(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 101796184 A
(43)申请公布日 2010.08.04

(21)申请号 200880105145.0

(22)申请日 2008.06.26

(30)优先权数据
60/937,616 2007.06.27 US

(85)PCT申请进入国家阶段日
2010.03.01

(86)PCT申请的申请数据
PCT/JP2008/001665 2008.06.26

(87)PCT申请的公布数据
WO2009/001562 EN 2008.12.31

(71)申请人 肿瘤疗法科学股份有限公司
地址 日本神奈川县

(72)发明人 富樫亮 飞田隆太郎 石崎优香
小沼亚希子

(74)专利代理机构 北京市柳沈律师事务所 11105
代理人 封新琴

(51)Int.Cl.
C12N 15/00(2006.01)
A61K 31/7105(2006.01)
A61K 48/00(2006.01)
A61P 35/00(2006.01)

权利要求书 2 页 说明书 35 页 序列表 69 页
附图 8 页

(54)发明名称

治疗癌症的组合物及方法

(57)摘要

本发明涉及一种通过施用针对选自C14orf78、MYBL2、UBE2S和UBE2T中的CX基因的双链核酸分子来治疗癌症的方法。此外，本发明还涉及包含双链核酸分子和编码它们的载体的产物，以及包含在上述方法中有效的分子或载体的组合物。本发明的方法适用于包括肺癌、乳腺癌、膀胱癌、食道癌、前列腺癌、胆管细胞癌和睾丸精原细胞瘤在内的癌症的治疗。

CN 101796184 A

1. 一种分离的双链核酸分子，它当被导入细胞时能够抑制选自 C14orf78、MYBL2、UBE2S 和 UBE2T 中的 CX 基因的表达以及表达该 CX 基因的细胞的细胞生长，并且该双链核酸分子以选自 SEQ ID NO :47 ～ 57 中的序列作为靶标。

2. 权利要求 1 所述的分离的双链核酸分子，其具有包含与选自 SEQ IDNO :47 ～ 57 中的靶序列对应的序列的有义链。

3. 权利要求 2 所述的双链核酸分子，其长度为约 19 ～约 25 个核苷酸。

4. 权利要求 1 所述的双链核酸分子，其由单一多核苷酸构成，该单一多核苷酸包含通过间插单链连接在一起的有义链和反义链。

5. 权利要求 4 所述的双链核酸分子，该双链核酸分子具有通式 5′ -[A]-[B]-[A′]-3′ ，其中 [A] 为包含与选自 SEQ ID NO :47 ～ 57 中的靶序列对应的序列的有义链，[B] 为由 3 ～ 23 个核苷酸构成的间插单链，且 [A′] 为包含与 [A] 互补的序列的反义链。

6. 权利要求 1 所述的双链核酸分子，其包含 3’突出端。

7. 表达权利要求 1 所述的双链核酸分子的载体。

8. 权利要求 7 所述的载体，其中，所述双链核酸分子具有通式 5’-[A]-[B]-[A’]-3’，其中 [A] 为包含与选自 SEQ ID NO :47 ～ 57 中的靶序列对应的序列的有义链，[B] 为由 3 ～ 23 个核苷酸构成的间插单链，且 [A′] 为包含与 [A] 互补的序列的反义链。

9. 一种用于治疗癌症的方法，该方法包括施用至少一种在过表达 CX 基因的细胞中抑制该 CX 基因表达的分离的双链核酸分子的步骤，其中，所述 CX 基因选自 C14orf78、MYBL2、UBE2S 和 UBE2T，所述双链核酸分子以选自 SEQ ID NO :47 ～ 57 中的序列为靶标。

10. 权利要求 9 所述的方法，其中有义链包含与选自 SEQ ID NO :47 ～ 57 中的靶序列对应的序列。

11. 权利要求 9 所述的方法，其中，所治疗的癌症选自下组 :

(i) 当所选择的 CX 基因为 C14orf78 时，选自胰腺癌、胆管细胞癌和非小细胞肺癌 ;

(ii) 当所选择的 CX 基因为 MYBL2 时，选自胰腺癌、非小细胞肺癌、小细胞肺癌、膀胱癌、食道癌和睾丸精原细胞瘤 ;

(iii) 当所选择的 CX 基因为 UBE2S 时，选自胰腺癌、乳腺癌、前列腺癌、小细胞肺癌、膀胱癌、胆管细胞癌和结肠癌 ;和

(iv) 当所选择的 CX 基因为 UBE2T 时，选自乳腺癌、非小细胞肺癌、小细胞肺癌、膀胱癌、胆管细胞癌、前列腺癌和食道癌。

12. 权利要求 9 所述的方法，其中施用多于一种所述双链核酸分子。

13. 权利要求 10 所述的方法，其中，所述双链核酸分子的长度为约 19 ～约 25 个核苷酸。

14. 权利要求 9 所述的方法，其中，所述双链核酸分子由单一多核苷酸构成，所述多核苷酸包含通过间插单链连接在一起的有义链和反义链。

15. 权利要求 14 所述的方法，其中，所述双链核酸分子具有通式 5′ -[A]-[B]-[A′]-3′ ，其中 [A] 为包含与选自 SEQ ID NO :47 ～ 57 中的靶序列对应的序列的有义链，[B] 为由 3 个～ 23 个核苷酸构成的间插单链，且 [A′] 为包含与 [A] 互补的序列的反义链。

16. 权利要求 9 所述的方法，其中，所述双链核酸分子包含 3’突出端。

17. 权利要求 9 所述的方法，其中，所述双链核酸分子由载体编码。

18. 权利要求 17 所述的方法，其中，由载体编码的双链核酸分子具有通式 5′ -[A]-[B]-[A′]-3′ ，[A] 为包含与选自 SEQ ID NO :47 ～ 57 中的靶序列对应的序列的有义链，[B] 为由 3 个～ 23 个核苷酸构成的间插单链，且 [A′] 为包含与 [A] 互补的序列的反义链。

19. 权利要求 9 所述的方法，其中，所述双链核酸分子包含在组合物中，所述组合物除了包含该分子外，还包含转染促进剂和药学上可接受的载体。

20. 一种用于治疗癌症的组合物，该组合物包括至少一种在过表达 CX 基因的细胞中抑制 CX 基因表达的分离的双链核酸分子，其中，所述 CX 基因选自 C14orf78、MYBL2、UBE2S 和 UBE2T，所述双链核酸分子以选自 SEQ ID NO :47 ～ 57 中的序列为靶标。

21. 权利要求 20 所述的组合物，其中，所述双链核酸分子具有包含与选自 SEQ ID NO : 47 ～ 57 中的靶序列对应的序列的有义链。

22. 权利要求 20 所述的组合物，其中，所治疗的癌症选自下组 :

(i) 当所选择的 CX 基因为 C14orf78 时，选自胰腺癌、胆管细胞癌和非小细胞肺癌 ;

(ii) 当所选择的 CX 基因为 MYBL2 时，选自胰腺癌、非小细胞肺癌、小细胞肺癌、膀胱癌、食道癌和睾丸精原细胞瘤 ;

(iii) 当所选择的 CX 基因为 UBE2S 时，选自胰腺癌、乳腺癌、胆管细胞癌、前列腺癌、小细胞肺癌、膀胱癌和结肠癌 ;和

(iv) 当所选择的 CX 基因为 UBE2T 时，选自乳腺癌、非小细胞肺癌、小细胞肺癌、膀胱癌、胆管细胞癌、前列腺癌和食道癌。

23. 权利要求 20 所述的组合物，该组合物包含多于一种所述双链核酸分子。

24. 权利要求 21 所述的组合物，其中，所述双链核酸分子的长度为约 19 ～约 25 个核苷酸。

25. 权利要求 20 所述的组合物，其中，所述双链核酸分子由单一多核苷酸构成，所述多核苷酸包含通过间插单链连接在一起的有义链和反义链。

26. 权利要求 25 所述的组合物，其中，所述双链核酸分子具有通式 5′ -[A]-[B]-[A′]-3′ ，其中 [A] 为包含与选自 SEQ ID NO :47 ～ 57 中的靶序列对应的序列的有义链序列，[B] 为由 3 个～ 23 个核苷酸构成的间插单链，且 [A′] 为包含与 [A] 互补的序列的反义链。

27. 权利要求 20 所述的组合物，其中，所述双链核酸分子包含 3’突出端。

28. 权利要求 20 所述的组合物，其中，所述双链核酸分子由载体编码，并包含在该组合物之中。

29. 权利要求 28 所述的组合物，其中，所述双链核酸分子具有通式 5′ -[A]-[B]-[A′]-3′ ，其中 [A] 为包含与选自 SEQ ID NO :47 ～ 57 中的靶序列对应的序列的有义链，[B] 为由 3 个～ 23 个核苷酸构成的间插单链，且 [A′] 为包含与 [A] 互补的序列的反义链。

30. 权利要求 20 所述的组合物，其中，该组合物包含转染促进剂和药学上可接受的载体。

治疗癌症的组合物及方法

技术领域

[0001] 相关申请的交叉引用

[0002] 本申请要求 2007 年 6 月 27 日提交的美国临时申请流水号 60/937,616 的权益，其全部内容援引并入本文。

[0003] 技术领域

[0004] 本发明涉及生物科学领域，更具体地涉及癌症研究领域。特别地，本发明涉及双链核酸分子，其抑制选自下组的 CX 基因的表达 :C14orf78、MYBL2、UBE2S 和 UBE2T 基因，还涉及包括所述双链核酸分子的组合物。本发明进一步涉及利用所述分子或组合物治疗癌症的方法。

背景技术

[0005] 胰腺癌（胰腺导管腺癌）

[0006] 在西方国家，胰腺导管腺癌 (PDAC) 在癌症死亡的主要原因中排名第四，是致死率最高的恶性肿瘤之一，5 年存活率仅 4% (DiMagno EP et al.，Gastroenterology 1999 Dec,117(6) :1464-84 ;Zervos EE et al.，Cancer Control2004Jan-Feb,11(1) :23-31 ; Jemal A et al.,CA Cancer J Clin 2003 Jan-Feb,53(1) :5-26)。仅在美国就约有 30700 名患者被诊断为胰腺癌，近 30000 名患者会死于胰腺癌 (Jemal A et al.,CA Cancer J Clin 2003Jan-Feb,53(1) :5-26)。由于大多数 PDAC 患者是在晚期被诊断的，现有的治疗方法均告无效。手术切除是目前唯一可能治愈的手段，但是 80% -90%经手术的 PDAC 患者会出现复发，而死于该病 (DiMagno EP et al.，Gastroenterology 1999 Dec,117(6) :1464-84 ; Zervos EE et al.,Cancer Control 2004 Jan-Feb,11(1) :23-31)。手术和化疗方面的一些手段，包括 5- 氟尿嘧啶 (5-FU) 或吉西他滨，结合或不结合放疗，可改善患者的生活质量 (DiMagno EP et al.,Gastroenterology 1999 Dec,117(6) :1464-84 ;Zervos EE et al.， Cancer Control 2004 Jan-Feb,11(1) :23-31)。但是，这些疗法对延长存活的作用有限，因为 PDAC 极具侵略性和化疗抗性。为了克服这种几乎令人绝望的形势，通过鉴定分子靶标来发展新的用于 PDAC 的分子疗法是当务之急。

[0007] 肺癌

[0008] 肺癌是世界上导致癌症死亡的主要原因，而非小细胞肺癌 (NSCLC) 占肺癌病例的近 80% (Greenlee RT et al.，CA Cancer J Clin 2001 Jan-Feb,51(1) :15-36)。许多与肺癌发生和发展相关的遗传学改变已有报道，但精确的分子机制仍然不清楚 (Sozzi G, Eur J Cancer 2001 Oct,37 Suppl 7 :S63-73)。在过去十年间，一系列新开发的化疗剂，如紫杉醇 (paclitaxel)、多西他赛 (docetaxel)、吉西他滨 (gemcitabine) 和长春瑞滨 (vinorelbine)，已开始为患有晚期肺癌患者的治疗提供了多种选择 ;但与基于顺铂的疗法相比，这些治疗方案对于存活率的改善差强人意 (Kelly K et al.,J Clin Oncol 2001 Jul 1,19(13) :3210-8 ;Schiller JH et al.，N Engl J Med 2002 Jan 10,346(2) :92-8)。因此，人们正在急切地探寻新的治疗策略，例如针对分子靶标的药物和抗体以及癌症疫苗。

[0009] 与其他类型的肺癌相比，小细胞肺癌 (SCLC) 更容易出现到诊断时已广泛播散的情况，具有高度的侵略性，临床上表现迅速生长、频繁浸润和转移等特征 (Ihde DC，N Engl J Med 1992 Nov 12，327(20) :1434-41)。SCLC 是常见的肺癌类型，通常被划归神经内分泌肺肿瘤的范围内，SCLC 的起源被认为是神经嵴。众所周知，SCLC 最初可能对化疗和放疗敏感，但不幸地，它们中的许多会产生对任何治疗的抵抗力。

[0010] 乳腺癌

[0011] 全世界每年有一百万妇女被诊断患有乳腺癌。雌激素受体 (ER)- 阳性乳腺癌通常预后更好，因为含抗雌激素试剂 (如他莫昔芬 (tamoxifen) 或托瑞米芬 (toremifene)) 的辅助激素疗法通常有效，无论年龄、绝经状态、腋淋巴结受累或肿瘤大小等因素如何。对于患有 ER- 阳性晚期乳腺癌的绝经后妇女的内分泌治疗，含非类固醇类第三代产品芳香酶抑制剂的雌激素剥夺治疗甚至比他莫昔芬更有效 (Nabholtz JM et al.，J Clin Oncol 2000 Nov 15，18(22) :3758-67 ;Mouridsen H et al.，J Clin Oncol 2001 May 15，19(10) : 2596-606)。虽然这些药剂具有重要的临床价值，但内分泌疗法的主要不足仍然是几乎普遍发生的化疗抗性。最初对内分泌疗法响应的 ER- 阳性乳腺癌中，大多数会获得对抗雌激素治疗的抗性，而转化成 ER- 阴性肿瘤。不幸的是，ER- 阴性乳腺癌往往会更具侵略性，而且对抗雌激素药无响应 (Goldhirsch A et al.，JClin Oncol 2003 Sep 1，21(17) : 3357-65，Epub 2003 Jul 7)。人们正在研究很多针对该疾病的靶向药物疗法，包括酪氨酸激酶抑制剂 (Gee JM et al.，Endocrinology 2003 Nov，144(11) :5105-17，Epub 2003 Aug 7 ;Moulder SL &Arteaga CL，Clin Breast Cancer 2003 Jun，4(2) :142-5 ;Okubo S et al.，Br JCancer 2004 Jan 12，90(1) :236-44 ;Schneeweiss A et al.，Anticancer Drugs 2004Mar，15(3) :235-8 ;Warburton C et al.，Clin Cancer Res 2004 Apr 1，10(7) : 2512-24)，然而迄今仅在少数的患者中取得了有希望的结果，而有些受者遭受了严重的不良反应。

[0012] 膀胱癌

[0013] 膀胱癌是人类群体中第二常见的泌尿生殖器肿瘤，全世界每年有 261，000 例新发病例。大多数膀胱癌表现为浅表性疾病，可能在 50％ -75％的病例中复发 (Heney NM et al.，J Urol 1983 Dec，130(6) :1083-6)。因此，这种癌症的进行患病率远超过其原发发生率。此外，尽管这些病例中仅 15％ -25％可能会发展，但病例的另外 25％在最初表现时就是浸润性的 (Kaye KW &Lange PH，J Urol 1982 Jul，128(1) :31-3)。因此，这种癌症需要高度关注。虽然根治性膀胱切除术当前视为局限性但肌肉浸润性膀胱癌的患者的常规治疗，然而这些患者中约有 50％在接受膀胱切除术后 2 年内出现转移，并随后死于该疾病 (Sternberg CN，Ann Oncol 1995Feb，6(2) :113-26)。

[0014] 食道癌

[0015] 食道中发生的癌症是世界范围流行的恶性肿瘤，特别是在太平洋国家。对于患有可切除的局部晚期食道癌的患者，手术仍然是常规治疗方法。根治性切除术在 50％病例中是可行的，然而切除术后局部病变或远端病变是常见的 (Tepper J，J Clin Oncol 2000 Feb，18(3) :453-4)。对于经手术的 III 期和 IV 期患者，5 年存活率仅约 30％。已有人尝试用一些辅助的多科性治疗控制局部疾病和全身疾病 (Coia LR et al.，J Clin Oncol，2000Feb，18(3) :455-62 ;Pouliquen X et al.，Ann Surg 1996 Feb，223(2) :127-33)。但

是，不能切除的和复发的食管癌对目前可用的化疗或放疗方案可能有抗性，而且这些方案对总存活率几乎没有明显优势。因此，需要开发新的有效治疗方法（例如分子靶向治疗），以提供更多的治疗方式。

[0016] 前列腺癌

[0017] 前列腺癌是男性中最常见的恶性肿瘤，而且是美国和欧洲的癌症相关死亡的第二位主要原因 (Gronberg H, Lancet 2003 Mar 8,361(9360) :859-64)，前列腺癌的发生率在大多数发达国家显著增加，很可能的原因是西方饮食方式的流行和老年人口的激增 (Hsing AW & Devesa SS, Epidemiol Rev 2001,23(1) :3-13 ;Feldman BJ & Feldman D, Nat Rev Cancer 20010ct,1(1) :34-45)。手术和放疗对局部疾病有效，但接受治疗后的前列腺癌患者有几乎 30％仍然受疾病复发的困扰 (Han M et al., J Urol 2001 Aug,166(2) :416-9 ; Isaacs Wet al., Cancer Cell 2002 Aug,2(2) :113-6)。大多数复发或晚期前列腺癌患者对雄激素消减疗法响应良好，因为前列腺癌在较早阶段通常是雄激素依赖性的。然而，这些患者常常会获得雄激素非依赖性表型，并且对雄激素消减疗法无响应或响应很差。对于晚期或复发的雄激素非依赖性前列腺癌，目前没有有效的抗癌药物或疗法。因此，迫切需要基于前列腺癌发生或激素抵抗性的分子机制来开发新的疗法。

[0018] 睾丸精原细胞瘤

[0019] 尽管睾丸生殖细胞肿瘤 (TGCT) 占男性中所有癌症的 1％ -2％，但是它们是 20-40 岁年龄群体的男性中最常见的癌症 (Chaganti, R.et al.Cancer Res.,60 :1475-1482, 2000.)，且在过去数十年中发生率显著增加 (Bergstorm, R., et al.JNatl.Cancer Inst., 88 :727-733,1996.,3 ;Zheng, T., et al.Int.J.Cancer,65 :723-729,1996.)。TGCT 分为两种主要的组织学类型，即精原细胞瘤（类似于未分化的生殖细胞）和非精原细胞瘤（可类似于胚胎和胚外两种组织，因为它们具有沿任一途径分化的能力）(Smiraglia, DJ., et al.Oncogene,21 :3909-3916,2002.)。精原细胞瘤是 TGCT 中最常见的组织学睾丸肿瘤，占所有 TGCT 的约 60％ -65％ (Richie, J.P.et al.Cambell′ s Urology Seventh Edition, pp2411-2452.Philadelphia :W.B Sauders Co.,1998)。目前，甲胎蛋白 (AFP)、人 β-亚单位绒毛膜促性腺激素 (HCG β) 和乳酸脱氢酶 (LDH) 已被用作 TGCT 的诊断肿瘤标记物 (Van Brussel, J.P.and Mikisch, G.H.J.BJU International,83 :910-917,1999)。然而，尚未鉴定出精原细胞瘤的特异性治疗靶标。

[0020] 胆管细胞癌

[0021] 胆管细胞癌是由胆管上皮细胞产生的恶性肿瘤，最初由 Durand-Fardel 于 1840 年描述。现今，其诊断和治疗仍然是难题。胆管细胞癌之所以难以诊断，部分是由于其相对罕见，还因为其临床上无症状直到晚期出现梗阻症状。在过去的三十年中，胆管细胞癌世界范围内的发病率上升。该疾病在流行方面有显著的地理差异，这很大程度上是因为区域环境的危险因素。手术切除仍然是唯一的根治疗法，一旦怀疑患有该疾病，则应优先改善诊断方法，以及判断临床分期以便进行切除。目前出现了实施积极手术治疗 (aggressive surgical management) 的新趋势，疾病后果有所改善。化疗、姑息性支架和放射线用于不可切除的患者、手术后有复发的患者以及拒绝外科手术的患者。最近利用联合全身化疗和新辅助放化疗的试验显示出前景，但还需要进一步研究。

[0022] 结肠癌

[0023] 结肠癌是发达国家中癌症死亡的主要原因。具体地，在美国每年报道超过130,000例结肠癌新发病例。结肠癌占所有癌症的约15％。其中，大约5％与通过遗传继承的基因缺陷直接相关。许多患者在癌症发作之前诊断为癌前结肠或直肠息肉。虽然许多小的结肠直肠息肉是良性的，但有些类型可发展成癌症。结直肠癌的最广泛使用的筛查试验是结肠镜检查。该方法用于使可疑的生长可视化，和/或做组织活检。通常，对组织活检物进行组织学检查，根据活检细胞的镜下表现给出诊断结果。然而，该方法是有局限性的，因为其给出的结果是主观性的，且不能用于癌前状态的极早期检出。开发敏感、特异且方便诊断的系统用于检测极早期结直肠癌或癌前病变是非常必要的，因为它可能最终消灭该疾病。

[0024] RNAi

[0025] 在细胞中不同种类的双链核酸分子可诱导RNA干扰作用，所述双链核酸分子包括短干扰RNA(siRNA)，如双链RNA(dsRNA)和短发夹RNA(shRNA)，以及短干扰DNA/RNA(siD/RNA)，如双链DNA/RNA(dsD/RNA)和短发夹DNA/RNA(shD/RNA)。在RNAi中，双链核酸分子的一条链具备与靶基因转录物(mRNA)中的核苷酸序列相同或基本相同的多核苷酸，而双链核酸分子的第二条链有与其互补的序列。不希望受理论的束缚，公认的事实是，一旦所述双链核酸分子被导入细胞，或者在细胞中由更长的双链核酸分子通过RNA酶III样酶(Rnase III like enzyme)而产生所述双链核酸分子，它就与一种蛋白复合物相互作用，该蛋白复合物称为RNA诱导的沉默复合物(RISC)。接着，RISC引导小双链核酸分子至mRNA，其中双链核酸分子的两条链分离，反义链与mRNA结合，核酸酶在双链核酸分子的反义链结合的位置处剪切mRNA(Hammond SM et al.，Nature 2000Mar 16,404(6775):293-6)。随后mRNA被细胞内的核酸酶进一步降解。短发夹类型已被证明具有很强的RNAi触发能力，在某些情况下可以比双链核酸分子更有效(Siolas D et al.，Nat Biotechnol 2005Feb,23(2):227-31,Epub 2004 Dec26)。shRNA可通过化学合成以及重组方法生成。

[0026] 近年来，一种利用基因特异性siRNA的癌症治疗新方法已在临床试验中得以实施(Bumcrot D et al.,Nat Chem Biol 2006 Dec,2(12):711-9)。RNAi在主要的技术平台中已获得了一席之地(Putral LN et al.,Drug News Perspect2006 Jul-Aug,19(6):317-24;Frantz S,Nat Rev Drug Discov 2006 Jul,5(7):528-9;Dykxhoorn DM et al.,Gene Ther 2006 Mar,13(6):541-52)。

[0027] 去端肽胶原(Atelocollagen)，一种新型siRNA投递工具

[0028] 胶原是在多种结缔组织中观察到的一种三螺旋纤维蛋白。通过胃蛋白酶处理获得的去端肽胶原所显示的免疫原性相当低，因为其没有抗原性涉及的端肽(Stenzel KH,et al.Annu.Rev.Biophys Bioeng.,1974;3:231-53)。此外，去端肽胶原可增强细胞摄取、核酸酶抗性、以及基因和寡核苷酸的持续释放(Ochiya T,et al.Curr.Gene Ther.,2001;1:31-52)。端胶原具有优异的性能，它被移植到体内时，显示低毒性和低免疫原性(Ochiya T,et al.Curr.Gene Ther.,2001;1:31-52;Sano A,et al.Adv.Drug Deliv.Rev.,2003;55:1651-77)。Ochiya等人最近的研究显示，去端肽胶原可用作siRNA的载体(Minakuchi Y,et al.Nucleic Acids Res.2004;32:e 109;Takeshita F,et al.ProcNatl Acad Sci USA.2005 August 23;102:12177-82)。

[0029] 非专利引用1:DiMagno EP et al.,Gastroenterology 1999 Dec,117(6):1464-84

[0030] 非专利引用 2 :Zervos EE et al.,Cancer Control 2004Jan-Feb,11(1) :23-31

[0031] 非专利引用 3 :Jemal A et al.,CA Cancer J Clin 2003Jan-Feb,53(1) :5-26

[0032] 非专利引用 4 :Greenlee RT et al.,CA Cancer J Clin 2001 Jan-Feb,51(1) :15-36

[0033] 非专利引用 5 :Sozzi G,Eur J Cancer 2001 Oct,37 Suppl 7 :S63-73

[0034] 非专利引用 6 :Kelly K et al.,J Clin Oncol 2001 Jul 1,19(13) :3210-8

[0035] 非专利引用 7 :Schiller JH et al.,N Engl J Med 2002 Jan 10,346(2) :92-8

[0036] 非专利引用 8 :Hide DC,N Engl J Med 1992 Nov 12,327(20) :1434-41

[0037] 非专利引用 9 :Nabholtz JM et al.,J Clin Oncol 2000 Nov 15,18(22) :3758-67

[0038] 非专利引用 10 :Mouridsen H et al.,J Clin Oncol 2001 May 15,19(10) :2596-606

[0039] 非专利引用 11 :Goldhirsch A et al.,J Clin Oncol 2003 Sep 1,21(17) :3357-65,Epub 2003 Jul7

[0040] 非专利引用 12 :Gee JM et al.,Endocrinology 2003 Nov,144(11) :5105-17,Epub 2003 Aug 7

[0041] 非专利引用 13 :Moulder SL & Arteaga CL,Clin Breast Cancer 2003Jun,4(2) :142-5

[0042] 非专利引用 14 :Okubo S et al.,Br J Cancer 2004 Jan 12,90(1) :236-44

[0043] 非专利引用 15 :Schneeweiss A et al.,Anticancer Drugs 2004 Mar,15(3) :235-8

[0044] 非专利引用 16 :Warburton C et al.,Clin Cancer Res 2004 Apr 1,10(7) :2512-24

[0045] 非专利引用 17 :Heney NM et al.,J Urol 1983 Dec,130(6) :1083-6

[0046] 非专利引用 18 :Kaye KW & Lange PH,J Urol 1982 Jul,128(1) :31-3

[0047] 非专利引用 19 :Sternberg CN,Ann Oncol 1995 Feb,6(2) :113-26

[0048] 非专利引用 20 :Tepper J,J Clin Oncol 2000 Feb,18(3) :453-4

[0049] 非专利引用 21 :Coia LR et al.,J Clin Oncol,2000 Feb,18(3) :455-62

[0050] 非专利引用 22 :Pouliquen X et al.,Ann Surg 1996 Feb,223(2) :127-33

[0051] 非专利引用 23 :Gronberg H,Lancet 2003 Mar 8,361(9360) :859-64

[0052] 非专利引用 24 :Hsing AW & Devesa SS,Epidemiol Rev 2001,23(1) :3-13

[0053] 非专利引用 25 :Feldman BJ & Feldman D,Nat Rev Cancer 2001 Oct,1(1) :34-45

[0054] 非专利引用 26 :Han M et al.,J Urol 2001 Aug,166(2) :416-9

[0055] 非专利引用 27 :Isaacs W et al.,Cancer Cell 2002Aug,2(2) :113-6

[0056] 非专利引用 28 :Chaganti,R.et al.Cancer Res.,60 :1475-1482,2000

[0057] 非专利引用 29 :Bergstorm,R.,et al.J Natl.Cancer Inst.,88 :727-733,1996.,3

[0058] 非专利引用 30 :Zheng,T.,et al.Int.J.Cancer,65 :723-729,1996.

[0059] 非专利引用 31 :Smiraglia,D.J.,et al.Oncogene,21 :3909-3916,2002.

[0060] 非专利引用 32 :Richie, J.P.et al.Cambell′ s Urology Seventh Edition, pp2411-2452.Philadelphia :W.B Sauders Co.,1998

[0061] 非专利引用 33 :Van Brussel, J.P.and Mikisch, G.H.J.BJU International,83 : 910-917,1999

[0062] 非专利引用 34 :Hammond SM et al.,Nature 2000 Mar 16,404(6775) :293-6

[0063] 非专利引用 35 :Siolas D et al.,Nat Biotechnol 2005 Feb,23(2) :227-31, Epub 2004 Dec 26

[0064] 非专利引用 36 :Bumcrot D et al.,Nat Chem Biol 2006Dec,2(12) :711-9

[0065] 非专利引用 37 :Putral LN et al.,Drug News Perspect 2006 Jul-Aug,19(6) : 317-24

[0066] 非专利引用 38 :Frantz S,Nat Rev Drug Discov 2006 Jul,5(7) :528-9

[0067] 非专利引用 39 :Dykxhoorn DM et al.,Gene Ther 2006 Mar,13(6) :541-52

[0068] 非专利引用 40 :Stenzel KH,et al.Annu.Rev.Biophys Bioeng.,1974 ;3 :231-53

[0069] 非专利引用 41 :Ochiya T,et al.Curr.Gene Ther.,2001 ;1 :31-52

[0070] 非专利引用 42 :Sano A,et al.Adv.Drug Deliv.Rev.,2003 ;55 :1651-77

[0071] 非专利引用 43 :Minakuchi Y,et al.Nucleic Acids Res.2004 ;32 :e109

[0072] 非专利引用 44 :Takeshita F,et al.Proc Natl Acad Sci USA.2005 August23 ; 102 :12177-82

发明内容

[0073] 本发明基于下述发现 :包含特定序列 (具体为 SEQ ID NO :47 ～ 57) 的双链核酸分子对于抑制包括胰腺癌、肺癌、乳腺癌、膀胱癌、食道癌、前列腺癌、睾丸精原细胞瘤、结肠癌和胆管细胞癌在内的多种癌细胞的细胞生长是有效的。具体地说,本发明提供了以 C14orf78、MYBL2、UBE2S 和 UBE2T 基因为靶标的小干扰 RNA(siRNA)。

[0074] 根据本发明的一个方面,双链核酸分子可被编码在载体中,在体内和体外由载体进行表达。

[0075] 本发明的双链核酸分子和载体具有抑制表达靶基因 (C14orf78、MYBL2、UBE2S 或 UBE2T 基因) 的细胞的细胞生长的能力。因此,本发明提供通过施用本发明的双链核酸分子或载体来抑制细胞生长并治疗癌症的方法。此类方法包括给受试者施用包含 1 种或多种所述双链核酸分子或载体的组合物。

[0076] 本发明的另一方面涉及包含至少 1 种本发明的双链核酸分子或载体的治疗用于癌症的组合物。

[0077] 发明的公开

[0078] 鉴定出的癌症治疗候选物质的模式

[0079] C14orf78 基因 ((Genbank 登录号 :XM_290629 ;SEQ ID NO :1) 编码一种分子量达 668kDa 的巨大蛋白质 (SEQ ID NO :2 ;以下称为“C14orf78 蛋白质”)。如前人所述, C14orf78 和 AHNAK1 蛋白质属于同一家族 (KomuroA 等,Proc Natl Acad Sci USA 2004 Mar 23,101(12) :4053-8, Epub 2004 Mar8)。这种大小的 AHNAK1 蛋白质是定位于间期核的分化相关蛋白质。近年的研究中报导了利用肾上腺素激动剂刺激心肌细胞能够激活膜结合

型AHNAK1蛋白质的磷酸化(Komuro A等,Proc Natl Acad Sci USA 2004 Mar23,101(12):4053-8,Epub 2004 Mar 8)。磷酸化的AHNAK1蛋白质能够与针对L型电位调节钙通道的两种不同亚基的抗体共沉淀,这提示该蛋白质与钙通道结合(Komuro A等,Proc Natl Acad Sci USA 2004 Mar 23,101(12):4053-8,Epub 2004 Mar 8)。

[0080] 另一项报导发现,在AHNAK1敲除小鼠中,表型完全没有检测到明显异常(Komuro A等,Proc Natl Acad Sci USA 2004 Mar 23,101(12):4053-8,Epub 2004 Mar 8),这表明,就目前看,AHNAK1不是细胞增殖和分化的关键因素。

[0081] MYBL2基因所编码的蛋白质(GenBank登录号:NM_002466;SEQ IDNO:3,编码SEQ ID NO:4)作为作用于细胞增殖、分化和凋亡的、涉及细胞周期进行的转录因子而发挥功能(Oh IH & Reddy EP,Oncogene 1999 May13,18(19):3017-33;Weston K,Curr Opin Genet Dev 1998Feb,8(1):76-81)。此外,MYBL2蛋白质还显示作为基因转录的激活因子或抑制因子而发挥作用(Klempnauer KH & Sippel AE,EMBO J 1987 Sep,6(9):2719-25;Biedenkapp H等,Nature 1988 Oct 27,335(6193):835-7;Nomura N等,Nucleic Acids Res 1988 Dec 9,16(23):11075-89)。以前有报导称MYBL2基因表达通过E2F依赖性机制而限制增殖细胞,而MYBL2蛋白质的活性在S期被CDK2/细胞周期蛋白A复合体所刺激(Robinson C等,Oncogene1996 May 2,12(9):1855-64)。MYBL2蛋白质在细胞分裂中的功能与其调控细胞周期蛋白B1基因表达的能力至少部分相关(Okada M等,EMBO J2002 Feb 15,21(4):675-84.)。

[0082] 由UBE2S基因(GenBank登录号:NM_014501;SEQ ID NO:5编码SEQID NO:6)和UBE2T基因(GenBank登录号:NM_014176;SEQ ID NO:7编码SEQ ID NO:8)编码的两种蛋白质具有一个泛素结合酶E2催化结构域,被认为是在蛋白质分解途径中起作用的泛素结合酶。由近年的研究可知:UBE2S蛋白质——一种推定泛素E2连接酶——特异地靶定pVHL(vonHippel-Lindau蛋白质)以进行分解作用,而UBE2S基因的过表达显著促进细胞增殖(Ohh M Cancer Cell 2006Aug,10(2):95-7;Jung CR等,Nat Med2006 Jul,12(7):809-16,Epub 2006Jul 2)。

[0083] pVHL发挥泛素连接酶E3复合体的底物识别模块的功能,该复合体在正常氧分压条件下将低氧诱导因子-1α(HIF-1α)泛素化。HIF-1α在正常氧分压范围被正常地降解,但在低氧下会逃离蛋白质分解机制。该HIF-1α的异常蓄积引起与代谢适应(诸如肿瘤血管新生)、细胞生存所需的代谢作用、以及细胞的增殖和分化等作用相关的靶基因的激活(Semenza GL,TrendsMol Med 2001 Aug,7(8):345-50;Pugh CW & Ratcliffe PJ,Nat Med 2003 Jun,9(6):677-84)。因而,基于UBE2S蛋白质的泛素途径介导的pVHL耗竭可引起HIF-1α异常蓄积,结果可能促进癌细胞增殖。

[0084] 蛋白质的泛素化要经过一条ATP依赖性的途径。第一步需要ATP,泛素藉由其C末端甘氨酸残基通过硫酯键结合到泛素激活酶(E1)上。然后,泛素通过转硫醇酯化反应(trans-thiol esterification)被转移至泛素结合酶(E2),进而被转移至靶蛋白质的赖氨酸残基的ε-氨基,该过程通常在泛素蛋白质连接酶(E3)帮助下进行。结合后的泛素可本身作为泛素化底物起作用,反复地进行泛素化,导致多泛素链的形成。多泛素化的靶蛋白质转移至26S蛋白酶体。泛素-26S蛋白酶体(UPS)途径是真核细胞中正常和错误折叠的细胞质或膜蛋白质的主要分解机制。

[0085] 定义

[0086] 本说明书中使用的术语“一个”、“一种”(a、an)和“该(the)”,如无特殊提及是指“至少1个”。

[0087] 在癌症中差异表达的基因在本说明书中统称为“CX基因”、“CX核酸”或“CX多核苷酸”,该基因编码的相应多肽称为“CX多肽”或“CX蛋白质”。在本发明中,CX基因选自:C14orf78基因(也称“C14orf78”,GenBank登录号:XM_290629、SEQ ID NO:1),其编码一种巨大蛋白质(以下称为“C14orf78蛋白质”,SEQ ID NO:2);MYBL2基因(也称“MYBL2”,GenBank登录号:NM_002466、SEQ ID NO:3),其编码具有SEQ ID NO:4的序列的蛋白质(以下称为“MYBL2蛋白质”);UBE2S基因(也称“UBE2S”,GenBank登录号:NM_014501、SEQ ID NO:5),其编码具有SEQ ID NO:6的序列的蛋白质(以下称为“UBE2S蛋白质”);和UBE2T基因(也称“UBE2T”,GenBank登录号:NM_014176、SEQ ID NO:7),其编码具有SEQ ID NO:8的序列的蛋白质(以下称为“UBE2T蛋白质”)。在本说明书中,这些CX基因可以称为“靶基因(们)”,其中包含至少1个靶序列。

[0088] 靶序列是CX基因内的核苷酸序列,如果本发明的双链核酸分子与之结合,则整个mRNA的翻译会受到抑制。在表达CX基因的细胞中,当包含与某个核苷酸序列序列对应的序列的双链多核苷酸可抑制CX基因的表达时,则可确定该序列为靶序列。根据本发明,发现以下的序列具有作为靶序列的功能。

[0089] C14orf78基因:

[0090] SEQ ID NO:1的核苷酸

[0091] 13846-13864(SEQ ID NO:47)、

[0092] 13909-13927(SEQ ID NO:48)、

[0093] 14001-14019(SEQ ID NO:49)、和

[0094] 14647-14665(SEQ ID NO:50);

[0095] MYBL2基因:

[0096] SEQ ID NO:3的核苷酸

[0097] 977-995(SEQ ID NO:51)、

[0098] 1938-1956(SEQ ID NO:52)、

[0099] 1940-1958(SEQ ID NO:53)、和

[0100] 1995-2013(SEQ ID NO:54);

[0101] UBE2S基因:

[0102] SEQ ID NO:5的核苷酸

[0103] 706-724(SEQ ID NO:55)和

[0104] 528-546(SEQ ID NO:56);以及

[0105] UBE2T基因:

[0106] SEQ ID NO:7的核苷酸

[0107] 148-166(SEQ ID NO:57)。

[0108] 在本说明书中使用的术语“生物体”是指由至少1个细胞构成的任意有生命的实体。活生物体可以是例如真核单细胞那样的简单生物,也可以是诸如包含人在内的哺乳动物那样的复杂生物。

[0109] 在本说明书中使用的术语“生物学样品”是指：生物整体或其组织、细胞或构成部分（例如体液，包括但不限于包括血液、粘液、淋巴液、滑液、脑脊液、唾液、羊水、脐带血、尿、阴道液和精液）的子集。“生物学样品”还指由生物整体或其细胞、组织或构成部分的子集制备得到的匀浆、裂解物、细胞培养物、或组织培养物、或其级分或部分。最后，“生物学样品”还指这样的培养基，其包含蛋白质或多核苷酸等细胞成分，例如已有生物在其中增殖的营养培养基或凝胶等。

[0110] 除非另有具体说明，术语“多核苷酸”和“寡核苷酸”在本说明书中可互换使用，并以它们的通用单字母符号表示。本术语适用于这样的核酸（核苷酸）多聚物，其中一个以上核酸通过酯键连接。多核苷酸或寡核苷酸可以由 DNA、RNA 或其组合构成。

[0111] 本说明书中使用的术语“分离的双链核酸分子”是指：抑制靶基因表达的核酸分子，包括例如小干扰 RNA(siRNA：例如双链核糖核酸 (dsRNA) 或小发夹 RNA(small hairpin RNA，shRNA)) 和小干扰 DNA/RNA(siD/R-NA：例如 DNA 和 RNA 的双链嵌合体 (dsD/R-NA) 或 DNA 和 RNA 的小发夹嵌合体 (shD/R-NA))。

