ile Cys Trp Ala Pro Phe His Ile Asp Arg Leu Phe Phe 275	280	285
Ser Phe Val Glu Glu Trp Thr Glu Ser Leu Ala Ala Val Phe Asn Leu 290	295	300
Val His Val Val Ser Gly Val Leu Phe Tyr Leu Ser Ser Ala Val Asn 305	310	315 320
Pro Ile Ile Tyr Asn Leu Leu Ser Arg Arg Phe Gln Ala Ala Phe Gln 325	330	335
Asn Val Ile Ser Ser Phe His Lys Gln Trp His Ser Gln His Asp Pro 340	345	350
Gln Leu Pro Pro Ala Gln Arg Asn Ile Phe Leu Thr Glu Cys His Phe 355	360	365
Val Glu Leu Thr Glu Asp Ile Gly Pro Gln Phe Leu Cys Gln Ser Ser 370	375	380
Val His Asn Ser His Leu Pro Thr Ala Leu Ser Ser Glu Gln Met Ser 385	390	395 400
Arg Thr Asn Tyr Gln Ser Phe His Phe Asn Lys Thr 405	410	

<210> 9

<211> 1594

<212> DNA

<213> 人类

<220>

<221> CDS

<222> (55)..(942)

<223> IGS4A 截短的DNA长型

<400> 9

ggctcagctt gaaacagagc ctcgtaccag gggaggctca ggccttggat ttta atg 57
Met

1

tca ggg atg gaa aaa ctt cag aat gct tcc tgg atc tac cag cag aaa	105
Ser Gly Met Glu Lys Leu Gln Asn Ala Ser Trp Ile Tyr Gln Gln Lys	
5	10
15	
cta gaa gat cca ttc cag aaa cac ctg aac agc acc gag gag tat ctg	153
Leu Glu Asp Pro Phe Gln Lys His Leu Asn Ser Thr Glu Glu Tyr Leu	
20	25
30	
gcc ttc ctc tgc gga cct cgg cgc agc cac ttc ttc ctc ccc gtg tct	201
Ala Phe Leu Cys Gly Pro Arg Arg Ser His Phe Phe Leu Pro Val Ser	
35	40
45	
gtg gtg tat gtg cca att ttt gtg gtg ggg gtc att ggc aat gtc ctg	249
Val Val Tyr Val Pro Ile Phe Val Val Gly Val Ile Gly Asn Val Leu	
50	55
60	65
gtg tgc ctg gtg att ctg cag cac cag gct atg aag acg ccc acc aac	297
Val Cys Leu Val Ile Leu Gln His Gln Ala Met Lys Thr Pro Thr Asn	
70	75
80	
tac tac ctc ttc agc ctg gcg gtc tct gac ctc ctg gtc ctg ctc ctt	345
Tyr Tyr Leu Phe Ser Leu Ala Val Ser Asp Leu Leu Val Leu Leu Leu	
85	90
95	
gga atg ccc ctg gag gtc tat gag atg tgg cgc aac tac cct ttc ttg	393
Gly Met Pro Leu Glu Val Tyr Glu Met Trp Arg Asn Tyr Pro Phe Leu	
100	105
110	
ttc ggg ccc gtg ggc tgc tac ttc aag acg gcc ctc ttt gag acc gtg	441
Phe Gly Pro Val Gly Cys Tyr Phe Lys Thr Ala Leu Phe Glu Thr Val	
115	120
125	
tgc ttc gcc tcc atc ctc agc atc acc acc gtc agc gtg gag cgc tac	489
Cys Phe Ala Ser Ile Leu Ser Ile Thr Thr Val Ser Val Glu Arg Tyr	
130	135
140	145
gtg gcc atc cta cac ccg ttc cgc gcc aaa ctg cag agc acc cgg cgc	537
Val Ala Ile Leu His Pro Phe Arg Ala Lys Leu Gln Ser Thr Arg Arg	
150	155
160	
cgg gcc ctc agg atc ctc ggc atc gtc tgg ggc ttc tcc gtg ctc ttc	585
Arg Ala Leu Arg Ile Leu Gly Ile Val Trp Gly Phe Ser Val Leu Phe	
165	170
175	
tcc ctg ccc aac acc agc atc cat ggc atc aag ttc cac tac ttc ccc	633
Ser Leu Pro Asn Thr Ser Ile His Gly Ile Lys Phe His Tyr Phe Pro	
180	185
190	
aat ggg tcc ctg gtc cca ggt tcg gcc acc tgt acg gtc atc aag ccc	681
Asn Gly Ser Leu Val Pro Gly Ser Ala Thr Cys Thr Val Ile Lys Pro	
195	200
205	
atg tgg atc tac aat ttc atc atc cag gtc acc tcc ttc cta ttc tac	729
Met Trp Ile Tyr Asn Phe Ile Ile Gln Val Thr Ser Phe Leu Phe Tyr	

```

210                215                220                225
ctc ctc ccc atg act gtc atc agt gtc ctc tac tac ctc atg gca ctc 777
Leu Leu Pro Met Thr Val Ile Ser Val Leu Tyr Tyr Leu Met Ala Leu
                230                235                240

aga cta aag aaa gac aaa tct ctt gag gca gat gaa ggg aat gca aat 825
Arg Leu Lys Lys Asp Lys Ser Leu Glu Ala Asp Glu Gly Asn Ala Asn
                245                250                255

att caa aga ccc tgc aga aaa tca gtc aac aag atg ctg tct ttg tgg 873
Ile Gln Arg Pro Cys Arg Lys Ser Val Asn Lys Met Leu Ser Leu Trp
                260                265                270

agg agt gga gtg aat ccc tgg ctg ctg tgt tca acc tcg tcc atg tgg 921
Arg Ser Gly Val Asn Pro Trp Leu Leu Cys Ser Thr Ser Ser Met Trp
                275                280                285

tgt cag gtg tct tct tct acc tgagctcagc tgtcaacccc attatctata 972
Cys Gln Val Ser Ser Ser Thr
290                295

acctactgtc tcgccgcttc caggcagcat tccagaatgt gatctcttct ttccacaaac 1032

agtggcactc ccagcatgac ccacagttgc cacctgcccc gcggaacatc ttcttgacag 1092

aatgccactt tgtggagctg accgaagata taggtcccca attcccatgt cagtcatcca 1152

tgcacaactc tcacctcca acagccctct ctagtgaaca gatgtcaaga acaaactatc 1212

aaagcttcca cttaacaaa acctgaattc tttcagagct gactctctc tatgectcaa 1272

aacttcagag aggaacatcc cataatgtat gccttctcat atgatattag agaggtagaa 1332

tggctcttac aactcatgta ccattgcta gttttttttt ttaataaac gtgaaaactg 1392

agagttagat ctggtttcaa aaccaagac tgectgattt ttagttatct ttccactatc 1452

ctaactgcct catgccctt cactagttca tgccaagaac gtgactggaa aggcattggca 1512

cctatacctt gattaatttc cattaatgga aatggttcgt cctgagtcac ctacgttccg 1572

agtcaggctg tcactcctac ta 1594

