



(12) 发明专利

(10) 授权公告号 CN 1525981 B

(45) 授权公告日 2012.07.04

(21) 申请号 02803431.7

(22) 申请日 2002.01.03

(30) 优先权数据

60/259,227 2001.01.03 US

60/284,547 2001.04.19 US

(85) PCT申请进入国家阶段日

2003.07.03

(86) PCT申请的申请数据

PCT/US2002/000198 2002.01.03

(87) PCT申请的公布数据

W02002/064631 EN 2002.08.22

(73) 专利权人 塞诺米克斯公司

地址 美国加利福尼亚

(72) 发明人 J·E·阿德勒 李晓东

L·斯塔谢夫斯基 S·奥康奈尔

S·佐祖利亚

(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 11038

代理人 程泳

(51) Int. Cl.

C07K 14/705(2006.01)

C12N 15/12(2006.01)

C12N 15/62(2006.01)

C07K 16/18(2006.01)

C12Q 1/68(2006.01)

C12N 5/10(2006.01)

G01N 33/50(2006.01)

(56) 对比文件

W0 9508627 A, 1995.03.30, 全文.

审查员 吴永庆

权利要求书 2 页 说明书 53 页

序列表 32 页 附图 7 页

(54) 发明名称

T1R 味觉受体及其编码基因

(57) 摘要

描述了最近鉴定的哺乳动物味觉细胞特异性 G 蛋白偶联受体以及编码该受体的基因和 cDNA。特别地,描述了在味觉信号传导中起作用的 T1R G 蛋白偶联受体以及其编码基因和 cDNA,同时还描述了用于分离该基因和用于分离与表达该受体的方法。也描述了用于在哺乳动物中表示特殊味觉刺激物的味觉感受的方法,以及用于产生在哺乳动物中引起预定味觉感受的新分子或分子组合的方法,和模仿一种或多种味觉的方法。此外,也公开了用于在哺乳动物中刺激或阻断味觉感受的方法。

1. 一种分离的核酸序列,其编码 SEQ ID NO :21 中的人类 T1R2 多肽,其中当所述多肽与 SEQ ID NO :17 中的人类 T1R1 多肽或 SEQ IDNO :4 中的人类 T1R 3 多肽一起表达时产生功能性味觉受体。

2. 权利要求 1 的分离的核酸序列,其由 SEQ ID NO :20 中的核酸序列组成。

3. 权利要求 1 的分离的核酸序列,其为 cDNA。

4. 权利要求 1 的分离的核酸序列,其为 mRNA。

5. 权利要求 1 的分离的核酸序列,其为基因组序列。

6. 权利要求 1 的分离的核酸序列,其有效地连接于某序列上,所述某序列促进所述味觉受体多肽在含有所述分离的核酸序列的细胞表面表达。

7. 权利要求 6 的分离的核酸序列,其中所述促进表面表达的序列是哺乳动物视紫红质多肽。

8. 权利要求 1 的分离的核酸序列,其有效地连接于编码可探测多肽的核酸序列上。

9. 权利要求 8 的分离的核酸序列,其中所述可探测多肽是绿色荧光多肽。

10. 权利要求 1 的分离的核酸序列,其有效地连接于至少一个核酸序列上,所述至少一个核酸序列在宿主细胞中调节由所述分离的核酸序列编码的多肽的转录和 / 或翻译。

11. 权利要求 10 的分离的核酸序列,其中所述调节转录的序列是可调控的启动子或组成型启动子。

12. 一种表达载体,其包含权利要求 1-11 中任一项的分离的核酸序列。

13. 权利要求 12 的表达载体,其选自质粒、噬菌粒、哺乳动物病毒、反转录病毒、噬菌体载体和能够整合入宿主细胞基因组中的线状或环状质粒。

14. 一种宿主细胞,其包含权利要求 1 或 2 的分离的核酸序列或权利要求 12 或 13 的表达载体。

15. 权利要求 14 的宿主细胞,其选自原核细胞和真核细胞。

16. 权利要求 15 的宿主细胞,其中所述真核细胞选自哺乳动物细胞、酵母细胞、昆虫细胞和两栖动物细胞。

17. 权利要求 16 的宿主细胞,其中所述哺乳动物细胞选自 CHO、HeLa、HEK-293、培养的细胞和外植块。

18. 权利要求 14-17 中任一项的宿主细胞,其表达与由所述宿主细胞表达的味觉受体多肽功能性相关的 G 蛋白。

19. 权利要求 18 的宿主细胞,其中所述 G 蛋白是 G α 15、G α 16、其它的混杂 G 蛋白或嵌合 G 蛋白。

20. 权利要求 17 的宿主细胞,其为 HEK-293 细胞。

21. 权利要求 14-17 中任一项的宿主细胞,其进一步表达其它的 G 蛋白偶联受体。

22. 一种人类味觉受体多肽,其由权利要求 1-11 中任一项的分离的核酸序列编码。

23. 一种筛选味觉调节化合物的方法,包括:使权利要求 22 的人类味觉受体多肽与推定的味觉调节化合物接触,并确定所述化合物是否特异性地结合所述味觉受体多肽和 / 或影响所述味觉受体多肽的功能活性。

24. 权利要求 23 的方法,其为一种结合测定法。

25. 权利要求 24 的方法,其为竞争性结合测定法。

26. 权利要求 24 的方法,其中所述味觉受体多肽表达于细胞上。
27. 权利要求 24 的方法,其中所述味觉受体多肽表达于细胞膜上。
28. 权利要求 24 的方法,其中所述味觉受体多肽固定于基质上。
29. 权利要求 24 的方法,其包括使用可探测标记。
30. 权利要求 29 的方法,其中所述标记是荧光团、酶或放射性部分。
31. 权利要求 23 的方法,其是一种功能测定法,该功能测定法检测推定的味觉调节化合物对所述味觉受体多肽的功能特性的影响。
32. 权利要求 31 的方法,其作为一种基于细胞的测定法。
33. 权利要求 31 的方法,其测量所述推定的味觉调节物对选自下列的参数的影响:递质释放、激素释放、转录、细胞代谢和胞内第二信使水平。
34. 权利要求 32 的方法,其检测选自 IP₃、Ca⁺⁺、Na⁺、cGMP 和 cAMP 的第二信使的变化。
35. 权利要求 34 的方法,其检测胞内 Na⁺ 或 Ca⁺⁺ 离子。
36. 权利要求 35 的方法,其使用离子敏感性指示剂或电压探针。
37. 权利要求 35 的方法,其使用离子敏感性或膜电压荧光指示剂。
38. 权利要求 32 的方法,其检测所述推定的味觉调节物对磷脂酶途径的影响。
39. 权利要求 34 的方法,其为免疫测定法。
40. 权利要求 32 的方法,其检测磷脂酰水解程度。
41. 权利要求 32 的方法,其测量所述推定的味觉调节物对胞内信号传导途径或信号传导蛋白的影响。
42. 权利要求 32 的方法,其为荧光偏振测定法。
43. 权利要求 32 的方法,其检测所述推定的味觉调节物对受体多肽/G 蛋白相互作用的影响。
44. 权利要求 43 的方法,其检测 G 蛋白结合或释放。
45. 权利要求 44 的方法,其检测 GTP 向 GDP 的转化或 GDP 向 GTP 的转化。
46. 权利要求 45 的方法,其使用含有经放射性标记的 GTP 的化合物。
47. 权利要求 23-46 中任一项的方法,其中所述推定的味觉调节化合物包含于化合物文库中。
48. 权利要求 47 的方法,其中所述文库是组合化学文库。
49. 权利要求 23-46 中任一项的方法,其作为一种高通量筛选测定法。

T1R 味觉受体及其编码基因

[0001] 相关申请

[0002] 本申请要求 2001 年 1 月 3 日提交的 U. S. Serial No. 60/259, 227 和 2001 年 4 月 19 日提交的 U. S. Serial No. 60/284, 547 的优先权, 此处整体引用两者作为参考文献。

[0003] 发明背景

发明领域

[0004] 本发明涉及最近鉴定的哺乳动物化学感受性 G 蛋白偶联受体、这种受体的家族和编码该受体的基因与 cDNA。更具体而言地, 本发明涉及最近鉴定的在味觉信号传导中起作用的哺乳动物化学感受性 G 蛋白偶联受体、这种受体的家族、编码该受体的基因与 cDNA 以及在味觉调节子的分析和发现中使用这种受体、基因与 cDNA 的方法。本发明特别提供了编码新的在下文标识为 T1R2 的人味觉受体的 DNA 序列和相应的受体多肽。

[0005] 相关领域描述

[0006] 味觉系统提供了有关外部世界化学组合物的感觉信息。认为哺乳动物具有至少 5 种基本的味觉形式: 甜、苦、酸、咸和 umami。参见如 Kawamura 等人, Umami 介绍: 一种基本的味觉 (Introduction to U mami :A Basic Taste) (1987); Kinnamon 等人, Ann. Rev. Physiol., 54 :715-31 (1992); Lindemann, Physiol. Rev., 76 :718-66 (1996); Stewart 等人, Am. J. Physiol., 272 :1-26 (1997)。每一种味觉形式被认为是由舌表面上的味觉受体细胞所表达的多种受体或一种独特的蛋白质受体来介导的 (Lindemann, Physiol. Rev., 76 :718-716 (1996))。识别苦、甜和 umami 味觉刺激物的味觉受体属于 G 蛋白偶联受体 (GPCR) 超家族 (Hoon 等人, Cell 96 :451 (1999); Adler 等人, Cell 100 :693 (2000))。(认为其他味觉形式通过离子通道介导。)

[0007] G 蛋白偶联受体介导了许多其他的生理学功能, 如内分泌功能、外分泌功能、心率、脂解作用和糖类代谢。对于一些这种受体的生物化学分析和分子克隆已揭示出许多有关这些受体功能的原理。例如, 美国专利 No. 5, 691, 188 描述了当配体结合到 GPCR 上时, 受体怎样经过构象变化从而导致异三聚体化的 G 蛋白的激活, 其中 G 蛋白的激活通过促进 Ga 亚基表面所结合的 GDP 被 GTP 的置换以及随后 Ga 亚基从 Gb 与 Gg 亚基的解离来进行。游离的 Ga 亚基和 Gbg 复合物激活了多种信号转导通路的下游元件。

[0008] 当前已知许多人和其他真核生物化学感受性受体的完整或部分序列。参见如 Pilpel, Y. 和 Lancet, D., Protein Science, 8 :969-977 (1999); Mombaerts, P., Annu. Rev. Neurosci., 22 :487-50 (1999)。也参见 EP0867508A2、US 5874243、WO 92/17585、WO 95/18140、WO97/17444、WO 99/67282。由于配体-受体相互作用的复杂性, 更具体而言地为味觉刺激物-受体相互作用的复杂性, 关于配体-受体识别的信息是缺乏的。

[0009] 作为甜和 umami 味觉受体的 GPCR 的鉴定和表征可允许发现新味觉刺激物的新方法。例如, 受体的可用性能够允许对于受体调节子的筛选。这种化合物能够调节味觉, 并可在食品工业中用于改善多种消费产品的味道; 例如, 通过新人工甜味剂的开发来改善低卡路里饮料的可口性。

[0010] 部分地,本发明提出了对于更好地理解化学感受性受体和化学刺激物之间相互作用的需要。其中,本发明(除了其它)也提供了新的化学感受性受体和使用这种受体的方法,以及编码这种受体的基因与 cDNA,从而鉴定可用于调节化学感受性转导如味觉的分子。

[0011] 发明概述

[0012] 本发明涉及 G 蛋白偶联受体的新家族和编码该受体的基因与 cDNA。认为受体主要地与甜味转导有关,但是也可以参与转导来自其他味觉形式的信号。

[0013] 本发明提供了在哺乳动物中(包括人)表示味觉感受和/或预测味觉感受的方法。优选地,这种方法可通过使用此处所公开的受体和编码该受体的基因来进行。

[0014] 所以,提供此处称为 T1Rs 的在味觉中起作用的哺乳动物 G 蛋白偶联受体新家族是本发明的一个目的。提供保留味觉刺激物结合活性的这种 T1Rs 的片段和变体为本发明的另一个目的。

[0015] 提供编码这种 T1Rs 或其片段或变体的核酸序列或分子是本发明的另外一个目的。

[0016] 提供包括编码这种 T1Rs 或其片段或变体的核酸序列的表达载体为本发明的另外一个目的,该核酸序列可操作地连接到至少一个调节序列上,如启动子、增强子或其他参与正或负的基因转录和/或翻译的序列。

[0017] 提供功能性地表达至少一种这种 T1Rs 或其片段或变体的人或非人细胞为本发明的另外一个目的。

[0018] 提供包括至少一种这种 T1Rs 的至少一个片段的 T1R 融合蛋白质或多肽为本发明的另外一个目的。

[0019] 提供分离的编码 T1R 多肽的核酸分子为本发明的另一个目的,该核酸分子含有与选自 SEQ ID NOS :1、2、3、9、11、13、15、16、20 的核酸序列具有至少 50%,优选的为 75%、85%、90%、95%、96%、97%、98%或 99%的一致性的核酸序列及其保守修饰的变体。

[0020] 提供分离的包含编码多肽的核酸序列的核酸分子为本发明的进一步的目的,该多肽具有与选自 SEQ ID NOS :4、10、12、14、17、21 的氨基酸序列具有至少 35~50%,优选的为 60%、75%、85%、90%、95%、96%、97%、98%或 99%的一致性的氨基酸序列及其保守修饰的变体,其中片段长度为至少 20 个,优选的为 40、60、80、100、150、200 或 250 个氨基酸。可选地,片段可以为与抗 T1R 抗体结合的抗原片段。

[0021] 提供分离的含有该片段变体的多肽为本发明另外的目的,其中有至多 10 个,优选的为 5、4、3、2 或 1 个氨基酸残基发生变异。

[0022] 提供这种 T1Rs 的激动剂或拮抗剂,或其片段或变体为本发明的另外一个目的。

[0023] 提供在哺乳动物包括人中表示味觉感受和/或预测味觉感受的方法为本发明的另外一个目的。优选地,这种方法可通过使用此处所公开的 T1Rs 或其片段或变体和编码这种 T1Rs 的基因或其片段或变体来进行。

[0024] 提供在哺乳动物中引起预定味觉感受的新的分子或分子组合为本发明的另外一个目的。这种分子或组合物可通过下述方法产生:测定已知分子或分子组合在哺乳动物中的味觉感受值;测定一种或多种未知分子或分子组合在哺乳动物中的味觉感受值;将一种或多种未知组合物在哺乳动物中的味觉感受值与一种或多种已知组合物在哺乳动物中的味觉感受值进行比较;选择出在哺乳动物中引起预定味觉感受的分子或分子组合;以及合

并两种或更多种未知分子或分子组合以形成在哺乳动物中引起预定味觉感受的分子或分子组合。合并步骤产生在哺乳动物中引起预定味觉感受的单一分子或分子组合。

[0025] 提供筛选一种或多种化合物以确定哺乳动物可探测的味觉是否存在的方法为本发明另外的目的,该方法包括:将该一种或多种化合物与至少一种公开的 T1Rs 或其片段或变体接触的步骤,其中哺乳动物优选地为人。

[0026] 提供用于模拟味觉的方法为本发明的另一个目的,该方法包括下列步骤:对于此处所公开的大量 T1Rs 或其片段或变体的每一个,优选的为人 T1Rs,确定 T1R 与味觉刺激物相互作用的程度;以及以一定量组合大量的化合物,其中每一个具有先前所确定的与一种或多种 T1Rs 的相互作用,它们总和在一起提供了模拟味觉特征的受体-刺激特征。味觉刺激物与 T1R 的相互作用可以使用任何这里所描述的结合或受体测定来确定。然后,可将大多数化合物组合形成混和物。如果需要,大多数化合物中的一种或多种可以共价组合。组合化合物基本上刺激至少 50%、60%、70%、75%、80% 或 90% 或者所有的基本上可被味觉刺激物刺激的受体。

[0027] 在本发明的另外一个方面中,提供了一种方法,其中将大量的标准化化合物进行抗大量 T1Rs 或其片段或变体的检验,从而确定每一个 T1Rs 与每一个标准化化合物相互作用的程度,因此对于每一个标准化化合物产生一个受体刺激特性。然后可将这些受体刺激特征储存在位于数据存储介质上的相关数据库中。该方法可以进一步包括提供味觉的想要的受体刺激特征;将想要的受体刺激特征与相关数据库进行比较;并确定最紧密地与想要的受体刺激特征相匹配的一种或多种标准化化合物的组合。该方法可以进一步包含在一种或多种所确定的组合中来组合标准化化合物以模拟味觉。

[0028] 提供用于在哺乳动物中表示特殊味觉刺激物的味觉感受的方法为本发明的进一步的目的,该方法包括下列步骤:提供所述脊椎动物 n 个 T1R 中每一个的定量刺激的代表值 $X_1 \sim X_n$, 其中 n 大于或等于 2;从该值中产生味觉的定量表示 (representation)。T1Rs 可以是这里所公开的味觉受体或者其片段或变体,所述表示可以构成 n 维空间中的点或组 (volume),可以构成图表或谱,还可以构成定量表示的矩阵。同样的,提供的步骤可以包含将大量重组生产的 T1Rs 或其片段或变体与检验组合物接触,并定量测量该组合物与该受体之间的相互作用。

[0029] 提供用于在哺乳动物中预测由一种或多种分子或分子组合生成的味觉感受的方法为本发明的另外一个目的,该分子或分子联合物在哺乳动物中产生未知的味觉感受,该方法包括下列步骤:提供该脊椎动物 n 个 T1Rs 中每一个的定量刺激的代表值 $X_1 \sim X_n$, 其中 n 大于或等于 2;对于一种或多种分子或分子组合,在哺乳动物中产生已知味觉;并从该值中产生出对于在哺乳动物中产生已知味觉感受的一种或多种分子或分子组合在哺乳动物中味觉感受的定量表示;提供该脊椎动物 n 个 T1Rs 中每一个的定量刺激的代表值 $X_1 \sim X_n$, 其中 n 大于或等于 2;对于一种或多种分子或分子组合,在哺乳动物中产生未知味觉感受;并从该值中产生出对于在哺乳动物中产生未知味觉感受的一种或多种分子或分子组合在哺乳动物中味觉感受的定量表示;然后预测哺乳动物中由在哺乳动物中产生未知味觉感受的一种或多种分子或分子组合产生的味觉感受,该预测是通过将对于在哺乳动物中产生未知味觉感受的一种或多种分子或分子组合的哺乳动物中味觉感受的定量表示与对于在哺乳动物中产生已知味觉感受的一种或多种分子或分子组合的哺乳动物中味觉感受的定量

表示进行比较来实现的。在该方法中所使用的 T1Rs 可以包括这里所公开的味觉受体或者其片段或变体。

[0030] 附图详述

[0031] 图 1a-1c 表示对于人 T1R2/T1R3 受体的功能数据。在图 1a 中显示了稳定表达 G α 15 的 HEK 细胞对于多种蔗糖浓度的细胞内钙反应,该 HEK 细胞用人 T1R2、T1R3 和 T1R2/T1R3 进行了瞬时转染。每一组相当于大约 1000 个汇合的、转染的和装载了钙染料的细胞。图 1b 中显示了由甜味觉抑制剂 gurmarin 对 T1R2/T1R3 活性的抑制。图 1c 中显示了对于 4 种甜味剂的 T1R2/T1R3 剂量反应和相关的心理物理学探测阈值 (X 轴圆圈)。

[0032] 图 2 表示对于大鼠 T1R2/T1R3 受体的功能数据。显示了对于 350mM 蔗糖、25mM 色氨酸、15mM 阿斯巴甜和 0.05% 应乐果甜蛋白的人 T1R2/T1R3 和大鼠 T1R2/T1R3 (以及混和的大鼠/人受体) 反应。大鼠 T1R2/T1R3 不对阿斯巴甜或应乐果甜蛋白有反应,这些对于啮齿动物是不可口的。

[0033] 图 3a-3c 表示对于人 T1R2/T1R3 受体的功能数据。图 3a 中显示了稳定表达 G α 15 的 HEK 细胞对于多种浓度 L-谷氨酸的细胞内钙响应,该 HEK 细胞用人 T1R1、T1R3 和 T1R1/T1R3 进行了瞬时转染。图 3b 中显示了由 IMP 所引起的 T1R1/T1R3 反应的增强。图 3c 中显示了对于 L-谷氨酸以及 L-谷氨酸加 0.2mM IMP 的 T1R1/T1R3 剂量反应,以及相关的心理物理学探测阈值 (X 轴圆圈)。

[0034] 图 4a-4b 表示免疫荧光和 FACS 实验,实验证明了将 PDZIP 肽 (SEQ ID No :1) 融合到人 T1R2 上可增强其在 HEK 细胞表面的表达。

[0035] 图 5 表示对于稳定表达 G α 15 和人 T1R1/T1R3 的细胞系的自动荧光成像数据。测定了在 0.5mM IMP 存在时的 L-谷氨酸剂量反应。

[0036] 图 6 表示对于稳定表达 G α 15 和人 T1R2/T1R3 的细胞系的自动荧光成像数据。显示了稳定细胞系的蔗糖、D-色氨酸、糖精和阿斯巴甜的剂量反应。

[0037] 发明详述

[0038] 本发明从而提供了分离的编码味觉细胞特异性 G 蛋白偶联受体 (“GPCR”) 的核酸分子及其编码的多肽。这些核酸分子及其编码的多肽是味觉细胞特异性 GPCRs 的 T1R 家族的成员。味觉细胞特异性 GPCRs 的 T1R 家族的成员在 Hoon 等人,Cell 96 :541-551 (1999)、WO 00/06592 和 WO 00/06593 中进行了鉴定,这些文献在此处整体引用作为参考。

[0039] 更具体而言地,本发明提供了编码味觉细胞特异性 GPCRs 新家族的核酸。这些核酸及其编码的受体被认为是味觉细胞特异性 GPCRs 的 “T1R” 家族的成员。在本发明特别的实施方案中,T1R 家族成员包括人 T1R1、T1R2 和 T1R3。根据上面的描述,不同的 T1R 组合物可能介导了甜和 umami 味觉。此外,人们确信 T1R 家族成员可以与其他 T1R 家族成员、其他味觉细胞特异性 GPCRs 或其组合协同起作用,从而影响化学感受性味觉转导。例如,人们确信 T1R1 和 T1R3 可能在相同的味觉受体细胞类型中共表达,这两种受体可以在物理上相互作用从而形成异二聚体化的味觉受体。可选择的,T1R1 和 T1R3 两者可以独立地结合到同类型的配体上,它们的组合结合可以导致特定的感觉到的味觉。

[0040] 由于核酸在味觉细胞中特异表达,因此这些核酸提供了用于鉴定味觉细胞的有价值的探针。例如,对于 T1R 多肽和蛋白质的探针可用于鉴定存在于叶状、轮廓状和蕈状乳头之中的味觉细胞,以及存在于味觉带 (geschmackstreifen)、口腔、胃肠上皮和会厌之中的

味觉细胞。它们也可作为用于产生味觉局部解剖图的工具,该图阐明了舌的味觉细胞与通向大脑味觉中心的味觉感觉神经元之间的相互关系。特别地,探测 T1Rs 的方法可用于鉴定对于甜味觉刺激物或其他特定形式味觉刺激物敏感的味觉细胞。此外,核酸及其编码的蛋白质可用作探针来剖析味觉诱导的行为。同样的,编码人 T1Rs 的基因的染色体定位可用于鉴定由 T1R 家族成员所引起的和与之有关的疾病、突变和性状。

[0041] 编码本发明的 T1R 蛋白质和多肽的核酸可以从多种来源分离获得,可根据 WO 00/035374 中所公开的方法采用遗传工程、扩增、合成和 / 或重组表达的方式获得,所述文献在此处整体引用作为参考。

[0042] 本发明也提供了筛选这些新的味觉细胞特异性 GPCRs 的调节子如活化剂、抑制剂、刺激物、增强剂、激动剂和拮抗剂的方法。这种味觉转导的调节子可用于味觉信号通路的药理学、化学和遗传学调节。这些筛选方法可用于鉴定味觉细胞活性的高亲和力和激动剂和拮抗剂。然后这些调节化合物可用在食品和制药工业中来定制味觉,例如调节食品或药物的甜味。

[0043] 因而,本发明提供了用于探测和表征味觉调节的测定方法,其中 T1R 家族成员充当了调节子对味觉转导效应的直接或间接的报道分子。GPCRs 在测定中可用于如测量体外、体内和来自体内的配体结合、离子浓度、膜电位、电流、离子流、转录、信号转导、受体 - 配体相互作用、第二信使浓度的变化。在一个实施方案中, T1R 家族成员通过附着于第二报道分子如绿色荧光蛋白之上可用作间接的报道分子 (参见如 Mistili & Spector, *Nature Biotechnology*, 15 :961-964 (1997))。在另一个实施方案中, T1R 家族成员可以在细胞中重组表达,并且通过 GPCRs 活性的味觉转导的调节可采用测量 Ca^{2+} 水平和其他细胞内信号如 cAMP、cGMP 或 IP3 的变化来测定。

[0044] 在某些实施方案中,将 T1R 多肽的一个结构域如细胞外、跨膜或细胞内结构域融合到一个异源多肽上,因此形成一种嵌合多肽,如具有 GPCR 活性的嵌合多肽。这种嵌合多肽可用于如在测定中鉴定 T1R 多肽的配体、激动剂、拮抗剂或者其他调节子。另外,这种嵌合多肽可用于创造具有新的配体结合特异性、调节模式、信号转导通路或其他此类性质的新味觉受体,或者用于创造具有配体结合特异性、调节模式、信号转导通路等特性的新组合的新味觉受体。

[0045] 在一个实施方案中, T1R 多肽在真核生物细胞中作为嵌合受体表达,该多肽具有异源的且有助于质膜运输或者经分泌路径成熟与定位的陪伴分子序列。可选的异源序列可以为视紫红质序列,如视紫红质的 N 末端片段。这种嵌合的 T1R 受体能够在任何真核生物细胞如 HEK-293 细胞中表达。优选地,这些细胞含有 G 蛋白,例如 G α 15 或 G α 16 或另一类型的混杂 G 蛋白,该混杂 G 蛋白能够将大范围的化学感受性 GPCRs 与细胞内信号通路或者与信号蛋白如磷脂酶 C 配对。可选择的,这些细胞可以表达嵌合的或变异的 G 蛋白从而产生功能性 T1R 味觉受体,所述嵌合或变异 G 蛋白是基于偶联 T1R 的能力挑选出来的。特别优选的变异 G 蛋白的例子包括在 2001 年 10 月 29 日提交的 U. S. 系列 No. 09/984, 292 中所公开的 G 蛋白变体,和在 U. S. 临时申请 No. 60/____、Attorney Docket No. 078003-0280737 中所公开的嵌合 G α 15 变体,这两篇文献都在此处整体引用作为参考。这些申请公开了 G 蛋白变体,其显示比一种熟知的混杂 G 蛋白 Ga15 能更好地偶联 T1Rs。这种嵌合受体在这种细胞中的活化可以通过使用任何标准方法来探测,如通过探测细胞中 FURA-2 依赖性荧

光来探测细胞内钙的变化。如果优选的宿主细胞不表达合适的 G 蛋白,它们可以用编码混杂 G 蛋白的基因转染,如在 US 申请系列 No. 60/243,770(此处整体引用作为参考)中所描述的那些 G 蛋白。

[0046] 味觉转导调节子的测定方法包括体外配体结合测定,该方法使用下列材料和方法:T1R 多肽、其部分,即细胞外结构域、跨膜区域或其组合或者含有一个或多个 T1R 家族成员的结构域的嵌合蛋白;表达 T1R 多肽、片段或融合蛋白的卵母细胞或组织培养细胞;T1R 家族成员的磷酸化和去磷酸化;结合到 GPCRs 上的 G 蛋白;配体结合测定;电压、膜电位和电导的变化;离子流测定;细胞内第二信使如 cGMP、cAMP 和肌醇三磷酸的变化;细胞内钙水平的变化;和神经递质的释放。

[0047] 此外,本发明提供了探测 T1R 核酸和蛋白质表达的方法,该方法使得能进行味觉转导调节的研究和味觉受体细胞的特殊鉴定。T1R 家族成员也提供了对于父子关系和法庭调查有用的核酸探针。T1R 基因也用作用于鉴定味觉受体细胞如叶状、蕈状、轮廓状、味觉带和会厌味觉受体细胞的核酸探针。T1R 受体也可用于产生用于鉴定味觉受体细胞的单克隆和多克隆抗体。可以采用下列技术来鉴定味觉受体细胞:如 mRNA 的反转录和扩增、总 RNA 或聚腺苷酸化 RNA 的分离、RNA 印迹法、点印迹法、原位杂交、RNA 酶保护、S1 消化、探测 DNA 微芯片阵列、蛋白质印迹法等。

[0048] 在功能上,T1R 多肽含有一个相关的七跨膜 G 蛋白偶联受体的家族,认为该受体参与了味觉转导,并可能与 G 蛋白相互作用从而介导了味觉信号转导(参见如 Fong, Cell Signal, 8:217(1996);Baldwin, Curr. Opin. Cell Biol., 6:180(1994))。在结构上,T1R 家族成员的核苷酸序列可能编码了含有细胞外结构域、七跨膜结构域和细胞质结构域的相关多肽。来自其他种类的相关 T1R 家族基因与 SEQ ID NO:1、2、3、9、11、13、15、16、20 或者其保守修饰的变体在至少大约 50 个核苷酸,可选的为 100、200、500 个或更多核苷酸长的区域具有至少大约 50%,可选的为 60%、70%、80%或 90%的核苷酸序列一致性;或者其编码的多肽与 SEQ ID NO:4、10、12、14、17、21 或者其保守修饰的变体在至少大约 25 个氨基酸,可选的为 50~100 个氨基酸长的区域具有至少大约 35~50%,可选的为 60%、70%、80%或 90%的氨基酸序列一致性。

[0049] 还鉴定了几个作为 T1R 家族成员的特征的共有氨基酸序列或结构域。例如,T1R 家族成员一般含有与 T1R 共有序列 1 和 2(分别为 SEQ ID NOs 18 和 19)具有至少大约 50%,可选的为 55%、60%、65%、70%、75%、80%、85%、90%、95-99%或更高一致性的序列。因而,这些保守结构域可用于鉴定 T1R 家族成员,其鉴定方法有一致性比较、特异性杂交或扩增或者用抗结构域的抗体进行特异性结合。这种 T1R 共有序列具有下列氨基酸序列:

[0050] T1R 家族共有序列 1:(SEQ ID NO:18)

[0051] (TR)C(FL)(RQP)R(RT)(SPV)(VERKT)FL(AE)(WL)(RHG)E

[0052] T1R 家族共有序列 2:(SEQ ID NO:19)

[0053] (LQ)P(EGT)(NRC)YN(RE)A(RK)(CGF)(VLI)T(FL)(AS)(ML)

[0054] 这些共有序列包括了那些在此处所描述的 T1R 多肽中所发现的序列,但是来自其他生物的 T1R 家族成员可以被认为包含与此处详细描述中所包含的一致序列具有大约 75%或更高一致性的一致序列。

[0055] T1R 核苷酸和氨基酸序列的特定区域可用于鉴定 T1R 家族成员的多态变体、种间

同系物和等位基因。该鉴定可在体外进行,例如在严格性杂交条件或 PCR(如用编码上面所鉴定的 T1R 共有序列的引物)下,或者使用计算机系统序列信息来与其他核苷酸序列进行比较。单一物种种群中的 T1R 基因的不同等位基因也可用于确定等位基因序列的差异是否和种群成员之间味觉感受的差异有关。经典的 PCR 型扩增和克隆技术可用于定向同源物(ortholog)的分离,例如当简并引物对于探测物种间相关基因是足够的时,这些物种一般具有比单一物种中 T1R 家族的平行进化同源成员更高水平的相关一致性。

[0056] 例如,简并引物 SAP 077(SEQ. ID NO. 5)和 SAP 0079(SEQ. ID NO. 6)可用于从不同哺乳动物基因组中扩增和克隆 T1R3 基因。相反地,单一物种中与 T1R3 有关的基因用序列模式识别软件进行最佳鉴定以寻找相关序列。一般地,T1R 家族成员的多态变体和等位基因的鉴定可以以下列方式进行,即比较含有大约 25 个或更多如 50-100 个氨基酸的氨基酸序列。大约至少 35-50%,可选的为 60%、70%、75%、80%、85%、90%、95-99% 或更高的氨基酸一致性一般可证明蛋白质为 T1R 家族成员的多态变体、种间同系物或等位基因。可用任何下面所讨论的序列比较算法进行序列比较。特异地结合到 T1R 多肽或者其保守区域上的抗体也可用于鉴定等位基因、种间同系物和多态变体。

[0057] T1R 基因的多态变体、种间同系物和等位基因可通过检查推定的 T1R 多肽的味觉细胞特异性表达来加以确认。一般地,具有这里所公开的氨基酸序列的 T1R 多肽可用作与推定 T1R 多肽比较的正对照从而证实 T1R 家族成员的多态变体或等位基因的鉴定。预期多态变体、等位基因和种间同系物保留了 G 蛋白偶联受体的七跨膜结构。对于进一步的详情,参见 WO 00/06592,其公开了相关的 T1R 家族成员 GPCR-B3,此处以与该公开内容相一致的方式引用其内容作为参考。此处将 GPCR-B3 受体称为 rT1R1 和 mT1R1。另外,参见 WO00/06593,其也公开了相关的 T1R 家族成员 GPCR-B4,此处以与该公开内容相一致的方式引用其内容作为参考。此处将 GPCR-B4 受体称为 rT1R2 和 mT1R2。

[0058] T1R 家族成员的核苷酸和氨基酸序列信息也可用于在计算机系统中构建味觉细胞特异性多肽的模型。随后,这些模型可用于鉴定能够激活或抑制 T1R 受体蛋白质的化合物。然后,这种调节 T1R 家族成员的活性的化合物可用于研究 T1R 基因和受体在味觉转导中的作用。

[0059] 本发明还提供了测定方法,优选的为高通量的测定方法来鉴定与 T1R 多肽相互作用和/或调节 T1R 多肽的分子。在大量的试验中,使用 T1R 家族成员的特殊结构域,例如细胞外、跨膜或者细胞内结构域或区域。在大量的实施方案中,可以将细胞外结构域、跨膜区域或者其组合结合到固体基质上,并用来如分离配体、激动剂、拮抗剂或者任何其他能够结合和/或调节 T1R 多肽活性的分子。

[0060] 在本发明的一个方面中,提供了一种新的 T1R 家族的人 GPCR 基因,称之为 hT1R3。hT1R3 基因从包括 GenBank 的 HTGS 部分的人基因组序列数据库来鉴定。在 SEQ. ID NOS 1-4 中提供了 hT1R3 的核苷酸和推定的氨基酸序列。hT1R3 受体在部分测序的 BAC 基因组克隆 RP5-89003(数据库存取号 AL 139287)中通过它与候选大鼠味觉受体 rT1R1(访问号 AF 127389)的序列相似性来鉴定。作为参考,预测的 hT1R3 和 rT1R1 蛋白质序列之间的成对一致性为大约 34%。与额外的 GPCR 家族 C(包括钙感受受体、推定的 V2R 外激素受体、GABA-B 受体、鱼味觉受体和 metabotropic 谷氨酸受体)的成员的序列比较表明,hT1R3 很可能属于由 T1R1 和第二个大鼠候选味觉受体(rT1R2,访问号 AF 127390)所定义的家族 C 亚群。

[0061] 本发明也提供了标明为 rT1R1 的大鼠味觉受体的人定向同源物,称之为 hT1R1。rT1R1 和 hT1R1 的基因产物具有大约 74% 的一致性。已经报道了小鼠基因 mT1R1, 参见 Hoon 等人, Cell, 96 :541-551 (2000), 它定位在与含有 hT1R1 的间隔同源的染色体间隔。此处分别将核苷酸和推定的 hT1R1 序列描述为 SEQ. ID NOS 15 和 16。

[0062] 虽然不希望受缚于任何特定理论,但是预测 T1R 家族受体依靠 mT1R3 与 Sac 基因座的连锁参与了甜味转导,所述影响甜味觉的 Sac 基因座位于染色体 4 的远端。也报道人 T1R3 位于 1p36.2-1p36.33, 该区域显示出与含有 Sac 和 T1R1 的小鼠间隔具有保守的同线性。然而, T1R 型受体可介导其他味觉形式,如苦、umami、酸和咸。

[0063] 多种保守的突变和替代被包含在本发明的范围之内。例如,在本领域的技术水平之内用已知的重组基因技术规程进行氨基酸替代,该重组基因技术包括 PCR、基因克隆、cDNA 的定点诱变、宿主细胞的转染和体外转录。然后,可以对于味觉细胞特异性 GPCR 功能的活性来筛选变体。

[0064] A. T1R 多肽的鉴定和表征

[0065] 本发明的 T1R 蛋白质和多肽的氨基酸序列可以通过其编码核酸序列的推定翻译来鉴定。这些多种氨基酸序列和编码核酸序列可以根据许多方法彼此或与其他序列进行比较。

[0066] 例如在序列比较中,一般一个序列作为参考序列,测试序列与之进行比较。当用序列比较算法时,将测试和参考序列输入计算机,如果需要,指明子序列坐标,指定序列算法程序参数。如下面对于 BLASTN 和 BLASTP 程序的描述,可以使用默认程序参数或者可以指定可选择的参数。然后,序列比较算法基于程序参数计算出测试序列相对于参考序列的序列一致性百分数。

[0067] 此处所用的“比较窗口”包括对于选自 20 至 600、通常大约 50 至大约 200、更通常大约 100 至大约 150 的邻近位置数中任一个的片段的参考,在其中,一个序列可以在两个序列最佳比对之后与具有同样邻近位置数的参考序列进行比较。比较序列的比对的方法在本领域中是公知的。比较序列的最佳比对可以通过下列方法来实施,例如 Smith&Waterman, Adv. Appl. Math. 2 :482 (1981) 的局部同源性算法、Needleman&Wunsch, J. Mol. Biol. 48 :443 (1970) 的同源性比对算法、Pearson & Lipman, Proc. Natl. Acad. Sci. USA 85 :2444 (1988) 的相似性搜索方法、这些算法的计算机执行 (GAP、BESTFIT、FASTA 和 TFASTA, Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI)、或者手工比对和肉眼检查 (参见如 Current Protocols in Molecular Biology (Ausubel 等人编, 1995 增刊))。

[0068] 适合于确定序列一致性和序列相似性百分比的算法的优选的例子为 BLAST 和 BLAST 2.0 算法,它们分别描述于 Altschul 等人, Nuc. Acids Res. 25 :3389-3402 (1977) 和 Altschul 等人, J. Mol. Biol. 215 :403-410 (1990)。进行 BLAST 分析的软件在国家生物技术信息中心 (National Center for Biotechnology Information) (<http://www.ncbi.nlm.nih.gov/>) 上是公共可获得的。该算法包括首先通过在查询序列中鉴定长度为 W 的短字串而鉴定高分值的序列对 (HSPs), 该短字串当与数据库序列中相同长度的字串比对时,或者匹配或者满足某个正的阈值分数 T。T 被称为邻近字串分数阈值 (Altschul 等人, Nuc. Acids Res. 25 :3389-3402 (1977) 和 Altschul 等人, J. Mol. Biol. 215 :403-410 (1990))。这些初始

的邻近字串命中值 (hit) 充当起始种子以发现含有该字串的更长 HSPs。字串命中值在每一序列的两个方向延伸,只要累积的比对分值可以增加。累积的分值对于核苷酸序列而言是用参数 M(一对匹配的残基的奖励分值;总是 > 0) 和 N(错配残基的惩罚分值;总值 < 0) 来计算的。对于氨基酸序列,用分数矩阵来计算累积的分值。字串命中值在两个方向的延伸在如下情况时终止:累积的比对分值从其最大所得值下降了量 X 时;累积分值由于一个或多个负分值残基比对的累积而达到零或低于零时;或者达到了任一序列的末端。BLAST 程序的参数 W、T 和 X 确定了比对的灵敏度和速度。BLASTN 程序(对于核苷酸序列)默认使用字长 (W) 为 11、期望 (E) 为 10、 $M = 5$ 、 $N = -4$ 和对两条链的比较。对于氨基酸序列,BLASTP 程序默认使用字长为 3 和期望 (E) 为 10,及 BLOSUM62 分值矩阵(参见 Henikoff&Henikoff, Proc. Natl. Acad. Sci. USA 89:10915(1989)) 比对 (B) 为 50、期望 (E) 为 10、 $M = 5$ 、 $N = -4$ 和对两条链的比较。

[0069] 有用的算法的另一个例子是 PILEUP。PILEUP 用渐进性的成对比对从一组相关序列中创建了多重的序列比对以显示相互关系和序列一致性百分比。它也标绘了显示用于创建比对的聚类 (clustering) 相互关系的所谓的“树”或“树形图 (dendrogram)”(参见如图 2)。PILEUP 使用 Feng&Doolittle, J Mol. Evol. 35:351-360(1987) 的渐进性比对方法的简化方法。所用的方法与由 Higgins&Sharp, CABIOS 5:151-153(1989) 描述的方法相似。该程序可比对多达 300 个序列,每一序列最长为 5,000 个核苷酸或氨基酸。该多重比对程序开始于对 2 个最相似序列的成对比对,产生了两个比对的序列的聚类。然后使该聚类与下一个最相关的序列或比对序列的聚类进行比对。2 个序列聚类是通过对 2 个单一序列的成对比对的简单延伸而比对的。最终的比对是通过一系列渐进的成对比对完成的。该程序是通过指定特定的序列并且在序列比较的区域指定它们的氨基酸或核苷酸坐标和指定程序的参数而运行的。利用 PILEUP,将参考序列与其他检验序列进行比较以用下面的参数确定序列一致性相互关系的百分比:默认的缺口权重 (3.00)、默认的缺口长度权重 (0.10) 和加权的末端缺口。PILEUP 可从 GCG 序列分析软件包如 7.0 版 (Devereaux 等人, Nuc. Acids Res. 12:387-395(1984)) 获得。由基因编码的是来自相应可读框的推定翻译。用 BLASTP 算法将这些蛋白质序列与公共序列数据库中所有已知蛋白质进行的比较揭示它们与 T1R 家族成员有强的同源性,每一个 T1R 家族序列都与至少一个已知的家族成员具有至少约 35~50% 的,且优选地为至少 55% 的、至少 60% 的、至少 65% 的和最优选地为至少 70% 的氨基酸一致性。

