

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12N 15/31



[12] 发明专利申请公开说明书

[21] 申请号 00816501.7

C12N 15/63 C07K 14/195

C07K 16/12 G01N 33/53

G01N 33/68 A61K 39/02

[43] 公开日 2003 年 3 月 12 日

[11] 公开号 CN 1402785A

[22] 申请日 2000.9.26 [21] 申请号 00816501.7

[30] 优先权

[32] 1999. 9. 30 [33] GB [31] 9923156.5

[86] 国际申请 PCT/EP00/09495 2000.9.26

[87] 国际公布 WO01/23416 英 2001.4.5

[85] 进入国家阶段日期 2002.5.30

[71] 申请人 史密丝克莱恩比彻姆生物有限公司

地址 比利时里克森萨特

[72] 发明人 J·通纳德

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 张广育 姜建成

权利要求书 3 页 说明书 48 页 序列表 10 页
PCT/RO/134 表 1 页 附图 45 页

[54] 发明名称 新化合物

[57] 摘要

本发明提供 BASB132 多肽和编码 BASB132 多肽的多核苷酸, 以及由重组技术产生这些多肽的方法。 另外还提供诊断、预防及治疗性应用。

BASB132多核苷酸序列的对比
与序列编号:1相同处用圆点表示, 间隔处用短横线表示。

```
*                20                *
Seqid1 : ATGACTCATCAGGATAATTCAAAAAAGTGG : 30
Seqid3 : ..... : -

                40                *                60
Seqid1 : TCATGGGCGAGAAATGGGGTTTCGCTGGC : 60
Seqid3 : ..... : -

                *                80                *
Seqid1 : GTTGCTGCACCGAATAAAGTCAGCCGAAA : 90
Seqid3 : ..... : -

                100                *                120
Seqid1 : CCACCAAAACCCACCAAAACGCAATGGCTA : 120
Seqid3 : ..... : -

                *                140                *
Seqid1 : ATCTATTTAATTGCTTAAATGTCGGAGCA : 150
Seqid3 : ..... : -

                160                *                180
```

I S S N 1 0 0 8 - 4 2 7 4

1. 一种分离的多肽，该多肽含有一段氨基酸序列，该氨基酸序列与选自：序列编号：2 和序列编号：4 的全长氨基酸序列分别具有至少 85% 的同一性。
- 5 2. 权利要求 1 的一种分离的多肽，其中的氨基酸序列与选自：序列编号：2 和序列编号：4 的全长氨基酸序列分别具有至少 95% 的同一性。
3. 权利要求 1 的多肽，该多肽包含一段选自序列编号：2 和序列编号：4 的氨基酸序列。
4. 序列编号：2 或序列编号：4 的一种分离的多肽。
- 10 5. 权利要求 1-4 之一的多肽的一种免疫原性片段，其中该免疫原性片段的免疫原性与序列编号：2 或序列编号：4 的多肽基本相同。
6. 权利要求 1-5 之一的一种多肽，其中该多肽是一种较大的融合蛋白的一部分。
7. 一种分离的多核苷酸，该多核苷酸可编码权利要求 1-6 之一的
- 15 一种多肽。
8. 一种分离的多核苷酸，该多核苷酸含有一段可编码一种多肽的核苷酸序列，该多肽与序列编号：2 或 4 的全长氨基酸序列分别具有至少 85% 的同一性；或与这种分离的多核苷酸互补的一种核苷酸序列。
9. 一种分离的多核苷酸，该多核苷酸含有一段核苷酸序列，该序列与编码序列编号：2 或 4 的全部编码区的一种核苷酸序列具有至少 85% 的同一性；或与这种分离的多核苷酸互补的一种核苷酸序列。
- 20 10. 一种分离的多核苷酸，该多核苷酸含有一段核苷酸序列，该序列与序列编号：1 或 3 的全长核苷酸序列分别具有至少 85% 的同一性；或与这种分离的多核苷酸互补的一种核苷酸序列。
- 25 11. 权利要求 7-10 之一的一种分离的多核苷酸，其中与序列编号：1 或 3 的同一性至少为 95%。
12. 一种分离的多核苷酸，该多核苷酸含有一段核苷酸序列，该序列可编码序列编号：2 或序列编号：4 的多肽。
13. 一种分离的多核苷酸，该多核苷酸含有序列编号：1 或序列编号：3 的多核苷酸。
- 30 14. 一种分离的多核苷酸，该多核苷酸含有一段核苷酸序列，该序列可编码序列编号：2、序列编号：4 的多肽，获得该多核苷酸的方法可

以是，在严格杂交条件下用一种具有序列编号：1 或序列编号：3 所述序列的标记探针或其片段对适当的库进行筛选。

15. 一种表达载体或一种活体重组微生物，该载体或该活体重组微生物含有权利要求 7-14 之一的一种分离的多核苷酸。

5 16. 含有权利要求 15 的表达载体的一种宿主细胞，或该宿主细胞的一种亚细胞组分或膜，该宿主细胞或其亚细胞组分或膜能够表达一种分离的多肽，该多肽含有一段氨基酸序列，该序列与选自序列编号：2 和序列编号：4 的氨基酸序列具有至少 85% 的同一性。

10 17. 一种方法，用于产生权利要求 1-6 的一种多肽，该方法包括，在足以产生所述多肽的条件下培养权利要求 16 的一种宿主细胞，以及从培养基中回收这种多肽。

15 18. 一种方法，用于表达权利要求 7-14 之一的一种多核苷酸，该方法包括，用至少含有一种所述多核苷酸的表达载体对一种宿主细胞进行转化，以及在足以表达任何一种所述多核苷酸的条件下培养该宿主细

胞。

19. 一种疫苗组合物，该组合物含有有效剂量的权利要求 1-6 之一的一种多肽和一种药学容许的载体。

20. 一种疫苗组合物，该组合物含有有效剂量的权利要求 7-14 之一的一种多核苷酸和一种有效的药用载体。

20 21. 权利要求 19 或 20 之一的一种疫苗组合物，其中该组合物至少含有粘膜炎莫拉氏菌以外的一种抗原。

22. 一种抗体，该抗体对权利要求 1-6 之一的多肽或免疫原性片段具有免疫特异性。

25 23. 一种诊断粘膜炎莫拉氏菌感染的方法，该方法包括，确定权利要求 1-6 之一的一种多肽或对该多肽具有免疫特异性的一种抗体在怀疑患有这种感染的动物的生物学样品中的存在情况。

24. 一种组合物在一种药剂的制备中的应用，该组合物含有免疫学有效剂量的权利要求 1-6 之一的一种多肽，该药剂用于在动物体内产生免疫应答。

30 25. 一种组合物在一种药剂的制备中的应用，该组合物含有免疫学有效剂量的权利要求 7-14 之一的一种多核苷酸，该药剂用于在动物体内产生免疫应答。

26. 一种治疗组合物，可用于对患有粘膜炎莫拉氏菌性疾病的人进行治疗，该组合物至少含有一种针对权利要求 1-6 之一的多肽的抗体，以及一种适当的药用载体。

新化合物

发明领域

- 5 本发明涉及多核苷酸（文中称为“BASB132 多核苷酸”）、这些多核苷酸编码的多肽（文中称为“BASB132”或“BASB132 多肽”），以及用于其产生的重组材料和方法。另一方面，本发明涉及这些多肽和多核苷酸的使用方法，其中包括抗细菌感染的疫苗。再一方面，本发明涉及检测某些病原体感染的诊断测定方法。

10

发明背景

- 粘膜炎莫拉氏菌（也叫粘膜炎布兰汉氏球菌）是一种常分离自人类上呼吸道的革兰氏阴性菌。它可引起数种疾病，主要是幼儿及儿童中耳炎和老年肺炎。它还能引起鼻窦炎和院内感染，有时还可导致侵袭性疾病。

- 15 从发病数量和可能引发的后遗症来看，中耳炎是一种重要的儿童期疾病。在美国，每年记录在案的病例都在 3,500,000 例以上，估计有 80% 的儿童在 3 岁以前至少经历过一次耳炎发病（Klein, JO(1994)Clin.Inf.Dis 19:823）。这种疾病若不经处理，或是转为慢性，将会导致暂时性（液体在中耳内累积时）或永久性（听神经损坏时）的听力丧失。在幼儿中，这种听力丧失可能会造成语言学习延迟。

- 20 从患有中耳炎的儿童的中耳内主要可分离出三种细菌：肺炎链球菌、未分类流感嗜血菌（NTHi）和粘膜炎莫拉氏菌。这些细菌存在于 60-90% 的病例中。关于近期研究的一篇综述显示，肺炎链球菌和 NTHi 各占中耳炎病例的约 30%，粘膜炎莫拉氏菌约占 15%（Murphy, TF(1996) Microbiol.Rev. 60:267）。从中耳内还可分离出其他细菌（B 型流感嗜血菌、酿脓链球菌等），但出现率要低得多（占病例的 2% 或更低）。

- 25 流行病学的资料显示，对中耳内发现的病原体而言，在上呼吸道建群是耳炎产生的必需前提；但引起该疾病还需要其他因素（Dickinson, DP et al. (1988) J.Infect.Dis. 158:205, Faden, HL et al. (1991) Ann. Otorhinol. Laryngol. 100:612）。这些因素的重要性在于可引发细菌经咽鼓管迁移到中耳内，然后启动炎症过程。目前对这些因素还不甚了解。据推测，例

如病毒性感染后的免疫系统暂时性异常可能会造成无法控制呼吸道建群 (Faden, HL et al (1994) J. Infect. Dis. 169:1312)。另一种解释是,暴露于环境因素会使一些儿童出现更重要的建群,所以这些儿童因中耳病原体的持续存在而变得容易患中耳炎 (Murphy, TF (1996) Microbiol. Rev. 60:267)。

有关抗粘膜炎莫拉氏菌的免疫应答的特性还了解得很少。从0岁生长至2岁的幼儿的鼻咽中顺序分离出菌株并进行分析,其结果表明,他们经常会感染新菌株,然后再将其清除。这说明建群后的儿童能产生抗这种细菌的有效免疫应答 (Faden, HL et al (1994) J. Infect. Dis. 169:1312)。

在对成人的测试中,大多数都能鉴定出杀菌性抗体 (Chapman, AJ et al. (1985) J. Infect. Dis. 151:878)。粘膜炎莫拉氏菌菌株在抵抗血清杀菌活性的能力上存在差异:一般而言,患病个体的分离物要比只建群的个体的分离物更具有抗性 (Hol, C et al. (1993) Lancet 341:1281, Jordan, KL et al. (1990) Am. J. Med. 88 (suppl. 5A): 28S)。因此,可将血清抗性看作细菌的一种毒力因素。在中耳炎痊愈儿童的血清中发现了一种调理素活性。

OMP B1 是一种 84 kDa 的蛋白,其表达受铁元素调控,它可被肺炎患者的血清识别 (Sethi, S, et al. (1995) Infect. Immun. 63:1516),除了该蛋白以及 UspA1 和 UspA2 (Chen D. et al. (1999), Infect. Immun. 67:1310) 之外,在人类中还未发现这些不同的免疫应答所针对的其他抗原。

目前已利用生化方法对粘膜炎莫拉氏菌表面的其他一些膜蛋白进行了特性鉴定,或是对这些膜蛋白在诱导保护性免疫方面的潜在作用进行了研究 (有关综述可参考 Murphy, TF (1996) Microbiol. Rev. 60:267)。

在小鼠肺炎模型中,针对其中一些膜蛋白 (UspA、CopB) 的抗体的存在有利于更快清除肺部感染。铜绿假单胞菌的孔蛋白能够有效抵抗动物模型中的这种细菌,有一种在粘膜炎莫拉氏菌菌株中高度保守的多肽 (OMP CD) 与孔蛋白具有同源性。

在过去几十年中,粘膜炎莫拉氏菌感染的频率大幅度提高。其原因是多重抗生素抗性菌株的出现和免疫系统薄弱人群的增加。分离出对某些或所有普通抗生素具有抗性的粘膜炎莫拉氏菌菌株已不再困难。这种现象使我们必须用新型抗微生物剂、疫苗、药物筛选方法和诊断测试方

法来满足对该生物的医疗需求。

发明概述

本发明涉及 BASB132，特别是 BASB132 多肽和 BASB132 多核苷酸，以及用于其产生的重组材料和方法。另一方面，本发明涉及这些多肽和多核苷酸的使用方法，其中特别包括微生物性疾病的预防和治疗方
5 法。再一方面，本发明涉及检测与微生物感染有关的疾病和病症的诊断测定方法，如检测 BASB132 多核苷酸或多肽的表达或活性的测定方法。
通过阅读下文以及本公开内容的其他部分，该领域的技术人员将很
10 容易了解在本发明的精神和范围内的不同变化和改进。

发明详述

本发明涉及 BASB132 多肽和多核苷酸，下文将对此做更详细的描述。具体地说，本发明涉及粘膜炎莫拉氏菌的 BASB132 多肽和多核苷酸，其氨基酸序列与大肠杆菌假定外膜蛋白 YtfN 同源。本发明特别涉及
15 分别具有序列编号：1 或 3 的核苷酸序列以及序列编号：2 或 4 的氨基酸序列的 BASB132。应该了解的是，下文“序列表”中列举的“DNA”序列只代表本发明的一项实施方案的一个范例，因为该领域的技术人员都会承认，这些序列一般都能够能够在多核苷酸，包括多聚核糖核苷酸，中有效
20 地使用。

多肽

一方面，本发明提供文中称为“BASB132”和“BASB132 多肽”的粘膜炎莫拉氏菌多肽及其具有生物学、诊断、预防、临床或治疗作用的
25 变异体，以及含有这些多肽的组合物。

本发明还提供：

(a) 一种分离的多肽，该多肽含有一段氨基酸序列，该序列与序列编号：2 或 4 的序列具有至少 85% 的同一性，优选的是至少 90% 的同一性，更优选的是至少 95% 的同一性，最优选的是至少 97-99% 的同一性或
30 完全相同；

(b) 一种由分离的多核苷酸编码的多肽，该多核苷酸包含一段多核苷酸序列，该序列与序列编号：1 或 3 的全长序列分别具有至少 85% 的

同一性，优选的是至少 90% 的同一性，更优选的是至少 95% 的同一性，再更优选的是至少 97-99% 的同一性或完全相同；或

- (c) 一种由分离的多核苷酸编码的多肽，该多核苷酸包含一段可编码一种多肽的多核苷酸序列，该多肽与序列编号：2 或 4 的氨基酸序列具有至少 85% 的同一性，优选的是至少 90% 的同一性，更优选的是至少 95% 的同一性，再更优选的则至少 97-99% 的同一性或完全相同。

序列编号：2 或 4 提供的 BASB132 多肽为粘膜莫拉氏菌 Mc2931 菌株 (ATCC43617) 的 BASB132 多肽。

- 本发明还提供 BASB132 多肽的一种免疫原性片段，即 BASB132 多肽中与含有序列编号：2 或 4 所示氨基酸序列的多肽具有相同或基本相同的免疫原性的一段连续部分；也就是说，该片段（如果需要，可与一种载体相联接）可诱发一种能够识别 BASB132 多肽的免疫应答。这种免疫原性片段可包括，例如，缺少 N 端前导序列和（或）跨膜区和（或）C 端锚定区的 BASB132 多肽。优选地，本发明所述 BASB132 的该免疫原性片段基本上包含一种多肽的所有细胞外区域，这种多肽与序列编号：2 或 4 的全长序列分别具有至少 85% 的同一性，优选的是至少 90% 的同一性，更优选的是至少 95% 的同一性，最优选的是至少 97-99% 的同一性。

- 片段是指一种多肽，该多肽具有一段氨基酸序列，该序列与本发明的任何一种多肽的任意部分氨基酸序列完全相同，但不是与全部氨基酸序列都完全相同。与 BASB132 多肽一样，片段可以是“独立的”，也可以包含在一种较大的多肽中，从而构成这种较大多肽的一部分或一个区域，最优选的是构成单一较大多肽的一段单一的连续区域。

- 优选的片段包括，例如，具有序列编号：2 或 4 的一部分氨基酸序列的截断多肽或其变体，如含有氨基端和（或）羧基端氨基酸序列的连续残基链。优选的片段还包括在宿主细胞内或由宿主细胞产生的本发明所述多肽的降解形式。其他优选片段还包括用结构或功能特性表征的片段，如含有 α -螺旋和 α -螺旋形成区、 β -折叠和 β -折叠形成区、转角和转角形成区、卷曲和卷曲形成区、亲水区、疏水区、 α 两性区、 β 两性区、挠性区、表面形成区、底物结合区及高抗原指数区的片段。

- 优选的片段还包括一种分离的多肽，该多肽含有一段至少有 15、20、30、40、50 或 100 个连续氨基酸的氨基酸序列，这些连续氨基酸来自序列编号：2 或 4 的氨基酸序列，优选的片段还包括另一种分离的多肽，该

多肽含有一段至少有 15、20、30、40、50 或 100 个连续氨基酸的氨基酸序列，这些连续氨基酸是由序列编号：2 或 4 的氨基酸序列截断或缺失而来。

本发明所述多肽的片段可通过肽合成方法用来产生相应的全长多肽；因此，这些片段可作为中间体来产生本发明的全长多肽。

特别优选的变异体是以任意组合方式取代、缺失或添加若干个、5-10 个、1-5 个、1-3 个、1-2 个或 1 个氨基酸的变异体。

本发明的多肽，或免疫原性片段，可以是“成熟”蛋白的形式，也可以是一种较大多肽的一部分，如前体或融合蛋白的一部分。通常可添加一种有益的额外氨基酸序列，该序列可含有分泌序列或前导序列、原序列、多组氨酸残基等有助于纯化的序列，或能够在重组生产过程中保持稳定的序列。此外，还可以考虑添加外源多肽，或脂类尾，或多核苷酸序列，以提高最终分子的免疫原性潜力。

一方面，本发明涉及可溶性基因工程融合蛋白，这些融合蛋白含有本发明的多肽或其片段，并含有不同亚类免疫球蛋白（IgG、IgM、IgA、IgE）的重链或轻链恒定区的不同部分。优选的免疫球蛋白是人类 IgG 的重链恒定区，特别是 IgG1 的重链恒定区，并且融合出现在铰链区。在一项特别实施方案中，去除 Fc 部分的方法只是简单地掺入一种能被凝血因子 Xa 剪切的酶切序列。

此外，本发明涉及利用基因工程方法制备这些融合蛋白的方法，以及这些融合蛋白在药物筛选、诊断和治疗中的应用。另一方面，本发明涉及编码这些融合蛋白的多核苷酸。有关融合蛋白的技术可参考，例如，国际专利申请 WO94/29458 和 WO94/22914。

可以用化学方法使蛋白结合，也可以作为重组融合蛋白来进行表达，以使表达系统的产量与非融合蛋白相比有所提高。融合配体的作用可以是协助提供辅助 T 细胞表位（免疫融合配体），优选的是能被人类识别的辅助 T 细胞表位，也可以是有助于以高于天然重组蛋白的产量来表达蛋白（表达增强子）。优选的是融合配体既具有免疫融合配体的作用，又具有表达增强配体的作用。

融合配体包括流感嗜血菌的蛋白 D 和流感病毒的非结构蛋白 NS1（血凝素）。另一种融合配体是被称为 LytA 的蛋白。优选的是使用该分子的 C 端部分。LytA 来自肺炎链球菌，该细菌可合成 N-乙酰-L-丙氨酸酰

胺酶，即 LytA 酰胺酶，（由 lytA 基因编码 {Gene, 43 (1986) 第 265-272 页}）这是一种可特异性降解肽聚糖骨架中特定化学键的自溶素。

LytA 蛋白的 C 端区域与胆碱或 DEAE 等某些胆碱类似物具有亲和性。这种特性被用来产生可用于融合蛋白表达的大肠杆菌 C-LytA 表达质粒。氨基

5 基端含有 C-LytA 片段的杂合蛋白的纯化方法已有所描述

{Biotechnology:10, (1992) 第 795-798 页}。也可以使用 LytA 分子中从第 178 位残基开始的 C 端重复部分，例如第 188-305 位残基区域。

本发明还包括上述多肽的变异体，即通过保守氨基酸取代将参照多肽的一种残基取代成另一种性质类似的残基而获得的不同多肽。典型的
10 保守取代是在 Ala、Val、Leu 和 Ile 之间；Ser 和 Thr 之间；酸性残基 Asp 和 Glu 之间；Asn 和 Gln 之间；以及碱性残基 Lys 和 Arg 之间；或芳香族残基 Phe 和 Tyr 之间进行。

本发明的多肽可采用任何适当的方式来制备。这些多肽包括，分离的天然多肽、重组产生的多肽、合成的多肽，或由这些方法的组合方法
15 产生的多肽。制备这些多肽的方法已为该领域所熟知。

本发明最优选的多肽是来源于粘膜莫拉氏菌的多肽，但由分类上处于相同属的其他生物获得的多肽也是优选的。本发明的多肽还可以从，例如，分类上处于相同科或相同目的其他生物获得。

20 多核苷酸

本发明的目标是提供一种可编码 BASB132 多肽的多核苷酸，特别是可编码文中指定的 BASB132 的多核苷酸。

在一项特别优选实施方案中，本发明的多核苷酸含有一段 BASB132 多肽编码区，该编码区含有序列编号：1 或 3 的一段全长基因序列，或其
25 变异体。

序列编号：1 或 3 的 BASB132 多核苷酸是来源于粘膜莫拉氏菌 Mc2931 菌株 (ATCC43617) 的 BASB132 多核苷酸。

另一方面，本发明提供分离的核酸分子，这些核酸分子可编码和（或）表达 BASB132 多肽和多核苷酸，特别是粘膜莫拉氏菌的
30 BASB132 多肽和多核苷酸，其中包括，例如，未加工的 RNAs、核酶 RNAs、mRNAs、cDNAs、基因组 DNAs、B 型 DNAs 和 Z 型 DNAs。本发明的其他实施方案还包括，具有生物学、诊断、预防、临床或治疗作

用的多核苷酸和多肽及其变异体和含有这些多核苷酸和（或）多肽的组合物。

另一方面，本发明涉及分离的多核苷酸，该多核苷酸至少包括一个全长基因，并且可编码具有序列编号：2或4所示推导氨基酸序列的

5 BASB132多肽，以及与其密切相关的多核苷酸及其变异体。

在另一项特别优选实施方案中，本发明的粘膜炎莫拉氏菌 BASB132多肽含有或包括序列编号：2或4的氨基酸序列或其变异体。

利用文中提供的信息，如序列编号：1或3所示多核苷酸序列，本发明的编码 BASB132多肽的多核苷酸可采用标准的克隆和筛选方法来产生，例如以粘膜炎莫拉氏菌 Catlin 细胞为原材料将细菌的染色体 DNA 片段克隆和测序，然后获得全长克隆。举例来说，为了获得本发明的多核苷酸序列，如序列编号：1或3的多核苷酸序列，典型的方法是利用由部分序列获得的放射性标记寡核苷酸，优选的是含有17个或更多核苷酸的寡核苷酸，对大肠杆菌或其他适当宿主中的粘膜炎莫拉氏菌 Catlin 菌株染色体 DNA 克隆库进行探测。然后用严格杂交条件鉴别出带有与探针序列完全相同的 DNA 的克隆。然后利用由最初的多肽或多核苷酸序列设计出的测序引物对杂交鉴定出的克隆分别测序，这样就可能使多核苷酸序列双向延伸，从而确定全长基因序列。这种测序方法可采用，例如，由质粒克隆制备的双链变性 DNA 来方便地实现。适当的技术在 Maniatis, T., 10 Fritsch, E.F.和 Sambrook et al., 分子克隆实验手册，第二版；Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989) 中有所描述。（特别是第1.90节“通过杂交进行筛选”和第13.70节“双链变性 DNA 模板的测序”）。也可通过基因组 DNA 的直接测序来获得全长基因序列。需要说明的是，序列编号：1或3的多核苷酸是在粘膜炎莫拉氏菌的 DNA 库中发现的。

此外，序列编号：1或3的 DNA 序列均包含一种可读框，可编码大致含有序列编号：2或4的氨基酸残基数量的一种蛋白，该蛋白的推导分子量可利用该领域技术人员熟知的氨基酸残基分子量值来加以计算。

在序列编号：1中，第1位核苷酸的起始密码子与开始于第5017位核苷酸的终止密码子之间的多核苷酸可编码序列编号：2的多肽。

在序列编号：3中，第2042位核苷酸的起始密码子与开始于第5014位核苷酸的终止密码子之间的多核苷酸可编码序列编号：4的多肽。

另一方面，本发明提供一种分离的多核苷酸，该多核苷酸含有或包括：

(a) 一段多核苷酸序列，该序列与序列编号：1 或 3 的全长序列分别具有至少 85% 的同一性，优选的是至少 90% 的同一性，更优选的是至少 95% 的同一性，再更优选的是至少 97-99% 的同一性或完全相同；或

(b) 一段多核苷酸序列，该序列可编码一种多肽，该多肽与序列编号：2 或 4 的全长氨基酸序列分别具有至少 85% 的同一性，优选的是至少 90% 的同一性，更优选的是至少 95% 的同一性，再更优选的是至少 97-99% 的同一性或 100% 相同。

利用一种方法可获得编码本发明所述多肽的多核苷酸，其中包括由粘膜莫拉氏菌以外的物种获得的同源物和直向同源物，该方法包括，在严格杂交条件下（如采用 45-65℃ 的温度和 0.1-1% 的 SDS 浓度）用一种含有或包括序列编号：1 或 3 所述序列或其片段的标记探针或可检测探针针对适当的库进行筛选；分离出含有所述多核苷酸序列的全长基因克隆和（或）基因组克隆。

本发明提供一种多核苷酸序列，该序列的全长序列与序列编号：1 或 3 的编码序列（可读框）完全相同。本发明还提供一种可编码成熟多肽或其片段的独立序列，以及可编码成熟多肽或其片段并且含有可读框吻合的另一编码序列的序列，另一编码序列可编码，例如，前导序列或分泌序列、前蛋白序列、原蛋白序列或前原蛋白序列。本发明的多核苷酸还可含有至少一种非编码序列，其中包括，例如，至少一种 5' 和 3' 非编码序列，如转录但不翻译的序列，终止信号（如 rho-依赖性和 rho-非依赖性终止信号）、核糖体结合位点、Kozak 序列、可保持 mRNA 稳定的序列、内含子以及多腺苷酸化信号，但不局限于此。多核苷酸序列还可含有编码额外氨基酸的额外编码序列。例如可编码一种有利于融合多肽纯化的标记序列。在本发明的某些实施方案中，该标记序列为 Gentz et al, Proc. Natl. Acad. Sci., USA 86:821-824 (1989) 描述的由 pQE 载体（Qiagen, Inc.）提供的六组氨酸肽，或 HA 肽标记（Wilson et al., Cell 37:767 (1984)），这两种标记均可用来纯化与其融合的多肽序列。本发明的多核苷酸还包括，但不局限于，含有一种结构基因以及调控该基因表达的天然相关序列的多核苷酸。