```

```

<210> 10
<211> 296
<212> PRT
<213> 人类

```

```

<400> 10
Met Ser Gly Met Glu Lys Leu Gln Asn Ala Ser Trp Ile Tyr Gln Gln
  1          5          10          15

Lys Leu Glu Asp Pro Phe Gln Lys His Leu Asn Ser Thr Glu Glu Tyr
  20          25          30

```

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

<210> 11

<211> 1594

<212> DNA

<213> 人类

<220>

<221> CDS

<222> (64)..(942)

<223> IGS4A 截短的DNA短型

<400> 11

```

ggctcagctt gaaacagagc ctcgtaccag gggaggctca ggccttggat tttaatgtca 60

ggg atg gaa aaa ctt cag aat gct tcc tgg atc tac cag cag aaa cta 108
Met Glu Lys Leu Gln Asn Ala Ser Trp Ile Tyr Gln Gln Lys Leu
  1             5             10             15

gaa gat cca ttc cag aaa cac ctg aac agc acc gag gag tat ctg gcc 156
Glu Asp Pro Phe Gln Lys His Leu Asn Ser Thr Glu Glu Tyr Leu Ala
             20             25             30

ttc ctc tgc gga cct cgg cgc agc cac ttc ttc ctc ccc gtg tct gtg 204
Phe Leu Cys Gly Pro Arg Arg Ser His Phe Phe Leu Pro Val Ser Val
             35             40             45

gtg tat gtg cca att ttt gtg gtg ggg gtc att ggc aat gtc ctg gtg 252
Val Tyr Val Pro Ile Phe Val Val Gly Val Ile Gly Asn Val Leu Val
             50             55             60

tgc ctg gtg att ctg cag cac cag gct atg aag acg ccc acc aac tac 300
Cys Leu Val Ile Leu Gln His Gln Ala Met Lys Thr Pro Thr Asn Tyr
             65             70             75

tac ctc ttc agc ctg gcg gtc tct gac ctc ctg gtc ctg ctc ctt gga 348
Tyr Leu Phe Ser Leu Ala Val Ser Asp Leu Leu Val Leu Leu Leu Gly
             80             85             90             95

atg ccc ctg gag gtc tat gag atg tgg cgc aac tac cct ttc ttg ttc 396
Met Pro Leu Glu Val Tyr Glu Met Trp Arg Asn Tyr Pro Phe Leu Phe
             100             105             110

ggg ccc gtg ggc tgc tac ttc aag acg gcc ctc ttt gag acc gtg tgc 444
Gly Pro Val Gly Cys Tyr Phe Lys Thr Ala Leu Phe Glu Thr Val Cys
             115             120             125

ttc gcc tcc atc ctc agc atc acc acc gtc agc gtg gag cgc tac gtg 492
Phe Ala Ser Ile Leu Ser Ile Thr Thr Val Ser Val Glu Arg Tyr Val
             130             135             140

gcc atc cta cac ccg ttc cgc gcc aaa ctg cag agc acc cgg cgc cgg 540
Ala Ile Leu His Pro Phe Arg Ala Lys Leu Gln Ser Thr Arg Arg Arg
             145             150             155

gcc ctc agg atc ctc ggc atc gtc tgg ggc ttc tcc gtg ctc ttc tcc 588
Ala Leu Arg Ile Leu Gly Ile Val Trp Gly Phe Ser Val Leu Phe Ser
             160             165             170             175

ctg ccc aac acc agc atc cat ggc atc aag ttc cac tac ttc ccc aat 636

```

Leu Pro Asn Thr Ser Ile His Gly Ile Lys Phe His Tyr Phe Pro Asn	
180	185 190
ggg tcc ctg gtc cca ggt tcg gcc acc tgt acg gtc atc aag ccc atg	684
Gly Ser Leu Val Pro Gly Ser Ala Thr Cys Thr Val Ile Lys Pro Met	
195	200 205
tgg atc tac aat ttc atc atc cag gtc acc tcc ttc cta ttc tac ctc	732
Trp Ile Tyr Asn Phe Ile Ile Gln Val Thr Ser Phe Leu Phe Tyr Leu	
210	215 220
ctc ccc atg act gtc atc agt gtc ctc tac tac ctc atg gca ctc aga	780
Leu Pro Met Thr Val Ile Ser Val Leu Tyr Tyr Leu Met Ala Leu Arg	
225	230 235
cta aag aaa gac aaa tct ctt gag gca gat gaa ggg aat gca aat att	828
Leu Lys Lys Asp Lys Ser Leu Glu Ala Asp Glu Gly Asn Ala Asn Ile	
240	245 250 255
caa aga ccc tgc aga aaa tca gtc aac aag atg ctg tct ttg tgg agg	876
Gln Arg Pro Cys Arg Lys Ser Val Asn Lys Met Leu Ser Leu Trp Arg	
260	265 270
agt gga gtg aat ccc tgg ctg ctg tgt tca acc tcg tcc atg tgg tgt	924
Ser Gly Val Asn Pro Trp Leu Leu Cys Ser Thr Ser Ser Met Trp Cys	
275	280 285
cag gtg tct tct tct acc tgagctcagc tgtcaacccc attatctata	972
Gln Val Ser Ser Ser Thr	
290	
acctactgtc tcgccgcttc caggcagcat tccagaatgt gatctcttct ttccacaaac	1032
agtggcactc ccagcatgac ccacagttgc cacctgcccc gcggaacatc ttcctgacag	1092
aatgccactt tgtggagctg accgaagata taggtcccca attcccatgt cagtcaccca	1152
tgcaacaactc tcacctccca acagccctct ctagtgaaca gatgtcaaga acaaaactatc	1212
aaagcttcca ctttaacaaa acctgaattc tttcagagct gactctcctc tatgcctcaa	1272
aacttcagag aggaacatcc cataatgtat gccttctcat atgatattag agaggtagaa	1332
tggctcttac aactcatgta ccattgcta gttttttttt ttttaataaac gtgaaaactg	1392
agagttagat ctggtttcaa aaccaagac tgectgattt ttagttatct ttccactatc	1452
ctaactgcct catgccctt cactagttca tgccaagaac gtgactggaa aggcattggca	1512
cctatacctt gattaatttc cattaatgga aatggttcgt cctgagtcac ctacgttccg	1572
agtcaggctg tcaactcctac ta	1594
<210> 12	
<211> 293	

