

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12N 15/31

C07K 14/35 A61K 38/16

C12N 15/62 G01N 33/569

C12Q 1/68 C12N 5/10

C12N 1/21

//(A61K39/04, C12N1/21,

1:19)

[12] 发明专利说明书

[21] ZL 专利号 96197639.X



[43] 授权公告日 2003 年 8 月 6 日

[11] 授权公告号 CN 1117149C

[22] 申请日 1996.8.30 [21] 申请号 96197639.X

[30] 优先权

[32] 1995.9.1 [33] US [31] 08/523,436

[32] 1995.9.22 [33] US [31] 08/533,634

[32] 1996.3.22 [33] US [31] 08/620,874

[32] 1996.6.5 [33] US [31] 08/659,683

[32] 1996.7.12 [33] US [31] 08/680,574

[86] 国际申请 PCT/US96/14674 1996.8.30

[87] 国际公布 WO97/09428 英 1997.3.13

[85] 进入国家阶段日期 1998.4.15

[71] 专利权人 科里克萨有限公司

地址 美国华盛顿

[72] 发明人 S·G·里德 Y·A·W·斯基凯

D·C·狄隆 A·坎泊斯-尼托

R·豪格顿

T·H·维德维克

D·R·特瓦德泽克

[56] 参考文献

WO9501440A 1995.01.12 A61K39/04

INFECTION AND IMMUNITY, 55 (6) 1987 - 06 -

01 D. B. Young et al, "Screening of a Recombinant Myco-

bacterial DNA Library with Polyclonal antiserum and

审查员 沈志红

[74] 专利代理机构 中国国际贸易促进委员会专利

商标事务所

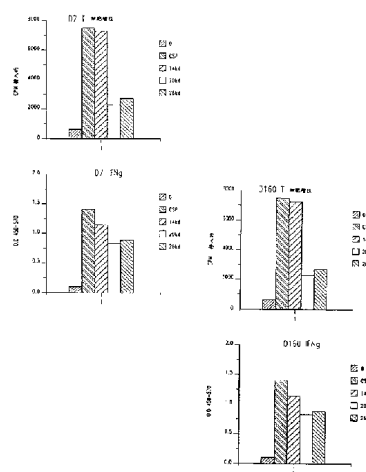
代理人 李 瑛

权利要求书 2 页 说明书 152 页 附图 3 页

[54] 发明名称 用于免疫治疗和诊断结核病的化合物和方法

[57] 摘要

本文公开了用于诱导针对结核病的保护性免疫的化合物和方法。所提供的化合物包括多肽以及编码这些多肽的 DNA 分子,所说的多肽含有一种或多种结核分枝杆菌蛋白质的至少一种免疫原性部分。这些化合物可以配制成供抗结核分枝杆菌感染之免疫的疫苗和/或药物组合物,或者可以用于诊断结核病。



1. 一种多肽，该多肽包含可溶性结核分枝杆菌抗原或仅在保守取代和/或修饰上不同的该抗原的变体的免疫原性部分，其中所说的抗原包含由选自下组的 DNA 序列编码的氨基酸序列：SEQ ID NO. 1，2，4-10，13-52，99 和 101 中所示的序列和这些序列的互补体。
2. 一种 DNA 分子，该分子包含编码按照权利要求 1 的多肽的核苷酸序列。
3. 一种表达载体，该载体包含按照权利要求 2 的 DNA 分子。
4. 一种宿主细胞，该宿主细胞由按照权利要求 3 的表达载体转化过。
5. 权利要求 4 的宿主细胞，其中所说的宿主细胞选自大肠杆菌、酵母和哺乳动物细胞。
6. 一种药物组合物，该组合物包含一种或多种按照权利要求 1 的多肽和生理学上可接受的载体。
7. 一种药物组合物，该组合物包含一种或多种按照权利要求 2 的 DNA 分子和生理学上可接受的载体。
8. 一种疫苗，该疫苗包含一种或多种按照权利要求 1 的多肽和一种非特异性免疫反应增强剂。
9. 权利要求 8 的疫苗，其中所说的非特异性免疫反应增强剂是佐剂。
10. 一种疫苗，该疫苗包含一种或多种按照权利要求 2 的 DNA 分子和一种非特异性免疫反应增强剂。
11. 按照权利要求 10 的疫苗，其中所说的非特异性免疫反应增强剂是佐剂。
12. 一种融合蛋白，该融合蛋白包含两种或多种按照权利要求 1 的多肽。
13. 一种融合蛋白，该融合蛋白包含一种或多种按照权利要求 1 的多肽和 ESAT-6。
14. 一种药物组合物，该组合物包含按照权利要求 12 或 13 的融合蛋白和生理学上可接受的载体。
15. 一种疫苗，该疫苗包含按照权利要求 12 或 13 的融合蛋白和一种

非特异性免疫反应增强剂。

16. 权利要求 15 的疫苗,其中所说的非特异性免疫反应增强剂是佐剂。

17. 一种诊断试剂盒,该试剂盒包含:

(a) 按照权利要求 1 的多肽; 和

(b) 足以把所说的多肽与患者的皮肤细胞接触的装置。

用于免疫治疗和诊断结核病的化合物和方法

技术领域

本发明总的来说涉及检测、治疗和预防结核分枝杆菌感染的方法。更具体地说本发明涉及包含结核分枝杆菌抗原，或其部分或变体的多肽，以及这些多肽在针对结核分枝杆菌感染的诊断和免疫接种上的用途。

发明背景

结核病是一种慢性传染病，一般由结核分枝杆菌感染引起。它在发展中国家是一种主要的疾病，在世界上发达地区也是一个日益严重的问题，每年有约 8 百万新病例和 3 百万人死亡。虽然感染可以在相当长一段时间内无症状，但是该疾病最常见地表现为急性肺炎，导致发热和非排痰性咳嗽。如果不进行治疗，则常常会出现严重的并发症并导致死亡。

虽然一般地可以采用多种抗生素控制结核病，但这样的治疗不足以阻止该疾病的传播。传染的个体可以是无症状的，但有时是传染性的。此外，虽然符合治疗方案是关键性的，但患者的行为难以监测。某些患者不完成治疗过程，这可以导致无效的治疗并产生药物抗性。

抑制结核病的传播需要有效的免疫接种和准确地早期诊断该疾病。当前，用活细菌接种是诱导保护性免疫最有效的方法。用于这一目的的最普通的分枝杆菌属是卡介苗 (BCG) 和牛型分枝杆菌的无毒菌株。然而，BCG 的安全性和有效性上存在争议，并且一些国家(如，美国)不接种一般公众。诊断一般利用皮试进行，这牵涉到真皮内接触结核菌素 PPD(纯化的蛋白质衍生物)。在注射之后 48-72 小时，抗原特异性 T-细胞反应在注射部位导致可测量的潜伏(incubation)，这表明接触到分枝杆菌抗原。然而，这一实验的灵敏度和特异性一直存在问题，用 BCG 接种的个体与感染的个体不能区别。

虽然巨噬细胞已显示出作为结核分枝杆菌免疫性的主要的效应细胞，

但T-细胞是这种免疫性的主要的诱导物。T-细胞在针对结核分枝杆菌感染的保护中的十分重要的作用由在爱滋病患者中结核分枝杆菌频繁发生说明, 因为CD4 T-细胞的耗竭与人免疫缺损病毒(HIV)感染相关。分枝杆菌属反应性CD4 T-细胞已显示出是 γ -干扰素(IFN- γ)的有力的生产者, 后者依次已显示出在小鼠中触发巨噬细胞的抗分枝杆菌作用。尽管IFN- γ 在人类中的作用还不太清楚, 但研究已表明1, 25-二羟基-维生素D3单独或与IFN- γ 或肿瘤坏死因子- α 一道激活人巨噬细胞以阻止结核分枝杆菌感染。此外, 已知IFN- γ 刺激人巨噬细胞产生1, 25-二羟基-维生素D3。同样地, IL-12已显示出在刺激对结核分枝杆菌感染的抗性中起作用。有关结核分枝杆菌感染的免疫学参见Chan和Kaufmann, 结核病: 病理, 预防和治疗, Boom(编者), ASM出版社, 华盛顿, DC, 1994。

因此, 本领域需要用于预防、治疗和检测结核病的改进的疫苗和方法。本发明满足了这一需要并进一步提供了其它相关优点。

发明概要

简言之, 本发明提供了用于预防和诊断结核病的化合物和方法。在一个方面, 本发明提供了一些多肽, 这些多肽包含可溶性结核分枝杆菌抗原或仅在保守取代和/或修饰上不同的该抗原的变体的免疫原性部分。在这一方面的一个实施方案中, 所说的可溶性抗原具有一种以下N端序列:

- (a) Asp-Pro-Val-Asp-Ala-Val-Ile-Asn-Thr-Thr-Cys-Asn-Tyr-Gly-Gln-Val-Val-Ala-Ala-Leu; (SEQ ID NO. 120)
- (b) Ala-Val-Glu-Ser-Gly-Met-Leu-Ala-Leu-Gly-Thr-Pro-Ala-Pro-Ser (SEQ ID NO. 121)
- (c) Ala-Ala-Met-Lys-Pro-Arg-Thr-Gly-Asp-Gly-Pro-Leu-Glu-Ala-Ala-Lys-Glu-Gly-Arg; (SEQ ID NO. 122)
- (d) Tyr-Tyr-Trp-Cys-Pro-Gly-Gln-Pro-Phe-Asp-Pro-Ala-Trp-Gly-Pro; (SEQ ID NO. 123)
- (e) Asp-Ile-Gly-Ser-Glu-Ser-Thr-Glu-Asp-Gln-Gln-Xaa-Ala-Val; (SEQ ID NO. 124)
- (f) Ala-Glu-Glu-Ser-Ile-Ser-Thr-Xaa-Glu-Xaa-Ile-Val-Pro;

(SEQ ID NO. 125)

- (g) Asp-Pro-Glu-Pro-Ala-Pro-Pro-Val-Pro-Thr-Thr-Ala-Ala-Ser-Pro-Pro-Ser; (SEQ ID NO. 126)
- (h) Ala-Pro-Lys-Thr-Tyr-Xaa-Glu-Glu-Leu-Lys-Gly-Thr-Asp-Thr-Gly; (SEQ ID NO. 127)
- (i) Asp-Pro-Ala-Ser-Ala-Pro-Asp-Val-Pro-Thr-Ala-Ala-Gln-Leu-Thr-Ser-Leu-Leu-Asn-Ser-Leu-Ala-Asp-Pro-Asn-Val-Ser-Phe-Ala-Asn; (SEQ ID NO. 128) 和
- (j) Xaa-Asp-Ser-Glu-Lys-Ser-Ala-Thr-Ile-Lys-Val-Thr-Asp-Ala-Ser; (SEQ ID NO. 134)
- (k) Ala-Gly-Asp-Thr-Xaa-Ile-Tyr-Ile-Val-Gly-Asn-Leu-Thr-Ala-Asp; (SEQ ID NO. 135) 或
- (l) Ala-Pro-Glu-Ser-Gly-Ala-Gly-Leu-Gly-Gly-Thr-Val-Gln-Ala-Gly; (SEQ ID NO. 136)

其中 Xaa 可以是任何氨基酸。

在一个相关的方面，本发明提供了一些多肽，这些多肽包含结核分枝杆菌抗原或仅在保守取代和/或修饰上不同的该抗原的变体的免疫原性部分，所说的抗原具有一种以下的 N 端序列：

- (m) Xaa-Tyr-Ile-Ala-Tyr-Xaa-Thr-Thr-Ala-Gly-Ile-Val-Pro-Gly-Lys-Ile-Asn-Val-His-Leu-Val; (SEQ ID NO. 137) 或
- (n) Asp-Pro-Pro-Asp-Pro-His-Gln-Xaa-Asp-Met-Thr-Lys-Gly-Tyr-Tyr-Pro-Gly-Gly-Arg-Arg-Xaa-Phe; (SEQ ID NO. 129),

其中 Xaa 可以是任何氨基酸。

在另一个实施方案中，所说的抗原包含可溶性结核分枝杆菌抗原或仅在保守取代和/或修饰上不同的该抗原的变体的免疫原性部分，其中所说的抗原包含由选自下组的 DNA 序列编码的氨基酸序列：SEQ ID NO. 1, 2, 4-10, 13-25, 52, 99 和 101 中所示的序列、这些序列的补体、以及在中等严格条件下与 SEQ ID NO. 1, 2, 4-10, 13-25, 52, 99 和 101 中所示的序列杂交的 DNA 序列或它们的补体。

在一个相关的方面，所说的多肽包含结核分枝杆菌抗原或仅在保守取

代和/或修饰上不同的该抗原的变体的免疫原性部分,其中所说的抗原包含由选自下组的DNA序列编码的氨基酸序列:SEQ ID NO. 26-51中所示的序列、这些序列的补体、和在中等严格条件下与SEQ ID NO. 26-51中所示的序列杂交的DNA序列或它们的补体。

在一个相关的方面,本发明提供了编码上述多肽的DNA序列,包含这些DNA序列的表达载体和用这样的表达载体转化或转染的宿主细胞。

另一方面,本发明提供了包含第一与第二发明多肽或者是发明多肽与已知的结核分枝杆菌抗原的融合蛋白。

在其它方面,本发明提供了药物组合物,这些组合物包含一种或多种上述多肽或者编码这些多肽的DNA分子与生理学上可接受的载体。本发明也提供了包含一种或多种以上描述的多肽和非特异性免疫反应增强剂(enhancer)的疫苗,以及包含一种或多种编码这些多肽的DNA序列和非特异性免疫反应增强剂的疫苗。

在另一方面,本发明提供了在患者中诱导保护性免疫的方法,该方法包括向患者施用有效量的一种或多种上述多肽。

在本发明的另一方面,本发明提供了用于在患者中检测结核病的方法和诊断试剂盒。所说的方法包括使患者的皮肤细胞与一种或多种以上的多肽接触,并检测在病人皮肤上的免疫反应。所说的诊断试剂盒包含一种或多种以上的多肽以及足以把所说的多肽与患者皮肤细胞接触的装置。

参照下列详细描述和附图,本发明的这些和其他方面会很清楚。本文所公开的所有参考文献与它们单个并入作为参考一样,以它们的整体由本文一并参考。

附图和序列识别号的简要描述

图1A和B说明实施例1中描述的14Kd、20Kd和26Kd抗原对分别来源于第一和第二结核分枝杆菌免疫供体的T-细胞的增殖和干扰素- γ 产生的刺激作用。

图2说明两种代表性多肽TbRa3和TbRa9对来源于结核分枝杆菌免疫个体的T-细胞的增殖和干扰素- γ 产生的刺激作用。

SEQ ID NO. 1是TbRa1的DNA序列。

SEQ ID NO. 2 是 TbRa10 的 DNA 序列。
SEQ ID NO. 3 是 TbRa11 的 DNA 序列。
SEQ ID NO. 4 是 TbRa12 的 DNA 序列。
SEQ ID NO. 5 是 TbRa13 的 DNA 序列。
SEQ ID NO. 6 是 TbRa16 的 DNA 序列。
SEQ ID NO. 7 是 TbRa17 的 DNA 序列。
SEQ ID NO. 8 是 TbRa18 的 DNA 序列。
SEQ ID NO. 9 是 TbRa19 的 DNA 序列。
SEQ ID NO. 10 是 TbRa24 的 DNA 序列。
SEQ ID NO. 11 是 TbRa26 的 DNA 序列。
SEQ ID NO. 12 是 TbRa28 的 DNA 序列。
SEQ ID NO. 13 是 TbRa29 的 DNA 序列。
SEQ ID NO. 14 是 TbRa2A 的 DNA 序列。
SEQ ID NO. 15 是 TbRa3 的 DNA 序列。
SEQ ID NO. 16 是 TbRa32 的 DNA 序列。
SEQ ID NO. 17 是 TbRa35 的 DNA 序列。
SEQ ID NO. 18 是 TbRa36 的 DNA 序列。
SEQ ID NO. 19 是 TbRa4 的 DNA 序列。
SEQ ID NO. 20 是 TbRa9 的 DNA 序列。
SEQ ID NO. 21 是 TbRaB 的 DNA 序列。
SEQ ID NO. 22 是 TbRaC 的 DNA 序列。
SEQ ID NO. 23 是 TbRaD 的 DNA 序列。
SEQ ID NO. 24 是 YYWCPG 的 DNA 序列。
SEQ ID NO. 25 是 AAMK 的 DNA 序列。
SEQ ID NO. 26 是 TbL-23 的 DNA 序列。
SEQ ID NO. 27 是 TbL-24 的 DNA 序列。
SEQ ID NO. 28 是 TbL-25 的 DNA 序列。
SEQ ID NO. 29 是 TbL-28 的 DNA 序列。
SEQ ID NO. 30 是 TbL-29 的 DNA 序列。
SEQ ID NO. 31 是 TbH-5 的 DNA 序列。

SEQ ID NO. 32 是 TbH-8 的 DNA 序列。
SEQ ID NO. 33 是 TbH-9 的 DNA 序列。
SEQ ID NO. 34 是 TbM-1 的 DNA 序列。
SEQ ID NO. 35 是 TbM-3 的 DNA 序列。
SEQ ID NO. 36 是 TbM-6 的 DNA 序列。
SEQ ID NO. 37 是 TbM-7 的 DNA 序列。
SEQ ID NO. 38 是 TbM-9 的 DNA 序列。
SEQ ID NO. 39 是 TbM-12 的 DNA 序列。
SEQ ID NO. 40 是 TbM-13 的 DNA 序列。
SEQ ID NO. 41 是 TbM-14 的 DNA 序列。
SEQ ID NO. 42 是 TbM-15 的 DNA 序列。
SEQ ID NO. 43 是 TbH-4 的 DNA 序列。
SEQ ID NO. 44 是 TbH-4-FWD 的 DNA 序列。
SEQ ID NO. 45 是 TbH-12 的 DNA 序列。
SEQ ID NO. 46 是 Tb38-1 的 DNA 序列。
SEQ ID NO. 47 是 Tb38-4 的 DNA 序列。
SEQ ID NO. 48 是 TbL-17 的 DNA 序列。
SEQ ID NO. 49 是 TbL-20 的 DNA 序列。
SEQ ID NO. 50 是 TbL-21 的 DNA 序列。
SEQ ID NO. 51 是 TbH-16 的 DNA 序列。
SEQ ID NO. 52 是 DPEP 的 DNA 序列。
SEQ ID NO. 53 是 DPEP 的推定的氨基酸序列。
SEQ ID NO. 54 是 DPV N-端抗原的蛋白质序列。
SEQ ID NO. 55 是 AVGS N-端抗原的蛋白质序列。
SEQ ID NO. 56 是 AAMK N-端抗原的蛋白质序列。
SEQ ID NO. 57 是 YYWC N-端抗原的蛋白质序列。
SEQ ID NO. 58 是 DIGS N-端抗原的蛋白质序列。
SEQ ID NO. 59 是 AEES N-端抗原的蛋白质序列。
SEQ ID NO. 60 是 DPEP N-端抗原的蛋白质序列。
SEQ ID NO. 61 是 APKT N-端抗原的蛋白质序列。

SEQ ID NO. 62 是 DPAS N-端抗原的蛋白质序列。
SEQ ID NO. 63 是 TbRa1 的推定的氨基酸序列。
SEQ ID NO. 64 是 TbRa10 的推定的氨基酸序列。
SEQ ID NO. 65 是 TbRa11 的推定的氨基酸序列。
SEQ ID NO. 66 是 TbRa12 的推定的氨基酸序列。
SEQ ID NO. 67 是 TbRa13 的推定的氨基酸序列。
SEQ ID NO. 68 是 TbRa16 的推定的氨基酸序列。
SEQ ID NO. 79 是 TbRa17 的推定的氨基酸序列。
SEQ ID NO. 70 是 TbRa18 的推定的氨基酸序列。
SEQ ID NO. 71 是 TbRa19 的推定的氨基酸序列。
SEQ ID NO. 72 是 TbRa24 的推定的氨基酸序列。
SEQ ID NO. 73 是 TbRa26 的推定的氨基酸序列。
SEQ ID NO. 74 是 TbRa28 的推定的氨基酸序列。
SEQ ID NO. 75 是 TbRa29 的推定的氨基酸序列。
SEQ ID NO. 76 是 TbRa2A 的推定的氨基酸序列。
SEQ ID NO. 77 是 TbRa3 的推定的氨基酸序列。
SEQ ID NO. 78 是 TbRa32 的推定的氨基酸序列。
SEQ ID NO. 79 是 TbRa35 的推定的氨基酸序列。
SEQ ID NO. 80 是 TbRa36 的推定的氨基酸序列。
SEQ ID NO. 81 是 TbRa4 的推定的氨基酸序列。
SEQ ID NO. 82 是 TbRa9 的推定的氨基酸序列。
SEQ ID NO. 83 是 TbRaB 的推定的氨基酸序列。
SEQ ID NO. 84 是 TbRaC 的推定的氨基酸序列。
SEQ ID NO. 85 是 TbRaD 的推定的氨基酸序列。
SEQ ID NO. 86 是 YYWCPG 的推定的氨基酸序列。
SEQ ID NO. 87 是 TbAAMK 的推定的氨基酸序列。
SEQ ID NO. 88 是 Tb38-1 的推定的氨基酸序列。
SEQ ID NO. 89 是 TbH-4 的推定的氨基酸序列。
SEQ ID NO. 90 是 TbH-8 的推定的氨基酸序列。
SEQ ID NO. 91 是 TbH-9 的推定的氨基酸序列。

SEQ ID NO. 92 是 TbH-12 的推定的氨基酸序列。

SEQ ID NO. 93 是 Tb38-1 肽 1 的氨基酸序列。

SEQ ID NO. 94 是 Tb38-1 肽 2 的氨基酸序列。

SEQ ID NO. 95 是 Tb38-1 肽 3 的氨基酸序列。

SEQ ID NO. 96 是 Tb38-1 肽 4 的氨基酸序列。

SEQ ID NO. 97 是 Tb38-1 肽 5 的氨基酸序列。

SEQ ID NO. 98 是 Tb38-1 肽 6 的氨基酸序列。

SEQ ID NO. 99 是 DPAS 的 DNA 序列。

SEQ ID NO. 100 是 DPAS 的推定的氨基酸序列。

SEQ ID NO. 101 是 DPV 的 DNA 序列。

SEQ ID NO. 102 是 DPV 的推定的氨基酸序列。

SEQ ID NO. 103 是 ESAT-6 的 DNA 序列。

SEQ ID NO. 104 是 ESAT-6 的推定的氨基酸序列。

SEQ ID NO. 105 是 TbH-8-2 的 DNA 序列。

SEQ ID NO. 106 是 TbH-9FL 的 DNA 序列。

SEQ ID NO. 107 是 TbH-9FL 的推定的氨基酸序列。

SEQ ID NO. 108 是 TbH-9-1 的 DNA 序列。

SEQ ID NO. 109 是 TbH-9-1 的推定的氨基酸序列。

SEQ ID NO. 110 是 TbH-9-4 的 DNA 序列。

SEQ ID NO. 111 是 TbH-9-4 的推定的氨基酸序列。

SEQ ID NO. 112 是 Tb38-1F2 IN 的 DNA 序列。

SEQ ID NO. 113 是 Tb38-1F2 RP 的 DNA 序列。

SEQ ID NO. 114 是 Tb37-FL 的推定的氨基酸序列。

SEQ ID NO. 115 是 Tb38-IN 的推定的氨基酸序列。

SEQ ID NO. 116 是 Tb38-1F3 的 DNA 序列。

SEQ ID NO. 117 是 Tb38-1F3 的推定的氨基酸序列。

SEQ ID NO. 118 是 Tb38-1F5 的 DNA 序列。

SEQ ID NO. 119 是 Tb38-1F6 的 DNA 序列。

SEQ ID NO. 120 是 DPV 的推定的 N-端氨基酸序列。

SEQ ID NO. 121 是 AVGS 的推定的 N-端氨基酸序列。

SEQ ID NO. 122 是 AAMK 的推定的 N-端氨基酸序列。
SEQ ID NO. 123 是 YYWC 的推定的 N-端氨基酸序列。
SEQ ID NO. 124 是 DIGS 的推定的 N-端氨基酸序列。
SEQ ID NO. 125 是 AAES 的推定的 N-端氨基酸序列。
SEQ ID NO. 126 是 DPEP 的推定的 N-端氨基酸序列。
SEQ ID NO. 127 是 APKT 的推定的 N-端氨基酸序列。
SEQ ID NO. 128 是 DPAS 的推定的 N-端氨基酸序列。
SEQ ID NO. 129 是 DPPD N-端抗原的蛋白质序列。
SEQ ID NO. 130-133 是四种 DPPD 溴化氰片段的蛋白质序列。
SEQ ID NO. 134 是 XDS 抗原的 N-端蛋白质序列。
SEQ ID NO. 135 是 AGD 抗原的 N-端蛋白质序列。
SEQ ID NO. 136 是 APE 抗原的 N-端蛋白质序列。
SEQ ID NO. 137 是 XYI 抗原的 N-端蛋白质序列。

发明详述

如上所述, 本发明总的来说涉及预防、治疗和诊断结核病的组合物和方法。本发明的组合物包含一些多肽, 这些多肽包含结核分枝杆菌抗原或仅在保守取代和/或修饰上不同的该抗原的变体的至少一种免疫原性部分。在本发明的范围内的多肽包括, 但不限于, 可溶性结核分枝杆菌抗原。"可溶性结核分枝杆菌抗原"是存在于结核分枝杆菌培养物滤液中的结核分枝杆菌源的蛋白质。如本文所使用的, 术语"多肽"包括任何长度的氨基酸链, 包括全长蛋白质(即, 抗原), 其中的氨基酸残基由共价肽键连接。这样, 包含上述一种抗原的免疫原性部分的多肽可以是完全由免疫原性部分组成的, 或者可以含有附加序列。所说的附加序列可以是来源于天然结核分枝杆菌抗原或者可以是异源的, 这样的序列可以是(但不需要是)免疫原性的。

本文使用的"免疫原性"指在患者(例如人)和/或生物样品中激发免疫反应(例如, 细胞免疫)的能力。具体地说, 免疫原性的抗原(或这种抗原的免疫原性部分或其它变体)能够在生物样品中刺激细胞增殖、白介素-12 产生和/或 γ -干扰素产生, 所说的生物样品包含一种或多种选自 T-细胞、NK

细胞、B细胞和巨噬细胞的细胞，这些细胞来源于结核分枝杆菌感染个体。包含一种或多种结核分枝杆菌抗原的至少一种免疫原性部分的多肽一般来说可以用于在患者中检测结核病和诱导针对结核病的保护性免疫。

本发明的组合物与方法也包括上述多肽的变体。如本文所使用的"变体"是仅在保守取代和/或修饰上不同于天然抗原(以便所述多肽诱导免疫反应的能力得到保留)的多肽。通过采用本文描述的代表性方法修饰一种上述多肽序列并评价修饰的多肽的免疫原性特性可以一般性地鉴别这样的变体。

"保守取代"是这样一种取代，其中一种氨基酸取代具有类似性质的另一种氨基酸，以便肽化学领域的技术人员可以期望多肽的二级结构与亲水性质实质上不变。一般来说，下组氨基酸代表保守取代：(1) ala、pro、gly、glu、asp、gln、asn、ser、thr；(2) cys、ser、tyr、thr；(3) val、ile、leu、met、ala、phe；(4) lys、arg、his；和(5) phe、tyr、trp、his。

变体也可以(或选择性地)是由例如氨基酸缺失或者添加(对多肽免疫原性性质，二级结构和亲水性质具有最小限度的影响)修饰的。例如，多肽可以连接到蛋白质N端的信号(或前导)序列上，后者共翻译或翻译后指导蛋白质的转移。所述多肽也可以连结到使多肽容易合成，纯化以及鉴定或增强多肽结合到固相支持体上的接头和其他序列(例如 poly-His)上。例如，多肽可以连结到免疫球蛋白Fc区上。

在一个相关的方面，本文公开了组合多肽。"组合多肽"是包含至少一种上述免疫原性部分和一种或多种附加免疫原性结核分枝杆菌序列(其经由肽键连接到单一的氨基酸链上)的多肽。所述的序列可以直接连接(即没有间插氨基酸)或通过不明显降低组成多肽的免疫原性特性的接头序列(例如，Gly-Cys-Gly)连接。

一般来说，结核分枝杆菌抗原，编码这种抗原的DNA序列，可以以多种方法的任何一种制备。例如，可溶性抗原可以用本领域技术人员已知的方法从结核分枝杆菌培养物滤液分离，包括阴离子交换、反相层析。可以采用例如本文描述的代表性方法就激发免疫反应(例如细胞免疫反应)评价纯化的抗原。可以利用例如传统的Edman化学之类的技术对免疫原性抗原

进行部分测序。参见Edman和Berg, 欧洲生物化学杂志, 80: 116-132, 1967。

也可以用编码抗原的DNA序列(已插入到表达载体中并在合适的宿主中表达)重组产生免疫原性抗原。可以通过用特异性抗可溶性结核分枝杆菌抗原产生的抗血清(例如兔)筛选合适的结核分枝杆菌表达文库来分离编码可溶性抗原的DNA分子。可以用从感染了结核分枝杆菌的病人获得的血清筛选合适的结核分枝杆菌基因组或者cDNA表达文库鉴别编码抗原(抗原可以是或者可以不是可溶性的)的DNA序列。这样的筛选一般可以利用本领域技术人员熟知的技术完成, 例如在Sambrook等, 分子克隆: 实验室手册, 冷泉港实验室, 冷泉港, NY, 1989中所描述的那些。

编码可溶性抗原的DNA序列也可以通过就与简并寡核苷酸(该寡核苷酸来源于分离的可溶性抗原的部分氨基酸序列)杂交的DNA序列筛选适当的结核分枝杆菌cDNA或基因组DNA文库来获得。可以如(例如)Sambrook等, 分子克隆: 实验室手册, 冷泉港实验室, 冷泉港, NY(和该文引用的参考文献)中的描述设计和合成用于这种筛选的简并寡核苷酸序列, 并且完成筛选。也可以使用聚合酶链反应(PCR), 用本领域已知的方法用上述寡核苷酸, 以从cDNA或基因组文库分离核酸探针。然后可以使用所分离的探针完成文库的筛选。

此外, 可以直接采用外周血单核细胞(PBMCs)或T-细胞系或来源于结核分枝杆菌免疫个体的克隆筛选来源于结核分枝杆菌的基因组或cDNA文库。一般来说, 用于这种筛选的PBMCs和/或T-细胞可以如以下的描述制备。直接文库筛选一般可以这样进行, 即测定表达的重组蛋白质库在来源于结核分枝杆菌免疫个体T-细胞中诱导增殖和/或干扰素- γ 产生的能力。此外, 如下所述, 可以基于抗体反应性首先选择潜在的T-细胞抗原。

不论是什么制备方法, 本文描述的抗原(和其免疫原性部分)(其可以是也可以不是可溶性的)具有诱导免疫原性反应的能力。更具体地说, 所说的抗原具有在来源于结核分枝杆菌免疫个体T-细胞、NK细胞、B细胞和/或巨噬细胞中诱导增殖和/或细胞因子产生(即干扰素- γ 和/或白介素-12产生)的能力。当然, 用于评价对抗原的免疫原性反应的细胞类型的选择取决于所需的反应。例如, 利用包含B细胞和/或巨噬细胞的制剂最容易评价白