[0070] B. 定义

[0071] 如在此处所用的,下面的术语具有归于它们的意思,除非另外说明。

[0072] “味觉细胞”包括组织成群体的形式以形成舌的味蕾的神经上皮细胞,如叶状、蕈状和轮廓状细胞(参见如 Roper 等人, Ann. Rev. Neurosci. 12:329-353(1989))。味觉细胞也发现于腭和其他组织,如食管和胃。

[0073] “T1R”指一种或多种 G 蛋白偶联受体家族的成员,该成员表达于如叶状、蕈状和轮廓状细胞的味觉细胞以及腭和食管的细胞中(参见如 Hoon 等人, Cell, 96:541-551(1999), 此处整体引用作为参考)。该家族的成员在 WO 00/06592 中也称为 GPCR-B3 和 TR1,以及在 WO 00/06593 中称为 GPCR-B4 和 TR2。GPCR-B3 在此也称为 rT1R1, GPCR-B4 称为 rT1R2。味觉受体细胞也可在形态学的基础上进行鉴定(参见如上述的 Roper),或者通过在味觉细胞

中特定表达的蛋白质的表达来鉴定。T1R 家族成员可具有充当甜味转导的受体或者在多种其他的味觉形式间进行辨别的能力。

[0074] “T1R”核酸编码具有七跨膜区域且具有“G 蛋白偶联受体活性”的 GPCRs 家族,如它们可响应于细胞外的刺激而结合到 G 蛋白上,并通过对如磷脂酶 C 和腺苷酸环化酶的刺激而促进第二信使如 IP₃、cAMP、cGMP 和 Ca²⁺ 的生产(对于 GPCRs 结构和功能的描述,参见如上述的 Fong 和上述的 Baldwin)。单个味觉细胞可含有许多独特的 T1R 多肽。

[0075] 术语“T1R”家族因此指多态变体、等位基因、突变型和种间同系物,它们:(1) 与 SEQ ID NOS:4、10、12、14、17 或 21 在约 25 个氨基酸,可选择地为 50-100 个氨基酸的窗口上具有至少约 35 ~ 50% 的氨基酸序列一致性,可选择地为约 60、75、80、85、90、95、96、97、98 或 99% 的氨基酸序列一致性;(2) 特异性地结合到抗免疫原的抗体上,该免疫原包含选自 SEQ ID NOS:4、10、12、14、17、21 及其保守修饰的变体的氨基酸序列;(3) 由在严格性杂交条件下与选自 SEQ ID NOS:1、2、3、9、11、13、15、16、20 及其保守修饰的变体的序列杂交的核酸分子(大小至少为约 100 个,可选地至少为约 500-1000 个核苷酸)编码;(4) 包含与选自 SEQ ID NOS:4、10、12、14、17 或 21 的氨基酸序列具有至少约 35 ~ 50% 的一致性的序列;或(5) 是由在严格性杂交条件下与编码 SEQ ID NOS:7、8 及其保守修饰的变体的简并引物集相同的序列杂交的引物扩增的。

[0076] 在拓扑学上,某些化学感受性 GPCRs 具有“N-末端结构域”、“细胞外结构域”、包含七跨膜区域的“跨膜结构域”和相应的胞质和胞外环、“胞质结构域”和“C-末端结构域”(参见如 Hoon 等人,Cell,96:541-551(1999);Buck&Axel,Cell,65:175-187(1991))。这些结构域可用本领域技术人员公知的技术进行结构鉴定,如鉴定疏水和亲水结构域的序列分析程序(参见如 Stryer,Biochemistry,(3rd ed. 1988);也参见许多基于因特网的序列分析程序的任何一种,如那些发现于 dot.imgen.bcm.tmc.edu 的)。这种结构域可用于构建嵌合蛋白质,也可用于本发明的体外测定,如配体结合测定。

[0077] “细胞外结构域”因此指 T1R 多肽从细胞膜伸出并暴露于细胞的细胞外表面的结构域。这种结构域一般包括暴露于细胞外表面的“N-末端结构域”,并可选择地可包括跨膜结构域暴露于细胞外表面的部分细胞外环,即在跨膜区域 2 和 3 之间、跨膜区域 4 和 5 之间以及跨膜区域 6 和 7 之间的环。

[0078] “N-末端结构域”区域起始于 N-末端并延伸至靠近跨膜结构域开始处的区域。更具体而言地,在本发明的一个实施方案中,该结构域起始于 N-末端并约终止于氨基酸位置 563 正负约 20 个氨基酸处的保守谷氨酸。这些细胞外结构域可用于可溶性和固相体外配体结合测定。另外,在下面描述的跨膜区域也可与细胞外结构域组合而结合配体,并因此也可用于体外配体结合测定。

[0079] 包含 7 个“跨膜区域”的“跨膜结构域”指 T1R 多肽位于质膜内部的结构域,且也可包括相应的胞质(细胞内)和细胞外环。在一个实施方案中,该区域相应于约起始于氨基酸位置 563 正负约 20 个氨基酸的保守谷氨酸残基处而约终止于位置 812 正负约 10 个氨基酸的保守酪氨酸残基处的 T1R 家族成员的结构域。该七跨膜区域和细胞外以及胞质环可用标准方法进行鉴定,如描述于 Kyte&Doolittle, J. Mol. Biol., 157:105-32(1982) 或上述的 Stryer 中的。

[0080] “胞质结构域”指 T1R 多肽面向细胞内侧的结构域,如“C 末端结构域”和跨膜结构

域的细胞内环,如在跨膜区域 1 和 2 之间的细胞内环、跨膜区域 3 和 4 之间的细胞内环和跨膜区域 5 和 6 之间的细胞内环。“C 末端结构域”指跨越最后的跨膜结构域的末端和蛋白质的 C-末端的区域,且该区域通常位于胞质内。在一个实施方案中,该区域起始于位置 812 正负约 10 个氨基酸的保守酪氨酸残基处并延伸到多肽的 C-末端。

[0081] 术语“配体结合区域”或“配体结合结构域”指源自化学感受性受体,具体而言地为味觉受体的序列,该序列一般包括至少受体的细胞外结构域。在一个实施方案中,配体结合区域的细胞外结构域可包括 N-末端结构域和(可选择的)跨膜结构域的部分,如跨膜结构域的细胞外环。配体结合区域能够结合配体,且更具体而言地为结合一种味觉刺激物。

[0082] 在检验调节 T1R 家族成员介导的味觉转导的化合物的测定中,短语“功能作用”包括对任何间接地或直接地受受体影响的参数如功能的、物理的和化学的作用的确定。它包括体外、体内和来自体内的配体结合、离子流变化、膜电位、电流、转录、G 蛋白结合、GPCR 磷酸化或脱磷酸化、信号转导、受体-配体相互作用、第二信使浓度(如 cAMP、cGMP、IP3 或细胞内 Ca^{2+}),且也包括其他的生理作用如神经递质或激素释放的增加或减少。

[0083] 在测定中,“确定功能作用”意思为对一个化合物的测定,其中该化合物增加或降低了间接或直接受 T1R 家族成员影响的参数,如功能的、物理的和化学的作用。这种功能作用可通过任何本领域技术人员公知的方法进行测量,如光谱学特征的变化(如荧光、吸收、折射率)、流体动力学(如形状)的、色谱的或可溶性性质、膜片钳、电压灵敏染料、全细胞电流、放射性同位素外向通量、诱导型标记、卵母细胞 T1R 基因表达的变化;组织培养细胞 T1R 表达;T1R 基因的转录激活;配体结合测定;电压、膜电位和电导变化;离子流分析;细胞内第二信使如 cAMP、cGMP 和肌醇三磷酸(IP3)的变化;细胞内钙水平的变化;神经递质的释放等。

[0084] T1R 基因或蛋白质的“抑制剂”、“激活剂”和“调节子”可互换使用,用于指用体外和体内测定鉴定的对味觉转导的抑制性、激活性或调节性分子,如配体、激动剂、拮抗剂及其同系物和模拟物。抑制剂是如结合、部分或全部地阻断刺激、降低、防止、延迟活化、失活、脱敏或减量调节味觉转导的化合物,如拮抗剂。激活剂是如结合、刺激、增加、开放、活化、促进、增强活化、致敏、或增量调节味觉转导的化合物,如激动剂。调节子包括那些如改变受体与下述物质相互作用的化合物:结合激活剂或抑制剂的细胞外蛋白质(如 ebnerin 和其他疏水载体家族的成员);G 蛋白;激酶(如视紫红质激酶和参与受体失活和脱敏中的 β 肾上腺素能受体激酶的同源物)和也使受体失活和脱敏的视紫红质抑制蛋白。调节子可包括 T1R 家族成员的遗传修饰形式(如具有改变的活性)以及天然存在和合成的配体、拮抗剂、激动剂、小的化学分子等。对抑制剂和激活剂的这种测定包括(如上所述)如在细胞或细胞膜上表达 T1R 家族成员、在味觉刺激物如甜味刺激物存在或不存在时应用推定的调节化合物,然后确定对味觉转导的功能作用。将包含用潜在的激活剂、抑制剂或调节子处理的 T1R 家族成员的样品或测定物与无抑制剂、激活剂或调节子的对照样品进行比较以检查调节的程度。对对照样品(未用调节子处理的)赋予 100% 的相对 T1R 活性值。在当 T1R 活性值相对于对照为约 80%,可选择的为 50%或 25-0%时,对 T1R 的抑制就实现了。对 T1R 的激活是在当 T1R 活性值相对于对照为 110%,可选择的为 150%,可选择的为 500%或 1000-3000%时实现的。

[0085] 如在此处所用的术语“纯化的”、“基本纯化的”和“分离的”指不含其他不同化合

物的状态,其中本发明的化合物与所述不同化合物在天然状态下是结合的,因而“纯化的”、“基本纯化的”和“分离的”对象按重量包含给定样品的至少 0.5%、1%、5%、10%或 20%,且最优选地为至少 50%或 75%的物质。在一个优选实施方案中,这些术语指本发明的化合物按重量包含给定样品的至少 95%的物质。如在此处所用的,术语“纯化的”、“基本纯化的”和“分离的”当涉及核酸或蛋白质时也指与在哺乳动物尤其是人体中天然发生的不同的纯化或浓缩状态。任何比天然发生于哺乳动物尤其是人体中的高的纯化或浓缩程度都在“分离的”意思之内,包括(1)从其他结合的结构或化合物中纯化或(2)与在哺乳动物尤其是人体中不是天然结合的结构和化合物结合。在此处描述的核酸或蛋白质或核酸或蛋白质类群可根据本领域技术人员所公知的多种方法和过程来分离,或与其天然时一般不结合的结构或化合物结合。

[0086] 如在此处所用的,术语“分离的”当涉及核酸或多肽时指与天然发生于哺乳动物尤其是人体中的不同的纯化或浓缩状态。任何比天然发生于哺乳动物尤其是人体中的高的纯化或浓缩程度都在此处所用的“分离的”意思之内,包括(1)从其他天然发生的结合的结构或化合物中纯化或(2)与在身体中通常是不结合的结构和化合物结合。在此处描述的核酸或多肽可根据本领域技术人员所公知的多种方法和过程来分离,或与其天然时一般不结合的结构或化合物结合。

[0087] 如在此处所用的,术语“进行扩增”和“扩增”指利用任何适当的扩增方法以生成或探测重组或天然表达的核酸,如在下面详细描述。例如,本发明提供了在体内或体外扩增(如通过聚合酶链反应,PCR)天然表达的(如基因组或 mRNA)或重组(如 cDNA)的本发明核酸(如本发明的味觉刺激物结合序列)的方法和试剂(如特定的简并寡核苷酸引物对)。

[0088] 术语“七跨膜受体”意思为一种属于跨膜蛋白质超家族的多肽,该蛋白质具有跨越质膜 7 次的 7 个结构域(因而,该 7 个结构域被称为“跨膜”或“TM”结构域 TM I ~ TMVII)。嗅觉和某些味觉受体的家族各自属于该超家族。七跨膜受体多肽具有相似的典型性一级、二级和三级结构,如在下面进一步详细讨论的。

[0089] 术语“文库”意思为一种不同核酸或多肽分子的混合物制剂,如重组生成的化学感受性,特别是味觉受体配体结合结构域的文库,该文库是用简并引物对核酸的扩增来生成的,或者是整合了扩增的配体结合结构域的载体的分离集合,或者是每个细胞随机用至少一个编码味觉受体的载体转染的细胞的混合物。

[0090] 术语“核酸”或“核酸序列”指单链或双链形式的脱氧核糖核苷酸或核糖核苷酸寡核苷酸。该术语包含含有天然核苷酸的已知类似物的核酸(即寡核苷酸)。该术语也包含具有合成主链的类核酸结构(参见如 *Oligonucleotides and Analogues, a Practical Approach*), ed. F. Eckstein, Oxford Univ. Press(1991); *Antisense Strategies*, *Annals of the N. Y. Academy of Sciences*, Vol. 600, Eds. Baserga 等人 (NYAS1992); *Milligan J. Med. Chem.* 36:1923-1937(1993); *Antisense Research and Applications*(1993, CRC Press), WO 97/03211; WO96/39154; Mata, *Toxicol. Appl. Pharmacol.* 144:189-197(1997); Strauss-Soukup, *Biochemistry* 36:8692-8698(1997); Samstag, *Antisense Nucleic Acid Drug Dev.* 6:153-156(1996))。

[0091] 除非另外指出的,特别的核酸序列也隐含其保守修饰的变体(如简并密码子替代)和互补序列以及明确指出的序列。特定地,简并密码子替代可通过生成如其一

个或多个选择的密码子的第三位被混合的碱基和 / 或脱氧肌苷残基替代的序列而实现 (Batzer 等人, *Nucleic Acid Res.*, 19:5081(1991); Ohtsuka 等人, *J. Biol. Chem.*, 260:2605-2608(1985); Rossolini 等人, *Mol. Cell. Probes*, 9:91-98(1994))。术语核酸是与基因、cDNA、mRNA、寡核苷酸和多核苷酸互换使用的。

[0092] 此处术语“多肽”、“肽”和“蛋白质”是互换使用的,用于指氨基酸残基的多聚体。该术语适用于氨基酸多聚体,其中一种或多种氨基酸残基是相应的天然存在的氨基酸的人工化学模拟物,也适用于天然存在的氨基酸多聚体和非天然存在的氨基酸多聚体。

[0093] 术语“原生质膜转运结构域”或简单的“转运结构域”意思为这样一种多肽结构域,当该结构域整合于多肽编码序列的氨基末端时可高效地“陪伴”或将杂交(“融合”)蛋白质“转运”到细胞质膜。例如,“转运结构域”可源自一种七跨膜受体牛视紫红质受体多肽的氨基末端。然而,可使用源自任何哺乳动物的视紫红质,也可使用其他的转运促进序列。因而,该转运结构域在将七跨膜融合蛋白质转运到质膜中是特别有效的,且包含氨基末端转运结构域的蛋白质(如味觉受体多肽)可比不具有该结构域的蛋白质更有效地被转运到质膜。然而,如果多肽的 N-末端结构域有结合活性,则其他转运结构域的使用是优选的。

[0094] 此处所描述的“转运结构域”、“配体结合结构域”和嵌合受体组合物也包括具有基本相当于示范序列的结构和活性的“类似物”或“保守的变体”和“模拟物”(“肽模拟物”)。因而,术语“保守的变体”或“类似物”或“模拟物”指具有修饰的氨基酸序列的多肽,它一般不改变多肽的(保守的变体的)结构和 / 或活性,如在此处所定义的。这包括氨基酸序列的保守修饰的变体,即对蛋白质活性非关键的那些残基的氨基酸替代、添加或删除,或用具有相似性质(如酸性、碱性、带正或负电荷的、极性或非极性的等)的残基替代氨基酸,从而对即使关键氨基酸的替代也一般不改变其结构和 / 或活性。

[0095] 更具体而言地,“保守修饰的变体”适用于氨基酸和核酸序列两者。关于特定的核酸序列,保守修饰的变体指那些编码相同的或基本相同的氨基酸序列的核酸,或者在核酸不编码氨基酸序列时指基本相同的序列。因为遗传密码的简并性,大量功能上相同的核酸编码任何给定的蛋白质。

[0096] 例如,密码子 GCA、GCC、GCG 和 GCU 都编码氨基酸丙氨酸。因而,在由密码子指定的丙氨酸的每一个位置,该密码子可改变为任何所描述的相应密码子而不改变所编码的多肽。

[0097] 这种核酸变异是“沉默变异”,是保守修饰的变异的一种。此处编码多肽的每个核酸序列也描述了核酸每一个可能的沉默变异。技术人员将认识到核酸中的每一个密码子(除一般为甲硫氨酸的唯一密码子的 AUG 和一般为色氨酸的唯一密码子的 TGG 外)可进行修饰以产生功能上相同的分子。因此,编码多肽的核酸的每一个沉默变异在每一个描述的序列中都是隐含的。

[0098] 提供功能上相似的氨基酸的保守替代表是本领域所公知的。例如,选择保守替代的示范性指南包括(原始残基之后为示范性的替代):ala/gly 或 ser、arg/lys、asn/gln 或 his、asp/glu、cys/ser、gln/asn、gly/asp、gla/ala 或 pro、his/asn 或 gln、ile/leu 或 val、leu/ile 或 val、lys/arg 或 gln 或 glu、met/leu 或 tyr 或 ile、phe/met 或 leu 或 tyr、ser/thr、thr/ser、trp/tyr、tyr/trp 或 phe、val/ile 或 leu。可选择的示范性指南使用如下的 6 个组,每组包含相互为保守替代的氨基酸:1) 丙氨酸(A)、丝氨酸(S)、苏氨酸(T);

2) 天冬氨酸 (D)、谷氨酸 (E) ;3) 天冬酰胺 (N)、谷氨酰胺 (Q) ;4) 精氨酸 (R)、赖氨酸 (I) ;5) 异亮氨酸 (I)、亮氨酸 (L)、甲硫氨酸 (M)、缬氨酸 (V) ;和 6) 苯丙氨酸 (F)、酪氨酸 (Y)、色氨酸 (W) ;(参见如 Creighton, Proteins, W. H. Freeman and Company (1984) ;Schultz 和 Schimer, Principles of Protein Structure, Springer-Vrlag (1979))。本领域的技术人员将理解上述的替代不是唯一可能的保守替代。例如,为了一些目的,可将所有带电荷的氨基酸认作是保守的替代,不论它们是正的还是负的。另外,在编码的序列中改变、添加或删除单个氨基酸或小百分比氨基酸的单独替代、删除或添加也可被认为是“保守修饰的变异”。

[0099] 术语“模拟物”和“肽模拟物”指一种合成的化学化合物,该化合物基本具有多肽的相同的结构和 / 或功能特征,如本发明的转运结构域、配体结合结构域或嵌合受体。该模拟物可全部由氨基酸的合成的非天然类似物组成,或可为部分的天然肽氨基酸和氨基酸的部分非天然类似物的嵌合分子。该模拟物也可包含任何量的天然氨基酸保守替代,只要这种替代也基本上不改变模拟物的结构和 / 或活性。

[0100] 与作为保守变体的本发明的多肽一样,常规的实验法将确定一个模拟物是否是在本发明的范围之内的,即其结构和 / 或功能基本上没有改变。多肽模拟物组合物可含有非天然结构成分的任何组合,该成分一般来自 3 个结构组 :a) 除天然酰胺键 (“肽键”) 键合之外的残基键合组 ;b) 替代天然存在的氨基酸残基的非天然残基 ;或 c) 诱导二级结构模拟的残基,即诱导或稳定二级结构如 β -转角、 γ -转角、 β -折叠、 α -螺旋构象等。多肽在所有或一些其残基是通过化学方法而不是天然的肽键连接时可被赋予模拟物的特征。单独的肽模拟物残基可以通过肽键、其他的化学键或偶联方式来连接,如戊二醛、N- 羟基琥珀酰亚胺酯、双功能顺丁烯二酰亚胺、N, N' - 二环己基碳二亚胺 (DCC) 或 N, N' - 二异丙基碳二亚胺 (DIC)。可替代常规的酰胺键 (“肽键”) 键合的连接基团包括如氧代亚甲基 (如 $-C(=O)-CH_2-$ 替代 $-C(=O)-NH-$)、氨基亚甲基 (CH_2-NH)、亚乙基、链烯 ($CH=CH$)、醚 (CH_2-O)、硫醚 (CH_2-S)、四唑 (CN_4)、噻唑、retroamide、硫代酰胺或酯 (参见如 Spatola, 氨基酸、肽和蛋白质的化学和生物化学 (Chemistry and Biochemistry of Amino Acids, Peptides and Proteins), Vol. 7, pp267-357, “肽主链修饰 (peptide Backbone Modifications)”, Marcell Dekker, NY (1983))。多肽也可通过含有替代天然存在的氨基酸残基的所有或一些非天然残基而被赋予模拟物的特征 ;非天然的残基是在科学和专利文献中充分描述的。

[0101] “标记”或“可探测的部分”是可通过光谱学、光化学、生物化学、免疫化学或化学方法进行探测的组合物。例如,有用的标记包括 ^{32}P 、荧光染料、电子致密试剂、酶 (如通常在 ELISA 中所用的)、生物素、洋地黄毒苷或半抗原和蛋白质,该半抗原和蛋白质可制备为可探测的,如通过向肽中整合放射性标记的,或者可用于探测与肽特异性反应的抗体。

[0102] “标记的核酸探针或寡核苷酸”是或者通过连接体或化学键共价地,或者通过离子、范德瓦耳力、静电或氢键非共价地结合到标记上的分子,从而使得探针的存在可通过探测结合到探针上的标记的存在而探测。

[0103] 如在此处所用的“核酸探针或寡核苷酸”定义为能够结合到目标互补序列核酸的核酸,该结合是通过一种或多种类型的化学键形成的,通常是通过互补碱基配对 (通常是通过氢键) 形成的。如在此处所用的探针可包括天然的 (即 A、G、C 或 T) 或修饰的碱基 (7- 脱氮鸟苷、肌苷等)。另外,探针中的碱基可通过除磷酸二酯键之外的键合来连接,只要其不干扰杂交。因而,例如探针可为肽核酸,其中的组成碱基是通过除磷酸二酯键键合之外

的肽键连接的。本领域的技术人员可理解依赖于杂交条件的严格性程度,探针可结合到与探针序列缺乏完全互补性的目标序列上。该探针可选择地直接用同位素、发色团、发光团(lumiphores)、色原进行标记,或间接地用如生物素进行标记,其后可在生物素上结合链霉抗生物素蛋白复合物。通过测定探针存在与否,人们可探测选定的序列或子序列存在与否。

[0104] 术语“异源的”当关于核酸的部分使用时表明核酸具有两个或多个子序列,该子序列在天然状态下来发现具有与此相同的相互关系。例如,核酸一般是重组地生产的,具有两个或多个来自不相关基因的序列,该序列被排列以形成新的功能核酸,如来自一个来源的启动子和来自另一个来源的编码区域。相似地,异源的蛋白质表明该蛋白质包含两个或多个子序列,该子序列在天然状态下未发现具有与此相同的相互关系(如融合蛋白质)。

[0105] “启动子”定义为指导核酸转录的核酸序列排列。如在此处所用的,启动子包括靠近转录起始位点的必需核酸序列,如在聚合酶 II 型启动子的情况下为 TATA 元件。启动子也可选择地包括远端的增强子或抑制物元件,它们可位于离开转录起始位点长达几千个碱基对的位置。“组成型的”启动子是在大多数环境和发育条件下起作用的启动子。“诱导型”启动子是在环境或发育调节下起作用的启动子。术语“可操作地连接的”指在核酸表达控制序列(如启动子或转录因子结合位点排列)与第二个核酸序列之间的功能键合,其中的表达控制序列指导相应于第二个序列的核酸的转录。

[0106] 如在此处所用的,“重组”指合成的或在体外操作的多核苷酸(如“重组多核苷酸”),指利用重组多核苷酸以在细胞或其他生物系统中生产基因产物的方法,或者指由重组多核苷酸编码的多肽(“重组蛋白质”)。“重组方法”也包含将具有来自不同来源的多种编码区域或结构域或启动子序列的核酸连接入表达盒或载体中,从而用于如融合蛋白质的诱导型或组成型表达,该融合蛋白质包含本发明的转运结构域和用本发明的引物扩增的核酸序列。

[0107] 短语“选择性(或特异性)杂交”指在严格性的杂交条件下一个分子仅与一个存在于复杂混合物(如总的细胞或文库 DNA 或 RNA)中的特别的核苷酸序列结合、形成双链或杂交。

[0108] 短语“严格性杂交条件”指在其中探针将与其一般在复杂混合物中的目标序列杂交而不与其他序列杂交的条件。严格性条件是序列依赖性的且在不同的环境下是不同的。较长的序列在较高的温度下特异性地杂交。对核酸杂交的详尽指南参见 Tijssen, 生物化学和分子生物学技术-用核酸探针杂交(Techniques in Biochemistry and Molecular Biology-Hybridisation with Nucleic Probes),“Overview of principles of hybridization and the strategy of nucleic acid assays”(1993)。通常,严格性条件选择为比在限定的离子强度 pH 中特异性序列的热熔解点(T_m)低约 5-10°C。该 T_m 是在平衡时(由于目标序列过量存在,因而在 T_m 时,50%的探针在平衡时杂交了)与目标互补的探针的 50%与目标序列杂交的温度(在限定的离子强度、pH 和核酸浓度时)。严格性条件是那些其中的盐浓度少于约 1.0M 钠离子,一般在 pH 7.0 ~ 8.3 为 0.01 ~ 1.0M 钠离子浓度(或其他盐),且其温度对于短的探针(如 10 ~ 50 个核苷酸)至少为约 30°C 而对于长的探针(如大于 50 个核苷酸)至少为约 60°C。严格性条件也可通过添加去稳定剂如甲酰胺而实现。对于选择性或特异性杂交,阳性信号应为背景的至少 2 倍,可选择地为背景杂交的 10 倍。示范性严格性杂交条件可为如下的:50%甲酰胺、 S_x SSC 和 1% SDS,在 42°C 温

育,或者 $S \times SSC$ 和 1% SDS,在 65°C 温育,并于 65°C 在 0.2x SSC 和 1% SDS 中洗涤。这样的杂交和洗涤步骤可进行如 1、2、5、10、15、30、60 或更多分钟。

[0109] 在严格性条件下没有相互杂交的核酸如果它们编码的多肽基本相关,则仍然是基本相关的。这发生于如当核酸的一个拷贝是用由遗传密码所允许的最大密码子简并性而生成的时。在这种情况下,该核酸一般在中等严格性杂交条件下进行杂交。示范性的“中等严格性杂交条件”包括在 40% 甲酰胺、1M NaCl、1% SDS 的缓冲液中于 37°C 进行杂交并于 45°C 在 1x SSC 中洗涤。这种杂交和洗涤步骤可进行如 1、2、5、10、15、30、60 或更多分钟。阳性杂交为背景的至少 2 倍。那些普通技术人员易于认识到替代的杂交和洗涤条件可用于提供相似严格性的条件。

[0110] “抗体”指包含来自免疫球蛋白基因或其片段的构架区域的多肽,该多肽特异性地结合并识别抗原。公认的免疫球蛋白基因包括 κ 、 λ 、 α 、 γ 、 δ 、 ϵ 、 μ 恒定区基因,以及无数的免疫球蛋白可变区基因。轻链划分为 κ 或者 λ 。重链划分为 γ 、 μ 、 α 、 δ 或 ϵ ,这依次分别定义了免疫球蛋白种类: IgG、IgM、IgA、IgD 和 IgE。

[0111] 示范性的免疫球蛋白(抗体)结构单元包含四聚体。每个四聚体包含两个相同的多肽链对,每一对具有一个“轻”链(约 25kDa)和一个“重”链(约 50-70kDa)。每一条链的 N-末端定义了约 100 ~ 110 个或更多氨基酸的可变区,该可变区主要负责抗原识别。术语可变的轻链(VL)和可变的重链(VH)分别指这些轻链和重链。

[0112] “嵌合抗体”是这样一种抗体分子,其中:(a) 恒定区或其部分被改变、替代或交换了,从而抗原结合位点(可变区)连接到不同的或改变的种类的恒定区、效应物功能和/或种类或赋予了该嵌合抗体新的性质的完全不同的分子上,如酶、毒素、激素、生长因子、药物等;或(b) 可变区或其部分进行改变、替代或与具有不同的或改变的抗原特异性的可变区进行交换。

[0113] “抗-T1R”抗体是特异性地结合由 T1R 基因、cDNA 或子序列编码的多肽的抗体或抗体片段。

[0114] 术语“免疫测定”是用抗体特异性地结合抗原的测定。免疫测定特征在于使用特定抗体的特异结合性质来分离、定位和/或定量抗原。

[0115] 当指蛋白质或肽时,短语“特异性地(或选择性地)结合”抗体或“特异性地(或选择性地)免疫反应”指一种结合反应,该反应确定蛋白质在蛋白质和其他生物制剂的异源群体中的存在。因而,在指定的免疫测定条件下,特异性的抗体对特定蛋白质的结合至少两倍于背景,且基本不以显著的量结合存在于样品中的其他蛋白质。在这种条件下对抗体的特异性结合可能需要一种根据其对特定蛋白质的特异性而选择的抗体。例如,从特定的物种如大鼠、小鼠或人而来的 T1R 家族成员的多克隆抗体可选择来获得那些多克隆抗体,该多克隆抗体特异性地与 T1R 多肽或其免疫原性部分进行免疫作用,而不与除 T1R 多肽的定向同源物或多态变体和等位基因之外的其他蛋白质进行免疫作用。该选择可通过将与 T1R 分子交叉反应的抗体从其他种类或其他 T1R 分子中减除而实现。也可选择那些仅识别 T1R GPCR 家族成员但不识别来自其他家族的 GPCRs 的抗体。多种免疫测定形式可用于选择与特定蛋白质发生特异性免疫反应的抗体。例如,固相 ELISA 免疫测定常规用于选择与蛋白质发生特异性免疫反应的抗体(对于可用于确定特定的免疫反应性的免疫测定形式和条件,参见如 Harlow&Lane, *Antibodies, A Laboratory Manual*, (1988))。一般地,特异性或选择

性反应将至少为背景信号或噪音的 2 倍,且更一般地为大于背景的 10 ~ 100 倍。

[0116] 短语“选择性地结合”指如上所述的核酸与另一种核酸“选择性地杂交”的能力,或者如上所述的一种抗体与蛋白质“选择性地(或特异地)结合”的能力。

[0117] 术语“表达载体”指任何目的在于表达本发明的核酸序列的重组表达系统,该表达是在体外或体内进行的、可以是组成型的或诱导型的、可在任何细胞包括原核的、酵母的、真菌的、植物的、昆虫的或哺乳动物的细胞中进行。该术语包括线状或环状表达系统。该术语包括保持为附加体或整合入宿主细胞基因组中的表达系统。该表达系统可具有自身复制能力或不具有自身复制能力,即在宿主细胞中仅驱动瞬时的表达。该术语包括仅含有重组核酸转录所需要的最少元件的重组表达盒。

[0118] “宿主细胞”意思指含有表达载体并支持该表达载体复制或表达的细胞。宿主细胞可为原核细胞如大肠杆菌(E. coli)或真核细胞如酵母、昆虫、两栖动物或哺乳动物细胞如 CHO、HeLa、HEK-293 等,如培养的细胞、外植体和体内的细胞。

[0119] A. T1R 多肽的分离及表达

[0120] 本发明的 T1Rs 或其片段或变体的分离及表达可如下面所描述的进行。可将 PCR 引物用于编码味觉受体配体结合区域的核酸的扩增,且可选择地生成这些核酸的文库。然后单独的表达载体或表达载体文库可用于感染或转染宿主细胞以使得这些核酸或文库进行功能性表达。这些基因和载体可在体外或体内制备及表达。技术人员将认识到改变和控制核酸表达的想要的表型可通过调节本发明的载体内的基因和核酸(如启动子、增强子等)的表达或活性而获得。可以使用任何已知的能增加或降低表达或活性的方法。本发明可与任何本领域公知的方法或规程结合实施,其中的方法和规程是在科学和专利文献中详尽描述的。

[0121] 本发明的核酸序列和其他用于实施本发明的核酸,不论是 RNA、cDNA、基因组 DNA、载体、病毒还是其杂交体,可从多种来源分离、基因工程改造、扩增和/或重组表达。可以使用任何重组表达系统,该系统除哺乳动物细胞之外包括如细菌、酵母、昆虫或植物系统。

[0122] 可选择地,这些核酸可利用众所周知的化学合成技术进行体外合成,如描述于 Carruthers, Cold Spring Harbor Symp. Quant. Biol. 47 :411-418(1982) ;Adams, Am. Chem. Soc. 105 :661(1983) ;Belousov, Nucleic Acids Res. 25 :3440-3444(1997) ;Frenkel, Free Radic. Biol. Med. 19 :373-380(1995) ;Blommers, Biochemistry 33 :7886-7896(1994) ;Narang, Meth. Enzymol. 68 :90(1979) ;Brown, Meth. Enzymol. 68 :109(1979) ;Beaucage, Tetra. Lett. 22 :1859(1981) ;美国专利 No. 4, 458, 066。然后双链 DNA 片段可通过合成互补链并在适当的条件下使链退火获得,或者通过用 DNA 聚合酶以适当的引物序列来添加互补链。

[0123] 对核酸进行操作的技术如在序列中生成突变、亚克隆、标记探针、测序、杂交等是在科学和专利文献中详尽描述的。参见如 Sambrook, ed., Molecular Cloning :a Laboratory manual (2nd ed.), Vols. 1-3, Cold Spring Harbor Laboratory(1989) ;Current Protocols in Molecular Biology, Ausubel, ed. John Wiley&Sons, Inc., New York(1997) ;生物化学和分子生物学实验室技术:用核酸探针进行杂交(Laboratory Techniques in Biochemistry and Molecular Biology :Hybridization With Nucleic Acid Probes), 部分 I, 理论和核酸制剂(Theory and Nucleic Acid Preparation), Tijssen, ed. Elsevier,

N. Y. (1993)。

[0124] 核酸、载体、病毒壳体、多肽等可通过大量本领域技术人员所公知的一般方法来分析和量化。这包括如分析生物化学技术如 NMR、分光光度法、放射显影法、电泳、毛细管电泳、高效液相色谱 (HPLC)、薄层色谱 (TLC) 和超扩散 (hyperdiffusion) 色谱, 多种免疫学方法如流体或凝胶沉淀素反应、免疫扩散、免疫电泳、放射免疫测定 (RIAs)、酶联免疫吸附测定 (ELISAs)、免疫荧光测定, DNA 印迹分析, RNA 印迹分析, 斑点印迹分析, 凝胶电泳 (如 SDS-PAGE), RT-PCT, 定量 PCR, 其他核酸或目标或信号扩增方法, 放射性标记, 闪烁计数和亲和层析。

[0125] 可将寡核苷酸引物用于扩增编码味觉受体配体结合区域的核酸片段。此处所描述的核酸也可用扩增技术进行克隆或定量测量。扩增方法也是本领域所公知的, 且包括如聚合酶链反应 PCR (PCR Protocols, a Guide to Methods and Applications, ed. Innis. Academic Press, N. Y. (1990) 和 PCR Strategies, ed. Innis. Academic Press, Inc., N. Y. (1995)); 连接酶链反应 (LCR) (参见如 Wu, Genomics 4:560(1989); Landegren, Science 241:1077(1988); Barringer, Gene 89:117(1990)); 转录扩增 (参见如 Kwok, Proc. Natl. Acad. Sci. USA 86:1173(1989)); 和自动维持序列扩增 (参见如 Guatelli, Proc. Natl. Acad. Sci. USA 87:1874(1990)); Q β 复制酶扩增 (参见如 Smith, J. Clin. Microbiol. 35:1477-1491(1997)); 自动 Q β 复制酶扩增测定 (参见如 Burg, Mol. Cell. Probes 10:257-271(1996)) 和其他 RNA 聚合酶介导的技术 (如 NASBA, Canguene, Mississauga, Ontario); 也参见 Berger, Methods Enzymol. 152:307-316(1987); Sambrook; Ausubel; 美国专利 Nos. 4,683,195 和 4,683,202; Sooknanan, Biotechnology 13:563-564(1995)。可将引物设计为保留“供体”7-膜受体的原始序列。可选择地, 该引物可编码作为保守替代的 (如用疏水残基替代疏水残基, 参见上述) 或功能性良性替代的 (如不防止质膜插入的、导致肽酶的切割、导致受体异常折叠等) 氨基酸残基。一旦扩增后, 可将该核酸单独地或者作为文库根据本领域公知的方法克隆, 如果需要, 可以用常规的分子生物学方法克隆入多种载体中的任何一个; 体外克隆扩增的核酸的方法描述于美国专利 No. 5,426,039。

[0126] 引物对可被设计为选择性地扩增 T1R 家族成员的配体结合区的。这些区域对不同的配体或味觉刺激物可有变化。因而, 作为一个味觉刺激物的最小结合区域对于第二个味觉刺激物可能太小了。因此, 可能要扩增包含不同细胞外结构域结构的不同大小的配体结合区域。

[0127] 设计简并引物对的范例是本领域所公知的。例如, 一致-简并杂交寡核苷酸引物 (Consensus-Degenerate hybrid Oligonucleotide Primer) (CODEHOP) 策略计算机程序是易从 <http://blocks.fhcr.org/codehop.html> 得到的, 且直接从 BlockMaker 多重序列比对站点连接以进行开始于一套相关蛋白质序列的杂交引物预测, 如已知的味觉受体配体结合区域 (参见如 Rose, Nucleic Acids Res. 26:1628-1635(1998); Singh, Biotechniques 24:318-319(1998))。

[0128] 合成寡核苷酸引物对的方法是本领域所公知的。可使用“天然”的碱基对或合成的碱基对。例如, 人工核碱基 (nucleobases) 的应用提供了通用的方法以处理引物序列和生成更复杂的扩增产物混合物。多种人工核碱基的家族能够通过内部键旋转而实现多

重氢键定向以提供简并分子识别的方法。将这些类似物整合入 PCR 引物的单一位置使得能生成扩增产物的复杂文库。参见如 Hoops, *Nucleic Acids Res.* 25 :4866-4871(1997)。非极性分子也可用于模拟天然 DNA 碱基的形状。对腺嘌呤的非氢键形状模拟可对胸腺嘧啶的非极性形状模拟进行有效和选择性的复制(参见如 Morales, *Nat. Struct. Biol.* 5 :950-954(1998))。例如,2 个简并碱基可为嘧啶碱基 6H,8H-3,4-二氢嘧啶 [4,5-c][1,2] 嗪-7-酮或嘌呤碱基 N6-甲氧基-2,6-二氨基嘌呤(参见如 Hill, *Proc. Natl. Acad. Sci. USA* 95 :4258-4263(1998))。本发明的示范性简并引物包含核碱基类似物 5'-二甲氧基三苯甲基-N-苯甲酰-2'-脱氧胞嘧啶核苷,3'-[(2-氰乙基)-(N,N-二异丙基)]-亚磷酰胺(序列中的术语“P”,见上述)。该嘧啶类似物与嘌呤包括 A 和 G 残基形成氢键。

[0129] 与在此处所公开的味觉受体基本相同的多态变体、等位基因和种间同系物可用上述的核酸探针进行分离。可选择地,可将表达文库用于克隆 T1R 多肽和其多态变体、等位基因和种间同系物,这是通过用抗 T1R 多肽的抗血清或纯化的抗体免疫探测表达的同系物而完成的,其中的抗血清或纯化的抗体也识别和选择性地结合 T1R 同系物。

[0130] 编码味觉受体配体结合区域的核酸可通过用简并引物对扩增(如 PCR)适当的核酸序列而生成。扩增的核酸可为来自任何细胞或组织的基因组 DNA 或源自味觉受体表达细胞的 mRNA 或 cDNA。

[0131] 在一个实施方案中,可构建融合于转运基因的、包含核酸的杂交蛋白质编码序列,该核酸编码 T1Rs。也提供的是包含其他化学感受性受体家族特别是味觉受体的味觉刺激物结合结构域和转运基元的杂交 T1Rs。这些核酸序列可被可操作地连接到转录或翻译控制序列上,如转录和翻译起始序列、启动子和增强子、转录和翻译终止子、聚腺苷酸化序列和其他对于将 DNA 转录为 RNA 有用的序列。在重组表达盒、载体和转基因的构建中,可采用启动子片段以指导想要的核酸在所有想要的细胞或组织中表达。

[0132] 在另一个实施方案中,融合蛋白质可包括 C-末端或 N-末端转运序列。进一步地,融合蛋白质可包含额外的元件以进行如蛋白质探测、纯化或其他应用。探测和纯化促进结构域包括如金属螯合肽如多组氨酸序列、组氨酸-色氨酸组件或其他允许在固定的金属上纯化的结构域;麦芽糖结合蛋白质;允许在固定的免疫球蛋白上纯化的蛋白质 A 结构域;或用于 FLAGS 延伸/亲和纯化系统的结构域(Immunex Corp, Seattle WA)。

[0133] 在转运结构域(为了有效的质膜表达)和新翻译多肽的剩余部分之间包含可切割的连接体序列如因子 Xa(参见如 Ottavi, *Biochimie* 80 :289-293(1998))、枯草杆菌蛋白酶识别基元(参见如 Polyak, *Protein Eng.* 10 :615-619(1997))、肠激酶(Invitrogen, San Diego, CA)等对于促进纯化是有用的。例如,一个构建体可包括一个核酸序列编码的连接到 6 个组氨酸残基上的多肽,随后是硫氧还蛋白、肠激酶切割位点(参见如 Williams, *Biochemistry* 34 :1787-1797(1995))和 C-末端转运结构域。组氨酸残基促进探测和纯化,而肠激酶切割位点提供了将想要的蛋白质从剩余的融合蛋白质中纯化的方法。涉及编码融合蛋白质的载体和融合蛋白质应用的技术是在科学和专利文献中详尽描述的,参见如 Kroll, *DNA Cell Biol.* 12 :441-53(1993)。