<212> PRT

<213> 人类

<400> 12

```

Met Glu Lys Leu Gln Asn Ala Ser Trp Ile Tyr Gln Gln Lys Leu Glu
  1           5           10           15

Asp Pro Phe Gln Lys His Leu Asn Ser Thr Glu Glu Tyr Leu Ala Phe
      20           25           30

Leu Cys Gly Pro Arg Arg Ser His Phe Phe Leu Pro Val Ser Val Val
      35           40           45

Tyr Val Pro Ile Phe Val Val Gly Val Ile Gly Asn Val Leu Val Cys
      50           55           60

Leu Val Ile Leu Gln His Gln Ala Met Lys Thr Pro Thr Asn Tyr Tyr
      65           70           75           80

Leu Phe Ser Leu Ala Val Ser Asp Leu Leu Val Leu Leu Leu Gly Met
      85           90           95

Pro Leu Glu Val Tyr Glu Met Trp Arg Asn Tyr Pro Phe Leu Phe Gly
      100          105          110

Pro Val Gly Cys Tyr Phe Lys Thr Ala Leu Phe Glu Thr Val Cys Phe
      115          120          125

Ala Ser Ile Leu Ser Ile Thr Thr Val Ser Val Glu Arg Tyr Val Ala
      130          135          140

Ile Leu His Pro Phe Arg Ala Lys Leu Gln Ser Thr Arg Arg Arg Ala
      145          150          155          160

Leu Arg Ile Leu Gly Ile Val Trp Gly Phe Ser Val Leu Phe Ser Leu
      165          170          175

Pro Asn Thr Ser Ile His Gly Ile Lys Phe His Tyr Phe Pro Asn Gly
      180          185          190

Ser Leu Val Pro Gly Ser Ala Thr Cys Thr Val Ile Lys Pro Met Trp
      195          200          205

Ile Tyr Asn Phe Ile Ile Gln Val Thr Ser Phe Leu Phe Tyr Leu Leu
      210          215          220

Pro Met Thr Val Ile Ser Val Leu Tyr Tyr Leu Met Ala Leu Arg Leu
      225          230          235          240

Lys Lys Asp Lys Ser Leu Glu Ala Asp Glu Gly Asn Ala Asn Ile Gln
      245          250          255

Arg Pro Cys Arg Lys Ser Val Asn Lys Met Leu Ser Leu Trp Arg Ser
      260          265          270

Gly Val Asn Pro Trp Leu Leu Cys Ser Thr Ser Ser Met Trp Cys Gln

```

275	280	285
Val Ser Ser Ser Thr 290		
<210> 13		
<211> 26		
<212> DNA		
<213> 人工序列		
<220>		
<223> 人工序列的描述: 简并引物		
<220>		
<221> 变异		
<222> (21)		
<223> A、C、G、或T		
<220>		
<221> 变异		
<222> (24)		
<223> A、C、G、或T		
<400> 13		
ctcatcttctcg cggtgggcrc ngyngg		26
<210> 14		
<211> 31		
<212> DNA		
<213> 人工序列		
<220>		
<223> 人工序列的描述: 简并引物		
<220>		
<221> 变异		
<222> (22)		
<223> C或肌昔		
<220>		
<221> 变异		
<222> (25)		
<223> A、C、G、或T		
<220>		
<221> 变异		
<222> (28)		
<223> A、C、G、或T		
<400> 14		
ggccaggcag cgctccgcgc tnarncyngc d		31

	<210> 15	
	<211> 22	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 人工序列的描述: 简并引物	
	<400> 15	
	gaartartag ccrcgrcagc cw	22
	<210> 16	
	<211> 27	
	<212> DNA	
	<213>人工序列	
	<220>	
	<223>人工序列描述: 引物	
	<400> 16	
	ccatccta at acgactcact atagggc	27
	<210> 17	
	<211> 23	
	<212> DNA	
	<213>人工序列	
	<220>	
	<223>人工序列描述: 引物	
43>	<400> 17	
	actcactata gggctcgagc ggc	23
	<210> 18	
	<211> 27	
	<212> DNA	
	<213>人工序列	
	<220>	
	<223>人工序列描述: 引物	
	<400> 18	
	ggatcccaaa taagaaaggg tagttgc	27
	<210> 19	
	<211> 29	
	<212> DNA	
	<213>人工序列	

<220>

<223>人工序列描述: 引物

<400> 19

aaagggtagt tgcgccacat ctcatagac

29

<210> 20

<211> 29

<212> DNA

<213>人工序列

<220>

<223>人工序列描述: 引物

<400> 20

aggtctatga gatgtggcgc aactaccct

29

<210> 21

<211> 30

<212> DNA

<213>人工序列

<220>

<223>人工序列描述: 引物

<400> 21

atgtggcgca actacccttt cttatttggg

30

<210> 22

<211> 26

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述: 简并引物

<400> 22

cggaagttgg cggacacgrv rttrta

26

<210> 23

<211> 27

<212> DNA

<213>人工序列

<220>

<223>人工序列描述: 引物

<400> 23

gctcagcttg aaacagagcc tcgtacc

27

<210> 24	
<211> 27	
<212> DNA	
<213>人工序列	
<220>	
<223>人工序列描述: 引物	
<400> 24	
ccatgtggat ctacaatttc atcatcc	27
<210> 25	
<211> 29	
<212> DNA	
<213>人工序列	
<220>	
<223>人工序列描述: 引物	
<400> 25	
aagacaaatc tcttgaggca gatgaaggg	29
<210> 26	
<211> 30	
<212> DNA	
<213>人工序列	
<220>	
<223>人工序列描述: 引物	
<400> 26	
gatgctgttt gtcttggtct tagtgtttgc	30
<210> 27	
<211> 29	
<212> DNA	
<213>人工序列	
<220>	
<223>人工序列描述: 引物	
<400> 27	
ggatgatgaa attgtagatc cacatgggc	29
<210> 28	
<211> 25	
<212> DNA	
<213>人工序列	
<220>	
<223>人工序列描述: 引物	