序列编号：2 或 4 所述 BASB132 多肽的核苷酸编码序列可分别与序

列编号：1 中第 1-5016 位核苷酸包含的多肽编码序列或序列编号：3 中第 1-5016 位核苷酸包含的多肽编码序列完全相同。作为选择，也可以是由于遗传密码冗余（简并性）而同样编码序列编号：2 或 4 所述多肽的序列。

5 文中使用的术语“编码多肽的多核苷酸”包括那些含有一种序列的多核苷酸，该序列可编码本发明的多肽，特别是细菌多肽，更明确地说是具有序列编号：2 或 4 所示氨基酸序列的粘膜炎莫拉氏菌 BASB132 多肽。该术语还包括那些含有连续或不连续的单一区域以及附加区域的多核苷酸（例如，被整合噬菌体、整合插入序列、整合载体序列、整合转座子序列打断的多核苷酸，或由于 RNA 编辑或基因组 DNA 重排而被打断的多核苷酸），其中，单一区域可编码本发明的多肽，附加区域也可以含有编码序列和（或）非编码序列。

本发明还涉及上述多核苷酸的变异体，这些变异体可编码一种具有序列编号：2 或 4 所述推导氨基酸序列的多肽的变异体。本发明所述多核苷酸的片段可用来，例如，合成本发明的全长多核苷酸。

其他的特别优选实施方案还包括可编码 BASB132 变异体的多核苷酸，这些变异体具有序列编号：2 或 4 所述 BASB132 多肽的氨基酸序列，并且以任意组合方式取代、修饰、缺失和（或）添加若干个、几个、5-10 个、1-5 个、1-3 个、2 个、1 个或 0 个氨基酸残基。其中，特别优选的是不改变 BASB132 多肽的特性和活性的沉默取代、添加和缺失。

本发明的其他优选实施方案还包括一些多核苷酸及其互补多核苷酸，这些多核苷酸的全长序列与具有序列编号：2 或 4 所述氨基酸序列的 BASB132 多肽的编码多核苷酸至少有 85% 同一性。作为选择，最优选的是含有一种区域的多核苷酸及其互补多核苷酸，该区域的全长序列与编码 BASB132 多肽的多核苷酸至少有 90% 的同一性。由此来看，特别优选的多核苷酸的全长序列与编码 BASB132 多肽的多核苷酸至少有 95% 的同一性。而在至少有 95% 的同一性的多核苷酸中，高度优选的是至少有 97% 的同一性的多核苷酸，其中，特别高度优选的是至少有 98% 的同一性和至少有 99% 的同一性的多核苷酸，而更优选的是至少有 99% 的同一性的多核苷酸。

优选的实施方案是可编码一些多肽的多核苷酸，这些多肽可保持与序列编号：1 或 3 所述 DNA 编码的成熟多肽基本相同的生物功能或活

性。

本发明的某些优选实施方案提供一些多核苷酸，这些多核苷酸能够与 BASB132 多核苷酸序列杂交，明确地说，是在严格条件下杂交，如序列编号：1 或 3 的多核苷酸。

- 5 本发明还涉及一些能够与文中提供的多核苷酸序列杂交的多核苷酸。因此，本发明特别涉及一些能够在严格条件下与文中提供的多核苷酸序列杂交的多核苷酸。文中使用的术语“严格条件”和“严格杂交条件”的含义是，只有在序列间至少有 95% 的同一性的条件下才发生杂交，优选的是至少有 97% 的同一性的条件下才发生杂交。严格杂交条件的具体实例是，在含有：50% 甲酰胺、5 × SSC (150mM NaCl、15mM 柠檬酸三钠)、50mM 磷酸钠 (pH7.6)、5 × Denhardt's 溶液、10% 硫酸葡聚糖和 20 微克/ml 剪切的变性鲑鱼精 DNA 的溶液中 42℃ 保温过夜，然后在约 65℃ 的 0.1 × SSC 中清洗杂交支持物。

- 杂交和清洗条件已众所周知，其实例可参考 Sambrook, et al., 分子克隆实验手册，第二版，Cold Spring Harbor, N.Y., (1989)，特别是其中的第 11 章。利用本发明的多核苷酸序列还可以进行溶液杂交。

- 本发明还提供一种多核苷酸，该多核苷酸含有或包括一种多核苷酸序列，该序列的获得方法是在严格杂交条件下用含有序列编号：1 或 3 所述多核苷酸序列或其片段的探针对含有序列编号：1 或 3 所述多核苷酸序列的完整基因的适当库进行筛选；然后分离出所述多核苷酸序列。可用来获得这种多核苷酸的片段包括，例如，文中其他地方详细描述探针和引物。

- 举例来说，按照文中有关本发明的多核苷酸测定方法的描述，本发明的多核苷酸可作为 RNA、cDNA 和基因组 DNA 的杂交探针用来分离编码 BASB132 的全长 cDNA 和基因组克隆，并且可用来分离与 BASB132 基因具有高度同一性的其他基因的 cDNA 和基因组克隆，特别是与 BASB132 基因具有高度序列同一性的其他基因的 cDNA 和基因组克隆。这些探针一般含有至少 15 个核苷酸残基或碱基对。优选的是含有至少 30 个核苷酸残基或碱基对，也可以含有至少 50 个核苷酸残基或碱基对。特别优选的是含有至少 20 个但又少于 30 个核苷酸残基或碱基对的探针。

为了合成寡核苷酸探针，可通过筛选方法用序列编号：1 或 3 的 DNA 序列分离出 BASB132 基因的一段编码区。然后用含有与本发明所述基因

互补的序列的标记寡核苷酸来筛选 cDNA 库、基因组 DNA 库或 mRNA 库，以确定与探针杂交的库成分。

有几种该领域技术人员熟知的方法可用来获得全长 DNAs 或延伸短 DNAs，例如基于 cDNA 末端快速扩增 (RACE) 的方法 (可参考，例如，
5 Frohman, et al., PNAS USA 85:8998-9002, 1988)。近期的改进技术，例如 MarathonTM 技术 (Clontech Laboratories Inc.)，能够显著简化对较长 cDNAs 的搜索。MarathonTM 技术是利用由选定组织提取的 mRNA 制备 cDNAs 并将“衔接”序列连接在每个末端上。然后用基因特异性和衔接序列特异性寡核苷酸引物的组合进行核酸扩增 (PCR)，以扩增出 DNA
10 中“丢失的”5'末端。然后用“嵌套”引物，即设计成可与扩增产物退火的引物 (通常是能够与衔接序列的其他 3'区退火的一个衔接序列特异性引物和能够与选定基因序列的其他 5'区退火的一个基因特异性引物)，重复 PCR 反应。然后对该反应的产物进行 DNA 测序分析，再将产物直接与已有的 DNA 连接而产生完整序列，或利用新序列信息设计出
15 5'引物再进行一次单独的全长 PCR，由此构建出全长 DNA。

按照文中有关本发明的多核苷酸测定方法的描述，可将本发明的多核苷酸和多肽作为研究试剂和研究材料用来，例如，探索疾病的治疗和诊断方法，尤其是人类的疾病。

在本发明的多核苷酸中，来源于序列编号：1 或 3 所述序列的寡核苷酸可在文中描述的方法中使用，但优选的是用于 PCR，以确定感染组织
20 内的细菌是否转录完整的或部分的本文所述多核苷酸。应该了解的是，这些序列还可用来诊断病原体所处的感染期和感染类型。

本发明还提供可编码一种多肽的多核苷酸，该多肽为成熟蛋白加上氨基端的或羧基端的或成熟多肽内部的额外氨基酸 (例如当成熟形式具
25 有一条以上的多肽链时)。这些序列可以在蛋白由前体变为成熟形式的加工过程中起一定作用，也可以用于蛋白运输，或是延长或缩短蛋白的半寿期，或是有利于测定或产生蛋白的操作，等等。额外氨基酸可以象体内的通常情况那样由细胞酶去除，以产生成熟蛋白。

本发明还提供所述各种多核苷酸的互补多核苷酸。优选的是这些互
30 补多核苷酸与各自的多核苷酸完全互补。

一种或多种原序列与成熟多肽融合成的前体蛋白可以是非活性的状态。这些非活性前体在原序列被去除后一般会活化。在活化前可以将部

分或所有原序列去除。这些前体通常称为蛋白原。

除利用常规的 A、G、C、T/U 来表示核苷酸外，还可利用术语“N”来描述本发明的某些多核苷酸。“N”的含义是，在 DNA 或 RNA 序列的该指定位置可以是四种 DNA 核苷酸或 RNA 核苷酸中的任意一种，优选的是不出现下述情况，即按照正确可读框读码时，N 与邻近核苷酸的组合可导致该可读框过早产生终止密码子。

总之，由本发明所述多核苷酸编码的蛋白可以是成熟蛋白、成熟蛋白加前导序列（可称为前蛋白）、含有非前蛋白前导序列的一种或多种原序列的成熟蛋白前体、或前蛋白原，即含有一种前导序列和一种或多种原序列的蛋白原前体，这些序列通常在产生活性多肽和成熟多肽的加工过程中被去除。

一方面，本发明提供所述多核苷酸的治疗或预防性应用，特别是用于遗传免疫。

将本发明所述多核苷酸应用于遗传免疫时，优选的是采用一种适当的递送方法，例如直接把质粒 DNA 注射到肌肉中（Wolff et al., Hum Mol Genet (1992) 1:363; Manthorpe et al., Hum. Gene Ther. (1983) 4:419）、递送与特定蛋白载体复合的 DNA（Wu et al., J Biol Chem. (1989) 264:16985）、DNA 与磷酸钙共沉淀（Benvenisty & Reshef, PNAS USA, (1986) 83:9551）、将 DNA 封装在不同形式的脂质体中（Kaneda et al., Science (1989) 243:375）、颗粒轰击（Tang et al., Nature (1992) 356:152, Eisenbraun et al., DNA Cell Biol (1993) 12:791）以及用克隆化逆转录病毒载体进行体内感染（Seeger et al., PNAS USA (1984) 81:5849）。

载体、宿主细胞、表达系统

本发明还涉及含有本发明的一种或多种多核苷酸的载体、用本发明的载体改造的基因工程宿主细胞，以及由重组技术产生本发明所述多肽的方法。也可采用无细胞翻译系统，利用由本发明所述 DNA 构建体获得的 RNAs 产生这些蛋白。

本发明的重组多肽可采用该领域技术人员所熟知的方法，由包含表达系统的基因工程宿主细胞来加以制备。因此，另一方面，本发明涉及含有本发明的一种或多种多核苷酸的表达系统、含有这些表达系统的基因工程宿主细胞，以及由重组技术产生本发明所述多肽的方法。

在利用重组方法产生本发明的多肽时，可对宿主细胞进行基因工程改造，以掺入表达系统或其部分或本发明的多核苷酸。将多核苷酸引入宿主细胞可采用 Davis, et al., 分子生物学基本方法, (1986) 和 Sambrook et al., 分子克隆实验手册, 第二版, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989) 等常规实验手册中描述的方法，如磷酸钙转染、DEAE-葡聚糖介导的转染、转位、显微注射、阳离子脂质介导的转染、电穿孔、转导、摩擦引入、冲击引入和感染。

适当宿主的代表性实例包括，细菌细胞，如链球菌、葡萄球菌、肠球菌、大肠杆菌、链霉菌、蓝细菌、枯草芽孢杆菌、脑膜炎奈瑟菌和粘膜炎莫拉氏菌；真菌细胞，如酵母菌、克鲁维酵母、酵母属、担子菌、白色念珠菌和曲霉菌；昆虫细胞，如果蝇 S2 细胞和秋粘虫 Sf9 细胞；动物细胞，如 CHO、COS、HeLa、C127、3T3、BHK、293、CV-1 和 Bowes 黑素瘤细胞；以及植物细胞，如裸子植物或被子植物的细胞。

有很多种表达系统可用来产生本发明的多肽。这些载体尤其包括来源于染色体、游离基因或病毒的载体，例如，来源于细菌质粒、噬菌体、转座子、酵母游离基因、插入元件、酵母染色体元件的载体、来源于诸如杆状病毒、乳头多瘤空泡病毒如 SV40、痘苗病毒、腺病毒、鸡痘病毒、假狂犬病病毒、小核糖核酸病毒、逆转录病毒和甲病毒等病毒的载体，以及来源于其组合物的载体，如粘粒和噬菌粒等来源于质粒与噬菌体遗传因子的载体。表达系统构建体可含有能够引发并调控表达的调控区。一般而言，能够维持、增殖或表达多核苷酸并且（或者）在宿主内表达多肽的任何适当系统或载体均可用于该表达。将适当 DNA 序列插入表达系统的方法可采用多种众所周知的常规技术中的任何一种，如 Sambrook et al., 分子克隆实验手册, (见上文) 中描述的技术。

在真核细胞的重组表达系统中，可将适当的分泌信号掺入到表达的多肽内，以使翻译的蛋白能够分泌到内质网腔、外周质间隙或细胞外环境中。这些信号可以是多肽的内源性信号，也可以是外源性信号。

由重组细胞培养物回收和纯化本发明所述多肽的方法已众所周知，其中包括硫酸铵或乙醇沉淀法、酸抽提法、阴离子或阳离子交换层析法、磷酸纤维素层析法、疏水相互作用层析法、亲和层析法、羟磷灰石层析法和外源凝集素层析法。最优选的是用金属离子亲和层析法 (IMAC) 来进行纯化。若多肽在胞内合成、分离和（或）纯化的过程中变性，可利

用众所周知的蛋白重折叠技术来产生活性构象。

表达系统也可以是病毒或细菌等活体重组微生物。可将目的基因插入到活体重组病毒或细菌的基因组中。用这种活体载体进行接种或体内感染会使抗原在体内表达，并且可诱导免疫应答。可用于此目的的病毒和细菌包括，例如，痘病毒（如痘苗病毒、鸡痘病毒、金丝雀痘病毒）、甲病毒（辛德毕斯病毒、塞姆利基森林病毒、委内瑞拉马脑脊髓炎病毒）、腺病毒、腺相关病毒、小核糖核酸病毒（脊髓灰质炎病毒、鼻病毒）、疱疹病毒（水痘带状疱疹病毒等）、李斯特杆菌、沙门菌、志贺氏菌、卡介菌。这些病毒和细菌可以是毒性的，为了获得活体疫苗，也可以使用由不同方式减毒的病毒和细菌。这些活体疫苗也构成本发明的一部分。

诊断、预防、血清型及突变测定

本发明还涉及本发明所述 BASB132 多核苷酸及多肽在作为诊断试剂方面的应用。在真核细胞中，特别是哺乳动物细胞，尤其是人类细胞中，检测 BASB132 多核苷酸和（或）多肽可作为诊断疾病、病期或被感染生物对药物的反应的一种方法。利用众所周知的技术以及文中提供的方法，可以在核酸或氨基酸水平上来检测真核细胞，特别是哺乳动物，尤其是人类，特别是被含有 BASB132 基因或蛋白的生物感染的人或怀疑被感染的人。

用于预后、诊断或其他分析的多肽和多核苷酸可由怀疑被感染的个体和（或）感染个体的身体材料获得。由任意的这些来源获得的多核苷酸，特别是 DNA 或 RNA，可直接用于检测，也可以在分析前先用酶学方法由 PCR 或其他任何扩增技术进行扩增。RNA，特别是 mRNA，以及 cDNA 和基因组 DNA 也可同样使用。对生物的选定多核苷酸进行扩增和基因型分析，由此可鉴定个体中存在的感染性或定居生物的物种和品系。将扩增产物与选自相关生物的参考序列基因型进行比较，优选的是与选自同属不同种或同种不同株系的生物的参考序列基因型进行比较，由其尺寸的变化可检测缺失和插入。将扩增的 DNA 与标记的 BASB132 多核苷酸序列杂交，由此可检测点突变。分别用 DNase 或 RNase 对 DNA 或 RNA 进行消化，或是检测解链温度或复性动力学的差异，由此可以从不完全匹配甚至明显错配的双螺旋体中鉴别出完全匹配或明显匹配的序

列。将多核苷酸片段在凝胶中的电泳迁移率与参考序列进行比较，由其变化可同样检测出多核苷酸序列的差异。这种检测方法可使用变性剂，也可以不使用变性剂。直接对 DNA 或 RNA 测序也可以检测多核苷酸的差异。可参考，例如，Myers et al., Science, 230:1242 (1985)。利用核酸酶保护测定法，如 RNase、V1 及 S1 保护测定法，或化学裂解法还可以检测特定部位的序列变化。可参考，例如，Cotton et al., Proc.Natl.Acad.Sci., USA, 85:4397-4401 (1985)。

在另一项实施方案中，可将含有 BASB132 核苷酸序列或其片段的寡核苷酸探针排成阵列，以有效地筛选，例如，基因突变、血清型、系统分类或鉴定。有关阵列技术的方法已众所周知，这些方法具有广泛的适用性，可用来解决多种分子遗传学问题，其中包括基因表达、遗传连锁和遗传变异性（可参考，例如，Chee et al., Science, 274:610 (1996)）。

另一方面，本发明涉及一种诊断试剂盒，该试剂盒含有：

(a) 本发明的一种多核苷酸，优选的是序列编号：1 或 3 的核苷酸序列，或其片段；

(b) 与(a)互补的核苷酸序列；

(c) 本发明的一种多肽，优选的是序列编号：2 或 4 的多肽，或其片段；或

(d) 抗本发明所述多肽的一种抗体，优选的是抗序列编号：2 或 4 的多肽。

应该了解的是，所有这些试剂盒中的(a)、(b)、(c)或(d)均可以包含一种基本成分。这种试剂盒尤其适用于疾病或疾病易感性的诊断。

本发明还涉及所述多核苷酸在作为诊断试剂方面的应用。对本发明中与疾病或致病性相关的多核苷酸的突变形式进行检测的方法，优选的是对序列编号：1 或 3 所示多核苷酸的突变形式进行检测的方法，可作为一种诊断工具，该方法可以由这种多核苷酸的低表达、过表达或表达水平变化的结果来对疾病诊断、病程预后、病期或疾病易感性判定进行补充或说明。有多种技术，如文中描述的方法，可以在多核苷酸水平上检测所述多核苷酸中含有突变的生物，特别是感染性生物。

还可以用多种技术对本发明所述多核苷酸和（或）多肽中含有突变或多态性（等位变异）的生物细胞进行检测，由此为，例如，血清分型做准备。举例来说，RT-PCR 方法可检测 RNA 中的突变。特别优选的是

将 RT-PCR 方法与诸如 GeneScan 等自动检测系统结合使用。为此也可以在 PCR 中使用 RNA、cDNA 或基因组 DNA。例如，可利用与 BASB132 多肽的编码多核苷酸互补的 PCR 引物来对突变进行鉴定和分析。

本发明还提供 5' 和 (或) 3' 端被去除 1、2、3 或 4 个核苷酸的引物。

- 5 这些引物尤其适用于扩增那些由个体的样品，如身体材料，分离的 BASB132 DNA 和 (或) RNA。这些引物可用来扩增那些分离自感染个体的多核苷酸，继而可通过不同的技术来阐明该多核苷酸的序列。这样就可以检测该多核苷酸序列中的突变，并可将其用于感染、感染期或感染过程的诊断和 (或) 预后，或用于感染物的血清分型和 (或) 分类。
- 10 本发明还提供一种方法，用于诊断疾病，优选的是细菌性感染，更优选的是由粘膜炎莫拉氏菌引起的感染，该方法包括，测定具有序列编号：1 或 3 所述序列的多核苷酸在个体的样品，如身体材料，中的表达水平是否提高。BASB132 多核苷酸的表达增加或降低情况可以用该领域熟知的任何一种多核苷酸定量方法来检测，如扩增、PCR、RT-PCR、RNase
- 15 保护、RNA 印迹、光谱测定和其他杂交方法。

- 此外，本发明的诊断测定法可检测 BASB132 多肽相对于正常对照组织样品的过表达，该方法可检测，例如，是否存在一种感染。可检测宿主样品，如身体材料，中的 BASB132 多肽含量的测定技术已为该领域技术人员所熟知。这些测定方法包括，放射免疫测定法、竞争性结合测定
- 20 法、蛋白印迹分析法、抗体夹层测定法、抗体检测法和 ELISA 测定法。

- 本发明的多核苷酸可用作多核苷酸阵列的成分，优选的是用于高密度阵列或网格。这些高密度阵列特别适用于诊断和预后。举例来说，可利用各含有不同基因并且另含有本发明的一种或多种多核苷酸的一组点来进行探查，如利用一种来自身体样品的探针进行杂交或核酸扩增，以
- 25 确定个体中是否存在特定的多核苷酸序列或相关序列。如果存在，则表明存在病原体，特别是存在粘膜炎莫拉氏菌，这种方法适用于疾病或病程的诊断和 (或) 预后。优选的网格可含有序列编号：1 或 3 所述多核苷酸序列的多种变异体。含有一种多核苷酸序列的多种变异体的网格也同样优选，这种多核苷酸序列可编码序列编号：2 或 4 的多肽序列。

30

抗体

本发明的多肽和多核苷酸，或其变异体，或表达这些多肽和多核苷

酸的细胞可作为免疫原用来产生分别对这些多肽或多核苷酸具有免疫特异性的抗体。术语“免疫特异性”的含义是，这些抗体对本发明所述多肽的亲和力基本上高于对先有技术中其他相关多肽的亲和力。

在某些优选实施方案中，本发明提供抗 BASB132 多肽或多核苷酸的
5 抗体。

产生抗本发明所述多肽或多核苷酸的抗体的方法是，利用常规方案将本发明的多肽和（或）多核苷酸，或二者或二者之一的带有表位的片段，或二者或二者之一的类似物，或表达二者或二者之一的细胞，向动物给药，优选的是向非人类的动物给药。单克隆抗体的制备可采用该领域已知的任何能够由连续细胞系培养物产生抗体的技术。其实例包括多种技术，例如 Kohler, G. and Milstein, C., *Nature* 256:495-497 (1975) ; Kozbor et al., *Immunology Today* 4:72 (1983) ; 《单克隆抗体与癌症治疗》，Alan R. Liss, Inc. (1985) 的 Cole et al., 第 77-96 页中描述的技术。

可采用单链抗体生产技术（美国专利号 4,946,778）来产生抗本发明
15 所述多肽或多核苷酸的单链抗体。也可利用转基因小鼠或其他生物或动物，如其他哺乳动物，来表达对本发明所述多肽或多核苷酸具有免疫特异性的人源化抗体。

作为选择，也可以用噬菌体展示技术来筛选与本发明所述多肽具有结合活性的抗体基因，筛选可以在具有抗 BASB132 特性的人淋巴细胞的
20 PCR 扩增 v-基因库中进行，也可以在天然库中进行（McCafferty et al., (1990) *Nature* 348, 552-554; Marks, et al., (1992) *Biotechnology* 10, 779-783）。还可以通过，例如，链改组方法来提高这些抗体的亲和力（Clackson et al., (1991) *Nature* 352:628）。

上述抗体可用来分离或鉴定那些可表达本发明所述多肽或多核苷酸的克隆，从而通过，例如，亲和层析来纯化这些多肽或多核苷酸。
25

因此，抗 BASB132 多肽或 BASB132 多核苷酸的抗体尤其可用来治疗感染，特别是细菌性感染。

在抗原性、表位或免疫学方面等价的多肽变异体构成本发明的一个特别方面。

30 优选的是对抗体或其变异体进行修饰，使其在个体中的免疫原性降低。举例来说，若个体为人类，则最优选的是将抗体“人源化”，例如可按照 Jone et al. (1986) *Nature* 321, 522-525 或 Tempest et al., (1991)

Biotechnology 9, 266-273 所述,将来源于杂交瘤的抗体的互补决定区移植到人单克隆抗体中。

拮抗剂和激动剂——测定和分子

5 本发明的多肽和多核苷酸还可用来评定,例如,细胞、无细胞制剂、化学库和天然产物混合物中的小分子底物与配体的结合。这些底物和配体可以是天然的,也可以是结构或功能模拟物。可参考,例如, Coligan et al., Current Protocols in Immunology 1(2): 第5章(1991)。

10 筛选方法可以是,通过与候选化合物直接或间接相连的标记,简单地测量该候选化合物与多肽或多核苷酸、带有多肽或多核苷酸的细胞或膜、或多肽的融合蛋白的结合。作为选择,筛选方法也可涉及与标记竞争物的竞争。此外,还可通过这些筛选方法,利用适合于含有本发明所述多肽或多核苷酸的细胞的检测系统,来测试候选化合物能否随多肽或多核苷酸的活化或抑制而产生一种信号。一般而言,在含有已知激动剂
15 的情况下可测定活化的抑制剂,而在含有候选化合物的情况下可测定激动剂对活化作用的影响。在不含有激动剂或抑制剂时,可利用组成型活性多肽和(或)组成型表达的多肽和多核苷酸来筛选反向激动剂或抑制剂,方法是测试候选化合物能否对多肽或多核苷酸的活化产生抑制作用,其结果需视情况而定。此外,该筛选方法的步骤可简单地包括,将
20 候选化合物与含有本发明所述多肽或多核苷酸的一种溶液相混合,以形成一种混合物,测量混合物中 BASB132 多肽和(或)多核苷酸的活性,并将测量结果与标准值进行比较。还可利用上文所述的融合蛋白,例如 Fc 部分与 BASB132 多肽的融合蛋白,来进行高通量的筛选测定,以鉴别本发明所述多肽的拮抗剂以及系统发生相关多肽或功能相关多肽的拮抗
25 剂(参考 D. Bennett et al., J Mol Recognition, 8:52-58(1995); 和 K. Johanson et al., J Biol Chem, 270(16):9459-9471(1995))。

还可利用本发明的多核苷酸、多肽以及能够与本发明所述多肽结合和(或)相互作用的抗体来设计筛选方法,以检测添加的化合物对细胞产生 mRNA 和(或)多肽的影响。例如可设计一种 ELISA 测定方法,用
30 于通过该领域已知的标准方法由单克隆抗体和多克隆抗体对多肽分泌水平或细胞结合水平进行测量。该方法可用于从适当处理的细胞或组织中发现一些可抑制或提高多肽产量的作用剂(分别称为拮抗剂或激动剂)。

本发明还提供一种筛选化合物的方法，用于鉴别那些可提高（激动剂）或阻断（拮抗剂）BASB132多肽或多核苷酸的功能的化合物，特别是抑菌性和（或）杀菌性化合物。该筛选方法可涉及高通量技术。举例来说，为了筛选激动剂或拮抗剂，可在不含或含有候选分子的情况下，

5 即不含或含有BASB132的可能激动剂或拮抗剂的情况下，将含有BASB132多肽及其标记底物或标记配体的合成反应混合物、细胞间隔，如细胞膜、细胞被膜或细胞壁，或其制剂进行保温。标记配体的结合量下降情况或标记底物的产物的产量下降情况可反映出候选分子促进或拮抗BASB132多肽的能力。最可能的优良拮抗剂是能够无偿结合的分子，

10 也就是在不影响BASB132多肽情况下发生结合的分子。能够良好结合并且，例如，提高底物生成产物的速率、加强信号转导，或提高化学通道活性的分子为激动剂。利用报告系统可促进对，例如，底物生成产物的速率、信号转导水平或化学通道活性水平的检测。可用于此方面的报告系统包括，但不局限于，该领域已知的比色法、测定标记底物向产物的

15 转化、可反映BASB132多核苷酸或多肽活性变化的报告基因，以及结合测定法。

测定BASB132激动剂的另一实例为竞争性测定法，该方法是在适当条件下将BASB132和一种潜在的激动剂与能够结合BASB132或重组BASB132的分子、天然底物或配体，或底物模拟物或配体模拟物相混合，

20 并进行竞争性抑制测定。可以用放射性或比色分析化合物等将BASB132标记，以便精确测量与结合分子结合的或转化成产物的BASB132分子的数量，由此来评定潜在拮抗剂的效力。

潜在的拮抗剂包括，例如，能够与本发明的多核苷酸和（或）多肽结合从而抑制或消除其活性或表达的有机小分子、肽、多肽和抗体。潜

25 在的拮抗剂还包括，能够与结合分子的相同位点结合但不产生由BASB132诱导的活性的有机小分子、肽、多肽如密切相关的蛋白，或抗体，它们使BASB132多肽和（或）多核苷酸不能进行结合，由此阻碍其功能或表达。

潜在的拮抗剂还包括一种小分子，这种小分子可结合并占据多肽结

30 合位点，从而阻碍其与细胞结合分子结合，继而阻碍正常生物学活性。这些小分子的实例包括，但不局限于，有机小分子、肽或肽样分子。其他潜在的拮抗剂还包括反义分子（有关这些分子的描述可参考 Okano, J.