[0112] 在本说明书中使用的术语“siRNA”，是指阻止靶 mRNA 翻译的双链 RNA 分子。可以使用将 siRNA 导入细胞的常规技术，包括用 DNA 作为模板转录 RNA 的那些方法。siRNA 包括 CX 有义核酸序列（也称“有义链”）、CX 反义核酸序列（也称“反义链”）或这两者。siRNA 可以这样构建，使单一转录物同时具有靶基因的有义序列及互补的反义序列，例如发夹。siRNA 可以是 dsRNA 或 shRNA。

[0113] 本说明书中使用的术语“dsRNA”是指两个 RNA 分子的构建体，这两个 RNA 分子包含彼此互补的序列，并且藉由所述互补序列退火在一起形成双链 RNA 分子。两条链的核苷酸序列不仅可以包含选自靶基因序列的蛋白编码序列的“有义”或“反义”RNA，还可以包含具有选自靶基因非编码区的核苷酸序列的 RNA 分子。

[0114] 在本说明书中使用的术语“shRNA”是指：具有茎 - 环结构的 siRNA，其包含彼此互补的第一区和第二区（即有义链和反义链）。两个区的互补程度和方向足以使两个区之间发生碱基配对，所述第一区和第二区通过环区连接在一起，而所述环是因为环区内的核苷酸（或核苷酸类似物）之间缺乏碱基配对而形成的。shRNA 的环区是介于有义链和反义链之间的单链区，也可以称作“间插单链 (intervening single-strand)”。

[0115] 在本说明书中使用的术语“siD/R-NA”是指：由 RNA 和 DNA 二者组成的双链多核苷酸分子，包括 RNA 和 DNA 的杂合体和嵌合体，其阻止靶 mRNA 的翻译。在本说明书中，杂合体表示这样的分子，其中由 DNA 组成的多核苷酸和由 RNA 组成的多核苷酸相互杂交形成双链分子；而嵌合体表示组成所述双链分子的链中的一条或两条可以同时含有 RNA 和 DNA。使用将 siD/R-NA 导入细胞的常规技术。在本发明中，这样的双链核酸分子可以指双链分子。siD/R-NA 包括 CX 有义核酸序列（也称为“有义链”）和 / 或 CX 反义核酸序列（也称作“反义链”）。siD/R-NA 可以这样构建，使单个转录物同时具有来自靶基因的有义核酸序列和互补反义核酸序列，例如发夹。siB/R-NA 可以是 dsD/R-NA 或 shD/R-NA。

[0116] 在本说明书中使用的术语“dsD/R-NA”是指这样两个分子的构建体，所述两个分子包含彼此互补的序列并且已经藉由所述互补序列退火在一起而形成双链多核苷酸分子。两条链的核苷酸序列可以不仅仅包含选自靶基因序列的蛋白编码序列的“有义”或“反义”多核苷酸序列，还可以包含具有选自靶基因非编码区的核苷酸序列的多核苷酸。组成 dsD/

R-NA 的两个分子中的一个或两个由 RNA 和 DNA 二者组成（嵌合体分子），或者一个分子由 RNA 组成而另一个由 DNA 组成（杂合双链）。

[0117] 在本说明书中使用的术语“shD/R-NA”是指：具有茎－环结构的 siD/R-NA，其包含彼此互补的第一区和第二区，即有义链和反义链。所述区的互补程度和方向足以使它们之间发生碱基配对，第一区和第二区通过环区连接，所述环是因为在环区内的核苷酸（或核苷酸类似物）之间缺乏碱基配对而形成的。shD/R-NA 的环区是介于有义链和反义链之间的单链区，也可以称作“间插单链”。

[0118] 本说明书中使用的“分离的核酸”是指：从本来的环境（例如自然出现时的自然环境）中被取出，与其自然状态相比发生了合成性的改变的核酸。在本发明中，分离的核酸包括 DNA、RNA 和它们的衍生物。

[0119] 本说明书中使用的术语“CX 基因相关疾病”是指以 CX 基因与相应正常组织相比较的过表达为特征的疾病，包括例如胰腺癌、肺癌、乳腺癌、膀胱癌、食道癌、前列腺癌、睾丸精原细胞瘤、结肠癌和胆管细胞癌。

[0120] 在本说明书中，所谓抑制细胞生长，是指与未处理细胞相比较，自然表达靶基因的细胞以更低的速度生长、或者存活率降低。细胞生长可以通过传统的增殖试验、例如使用细胞分析仪 (cell analyzer)1000 进行的试验来测定。

[0121] 概要

[0122] 在非哺乳动物细胞中，双链 RNA(dsRNA) 显示针对基因表达具有强烈而特异性的沉默效应，这称作 RNA 干扰 (RNAi)(Sharp PA, Genes Dev 1999Jan 15,13(2) :139-41)。dsRNA 被一种含有 RNA 酶 III 基序的酶加工成 20-23 个核苷酸的 dsRNA，称作小干扰 RNA(siRNA)。siRNA 与一种多组分核酸酶复合物一起特异性靶向与之互补的 mRNA(Hammond SM 等, Nature 2000Mar 16,404(6775) :293-6 ;Hannon GJ, Nature 2002 Jul 11, 418(6894) :244-51)。在哺乳动物细胞中，由 20 或 21 聚体 dsRNA 构成、具有 19 个互补核苷酸和 3’ 末端非互补胸苷或尿苷二聚体的 siRNA，显示具有基因特异性敲低 (knockdown) 作用，同时不会导致基因表达的全局改变 (Elbashir SM等, Nature 2001 May 24,411(6836) : 494-8)。此外，含有小核 RNA(snRNA)U6 或聚合酶 III H1-RNA 启动子的质粒可有效产生这种能募集短 RNA 的 III 族 RNA 聚合酶 III，从而能够组成型抑制其靶 mRNA(Miyagishi M & Taira K, Nat Biotechnol 2002 May,20(5) :497-500 ;Brummelkamp TR 等, Science 2002Apr 19,296(5567) :550-3, Epub 2002 Mar 21)。

[0123] 细胞增殖抑制方法是本发明的特色。细胞增殖因针对 CX 基因的双链核酸分子与细胞接触而受到抑制。在 CX 基因中，C14orf78 在 18 例临床胰腺癌中的 11 例、25 例临床胆管细胞癌中的 14 例、37 例非小细胞肺癌中的 10 例中过表达 (T/N 比＞＝ 5) ;MYBL2 被发现多种多样的癌症中过表达，即：在 34 例临床膀胱癌中的 18 例、64 例食道癌中的 29 例、非小细胞肺癌 (NSCLC)37 例中的 18 例、18 例临床胰腺癌中的 6 例和 15 例小细胞肺癌 (SCLC) 中的 14 例中上调（比例＞＝ 5) ;UBE2S 在全部 SCLC 病例、34 例膀胱癌中的 29 例、81 例乳腺癌中的 27 例、59 例前列腺癌中的 18 例、48 例结肠癌中的 11 例、25 例胆管细胞癌中的 9 例和 18 例胰腺癌中的 12 例中过表达 ;UBE2T 也在各种类型的肿瘤，即 25 例胆管细胞癌中的 12 例、25 例 SCLC 中的 12 例、34 例膀胱癌中的 23 例、81 例乳腺癌中的 28 例、37 例 NSCLC 中的 13 例、64 例食道癌中的 14 例和 59 例前列腺癌中的 15 例中显示增加的表达（表 2)。可

以使用针对各靶基因的本发明双链核酸分子来抑制表达 CX 基因的细胞的增殖。

[0124] 所述方法被用于改变 CX 基因表达上调（例如作为细胞癌化的结果）的细胞中的基因表达。靶细胞中双链核酸分子与 CX 基因转录物的结合导致该细胞中 CX 蛋白质产生减少，细胞生长被抑制。

[0125] 双链核酸分子

[0126] 针对 CX 基因的双链核酸分子——该分子与靶 mRNA 杂交——与 CX 基因的正常单链 mRNA 转录物结合，阻碍翻译、抑制蛋白质表达，从而减少或抑制该基因编码的 CX 蛋白质的产生。PK-1 和 Panc.02.03 胰腺癌细胞系中的 C14orf78 的表达被 4 种不同的 dsRNA 所抑制（图 2a、b），NSCLC 细胞系 (H358) 与食道癌细胞系 (TE-9) 中的 MYBL2 的表达被 4 种不同的 dsRNA 所抑制（图 3a、b）；乳腺癌细胞系 (MCF7)、胰腺癌细胞系 (PK-1) 和膀胱癌细胞系 (SW780) 中的 UBE2S 的表达被两种不同的 dsRNA 所抑制（图 4a-c），乳腺癌细胞系 (MCF7)、NSCLC 细胞系 (A549)、膀胱癌细胞系 (SW780) 和前列腺癌细胞系 (DU-145) 中的 UBE2T 表达被 1 种 dsRNA 所抑制（图 5a-d）。

[0127] 因此，本发明提供分离的双链核酸分子，其具有当将其导入表达 CX 基因的细胞时可抑制该基因表达的性质。双链核酸分子的靶序列可以采用以下所述的 siRNA 设计算法来设计。

[0128] C14orf78 靶序列包括例如：

[0129] SEQ ID NO:1 的核苷酸

[0130] 13846-13864(SEQ ID NO:47)、

[0131] 13909-13927(SEQ ID NO:48)、

[0132] 14001-14019(SEQ ID NO:49) 或

[0133] 14647-14665(SEQ ID NO:50)；

[0134] MYBL2 靶序列包括例如 SEQ ID NO:3 的核苷酸

[0135] 977-995(SEQ ID NO:51)、

[0136] 1938-1956(SEQ ID NO:52)、

[0137] 1940-1958(SEQ ID NO:53) 或

[0138] 1995-2013(SEQ ID NO:54)；

[0139] UBE2S 靶序列包括例如 SEQ ID NO:5 的核苷酸

[0140] 706-724(SEQ ID NO:55) 或

[0141] 528-546(SEQ ID NO:56)；

[0142] UBE2T 靶序列包括例如 SEQ ID NO:7 的核苷酸

[0143] 148-166(SEQ ID NO:57)。

[0144] 具体地，本发明提供以下的双链核酸分子 [1] ～ [17]。

[0145] [1] 一种分离的双链核酸分子，其当被导入细胞时，抑制 CX 基因的表达以及表达该 CX 基因的细胞的生长，所述 CX 基因选自 C14orf78、MYBL2、UBE2S 和 UBE2T，该分子包含有义链和与之互补的反义链，所述有义链和反义链彼此杂交形成双链核酸分子，并且该分离的双链核酸分子以选自 SEQ ID NO:47 ～ 57 中的序列为靶标；

[0146] [2][1] 的分离的双链核酸分子，其中，上述有义链包含与选自 SEQ IDNO:47 ～ 57 中的靶序列对应的序列；

[0147] [3][2] 的双链核酸分子，其长度小于约 100 个核苷酸；

[0148] [4][3] 的双链核酸分子，其长度小于约 75 个核苷酸；

[0149] [5][4] 的双链核酸分子，其长度小于约 50 个核苷酸；

[0150] [6][5] 的双链核酸分子，其长度小于约 25 个核苷酸；

[0151] [7][6] 的双链核酸分子，其长度为约 19 ～约 25 个核苷酸；

[0152] [8][1] 的双链核酸分子，其由单一多核苷酸构成，所述多核苷酸包含通过间插单链连接在一起的有义链和反义链；

[0153] [9][8] 的双链核酸分子，其具有通式 5′ -[A]-[B]-[A′]-3′ ，其中，[A] 为包含与选自 SEQ ID NO :47 ～ 57 中的靶序列对应的序列的有义链，[B] 为由 3 个～ 23 个核苷酸构成的间插单链，[A′] 为包含 [A] 的互补序列的反义链；

[0154] [10][1] 的双链核酸分子，其包含 RNA；

[0155] [11][1] 的双链核酸分子，其包含 DNA 和 RNA；

[0156] [12][11] 的双链核酸分子，其是 DNA 多核苷酸与 RNA 多核苷酸的杂合体；

[0157] [13][12] 的双链核酸分子，其中，有义链和反义链分别由 DNA 和 RNA 构成；

[0158] [14][11] 的双链核酸分子，其是 DNA 与 RNA 的嵌合体；

[0159] [15][14] 的双链核酸分子，其中，有义链中靶序列或其互补序列的 5′ 末端的一个侧翼区域和 / 或反义链中靶序列或互补序列的 3′ 末端的一个侧翼区域是 RNA 构成的；

[0160] [16][15] 的双链核酸分子，其中所述侧翼区由 9 ～ 13 个核苷酸构成 ;和

[0161] [17][1] 的双链核酸分子，其包含 3′ 突出端。

[0162] 本发明的双链核酸分子将在下文中详细描述。

[0163] 设计具有抑制细胞内靶基因表达能力的双链核酸分子的方法是已知的（见例如，美国专利 No.6,506,559，本文引用其全部内容作为参考）。例如，可以从 Ambion 网站 (http://www.ambion.com/techlib/misc/siRNA_finder.html) 获得用于设计 siRNA 的计算机程序。该计算机程序可以根据如下的方案选择双链核酸分子的靶核苷酸序列。

[0164] 靶点的选择

[0165] 1. 从转录物的 AUG 起始密码子开始向下游扫描搜寻 AA 双核苷酸序列。记录每个 AA 的出现及其 3’ 侧相邻的 19 个核苷酸作为潜在 siRNA 靶点。Tuschl 等建议避免针对 5’ 和 3’ 非翻译区 (UTR) 以及邻近起始密码子的区域 (75 个碱基之内) 设计 siRNA，因为这些区域可能更富含调节蛋白质的结合位点，而 UTR 结合蛋白和 / 或翻译起始复合物可能干扰 siRNA 核酸内切酶复合物的结合。

[0166] 2. 将潜在靶点与合适的基因组数据库（人、小鼠、大鼠等）进行比较，将任何与其他编码序列显著同源的靶序列排除在考虑之外。主要使用 BLAST(Altschul SF 等，Nucleic Acids Res 1997 Sep 1,25(17) :3389-402)，其可见于 NCBI 服务器：!! HYPERLINK " http://www.ncbi.nlm.nih.gov/BLAST/ " ¶ www.ncbi.nlm.nih.gov/BLAST/ ⊥。

[0167] 3. 选择合格的靶序列用于合成。通常沿着待评估的基因的长度选择几个靶序列。

[0168] 根据该规程，本发明的分离的双链核酸分子的靶序列如下地设计。

[0169] 对于 C14orf78 基因，SEQ ID NO :1 的核苷酸

[0170] 13846-13864(SEQ ID NO :47)、

[0171] 13909-13927(SEQ ID NO:48)、

[0172] 14001-14019(SEQ ID NO:49)、和

[0173] 14647-14665(SEQ ID NO:50)；

[0174] 对于 MYBL2 基因，SEQ ID NO:3 的核苷酸

[0175] 977-995(SEQ ID NO:51)、

[0176] 1938-1956(SEQ ID NO:52)、

[0177] 1940-1958(SEQ ID NO:53)、和

[0178] 1995-2013(SEQ ID NO:54)；

[0179] 对于 UBE2S 基因，SEQ ID NO:5 的核苷酸

[0180] 706-724(SEQ ID NO:55) 和

[0181] 528-546(SEQ ID NO:56)；以及

[0182] 对于 UBE2T 基因，SEQ ID NO:7 的核苷酸

[0183] 148-166(SEQ ID NO:57)。

[0184] 对于以上述的靶序列作为靶标的双链核酸分子，分别研究了其抑制表达靶基因的细胞的生长的能力。因此，本发明提供以选自下组中的任何序列作为靶标的双链核酸分子。

[0185] 对于 C14orf78 基因：SEQ ID NO:1 的核苷酸

[0186] 13846-13864(SEQ ID NO:47)、

[0187] 13909-13927(SEQ ID NO:48)、

[0188] 14001-14019(SEQ ID NO:49)、和

[0189] 14647-14665(SEQ ID NO:50)；

[0190] 对于 MYBL2 基因，SEQ ID NO:3 的核苷酸

[0191] 977-995(SEQ ID NO:51)、

[0192] 1938-1956(SEQ ID NO:52)、

[0193] 1940-1958(SEQ ID NO:53)、和

[0194] 1995-2013(SEQ ID NO:54)；

[0195] 对于 UBE2S 基因，SEQ ID NO:5 的核苷酸

[0196] 706-724(SEQ ID NO:55) 和

[0197] 528-546(SEQ ID NO:56)；以及

[0198] 对于 UBE2T 基因，SEQ ID NO:7 的核苷酸

[0199] 148-166(SEQ ID NO:57)。

[0200] 本发明的双链核酸分子针对单个靶 CX 基因序列，或者可以针对多个靶 CX 基因序列。

[0201] 以上述的 CX 基因靶序列中之一为靶标的本发明的双链核酸分子，包括这样的分离的多核苷酸，它们包含与靶序列的核酸序列对应的序列和 / 或靶序列的互补序列。例如，以上述靶序列为靶标的双链核酸分子包含与靶序列及其互补体对应的核苷酸序列。在本发明中，当双链核酸分子包含 RNA、或者由 RNA 构成时，靶序列中核苷酸 t(胸腺嘧啶) 被替换成 u(尿嘧啶)。以 C14orf78 基因为靶标的寡核苷酸的例子包括这样的寡核苷酸，它们包含与 SEQ ID NO:1 的 13846-13864(SEQ ID NO:47)、13909-13927(SEQ IDNO:48)、14001-14019(SEQ ID NO:49) 或 14647-14665(SEQ ID NO:50) 的核苷酸序列对应的序列和

这些核苷酸的互补序列；以 MYBL2 基因为靶标的多核苷酸包括这样的多核苷酸，它们包含与 SEQ ID NO：3 的 977-995(SEQ ID NO：51)、1938-1956(SEQ ID NO：52)、1940-1958(SEQ ID NO：53) 或 1995-2013(SEQ ID NO：54) 的核苷酸序列对应的序列和这些核苷酸的互补序列；以 UBE2S 基因为靶标的多核苷酸包括这样的多核苷酸，所述多核苷酸包含与 SEQ ID NO：5 的 706-724(SEQ ID NO：55) 或 528-546(SEQ IDNO：56) 的核苷酸序列对应的序列和这些核苷酸的互补序列；以 UBE2T 基因为靶标的多核苷酸包括这样的多核苷酸，它们包含与 SEQ ID NO：7 的 148-166(SEQ ID NO：57) 的核苷酸序列相当的序列和这些核苷酸的互补序列。但是，本发明不限于这些实例，上述的核酸序列中的轻微修饰是可以接受的，只要经过修饰的分子能够保持抑制 CX 基因的表达的能力即可。在本说明书中，核酸序列中的“轻微修饰”表示对该序列进行一个、两个或数个核酸的取代、缺失、添加或插入。

[0202] 根据本发明，对于本发明的双链核酸分子，可以采用实施例中采用的方法来测试其能力。在实施例中，对于包含 CX 基因 mRNA 的各个部分的有义链或与之互补的反义链的双链核酸分子，采用常规方法在体外研究了它们减少癌细胞中 CX 基因产物的产生的能力（例如使用胰腺癌细胞 PK-1 细胞系和 Panc.02.03 细胞系、肺癌细胞 H358 细胞系和 A549 细胞系、食道癌细胞 TE-9 细胞系、乳腺癌细胞 MCF-7 细胞系、膀胱癌细胞 SW780 细胞系以及前列腺癌细胞 DU145 细胞系）。此外，例如，与在不存在候选分子的条件下培养出的细胞相比较，与候选双链核酸分子接触的细胞中 CX 基因产物的减少，可以通过例如使用针对 CX 蛋白质的抗体的 western 印迹分析或实施例 1 的项目“半定量的 RT-PCR”中所述的使用针对 CX mRNA 的引物的 RT-PCR 来进行检测。对于在体外细胞测定中减少 CX 基因产物的产生的序列，可以进一步研究其对细胞生长的抑制效果。此外，对于在体外细胞测定中抑制细胞生长的序列，可以使用患有癌症的动物，例如裸小鼠异种移植物模型来研究在体内的能力，以确认 CX 产物产生的减少和癌细胞生长的降低。

[0203] 当分离的多核苷酸为 RNA 或其衍生物时，应在核苷酸序列中将碱基“t”替换为“u”。本说明书中使用的术语“互补”是指多核苷酸的核苷酸单位之间形成 Watson-Crick 或 Hoogsteen 碱基配对；而术语“结合”是指两个多核苷酸之间的物理或化学相互作用。当多核苷酸包含修饰的核苷酸和 / 或非磷酸二酯键连接时，这些多核苷酸也可以相同的方式彼此结合。通常，互补的多核苷酸序列在适宜条件下杂交，形成基本上不含或完全不含错配的稳定双链体。此外，本发明的分离的多核苷酸的有义链和反义链可以通过杂交形成双链核酸分子或发夹环结构。在优选的实施方案中，这样的双链体的平均每 10 个匹配中含有的错配不超过 1 个。在特别优选的实施方案中，双链体的链完全互补，这样的双链体不含错配。

[0204] 对于 C14orf78，多核苷酸长度小于 15958 个核苷酸，对于 MYBL2，多核苷酸长度小于 2731 个核苷酸，对于 UBE2S，多核苷酸长度小于 1207 个核苷酸，或者，对于 UBE2T，多核苷酸长度小于 927 个核苷酸。例如，对于所有的所述基因，多核苷酸的长度小于 500、200、100、75、50 或 25 个核苷酸。本发明的分离的多核苷酸可用于形成针对 CX 基因的双链核酸分子或者制备编码双链核酸分子的模板 DNA。当该多核苷酸用于形成双链核酸分子时，该多核苷酸可长于 19 个核苷酸，优选 21 个核苷酸，更优选长度为约 19 个～ 25 个核苷酸之间。

[0205] 本发明的双链核酸分子可以包含一个或多个修饰的核苷酸和 / 或非磷酸二酯键连接。本领域熟知的化学修饰能够增加所述双链核酸分子的稳定性、利用度 (availability) 和 / 或细胞对该双链核酸分子的摄取。本领域技术人员将知道本分子中

可以引入的其它类型的化学修饰 (WO03/070744 ;WO2005/045037)。在一个实施方案中,可以进行修饰来提供更好的抗降解性或更好的摄取。这些修饰的实例包括硫代磷酸酯键连接、2'-O- 甲基核糖核苷酸 (特别是在双链核酸分子的有义链上)、2'- 脱氧 - 氟代核糖核苷酸、2'- 脱氧 - 核糖核苷酸、"通用碱基"核苷酸、5'-C- 甲基核苷酸和倒位脱氧脱碱基残基的引入 (inverted deoxyabasic residue incorporation)(US20060122137)。在另一个实施方案中,可以使用修饰来增强双链核酸分子的稳定性或增加寻靶效率。修饰包括双链核酸分子两条互补链之间的化学交联、双链核酸分子一条链的 3' 或 5' 末端的化学修饰、糖修饰、核碱基修饰和 / 或骨架修饰、2- 氟代修饰的核糖核苷酸和 2'- 脱氧核糖核苷酸 (WO2004/029212)。在另一个实施方案中,修饰可以用于增加或减少针对靶 mRNA 中和 / 或互补双链核酸分子链中互补核苷酸的亲和力 (WO2005/044976)。例如,未修饰的嘧啶核苷酸可以用 2- 硫、5- 炔基 (5-alkynyl)、5- 甲基或 5- 丙炔基 (5-propynyl) 嘧啶取代。另外,未修饰的嘌呤可以用 7- 脱氮 (7-deza)、7- 烷基或 7- 烯基嘌呤取代。在另一个实施方案中,当双链核酸分子是具有 3' 突出端的双链核酸分子时,可以把 3'- 末端核苷酸的突出核苷酸替换成脱氧核糖核苷酸 (Elbashir SM 等,Genes Dev 2001 Jan 15,15(2) :188-200)。关于进一步的细节,可以利用公开文献如 US20060234970。本发明不限于这些实例,可以将任何已知化学修饰应用于本发明的双链核酸分子,只要所得分子保留抑制靶基因表达的能力。

[0206] 此外,本发明的双链核酸分子可以包含 DNA 和 RNA 二者,例如 dsD/R-NA 或 shD/R-NA。具体而言,由 DNA 链与 RNA 链形成的杂合多核苷酸或 DNA-RNA 嵌合体多核苷酸表现出提高的稳定性。可以形成 DNA 和 RNA 的混合,即由 DNA 链 (多核苷酸) 和 RNA 链 (多核苷酸) 组成的杂合型双链核酸分子、或在任一单链 (多核苷酸) 或两条单链 (多核苷酸) 上同时包含 DNA 和 RNA 的嵌合型双链核酸分子,诸如此类,来增强所述双链核酸分子的稳定性。DNA 链和 RNA 链的杂合体可以是这样的杂合体,其中有义链是 DNA 而反义链是 RNA,或者相反,只要它在导入表达靶基因的细胞中后具有抑制靶基因表达的活性。优选地,有义链多核苷酸是 DNA 而反义链多核苷酸是 RNA。同样,嵌合型双链核酸分子可以是其中的有义链和反义链均由 DNA 和 RNA 组成,或者有义链或反义链中的任一条由 DNA 和 RNA 组成,只要在导入表达靶基因的细胞中时,所述双链核酸分子具有抑制该基因表达的活性即可。为了提高双链核酸分子的稳定性,分子中优选包含尽可能多的 DNA ;而为了诱导靶基因的表达抑制,要求分子在一定范围内是 RNA,以充分地诱导表达抑制。作为嵌合体双链核酸分子的一个优选实例,双链核酸分子的上游部分区域 (即位于有义链或反义链内的靶序列或其互补序列之侧翼的区域) 为 RNA。优选地,所述上游部分区表示有义链的 5' 侧 (5' 端) 和反义链的 3' 侧 (3' 端)。

[0207] 就是说,在某些实施方案中,反义链 3' 末端侧翼的区域由 RNA 组成,或者有义链 5' 末端侧翼的区域和反义链 3' 末端侧翼的区域均由 RNA 组成。例如,本发明的嵌合型或杂合型双链核酸分子包含以下组合。

[0208] 有义链 :

[0209] 5'-[---DNA---]-3'

[0210] 3'-(RNA)-[DNA]-5'

[0211] :反义链 ;

[0212] 有义链 :

[0213] 5’-(RNA)-[DNA]-3’

[0214] 3’-(RNA)-[DNA]-5’

[0215] ;反义链;以及

[0216] 有义链：

[0217] 5’-(RNA)-[DNA]-3’

[0218] 3’-(---RNA---)-5’

[0219] ;反义链。

[0220] 上游部分区优选是由 9-13 个核苷酸组成的域，从双链核酸分子的有义链或反义链内的靶序列或其互补序列的末端算起。此外，这种嵌合型双链核酸分子的优选实例包括这样的实例：链长为 19-21 个核苷酸，其中多核苷酸的至少上游的半区（对于有义链为 5’侧区域而对于反义链为 3’侧区域）是 RNA，而另一半是 DNA。在这样的嵌合型双链核酸分子中，抑制靶基因表达的效果比反义链整体为 RNA 时强得多 (US20050004064)。

[0221] 在本发明中，双链核酸分子可以形成发夹结构，例如短发夹 RNA(shRNA) 以及由 DNA 与 RNA 组成的短发夹 (shD/R-NA)。shRNA 或 shD/R-NA 是 RNA 序列或 RNA 和 DNA 的混合序列，其形成紧密的发夹转弯，可以用来通过 RNA 干扰来沉默基因表达。shRNA 或 shD/R-NA 在单一链上包含有义靶标序列和反义靶标序列，其中所述序列被环序列隔开。通常，发夹结构被细胞机制切割成 dsRNA 或 dsD/R-NA，所述 dsRNA 或 dsD/R-NA 进一步与 RNA 诱导的沉默复合物 (RNA-induced silencing complex，RISC) 结合。这种复合物结合并切割与所述 dsRNA 或 dsD/R-NA 的靶标序列匹配的 mRNA。

[0222] 为了形成发夹环结构，可以在有义序列与反义序列之间设置由任意核苷酸序列构成的环序列。因此，本发明还提供具有通式 5’-[A]-[B]-[A’]-3’ 的双链核酸分子。式中，[A] 为有义链，包含与靶序列对应的序列；[B] 为间插单链；[A′] 为反义链，包含 [A] 的互补序列。靶序列例如可以选自以下的组。

[0223] 对于 C14orf78，SEQ ID NO:1 的核苷酸

[0224] 13846-13864(SEQ ID NO:47)、

[0225] 13909-13927(SEQ ID NO:48)、

[0226] 14001-14019(SEQ ID NO:49)、或

[0227] 14647-14665(SEQ ID NO:50)；

[0228] 对于 MYBL2，SEQ ID NO:3 的核苷酸

[0229] 977-995(SEQ ID NO:51)、

[0230] 1938-1956(SEQ ID NO:52)、

[0231] 1940-1958(SEQ ID NO:53)、或

[0232] 1995-2013(SEQ ID NO:54)；

[0233] 对于 UBE2S，SEQ ID NO:5 的核苷酸

[0234] 706-724(SEQ ID NO:55)、或

[0235] 528-546(SEQ ID NO:56)；

[0236] UBE2T 的 SEQ ID NO:7 的核苷酸

[0237] 148-166(SEQ ID NO:57)。

[0238] 本发明不限于这些实例，在双链核酸分子保持抑制作为靶标的 CX 基因的表达的

能力的前提下，[A] 中的靶序列可以是在这些实例的序列的基础上修饰而得到的序列。区域 [A] 与 [A′] 杂交而形成由区域 [B] 构成的环。间插单链部分 [B]、即环序列，长度优选可以为 3 ～ 23 个核苷酸。例如，环序列可以选自由以下的序列组成的组 (http://www.ambion.com/techlib/tb/tb_506.html)。此外，由 23 个核苷酸组成的环序列也提供活性 siRNA(Jacque，J.-M.，等人，(2002)Nature 418 :435-8.) :

[0239] CCC、CCACC、或 CCACACC :Jacque JM 等，Nature 2002 Jul 25,418(6896) :435-8，Epub 2002 Jun 26 ;

[0240] UUCG :Lee NS 等，Nat Biotechnol 2002 May,20(5) :500-5 ;FruscoloniP 等，Proc Natl Acad Sci USA 2003 Feb 18,100(4) :1639-44，Epub 2003 Feb10 ;和

[0241] UUCAAGAGA :Dykxhoorn DM 等，Nat Rev Mol Cell Biol 2003 Jun,4(6) :457-67.

[0242] 代表性的、优选的具有本发明的发夹环结构的双链核酸分子如下所示。在下面的结构中，环序列可以选自由 AUG、CCC、UUCG、CCACC、CTCGAG、AAGCUU、CCACACC 和 UUCAAGAGA 组成的组，但本发明不限于此。

[0243] gauaugccaucccagauuu-[B]-aaaucugggauggcauauc (靶序列 SEQ ID NO :47) ;

[0244] gucaaauuccccaaauuaa-[B]-uuaauuuggggaauuugac (靶序列 SEQ ID NO :48) ;

[0245] guguccagaggccaauauu-[B]-aauauuggccucuggacac (靶序列 SEQ ID NO :49) ;

[0246] ggcagggcuccaaaagaca-[B]-ugucuuuuggagcccugcc (靶序列 SEQ ID NO :50) ;

[0247] ggagcccaucgguacagau-[B]-aucuguaccgaugggcucc (靶序列 SEQ ID NO :51) ;

[0248] cggcggagccccaucaaga-[B]-ucuugaugggcuccgccg (靶序列 SEQ ID NO :52) ;

[0249] gcggagccccaucaagaaa-[B]-uuucuugaugggcuccgc (靶序列 SEQ ID NO :53) ;

[0250] gaugugaagcugaugaugu-[B]-acaucaucagcuucacauc (靶序列 SEQ ID NO :54) ;

[0251] ugcugaccaucaagugccu-[B]-aggcacuugauggucagca (靶序列 SEQ ID NO :55) ;

[0252] ccauaugcuggaggucugu-[B]-acagaccuccagcauaugg (靶序列 SEQ ID NO :56) ;和

[0253] agagagagcugcacauguu-[B]-aacaugugcagcucucucu (靶序列 SEQ ID NO :57)。

[0254] 此外，为了增强双链核酸分子的抑制活性，可以将核苷酸“u”添加到靶序列的反义链的 3’ 末端，作为 3’ 突出端。添加的“u”的数目为至少 2 个，通常为 2-10 个，优选为 2-5 个。添加的“u”在双链核酸分子的反义链的 3’ 末端形成单链。

[0255] 对于双链核酸分子的制备方法没有特殊限制，但优选采用本技术领域公知的化学合成方法。根据化学合成方法，分别合成有义和反义单链多核苷酸，然后采用适当的方法使它们相互退火，来获得双链核酸分子。退火的具体例子包括 :将合成的单链多核苷酸以至少约 3 ： 7、更优选约 4 ： 6、最优选基本上等摩尔量 (即约 5 ： 5 的摩尔比) 的摩尔比混合。然后，将混合物加热到双链核酸分子解离的温度，再缓慢冷却。经退火的双链多核苷酸可以采用本技术领域公知的常用方法来纯化。纯化方法的例子包括 :利用琼脂糖凝胶电泳的方法，或者任选地除去剩余的单链多核苷酸 (例如利用适当的酶进行降解) 的方法。

[0256] 位于 CX 序列们的侧翼的调节序列们可以是相同的或不同的，以使得它们的表达能够独立地，或者以时间性或空间性方式被调控。通过将 CX 基因模板克隆到载体 (例如包含来源于小分子核 RNA(snRNA)U6 或来自人 H1RNA 启动子的 RNA polIII 转录单元的载体) 中，可以在细胞内转录双链核酸分子。

[0257] 载体

[0258] 本发明还包括含有一种或多种本文所述的双链核酸分子的载体、以及包含该载体的细胞。本发明的载体优选以可表达的形式编码本发明的双链核酸分子。在本说明书中，术语“以可表达的形式”，是指该载体当被导入细胞中时，将表达该分子。在优选的实施方案中，载体包含双链核酸分子表达所必需的调节元件。本发明的这种载体可以用于产生本发明的双链核酸分子，也可以直接用作癌症治疗的活性成分。

[0259] 本发明的载体可以通过，例如，将 CX 序列克隆到表达载体中，以使得调节序列与 CX 序列令两条链均可表达（通过 DNA 分子转录）的方式可操作连接，来加以制备 (Lee NS 等，Nat Biotechnol 2002 May，20(5)：500-5)。例如，与 mRNA 反义的 RNA 分子由第一启动子（例如位于克隆 DNA 的 3’末端侧翼的启动子序列）转录，对 mRNA 而言为有义链 RNA 分子由第二启动子（例如位于克隆 DNA 的 5’末端侧翼的启动子序列）转录。有义和反义链在体内杂交，产生用于沉默该基因的双链核酸分子构建体。或者，使用分别编码双链核酸分子之有义链和反义链的两个载体构建体来分别表达有义链和反义链，随后形成双链核酸分子构建体。而且，克隆的序列可编码具有二级结构（例如发夹）的构建体，即，载体的单一转录物同时包含靶基因的有义序列和互补反义序列二者。

[0260] 也可以配置本发明的载体使得其实现在靶细胞的基因组中的稳定插入（关于同源重组盒载体的说明，参见 Thomas KR & Capecchi MR，Cell 1987，51：503-12)。可以参考例如：Wolff 等，Science 1990，247：1465-8；美国专利第 5,580,859 号、第 5,589,466 号、第 5,804,566 号、第 5,739,118 号、第 5,736,524 号、第 5,679,647 号和 WO 98/04720。基于 DNA 的输送技术的例子包括：“裸 DNA”、辅助（布比卡因 (bupivicaine)、聚合物、肽介导型）输送、阳离子脂质复合体、和粒子介导型输送（“基因枪”）或压力介导型输送（例如参照美国专利第 5,922,687 号）。

[0261] 本发明的载体还可以是例如病毒性载体或细菌性载体。表达载体的例子包括牛痘或鸡痘等减毒病毒宿主（参照美国专利第 4,722,848 号）。该策略包含例如使用牛痘病毒作为载体来表达编码双链核酸分子的核苷酸序列。重组牛痘病毒在被导入表达靶基因的细胞时即表达该分子并由此抑制细胞的增殖。可使用的载体的其它例子包括卡介苗 (BCG)。BCG 载体在 Stover 等，Nature 1991，351：456-60 中有记载。多种多样的其它载体可用于双链核酸分子的治疗性施用与生产，例子包括腺病毒载体和腺随伴病毒载体、逆转录病毒载体、伤寒沙门氏菌 (Salmonella typhi) 载体、脱毒的炭疽毒素载体等。可以参照例如 Shata 等，Mol Med Today 2000，6：66-71；Shedlock 等，J Leukoc Biol 2000，68：793-806；以及 Hipp 等，In Vivo 2000，14：571-85。

[0262] 癌症的治疗方法

[0263] 在本发明中，构建了针对 C14orf78 的 4 种不同的 dsRNA、针对 MYBL2 的 4 种不同的 dsRNA、针对 UBE2S 的 2 种不同的 dsRNA 以及针对 UBE2T 的 1 种 dsRNA，来研究它们抑制细胞生长的能力。针对 C14orf78 的所有 4 种 dsRNA 在 PK-1 和 Panc.02.03 等表达该基因的细胞中均有效地敲低了该基因表达，与细胞增殖的抑制相一致（图 2a、b)，另一方面，这些 dsRNA 在不表达 C14orf78 的细胞系 SK-BR-3 中未见显著变化（图 2c)。针对 MYBL2 的所有 4 种 dsRNA 在 NSCLC 细胞系 (H358) 和食道癌细胞系 (TE-9) 等表达该基因的细胞中均显著地减少了其表达水平和细胞生长活性（图 3a、b)，而在不表达 MYBL2 的细胞系正常小气道上皮细胞 (SAEC) 中，没有可检出的生长抑制（图 3c)。针对 UBE2S 的 2 种 dsRNA 在例如乳

腺癌细胞系(MCF7)、胰腺癌细胞系(PK-1)和膀胱癌(SW780)细胞系等表达该基因的细胞中均显著地减少了其表达水平和细胞存活力(图4a-c);针对UBE2T的1种dsRNA在例如乳腺癌(MCF7)、NSCLC(A549)、膀胱癌(SW780)和前列腺癌(DU-145)细胞系等表达该基因的细胞中,有效地抑制了该基因的表达(图5a-d),而在不表达UBE2S和UBE2T这两者的细胞系HMEC(正常乳房上皮细胞)中,没有可检测出的生长抑制(图4d、5e)。由此可见,使用针对各CX基因的所有dsRNA进行的处理在体内均有效地抑制了癌症的进展(图6a和b)。

[0264] 本发明的双链核酸分子和载体的这种抑制癌细胞的细胞增殖的能力,提示它们可以在癌症治疗方法中使用。因此,本发明提供了治疗以CX基因过表达为特征的癌症患者的方法,其是通过施用针对CX基因的双链核酸分子或表达该分子的载体来实施的。

[0265] 实际上,已经确认,与对应的正常组织相比较,CX基因在癌组织中过表达。例如,在临床样品中,C14orf78在18例胰腺癌中的11例、25例胆管细胞癌中的14例和37例非小细胞肺癌中的10例中过表达(T/N比>=5);MYBL2被发现在很多种癌症中过表达:在18例胰腺癌中的6例、34例临床膀胱癌中的18例、64例食道癌中的29例、37例非小细胞肺癌(NSCLC)中的18例和15例小细胞肺癌(SCLC)中的14例中上调(比例>=5);UBE2S在临床样品中过表达:在全部SCLC病例、34例膀胱癌中的29例、81例乳腺癌中的27例、25例胆管细胞癌中的9例、59例前列腺癌中的18例、48例结肠癌中的11例和18例胰腺癌中的12例中过表达。而且,一种与UBE2S类似的蛋白质:泛素E2连接酶样UBE2T基因,也在多种类型的癌中显示表达增大,即,25例胆管细胞癌中的12例、15例SCLC中的12例、34例膀胱癌中的23例、81例乳腺癌中的28例、37例NSCLC中的13例、64例食道癌中的14例和59例前列腺癌中的15例(表2)。

[0266] 在本发明中鉴定出了多种CX基因,通过抑制它们的表达水平,能够诱导出抑制细胞生长或细胞增殖的效果。通过抑制这些基因的表达,可以抑制表达该基因的细胞的细胞生长。有报告称,本发明的CX基因在以下的数种癌症中上调。

[0267] C14orf78

[0268] 胰腺癌(WO2004/31412)