[0134] 包含配体结合结构域编码序列的单个表达载体或表达载体文库可通过多种常规的技术而引入到细胞的基因组或细胞质或细胞核中并表达,该技术在科学和专利文献中有详尽描述。参见如 Roberts, *Nature* 328 :731(1987); 上述的 Berger; Schneider, *Protein*

Expr. Purif. 6435 :10(1995) ;Sambrook ;Tijssen ;Ausubel。来自生物学试剂和实验设备的生产商的产品信息也提供了关于已知生物学方法的信息。该载体可从天然来源分离、从如 ATCC 或 GenBank 文库的来源获得或通过合成或重组的方法制备。

[0135] 核酸可在细胞内稳定或瞬时表达的表达盒、载体或病毒中表达（如附加体表达系统）。可将选择性标记整合到表达盒和载体中以赋予转化的细胞和序列选择性的表型。例如，选择性标记可编码附加体的维持和复制，从而整合入宿主基因组不是必需的。例如，标记可编码抗生素抗性（如氯霉素、卡那霉素、G418、博来霉素、潮霉素）或除草剂抗性（如氯磺隆 (chlorosulfuron) 或 Basta）以允许对那些用想要的 DNA 转化了的细胞进行选择（参见如 Blondelet-Rouault, Gene 190 :315-317(1997) ;Aubrecht, J. Pharmacol. Exp. Ther. 281 :992-997(1997)）。因为赋予对物质如新霉素或潮霉素抗性的选择性标记基因仅能用于组织培养中，所以化学抗性基因也用作体外或体内的选择性标记。

[0136] 嵌合的核酸序列可编码任何七跨膜多肽中的 T1R 配体结合结构域。因为七跨膜受体多肽具有相似的一级序列以及二级和三级结构，所以其结构上的结构域（如细胞外结构域、TM 结构域、胞质结构域等）易于通过序列分析而鉴定。例如，同源性模拟、傅立叶分析和螺旋周期性探测可鉴定和表征七跨膜受体序列的 7 个结构域。快速傅立叶变换 (FFT) 算法可用于估计主要的周期，该周期表征了分析的序列的疏水性和变异性。周期性探测增强和 α 螺旋周期性指数可通过如 Donnelly, Protein Sci. 2 :55-70(1993) 所述的来进行。其他比对和模拟算法是本领域众所周知的，参见如 Peitsh, Receptors Channels 4 :161-164(1996) ;Kyte&Doolittle, J. Md. Bio. 157 :105-132(1982) ;Cronet, Protein Eng. 6 :59-64(1993)（同源性和“发现模拟”）；<http://bioinfo.weizmann.ac.il/>。

[0137] 本发明不仅包括具有特定的核酸和氨基酸序列的 DNA 和蛋白质，也包括 DNA 片段特别是如 40、60、80、100、150、200 或 250 个或更多核苷酸的片段，以及如 10、20、30、50、70、100 或 150 个或更多氨基酸的蛋白质片段。可选择地，该核酸序列可编码抗原性多肽，该多肽能够结合抗 T1R 家族成员的抗体。进一步地，本发明的蛋白质片段可选择地为抗原性片段，该片段能够结合抗 T1R 家族成员的抗体。

[0138] 也涉及的是一种嵌合蛋白质，该蛋白质包含此处所描述的一个或至少一个 T1R 多肽的至少 10、20、30、50、70、100 或 150 个或更多氨基酸，这些氨基酸是与代表另一种 GPCR 的整体或部分的额外氨基酸相连结的，其中 GPCR 优选地为七跨膜家族的一个成员。这些嵌合体可从本发明的受体和另一种 GPCR 来制备，或可通过组合两种或多种本发明的受体而制备。在一个实施方案中，嵌合体的一个部分相应于或源自本发明的 T1R 多肽的细胞外结构域。在另一个实施方案中，嵌合体的一个部分相应于或源自此处所描述的 T1R 多肽的细胞外结构域和一个或多个跨膜结构域，且剩余的部分可来自另一种 GPCR。嵌合受体是本领域众所周知的，且创造该受体的技术和用于在其中整合的 G 蛋白 - 偶联受体的结构域或片段的选择和边界也是公知的。因而，本领域技术人员的这些知识可易于用于创造这种嵌合受体。这种嵌合受体的用途可提供如一种此处特定地公开的受体的味觉选择特征，该特征与另一种受体（如在现有技术的测定系统中所用的公知的受体）的信号转导特征偶联。

[0139] 例如，可将一个结构域共价地连接到异源蛋白质上，该结构域可以是例如配体结合结构域、细胞外结构域、跨膜结构域、跨膜结构域、胞质结构域、N-末端结构域、C-末端结构域或其任意组合。例如，T1R 细胞外结构域可连接到异源 GPCR 跨膜结构域上，或异源的

GPCR 胞外结构域可连接到 T1R 跨膜结构域上。其他选择的异源蛋白质可包括如绿色荧光蛋白、 β -gal、谷氨酸 (glutamate) 受体和视紫红质前序列。

[0140] 也在本发明范围之内的是表达本发明的 T1Rs、片段或变体的宿主细胞。为了获得克隆的基因或核酸的高水平表达,如编码本发明的 T1Rs、片段或变体的 cDNA,技术人员一般将目标核酸序列亚克隆入表达载体中,该表达载体含有指导转录的强启动子、转录/翻译终止子和(如果对于编码蛋白质的核酸时)转录起始的核糖体结合位点。适当的细菌启动子是本领域公知的并描述于如 Sambrook 等人。然而,细菌和真核表达系统均可使用。

[0141] 可使用任何众所周知的将外源核苷酸序列引入宿主细胞的程序。这些包括磷酸钙转染、1,5-二甲基-1,5-二氮十一亚甲基聚甲溴化物、原生质体融合、电穿孔、脂质体、微注射、原生质载体、病毒载体和任何其他众所周知的将克隆的基因组 DNA、cDNA、合成的 DNA 或其他外源遗传材料引入宿主细胞的方法(参见如 Sambrook 等人)。仅仅需要所用的特殊的基因工程程序能够成功地将至少一个核酸分子引入到能够表达目标 T1R、片段或变体的宿主细胞中。

[0142] 在表达载体引入到细胞中后,将转染的细胞培养于有利于目标受体、片段或变体表达的条件下,然后将它们用标准技术从培养物中回收。这种技术的例子是本领域众所周知的。参见如 W000/06593,它是以与本公开内容一致的方式引用作为参考的。

[0143] B. T1R 多肽的探测

[0144] 除了用核酸杂交技术探测 T1R 基因和基因表达之外,人们也可用免疫测定来探测 T1Rs,如鉴定味觉受体细胞和 T1R 家族成员的变体。免疫测定也可用于定性或定量地分析 T1Rs。对可采用的技术的一般综述可发现于 Harlow&Lane, *Antibodies: A Laboratory Manual* (1988)。

[0145] 1. 针对 T1R 家族成员的抗体

[0146] 生产特异性地与 T1R 家族成员反应的多克隆和单克隆抗体的方法是本领域的技术人员公知的(参见如 Coligan, *Current Protocols in Immunology* (1991); 上述的 Harlow&Lane; Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed. 1986); 和 Kohler&Milstein, *Nature*, 256:495-497 (1975))。这种技术包括通过从噬菌体或类似载体中的重组抗体文库中选择抗体的抗体制备,以及通过对兔或小鼠进行免疫而制备多克隆和单克隆抗体(参见如 Huse 等人, *Science*, 246:1275-1281 (1989); Ward 等人, *Nature*, 341:544-546 (1989))。

[0147] 大量包含 T1R 的免疫原可用于生产与 T1R 家族成员特异性反应的抗体。例如,重组 T1R 多肽或其抗原性片段可如此处所描述地进行分离。适当的抗原性区域包括如用于鉴定 T1R 家族成员的一致序列。重组蛋白质可如上所述表达于真核或原核细胞中,并如在上文一般描述的进行纯化。重组蛋白质是用于生产单克隆或多克隆抗体的优选免疫原。可选择地,源自此处所公开的序列并与载体蛋白质缀合的合成肽可用作免疫原。天然存在的蛋白质也可以以纯的或不纯的形式使用。然后将该产物注射入能够生产抗体的动物中。可生成单克隆或多克隆抗体以随后用于免疫测定以测量蛋白质。

[0148] 生产多克隆抗体的方法是本领域技术人员所公知的。例如,将小鼠(如 BALB/C 小鼠)或兔的近交品系用标准的佐剂如弗氏佐剂和标准的免疫接种规程用蛋白质进行免疫接种。动物对免疫原制剂的免疫反应是通过采取血样和确定对 T1R 的反应性效价而监

控的。当获得了适当高效价的对免疫原的抗体时,从动物中收集血液并制备抗血清。如果需要,可对抗血清进行进一步的分级分离以富集对蛋白质起反应的抗体(参见上述的 Harlow&Lane)。

[0149] 单克隆抗体可通过多种本领域技术人员所熟悉的技术获得。简要地,通常可通过与骨髓瘤细胞融合而使来自用想要的抗原免疫接种的动物的脾细胞永生化(参见 Kohler&Milstein, Eur. J. Immunol., 6:511-519(1976))。可选择的永生化方法包括用 EB 病毒、癌基因或反转录病毒进行转化或用其他本领域众所周知的方法。将从单一永生化细胞发生的集落进行筛选以生产具有想要的对抗原的特异性和亲和力的抗体,且由这种细胞生产的单克隆抗体的产量可通过多种途径来增强,包括注射入脊椎动物宿主的腹膜腔内。可选择地,人们可通过根据一般规程筛选来自人 B 细胞的 DNA 文库而分离编码单克隆抗体或其结合片段的 DNA 序列,该规程概述于 Huse 等人, Science, 246:1275-1281(1989)。

[0150] 在一个免疫测定中,如免疫原固定于固体基质上的固相免疫测定中,收集单克隆抗体和多克隆血清并测定其对免疫原蛋白质的效价。一般地,将效价为 104 或更高的多克隆抗血清选择出来,并用竞争性结合免疫测定来检验其对非-T1R 多肽,或者甚至其他 T1R 家族成员或来自其他生物的其他相关蛋白质的交叉反应性。特异性的多克隆抗血清和单克隆抗体通常以至少约 0.1mM 的 Kd 结合,更通常地为至少 1pM,可选择地为至少约 0.1pM 或更好,更通常地为 0.01pM 或更好。

[0151] 一旦 T1R 家族成员特异性抗体为可获得的,那么单独的 T1R 蛋白质和蛋白质片段可通过多种免疫测定方法进行探测。对于免疫学和免疫测定程序的综述,参见 Basic and Clinical Immunology(Stites&Terr eds., 7th ed. 1991)。此外,本发明的免疫测定可以以任何几个形式来进行,这在 Enzyme Immunoassay(Maggio, ed., 1980)和上述的 Harlow&Lane 中有详尽的综述。

[0152] 2. 免疫学结合测定

[0153] T1R 蛋白质、片段和变体可用大量公认的免疫学结合测定中的任何一个进行探测和/或定量(参见如美国专利 4,366,241;4,376,110;4,517,288 和 4,837,168)。对于一般的免疫测定的综述,也参见 Methods in Cell Biology:Antibodies in Cell Biology,卷 37(Asai, ed. 1993);Basic and Clinical Immunology(Stites&Terr eds., 7th ed. 1991)。免疫学结合测定(或免疫测定)一般使用特异性结合到选择的蛋白质或抗原(在这个情况下为 T1R 家族成员或其抗原性序列)上的抗体。抗体(如抗-T1R)可通过大量本领域技术人员众所周知且如上所述的方法的任何一种来生产。

[0154] 免疫测定经常也用标记试剂以特异性地结合并标记由抗体和抗原形成的复合物。该标记试剂自身可为包含抗体/抗原复合物的部分的一种。因而,标记试剂可为标记的 T1R 多肽或标记的抗-T1R 抗体。可选择地,该标记试剂可为第三种特异性地结合到抗体/T1R 复合物的部分如二级抗体(二级抗体一般是对一级抗体从中获得的物种的抗体具有特异性的)。其他能够特异性结合免疫球蛋白恒定区的蛋白质如蛋白质 A 或蛋白质 G 也可用作标记试剂。这些蛋白质显示与来自多种物种的免疫球蛋白恒定区的强的非免疫原性反应性(参见如 Kronval 等人, J. Immunol. 111:1401-1406(1973);Akerstrom 等人, J. Immunol. 135:2589-2542(1985))。该标记试剂可用可探测的部分如生物素进行修饰,其他的分子如链霉抗生物素蛋白可结合到该部分上。多种可探测的部分是本领域技术人员众所周知的。

[0155] 在整个测定过程中,在每个试剂组合之后均需要温育和 / 或洗涤步骤。温育步骤可为 5 秒钟~几小时,可选择地为约 5 分钟~约 24 小时。然而,温育时间将依赖于测定形式、抗原、溶液体积、浓度等。通常,该测定将在环境温度进行,尽管它们可在如 10℃~40℃ 的温度范围内进行。

[0156] a. 非竞争性测定形式

[0157] 在样品中探测 T1R 多肽的免疫测定可为竞争性的或非竞争性的。非竞争性的免疫测定是其中抗原的量是直接测量的测定。如在一个优选的“三明治 (sandwich)”测定中,抗-T1R 抗体可直接结合到固体基质上,在其上它们得以固定。这些固定的抗体然后捕获存在于检验样品中的 T1R 多肽。这样固定的 T1R 多肽然后由标记试剂如带有标记的二级 T1R 抗体进行结合。可选择地,二级抗体可缺乏标记,但它将由标记的三级抗体进行结合,该三级抗体是对二级抗体从中获得的物种的抗体具有特异性的。二级或三级抗体一般可用可探测的部分如生物素进行修饰,其他的分子如链霉抗生物素蛋白可特异性地结合到该部分上以提供可探测的部分。

[0158] b. 竞争性测定形式

[0159] 在竞争性测定中,存在于样品中的 T1R 多肽的量是通过测量被置换的已知的添加的 (外源的) T1R 多肽的量而间接测量的,其中的多肽是由存在于样品中的未知 T1R 多肽从抗-T1R 抗体中置换 (竞争掉) 的。在一个竞争性测定中,将已知量的 T1R 多肽添加入样品中,然后使样品与特异性结合 T1R 的抗体接触。结合到抗体的外源 T1R 多肽的量与存在于样品中的 T1R 多肽的浓度成反比。在一个特别优选的实施方案中,该抗体是固定到固体基质上的。结合到抗体上的 T1R 多肽的量可通过测量存在于 T1R/ 抗体复合物中的 T1R 多肽的量而确定,或者可选择地通过测量剩余的未复合的蛋白质的量而确定。T1R 多肽的量可通过提供标记的 T1R 分子而探测。

[0160] 半抗原抑制测定是另一个优选的竞争性测定。在该测定中,已知的多肽是固定于固体基质上的。将抗-T1R 抗体的已知量添加到样品中,然后使样品与固定的 T1R 接触。结合到已知固定的 T1R 多肽的抗-T1R 抗体的量与存在于样品中的 T1R 多肽的量成反比。再次地,固定的抗体的量可通过探测抗体的固定比例或剩余在溶液中的抗体的比例来探测。该探测在抗体进行了标记时可为直接的,或者通过随后添加标记的部分而为间接的,该部分特异性地结合上述的抗体。

[0161] c. 交叉反应性的确定

[0162] 竞争性结合形式的免疫测定也可用于交叉反应性的确定。例如,至少部分地由此处所公开的核酸序列编码的蛋白质可固定于固体基质上。将与抗血清对固定的抗原的结合进行竞争的蛋白质 (如 T1R 多肽和同系物) 添加到测定体系中。该添加的蛋白质与抗血清对固定的抗原的结合进行竞争的能力是与由此处所公开的核酸序列编码的 T1R 多肽与其自身竞争的能力进行比较的。利用标准的计算对上述蛋白质的交叉反应性百分数进行了计算。将那些与上面列出的每一种添加的蛋白质具有少于 10% 的交叉反应性的抗血清选择并收集。交叉反应的抗体是可选择地通过用添加的慎重考虑的蛋白质 (如远缘相关的同系物) 的免疫吸附从收集的抗血清中去除的。另外,包含代表保守基元的氨基酸序列的肽可用于交叉反应性的确定,其中的保守基元用于鉴定 T1R 家族的成员。

[0163] 然后将免疫吸附和收集的抗血清用于上述的竞争性结合免疫测定中以将第二个

蛋白质与免疫原蛋白质（即由此处所公开的核酸序列所编码的 T1R 多肽）进行比较，其中的第二种蛋白质被认为可能是 T1R 家族成员的等位基因或多态变体。为了进行该比较，将两种蛋白质各自在广泛的浓度范围内进行测定，且确定了能抑制抗血清与固定的蛋白质 50% 的结合所需的每种蛋白质的量。如果抑制 50% 的结合所需的第二种蛋白质的量比抑制 50% 的结合所需的由此处公开的核酸序列编码的蛋白质的量少 10 倍，那么即认为该第二种蛋白质特异性地结合抗 T1R 免疫原的多克隆抗体。

[0164] 抗 T1R 保守基元的抗体也可用于制备仅仅特异性地结合 T1R 家族的 GPCRs 而不结合来自其他家族的 GPCRs 的抗体。

[0165] 特异性结合 T1R 家族的特定成员的多克隆抗体可通过用其他的 T1R 家族成员减除交叉反应性抗体而制备。物种特异性多克隆抗体可用相似的途径制备。例如，对人 T1R1 特异性的抗体可通过减除那些与定向同源序列如大鼠 T1R1 或小鼠 T1R1 交叉反应的抗体而制备。

[0166] d. 其他测定形式

[0167] 蛋白质印迹（免疫印迹）分析用于探测和量化 T1R 多肽在样品中的存在。该技术一般包含通过凝胶电泳在分子量基础上分离样品蛋白质、将分离的蛋白质转移到适当的固体基质上（如硝酸纤维素滤膜、尼龙滤膜或衍生的尼龙滤膜）及将样品与特异性结合 T1R 多肽的抗体进行温育。抗-T1R 多肽抗体特异性地结合固体基质上的 T1R 多肽。这些抗体可直接标记或可选择地随后用标记的抗体（如标记的绵羊抗小鼠抗体）进行探测，该标记的抗体特异性地结合抗-T1R 抗体。

[0168] 其他的测定形式包括脂质体免疫测定（LIA），该测定使用设计为结合特定分子（如抗体）和释放所包装的试剂或标记的脂质体。然后释放的化学药剂可根据标准的技术而探测（参见 Monroe 等人，*Amer. Clin. Prod. Rev.* 5 :34-41 (1986)）。

[0169] e. 非特异性结合的减少

[0170] 本领域的技术人员将理解经常期望将免疫测定中的非特异性结合减到最小。特别地，在测定包含固定于固体基质上的抗原或抗体时，即期望对底物非特异性结合的量减到最小。减小这种非特异性结合的方法是本领域技术人员众所周知的。一般地，该技术包含用蛋白质组合物包被底物。具体而言地，蛋白质组合物如牛血清白蛋白（BSA）、脱脂奶粉和明胶是广泛使用的，其中奶粉是最优选的。

[0171] f. 标记

[0172] 在测定中所用的特殊的标记或可探测的基团不是本发明的关键方面，只要它不显著地干扰测定中所用的抗体的特异性结合。可探测的基团可为任何具有可探测的物理或化学性质的材料。这种可探测的标记已经在免疫测定领域得到充分发展，并且通常这种方法中的大多数标记都可应用于本发明。因而，标记是任何用光谱学、光化学、生物化学、免疫化学、电学、光学或化学方法可探测的组合物。本发明中有用的标记包括磁珠（如 DYNABEADSTM）、荧光染料（如异硫氰酸荧光素、德克萨斯红、罗丹明等）、放射性标记（如 ^3H 、 ^{125}I 、 ^{35}S 、 ^{14}C 或 ^{32}P ）、酶（如辣根过氧化物酶、碱性磷酸酶和其他通常用于 ELISA 的）和比色法标记如胶体金或有色玻璃或塑料珠（如聚苯乙烯、聚丙烯、胶乳等）。

[0173] 该标记可根据本领域众所周知的方法直接或间接与测定的想要成分偶联。如上所述，可使用多种标记，对标记的选择依赖于需要的灵敏度、与化合物缀合的容易程度、稳定

性需要、可获得的仪器和处理规定。

[0174] 非放射性的标记经常通过间接方法附着。通常,配体分子(如生物素)是共价结合到分子上的。该配体然后结合到另一个分子(如链霉抗生物素蛋白),该分子本身是可探测的,或者是共价地结合到信号系统上的,如可探测的酶、荧光化合物或化学发光化合物上。该配体及其目标可用于与识别 T1R 多肽的抗体或识别抗-T1R 的二级抗体的任何组合。

[0175] 该分子也可直接缀合到信号生成化合物上,如通过与酶或荧光团缀合。作为标记的目标酶主要为水解酶,具体而言地为磷酸酶、酯酶和糖苷酶,或者氧化酶(oxidotase),具体而言地为过氧化物酶。荧光化合物包括荧光蛋白及其衍生物、罗丹明及其衍生物、丹磺酰、伞形酮等。化学发光化合物包括萤光素和 2,3-二氢二氮杂萘二酮类(2,3-dihydrophthalazinediones),如鲁米诺。对于可用的多种标记或信号生产系统,参见美国专利 No. 4,391,904。

[0176] 探测标记的方法是本领域技术人员众所周知的。因而,例如当标记是放射性标记时,探测手段包括闪烁计数器或放射自显影术中的照相胶卷。当标记是荧光标记时,它可通过用适当波长的光激发荧光染料并探测结果所得的荧光。该荧光可用肉眼探测、用照相胶卷的方法探测、用电子探测器如电荷偶联装置(CCDs)或光电倍增管等进行探测。相似地,酶标记可通过提供适当的酶底物并探测结果所得的反应产物而探测。最后简单的比色标记可简单地通过观察与标记有关的颜色而探测。因而,在多种量油计(dipstick)测定中,缀合的金经常显示粉色,而多种缀合的珠子显示珠子的颜色。

[0177] 一些测定形式不需要使用标记的成分。例如,凝集测定可用于探测目标抗体的存在。在这种情况下,抗原包被的颗粒被包含目标抗体的样品凝集。在这种形式中,没有一种成分需要进行标记,且目标抗体的存在是通过简单的肉眼观测探测的。

[0178] C. 调节子的探测

[0179] 在下面描述了在体外或体内用于确定检验的化合物是否特异性地结合本发明的化学感受性受体的组合物和方法。可对细胞生理学的许多方面进行监控以估计配体结合到本发明的 T1R 多肽的作用。这些测定可在表达化学感受性受体的完整细胞上、在透化的细胞上或在由标准方法生产的膜组分上进行。

[0180] 味觉受体结合味觉刺激物并起始化学刺激物向电信号的转导。活化的或抑制的 G 蛋白将改变目标酶、通道和其他效应蛋白质的性质。一些例子为视觉系统中的转导素对 cGMP 磷酸二酯酶的活化、刺激性 G 蛋白对腺苷酸环化酶的活化、Gq 和其他同类 G 蛋白对磷脂酶 C 的活化及 Gi 和其他 G 蛋白对各种通道的调节。也可检查下游的结果如由磷脂酶 C 生成二酰甘油和 IP3 及接着 IP3 对钙的动员。

[0181] 测定中的 T1R 蛋白质或多肽一般选自具有 SEQ ID NOS :4、10、12、14、17、21 或其片段或保守修饰的变体的序列的多肽。可选择地,该片段和变体可为结合到抗-T1R 抗体上的抗原性片段和变体。

[0182] 可选择地,测定中的 T1R 蛋白质或多肽可源自真核生物宿主细胞,并可包括与 SEQ ID NOS :4、10、12、14、17、21 或其片段或保守修饰的变体具有氨基酸序列一致性的氨基酸子序列。通常,该氨基酸序列一致性将为至少 35 ~ 50%,或可选择地为 75%、85%、90%、95%、96%、97%、98%或 99%。可选择地,测定中的 T1R 蛋白质或多肽可包含 T1R 蛋白质的一个结构域,如细胞外结构域、跨膜区域、跨膜结构域、胞质结构域、配体结合结构域等。进

一步地,如上所述,该 T1R 蛋白质或其结构域可共价地连接到异源蛋白质上以创建用于此处所描述的测定中的嵌合蛋白质。

[0183] T1R 受体活性的调节子是用上述重组的或天然存在的 T1R 蛋白质或多肽进行检验的。可将重组的或者天然存在的 T1R 蛋白质分离,然后在细胞中表达、在源自细胞的膜中表达、在组织或动物中表达。例如,可使用舌切片、从舌解离的细胞、转化的细胞或膜。该调节可利用此处所描述的体外或体内测定的任何一种来检验。

[0184] 1. 体外结合测定

[0185] 味觉转导也可利用本发明的 T1R 多肽以可溶性的或固体状态的反应进行检查。在一个特殊的实施方案中, T1R 配体结合结构域可在体外用于可溶性的或固体状态的反应以测定配体的结合。

[0186] 例如,预测 T1R 的 N-末端结构域参与配体的结合。更具体而言地, T1R 属于 GPCR 亚家族,该亚家族特征在于具有大的约 600 个氨基酸的细胞外 N-末端区段。认为这些 N-末端区段至少部分地形成配体结合结构域,因此可用于生物化学测定中以鉴定 T1R 激动剂和拮抗剂。该配体结合结构域也可含有另外的细胞外结构域部分,如跨膜结构域的细胞外环。对其他与 T1Rs 相关的 GPCRs,如 metabotropic 谷氨酸受体已采用了相似的测定(参见如 Han 和 Hampson, J. Biol. Chem. 274:10008-10013(1999))。这些测定可包含置换放射性或荧光标记的配体、测量内部荧光的变化或蛋白水解易感性的变化等。

[0187] 结合到本发明的 T1R 多肽的配体可在溶液中、在(可选择地附着到固相上)双层膜中、在脂单层膜或小泡中进行检验。调节子的结合可用如光谱学特征(如荧光、吸收、折射率)、流体动力学(如形状)、色谱或溶解性性质的变化来检验。本发明优选的结合测定是使用重组的可溶性 N-末端 T1R 结构域的生物化学结合测定。

[0188] 也可检查受体-G蛋白的相互作用。例如,可检查 G 蛋白与受体的结合,或它从受体的释放。更具体而言地,在不存在 GTP 时,激活剂可导致 G 蛋白(所有三个亚基)与受体的紧密复合物的形成。该复合物可用上述的多种途径进行探测。可对这种测定进行修饰以搜索抑制剂,如通过在不存在 GTP 时向受体和 G 蛋白质添加一种激活剂,其中的受体和 G 蛋白形成紧密的复合物,然后通过查看受体-G蛋白复合物的解离来筛选抑制剂。在 GTP 存在时,G蛋白的 α 亚基从其他 2 个 G 蛋白亚基的释放充当激活的判别准则。活化的或抑制的 G 蛋白将反过来改变目标酶、通道和其他效应蛋白质的性质。

[0189] 在本发明的另一个实施方案中,可以采用 GTP γ S 测定。如上所述,在 GPCR 活化后,G蛋白复合物的 G α 亚基即受到刺激以将结合的 GDP 交换为 GTP。对 G 蛋白交换活性的配体介导的刺激可在生物化学测定中进行测量,该测定测量在推定的配体存在时添加的放射性标记的 GTP γ 35S 与 G 蛋白的结合。一般地,含有目标化学感受性受体的膜是与 G 蛋白复合物混合的。将潜在的抑制剂和/或激活剂和 GTP γ S 添加入测定体系中,并测量 GTP γ S 与 G 蛋白的结合。该结合可通过液体闪烁计数或任何其他本领域公知的方法来测量,包括闪烁亲近测定法(SPA)。在其他的测定形式中,可使用荧光标记的 GTP γ S。

[0190] 2. 荧光偏振测定

[0191] 在另一个实施方案中,基于荧光偏振(“FP”)的测定可用于探测和监控配体的结合。荧光偏振是一种通用的实验室技术,该技术用于测量平衡结合、核酸杂交和酶活性。荧光偏振测定由于它不需要分离步骤如离心、过滤、色谱、沉淀或电泳而是均一的。这些测定

是实时地直接在溶液中进行的,且不需要固定相。偏振值可重复地并在添加试剂之后进行测量,这是因为偏振的测定是快速的且不破坏样品。通常,该技术可用于测量低的皮摩尔至微摩尔水平的荧光团偏振值。本部分描述了荧光偏振怎样以简单和定量的途径测量配体与本发明的 T1R 多肽的结合。

[0192] 当将荧光标记的分子用平面偏振光进行激发时,它释放具有与其摩尔旋光度成反比的偏振程度的光。大的荧光标记分子在激发状态(在荧光蛋白的情况下为 4 纳秒)保持相对的静止,且光的偏振在激发和发射之间保持相对的恒定。小的荧光标记分子在激发状态快速旋转,且偏振在激发和发射之间显著地改变。因此,小分子具有低的偏振值而大分子具有高的偏振值。例如,单链荧光蛋白标记的寡核苷酸具有相对低的偏振值,但当它与互补链杂交之后,则具有较高的偏振值。当用 FP 来探测和监控味觉刺激物的结合时(该结合可激活或抑制本发明的化学感受性受体),可以使用荧光标记的味觉刺激物或自发荧光的味觉刺激物。

[0193] 荧光偏振(P)定义为:

$$[0194] \quad P = \frac{Int_{\parallel} - Int_{\perp}}{Int_{\parallel} + Int_{\perp}}$$

[0195] 其中 $\parallel$ 是与激发光平面平行的发射光的强度,且 $\perp$ 是与激发光平面垂直的发射光的强度。作为光强度比例的 P 是一个无量纲的数。例如,Beacon®和 Beacon 2000™系统可连同这些测定使用。这种系统一般以毫偏振单位来表示偏振(1 偏振单位 = 1000mP 单位)。

[0196] 摩尔旋光度和大小的关系由 Perrin 方程描述,读者可参考 Jolley, M. E. (1991) 在杂志 Journal of Analytical Toxicology, PP. 236-240 的文章,该文章对方程做了详尽的解释。总之, Perrin 方程表明偏振是直接与旋转弛豫时间(rotational relaxation time)成反比的,该时间为一个分子旋转约 68.5° 所需的时间。旋转弛豫时间通过下面的方程与粘度(η)、绝对温度(T)、分子体积(V)和气体常数(R)相关:

[0197]

$$\text{旋转弛豫时间} = \frac{3\eta V}{RT}$$

[0198] 旋转弛豫时间对于小分子(如荧光蛋白)是小的(≈ 1 纳秒),而对于大分子(如免疫球蛋白)是大的(≈ 100 纳秒)。如果粘度和温度保持恒定,那么旋转弛豫时间和由此的偏振则直接与分子体积相关。分子体积的变化可归因于与其他分子的相互作用、解离、多聚化、降解、杂交或荧光标记分子的构象变化。例如,荧光偏振已用于测量由蛋白酶、DNases 和 RNases 对大的荧光蛋白标记的多聚体的酶切割。它也已用于测量蛋白质/蛋白质相互作用、抗体/抗原结合和蛋白质/DNA 结合的平衡结合。

[0199] 3. 固相和可溶性的高生产量的测定

[0200] 在另外一个实施方案中,本发明提供了用 T1R 多肽的可溶性测定;或者表达 T1R 多肽的细胞或组织。在另一个实施方案中,本发明以高生产量的形式提供了基于固相的体外测定,其中 T1R 多肽或表达 T1R 多肽的细胞或组织是附着在固相基质上的。

[0201] 在本发明的一个高生产量的测定中,在一天内筛选达几千个不同的调节子或配体是可能的。特别地,微量滴定板上的每一个孔都可用于进行对选择的潜在调节子的单独测

定,或者如果要观察浓度或温育时间的效应,每 5-10 个孔可检验一个调节子。因而,一个标准的微量滴定板可测定约 100 (如 96) 个调节子。如果使用 1536 个孔的板,那么在一个板上可容易地测定约 1000 ~ 约 1500 个不同的化合物。也可能在每个板孔中测定多个混合物。每天测定几个不同的板是可能的;利用本发明的整合系统测定筛选多达约 6,000 ~ 20,000 个不同的化合物是可能的。

[0202] 目标分子可直接或间接地通过共价键或非共价键如通过标记而结合到固相成分上。该标记可为多种成分的任何一种。通常,将结合标记的分子(标记粘合剂)固定在固体基质上,然后将标记的目标分子(如目标味觉转导分子)通过标记和标记粘合剂的相互作用而附着于固体基质上。

[0203] 基于在文献中详尽描述的分子间相互作用,可采用大量的标记和标记粘合剂。例如,当标记具有天然的粘合剂如生物素、蛋白质 A 或蛋白质 G 时,它可与适当的标记粘合剂(抗生物素蛋白、链霉抗生物素蛋白、neutravidin、免疫球蛋白的 Fc 区域等)一起使用。对具有天然粘合剂如生物素的分子的抗体也是广泛可用的,并且也是适当的标记粘合剂(参见 SIGMA Immunochemicals 1998 catalogue SIGMA, St. Louis MO)。

[0204] 相似地,任何半抗原或抗原化合物可与适当的抗体一起使用以形成标记/标记粘合剂对。数千种特异性的抗体是商业可获得的且许多额外的抗体在文献中有描述。例如,在一个普通的形式中,标记是一级抗体而标记粘合剂是识别该一级抗体的二级抗体。除抗体-抗原相互作用之外,受体-配体相互作用也适合于做标记和标记粘合剂对。例如,细胞膜受体的激动剂和拮抗剂(如细胞受体-配体相互作用如运铁蛋白、c-kit、病毒受体配体、细胞因子受体、趋化因子受体、白细胞介素受体、免疫球蛋白受体和抗体、钙粘着蛋白(cadherein) 家族、整联蛋白家族、选择蛋白家族等;参见如 Pigott & Power, The Adhesion Molecule Facts Book I (1993))。相似地,毒素和毒液、病毒抗原决定部位、激素(如阿片剂、类固醇等)、细胞内受体(如介导多种小配体作用的,该配体包括类固醇、甲状腺激素、类视黄醇和维生素 D;肽)、药物、凝集素、糖类、核酸(线状和环状多聚体构型两者)、寡糖、蛋白质、磷脂和抗体都可与多种细胞受体相互作用。

[0205] 合成的多聚体如聚氨基甲酸酯、聚酯、聚碳酸酯、聚脲、聚酰胺、聚乙烯亚胺、聚芳撑硫化物(polyarylene sulfides)、聚硅氧烷、聚酰亚胺和聚乙酸酯也可形成适当的标记或标记粘合剂。许多其他的标记/标记粘合剂对也可用于此处所描述的测定系统中,基于本公开内容的综述,这对技术人员显而易见的。

[0206] 普通的连接体如肽、聚醚等也可充当标记,并包括多肽序列,如约 5 ~ 200 个氨基酸的聚 gly 序列。这种灵活的连接体是本领域技术人员所公知的。例如,聚乙二醇连接体可得自 Shearwater Polymers, Inc. Huntsville, Alabama。这些连接体可选择地具有酰胺键、硫氢键或异功能键。

[0207] 标记粘合剂是用多种目前可获得的方法的任何一种固定到固体基质上的。固体基质通常是通过将基质的所有或一部分暴露于化学试剂中而衍生或功能化的,这样即在其表面固定了一个可与标记粘合剂的一部分进行反应的化学基团。例如,适合于附着较长链部分的基团可包括胺、羟基、巯基和羧基基团。氨烷基硅烷和羟烷基硅烷可用于功能化多种表面,如玻璃表面。这种固相生物多聚体阵列的构建在文献中有详尽描述。参见如 Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963) (描述如多肽的固相成

分合成);Geysen 等人, J. Immun. Meth., 102:259-274(1987)(描述在针上的固相合成);Frank&Doring, Tetrahedron, 44:6031-6040(1988)(描述多种肽序列在纤维素盘上的合成);Fodor 等人, Science, 251:767-777(1991);Sheldon 等人, Clinical Chemistry, 39(4):718-719(1993);和 Kozal 等人, Nature Medicine, 2(7):753-759(1996)(均描述固定在固体基质上的生物多聚体阵列)。在基质上固定标记粘合剂的非化学方法包括其他的普通方法,如加热、用 UV 辐射的交联等。

[0208] 4. 基于计算机的测定

[0209] 对调节 T1R 多肽活性的化合物的另一种测定包括计算机辅助的化合物设计,其中将计算机系统用于生成 T1R 多肽的三维结构,该结构是基于由其氨基酸序列编码的结构信息的。输入的氨基酸序列与计算机程序中预先设定的算法直接和积极地相互作用以得到蛋白质的二级、三级和四级结构模型。然后再检查该蛋白质的结构模型以鉴定具有结合如配体能力的结构区域。然后将这些区域用于鉴定结合该蛋白质的配体。

[0210] 蛋白质的三维结构模型是通过将至少 10 个氨基酸残基的蛋白质氨基酸序列或相应的编码 T1R 多肽的核酸序列输入到计算机程序中而生成的。编码 T1R 多肽的核苷酸序列或其氨基酸序列可为任何此处所公开的序列和其保守修饰的形式。

[0211] 氨基酸序列代表了蛋白质的一级序列或子序列,该序列编码了蛋白质的结构信息。将氨基酸序列的至少 10 个残基(或编码 10 个氨基酸的核苷酸序列)从计算机键盘或计算机可读的介质输入到计算机系统中,其中的介质包括但不限于电子存储介质(如磁盘、磁带、盒式磁带和芯片)、光学介质(如 CD ROM)、由英特网站点及 RAM 分发的信息。然后该蛋白质的三维结构模型可利用本领域技术人员公知的软件,通过氨基酸序列和计算机系统的相互作用来生成。

[0212] 氨基酸序列代表了编码形成目标蛋白质的二级、三级和四级结构所需的信息的一级结构。软件查看由一级序列编码的某些参数以生成结构模型。这些参数被称为“能量条件(energy terms)”,且主要包括静电势、疏水势、溶剂可及表面和氢键。二级能量条件包括范德瓦耳斯势。生物学分子以累积的方式形成使能量条件最小化的结构。因此该计算机程序用这些由氨基酸序列的一级结构编码的条件来创建二级结构模型。

[0213] 然后由二级结构编码的蛋白质的三级结构在二级结构的能量条件的基础上形成。在这时用户可输入额外的变量,如该蛋白质是膜结合的还是可溶性的、它在身体内的定位及其细胞定位如胞质、表面或细胞核。这些变量与二级结构的能量条件一起用于形成三级结构的模型。在三级结构的建模中,计算机程序使二级结构的疏水面与相似的匹配,并使二级结构的亲水面与相似的匹配。

[0214] 一旦生成了该结构,即可通过计算机程序来鉴定潜在的配体结合区域。潜在配体的三维结构是如上所述通过将化合物的氨基酸或核苷酸序列或化学分子式输入而生成的。然后将潜在配体的三维结构与 T1R 多肽的结构进行比较以鉴定结合蛋白质的配体。蛋白质和配体之间的结合亲和力是用能量条件确定的,从而确定哪个配体具有增强的结合蛋白质的可能性。

[0215] 计算机系统也用于筛选 T1R 基因的突变、多态变体、等位基因和种间同系物。这种突变可与疾病状况或遗传性状相关。如上所述, GeneChip™ 和相关的技术也可用于筛选突变、多态变体、等位基因和种间同系物。一旦鉴定了变体,诊断测定即可用于鉴定具有这种

突变基因的病人。对突变的 T1R 基因的鉴定包含接受 T1R 基因或其保守修饰形式的第一个核酸或氨基酸序列的输入。如上所述将序列输入到计算机系统中。然后将第一个核酸或氨基酸序列与第二个（与第一个序列基本一致的）核酸或氨基酸序列进行比较。以上述的方式将第二个序列输入到计算机系统中。一旦对第一个和第二个序列进行了比较，即可鉴定序列间的核苷酸或氨基酸差异。这种序列代表多种 T1R 基因中的等位基因差异和与疾病状态和遗传性状有关的突变。

[0216] 5. 基于细胞的结合测定

[0217] 在一个实施方案中，T1R 蛋白质或多肽是作为与异源的伴随序列的嵌合受体在真核细胞中表达的，该伴随序列促进其成熟及在分泌途径中的定位。这种嵌合 T1R 多肽可表达于任何真核细胞如 HEK-293 细胞中。优选地，该细胞包含功能性 G 蛋白，如 G α 15，该蛋白质能够使嵌合受体与细胞内信号通路或信号蛋白质如磷脂酶 C 偶联。这种嵌合受体在这种细胞内的活化可用任何标准的方法探测，如通过探测细胞内钙的变化，其中钙的变化是通过探测细胞中 FURA-2 依赖性的荧光而实现的。

[0218] 活化的 GPCR 受体变为使受体的 C- 末端尾部（也可能其他位点）磷酸化的激酶的底物。因而，激活剂将促进 32P 从 γ - 标记的 GTP 转移到受体，这将用闪烁计数器进行测定。C- 末端尾部的磷酸化将促进视紫红质抑制蛋白样蛋白质的结合及干扰 G 蛋白的结合。激酶 / 视紫红质抑制蛋白途径在许多 GPCR 受体的脱敏中起关键的作用。例如，调节味觉受体保持活性的持续时间的化合物将用作延长想要的味觉或切断讨厌的味觉的手段。对于 GPCR 信号转导和测定信号转导方法的一般综述，参见如 *Methods in Enzymology*, 卷 237 和 238 (1994) 和卷 96 (1983) ; Bourne 等人, *Nature*, 10 : 349 : 117-27 (1991) ; Bourne 等人, *Nature*, 348 : 125-32 (1990) ; Pitcher 等人, *Annu. Rev. Biochem.*, 67 : 653-92 (1998)。