Neurochem. 56: 560 (1991); 利用寡脱氧核苷酸作为基因表达的反义抑制剂, CRC Press, Boca Raton, FL (1988))。优选的潜在拮抗剂包括, 与 BASB132 及其变异体相关的化合物。

另一方面, 本发明涉及可溶性基因工程融合蛋白, 这些融合蛋白含有本发明的多肽或其片段, 并含有不同亚类免疫球蛋白 (IgG、IgM、IgA、IgE) 的重链或轻链恒定区的不同部分。优选的免疫球蛋白是人类 IgG 的重链恒定区, 特别是 IgG1 的重链恒定区, 并且融合出现在铰链区。在一项特别实施方案中, 去除 Fc 部分的方法只是简单地掺入一种能被凝血因子 Xa 剪切的酶切序列。此外, 本发明还涉及通过基因工程方法制备这些融合蛋白的方法, 以及这些融合蛋白在药物筛选、诊断和治疗中的应用。另一方面, 本发明涉及编码这些融合蛋白的多核苷酸。有关融合蛋白技术的实例可参考国际专利申请 WO94/29458 和 WO94/22914。

文中提供的每种多核苷酸序列都可用来开发和研制抗菌性化合物。这些多核苷酸被表达后, 其编码的蛋白可作为标的物用于抗菌药物的筛选。此外, 还可利用编码这些蛋白的氨基端区域的多核苷酸序列, 或编码相应 mRNA 的 Shine-Delgarno 序列或其他促翻译序列的多核苷酸序列来构建反义序列, 以调控目的编码序列的表达。

本发明还提供所述多肽、多核苷酸、激动剂或拮抗剂的一种应用, 即用来干扰一种或多种病原体与导致感染后遗症的真核宿主之间的最初物理相互作用, 优选的是与哺乳动物宿主之间的最初物理相互作用。详细地说, 本发明的这些分子可用来: 阻碍细菌, 特别是革兰氏阳性菌和 (或) 革兰氏阴性菌, 粘附于留置装置上或伤口内的真核细胞外基质蛋白, 优选的是哺乳动物细胞外基质蛋白; 阻断真核细胞外基质蛋白, 特别是哺乳动物细胞外基质蛋白, 与介导组织损伤的细菌 BASB132 蛋白之间的细菌性吸附; 以及 (或者) 阻断感染的正常发病进程, 这些感染由留置装置移植或其他外科技术以外的因素引起。

再一方面, 本发明提供 BASB132 的激动剂和拮抗剂, 优选的是抑菌性或杀菌性激动剂和拮抗剂。

本发明的拮抗剂和激动剂可用于, 例如, 疾病的预防、抑制和 (或) 治疗。

另一方面, 本发明涉及本发明所述多肽的拟表位(mimotope)。拟表位是指与天然肽 (在序列或结构上) 非常相似的一种肽序列, 它可被能

够识别天然肽的抗体所识别；或可诱发出一些抗体，这些抗体与适当载体结合后可识别天然的肽。

通过特定氨基酸的添加、缺失或取代可设计出用于特定目的的肽拟表位。因此，可以对肽进行，使其便于与一种蛋白载体结合。例如，对一些化学结合方法而言，希望在肽的末端引入一个半胱氨酸。而对与蛋白载体结合的肽而言，希望在肽的结合端的远端引入一种疏水末端，使该肽的未结合游离末端能与载体蛋白表面保持结合。这样就可获得在构象上与完整天然肽分子最为接近的肽。举例来说，可以对肽进行改造，使其具有一个 N-端半胱氨酸和一种疏水的 C-端酰胺化尾。作为选择，也可以添加或取代成一个或多个 D-型氨基酸异构体，以产生有利的衍生物，例如可提高肽的稳定性的衍生物。

作为选择，还可通过噬菌体展示等技术（EP0552267B1），用自身能够与本发明的多肽结合的抗体来鉴别肽拟表位。利用该技术可产生大量能够模拟天然肽结构、从而能够与抗-天然肽的抗体结合的肽序列，但这些肽序列本身不一定与天然多肽有明显的序列同源性。

疫苗

另一方面，本发明涉及一种方法，用于在个体中，特别是在哺乳动物中，优选的是在人体中，诱导一种免疫应答，该方法包括，用足以产生抗体和（或）T 细胞免疫应答的 BASB132 多核苷酸和（或）多肽或其片段或变体对个体进行接种，以保护该个体不受感染，特别是不受细菌性感染，最明确地说是不受粘膜莫拉氏菌感染。另外还提供通过这种免疫应答来延缓细菌复制的方法。再一方面，本发明还涉及一种在个体中诱导免疫应答的方法，该方法包括，用一种可指导 BASB132 多核苷酸和（或）多肽或其片段或变体表达的核酸载体、核酸序列或核酶向所述个体给药，使 BASB132 多核苷酸和（或）多肽或其片段或变体能够在体内表达，以便诱导一种免疫应答，如抗体免疫应答和（或）T 细胞免疫应答，其中包括，例如，产细胞因子 T 细胞或细胞毒性 T 细胞，从而无论该个体是否已形成该疾病，都可以保护该个体免得疾病，优选的是保护人类免得疾病。基因给药方法的实例是将基因作为颗粒的包被或以其他方式快速轰击到目的细胞内。所述核酸载体包括 DNA、RNA、核酶、修饰的核酸、DNA/RNA 杂合体、DNA-蛋白复合物或 RNA-蛋白

复合物。

另一方面，本发明涉及一种免疫组合物，该组合物被引入一种能够在体内产生免疫应答的个体时，优选的是被引入人体时，可在该个体内诱导一种抗 BASB132 多核苷酸和（或）其编码的多肽的免疫应答，其中该组合物含有一种重组 BASB132 多核苷酸和（或）其编码的多肽，以及（或者）可编码或表达该 BASB132 多核苷酸、其编码的多肽，或本发明的其他多肽抗原的 DNA 和（或）RNA。这种免疫应答可用于治疗和预防，并且可采用抗体免疫和（或）细胞免疫的形式，例如由 CTL 或 CD4+ T 细胞产生的细胞免疫。

10 BASB132 多肽或其片段可以与辅助蛋白或化学半分子(moiety)相融合，这些辅助蛋白或化学半分子可以自身产生抗体，也可以自身不产生抗体，它们可稳定第一蛋白，并且可产生一种具有抗原性和（或）免疫原性的融合或修饰蛋白，优选的是产生一种具有保护特性的融合或修饰蛋白。优选的是这种融合重组蛋白另含有一种抗原性辅助蛋白，如来源于流感嗜血菌的脂蛋白 D、谷胱甘肽-S-转移酶（GST）或 β -半乳糖苷酶，15 或其他任何能够稳定蛋白并利于其产生和纯化的较大的辅助蛋白。此外，辅助蛋白可以对接受该蛋白的生物的免疫系统进行广泛刺激，因而可作为一种佐剂。辅助蛋白可以联接在第一蛋白的氨基端或羧基端。

在本发明的疫苗组合物中，BASB132 多肽和（或）多核苷酸、或其片段或拟表位或变异体可以由一种载体来提供，如上文描述的活细菌载体等活体重组载体。

BASB132 多肽也可使用适当的非活体载体，例如细菌的外膜小泡或“大泡”。OM 大泡来源于革兰氏阴性菌双层膜的外膜，据资料显示，OM 大泡在多种革兰氏阴性菌中都存在（Zhou, L et al. 1998. FEMS Microbiol. Lett. 163:223-228），其中包括沙眼衣原体和鹦鹉热衣原体。据25 报道，可产生大泡的部分细菌变异体包括：百日咳博德特氏菌、博氏疏螺旋体、马尔他布氏杆菌、绵羊布氏杆菌、大肠杆菌、流感嗜血菌、嗜肺军团菌、粘膜炎莫拉氏菌、淋病奈瑟球菌、脑膜炎奈瑟菌、铜绿假单胞菌和小肠结肠炎耶尔森菌。

30 大泡的优点是能提供天然构象的外膜蛋白，因而特别适用于疫苗。在疫苗应用方面，还可以对大泡进行改良，方法是对细菌进行改造，以改变外膜的一种或多种分子的表达。举例来说，利用这种方法可引入

BASB132 多肽等所需免疫原性蛋白或对其表达进行增量调节（如改变启动子）。另外，还可以对无关外膜分子（如非保护性抗原或可变的优势免疫蛋白）或有害外膜分子（如 LPS 等毒性分子或潜在的自身免疫应答诱导剂）的表达进行减量调节。这些方法将在下文做更详细地描述。

- 5 BASB132 基因的非编码侧翼区含有对该基因表达十分重要的调控元件。这种调控可以在转录和翻译水平上进行。这些区域无论处在基因可读框的上游还是下游，都可通过 DNA 测序来获得其序列。该序列信息能用来确定可能的调控基元，如不同的启动子元件、终止子序列、可诱导序列元件、阻抑物、负责相变异的元件、SD 序列、具有潜在二级结构并参与调控的区域，以及其他类型的调控基元或序列。该序列构成本发明的另一方面。

- 15 该序列信息可用来调变 BASB132 基因的天然表达。基因表达的增量调节可通过改变启动子、SD 序列、潜在的阻抑物或操纵子序列或其他任何有关元件来实现。也可通过类似的修饰来实现对表达的减量调节。作为选择，也可以改变相变异序列，使基因表达受相变异的调控或是脱离这种调控。另一方法是用一种或多种能够实现受控表达的可诱导元件来控制基因的表达。这种调控的实例包括，但不局限于，用温度的改变进行诱导、添加诱导剂底物，如选定的糖类或其衍生物、微量元素、维生素、辅助因子、金属离子，等等。

- 20 有几种不同的方法可用来引入上述修饰。对参与基因表达的序列进行修饰的体内实现方法是，进行随机诱变，然后筛选预期表型。另一种方法的步骤包括，分离出目的区域，并通过随机诱变或定点取代、插入或缺失诱变对其进行修饰。然后用同源重组方法将修饰区域重新引入细菌基因组，并评定对基因表达的影响。另一种方法是，根据目的区域的序列信息，将所有或部分天然调控序列替换或去除。这种方法需要分离出靶调控区并进行修饰，使其含有另一种基因的调控元件、来自不同基因的组合调控元件、合成调控区或其他任何调控区，或将其野生型调控序列的选定部分去除。然后用同源重组方法将这些修饰序列重新引入细菌基因组。可用于对基因表达进行增量调节的部分优选启动子包括，来
25 源于脑膜炎奈瑟菌或淋病奈瑟球菌的启动子 porA、porB、lbpB、tbpB、p110、1st、hpuAB；ompCD、copB、lbpB、ompE、UspA1；UspA2；来源于粘膜炎莫拉氏菌的 TbpB；来源于流感嗜血菌的 p1、p2、p4、p5、p6、
30

lpD、tbpB、D15、Hia、Hmw1、Hmw2。

5 在一项实施例中，基因表达的调变可通过将其启动子与一种强启动子互换的方法来实现（包括分离该基因的上游序列，对该序列进行体外修饰，并用同源重组方法将其重新引入基因组）。在细菌及细菌脱落（或产生）的外膜小泡中均可以实现对表达的增量调节。

10 其他实施例中描述的方法可用来产生一些重组细菌菌株，这些重组细菌菌株的特性更适合于疫苗应用。这些菌株包括，但不局限于，减毒菌株、选定抗原的表达水平增强的菌株、干扰免疫应答的基因被敲除（或表达水平降低）的菌株、优势免疫蛋白的表达被调变的菌株、外膜小泡的脱落被调变的菌株。

15 因此，本发明还提供一种修饰的 BASB132 基因上游区域，该区域含有一种可改变外膜 BASB132 蛋白表达水平的异种调控元件。该区域构成本发明的一个方面，其中包括 BASB132 基因的上游序列。该区域从紧接着 BASB132 基因的上游位置开始，通常持续到该基因上游的一个离 ATG 起始密码子不足约 1000bp 的位置。当基因处在多顺反子序列（操纵子）中时，该区域可以从紧接着该基因的上一位置开始，也可以从紧接着该操纵子的第一个基因的上一位置开始。优选的是，这种修饰的上游区域在 ATG 上游 500-700bp 之间含有一种异种启动子。

20 因此，本发明提供一种位于修饰的细菌大泡内的 BASB132 多肽。本发明还提供改良的宿主细胞，这些宿主细胞可产生非活性膜泡载体。本发明还提供含有 BASB132 基因的核酸载体，该 BASB132 基因含有一段修饰的上游区域，该区域中含有一种异种调控元件。

本发明还提供所述宿主细胞及细菌大泡的制备方法。

25 本发明还提供一些组合物，特别是疫苗组合物，以及这些组合物的制备方法，这些组合物中含有本发明的多肽和（或）多核苷酸以及免疫刺激性 DNA 序列，如 Sato, Y. et al. Science 273:352(1996)中描述的 DNA 序列。

30 此外，本发明还提供一些方法，这些方法是将已证明能够编码细菌表面蛋白非可变区的所述多核苷酸或其特定片段构建多核苷酸构建体，并利用这些构建体在粘膜莫拉氏菌感染的动物模型中进行遗传免疫实验。这些实验特别适用于鉴别那些可激发预防性或治疗性免疫应答的蛋白表位。相信该方法还可随后用于从成功抵抗或清除感染的动物的

必需器官中制备具有特定用途的单克隆抗体，以获得可用于哺乳动物，特别是人类的细菌性感染，特别是粘膜炎症莫拉氏菌感染的预防药或治疗方法。

5 本发明还提供一种疫苗制剂，该疫苗制剂含有本发明的一种免疫原性重组多肽和（或）多核苷酸，同时含有一种适当的载体，如药学容许的载体。由于多肽和多核苷酸在胃内可被降解，因而优选的是均以胃肠外方式给药，包括皮下给药、肌内给药、静脉内给药或真皮内给药。适于胃肠外给药的制剂包括水性和非水性的无菌注射液，其中可含有抗氧化剂、缓冲液、抑菌化合物以及使制剂与个体的体液等渗的溶液，优选的是与血液等渗的溶液；以及水性和非水性的无菌悬浮液，其中可含有悬浮剂或增稠剂。该制剂可以在诸如密封安瓿瓶和小药水瓶等单位剂量或多剂量容器中提供，并且可在冻干条件下保存，在使用前只需直接添加无菌流质载体。

15 为了提高制剂的免疫原性，本发明的疫苗制剂还可含有佐剂系统。优选的是用优选佐剂系统来诱发 TH1 型应答。

免疫应答可大致分为两大类，即体液介导的免疫应答和细胞介导的免疫应答（传统的鉴定方法是保护机制分别为抗体效应剂和细胞效应剂）。这两类应答被称为 TH1 型应答（细胞介导的应答）和 TH2 型免疫应答（体液应答）。

20 TH1 型免疫应答的特征在于，可产生抗原特异性单倍体细胞毒性 T 淋巴细胞以及天然杀伤细胞应答。小鼠 TH1 型应答的特征是产生 IgG2a 亚型抗体，而相应的人类 TH1 型应答则产生 IgG1 型抗体。TH2 型免疫应答的特征是产生广泛的同种免疫球蛋白，小鼠中产生的免疫球蛋白包括 IgG1、IgA 和 IgM。

25 细胞因子被认为是产生这两种免疫应答类型的驱动力。高水平的 TH1 型细胞因子倾向于产生针对给定抗原的细胞介导的免疫应答，而高水平的 TH2 型细胞因子倾向于产生针对抗原的体液免疫应答。

TH1 型免疫应答与 TH2 型免疫应答的区别并不是绝对的。实际上，个体可维持 TH1 占优势的免疫应答，也可以维持 TH2 占优势的免疫应答。但为了方便，通常是根据 Mosmann 和 Coffman (Mosmann, T.R. and Coffman, R.L. (1989) TH1 与 TH2 细胞：不同的淋巴因子分泌方式导致不同的功能特性。Annual Review of Immunology, 7, p145-173) 对鼠 CD4+

ve T 细胞克隆的描述来研究细胞因子家族。根据传统观点, TH1 型应答与 T-淋巴细胞产生的 INF- γ 和 IL-2 细胞因子有关。但是, 常与诱导 TH1 型免疫应答直接相关的其他细胞因子不由 T 细胞产生, 如 IL-12。相比之下, TH2 型应答与 IL-4、IL-5、IL-6 和 IL-13 的分泌有关。

5 已知某些疫苗佐剂特别适用于刺激 TH1 型或 TH2 型细胞因子应答。能够在接种或感染后显示免疫应答 TH1:TH2 平衡的最佳方法通常包括, 在用抗原重新刺激后, 直接对 T 淋巴细胞产生的 TH1 或 TH2 细胞因子进行体外测量, 以及 (或者) 对抗原特异性抗体应答的 IgG1:IgG2a 比率进行测量。

10 因此, 用抗原对分离的 T 细胞群体重新进行体外刺激时, TH1 型佐剂可优先刺激这些 T 细胞产生高水平的 TH1 型细胞因子, 并且可促进 CD8⁺ 细胞毒性 T 淋巴细胞的产生以及与 TH1 型相关的抗原特异性同种免疫球蛋白应答的产生。

可优先刺激 TH1 型细胞应答的佐剂在国际专利申请 WO94/00153 和
15 WO95/17209 中有所描述。

3-脱氧-酰化单磷酸脂 A (3D-MPL) 就是这样一种佐剂。它公开于 GB2220211 (Ribi)。从化学角度来看, 它是带有 4、5 或 6 个酰化链的 3-脱氧-酰化单磷酸脂 A 的混合物, 这种混合物由 Ribi Immunochem, Montana 制造。3-脱氧-酰化单磷酸脂 A 的优选形式公开于欧洲专利
20 0689454B1 (SmithKline Beecham Biologicals SA)。

优选的是, 3D-MPL 的颗粒足够小, 使其能够通过 0.22 微米滤膜过滤除菌 (欧洲专利 0689454)。3D-MPL 的提供量为每剂 10 μ g-100 μ g, 优选的是每剂 25-50 μ g, 其中抗原的提供量一般为每剂 2-50 μ g。

另一种优选佐剂包括 QS21, 它是用 HPLC 方法由 Quillaja Saponaria
25 Molina 的树皮纯化出的一种无毒成分。作为选择, 可将 QS21 与 3-脱氧-酰化单磷酸脂 A (3D-MPL) 相混合, 也可选择再与一种载体混合。

美国专利 5,057,540 公开了产生 QS21 的方法。

有关含有 QS21 的反应原性佐剂, 前人已有所描述 (WO96/33739)。当与一种抗原混合时, 这些含有 QS21 和胆固醇的制剂表现为一种有效的
30 TH1 刺激佐剂。

可优先刺激 TH1 型细胞应答的其他佐剂还包括免疫调制寡核苷酸, 如 WO96/02555 中公开的未甲基化 CpG 序列。

还可考虑将如上所述的不同 TH1 刺激佐剂的组合作为优先产生 TH1 型细胞应答的刺激佐剂。例如 QS21 可以与 3D-MPL 配制在一起。

QS21:3D-MPL 的典型比率为 1:10 到 10:1; 优选的是 1:5 到 5:1, 通常情况下基本为 1:1。产生最佳协同作用的优选 3D-MPL:QS21 比率为 2.5:1

5 到 1:1。

优选的是, 本发明的疫苗组合物中还含有一种载体。该载体可以是一种水包油乳剂, 或一种铝盐, 如磷酸铝或氢氧化铝。

10 优选的水包油乳剂可含有一种可代谢的油, 如角鲨烯、 α -生育酚和 Tween80。在一项特别优选的实施方案中, 本发明所述疫苗组合物中的抗原可以在这种乳剂中与 QS21 和 3D-MPL 混合。此外, 该水包油乳剂可含有 Span85 和 (或) 卵磷脂和 (或) 三辛酸甘油酯。

当用于人类给药时, 疫苗中的 QS21 及 3D-MPL 含量通常为每剂 1 μ g-200 μ g, 如每剂 10-100 μ g, 优选的是每剂 10 μ g -50 μ g。水包油乳剂一般可含有 2-10%的角鲨烯、2-10%的 α -生育酚和 0.3-3%的 Tween80。角鲨烯: α -生育酚的优选比率为等于或小于 1, 因为这样的乳剂更加稳定。此外还可含有 1%的 Span85。本发明的疫苗可另含有一种稳定剂, 这在某些情况下是有益的。

20 优选地, 无毒水包油乳剂可以在一种水性载体中含有一种无毒的油, 如角鲨烷或角鲨烯, 以及一种乳化剂, 如 Tween80。该水性载体可以是, 例如, 磷酸缓冲盐溶液。

在 WO95/17210 中描述了一种特别有效的佐剂, 该佐剂是在一种水包油乳剂中含有 QS21、3D-MPL 和生育酚。

25 本发明还提供一种多价疫苗组合物, 该疫苗组合物同时含有本发明的疫苗制剂和其他抗原, 特别是可用来治疗癌症、自身免疫疾病及其相关病情的抗原。这种多价疫苗组合物可含有上文所述的一种 TH-1 诱导佐剂。

30 尽管本发明是依据某些 BASB132 多肽和多核苷酸来进行描述, 但应该了解的是, 本发明还包括天然的多肽和多核苷酸的片段, 以及带有添加、缺失或取代的类似多肽和多核苷酸, 这些添加、缺失或取代基本不影响重组多肽或多核苷酸的免疫原性。

组合物、试剂盒与给药方法

另一方面，本发明提供含有 BASB132 多核苷酸和（或）BASB132 多肽的组合物，用于向细胞或多细胞生物给药。

本发明还涉及一些组合物，这些组合物含有文中描述的一种多核苷酸和（或）多肽或其激动剂或拮抗剂。本发明的多肽和多核苷酸可以与
5 未灭菌或灭菌的一种或多种载体形成组合物，以供细胞、组织或生物使用，例如适于向个体给药的药用载体。这些组合物可含有例如介质添加剂或药学有效剂量的本发明所述多肽和（或）多核苷酸，以及一种药学容许的载体或赋形剂。这些载体包括，但不局限于，盐溶液、缓冲盐溶液、葡糖、水、甘油、乙醇及其组合。制剂需适应其给药方式。本发明
10 还涉及含有一个或多个容器的诊断及药用的包和试剂盒，容器中装有一种或多种本发明所述组合物成分。

本发明的多肽、多核苷酸和其他化合物可以单独使用，或与诸如治疗化合物等其他化合物结合使用。

药用组合物可采用任何有效而又便利的方式进行给药，其中尤其包
15 括，例如，通过局部、口腔、肛门、阴道、静脉内、腹膜内、肌内、皮下、鼻内或真皮内途径进行给药。

当用于治疗或预防时，可将活性剂作为可注射的组合物向个体给药，例如无菌的水性分散体，优选的是等渗的无菌的水性分散体。

另一方面，本发明提供一些药用组合物，这些组合物含有药学有效
20 剂量的多肽和（或）多核苷酸，例如可溶形式的本发明所述多肽和（或）多核苷酸，以及显效肽或拮抗肽或小分子化合物，同时含有一种药学容许的载体或赋形剂。这些载体包括，但不局限于，盐溶液、缓冲盐溶液、葡糖、水、甘油、乙醇及其组合。本发明还涉及含有一个或多个容器的药用包和药用试剂盒，容器中装有一种或多种上述组合物成分。本发明的
25 多肽、多核苷酸和其他化合物可以单独使用，或与诸如治疗化合物等其他化合物结合使用。

组合物需适应其给药途径，如通过全身或口腔途径给药。优选的全身给药方式包括注射，一般为静脉内注射。也可采用其他注射方法，如皮下注射、肌内注射或腹膜内注射。全身给药的替代方法还包括，利用
30 诸如胆汁盐或梭链孢酸或其他去污剂等渗透剂进行穿粘膜给药和穿皮给药。此外，若能将本发明的多肽或其他化合物配制成肠溶制剂或胶囊制剂，则也可以进行口服给药。也可将这些化合物以油膏剂、糊剂、凝胶、

溶液、粉末等形式进行局部和（或）局限性给药。

当给药于哺乳动物，特别是给药于人类时，活性剂的日剂量应为0.01mg/kg-10mg/kg，通常约为1mg/kg。医生必须确定最适合某一个体的实际剂量，该剂量随特定个体的年龄、体重和反应而有所不同。以上提供的剂量为典型的平均剂量。当然，在个别情况下应采用更高或更低的剂量，这些情况也处在本发明的范围内。

所需剂量范围取决于选择的肽、给药途径、制剂性质、主体状况的特性和主治医师的判断。但适当的剂量范围为0.1-100 µg/kg 主体。

为方便起见，疫苗组合物可以是可注射的形式。可以用传统的佐剂来加强免疫应答。适于接种的单位剂量为0.5-5µg/kg的抗原，优选的是以该剂量给药1-3次，间隔1-3周。以所述剂量给药时，本发明的化合物不会产生可阻碍将其给药于适当个体的毒副作用。

但考虑到可用化合物的多样性和不同给药途径的效率差异，所需剂量会有很大的不同。举例来说，与静脉内注射给药相比，口服给药需要更大的剂量。利用该领域熟知的以常规经验为依据的优化方法可对这些剂量水平的变化进行调整。

序列库、有形介质中的序列以及算法

多核苷酸与多肽的序列是一种重要的信息资源，可用来确定其二维及三维结构，以及鉴别具有类似同源性的其他序列。实现这些方法的最便利途径是，将序列储存在可被计算机读取的介质中，然后用已知的大分子结构程序中的存储数据进行分析，或利用众所周知的搜索工具，如GCG程序包，对序列库进行搜索。

本发明还提供字符序列或字符串的分析方法，特别是遗传序列或编码的蛋白序列的分析方法。优选的序列分析方法包括，例如，序列同源性分析方法，如同一性或相似性分析方法，DNA、RNA及蛋白结构分析方法、序列装配方法、分支分析方法、序列基元分析方法、可读框判定方法、核酸碱基调用方法、密码子选择分析方法、核酸碱基修剪方法，以及序列层析谱峰值分析方法。