[0269] MYBL2

[0270] 膀胱癌(WO2006/085684)

[0271] 食道癌(WO2007/013671)

[0272] NSCLC(WO2004/031413)

[0273] 胰腺癌(WO2004/31412)

[0274] SCLC(WO2007/013665)

[0275] 睾丸精原细胞瘤(WO2004/031410)

[0276] UBE2S

[0277] 膀胱癌(WO2006/085684)

[0278] 乳腺癌(WO2005/028676)

[0279] 胰腺癌(WO2004/31412)

[0280] 前列腺癌(WO2004/031414)

[0281] SCLC(WO2007/013665)

[0282] UBE2T

[0283] 膀胱癌 (WO2006/085684)

[0284] 乳腺癌 (WO2005/028676)

[0285] 食道癌 (WO2007/013671)

[0286] NSCLC(WO2004/031413)

[0287] SCLC(WO2007/031413)

[0288] 因此，在优选的实施方案中，本发明提供通过抑制选自 C14orf78、MYB2L、UBE2S 和 UBE2T 中的 CX 基因来治疗或预防这些癌症的方法。

[0289] 例如，本发明提供治疗选自胰腺癌、胆管细胞癌和非小细胞肺癌中的癌症的方法，该方法包括施用至少一种分离的双链核酸分子的步骤，所述双链核酸分子包含有义链和与之互补的反义链，二者相互杂交而形成该双链核酸分子，并且其中该有义链包含与选自 SEQ ID No :47 ～ 50（就 C14orf78 而言）中的靶序列对应的序列。

[0290] 本发明还提供治疗选自胰腺癌、非小细胞肺癌、小细胞肺癌、膀胱癌、食道癌和睾丸精原细胞瘤中的癌症的方法，该方法包括施用至少一种分离的双链核酸分子的步骤，所述双链核酸分子包含有义链和与之互补的反义链，二者相互杂交而形成该双链核酸分子，并且其中该有义链包含与选自 SEQ ID No :51 ～ 54（就 MYBL2 而言）中的靶序列对应的序列。

[0291] 另一方面，本发明还提供治疗选自胰腺癌、乳腺癌、小细胞肺癌、膀胱癌、胆管细胞癌、结肠癌和前列腺癌中的癌症的方法，该方法包括施用至少一种分离的双链核酸分子的步骤，所述双链核酸分子包含有义链和与之互补的反义链，二者相互杂交而形成该双链核酸分子，并且其中该有义链包含选自 SEQ ID No :55 ～ 56（就 UBE2S 而言）中的靶序列。

[0292] 而且，本发明还提供治疗选自乳腺癌、非小细胞肺癌、小细胞肺癌、膀胱癌、胆管细胞癌、前列腺癌和食道癌中的癌症的方法，该方法包括施用至少一种分离的双链核酸分子的步骤，所述双链核酸分子包含有义链和与之互补的反义链，二者相互杂交而形成该双链核酸分子，并且其中该有义链包含靶序列 SEQ ID No :57（就 UBE2T 而言）。

[0293] 因此，本发明的方法可以用于抑制罹患或有风险患上 CX 基因相关疾病例如胰腺癌、非小细胞肺癌、小细胞肺癌、乳腺癌、膀胱癌、食道癌、前列腺癌、结肠癌和 / 或胆管细胞癌等的患者中的 CX 基因表达。优选地，针对 C14orf78 的双链核酸分子与表达它们的载体可以在胰腺癌、胆管细胞癌和 / 或非小细胞肺癌的治疗中使用，针对 MYBL2 的双链核酸分子与表达它们的载体可以在膀胱癌、食道癌、睾丸精原细胞瘤、非小细胞肺癌、胰腺癌和 / 或小细胞肺癌的治疗中使用，针对 UBE2S 的双链核酸分子与表达它们的载体可以在小细胞肺癌、膀胱癌、乳腺癌、胆管细胞癌、前列腺癌、结肠癌和 / 或胰腺癌的治疗中使用，针对 UBE2T 的双链核酸分子与表达它们的载体可以在胆管细胞癌、非小细胞肺癌、小细胞肺癌、膀胱癌、乳腺癌、食道癌和 / 或前列腺癌的治疗中使用。

[0294] 具体地，本发明提供以下 [1] ～ [29] 的方法。

[0295] [1] 用于治疗癌症的方法，该方法包括在过表达 CX 基因的细胞中施用至少一种抑制该基因表达的分离的双链核酸分子的步骤；其中所述 CX 基因选自 C14orf78、MYBL2、UBE2S 和 UBE2T，所述分子包含有义链及与之互补的反义链，二者相互杂交而形成所述双链核酸分子，且所述双链核酸分子以选自 SEQ ID NO :47 ～ 57 中的序列为靶标；

[0296] [2][1] 的方法，其中，所述有义链包含与选自 SEQ ID NO :47 ～ 57 中的靶序列对

应的序列；

[0297] [3][1] 的方法，其中，所述细胞为癌细胞；

[0298] [4][1] 的方法，其中，所治疗的癌症选自胰腺癌、肺癌、乳腺癌、膀胱癌、食道癌、睾丸精原细胞瘤、前列腺癌、结肠癌或胆管细胞癌；

[0299] [5][4] 的方法，其中，所述肺癌是非小细胞肺癌或小细胞肺癌；

[0300] [6][1] 的方法，其中，当选择的 CX 基因是 C14orf78 时，要治疗的癌症选自胰腺癌、胆管细胞癌或非小细胞肺癌；

[0301] [7][1] 的方法，其中，当选择的 CX 基因是 MYBL2 时，要治疗的癌是胰腺癌、非小细胞肺癌、小细胞肺癌、膀胱癌、食道癌或睾丸精原细胞瘤；

[0302] [8][1] 的方法，其中，当选择的 CX 基因是 UBE2S 时，所治疗的癌选自胰腺癌、乳腺癌、小细胞肺癌、膀胱癌、胆管细胞癌、前列腺癌或结肠癌；

[0303] [9][1] 的方法，其中，当选择的 CX 基因是 UBE2T 时，所治疗的癌是乳腺癌、非小细胞肺癌、小细胞肺癌、膀胱癌、胆管细胞癌、前列腺癌或食道癌；

[0304] [10][1] 的方法，其中，施用多种双链核酸分子；

[0305] [11][10] 的方法，其中，所述多种双链核酸分子以同一基因为靶标；

[0306] [12][2] 的方法，其中，所述双链核酸分子具有小于约 100 个核苷酸的长度。

[0307] [13][12] 的方法，其中，双链核酸分子具有小于 75 个核苷酸的长度；

[0308] [14][13] 的方法，其中，双链核酸分子具有小于约 50 个核苷酸的长度；

[0309] [15][14] 的方法，其中，双链核酸分子具有小于约 25 个核苷酸的长度；

[0310] [16][15] 的方法，其中，双链核酸分子具有约 19 个～约 25 个核苷酸的度；

[0311] [17][1] 的方法，其中，双链核酸分子由单一多核苷酸构成，所述多核苷酸包含通过间插单链连接在一起的有义链和反义链；

[0312] [18][17] 的方法，其中，双链核酸分子具有通式 5′ -[A]-[B]-[A′]-3′ ，其中 [A] 为有义链，包含与选自 SEQ ID NO :47 ～ 57 中的靶序列对应的序列；[B] 为由 3 个～ 23 个核苷酸构成的间插单链，[A’] 为反义链，包含与 [A] 互补的序列；

[0313] [19][1] 的方法，其中，双链核酸分子包含 RNA；

[0314] [20][1] 的方法，其中，双链核酸分子包含 DNA 与 RNA 二者；

[0315] [21][20] 的方法，其中，双链核酸分子是 DNA 多核苷酸与 RNA 多核苷酸的杂合体；

[0316] [22][21] 的方法，其中，有义链和反义链多核苷酸分别由 DNA 和 RNA 构成；

[0317] [23][20] 的方法，其中，双链核酸分子为 DNA 与 RNA 的嵌合体；

[0318] [24][23] 的方法，其中，位于有义多核苷酸和 / 或反义多核苷酸的 5’末端侧翼的区域由 RNA 构成；

[0319] [25][24] 的方法，其中，所述侧翼区由 9 ～ 13 个核苷酸构成；

[0320] [26][1] 的方法，其中，双链核酸分子包含 3’突出端；

[0321] [27][1] 的方法，其中，双链核酸分子由载体编码；

[0322] [28][27] 的方法，其中，由载体编码的双链核酸分子具有通式 5′ -[A]-[B]-[A′]-3′ ，其中，[A] 为有义链，包含与选自 SEQ ID NO :47 ～ 57 中的靶序列对应的序列；[B] 为由 3 个～ 23 个核苷酸构成的间插单链，[A’] 为反义链，包含与 [A] 互补的序列；以及

[0323] [29][1] 的方法，双链核酸分子包含于组合物中，该组合物中除了包含该双链核酸分子之外，还饱和转染促进剂和药学可接受的载体。

[0324] 本发明的方法将在下文中具体描述。

[0325] 通过使针对 CX 基因的双链核酸分子、表达上述分子的载体或包含其的组合物与表达 CX 基因的细胞接触，抑制所述细胞的生长。进一步使细胞与转染剂接触。合适的转染剂是本技术领域公知的。短语“抑制细胞生长”意思是与没有暴露于该分子的细胞相比，细胞的增殖速度较低，或存活力下降。细胞生长可以采用本技术领域公知的方法来测定，利用使用细胞分析仪 (Cell Analyzer)1000 和 MTT 细胞增殖试验等。

[0326] 采用本发明的方法可以抑制任何种类的细胞的生长，只要细胞表达或过表达本发明的双链核酸分子的靶基因。代表性的细胞包括癌细胞，更具体地是胰腺癌细胞、肺癌细胞、乳腺癌细胞、膀胱癌细胞、食道癌细胞、前列腺癌细胞、睾丸精原细胞瘤细胞、结肠癌细胞和胆管细胞癌细胞。

[0327] 因此，对于已罹患或有风险发生 C14orf78、MYBL2、UBE2S 或 UBE2T 相关的疾病的患者，可以通过施用至少一种本发明的双链核酸分子、至少一种表达至少一种所述分子的载体、或者至少一种包含至少一种所述分子的组合物来进行治疗。例如，癌症患者，具体地说是胰腺癌、肺癌、乳腺癌、膀胱癌、食道癌、前列腺癌、睾丸精原细胞瘤、结肠癌和 / 或胆管细胞癌患者，可以采用本发明的方法来进行治疗。癌症的类型可以根据诊断的肿瘤的特定类型，采用常规方法来确定。胰腺癌可以通过例如核磁共振成像、计算机轴向断层扫描、超声或活检来诊断。肺癌可以通过例如胸部 X 放射照相、计算机断层扫描、核磁共振成像、支气管镜检、针刺活检或骨扫描来诊断。乳腺癌可以通过例如临床检查、成像程序（乳房 X 射线扫描像、乳腺超声、核磁共振成像等）或活检来诊断。膀胱癌可以通过例如 NMP22（注册商标）、BladderChek（注册商标）、尿液分析、尿细胞学检查或尿培养来诊断。食道癌可以通过例如针穿刺吸引、活检、血液检查或成像检查、食道镜检查来诊断。睾丸精原细胞瘤或前列腺癌可以通过例如直肠指诊、直肠超声波、前列腺特异抗原 (PSA) 和前列腺酸性磷酸酶 (PAP) 测试、肿瘤活检或骨扫描来诊断。胆管细胞癌可以通过例如肝肥大、断层扫描、超声波或活检来诊断。结肠癌可以通过例如血便、结肠镜检查、可屈式乙状结肠镜检查、CEA 测定、双重对比造影钡餐灌肠 CT 扫描、断层扫描或活检来诊断。更优选地，通过 RT-PCR 或免疫测定检测患者来源的活检物中的 CX 基因表达，来选择依照本发明的方法治疗的患者。优选地，在实施本发明的治疗前，通过例如免疫化学的分析或 RT-PCR 等本技术领域公知的方法来确认受试者来源的活检组织中 CX 基因的过表达。

[0328] 根据本发明的抑制细胞生长并由此治疗癌症的方法，在施用多种双链核酸分子（或者表达它们的载体或包含它们的组合物）时，各分子可以靶向同一靶序列，或者靶向同一 CX 基因内或不同 CX 基因中的不同靶序列。例如，上述方法可以利用针对 1、2、3 或 4 个 CX 基因的多种双链核酸分子。或者，例如，上述方法可以利用针对同一 CX 基因内的 1、2、3、4、5 个或更多靶序列的多种双链核酸分子。

[0329] 为了抑制细胞生长，可以将本发明的双链核酸分子以能够实现与对应的 mRNA 转录物结合的形式直接导入细胞。或者，如上文所述，可将编码双链核酸分子的 DNA 作为载体导入细胞。为了将双链核酸分子和载体导入细胞，可以使用 FuGENE Roche diagnostics)、Lipofectamine 2000(Invitrogen)、Oligofectamine(Invitrogen) 和 Nucleofector(Wako

pure Chemical）等转染促进剂。

[0330] 当治疗导致临床效益，诸如受试者中 CX 基因表达的减少、癌症的尺寸的减少、或癌症的扩张、患病率 (prevalence)、转移潜力的抑制等，可以判定该治疗是有效的。当预防性地适用治疗时，“有效”是指其延迟或防止癌症形成，或者预防或缓解癌症的临床症状。有效性可以结合特定肿瘤类型的任何公知的诊断或治疗方法来加以确定。

[0331] 我们认为，本发明的双链核酸分子以半化学计量式的量来降解 (C14orf78、MYBL2、UBE2S 或 UBE2T 的）靶 mRNA。不希望受限于任何理论，我们认为本发明的双链核酸分子像催化剂那样导致靶 mRNA 的降解。因此，与常规的癌症疗法相比，为实施治疗效应而需要输送到癌症的位置或其附近的双链核酸分子要少的多。

[0332] 本领域技术人员在考虑了受试者的体重、年龄、性别、疾病类型、症状及其它条件、施用途径、以及是局部施用还是全身施用等因素的基础上，能够容易地确定本发明的双链核酸分子的有效量。典型地，本发明的双链核酸分子的有效量包括：在癌症部位或其附近约 1 纳摩尔 (nM) ～约 100nM，优选为约 2nM ～约 50nM，更优选为约 2.5nM ～ 10nM 的细胞内浓度。我们考虑也可以施用更多量或更少量的双链核酸分子。

[0333] 本发明的方法可以用于抑制癌症，例如胰腺癌、肺癌、乳腺癌、膀胱癌、食道癌、前列腺癌、睾丸精原细胞瘤、结肠癌和胆管细胞癌的生长或转移。具体地，包含 C14orf78 的靶序列（即 SEQ ID NO：47 ～ 50）的双链核酸分子是胰腺癌、胆管细胞癌和非小细胞肺癌的治疗所特别优选的；包含 MYBL2 的靶序列（即 SEQ ID NO：51 ～ 54）的双链核酸分子是胰腺癌、非小细胞肺癌、小细胞肺癌、膀胱癌、食道癌、和睾丸精原细胞瘤的治疗所特别优选的；包含 UBE2S 的靶序列（即 SEQ ID NO：55 和 56）的双链核酸分子是胰腺癌、乳腺癌、小细胞肺癌、膀胱癌、胆管细胞癌、前列腺癌和结肠癌的治疗所特别优选的；包含 UBE2T 的靶序列（即 SEQ ID NO：55）的双链核酸分子是乳腺癌、胆管细胞癌、非小细胞肺癌、小细胞肺癌、膀胱癌、前列腺癌和食道癌的治疗所特别优选的。

[0334] 为了治疗癌症，还可以将本发明的双链核酸分子与不同于所述双链核酸分子的药剂组合来施用于受试者。或者，本发明的双链核酸分子还可以与其它旨在用于癌症治疗的治疗方法组合来施用于患者。例如，本发明的双链核酸分子还可以与现行使用的治疗癌症或预防癌症转移的治疗方法（例如，放射疗法，外科手术，以及使用化疗剂，如顺铂、卡铂、环磷酰胺、5-氟尿嘧啶、阿霉素、柔红霉素 (daunorubicin) 或他莫昔芬 (tamoxifen) 等的治疗）组合来进行施用。

[0335] 在本发明的方法中，双链核酸分子可以以与投递试剂相组合的裸双链核酸分子的形式，或者以表达双链核酸分子的重组质粒或者病毒载体的形式施用于受试者。

[0336] 用于与本发明的双链核酸分子组合施用的合适的投递试剂包括 MirusTransit TKO 脂溶性试剂、LipoTrust™ SR、Lipofectin、LipofectAMINE、Cellfectin、聚阳离子（例如聚赖氨酸）、脂质体、或胶原；或去端肽胶原。优选的投递试剂是脂质体。

[0337] 脂质体能够帮助将双链核酸分子投递至特定的组织如视网膜或肿瘤组织中，还能够增加双链核酸分子的血中半衰期。适合在本发明中使用的脂质体是由常规的囊泡形成性脂质 (vesile-forming lipids) 形成的，囊泡形成性脂质通常包括中性或带负电荷的磷脂，以及甾醇，诸如胆固醇。对一些因素的考虑通常可以为脂质的选择提供指导，如脂质体大小以及在血流中的脂质体的半衰期等。有多种制备脂质体的方法是公知的，例如 Szoka

等，Ann Rev Biophys Bioeng 1980，9：467；美国专利第 4,235,871 号；第 4,501,728 号；第 4,837,028 号；第 5,019,369 号；它们全部援引并入本说明书。

[0338] 优选地，包被本发明的双链核酸分子的脂质体包含配体分子，它能够将脂质体输送至癌症部位。优选的配体是与肿瘤或血管内皮细胞中常见的受体结合的配体，例如与肿瘤抗原或内皮细胞表面抗原结合的单克隆抗体。

[0339] 特别优选的是，包被本发明的双链核酸分子的脂质体经过了修饰以免被单核巨噬细胞和网状内皮系统清除，例如，其结构的表面结合有调理作用抑制部分 (opsonization inhibition moities)。在一个实施方案中，本发明的脂质体可以同时包括调理作用抑制部分和配体。

[0340] 用于制备本发明的脂质体的调理作用抑制部分通常是与脂质体膜结合的大型亲水性聚合物。如在本说明书中所使用的，例如，当调理作用抑制部分化学地或物理地搭接在膜上时，例如通过脂溶性锚 (anchor) 插入膜本身，或者通过与膜脂质的活性基团直接结合，则调理作用抑制部分与脂质体膜“结合”。这些调理作用抑制性亲水性聚合物形成保护性表层，显著地减少单核巨噬细胞系统（“MMS”）和网状内皮系统（”RES”）对脂质体的摄入；在例如美国专利第 4,920,016 号中对此有记载，后者的全部公开内容援引并入本说明书。因此，与未修饰的脂质体相比，被调理作用抑制部分修饰过的脂质体能够保留在血液循环中的时间显著更长。出于以上理由，上述脂质体有时也被称为“隐形”(stealth) 脂质体。

[0341] 已知隐形脂质体蓄积在依靠多孔性或“渗漏性”微血管系统供给的组织中。因此，在以这样的微血管系统缺损为特征的靶组织，例如实体瘤中，这些脂质体会高效率地蓄积。参见 Gabizon 等，Proc Natl Acad Sci USA 1988，18：6949-53。此外，由于 RES 的摄入的减少，防止隐形脂质体在肝脏和脾脏中的显著蓄积，从而降低隐形脂质体的毒性。因此，用调理作用抑制部分修饰的本发明的脂质体能够将本发明的双链核酸分子输送至肿瘤细胞。

[0342] 适用于修饰脂质体的调理作用抑制部分优选为分子量约 500 ～约 40,000 道尔顿、更优选约 2,000 ～约 20,000 道尔顿的水溶性聚合物。这样的聚合物中包括聚乙二醇 (PEG) 或聚丙二醇 (PPG) 衍生物；例如，甲氧基 PEG 或 PPG、和 PEG 或 PPG 硬脂酸酯；合成聚合物如聚丙烯酰胺或聚 N- 乙烯基吡咯烷酮；直链状、分枝状或者树状聚酰胺胺 (polyamidoamine)；聚丙烯酸；多元醇，诸如化学结合有羧基或氨基的聚乙烯醇和聚木糖醇，以及神经节苷脂，诸如神经节苷脂 GM_1。PEG、甲氧基 PEG、或甲氧基 PPG 或其衍生物的共聚体也是适合的。此外，抑制调理作用的聚合物可以是 PEG 与聚氨基酸、多糖、聚酰胺胺、聚亚乙基胺或者多核苷酸中的任何一种的嵌段共聚物。抑制调理作用的聚合物可以是含有氨基酸或羧酸的天然多糖，例如半乳糖醛酸、葡糖醛酸、甘露糖醛酸、透明质酸、果胶酸、神经氨酸、海藻酸、角叉胶；氨基化多糖类或寡糖（直链状或者分枝状）；或者羧基化的多糖或寡糖，例如，与碳酸的衍生物反应而生成了羧基结合的羧基化多糖类或寡糖。

[0343] 优选地，调理作用抑制部分为 PEG、PPG 或其衍生物。用 PEG 或 PEG 衍生物修饰的脂质体有时称为“PEG 化脂质体”。

[0344] 调理作用抑制部分可以通过许多公知技术中的任何一种结合到脂质体膜上。例如，PEG 的 N- 羟基琥珀酰亚胺酯能够与磷脂酰乙醇胺脂溶性锚 (lipid-soluble anchor) 结合，然后再结合到膜上。相似地，可以通过还原性氨基化，用硬脂酰胺脂溶性锚来将右旋糖酐聚合物衍生化，所述还原性氨基化中使用 $Na(CN)BH_3$ 和混合溶剂，如 60℃的四氢呋喃与水

的 30 ： 12 比例混合物。

[0345] 上文讨论了表达本发明的双链核酸分子的载体。这样的表达至少一种本发明的双链核酸分子的载体也可以直接施用或者与 Mirus Transit LT1 脂溶性试剂、LipoTrust（商标）SR、Lipofectin、LipofectAMINE、Cellfectin、聚阳离子（例如聚赖氨酸）或脂质体、或胶原、去端肽胶原 (Atelocollagen) 等合适的投递试剂组合施用。将表达本发明的双链核酸分子的重组病毒载体输送到患者的癌症区域的方法在本技术领域的技术范围内。

[0346] 本发明的双链核酸分子可以通过适于将双链核酸分子输送到癌症区域的任何手段来施用给受试者。例如，双链核酸分子可以通过基因枪、电穿孔、或者其它合适的非消化道或肠内施用途径来施用。

[0347] 合适的非消化道包括静脉内施用（例如静脉内推注、静脉内输注、动脉内推注、动脉内输注和针对血管网络的导管滴注），组织周围和组织内注射（例如肿瘤周围和肿瘤内注射、视网膜内注射或视网膜下注射），皮下注射或沉积，包括皮下输注（例如利用浸透压泵），直接施加到癌症区域或其附近，例如借助导管或其它放置装置（例如，包含多孔性、非多孔性或明胶状材料的视网膜片剂 (retinal pellet) 或栓剂或植入物），和吸入。优选通过注射或输注将双链核酸分子或载体施用到癌症部位或其附近。

[0348] 本发明的双链核酸分子可以以单一剂量或分多个剂量来施用。当本发明的双链核酸分子的施用为输注方式时，输注可以是单个持续剂量，或者通过多次输注来施用。优选的是将药剂直接注射到癌症部位或其附近的组织中。特别优选的是将药剂多次注射到癌症部位或其附近的组织中。

[0349] 为了给受试者施用本发明的双链核酸分子，本领域技术人员可以容易地确定合适的剂量方案。例如，双链核酸分子可以一次性施用给受试者，例如以单次注射或者沉积的形式施用到癌症部位或其附近。或者，双链核酸分子可以在约 3 ～约 28 日、更优选约 7 ～约 10 日的期间内每日一次或二次地施用给受试者。在优选的剂量方案中，双链核酸分子可以在 7 日期间内一日一次地注射到癌症部位或其附近。当剂量方案包括多次施用时，应该理解的是，给受试者施用的双链核酸分子的有效量，可以包含在整个剂量方案中施用的该双链核酸分子的总量。

[0350] 组合物

[0351] 而且，本发明还提供包含至少一种本发明的双链核酸分子或编码该分子的载体的药物组合物。特别是，本发明提供以下 [1] ～ [29] 的组合物：

[0352] [1] 一种用于治疗癌症的组合物，其包含至少一种抑制过表达 CX 基因的细胞中的该基因表达的分离的双链核酸分子；其中所述 CX 基因选自 C14orf78、MYBL2、UBE2S 和 UBE2T，且该分子包含有义链和与之互补的反义链，二者相互杂交形成该双链核酸分子，且该分子以选自 SEQ ID NO：47 ～ 57 中的序列为靶标；

[0353] [2][1] 的用于治疗癌症的组合物，其中，所述有义链包含与选自 SEQ IDNO：47 ～ 57 中的靶序列对应的序列；

[0354] [3][1] 的组合物，其中所述细胞为癌细胞。

[0355] [4][1] 的组合物，其中，所治疗的癌症选自胰腺癌、肺癌、乳腺癌、膀胱癌、食道癌、前列腺癌、睾丸精原细胞瘤、结肠癌和胆管细胞癌；

[0356] [5][4] 的组合物，其中，所述肺癌为非小细胞肺癌或小细胞肺癌；

[0357] [6][1] 的组合物，其中，当所选择的 CX 基因为 C14orf78 时，所治疗的癌选自胰腺癌、胆管细胞癌或非小细胞肺癌；

[0358] [7][1] 的组合物，其中，当所选择的 CX 基因为 MYBL2 时，所治疗的癌选自胰腺癌、非小细胞肺癌、小细胞肺癌、膀胱癌、食道癌或睾丸精原细胞瘤；

[0359] [8][1] 的组合物，其中，当所选择的 CX 基因为 UBE2S 时，所治疗的癌选自胰腺癌、乳腺癌、小细胞肺癌、膀胱癌、结肠癌、胆管细胞癌或前列腺癌；

[0360] [9][1] 的组合物，其中，当所选择的 CX 基因为 UBE2T 时，所治疗的癌选自乳腺癌、胆管细胞癌、非小细胞肺癌、小细胞肺癌、膀胱癌、前列腺癌或食道癌；

[0361] [10][1] 的组合物，其中，所述组合物包含多种所述双链核酸分子；

[0362] [11][10] 的组合物，其中，多种双链核酸分子以同一基因为靶标；

[0363] [12][2] 的组合物，其中，双链核酸分子长度小于约 100 个核苷酸；

[0364] [13][12] 的组合物，其中，双链核酸分子长度小于约 75 个核苷酸；

[0365] [14][13] 的组合物，其中，双链核酸分子长度小于约 50 个核苷酸；

[0366] [15][14] 的组合物，其中，双链核酸分子长度小于约 25 个核苷酸；

[0367] [16][15] 的组合物，其中，双链核酸分子长度约 19 个～约 25 个核苷酸；

[0368] [17][2] 的组合物，其中，双链核酸分子由单一多核苷酸构成，所述多核苷酸包含通过间插单链连接在一起的有义链和反义链；

[0369] [18][17] 的组合物，其中，双链核酸分子具有通式 5′ -[A]-[B]-[A′]-3′ ，其中 [A] 为包含与选自 SEQ ID NO:47 ～ 57 中的靶序列对应的序列的有义链，[B] 为由 3 个～ 23 个核苷酸构成的间插单链，[A′] 为包含 [A] 的互补序列的反义链；

[0370] [19][2] 的组合物，其中，双链核酸分子包含 RNA；

[0371] [20][2] 的组合物，其中，双链核酸分子包含 DNA 和 RNA；

[0372] [21][20] 的组合物，其中，双链核酸分子为 DNA 多核苷酸与 RNA 多核苷酸的杂合体；

[0373] [22][21] 的组合物，其中，有义链和反义链多核苷酸分别由 DNA 和 RNA 构成；

[0374] [23][20] 的组合物，其中，双链核酸分子为 DNA 与 RNA 的嵌合体；

[0375] [24][23] 的组合物，其中，位于所述有义与反义多核苷酸中之一或两者的 5’末端之侧翼的至少一个区域是由 RNA 构成的；

[0376] [25][24] 的组合物，其中，侧翼区由 9 ～ 13 个核苷酸构成；

[0377] [26][2] 的组合物，其中，双链核酸分子包含 3’突出端；

[0378] [27][2] 的组合物，其中，双链核酸分子由载体编码，且包含于组合物中；

[0379] [28][27] 的组合物，其中，双链核酸分子具有通式 5′ -[A]-[B]-[A′]-3′ ，其中 [A] 为包含与选自 SEQ ID NO:47 ～ 57 中的靶序列对应的序列的有义链，[B] 为由 3 个～ 23 个核苷酸构成的间插单链，[A′] 为包含 [A] 的互补序列的反义链；

[0380] [29][2] 的组合物，其中该组合物包含转染促进剂和药学可接受的载体。

[0381] 对于本发明的双链核酸分子，在施用给受试者之前，优选采用本技术领域公知的技术将其配制成药物组合物的形式。本发明的药物组合物的特征在于至少是灭菌和无热源的。如在本说明书中使用的，“药物制剂”包括人用和兽医用的制剂。本发明的药物组合物的制备方法在本技术领域的技术范围内，例如 Remington′ s Pharmaceutical Science，

17th ed.,Mack PublishingCompany,Easton,Pa.(1985) 所记载的,将其全体援引并入本说明书。

[0382] 本药学制剂包含至少一种本发明的双链核酸分子或编码它们的载体(例如0.1-90%重量),或者分子的生理上可接受的盐,与生理上可接受的载体基质相混合。优选的生理上可接受的载体基质是水、缓冲液(buffered water)、生理盐水、0.4%食盐水、0.3%甘氨酸、透明质酸等。

[0383] 根据本发明,上述组合物可以包含多种所述双链核酸分子,这些分子可以各自靶向同一靶序列,或者靶向同一 CX 基因中或不同 CX 基因内的不同靶序列。例如,所述组合物可以包含靶向 1、2、3 或 4 个 CX 基因的多种双链核酸分子。或者,例如,所述组合物可以包含靶向同一 CX 基因内的 1 个、2 个、3 个、4 个、5 个或更多个靶序列的双链核酸分子。

[0384] 而且,本组合物可以包含编码 1 个或多个双链核酸分子的载体。例如,所述载体可以编码 1 种、2 种或数种本双链核酸分子。或者,本组合物可以包含多种载体,而各载体编码不同的双链核酸分子。

[0385] 而且,本双链核酸分子可以作为脂质体的形式包含在本组合物中。脂质体的具体情况可以参照“癌症治疗方法”项。

[0386] 此外,本发明的药物组合物中还可以包含传统的药用赋形剂和/或添加剂。合适的药用赋形剂包括稳定化剂、抗氧化剂、浸透压调节剂、缓冲剂、和 pH 调节剂。合适的添加剂包括:生理学生物相容性的缓冲液(例如氨基丁三醇盐酸盐)、补加螯合剂(例如 DTPA 或 DTPA- 双酰胺等),或钙螯合剂复合物(例如、钙 DTPA、CaNaDTPA- 双酰胺),或者,任选的,补加钙或钠盐(例如氯化钙、抗坏血酸酸钙、葡糖酸钙或乳酸钙)。本发明的药物组合物可以进行包装以便作为液体使用,或者也可以加以冷冻干燥。

[0387] 对于固体组合物,可以使用常规的无毒固态载体。例如,药物级的甘露醇、乳酸、淀粉、硬脂酸镁、糖精钠、滑石、纤维素、葡萄糖、蔗糖、碳酸镁等。

[0388] 例如,用于口服施用的固体药物组合物中可以包含上述列举的任意载体和赋形剂,以及 10-95%,优选 25-75%的本发明的一种或多种双链核酸分子。用于气雾剂(吸入)施用的药物组合物可以包含 0.01-20 重量%、优选 1-10 重量%的包被于上述脂质体中的一种或多种本发明的双链核酸分子,以及推进剂。还可以根据需要包含载体,例如用于鼻内投递的卵磷脂等。

[0389] 除上述之外,本组合物中还可以包含其它药学活性成分,只要它们不抑制本双链核酸分子的体内功能。例如,上述组合物中可以包含常规用于癌症治疗的化疗药物。

[0390] 在其它实施方案中,本发明还提供本发明的双链核酸分子在制备用于治疗表达 CX 基因的癌症的药物组合物中的用途。例如,本发明涉及下述双链核酸分子在制备用于治疗表达 CX 基因的癌的药物组合物中的用途:该双链核酸分子在过表达 CX 基因的细胞内抑制 CX 基因的表达,其中 CX 基因选自 C14orf78、MYBL2、UBE2S 和 UBE2T,该分子包含有义链和与之互补的反义链,所述有义链和与之互补的反义链彼此杂交而形成双链核酸分子,且该分子以选自 SEQ ID NO:47-57 中的序列为靶标。

[0391] 此外,本发明还提供生产用于治疗表达 CX 基因的癌症的药物组合物的方法或工艺。该方法或工艺包括将药学或者生理学可接受的载体与在过表达 CX 基因的细胞内抑制该 CX 基因的表达的双链核酸分子一起制剂化的步骤,其中 CX 基因选自 C14orf78、MYBL2、

UBE2S 和 UBE2T，该分子包含有义链和与之互补的反义链作为活性成分，所述有义链和与之互补的反义链彼此杂交而形成双链核酸分子，且该分子以选自 SEQ ID NO :47-57 中的序列为靶标。

[0392] 在另一实施方案中，本发明还提供生产用于治疗表达 CX 基因的癌的药学组合物的方法或工艺，其中该方法或工艺包括施用药学或生理学可接受的载体与活性成分的步骤，其中活性成分是在过表达 CX 基因的细胞中抑制 CX 基因的表达的双链核酸分子，其中所述 CX 基因选自 C14orf78、MYBL2、UBE2S 和 UBE2T，该分子包含有义链和与之互补的反义链，所述有义链和与之互补的反义链彼此杂交而形成双链核酸分子，且该分子以选自 SEQ IDNO :47-57 中的序列为靶标。

[0393] 附图简要说明

[0394] [图 1] 经筛选为治疗靶标候选物的 4 种基因的谱图。通过 RT-PCR 分析进行筛选，检测表达靶基因的细胞。(a) :C14orf78，(b) :MYBL2，(c) :UBE2S，(d) :UBE2T。

[0395] [图 2] 经优化的 siRNA 序列针对 C14orf78 基因的 RNAi 活性的测定。利用内源性表达 C14orf78 基因的细胞 PK-I(a) 和 Panc.02.03(b)，评估了 siRNA 的基因沉默活性、生长抑制作用和非特异性细胞死亡诱导能力。(c) 利用 SK-BR-3(一种表达低水平或不表达 C14orf78 基因的细胞系) 评估了 RNAi 反应的特异性。

[0396] [图 3] 经优化的 siRNA 序列针对 MYBL2 基因的 RNAi 活性的测定。利用内源性表达 MYBL2 基因的细胞 H358(a) 和 TE-9(b)，评估了 siRNA 的基因沉默活性、生长抑制作用和非特异性细胞死亡诱导能力。(c) 利用 SAEC(一种表达低水平或不表达 MYB L2 基因的细胞系) 评估了 RNAi 反应的特异性。

[0397] [图 4] 经优化的 siRNA 序列针对 UBE2S 基因的 RNAi 活性的测定。利用内源性表达 UBE2S 基因的细胞 MCF-7(a)、PK-I(b) 和 SW780(c) 评估了 siRNA 的基因沉默活性、生长抑制作用和非特异性细胞死亡诱导能力。(d) 利用 HMEC(一种表达低水平或不表达 UBE2S 基因的细胞系) 评估了 RNAi 反应的特异性。

[0398] [图 5-1] 经优化的 siRNA 序列针对 UBE2T 基因的 RNAi 活性的测定。利用内源性表达 UBE2T 基因的细胞 MCF-7(a) 和 A549(b)，评估了 siRNA 的基因沉默活性、生长抑制作用和非特异性细胞死亡诱导能力。

[0399] [图 5-2] 经优化的 siRNA 序列针对 UBE2T 基因的 RNAi 活性的测定。利用内源性表达 UBE2T 基因的细胞 SW780(c) 和 DU 145(d)，评估了 siRNA 的基因沉默活性、生长抑制作用和非特异性细胞死亡诱导能力。(e) 利用 HMEC(表达低水平或不表达 UBE2T 基因的细胞系) 评估了 RNAi 反应的特异性。

[0400] [图 6-1] 各 siRNA 针对 4 种靶基因的体内抗肿瘤活性。(a) 通过瘤内注射，向异种移植小鼠施以 LipoTrust(TM)SR 包埋的各 MYBL2siRNA(C7、C13 和 C15) 或萤光素酶 siRNA(作为对照)。各 MYBL2siRNA 显著抑制了第 7 天时的相对肿瘤大小。这些实验进行了 5 次重复。误差线代表均值 +/-SD。* 和 ** 分别意味着 $p < 0.05$ 和 $p < 0.01$(Student 氏 t 检验)。

[0401] [图 6-2] 各 siRNA 的针对 4 种靶基因的体内抗肿瘤活性。(b) 通过瘤内注射，向异种移植小鼠施以去端肽胶原与针对 MYBL2(C16)、C14orf78(C8、C10、C11 和 C24)、UBE2S(C8 和 C9)、UBE2T(C10) 的各 siRNA 所成的复合物以及萤光素酶 (对照)。针对 MYBL2、C14orf78、

UBE2S 和 UBE2T 的各 siRNA 显著抑制了第 7 天时的相对肿瘤大小。误差线代表均值 +/-SD。* 和 ** 分别意味着 $p < 0.05$ 和 $p < 0.01$(Student 氏 t 检验)。

实施例

[0402] 以下的实施例中将进一步描述本发明,它们并不限制权利要求书中所述的本发明的范围。

[0403] [实施例 1] 一般方法

[0404] 组织制备

[0405] 经患者术前知情同意后,从经受手术切除的 34 例患者(膀胱癌)、25 例患者(胆管细胞癌)、48 例患者(结肠癌)、64 例患者(食道癌)、59 例患者(前列腺癌)、15 例患者 (SCLC)、18 例患者(胰腺癌)、37 例患者 (NSCLC)、81 例患者(乳腺癌),获得了临床膀胱癌、胆管细胞癌、结肠癌、食道癌、前列腺癌、小细胞肺癌 (SCLC)、胰腺癌、非小细胞肺癌 (NSCLC) 和乳腺癌样本。

[0406] cDNA 微阵列

[0407] cDNA 微阵列载玻片的制备已在别处描述 (Zembutsu H et al.,Cancer Res2002 Jan 15,62(2) :518-2. ;Nishidate T et al.,Int J Oncol 2004 Oct,25(4) :797-819)。为了分析不同的癌症表达谱,本发明人制备了重复的两组载玻片,含 23,040 个(结肠癌、软组织肉瘤和睾丸精原细胞瘤、前列腺癌)或 27,648 个(乳腺癌和膀胱癌)或 36,864 个(胰腺癌、NSCLC、SCLC 和食道癌)cDNA 斑点以降低实验波动。简言之,从有肿瘤的患者和从相应的正常组织中提取总 RNA,用于癌症表达分析。进行基于 T7 的 RNA 扩增以获得足够量的 RNA 用于微阵列实验。取扩增后的 RNA 的等分试样,使用足够量的 Cy5-dCTP 或 Cy3-dCTP(Amersham Biosciences,Buckinghamshire,UnitedKingdom) 通过反转录加以标记。

[0408] 如前人所描述的进行杂交、洗涤和检测 (Zembutsu H et al.,Cancer Res2002 Jan 15,62(2) :518-27 ;Nishidate T et al., Int J Oncol 2004 Oct,25(4) :797-819)。为了检测在癌症(胰腺癌、NSCLC 和乳腺癌)中普遍上调的基因,首先筛选了所有基因在微阵列上的整体表达模式,选择出在 > 20% 的所检查癌症病例中存在的表达比例 > 5.0 的那些基因。最后,为了获得对目标癌症高度特异的治疗靶标,本发明人通过参照关于正常人类组织表达的内部数据库,选择了在正常组织中不表达的基因。

