



(12) 发明专利

(10) 授权公告号 CN 101932595 B

(45) 授权公告日 2016. 02. 10

(21) 申请号 200880109156. 6

(22) 申请日 2008. 09. 25

(30) 优先权数据

60/975, 104 2007. 09. 25 US

60/989, 840 2007. 11. 22 US

60/989, 841 2007. 11. 22 US

(85) PCT国际申请进入国家阶段日

2010. 03. 23

(86) PCT国际申请的申请数据

PCT/NZ2008/000248 2008. 09. 25

(87) PCT国际申请的公布数据

W02009/041831 EN 2009. 04. 02

(73) 专利权人 田园温室气体研究有限公司

地址 新西兰惠灵顿

(72) 发明人 E·H·阿特曼 G·T·爱特伍德

S·C·利伊 W·J·凯利

R·S·罗尼姆斯 D·李 Z·孔

L·R·斯考菲尔德 D·德

C·M·图提尔 C·桑 C·D·穆恩

H·J·派特鲁斯

(74) 专利代理机构 上海专利商标事务所有限公

司 31100

代理人 韦东 张静

(51) Int. Cl.

C07K 14/01(2006. 01)

C07H 21/04(2006. 01)

C12N 7/01(2006. 01)

(56) 对比文件

Attwood, GT 等. Analysis of the
Methanobrevibacter ruminantium draft
genome: understanding methanogen biology
to inhibit their action in the rumen.
《Australian Journal of Experimental
Agriculture》. 2008, 第 48 卷 (第 2 期), 图 2b,
摘要.

审查员 田颖

权利要求书2页 说明书35页

序列表90页 附图32页

(54) 发明名称

噬菌体 ϕ_{mru} 多核苷酸和多肽及其应用

(57) 摘要

本发明包括含有噬菌体诱导、噬菌体颗粒和噬菌体基因组的噬菌体 ϕ_{mru} 还包括噬菌体多肽、编码这些多肽的多核苷酸、包含这些多核苷酸的表达载体和包含这些载体的宿主细胞。本发明还包括使用所公开的噬菌体、多肽、多核苷酸、表达载体或宿主细胞来检测、靶向、通透并抑制微生物细胞, 尤其是产甲烷菌细胞的组合物和方法。

1. 一种分离的多肽,其氨基酸序列组成为 SEQ ID NO:63,所述多肽是产甲烷菌细胞的抑制剂和 / 或检测噬菌体 ϕ mru 的生物标记物。

2. 一种分离的多核苷酸,其编码权利要求 1 的多肽。

3. 一种分离的多核苷酸,其核苷酸序列组成为 SEQ ID NO:136,其编码产甲烷菌细胞的抑制剂和 / 或检测噬菌体 ϕ mru 的生物标记物。

4. 一种载体,其包含编码权利要求 1 所述多肽的多核苷酸。

5. 一种包含权利要求 2 或 3 所述多核苷酸的载体。

6. 一种包含权利要求 4 或 5 所述载体的宿主细胞。

7. 一种经遗传修饰来编码权利要求 1 所述多肽的宿主细胞。

8. 一种经遗传修饰包含权利要求 2 或 3 所述多核苷酸的宿主细胞。

9. 如权利要求 6-8 中任一项所述的宿主细胞,其特征在于,所述细胞是原核细胞。

10. 如权利要求 9 所述的宿主细胞,其特征在于,所述细胞是大肠杆菌。

11. 如权利要求 6-8 中任一项所述的宿主细胞,其特征在于,所述细胞是产甲烷菌。

12. 如权利要求 11 所述的宿主细胞,其特征在于,所述细胞是反刍甲烷短杆菌。

13. 一种偶联物分子,其包含权利要求 1 所述多肽。

14. 如权利要求 13 所述的偶联物分子,其特征在于,所述分子还包含抗甲烷生成化合物、信号序列、抗体或抗体片段、肽核酸、抗微生物肽或抗生素。

15. 一种包含权利要求 1 所述多肽的融合分子。

16. 如权利要求 15 所述的融合分子,其特征在于,所述分子还包含信号序列、抗体或抗体片段、肽核酸或抗微生物肽。

17. 一种与权利要求 1 所述多肽结合的抗体或抗体片段,其为检测噬菌体 ϕ mru 的生物标记物。

18. 如权利要求 17 所述的抗体或抗体片段,其特征在于,所述抗体是多克隆的或所述抗体片段来自多克隆抗体。

19. 如权利要求 17 所述的抗体或抗体片段,其特征在于,所述抗体是单克隆的或所述抗体片段来自单克隆抗体。

20. 一种分离的 ϕ mru 噬菌体,其包含权利要求 1 所述的多肽。

21. 一种分离的 ϕ mru 噬菌体,其包含权利要求 2 或 3 所述的多核苷酸。

22. 一种包含权利要求 1 所述多肽的药物组合物。

23. 一种包含权利要求 2 或 3 所述多核苷酸的药物组合物。

24. 一种包含权利要求 20 或 21 所述噬菌体的药物组合物。

25. 一种抑制产甲烷菌细胞的方法,所述方法包括 :a) 产生或分离权利要求 1 所述的多肽 ;和 b) 将所述细胞与权利要求 1 所述的多肽相接触。

26. 如权利要求 25 所述的方法,其特征在于,所述细胞是反刍甲烷短杆菌。

27. 如权利要求 26 所述的方法,其特征在于,所述细胞是反刍甲烷短杆菌菌株 M1^T即 DSM1093。

28. 一种抑制产甲烷菌细胞的方法,所述方法包括 :a) 产生或分离权利要求 13 或 14 所述的偶联物分子 ;和 b) 将所述细胞与权利要求 13 或 14 所述的偶联物分子相接触。

29. 如权利要求 28 所述的方法,其特征在于,所述细胞是反刍甲烷短杆菌。

30. 如权利要求 29 所述的方法,其特征在于,所述细胞是反刍甲烷短杆菌菌株 M1^T即 DSM1093。

31. 一种抑制产甲烷菌细胞的方法,所述方法包括:a) 产生或分离权利要求 15 或 16 所述的融合分子;和 b) 将所述细胞与权利要求 15 或 16 所述的融合分子相接触。

32. 如权利要求 31 所述的方法,其特征在于,所述细胞是反刍甲烷短杆菌。

33. 如权利要求 32 所述的方法,其特征在于,所述细胞是反刍甲烷短杆菌菌株 M1^T即 DSM1093。

34. 一种抑制产甲烷菌细胞的方法,所述方法包括:a) 产生或分离权利要求 20 或 21 所述的噬菌体;和 b) 将所述细胞与权利要求 20 或 21 所述的噬菌体相接触。

35. 如权利要求 34 所述的方法,其特征在于,所述细胞是反刍甲烷短杆菌。

36. 如权利要求 35 所述的方法,其特征在于,所述细胞是反刍甲烷短杆菌菌株 M1^T即 DSM1093。

噬菌体 ϕ mru 多核苷酸和多肽及其应用

[0001] 相关申请

[0002] 本申请要求 2007 年 9 月 25 日提交的美国申请号 60/975,104, 2007 年 9 月 22 日提交的美国申请号 60/989,840 和 2007 年 11 月 22 日提交的美国申请号 60/989,841 的优先权, 将上述所有的内容整体纳入本文作参考。

技术领域

[0003] 本发明涉及将抑制分子递送至微生物细胞, 具体说是产甲烷菌细胞的组合物和方法。具体说, 本发明涉及新鉴定的噬菌体 ϕ mru, 包括噬菌体诱导、噬菌体颗粒、噬菌体基因组、噬菌体多肽以及编码这些多肽的多核苷酸。本发明还涉及用于产生这些多肽的表达载体和宿主细胞。本发明还涉及使用所公开的噬菌体、多肽、多核苷酸、表达载体和宿主细胞来探测、靶向并抑制微生物细胞, 尤其是产甲烷菌细胞的方法。

[0004] 发明背景

[0005] 在新西兰, 农业活动是温室气体排放的主要来源。因此, 减少农业温室气体排放对于新西兰完成京都议定书的义务是重要的。议定书要求在首次承诺期 (2008-2012) 结束前温室气体减排至 1990 年水平。为此, 农业部门和新西兰政府成立了田园温室气体研究联合会 (PGGRC) 寻求降低新西兰的农业温室气体排放的方法。

[0006] PGGRC 活动的一个重要部分是研究减少来自新西兰放牧反刍动物的甲烷排放量。出于两个原因, 减少反刍动物甲烷排放量具有商业利益。首先, 不履行京都议定书承诺将迫使政府购买碳排放额度。目前预计该项将花费三亿五千万美金。第二, 甲烷的产生导致瘤胃中产生的总能量损耗 8-12%。可利用这种能量来提高反刍动物的生产性能。

[0007] 甲烷在瘤胃中由称为产甲烷菌的微生物产生, 这种微生物属于古细菌 (Archaea) 界广古菌门 (euryarchaeota) 的一部分。多数产甲烷菌依靠 CO_2 和 H_2 作为它们唯一的能量来源, 但是一些能使用乙酸或甲基化合物生长。瘤胃中中存在多种不同属的产甲烷古菌, 但是甲烷短杆菌属中的种, 尤其是反刍甲烷短杆菌 (*M. ruminantium*) 和史氏甲烷短杆菌 (*M. smithii*) 被认为是新西兰反刍动物中主要的产甲烷菌, 反刍甲烷短杆菌目前是 PGGRC 基金资助基因组测序项目的研究对象。该项目第一次对于瘤位产甲烷菌的基因组进行测序, 旨在对甲烷短杆菌 (*Methanobrevibacter*) 的生物学建立更好的理解, 以发现抑制甲烷产生的靶点。

[0008] 减少瘤胃中甲烷产生需要抑制产甲烷菌或使其甲烷生成通路失活。一种抑制甲烷产生的方法是将特异性抑制分子递送至产甲烷菌细胞中。可通过 (例如) 使用特异性靶向产甲烷菌的试剂如噬菌体达到这种目的。已对数种非瘤胃产甲烷菌的噬菌体进行表征, 但是没有报道关于能够感染或裂解瘤胃产甲烷菌的噬菌体。因此, 鉴定能够感染产甲烷菌细胞和 / 或递送抑制剂的噬菌体极具优势。

发明内容

[0009] 本发明涉及一种分离的噬菌体 ϕ mru, 包括本文详述的以整体或部分形式生产的

噬菌体颗粒和 / 或噬菌体基因组, 以及分离的噬菌体多核苷酸和多肽。

[0010] 本发明还涉及一种分离的多肽, 其包含至少一个选自 SEQ ID NO :1-69 的噬菌体氨基酸序列。在一个具体方面, 所述多肽包含选自 SEQ ID NO :2-5 和 62-68 的氨基酸序列。另一方面, 所述多肽包含 SEQ ID NO :63 的氨基酸序列。在另一方面, 所述多肽为片段, 例如包含至少一个延伸自 SEQ ID NO :63 中残基 32-186 的氨基酸序列。

[0011] 本发明还涉及一种分离的多核苷酸, 其包含至少一种噬菌体多肽的编码序列。在一个方面, 所述多核苷酸包含至少一个选自 SEQ ID NO :1-69 的氨基酸序列的编码序列。在一个具体方面, 所述多核苷酸包含选自 SEQ ID NO :2-5 和 62-68 的序列的编码序列。在另一方面, 所述多核苷酸包含 SEQ ID NO :63 的编码序列。在另一方面, 所述多核苷酸包含某编码序列的片段, 该编码序列编码, 例如, 至少一个延伸自 SEQ ID NO :63 中残基 32-186 的氨基酸序列。

[0012] 在另一方面, 本发明涉及一种分离的多核苷酸, 其包含选自 SEQ ID NO :74-142 的噬菌体核酸序列。在特定方面, 所述多核苷酸包含选自 SEQ ID NO :75-78 和 135-141, 或更具具体是 SEQ ID NO :136 的核酸序列。在另一方面, 所述多核苷酸是包含例如延伸自 SEQ ID NO :136 中核苷酸 94-558 的核酸序列的片段或寡核苷酸。此外, 本发明包括一种分离的多核苷酸或其片段, 该多核苷酸或其片段能与 SEQ ID NO :74-142 中任一核酸序列杂交。本发明还包括一种分离的多核苷酸, 该多核苷酸含有任一所述核酸序列的互补物、反向互补物、反向序列或其片段。

[0013] 本发明涉及一种表达载体, 其包含含有至少一种噬菌体多肽的编码序列的多核苷酸。在一个方面, 所述表达载体包含至少一个选自 SEQ ID NO :1-69 的氨基酸序列的编码序列。在一个特定方面, 所述表达载体包含 SEQ ID NO :2-5 和 62-68 中至少一个氨基酸序列的编码序列。在另一方面, 所述表达载体包含至少一个 SEQ ID NO :63 的氨基酸序列的编码序列。在另一方面, 所述表达载体包含至少一个延伸自 SEQ ID NO :63 中残基 32-186 的氨基酸序列的编码序列。

[0014] 在一个具体方面, 本发明涉及以完整或部分形式产生如本文详述的噬菌体 ϕ mrU 的表达载体。具体说, 所述表达载体可产生噬菌体颗粒、噬菌体基因组或修饰的噬菌体, 包括其任何变化形式、衍生物、变体或片段。

[0015] 本发明还涉及包含至少一种表达载体的宿主细胞, 例如微生物宿主细胞。

[0016] 本发明尤其涉及本文公开的多肽或多核苷酸的抗体。在某些方面, 所述抗体针对至少一个选自 SEQ ID NO :1-69 的多肽序列或其片段。在另一方面, 所述抗体针对选自 SEQ ID NO :74-142 的多核苷酸的至少一个片段或其互补或修饰序列。在另一方面, 所述抗体包含与至少一种细胞抑制剂, 例如本文详述的抗-甲烷生成化合物 (如, 溴代乙磺酸)、抗体和抗体片段、裂解酶、肽核酸、抗微生物肽和其他抗生素产生的一种或多种融合物或偶联物。

[0017] 本发明还涉及例如 SEQ ID NO :1-69 中至少一项的修饰的噬菌体多肽, 其包括如本文所述的生物学活性变化形式、片段、变体和衍生物。本发明还涉及例如针对 SEQ ID NO :1-69 中至少一项的修饰的抗体, 其包括如本文所述的生物学活性变化形式、片段、变体和衍生物。本发明还涉及编码这些修饰多肽的多核苷酸, 以及所公开多核苷酸的变化形式、片段、变体和衍生物, 包含这些多核苷酸的表达载体以及包含这些载体的宿主细胞。在具体方面, 本发明的组合物和方法使用这些修饰的多核苷酸或多肽, 或对应的表达载体或宿主细

胞。

[0018] 此外,本发明涉及噬菌体多肽,如SEQ ID NO:1-69中至少一种或其修饰序列,包括与至少一种细胞抑制剂,如本文详述的抗-甲烷生成化合物(如,溴代乙磺酸)、抗体和抗体片段、裂解酶、肽核酸、抗微生物肽和其他抗生素形成的融合物或偶联物。

[0019] 本发明涉及一种组合物,该组合物包含分离多肽如SEQ ID NO:1-69中至少一种或其修饰序列。本发明还涉及一种组合物,该组合物包含针对,例如SEQ ID NO:1-69中至少一种或其修饰序列的抗体。本发明还涉及一种组合物,该组合物包含分离的多核苷酸,如SEQ ID NO:74-142中至少一种或其互补或修饰序列。本发明还涉及一种组合物,该组合物包含本发明表达载体或本发明含有表达载体的宿主细胞。所述组合物可包含本文所述生物学活性变化形式、片段、变体和衍生物中的任何一种。所述组合物可包含至少一种细胞抑制剂(如,作为融合物或偶联物)并可配制成,例如,药物组合物或食品补充剂,具体是反刍动物饲料组分。

[0020] 本发明还涉及作为按照所公开方法靶向和/或抑制微生物细胞,特别是产甲烷菌细胞的试剂盒的一部分的本发明组合物。该试剂盒包括:a)至少一种本文所列的组合物;以及b)任选地包括,用其(例如)靶向细胞或抑制产甲烷菌或其他微生物的细胞生长或复制的使用说明书。

[0021] 本发明涉及一种产生噬菌体的方法,所述方法包括:a)在适合产生噬菌体条件下培养表达载体或含有表达载体的宿主细胞,所述表达载体包含噬菌体基因组的至少一部分;以及b)从培养物中回收该噬菌体。在特定方面,该噬菌体包含至少一个选自SEQ ID NO:1-69或其修饰序列的多肽。在另一些方面,该噬菌体包含至少一个选自SEQ ID NO:74-142或其修饰序列的多核苷酸。

[0022] 本发明还涉及一种产生噬菌体多肽的方法,所述方法包括:a)在适合多肽表达条件下培养表达载体或含有表达载体的宿主细胞,所述表达载体包含至少一种噬菌体多肽的编码序列的至少一部分;以及b)从培养物中回收该多肽。在特定方面,所述多肽包含至少一个选自SEQ ID NO:1-69或其修饰序列的氨基酸序列。

[0023] 此外,本发明还涉及一种产生噬菌体多肽的方法,该多肽如SEQ ID NO:1-69中至少一种,包括与至少一种细胞抑制剂,如本文详述的抗-甲烷生成化合物(如,溴代乙磺酸)、抗体和抗体片段、裂解酶、肽核酸、抗微生物肽和其他抗生素形成的融合物或偶联物。此类方法包括:a)在适合多肽表达的条件下培养表达载体或包含表达载体的宿主细胞,所述表达载体包含至少一种噬菌体多肽的编码序列;b)形成噬菌体融合物或偶联物(如,通过表达融合序列或与细胞抑制剂化学偶联);以及c)回收该融合物或偶联物。在特定方面,所述多肽包含至少一个选自SEQ ID NO:1-69或其修饰序列的氨基酸序列。

[0024] 此外,本发明涉及一种抑制微生物细胞(如抑制其生长或复制),具体是产甲烷菌细胞的方法,该方法包括:a)任选产生或分离至少一种噬菌体多肽;以及b)将所述细胞与所述噬菌体多肽接触。在一个特定方面,所述多肽包含至少一个选自SEQ ID NO:1-69或其修饰序列的氨基酸序列。

[0025] 此外,本发明还包括一种抑制微生物细胞(如抑制其生长或复制),具体是产甲烷菌细胞的方法,该方法包括:a)任选产生或分离至少一种噬菌体多肽,其还包含至少一种细胞抑制剂;以及b)将所述细胞与所述噬菌体多肽接触。在一个特定方面,所述多肽包含

至少一个选自 SEQ ID NO :1-69 或其修饰序列的氨基酸序列。

[0026] 本发明还涉及一种检测和 / 或测定噬菌体水平或对应的噬菌体多肽或多核苷酸水平的方法,该方法包括:1) 将来自对象的样品与抗体接触,所述抗体针对噬菌体多肽(如,SEQ ID NO :1-69 中至少一种或其修饰序列)或对应的多核苷酸;和 2) 检测与样品中多肽或多核苷酸形成的抗体复合物的存在或水平。此类方法还可用于检测和 / 或测定微生物细胞,具体是产甲烷菌细胞的水平。

[0027] 本发明还涉及一种检测和 / 或测定噬菌体水平或对应的噬菌体多核苷酸(如噬菌体编码序列)水平的方法,该方法包括:1) 将来自对象的样品与互补多核苷酸(如与 SEQ ID NO :74-142 中任一项或其修饰序列互补的序列)接触;以及 2) 测定与样品中噬菌体多核苷酸形成的杂交复合物的存在或水平。此类方法还可用于检测和 / 或测定微生物细胞,具体是产甲烷菌细胞的水平。

[0028] 在特定方面,本发明的方法使用体内或体外表达组件。在其它方面,本发明使用重组、合成或半合成方法产生的多肽,或内源性方法产生的多肽。

[0029] 本发明的其它方面和实施方式如下文所述。

[0030] 附图简要说明

[0031] 用本发明特定实施方式和附图作为参考对本发明进行描述。

[0032] 图 1A-1B. 反刍甲烷短杆菌原噬菌体 ϕ mru,显示推定整合位点序列 attL 和 attR(图 1A)和预测的噬菌体功能模块和基因结构(图 1B)。

[0033] 图 1C :用灭菌空气(氧压)进行噬菌体诱导。

[0034] 图 1D :用丝裂霉素 C 对噬菌体进行初始诱导。

[0035] 图 1E :诱导(氧气刺激)和未诱导的反刍甲烷短杆菌的 PCR 扩增产物的琼脂糖凝胶电泳。第 2 和第 4 泳道:英骏公司(Invitrogen)的 1kb DNA 梯标记物。第 1 和第 3 泳道分别代表使用引物对 R1F-L2R 对分离自诱导和未诱导的反刍甲烷短杆菌培养物的 DNA 进行 PCR 的产物。

[0036] 图 2. 原噬菌体 ϕ mru 开放阅读框注释和注解。

[0037] 图 3. 原噬菌体 ϕ mru 开放阅读框注释、功能预测和注解。

[0038] 图 4A-4B. 反刍甲烷短杆菌原噬菌体 ϕ mru 序列信息,包括噬菌体 ϕ mru 的编码序列(图 4A)和氨基酸序列(图 4B)。

[0039] 图 5. 噬菌体 ϕ mru ORF 2058 与马尔堡甲烷杆菌 (*M. marburgensis*) 的 PeiP 和沃氏甲烷杆菌 (*M. wolfeii*) 的 PeiW 的序列比对。

[0040] 图 6 :使用 LogoBar 创建的来自反刍甲烷短杆菌的信号肽序列的蛋白质序列标识,显示了核心共有信号。

[0041] 图 7 :ORF2058 对反刍甲烷短杆菌静息细胞的抑制效应。

[0042] 图 8 :ORF2058 对反刍甲烷短杆菌细胞生长和甲烷产生的抑制效应。

具体实施方式

[0043] 定义

[0044] 如本文所述,“改变的”编码噬菌体多肽的核酸序列包括那些具有不同的核苷酸缺失、插入或取代,得到优选编码相同或功能等同多肽的多核苷酸的核酸。编码的多肽或抗体

也可以是“改变的”，并包含氨基酸残基的缺失、插入或取代，产生沉默改变并得到功能等同的多肽。只要多肽的生物学活性（如细胞结合、细胞通透或细胞裂解）或免疫学活性（如一个或多个抗体结合位点）得以保留，可在残基的极性、电荷、溶解性、疏水性、亲水性和/或两亲特性相似的基础上，有意进行氨基酸取代。例如，带负电的氨基酸可以包括天冬氨酸和谷氨酸；带正电的氨基酸可以包括赖氨酸和精氨酸；具有亲水性值相似的不带电极性头部基团的氨基酸包括亮氨酸、异亮氨酸和缬氨酸，甘氨酸和丙氨酸，天冬酰胺和谷氨酰胺，丝氨酸和苏氨酸以及苯丙氨酸和酪氨酸。

[0045] 如本文所用“氨基酸序列”指寡肽、肽、多肽或蛋白质序列及其任何片段，并且指天然产生的、重组的、合成或半合成的分子。本发明的序列包含至少 5、10、15、20、25、30、35、40、45、50、100、150、200、250 个氨基酸，优选至少 5-10、10-20、20-30、30-40、40-50、50-100、100-150、150-200 或 200-250 或 250-4000 个氨基酸，并且优选保持原始序列的生物学活性（如细胞结合、细胞通透或细胞裂解）或免疫学活性（如一个或多个抗体结合位点）。当本文引用“氨基酸序列”指代天然产生多肽分子的氨基酸序列、氨基酸序列，以及类似术语时，并不意于将氨基酸序列限制为与全长分子相关的完整原始氨基酸序列。

[0046] 本文所用“扩增”指产生核酸序列的其它拷贝，通常使用本发明熟知的聚合酶链反应 (PCR) 技术操作 (Dieffenbach, C. W. 和 G. S. Dveksler (1995) 《PCR 引物, 实验室手册, 冷泉港出版社, 纽约普莱恩维尤》(PCR Primer, a Laboratory Manual, Cold Spring Harbor Press, Plainview, NY))。

[0047] 术语“抗体”应理解为最广泛的可能含义，并意于包含完整的单克隆抗体和多克隆抗体。也意于覆盖抗体片段和衍生物，只要其显示生物学活性。抗体包括免疫球蛋白分子和免疫球蛋白 (Ig) 分子的活性部分，即包含特异性结合抗原（与其免疫反应）的抗原结合位点的分子。这些包括但不限于：多克隆抗体、单克隆抗体、嵌合抗体、单链抗体、Fc、Fab、Fab' 和 Fab₂ 片段和 Fab 表达文库。

[0048] 抗体分子涉及 IgG、IgM、IgA、IgE 和 IgD 的任何种类，它们之间的差异在于分子中重链的特性。它们也包括亚类，如 IgG1、IgG2 和其它。轻链可以是 κ 链或 λ 链。本文提到抗体时包括所有类、亚类和类型。还包括嵌合抗体，例如对于不止一种来源如一种或多种小鼠、人或反刍动物序列具有特异性的单克隆抗体或其片段。还包括羊驼抗体或纳米抗体。应理解，本文中提到的“抗体”或任何类似术语包括完整抗体，及其任何片段、变化形式、衍生物或变体。

[0049] 如本文所用术语“生物学活性”或“功能性”指保持一种或多种结构、免疫学或生物化学功能（如细胞结合、细胞通透或细胞裂解）的多肽。

[0050] 本文所用术语“细胞抑制剂”或“抑制剂”指降低或阻碍微生物细胞，尤其是产甲烷菌细胞生长或复制的试剂。细胞抑制剂可用于降低或阻碍如细胞分裂。抑制剂可减少或阻断例如 DNA 合成、RNA 合成、蛋白质合成或翻译后修饰。抑制剂也可降低或阻断参与甲烷生成通路的酶的活性。抑制剂还可靶向细胞，以便免疫系统组件识别。抑制细胞还包括细胞杀伤和细胞死亡，例如通过裂解、凋亡、坏死等实现。有用的抑制剂包括但不限于：如本文详述的抗甲烷生成化合物（如溴代乙磺酸）、抗体和抗体片段、裂解酶、肽核酸、抗微生物肽和其他抗生素。

[0051] 本文所用术语“互补的”或“互补性”指在允许的盐和温度条件下多核苷酸通过

碱基配对自然结合。对于序列 A-G-T, 互补序列是 T-C-A, 反向互补是 A-C-T 而反向序列是 T-G-A。两条单链分子间的互补性可以是部分的, 即仅有一些核酸结合, 或者为完全互补, 即当单链分子间存在完全 互补性时。核酸链间的互补程度对于核酸链间杂交的效率和强度有显著性影响。这对于依赖核酸链结合的扩增反应以及 PNA 分子的设计和使用尤其重要。

[0052] 本文所用术语“衍生物”指噬菌体多肽的编码核酸或与之互补的核酸的化学修饰形式。此类修饰包括例如, 用烷基、酰基或氨基取代氢。在优选方面, 核酸衍生物编码的多肽保留天然分子的生物学或免疫学功能。

[0053] 衍生多肽是通过糖基化、peg 化或任何相似过程修饰的多肽, 它保留了其来源序列的一种或多种生物学功能(如细胞结合、细胞通透或细胞裂解)或免疫学功能。

[0054] 本文所用术语“同源”指一定程度的互补性。可以是部分同源(即, 1 相同)或完全同源(即, 100% 相同)。使用功能性术语“基本同源”指代至少部分抑制相同序列与靶核酸杂交的部分互补序列。可在低严谨性条件下使用杂交实验(Southern 或 northern 印迹, 溶液杂交等)检验对完全互补序列与靶序列杂交的抑制。基本同源的序列或杂交探针在低严谨性条件下将竞争并抑制完全同源序列与靶序列的结合。这并不是说低严谨性条件允许非特异性结合; 低严谨性条件需要两条序列的相互结合是特异性(选择性)相互作用。

[0055] 如本文所用术语“杂交”指核酸链通过碱基配对与互补链结合的任何过程。

[0056] 本文所用“插入”或“加入”指对氨基酸或核苷酸序列进行改变, 与天然产生的分子相比, 导致加入一个或多个氨基酸残基或核苷酸。

[0057] 本文所用“产甲烷菌”指产生甲烷气体的微生物, 包括甲烷短杆菌(*Methanobrevibacter*)、甲烷嗜热杆菌(*Methanothermobacter*)、甲烷微菌(*Methanomicrobium*)、甲烷杆菌(*Methanobacterium*)和甲烷八叠球菌(*Methanosarcina*)。具体的产甲烷菌包括但不限于: 反刍甲烷短杆(*Methanobrevibacter ruminantium*)、史氏甲烷短杆菌(*Methanobrevibacter smithii*)、酸性甲烷短杆菌(*Methanobrevibacter acididurans*)、巢氏甲烷短杆菌(*Methanobrevibacter thaueri*)、布氏甲烷杆菌(*Methanobacterium bryantii*)、甲酸甲烷杆菌(*Methanobacterium formicicum*)、马尔堡甲烷嗜热杆菌(*Methanothermobacter marburgensis*)、沃氏甲烷嗜热杆菌(*Methanothermobacter wolfeii*)、斯氏甲烷球形菌(*Methanosphaerastadtmanae*)、运动甲烷微菌(*Methanomicrobium mobile*)、巴氏甲烷八叠球菌(*Methanosarcina barkeri*)、梅氏甲烷八叠球菌(*Methanosarcina mazei*)、布氏拟甲烷球菌(*Methanococcoides burtonii*)和泰氏甲烷叶菌(*Methanolobustaylorii*)。所有产甲烷菌的属和种均在此术语涵盖范围内。

[0058] 本文所用“微生物细胞”指天然产生或遗传修饰的微生物细胞, 包括古细菌如产甲烷菌、嗜盐微生物和嗜酸嗜热菌, 以及真细菌, 如蓝藻、螺旋体、变形细菌以及革兰阳性和革兰阴性细菌。

[0059] 术语“修饰的”指如本文所述的更改的序列和序列片段、变体和衍生物。

[0060] 本文所用“核酸序列”或“核苷酸序列”指多核苷酸序列、寡核苷酸或其片段, 以及天然、重组、合成或半合成来源的 DNA 或 RNA, 它们可以是单链或双链, 并可代表正义或反义链以及编码或非编码区。本发明的序列最优选包括包含至少 12、15、30、45、60、75、90、105、120、135、150、300、450、600、750 个核苷酸, 优选至少 15-30、30-60、60-90、90-120、

120-150、150-300、300-450、450-600 或 600-750 个核苷酸或至少 1000 个核苷酸或至少 1500 个核苷酸的多肽编码序列。应理解的是本文中每一处提及“核酸序列”或“核苷酸序列”将包括原始、全长序列,及其任何互补序列、片段、变化形式、衍生物或变体。

[0061] 术语“寡核苷酸”指包含至少 6、8、10、12、15、18、21、25、27、30 或 36 个核苷酸或至少 12-36 个核苷酸或至少 15-30 个核苷酸的核酸序列,可用于例如 PCR 扩增、测序或杂交实验。如本文所用,“寡核苷酸”基本等同于“扩增物”、“引物”、“寡聚物”、“寡核苷酸”以及“探针”,如本领域通常定义的那样。

[0062] 本文所用“多肽”指本发明的分离多肽,其获自任何物种,优选微生物,以及任何天然、合成、半合成或重组的来源。具体说,噬菌体多肽可获自产甲烷菌细胞,如甲烷短杆菌 (*Methanobrevibacter*) 细胞,具体是反刍甲烷短杆菌或史氏甲烷短杆菌细胞。对于重组产生,本发明多肽可获自微生物或真核细胞,例如大肠杆菌 (*Escherichia*)、链球菌 (*Streptomyces*)、芽孢杆菌 (*Bacillus*)、沙门菌 (*Salmonella*)、酵母、昆虫细胞如果蝇细胞、动物细胞如 COS 和 CHO 细胞或植物细胞。应理解的是本文中每一处提及“多肽”将包括原始、全长序列,及其任何片段、变化形式、衍生物或变体。

[0063] 本文中所用单数或复数形式的术语“多核苷酸”一般指任何核酸序列,例如,任何聚核糖核苷酸或聚脱氧核糖核苷酸,它可以是未修饰的 RNA 或 DNA,或者修饰的 RNA 或 DNA。包括但不限于:单链和双链 DNA,包括单链和双链区的 DNA,单链和双链 RNA 以及包括单链和双链区的 RNA,包含 DNA 和 RNA 的杂交分子(可为单链或(更典型的)双链或包含单链和双链区)。也包括含有 RNA 或 DNA 或者 RNA 和 DNA 的三链区。具体说,包括 mRNA、cDNA 和基因组 DNA 及其任何片段。该术语包括含有一个或多个修饰碱基如含氟碱基或非常见碱基如肌苷的 DNA 和 RNA。本发明的多核苷酸可涵盖编码或非编码序列,或正义或反义序列或 iRNA 如 siRNA。应理解的是本文中每一处提及“多核苷酸”或类似术语将包括全长序列,及其任何互补序列、片段、变化形式、衍生物或变体。

[0064] 本文所用“肽核酸”或“PNA”指反义分子或抗-基因试剂,其包含通过肽主链连接的碱基。

[0065] 本文所用术语“反刍动物”指具有瘤胃作为特殊类型消化器官的动物。反刍动物包括但不限于:牛、绵羊、山羊、水牛、麋鹿、羚羊、北美驯鹿和鹿。

[0066] 本文所用术语“严谨条件”或“严谨性”指核酸、盐和温度限定的杂交条件。这些条件为本领域所熟知,可更改这些条件以鉴定或检测相同或相关的多核苷酸序列。参见例如 Sambrook, J. 等 (1989)《分子克隆,实验室手册》冷泉港出版社,纽约普莱恩维尤 (*Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview, NY), 以及 Ausubel, F. M. 等 (1989) 新编分子生物学实验指南,约翰韦利森出版社,纽约 (*Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY)。数种包含低或高严谨性的相等条件取决于以下因素,如序列的长度和特性 (DNA、RNA、碱基组成)、靶点特性 (DNA、RNA、碱基组成)、环境 (在溶液或固定于固体基材上)、盐和其它成分 (如甲酰胺、硫酸右旋糖苷和/或聚乙二醇) 的浓度和反应温度 (从低于探针解链温度 5°C 至低于解链温度约 20°C -25°C 的范围内)。可改变一种或多种因素以产生不同于或等同于上述条件的低或高严谨性。

[0067] 术语“对象”包括人或非人动物。非人动物包括但不限于:鸟类和哺乳动物,如反刍动物,具体是小鼠、兔、猫、狗、猪、绵羊、山羊、牛和马。

[0068] 本文所用术语“基本纯化”或“分离”指从细胞、重组或合成环境中移出的核酸或氨基酸序列,并且至少至少 60%,优选 75%并最优选至少 90%或至少 99%不含它们在细胞、重组或合成环境中相关联的其它组分。

[0069] 本文所定义“转化”指外源 DNA 进入并改变接受细胞的过程。可使用本领域所熟知多种方法在天然或人工条件下进行转化。转化可依赖将外源核酸序列插入原核或真核宿主细胞的任何已知方法。基于需转化的宿主细胞类型选择方法,可包括但不限于病毒感染、电穿孔、热休克、脂质转染和粒子轰击。此类“转化”细胞包括稳定转化细胞,其中插入的 DNA 能够作为自主复制质粒或作为宿主染色体的一部分进行复制。它们还包括在限定时间期间瞬时表达插入的 DNA 或 RNA 的细胞。

[0070] 如本文所用多肽“变体”指一个或多个氨基酸改变的氨基酸序列。改变变体多核苷酸的一个或多个核苷酸。变体可导致“保守性”改变,其中取代的氨基酸具有相似结构或化学特性,如用异亮氨酸取代亮氨酸。更罕见的,变体可导致“非保守性”变化,如用色氨酸替换甘氨酸。类似的小变异还可包括氨基酸缺失或插入或两者均有。使用本领域熟知计算机程序如 LASERGENE 软件 (DNASTAR) 可发现如何确定可对哪一氨基酸取代、插入或缺失而不破坏生物学或免疫学活性的指南。

[0071] 本发明还涵盖保留多肽的至少一种生物学活性(如细胞结合、细胞通透或细胞裂解)或免疫学活性的变体。优选的变体是具有基本相同或功能等同序列的变体,例如与公开序列至少 80%和更优选至少 90%序列相同的变体。最优选的变体是与本文公开序列至少 95%,至少 97%,至少 98%,至少 99%,至少 99.5%,至少 99.8%或至少 99.9%序列相同的变体。如下所述通过对需比较的两条序列比对、确定比对部分的相同残基数目、除以本发明(查询)序列总残基数并乘以 100,从而确定相同性百分数。有用的比对程序是 AlignX(载体 NTI)。

[0072] 发明详述

[0073] 甲烷在反刍动物前肠中由产甲烷菌产生,它是瘤胃系统中碳的终末还原产物。已对甲烷产生通路的多个步骤进行了阐述,主要来自于对非瘤胃产甲烷菌的研究,但是尚不了解对于使产甲烷菌在瘤胃中生长和持续存在的适应性。反刍甲烷短杆菌(*Methanobrevibacter ruminantium*)是新西兰反刍动物体内主要的产甲烷菌。如本文所述,已经对反刍甲烷短杆菌的基因组进行测序,显示大小约 3.0Mb, GC 含量为 33.68%。一个出乎意料的发现是反刍甲烷短杆菌基因组包括原噬菌体序列(称为 ϕ mru),该序列具有编码噬菌体整合、DNA 复制和包装、衣壳蛋白质、裂解和溶原性转换功能的独特功能模块。

[0074] 在进行高通量测序时,发现基因组的 30-40kb 区域在测序克隆中过度出现,因此发现反刍甲烷短杆菌噬菌体。这表明基因组的一部分以高于通常水平的拷贝数出现,这可能归因于驻留噬菌体的复制。对过度出现的区域进行研究,预测开放阅读框的详细生物信息学分析表明它含有噬菌体样基因。已证明噬菌体序列(溶原性转换)远端发现的低 GC 区含有预测的 DNA 甲基化系统(dnd),它可能为宿主或外源 DNA 的提供其它修饰。本文详述了反刍甲烷短杆菌原噬菌体序列。在本发明的不同方面,可以使用原噬菌体多核苷酸和多肽作为在瘤胃中抑制产甲烷菌和/或甲烷生成的手段,并进一步阐明反刍甲烷短杆菌在甲烷形成中的作用。

[0075] 因此,本发明包括噬菌体多肽,包括含有至少 SEQ ID NO:1-69 中至少一种或其片

段、变体和衍生物的那些多肽。本发明还包括这些多肽在靶向和抑制微生物细胞,尤其是产甲烷菌细胞中的应用。本发明还包括该多肽在抑制此类细胞生长或复制中的应用。可表达本发明的多肽并用于不同的实验以测定其生物学活性。该多肽可用于大规模合成和分离实验方案,例如,用于商业生产。可使用此类多肽产生抗体,以分离相应的氨基酸序列 和对氨基酸序列水平进行定量测定。本发明的多肽还可用作组合物,例如药物组合物和食品补充剂,如反刍动物饲料组分。本发明的多肽还具有健康益处。在健康相关方面,产甲烷菌抑制剂可用于恢复通常以甲烷形式从个体流失的能量。在特定方面,缓释瘤胃装置可与本发明的多肽和组合物(如药物组合物和食品补充剂)联用。

[0076] 本发明多肽包含选自下组的至少一个序列:(a) 包含至少一个选自 SEQ ID NO: 1-69 或其片段、变体或衍生物的氨基酸序列的多肽;(b) 包含至少一个选自 SEQ ID NO: 1-69 和其片段、变体的氨基酸序列的功能性结构域的多肽;以及(c) 包含的选自 SEQ ID NO: 1-69 或其变体或衍生物的至少一个氨基酸序列的至少一定数目的毗连残基的多肽。在一个实施方式中,本发明包括含有 SEQ ID NO: 1-69 中至少一个氨基酸序列的分离的多肽。所有这些序列统称本发明的多肽。

[0077] 本发明还包括编码至少一种噬菌体多肽,包括 SEQ ID NO: 1-69 及其片段、变体和衍生物的多核苷酸。本发明还包括这些多核苷酸在制备靶向和抑制微生物细胞尤其是产甲烷菌细胞的表达载体和宿主细胞中的应用。本发明还包括多核苷酸在抑制此类细胞生长或复制中的应用。本发明的分离的多核苷酸还可用于对或多或少的相关噬菌体进行基因组作图、物理作图和基因克隆。通过本领域熟知技术如狭缝印迹技术或微阵列技术,可使用根据本发明多核苷酸设计的探针在细胞中具有充分同源 DNA 和 RNA 序列的任何生物体中检测基因的存在和表达模式。使用本发明多核苷酸设计的引物可用于测序和 PCR 扩增。本发明的多核苷酸还可用作组合物,例如药物组合物和食品补充剂,如,反刍动物饲料组分。本发明的多核苷酸还具有健康益处。对于此类益处,多核苷酸可作为表达载体或含表达载体的宿主细胞提供。在特定方面,缓释瘤胃装置可与本发明的多核苷酸、载体、宿主细胞和组合物(如药物组合物和食品补充剂)联用。

[0078] 本发明多核苷酸包含选自下组的至少一个序列:(a) 包含至少一个选自 SEQ ID NO: 1-69 或其片段或变体的氨基酸序列的编码序列的序列;(b) 至少一个选自 SEQ ID NO: 1-69 或其片段或变体的氨基酸序列的编码序列的互补序列、反向序列以及反向互补序列;(c) 至少一个选自 SEQ ID NO: 1-69 或其片段或变体的氨基酸序列的编码序列中所含的开放阅读框;(d) 至少一个选自 SEQ ID NO: 1-69 或其片段或变体的氨基酸序列的编码序列的功能性结构域;以及(e) 含有至少一个选自 SEQ ID NO: 1-69 或其变体的氨基酸序列的编码序列的至少一定数目的毗连残基的序列。在一个方面,本发明包括包含至少一个选自 SEQ ID NO: 1-69 的氨基酸序列的编码序列的分离多核苷酸。