本发明提供一种用计算机进行同源性鉴定的方法。该方法的步骤包括：在可被计算机读取的介质中提供一种含有本发明所述多核苷酸序列的第一多核苷酸序列；并将所述第一多核苷酸序列至少与另一种多核苷

酸或多肽序列进行比较，以确定同源性。

本发明还提供一种用计算机进行同源性鉴定的方法，该方法的步骤包括：在可被计算机读取的介质中提供一种含有本发明所述多肽序列的第一多肽序列；并将所述第一多肽序列至少与另一种多核苷酸或多肽序列进行比较，以确定同源性。

该申请书引用的所有出版物和参考文献均在此全面引入作为参考，其中包括专利和专利申请，但不局限于此，这等同于对每个出版物或参考文献都特别单独标明要在此完整引入作为参考。在该申请书中要求优先权的所有专利申请也在此全面引入作为参考，其方式与上述出版物和参考文献的方式相同。

定义

在该领域中，“同一性”是指通过序列比较而确定的两种或多种多肽序列之间的关系、或两种或多种多核苷酸序列之间的关系，这种关系视具体情况而定。在该领域中，“同一性”还表示通过序列串之间的匹配情况而确定的多肽或多核苷酸序列之间的序列相关程度，这种相关程度同样视具体情况而定。利用已知的方法能够很容易地计算出“同一性”，这些方法包括，但不局限于，《计算分子生物学》，Lesk, A.M., ed., Oxford University Press, New York, 1988；《生物计算：信息学与基因组计划》，Smith, D.W., ed., Academic Press, New York, 1993；《序列数据的计算机分析》，Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994；《分子生物学中的序列分析》，von Heine, G., Academic Press, 1987；以及《序列分析入门》，Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991；和 Carillo, H., and Lipman, D., SIAM J. Applied Math., 48:1073 (1988) 中描述的方法。计算同一性的方法可以给出测试序列间最大的匹配程度。此外，计算同一性的方法还被编写成公用的计算机程序。能够计算两个序列间同一性的计算机程序包括，但不局限于，GCG 程序包中的 GAP 程序 (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984))、BLASTP、BLASTN (Altschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990))，以及 FASTA (Pearson and Lipman Proc. Natl. Acad. Sci. USA 85:2444-2448 (1988))。由 NCBI 和其他资源可获得共用的 BLAST 系列程序 (BLAST 手册, Altschul, S., et al., NCBI

NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990))。还可使用众所周知的 Smith Waterman 算法来计算同一性。

用于多肽序列比较的参数包括:

5 算法: Needleman and Wunsch, J. Mol Biol. 48:443-453 (1970)

对比矩阵: Henikoff and Henikoff, Proc. Natl. Acad. Sci. USA, 89:10915-10919 (1992) 中的 BLOSSUM62

间隔损失: 8

间隔长度损失: 2

10 采用这些参数的有效程序有 Genetics Computer Group, Madison WI 的公用“gap”程序。以上参数是进行肽比较的默认值(对末端间隔不计算损失)。

用于多核苷酸序列比较的参数包括:

算法: Needleman and Wunsch, J. Mol Biol. 48:443-453 (1970)

15 对比矩阵: 匹配=+10, 错配=0

间隔损失: 50

间隔长度损失: 3

可用程序: Genetics Computer Group, Madison WI 的“gap”程序。
这些参数是进行核酸比较的默认值。

20 下文(1)和(2)分别为多核苷酸和多肽的“同一性”的优选含义。

(1)多核苷酸实施方案还包括一种分离的多核苷酸,该多核苷酸含有一段多核苷酸序列,该序列与序列编号:1的参考序列至少有50、60、70、80、85、90、95、97或100%的同一性,其中,所述多核苷酸序列可以与序列编号:1的参考序列完全相同,也可以与参考序列相比含有多至特定整数个核苷酸改变,其中,所述改变可选自,至少有一个核苷酸缺失、至少有一个核苷酸被取代,包括转换和颠换,或至少插入一个核苷酸,并且该改变可出现在参考核苷酸序列的5'或3'端位置或其间的任意位置,可单个地散布在参考序列的核苷酸中,也可分布在参考序列内的一个或多个连续基团中,并且所述核苷酸改变数量的计算方法是,序
25 列编号:1的总核苷酸数乘以确定同一性的整数除以100所得的百分率,
30 然后用序列编号:1的总核苷酸数减去所得结果,或:

$$n_n \leq x_n - (x_n \cdot y),$$

其中 n_n 为核苷酸改变数量, x_n 为序列编号: 1 的总核苷酸数, y 用 0.50 表示 50%, 用 0.60 表示 60%, 用 0.70 表示 70%, 用 0.80 表示 80%, 用 0.85 表示 85%, 用 0.90 表示 90%, 用 0.95 表示 95%, 用 0.97 表示 97%, 用 1.00 表示 100%, 而 $\cdot$ 为乘法运算符, 其中 x_n 和 y 的任何非整数结果都先舍成最接近的整数, 然后用 x_n 减去该整数。通过改变序列编号: 2 所述多肽的多核苷酸编码序列, 可以在该编码序列内产生无义、错义或移码突变, 从而由此改变该多核苷酸编码的多肽。

举例来说, 本发明的多核苷酸序列可以与序列编号: 1 的参考序列完全相同, 即 100% 相同, 也可以与参考序列相比而含有多至特定整数个核苷酸改变, 从而使同一性百分率小于 100%。这些改变可选自, 至少有一个核苷酸缺失、至少有一个核苷酸被取代, 包括转换和颠换, 或至少插入一个核苷酸, 其中该改变可出现在参考多核苷酸序列的 5' 或 3' 端位置或其间的任意位置, 可单个地散布在参考序列的核酸中, 也可分布在参考序列内的一个或多个连续基团中。特定同一性百分率的核苷酸改变数量的计算方法是, 序列编号: 1 的总核苷酸数乘以确定同一性的整数除以 100 所得的百分率, 然后用序列编号: 1 的总核苷酸数减去所得结果, 或:

$$n_n \leq x_n - (x_n \cdot y),$$

其中 n_n 为核苷酸改变数量, x_n 是序列编号: 1 的总核苷酸数, y 是, 例如, 用 0.70 表示 70%, 用 0.80 表示 80%, 用 0.85 表示 85% 等等, $\cdot$ 为乘法运算符, 其中 x_n 和 y 的任何非整数结果都先舍成最接近的整数, 然后用 x_n 减去该整数。

(2) 多肽实施方案还包括一种分离的多肽, 该多肽含有一段与序列编号: 2 的参考多肽序列至少有 50、60、70、80、85、90、95、97 或 100% 同一性的多肽, 其中所述多肽序列可以与序列编号: 2 的参考序列完全相同, 也可以与参考序列相比而含有多至特定整数个氨基酸改变, 其中, 所述改变可选自, 至少有一个氨基酸缺失、至少有一个氨基酸被取代, 包括保守和非保守取代, 或至少插入一个氨基酸, 并且该改变可出现在

参考多肽序列的氨基端或羧基端位置或其间的任意位置，可单个地散布在参考序列的氨基酸中，也可分布在参考序列内的一个或多个连续基团中，并且所述氨基酸改变数量的计算方法是，序列编号：2 的总氨基酸数乘以确定同一性的整数除以 100 所得的百分率，然后用序列编号：2 的总氨基酸数减去所得结果，或：

$$n_a \leq x_a - (x_a \cdot y),$$

其中 n_a 为氨基酸改变数量， x_a 为序列编号：2 的总氨基酸数， y 用 0.50 表示 50%，用 0.60 表示 60%，用 0.70 表示 70%，用 0.80 表示 80%，用 0.85 表示 85%，用 0.90 表示 90%，用 0.95 表示 95%，用 0.97 表示 97%，用 1.00 表示 100%，而 $\cdot$ 为乘法运算符，其中 x_a 和 y 的任何非整数结果都先舍成最接近的整数，然后用 x_a 减去该整数。

举例来说，本发明的多肽序列可以与序列编号：2 的参考序列完全相同，即 100% 相同，也可以与参考序列相比而含有多至特定整数个氨基酸改变，从而使同一性百分率小于 100%。这些改变可选自，至少有一个氨基酸缺失、至少有一个氨基酸被取代，包括保守和非保守取代，或至少插入一个氨基酸，并且该改变可出现在参考多肽序列的氨基端或羧基端位置或其间的任意位置，可单个地散布在参考序列的氨基酸中，也可分布在参考序列内的一个或多个连续基团中。特定同一性百分率的氨基酸改变数量的计算方法是，序列编号：2 的总氨基酸数乘以确定同一性的整数除以 100 所得的百分率，然后用序列编号：2 的总氨基酸数减去所得结果，或：

$$n_a \leq x_a - (x_a \cdot y),$$

其中 n_a 为氨基酸改变数量， x_a 是序列编号：2 的总氨基酸数， y 是，例如，用 0.70 表示 70%，用 0.80 表示 80%，用 0.85 表示 85% 等等， $\cdot$ 为乘法运算符，其中 x_a 和 y 的任何非整数结果都先舍成最接近的整数，然后用 x_a 减去该整数。

文中使用的关于生物的“个体”是指一种多细胞真核生物，其中包括，但不局限于，后生动物、哺乳动物、ovid、牛、猿、灵长类动物和

人。

“分离的”是指由天然状态被“人工”改变的，即如果天然存在，则已发生改变或已由最初环境转移，或两种情况都发生。举例来说，在活体生物中天然存在的多核苷酸或多肽不属于本文使用的术语“分离的”，但由其天然共存物分离出的相同多核苷酸或多肽则属于“分离的”。此外，通过转化、遗传操作或其他任何重组方法被引入生物体的多核苷酸或多肽即使在该生物体中仍然存在，也是属于“分离的”，其中该生物体可以有生命或无生命。

“多核苷酸”通常指任何多聚核糖核苷酸或多聚脱氧核糖核苷酸，它可以是含有单链和双链区的未修饰或修饰的 RNA 或 DNA。

“变异体”是指与参考多核苷酸或多肽不同，但仍保持基本特性的多核苷酸或多肽。典型的多核苷酸变异体与其参考多核苷酸的核苷酸序列不同。变异体中的核苷酸序列变化可以改变或不改变由参考多核苷酸编码的多肽的氨基酸序列。如下文所述，核苷酸变化的结果可以是参考序列编码的多肽中的氨基酸取代、添加、缺失、融合和截断。典型的多肽变异体与其参考多肽的氨基酸序列不同。其差异通常有限，因而参考多肽与其变异体的序列整体上非常相似，并且有许多区域完全相同。变异体与参考多肽在氨基酸序列上的差异可以是一个或多个取代、添加、缺失的任意组合。取代或插入的氨基酸残基可以被遗传密码编码，也可以不能被遗传密码编码。多核苷酸或多肽的变异体可以是天然的，如等位变异体，也可以是不知是否天然的变异体。多核苷酸和多肽的非天然变异体可通过诱变技术或直接合成的方法来获得。

“疾病”是指由细菌感染引起或与其相关的任何疾病，其中包括，例如，幼儿及儿童中耳炎、老年肺炎、鼻窦炎、院内感染以及侵袭性疾病、慢性中耳炎伴听力丧失、中耳内液体累积、听神经损伤、语言学习延迟、上呼吸道感染和中耳炎症。

实施例

除详细描述的地方外，以下实施例均采用该领域技术人员熟知的和常规的标准技术来实现。这些实施例均作为说明，而不是限制本发明。

实施例 1: 来源于粘膜炎莫拉氏菌 ATCC 43617 菌株的 BASB132

基因的 DNA 序列

A: 粘膜炎症莫拉氏菌中的 BASB132

序列编号: 1 显示来源于粘膜炎症莫拉氏菌 ATCC 43617 菌株 (也称为 MC2931 菌株) 的 BASB132 基因的 DNA 序列。序列编号: 2 显示 BASB132 多核苷酸序列的翻译结果。

B: 粘膜炎症莫拉氏菌 43617 菌株中的 BASB132

确定粘膜炎症莫拉氏菌 ATCC 43617 菌株中的 BASB132 基因的序列。为此, 用 Big Dyes 试剂盒 (Applied biosystems) 对含有截断形式的粘膜炎症莫拉氏菌 ATCC 43617 菌株 BASB132 基因编码区 (相当于 3' 端的最后 2977 个核苷酸) 的质粒 DNA (见实施例 2A) 进行 DNA 测序, 并按照供应商提供的条件在 ABI 373/A DNA 测序仪上进行分析, 所用引物是对粘膜炎症莫拉氏菌 BASB132 基因有特异性的 oli5 YtfN (5'-CCC ACC GTT TGC CCT TCA AT-3') [序列编号: 5]、oli6 YtfN (5'-CAT TTT TCC CGA GCA TTC AAA C-3') [序列编号: 6] 和 oli7 YtfN (5'-AAT CAG CCA CTT ATC GCC AC-3') [序列编号: 7], 以及对载体有特异性的 M13 通用测序引物 (5'-GTA AAA CGA CGG CCA GT-3') [序列编号: 8] 和 M13 反向测序引物 (5'-CAG GAA ACA GCT ATG AC-3') [序列编号: 9]。结果是获得了分别称为序列编号: 3 和序列编号: 4 的多核苷酸序列及其推导多肽序列。

利用 DNASTAR 软件包的 MegAlign 程序对序列编号: 1 和 3 的多核苷酸序列进行比较, 结果显示于图 1; 成对比较的同一性结果显示, 两种 BASB132 多核苷酸基因序列在重叠区 100% 相同。利用该 MegAlign 程序对序列编号: 2 和 4 的多肽序列进行比较, 结果显示于图 2; 成对比较的同一性结果显示, 两种 BASB132 多肽序列在重叠区 100% 相同。

实施例 2: 构建质粒以表达截断形式的重组 BASB132

A: 截断 BASB132 的克隆

正向扩增引物 oli1 YtfN (5'-TCA TGA ATG ACT CAG GCA AAG-3') [序列编号: 10] 和反向扩增引物 oli2 YtfN (5'-AGA TCT AAA CTT CCA ACG ATA AAT C-3') [序列编号: 11] 中分别设计了 *Bsp*HI 和 *Bgl*II 限制位点, 使 PCR 产物能够定向克隆到大肠杆菌表达质粒 pQE60 中, 从

而可将截断形式的 BASB132 蛋白作为 C-端含有(His)₆ 亲和层析标记的融合蛋白来加以表达。由于基因较大, 所以用两组并行的 PCR 实验来扩增整个编码序列, 第一组扩增使用 oli1 YtfN (5'-TCA TGA ATG ACT CAG GCA AAG-3') [序列编号: 10] 和 oli3 YtfN (5'-AAA CAA ATC GCA CCC ACG CC-3') [序列编号: 12], 第二组扩增使用 oli2 YtfN (5'-AGA TCT AAA CTT CCA ACG ATA AAT C-3') [序列编号: 11] 和 oli4 YtfN (5'-ACA AAT TGC AGC GCA TTG TTG G-3') [序列编号: 13]。扩增的两组片段都具有一段含 *Bst*XI 单一限制位点的公用区。首先用 Top10 细菌按厂商说明将 BASB132 的 PCR 产物引入到 pCRIITOPO 克隆载体中 (Invitrogen)。这种中间构建体有利于进一步克隆到表达载体中。利用限制酶分析方法筛选出含有 BASB132 DNA 插入序列的转化体。约取 20μl 消化后的反应物试样进行琼脂糖凝胶电泳分析 (Tris-醋酸盐-EDTA (TAE) 缓冲液中含有 0.8% 的琼脂糖)。凝胶电泳结束后用溴化乙锭染色, 然后用 UV 光显示 DNA 片段。将 DNA 分子量标准 (1 Kb 梯形分子量标准, Life Technologies) 与测试样品并行电泳, 以估计 DNA 片段的大小。然后根据厂商建议, 用 *Bsp*HI 与 *Bst*XI 或 *Bst*XI 与 *Bg*/II 限制酶 (Life Technologies) 依次完全消化从每次克隆的选定转化体纯化出的质粒。消化的 DNA 片段用硅胶旋转柱纯化, 然后与 pQE60 质粒连接。

20 **B: PCR-阳性转化体的表达分析**

为了制备用于连接的表达质粒 pQE60, 用 *Nco*I 和 *Bg*/II 将其进行类似的完全消化, 然后按厂商说明用牛小肠磷酸酶 (CIP, ~0.02 单位/pmol 5'末端, Life Technologies) 进行处理, 以避免自身连接。连接反应使用了约 5 倍摩尔数过量的消化片段来制备载体。用该领域熟知的方法以 T4 DNA 连接酶 (~2.0 单位/反应, Life Technologies) 进行~20μl 的标准连接反应 (~16℃, ~16 小时)。按该领域熟知的方法用连接反应的试样 (~5μl) 对电感受态细胞进行转化。再将转化细胞置于~1.0 ml LB 培养基中 37℃ 生长~2-3 小时, 然后铺在含有氨苄青霉素 (100 μg/ml) 的 LB 琼脂平板上。添加的抗生素可用于筛选。将平板 37℃ 过夜保温~16 小时。用无菌牙签挑取 ApR 单菌落, 并对新鲜 LB ApR 平板进行“点状”接种, 同时对~1.0 ml LB ApR 液体培养基接种。在普通培养箱 (用于平板) 或振荡水浴中将接种的平板和液体培养基 37℃ 保温过夜。然后进行限制酶分

析，以证实转化体中含有 BASB132 DNA 插入序列。约取 20 μ l 消化后的反应物试样进行琼脂糖凝胶电泳分析（Tris-醋酸盐-EDTA（TAE）缓冲液中含有 0.8% 的琼脂糖）。凝胶电泳结束后用溴化乙锭染色，然后用 UV 光显示 DNA 片段。将 DNA 分子量标准（1 Kb 梯形分子量标准，Life Technologies）与测试样品并行电泳，以估计 DNA 片段的大小。产生预期大小 DNA 片段的转化体即为含有 BASB132 表达构建体的菌株。然后对含有表达质粒的菌株进行重组 BASB132 的可诱导表达分析。

C: PCR-阳性转化体的表达分析

10 然后用该领域熟知的方法将重组质粒的 DNA 试样（~1 μ l）转化到感受态 M15（pREP4）细菌中。将转化细胞置于~1.0 ml LB 培养基中 37 $^{\circ}$ C 生长~2-3 小时，然后铺在含有氨苄青霉素（100 μ g/ml）和卡那霉素（30 μ g/ml）的 LB 琼脂平板上。添加的抗生素可用于筛选。将平板 37 $^{\circ}$ C 过夜保温~16 小时。用无菌牙签挑取 ApR 单菌落，并接种到~5.0 ml 的 LB ApR
15 KmR 液体培养基中。将液体培养基 37 $^{\circ}$ C 振荡（~250 rpm）保温过夜。再把过夜的接种培养物的试样（~1.0 ml）接种到含有~25 ml LB Ap 培养基的 125 ml 锥形瓶中，然后 37 $^{\circ}$ C 振荡（~250 rpm）培养，直到培养物的混浊度达到 O.D.600 为~0.5，即对数中期（一般约为 1.5-2.0 小时）。此时将大约一半的培养物（~12.5 ml）转到另一个 125 ml 锥形瓶中，并通过
20 添加终浓度为 1.0 mM 的 IPTG（在无菌水中配制的 1.0 M 储液，Sigma）来诱导重组 BASB132 蛋白的表达。将 IPTG 诱导的和未诱导的培养物都于 37 $^{\circ}$ C 及振荡条件下再保温~4 小时。诱导期结束后，从诱导的和未诱导的培养物中取样（~1.0 ml），并在微量离心机中室温离心~3 分钟，以收集细胞。将细胞沉淀分别悬浮在~50 μ l 无菌水中，然后与等体积的含有
25 2-巯基乙醇的 2 $\times$ Laemmli SDS-PAGE 样品缓冲液相混合，并在沸水浴中煮~3 分钟，以使蛋白变性。将 IPTG 诱导的和未诱导的粗制细胞裂解物以相同的体积（~15 μ l）加样到 12% Tris/甘氨酸聚丙烯酰胺凝胶上（1 mm 厚的 Mini-gels, Novex），一式两份。在常规条件下用标准的 SDS/Tris/甘氨酸电泳缓冲液（BioRad）将诱导的和未诱导的裂解物样品及预染的
30 分子量标准（SeeBlue, Novex）一同进行电泳。电泳结束后，一块胶用考马斯亮蓝 R250（BioRad）染色，然后脱色，以显示可由 IPTG 诱导而新产生的 BASB132 蛋白（图 3A）。另一块胶则用 BioRad Mini-Protein II

印迹装置和 Towbin's 甲醇 (20%) 转移缓冲液在 PVDF 膜 (0.45 微米孔径, Novex) 上 4℃ 电印迹~2 小时。膜的阻断和抗体保温是按该领域熟知的方法来进行。利用抗 RGS(His)₃ 单克隆抗体和 HRP 偶联的兔抗鼠二抗 (QiaGen) 来验证 BASB132 重组蛋白及其表达 (图 3B)。用 ABT 不溶底物或 Hyperfilm 和 Amersham ECL 化学发光系统来显示抗-His 抗体的反应模式。

实施例 3: 重组 BASB132 的产生

菌株

10 用含有粘膜炎莫拉氏菌 BASB132 编码质粒 (pQE60) 的大肠杆菌 M15 重组表达菌株 (PREP4) 来产生用于纯化重组蛋白的细胞团。在含有 100 µg/ml 氨苄青霉素 (“Ap”) 和 30 µg/ml 卡那霉素 (“Km”) 的 LB 琼脂平板上培养表达菌株, 以确保 pQE60 和 pREP4 得以维持。用于 -80℃ 冷冻保存的菌株是在含有相同浓度抗生素的 LB 培养基中增殖, 然后与等体积的含有 30% (w/v) 甘油的 LB 培养基相混合。

培养基

用于产生重组蛋白的发酵培养基含有 2 × YT 培养基 (Difco) 和 100 µg/ml Ap 及 30 µg/ml Km。在发酵罐中使用的培养基添加了 0.25 ml/L 消泡剂 (Antifoam 204, Sigma)。在发酵罐中添加 IPTG (异丙基-β-D-硫代吡喃半乳糖苷) (终浓度 1 mM), 以诱导 BASB132 重组蛋白的表达。

发酵

25 用 0.3 ml 快速融解的冷冻培养物, 或选择性琼脂平板培养基上的数个菌落, 对含有 50 ml 工作容积的 500-ml 接种锥形瓶进行接种, 并在 150 rpm 的振荡平台上 (Innova2100, New Brunswick Scientific) 于 37±1℃ 保温约 12 小时。然后用该接种培养物对含有 2 × YT 培养基和 Ap 抗生素的工作容积为 5-L 的发酵罐进行接种。发酵罐 (Bioflo 3000, New Brunswick Scientific) 在 37±1℃、曝气速率 0.2-0.4 VVM 以及 Rushton 叶轮转速 250 rpm 的条件下运转。接种锥形瓶与发酵罐中的 pH 值均不做调整。发酵其间, 发酵罐中的 pH 值为 6.5-7.3。当培养物达到对数中期时 (~0.7 O.D.600 单位), 把 IPTG (溶于无菌水的 1.0 M 储液) 添加到发酵罐中。细胞诱

导 2-4 小时,然后用 28RS Heraeus(Sepatech)或 RC5C 超速离心机(Sorvall Instruments) 离心以收集细胞。细胞浆在处理前保存于-20℃。

5 化学剂与材料

咪唑和 Triton X-100 购于 Merck 公司。盐酸胍购于 Fluka 公司。抑肽酶来自 Sigma 化学制品公司。尿素和 AEBSF 来自 ICN-Biochemicals 公司。其他所有化学剂均为试剂纯或更高程度。

10 Ni-NTA Superflow 树脂和不含 BSA 的 Penta-His 抗体购于 QiaGen 公司。MicroBCA 测定剂来自 Pierce 公司; Amicon 3 滤膜来自 Millipore 公司。透析膜 (MWCO12-14000) 购于 MFPI, USA。分子量标准 (梯形标准) 购于 Life-technologies。

实施例 4: 由大肠杆菌纯化重组 BASB132

15 提取-纯化

把 2750 ml IPTG 诱导的培养物 (~4 小时, OD₆₂₀=0.5) 获得的细胞浆重悬在 110 ml 含有 6 M 盐酸胍的 100 mM NaH₂PO₄, 10 mM Tris-HCl, pH 8 缓冲液中 (缓冲液 A), 室温放置 1 小时, 然后 10,000g 离心 20 分钟。将上清液加到用缓冲液 A 平衡的 Ni-NTA Superflow 树脂上, 立即收
20 集。用含有 8 M 尿素的 100 mM NaH₂PO₄, 10 mM Tris-HCl, pH 6.3 缓冲液 (缓冲液 C) 将树脂清洗两次。洗脱是用 4 × 1.5 ml 含有 8 M 尿素的 100 mM NaH₂PO₄, 10 mM Tris-HCl, pH 5.9 缓冲液 (缓冲液 D), 然后用 4 × 1.5 ml 含有 8 M 尿素的 100 mM NaH₂PO₄, 10 mM Tris-HCl, pH 4.5 缓冲液 (缓冲液 E)。洗脱组分用 25% 体积 pH7.5 的 200 mM 磷酸缓冲液
25 中和。将含有 BASB132 蛋白的组分合并, 并对含有 8 M 尿素的 100 mM NaH₂PO₄, pH7.4 缓冲液透析, 然后对含有 4 M 尿素的该缓冲液透析, 再对含有 2 M 尿素的该缓冲液透析, 最后对含有 0.1% Triton-X100 的 pH7.4 的 PBS 缓冲液透析三次。

由于在透析过程中产生沉淀, 所以溶液呈混浊状 (溶液 A)。离心
30 (3000 rpm, 10 分钟) 两次后, 溶液变得澄清 (溶液 B)。利用 MicroBCA 测定剂对溶液 A 和溶液 B 的纯化 BASB132 蛋白进行定量。溶液 A 的结果为 170 µg/ml。获得的纯化澄清蛋白 (溶液 B) 为 1.65 mg, 终浓度为

75 µg/ml。如图 4-A 所示，纯化的 BASB132 蛋白在 SDS-PAGE 分析中显示为一条迁移率约 100 kDa（估计的相对分子量）的主带。纯度估计在 70%以上。BASB132 蛋白与一种针对 6-组氨酸基元诱发的小鼠单克隆抗体具有反应活性（图 4-B）。

5

实施例 5：产生抗重组 BASB132 的抗血清

用纯化的重组 BASB132 蛋白对兔子接种可产生抗 BASB132 蛋白的多价抗血清。

10 用纯化的重组 BASB132 蛋白对小鼠接种也可产生抗 BASB132 蛋白的多价抗血清。

在首次免疫前（免疫前血）和最后一次免疫后取动物血。

通过 ELISA 方法用纯化的重组 BASB132 蛋白测定抗 BASB132 蛋白的效价。该效价的定义是用 XL Fit 软件由 4-参数逻辑模型计算出的中点效价。