[0219] T1R 调节可通过将用推定的 T1R 调节子处理的 T1R 多肽的反应与未处理的对照样品的反应进行比较而测定。这种推定的 T1R 调节子可包括抑制或激活 T1R 多肽活性的味觉刺激物。在一个实施方案中，指定对照样品（未用激活剂或抑制剂进行处理的）具有 100 的相对 T1R 活性值。对 T1R 多肽的抑制是在当相对于对照的 T1R 活性值为约 90%，可选择地为 50%，可选择地为 25-0% 时而获得的。T1R 多肽的活化是在当相对于对照的 T1R 活性值为约 110%，可选择地为 150%、200-500% 或 1000-2000% 时而获得的。

[0220] 离子流的变化可通过确定表达 T1R 多肽的细胞或膜的离子极化（即电势）的变化而测定。一种确定细胞极化中变化的方法是通过用电压钳和膜片钳技术测量电流的变化（因而测量极化中的变化）（参见如“细胞附着的”模式、“翻转的（inside-out）”模式和“全细胞”模式，如 Ackerman 等人, *New Engl. J. Med.*, 336 : 1575-1595 (1997)）。全细胞电流可用标准方法方便地确定。其他已知的测定包括：放射性标记的离子流测定和用电压敏感性染料的荧光测定（参见如 Vestergaard-Bogind 等人, *J. Membrane Biol.*, 88 : 67-75 (1988) ; Gonzales & Tsien, *Chem. Biol.*, 4 : 269-277 (1997) ; Daniel 等人, *J. Pharmacol. Meth.*, 25 : 185-193 (1991) ; Holevinsky 等人, *J. Membrane Biol.*, 137 : 59-70 (1994)）。通常，要检验的化合物是以 1pM ~ 100mM 的范围存在的。

[0221] 检验的化合物对多肽功能的作用可通过检查任何上述的参数而测量。任何影响 GPCR 活性的适当生理学变化都可用于测定检验的化合物对本发明的多肽的影响。当用完整的细胞或动物来确定功能结果时，人们也可测量多种作用，如递质的释放、激素的释放、对

已知或未表征的遗传标记的转录变化（如 RNA 印迹）、细胞代谢如细胞生长的变化或 pH 变化，以及细胞内第二信使如 Ca^{2+} 、IP₃、cGMP 或 cAMP 的变化。

[0222] 优选的 GPCRs 测定物包括加载了离子或电压敏感性染料以报道受体活性的细胞。确定这种受体活性的测定也可用其他 G 蛋白偶联受体的已知激动剂和拮抗剂作为负或正对照以测定检验的化合物的活性。除了鉴定调节化合物（如激动剂、拮抗剂）的测定中，胞质中离子水平或膜电压的变化将分别用离子敏感性或膜电压荧光指示剂来监控。除了可采用的离子敏感性指示剂和电压探针外，其它可用的公开于 Molecular Probes 1997Catalog。对于 G 蛋白偶联的受体，混杂的 G 蛋白如 G α 15 和 G α 16 可用于选择的测定中（Wilkie 等人，Proc. Nat'l Acad. Sci, 88 :10049-10053(1991)）。这种混杂的 G 蛋白允许大量受体的偶联。

[0223] 受体的活化一般起始了随后的细胞内事件，如第二信使如 IP₃ 的增加，这释放了细胞内的钙离子贮藏。一些 G 蛋白偶联受体的活化通过磷脂酶 C 介导的磷脂酰肌醇的水解而刺激肌醇三磷酸（IP₃）的形成（Berridge&Irvine, Nature, 312 :315-21(1984)）。IP₃ 反过来刺激细胞内钙离子贮藏的释放。因而，胞质钙离子水平的变化或第二信使如 IP₃ 水平的变化可用于估计 G 蛋白偶联受体的功能。作为细胞内贮藏和通过离子通道的活化两者贡献的结果，表达这种 G 蛋白偶联受体的细胞可显示增加的胞质钙水平，在该情况下想要的是进行这种在无钙的缓冲液中辨别由内部贮藏的钙释放导致的荧光反应的测定，尽管这不是必需的，在其中的缓冲液中可选择地补充了整合剂如 EGTA。

[0224] 其他测定可包括确定受体的活性，该受体当活化时通过激活或抑制如腺苷酸环化酶的活性而导致细胞内环核苷酸如 cAMP 或 cGMP 水平的变化。存在环核苷酸门控离子通道，如通过结合 cAMP 或 cGMP 而活化的阳离子可透过的杆状光受体细胞通道和嗅觉神经元通道（参见如 Altenhofen 等人，Proc. Nat'l Acad. Sci, 88 :9868-9872(1991) 和 Dhallan 等人，Nature, 347 :184-187(1990)）。在受体的活化导致环核苷酸水平降低的情况下，可能优选的是在向本测定中的细胞内添加受体激活化合物之前，将细胞暴露于增加细胞内环核苷酸水平的试剂如毛喉素中。用于这类测定的细胞可通过宿主细胞的共转染而制备，其中该宿主细胞用编码环核苷酸门控离子通道、GPCR 磷酸酶的 DNA 和编码受体的 DNA（如某些谷氨酸受体、毒蝇碱性乙酰胆碱受体、多巴胺受体、5-羟色胺受体等）共转染，所述受体当活化时，导致胞质中环核苷酸水平的变化。

[0225] 在一个优选的实施方案中，T1R 多肽活性是通过在异源细胞中表达 T1R 基因而测量的，该细胞具有将受体连接到磷脂酶 C 信号转导通路中的混杂 G 蛋白（参见 Offermanns & Simon, J. Biol. Chem., 270 :15175-15180(1995)）。可选择地，该细胞系是 HEK-293（它并不天然地表达 T1R 基因），且混杂的 G 蛋白为 G α 15（上述的 Offermanns&Simon）。对味觉转导的调节是通过测量细胞内 Ca^{2+} 水平的变化而测定的，该 Ca^{2+} 水平响应于通过给予与 T1R 多肽有关的分子而对 T1R 信号转导通路的调节。 Ca^{2+} 水平的变化可选择地用荧光 Ca^{2+} 指示剂染料和荧光测定成像来测量。

[0226] 在一个实施方案中，细胞内 cAMP 或 cGMP 的变化可用免疫测定来测量。在 Offermanns&Simon, J. Biol. Chem., 270 :15175-15180(1995) 中描述的方法可用于确定 cAMP 的水平。同样地，在 Felley-Bosco 等人，Am. J. Resp. Cell and Mol. Biol., 11 :159-164(1994) 中描述的方法可用于确定 cGMP 的水平。进一步地，测量 cAMP 和 / 或 cGMP

的测定试剂盒描述于美国专利 4, 115, 538 中, 此处引用作为参考。

[0227] 在另一个实施方案中, 磷脂酰肌醇 (PI) 水解可根据美国专利 5, 436, 128 进行分析, 此处引用作为参考。简要地, 该测定包括用 ^3H - 肌醇对细胞进行 48 或更多小时的标记。用检验化合物对标记的细胞处理 1 小时。使处理的细胞裂解并在氯仿-甲醇-水中抽提, 其后用离子交换色谱分离肌醇磷酸并用闪烁计数进行定量。折叠刺激是通过计算在激动剂存在时的 cpm 与在缓冲液对照存在时的 cpm 的比例来确定的。同样地, 折叠抑制是通过计算在拮抗剂存在时的 cpm 与在缓冲液对照 (该对照中可含有或不含有激动剂) 存在时的 cpm 的比例来确定的。

[0228] 在另一个实施方案中, 可测量转录水平以估计检验化合物对信号转导的作用。使含有目标 T1R 多肽的宿主细胞与检验化合物接触足够的时间以引起任何相互作用, 然后测量基因表达的水平。引起这种相互作用的时间可根据经验确定, 如通过进行一个时程实验并测量作为时间的函数的转录水平。转录的量可通过用任何本领域技术人员公知的适当方法来测量。例如, 目标蛋白质的 mRNA 表达可用 RNA 印迹来探测, 或者其多肽产物可用免疫测定来鉴定。可选择地, 可以利用如在美国专利 5, 436, 128 中描述的用报道基因的基于转录的测定, 此处引用作为参考。该报道基因可为如氯霉素乙酰转移酶、萤光素酶、 3 - 半乳糖苷酶和碱性磷酸酶。此外, 可通过附着于第二报道基因如绿色荧光蛋白上而使目标蛋白质用作间接的报道基因 (参见如 Mistili & Spector, *Nature Biotechnology*, 15: 961-964 (1997))。

[0229] 然后将转录的量或者与在检验化合物不存在时的相同细胞中转录的量进行比较, 或者与缺乏 T1R 多肽的基本相同的细胞中转录的量进行比较。基本相同的细胞可源自相同的细胞, 从所述相同细胞制备了重组细胞但该细胞未通过引入异源 DNA 而修饰。转录量的任何差异均显示检验化合物以某些方式改变了目标 T1R 多肽的活性。

[0230] 6. 表达化学感受性受体的转基因非人动物

[0231] 表达一种或多种本发明的化学感受性受体序列的非人动物也可用于受体分析。这种表达可用于确定检验化合物在体外是否特异性地结合哺乳动物味觉跨膜受体多肽, 这是通过使用被核酸序列稳定或瞬时转染的非人动物与检验化合物接触而实现的, 其中的核酸序列编码化学感受性受体或其配体结合区域, 也可用于确定动物是否通过特异性地结合受体多肽而与检验化合物进行反应。

[0232] 用本发明的载体转染或感染的动物特别可用于对味觉刺激物 / 配体的鉴定和表征的测定中, 其中的味觉刺激物 / 配体可结合特定的或一套受体。这种表达人化学感受性受体序列的载体感染的动物可用于味觉刺激物的体内筛选和其对如细胞生理学 (如对味觉神经元)、对 CNS 或行为的作用中。

[0233] 感染 / 表达核酸和载体 (单个的或作为文库的) 的方法是本领域众所周知的。多种单独的细胞、器官或完整动物参数可通过多种方法进行测量。本发明的 T1R 序列可如通过用感染介质 (如腺病毒表达载体) 进行递送而表达于动物味觉组织中。

[0234] 内源的化学感受性受体基因可保持有功能的, 且野生型 (天然的) 活性仍然可存在。在其他的情形中, 当想要所有的化学感受性受体活性都由引入的外源杂交受体引起时, 则基因敲除品系是优选的。构建非人的转基因动物具体而言地为转基因小鼠, 以及筛选和制备重组构建体以生成转化的细胞的方法是本领域众所周知的。

[0235] “基因敲除”细胞和动物的构建是基于以下前提的,即特定基因在哺乳动物中的表达水平可通过向基因组引入新的 DNA 序列而降低或完全取消,该 DNA 序列起中断要抑制的基因的 DNA 序列的一些部分的作用。同样地,“基因陷阱插入 (gene trap insertion)”可用于破坏宿主基因,且小鼠胚胎干细胞 (ES) 可用于生产基因敲除的转基因动物 (参见如 Holzschu, Transgenic Res 6 :97-106(1997))。外源序列的插入一般是通过在互补核酸序列之间的同源重组实现的。外源序列是要修饰的目标基因的一部分,如外显子、内含子或转录调节序列,或任何能够影响目标基因表达水平的基因组序列;或其组合。在多能胚胎干细胞中通过同源重组的基因寻靶可允许人们精确地修饰目标基因组序列。任何技术都可用于创建、筛选、繁殖基因敲除动物,如参见 Bijvoet, Hum. Mol. Genet. 7 :53-62(1998); Moreadith, J. Mol. Med. 75 :208-216(1997); Tojo, Cytotechnology 19 :161-165(1995); Mudgett, Methods Mol. Biol. 48 :167-184(1995); Longo, Transgenic Res. 6 :321-328(1997); 美国专利 Nos. 5,616,491; 5,464,764; 5,631,153; 5,487,922; 5,627,059; 5,272,071; WO 91/09955; WO 93/09222; WO96/29411; WO 95/31560; WO 91/12650。

[0236] 本发明的核酸也可用作生产“基因敲除”的人细胞及其后裔的试剂。同样地,本发明的核酸也可用作生产小鼠中“基因引入 (knock-ins)”的试剂。人或大鼠 T1R 基因序列可置换小鼠基因组中的定向同源 T1R。在这个途径中,表达人或大鼠 T1R 的小鼠即可生产。然后该小鼠可用于分析人或大鼠 T1R 的功能,并鉴定这种 T1Rs 的配体。

[0237] D. 调节子

[0238] 作为 T1R 家族成员的调节子而检验的化合物可为任何小的化学化合物或生物学实体,如蛋白质、糖类、核酸或脂类。可选择地,调节子可为 T1R 基因的遗传改变形式。一般地,检验化合物将为小的化学分子和肽。基本上任何化合物都可以用作本发明测定中的潜在调节子或配体,尽管大多数时候使用可溶解于水或有机 (尤其是基于 DMSO 的) 溶液中的化合物。该测定通过使测定步骤自动化和从任何方便的来源为测定提供化合物而设计为能够筛选大的化学文库,这一般是平行进行的 (如以微量滴定形式在机器测定的微量滴定板上)。将理解的是有许多化学化合物的供应商,包括 Sigma (St. Louis, MO)、Aldrich (St. Louis, MO)、Sigma-Aldrich (St. Louis, MO)、Fluka Chemika-Biochemica Analytika (Buchs, 瑞士) 等。

[0239] 在一个优选的实施方案中,高生产量的筛选方法涉及提供含有大量潜在药物组合 (潜在的调节子或配体化合物) 的组合化学或肽文库。然后这种“组合化学文库”或“肽文库”在此处所描述的一种或多种测定中进行筛选,从而鉴定那些显示想要的特征活性的文库成员 (特别是化学种类或亚类)。这样鉴定的化合物可充当常规的“引导化合物 (lead compound)”或自身用作潜在的或实际的药物。

[0240] 组合化学文库是通过组合大量的化学“模块 (building block)”如试剂而得到的不同化学化合物的集合,该化合物是由化学合成或生物合成生成的。例如,线性组合化学文库如多肽文库是通过以各种可能的途径对于给定的混合物长度 (即多肽化合物中的氨基酸数目) 组合一套化学模块 (氨基酸) 而形成的。通过这种化学模块的组合混合可以合成数百万化学化合物。

[0241] 组合化学文库的制备和筛选是本领域技术人员众所周知的。这种组合化学文库包

括,但不局限于肽文库(参见如美国专利 5,010,175, Furka, Int. J. Pept. Prot. Res., 37: 487-493(1991) 和 Houghton 等人, Nature, 354:84-88(1991))。也可采用其他生成化学多样性文库的化学品。这种化学品包括,但不局限于:类肽(peptoids)(如 PCT Publication No. WO 91/19735)、编码的肽(如 PCT Publication No. WO 93/20242)、随机生物寡聚体(如 PCT Publication No. WO 92/00091)、苯并二氮杂(如美国专利 No. 5,288,514)、diversomers 如乙内酰脲、苯并二氮杂和二肽(Hobbs 等人, Proc. Nat. Acad. Sci., 90: 6909-6913(1993))、插烯多肽(Hagihara 等人, J. Amer. Chem. Soc., 114:6568(1992))、具有葡萄糖支架的非肽肽模拟物(peptidomimetics)(Hirschmann 等人, J. Amer. Chem. Soc., 114:9217-9218(1992))、小化合物文库的类似有机合成(Chen 等人, J. Amer. Chem. Soc., 116:2661(1994))、寡聚氨基甲酸酯(oligocarbamates)(Cho 等人, Science, 261:1303(1993))、肽酰磷酸酯(Campbell 等人, J. Org. Chem., 59:658(1994))、核酸文库(Ausubel, Berger 和 Sambrook, 均为上述的)、肽核酸文库(美国专利 5,539,083)、抗体文库(Vaughn 等人, Nature Biotechnology, 14(3):309-314(1996) 和 PCT/US96/10287)、糖类文库(Liang 等人, Science, 274:1520-1522(1996) 和美国专利 5,593,853)、小的有机分子文库(苯并二氮杂, Baum, C&EN, 1 月 18, 33 页(1993); thiazolidinones 和 metathiazanones, 美国专利 5,549,974; pynrolidines, 美国专利 5,525,735 和 5,519,134; 吗啉代化合物, 美国专利 5,506,337; 苯并二氮杂, 美国专利 5,288,514 等)

[0242] 制备组合文库的装置是商业上可获得的(参见如 357MPS、390MPS(Advanced Chem Tech, Louisville KY)、Symphony(Rainin, Woburn, MA)、433A(Applied Biosystems, Foster City, CA)、9050Plus(Millipore, Bedford, MA))。此外,许多组合文库自身是商业上可获得的(参见如 ComGenex, Princeton, NJ; Tripos, Inc., St. Louis, MO; 3D Pharmaceuticals, Exton, PA; Martek Biosciences; Columbia, MD 等)。

[0243] 在本发明的一个方面, T1R 调节子可用于任何食品产品、糖果、药物组合物或其成分以借此以想要的方式调节产品、组合物或成分的味道。例如,可添加增强甜味感觉的 T1R 调节子以使产品或组合物变甜,而可添加阻挡讨厌的味觉的 T1R 调节子以改善产品或组合物的味道。

[0244] E. 表示和预测味觉感受的方法

[0245] 本发明也优选地提供了在哺乳动物包括人中表示味觉感受和/或预测味觉感受的方法。优选地,这种方法可用此处所公开的受体和编码该 T1R 多肽的基因来进行。

[0246] 在本发明中也涉及的是一种筛选一种或多种能产生哺乳动物可探测的味觉的化合物的方法,该方法包含:使该一种或多种化合物与公开的受体接触,其中的哺乳动物优选地为人。在本发明中也涉及的是一种在哺乳动物中表示特殊味道的味觉感受的方法,该方法包含以下步骤:提供代表该脊椎动物的 n 个化学感受性受体每一个的定量刺激的值 X_1-X_n , 其中 n 大于或等于 2; 及从该值生成味觉感受的定量表示。化学感受性受体可为此处所公开的化学感受性受体,该表示可构成 n- 维空间中的一个点或一个组,可构成一个图表或一个谱,且可构成一个定量表示矩阵。同样地,提供的步骤可包含使大多数重组生产的化学感受性受体与检验组合物接触并定量测量该组合物与该受体的相互作用。

[0247] 在本发明中也涉及的是一种在哺乳动物中预测味觉感受的方法,该味觉感受是由在哺乳动物中产生未知味觉感受的一种或多种分子或分子组合生成的,该方法包含以下步

骤:对于在哺乳动物中产生已知味觉感受的一种或多种分子或分子组合,提供代表该脊椎动物的 n 个化学感受性受体每一个的定量刺激的值 X_1-X_n ,其中 n 大于或等于2;及从该值生成能在哺乳动物产生已知味觉感受的一种或多种分子或分子组合在哺乳动物中味觉感受的定量表示,对于在哺乳动物中产生未知味觉感受的一种或多种分子或分子组合,提供代表该脊椎动物的 n 个化学感受性受体每一个的定量刺激的值 X_1-X_n ,其中 n 大于或等于2;及从该值生成能在哺乳动物产生未知味觉感受的一种或多种分子或分子组合在哺乳动物中味觉感受的定量表示,以及通过将能在哺乳动物产生未知味觉感受的一种或多种分子或分子组合的味觉感受的定量表示与能在哺乳动物产生已知味觉感受的一种或多种分子或分子组合的味觉感受的定量表示进行比较,从而预测哺乳动物中由在哺乳动物中产生未知味觉感受的一种或多种分子或分子组合生成的味觉感受。在该方法中所用的化学感受性受体可包括此处所公开的化学感受性受体。

[0248] 在另一个实施方案中,生成了在哺乳动物中引起预定味觉感受的新分子或分子组合,这通过以下步骤实现:即在哺乳动物中确定上述的已知分子或分子组合的味觉感受值;在哺乳动物中确定上述的未知分子或分子组合的味觉感受值;将哺乳动物中对一种或多种未知组合物的味觉感受值与哺乳动物中对一种或多种已知组合物的味觉感受值进行比较;选择在哺乳动物中引起预定味觉感受的分子或分子组合;组合两种或多种未知分子或分子组合以形成在哺乳动物中引起预定味觉感受的分子或分子组合。该组合步骤产生了在哺乳动物中引起预定味觉感受的单一分子或分子组合。

[0249] 在本发明的另一个实施方案中,提供了一种模拟味觉的方法,该方法包含以下步骤:对于大多数克隆的化学感受性受体(优选地为人受体)中的每一个,确定受体与味觉刺激物相互作用的程度;以一定量组合大量的化合物使得它们一起提供模拟味觉刺激物特点的受体刺激特点,其中所述化合物中的每一个都具有预先确定的与一种或多种受体相互作用的性质。味觉刺激物与化学感受性受体的相互作用可用此处所描述的任何结合或报道测定进行确定。然后组合大量化合物以形成混合物。如果需要,大量化合物的一种或多种可共价结合。组合的化合物基本刺激至少75%、80%或90%的受体,其中该受体基本是由味觉刺激物刺激的。

[0250] 在本发明另一个优选实施方案中,将大量标准化合物对大量化学感受性受体进行了检验以确定每一受体与每一标准化合物相互作用的程度,并从而生成对每一标准化合物的受体刺激特点。然后这些受体刺激特点可在数据储存介质上贮藏于相关的数据库中。该方法可进一步包含提供对于味觉想要的受体刺激特点;将想要的受体刺激特点与相关数据库进行比较;以及确定一种或多种与想要的受体刺激特点最匹配的标准化合物的组合。该方法可进一步包含在一种或多种确定的组合中组合标准化合物以刺激味觉。

[0251] F. 试剂盒

[0252] T1R 基因及其同系物对鉴定化学感受性受体细胞、对法医和父子关系确定及对检查味觉转导都是有用的工具。特异性地与 T1R 核酸杂交的 T1R 家族成员特异性试剂如 T1R 探针和引物,以及特异性地结合 T1R 多肽的 T1R 特异性试剂如 T1R 抗体用于检查味觉细胞表达和味觉转导调节。

[0253] 包括多种技术的在样品中确定 T1R 家族成员的 DNA 和 RNA 是否存在的核酸测定是本领域技术人员所公知的,如 DNA 印迹分子、RNA 印迹分析、斑点印迹、RNase 保护、S1 分析、

如 PCR 的扩增技术和原位杂交。例如, 在原位杂交中, 目标核酸是这样从其细胞环境中释放的, 从而它可用于细胞内的杂交, 而同时保持了细胞的形态学以进行随后的描述和分析。下面的文章提供了对原位杂交技术的综述: Singer 等人, *Biotechniques*, 4:230-250 (1986); Haase 等人, *Methods in Virology*, vol. VII, pp. 189-226 (1984); 和 *Nucleic Acid Hybridization: A Practical Approach* (Names 等人, eds. 1987)。此外, T1R 多肽可用上述多种免疫测定技术进行探测。检验样品一般是与正对照 (如表达重组 T1R 多肽的样品) 和负对照进行比较的。

[0254] 本发明也提供了筛选 T1R 家族成员的调节子的试剂盒。这种试剂盒可从易于获得的原料和试剂制备。例如, 这种试剂盒可包含下述材料的任何一种或多种: T1R 核酸或蛋白质、反应试管和检验 T1R 活性的说明书。可选择地, 该试剂盒含有生物学活性的 T1R 受体。依赖于试剂盒的预定用户和用户的特殊需要, 根据本发明可制备多种试剂盒和成分。

实施例

[0255] 在此处所给出的蛋白质序列中, 单字母密码 X 或 Xaa 指 20 个常用氨基酸的任何一个。在此处所给出的 DNA 序列中, 单字母密码 N 或 n 指 4 个常用核苷酸碱基 A、T、C 或 G 的任何一个。

[0256] 实施例 1-hT1R3

[0257] 在下面提供 hT1R3 基因组 DNA, 以 SEQ ID NO 1 和 SEQ ID NO 2 表示, 其预测的编码序列 (cds) 以黑体字显示。在 5' 和 3' 毗连群 (contig) 之间的间断以 elipses (‘.....’) 显示。由 hT1R3 预测的 cds 描述于 SEQ ID NO 3 中。最后, 用氨基酸的单字母密码提供优选的预测 hT1R3 氨基酸序列为 SEQ ID NO 4。

[0258] hT1R3 基因组 DNA-5' 毗连群 (SEQ ID NO 1)

[0259] AGCCTGGCAGTGGCCTCAGGCAGAGTCTGACGCGCACAACTTTTCAGGCCAGGAAGCGAGGACACCACTGGGGCCCCAGGGTGTGGCAAGTGAGGATGGCAAGGGTTTTGCTAAACAAATCCTCTGCCCGCTCCCCGCCCGGGCTCACTCCATGTGAGGCCAGTCGGGGCAGCCACCTGCCGTGCCTGTTGGAAGTTGCC
T C T G C C **ATGCTGGGCCCTGCTGTCCTGGGCCTCAGCCTCTGGGCTCTCCTG
CACCTGGGACGGGGGCCCCATTGTGCCTGTCACAGCAACTTAGGATGAAG
GGGGACTACGTGCTGGGGGGGCTGTTCCCCCTGGGCGAGGCCGAGGAGGC
TGGCCTCCGCAGCCGGACACGGCCCAGCAGCCCTGTGTGCACCAG** GTACAGAGGTG
GGACGGCCTGGGTCGGGGTCAGGGTGACCAGGTCTGGGGTGCTCCTGAGCTGGGGCCGAGGTGGCCATCTGCGGTTC
TGTGTGGCCCCAG **GTTCT**
**CCTCAAACGGCCTGCTCTGGGCACTGGCCATGAAAATGGCCGTGGAGGAGA
TCAACAACAAGTCGGATCTGCTGCCCGGGCTGCGCCTGGGCTACGACCTCT
TTGATACGTGCTCGGAGCCTGTGGTGGCCATGAAGCCCAGCCTCATGTTCT
GGCCAAGGCAGGCAGCCGCGACATCGCCGCCTACTGCAACTACACGCAGTA
CCAGCCCCGTGTGCTGGCTGTCATCGGGCCCCACTCGTCAGAGCTCGCCAT
GGTCACCGGCAAGTTCTTCAGCTTCTTCTCATGCCCA** GTGGGGCGCCCCCACCATCAC
CCACCCCAACCAACCCCTGCCCCGTGGGAGCCCCCTTGTGTCAGGAGAATGC (SEQ ID NO :1)

[0260] hT1R3 基因组 DNA-3' 毗连群 (SEQ ID NO 2)

[0261] TACATGCACCCCACCCAGCCCTGCCCTGGGAGCCCTGTGTC
AGAAGATGCTCTTGGCCTTGCA **GGTCAGCTACGGTGCTAGCATGGAGCTGCTG**
AGCGCCCGGGAGACCTTCCCCTCCTTCTTCCGCACCGTGCCCAGCGACCGT
GTGCAGCTGACGGCCGCCGCGGAGCTGCTGCAGGAGTTCGGCTGGAAGT
GGTGGCCGCCCTGGGCAGCGACGACGAGTACGGCCGGCAGGGCCTGAGC
ATCTTCTCGGCCCTGGCCGCGGCACGCGGCATCTGCATCGCGCACGAGGG
CCTGGTGCCGCTGCCCCGTGCCGATGACTCGCGGCTGGGGAAGGTGCAGG
ACGTCCTGCACCAGGTGAACCAGAGCAGCGTGCAAGGTGGTGCTGCTGTTCC
CCTCCGTGCACGCCGCCACGCCCTCTTCAACTACAGCATCAGCAGCAGGC
TCTGCCCAAGGTGTGGGTGGCCAGCGAGGCCTGGCTGACCTCTGACCTGG
TCATGGGGCTGCCCGGCATGGCCAGATGGGCACGGTGCTTGGCTTCTCC
AGAGGGGTGCCAGCTGCACGAGTTCCCCCAGTACGTGAAGACGCACCTG
GCCCTGGCCACCGACCCGGCCTTCTGCTCTGCCCTGGGCGAGAGGGAGCA
GGGTCTGGAGGAGGACGTGGTGGGCCAGCGCTGCCCGCAGTGTGACTGCA
TCACGCTGCAGAACGTGAGCGCAGGGCTAAATCACCACCAGACGTTCTCTG
TCTACGCAGCTGTGTATAGCGTGGCCCAGGCCCTGCACAACACTCTTCAGTG
CAACGCCTCAGGCTGCCCCGCGCAGGACCCCGTGAAGCCCTGGCAG
GTGAGCCCCGGGAGATGGGGGTGTGCTGTCCTCTGCATGTGCCAGGCCACCAGGCA
CGGCCACCACGCCTGAGCTGGAGGTGGCTGGCGGCTCAGCCCCGTCCCCCGCCCGCA
G **CTCCTGGAGAACATGTACAACCTGACCTTCCACGTGGGCGG**
GCTGCCGCTGCGGTTCGACAGCAGCGGAAACGTGGACATGGAGTACGACCT
GAAGCTGTGGGTGTGGCAGGGCTCAGTGCCCAGGCTCCACGACGTGGGCA
GGTTCAACGGCAGCCTCAGGACAGAGCGCCTGAAGATCCGCTGGCACACGT
CTGACAACCAG GTGAGGTGAGGGTGGGTGTGCCAGGCGTGCCCGTGG
TAGCCCCCGCGGCAGGGCGCAGCCTGGGGGTGGGGGCCGTTCCAGTCTCCCGT
GGGCATGCCAGCCGAGCAGAGCCAGACCCAGGCCTGTGCGCAG **AAGCC**
CGTGTCCCGGTGCTCGCGGCAGTGCCAGGAGGGCCAGGTGCGCCGGGTCA
AGGGGTTCCACTCCTGCTGCTACGACTGTGTGGACTGCGAGGCGGGCAGCT
ACCGGCAAACCCAG GTGAGCCGCCTTCCCGGCAGGCGGGGGTGGGAA
CGCAGCAGGGGAGGGTCCTGCCAAGTCCTGACTCTGAGACCAGAGCCCACAGGGT
ACAAGACGAACACCCAGCGCCCTTCTCCTCTCTCACAG **ACGACATCGCCT**

41

CCTCAAACGGCCTGCTCTGGGCACTGGCCATGAAAATGGCCGTGGAGGAGA
TCAACAACAAGTCGGATCTGCTGCCCCGGGCTGCGCCTGGGCTACGACCTCT
TTGATACGTGCTCGGAGCCTGTGGTGGCCATGAAGCCCAGCCTCATGTTCTT
GGCCAAGGCAGGCAGCCGCGACATCGCCGCCTACTGCAACTACACGCAGTA
CCAGCCCCGTGTGCTGGCTGTTCATCGGGCCCCACTCGTCAGAGCTCGCCAT
GGTCACCGGCAAGTTCTTCAGCTTCTTCCTCATGCCCA GTGGGGCGCCCCCACCATCACC
CACCCCCAACCAACCCCTGCCCCGTGGGAGCCCTTGTGTGAGGAGAATGCTACATGCACCCACCCAGCCCTGCCCT
GGGAGCCCTGTGTGAGAAGATGCTCTTGGCCTTGCA GGTCAGCTACGGTGCTAGCATGGAGCTG
CTGAGCGCCCGGGAGACCTTCCCCTCCTTCTTCCGCACCGTGCCCAGCGAC
CGTGTGCAGCTGACGGCCGCGCGGAGCTGCTGCAGGAGTTCCGGCTGGAA
CTGGGTGGCCGCCCTGGGCAGCGACGACGAGTACGGCCGGCAGGGCCTGA
GCATCTTCTCGGCCCTGGCCGCGGCACGCGGCATCTGCATCGCGCACGAG
GGCCTGGTGCCGCTGCCCCGTGCCGATGACTCGCGGCTGGGGAAGGTGCA
GGACGTCCTGCACCAGGTGAACCAGAGCAGCGTGACAGGTGGTGCTGCTGTT
CGCCTCCGTGCACGCCGCCACGCCCTCTTCAACTACAGCATCAGCAGCAG
GCTCTCGCCCAAGGTGTGGGTGGCCAGCGAGGCCTGGCTGACCTCTGACCT
GGTCATGGGGCTGCCCGGCATGGCCAGATGGGCACGGTGCTTGGCTTCT
CCAGAGGGGTGCCAGCTGCACGAGTTCCCCCAGTACGTGAAGACGCACCT
GGCCCTGGCCACCGACCCGGCCTTCTGCTGTGCCCTGGGCGAGAGGGAGC
AGGGTCTGGAGGAGGACGTGGTGGGCCAGCGCTGCCCGCAGTGTGACTGC
ATCACGCTGCAGAACGTGAGCGCAGGGCTAAATCACCACCAGACGTTCTCT
GTCTACGCAGCTGTGTATAGCGTGGCCAGGCCCTGCACAACACTCTTCAGT
GCAACGCCTCAGGCTGCCCCGCGCAGGACCCCGTGAAGCCCTGGCAG
GTGAGCCCCGGGAGATGGGGGTGTGCTGTCTCTGCATGTGCCAGGCCACCAGGCA
CGGCCACCACGCCTGAGCTGGAGGTGGCTGGCGGCTCAGCCCCGTCCCCCGCCCGCA
G CTCCTGGAGAACATGTACAACCTGACCTTCCACGTGGGCG
GGCTGCCGCTGCGGTTGACAGCAGCGGAAACGTGGACATGGAGTACGAC
CTGAAGCTGTGGGTGTGGCAGGGCTCAGTGCCCAGGCTCCACGACGTGGG
CAGGTTCAACGGCAGCCTCAGGACAGAGCGCCTGAAGATCCGCTGGCACAC
GTCTGACAACCAG GTGAGGTGAGGGTGGGTGTGCCAGGCGTGCCCG
TGGTAGCCCCCGCGGCAGGGCGCAGCCTGGGGGTGGGGGCGGTTCCAGTCTCC
CGTGGGCATGCCAGCCGAGCAGAGCCAGACCCAGGCCTGTGCGCAG AAG
CCCGTGTCCCGGTGCTCGCGGCAGTGCCAGGAGGGCCAGGTGCGCCGGGT
CAAGGGGTTCCACTCCTGCTGCTACGACTGTGTGGACTGCGAGGCGGGCAG
CTACCGGCAAAACCCAG GTGAGCCGCCTTCCCGCAGGCGGGGTGGGAACGCAGCAGGGGAGGGTC

CTGCCAAGTCCTGACTCTGAGACCAGAGCCACAGGGTACAAGACGAACACCCAGCGCCCTTCTCCTCTCTCACA
G **ACGACATCG**
CCTGCACCTTTTGTGGCCAGGATGAGTGGTCCCCGGAGCGAAGCACACGCT
GCTTCCGCCGCAGGTCTCGGTTCTTGGCATGGGGCGAGCCGGCTGTGCTGC
TGCTGCTCCTGCTGCTGAGCCTGGCGCTGGGCCTTGTGCTGGCTGCTTTGG
GGCTGTTTCGTTACCATCGGGACAGCCCACTGGTTCAGGCCTCGGGGGGGC
CCCTGGCCTGCTTTGGCCTGGTGTGCCTGGGCCTGGTCTGCCTCAGCGTCC
TCCTGTTCCCTGGCCAGCCAGCCCTGCCCGATGCCTGGCCCAGCAGCCCT
TGTCCACCTCCCGCTCACGGGCTGCCTGAGCACACTCTTCTGTCAGGCGG
CCGAGATCTTCGTGGAGTCAGAACTGCCTCTGAGCTGGGCAGACCGGCTGA
GTGGCTGCCTGCGGGGGCCCTGGGCCTGGCTGGTGGTGGCTGCTGGCCATG
CTGGTGGAGGTCGCACTGTGCACCTGGTACCTGGTGGCCTTCCCGCCGGAG
GTGGTGACGGACTGGCACATGCTGCCCACGGAGGCGCTGGTGCCTGCCG
CACACGCTCCTGGGTGAGCTTCGGCCTAGCGCACGCCACCAATGCCACGCT
GGCCTTTCTCTGCTTCCTGGGCACTTTCTGGTGGGAGCCAGCCGGGCTG
CTACAACCGTGCCCGTGGCCTCACCTTTGCCATGCTGGCCTACTTCATCACC
TGGGTCTCCTTTGTGCCCCCTCCTGGCCAATGTGCAGGTGGTCTCAGGCCC
GCCGTGCAGATGGGCGCCCTCCTGCTCTGTGTCCTGGGCATCCTGGCTGCC
TTCCACCTGCCCAGGTGTTACCTGCTCATGCGGCAGCCAGGGCTCAACACC
CCCGAGTTCTTCTGGGAGGGGGCCCTGGGGATGCCCAAGGCCAGAATGA
CGGGAACACAGGAAATCAGGGGAAACATGAGTGA CCCAACCCTGTGATCTCAGCCCCGGTGA
ACCCAGACTTAGCTGCGATCCCCCAAGCCAGCAATGACCCGTGTCTCGCTACAGAGACCCTCCCGCTCTAGGTTC
TGACCCAGGTTGTCTCCTGACCCTGACCCACAGTGAGCCCTAGGCCTGGAGCACGTGGACACCCCTGTGACCATC
(SEQ ID NO 20)

[0264] hT1R3 预测的 cds (SEQ ID NO 3)

[0265] ATGCTGGGCCCTGCTGTCTTGGGCCTCAGCCTCTGGGCTCTCCTGCACCCTGGGACGGGGGCCCCATTG
TGCCTGTCACAGCAACTTAGGATGAAGGGGACTACGTGCTGGGGGGGCTGTTCCCCCTGGGCGAGGCCGAGGAGGC
TGGCCTCCGCAGCCGGACACGGCCCAGCAGCCCTGTGTGCACCAGGTTCTCCTCAAACGGCCTGCTCTGGGCACTGG
CCATGAAAATGGCCGTGGAGGAGATCAACAACAAGTCGGATCTGCTGCCCGGGCTGCGCCTGGGCTACGACCTCTTT
GATACGTGCTCGGAGCCTGTGGTGGCCATGAAGCCCAGCCTCATGTTCTGGCCAAGGCAGGCAGCCGCGACATCGC
CGCCTACTGCAACTACACGCAGTACCAGCCCCGTGTGCTGGCTGTGTCATCGGGCCCCACTCGTCAGAGCTCGCCATGG
TCACCGGCAAGTTCTTCAGCTTCTTCTCATGCCCCAGGTCAGCTACGGTGCTAGCATGGAGCTGCTGAGCGCCCGG
GAGACCTTCCCCTCCTTCTTCCGCACCGTGCCAGCGACCGTGTGCAGCTGACGGCCCGCGGAGCTGCTGCAGGA
GTTTCGGCTGGAATGGGTGGCCGCCCTGGGCAGCGACGACGAGTACGGCCGGCAGGGCCTGAGCATCTTCTCGGCCC
TGGCCGCGGCACGCGGCATCTGCATCGCGCACGAGGGCCTGGTGGCCGTGCCCCGTGCCGATGACTCGCGGCTGGGG
AAGGTGCAGGACGTCTGCACCAGGTGAACCAGAGCAGCGTGCAGGTGGTGGCTGCTGTTTCGCCTCCGTGCACGCCGC

CCACGCCCCCTCTTCAACTACAGCATCAGCAGCAGGCTCTCGCCCAAGGTGTGGGTGGCCAGCGAGGCCTGGCTGACCT
CTGACCTGGTTCATGGGGCTGCCCGCATGGCCCAGATGGGCACGGTGTCTTGGCTTCCTCCAGAGGGGTGCCAGCTG
CACGAGTTCCCCCAGTACGTGAAGACGCACCTGGCCCTGGCCACCGACCCGGCCTTCTGCTCTGCCCTGGGCGAGAG
GGAGCAGGGTCTGGAGGAGGACGTGGTGGGCCAGCGCTGCCCGCAGTGTGACTGCATCACGCTGCAGAACGTGAGCG
CAGGGCTAAATCACCACCAGACGTTCTCTGTCTACGCAGCTGTGTATAGCGTGGCCCAGGCCCTGCACAACACTCTT
CAGTGC AACGCCCTCAGGCTGCCCGCGCAGGACCCCGTGAAGCCCTGGCAGCTCCTGGAGAACATGTACAACCTGAC
CTTCCACGTGGGCGGGCTGCCGCTGCGGTTTCGACAGCAGCGAAACGTGGACATGGAGTACGACCTGAAGCTGTGGG
TGTGGCAGGGCTCAGTGCCAGGCTCCACGACGTGGGCAGGTTCAACGGCAGCCTCAGGACAGAGCGCCTGAAGATC
CGCTGGCACACGTCTGACAACCAGAAGCCCGTGTCCCGGTGCTCGCGGCAGTGCCAGGAGGGGCCAGGTGCGCCGGGT
CAAGGGGTTCCTACTCCTGCTGCTACGACTGTGTGGACTGCGAGGCGGGCAGCTACCGGCAAACCCAGACGACATCG
CCTGCACCTTTTGTGGCCAGGATGAGTGGTCCCCGGAGCGAAGCACACGCTGCTTCCGCCGAGGTCTCGGTTCTCTG
GCATGGGGCGAGCCGGCTGTGCTGCTGCTGCTCCTGCTGCTGAGCCTGGCGCTGGGCCTTGTGCTGGCTGCTTTGGG
GCTGTTTCGTTACCATCGGGACAGCCCACTGGTTTCAGGCCTCGGGGGGGCCCTGGCCTGCTTTGGCCTGGTGTGCC
TGGGCCTGGTCTGCCTCAGCGTCTCCTCTGTTCCCTGGCCAGCCCAGCCCTGCCCGATGCCCTGGCCCAGCAGCCCTTG
TCCCACCTCCCGCTCACGGGCTGCCGTGAGCACACTCTTCTGTCAGGCGGGCCGAGATCTTCGTGGAGTCAGAACTGCC
TCTGAGCTGGGCAGACCGGCTGAGTGGCTGCCTGCGGGGGCCCTGGGCCTGGCTGGTGGTGTGCTGCTGGCCATGCTGG
TGGAGGTCGCACTGTGCACCTGGTACCTGGTGGCCTTCCCGCCGGAGGTGGTGACGGA CTGGCACATGCTGCCACG
GAGGCGCTGGTGCCTGCGGCACACGCTCCTGGGTGAGCTTCGGCCTAGCGCACGCCACCAATGCCACGCTGGCCTT
TCTCTGCTTCCTGGGCACCTTCTCTGGTGGCGAGCCAGCCGGGCTGCTACAACCGTGCCCGTGGCCTCACCTTTGCCA
TGCTGGCCTACTTCATCACCTGGGTCTCCTTTGTGCCCCCTCCTGGCCAATGTGCAGGTGGTCTCAGGCCCGCCGTG
CAGATGGGCGCCCTCCTGCTCTGTGTCTGGGCATCCTGGCTGCCTTCCACCTGCCAGGTGTTACCTGCTCATGCG
GCAGCCAGGGCTCAACACCCCCGAGTTCTTCTGGGAGGGGGCCCTGGGGATGCCCAAGGCCAGAATGACGGGAACA
CAGGAAATCAGGGGAAACATGAGTGA (SEQ ID NO 3)