介素-12 的产生。结核分枝杆菌免疫个体被认为是这样的个体，其被认为借助于发动对结核分枝杆菌的有效的 T-细胞反应对结核病的发生具有抗性（即基本上没有疾病症状）。这样的个体可以基于强阳性（即大于约 10 mM 直径硬结）结核病蛋白质（PPD）真皮内皮肤试验反应和缺乏结核病的任何征兆或症状来鉴别。来源于结核分枝杆菌免疫个体的 T-细胞、NK 细胞、B 细胞和巨噬细胞可以用本领域普通技术人员已知的方法制备。例如，PBMCs（即，外周血单核细胞）制剂不需进一步分离组分细胞就可以使用。PBMCs 一般可以采用例如通过 Ficoll™（Winthrop 实验室，NY）的密度离心制备。用于本文描述的测定中的 T-细胞也可以直接从 PBMCs 纯化。此外，可以使用抗分枝杆菌蛋白质反应性的富 T-细胞系或者对单个分枝杆菌蛋白质反应性的 T-细胞克隆。这样的 T-细胞克隆可以通过，例如，将来源于结核分枝杆菌免疫个体的 PBMCs 与分枝杆菌蛋白质一起温育 2-4 周生成。这使得可以仅仅扩增分枝杆菌蛋白质-特异性 T-细胞，导致产生由这样的细胞单独组成的系。然后，采用本领域普通技术人员已知的方法，这些细胞可被克隆和用个体蛋白质试验，以更精确地限定个体 T-细胞特异性。一般来说，这样一些抗原被认为是免疫原性的，其在采用来源于结核分枝杆菌免疫个体的 T-细胞、NK 细胞、B 细胞和/或巨噬细胞进行的增殖和/或细胞因子产生（即干扰素- γ 和/或白介素-12 的产生）测定中试验为阳性。这样的测定可以采用例如以下所描述的代表性方法完成。这些抗原的免疫原性部分可以用类似的测定来鉴别，并且可以存在于本文所描述的多肽内。

多肽（例如，免疫原性抗原，或其部分或其它变体）诱导细胞增殖的能力通过使细胞（例如，T-细胞和/或 NK 细胞）与所说的多肽接触并测定所说细胞的增殖评价。一般来说，足以评价约 10^5 个细胞的多肽的量在约 10ng/mL 到约 100 μ g/mL 的范围内，优选地是约 10 μ g/mL。多肽与细胞的温育典型地在 37 $^{\circ}$ C 下进行约六天。在与多肽一起温育后，就增殖性反应测定细胞，这可以通过本领域普通技术人员已知的方法来评价，例如将细胞露置在放射性标记胸苷的脉冲下，并测定标记对细胞 DNA 的掺入。一般来说，导致产生在基底（即，对没有多肽时培养 of 细胞所观察到的增殖）之上的增殖至少增加 3 倍的多肽被认为能够诱导增殖。

多肽刺激细胞中干扰素- γ 和/或白介素-12 产生的能力可以通过使细胞

与所说的多肽接触并测量由细胞产生的干扰素- γ 或白介素-12 的水平来评价。一般来说, 足以评价约 10^5 个细胞的多肽的量在约 10ng/mL 到约 100 μ g/mL 的范围内, 优选地是约 10 μ g/mL。多肽可以(但是不必)固定化在固相支持体上, 如小珠或可被生物降解的微球体, 如在美国专利 4,897,268 和 5,075,109 中所描述的那些。多肽与细胞的温育典型地在 37 $^{\circ}$ C 下进行约六天。在与多肽一起温育后, 就干扰素- γ 和/或白介素-12(或者它们的一个或多个亚单位)测定细胞, 这可以通过本领域普通技术人员已知的方法来评价, 例如酶联免疫吸附测定(ELISA)或者在 IL-12 P70 亚单位的情况下的生物测定(例如测量 T-细胞增殖的测定)。一般来说, 导致产生每 mL 培养上清液(每 mL 包含 10^4 - 10^5 个 T-细胞)至少 50 pg 干扰素- γ 的多肽被认为能够刺激干扰素- γ 的产生。刺激每 10^5 个巨噬细胞或 B 细胞(或每 3×10^5 个 PBMC)产生 10 pg/mL IL-12 P70 亚单位和/或至少 100 pg/mL IL-12 P40 亚单位的多肽被认为能够刺激 IL-12 的产生。

一般来说, 免疫原性抗原是这样一些抗原, 其刺激来源于至少约 25% 的结核分枝杆菌免疫个体的 T-细胞、NK 细胞、B 细胞和/或巨噬细胞的增殖和/或细胞因子产生(即, 干扰素- γ 和/或白介素-12 产生)。在这些免疫原性抗原中, 具有较好治疗性质的多肽可以基于在上述测定中的反应程度和基于观察到其反应的个体的百分比来辨别。此外, 基于较好治疗性质的抗原将不在体外刺激来源于约 25% 以上的非结核分枝杆菌免疫个体的细胞的增殖和/或细胞因子产生, 从而消除非特异性(由于结核分枝杆菌-反应性细胞)的反应。在高百分比的来源于结核分枝杆菌免疫个体的 T-细胞、NK 细胞、B 细胞和/或巨噬细胞制剂中诱导反应的那些抗原(其它个体的细胞制剂中具有低的发生率)具有较好的治疗性质。

具有较好治疗性质的抗原也可以基于它们(当作为疫苗施用)在实验动物中降低结核分枝杆菌感染的严重程度的能力来鉴别。用于实验动物上的合适的疫苗制剂在以下详细描述。效能可以基于实验感染后抗原提供至少约 50 % 的细菌数降低和/或至少约 40% 的死亡率降低的能力来确定。合适的实验动物包括小鼠、豚鼠和灵长类。

具有较好诊断性质的抗原一般可以基于在具有活动性结核病之个体上进行的真皮内皮肤试验中, 而不是在未被结核分枝杆菌感染之个体上进行的

的试验中激发反应的能力来鉴别。皮肤试验一般也可以按以下进行,至少 5 mm 的硬结反应被认为是阳性的。

也可以采用本领域已知的技术(例如在 Paul, 基础免疫学, 第三版, Raven 出版社, 1993, pp. 243-247 和该文引用的参考文献中描述的那些技术)制备和鉴别本文描述的抗原的免疫原性部分。这样的技术包括就免疫原性特性筛选天然抗原的多肽部分。一般可以将本文所描述的代表性增殖和细胞因子产生测定用于这些筛选。多肽的免疫原性部分是这样的部分,其在这样的代表性测定中产生实质上类似由全长抗原产生的免疫反应(例如,增殖、干扰素- γ 产生和/或白介素-12 产生)。换句话说,在本文描述的模式增殖测定中,抗原的免疫原性部分产生至少约 20%,优选地约 100 % 的由全长抗原所诱导的增殖。在本文描述的模式测定中,免疫原性部分也可以或者可替性地刺激至少约 20%,优选地约 100 % 的由全长抗原所诱导的干扰素- γ 和/或白介素-12 产生。

结核分枝杆菌抗原的部分和其它变体可以用合成或者重组方法产生。利用本领域普通技术人员已知的技术,可以产生具有少于约 100 个氨基酸,一般少于约 50 个氨基酸的合成多肽。例如,这些多肽可以用任何通过商业途径可获得的固相技术合成,如 Merrifield 固相合成法,其中氨基酸依次添加到增长的氨基酸链上。参见 Merrifield, 美国化学会杂志, 8: 2149-2146, 1963。用于多肽的自动合成的设备是可通过商业途径从供应商(如应用生物系统公司, Foster City, CA)获得的,并且可以按照制造厂商的说明操作。一般可以用标准的诱变技术(如寡核苷酸定点特异性诱变)制备天然抗原的变体。也可以用标准的技术除去 DNA 序列的片段,以便可以制备截短的多肽。

采用本领域技术人员熟知的各种技术,可以容易地从编码多肽的 DNA 序列制备包含天然抗原部分和/或变体的重组多肽。例如,将重组蛋白质分泌到培养基中合适的宿主/载体系统的上清液可以首先采用市售的滤器浓缩。在浓缩之后,可以将浓缩液用于合适的纯化基质(如亲和性基质或离子交换树脂)上。最后,可以使用一个或多个反相 HPLC 步骤,以进一步纯化重组蛋白质。

本领域普通技术人员已知的各种表达载体的任何一种都可以用于表达

本文所描述的重组多肽。表达可以在任何合适的宿主细胞中进行，所述的宿主细胞已用含有编码重组多肽的 DNA 分子的表达载体转化或转染过。合适的宿主细胞包括原核生物，酵母和高级真核细胞。优选地，使用的宿主细胞是大肠杆菌，酵母或哺乳动物细胞系，如 COS 或 CHO。以这一方式表达的 DNA 序列可以编码天然存在的抗原，天然存在的抗原的部分，或者其它变体。

一般来说，不论采用哪一种制备方法，本文所公开的多肽实质上以纯化的形式得以制备。优选地，所述多肽的纯度为至少大约 80%，更优选地至少大约 90%，最优选地至少大约 99%。在以下详细描述某些优选的实施方案中，将基本上纯化的多肽掺入到药物组合物中或者疫苗中，以用于一种或多种本文公开的方法中。

在某些特定的实施方案中，主题发明公开了一些多肽，这些多肽包含可溶性结核分枝杆菌抗原或仅在保守取代和/或修饰上不同的其变体的至少一种免疫原性部分，所说的抗原具有一种以下的 N 端序列：

- (a) Asp-Pro-Val-Asp-Ala-Val-Ile-Asn-Thr-Thr-Cys-Asn-Tyr-Gly-Gln-Val-Val-Ala-Ala-Leu; (SEQ ID NO. 120)
- (b) Ala-Val-Glu-Ser-Gly-Met-Leu-Ala-Leu-Gly-Thr-Pro-Ala-Pro-Ser (SEQ ID NO. 121)
- (c) Ala-Ala-Met-Lys-Pro-Arg-Thr-Gly-Asp-Gly-Pro-Leu-Glu-Ala-Ala-Lys-Glu-Gly-Arg; (SEQ ID NO. 122)
- (d) Tyr-Tyr-Trp-Cys-Pro-Gly-Gln-Pro-Phe-Asp-Pro-Ala-Trp-Gly-Pro; (SEQ ID NO. 123)
- (e) Asp-Ile-Gly-Ser-Glu-Ser-Thr-Glu-Asp-Gln-Gln-Xaa-Ala-Val; (SEQ ID NO. 124)
- (f) Ala-Glu-Glu-Ser-Ile-Ser-Thr-Xaa-Glu-Xaa-Ile-Val-Pro; (SEQ ID NO. 125)
- (g) Asp-Pro-Glu-Pro-Ala-Pro-Pro-Val-Pro-Thr-Thr-Ala-Ala-Ser-Pro-Pro-Ser; (SEQ ID NO. 126)
- (h) Ala-Pro-Lys-Thr-Tyr-Xaa-Glu-Glu-Leu-Lys-Gly-Thr-Asp-Thr-Gly; (SEQ ID NO. 127)

- (i) Asp-Pro-Ala-Ser-Ala-Pro-Asp-Val-Pro-Thr-Ala-Ala-Gln-Leu-Thr-Ser-Leu-Leu-Asn-Ser-Leu-Ala-Asp-Pro-Asn-Val-Ser-Phe-Ala-Asn; (SEQ ID NO. 128) 和
- (j) Xaa-Asp-Ser-Glu-Lys-Ser-Ala-Thr-Ile-Lys-Val-Thr-Asp-Ala-Ser; (SEQ ID NO. 134)
- (k) Ala-Gly-Asp-Thr-Xaa-Ile-Tyr-Ile-Val-Gly-Asn-Leu-Thr-Ala-Asp; (SEQ ID NO. 135) 或
- (l) Ala-Pro-Glu-Ser-Gly-Ala-Gly-Leu-Gly-Gly-Thr-Val-Gln-Ala-Gly; (SEQ ID NO. 136)

其中 Xaa 可以是任何氨基酸, 优选地是半胱氨酸残基。编码以上标记有 (g) 的抗原的 DNA 序列在 SEQ ID NO. 52 中给出, 由 SEQ ID NO. 52 编码的多肽在 SEQ ID NO. 53 中给出。编码以上标记有 (a) 的抗原的 DNA 序列在 SEQ ID NO. 101 中给出, 其推定的氨基酸序列在 SEQ ID NO. 102 中给出。相应于以上抗原 (d) 的 DNA 序列在 SEQ ID NO. 24 中给出, 相应于以上抗原 (c) 的 DNA 序列在 SEQ ID NO. 25 中给出, 相应于以上抗原 (I) 的 DNA 序列在 SEQ ID NO. 99 中给出, 其推定的氨基酸序列在 SEQ ID NO. 100 中给出。

在另一个特定的实施方案中, 主题发明公开了一些多肽, 这些多肽包含具有一种以下的 N 端序列的结核分枝杆菌抗原, 或仅在保守取代和/或修饰上不同的该抗原的变体的至少一种免疫原性部分:

- (m) Xaa-Tyr-Ile-Ala-Tyr-Xaa-Thr-Thr-Ala-Gly-Ile-Val-Pro-Gly-Lys-Ile-Asn-Val-His-Leu-Val; (SEQ ID NO. 137) 或
- (n) Asp-Pro-Pro-Asp-Pro-His-Gln-Xaa-Asp-Met-Thr-Lys-Gly-Tyr-Tyr-Pro-Gly-Gly-Arg-Arg-Xaa-Phe; (SEQ ID NO. 129),

其中 Xaa 可以是任何氨基酸, 优选地是半胱氨酸残基。

在其它特定的实施方案中, 主题发明公开了一些多肽, 这些多肽包含可溶性的结核分枝杆菌抗原 (或这种抗原的变体) 的至少一种免疫原性部分, 所述抗原 (或其变体) 包含由以下序列编码的一种或多种氨基酸序列:

- (a) SEQ ID NO. 1, 2, 4-10, 13-25 和 52 的 DNA 序列, (b) 这些 DNA 序列的补体, 或 (c) 实质上同源 (a) 或 (b) 中的序列的 DNA 序列。

在其他特定的实施方案中, 主题发明公开了一些多肽, 这些多肽包含

结核分枝杆菌抗原(或这种抗原的变体)的至少一种免疫原性部分, 所述抗原(或其变体)可以是也可以不是可溶性的, 其包含由以下序列编码的一种或多种氨基酸序列: (a) SEQ ID NO. 26-51 的 DNA 序列, (b) 这些 DNA 序列的补体, 或 (c) 实质上同源于(a)或(b)中的序列的 DNA 序列。

在以上讨论的特定的实施方案中, 结核分枝杆菌抗原包括由实质上同源于本文特别提出的一种或多种 DNA 序列的 DNA 序列编码的变体。本文使用的 实质上的同源性 指在中等严格条件下能够杂交的 DNA 序列。合适的中等严格条件包括在 5X SSC, 0.5%SDS, 1.0 mM EDTA (pH 8.0) 溶液中预洗涤; 在 50 °C-65 °C, 5X SSC 下杂交一夜, 或者在杂交物种同源的情况下在 45 °C, 5X SSC 下杂交; 接着在 65 °C 下洗涤两次, 每次以包含 0.1 %SDS 的 2X, 0.5X 和 0.2X SSC 洗涤 20 分钟。这样的杂交 DNA 序列也在本发明的范围内, 由于密码简并, 编码由杂交 DNA 序列编码的免疫原性多肽的核苷酸序列也是如此。

在一个相关的方面, 本发明提供了一些融合蛋白以及这些融合蛋白的变体, 所说的融合蛋白包含第一与第二发明多肽或者是本发明多肽与已知的结核分枝杆菌抗原的融合蛋白, 所述抗原如以上描述的 38 kD 抗原或 ESAT-6 (SEQ ID NO. 98 和 99)。本发明的融合蛋白也可以包含在所说的第一和第二多肽之间的接头肽。

利用已知的 DNA 重组技术将分离的编码第一和第二多肽的 DNA 序列装配到适当的表达载体中, 由此来构建编码本发明的融合蛋白的 DNA 序列。将具有或不具有肽接头的编码第一多肽的 DNA 序列的 3' 末端连接到编码第二多肽的 DNA 序列的 5' 末端, 以便这些序列的读框处于可以使两种 DNA 序列的 mRNA 翻译成保持第一和第二多肽两者的生物学活性的单一融合蛋白的状态。

肽接头序列可以用于通过足以保证各多肽折叠成其二级和四级结构的距离分离第一和第二多肽。采用本领域熟知的标准技术将这样一种肽接头序列掺入到融合蛋白中。可以基于下列因素选择合适的肽接头序列: (1) 它们采取柔性延伸构象的能力; (2) 它们不采取二级结构(其可以与第一和第二多肽上的功能性表位相互作用)的能力; 和 (3) 可以与多肽的功能性表位进行反应的疏水或带电残基的缺乏。优选的肽接头序列包括 Gly、Asn 和

Ser 残基。其它接近中性的氨基酸, 如 Thr 和 Ala 也可以用于接头序列。可以有利地用作接头序列的氨基酸序列包括在 Maratea 等, 基因, 40: 39-46, 1985; Murphy 等, 美国科学院学报, 83: 8258-8562, 1986; 美国专利 4, 935, 233 和美国专利 4, 751, 180 中公开的那些。所说的接头序列长度可以从 1 到约 50 个氨基酸。当第一和第二多肽具有可以用来分离功能域和阻止空间位阻的非必需 N 端氨基酸区时, 肽接头序列是不需要的。

将连接的 DNA 序列可操作地连接到合适的转录或翻译调节元件上。在 DNA 的表达中起作用的调节元件仅被定位在编码第一多肽的 DNA 序列的 5'。同样地, 结束翻译所需的终止密码子和转录终止信号仅存在于编码第二多肽的 DNA 序列的 3'。

另一方面, 本发明提供了利用一种或多种以上的多肽或融合蛋白(或者编码这些多肽的 DNA 分子)在患者中诱导针对结核病的保护性免疫的方法。如本文所使用的, "患者"指任何温血动物, 优选地是人。患者可以是有疾病痛苦的, 或者可以是无可检测的疾病和/或感染的。换句话说, 保护性免疫可被诱导来预防或治疗结核病。

在这一方面, 所说的多肽、融合蛋白或 DNA 分子一般地存在于药物组合物和/或疫苗内。药物组合物可以包含一种或多种多肽, 每种多肽可以含有一种或多种以上的序列(或其变体)以及生理学上可接受的载体。疫苗可以包含一种或多种以上的多肽和一种非特异性免疫反应增强剂, 如佐剂或者脂质体(多肽掺入其中)。这样的药物组合物和疫苗也可以含有其它结核分枝杆菌抗原(掺入组合的多肽中或者存在于独立的多肽内)。

此外, 疫苗可以含有编码一种或多种以上描述的多肽的 DNA 序列, 以便多肽在原位产生。在这样的疫苗中, 所说的 DNA 可以存在于任何本领域普通技术人员已知的各种递送系统内, 包括核酸表达系统, 细菌和病毒的表达系统。合适的核酸表达系统含有在患者中表达所必需的 DNA 序列(例如合适的启动子和终止信号)。细菌递送系统涉及施用在其细胞表面上表达多肽的免疫原性部分的细菌(如 *Bacillus-Calmette-Guerrin*)。在一个优选的实施方案中, 可以利用病毒表达系统(例如, 疫苗或其它痘病毒, 逆转录病毒, 或腺病毒)引入所说的 DNA, 这可以包括利用非致病的(缺损的)复制感受态病毒。将 DNA 掺入到这样的表达系统中的技术是本领域普通技术人员

已知的。DNA 也可以是"裸"的,如在 Ulmer 等, 科学 259: 1745-1749, 1993 中描述的和在 Cohen, 科学 259: 1691-1692, 1993 中所综述的。裸 DNA 的摄取可以通过把 DNA 涂布在可被生物降解的小珠(其被有效地运输至细胞中)上来增加。

在相关的方面, 以上描述的 DNA 疫苗可以与本发明的多肽和已知的结核分枝杆菌抗原(如以上描述的 38 kD 抗原)同时施用或依次施用。例如, 编码本发明的多肽的 DNA (以上描述的"裸"的或送递系统中的)的施用可以在施用抗原之后进行, 以便提高疫苗的保护性免疫作用。

施用的路线和频率以及剂量将随不同的个体变化, 并且可以相似于当前所使用的利用 BCG 的免疫的那些。一般来说, 所说的药物组合物和疫苗可以由注射(例如, 皮内, 肌内, 静脉内或皮下), 鼻内(例如, 吸入)或口头施用。1 到 3 剂可以施用 1-36 周期限。优选地, 施用 3 剂, 间隔为 3-4 个月, 此后, 可以周期性地给予加强免疫接种。可替的方案对个体患者可能是合适的。合适的剂量是, 当如上所述施用时, 能够在免疫过的患者中产生免疫反应, 在至少 1-2 年内保护患者不受结核分枝杆菌感染的多肽或 DNA 的量。一般来说, 在药剂中存在的多肽的量(或者由药剂中 DNA 原位产生的量)在每 kg 宿主约 1pg 到约 100 mg 的范围内, 典型地在约 10pg 到约 1mg 的范围内, 优选地在约 100 pg 至约 1 μ g 的范围内。合适的剂量将随患者的情况而变化, 但典型地在约 0.1mL 到约 5mL 之间。

虽然本领域普通技术人员已知的任何合适的载体都可以用于本发明的药物组合物中, 但载体的类型将随施用方式变化。对于肠胃外施用(如皮下注射), 载体优选地包括水, 盐水, 醇, 脂肪, 蜡或缓冲液。对于口服施用, 可以使用任何上述载体或固相载体, 如甘露糖醇, 乳糖, 淀粉, 硬脂酸镁, 糖精钠, 滑石, 纤维素, 葡萄糖, 蔗糖, 和碳酸镁。可被生物降解的微球体(例如 polylactic galactide)也可以用作本发明的药物组合物的载体。合适的可被生物降解的微球体在例如美国专利 4,897,268 和 5,075,109 中公开。

各种佐剂中的任何一种都可以用于本发明的疫苗中, 以非特异性地增强免疫反应。大多数佐剂含有用来保护抗原使之免受迅速分解代谢的物质(如氢氧化铝或矿物油)和免疫反应的非特异性刺激物(如脂质 A, *Bordetella*

pertussis 或结核分枝杆菌)。合适的佐剂是市售的, 例如, 弗氏不完全佐剂和弗氏完全佐剂(Difco 实验室)和 Merck 佐剂 65(Merck and Company, Inc., Rahway, NJ.)。其它合适的佐剂包括明矾, 可被生物降解的微球体, 单磷酸脂 A 和 quil A。

在另一方面, 本发明提供了利用一种或多种以上描述的多肽采用皮试诊断结核病的方法。如本文所使用的, "皮试"是直接在患者上进行的任何试验, 其中, 在真皮内注射一种或多种以上描述的多肽后, 测定迟发型超敏反应(DTH)(例如膨胀, 变红或皮炎)。这样的注射可以用任何足以使多肽与患者皮肤细胞接触的合适装置完成, 如结核菌素注射器或 1mL 注射器。优选地, 在注射之后至少 48 小时, 更优选地在 48-72 小时后测定反应。

DTH 反应是细胞介导的免疫反应, 其在以前已接触过试验抗原(即, 所使用多肽或其变体的免疫原性部分)的患者中较强。采用标尺可视觉测量该反应。一般来说, 直径大于约 0.5 cm, 优选地大于约 1.0 cm 的反应是阳性反应, 指示结核病感染, 其可以也可以不表明活动性疾病。

就用于皮试而言, 本发明的多肽优选地配制成含有多肽和以上描述的生理学上可接受的载体的药物组合物。这样的组合物典型地含有一种或多种以上的多肽, 其含量在每 0.1mL 体积约 1 μ g 到约 100 μ g, 优选地约 10 μ g 到约 50 μ g 的范围内。优选地, 用于这样的药物组合物中的载体是具有适当的防腐剂(如苯酚和/或吐温 80™)的盐水溶液。

在一个优选的实施方案中, 用于皮试的多肽是具有足够大小的, 以便其在反应期间保持在注射部位。一般来说, 至少 9 个氨基酸的多肽在长度上是足够的。优选地是所述多肽在几小时内被巨噬细胞裂解, 以使得可以传递至 T-细胞。这样的多肽可以含有一种或多种上述序列和/或其它免疫原性或非免疫原性序列的重复序列。

以说明性的方式但不以限制性的方式给出下列实施例。

实施例

实施例 1

来源于结核分枝杆菌培养物滤液的多肽的纯化和特征确定

这一例子说明从培养物滤液制备结核分枝杆菌可溶性多肽的方法。除

非有其它方式注明, 下列例子的所有百分比都是重量/体积百分比。

于 37 °C 在无菌 GAS 培养基中培养结核分枝杆菌 (H37Ra , ATCC No. 25177 或 H37Rv , ATCC No. 25618) 14 天。然后经 0.45 μ 滤器将培养基真空过滤(留下大批细胞)到无菌的 2.5L 瓶中。接着经 0.2 μ 滤器将培养基过滤到无菌的 4L 瓶中。向培养物滤液中加入 NaN₃, 使其浓度达 0.04%。然后将瓶置于 4 °C 的冷室中。

通过将滤液置于已高压灭菌的 12L 贮器中, 并将滤液供入 400ml Amicon 搅拌池中浓缩培养物滤液, 该搅拌池已以乙醇冲洗过, 并且被包含 10,000 kDa MWC0 膜。使用氮气使压力保持在 60 psi。这一过程使 12 L 体积减少到约 50ml。

然后, 采用 8,000 kDa MWC0 纤维素酯膜将培养物滤液对 0.1 %碳酸氢铵透析, 两次更换碳酸氢铵溶液。接着由通过商业途径可获得的 BCA 测定法(Pierce, Rockford, IL)测定蛋白质浓度。

然后将透析培养物滤液进行冻干, 把多肽重悬于蒸馏水中。然后, 将多肽对 0.01 mM 1,3 双[三(羟甲基)-甲氧基]丙烷, pH 7.5 (Bis-Tris 丙烷缓冲液) (阴离子交换层析的起始条件) 透析。利用在 POROS 146 II Q/M 阴离子交换柱 4.6 mm $\times$ 100 mm (Perseptive BioSystems, Framingham, MA) 上的凝胶预熔融 (profusion) 层析完成分级分离, 所述交换柱已在 0.01 mM Bis-Tris 丙烷缓冲液 (pH7.5) 中平衡过。用在上述缓冲液系统中的 0-0.5 M NaCl 梯度洗脱多肽。在 220 nm 波长下监测柱洗脱液。

将从离子交换柱洗脱的多肽收集物对蒸馏水透析并冻干。将所形成的物质溶解到在水中的 0.1%三氟乙酸(TFA) (pH 1.9) 中, 并且在 Delta-Pak C18 柱 (Waters, Milford, MA, 300 埃孔径大小, 5 微米颗粒大小 (3.9 X150 mm)) 上纯化该多肽。用从 0 到 60%稀释缓冲液(在乙腈中的 0.1 %TFA) 线性梯度液从柱中洗脱多肽。流速是 0.75ml/分钟, 在 214nm 监测 HPLC 洗脱液。收集包含洗脱的多肽的组分, 使单个样品纯度最大。获得约 200 个纯化的多肽。

然后, 就在 PBMC 制剂中诱导 T-细胞增殖的能力筛选纯化的多肽。将 PBMC (来源于称为 PPD 皮试阳性的供体, 并且其 T-细胞表现出应答 PPD 和粗的 MTB 可溶性蛋白质的增殖) 在包含 RPMI 1640 (补充有 10%收集的人血清和

50 $\mu\text{g/mL}$ 庆大霉素)的培养基中培养。双份以 0.5 至 10 $\mu\text{g/mL}$ 的浓度添加纯化的多肽。96-孔园底平板中以 200 μl 体积培养 6 天后,从各孔除去 50 μl 培养基,以测定 IFN- γ 水平,如以下所述。接着用 1 μCi /孔含氘胸苷脉冲平板另外的 18 小时,收获,并用气体闪烁计数器测定氘摄取。在两个重复中产生的增殖高于在单独的培养基中培养的细胞上观察到的增殖的 3 倍的组分被认为是阳性的。

用酶联免疫吸附测定(ELISA)测定 IFN- γ 。在室温下用在 PBS 中的针对人 IFN- γ (PharMingen, San Diego, CA)的小鼠单克隆抗体涂布 ELISA 平板 4 小时。然后在室温下用包含 5% (WN) 脱脂干奶的 PBS 阻断各孔。接着用 PBS/0.2% TWEEN-20 洗涤平板 6 次,将在 ELISA 平板上的以培养基 1:2 稀释的样品在室温下过夜温育。再次洗涤平板,向各孔中添加以 PBS/10% 正常山羊血清 1:3000 稀释的多克隆兔抗-人 IFN- γ 血清。然后在室温下温育平板两小时,洗涤,加入以 PBS/5% 脱脂干奶 1:2000 稀释的辣根过氧化物酶-偶联的抗-兔 IgG (西格玛化学公司,圣路易斯, MO)。在室温下进一步温育 2 小时后,洗涤平板,并加入 TMB 底物。20 分钟后用 1N 硫酸终止反应。用 570 nm 为参照波长,在 450 nm 测定光密度。在两个重复中导致给出的 OD 高于在单独的培养基中培养的平均 OD 加 3 个标准偏差的组分被认为是阳性的。

为了测序,将多肽单个地干燥到 Biobrene (Perkin Elmer/Applied BioSystems Division, Foster City, CA)处理过的玻璃纤维过滤器上。将具有多肽的过滤器装到 Perkin Elmer/Applied BioSystems Division Procise 492 蛋白质测序仪上。从氨基端测序多肽,并且用传统的 Edman 化学法。通过把 PTH 氨基酸衍生物的保留时间与适当的 PTH 衍生物标准比较,确定各多肽的氨基酸序列。

利用以上描述的方法,分离到具有下列 N 端序列的抗原:

- (a) Asp-Pro-Val-Asp-Ala-Val-Ile-Asn-Thr-Thr-Xaa-Asn-Tyr-Gly-Gln-Val-Val-Ala-Ala-Leu; (SEQ ID NO. 54)
- (b) Ala-Val-Glu-Ser-Gly-Met-Leu-Ala-Leu-Gly-Thr-Pro-Ala-Pro-Ser; (SEQ ID NO. 55)
- (c) Ala-Ala-Met-Lys-Pro-Arg-Thr-Gly-Asp-Gly-Pro-Leu-

- Gln-Ala-Ala-Lys-Glu-Gly-Arg; (SEQ ID NO. 56)
- (d) Tyr-Tyr-Trp-Cys-Pro-Gly-Gln-Pro-Phe-Asp-Pro-Ala-Trp-Gly-Pro; (SEQ ID NO. 57)
- (e) Asp-Ile-Gly-Ser-Glu-Ser-Thr-Glu-Asp-Gln-Gln-Xaa-Ala-Val; (SEQ ID NO. 58)
- (f) Ala-Glu-Glu-Ser-Ile-Ser-Thr-Xaa-Glu-Xaa-Ile-Val-Pro; (SEQ ID No. 59)
- (g) Asp-Pro-Glu-Pro-Ala-Pro-Pro-Val-Pro-Thr-Ala-Ala-Ala-Ala-Pro-Pro-Ala; (SEQ ID NO. 60) 和
- (h) Ala-Pro-Lys-Thr-Tyr-Xaa-Glu-Glu-Leu-Lys-Gly-Thr-Asp-Thr-Gly; (SEQ ID NO. 61)

其中 Xaa 可以是任何氨基酸。

除以上所描述的方法之外, 通过使用微内径柱 HPLC 纯化步骤分离到另外的抗原。具体地说, 在 Aquapore C18 柱 (Perkin Elmer/Applied Biosystems Division, Foster City, CA) 上纯化包含以上描述的层析纯化步骤的抗原混合物的 20 μ l 组分, 所说的柱具有 7 微米孔径大小, 柱规格为 1 mmX100 mm, 在 Perkin Elmer/Applied Biosystems Division 172 HPLC 型中。以 80 μ l/分钟的流速, 在水 (0.05 %TFA) 中的乙腈 (含 0.05 %TFA) 的 1%/分钟的线性梯度液从柱上洗脱各组分。在 250 nm 下监测洗脱液。原组分被分离成 4 个主要的峰加其他小的组分, 并且获得显示出具有 12.054 Kd 分子量 (由质谱测得) 和具有以下 N 端序列的多肽:

- (i) Asp-Pro-Ala-Ser-Ala-Pro-Asp-Val-Pro-Thr-Ala-Ala-Gln-Gln-Thr-Ser-Leu-Leu-Asn-Asn-Leu-Ala-Asp-Pro-Asp-Val-Ser-Phe-Ala-Asp (SEQ ID NO. 62)。

采用以上所述的测定法, 这一多肽显示出在 PBMC 制剂中诱导增殖和 IFN- γ 产生。

按照以下所述从结核分枝杆菌培养物滤液分离另外的可溶性抗原。结核分枝杆菌培养物滤液按照以上描述的方法制备。在 pH 5.5 下对 Bis-Tris 丙烷缓冲液透析后, 用在 Poros QE 柱 4.6 X100 mm (Perseptive Biosystems) 上的阴离子交换层析完成分级分离, 所述柱在 Bis-Tris 丙烷缓冲液 (pH 5.5)

中平衡过。以 10ml/分钟的流速，用在上述缓冲系统中的线性 0-1.5M NaCl 梯度液洗脱多肽。在 214nm 下检测柱洗脱液。

收集从离子交换柱洗脱的组分，并采用 Poros R2 柱 4.6X100 mm (Perseptive Biosystems) 进行反相层析。以 5ml/分钟的流速，用 0-100 % 乙腈(0.1%TFA) 的线性梯度液从柱上洗脱多肽，在 214 nm 监测洗脱液。

将包含洗脱的多肽的组分冷干，并重悬于 80 μ l 0.1 %TFA 水溶液中，并再在 Vydac C4 柱 4.6 X150 mm(Western Analytical, Temecula, CA) 上，以 2ml/分钟的流速，用 0-100 %乙腈(0.1 %TFA) 线性梯度液进行反相层析。在 214 nm 监测洗脱液。

具有生物活性的组分被分离成一个主要的峰加其它小组分。这一峰的 PVDF 膜上的 Western 印迹揭示分子量为 14Kd, 20Kd 和 26Kd 的三个主要带。确定了这些多肽分别具有下列 N 端序列:

- (j) Xaa-Asp-Ser-Glu-Lys-Ser-Ala-Thr-Ile-Lys-Val-Thr-Asp-Ala-Ser; (SEQ ID NO. 134)
- (k) Ala-Gly-Asp-Thr-Xaa-Ile-Tyr-Ile-Val-Gly-Asn-Leu-Thr-Ala-Asp; (SEQ ID NO. 135) 和
- (l) Ala-Pro-Glu-Ser-Gly-Ala-Gly-Leu-Gly-Gly-Thr-Val-Gln-Ala-Gly; (SEQ ID NO. 136), 其中 Xaa 可以是任何氨基酸。

采用以上所述的测定法，这些多肽显示出在 PBMC 制剂中诱导增殖和 IFN- γ 产生。图 1A 和 B 分别显示了使用第一和第二供体的 PMBC 制剂进行的这种测定的结果。

通过采用 32 P 末端标记的简并寡核苷酸(相应于 N 端序列并含有结核分枝杆菌密码子偏倚) 筛选结核分枝杆菌基因组文库获得编码以上指定为(a), (c), (d)和(g)的抗原的 DNA 序列。采用相应于以上抗原(a)的探针进行的筛选鉴别具有 SEQ ID NO. 101 所示的序列的克隆。由 SEQ ID NO. 101 编码的多肽在 SEQ ID NO. 102 中给出。采用相应于以上抗原(g)的探针进行的筛选鉴别具有 SEQ ID NO. 52 所示的序列的克隆。由 SEQ ID NO. 52 编码的多肽在 SEQ ID NO. 53 中给出。采用相应于以上抗原(d)的探针进行的筛选鉴别具有 SEQ ID NO. 24 所示的序列的克隆。采用相应于以上抗原(c)的探针进行的筛选鉴别具有 SEQ ID NO. 25 所示的序列的克隆。

采用 DNA STAR 系统, 将以上氨基酸序列与基因库中的已知氨基酸序列比较。所检索的数据库含有大约 173,000 种蛋白质, 并且是 Swiss, PIR 数据库以及翻译的蛋白质序列 (版本 87) 的组合。对抗原 (a)-(h) 和 (l), 没有检测到与所说的氨基酸序列的明显的同源性。

发现抗原 (i) 的氨基酸序列同源与麻风分枝杆菌的序列。利用从 GENBANK 获得的序列从基因组 DNA 扩增全长麻风分枝杆菌序列。然后, 将这一序列用于筛选以下实施例 2 中描述的结核分枝杆菌文库, 获得全长拷贝的结核分枝杆菌的同系物 (SEQ ID NO. 99)。

发现抗原 (j) 的氨基酸序列同源与从 DNA 序列翻译的已知结核分枝杆菌蛋白质。就发明者所知, 这一蛋白质以前还没有显示出具有 T-细胞刺激性。发现抗原 (k) 的氨基酸序列与麻风分枝杆菌的序列相关。

在以上描述的增殖与 IFN- γ 测定中, 利用三个 PPD 阳性供体, 以上所提供的代表性抗原的结果在表 1 中给出:

表 1
PBMC 增殖和 IFN- γ 测定的结果

序列	增殖	IFN- γ
(a)	+	-
(c)	+++	+++
(d)	++	++
(g)	+++	+++
(h)	+++	+++

在表 1 中, 给出 2 和 4 之间的刺激指数 (SI) 的反应 (与在单独的培养基培养的细胞比较) 记录为+, 在 1 μ g 或更低的浓度下的 4-8 或 2-4 的 SI 记录为++, 大于 8 的 SI 记录为+++。发现序列 (i) 的抗原在增殖和 IFN- γ 测定两者中, 对一种供体具有高的 SI (+++), 对两种其它供体具有较低的 SI (++和+)。这些结果表明这些抗原有能力诱导增殖和/或干扰素- γ 产生。

实施例 2

使用病人血清分离结核分枝杆菌抗原

这一例子说明通过用结核分枝杆菌感染个体的血清筛选从结核分枝杆菌溶解产物分离抗原的方法。

将干燥的结核分枝杆菌 H37Ra (Difco 实验室) 添加至 2% NP40 溶液中, 此外, 匀浆和超声处理三次。在 13,000 rpm 下在微量离心管中离心所形成的悬浮液, 将上清液通过 0.2 微米注射滤器。将滤液结合到 Macro Prep DEAE 小珠 (BioRad, Hercules, CA) 上。用 20mM Tris (pH 7.5) 充分洗涤小珠, 结合的蛋白质以 1M NaCl 洗脱。将 NaCl 洗脱液对 10mM Tris (pH 7.5) 透析一夜。在室温下用 0.05 mg/ml 的 DNase 和 RNase 处理透析溶液 30 分钟, 然后于室温在 pH 4.5 下用 0.5 U/mg α -D-甘露糖苷酶处理。在返回到 pH 7.5 后, 在 Bio Scale-Q-20 柱 (BioRad) 上经 FPLC 分级分离该物质。将组分合并到九个池中, 在 Centriprep 10 (Amicon, Beverley, MA) 中浓缩, 并且采用结核分枝杆菌感染病人的血清 (其与本发明的其它抗原不发生免疫反应) 就血清学活性经 Western 印迹筛选。

将反应性最强的组分在 SDS-PAGE 上进行分析, 并转移到 PVDF 上。切下约 85 Kd 的带, 产生以下序列:

(m) Xaa-Tyr-Ile-Ala-Tyr-Xaa-Thr-Thr-Ala-Gly-Ile-Val-Pro-Gly-Lys-Ile-Asn-Val-His-Leu-Val; (SEQ ID NO. 137), 其中 Xaa 可以是任何氨基酸。

这些序列与以上描述的基因库中的那些序列的比较揭示出与已知的序列没有明显的同源性。

实施例 3

制备编码结核分枝杆菌抗原的 DNA 序列

这一例子说明通过用从结核分枝杆菌感染病人获得的血清或者用抗结核分枝杆菌抗原产生的抗血清筛选结核分枝杆菌表达文库, 制备编码结核分枝杆菌抗原的 DNA 序列的方法。

A. 用兔抗血清制备结核分枝杆菌可溶性抗原

从结核分枝杆菌菌株 H37Ra 分离基因组 DNA。随机剪切该 DNA, 并用用于 Lambda ZAP 表达系统 (Stratagene, La Jolla, CA) 构建表达文库。

通过用结核分枝杆菌培养物的浓缩上清液免疫兔产生抗结核分枝杆菌菌株 H37Ra, H37Rv 和 Erdman 的分泌蛋白质的兔抗血清。具体地说, 首先用 200 μ g 在含有 100 μ g 胞壁酰二肽的 2ml 总体积中的蛋白质抗原 (Calbiochem, La Jolla, CA) 和 1ml 弗氏不完全佐剂皮下免疫兔。四周后, 用在弗氏不完全佐剂中的 100 μ g 抗原皮下加强免疫兔。最后, 在四周后用 50 μ g 蛋白质抗原静脉内免疫兔。如 Sambrook 等, 分子克隆: 实验室手册, 冷泉港实验室, 冷泉港, NY, 1989 中的描述将抗血清用于筛选表达文库。纯化表达免疫反应性抗原的噬菌体噬斑。噬斑的噬粒得到救援, 结核分枝杆菌克隆的核苷酸序列被推定。

纯化了 32 个克隆。在这些克隆中, 25 个代表在结核分枝杆菌中以前没有鉴别过的序列。表达重组抗原, 纯化的抗原用于实施例 1 描述的免疫学分析中。如 Skeiky 等, 实验医学杂志, 181: 1527-1537, 1995 中所述用 IPTG 诱导蛋白质, 并经凝胶洗脱纯化。在这一筛选中鉴别的 DNA 分子的代表性部分序列在 SEQ ID NO. 1-25 中给出。相应的预言的氨基酸序列在 SEQ ID NO. 63-87 中给出。

基于采用以上所述的数据库将这些序列与基因库中的已知序列比较, 发现下文中称为 TbRA2A、TbRA16、TbRA18 和 TbRA29 (SEQ ID NO. 76、68、70、75) 的克隆显示出与以前在麻风分枝杆菌中而不是在结核分枝杆菌中鉴别的序列的某些同源性。TbRA11、TbRA26、TbRA28 和 TbDPEP (SEQ ID NO. 65、73、74、53) 以前在结核分枝杆菌中已鉴定过。对 TbRA1、TbRA3、TbRA4、TbRA9、TbRA10、TbRA13、TbRA17、TbRA19、TbRA29、TbRA32、TbRA36 和重叠克隆 TbRA35 和 TbRA12 (分别为 SEQ ID NO. 63、77、81、82、64、67、69、71、75、78、80、79、66) 没有发现明显的同源性。克隆 TbRa24 与克隆 TbRa29 重叠。

采用来源于几种不同的结核分枝杆菌免疫患者的 T-细胞制剂, 在代表性的重组抗原上进行的 PBMC 增殖和干扰素- γ 测定的结果分别示于表 2 和 3 中。

表 2
PBMC 增殖对代表性可溶性抗原的结果

抗原	患者												
	1	2	3	4	5	6	7	8	9	10	11	12	13
TbRa1	-	-	±	++	-	-	±	±	-	-	+	±	-
TbRa3	-	±	++	-	±	-	-	++	±	-	-	-	-
TbRa9	-	-	nt	nt	++	++	nt	nt	nt	nt	nt	nt	nt
TbRa10	-	-	±	±	±	+	nt	±	-	+	±	±	-
TbRa11	±	±	+	++	++	+	nt	-	++	++	++	±	nt
TbRa12	-	-	+	+	±	++	+	±	±	-	+	-	-
TbRa16	nt	nt	nt	nt	-	+	nt	nt	nt	nt	nt	nt	nt
TbRa24	nt	nt	nt	nt	-	-	nt	nt	nt	nt	nt	nt	nt
TbRa26	-	+	nt	nt	-	-	nt	nt	nt	nt	nt	nt	nt
TbRa29	nt	nt	nt	nt	-	-	nt	nt	nt	nt	nt	nt	nt
TbRa35	++	nt	++	++	++	++	nt	++	++	++	++	++	nt
TbRaB	nt	nt	nt	nt	-	-	nt	nt	nt	nt	nt	nt	nt
TbRaC	nt	nt	nt	nt	-	-	nt	nt	nt	nt	nt	nt	nt
TbRaD	nt	nt	nt	nt	-	-	nt	nt	nt	nt	nt	nt	nt
AAMK	-	-	±	-	-	-	nt	-	-	-	nt	±	nt
YY	-	-	-	-	-	-	nt	-	-	-	nt	+	nt
DPEP	-	+	-	++	-	-	nt	++	±	+	±	±	nt
Control	-	-	-	-	-	-	-	-	-	-	-	-	-

nt = 未试验

表 3
PBMC 的干扰素- γ 产生对代表性可溶性抗原的结果

抗原	患者												
	1	2	3	4	5	6	7	8	9	10	11	12	13
TbRa1	+	++		+++	+	-		±	-	-	+	±	-
TbRa3	-	±	++	-	±	-	-	++	±	-	-	-	-
TbRa9	++	+	nt	nt	++	-	nt	nt	nt	nt	nt	nt	nt
TbRa10	+	+	±	±	±	+	nt	±	-	+	±	±	-
TbRa11		±	+	++	++	+	nt	-	++	++	++	±	nt
TbRa12	-	-	+	+	±	+++	+	±	±	-	+	-	-
TbRa16	nt	nt	nt	nt	+	+	nt	nt	nt	nt	nt	nt	nt
TbRa24	nt	nt	nt	nt	+	-	nt	nt	nt	nt	nt	nt	nt
TbRa26	++	++	nt	nt	+	+	nt	nt	nt	nt	nt	nt	nt
TbRa29	nt	nt	nt	nt	+	-	nt	nt	nt	nt	nt	nt	nt
TbRa35	++	nt	++	++	+++	+++	nt	++	++	+++	+++	++	nt
TbRaB	nt	nt	nt	nt	++	+	nt	nt	nt	nt	nt	nt	nt
TbRaC	nt	nt	nt	nt	+	+	nt	nt	nt	nt	nt	nt	nt
TbRaD	nt	nt	nt	nt	+	+	nt	nt	nt	nt	nt	nt	nt
AAMK	-	-	±	-	-	-	nt	-	-	-	nt	±	nt
YY	-	-	-	-	-	-	nt	-	-	-	nt	+	nt
DPEP	+	+	+	+++	+	-	nt	+++	±	+	±	±	nt
对照	-	-	-	-	-	-	-	-	-	-	-	-	-

在表 2 和 3 中, 给出 1.2 和 2 (与在单独的培养基中培养的细胞比较) 之间的刺激指数(SI) 的反应记录为 $\pm$, 2-4 的 SI 记录为+, 在 1 μ g 或更低的浓度下的 4-8 或 2-4 的 SI 记录为++, 大于 8 的 SI 记录为+++。此外, 对两种以上的抗原, 浓度对增殖和干扰素- γ 产生的影响示于附图中。对增殖和干扰素- γ 产生两者而言, TbRa3 记录为++, TbRa9 记录为+。

这些结果说明, 这些可溶性抗原可以在来源于结核分枝杆菌免疫个体的 T-细胞中诱导增殖和/或干扰素- γ 产生。

B. 使用病人血清鉴别编码结核分枝杆菌抗原的 DNA 序列

采用从患活动性结核病的患者获得的血清库筛选以上描述的基因组 DNA 文库和另外的 H37Rv 文库。为了制备 H37Rv 文库, 分离结核分枝杆菌菌株 H37Rv 基因组的 DNA, 进行部分 Sau3A 消化, 并用于采用 Lambda Zap 表达系统(Stratagene, La Jolla, Ca)构建表达文库。将三种不同库的血清(各含有从患有活动性肺部或胸膜疾病的个体获得的血清)用于表达筛选。有关在 ELISA 和免疫印迹方式两者中与 H37Ra 溶解产物的相对反应性, 这些库被指定为 TbL、TbM 和 TbH(即, TbL =低反应性, TbM =中等反应性和 TbH =高反应性)。也使用了来自活动性肺结核病的七个患者血清的四个库。所有血清缺乏与重组 38 kD 结核分枝杆菌 H37Ra 磷酸盐-结合蛋白的增加了的反应性。

所有库用大肠杆菌溶解产物预吸附, 并用于如 Sambrook 等, 分子克隆: 实验室手册, 冷泉港实验室, 冷泉港, NY, 1989 中所述筛选 H37Ra 和 H37Rv 表达文库。纯化表达免疫反应性抗原的噬菌体噬斑。噬斑的噬粒得到救援, 结核分枝杆菌克隆的核苷酸序列被推定。

纯化了 32 个克隆。在这些克隆中, 31 个代表在人结核分枝杆菌中以前没有鉴别过的序列。所鉴别的 DNA 分子的代表性序列在 SEQ ID NO. 26-51 和 105 中给出。在这些克隆中, TbH-8 和 TbH-8-2 (SEQ ID NO. 105) 是相同克隆的非邻接 DNA 序列, TbH-4 (SEQ ID NO. 43) 和 TbH-4-FWD (SEQ ID NO. 44) 是相同克隆的非邻接序列。此后鉴别为 Tb38-1、TbH-4、TbH-8、TbH-9、和 TbH-12 的抗原的氨基酸序列在 SEQ ID NO. 88-92 中显示。利用以上确定的数据库将这些序列与基因库中的已知序列的比较揭示出, 对 TbH-4、TbH-8、TbH-9 和 TbM-3 没有明显的同源性, 虽然对 TbH-9 发现了

弱的同源性。发现 TbH-12 同源于以前在副结核分枝杆菌 (Acc. S28515) 中鉴定的 34 kD 抗原蛋白质。发现 Tb38-1 位于以前在牛型分枝杆菌 (Acc. No. U34848) 和结核分枝杆菌中鉴别的抗原 ESAT-6 开放读框上游 34 个碱基对 (Sorensen 等, 感染免疫学, 63: 1710-1717, 1995)。

将来源于 Tb38-1 和 TbH-9 (两者都是从 H37Ra 文库分离的) 的探针用于鉴别 H37Rv 文库中的克隆。Tb38-1 杂交到 Tb38-1F2、Tb38-1F3、Tb38-1F5 和 Tb38-1F6 (SEQ ID NO. 112、113、116、118 和 119)。SEQ ID NO. 112 和 113 是来源于克隆 Tb38-1F2 的非邻接序列, 推定了 Tb38-1F2 中的两个开放读框; 一个相应于 Tb37FL (SEQ ID NO. 114), 第二个 (部分序列) 可以是 Tb38-1 的同系物, 并称为 Tb38-IN (SEQ ID NO. 115)。Tb38-1F3 的推定的氨基酸序列在 SEQ ID NO. 117 中给出。TbH-9 探针鉴别了 H37Rv 文库中的三个克隆: TDH-9-FL (SEQ ID NO. 106), 其可以是 TbH-9 (R37Ra) 的同系物; TbH-9-1 (SEQ ID NO. 108) 和 TbH-9-4 (SEQ ID NO. 110), 所有这些都是 TbH-9 的高度相关序列。这三个克隆的推定的氨基酸序列在 SEQ ID NO. 107、109 和 111 中给出。

在 Tb38-1, ESAT-6 和其它代表性重组抗原上进行的 T-细胞测定的结果分别示于以下表 4A、B 和 5 中。

表 4A
PBMC 增殖对代表性抗原的结果

抗原	供体										
	1	2	3	4	5	6	7	8	9	10	11
Tb38.1	+++	+	-	-	-	++	-	+	-	++	+++
ESAT-6	+++	+	+	+	-	+	-	+	+	++	+++
TbH-9	++	++	-	++	±	±	++	++	++	++	++

表 4B
PBMC 的干扰素- γ 产生对代表性抗原的结果

抗原	供体										
	1	2	3	4	5	6	7	8	9	10	11
Tb38.1	+++	+	-	+	+	+++	-	++	-	+++	+++
ESAT-6	+++	+	+	+	+-	+	-	+	+	+++	+++
TbH-9	++	++	-	+++	±	±	+++	+++	++	+++	++

表 5
T - 细胞反应对代表性抗原的总结

抗原	增殖			干扰素- γ			总共
	患者 4	患者 5	患者 6	患者 4	患者 5	患者 6	
TbH9	++	++	++	+++	++	++	13
TbM7	-	+	-	++	+	-	4
TbH5	-	+	+	++	++	++	8
TbL23	-	+	$\pm$	++	++	+	7.5
TbH4	-	++	$\pm$	++	++	$\pm$	7
- 对照	-	-	-	-	-	-	0

这些结果说明，本发明的结核分枝杆菌抗原和 ESAT-6 两者都可以在来源于结核分枝杆菌免疫个体的 T-细胞中诱导增殖和/或干扰素- γ 产生。就发明人所知，ESAT-6 以前没有显示出刺激人免疫反应。

采用实施例 4 所述的方法构建了覆盖抗原 Tb38-1 的氨基酸序列的一组六个重叠肽。这些肽序列(此后称为 pep1-6)分别在 SEQ ID NO. 93-98 中显示。采用这些肽进行的 T-细胞测定的结果示于表 6 和 7 中。这些结果证实，在 Tb38-1 内的存在和帮助定位 T-细胞表位能够在来源于结核分枝杆菌免疫个体的 T-细胞中诱导增殖和干扰素- γ 产生。

表 6
PBMC 增殖对 T_B38 - 1 肽的结果

肽	患者												
	1	2	3	4	5	6	7	8	9	10	11	12	13
pep1	-	-	-	-	±	-	-	-	-	±	-	-	+
pep2	±	-	-	-	±	-	-	-	±	±	-	-	+
pep3	-	-	-	-	-	-	-	-	±	-	-	-	±
pep4	++	-	-	-	-	-	+	-	±	±	-	-	+
pep5	++	±	-	-	-	-	+	-	±	-	-	-	+
pep6	-	++	-	-	-	-	±	-	±	+	-	-	+
对照	-	-	-	-	-	-	-	-	-	-	-	-	-

表 7
PBMC 的干扰素- γ 产生对 T_H38 - 1 肽的结果

肽	患者												
	1	2	3	4	5	6	7	8	9	10	11	12	13
pep1	+	-	-	-	±	-	-	-	-	±	-	-	+
pep2		-	-	-	±	-	-	-	±	±	-	-	+
pep3	-	-	-	-	-	-	-	-	±	-	-	-	±
pep4	++	-	-	-	-	-	+	-	±	±	-	-	+
pep5	++	±	-	-	-	-	+	-	±	-	-	-	+
pep6	+	++	-	-	-	-	±	-	±	+	-	-	+
对照	-	-	-	-	-	-	-	-	-	-	-	-	-

实施例 4

来源于结核菌素纯化蛋白质衍生物的多肽的纯化和特征确定

按照以下所述从结核菌素纯化蛋白质衍生物 (PPD) 分离结核分枝杆菌多肽。

按进行某些修改的出版的方法 (Seibert, F 等, 结核菌素纯化蛋白质衍生物。大量制备和分析标准。美国结核病评论 44: 9-25, 1941) 制备 PPD。

于 37 °C 下在摇瓶中用合成培养基培养结核分枝杆菌 Rv 菌株 6 周。然后将含有细菌生长物的瓶子用水蒸汽加热到 100 °C 3 小时。用 0.22 μ 滤器无菌过滤培养物, 采用 3 kD 切断膜浓缩 20 倍。用 50 % 硫酸铵溶液沉淀蛋白质一次, 用 25% 硫酸铵溶液沉淀 8 次。通过反相液相层析 (RP-HPLC) 分级分离所形成的蛋白质 (PPD), 所说的层析采用在 Biocad HPLC 系统 (Perseptive Biosystems, Framingham, MA) 中的 C18 柱 (7.8 X300 mm; Waters, Milford, MA)。用 0-100 % 线性梯度缓冲液 (在乙腈中的 0.1 % TFA) 从柱中洗脱组分。流速是 10ml/分钟, 在 214 nm 和 280 nm 下监测洗脱液。

收集六个组分, 干燥, 悬浮在 PBS 中, 并在结核分枝杆菌感染豚鼠中就诱导迟发型超敏 (DTH) 反应分别进行试验。发现一个组分诱导强的 DTH 反应, 接着在微内径 Vydac C18 柱 (Cat. No. 218TP5115) 上进一步经 RP-HPLC 分级分离, 所说的柱在 Perkin Elmer/Applied Biosystems Division 172 HPLC 型中。以 5-100% 线性梯度缓冲液 (在乙腈中的 0.05 % TFA) 洗脱各组分, 流速为 80 μ l/分钟。在 215 nm 监测洗脱液。收集八个组分, 在结核分枝杆菌感染豚鼠中试验对 DTH 的诱导。发现一个组分诱导约 16mm 硬结的强 DTH。其它组分不诱导可检测的 DTH。将阳性组分进行 SDS-PAGE 凝胶电泳, 发现其含有 12 kD 分子量的一单一蛋白质带。

如以上的描述, 用 Perkin Elmer/Applied Biosystems Division Procise 492 蛋白质测序仪从氨基末端对这一多肽 (此后称作 DPPD) 进行测序, 发现其具有 SEQ ID NO. 129 中显示的 N 端序列。这一序列与以上描述的基因库中的已知序列的比较揭示没有已知的同系物。分离到 DPPD 的四个溴化氰片段, 发现其具有 SEQ ID NO. 130-133 中显示的序列。

如实施例 1 的描述测定抗原 DPPD 刺激人 PBMC 增殖和产生 IFN- γ 的能力。如表 8 所示, 发现 DPPD 刺激增殖和激发大量 IFN- γ 的产生; 高于由市

售的 PPD 所激发的。

表 8
增殖和干扰素- γ 测定对 DPPD 的结果

PBMC 供体	刺激物	增殖 (CPM)	IFN- γ (OD ₄₅₀)
A	介质	1,089	0.17
	PPD (市售的)	8,394	1.29
	DPPD	13,451	2.21
B	介质	450	0.09
	PPD (市售的)	3,929	1.26
	DPPD	6,184	1.49
C	介质	541	0.11
	PPD (市售的)	8,907	0.76
	DPPD	23,024	>2.70

实施例 5

合成多肽的合成

可以采用由 HPTU (O-苯并三唑-N, N, N', N'-四甲基脲鎓(uronium) 六氟磷酸盐) 活化的 FMOC 化学在 Millipore 9050 肽合成仪上合成多肽。Gly-Cys-Gly 序列可以连接到肽的氨基末端, 以提供所述肽的缀合或标记方法。可以采用下列切割混合物从固相支持体上切割肽: 三氟乙酸: 乙烷二硫酚: 苯硫基甲烷: 水: 苯酚(40: 1: 2: 2: 3)。在切割两小时后, 可以在冷的甲基-叔丁基醚中沉淀所说的肽。然后, 肽沉淀可以溶解在含 0.1 % 三氟乙酸(TFA) 的水中, 并且在经 C18 反相 HPLC 纯化之前冷干。在水(含 0.1 %TFA) 中的 0-60 %乙腈(含 0.1 %TFA) 梯度液可以用于洗脱肽。在纯组分的冷干后, 可以采用电喷射质谱测定法和氨基酸分析确定肽的特征。

从以上所述可以清楚看到, 虽然为说明的目的, 本文描述了本发明的特定的实施方案, 但是可以进行各种修改而不背离本发明的精神和范围。

序 列 表

(1) 一般信息:

- (i) 申请人: Corixa 公司
- (ii) 发明名称: 用于免疫治疗和诊断结核病的化合物和方法
- (ii) 序列数: 137 个
- (iv) 通讯地址:
 - (A) 收信人: SEED 和 BERRY LLP
 - (B) 街道: 6300 哥伦比亚中心, 第五大街 701 号
 - (C) 城市: Seattle
 - (D) 州: 华盛顿
 - (E) 国家: 美国
 - (F) ZIP : 98104-7092
- (v) 计算机可读形式:
 - (A) 介质类型: 软盘
 - (B) 计算机: IBM PC 兼容机
 - (C) 操作系统: PC-DOS/MS-DOS
 - (D) 软件: PatentIn Release#1.0, 版本#1.30
- (vi) 当前申请的数据:
 - (A) 申请号:
 - (B) 申请日: 1996-8-27
 - (C) 分类号:
- (viii) 律师/代理人信息:
 - (A) 姓名: Maki, David J.
 - (B) 登记号: 31.392
 - (C) 证书号: 210121.417PC
- (ix) 电讯信息:
 - (A) 电话: (206) 622-4900
 - (B) 传真: (206) 682-6031

(2) SEQ ID NO: 1 的信息:

(i) 序列特征:

(A) 长度: 766 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 1:

```

CGAGGCACCG GTAGTTTGAA CCAAACGCAC AATCGACGGG CAAACGAACG GAAGAACACA      60
ACCATGAAGA TGGTGAAATC GATCGCCGCA GGTCTGACCG CCGCGGCTGC AATCGGCGCC      120
GCTGCGGCCG GTGTGACTTC GATCATGGCT GGCGGCCCGG TCGTATACCA GATGCAGCCG      180
GTCGTCTTCG GCGCGCCACT GCCGTTGGAC CCGGCATCCG CCCCTGACGT CCCGACCGCC      240
GCCCAGTTGA CCAGCCTGCT CAACAGCCTC GCCGATCCCA ACGTGTCGTT TGCGAACAAG      300
GGCAGTCTGG TCGAGGGCGG CATCGGGGGC ACCGAGGCGC GCATCGCCGA CCACAAGCTG      360
AAGAAGGCCG CCGAGCACGG GGATCTGCCG CTGTGTTCA GCGTGACGAA CATCCAGCCG      420
GCGGCCGCCG GTTCGGCCAC CGCCGACGTT TCCGTCTCGG GTCCGAAGCT CTCGTGCGCG      480
GTCACGCAGA ACGTCACGTT CGTGAATCAA GGCGGCTGGA TGCTGTCACG CGCATCGGCG      540
ATGGAGTTGC TGCAGGCCGC AGGGNAACTG ATTGGCGGGC CGGNTTCAGC CCGCTGTTCA      600
GCTACGCCGC CCGCCTGGTG ACGCGTCCAT GTCGAACACT CGCGCGTGTA GCACGGTGCG      660
GTNTGCGCAG GGNCGCACGC ACCGCCCGGT GCAAGCCGTC CTCGAGATAG GTGGTGNCTC      720
GNCACCAGNG ANCACCCCN NNTCGNCNNT TCTCGNTGNT GNATGA                        766

```

(2) SEQ ID NO: 2 的信息:

(i) 序列特征:

(A) 长度: 752 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 2:

ATGCATCACC ATCACCATCA CGATGAAGTC ACGGTAGAGA CGACCTCCGT CTTCCGCGCA	60
GACTTCCTCA GCGAGCTGGA CGCTCCTGCG CAAGCGGGTA CGGAGAGCGC GGTCTCCGGG	120
GTGGAAGGGC TCCCGCCGGG CTCGGCGTTG CTGGTAGTCA AACGAGGCC CAACGCCGGG	180
TCCCGGTTCC TACTCGACCA AGCCATCAG TCGGCTGGTC GGCATCCCGA CAGCGACATA	240
TTTCTCGACG ACGTGACCGT GAGCCGTCGC CATGCTGAAT TCCGGTTGGA AAACAACGAA	300
TTCAATGTCG TCGATGTCGG GAGTCTCAAC GGCACCTACG TCAACCGCGA GCCCGTGGAT	360
TCGGCGGTGC TGGCGAACGG CGACGAGGTC CAGATCGGCA AGCTCCGGTT GGTGTTCTTG	420
ACCGGACCCA AGCAAGGCGA GGATGACGGG AGTACCGGGG GCCCGTGAGC GCACCCGATA	480
GCCCCGCGCT GGCCGGGATG TCGATCGGGG CGGTCCTCCG ACCTGCTACG ACCGGATTTT	540
CCCTGATGTC CACCATCTCC AAGATTCGAT TCTTGGGAGG CTTGAGGGTC NGGGTGACCC	600
CCCCGCGGGC CTCATTNCGG GGTNTCGGCN GGTTCACCC CNTACCNACT GCCNCCGGN	660
TTGCNAATTC NTTCTTCNCT GCCCNAAAG GGACCNTTAN CTTGCCGCTN GAAANGGTNA	720
TCCNGGGCCC NTCCTNGAAN CCCCNTCCCC CT	752