<400> 28 tgtggagaag tctctcaaag tgtgg	25
<210> 29 <211> 29 <212> DNA <213>人工序列	
<220> <223>人工序列描述: 引物	
<400> 29 tagtaggagt gacagcctga ctcggaacg	29
<210> 30 <211> 30 <212> DNA <213>人工序列	
<220> <223>人工序列描述: 引物	
<400> 30 aacgtagatg actcaggacg aaccatttcc	30
<210> 31 <211> 22 <212> DNA <213>人工序列	
<220> <223>人工序列描述: 引物	
<400> 31 tcgtaccagg ggaggctcag gc	22
<210> 32 <211> 23 <212> DNA <213>人工序列	
<220> <223>人工序列描述: 引物	
<400> 32 cctcttcagc ctggcggtct ctg	23
<210> 33 <211> 22 <212> DNA <213>人工序列	

<220>

<223>人工序列描述: 引物

<400> 33

ggaggcgaag cacacggtct ca

22

<210> 34

<211> 34

<212> DNA

<213>人工序列

<220>

<223>人工序列描述: 引物

<220>

<221> misc_binding

<222> (1)

<223> 用6-羧基荧光素标记

<220>

<221> misc_binding

<222> (34)

<223> 用N, N, N', N'-四甲基-6-羧基罗丹明标记

<400> 34

agatgtggcg caactaccct ttcttggtcg ggcc

34

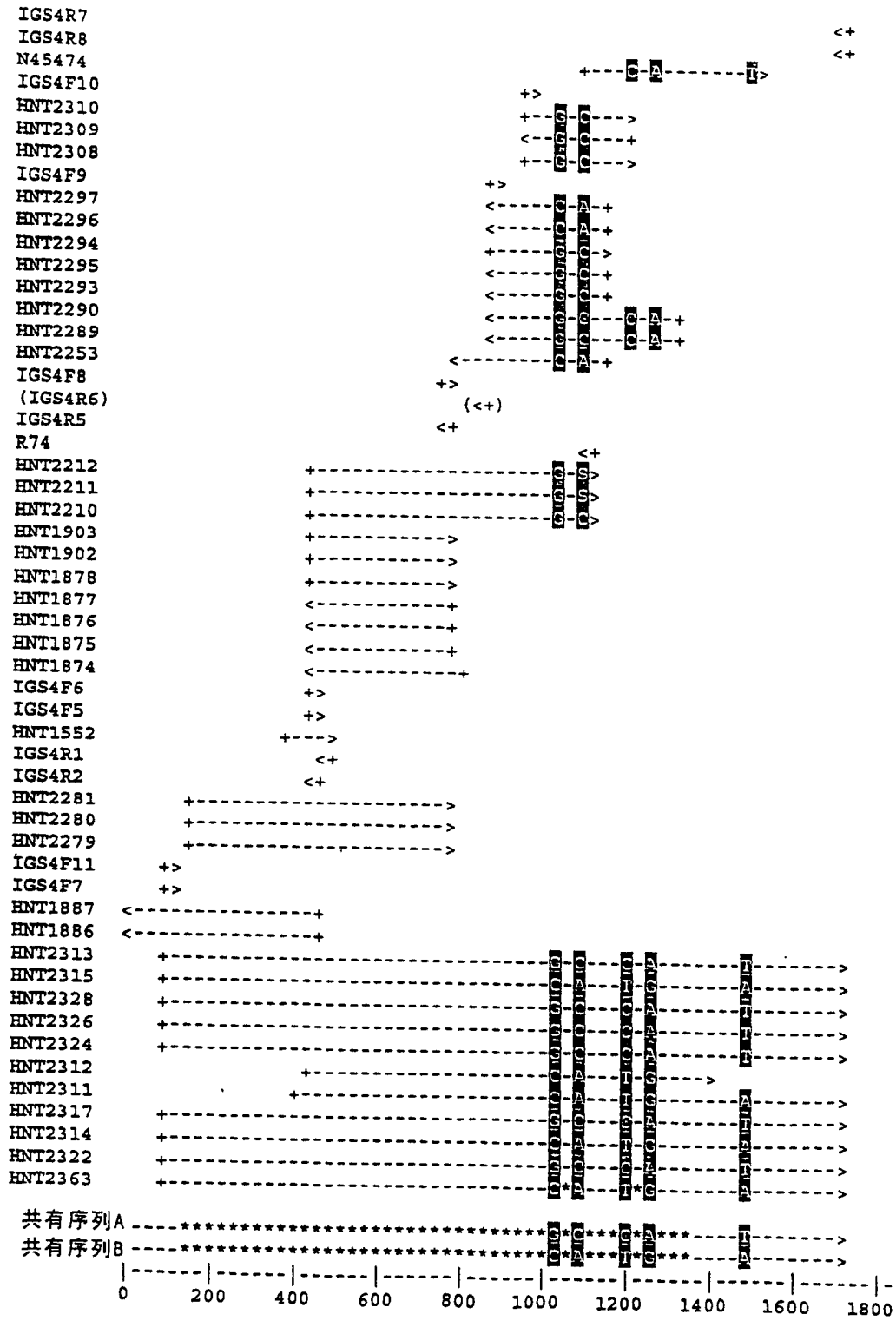


图 1

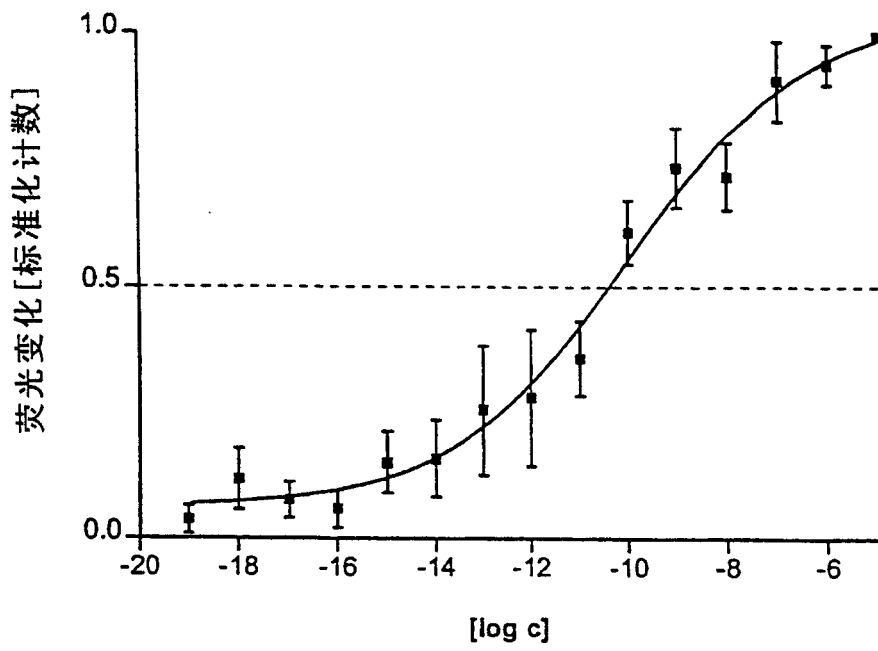


图 3a

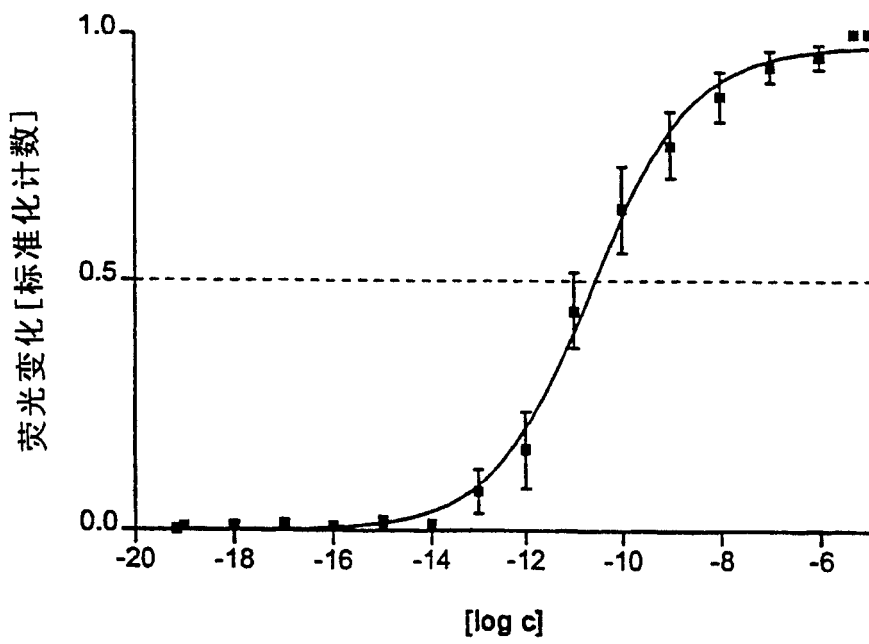


图 3b

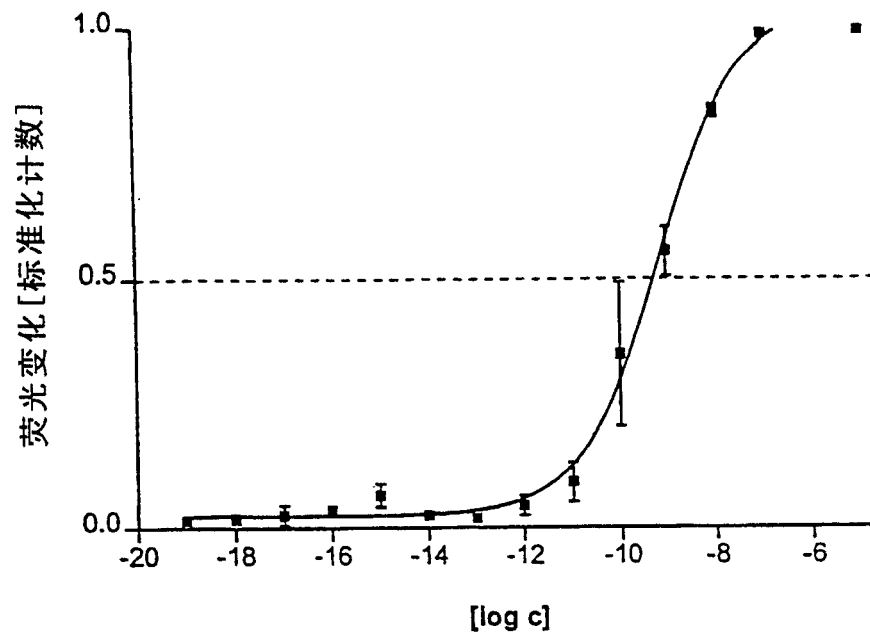


图 3c

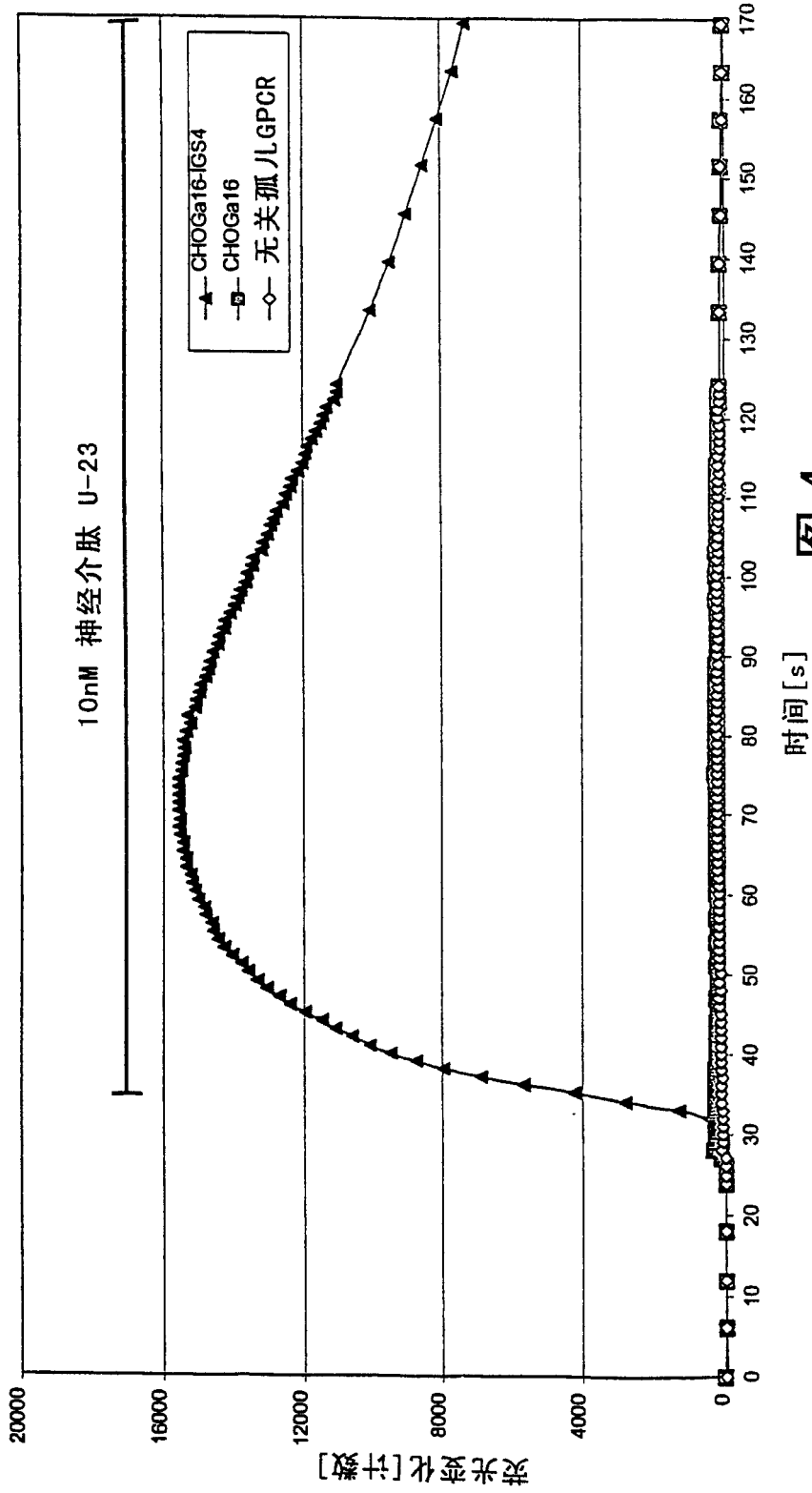
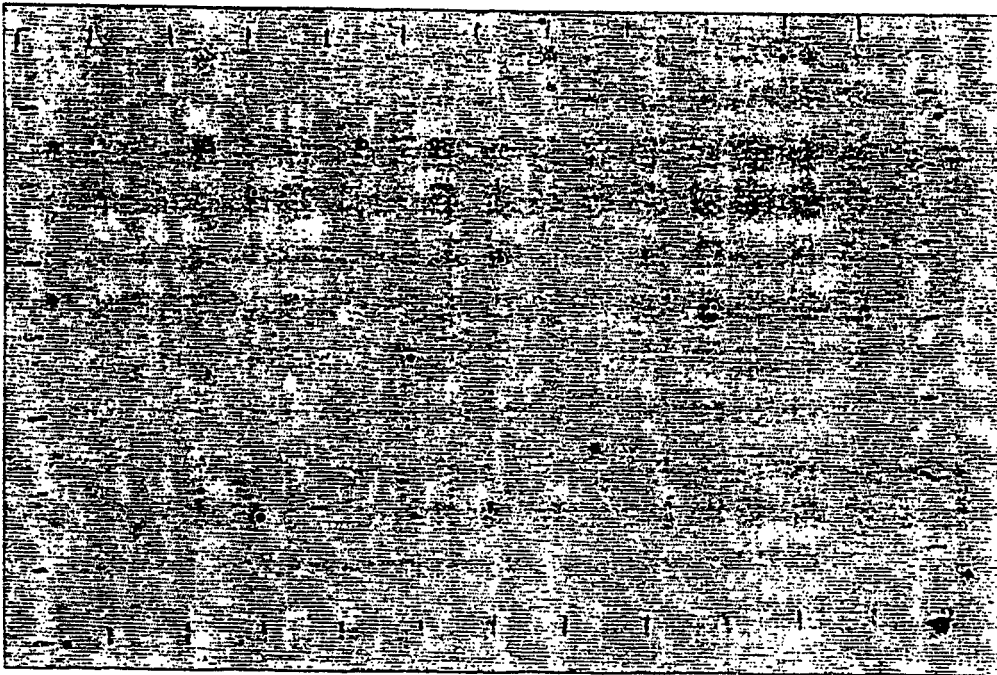