[0079] 本发明多核苷酸包含选自下组的至少一个序列:(a) 包含至少一个选自 SEQ ID NO: 74-142 或其片段或变体的核酸序列的序列;(b) 至少一个选自 SEQ ID NO: 74-142 或其片段或变体的核酸序列的编码序列的互补序列、反向序列以及反向互补序列;(c) 选自 SEQ ID NO: 74-142 或其片段或变体的核酸序列中所含的开放阅读框;(d) 至少一个选自 SEQ ID NO: 74-142 或其片段或变体的核酸序列的编码序列的功能性结构域;以及(e) 含有至少一个选自 SEQ ID NO: 74-142 或其变体的核酸序列的至少一定数目的毗连残基的序列。还提

供获自任何公开序列的寡核苷酸探针和引物及其变体。所有这些多核苷酸和寡核苷酸在本文中统称为本发明的多核苷酸。

[0080] 本领域熟练技术人员应当理解的是,作为遗传密码简并性的结果,可产生大量本发明多肽的编码核苷酸序列,其中某些序列与任何已知和天然产生的基因的核苷酸序列同源性很小。因此,本发明考虑到各个和每个可能的核苷酸序列变异,这些变异可通过根据可能密码子选择密码子组合而产生。根据应用于天然产生的氨基酸序列的标准细菌三联遗传密码产生这些组合,所有此类变异都被视作已具体公开。

[0081] 编码噬菌体多肽或其片段或变体的核苷酸序列优选在适当选择的严谨条件下能够与天然产生序列的核苷酸序列杂交。然而,具有优势的是产生编码多肽或其片段或衍生物的、具有明显不同密码子使用的核苷酸序列。根据宿主使用特定密码子的频率,可选择密码子以提高特定原核或真核宿主中的多肽表达速率。例如,按照已知方法,可根据大肠杆菌表达优化密码子。改变多肽及其衍生物的编码核苷酸序列而不改变编码的氨基酸序列的其它原因包括产生具有更有利特性,如比天然产生序列产生的转录物具有更长半衰期的 RNA 转录物。

[0082] 本发明还包括完全通过合成化学产生编码多肽或其片段或变体的 DNA 序列或其片段。在合成序列产生后,利用本领域熟知试剂,可将其插入到任何可用的表达载体和细胞系统中。而且,可使用合成化学在编码多肽或其任何变体或片段的序列中引入突变。本发明还包括在如 Wahl, G. M 和 S. L. Berger (1987 ;Methods Enzymol. 152 :399-407) 以及 Kimmel, A. R. (1987 ;Methods Enzymol. 152 :507-511) 所教导的各种严谨条件下,能与要求权利的核苷酸序列杂交的多核苷酸序列,尤其是 SEQ ID NO :74-149 中显示的那些及其互补序列。

[0083] 本领域熟知并常用的 DNA 测序方法可用于实践本发明的任何实施方式。这类方法可使用如 DNA 聚合酶 I 的克列诺片段, SEQUENASE (美国生物化学集团 (U. S. Biochemical Corp.), 俄亥俄州克里富兰) (U. S. Biochemical Corp, Cleveland, OH), Taq 聚合酶 (珀金埃尔默) (Perkin Elmer), 热稳定 T7 聚合酶 (安法马西亚生物技术公司, 新泽西州皮斯卡塔韦) (Amersham Pharmacia Biotech Piscataway, NJ), 或者聚合酶和校对核酸外切酶的组合, 如生命技术公司 (Life Technologies) (马里兰州盖瑟斯堡) 销售的 ELONGASE 扩增系统中找到的那些。优选地, 该方法通过机器自动化进行, 如汉密尔顿微型实验室 2200 (Hamilton Micro Lab 2200) (内华达州雷诺汉密尔顿) (Hamilton, Reno, NV), 派而特热循环仪 (Peltier Thermal Cycler) (PTC200 ;MJ 研究公司, 麻省沃特敦) (MJ Research, Watertown, MA) ;ABI 催化仪 (ABI Catalyst) 以及 373 和 377DNA 测序仪 (珀金埃尔默) (Perkin Elmer) 或基因组测序仪 20™ (罗氏诊断公司) (Roche Diagnostics)。

[0084] 可使用部分核苷酸序列并利用本领域所知检测上游序列如启动子和调控元件以及下游元件如终止子和非编码 RNA 结构的方法, 对多肽的编码核酸序列进行扩展。例如, 可使用的一种方法“位点限制性”PCR 使用通用引物以获取与已知基因座相邻的未知序列 (Sarkar, G. (1993) PCR Methods Applic. 2 :318-322)。具体说, 在接头序列引物和已知区域特异性引物存在的条件下对基因组 DNA 进行首次扩增。然后对扩增序列进行第二轮 PCR, 其中使用相同的接头引物和在第一引物内的另一特异性引物。用合适的 RNA 聚合酶对每一轮 PCR 产物进行转录, 用逆转录酶进行测序。

[0085] 另一个有用的方法是反向 PCR, 也称为 IPCR (参见例如, Ochman H, Gerber AS,

Hartl DL. *Genetics*. 1988 年 11 月 ;120(3) :621-3)。当仅知道靶 DNA 的一个内部序列时,可使用反向 PCR。反向 PCR 方法包括使用限制性核酸内切酶切割的 DNA 进行一系列消化和自连。此种切割在未知序列的任一末端产生已知序列。按照该方法,通过限制性核酸内切酶消化将靶 DNA 轻微切割成数千碱基的较小片段。然后在低浓度下诱导自连,引起磷酸主链重新形成并产生环状 DNA 连接产物。然后用已知核酸内切酶限制性消化靶 DNA。由此在已知内部序列内产生切割,产生具有已知末端序列的线性产物。然后可利用该产物和与已知内部序列互补的引物进行标准的 PCR。

[0086] 市售毛细管电泳系统可用于分析测序或 PCR 产物的核苷酸序列的大小或对其进行确认。具体说,毛细管测序可使用可流动聚合物进行电泳分离,激光激发的四种不同荧光染料(每种对应一种核苷酸),用电荷耦合器件照相机探测发射波长。使用合适的软件(如基因组分型和序列导航者,珀金埃尔默公司)(GENOTYPER and Sequence NAVIGATOR, Perkin Elmer) 可将输出/光强度转化为电信号,从上样到计算机分析和电子数据显示均可在计算机控制下进行。毛细管电泳尤其优选用于对特定样品中少量存在的小片段 DNA 进行测序。

[0087] 近年来,焦磷酸测序成为一种有用的测序方法。参见例如 Ronaghi, M. 等 1996. 通过探测焦磷酸释放进行实时 DNA 测序 (Real-time DNA sequencing using detection of pyrophosphate release). *Anal. Biochem.* 242 :84-89 ;Ronaghi, M. 等 1998. 基于实时焦磷酸的测序方法 (A sequencing method based on real-time pyrophosphate). *Science* 281 :363-365 ;Ronaghi, M. 等 1999. 通过焦磷酸测序分析 DNA 二级结构 (Analyses of secondary structures in DNA by pyrosequencing). *Anal. Biochem.* 267 :65-71 ; Ronaghi 2001. *Genome Res.* 11 卷, 1 期, 3-11 ;Nyrén, 焦磷酸测序历史 (The history of pyrosequencing). *Methods Mol Biol.* 2007 ;373 :1-14。焦磷酸测序的优势是准确、灵活、平行处理并容易自动化。而且,该技术摒弃了对标记引物、标记核苷酸和凝胶电泳的需要。根据该方法,聚合酶催化将核苷酸掺入核酸链。掺入的结果是焦磷酸分子释放并由硫酸化酶转化为 ATP。萤光素酶反应产生光,其间萤光素分子被氧化。在加入每种核苷酸后,进行洗涤步骤以允许反复添加。核苷酸降解酶持续降解核苷酸从而允许加入后续核苷酸。已经成功将焦磷酸测序作为平台用于大规模测序,包括基因组和巨大基因组分析(参见例如,来自 454 罗氏生命科学 (Life Sciences/Roche) 的基因组测序仪 FLX™)。

[0088] 研发 SOLiD™ 系统用于测序(参见例如应用生物系统公司, SOLiD™ 系统应用说明书,加州福斯特城)(Applied Biosystems. Application Fact Sheet for the SOLiD™ System. Foster City, CA)。该方法基于将染料标记的寡核苷酸依次连接于磁珠连接的克隆扩增 DNA 片段。在该方法中,通过测量连续连接产生 DNA 序列。连接反应基于探针识别而非依次加入,因此不易于累计错误。化学性质大大减少了错误插入或缺失的可能性。连接步骤和用磷酸酶处理未连接探针能防止移相。此外,经历 7 个循环的连接后,原始引物从模板上脱离,新引物杂交以开始对 n-1 位置的查询。使用这种“重置”相允许降低系统性噪声并允许更长的阅读长度。此外,使用两碱基编码以区分测量误差与真正的多态性。单一位点的改变被鉴定为随机误差,并在数据分析中可被软件去除。作为分析平台, SOLiD™ 系统可应用于大规模测序,数字基因表达, ChIP 和甲基化研究,并在探测基因组变异上尤其有用。

[0089] 在本发明的另一实施方式中,编码多肽的多核苷酸或其片段可用于重组 DNA 分子以在合适宿主细胞中指导多肽或其片段或变体的表达。由于遗传编码的固有简并性,可产

生编码基本相同或功能等同的氨基酸序列的其它 DNA 序列,这些序列可用于克隆和表达噬菌体多肽。可使用本领域通常所知方法对本发明的核苷酸序列进行工程改造以根据多种原因改变氨基酸编码序列,包括但不限于为了改变克隆、加工和 / 或基因产物表达而进行的变化。可使用通过随机片段化进行的 DNA 改组以及对基因片段和合成寡核苷酸的 PCR 组装对核苷酸序列进行工程改造。例如,可使用定位诱变插入新的限制性位点,改变糖基化形式,改变密码子偏好,引入突变等等。

[0090] 在本发明的另一实施方式中,可将天然、修饰或重组的编码多肽的核酸序列连接至异源序列以编码融合蛋白。例如,编码可被市售抗体识别的 嵌合序列可能是有用的。也可工程改造融合蛋白使其在本发明的多肽和异源蛋白序列间含有切割位点,以便于切割并纯化该多肽,与异源部分分离。

[0091] 在另一实施方式中,可使用本领域熟知化学方法合成完整或部分形式的多肽编码序列(参见 Caruthers, M. H. 等 (1980) Nucl. Acids Res. Symp. Ser. 215-223, Horn, T. 等 (1980) Nucl. Acids Res. Symp. Ser. 225-232)。或者,可使用合成氨基酸序列或其片段的化学方法产生多肽本身。例如,可使用多种固相合成技术 (Roberge, J. Y. 等 (1995) Science 269 :202-204 ; Merrifield J. (1963) J. Am. Chem. Soc. 85 :2149-2154) 进行多肽合成,使用例如 ABI 431A 肽合成仪(珀金埃尔默公司)进行自动合成。可用化学方法合成多肽的各种片段,并用化学方法组合产生全长分子。

[0092] 可用制备型高效液相色谱分离新合成的多肽(如, Creighton, T. (1983) 蛋白质结构和分子原则, WH 富瑞曼公司, 纽约 (Proteins Structures and Molecular Principles, WH Freeman and Co., New York, NY))。通过氨基酸分析和测序可确认合成多肽的组成(如, 埃得曼降解步骤 (Edman degradation procedure ; Creighton), 同上)。此外,多肽或其任何部分的氨基酸序列可在直接合成中更改和 / 或通过化学方法与来自其它蛋白质或其部分的序列组合以产生变体分子。

[0093] 为了表达生物学活性多肽,可将编码多肽或功能等同物的核苷酸序列插入合适的表达载体,即包含插入序列转录和翻译必需元件的载体。可用本领域熟练技术人员熟知的方法构建包含多肽编码序列以及合适的转录和翻译控制元件的表达载体。

[0094] 这些方法包括体外重组 DNA 技术、合成技术和体内遗传重组。

[0095] Sambrook, J. 等 (2001) 《分子克隆, 实验室手册》冷泉港出版社, 纽约普莱恩维尤 (Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, NY), 以及 Ausubel, F. M. 等 (2007) 新编分子生物学实验指南, 约翰韦利森出版社, 纽约 (Current Protocols in Molecular Biology, John Wiley & Sons, New York, NY) 对此类技术进行了描述。

[0096] 可使用多种表达载体 / 宿主系统包含并表达本发明多肽的编码序列。它们包括但不限于: 微生物体如重组噬菌体、质粒或粘粒 DNA 表达载体转化的细菌; 酵母表达载体转化的酵母; 病毒表达载体(如杆状病毒)转化的昆虫细胞系统; 病毒表达载体(如, 花椰菜花叶病毒 CaMV; 烟草花叶病毒 TMV) 或细菌表达载体(如, Ti 或 pBR322 质粒)转化的植物细胞; 或动物细胞系统。对于细菌, 有用的质粒包括英骏公司 (Invitrogen) 的 pET、pRSET、pTrcHis2 和 pBAD 质粒; 诺瓦基公司 (Novagen) 的 pET 和 pCDF 质粒以及西格玛奥德里奇公司 (Sigma-Aldrich) 的 Director™ 质粒。对于产甲烷菌, 有用的质粒包括但不限于 pME2001、

pMV15 和 pMP1。具体说,可将大肠杆菌 (*Escherichiacoli*) 与表达载体 pET 一起使用。所用表达载体或宿主细胞并不限制本发明。

[0097] “控制元件”和“调节序列”是载体的非翻译区—增强子、启动子、5′ 和 3′ 非翻译区——它们与宿主细胞蛋白质相互作用,进行转录和翻译。此类元件的强度和特异性可能不同。根据所使用载体系统和宿主,可使用任何数量的合适的转录和翻译元件,包括组成型和诱导型启动子。例如,当克隆进细菌系统时,可使用诱导型启动子,如 BLUESCRIPT 噬菌粒(司查塔基公司,加州拉霍亚)(Stratagene, LaJolla, CA) 或 pSPORT1 质粒(生命技术公司)(Life Technologies) 的杂交 lacZ 启动子等。可在昆虫细胞中使用杆状病毒多角体蛋白启动子。可将来源于植物细胞基因组(如,热休克、RUBISCO 和存储蛋白质基因)的启动子或增强子或来自于植物病毒的启动子或增强子(如病毒启动子或前导序列)克隆到载体中。

[0098] 在细菌系统中,可根据多肽的指定应用对许多表达载体进行选择。例如,当需要大量多肽时,可使用指导高水平表达易于纯化的融合蛋白的载体。此类载体包括但不限于:多功能大肠杆菌克隆和表达载体如 BLUESCRIPT(司查塔基公司),其中可将多肽编码序列连接入载体,与 β -半乳糖苷酶的氨基末端 Met 和后续 7 个残基在相同读框内,从而产生杂合蛋白;pIN 载体 (Van Heeke, G. 和 S. M. Schuster (1989) *J. Biol. Chem.* 264 :5503-5509) 等。

[0099] 也可采用 pGEX 载体(普洛麦格公司,威斯康星州麦迪逊)(Promega, Madison, WI) 表达多肽与谷胱甘肽-S-转移酶 (GST) 的融合蛋白。通常,这类融合蛋白可溶,并容易通过以下方法由裂解细胞纯化:吸附于谷胱甘肽琼脂糖珠上,然后在游离谷胱甘肽的存在下洗脱。可设计此类系统产生的蛋白质以使其包含肝素、凝血酶或 Xa 因子蛋白酶切割位点,从而使感兴趣的克隆多肽可按需从 GST 部分释放出来。在酿酒酵母 (*Saccharomyces cerevisiae*) 中,可使用包含组成型或诱导型启动子的多种载体,如 α 因子、醇氧化酶和 PGH。参见 Ausubel 等(同上)和 Grant 等 (1987) *Methods Enzymol.* 153 : 516-544 的综述。

[0100] 也可使用特异性起始信号达到更有效的翻译本发明多肽编码序列的目的。此类信号包括 ATG 启动密码子和相邻序列。当多肽编码序列、其起始密码子和上游序列插入到合适表达载体的情况下,无需另外的转录或翻译控制信号。然而,当仅插入编码序列或其片段时,需提供外源的翻译控制信号包括 ATG 起始密码子。而且,起始密码子应当位于正确的阅读框中,以保证完整插入物的翻译。外源翻译控制元件和起始密码子可有多种天然或合成来源。通过包括适合所用特定细胞体系的增强子可增强表达效率,如文献中所述 (Scharf, D. 等, (1994) *Results Probl. Cell Differ.* 20 :125-162)。

[0101] 此外,可根据宿主细胞株调节插入序列表达或以所期望的方式加工表达多肽的能力对其进行选择。此类序列的修饰包括但不限于:乙酰化、羧化、糖基化、磷酸化、脂化和酰基化。也可使用切割多肽“前体”形式的翻译后加工以利于正确的插入、折叠和/或功能。可从美国典型培养物保藏中心(马里兰州贝塞斯达)(American Type Culture Collection(ATCC;Bethesda, MD) 获取具有特定细胞机器或翻译后活性的特征性机制的不同宿主细胞,选择以保证正确的序列修饰和加工。具体宿主细胞包括但不限于:产甲烷菌细胞,如产甲烷短杆菌细胞,具体说,反刍甲烷短杆菌或史氏甲烷短杆菌细胞。感兴趣的宿主细胞还包括例如,红酵母 (*Rhodotorula*)、短梗霉菌 (*Aureobasidium*)、酿酒酵

母 (*Saccharomyces*)、掷孢酵母 (*Sporobolomyces*)、假单胞菌 (*Pseudomonas*)、欧文氏菌 (*Erwinia*) 和黄杆菌 (*Flavobacterium*) ;或其它此类有机体如埃希菌 (*Escherichia*)、乳酸杆菌 (*Lactobacillus*)、芽胞杆菌 (*Bacillus*)、链霉菌 (*Streptomyces*) 等。特定的宿主细胞包括尤其适于本发明使用的大肠杆菌 (*Escherichia coli*)、酿酒酵母 (*Saccharomyces cerevisiae*)、苏云金芽孢杆菌 (*Bacillus thuringiensis*)、枯草芽孢杆菌 (*Bacillus subtilis*)、变青链霉菌 (*Streptomyces lividans*) 等。

[0102] 数种技术可将核酸引入体外培养的真核细胞中。它们包括化学方法 (Felgner 等, *Proc. Natl. Acad. Sci., USA*, 84 :74137417(1987) ;Bothwell 等, 真核基因克隆和分析方法 (*Methods for Cloning and Analysis of Eukaryotic Genes*), 编, 约翰和巴特勒出版公司, 麻省波士顿 (Jones and Bartlett Publishers Inc., Boston, Mass.) (1990) ;Ausubel 等, 分子生物学简明实验方案 (*Short Protocols in Molecular Biology*), 约翰韦利森公司, 纽约 (John Wiley and Sons, New York, NY) (1992) ;以及 Farhood, *Ann. NY Acad. Sci.*, 716 :2334(1994)), 使用原生质体 (Bothwell, 同上) 或电脉冲 (Vatteroni 等, *Mutn. Res.*, 291 :163169(1993) ;Sabelnikov, *Prog. Biophys. Mol. Biol.*, 62 :119 152(1994) ;Bothwell 等, 同上 ;以及 Ausubel 等, 同上), 使用减毒病毒 (Davis 等, *J. Virol.* 1996, 70(6), 3781 3787 ;Brinster 等 *J. Gen. Virol.* 2002, 83(Pt 2), 369 381 ;Moss, *Dev. Biol. Stan.*, 82 :55 63(1994) ;以及 Bothwell 等, 同上), 以及物理方法 (Fynan 等, 同上 ;Johnston 等, *Meth. Cell Biol.*, 43(PtA) :353 365(1994) ;Bothwell 等, 同上 ;以及 Ausubel 等, 同上)。

[0103] 可通过以下方法将核酸成功递送至动物组织 :使用阳离子脂质体 (Watanabe 等, *Mol. Reprod. Dev.*, 38 :268 274(1994)), 直接将裸露的 DNA 或 RNA 注射入动物肌肉组织 (Robinson 等, *Vacc.*, 11 :957 960(1993) ;Hoffman 等, *Vacc.* 12 :1529 1533(1994) ;Xiang 等, *Virol.*, 199 :132 140(1994) ;Webster 等, *Vacc.*, 12 :1495 1498(1994) ;Davis 等, *Vacc.*, 12 :1503 1509(1994) ;Davis 等, *Hum. Molec. Gen.*, 2 :1847 1851(1993) ;Dalemans 等, *Ann NY Acad. Sci.* 1995, 772, 255 256. Conry 等, *Cancer Res.* 1995, 55(7), 1397-1400) 和胚胎 (Naito 等, *Mol. Reprod. Dev.*, 39 :153 161(1994) ;和 Burdon 等, *Mol. Reprod. Dev.*, 33 :436 442(1992)), 肌肉内注射自我复制的 RNA 疫苗 (Davis 等, *J. Virol.* 1996, 70(6), 3781 3787 ;Balasuriya 等, *Vaccine* 2002, 20(1112), 1609 1617) 或者用“基因枪”技术皮内注射 DNA (Johnston 等, 同上)。

[0104] 用蛋白质特异性多克隆或单克隆抗体检测和测定本发明多肽表达的各种实验方案为本领域所知。例子包括酶联免疫吸附实验 (ELISA)、放射性免疫 (RIA) 和荧光激活的细胞分选 (FACS)。可使用与多肽上两个非干扰表位具有反应性的单克隆抗体进行双位点单克隆免疫实验, 也可使用竞争结合实验。这些和其它实验参见 Hampton, R. 等 (1990 ;血清学方法, 实验室手册 (*Serological Methods, a laboratory Manual*), APS 出版社, 明尼苏达州圣保罗) 以及 Maddox, D. E. 等 (1983 ;*J. Exp. Med.* 158 :1211-1216) 等。

[0105] 本领域熟练技术人员了解多种标记和偶联技术, 可用于多种核酸和氨基酸实验。产生标记的杂交或 PCR 探针用于探测多核苷酸相关序列的方法包括寡核苷酸标记、缺口平移、末端标记或使用标记核苷酸进行 PCR 扩增。或者, 可将多肽或其任何片段或变体的编码序列克隆入载体用于产生 mRNA 探针。此类载体为本领域所知并可购得, 可在加入合适 RNA 聚合酶如 T7、T3 或 SP6 以及标记的核苷酸后用于体外合成 RNA 探针。可利用安法马西亚

生物技术公司、普洛麦格公司和美国生物化学公司 (AmershamPharmacia Biotech, Promega and US Biochechemical) 生产的多种市售试剂盒进行这些步骤。易于探测的合适报告分子或标记包括放射性核素、酶、荧光物质、化学发光物质或发色剂以及底物、辅因子、抑制剂、磁性颗粒等。

[0106] 可在适于表达和从培养物中回收多肽的条件下,复制表达载体或用表达载体转化的宿主细胞。培养物可包含体外或体内表达组件。体外表达组件包括兔网织红细胞裂解物、大肠杆菌裂解物和麦芽提取物的体外表达组件,例如,英骏公司 (Invitrogen) 的 Expressway™或 RiPs 系统,英特龙生物技术公司 (iNtRON Biotechnology) 的 Genelator™系统,诺瓦基公司 (Novagen) 的 EcoPro™或 STP3™系统,普洛麦格公司 (Promega) 的 **TNT®** 快速偶联系统和凯杰公司 (QIAGEN) 的 EasyXpress 系统。产自培养物的多肽可以是分泌的或细胞内包含的多肽,这取决于序列和 / 或所用载体。在某些特定方面,可设计编码噬菌体多肽的表达载体使其包含指导多肽通过原核或真核细胞膜分泌的信号序列。

[0107] 其它构建物可包含利于多肽纯化的氨基酸结构域。此类结构域包括但不限于:可在固化金属上进行纯化的金属螯合肽如组氨酸-色氨酸(如 6X-HIS (SEQ ID NO:150)) 模块,可在固定免疫球蛋白上进行纯化的蛋白质 A 结构域,以及在 **FLAG®** 扩展 / 亲和纯化系统(因缪耐斯集团,华盛顿州西雅图)(Immunex Corp., Seattle, WA) 中所用结构域。可用的表位标签包括 **3XFLAG®**、HA、VSV-G、V5、HSV、GST、GFP、MBP、GAL4 和 β -半乳糖苷酶。有用的质粒包括含有生物素标签的质粒(如,普洛麦格公司的 PinPoint™质粒),含有钙调蛋白结合蛋白的质粒(如司查塔基公司的 pCAL 质粒),含有链霉亲和素结合肽的质粒(如司查塔基公司的 InterPlay™质粒),含有 c-myc 或 **FLAG®** 标签的质粒(如西格玛奥德里奇公司的免疫沉淀质粒)或含有组氨酸标签的质粒(如凯杰公司的 QIAExpress 质粒)。

[0108] 为了利于纯化,表达载体可包含可切割的接头序列,如 Xa 因子或肠激酶(英骏公司,加州圣地亚哥)的特异性接头序列。例如,载体在纯化结构域和多肽间可包含一个或多个接头。一种此类表达载体能够用于表达包含本发明多肽的融合蛋白,并提供编码硫氧还蛋白或肠激酶切割位点前 6 个组氨酸的核酸。组氨酸残基利于在 IMAC(固定金属离子亲和色谱柱,如 Porath, J 等 (1992) Prot. Exp. Purif. 3:263-281 所述)上纯化,而肠激酶切割位点则提供一种从融合蛋白中纯化多肽的方法。Kroll, D. J. 等 (1993; DNA Cell Biol. 12: 441-453) 提供了关于包含融合蛋白的载体的讨论。

[0109] 可使用本领域通常所知方法产生本发明抗体,用于纯化或诊断技术。具体说,根据通常所知的实验方案,可使用多肽或多核苷酸产生抗体。此类抗体可包括但不限于:多克隆、单克隆、嵌合和单链抗体、Fab 片段和 Fab 表达文库产生的片段。本发明尤其优选使用中和抗体(即抑制功能的那些抗体)。

[0110] 为了产生抗体,可将具有免疫原性的多肽、多核苷酸或其任何片段注射入不同的宿主,包括山羊、兔、大鼠、人等,以进行免疫。根据宿主种类,可使用不同的佐剂以提高免疫应答。此类佐剂包括但不限于:福氏佐剂、矿物凝胶如氢氧化铝和表面活性物质如溶血卵磷脂、普流罗尼克多元醇、聚阴离子、肽、油乳剂、钥孔血蓝素和二硝基酚。用于人的佐剂中,尤其优选 BCG(卡介苗)和短小棒状杆菌 (Corynebacterium parvum)。

[0111] 用于诱导抗体的多肽或片段优选具有包含至少 5 个氨基酸,更优选 10 个氨基酸的氨基酸序列。优选的是,它们与天然蛋白质氨基酸序列的一部分相同,以及它们可包含小的

天然产生分子的完整氨基酸序列。可将氨基酸短臂与另一蛋白质如钥孔血蓝素以及抗嵌合分子的抗体的短臂融合。

[0112] 使用通过连续培养细胞系产生抗体分子的任何技术制备单克隆抗体。这些技术包括但不限于杂交瘤技术,人 B- 细胞杂交瘤技术和 EBV- 杂交瘤技术 (Kohler, G. 等 (1975) *Nature* 256 :495-497 ;Kozbor, D. 等 (1985) *J. Immunol. Methods* 81 :31-42 ;Cote, R. J 等 (1983) *Proc. Natl. Acad. Sci.* 80 :2026-2030 ;Cole, S. P. 等 (1984) *Mol. Cell Biol.* 62 :109-120)。如文献报道,也可通过诱导淋巴细胞在体内生产抗体,或通过筛选免疫球蛋白文库或高特异性结合试剂板块产生抗体 (Orlandi, R. 等 (1989) *Proc. Natl. Acad. Sci.* 86 :3833-3837 ;Winter, G. 等 (1991) *Nature* 349 :293-299)。

[0113] 此外,也可使用某些技术产生“嵌合抗体”,如将抗体基因组合以获得具有合适抗原特异性和生物学活性的分子 (Morrison, S. L. 等 (1984) *Proc. Natl. Acad. Sci.* 81 :6851-6855 ;Neuberger, M. S. 等 (1984) *Nature* 312 :604-608 ;Takeda, S. 等 (1985) *Nature* 314 :452-454)。或者,可通过本领域所知方法改进用于产生单链抗体的技术,以产生特异性单链抗体。具有相关特异性但独特型组成不同的抗体可通过随机组合免疫球蛋白文库的链改组产生 (Burton D. R. (1991) *Proc. Natl. Acad. Sci.* 88 :11120-3)。

[0114] 本发明相关领域的技术人员理解术语“双抗体”和“三抗体”。存在包含重链可变结构域 (VH) 通过短肽接头与轻链可变结构域 (VL) 连接的分子,该短肽接头太短以至相同链上的两个结构域间无法配对。这促使与一条或多条其它链的互补结构域配对,并促进形成具有两个或多个功能性抗原结合位点的二聚体或三聚体分子。所得的抗体分子可以是单特异性或多特异性的 (如,双抗体的情况下具有双重特异性)。使用本发明涉及领域的标准方法,如 Todorovska 等 (在癌症靶向中双抗体、三抗体和四抗体的设计 and application of diabodies, triabodies and tetrabodies for cancer targeting). *J. Immunol. Methods.* 2001 年 2 月 1 日 ;248 (1-2) :47-66) 所述,从两种或多种抗体产生此类抗体分子。

[0115] 也可产生包含特异性结合位点的抗体片段。例如,此类片段包括但不限于:可通过胃蛋白酶消化抗体分子产生的 $F(ab')_2$ 片段,以及通过还原 $F(ab')_2$ 片段的二硫桥产生的 Fab 片段。或者,可构建 Fab 表达文库以快速简便的鉴定具有所需特异性的单克隆 Fab 片段 (Huse, W. D. 等 (1989) *Science* 254 :1275-1281)。

[0116] 可使用不同免疫实验来筛选鉴定具有结合特异性的抗体。本领域熟知使用具有确定特异性的多克隆或单克隆抗体进行竞争结合或免疫放射实验的多种实验方法。此类免疫实验通常包括测定多肽或多核苷酸和其特异性抗体间形成的复合物。优选使用与两个非干扰表位具有反应性的单克隆抗体的双位点单克隆免疫实验,但也可使用竞争性结合实验 (Maddox, 同上)。

[0117] 本文所述噬菌体多肽具有靶向、通透和 / 或抑制细胞的能力,还可作为载体分子用于将其它抑制剂分子递送入微生物细胞。将化合物偶联至氨基酸的化学方法已经得以很好的发展,可将多种不同的分子类型与多肽连接。最常见的偶联方法依赖于游离氨基 (α -氨基或 Lys)、巯基 (Cys) 或羧酸基团 (Asp、Glu 或 α -羧基) 的存在。可使用偶联方法通过羧基-或氨基-末端残基将多肽连接至细胞抑制剂。在一些情况下,序列包含多个可与所选化学物质进行反应的残基。可使用该方法生产包含多于一种细胞抑制剂的多聚物。

或者,可缩短或选择多肽,使反应性残基位于序列的氨基或羧基末端。

[0118] 例如,在多肽合成中可使用 N- α -Fmoc-N ϵ -1-(4,4-二甲基-2,6-二氧代环亚己-1-基-3-甲基丁基)-L-赖氨酸,将报告分子如荧光素特异性掺入到赖氨酸残基(Ono 等,1997)。合成后用胍处理去除 4,4-二甲基-2,6-二氧代环亚己-1-基,偶联 5-和 6-羧基荧光素琥珀酰亚胺酯。因此,通过多肽序列中包含赖氨酸残基,然后与合适的衍生化细胞抑制剂反应,能够完成抑制剂分子与噬菌体多肽的偶联。

[0119] 还可使用 EDC(盐酸 1-乙基-3-(3-二甲基氨基丙基)碳二亚胺)或碳二亚胺偶联方法。碳二亚胺可激活天门冬氨酸和谷氨酸的侧链羧基基团,以及羧基末端基团,使其成为伯胺偶联的反应性位点。将活化的多肽与细胞抑制剂混合以产生最终的偶联物。如果细胞抑制剂首先被活化,那么 EDC 法将通过 N 末端 α 胺并且可能通过 Lys 侧链中的胺(如果存在的话)偶联细胞抑制剂。

[0120] 间-马来酰亚胺基苯甲酰基-N-羟基琥珀酰亚胺酯(MBS)是异双功能试剂,可通过半胱氨酸将多肽连接于细胞抑制剂。半胱氨酸残基的巯基基团参与偶联。如果所选序列不包含 Cys,通常将 Cys 残基放置于 N-或 C-末端以便在高度控制条件下连接多肽与细胞抑制剂。出于合成的目的,将半胱氨酸置于多肽的 N 末端是有帮助的。MBS 尤其适合用于本发明。

[0121] 戊二醛可用作通过其氨基将两个化合物连接的双功能偶联试剂。戊二醛在多肽和细胞抑制剂之间提供了高度灵活的间隔物,以利于呈现。戊二醛是反应性非常高的化合物,与 Cys、Tyr 和 His 进行有限程度的反应。当多肽在其氨基末端仅包含一个游离氨基时,戊二醛偶联方法尤其有用。当多肽包含不止一个游离氨基时,可形成大型多聚复合物。

[0122] 在一个方面,本发明的多肽可融合于(如通过框内克隆)或连接于(如通过化学偶联)细胞抑制剂,如抗微生物试剂。其中包括抗微生物肽,例如杀菌/通透性增加蛋白、阳离子抗微生物蛋白、溶菌酶、乳铁蛋白和犬科抗菌肽(cathelicidins)(如来自中性粒细胞,参见例如 Hancock 和 Chapple,1999, Antimicrob. Agents Chemother. 43:1317-1323; Ganz 和 Lehrer,1997, Curr. Opin. Hematol. 4:53-58; Hancock 等,1995, Adv. Microb. Physiol. 37:135-175)。抗微生物肽还包括防御素(如,来自上皮细胞或中性粒细胞)和血小板杀微生物蛋白(参见例如, Hancock 和 Chapple,1999, Antimicrob. Agents Chemother. 43:1317-1323)。其它抗微生物肽包括但不限于:短杆菌肽 S、杆菌肽、多粘菌素 B、蛭素、白细胞抗菌肽(bactenecin)(如,牛白细胞抗菌肽),牛蛙皮肤抗菌肽(ranalexin)、天蚕素 A(cecropin A)、尹杜抗菌肽(indolicidin)(如牛尹杜抗菌肽)和乳酸链球菌素(如细菌乳酸链球菌素)。

[0123] 抗微生物试剂还包括离子载体(ionophores),它利于离子(如钠)的穿越脂质屏障如细胞膜进行传播。尤其适合本发明的两个离子载体化合物是 RUMENSINTM(礼来公司)(Eli Lilly)和拉沙里菌素(Lasalocid)(罗氏公司)(Hoffman LaRoche)。其它离子载体包括但不限于:盐霉素、阿伏霉素、阿瑞辛(aridcin)和阿克拉宁(actaplanin)。其他抗微生物剂包括青霉素、莫能菌素TM和阿奇霉素、甲硝唑、链霉素、卡那霉素和青霉素,以及 β -内酰胺类、氨基糖苷类、大环内酯类、氯霉素、新生霉素、利福平和氟喹诺酮(参见例如 Horn 等,2003, Applied Environ. Microbiol. 69:74-83; Eckburg 等,2003, Infection Immunity 71:591-596; Gijzen 等,1991, Applied Environ. Microbiol. 57:1630-1634; Bonelo 等,

1984, FEMS Microbiol. Lett. 21: 341-345; Huser 等, 1982, Arch. Microbiol. 132: 1-9; Hilpert 等, 1981, Zentbl. Bakteriол. Mikrobiol. Hyg. 1 Abt Orig. C 2: 21-31)。

[0124] 尤其有用的抑制剂是阻断或干扰甲烷生成的化合物, 包括溴代乙磺酸, 如, 2- 溴代乙磺酸 (BES) 或其盐, 例如钠盐。钼酸钠 (Mo) 是硫酸还原抑制剂, 可与溴代乙磺酸一起使用。其它抗甲烷生成化合物包括但不限于: 硝酸、甲酸、甲基氟、三氯甲烷、水合氯醛、亚硫酸钠、乙烯和不饱和烃、乙炔, 脂肪酸如亚油酸, 顺式油酸, 饱和脂肪酸, 如山萘酸 (behenic) 和硬脂酸, 以及陆马嗪 (lumazine) (例如, 2, 4- 二羟基蝶啶)。其它化合物包括 3- 溴丙磺酸盐 (BPS)、丙炔酸和 2- 丁炔酸乙酯。

[0125] 抗微生物剂还包括裂解酶, 包括噬菌体溶菌酶、内溶素、溶菌酶、溶素、噬菌体溶素、溶肠壁素、胞壁质酶和病毒溶素。有用的酶具有水解细菌细胞壁中特定键的能力。特定的裂解酶包括但不限于: 葡萄糖苷酶, 它水解肽聚糖中氨基糖 (如 N- 乙酰基胞壁酸和 N- 乙酰基葡萄糖胺) 间的糖苷键; 酰胺酶, 它切割多糖链和交联肽之间的 N- 乙酰胞壁酰-L- 丙氨酸的酰胺连接; 以及内肽酶, 它水解肽间连接 (如, 半胱氨酸内肽酶), 和内异肽酶, 它攻击来自甲烷杆菌科 (Methanobacteriaceae) 的产甲烷菌的假胞壁质。

[0126] 本文详述的 ORF 2058 或 ORF 2055 编码的多肽可用作瘤胃产甲烷菌特异性裂解酶。可以从新鲜 **qmru** 裂解的反刍甲烷短杆菌细胞中制备天然的酶。或者, 可将 ORF 2058 或 ORF 2055 克隆入表达载体并在异源宿主如大肠杆菌中表达。早前利用 PeiP 和 PeiW 完成这项工作, 并且在还原条件下该重组蛋白显示对抗甲烷嗜热杆菌 (Methanothermobacter) 细胞的活性 (Luo 等, 2002)。可将 ORF 2058 或 ORF 2055 裂解酶或任何其它裂解酶用于组合物, 例如, 作为反刍动物的饲料添加剂, 或者可将其掺入缓释胶囊或药丸装置用于在瘤胃中长时间递送。可将裂解酶与其它产甲烷菌抑制剂组合或依次使用以避免宿主产甲烷菌适应并对酶耐受。也可在酶中使用随机和 / 或靶向突变以避免适应。裂解 / 溶原开关组件 (如 ORF 1981 和 ORF 1983-ORF1986) 可以与裂解酶相似的方式使用。

[0127] 此外, PNA 也包括作为抗微生物试剂。PNA 是肽-核酸杂交物, 其中磷酸主链被获自 N-(2- 氨基乙基)- 甘氨酸单位的非手性中性主链取代 (参见例如优卡生物科学系列 (Eurekah Bioscience Collection). 人端粒酶的 PNA 和寡核苷酸抑制剂 (PNA and Oligonucleotide Inhibitors of Human Telomerase). G. Gavory 和 S. Balasubramanian, 兰登生物科学公司 (Landes Bioscience), 2003)。通过亚甲基连接将 A、G、T、C 碱基连接于主链的氨基氮上 (P. E. Nielsen 等, Science 1991. 254: 1497-1500; M. Egholm 等, Nature 1993. 365: 566-568)。PNA 以高特异性与互补序列结合, 并且相对类似的 DNA 或 RNA, PNA 具有更高亲合力 (M. Egholm 等, 同上)。与对应的 DNA/DNA 或 DNA/RNA 双链体相比, PNA/DNA 或 PNA/RNA 杂交物具有更高的热稳定性 (M. Egholm 等, 同上)。由于非天然酰胺主链不被核酶或蛋白酶识别, 因而 PNA 还具有高的化学和生物学稳定性 (V. Demidov 等, Biochem Pharmacol 1994. 48: 1310-1313)。通常 PNA 的长度至少为 5 个碱基, 并包含末端赖氨酸。可将 PNA PEG 化以进一步延长其寿命 (Nielsen, P. E. 等 (1993) Anticancer Drug Des. 8: 53-63)。

[0128] 在一个特定方面, 本发明的多肽可与细胞抑制剂如抗体或其片段融合 (如通过框架内克隆) 或连接 (如通过化学偶联)。抗体或抗体片段可靶向微生物细胞, 或者特别是产甲烷菌细胞, 或一种或多种细胞组件。例如, 可靶向细胞表面蛋白质, 如胞外受体。包括免

疫球蛋白分子和免疫球蛋白 (Ig) 分子的免疫活性部分,即包含特异性结合抗原(与其免疫反应)的抗原结合位点的分子。

[0129] 本发明的多肽在靶向微生物细胞,具体说是产甲烷菌细胞中尤其有用。在特定方面,可利用该多肽与细胞壁或膜连接或结合,通透细胞和/或抑制细胞的生长或复制。同样,可使用该多肽瞬时或长时间附着于细胞,或通透细胞壁或膜和/或在胞内环境中累积。应当理解的是,本发明的噬菌体多肽,以及对应的多核苷酸、表达载体、宿主细胞和抗体可用于靶向多种微生物,例如,反刍甲烷短杆菌和史氏甲烷短杆菌,前者是反刍动物中的主要产甲烷菌,后者是人体的主要产甲烷菌。为了实施靶向,可将微生物细胞与噬菌体多肽接触,如本文详述,所述多肽可分离自一种或多种天然资源,或产自表达载体和/或宿主细胞,或来自于合成或半合成化学方法。为了增强通透性,可将该多肽与一种或多种信号序列融合或连接(预测共有序列):

[0130] [ML]KKKK[K]{0,1}X{0,9}[IL][IFL][IL][IL][IS][LIA]X{0,4}[LIVF][LIAV][LI][ILV][LAIV][ILFV][LIVF][SAL][ILV][GSA][AS][VAI][SA]A(SEQ ID NO:151),参见图6)。还参见 Pérez-Bercoff, **A.**, Koch, J. 和 Bürglin, T. R. (2006) LogoBar: 利用柱状图观察蛋白质标识和缺口 (LogoBar: bar graph visualization of protein logos with gaps). *Bioinformatics* 22, 112-114。在特定方面,将该多肽以本文详述的组合物形式递送给对象,例如通过用于反刍动物的缓释装置递送。