15 另外还按下文实施例 7 所述，用蛋白印迹方法以抗血清为一抗对蛋白进行鉴定。蛋白印迹的结果显示，在免疫动物的血清内存在抗 BASB132 抗体。

实施例 6：免疫学鉴定：BASB132 的表面暴露情况

20 利用被福尔马林杀死的粘膜炎莫拉氏菌全细胞以 ELISA 方法测定抗 BASB132 蛋白的效价。该效价的定义是用 XL Fit 软件由 4-参数逻辑模型计算出的中点效价。

实施例 7：免疫学鉴定：蛋白质印迹分析

25 在 Muller Hinton 琼脂平板上将包括 ATCC43617 菌株在内的几种粘膜炎莫拉氏菌菌株以及来自不同地理区的临床分离物于 36℃培育 24 小时。用几个菌落来接种培养基。培养物生长到 A620 约为 0.6 时，离心收集细胞。然后将细胞浓缩并溶解在 PAGE 样品缓冲液中。然后在 4-20% 聚丙烯酰胺凝胶上对溶解的细胞进行解析，并将分离的蛋白电泳转移到
30 PVDF 膜上。然后用饱和缓冲液对 PVDF 膜进行预处理。随后的所有保温步骤都使用这种预处理缓冲液。

用免疫前血清或兔或鼠的免疫血清与 PVDF 膜保温。然后清洗 PVDF

膜。

用生物素标记的绵羊抗兔或抗鼠 Ig 与 PVDF 膜保温，再与抗生蛋白链菌素-过氧化物酶保温。然后用清洗缓冲液将 PVDF 膜清洗 3 次，再用 4-氯-1-萘酚显色。

5

实施例 8: 免疫学鉴定: 杀菌活性

检测抗-BASB132 抗体的补体介导的细胞毒性，以确定按上文所述制备的 BASB132 蛋白抗血清的疫苗效力。检测免疫前血清与抗-BASB132 抗血清的补体介导的杀粘膜炎莫拉氏菌活性。在平板上培育粘膜炎莫拉氏菌菌株。将几个菌落接种到液体培养基中。使培养物增殖，并在 A620 约为 0.4 时收集细胞。清洗一次，然后将沉淀悬浮并稀释。

96 孔板的第一个孔中附着免疫前血清和抗-BASB132 血清，同一行的其他孔中附着一系列稀释度的血清。然后添加稀释的活体粘膜炎莫拉氏菌，并将混合物保温。按照在毒性测定试验中预先确定的有效稀释度将补体添加到每个孔中。

每次测试都包括补体对照（不含血清但含有活性或灭活补体源的孔）、阳性对照（含有血清和已知效价的杀菌性抗体的孔）、培养物对照（不含血清和补体的孔）和血清对照（不含补体的孔）。

测量兔或鼠抗血清的杀菌活性（杀死 50% 的同源菌株）。

20

实施例 9: 抗 BASB132 抗体在人类恢复期血清中的存在情况

除一抗采用来源于被粘膜炎莫拉氏菌感染的儿童的人类血清之外，按上文实施例 7 所述对纯化的重组 BASB132 进行蛋白印迹分析。

实施例 10: BASB132 疫苗的效力: 提高小鼠肺部的粘膜炎莫拉氏菌清除率

该小鼠模型的依据是对接种小鼠进行标准的鼻内诱发，然后分析粘膜炎莫拉氏菌对肺的侵袭。

用 BASB132 疫苗对成组小鼠进行免疫。在促升接种后，将细菌悬浮液滴注到麻醉状态的小鼠的鼻孔内，以此来激发小鼠。在激发后 30 分钟到 24 小时之间处死小鼠，在无菌条件下取出肺脏并分别匀浆。把匀浆的稀释物铺在琼脂平板上，然后计算生长的菌落，并由此确定 CFU/肺的

log10 加权平均数。计算各组 CFU/肺的 log10 加权平均数的算术平均值和标准偏差。用统计学方法对结果进行分析。

该实验中的各组小鼠是用 BASB132 或死亡的粘膜炎莫拉氏菌全细胞制剂 (kwc) 进行免疫或进行假免疫。

5

实施例 11: 抗 BASB132 抗血清对粘膜炎莫拉氏菌粘附于细胞的抑制

该试验可测量抗 BASB132 血清对莫拉氏菌粘附于上皮细胞的抑制能力。该活性可阻碍莫拉氏菌在鼻咽部建群。

- 10 将 1 体积细菌与 1 体积免疫前血清或抗 BASB132 免疫血清的稀释物置于冰上保温。然后把该混合物添加到含有为去除痕量抗生素而用培养基清洗一次的铺满细胞培养物的 24 孔板中。将板离心并保温。然后温和清洗每个孔。清洗完后, 在孔中加入甘胆酸钠。将板保温, 然后刮出细胞层并匀浆。把匀浆的稀释物铺在琼脂平板上并保温。对每个平板的菌落计数, 并计算出每个孔中存在的细菌数。
- 15

保藏材料

- 含有粘膜炎莫拉氏菌 Catlin 菌株的保藏物已于 1997 年 6 月 21 日保藏在美国典型培养物保藏中心 (文中称为 “ATCC”), 保藏号为 43617。
- 20 被描述成粘膜炎布兰汉氏球菌 (Frosch and Kolle) 的该保藏物是一种由粘膜炎莫拉氏菌分离物构建的冻干形式的 1.5-2.9 kb 插入序列库, 其中的粘膜炎莫拉氏菌分离物是从患有慢性支气管炎的一名煤矿工人的穿气管吸出物中获得。该保藏物在 Antimicrob. Agents Chemother. 21:506-598 (1982) 中有所描述。

- 25 该粘膜炎莫拉氏菌菌株保藏物在文中称为 “保藏菌株” 或 “保藏菌株的 DNA”。

该保藏菌株含有全长 BASB132 基因。

- 含有插入到 pQE30 中的粘膜炎莫拉氏菌 DNA 的 pMC-D15 载体保藏物已于 1999 年 2 月 12 日保藏在美国典型培养物保藏中心 (ATCC), 保藏号为 207105。
- 30

该保藏菌株/克隆中包含的多核苷酸序列及其编码的所有多肽的氨基酸序列如果与文中关于序列的任何描述有任何冲突, 则以前者为准。

该保藏菌株的保藏符合“国际承认用于专利程序的微生物保藏的布达佩斯条约”的规定。本申请获得专利权时，该保藏菌株即可以不可撤销的方式不受限制地或无条件地向公众发放。提供该保藏菌株只是为该领域的技术人员提供方便，而不是认为保藏是申请专利的必需条件，例如 35 U.S.C. §112 即要求该条件。

序列信息

BASB132 的多核苷酸及多肽序列

5 序列编号：1

来源于 ATCC43617 菌株的粘膜炎莫拉氏菌 BASB132 多核苷酸序列

```
ATGACTCATCAGGATAATTCAAAAACTCGTCATCGGCAGAAAATGGGGTTTCTGCTGGCGTTGCTGCACCGACTAAAA
CTCAGCCCAAACCAACCAACCCACCAACCAACCAATGGCTAATCTATTTAATTCGCTTAATTGTCGGAGCAATTATTGC
CGCTTTATTGTTGTTGGCGGTGTTATTTGCCATGACCAACAGCGAGTCAGGTTCAAAGTTTTTAATCGAAAAAATTGCG
TTAGAAACTGGCACTAAGCTTAAGTACAGCGAAGGCTCAATTCGCCATGGAGTTTGGGTGCAAGATGTCAAATCGCTC
AAAGTGAAGATATTACCATCACCATTAAATCGTGCCATGTACAGCTTGGGTGGCGAGCCTTGTTGCTCGCCAAGTGCA
TTTGGTCAATCCTAAGATTGATAAAGTTTATGTGACAAACACCAAGCCATCAACAGGCGAACCCCTTGATTATGCGACC
ATCAACCTACCAAGTGACGCTTAAGCTTGAAAATGCCAAAGTCAATGAAATTATCTATGACCAAGTGGATTCTGAGCCTG
TCGTACTGCATCATATCGCATTTGATCAGCATCATGGGCAGATTCAACAGTTAAAATTGATAACGCCATGCTAAGCTA
TGGTGATGATATTAATATCAGCCATGCCACTGGTGGAAATTGATTTAACAGGTCATTATCCGCTGTCGTTGTCGGCAGAT
GTGCATATTTTGGCACTTGATGATGCGTATTTTGATACCTTTGTCGGTGAAAGCAGGCGGTAGCCTTAAGCGTACCGTTG
GTACACTCACTGGCAAATATAATCAGCATCATGTGACAGGCGAGCTTTATCGCTCAAGGGTTGGATAAAAACTCACCTTT
TAGCGCAGCCTTGATTTTGATGAAGTACGATTGCCTTATGCTGACAGTCAAATATTTTACTAAAAAATGGCTCTATC
ATCGCTGATGGCGTCATCTCAAATATCGAGCTACGCATTAACACTGAGTTATCCGCCAAAGATATTCCTGATGGGCATT
ATCACGGTCGTGGAATTATTCGTGGCAGTACCATGCAAATCCCATATTTGCAGGCTGATACTGCCAATGGTACTTTGGT
GGCAACGGGTGATATGACTTGGGAAGATGGCTATGAGCTTGATGCCACCATTACAGCAGACGGCTATCGTATCCGTGAA
GAGATGCCAAGTGATTATCATGAATATAGAGCCTATCTACCTAAGGTTTTGACAGGCTCACTTGGGGTTAAGTATACGC
TATTAGACAAAGCCAGTCATGATACTCGGTTTGAGTTTGATCTGAATCAAAAAGACGGTGAACGCATTCAAGCGACGCT
GGCTCAAATCAACAGAGTGATCATGAGCCTTGGCGTATTGATGCGACTTGGGCAAATCTAATCCGCCATGATATCCCA
CAAATTGGCGAGATTATAGCCGCTCAGGTCAGGCATCAGTTCGTTTGGAGAACGGACATACCTATATTAACGCCTCTG
CTGACATTGTTAAACTTAATGCCGTCCCAAGCGGATCATATCATGTGCAAGCAAATATTGAGCAAAATCAGCACTTGCA
CTTGACAGATTTTAACTATCAAGGTGTGATGGCGGAGTTGACAGGTACTGGCAGGGTTGATTTTGGCACTGCCCAAAGA
CCGCTTGAGTGGCAGCTTGATGTCACTGCCAATCCAGTCAAACCCCAATGCGTATTTTCAAACACCCCAATCAAACACCGT
TTGAGCAAATTTAGGCAGAATAATTGCTTCAGGGCGTCTGCGTGAGATAGATGGTGTGTCAATTCATGACATTAAAGT
GGATGACAGTGATTTAACGCTTTATTAATGACGGCAAGCAAGTACATCTGATGGGTCAAGGCGTCAGCAAGATTAGG
CTACAAGATGGACAAATTAGCCATCTTAAGGCAAATTTTGATGGCAAGCTTGCTCAAGATATCTTACCGCAGGTAGCTG
ACTCAAGCATTGGATTGGATATTGAAGGTGATTTGAATGATTTGACCATCACACGGGCAGTCATGATGAATGACTCAGG
CAAAGTATCGCTTGCAAGTCTGCTTGATTAAATGACGGTATTATTGGCAAATTCAGGGCGATTAGATGAGGTGGAT
ACCGCCGCTTTTGTGATGATGACAATTTAGTTGCGATCATCACAGGCGATCTTGCCACATCAGGCAGATATCGTGATG
GCGTGATGGATGATATCGCACTCACTTTTGATGGTCAAGTGATTAACAATTCAATCCCAAATGGTCATCTAAGTATTGA
TGCGACAGGTTCTGGCAATCGCTTTATGGTCAATCATCTAAGACATGATGGCACAGCAGGCACGCTGAATGCGACAGGC
TGGGTAGATATTAGCCAAGGTGCTGCATGGCAGCTAGAGGCTGATATGAATTCATTGAATTTGGGTGCATTTATCCAGT
CGCTAGATACTGACTTAAACGGTACAATTAGCTTGCAAGCAACTGGCAAGAGTCTCATCAAGTCATTGATATCAGTAA
TTTGGATATTACAGGTAGCTACAATAATCAGCCACTTATCGCCACAGGCAGCTTATATGCCAAGCTTTCCCTACCTAAA
GATTTGGCTGGGTATATCCAAAGTATCAAGCAAGCAAGCCGCCACCGACTTCGTCTGATGATTTATTGGCACTAAAAA
ATCGTATTGATAACAACGCACGCCAAACCCAAAATATCGTTCATAAACTCAATGCTGATAATCTACAAGTCCGCATTGG
CAATAATCATTTGGCCATATCAGGCGATGAAAGACAGTTGACCGCCAGTGTGAATGTGATTGATCTTGGGCAGCTGATT
GATACCGCAAGTGGTGGGATTCAAGGCGGTGTGATTGTCGTCAATGATCATCACGCTTGCCAACTTTATATATTGATG
CCAGTGTCTCATCACTCAGATTTGCCAATATCACCATCCAAAACGCCCAAGCTATCGGTAAATTTGTCAATCTGGGTAA
TAGCGAAAGCCAGCTGTTGGTTTCAAGGTGATGATATTATCGTGATGGGACGCACCATTAATCTGCTCGGATGGATTTT
AGTGGCACAGAAGCCGATCATATCTTATCAATTTCAACCAAAAGCGGTGATATCGAAGCATCTATGCATATTGATGGGG
CATTTAATCGCAACAATATGCGATATCATGGTGTGTTGCTGATGGTGTGTCAGAGTGGATTGGCAAAATGTCACA
GCGTCAGCCGACCGAGTTTAGCTATGGATTGGATGATAATAGCTTACAAATTGCAGCGCATTGTTGGCAATCGGCACAT
```

ATCCAAGGCGATGGCGTGGGTGCGATTTGTTTACAAGATACATTAAAGTTATACACCGCAATCAGGGAATGTCAATCTTA
 TTATCCAAAATCTTGATACTCAAGTACTATTGGCGGCTCTACCCAGTGATATTCATTGGAAATCCATGCTCAATGGTCG
 CATTAAAGCAATGTGGCAGGCAGGTCAATCACCTTTGGTTGATGCTGTACTGTATTCTGATGATGGTACAATTGGCTTA
 ACTCAAGAAGAGACAGGTTATATTGAGATGCCATATCAGCGTGCGTCTGTGATTGCCAAAAGTGTGATAATGGCTTAA
 AGTAAGAACCAGATATCTTGGGTACGGCAGGTCTGGTTATGCCGATGTAATCATCAACCCCAAGCAACAGACAAGCC
 CATTTCTGGTGCGCTTGTATGAATGATCTTAATTTGGCGGTATTGCGACCATTTTCCCGAGCATTCAAACATTGAGC
 GGTAAAGTCAGCTTGGCAGGTGGCTTAGGTGGTACATTATCCAAACCATTGTTTTATGGTAATGCCAATTTAAAGAATG
 GTGCATTAGCGATTGTTGGCGTACCAATGCCATTCTGATGTTGATGCCACGGTGAACATACGAGGTACTAAAGCACG
 CTTAGACGGTAGATTTATGGGTGGTGAGGGTCATGGTGTGCTTTATGGTGAAATGGATTGGGCAGAAGAGCTGTATGCT
 CGCCTTGGCGTATTTGGTGAAAATCTGACCGTCAGCCAGCCACCCTTGTGACGGCACAAATCAGCCCTGAGCTGGAAG
 TGATTATCAGACCTTTCCAAAATTTGTAGATGTTCAAGGCGTTGTTAGTATCCCATCAGCAACCATTCTGCCGCCAGA
 GGCAACCGCAGATATCGTAACCGAATCCCAGATGTATCAGTCATTGACCGCGTATCAGAGTAATATTGACCAGATT
 TTAAGTAGAGCGAAACCTTGGGATATTAATGCAAATATCGGTGTTGATTGGGTAATGAGATTGAATTCGTGGATTG
 GGGCGGTATTACCATTTGGCAGGCGCGATACATTTGACTCAGTCAGGACAAGGTGCAATGCAAGCTCTTGGCGTGGTTCA
 GGTTCCTAAACGCACAAAATCGATGTCATCGGTCAAAATTTGGATTTAAATTATGCCCAAATTCGTTTGTATGGCGAC
 ATGCTTAATCCACGCTTATCTATTGAGGGTGAAAACAAATTGAAGGGCAAACGGTGGGCGTTCGATTCTGTGGAACGG
 CATCAATCCAAACATTACCGTATTTAATGATGCAGGTTTGGATGAATACCAAGCAATGAATGCACTGGTGACAGGACG
 CATTAGCGAATCAAGTGATTTGGGTATCAGAGCAGGGTTTTCGCTCACAAGTAACCAATCACCTTGCTGCTGCTGGC
 TTAAGCTTAGGTTTGTGAGGAACAGAGATTTAACCAACCAATTTGGTCACGCATTTGGTTTTAGAGTTTGACCATTG
 ATGCCTCGGGCAGCTCAGATGATACCAATGTTAATGTTACAGGCTATATTACACAGATTGTATTGCGTTATGGCGT
 GGGTGTGTTAATGCTGAATCAACCTTATCCATGCGTTATCAATTGACACGCCGTGTTTATATTGAAGCAACAGTGCC
 GCTGAAAATATGGTTGATGTGATTATCGTTGGAAGTTTAG

序列编号：2

由序列编号：1的多核苷酸序列推导出的粘膜炎莫拉氏菌 BASB132

多肽序列

MTHQDNSKNSSSAENGVSAGVAAPTQKPKPPNPPKQWLIYLIRLIVGAI I AALLLLAVLFAMTNSSESGSKFLIEKIAL
 ETGTKLYSEGSIRHGWWQDVKIAQSEDIITITINRAYVQLGWRALFARQVHLVNPKIDKVYVNTKPTSTGEPFYATIN
 LPVTLKLENAKVNEI IYDQVDEPVLHHAIFDHASWADSTVKIDNAML SYGDDINISHATGGIDLTGHYPLSLSDVHI
 LALDDAYFDTL SVKAGGSLKRTVGTLTGKYNQHKVGTGSFIAQGLDKNSPFSARLDFDEVRLPYADSQNILLKNGSIIADG
 VISNIELRINTELSAKDIPDGHYHGRGIIRGSTMQIPYLOADTANGTLVATGDMTWEDGYELDATITADGYRIREEMPSD
 YHEYRAYLPKVL TGSLGVKYLTDKASHDTRFEFDLNQKDGRIQATLAQNQQSDHEPWRIDATWANLIRHDIPQIGEIH
 SRSGQASVRLNGHTYINASADIVKLNVPVSGSYHVQANIEQNQHLHLTDFNYQGVMGELTGTGRVDFATAQRPLEWQLD
 VTANPVKPNAYFQTPNQTPFEQISGRIIASGRLREIDGVSIDIKVDDSDLTALLNDGKQVHLMGQGVSKIRLQDQGISH
 LKANFDGKLAQDILPQVADSSIGLDIEGDLNDLTITRAVMNDSGKVS LAGRLGLNDGIHWQIQGRLDEVDTAAAFVDDDN
 LVAIITGDLATSGRYRDGVMDIALTFDQGVINNSIPNGHLSIDATGSGNRFMVNHLRHDGTAGTLNATGWVDISQGAW
 QLEADMNSLNLGAFIQSLDLDLNGTIQLAGNWQESHQVIDISNLDITGSYNNQPLIATGSLYAKLSLPKDLAGYIQSIKQ
 ASRPPTSSDDLALKNRIDNNARQTQNI VHKLNADNLQVRIGNNHLAISGDERQLTASVNVIDLGLIDTASGAIQGGVI
 VVNDHHALPTLYIDASVSSLRFANITIQNAQAIGRIVNLGNSESQLLVQGGDIIVMGRTIKSARMDFSGTEADHILSIST
 KSGDIEASMHI DGAFNRNMRYHGVLSDFVKSGFGKMSQRQPTFSYGLDDNSLQIAAHCWQSAHIQGDGVGAICLQDT
 LSYPQSGNVNLI IQNLDTQVLLAALPSDIHWKSMNLGRIKAMWQAGQSPLVDAVLYSDDGTIGLTQEETGYIEMPYQRA
 SVIAKSVNDGLKVRTDILGTAGRGYADV IINPKQTDKPI SGALVMNDLNLAVLRPFPPSIQTL SGKVS LAGGLGGT LSKP
 LFGYANLKNGLAI VGVPMPI SDVDATVNI RGTKARLDGRFMGEGHGVLYGEMDWAEEYARLVGFGENLTVSQPPLV
 TAQISPELEV IIRFPQKFVDVQGVVSIPSATIRPPEATADIVTESPDVSVIDRRITGNIDQILSRAPWDINANIGVDLG
 NEIEFRGFGAVLPLAGAIHLTQSGQGAMQALGVVQVSKRTKIDVIGQNLDLNYAQIRFDGMDLNPRLSIEGEKQIEGQTV
 GVRIRGTASNPNITVFNDAGLDEYQAMNALVTGRISSESDLGITEQGFVSQVTNHLAAAGLSLGLSGTRDLTNQIGHAFG
 FQSLTIDASGSSDOTNVNVTGYITPDLYLRYGVGVFNAESTLSMRYQLTRRVYIEATSAENMVDV IYRWKF

序列编号：3

来源于 ATCC43617 菌株的粘膜炎莫拉氏菌 BASB132 多核苷酸序列

```

ATGAATGACTCAGGCAAAGTATCGCTTGCGCTCGTCTTGGATTAAATGACGGTATTCATTGGCAAATTCAGGGCGAT
TAGATGAGGTGGATACCGCGCTTTTGTGATGATGACAATTTAGTTGCGATCATCACAGGCGATCTTGCCACATCAGG
CAGATATCGTGATGGCGTGATGGATGATATCGCACTCACTTTTGATGGTCAAGTGATTAACAATTCATCCCAATGGT
CATCTAAGTATTGATGCGACAGGTTCTGGCAATCGCTTTATGGTCAATCATCTAAGACATGATGGCACAGCAGGCACGC
TGAATGCGACAGGCTGGGTAGATATTAGCCAAGGTGCTGCATGGCAGCTAGAGGCTGATATGAATTCATTGAATTTGGG
TGCATTTATCCAGTCGCTAGATACTGACTTAAACGGTACAATTCAGCTTGCCAGGCAACTGGCAAGAGTCTCATCAAGTC
ATTGATATCAGTAATTTGGATATTACAGGTAGCTACAATAATCAGCCACTTATCGCCACAGGCAGCTTATATGCCAAGC
TTTCCCTACCTAAAGATTTGGCTGGGTATATCCAAAGTATCAAGCAAGCAAGCCGCCACCGACTTCGCTGATGATTT
ATTGGCACTAAAAATCGTATTGATAACAACGCACGCCAAACCCAAAATATCGTTCATAAACTCAATGCTGATAATCTA
CAAGTCCGCATTGGCAATAATCATTGGCCATATCAGGCGATGAAAGACAGTTGACCGCCAGTGTGAATGTGATTGATC
TTGGGCAGCTGATTGATACCGCAAGTGGTGGCATTCAAGGCGGTGTGATTGTCTGCAATGATCATCACGCTTTGCCAAC
TTTATATATTGATGCCAGTGTCTCATCTCAGATTTGCCAATATCACCATCCAAAACGCCAAGCTATCGGTAAAATT
GTCAATCTGGGTAATAGCGAAAGCCAGCTGTTGGTTCAGGGTGATGATATTATCGTGATGGGACGCCACCTAAATCTG
CTCGGATGGATTTTAGTGGCACAGAAGCCGATCATATCTTATCAATTTCAACCAAAAGCGGTGATATCGAAGCATCTAT
GCATATTGATGGGGCATTTAATCGCAACAATATGCGATATCATGGTGTTTTGTCTGATGGTTTTGTCAAGAGTGGATTT
GGCAAAATGTCACAGCGTCAGCCGACCGAGTTTAGCTATGGATTGGATGATAATAGCTTTACAAATTCAGCGCATTGTT
GGCAATCGGCACATATCCAAGGCGATGGCGTGGGTGCGATTTGTTTACAAGATACATTAAAGTTATACACCGCAATCAGG
GAATGTCAATCTTATTATCCAAAATCTTGATACTCAAGTACTATTGGCGGCTCTACCCAGTGATATTCATTGGAAATCC
ATGCTCAATGGTCGCATTAAGGCAATGTGGCAGGCAGGTCAATCACCTTTGGTTGATGCTGTACTGTATCTCATGATG
GTACAATTGGCTTAACTCAAGAAGAGACAGGTTATATTGAGATGCCATATCAGCGTGCGTCTGTGATTGCCAAAAGTGT
TGATAATGGCTTAAAGGTAAGAACCGATATCTTGGGTACGGCAGGTGCGTGGTTATGCCGATGTAATCATCAACCCCAAG
CAAACAGACAAGCCCATTTCTGGTGCGCTTGTATGAATGATCTTAATTTGGCGGTATTGCGACCATTTTCCCGAGCA
TTCAAACATTGAGCGGTAAAGTCAGCTTGGCAGGTGGCTTAGGTGGTACATTATCCAAACCATTTGTTTTATGGTAATGC
CAATTTAAAGAATGGTGATTAGCGATTGTTGGCGTACCAATGCCCATTTCTGATGTTGATGCCACGGTGAACATACGA
GGTACTAAAGCACGCTTAGACGGTAGATTTATGGGTGGTGAGGGTCATGGTGTGCTTTATGGTGAAATGGATTGGGCAG
AAGAGCTGTATGCTCGCCTTGGCGTATTTGGTGAAAATCTGACCGTCAGCCAGCCACCGCTTGTGACGGCACAATCAG
CCCTGAGCTGGAAGTGATTATCAGACCTTTCAAAAATTTGTAGATGTTCAAGGCGTTGTTAGTATCCCATCAGCAACC
ATTCGTCGCCAGAGGCAACCGCAGATATCGTAACCGAATCCCAGATGTATCAGTCATTGACCGCGGTATCACAGGTA
ATATTGACCAGATTTTAAGTAGAGCGAAACCTTGGGATATTAATGCAAAATATCGGTGTTGATTGGGTAATGAGATTGA
ATTCGTGGATTGGGGCGGTATTACCATTTGGCAGGCGCGATACATTGACTCAGTCAGGACAAGGTGCAATGCAAGCT
CTTGGCGTGGTTCAAGTTTCTAAACGCACAAAATCGATGTCTCATCGGTCAAAATTTGGATTAAATTTATGCCAAATTC
GTTTTGATGGCGACATGCTTAATCCACGCTTATCTATTGAGGGTGAAAAACAAATTGAAGGGCAACCGGTGGGCGTTTCG
CATTCGTGGAACGGCATCAATCCAAACATTACCGTATTTAATGATGCAGGTTTGGATGAATACCAAGCAATGAATGCA
CTGGTGACAGGACGCATTAGCGAATCAAGTGATTGGGTATCACAGAGCAGGGTTTTTCGCTCACAAGTAACCAATCACC
TTGCTGCTGCTGGCTTAAGCTTAGGTTTGTGAGGAACAGAGATTTAACCACCAAAATTGGTCACGCATTGGTTTTCA
GAGTTTGACCATTGATGCCTCGGGCAGCTCAGATGATACCAATGTTAATGTTACAGGCTATATTACACAGATTTGTAT
TTGCGTTATGGCGTGGGTGTGTTAATGCTGAATCAACCTTATCCATGCGTTATCAATTGACACGCCGTGTTTATATTG
AAGCAACAGTGCCGCTGAAAATATGGTTGATGTGATTTATCGTTGGAAGTTTAG

```

序列编号：4

- 5 由序列编号：3 的多核苷酸序列推导出的粘膜炎莫拉氏菌 BASB132 多肽序列

```

MNDGKVLGRLGLNDGIHWQIQRLDEVDTAAFDVDDNLVAIITGDLATSGRYRDGVMDDIALTFDQVINNSIPNGH
LSIDATGSGNRFMVNHLRHDGTAGTLNATGWVDISQGAAWQLEADMNSLNLGAFIQSLDLDLNGTIQLAGNWQESHQVID

```

ISNLDITGSYNNQPLIATGSLYARLSLEPKDLAGYIQSIKQASRPPTSSDDLLALKNRIDNNARQTQNIHVHKLNADNLQVR
 IGNNHLAISGDERQLTASVNVNIDLGQLIDTASGAIQGGVIVVNDHHALEPTLYIDASVSSLRFANITIONAQAIGKIVNLG
 NSESQLLVQGGDIIVMGRITIKSARMDFSGTEADHILSISTKSGDIEASMRHIDGAFNRNMMRYHGVLSDFVKSFGKMSQ
 RQPTFEFSYGLDDNSLQIAAHCWQSAHIQGDGVGAICLQDTLSYTPQSGNVNLI IQNLDTQVLLAALPSDIHWKSMNLNGRI
 KAMWQAGQSPLVDAVLYSDDGTIGLTQEETGYIEMPYQRASVIAKSVDNGLKVRTDILGTAGRGYADVIINPKQTDKPIS
 GALVMNDLNLAVLRPFPSIQTLGKVS LAGGLGGTLSKPLFYGNANLKNLALAI VGVPMPI SDVDATVNI RGT KARLDG
 RFMGGEHGVLYGEMDWAEELYARLGVFGENLTVSQPPLVTAQISPELEVII RPFQKFVDVQGVVSI PSATIRPPEATAD
 IVTESPDVSVIDRRITGNIDQILSRAKPWDINANIGVDLGNEIEFRGFGAVLPLAGAIHLTQSGQGAMQALGVVQVSKRT
 KIDVIGQNLDLNYAQIREFDGDMNLNPRLSIEGEKQIEGQTVGVRI RGTASNPNITVFNDAGLDEYQAMNALVTGRISSESSD
 LGITEQGFERSQVTNHLAAAGLSGLSGTRDLTNQIGHAFGFQSLTIDASGSSDDTNVNV TGYITPDLYLRYGVGVFNAES
 TLSMRYQLTRRVYIEATSAAENMVDVIYRWKF

序列编号: 5

CCC ACC GTT TGC CCT TCA AT

5

序列编号: 6

CAT TTT TCC CGA GCA TTC AAA C

序列编号: 7

AAT CAG CCA CTT ATC GCC AC

10

序列编号: 8

GTA AAA CGA CGG CCA GT

序列编号: 9

CAG GAA ACA GCT ATG AC

15

序列编号: 10

TCA TGA ATG ACT CAG GCA AAG

20

序列编号: 11

AGA TCT AAA CTT CCA ACG ATA AAT C

序列编号: 12

AAA CAA ATC GCA CCC ACG CC

25

序列编号: 13

ACA AAT TGC AGC GCA TTG TTG G

序列表

<110> 史密丝克莱恩比彻姆生物有限公司(SmithKline Beecham Biologicals S.A.)