图 4

图 5

1. A1

A12



2. H1

H12

使用人IGS4探针的人多组织表达阵列

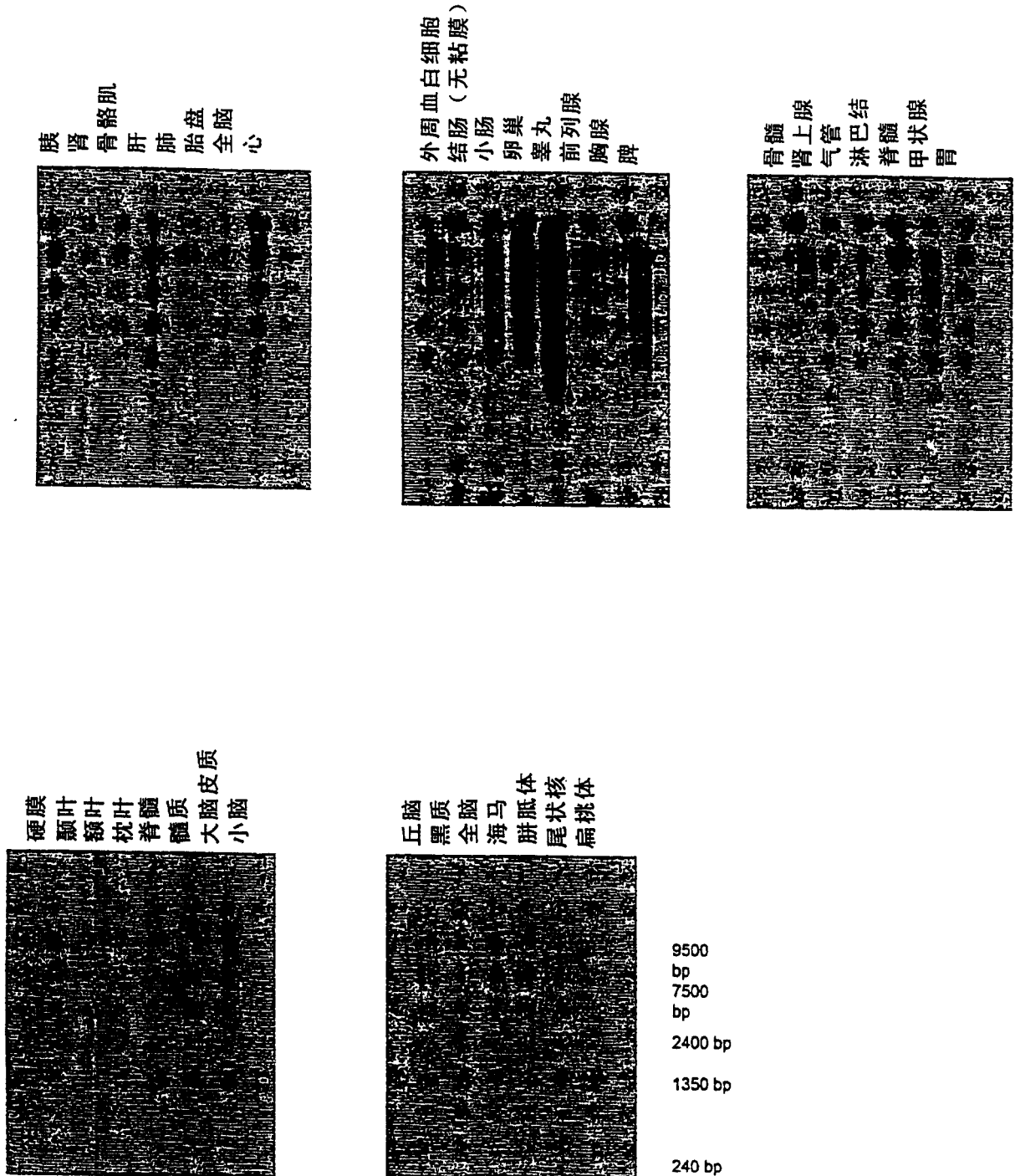
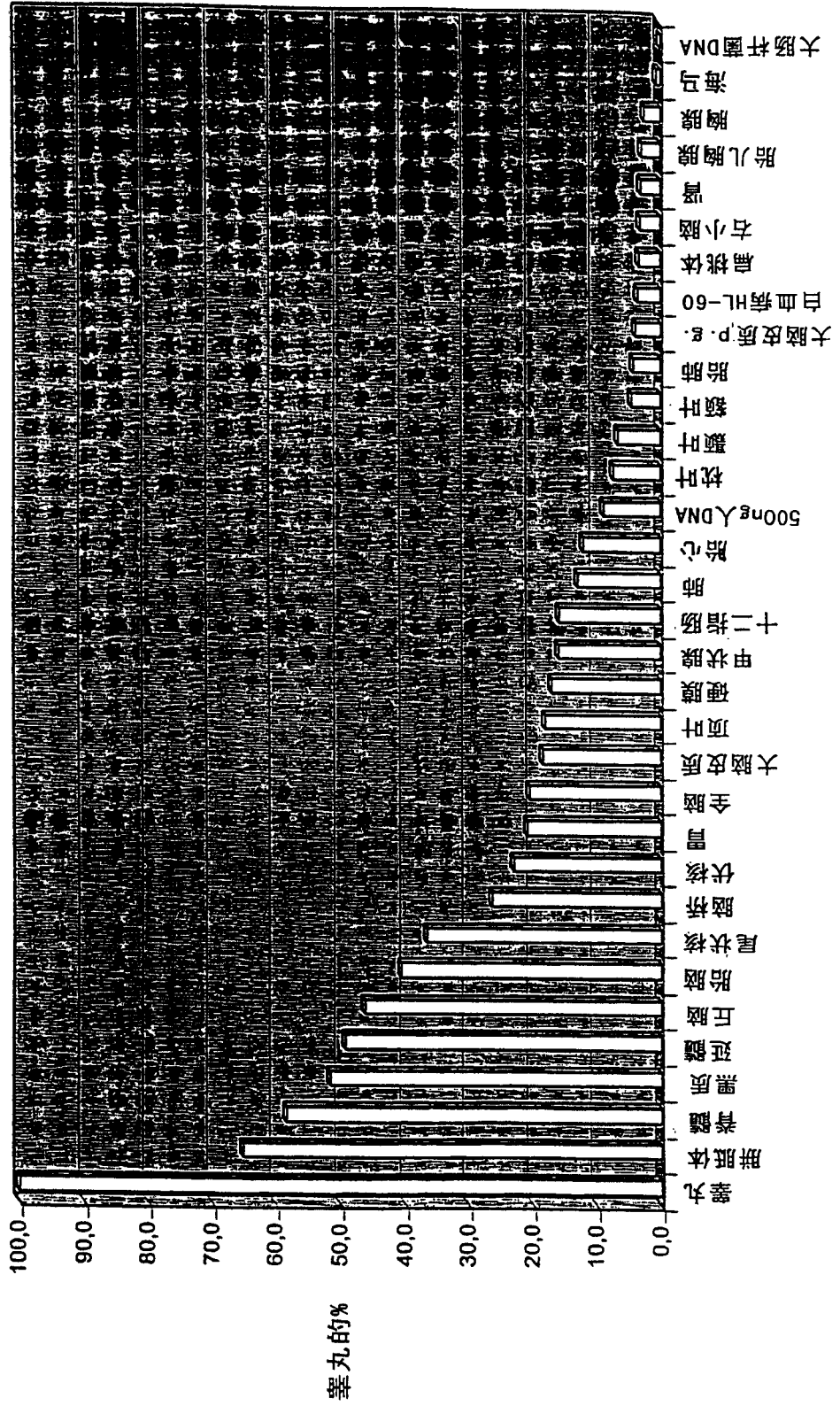