[0131] 在某些实施方式中,将该多肽与细胞抑制剂融合或连接,所述细胞抑制剂包括例如,抗-产甲烷生成化合物(如溴代乙磺酸),抗体或抗体片段,裂解酶,肽核酸,抗微生物肽或其它抗生素。多肽抑制剂以组合物的形式递送给对象以抑制微生物细胞,具体是产甲烷菌细胞的生长和/或复制。该组合物包含,例如:a) 分离的噬菌体,噬菌体颗粒,噬菌体基因组或其变化形式、片段、变体或衍生物;b) 分离的噬菌体多肽或其变化形式、片段、变体或衍生物;c) 分离的多核苷酸或其变化形式、片段、变体或衍生物;d) 包含该多核苷酸的表达载体;或 e) 包含该表达载体的宿主细胞。根据公开的方法,本发明的组合物可专门包装为试剂盒的一部分,用于靶向、通透和/或抑制微生物细胞,尤其是产甲烷菌细胞。该试剂盒包括至少一种本文所列组合物;以及用于靶向或通透产甲烷菌或其它微生物细胞,或抑制其细胞生长或复制的使用说明。

[0132] 在另一实施方式中,本发明涉及上述任何方法所用的与药学上可接受载体联用的药物组合物。此类药物组合物可包含噬菌体多肽和细胞抑制剂的组合。或者,药物组合物可包含本文详述的表达载体或宿主细胞。可单独给予该组合物或与至少一种其它药剂,如稳定化合物联合给予,可利用任何无菌的生物相容性药学载体,包括但不限于:盐水、缓冲盐水、右旋糖和水给与该组合物。该组合物可单独给予对象或与其它药剂、药物(如抗微生物药物)或激素联合给予。

[0133] 除了活性成分,这些药物组合物还可包含合适的药学上可接受的载体,包括利于将活性化合物加工成药学可用制剂的赋形剂和辅助剂。在最新版 的雷明顿药物科学 (Remington's Pharmaceutical Sciences) (宾夕法尼亚州伊斯顿的麦克出版公司 (Maack Publishing Co., Easton, PA)) 可找到制剂和给药技术的更多细节。本发明使用的药物组合物可通过任何途径给予,包括但不限于:口服、静脉内、肌肉内、动脉内、髓内、鞘内、心室内、透皮、皮下、腹膜内、鼻内、肠内、局部、舌下或直肠方式。

[0134] 可使用本领域熟知的药学上可接受的载体,以口服给药合适剂量,配制用于口服给药的药物组合物。此类载体使药物组合物可以配制为片剂、丸剂、糖衣剂、胶囊剂、液体、凝胶剂、糖浆、浆液、混悬剂等,便于对象摄取。可通过以下方法获得口服用途的药物制剂:将活性化合物与固定赋形剂混合,任选将所得混合物研磨,在加入合适的辅助剂后将该混合物加工成颗粒,获得片剂或糖衣剂芯体。合适的赋形剂是糖或蛋白质填充物,如糖,包括乳糖、蔗糖、甘露醇或山梨糖醇;来自玉米、小麦、稻米、土豆或其它植物的淀粉;纤维素,如甲基纤维素、羟丙基甲基-纤维素或羧甲基纤维素钠;胶质物,包括阿拉伯胶和黄芪胶;以及蛋白质如明胶和胶原。需要的话,可加入崩解剂或增溶剂,如交联聚乙烯吡咯烷酮、琼脂、藻酸或其盐,如藻酸钠。

[0135] 可口服使用的其它药物制剂包括明胶制成的压接(push-fit)胶囊,以及明胶和包衣剂(coating)(如甘油或山梨糖醇)制成的密封软胶囊。压接胶囊可包含与填充剂或粘合剂,如乳糖或淀粉;润滑剂,如滑石粉或硬脂酸镁以及任选稳定剂混合的活性组分。

[0136] 在软胶囊中,活性化合物可溶解或悬浮于含有或不含稳定剂的合适液体,如脂肪油、液体或液体聚乙二醇中。

[0137] 糖衣剂芯体可与合适的包衣剂,如浓缩糖溶液联合使用,其中还可包含阿拉伯胶、滑石粉、聚乙烯吡咯烷酮、卡波姆凝胶、聚乙二醇和/或二氧化钛、漆溶液和合适的有机溶剂或溶剂混合物。可将染料或颜料加入片剂或糖衣剂包衣中,用于标示产品或表征活性化合物的量,即剂量。

[0138] 可用水性溶液,优选生理相容性缓冲液,如汉克斯溶液、林格溶液或生理缓冲盐水配制适合胃肠道外给药的药物制剂。水性注射悬液可包含增加悬液粘度的物质,如羧甲基纤维素钠、山梨糖醇或右旋糖苷。此外,活性化合物的悬液可制备为合适的油性注射悬液。合适的亲脂性溶剂或运载体包括脂肪油如芝麻油,或合成脂肪酸酯如油酸乙酯或甘油三酯,或脂质体。也可使用非脂质聚阳离子氨基酸聚合物进行递送。悬液可任选包含合适的稳定剂或增加化合物溶解度以制备高浓缩溶液的试剂。对于局部或鼻腔给药,在制剂中使用适合要渗透的特定屏障的渗透剂。本领域通常了解此类渗透剂。

[0139] 本发明的药物组合物可以本领域所知方式制造,如,通过常规的混合、溶解、造粒、制造糖衣剂、水飞、乳化、包胶、包封或冻干工艺的方法。可以盐的形式提供药物组合物,所述盐可用多种酸形成,包括但不限于:盐酸、硫酸、乙酸、乳酸、酒石酸、苹果酸、琥珀酸等。在水性或其它质子溶剂中盐比对应的游离碱形式更易于溶解。在其它情况下,优选制剂可以是包含下列任何或所有成分的冻干粉末:1-50mM 组氨酸、0.1%-2%蔗糖和2-7%甘露醇,pH范围为4.5-5.5,在使用前与缓冲液组合。药物组合物制备后,可将其放置于合适的容器中,并贴上用于治疗所示病症的标签。对于本发明组合物的给药,此标签可包含给药量、频率和方法。

[0140] 适用于本发明的药物组合物包括包含有效量活性组分以达到预期目的的组合物。对于任何化合物,最初可在细胞实验中,如微生物细胞或具体说,在产甲烷菌细胞中,或在动物模型中,通常是小鼠、兔、狗或猪或反刍动物如绵羊、牛、鹿或山羊中估计治疗有效剂量。还可使用动物模型确定合适的浓度范围和给药途径。此类信息还可用于确定有用的给药剂量或途径。通常给药剂量可以是0.1-100,000微克,甚至总剂量约为1g或更多,这取决于给药途径。文献中提供了有关具体剂量和递送方法的指南,本领域实施人员可获得这

些指南。与递送多肽相比,本领域熟练技术人员将使用不同制剂递送多核苷酸。相似地,多核苷酸或多肽的递送对于特定细胞、病症、位置等而言是不同的。

[0141] 本领域已知基于噬菌体的疗法,并且已公开此类组合物的制备方法。已描述噬菌体疗法可用于靶向葡萄球菌 (*Staphylococcus*) (如金黄色葡萄球菌 (*S. aureus*))、假单胞菌 (*Pseudomonas*) (如绿脓假单胞菌 (*P. aeruginosa*))、大肠杆菌 (*Escherichia*) (如大肠杆菌 (*E. coli*))、克雷伯菌 (*Klebsiella*) (如臭鼻克雷伯菌 (*K. ozaenae*))、鼻硬结克雷伯氏菌 (*K. rhinoscleromatis scleromatis*) 和肺炎克雷伯菌 (*K. pneumonia*))、变形杆菌 (*Proteus*)、沙门氏菌 (*Salmonella*)、志贺氏菌 (*Shigella*) (参见例如 Carlton, R. M. (1999). *Archivum Immunologiae et Therapiae Experimentalis*, 47:267-274; Liu, J. 等 (2004). *Nat. Biotechnol.* 22, 185-191; Projan, S. (2004). *Nat. Biotechnol.* 22, 167-168; Sulakvelidze, A., Alavidze, Z. 和 Morris, J. G. (2001). *Antimicrobial Agents and Chemotherapy*, 45(3):649-659; Weber-Dabrowska, Mulczyk, M. 和 Gorski, A. (2000). *Archivum Immunologiae et Therapiae Experimentalis*, 48:547-551)。噬菌体疗法比传统抗微生物剂具有先天的优势,因为噬菌体高度特异,不影响人体中正常微生物;噬菌体不感染真核细胞,不具有已知的严重副作用;噬菌体可定位于感染位置;并且噬菌体可以指数级复制,因此这种治疗仅需小剂量,成本通常很低(参见例如 Sulakvelidze 等,同上)。关于当前的综述,参见 Fischetti VA, Nelson D, Schuch R. 重新使用噬菌体疗法:部分大于总和? (Reinventing phage therapy:are the parts greater than the sum?) *Nat Biotechnol.* 2006 年 12 月;24(12):1508-11。

[0142] 基于肽和多肽的疗法已作描述,例如白介素融合毒素 (denileukindifitox)、奥曲肽、伐普肽、兰瑞肽、RC-3940 系列肽、达必佳、醋酸亮丙瑞林 (lupron)、瑞林、西曲瑞克(参见例如, Lu 等, 2006, *AAPS J* 8:E466-472), 亥莫西汀 (hemocidins), 司他弗平 (staphopains) (参见例如, Dubin 等, 2005, *Acta Biochemica Polonica*, 52:633-638), 以及吲哚西汀 (indolicidin), 防御素, 羊毛硫抗生素, 微西汀 B17 (microcidin B17), 富组蛋白 (histatins) 和吗加宁 (maganin) (参见例如, Yeaman 和 Yount, 2003, *Pharmacol Rev* 55:27-55)。在 Degim 等, 2007, *Curr Pharm Des* 13:99-117 和 Shai 等, 2006, *Curr Prot Pept Sci*, 7:479-486 中能够找到肽和多肽疗法的一般性指南。最近批准的基于肽的药物包括 Hematide™ (合成的基于肽的红细胞生成刺激剂, 埃非麦克司公司 (Affymax, Inc.)), 艾塞那肽 (Exenatide) (合成的肠促胰岛素类似物 (exendin)-4, 爱麦林/礼来公司 (Amylin/Eli Lilly)), Natreacor (奈西立肽 (nesiritide), 奈西立肽, 司可奥斯公司 (Scios)), 普来纳西 (Plenaxis) (阿巴瑞克, 普瑞西斯制药公司 (Praecis Pharmaceuticals)) 和 SecreFlo (促胰液素, 瑞普里根公司 (Repligen))。

[0143] 实施者根据治疗对象相关的因素确定准确剂量。对剂量和给药进行调整以提供足够水平的活性试剂或维持所需效果。需要考虑的因素包括疾病状况的严重性、对象的总体健康状况、对象的年龄、体重和性别、饮食、给药的时间和频率、联合用药、对治疗的反应敏感性以及耐受/响应。根据特定制剂的半衰期和清除率, 长效药物组合物可每 3-4 天, 每周或每两周给予一次。

[0144] 本发明组合物(如药物组合物)尤其有用的是缓释配方或原理。例如, 瘤胃内装置包括但不限于: 新西兰埃格瑞饲料公司 (Agri-Feeds Ltd., New Zealand) 时间胶囊™丸

范围,最初由新西兰爱吉研究公司开发 (AgResearchLtd., New Zealand),如 WO 95/19763 和 NZ 278977 所公开的,以及新西兰奥克兰的纽法姆公司的分支机构纽法姆健康和科学公司的 CAPTEC (Nufarm Health&Sciences, a division of Nufarm Ltd., Auckland, New Zealand),如 AU 35908178, PCT/AU81/100082 和 Laby 等,1984, Can. J. Anim. Sci. 64 (增刊) 337-8 所公开的,将所有这些参考文献纳入本文作参考。作为特定的例子,该装置可包括弹簧和柱塞,迫使组合物穿过同末端的孔。

[0145] 作为另一实施方式,本发明涉及作为水补充物的组合物如进水组合物,和食物补充剂如反刍动物饲料组分,用于上述任何方法。在特定方面,食物补充剂包含至少一种可食用的植物材料以及本发明的肽或多肽。或者,食物补充剂包含至少一种可食用的植物材料以及本文公开的肽或多肽,或编码本文公开的肽或多肽的多核苷酸,例如,表达载体或含有表达载体的宿主细胞的形式。具体说,组合物还包含与所得序列融合或连接的细胞抑制剂。优选的植物材料包括干草、草、谷物或粗磨粉中的任一种,例如,豆科干草、禾本科干草、玉米青贮、禾本青贮、豆科青贮、玉米粒、燕麦、大麦、酿酒谷物、啤酒谷物、大豆粗磨粉和棉籽粗磨粉。具体说,禾本青贮可用作反刍动物的食物组合物。可对植物材料进行遗传改造,使其包含本发明的一种或多种组分,如一种或多种多肽或肽、多核苷酸或载体。

[0146] 在另一实施方式中,本发明多肽或多核苷酸的抗体可在监测微生物水平的实验种用于测定该微生物,尤其是产甲烷菌的存在。同上所述可以相同的方法制备用于诊断目的的抗体。诊断实验包括使用抗体和标签检测人 体体液或细胞或组织提取物中多肽的方法。可使用修饰或未修饰的抗体,通过共价或非共价方式将报告分子与其结合进行标记。可使用本领域所知的多种不同的报告分子,上文对其中几个进行了描述。

[0147] 本领域已知用于测量多肽或多核苷酸水平的多种实验方案 (如 ELISA、RIA、FACS 和印迹法,如 Southern、Northern、Western 印迹),为测定微生物尤其是产甲烷菌的存在或水平提供了基础。通过在适合形成复合物的条件下将来自正常对象如正常人或反刍动物的体液或细胞提取物与抗体组合,从而确定正常或标准水平。利用各种方法可对标准复合物的形成量进行定量,但优选分光光度法。将对象、对照和治疗样品 (如来自治疗对象的样品) 中表达的多肽或多核苷酸的量与标准值进行比较。标准值和对象值之间的偏差确定了用于测定微生物存在或水平的参数。

[0148] 在本发明的一个特定实施方式中,通过使用特定杂交和 / 或扩增技术,可将多核苷酸用于诊断目的。可使用的多核苷酸包括寡核苷酸、互补 RNA 和 DNA 分子和 PNA。多核苷酸可用于检测或定量测定样品中的基因表达,其中表达与微生物的存在或水平相关。可使用诊断实验区分微生物水平的不存在、存在和改变,并在治疗干预中监测水平。

[0149] 在一个方面,可利用与 PCR 探针的杂交鉴定核酸序列,尤其是基因组序列,其编码本发明的多肽。无论探针产生于高度特异性的区域,如 5' 调节区的 10 个独特核苷酸,或者较低特异性的区域,如 3' 编码区,探针的特异性以及杂交或扩增的严谨性 (最大、高、中等或低) 将决定该探针是否仅仅识别天然产生的序列、等位基因或相关序列。探针也可用于检测相关序列,并应该优选包含来自任何编码序列的至少 50% 核苷酸。本发明的杂交探针可以是 DNA 或 RNA,并来源于核苷酸序列 SEQ ID NO:74-142 或其互补或修饰序列,或来自于包含天然产生序列的启动子、增强元件和内含子的基因组序列。

[0150] 用于产生特异性 DNA 杂交探针的方法包括将核酸序列克隆入载体以产生 mRNA 探

针。本领域了解此类载体,它们可购得,并且通过加入合适的 RNA 聚合酶和合适的标记核苷酸可将它们用于体外合成 RNA 探针。可通过多种报告基团对杂交探针进行标记,这些报告基团例如放射性核素如 ^{32}P 或 ^{35}S ,或酶标记,如通过亲和素/生物素偶联系统偶联至探针的碱性磷酸酶等。多核苷酸可用于 Southern 或 northern 分析,点印迹或其它基于膜的技术;用于 PCR 技术;或者用于试纸条、狭缝、ELISA 实验或微阵列,用来自对象活检的液体或组织检测微生物的存在和水平。本领域熟知此类定性或定量的方法。

[0151] 在一个特定方面,核酸序列可用于各种标准方法标记的实验,并在适合杂交和/或扩增的条件下加入来自对象的液体或组织。孵育合适的时间后,洗涤样品,对信号进行定量并与标准值比较。如果受试样品的信号量与可比较的对照样品的信号量相比发生了显著的改变,那么样品中出现核苷酸序列水平的改变表明微生物的出现或水平。还可使用此类实验评价特定治疗方案在动物研究、临床实验或者监测对象治疗中的效果。

[0152] 为了提供微生物出现或水平的诊断基础,需确定表达的正常或标准特征。通过在适合杂交和/或扩增的条件下,将来自正常对象的体液或细胞提取物与多核苷酸或其片段混合,来完成这项实验。通过比较获自正常对象的值与获自使用已知量基本纯化多核苷酸的实验中的值,可对标准水平进行定量。可将来自正常样品的标准值与来自针对微生物生长进行治疗的对象的值进行比较。利用标准值和对象值之间的偏差测定微生物的存在或水平。

[0153] 一旦鉴定了微生物并开始了治疗方案,则在常规基础上进行重复的杂交和/或扩增实验以评价相对于正常对象中观察到的表达水平,对象中的表达水平是否开始降低。来自连续实验的结果可用于显示从数天到数月的治疗期间的效果。

[0154] 由核酸序列设计的寡核苷酸的特定诊断应用可包括使用 PCR。此类寡聚体可以是化学合成、酶学方法产生或体外产生的。寡聚体优选由两个核苷酸序列组成,一个为正义方向 ($5' \rightarrow 3'$),另一个为反义方向 ($3' \rightarrow 5'$),在为鉴定具体核苷酸序列或条件的优化条件下使用。在用于检测和/或定量测定紧密相关的 DNA 或 RNA 序列的较低严谨性条件下,使用相同的两种寡聚物,巢式寡聚物集合或寡聚物简并库。

[0155] 可用于定量测定表达的方法包括放射性标记或生物素化核苷酸,对照 核酸共扩增以及可通过插值获得实验结果插值的标准曲线 (Melby, P. C. 等 (1993) *J. Immunol. Methods*, 159 :235-244 ;Duplaa, C. 等 (1993) *Anal. Biochem.* 229-236)。可通过以 ELISA 形式进行实验加快多个样品的定量速度,其中感兴趣的寡聚物以各种稀释度出现,通过分光光度法或比色法进行快速定量。

[0156] 在另外的实施方式中,源自本发明所述任何多核苷酸的寡核苷酸或更长的片段可用作微阵列中的靶点。可使用微阵列同时监测大量基因的表达水平(产生转录物图像),并鉴定遗传变体、突变和多态性。可使用该信息确定基因功能,了解疾病的遗传学基础,诊断疾病,开发并监测治疗剂活性。在一个实施方式中,根据本领域所知方法制备和使用微阵列,如 PCT 申请 WO 95/11995(Chee 等), Lockhart, D. J. 等 (1996 ;*Nat. Biotech.* 14 : 1675-1680) 以及 Schena, M. 等 (1996 ;*Proc. Natl. Acad. Sci.* 93 :10614-10619) 所述。

[0157] 在一个方面,可使用化学偶联步骤和喷墨打印应用设备,如 PCT 申请 WO 95/251116(Baldeschweiler 等)所述在微阵列表面合成寡核苷酸。在另一方面,可使用类似于斑点或狭缝印迹的“网状”阵列 (HYBRIDOT 设备,生命科技公司 (Life

Technologies)), 利用真空系统、热、UV、机械或化学连接步骤将 cDNA 片段或寡核苷酸安置和连接至基材表面。在另一方面, 可手工或利用可获得的设备、材料或机器 (包括多道移液器或机器人仪器; 布林克曼公司, 韦斯特伯里, 纽约州 (Brinkmann, Westbury, N. Y.)) 生产阵列, 该阵列可包括例如 24、48、96、384、1024、1536 或 6144 或更多个点或孔 (如多孔板), 或从 2-1,000,000 的任何倍数, 以便充分利用市售仪器。

[0158] 为了用微阵列进行样品分析, 从生物制品中抽提多核苷酸。可从任何体液 (血液、尿液、唾液、痰、胃液等), 培养细胞, 活检物或其它组织制备物中获取生物样品。为了生产探针, 使用抽提自样品的多核苷酸产生与阵列上核酸互补的核酸序列。如果微阵列由 cDNA 组成, 则反义 RNA 是合适的探针。因此, 在一个方面, 使用 mRNA 产生 cDNA, 进而在荧光核苷酸存在下用 cDNA 产生片段或反义 RNA 探针。将这些荧光标记的探针与微阵列孵育, 以使探针序列与微阵列的 cDNA 寡核苷酸杂交。在另一方面, 用作探针的核酸序列可包含使用杂交技术领域熟知的限制酶、PCR 技术和寡聚物标记试剂盒 (安法马西亚生物技术公司 (Amersham Pharmacia Biotech)) 产生的多核苷酸、片段和互补或反义序列。

[0159] 在本发明的另一实施方式中, 可利用本发明的多肽或其功能性或免疫原性片段或寡肽通过任何药物筛选技术筛选化合物文库。此类筛选所用片段可在溶液中游离, 附加于固体支持物、携带于细胞表面或定位于细胞内部。可检测多肽和受试物间的结合复合物的形成。

[0160] 公开的 PCT 申请 WO 84/03564 描述了一种药物筛选技术, 它可用于高通量筛选对感兴趣多肽具有合适亲和力的化合物。在该方法中, 在固体基质如塑料针或一些其它表面上合成大量不同的小受试化合物。受试化合物与多肽或其片段反应, 然后洗涤。然后利用本领域熟知方法检测结合的多肽。还可将纯化的多肽直接包被在板上, 用于上述药物筛选技术。或者, 可使用非中和抗体捕获多肽并将其固定到固体支持物上。

[0161] 在另一项技术中, 可使用竞争性药物筛选实验, 其中能够结合多肽的中和抗体与受试化合物特异性竞争结合该多肽。以此方式, 可使用该抗体检测与该抗体具有相同的一个或多个抗原结合位点的受试化合物的存在。

实施例

[0162] 本文所述实施例用以解释本发明的实施方式。其它实施方式、方法和分析类型在分子诊断领域的一般技术人员技术范围内, 无需在此赘述。其它本领域范围内的实施方式也应作为本发明的一部分。

[0163] 实施例 1: 基因组大小估测

[0164] 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*) 菌株 M1^T (DSM1093) 生长于 BY+ 培养基中 (基本培养基, Joblin 等, 1990), 该培养基由 [g/l] NaCl (1)、KH₂PO₄ (0.5)、(NH₄)₂SO₄ (0.25)、CaCl₂·2H₂O (0.13)、MgSO₄·7H₂O (0.2)、K₂HPO₄ (1)、澄清反刍液 (300ml)、dH₂O (360ml)、NaHCO₃ (5)、刃天青 (0.2ml)、L- 盐酸半胱氨酸 (0.5)、酵母提取物 (2) 以及巴赫 (Balch) 微量元素溶液 (10ml) (加入微量元素; Balch 等, 1979) 组成, 由 (g/l) 三乙酸腈 (1.5)、MgSO₄·7H₂O (3)、MnSO₄·H₂O (0.5)、NaCl (1)、FeSO₄·7H₂O (0.1)、CoCl₂·6H₂O (0.1)、CaCl₂ (0.1)、ZnSO₄·7H₂O (0.1)、CuSO₄·5H₂O (0.01)、AlK(SO₄)₂·12H₂O (0.01)、H₃BO₃ (0.01)、Na₂MoO₄·2H₂O (0.01)、NiSO₄·6H₂O (0.03)、Na₂SeO₃ (0.02) 和 Na₂WO₄·2H₂O (0.02) 组成。在液氮

中冷冻细胞团并用预冷的无菌臼和杵研磨,抽提基因组 DNA。将细胞匀浆液包埋于琼脂糖塞中,后续操作在栓中进行以减少对基因组 DNA 的物理剪切。用限制性核酸内切酶进行消化,用脉冲场电泳 (PFGE) 分离 DNA 片段。

[0165] 实施例 2 :DNA 克隆和测序

[0166] 安进 科 生 物 科 学 公 司 (美 国 麻 省) (Agencourt Biosciences Corporation (Massachusetts, USA)) 使用随机鸟枪克隆方案 (Fleischmann 等, 1995), 马克基因公司 (美国马里兰州洛克维尔) (Macrogen Corporation (Rockville, MD, USA)) 使用焦磷酸测序对反刍甲烷短杆菌的基因组 DNA 进行测序。简单的说, 通过随机物理破坏基因组 DNA 和电泳分离片段在大肠杆菌中构建反刍甲烷短杆菌的 DNA 文库。从凝胶中回收 40Kb 范围内的大片段, 并用于产生大的插入 f 粘粒文库。回收 2-4kb 范围的 DNA 片段并用于产生小的插入质粒文库。培养大和小的插入文库所得的克隆, 回收其 f 粘粒或质粒 DNA 并用高通量测序技术进行测序。对足够数量的克隆进行测序以达到对反刍甲烷短杆菌基因组理论上 8 倍的覆盖。在随机剪切的基因组 DNA 片段上进行焦磷酸测序, 并达到理论上 10 倍的覆盖。

[0167] 实施例 3 :序列组装和原噬菌体注释

[0168] 比对 DNA 序列以找到交叠序列, 并利用帕拉赛基因组组装者 (Paracel Genome Assembler) (帕拉赛公司, 加利福尼亚州, 美国) (Paracel Inc, CA, USA) 和 Staden 软件包 (Staden 等, 1998), 与来自标准和反向 PCR 的序列组合组装成毗连序列 (毗连群)。使用开放阅读框 (ORF) 寻找器 GLIMMER (基因定位插值马尔可夫模型 ER) (Gene Locator Interpolated Markov Model ER, Delcher 等, 1999) 分析毗连群, 利用缺口 BLASTP (基础本地比对搜索工具) (Basic Local Alignment Search Tool (Altschul 等, 1997) 在国家生物技术信息中心 (National Center for Biotechnology Information) (NCBI) 的非冗余核苷酸和蛋白质数据库中对每一 ORF 进行分析。通过以随机方式人工连接来自 8 倍草图噬菌体序列的毗连群产生“假分子”, 并提交给基因组研究所 (The Institute for Genomic Research) (TIGR, 美国华盛顿特区) (TIGR, DC, USA) 进行自动注释。用 GLIMMER 对来自 10 倍焦磷酸测序组装的毗连群再次进行分析, 并且用 GAMOLA (复杂现场 Blast DNA 序列全局注释 (Global Annotation of Multiplexed On-site Blasted DNA sequences); Altermann 和 Klaenhammer, 2003) 对 ORF 进行自动注释。用直系同源蛋白聚类 (COG) 数据库根据功能对 ORF 进行分类 (阈值 $1e-02$) (<http://www.pnas.org/cgi/content/full/102/11/3906>; Tatusov 等, 2001)。

[0169] 分别利用全局和局部比对 (<http://pfam.wustl.edu>) 以及标准和片段模式 TIGRFAM HMM 模型 (<http://www.tigr.org/TIGRFAMs>) (阈值 $1e-02$), 使用 PFAM HMM 和 TIGRFAM 文库通过 HMMER (<http://hmmer.wustl.edu>) 对蛋白基序进行确定。用 TRNASCAN-SE 鉴定 tRNA (Lowe 和 Eddy, 1997), 用 KODON 软件包 (应用数学公司, 美国得克萨斯州奥斯汀) (Applied Maths, Austin, TX, USA) 和 REPUTER (Kurtz 和 Schleiermacher, 1999) 鉴定核苷酸重复。随后对自动注释进行人工校验。使用 GENEWIZ (Jensen 等, 1999) 构建可视化基因组图集, 用定制的内部开发算法产生潜在的数据结构。利用内部开发的软件 (Pathway Voyager; Altermann 和 Klaenhammer, 2005), 从预测的反刍甲烷短杆菌 ORF 组和 KEGG (京都基因与基因组百科全书) (Kyoto Encyclopedia of Genes and Genomes, Kanehisa 等, 2004) 在线数

数据库重建通路。

[0170] 实施例 4 :测序结果和分析

[0171] 通过限制性酶消化基因组 DNA 和通过 PFGE 测定片段大小进行反刍甲烷短杆菌基因组大小的估测,表明单条染色体大小约 2.5-2.9Mb。大和小插入克隆的初始测序(6 倍草图覆盖)以及将序列组装为毗连群表明基因组中 40Kb 的区域出现频率非常高(over-represented)(> 20 倍),特别是在小插入物文库中。可能的原因是在用于 DNA 提取的培养物的生长过程中,高拷贝数的质粒(尽管没有鉴定到染色体外 DNA)或溶源性细菌噬菌体复制。因为这种巨大序列偏向,所以仅对小插入克隆进行再次测序(覆盖 2 倍理论基因组),从 Sanger 测序中产生最终的 8 倍覆盖。将 8 倍草图阶段(phase)序列组装为通过 105 个支架连接的 756 个毗连群。进一步进行焦磷酸测序,以实现额外约 10 倍的覆盖,将这些序列掺入组装物使毗连群的数目降至 27。使用反向和长范围 PCR 技术的后续缺口连接过程使毗连群数目降至 14,保留一个错误组装。

[0172] 在高通量测序阶段,观察到序列倾向于覆盖与低 G+C 区(~ 12 Kb)紧紧相邻的 G+C 含量显著较高区域(~ 50 Kb)的方向。通过 GAMOLA 和 GeneWiz 对基因组序列进行分析发现了与巨大低 -GC 穗(spikes)紧邻的显著的高 -GC 区。对高 G+C 区的详细分析揭示了存在与噬菌体相关整合酶、噬菌体末端酶大亚基、噬菌体门户蛋白、噬菌体衣壳蛋白以及预测的用作噬菌体细胞溶素的肽酶相似的基因产物(图 3)。这些基因产物用作预测的反刍甲烷短杆菌原噬菌体(记做 ϕ mru)的整体结构的锚定位点。基于 DNA 二级结构的分析,鉴定出可能的噬菌体整合位点 attL 和 attR(图 1A)。att 位点的噬菌体整合似乎破坏了 ORF 1980 和 2069 编码的推定膜蛋白,并且该基因可能具有 ϕ mru 噬菌体基因组的原始整合位点 attB。

[0173] 基于公认的噬菌体基因组模块化结构以及与序列相似性和功能数据库,确定 ϕ mru 的一般结构(图 1B)和 DNA 序列(图 4A)。参见例如,Altermann E, Klein JR, Henrich B. 格氏乳杆菌温和噬菌体 (ϕ)adh 基因组的一级结构和特征 (Primary structure and features of the genome of the *Lactobacillus gasseri* temperate bacteriophage (ϕ)adh). *Gene*. 1999 年 8 月 20 日 ;236(2) :333-46 ;Desiere F, Lucchini S, Canchaya C, Ventura M, Brüssow H. Antonie Van Leeuwenhoek. 乳酸细菌中的噬菌体和原噬菌体的比较基因组学 (Comparative genomics of phages and prophages in lactic acid bacteria). 2002 年 8 月 ;82(1-4) :73-91。成功将预测的 ϕ mru 噬菌体 ORF 组分类为编码噬菌体整合、DNA 复制和包装、噬菌体结构蛋白以及裂解盒的模块,并且对约 40% 噬菌体 ORF 进行了功能表征。与大量直接或非直接重复侧接并且确定位于 DNA 复制模块内部的非编码区中的终止子样结构(244bp)被表征为推定的 DNA 复制起点。经预测,噬菌体基因组序列内部的数个基因位于反义链上,它们与低 -GC 区同时出现。有待确定的是这些基因是否使噬菌体功能失活或表明噬菌体基因组内的错误组装。

[0174] 发现预测的噬菌体溶素和 attR 之间的低 -GC 区具有 DNA 硫修饰系统, dnd (电泳期间降解),包括 II 型限制性 m6 腺嘌呤 DNA 甲基转移酶和可能是 dnd 系统特异性的转录调节子。而且,在噬菌体基因组内部或者与噬菌体基因组侧接处鉴定到非编码 RNA 结构。在预测的 DNA 复制模块内部,鉴定到 rbcL。rbcL 代表来自莱茵衣藻 (*Chlamydomonas reinhardtii*) 的 5' TRRNA 稳定元件。该家族被认为参与了 rbcL 基因的稳定, rbcL 基因编码 1,5-二磷酸核酮糖羧化酶的大亚基。该家族的突变可导致转录物降解加速 50 倍。

[0175] 鉴定出与噬菌体基因组侧接的三组 I 型内含子结构。I 组催化内含子是大的自我剪切核酶。在广泛的有机体中,它们催化自身从 mRNA、tRNA 和 rRNA 前体上剪切。核心二级结构由 9 个配对区域组成 (P1-P9)。它们主要折叠成为两个结构域 -P4-P6 结构域 (由堆叠的 P5、P4、P6 和 P6a 螺旋组成) 和 P3-P9 结构域 (由 P8、P3、P7 和 P9 螺旋组成)。该家族的标注 (mark-up) 二级结构仅代表该保守核心。I 组催化内含子通常具有插入环区域的长 ORF。这些非编码 RNA 结构位于 ORF 1980 (SEQ ID NO :74) 上游, ORF2065 (SEQ ID NO :141) 和 attR 的下游以及 ORF 2069 (SEQ ID NO :142) 上游的非编码区。

[0176] 实施例 5A :噬菌体基因

[0177] 在反刍甲烷短杆菌基因组序列内发现原噬菌体的序列是预料之外的。此前未有关于反刍甲烷短杆菌菌株 M1 (DSM 1093) 对裂解或溶原性噬菌体易感性的报道,尽管有报道称鉴定出其它甲烷短杆菌种的噬菌体 (Baresi 和 Bertani, 1984 ;Knox 和 Harris, 1986)。 ϕ mrur 原噬菌体序列的 G+C 含量显著高于周围反刍甲烷短杆菌基因组,表明它起源于另一有机体。观察到的同源性水平未表明其起源的明显宿主,这表明 ϕ mrur 不同于迄今遇见的其它任何噬菌体。

[0178] ϕ mrur DNA 序列插入到预测的反刍甲烷短杆菌膜蛋白内,并侧接于具有与 attL 和 attR 位点相一致的二级结构的 DNA 序列。尽管与其它已知蛋白质缺少强烈同源性,但在 ϕ mrur 内部可鉴定出噬菌体的所有功能性特征模块。该序列的一个有趣特征是 3' 末端存在低 G+C 区,这与参与 DNA 硫修饰系统 (dnd) 的蛋白质显示出同源性。这些基因位于 ϕ mrur 的 attR 依附位点上游,因此可能在噬菌体整合过程中被带进反刍甲烷短杆菌基因组中。该区域编码多种 dnd 相关的 ORF (dnd1、2 和 3),以及 II 型甲基酶亚基和推定的转录调节子。

[0179] dnd 表型促使其 DNA 在电泳过程中降解。对各 dnd 的 ORF 功能进行分析表明宿主基因组中掺入了硫或含硫物质。还发现 Dnd 表型存在于不同来源和不同栖息地的各种细菌种类的 DNA 中。在代表不同属的多种细菌基因组和海洋生物的 eDNA 中,发现组织相似的基因簇,这表明这种修饰是一种广泛的现象。体内 (35)S 标记实验显示 Dnd 表型和 DNA 硫修饰同时发生于多种代表性细菌的基因组中 (Zhou X, He X, Liang J, Li A, Xu T, Kieser T, Helmann JD, Deng Z. 新型 DNA 硫修饰 (A novel DNA modification by sulphur). Mol Microbiol. 2005 年 9 月 ;57 (5) :1428-38)。

[0180] II 型 R\M 系统是最简单最流行的。甲基转移酶和核酸内切酶编码为两个独立蛋白质并独立作用,而非以复合物的形式工作。没有特异性蛋白质。两种蛋白识别相同的识别位点因此活性竞争。甲基转移酶以单体形式作用,一次对双连体中的一条链甲基化。核酸内切酶以同源二聚体形式作用,促进两条链的切割。切割在识别序列附近或内部的预定位点发生。在这一点上,目前尚不清楚预测的功能如何与 dnd 系统共同作用。但是,清楚的是噬菌体能够用作基因递送运载体。具体说,能够利用溶原性转化区作为基因置换的基因座。

[0181] dnd 系统可能被噬菌体转运至反刍甲烷短杆菌中。同样,尚不了解 ϕ mrur dnd 系统在保护或修饰反刍甲烷短杆菌或外来 DNA 中的作用。 ϕ mrur 序列的另一个有趣的特性是反义链上编码的 ORF 数目。这些 ORF 对应低 GC 区并与来自多种有机体的蛋白质具有低 BLAST 吻合度。这表明在整合入反刍甲烷短杆菌后,这些基因在 ϕ mrur 基因组中累积。尚不清楚这些 ORF 是否代表插入正在累积,并最终导致噬菌体失活和驯化,或 ϕ mrur 是否完全激活。

[0182] 对于减少甲烷特别重要的一个 ϕ mrur 基因是位于裂解盒中的 ORF2058。ORF

2058 被注释为肽酶,并与肽酶 C39 家族蛋白具有蛋白家族 (Pfam) 匹配性 (Protein Family (Pfam)match) (评分: -13.7, E 值: 0.00054)。这些蛋白质是半胱氨酸肽酶,并且是 MEROPS 肽酶数据库定义的更大 CA 肽酶家族的一部分 (Rawlings 等,2006)。C39 肽酶家族常常与 ABC 转运蛋白相关联,在细菌素输出和加工中用作成熟蛋白酶。CA 肽酶家族还包括病毒半胱氨酸内肽酶,如切割古菌细胞壁交联肽的 C71 古菌噬菌体内异肽酶。属于甲烷杆菌科的产甲烷古菌的细胞壁包含假胞壁质的平行链,一种被肽交联的 N-乙酰基-L-塔罗糖胺糖醛酸 (N-acetyl-L-talosaminurinic acid) 的聚合物。C71 假胞壁质内异肽酶能够切割古菌细胞壁肽交联并能裂解细胞。

[0183] 基于来自马尔堡甲烷嗜热杆菌 (*Methanothermobacter marburgensis*) 噬菌体 Ψ M2 (图 3) 的假胞壁质内异肽酶的位置和同线性,ORF 2058 可具有产甲烷菌溶素基因的角色,其编码参与子代噬菌体释放前的细胞裂解的裂解酶。将 ORF 2058 与来自马尔堡甲烷杆菌 (*M. marburgensis*) 的 PeiP 和来自沃氏甲烷杆菌 (*M. wolfeii*) 的 PeiW (图 5) 进行比对,显示这些蛋白质间的总体同源性低。然而,参与内异肽酶催化三联体的组氨酸和天冬氨酸残基具有保守性,另外 ORF 2058 的半胱氨酸残基位置靠近 PeiP 和 PeiW 的保守半胱氨酸,它形成了催化三联体的第三个保守位点 (Makarova 等,1999, Luo 等,2002)。而且,在 ORF 2058 中也找到了环绕 PeiP 和 PeiW 中催化性 His 残基的 Gly-His-Tyr 基序。这些发现表明 ORF 2058 是 ϕ mru 溶菌素基因,功能是在噬菌体裂解循环中裂解反刍甲烷短杆菌细胞。ORF 2058 以及 PeiP 和 PeiW 间的差异可能反应了不同的古菌细胞壁肽交联,因此反应了不同的肽酶底物特异性。

[0184] 实施例 5B:噬菌体诱导

[0185] 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*) 菌株 M1^T (DSM1093) 生长于 BY+ 培养基中 (基本培养基, Joblin 等,1990),该培养基由 [g/l] NaCl (1)、KH₂PO₄ (0.5)、(NH₄)₂SO₄ (0.25)、CaCl₂·2H₂O (0.13)、MgSO₄·7H₂O (0.2)、K₂HPO₄ (1)、澄清反刍液 (300ml)、dH₂O (360ml)、NaHCO₃ (5)、刃天青 (0.2ml)、L-盐酸半胱氨酸 (0.5)、酵母提取物 (2) 以及巴赫微量元素溶液 (10ml) (加入微量元素; Balch 等,1979) 组成,由 (g/l) 三乙酸胺 (1.5)、MgSO₄·7H₂O (3)、MnSO₄·H₂O (0.5)、NaCl (1)、FeSO₄·7H₂O (0.1)、CoCl₂·6H₂O (0.1)、CaCl₂ (0.1)、ZnSO₄·7H₂O (0.1)、CuSO₄·5H₂O (0.01)、AlK(SO₄)₂·12H₂O (0.01)、H₃BO₃ (0.01)、Na₂MoO₄·2H₂O (0.01)、NiSO₄·6H₂O (0.03)、Na₂SeO₃ (0.02) 和 Na₂WO₄·2H₂O (0.02) 组成。

[0186] 当波长 600 (OD₆₀₀) 时测量的光密度 (OD) 介于 0.10-0.14 时,分别用 1ml 和 2ml 灭菌空气 (~160 至 320 μ l 氧气) (图 1C) 和 2 μ g/ml 丝裂霉素 C (图 1D) 刺激反刍甲烷短杆菌。两种刺激均可观察到典型裂解曲线,空气刺激的潜伏时间为 ~90 分钟。丝裂霉素 C 刺激的最初结果表明非常短的潜伏期。为了校验从宿主基因组中切除噬菌体,设计两条寡核苷酸,分别面向两个噬菌体附着位点 (R1F: caaagagagattaaagaagcagacg; SEQ ID NO: 146 以及 L2Ragtagtggttggaatcagtgaaaagg; SEQ ID NO: 147)。如果噬菌体基因组在切除后再次环化,这对引物则能产生扩增子。

[0187] 图 1E 描绘了当用空气刺激反刍甲烷短杆菌时,最初的切除实验。诱导后发现清楚明确的具有期望大小的扩增子,表明成功的切割和再环化。在未诱导的反刍甲烷短杆菌细胞中也发现相似但微弱的条带,表明在正常非刺激的生长中, ϕ mru 具有自发切割的能力。