[0409] 细胞系和细胞培养

[0410] 本发明人制备了肺癌、乳腺癌、胰腺癌和正常上皮细胞系,把它们维持在充分的培养基中,用于离体分析和提取 mRNA 以评估靶基因表达水平。肺癌细胞系 :A549、EBC-1、H1373、H1435、H1650、H1666、H1781、H1793、H2170、H226、H358、H520、H522、H596、PC-14、SK-LU-1、SW900 和 SBC5 ;乳腺癌细胞系 :BT-20、BT-474、BT-549、HCC1143、HCC1500、HCC 1599、HCC 1937、MCF-7、MDA-MB-453、MDA-MB-453S、SK-BR-3、T47D 和 ZR-75-1 ;胰腺癌细胞系 :capan-1、capan-2、HPAF-II、KLM-1、KP-1N、MiaPaCa-2、Panc.02.03、PK-1、PK-45P、PK-59、PK-9、SUIT-2 和 Panc-1 ;以及正常上皮细胞系 :小气道上皮细胞 (SAEC) 和乳腺上皮细胞 (HMEC)。

[0411] 半定量 RT-PCR

[0412] 对于所选择的基因,利用半定量 RT-PCR 实验,评估了它们在正常器官(心、肝、肺

和肾）、癌细胞系、相应正常组织和正常上皮细胞系中的表达水平。具体地，使用寡核苷酸 d(T)16 引物 (Roche) 和 Superscript II(Invitrogen)，将来自每种细胞系、正常器官、及 siRNA 感染细胞的 3-mc g mRNA 等分试样反转录成单链 cDNA。α-肌动蛋白 (ACTB)、β2 微球蛋白 (β2MG) 和微管蛋白 α3(TUB A3) 的表达分别充当肺癌、乳腺癌和胰腺癌的内部对照。使用干扰素诱导的跨膜蛋白质 1(IFITM1) 作为各 siRNA 脱靶活性的标志。对 PCR 反应的循环数进行优化，以确保产物强度在扩增的线性相内。对各 cDNA 混合物进行稀释，随后使用如下引物对进行 PCR 扩增：

[0413] C14orf78：

[0414] 正向引物：5′ -GAGAAGGAAGAGGGTGAACTGAT-3′ (SEQ ID NO:9)；

[0415] 反向引物：5′ -CAGTGGACATGGATAGATGAGAA-3′ (SEQ ID NO:10)；

[0416] MYBL2：

[0417] 正向引物：5′ -GAAGCCACTTCACGACACCT-3′ (SEQ ID NO:11)；

[0418] 反向引物：5′ -ATCCTAAGCAGGGTCTGAGATG-3′ (SEQ ID NO:12)；

[0419] UBE2S：

[0420] 正向引物：5′ -TACTTCCTGACCAAGATCTTCCA-3′ (SEQ ID NO:13)；

[0421] 反向引物：5′ -TTAGAGACAGAGTTGGAGGGAGG-S′ (SEQ ID NO:14)；

[0422] UBE2T：

[0423] 正向引物：5′ -CAAATATTAGGTGGAGCCAACAC-3′ (SEQ ID NO:15)；

[0424] 反向引物：5′ -TAGATCACCTTGGCAAAGAACAC-3′ (SEQ ID NO:16)；

[0425] ACTB：

[0426] 正向引物：5′ -AGGATGCAGAAGGAGATCAC-3′ (SEQ ID NO:17)；

[0427] 反向引物：5′ -AGAAAGGGTGTAACGCAACT-3′ (SEQ ID NO:18)；

[0428] β2MG：

[0429] 正向引物：5′ -CACCCCCACTGAAAAAGATGA-3′ (SEQ ID NO:19)；

[0430] 反向引物：5′ -TACCTGTGGAGCAACCTGC-3′ (SEQ ID NO:20)；

[0431] TUBA3：

[0432] 正向引物：5′ -AAGGATTATGAGGAGGTTGGTGT-3′ (SEQ ID NO:21)；

[0433] 反向引物：5′ -CTTGGGTCTGTAACAAAGCATTC-3′ (SEQ ID NO:22)；

[0434] IFITM1：

[0435] 正向引物：5′ -GATCAACATCCACAGCGAGA-3′ (SEQ ID NO:23)；

[0436] 反向引物：5′ -TGTCACAGAGCCGAATACCA-3′ (SEQ ID NO:24)。

[0437] RNAi 实验

[0438] 利用 Lipofectamine 2000™(Invitrogen)，将针对 4 种候选基因 (C14orf78、MYBL2、UBE2S 和 UBE2T) 的 dsRNA 寡核苷酸以 10pmol/ 孔转染至 96 孔微孔板 (Becton Dickinson) 上的表达靶基因的癌细胞和对照细胞中。各细胞系的培养细胞的起始浓度不同。例如，PK-1(3,000-4,000 个细胞 / 孔)、SK-BR-3(4,000 个细胞 / 孔)、H358(5,000-6,000 个细胞 / 孔)、SAEC(9,000 个细胞 / 孔)、MCF-7(2,500-3,500 个细胞 / 孔) 以及 HMEC(7,000 个细胞 / 孔)。使用 SiControl I(Dharmacon) 作为阴性对照，以避免错误解读由与 siRNA 特异性无关的原因引起的细胞死亡。用 SiTox(Dharmacon) 作为阳性对照以确认转染效率。

对于各候选靶序列，检测了不同序列的基因特异性 siRNA，以优化所述序列作为治疗药物。转染后，考察了各 siRNA 对癌细胞的生长阻碍作用。用 RT-PCR 分析了 siRNA 敲低靶基因的能力；通过监测 IFITM1 的上调——IFITM1 是由常见双链 RNA 感染引发的干扰素应答的指标——确认了 siRNA 的脱靶活性。

[0439] 体内 siRNA 处理

[0440] 将经筛选的 4 种针对 MYBL2 基因的 siRNA(C7、C13 或 C15) 包裹在 LipoTrust™SR(Hokkaido System Science) 的脂质结构中，每三天将其瘤内注射到 H358 异种移植小鼠体内。简言之，将每种 siRNA 50mc g/mL 与 0.5mcmol/mL LipoTrast™SR 混合，轻柔超声处理以形成包埋所需 siRNA 的脂质体。将 400mc L 脂质体 /siRNA 用于皮下移植了人肺癌细胞的小鼠的癌症治疗。每天监测肿瘤进展的减少。或者，利用去端肽胶原 (AteloGene™, KOKEN) 作为载体，评估经筛选的针对 C14orf78(C8、C10、C11 和 C24)；MYBL2(C16)；UBE2S(C8 和 C9) 和 UBE2T(C10) 的 siRNA 序列的治疗潜力。使用旋转器 (4rpm) 于 4℃使等体积 AteloGene™ 和 10mc M siRNA 彼此轻柔混合 20 分钟。接着将混合物于 4℃离心 (10,000rpm)1 分钟以去掉泡沫。每三天将 200mc L 的混合物瘤内注射到小鼠肩部的肿瘤中。在两种情况下都在第一次注射后第 7 天评估 siRNA 的抗癌作用。

[0441] 细胞增殖测定

[0442] 从 siRNA 转染起 48h、72h、96h 或 120h 后，利用 IN Cell Analyzer 1000(GE Healthcare Bio-Science KK) 评估了用钙黄绿素可视化的活细胞的浓度。

[0443] [实施例 2] 筛选在临床癌症样本中上调而在正常器官中无表达或低表达的基因

[0444] 如以前所述，进行了 cDNA 微阵列分析 (Zembutsu H et al.,Cancer Res2002 Jan 15,62(2)：518-27；Nishidate T et al., Int J Oncol 2004 Oct,25(4)：797-819)。通过癌组织和相应正常上皮之间表达图式的比较，选择了在临床癌组织中普遍上调的基因。接着，进行半定量 RT-PCR 分析以选择这样的癌症特异性基因，其被检测到在癌细胞系中高表达，但在相应正常器官和正常生命器官中则没有高表达 (图 1)。将正常器官中高表达的基因排除，以防当用这样的基因作为治疗中抑制的靶基因时，可能诱发致死性的副作用。

[0445] [实施例 3] 针对候选物的定制 siRNA 的设计

[0446] 利用在 Ambion 公司网站 (http://www.ambion.com/techlib/misc/siRNA_finder.html)(Tuschl T et al.,Genes Dev 1999Dec 15,13(24)：3191-7) 上提供的 siRNA 设计工具设计了针对各候选基因的 siRNA 序列，以选择 siRNA 的候选序列。将每个 siRNA 导入癌细胞和对照细胞，评估它们的相对细胞生存力，以获得对于抑制细胞生长最有效的序列 (表 1)。

[0447] [表 1]

[0448] 针对 4 种候选基因设计的 siRNA 序列

[0449]

靶基因	siRNA名	链	序列	SEQ ID NO
C14orf78	C8	靶	5'-GATATGCCATCCCAGATTT-3'	47
		有义	5'-GAUAUGCCAUCCCAGAUUUUU-3'	25
		反义	5'-AAAUCUGGGAUGGCAUAUCUU-3'	26
	C10	靶	5'-GTCAAATTCCCCAAATTAA-3'	48
		有义	5'-GUCAAAUUCCCCAAAUUAAUU-3'	27
		反义	5'-UUAAUUUGGGGAAUUUGACUU-3'	28
	C11	靶	5'-GTGTCCAGAGGCCAATATT-3'	49
		有义	5'-GUGUCCAGAGGCCAAUAUUUU-3'	29
		反义	5'-AAUAUUGGCCUCUGGACACUU-3'	30
	C24	靶	5'-GGCAGGGCTCCAAAAGACA-3'	50
		有义	5'-GGCAGGGCUCCAAAAGACAUU-3'	31
		反义	5'-UGUCUUUUGGAGCCCUGCCUU-3'	32
MYBL2	C7	靶	5'-GGAGCCCATCGGTACAGAT-3'	51
		有义	5'-GGAGCCCAUCGGUACAGAUUU-3'	33
		反义	5'-AUCUGUACCGAUGGGCUCCUU-3	34
	C13	靶	5'-CGGCGGAGCCCCATCAAGA-3'	52
		有义	5'-CGGCGGAGCCCCAUCAAGAUU-3'	35
		反义	5'-UCUUGAUGGGGCUCCGCCGUU-3'	36
	C15	靶	5'-GCGGAGCCCCATCAAGAAA-3'	53
		有义	5'-GCGGAGCCCCAUCAAGAAAUU-3'	37
		反义	5'-UUUCUUGAUGGGGCUCCGCUU-3'	38
	C16	靶	5'-GATGTGAAGCTGATGATGT-3'	54
		有义	5'-GAUGUGAAGCUGAUGAUGUUU-3'	39
		反义	5'-ACAUCAUCAGCUUCACAUCUU-3'	40
UBE2S	C8	靶	5'-TGCTGACCATCAAGTGCCT-3'	55
		有义	5'-UGCUGACCAUCAAGUGCCUUU-3'	41
		反义	5'-AGGCACUUGAUGGUCAGCAUU-3	42
	C9	靶	5'-CCATATGCTGGAGGTCTGT-3'	56
		有义	5'-CCAUAUGCUGGAGGUCUGUUU-3'	43
		反义	5'-ACAGACCUCCAGCAUAUGGUU-3'	44
UBE2T	C10	靶	5'-AGAGAGAGCTGCACATGTT-3'	57
		有义	5'-AGAGAGAGCUGCACAUGUUUU-3'	45
		反义	5'-AACAUGUGCAGCUCUCUCUUU-3'	46

[0450] [实施例 4] 基因特异性 siRNA 的优化及其沉默特异性的评估

[0451] C14orf78 是胰腺癌的治疗靶标，因为其在临床样本中过表达 (T/N 比值＞＝ 5)；18 例胰腺癌样本中的 11 例，25 例胆管细胞癌中的 14 例，37 例非小细胞肺癌中的 10 例（表 2）。所有对 C14orf78(C8、C10、C11 和 C24) 最优化的 siRNA 都在抑制细胞增殖的同时，有效地敲低了 PK-I 和 Panc.02.03 中的基因表达（图 2a,b）。本发明人利用针对该基因的双链 RNA(dsRNA) 进一步考察了干扰素途径的激活。干扰素诱导的跨膜蛋白 1(IFITM1) 是干扰素应答——其结果是双链 RNA 的感染引起不希望的非特异性细胞死亡——的指标。在本发明中，IFITM1 的表达几乎一致地无变化（图 2a，b）。此外，SK-BR-3（一种表达低水平或不表达 C14orf78 基因的细胞系）被 siRNA 感染后增殖未显示明显变化（图 2c）。因此证实了本发明 siRNA 针对 C14orf78 的特异性。

[0452] MYB L2 基因在多种癌症中显示出过表达。具体地说，该基因在临床样本中是上调的（比值＞＝5）；34 例膀胱癌中的 18 例，64 例食道癌中的 29 例，37 例非小细胞肺癌 (NSCLC) 中的 18 例，18 例胰腺癌中的 6 例，以及 15 例小细胞肺癌 (SCLC) 中的 14 例（表 2）。此外，据报道 MYBL2 基因在睾丸精原细胞瘤中也是上调的 (WO2004/031410)。最近的一项报道显示，MYBL2 蛋白作为细胞周期进程中涉及的转录因子而发挥作用 (Garcia P & FramptonJ, J Cell Sci 2006 Apr 15,119(Pt 8):1483-93, Epub 2006Mar 21)。用 cDNA 微阵列获得的表达图式和以前关于 MYBL2 的报道提示，该基因的过表达刺激细胞增殖，促进多种类型癌症的癌发生或肿瘤发展。所有经筛选的针对 MYBL2 的 siRNA(C7、C13、C15 和 C16) 都显著降低了 NSCLC 细胞系 (H358) 和食道癌细胞系 (TE-9) 中该基因的表达水平和细胞生长（图 3a,b），而由 siRNA 诱导的生长抑制是相当严格的，仅限于特异性的 siRNA。实际上，未观察到可检测的干扰素应答的激活（图 3a,b）。此外，在正常小气道上皮细胞 (SAEC) 中也未能观察到可检测的生长抑制，SAEC 是 MYBL2 非表达细胞系（图 3c）。因此，MYBL2 基因是 siRNA 治疗的极好靶标，不仅可用于 NSCLC，还可用于 SCLC、食道癌、膀胱癌、睾丸精原细胞瘤和胰腺癌。所以，本发明的 MYBL2- 特异性 siRNA 是治疗这些癌症的有力工具。

[0453] UBE2S 基因在临床样本中过表达；所有 SCLC 病例，34 例膀胱癌中的 29 例，81 例乳腺癌中的 27 例，25 例胆管细胞癌中的 9 例，59 例前列腺癌中的 18 例，48 例结肠癌中的 11 例，以及 18 例胰腺癌中的 12 例（表 2）。正如编码泛素样 E2 连接酶蛋白的 UBE2S 基因一样，UBE2T 基因也在多种类型的癌症中显示出表达增加，即，在 25 例胆管细胞癌中的 12 例，15 例 SCLC 中的 12 例，34 例膀胱癌中的 23 例，81 例乳腺癌中的 28 例，37 例 NSCLC 中的 13 例，64 例食道癌中的 14 例，以及 59 例前列腺癌中的 15 例（表 2）。选择的针对 UBE2S 的 siRNA(C8 和 C9) 显著降低了乳腺癌细胞系 (MCF7)、胰腺癌细胞系 (PK-I) 和膀胱癌细胞系 (SW780) 中该基因的表达水平和细胞生存力（图 4a-c）。未能观察到干扰素应答的激活（图 4a-c）。因此，双链 RNA 感染引起的不希望的非特异性细胞死亡似乎不是由本发明 siRNA 诱导的。同样地，对 UBE2T 的 siRNA(C10) 有效地抑制了乳腺癌 (MCF7)、NSCLC(A549)、膀胱癌 (SW780) 和前列腺癌 (DU-145) 中的基因表达（图 5-1a-b、5-2c-d）。此外，对于既不表达 UBE2S 也不表达 UBE2T 的细胞系 HMEC（正常乳腺上皮细胞）而言，也未能观察到可检测的生长抑制（图 4d,5-2e）。因此，UBE2S 是用于包括 SCLC、乳腺、胰腺、膀胱、结肠、胆管细胞和前列腺癌症在内的很多种癌症的治疗靶标；UBE2T 是用于肺、膀胱、乳腺、胆管细胞、食道和前列腺癌症的靶标。

[0454] [表 2]

[0455] 来自 cDNA 微阵列数据库的临床癌组织中经筛选的基因的过表达 (T/N 比值＞＝5) 频率

[0456]

基因	膀胱癌	乳腺癌	胆管细胞癌	结肠癌	食道癌	NSCLC	胰腺癌	前列腺癌	SCLC
C14orf78	0/34	1/81	14/25	1/48	0/19	10/37	11/18	3/59	1/15
MYBL2	18/34	11/81	5/25	9/48	29/64	18/37	6/18	6/59	14/15

基因	膀胱癌	乳腺癌	胆管细胞癌	结肠癌	食道癌	NSCLC	胰腺癌	前列腺癌	SCLC
UBE2S	29/34	27/81	9/25	11/48	2/64	1/37	12/18	18/59	15/15
UBE2T	23/34	28/81	12/25	8/48	14/64	13/37	3/18	15/59	12/15

[0457] [实施例 5]针对靶基因筛选的 siRNAs 的体内疗效

[0458] 利用体内模型评估了经筛选的siRNA的治疗可用性(therapeuticavailability)。将MYBL2siRNA(C7、C 13、C15和C 16)包埋到商业脂质体或去端肽胶原中,并瘤内注射到移植了H358细胞的裸鼠体内。通过每天监测肿瘤大小的转变,评估这些siRNA的治疗效力。与对照相比,以LipoTrust™SR包埋的MYBL2 siRNA(C7、C13和C15)处理的第7天肿瘤大小显著受抑制(*p < 0.05,**p < 0.01;student 氏 t 测验)(图 6-1a)。另一方面,将去端肽胶原与针对MYBL2(C 16)、C14orf78(C8、C10、C11和C24)、UBE2S(C8和C9)和UBE2T(C10)的siRNA形成的复合物瘤内注射到肿瘤模型小鼠体内后,与对照 siRNA 相比,其发挥显著的肿瘤生长消除作用。显著性差异和 +/-SD 也用 Student 氏 t 检验计算(*p < 0.05;**p < 0.01)(图 6-2b)。因此,筛选的所有针对 C14orf78、MYBL2、UBE2S 和 UBE2T 的 siRNA 都可能是用于多种癌症的有希望的治疗试剂。

[0459] 讨论

[0460] 近年来,一种利用基因特异性 siRNA 的癌症治疗新手段正在临床试验中使用(Bumcrot D et al.,Nat Chem Biol 2006 Dec,2(12):711-9)。RNAi 似乎已在主要的技术平台中已获得了一席之地(Putral LN et al.,Drug NewsPerspect 2006 Jul-Aug,19(6):317-24;Frantz S,Nat Rev Drug Discov 2006 Jul,5(7):528-9;Dykxhoorn DM et al.,Gene Ther 2006 Mar,13(6):541-52)。

[0461] 如先前所描述的(参见一般方法),本发明人鉴定了在癌症中特异性表达而不在正常器官中表达的基因。在使用本发明的双链核酸分子治疗的病例中,由于靶基因的表达图式以相当排他性(exclusive)的方式对癌症高度特异,所以不会引起严重的副作用。因此,本发明的靶向癌症特异性基因的双链核酸分子可以作为有力的工具,用来开发无任何副作用的抗癌药。

[0462] C14orf78 蛋白是一种由 6,287 个氨基酸残基构成的巨大膜蛋白,具有 PDZ 区。C14orf78 蛋白有一种家族蛋白称为 AHNAK1 蛋白,其 PDZ 域与 L- 型电压调节钙通道的亚单元结合。因此,已有人预测,C14orf78 蛋白的 PDZ 域与许多通道蛋白(包括涉及钙转运的通道蛋白)的 C 末端残基相互作用(Komuro A et al.,Proc Natl Acad Sci USA 2004Mar 23,101(12):4053-8,Epub 2004 Mar 8)。上文已经提到,AHNAK1 缺失小鼠表型显示无异常,因此断定,AHNAK1 蛋白对于细胞的发育或增殖而言不是必需的。然而,没有关于 C14orf78 基因敲除小鼠的表型的报道(Komuro A et al.,Proc NatlAcad Sci USA 2004 Mar 23,101(12):4053-8,Epub 2004 Mar 8)。因此,尚不清楚 C14orf78 蛋白是否在细胞的生长发育中起重要的作用。在本发明中,C14orf78 蛋白被证明是胰腺癌细胞系的细胞生长或存活的关键因素。为了治疗恶性 PDAC,本发明提供了包含靶向 C14orf78 基因的 siRNA 的治疗剂。

[0463] 在通过全基因组 cDNA 微阵列鉴定的许多过表达基因中(Kikuchi T et al.,

Oncogene 2003 Apr 10,22(14) :2192-205),选出了 MYBL2 基因用于进一步的详细分析,因为通过 cDNA 微阵列检测到了癌细胞中明显的信号强度(与正常肺相比超出 5 倍)。考虑到治疗的副作用,欲用于癌症治疗的 siRNA 靶标的分子在正常成年人组织中限制性表达是一个重要因素。此外,各种临床癌症的基因表达图式的内部数据库揭示了 MYBL2 基因在膀胱癌、食道癌、NSCLC、SCLC、胰腺癌(参见“结果”)和软组织肉瘤(数据未显示)以及所述睾丸肿瘤(参见“结果”)中显著过表达(比值>=5)。先前的关于 MYBL2 缺失 (-/-) 小鼠的研究证明了 MYBL2 蛋白对于胚胎发育是必要的 ;小鼠在约 E4.5 时死亡 (Tanaka Y et al., J Biol Chem 1999 Oct 1,274(40) :28067-70)。在正常成年人组织中几乎检测不到 MYBL2 基因表达,而在胚胎组织和癌症中检测到了大量表达。因此,MYBL2 可能与癌发生和肿瘤的发生有关,可能作为治疗许多种癌症的极好靶标,同时产生副作用的风险较低。

[0464] SMART 程序 (http://smart.embl-heidelberg.de/) 预测,UBE2T 和 UBE2S 蛋白都包含 UBCc 域(泛素缀合酶 E2,催化域同源物),说明这两种蛋白具有潜在的经由单泛素化方式的 E2 泛素酶活性,且与乳腺癌的肿瘤发生有关。许多先前的研究报道了 E3 连接酶的失调控导致癌症发生 (Yen L et al., Cancer Res 2006 Dec 1,66(23) :11279-86 ;Ohh M, Neoplasia 2006 Aug,8(8) :623-9 ;Lisztwan J et al., Genes Dev 1999Jul 15,13(14) :1822-33),只有少数报道提出,癌症发生中可能涉及 E2 连接酶 (Jung CR et al., Nat Med 2006 Jul,12(7) :809-16,Epub 2006 Jul 2 ;Okamoto Y et al.,Cancer Res 2003 Jul 15, 63(14) :4167-73)。先前的研究报道,UBE2 家族蛋白 (UBE2) 是推定的泛素缀合酶 (E2 连接酶),其有助于蛋白质降解途径。然而,UBE2 在癌症中起作用的细节仍然未知,有待于研究揭示是否它们仅具有蛋白质降解途径中的 E2 连接酶活性,或者还具有其他体内性质。

[0465] 工业实用性

[0466] 本发明人们显示了细胞生长被特异性靶向 C14orf78、MYBL2、UBE2S 和 UBE2T 基因的双链核酸分子所抑制。因此,这些新的双链核酸分子是开发抗癌药物的有用的候选物。例如,阻碍 C14orf78、MYBL2、UBE2S 或 UBE2T 蛋白质表达或阻断其活性的试剂可能发挥抗癌剂的治疗功效,特别是用于治疗肺癌、乳腺癌、膀胱癌、胆管细胞癌、食道癌、前列腺癌、前列腺癌或睾丸精原细胞瘤的抗癌剂。

[0467] 对于本发明,参照其特定实施方案进行了具体描述,但这里,本领域技术人员完全能够理解的是,在不脱离本发明的精神和范围的前提下,可以进行各种变更和修饰。本说明书中提及的全部出版物、专利申请、专利和其它参考文献均全部援引并入本说明书。如有矛盾,以本说明书包括定义在内为准。而且,材料、方法和实施例仅用于说明目的,不意在限制。

序列表

<110> 肿瘤疗法科学股份有限公司 (ONCOTHERAPY SCIENCE, INC.)

<120> 治疗癌症的组合物及方法

<130>ONC-A0709P

<150>US 60/937,616
<151>2007-06-27

<160>57

<170>PatentIn version 3.4

<210>1
<211>15958
<212>DNA
<213> 人类

<220>
<221>CDS
<222>(1)..(15147)

<400>1

```
atg ccc aag ttc aag atg cca ttg ttc ggg gcg tca gcc cca ggc aag       48
Met Pro Lys Phe Lys Met Pro Leu Phe Gly Ala Ser Ala Pro Gly Lys
1               5                  10                  15
tcc atg gag gcc tcg gtg gat gtg tct gcg ccg aag gtg gag gcc gac       96
Ser Met Glu Ala Ser Val Asp Val Ser Ala Pro Lys Val Glu Ala Asp
            20                  25                  30
gtg agc ctc ctc tcc atg cag ggg gac ctc aag acc act gac ctc agc      144
Val Ser Leu Leu Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Leu Ser
        35                  40                  45
gtc cag acc cct tcc gct gac ctg gag gtc cag gat ggc caa gtg gat      192
Val Gln Thr Pro Ser Ala Asp Leu Glu Val Gln Asp Gly Gln Val Asp
    50                  55                  60
gtg aaa ctt ccg gag ggc ccc ctg ccc gag gga gcc agc ctc aaa ggg      240
Val Lys Leu Pro Glu Gly Pro Leu Pro Glu Gly Ala Ser Leu Lys Gly
```

65 70 75 80

cac ctg ccc aag gtg cag agg ccc agt ttg aag atg ccc aaa gtg gac 288
His Leu Pro Lys Val Gln Arg Pro Ser Leu Lys Met Pro Lys Val Asp
85 90 95

ctc aag ggc ccc aag ctg gac ctg aaa ggc ccc aag gcg gaa gtg aca 336
Leu Lys Gly Pro Lys Leu Asp Leu Lys Gly Pro Lys Ala Glu Val Thr
100 105 110

gcc ccc gat gtg aag atg tct ctg tcc agc atg gag gtg gac gtc cag 384
Ala Pro Asp Val Lys Met Ser Leu Ser Ser Met Glu Val Asp Val Gln
115 120 125

gcc ccg aga gca aag ctg gat ggt gcg cgg ctg gag ggg gac ctg tcc 432
Ala Pro Arg Ala Lys Leu Asp Gly Ala Arg Leu Glu Gly Asp Leu Ser
130 135 140

ctg gcc gac aag gag gtg act gcc aaa gac agc aag ttc aaa atg ccc 480
Leu Ala Asp Lys Glu Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro
145 150 155 160

aag ttc aag atg cca tca ttc ggg gtg tcg gcc cca ggc aag tcc atg 528
Lys Phe Lys Met Pro Ser Phe Gly Val Ser Ala Pro Gly Lys Ser Met
165 170 175

gag gac tcg gtg gat gtg tct gcg ccg aag gtg gag gcc gac gtg agc 576
Glu Asp Ser Val Asp Val Ser Ala Pro Lys Val Glu Ala Asp Val Ser
180 185 190

ctc tcc tcc atg cag ggg gac ctc aag gcc act gac ctc agc att cag 624
Leu Ser Ser Met Gln Gly Asp Leu Lys Ala Thr Asp Leu Ser Ile Gln
195 200 205

ccc cct tcc gct gac ctg gag gtc cag gct ggc caa gtg gat gtg aaa 672
Pro Pro Ser Ala Asp Leu Glu Val Gln Ala Gly Gln Val Asp Val Lys
210 215 220

ctt ccg gag ggc cct gtg ccc gag gga gcc ggc ccc aaa gtg cac ctg 720
Leu Pro Glu Gly Pro Val Pro Glu Gly Ala Gly Pro Lys Val His Leu
225 230 235 240

ccc aaa gtg gag atg ccc agt ttc aag atg ccc aaa gtg gac ctc aag 768
Pro Lys Val Glu Met Pro Ser Phe Lys Met Pro Lys Val Asp Leu Lys
245 250 255

ggc ccc cag ata gat gtt aag ggc ccc aag ctg gac ctg aaa ggc ccc 816
Gly Pro Gln Ile Asp Val Lys Gly Pro Lys Leu Asp Leu Lys Gly Pro
260 265 270

aag gcg gaa gtg aca gcc ccc gat ggc gag gtg tct ctg ccc agc atg 864
Lys Ala Glu Val Thr Ala Pro Asp Gly Glu Val Ser Leu Pro Ser Met

275 280 285

gag gtg gat gtc cag gcc cag aag gcc aag ctg gat ggt gcg tgg ctg 912
Glu Val Asp Val Gln Ala Gln Lys Ala Lys Leu Asp Gly Ala Trp Leu
290 295 300

gag ggg gac ctg tcc ctg gcc gac aag gac gtg act gcc aaa gac agc 960
Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp Val Thr Ala Lys Asp Ser
305 310 315 320

aag ttc aaa atg ccc aag ttc aag atg ccg tcg ttc ggg gta tcg gcc 1008
Lys Phe Lys Met Pro Lys Phe Lys Met Pro Ser Phe Gly Val Ser Ala
325 330 335

cca ggg aag tcc atc aag gcc ttg gtg gat gtg tct gca ccc aag gtg 1056
Pro Gly Lys Ser Ile Lys Ala Leu Val Asp Val Ser Ala Pro Lys Val
340 345 350

gag gcc gac ctg agt ctc ccc tcc atg cag ggg gac ctg aag acc act 1104
Glu Ala Asp Leu Ser Leu Pro Ser Met Gln Gly Asp Leu Lys Thr Thr
355 360 365

gac ctc agc att cag cct gct tct act gac ctg aag gtc cag gct gac 1152
Asp Leu Ser Ile Gln Pro Ala Ser Thr Asp Leu Lys Val Gln Ala Asp
370 375 380

cag gtg gat gtg aag ctc ccg gag ggc cac ctg ccc gag gga gct ggc 1200
Gln Val Asp Val Lys Leu Pro Glu Gly His Leu Pro Glu Gly Ala Gly
385 390 395 400

ctt aaa ggg cac ttg ccc aag gtg gag atg ccc agt ttc aag atg ccc 1248
Leu Lys Gly His Leu Pro Lys Val Glu Met Pro Ser Phe Lys Met Pro
405 410 415

aaa gtg gcc ctc aag ggc ccc cag gtg gac gtc aag ggc ccc aag ctg 1296
Lys Val Ala Leu Lys Gly Pro Gln Val Asp Val Lys Gly Pro Lys Leu
420 425 430

gac ctg aaa agc ccc aag gcg gaa gtc aca gcc cct gat gtg gag gtg 1344
Asp Leu Lys Ser Pro Lys Ala Glu Val Thr Ala Pro Asp Val Glu Val
435 440 445

tct ctg ccc agc gtg gag gtg gac gtc gag gcc ccg gga gcc aag ctg 1392
Ser Leu Pro Ser Val Glu Val Asp Val Glu Ala Pro Gly Ala Lys Leu
450 455 460

gac agt gcg cgg ctg gag ggg gaa ctg tcc ctg gcc gac aag gat gtg 1440
Asp Ser Ala Arg Leu Glu Gly Glu Leu Ser Leu Ala Asp Lys Asp Val
465 470 475 480

act gcc aaa gac agc agg ttc aaa atg ccc aag ttc aag atg cca tcg 1488
Thr Ala Lys Asp Ser Arg Phe Lys Met Pro Lys Phe Lys Met Pro Ser

485 490 495

ttc ggg gcg tca gcc cca ggc aag tcc atc gag gcc tcg gtg gat gtg 1536
Phe Gly Ala Ser Ala Pro Gly Lys Ser Ile Glu Ala Ser Val Asp Val
500 505 510

tct gca ccc aaa gtg gag gcc gac gtg agt ctc ccc tcc atg cag ggg 1584
Ser Ala Pro Lys Val Glu Ala Asp Val Ser Leu Pro Ser Met Gln Gly
515 520 525

gac ctc aag acc act gac ctc agc att cag ccc cct tcc gct gac ctg 1632
Asp Leu Lys Thr Thr Asp Leu Ser Ile Gln Pro Pro Ser Ala Asp Leu
530 535 540

gag gtc cac gct ggc cag gtg gac gtg aag ctc ctg gag ggc cac gtg 1680
Glu Val His Ala Gly Gln Val Asp Val Lys Leu Leu Glu Gly His Val
545 550 555 560

cct gag gga gcc ggc ttc aaa ggg cac ctg ccc aag gtg cag atg cct 1728
Pro Glu Gly Ala Gly Phe Lys Gly His Leu Pro Lys Val Gln Met Pro
565 570 575

agt ttg aag atg ccc aaa gtg gac ctc aag ggc ccc cag gtg gaa gtc 1776
Ser Leu Lys Met Pro Lys Val Asp Leu Lys Gly Pro Gln Val Glu Val
580 585 590

agg ggc ccc aag ctg gac ctg aaa ggt cat aag gca gag gtg acg gcc 1824
Arg Gly Pro Lys Leu Asp Leu Lys Gly His Lys Ala Glu Val Thr Ala
595 600 605

cac gaa gtg gct gtg tct ctg ccc agt gtg gag gtg gac atg cag gcc 1872
His Glu Val Ala Val Ser Leu Pro Ser Val Glu Val Asp Met Gln Ala
610 615 620

ccg gga gcc aag ttg gat ggc gca cag ctg gac ggg gac ctg tcc ctg 1920
Pro Gly Ala Lys Leu Asp Gly Ala Gln Leu Asp Gly Asp Leu Ser Leu
625 630 635 640

gct gac aag gac gtg act gcc aaa gac agc aag ttc aaa atg ccc aag 1968
Ala Asp Lys Asp Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys
645 650 655

ttc aag atg ccg tcg ttc ggg gtg tct gcc cca ggc aag tcc att gag 2016
Phe Lys Met Pro Ser Phe Gly Val Ser Ala Pro Gly Lys Ser Ile Glu
660 665 670

gcc tcc gtg gac ctg tct gca ccc aag gtg gag gcc gac atg agc ctc 2064
Ala Ser Val Asp Leu Ser Ala Pro Lys Val Glu Ala Asp Met Ser Leu
675 680 685

ccc tcc atg cag ggg gac ctc aag acc act gac ctc agc att cag ccc 2112
Pro Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Leu Ser Ile Gln Pro

690 695 700

cct tcc act gac ctg gag ctc cag gct ggc caa ttg gac gtg aaa ctc 2160
Pro Ser Thr Asp Leu Glu Leu Gln Ala Gly Gln Leu Asp Val Lys Leu
705 710 715 720

cca gag ggc ccc gtg ccc gag gga gcc ggc ctc aaa ggg cac ctg ccc 2208
Pro Glu Gly Pro Val Pro Glu Gly Ala Gly Leu Lys Gly His Leu Pro
725 730 735

aag ctg cag atg ccc agt ttc aag gtg ccc aaa gtg gac ctc aag ggc 2256
Lys Leu Gln Met Pro Ser Phe Lys Val Pro Lys Val Asp Leu Lys Gly
740 745 750

cct gaa ata gac atc aag ggc ccc aag ctg gac cta aaa gac ccc aag 2304
Pro Glu Ile Asp Ile Lys Gly Pro Lys Leu Asp Leu Lys Asp Pro Lys
755 760 765

gtg gaa gtg aca gcc cct gat gtg gag gtt tct ctg ccc agc gtg gag 2352
Val Glu Val Thr Ala Pro Asp Val Glu Val Ser Leu Pro Ser Val Glu
770 775 780

gtg gat gtc gag gcc cca gga gcc aag ctg gat ggt gga cgg ctg gag 2400
Val Asp Val Glu Ala Pro Gly Ala Lys Leu Asp Gly Gly Arg Leu Glu
785 790 795 800

gag gac atg tcc ctg gcc gac aag gac ttg act acc aaa gac agc aag 2448
Glu Asp Met Ser Leu Ala Asp Lys Asp Leu Thr Thr Lys Asp Ser Lys
805 810 815

ttc aaa atg ccc aag ttc aag atg ccg tcg ttc ggg gtg tct gcc cca 2496
Phe Lys Met Pro Lys Phe Lys Met Pro Ser Phe Gly Val Ser Ala Pro
820 825 830

ggc aag tcc atc gag gcc tca gtg gat gtg tct gcg ccg aag gtg gag 2544
Gly Lys Ser Ile Glu Ala Ser Val Asp Val Ser Ala Pro Lys Val Glu
835 840 845

gcc gac gtg agc ctc ccc tcc atg cag ggg gac ctc aag gcc act gac 2592
Ala Asp Val Ser Leu Pro Ser Met Gln Gly Asp Leu Lys Ala Thr Asp
850 855 860

ctg agc ata cag ccc cct tct gct gac ctg gag gtc cag gct ggc caa 2640
Leu Ser Ile Gln Pro Pro Ser Ala Asp Leu Glu Val Gln Ala Gly Gln
865 870 875 880

gtg gac gtg aaa ctc cca gag ggc cct gtg tcc gag gga gcc ggc ctc 2688
Val Asp Val Lys Leu Pro Glu Gly Pro Val Ser Glu Gly Ala Gly Leu
885 890 895

aaa ggg cac ctg ccc aaa gtg cag atg ccc agt ttc aag atg ccc aaa 2736
Lys Gly His Leu Pro Lys Val Gln Met Pro Ser Phe Lys Met Pro Lys

900 905 910

gtg gac ctc aag ggg ccc cag ata gat gtt aag ggc ccc aag ctg gac 2784
Val Asp Leu Lys Gly Pro Gln Ile Asp Val Lys Gly Pro Lys Leu Asp
915 920 925

ctg aaa ggc ccc aag gtg gaa gtg aca gcc ccc gat gtg aag atg tct 2832
Leu Lys Gly Pro Lys Val Glu Val Thr Ala Pro Asp Val Lys Met Ser
930 935 940

ctg tcc agc atg gag gtg gac gtc cag gcc ccg aga gca aag ctg gat 2880
Leu Ser Ser Met Glu Val Asp Val Gln Ala Pro Arg Ala Lys Leu Asp
945 950 955 960

ggt gcg cag ctg gag ggg gac ctg tcc ctg gcc gac aag gcg gtg act 2928
Gly Ala Gln Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Ala Val Thr
965 970 975

gcc aaa gac agc aag ttc aaa atg ccc aag ttc aag atg cca tca ttt 2976
Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met Pro Ser Phe
980 985 990

ggg gtg tcg gcc cca ggc aag tcc atc gag gcc tcg gtg gat gtg tct 3024
Gly Val Ser Ala Pro Gly Lys Ser Ile Glu Ala Ser Val Asp Val Ser
995 1000 1005

gag ccg aag gtg gaa gct gat gtg agc ctc ccc tcc atg cag ggg 3069
Glu Pro Lys Val Glu Ala Asp Val Ser Leu Pro Ser Met Gln Gly
1010 1015 1020

gac ctg aag acc act gac ctc agc att cag tcc cct tcc gcc gac 3114
Asp Leu Lys Thr Thr Asp Leu Ser Ile Gln Ser Pro Ser Ala Asp
1025 1030 1035

ctg gag gtc cag gct ggc caa gtg aac gtg aaa ctc ccg gag ggc 3159
Leu Glu Val Gln Ala Gly Gln Val Asn Val Lys Leu Pro Glu Gly
1040 1045 1050

ccc ctt ccc gag gga gcc ggc ttc aaa ggg cac ctc ccc aag gtg 3204
Pro Leu Pro Glu Gly Ala Gly Phe Lys Gly His Leu Pro Lys Val
1055 1060 1065

cag atg ccc agt ttg aag atg ccc aaa gtg gcc ctc aag ggc ccc 3249
Gln Met Pro Ser Leu Lys Met Pro Lys Val Ala Leu Lys Gly Pro
1070 1075 1080

cag atg gac gtc aag ggc ccc aag ctg gac ctg aaa ggc ccc aag 3294
Gln Met Asp Val Lys Gly Pro Lys Leu Asp Leu Lys Gly Pro Lys
1085 1090 1095

gcg gag gtg atg gcc ccc gac gtg gag gtg tct ctg ccc agc gtg 3339
Ala Glu Val Met Ala Pro Asp Val Glu Val Ser Leu Pro Ser Val