(2) SEQ ID NO: 3 的信息:

(i) 序列特征:

(A) 长度: 813 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 3:

CATATGCATC ACCATCACCA TCACACTTCT AACCGCCCAG CGCGTCGGGG GCGTCGAGCA	60
CCACGCGACA CCGGGCCCGA TCGATCTGCT AGCTTGAGTC TGGTCAGGCA TCGTCGTCAG	120
CAGCGCGATG CCCTATGTTT GTCGTGCGACT CAGATATCGC GGCAATCCAA TCTCCCGCCT	180
GCGGCCGGCG GTGCTGCAAA CTA CTCCCGG AGGAATTTG ACGTGCGCAT CAAGATCTTC	240
ATGCTGGTCA CGGCTGTGCT TTTGCTCTGT TGTCGGGTG TGGCCACGGC CGCGCCCAAG	300
ACCTACTGCG AGGAGTTGAA AGGCACCGAT ACCGGCCAGG CGTGCCAGAT TCAAATGTCC	360
GACCCGGCCT ACAACATCAA CATCAGCCTG CCCAGTTACT ACCCCGACCA GAAGTCGCTG	420
GAAAATTACA TCGCCCAGAC GCGCGACAAG TTCCTCAGCG CGGCCACATC GTCCACTCCA	480
CGCGAAGCCC CCTACGAATT GAATATCACC TCGGCCACAT ACCAGTCCGC GATACCGCCG	540
CGTGGTACGC AGGCCGTGGT GCTCAMGGTC TACCACAACG CCGGCGGCAC GCACCCAACG	600
ACCACGTACA AGGCCTTCGA TTGGGACCAG GCCTATCGCA AGCCAATCAC CTATGACACG	660
CTGTGGCAGG CTGACACCGA TCCGCTGCCA GTCGTCTTCC CCATTGTTGC AAGGTGAACT	720
GAGCAACGCA GACCGGGACA ACWGGTATCG ATAGCCGCCN AATGCCGGCT TGGAACCCNG	780
TGAAATTATC ACAACTTCGC AGTCACNAAA NAA	813

(2) SEQ ID NO: 4 的信息:**(i) 序列特征:**

- (A) 长度: 447 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 4:

CGGTATGAAC ACGGCCGCGT CCGATAACTT CCAGCTGTCC CAGGGTGGGC AGGGATTTCGC	60
CATTCCGATC GGGCAGGCGA TGGCGATCGC GGGCCAGATC CGATCGGGTG GGGGGTCACC	120
CACCGTTCAT ATCGGGCCTA CCGCCTTCCT CGGCTTGGGT GTTGTGACA ACAACGGCAA	180
CGGCGCACGA GTCCAACGCG TGGTCGGGAG CGCTCCGGCG GCAAGTCTCG GCATCTCCAC	240
CGGCGACGTG ATCACC GCGG TCGACGGCGC TCCGATCAAC TCGGCCACCG CGATGGCGGA	300
CGCGCTTAAC GGGCATCATC CCGGTGACGT CATCTCGGTG AACTGGCAAA CCAAGTCGGG	360
CGGCACGCGT ACAGGGAACG TGACATTGGC CGAGGGACCC CCGGCCTGAT TTCGTCGYGG	420
ATACCACCCG CCGGCCGGCC AATTGGA	447

(2) SEQ ID NO: 5 的信息:

(i) 序列特征:

- (A) 长度: 604 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 5:

GTCCCACTGC GGTGCGCGAG TATGTCGCCC AGCAAATGTC TGGCAGCCGC CCAACGGAAT	60
CCGGTGATCC GACGTCGAG GTTGTGGAAC CCGCCGCCGC GGAAGTATCG GTCCATGCCT	120
AGCCCGGCGA CGGCGAGCGC CGGAATGGCG CGAGTGAGGA GGCGGGCAAT TTGGCGGGGC	180
CCGGCGACGG NGAGCGCCGG AATGGCGCGA GTGAGGAGGT GGNCAGTCAT GCCAGNGTG	240
ATCCAATCAA CCTGNATTCG GNCTGNNGGN CCATTTGACA ATCGAGGTAG TGAGCGCAAA	300
TGAATGATGG AAAACGGGNG GNGACGTCCG NTGTTCTGGT GGTGNTAGGT GNCTGNCTGG	360

NGTNGNGGNT ATCAGGATGT TCTTCGNCGA AANCTGATGN CGAGGAACAG GGTGTNCCCG	420
NNANNCCNAN GNGTCCNAN CCCNNNTCC TCGNCGANAT CANANAGNCG NTTGATGNGA	480
NAAAAGGGTG GANCAGNNNN AANTNGNGGN CCNAANAANC NNNANNGNNG NNAGNTNGNT	540
NNNTNTTNNC ANNNNNNTG NNGNNGNCCN NNNCAANCNN NTNNNNGNAA NNGGNTTNTT	600
NAAT	604

(2) SEQ ID NO: 6 的信息:

(i) 序列特征:

(A) 长度: 633 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 6:

TTGCANGTCG AACCACCTCA CTAAAGGGAA CAAAAGCTNG AGCTCCACCG CGGTGGCGGC	60
CGCTCTAGAA CTAGTGKATM YYYCKGGCTG CAGSAATYCG GYACGAGCAT TAGGACAGTC	120
TAACGGTCCT GTTACGGTGA TCGAATGACC GACGACATCC TGCTGATCGA CACCGACGAA	180
CGGGTGCGAA CCCTCACCCT CAACCGGCCG CAGTCCCGYA ACGCGCTCTC GCGGGCGCTA	240
CGGGATCGGT TTTTCGCGGY GTTGGYCGAC GCCGAGGYCG ACGACGACAT CGACGTCGTC	300
ATCCTCACCG GYGCCGATCC GGTGTTCTGC GCCGGACTGG ACCTCAAGGT AGCTGGCCGG	360
GCAGACCGCG CTGCCGACA TCTACCGCG GTGGGCGGCC ATGACCAAGC CGGTGATCGG	420
CGCGATCAAC GCGCCGCGG TCACCGGCGG GCTCGAACTG GCGCTGTACT GCGACATCCT	480
GATCGCCTCC GAGCACGCC GCTTCGNCGA CACCCACGCC CGGGTGGGGC TGCTGCCAC	540
CTGGGGACTC AGTGTGTGCT TGCCGCAAAA GGTGGCATC GGNTGGGCC GGTGGATGAG	600
CCTGACCGGC GACTACCTGT CCGTGACCGA CGC	633

(2) SEQ ID NO: 7 的信息:

(i) 序列特征:

(A) 长度: 1362 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 7:

CGACGACGAC GGC GCCGAG AGCGGGCGCG AACGGCGATC GACGCGGCC TGGCCAGAGT	60
CGGCACCACC CAGGAGGGAG TCGAATCATG AAATTTGTCA ACCATATTGA GCCCGTCGCG	120
CCCCGCCGAG CCGGCGGGCG GGTGCGCCGAG GTCTATGCCG AGGCCCGCCG CGAGTTCGGC	180
CGGCTGCCCC AGCGCTCGC CATGCTGTCC CCGGACGAGG GACTGCTCAC CGCCGGCTGG	240
GCGACGTTGC GCGAGACACT GCTGGTGGGC CAGGTGCCGC GTGGCCGCAA GGAAGCCGTC	300
GCCGCCGCCG TCGCGGCCAG CCTGCGCTGC CCCTGGTGCG TCGACGCACA CACCACCATG	360
CTGTACGCGG CAGGCCAAAC CGACACCGCC GCGGCGATCT TGGCCGGCAC AGCACCTGCC	420
GCCGGTGACC CGAACGCGCC GTATGTGGCG TGGCGGCAG GAACCGGGAC ACCGGCGGGA	480
CCGCCGGCAC CGTTCGGCCC GGATGTCGCC GCCGAATACC TGGGCACCGC GGTGCAATTC	540
CACTTCATCG CACGCCTGGT CCTGGTGCTG CTGGACGAAA CCTTCCTGCC GGGGGGCCCG	600
CGCGCCCAAC AGCTCATGCG CCGCGCCGGT GGA CTGGTGT TCGCCCGCAA GGTGCGCGCG	660
GAGCATCGGC CGGGCCGCTC CACCCGCCGG CTCGAGCCGC GAACGCTGCC CGACGATCTG	720
GCATGGGCAA CACCGTCCGA GCCCATAGCA ACCGCGTTCG CCGCGCTCAG CCACCACCTG	780
GACACCGCGC CGCACCTGCC GCCACCGACT CGTCAGGTGG TCAGGCGGGT CGTGGGGTCG	840
TGGCACGGCG AGCCAATGCC GATGAGCAGT CGCTGGACGA ACGAGCACAC CGCCGAGCTG	900

CCCGCCGACC TGCACGCGCC CACCCGTCTT GCCCTGCTGA CCGGCCTGGC CCCGCATCAG	960
GTGACCGACG ACGACGTCGC CGCGGCCCGA TCCCTGCTCG ACACCGATGC GGCGCTGGTT	1020
GGCGCCCTGG C̃CTGGGCCGC CTTACCGCC GCGCGGCGCA TCGGCACCTG GATCGGCGCC	1080
GCCGCCGAGG GCCAGGTGTC GCGGCAAAAC CCGACTGGGT GAGTGTGCGC GCCCTGTCGG	1140
TAGGGTGTCA TCGCTGGCCC GAGGGATCTC GCGGCGGCGA ACGGAGGTGG CGACACAGGT	1200
GGAAGCTGCG CCCACTGGCT TGCGCCCAA CGCCGTCGTG GCGGTTGGT TGGCCGCACT	1260
GGCCGATCAG GTCGGCGCCG GCCCTTGCC GAAGGTCCAG CTCAACGTGC CGTCACCGAA	1320
GGACCGGACG GTCACCGGG GTCACCCTGC GCGCCAAGG AA	1362

(2) SEQ ID NO: 8 的信息:

(i) 序列特征:

(A) 长度: 1458 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 8:

GCGACGACCC CGATATGCCG GGCACCGTAG CGAAAGCCGT CGCCGACGCA CTCGGGCGCG	60
GTATCGCTCC CGTTGAGGAC ATTCAGGACT GCGTGAGGC CCGGCTGGGG GAAGCCGGTC	120
TGGATGACGT GGCCCGTGTT TACATCATCT ACCGGCAGCG GCGCGCCGAG CTGCGGACGG	180
CTAAGGCCTT GCTCGGCGTG CGGGACGAGT TAAAGCTGAG CTTGGCGGCC GTGACGGTAC	240
TGCGCGAGCG CTATCTGCTG CACGACGAGC AGGGCCGGCC GGCCGAGTCG ACCGGCGAGC	300
TGATGGACCG ATCGGCGCGC TGTGTCGCGG CGGCCGAGGA CCAGTATGAG CCGGGCTCGT	360

CGAGGCGGTG GGCCGAGCGG TTCGCCACGC TATTACGCAA CCTGGAATTC CTGCCGAATT	420
CGCCCACGTT GATGAACTCT GGCACCGACC TGGGACTGCT CGCCGGCTGT TTTGTTCTGC	480
CGATTGAGGA TTCGCTGCAA TCGATCTTTG CGACGCTGGG ACAGGCCGCC GAGCTGCAGC	540
GGGCTGGAGG CGGCACCGGA TATGCGTTCA GCCACCTGCG ACCCGCCGGG GATCGGGTGG	600
CCTCCACGGG CGGCACGGCC AGCGGACCGG TGTCGTTTCT ACGGCTGTAT GACAGTGCCG	660
CGGGTGTGGT CTCCATGGGC GGTGCCCGC GTGGCGCCTG TATGGCTGTG CTTGATGTGT	720
CGCACCCGGA TATCTGTGAT TTCGTCACCG CCAAGGCCGA ATCCCCAGC GAGCTCCCGC	780
ATTTCAACCT ATCGGTTGGT GTGACCGACG CGTTCCTGCG GGCCGTCGAA CGCAACGGCC	840
TACACCGGCT GGTCAATCCG CGAACCGCA AGATCGTCGC GCGGATGCC GCGCCGAGC	900
TGTTGACGC CATCTGCAA GCGCGCACG CCGGTGGCGA TCCCGGGCTG GTGTTTCTCG	960
ACACGATCAA TAGGGCAAAC CCGGTGCCG GGAGAGGCCG CATCGAGGCG ACCAACCCGT	1020
GCGGGGAGGT CCCACTGCTG CCTTACGAGT CATGTAATCT CGGCTCGATC AACCTGCCC	1080
GGATGCTCGC CGACGGTCGC GTCGACTGGG ACCGGCTCGA GGAGGTCGCC GGTGTGGCGG	1140
TGCGGTTCCCT TGATGACGTC ATCGATGTCA GCCGCTACCC CTTCCCCGAA CTGGGTGAGG	1200
CGGCCCCGCG CACCCGCAAG ATCGGGCTGG GAGTCATGGG TTTGGCGGAA CTGCTTGCCG	1260
CACTGGGTAT TCCGTACGAC AGTGAAGAAG CCGTGCGGTT AGCCACCCGG CTCATGCGTC	1320
GCATACAGCA GCGGGCGCAC ACGGCATCGC GGAGGCTGGC CGAAGAGCGG GGCGCATTCC	1380
CGGCGTTCAC CGATAGCCGG TTCGCGCGGT CGGGCCCGAG GCGCAACGCA CAGGTACCT	1440
CCGTCGCTCC GACGGGCA	1458

(2) SEQ ID NO: 9 的信息:

(i) 序列特征:

(A) 长度: 862 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 9:

```

ACGGTGTAAAT C̄GTGCTGGAT CTGGAACCGC GTGGCCCGCT ACCTACCGAG ATCTACTGGC      60
GGCGCAGGGG GCTGGCCCTG GGCATCGCGG TCGTCGTAGT CGGGATCGCG GTGGCCATCG      120
TCATCGCCTT CGTCGACAGC AGCGCCGGTG CCAAACCGGT CAGCGCCGAC AAGCCGGCCT      180
CCGCCCAGAG CCATCCGGGC TCGCCGGCAC CCCAAGCACC CCAGCCGGCC GGGCAAACCG      240
AAGGTAACGC CGCCGCGGCC CCGCCGCAGG GCCAAAACCC CGAGACACCC ACGCCCACCG      300
CCGCGGTGCA GCCGCCGCCG GTGCTCAAGG AAGGGGACGA TTGCCCGAT TCGACGCTGG      360
CCGTCAAAGG TTTGACCAAC GCGCCGCAGT ACTACGTCGG CGACCAGCCG AAGTTCACCA      420
TGGTGGTCAC CAACATCGGC CTGGTGTCTT GTAAACGCGA CGTTGGGGCC GCGGTGTTGG      480
CCGCCTACGT TTA CTGCTG GACAACAAGC GGTGTGGTC CAACCTGGAC TGCGCGCCCT      540
CGAATGAGAC GCTGGTCAAG ACGTTTTCCC CCGGTGAGCA GGTAACGACC GCGGTGACCT      600
GGACCGGGAT GGGATCGGCG CCGCGCTGCC CATTGCCGCG GCCGGCGATC GGGCCGGGCA      660
CCTACAATCT CGTGGTACAA CTGGGCAATC TGCCTCGCT GCCGGTCCG TTCATCCTGA      720
ATCAGCCGCC GCCGCCGCC GGGCCGGTAC CCGCTCCGGG TCCAGCGCAG GCGCCTCCGC      780
CGGAGTCTCC CGCGCAAGGC GGATAATTAT TGATCGCTGA TGGTCGATTC CGCCAGCTGT      840
GACAACCCCT CGCCTCGTGC CG      862

```

(2) SEQ ID NO: 10 的信息:

(i) 序列特征:

(A) 长度: 622 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 10:

TTGATCAGCA CCGGCAAGGC GTCACATGCC TCCCTGGGTG TGCAGGTGAC CAATGACAAA	60
GACACCCCGG GCGCCAAGAT CGTCGAAGTA GTGGCCGGTG GTGCTGCCGC GAACGCTGGA	120
GTGCCGAAGG GCGTCGTTGT CACCAAGGTC GACGACCGCC CGATCAACAG CGCGGACGCG	180
TTGGTTGCCG CCGTGCGGTC CAAAGCGCCG GCGGCCACGG TGGCGCTAAC CTTTCAGGAT	240
.CCCTCGGGCG GTAGCCGCAC AGTGCAAGTC ACCCTCGGCA AGGCGGAGCA GTGATGAAGG	300
TCGCCGCGCA GTGTTCAAAG CTCGGATATA CGGTGGCACC CATGGAACAG CGTGCGGAGT	360
TGGTGGTTGG CCGGGCACTT GTCGTCGTCG TTGACGATCG CACGGCGCAC GCGATGAAG	420
ACCACAGCGG GCCGCTTGTC ACCGAGCTGC TCACCGAGGC CGGGTTTGTT GTCGACGGCG	480
TGGTGGCGGT GTCGGCCGAC GAGGTCGAGA TCCGAAATGC GCTGAACACA GCGGTGATCG	540
GCGGGGTGGA CCTGGTGGTG TCGGTCGGCG GGACCGNGT GACGNCTCGC GATGTCACCC	600
CGGAAGCCAC CCGNGACATT CT	622

(2) SEQ ID NO: 11 的信息:

(i) 序列特征:

(A) 长度: 1200 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 11:

GGCGCAGCGG TAAGCCTGTT GGCCGCCGGC ACACTGGTGT TGACAGCATG CGGCGGTGGC	60
ACCAACAGCT CGTCGTCAGG CGCAGGCGGA ACGTCTGGGT CGGTGCACTG CGGCGGCAAG	120
AAGGAGCTCC ACTCCAGCGG CTCGACCGCA CAAGAAAATG CCATGGAGCA GTTCGTCTAT	180
GCCTACGTGC GATCGTGCCC GGGCTACACG TTGGACTACA ACGCCAACGG GTCCGGTGCC	240
GGGGTGACCC AGTTTCTCAA CAACGAAACC GATTTGCGCG GCTCGGATGT CCCGTTGAAT	300
CCGTCGACCG GTCAACCTGA CCGGTCGGCG GAGCGGTGCG GTTCCCCGGC ATGGGACCTG	360
CCGACGGTGT TCGGCCCGAT CGCGATCACC TACAATATCA AGGGCGTGAG CACGCTGAAT	420
CTTGACGGAC CCACTACCGC CAAGATTTTC AACGGCACCA TCACCGTGTG GAATGATCCA	480
CAGATCCAAG CCCTCAACTC CGGCACCGAC CTGCCGCCAA CACCGATTAG CGTTATCTTC	540
CGCAGCGACA AGTCCGGTAC GTCGGACAAC TTCCAGAAAT ACCTCGACGG TGTATCCAAC	600
GGGGCGTGGG GCAAAGGCGC CAGCGAAACG TTCAGCGGGG GCGTCGGCGT CGGCGCCAGC	660
GGGAACAACG GAACGTCGGC CCTACTGCAG ACGACCGACG GGTGATCAC CTACAACGAG	720
TGGTCGTTTG CGGTGGGTAA GCAGTTGAAC ATGGCCCAGA TCATCACGTC GGCGGGTCCG	780
GATCCAGTGG CGATCACCAC CGAGTCGGTC GGTAAGACAA TCGCCGGGGC CAAGATCATG	840
GGACAAGGCA ACGACCTGGT ATTGGACACG TCGTCGTTCT ACAGACCCAC CCAGCCTGGC	900
TCTTACCCGA TCGTGCTGGC GACCTATGAG ATCGTCTGCT CGAAATACCC GGATGCGACG	960
ACCGGTACTG CGGTAAGGGC GTTTATGCAA GCCGCGATTG GTCCAGGCCA AGAAGGCCTG	1020
GACCAATACG GCTCCATTCC GTTGCCCAA TCGTTCCAAG CAAAATTGGC GGCCGCGGTG	1080
AATGCTATTT CTTGACCTAG TGAAGGGAAT TCGACGGTGA GCGATGCCGT TCCGACGGTA	1140
GGGTCGCAAT TTGGGCCGTA TCAGCTATTG CGGCTGCTGG GCCGAGGCGG GATGGGCGAG	1200

(2) SEQ ID NO: 12 的信息:

(i) 序列特征:

(A) 长度: 1155 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 12:

GCAAGCAGCT G̃CAGGTCGTG CTGTTGACG AACTGGGCAT GCCGAAGACC AAACGCACCA	60
AGACCGGCTA CACCACGGAT GCCGACGCGC TGCAGTCGTT GTTCGACAAG ACCGGGCATC	120
CGTTTCTGCA ACATCTGCTC GCCCACC GCG ACGTCACCCG GCTCAAGGTC ACCGTCGACG	180
GGTTGCTCCA AGCGGTGGCC GCCGACGGCC GCATCCACAC CACGTTCAAC CAGACGATCG	240
CCGCGACCGG CCGGCTCTCC TCGACCGAAC CCAACCTGCA GAACATCCCG ATCCGCACCG	300
ACGCGGGCCG GCGGATCCGG GACGCGTTG TGGTCGGGGA CGGTTACGCC GAGTTGATGA	360
CGGCCGACTA CAGCCAGATC GAGATGCGGA TCATGGGGCA CCTGTCCGGG GACGAGGGCC	420
TCATCGAGGC GTTCAACACC GGGGAGGACC TGTATTGTT CGTCGCGTCC CGGGTGTTG	480
GTGTGCCCAT CGACGAGGTC ACCGGCGAGT TCGGGCGCCG GGTCAAGGCG ATGTCCTACG	540
GGCTGGTTTA CGGGTTGAGC GCCTACGGCC TGTCGCAGCA GTTGAAAATC TCCACCGAGG	600
AAGCCAACGA GCAGATGGAC GCGTATTTG CCGATTGCG CGGGGTGCGC GACTACCTGC	660
GCGCCGTAGT CGAGCGGGCC CGCAAGGACG GCTACACCTC GACGGTGCTG GGCCGTCGCC	720
GCTACCTGCC CGAGCTGGAC AGCAGCAACC GTCAAGTGCG GGAGGCCGCC GAGCGGGCGG	780
CGCTGAACGC GCCGATCCAG GGCAGCGCGG CCGACATCAT CAAGGTGGCC ATGATCCAGG	840
TCGACAAGGC GCTCAACGAG GCACAGCTGG CGTCGCGCAT GCTGCTGCAG GTCCACGACG	900
AGCTGCTGTT CGAAATCGCC CCCGGTGAAC GCGAGCGGGT CGAGGCCCTG GTGCGCGACA	960

AGATGGGCGG CGCTTACCCG CTCGACGTCC CGCTGGAGGT GTCGGTGGGC TACGGCCGCA	1020
GCTGGGACGC GGCGGCGCAC TGAGTGCCGA GCGTGATCT GGGGCGGGAA TTCGGCGATT	1080
TTTCCGCCCT GAGTTCACGC TCGGCGCAAT CGGGACCGAG TTTGTCCAGC GTGTACCCGT	1140
CGAGTAGCCT CGTCA	1155

(2) SEQ ID NO: 13 的信息:

(i) 序列特征:

(A) 长度: 1771 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 13:

GAGCGCCGTC TGGTGTTTGA ACGGTTTTAC CGGTCGGCAT CGGCACGGGC GTTGCCGGGT	60
TCGGGCCTCG GGTGGCGAT CGTCAACAG GTGGTGCTCA ACCACGGCGG ATTGCTGCGC	120
ATCGAAGACA CCGACCCAGG CGGCCAGCCC CCTGGAACGT CGATTTACGT GCTGCTCCCC	180
GGCCGTGCGA TGCCGATTCC GCAGCTTCCC GGTGCGACGG CTGGCGCTCG GAGCACGGAC	240
ATCGAGAACT CTCGGGGTTC GGCGAACGTT ATCTCAGTGG AATCTCAGTC CACGCGCGCA	300
ACCTAGTTGT GCAGTTACTG TTGAAAGCCA CACCCATGCC AGTCCACGCA TGGCCAAGTT	360
GGCCCGAGTA GTGGGCCTAG TACAGGAAGA GCAACCTAGC GACATGACGA ATCACCACG	420
GATTTCGCCA CCGCCGACG AGCCGGGAAC CCCAGGTTAT GCTCAGGGGC AGCAGCAAAC	480
GTACAGCCAG CAGTTCGACT GGC GTTACCC ACCGTCCCCG CCCCCGACG CAACCCAGTA	540

CCGTCAACCC TACGAGGCGT TGGGTGGTAC CCGGCCGGGT CTGATACCTG GCGTGATTCC	600
GACCATGACG CCCCTCCTG GGATGGTTCG CCAACGCCCT CGTGCAGGCA TGTGGCCAT	660
CGGCGCGGTG ACGATAGCGG TGGTGTCCGC CGGCATCGGC GCGCGGCCG CATCCCTGGT	720
CGGGTTCAAC CGGGCACCCG CCGGCCCCAG CGGCGGCCCA GTGGCTGCCA GCGCGGCGCC	780
AAGCATCCCC GCAGCAAACA TGCCGCCGGG GTCGGTCGAA CAGGTGGCGG CCAAGGTGGT	840
GCCCAGTGTC GTCATGTTGG AAACCGATCT GGGCCGCCAG TCGGAGGAGG GCTCCGGCAT	900
CATTCTGTCT GCCGAGGGGC TGATCTTGAC CAACAACCAC GTGATCGCGG CGGCCGCCAA	960
GCCTCCCCTG GGCAGTCCGC CGCCGAAAAC GACGGTAACC TTCTCTGACG GGCAGACCGC	1020
ACCCTTCACG GTGGTGGGGG CTGACCCAC CAGTGATATC GCCGTGCTCC GTGTTCAAGG	1080
CGTCTCCGGG CTCACCCCGA TCTCCCTGGG TTCCTCTCG GACCTGAGGG TCGGTCAGCC	1140
GGTGCTGGCG ATCGGGTCGC CGCTCGGTTT GGAGGGCACC GTGACCACGG GGATCGTCAG	1200
CGCTCTCAAC CGTCCAGTGT CGACGACCGG CGAGGCCGGC AACCAGAACA CCGTGCTGGA	1260
CGCCATTCAG ACCGACGCCG CGATCAACCC CGGTAAGTCC GGGGGCGCGC TGGTGAACAT	1320
GAACGCTCAA CTCGTGGAG TCAACTCGGC CATTGCCACG CTGGGCGCGG ACTCAGCCGA	1380
TGCGCAGAGC GGCTCGATCG GTCTCGGTTT TGCGATTCCA GTCGACCAGG CCAAGCGCAT	1440
CGCCGACGAG TTGATCAGCA CCGCAAGGC GTCACATGCC TCCCTGGGTG TGCAGGTGAC	1500
CAATGACAAA GACACCCCGG GCGCCAAGAT CGTCGAAGTA GTGGCCGGTG GTGCTGCCGC	1560
GAACGCTGGA GTGCCGAAGG GCGTCGTTGT CACCAAGGTC GACGACCGCC CGATCAACAG	1620
CGCGGACCGG TTGGTTGCCG CCGTGCGGTC CAAAGCGCCG GCGCCACGG TGGCGCTAAC	1680
CTTTCAGGAT CCCTCGGGCG GTAGCCGCAC AGTGCAAGTC ACCCTCGGCA AGGCGGAGCA	1740
GTGATGAAGG TCGCCGCGCA GTGTTCAAAG C	1771

(2) SEQ ID NO: 14 的信息:

(i) 序列特征:

(A) 长度: 1058 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 14:

CTCCACCGCG GTGGCGGCCG CTCTAGAACT AGTGGATCCC CCGGGCTGCA GGAATTCGGC	60
ACGAGGATCC GACGTCGCAG GTTGTGCAAC CCGCCGCCGC GGAAGTATCG GTCCATGCCT	120
AGCCCGGGCGA CGGCGAGCGC CGGAATGGCG CGAGTGAGGA GGCGGGCAAT TTGGCGGGGC	180
CCGGCGACGG CGAGCGCCGG AATGGCGCGA GTGAGGAGGC GGGCAGTCAT GCCCAGCGTG	240
ATCCAATCAA CCTGCATTCG GCCTGCGGGC CCATTTGACA ATCGAGGTAG TGAGCGCAAA	300
TGAATGATGG AAAACGGGCG GTGACGTCCG CTGTTCTGGT GGTGCTAGGT GCCTGCCTGG	360
CGTTGTGGCT ATCAGGATGT TCTTCGCCGA AACCTGATGC CGAGGAACAG GGTGTTCCCG	420
TGAGCCCGAC GGGCTCCGAC CCCGCGCTCC TCGCCGAGAT CAGGCAGTCG CTTGATGCGA	480
CAAAAGGGTT GACCAGCGTG CACGTAGCGG TCCGAACAAC CGGGAAAGTC GACAGCTTGC	540
TGGGTATTAC CAGTGCCGAT GTCGACGTCC GGGCCAATCC GCTCGCGGCA AAGGGCGTAT	600
GCACCTACAA CGACGAGCAG GGTGTCCCGT TTCGGGTACA AGGCGACAAC ATCTCGGTGA	660
AACTGTTCTGA CGACTGGAGC AATCTCGGCT CGATTTCTGA ACTGTCAACT TCACGCGTGC	720
TCGATCCTGC CGCTGGGGTG ACGCAGCTGC TGTCCGGTGT CACGAACCTC CAAGCGCAAG	780
GTACCGAAGT GATAGACGGA ATTTGACCA CCAAATCAC CGGGACCATC CCCGCGAGCT	840

CTGTCAAGAT GCTTGATCCT GGCGCCAAGA GTGCAAGGCC GCGACCGTG TGGATTGCCC	900
AGGACGGCTC GCACCACCTC GTCCGAGCGA GCATCGACCT CGGATCCGGG TCGATTCAGC	960
TCACGCAGTC GAAATGGAAC GAACCCGTCA ACGTCGACTA GGCCGAAGTT GCGTCGACGC	1020
GTTGNTCGAA ACGCCCTTGT GAACGGTGTC AACGGNAC	1058

(2) SEQ ID NO: 15 的信息:

(i) 序列特征:

(A) 长度: 542 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 15:

GAATTCGGCA CGAGAGGTGA TCGACATCAT CGGGACCAGC CCCACATCCT GGGAACAGGC	60
GGCGGGCGGAG GCGGTCCAGC GGGCGCGGGA TAGCGTCGAT GACATCCGCG TCGCTCGGGT	120
CATTGAGCAG GACATGGCCG TGGACAGCGC CGGCAAGATC ACCTACCGCA TCAAGCTCGA	180
AGTGTCGTTT AAGATGAGGC CGGCGCAACC GCGCTAGCAC GGGCCGGCGA GCAAGACGCA	240
AAATCGCACG GTTTGCGGTT GATTCGTGCG ATTTTGTGTC TGCTCGCCGA GGCCTACCAG	300
GCGCGGCCCA GGTCCGCGTG CTGCCGTATC CAGGCGTGCA TCGCGATTCC GGCGGCCACG	360
CCGGAGTTAA TGCTTCGCGT CGACCCGAAC TGGGCGATCC GCCGGNGAGC TGATCGATGA	420
CCGTGGCCAG CCCGTCGATG CCCGAGTTGC CCGAGGAAAC GTGCTGCCAG GCCGGTAGGA	480
AGCGTCCGTA GGCGGGCGGTG CTGACCGGCT CTGCCTGCGC CCTCAGTGCG GCCAGCGAGC	540
GG	542

(2) SEQ ID NO: 16 的信息:

(i) 序列特征:

(A) 长度: 913 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 16:

CGGTGCCGCC CGCGCCTCCG TTGCCCCCAT TGCCGCCGTC GCCGATCAGC TGCGCATCGC	60
CACCATCACC GCCTTTGCCG CCGGCACCGC CGGTGGCGCC GGGGCCGCCG ATGCCACCGC	120
TTGACCCTGG CCGCCGGCGC CGCATTGCC ATACAGCACC CCGCCGGGGG CACCGTTACC	180
GCCGTCGCCA CCGTCGCCGC CGTGCCGTT TCAGGCCGGG GAGGCCGAAT GAACGCCCGC	240
CAAGCCCGCC GCCGGCACCG TTGCCGCCTT TTCCGCCCGC CCCGCCGGCG CCGCCAATTG	300
CCGAACAGCC AMGCACCGTT GCCGCCAGCC CCGCCGCCGT TAACGGCGCT GCCGGGCGCC	360
GCCGCCGGAC CCGCCATTAC CGCCGTTCCC GTTCGGTGCC CCGCCGTTAC CGGCGCCGCC	420
GTTTGCCGCC AATATTCGGC GGGCACCGCC AGACCCGCCG GGGCCACCAT TGCCGCCGGG	480
CACCGAAACA ACAGCCCAAC GGTGCCGCCG GCCCGGCCGT TTGCCGCCAT CACCGGCCAT	540
TCACCGCCAG CACCGCCGTT AATGTTTATG AACCCGGTAC CGCCAGCGCG GCCCCTATTG	600
CCGGGCGCCG GAGNGCGTGC CCGCCGGCGC CGCCAACGCC CAAAAGCCCG GGGTTGCCAC	660
CGGCCCCGCC GGACCCACCG GTCCCGCGA TCCCCCGTT GCCGCCGGTG CCGCCGCCAT	720
TGGTGCTGCT GAAGCCGTTA GCGCCGGTTC CGCSGGTTCC GGCGGTGGCG CCNTGGCCGC	780
CGGCCCCGCC GTTGCCGTAC AGCCACCCCC CGGTGGCGCC GTTGCCGCCA TTGCCGCCAT	840
TGCCGCCGTT GCCGCCATTG CCGCCGTTCC CGCCGCCACC GCCGGNTTGG CCGCCGGCGC	900
CGCCGGCGGC CGC	913

(2) SEQ ID NO: 17 的信息:

(i) 序列特征:

(A) 长度: 1872 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 17:

GACTACGTTG GTGTAGAAAA ATCCTGCCGC CCGGACCCTT AAGGCTGGGA CAATTTCTGA	60
TAGCTACCCC GACACAGGAG GTTACGGGAT GAGCAATTCG CGCCGCCGCT CACTCAGGTG	120
GTCATGGTTG CTGAGCGTGC TGGCTGCCGT CGGGCTGGGC CTGGCCACGG CGCCGGCCCA	180
GGCGGCCCCG CCGGCCTTGT CGCAGGACCG GTTCGCCGAC TTCCCCGCGC TGCCCCTCGA	240
CCCGTCCGCG ATGGTCGCCC AAGTGGCGCC ACAGGTGGTC AACATCAACA CCAAAGTGGG	300
CTACAACAAC GCCGTGGGCG CCGGGACCGG CATCGTCATC GATCCCAACG GTGTCGTGCT	360
GACCAACAAC CACGTGATCG CGGGCGCCAC CGACATCAAT GCGTTCAGCG TCGGCTCCGG	420
CCAAACCTAC GGCCTCGATG TGGTCGGGTA TGACCGCACC CAGGATGTCG CGGTGCTGCA	480
GCTGCGCGGT GCCGGTGGCC TGCCGTCCGC GGCGATCGGT GGC GGCGTCG CGGTTGGTGA	540
GCCCGTCGTC GCGATGGGCA ACAGCGGTGG GCAGGGCGGA ACGCCCCGTG CGGTGCCTGG	600
CAGGGTGGTC GCGCTCGGCC AAACCGTGCA GGCCTCGGAT TCGCTGACCG GTGCCGAAGA	660
GACATTGAAC GGGTTGATCC AGTTCGATGC CGCAATCCAG CCCGGTGATT CGGGCGGGCC	720
CGTCGTCAAC GGCCTAGGAC AGGTGGTCGG TATGAACACG GCCGCGTCCG ATAACCTCCA	780
GCTGTCCCAG GGTGGGCAGG GATTCGCCAT TCCGATCGGG CAGGCGATGG CGATCGCGGG	840
CCAAATCCGA TCGGGTGGGG GGTCAACCCAC CGTTCATATC GGGCCTACCG CCTTCCTCGG	900

CTTGGGTGTT GTCGACAACA ACGGCAACGG CGCACGAGTC CAACGCGTGG TCGGAAGCGC	960
TCCGGCGGCA AGTCTCGGCA TCTCCACCGG CGACGTGATC ACCGCGGTGG ACGGCGCTCC	1020
GATCAACTCG GCCACCGCGA TGGCGGACGC GCTTAACGGG CATCATCCCG GTGACGTCAT	1080
CTCGGTGAAC TGGCAAACCA AGTCGGGCGG CACGCGTACA GGGAACTGA CATTGGCCGA	1140
GGGACCCCCG GCCTGATTG TCGCGGATAC CACCCGCCGG CCGGCCAATT GGATTGGCGC	1200
CAGCCGTGAT TGCCGCGTGA GCCCCGAGT TCCGTCTCCC GTGCGCGTGG CATTGTGGAA	1260
GCAATGAACG AGGCAGAACA CAGCGTTGAG CACCCTCCCG TGCAGGGCAG TTACGTCGAA	1320
GGCGGTGTGG TCGAGCATCC GGATGCCAAG GACTTCGGCA GCGCCGCCGC CCTGCCCGCC	1380
GATCCGACCT GGTTTAAGCA CGCCGTCTTC TACGAGGTGC TGGTCCGGGC GTTCTTCGAC	1440
GCCAGCGCGG ACGGTTCCGN CGATCTGCGT GGACTCATCG ATCGCCTCGA CTACCTGCAG	1500
TGGCTTGGA TCGACTGCAT CTGTTGCCGC CGTTCCTACG ACTCACCGCT GCGCGACGGC	1560
GGTTACGACA TTCGCGACTT CTACAAGGTG CTGCCCGAAT TCGGCACCGT CGACGATTTT	1620
GTCGCCCTGG TCGACACCGC TCACCGGCGA GGTATCCGCA TCATCACCGA CCTGGTGATG	1680
AATCACACCT CGGAGTCGCA CCCCTGGTTT CAGGAGTCCC GCCGCGACCC AGACGGACCG	1740
TACGGTGACT ATTACGTGTG GAGCGACACC AGCGAGCGCT ACACCGACGC CCGGATCATC	1800
TTCGTCGACA CCGAAGAGTC GAACTGGTCA TTCGATCCTG TCCGCCGACA GTTNCCTACTG	1860
GCACCGATTG TT	1872

(2) SEQ ID NO: 18 的信息:

(i) 序列特征:

- (A) 长度: 1482 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链

(D) 拓扑结构: 线型**(xi) 序列描述: SEQ ID NO: 18:**

CTTCGCCGAA ACCTGATGCC GAGGAACAGG GTGTTCCCGT GAGCCCGACG GCGTCCGACC	60
CCGCGCTCCT CGCCGAGATC AGGCAGTCGC TTGATGCGAC AAAAGGGTTG ACCAGCGTGC	120
ACGTAGCGGT CCGAACAACC GGGAAAGTCG ACAGCTTGCT GGGTATTACC AGTGCCGATG	180
TCGACGTCCG GGCCAATCCG CTCGCGGCAA AGGGCGTATG CACCTACAAC GACGAGCAGG	240
GTGTCCCGTT TCGGGTACAA GGCACACA TCTCGGTGAA ACTGTTGAC GACTGGAGCA	300
ATCTCGGCTC GATTTCTGAA CTGTCAACTT CACGCGTGCT CGATCCTGCC GCTGGGGTGA	360
CGCAGCTGCT GTCCGGTGTC ACGAACCTCC AAGCGCAAGG TACCGAAGTG ATAGACGGAA	420
TTTCGACCAC CAAAATCACC GGGACCATCC CCGCGAGCTC TGCAAGATG CTTGATCCTG	480
GCGCCAAGAG TGCAAGGCCG GCGACCGTGT GGATTGCCCA GGACGGCTCG CACCACCTCG	540
TCCGAGCGAG CATCGACCTC GGATCCGGGT CGATTCAGCT CACGCAGTCG AAATGGAACG	600

AACCCGTCAA CGTCGACTAG GCCGAAGTTG CGTCGACGCG TTGCTCGAAA CGCCCTTG TG	660
AACGGTGTCA ACGGCACCCG AAAACTGACC CCCTGACGGC ATCTGAAAAT TGACCCCTA	720
GACCGGGCGG TTGGTGGTTA TTCTTCGGTG GTTCCGGCTG GTGGGACGCG GCCGAGGTCG	780
CGGTCTTTGA GCCGGTAGCT GTCGCCTTTG AGGGCGACGA CTTCAGCATG GTGGACGAGG	840
CGGTCGATCA TGGCGGCAGC AACGACGTCG TCGCCGCCGA AAACCTCGCC CCACCGGCCG	900
AAGGCCTTAT TGGACGTGAC GATCAAGCTG GCCCGCTCAT ACCGGGAGGA CACCAGCTGG	960
AAGAAGAGGT TGGCGGCCTC GGGCTCAAAC GGAATGTAAC CGACTTCGTC AACCACCAGG	1020
AGCGGATAGC GGCCAAACCG GGTGAGTTCT GCGTAGATGC GCCCGGCGTG GTGAGCCTCG	1080
GCGAACCGTG CTACCCATTC GGCGGCGGTG GCGAACAGCA CCCGATGACC GGCCTGACAC	1140
GCGCGTATCG CCAGGCCGAC CGCAAGATGA GTCTTCCCGG TGCCAGGCGG GGCCAAAAA	1200
CACGACGTTA TCGCGGGCGG TGATGAAATC CAGGGTGCCC AGATGTGCGA TGGTGTCGCG	1260
TTTGAGGCCA CGAGCATGCT CAAAGTCGAA CTCTTCCAAC GACTTCCGAA CCGGGAAGCG	1320
GGCGGCGCGG ATGCGGCCCT CACCACCATG GGAATCCCGG GCTGACACTT CCCGCTGCAG	1380
GCAGGCGGCC AGGTATTCTT CGTGGCTCCA GTTCTCGGCG CGGGCGCGAT CGGCCAGCCG	1440
GGACACTGAC TCACGCAGGG TGGGAGCTTT CAATGCTCTT GT	1482

(2) SEQ ID NO: 19 的信息:

(i) 序列特征:

(A) 长度: 876 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 19:

GAATTCGGCA CGAGCCGGCG ATAGCTTCTG GGCCGCGGCC GACCAGATGG CTCGAGGGTT	60
CGTGCTCGGG GCCACCGCCG GCGCACACAC CCTGACCGGT GAGGGCCTGC AACACGCCGA	120
CGGTCACTCG TTGCTGCTGG ACGCCACCAA CCCGGCGGTG GTTGCCTACG ACCCGGCCTT	180
CGCCTACGAA ATCGGCTACA TCGNGGAAAG CGGACTGGCC AGGATGTGCG GGGAGAACCC	240
GGAGAACATC TTCTTCTACA TCACCGTCTA CAACGAGCCG TACGTGCAGC CGCCGGAGCC	300
GGAGAACTTC GATCCCGAGG GCGTGCTGGG GGGTATCTAC CGNTATCACG CGGCCACCGA	360
GCAACGCACC AACAAGNGC AGATCCTGGC CTCGGGGGTA GCGATGCCCG CGGCGCTGCG	420
GGCAGCACAG ATGCTGGCCG CCGAGTGGGA TGTCGCCGCC GACGTGTGGT CGGTGACCAG	480
TTGGGGCGAG CTAAACCGCG ACGGGGTGGT CATCGAGACC GAGAAGCTCC GCCACCCCGA	540
TCGGCCGGCG GGCGTGCCCT ACGTGACGAG AGCGCTGGAG AATGCTCGGG GCCCGGTGAT	600
CGCGGTGTCT GACTGGATGC GCGCGGTCCC CGAGCAGATC CGACCGTGGG TGCCGGGCAC	660
ATACCTCACG TTGGGCACCG ACGGGTTCGG TTTTCCGAC ACTCGGCCCC CCGGTGCTCG	720
TTACTTCAAC ACCGACGCCG AATCCCAGGT TGGTCGCGGT TTTGGGAGGG GTTGGCCGGG	780
TCGACGGGTG AATATCGACC CATTGCGTGC CGGTGCTGGG CCGCCCGCCC AGTTACCCGG	840
ATTGACGAA GGTGGGGGGT TGCGCCCGAN TAAGTT	876

(2) SEQ ID NO: 20 的信息:

(i) 序列特征:

(A) 长度: 1021 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 20:

ATCCCCCGG GCTGCAGGAA TTCGGCACGA GAGACAAAAT TCCACGCGTT AATGCAGGAA	60
CAGATTCATA ACGAATTCAC AGCGGCACAA CAATATGTCG CGATCGCGGT TTATTTGAC	120
AGCGAAGACC TGCCGCAGTT GGCGAAGCAT TTTTACAGCC AAGCGGTCGA GGAACGAAAC	180
CATGCAATGA TGCTCGTGCA ACACCTGCTC GACCGCGACC TTCGTGTCGA AATTCCCGGC	240
GTAGACACGG TGCGAAACCA GTTCGACAGA CCCCOCGAGG CACTGGCGCT GGCCTCGAT	300
CAGGAACGCA CAGTCACCGA CCAGGTCGGT CGGCTGACAG CGGTGGCCCG CGACGAGGGC	360
GATTTCTCG GCGAGCAGTT CATGCAGTGG TTCTTGACAG AACAGATCGA AGAGGTGGCC	420
TTGATGGCAA CCCTGGTGCG GGTTGCCGAT CGGGCCGGGG CCAACCTGTT CGAGCTAGAG	480
AACTTCGTCG CACGTGAAGT GGATGTGGCG CCGGCCGCAT CAGGCGCCCC GCACGCTGCC	540
GGGGGCCGCC TCTAGATCCC TGGGGGGGAT CAGCGAGTGG TCCCGTTCGC CCGCCCGTCT	600
TCCAGCCAGG CCTTGGTGCG GCCGGGGTGG TGAGTACCAA TCCAGGCCAC CCCGACCTCC	660
CGGNAAAAGT CGATGTCCTC GTACTCATCG ACGTTCCAGG AGTACACCGC CCGGCCCTGA	720
GCTGCCGAGC GGTCAACGAG TTGCGGATAT TCCTTTAAG CAGGCAGTGA GGGTCCCACG	780
GCGGTTGGCC CGACCGCCGT GGCCGCACTG CTGGTCAGGT ATCGGGGGGT CTTGGCGAGC	840
AACAACGTCG GCAGGAGGGG TGGAGCCCGC CGGATCCGCA GACCGGGGGG GCGAAAACGA	900
CATCAACACC GCACGGGATC GATCTGCGGA GGGGGGTGCG GGAATACCGA ACCGGTGTAG	960
GAGCGCCAGC AGTTGTTTTT CCACCAGCGA AGCGTTTTCG GGTTCATCGGN GGCNNTTAAG	1020
T	1021

(2) SEQ ID NO: 21 的信息:

(i) 序列特征:

(A) 长度: 321 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 21:

```

CGTGCCGACG AACGGAAGAA CACAACCATG AAGATGGTGA AATCGATCGC CGCAGGTCTG      60
ACCGCCGCGG CTGCAATCGG CGCCGCTGCG GCCGGTGTGA CTTGATCAT GGCTGGCGGN      120
CCGGTCGTAT ACCAGATGCA GCCGGTCGTC TTCGGCGCGC CACTGCCGTT GGACCCGGNA      180
TCCGCCCTG ANGTCCTGAC CGCCGCCAG TGGACCAGNC TGCTCAACAG NCTCGNCGAT      240
CCCAACGTGT CGTTTGNGAA CAAGGGNAGT CTGGTCGAGG GNGGNATCGG NGGNANCGAG      300
GGNGNGNATC GNCGANACA A                                          321

```

(2) SEQ ID NO: 22 的信息:

(i) 序列特征:

(A) 长度: 373 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 22:

```

TCTTATCGGT TCCGGTTGGC GACGGGTTTT GGGNGCGGGT GGTTAACCG CTCGCCAGC      60
CGATCGACGG GCGCGGAGAC GTCGACTCCG ATACTCGGCG CGCGCTGGAG CTCCAGGCGC      120
CCTCGGTGGT GNACCGGCAA GCGGTGAAGG AGCCGTTGNA GACCGGGATC AAGGCGATTG      180

```

ACGCGATGAC CCCGATCGGC CGCGGGCAGC GCCAGCTGAT CATCGGGGAC CGCAAGACCG	240
GCAAAAACCG CCGTCTGTGT CGGACACCAT CCTCAAACCA GCGGGAAGAA CTGGGAGTCC	300
GGTGGATCCC AAGAAGCAGG TGCCTTGTG TATACGTTGG CCATCGGGCA AGAAGGGGAA	360
CTTACCATCG CCG	373

(2) SEQ ID NO: 23 的信息:

(i) 序列特征:

- (A) 长度: 352 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 23:

GTGACGCCGT GATGGGATTC CTGGGCGGGG CCGGTCCGCT GGCGGTGGTG GATCAGCAAC	60
TGGTTACCCG GGTGCCGCAA GGCTGGTCGT TTGCTCAGGC AGCCGCTGTG CCGGTGGTGT	120
TCTTGACGGC CTGGTACGGG TTGGCCGATT TAGCCGAGAT CAAGGCGGGC GAATCGGTGC	180
TGATCCATGC CCGTACCGGC GGTGTGGGCA TGGCGGCTGT GCAGCTGGCT CGCCAGTGGG	240
GCGTGGAGGT TTTCGTCACC GCCAGCCGTG GNAAGTGGGA CACGCTGCGC GCCATNGNGT	300
TTGACGACGA NCCATATCGG NGATTCCNC ACATNCGAAG TTCCGANGGA GA	352

(2) SEQ ID NO: 24 的信息:

(i) 序列特征:

- (A) 长度: 726 个碱基对
- (B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 24:

GAAATCCGCG TTCATTCCGT TCGACCAGCG GCTGGCGATA ATCGACGAAG TGATCAAGCC	60
GCGGTTCGCG GCGCTCATGG GTCACAGCGA GTAATCAGCA AGTTCTCTGG TATATCGCAC	120
CTAGCGTCCA GTTGCTTGCC AGATCGCTTT CGTACCGTCA TCGCATGTAC CGGTTGCGGT	180
GCCGCACGCT CATGCTGGCG GCGTGCATCC TGGCCACGGG TGTGGCGGGT CTCGGGGTCG	240
GCGCGCAGTC CGCAGCCCAA ACCGCGCCGG TGCCCGACTA CTACTGGTGC CCGGGGCAGC	300
CTTTGACCCC CGCATGGGGG CCCAACTGGG ATCCCTACAC CTGCCATGAC GACTTCCACC	360
GCGACAGCGA CGGCCCCGAC CACAGCCGCG ACTACCCCGG ACCCATCCTC GAAGGTCCCG	420
TGCTTGACGA TCCCGGTGCT GCGCCGCCGC CCCC GGCTGC CGGTGGCGGC GCATAGCGCT	480
CGTTGACCGG GCCGCATCAG CGAATACGCG TATAAACCCG GCGGTGCCCC CGGCAAGCTA	540
CGACCCCCCG CGGGGCAGAT TTACGCTCCC GTGCCGATGG ATCGCGCCGT CCGATGACAG	600
AAAATAGGCG ACGGTTTTTG CAACCGCTTG GAGGACGCTT GAAGGGAACC TGTCATGAAC	660
GGCGACAGCG CCTCCACCAT CGACATCGAC AAGGTTGTTA CCCGCACACC CGTTCGCCGG	720
ATCGTG	726

(2) SEQ ID NO: 25 的信息:

(i) 序列特征:

(A) 长度: 580 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 25:

CGCGACGACG ACGAACGTCG GGCCACCAC CGCCTATGCG TTGATGCAGG CGACCGGGAT	60
GGTCGCCGAC CATATCCAAG CATGCTGGGT GCCCACTGAG CGACCTTTTG ACCAGCCGGG	120
CTGCCCCGATG GCGGCCCGGT GAAGTCATTG CGCCGGGGCT TGTGCACCTG ATGAACCCGA	180
ATAGGGAACA ATAGGGGGGT GATTTGGCAG TTCAATGTCG GGTATGGCTG GAAATCCAAT	240
GGCGGGGCAT GCTCGGCGCC GACCAGGCTC GCGCAGGCGG GCCAGCCCGA ATCTGGAGGG	300
AGCACTCAAT GCGGCGGATG AAGCCCCGGA CCGGCGACGG TCCTTTGGAA GCAACTAAGG	360
AGGGGCGCGG CATTGTGATG CGAGTACCAC TTGAGGGTGG CGGTCGCCTG GTCGTCGAGC	420
TGACACCCGA CGAAGCCGCC GCACTGGGTG ACGAACTCAA AGGCGTTACT AGCTAAGACC	480
AGCCCAACGG CGAATGGTCG GCGTTACGCG CACACCTTCC GGTAGATGTC CAGTGTCTGC	540
TCGGCGATGT ATGCCCAGGA GAACTCTTGG ATACAGCGCT	580

(2) SEQ ID NO: 26 的信息:

(i) 序列特征:

- (A) 长度: 160 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 26:

AACGGAGGCG CCGGGGGTTT TGGCGGGGCC GGGGCGGTGCG GCGGCAACGG CGGGGCCGGC	60
GGTACCGCCG GGTGTTCGG TGTCGGCGGG GCCGGTGGGG CCGGAGGCAA CGGCATCGCC	120
GGTGTACGCG GTACGTCGGC CAGCACACCG GGTGGATCCG	160

(2) SEQ ID NO: 27 的信息:

(i) 序列特征:

(A) 长度: 272 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 27:

GACACCGATA CGATGGTGAT GTACGCCAAC GTTGTCGACA CGCTCGAGGC GTTCACGATC	60
CAGCGCACAC CCGACGGCGT GACCATCGGC GATGCGGCCC CGTTCGCGGA GGCGGCTGCC	120
AAGGCGATGG GAATCGACAA GCTGCGGGTA ATTCATACCG GAATGGACCC CGTCGTCGCT	180
GAACGCGAAC AGTGGGACGA CGGCAACAAC ACGTTGGCGT TGGCGCCCGG TGTCGTTGTC	240
GCCTACGAGC GCAACGTACA GACCAACGCC CG	272

(2) SEQ ID NO: 28 的信息:

(i) 序列特征:

(A) 长度: 317 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 28:

GCAGCCGGTG GTTCTCGGAC TATCTGCGCA CGGTGACGCA GCGCGACGTG CGCGAGCTGA	60
AGCGGATCGA GCAGACGGAT CGCCTGCCGC GGTTCATGCG CTACCTGGCC GCTATCACCG	120
CGCAGGAGCT GAACGTGGCC GAAGCGGCGC GGGTCATCGG GGTGACGCG GGGACGATCC	180

GTTCTGGATCT GGCCTGGTTC GAGACGGTCT ATCTGGTACA TCGCCTGCCC GCCTGGTCGC	240
GGAATCTGAC CGCGAAGATC AAGAAGCGGT CAAAGATCCA CGTCGTCGAC AGTGGCTTCG	300
CGGCCTGGTT GCGCGGG	317

(2) SEQ ID NO: 29 的信息:

(i) 序列特征:

- (A) 长度: 182 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 29:

GATCGTGGAG CTGTCGATGA ACAGCGTTGC CGGACGCGCG GCGGCCAGCA CGTCGGTGTA	60
GCAGCGCCGG ACCACCTCGC CGGTGGGCAG CATGGTGATG ACCACGTCGG CCTCGGCCAC	120
CGCTTCGGGC GCGCTACGAA ACACCGCGAC ACCGTGCGCG GCGGCGCCGG ACGCCGCCGT	180
GG	182

(2) SEQ ID NO: 30 的信息:

(i) 序列特征:

- (A) 长度: 308 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 30:

GATCGCGAAG TTTGGTGAGC AGGTGGTCGA CGCGAAAGTC TGGGCGCCTG CGAAGCGGGT	60
CGGCGTTCAC GAGGCGAAGA CACGCCTGTC CGAGCTGCTG CGGCTCGTCT ACGGCGGGCA	120
GAGGTTGAGA TTGCCCGCCG CGGCGAGCCG GTAGCAAAGC TTGTGCCGCT GCATCCTCAT	180
GAGACTCGGC GGTTAGGCAT TGACCATGGC GTGTACCGCG TGCCCGACGA TTTGGACGCT	240
CCGTTGTCAG ACGACGTGCT CGAACGCTTT CACCGGTGAA GCGCTACCTC ATCGACACCC	300
ACGTTTGG	308

(2) SEQ ID NO: 31 的信息:

(i) 序列特征:

- (A) 长度: 267 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 31:

CCGACGACGA GCAACTCACG TGGATGATGG TCGGCAGCGG CATTGAGGAC GGAGAGAATC	60
CGGCCGAAGC TGCCGCGCGG CAAGTGCTCA TAGTGACCGG CCGTAGAGGG CTCCCCGAT	120
GGCACCGGAC TATTCTGGTG TGCCGCTGGC CGGTAAGAGC GGGTAAAAGA ATGTGAGGGG	180
ACACGATGAG CAATCACACC TACCGAGTGA TCGAGATCGT CGGGACCTCG CCCGACGGCG	240
TCGACGCGGC AATCCAGGGC GGTCTGG	267

(2) SEQ ID NO: 32 的信息:

(i) 序列特征:

- (A) 长度: 189 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 32:

CTCGTGCCGA AAGAATGTGA GGGGACACGA TGAGCAATCA CACCTACCGA GTGATCGAGA	60
TCGTGCGGGAC CTCGCCCCGAC GCGGTCGACG CGGCAATCCA GGGCGGTCTG GCCCGAGCTG	120
CGCAGACCAT GCGCGCGCTG GACTGGTTCG AAGTACAGTC AATTCGAGGC CACCTGGTCG	180
ACGGAGCGG	189

(2) SEQ ID NO: 33 的信息:

(i) 序列特征:

(A) 长度: 851 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 33:

CTGCAGGGTG GCGTGGATGA GCGTCACCGC GGGGCAGGCC GAGCTGACCG CCGCCCAGGT	60
CCGGGTTGCT GCGGCGGCCT ACGAGACGGC GTATGGGCTG ACGGTGCCCC CGCCGGTGAT	120
CGCCGAGAAC CGTGCTGAAC TGATGATTCT GATAGCGACC AACCTCTTGG GGCAAAACAC	180
CCCGGCGATC GCGGTCAACG AGGCCGAATA CGGCGAGATG TGGGCCCAAG ACGCCGCCGC	240
GATGTTTGGC TACGCCGCGG CGACGGCGAC GGCACGGCG ACGTTGCTGC CGTTCGAGGA	300
GGCGCCGGAG ATGACCAGCG CGGGTGGGCT CCTCGAGCAG GCCGCCGCGG TCGAGGAGGC	360
CTCCGACACC GCCGCGGCGA ACCAGTTGAT GAACAATGTG CCCCAGGCGC TGAAACAGTT	420
GGCCCAGCCC ACGCAGGGCA CCACGCCTTC TTCCAAGCTG GGTGGCCTGT GGAAGACGGT	480
CTCGCCGCAT CGGTGCGCGA TCAGCAACAT GGTGTGATG GCCAACAACC ACATGTCGAT	540
GACCAACTCG GGTGTGTGCA TGACCAACAC CTTGAGCTCG ATGTTGAAGG GCTTTGCTCC	600

GGCGGCGGCC GCCCAGGCCG TGCAAACCGC GGCACAAAAC GGGGTCCGGG CGATGAGCTC	660
GCTGGGCAGC TCGCTGGGTT CTTCCGGTCT GGGCGGTGGG GTGGCCGCCA ACTTGGGTCG	720
GGCGGCCTCG GTACGGTATG GTCACCGGGA TGGCGGAAAA TATGCANAGT CTGGTCGGCG	780
GAACGGTGGT CCGGCGTAAG GTTTACCCCC GTTTTCTGGA TCGGTGAAC TTCGTCAACG	840
GAAACAGTTA C	851

(2) SEQ ID NO: 34 的信息:

(i) 序列特征:

- (A) 长度: 254 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 34:

GATCGATCGG GCGGAAATTT GGACCAGATT CGCCTCCGGC GATAACCCAA TCAATCGAAC	60
CTAGATTTAT TCCGTCCAGG GGCCCAGTA ATGGCTCGCA GGAGAGGAAC CTTACTGCTG	120
CGGGCACCTG TCGTAGGTCC TCGATACGGC GGAAGGCGTC GACATTTTCC ACCGACACCC	180
CCATCCAAAC GTTCGAGGGC CACTCCAGCT TGTGAGCGAG GCGACGCAGT CGCAGGCTGC	240
GCTTGGTCAA GATC	254

(2) SEQ ID NO: 35 的信息:

(i) 序列特征:

- (A) 长度: 408 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链

(D) 拓扑结构: 线型**(xi) 序列描述: SEQ ID NO: 35:**

CGGCACGAGG ATCCTGACCG AAGCGGCCGC CGCCAAGGCG AAGTCGCTGT TGGACCAGGA	60
GGGACGGGAC GATCTGGCGC TCGGATCGC GGTTCAGCCG GGGGGGTGCG CTGGATTGCG	120
CTATAACCTT TTCTTCGACG ACCGGACGCT GGATGGTGAC CAAACCGCGG AGTTCGGTGG	180
TGTCAGGTTG ATCGTGGACC GGATGAGCGC GCCGTATGTG GAAGGCGCGT CGATCGATTT	240
CGTCGACACT ATTGAGAAGC AAGGNTTCAC CATCGACAAT CCCAACGCCA CCGGCTCCTG	300
CGCGTGCGGG GATTCGTTCA ACTGATAAAA CGCTAGTACG ACCCCGCGGT GCGCAACACG	360
TACGAGCACA CCAAGACCTG ACCGCGCTGG AAAAGCAACT GAGCGATG	408

(2) SEQ ID NO: 36 的信息:**(i) 序列特征:****(A) 长度: 181 个碱基对****(B) 类型: 核酸****(C) 链型: 单链****(D) 拓扑结构: 线型****(xi) 序列描述: SEQ ID NO: 36:**

GCGGTGTCGG CGGATCCGGC GGGTGGTTGA ACGGCAACGG CGGGGCCGGC GGGGCCGGCG	60
GGACCGGCGC TAACGGTGGT GCCGGCGGCA ACGCCTGGTT GTTCGGGGCC GGCGGGTCCG	120
GCGGNGCCGG CACCAATGGT GGNGTCGGCG GGTCCGGCGG ATTTGTCTAC GGCAACGGCG	180
G	181

(2) SEQ ID NO: 37 的信息:

(i) 序列特征:

- (A) 长度: 290 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 37:

```

GCGGTGTCGG CGGATCCGGC GGGTGGTTGA ACGGCAACGG CGGTGTCGGC GGCCGGGGCG      60
GCGACGGCGT CTTTGCCGGT GCCGGCGGCC AGGGCGGCCT CGGTGGGCAG GGCGGCAATG      120
GCGGCGGCTC CACCGGCGGC AACGGCGGTC TTGGCGGCGC GGGCGGTGGC GGAGGCAACG      180
CCCCGGACGG CGGCTTCGGT GGCAACGGCG GTAAGGGTGG CCAGGGCGGN ATTGGCGGCG      240
GCACTCAGAG CGCGACCGGC CTCGGNGGTG ACGGCGGTGA CGGCGGTGAC      290