[0188] 实施例 5C:裂解酶生物实验

[0189] ORF 2058 编码的多肽可用作瘤胃产甲烷菌特异性裂解酶,并且亚克隆入大肠杆菌表达载体用于产生重组蛋白。使用引物 Mbbrum11for22(1122For, cac cat ggt tag att cag cag aga c;SEQ ID NO:148) 和 Mbbrum 11rev22(1122Rev,tca tgc agg aca gac aac ata gta g;SEQ ID NO:149) 通过 PCR 在 150 μ L 反应体系中扩增 ORF 2058,该反应体系含 121.5ng 反刍甲烷短杆菌菌株 M1 基因组 DNA;0.2 μ M 1122For 和 1122Rev 引物;15 μ L Accuprime Pfx 缓冲液(含 dNTP,英骏公司(InVitrogen));2.4 μ L Accuprime Pfx(英骏公司)。PCR 条件为 95 $^{\circ}$ C 2 分钟初始变性,然后进行 35 个下述循环:95 $^{\circ}$ C 15 秒、55 $^{\circ}$ C 30 秒和 68 $^{\circ}$ C 40 秒。无需最后的延伸步骤。纯化 PCR 产物,并用 Nanodrop(赛默飞世尔科技公司,美国佐治亚州)定量。

[0190] ORF 2058 克隆:根据厂商推荐,将 PCR- 扩增 ORF 2058 克隆入 pET 100 或 pET 151-D Topo 载体(英骏公司)中,并转化到化学感受态 TOP 10 细胞(英骏公司)中。用菌落 PCR 分析转化子,纯化质粒 DNA 并测序。选择具有匹配 ORF 2058 的 DNA 序列的克隆。

[0191] ORF 2058 表达:利用电穿孔将来自含有经校验的 ORF 2058 插入物的克隆的质粒 DNA 转化到电感受态 BL21* 或 Rosetta 2 细胞中。发现表达可溶性 ORF 2058 蛋白的最佳生长条件是 LB 培养基,在吸光度 600 为 0.48-0.6 时用 0.5mM IPTG 进行诱导,并在 30 $^{\circ}$ C 继续培养 6 小时。然后离心收集细胞并冻存于 -20 $^{\circ}$ C。

[0192] 细胞裂解:融解细胞团块,并重悬于下列缓冲液中(pH 7.5):300mMNaCl、2mM DTT、10mM 咪唑、20mM Tris、20%甘油、1% Triton-X、5mM CaCl₂和 10mM MgCl₂。加入溶菌酶终浓度为 1mg/ml,然后在冰上轻轻搅拌 30min。加入 DNA 酶 I 和 RNA 酶 I,终浓度均为 5 μ g/ml,然后在冰上轻轻搅拌 30 分钟。将细胞裂解物在 12,000rpm 离心 15 分钟,用 0.8 μ m 滤器过滤粗提裂解物。

[0193] 镍亲和层析:将来自细胞裂解步骤的过滤上清液施加于 80mL 镍亲和柱,并用含 20mM-250mM咪唑的下列缓冲液(pH 7.5)梯度洗脱:300mMNaCl、2mM DTT、20mM Tris和 20%甘油。用装有 10,000kDa 分子量截止膜的密理博超滤小室浓缩从层析柱上洗脱的含有所表达 ORF 2058 蛋白的组分。利用下列洗脱缓冲液从镍柱上洗脱大肠杆菌 BL21* 细胞中表达的 pET100 中的 ORF 2058 构建物:pH 8.2(20mM Tris、250mM咪唑、300mMNaCl、10mM b- 巯基乙醇、10%甘油),并且在另外加入甘油和二硫苏糖醇的缓冲液中储存酶,该缓冲液的终浓度为 40%甘油,1mM 二硫苏糖醇,pH8.2。

[0194] 脱盐:使用含下列缓冲液(pH 7.0):20mM MOPS、1mM DTT、300mMNaCl 和 20%甘油的 250mL BioGel P6DG(伯乐公司,美国加州)(BioRad,CA,USA) 柱对 Rosetta 2 细胞中 pET 151 构建物表达的浓缩蛋白进行脱盐。如上所述浓缩来自柱的组分,将最终样品过滤并用液氮速冻,然后储存于 -20 $^{\circ}$ C。

[0195] 裂解静息的反刍甲烷短杆菌细胞:在含 BY+ 培养基的亨盖特试管(Hungate tube)中将 5ml 反刍甲烷短杆菌 M1(DSM 1093)的培养物培养至对数后期,室温下 5,000x 离心 30 分钟收集亨盖特试管中的细胞。将细胞移至无氧室中(95% CO₂-5% H₂,科艺实验室产品公司,美国密歇根州)(CoyLaboratory Products, MI, USA),弃去上清,将来 10ml 培养物的细胞重悬于 1ml MOPS 缓冲液 pH 6.8(50mM MOPS,5mM CaCl₂,1mM 二硫苏糖醇)中。用另外的 MOPS 缓冲液稀释细胞悬液,将其 OD(600nm)调整至 $\sim$ 0.12。

[0196] 将标准细胞悬液(50 μ l)分配至微孔板中,加入不同浓度的 ORF 2058 裂解酶(由

pET 100 构建物制备),用缓冲液将反应总体积补至 250ml。在 37℃ 孵育细胞和蛋白混合物,并记录 OD 读值。在静息反刍甲烷短杆菌细胞中加入酶(每个实验加入的酶 μg 数)的效果如图 7 所示。与未加入酶的对照细胞相比,酶的加入以剂量依赖方式降低悬浮细胞的 OD 600nm 读数。这表明 ORF 2058 裂解酶能够在无氧条件下攻击和裂解反刍甲烷短杆菌静息细胞。

[0197] 裂解生长的反刍甲烷短杆菌细胞:反刍甲烷短杆菌生长于 RM02 培养基中。RM02 培养基由下列组分组成(g/L): KH_2PO_4 (1.4)、 $(\text{NH}_4)_2\text{SO}_4$ (0.6)、KCl(1.5)、微量元素溶液 SL10(1ml)、硒/钨溶液(1ml)、0.1% (w/v) 刃天青溶液(4 滴)。将组分混合并在无氧 100% CO_2 的条件下煮沸,用 100% CO_2 鼓泡并在冰浴中冷却。冷却后加入 NaHCO_3 (4.2g) 和 L- 盐酸半胱氨酸 $\cdot\text{H}_2\text{O}$ (0.5g),将 9.5ml 培养基分配至亨盖特管中,用 100% CO_2 对试管进行通气。将试管高压灭菌,使用前避光储存 24h。接种前,加入 NoSubRFV(每管 0.5ml,包含底物、酵母提取物、维生素)。接种后用 80% CO_2 /20% H_2 通气至 25lb/in²。反刍甲烷短杆菌生长至对数中期(OD 600nm ~ 0.1),这时加入不同浓度的 ORF 2058 裂解酶(由 pET 151D Topo 克隆制备)。继续孵育培养物并记录 OD 读数。加入酶对反刍甲烷短杆菌生长和甲烷形成(以括号表示生长 217 小时后相对未加入对照的%甲烷产生)的影响如图 8 所示。结果显示 ORF 2058 裂解酶以剂量依赖方式显著影响反刍甲烷短杆菌生长,在加入 2 小时内降低生长培养物的 OD 600。两次最高水平的酶加入也减少了甲烷形成,与加入氯仿(100 μl /10ml 培养物)的程度相似。

[0198] 实施例 6:概述

[0199] 反刍甲烷短杆菌基因组测序中的一个意外发现是存在原噬菌体序列。对于基因组序列的分析鉴定出异常高 GC 含量区,其中包含多种噬菌体相关的基因。进一步的生物信息学分析鉴定出预测的原噬菌体的整体结构,并称其为 ϕmru 。将约 40% 的噬菌体基因指定为不同的功能组,包括噬菌体整合、DNA 复制和包装、噬菌体结构蛋白质和裂解。发现侧接噬菌体基因组的 DNA 序列代表潜在的噬菌体整合位点(attL 和 attR)。

[0200] 噬菌体似乎将其本身插入到反刍甲烷短杆菌的推定膜蛋白中,该蛋白可能含有 ϕmru 噬菌体基因组的原始产甲烷菌整合位点,attB。而且,发现于 DNA 复制模块内的终止子样结构被认为代表了噬菌体 DNA 复制起点。噬菌体基因组 3' 末端的低 -GC 区含有似乎是 DNA 修饰系统的组件,包括可能控制 dnd 系统表达的基因。这些基因可能在噬菌体整合的过程中被带入反刍甲烷短杆菌的基因组中,并且其在修饰噬菌体、宿主或外源 DNA 方面的角色还有待阐明。反刍甲烷短杆菌对于 dnd 系统的保留表明它使宿主受益。然而, ϕmru dnd 系统在反刍甲烷短杆菌或外源 DNA 修饰中的作用还有待研究。

[0201] ϕmru 序列的另一有趣特点是对应低 GC 区且与各种有机体蛋白质匹配性较低的反义链编码基因的数目。这表明这些基因在其整合入反刍甲烷短杆菌之后在 ϕmru 的基因组中积累,这些基因可能代表正在进行之中的插入建立,也许最终导致噬菌体失活和噬菌体驯化。与反刍甲烷短杆菌基因组相比 ϕmru 噬菌体序列的高 GC 含量表明它起源于另一有机体。然而,之前的宿主并未明确,因为与其它迄今为止已知的噬菌体相比 ϕmru 蛋白质显示出某种独特性。

[0202] 关于减少甲烷的尤其让人感兴趣的 ϕmru 基因是位于裂解盒中的那些基因。具体说,一个基因编码与 C39 肽酶家族类似的蛋白质。该肽酶家族包括病毒半胱氨酸内肽酶等,

如 C71 古菌噬菌体内异肽酶,该酶切割组成甲烷短杆菌细胞壁的假胞壁质中的交联肽。根据噬菌体基因组内的基因位置以及来自其它非反刍产甲烷菌噬菌体基因组的假胞壁质内异肽酶的同线性,该基因可用作编码参与噬菌体后代释放前细胞裂解中的裂解酶的溶素基因。对该基因和它编码的酶尤感兴趣,因为可能在具有相似细胞壁的反刍甲烷短杆菌和其它反刍产甲烷菌中有控制机制。

[0203] 反刍噬菌体及其参与裂解宿主细胞的酶有可能用来控制瘤胃中产甲烷菌群体和其它群落成员(细菌、原生动物和真菌)。此外,通过研究噬菌体的生命周期,有可能鉴定出易于受噬菌体蛋白质抑制的关键的宿主酶靶点。本发明人对牛、绵羊和鹿中反刍噬菌体的组成进行过调查研究,显示噬菌体数量和类型随时间的变化。还鉴定出受噬菌体影响的新西兰产甲烷菌分离物。利用纯产甲烷菌培养物评价噬菌体裂解酶,开发基于培养物和 PCR 的技术以筛选新噬菌体。纯化自瘤胃样品的噬菌体可进行随机 DNA 测序分析,从而发现噬菌体酶。

[0204] 噬菌体或其酶在用于降低甲烷排放的缓解技术中具有多种优势。噬菌体是瘤胃微生物群落中的天然成员,因此,不会被视作抗生素治疗(并可更容易地克服任何管理限制)。噬菌体通常对窄范围的宿主具有特异性,因此可以选择性靶向产甲烷菌。现在认为噬菌体治疗是对抗生素耐受型有机体的一种治疗,一般认为是安全的。一旦产生,噬菌体通常相对稳定。将对噬菌体易感的产甲烷菌菌株引入瘤胃可能具有长期益处,尤其是在早期进行接种的情况下(如羊羔或和牛犊)。已知某些产甲烷菌包含噬菌体基因组、对裂解噬菌体易感、或发生自溶(提示有裂解酶),包括史氏甲烷短杆菌(*Methanobrevibacter smithii*) (菌株 PS)、布氏甲烷杆菌(*Methanobacterium bryantii*) 和甲烷短杆菌菌株 MF-1。用噬菌体抑制引起农业问题的有机体的一个著名实例是使用噬菌体靶向大肠杆菌 0157:H7。

[0205] 选择反刍甲烷短杆菌进行基因组测序是因为它在各种饲养条件下的反刍动物瘤胃中广泛存在(基于培养和分子检测数据),容易获得培养物,易于进行实验室常规培养以及可获得关于这种有机体的大量的前期研究数据和背景文献。本发明提供关于反刍甲烷短杆菌基因组的重要数据,并且对瘤胃内噬菌体进行详细构图。 ϕ mrul 原噬菌体序列为抑制反刍甲烷短杆菌以及未来的有助于确定基因功能的遗传学操纵提供了具体试剂。可利用该噬菌体阻断产甲烷菌中的保守功能/组件以防止或减少瘤胃中的甲烷形成。

[0206] 参考文献

[0207] Altermann E, Klaenhammer TR(2005) 通路探险人:用京都基因和基因组(KEGG)百科全书数据库进行通路作图(PathwayVoyager: pathway mapping using the Kyoto Encyclopedia of Genes and Genomes(KEGG) database). BMC Genomics 6:60-66。

[0208] Altermann, E Klaenhammer TR(2003) GAMOLA:序列注释以及对原核基因组草图和完成序列进行分析的新的本地解决方案(GAMOLA: a new local solution for sequence annotation and analyzing draft and finished prokaryotic genomes). OMICS: A journal of integrative biology 7, 161-169。

[0209] Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ(1997), 缺口 BLAST 和 PSI-BLAST:新一代的蛋白数据库检索程序(Gapped BLAST and PSI-BLAST: a new generation of protein database search programs). Nucleic Acids Research 25, 3389-3402。

[0210] Balch WE, Fox GE, Magrum LJ, Woese CR, Wolfe RS(1979) 产甲烷菌:一个独特生物学群体的重新评估 (Methanogens:reevaluation of a uniquebiological group). Microbiological Reviews 43,260-296。

[0211] Baresi, L. 和 Bertani, G. 1984. 分离一种产甲烷菌的细菌噬菌体 (Isolation of a bacteriophage for a methanogenic bacterium), 刊于《美国微生物协会年会摘要》 (Abstracts of the Annual Meeting of the American Societyfor Microbiology), 华盛顿特区, 美国微生物协会 (Washington DC:AmericanSociety for Microbiology), 第 133 页。

[0212] Bickle, T.A. 和 D.H.Kruger. 1993, DNA 限制性生物学 (Biology ofDNA restriction). Microbiol. Rev. 57 :434-450。

[0213] Bult CJ 等 (1996) 产甲烷古细菌詹氏产甲烷菌的完整基因组序列 (Complete genome sequence of the methanogenic archaeon, *Methanococcusjannaschii*). Science 273,1058-1073.

[0214] Coutinho PM, Henrissat B(1999) 碳水化合物-活性酶:整合数据方案. “碳水化合物生物工程最新进展” (Carbohydrate-active enzymes:anintegrated database approach. ‘Recent Advances in CarbohydrateBioengineering’) (HJ Gilbert,G Davies, B Henrissat 和 B Svensson 编) 3-12 皇家化学会, 剑桥 (The Royal Society of Chemistry, Cambridge) (碳水化合物活性酶数据库 (Carbohydrate Active Enzymes database), <http://www.cazy.org/>)。

[0215] Delcher AL, Harmon D, Kasif S, White O, Salzberg SL(1999) 用 GLIMMER 提高微生物基因识别 (Improved microbial gene identification withGLIMMER). Nucleic Acids Research 27,4636-4641。

[0216] Fleischmann, RD 等 (1995) 流感嗜血杆菌 Rd 的全基因组测序和组装 (Whole-genome random sequencing and assembly of *Haemophilus influenzae*Rd). Science 269,496-512。

[0217] Fricke WF, Seedorf H, Henne A, Kruer M, Liesegang H, Hedderich R, Gottschalk G, Thauer RK(2006) 斯氏甲烷球形菌基因组序列揭示为什么人小肠古菌仅限于用甲醇和 H₂ 形成甲烷和合成 ATP(The genome sequenceof *Methanosphaera stadtmanae* reveals why this human intestinal archaeon isrestricted to methanol and H₂ for methane formation and ATP synthesis). Journal of Bacteriology 188,642-658.

[0218] Godde JS, Bickerton A(2006) 称为 CRISPR 的重复性 DNAe 及其相关基因:原核动物间水平转移的证据 (The repetitive DNAe called CRISPRs andtheir associated genes:evidence of horizontal transfer among prokaryotes.) Journal of Molecular Evolution 62,718-729。

[0219] Haft DH, Selengut J, Mongodin EF, Nelson KE(2005) 原核基因组中 45 个 CRISPR 相关 (Cas) 蛋白质家族和多种 CRISPR/Cas 亚型的指南 (a guild of45 CRISPR-associated(Cas)protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes). PLoS Computational Biology1:474-483。

[0220] Jansen R, Embden JD, Gaastera W, Schouls LM(2002) 原核细胞中 DNA 重复相

关基因的鉴定 (Identification of genes that are associated with DNA repeats in prokaryotes). *Molecular Microbiology* 43, 1565-1575.

[0221] Jansen R, van Embden JD, Gastra W, Schouls LM (2002) 在原核细胞中鉴定新颖序列重复家族 (Identification of a novel family of sequence repeats among prokaryotes). *OMICS: A journal of integrative biology* 6, 23-33.

[0222] Jensen LJ, Friis C, Ussery DW (1999) 微生物基因组的三种观点 (Three views of microbial genomes). *Research in Microbiology* 150, 773-777.

[0223] Joblin KN, Naylor GE, Williams AG (1990) 史氏甲烷短杆菌对厌氧反刍真菌木聚糖水解活性的影响 (Effect of *Methanobrevibacter smithii* on xylolytic activity of anaerobic ruminal fungi). *Applied and Environmental Microbiology* 56, 2287-2295.

[0224] Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M (2004) 解密基因组的KEGG资源 (The KEGG resource for deciphering the genome). *Nucleic Acids Research* 32, D277-D280.

[0225] Kiener, A., König, H., Winter, J. 和 Leisinger, T. 1987. 纯化并利用沃氏甲烷杆菌的假胞质壁内肽酶裂解热自养甲烷杆菌 (Purification and use of *Methanobacterium wolfei* pseudomurein endopeptidase for lysis of *Methanobacterium thermoautotrophicum*). *J. Bacteriol.* 169, 1010-1016.

[0226] Knox, M. R. 和 Harris, J. E. 1986. 分离并表征史氏甲烷短杆菌的噬菌体 (Isolation and characterisation of a bacteriophage of *Methanobrevibacter smithii*.) 第十四届国际微生物大会摘要 (Abstracts of the XIV International Congress on Microbiology), 曼彻斯特: 国际微生物协会联盟 (Manchester: International Union of Microbiological Societies).

[0227] Kurtz S, Schleiermacher C (1999) REPuter: 完整基因组最大重复的快速计算 (REPuter: fast computation of maximal repeats in complete genomes). *Bioinformatics* 15, 426-427.

[0228] Lowe TM, Eddy SR (1997) tRNAscan-SE: 改进对基因组序列中转移 RNA 基因进行检测的程序 (tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence). *Nucleic Acids Research* 25, 955-964.

[0229] Loenen, W. 和 N. Murray. 1986. 噬菌体 λ 的限制性缓解蛋白 (Ral) 增强修饰 (Modification enhancement by restriction alleviation protein (Ral) of bacteriophage lambda.) *J. Mol. Biol.* 190: 11-22.

[0230] Lucchini, S., F. Desiere, 和 H. Brussow. 1999. 嗜热链球菌噬菌体种类的比较基因组学支持模块化进化理论 (Comparative genomics of *Streptococcus thermophilus* phage species supports a modular evolution theory). *J. Virol.* 73: 8647-8656.

[0231] Luo, Y. N., Pfister, P., Leisinger, T. 和 Wasserfallen, A. 2002. 假胞壁质内异肽酶 PeiW 和 PeiP, 甲烷嗜热杆菌菌株产生的新颖蛋白酶家族的两个中度相关的成员 (Pseudomurein endoisopeptidases PeiW and PeiP, two moderately related members of a novel family of proteases produced in *Methanothermobacter* strains), *FEMS*

Microbiol. Lett. 208, 47-51.

[0232] Makarova, K. S., Aravind, L. 和 Koonin, E. V. 1999. 与动物转谷氨酰胺酶同源的古细菌、细菌和真核蛋白的超家族 (A superfamily of archaeal, bacterial, and eukaryotic proteins homologous to animal transglutaminases). Protein Sci. 8, 1714-1719.

[0233] Makarova KS, Grishin NV, Shabalina, SA, Wolf YI, Koonin EV (2006) 原核细胞中推定的基于 RNA 干扰的免疫系统: 用计算机分析预测酶机器、真核 RNAi 功能类推和假拟作用机制 (A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action). Biology Direct 1:7-32.

[0234] 新西兰统计 2005 (www.stats.govt.nz).

[0235] 新西兰温室气体目录 (New Zealand's Greenhouse Gas Inventory) 1990-2004. 国家目录报告和一般报告格式 (The National Inventory Report and Common Reporting Format). (2006) 环境部 (Ministry for the Environment). <http://www.mfe.govt.nz/publications/climate/nir-apr06/nir-apr06.pdf>.

[0236] Reeve JN, Nolling J, Morgan RM, Smith DR (1997) 甲烷生成: 基因、基因组、谁是第一个? (Methanogenesis: genes, genomes and who's on first?) Journal of Bacteriology 179, 5975-5986.

[0237] Rawlings, N. D., Morton, F. R. 和 Barrett, A. J. 2006. MEROPS: 肽酶数据库 (MEROPS: the peptidase database). Nucleic Acids Res. 34, D270-D272.

[0238] Samuel BS, Hansen EE, Manchester JK, Coutinho PM, Henriissat B, Fulton R, Latreille P, Kim K, Wilson RK, Gordon JI (2007) 史氏甲烷短杆菌对人肠道的基因组适应 (Genomic adaptations of *Methanobrevibacter smithii* to the human gut). Proceedings of the National Academy of Sciences USA 104, 10643-10648.

[0239] Smith DR 等 (1997) 热自养甲烷杆菌 H 的完整基因组序列: 功能分析和比较基因组学 (Complete genome sequence of *Methanobacterium thermoautotrophicum* ΔH: Functional analysis and comparative genomics). Journal of Bacteriology 179, 7135-7155.

[0240] Smith PH, Hungate RE (1958) 反刍甲烷短杆菌的分离和表征 (Isolation and characterization of *Methanobacterium ruminantium* n. sp.) Journal of Bacteriology 75, 713-718.

[0241] Staden R, Beal KF, Bonfield JK (1998) 司塔登软件包 (The Staden Package). 分子生物学方法: 生物信息学方法和实验指南 (Methods in Molecular Biology: Bioinformatics Methods and Protocols) 132, 115-130.

[0242] Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS, Kiryutin B, Galperin MY, Fedorova ND, Koonin EV (2001) COG 数据库: 来自完整基因组的蛋白质的系统发育分类的新进展 (The COG database: new developments in phylogenetic classification of proteins from complete genomes) Nucleic Acids Research 29, 22-28.

[0243] 将上述说明书提及的所有公开物和专利全文纳入本文作参考。

[0244] 如果参考上述说明书中提及具有已知等同形式的内容,那么将这些等同形式也纳入本文,就好像在本文中单独列出那样。虽然已结合具体的优选实施方案对本发明作了描述,但应了解,如权利要求所述,本发明不应过分地受限于这些具体实施方案。应当理解的是在不背离本发明的构思和范围的情况下,可对本发明作进一步的修改。

序列表

<110> 田园温室气体研究有限公司 (PASTORAL GREENHOUSE GAS RESEARCH LTD)

<120> 噬菌体 ϕ mru 多核苷酸和多肽及其应用

<130>CDH 229498PCTPR

<140>PCT/NZ08/000248

<141>2008-09-25

<150>60/989, 841

<151>2007-11-22

<150>60/989, 840

<151>2007-11-22

<150>60/975, 104

<151>2007-09-25

<160>151

<170>PatentIn 3.5 版

<210>1

<211>258

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>1

Val Arg Asn Met Lys Asn Lys Ser Leu Ile Leu Ile Ser Leu Leu Leu

1 5 10 15

Leu Ile Thr Ile Ile Ser Ile Gly Ser Val Val Ala Thr Asp Asn Glu

20 25 30

Glu Ile Asn Met Asp Asn Ile Asn Asn Ile Asp Asn Asn Glu Asp Ile

35 40 45

Ala Asn Ile Asp Asn Val Asp Asn Val Asp Asn Ser Asn Ile Asn Asn

50 55 60

Pro Thr Asp Ile Arg Ile Asp Asn Ser Asn Leu Asn Arg Glu Thr Glu

65 70 75 80

Leu Asp Ser Asn Leu Asn Lys Ser Asn Gln Ile Arg Glu Asp Glu Leu

85 90 95

Glu Gln Ser Asn Ala Lys Ser Asn Leu Lys Ser Ser Lys Leu Ser Ser

100 105 110

Thr Ile Thr Val Asp Gly Ser Asp Glu Asn Gln Met Ser Asn Pro Thr

115 120 125

Ile Gln Ser Ala Ile Asp Ser Ala Asn Ala Gly Asp Thr Ile Ile Ile

130 135 140

Thr Gly Lys Ser Tyr Val His Cys His Phe Ile Val Asn Lys Pro Leu

145	150	155	160
Thr Ile Ile Ser Glu Ile Gly Thr Ser Met Ser Pro Cys Pro Ser Asn			
	165	170	175
Thr Lys Gly Ser Gly Ala His Gly Ile Phe Tyr Ile Ser Pro Glu Ala			
	180	185	190
Ser Gly Thr Val Leu Lys Gly Phe Asn Leu Thr Asn Thr Tyr Gly Asp			
	195	200	205
Tyr Asp Asp Tyr Gly Ile Leu Ile Arg Gly Ala Glu Asn Val Glu Ile			
	210	215	220
Ile Asn Cys Thr Ile Asn Thr Val Ser Asp Gly Asp Gly Ile Arg Ile			
225	230	235	240
Glu Asn Ala Thr Asn Thr Lys Ile Ala Asp Cys Leu Ile Lys Asp Ser			
	245	250	255
Asn Ile			
<210>2			
<211>255			
<212>PRT			
<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)			
<400>2			
Met Lys Thr Phe Arg Gln Gln Leu Leu Glu Asp Pro Glu Phe Gln Asn			
1	5	10	15
Tyr Leu Leu Gln Arg Pro Asn Leu Thr Glu Ser Ser Leu Gln Ser Tyr			
	20	25	30
Leu Asn Ala Ala Thr Asn Phe Val Arg Phe Thr Gly Glu Pro Phe Tyr			
	35	40	45
Lys Thr Val His Glu Leu Arg Ser Gln Gln Asn Asp Arg Ile Glu Asn			
	50	55	60
Asn Ile Ile Ile Arg Phe Asn Pro Asn Gln Ser Arg Ile Asn Ile Met			
65	70	75	80
Gln Phe Glu Phe Ile Glu Tyr Leu Lys Gly Arg Gly Cys Thr Clu Val			
	85	90	95
Ser Ile Asp Ser Tyr Val Arg Tyr Met Arg Thr Ile Leu Ser Thr Leu			
	100	105	110
Gly Ile Ile Leu Pro Lys Ser Pro Lys Leu Asp Asp Thr Pro Gln Asp			
	115	120	125
Trp Tyr Leu Leu Thr Lys Asp Asp Ile Lys Tyr Val Leu Asp Thr Ala			
	130	135	140
Asn Leu Gln Tyr Lys Ala Val Ile Asn Phe Ala Ala Val Thr Gly Leu			
145	150	155	160

Arg	Val	Arg	Asp	Met	Arg	Ser	Leu	Thr	Ile	Lys	Asp	Phe	Met	Thr	Ala	
				165					170					175		
Thr	Glu	Glu	Tyr	His	Gly	Cys	Thr	Glu	Val	Glu	Asp	Phe	Leu	Asp	Ser	
				180					185					190		
Ala	Pro	Asp	Gly	Met	Ile	Gly	Phe	Trp	Glu	Leu	Phe	Pro	Gln	Lys	Thr	
				195					200					205		
Arg	Lys	Phe	Arg	Leu	Pro	Cys	Lys	Val	Cys	Asn	Thr	Pro	Glu	Ser	Ser	
				210					215					220		
Asp	Leu	Leu	Leu	Phe	Ser	Leu	Asn	Glu	Arg	Val	Lys	Tyr	Phe	Glu	Trp	
				225					230					235		
Lys	Asn	Glu	Lys	Asp	Gly	Thr	Asp	Leu	Lys	Ile	Thr	Lys	Asn	Asp		
				245					250					255		

<210>3

<211>98

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>3

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

Lys Ile

<210>4

<211>255

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>4

Met	Gly	Met	Lys	Lys	Val	Lys	Ile	Glu	Ser	Lys	Ala	Lys	Asn	Asn	Met	
1	5				10				15							
Thr	Arg	Asn	Glu	Lys	Leu	Phe	Tyr	Lys	Leu	Tyr	Asp	Asp	Leu	Tyr	Asn	
				20					25					30		

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

<210>5

<211>228

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>5

Met Lys Thr Gln Asp Leu Ile Asn Ile Ile Asn Asp Glu Glu Ser Pro		
1	5	10
Val Phe Leu Asn Arg Glu Val Phe Glu Met Asp Tyr Val Pro Asp Ile		
20	25	30
Tyr Lys Tyr Arg Asp Glu Gln Leu Ala Lys Met Ala Met Tyr Cys Asn		
35	40	45

Ser Ile Pro Asp Asn Ile Ala Pro Lys Asn Leu Gln Leu Cys Gly Gly
 50 55 60
 Asn Ala Thr Gly Lys Thr Thr Thr Leu Lys Gln Phe Phe Lys Met Leu
 65 70 75 80
 Asn Glu Ala Phe Pro Asn Ile Val Thr Val Tyr Ile Asn Cys Gln Leu
 85 90 95
 Phe Asn Thr Glu Asn Thr Val Tyr Gly Lys Ile Tyr Asn Lys Leu Tyr
 100 105 110
 Gly Val Lys Gly Ser Ile Asn Gly Lys Ser Asn Thr Met Leu Phe Asp
 115 120 125
 Lys Ile Val Ala Arg Leu Lys Lys Glu Asn Lys Ile Leu Ile Ile Gly
 130 135 140
 Leu Asp Asp Phe Asp Ser Phe Lys Ser Arg Asp Gly Leu Asn Lys Met
 145 150 155 160
 Leu Tyr Asn Phe Leu Arg Ile His Glu Ala Glu Glu Gly Ile Gln Ile
 165 170 175
 Cys Ile Phe Thr Val Ser Asn Lys Gly Glu Ser Glu Ser Leu Leu Leu
 180 185 190
 Pro Ser Arg Gln Ser Ser Thr Gly Phe Arg Tyr Ser Leu Thr Ser Thr
 195 200 205
 Pro Trp Ser Arg Cys Thr Thr Tyr Trp Thr Thr Gly Ala Pro Ser Val
 210 215 220
 Ser Ile Leu Ala
 225
 <210>6
 <211>255
 <212>PRT
 <213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)
 <400>6
 Met Phe Leu Glu Lys Cys Asp Gly Lys Asp Thr Ile Ser Met Ser Leu
 1 5 10 15
 Gln Glu Lys Met Asn Leu Ile Leu Glu Thr Met Glu Ser Lys Gly Ser
 20 25 30
 Pro Phe Ile Ser Cys Ile His Cys Asn Ile Pro Ile Asn Glu Ala Glu
 35 40 45
 Glu Trp Tyr Lys Asn Gly Glu Ile Gly Asp Gln Asp Phe Ile Asn Phe
 50 55 60
 Tyr Asp Asp Val Asn Leu Ile Glu Glu Ser Phe Gly Phe Glu Ile Tyr
 65 70 75 80

Lys Lys Ser Glu Tyr Pro Thr Leu His Thr Gln Ser Ile Asn Gln Ile
 90 95
 Ala Ser Thr Tyr Pro Met Asn Arg Thr Gln Asn Glu Lys Thr Pro Leu
 100 105 110
 Phe Leu Arg Arg Glu Arg Lys Leu Tyr Glu Ile Thr Asn Ile Phe Lys
 115 120 125
 Ser His Ser Thr Asn Glu Ile Phe Ile Ser Met Asp Ser Lys Ser Lys
 130 135 140
 Leu Lys His Glu Ile Lys Tyr Asp Phe Thr Leu Lys Glu Leu Asn Glu
 145 150 155 160
 Ile Phe Lys Asn Tyr Leu Glu Glu Asp Cys Ser Ile Tyr Ile Leu Asn
 165 170 175
 Asp Asn Arg Ala Phe Val Met Thr Leu Gly His Phe Gln Phe Glu Phe
 180 185 190
 Asp Val Phe Gly Ser Ala Lys Glu Ser Tyr Val Phe Asp Val Glu Ile
 195 200 205
 Asp Asp Glu Lys Tyr Asp Lys Ile Phe Ile Arg Ser Ser Tyr Tyr Phe
 210 215 220
 Asn Ile Pro Val Asp Glu Leu Asp Gly Leu Ala Lys Ile Leu Lys Gln
 225 230 235 240
 Lys Arg Ile Ile Glu Glu Gly Phe Phe Glu Gly Gln Phe Leu Tyr
 245 250 255

<210>7

<211>68

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>7

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

Lys Gly Asp Ala

65

<210>8

<211>51

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>8

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

<210>9

<211>77

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>9

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

<210>10

<211>53

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>10

```
Met Ser Met Leu Ala Asp Phe Glu Pro Ala Arg Leu His Lys Arg Thr
1           5           10           15
Trp Ala Glu Arg His Asp Val Glu Ile Leu Ala Val Ile Cys Leu Ala
           20           25           30
Ile Ser Ile Ala Met Leu Leu Leu Phe Phe Ala Leu Ala Glu Pro Thr
           35           40           45
Val Ala Gly Val Ile
           50
```

<210>11

<211>150

<212>PRT

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>11

```

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

```

<210>12

<211>251

<212>PRT

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>12

```

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

```

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

<210>13

<211>76

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>13

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

<210>14

<211>255

<212>PRT

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>14

```

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

```

<210>15

<211>143

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>15

```

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

```

<210>16

<211>134

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>16

```

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

```

Val Lys Leu Lys His Ile Arg Arg Met Tyr Glu Thr Ala Leu Lys His
115 120 125

Ser Phe Glu Ala Ile Lys
130

<210>17

<211>91

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>17

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

<210>18

<211>77

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>18

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

<210>19

<211>49

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>19

```
Met Asn Tyr Arg Glu Trp Val Ala Ser Gln Phe Glu Leu Arg Met Thr
1           5           10           15
Glu Asp Gly Arg Val Cys Phe Lys Thr Pro Cys Ser Tyr Tyr Thr Phe
          20           25           30
Ser Lys Glu Asp Phe Glu Ile Ile Arg Glu Met Phe Leu Asn Phe Glu
        35           40           45
```

Ser

<210>20

<211>57

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>20

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

<210>21

<211>111

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>21

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

Tyr Met Ala Ile Gln Lys Glu Met Lys Arg Ile Gly Leu Lys Arg
 100 105 110

<210>22

<211>183

<212>PRT

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>22

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

<210>23

<211>77

<212>PRT

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>23

Met Ile Tyr Asp Lys Ala Thr Val Thr Ser Leu Ile Val Ala Ile Leu
 1 5 10 15
 Leu Pro Leu Met Ser Met Leu Gly Ile Gly Glu Leu Thr Gln Asn Tyr

20	25	30
Ile Leu Ala Ile Val Ser Gly Met	Ile Ala Leu Val Val Trp Tyr Tyr	
35	40	45
Asn Glu Lys His Asn Ser Asp Leu Val Ser Gly Thr Thr Lys Cys Asp		
50	55	60
Cys Glu Leu Cys Tyr Gly Gly Asp Asp Glu Ala Leu Ile		
65	70	75

<210>24

<211>112

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>24

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

<210>25

<211>204

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>25

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

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

<210>26

<211>152

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>26

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

Ala Val Leu Ser Tyr Ile Lys Trp Met Lys Thr Gln Glu Glu Tyr Tyr
 130 135 140
 Lys Lys Trp Arg His Tyr His Asp
 145 150
 <210>27
 <211>218
 <212>PRT
 <213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)
 <400>27
 Met Ile Asn Glu Arg Ile Arg Pro Tyr Leu Asn Phe Ser Phe Asn Asp
 1 5 10 15
 Lys Lys Val Ile Phe Val Thr Leu Phe Val Val Ser Asn Leu Ile Ser
 20 25 30
 Asn Leu Leu Ala Ile Lys Val Phe Asn Leu Gly Phe Trp Gly Leu Thr
 35 40 45
 Thr Asp Cys Gly Asn Leu Leu Phe Pro Leu Gly Tyr Leu Met Ala Asp
 50 55 60
 Val Ile Thr Glu Val Tyr Gly Glu Arg Thr Ala Arg Arg Val Ile Leu
 65 70 75 80
 Leu Gly Leu Phe Ala Asn Ile Leu Leu Ile Val Ala Thr Thr Leu Thr
 85 90 95
 Val Tyr Met Pro Tyr Pro Ser Tyr Trp Thr Gly Gln Gly Ala Tyr Ala
 100 105 110
 Tyr Met Phe Gly Phe Thr Pro Arg Ile Val Leu Ala Gly Phe Ile Ala
 115 120 125
 Tyr Leu Val Gly Gln Phe Val Asn Ala Arg Leu Met Val Leu Ile Lys
 130 135 140
 Lys Trp Thr Asn Ser Lys Tyr Leu Phe Met Arg Thr Ile Gly Ser Thr
 145 150 155 160
 Leu Gly Gly Glu Leu Cys Asp Ser Cys Ile Cys Ser Ser Ile Ala Tyr
 165 170 175
 Tyr Gly Ile Val Pro Asn Ser Gly Ile Leu Leu Phe Ile Leu Met Gln
 180 185 190
 Tyr Val Val Lys Val Thr Trp Glu Val Val Met Gln Pro Leu Thr Tyr
 195 200 205
 Lys Ser Ile Ala Trp Ala Arg Lys Asp Gly
 210 215
 <210>28
 <211>194

<212>PRT

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>28

```

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

```

Glu Lys

<210>29

<211>255

<212>PRT

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>29

```

Met Lys Asp Ile Val Asn His Tyr Gly Tyr Leu Ser Pro Tyr Lys Pro
1           5           10           15
Thr Ile Arg Ser Asp Ser Lys Ala Lys Asn Lys Phe Lys Leu Asn Glu
           20           25           30
Pro Tyr Arg Gly Gln Met Leu Ser Ala Gly Ala Gly Gly Ser Ile Met
           35           40           45

```


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

<210>30

<211>261

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>30

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

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

<210>31

<211>255

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>31

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

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

<210>32

<211>254

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>32

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

Phe	Lys	Tyr	Cys	Lys	Asp	Ser	Tyr	Met	Lys	Gly	Arg	Glu	Ala	Thr	Glu	
				85					90					95		
Arg	Arg	Phe	Asn	Arg	Lys	Leu	Glu	Asn	Ile	Ser	Met	Lys	Ala	Asp	Val	
				100					105					110		
Asn	Ile	Thr	Arg	Leu	Glu	Asp	Leu	Phe	Lys	Pro	Asp	Pro	Thr	Ile	Arg	
				115					120					125		
Tyr	Asn	Leu	Asn	Asn	Lys	Val	Phe	Gln	Ala	Ser	Ala	His	Thr	Met	Asp	
				130					135					140		
Arg	Val	Asp	Asn	Arg	Ile	Met	Glu	Asn	Ile	Thr	Gln	Ser	Tyr	Asp	Asp	
				145					150					155		
Gly	Leu	Gly	Ile	Asp	Glu	Ala	Lys	Asp	Arg	Leu	Thr	Val	Glu	Tyr	Asn	
				165					170					175		
Gly	Leu	Lys	Ser	Trp	Glu	Ala	Gln	Arg	Ile	Ala	Arg	Thr	Glu	Ile	Asn	
				180					185					190		
Ser	Ala	Gln	Asn	Asp	Gly	Ala	Phe	Asp	Val	Tyr	Gly	Glu	Leu	Gly	Val	
				195					200					205		
Glu	Tyr	His	Gln	Trp	Trp	Thr	Ala	Gln	Asp	Glu	Arg	Val	Arg	Glu	Thr	
				210					215					220		
Pro	Gln	Ala	Asp	His	Arg	Glu	Leu	His	Gly	Lys	Ile	Val	Lys	Val	Gly	
				225					230					235		
Asn	Ser	Phe	Ser	Asn	Gly	Leu	Gln	Tyr	Pro	Gly	Asp	Arg	Thr			
				245					250							

<210>33

<211>87

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>33

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

<210>34

<211>179

<212>PRT

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>34

```

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

```

Lys Glu Phe

<210>35

<211>96

<212>PRT

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>35

```

Met Asp Ala Ile Asn Val Ile Asn Gln Asn Lys Ile Leu Val Asp Val
1           5           10           15
Leu Tyr Arg Gly Thr Val Asn Leu Ile Asp Ile Val Ile Gly Asp Ala
           20           25           30
Leu Val Tyr Asp Asn Pro Thr Val Val Val Lys Cys Tyr Thr Thr Asp
           35           40           45

```

Leu Ala Phe Ala Thr Thr Glu Ile Asp Glu Ile Val Leu Glu Asn Glu
 50 55 60
 Glu Phe Glu Leu Leu Ile Thr Tyr Glu Asp Gly Glu Phe Ser Ile Leu
 65 70 75 80
 Leu Lys Ser His Asn Leu Gly Glu Leu Gly Tyr Ile Glu Trp Val Ile
 85 90 95

<210>36

<211>101

<212>PRT

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>36

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

<210>37

<211>185

<212>PRT

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>37

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

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

<210>38

<211>63

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>38

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

<210>39

<211>93

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>39

Met Asp Glu Phe Val Glu Thr Leu Phe Asp Thr Tyr Trp Lys Val Asn			
1	5	10	15
Glu Asn Gly Glu Tyr Met Ser Leu Thr Asp Cys Gly Asp Phe Tyr Ile			
	20	25	30
Ala Lys Val Ala Pro Cys Val Arg Asn Trp Ser Ile Val Ile Glu Cys			
	35	40	45