1100 1105 1110

gag gtg gac gtc gag gct cca gga gcc aag ctg gac agt gtg cgg 3384
Glu Val Asp Val Glu Ala Pro Gly Ala Lys Leu Asp Ser Val Arg
1115 1120 1125

ctg gag ggt gac ctg tcc ctg gcc gac aag gat gtg act gcc aaa 3429
Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp Val Thr Ala Lys
1130 1135 1140

gac agc aag ttc aaa atg ccc aag ttc aag atg ccg tcg ttc ggg 3474
Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met Pro Ser Phe Gly
1145 1150 1155

gtg tct gcc cca ggc aag tcc atc gag gcc tcg gtg gat gtg tct 3519
Val Ser Ala Pro Gly Lys Ser Ile Glu Ala Ser Val Asp Val Ser
1160 1165 1170

gcg ccg aag gtg gag gcc gaa gtg agc ctc ccc tcc atg cag ggg 3564
Ala Pro Lys Val Glu Ala Glu Val Ser Leu Pro Ser Met Gln Gly
1175 1180 1185

gac ctc aag acc acg gac ctc tgc att ccg ctc cct tct gca gac 3609
Asp Leu Lys Thr Thr Asp Leu Cys Ile Pro Leu Pro Ser Ala Asp
1190 1195 1200

ctg gtg gtc cag gct ggc caa gtg gac atg aag ctc ccg gag ggc 3654
Leu Val Val Gln Ala Gly Gln Val Asp Met Lys Leu Pro Glu Gly
1205 1210 1215

cag gtg ccc gag gga gcc ggc ctc aaa ggg cac ttg ccc aag gtg 3699
Gln Val Pro Glu Gly Ala Gly Leu Lys Gly His Leu Pro Lys Val
1220 1225 1230

gat atg ccc agt ttc aag atg ccc aaa gtg gac ctc aag ggc ccc 3744
Asp Met Pro Ser Phe Lys Met Pro Lys Val Asp Leu Lys Gly Pro
1235 1240 1245

cag aca gat gtt aag ggc gcc aag ctg gac ctg aaa ggc ccc aag 3789
Gln Thr Asp Val Lys Gly Ala Lys Leu Asp Leu Lys Gly Pro Lys
1250 1255 1260

gcg gaa gtg aca gcc ccc gat gtc gag gtg tct ctg ccc agc atg 3834
Ala Glu Val Thr Ala Pro Asp Val Glu Val Ser Leu Pro Ser Met
1265 1270 1275

gag gtg gat gtc cag gcc cag aag gct aag ctg gat ggt gcg cgg 3879
Glu Val Asp Val Gln Ala Gln Lys Ala Lys Leu Asp Gly Ala Arg
1280 1285 1290

ctg gag gga gac ctg tcc ctg gcc gac aag gac atg act gcc aaa 3924
Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp Met Thr Ala Lys

1295 1300 1305

gac agc aag ttc aaa atg ccc aaa ttc aag atg ccg tcg ttc ggg 3969
Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met Pro Ser Phe Gly
1310 1315 1320

gta tcg gcc cca ggg agg tcc atc gag gcc tcg gtg gat gtg cct 4014
Val Ser Ala Pro Gly Arg Ser Ile Glu Ala Ser Val Asp Val Pro
1325 1330 1335

gca ccc aag gtg gag gcc gac gtg agt ctc ccc tcc atg cag ggg 4059
Ala Pro Lys Val Glu Ala Asp Val Ser Leu Pro Ser Met Gln Gly
1340 1345 1350

gac ctg aag acc act gac ctc agc att cag ccc cct tct gcc gac 4104
Asp Leu Lys Thr Thr Asp Leu Ser Ile Gln Pro Pro Ser Ala Asp
1355 1360 1365

ctg aag gtc cag act ggc cag gtg gat gtg aag ctc ccg gag ggc 4149
Leu Lys Val Gln Thr Gly Gln Val Asp Val Lys Leu Pro Glu Gly
1370 1375 1380

cac gtg ccc gag gga gct ggc ctc aaa ggg cac ctg ccc aag gtg 4194
His Val Pro Glu Gly Ala Gly Leu Lys Gly His Leu Pro Lys Val
1385 1390 1395

gag atg ccc agt ttg aag atg ccc aaa gtg gac ctc aag ggc ccc 4239
Glu Met Pro Ser Leu Lys Met Pro Lys Val Asp Leu Lys Gly Pro
1400 1405 1410

cag gtg gac atc aag ggc ccc aaa ctg gac cta aaa gac ccc aag 4284
Gln Val Asp Ile Lys Gly Pro Lys Leu Asp Leu Lys Asp Pro Lys
1415 1420 1425

gtg gaa atg aga gtc ccc gat gtc gag gtg tct ctg ccc agc atg 4329
Val Glu Met Arg Val Pro Asp Val Glu Val Ser Leu Pro Ser Met
1430 1435 1440

gag gtg gac gtc cag gcc cca aga gcc aag ctg gat agt gcg cat 4374
Glu Val Asp Val Gln Ala Pro Arg Ala Lys Leu Asp Ser Ala His
1445 1450 1455

ctg cag ggg gac ctg acc ctg gcc aac aag gac ctg act acc aaa 4419
Leu Gln Gly Asp Leu Thr Leu Ala Asn Lys Asp Leu Thr Thr Lys
1460 1465 1470

gac agc aag ttc aaa atg ccc aag ttc aag atg ccg tcg ttt ggg 4464
Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met Pro Ser Phe Gly
1475 1480 1485

gtg tct gcc cca ggc aag tcc atc gag gcc tcg gtg gat gtg tct 4509
Val Ser Ala Pro Gly Lys Ser Ile Glu Ala Ser Val Asp Val Ser

1490 1495 1500

cca ccc aag gtg gag gcc gac atc aag ggc ccc aag ctg gac cta 4554
Pro Pro Lys Val Glu Ala Asp Ile Lys Gly Pro Lys Leu Asp Leu
1505 1510 1515

aaa gac ccc aag gtg gaa gtg aca gcc cct gat gtg gag gtg tct 4599
Lys Asp Pro Lys Val Glu Val Thr Ala Pro Asp Val Glu Val Ser
1520 1525 1530

ctg ccc agc gtg gag gtg gac gtc aag gcc cca gga gcc aag ctg 4644
Leu Pro Ser Val Glu Val Asp Val Lys Ala Pro Gly Ala Lys Leu
1535 1540 1545

gat ggt gcg cgg ctg gag ggg gac atg tcc ctg gcc gac aag gac 4689
Asp Gly Ala Arg Leu Glu Gly Asp Met Ser Leu Ala Asp Lys Asp
1550 1555 1560

gtg act gcc aaa gac agc aag ttc aaa atg ccc aag ttc aag atg 4734
Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
1565 1570 1575

ctg tcg ttt ggg gtg tct gcc ctt ggc aag tcc atc gag gcc tca 4779
Leu Ser Phe Gly Val Ser Ala Leu Gly Lys Ser Ile Glu Ala Ser
1580 1585 1590

gcg gat gtg tct gcg ttg aag gtg gag gcc gac gtg agc ctc ccc 4824
Ala Asp Val Ser Ala Leu Lys Val Glu Ala Asp Val Ser Leu Pro
1595 1600 1605

tcc atg cag ggg gac ctc aag acc act gac ctc agc gtt cag ccc 4869
Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Leu Ser Val Gln Pro
1610 1615 1620

cct tcc gct gac ctg gag gtc cag gct ggc caa gtg gat gtg aaa 4914
Pro Ser Ala Asp Leu Glu Val Gln Ala Gly Gln Val Asp Val Lys
1625 1630 1635

ctc cca gag ggc ccc gtg ccg gag gga gcc ggc ctc aaa ggg cac 4959
Leu Pro Glu Gly Pro Val Pro Glu Gly Ala Gly Leu Lys Gly His
1640 1645 1650

ctg ccc aag ctg cag atg ccc agt ttc aag atg ccc aaa gta gat 5004
Leu Pro Lys Leu Gln Met Pro Ser Phe Lys Met Pro Lys Val Asp
1655 1660 1665

ctc aag ggc ccc cag ata gat gtc aag ggc ccc aag ctg gac ctg 5049
Leu Lys Gly Pro Gln Ile Asp Val Lys Gly Pro Lys Leu Asp Leu
1670 1675 1680

aaa ggc ccc aag acg gac gtg atg gcc ccc gac gtg gag gtg tct 5094
Lys Gly Pro Lys Thr Asp Val Met Ala Pro Asp Val Glu Val Ser

1685 1690 1695

cag ccc agc gtg gag gtg gat gtc gag gcc ccg gga gcc aag ctg 5139
Gln Pro Ser Val Glu Val Asp Val Glu Ala Pro Gly Ala Lys Leu
1700 1705 1710

gat ggt gcg tgg ctg gag ggg gac ctg tct gtg gcg gac aag gat 5184
Asp Gly Ala Trp Leu Glu Gly Asp Leu Ser Val Ala Asp Lys Asp
1715 1720 1725

gtg act acc aaa gac agc agg ttc aaa att ccc aag ttc aag atg 5229
Val Thr Thr Lys Asp Ser Arg Phe Lys Ile Pro Lys Phe Lys Met
1730 1735 1740

ccg tca ttc ggg gtg tct gcc cca ggc aag tcc atc gag gcc tcg 5274
Pro Ser Phe Gly Val Ser Ala Pro Gly Lys Ser Ile Glu Ala Ser
1745 1750 1755

gtg gat gtg tct gcg ccg aag gtg gag gcc gac ggg agc ctc tcc 5319
Val Asp Val Ser Ala Pro Lys Val Glu Ala Asp Gly Ser Leu Ser
1760 1765 1770

tcc atg cag ggg gac ctc aag gcc act gac ctc agc att cag ccc 5364
Ser Met Gln Gly Asp Leu Lys Ala Thr Asp Leu Ser Ile Gln Pro
1775 1780 1785

cct tcc gct gac ctg gag gtc cag gct ggc caa gtg gac gtg aaa 5409
Pro Ser Ala Asp Leu Glu Val Gln Ala Gly Gln Val Asp Val Lys
1790 1795 1800

ctc cca gag ggc cct gtg ccg gag gga gcc ggc ctc aaa ggg cac 5454
Leu Pro Glu Gly Pro Val Pro Glu Gly Ala Gly Leu Lys Gly His
1805 1810 1815

ctg ccc aag gtg cag atg ccc agt ttc aag atg cct gaa atg gac 5499
Leu Pro Lys Val Gln Met Pro Ser Phe Lys Met Pro Glu Met Asp
1820 1825 1830

ctc aag ggc ccc cag cta gat gtc aag ggc ccc aag ctg gac ctg 5544
Leu Lys Gly Pro Gln Leu Asp Val Lys Gly Pro Lys Leu Asp Leu
1835 1840 1845

aaa ggc ccc aag gcg gaa gtg aca gcc ccc gat gtg gag atg tct 5589
Lys Gly Pro Lys Ala Glu Val Thr Ala Pro Asp Val Glu Met Ser
1850 1855 1860

ctg tcc agc atg gag gtg gac gtc cag gcc ccg aga gca aag ctg 5634
Leu Ser Ser Met Glu Val Asp Val Gln Ala Pro Arg Ala Lys Leu
1865 1870 1875

gat ggt gcg cgg ctg gag ggg gac ctg tcc ctg gcc gac aag ggt 5679
Asp Gly Ala Arg Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Gly

1880 1885 1890

gtg aca gcc aaa gat agc aag ttc aaa atg ccc aag ttc aag atg 5724
Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
1895 1900 1905

cca tca ttc agg gtg tcg gcc cca ggc gag tcc atc gag gcg ttg 5769
Pro Ser Phe Arg Val Ser Ala Pro Gly Glu Ser Ile Glu Ala Leu
1910 1915 1920

gtg gat gtg tct gag ctg aag gtg gaa gcc gac atg agc ctc ccc 5814
Val Asp Val Ser Glu Leu Lys Val Glu Ala Asp Met Ser Leu Pro
1925 1930 1935

tcc atg caa ggg gac ctt aag acc act gac atc agc att cag ccc 5859
Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Ile Ser Ile Gln Pro
1940 1945 1950

ccc tct gcc caa ctg gag gtc cag gct ggc cag gtg gat gtg aaa 5904
Pro Ser Ala Gln Leu Glu Val Gln Ala Gly Gln Val Asp Val Lys
1955 1960 1965

ctc cca gag ggc cac gtt ccc gag gga gcc ggc ctc aaa ggg cac 5949
Leu Pro Glu Gly His Val Pro Glu Gly Ala Gly Leu Lys Gly His
1970 1975 1980

ctg ccc aag ctg cag atg ccc agt ttc aag atg cct gaa gtg gac 5994
Leu Pro Lys Leu Gln Met Pro Ser Phe Lys Met Pro Glu Val Asp
1985 1990 1995

ctc aag ggc ccc cag ata gat gtt aag ggc ccc aac gtg gac ctg 6039
Leu Lys Gly Pro Gln Ile Asp Val Lys Gly Pro Asn Val Asp Leu
2000 2005 2010

aaa ggc ccc aag gcg gaa gtg aca gcc ccc gat gtg aag atg tct 6084
Lys Gly Pro Lys Ala Glu Val Thr Ala Pro Asp Val Lys Met Ser
2015 2020 2025

ctg tcc agc atg gag gtg gac gtc cag gcc ccg aga gca aag ctg 6129
Leu Ser Ser Met Glu Val Asp Val Gln Ala Pro Arg Ala Lys Leu
2030 2035 2040

gat ggt gcg cgg ctg gag ggg gac ctg tcc ctg gcc gac aag ggc 6174
Asp Gly Ala Arg Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Gly
2045 2050 2055

atg aca gcc aaa gac agc aag ttc aaa atg ccc aag ttc aag atg 6219
Met Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
2060 2065 2070

ccg tca ttc ggg gtg tcg gcc cca ggc aag tcc atc gag gcc tcg 6264
Pro Ser Phe Gly Val Ser Ala Pro Gly Lys Ser Ile Glu Ala Ser

2075 2080 2085

gtg gat gtg tct gag ctg aag gtg gaa gct gac ggg agc ttc ccc 6309
Val Asp Val Ser Glu Leu Lys Val Glu Ala Asp Gly Ser Phe Pro
2090 2095 2100

tcc atg caa ggg gat ctt aag acc act gac atc cgc att cag ccc 6354
Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Ile Arg Ile Gln Pro
2105 2110 2115

ccc tcc gcc caa ctg gag gtc cag gct ggc cag gtg gac gtg aaa 6399
Pro Ser Ala Gln Leu Glu Val Gln Ala Gly Gln Val Asp Val Lys
2120 2125 2130

ctc cca gag ggc cac gtt ccc gag gga gcc ggc ctc aaa ggg cac 6444
Leu Pro Glu Gly His Val Pro Glu Gly Ala Gly Leu Lys Gly His
2135 2140 2145

ctg ccc aag gtg cag atg ccc agt ttc aag atg ccc aaa gtg gat 6489
Leu Pro Lys Val Gln Met Pro Ser Phe Lys Met Pro Lys Val Asp
2150 2155 2160

ctc aag ggc ccc cag ata gac gtc aag ggc ccc aag ctg gac ctg 6534
Leu Lys Gly Pro Gln Ile Asp Val Lys Gly Pro Lys Leu Asp Leu
2165 2170 2175

aaa ggc ccc aag gcg gag gtg acg gcc ccc gac gtg gag gtg tct 6579
Lys Gly Pro Lys Ala Glu Val Thr Ala Pro Asp Val Glu Val Ser
2180 2185 2190

ctg ccc agc gtg gag gtg gac gtc gag gcc ccg aga gca aag ctg 6624
Leu Pro Ser Val Glu Val Asp Val Glu Ala Pro Arg Ala Lys Leu
2195 2200 2205

gat ggt gca cgg ctg gag ggt gac ctg tcc ctg gcc gac aag gat 6669
Asp Gly Ala Arg Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp
2210 2215 2220

gtg act gcc aaa gac agc aag ttc aaa atg ccc aag ttc aag atg 6714
Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
2225 2230 2235

ccg tcg ttc ggg gtg tct gcc cca ggc aag tcc att gag gtc tcg 6759
Pro Ser Phe Gly Val Ser Ala Pro Gly Lys Ser Ile Glu Val Ser
2240 2245 2250

gtg gat gtg tct gcg ccg aag gtg gag gcc gaa gtg agc ctc ccc 6804
Val Asp Val Ser Ala Pro Lys Val Glu Ala Glu Val Ser Leu Pro
2255 2260 2265

tcc atg cag ggg gac ctg aag acc act gac atc agc att gag ccc 6849
Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Ile Ser Ile Glu Pro

2270 2275 2280

ccc tct gcc caa ctg gag gtc cag gct ggc cag gtg gac ctg aag 6894
Pro Ser Ala Gln Leu Glu Val Gln Ala Gly Gln Val Asp Leu Lys
2285 2290 2295

ctc cca gag ggc cac gtt ccc gag gga gct ggc ctc aaa ggg cac 6939
Leu Pro Glu Gly His Val Pro Glu Gly Ala Gly Leu Lys Gly His
2300 2305 2310

ctg ccc aag ttg cag atg ccc agt ttc aag atg ccc aaa gta gat 6984
Leu Pro Lys Leu Gln Met Pro Ser Phe Lys Met Pro Lys Val Asp
2315 2320 2325

cgc aag gga ccc cag ata gat gtc aag ggc ccc aag ctg gac ctg 7029
Arg Lys Gly Pro Gln Ile Asp Val Lys Gly Pro Lys Leu Asp Leu
2330 2335 2340

aaa ggc ccg aag acg gac gtg acg gcc ccc gac gtg gag gtg tct 7074
Lys Gly Pro Lys Thr Asp Val Thr Ala Pro Asp Val Glu Val Ser
2345 2350 2355

cag ccc ggc atg gag gtg gat gtc gag gcc cca gga gcc aag ttg 7119
Gln Pro Gly Met Glu Val Asp Val Glu Ala Pro Gly Ala Lys Leu
2360 2365 2370

gat ggt gca cgg ctg gag ggg gac ctg tcc ctg gcc gac aag gat 7164
Asp Gly Ala Arg Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp
2375 2380 2385

gtg act gcc aaa gac agc aag ttc aaa atg ccc aag ttc aag atg 7209
Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
2390 2395 2400

ccg tcg ttc ggg gtg tct gcc cca ggc aag tcc att gag gtc ttg 7254
Pro Ser Phe Gly Val Ser Ala Pro Gly Lys Ser Ile Glu Val Leu
2405 2410 2415

gtg gat gtg tct gcg cca aag gtg gag gcc gac ctg agc ctc ccc 7299
Val Asp Val Ser Ala Pro Lys Val Glu Ala Asp Leu Ser Leu Pro
2420 2425 2430

tcc atg cag ggg gac ctg aag aac act gac atc agc att gag ccc 7344
Ser Met Gln Gly Asp Leu Lys Asn Thr Asp Ile Ser Ile Glu Pro
2435 2440 2445

ccc tct gcc caa ctg gag gtc cag gct ggc cag gtg gac gtg aag 7389
Pro Ser Ala Gln Leu Glu Val Gln Ala Gly Gln Val Asp Val Lys
2450 2455 2460

ctc cca gag ggc cac gtt ctc gag gga gct ggc ctc aaa ggg cac 7434
Leu Pro Glu Gly His Val Leu Glu Gly Ala Gly Leu Lys Gly His

2465 2470 2475

ctg ccc aag ttg cag atg ccc agt ttc aag atg ccc aaa gta gat 7479
Leu Pro Lys Leu Gln Met Pro Ser Phe Lys Met Pro Lys Val Asp
2480 2485 2490

cgc aag ggc ccc cag ata gac atc aag ggc ccc aag ctg gac ctg 7524
Arg Lys Gly Pro Gln Ile Asp Ile Lys Gly Pro Lys Leu Asp Leu
2495 2500 2505

aaa ggc ccg aag atg gat gtg acg gcc ccc gac gtg gag gtg tct 7569
Lys Gly Pro Lys Met Asp Val Thr Ala Pro Asp Val Glu Val Ser
2510 2515 2520

cag ccc agc atg gag gtg gac gtc gag gcc cca gga gcc aag ttg 7614
Gln Pro Ser Met Glu Val Asp Val Glu Ala Pro Gly Ala Lys Leu
2525 2530 2535

gat ggt gca cgg ctg gag ggg gac ctg tcc ctg gcc gac aag gat 7659
Asp Gly Ala Arg Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp
2540 2545 2550

gtg act gcc aaa gac agc aag ttc aaa atg ccc aaa ttc aag atg 7704
Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
2555 2560 2565

ccg tcg tac agg gcg tct gcc cca ggc aag tcc atc cag gcc tcg 7749
Pro Ser Tyr Arg Ala Ser Ala Pro Gly Lys Ser Ile Gln Ala Ser
2570 2575 2580

gtg gat gtg tct gcg ccg aag gcg gag gcc gac gtg agc ctc ccc 7794
Val Asp Val Ser Ala Pro Lys Ala Glu Ala Asp Val Ser Leu Pro
2585 2590 2595

tcc atg cag ggg gac ctc aag acc act gac ctc agc att cag ctc 7839
Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Leu Ser Ile Gln Leu
2600 2605 2610

cct tct gtg gac ctg gag gtc cag gct ggc cag gtg gac gtg aag 7884
Pro Ser Val Asp Leu Glu Val Gln Ala Gly Gln Val Asp Val Lys
2615 2620 2625

ctc ccg gag ggc cac gtg ccc gag gga gct ggc ctc aaa ggg cac 7929
Leu Pro Glu Gly Hi s Val Pro Glu Gly Ala Gly Leu Lys Gly His
2630 2635 2640

ctg ccc aag gtg gag atg ccc agt ttc aag atg ccc aaa gtg gac 7974
Leu Pro Lys Val Glu Met Pro Ser Phe Lys Met Pro Lys Val Asp
2645 2650 2655

ctc aag agc ccc cag gtg gac atc aag ggc ccc aag ctg gac cta 8019
Leu Lys Ser Pro Gln Val Asp Ile Lys Gly Pro Lys Leu Asp Leu

2660 2665 2670

aaa gtc ccc aag gcg gaa gtg aca gtc cct gat gtg gag gtg tct 8064
Lys Val Pro Lys Ala Glu Val Thr Val Pro Asp Val Glu Val Ser
2675 2680 2685

ctg ccc agc gtg gag gtg gac gtc cag gcc ccg aga gcc aag ctg 8109
Leu Pro Ser Val Glu Val Asp Val Gln Ala Pro Arg Ala Lys Leu
2690 2695 2700

gat ggt gcg cgg ctg gag ggg gac ctg tcc ctg gct gaa aag gat 8154
Asp Gly Ala Arg Leu Glu Gly Asp Leu Ser Leu Ala Glu Lys Asp
2705 2710 2715

gtg act gcc aaa gac agc aag ttc aaa atg ccc aag ttc aag atg 8199
Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
2720 2725 2730

ccc tcc ttc ggg gtg tcg gcc cca ggc agg tcc atc gag gcc tcg 8244
Pro Ser Phe Gly Val Ser Ala Pro Gly Arg Ser Ile Glu Ala Ser
2735 2740 2745

ctg gat gtg tct gcg ccg aag gtg gag gcc gac gtg agc ctc tcc 8289
Leu Asp Val Ser Ala Pro Lys Val Glu Ala Asp Val Ser Leu Ser
2750 2755 2760

tcc atg cag ggg gac ctc aag gcc act gac ctc agc att cag ccc 8334
Ser Met Gln Gly Asp Leu Lys Ala Thr Asp Leu Ser Ile Gln Pro
2765 2770 2775

cct tcc gct gac ctg gag gtc cag gct gtc caa gtg gat gtg gaa 8379
Pro Ser Ala Asp Leu Glu Val Gln Ala Val Gln Val Asp Val Glu
2780 2785 2790

ctc ctg gag ggc ccc gtg ccc gag gga gcc ggc ctc aaa ggg cac 8424
Leu Leu Glu Gly Pro Val Pro Glu Gly Ala Gly Leu Lys Gly His
2795 2800 2805

ctg ccc aaa gtg gag atg ccc agt tta aag acg ccc aaa gtg gac 8469
Leu Pro Lys Val Glu Met Pro Ser Leu Lys Thr Pro Lys Val Asp
2810 2815 2820

ctc aag ggc ccc cag ata gat gtt aag ggc ccc aag ctg gac ctg 8514
Leu Lys Gly Pro Gln Ile Asp Val Lys Gly Pro Lys Leu Asp Leu
2825 2830 2835

aaa ggc ccc aag gca gaa gtg aga gtc ccc gat gtc gag gtg tct 8559
Lys Gly Pro Lys Ala Glu Val Arg Val Pro Asp Val Glu Val Ser
2840 2845 2850

ctg ccc agc gtg gag gtg gat gtc cag gcc ccg aag gcc aag ctg 8604
Leu Pro Ser Val Glu Val Asp Val Gln Ala Pro Lys Ala Lys Leu

2855 2860 2865

gat gct ggg cgg ctg gag gga gac ctg tcc ctg gct gac aag gac 8649
Asp Ala Gly Arg Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp
2870 2875 2880

gtg act gcc aaa gac agc aag ttc aaa atg ccc aaa ttc aag atg 8694
Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
2885 2890 2895

ccg tca ttc agg gta tcg gcc cca ggg aag tcc atg gag gcc tcg 8739
Pro Ser Phe Arg Val Ser Ala Pro Gly Lys Ser Met Glu Ala Ser
2900 2905 2910

gtg gat gtg tct gca ccc aag gtg gaa gcc gat gtg agt ctc ccc 8784
Val Asp Val Ser Ala Pro Lys Val Glu Ala Asp Val Ser Leu Pro
2915 2920 2925

tcc atg cag ggg gac ctg aag acc act gac ctc agc att cag ccc 8829
Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Leu Ser Ile Gln Pro
2930 2935 2940

cct tct gcc gac ctg aag gtc cag gct ggc cag atg gat gtg aag 8874
Pro Ser Ala Asp Leu Lys Val Gln Ala Gly Gln Met Asp Val Lys
2945 2950 2955

ctc ccg gag ggc cag gtg ccc gag gga gcc ggc ctc aaa gag cac 8919
Leu Pro Glu Gly Gln Val Pro Glu Gly Ala Gly Leu Lys Glu His
2960 2965 2970

ctg ccc aag gtg gag atg ccc agt ttg aag atg ccc aaa gtg gac 8964
Leu Pro Lys Val Glu Met Pro Ser Leu Lys Met Pro Lys Val Asp
2975 2980 2985

ctc aag ggc ccc cag gtg gac atc aag ggc ccc aag ctg gac cta 9009
Leu Lys Gly Pro Gln Val Asp Ile Lys Gly Pro Lys Leu Asp Leu
2990 2995 3000

aaa gtc tcc aag gcg gaa gtc aca gcc cct gat gtg gag gtg tct 9054
Lys Val Ser Lys Ala Glu Val Thr Ala Pro Asp Val Glu Val Ser
3005 3010 3015

ctg ccc agc gtg gag gtg gac gtc cag gcc cca aga gcc aaa ctg 9099
Leu Pro Ser Val Glu Val Asp Val Gln Ala Pro Arg Ala Lys Leu
3020 3025 3030

gat agt gca cag ctg gag ggg gac ctg tcc ctg gcc gac aag gat 9144
Asp Ser Ala Gln Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp
3035 3040 3045

gtg act gcc aaa gac agc aaa ttc aaa atg ccc aag ttc aag atg 9189
Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met

3050 3055 3060

ccg tca ttt ggg gtg tct gcc cca ggc aag tcc att gag gcc tcg 9234
Pro Ser Phe Gly Val Ser Ala Pro Gly Lys Ser Ile Glu Ala Ser
3065 3070 3075

gtg cac gtg tct gca ccc aag gtg gag gcc gat gtg agt ctc ccc 9279
Val His Val Ser Ala Pro Lys Val Glu Ala Asp Val Ser Leu Pro
3080 3085 3090

tcc atg cag ggg gac ctc aag acc act gac ctc agc att cag ccc 9324
Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Leu Ser Ile Gln Pro
3095 3100 3105

cat tct gcc gac ctg acg gtc caa gct cgc cag gtg gac atg aaa 9369
His Ser Ala Asp Leu Thr Val Gln Ala Arg Gln Val Asp Met Lys
3110 3115 3120

ctc ctg gag ggc cac gtg ccc gag gaa gcc ggc ctc aaa gga cac 9414
Leu Leu Glu Gly His Val Pro Glu Glu Ala Gly Leu Lys Gly His
3125 3130 3135

ctg ccc aag gtg cag atg ccc agt ttc aag atg ccc aaa gtc gac 9459
Leu Pro Lys Val Gln Met Pro Ser Phe Lys Met Pro Lys Val Asp
3140 3145 3150

ctc aag ggc cct gaa ata gac atc aag ggc ccc aag ctg gac cta 9504
Leu Lys Gly Pro Glu Ile Asp Ile Lys Gly Pro Lys Leu Asp Leu
3155 3160 3165

aaa gac ccc aag gtg gaa gtg aca gcc cct gat gtg gag gtt tct 9549
Lys Asp Pro Lys Val Glu Val Thr Ala Pro Asp Val Glu Val Ser
3170 3175 3180

ctg ccc agc gtg gag gtg gac gtc gag gcc cca gga gcc aag ctg 9594
Leu Pro Ser Val Glu Val Asp Val Glu Ala Pro Gly Ala Lys Leu
3185 3190 3195

gat ggt gcg cgg ctg gag ggg gac ctg tcc ctg gcc gac aag gac 9639
Asp Gly Ala Arg Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp
3200 3205 3210

atg acg gcc aaa gac agc aag ttc aaa atg ccc aag ttc aag atg 9684
Met Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
3215 3220 3225

ccg tcg ttc ggg gtg tct gcc cca ggc aag tcc atg gag gca tca 9729
Pro Ser Phe Gly Val Ser Ala Pro Gly Lys Ser Met Glu Ala Ser
3230 3235 3240

gtg gat gtg acc gcg cca aag gtg gag gcc gac gtg agc ctc cct 9774
Val Asp Val Thr Ala Pro Lys Val Glu Ala Asp Val Ser Leu Pro

3245 3250 3255

tcc atg cag ggg gac ctc aag gcc act gac ctc agc gtt cag ccc 9819
Ser Met Gln Gly Asp Leu Lys Ala Thr Asp Leu Ser Val Gln Pro
3260 3265 3270

cct tcc gct gac ctg gag gtc cag gct ggc caa gtg gac gtg aaa 9864
Pro Ser Ala Asp Leu Glu Val Gln Ala Gly Gln Val Asp Val Lys
3275 3280 3285

ctc cca gag ggc ccc gtg ccc gag gga gcc agc ctc aaa ggg cac 9909
Leu Pro Glu Gly Pro Val Pro Glu Gly Ala Ser Leu Lys Gly His
3290 3295 3300

ctg ccc aag gtg cag atg ccc agt ttc aag atg ccc aaa gtg gac 9954
Leu Pro Lys Val Gln Met Pro Ser Phe Lys Met Pro Lys Val Asp
3305 3310 3315

ctc aag ggc ccc cag ata gat gtt aag ggc ccc aag ctg gac ctg 9999
Leu Lys Gly Pro Gln Ile Asp Val Lys Gly Pro Lys Leu Asp Leu
3320 3325 3330

aaa ggc ccc aag gcg gaa gtg aca gcc cct gat gtg aag atg tct 10044
Lys Gly Pro Lys Ala Glu Val Thr Ala Pro Asp Val Lys Met Ser
3335 3340 3345

ctg tcc agc atg gag gtg gac gtc cag gcc ccg aga gca aag ctg 10089
Leu Ser Ser Met Glu Val Asp Val Gln Ala Pro Arg Ala Lys Leu
3350 3355 3360

gat ggt gtg cag ctg gag ggg gac ctg tcc ctg gcc gac aag gat 10134
Asp Gly Val Gln Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp
3365 3370 3375

gtg act gcc aaa gac agc aag ttc aaa atg ccc aag ttc aag atg 10179
Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
3380 3385 3390

cca tca ttc ggg gtg tcg gcc cca ggc aag tcc atg gag gcg tcc 10224
Pro Ser Phe Gly Val Ser Ala Pro Gly Lys Ser Met Glu Ala Ser
3395 3400 3405

gtg gat gtg tct gag ctg aag gcg aaa gcc gac gtg agc ctc ccc 10269
Val Asp Val Ser Glu Leu Lys Ala Lys Ala Asp Val Ser Leu Pro
3410 3415 3420

tcc atg cag ggg gac ctc aag acc act gac ctc agc att cag tcc 10314
Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Leu Ser Ile Gln Ser
3425 3430 3435

cct tcc gcc gac ctg gag gtc cag gct ggc caa gtg gac gtg aaa 10359
Pro Ser Ala Asp Leu Glu Val Gln Ala Gly Gln Val Asp Val Lys

3440 3445 3450

ctc ccg gag ggc ccc ctg ccc aag gga gcc ggc ctc aaa ggg cac 10404
Leu Pro Glu Gly Pro Leu Pro Lys Gly Ala Gly Leu Lys Gly His
3455 3460 3465

ctc ccc aag gtg cag atg ccc tgt ttg aag atg ccc aaa gtg gcc 10449
Leu Pro Lys Val Gln Met Pro Cys Leu Lys Met Pro Lys Val Ala
3470 3475 3480

ctc aag ggc ccc cag gtg gat gtc aag ggc ccc aag ctg gac ctg 10494
Leu Lys Gly Pro Gln Val Asp Val Lys Gly Pro Lys Leu Asp Leu
3485 3490 3495

aaa ggc ccc aag gcg gat gtg atg acc ccc gtc gtg gag gtg tct 10539
Lys Gly Pro Lys Ala Asp Val Met Thr Pro Val Val Glu Val Ser
3500 3505 3510

ctg ccc agc atg gag gtg gac gtc gag gcc ccg gga gcc aag ctg 10584
Leu Pro Ser Met Glu Val Asp Val Glu Ala Pro Gly Ala Lys Leu
3515 3520 3525

gac agt gtg cgg ctg gag ggt gac ctg tcc cta gcc gac aag gac 10629
Asp Ser Val Arg Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp
3530 3535 3540

atg act gcc aaa gac agc aag ttc aaa atg ccc aag ttc aag atg 10674
Met Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
3545 3550 3555

ccg tcg ttc ggg gtg tct gcc cca ggc aag tcc atc gag gcc tcg 10719
Pro Ser Phe Gly Val Ser Ala Pro Gly Lys Ser Ile Glu Ala Ser
3560 3565 3570

ttg gat gtg tct gcg ctg aag gtg gag gct gac gtg agc ctc ccc 10764
Leu Asp Val Ser Ala Leu Lys Val Glu Ala Asp Val Ser Leu Pro
3575 3580 3585

tcc atg cag ggg gac ctg aag acc act cac ctc agc att cag ccc 10809
Ser Met Gln Gly Asp Leu Lys Thr Thr His Leu Ser Ile Gln Pro
3590 3595 3600

cct tcc gct gat ctg gag gtc cag gct ggc caa gag gat gtg aaa 10854
Pro Ser Ala Asp Leu Glu Val Gln Ala Gly Gln Glu Asp Val Lys
3605 3610 3615

ctc cca gag ggc cct gtg cat gag gga gcc ggc ctc aaa ggg cac 10899
Leu Pro Glu Gly Pro Val His Glu Gly Ala Gly Leu Lys Gly His
3620 3625 3630

ctg ccg aag ctg cag atg ccc agt ttc aag gta ccc aaa gtg gac 10944
Leu Pro Lys Leu Gln Met Pro Ser Phe Lys Val Pro Lys Val Asp

3635 3640 3645

ctc aag ggt ccc cag ata gac gtt aat gtc ccc aag ctg gac ctg 10989
Leu Lys Gly Pro Gln Ile Asp Val Asn Val Pro Lys Leu Asp Leu
3650 3655 3660

aaa ggc ccc aag gtg gag gtg acg tcc ccc aac ctg gac gtg tct 11034
Lys Gly Pro Lys Val Glu Val Thr Ser Pro Asn Leu Asp Val Ser
3665 3670 3675

ctg ccc agc atg gag gtg gac atc caa gcc cca gga gcc aag ctg 11079
Leu Pro Ser Met Glu Val Asp Ile Gln Ala Pro Gly Ala Lys Leu
3680 3685 3690

gac agt acg cgg ctg gag ggg gac ctg tcc ctg gct gac aag gac 11124
Asp Ser Thr Arg Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp
3695 3700 3705

gtg act gcc aaa gac agc aag ttc aaa atg ccc aag ttc aag atg 11169
Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
3710 3715 3720

cca tcc ttt ggg atg ttg tcc cca ggc aag tcc atc gag gtc tcg 11214
Pro Ser Phe Gly Met Leu Ser Pro Gly Lys Ser Ile Glu Val Ser
3725 3730 3735

gtg gat gtg tct gcg cca aag atg gag gcc gac atg agc att ccc 11259
Val Asp Val Ser Ala Pro Lys Met Glu Ala Asp Met Ser Ile Pro
3740 3745 3750

tcc atg cag ggg gac ctc aag acc act gac ctc cgc att cag gcc 11304
Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Leu Arg Ile Gln Ala
3755 3760 3765

cct tcc gcc gac ctg gag gtc cag gct ggc cag gtg gac ttg aaa 11349
Pro Ser Ala Asp Leu Glu Val Gln Ala Gly Gln Val Asp Leu Lys
3770 3775 3780

ctt cca gaa ggc cac atg ccc gag gta gcc ggc ctc aaa ggg cac 11394
Leu Pro Glu Gly His Met Pro Glu Val Ala Gly Leu Lys Gly His
3785 3790 3795

ctg ccc aag gtg gag atg ccc agt ttc aag atg ccc aaa gtg gac 11439
Leu Pro Lys Val Glu Met Pro Ser Phe Lys Met Pro Lys Val Asp
3800 3805 3810

ctc aag ggc ccc cag gtg gac gtc aag ggc ccc aag ctg gac ctg 11484
Leu Lys Gly Pro Gln Val Asp Val Lys Gly Pro Lys Leu Asp Leu
3815 3820 3825

aaa ggc cca aag gca gag gtg atg gcc ccc gat gtg gag gtg tct 11529
Lys Gly Pro Lys Ala Glu Val Met Ala Pro Asp Val Glu Val Ser

3830 3835 3840

ctg ccc agc gtg gag acg gat gtc cag gcc cca gga tcc atg ctg 11574
Leu Pro Ser Val Glu Thr Asp Val Gln Ala Pro Gly Ser Met Leu
3845 3850 3855

gat ggt gcg cgg ctt gag ggg gac ctg tcc ctg gcc cac gag gat 11619
Asp Gly Ala Arg Leu Glu Gly Asp Leu Ser Leu Ala His Glu Asp
3860 3865 3870

gta gct ggg aaa gac agt aag ttt caa gga cca aaa ctg agc acg 11664
Val Ala Gly Lys Asp Ser Lys Phe Gln Gly Pro Lys Leu Ser Thr
3875 3880 3885

tct ggt ttt gaa tgg tcg tca aag aaa gtt tcc atg tct tcc tct 11709
Ser Gly Phe Glu Trp Ser Ser Lys Lys Val Ser Met Ser Ser Ser
3890 3895 3900

gaa atc gaa gga aat gtt aca ttc cat gag aag act tcc aca ttt 11754
Glu Ile Glu Gly Asn Val Thr Phe His Glu Lys Thr Ser Thr Phe
3905 3910 3915

ccc att gtg gaa tct gtt gtt cat gaa ggt gat ctt cat gat cca 11799
Pro Ile Val Glu Ser Val Val His Glu Gly Asp Leu His Asp Pro
3920 3925 3930

tct cgc gat ggt aac ttg ggg ctt gct gtt gga gaa gtt gga atg 11844
Ser Arg Asp Gly Asn Leu Gly Leu Ala Val Gly Glu Val Gly Met
3935 3940 3945

gat tcg aag ttt aag aaa ctg cat ttt aaa gtg ccc aaa gtt tca 11889
Asp Ser Lys Phe Lys Lys Leu His Phe Lys Val Pro Lys Val Ser
3950 3955 3960

ttt tct tct acc aaa act cct aaa gat agt tta gtc cca ggt gca 11934
Phe Ser Ser Thr Lys Thr Pro Lys Asp Ser Leu Val Pro Gly Ala
3965 3970 3975

aag tct agc ata ggt ctt tcc acg att cct tta tca tct tca gaa 11979
Lys Ser Ser Ile Gly Leu Ser Thr Ile Pro Leu Ser Ser Ser Glu
3980 3985 3990

tgc tca agt ttt gaa tta caa cag gtt tcg gct tgt tca gag cca 12024
Cys Ser Ser Phe Glu Leu Gln Gln Val Ser Ala Cys Ser Glu Pro
3995 4000 4005

tcc atg cag atg cct aag gtg ggt ttt gct ggg ttt cca tca tcc 12069
Ser Met Gln Met Pro Lys Val Gly Phe Ala Gly Phe Pro Ser Ser
4010 4015 4020

cgg ctt gat ctc act ggt cct cac ttt gaa tct tct att ctc tct 12114
Arg Leu Asp Leu Thr Gly Pro His Phe Glu Ser Ser Ile Leu Ser