```

(2) SEQ ID NO: 38 的信息:

(i) 序列特征:

- (A) 长度: 34 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 38:

```

GATCCAGTGG CATGGNGGGT GTCAGTGGAA GCAT      34

```

(2) SEQ ID NO: 39 的信息:

(i) 序列特征:

- (A) 长度: 155 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 39:

GATCGCTGCT CGTCCCCCCC TTGCCGCCGA CGCCACCGGT CCCACCGTTA CCGAACAAGC	60
TGGCGTGGTC GCCAGCACCC CCGGCACCGC CGACGCCGGA GTCGAACAAT GGCACCGTCG	120
TATCCCCACC ATTGCCGCCG GNCCCACCGG CACCG	155

(2) SEQ ID NO: 40 的信息:

(i) 序列特征:

- (A) 长度: 53 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 40:

ATGGCGTTCA CGGGGCGCCG GGGACCGGC AGCCCGGNGG GGCCGGGGGG TGG	53
---	----

(2) SEQ ID NO: 41 的信息:

(i) 序列特征:

- (A) 长度: 132 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 41:

GATCCACCGC GGGTGCAGAC GGTGCCCGCG GCGCCACCCC GACCAGCGGC GGCAACGGCG	60
GCACCGGCGG CAACGGCGCG AACGCCACCG TCGTCGGNGG GGCCGGCGGG GCCGGCGGCA	120
AGGGCGGCAA CG	132

(2) SEQ ID NO: 42 的信息:

(i) 序列特征:

(A) 长度: 132 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 42:

GATCGGCGGC CGGNACGGNC GGGGACGGCG GCAAGGGCGG NAACGGGGGC GCCGNAGCCA	60
CCNGCCAAGA ATCCTCCGNG TCCNCCAATG GCGCGAATGG CGGACAGGGC GGCAACGGCG	120
GCANCGGCGG CA	132

(2) SEQ ID NO: 43 的信息:

(i) 序列特征:

(A) 长度: 702 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 43:

CGGCACGAGG ATCGGTACCC CGCGGCATCG GCAGCTGCCG ATTCGCCGGG TTTCCCACC	60
CGAGGAAAGC CGCTACCAGA TGGCGCTGCC GAAGTAGGGC GATCCGTTCC CGATGCCGGC	120
ATGAACGGGC GGCATCAAAT TAGTGAGGA ACCTTTCAGT TTAGCGACGA TAATGGCTAT	180
AGCACTAAGG AGGATGATCC GATATGACGC AGTCGCAGAC CGTGACGGTG GATCAGCAAG	240
AGATTTTGAA CAGGGCCAAC GAGGTGGAGG CCCCAGTGGC GGACCCACCG ACTGATGTCC	300
CCATCACACC GTGCGAACTC ACGGNGGNTA AAAACGCCGC CCAACAGNTG GTNTTGTCCG	360
CCGACAACAT GCGGGAATAC CTGGCGGCCG GTGCCAAAGA GCGGCAGCGT CTGGCGACCT	420
CGCTGCGCAA CGCGGCCAAG GNGTATGGCG AGGTTGATGA GGAGGCTGCG ACCGCGCTGG	480

ACAACGACGG CGAAGGAACT GTGCAGGCAG AATCGGCCGG GGCCGTCGGA GGGGACAGTT	540
CGGCCGAACT AACCGATACG CCGAGGGTGG CCACGGCCGG TGAACCCAAC TTCATGGATC	600
TCAAAGAAGC GGCAAGGAAG CTCGAAACGG GCGACCAAGG CGCATCGCTC GCGCACTGNG	660
GGGATGGGTG GAACACTTNC ACCCTGACGC TGCAAGGCCA CG	702

(2) SEQ ID NO: 44 的信息:

(i) 序列特征:

- (A) 长度: 298 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 44:

GAAGCCGCAG CGCTGTCGGG CGACGTGGCG GTCAAAGCGG CATCGCTCGG TGGCGGTGGA	60
GGCGGCGGGG TGCCGTCGGC GCCGTTGGGA TCCGCGATCG GGGGCGCCGA ATCGGTGCGG	120
CCCCTGGCG CTGGTGACAT TGCCGGCTTA GGCCAGGGAA GGGCCGGCGG CGGCGCCGCG	180
CTGGGCGGCG GTGGCATGGG AATGCCGATG GGTGCCGCGC ATCAGGGACA AGGGGGCGCC	240
AAGTCCAAGG GTTCTCAGCA GGAAGACGAG GCGCTCTACA CCGAGGATCC TCGTGCCG	298

(2) SEQ ID NO: 45 的信息:

(i) 序列特征:

- (A) 长度: 1058 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 45:

CGGCACGAGG ATCGAATCGC GTCGCCGGGA GCACAGCGTC GCACTGCACC AGTGGAGGAG	60
CCATGACCTA CTCGCCGGGT AACCCCGGAT ACCCGCAAGC GCAGCCCGCA GGCTCCTACG	120
GAGGCGTCAC ACCCTCGTTC GCCCAGCCG ATGAGGGTGC GAGCAAGCTA CCGATGTACC	180
TGAACATCGC GGTGGCAGTG CTCGGTCTGG CTGCGTACTT CGCCAGCTTC GGCCCAATGT	240
TCACCCTCAG TACCGAACTC GGGGGGGGTG ATGGCGCAGT GTCCGGTGAC ACTGGGCTGC	300
CGGTGCGGGT GGCTCTGCTG GCTGCGCTGC TTGCCGGGGT GGTTCCTGGT CCTAAGGCCA	360
AGAGCCATGT GACGGTAGTT GCGGTGCTCG GGGTACTCGG CGTATTTCTG ATGGTCTCGG	420
CGACGTTTAA CAAGCCCAGC GCCTATTGGA CCGGTTGGGC ATTGTGGGTT GTGTTGGCTT	480
TCATCGTGTT CCAGGCGGTT GCGGCAGTCC TGGCGCTCTT GGTGGAGACC GCGCTATCA	540
CCGCGCCGGC GCCGCGGCC AAGTTCGACC CGTATGGACA GTACGGGCGG TACGGGCAGT	600
ACGGGCAGTA CGGGGTGCAG CCGGGTGGT ACTACGGTCA GCAGGGTGCT CAGCAGGCCG	660
CGGGA CTGCA GTCGCCCGGC CCGCAGCAGT CTCCGCAGCC TCCCGGATAT GGGTCGCAGT	720
ACGGCGGCTA TTCGTCCAGT CCGAGCCAAT CGGGCAGTGG ATACACTGCT CAGCCCCCGG	780
CCCAGCCGCC GCGGCAGTCC GGGTCGCAAC AATCGCACCA GGGCCCATCC ACGCCACCTA	840
CCGGCTTTCC GAGCTTCAGC CCACCACCAC CGGTCAGTGC CGGGACGGGG TCGCAGGCTG	900
GTTCCGGCTCC AGTCAACTAT TCAAACCCCA GCGGGGGCGA GCAGTCGTG TCCCCGGGG	960
GGGCGCCGGT CTAACCGGGC GTTCCCGCGT CCGGTCGCGC GTGTGCGCGA AGAGTGAACA	1020
GGGTGTCAGC AAGCGCGGAC GATCCTCGTG CCGAATTC	1058

(2) SEQ ID NO: 46 的信息:

(i) 序列特征:

(A) 长度: 327 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 46:

CGGCACGAGA GACCGATGCC GCTACCCTCG CGCAGGAGGC AGGTAATTC GAGCGGATCT	60
CCGGCGACCT GAAAACCCAG ATCGACCAGG TGGAGTCGAC GGCAGGTTCTG TTGCAGGGCC	120
AGTGGCGCGG CGCGGCGGGG ACGGCCGCCC AGGCCGCGGT GGTGCGCTTC CAAGAAGCAG	180
CCAATAAGCA GAAGCAGGAA CTCGACGAGA TCTCGACGAA TATTCGTCAG GCCGGCGTCC	240
AATACTCGAG GGCCGACGAG GAGCAGCAGC AGGCGCTGTC CTCGCAAATG GGCTTCTGAC	300
CCGCTAATAC GAAAAGAAAC GGAGCAA	327

(2) SEQ ID NO: 47 的信息:

(i) 序列特征:

(A) 长度: 170 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 47:

CGGTCGCGAT GATGGCGTTG TCGAACGTGA CCGATTCTGT ACCGCCGTCG TTGAGATCAA	60
CCAACAACGT GTTGGCGTCG GCAAATGTGC CGNACCCGTG GATCTCGGTG ATCTTGTTCT	120
TCTTCATCAG GAAGTCACA CCGGCCACCC TGCCCTCGGN TACCTTTCGG	170

(2) SEQ ID NO: 48 的信息:

(i) 序列特征:

- (A) 长度: 127 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 48:

GATCCGGCGG CACGGGGGGT GCCGGCGGCA GCACCGCTGG CGCTGGCGGC AACGGCGGGG	60
CCGGGGGTGG CGGCGGAACC GGTGGGTTGC TCTTCGGCAA CGGCGGTGCC GGCGGGCACG	120
GGGCCGT	127

(2) SEQ ID NO: 49 的信息:

(i) 序列特征:

- (A) 长度: 81 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 49:

CGGCGGCAAG GGCGGCACCG CCGGCAACGG GAGCGGCGCG GCCGGCGGCA ACGGCGGCAA	60
CGGCGGCTCC GGCCTCAACG G	81

(2) SEQ ID NO: 50 的信息:

(i) 序列特征:

- (A) 长度: 149 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 50:

GATCAGGGCT GGCCGGCTCC GGCCAGAAGG GCGGTAACGG AGGAGCTGCC GGATTGTTTG	60
GCAACGGCGG GGCCGGNGGT GCCGGCGCGT CCAACCAAGC CGGTAACGGC GGNGCCGGCG	120
GAAACGGTGG TGCCGGTGGG CTGATCTGG	149

(2) SEQ ID NO: 51 的信息:

(i) 序列特征:

- (A) 长度: 355 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 51:

CGGCACGAGA TCACACCTAC CGAGTGATCG AGATCGTCGG GACCTCGCCC GACGGTGTCG	60
ACGCGGNAAT CCAGGGCGGT CTGGCCCGAG CTGCGCAGAC CATGCGCGCG CTGGACTGGT	120
TCGAAGTACA GTCAATTCGA GGCCACCTGG TCGACGGAGC GGTCGCGCAC TTCCAGGTGA	180
CTATGAAAGT CGGCTTCCGC CTGGAGGATT CCTGAACCTT CAAGCGCGGC CGATAACTGA	240
GGTGCATCAT TAAGCGACTT TTCCAGAACA TCCTGACGCG CTCGAAACGC GGTTCAGCCG	300
ACGGTGGCTC CGCCGAGGCG CTGCCTCCAA AATCCCTGCG ACAATTCGTC GGCGG	355

(2) SEQ ID NO: 52 的信息:

(i) 序列特征:

- (A) 长度: 999 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 52:

ATGCATCACC ATCACCATCA CATGCATCAG GTGGACCCCA ACTTGACACG TCGCAAGGGA	60
CGATTGGCGG CACTGGCTAT CGCGGCGATG GCCAGCGCCA GCCTGGTGAC CGTTGCGGTG	120
CCCGCGACCG CCAACGCCGA TCCGGAGCCA GCGCCCCCGG TACCCACAAC GGCCGCCTCG	180
CCGCCGTGCA CCGCTGCAGC GCCACCCGCA CCGGCGACAC CTGTTGCCCC CCCACCACCG	240
GCCGCCGCCA ACACGCCGAA TGCCCAGCCG GCGGATCCCA ACGCAGCACC TCCGCCGGCC	300
GACCCGAACG CACCGCCGCC ACCTGTCATT GCCCCAAACG CACCCCAACC TGTCCGGATC	360
GACAACCCGG TTGGAGGATT CAGCTTCGCG CTGCCTGCTG GCTGGGTGGA GTCTGACGCC	420
GCCCCATTCT ACTACGGTTC AGCACTCCTC AGCAAAACCA CCGGGGACCC GCCATTTCCC	480
GGACAGCCGC CGCCGGTGGC CAATGACACC CGTATCGTGC TCGGCCGGCT AGACCAAAAG	540
CTTTACGCCA GCGCCGAAGC CACCGACTCC AAGGCCGCGG CCCGGTTGGG CTCGGACATG	600
GGTGAGTTCT ATATGCCCTA CCCGGGCACC CGGATCAACC AGGAAACCGT CTCGCTCGAC	660
GCCAACGGGG TGTCTGGAAG CGCGTCGTAT TACGAAGTCA AGTTCAGCGA TCCGAGTAAG	720
CCGAACGGCC AGATCTGGAC GGGCGTAATC GGCTCGCCCG CGGCGAACGC ACCGGACGCC	780
GGGCCCCCTC AGCGCTGGTT TGTGGTATGG CTCGGGACCG CCAACAACCC GGTGGACAAG	840
GGCGCGGCCA AGGCGCTGGC CGAATCGATC CGGCCTTTGG TCGCCCCGCC GCCGGCGCCG	900
GCACCGGCTC CTGCAGAGCC CGCTCCGGCG CCGGCGCCGG CCGGGGAAGT CGCTCCTACC	960
CCGACGACAC CGACACCGCA GCGGACCTTA CCGGCCTGA	999

(2) SEQ ID NO: 53 的信息:**(i) 序列特征:**

(A) 长度: 332 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 53:

```

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

```

```

Gly Thr Arg Ile Asn Gln Glu Thr Val Ser Leu Asp Ala Asn Gly Val
 210                      215                      220

Ser Gly Ser Ala Ser Tyr Tyr Glu Val Lys Phe Ser Asp Pro Ser Lys
 225                      230                      235                      240

Pro Asn Gly Gln Ile Trp Thr Gly Val Ile Gly Ser Pro Ala Ala Asn
                245                      250                      255

Ala Pro Asp Ala Gly Pro Pro Gln Arg Trp Phe Val Val Trp Leu Gly
                260                      265                      270

Thr Ala Asn Asn Pro Val Asp Lys Gly Ala Ala Lys Ala Leu Ala Glu
 275                      280                      285

Ser Ile Arg Pro Leu Val Ala Pro Pro Pro Ala Pro Ala Pro Ala Pro
 290                      295                      300

Ala Glu Pro Ala Pro Ala Pro Ala Pro Ala Gly Glu Val Ala Pro Thr
 305                      310                      315                      320

Pro Thr Thr Pro Thr Pro Gln Arg Thr Leu Pro Ala
                325                      330

```

(2) SEQ ID NO: 54 的信息:

(i) 序列特征:

- (A) 长度: 20 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 54:

```

Asp Pro Val Asp Ala Val Ile Asn Thr Thr Xaa Asn Tyr Gly Gln Val
1           5           10           15
Val Ala Ala Leu
           20

```

(2) SEQ ID NO: 55 的信息:

(i) 序列特征:

(A) 长度: 15 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 55:

```

Ala Val Glu Ser Gly Met Leu Ala Leu Gly Thr Pro Ala Pro Ser
1           5           10           15

```

(2) SEQ ID NO: 56 的信息:

(i) 序列特征:

(A) 长度: 19 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 56:

```

Ala Ala Met Lys Pro Arg Thr Gly Asp Gly Pro Leu Glu Ala Ala Lys
1           5           10           15
Glu Gly Arg

```

(2) SEQ ID NO: 57 的信息:

(i) 序列特征:

(A) 长度: 15 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 57:

Tyr	Tyr	Trp	Cys	Pro	Gly	Gln	Pro	Phe	Asp	Pro	Ala	Trp	Gly	Pro
1				5					10				15	

(2) SEQ ID NO: 58 的信息:

(i) 序列特征:

(A) 长度: 14 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 58:

Asp	Ile	Gly	Ser	Glu	Ser	Thr	Glu	Asp	Gln	Gln	Xaa	Ala	Val
1				5				10					

(2) SEQ ID NO: 59 的信息:

(i) 序列特征:

(A) 长度: 13 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 59:

Ala Glu Glu Ser Ile Ser Thr Xaa Glu Xaa Ile Val Pro
1 5 10

(2) SEQ ID NO: 60 的信息:

(i) 序列特征:

(A) 长度: 17 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 60:

Asp Pro Glu Pro Ala Pro Pro Val Pro Thr Ala Ala Ala Ala Pro Pro
1 5 10 15

Ala

(2) SEQ ID NO: 61 的信息:

(i) 序列特征:

(A) 长度: 15 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 61:

Ala Pro Lys Thr Tyr Xaa Glu Glu Leu Lys Gly Thr Asp Thr Gly
1 5 10 15

(2) SEQ ID NO: 62 的信息:

(i) 序列特征:

(A) 长度: 30 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 62:

```

Asp Pro Ala Ser Ala Pro Asp Val Pro Thr Ala Ala Gln Gln Thr Ser
1           5           10           15
Leu Leu Asn Asn Leu Ala Asp Pro Asp Val Ser Phe Ala Asp
           20           25           30

```

(2) SEQ ID NO: 63 的信息:

(i) 序列特征:

(A) 长度: 187 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 63:

```

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

```

```

Ile Ala Asp His Lys Leu Lys Lys Ala Ala Glu His Gly Asp Leu Pro
  115                      120                      125

Leu Ser Phe Ser Val Thr Asn Ile Gln Pro Ala Ala Ala Gly Ser Ala
  130                      135                      140

Thr Ala Asp Val Ser Val Ser Gly Pro Lys Leu Ser Ser Pro Val Thr
  145                      150                      155                      160

Gln Asn Val Thr Phe Val Asn Gln Gly Gly Trp Met Leu Ser Arg Ala
                      165                      170                      175

Ser Ala Met Glu Leu Leu Gln Ala Ala Gly Xaa
          180                      185

```

(2) SEQ ID NO: 64 的信息:

(i) 序列特征:

- (A) 长度: 148 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 64:

```

Asp Glu Val Thr Val Glu Thr Thr Ser Val Phe Arg Ala Asp Phe Leu
 1                      5                      10                      15

Ser Glu Leu Asp Ala Pro Ala Gln Ala Gly Thr Glu Ser Ala Val Ser
          20                      25                      30

Gly Val Glu Gly Leu Pro Pro Gly Ser Ala Leu Leu Val Val Lys Arg
          35                      40                      45

Gly Pro Asn Ala Gly Ser Arg Phe Leu Leu Asp Gln Ala Ile Thr Ser
          50                      55                      60

```

Ala Gly Arg His Pro Asp Ser Asp Ile Phe Leu Asp Asp Val Thr Val
 65 70 75 80
 Ser Arg Arg His Ala Glu Phe Arg Leu Glu Asn Asn Glu Phe Asn Val
 85 90 95
 Val Asp Val Gly Ser Leu Asn Gly Thr Tyr Val Asn Arg Glu Pro Val
 100 105 110
 Asp Ser Ala Val Leu Ala Asn Gly Asp Glu Val Gln Ile Gly Lys Leu
 115 120 125
 Arg Leu Val Phe Leu Thr Gly Pro Lys Gln Gly Glu Asp Asp Gly Ser
 130 135 140
 Thr Gly Gly Pro
 145

(2) SEQ ID NO: 65 的信息:

(i) 序列特征:

- (A) 长度: 230 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 65:

Thr Ser Asn Arg Pro Ala Arg Arg Gly Arg Arg Ala Pro Arg Asp Thr
 1 5 10 15
 Gly Pro Asp Arg Ser Ala Ser Leu Ser Leu Val Arg His Arg Arg Gln
 20 25 30
 Gln Arg Asp Ala Leu Cys Leu Ser Ser Thr Gln Ile Ser Arg Gln Ser
 35 40 45
 Asn Leu Pro Pro Ala Ala Gly Gly Ala Ala Asn Tyr Ser Arg Arg Asn
 50 55 60

```

Phe Asp Val Arg Ile Lys Ile Phe Met Leu Val Thr Ala Val Val Leu
65              70              75              80
Leu Cys Cys Ser Gly Val Ala Thr Ala Ala Pro Lys Thr Tyr Cys Glu
85              90              95
Glu Leu Lys Gly Thr Asp Thr Gly Gln Ala Cys Gln Ile Gln Met Ser
100            105            110
Asp Pro Ala Tyr Asn Ile Asn Ile Ser Leu Pro Ser Tyr Tyr Pro Asp
115            120            125
Gln Lys Ser Leu Glu Asn Tyr Ile Ala Gln Thr Arg Asp Lys Phe Leu
130            135            140
Ser Ala Ala Thr Ser Ser Thr Pro Arg Glu Ala Pro Tyr Glu Leu Asn
145            150            155            160
Ile Thr Ser Ala Thr Tyr Gln Ser Ala Ile Pro Pro Arg Gly Thr Gln
165            170            175
Ala Val Val Leu Xaa Val Tyr His Asn Ala Gly Gly Thr His Pro Thr
180            185            190
Thr Thr Tyr Lys Ala Phe Asp Trp Asp Gln Ala Tyr Arg Lys Pro Ile
195            200            205
Thr Tyr Asp Thr Leu Trp Gln Ala Asp Thr Asp Pro Leu Pro Val Val
210            215            220
Phe Pro Ile Val Ala Arg
225            230

```

(2) SEQ ID NO: 66 的信息:

(i) 序列特征:

(A) 长度: 132 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 66:

```

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

```

(2) SEQ ID NO: 67 的信息:

(i) 序列特征:

(A) 长度: 100 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 67:

```

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

```

(2) SEQ ID NO: 68 的信息:

(i) 序列特征:

(A) 长度: 163 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 68:

```

Met Thr Asp Asp Ile Leu Leu Ile Asp Thr Asp Glu Arg Val Arg Thr
1           5           10           15

Leu Thr Leu Asn Arg Pro Gln Ser Arg Asn Ala Leu Ser Ala Ala Leu
          20           25           30

Arg Asp Arg Phe Phe Ala Xaa Leu Xaa Asp Ala Glu Xaa Asp Asp Asp
          35           40           45

Ile Asp Val Val Ile Leu Thr Gly Ala Asp Pro Val Phe Cys Ala Gly
          50           55           60

Leu Asp Leu Lys Val Ala Gly Arg Ala Asp Arg Ala Ala Gly His Leu
65           70           75           80

Thr Ala Val Gly Gly His Asp Gln Ala Gly Asp Arg Arg Asp Gln Arg
          85           90           95

Arg Arg Gly His Arg Arg Ala Arg Thr Gly Ala Val Leu Arg His Pro
          100          105          110

Asp Arg Leu Arg Ala Arg Pro Leu Arg Arg His Pro Arg Pro Gly Gly
          115          120          125

Ala Ala Ala His Leu Gly Thr Gln Cys Val Leu Ala Ala Lys Gly Arg
          130          135          140

His Arg Xaa Gly Pro Val Asp Glu Pro Asp Arg Arg Leu Pro Val Arg
145           150           155           160

Asp Arg Arg

```

(2) SEQ ID NO: 69 的信息:

(i) 序列特征:

(A) 长度: 344 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 69:

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

```

Thr Arg Gln Val Val Arg Arg Val Val Gly Ser Trp His Gly Glu Pro
                245                250                255

Met Pro Met Ser Ser Arg Trp Thr Asn Glu His Thr Ala Glu Leu Pro
                260                265                270

Ala Asp Leu His Ala Pro Thr Arg Leu Ala Leu Leu Thr Gly Leu Ala
                275                280                285

Pro His Gln Val Thr Asp Asp Asp Val Ala Ala Ala Arg Ser Leu Leu
                290                295                300

Asp Thr Asp Ala Ala Leu Val Gly Ala Leu Ala Trp Ala Ala Phe Thr
305                310                315                320

Ala Ala Arg Arg Ile Gly Thr Trp Ile Gly Ala Ala Ala Glu Gly Gln
                325                330                335

Val Ser Arg Gln Asn Pro Thr Gly
                340

```

(2) SEQ ID NO: 70 的信息:

(i) 序列特征:

- (A) 长度: 485 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 70:

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

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

Ser Arg Phe Ala Arg Ser Gly Pro Arg Arg Asn Ala Gln Val Thr Ser
 465 470 475 480

Val Ala Pro Thr Gly
 485

(2) SEQ ID NO: 71 的信息:

(i) 序列特征:

(A) 长度: 267 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 71:

Gly Val Ile Val Leu Asp Leu Glu Pro Arg Gly Pro Leu Pro Thr Glu
 1 5 10 15

Ile Tyr Trp Arg Arg Arg Gly Leu Ala Leu Gly Ile Ala Val Val Val
 20 25 30

Val Gly Ile Ala Val Ala Ile Val Ile Ala Phe Val Asp Ser Ser Ala
 35 40 45

Gly Ala Lys Pro Val Ser Ala Asp Lys Pro Ala Ser Ala Gln Ser His
 50 55 60

Pro Gly Ser Pro Ala Pro Gln Ala Pro Gln Pro Ala Gly Gln Thr Glu
 65 70 75 80

Gly Asn Ala Ala Ala Ala Pro Pro Gln Gly Gln Asn Pro Glu Thr Pro
 85 90 95

Thr Pro Thr Ala Ala Val Gln Pro Pro Pro Val Leu Lys Glu Gly Asp
 100 105 110

Asp Cys Pro Asp Ser Thr Leu Ala Val Lys Gly Leu Thr Asn Ala Pro
 115 120 125

Gln Tyr Tyr Val Gly Asp Gln Pro Lys Phe Thr Met Val Val Thr Asn

```

130              135              140
Ile Gly Leu Val Ser Cys Lys Arg Asp Val Gly Ala Ala Val Leu Ala
145      -      150              155              160

Ala Tyr Val Tyr Ser Leu Asp Asn Lys Arg Leu Trp Ser Asn Leu Asp
      165              170              175

Cys Ala Pro Ser Asn Glu Thr Leu Val Lys Thr Phe Ser Pro Gly Glu
      180              185              190

Gln Val Thr Thr Ala Val Thr Trp Thr Gly Met Gly Ser Ala Pro Arg
      195              200              205

Cys Pro Leu Pro Arg Pro Ala Ile Gly Pro Gly Thr Tyr Asn Leu Val
      210              215              220

Val Gln Leu Gly Asn Leu Arg Ser Leu Pro Val Pro Phe Ile Leu Asn
225              230              235              240

Gln Pro Pro Pro Pro Pro Gly Pro Val Pro Ala Pro Gly Pro Ala Gln
      245              250              255

Ala Pro Pro Pro Glu Ser Pro Ala Gln Gly Gly
      260              265

```

(2) SEQ ID NO: 72 的信息:

(i) 序列特征:

- (A) 长度: 97 个氨基酸
(B) 类型: 氨基酸
(C) 链型: 单链
(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 72:

```

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

```

(2) SEQ ID NO: 73 的信息:

(i) 序列特征:

(A) 长度: 364 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 73:

Gly Ala Ala Val Ser Leu Leu Ala Ala Gly Thr Leu Val Leu Thr Ala
 1 5 10 15
 Cys Gly Gly Gly Thr Asn Ser Ser Ser Ser Gly Ala Gly Gly Thr Ser
 20 25 30
 Gly Ser Val His Cys Gly Gly Lys Lys Glu Leu His Ser Ser Gly Ser
 35 40 45
 Thr Ala Gln Glu Asn Ala Met Glu Gln Phe Val Tyr Ala Tyr Val Arg
 50 55 60
 Ser Cys Pro Gly Tyr Thr Leu Asp Tyr Asn Ala Asn Gly Ser Gly Ala
 65 70 75 80
 Gly Val Thr Gln Phe Leu Asn Asn Glu Thr Asp Phe Ala Gly Ser Asp
 85 90 95
 Val Pro Leu Asn Pro Ser Thr Gly Gln Pro Asp Arg Ser Ala Glu Arg
 100 105 110
 Cys Gly Ser Pro Ala Trp Asp Leu Pro Thr Val Phe Gly Pro Ile Ala
 115 120 125
 Ile Thr Tyr Asn Ile Lys Gly Val Ser Thr Leu Asn Leu Asp Gly Pro
 130 135 140
 Thr Thr Ala Lys Ile Phe Asn Gly Thr Ile Thr Val Trp Asn Asp Pro
 145 150 155 160
 Gln Ile Gln Ala Leu Asn Ser Gly Thr Asp Leu Pro Pro Thr Pro Ile
 165 170 175
 Ser Val Ile Phe Arg Ser Asp Lys Ser Gly Thr Ser Asp Asn Phe Gln
 180 185 190
 Lys Tyr Leu Asp Gly Val Ser Asn Gly Ala Trp Gly Lys Gly Ala Ser
 195 200 205
 Glu Thr Phe Ser Gly Gly Val Gly Val Gly Ala Ser Gly Asn Asn Gly
 210 215 220
 Thr Ser Ala Leu Leu Gln Thr Thr Asp Gly Ser Ile Thr Tyr Asn Glu
 225 230 235 240
 Trp Ser Phe Ala Val Gly Lys Gln Leu Asn Met Ala Gln Ile Ile Thr
 245 250 255

```

Ser Ala Gly Pro Asp Pro Val Ala Ile Thr Thr Glu Ser Val Gly Lys
      260                      265                      270

Thr Ile Ala Gly Ala Lys Ile Met Gly Gln Gly Asn Asp Leu Val Leu
      275                      280                      285

Asp Thr Ser Ser Phe Tyr Arg Pro Thr Gln Pro Gly Ser Tyr Pro Ile
      290                      295                      300

Val Leu Ala Thr Tyr Glu Ile Val Cys Ser Lys Tyr Pro Asp Ala Thr
      305                      310                      315                      320

Thr Gly Thr Ala Val Arg Ala Phe Met Gln Ala Ala Ile Gly Pro Gly
      325                      330                      335

Gln Glu Gly Leu Asp Gln Tyr Gly Ser Ile Pro Leu Pro Lys Ser Phe
      340                      345                      350

Gln Ala Lys Leu Ala Ala Ala Val Asn Ala Ile Ser
      355                      360

```

(2) SEQ ID NO: 74 的信息:

(i) 序列特征:

(A) 长度: 309 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 74:

Gln Ala Ala Ala Gly Arg Ala Val Arg Arg Thr Gly His Ala Glu Asp
 1 5 10 15
 Gln Thr His Gln Asp Arg Leu His His Gly Cys Arg Arg Ala Ala Val
 20 25 30
 Val Val Arg Gln Asp Arg Ala Ser Val Ser Ala Thr Ser Ala Arg Pro
 35 40 45
 Pro Arg Arg His Pro Ala Gln Gly His Arg Arg Arg Val Ala Pro Ser
 50 55 60
 Gly Gly Arg Arg Arg Pro His Pro His His Val Gln Pro Asp Asp Arg
 65 70 75 80
 Arg Asp Arg Pro Ala Leu Leu Asp Arg Thr Gln Pro Ala Glu His Pro
 85 90 95
 Asp Pro His Arg Arg Gly Pro Ala Asp Pro Gly Arg Val Arg Gly Arg
 100 105 110
 Gly Arg Leu Arg Arg Val Asp Asp Gly Arg Leu Gln Pro Asp Arg Asp
 115 120 125
 Ala Asp His Gly Ala Pro Val Arg Gly Arg Gly Pro His Arg Gly Val
 130 135 140
 Gln His Arg Gly Gly Pro Val Phe Val Arg Arg Val Pro Gly Val Arg
 145 150 155 160
 Cys Ala His Arg Arg Gly His Arg Arg Val Ala Ala Pro Gly Gln Gly
 165 170 175
 Asp Val Leu Arg Ala Gly Leu Arg Val Glu Arg Leu Arg Pro Val Ala
 180 185 190
 Ala Val Glu Asn Leu His Arg Gly Ser Gln Arg Ala Asp Gly Arg Val
 195 200 205
 Phe Arg Pro Ile Arg Arg Gly Ala Arg Leu Pro Ala Arg Arg Ser Arg
 210 215 220
 Ala Gly Pro Gln Gly Arg Leu His Leu Asp Gly Ala Gly Pro Ser Pro
 225 230 235 240
 Leu Pro Ala Arg Ala Gly Gln Gln Gln Pro Ser Ser Ala Gly Gly Arg
 245 250 255