Asn Cys Phe Cys Phe His Cys Lys Glu Phe Val Tyr His Glu Asn Gly
 50 55 60
 Ala Ile Leu Glu Ile Gly Met Glu Ile Ser Ser Leu Tyr Leu Ser Gln
 65 70 75 80
 Met Glu Ile Lys Asp Leu Lys Ile Tyr Met Ile Asp Cys
 85 90

<210>40

<211>119

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>40

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

<210>41

<211>57

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<220>

<221>MOD_RES

<222>(42)..(42)

<223> 任何氨基酸

<400>41

Met Gln Leu Val Val Glu Gly Glu Asn Met Glu Cys Pro Cys Asp Asn
 1 5 10 15
 Cys Glu Met Leu Arg Glu Asn Glu Pro Ile Lys Val Ile Asp Trp Lys

	20		25		30
Ser Asn Ser Pro Phe Gly Asn Gly Ile Xaa Ile His Phe Trp Arg Glu					
35		40		45	
Arg Leu Gln Asp Val Gly Tyr Asn Gly					
50		55			
<210>42					
<211>116					
<212>PRT					
<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)					
<400>42					
Met Glu Leu Met Thr Arg Glu Met Glu Gly Lys Leu Lys Ser Phe Pro					
1	5		10		15
Phe Tyr Ser Gln Asp Gly Lys Gly Asp Asp Ala Ile Val Val Met Lys					
20		25		30	
Phe Phe Asn Pro Tyr Gly Leu Gly Thr Trp Tyr Val Leu Glu Ala Glu					
35		40		45	
Lys Gln Glu Asn Gly Asp Tyr Leu Phe Phe Gly Tyr Val Glu Ser Pro					
50		55		60	
Ile Thr Pro Glu Phe Asn Glu Tyr Gly Tyr Phe Ser Leu Ser Glu Leu					
65		70		75	80
Glu Asn Leu Lys Ile Pro Ile Lys Ile Asn Gly Ile Thr Val Ser Tyr					
85		90		95	
Gly Arg Ile Glu Arg Asp Leu Tyr Phe Glu Arg Val Arg Ile Gly Asp					
100		105		110	
Ile Ile Gly Asn					
115					
<210>43					
<211>74					
<212>PRT					
<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)					
<400>43					
Met Phe Gly Gln Lys Lys Glu Phe Val Lys Lys Met Tyr His Val Gly					
1	5		10		15
Asp Val Val Glu Leu Val His Met Asp Asp Ala Gln Ala Pro Pro Ser					
20		25		30	
Gly Thr Arg Gly Glu Ile Leu Phe Val Asp Asp Ile Gly Gln Ile His					
35		40		45	
Val Arg Trp Glu Asn Gly Ser Gly Leu Ala Leu Ile Tyr Gly Glu Asp					
50		55		60	

Arg Phe Lys Val Val Glu Arg Lys Gly Glu

65

70

<210>44

<211>50

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>44

Met Asp Leu Leu Tyr Leu Tyr Asp Asp Leu Thr Ala Arg Arg Glu Val

1

5

10

15

Tyr Asp Ser Val Gly Leu Ser Phe Val Val Lys Tyr Lys Phe Ser Ser

20

25

30

Arg Lys Glu Ala Gln Asp Phe Ala Leu Lys Tyr Gly Ala Glu Leu Ile

35

40

45

Glu Glu

50

<210>45

<211>62

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>45

Met Ile Glu Arg Arg Arg Leu Gly Met Lys Tyr Asp Ile Phe Thr Ile

1

5

10

15

Leu Asp Glu Ile Ser Arg Lys Leu Asp Asp Gly Glu Leu Ser Asp Glu

20

25

30

Gln Val Asp Phe Leu Leu Gln Met Glu Ile Leu Val Glu Glu Gly Thr

35

40

45

Ile Thr Asp Glu Gln Ala Gln Asp Val Met Asn Gly Asp Tyr

50

55

60

<210>46

<211>55

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>46

Met Met Lys Leu Ser Leu Lys Glu Leu Gly Glu Glu Ile Glu Met Ile

1

5

10

15

Leu Ala Glu Gly Gly Leu Thr Phe Asp Gln Val Asp Tyr Leu Leu Tyr

20

25

30

Leu Glu Thr Cys Ile Ala Asp Gly Ser Ile Thr Glu Glu Gln Lys Arg

35

40

45

Glu Ile Ile Cys Arg Asp Phe

50

55

<210>47

<211>77

<212>PRT

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>47

Met Val Arg Glu Gln Glu Arg Leu Ile Met Tyr Leu Glu Ile Asp Glu

1

5

10

15

Val Lys Cys Glu Asn Ile Asp Arg Ile Glu Phe Asp Asp Val Ala Met

20

25

30

Glu Ile Val Leu Thr Asp Glu Lys Val Tyr Glu Arg Ile Lys Arg Trp

35

40

45

Leu Lys Ser Asn Glu Ile Asp Tyr Asp Cys Arg Glu Asp Arg Tyr Phe

50

55

60

Ala Asn Leu Ile Glu Tyr Val Ile Arg Ile Thr Trp Trp

65

70

75

<210>48

<211>71

<212>PRT

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>48

Met Val Thr Val Arg Asp Val Gly Phe Thr Ile Glu Glu Arg Phe Phe

1

5

10

15

Leu Thr Ala Gln Glu Leu Glu Tyr Ser Glu Val Gly Glu Glu His Glu

20

25

30

Ser Val Ile Asp Arg Ala Ile Ala Leu Leu Tyr Thr Lys Leu Arg Thr

35

40

45

Arg Asp Phe Glu Phe Thr Asp Glu Glu Arg Glu Leu Leu Glu Asp Ala

50

55

60

Phe Val Ile Val Ser Asp Gln

65

70

<210>49

<211>255

<212>PRT

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>49

Met Ala Lys Glu Asn Val Ile Asp Tyr Lys Ile Glu Arg Gln Asn Asp

1

5

10

15

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

<210>50

<211>231

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>50

Met Val Val Lys Cys Pro Asn Cys Phe Ser Pro Arg Val Ser Lys Cys		
1	5	10
Glu Asp Thr Asn Ile Lys Trp Gln Cys Asp Lys Cys Lys Cys Lys Phe		
20	25	30

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

<210>51

<211>248

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>51

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

Ala Thr Thr Arg Ile Thr Asp Glu Glu Glu Met Glu Lys Ala Leu Asn			
65	70	75	80
Gly Glu Tyr Thr Gly Tyr Ser Ile Thr Thr Val Ser Lys Lys Phe Ala			
	85	90	95
Asp Lys Gln Ile Gln Leu Pro Arg Arg Val Leu Met Lys Asp Ile Lys			
	100	105	110
Asp Pro Val Gly Phe Thr Ile Ser Leu Val Arg Lys Pro Cys Val Arg			
	115	120	125
Gly Ala Lys Phe Cys Ser Met Lys Glu Asp Ile Glu Asn Gly Asp Val			
	130	135	140
Val Ser Glu Asn Ile Asp Asp Lys Leu Glu Glu Glu Thr Lys Gly Phe			
145	150	155	160
Val Gln Ser Ile Lys Gly Ile Phe Asn Lys Glu Asp Lys Asp Glu Asp			
	165	170	175
Lys Asn Pro Glu Asp Glu Asp Ile Glu Leu Asp Ile Lys Ala Ile Val			
	180	185	190
Asp Glu Val Thr Lys Asp Phe Val Asn Thr Asp Asp Phe Glu Thr Phe			
	195	200	205
Lys Asn Glu Leu Glu Lys Ala Leu Ser Asp Lys Phe Glu Thr Leu Gly			
	210	215	220
Ala Glu Leu Phe Lys Ser Leu Lys Lys Ser Leu Glu Lys Asp Lys Ala			
225	230	235	240
Glu Glu Ala Lys Lys Ser Gly Arg			
	245		

<210>52

<211>280

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>52

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

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

<210>53

<211>147

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>53

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

〈213〉 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>55

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

145

<210>56

<211>169

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>56

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

	100		105		110
Lys Val Asp Asp Ser Met Pro Asn Arg Phe Phe His Lys Leu Glu Phe					
	115		120		125
Glu Thr Leu Tyr Pro Val Gly Glu Val Thr Val Val Gly Lys Ser Glu					
	130		135		140
Arg Ile Pro Ala Thr Ser Ile Ile Phe Asn Phe Val Tyr Val Val Asp					
145		150		155	160
Trp Leu Lys Cys Asn Arg Arg Tyr Asp					
	165				
<210>57					
<211>255					
<212>PRT					
<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)					
<400>57					
Met Gly Ile Arg Val Val Gly Met Lys Glu Glu Ala Arg Tyr Gly Val					
1	5		10		15
Ala Glu Ser Ala Pro Asp Phe His Gln Glu Val Ser Lys Ala Lys Ala					
	20		25		30
Ser Leu Asn Ser Thr Pro Asn Thr Lys Ser Ser Gly Ser Arg Met Lys					
	35		40		45
Lys Lys Ala Arg Ala Gly Val Tyr Lys Pro Thr Ala Asn Ile Glu Gly					
	50		55		60
Glu Val Asp Leu Lys Arg Ile Gly His Tyr Leu Lys Ala Phe Leu Asp					
65		70		75	80
Asn Tyr His Phe Thr Asp Gly Gly Ser Asn Pro Asn Val His Glu Phe					
	85		90		95
Trp Gly Gly Glu Asn Asn Lys Leu Ser Ser Phe Thr Leu Trp Val Thr					
	100		105		110
Phe Asp Ile Phe Glu Lys Thr Ile Val Gly Ser Leu Leu Asp Asn Phe					
	115		120		125
Lys Met Glu Val Ser Asp Glu Tyr Met Lys Phe Thr Ala Asp Phe Val					
	130		135		140
Tyr Lys Thr Glu Glu Ser Asp Glu Ile Glu Asn Ile Glu Leu Tyr Lys					
145		150		155	160
Val Lys Leu Leu Asp Gly Asp Trp Ala Leu Met Phe Tyr Asp Val Ser					
	165		170		175
Val Glu Ile Asp Glu Asn Ala Pro Pro Gly Ile Val Ser Ser Phe Ser					
	180		185		190
Phe Asp Gly Lys Asn Asn Ile Asn Val Asp Lys Thr Ile Gly Leu Gly					

195	200	205
Ser Arg Gly Pro Gln Arg Lys Ala Ala Ala Gln Gly Arg Asp Ile Ser		
210	215	220
Ile Ser Phe Val Ser Thr Leu Glu Arg Glu Thr Leu Glu Leu Ile Gln		
225	230	235
Lys Ala Glu Tyr Gly Glu Val Gly Thr Glu Pro Ser Glu Cys Lys		
245	250	255

<210>58

<211>146

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>58

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

Phe His

145

<210>59

<211>255

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>59

Met Pro Ser Ser Asn Val Met Asn Ile Ile Val Lys Ala Glu Asp Met		
1	5	10
		15

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

<210>60

<211>274

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>60

Met Ala Ser Val Thr Lys Tyr Pro Ser Asn Val Ser Gln Thr Thr Gly			
1	5	10	15
Gly Lys Phe Val Ser Phe Ser Asn Leu Ala Asn Ile Lys Asn Asn Ala			
20	25	30	

Asp Gly Ala His Ala Val Ser Ser Val Leu Ile Lys Ser Lys Lys Gln		
35	40	45
Ser Pro Asn Arg Pro Ser Thr Val Ser Cys Lys Gly Phe Gly Phe Ser		
50	55	60
Leu Pro Glu Gly Ala Glu Pro Thr Lys Ile Thr Val Thr Tyr Arg His		
65	70	75
Arg Lys Asn Ala Gly Ser Asp Tyr Ser Ser Lys Asn Lys Thr His Ile		
85	90	95
Cys Asn Ile Gly Gly Pro Thr Ile Ser Leu Leu Gly Val Ser Gly Phe		
100	105	110
Ser Ser Lys Gly Ser Gly Cys Thr Thr Thr Met Thr Thr His Thr Lys		
115	120	125
Ala Phe Ser Val His Gly Lys Leu Ser Arg Ala Gln Val Asn Ser Ala		
130	135	140
Asn Phe Gly Val Lys Leu Asp Tyr Pro Thr Asn Ser Asn Thr Tyr Asn		
145	150	155
Gly Tyr Met Arg Ile Ser Tyr Val Arg Val Thr Val Glu Tyr Ile Thr		
165	170	175
Ser Gln Tyr Ser Val Ser Val Lys His Val Ser Gly Thr Tyr Glu Asp		
180	185	190
Glu Asp Tyr Val Val Ser Leu Gly Ile Ser Asn Lys Asn Leu Thr Ser		
195	200	205
Tyr Asn Pro Thr Cys Thr Leu Thr Val Pro Ala Gly Phe Thr Tyr Lys		
210	215	220
Gly Val Thr Gly Ala Ala Thr Gly Thr Val Thr Lys Val Asn Asn Arg		
225	230	235
Thr Phe Ser Trp Asn Pro Gln Leu Gln Gly Arg Ala Gly Ser Arg Gln		
245	250	255
Ile Ser Leu Ala Phe Glu Pro Asn Val Thr Phe Pro Glu Gly Thr Asp		
260	265	270
Ser Tyr		
<210>61		
<211>255		
<212>PRT		
<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)		
<400>61		
Met Gly Ile Ala Ile Val Val Met Asp Asn Glu Glu Asn Phe Leu Gln		
1	5	10
Phe Leu Asp Pro Asp Leu Cys Thr Ile Asn Glu Thr Ile Glu Glu Leu		

Met	Val	Glu	Lys	Ile	Thr	Val	Ser	Pro	Gln	Glu	Val	Arg	Gly	Tyr	Gly
1				5					10					15	
Asn	Val	Val	Asp	Glu	Lys	Glu	Leu	Glu	Asp	Tyr	Gly	Ser	Tyr	Arg	Cys
			20					25					30		
Asp	Val	Ser	Glu	Ser	Ser	Glu	Val	Ile	Lys	Gly	Val	Glu	Glu	Arg	Ile

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

50	55	60
Glu Pro Asn Tyr Val	Thr Leu Thr Ser Val	Ala Arg Asn Pro Leu Val
65	70	75
Met Asp Tyr Gln Asn	Thr Asn Tyr Thr Cys Cys	Pro Thr Ser Leu Ser
85	90	95
Leu Ala Ser Gln Met	Leu Tyr His Tyr Lys	Ser Glu Ser Glu Cys Ala
100	105	110
Lys Ala Leu Gly Thr	Ser Lys Gly Ser Gly	Thr Ser Pro Ala Gln Leu
115	120	125
Ile Ala Asn Ala Pro	Lys Leu Gly Phe Lys	Ile Ile Pro Ile Lys Arg
130	135	140
Asp Ser Lys Glu Val	Lys Lys Tyr Leu Lys	Lys Gly Phe Pro Val Ile
145	150	155
Cys His Trp Gln Val	Asn Gln Ser Arg Asn	Cys Lys Gly Asp Tyr Thr
165	170	175
Gly Asn Phe Gly His	Tyr Gly Leu Ile Trp	Asp Met Thr Ser Thr His
180	185	190
Tyr Val Val Ala Asp	Pro Ala Lys Gly Val	Asn Arg Lys Tyr Lys Phe
195	200	205
Ser Cys Leu Asp Asn	Ala Asn Lys Gly Tyr	Arg Gln Asn Tyr Tyr Val
210	215	220
Val Cys Pro Ala		
225		
<210>64		
<211>749		
<212>PRT		
<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)		
<400>64		
Met Lys Lys His Cys	Phe Tyr Phe Leu Gly	Asp Ser Phe Ala Asp Ile
1	5	10
Cys Asn Glu Ala Met	Phe Cys Glu Lys His	Leu Val Glu Gly Asn Tyr
20	25	30
Leu Asp Ser Ile Ile	Arg Ala Gly Lys Ala	Ser Glu Ile Ile Thr Val
35	40	45
Asn Ile Cys Glu Leu	Glu Gly Gln Asp Gly	Leu Ile Ser Ser Gly Gln
50	55	60
Lys Lys Arg Leu Glu	Met Leu Gly Tyr Lys	Gly Ile Ile Ser Tyr Asp
65	70	75
Ile Tyr Lys Arg Leu	Asn His Ile Arg Lys	Ile Arg Asn Lys Ala Val

				85				90				95			
His	Gly	His	Leu	Ser	Asp	Ile	Glu	Asp	Asn	Ala	Asn	Ile	Leu	His	Ala
				100				105				110			
Tyr	Leu	Tyr	Leu	Ile	Cys	Ala	Tyr	Phe	Tyr	Lys	Glu	Tyr	Arg	Asp	Thr
				115				120				125			
Asn	Phe	Ser	Ala	Glu	Asp	Tyr	Thr	Gly	Pro	Ile	Met	Asp	Ile	Ala	Ser
				130				135				140			
Lys	Pro	Lys	Glu	Thr	Ala	Ser	Glu	Thr	Ser	Glu	Asp	Asn	Glu	Asn	Ile
145				150				155				160			
Gly	Glu	Phe	Ile	Ser	Ser	Pro	Leu	Asp	Asp	Tyr	Leu	Phe	Glu	Lys	Tyr
				165				170				175			
Asp	Asp	Ser	Tyr	Leu	Leu	Asn	Glu	Leu	Ser	Lys	Leu	Lys	Asp	Ser	Ser
				180				185				190			
Lys	Glu	Ala	Val	Glu	Asp	Asp	Asn	Leu	Ser	Glu	Phe	Lys	Glu	Tyr	Leu
				195				200				205			
His	Ile	Asp	Arg	Ser	Ile	Gln	Glu	Asp	Phe	Leu	Lys	Ala	Leu	Asn	Arg
210				215				220							
Ala	Thr	Ser	Phe	Asn	Ser	Ser	His	Leu	Ile	Met	Leu	Cys	Gly	Ser	Val
225				230				235				240			
Gly	Asp	Gly	Lys	Ser	His	Leu	Ile	Ala	Asn	Leu	Lys	Lys	Lys	Asn	Pro
				245				250				255			
Glu	Leu	Phe	Asn	Gln	Phe	Ala	Ile	His	Tyr	Asp	Ala	Thr	Glu	Ser	Phe
				260				265				270			
Asp	Pro	Glu	Lys	Asn	Ala	Ile	Asp	Thr	Leu	Ala	Ser	Val	Leu	Glu	Pro
				275				280				285			
Phe	Asn	Asp	Asn	Asn	Leu	Asn	Asn	Ser	Thr	Glu	Lys	Leu	Ile	Leu	Ala
290				295				300							
Ile	Asn	Leu	Gly	Val	Leu	Asn	Asn	Phe	Leu	Glu	Ser	Ser	Tyr	Ala	Asn
305				310				315				320			
Glu	Asp	Tyr	Thr	Lys	Leu	Lys	Leu	Ile	Ile	Glu	Glu	Ala	Asn	Ile	Phe
				325				330				335			
Glu	Ser	Asn	Glu	Val	Ser	Asp	Asn	Ile	Tyr	Gly	Asp	Lys	Val	Ser	Phe
				340				345				350			
Val	Thr	Phe	Ser	Asp	Tyr	Asn	Met	Phe	Glu	Leu	Asn	Asp	Asp	Glu	Asn
				355				360				365			
Ser	Asn	Tyr	Thr	Ser	Ser	Lys	Tyr	Ile	Ser	Ser	Leu	Phe	Asn	Lys	Ile
370				375				380							
Thr	Gln	Lys	Glu	Asp	Thr	Asn	Pro	Phe	Tyr	Val	Ala	Tyr	Leu	Lys	Asp
385				390				395				400			

Lys Asp Ser His Phe Ile Asn Pro Ile Ile Tyr Asn Tyr Glu Met Leu			
405	410	415	
Met Asp Glu Glu Val Gln Lys Thr Ile Ile Asp Tyr Leu Ile Lys Ile			
420	425	430	
Phe Ile Lys Tyr Arg Lys Ile Ile Ser Thr Arg Asp Leu Leu Asn Phe			
435	440	445	
Ile Tyr Glu Ile Ile Val Pro Pro Glu Phe Leu Lys Ser Glu Asp Leu			
450	455	460	
Asp Asn Ile Asn Asp Phe Met Asp Tyr Ser Leu Pro Asn Leu Leu Phe			
465	470	475	480
Gly Tyr Pro Glu Arg Ser Asp Leu Leu Lys Leu Cys Asn Glu Leu Asp			
485	490	495	
Pro Thr Leu His Arg Asn Glu Ser Leu Asp Lys Phe Ile Ile Asp Leu			
500	505	510	
Asn Ile Asn Asp Asp Thr Glu Lys Ile Leu Asn Arg Tyr Phe Asp Phe			
515	520	525	
Thr Arg Phe Asn Phe Leu Glu Glu Tyr Gly Glu Tyr Leu Val Asp Phe			
530	535	540	
Arg Glu Phe Asn Asn Ser Glu Lys Glu Lys Val Thr Asn Ile Leu Ile			
545	550	555	560
Arg Phe Ala Val Phe Tyr Gly Lys Ser Ile Ile Lys Asn Asn Phe Lys			
565	570	575	
Asp Lys Val Tyr Leu Asn Tyr Leu Lys Tyr Leu Tyr Ala Tyr Asn Thr			
580	585	590	
Gln Ser His Lys Asp Tyr Lys Tyr Leu Phe Thr Glu Val Lys Asp Ala			
595	600	605	
Ile Phe Asn Trp Lys Gly Ser Tyr Lys Lys Asn Thr Ile Cys Ile Asp			
610	615	620	
Thr Leu Asp Ser Phe Lys Val Tyr Lys Asn Leu Lys Leu Lys Pro Ser			
625	630	635	640
Val Asp Lys Phe Glu Lys Ser Leu Leu Asp Gly Leu Phe Leu Gly Asn			
645	650	655	
Arg Phe Lys Thr Asp Ile Lys Ile Tyr Phe Ser Val Glu Ser Asn Lys			
660	665	670	
Lys Lys Ile Pro Leu Asn Val Asp Phe Ser Leu Tyr Gln Tyr Ile Met			
675	680	685	
Lys Leu Tyr Asn Gly Phe Lys Pro Asn Gln Ser Asp Lys Asp Asp Leu			
690	695	700	
Ile Ile Leu Asp Glu Phe Ile Asn Asn Leu Leu Asp Glu Asp Thr Asp			

705	710	715	720
Asp Asp Leu Tyr Val	Ile Ser Leu Glu Thr Tyr Glu Glu Phe Leu Phe		
	725	730	735
Glu Ser Asn Asp Phe Gly Thr Phe Glu Phe Lys Arg Gly			
	740	745	
<210>65			
<211>456			
<212>PRT			
<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)			
<400>65			
Met Asp Phe Ser Glu Asn Tyr Asn Ile Leu Leu Lys Gln Met Thr Cys			
1	5	10	15
Asp Val Asn Lys Arg Lys Leu Ile His Gln Ile Asn Gln Asn Ser Pro			
	20	25	30
Leu Leu Pro Phe Lys Thr Asn Thr Pro Lys Lys Ala Asn Phe Glu Asn			
	35	40	45
Gly Phe Asp Ile Ile Leu Gly Glu Leu Ser Arg Ile Leu Leu Asn Lys			
	50	55	60
Thr Ile Glu Lys Asn Phe Lys Leu Asp Asn Ile Val Ser Asn Leu Ile			
65	70	75	80
Asp Asn Asn Ile Glu Ile Glu Asp Gly Thr Lys Glu Tyr Ile Thr Lys			
	85	90	95
Leu Leu Asn Glu Tyr Leu Phe Asp Glu Lys Asn Asp Leu Lys Ile Ser			
	100	105	110
His Pro Asn Leu Tyr Leu Tyr Ile Pro Leu Ser Asn Asn Lys Ser Ser			
	115	120	125
Asn Gly Glu Gln Glu Val Ala Leu Phe Leu Arg Asp Ile Phe Cys Lys			
	130	135	140
Asn Asn Gln Asn Leu Ile Asn Phe Phe Glu Ser Tyr Asp Ser Asn His			
145	150	155	160
Ile Ile Leu Asn Leu Ile Leu Lys Asn Thr Pro Asn Leu His His Lys			
	165	170	175
Ile Thr Glu Thr Lys Tyr Val Ile His Phe Glu Glu Ile Ala Asn Leu			
	180	185	190
Phe Asn Glu Asp Ile Asn Tyr Ala Ile Leu Tyr Lys Lys Phe Phe Met			
	195	200	205
Glu Asn Ile Gly Asn Ile Phe Ala Tyr Tyr Tyr Phe Phe Tyr Ile Ser			
	210	215	220
Gln Leu Ile Leu Lys Ile Ser Lys Gly Phe Asn Asp Asn Asn Glu Phe			

225	230	235	240
Glu Lys Leu Tyr Tyr Leu Leu Asp Trp Glu Ser Ala Ser Lys Asn Arg			
	245	250	255
Lys Ser Leu Asn Ser Tyr Ser Leu Leu Lys His His Ser Lys Pro Leu			
	260	265	270
Tyr Ala Lys Met Ala Val Ile Asp Gln Ile Asn Thr Leu Leu Gly Thr			
	275	280	285
Asn Asn Leu Leu Glu Lys Asp Ile Ser Glu Tyr Phe Asn Asn Leu Asp			
	290	295	300
Ile Asn Ser Lys Asn Asn Phe Leu His Phe Leu Lys Lys Trp Val Ser			
305	310	315	320
Asp Tyr Arg Tyr Val Arg Asn Phe Asp Asp Lys Glu Leu Pro Asp Asn			
	325	330	335
Leu Leu Glu Leu Thr Glu Ile Leu Phe Glu Ser Leu Lys Asn Glu Lys			
	340	345	350
Leu Gly Val Asp Gly Ala Val Gln Ser Arg Tyr Ala Leu Asn Leu Glu			
	355	360	365
Asp Ile Ala Lys Lys Tyr Leu Leu Lys Arg Arg Gly Ser Tyr Gly Tyr			
	370	375	380
Val Leu Asn Ile Asn Arg Asp Met Leu Leu Val Leu Thr Ala Leu Cys			
385	390	395	400
Val Lys Asp Lys Lys Ile Lys Leu Asn Gln Leu Phe Ile Glu Phe Glu			
	405	410	415
Lys Arg Gly Val Tyr Phe Asp Lys Tyr Ser Lys Glu Glu Val Val Asn			
	420	425	430
Phe Leu Thr Lys Leu Asn Leu Ile Asp Lys Lys Ser Asp Ser Gly Asp			
	435	440	445
Ala Gln Tyr Val Lys Pro Val Leu			
450	455		
<210>66			
<211>1740			
<212>PRT			
<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)			
<400>66			
Met Leu Asn Gln Phe Tyr Asp Tyr Leu Ser Asn Lys Leu Leu Asn Tyr			
1	5	10	15
Phe Asp Asp Thr Lys Ile Leu Ser Gly Glu Lys Phe Phe Ile Ser Phe			
	20	25	30
Asp Glu Asp Asp Gln Ile Met Ser Phe Tyr Asn Ser Leu Arg Ser Ile			

35					40					45						
Ala	Glu	Thr	Asn	Phe	Ser	Cys	Ser	Glu	Phe	Ile	Tyr	Val	His	Thr	Ile	
50					55					60						
Ser	Gly	Lys	Glu	Tyr	Asn	Thr	Tyr	Ser	Ile	Asn	Ile	Asn	Gly	Val	Lys	
65					70					75					80	
Phe	Val	Ile	Ser	Glu	Ser	Leu	Thr	Ile	Asn	Val	Asp	Phe	Leu	Val	Thr	
85					90					95						
Leu	Arg	Asn	Gln	Val	Thr	Ser	Gln	Glu	Gly	Val	Trp	Lys	Asp	Thr	Ala	
100					105					110						
Leu	Leu	Val	Ile	Cys	Asn	Glu	Ala	Ile	Asp	Ser	Ile	Gly	Lys	Gly	Met	
115					120					125						
Arg	Asn	Leu	Gln	Lys	Glu	Gly	Met	Pro	Leu	Asn	Val	Lys	Ser	Ile	Ser	
130					135					140						
Lys	Asn	Leu	Glu	Asp	Glu	Ile	Asn	Asp	Ser	Gln	Ile	Leu	Asn	Tyr	Ser	
145					150					155					160	
Asp	Lys	Gln	Ile	Ala	Lys	Phe	Ser	Leu	Asn	Ile	Gln	Glu	Glu	Glu	Leu	
165					170					175						
Phe	Gln	Thr	Thr	Leu	Trp	Asp	Tyr	Glu	Thr	Ile	Leu	Ser	Ile	Ile	Asn	
180					185					190						
Lys	Gly	Phe	Val	Ser	Asp	Glu	Asp	Leu	Arg	Glu	Leu	Asn	Leu	Phe	Lys	
195					200					205						
Asp	Asp	Gln	Leu	Asn	Gln	Asn	Ser	Pro	Gln	Lys	Met	Leu	Lys	Arg	Leu	
210					215					220						
Lys	Glu	Asn	Tyr	Asp	Thr	Phe	Asn	Glu	Val	Asn	Lys	Phe	Ser	Gln	Tyr	
225					230					235					240	
Gly	Asp	Lys	Lys	Glu	Gln	Leu	Lys	Asn	Met	Phe	Thr	Asp	Ser	Gly	Val	
245					250					255						
Ser	Ile	Leu	Ser	Lys	Asp	Asp	Trp	Tyr	Lys	Ala	Glu	Trp	Lys	Met	Val	
260					265					270						
Lys	Lys	Ser	Lys	Asp	Asp	Phe	Ile	Asn	Gln	Gln	Asn	Pro	Leu	Asn	Tyr	
275					280					285						
Asn	Glu	Asn	Leu	Glu	Lys	Ile	Thr	Glu	Asn	Gly	Leu	Asn	Tyr	Trp	Glu	
290					295					300						
Ile	Pro	Asn	Ser	Phe	Thr	Lys	Thr	Gly	Lys	Arg	Lys	Arg	Asn	Ile	Ile	
305					310					315					320	
Val	Phe	Asn	Pro	Asn	His	Ser	His	Glu	Val	Ser	Leu	Lys	Phe	Ser	Phe	
325					330					335						
Asp	Gln	Ile	Leu	Ser	Asn	Ser	Phe	Leu	Asn	Thr	Asn	Ser	Lys	Lys	Phe	
340					345					350						

Thr	Ile	Ala	Arg	Gly	Lys	Ser	Leu	Ile	Val	Asn	Phe	Thr	Leu	Asp	Ser			
		355					360					365						
Ser	Glu	Pro	Ile	Phe	Lys	Thr	Ile	Lys	Tyr	Lys	His	Lys	Asn	Glu	Asn			
		370					375				380							
Ile	Ser	Glu	Phe	Thr	Phe	Asn	Ile	Val	Val	Leu	Asn	Phe	Glu	Pro	Glu			
385					390					395				400				
Ile	Phe	Asn	Ser	Ile	Lys	Ser	Arg	Phe	Ser	Val	Asn	Val	Lys	Ser	Lys			
			405					410					415					
Gln	Ile	Ile	Val	Thr	Asn	Asp	Glu	Asp	Ser	Phe	Asp	Ile	Val	Phe	Gly			
			420					425					430					
Thr	Gly	Ser	Lys	Glu	Ile	Glu	Lys	Leu	Ile	Lys	Glu	Asn	Gly	Glu	Lys			
		435					440					445						
Leu	Tyr	Leu	Tyr	Asp	Asp	Glu	Ser	Leu	Ile	Ile	Ser	Glu	Gln	Ser	Pro			
	450					455					460							
Ala	Trp	Asn	Asp	Gly	Lys	Leu	Ser	Phe	Lys	Leu	Tyr	Lys	Asp	Asn	Asn			
465					470					475				480				
Tyr	Ala	Pro	Phe	Leu	Ile	Lys	Glu	Lys	Ser	Lys	Lys	Thr	Leu	Pro	Val			
			485					490					495					
Asn	Ser	Tyr	Val	Ile	Trp	Asn	Leu	Lys	Arg	Arg	Asn	Met	Glu	Asn	Phe			
		500					505					510						
Ile	Phe	Asn	Gly	Val	Lys	Ala	Val	Gln	Asp	Val	Asn	Ser	Phe	Tyr	Leu			
		515					520					525						
Val	Glu	Glu	Phe	Lys	Glu	Phe	Leu	Lys	Met	Glu	Arg	Glu	Ile	Ile	Lys			
	530					535					540							
Gln	Asp	Ile	Phe	Tyr	Ala	Lys	Arg	Asn	Ile	Asp	Gly	Ser	Leu	Glu	Lys			
545					550					555				560				
Ile	Glu	Val	Ser	Phe	Ser	Asn	Glu	Leu	Glu	Thr	Ala	Tyr	Met	Asp	Ile			
			565					570				575						
Phe	Asn	Tyr	Tyr	Lys	Thr	Phe	Asp	Asp	Ser	Pro	Glu	Asp	Asn	Leu	Pro			
		580					585					590						
Ser	Leu	Val	Tyr	Leu	Asn	Asp	Asp	Leu	Lys	Glu	Leu	Tyr	Lys	Lys	Phe			
		595					600					605						
Ile	Thr	Ile	Phe	Asn	Lys	Glu	Ile	Ser	Glu	Ile	Glu	Glu	Asn	Ser	Ile			
	610					615					620							
Leu	Ser	Asp	Phe	Lys	Tyr	Lys	Lys	Asn	Leu	Leu	Lys	Leu	Gly	Arg	Ile			
625					630					635				640				
Glu	Thr	Asp	Asn	Lys	Ile	Met	Tyr	Ser	Pro	Leu	Ser	Pro	Leu	Asn	Ile			
			645					650				655						
Ala	Tyr	Gln	Leu	Glu	Val	Ser	Lys	Gln	Cys	Gly	Asn	Glu	Asp	Leu	Ser			

660	665	670
Val Asn Ile Leu Glu Arg Leu	Val Pro Asn Asn Leu Ile	Pro Tyr Ile
675	680	685
Cys Ser Asp Asp Gly Lys Glu	Leu Phe Arg Pro Ile Tyr	Gln Glu Glu
690	695	700
Ala His Glu Trp Leu Ile Tyr	Glu Lys Ser Glu Glu Val	Ser Ile Gly
705	710	715
Thr Thr Asn Val Phe Ile Ser	Asn Val Val Thr Glu Lys	Leu Asn Gln
725	730	735
Phe Val Lys His Phe Asn Tyr	Leu Phe Ser Phe Asn Asn	Ser Ser Pro
740	745	750
Ile Lys Ile Asn Leu Ile Asn	Ile Lys Asp Asp Lys Glu	Val Val Lys
755	760	765
Gly Val Phe Asn Phe Ile Arg	Ser Arg Leu Pro Asp Lys	Thr Lys Thr
770	775	780
Lys Lys Val Ile Pro Val Glu	Ile Asn Ile Tyr Asn Asp	Ala Glu Lys
785	790	795
Ser Ser Phe Asp Asn Leu Phe	Asp Cys Gln Ser Glu Ile	Gln Leu Leu
805	810	815
Glu Glu Phe Gly Ile Lys Lys	Leu Lys Ser Asp Ile Phe	Asp Pro Ile
820	825	830
Asp Ile Ile His Met Ile Gln	Asn Asn Ile Ser Tyr Tyr	Lys His Pro
835	840	845
Phe Lys Lys Glu Glu Tyr Glu	Tyr Ala His Leu Ser Phe	Tyr Lys Val
850	855	860
Lys Ser His Asn Asn Ile Ala	Asn Asp Asn Met Asp Lys	Ile Glu Thr
865	870	875
Gly Leu Ser Leu Asn Gly Leu	Leu Ser Ser Val Thr Ser	Thr Thr Lys
885	890	895
His Ser Glu Tyr Arg Thr Gly	Phe Gly Thr Asn Asn Ile	Leu Asn Met
900	905	910
Ser Asn Pro Leu Ile Lys Thr	Val Ile Asn Leu Asn Glu	Leu Val Glu
915	920	925
Asn Ser Lys Asn Phe Gly Lys	Asn Thr Tyr Ser Lys Asn	Lys Ser Val
930	935	940
Ile Thr Thr Val Glu Leu Glu	Glu Asp Asn Ile Glu Glu	Leu Tyr Asp
945	950	955
Lys Ser His Trp Val Thr Phe	Ile Glu Pro Thr Phe Gly	Ile Glu Tyr
965	970	975

Phe Asp Ser Ser Asp Ser Asn Leu Ile Ile Ile His Tyr Ser Asp Gln		
980	985	990
Tyr Ser Ser Ser Ser Lys Tyr Asp Thr Ile Thr Val Thr Asn Lys Ser		
995	1000	1005
Thr Gln Tyr Glu Glu Ile Ile Arg Asp Phe Leu Gln Ser Lys Tyr		
1010	1015	1020
Val Lys Val Thr Asp Glu Glu Leu Tyr Asp Val Ile Lys Met Phe		
1025	1030	1035
Asn Ser Ile Asn Gly Glu Trp Leu Leu Arg Val Ile Ser Asn Ser		
1040	1045	1050
Gly His Tyr Asp Arg Glu Lys Leu Ser Ile Ile Ser Ala Ile Lys		
1055	1060	1065
Tyr Cys Leu Ser Ile Leu Asp His Lys Asp Ile Val Trp Ile Pro		
1070	1075	1080
Val Ser Met Glu Glu Ile Leu Arg Ile Ala Gly Asn Val Lys Leu		
1085	1090	1095
Asp Lys Asn Lys Gly Ile Phe Asp Ser Lys Leu Ile Lys Gly Asn		
1100	1105	1110
His Ser Asp Asp Leu Leu Phe Ile Gly Val Lys Phe Asn Glu Asp		
1115	1120	1125
Asn Arg Ile Glu Val Ile Phe Tyr Pro Ile Glu Val Lys Ile Gly		
1130	1135	1140
Leu Asn Asn Ala Ser Thr Ile Lys Lys Gly Lys Ser Gln Leu Asp		
1145	1150	1155
Asn Thr Tyr Lys Leu Leu Lys Thr Gln Leu Gln Asn Ile Asn Val		
1160	1165	1170
Glu Asn Ser Glu Phe Arg Asn Lys Phe Phe Arg Asn Phe Phe Ile		
1175	1180	1185
Gln Ile Leu Leu Ser Asn Glu Gln Lys Leu Val Thr Asn His Ile		
1190	1195	1200
Trp Asp Glu Lys Gly Leu Asp Arg Ile Glu Glu Phe Lys Ala Glu		
1205	1210	1215
Leu Leu Asn Asp Glu Tyr Asp Ile Leu Tyr Gly Leu Glu Glu Tyr		
1220	1225	1230
Ile Gly Lys Gly Ser Leu Val Ser Phe Lys Lys Glu Ser His His		
1235	1240	1245
Ile Ser Ile Tyr Met Asp Val Asp Lys Gln Val Ile Glu Leu Pro		
1250	1255	1260
Glu Asp Phe Ala Tyr Tyr Gly Leu Ala Thr Pro Ile Arg Glu Ile		

1265	1270	1275
His Asp Glu Ile Gln Ser Asp	Asn Thr Asp Ile Leu	Ala Glu Thr
1280	1285	1290
Leu Leu Ser His Val Asp Ile	Ser Glu Ile Arg Ala	Lys Asn Asn
1295	1300	1305
Asp Ile Cys Asp Ser Asn Glu	Asp Met Ser Ile Asp	Asp Asp Phe
1310	1315	1320
Asp Asn Leu Ser Glu Phe Glu	Asp Ser Phe Ile Glu	Glu Glu Ser
1325	1330	1335
Glu Ile Ser Glu Glu Pro Asp	Glu Glu Leu Thr Asp	Ala Thr Ser
1340	1345	1350
Ser Ser Asp Asn Glu Ser Ile	Glu Asn Ile Gly Glu	Ser Pro Ser
1355	1360	1365
Lys Ile Ser Asn Val Arg Ala	Leu Ile Gly Thr Gln	Lys Gly Tyr
1370	1375	1380
Asn His Lys Val Tyr Trp Glu	Phe Gly His Pro Ser	Leu Ala Asn
1385	1390	1395
Arg His Met Leu Ile Gln Gly	Lys Ser Gly Gln Gly	Lys Thr Tyr
1400	1405	1410
Phe Ile Gln Arg Met Leu Lys	Glu Leu Ser Ile Gln	Gly Ile Pro
1415	1420	1425
Ser Ile Ile Ile Asp Tyr Thr	Asp Gly Phe Lys Pro	Ser Gln Leu
1430	1435	1440
Glu Pro Asn Phe Lys Asp Ser	Leu Gly Asp Lys Ile	Ser Gln Tyr
1445	1450	1455
Phe Val Val Lys Glu Asn Phe	Pro Ile Asn Pro Phe	Lys Arg Asn
1460	1465	1470
Thr Ile Met Ile Asp Lys Asp	Ile Phe Ile Glu Glu	Asp Asn Ser
1475	1480	1485
Thr Ile Ala Ser Arg Phe Lys	Ser Ile Ile Asn Ser	Val Tyr Gly
1490	1495	1500
Leu Gly Ile Gln Gln Ser Asn	Thr Leu Tyr Gln Thr	Val Leu Asp
1505	1510	1515
Cys Leu Asp Lys Tyr Asp Asp	Asn Phe Asp Leu Asn	Ile Leu Lys
1520	1525	1530
Glu Glu Ile Leu Lys Asp Glu	Ser Asn Ser Ala Gln	Thr Val Leu
1535	1540	1545
Asn Lys Leu Asn Glu Leu Leu	Asp Lys Asn Pro Phe	Ala Ser Thr
1550	1555	1560