<120> 来源于粘膜莫拉氏菌的免疫原性多肽及其应用

<130> BM45417

<160> 13

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 5019

<212> DNA

<213> 粘膜莫拉氏菌(*Moraxella catarrhalis*)

<400> 1

atgactcatc	aggataattc	aaaaaactcg	tcatcggcag	aaaatgggggt	ttctgctggc	60
gttgctgcac	cgactaaaac	tcagcccaaa	ccaccaaac	cacccaaacc	acaatggcta	120
atctatttaa	ttcgcttaat	tgctggagca	attattgccg	ctttattgtt	gttggcgggtg	180
ttatttgcca	tgaccaacag	cgagtcaggt	tcaaagtttt	taatcgaaaa	aattgcggtta	240
gaaactggca	ctaagcttaa	gtacagcgaa	ggctcaattc	gccatggagt	ttgggtgcaa	300
gatgtcaaaa	tcgctcaaag	tgaagatatt	accatcacca	ttaatcgtgc	ctatgtacag	360
cttgggtggc	gagccttggt	tgctcgccaa	gtgcatttgg	tcaatcctaa	gattgataaa	420
gtttatgtga	caaacaccaa	gccatcaaca	ggcgaaccct	ttgattatgc	gaccatcaac	480
ctaccagtga	cgcttaagct	tgaaaatgcc	aaagtcaatg	aaattatcta	tgaccaagtg	540
gattctgagc	ctgtcgtact	gcatacatc	gcatttgatc	acgcatacat	ggcagattca	600
acagttaaaa	ttgataacgc	catgctaagc	tatggtgatg	atattaatat	cagccatgcc	660
actgggtggaa	ttgatttaac	aggtcattat	ccgctgtcgt	tgctggcaga	tgtgcatatt	720
ttggcacttg	atgatgcgta	ttttgatact	ttgtcgggtga	aagcaggcgg	tagccttaag	780
cgtaccgttg	gtacactcac	tggcaaatat	aatcagcatc	atgtgacagg	cagctttatc	840
gctcaagggt	tggataaaaa	ctcacctttt	agcgcacgcc	ttgattttga	tgaagtacga	900
ttgccttatg	ctgacagtca	aaatatttta	ctaaaaaatg	gctctatcat	cgctgatggc	960
gtcatctcaa	atatcgagct	acgcattaac	actgagttat	ccgccaaaga	tattcctgat	1020
gggcattatc	acggctcgtg	aattattcgt	ggcagtacca	tgcaaatccc	atatttgacg	1080
gctgatactg	ccaatggtac	tttgggtggc	acgggtgata	tgacttggga	agatggctat	1140
gagcttgatg	ccaccattac	agcagacggc	tatcgtatcc	gtgaagagat	gccaagtgat	1200
tatcatgaat	atagagccta	tctacctaag	gttttgacag	gctcacttgg	ggttaagtat	1260
acgctatttag	acaaagccag	tcatagatac	cggcttgagt	ttgatctgaa	tcaaaaagac	1320
gggtgaacgca	ttcaagcgac	gctggctcaa	aatcaacaga	gtgatcatga	gccttggcgt	1380
attgatgcga	cttggggcaaa	tctaatecgc	catgatattc	cacaaattgg	cgagattcat	1440
agccgctcag	gtcagggcatc	agttcgtttg	gagaacggac	atacctatat	taacgcctct	1500
gctgacattg	ttaaacttaa	tgccgtccca	agcggatcat	atcatgtgca	agcaaatatt	1560
gagcaaaaatc	agcacttgca	cttgacagat	tttaactatc	aagggtgtgat	gggcgagttg	1620
acagggtactg	gcagggttga	ttttgcgact	gcccaaagac	cgcttgagtg	gcagcttgat	1680
gtcactgccca	atccagtcaa	acccaatgcg	tattttcaaa	cacccaatca	aacaccgttt	1740
gagcaaaattt	caggcagaat	aattgcttca	gggcgtctgc	gtgagataga	tgggtgtgtca	1800
attcatgaca	ttaaagtgga	tgacagtgat	ttaaccgctt	tattaaatga	cggcaagcaa	1860
gtacatctga	tgggtcaagg	cgtcagcaag	attaggctac	aagatggaca	aattagccat	1920
cttaaggcaa	attttgatgg	caagcttgct	caagatatct	taccgcagggt	agctgactca	1980
agcattggat	tggatattga	aggtgatttg	aatgatttga	ccatcacacg	ggcagtcacg	2040
atgaatgact	caggcaaagt	atcgcttgca	ggcgtctctg	gattaaatga	cggtattcat	2100
tggcaaattc	aagggcgatt	agatgaggtg	gataccgccg	cttttggtga	tgatgacaat	2160
ttagttgcga	tcatcacagg	cgatcttgcc	acatcaggca	gatatcgtga	tggcgtgatg	2220
gatgatatcg	cactcacttt	tgatggtcaa	gtgatataca	attcaatccc	aaatggtcac	2280
ctaagtattg	atgcgacagg	ttctggcaat	cgctttatgg	tcaatcatct	aagacatgat	2340
ggcacagcag	gcacgctgaa	tgcgacaggc	tgggtagata	ttagccaagg	tgctgcatgg	2400
cagctagagg	ctgatatgaa	ttcattgaat	ttgggtgcac	ttatccagtc	gctagatact	2460
gacttaaacg	gtacaattca	gcttgccaggc	aactggcaag	agtctcatca	agtcattgat	2520

```

atcagtaatt tggatattac aggtagctac aataatcagc cacttatcgc cacaggcagc 2580
ttatatgcc aagctttccct acctaaagat ttggctgggt atatccaaag tatcaagcaa 2640
gcaagccgcc caccgacttc gtctgatgat ttattggcac taaaaaatcg tattgataac 2700
aacgcacgcc aaacccaaaa tatcgttcat aaactcaatg ctgataatct acaagtccgc 2760
attggcaata atcatttggt catatcaggg gatgaaagac agttgaccgc cagtgtgaat 2820
gtgattgac ttgggcagct gattgatacc gcaagtgggt cgattcaagg cgggtgtgatt 2880
gtcgtcaatg atcatcacgc tttgccaaact ttatatattg atgccagtgt ctcactactc 2940
agatttgcca atatcaccat ccaaaacgcc caagctatcg gtaaaattgt caatctgggt 3000
aatagcgaaa gccagctgtt gggtcaggg gatgataatta tcgtgatggg acgcaccatt 3060
aaatctgctc ggatggattt tagtggcaca gaagccgac atatcttata aatttcaacc 3120
aaaagcgggtg atatcgaagc atctatgcat attgatgggg catttaatcg caacaatatg 3180
cgatatcatg gtgttttgct tgatggtttt gtcaagagtg gatttggcaa aatgtcacag 3240
cgtcagccga ccgagtttag ctatggattg gatgataata gcttaccat tgcagcgcac 3300
tgttggcaat cggcacatat ccaaggcgat ggctgtgggt cgatttgttt acaagataca 3360
ttaagttata caccgcaatc agggatgtgc aatcttatta tccaaaatct tgatactcaa 3420
gtactatttg cggctctacc cagtgatatt cattggaaat ccattgctca tggctgcatt 3480
aaggcaatgt ggcaggcagg tcaatcacct ttggtgatg ctgtactgta ttctgatgat 3540
ggtaacaattg gcttaactca agaagagaca ggttatattg agatgcata tcagcgtgcg 3600
tctgtgattg ccaaaagtgt tgataatggc ttaaaggtaa gaaccgatat cttgggtacg 3660
gcaggtcgtg gttatgcga tgtaatcat aacccaagc aaacagacaa gccatttct 3720
ggtgcgcttg ttatgaatga tcttaatttg gcggtattgc gaccattttt cccgagcatt 3780
caaacattga gcggtaaagt cagcttggca ggtggcttag gtggtacatt atccaaacca 3840
ttgttttatg gtaatgccaa tttaaagaat ggtgcattag cgattgttg cgtaccaatg 3900
cccatttctg atgttgatgc cacgtgaac atacgaggta ctaaagcacg cttagacggg 3960
agatttatgg gtggtgaggg tcatggtgtg ctttatgggt aaatggattg ggcagaagag 4020
ctgtatgctc gccttgccgt atttggtgaa aatctgaccg tcagccagcc accgcttggt 4080
acggcacaaa tcagccctga gctggaagt attatcacac ctttccaaa attttagat 4140
gttcaaggcg ttgttagtat cccatcacga accattcgtc cgccagaggc aaccgcagat 4200
atcgtaaacc aatcccaga tgtatcagtc attgaccgcc gtatcacagg taatattgac 4260
cagattttta gtagagcgaa accttgggat attaatgcaa atatcggtgt tgatttgggt 4320
aatgagattg aatttcgtgg atttggggcg gtattaccat tggcaggcgc gatacatttg 4380
actcagtcag gacaaggtgc aatgcaagct cttggcgtgg ttcaggtttc taaacgcaca 4440
aaaatcgatg tcatcggtca aaatttggat ttaaattatg cccaaattcg ttttgatggc 4500
gacatgctta atccacgctt atctattgag ggtgaaaaac aaattgaagg gcaaacggtg 4560
ggcgttcgca ttcgtggaac ggcacaaat ccaaacatta ccgtatttaa tgatgcaggt 4620
ttggatgaat accaagcaat gaatgcactg gtgacaggac gcattagcga atcaagtgat 4680
ttgggtatca cagagcaggg ttttcgtcca caagtaacca atcacctgc tgctgctggc 4740
ttaagcttag gttgtcagg aaccagagat ttaaccaacc aaattggtca cgcatttggg 4800
tttcagagtt tgaccattga tgcctcgggc agctcagatg ataccaatgt taatgttaca 4860
ggctatatta caccagattt gtatttgcgt tatggcgtgg gtgtgtttaa tgctgaatca 4920
accttatcca tgcgttatca attgacacgc cgtgtttata ttgaagcaac cagtgccgct 4980
gaaaatatgg ttgatgtgat ttatcgttgg aagttttag 5019

```

<210> 2

<211> 1672

<212> PRT

<213> 粘膜莫拉氏菌(*Moraxella catarrhalis*)

<400> 2

```

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

```

```

Thr Ile Asn Arg Ala Tyr Val Gln Leu Gly Trp Arg Ala Leu Phe Ala
      115              120              125
Arg Gln Val His Leu Val Asn Pro Lys Ile Asp Lys Val Tyr Val Thr
      130              135              140
Asn Thr Lys Pro Ser Thr Gly Glu Pro Phe Asp Tyr Ala Thr Ile Asn
145              150              155              160
Leu Pro Val Thr Leu Lys Leu Glu Asn Ala Lys Val Asn Glu Ile Ile
      165              170              175
Tyr Asp Gln Val Asp Ser Glu Pro Val Val Leu His His Ile Ala Phe
      180              185              190
Asp His Ala Ser Trp Ala Asp Ser Thr Val Lys Ile Asp Asn Ala Met
      195              200              205
Leu Ser Tyr Gly Asp Asp Ile Asn Ile Ser His Ala Thr Gly Gly Ile
      210              215              220
Asp Leu Thr Gly His Tyr Pro Leu Ser Leu Ser Ala Asp Val His Ile
225              230              235              240
Leu Ala Leu Asp Asp Ala Tyr Phe Asp Thr Leu Ser Val Lys Ala Gly
      245              250              255
Gly Ser Leu Lys Arg Thr Val Gly Thr Leu Thr Gly Lys Tyr Asn Gln
      260              265              270
His His Val Thr Gly Ser Phe Ile Ala Gln Gly Leu Asp Lys Asn Ser
      275              280              285
Pro Phe Ser Ala Arg Leu Asp Phe Asp Glu Val Arg Leu Pro Tyr Ala
      290              295              300
Asp Ser Gln Asn Ile Leu Leu Lys Asn Gly Ser Ile Ile Ala Asp Gly
305              310              315              320
Val Ile Ser Asn Ile Glu Leu Arg Ile Asn Thr Glu Leu Ser Ala Lys
      325              330              335
Asp Ile Pro Asp Gly His Tyr His Gly Arg Gly Ile Ile Arg Gly Ser
      340              345              350
Thr Met Gln Ile Pro Tyr Leu Gln Ala Asp Thr Ala Asn Gly Thr Leu
      355              360              365
Val Ala Thr Gly Asp Met Thr Trp Glu Asp Gly Tyr Glu Leu Asp Ala
      370              375              380
Thr Ile Thr Ala Asp Gly Tyr Arg Ile Arg Glu Glu Met Pro Ser Asp
385              390              395              400
Tyr His Glu Tyr Arg Ala Tyr Leu Pro Lys Val Leu Thr Gly Ser Leu
      405              410              415
Gly Val Lys Tyr Thr Leu Leu Asp Lys Ala Ser His Asp Thr Arg Phe
      420              425              430
Glu Phe Asp Leu Asn Gln Lys Asp Gly Glu Arg Ile Gln Ala Thr Leu
      435              440              445
Ala Gln Asn Gln Gln Ser Asp His Glu Pro Trp Arg Ile Asp Ala Thr
      450              455              460
Trp Ala Asn Leu Ile Arg His Asp Ile Pro Gln Ile Gly Glu Ile His
465              470              475              480
Ser Arg Ser Gly Gln Ala Ser Val Arg Leu Glu Asn Gly His Thr Tyr
      485              490              495
Ile Asn Ala Ser Ala Asp Ile Val Lys Leu Asn Ala Val Pro Ser Gly
      500              505              510
Ser Tyr His Val Gln Ala Asn Ile Glu Gln Asn Gln His Leu His Leu
      515              520              525
Thr Asp Phe Asn Tyr Gln Gly Val Met Gly Glu Leu Thr Gly Thr Gly
      530              535              540
Arg Val Asp Phe Ala Thr Ala Gln Arg Pro Leu Glu Trp Gln Leu Asp
545              550              555              560
Val Thr Ala Asn Pro Val Lys Pro Asn Ala Tyr Phe Gln Thr Pro Asn
      565              570              575
Gln Thr Pro Phe Glu Gln Ile Ser Gly Arg Ile Ile Ala Ser Gly Arg
      580              585              590
Leu Arg Glu Ile Asp Gly Val Ser Ile His Asp Ile Lys Val Asp Asp
      595              600              605
Ser Asp Leu Thr Ala Leu Leu Asn Asp Gly Lys Gln Val His Leu Met

```

610	615	620
Gly Gln Gly Val Ser Lys Ile Arg Leu Gln Asp Gly Gln Ile Ser His		
625	630	635
Leu Lys Ala Asn Phe Asp Gly Lys Leu Ala Gln Asp Ile Leu Pro Gln		640
	645	650
Val Ala Asp Ser Ser Ile Gly Leu Asp Ile Glu Gly Asp Leu Asn Asp		655
	660	665
Leu Thr Ile Thr Arg Ala Val Met Met Asn Asp Ser Gly Lys Val Ser		670
	675	680
Leu Ala Gly Arg Leu Gly Leu Asn Asp Gly Ile His Trp Gln Ile Gln		685
	690	695
Gly Arg Leu Asp Glu Val Asp Thr Ala Ala Phe Val Asp Asp Asp Asn		700
705	710	715
Leu Val Ala Ile Ile Thr Gly Asp Leu Ala Thr Ser Gly Arg Tyr Arg		720
	725	730
Asp Gly Val Met Asp Asp Ile Ala Leu Thr Phe Asp Gly Gln Val Ile		735
	740	745
Asn Asn Ser Ile Pro Asn Gly His Leu Ser Ile Asp Ala Thr Gly Ser		750
	755	760
Gly Asn Arg Phe Met Val Asn His Leu Arg His Asp Gly Thr Ala Gly		765
	770	775
Thr Leu Asn Ala Thr Gly Trp Val Asp Ile Ser Gln Gly Ala Ala Trp		780
785	790	795
Gln Leu Glu Ala Asp Met Asn Ser Leu Asn Leu Gly Ala Phe Ile Gln		800
	805	810
Ser Leu Asp Thr Asp Leu Asn Gly Thr Ile Gln Leu Ala Gly Asn Trp		815
	820	825
Gln Glu Ser His Gln Val Ile Asp Ile Ser Asn Leu Asp Ile Thr Gly		830
	835	840
Ser Tyr Asn Asn Gln Pro Leu Ile Ala Thr Gly Ser Leu Tyr Ala Lys		845
	850	855
Leu Ser Leu Pro Lys Asp Leu Ala Gly Tyr Ile Gln Ser Ile Lys Gln		860
865	870	875
Ala Ser Arg Pro Pro Thr Ser Ser Asp Asp Leu Leu Ala Leu Lys Asn		880
	885	890
Arg Ile Asp Asn Asn Ala Arg Gln Thr Gln Asn Ile Val His Lys Leu		895
	900	905
Asn Ala Asp Asn Leu Gln Val Arg Ile Gly Asn Asn His Leu Ala Ile		910
	915	920
Ser Gly Asp Glu Arg Gln Leu Thr Ala Ser Val Asn Val Ile Asp Leu		925
	930	935
Gly Gln Leu Ile Asp Thr Ala Ser Gly Ala Ile Gln Gly Gly Val Ile		940
945	950	955
Val Val Asn Asp His His Ala Leu Pro Thr Leu Tyr Ile Asp Ala Ser		960
	965	970
Val Ser Ser Leu Arg Phe Ala Asn Ile Thr Ile Gln Asn Ala Gln Ala		975
	980	985
Ile Gly Lys Ile Val Asn Leu Gly Asn Ser Glu Ser Gln Leu Leu Val		990
	995	1000
Gln Gly Asp Asp Ile Ile Val Met Gly Arg Thr Ile Lys Ser Ala Arg		1005
	1010	1015
Met Asp Phe Ser Gly Thr Glu Ala Asp His Ile Leu Ser Ile Ser Thr		1020
1025	1030	1035
Lys Ser Gly Asp Ile Glu Ala Ser Met His Ile Asp Gly Ala Phe Asn		1040
	1045	1050
Arg Asn Asn Met Arg Tyr His Gly Val Leu Ser Asp Gly Phe Val Lys		1055
	1060	1065
Ser Gly Phe Gly Lys Met Ser Gln Arg Gln Pro Thr Glu Phe Ser Tyr		1070
	1075	1080
Gly Leu Asp Asp Asn Ser Leu Gln Ile Ala Ala His Cys Trp Gln Ser		1085
	1090	1095
Ala His Ile Gln Gly Asp Gly Val Gly Ala Ile Cys Leu Gln Asp Thr		1100
1105	1110	1115
		112

```

Leu Ser Tyr Thr Pro Gln Ser Gly Asn Val Asn Leu Ile Ile Gln Asn
      1125      1130      1135
Leu Asp Thr Gln Val Leu Leu Ala Ala Leu Pro Ser Asp Ile His Trp
      1140      1145      1150
Lys Ser Met Leu Asn Gly Arg Ile Lys Ala Met Trp Gln Ala Gly Gln
      1155      1160      1165
Ser Pro Leu Val Asp Ala Val Leu Tyr Ser Asp Asp Gly Thr Ile Gly
      1170      1175      1180
Leu Thr Gln Glu Glu Thr Gly Tyr Ile Glu Met Pro Tyr Gln Arg Ala
      1185      1190      1195      120
Ser Val Ile Ala Lys Ser Val Asp Asn Gly Leu Lys Val Arg Thr Asp
      1205      1210      1215
Ile Leu Gly Thr Ala Gly Arg Gly Tyr Ala Asp Val Ile Ile Asn Pro
      1220      1225      1230
Lys Gln Thr Asp Lys Pro Ile Ser Gly Ala Leu Val Met Asn Asp Leu
      1235      1240      1245
Asn Leu Ala Val Leu Arg Pro Phe Phe Pro Ser Ile Gln Thr Leu Ser
      1250      1255      1260
Gly Lys Val Ser Leu Ala Gly Gly Leu Gly Gly Thr Leu Ser Lys Pro
      1265      1270      1275      128
Leu Phe Tyr Gly Asn Ala Asn Leu Lys Asn Gly Ala Leu Ala Ile Val
      1285      1290      1295
Gly Val Pro Met Pro Ile Ser Asp Val Asp Ala Thr Val Asn Ile Arg
      1300      1305      1310
Gly Thr Lys Ala Arg Leu Asp Gly Arg Phe Met Gly Gly Glu Gly His
      1315      1320      1325
Gly Val Leu Tyr Gly Glu Met Asp Trp Ala Glu Glu Leu Tyr Ala Arg
      1330      1335      1340
Leu Gly Val Phe Gly Glu Asn Leu Thr Val Ser Gln Pro Pro Leu Val
      1345      1350      1355      136
Thr Ala Gln Ile Ser Pro Glu Leu Glu Val Ile Ile Arg Pro Phe Gln
      1365      1370      1375
Lys Phe Val Asp Val Gln Gly Val Val Ser Ile Pro Ser Ala Thr Ile
      1380      1385      1390
Arg Pro Pro Glu Ala Thr Ala Asp Ile Val Thr Glu Ser Pro Asp Val
      1395      1400      1405
Ser Val Ile Asp Arg Arg Ile Thr Gly Asn Ile Asp Gln Ile Leu Ser
      1410      1415      1420
Arg Ala Lys Pro Trp Asp Ile Asn Ala Asn Ile Gly Val Asp Leu Gly
      1425      1430      1435      144
Asn Glu Ile Glu Phe Arg Gly Phe Gly Ala Val Leu Pro Leu Ala Gly
      1445      1450      1455
Ala Ile His Leu Thr Gln Ser Gly Gln Gly Ala Met Gln Ala Leu Gly
      1460      1465      1470
Val Val Gln Val Ser Lys Arg Thr Lys Ile Asp Val Ile Gly Gln Asn
      1475      1480      1485
Leu Asp Leu Asn Tyr Ala Gln Ile Arg Phe Asp Gly Asp Met Leu Asn
      1490      1495      1500
Pro Arg Leu Ser Ile Glu Gly Glu Lys Gln Ile Glu Gly Gln Thr Val
      1505      1510      1515      152
Gly Val Arg Ile Arg Gly Thr Ala Ser Asn Pro Asn Ile Thr Val Phe
      1525      1530      1535
Asn Asp Ala Gly Leu Asp Glu Tyr Gln Ala Met Asn Ala Leu Val Thr
      1540      1545      1550
Gly Arg Ile Ser Glu Ser Ser Asp Leu Gly Ile Thr Glu Gln Gly Phe
      1555      1560      1565
Arg Ser Gln Val Thr Asn His Leu Ala Ala Ala Gly Leu Ser Leu Gly
      1570      1575      1580
Leu Ser Gly Thr Arg Asp Leu Thr Asn Gln Ile Gly His Ala Phe Gly
      1585      1590      1595      160
Phe Gln Ser Leu Thr Ile Asp Ala Ser Gly Ser Ser Asp Asp Thr Asn
      1605      1610      1615
Val Asn Val Thr Gly Tyr Ile Thr Pro Asp Leu Tyr Leu Arg Tyr Gly

```

1620 1625 1630
 Val Gly Val Phe Asn Ala Glu Ser Thr Leu Ser Met Arg Tyr Gln Leu
 1635 1640 1645
 Thr Arg Arg Val Tyr Ile Glu Ala Thr Ser Ala Ala Glu Asn Met Val
 1650 1655 1660
 Asp Val Ile Tyr Arg Trp Lys Phe
 1665 1670

<210> 3

<211> 2979

<212> DNA

<213> 粘膜炎莫拉氏菌(*Moraxella catarrhalis*)