Arg Ala Gly Gly Ala Glu Arg Ala Asp Pro Gly Gln Arg Gly Arg His
 260 265 270

His Gln Gly Gly His Asp Pro Gly Arg Gln Gly Ala Gln Arg Gly Thr
 275 280 285

Ala Gly Val Ala His Ala Ala Ala Gly Pro Arg Arg Ala Ala Val Arg
 290 295 300

Asn Arg Pro Arg Arg
 305

(2) SEQ ID NO: 75 的信息:

(i) 序列特征:

- (A) 长度: 580 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 75:

Ser Ala Val Trp Cys Leu Asn Gly Phe Thr Gly Arg His Arg His Gly
 1 5 10 15

Arg Cys Arg Val Arg Ala Ser Gly Trp Arg Ser Ser Asn Arg Trp Cys
 20 25 30

Ser Thr Thr Ala Asp Cys Cys Ala Ser Lys Thr Pro Thr Gln Ala Ala
 35 40 45

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

Arg Phe Arg Ser Phe Pro Val Arg Arg Leu Ala Leu Gly Ala Arg Thr
 65 70 75 80

Ser Arg Thr Leu Gly Val Arg Arg Thr Leu Ser Gln Trp Asn Leu Ser
 85 90 95

Pro Arg Ala Gln Pro Ser Cys Ala Val Thr Val Glu Ser His Thr His
 100 105 110

Ala Ser Pro Arg Met Ala Lys Leu Ala Arg Val Val Gly Leu Val Gln
 115 120 125

Glu Glu Gln Pro Ser Asp Met Thr Asn His Pro Arg Tyr Ser Pro Pro
 130 135 140

Pro Gln Gln Pro Gly Thr Pro Gly Tyr Ala Gln Gly Gln Gln Gln Thr
 145 150 155 160

Tyr Ser Gln Gln Phe Asp Trp Arg Tyr Pro Pro Ser Pro Pro Pro Gln
 165 170 175

Pro Thr Gln Tyr Arg Gln Pro Tyr Glu Ala Leu Gly Gly Thr Arg Pro
 180 185 190

Gly Leu Ile Pro Gly Val Ile Pro Thr Met Thr Pro Pro Pro Gly Met
 195 200 205

Val Arg Gln Arg Pro Arg Ala Gly Met Leu Ala Ile Gly Ala Val Thr
 210 215 220

Ile Ala Val Val Ser Ala Gly Ile Gly Gly Ala Ala Ala Ser Leu Val
 225 230 235 240

Gly Phe Asn Arg Ala Pro Ala Gly Pro Ser Gly Gly Pro Val Ala Ala
 245 250 255

Ser Ala Ala Pro Ser Ile Pro Ala Ala Asn Met Pro Pro Gly Ser Val
 260 265 270

Glu Gln Val Ala Ala Lys Val Val Pro Ser Val Val Met Leu Glu Thr
 275 280 285

Asp Leu Gly Arg Gln Ser Glu Glu Gly Ser Gly Ile Ile Leu Ser Ala
 290 295 300

Glu Gly Leu Ile Leu Thr Asn Asn His Val Ile Ala Ala Ala Ala Lys
 305 310 315 320

Pro Pro Leu Gly Ser Pro Pro Pro Lys Thr Thr Val Thr Phe Ser Asp
 325 330 335

Gly Arg Thr Ala Pro Phe Thr Val Val Gly Ala Asp Pro Thr Ser Asp
 340 345 350
 Ile Ala Val Val Arg Val Gln Gly Val Ser Gly Leu Thr Pro Ile Ser
 355 360 365
 Leu Gly Ser Ser Ser Asp Leu Arg Val Gly Gln Pro Val Leu Ala Ile
 370 375 380
 Gly Ser Pro Leu Gly Leu Glu Gly Thr Val Thr Thr Gly Ile Val Ser
 385 390 395 400
 Ala Leu Asn Arg Pro Val Ser Thr Thr Gly Glu Ala Gly Asn Gln Asn
 405 410 415
 Thr Val Leu Asp Ala Ile Gln Thr Asp Ala Ala Ile Asn Pro Gly Asn
 420 425 430
 Ser Gly Gly Ala Leu Val Asn Met Asn Ala Gln Leu Val Gly Val Asn
 435 440 445
 Ser Ala Ile Ala Thr Leu Gly Ala Asp Ser Ala Asp Ala Gln Ser Gly
 450 455 460
 Ser Ile Gly Leu Gly Phe Ala Ile Pro Val Asp Gln Ala Lys Arg Ile
 465 470 475 480
 Ala Asp Glu Leu Ile Ser Thr Gly Lys Ala Ser His Ala Ser Leu Gly
 485 490 495
 Val Gln Val Thr Asn Asp Lys Asp Thr Pro Gly Ala Lys Ile Val Glu
 500 505 510
 Val Val Ala Gly Gly Ala Ala Ala Asn Ala Gly Val Pro Lys Gly Val
 515 520 525

Val Val Thr Lys Val Asp Asp Arg Pro Ile Asn Ser Ala Asp Ala Leu
 530 535 540
 Val Ala Ala Val Arg Ser Lys Ala Pro Gly Ala Thr Val Ala Leu Thr
 545 550 555 560
 Phe Gln Asp Pro Ser Gly Gly Ser Arg Thr Val Gln Val Thr Leu Gly
 565 570 575
 Lys Ala Glu Gln
 580

(2) SEQ ID NO: 76 的信息:

(i) 序列特征:

- (A) 长度: 233 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 76:

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

```

Leu Gly Ile Thr Ser Ala Asp Val Asp Val Arg Ala Asn Pro Leu Ala
      85                      90                      95
Ala Lys Gly Val Cys Thr Tyr Asn Asp Glu Gln Gly Val Pro Phe Arg
      100                    105                    110
Val Gln Gly Asp Asn Ile Ser Val Lys Leu Phe Asp Asp Trp Ser Asn
      115                    120                    125
Leu Gly Ser Ile Ser Glu Leu Ser Thr Ser Arg Val Leu Asp Pro Ala
      130                    135                    140
Ala Gly Val Thr Gln Leu Leu Ser Gly Val Thr Asn Leu Gln Ala Gln
      145                    150                    155                    160
Gly Thr Glu Val Ile Asp Gly Ile Ser Thr Thr Lys Ile Thr Gly Thr
      165                    170                    175
Ile Pro Ala Ser Ser Val Lys Met Leu Asp Pro Gly Ala Lys Ser Ala
      180                    185                    190
Arg Pro Ala Thr Val Trp Ile Ala Gln Asp Gly Ser His His Leu Val
      195                    200                    205
Arg Ala Ser Ile Asp Leu Gly Ser Gly Ser Ile Gln Leu Thr Gln Ser
      210                    215                    220
Lys Trp Asn Glu Pro Val Asn Val Asp
      225                    230

```

(2) SEQ ID NO: 77 的信息:

(i) 序列特征:

- (A) 长度: 66 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 77:

```

Val Ile Asp Ile Ile Gly Thr Ser Pro Thr Ser Trp Glu Gln Ala Ala
1           5           10           15
Ala Glu Ala Val Gln Arg Ala Arg Asp Ser Val Asp Asp Ile Arg Val
20           25           30
Ala Arg Val Ile Glu Gln Asp Met Ala Val Asp Ser Ala Gly Lys Ile
35           40           45
Thr Tyr Arg Ile Lys Leu Glu Val Ser Phe Lys Met Arg Pro Ala Gln
50           55           60
Pro Arg
65

```

(2) SEQ ID NO: 78 的信息:

(i) 序列特征:

(A) 长度: 69 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 78:

```

Val Pro Pro Ala Pro Pro Leu Pro Pro Leu Pro Pro Ser Pro Ile Ser
1           5           10           15
Cys Ala Ser Pro Pro Ser Pro Pro Leu Pro Pro Ala Pro Pro Val Ala
20           25           30
Pro Gly Pro Pro Met Pro Pro Leu Asp Pro Trp Pro Pro Ala Pro Pro
35           40           45
Leu Pro Tyr Ser Thr Pro Pro Gly Ala Pro Leu Pro Pro Ser Pro Pro
50           55           60
Ser Pro Pro Leu Pro
65

```

(2) SEQ ID NO: 79 的信息:

(i) 序列特征:

(A) 长度: 355 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 79:

Met Ser Asn Ser Arg Arg Arg Ser Leu Arg Trp Ser Trp Leu Leu Ser
1 5 10 15

Val Leu Ala Ala Val Gly Leu Gly Leu Ala Thr Ala Pro Ala Gln Ala
20 25 30

Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
35 40 45

Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Ala Pro Gln Val Val
50 55 60

Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
65 70 75 80

Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val
85 90 95

Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln
100 105 110

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

(2) SEQ ID NO: 80 的信息:

(i) 序列特征:

(A) 长度: 205 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 80:

Ser Pro Lys Pro Asp Ala Glu Glu Gln Gly Val Pro Val Ser Pro Thr
1 5 10 15

Ala Ser Asp Pro Ala Leu Leu Ala Glu Ile Arg Gln Ser Leu Asp Ala
20 25 30

Thr Lys Gly Leu Thr Ser Val His Val Ala Val Arg Thr Thr Gly Lys
35 40 45

Val Asp Ser Leu Leu Gly Ile Thr Ser Ala Asp Val Asp Val Arg Ala
50 55 60

Asn Pro Leu Ala Ala Lys Gly Val Cys Thr Tyr Asn Asp Glu Gln Gly
65 70 75 80

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

Asp Trp Ser Asn Leu Gly Ser Ile Ser Glu Leu Ser Thr Ser Arg Val
100 105 110

Leu Asp Pro Ala Ala Gly Val Thr Gln Leu Leu Ser Gly Val Thr Asn
115 120 125

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

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

145	150	155	160
Ala Lys Ser Ala Arg Pro Ala Thr Val Trp Ile Ala Gln Asp Gly Ser			
	165	170	175
His His Leu Val Arg Ala Ser Ile Asp Leu Gly Ser Gly Ser Ile Gln			
	180	185	190
Leu Thr Gln Ser Lys Trp Asn Glu Pro Val Asn Val Asp			
	195	200	205

(2) SEQ ID NO: 81 的信息:

(i) 序列特征:

(A) 长度: 286 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 81:

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

```

Ala Thr Glu Gln Arg Thr Asn Lys Xaa Gln Ile Leu Ala Ser Gly Val
    115                      120                      125

Ala Met Pro Ala Ala Leu Arg Ala Ala Gln Met Leu Ala Ala Glu Trp
    130                      135                      140

Asp Val Ala Ala Asp Val Trp Ser Val Thr Ser Trp Gly Glu Leu Asn
    145                      150                      155                      160

Arg Asp Gly Val Val Ile Glu Thr Glu Lys Leu Arg His Pro Asp Arg
                      165                      170                      175

Pro Ala Gly Val Pro Tyr Val Thr Arg Ala Leu Glu Asn Ala Arg Gly
    180                      185                      190

Pro Val Ile Ala Val Ser Asp Trp Met Arg Ala Val Pro Glu Gln Ile
    195                      200                      205

Arg Pro Trp Val Pro Gly Thr Tyr Leu Thr Leu Gly Thr Asp Gly Phe
    210                      215                      220

Gly Phe Ser Asp Thr Arg Pro Ala Gly Arg Arg Tyr Phe Asn Thr Asp
    225                      230                      235                      240

Ala Glu Ser Gln Val Gly Arg Gly Phe Gly Arg Gly Trp Pro Gly Arg
                      245                      250                      255

Arg Val Asn Ile Asp Pro Phe Gly Ala Gly Arg Gly Pro Pro Ala Gln
                      260                      265                      270

Leu Pro Gly Phe Asp Glu Gly Gly Gly Leu Arg Pro Xaa Lys
    275                      280                      285

```

(2) SEQ ID NO: 82 的信息:

(i) 序列特征:

(A) 长度: 173 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 82:

```

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

```

(2) SEQ ID NO: 83 的信息:

(i) 序列特征:

(A) 长度: 107 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 83:

```

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

```

(2) SEQ ID NO: 84 的信息:

(i) 序列特征:

(A) 长度: 125 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 84:

```

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

```

(2) SEQ ID NO: 85 的信息:

(i) 序列特征:

(A) 长度: 117 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 85:

```

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

```

(2) SEQ ID NO: 86 的信息:

(i) 序列特征:

(A) 长度: 103 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 86:

```

Met Tyr Arg Phe Ala Cys Arg Thr Leu Met Leu Ala Ala Cys Ile Leu
1           5           10           15
Ala Thr Gly Val Ala Gly Leu Gly Val Gly Ala Gln Ser Ala Ala Gln
20           25           30
Thr Ala Pro Val Pro Asp Tyr Tyr Trp Cys Pro Gly Gln Pro Phe Asp
35           40           45
Pro Ala Trp Gly Pro Asn Trp Asp Pro Tyr Thr Cys His Asp Asp Phe
50           55           60

```

```

His Arg Asp Ser Asp Gly Pro Asp His Ser Arg Asp Tyr Pro Gly Pro
65              70              75              80

Ile Leu Glu Gly Pro Val Leu Asp Asp Pro Gly Ala Ala Pro Pro Pro
            85              90              95

Pro Ala Ala Gly Gly Gly Ala
            100

```

(2) SEQ ID NO: 87 的信息:

(i) 序列特征:

(A) 长度: 88 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 87:

```

Val Gln Cys Arg Val Trp Leu Glu Ile Gln Trp Arg Gly Met Leu Gly
1              5              10              15

Ala Asp Gln Ala Arg Ala Gly Gly Pro Ala Arg Ile Trp Arg Glu His
            20              25              30

Ser Met Ala Ala Met Lys Pro Arg Thr Gly Asp Gly Pro Leu Glu Ala
            35              40              45

Thr Lys Glu Gly Arg Gly Ile Val Met Arg Val Pro Leu Glu Gly Gly
            50              55              60

Gly Arg Leu Val Val Glu Leu Thr Pro Asp Glu Ala Ala Ala Leu Gly
65              70              75              80

Asp Glu Leu Lys Gly Val Thr Ser
            85

```

(2) SEQ ID NO: 88 的信息:

(i) 序列特征:

(A) 长度: 95 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 88:

```

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

```

(2) SEQ ID NO: 89 的信息:

(i) 序列特征:

(A) 长度: 166 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 89:

```

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

```

(2) SEQ ID NO: 90 的信息:**(i) 序列特征:****(A) 长度: 5 个氨基酸**

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 90:

```
Arg Ala Glu Arg Met
1           5
```

(2) SEQ ID NO: 91 的信息:

(i) 序列特征:

(A) 长度: 263 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 91:

```
Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala Ala
1           5           10          15
```

```
Gln Val Arg Val Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu Thr
20          25          30
```

```
Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile Leu
35          40          45
```

```
Ile Ala Thr Asn Leu Leu Gly Gln Asn Thr Pro Ala Ile Ala Val Asn
50          55          60
```

```
Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met Phe
65          70          75          80
```

```
Gly Tyr Ala Ala Ala Thr Ala Thr Ala Thr Leu Leu Pro Phe
85          90          95
```

```
Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gln Ala
100         105         110
```

Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Gln Leu Met
 115 120 125
 Asn Asn Val Pro Gln Ala Leu Lys Gln Leu Ala Gln Pro Thr Gln Gly
 130 135 140
 Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser Pro
 145 150 155 160
 His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His Met
 165 170 175
 Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser Met
 180 185 190
 Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala Gln Ala Val Gln Thr Ala
 195 200 205
 Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu Gly
 210 215 220
 Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala Ala
 225 230 235 240
 Ser Val Arg Tyr Gly His Arg Asp Gly Gly Lys Tyr Ala Xaa Ser Gly
 245 250 255
 Arg Arg Asn Gly Gly Pro Ala
 260

(2) SEQ ID NO: 92 的信息:

(i) 序列特征:

- (A) 长度: 303 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 92:

```

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

```

Gln Gln Gly Ala Gln Gln Ala Ala Gly Leu Gln Ser Pro Gly Pro Gln
 195 200 205
 Gln Ser Pro Gln Pro Pro Gly Tyr Gly Ser Gln Tyr Gly Gly Tyr Ser
 210 215 220
 Ser Ser Pro Ser Gln Ser Gly Ser Gly Tyr Thr Ala Gln Pro Pro Ala
 225 230 235 240
 Gln Pro Pro Ala Gln Ser Gly Ser Gln Gln Ser His Gln Gly Pro Ser
 245 250 255
 Thr Pro Pro Thr Gly Phe Pro Ser Phe Ser Pro Pro Pro Pro Val Ser
 260 265 270
 Ala Gly Thr Gly Ser Gln Ala Gly Ser Ala Pro Val Asn Tyr Ser Asn
 275 280 285
 Pro Ser Gly Gly Glu Gln Ser Ser Ser Pro Gly Gly Ala Pro Val
 290 295 300

(2) SEQ ID NO: 93 的信息:

(i) 序列特征:

- (A) 长度: 28 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 93:

Gly Cys Gly Glu Thr Asp Ala Ala Thr Leu Ala Gln Glu Ala Gly Asn
 1 5 10 15
 Phe Glu Arg Ile Ser Gly Asp Leu Lys Thr Gln Ile
 20 25

(2) SEQ ID NO: 94 的信息:

(i) 序列特征:

(A) 长度: 16 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 94:

Asp	Gln	Val	Glu	Ser	Thr	Ala	Gly	Ser	Leu	Gln	Gly	Gln	Trp	Arg	Gly
1				5				10						15	

(2) SEQ ID NO: 95 的信息:

(i) 序列特征:

(A) 长度: 27 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 95:

Gly	Cys	Gly	Ser	Thr	Ala	Gly	Ser	Leu	Gln	Gly	Gln	Trp	Arg	Gly	Ala
1				5				10						15	

Ala	Gly	Thr	Ala	Ala	Gln	Ala	Ala	Val	Val	Arg
			20					25		

(2) SEQ ID NO: 96 的信息:

(i) 序列特征:

(A) 长度: 27 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 96:

Gly Cys Gly Gly Thr Ala Ala Gln Ala Ala Val Val Arg Phe Gln Glu
1 5 10 15

Ala Ala Asn Lys Gln Lys Gln Glu Leu Asp Glu
20 25

(2) SEQ ID NO: 97 的信息:

(i) 序列特征:

(A) 长度: 27 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 97:

Gly Cys Gly Ala Asn Lys Gln Lys Gln Glu Leu Asp Glu Ile Ser Thr
1 5 10 15

Asn Ile Arg Gln Ala Gly Val Gln Tyr Ser Arg
20 25

(2) SEQ ID NO: 98 的信息:

(i) 序列特征:

(A) 长度: 28 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 98:

Gly Cys Gly Ile Arg Gln Ala Gly Val Gln Tyr Ser Arg Ala Asp Glu
1 5 10 15

Glu Gln Gln Gln Ala Leu Ser Ser Gln Met Gly Phe
20 25

(2) SEQ ID NO: 99 的信息:

(i) 序列特征:

(A) 长度: 507 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 99:

ATGAAGATGG TGAAATCGAT CGCCGCAGGT CTGACCGCCG CGGCTGCAAT CGGCGCCGCT	60
GCGGCCGGTG TGACTTCGAT CATGGCTGGC GGCCCGGTCTG TATACCAGAT GCAGCCGGTC	120
GTCTTCGGCG CGCCACTGCC GTTGGACCCG GCATCCGCCC CTGACGTCCC GACCGCCGCC	180
CAGTTGACCA GCCTGCTCAA CAGCCTCGCC GATCCCAACG TGTCGTTTGC GAACAAGGGC	240
AGTCTGGTCG AGGGCGGCAT CGGGGGCACC GAGGCGCGCA TCGCCGACCA CAAGCTGAAG	300
AAGGCCGCCG AGCACGGGGA TCTGCCGCTG TCGTTCAGCG TGACGAACAT CCAGCCGGCG	360
GCCGCCGGTT CGGCCACCGC CGACGTTTCC GTCTCGGGTC CGAAGCTCTC GTCGCCGGTC	420
ACGCAGAACG TCACGTTCTG GAATCAAGGC GGCTGGATGC TGTCACGCGC ATCGGCGATG	480
GAGTTGCTGC AGGCCGCAGG GAACTGA	507

(2) SEQ ID NO: 100 的信息:

(i) 序列特征:

(A) 长度: 168 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 100:

```

Met Lys Met Val Lys Ser Ile Ala Ala Gly Leu Thr Ala Ala Ala Ala
1           5           10           15

Ile Gly Ala Ala Ala Ala Gly Val Thr Ser Ile Met Ala Gly Gly Pro
          20           25           30

Val Val Tyr Gln Met Gln Pro Val Val Phe Gly Ala Pro Leu Pro Leu
          35           40           45

Asp Pro Ala Ser Ala Pro Asp Val Pro Thr Ala Ala Gln Leu Thr Ser
          50           55           60

Leu Leu Asn Ser Leu Ala Asp Pro Asn Val Ser Phe Ala Asn Lys Gly
65           70           75           80

Ser Leu Val Glu Gly Gly Ile Gly Gly Thr Glu Ala Arg Ile Ala Asp
          85           90           95

His Lys Leu Lys Lys Ala Ala Glu His Gly Asp Leu Pro Leu Ser Phe
          100          105          110

Ser Val Thr Asn Ile Gln Pro Ala Ala Ala Gly Ser Ala Thr Ala Asp
          115          120          125

Val Ser Val Ser Gly Pro Lys Leu Ser Ser Pro Val Thr Gln Asn Val
          130          135          140

Thr Phe Val Asn Gln Gly Gly Trp Met Leu Ser Arg Ala Ser Ala Met
145           150           155           160

Glu Leu Leu Gln Ala Ala Gly Asn
          165

```

(2) SEQ ID NO: 101 的信息:

(i) 序列特征:

(A) 长度: 500 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 101:

```

CGTGGCAATG TCGTTGACCG TCGGGGCCGG GGTGCGCTCC GCAGATCCCG TGGACGCGGT      60
CATTAACACC ACCTGCAATT ACGGGCAGGT AGTAGCTGCG CTCAACGCGA CGGATCCGGG      120
GGCTGCCGCA CAGTTCAACG CCTCACCGGT GGCGCAGTCC TATTTGCGCA ATTTCTCGC      180
CGCACCGCCA CCTCAGCGCG CTGCCATGGC CGCGCAATTG CAAGCTGTGC CGGGGGCGGC      240
ACAGTACATC GGCCTTGTCG AGTCGGTTGC CGGCTCCTGC AACAACTATT AAGCCCATGC      300
GGGCCCCATC CCGCGACCCG GCATCGTCGC CGGGGCTAGG CCAGATTGCC CCGCTCCTCA      360
ACGGGGCCGA TCCCGCGACC CGGCATCGTC GCCGGGGCTA GGCCAGATTG CCCCCTCCT      420
CAACGGGGCCG CATCTCGTGC CGAATTCCTG CAGCCCGGGG GATCCACTAG TTCTAGAGCG      480
GCCGCCACCG CGGTGGAGCT                                          500

```

(2) SEQ ID NO: 102 的信息:

(i) 序列特征:

(A) 长度: 96 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 102:

```

Val Ala Met Ser Leu Thr Val Gly Ala Gly Val Ala Ser Ala Asp Pro
1           5           10           15

Val Asp Ala Val Ile Asn Thr Thr Cys Asn Tyr Gly Gln Val Val Ala
          20           25           30

Ala Leu Asn Ala Thr Asp Pro Gly Ala Ala Ala Gln Phe Asn Ala Ser
          35           40           45

Pro Val Ala Gln Ser Tyr Leu Arg Asn Phe Leu Ala Ala Pro Pro Pro
50           55           60

```

Gln Arg Ala Ala Met Ala Ala Gln Leu Gln Ala Val Pro Gly Ala Ala
65 70 75 80

Gln Tyr Ile Gly Leu Val Glu Ser Val Ala Gly Ser Cys Asn Asn Tyr
85 90 95

(2) SEQ ID NO: 103 的信息:

(i) 序列特征:

- (A) 长度: 154 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 103:

ATGACAGAGC AGCAGTGGAA TTTCGCGGGT ATCGAGGCCG CGGCAAGCGC AATCCAGGGA 60
AATGTCACGT CCATTCATTC CCTCCTTGAC GAGGGGAAGC AGTCCCTGAC CAAGCTCGCA 120
GCGGCCTGGG GCGGTAGCGG TTCGGAAGCG TACC 154

(2) SEQ ID NO: 104 的信息:

(i) 序列特征:

- (A) 长度: 51 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 104:

Met Thr Glu Gln Gln Trp Asn Phe Ala Gly Ile Glu Ala Ala Ala Ser
1 5 10 15
Ala Ile Gln Gly Asn Val Thr Ser Ile His Ser Leu Leu Asp Glu Gly
20 25 30

Lys Gln Ser Leu Thr Lys Leu Ala Ala Ala Trp Gly Gly Ser Gly Ser
 35 40 45
 Glu Ala Tyr
 50

(2) SEQ ID NO: 105 的信息:

(i) 序列特征:

- (A) 长度: 282 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 105:

CGGTCGCGCA CTTCCAGGTG ACTATGAAAG TCGGCTTCCG NCTGGAGGAT TCCTGAACCT	60
TCAAGCGCGG CCGATAACTG AGGTGCATCA TTAAGCGACT TTTCCAGAAC ATCCTGACGC	120
GCTCGAAACG CGGCACAGCC GACGGTGGCT CCGNCGAGGC GCTGNCTCCA AAATCCCTGA	180
GACAATTCGN CGGGGGCGCC TACAAGGAAG TCGGTGCTGA ATTCGNCNG TATCTGGTCG	240
ACCTGTGTGG TCTGNAGCCG GACGAAGCGG TGCTCGACGT CG	282

(2) SEQ ID NO: 106 的信息:

(i) 序列特征:

- (A) 长度: 1565 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 106:

GTATGCGGCC ACTGAAGTCG CCAATGCGGC GCGGCCAGC TAAGCCAGGA ACAGTCGGCA	60
CGAGAAACCA CGAGAAATAG GGACACGTAA TGGTGGATTT CGGGGCGTTA CCACCGGAGA	120
TCAACTCCGC GAGGATGTAC GCCGGCCCGG GTTCGGCCTC GCTGGTGGCC GCGGCTCAGA	180
TGTGGGACAG CGTGGCGAGT GACCTGTTTT CGGCCGCGTC GGC GTTTCAG TCGGTGGTCT	240
GGGGTCTGAC GGTGGGGTCG TGGATAGGTT CGTCGGCGGG TCTGATGGTG GCGGCGGCCT	300
CGCCGTATGT GCGTGGATG AGCGTCACCG CGGGGCAGGC CGAGCTGACC GCCGCCAGG	360
TCCGGGTTGC TCGGGCGGCC TACGAGACGG CGTATGGGCT GACGGTGCCC CCGCCGGTGA	420
TCGCCGAGAA CCGTGCTGAA CTGATGATTC TGATAGCGAC CAACCTCTTG GGGCAAAACA	480
CCCCGGCGAT CGCGGTCAAC GAGGCCGAAT ACGCGAGAT GTGGGCCCAA GACGCCGCCG	540
CGATGTTTGG CTACGCCGCG GCGACGGCGA CGGCGACGGC GACGTTGCTG CCGTTCGAGG	600
AGGCGCCGGA GATGACCAGC GCGGGTGGGC TCCTCGAGCA GGCCGCCGCG GTCGAGGAGG	660
CCTCCGACAC CGCCGCGGCG AACCAGTTGA TGAACAATGT GCCCCAGGCG CTGCAACAGC	720
TGGCCAGCC CACGCAGGGC ACCACGCCTT CTTCCAAGCT GGGTGGCCTG TGGAAGACGG	780
TCTCGCCGCA TCGGTCGCCG ATCAGCAACA TGGTGTCAAT GGCCAACAAC CACATGTCAA	840
TGACCAACTC GGGTGTGTCA ATGACCAACA CCTTGAGCTC GATGTTGAAG GGCTTTGCTC	900
CGGCGGCGGC CGCCAGGCC GTGCAAACCG CGGCGCAAAA CGGGGTCCGG GCGATGAGCT	960
CGCTGGGCAG CTCGCTGGGT TCTTCGGGTC TGGGCGGTGG GGTGGCCGCC AACTTGGGTC	1020
GGGCGGCCTC GGTGGTTCG TTGTCGGTGC CGCAGGCCTG GGCCGCGGCC AACCAGGCAG	1080
TCACCCCGGC GGC GCGGGCG CTGCCGCTGA CCAGCCTGAC CAGCGCCGCG GAAAGAGGGC	1140
CCGGGCAGAT GCTGGGCGGG CTGCCGGTGG GGCAGATGGG CGCCAGGGCC GGTGGTGGGC	1200
TCAGTGGTGT GCTGCGTGTT CCGCCGCGAC CCTATGTGAT GCCGCATTCT CCGGCGGCCG	1260
GCTAGGAGAG GGGGCGCAGA CTGTCGTTAT TTGACCAGTG ATCGGCGGTC TCGGTGTTTC	1320

```

CGCGGCCGGC TATGACAACA GTCAATGTGC ATGACAAGTT ACAGGTATTA GGTCCAGGTT    1380
CAACAAGGAG ACAGGCAACA TGGCCTCACG TTTTATGACG GATCCGCACG CGATGCGGGA    1440
CATGGCGGGC CGTTTTGAAG TGCACGCCCA GACGGTGGAG GACGAGGCTC GCCGGATGTG    1500
GGCGTCCGCG CAAAACATTT CCGGTGCGGG CTGGAGTGGC ATGGCCGAGG CGACCTCGCT    1560
AGACA                                             1565

```

(2) SEQ ID NO: 107 的信息:

(i) 序列特征:

(A) 长度: 391 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 107:

```

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

```

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

Ala Leu Pro Leu Thr Ser Leu Thr Ser Ala Ala Glu Arg Gly Pro Gly
 340 345 350

Gln Met Leu Gly Gly Leu Pro Val Gly Gln Met Gly Ala Arg Ala Gly
 355 360 365

Gly Gly Leu Ser Gly Val Leu Arg Val Pro Pro Arg Pro Tyr Val Met
 370 375 380

Pro His Ser Pro Ala Ala Gly
 385 390

(2) SEQ ID NO: 108 的信息:

(i) 序列特征:

- (A) 长度: 259 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 108:

ACCAACACCT TGA CT CNAT GTTGAAGGGC TTAGCTCCGG CGGCGGCTCA GGCCGTGGAA	60
ACCGCGGCGG AAAACGGGGT CTGGGCAATG AGCTCGCTGG GCAGCCAGCT GGGTTCGTCTG	120
CTGGGTTCTT CGGGTCTGGG CGCTGGGGTG GCCGCCAACT TGGGTCGGGC GGCCTCGGTC	180
GGTTCGTTGT CGGTGCCGCC AGCATGGGCC GCGGCCAACC AGGCGGTCAC CCCGGCGGCG	240
CGGGCGCTGC CGCTGACCA	259

(2) SEQ ID NO: 109 的信息:

(i) 序列特征:

(A) 长度: 86 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 109:

```

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

```

(2) SEQ ID NO: 110 的信息:

(i) 序列特征:

(A) 长度: 1109 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 110:

TACTTGAGAG AATTTGACCT GTTGCCGACG TTGTTTGCTG TCCATCATTG GTGCTAGTTA	60
TGGCCGAGCG GAAGGATTAT CGAAGTGGTG GACTTCGGGG CGTTACCACC GGAGATCAAC	120
TCCGCGAGGA TGTACGCCGG CCCGGGTTCTG GCCTCGCTGG TGGCCGCCGC GAAGATGTGG	180
GACAGCGTGG CGAGTGACCT GTTTTCGGCC GCGTCGGCGT TTCAGTCGGT GGTCTGGGGT	240
CTGACGACGG GATCGTGGAT AGGTTCTGTCG GCGGGTCTGA TGGTGGCGGC GGCCTCGCCG	300
TATGTGGCGT GGATGAGCGT CACCGCGGGG CAGGCCGAGC TGACCGCCGC CCAGGTCCGG	360
GTTGCTGCGG CGGCCTACGA GACGGCGTAT GGGCTGACGG TGCCCCCGCC GGTGATCGCC	420
GAGAACCGTG CTGAACTGAT GATTCTGATA GCGACCAACC TCTTGGGGCA AAACACCCCG	480
GCGATCGCGG TCAACGAGGC CGAATACGGG GAGATGTGGG CCCAAGACGC CGCCGCGATG	540
TTTGGCTACG CCGCCACGGC GCGGACGGCG ACCGAGGCGT TGCTGCCGTT CGAGGACGCC	600
CCACTGATCA CCAACCCCGG CGGGCTCCTT GAGCAGGCCG TCGCGGTCGA GGAGGCCATC	660
GACACCGCCG CGGCGAACCA GTTGATGAAC AATGTGCCCC AAGCGCTGCA ACAACTGGCC	720
CAGCCCACGA AAAGCATCTG GCCGTTTCGAC CAACTGAGTG AACTCTGGAA AGCCATCTCG	780
CCGCATCTGT CGCCGCTCAG CAACATCGTG TCGATGCTCA ACAACCACGT GTCGATGACC	840
AACTCGGGTG TGTCAATGGC CAGCACCTTG CACTCAATGT TGAAGGGCTT TGCTCCGGCG	900
GCGGCTCAGG CCGTGGAAC CGCGGCGCAA AACGGGGTCC AGGCGATGAG CTCGCTGGGC	960
AGCCAGCTGG GTTCGTGCT GGGTTCTTCG GGTCTGGGCG CTGGGGTGGC CGCCAACTTG	1020
GGTCGGGCGG CCTCGGTCGG TTCGTTGTCTG GTGCCGAGG CCTGGGCGGC GGCCAACCAG	1080
GCGGTCACCC CGGCGGCGCG GGCCTGCC	1109

(2) SEQ ID NO: 111 的信息:

(i) 序列特征:

(A) 长度: 341 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 111:

```

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

```

Gln Asn Thr Pro Ala Ile Ala Val Asn Glu Ala Glu Tyr Gly Glu Met
 130 135 140

Trp Ala Gln Asp Ala Ala Ala Met Phe Gly Tyr Ala Ala Thr Ala Ala
 145 150 155 160

Thr Ala Thr Glu Ala Leu Leu Pro Phe Glu Asp Ala Pro Leu Ile Thr
 165 170 175

Asn Pro Gly Gly Leu Leu Glu Gln Ala Val Ala Val Glu Glu Ala Ile
 180 185 190

Asp Thr Ala Ala Ala Asn Gln Leu Met Asn Asn Val Pro Gln Ala Leu
 195 200 205

Gln Gln Leu Ala Gln Pro Thr Lys Ser Ile Trp Pro Phe Asp Gln Leu
 210 215 220

Ser Glu Leu Trp Lys Ala Ile Ser Pro His Leu Ser Pro Leu Ser Asn
 225 230 235 240

Ile Val Ser Met Leu Asn Asn His Val Ser Met Thr Asn Ser Gly Val
 245 250 255

Ser Met Ala Ser Thr Leu His Ser Met Leu Lys Gly Phe Ala Pro Ala
 260 265 270

Ala Ala Gln Ala Val Glu Thr Ala Ala Gln Asn Gly Val Gln Ala Met
 275 280 285

Ser Ser Leu Gly Ser Gln Leu Gly Ser Ser Leu Gly Ser Ser Gly Leu
 290 295 300

Gly Ala Gly Val Ala Ala Asn Leu Gly Arg Ala Ala Ser Val Gly Ser
 305 310 315 320

Leu Ser Val Pro Gln Ala Trp Ala Ala Ala Asn Gln Ala Val Thr Pro
 325 330 335

Ala Ala Arg Ala Leu
 340

(2) SEQ ID NO: 112 的信息:

(i) 序列特征:

(A) 长度: 1256 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 112:

CATCGGAGGG AGTGATCACC ATGCTGTGGC ACGCAATGCC ACCGGAGNTA AATACCGCAC	60
GGCTGATGGC CGGCGCGGGT CCGGCTCCAA TGCTTGCGGC GGCCGCGGGA TGGCAGACGC	120
TTTCGGCGGC TCTGGACGCT CAGGCCGTCG AGTTGACCGC GCGCCTGAAC TCTCTGGGAG	180
AAGCCTGGAC TGGAGGTGGC AGCGACAAGG CGCTTGCGGC TGCAACGCCG ATGGTGGTCT	240
GGCTACAAAC CGCGTCAACA CAGGCCAAGA CCCGTGCGAT GCAGGCGACG GCGCAAGCCG	300
CGGCATACAC CCAGGCCATG GCCACGACGC CGTCGCTGCC GGAGATCGCC GCCAACCACA	360
TCACCCAGGC CGTCCTTACG GCCACCAACT TCTTCGGTAT CAACACGATC CCGATCGCGT	420
TGACCGAGAT GGATTATTTC ATCCGTATGT GGAACCAGGC AGCCCTGGCA ATGGAGGTCT	480
ACCAGGCCGA GACCGCGGTT AACACGCTTT TCGAGAAGCT CGAGCCGATG GCGTCGATCC	540
TTGATCCCGG CGCGAGCCAG AGCACGACGA ACCCGATCTT CGGAATGCCC TCCCCTGGCA	600
GCTCAACACC GGTGCGCCAG TTGCCGCCGG CGGCTACCCA GACCCTCGGC CAACTGGGTG	660
AGATGAGCGG CCCGATGCAG CAGCTGACCC AGCCGCTGCA GCAGGTGACG TCGTTGTTCA	720
GCCAGGTGGG CGGCACCGGC GCGGGCAACC CAGCCGACGA GGAAGCCGCG CAGATGGGCC	780
TGCTCGGCAC CAGTCCGCTG TCGAACCATC CGCTGGCTGG TGGATCAGGC CCCAGCGCGG	840
GCGCGGGCCT GCTGCGCGCG GAGTCGCTAC CTGGCGCAGG TGGGTCGTTG ACCCGCACGC	900
CGCTGATGTC TCAGCTGATC GAAAAGCCGG TTGCCCCCTC GGTGATGCCG GCGGCTGCTG	960
CCGGATCGTC GGCGACGGGT GGCGCCGCTC CGGTGGGTGC GGGAGCGATG GGCCAGGGTG	1020

CGCAATCCGG CGGCTCCACC AGGCCGGGTC TGGTCGCGCC GGCACCGCTC GCGCAGGAGC	1080
GTGAAGAAGA CGACGAGGAC GACTGGGACG AAGAGGACGA CTGGTGAGCT CCCGTAATGA	1140
CAACAGACTT CCGGCCACC CGGGCCGGAA GACTTGCCAA CATTTTGGCG AGGAAGGTAA	1200
AGAGAGAAAG TAGTCCAGCA TGGCAGAGAT GAAGACCGAT GCCGCTACCC TCGCGC	1256

(2) SEQ ID NO: 113 的信息:

(i) 序列特征:

(A) 长度: 432 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 113:

CTAGTGGATG GGACCATGGC CATTTTCTGC AGTCTCACTG CCTTCTGTGT TGACATTTTG	60
GCACGCCGGC GGAAACGAAG CACTGGGGTC GAAGAACGGC TCGCTGCCA TATCGTCCGG	120
AGCTTCCATA CCTTCGTGCG GCCGGAAGAG CTTGTCTAG TCGGCCGCCA TGACAACCTC	180
TCAGAGTGCG CTCAAACGTA TAAACACGAG AAAGGGCGAG ACCGACGGAA GGTGGAATC	240
GCCCGATCCC GTGTTTCGCT ATTCTACGCG AACTCGGCGT TGCCCTATGC GAACATCCCA	300
GTGACGTTGC CTTCGGTCGA AGCCATTGCC TGACCGGCTT CGCTGATCGT CCGCGCCAGG	360
TTCTGCAGCG CGTTGTTTCTAG CTCGGTAGCC GTGGCGTCCC ATTTTGTCTG GACACCCTGG	420
TACGCCTCCG AA	432

(2) SEQ ID NO: 114 的信息:

(i) 序列特征:

(A) 长度: 368 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 114:

```

Met Leu Trp His Ala Met Pro Pro Glu Xaa Asn Thr Ala Arg Leu Met
1           5           10           15

Ala Gly Ala Gly Pro Ala Pro Met Leu Ala Ala Ala Ala Gly Trp Gln
          20           25           30

Thr Leu Ser Ala Ala Leu Asp Ala Gln Ala Val Glu Leu Thr Ala Arg
          35           40           45

Leu Asn Ser Leu Gly Glu Ala Trp Thr Gly Gly Gly Ser Asp Lys Ala
          50           55           60

Leu Ala Ala Ala Thr Pro Met Val Val Trp Leu Gln Thr Ala Ser Thr
          65           70           75           80

Gln Ala Lys Thr Arg Ala Met Gln Ala Thr Ala Gln Ala Ala Ala Tyr
          85           90           95

Thr Gln Ala Met Ala Thr Thr Pro Ser Leu Pro Glu Ile Ala Ala Asn
          100          105          110

His Ile Thr Gln Ala Val Leu Thr Ala Thr Asn Phe Phe Gly Ile Asn
          115          120          125

Thr Ile Pro Ile Ala Leu Thr Glu Met Asp Tyr Phe Ile Arg Met Trp
          130          135          140

Asn Gln Ala Ala Leu Ala Met Glu Val Tyr Gln Ala Glu Thr Ala Val
          145          150          155          160

Asn Thr Leu Phe Glu Lys Leu Glu Pro Met Ala Ser Ile Leu Asp Pro
          165          170          175

```

Gly Ala Ser Gln Ser Thr Thr Asn Pro Ile Phe Gly Met Pro Ser Pro
 180 185 190

Gly Ser Ser Thr Pro Val Gly Gln Leu Pro Pro Ala Ala Thr Gln Thr
 195 200 205

Leu Gly Gln Leu Gly Glu Met Ser Gly Pro Met Gln Gln Leu Thr Gln
 210 215 220

Pro Leu Gln Gln Val Thr Ser Leu Phe Ser Gln Val Gly Gly Thr Gly
 225 230 235 240

Gly Gly Asn Pro Ala Asp Glu Glu Ala Ala Gln Met Gly Leu Leu Gly
 245 250 255

Thr Ser Pro Leu Ser Asn His Pro Leu Ala Gly Gly Ser Gly Pro Ser
 260 265 270

Ala Gly Ala Gly Leu Leu Arg Ala Glu Ser Leu Pro Gly Ala Gly Gly
 275 280 285

Ser Leu Thr Arg Thr Pro Leu Met Ser Gln Leu Ile Glu Lys Pro Val
 290 295 300

Ala Pro Ser Val Met Pro Ala Ala Ala Ala Gly Ser Ser Ala Thr Gly
 305 310 315 320

Gly Ala Ala Pro Val Gly Ala Gly Ala Met Gly Gln Gly Ala Gln Ser
 325 330 335

Gly Gly Ser Thr Arg Pro Gly Leu Val Ala Pro Ala Pro Leu Ala Gln
 340 345 350

Glu Arg Glu Glu Asp Asp Glu Asp Asp Trp Asp Glu Glu Asp Asp Trp
 355 360 365

(2) SEQ ID NO: 115 的信息:

(i) 序列特征:

(A) 长度: 12 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 115:

Met	Ala	Glu	Met	Lys	Thr	Asp	Ala	Ala	Thr	Leu	Ala
1				5					10		

(2) SEQ ID NO: 116 的信息:

(i) 序列特征:

(A) 长度: 396 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 116:

GATCTCCGGC GACCTGAAAA CCCAGATCGA CCAGGTGGAG TCGACGGCAG GTTCGTTGCA	60
GGGCCAGTGG CGCGGCGCGG CGGGGACGGC CCCCCAGGCC GCGGTGGTGC GCTTCCAAGA	120
AGCAGCCAAT AAGCAGAAGC AGGAACTCGA CGAGATCTCG ACGAATATTC GTCAGGCCGG	180
CGTCCAATAC TCGAGGGCCG ACGAGGAGCA GCAGCAGGCG CTGTCCTCGC AAATGGGCTT	240
CTGACCCGCT AATACGAAAA GAAACGGAGC AAAACATGA CAGAGCAGCA GTGGAATTTT	300
GCGGGTATCG AGGCCGCGGC AAGCGCAATC CAGGGAAATG TCACGTCCAT TCATTCCCTC	360
CTTGACGAGG GGAAGCAGTC CCTGACCAAG CTCGCA	396

(2) SEQ ID NO: 117 的信息:

(i) 序列特征:

(A) 长度: 80 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 117:

```

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

```

(2) SEQ ID NO: 118 的信息:

(i) 序列特征:

(A) 长度: 387 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 118:

GTGGATCCCG ATCCCGTGTT TCGCTATTCT ACGCGAACTC GGCGTTGCCC TATGCGAACA	60
TCCCAGTGAC GTTGCCTTCG GTCGAAGCCA TTGCCTGACC GGCTTCGCTG ATCGTCCGCG	120
CCAGGTTCTG CAGCGCGTTG TTCAGCTCGG TAGCCGTGGC GTCCCATTTT TGCTGGACAC	180
CCTGGTACGC CTCCGAACCG CTACCGCCCC AGGCCGCTGC GAGCTTGGTC AGGGACTGCT	240
TCCCCTCGTC AAGGAGGGAA TGAATGGACG TGACATTTCC CTGGATTGCG CTTGCCGCGG	300
CCTCGATACC CGCGAAATTC CACTGCTGCT CTGTCATGTT TTTGCTCCGT TTCTTTTCGT	360
ATTAGCGGGT CAGAAGCCCA TTTGCGA	387

(2) SEQ ID NO: 119 的信息:

(i) 序列特征:

- (A) 长度: 272 个碱基对
- (B) 类型: 核酸
- (C) 链型: 单链
- (D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 119:

CGGCACGAGG ATCTCGGTTG GCCCAACGGC GCTGGCGAGG GCTCCGTTCC GGGGGCGAGC	60
TGCGCGCCGG ATGCTTCCTC TGCCCGCAGC CGCGCCTGGA TGGATGGACC AGTTGCTACC	120
TTCCCGACGT TTCGTTCCGT GTCTGTGCGA TAGCGGTGAC CCCGGCGCGC ACGTCGGGAG	180
TGTTGGGGGG CAGGCCGGGT CGGTGGTTCG GCCGGGGACG CAGACGGTCT GGACGGAACG	240
GGCGGGGGTT CGCCGATTGG CATCTTTGCC CA	272

(2) SEQ ID NO: 120 的信息:

(i) 序列特征:

- (A) 长度: 20 个氨基酸

- (B) 类型: 氨基酸
(C) 链型:
(D) 拓扑结构: 线型
(xi) 序列描述: SEQ ID NO: 120:

Asp Pro Val Asp Ala Val Ile Asn Thr Thr Cys Asn Tyr Gly Gln Val
1 5 10 15
Val Ala Ala Leu
 20

(2) SEQ ID NO: 121 的信息:

- (i) 序列特征:
(A) 长度: 15 个氨基酸
(B) 类型: 氨基酸
(C) 链型:
(D) 拓扑结构: 线型
(xi) 序列描述: SEQ ID NO: 121:

Ala Val Glu Ser Gly Met Leu Ala Leu Gly Thr Pro Ala Pro Ser
1 5 10 15

(2) SEQ ID NO: 122 的信息:

- (i) 序列特征:
(A) 长度: 19 个氨基酸
(B) 类型: 氨基酸
(C) 链型:
(D) 拓扑结构: 线型
(xi) 序列描述: SEQ ID NO: 122:

Ala Ala Met Lys Pro Arg Thr Gly Asp Gly Pro Leu Glu Ala Ala Lys
1 5 10 15

Glu Gly Arg

(2) SEQ ID NO: 123 的信息:

(i) 序列特征:

(A) 长度: 15 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 123:

Tyr Tyr Trp Cys Pro Gly Gln Pro Phe Asp Pro Ala Trp Gly Pro
1 5 10 15

(2) SEQ ID NO: 124 的信息:

(i) 序列特征:

(A) 长度: 14 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 124:

Asp Ile Gly Ser Glu Ser Thr Glu Asp Gln Gln Xaa Ala Val
1 5 10

(2) SEQ ID NO: 125 的信息:

(i) 序列特征:

(A) 长度: 13 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 125:

Ala Glu Glu Ser Ile Ser Thr Xaa Glu Xaa Ile Val Pro
1 5 10

(2) SEQ ID NO: 126 的信息:

(i) 序列特征:

(A) 长度: 17 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 126:

Asp Pro Glu Pro Ala Pro Pro Val Pro Thr Thr Ala Ala Ser Pro Pro
1 5 10 15
Ser

(2) SEQ ID NO: 127 的信息:

(i) 序列特征:

(A) 长度: 15 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 127:

Ala Pro Lys Thr Tyr Xaa Glu Glu Leu Lys Gly Thr Asp Thr Gly
1 5 10 15

(2) SEQ ID NO: 128 的信息:

(i) 序列特征:

(A) 长度: 30 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 128:

Asp	Pro	Ala	Ser	Ala	Pro	Asp	Val	Pro	Thr	Ala	Ala	Gln	Leu	Thr	Ser
1				5				10					15		
Leu	Leu	Asn	Ser	Leu	Ala	Asp	Pro	Asn	Val	Ser	Phe	Ala	Asn		
		20					25						30		

(2) SEQ ID NO: 129 的信息:

(i) 序列特征:

(A) 长度: 22 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 129:

Asp	Pro	Pro	Asp	Pro	His	Gln	Xaa	Asp	Met	Thr	Lys	Gly	Tyr	Tyr	Pro
1				5				10					15		
Gly	Gly	Arg	Arg	Xaa	Phe										
				20											

(2) SEQ ID NO: 130 的信息:

(i) 序列特征:

(A) 长度: 7 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 130:

Asp Pro Gly Tyr Thr Pro Gly
1 5

(2) SEQ ID NO: 131 的信息:

(i) 序列特征:

(A) 长度: 10 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑结构: 线型

(ix) 特征:

(D) 其它信息: /注= 第二残基可以是Pro 或者 Thr

(xi) 序列描述: SEQ ID NO: 131:

Xaa Xaa Gly Phe Thr Gly Pro Gln Phe Tyr
1 5 10

(2) SEQ ID NO: 132 的信息:

(i) 序列特征:

(A) 长度: 9 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑结构: 线型

(ix) 特征:

(D) 其它信息: /注= 第三残基可以是Gln 或者 Leu

(xi) 序列描述: SEQ ID NO: 132:

Xaa Pro Xaa Val Thr Ala Tyr Ala Gly
1 5

(2) SEQ ID NO: 133 的信息:

(i) 序列特征:

(A) 长度: 9 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 133:

Xaa Xaa Xaa Glu Lys Pro Phe Leu Arg
1 5

(2) SEQ ID NO: 134 的信息:

(i) 序列特征:

(A) 长度: 15 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 134:

Xaa Asp Ser Glu Lys Ser Ala Thr Ile Lys Val Thr Asp Ala Ser
1 5 10 15

(2) SEQ ID NO: 135 的信息:

(i) 序列特征:

(A) 长度: 个氨基酸

(B) 类型: 15 氨基酸

(C) 链型:

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 135:

Ala Gly Asp Thr Xaa Ile Tyr Ile Val Gly Asn Leu Thr Ala Asp
1 5 10 15

(2) SEQ ID NO: 136 的信息:

(i) 序列特征:

(A) 长度: 15 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 136:

Ala Pro Glu Ser Gly Ala Gly Leu Gly Gly Thr Val Gln Ala Gly
1 5 10 15

(2) SEQ ID NO: 137 的信息:

(i) 序列特征:

(A) 长度: 21 个氨基酸

(B) 类型: 氨基酸

(C) 链型:

(D) 拓扑结构: 线型

(xi) 序列描述: SEQ ID NO: 137:

Xaa Tyr Ile Ala Tyr Xaa Thr Thr Ala Gly Ile Val Pro Gly Lys Ile
1 5 10 15

Asn Val His Leu Val
20

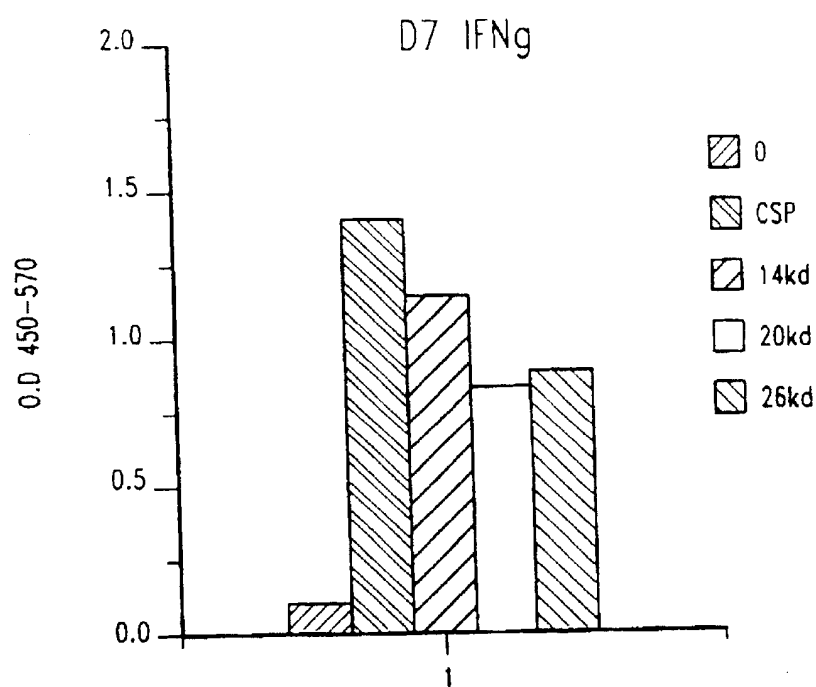
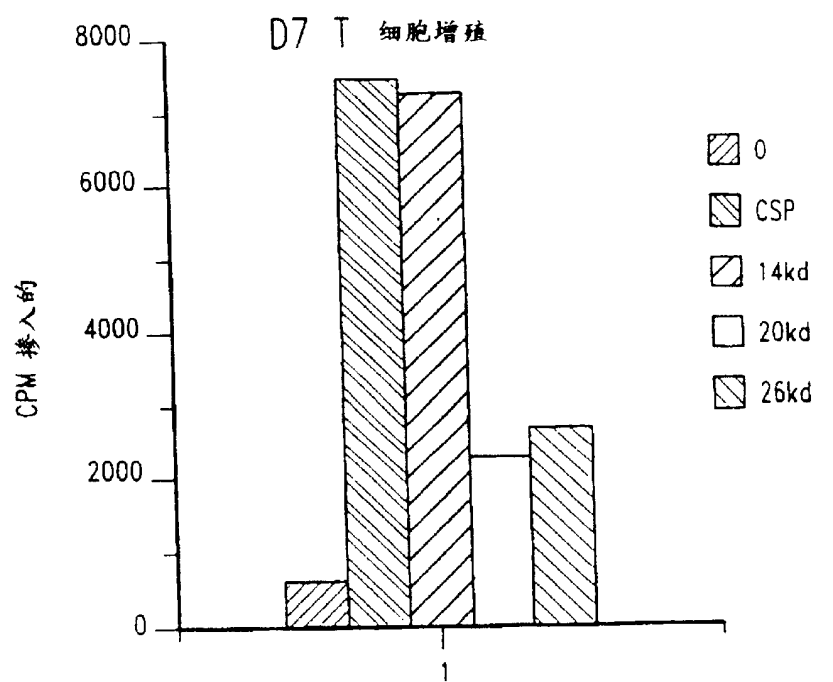


图 1A

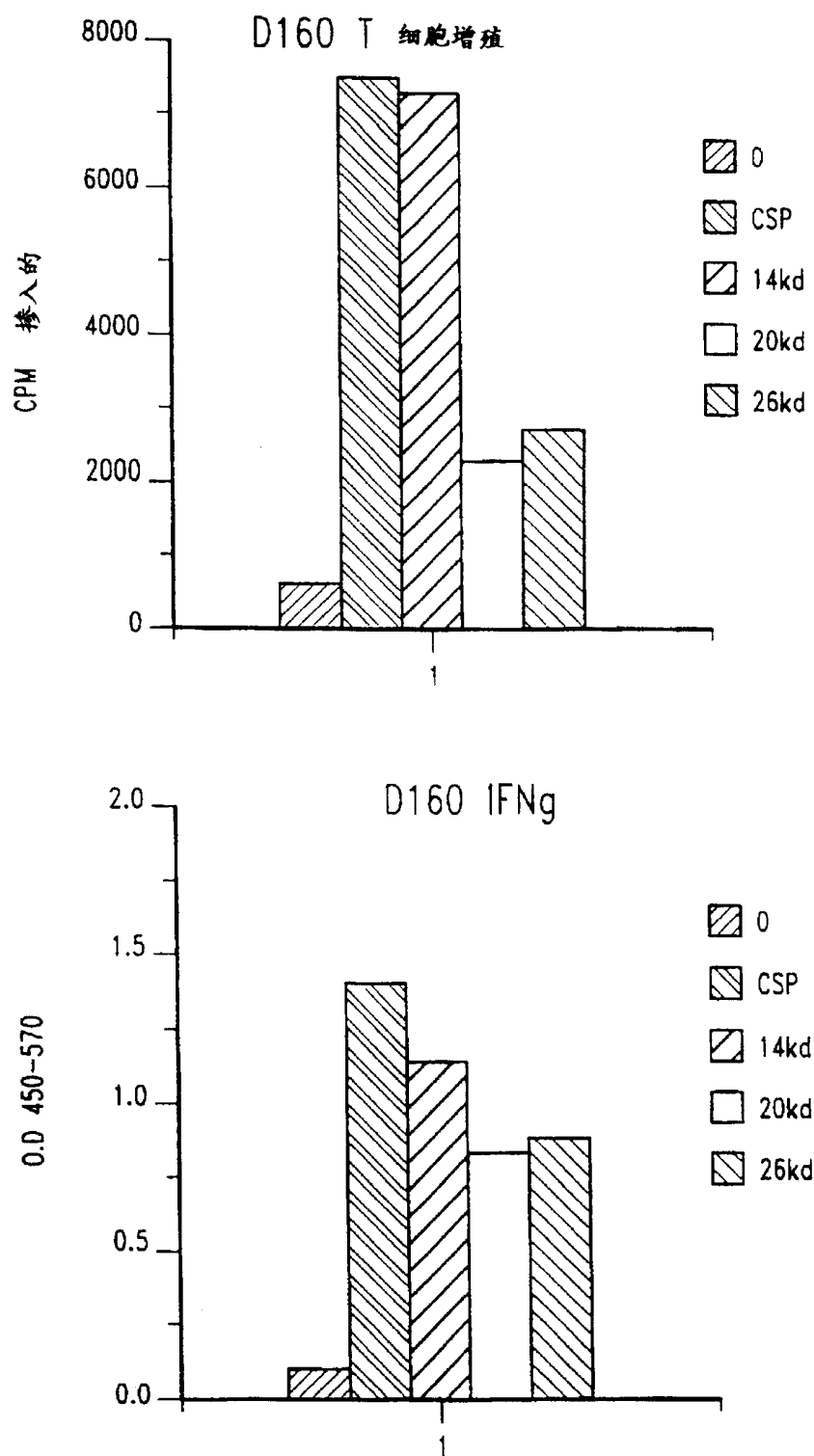


图 1B

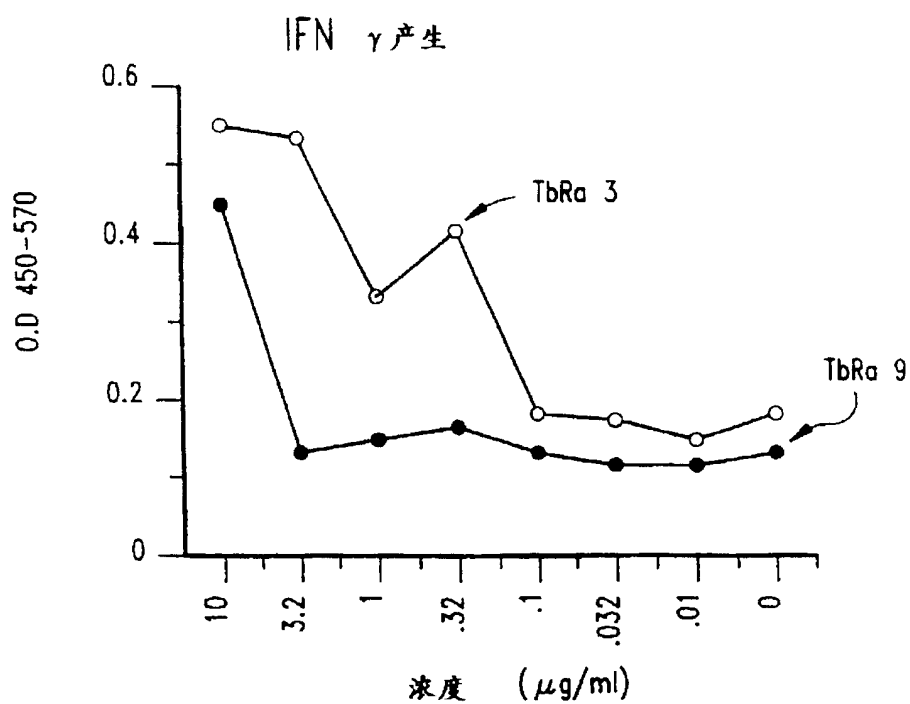
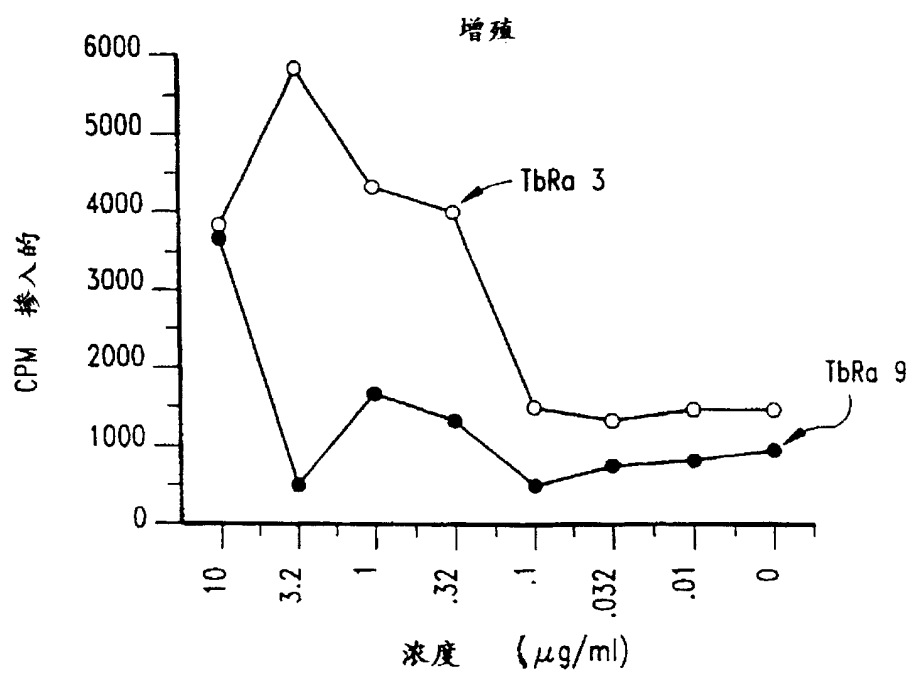


图 2