4025 4030 4035

ccc tgt gag gat gtt aca ctt aca aaa tac cag gtg act gtt ccc 12159
Pro Cys Glu Asp Val Thr Leu Thr Lys Tyr Gln Val Thr Val Pro
4040 4045 4050

aga gct gcc ttg gcc cct gag ctt gct ctg gaa att cct tct ggg 12204
Arg Ala Ala Leu Ala Pro Glu Leu Ala Leu Glu Ile Pro Ser Gly
4055 4060 4065

tct cag gct gat att cct ctt ccc aag aca gag tgc tcc act gac 12249
Ser Gln Ala Asp Ile Pro Leu Pro Lys Thr Glu Cys Ser Thr Asp
4070 4075 4080

ctg cag cct cca gag gga gtt cca aca tct caa gct gag agt cac 12294
Leu Gln Pro Pro Glu Gly Val Pro Thr Ser Gln Ala Glu Ser His
4085 4090 4095

tct ggc cca ctg aat tcc atg att cct gtt tct ctt ggt cag gta 12339
Ser Gly Pro Leu Asn Ser Met Ile Pro Val Ser Leu Gly Gln Val
4100 4105 4110

tct ttt cct aaa ttc tat aaa cca aag ttt gtg ttt tca gtc ccc 12384
Ser Phe Pro Lys Phe Tyr Lys Pro Lys Phe Val Phe Ser Val Pro
4115 4120 4125

caa atg gca gtt cct gag gga gac cta cat gca gca gtg ggt gcc 12429
Gln Met Ala Val Pro Glu Gly Asp Leu His Ala Ala Val Gly Ala
4130 4135 4140

cca gtc atg tct cct ctt agc cct gga gaa aga gtg cag tgc ccc 12474
Pro Val Met Ser Pro Leu Ser Pro Gly Glu Arg Val Gln Cys Pro
4145 4150 4155

ttg cca agc acc cag ctg cca tcc cca ggc acc tgt gtg tct cag 12519
Leu Pro Ser Thr Gln Leu Pro Ser Pro Gly Thr Cys Val Ser Gln
4160 4165 4170

ggc cca gaa gag ctt gtg gcc tcc ttg cag aca tca gta gtg gcc 12564
Gly Pro Glu Glu Leu Val Ala Ser Leu Gln Thr Ser Val Val Ala
4175 4180 4185

cct gga gaa gcc cct tct gaa gat gct gac cac gaa ggg aaa ggg 12609
Pro Gly Glu Ala Pro Ser Glu Asp Ala Asp His Glu Gly Lys Gly
4190 4195 4200

agt ccc ttg aaa atg cct aag att aag ctt cca tca ttt agg tgg 12654
Ser Pro Leu Lys Met Pro Lys Ile Lys Leu Pro Ser Phe Arg Trp
4205 4210 4215

tcc ccg aag aag gaa aca ggg cca aag gtg gac cca gaa tgc agc 12699
Ser Pro Lys Lys Glu Thr Gly Pro Lys Val Asp Pro Glu Cys Ser

4220 4225 4230

gtg gag gac tca aaa ctc agc ctg gtt tta gac aag gat gaa gtg 12744
Val Glu Asp Ser Lys Leu Ser Leu Val Leu Asp Lys Asp Glu Val
4235 4240 4245

gcc ccg cag tct gcc atc cac atg gat ctg cct cct gag agg gat 12789
Ala Pro Gln Ser Ala Ile His Met Asp Leu Pro Pro Glu Arg Asp
4250 4255 4260

gga gag aag ggg agg agc aca aag cct ggc ttt gcc atg cca aaa 12834
Gly Glu Lys Gly Arg Ser Thr Lys Pro Gly Phe Ala Met Pro Lys
4265 4270 4275

ctt gca ctt ccc aaa atg aag gct tct aag agt ggg gtc agc ctg 12879
Leu Ala Leu Pro Lys Met Lys Ala Ser Lys Ser Gly Val Ser Leu
4280 4285 4290

cca cag aga gac gtg gat cct tcc ctt tct agt gcc aca gca ggg 12924
Pro Gln Arg Asp Val Asp Pro Ser Leu Ser Ser Ala Thr Ala Gly
4295 4300 4305

ggt agc ttt caa gac aca gaa aag gcc agc agt gac ggt ggt agg 12969
Gly Ser Phe Gln Asp Thr Glu Lys Ala Ser Ser Asp Gly Gly Arg
4310 4315 4320

gga gga ctt ggt gca aca gca agt gcc aca gga agt gag ggt gtg 13014
Gly Gly Leu Gly Ala Thr Ala Ser Ala Thr Gly Ser Glu Gly Val
4325 4330 4335

aac ctc cac cgg cca cag gtc cac att ccc agt ttg ggc ttt gcc 13059
Asn Leu His Arg Pro Gln Val His Ile Pro Ser Leu Gly Phe Ala
4340 4345 4350

aaa cct gat ctc aga tcc tcc aag gcc aag gtg gag gtg agc cag 13104
Lys Pro Asp Leu Arg Ser Ser Lys Ala Lys Val Glu Val Ser Gln
4355 4360 4365

cct gaa gct gac ctg cct ctt ccc aaa cat gat ctg tct acc gaa 13149
Pro Glu Ala Asp Leu Pro Leu Pro Lys His Asp Leu Ser Thr Glu
4370 4375 4380

ggt gac agc aga gga tgt ggg ctc ggg gat gtc cca gtg agc cag 13194
Gly Asp Ser Arg Gly Cys Gly Leu Gly Asp Val Pro Val Ser Gln
4385 4390 4395

cct tgt ggg gag ggg ata gcc ccc aca cct gaa gat ccc ctc cag 13239
Pro Cys Gly Glu Gly Ile Ala Pro Thr Pro Glu Asp Pro Leu Gln
4400 4405 4410

cca tcc tgt aga aaa cca gat gct gaa gtc ctc aca gtg gaa agc 13284
Pro Ser Cys Arg Lys Pro Asp Ala Glu Val Leu Thr Val Glu Ser

4415 4420 4425

cca gag gag gaa gcc atg acc aag tac tcg cag gaa agc tgg ttt 13329
Pro Glu Glu Glu Ala Met Thr Lys Tyr Ser Gln Glu Ser Trp Phe
4430 4435 4440

aaa atg ccc aag ttc cgc atg ccc agc ctt agg cgc tct ttc agg 13374
Lys Met Pro Lys Phe Arg Met Pro Ser Leu Arg Arg Ser Phe Arg
4445 4450 4455

gac aga ggc ggg gct gga aag ctg gaa gtg gct cag aca cag gca 13419
Asp Arg Gly Gly Ala Gly Lys Leu Glu Val Ala Gln Thr Gln Ala
4460 4465 4470

ccg gca gca aca ggg ggt gaa gca gca gct aaa gtc aaa gag ttc 13464
Pro Ala Ala Thr Gly Gly Glu Ala Ala Ala Lys Val Lys Glu Phe
4475 4480 4485

ctt gtt tct ggg tca aac gtg gag gca gct atg tcc cta cag ctc 13509
Leu Val Ser Gly Ser Asn Val Glu Ala Ala Met Ser Leu Gln Leu
4490 4495 4500

cca gag gca gat gca gaa gtg aca gct tct gag agc aaa tca tcc 13554
Pro Glu Ala Asp Ala Glu Val Thr Ala Ser Glu Ser Lys Ser Ser
4505 4510 4515

aca gat att cta agg tgt gat ctt gac agc aca ggc ttg aag ctg 13599
Thr Asp Ile Leu Arg Cys Asp Leu Asp Ser Thr Gly Leu Lys Leu
4520 4525 4530

cac ctc tcc act gct ggg atg act ggg gat gag ctt tcc act tct 13644
His Leu Ser Thr Ala Gly Met Thr Gly Asp Glu Leu Ser Thr Ser
4535 4540 4545

gag gtc agg atc cat cca tcc aaa gga cct ctc cct ttt cag atg 13689
Glu Val Arg Ile His Pro Ser Lys Gly Pro Leu Pro Phe Gln Met
4550 4555 4560

cct ggc atg agg ctt cca gaa acc cag gtt ctt cca gga gaa ata 13734
Pro Gly Met Arg Leu Pro Glu Thr Gln Val Leu Pro Gly Glu Ile
4565 4570 4575

gat gag act cct ctt tcc aag cca gga cat gac ctt gcc agc atg 13779
Asp Glu Thr Pro Leu Ser Lys Pro Gly His Asp Leu Ala Ser Met
4580 4585 4590

gag gat aaa aca gag aaa tgg tct tcc cag cct gaa ggt cca ctt 13824
Glu Asp Lys Thr Glu Lys Trp Ser Ser Gln Pro Glu Gly Pro Leu
4595 4600 4605

aaa ttg aaa gct tca agt act gat atg cca tcc cag att tct gtg 13869
Lys Leu Lys Ala Ser Ser Thr Asp Met Pro Ser Gln Ile Ser Val

4610 4615 4620

gtt aat gtg gat caa ctg tgg gaa gat tct gtc cta act gtc aaa 13914
Val Asn Val Asp Gln Leu Trp Glu Asp Ser Val Leu Thr Val Lys
4625 4630 4635

ttc ccc aaa tta atg gta cca agg ttc tcc ttc cct gcc ccc agc 13959
Phe Pro Lys Leu Met Val Pro Arg Phe Ser Phe Pro Ala Pro Ser
4640 4645 4650

tca gag gat gat gtg ttc atc ccc act gtg agg gaa gtg cag tgt 14004
Ser Glu Asp Asp Val Phe Ile Pro Thr Val Arg Glu Val Gln Cys
4655 4660 4665

cca gag gcc aat att gat aca gcc ctt tgt aag gaa agt ccg ggg 14049
Pro Glu Ala Asn Ile Asp Thr Ala Leu Cys Lys Glu Ser Pro Gly
4670 4675 4680

ctc tgg gga gcc agc atc ctg aag gca ggt gct ggg gtc cct ggg 14094
Leu Trp Gly Ala Ser Ile Leu Lys Ala Gly Ala Gly Val Pro Gly
4685 4690 4695

gag cag cct gtg gac ctt aac ctg cct ttg gaa gct ccc cca att 14139
Glu Gln Pro Val Asp Leu Asn Leu Pro Leu Glu Ala Pro Pro Ile
4700 4705 4710

tca aag gtc aga gtg cat att cag ggt gct cag gtt gaa agt caa 14184
Ser Lys Val Arg Val His Ile Gln Gly Ala Gln Val Glu Ser Gln
4715 4720 4725

gag gtc act ata cac agc ata gtg aca cca gag ttt gta gat ctc 14229
Glu Val Thr Ile His Ser Ile Val Thr Pro Glu Phe Val Asp Leu
4730 4735 4740

tca gta ccc agg act ttt tcc act cag att gtg cgg gaa tca gag 14274
Ser Val Pro Arg Thr Phe Ser Thr Gln Ile Val Arg Glu Ser Glu
4745 4750 4755

atc ccc acg tca gag att caa aca cct tcg tac gga ttt tcc tta 14319
Ile Pro Thr Ser Glu Ile Gln Thr Pro Ser Tyr Gly Phe Ser Leu
4760 4765 4770

tta aaa gtg aaa atc cca gag ccc cac acg cag gct aga gtg tac 14364
Leu Lys Val Lys Ile Pro Glu Pro His Thr Gln Ala Arg Val Tyr
4775 4780 4785

aca aca atg act caa cac tct agg act cag gag ggc aca gaa gag 14409
Thr Thr Met Thr Gln His Ser Arg Thr Gln Glu Gly Thr Glu Glu
4790 4795 4800

gct ccc ata caa gcc acc cca gga gta gac tcc att tct gga gat 14454
Ala Pro Ile Gln Ala Thr Pro Gly Val Asp Ser Ile Ser Gly Asp

4805 4810 4815

ctc cag cct gac act gga gaa cca ttt gag atg atc tct tcc agc 14499
Leu Gln Pro Asp Thr Gly Glu Pro Phe Glu Met Ile Ser Ser Ser
4820 4825 4830

gtc aat gta ctg gga cag caa aca ctc aca ttt gaa gtt cct tct 14544
Val Asn Val Leu Gly Gln Gln Thr Leu Thr Phe Glu Val Pro Ser
4835 4840 4845

ggc cac cag ctt gca gac agc tgt tca gat gag gag cca gca gaa 14589
Gly His Gln Leu Ala Asp Ser Cys Ser Asp Glu Glu Pro Ala Glu
4850 4855 4860

att ctt gag ttt ccc cct gat gat agc caa gag gca acc aca cca 14634
Ile Leu Glu Phe Pro Pro Asp Asp Ser Gln Glu Ala Thr Thr Pro
4865 4870 4875

ctg gca gat gaa ggc agg gct cca aaa gac aaa cca gaa agt aaa 14679
Leu Ala Asp Glu Gly Arg Ala Pro Lys Asp Lys Pro Glu Ser Lys
4880 4885 4890

aaa tct ggt ctg ctc tgg ttt tgg ctt cca aac att ggg ttt tcc 14724
Lys Ser Gly Leu Leu Trp Phe Trp Leu Pro Asn Ile Gly Phe Ser
4895 4900 4905

tct tct gtt gat gag aca ggt gtt gat tcc aaa aat gac gtc cag 14769
Ser Ser Val Asp Glu Thr Gly Val Asp Ser Lys Asn Asp ValGln
4910 4915 4920

aga tct gct ccc att caa aca cag cct gag gca cga cca gag gca 14814
Arg Ser Ala Pro Ile Gln Thr Gln Pro Glu Ala Arg Pro Glu Ala
4925 4930 4935

gaa ctg cct aaa aaa cag gag aag gca ggc tgg ttc cga ttt ccc 14859
Glu Leu Pro Lys Lys Gln Glu Lys Ala Gly Trp Phe Arg Phe Pro
4940 4945 4950

aaa tta ggg ttc tcc tca tct cct acc aag aaa agc aaa agc acc 14904
Lys Leu Gly Phe Ser Ser Ser Pro Thr Lys Lys Ser Lys Ser Thr
4955 4960 4965

gaa gat ggg gca gag ctg gaa gaa caa aaa ctt caa gaa gaa aca 14949
Glu Asp Gly Ala Glu Leu Glu Glu Gln Lys Leu Gln Glu Glu Thr
4970 4975 4980

atc acg ttt ttt gat gcc cga gaa agt ttc tcc cct gaa gag aag 14994
Ile Thr Phe Phe Asp Ala Arg Glu Ser Phe Ser Pro Glu Glu Lys
4985 4990 4995

gaa gag ggt gaa ctg atc ggg cct gtg ggc act ggg ctg gac tcc 15039
Glu Glu Gly Glu Leu Ile Gly Pro Val Gly Thr Gly Leu Asp Ser

5000 5005 5010

aga gtg atg gtg aca tcc gcg gca aga aca gag tta atc ctg ccc 15084
Arg Val Met Val Thr Ser Ala Ala Arg Thr Glu Leu Ile Leu Pro
5015 5020 5025

gag cag gac aga aaa gct gac gat gaa agc aaa ggg tca ggc ctg 15129
Glu Gln Asp Arg Lys Ala Asp Asp Glu Ser Lys Gly Ser Gly Leu
5030 5035 5040

gga cca aat gaa ggc tga gaggtatggc tcatcagtac aagagagatg 15177
Gly Pro Asn Glu Gly
5045

caaaaaacta agttggaaag taaaggctac acacacatat ggagcacccc atcccacagc 15237
acattacatc cacctcactt cacagaacgg agaacagagc agaaatgacc agaacacctt 15297
tgtcaccatc acacagccct cctaaaatgg aaccaaagct tcccagctcc ctcaaagctt 15357
tggatgcaaa gaaggcaccc tgacttccac aagacaccag aattcacacg gtactcagag 15417
gcactgctgg ggaagtttgt tggtctttat tagataaatt tccagagacc tgtccataat 15477
acccaacaga acatgactgt ttctttgagg aaagggttat aatgtctgtg gtgtacaagt 15537
cgtttttggt ataacttctt tcctgctgct gctgcttccc ggcaaacata gttttcctat 15597
ttcaggcaga gtgcggtata ttccaggaaa cactgtttcc tactcactta gcttacttct 15657
ttgttgaatg cctcactaat ggcaagtttc aagatgtttt gggtgacaat gcacacatgc 15717
tgggcaaaag ggtgatggcc agtggctggc agctgggcca gcagaagcta ggacatctgt 15777
gagttgtcat tctcatctat ccatgtccac tggcctgcca gcatccgcca gtgccttgcc 15837
agtgtgcacg gtcccacact gtggcccctg agtcccctaa tgtacacgct gcagccagaa 15897
tgcagatgga gctggcttgg ctgttccctg gatgggcaat aaagaaagtg ctgcatccca 15957
t
15958

<210>2
<211>5048
<212>PRT
<213> 人类

<400>2
Met Pro Lys Lye Lys Met Pro Leu Phe Gly Ala Ser Ala Pro Gly Lys
1 5 10 15
Ser Met Glu Ala Ser Val Asp Val Ser Ala Pro Lys Val Glu Ala Asp
20 25 30
Val Ser Leu Leu Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Leu Ser
35 40 45
Val Gln Thr Pro Ser Ala Asp Leu Glu Val Gln Asp Gly Gln Val Asp
50 55 60

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

370 375 380

Gln Val Asp Val Lys Leu Pro Glu Gly His Leu Pro Glu Gly Ala Gly

385 390 395 400

Leu Lys Gly His Leu Pro Lys Val Glu Met Pro Ser Phe Lys Met Pro

405 410 415

Lys Val Ala Leu Lys Gly Pro Gln Val Asp Val Lys Gly Pro Lys Leu

420 425 430

Asp Leu Lys Ser Pro Lys Ala Glu Val Thr Ala Pro Asp Val Glu Val

435 440 445

Ser Leu Pro Ser Val Glu Val Asp Val Glu Ala Pro Gly Ala Lys Leu

450 455 460

Asp Ser Ala Arg Leu Glu Gly Glu Leu Ser Leu Ala Asp Lys Asp Val

465 470 475 480

Thr Ala Lys Asp Ser Arg Phe Lys Met Pro Lys Phe Lys Met Pro Ser

485 490 495

Phe Gly Ala Ser Ala Pro Gly Lys Ser Ile Glu Ala Ser Val Asp Val

500 505 510

Ser Ala Pro Lys Val Glu Ala Asp Val Ser Leu Pro Ser Met Gln Gly

515 520 525

Asp Leu Lys Thr Thr Asp Leu Ser Ile Gln Pro Pro Ser Ala Asp Leu

530 535 540

Glu Val His Ala Gly Gln Val Asp Val Lys Leu Leu Glu Gly His Val

545 550 555 560

Pro Glu Gly Ala Gly Phe Lys Gly His Leu Pro Lys Val Gln Met Pro

565 570 575

Ser Leu Lys Met Pro Lys Val Asp Leu Lys Gly Pro Gln Val Glu Val

580 585 590

Arg Gly Pro Lys Leu Asp Leu Lys Gly His Lys Ala Glu Val Thr Ala

595 600 605

His Glu Val Ala Val Ser Leu Pro Ser Val Glu Val Asp Met Gln Ala

610 615 620

Pro Gly Ala Lys Leu Asp Gly Ala Gln Leu Asp Gly Asp Leu Ser Leu

625 630 635 640

Ala Asp Lys Asp Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys

645 650 655

Phe Lys Met Pro Ser Phe Gly Val Ser Ala Pro Gly Lys Ser Ile Glu

660 665 670

Ala Ser Val Asp Leu Ser Ala Pro Lys Val Glu Ala Asp Met Ser Leu

675 680 685

Pro Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Leu Ser Ile Gln Pro
690 695 700
Pro Ser Thr Asp Leu Glu Leu Gln Ala Gly Gln Leu Asp Val Lys Leu
705 710 715 720
Pro Glu Gly Pro Val Pro Glu Gly Ala Gly Leu Lys Gly His Leu Pro
725 730 735
Lys Leu Gln Met Pro Ser Phe Lys Val Pro Lys Val Asp Leu Lys Gly
740 745 750
Pro Glu Ile Asp Ile Lys Gly Pro Lys Leu Asp Leu Lys Asp Pro Lys
755 760 765
Val Glu Val Thr Ala Pro Asp Val Glu Val Ser Leu Pro Ser Val Glu
770 775 780
Val Asp Val Glu Ala Pro Gly Ala Lys Leu Asp Gly Gly Arg Leu Glu
785 790 795 800
Glu Asp Met Ser Leu Ala Asp Lys Asp Leu Thr Thr Lys Asp Ser Lys
805 810 815
Phe Lys Met Pro Lys Phe Lys Met Pro Ser Phe Gly Val Ser Ala Pro
820 825 830
Gly Lys Ser Ile Glu Ala Ser Val Asp Val Ser Ala Pro Lys Val Glu
835 840 845
Ala Asp Val Ser Leu Pro Ser Met Gln Gly Asp Leu Lys Ala Thr Asp
850 855 860
Leu Ser Ile Gln Pro Pro Ser Ala Asp Leu Glu Val Gln Ala Gly Gln
865 870 875 880
Val Asp Val Lys Leu Pro Glu Gly Pro Val Ser Glu Gly Ala Gly Leu
885 890 895
Lys Gly His Leu Pro Lys Val Gln Met Pro Ser Phe Lys Met Pro Lys
900 905 910
Val Asp Leu Lys Gly Pro Gln Ile Asp Val Lys Gly Pro Lys Leu Asp
915 920 925
Leu Lys Gly Pro Lys Val Glu Val Thr Ala Pro Asp Val Lys Met Ser
930 935 940
Leu Ser Ser Met Glu Val Asp Val Gln Ala Pro Arg Ala Lys Leu Asp
945 950 955 960
Gly Ala Gln Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Ala Val Thr
965 970 975
Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met Pro Ser Phe
980 985 990
Gly Val Ser Ala Pro Gly Lys Ser Ile Glu Ala Ser Val Asp Val Ser

995 1000 1005

Glu Pro Lys Val Glu Ala Asp Val Ser Leu Pro Ser Met Gln Gly
1010 1015 1020

Asp Leu Lys Thr Thr Asp Leu Ser Ile Gln Ser Pro Ser Ala Asp
1025 1030 1035

Leu Glu Val Gln Ala Gly Gln Val Asn Val Lys Leu Pro Glu Gly
1040 1045 1050

Pro Leu Pro Glu Gly Ala Gly Phe Lys Gly His Leu Pro Lys Val
1055 1060 1065

Gln Met Pro Ser Leu Lys Met Pro Lys Val Ala Leu Lys Gly Pro
1070 1075 1080

Gln Met Asp Val Lys Gly Pro Lys Leu Asp Leu Lys Gly Pro Lys
1085 1090 1095

Ala Glu Val Met Ala Pro Asp Val Glu Val Ser Leu Pro Ser Val
1100 1105 1110

Glu Val Asp Val Glu Ala Pro Gly Ala Lys Leu Asp Ser Val Arg
1115 1120 1125

Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp Val Thr Ala Lys
1130 1135 1140

Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met Pro Ser Phe Gly
1145 1150 1155

Val Ser Ala Pro Gly Lys Ser Ile Glu Ala Ser Val Asp Val Ser
1160 1165 1170

Ala Pro Lys Val Glu Ala Glu Val Ser Leu Pro Ser Met Gln Gly
1175 1180 1185

Asp Leu Lys Thr Thr Asp Leu Cys Ile Pro Leu Pro Ser Ala Asp
1190 1195 1200

Leu Val Val Gln Ala Gly Gln Val Asp Met Lys Leu Pro Glu Gly
1205 1210 1215

Gln Val Pro Glu Gly Ala Gly Leu Lys Gly His Leu Pro Lys Val
1220 1225 1230

Asp Met Pro Ser Phe Lys Met Pro Lys Val Asp Leu Lys Gly Pro
1235 1240 1245

Gln Thr Asp Val Lys Gly Ala Lys Leu Asp Leu Lys Gly Pro Lys
1250 1255 1260

Ala Glu Val Thr Ala Pro Asp Val Glu Val Ser Leu Pro Ser Met
1265 1270 1275

Glu Val Asp Val Gln Ala Gln Lys Ala Lys Leu Asp Gly Ala Arg
1280 1285 1290

Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp Met Thr Ala Lys
1295 1300 1305
Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met Pro Ser Phe Gly
1310 1315 1320
Val Ser Ala Pro Gly Arg Ser Ile Glu Ala Ser Val Asp Val Pro
1325 1330 1335
Ala Pro Lys Val Glu Ala Asp Val Ser Leu Pro Ser Met Gln Gly
1340 1345 1350
Asp Leu Lys Thr Thr Asp Leu Ser Ile Gln Pro Pro Ser Ala Asp
1355 1360 1365
Leu Lys Val Gln Thr Gly Gln Val Asp Val Lys Leu Pro Glu Gly
1370 1375 1380
His Val Pro Glu Gly Ala Gly Leu Lys Gly His Leu Pro Lys Val
1385 1390 1395
Glu Met Pro Ser Leu Lys Met Pro Lys Val Asp Leu Lys Gly Pro
1400 1405 1410
Gln Val Asp Ile Lys Gly Pro Lys Leu Asp Leu Lys Asp Pro Lys
1415 1420 1425
Val Glu Met Arg Val Pro Asp Val Glu Val Ser Leu Pro Ser Met
1430 1435 1440
Glu Val Asp Val Gln Ala Pro Arg Ala Lys Leu Asp Ser Ala His
1445 1450 1455
Leu Gln Gly Asp Leu Thr Leu Ala Asn Lys Asp Leu Thr Thr Lys
1460 1465 1470
Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met Pro Ser Phe Gly
1475 1480 1485
Val Ser Ala Pro Gly Lys Ser Ile Glu Ala Ser Val Asp Val Ser
1490 1495 1500
Pro Pro Lys Val Glu Ala Asp Ile Lys Gly Pro Lys Leu Asp Leu
1505 1510 1515
Lys Asp Pro Lys Val Glu Val Thr Ala Pro Asp Val Glu Val Ser
1520 1525 1530
Leu Pro Ser Val Glu Val Asp Val Lys Ala Pro Gly Ala Lys Leu
1535 1540 1545
Asp Gly Ala Arg Leu Glu Gly Asp Met Ser Leu Ala Asp Lys Asp
1550 1555 1560
Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
1565 1570 1575
Leu Ser Phe Gly Val Ser Ala Leu Gly Lys Ser Ile Glu Ala Ser

1580 1585 1590

Ala Asp Val Ser Ala Leu Lys Val Glu Ala Asp Val Ser Leu Pro

1595 1600 1605

Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Leu Ser Val Gln Pro

1610 1615 1620

Pro Ser Ala Asp Leu Glu Val Gln Ala Gly Gln Val Asp Val Lys

1625 1630 1635

Leu Pro Glu Gly Pro Val Pro Glu Gly Ala Gly Leu Lys Gly His

1640 1645 1650

Leu Pro Lys Leu Gln Met Pro Ser Phe Lys Met Pro Lys Val Asp

1655 1660 1665

Leu Lys Gly Pro Gln Ile Asp Val Lys Gly Pro Lys Leu Asp Leu

1670 1675 1680

Lys Gly Pro Lys Thr Asp Val Met Ala Pro Asp Val Glu Val Ser

1685 1690 1695

Gln Pro Ser Val Glu Val Asp Val Glu Ala Pro Gly Ala Lys Leu

1700 1705 1710

Asp Gly Ala Trp Leu Glu Gly Asp Leu Ser Val Ala Asp Lys Asp

1715 1720 1725

Val Thr Thr Lys Asp Ser Arg Phe Lys Ile Pro Lys Phe Lys Met

1730 1735 1740

Pro Ser Phe Gly Val Ser Ala Pro Gly Lys Ser Ile Glu Ala Ser

1745 1750 1755

Val Asp Val Ser Ala Pro Lys Val Glu Ala Asp Gly Ser Leu Ser

1760 1765 1770

Ser Met Gln Gly Asp Leu Lys Ala Thr Asp Leu Ser Ile Gln Pro

1775 1780 1785

Pro Ser Ala Asp Leu Glu Val Gln Ala Gly Gln Val Asp Val Lys

1790 1795 1800

Leu Pro Glu Gly Pro Val Pro Glu Gly Ala Gly Leu Lys Gly His

1805 1810 1815

Leu Pro Lys Val Gln Met Pro Ser Phe Lys Met Pro Glu Met Asp

1820 1825 1830

Leu Lys Gly Pro Gln Leu Asp Val Lys Gly Pro Lys Leu Asp Leu

1835 1840 1845

Lys Gly Pro Lys Ala Glu Val Thr Ala Pro Asp Val Glu Met Ser

1850 1855 1860

Leu Ser Ser Met Glu Val Asp Val Gln Ala Pro Arg Ala Lys Leu

1865 1870 1875

Asp Gly Ala Arg Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Gly
1880 1885 1890

Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
1895 1900 1905

Pro Ser Phe Arg Val Ser Ala Pro Gly Glu Ser Ile Glu Ala Leu
1910 1915 1920

Val Asp Val Ser Glu Leu Lys Val Glu Ala Asp Met Ser Leu Pro
1925 1930 1935

Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Ile Ser Ile Gln Pro
1940 1945 1950

Pro Ser Ala Gln Leu Glu Val Gln Ala Gly Gln Val Asp Val Lys
1955 1960 1965

Leu Pro Glu Gly His Val Pro Glu Gly Ala Gly Leu Lys Gly His
1970 1975 1980

Leu Pro Lsy Leu Gln Met Pro Ser Phe Lys Met Pro Glu Val Asp
1985 1990 1995

Leu Lys Gly Pro Gln Ile Asp Val Lys Gly Pro Asn Val Asp Leu
2000 2005 2010

Lys Gly Pro Lys Ala Glu Val Thr Ala Pro Asp Val Lys Met Ser
2015 2020 2025

Leu Ser Ser Met Glu Val Asp Val Gln Ala Pro Arg Ala Lys Leu
2030 2035 2040

Asp Gly Ala Arg Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Gly
2045 2050 2055

Met Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
2060 2065 2070

Pro Ser Phe Gly Val Ser Ala Pro Gly Lys Ser Ile Glu Ala Ser
2075 2080 2085

Val Asp Val Ser Glu Leu Lys Val Glu Ala Asp Gly Ser Phe Pro
2090 2095 2100

Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Ile Arg Ile Gln Pro
2105 2110 2115

Pro Ser Ala Gln Leu Glu Val Gln Ala Gly Gln Val Asp Val Lys
2120 2125 2130

Leu Pro Glu Gly His Val Pro Glu Gly Ala Gly Leu Lys Gly His
2135 2140 2145

Leu Pro Lys Val Gln Met Pro Ser Phe Lys Met Pro Lys Val Asp
2150 2155 2160

Leu Lys Gly Pro Gln Ile Asp Val Lys Gly Pro Lys Leu Asp Leu

2165 2170 2175

Lys Gly Pro Lys Ala Glu Val Thr Ala Pro Asp Val Glu Val Ser

2180 2185 2190

Leu Pro Ser Val Glu Val Asp Val Glu Ala Pro Arg Ala Lys Leu

2195 2200 2205

Asp Gly Ala Arg Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp

2210 2215 2220

Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met

2225 2230 2235

Pro Ser Phe Gly Val Ser Ala Pro Gly Lys Ser Ile Glu Val Ser

2240 2245 2250

Val Asp Val Ser Ala Pro Lys Val Glu Ala Glu Val Ser Leu Pro

2255 2260 2265

Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Ile Ser Ile Glu Pro

2270 2275 2280

Pro Ser Ala Gln Leu Glu Val Gln Ala Gly Gln Val Asp Leu Lys

2285 2290 2295

Leu Pro Glu Gly His Val Pro Glu Gly Ala Gly Leu Lys Gly His

2300 2305 2310

Leu Pro Lys Leu Gln Met Pro Ser Phe Lys Met Pro Lys Val Asp

2315 2320 2325

Arg Lys Gly Pro Gln Ile Asp Val Lys Gly Pro Lys Leu Asp Leu

2330 2335 2340

Lys Gly Pro Lys Thr Asp Val Thr Ala Pro Asp Val Glu Val Ser

2345 2350 2355

Gln Pro Gly Met Glu Val Asp Val Glu Ala Pro Gly Ala Lys Leu

2360 2365 2370

Asp Gly Ala Arg Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp

2375 2380 2385

Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met

2390 2395 2400

Pro Ser Phe Gly Val Ser Ala Pro Gly Lys Ser Ile Glu Val Leu

2405 2410 2415

Val Asp Val Ser Ala Pro Lys Val Glu Ala Asp Leu Ser Leu Pro

2420 2425 2430

Ser Met Gln Gly Asp Leu Lys Asn Thr Asp Ile Ser Ile Glu Pro

2435 2440 2445

Pro Ser Ala Gln Leu Glu Val Gln Ala Gly Gln Val Asp Val Lys

2450 2455 2460

Leu Pro Glu Gly His Val Leu Glu Gly Ala Gly Leu Lys Gly His
2465 2470 2475
Leu Pro Lys Leu Gln Met Pro Ser Phe Lys Met Pro Lys Val Asp
2480 2485 2490
Arg Lys Gly Pro Gln Ile Asp Ile Lys Gly Pro Lys Leu Asp Leu
2495 2500 2505
Lys Gly Pro Lys Met Asp Val Thr Ala Pro Asp Val Glu Val Ser
2510 2515 2520
Gln Pro Ser Met Glu Val Asp Val Glu Ala Pro Gly Ala Lys Leu
2525 2530 2535
Asp Gly Ala Arg Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp
2540 2545 2550
Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
2555 2560 2565
Pro Ser Tyr Arg Ala Ser Ala Pro Gly Lys Ser Ile Gln Ala Ser
2570 2575 2580
Val Asp Val Ser Ala Pro Lys Ala Glu Ala Asp Val Ser Leu Pro
2585 2590 2595
Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Leu Ser Ile Gln Leu
2600 2605 2610
Pro Ser Val Asp Leu Glu Val Gln Ala Gly Gln Val Asp Val Lys
2615 2620 2625
Leu Pro Glu Gly His Val Pro Glu Gly Ala Gly Leu Lys Gly His
2630 2635 2640
Leu Pro Lys Val Glu Met Pro Ser Phe Lys Met Pro Lys Val Asp
2645 2650 2655
Leu Lys Ser Pro Gln Val Asp Ile Lys Gly Pro Lys Leu Asp Leu
2660 2665 2670
Lys Val Pro Lys Ala Glu Val Thr Val Pro Asp Val Glu Val Ser
2675 2680 2685
Leu Pro Ser Val Glu Val Asp Val Gln Ala Pro Arg Ala Lys Leu
2690 2695 2700
Asp Gly Ala Arg Leu Glu Gly Asp Leu Ser Leu Ala Glu Lys Asp
2705 2710 2715
Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
2720 2725 2730
Pro Ser Phe Gly Val Ser Ala Pro Gly Arg Ser Ile Glu Ala Ser
2735 2740 2745
Leu Asp Val Ser Ala Pro Lys Val Glu Ala Asp Val Ser Leu Ser

2750 2755 2760

Ser Met Gln Gly Asp Leu Lys Ala Thr Asp Leu Ser Ile Gln Pro

2765 2770 2775

Pro Ser Ala Asp Leu Glu Val Gln Ala Val Gln Val Asp Val Glu

2780 2785 2790

Leu Leu Glu Gly Pro Val Pro Glu Gly Ala Gly Leu Lys Gly His

2795 2800 2805

Leu Pro Lys Val Glu Met Pro Ser Leu Lys Thr Pro Lys Val Asp

2810 2815 2820

Leu Lys Gly Pro Gln Ile Asp Val Lys Gly Pro Lys Leu Asp Leu

2825 2830 2835

Lys Gly Pro Lys Ala Glu Val Arg Val Pro Asp Val Glu Val Ser

2840 2845 2850

Leu Pro Ser Val Glu Val Asp Val Gln Ala Pro Lys Ala Lys Leu

2855 2860 2865

Asp Ala Gly Arg Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp

2870 2875 2880

Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met

2885 2890 2895

Pro Ser Phe Arg Val Ser Ala Pro Gly Lys Ser Met Glu Ala Ser

2900 2905 2910

Val Asp Val Ser Ala Pro Lys Val Glu Ala Asp Val Ser Leu Pro

2915 2920 2925

Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Leu Ser Ile Gln Pro

2930 2935 2940

Pro Ser Ala Asp Leu Lys Val Gln Ala Gly Gln Met Asp Val Lys

2945 2950 2955

Leu Pro Glu Gly Gln Val Pro Glu Gly Ala Gly Leu Lys Glu His

2960 2965 2970

Leu Pro Lys Val Glu Met Pro Ser Leu Lys Met Pro Lys Val Asp

2975 2980 2985

Leu Lys Gly Pro Gln Val Asp Ile Lys Gly Pro Lys Leu Asp Leu

2990 2995 3000

Lys Val Ser Lys Ala Glu Val Thr Ala Pro Asp Val Glu Val Ser

3005 3010 3015

Leu Pro Ser Val Glu Val Asp Val Gln Ala Pro Arg Ala Lys Leu

3020 3025 3030

Asp Ser Ala Gln Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp

3035 3040 3045

Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
3050 3055 3060

Pro Ser Phe Gly Val Ser Ala Pro Gly Lys Ser Ile Glu Ala Ser
3065 3070 3075

Val His Val Ser Ala Pro Lys Val Glu Ala Asp Val Ser Leu Pro
3080 3085 3090

Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Leu Ser Ile Gln Pro
3095 3100 3105

His Ser Ala Asp Leu Thr Val Gln Ala Arg Gln Val Asp Met Lys
3110 3115 3120

Leu Leu Glu Gly His Val Pro Glu Glu Ala Gly Leu Lys Gly His
3125 3130 3135

Leu Pro Lys Val Gln Met Pro Ser Phe Lys Met Pro Lys Val Asp
3140 3145 3150

Leu Lys Gly Pro Glu Ile Asp Ile Lys Gly Pro Lys Leu Asp Leu
3155 3160 3165

Lys Asp Pro Lys Val Glu Val Thr Ala Pro Asp Val Glu Val Ser
3170 3175 3180

Leu Pro Ser Val Glu Val Asp Val Glu Ala Pro Gly Ala Lys Leu
3185 3190 3195

Asp Gly Ala Arg Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp
3200 3205 3210

Met Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
3215 3220 3225

Pro Ser Phe Gly Val Ser Ala Pro Gly Lys Ser Met Glu Ala Ser
3230 3235 3240

Val Asp Val Thr Ala Pro Lys Val Glu Ala Asp Val Ser Leu Pro
3245 3250 3255

Ser Met Gln Gly Asp Leu Lys Ala Thr Asp Leu Ser Val Gln Pro
3260 3265 3270

Pro Ser Ala Asp Leu Glu Val Gln Ala Gly Gln Val Asp Val Lys
3275 3280 3285

Leu Pro Glu Gly Pro Val Pro Glu Gly Ala Ser Leu Lys Gly His
3290 3295 3300

Leu Pro Lys Val Gln Met Pro Ser Phe Lys Met Pro Lys Val Asp
3305 3310 3315

Leu Lys Gly Pro Gln Ile Asp Val Lys Gly Pro Lys Leu Asp Leu
3320 3325 3330

Lys Gly Pro Lys Ala Glu Val Thr Ala Pro Asp Val Lys Met Ser

3335 3340 3345

Leu Ser Ser Met Glu Val Asp Val Gln Ala Pro Arg Ala Lys Leu
3350 3355 3360

Asp Gly Val Gln Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp
3365 3370 3375

Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
3380 3385 3390

Pro Ser Phe Gly Val Ser Ala Pro Gly Lys Ser Met Glu Ala Ser
3395 3400 3405

Val Asp Val Ser Glu Leu Lys Ala Lys Ala Asp Val Ser Leu Pro
3410 3415 3420

Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Leu Ser Ile Gln Ser
3425 3430 3435