Asp	Phe	Asp	Trp	Ser	Val	Leu	Asp	Asn	Lys	Asp	Gly	Lys	Val	Tyr
1565						1570		1575						
Ile	Ile	Gln	Leu	Thr	Ala	Leu	Ser	Lys	Asp	Ile	Gln	Thr	Ile	Ile
1580						1585		1590						
Thr	Glu	Phe	Ile	Leu	Trp	Asp	Leu	Trp	Asn	Tyr	Lys	Leu	Thr	Asn
1595						1600		1605						
Gly	Ser	Glu	Asp	Asn	Pro	Phe	Ile	Val	Val	Leu	Asp	Glu	Ala	His
1610						1615		1620						
Asn	Leu	Asp	Phe	Ser	Asn	Asp	Ser	Pro	Cys	Ser	Lys	Ile	Leu	Lys
1625						1630		1635						
Glu	Gly	Arg	Lys	Phe	Gly	Trp	Ser	Gly	Trp	Phe	Ala	Thr	Gln	Ser
1640						1645		1650						
Val	Lys	Gly	Ser	Met	Lys	Ile	Asp	Glu	Ile	Ala	Lys	Leu	Glu	Asn
1655						1660		1665						
Ala	Asp	Glu	Lys	Ile	Tyr	Phe	His	Pro	Thr	Asp	Val	Ser	Thr	Ile
1670						1675		1680						
Ala	Lys	Asp	Leu	Ser	Lys	Asp	Asn	Glu	Asp	Lys	Lys	Ile	Tyr	Glu
1685						1690		1695						
Lys	Glu	Leu	Ser	Gln	Leu	Thr	Lys	Gly	Tyr	Cys	Ile	Val	Gln	Gly
1700						1705		1710						
Ser	Ala	Ile	Asp	Ser	Ser	Gly	Asn	Leu	Tyr	Gln	Pro	Asn	Pro	Val
1715						1720		1725						
Thr	Val	Lys	Ile	Glu	Glu	Ile	Ser	Phe	Asp	Glu	Asn			
1730						1735		1740						

<210>67

 $\langle 211 \rangle 1054$ $\langle 212 \rangle$ PRT

〈213〉 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

 $\langle 400 \rangle_{67}$

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

Asn Ile Gly Asp Glu Gly His Pro Val Glu Phe Val Leu Glu Lys Asp			
85	90	95	
Gly Lys Asp Tyr Val Ile Ile Glu Leu Lys Gly Thr Thr Tyr Lys Asp			
100	105	110	
Leu Thr Lys Arg Arg Pro Gly Gln Gln Ser Pro Val Glu Gln Ala Thr			
115	120	125	
Asn Tyr Ala Ser Ala Lys Lys Glu Thr Glu Trp Ala Thr Val Ser Asn			
130	135	140	
Tyr Asp Glu Phe Arg Phe Phe Asn Pro Thr Ala Arg Asp Asn Tyr Ile			
145	150	155	160
Ser Phe Lys Phe Arg Gln Leu Lys Asp Leu Glu Ile Phe Lys Lys Phe			
165	170	175	
Leu Leu Val Phe Ser Lys Phe Ser Leu Ile Asp Glu Asp Ile Pro Lys			
180	185	190	
Lys Leu Leu Asn Glu Thr Lys Val Ile Glu Arg Glu Leu Glu Asn Glu			
195	200	205	
Phe Tyr Gln Leu Tyr Ser Asp Thr Arg Leu Met Ile Ile Lys Glu Leu			
210	215	220	
Glu Tyr Ser Ser Glu Asp Ile Asn Arg Ile Glu Ala Ile Lys Leu Ser			
225	230	235	240
Gln Ile Ile Leu Asn Arg Phe Ile Phe Leu Cys Phe Ala Glu Asp Leu			
245	250	255	
Ala Leu Met Glu Glu Glu Thr Thr Ala Asp Val Leu Leu Thr Pro Leu			
260	265	270	
Lys His Arg Asn Leu Ile Gly Asn Thr Met Trp Asn Arg Leu Asn Glu			
275	280	285	
Leu Phe Ile Phe Ala Asn Gln Gly Asn Lys His Arg Arg Ile Pro Ala			
290	295	300	
Phe Asn Gly Gly Leu Phe Glu Asp Asp Leu Ser Asn Leu Lys Ile Arg			
305	310	315	320
Asp Glu Ile Glu Asp Arg Ser Phe Phe Glu Asn Trp Asn Leu Lys Glu			
325	330	335	
Asp Phe Glu Asp Lys Tyr Glu Asp Ile Ala Lys Leu Ile Gly Val Tyr			
340	345	350	
Lys Asp Thr Leu Asn Pro Ile Phe Ile Asn Leu Leu Ile Ile Ser Thr			
355	360	365	
Tyr Asp Phe Asp Ser Glu Leu Asp Val Asn Ile Leu Gly His Ile Phe			
370	375	380	
Glu Asn Ser Ile Ser Asp Ile Glu Glu Leu Lys Asn Asp Asn Gln Glu			

385	390	395	400
Gln Arg Lys Lys Asp Gly Val Tyr Tyr Thr Pro Glu Tyr Ile Thr Asp			
	405	410	415
Tyr Ile Cys Arg Asn Thr Ile Ile Pro Tyr Leu Ser Ile Ser Gly Lys			
	420	425	430
Ala Ser Thr Val His Glu Leu Leu Tyr Glu Tyr Glu Ser Ser Asn Ser			
	435	440	445
Leu Asp Val Leu Asp Ser Lys Leu Thr Asn Ile Lys Val Leu Asp Pro			
	450	455	460
Ala Cys Gly Ser Gly Ser Met Leu Asn Lys Ser Val Asp Ile Leu Phe			
465	470	475	480
Glu Ile His Glu Ala Leu His Ala Ser Lys Tyr Ala Gly Asp Ser Ser			
	485	490	495
Leu Asp Arg Phe Phe Asp Ser Leu Glu Lys Arg Lys Glu Ile Ile Ser			
	500	505	510
Asn Asn Ile Tyr Gly Val Asp Leu Asn Glu Glu Ser Val Glu Ile Thr			
	515	520	525
Lys Leu Ser Leu Phe Leu Lys Leu Ala Thr Thr Val Gly Leu Lys Glu			
	530	535	540
Gly Phe Gln Leu Pro Ser Leu Asp Lys His Ile Lys Cys Gly Asp Ser			
545	550	555	560
Leu Val Asp Asp Glu Ser Ile Ala Gly Asn Lys Ala Phe Asn Trp Tyr			
	565	570	575
Glu Ser Phe Ser Glu Val Phe Glu Ser Gly Gly Phe Asp Ile Ile Val			
	580	585	590
Gly Asn Pro Pro Tyr Val Asp Ile Lys Glu Met Asp Glu Lys Thr Ala			
	595	600	605
Lys Tyr Ile Phe Asp Asn Tyr Glu Thr Ser Phe Asn Arg Ile Asn Leu			
	610	615	620
Tyr Ser Thr Phe Val Glu Lys Ser Tyr Tyr Leu Leu Lys Asn Glu Gly			
625	630	635	640
Ile Phe Ser Phe Ile Met Pro Asn Ser Ile Leu Phe Asn Ser Thr Tyr			
	645	650	655
Ser Lys Ile Arg Glu Leu Ile Leu Asn Asn Thr Ser Ile Leu Asn Ile			
	660	665	670
Val Arg Thr Ser Asp Asp Val Phe Lys Asp Ala Lys Val Glu Pro Ile			
	675	680	685
Ile Leu Ile Phe Lys Lys Gly Tyr Asp Glu Gly Asn Lys Thr Lys Ile			
	690	695	700

Leu Ile Lys Lys Asp Asp Met Asp Glu Ile Pro Ile Asn Asn Tyr Ser			
705	710	715	720
Glu His Phe Phe Thr Gln Glu Arg Trp Phe Glu Asn Asn Ser Ile Ile			
	725	730	735
Asn Ile Phe Ser Asp Asp Phe Thr Phe Asp Leu Leu Lys Lys Ile Asp			
	740	745	750
Gly Asn Asn Glu Arg Leu Ile Asp Tyr Cys Asp Phe Ser Leu Gly Leu			
	755	760	765
Thr Pro Tyr Asp Lys Tyr Lys Gly Met Ser Glu Asp Ile Ile Lys Asn			
	770	775	780
Arg Lys Phe His Ser Lys Ile Lys Leu Asp Asp Thr Phe Lys Glu Leu			
785	790	795	800
Leu Asp Gly Ser Asp Ile Thr Arg Tyr Asn Val Lys Trp Gly Glu Lys			
	805	810	815
Glu Tyr Ile Lys Tyr Gly Asp Trp Leu Gly Ala Pro Arg Glu Glu Lys			
	820	825	830
Phe Phe Lys Asn Pro Arg Ile Leu Ile Arg Gln Ile Leu Ser Ile Ala			
	835	840	845
Pro Lys Glu Ser Arg Lys Arg Ile Phe Ala Ala Tyr Thr Glu Glu Glu			
	850	855	860
Leu Tyr Asn Ala Gln Ile Ala Phe Asn Leu Val Leu Lys Glu Gly Phe			
865	870	875	880
Asp Asp Lys Asn Leu Leu Lys Tyr Phe Leu Gly Ile Ile Asn Ser Lys			
	885	890	895
Met Met Thr Trp Tyr Tyr Glu Glu Arg Phe Met Asp Lys Asn Lys Lys			
	900	905	910
Asn Phe Ala Lys Ile Leu Ile Glu Asn Ala Lys Asn Leu Pro Val Ile			
	915	920	925
Ile Asn Ser Asn Phe Leu Asp Glu Ile Val Ser Asn Val Asp Ser Ile			
	930	935	940
Ile Glu Leu Asn Lys Glu Phe Tyr Ser Val Arg Asn Ala Phe Gln Thr			
945	950	955	960
Trp Leu Lys Ile Glu Phe Glu Ile Glu Lys Leu Ser Lys Lys Leu Glu			
	965	970	975
Asn Tyr Tyr Asp Leu Asn Phe Glu Glu Phe Leu Lys Glu Ile Lys Lys			
	980	985	990
Lys Lys Val Val Ile Arg Pro Asn Gln Ile Gln Asp Leu Ser Glu Leu			
	995	1000	1005
Phe Asn Glu Ser Leu Gly Lys Ile Glu Tyr Leu Gln Arg Glu Ile			

1010	1015	1020
Lys Glu Ala Asp Glu Lys	Ile Asn Leu Leu Val	Tyr Glu Leu Tyr
1025	1030	1035
Gly Leu Asn His Glu Glu	Ile Glu Ile Ile Glu	Asn Ser Phe Asn
1040	1045	1050
Asp		
<210>68		
<211>86		
<212>PRT		
<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)		
<400>68		
Met Asp Asp Glu Thr Leu Ile Leu Ile Glu Tyr Ile Arg Asn Ala Pro		
1	5	10
Thr Arg Glu Met Val Leu Lys Ser Phe Glu Gly Val Asp Phe Ile Arg		
20	25	30
Pro Ile Gln Ile Ser Arg Lys Thr Gly Ile His Pro Asn Asn Val Ser		
35	40	45
Lys Lys Leu Lys Asp Leu Arg Glu His Glu Leu Val Tyr Val Ile Asn		
50	55	60
Pro Glu Tyr His Val Pro Lys Leu Tyr Arg Leu Thr Glu Lys Gly Lys		
65	70	75
Asn Met Leu Gln Phe Leu		80
85		
<210>69		
<211>255		
<212>PRT		
<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)		
<400>69		
Met Asn Lys Lys Ile Ile Leu Ser Leu Leu Leu Val Leu Leu Val Ala		
1	5	10
Ile Ser Val Ser Ala Val Ala Ala Ala Asp Ala Asp Val Thr Tyr Ile		
20	25	30
Asn Asp Ala Ala Asp Val Asp Asp Val Ala Asp Glu Lys Val Ala Pro		
35	40	45
Leu Thr Ala Ser Ala Asp Ala Gln Asp Ile Gln Thr Lys Leu Asp Asn		
50	55	60
Ala Lys Pro Gly Asp Thr Ile Glu Leu Glu Asn Lys Thr Tyr Asp Val		
65	70	75
Asp Thr Thr Phe Asn Val Thr Lys Gln Val Thr Ile Lys Gly Gln Asp		80

	85	90	95
Thr Val Ile Lys Ala Ser Gly Ala Ser Gln Gly Gly Ser Gly Ala Leu			
	100	105	110
Phe Ile Ala Asn Glu Ala Gly Thr Ala Phe Glu Gly Ile Thr Phe Ile			
	115	120	125
Asn Thr Asp Gly His Lys Asn Tyr Gly Glu Gln Val Ser Gly Tyr Ala			
	130	135	140
Ile Gln Leu Ala Ile Glu Asn Gly Thr Val Asp Asn Cys Lys Phe Ile			
145	150	155	160
Asp Trp Ser Ser Gly Val Tyr Gly Lys Gly Ala Ser Phe Cys Ser Ile			
	165	170	175
Thr Asn Ser Tyr Phe Asn Gly Ser Ser Glu Gln Val Thr Asn Gly Gly			
	180	185	190
Lys Lys Glu Tyr Gly Thr Lys Ala Ile Asn Leu Met Gly Ser His Asp			
	195	200	205
Ile Thr Val Thr Gly Cys Thr Phe Glu Gly Gln Val Leu Asp Ala Ile			
210	215	220	
Ser Ile Ala Ser Asn Ser Gly Asn Asn Ile Met Thr Asp Asn Thr Phe			
225	230	235	240
Ile Asp Asn Cys Tyr Ala Ile Tyr Phe Gly Gly Ala Ser Thr Gln			
	245	250	255

<210>70

<211>208

<212>PRT

<213> 马尔堡甲烷短杆菌 (Methanobrevibacter marburgensis)

<400>70

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

	100		105		110
Gly Ile Asn Cys Thr Asp Ala Cys Gln Leu Phe Lys Pro Val Ile Glu					
	115		120		125
Gly Leu Gly Tyr Ser Val Arg Ile Glu His Val Lys Val Arg Cys Asn					
	130		135		140
Asp Asn Lys Trp Tyr Gly His Tyr Phe Leu Arg Val Ala Gly Lys Glu					
145		150		155	160
Leu Ala Ser Val Ser Leu Pro Ser Glu Arg Trp Thr Val Trp Asp Tyr					
	165		170		175
Val Ser Ala Thr Lys Thr Gly Arg Pro Leu Gly Ala Pro Cys Cys Ser					
	180		185		190
Arg Gly Ile Gln His Leu Gly Trp Gly Ile Val Ser Pro Lys His Asp					
	195		200		205
<210>71					
<211>202					
<212>PRT					
<213> 沃氏甲烷短杆菌 (Methanobrevibacter wolfeii)					
<400>71					
Arg Arg Tyr Glu Asp Phe Val Arg Ile Asn Gly Arg Glu Pro Asn Ile					
1	5		10		15
Ile Tyr Leu Glu Gln Gly Lys Ser Asp His Val Ser Leu Gly Thr Phe					
	20		25		30
Lys Asp Met Leu Arg Arg Tyr Lys Asp Phe Val Arg Ile Asn Gly Arg					
	35		40		45
Glu Pro Asn Tyr Ile Ser Ile Gln Pro Gln Pro Ser Leu Lys Gly His					
	50		55		60
Trp Thr Thr Lys Val Ile Glu Lys Ile Gly Thr Phe His Asp Ala Thr					
65		70		75	80
Ser Leu Tyr Glu Arg Val Lys Lys Thr Cys Lys Tyr Lys Tyr Tyr Tyr					
	85		90		95
Asn Asp Gln Val Pro Asn His Val Ala Val Met Arg Met Thr Thr Ser					
	100		105		110
Gly Ile Asn Cys Thr Asp Ala Cys Gln Leu Phe Ser Lys Val Leu Glu					
	115		120		125
Glu Met Gly Tyr Glu Val Lys Ile Glu His Val Arg Val Lys Cys Asn					
	130		135		140
Asp Gly Lys Trp Tyr Gly His Tyr Leu Leu Arg Val Gly Gly Phe Glu					
145		150		155	160
Leu Lys Asp Gly Thr Ile Trp Asp Tyr Val Ser Ala Thr Lys Thr Gly					

	165	170	175
Arg Pro Leu Gly Val Pro Cys Cys Thr Ala Gly Phe Gln His Leu Gly			
	180	185	190
Trp Gly Ile Val Gly Pro Val Tyr Asp Lys			
	195	200	

<210>72

<211>155

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>72

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

<210>73

<211>212

<212>PRT

<213> 人工序列

<220>

<223> 人工序列描述 :合成的共有序列

<220>

<221>MOD_RES

<222>(1).. (1)

<223> 任何氨基酸
<220>
<221>MOD_RES
<222>(4)..(5)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222>(7)..(8)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222>(15)..(15)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222>(20)..(20)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222>(23)..(28)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222>(32)..(32)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222>(43)..(43)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222>(46)..(47)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222>(49)..(49)
<223> 任何氨基酸
<220>
<221>MOD_RES

<222> (53).. (54)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222> (56).. (56)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222> (58).. (58)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222> (60).. (69)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222> (72).. (72)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222> (74).. (74)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222> (80).. (80)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222> (87).. (87)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222> (92).. (92)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222> (103).. (103)
<223> 任何氨基酸
<220>

<221>MOD_RES
<222>(106)..(107)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222>(115)..(115)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222>(127)..(128)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222>(132)..(132)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222>(149)..(149)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222>(156)..(156)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222>(162)..(162)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222>(165)..(166)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222>(168)..(174)
<223> 任何氨基酸
<220>
<221>MOD_RES
<222>(191)..(191)
<223> 任何氨基酸

<220>

<221>MOD_RES

<222>(196).. (196)

<223> 任何氨基酸

<220>

<221>MOD_RES

<222>(198).. (198)

<223> 任何氨基酸

<220>

<221>MOD_RES

<222>(207).. (207)

<223> 任何氨基酸

<220>

<221>MOD_RES

<222>(209).. (209)

<223> 任何氨基酸

<220>

<221>MOD_RES

<222>(212).. (212)

<223> 任何氨基酸

<400>73

Xaa Arg Tyr Xaa Xaa Phe Xaa Xaa Ile Asn Gly Arg Glu Pro Xaa Val

1 5 10 15

Ile Phe Ile Xaa Asn Gly Xaa Xaa Xaa Xaa Xaa Xaa Val Ser Leu Xaa

20 25 30

Thr Phe Lys Asp Met Leu Lys Arg Tyr Lys Xaa Phe Leu Xaa Xaa Asn

35 40 45

Xaa Arg Glu Pro Xaa Xaa Ile Xaa Ile Xaa Pro Xaa Xaa Xaa Xaa Xaa

50 55 60

Xaa Xaa Xaa Xaa Xaa Glu Met Xaa Lys Xaa Leu Gly Thr Phe Lys Xaa

65 70 75 80

Ser Ala Thr Ser Leu Tyr Xaa Leu Val Ala Lys Xaa Cys Lys Tyr Lys

85 90 95

Phe Tyr Tyr Asn Asp Gln Xaa Pro Asn Xaa Xaa Ala Val Lys Lys Met

100 105 110

Leu Thr Xaa Gly Ile Asn Cys Thr Asp Ala Cys Gln Leu Phe Xaa Xaa

115 120 125

Val Ile Glu Xaa Leu Gly Tyr Ser Val Lys Ile Glu His Val Lys Val

130 135 140

Lys Cys Asn Asp Xaa Lys Trp Tyr Gly His Tyr Xaa Leu Arg Val Ala			
145	150	155	160
Gly Xaa Glu Leu Xaa Xaa Val Xaa Xaa Xaa Xaa Xaa Xaa Thr Val			
	165	170	175
Trp Asp Tyr Val Ser Ala Thr Lys Thr Gly Arg Pro Leu Gly Xaa Pro			
	180	185	190
Cys Cys Ser Xaa Gly Xaa Gln His Leu Gly Trp Gly Ile Val Xaa Pro			
	195	200	205
Xaa His Asp Xaa			
210			
<210>74			
<211>264			
<212>DNA			
<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)			
<400>74			
gtgagaaata tgaagaataa gagtttaata ttaatttctt tattattact gattacaata	60		
ataagcatag gatctgttgt tgcaacggat aatgaagaaa ttaatatgga taatataaat	120		
aatattgata ataatgagga tategcta attgataatg tcgataatgt cgataactcc	180		
aatataaaca atccaactga cataagaata gacaattcaa acctaaatag agaaacagaa	240		
ctagattcaa atttaaataa atct	264		
<210>75			
<211>255			
<212>DNA			
<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)			
<400>75			
atgaagactt ttagacaaca gctattggaa gaccctgaat tccagaacta cctgttgcag	60		
agaccgaacc tgacggagag cagcctgcag tcatacctca atgccgccac caacttcgtg	120		
aggttcacag gggagccgtt ctacaagacc gtgcatgagc tcagaagcca gcagaacgat	180		
aggattgaga acaacatcat cataaggttc aaccggaacc agtcaaggat aaacatcatg	240		
cagtttgagt tcata	255		
<210>76			
<211>255			
<212>DNA			
<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)			
<400>76			
atgaatacat taaaaataga atgttcaaaa gatgattata tcataaagac tgcaaaggca	60		
aattctgaaa atgagttgaa agttgaagtt cctgagaact ggaactgcga ctatgtgaat	120		
gctgtcctgt gggaagagga catctgcgaa gtgcttgaaa agggatgatga aagaatgctt	180		
ttgattccta tgtgcggtga actgcttctt gaaggagtgc aggaagatga atacataaag	240		

tacatttgcc ttcct	255
<210>77	
<211>255	
<212>DNA	
<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)	
<400>77	
atgggaatga aaaaagttaa gatcgaaagt aaagcaaaaa ataatatgac aagaaacgaa	60
aaactcttct ataaattata tgacgattta tataatgaag atagattaat tttattctcc	120
acatatTTTA acatatatga taaatttatt tttaaaaaag atataatcca ttatgtttta	180
atgaactact ctgaaaatga aattattgaa gcaatgaaaa aaattgatga aattcagtca	240
gaaggcattg atatt	255
<210>78	
<211>255	
<212>DNA	
<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)	
<400>78	
atgaaaacac aagatctaata taatataata aatgatgagg aatctcctgt attcctcaac	60
agggaagtct ttgagatgga ctatgtgccg gacatctaca aatacagga cgagcagctg	120
gcgaaaatgg cgatgtactg caattcaata cctgacaaca tagctcccaa gaacctgcaa	180
ttgtgcggag gcaatgcgac aggaaagacc acaacattaa agcagttctt caagatgttg	240
aacgaggctt ttcca	255
<210>79	
<211>255	
<212>DNA	
<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)	
<400>79	
atgttttttg aaaaatgcga tggaaaagat acaatcagca tgtctcttca agaaaagatg	60
aatcttatat tggaaactat ggaaagcaaa ggaagcccct tcatatcctg tattcattgt	120
aacataccta taaatgaagc agaagaatgg tacaaaaatg gagaaatagg agatcaggac	180
tttataaatt tctatgatga tgtaaacttg attgaagaaa gttttgggtt tgaaatttat	240
aaaaaatcag aatat	255
<210>80	
<211>207	
<212>DNA	
<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)	
<400>80	
ttgataatTT tccttaaact aaaatttggc cgtgtaatca tggaaaaatt aatcgaaatt	60
gacggagttc aatacacaga agcccaaate agaagggtt tagccattga acgagacgtg	120
cgctctccta atttcgttga tatgttattg ggaaagatca aacctagcga acttgcaagc	180

agagtttctg aaaaaggtga tgcctaa	207
<210>81	
<211>156	
<212>DNA	
<213> 反刍甲烷短杆菌 (<i>Methanobrevibacter ruminantium</i>)	
<400>81	
atgtgctatg tgggaaacac caggacattg gtctatcaca cagaggactg cttctgcaac	60
cactggctgt tgaacgagaa caagaccatt ctagaggaaa agcctgtaga catgaaacct	120
tgcagcttct gcaaaccaca gtttgacact gaatag	156
<210>82	
<211>234	
<212>DNA	
<213> 反刍甲烷短杆菌 (<i>Methanobrevibacter ruminantium</i>)	
<400>82	
gtgttgata tggttgcaga gatggttgaa aacatcagga aaggagaagg agatggatat	60
tccatctacc ctcccttttc ctgcattggtt ttcctacaag gcaaaaagta ttccgaatgc	120
tgctgcaagg ctgaggcaag ggatcagaag tttgcacttg tgaaccttgt aggcttcaga	180
agagaggatg tcaaggtcat agaccaagg accaacgagg agctgttcgt atga	234
<210>83	
<211>162	
<212>DNA	
<213> 反刍甲烷短杆菌 (<i>Methanobrevibacter ruminantium</i>)	
<400>83	
atgagcatgc ttgcagactt cgagcctgca aggctccaca agaggacatg ggctgaaagg	60
catgatgtcg aaatccttgc tgtcatatgc cttgccataa gcattgcaat gctacttctt	120
ttctttgcgc ttgcagagcc gactgttgca ggagtgattt ag	162
<210>84	
<211>255	
<212>DNA	
<213> 反刍甲烷短杆菌 (<i>Methanobrevibacter ruminantium</i>)	
<400>84	
atgactaagg aatttgagga ctatcatgaga aggaacacag gactgcttgt tttcttgaga	60
tgggacacag tggcatatct gaaatacctt gagtcccatt atgacgaaag gaagtacgaa	120
tgcgcttaca gattgtttgga ggcaatcgac aatctctttg acttctacca gataaccttc	180
agcaaaaaga ctgaaaggga aattcctgac cagctttttg aaaaggacaa gataaacaaa	240
ggcttccctt cacac	255
<210>85	
<211>255	
<212>DNA	

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>85

atgaatgaga tagttactac aactaacgag aataatgtac ctgtcgacgt tgactatgcc	60
atcgaggaat ggaaggcata ccagcgattg accagagagc tattggacga aacagattat	120
cagacacaca gaggcaggaa gtacaagacc aagagcgctt ggcagaagta cgcacgtgcc	180
ttcaacataa acacacagat aatcgacaag gaaatcgtca agaatgacaa gggaatcgtc	240
attgaggctg agtat	255

<210>86

<211>231

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>86

atgattgagt gcatacagga agaggagac ttcacggatt gggaggttcc ttcctcctct	60
tccgacatca agtacatagt ttcagtggat gatgaggga accgtgtctg cagctgtcct	120
gacttctatt acaggaagtc caggatgaac cctcacatct caaacctga atcctattgc	180
aaacacatca gacaagtctt ggaggaggac aatagattgc aaatgctcta g	231

<210>87

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>87

atggataata taaataagac aaaaacaagt ttagctaagt ttgaagagtt tttcagcact	60
gtatacaagg acgaagtcac ggaagttttg gaaaagtatc ctgaggaaag gacattgggt	120
gtggactatg agaatttgg aatgtttgat cctgatttgg cagacttggt gattgaaaag	180
cctgacgaag taatcgacg ttcacagaaa gcaatcaaga acattgaccc attgatgaaa	240
gatccaaagt tggac	255

<210>88

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>88

atggctaaca agattagagt taatctcacg gttgacccta atttgtggca attagcaaaa	60
gacaagttac cttgcagcag aagcgaattc tttgagaatc agttgaagat gttcttaggg	120
attgaagatg acgagtcgga aatcattaag gacatccaaa ccaaagaaaa tgagattaac	180
gcactaagag acaagttatg tcatgttcgc aaatcaaagc agttgaaatt ggaatcaaac	240
aaatctatgg agaag	255

<210>89

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>89

atgacaattg gaagtttggg caatttttggg aagtcagatg gagaaaacat gaaccctgaa	60
gatttcgact gttctgtttt ctttgaaatg tacaaggcac tttttgagat attggatgtt	120
gaagttggca gcttcgctga actcttggat gtgtacaaga atgtcgagat ggattacacc	180
ttgaagagac atgccctcaa gcagaaggaa atcctttact ggttcaatac cgactggaag	240
gaggaacttg gaaag	255

<210>90

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>90

atgaatgtga aaacagtcac gaatgacctt ataggattgt caaaggagtt cgagggagtg	60
gaatatgaga ttgaatccaa aaattcaatc tacttctatt cctttccaaa gtacatgaag	120
gaaggcattg tgattctgaa atattccgca atctatgacc tgcacaccat attgaaaggc	180
atggatggaa tcatagtggg cattctggag gttgaggaca atcctggtga tgagaagaag	240
gatttgctat atgtg	255

<210>91

<211>234

<212>DNA

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>91

atgattagcg acgaatggga agaggaatac tatgtcaagg tcaacgagga actggagcag	60
gtggaagtca ggtttttcag caaggctgac aggctcgtct ttgctcaatc ctactcttct	120
tcatcttctt tttcctttga ggaggctgaa atcctgtgag acaggatatg cgacatactg	180
acaacagatt tgggaaatgc caagtattat ctaggaggaa aagatgaatt atag	234

<210>92

<211>150

<212>DNA

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>92

atgaattata gggaatgggt tgcaagccag ttcgagctga ggatgactga ggacggaagg	60
gtttgcttca agacaccatg cagctactac acattttcaa aggaggactt tgagataata	120
agggagatgt tccttaattt cgagtctga	150

<210>93

<211>174

<212>DNA

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>93

atggatagaa taaaggaatt gaatgtatgc ggaacctgca agcacagcca tttgattcca	60
gacatcaacg gagagatagc agtgaacatt tgcaggatag gttcagaggc tgtcaacaag	120
gacggaggac ttacatattg cgtagactgg acaccaagga ggaagctgtc atga	174
<210>94	
<211>336	
<212>DNA	
<213> 反刍甲烷短杆菌 (<i>Methanobrevibacter ruminantium</i>)	
<400>94	
atgttaagta aaaaagaagc catacagatg acgctggaca atgagaagca ctatcctgtt	60
aagtgcgaagt actgcggaag gccgttcacc aagtctcaca acaggcagat gtactgttca	120
gacagttgca gacggaatgc cttgagggaa cagaaggcaa gataccaggc taaaaggagg	180
ctaaaaataa agcagaaagt gctgattgta gatgaatata agaaatacgg tttgggtagc	240
tatggaacaa gtgcgaacgg acacagaaag aacaattttt cacatgagta catggctatt	300
caaaaggaga tgaaaagaat aggattgaag agataa	336
<210>95	
<211>261	
<212>DNA	
<213> 反刍甲烷短杆菌 (<i>Methanobrevibacter ruminantium</i>)	
<400>95	
atgtattttg ctaaattttg tcctaactgt ggaaataaag tagaagaaaa tgataaatc	60
tgcattttatt gtggaaataa actaagagtt ataattcctg aaaaaaagt gaaaagaagc	120
tcaaatagta ttaatgatga aaaaactagc aaatatgttg aagtcattga tgggctaattg	180
agatataaag tttttccatc aagtttacct gtgaaatata ttatttataa agtgaattat	240
ggaacaacat ctgatgaaat a	261
<210>96	
<211>234	
<212>DNA	
<213> 反刍甲烷短杆菌 (<i>Methanobrevibacter ruminantium</i>)	
<400>96	
atgatttatg ataaagcgac agttacaagc ctcatagtgg ctatcctatt gccattgatg	60
agcatgttag gaatcggaga attgactcag aactacatat tggcaatagt gagcggtagt	120
atagcccttg tggtttggtg ctacaacgag aagcacaaca gcgatttggt gagcggaacc	180
accaagtgcg actgcgagct ttgctatggg ggagacgatg aagcacttat atga	234
<210>97	
<211>255	
<212>DNA	
<213> 反刍甲烷短杆菌 (<i>Methanobrevibacter ruminantium</i>)	
<400>97	
atgaagcact tatatgagat aatcccttac aggaggactg tttggattac agggttcctg	60

aagacaacag tctcctcggc aatgataacc actggcgtag tgatattggt caacagcata	120
accgagcacc cttacttcat ggagtgggac gagataggaa tcgttcttgg aatcggttcc	180
atcaccateg cctgcattta catagegatg atagacagat ggaaggaacg caggaagaag	240
gaggagcttg acacg	255

<210>98

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>98

atgatagaaa tcagcaccat aaagatcacc gacataaagc ctgccgaata caatccgagg	60
ataatgagcc agcttgaaca cacaagctc aggaactcca tggagacatt cggagtgggt	120
gaccaataa taatcaacct gaagaacaac cacatcatag gagggcacca aaggtacgag	180
gtccttctgg acaagtccat ggaggacaac gagttcataa aggagctcca cctcatcagg	240
cttggagatg taggc	255

<210>99

<211>330

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>99

atggacgtcg acaacctgct gggatatgac aaggtactgg aattgaggtc tttattggag	60
gaggtcacgg acaaggtgat acccgatgg cacaagaacc gcggaatcaa ggacttcaag	120
cagatgtgcc aggactacaa tttcgtcagc ataagcggct ggagaaacga ggacgtcaag	180
gatgaccagt tcatccattt cgtcaggcat gcacacagga acggctgcag gattcacgga	240
ttgggattga cacggagaaa gggtctcgac agggttccct tcgacagcgt cgacagcagc	300
agctggctcc agacatatt gtatgcaagg	330

<210>100

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>100

atgattaatg aaaggataag acctattttg aatttttagct tcaatgacaa gaaagtcatt	60
ttcgtcacat tgttcgttgt gagcaatctg ataagcaacc tgtagccat caaggttttc	120
aacttgggat tttggggatt gacaaccgat tgcggaaatc tcctgttccc gtaggatac	180
cttatggcag acgtgattac agaagtctat ggagagagga cagctcgaag ggtcatattg	240
cttgggctct ttgca	255

<210>101

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>101

atgcccgaac cttgggaaag gcaaagggac gagaacggaa agcttgagcc aataaaggca	60
ttcgagtact tcaccgagta cctgacaatg gacaagccac gaagcatgag ggttctgtgc	120
gagaggctcg gcaagaaaga tgggtatatt aggcaacttc atgcctactc atccacatgg	180
aactgggtcg aaagggcaga ggcatacgat gaacacataa tattgaagaa acggttaagg	240
aaagagaagt ttac	255

<210>102

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>102

gtgaaggaca ttgtcaatca ttacggctac ctttcaccgt acaagccgac catcaggtct	60
gatagtaagg ctaaaaacaa gttcaagctc aatgagcctt atcgtggaca gatgttgagt	120
gcaggtgctg gcggttcaat catgggatac ggtgcaggtc tccttattgt cgatgaccct	180
atcaagaacg tcgctgaagc cgaatccaag gtacgtcagg caaagctcaa ggactgggtg	240
ggaggaacca tcaag	255

<210>103

<211>273

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>103

atgagtaaga aacaagaaat gatgagaatt gaaaggttaa aacactatgc atatcaaaca	60
ggactaataa tccctatitt taaaaatcat gatttaatta aaaaaataga aaatggaaaa	120
ataactaata ctgacgaaat aaagatatat attgaggaaa atgaaaagca gataaaaaga	180
agaagagaat ttataagtat tatttatgat aattgcaaat attttaaata cgacagcata	240
tggtataaat taatatccaa agtaaataat ttt	273

<210>104

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>104

ttgagtaaga atgctaaagc agacgccttt gttgtgacca ctgaagacgg atcctatgac	60
attgtcgatg cagatgtctt ggagaggtat gcaataaaga gcaaatctga cgagacagga	120
agcaagcagt tgaagactga tggttgggaa tatgatgata cattgcttga gccgttgtat	180
gaccatttgc agtttgttga gctattggag ataaacacat atcatgagaa ctgtgtcgat	240
gttggttgc aa gggac	255

<210>105

<211>252

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>105

atgatggcgc tcaaggtgcg tcattcaaag aggcagattc agaatatcaa gagagaatac	60
agaaggcggtt tgatactcga agagcagtcg agtagggaga ttgcggactt cttcagaagg	120
cttgagagga aaatccacaa ggtcatggat gagcactggg aaagcgaatt gggctctttc	180
catctgaaca aggtttcaga catcatacag gacagcaggc aggagtacta tgacatcctg	240
ttcaagtact gc	252

<210>106

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>106

atgataattg aaattcctgg aaatgacagg acagaagaaa taaatttgcc gaacggacag	60
tttgtactga taacctactt gcaggaaaat gacatgatgt cattgcctga cgggaagtat	120
gtctgcccatt tcagaatgat tcagcttttc actgaagagg gcggtgaact gattggagag	180
tgcatggagg aaaatcctca ctacaacact cgcttctaca atactgtctt tgagcatttg	240
gatgaaggaa ttgag	255

<210>107

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>107

atgttaagag tcgtagagag aacctattat caacaagaag aaatcaagac actagactgt	60
cgaattagag aggcaggagt caacacttac tcccttgctc gtcaaggagc agttgactat	120
cctacctacg aatcctacaa cgaggtgaga gaggaaagga taaaagaggc taaggaaaaa	180
tacggtgaat cctactacta tctctggaga gacgttgaga cctacttcta ttacttagga	240
agattcttta gtgat	255

<210>108

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>108

atggacgcga taaacgttat caacaaaac aagatattag tagatgtttt gtaccgtggg	60
actgtaaacc tgattgacat cgtcatcggt gatgcattgg tctatgacaa tccgacagtt	120
gtagtcaagt gctacaccac agatttggct ttgcaacca cagagattga tgagattgta	180
ttagagaacg aagaatttga gttattaata acttacgagg acggagagtt ctcaattttg	240
ttgaagtecc acaat	255

<210>109

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>109

atggtttttg tcgatttgtg cgagaacgaa atagcagaca tggtcgagag cttctatcgc	60
aatggagacg gtagcagcag gatattgacc aacagagcag ttggagagat agtggagcac	120
tactgttctt ttgaagtaga cggtagaate accactaatc tccgagattt cctattatat	180
agtgttgtct tatatgacac cattggagag gctgtcgatg acggagtcaa ccttgaggaa	240
gtgttcatca tagaa	255

<210>110

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>110

gtggtgaatt ggttgaagc tattggagat aacttctcag tggattatct ccttttagca	60
ttattctcta gtggagattt gatattgggt gccatcggtc tcaattctta tgggtgcatc	120
tcacctgaga atgtaaggga actggtgatt gactatataa gttacaggaa agttgatatt	180
ttttggagac atctaaggag acctagaatg agttttgaag attatgtctt agacaatttt	240
gaagagatgg aaact	255

<210>111

<211>192

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>111

atgacaatga aagttacttt tgaagatgag aacggtgaga gaacagtcga gttcggcgat	60
gacgtcgatt tcgttttaaat cgaaagcgat gatgacggaa acattgagat tcgcgaagga	120
gactgggaat tggacggaga ttccgacgat gattgggaag agtacgatga ctgggacgag	180
gaggagtctt ga	192

<210>112

<211>279

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>112

gtggacgaat tcgttgagac cctgttcgac acctactgga aggtgaatga gaacggagag	60
tacatgagcc ttaccgattg cggtagtttc tacatcgcaa aggttgctcc ctgtgttcgc	120
aactggagca tagtgataga atgcaattgc ttctgttttc attgcaagga atttgtctac	180
catgagaacg gagcaatact tgagataggt atggagatta gctctcttta cctaagccaa	240
atggagataa aagatttgaa gatatacatg atagattgt	279

<210>113

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>113

atgttttttg	aaggaatatt	tgaacaagac	ggagagaatg	taagagaaca	ggtgatttac	60
tggagaaagt	ccaatcaggt	acacaactgg	ttcgtagtca	acgctcagga	cggagaggac	120
aactgccaac	cacactcagt	aagcagagaa	cagtttagagg	aactaagaga	cctatgcaga	180
gcagtgctcg	cagacaatga	caaggcagag	gaacttctcc	caacaagacc	tggtttcttc	240
tttgggtgcga	tagac					255

<210>114

<211>174

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>114

ttgcagttgg	tggtagaagg	tgagaatatg	gaatgccctt	gcgataattg	tgagatgtta	60
agagagaacg	aaccaatcaa	ggtgattgac	tggaagtcca	attctccatt	tggaaatggg	120
atttktattc	acttttggag	ggaacggctt	caggatgttg	gctataatgg	gtga	174

<210>115

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>115

atggaactga	tgactaggga	aatggaaggg	aaactcaa	aat ccttcccttt	ctattcccaa	60
gacggaaaag	gagatgacgc	gatagtgggtg	atgaagttct	tcaaccctta	cggattggga	120
acttgggtatg	tattggaggc	agagaaacag	gagaatgggtg	attacctttt	cttcggatat	180
gttgaatcac	cgataacacc	tgaattcaac	gagtacgggt	acttctcatt	gtctgagttg	240
gagaacctta	agata					255

<210>116

<211>225

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>116

atgttcggac	agaagaaaga	atttgtcaaa	aagatgtatc	atgtcggaga	cgttgtcgaa	60
ctggttcata	tggacgatgc	acaggctcca	cctagtggaa	ccagaggaga	aatactcttc	120
gtagacgata	ttggtcaaat	ccatgtgaga	tgggagaacg	gctcaggtct	ggcactaata	180
tatggtgagg	acaggttcaa	agttgtagag	aggaaaggag	aatag		225