[0266] hT1R3 推定の氨基酸序列 (SEQ ID NO 4)

[0267] MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGFLPLGEAEEAGLSRTRPSSPVCTRFSSNG
LLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKAGSRDIAAYCNYTQYQPRVLAVIGPHS
SELAMVTGKFFSFFLMPQVSYGASMELLSARETFPSFFRTVPSDRVQLTAAAEELLQEFGWNWVAALGSDDEYGRQGL
SIFSALAAARGICIAHEGLVPLPRADDSRLGKVQDVLHQVNQSSVQVLLFASVHAHAHALFNYSISSRLSPKVWVAS
EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQGLEEDVVGQRCPQCDCIT
LQNSAGLNHHQTFSVYAAVYSVAQALH NTLQCNASGCPAQDPVKPWQLLENMYNLTfHVGGPLPLRFDSSGNVDME
YDLKLWVWQGSVPRLHDVGRFNGSLRTERLKIrwHTSDNqKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYR
QNPDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPVLLLLLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLA
CFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESELPLSWADRLSGCLRGPWAWLV
VLLAMLVEVALCTWYLVAFPPEVVDWHMLPTEALVHCRTRSWVSFGLAHATNATLAFCLFLGTFLVRSQPGCYNRA
RGLTFAMLAY FITWVSFVPLLANVQVVLRAVQMGALLLCVLGILAAFHLPrcYLLMRQPGLNTPEFFLGGGPGDA
QGQNDGNTGNQGKHE (SEQ IDNO 4)

[0268] 实施例 2-rT1R3 和 mT1R 3

[0269] 通过 PCR 扩增用基于人 T1R3 序列的简并引物从基因组 DNA 中分离了大鼠和小鼠 T1R3 基因的区段。简并引物 SAP077(5'-CGNTTYYTNGCNTGGGGNGARCC-3';SEQ ID NO

5) 和 SAP079 (5' -CGNGCNCGRTRTARCANCCNGG-3'; SEQ ID NO 6) 是分别与人的 T1R3 残基 RFLAWGEPA (相应于 SEQ ID NO 7) 和 PGCYNRAR (相应于 SEQ ID NO 8) 互补的。将 PCR 产物克隆并测序。质粒 SAV115 携带有小鼠 T1R3 基因的克隆片段, SAV118 携带有大鼠基因的一个片段。下面显示的这些序列明显地代表人 T1R3 的在啮齿类中的对应物, 因为小鼠的片段与人 T1R3 的相应片段有 74% 的一致性, 且大鼠的片段与人 T1R3 的相应片段有 80% 的一致性。小鼠与大鼠的片段有 88% 的一致性。没有其他的数据库序列与这些 T1R3 片段有大于 40% 的一致性。

[0270] 有义方向的 SAV115 小鼠 T1R3 片段 (序列对应于移去的简并引物) (SEQ ID NO 9)

[0271] GTGCTGTCACTCCTCCTGCTGCTTTGCCTGGTGCTGGGTCTAGCACTGGCTGCTCTGGGGCTCTCTGCTCACCACCTGGGACAGCCCTCTTGTCCAGGCCTCAGGCGGCTCACAGTTCTGCTTTGGCCTGATCTGCCTAGGCCTCTTCTGCCTCAGTGTCTTCTGTTCCAGGACGGCCAAGCTCTGCCAGCTGCCCTTGACACAACAACCAATGGCTCACCTCCCTCTCACAGGCTGCCTGAGCACACTCTTCCTGCAAGCAGCTGAGACCTTTGTGGAGTCTGAGCTGCCACTGAGCTGGGCAAACTGGCTATGCAGCTACCTTCGGGACTCTGGCCTGCTAGTGGTACTGTTGGCCACTTTTGTGGAGGCAGCACTATGTGCCTGGTATTTGACCGCTTCACCAGAAGTGGTGACAGACTGGTCAGTGCTGCCCACAGAGGTACTGGAGCACTGCCACGTGCGTTCTGGGTCAACCTGGGCTTGGTGACATCACCAATGCAATGGTAGCT TTTCTCTGCTTTCTGGGCACTTTCTGCTGTTACAAGACCAG (SEQ ID NO 9)

[0272] mT1R3 片段, 推定的氨基酸序列 (SEQ ID NO 10)

[0273] VLSLLLLLCLVLGLALALGLSVHHWDSPLVQASGGSQFCFLICLGLFCLSVLLFPGRPSSASCLAQQPMAHLPLTGCLSTLFLQAAETFVESELPLSWANWLCSYLRDSSLVLLATFVEAALCAWYLTASPEVVTDWSVLPT EVLEHCHVRSWVNLGLVHITNAMVAFLCFLGTFLVQDQ (SEQ ID NO 10)

[0274] 有义方向的 SAV118 大鼠 T1R3 片段 (序列对应于移去的简并引物) (SEQ ID NO 11)

[0275] GTGCTGTCACTTCTCCTGCTGCTTTGCCTGGTGCTGGGCCTGACACTGGCTGCCCTGGGGCTCTTTGCTCACTACTGGGACAGCCCTCTTGTTCAGGCCTCAGGTGGGTCACTGTTCTGCTTTGGCCTGATCTGCCTAGGCCTCTTCTGCCTCAGTGTCTTCTGTTCCAGGACGACCACGCTCTGCCAGCTGCCCTTGCCCAACAACCAATGGCTCACCTCCCTCTCACAGGCTGCCTGAGCACACTCTTCCTGCAAGCAGCCGAGATCTTTGTGGAGTCTGAGCTGCCACTGAGTTGGGCAAACTGGCTCTGCAGCTACCTTCGGGGCCCCCTGGGCTTGGCTGGTGGTACTGCTGGCCACTCTTGTGGAGGCTGCACTATGTGCCTGGTACTTGATGGCTTTCCCTCCAGAGGTGGTGACAGATTGGCAGGTGCTGCCCACGGAGGTACTGGAACACTGCCGCATGCGTTCTGGGTGAGCCTGGGCTTGGTGACATCACCAATGCAGGGGTAGCTTTCTCTGCTTTCTGGGCACTTTCTGCTGTTACAAGCCAG (SEQ ID NO 11)

[0276] rT1R3 片段, 推定的氨基酸序列 (SEQ ID NO 12)

[0277] VLSLLLLLCLVLGLTLAALGLFVHYWDSPLVQASGGSFLFCFLICLGLFCLSVLLFPGRPRSASCLAQQPMAHLPLTGCLSTLFLQAAEIFVESELPLSWANWLCSYLRGPWAWLVLLATLVEAALCAWYLMAFPPEVVTDWQVL PTEVLEHCRMRSWVSLGLVHITNAGVAFLCFLGTFLVQSQ (SEQ ID NO 12)

[0278] 实施例 3-rT1R3 的克隆

[0279] 将上面鉴定为 SEQ ID NOs 9 和 11 的 mT1R3 和 rT1R3 区段用于筛选源自大鼠味觉组织的 cDNA 文库。对一个阳性克隆进行测序并发现该克隆含有在下面以 SEQ ID NO 13 给出的全长 rT1R3 序列。与 mT1R3 和 rT1R3 部分序列和与全长 hT1R3 序列的序列比较确定该 cDNA 代表 hT1R3 的大鼠对应物。例如, rT1R3 和 hT1R3 之间的成对氨基酸一致性为约 72%, 而公共 DNA 序列数据库中最相关的注释的序列与 rT1R3 仅有约 33% 的一致性。

[0280] rT1R3 预测的 cds (SEQ. ID NO. 13)

[0281] ATGCCGGGTTTGGCTATCTTGGGCCCTCAGTCTGGCTGCTTTCCTGGAGCTTGGGATGGGGTCCTCTTTG
TGTCTGTCACAGCAATTCAAGGCACAAGGGGACTATATATTGGGTGGACTATTTCCCTGGGCACAAGTGGGAGGC
CACTCTCAACCAGAGAACACAGCCCCAACGGCATCCTATGTACCAGGTTCTCGCCCCCTTGGTTTGTTCCTGGCCATGG
CTATGAAGATGGCTGTAGAGGAGATCAACAATGGATCTGCCTTGCTCCCTGGGCTGCGACTGGGCTATGACCTGTTT
GACACATGCTCAGAGCCAGTGGTCACCATGAAGCCCAGCCTCATGTTTCATGGCCAAGGTGGGAAGTCAAAGCATTGC
TGCCTACTGCAACTACACACAGTACCAACCCCGTGTGCTGGCTGTCATTGGTCCCCACTCATCAGAGCTTGCCCTCA
TTACAGGCAAGTTCTTCAGCTTCTTCCTCATGCCACAGGTCAGCTATAGTGCCAGCATGGATCGGCTAAGTGACCGG
GAAACATTTCCATCCTTCTTCCGCACAGTGGCCAGTGACCGGGTGCAGCTGCAGGCCGTTGTGACACTGTTGCAGAA
TTTCAGCTGGAAGTGGGTGGCTGCCTTAGGTAGTGTATGACTATGGCCGGGAAGGTCTGAGCATCTTTTCTGGTC
TGGCCAACCTCACGAGGTATCTGCATTGCACACGAGGGCCTGGTGCCACAACATGACACTAGTGGCCAACAATTGGGC
AAGGTGGTGGATGTGCTACGCCAAGTGAACCAAAGCAAAGTACAGGTGGTGGTGTGCTGTTTGCATCTGCCCGTGTGT
CTACTCCCTTTTTAGCTACAGCATCCTTCATGACCTCTCACCCAAGGTATGGGTGGCCAGTGAGTCCTGGCTGACCT
CTGACCTGGTCATGACACTTCCCAATATTGCCCCGTGTGGGCACTGTTCTTGGGTTTCTGCAGCGCGGTGCCCTACTG
CCTGAATTTTCCATTATGTGGAGACTCGCCTTGCCCTAGCTGCTGACCCAACATTCTGTGCCTCCCTGAAAGCTGA
GTTGGATCTGGAGGAGCGCGTGATGGGGCCACGCTGTTCACAATGTGACTACATCATGCTACAGAACCTGTATCTG
GGCTGATGCAGAACCTATCAGCTGGGCAGTTGCACCACCAAATATTTGCAACCTATGCAGCTGTGTACAGTGTGGCT
CAGGCCCTTCACAACACCCTGCAGTGCAATGTCTCACATTGCCACACATCAGAGCCTGTTCAACCCTGGCAGCTCCT
GGAGAACATGTACAATATGAGTTTCCGTGCTCGAGACTTGACACTGCAGTTTGATGCCAAAGGGAGTGTAGACATGG
AATATGACCTGAAGATGTGGGTGTGGCAGAGCCCTACACCTGTACTACATACTGTAGGCACCTTCAACGGCACCCCTT
CAGCTGCAGCACTCGAAAATGTATTGGCCAGGCAACCAGGTGCCAGTCTCCCAGTGCTCCCGGCAGTGCAAAGATGG
CCAGGTGCGCAGAGTAAAGGGCTTTCATTCTGCTGCTATGACTGTGTGGACTGCAAGGCAGGGAGCTACCGGAAGC
ATCCAGATGACTTCACCTGTACTCCATGTGGCAAGGATCAGTGGTCCCCAGAAAAAAGCACAACTGCTTACCTCGC
AGGCCCAAGTTTCTGGCTTGGGGGGAGCCAGCTGTGCTGTCACTTCTCCTGCTGCTTTGCCTGGTGTGCTGGGCCTGAC
ACTGGCTGCCCTGGGGCTCTTTGTCCACTACTGGGACAGCCCTCTTGTTTCAGGCCTCAGGTGGGTCACTGTTCTGCT
TTGGCCTGATCTGCCTAGGCCTCTTCTGCCTCAGTGTCTTCTGTTCCCAGGACGACCACGCTCTGCCAGCTGCCTT
GCCCAACAACCAATGGCTCACCTCCCTCTCACAGGCTGCCTGAGCACACTCTTCCTGCAAGCAGCCGAGATCTTTGT
GGAGTCTGAGCTGCCACTGAGTTGGGCAAACTGGCTCTGCAGCTACCTTCGGGGCCCCCTGGGCTTGGCTGGTGGTAC
TGCTGGCCACTCTTGTGGAGGCTGCACTATGTGCCTGGTACTTGATGGCTTTCCCTCCAGAGGTGGTGACAGATTGG
CAGGTGCTGCCCACGGAGGTAAGTGAACACTGCCGCATGCGTTCCTGGGTCAGCCTGGGCTTGGTGACATCACCAA
TGCAGTGTTAGCTTTCCTCTGCTTTCTGGGCACTTTCCTGGTACAGAGCCAGCCTGGTCGCTATAACCGTGCCCCGTG
GCCTCACCTTCGCCATGCTAGCTTATTTTCATCATCTGGGTCTCTTTGTGCCCCCTCCTGGCTAATGTGCAGGTGGCC
TACCAGCCAGCTGTGCAGATGGGTGCTATCTTATTCTGTGCCCTGGGCATCCTGGCCACCTTCCACCTGCCCCAATG
CTATGTACTTCTGTGGCTGCCAGAGCTCAACACCCAGGAGTTCTTCCTGGGAAGGAGCCCCAAGGAAGCATCAGATG
GGAATAGTGGTAGTAGTGAGGCAACTCGGGGACACAGTGAATGA (SEQ ID NO 13)

[0282] rT1R3 推定的氨基酸序列 (SEQ. ID NO. 14)

[0283] MPGLAILGLSLAAFLELGMGSSLCLSQQFKAQGDYILGGLFPLGTTEEATLNQRTQPNGILCTRFSPGL
LFLAMAMKMAVEEINNGSALLPGLRLGYDLFDTCSEPVVTMKPSLMFMAKVGSSQIAAYCNYTQYQPRVLAVIGPHS
SELALITGKFFSFFLMPQVSYSASMDRLSDRETFPSPFFRTVPSPDRVQLQAVVTLLQNFSWNWVAALGSDDDYGREGL

SIFSGLANSRGICIAHEGLVPQHDTSGQQLGKVVDVLRQVNQSKVQVVVLFASARAVYSLFSYSILHDLSPKVWVAS
ESWLTSDLVMTLPNIARVGTVLGFLQRGALLPEFSHYVETRLALAADPTFCASLKAELDLEERVMPGRCSQCDYIML
QNLSSGLMQNLSAGQLHHQIFATYAAVYSVAQALHNTLQCNVSHCHTSEPVPWQLLENMYNMSFRARDLTLQFDAK
GSVDMEYDLKMWWVQSPTPVLHTVGTENGTLQLQHSKMYWPGNQVPVSQCSRQCKDGQVRRVKGFHSCCYDCVDCKA
GSYRKHPDDFTCTPCGKDQWSPEKSTTCLPRRPKFLAWGEPVLSLLLLLCLVLGLTLAALGLFVHYWDSPLVQASG
GSLFCFGLICLGLFCLSVLLFPGRPRSASCLAQQPMAHLPLTGCLSTLFLQAAEIFVESELPLSWANWLCSYLRGPW
AWLVVLLATLVEAALCAWYLMAFPPEVVTDWQVLPTEVLEHCRMRSWVSLGLVHITNAVLAFLCFLGTFLVQSQPGR
YNRARGLTFAMLAYFIIWVSFVPLLANVQVAYQPAVQMGAILFCALGILATFHLPKCYVLLWLPELNTQEFFLGRSP
KEASDGNSSSEATRHS (SEQ ID NO 14)

[0284] 实施例 4-mT1R3 的表达

[0285] 将上述包含于 SAV115 中的小鼠 T1R3 片段用 M13 正向和 M13 反向引物进行 PCR 扩增,然后凝胶纯化。将 T1R3DNA 模板置于体外转录标记反应体系中,其中洋地黄毒苷标记的 UTP 被整合入反义 cRNA 探针中。使该探针与含有轮廓状乳头的成年小鼠味觉组织进行杂交。T1R3 原位杂交和探测是根据 Schaeren-Wiemers 等人, Histochemistry, 100 : 431-400 (1993) 的规程进行的。简要地,将新鲜的冷冻小鼠舌做 14 μ m 的切片并准备杂交。使 200ng/mL 的反义洋地黄毒苷 T1R3 探针在 72°C 杂交 14 个小时。杂交后用 0.2xSSC 于 72°C 洗涤。洋地黄毒苷的探测是通过与 anti-DIG 碱性磷酸酶抗体的 1:5000 稀释液进行温育并随后在 NBT/BCIP 中与磷酸酶的 12 个小时的反应而实现。

[0286] 实施例 5-hT1R1

[0287] 大鼠味觉受体的人定向同源物 (数据库访问号 AL159177) 在下面提供为 SEQ ID NO 15,指定为 rT1R1。预测的 cds 以粗体显示,且一些内含子序列间隔表示为一组 N。核苷酸和推定翻译的 hT1R1 序列也分别在此处描述为 SEQ ID NO 16 和 17。

[0288] hT1R1 基因组 DNA (SEQ ID NO 15)

[0289] GAGAATCTCGCGAGATCCCGTCCGTCGCCCCGCTGCCCTCCCAGCTGCCGAAAAGAGGGGCCTCCG
AGCCGCCGGCGCCCTCTGCCGGAACCTCCGGAAGCACACTAGGAQGTTCAGCCGATCTGGTCGAGGGGCTCCA
CGGAGGACTCCATTTACGTTACGCAAATTCCTACCCAGCCGGCCGAGAGAGAAAAGCCAGAAACCTCGCGACC
AGCCATGGGCCACCTCTCCGAAAAACACCGGGATATTTTTTTTCTCCTGCAGAAAAAGCTTTAGGATTGGCAGT
TTAAACAAAACATGTCTATTTGCATACCTTCGGTTTGCATGCATTTGTTTCGAAGTGAGCAACCCTGGGTAACAA
GGCGAAAGTATATGACAATTTGCTCAGAATCTTAATGTCAGAAACTGGAGACTGGGGCAGGGGGGTGTCGACTC
AAAGCTGTGTCTCATTTAGTAACTGAGGCCAGGTAAAAAGTTCTGAAACCTCGCAACACCCGGAGAAATTGTG
TTCCAGCCTCCCACCTCGCCCCAAAATGCCAGAGCTCCTTTTCTAAGCCAGGTGAAGTCACAGAGCGTGGACAGA
ACCCACAACCGTCCAGAGGAAGGGTCACTGGGTGCCACCTGGTTTGCATCTGTGCCTTCGTCCTGCCAGTTCCT
GAGTGGGACCGCAGGCCCGGAATGTCAAGGCAAACAGTCCTGCTTCAGCCACTGGGCTCCAGTCCCACCCCTTT
TGGGGGCCTGAAGTTAGGAAGCATCCGGCAGCTGCCTTCTATTTAAGCAACTGGCCTCCTTAGAGGCCACTCCT
TGGCCATGCCAGGCGCGGGCATCTGGCCAGC :**ATGCTGCTCTGCACGGCTCGCCTGGTCCGGCCTG**
CAGCTTCTCATTTCTGCTGCTGGGCCTTTGCCTGCCATAGCACGGAGTCTT
CTCCTGACTTCACCCTCCCCGGAGATTACCTCCTGGCAGGCCTGTTCCCTCT
CCATTCTGGCTGTCTGCAGGTGAGGCACAGACCCGAGGTGACCCTGTGTGA CAG
GTGAGTGAGGGGCCAGCAGAGCCACACTTAGTGGGACCCCTGGCTATAGGGCCCCTCTGGCTGCCATCCTCC

AAACAGGACCTTGCCTCTGCCTTTGCCCCCTTGAACCTGTCCCCAGGCCTTGTTTCATCAATCCACTTGCCACCT
AAGTGCTGGCTAGACCTTCCTAGACACTTCGGCCAGTTTCCAATTATTTACCCCTTGCTGTTAGAATGTNNN
NNNAATTCCTTAAACTAAATTTCTCACT
TTCTCTCTCTCTCTGGAAAACACTGACTAATGTAGCAGGTTTCTCTGCTCCAGGACTTCAGGACCTTTTCGA
TGCTAATAAGTTTCTCCATCAGGGCCAGCTTGTTCTCCTACTGAGCTTGAGAGCCCTTGTTGAAGTTGTGG
TTTGGGGGACTGGACCGATGACCTCAAAGGTTCCCTTTGCTCCCAAGCCTCAGAGTCTAGGAGGCCAGAGGG
TCTCAGCAGGCCTTTGTCTTCTCAGCTGTCTCTTACTGGCTTTCTCCACAG **GTCTTGTAGCTTCA**
ATGAGCATGGCTACCACCTCTTCCAGGCTATGCGGCTTGGGGTTGAGGAGA
TAAACAACCTCCACGGGCCCTGCTGCCCAACATCACCTGGGGTACCAGCTGT
ATGATGTGTGTTCTGACTCTGCCAATGTGTATGCCACGCTGAGAGTGCTCTC
CCTGCCAGGGCAACACCACATAGAGCTCCAAGGAGACCTTCTCCACTATTCC
CCTACGGTGCTGGCAGTGATTGGGCCTGACAGCACCAACCGTGCTGCCACC
ACAGCCGCCCTGCTGAGCCCTTCTCTGGTGCCCATG GTAAGCTGGAGCCTCAGACCT
TTGCCCATCTCCCTTCAGGCAAGTCTGGGNN
NNNNNGCCACCATGCCCGGCTAATTTTTTTTGTATTTTGTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGCT
GGTCGCAAACTCCTAACCTCGTGATCCACCCACCTCGGCCTCCCAATGTGCTGGGATTACAGGTGTGAGCCAC
TGCACCCGGCCATAATGTATTAATATAATAAAATAATTATACAACCTCACCATATGTAGAATCAGTGGGAGCC
CTGAGCTTGTTTTCTACAACCTAGATGGTCCCCTCTGGGGGTGATGGGAGACAGTGACAGATCATCAGACATT
AGATTCTCATAAGTAGCGTGCAACCCAGATCCCTCGCATGTGCAGTTCACAGTAGGGTTCAAGCTCCTACAAG
AATCTGATGCTGCTGCTGATCTGACAGGAGGGGAGCAGCTGTAAATACAGATGAAGCTTCGCTTACTCACAG
CTGCTCACCTCCTCCTGTGAGGCCCGGTTCTAACAGGCCACTGACCTAACTTCTGCCCTGACCTACACATGC
TTCTCTTCTCCTTGCAAACTGCCTCCAGTGGAAGTCCCTGAAGGTCCCCAAACACACGGGACTATTTCACTC
CTATGCAGGTTTTGTCTCCTTTGCTTGGAATGCATCCCCTCACCCCTTGTCCTCCAGGAGATTCCCACCCCT
CCCCCAGAACCTGCCCCAGTGGAGCCTTCGCAGGTGATTTGTCAGTTTCACAGGCTGAGGGGTGCTCTCCTG
GTCTCCCCGGCTCCCTGTATCCCCACACCCAGCACAGGGCCAGGCACTGGGGGGGCCTTCAGTGGAGACTGA
AATGGCTGAACGGGACCTCCCATAG **ATTAGCTATGCGGCCAGCAGCGAGACGCTC**
AGCGTGAAGCGGCAGTATCCCTCTTCTCTGCGCACCATCCCCAATGACAAGT
ACCAGGTGGAGACCATGGTGCTGCTGCTGCAGAAGTTCGGGTGGACCTGGA
TCTCTCTGGTTGGCAGCAGTGACGACTATGGGCAGCTAGGGGTGCAGGCAC

TGGAGAACCAGGCCACTGGTCAGGGGATCTGCATTGCTTTCAAGGACATCAT
GCCCTTCTCTGCCCAGGTGGGCGATGAGAGGATGCAGTGCCTCATGCGCCA
CCTGGCCCAGGCCGGGGGCCACCGTCGTGGTTGTTTTTTCCAGCCGGCAGTT
GGCCAGGGTGTTTTTCGAGTCCGTGGTGCTGACCAACCTGACTGGCAAGGT
GTGGGTGCGCTCAGAAGCCTGGGCCCTCTCCAGGCACATCACTGGGGTGCC
CGGGATCCAGCGCATTGGGATGGTGCTGGGCGTGGCCATCCAGAAGAGGG
CTGTCCCTGGCCTGAAGGCGTTTGAAGAAGCCTATGCCCGGGCAGACAAGA
AGGCCCTAGGCCTTGCCACAAGGGCTCCTGGTGCAGCAGCAATCAGCTCT
GCAGAGAATGCCAAGCTTTCATGGCACACACGATGCCCAAGCTCAAAGCCTT
CTCCATGAGTTCTGCCTACAACGCATACCGGGCTGTGTATGCGGTGGCCCAT
GGCCTCCACCAGCTCCTGGGCTGTGCCTCTGGAGCTTGTTCCAGGGGGCCGA
GTCTACCCCTGGCAG GTAAGAGAGCCCACCCAGCACCTCCTGTCAGGGAGAACAGCCAATCCTGAGATG
AGCAGAGTGGGCACTCTCCGGTCACTCTAAATGCCAAGGGGGATAAATGCCACTAAGTTGAGGTTTTTTGTTTTGTT
TTGTTTTGTTTTTTGAGACAGTCTGGCTCTGTCACCCAGGCTGCAGTGTAGTGATGCGATCTCGGCTCTCTGCAACT
TCCACCTCCTGGGTTCAAGTGATTCTCTTGCCCTCGGCCTCCTGAGTAGCTGGGATTACAGGCACCCACCACCATGCC
TGGATAATTTTTCTTTCTTTTTTTTTTTTTTTGAGATAGAGTCTCGCTCTGTTGCCAGGCTGGAATGCAGTGGTGC
GATCTTGGCTCACTGTGAGCTCCGCCTCCCAGGTTCACTCCATCCCCCTGCCTCAGCCTCCCAAGTAGGTGGGACTA
CGGGCGCCCGCCACCACGCCAGCTAATTTTTTTTGTATTTTGTAGTAGAGACGGGGTTTACCATGTTAGCCAGGAT
GGTCTCAATCTCCTGACCTTGTATCCGCCACCTCGTCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCA
CCCGGCCTAATTTTTGTATTTTGTAGTAGAGATGGGGTTTACCATGTTGGCCAGGCTGGTCTCGAACTCCTGGCATC
AAGTGATCCTCCTGCTTCGGCCTCCCAAAGTGCTGGGATTACAGGCATTAGCTCTCTTCTCTTAGACAGATCTTTCT
CTCTGATCCTTGCCCTTCTCTACCCACTGTGTCTTGAAGTGTCAAGTGATAAGATCCAGGGCTAAACTGTCTGTA
AAGGAGTGTTTGTAGAGGCCTCCTCTCAGGAGGTTGGTGGGGAAGATTGAGGGGCTTCCTAAGAAGGAAGGGACGA
GACCTTCCTGATGGGCTGAAACCACCAGGACGGAACCCAGGAAGGCCCCAGGCCCTTGCTTCTGGGACCATGTGGG
TCTGTGCTGTCTGTGGTGGCTTCATGATACGCGTTTCTTTCAG **CTTTTG**

GAGCAGATCCACAAGGTGCATTTCTTCTACACAAGGACACTGTGGCGTTTA
ATGACAACAGAGATCCCCTCAGTAGCTATAACATAATTGCCTGGGACTGGAA
TGGACCCAAGTGGACCTTCACGGTCCTCGGTTCTCCACATGGTCTCCAGTT G
CAGCTAAACATAAATGAGACCAAATCCAGTGGCACGGAAAGGACAACCAG
TAATGGGGATGTGGCTACTACCATGTAAGTGGCTTATGGGCAACCTAGAGCCTGGGGGTGATGCTGACACAGTG
TACAGGGAGCAGGAGGGGGGGCCCCAGGGGTCCAGCTGCCACCACTCTACCCATCCTGGCCAGGGAAGCAGGGAA
GACACTCCGTAGGCGAGTGTGCAGATGCCCTGGGGCGGAAGTTCACACGACCAGGGGCCCTGCCCTGGGAGTGA
GCCCTGAGGGCAGATGCACAGAGATTCTGTTTTCTGTTCCACATGTGAGCTGTCCTTTGACTTGGGCCCTACG
TGTGGCCCCCTCTGGCTTCTTACAG **GTGCCTAAGTCTGTGTGTTCCAGCGACTGTCTTGAAG**
GGCACCAGCGAGTGGTTACGGGTTTCCATCACTGCTGCTTTGAGTGTGTGCC
CTGTGGGGCTGGGACCTTCCTCAACAAGAGTG GTGAGTGGGCAATGGAGCAGGCGAGCTACC

CAGCACTCCCGGGGGCTGCACGGTGGAGGGAGGGCCTCCCTTGGGGCCCCATGTGCCCTGCCCCAGAACCAAGGCCCA
GTCAGTGGGCTGCCAGTTAGCTTCAGGTTGGAGGACACCTGCTACCAGACAGAATTCTGATCAAGAGAATCAGCCAC
TGGGTGCGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGTGGATCACTTGAGGTCGGGAGTTC
GAGACCAGCCTGGCCAACATGGTGAAACCCCATCTCTACCAAAAAATATAAAAAATTAGCTGGGTGTGGTGGCGCGTG
CCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCACTTGAACCCAGGAGGCGGAGGTTGCAGTGAGCCAA
GATGCATTCCAGCCTGGACCACAAAGCGAGAATTCGTCCCCCAAAAAAGAAAGGAGGCCGGGCGCGGTGGCTCAC
ACCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGTGGATCACCTGAGGTCAGGAGTTCGAGACCAGCCTGACCAA
CATGGTGAAACCCCATCTCTACTAAAAATACAAAAAAGTTAGCCGGGCGTTGTGGCGTGTGCCTGTAATTCCAGCT
ACTCGGGAGGCTGAGGCAGGAGAATTGCTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCAAGATTGCACCATTGCA
CTCCAGCCTGGGCGACAAGAGAAAACTCTGTCTCAAAAAAAGAAAGAAAGAAAGAATTAGCCAACTGAAAGCCT
TAGACTGAGGTGTGTCTCTGTTAGAGAGCTGTCAACAACTCCTACAAAAGCAGTCGTATCCTGAATTCAACCTC
TTTCTCTAAATGAATATAGCTATTGTTCCCTTTGTGCCCTCTTGTCCCTACTGTCCCTTCTGTTGCCCATGCCAAAGA
CAGCTAGCTCCTTGAACAGCTTGGCCTGAATACAGATACTAGCGTGTCTGCAGCAGAGAAAAAACAGCATTCCCCA
TCCAGAAATGCAAGGTCAAGAACAGAGAGCAAAATTAGGTAGCTAAGGACTCAGGTCCTTAGTTGGTGTCCAGGGGCC
ACATTCTTTCCCTTTCACCATCTCTGTAGGGACAGGAATACTTCCCTTCTGTCTCAGAGGGTCAGGACTCAGAGAA
ACCACAGAGCAGCAGCTCAGGAAAGTGGTTCATGGAAATGCTGGCAAGAGAGAGGGGTACAATGCCCTCCCTTGG
GAGCAGGCTGCTCCCATCAGATCGTAACCTCTCTGGTATGTGGGCAGAGCTACCAGGTAAAGGTCCTCCCTAGGGT
TTGCAAAACCTCATGGGATCATGAGCCATACAGAACCGACCTGTGTGTCTCCAGAGTCTGTAATTAACACAGGCA
TTTTGAGGAAATGCGTGGCCTCAGGCCCCACTCCCGGCTACCCCCATCCCACTATGCCTAGTATAGTCTAGCTGCC
CTGGTACAATTCTCCAGTATCTTGCAGGCCCTATTTCTATTCTACTCTGCTCATCTGGCTCTCAGGAACCTT
CTTGGCCTTCCCTTTCAG **ACCTCTACAGATGCCAGCCTTGTGGGAAAGAAGAGTGGGCA**

51

CCAGGCACATCACTGGGGTGGCCGGATCCAGCGCATTTGGGATGGTGCTGGGCGTGGCCATCCAGAAGAGGGCTGTC
CCTGGCCTGAAGGCGTTTGAAGAAGCCTATGCCCGGGCAGACAAGAAGGCCCTAGGCCTTGCCACAAGGGCTCCTG
GTGCAGCAGCAATCAGCTCTGCAGAGAATGCCAAGCTTTCATGGCACACACGATGCCCAAGCTCAAAGCCTTCTCCA
TGAGTTCTGCCTACAACGCATACCGGGCTGTGTATGCGGTGGCCCATGGCCTCCACCAGCTCCTGGGCTGTGCCTCT
GGAGCTTGTTCCAGGGGCCGAGTCTACCCCTGGCAGCTTTTGGAGCAGATCCACAAGGTGCATTTCTTCTACACAA
GGACACTGTGGCGTTTAATGACAACAGAGATCCCCCTAGTAGCTATAACATAATTGCCTGGGACTGGAATGGACCCA
AGTGGACCTTCACGGTCCTCGGTTCTCCACATGGTCTCCAGTTCAGCTAAACATAAATGAGACCAAAATCCAGTGG
CACGGAAAGGACAACCAGGTGCCTAAGTCTGTGTGTTCCAGCGACTGTCTTGAAGGGCACCAGCGAGTGGTTACGGG
TTTCCATCACTGCTGCTTTGAGTGTGTGCCCTGTGGGGCTGGGACCTTCCTCAACAAGAGTGACCTCTACAGATGCC
AGCCTTGTGGGAAAGAAGAGTGGGCACCTGAGGGAAGCCAGACCTGCTTCCCGCGCACTGTGGTGTTTTGGCTTTG
CGTGAGCACACCTCTTGGGTGCTGCTGGCAGCTAACACGCTGCTGCTGCTGCTGCTGCTGGGACTGCTGGCCTGTT
TGCCTGGCACCTAGACACCCCTGTGGTGAGGTCAGCAGGGGGCCGCTGTGCTTTCTTATGCTGGGCTCCCTGGCAG
CAGGTAGTGGCAGCCTCTATGGCTTCTTTGGGGAACCCACAAGGCCTGCGTGCTTGCTACGCCAGGCCCTCTTTGCC
CTTGGTTTACCATCTTCTGTCTGCTGACAGTTTCGCTCATTCCTCAACTAATCATCATCTTCAAGTTTCCACCAA
GGTACCTACATTCTACCACGCCTGGGTCCAAAACACGGTGCTGGCCTGTTTGTGATGATCAGCTCAGCGGCCAGC
TGCTTATCTGTCTAACTTGGCTGGTGGTGAGACCCACTGCCTGCTAGGGAATACCAGCGCTTCCCCATCTGGTG
ATGCTTGAGTGACAGAGACCAACTCCCTGGGCTTCATACTGGCCTTCCTCTACAATGGCCTCCTCTCCATCAGTGC
CTTTGCCTGCAGCTACCTGGGTAAGGACTTGCCAGAGAACTACAACGAGGCCAAATGTGTCACCTTCAGCCTGCTCT
TCAACTTCGTGTCCTGGATCGCCTTCTTACCACGGCCAGCGTCTACGACGGCAAGTACCTGCCTGCGGCCAACATG
ATGGCTGGGCTGAGCAGCCTGAGCAGCGGCTTCGGTGGGTATTTTCTGCCTAAGTGCTACGTGATCCTCTGCCGCCC
AGACCTCAACAGCACAGAGCACTTCCAGGCCTCCATTAGGACTACACGAGGCGCTGCGGCTCCACCTGA (SEQ ID
NO 16)

[0292] hT1R1 推定的氨基酸序列 (SEQ ID NO 17)

[0293] MLLCTARLVGLQLLISCCWAFACHSTESSPDFTLPGDYLLAGLFPLHSGCLQVRHRPEVTLCDRSCSFN
EHGYHLFQAMRLGVEEINNSTALLPNI TLGYQLYDVCSDSANVYATLRVLSLPGQHHLIELQGDLLHYSPTVLAVIGP
DSTNRAATTAALLSPFLVPMISYAASSETLSVKRQYPSFLRTIPNDKYQVETMVL LLQKFGWTWISLVGSSDDYGQL
GVQALENQATGGGICIAFKDIMPFAQVQDERMQCLMRHLAQAGATVVVVVFSSRQLARVFFESVVLTNLTGKVVVAS
EAWALSRHITGVPGIQRIGMVLGVAIQKRAVPGLKAFEEAYARADKKAPRPCHKGSWCSSNQLCRECQAFMAHTMPK
LKAFSMSSAYNAYRAVYAVAHGLHQLLGCSAGCSRGRVYPWQLEQIHKVHFLHLDTVAFNDNRDPLSSYNI IAW
DWNPGKWTFTVLGSSTWSPVQLNINETKIQWHGKDNQVPKSVCSDDCLEGHQRVVTGFHHCCFECVPCGAGTFLNKS
DLYRCQPCGKEEWAPEGSQTCFPRTVVFLALREHTSWVLLAANTLLLLLLLTAGLFAWHLDTPVVRSAGGRLCFLM
LGSLAAGSGSLYGFFGEPTRPACLLRQALFALGFTIFLSCLTVRSFQLIIIFKFSTKVPTFYHAWVQNHGAGLFVMI
SSAAQLLICTLWLVVWTPLPAREYQRFPFLVMLECTETNSLGFILAFLYNGLLSISAFACSYLGKDLPENYNEAKCV
TFSLLFNFSWIAFFTTASVYDGKYLPAANMMAGLSSSLSSGFGGYFLPKCYVILCRPDLNSTEHFQASIQDYTRRCG
ST (SEQ ID NO 17)

[0294] 实施例 6-hT1R2

[0295] 大鼠味觉受体的人定向同源物的预测的 cds 在下面提供为 SEQ ID NO 20, 指定为
rT1R2。hT1R2 序列的推定翻译序列也在此处描述为 SEQ ID NO 21。根据本发明, 通过 DNA
印迹将 hT1R2 的最初两个外显子鉴定为在 PAC 中。外显子 1 是在 DNA 印迹实验中鉴定的

BamH I/BgIII 片段中分离的,而外显子 2 是在跨越外显子 1 和外显子 3 的 PCR 产物中分离的。对 rT1R2 序列与这最初两个外显子的比较确定该两个外显子编码 rT1R2 的人对应物的 N- 末端。例如,由两个外显子编码的 hT1R2N- 末端序列和 rT1R2 的相应区域之间的成对氨基酸一致性为约 72%,而公共 DNA 序列数据库中最相关的注释的序列与 hT1R2 仅有约 48% 的一致性。

[0296] hT1R2 预测的 cds (SEQ ID NO 20)

[0297] ATGGGGCCCAGGGCAAAGACCATCTGCTCCCTGTTCTTCCTCCTATGGGTCTGGCTGAGCCGGCTGAG
AACTCGGACTTCTACCTGCCTGGGGATTACCTCCTGGGTGGCCTCTTCTCCCTCCATGCCAACATGAAGGGCATTGT
TCACCTTAACCTTCTGCAGGTGCCCATGTGCAAGGAGTATGAAGTGAAGGTGATAGGCTACAACCTCATGCAGGCCA
TGCGCTTCGCGGTGGAGGAGATCAACAATGACAGCAGCCTGCTGCCTGGTGTGCTGCTGGGCTATGAGATCGTGGAT
GTGTGCTACATCTCCAACAATGTCCAGCCGGTGCTCTACTTCCTGGCACACGAGGACAACCTCCTTCCCATCCAAGA
GGACTACAGTAACTACATTTCCCGTGTGGTGGCTGTCAATTGGCCCTGACAACTCCGAGTCTGTCACTGCTGGCCA
ACTTCCTCTCCCTATTTCTCCTTCCACAGATCACCTACAGCGCCATCAGCGATGAGCTGCGAGACAAGGTGCGCTTC
CCGGCTTTGCTGCGTACCACACCCAGCGCCGACCAACGTCGAGGCCATGGTGCAGCTGATGCTGCACTTCCGCTG
GAACTGGATCATTGTGCTGGTGAGCAGCGACACCTATGGCCGCGACAATGGCCAGCTGCTTGGCGAGCGCGTGGCCC
GGCGCGACATCTGCATCGCCTTCCAGGAGACGCTGCCACACTGCAGCCCAACCAGAACATGACGTCAGAGGAGCGC
CAGCGCCTGGTGACCATTGTGGACAAGCTGCAGCAGAGCACAGCGCGCGTCTGGTCTGTCTCGCCCGACCTGAC
CCTGTACCACTTCTTCAATGAGGTGCTGCGCCAGAACTTCACGGGCGCCGTGTGGATCGCCTCCGAGTCTTGGGCCA
TCGACCCGGTCTGCACAACCTCACGGAGCTGGGCCACTTGGGCACCTTCTTGGGCATCACCATCCAGAGCGTGCCC
ATCCCGGGCTTCAGTGAGTTCCGCGAGTGGGGCCACAGGCTGGGCGCCACCCCTCAGCAGGACCAGCCAGAGCTA
TACCTGCAACCAGGAGTGCGACAACCTGCCTGAACGCCACCTTGTCTTCAACACCATTTCTCAGGCTCTCTGGGGAGC
GTGTGCTCTACAGCGTGTACTCTGCGGTCTATGCTGTGGCCCATGCCCTGCACAGCCTCCTCGGCTGTGACAAAAGC
ACCTGCACCAAGAGGGTGGTCTACCCCTGGCAGCTGCTTGAGGAGATCTGGAAGGTCAACTTCACTCTCCTGGACCA
CCAAATCTTCTTCGACCCGCAAGGGGACGTGGCTCTGCACTTGGAGATTGTCCAGTGGCAATGGGACCGGAGCCAGA
ATCCCTTCCAGAGCGTCGCCTCTACTACCCCTGCAGCGACAGCTGAAGAACATCCAAGACATCTCCTGGCACACC
GTCAACAACACGATCCCTATGTCCATGTGTTCCAAGAGGTGCCAGTCAGGGCAAAAGAAGAAGCCTGTGGGCATCCA
CGTCTGCTGCTTCGAGTGCATCGACTGCCTTCCCGGCACCTTCTCAACCACACTGAAGATGAATATGAATGCCAGG
CCTGCCCCGAATAACGAGTGGTCTACCAGAGTGAGACCTCCTGCTTCAAGCGGCAGCTGGTCTTCTGGAATGGCAT
GAGGCACCCACCATCGCTGTGGCCCTGCTGGCCGCCCTGGGCTTCTCAGCACCCCTGGCCATCCTGGTGATATTCTG
GAGGCACTTCCAGACACCCATAGTTTCGCTCGGCTGGGGGCCCCATGTGCTTCTTGATGCTGACACTGCTGCTGGTGG
CATAATGGTGGTCCCGGTGTACGTGGGGCCGCCAAGTCTCCACCTGCCTCTGCCGCCAGGCCCTCTTTCCCCTC
TGCTTCACAATTTGCATCTCCTGTATCGCCGTGCGTTCTTCCAGATCGTCTGCGCCTTCAAGATGGCCAGCCGCTT
CCCACGCGCCTACAGCTACTGGGTCCGCTACCAGGGGGCCCTACGTCTCTATGGCATTTATCACGGTACTCAAAATGG
TCATTGTGGTAATTGGCATGCTGGCCACGGGCCTCAGTCCCACCAACCCGTAAGTACCCCCGATGACCCCAAGATCACA
ATTGTCTCCTGTAAACCCCAACTACCGCAACAGCCTGTGTTCAACACCAGCCTGGACCTGCTGCTCTCAGTGGTGGG
TTTCAGCTTCGCCTACATGGGCAAAGAGCTGCCCACCAACTACAACGAGGCCAAGTTCATCACCCCTCAGCATGACCT
TCTATTTACCTCATCCGTCTCCCTCTGCACCTTCATGTCTGCCTACAGCGGGGTGCTGGTCAACCATCGTGGACCTC
TTGGTCACTGTGCTCAACCTCCTGGCCATCAGCCTGGGCTACTTCGGCCCCAAGTGCTACATGATCCTCTTCTACCC
GGAGCGCAACACGCCCCGCTACTTCAACAGCATGATCCAGGGCTACACCATGAGGAGGGACTAG (SEQ ID NO. 20)