Pro Ser Ala Asp Leu Glu Val Gln Ala Gly Gln Val Asp Val Lys
3440 3445 3450

Leu Pro Glu Gly Pro Leu Pro Lys Gly Ala Gly Leu Lys Gly His
3455 3460 3465

Leu Pro Lys Val Gln Met Pro Cys Leu Lys Met Pro Lys Val Ala
3470 3475 3480

Leu Lys Gly Pro Gln Val Asp Val Lys Gly Pro Lys Leu Asp Leu
3485 3490 3495

Lys Gly Pro Lys Ala Asp Val Met Thr Pro Val Val Glu Val Ser
3500 3505 3510

Leu Pro Ser Met Glu Val Asp Val Glu Ala Pro Gly Ala Lys Leu
3515 3520 3525

Asp Ser Val Arg Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp
3530 3535 3540

Met Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
3545 3550 3555

Pro Ser Phe Gly Val Ser Ala Pro Gly Lys Ser Ile Glu Ala Ser
3560 3565 3570

Leu Asp Val Ser Ala Leu Lys Val Glu Ala Asp Val Ser Leu Pro
3575 3580 3585

Ser Met Gln Gly Asp Leu Lys Thr Thr His Leu Ser Ile Gln Pro
3590 3595 3600

Pro Ser Ala Asp Leu Glu Val Gln Ala Gly Gln Glu Asp Val Lys
3605 3610 3615

Leu Pro Glu Gly Pro Val His Glu Gly Ala Gly Leu Lys Gly His
3620 3625 3630

Leu Pro Lys Leu Gln Met Pro Ser Phe Lys Val Pro Lys Val Asp
3635 3640 3645
Leu Lys Gly Pro Gln Ile Asp Val Asn Val Pro Lys Leu Asp Leu
3650 3655 3660
Lys Gly Pro Lys Val Glu Val Thr Ser Pro Asn Leu Asp Val Ser
3665 3670 3675
Leu Pro Ser Met Glu Val Asp Ile Gln Ala Pro Gly Ala Lys Leu
3680 3685 3690
Asp Ser Thr Arg Leu Glu Gly Asp Leu Ser Leu Ala Asp Lys Asp
3695 3700 3705
Val Thr Ala Lys Asp Ser Lys Phe Lys Met Pro Lys Phe Lys Met
3710 3715 3720
Pro Ser Phe Gly Met Leu Ser Pro Gly Lys Ser Ile Glu Val Ser
3725 3730 3735
Val Asp Val Ser Ala Pro Lys Met Glu Ala Asp Met Ser Ile Pro
3740 3745 3750
Ser Met Gln Gly Asp Leu Lys Thr Thr Asp Leu Arg Ile Gln Ala
3755 3760 3765
Pro Ser Ala Asp Leu Glu Val Gln Ala Gly Gln Val Asp Leu Lys
3770 3775 3780
Leu Pro Glu Gly His Met Pro Glu Val Ala Gly Leu Lys Gly His
3785 3790 3795
Leu Pro Lys Val Glu Met Pro Ser Phe Lys Met Pro Lys Val Asp
3800 3805 3810
Leu Lys Gly Pro Gln Val Asp Val Lys Gly Pro Lys Leu Asp Leu
3815 3820 3825
Lys Gly Pro Lys Ala Glu Val Met Ala Pro Asp Val Glu Val Ser
3830 3835 3840
Leu Pro Ser Val Glu Thr Asp Val Gln Ala Pro Gly Ser Met Leu
3845 3850 3855
Asp Gly Ala Arg Leu Glu Gly Asp Leu Ser Leu Ala His Glu Asp
3860 3865 3870
Val Ala Gly Lys Asp Ser Lys Phe Gln Gly Pro Lys Leu Ser Thr
3875 3880 3885
Ser Gly Phe Glu Trp Ser Ser Lys Lys Val Ser Met Ser Ser Ser
3890 3895 3900
Glu Ile Glu Gly Asn Val Thr Phe His Glu Lys Thr Ser Thr Phe
3905 3910 3915
Pro Ile Val Glu Ser Val Val His Glu Gly Asp Leu His Asp Pro

3920 3925 3930

Ser Arg Asp Gly Asn Leu Gly Leu Ala Val Gly Glu Val Gly Met

3935 3940 3945

Asp Ser Lys Phe Lys Lys Leu His Phe Lys Val Pro Lys Val Ser

3950 3955 3960

Phe Ser Ser Thr Lys Thr Pro Lys Asp Ser Leu Val Pro Gly Ala

3965 3970 3975

Lys Ser Ser Ile Gly Leu Ser Thr Ile Pro Leu Ser Ser Ser Glu

3980 3985 3990

Cys Ser Ser Phe Glu Leu Gln Gln Val Ser Ala Cys Ser Glu Pro

3995 4000 4005

Ser Met Gln Met Pro Lys Val Gly Phe Ala Gly Phe Pro Ser Ser

4010 4015 4020

Arg Leu Asp Leu Thr Gly Pro His Phe Glu Ser Ser Ile Leu Ser

4025 4030 4035

Pro Cys Glu Asp Val Thr Leu Thr Lys Tyr Gln Val Thr Val Pro

4040 4045 4050

Arg Ala Ala Leu Ala Pro Glu Leu Ala Leu Glu Ile Pro Ser Gly

4055 4060 4065

Ser Gln Ala Asp Ile Pro Leu Pro Lys Thr Glu Cys Ser Thr Asp

4070 4075 4080

Leu Gln Pro Pro Glu Gly Val Pro Thr Ser Gln Ala Glu Ser His

4085 4090 4095

Ser Gly Pro Leu Asn Ser Met Ile Pro Val Ser Leu Gly Gln Val

4100 4105 4110

Ser Phe Pro Lys Phe Tyr Lys Pro Lys Phe Val Phe Ser Val Pro

4115 4120 4125

Gln Met Ala Val Pro Glu Gly Asp Leu His Ala Ala Val Gly Ala

4130 4135 4140

Pro Val Met Ser Pro Leu Ser Pro Gly Glu Arg Val Gln Cys Pro

4145 4150 4155

Leu Pro Ser Thr Gln Leu Pro Ser Pro Gly Thr Cys Val Ser Gln

4160 4165 4170

Gly Pro Glu Glu Leu Val Ala Ser Leu Gln Thr Ser Val Val Ala

4175 4180 4185

Pro Gly Glu Ala Pro Ser Glu Asp Ala Asp His Glu Gly Lys Gly

4190 4195 4200

Ser Pro Leu Lys Met Pro Lys Ile Lys Leu Pro Ser Phe Arg Trp

4205 4210 4215

Ser Pro Lys Lys Glu Thr Gly Pro Lys Val Asp Pro Glu Cys Ser
4220 4225 4230

Val Glu Asp Ser Lys Leu Ser Leu Val Leu Asp Lys Asp Glu Val
4235 4240 4245

Ala Pro Gln Ser Ala Ile His Met Asp Leu Pro Pro Glu Arg Asp
4250 4255 4260

Gly Glu Lys Gly Arg Ser Thr Lys Pro Gly Phe Ala Met Pro Lys
4265 4270 4275

Leu Ala Leu Pro Lys Met Lys Ala Ser Lys Ser Gly Val Ser Leu
4280 4285 4290

Pro Gln Arg Asp Val Asp Pro Ser Leu Ser Ser Ala Thr Ala Gly
4295 4300 4305

Gly Ser Phe Gln Asp Thr Glu Lys Ala Ser Ser Asp Gly Gly Arg
4310 4315 4320

Gly Gly Leu Gly Ala Thr Ala Ser Ala Thr Gly Ser Glu Gly Val
4325 4330 4335

Asn Leu His Arg Pro Gln Val His Ile Pro Ser Leu Gly Phe Ala
4340 4345 4350

Lys Pro Asp Leu Arg Ser Ser Lys Ala Lys Val Glu Val Ser Gln
4355 4360 4365

Pro Glu Ala Asp Leu Pro Leu Pro Lys His Asp Leu Ser Thr Glu
4370 4375 4380

Gly Asp Ser Arg Gly Cys Gly Leu Gly Asp Val Pro Val Ser Gln
4385 4390 4395

Pro Cys Gly Glu Gly Ile Ala Pro Thr Pro Glu Asp Pro Leu Gln
4400 4405 4410

Pro Ser Cys Arg Lys Pro Asp Ala Glu Val Leu Thr Val Glu Ser
4415 4420 4425

Pro Glu Glu Glu Ala Met Thr Lys Tyr Ser Gln Glu Ser Trp Phe
4430 4435 4440

Lys Met Pro Lys Phe Arg Met Pro Ser Leu Arg Arg Ser Phe Arg
4445 4450 4455

Asp Arg Gly Gly Ala Gly Lys Leu Glu Val Ala Gln Thr Gln Ala
4460 4465 4470

Pro Ala Ala Thr Gly Gly Glu Ala Ala Ala Lys Val Lys Glu Phe
4475 4480 4485

Leu Val Ser Gly Ser Asn Val Glu Ala Ala Met Ser Leu Gln Leu
4490 4495 4500

Pro Glu Ala Asp Ala Glu Val Thr Ala Ser Glu Ser Lys Ser Ser

4505 4510 4515

Thr Asp Ile Leu Arg Cys Asp Leu Asp Ser Thr Gly Leu Lys Leu

4520 4525 4530

His Leu Ser Thr Ala Gly Met Thr Gly Asp Glu Leu Ser Thr Ser

4535 4540 4545

Glu Val Arg Ile His Pro Ser Lys Gly Pro Leu Pro Phe Gln Met

4550 4555 4560

Pro Gly Met Arg Leu Pro Glu Thr Gln Val Leu Pro Gly Glu Ile

4565 4570 4575

Asp Glu Thr Pro Leu Ser Lys Pro Gly His Asp Leu Ala Ser Met

4580 4585 4590

Glu Asp Lys Thr Glu Lys Trp Ser Ser Gln Pro Glu Gly Pro Leu

4595 4600 4605

Lys Leu Lys Ala Ser Ser Thr Asp Met Pro Ser Gln Ile Ser Val

4610 4615 4620

Val Asn Val Asp Gln Leu Trp Glu Asp Ser Val Leu Thr Val Lys

4625 4630 4635

Phe Pro Lys Leu Met Val Pro Arg Phe Ser Phe Pro Ala Pro Ser

4640 4645 4650

Ser Glu Asp Asp Val Phe Ile Pro Thr Val Arg Glu Val Gln Cys

4655 4660 4665

Pro Glu Ala Asn Ile Asp Thr Ala Leu Cys Lys Glu Ser Pro Gly

4670 4675 4680

Leu Trp Gly Ala Ser Ile Leu Lys Ala Gly Ala Gly Val Pro Gly

4685 4690 4695

Glu Gln Pro Val Asp Leu Asn Leu Pro Leu Glu Ala Pro Pro Ile

4700 4705 4710

Ser Lys Val Arg Val His Ile Gln Gly Ala Gln Val Glu Ser Gln

4715 4720 4725

Glu Val Thr Ile His Ser Ile Val Thr Pro Glu Phe Val Asp Leu

4730 4735 4740

Ser Val Pro Arg Thr Phe Ser Thr Gln Ile Val Arg Glu Ser Glu

4745 4750 4755

Ile Pro Thr Ser Glu Ile Gln Thr Pro Ser Tyr Gly Phe Ser Leu

4760 4765 4770

Leu Lys Val Lys Ile Pro Glu Pro His Thr Gln Ala Arg Val Tyr

4775 4780 4785

Thr Thr Met Thr Gln His Ser Arg Thr Gln Glu Gly Thr Glu Glu

4790 4795 4800

Ala Pro Ile Gln Ala Thr Pro Gly Val Asp Ser Ile Ser Gly Asp
4805 4810 4815
Leu Gln Pro Asp Thr Gly Glu Pro Phe Glu Met Ile Ser Ser Ser
4820 4825 4830
Val Asn Val Leu Gly Gln Gln Thr Leu Thr Phe Glu Val Pro Ser
4835 4840 4845
Gly His Gln Leu Ala Asp Ser Cys Ser Asp Glu Glu Pro Ala Glu
4850 4855 4860
Ile Leu Glu Phe Pro Pro Asp Asp Ser Gln Glu Ala Thr Thr Pro
4865 4870 4875
Leu Ala Asp Glu Gly Arg Ala Pro Lys Asp Lys Pro Glu Ser Lys
4880 4885 4890
Lys Ser Gly Leu Leu Trp Phe Trp Leu Pro Asn Ile Gly Phe Ser
4895 4900 4905
Ser Ser Val Asp Glu Thr Gly Val Asp Ser Lys Asn Asp Val Gln
4910 4915 4920
Arg Ser Ala Pro Ile Gln Thr Gln Pro Glu Ala Arg Pro Glu Ala
4925 4930 4935
Glu Leu Pro Lys Lys Gln Glu Lys Ala Gly Trp Phe Arg Phe Pro
4940 4945 4950
Lys Leu Gly Phe Ser Ser Ser Pro Thr Lys Lys Ser Lys Ser Thr
4955 4960 4965
Glu Asp Gly Ala Glu Leu Glu Glu Gln Lys Leu Gln Glu Glu Thr
4970 4975 4980
Ile Thr Phe Phe Asp Ala Arg Glu Ser Phe Ser Pro Glu Glu Lys
4985 4990 4995
Glu Glu Gly Glu Leu Ile Gly Pro Val Gly Thr Gly Leu Asp Ser
5000 5005 5010
Arg Val Met Val Thr Ser Ala Ala Arg Thr Glu Leu Ile Leu Pro
5015 5020 5025
Glu Gln Asp Arg Lys Ala Asp Asp Glu Ser Lys Gly Ser Gly Leu
5030 5035 5040
Gly Pro Asn Glu Gly
5045

<210>3
<211>2731
<212>DNA
<213> 人类

<220>
<221>CDS
<222>(216)..(2318)

<400>3

```
gcgcttggcg ggagatagaa aagtgcttca acccgcgccg gcggcgactg cagttcctgc      60
gagcgaggag cgcgggacct gctgacacgc tgacgccttc gagcgcggcc cggggcccgg     120
agcggccgga gcagcccggg tcctgacccc ggcccggctc ccgctccggg ctctgccggc     180
gggcgggcga gcgcggcgcg gtccgggccg ggggg atg tct cgg cgg acg cgc         233
                                        Met Ser Arg Arg Thr Arg
                                        1               5
tgc gag gat ctg gat gag ctg cac tac cag gac aca gat tca gat gtg         281
Cys Glu Asp Leu Asp Glu Leu His Tyr Gln Asp Thr Asp Ser Asp Val
            10                  15                  20
ccg gag cag agg gat agc aag tgc aag gtc aaa tgg acc cat gag gag         329
Pro Glu Gln Arg Asp Ser Lys Cys Lys Val Lys Trp Thr His Glu Glu
        25                  30                  35
gac gag cag ctg agg gcc ctg gtg agg cag ttt gga cag cag gac tgg         377
Asp Glu Gln Leu Arg Ala Leu Val Arg Gln Phe Gly Gln Gln Asp Trp
    40                  45                  50
aag ttc ctg gcc agc cac ttc cct aac cgc act gac cag caa tgc cag         425
Lys Phe Leu Ala Ser His Phe Pro Asn Arg Thr Asp Gln Gln Cys Gln
55                  60                  65                  70
tac agg tgg ctg aga gtt ttg aat cca gac ctt gtc aag ggg cca tgg         473
Tyr Arg Trp Leu Arg Val Leu Asn Pro Asp Leu Val Lys Gly Pro Trp
                75                  80                  85
acc aaa gag gaa gac caa aaa gtc atc gag ctg gtt aag aag tat ggc         521
Thr Lys Glu Glu Asp Gln Lys Val Ile Glu Leu Val Lys Lys Tyr Gly
            90                  95                  100
aca aag cag tgg aca ctg att gcc aag cac ctg aag ggc cgg ctg ggg         569
Thr Lys Gln Trp Thr Leu Ile Ala Lys His Leu Lys Gly Arg Leu Gly
        105                 110                 115
aag cag tgc cgt gaa cgc tgg cac aac cac ctc aac cct gag gtg aag         617
Lys Gln Cys Arg Glu Arg Trp His Asn His Leu Asn Pro Glu Val Lys
    120                 125                 130
aag tct tgc tgg acc gag gag gag gac cgc atc atc tgc gag gcc cac         665
Lys Ser Cys Trp Thr Glu Glu Glu Asp Arg Ile Ile Cys Glu Ala His
135                 140                 145                 150
```

aag gtg ctg ggc aac cgc tgg gcc gag atc gcc aag atg ttg cca ggg 713
Lys Val Leu Gly Asn Arg Trp Ala Glu Ile Ala Lys Met Leu Pro Gly
155 160 165
agg aca gac aat gct gtg aag aat cac tgg aac tct acc atc aaa agg 761
Arg Thr Asp Asn Ala Val Lys Asn His Trp Asn Ser Thr Ile Lys Arg
170 175 180
aag gtg gac aca gga ggc ttc ttg agc gag tcc aaa gac tgc aag ccc 809
Lys Val Asp Thr Gly Gly Phe Leu Ser Glu Ser Lys Asp Cys Lys Pro
185 190 195
cca gtg tac ttg ctg ctg gag ctc gag gac aag gac ggc ctc cag agt 857
Pro Val Tyr Leu Leu Leu Glu Leu Glu Asp Lys Asp Gly Leu Gln Ser
200 205 210
gcc cag ccc acg gaa ggc cag gga agt ctt ctg acc aac tgg ccc tcc 905
Ala Gln Pro Thr Glu Gly Gln Gly Ser Leu Leu Thr Asn Trp Pro Ser
215 220 225 230
gtc cct cct acc ata aag gag gag gaa aac agt gag gag gaa ctt gca 953
Val Pro Pro Thr Ile Lys Glu Glu Glu Asn Ser Glu Glu Glu Leu Ala
235 240 245
gca gcc acc aca tcg aag gaa cag gag ccc atc ggt aca gat ctg gac 1001
Ala Ala Thr Thr Ser Lys Glu Gln Glu Pro Ile Gly Thr Asp Leu Asp
250 255 260
gca gtg cga aca cca gag ccc ttg gag gaa ttc ccg aag cgt gag gac 1049
Ala Val Arg Thr Pro Glu Pro Leu Glu Glu Phe Pro Lys Arg Glu Asp
265 270 275
cag gaa ggc tcc cca cca gaa acg agc ctg cct tac aag tgg gtg gtg 1097
Gln Glu Gly Ser Pro Pro Glu Thr Ser Leu Pro Tyr Lys Trp Val Val
280 285 290
gag gca gct aac ctc ctc atc ccc gct gtg ggt tct agc ctc tct gaa 1145
Glu Ala Ala Asn Leu Leu Ile Pro Ala Val Gly Ser Ser Leu Ser Glu
295 300 305 310
gcc ctg gac ttg atc gag tcg gac cct gat gct tgg tgt gac ctg agt 1193
Ala Leu Asp Leu Ile Glu Ser Asp Pro Asp Ala Trp Cys Asp Leu Ser
315 320 325
aaa ttt gac ctc cct gag gaa cca tct gca gag gac agt atc aac aac 1241
Lys Phe Asp Leu Pro Glu Glu Pro Ser Ala Glu Asp Ser Ile Asn Asn
330 335 340
agc cta gtg cag ctg caa gcg tca cat cag cag caa gtc ctg cca ccc 1289
Ser Leu Val Gln Leu Gln Ala Ser His Gln Gln Gln Val Leu Pro Pro
345 350 355

cgc cag cct tcc gcc ctg gtg ccc agt gtg acc gag tac cgc ctg gat 1337
Arg Gln Pro Ser Ala Leu Val Pro Ser Val Thr Glu Tyr Arg Leu Asp
360 365 370

ggc cac acc atc tca gac ctg agc cgg agc agc cgg ggc gag ctg atc 1385
Gly His Thr Ile Ser Asp Leu Ser Arg Ser Ser Arg Gly Glu Leu Ile
375 380 385 390

ccc atc tcc ccc agc act gaa gtc ggg ggc tct ggc att ggc aca ccg 1433
Pro Ile Ser Pro Ser Thr Glu Val Gly Gly Ser Gly Ile Gly Thr Pro
395 400 405

ccc tct gtg ctc aag cgg cag agg aag agg cgt gtg gct ctg tcc cct 1481
Pro Ser Val Leu Lys Arg Gln Arg Lys Arg Arg Val Ala Leu Ser Pro
410 415 420

gtc act gag aat agc acc agt ctg tcc ttc ctg gat tcc tgt aac agc 1529
Val Thr Glu Asn Ser Thr Ser Leu Ser Phe Leu Asp Ser Cys Asn Ser
425 430 435

ctc acg ccc aag agc aca cct gtt aag acc ctg ccc ttc tcg ccc tcc 1577
Leu Thr Pro Lys Ser Thr Pro Val Lys Thr Leu Pro Phe Ser Pro Ser
440 445 450

cag ttt ctg aac ttc tgg aac aaa cag gac aca ttg gag ctg gag agc 1625
Gln Phe Leu Asn Phe Trp Asn Lys Gln Asp Thr Leu Glu Leu Glu Ser
455 460 465 470

ccc tcg ctg aca tcc acc cca gtg tgc agc cag aag gtg gtg gtc acc 1673
Pro Ser Leu Thr Ser Thr Pro Val Cys Ser Gln Lys Val Val Val Thr
475 480 485

aca cca ctg cac cgg gac aag aca ccc ctg cac cag aaa cat gct gcg 1721
Thr Pro Leu His Arg Asp Lys Thr Pro Leu His Gln Lys His Ala Ala
490 495 500

ttt gta acc cca gat cag aag tac tcc atg gac aac act ccc cac acg 1769
Phe Val Thr Pro Asp Gln Lys Tyr Ser Met Asp Asn Thr Pro His Thr
505 510 515

cca acc ccg ttc aag aac gcc ctg gag aag tac gga ccc ctg aag ccc 1817
Pro Thr Pro Phe Lys Asn Ala Leu Glu Lys Tyr Gly Pro Leu Lys Pro
520 525 530

ctg cca cag acc ccg cac ctg gag gag gac ttg aag gag gtg ctg cgt 1865
Leu Pro Gln Thr Pro His Leu Glu Glu Asp Leu Lys Glu Val Leu Arg
535 540 545 550

tct gag gct ggc atc gaa ctc atc atc gag gac gac atc agg ccc gag 1913
Ser Glu Ala Gly Ile Glu Leu Ile Ile Glu Asp Asp Ile Arg Pro Glu
555 560 565

aag cag aag agg aag cct ggg ctg cgg cgg agc ccc atc aag aaa gtc 1961
Lys Gln Lys Arg Lys Pro Gly Leu Arg Arg Ser Pro Ile Lys Lys Val
570 575 580

cgg aag tct ctg gct ctt gac att gtg gat gag gat gtg aag ctg atg 2009
Arg Lys Ser Leu Ala Leu Asp Ile Val Asp Glu Asp Val Lys Leu Met
585 590 595

atg tcc aca ctg ccc aag tct cta tcc ttg ccg aca act gcc cct tca 2057
Met Ser Thr Leu Pro Lys Ser Leu Ser Leu Pro Thr Thr Ala Pro Ser
600 605 610

aac tct tcc agc ctc acc ctg tca ggt atc aaa gaa gac aac agc ttg 2105
Asn Ser Ser Ser Leu Thr Leu Ser Gly Ile Lys Glu Asp Asn Ser Leu
615 620 625 630

ctc aac cag ggc ttc ttg cag gcc aag ccc gag aag gca gca gtg gcc 2153
Leu Asn Gln Gly Phe Leu Gln Ala Lys Pro Glu Lys Ala Ala Val Ala
635 640 645

cag aag ccc cga agc cac ttc acg aca cct gcc cct atg tcc agt gcc 2201
Gln Lys Pro Arg Ser His Phe Thr Thr Pro Ala Pro Met Ser Ser Ala
650 655 660

tgg aag acg gtg gcc tgc ggg ggg acc agg gac cag ctt ttc atg cag 2249
Trp Lys Thr Val Ala Cys Gly Gly Thr Arg Asp Gln Leu Phe Met Gln
665 670 675

gag aaa gcc cgg cag ctc ctg ggc cgc ctg aag ccc agc cac aca tct 2297
Glu Lys Ala Arg Gln Leu Leu Gly Arg Leu Lys Pro Ser His Thr Ser
680 685 690

cgg acc ctc atc ttg tcc tga ggtgttgagg gtgtcacgag cccattctca 2348
Arg Thr Leu Ile Leu Ser
695 700

tgtttacagg ggttgtgggg gcagaggggg tctgtgaatc tgagagtcat tcaggtgacc 2408
tcctgcaggg agccttctgc caccagcccc tccccagact ctcaggtgga ggcaacaggg 2468
ccatgtgctg ccctgttgcc gagcccagct gtgggcggct cctggtgcta acaacaaagt 2528
tccacttcca ggtctgcctg gttccctccc caaggccaca gggagctccg tcagcttctc 2588
ccaagcccac gtcaggcctg gcctcatctc agaccctgct taggatgggg gatgtggcca 2648
ggggtgctcc tgtgctcacc ctctcttggt gcatttttt ggaagaataa aattgcctct 2708
c t c t t a a a a a a a a a a a a a a a a a 2731

<210>4
<211>700

<212>PRT
<213> 人类

<400>4

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

275 280 285

Pro Tyr Lys Trp Val Val Glu Ala Ala Asn Leu Leu Ile Pro Ala Val

290 295 300

Gly Ser Ser Leu Ser Glu Ala Leu Asp Leu Ile Glu Ser Asp Pro Asp

305 310 315 320

Ala Trp Cys Asp Leu Ser Lys Phe Asp Leu Pro Glu Glu Pro Ser Ala

325 330 335

Glu Asp Ser Ile Asn Asn Ser Leu Val Gln Leu Gln Ala Ser His Gln

340 345 350

Gln Gln Val Leu Pro Pro Arg Gln Pro Ser Ala Leu Val Pro Ser Val

355 360 365

Thr Glu Tyr Arg Leu Asp Gly His Thr Ile Ser Asp Leu Ser Arg Ser

370 375 380

Ser Arg Gly Glu Leu Ile Pro Ile Ser Pro Ser Thr Glu Val Gly Gly

385 390 395 400

Ser Gly Ile Gly Thr Pro Pro Ser Val Leu Lys Arg Gln Arg Lys Arg

405 410 415

Arg Val Ala Leu Ser Pro Val Thr Glu Asn Ser Thr Ser Leu Ser Phe

420 425 430

Leu Asp Ser Cys Asn Ser Leu Thr Pro Lys Ser Thr Pro Val Lys Thr

435 440 445

Leu Pro Phe Ser Pro Ser Gln Phe Leu Asn Phe Trp Asn Lys Gln Asp

450 455 460

Thr Leu Glu Leu Glu Ser Pro Ser Leu Thr Ser Thr Pro Val Cys Ser

465 470 475 480

Gln Lys Val Val Val Thr Thr Pro Leu His Arg Asp Lys Thr Pro Leu

485 490 495

His Gln Lys His Ala Ala Phe Val Thr Pro Asp Gln Lys Tyr Ser Met

500 505 510

Asp Asn Thr Pro His Thr Pro Thr Pro Phe Lys Asn Ala Leu Glu Lys

515 520 525

Tyr Gly Pro Leu Lys Pro Leu Pro Gln Thr Pro His Leu Glu Glu Asp

530 535 540

Leu Lys Glu Val Leu Arg Ser Glu Ala Gly Ile Glu Leu Ile Ile Glu

545 550 555 560

Asp Asp Ile Arg Pro Glu Lys Gln Lys Arg Lys Pro Gly Leu Arg Arg

565 570 575

Ser Pro Ile Lys Lys Val Arg Lys Ser Leu Ala Leu Asp Ile Val Asp

580 585 590

Glu Asp Val Lys Leu Met Met Ser Thr Leu Pro Lys Ser Leu Ser Leu
595 600 605
Pro Thr Thr Ala Pro Ser Asn Ser Ser Ser Leu Thr Leu Ser Gly Ile
610 615 620
Lys Glu Asp Asn Ser Leu Leu Asn Gln Gly Phe Leu Gln Ala Lys Pro
625 630 635 640
Glu Lys Ala Ala Val Ala Gln Lys Pro Arg Ser His Phe Thr Thr Pro
645 650 655
Ala Pro Met Ser Ser Ala Trp Lys Thr Val Ala Cys Gly Gly Thr Arg
660 665 670
Asp Gln Leu Phe Met Gln Glu Lys Ala Arg Gln Leu Leu Gly Arg Leu
675 680 685
Lys Pro Ser His Thr Ser Arg Thr Leu Ile Leu Ser
690 695 700

<210>5
<211>1207
<212>DNA
<213> 人类

<220>
<221>CDS
<222>(369)..(1037)

<400>5
cctctccgcc acttccctcg cttctgacca tagtttgcgg ggaagggagc gagcgcgtcg 60
aaaaccaagg aacgtgcgcg ctgacgtcac ggttgaggct cggagctgag gggccgcgga 120
gggcgtggcc tgcgggcggt tataaagagg cagtggtgcg cgcgcggccg gctcagtgct 180
gccgggcacc ggggcggcgg gttggtctac gctgtgcgcg gcggacgtcg gaggcagcgg 240
ggagcggagc ggggccgccg gggcctctcc agggccgcag cggcagcagt tgggcccccc 300
gccccggccg gcggaccgaa gaacgcagga agggggccgg ggggacccgc ccccggccgg 360
ccgcagcc atg aac tcc aac gtg gag aac cta ccc ccg cac atc atc cgc 410
Met Asn Ser Asn Val Glu Asn Leu Pro Pro His Ile Ile Arg
1 5 10
ctg gtg tac aag gag gtg acg aca ctg acc gca gac cca ccc gat ggc 458
Leu Val Tyr Lys Glu Val Thr Thr Leu Thr Ala Asp Pro Pro Asp Gly
15 20 25 30
atc aag gtc ttt ccc aac gag gag gac ctc acc gac ctc cag gtc acc 506
Ile Lys Val Phe Pro Asn Glu Glu Asp Leu Thr Asp Leu Gln Val Thr

35 40 45

atc gag ggc cct gag ggg acc cca tat gct gga ggt ctg ttc cgc atg 554
Ile Glu Gly Pro Glu Gly Thr Pro Tyr Ala Gly Gly Leu Phe Arg Met
50 55 60

aaa ctc ctg ctg ggg aag gac ttc cct gcc tcc cca ccc aag ggc tac 602
Lys Leu Leu Leu Gly Lys Asp Phe Pro Ala Ser Pro Pro Lys Gly Tyr
65 70 75

ttc ctg acc aag atc ttc cac ccg aac gtg ggc gcc aat ggc gag atc 650
Phe Leu Thr Lys Ile Phe His Pro Asn Val Gly Ala Asn Gly Glu Ile
80 85 90

tgc gtc aac gtg ctc aag agg gac tgg acg gct gag ctg ggc atc cga 698
Cys Val Asn Val Leu Lys Arg Asp Trp Thr Ala Glu Leu Gly Ile Arg
95 100 105 110

cac gta ctg ctg acc atc aag tgc ctg ctg atc cac cct aac ccc gag 746
His Val Leu Leu Thr Ile Lys Cys Leu Leu Ile His Pro Asn Pro Glu
115 120 125

tct gca ctc aac gag gag gcg ggc cgc ctg ctc ttg gag aac tac gag 794
Ser Ala Leu Asn Glu Glu Ala Gly Arg Leu Leu Leu Glu Asn Tyr Glu
130 135 140

gag tat gcg gct cgg gcc cgt ctg ctc aca gag atc cac ggg ggc gcc 842
Glu Tyr Ala Ala Arg Ala Arg Leu Leu Thr Glu Ile His Gly Gly Ala
145 150 155

ggc ggg ccc agc ggc agg gcc gaa gcc ggt cgg gcc ctg gcc agt ggc 890
Gly Gly Pro Ser Gly Arg Ala Glu Ala Gly Arg Ala Leu Ala Ser Gly
160 165 170

act gaa gct tcc tcc acc gac cct ggg gcc cca ggg ggc ccg gga ggg 938
Thr Glu Ala Ser Ser Thr Asp Pro Gly Ala Pro Gly Gly Pro Gly Gly
175 180 185 190

gct gag ggt ccc atg gcc aag aag cat gct ggc gag cgc gat aag aag 986
Ala Glu Gly Pro Met Ala Lys Lys His Ala Gly Glu Arg Asp Lys Lys
195 200 205

ctg gcg gcc aag aaa aag acg gac aag aag cgg gcg ctg cgg cgg ctg 1034
Leu Ala Ala Lys Lys Lys Thr Asp Lys Lys Arg Ala Leu Arg Arg Leu
210 215 220

tag tgggctctct tcctccttcc accgtgaccc caacctctcc tgtcccctcc 1087

ctccaactct gtctctaagt tatttaaatt atggctgggg tcggggaggg tacaggggggc 1147

actgggacct ggatttgttt ttctaaataa agttggaaaa gcagaaaaaa aaaaaaaaaa 1207

<210>6

<211>222
<212>PRT
<213> 人类
<400>6

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

<210>7
<211>927
<212>DNA
<213> 人类

<220>

<221>CDS

<222>(126)..(719)

<400>7

cgcgcagcgc tggtaccccg ttggtccgcg cgttgctgcg ttgtgagggg tgtcagctca 60

gtgcatccca ggcagctctt agtgtggagc agtgaactgt gtgtggttcc ttctacttgg 120

ggatc atg cag aga gct tca cgt ctg aag aga gag ctg cac atg tta gcc 170

Met Gln Arg Ala Ser Arg Leu Lys Arg Glu Leu Hi s Met Leu Ala

1 5 10 15

aca gag cca ccc cca ggc atc aca tgt tgg caa gat aaa gac caa atg 218

Thr Glu Pro Pro Pro Gly Ile Thr Cys Trp Gln Asp Lys Asp Gln Met

20 25 30

gat gac ctg cga gct caa ata tta ggt gga gcc aac aca cct tat gag 266

Asp Asp Leu Arg Ala Gln Ile Leu Gly Gly Ala Asn Thr Pro Tyr Glu

35 40 45

aaa ggt gtt ttt aag cta gaa gtt atc att cct gag agg tac cca ttt 314

Lys Gly Val Phe Lys Leu Glu Val Ile Ile Pro Glu Arg Tyr Pro Phe

50 55 60

gaa cct cct cag atc cga ttt ctc act cca att tat cat cca aac att 362

Glu Pro Pro Gln Ile Arg Phe Leu Thr Pro Ile Tyr His Pro Asn Ile

65 70 75

gat tct gct gga agg att tgt ctg gat gtt ctc aaa ttg cca cca aaa 410

Asp Ser Ala Gly Arg Ile Cys Leu Asp Val Leu Lys Leu Pro Pro Lys

80 85 90 95

ggt gct tgg aga cca tcc ctc aac atc gca act gtg ttg acc tct att 458

Gly Ala Trp Arg Pro Ser Leu Asn Ile Ala Thr Val Leu Thr Ser Ile

100 105 110

cag ctg ctc atg tca gaa ccc aac cct gat gac ccg ctc atg gct gac 506

Gln Leu Leu Met Ser Glu Pro Asn Pro Asp Asp Pro Leu Met Ala Asp

115 120 125

ata tcc tca gaa ttt aaa tat aat aag cca gcc ttc ctc aag aat gcc 554

Ile Ser Ser Glu Phe Lys Tyr Asn Lys Pro Ala Phe Leu Lys Asn Ala

130 135 140

aga cag tgg aca gag aag cat gca aga cag aaa caa aag gct gat gag 602

Arg Gln Trp Thr Glu Lys His Ala Arg Gln Lys Gln Lys Ala Asp Glu

145 150 155

gaa gag atg ctt gat aat cta cca gag gct ggt gac tcc aga gta cac 650

Glu Glu Met Leu Asp Asn Leu Pro Glu Ala Gly Asp Ser Arg Val His

160 165 170 175

aac tca aca cag aaa agg aag gcc agt cag cta gta ggc ata gaa aag 698
Asn Ser Thr Gln Lys Arg Lys Ala Ser Gln Leu Val Gly Ile Glu Lys
180 185 190

aaa ttt cat cct gat gtt tag gggacttgtc ctggttcatc ttagttaatg 749
Lys Phe His Pro Asp Val
195

tgttctttgc caaggtgatc taagttgcct accttgaatt ttttttaaa tatatttgat 809
gacataattt ttgtgtagtt tatttatctt gtacatatgt attttgaaat cttttaaacc 869
tgaaaaataa atagtcattt aatgttgaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaa 927