图 6

图7

hu-IGS4表达分析 (MTE印迹)



IGS4 QPCR

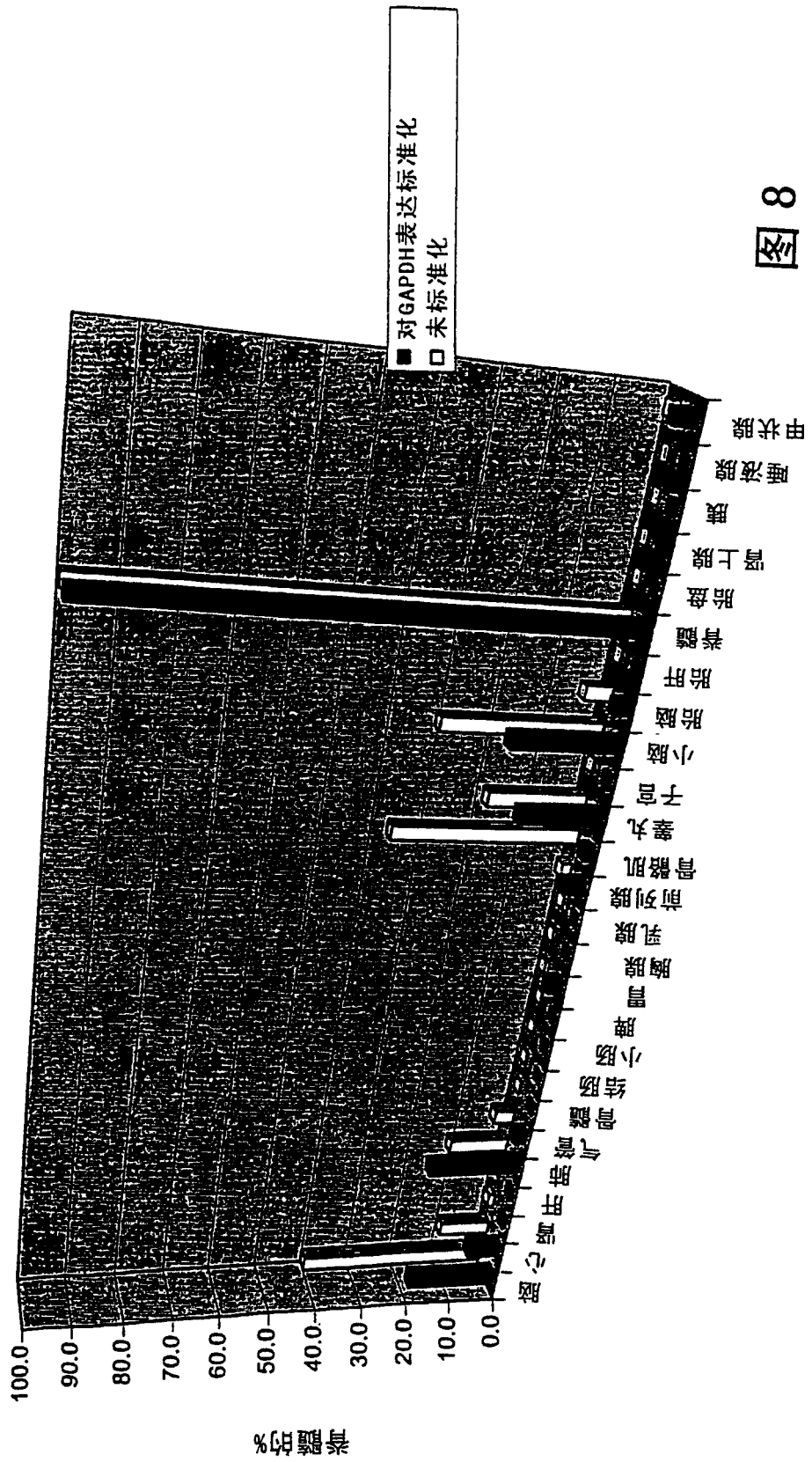


图 8

专利名称(译)	新的人G蛋白偶联受体		
公开(公告)号	CN1824777A	公开(公告)日	2006-08-30
申请号	CN200510092639.8	申请日	2000-09-25
[标]申请(专利权)人(译)	索尔瓦药物有限公司		
申请(专利权)人(译)	索尔瓦药物有限公司		
当前申请(专利权)人(译)	索尔瓦药物有限公司		
[标]发明人	W德里尔斯尼德尔 C比戈尔 C劳肯 G尼斯 J维尼玛		
发明人	W·德里尔斯尼德尔 C·比戈尔 C·劳肯 G·尼斯 J·维尼玛		
IPC分类号	C12N15/12 C07K14/435 C12N15/63 C12N15/81 C12N15/85 C12P21/02 C07K16/18 C07K16/28 A61K38/17 A61P25/00 A61P9/00 A61P1/00 G01N33/53		
代理人(译)	刘晓东		
优先权	1013140 1999-09-24 NL 1999203140 1999-09-24 EP 2000202683 2000-07-28 EP 60/222047 2000-07-31 US		
外部链接	Espacenet SIPO		

摘要(译)

本发明的多核苷酸和多肽涉及G蛋白偶联受体家族，称为IGS4家族。本发明还涉及抑制或激活这些多核苷酸和多肽的作用，包含所述多核苷酸的载体、包含这种载体的宿主细胞、和IGS4基因过度表达、错误表达、表达不足、或受遏制(敲除动物)的转基因动物。本发明还涉及用于筛选能够作为IGS4家族的激动剂或拮抗剂的化合物的方法，以及这些IGS4多肽、多核苷酸、激动剂、或拮抗剂的用途。本发明的优选用途涉及神经系统、胃肠系统、心血管系统、骨骼肌、和/或甲状腺的紊乱，和/或肺病、免疫学疾病、和泌尿生殖系统紊乱。本发明还涉及IGS4多肽的相关配体 - 称为神经介肽U的神经肽 - 的鉴定。