[0298] hT1R2 推定的氨基酸序列 (SEQ ID NO 21)

[0299] MGPRAKTICSLFFLLWVLAEPANSDFYLP GDYLLGGFLSLHANMKGIVHLNFLQVPMCKEYEVKVIGY
NLMQAMRFAVEEINNDSSLLPGVLLGYEIVDVCYISNNVQPVLYFLAHEDNLLPIQEDYSNYISRVVAVIGPDNSES
VMTVANFLSLFLLPQITYSAISDEL RDKVRFPALLRTTPSADHHVEAMVQLMLHFRWNWIIVLVSSDTYGRDNGQLL
GERVARRDICIAFQETLPTLQPNQNM TSEERQRLVTIVDKLQQSTARVVVVFSPDLTYHFFNEVLRQNFTGAVWIA
SESWAIDPVLHNLTELGH LGTFLGITIQSVPIPGFSEFREWGPQAGPPPLSRTSQSYTCNQECDNCLNATLSFNTIL
RLSGERVVYSVYSAVYAVAHALHSLGCDKSTCTKR VVYPWQLLEEIWKNFTLLDHQIFFDPQGDVALHLEIVQWQ
WDRSQNPFSVASYYPLQRQLKNIQDISWHTVNNTIPMSMCSKRCQSGQKKKPVGIHVCCFECIDCLPGTFLNHTED
EYECQACPNNEWSYQSETSCFKRQLVFLEWHEPTIAVALLAALGFLSTLAILVIFWRHFQTPIVR SAGGPMCFMLL
TLLLVAYMVVPVYVGPPKVSTCLCRQALFPLCFTICISCI AVRSFQIVCAFKMASRFPRAYSYWVRYQGPYVSMAFI
TVLKMVIVVIGMLATGLSPTTRTDPDDPKITIVSCNP NYRNSLLFNTSLDLLLSVVGFSFAYMGKELPTNYNEAKFI
TLSMTFYFTSSVSLCTFMSAYSGLVTIVDLLVTVLNLLAISLG YFGPKCYMILFYPERNTPAYFNSMIQGYTMRRD
(SEQ ID NO. 21)

[0300] 实施例 7

[0301] T1Rs 在异源细胞中异源表达的方法

[0302] 使稳定地表达 G α 15 的 HEK-293 衍生细胞 (Chandrashekar 等人, Cell 100 (6) : 703-11 (2000)) 于 37°C 生长并维持于 Dulbecco's Modified Eagle Medium (DMEM, Gibco BRL) 上, 该培养基补充了 10% 的 FBS、MEM 非必需氨基酸 (Gibco BRL) 和 3 μ g/ml 的杀稻瘟素。对于钙 - 成像实验, 首先将细胞接种于 24- 孔的组织培养平板上 (约每孔 10 万个细胞), 并用 Mirus TransIt-293 (PanVera) 通过脂质转染法进行转染。为了使谷氨酸诱导的和葡萄糖诱导的脱敏减到最小, 在转染后约 24 小时用低葡萄糖的 DMEM/GlutaMAX (Gibco BRL) 置换补加的 DMEM。24 个小时后, 在室温使细胞被钙染料 Fluo-4 (Molecular Probes) 装载 1.5 个小时, 该染料在 Dulbecco's PBS 缓冲液 (DPBS, Gibco BRL) 中为 3 μ M。在用 DPBS 置换之后, 通过添加 200 μ l 补充了味觉刺激物的 DPBS 于室温进行了刺激。钙的活动化是用 Imaging Workbench 4.0 软件 (Axon) 在 Axiovert S100TV 显微镜 (Zeiss) 上监控的。T1R1/T1R3 和 T1R2/T1R3 反应是显著地瞬时和不同时的, 其中钙的释放很少持续 15 秒之上。因而反应细胞的数目在整个过程中是相对恒定的; 因此, 通过在固定的时间点人工计数反应细胞的数目来测定细胞反应的数量, 一般是在刺激添加 30 秒之后测定的。

[0303] 实施例 8

[0304] 人 T1R2/T1R3 发挥甜味受体的功能

[0305] 用人 T1R2、T1R3 和 T1R2/T1R3 瞬时转染稳定表达 G α 15 的 HEK 细胞, 并测定响应于蔗糖浓度的增加的细胞内钙的增加 (图 1(a))。同样地, 对几种甜味刺激物确定了 T1R2/T1R3 的剂量反应 (图 1(b))。对于不同的增甜剂反应细胞的最大百分比是不同的, 在 10-30% 范围内。为清楚起见, 使剂量反应对反应细胞的最大百分比进行了标准化。图 1 中的值代表 4 种独立反应的平均值 $\pm$ 标准差。X- 轴的圆标记了由味觉检验确定的心理物理学探测阈值。Gurmarin (10g/l 过滤的 *Gymnema sylvestre* 水提取物的 50- 倍稀释液) 抑制 T1R2/T1R3 对 250mM 蔗糖的反应, 但不抑制内源 β 2- 肾上腺素能受体对 20 μ M 异丙基肾上腺素的反应 (图 1(b))。图 1(c) 含有 T1R2/T1R3 共表达细胞系对不同增甜剂 (蔗糖、阿斯巴甜、色氨酸和糖精) 的标准化的反应。

[0306] 实施例 9

[0307] 鼠 T1R2/T1R3 发挥甜味受体的功能

[0308] 用 hT1R2/hT1R3、rT1R2/rT1R3、hT1R2/rT1R3 和 rT1R2/hT1R3 瞬时转染稳定表达 $G_{\alpha 15}$ 的 HEK 细胞。然后测定这些转染的细胞响应于 350mM 蔗糖、25mM 色氨酸、15mM 阿斯巴甜和 0.05 应乐果甜蛋白的细胞内钙的增加。蔗糖和阿斯巴甜的结果包含于图 2 中,且显示 rT1R2/rT1R3 也发挥甜味受体的功能。同样地,这些结果表明 T1R2 可控制 T1R2/T1R3 配体特异性。

[0309] 实施例 10

[0310] 人 T1R1/T1R3 发挥 umami 受体的功能

[0311] [0227] 用人 T1R1、T1R3 和 T1R1/T1R3 瞬时转染稳定表达 $G_{\alpha 15}$ 的 HEK 细胞,并测定响应于谷氨酸浓度增加(图 3(a))和响应于 0.5mM 谷氨酸、0.2mM IMP 和 0.5mM 谷氨酸加 0.2mM IMP(图 3(b))的细胞内钙的增加。在 0.2mM IMP 存在与不存在时对谷氨酸确定了人 T1R1/T1R3 的剂量反应(图 3(c))。反应细胞的最大百分比对于谷氨酸为约 5%,且对于谷氨酸加 IMP 为约 10%。为清楚起见,使剂量反应对反应细胞的最大百分比进行了标准化。该值代表 4 种独立反应的平均值 $\pm$ 标准差。X-轴的圆标记了由味觉检验确定的心理物理学探测阈值。

[0312] 实施例 11

[0313] 作为输出序列的 PDZIP

[0314] 将六残基 PDZIP 序列(SVSTW(SEQ ID NO:22))与 hT1R2 的 C-末端融合,且将嵌合受体(即 hT1R2-PDZIP)转染入 HEK-293 宿主细胞。然后对 hT1R2 的表面表达用免疫荧光和 FACS 扫描数据进行监控。如在图 6A 和 6B 中所显示的,PDZIP 序列的包含增加了 hT1R2-PDZIP 相对于 hT1R2 的表面表达。

[0315] PKZIP 序列

[0316] SVSTVV(SEQ ID NO:22)

[0317] 更特定地,图 4 表示 myc- 标记的 hT1R2 的免疫荧光染色,证明 PDZIP 显著地增加了质膜上 hT1R2 蛋白质的量。图 4B 表示证明了同样结果的 FACS 分析数据。——表达 myc- 标记的 hT1R2 的细胞以虚线显示,且表达 myc- 标记的 hT1R2-PDZIP 的细胞以实线显示。

[0318] 实施例 12

[0319] 稳定共表达 T1R1/T1R3 或 T1R2/T1R3 的细胞系的生成

[0320] 通过将线性化的源自 PEAK10(Edge Biosystems)的载体(含有 pCDNA3.1)和源自 ZE0 的载体(Invitrogen)转染入 $G_{\alpha 15}$ 表达细胞系中而生成了稳定共表达人 T1R2/T1R3 或 T1R1/T1R3 的人细胞系,所述载体分别含有 hT1R1 或 hT1R2 表达构建体(对于 T1R1 为质粒 SAV2485,对于 T1R2 为 SAV2486)和 hT1R3(对于 T1R3 为质粒 SXV550)。特定地,T1R2/T1R3 稳定细胞系是通过将线性化的 SAV2486 和 SXV550 共转染入稳定表达 $G_{\alpha 15}$ 的 Aurora Bioscience' sHEK-293 细胞系中而生产的。T1R1/T1R3 稳定细胞系是通过将线性化的 SAV2485 和 SXV550 共转染入稳定表达 $G_{\alpha 15}$ 的相同 HEK-293 细胞系中而生产的。在 SAV2485/SXV550 和 SAV2486/SXV550 转染后,将嘌呤霉素抗性的和 zeocin- 抗性的群体选择、扩增并用钙成像检验其对甜或 umami 味刺激物的反应。在 0.0005mg/ml 嘌呤霉素(CALBIOCHEM)和 0.1mg/ml zeocin(Invitrogen)中于 37°C 从低葡萄糖的 DMEM 中选择细胞,该 DMEM 补充

了 GlutaMAX、10% 透析的 FBS 和 0.003mg/ml 杀稻瘟素。将抗性群体进行了扩增,且将其对甜味刺激物的反应用荧光显微镜进行了评估。为了在 VIPR-II 仪器 (Aurora Biosciences) 上进行自动荧光测定成像,首先在 96-孔的平板上接种了 T1R2/T1R3 稳定细胞 (每孔约 15,000 个细胞)。24 个小时之后,在室温使细胞被钙染料 fluo-3-AM (Molecular Probes) 装载 1 个小时,该染料在 PBS 中为 0.005mM。在用 70ml PBS 置换之后,通过添加 70ml 补充了味觉刺激物的 PBS 于室温进行了刺激。对在化合物添加后 20 ~ 30 秒的荧光 (480nm 激发和 535nm 发射波长) 反应进行了平均,对在化合物添加之前测量的背景荧光进行了校正,并使其对一种钙离子载体 0.001mM 的离子霉素 (CALBIOCHEM) 的反应进行了标准化。

[0321] 然后观察到当这些细胞系与甜味或 umami 接触时,对于活性克隆一般 80-100% 的细胞对味觉刺激物有反应。出乎意料地,单独细胞反应的等级显著地大于瞬时转染细胞的反应。

[0322] 基于该观察,本发明者用上述的 Aurora Bioscience' sVIPR 仪器通过自动荧光测定成像检验了 T1R 稳定细胞系的活性。2 个 T1R1/T1R3 和 1 个 T1R2/T1R3 细胞系的反应分别显示于图 5 和图 6 中。

[0323] 显著地,反应细胞增加的数目和增加的反应等级一起导致相对于瞬时转染的细胞的大于 10 倍的活性增加。(通过比较的方式,对于用 T1R2/T1R3 瞬时转染的细胞的离子霉素反应的百分数在最适条件下约为 5%)。此外,对于稳定表达的人 T1R2/T1R3 和 T1R1/T1R3 的剂量反应与人味觉探测阈值相关。这些稳定细胞系的强 T1R 活性提示它们非常适合用于对化学文库的高生产量的筛选以鉴定化合物,如小分子,该化合物调节甜或 umami 味觉受体并因此调节、增强、阻断或模拟甜或 umami 味觉。

[0324] 虽然前述的详细说明书描述了本发明的几个实施方案,但应理解上述的说明书仅仅是说明性的,且不限公开的本发明。本发明仅通过后面的权利要求进行限制。

[0001]

序列表

<110> ADLER, JON ELLIOT
LI, XIAODONG
STASZEWSKI, LENA
O'CONNELL, SHAWN
ZOZULYA, SERGEY

<120> T1R味觉受体及其编码基因

<130> 078003-0280681

<140> 10/035,045

<141> 2002-01-03

<150> 60/259,227

<151> 2001-01-03

<150> 60/284,547

<151> 2001-04-19

<160> 24

<170> PatentIn Ver. 2.1

<210> 1

<211> 876

<212> DNA

<213> 智人

<400> 1

```
agcctggcag tggcctcagg cagagtctga cgcgcacaaa ctttcaggcc caggaagcga 60
ggacaccact ggggccccag ggtgtggcaa gtgaggatgg caagggtttt gctaaacaaa 120
tcctctgccc gctccccgcc ccgggctcac tccatgtgag gccccagtcg gggcagccac 180
ctgccgtgcc tgttggaagt tgcctctgcc atgctgggcc ctgctgtcct gggcctcagc 240
ctctgggctc tcctgcaccc tgggacgggg gccccattgt gcctgtcaca gcaacttagg 300
atgaaggggg actacgtgct gggggggctg ttccccctgg gcgaggccga ggaggctggc 360
ctccgcagcc ggacacggcc cagcagccct gtgtgcacca ggtacagagg tgggacggcc 420
tgggtcgggg tcagggtgac caggtctggg gtgctcctga gctggggccg aggtggccat 480
ctgcggttct gtgtggcccc aggttctcct caaacggcct gctctgggca ctggccatga 540
aaatggccgt ggaggagatc aacaacaagt cggatctgct gcccgggctg cgcctgggct 600
acgacctctt tgatacgtgc tcggagcctg tgggtggccat gaagcccagc ctcatgttcc 660
tggccaaggc aggcagccgc gacatcgccg cctactgcaa ctacacgcag taccagcccc 720
gtgtgctggc tgtcatcggg cccactcgt cagagctcgc catggtcacc ggcaagtctt 780
tcagcttctt cctcatgcc cagtggggcg cccccacca tcaccaccc ccaaccaacc 840
cctgccccgt gggagcccct tgtgtcagga gaatgc 876
```

<210> 2

[0002]

<211> 2687

<212> DNA

<213> 智人

<400> 2

```

tacatgcacc ccacccagcc ctgccctggg agccctgtgt cagaagatgc tcttggcctt 60
gcaggtcagc tacggtgcta gcatggagct gctgagcgcc cgggagacct tcccctcctt 120
cttccgcacc gtgcccagcg accgtgtgca gctgacggcc gccgcggagc tgctgcagga 180
gttcggcttg aactgggtgg ccgccctggg cagcgacgac gactacggcc ggcagggcct 240
gagcatcttc tcggccctgg ccgcggcacg cggcatctgc atcgcgcacg agggccttgt 300
gccgtgccc cgtgccgatg actcgcggct ggggaagggt caggacgtcc tgcaccaggt 360
gaaccagagc agcgtgcagg tgggtgctgt gttcgcctcc gtgcacgccg cccacgccct 420
cttcaactac agcatcagca gcaggctctc gcccaggtg tgggtggcca gcgaggcctg 480
gctgacctct gacctggtca tgggctgcc cggcatggcc cagatgggca cgggtgcttg 540
cttctccag aggggtgccc agctgcacga gttccccag tacgtgaaga cgcacctggc 600
cctggccacc gaccggcct tctgctctgc cctggcgag agggagcagg gtctggagga 660
ggacgtggtg ggccagcgt gccgcagtg tgactgcatc acgtgcaga acgtgagcgc 720
agggctaaat caccaccaga cgttctctgt ctacgcagct gtgtatagcg tggcccaggc 780
cctgcacaac actcttcagt gcaacgcctc aggtgcccc gcgcaggacc ccgtgaagcc 840
ctggcaggtg agcccgggag atgggggtgt gctgtcctct gcatgtgccc aggccaccag 900
gcacggccac cagcctgag ctggaggtgg ctggcggtc agccccgtcc cccgccgca 960
gctcctggag aacatgtaca acctgacctt ccacgtgggc gggctgccgc tgcggttcga 1020
cagcagcgga aacgtggaca tggagtacga cctgaagctg tgggtgtggc agggctcagt 1080
gccaggtc cagcaggtgg gcaggttcaa cggcagcctc aggacagagc gcctgaagat 1140
ccgtggcac acgtctgaca accaggtgag gtgaggtgg gtgtgccagg cgtgcccgtg 1200
gtagcccccg cggcaggcg cagcctgggg gtgggggccc ttccagtctc ccgtgggcat 1260
gccacgcca gcagagccag accccaggcc tgtcgcaga agcccggtgc ccgtgctcg 1320
cggcagtgcc aggaggcca ggtcgcggg gtcaagggt tccactcctg ctgctacgac 1380
tgtgtggact gcgaggcggg cagctaccgg caaaaccag gtgagccgc ttccggcgag 1440
gcgggggtgg gaacgcagca ggggagggtc ctgccaagtc ctgactctga gaccagagcc 1500
cacagggtac aagacgaaca cccagcgccc ttctctctc tcacagacga catgcctgc 1560
accttttgtg gccaggatga gtggtcccc gagcgaagca cacgtgctt ccgccgagg 1620
tctcggttcc tggcatgggg cgagccggct gtgctgtgc tgctcctgct gctgagcctg 1680
gcgctgggcc ttgtgctggc tgctttgggg ctgttcgttc accatcgga cagccactg 1740
gttcaggcct cgggggggcc cctggcctgc tttggcctgg tgtgcctggg cctggctctg 1800
ctcagcgctc tctgttccc tggccagccc agccctgccc gatgcctggc ccagcagccc 1860
ttgtcccacc tcccgtcac gggctgctg agcacactct tctgcaggc ggccgagatc 1920
ttcgtggagt cagaactgcc tctgagctgg gcagaccggc tgagtggctg cctgcggggg 1980
ccctgggct ggctggtggt gctgctggcc atgctggtg aggtcgact gtgcacctg 2040
tacctggtg cttcccgc ggaggtggt acggactggc acatgctgcc cacggaggcg 2100
ctggtgcact gccgcacacg ctctgggtc agcttcggcc tagcgcacgc caccaatgcc 2160
acgtggcct ttctctgct cctgggcact ttctggtgc ggagccagcc gggctgctac 2220
aacctgccc gtggcctcac cttgccatg ctggcctact tcatcacctg ggtctcctt 2280
gtgcccctcc tggccaatgt gcaggtggtc ctaggccc cgtgcagat gggcgccctc 2340
ctgctctgt tctgggcat cctggctgcc ttccacctgc ccaggtgtta cctgctcatg 2400
cggcagccag ggtcaacac cccgagttc ttctgggag ggggcccctg ggatgccccaa 2460
ggccagaat acgggaacac aggaatcag gggaaacat agtgacccaa cctgtgatc 2520
tcagccccg tgaaccaga cttagctgcg atcccccca agccagcaat gaccctgtc 2580
tcgtacaga gaccctccc ctctaggttc tgacccagg ttgtctcctg accctgacct 2640

```

[0003]

cacagtggagc cctaggcctg gagcacgtgg acacccctgt gaccatc

2687

<210> 3

<211> 2559

<212> DNA

<213> 智人

<400> 3

atgctggggc ctgctgtcct gggcctcagc ctctgggctc tcctgcaccc tgggacgggg 60
 gccccattgt gcctgtcaca gcaacttagg atgaaggggg actacgtgct gggggggctg 120
 ttccccctgg gcgaggccga ggaggctggc ctccgcagcc ggacacggcc cagcagccct 180
 gtgtgcacca ggttctcctc aaacggcctg ctctgggcac tggccatgaa aatggccgtg 240
 gaggagatca acaacaagtc ggatctgctg cccgggctgc gcctgggcta cgacctctt 300
 gatacgtgct cggagcctgt ggtggccatg aagcccagcc tcatgttctt ggccaaggca 360
 ggacagccgc acatcgccgc ctactgcaac tacacgcagt accagccccg tgtgtgtggt 420
 gtcacggggc cccactcgtc agagctcgcc atggtcaccg gcaagtctt cagcttcttc 480
 ctcatgcccc aggtcagcta cgggtgctagc atggagctgc tgagcgcccg ggagaccttc 540
 cccctcttct tccgcaccgt gccacgcgac cgtgtgcagc tgacggccgc cgcggagctg 600
 ctgcaggagt tcggctggaa ctgggtggcc gccctgggca gcgacgacga gtacggccgg 660
 cagggcctga gcattctctc ggccctggcc gcggcacgcg gcattctcat cgcgcacgag 720
 ggccctggtg cgtgccccg tgccgatgac tcgcggtctg ggaagggtga ggacgtcctg 780
 caccaggatga accagagcag cgtgcagggt gtgctgtgtg tcgcctccgt gcacgcccgc 840
 cacgccctct tcaactacag catcagcagc aggtctctgc ccaagggtgt ggtggccagc 900
 gaggcctggc tgacctctga cctggtcatg gggtgcccc gcattggcca gatgggcacg 960
 gtgcttggtt tcctccagag ggttgcccag ctgcacgagt tccccagta cgtgaagacg 1020
 cacctggccc tggccaccga cccggccttc tgctctgccc tgggcgagag ggagcagggt 1080
 ctggaggagg acgtgtgtgg ccagcgctgc ccgagtggt actgcatcac gctgcagaac 1140
 gtgagcgagc ggctaaatca ccaccagacg ttctctgtct acgcagctgt gtatagcgtg 1200
 gccacggccc tgcacaacac tcttcagtgc aacgcctcag gctgccccgc gcaggacccc 1260
 gtgaagccct ggcagctcct ggagaacatg tacaacctga ctttcacgt gggcgggctg 1320
 ccgctgcggt tcgacagcag cggaaacgtg gacatggagt acgacctgaa gctgtgggtg 1380
 tggcagggtc cagtgcaccg gctccacgac gtgggcagggt tcaacggcag cctcaggaca 1440
 gagcgccctg agatccgctg gcacacgtct gacaaccaga agcccgtgtc ccggtgctcg 1500
 cggcagtgcc aggaggccca ggtgcgccgg gtcaagggggt tccactcctg ctgctacgac 1560
 tgtgtggact gcgaggcggg cagctaccgg caaaaccagc acgacatcgc ctgcacctt 1620
 tgtggccagg atgagtggc cccggagcga agcacacgct gcttccgccc caggctctcg 1680
 ttccctggcat ggggcgagcc ggctgtgctg ctgctgctcc tgctgtgtag cctggcgctg 1740
 ggcccttgctc tggctgcttt ggggctgttc gttaccatc gggacagccc actggttcag 1800
 gcctcggggg ggccccctggc ctgctttggc ctggtgtgct tgggcctggt ctgcctcagc 1860
 gtccctctgt tccctggcca gccacgcct gcccgatgcc tggcccagca gcccttgctc 1920
 cacctccgcg tcacgggctg cctgagcaca ctcttctgc aggcggccga gatcttcgtg 1980
 gagtcagaac tgcctctgag ctgggcagac cggctgagtg gctgcctgcg ggggccctgg 2040
 gcctggctgg tgggtgctgt ggccatgctg gtggaggtcg cactgtgcac ctggtacctg 2100
 gtggccttcc cgcggagggt ggtgacggac tggcacatgc tgcccacgga ggcgctgggt 2160
 cactgcgcga cacgtcctg ggtcagcttc ggccatagcg acgccacca tgccacgctg 2220
 gcccttctct gcttctggg cactttctg gtgcggagcc agccgggctg ctacaaccgt 2280
 gcccggtggc tcaccttgc catgctggcc tacttcatca cctgggtctc ctttgtgccc 2340
 ctccctggcca atgtgcagggt ggtcctcagg cccgcccgtc agatggggcg cctcctgctc 2400

[0004]

tgtgtcctgg gcatcctggc tgccttccac ctgccaggt gttacctgct catgcggcag 2460
 ccagggtca acacccccga gttcttcctg ggagggggcc ctggggatgc ccaaggccag 2520
 aatgacggga acacaggaaa tcaggggaaa catgagtga 2559

<210> 4

<211> 852

<212> PRT

<213> 智人

<400> 4

Met Leu Gly Pro Ala Val Leu Gly Leu Ser Leu Trp Ala Leu Leu His
 1 5 10 15

Pro Gly Thr Gly Ala Pro Leu Cys Leu Ser Gln Gln Leu Arg Met Lys
 20 25 30

Gly Asp Tyr Val Leu Gly Gly Leu Phe Pro Leu Gly Glu Ala Glu Glu
 35 40 45

Ala Gly Leu Arg Ser Arg Thr Arg Pro Ser Ser Pro Val Cys Thr Arg
 50 55 60

Phe Ser Ser Asn Gly Leu Leu Trp Ala Leu Ala Met Lys Met Ala Val
 65 70 75 80

Glu Glu Ile Asn Asn Lys Ser Asp Leu Leu Pro Gly Leu Arg Leu Gly
 85 90 95

Tyr Asp Leu Phe Asp Thr Cys Ser Glu Pro Val Val Ala Met Lys Pro
 100 105 110

Ser Leu Met Phe Leu Ala Lys Ala Gly Ser Arg Asp Ile Ala Ala Tyr
 115 120 125

Cys Asn Tyr Thr Gln Tyr Gln Pro Arg Val Leu Ala Val Ile Gly Pro
 130 135 140

His Ser Ser Glu Leu Ala Met Val Thr Gly Lys Phe Phe Ser Phe Phe
 145 150 155 160

Leu Met Pro Gln Val Ser Tyr Gly Ala Ser Met Glu Leu Leu Ser Ala
 165 170 175

Arg Glu Thr Phe Pro Ser Phe Phe Arg Thr Val Pro Ser Asp Arg Val
 180 185 190

Gln Leu Thr Ala Ala Ala Glu Leu Leu Gln Glu Phe Gly Trp Asn Trp
 195 200 205

[0005]

Val	Ala	Ala	Leu	Gly	Ser	Asp	Asp	Glu	Tyr	Gly	Arg	Gln	Gly	Leu	Ser	210	215	220	
Ile	Phe	Ser	Ala	Leu	Ala	Ala	Ala	Arg	Gly	Ile	Cys	Ile	Ala	His	Glu	225	230	235	240
Gly	Leu	Val	Pro	Leu	Pro	Arg	Ala	Asp	Asp	Ser	Arg	Leu	Gly	Lys	Val	245	250	255	
Gln	Asp	Val	Leu	His	Gln	Val	Asn	Gln	Ser	Ser	Val	Gln	Val	Val	Leu	260	265	270	
Leu	Phe	Ala	Ser	Val	His	Ala	Ala	His	Ala	Leu	Phe	Asn	Tyr	Ser	Ile	275	280	285	
Ser	Ser	Arg	Leu	Ser	Pro	Lys	Val	Trp	Val	Ala	Ser	Glu	Ala	Trp	Leu	290	295	300	
Thr	Ser	Asp	Leu	Val	Met	Gly	Leu	Pro	Gly	Met	Ala	Gln	Met	Gly	Thr	305	310	315	320
Val	Leu	Gly	Phe	Leu	Gln	Arg	Gly	Ala	Gln	Leu	His	Glu	Phe	Pro	Gln	325	330	335	
Tyr	Val	Lys	Thr	His	Leu	Ala	Leu	Ala	Thr	Asp	Pro	Ala	Phe	Cys	Ser	340	345	350	
Ala	Leu	Gly	Glu	Arg	Glu	Gln	Gly	Leu	Glu	Glu	Asp	Val	Val	Gly	Gln	355	360	365	
Arg	Cys	Pro	Gln	Cys	Asp	Cys	Ile	Thr	Leu	Gln	Asn	Val	Ser	Ala	Gly	370	375	380	
Leu	Asn	His	His	Gln	Thr	Phe	Ser	Val	Tyr	Ala	Ala	Val	Tyr	Ser	Val	385	390	395	400
Ala	Gln	Ala	Leu	His	Asn	Thr	Leu	Gln	Cys	Asn	Ala	Ser	Gly	Cys	Pro	405	410	415	
Ala	Gln	Asp	Pro	Val	Lys	Pro	Trp	Gln	Leu	Leu	Glu	Asn	Met	Tyr	Asn	420	425	430	
Leu	Thr	Phe	His	Val	Gly	Gly	Leu	Pro	Leu	Arg	Phe	Asp	Ser	Ser	Gly	435	440	445	
Asn	Val	Asp	Met	Glu	Tyr	Asp	Leu	Lys	Leu	Trp	Val	Trp	Gln	Gly	Ser	450	455	460	

[0006]

Val	Pro	Arg	Leu	His	Asp	Val	Gly	Arg	Phe	Asn	Gly	Ser	Leu	Arg	Thr	465	470	475	480
Glu	Arg	Leu	Lys	Ile	Arg	Trp	His	Thr	Ser	Asp	Asn	Gln	Lys	Pro	Val	485	490	495	
Ser	Arg	Cys	Ser	Arg	Gln	Cys	Gln	Glu	Gly	Gln	Val	Arg	Arg	Val	Lys	500	505	510	
Gly	Phe	His	Ser	Cys	Cys	Tyr	Asp	Cys	Val	Asp	Cys	Glu	Ala	Gly	Ser	515	520	525	
Tyr	Arg	Gln	Asn	Pro	Asp	Asp	Ile	Ala	Cys	Thr	Phe	Cys	Gly	Gln	Asp	530	535	540	
Glu	Trp	Ser	Pro	Glu	Arg	Ser	Thr	Arg	Cys	Phe	Arg	Arg	Arg	Ser	Arg	545	550	555	560
Phe	Leu	Ala	Trp	Gly	Glu	Pro	Ala	Val	Leu	Leu	Leu	Leu	Leu	Leu	Leu	565	570	575	
Ser	Leu	Ala	Leu	Gly	Leu	Val	Leu	Ala	Ala	Leu	Gly	Leu	Phe	Val	His	580	585	590	
His	Arg	Asp	Ser	Pro	Leu	Val	Gln	Ala	Ser	Gly	Gly	Pro	Leu	Ala	Cys	595	600	605	
Phe	Gly	Leu	Val	Cys	Leu	Gly	Leu	Val	Cys	Leu	Ser	Val	Leu	Leu	Phe	610	615	620	
Pro	Gly	Gln	Pro	Ser	Pro	Ala	Arg	Cys	Leu	Ala	Gln	Gln	Pro	Leu	Ser	625	630	635	640
His	Leu	Pro	Leu	Thr	Gly	Cys	Leu	Ser	Thr	Leu	Phe	Leu	Gln	Ala	Ala	645	650	655	
Glu	Ile	Phe	Val	Glu	Ser	Glu	Leu	Pro	Leu	Ser	Trp	Ala	Asp	Arg	Leu	660	665	670	
Ser	Gly	Cys	Leu	Arg	Gly	Pro	Trp	Ala	Trp	Leu	Val	Val	Leu	Leu	Ala	675	680	685	
Met	Leu	Val	Glu	Val	Ala	Leu	Cys	Thr	Trp	Tyr	Leu	Val	Ala	Phe	Pro	690	695	700	
Pro	Glu	Val	Val	Thr	Asp	Trp	His	Met	Leu	Pro	Thr	Glu	Ala	Leu	Val	705	710	715	720
His	Cys	Arg	Thr	Arg	Ser	Trp	Val	Ser	Phe	Gly	Leu	Ala	His	Ala	Thr				

[0007]

725 730 735
Asn Ala Thr Leu Ala Phe Leu Cys Phe Leu Gly Thr Phe Leu Val Arg
740 745 750
Ser Gln Pro Gly Cys Tyr Asn Arg Ala Arg Gly Leu Thr Phe Ala Met
755 760 765
Leu Ala Tyr Phe Ile Thr Trp Val Ser Phe Val Pro Leu Leu Ala Asn
770 775 780
Val Gln Val Val Leu Arg Pro Ala Val Gln Met Gly Ala Leu Leu Leu
785 790 795 800
Cys Val Leu Gly Ile Leu Ala Ala Phe His Leu Pro Arg Cys Tyr Leu
805 810 815
Leu Met Arg Gln Pro Gly Leu Asn Thr Pro Glu Phe Phe Leu Gly Gly
820 825 830
Gly Pro Gly Asp Ala Gln Gly Gln Asn Asp Gly Asn Thr Gly Asn Gln
835 840 845
Gly Lys His Glu
850

<210> 5

<211> 23

<212> DNA

<213> 人工序列

<220>

<223> 人工序列描述: 引物

<220>

<221> 修饰的碱基

<222> (3)

<223> a, t, c, g, 其它或未知

<220>

<221> 修饰的碱基

<222> (9)

<223> a, t, c, g, 其它或未知

<220>

<221> 修饰的碱基

<222> (12)

<223> a, t, c, g, 其它或未知

[0008]

<220>	
<221> 修饰的碱基	
<222> (18)	
<223> a, t, c, g, 其它或未知	
<400> 5	
cgnttyytng cntggggnga rcc	23
<210> 6	
<211> 23	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述: 引物	
<220>	
<221> 修饰的碱基	
<222> (3)	
<223> a, t, c, g, 其它或未知	
<220>	
<221> 修饰的碱基	
<222> (6)	
<223> a, t, c, g, 其它或未知	
<220>	
<221> 修饰的碱基	
<222> (18)	
<223> a, t, c, g, 其它或未知	
<220>	
<221> 修饰的碱基	
<222> (21)	
<223> a, t, c, g, 其它或未知	
<400> 6	
cgngcncgrt trtarcanc ngg	23
<210> 7	
<211> 9	
<212> PRT	
<213> 智人	
<400> 7	
Arg Phe Leu Ala Trp Gly Glu Pro Ala	

[0009]

1 5

<210> 8

<211> 8

<212> PRT

<213> 智人

<400> 8

Pro Gly Cys Tyr Asn Arg Ala Arg

1 5

<210> 9

<211> 552

<212> DNA

<213> Mus sp.

<400> 9

```
gtgctgtcac tcctcctgct gctttgcctg gtgctgggct tagcactggc tgctctgggg 60
ctctctgtcc accactggga cagccctctt gtccaggcct caggcggctc acagttctgc 120
tttggcctga tctgcctagg cctcttctgc ctacagtgtcc ttctgttccc aggacggcca 180
agctctgcca gctgccttgc acaacaacca atggctcacc tccctctcac aggtgcctg 240
agcacactct tcctgcaagc agctgagacc tttgtggagt ctgagctgcc actgagctgg 300
gcaaactggc tatgcagcta cttcgggac tctggcctgc tagtggtact gttggccact 360
tttgtggagg cagcactatg tgcctgggat ttgaccgctt caccagaagt ggtgacagac 420
tggtcagtgc tgcccacaga ggtactggag cactgccacg tgcgttcctg ggtcaacctg 480
ggcttggtgc acatcaccaa tgcaatggta gcttttctct gctttctggg cactttcctg 540
gtacaagacc ag 552
```

<210> 10

<211> 184

<212> PRT

<213> Mus sp.

<400> 10

Val Leu Ser Leu Leu Leu Leu Cys Leu Val Leu Gly Leu Ala Leu

1 5 10 15

Ala Ala Leu Gly Leu Ser Val His His Trp Asp Ser Pro Leu Val Gln

20 25 30

Ala Ser Gly Gly Ser Gln Phe Cys Phe Gly Leu Ile Cys Leu Gly Leu

35 40 45

Phe Cys Leu Ser Val Leu Leu Phe Pro Gly Arg Pro Ser Ser Ala Ser

50 55 60

[0010]

Cys Leu Ala Gln Gln Pro Met Ala His Leu Pro Leu Thr Gly Cys Leu
65 70 75 80

Ser Thr Leu Phe Leu Gln Ala Ala Glu Thr Phe Val Glu Ser Glu Leu
85 90 95

Pro Leu Ser Trp Ala Asn Trp Leu Cys Ser Tyr Leu Arg Asp Ser Gly
100 105 110

Leu Leu Val Val Leu Leu Ala Thr Phe Val Glu Ala Ala Leu Cys Ala
115 120 125

Trp Tyr Leu Thr Ala Ser Pro Glu Val Val Thr Asp Trp Ser Val Leu
130 135 140

Pro Thr Glu Val Leu Glu His Cys His Val Arg Ser Trp Val Asn Leu
145 150 155 160

Gly Leu Val His Ile Thr Asn Ala Met Val Ala Phe Leu Cys Phe Leu
165 170 175

Gly Thr Phe Leu Val Gln Asp Gln
180

<210> 11

<211> 558

<212> DNA

<213> Rattus sp.

<400> 11

```
gtgctgtcac ttctcctgct gctttgcctg gtgctgggcc tgacactggc tgccttgggg 60
ctctttgtcc actactggga cagccctctt gttcaggcct cagggtgggtc actgttctgc 120
tttggcctga tctgcctagg cctcttctgc ctcagtgtcc ttctgttccc aggacgacca 180
cgctctgcca gctgccttgc ccaacaacca atggctcacc tccctctcac aggctgcctg 240
agcacactct tcctgcaagc agccgagatc tttgtggagt ctgagctgcc actgagttgg 300
gcaaactggc tctgcagcta ccttcggggc ccctgggctt ggctggtggt actgctggcc 360
actcttgtgg aggtgctact atgtgectgg tacttgatgg ctttccctcc agaggtggtg 420
acagattggc aggtgctgcc cacggaggta ctggaacact gccgcatgcg ttcctgggtc 480
agcctgggct tgggtgcacat caccaatgca ggggtagctt tcctctgctt tctgggcact 540
ttcctgttac aaagccag 558
```

<210> 12

<211> 186

<212> PRT

<213> Rattus sp.

<400> 12

[0011]

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

<210> 13

<211> 2577

<212> DNA

<213> Rattus sp.