<210>117

<211>153

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>117	
atggatttgt tgtatcttta tgatgacctg actgcaagga gagaagtcta cgattcagtc	60
ggattgagtt ttgtgggtcaa gtataaatte tcatctcgta aggaagcaca ggatttcgca	120
ttgaagtacg gtgcagaact gattgaggaa tga	153
<210>118	
<211>189	
<212>DNA	
<213> 反刍甲烷短杆菌 (<i>Methanobrevibacter ruminantium</i>)	
<400>118	
atgattgaac ggaggagatt aggaatgaag tacgatattt ttactatctt ggatgaaatt	60
tcaaggaagt tagatgacgg tgaattatcc gatgagcagg ttgatttcct gttgcagatg	120
gagattctag tagaggaagg cactatcaca gacgagcagg cacaagatgt catgaatgga	180
gattattaa	189
<210>119	
<211>168	
<212>DNA	
<213> 反刍甲烷短杆菌 (<i>Methanobrevibacter ruminantium</i>)	
<400>119	
atgatgaaat taagtttgaa ggaactcggg gaggaaattg agatgatact tgcagaaggt	60
ggattaacct ttgatcaagt cgattatctg ttgtacttgg agacctgcat tgcagacgga	120
agcataacag aggagcagaa gagagaaatt atctgtaggg acttctaa	168
<210>120	
<211>234	
<212>DNA	
<213> 反刍甲烷短杆菌 (<i>Methanobrevibacter ruminantium</i>)	
<400>120	
ttggtgagag aacaggagag attgattatg tatttggaaa ttgacgaggt caagtgcgaa	60
aacattgaca gaattgagtt tgacgatgtc gcaatggaaa tcgtcctgac agacgaaaag	120
gtgtacgaga gaataaagag atggctcaag tccaatgaga ttgactacga ttgcagggaa	180
gaccggtatt tcgccaatct catagaatat gtcattagga taacttgggtg gtaa	234
<210>121	
<211>216	
<212>DNA	
<213> 反刍甲烷短杆菌 (<i>Methanobrevibacter ruminantium</i>)	
<400>121	
gtggtaaccg tgagagatgt cggattcacc attgaggaac ggttcttcct tactgctcag	60
gaattggagt attccgaagt gggagaggaa cacgagagcg tcattgacag agccatcgca	120
ttgctgtaca cgaagctcag aaccagagac ttcgagttca ccgacgaaga acgggaacte	180
ttggaggatg cttttgtcat tgtcagcgac caatag	216

<210>122

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>122

atggctaaag aaaatgttat agattataag attgagcgac aaaatgataa cacttggagt	60
tatttatatg taactgaaag agggagagga aatatcattg cttcatcttt tggagaattg	120
aggcagaaag ttttaaaaag aggtttgcct tggaatgata tatctaacct attcactaaa	180
aagagttctg attctaattgt tcgtcgcgaa tctgtaaaga atattgatga tgaatctgtt	240
ttggctgatg tcgct	255

<210>123

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>123

atggtagtca aatgtcctaa ttgttttagt ccacgtgttt caaaatgtga agataactaat	60
attaaatggc aatgtgataa atgcaaatgc aaatttaacc atgggtgcttt cgatattaat	120
gctgaaatgg agaaagtaga acaattaaca atagaaaaaa ttgaaagaga acgagaaaga	180
actgaacaat ttgaaagagc tattaaagaa gcaaaggaac aatttgaaag agaaagaact	240
gaacaatttg aaaga	255

<210>124

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>124

ttgaccgagg gtcagataag gcagattgca catgagtatc tggcaaatta cagccttgct	60
gacaagaacc atgagttctt cgagaccaga gaagttatag gagttcctgt ggagtcatat	120
ataacaaacg agcctataag cctcaagggc ttggacggaa cagtcaatga gtatccaaaa	180
ggaacctgga tagccaccac aaggataact gatgaagagg agatggaaaa ggcactcaat	240
ggagagtaca ctgga	255

<210>125

<211>330

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>125

atgggaaacg aagctacttt aaaccaattg gtaaaccgagc aggaaaaggc agtcttcaag	60
tccatgagga ctgacatgga gacaggaaag gctgtattga acgtcgagca gctagggttac	120
ttccttagag aagcaacatt agacaatata attttaagag atgcggactt caagctgatg	180
aatcattca agaagcatct caacagggtg ggaataaacg gaagggtcct cacaaacgga	240

tatgacgtga acggcgagac cgaccctgag attcctgcgg ctgacgtcga cttcggagca	300
aacgagtttg acgtcaagaa gctcaaggca	330
<210>126	
<211>255	
<212>DNA	
<213> 反刍甲烷短杆菌 (<i>Methanobrevibacter ruminantium</i>)	
<400>126	
atgcggaggc gttgcctgaa ctctccagag cataatggta tgatttcgca tatcattgtg	60
cttctcattt gctttatagg attggttgag gcgatactga tggcattggg tgattgggag	120
gacttggcaa tatecgttcg caagtctcct agaaagcttt ataatgtttt gaaggatgag	180
ttaggtcttc ctgaatggaa cgaactgtct gtgattgaaa ggagaagcat gaagaaaagg	240
tatgctgtca taaga	255
<210>127	
<211>255	
<212>DNA	
<213> 反刍甲烷短杆菌 (<i>Methanobrevibacter ruminantium</i>)	
<400>127	
atgacttggg ttggtacgga agatgtcacc gagttcacag gagtcaagcc tcagacattc	60
aggttcgaga agggagacac ctccagcctg gaaacattgc ttgagaagtg gattctgcag	120
gcagaaggac ttataatata ctactgcaac tacgatttca atgacttggg ggagatacct	180
ccagcagttg ttaatgtctg ccttaggctc actgcaaaca tggtcgcatt ggcacaggca	240
aggaaggaca ctcct	255
<210>128	
<211>258	
<212>DNA	
<213> 反刍甲烷短杆菌 (<i>Methanobrevibacter ruminantium</i>)	
<400>128	
ttggtgaaat tgcagattga cgttgaggaa ctcaagccat tggagccaag gttcaaaaag	60
gttgccaaaa ggacagttgt gttgactgca aatgaattgc agagaaacct caagaagttg	120
agtcctgttg atcatggaag gcttcagggc tcttgggtaa tcttccagac aggagaattg	180
gaaaggactg tgaaaagcag tgcaaagtat gccattttcg tgaatgacgg aacaggactg	240
tacggtcctt tgggtcat	258
<210>129	
<211>333	
<212>DNA	
<213> 反刍甲烷短杆菌 (<i>Methanobrevibacter ruminantium</i>)	
<400>129	
atgagattcg taaatactgc ttgccttggt cctcagactg tcaaggcata tcttgaaaagg	60
gaaatctgcg aaggagggtt gcttgaggat gtagagacac tcattccgtc cgtgaacagc	120

gacgttcctg ttgacccacc tgcgatatgg atagtcacag accccactac cagatgggtct 180
 ggcagtcagc caaatctctc aaacaagata gctatgtcag tccctttcga gttcgtatgc 240
 gtggaataca gcgatgactt ggaagaggct gagatattgg gaataagcct agccagcagg 300
 gttggctcaa gcctgatgaa gaacttcaac aag 333

<210>130

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>130

atgggaattc gtgtttagg aatgaaggaa gaggcaaggt atggagttgc ggagtcagcg 60
 ccggacttcc atcaggaagt tagcaaggca aaggcttcc tgaactccac tccgaacaca 120
 aagtcaagcg gctcaaggat gaagaagaag gcacgtgcag gcgtgtacaa gcctactgcc 180
 aacatcgaag gtgaagttga cttgaagagg ataggacatt atctcaaggc tttcctggac 240
 aattaccatt ttact 255

<210>131

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>131

atggttgtag ttaagaaatc cgatatatta aaggcgctaa aaaagattga aaaggtgaag 60
 attgaggctt tggacggaga cgagatgtac ttgagaccgt tgtcccaagc cgagatcaac 120
 gaggtcgacg agattgagcg aaaggctatg ggaatcttcg agaccaacga gaccgcacac 180
 aggggaagaa ggcagaagcc taagagtgtg gttgagagca aaggaaagat aaatctcgaa 240
 ttgcagcaga aggca 255

<210>132

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>132

ttgccgtcaa gcaatgtaat gaacataata gtcaaggcag aggacatggc atcatctgtc 60
 gcccaaaagg ttgaaaacag cttcaggaaa ttgggaaaca caatagacag caccttcacc 120
 acatctcttt caaacaccaa gttcaatcag gaactcactt cttttggaac agacttggac 180
 aaggtaaccc aaaggctgaa gcaggtaggc gtgaacggtc agtccagctt caaccagctt 240
 acaaatgctg aaagg 255

<210>133

<211>312

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>133

atggcaagtg ttacaaaata tccgagcaac gtttcacaga ccactggagg aaagttcgtg	60
tccttcagca atctggcaaa cataaagaac aatgctgacg gagcgcatgc cgtgagcagt	120
gttcttatca aaagcaagaa gcagtctcca aacagaccgt ccacagtctc atgcaaaggc	180
ttcggattca gccttcctga aggtgctgaa cctactaaaa tcacagtaac ttataggcat	240
aggaagaatg ctggaagega ctacagctca aagaacaaga ctcacatatg caacatcgga	300
ggacctacaa tc	312

<210>134

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>134

atgggcatag ctattgttgt tatggacaat gaggagaact tcctgcaatt ccttgaccct	60
gatttatgca ccatcaatga gaccatagag gaattgggct tgaggacttt ggagttcaac	120
tacaagtttc aggactatgt tgaggacagg gatcttttca ggataggtaa caagatctgg	180
attagcaatt ccagagcctt ggaggactgc ctttatgtga taaacactcc tgtggagaac	240
agcgtctatc aggag	255

<210>135

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>135

atggctgaga agattacagt gagtcctcag gaagtcagag gatacggaaa tgttgttgat	60
gagaaggaat tggaagatta cggaagctac aggtgtgatg tgagcgagag ctcggagggtg	120
atcaaggagag ttgaggaaag gatattcagc gtttctggtg ttctgtctcc tgctctgagc	180
attgccaatg ttactgagga cactcgcagg ggcagggtgcg cccacatctc cgccctcattt	240
gaggatgggg aaggc	255

<210>136

<211>381

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>136

atggttagat tcagcagaga catgctccag gacggagcga agagaatgtt caagtggcta	60
agaaagggcg aagggttgcc taactacttg ataattgtatg acatggacag gaataaggag	120
tataagttgg ttccaaagga atatgcagga ctgtatgagt ccagaaacat attctggatt	180
aagaacggaa gggagcctaa ctatgttaca ctgacttccg ttgcaaggaa tcctcttgtg	240
atggactacc agaacaccaa ttacacctgt tgcccaacca gtttgtccct tgcctcacia	300
atgctatate actataagtc agaaagtga tgcgctaagg ctttggaac tagcaagggc	360
agtggaaaca gccctgcaca g	381

<210>137

<211>2250

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>137

atgaaaaaac attgctttta ttttttaggg gacagttttg ccgatatatg caatgaagcc	60
atgtttttgtg aaaaacattt agtagaagga aactatttgg attcaattat ccgtgcagga	120
aaggcttcag agataataac tgtaaacatt tgtgaacttg aaggtaaga tggcttaata	180
agctctggtc agaagaaaag attggaaatg cttggatata agggatatcat ttcttatgat	240
atttataaaa gattaaacca ttttcgtaaa attagaaata aggctgtcca tgggcattta	300
agtgatattg aagacaatgc aaatattctg catgcctatc tgtatcctaat ttgcgcatat	360
ttctataagg aatatagaga cactaatttt tcagcgggaag attatacggg ccctattatg	420
gacattgcct ctaagcctaa agagactgct tcagagactt cagaggataa tgaaaatatt	480
ggagagttta tttcaagtcc attggatgat tatctttttg aaaagtatga tgacagttac	540
ctgttaaagt aactgtctaa acttaaggac tcttcaaagg aagctgttga agacgataac	600
ttaagcgaat ttaaggaata tcttcatatt gacagatcta ttcaagaaga ttttttaaag	660
gcattgaaca gagccactag ttttaattca tctcatctga ttatgctttg cggtagtggt	720
ggagatggga aatcacactt gattgcaaat ttaaaaaaga aaaaccctga actctttaat	780
caatttgcta tccattatga tgctacagag agttttgatc ctgaaaagaa tgcaatagac	840
actttagcct cagtattgga acctttcaat gacaataatc ttaacaattc gacagaaaaa	900
cttatttttag ccatttaactt ggggtgtattg aacaatttcc tggaatcatc ttatgctaatt	960
gaagattaca ctaaactcaa gttaattata gaagaagcga acatatattga atctaattgag	1020
gtttcagata acatttatgg agataagggt agttttgtca ctttcagtga ttacaacatg	1080
tttgaattga atgatgatga aaattcaaac tacacatcat caaagtacat ctcttcccta	1140
tttaataaga taacccaaaa ggaggataga aacccttttt atgttgcata tctcaaggat	1200
aaggactctc actttatcaa tcttataatc tataattatg agatgctgat ggatgaggag	1260
gtccagaaga caatcattga ttatttaatc aagattttca taaaatacag aaaaatcatt	1320
tcaactagag acttattgaa ctctatttat gagataatag ttcctccgga attcctgaaa	1380
agtgaggatt tggacaatat caatgatttc atggactatt cattgcctaa tttactcttc	1440
ggatatccag aaaggtcaga tttattgaag ctatgcaatg aattggatcc gacattgcat	1500
cgtaatgaat ctttagataa gttcataata gacttgaaca ttaatgacga tactgaaaag	1560
atattaaatc gttattttga tttcactcgt ttttaatttcc ttgaagagta tggagagtat	1620
ctggttgatt ttagggagtt taacaattct gaaaaagaga aggtcaccaa tattctaatt	1680
aggtttgcag tattctatgg aaagagcatt atcaagaata atttcaagga taaggtttat	1740
ttaaattatt taaaatatct gtatgcctat aatacacagt ctcataagga ttacaaatat	1800
ctgttccactg aagttaagga cgctattttt aattggaaag gtccctataa gaagaatact	1860
atatgtattg acactttgga ttcattttaa gtgtataaaa attttaaatt aaaaccatct	1920
gttgataaat ttgaaaaaag tttattagat ggtctctttt taggaaatag attttaaaca	1980
gatataaaaa tttatttttc agttgaatct aacaaaaaga aaataccttt aaatgttgat	2040
ttttcattat atcaatacat aatgaaatta tataatgggt ttaagcctaa tcaatcagat	2100

aaagacgatt taataatfff agatgaatff attaataatt tattagatga agatacagat	2160
gatgatffat atgtaattag tttagaaact tatgaagaat tffttatffga gtcaaatgat	2220
tttggcactt ttgaatffta gaggggttaa	2250

<210>138

<211>1372

<212>DNA

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>138

atggatffff cagaaaatta taatatatta ttaaaacaaa tgacttgtga tgtaaataaa	60
aggaaattaa tacatcagat taatcaaaat tctccattgt tacctffftaa aactaatact	120
cctaaaaaag ctaatffftga aaatggffff gatataatat tgggtgaatt atccagaatc	180
ttattaaata aaactattga aaaaaatfff aaattagata atattgtfftc aaatffaatt	240
gataataata ttgaaattga agatggaact aaggagtata taacaaaatt gctaaatgaa	300
tatffttctg atgaaaaaaa tgatfftaaaa atatctcatc cgaatffftga tttatatatt	360
cctctffftcaa ataataaaag ttctaattgga gaacaggaag tagcattgtt tffaagagat	420
atffftffftga agaataatca aaatcttatt aactffftfftg aatcatatga ttcaaatcac	480
attatcttaa atffaatfff aaaaaatact cctaattffac atcataaaat aactgaaact	540
aagtatgtaa ttcattffftga ggaaattgca aatfftatffa atgaagatat caattatgca	600
atactffftata agaaatffftt tatggaaaac attggtaata tfffttgctta ttactatfftc	660
tffftatatct ctcaattaat tctaaaaaatt tctaaaggct ttaatgataa taatgaatff	720
gaaaaattgt actatffftatt ggactgggaa tctgcaagta aaaatcgtaa atcattaaat	780
agttatagtt tactaaaaca tcatcttaaa ccattatatg caaaaatggc agtcatagat	840
caaataaata cactattagg tacaaataat ttattggaaa aagatattfftc agaataatff	900
aataatffftgg acatcaatffc taaaaataat tfffttacatt ttcttaaaaa atgggffftca	960
gattataggt atgtaaggaa tffftgatgat aaagaattac cagataatff attagaatta	1020
actgaaatffc tfffttgaaag tfftaaaaaat gaaaaactgg gtgtggatgg agctgtacaa	1080
tctaggtacg cffftaaatct tgaagatata gctaaaaaatt atffftgttaaa gcgtagaggt	1140
tcgtatggtt atgtattaaa tattaataga gatatgtffgc ttgtffftaac tgcattatgt	1200
gttaaagata agaaatffa gttaaatcaa ctatffftattg aatfftgagaa aagaggagtt	1260
tatffftgaca aatattcaaa agaagaagta gtaaatffftt taactaaatt aaatffaatt	1320
gacaagaaaa gtgatagtgga agatgctcaa tatgttaaac cagttffftatg at	1372

<210>139

<211>5223

<212>DNA

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>139

atgttaaacc agttffftatga ttatctatct aataaattac taaattatff ttgatgatact	60
aagatffftaa gtggtgaaaa atfffttcatt agttffftgatg aagatgacca aataatgtct	120
tffftataata gtttaagaag tattgcagaa actaatffftt cffftgtffftga attcatatat	180

gtccatacta	tttcaggaaa	ggaatataat	acttattcca	taaacattaa	tgaggtcaaa	240
tttgtcattt	ctgaaagttt	aacaattaac	gtagatttct	tagttacttt	aagaaatcaa	300
gttacttctc	aagagggtgt	ttggaaggat	actgcattat	tagttatctg	taatgaagca	360
atagatagta	ttggaaaagg	tatgagaaat	ttacagaaag	aaggaatgcc	attaaatgta	420
aaatctattt	caaaaaattt	agaggatgaa	attaacgatt	ctcaaattct	taactattcc	480
gataagcaaa	tcgctaagtt	ttctttgaat	attcaagaag	aagaattggt	tcaaacaaca	540
ttatgggatt	atgaaactat	cttatcaata	atcaacaagg	gttttggttag	tgatgaagat	600
ttaagagagt	tgaatctttt	taaagatgac	caactaaatc	aaaattctcc	tcaaagatg	660
ttaaaaagat	taaaagaaaa	ttatgacaca	tttaatgaag	taaataaatt	ttcacaatat	720
ggagataaaa	aagaacagct	taaaaatatg	tttactgatt	ctggagtttc	tattttaagt	780
aaagatgact	ggtataaagc	agaatggaaa	atgggttaaaa	aatcaaagga	tgactttatt	840
aatcaacaaa	atccttttaa	ttacaatgaa	aatcttgaaa	aatcacaga	aatggctta	900
aattattggg	aaattcctaa	ctcatttact	aaaactggaa	aaagaaaacg	aaatatcatt	960
gtttttaatc	caaatcattc	tcatgaagtt	agttttaagt	tcagttttga	tcagattctt	1020
agtaattcat	tccttaatac	aaattcaaaa	aaattcacta	ttgcaagagg	caaatcatta	1080
atagtcaatt	ttactctgga	ttctagtga	ccaattttca	aaactataaa	atataaacat	1140
aaaaatgaaa	atatctctga	atttactttt	aatattgttg	ttttgaattt	tgagcctgaa	1200
atttttaatt	ctattaaatc	tcgttttagt	gtcaatgtca	aatctaaaca	gattattgta	1260
acaaatgatg	aggatagttt	tgatattggt	tttggtactg	gttcaaaaga	aatagagaaa	1320
ttaattaaag	aaaatgggtga	aaagttatat	ctttatgatg	atgaatcttt	aataatttct	1380
gaacaatctc	ctgcttggaa	tgatggaaaa	ttaagtttta	aattatataa	agacaataat	1440
tatgccccat	ttttgattaa	agaaaaatct	aagaaaacac	tgccagttaa	ttcctatgta	1500
atttgggaatt	tgaaaagaag	aaatatggaa	aactttattt	ttaatggagt	aaaagctggt	1560
caagatgtaa	acagtttcta	tctcgtagaa	gaattttaag	aattttctaaa	aatggaacga	1620
gaaatcatta	aacaggacat	attttatgct	aaaaggaata	ttgatgggtc	tttagaaaaa	1680
attgaagttt	cattcagtaa	cgagttagaa	actgcttaca	tgacatatt	taattattat	1740
aagacttttg	atgattctcc	agaagacaat	cttccaagtt	tagttttatt	aaatgatgat	1800
ttaaaagagt	tatataagaa	atttataact	atcttcaata	aggaaatata	agaaattgaa	1860
gagaattcta	ttttgtctga	ttttaaatat	aaaaagaatt	tgcttaagtt	aggctcgtatt	1920
gaaactgata	acaaaataat	gtattctcca	ttatctccac	taaatatcgc	ttatcaatta	1980
gaagtttcta	aacaatgcgg	aaatgaggat	ttatccgtaa	atatattaga	acggtttagtt	2040
ccaaataact	taattcctta	tatttgttca	gatgatggaa	aagaattggt	tagacctatt	2100
tatcaagaag	aagctcatga	atggttgatt	tatgagaaaa	gtgaagaagt	atcaattgga	2160
acaacaaatg	tatttatttc	taatgttggt	actgaaaaac	ttaatcaatt	tgtcaaacat	2220
tttaattact	tatttagttt	taacaactcc	tcaccaatta	aaattaactt	aatcaatata	2280
aaagatgata	aagaagttgt	gaaaggagta	tttaatttta	ttagatctag	attgcctgat	2340
aaaactaaaa	caaaaaaagt	tattcctgtt	gaaatcaata	tttacaacga	cgtgaaaaag	2400
agttcttttg	acaattttatt	tgattgtcag	tctgaaattc	agttactgga	agaatttgga	2460
attaaaaaat	taaaatcaga	tatttttgat	cctatagata	ttatccatat	gattcaaaaat	2520

aatatttcat	attataaaca	tccattttaa	aaagaagaat	atgaatatgc	tcattcttct	2580
ttctataaag	tcaaatcaca	taacaatatt	gcaaatgaca	atatggataa	aattgaaact	2640
ggattatctt	taaacgggtt	attatcttca	gttacatcaa	ctactaaaca	ttcagaatat	2700
agaacgggtt	tcggtacaaa	caatatTTTT	aatatgagta	atcctcttat	taaaacagtg	2760
attaacttga	acgaacttgt	tgagaatagc	aagaactttg	gaaaaaatac	ttattcaaaa	2820
aataaatcag	ttatcactac	tgttgaatta	gaagaagaca	atattgaaga	attgtacgac	2880
aaatctcatt	gggtaacatt	tattgaacct	acatttggta	tcgaatactt	tgacagttct	2940
gatagtaatt	tgatcattat	tcattacagc	gatcaatata	gttcctctag	caagtatgat	3000
actattactg	taactaataa	atcaacacaa	tatgaggaaa	ttataagaga	tttcctccaa	3060
tctaaatatg	taaaagttac	agatgaagaa	ttgtatgatg	taataaagat	gtttaattca	3120
attaatgggtg	aatggttact	tagagtaatt	tctaattccg	gccattatga	tagggaaaaa	3180
ttaagtatta	tttctgctat	taaatattgc	ttatccattt	tggaccataa	agatattggt	3240
tggattccag	tttctatgga	agaaatctta	agaattgcag	gaaatgttaa	attagataaa	3300
aataaaggaa	tttttgactc	taagtttaata	aaaggaaatc	atagtgatga	tctgctattc	3360
attgggtgtaa	aattcaatga	agataataga	atcgaagtta	tattctaccc	aatcgaagtt	3420
aaaatagggtt	tgaataatgc	ttcaactatt	aaaaaaggta	aaagtcaatt	agataaact	3480
tataaacttc	ttaaaactca	actacaaaat	attaatgtag	agaattctga	atttagaaat	3540
aaattcttta	gaaatttctt	catccaaata	ttattgtcca	atgagcaaaa	attagtaact	3600
aatcatattt	gggatgaaaa	aggattggat	agaattgaag	aattcaaagc	agaactatta	3660
aatgatgaat	atgatatttt	atatggttta	gaagagtaca	ttggtaaagg	ttcttttggt	3720
tcattttaaaa	aagaatctca	tcacatctct	atctatatgg	atgttgataa	gcaagtaatt	3780
gagttaccag	aagattttgc	ttattacggt	ttagccactc	ctataagaga	aatccatgat	3840
gagattcagt	cagacaatac	agatattctt	gcagaaacat	tgttatcaca	tgtagatata	3900
agtgaatta	gagcaaagaa	taatgatata	tgcgattcaa	atgaggatat	gtcaatagat	3960
gatgattttg	ataattttaag	tgaatttgag	gatagtttta	ttgaagaaga	atcagaaatt	4020
tctgaagagc	ctgatgaaga	attaactgat	gcaacttcaa	gcagtataa	tgaatcaatt	4080
gagaatattg	gagaatctcc	ttctaaaatt	tctaattgtca	gagcattaat	tggaactcag	4140
aagggttata	atcataaagt	ttattgggaa	tttgacatc	cttcttttagc	aaatagggcat	4200
atgttaattc	aaggaaaatc	tggacaagggt	aaaacatatt	tcattcaacg	aatgttaaaaa	4260
gagttgtcta	ttcaaggcat	tcctagcata	attattgatt	atactgatgg	attcaagcct	4320
tctcaattag	aacctaat	ttaaagactct	ttaggagata	aaataagcca	atattttggt	4380
gtaaaagaaa	atttcccaat	aaatccattt	aaacgaaaca	ctatcatgat	tgataaggac	4440
atattttattg	aagaagacaa	cagcactatt	gcgagccgat	ttaaatctat	cattaattct	4500
gtttatgggc	taggtattca	acagtccaat	actttatatac	aaacagtttt	agattgcttg	4560
gataagtatg	atgataat	tgatttgaat	atcctaaaag	aagaaattct	aaaagacgaa	4620
tcaaatagtg	cacaaacagt	cttgaataaa	ttaaatgagt	tattggataa	aaatccgttc	4680
gcattccactg	atttcgactg	gtctgtctta	gataataaag	acggtaaagt	ttacattatc	4740
caattaactg	cactttctaa	ggatatacag	acaattataa	ctgaattcat	cctttgggat	4800
ttatggaatt	acaaattaac	aaatggtagt	gaagataatc	ctttcattgt	tgttttagat	4860

gaagctcata atcttgattt ttccaatgac tctccatgta gtaaaattct taaagaaggc	4920
aggaaatttg gatggctctgg gtggtttgca acacaatctg ttaaaggatc tatgaaaatt	4980
gatgaaatag ctaaattaga gaatgcagat gagaagatat atttccatcc aacggatgtt	5040
tcaacaatag ctaaagattt gtcaaaagac aatgaagata aaaaaatata tgaaaaagag	5100
ttatctcaat tgacaaaagg atattgtatt gttcaagggt ctgcaataga ttcaagtggc	5160
aatttgtatc agccaaatcc tgtaactgta aaaattgagg agataagttt tgatgaaaat	5220
taa	5223

<210>140

<211>3165

<212>DNA

<213> 反刍甲烷短杆菌 (*Methanobrevibacter ruminantium*)

<400>140

atggaaagta aagacagaat agagaatact attttcaatg ataaaagaat ggatgcttat	60
atagacaaat attttgaaga gtttactctg agttctgaac aagaagaggc actgaatatt	120
tggatagata aattaaataa cgatcaatta acaagtgaag aagggaatta tcataatttt	180
tttgaaatta ttcttgaaga ttacttgggt tataaacgtt ctgatgttaa acatgaagag	240
aatattgggtg atgaaggcca tcctgtagag ttgtatttag aaaaagatgg aaaagattat	300
gtgattatag aactaaaagg aacaacttat aaagacctca ctaaaagaag acctggacag	360
caatcaccag tagagcaagc tacaatttat gctagtgcta aaaaagaaac tgaatgggct	420
acagtttcaa attatgatga atttagattt tttaatccaa cagcaagaga taattatatt	480
tcattcaagt ttaggcaatt gaaagatttg gaaatcttta aaaaattttt attagtattc	540
agtaaatattt cacttatcga tgaagacata cctaaaaaat tacttaatga aacaaaagtc	600
attgaaagag aattagaaaa tgaattttat caattgtata gtgacactag attaattgatt	660
attaaggaat tggaatattc ttcagaagat attaatagga ttgaagcaat aaaattatct	720
caaataatat taaatagatt tattttcctt tgttttgcag aagatttagc gcttatggag	780
gaagagacaa ctgctgatgt attattaaca cctttaaagc atagaaattt aatcggaaat	840
acaatgtgga ataggttaaa tgaattattt atttttgcaa atcaaggaaa taaacatagg	900
agaataacctg catttaatgg aggattattt gaggatgatt tatctaattt gaaaataagg	960
gatgaaattg aagataggtc ttcttttgaa aattggaatt taaaagaaga ttttgaagac	1020
aaatatgagg atattgctaa actaattgga gtttataaag atactcttaa tccaattttc	1080
attaattttat taataatttc tacatatgac tttagttcag aacttgatgt aaatatctta	1140
ggacatatatt ttgaaaacag tatttagcgat attgaagagt taaaaaatga taatcaagag	1200
caaagaaaaa aagatggagt gtattacact ccagaatata ttacagatta tttttgtaga	1260
aatacaatta ttccatattt aagtatttct ggtaaagcca gcacagttca tgaattatta	1320
tacgaatatg aatcatctaa ttcatggat gttttagatt caaaattaac taacatcaaa	1380
gttctagatc ctgcctgtgg tagtggggagc atgttaaata aatcagtaga tatattattt	1440
gaaattcatg aagcattaca tgcaagtaaa tatgctggag attcatcttt agatagattt	1500
tttgatagtt tagaaaaacg aaaggaaatt ataagtaata atatttatgg gggtgatttg	1560
aacgaggagt ccgttgaaat tactaaatta tccttattcc taaaattagc aactactgtt	1620

```

ggacttaaag aagggtttca acttcctagt ttagataaac atattaaatg tggggattca 1680
ttagtggatg atgagtcaat tgccggaaat aaagcattta actggtatga atcattttca 1740
gaagtttttg aaagtggcgg atttgatatt attgtaggaa atcctccgta tgtggatatt 1800
aaagaaatgg atgaaaaaac agcaaaatac atatttgata attatgaaac ttcttttaac 1860
aggataaatt tatattctac atttgtggaa aaaagttatt atttgctaaa aaatgaaggg 1920
atTTTTtcat ttatcatgcc taattccatt ttatttaatt ctacttactc taaaatcaga 1980
gaactaattc ttaataatac atccatatta aatattgtaa gaacttctga tgatgttttt 2040
aaagatgcta aagtggaaacc tattatTTTta atcttcaaaa aaggatatga tgaaggaaat 2100
aagactaaga tacttataaa aaaagatgat atggatgaaa ttccaataaa taattattca 2160
gaacatttct ttacacaaga aagatggttt gaaaataatt caataattaa tattttttca 2220
gatgatttta cttttgattt attaaagaaa attgatggaa ataatgaaag attaattgat 2280
tattgtgatt ttagtttagg tttaactcct tatgataaat ataaaggatg gagtgaggat 2340
attattaaaa atagaaaatt ccattcaaaa ataaaacttg acgatacttt taaagaatta 2400
ttggatgggt ccgatattac aagatataat gtaaaatggg gcgaaaaaga atatattaaa 2460
tatggggact ggtaggagc acctcgggaa gaaaaattct ttaaaaatcc tagaatttta 2520
ataagacaga ttttatctat agctcctaag gaatctcgaa aaaggatttt tgccgcttac 2580
acagaagaag aattatataa cgctcaaatt gcttttaatt tagtattgaa agaaggcttt 2640
gatgataaaa atttattgaa atatttttta ggaataatta attctaaaat gatgacttgg 2700
tattatgaag aaagatttat ggataaaaat aaaaagaatt ttgcaaaaat tttgattgaa 2760
aatgctaaaa atcttcagat cattattaat tcaaattttt tagatgagat agtatctaata 2820
gtggattcaa taatagagtt aaataaagag ttttatagtg taagaaatgc atttcaaaca 2880
tggtttaaaa tagaatttga aattgaaaaa ctctctaaga aactagaaaa ctattatgat 2940
ttaaactttg aagaattctt aaaagagata aaaaagaaaa aagtagtcat tagaccaaat 3000
caaatacagg acttgtccga attatttaata gaaagtttag gaaaaataga atatctgcaa 3060
agagagatta aagaagcaga cgaaaaaatt aacctactgg tctatgaatt atatggttta 3120
aatcatgaag aatagaaat tatagaaaat agctttaatg attaa 3165

```

<210>141

<211>261

<212>PRT

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>141

```

Ala Thr Gly Gly Ala Cys Gly Ala Thr Gly Ala Ala Ala Cys Ala Thr
1           5           10          15
Thr Ala Ala Thr Cys Cys Thr Ala Ala Thr Thr Gly Ala Ala Thr Ala
20          25          30
Cys Ala Thr Cys Ala Gly Ala Ala Ala Cys Gly Cys Gly Cys Cys Ala
35          40          45
Ala Cys Ala Ala Gly Ala Gly Ala Ala Ala Thr Gly Gly Thr Gly Cys
50          55          60

```

Thr Thr Ala Ala Gly Thr Cys Ala Thr Thr Thr Gly Ala Ala Gly Gly	
65	70
Ala Gly Thr Gly Gly Ala Cys Thr Thr Thr Ala Thr Ala Ala Gly Ala	
	85
Cys Cys Ala Ala Thr Cys Cys Ala Ala Ala Thr Thr Thr Cys Thr Ala	
	100
Gly Gly Ala Ala Ala Ala Cys Ala Gly Gly Cys Ala Thr Thr Cys Ala	
	115
Cys Cys Cys Ala Ala Ala Cys Ala Ala Thr Gly Thr Cys Ala Gly Cys	
	130
Ala Ala Gly Ala Ala Ala Thr Thr Ala Ala Ala Gly Gly Ala Thr Cys	
	145
Thr Thr Cys Gly Cys Gly Ala Ala Cys Ala Thr Gly Ala Ala Cys Thr	
	165
Gly Gly Thr Cys Thr Ala Thr Gly Thr Cys Ala Thr Ala Ala Ala Thr	
	180
Cys Cys Thr Gly Ala Gly Thr Ala Cys Cys Ala Thr Gly Thr Thr Cys	
	195
Cys Thr Ala Ala Ala Cys Thr Thr Thr Ala Cys Ala Gly Gly Cys Thr	
	210
Thr Ala Cys Thr Gly Ala Ala Ala Ala Gly Gly Gly Thr Ala Ala Ala	
	225
Ala Ala Thr Ala Thr Gly Cys Thr Thr Cys Ala Ala Thr Thr Cys Thr	
	245
Thr Gly Thr Ala Gly	
	260

<210>142

<211>255

<212>DNA

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)

<400>142

atgaataaaa aaattatctt atccctcctt ttagtattat tagtagctat ttctgtctct	60
gcagttgcag cagcagatgc tgatgtcaca tatataaacg atgctgcaga ttagacgat	120
gttgacagc aaaaagttgc tcctcttaca gctagtgtg atgcacaaga catccaaact	180
aagcttgata atgctaaacc tggagacaca attgaattag aaaacaagac atatgacgtt	240
gataacaacat ttaat	255

<210>143

<211>40

<212>DNA

<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)	
<400>143	
tttttaattg attgtaataa ttaattattc tgggtctgac	40
<210>144	
<211>80	
<212>DNA	
<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)	
<400>144	
aaacctttat ggaagtcgag taagtattgg tacgtattat aagtataaat acgtatccta	60
tacttacttt attaaaattt	80
<210>145	
<211>42	
<212>DNA	
<213> 反刍甲烷短杆菌 (Methanobrevibacter ruminantium)	
<400>145	
cagaataagg ataataatta attatattgt ttttattttt tt	42
<210>146	
<211>25	
<212>DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述 :合成引物	
<400>146	
caaagagaga tttaaagaagc agacg	25
<210>147	
<211>25	
<212>DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述 :合成引物	
<400>147	
agtagtggtg gaatcagtga aaagg	25
<210>148	
<211>25	
<212>DNA	
<213> 人工序列	
<220>	
<223> 人工序列描述 :合成引物	
<400>148	

caccatgggtt agattcagca gagac 25
<210>149
<211>25
<212>DNA
<213> 人工序列
<220>
<223> 人工序列描述 :合成引物
<400>149
tcatgcagga cagacaacat agtag 25
<210>150
<211>6
<212>PRT
<213> 人工序列
<220>
<223> 人工序列描述 :合成的 6xHis 标签
<400>150
His His His His His His
1 5
<210>151
<211>39
<212>PRT
<213> 人工序列
<220>
<223> 人工序列描述 :合成的共有序列
<220>
<221>MOD_RES
<222>(1).. (1)
<223>Met 或 Leu
<220>
<221>MOD_RES
<222>(6).. (6)
<223> 可能或可能不存在
<220>
<221>MOD_RES
<222>(7).. (15)
<223> 任何氨基酸 ;该区域可包括 0-9 个残基
<220>
<221>MOD_RES
<222>(16).. (16)

<223>Ile 或 Leu
<220>
<221>MOD_RES
<222>(17).. (17)
<223>Ile、Phe 或 Leu
<220>
<221>MOD_RES
<222>(18).. (19)
<223>Ile 或 Leu
<220>
<221>MOD_RES
<222>(20).. (20)
<223>Ile 或 Ser
<220>
<221>MOD_RES
<222>(21).. (21)
<223>Leu、Ile 或 Ala
<220>
<221>MOD_RES
<222>(22).. (25)
<223> 任何氨基酸 ;该区域可包括 0-4 个残基
<220>
<221>MOD_RES
<222>(26).. (26)
<223>Leu、Ile、Val 或 Phe
<220>
<221>MOD_RES
<222>(27).. (27)
<223>Leu、Ile、Ala 或 Val
<220>
<221>MOD_RES
<222>(28).. (28)
<223>Leu 或 Ile
<220>
<221>MOD_RES
<222>(29).. (29)
<223>Ile、Leu 或 Val
<220>
<221>MOD_RES

<222>(30).. (30)

<223>Leu、Ala、Ile 或 Val

<220>

<221>MOD_RES

<222>(31).. (32)

<223>Ile、Leu、Phe 或 Val

<220>

<221>MOD_RES

<222>(33).. (33)

<223>Ser、Ala 或 Leu

<220>

<221>MOD_RES

<222>(34).. (34)

<223>Ile、Leu 或 Val

<220>

<221>MOD_RES

<222>(35).. (35)

<223>Gly、Ser 或 Ala

<220>

<221>MOD_RES

<222>(36).. (36)

<223>Ala 或 Ser

<220>

<221>MOD_RES

<222>(37).. (37)

<223>Val、Ala 或 Ile

<220>

<221>MOD_RES

<222>(38).. (38)