<400> 3

```

atgaatgact caggcaaaagt atcgcttgca ggctgctctg gattaaatga cgggtattcat      60
tggcaaatc aagggcgatt agatgagggt gataccgccg cttttgttga tgatgacaat      120
ttagttgcga tcatcacagg cgatcttgcc acatcaggca gatatcgtga tggcgtgatg      180
gatgatatcg cactcacttt tgatgggtcaa gtgattaaca attcaatccc aaatggatcat      240
ctaagtattg atgcgacagg ttctggcaat cgctttatgg tcaatcatct aagacatgat      300
ggcacagcag gcacgctgaa tgcgacaggc tgggtagata ttagccaagg tgctgcatgg      360
cagctagagg ctgatatgaa ttcatatgaat ttgggtgcat ttatccagtc gctagatact      420
gacttaaacg gtacaattca gcttgacagg aactggcaag agtctcatca agtcattgat      480
atcagtaatt tggatattac aggtagctac aataatcagc cacttatcgc cacaggcagc      540
ttatatgcc aagctttccct acctaaagat ttggctgggt atatccaaag tatcaagcaa      600
gcaagccgcc caccgacttc gtctgatgat ttattggcac taaaaaatcg tattgataac      660
aacgcacgcc aaacccaaaa tatcggttcat aaactcaatg ctgataatct acaagtcgcg      720
attggcaata atcattttggc catatcaggc gatgaaagac agttgaccgc cagtgtgatt      780
gtgattgatc ttgggcagct gattgatacc gcaagtggtg cgattcaagg cgggtgtgatt      840
gtcgtcaatg atcatcacgc tttgccaaact ttatatattg atgccagtgt ctcatcactc      900
agatttgcca atatcaccat ccaaaacgcc caagctatcg gtaaaattgt caatctgggt      960
aatagcgaag gccagctggt gggtcagggt gatgatatta tcgtgatggg acgcaccatt     1020
aaatctgctc ggatggattt tagtggcaca gaagccgatc atatcttatt aatttcaacc     1080
aaaagcgggtg atatcgaagc atctatgcat attgatgggg catttaatcg caacaatatg     1140
cgatatcatg gtgtttttgtc tgatggtttt gtcaagagtg gatttggcaa aatgtcacag     1200
cgtcagccga ccgagtttag ctatggattg gatgataata gcttacaagt tgcagcgcag     1260
tggttgcaat cggcacatat ccaaggcgat ggctgggggt cgatttgggt acaagataca     1320
ttaagttata caccgcaatc agggaaatgtc aatcttatta tccaaaatct tgatactcaa     1380
gtactatttg cggctctacc cagtgatatt cattggaaat ccatgctcaa tggtcgcatt     1440
aaggcaatgt ggcaggcagg tcaatcacct ttgggtgatg ctgtactgta ttctgatgat     1500
ggtacaattg gcttaactca agaagagaca ggttatattg agatgccata tcagcgtgag     1560
tctgtgattg ccaaaagtgt tgataatggc ttaaaggtaa gaaccgatat cttgggtacg     1620
gcaggtcgtg gttatgccga tgtaatcatc aaccccaagc aaacagacaa gcccatctct     1680
ggtgctgctg ttatgaatga tcttaatttg cgggtattgc gaccattttt cccgagcatt     1740
caaacattga gcggtaaagt cagcttggca ggtggttagt gtggtacatt atccaaacca     1800
ttgttttatg gtaatgcca tttaaagaat ggtgcattag cgattgttgg cgtaccaatg     1860
cccatttctg atgttgatgc caaggtgaac atacgaggta ctaaagcacg cttagacggt     1920
agatttatgg gtggtagagg tcatgggtgt ctttatgggt aaatggattg ggcagaagag     1980
ctgttatgctc gccttggcgt atttgggtgaa aatctgaccg tcagccagcc accgcttggt     2040
acggcacaaa tcagccctga gctggaagtg attatcagac ctttccaaaa attttagatg     2100
gttcaaggcg ttgttagtat cccatcagca accattcgct cgccagaggc aaccgcagat     2160
atcgtaaccg aatccccaga tgtatcagtc attgaccgcc gtatcacagg taatattgac     2220
cagattttta gtagagcgaa accttgggat attaatgcaa atatcggtgt tgatttgggt     2280
aatgagattg aatttcgtgg atttggggcg gtattacat tggcaggcgc gatacatttg     2340
actcagtcag gacaagggtg aatgcaagct cttggcgtgg ttcagggttc taaacgcaca     2400
aaaatcgatg tcatcggtca aaatttggat ttaaattatg cccaaattcg ttttgatggc     2460
gacatgctta atccacgctt atctattgag ggtgaaaaac aaattgaagg gcaaacgggtg     2520
ggcgttcgca ttcgtggaac ggcatcaaat ccaaacatta ccgtatttaa tgatgcagggt     2580
ttggatgaat accaagcaat gaatgcaact gtgacaggac gcattagcga atcaagtgat     2640
ttgggtatca cagagcaggg ttttcgctca caagtaacca atcaccttgc tgctgctggc     2700
ttaagcttag gtttgtcagg aaccagagat ttaaccaacc aaattggtca cgcatttgggt     2760
tttcagagtt tgaccattga tgcctcgggc agctcagatg ataccaatgt taatgttaca     2820
ggctatatta caccagattt gtatttgcgt tatggcgtgg gtgtgtttaa tgctgaatca     2880
accttatoca tgcgttatca attgacacgc cgtgtttata ttgaagcaac cagtgccgct     2940

```

gaaaatatgg ttgatgtgat ttatcgttgg aagtttttag

2979

<210> 4

<211> 992

<212> PRT

<213> 粘膜炎莫拉氏菌(*Moraxella catarrhalis*)

<400> 4

```

Met Asn Asp Ser Gly Lys Val Ser Leu Ala Gly Arg Leu Gly Leu Asn
 1          5          10          15
Asp Gly Ile His Trp Gln Ile Gln Gly Arg Leu Asp Glu Val Asp Thr
          20          25          30
Ala Ala Phe Val Asp Asp Asp Asn Leu Val Ala Ile Ile Thr Gly Asp
          35          40          45
Leu Ala Thr Ser Gly Arg Tyr Arg Asp Gly Val Met Asp Asp Ile Ala
50          55          60
Leu Thr Phe Asp Gly Gln Val Ile Asn Asn Ser Ile Pro Asn Gly His
65          70          75          80
Leu Ser Ile Asp Ala Thr Gly Ser Gly Asn Arg Phe Met Val Asn His
          85          90          95
Leu Arg His Asp Gly Thr Ala Gly Thr Leu Asn Ala Thr Gly Trp Val
          100          105          110
Asp Ile Ser Gln Gly Ala Ala Trp Gln Leu Glu Ala Asp Met Asn Ser
          115          120          125
Leu Asn Leu Gly Ala Phe Ile Gln Ser Leu Asp Thr Asp Leu Asn Gly
130          135          140
Thr Ile Gln Leu Ala Gly Asn Trp Gln Glu Ser His Gln Val Ile Asp
145          150          155          160
Ile Ser Asn Leu Asp Ile Thr Gly Ser Tyr Asn Asn Gln Pro Leu Ile
          165          170          175
Ala Thr Gly Ser Leu Tyr Ala Lys Leu Ser Leu Pro Lys Asp Leu Ala
          180          185          190
Gly Tyr Ile Gln Ser Ile Lys Gln Ala Ser Arg Pro Pro Thr Ser Ser
          195          200          205
Asp Asp Leu Leu Ala Leu Lys Asn Arg Ile Asp Asn Asn Ala Arg Gln
210          215          220
Thr Gln Asn Ile Val His Lys Leu Asn Ala Asp Asn Leu Gln Val Arg
225          230          235          240
Ile Gly Asn Asn His Leu Ala Ile Ser Gly Asp Glu Arg Gln Leu Thr
          245          250          255
Ala Ser Val Asn Val Ile Asp Leu Gly Gln Leu Ile Asp Thr Ala Ser
          260          265          270
Gly Ala Ile Gln Gly Gly Val Ile Val Val Asn Asp His His Ala Leu
275          280          285
Pro Thr Leu Tyr Ile Asp Ala Ser Val Ser Ser Leu Arg Phe Ala Asn
290          295          300
Ile Thr Ile Gln Asn Ala Gln Ala Ile Gly Lys Ile Val Asn Leu Gly
305          310          315          320
Asn Ser Glu Ser Gln Leu Leu Val Gln Gly Asp Asp Ile Ile Val Met
          325          330          335
Gly Arg Thr Ile Lys Ser Ala Arg Met Asp Phe Ser Gly Thr Glu Ala
          340          345          350
Asp His Ile Leu Ser Ile Ser Thr Lys Ser Gly Asp Ile Glu Ala Ser
          355          360          365
Met His Ile Asp Gly Ala Phe Asn Arg Asn Asn Met Arg Tyr His Gly
370          375          380
Val Leu Ser Asp Gly Phe Val Lys Ser Gly Phe Gly Lys Met Ser Gln
385          390          395          400
Arg Gln Pro Thr Glu Phe Ser Tyr Gly Leu Asp Asp Asn Ser Leu Gln
          405          410          415
Ile Ala Ala His Cys Trp Gln Ser Ala His Ile Gln Gly Asp Gly Val
          420          425          430
Gly Ala Ile Cys Leu Gln Asp Thr Leu Ser Tyr Thr Pro Gln Ser Gly

```

435	440	445
Asn Val Asn Leu Ile Ile	Gln Asn Leu Asp Thr	Gln Val Leu Leu Ala
450	455	460
Ala Leu Pro Ser Asp Ile	His Trp Lys Ser Met	Leu Asn Gly Arg Ile
465	470	475
Lys Ala Met Trp Gln Ala	Gly Gln Ser Pro Leu	Val Asp Ala Val Leu
485	490	495
Tyr Ser Asp Asp Gly Thr	Ile Gly Leu Thr Gln	Glu Glu Thr Gly Tyr
500	505	510
Ile Glu Met Pro Tyr Gln	Arg Ala Ser Val Ile	Ala Lys Ser Val Asp
515	520	525
Asn Gly Leu Lys Val Arg	Thr Asp Ile Leu Gly	Thr Ala Gly Arg Gly
530	535	540
Tyr Ala Asp Val Ile Ile	Asn Pro Lys Gln Thr	Asp Lys Pro Ile Ser
545	550	555
Gly Ala Leu Val Met Asn	Asp Leu Asn Leu Ala	Val Leu Arg Pro Phe
565	570	575
Phe Pro Ser Ile Gln Thr	Leu Ser Gly Lys Val	Ser Leu Ala Gly Gly
580	585	590
Leu Gly Gly Thr Leu Ser	Lys Pro Leu Phe Tyr	Gly Asn Ala Asn Leu
595	600	605
Lys Asn Gly Ala Leu Ala	Ile Val Gly Val Pro	Met Pro Ile Ser Asp
610	615	620
Val Asp Ala Thr Val Asn	Ile Arg Gly Thr Lys	Ala Arg Leu Asp Gly
625	630	635
Arg Phe Met Gly Gly Glu	Gly His Gly Val Leu	Tyr Gly Glu Met Asp
645	650	655
Trp Ala Glu Glu Leu Tyr	Ala Arg Leu Gly Val	Phe Gly Glu Asn Leu
660	665	670
Thr Val Ser Gln Pro Pro	Leu Val Thr Ala Gln	Ile Ser Pro Glu Leu
675	680	685
Glu Val Ile Ile Arg Pro	Phe Gln Lys Phe Val	Asp Val Gln Gly Val
690	695	700
Val Ser Ile Pro Ser Ala	Thr Ile Arg Pro Pro	Glu Ala Thr Ala Asp
705	710	715
Ile Val Thr Glu Ser Pro	Asp Val Ser Val Ile	Asp Arg Arg Ile Thr
725	730	735
Gly Asn Ile Asp Gln Ile	Leu Ser Arg Ala Lys	Pro Trp Asp Ile Asn
740	745	750
Ala Asn Ile Gly Val Asp	Leu Gly Asn Glu Ile	Glu Phe Arg Gly Phe
755	760	765
Gly Ala Val Leu Pro Leu	Ala Gly Ala Ile His	Leu Thr Gln Ser Gly
770	775	780
Gln Gly Ala Met Gln Ala	Leu Gly Val Val Gln	Val Ser Lys Arg Thr
785	790	795
Lys Ile Asp Val Ile Gly	Gln Asn Leu Asp Leu	Asn Tyr Ala Gln Ile
805	810	815
Arg Phe Asp Gly Asp Met	Leu Asn Pro Arg Leu	Ser Ile Glu Gly Glu
820	825	830
Lys Gln Ile Glu Gly Gln	Thr Val Gly Val Arg	Ile Arg Gly Thr Ala
835	840	845
Ser Asn Pro Asn Ile Thr	Val Phe Asn Asp Ala	Gly Leu Asp Glu Tyr
850	855	860
Gln Ala Met Asn Ala Leu	Val Thr Gly Arg Ile	Ser Glu Ser Ser Asp
865	870	875
Leu Gly Ile Thr Glu Gln	Gly Phe Arg Ser Gln	Val Thr Asn His Leu
885	890	895
Ala Ala Ala Gly Leu Ser	Leu Gly Leu Ser Gly	Thr Arg Asp Leu Thr
900	905	910
Asn Gln Ile Gly His Ala	Phe Gly Phe Gln Ser	Leu Thr Ile Asp Ala
915	920	925
Ser Gly Ser Ser Asp Asp	Thr Asn Val Asn Val	Thr Gly Tyr Ile Thr
930	935	940

<400> 9
caggaaacag ctatgac 17

<210> 10	
<211> 21	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物	
<400> 10	
tcatgaatga ctcaggcaaa g	21
<210> 11	
<211> 25	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物	
<400> 11	
agatctaaac ttccaacgat aaatc	25
<210> 12	
<211> 20	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物	
<400> 12	
aaacaaatcg caccacgcc	20
<210> 13	
<211> 22	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物	
<400> 13	
acaaattgca gcgcattggt gg	22

申请人或代理人档案号	国际申请号 PCT/EP00/09495
------------	----------------------

关于微生物保藏的说明
(细则 13 之二)

A. 对说明书 第 42 页第 17 行 — 第 43 页第 2 行 所述的微生物的说明	
B. 保藏事项	
保藏单位名称	美国典型培养物保藏中心
保藏单位地址 (包括邮政编码和国名) 10801 University Blvd. Manassas, Virginia 20110-2209 USA	
保藏日期	1997 年 6 月 21 日和 1999 年 2 月 12 日
保藏编号	43617 和 207105
C. 补充说明(必要时)	
对于寻求欧洲专利的指定, 保藏的微生物样品将在欧洲专利被授予或者本申请被驳回或撤回或被视为撤回之日向公众发放, 且发放对象限于请求人指定的专家。	
D. 本说明是为下列指定国作的(如果说明不是为所有指定国而作的)	
E. 补充说明(必要时)	
下列说明将随后向国际局提供(写出说明的类别, 例如: “保藏的编号”)	

由受理局填写
☐ 本页已经和国际申请一起收到
授权官员

由国际局填写
☐ 国际局收到本页日期:
授权官员

图 1: BASB132多核苷酸序列的对比.
与序列编号:1相同处用圆点表示. 间隔处用短线表示.

*	20	*
Seqid1	: ATGACTCATCAGGATAATTCAAAAAACTCG	: 30
Seqid3	: -----	: -
	40	*
Seqid1	: TCATCGGCAGAAAATGGGGTTTCTGCTGGC	: 60
Seqid3	: -----	: -
	*	80
Seqid1	: GTTGCTGCACCGACTAAAACTCAGCCCAA	: 90
Seqid3	: -----	: -
	100	*
Seqid1	: CCACCAAACCCACCCAAACCACAATGGCTA	: 120
Seqid3	: -----	: -
	*	140
Seqid1	: ATCTATTTAATTCGCTTAATTGTCGGAGCA	: 150
Seqid3	: -----	: -
	160	*
		180

Seqid1 : ATTATTGCCGCTTTATTGTTGTTGGCGGTG : 180

Seqid3 : ----- :

* 200 *

Seqid1 : TTATTTGCCATGACCAACAGCGAGTCAGGT : 210

Seqid3 : ----- : -

220 * 240

Seqid1 : TCAAAGTTTTTAATCGAAAAATTGCGTTA : 240

Seqid3 : ----- : -

* 260 *

Seqid1 : GAAACTGGCACTAAGCTTAAGTACAGCGAA : 270

Seqid3 : ----- : -

280 * 300

Seqid1 : GGCTCAATTCGCCATGGAGTTTGGGTGCAA : 300

Seqid3 : ----- : -

* 320 *

Seqid1 : GATGTCAAATCGCTCAAAGTGAAGATATT : 330

Seqid3 : ----- : -

	340	*	360
Seqid1	: ACCATCACCATTAATCGTGCCTATGTACAG	:	360
Seqid3	: -----	:	-

	*	380	*
Seqid1	: CTTGGGTGGCGAGCCTTGTTTGCTCGCCAA	:	390
Seqid3	: -----	:	-

	400	*	420
Seqid1	: GTGCATTTGGTCAATCCTAAGATTGATAAA	:	420
Seqid3	: -----	:	-

	*	440	*
Seqid1	: GTTTATGTGACAAACACCAAGCCATCAACA	:	450
Seqid3	: -----	:	-

	460	*	480
Seqid1	: GGCGAACCCTTTGATTATGCGACCATCAAC	:	480
Seqid3	: -----	:	-

	*	500	*
Seqid1	: CTACCAGTGACGCTTAAGCTTGAAAATGCC	:	510

Seqid3 : ----- : -

520 * 540

Seqid1 : AAAGTCAATGAAATTATCTATGACCAAGTG : 540

Seqid3 : ----- : -

* 560 *

Seqid1 : GATTCTGAGCCTGTCGTACTGCATCATATC : 570

Seqid3 : ----- : -

580 * 600

Seqid1 : GCATTTGATCACGCATCATGGGCAGATTCA : 600

Seqid3 : ----- : -

* 620 *

Seqid1 : ACAGTTAAAATTGATAACGCCATGCTAAGC : 630

Seqid3 : ----- : -

640 * 660

Seqid1 : TATGGTGATGATATTAATATCAGCCATGCC : 660

Seqid3 : ----- : -

	*	680	*
Seqid1	:	ACTGGTGG AATTGATTTAACAGGTCATTAT	: 690
Seqid3	:	-----	: -
	700	*	720
Seqid1	:	CCGCTGTCGTTGTCGGCAGATGTGCATATT	: 720
Seqid3	:	-----	: -
	*	740	*
Seqid1	:	TTGGCACTTGATGATGCGTATTTTGATACT	: 750
Seqid3	:	-----	: -
	760	*	780
Seqid1	:	TTGTCGGTGAAAGCAGGCGGTAGCCTTAAG	: 780
Seqid3	:	-----	: -
	*	800	*
Seqid1	:	CGTACCGTTGGTACACTCACTGGCAAATAT	: 810
Seqid3	:	-----	: -
	820	*	840

Seqid1 : AATCAGCATCATGTGACAGGCAGCTTTATC : 840

Seqid3 : ----- : -

* 860 *

Seqid1 : GCTCAAGGGTTGGATAAAACTCACCTTTT : 870

Seqid3 : ----- : -

880 * 900

Seqid1 : AGCGCACGCCTTGATTTTGATGAAGTACGA : 900

Seqid3 : ----- : -

* 920 *

Seqid1 : TTGCCTTATGCTGACAGTCAAAATATTTTA : 930

Seqid3 : ----- : -

940 * 960

Seqid1 : CTAAAAAATGGCTCTATCATCGCTGATGGC : 960

Seqid3 : ----- : -

* 980 *

Seqid1 : GTCATCTCAAATATCGAGCTACGCATTAAC : 990

```
Seqid3 : ----- : -

          1000          *          1020
Seqid1 : ACTGAGTTATCCGCCAAAGATATTCCTGAT : 1020
Seqid3 : ----- : -

          *          1040          *
Seqid1 : GGGCATTATCACGGTCGTGGAATTATTCGT : 1050
Seqid3 : ----- : -

          1060          *          1080
Seqid1 : GGCAGTACCATGCAAATCCCATATTTGCAG : 1080
Seqid3 : ----- : -

          *          1100          *
Seqid1 : GCTGATACTGCCAATGGTACTTTGGTGGCA : 1110
Seqid3 : ----- : -

          1120          *          1140
Seqid1 : ACGGGTGATATGACTTGGGAAGATGGCTAT : 1140
Seqid3 : ----- : -
```

* 1160 *

Seqid1 : GAGCTTGATGCCACCATTACAGCAGACGGC : 1170

Seqid3 : ----- : -

1180 * 1200

Seqid1 : TATCGTATCCGTGAAGAGATGCCAAGTGAT : 1200

Seqid3 : ----- : -

* 1220 *

Seqid1 : TATCATGAATATAGAGCCTATCTACCTAAG : 1230

Seqid3 : ----- : -

1240 * 1260

Seqid1 : GTTTTGACAGGCTCACTTGGGGTTAAGTAT : 1260

Seqid3 : ----- : -

* 1280 *

Seqid1 : ACGCTATTAGACAAAGCCAGTCATGATACT : 1290

Seqid3 : ----- : -

1300 * 1320

Seqid1 : CGGTTTGAGTTTGATCTGAATCAAAAAGAC : 1320

Seqid3 : ----- : -

* 1340 *

Seqid1 : GGTGAACGCATTCAAGCGACGCTGGCTCAA : 1350

Seqid3 : ----- : -

1360 * 1380

Seqid1 : AATCAACAGAGTGATCATGAGCCTTGGCGT : 1380

Seqid3 : ----- : -

* 1400 *

Seqid1 : ATTGATGCGACTTGGGCAAATCTAATCCGC : 1410

Seqid3 : ----- : -

1420 * 1440

Seqid1 : CATGATATTCCACAAATTGGCGAGATTCAT : 1440

Seqid3 : ----- : -

* 1460 *

Seqid1 : AGCCGCTCAGGTCAGGCATCAGTTCGTTTG : 1470

Seqid3 : ----- : -

1480 * 1500

Seqid1 : GAGAACGGACATACCTATATTAACGCCTCT : 1500

Seqid3 : ----- : -

* 1520 *

Seqid1 : GCTGACATTGTTAACTTAATGCCGTCCCA : 1530

Seqid3 : ----- : -

1540 * 1560

Seqid1 : AGCGGATCATATCATGTGCAAGCAAATATT : 1560

Seqid3 : ----- : -

* 1580 *

Seqid1 : GAGCAAAATCAGCACTTGCACTTGACAGAT : 1590

Seqid3 : ----- : -

1600 * 1620

Seqid1 : TTAACTATCAAGGTGTGATGGGCGAGTTG : 1620

Seqid3 : ----- : -

* 1640 *

Seqid1 : ACAGGTACTGGCAGGGTTGATTTTGCGACT : 1650

Seqid3 : ----- : -

1660 * 1680

Seqid1 : GCCCAAAGACCGCTTGAGTGGCAGCTTGAT : 1680

Seqid3 : ----- : -

* 1700 *

Seqid1 : GTCACTGCCAATCCAGTCAAACCCAATGCG : 1710

Seqid3 : ----- : -

1720 * 1740

Seqid1 : TATTTTCAAACACCCAATCAAACACCGTTT : 1740

Seqid3 : ----- : -

* 1760 *

Seqid1 : GAGCAAATTTTCAGGCAGAATAATTGCTTCA : 1770

Seqid3 : ----- : -

1780 * 1800

Seqid1 : GGGCGTCTGCGTGAGATAGATGGTGTGTCA : 1800

Seqid3 : ----- : -

* 1820 *

Seqid1 : ATTCATGACATTAAAGTGGATGACAGTGAT : 1830

Seqid3 : ----- : -

1840 * 1860

Seqid1 : TTAACCGCTTTATTAAATGACGGCAAGCAA : 1860

Seqid3 : ----- : -

* 1880 *

Seqid1 : GTACATCTGATGGGTCAAGGCGTCAGCAAG : 1890

Seqid3 : ----- : -

1900 * 1920

Seqid1 : ATTAGGCTACAAGATGGACAAATTAGCCAT : 1920

Seqid3 : ----- : -

* 1940 *

Seqid1 : CTTAAGGCAAATTTTGATGGCAAGCTTGCT : 1950

Seqid3 : ----- :

1960 * 1980

Seqid1 : CAAGATATCTTACCGCAGGTAGCTGACTCA : 1980

Seqid3 : ----- : -

* 2000 *

Seqid1 : AGCATTGGATTGGATATTGAAGGTGATTTG : 2010

Seqid3 : ----- : -

2020 * 2040

Seqid1 : AATGATTTGACCATCACACGGGCAGTCATG : 2040

Seqid3 : ----- : -

* 2060 *

Seqid1 : ATGAATGACTCAGGCAAAGTATCGCTTGCA : 2070

Seqid3 : : 30

2080 * 2100

Seqid1 : GGTCGTCTTGGATTAAATGACGGTATTCAT : 2100

Seqid3 : : 60

* 2120 *

Seqid1 : TGGCAAATTCAAGGGCGATTAGATGAGGTG : 2130

Seqid3 : : 90

2140 * 2160

Seqid1 : GATACCGCCGCTTTTGTTGATGATGACAAT : 2160

Seqid3 : : 120

* 2180 *

Seqid1 : TTAGTTGCGATCATCACAGGCGATCTTGCC : 2190

Seqid3 : : 150

2200 * 2220

Seqid1 : ACATCAGGCAGATATCGTGATGGCGTGATG : 2220

Seqid3 : : 180

* 2240 *

Seqid1 : GATGATATCGCACTCACTTTTGATGGTCAA : 2250

Seqid3 : : 210

2260 * 2280

Seqid1 : GTGATTAACAATTCAATCCCAAATGGTCAT : 2280

Seqid3 : : 240

* 2300 *

Seqid1 : CTAAGTATTGATGCGACAGGTTCTGGCAAT : 2310

Seqid3 : : 270

2320 * 2340

Seqid1 : CGCTTTATGGTCAATCATCTAAGACATGAT : 2340

Seqid3 : : 300

* 2360 *

Seqid1 : GGCACAGCAGGCACGCTGAATGCGACAGGC : 2370

Seqid3 : : 330

2380 * 2400

Seqid1 : TGGGTAGATATTAGCCAAGGTGCTGCATGG : 2400

Seqid3 : : 360

* 2420 *

Seqid1 : CAGCTAGAGGCTGATATGAATTCATTGAAT : 2430

Seqid3 : : 390

	2440	*	2460
Seqid1	: TTGGGTGCATTTATCCAGTCGCTAGATACT	:	2460
Seqid3	:	:	420
	*	2480	*
Seqid1	: GACTTAAACGGTACAATTCAGCTTGCAGGC	:	2490
Seqid3	:	:	450
	2500	*	2520
Seqid1	: AACTGGCAAGAGTCTCATCAAGTCATTGAT	:	2520
Seqid3	:	:	480
	*	2540	*
Seqid1	: ATCAGTAATTTGGATATTACAGGTAGCTAC	:	2550
Seqid3	:	:	510
	2560	*	2580
Seqid1	: AATAATCAGCCACTTATCGCCACAGGCAGC	:	2580
Seqid3	:	:	540