<400> 13

atgccgggtt tggctatctt gggcctcagt ctggctgctt tcctggagct tgggatgggg 60
 tcctctttgt gtctgtcaca gcaattcaag gcacaagggg actatatatt gggtggacta 120
 tttcccttgg gcacaactga ggaggccact ctcaaccaga gaacacagcc caacggcatc 180
 ctatgtacca ggttctcgcc ccttggtttg ttcttgcca tggctatgaa gatggctgta 240
 gaggagatca acaatggatc tgccttgctc cctgggctgc gactgggcta tgacctgttt 300
 gacacatgct cagagccagt ggtcaccatg aagcccagcc tcatgttcat ggccaaggtg 360

[0012]

```

ggaagtcaaa gcattgctgc ctactgcaac tacacacagt accaaccocg tgtgctggct 420
gtcattgggc cccactcatc agagcttgcc ctcattacag gcaagttctt cagcttcttc 480
ctcatgccac aggtcagcta tagtgccagc atggatcggc taagtgaccg ggaacattt 540
ccatccttct tccgcacagt gccagtgac cgggtgcagc tgcaggccgt tgtgacactg 600
ttgcagaatt tcagctggaa ctgggtggct gccttaggta gtgatgatga ctatggccgg 660
gaaggtctga gcatcttttc tggctggcc aactcacgag gtatctgcat tgcacacgag 720
ggcctgggtc cacaacatga cactagtggc caacaattgg gcaaggtggg ggatgtgcta 780
cgccaagtga accaaagcaa agtacagggt gtgggtgctgt ttgcatctgc ccgtgctgtc 840
tactcccttt ttagctacag catccttcat gacctctcac ccaaggtatg ggtggccagt 900
gagtcctggc tgacctctga cctggctcat acacttccca atattgccc tgtgggcaact 960
gttcttgggt ttctgcagcg cgggtgcccta ctgcctgaat tttccatta tgtggagact 1020
cgccttgccc tagctgctga cccaacattc tgtgcctccc tgaaagctga gttggatctg 1080
gaggagcgcg tgatggggcc acgctgttca caatgtgact acatcatgct acagaacctg 1140
tcacttgggc tgatgcagaa cctatcagct gggcagttgc accaccaa atttgcaacc 1200
tatgcagctg tgtacagtgt ggctcaggcc cttcacaaca ccctgcagtg caatgtctca 1260
cattgccaca catcagagcc tgttcaacct tggcagctcc tggagaacat gtacaatatg 1320
agtttccgtg ctgcagactt gacactgcag tttgatgcca aaggaggtgt agacatggaa 1380
tatgacctga agatgtgggt gtggcagagc cctacacctg tactacatac ttaggcacc 1440
ttcaacggca ccctcagct gcagcactcg aaaatgtatt ggccaggcaa ccaggtgcca 1500
gtctccagct gctccggca gtgcaaagat ggccagggtc gcagagtaaa gggctttcat 1560
tcctgctgct atgactgtgt ggactgcaag gcaggagct accggaagca tccagatgac 1620
ttcacctgta ctccatgtg caaggatcag tgggtcccg aaaaaagcac aacctgctta 1680
cctgcaggc ccaagtttct ggcttgggg gagccagctg tgctgtcact tctcctgctg 1740
ctttgcctgg tgctgggct gacactggct gccctggggc tctttgtcca ctactgggac 1800
agccctcttg ttcaggcctc aggtgggtca ctgttctgct ttggcctgat ctgcctaggc 1860
ctctctgccc tcagtgtcct tctgttccca ggacgaccac gctctgccag ctgccttgcc 1920
caacaaccaa tggtcacct ccctctcaca ggctgcctga gcacactctt cctgcaagca 1980
gccgagatct ttgtggagtc tgagctgcca ctgagttggg caaactggct ctgcagctac 2040
cttcggggcc cctgggcttg gctgggtgta ctgctggcca ctcttgtgga ggctgacta 2100
tgtgcctgg acttgatggc tttccctcca gaggtgggtga cagattggca ggtgctgccc 2160
acggaggtac tggaacactg ccgcatgcgt tcctgggtca gcctgggctt ggtgcacatc 2220
accaatgcag tgtagcttt cctctgcttt ctgggcactt tcctggtaca gagccagcct 2280
ggtcgctata accgtgccc tgccctcacc ttcccatgc tagcttattt catcatctgg 2340
gtctcttttg tgccctcct ggctaattgt caggtggcct accagccagc tgtgcagatg 2400
gggtctatct tattctgtgc cctgggcac ctggccacct tccacctgcc caaatgctat 2460
gtacttctgt ggctgccaga gctcaacacc caggagttct tcctgggaag gagcccaag 2520
gaagcatcag atgggaatag tggtagtagt gaggcaactc ggggacacag tgaatga 2577

```

<210> 14

<211> 858

<212> PRT

<213> Rattus sp.

<400> 14

Met Pro Gly Leu Ala Ile Leu Gly Leu Ser Leu Ala Ala Phe Leu Glu

1

5

10

15

Leu Gly Met Gly Ser Ser Leu Cys Leu Ser Gln Gln Phe Lys Ala Gln

[0013]

	20	25	30
Gly Asp Tyr Ile Leu Gly Gly Leu Phe Pro Leu Gly Thr Thr Glu Glu	35	40	45
Ala Thr Leu Asn Gln Arg Thr Gln Pro Asn Gly Ile Leu Cys Thr Arg	50	55	60
Phe Ser Pro Leu Gly Leu Phe Leu Ala Met Ala Met Lys Met Ala Val	65	70	75
Glu Glu Ile Asn Asn Gly Ser Ala Leu Leu Pro Gly Leu Arg Leu Gly	85	90	95
Tyr Asp Leu Phe Asp Thr Cys Ser Glu Pro Val Val Thr Met Lys Pro	100	105	110
Ser Leu Met Phe Met Ala Lys Val Gly Ser Gln Ser Ile Ala Ala Tyr	115	120	125
Cys Asn Tyr Thr Gln Tyr Gln Pro Arg Val Leu Ala Val Ile Gly Pro	130	135	140
His Ser Ser Glu Leu Ala Leu Ile Thr Gly Lys Phe Phe Ser Phe Phe	145	150	155
Leu Met Pro Gln Val Ser Tyr Ser Ala Ser Met Asp Arg Leu Ser Asp	165	170	175
Arg Glu Thr Phe Pro Ser Phe Phe Arg Thr Val Pro Ser Asp Arg Val	180	185	190
Gln Leu Gln Ala Val Val Thr Leu Leu Gln Asn Phe Ser Trp Asn Trp	195	200	205
Val Ala Ala Leu Gly Ser Asp Asp Asp Tyr Gly Arg Glu Gly Leu Ser	210	215	220
Ile Phe Ser Gly Leu Ala Asn Ser Arg Gly Ile Cys Ile Ala His Glu	225	230	235
Gly Leu Val Pro Gln His Asp Thr Ser Gly Gln Gln Leu Gly Lys Val	245	250	255
Val Asp Val Leu Arg Gln Val Asn Gln Ser Lys Val Gln Val Val Val	260	265	270
Leu Phe Ala Ser Ala Arg Ala Val Tyr Ser Leu Phe Ser Tyr Ser Ile	275	280	285

[0014]

Leu His Asp Leu Ser Pro Lys Val Trp Val Ala Ser Glu Ser Trp Leu			
290	295	300	
Thr Ser Asp Leu Val Met Thr Leu Pro Asn Ile Ala Arg Val Gly Thr			
305	310	315	320
Val Leu Gly Phe Leu Gln Arg Gly Ala Leu Leu Pro Glu Phe Ser His			
	325	330	335
Tyr Val Glu Thr Arg Leu Ala Leu Ala Ala Asp Pro Thr Phe Cys Ala			
	340	345	350
Ser Leu Lys Ala Glu Leu Asp Leu Glu Glu Arg Val Met Gly Pro Arg			
	355	360	365
Cys Ser Gln Cys Asp Tyr Ile Met Leu Gln Asn Leu Ser Ser Gly Leu			
	370	375	380
Met Gln Asn Leu Ser Ala Gly Gln Leu His His Gln Ile Phe Ala Thr			
	385	390	395
Tyr Ala Ala Val Tyr Ser Val Ala Gln Ala Leu His Asn Thr Leu Gln			
	405	410	415
Cys Asn Val Ser His Cys His Thr Ser Glu Pro Val Gln Pro Trp Gln			
	420	425	430
Leu Leu Glu Asn Met Tyr Asn Met Ser Phe Arg Ala Arg Asp Leu Thr			
	435	440	445
Leu Gln Phe Asp Ala Lys Gly Ser Val Asp Met Glu Tyr Asp Leu Lys			
	450	455	460
Met Trp Val Trp Gln Ser Pro Thr Pro Val Leu His Thr Val Gly Thr			
	465	470	475
Phe Asn Gly Thr Leu Gln Leu Gln His Ser Lys Met Tyr Trp Pro Gly			
	485	490	495
Asn Gln Val Pro Val Ser Gln Cys Ser Arg Gln Cys Lys Asp Gly Gln			
	500	505	510
Val Arg Arg Val Lys Gly Phe His Ser Cys Cys Tyr Asp Cys Val Asp			
	515	520	525
Cys Lys Ala Gly Ser Tyr Arg Lys His Pro Asp Asp Phe Thr Cys Thr			
	530	535	540

[0015]

Pro Cys Gly Lys Asp Gln Trp Ser Pro Glu Lys Ser Thr Thr Cys Leu			
545	550	555	560
Pro Arg Arg Pro Lys Phe Leu Ala Trp Gly Glu Pro Ala Val Leu Ser			
	565	570	575
Leu Leu Leu Leu Leu Cys Leu Val Leu Gly Leu Thr Leu Ala Ala Leu			
	580	585	590
Gly Leu Phe Val His Tyr Trp Asp Ser Pro Leu Val Gln Ala Ser Gly			
	595	600	605
Gly Ser Leu Phe Cys Phe Gly Leu Ile Cys Leu Gly Leu Phe Cys Leu			
	610	615	620
Ser Val Leu Leu Phe Pro Gly Arg Pro Arg Ser Ala Ser Cys Leu Ala			
	625	630	635
Gln Gln Pro Met Ala His Leu Pro Leu Thr Gly Cys Leu Ser Thr Leu			
	645	650	655
Phe Leu Gln Ala Ala Glu Ile Phe Val Glu Ser Glu Leu Pro Leu Ser			
	660	665	670
Trp Ala Asn Trp Leu Cys Ser Tyr Leu Arg Gly Pro Trp Ala Trp Leu			
	675	680	685
Val Val Leu Leu Ala Thr Leu Val Glu Ala Ala Leu Cys Ala Trp Tyr			
	690	695	700
Leu Met Ala Phe Pro Pro Glu Val Val Thr Asp Trp Gln Val Leu Pro			
	705	710	715
Thr Glu Val Leu Glu His Cys Arg Met Arg Ser Trp Val Ser Leu Gly			
	725	730	735
Leu Val His Ile Thr Asn Ala Val Leu Ala Phe Leu Cys Phe Leu Gly			
	740	745	750
Thr Phe Leu Val Gln Ser Gln Pro Gly Arg Tyr Asn Arg Ala Arg Gly			
	755	760	765
Leu Thr Phe Ala Met Leu Ala Tyr Phe Ile Ile Trp Val Ser Phe Val			
	770	775	780
Pro Leu Leu Ala Asn Val Gln Val Ala Tyr Gln Pro Ala Val Gln Met			
	785	790	795
Gly Ala Ile Leu Phe Cys Ala Leu Gly Ile Leu Ala Thr Phe His Leu			

[0016]

805	810	815
Pro Lys Cys Tyr Val Leu Leu Trp Leu Pro Glu Leu Asn Thr Gln Glu		
820	825	830
Phe Phe Leu Gly Arg Ser Pro Lys Glu Ala Ser Asp Gly Asn Ser Gly		
835	840	845
Ser Ser Glu Ala Thr Arg Gly His Ser Glu		
850	855	

<210> 15

<211> 8194

<212> DNA

<213> 智人

<220>

<221> 修饰的碱基

<222> (1251).. (1300)

<223> a, t, c, g, 其它或未知

<220>

<221> 修饰的碱基

<222> (1951).. (2000)

<223> a, t, c, g, 其它或未知

<400> 15

```

gagaatctcg cgagatcccg tcggtccgcc ccgctgccct cccagctgcc gaaaagaggg 60
gcctccgagc cgccggcgcc ctctgccggc aacctcogga agcacactag gaggttccag 120
ccgatctggt cgaggggctc cacggaggac tccatttacg ttacgcaaat tccctacccc 180
agccggccgg agagagaaag ccagaaacct cgcgaccagc catggggccac ctctccggaa 240
aaacaccggg atattttttt tctcctgcag aaaaagcttt aggattggca gtttaaacia 300
aacatgtcta ttgtcatacc ttcggtttgc atgcatttgt ttcgaagtga gcaaccctgg 360
gtaacaaggc gaaagtatat gacaatttgc tcagaatctt aatgtcagaa aactggagac 420
tggggcaggg ggggtgtcgac tcaaagctgt gtctcattta gtaaaactgag gccaggttaa 480
aaagttctga aacctcgcaa caccgggaga aattgtgttc cagcctccca cctcgcccca 540
aatgcccaga gctccttttc taagccaggc gaagtcacag agcgtggaca gaaccacaa 600
ccgtccagag gaagggtcac tgggtgccac ctggtttgca tctgtgcctt cgtcctgccc 660
agttcctgag tgggaccgca ggcccggat gtcaaggcaa acagtcctgc ttcagccact 720
gggctccagt cccaccctt ttgggggcct gaagttagga agcatccggc agctgccttc 780
tatttaagca actggcctcc ttagaggcca ctcttggtgc atgccaggcg cgggcatctg 840
gccagcatgc tgetctgcac ggctcgctg gtcggcctgc agcttctcat ttctgtctgc 900
tgggccttig cctgcatag cacggagtct tctcctgact tcaccctccc cggagattac 960
ctcctggcag gcctgttccc tctccattct ggctgtctgc aggtgaggca cagaccggag 1020
gtgaccctgt gtgacagggt agtgaggggc cagcagagcc acacttagtg ggacccttg 1080
ctatagggcc cctctggctg ccatcctcca aacaggacct tgcctctgcc ttgcccctt 1140
gaactgtccc caggccttgt tcatcaatcc acttgccacc taagtgtctg ctagaccttc 1200
ctagacactt cggccagttt ccaattattt cacccttgct gttagaatgt nnnnnnnnnn 1260

```

[0017]

```

nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn aattccttaa actaaatttc 1320
tcactttctc tctctctctg gaaaacactg actaatgtag caggtttctc tgctccagga 1380
cttcaggacc ttttcgatgc taataagttt ctccatcagg gccagcttgt tctcctact 1440
gagcttgaga gcccttgttg aagtgttgtt ttgggggact ggaccgatga cctcaaaggt 1500
tccccttgct cccaagcctc agagtctagg aggccagagg gtctcagcag gcctttgtcc 1560
ttctcagctg tctcttactg gctttctcca caggtcttgt agcttcaatg agcatggcta 1620
ccacctcttc caggetatgc ggcttggggt tgaggagata aacaactcca cgccctgct 1680
gcccacatc accctggggt accagctgta tgatgtgtgt tctgactctg ccaatgtgta 1740
tgccacgctg agagtgtctt ccctgccagg gcaacaccac atagagctcc aaggagacct 1800
tctccactat tcccctacgg tgctggcagt gattgggcct gacagcacca accgtgctgc 1860
caccacagcc gccctgctga gccctttcct ggtgcccatt gtaagctgga gcctcagacc 1920
tttgccatc tccctcagg caagtctggg nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1980
nnnnnnnnnn nnnnnnnnnn gccaccatgc cgggctaatt tttttgtatt tttagtagag 2040
acggggtttc accgtgttag ccagcttgtt cgcaactcc taacctcgtg atccaccac 2100
ctcggcctcc caatgtgctg ggattacagg tgtgagccac tgcaccggc cataatgtat 2160
taatataata aaataattat acaactcacc ataatgtaga atcagtggga gccctgagct 2220
tgttttctca caactagatg gtccatctg ggggtgatgg gagacagtga cagatcatca 2280
gacattagat tctcataagt agcgtgcaac ccagatccct cgcatgtgca gttcacagta 2340
gggttcaagc tctacaaga atctgatgct gctgctgac tgacaggagg ggagcagctg 2400
taaatacaga tgaagcttcg ctactcacc agctgctcac ctctcctgt gaggcccggt 2460
tcctaacagg ccactgacct aacttctgcc ctgacctaca catgcttctc ttcttcttg 2520
caaactgcct ccagtgaag tccctgaagg tcccaaaca cacgggacta tttcactcct 2580
atgcaggttt tgtctccttt gcttgaatg catcccctca ccccttgtcc ccaggcagat 2640
tcccaccct cccccagaac ctgcccaggt ggagccttcg caggtgattt gtcagtttca 2700
caggtgagg ggtgctctcc tggctcccc ggctccctgt atcccacac ccagcacagg 2760
gccaggcact gggggggcct tcagtggaga ctgaaatggc tgaacgggac ctcccataga 2820
ttagctatgc ggccagcagc gagacgctca gcgtgaagcg gcagtatccc tcttctctgc 2880
gcaccatccc caatgacaag taccaggtgg agaccatggt gctgctgctg cagaagttcg 2940
ggtggacctg gatctctctg gttggcagca gtgacgacta tgggcagcta ggggtgcagg 3000
cactggagaa ccaggccact ggtcagggga tctgcattgc tttcaaggac atcatgccct 3060
tctctgcca ggtggcgat gagaggatgc agtgccctcat gcgccacctg gccaggccg 3120
gggccaccgt cgtggttgtt tttccagcc ggcaagtggc cagggtgttt ttcgagtcg 3180
tggtgctgac caactgact ggcaaggtgt gggctgcctc agaagcctgg gccctctcca 3240
ggcacatcac tggggtgcc gggatccagc gcattgggat ggtgctgggc gtggccatcc 3300
agaagagggc tgtccctggc ctgaaggcgt ttgaagaagc ctatgcccg gcagacaaga 3360
aggcccctag gccttgccac aagggtcct ggtgcagcag caatcagctc tgcagagaat 3420
gccaagcttt catggcacac acgatgcca agctcaaagc cttctccatg agttctgcct 3480
acaacgcata cgggctgtg tatcggtgg cccatggcct ccaccagctc ctgggctgtg 3540
cctctggagc ttgttccagg ggccgagtct acccctggca ggtaagagag cccacccag 3600
cacctcctgt cagggagaac agccaatcct gagatgagca gagtgggcac tctccggtca 3660
ctctaaatgc caaggggat aaatgccact aactgaggt tttttgttt gttttgttt 3720
gttttttgag acagtctggc tctgtcacc agctgcagt gtagtgatgc gatctcggt 3780
ctctgcaact tccacctct gggttcaagt gattctcttg cctcggcctc ctgagtagct 3840
gggattacag gcacccacca ccatgcctgg ataattttt tttcttttt ttttttttg 3900
agatagagtc tcgctctgtt gccaggtg gaatgcagt gtgcgatctt ggctcactgt 3960
gagctccgcc tcccaggttc actccattcc cctgcctcag cctcccaagt aggtgggact 4020
acgggcgccc gccaccacgc ccagctaatt tttttgtat tttgagtaga gacggggtt 4080
caccatgtta gccaggatgg tctcaatctc ctgacctgt catccgcca cctcgtcctc 4140
ccaaagtgtt gggattacag gcgtgagcca ccgaccccg cctaattttt gtatttttag 4200

```

[0018]

```

tagagatggg gtttcacccat gttggccagg ctggtctcga actcctggca tcaagtgatc 4260
ctctgcttc ggctcccaa agtctggga ttacaggcat tagctctctt ctcttagaca 4320
gatctttctc tctgacctt gccttctctc acccactgtg tcttgaagt gtcaagtgat 4380
aagatccagg gctaaaactg tctgtaaagg agtgtttgtt agaggcctcc tctcaggagg 4440
ttggtgggga agattgagg gcttcctaag aaggaaggga cgagacctc ctgatgggct 4500
gaaaccacca ggacggaaac ccaggaaggc ccaggccct tgcttctggg accatgtggg 4560
tctgtgctgt ctgtggtggc ttcatgata gcgtttcttt cagcttttgg agcagatcca 4620
caaggtgcat ttcttctac acaaggacac tgtggcggtt aatgacaaca gagatccct 4680
cagtagctat aacataattg cctgggactg gaatggaccc aagtggacct tcacggctct 4740
cggttctcc acatggtctc cagttcagct aaacataaat gagacaaaaa tccagtggca 4800
cggaaaggac aaccaggtaa tggggatgtg gctactcacc atgtaactgg ctatgggca 4860
acctagagcc tgggggtgat gctgacacag tgtacaggga gcaggagggg ggccccaggg 4920
gtccagctgc caccactcta cccatcctgg ccagggaagc agggaagaca ctccgtaggc 4980
gagtgtgcag atgccctggg gcggaagttc acacgaccag gggccctgcc ctgggagtga 5040
gccctgaggg cagatgcaca gagattctgt tttctgttcc acatgtgagc tgcctttga 5100
cttgggcccc tacgtgtggc cctctggct tcttacaggt gcctaagtct gtgtgttcca 5160
gcgactgtct tgaagggcac cagcgagtgg ttacgggtt ccatcactgc tgctttgagt 5220
gtgtgccctg tggggctggg accttcctca acaagagtgg tgagtgggca atggagcagg 5280
cgagctaccc agcactcccg ggggctgcac ggtggaggga gggcctcct tgggccccat 5340
gtgccctgcc ccagaaccaa ggccagtc aagggtgcc agtagcttc aggttggagg 5400
acacctgcta ccagacagaa ttctgatcaa gagaatcagc cactgggtgc ggtggctcat 5460
gcctgtaatc ccagcacttt gggaggctga ggcgggtgga tcacttgagg tcgggagtgc 5520
gagaccagcc tggccaacat ggtgaaaccc catctctacc aaaaatataa aaaattagct 5580
gggtgtgtg gcgcgtgcct gtaatcccag ctactcgga ggctgaggca ggagaatcac 5640
ttgaaccag gaggcggagg ttgcagtga ccaagatgca ttccagcctg gaccacaaag 5700
cgagaattcg tcccccaaa aaaagaaagg aggcgggagc cgggtggtca cacctgtaat 5760
cccagcactt tgggaggccg aggtgggtgg atcacctgag gtcaggagt cagagaccagc 5820
ctgaccaaca tggtgaaacc ccatctctac taaaaataca aaaaaagta gccggcggtt 5880
gtggcgtgtg cctgtaattc cagctactcg gagggtgag gcaggagaat tgcttgaacc 5940
cgggaggcgg aggttgagc gagccaagat tgcaccattg cactccagcc tgggcgacaa 6000
gagaaaaact ctgtctcaaa aaaaaagaaa gaaagaaaga attagccaac tgaagcctt 6060
agactgaggt gtgtcctctg ttagagagct gtcatacaaa ctctacaaa agcagtcgta 6120
tcctgaattc aaactcttc tctaaatgaa tatagctatt gttcccttg tgccctctg 6180
tcctactgtc cttctgttg cccatgccaa agacagctag ctcttgaaac agcttggcct 6240
gaatacagat actagcgtgt ctgcagcaga gaaaaaaca gcattccca tccagaaatg 6300
caaggtcaag aacagagagc aaattaggt gctaaggact caggctctta gttggtgtcc 6360
aggggccaca ttctttctt tcaccatctc ttagggaca ggaatactc cttctgtcc 6420
tcagagggtc aggactcaga gaaaccacag agcagcagct caggaaagt gtcatggaa 6480
atgctggcaa gagagagggt ttacatgcc ctcccttggg agcaggctgc tcccatcaga 6540
tcgtaacctc tctggtatgt ggcagagct accaggttaa ggtcctcct agggtttga 6600
aaacctcat gggatcatga gccatacaga accgacctgt gtgtctccag agtctgtaat 6660
taacacaggc attttagga aatgcgtggc ctcaggcccc actccggct accccatcc 6720
cactatgcct agtatagtct agctgccctg gtacaattct ccagtatct tgccaggccc 6780
tatttcctat tcctactctg ctcatctggc tctcaggaa cttcttggc ttcccttca 6840
gacctctaca gatccagcc ttgtgggaaa gaagagtggg cacctgaggg aagccagacc 6900
tgcttcccgc gactgtggt gtttttggt ttgcgtgagc acacctctg ggtgctgtg 6960
gcagctaaca cgctgctgt gctgctgtg cttgggactg ctggcctgtt tgctggcac 7020
ctagacaccc ctgtggtgag gtcagcaggg ggcgcctgt gctttctat gctgggctcc 7080
ctggcagcag gtagtggcag cctctatggc ttctttggg aaccacaag gcctgcgtgc 7140

```

[0019]

```

ttgtacgcc aggccctctt tggccttggg ttcaccatct tctgtctctg cctgacagtt 7200
cgctcattcc aactaatcat catcttcaag tttccacca aggtacctac attctaccac 7260
gcctgggtcc aaaaccacgg tgcgtggctg tttgtgatga tcagctcagc ggcccagctg 7320
cttatctgtc taacttggct ggtgggtgg accccactgc ctgctaggga ataccagcgc 7380
ttccccatc tggatgatct tgagtgcaca gagaccaact ccctgggctt catactggcc 7440
ttcctctaca atggcctcct ctccatcagt gcctttgcct gcagctacct gggttaaggac 7500
ttgccagaga actacaacga ggccaaatgt gtcacattca gcctgctctt caacttcgtg 7560
tcctggatcg ccttcttcac cacggccagc gtctacgacg gcaagtacct gcctgcggcc 7620
aacatgatgg ctgggctgag cagcctgagc agcggtctcg gtgggtattt tctgcctaag 7680
tgctacgtga tctctgccc cccagacctc aacagcacag agcacttcca ggccctcatt 7740
caggactaca cgaggcgtg cggtccacc tgaccagtgg gtcagcaggc acggctggca 7800
gccttctctg cctgagggt cgaaggtcga gcaggccggg ggtgtccggg aggtcttttg 7860
gcatcgcggt ctggggttg gacgtgtaag cgctgggag agcctagacc aggctccggg 7920
ctgccaataa agaagtgaag tgcgtatctg gtctcctgtc gtgggagagt gtgaggtgta 7980
acggattcaa gtctgaacc agagcctgga aaaggctgac cgccagatt gacgttgcta 8040
ggcaactccg gaggcgggcc cagcgccaaa agaacagggc gaggcgtcgt ccccgcatcc 8100
cattggcgt tctctgcggg gccccgccct cgggggccgg agctagaagc tctacgctc 8160
cgaggcgcac ctctggcct gcacgctttg acgt 8194

```

<210> 16

<211> 2526

<212> DNA

<213> 智人

<400> 16

```

atgtgtctct gcacggctcg cctggctggc ctgcagcttc tcatttctg ctgctgggcc 60
tttgctgcc atagcacgga gtcttctct gacttcaccc tccccggaga ttacctctg 120
gcaggcctgt tcctctcca ttctggctgt ctgcaggta ggacagacc cgaggtgacc 180
ctgtgtgaca ggtctttag ctcaatgag catggctacc acctctcca ggctatgcg 240
cttggggttg aggagataa caactccag gcctgtgct ccaacatcac cctgggttac 300
cagctgtatg atgtgtgtc tgactctgcc aatgtgtat ccacgtgag agtgcctcc 360
ctgccagggc aacaccacat agagctcaa ggagacctc tccactatc ccctacgtg 420
ctggcagtga ttggcctga cagcaccaac cgtgtgcca ccacagccg cctgctgagc 480
ccttctctgg tgccatgat tagctatgcg gccagcagc agacgtcag cgtgaagcgg 540
cagtatccct ctctctgag caccatccc aatgacaagt accagggtga gaccatggtg 600
ctgctgtgc agaagttcgg gtggacctg atctctctg ttggcagcag tgacgactat 660
gggcagctag ggtgcaggc actggagaac caggccactg gtcaggggat ctgcattgct 720
ttcaaggaca tcatgccct ctctgccag gtggcgatg agaggatgca gtgcctcatg 780
cgccacctgg cccaggccgg ggccaccgtc gtggttgtt tttccagccg gcagttggcc 840
agggtgttt tgcagtcct ggtgctgacc aacctgactg gcaagggtg ggtcgccctc 900
gaagcctggg ccctctccag gcacatcact ggggtgccg ggatccagc cattgggatg 960
gtgctgggcg tggccatcca gaagagggtc gtccctggc tgaaggcgt tgaagaagcc 1020
tatgcccggg cagacaagaa ggcccctagg ccttgccaca agggctcctg gtgcagcagc 1080
aatcagctct gcagagaatg ccaagcttct atggcacaca cgatgcccac gctcaaagcc 1140
ttctccatga gttctgccta caacgcatac cgggctgtgt atcggtggc ccatggcctc 1200
caccagctcc tgggtgtgc ctctggagct gttccaggg gccgagtcta cccctggcag 1260
cttttggagc agatccacaa ggtgcatttc cttctacaca aggacactgt ggcgtttaat 1320
gacaacagag atccccctag tagctataac ataattgcct gggactggaa tggacccaag 1380

```

[0020]

```

tggacattca cggtcctcgg ttctccaca tggctccag ttcagctaaa cataaatgag 1440
acaaaaatcc agtggcacgg aaaggacaac cagggtccta agtctgtgtg ttccagcgac 1500
tgtcttgaag ggcaccagcg agtggttacg ggtttccatc actgctgctt tgagtgtgtg 1560
ccctgtgggg ctgggacctt cctcaacaag agtgacctct acagatgcca gccttgtggg 1620
aaagaagagt gggcacctga ggaagccag acctgcttcc cgcgcactgt ggtgtttttg 1680
gctttgcgtg agcacacctc ttgggtgctg ctggcagcta acacgtgct gctgctgctg 1740
ctgcttggga ctgctggcct gttgcctgg cacctagaca cccctgtggt gaggtcagca 1800
gggggcccgc tgtcttttct tatgctgggc tccctggcag caggtagtgg cagcctctat 1860
ggcttctttg gggaacccac aaggcctgcg tgcttgctac gccaggccct cttgcccctt 1920
ggtttcacca tcttcctgtc ctgcctgaca gttcgtcat tccaactaat catcatcttc 1980
aagttttcca ccaaggtacc tacattctac cagcctggg tccaaaacca cgggtgctggc 2040
ctgtttgtga tgatcagctc agcggcccag ctgcttatct gtctaacttg gctggtggtg 2100
tggaccccac tgctgctag ggaataccag cgcttcccc atctggtgat gcttgagtgc 2160
acagagacca actccctggg cttcatactg gccttcctct acaatggcct cctctccatc 2220
agtgcctttg cctgcagcta cctgggtaag gacttgccag agaactaaa cgaggccaaa 2280
tgtgtcacct tcagcctgct cttcaacttc gtgtcctgga tcgccttctt caccacggcc 2340
agcgtctacg acggaagta cctgcctgcg gccaacatga tggctgggct gagcagcctg 2400
agcagcggct tcggtgggta tttctgcct aagtgtctag tgatcctctg ccgccagac 2460
ctcaacagca cagagcactt ccaggcctcc attcaggact acacgaggcg ctgcggctcc 2520
acctga 2526

```

<210> 17
 <211> 841
 <212> PRT
 <213> 智人

<400> 17
 Met Leu Leu Cys Thr Ala Arg Leu Val Gly Leu Gln Leu Leu Ile Ser
 1 5 10 15
 Cys Cys Trp Ala Phe Ala Cys His Ser Thr Glu Ser Ser Pro Asp Phe
 20 25 30
 Thr Leu Pro Gly Asp Tyr Leu Leu Ala Gly Leu Phe Pro Leu His Ser
 35 40 45
 Gly Cys Leu Gln Val Arg His Arg Pro Glu Val Thr Leu Cys Asp Arg
 50 55 60
 Ser Cys Ser Phe Asn Glu His Gly Tyr His Leu Phe Gln Ala Met Arg
 65 70 75 80
 Leu Gly Val Glu Glu Ile Asn Asn Ser Thr Ala Leu Leu Pro Asn Ile
 85 90 95
 Thr Leu Gly Tyr Gln Leu Tyr Asp Val Cys Ser Asp Ser Ala Asn Val
 100 105 110

[0021]

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

[0022]

370	375	380
Ser Ala Tyr Asn Ala Tyr Arg Ala Val Tyr Ala Val Ala His Gly Leu 385	390	395 400
His Gln Leu Leu Gly Cys Ala Ser Gly Ala Cys Ser Arg Gly Arg Val 405	410	415
Tyr Pro Trp Gln Leu Leu Glu Gln Ile His Lys Val His Phe Leu Leu 420	425	430
His Lys Asp Thr Val Ala Phe Asn Asp Asn Arg Asp Pro Leu Ser Ser 435	440	445
Tyr Asn Ile Ile Ala Trp Asp Trp Asn Gly Pro Lys Trp Thr Phe Thr 450	455	460
Val Leu Gly Ser Ser Thr Trp Ser Pro Val Gln Leu Asn Ile Asn Glu 465	470	475 480
Thr Lys Ile Gln Trp His Gly Lys Asp Asn Gln Val Pro Lys Ser Val 485	490	495
Cys Ser Ser Asp Cys Leu Glu Gly His Gln Arg Val Val Thr Gly Phe 500	505	510
His His Cys Cys Phe Glu Cys Val Pro Cys Gly Ala Gly Thr Phe Leu 515	520	525
Asn Lys Ser Asp Leu Tyr Arg Cys Gln Pro Cys Gly Lys Glu Glu Trp 530	535	540
Ala Pro Glu Gly Ser Gln Thr Cys Phe Pro Arg Thr Val Val Phe Leu 545	550	555 560
Ala Leu Arg Glu His Thr Ser Trp Val Leu Leu Ala Ala Asn Thr Leu 565	570	575
Leu Leu Leu Leu Leu Leu Gly Thr Ala Gly Leu Phe Ala Trp His Leu 580	585	590
Asp Thr Pro Val Val Arg Ser Ala Gly Gly Arg Leu Cys Phe Leu Met 595	600	605
Leu Gly Ser Leu Ala Ala Gly Ser Gly Ser Leu Tyr Gly Phe Phe Gly 610	615	620
Glu Pro Thr Arg Pro Ala Cys Leu Leu Arg Gln Ala Leu Phe Ala Leu 625	630	635 640

[0023]

Gly Phe Thr Ile Phe Leu Ser Cys Leu Thr Val Arg Ser Phe Gln Leu
645 650 655

Ile Ile Ile Phe Lys Phe Ser Thr Lys Val Pro Thr Phe Tyr His Ala
660 665 670

Trp Val Gln Asn His Gly Ala Gly Leu Phe Val Met Ile Ser Ser Ala
675 680 685

Ala Gln Leu Leu Ile Cys Leu Thr Trp Leu Val Val Trp Thr Pro Leu
690 695 700

Pro Ala Arg Glu Tyr Gln Arg Phe Pro His Leu Val Met Leu Glu Cys
705 710 715 720

Thr Glu Thr Asn Ser Leu Gly Phe Ile Leu Ala Phe Leu Tyr Asn Gly
725 730 735

Leu Leu Ser Ile Ser Ala Phe Ala Cys Ser Tyr Leu Gly Lys Asp Leu
740 745 750

Pro Glu Asn Tyr Asn Glu Ala Lys Cys Val Thr Phe Ser Leu Leu Phe
755 760 765

Asn Phe Val Ser Trp Ile Ala Phe Phe Thr Thr Ala Ser Val Tyr Asp
770 775 780

Gly Lys Tyr Leu Pro Ala Ala Asn Met Met Ala Gly Leu Ser Ser Leu
785 790 795 800

Ser Ser Gly Phe Gly Gly Tyr Phe Leu Pro Lys Cys Tyr Val Ile Leu
805 810 815

Cys Arg Pro Asp Leu Asn Ser Thr Glu His Phe Gln Ala Ser Ile Gln
820 825 830

Asp Tyr Thr Arg Arg Cys Gly Ser Thr
835 840

<210> 18

<211> 14

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述：一致序列

[0024]

<220>
<221> MOD_RES
<222> (1)
<223> Thr或Arg

<220>
<221> MOD_RES
<222> (3)
<223> Phe或Leu

<220>
<221> MOD_RES
<222> (4)
<223> Arg, Gln或Pro

<220>
<221> MOD_RES
<222> (6)
<223> Arg或Thr

<220>
<221> MOD_RES
<222> (7)
<223> Ser, Pro或Val

<220>
<221> MOD_RES
<222> (8)
<223> Val, Glu, Arg, Lys或Thr

<220>
<221> MOD_RES
<222> (11)
<223> Ala或Glu

<220>
<221> MOD_RES
<222> (12)
<223> Trp或Leu

<220>
<221> MOD_RES
<222> (13)
<223> Arg, His或Gly

<400> 18
Xaa Cys Xaa Xaa Arg Xaa Xaa Xaa Phe Leu Xaa Xaa Xaa Glu
1 5 10

[0025]

<210> 19
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 人工序列描述: 一致序列

<220>
<221> MOD_RES
<222> (1)
<223> Leu或Gln

<220>
<221> MOD_RES
<222> (3)
<223> Glu, Gly或Thr

<220>
<221> MOD_RES
<222> (4)
<223> Asn, Arg或Cys

<220>
<221> MOD_RES
<222> (7)
<223> Arg或Glu

<220>
<221> MOD_RES
<222> (9)
<223> Arg或Lys

<220>
<221> MOD_RES
<222> (10)
<223> Cys, Gly或Phe

<220>
<221> MOD_RES
<222> (11)
<223> Val, Leu或Ile

<220>
<221> MOD_RES
<222> (13)
<223> Phe或Leu

[0026]

<220>
 <221> MOD_RES
 <222> (14)
 <223> Ala或Ser

<220>
 <221> MOD_RES
 <222> (15)
 <223> Met或Leu

<400> 19
 Xaa Pro Xaa Xaa Tyr Asn Xaa Ala Xaa Xaa Xaa Thr Xaa Xaa Xaa
 1 5 10 15

<210> 20
 <211> 3563
 <212> DNA
 <213> 智人

<400> 20
 agcctggcag tggcctcagg cagagtctga cgcgcacaaa ctttcaggcc caggaagcga 60
 ggacaccact ggggccccag ggtgtggcaa gtgaggatgg caagggtttt gctaaacaaa 120
 tcctctgccc gctccccgcc ccgggctcac tccatgtgag gccccagtcg gggcagccac 180
 ctgccgtgcc tgttggaagt tgccctctgc atgctgggcc ctgctgtcct gggcctcagc 240
 ctctgggctc tcctgcaccc tgggacgggg gccccattgt gcctgtcaca gcaacttagg 300
 atgaaggggg actacgtgct gggggggctg ttccccctgg gcgaggccga ggaggctggc 360
 ctccgcagcc ggacacggcc cagcagccct gtgtgcacca ggtacagagg tgggacggcc 420
 tgggtcgggg tcagggtgac caggtctggg gtgctcctga gctggggccg aggtggccat 480
 ctgcggttct gtgtggcccc aggttctcct caaacggcct gctctgggca ctggccatga 540
 aaatggccgt ggaggagatc aacaacaagt cggatctgct gcccgggctg cgcctgggct 600
 acgacctctt tgatactgc tcggagcctg tgggtggccat gaagcccagc ctcatgttcc 660
 tggccaaggc aggcagccgc gacatcgccg cctactgcaa ctacacgcag taccagcccc 720
 gtgtgctggc tgtcatggg cccactcgt cagagctcgc catggtcacc ggcaagttct 780
 tcagcttctt cctcatgccc cagtggggcg cccccacca tcaccacccc ccaaccaacc 840
 cctgccccgt gggagcccc tgtgtcagga gaatgctaca tgcacccac ccagccctgc 900
 cctgggagcc ctgtgtcaga agatgctctt ggccttgag gtcagctacg gtgctagcat 960
 ggagctgctg agcggccggg agacctccc ctctctcttc cgcaccgtgc ccagcgaccg 1020
 tgtgcagctg acggccggcg cggagctgct gcaggagtgc ggcctggaact ggggtgcccg 1080
 cctgggcagc gacgacgagt acggccggca gggcctgagc atcttctcgg ccctggccgc 1140
 ggcacgcggc atctgcatcg cgcacagagg cctggtgccg ctgccccgtg ccgatgactc 1200
 gcggctgggg aaggtgcagg acgtcctgca ccaggtgaac cagagcagcg tgcaggtggt 1260
 gctgtgtgtc gcctccgtgc acgcccacca cgcctcttcc aactacagca tcagcagcag 1320
 gctctcgccc aaggtgtggg tggccagcga ggcctggctg acctctgacc tggatcatgg 1380
 gctgcccggc atggcccaga tgggacgggt gcttggcttc ctccagaggg gtgcccagct 1440
 gcacgagttc cccagtagc tgaagacgca cctggccctg gccaccgacc cggccttctg 1500
 ctctgccctg ggcgagaggg agcagggtct ggaggaggac gtggtgggcc agcgtgccc 1560
 gcaggtgac tgcatacgc tgcagaacgt gagcgcaggg ctaaatcacc accagacgtt 1620
 ctctgtctac gcagctgtgt atagctggc ccaggccctg cacaacactc ttcagtgcga 1680

[0027]

```

cgctcaggc tgcgcgcgc agggcccggt gaagccctgg caggtgagcc cgggagatgg 1740
gggtgtgctg tctctgcat gtgccaggc caccaggcac ggccaccacg cctgagctgg 1800
aggtggctgg cggctcagcc ccgtcccccg cccgcagctc ctggagaaca tgtacaacct 1860
gaccttcac gtgggcgggc tgcgctgctg gtctgacagc agcggaaacg tggacatgga 1920
gtacgacctg aagctgtggg tgtggcaggc ctgagtgccc aggtccacg acgtgggcag 1980
gttcaacggc agcctcagga cagagcgct gaagatccgc tggcacacgt ctgacaacca 2040
ggtgaggtga ggggtgggtgt gccaggcgtg cccgtggtag ccccgcggc agggcgagc 2100
ctgggggtgg gggccgttcc agtctccgt gggcatgccc agccgagcag agccagaccc 2160
caggcctgtg cgcagaagcc cgtgtcccg tgctcgcggc agtgccagga gggccagggt 2220
cgccgggtca aggggttcca ctctgctgc tacgactgtg tggactgca ggcgggcagc 2280
taccggcaaa acccaggtga gccgccttcc cggcaggcgg ggggtgggaac gcagcagggg 2340
agggtcctgc caagtcctga ctctgagacc agagcccaca ggtacaaga cgaacacca 2400
gcgcccttct cctctctcac agacgacatc gcctgcacct ttgtggcca ggatgagtgg 2460
tccccggagc gaagcacacg ctgcttccgc cgcaggtctc ggttcttggc atggggcgag 2520
ccggctgtgc tgcgtgctgt cctgctgtg agcctggcgc tgggccttgt gctggctgct 2580
ttggggctgt tcgttcacca tcgggacagc cacttggtt aggcctcggg ggggccccctg 2640
gcctgctttg gcctggtgtg cctgggcctg gtctgcctca gcgtcctct gttccctggc 2700
cagcccagcc ctgcccgatg cctggcccag cagcccttgt cccacctccc gtcacgggc 2760
tgcttgagca cactcttct gcaggcggcc gagatcttcg tggagtcaga actgcctctg 2820
agctgggcag accggtgag tggctgcctg cgggggccc cggcctggct ggtggtgtg 2880
ctggccatgc tgggtggagt cgcactgtgc acctggtacc tgggtggcctt cccgcccggag 2940
gtggtgacgg actggcacat gctgccacg gagcgctgg tgcactgca cacacgtcc 3000
tgggtcagct tcggcctagc gcagccacc aatgccacg tggccttct ctgcttctg 3060
ggcactttcc tgggtcggag ccagccgggc tgctacaacc gtgccctgg cctcacctt 3120
gccatgctgg cctacttcat cacctgggtc tcctttgtg ccctcctggc caatgtgcag 3180
gtggtcctca gggccgccgt gcagatgggc gccctcctg tctgtgtct gggcatcctg 3240
gctgccttcc acctgccag gtgttacct ctcatcggc agccagggt caacaccccc 3300
gagttcttcc tgggaggggg ccttgggat gcccaaggc agaatgacg gaacacagga 3360
aatcagggga aacatgagt acccaacct gtgatctcag ccccggtgaa ccagactta 3420
gtgcgatcc ccccaagcc agcaatgacc cgtgtctgc tacagagacc ctccgctct 3480
aggttctgac ccaggttgt ctctgaccc tgacccaca gtgagcccta ggcctggagc 3540
acgtggacac ccctgtgacc atc 3563