<210>8
<211>197
<212>PRT
<213> 人类

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

180 185 190

Phe His Pro Asp Val
195

<210>9
<211>23
<212>DNA
<213> 人工的

<220>
<223> 人工合成的用于 RT-PCR 的引物序列

<400>9
gagaaggaag agggtgaact gat 23

<210>10
<211>23
<212>DNA
<213> 人工的

<220>
<223> 人工合成的用于 RT-PCR 的引物序列

<400>10
cagtggacat ggatagatga gaa 23

<210>11
<211>20
<212>DNA
<213> 人工的
<220>
<223> 人工合成的用于 RT-PCR 的引物序列

<400>11
gaagccactt cacgacacct 20

<210>12
<211>22
<212>DNA

<213> 人工的

<220>
<223> 人工合成的用于 RT-PCR 的引物序列

<400>12
atcctaagca gggtctgaga tg 22

<210>13
<211>23
<212>DNA
<213> 人工的

<220>
<223> 人工合成的用于 RT-PCR 的引物序列

<400>13
tacttcctga ccaagatctt cca 23

<210>14
<211>23
<212>DNA
<213> 人工的

<220>
<223> 人工合成的用于 RT-PCR 的引物序列

<400>14
ttagagacag agttggaggg agg 23

<210>15
<211>23
<212>DNA
<213> 人工的

<220>
<223> 人工合成的用于 RT-PCR 的引物

<400>15

caaatattag gtggagccaa cac 23

<210>16
<211>23
<212>DNA
<213> 人工的

<220>
<223> 人工合成的用于 RT-PCR 的引物

<400>16
tagatcacct tggcaaagaa cac 23

<210>17
<211>20
<212>DNA
<213> 人工的
<220>
<223> 人工合成的用于 RT-PCR 的引物序列

<400>17
aggatgcaga aggagatcac 20

<210>18
<211>20
<212>DNA
<213> 人工的

<220>
<223> 人工合成的用于 RT-PCR 的引物序列

<400>18
agaaagggtg taacgcaact 20

<210>19
<211>21
<212>DNA
<213> 人工的

<220>
<223> 人工合成的用于 RT-PCR 的引物序列

<400>19
caccccccact gaaaaagatg a 21

<210>20
<211>19
<212>DNA
<213> 人工的

<220>
<223> 人工合成的用于 RT-PCR 的引物序列

<400>20
tacctgtgga gcaacctgc 19

<210>21
<211>23
<212>DNA
<213> 人工的

<220>
<223> 人工合成的用于 RT-PCR 的引物序列

<400>21
aaggattatg aggaggttgg tgt 23

<210>22
<211>23
<212>DNA
<213> 人工的

<220>
<223> 人工合成的用于 RT-PCR 的引物序列

<400>22
cttgggtctg taacaaagca ttc 23

<210>23
<211>20
<212>DNA
<213> 人工的

<220>
<223> 人工合成的用于 RT-PCR 的引物序列

<400>23
gatcaacatc cacagcgaga 20

<210>24
<211>20
<212>DNA
<213> 人工的

<220>
<223> 人工合成的用于 RT-PCR 的引物序列

<400>24
tgtcacagag ccgaatacca 20

<210>25
<211>21
<212>RNA
<213> 人工的

<220>
<223> 人工合成的用于 dsRNA 的寡核苷酸

<400>25
gauaugccau cccagauuuu u 21

<210>26
<211>21
<212>RNA
<213> 人工的

<220>

<223> 人工合成的用于 dsRNA 的寡核苷酸

<400>26
aaaucuggga uggcauaucu u 21

<210>27
<211>21
<212>RNA
<213> 人工的

<220>
<223> 人工合成的用于 dsRNA 的寡核苷酸

<400>27
gucaaauucc ccaaauuaau u 21

<210>28
<211>21
<212>RNA
<213> 人工的

<220>
<223> 人工合成的用于 dsRNA 的寡核苷酸

<400>28
uuaauuuggg gaauuugacu u 21

<210>29
<211>21
<212>RNA
<213> 人工的

<220>
<223> 人工合成的用于 dsRNA 的寡核苷酸
<400>29
guguccagag gccaauauuu u 21

<210>30
<211>21

<212>RNA
<213> 人工的

<220>
<223>′ 人工合成的用于 dsRNA 的寡核苷酸

<400>30
aauauuggcc ucuggacacu u 21

<210>31
<211>21
<212>RNA
<213> 人工的

<220>
<223> 人工合成的用于 dsRNA 的寡核苷酸

<400>31
ggcagggcuc caaaagacau u 21

<210>32
<211>21
<212>RNA
<213> 人工的

<220>
<223> 人工合成的用于 dsRNA 的寡核苷酸

<400>32
ugucuuuugg agcccugccu u 21

<210>33
<211>21
<212>RNA
<213> 人工的

<220>
<223> 人工合成的用于 dsRNA 的寡核苷酸

<400>33
ggagcccauc gguacagauu u 21

<210>34
<211>21
<212>RNA
<213> 人工的

<220>
<223> 人工合成的用于 dsRNA 的寡核苷酸

<400>34
aucuguaccg augggcuccu u 21

<210>35
<211>21
<212>RNA
<213> 人工的

<220>
<223> 人工合成的用于 dsRNA 的寡核苷酸
<400>35
cggcggagcc ccaucaagau u 21

<210>36
<211>21
<212>RNA
<213> 人工的

<220>
<223> 人工合成的用于 dsRNA 的寡核苷酸

<400>36
ucuugauggg gcuccgccgu u 21

<210>37
<211>21
<212>RNA
<213> 人工的

<220>
<223> 人工合成的用于 dsRNA 的寡核苷酸

<400>37
gcggagcccc aucaagaaau u 21

<210>38
<211>21
<212>RNA
<213> 人工的

<220>
<223>′ 人工合成的用于 dsRNA 的寡核苷酸

<400>38
uuucuugaug gggcuccgcu u 21

<210>39
<211>21
<212>RNA
<213> 人工的

<220>
<223> 人工合成的用于 dsRNA 的寡核苷酸

<400>39
gaugugaagc ugaugauguu u 21

<210>40
<211>21
<212>RNA
<213> 人工的

<220>
<223> 人工合成的用于 dsRNA 的寡核苷酸

<400>40
acaucaucag cuucacaucu u 21

<210>41
<211>21
<212>RNA
<213> 人工的

<220>
<223> 人工合成的用于 dsRNA 的寡核苷酸

<400>41
ugcugaccau caagugccuu u 21

<210>42
<211>21
<212>RNA
<213> 人工的

<220>
<223> 人工合成的用于 dsRNA 的寡核苷酸

<400>42
aggcacuuga uggucagcau u 21

<210>43
<211>21
<212>RNA
<213> 人工的

<220>
<223> 人工合成的用于 dsRNA 的寡核苷酸

<400>43
ccauaugcug gaggucuguu u 21

<210>44
<211>21
<212>RNA
<213> 人工的

<220>
<223> 人工合成的用于 dsRNA 的寡核苷酸

<400>44
acagaccucc agcauauggu u 21

<210>45
<211>21
<212>RNA
<213> 人工的

<220>
<223> 人工合成的用于 dsRNA 的寡核苷酸

<400>45
agagagagcu gcacauguuu u 21

<210>46
<211>21
<212>RNA
<213> 人工的

<220>
<223> 人工合成的用于 dsRNA 的寡核苷酸

<400>46
aacaugugca gcucucucuu u 21

<210>47
<211>19
<212>DNA
<213> 人工的

<220>
<223> 靶序列

<400>47
gatatgccat cccagattt 19
<210>48

<211>19
<212>DNA
<213> 人工的

<220>
<223> 靶序列

<400>48
gtcaaattcc ccaaattaa 19

<210>49
<211>19
<212>DNA
<213> 人工的

<220>
<223> 靶序列

<400>49
gtgtccagag gccaatatt 19

<210>50
<211>19
<212>DNA
<213> 人工的

<220>
<223> 靶序列

<400>50
ggcagggctc caaaagaca 19

<210>51
<211>19
<212>DNA
<213> 人工的

<220>
<223> 靶序列

<400>51
ggagcccatc ggtacagat 19

<210>52
<211>19
<212>DNA
<213> 人工的

<220>
<223> 靶序列

<400>52
cggcggagcc ccatcaaga 19

<210>53
<211>19
<212>DNA
<213> 人工的

<220>
<223> 靶序列

<400>53
gcggagcccc atcaagaaa 19
<210>54
<211>19
<212>DNA
<213> 人工的

<220>
<223> 靶序列

<400>54
gatgtgaagc tgatgatgt 19

<210>55
<211>19
<212>DNA

<213> 人工的

<220>
<223> 靶序列

<400>55
tgctgaccat caagtgcct 19

<210>56
<211>19
<212>DNA
<213> 人工的

<220>
<223> 靶序列

<400>56
ccatatgctg gaggtctgt 19

<210>57
<211>19
<212>DNA
<213> 人工的

<220>
<223> 靶序列

<400>57
agagagagct gcacatgtt 19

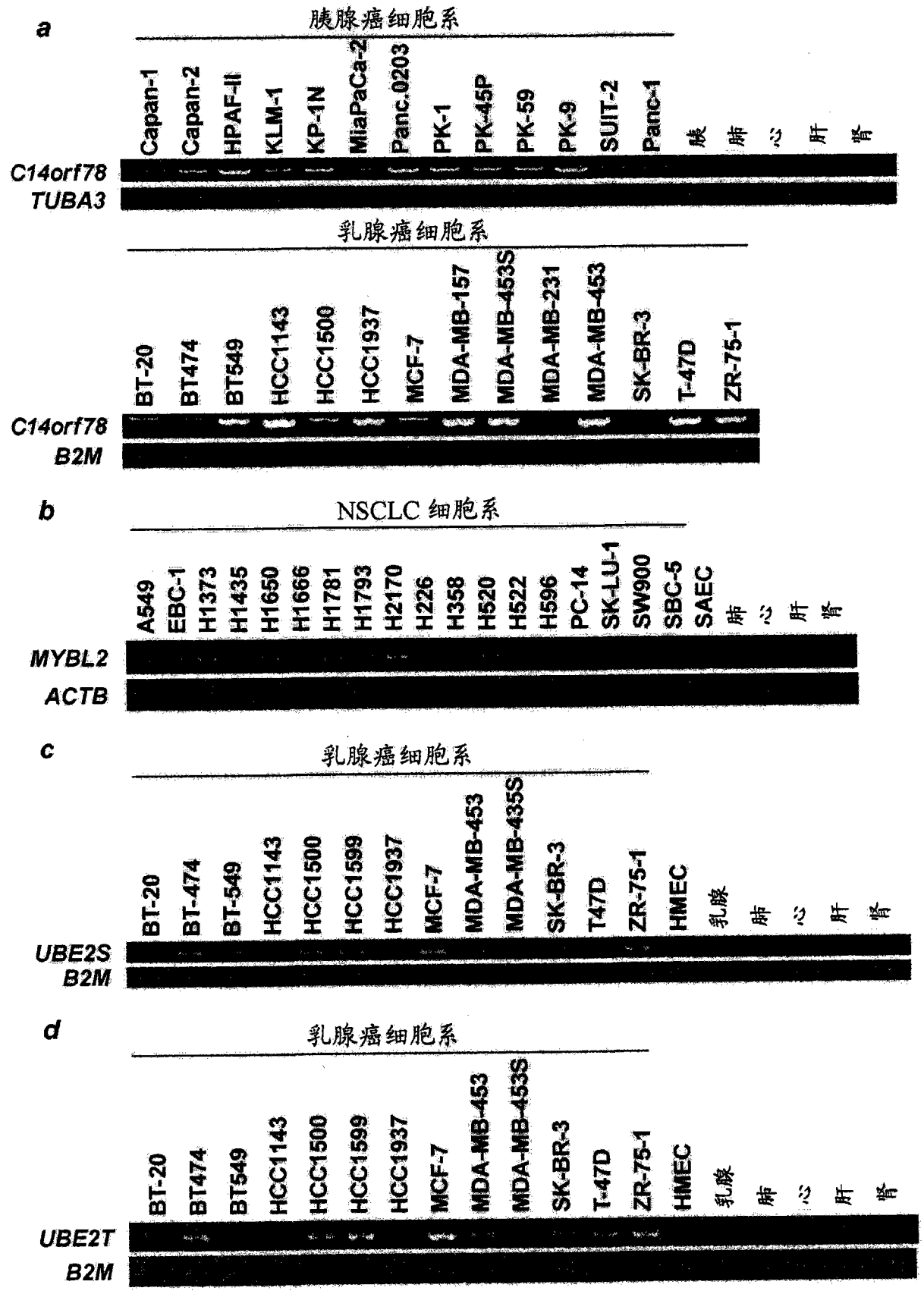

a
胰腺癌细胞系
Capan-1
Capan-2
HPAF-II
KLM-1
KP-1N
MiaPaCa-2
Panc.0203
PK-1
PK-45P
PK-59
PK-9
SUIT-2
Panc-1
胰
肺
心
肝
肾
C14orf78
TUBA3
乳腺癌细胞系
BT-20
BT474
BT549
HCC1143
HCC1500
HCC1937
MCF-7
MDA-MB-157
MDA-MB-453S
MDA-MB-231
MDA-MB-453
SK-BR-3
T-47D
ZR-75-1
C14orf78
B2M
b
NSCLC 细胞系
A549
EBC-1
H1373
H1435
H1650
H1666
H1781
H1793
H2170
H226
H358
H520
H522
H596
PC-14
SK-LU-1
SW900
SBC-5
SAEC
肺
心
肝
肾
MYBL2
ACTB
c
乳腺癌细胞系
BT-20
BT-474
BT-549
HCC1143
HCC1500
HCC1599
HCC1937
MCF-7
MDA-MB-453
MDA-MB-435S
SK-BR-3
T47D
ZR-75-1
HMEC
乳腺
肺
心
肝
肾
UBE2S
B2M
d
乳腺癌细胞系
BT-20
BT474
BT549
HCC1143
HCC1500
HCC1599
HCC1937
MCF-7
MDA-MB-453
MDA-MB-453S
SK-BR-3
T-47D
ZR-75-1
HMEC
乳腺
肺
心
肝
肾
UBE2T
B2M

图 1

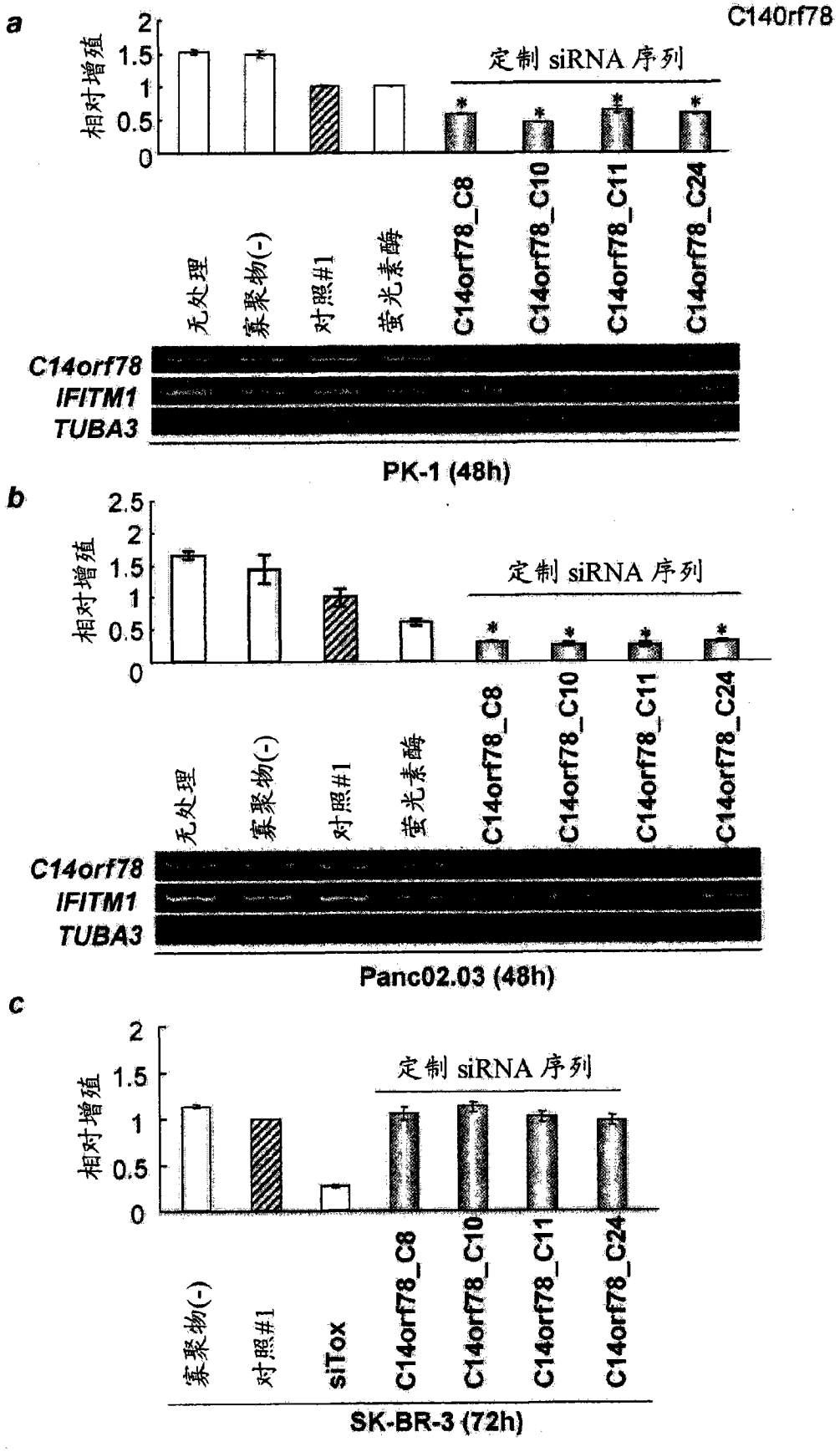

a
C14orf78
相对增殖
定制 siRNA 序列
无处理
寡聚物(-)
对照#1
萤光素酶
C14orf78_C8
C14orf78_C10
C14orf78_C11
C14orf78_C24
C14orf78
IFITM1
TUBA3
PK-1 (48h)
b
相对增殖
定制 siRNA 序列
无处理
寡聚物(-)
对照#1
萤光素酶
C14orf78_C8
C14orf78_C10
C14orf78_C11
C14orf78_C24
C14orf78
IFITM1
TUBA3
Panc02.03 (48h)
c
相对增殖
定制 siRNA 序列
寡聚物(-)
对照#1
siTox
C14orf78_C8
C14orf78_C10
C14orf78_C11
C14orf78_C24
SK-BR-3 (72h)

图 2

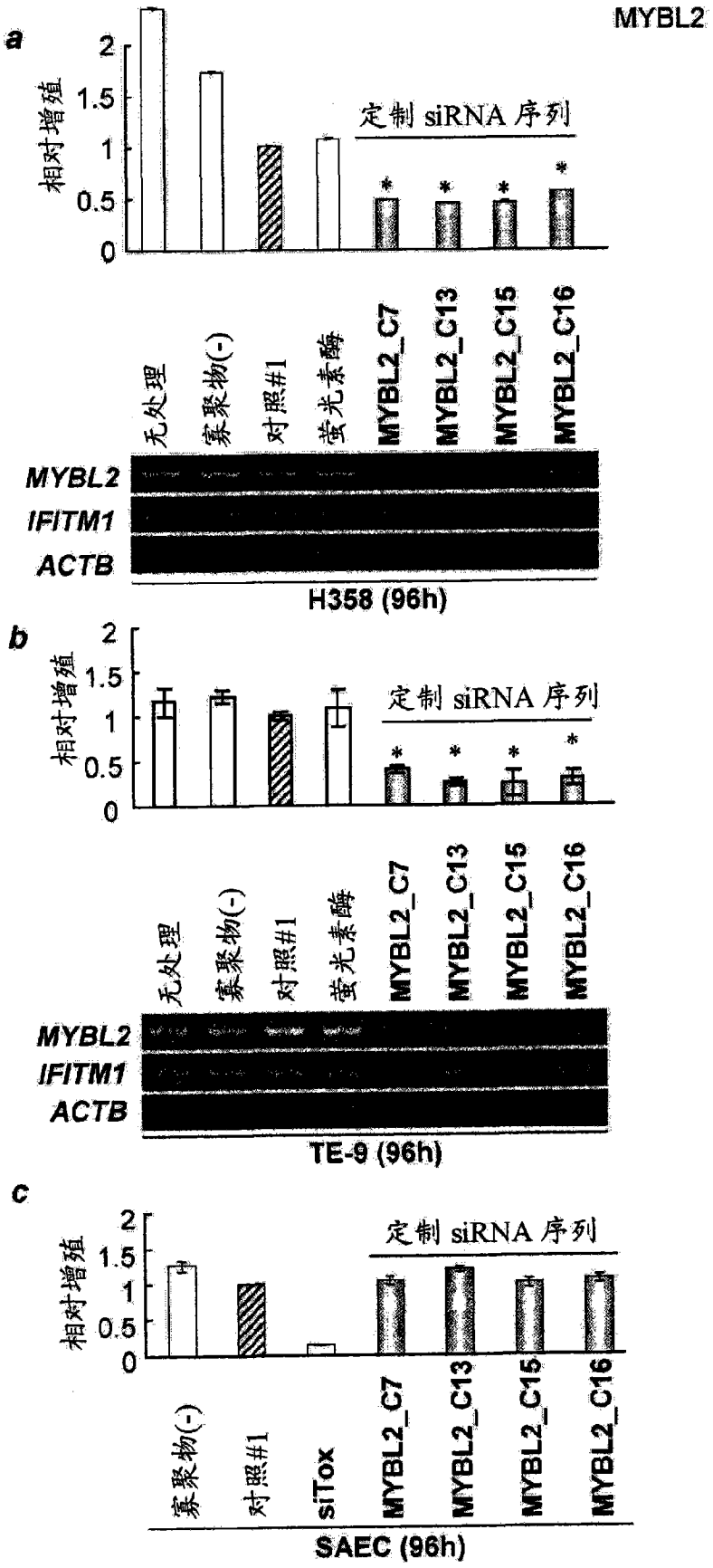

a
MYBL2
相对增殖
2
1.5
1
0.5
0
定制 siRNA 序列
*
*
*
*
无处理
寡聚物(-)
对照#1
萤光素酶
MYBL2_C7
MYBL2_C13
MYBL2_C15
MYBL2_C16
MYBL2
IFITM1
ACTB
H358 (96h)
b
相对增殖
2
1.5
1
0.5
0
定制 siRNA 序列
*
*
*
*
无处理
寡聚物(-)
对照#1
萤光素酶
MYBL2_C7
MYBL2_C13
MYBL2_C15
MYBL2_C16
MYBL2
IFITM1
ACTB
TE-9 (96h)
c
相对增殖
2
1.5
1
0.5
0
定制 siRNA 序列
寡聚物(-)
对照#1
siTox
MYBL2_C7
MYBL2_C13
MYBL2_C15
MYBL2_C16
SAEC (96h)

图 3

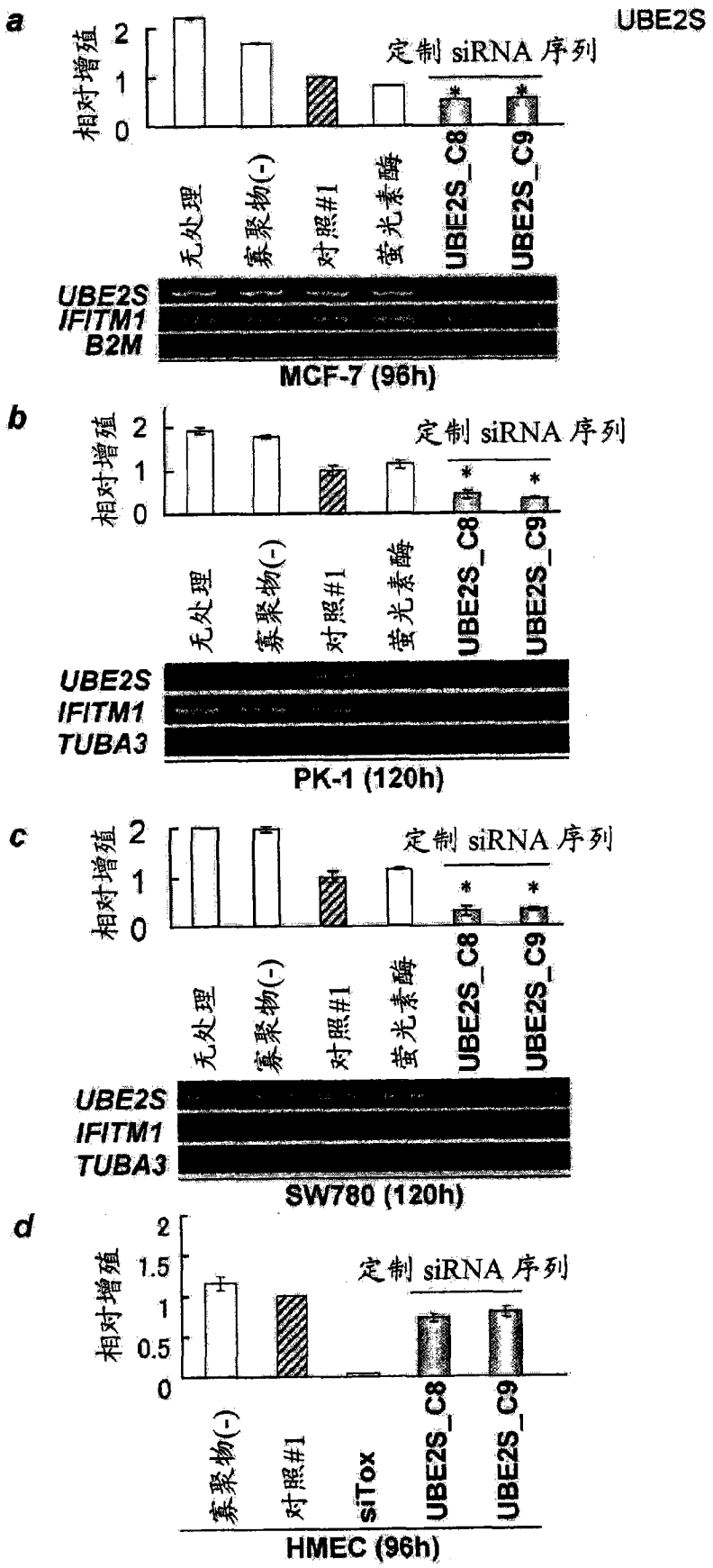

a
UBE2S
相对增殖
2
1
0
定制 siRNA 序列
*
*
无处理
寡聚物(-)
对照#1
萤光素酶
UBE2S_C8
UBE2S_C9
UBE2S
IFITM1
B2M
MCF-7 (96h)
b
相对增殖
2
1
0
定制 siRNA 序列
*
*
无处理
寡聚物(-)
对照#1
萤光素酶
UBE2S_C8
UBE2S_C9
UBE2S
IFITM1
TUBA3
PK-1 (120h)
c
相对增殖
2
1
0
定制 siRNA 序列
*
*
无处理
寡聚物(-)
对照#1
萤光素酶
UBE2S_C8
UBE2S_C9
UBE2S
IFITM1
TUBA3
SW780 (120h)
d
相对增殖
2
1.5
1
0.5
0
定制 siRNA 序列
寡聚物(-)
对照#1
siTox
UBE2S_C8
UBE2S_C9
HMEC (96h)

图 4

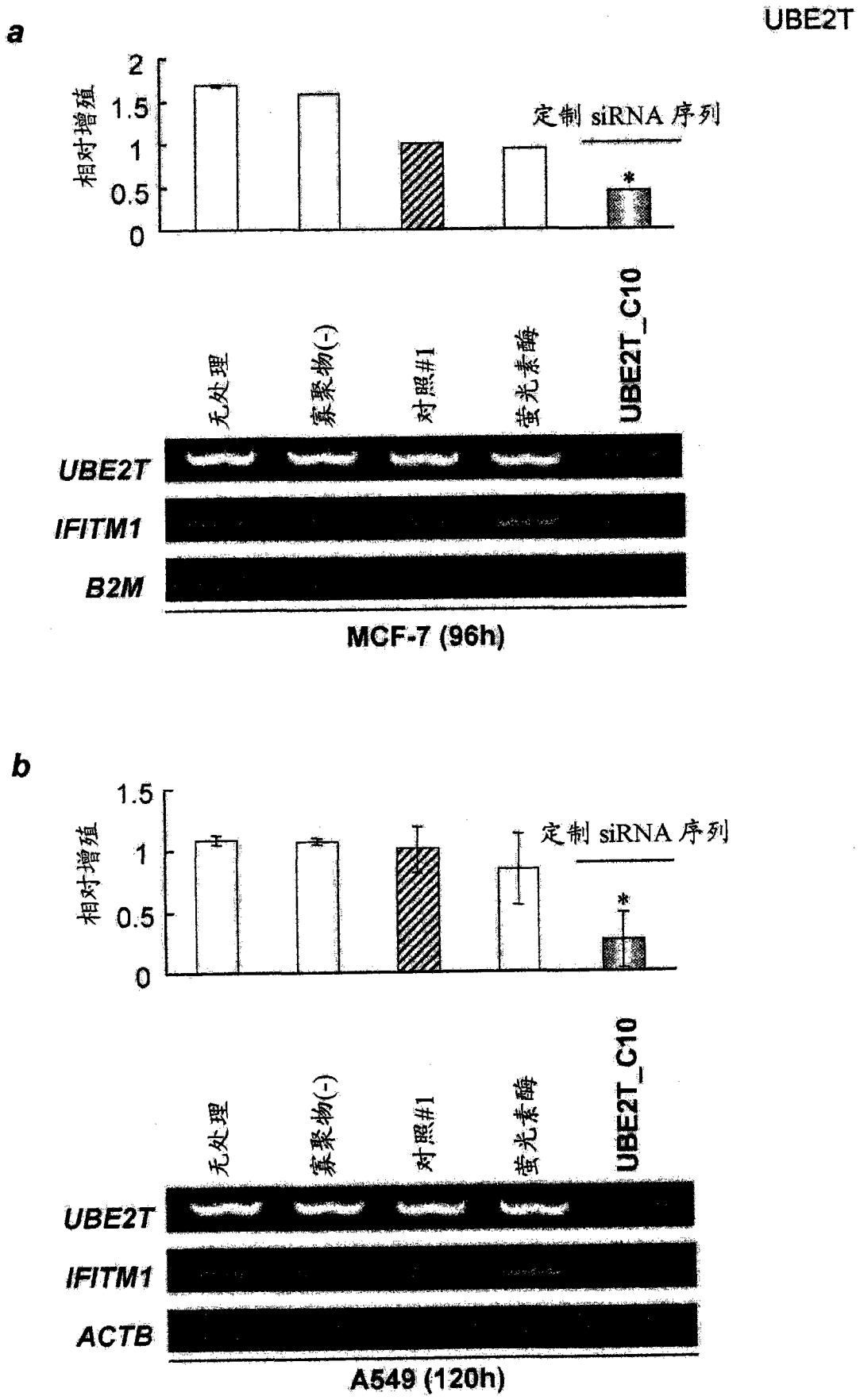

a
UBE2T
相对增殖
2
1.5
1
0.5
0
定制 siRNA 序列
*
无处理
寡聚物(-)
对照#1
萤光素酶
UBE2T_C10
UBE2T
IFITM1
B2M
MCF-7 (96h)
b
相对增殖
1.5
1
0.5
0
定制 siRNA 序列
*
无处理
寡聚物(-)
对照#1
萤光素酶
UBE2T_C10
UBE2T
IFITM1
ACTB
A549 (120h)

图 5-1

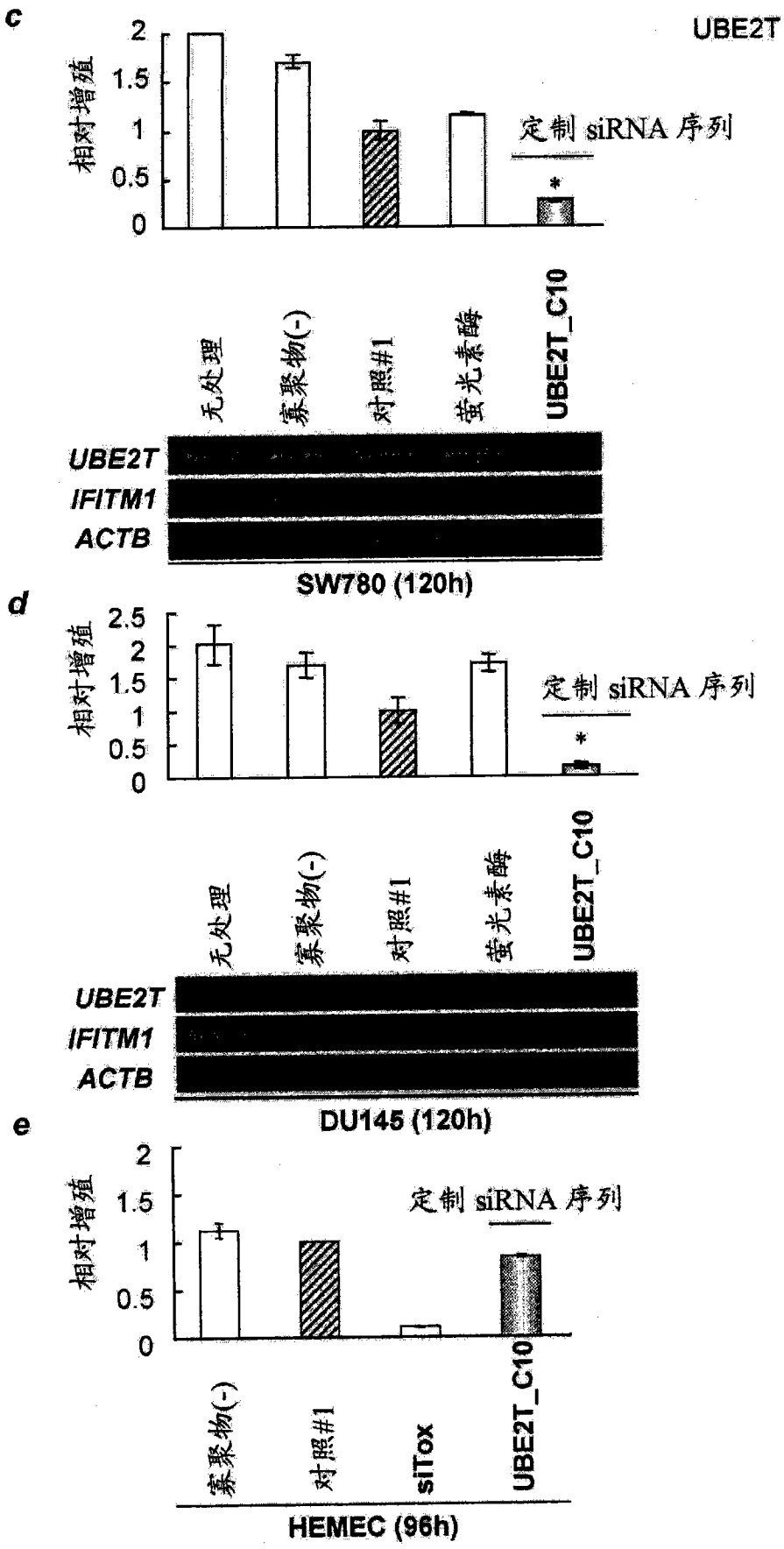

c
UBE2T
相对增殖
2
1.5
1
0.5
0
定制 siRNA 序列
*
无处理
寡聚物(-)
对照#1
萤光素酶
UBE2T_C10
UBE2T
IFITM1
ACTB
SW780 (120h)
d
相对增殖
2.5
2
1.5
1
0.5
0
定制 siRNA 序列
*
无处理
寡聚物(-)
对照#1
萤光素酶
UBE2T_C10
UBE2T
IFITM1
ACTB
DU145 (120h)
e
相对增殖
2
1.5
1
0.5
0
定制 siRNA 序列
寡聚物(-)
对照#1
siTox
UBE2T_C10
HEMEC (96h)

图 5-2

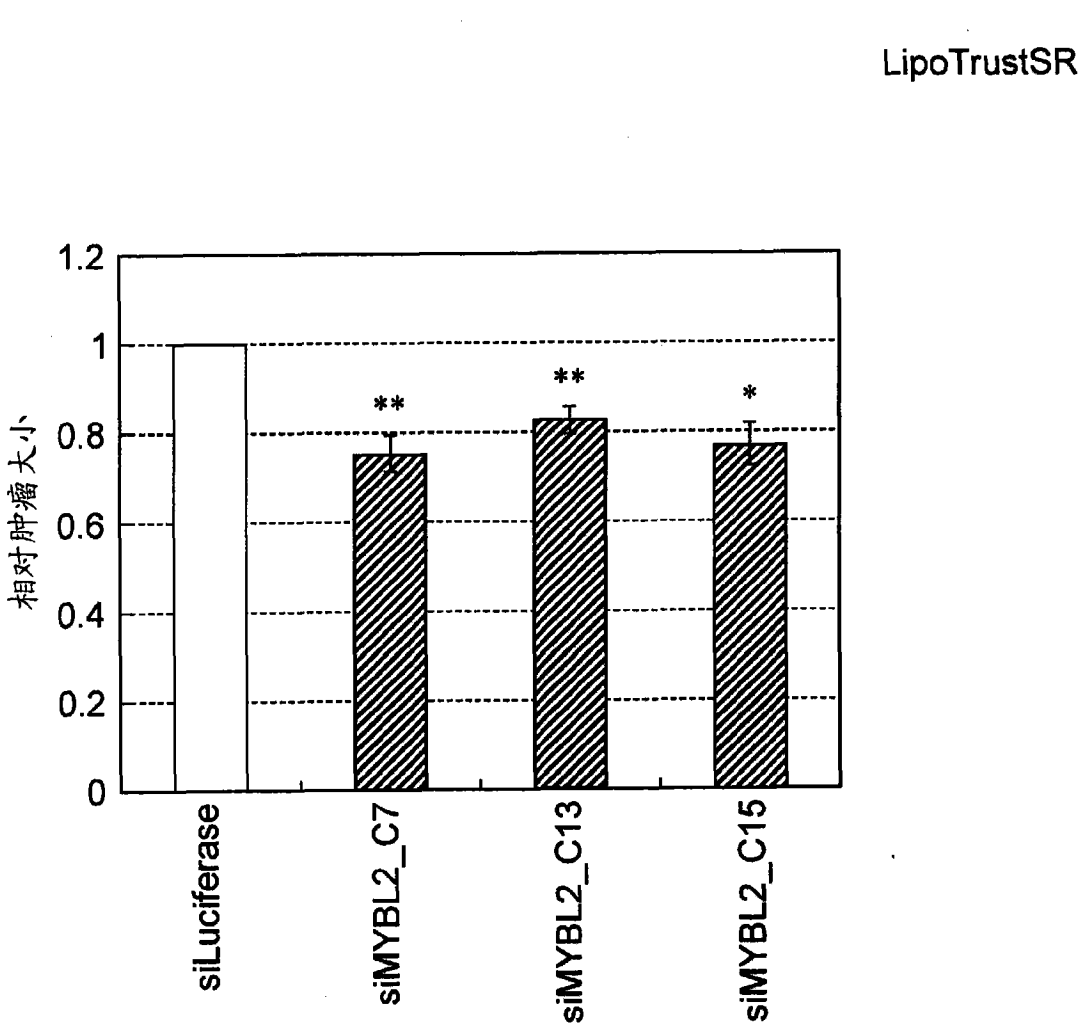

a
LipoTrustSR
相对肿瘤大小
1.2
1
0.8
0.6
0.4
0.2
0
**
**
*
siLuciferase
siMYBL2_C7
siMYBL2_C13
siMYBL2_C15

图 6-1

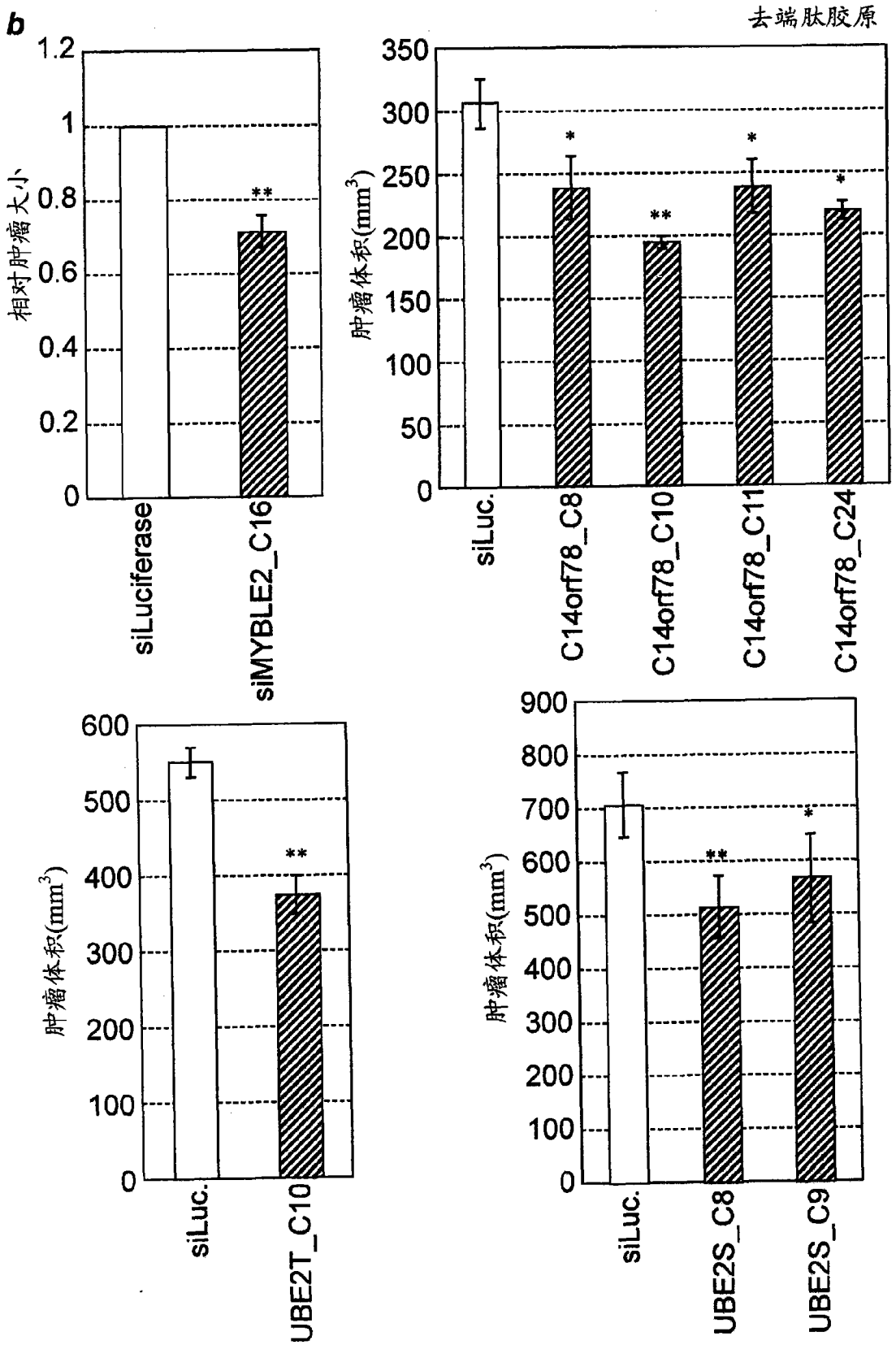

b
去端肽胶原
相对肿瘤大小
1.2
1
0.8
0.6
0.4
0.2
0
**
siLuciferase
siMYBLE2_C16
肿瘤体积(mm3)
350
300
250
200
150
100
50
0
*
**
*
*
siLuc.
C14orf78_C8
C14orf78_C10
C14orf78_C11
C14orf78_C24
肿瘤体积(mm3)
600
500
400
300
200
100
0
**
siLuc.
UBE2T_C10
肿瘤体积(mm3)
900
800
700
600
500
400
300
200
100
0
**
*
siLuc.
UBE2S_C8
UBE2S_C9

图 6-2