	*	2600	*
Seqid1	:	TTATATGCCAAGCTTTCCTACCTAAAGAT	: 2610
Seqid3	:		: 570
		2620	*
			2640
Seqid1	:	TTGGCTGGGTATATCCAAAGTATCAAGCAA	: 2640
Seqid3	:		: 600
	*	2660	*
Seqid1	:	GCAAGCCGCCACCGACTTCGTCTGATGAT	: 2670
Seqid3	:		: 630
		2680	*
			2700
Seqid1	:	TTATTGGCACTAAAAATCGTATTGATAAC	: 2700
Seqid3	:		: 660
	*	2720	*
Seqid1	:	AACGCACGCCAAACCCAAAATATCGTTCAT	: 2730
Seqid3	:		: 690
		2740	*
			2760
Seqid1	:	AAACTCAATGCTGATAATCTACAAGTCCGC	: 2760

Seqid3 : : 720

* 2780 *

Seqid1 : ATTGGCAATAATCATTGCGCCATATCAGGC : 2790

Seqid3 : : 750

2800 * 2820

Seqid1 : GATGAAAGACAGTTGACCGCCAGTGTGAAT : 2820

Seqid3 : : 780

* 2840 *

Seqid1 : GTGATTGATCTTGGGCAGCTGATTGATACC : 2850

Seqid3 : : 810

2860 * 2880

Seqid1 : GCAAGTGGTGCGATTCAAGGCGGTGTGATT : 2880

Seqid3 : : 840

* 2900 *

Seqid1 : GTCGTCAATGATCATCACGCTTTGCCAACT : 2910

Seqid3 : : 870

	2920	*	2940
Seqid1	: TTATATATTGATGCCAGTGTCTCATCACTC	:	2940
Seqid3	:	:	900
	*	2960	*
Seqid1	: AGATTTGCCAATATCACCATCCAAAACGCC	:	2970
Seqid3	:	:	930
	2980	*	3000
Seqid1	: CAAGCTATCGGTAAAATTGTCAATCTGGGT	:	3000
Seqid3	:	:	960
	*	3020	*
Seqid1	: AATAGCGAAAGCCAGCTGTTGGTTCAGGGT	:	3030
Seqid3	:	:	990
	3040	*	3060
Seqid1	: GATGATATTATCGTGATGGGACGCACCATT	:	3060
Seqid3	:	:	1020
	*	3080	*

```

Seqid1 : AAATCTGCTCGGATGGATTTTAGTGGCACA : 3090
Seqid3 : ..... : 1050

          3100          *          3120
Seqid1 : GAAGCCGATCATATCTTATCAATTTCAACC : 3120
Seqid3 : ..... : 1080

          *          3140          *
Seqid1 : AAAAGCGGTGATATCGAAGCATCTATGCAT : 3150
Seqid3 : ..... : 1110

          3160          *          3180
Seqid1 : ATTGATGGGGCATTTAATCGCAACAATATG : 3180
Seqid3 : ..... : 1140

          *          3200          *
Seqid1 : CGATATCATGGTGTTTTGTCTGATGGTTTT : 3210
Seqid3 : ..... : 1170

          3220          *          3240
Seqid1 : GTCAAGAGTGGATTTGGCAAAATGTCACAG : 3240
Seqid3 : ..... : 1200

```

	*	3260	*
Seqid1	:	CGTCAGCCGACCGAGTTTAGCTATGGATTG	: 3270
Seqid3	:		: 1230
	3280	*	3300
Seqid1	:	GATGATAATAGCTTACAAATTGCAGCGCAT	: 3300
Seqid3	:		: 1260
	*	3320	*
Seqid1	:	TGTTGGCAATCGGCACATATCCAAGGCGAT	: 3330
Seqid3	:		: 1290
	3340	*	3360
Seqid1	:	GGCGTGGGTGCGATTTGTTTACAAGATACA	: 3360
Seqid3	:		: 1320
	*	3380	*
Seqid1	:	TTAAGTTATACACCGCAATCAGGGAATGTC	: 3390
Seqid3	:		: 1350

3400 * 3420
Seqid1 : AATCTTATTATCCAAAATCTTGATACTCAA : 3420
Seqid3 : : 1380

* 3440 *
Seqid1 : GTACTATTGGCGGCTCTACCCAGTGATATT : 3450
Seqid3 : : 1410

3460 * 3480
Seqid1 : CATTGGAAATCCATGCTCAATGGTCGCATT : 3480
Seqid3 : : 1440

* 3500 *
Seqid1 : AAGGCAATGTGGCAGGCAGGTCAATCACCT : 3510
Seqid3 : : 1470

3520 * 3540
Seqid1 : TTGGTTGATGCTGTACTGTATTCTGATGAT : 3540
Seqid3 : : 1500

* 3560 *
Seqid1 : GGTACAATTGGCTTAACTCAAGAAGAGACA : 3570

Seqid3 : : 1530

3580 * 3600

Seqid1 : GGTTATATTGAGATGCCATATCAGCGTGCG : 3600

Seqid3 : : 1560

* 3620 *

Seqid1 : TCTGTGATTGCCAAAAGTGTTGATAATGGC : 3630

Seqid3 : : 1590

3640 * 3660

Seqid1 : TTAAAGGTAAGAACCGATATCTTGGGTACG : 3660

Seqid3 : : 1620

* 3680 *

Seqid1 : GCAGGTCGTGGTTATGCCGATGTAATCATC : 3690

Seqid3 : : 1650

3700 * 3720

Seqid1 : AACCCAAGCAAACAGACAAGCCCATTCT : 3720

Seqid3 : : 1680

	*	3740	*
Seqid1	:	GGTGCGCTTGTTATGAATGATCTTAATTTG	: 3750
Seqid3	:		: 1710
		3760	*
			3780
Seqid1	:	GCGGTATTGCGACCATTTTCCCGAGCATT	: 3780
Seqid3	:		: 1740
	*	3800	*
Seqid1	:	CAAACATTGAGCGGTAAAGTCAGCTTGGCA	: 3810
Seqid3	:		: 1770
		3820	*
			3840
Seqid1	:	GGTGGCTTAGGTGGTACATTATCCAAACCA	: 3840
Seqid3	:		: 1800
	*	3860	*
Seqid1	:	TTGTTTTATGGTAATGCCAATTAAAGAAT	: 3870
Seqid3	:		: 1830
		3880	*
			3900

Seqid1 : GGTGCATTAGCGATTGTTGGCGTACCAATG : 3900

Seqid3 : : 1860

* 3920 *

Seqid1 : CCCATTTCTGATGTTGATGCCACGGTGAAC : 3930

Seqid3 : : 1890

3940 * 3960

Seqid1 : ATACGAGGTACTAAAGCACGCTTAGACGGT : 3960

Seqid3 : : 1920

* 3980 *

Seqid1 : AGATTTATGGGTGGTGAGGGTCATGGTGTG : 3990

Seqid3 : : 1950

4000 * 4020

Seqid1 : CTTTATGGTGAAATGGATTGGGCAGAAGAG : 4020

Seqid3 : : 1980

* 4040 *

Seqid1 : CTGTATGCTCGCCTTGGCGTATTTGGTGAA : 4050

Seqid3 : : 2010

4060 * 4080

Seqid1 : AATCTGACCGTCAGCCAGCCACCGCTTGTG : 4080

Seqid3 : : 2040

* 4100 *

Seqid1 : ACGGCACAAATCAGCCCTGAGCTGGAAGTG : 4110

Seqid3 : : 2070

4120 * 4140

Seqid1 : ATTATCAGACCTTTCCAAAAATTTGTAGAT : 4140

Seqid3 : : 2100

* 4160 *

Seqid1 : GTTCAAGGCGTTGTTAGTATCCCATCAGCA : 4170

Seqid3 : : 2130

4180 * 4200

Seqid1 : ACCATTCGTCCGCCAGAGGCAACCGCAGAT : 4200

Seqid3 : : 2160

* 4220 *

Seqid1 : ATCGTAACCGAATCCCCAGATGTATCAGTC : 4230

Seqid3 : : 2190

4240 * 4260

Seqid1 : ATTGACCGCCGTATCACAGGTAATATTGAC : 4260

Seqid3 : : 2220

* 4280 *

Seqid1 : CAGATTTTAAGTAGAGCGAAACCTTGGGAT : 4290

Seqid3 : : 2250

4300 * 4320

Seqid1 : ATTAATGCAAATATCGGTGTTGATTTGGGT : 4320

Seqid3 : : 2280

* 4340 *

Seqid1 : AATGAGATTGAATTTTCGTGGATTTGGGGCG : 4350

Seqid3 : : 2310

4360 * 4380

Seqid1 : GTATTACCATTGGCAGGCGCGATACATTTG : 4380

Seqid3 : : 2340

* 4400 *

Seqid1 : ACTCAGTCAGGACAAGGTGCAATGCAAGCT : 4410

Seqid3 : : 2370

4420 * 4440

Seqid1 : CTTGGCGTGGTTCAGGTTTCTAAACGCACA : 4440

Seqid3 : : 2400

* 4460 *

Seqid1 : AAAATCGATGTCATCGGTCAAATTTGGAT : 4470

Seqid3 : : 2430

4480 * 4500

Seqid1 : TTAAATTATGCCCAAATTCGTTTTGATGGC : 4500

Seqid3 : : 2460

* 4520 *

Seqid1 : GACATGCTTAATCCACGCTTATCTATTGAG : 4530

Seqid3 : : 2490

	4540	*	4560
Seqid1	: GGTGAAAAACAAATTGAAGGGCAAACGGTG	:	4560
Seqid3	:	:	2520
	*	4580	*
Seqid1	: GCGGTTTCGCATTCGTGGAACGGCATCAAAT	:	4590
Seqid3	:	:	2550
	4600	*	4620
Seqid1	: CCAAACATTACCGTATTTAATGATGCAGGT	:	4620
Seqid3	:	:	2580
	*	4640	*
Seqid1	: TTGGATGAATACCAAGCAATGAATGCACTG	:	4650
Seqid3	:	:	2610
	4660	*	4680
Seqid1	: GTGACAGGACGCATTAGCGAATCAAGTGAT	:	4680
Seqid3	:	:	2640

	*	4700	*
Seqid1	:	TTGGGTATCACAGAGCAGGGTTTTCGCTCA	: 4710
Seqid3	:		: 2670
	4720	*	4740
Seqid1	:	CAAGTAACCAATCACCTTGCTGCTGCTGGC	: 4740
Seqid3	:		: 2700
	*	4760	*
Seqid1	:	TTAAGCTTAGGTTTGTCTCAGGAACCAGAGAT	: 4770
Seqid3	:		: 2730
	4780	*	4800
Seqid1	:	TTAACCAACCAAATTGGTCACGCATTTGGT	: 4800
Seqid3	:		: 2760
	*	4820	*
Seqid1	:	TTTCAGAGTTTGACCATTGATGCCTCGGGC	: 4830
Seqid3	:		: 2790
	4840	*	4860
Seqid1	:	AGCTCAGATGATACCAATGTTAATGTTACA	: 4860

Seqid3 : : 2820

* 4880 *

Seqid1 : GGCTATATTACACCAGATTTGTATTTGCGT : 4890

Seqid3 : : 2850

4900 * 4920

Seqid1 : TATGGCGTGGGTGTGTTTAATGCTGAATCA : 4920

Seqid3 : : 2880

* 4940 *

Seqid1 : ACCTTATCCATGCGTTATCAATTGACACGC : 4950

Seqid3 : : 2910

4960 * 4980

Seqid1 : CGTGTTTATATTGAAGCAACCAGTGCCGCT : 4980

Seqid3 : : 2940

* 5000 *

Seqid1 : GAAAATATGGTTGATGTGATTTATCGTTGG : 5010

Seqid3 : : 2970

Seqid1	:	AAGTTT'TAG	:	5019
Seqid3	:		:	2979

图 2: BASB132多肽序列的对比.

与序列编号: 2相同处用圆点表示. 间隔处用短线表示.

	*	20	*
Seqid2	:	MTHQDNSKNSSSAENGVSAGVAAPT	30
Seqid4	:	-----	-
	40	*	60
Seqid2	:	PPNPPKPQWLIYLIRLIVGAI	60
Seqid4	:	-----	-
	*	80	*
Seqid2	:	LFAMTNSESGSKFLIEKIALETG	90
Seqid4	:	-----	-
	100	*	120
Seqid2	:	GSIRHGVWVQDVKIAQSE	120
Seqid4	:	-----	-
	*	140	*
Seqid2	:	LGWRALFARQVHLVNP	150
Seqid4	:	-----	-
	160	*	180

Seqid2 : GEPFDYATINLPVTLKLENAKVNEIIYDQV : 180

Seqid4 : ----- :

* 200 *

Seqid2 : DSEPVLHHIAFDHASWADSTVKIDNAMLS : 210

Seqid4 : ----- : -

220 * 240

Seqid2 : YGDDINISHATGGIDLTGHYPLSLADVHI : 240

Seqid4 : ----- : -

* 260 *

Seqid2 : LALDDAYFDTLSVKAGGSLKRTVGTLTGKY : 270

Seqid4 : ----- : -

280 * 300

Seqid2 : NQHHVTGSFIAQGLDKNSPFSARLDFDEV : 300

Seqid4 : ----- : -

* 320 *

Seqid2 : LPYADSQNILLKNGSIIADGVISNIELRIN : 330

Seqid4 : ----- : -

	340	*	360
Seqid2	: TELSAKDIPDGHYHGRGIIRGSTMQIPYLQ	:	360
Seqid4	: -----	:	-
	*	380	*
Seqid2	: ADTANGTLVATGDMTWEDGYELDATITADG	:	390
Seqid4	: -----	:	-
	400	*	420
Seqid2	: YRIREEMPSDYHEYRAYLPKVLTGSLGVKY	:	420
Seqid4	: -----	:	-
	*	440	*
Seqid2	: TLLDKASHDTRFEFDLNQKDGERIQATLAQ	:	450
Seqid4	: -----	:	-
	460	*	480
Seqid2	: NQQSDHEPWRIDATWANLIRHDIPQIGEIH	:	480
Seqid4	: -----	:	-
	*	500	*
Seqid2	: SRSGQASVRLENGHTYINASADIVKLNAV	:	510

Seqid4 : ----- : -

520 * 540

Seqid2 : SGSYHVQANIEQNQHLLTDFNYQGVMGEL : 540

Seqid4 : ----- : -

* 560 *

Seqid2 : TGTGRVDFATAQRPLEWQLDVTANPVKPNA : 570

Seqid4 : ----- : -

580 * 600

Seqid2 : YFQTPNQTPFEQISGRRIASGRLREIDGVS : 600

Seqid4 : ----- : -

* 620 *

Seqid2 : IHDIKVDDSDLTALLNDGKQVHLMGQGVSK : 630

Seqid4 : ----- : -

640 * 660

Seqid2 : IRLQDGQISHLKANFDGKLAQDILPQVADS : 660

Seqid4 : ----- : -

	*	680	*
Seqid2	:	SIGLDIEGDLNDLTITRAVMMND SGKVSLA	: 690
Seqid4	:	-----	: 10
		700	*
Seqid2	:	GRLGLNDGIHWQIQGR LDEVDTAA FVDDDN	: 720
Seqid4	:		: 40
		740	*
Seqid2	:	LVAIITGDLATSGRYRDGVMDDIALTFDGQ	: 750
Seqid4	:		: 70
		760	*
Seqid2	:	VINNSIPNGHLSIDATGSGNR FMVNHLRHD	: 780
Seqid4	:		: 100
		800	*
Seqid2	:	GTAGTLNATGWVDISQGA AWQLEADMNSLN	: 810
Seqid4	:		: 130
		820	*
			840

Seqid2 : LGAFIQSLDLDLNGTIQLAGNWQESHQVID : 840

Seqid4 : : 160

* 860 *

Seqid2 : ISNLDITGSYNNQPLIATGSLYAKLSLPKD : 870

Seqid4 : : 190

880 * 900

Seqid2 : LAGYIQSIKQASRPPTSSDDLALKNRIDN : 900

Seqid4 : : 220

* 920 *

Seqid2 : NARQTQNIIVHKLNADNLQVRIGNNHLAISG : 930

Seqid4 : : 250

940 * 960

Seqid2 : DERQLTASVNVIDLGQLIDTASGAIQGGVI : 960

Seqid4 : : 280

* 980 *

Seqid2 : VVNDHHALPTLYIDASVSSLRFANITIONA : 990

```

Seqid4 : ..... : 310

          1000          *          1020
Seqid2 : QAIGKIVNLGNSESQLLVQGDDIIVMGRTI : 1020
Seqid4 : ..... : 340

          *          1040          *
Seqid2 : KSARMDFSGTEADHILSISTKSGDIEASMH : 1050
Seqid4 : ..... : 370

          1060          *          1080
Seqid2 : IDGAFNRNNMRYHGVLSDGFEVKSGFGKMSQ : 1080
Seqid4 : ..... : 400

          *          1100          *
Seqid2 : RQPTEFSYGLDDNSLQIAAHCWQSAHIQGD : 1110
Seqid4 : ..... : 430

          1120          *          1140
Seqid2 : GVGAIQLQDTLSYTPQSGNVNLIQNLDTO : 1140
Seqid4 : ..... : 460

```

	*		*
Seqid2	: VLLAALPSDIHWKSMNLNGRIKAMWQAGQSP	:	1170
Seqid4	:	:	490
	1180	*	1200
Seqid2	: LVDAVLYSDDGTIGLTQEETGYIEMPYQRA	:	1200
Seqid4	:	:	520
	*	1220	*
Seqid2	: SVIAKSVDNGLKVRTDILGTAGRGYADVII	:	1230
Seqid4	:	:	550
	1240	*	1260
Seqid2	: NPKQTDKPISGALVMNDLNLAVLRPFFPSI	:	1260
Seqid4	:	:	580
	*	1280	*
Seqid2	: QTLSGKVSLAGGLGGTLSKPLFYGNANLKN	:	1290
Seqid4	:	:	610
	1300	*	1320

Seqid2 : GALAIVGVMPISDV DATVNIRGTKARLDG : 1320

Seqid4 : : 640

* 1340 *

Seqid2 : RFMGGEHGVLYGEMDWAEE LYARLG VFGE : 1350

Seqid4 : : 670

1360 * 1380

Seqid2 : NLTVSQPPLVTAQISPELEVIIRPFQKFVD : 1380

Seqid4 : : 700

* 1400 *

Seqid2 : VQGVVSIPSATIRPPEATADIVTESPDVSV : 1410

Seqid4 : : 730

1420 * 1440

Seqid2 : IDRRITGNIDQILSRAKPWDINANIGVDLG : 1440

Seqid4 : : 760

* 1460 *

Seqid2 : NEIEFRGFGAVLPLAGAIHLTQSGQGAMQA : 1470

Seqid4 : : 790

	1480	*	1500
Seqid2	: LGVVQVSKRTKIDVIGQNLDLNYAQIRFDG	:	1500
Seqid4	:	:	820
		*	1520
Seqid2	: DMLNPRLSIEGEKQIEGQTVGVRIRGTASN	:	1530
Seqid4	:	:	850
	1540	*	1560
Seqid2	: PNITVFNDAGLDEYQAMNALVTGRISSESD	:	1560
Seqid4	:	:	880
		*	1580
Seqid2	: LGITEQGFRSQVTNHLAAAGLSLGLSGTRD	:	1590
Seqid4	:	:	910
	1600	*	1620
Seqid2	: LTNQIGHAFGFQSLTIDASGSSDDTNVNVT	:	1620
Seqid4	:	:	940

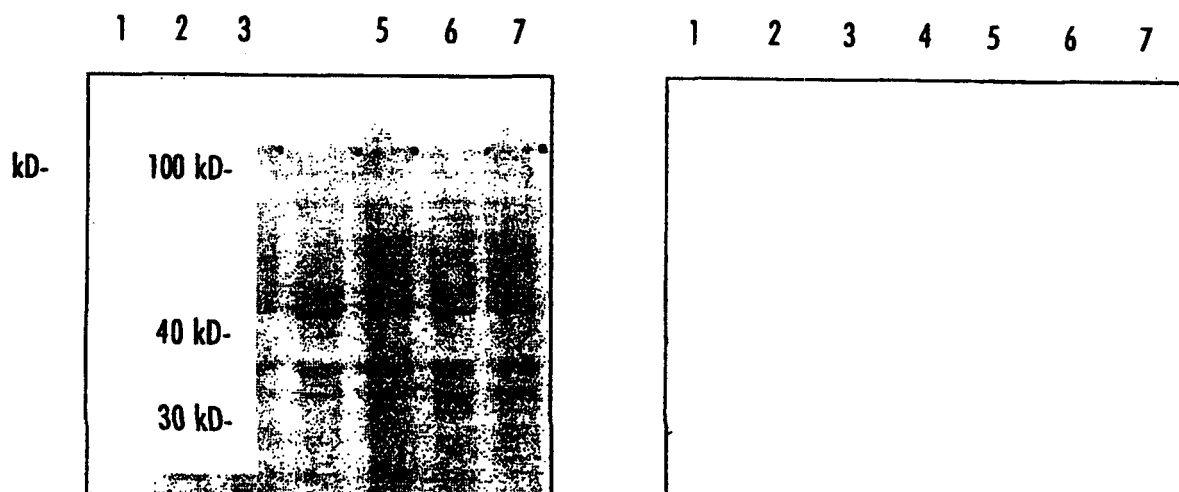
```

                                *           1640           *
Seqid2  :  GYITPDLYLRYGVGVFNAESTLSMRYQLTR  :  1650
Seqid4  :  ..... :  970

                                1660           *
Seqid2   :    RVYIEATSAAENMVDVIYRWKF      :  1672
Seqid4   :    ..... :  992
```

图 3A: BASB132的小规模表达分析: 考马斯染色的SDS-聚丙烯酰胺凝胶.

图 3B: BASB132的小规模表达分析: 蛋白印迹 (抗-His抗体)



- 1: MW (kD)
2: clone 14, t=0h (no induction)
3: clone 14, t=4h (after induction)
4: clone 17, t=0h
5: clone 17, t=4h
6: clone 21, t=0h
7: clone 21, t=4h

图 3A

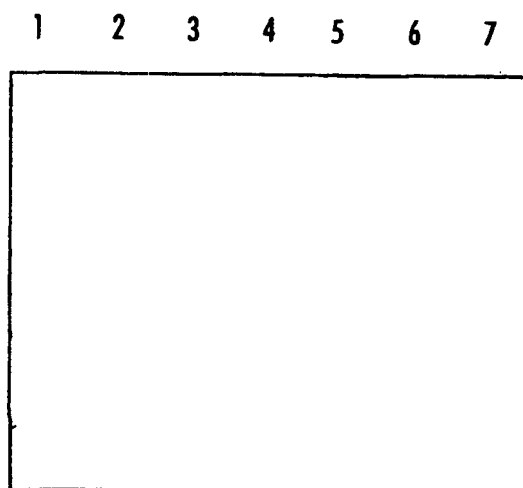


图 3B

图 4A: 纯化BASB132的经考马斯染色的聚丙烯酰胺凝胶

图 4B: 纯化BASB132的蛋白印迹(抗-His抗体)

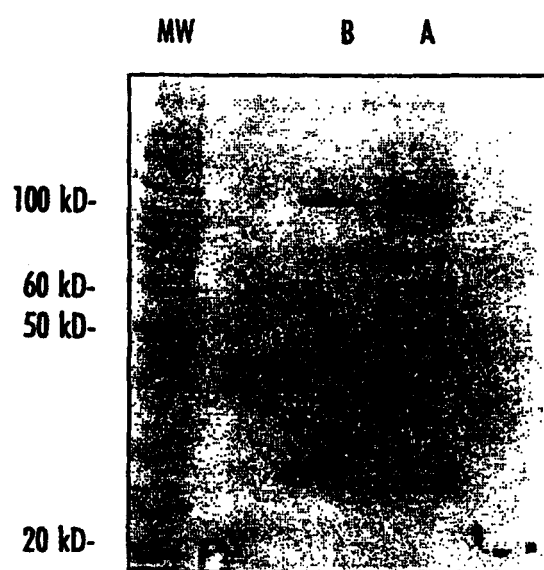


图 4A

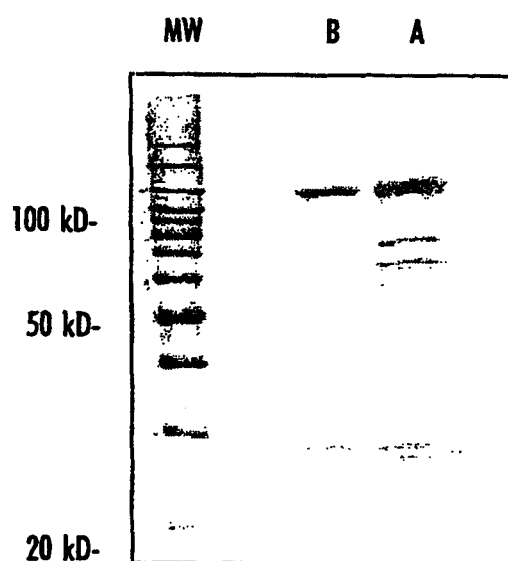


图 4B

专利名称(译)	新化合物		
公开(公告)号	CN1402785A	公开(公告)日	2003-03-12
申请号	CN00816501.7	申请日	2000-09-26
[标]申请(专利权)人(译)	史密丝克莱恩比彻姆生物有限公司		
申请(专利权)人(译)	史密丝克莱恩比彻姆生物有限公司		
当前申请(专利权)人(译)	史密丝克莱恩比彻姆生物有限公司		
[标]发明人	J·通纳德		
发明人	J·通纳德		
IPC分类号	C12N15/09 A61K39/00 A61K39/04 A61K39/395 A61K48/00 A61P31/04 C07K14/195 C07K14/21 C07K16/12 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12P21/02 C12N15/31 C12N15/63 G01N33/53 G01N33/68 A61K39/02		
CPC分类号	C07K2319/00 A61K2039/53 C07K14/212 A61K2039/505 A61K39/00		
代理人(译)	姜建成		
优先权	1999023156 1999-09-30 GB		
外部链接	Espacenet SIPO		

摘要(译)

本发明提供BASB132多肽和编码BASB132多肽的多核苷酸，以及由重组技术产生这些多肽的方法。另外还提供诊断、预防及治疗性应用。

BASB132多核苷酸序列的对比。

与序列编号-1相同处用圆点表示，不同处用短横线表示。

```

      20                                     *
Seqid1 : ATGACTCATCAGGAZAAATTCAAAAAACTCG : 30
Seqid3 : ----- : -

      40                                     *
Seqid1 : TCATCGGCAGAAAACTGGGTTTCTGCTGGC : 50
Seqid3 : ----- : -

      60                                     *
Seqid1 : GTTCCTGCAACCGACTAAAACTCAGCCCAAA : 70
Seqid3 : ----- : -

      80                                     *
Seqid1 : CCACCAAAACCCACCCCAAAACCAATGCTTA : 90
Seqid3 : ----- : -

     100                                     *
Seqid1 : ATCTATTTAATTCGCTTAATGTCGGAGCA : 110
Seqid3 : ----- : -

     120                                     *
Seqid1 : ----- : 130
Seqid3 : ----- : -

     140                                     *
Seqid1 : ----- : 150
Seqid3 : ----- : -

     160                                     *
Seqid1 : ----- : 170
Seqid3 : ----- : -

```