```

<210> 21

<211> 839

<212> PRT

<213> 智人

<400> 21

```

Met Gly Pro Arg Ala Lys Thr Ile Cys Ser Leu Phe Phe Leu Leu Trp
  1             5             10             15

```

```

Val Leu Ala Glu Pro Ala Glu Asn Ser Asp Phe Tyr Leu Pro Gly Asp
          20             25             30

```

```

Tyr Leu Leu Gly Gly Leu Phe Ser Leu His Ala Asn Met Lys Gly Ile
      35             40             45

```

[0028]

Val	His	Leu	Asn	Phe	Leu	Gln	Val	Pro	Met	Cys	Lys	Glu	Tyr	Glu	Val
50						55					60				
Lys	Val	Ile	Gly	Tyr	Asn	Leu	Met	Gln	Ala	Met	Arg	Phe	Ala	Val	Glu
65					70					75					80
Glu	Ile	Asn	Asn	Asp	Ser	Ser	Leu	Leu	Pro	Gly	Val	Leu	Leu	Gly	Tyr
				85					90					95	
Glu	Ile	Val	Asp	Val	Cys	Tyr	Ile	Ser	Asn	Asn	Val	Gln	Pro	Val	Leu
			100						105					110	
Tyr	Phe	Leu	Ala	His	Glu	Asp	Asn	Leu	Leu	Pro	Ile	Gln	Glu	Asp	Tyr
		115						120					125		
Ser	Asn	Tyr	Ile	Ser	Arg	Val	Val	Ala	Val	Ile	Gly	Pro	Asp	Asn	Ser
		130				135						140			
Glu	Ser	Val	Met	Thr	Val	Ala	Asn	Phe	Leu	Ser	Leu	Phe	Leu	Leu	Pro
145					150					155					160
Gln	Ile	Thr	Tyr	Ser	Ala	Ile	Ser	Asp	Glu	Leu	Arg	Asp	Lys	Val	Arg
				165					170					175	
Phe	Pro	Ala	Leu	Leu	Arg	Thr	Thr	Pro	Ser	Ala	Asp	His	His	Val	Glu
			180						185					190	
Ala	Met	Val	Gln	Leu	Met	Leu	His	Phe	Arg	Trp	Asn	Trp	Ile	Ile	Val
		195					200					205			
Leu	Val	Ser	Ser	Asp	Thr	Tyr	Gly	Arg	Asp	Asn	Gly	Gln	Leu	Leu	Gly
		210				215					220				
Glu	Arg	Val	Ala	Arg	Arg	Asp	Ile	Cys	Ile	Ala	Phe	Gln	Glu	Thr	Leu
225					230					235					240
Pro	Thr	Leu	Gln	Pro	Asn	Gln	Asn	Met	Thr	Ser	Glu	Glu	Arg	Gln	Arg
			245						250					255	
Leu	Val	Thr	Ile	Val	Asp	Lys	Leu	Gln	Gln	Ser	Thr	Ala	Arg	Val	Val
			260					265					270		
Val	Val	Phe	Ser	Pro	Asp	Leu	Thr	Leu	Tyr	His	Phe	Phe	Asn	Glu	Val
		275					280					285			
Leu	Arg	Gln	Asn	Phe	Thr	Gly	Ala	Val	Trp	Ile	Ala	Ser	Glu	Ser	Trp
		290				295					300				
Ala	Ile	Asp	Pro	Val	Leu	His	Asn	Leu	Thr	Glu	Leu	Gly	His	Leu	Gly

[0029]

305	310	315	320
Thr Phe Leu Gly	Ile Thr Ile Gln Ser Val	Pro Ile Pro Gly Phe Ser	
	325	330	335
Glu Phe Arg Glu Trp Gly Pro Gln Ala Gly Pro Pro Pro Leu Ser Arg			
	340	345	350
Thr Ser Gln Ser Tyr Thr Cys Asn Gln Glu Cys Asp Asn Cys Leu Asn			
	355	360	365
Ala Thr Leu Ser Phe Asn Thr Ile Leu Arg Leu Ser Gly Glu Arg Val			
	370	375	380
Val Tyr Ser Val Tyr Ser Ala Val Tyr Ala Val Ala His Ala Leu His			
	385	390	395
Ser Leu Leu Gly Cys Asp Lys Ser Thr Cys Thr Lys Arg Val Val Tyr			
	405	410	415
Pro Trp Gln Leu Leu Glu Glu Ile Trp Lys Val Asn Phe Thr Leu Leu			
	420	425	430
Asp His Gln Ile Phe Phe Asp Pro Gln Gly Asp Val Ala Leu His Leu			
	435	440	445
Glu Ile Val Gln Trp Gln Trp Asp Arg Ser Gln Asn Pro Phe Gln Ser			
	450	455	460
Val Ala Ser Tyr Tyr Pro Leu Gln Arg Gln Leu Lys Asn Ile Gln Asp			
	465	470	475
Ile Ser Trp His Thr Val Asn Asn Thr Ile Pro Met Ser Met Cys Ser			
	485	490	495
Lys Arg Cys Gln Ser Gly Gln Lys Lys Lys Pro Val Gly Ile His Val			
	500	505	510
Cys Cys Phe Glu Cys Ile Asp Cys Leu Pro Gly Thr Phe Leu Asn His			
	515	520	525
Thr Glu Asp Glu Tyr Glu Cys Gln Ala Cys Pro Asn Asn Glu Trp Ser			
	530	535	540
Tyr Gln Ser Glu Thr Ser Cys Phe Lys Arg Gln Leu Val Phe Leu Glu			
	545	550	555
Trp His Glu Ala Pro Thr Ile Ala Val Ala Leu Leu Ala Ala Leu Gly			
	565	570	575

[0030]

Phe Leu Ser Thr Leu Ala Ile Leu Val Ile Phe Trp Arg His Phe Gln			
580	585	590	
Thr Pro Ile Val Arg Ser Ala Gly Gly Pro Met Cys Phe Leu Met Leu			
595	600	605	
Thr Leu Leu Leu Val Ala Tyr Met Val Val Pro Val Tyr Val Gly Pro			
610	615	620	
Pro Lys Val Ser Thr Cys Leu Cys Arg Gln Ala Leu Phe Pro Leu Cys			
625	630	635	640
Phe Thr Ile Cys Ile Ser Cys Ile Ala Val Arg Ser Phe Gln Ile Val			
645	650	655	
Cys Ala Phe Lys Met Ala Ser Arg Phe Pro Arg Ala Tyr Ser Tyr Trp			
660	665	670	
Val Arg Tyr Gln Gly Pro Tyr Val Ser Met Ala Phe Ile Thr Val Leu			
675	680	685	
Lys Met Val Ile Val Val Ile Gly Met Leu Ala Thr Gly Leu Ser Pro			
690	695	700	
Thr Thr Arg Thr Asp Pro Asp Asp Pro Lys Ile Thr Ile Val Ser Cys			
705	710	715	720
Asn Pro Asn Tyr Arg Asn Ser Leu Leu Phe Asn Thr Ser Leu Asp Leu			
725	730	735	
Leu Leu Ser Val Val Gly Phe Ser Phe Ala Tyr Met Gly Lys Glu Leu			
740	745	750	
Pro Thr Asn Tyr Asn Glu Ala Lys Phe Ile Thr Leu Ser Met Thr Phe			
755	760	765	
Tyr Phe Thr Ser Ser Val Ser Leu Cys Thr Phe Met Ser Ala Tyr Ser			
770	775	780	
Gly Val Leu Val Thr Ile Val Asp Leu Leu Val Thr Val Leu Asn Leu			
785	790	795	800
Leu Ala Ile Ser Leu Gly Tyr Phe Gly Pro Lys Cys Tyr Met Ile Leu			
805	810	815	
Phe Tyr Pro Glu Arg Asn Thr Pro Ala Tyr Phe Asn Ser Met Ile Gln			
820	825	830	

[0031]

Gly Tyr Thr Met Arg Arg Asp

835

<210> 22

<211> 6

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述: 合成的PDZIP肽序列

<400> 22

Ser Val Ser Thr Val Val

1

5

<210> 23

<211> 2520

<212> DNA

<213> 智人

<400> 23

```

atggggccca gggcaaagac catctgtccc ctgttcttcc tcctatgggt cctggctgag 60
ccggctgaga actcggactt ctacctgcct ggggattacc tcctgggtgg cctcttctcc 120
ctccatgcca acatgaaggc cattgttcac cttaacttcc tgcagggtcc catgtgcaag 180
gagtatgaag tgaaggatgat aggtacaac ctcatgcagg ccatgcgctt cgcgggtggag 240
gagatcaaca atgacagcag cctgctgcct ggtgtgctgc tgggctatga gatcgtggat 300
gtgtgtctaca tctccaacaa tgtccagccg gtgctctact tcctggcaca cgaggacaac 360
ctccttccca tccaagagga ctacagtaac tacatttccc gtgtgggtgg tgtcattggc 420
cctgacaact ccgagtctgt catgactgtg gccaaactcc tctccctatt tctccttcca 480
cagatcacct acagcgccat cagcgatgag ctgcgagaca aggtgcgctt cccggctttg 540
ctgcgtacca caccagcgc cgaccaccac gtcgaggcca tgggtgcagct gatgctgcac 600
ttccgctgga actggatcat tgtgctgggt agcagcgaca cctatggccg cgacaatggc 660
cagctgcttg gcgagcgctt ggcccggcgc gacatctgca tcgccttcca ggagacgtg 720
cccacactgc agcccaacca gaacatgacg tcagaggagc gccagcgctt ggtgaccatt 780
gtggacaagc tgcagcagag cacagcgccg gtcgtggtcg tgttctcgcc cgacctgacc 840
ctgtaccact tcttcaatga ggtgctgcgc cagaacttca cgggcgccgt gtggatcgcc 900
tccgagtcct gggccatcga cccggctcct cacaacctca cggagctggg ccacttgggc 960
accttctctg gcatcaccat ccagagcgtg cccatcccgg gcttcagtga gttccgcgag 1020
tggggcccac aggtctggcc gccacccttc agcaggacca gccagagcta tacctgcaac 1080
caggagtgcg acaactgcct gaacgccacc ttgtccttca acaccattct caggctctct 1140
ggggagcgtg tcgtctacag cgtgtactct gcggtctatg ctgtggccca tgccctgcac 1200
agcctcctcg gctgtgacaa aagcacctgc accaagaggg tggcttacct ctggcagctg 1260
cttgaggaga tctggaaggt caacttact ctcttgacc accaaatctt cttcgaccgg 1320
caaggggacg tggctctgca ctgggagatt gtccagtggc aatgggaccg gagccagaat 1380
cccttccaga gcgtcgcttc ctactacccc ctgcagcgac agctgaagaa catccaagac 1440
atctcctggc acaccgtcaa caacacgac cctatgtcca tgtgttccaa gaggtgccag 1500
tcagggcaaa agaagaagcc tgtgggcacg cacgtctgct gcttcgagtg catcgactgc 1560

```

[0032]

```

cttccccgga ccttcctcaa ccacactgaa gatgaatatg aatgccaggc ctgcccgaat 1620
aacgagtggt cctaccagag tgagacctcc tgcttcaagc ggcagctggt cttcctggaa 1680
tgcatgagg caccacccat cgctgtggcc ctgctggccg ccctgggctt cctcagcacc 1740
ctggccatcc tggatgatatt ctggaggcac ttccagacac ccatagttcg ctgggctggg 1800
ggccccatgt gcttcctgat gctgacctg ctgctgggtg catacatggt ggtcccgggtg 1860
tacgtggggc gcccgaagg ctccacctgc ctctgccgcc aggcctctt tcccctctgc 1920
ttcacaattt gcatctcctg tatcgccgtg cgcttcttcc agatcgtctg cgccttcaag 1980
atggccagcc gcttcccacg cgcctacagc tactgggtcc gctaccaggg gccctacgtc 2040
tctatggcat ttatcacggt actcaaatg gtcattgtgg taattggcat gctggccacg 2100
ggcctcagtc ccaccaccg tactgacccc gatgaaccca agatcacaat tgtctcctgt 2160
aaccccaact accgcaacag cctgctgttc aacaccagcc tggacctgct gctctcagtg 2220
gtgggtttca gcttcgcta catgggcaaa gagctgcccc ccaactacaa cgaggccaag 2280
ttcatcacc tcagcatgac cttctatttc acctcatccg tctccctctg caccttcatg 2340
tctgcctaca gcgggggtgct ggtcaccatc gtggacctct tggtcactgt gctcaacctc 2400
ctggccatca gcctgggcta cttcgcccc aagtgtctaca tgatcctctt ctaccggag 2460
cgcaacacgc ccgcctactt caacagcatg atccagggt acaccatgag gagggactag 2520

```

<210> 24

<211> 6

<212> PRT

<213> 人工序列

<220>

<223> 人工序列描述: 6-His标签

<400> 24

His His His His His His

1

5

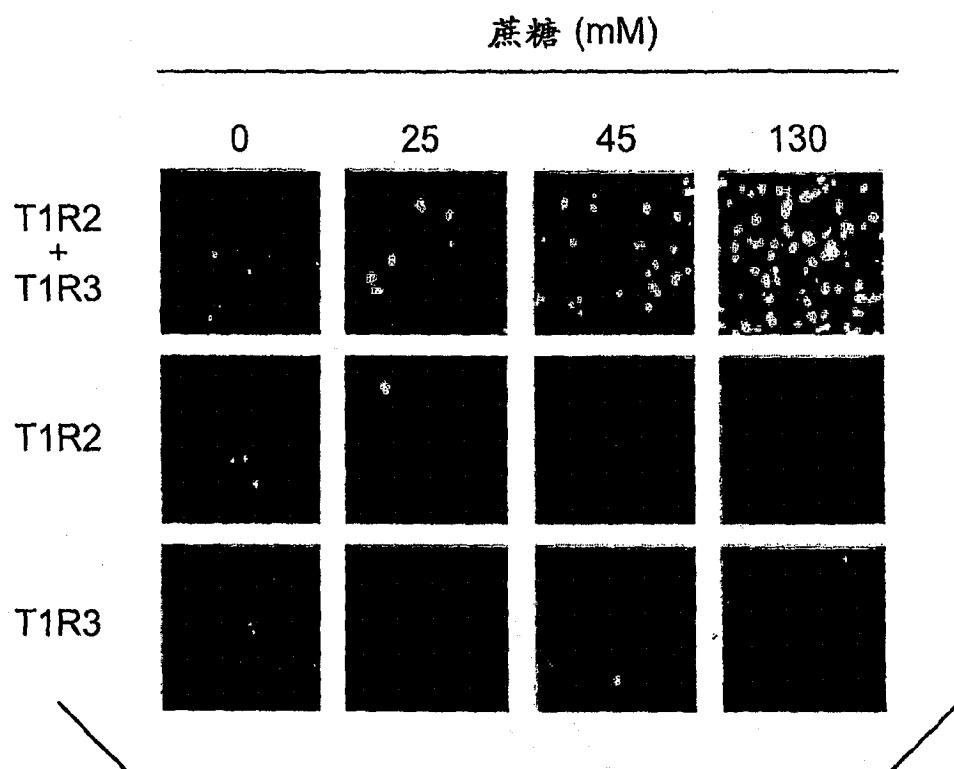


图 1A

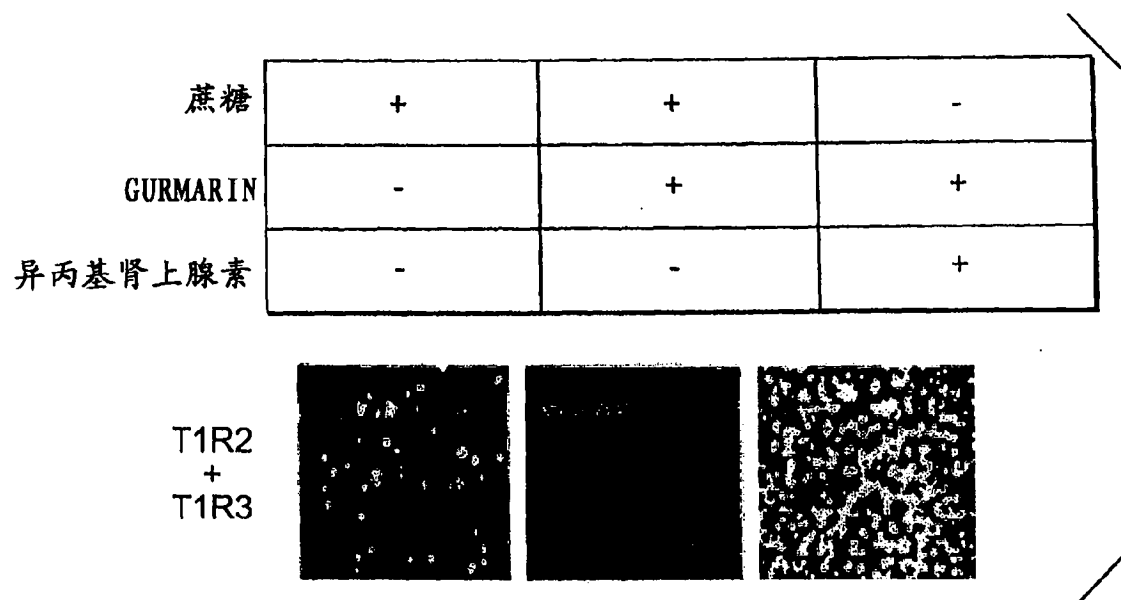


图 1B

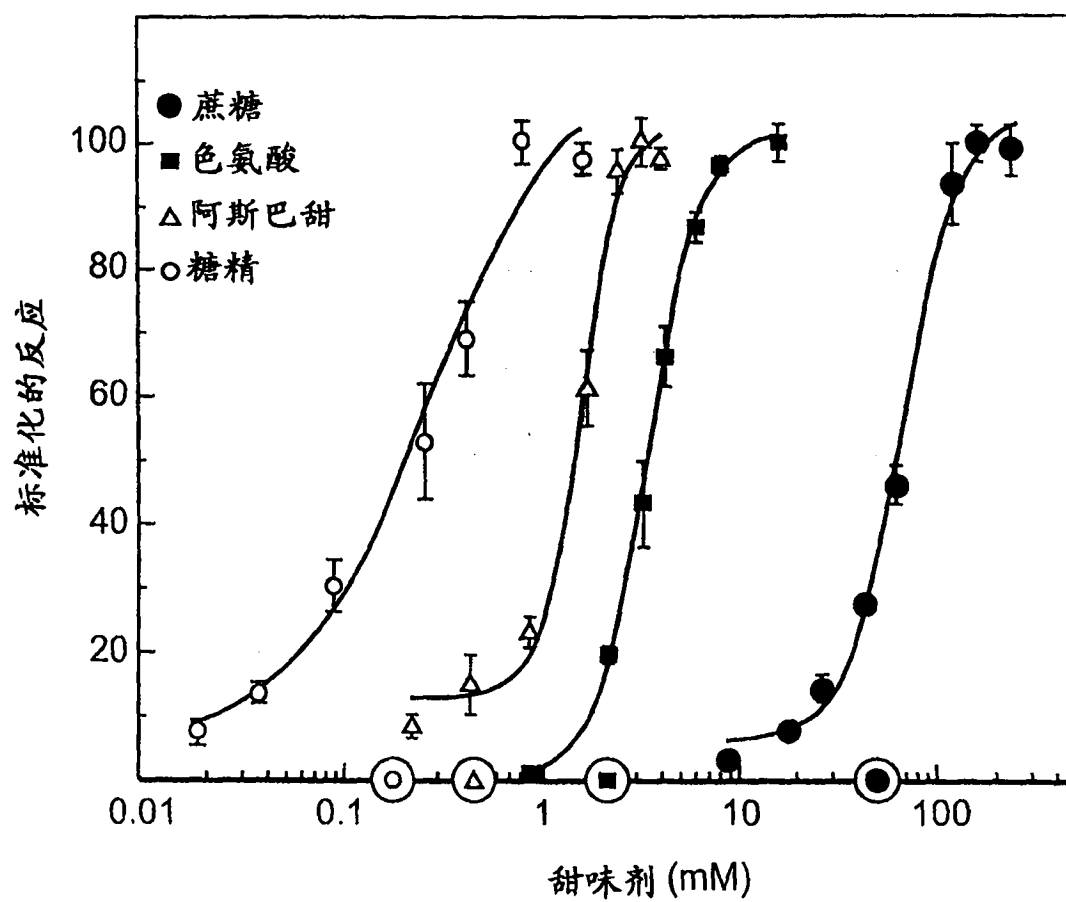


图 1C

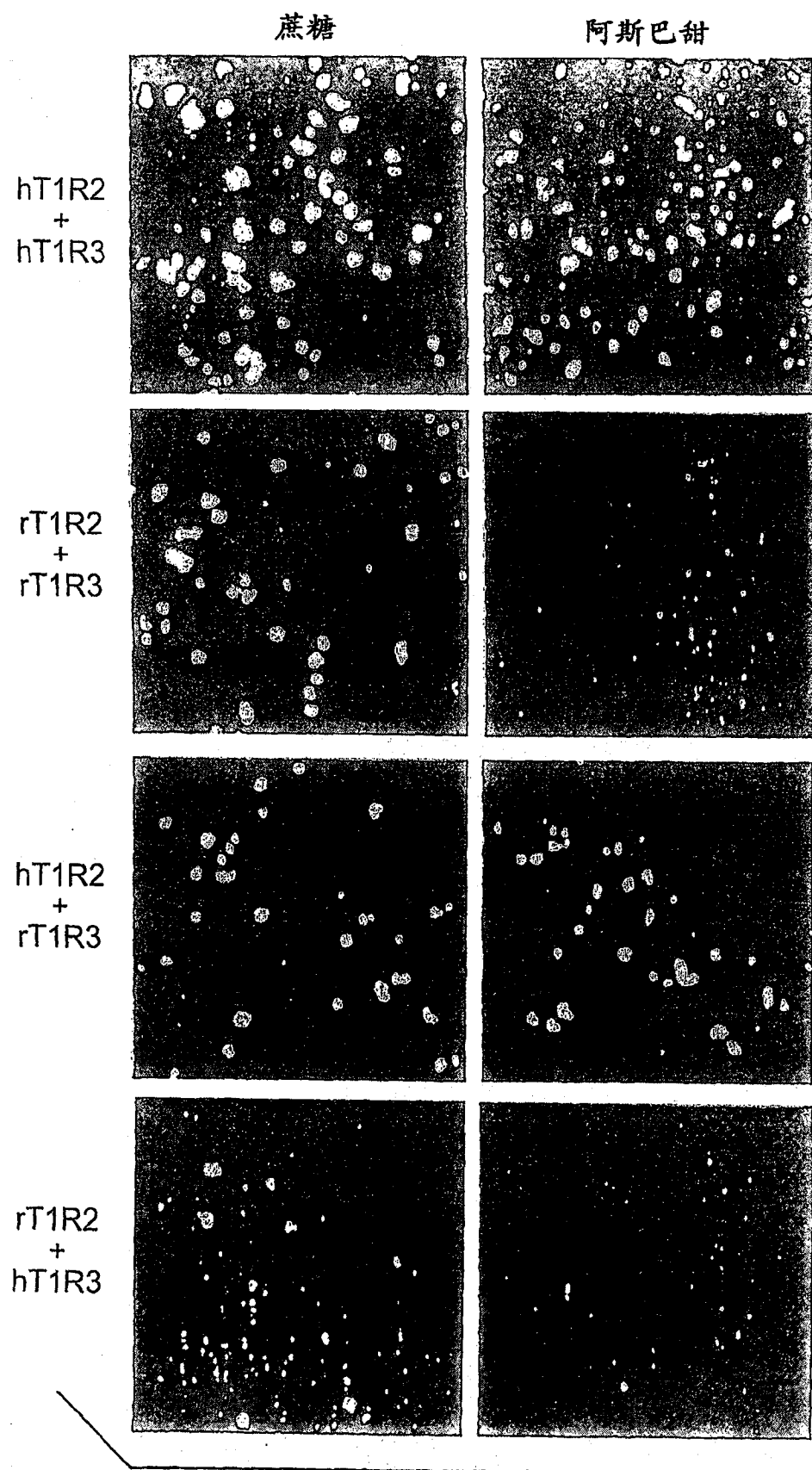


图 2

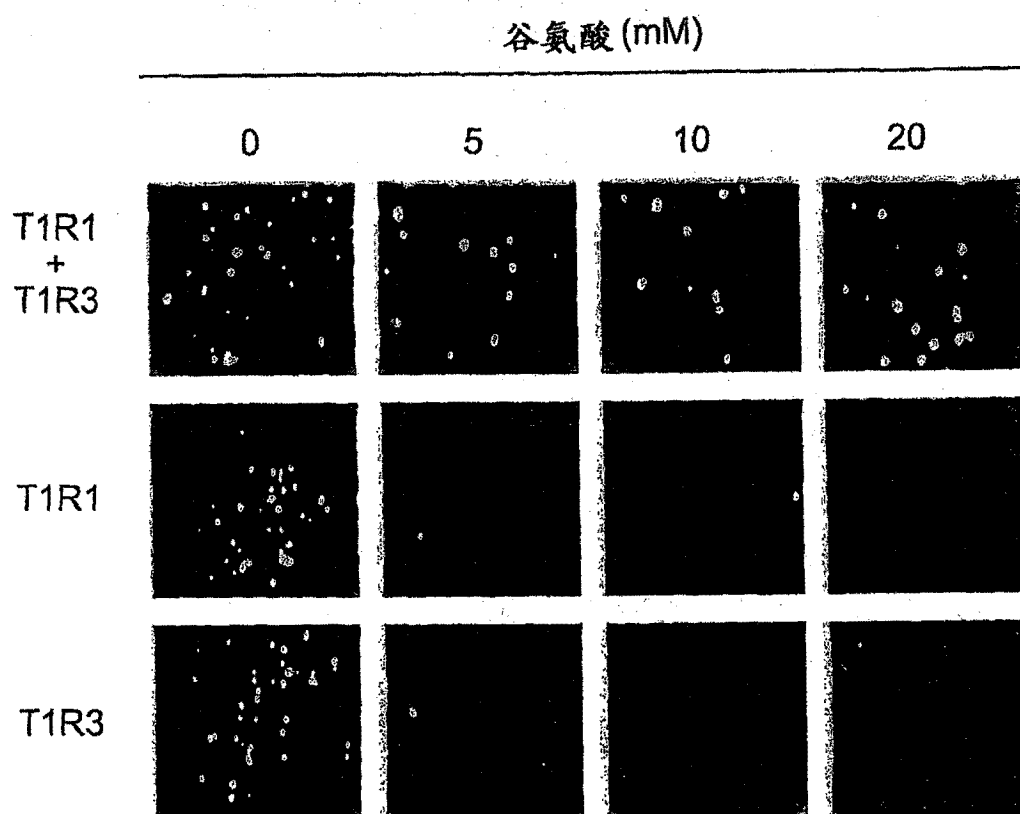


图 3A

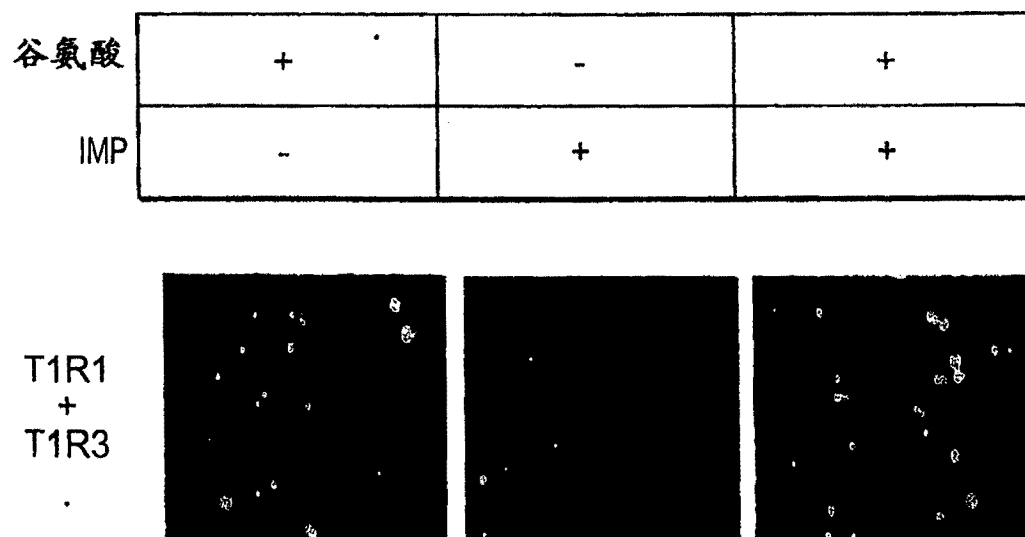


图 3B

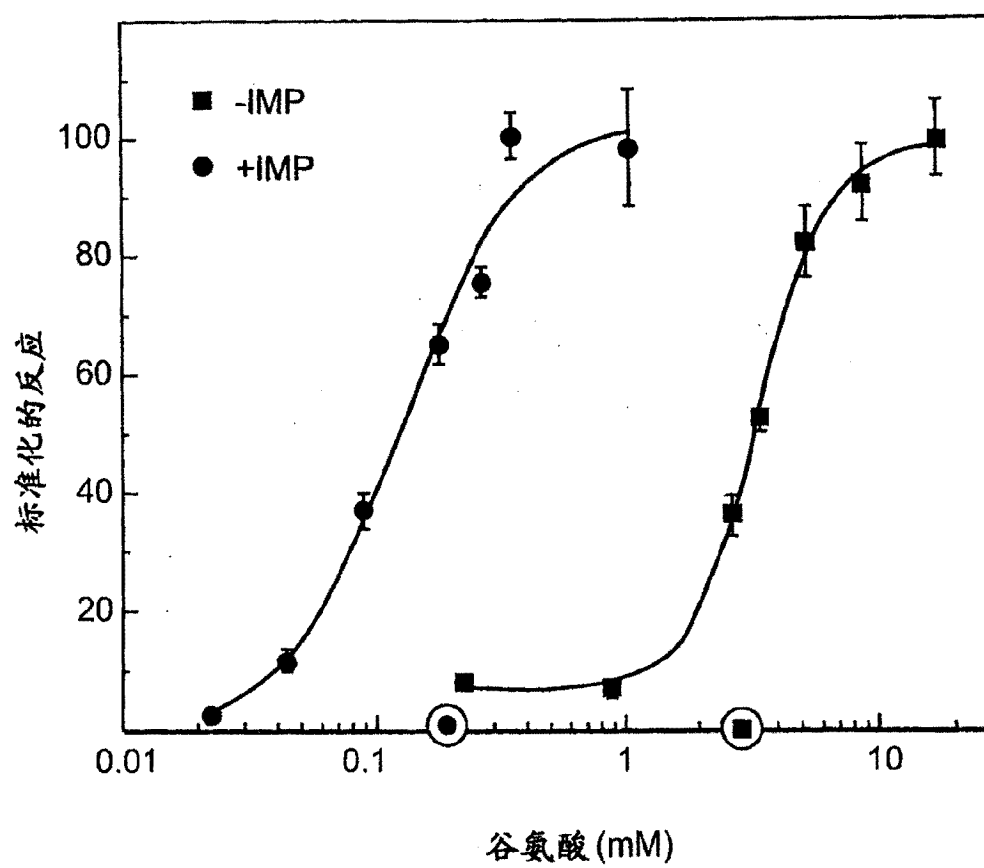


图 3C

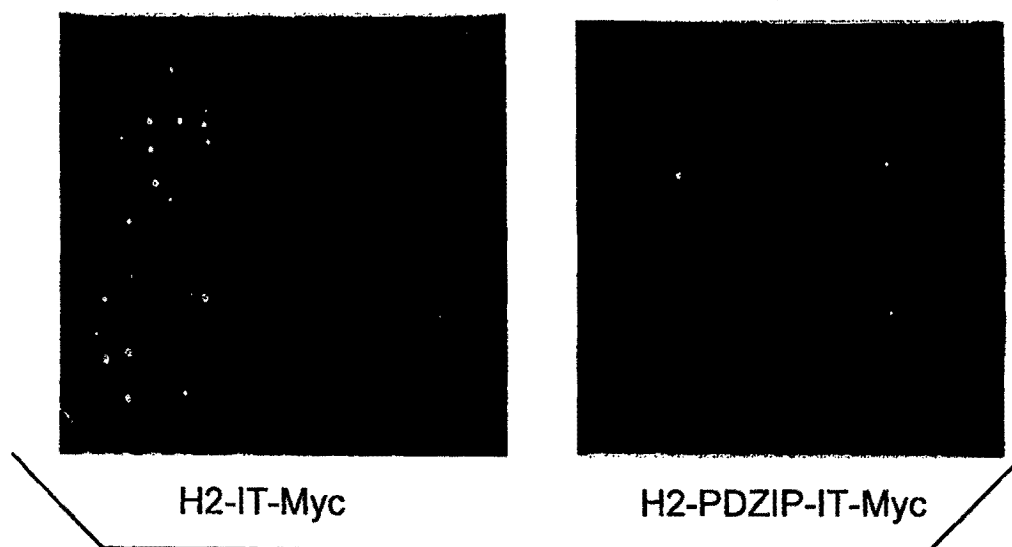


图 4A

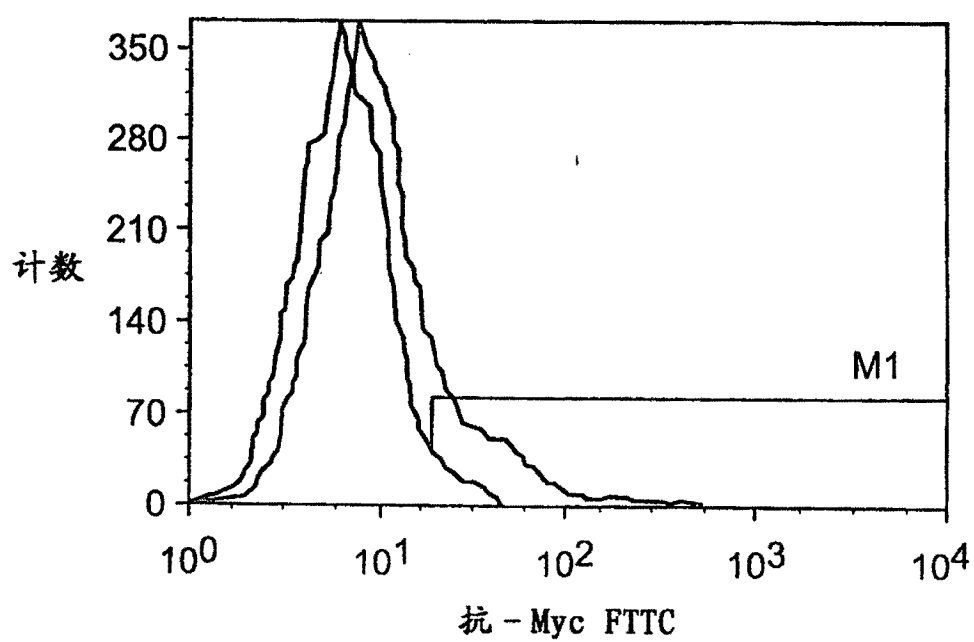


图 4B

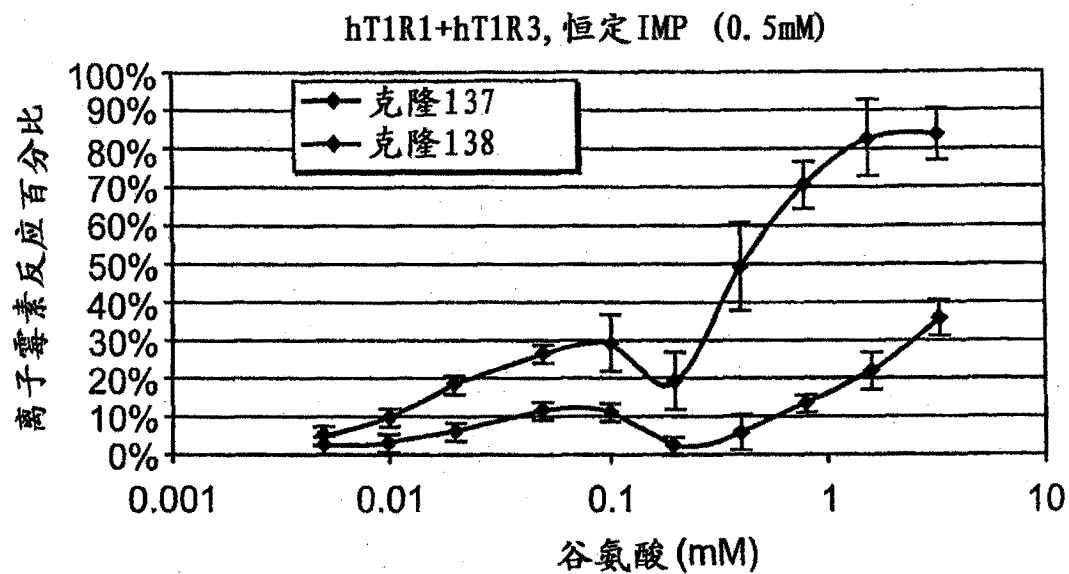


图 5

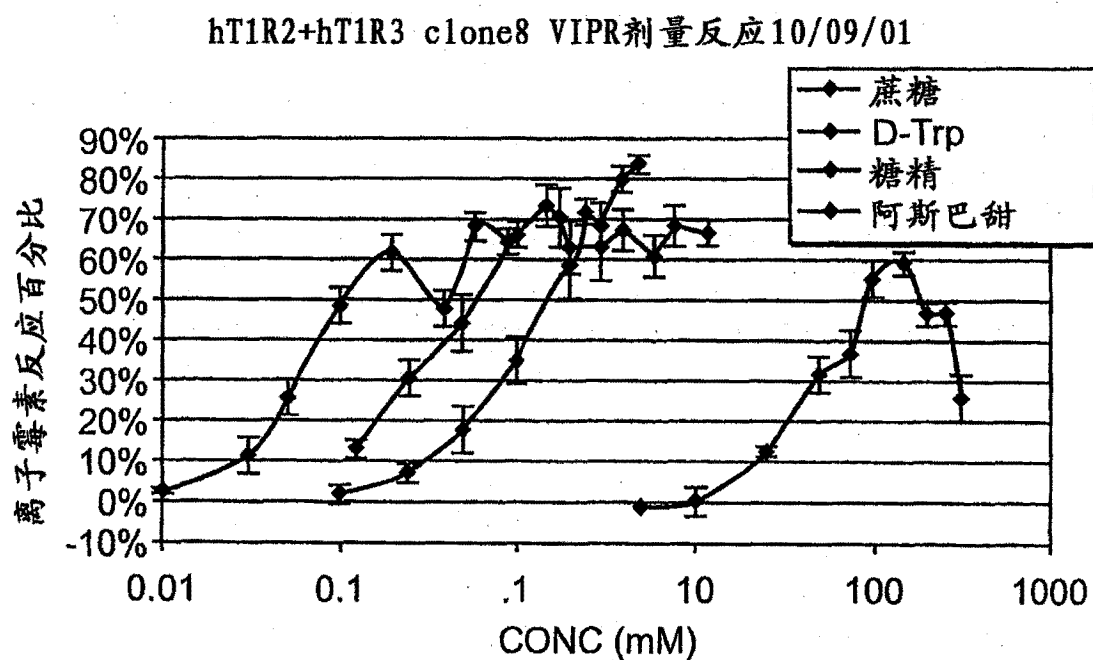


图 6

专利名称(译)	T1R味觉受体及其编码基因		
公开(公告)号	CN1525981B	公开(公告)日	2012-07-04
申请号	CN02803431.7	申请日	2002-01-03
[标]申请(专利权)人(译)	塞诺米克斯公司		
申请(专利权)人(译)	塞诺米克斯公司		
当前申请(专利权)人(译)	塞诺米克斯公司		
[标]发明人	JE阿德勒 李晓东 L斯塔谢夫斯基 S奥康奈尔 S佐祖利亚		
发明人	J·E·阿德勒 李晓东 L·斯塔谢夫斯基 S·奥康奈尔 S·佐祖利亚		
IPC分类号	C07K14/705 C12N15/12 C12N15/62 C07K16/18 C12Q1/68 C12N5/10 G01N33/50 A23L1/305 C07K16/28 C07K19/00 C12M1/00 C12N1/15 C12N1/19 C12N1/21 C12N15/09 C12Q1/02 G01N33/15 G01N33/53 G01N37/00		
CPC分类号	C12Q1/6883 C07K2319/00 G01N33/6872 G01N33/5041 G01N2333/726 C07K14/705 G16B5/00 C07K14/723 G01N33/74		
代理人(译)	程泳		
审查员(译)	吴永庆		
优先权	60/259227 2001-01-03 US 60/284547 2001-04-19 US		
其他公开文献	CN1525981A		
外部链接	Espacenet SIPO		

摘要(译)

描述了最近鉴定的哺乳动物味觉细胞特异性G蛋白偶联受体以及编码该受体的基因和cDNA。特别地，描述了在味觉信号传导中起作用的T1R G蛋白偶联受体及其编码基因和cDNA，同时还描述了用于分离该基因和用于分离与表达该受体的方法。也描述了用于在哺乳动物中表示特殊味觉刺激物的味觉感受的方法，以及用于产生在哺乳动物中引起预定味觉感受的新分子或分子组合的方法，和模仿一种或多种味觉的方法。此外，也公开了用于在哺乳动物中刺激或阻断味觉感受的方法。

$$Int_{\Pi} - Int_{\perp}$$

$$P = Int_{\Pi} + Int_{\perp}$$