<223>Ser 或 Ala

<400>151

Xaa Lys Lys Lys Lys Lys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

1

5

10

15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

20

25

30

Xaa Xaa Xaa Xaa Xaa Xaa Ala

35

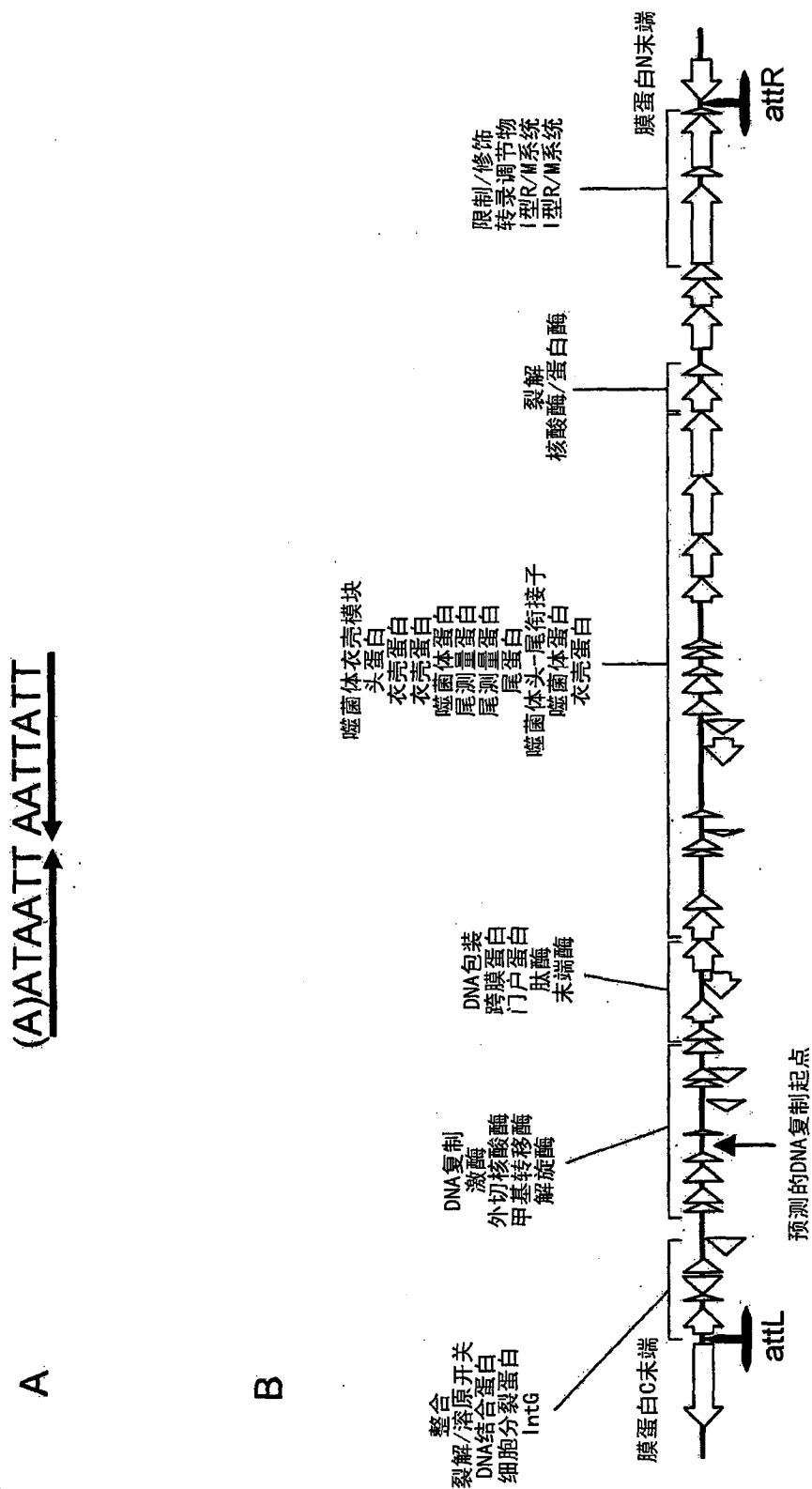


图 1

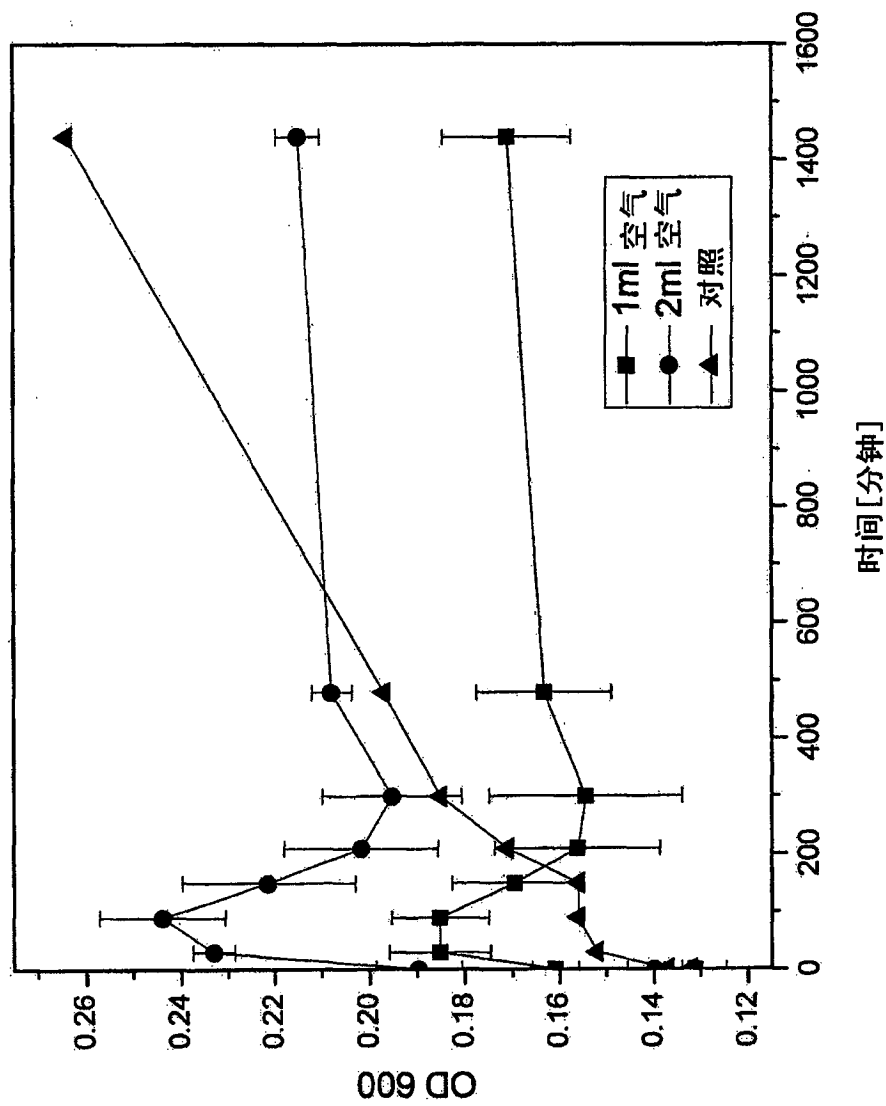


图 1C

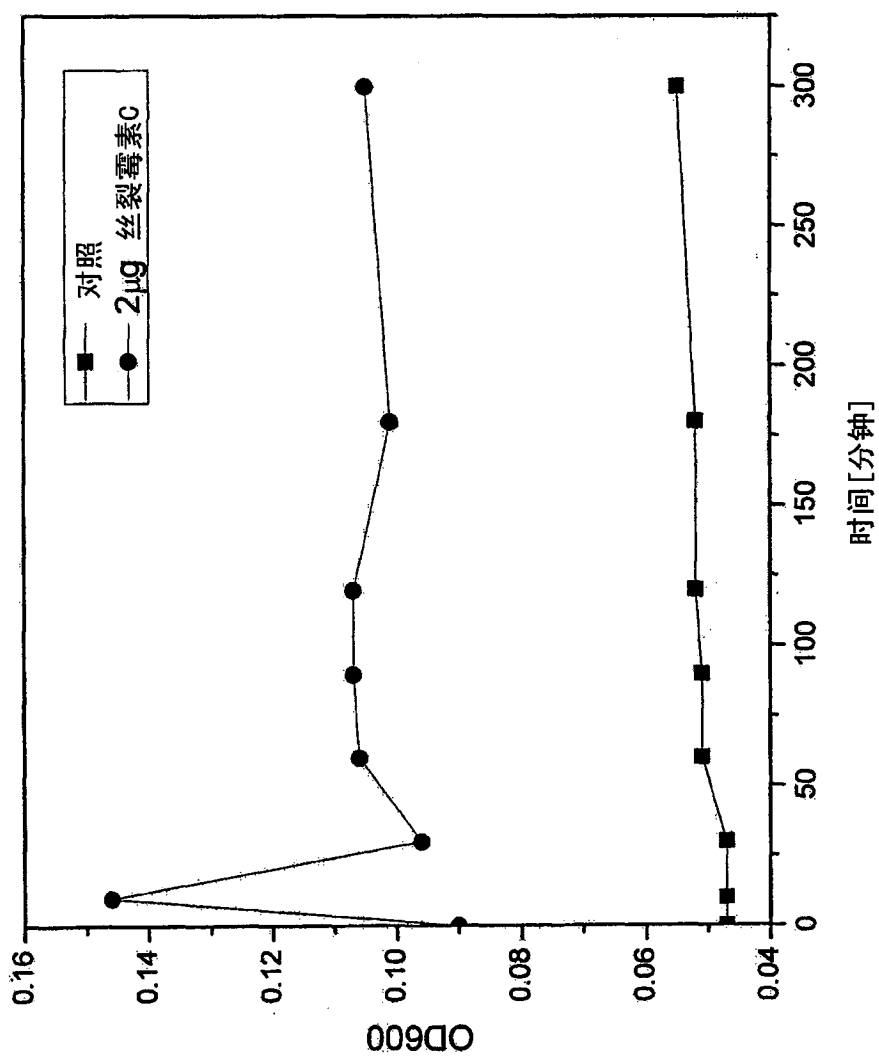


图 1D

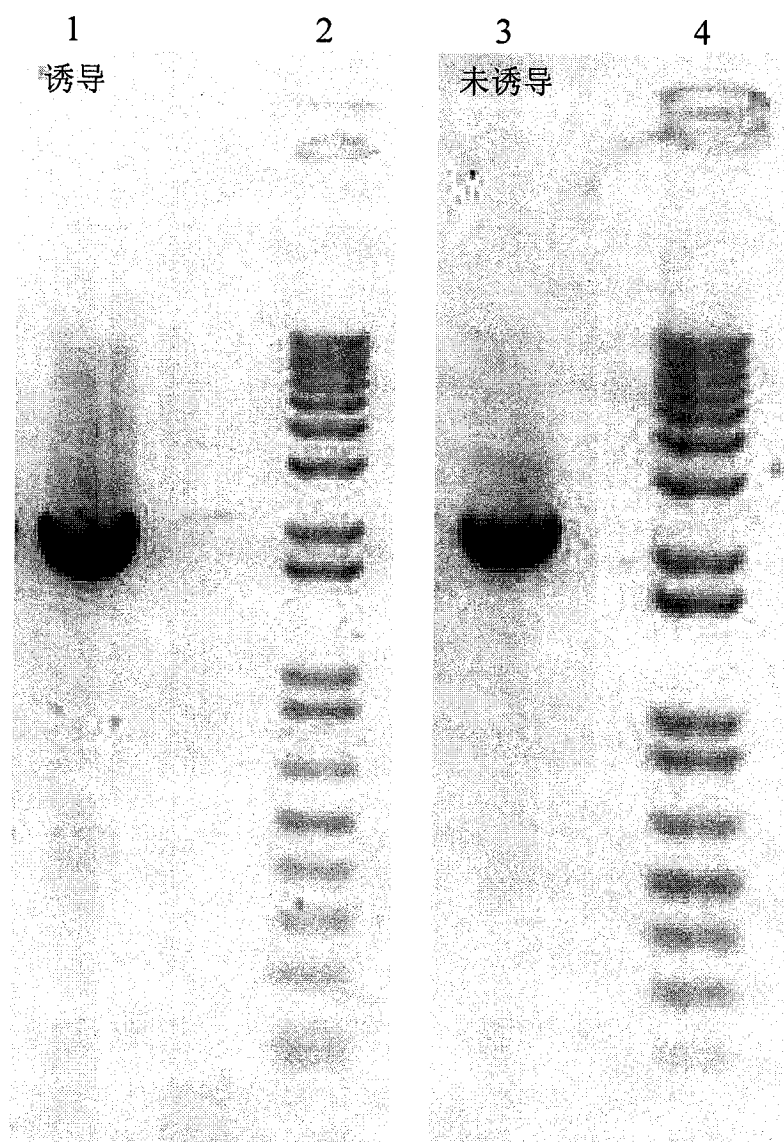


图 1E

原噬菌体 ϕ mru 开放阅读框的注释和注解

ORF 编号	注释	注解
896	可能的顿挫型感染机制, CAAx 氨基末端蛋白酶家族	噬菌体防御机制
897+898	转录调节物	可能作用于 ORF896, 很可能是移码
1451	预测的噬菌体相关性位点特异性整合酶	1450 显示与 DNA 引发酶具有相似性, 1452 显示与噬菌体蛋白具有相似性, 基因组同线性提示有基因簇, 即可动遗传因子的整体出现
1572	可能的顿挫型感染机制, CAAx 氨基末端蛋白酶家族	噬菌体防御机制
1578	可能的顿挫型感染机制, CAAx 氨基末端蛋白酶家族	噬菌体防御机制
1579	可能的顿挫型感染机制, CAAx 氨基末端蛋白酶家族	噬菌体防御机制
1783	噬菌体结构蛋白	基因簇 1776-1789 中的许多基因显示与噬菌体蛋白具有相似性。1989 位于毗连群边界处, 这表明这种噬菌体残迹的连续性。
1774+1775	类似于限制性内切核酸酶的 GTP 酶亚基	两种基因类似 (噬菌体特异性) R/M 系统的一部分: 根据基因同线性, 预测基因簇 1772-1775 参与噬菌体防御。
1959	可能的顿挫型感染机制, CAAx 氨基末端蛋白酶家族	噬菌体防御机制

图 2

133

原噬菌体 ϕ mr_u 开放阅读框的注释、预测功能和注解

GAMOLA ORF	注释	预测功能	注解
侧接基因组区, 基因组整合基因座			
1980	富含 asn/thr 的大蛋白质家族	细胞粘附	噬菌体整合破坏的表面蛋白, N 末端对应于 ORF 2069
	attL	核心以粗体表示	Tttttaattgattgtaataattatttcttggtotgao (SEQ ID NO:143)
整合裂解/溶原开关			
1981	整合酶	噬菌体整合酶	
1983	构建的假拟蛋白		
1984	构建的假拟蛋白	DNA 相互作用	可能带有锌结合结构域和裂解-溶原开关的一部分
1985	细胞分裂控制蛋白	DNA 相互作用	可能是裂解-溶原开关的一部分
DNA 复制			
1988	构建的假拟蛋白	DNA 复制	DNA 解旋酶或外切核酸酶, 相似
1989	构建的假拟蛋白	DNA 复制	
1990	未知	DNA 复制	
1991	未知	DNA 复制	
1992	未知	DNA 复制	预测为 1 TMH (跨膜螺旋), 可能是膜蛋白
1993	外切核酸酶	DNA 复制	类似于外切核酸酶
1995	构建的假拟蛋白	DNA 复制	
1996	构建的假拟蛋白	DNA 复制	
1998	ATP 依赖性 DNA 解旋酶	DNA 复制	参与复制控制的 ATP 酶结构域
1999	构建的假拟蛋白	DNA 复制	

图 3-1

GAMOLA ORF	注释	预测功能	注解
	终止子结构	类似于复制起点	AAACGTTTATGGAAG TCGAGTAAGTATTGGTAGGTATTATA AGTA TAAATACGTATCGTATAGTTTAA TTAAAAATTI (SEQ ID NO:144)
2000	构建的假拟蛋白	DNA 复制	
2001	未知	DNA 复制	
2002	构建的假拟蛋白	DNA 复制	
2003	未知	DNA 复制	
2004	未知	DNA 复制	
2005	噬菌体关联蛋白	DNA 复制	
2006	构建的假拟蛋白	反义方向	2007 的相邻终止子结构
2007	构建的假拟蛋白	DNA 复制	预测为 2 TMH; 可能是胞质膜蛋白
2008	通透酶	DNA 复制	预测为 2 TMH; 与通透酶超家族相似; 2007 和 2008 可能有移码
2010	ParB- 样分配蛋白 (partition protein)	DNA 复制	切割单链 DNA 并在富 AT 区域切开超螺旋质粒 DNA
2012	构建的假拟蛋白	DNA 复制	
2013	完整的膜蛋白	DNA 复制	预测为 5 TMH
DNA 包装			
2014	构建的假拟蛋白	DNA 包装	末端酶小亚基; 与 2015 竞争
2016	末端酶, 大亚基	DNA 包装	
2017	构建的假拟蛋白	反义方向	
2018	门户蛋白	DNA 包装	
噬菌体衣壳模块			
2019	小头蛋白	结构模块	头形态发生蛋白
2021	构建的假拟蛋白	结构模块	
2022	噬菌体关联蛋白	结构模块	也类似于 ATP 酶

图 3-2

GAMOLA ORF	注释	预测功能	注解
2023	未知	结构模块	
2024	构建的假拟蛋白	结构模块	
2025	构建的假拟蛋白	结构模块	预测为 1 TMH
2026	未知	结构模块	
2028	构建的假拟蛋白	结构模块	
2029	未知	结构模块	
2031	支架蛋白	结构模块	类似于支架蛋白结构域
2032	构建的假拟蛋白	结构模块	
2033	未知	结构模块	
2034	未知	结构模块	
2035	构建的假拟蛋白	结构模块	
2036	未知	结构模块	
2037	未知	结构模块	
2038	噬菌体头-尾衔接子	反义方向	与衔接子结构域相似
2039	噬菌体相关蛋白	反义方向	预测为 1 TMH
2040	噬菌体衣壳蛋白	结构模块	与结构蛋白相似
2041	大噬菌体衣壳蛋白	结构模块	
2042	构建的假拟蛋白	结构模块	预测为 1 TMH, 可能的信号序列
2043	构建的假拟蛋白	结构模块	
2044	噬菌体相关蛋白	结构模块	
2045	构建的假拟蛋白	结构模块	
2046	Gna_B 胶原结合表面蛋白	结构模块	预测参与噬菌体停靠
2047	构建的假拟蛋白	结构模块	
2049	噬菌体尾测量蛋白	结构模块	相似, 可能由 ORF2051 移码得来
2052	噬菌体尾蛋白	结构模块	噬菌体尾组件

图 3-3

GAMOLA ORF	注释	预测功能	注解
2053	小噬菌体结构蛋白	结构模块	
裂解			
2055	噬菌体相关性表面蛋白	裂解	类似于亲密素/根据同线性可能用作穴蛋白
2058	肽酶		根据位置/同形性预测为溶素
溶原转变			
2061	dnd 系统相关蛋白 3	DNA 硫修饰系统, dnd (电泳期 间降解)	
2062	dnd 系统相关蛋白 1	DNA 硫修饰系统, dnd (电泳期 间降解)	
2063	dnd 系统相关蛋白 2	DNA 硫修饰系统, dnd (电泳期 间降解)	
2064	II 型限制性 m6 腺嘌呤 DNA 甲基转移酶	甲基转移酶	
噬菌体整合位点, 基因组整合基因座			
	attR		Cagaataaggataataataattatattgtttttttttttt (SEQ ID NO: 145), 终止子结构的一部分
2069	富含 asn/thr 的大蛋白质家族	细胞粘附	噬菌体整合破坏的表面蛋白, C 末端对应于 ORF 1980

图 3-4

GAMOLA ORF	核苷酸序列/描述	SEQ ID NO:
侧接基因组区, 基因组整合基因座		
1980	gtgagaaatatgaagaataaagaggtttaattcttattactgattacaataaagaacataggaatctgttcttgcaacgga taatgaagaaataataggataaataaataatttgataaataatgagatatcgctaatattgataatgtcgataatgtcgataact ccaatataaacaatccaactgacataagaatagacaattcaaacctaaatagagaaacagaaactagattcaaatttaataaataatct	74
整合裂解/溶原开关		
1981	atgaagacttttagacaacagctattggaagaccctgaattccagaactacctgttgacagacccgaacctgacggagagcagcctgca gtcatacctcaatgccgcccaactctcgtaggttcacaggggagccgttctacaagaccgtgcattgagctcagaagccagcagaacg ataggattgagaacaacatcatcataagggttcaaccgcgaaccagtcaggataaaacatcatgcagtttgagttcata	75
1983	atgaatacattaaaaatagaatgttcaaaagatgatataatcataaaagactgcaaaagcgaattctgaaaaatgagttgaaagt tcctgagaactggaactcgactatgtgaatgctgctgtggaagagagacatctgcgaagtgttgaaaaagggtgatgaaagaatgc ttttgattccctatgtgcgtgactgcttcttgaggagtgaggaagatgaatacataaagtacatttgcccttccct	76
1984	atgggaatgaaaaaagttaagatcgaaagtaaaagcaaaaaataatgacaagaaacgaaaaactctctataaaattatatgacgattt atataatgaagatagatttaattttattctccacataatttaacataatgataaaatttttttaaaaaagataatccattatgttt taatgaactactctgaaaaatgaattattgaaagcaatgaaaaaatttgataattcagtcagaaggcattgatatt	77
1985	atgaaaacacaagatctaatataataaataatgataaaatgaggaatctcctgtattcctcaacaggggaagtcttgagatggactatgtgcc ggacatctacaaaacagggagcagcagctggcgaaaattggcgatgtactgcaattcaatacctgacaaacatagctcccaagaacctgc aattgtgcggaggcaatgcgacaggaagagaccacaacattaaagcagttcttcaagatgttgaaacgaggttttcca	78
DNA复制		
1988	atgtttttggaaaaatgcgatggaaaagatacaaatcagcatgtctcttcaagaaaaagatgaatcttatattggaaactatggaaagcaa aggaagcccttcataatcctgtattcattgttaacatacctataaaatgaagcagaagaatggtacaaaaatggaaaaataggagatcagg actttataaaatttctatgatgatgataaacttgattgaagaaagtgttggtttgaaatttataaaaaatcagaatat	79
1989	ttgataaattttccttaaaactaaaaatttgacgtgtaatcatggaaaaatgaatcgaaattgacggagttcaatacacagaagcccaaat cagaaggccttttagccattgaacgagacgtgcgctcctcctaatttcgttgatatgttttggaagaagatcaaacctagcgaacttgcaa gcagagtttctgaaaaaagtgatgcctaa	80
1990	atgtgctatgtgggaaacacacaggaacttggtctatcacacagaggaactgcttctgcaacacactggctgttgaaacgagaacaagaccat tctagaggaagaaagcctgtagacatgaaaccttgacgttctgcaaacacacagtttgacactgaaatag	81
1991	gtgttgatattggttcagagatggttgaaaaacatcaggaagaggaagagatggatatccatctacccctccttttctcctgattgt tttccctacaaggcaaaaagtattccgaatgctgctgcaaggtgaggaaggatcagaagtttgcaattgtgaaccttgtaggcttca gaagagaggtgtcaaggtcatagaccacaaggaccacagagaggtgttcgtatga	82

图 4A-1

GAMOLA ORF	核苷酸序列/描述	SEQ ID NO:
1992	atgagcatgcttgagagacttgcagcctgcaggctccacaaggagacatgggctgaaaggcatgatgtcgaaatccttgcgtcatatg ccttgccataagcattgcaatgctactcttcttcttgcttgagagccgactgttgaggagtgatttag	83
1993	atgactaaggaaatttgaggacttcatgagaaggaaacacagggactgcttgtttcttgagatgggacacagtgcatatctgaaataacct tgagtcctattatgacgaaaggaaatgccttacagattgttgaggcaatgcacaatctctttgacttctaccagataacct tcagcaaaaagactgaaagggaattctgcaccagcttttgaaaaggacaagataaacaaggcttcccttcacac	84
1995	atgaatgagatgagtactacaactaacgagaataatgtacctgtgcagcttgactatgccatcgaggaatggaaggcataccagcgatt gaccagagactattggagaaacagattatcagacacacagagcgaggaagtacaagccaagagcgcttgccagaagatcgcacgtg ccttcaacataaacacacagataaatcgacaaggaaatcgtaagaataacgtaagggaatcgctcattgagctgagtat	85
1996	atgattgagtgatacaggaagaggagacttcacggattgggaggttctctcctcctctccgacatcaagtcacatagtttcagtgga tgatgagggaaacctgttctgcagctgctgcactctattacaggaagtcaggatgaaccttcacatctcaaaccttgaatcctatt gcaaacacatcagacaagctctggaggaggacaatagattgcaaatgctctag	86
1998	atggataataataaagacaaaacagtttagtctaagtttgaaagagtttttcagcactgtatacaaggacgaagtcatggaaagtttt ggaaaagtatcctgaggaaaggacattggttgtagctatgagaatttgaaaattgttgactcctgatttggcagacttgttgattgaaa agcctgacgaagttaatcgagcttcacagaaagcaatcaagaacattgacctttagtgaagatccaaaagttggac	87
1999	atggtaacaagattagagtttaattctcacggttgacctaatgttgccaattagcaaaagacaagttaccttgcagcagaagcgaatt ctttgagaatcagttgaagatgttcttagggattgaagatgcagagtcggaaatcattaaaggacatccaaaccaaagaaaatgagatta acgcaataaggacaaagttatgtcatgttcgcaaatcaagcagttgaaattggaatcaacaatactatggagaag	88
2000	atgacactaaggaaatttgacaattttggaaagtgcagatgggagaaacatgaaccttgaaagatttcgactgttctgtttctttgaaat gtacaaggcacttttgagatattggatgttgagttggaagttggcagcttcgctgaacttcgagatgtgtacaagaatgtcgagatggattaca ccttgaagagacatgccctcaagcagaaggaaatcctttactggttcaatccgacttgaaggaggaaacttggaaaag	89
2001	atgaatgtgaaacagtcacgaatgaccttataggattgtcaaaaggagttcgagggagtggaatatgagattgaaatccaaaattcaat ctacttctattccttccaaagtaacatgaaggaaggcattgtgatctgaaatatccgcaatctatgacctgcacacccatattgaaaag gcatggatggaatcatagtggacattctggagggttgaggacaatcctggtgatgagaagaagatttgcctatatgtg	90
2002	atgattagcagaagaatggaaagaggaatactatgtcaaggtcaacgaggaactggagcaggttggaagtcaggtttttcagcaaggtcga caggctcgtcttctgctcaatcctactcttctcatcttcttcttcttggagggttgaatcctgtgcgacaggatatgcgacatac tgacaaacgatttgggaaatgccaaagtattatctaggaggaaagatgaattatag	91
2003	atgaattatagggaatgggttgcaagccagttcgagctgaggtgactgaggacggaagggttgccttcaagacaccatgcagctacta cacattttcaaggaggagactttgagataataaaggaggagatttctctaatttcgagtcctga	92
2004	atggatagaataaaggaaattgaaatgtatcggaacctgcgaacagcaccatttgatttcagagacatcaacggagagatagcagtgaaacat ttgcaggaataggttcagaggctgtcaacaaggcaggagacttacatatcttgcgttagactggacaccaaaggaggaagctgtcatga	93
2005	atgttaagtaaaaaagaagccatacagatgcgctggacaattgagaagcactatccgtt aagtgcaagtactcggaagacgcttcaccaagctccacaacaggcgagatgtactgttca gacagttgcagcggaaatgccttgagggaaacagaagggaagataccagggtcaaaagaggg	94

图 4A-2

GAMOLA ORF	核苷酸序列/描述	SEQ ID NO:
	ctaaaaataaagcagaaagtgctgattgtagatgaatacaagaataacggttttgggtagc tatggaacagtgcggaacgacacagaaagaaacaaattttcacatgagtagacatggctatt caaaaggagatgaaagaataggattgaagagataa	
2006	atgtatttggctaaattttgtcctaactgtgaaataaagtagaagaataatgataaattctgcattttgtgaaataaaactaaagagt tataatttcctgaaaaaaagtgaagaagctcaaatagatttaataatgaaaaaaactagcaaatatgttgaagtcattgatgggctaa tgagatataaagttttccatcaagtttacctgtgaaatatattattataaagtgaattatgaaacaacatctgatgaaata	95
2007	atgatttatgataaagcgacagttacaagcctcatagtgctatcctattgocattgatgagcatgttaggaatcggaagaattgacitca gaactacatattggcaatagtgagcggtagatagcccttggtttggtactacaacgagaagcacaacagcagatttggtagcgga ccaccaagtcgagactcgagctttgctatggggagacgatgaagcacttatatga	96
2008	atgaagcacttatagataaatcccttacaggagactgtttggattacaggttccctgaagacaacagctcctcggcaatgataaac cactggcgtagatattgttcaacagcataaccgagcacccttacttctcaggtggacgagataggaatcgttcttggaaatcgttt ccatcaccatcgctgcatttacatagcgtatgacagatggaaggaaacgaggaaggaggtttgacacg	97
2010	atgatagaatacagcaccataaaagatcacgacataaaagcctgcgaataacaaatccgaggaataatgagccagcttgaacacacaaaagct caggaaactccatggagacattcggagtggtgacccaataataatcaacctgaagaacaaccacatcatagaggagccaccaaaaggtacg aggtccttctggacaagtcctatggaggacaacgagttcataaaggagctccacctcatcaggcttggagatgtaggc	98
2012	atggacgtcgacaacctgctgggatatgacaaggtactggaattgaggtctttattggaggaggtcacgacaaggatgatacccgatg gcacaagaaccgggaatcaaggacttcaagcagatgtgccaggactacaatttcgtcagcataaagcggctggagaacgaggaacgtca aggatgaccagttccatccatttcgtcagcagatgcacacaggaacggctgcaggattcacggattgggattgacacggagaagggtctc gacagggttcccttcgacagcgtcgacacagcagcagctggctccagaccatattgtatgcaagg	99
2013	atgattaatgaaaggataaagaccttatttgaatttttagcttcaatgacaagaagtcattttcgtcacattgttctgtgagcaatct gataagcaacctgttagccatcaagggttttcaacttgggattttgggattgacaacccgattgcggaataatctcctgttcccgttaggat accttatggcagacgtgattacagaagctctatggagagaggacagctcgaagggtcatattgcttgggctcttttgca	100
DNA 包装		
2014	atgcccgaaaccttgggaaaggcaaaaggacgagaaacggaaagcttgagccaataaaggcattcgagtacttaccgagtacctgacaat ggacaagccacgaagcatgagggttctgtgcgagaggctcggaagaagaatgggtatattaggcaacttcatgacctactcatccacat ggaactgggtcgaaaggcgagggcatacagtagaacacataatattgaagaaacgggttaaggaaagagagagttttac	101
2016	gtgaaggacattgtcaatcattacggctaccttccaccgtacaagccgacacatcaggtctgtagtaaggtcaaaacaaagttcaagct caatgagccttatcgtggacagatgttgagtgagggtgctggcgggttcaatcatgggatacgggtgcaggtctccttattgtcgatgacc ctatcaagaacgtcgctgaagccgaatccaaaggtagctcaggcaaaagctcaaggactggtgggagggaacccatcaag	102
2017	atgagtgaagaacaagaatgatgagaattgaaagggttaaaacactatgcataatcaaacaggactaataatccctattttttaaataatca tgatttaataaaaaaatagaaaatgaaaaataactaactgacgaaaataagatatattgaggaaaatgaaaagcagataaaaa gaagaagagaaattataagattattattatgataattgcaaatattttaaatacgacagcatatgtttataaaataatccaaagtaaat	103

图 4A-3

GAMOLA ORF	核苷酸序列/描述	SEQ ID NO:
	aatttt	
2018	ttgagtaagaatgctaaagcagacgacccctttgtgtgaccactgaagcggatcctatgacattgtcgatgcagatgtcttgagagaggtatgcaataaagagcgaatctgacgagacaggaagcaagcagttgaagactgatggttggaatatgatgatacatgtcttgagccgttgtatgacccattgcagttgtgtgagctattggagataaacacatatcatgagaactgtgtcgatgtgtgtgcaaggac	104
	噬菌体衣壳模块	
2019	atgatggcgtcaaggtgcgtcattcaaaaggcagattcagaatatcaagagagaatacagaaggcgtttgatatactcgaagagcagtcagtagggagattgcggacttcttcagaaggcttgagaggaaaaatccacaaggctcatggatgagcactgggaaagcgaattgggtcttttccatctgaacaaaggcttcagacatcatacaggacagcaggcaggagtactatgacatcctgttccaagtactgc	105
2021	atgataattgaaattcctggaaatgacaggacagaaataaaatttgcgaacggacagtttgcactgataaacctacttgcaggaaaaatgacatgatgtcattgcctgacgggaagtgtgtgcccattcagaatgatcagcttttccactgaagaggcgggtgaactgattggagagtgcattgaggaaaatcctcactacaacactcgttctacaatactgtcttgcagcatttggatgaaggaaattgag	106
2022	atgttaagagtcgtagagagaacctatatacaacaagaagaaatcaagacactagactgtcgaattagagaggcaggagtcacacacttactcccttgcctcgaaggagcagttgactatcctactcgaatcctcaacgaggtgagagggaaagataaaagaggctaaggaaaataacggtgaatcctactactatcactggagagacgttggagacctacttacttaggaagattcttagtgat	107
2023	atggacgcgataaacgttatcaacccaaacaagataattagtagatgtttgtaccgtggactgtaaacctgattgacatcgctcatcgttgatgcattggtctatgacaaatccgacagttgtagtcgaagtgctacacacagagatttgcttttgcaaccacagagattgatgagattgtattagagaacgaagaatttgattgatttaataaacttacgaggacggaggttctcaattttgttgaagtcccaacaat	108
2024	atggttttgcgtcgatttgcgagaaacgaataagcagacatggtcgaagcttctatcgcaatggagacggtagcagcaggatatggaccaacagacagcttgagagatagtgagcactactgttcccttgaagtacggttagaatcaccactaatctccgagatttccctattatatagtgtgtcttatatgacacccattggagaggtgtcgtatgacggagtcacaccttgaggaggttccatcatagaa	109
2025	gtggtgaattggtgaaggctattggagataaacttctcagtggtattatctccttttagcattattctctagtggagatttgatatgtgttgccatcgttctcaattcttatggtgtcatctcaacctgagaatgtaaagggaactggtgattgactataatagttacaggaaagtgtgatatgttttggagacatctaaggagacccatagatttgaagattatgtcttagacaattttgaaagagatggaaact	110
2026	atgacaaatgaagttacttttgaagatgagaacggtgagagaacagtcgagttcggcgatgacgtcgatttctgttttaatcgaaagcgaatgatgacggaaacattgagattcgcgaaggagactgggaattggacggagattccgacgatgattgggaagagtagcagtgaggacgaggaggattctga	111
	gtggacgaattcgttgagaccctgttcgacacactactggaagggtgaatgagaacggagagtagacatgagccttaccgattgcggtgatttctacatcgcaagggtgtctccctgtgttcgcaactggagcatagtgatagaatgcaattgcttctgttttcatgcaaggaaattgtctacccatgagaacggagcaatacttgagataggtatggagattagctctcttctacctaagccaaatggagataaaaagatttgaagatatacatgatagattgt	112
2028	atgtttttggaaggaaatattgaacaagacggagagaaatgaagagaacaggtgatttactggagaaaagtccaatcaggtacaggtacacaactgttcgtagtcacgcctcaggacggagagggacaactgccaaaccacactcagtaagcagagaacagtttagaggaaactaagagacctatgca	113

图 4A-4

GAMOLA ORF	核苷酸序列/描述	SEQ ID NO:
	gagcagtgctcgagacaatgacaaggcagaggaacttctccacaagacactgggttcttcttcttggcgcatagac	
2029	ttgcagttgggtgtagaaggtgagataatggaatgcccttgcgataatgtgagatgttaagagagaacgaacacaaatcaaggtgattga ctggaagtccaattctccatttggaaatgggatttktattcacttttggagggaacggcttcaggatgttggctataatgggtga	114
2031	atggaactgatgactagggaatggaaggaaactcaaatccttcccttcttctattcccaagacggaaaaagagatgacgcgatatgggt gatgaagtcttcaacccttacggattgggaacttggatgtattggaggcagagaacacaggaagatgggtattaccttcttcttcggat atgttgatcaccgataaacacctgaattcaacgagtcgggttacttctcatgtctgagttggagaaccttaagata	115
2032	atgttcggacagagaagaatgttcaaaaagatgtatcatgtcggagacgttgcgaactgggttcataatggacgatgcacaggtccc acctagtgaaccagagagagaataactcttcgtagacgataattggtcaaatccatgtgagatgggagaacggctcagggtctggcactaa tatatgggtgaggacaggttcaaaagtgtgagagaggaaggaataag	116
2033	atggatttgggttatctttatgatgacctgactgcaaggagagagaagtcctacgattcagtcggattgagtttgggtcaagtataaaatt ctcatctcgtaaaggacacaggttgcgattgaaatcgggtcgagaactgattgaggaatga	117
2034	atgattgaacggagagattaggaatgaagtacgataatttttactatcttggatgaaatttcaaggaaagttagatgacggtgaattatc cgatgagcaggttgatttccctgttcagatggagattctagtagaggaaggcactatcacagacgagcaggcacaaagatgtcatgaatg gagattattaa	118
2035	atgatgaaattaaagttgaaggaaactcgggtgaggaaattgagatgatacttgcagaaggtggattaaaccttgcagtcgattatct gttgacttggagacctgcattcgacgcgggaagcataaacagaggagcagaagagagaaattatctgtagggaacttctaa	119
2036	ttggtgagagaacacaggagagattgattatgtatttggaaattgacgaggtcaagtcgaaaaacattgcagaaattgagttgacgatgt cgcaatggaaatcgtcctgacagacgaaaaaggtgtacgagagataaagagatggctcaagtcgaatgagattgactacgattgcaggg aagaccggtatttccgccaatctcatagaatatgtcattaggataacttgggtgtaa	120
2037	gtggtaacctgtagagatgacgattcaccattgaggaaacgggttcttcttactgctcaggaattggagatttccgaaagtgaggagga acacgagagcgtcattgacagagcctcgcattgctgtacacgaagctcagaaccagagacttcgagttcaccgacgaagaacgggaac tcttggaggatgcttttgcattgtcagcgaccaatag	121
2038	atggctaaagaaaaatgttatagattataagattgagcgacaaaaatgataacacttgaggttatttatatgtaactgaaagaggagagg aaatatcattgcttcatcttcttggagaattgaggcagaaaagttttaaagaaggttgccttggaatgatataatctaactattcacta aaaagagttctgattctaatgttctgcgcgaatctgtaaaagaatatgtgatgaatctgttttgggtgatgtcgct	122
2039	atggtagtcaaaatgtcctaattgttttagtccacgtgttcaaaaatgtgaagataactaataattaaatggcaatgtgataaatgcaaatg caaatttaaccatgggtgcttctgatatataatgctgaaatggagaaaagtgaacaattaaacaatagaaaaaaattgaaagagaacgagaaa gaactgaacaaatttgaaagagctattaaagaagcaaaaggaacaatttgaagagagaagaactgaacaaatttgaaaga	123
2040	ttgaccgagggtcagataaggcagattgcacatgagtatctggcaaatcacgcttgcgacaagaacacatgagttcttcgagaccag agaagtataagaggttcctgtgggtcattataaacaacgagcctataagcctcaagggttggacggaacagtcgaatgagtatccaa aaggaaacctggatagccaccacaagataactgatgaagaggagatggaagggcactcaatggagagtcactgga	124
2041	atgggaaacgaagctactttaaaccaattggtaaacgagcaggaaggaagcagcttcaagtcctatgagactgacatggagacagga ggctgtattgaacgctcagcagcgtaggttacttctttagagaagcaacattagacaatacaattttaagagatgcggacttcaagctga	125

图 4A-5

GAMOLA ORF	核苷酸序列/描述	SEQ ID NO:
	tgaatcattcaagaagcatctcaacagggtagaataaaacggaagggtcctcaaaaaggatgacgtgaacggcgagaccgaccct gagattcgtcggtgacgtcgacttcgagcaaacaggtggacgtcaagaagctcaaggca	
2042	atgcygaggggttgcctgaactctccagagcaataatggatgatttcgcataatcttgcctctcatttgcctttataggattggttga ggcgatactgattgggttggattgggagggattggcaaatatccgttcgaagctctctagaaaagctttataatggttttgaaggatg agttagggtcttctgaatggaacgaactgtctgtgattgaaggagaaagcatgaagaaaaggtatgctgtcataaga	126
2043	atgacttggattggtacggaagatgtcatcgagttcacaggagtcaggctcagacattcaggttcgagaaggagagacacctccagcct ggaaacattgcttgagaagtggattctcgaggcagaaggactataaatcctactgcaactacgatttcaatgacttggaggagatac ctccagcagttgttaattgtctgcttaggctcactgcaaacatggtcgcattggcagcaggaaggacactcct	127
2044	ttggtgaaattgcagattgacgttgaggaaactcaagccattggagccaaaggttccaaaaggacagttgtgtgactgc aaatgaattgcagagaaacctcaagaagttgagtcctgtggatcatggaaggtctcagggtcttgggtaatcttccagacaggagaat tggaaggactgtgaaaagcagtgcaaaagtatgccatttctgtgaatgacggaacaggaactgtacggtcctttgggtcat	128
2045	atgagattcgtaaatactgcttgccttgcctcagactgtcaaggcatatcttgaaaaggaaatctgcgaaggagggttgccttgagga tgtagagacactcatctcgtccgtgaacagcgagcttccgttgaccacactcgatagatagtcagcaccacactaccagatggt ctggcagtcagccaaatctctcaacaagatagctatgtcagtccttccagttcgtatcggtggaatacagcgatgacttggaaagag gctgagatatgggaataagcctagccagcaggttggctcaagcctgatgaagaaacttcaacaag	129
2046	atgggaattcgtgtttaggaatgaaggaaaggaggaaggtatggagttgcggagtcagcggacttccatcaggaagttagcaaggc aaaggcttcttgactccactccgaacacaaagtcagggtcgaagtgaaagaaggcagcgtgcaggtgtacaagcctactg ccaacatcgaaaggtgaagttgacttgaagaggtataggacattatctcaaggcttctcctggacaattaccattttact	130
2047	atggttgtagttaagaaatccgatatattttaagggttaaaaaagattgaaaaggtgaagattgaggtttggacgggagacgagatgta cttgagaccgttgcctccaaagccgagatcaacgaggtcgacgagattgaggcaaaaggtatggaaatcttccgagacccaacgagaccgcac acaggggaagaaagcagaaagcctaagagtggttgagagcaaaaggaaagataaactcgaattgcagcagaaggca	131
2049	ttgccgtcaagcaatgtaataacataatagtcaggcagagagacatggcatcatctgcgcccacaaaggttgaacacagcttcaggaa attgggaaacacataagacagcaccttcaccacatctcttcaaacaccaagttcaatcaggaaactcacttcttttggacagacttgg acaaggtaacccaaaggctgaagcaggttaggtgaacggtcagtcacagcttcaacagcttacaatgctgaaagg	132
2052	atggcaagttttacaaaatatccgagcaacgtttcacagaccactggaggaaggttcgttccctcagcaatctggcaaacataaagaa caatgctgacggagcgcatgccgtgagcagtggttcttatcaaaagcaagaagcagcttccaaacagacccgtccacagctctcatgcaaaag gcttcggattcagccttccgtgaaggtgctgaacctactaaaaatcacagtaacttataggcataggaagaatgcttggaaagcgactacagc tcaaaagaacaaagactcacatatgcaacatcggagggacctacaatc	133
2053	atgggcatagctattgttggatggacaatgaggagaacttccctgcaattccttgaccctgatttatgcaccatcaatgagaccataga ggaaattgggcttgaggacttggagttcaactacaagtttcaggactatgttgaggacagggatcttcttcaggataggttaacaagatctt ggattagcaattccagagcttggaggactgcttattgtataaacactcctgtggagaacagcgtctctatcaggag	134
	裂解	

图 4A-6

GAMOLA ORF	核苷酸序列/描述	SEQ ID NO:
2055	atggtcgagaagattacagtgagtcctcaggaagtcagaggtacggaatgtgttgatgagaaggaattggaagattacggaagcta caggtgtgatgtgagcgagagctcgaggtgatcaaggaggttgaggaaggtatcagcggttctggtgttcctgctcctgctga gcattgccaatgttactgaggaactcgagggcgaggtgcccacatctcgcctcatcttgaggatggggaaggg	135
2058	atggttagattcagcagagacatgctccagggcgaggaagatgttcaagtggctaaagaaagggcgaaggttgcttaactactt gataatgtatgacatggacaggaataaggagataaagtgggtccaaaggaataatgcaggaactgtatgagtcagaaacataattctgga ttaagaacggaaggagcctaactatgttacactgactccgttgcaaggaaatcctcttgatgactaccagaaacacccaattacaccc tgttgcccaaccagttgtcccttgccctcacaaatgctatatcactataagtcagaaagtgaatgcgctaaggcctttgggaactagcaa gggcagtggaacaagccctgcacag	136
	溶原转变	
2061	atgaaaaaacattgcttttatttttaggggacagtttgcgatatatgcaatgaagcc atggtttgtgaaaaacatttagtagaaggaactatttgggtatcaattatccgtgcagga aaggcttcagagataaactgtaaacatttgtgaacttggaaggtcaagatggcttaata agctctggtcagaagaagattggaaatgcttgatataaagggtatcatttcttatgat atttataaagattaaaccatattcgtaaaattagaaataaggctgccatggcatttaagtgtatgaagacaatgcaaatattct gcatgctatctgtatctaatgtgcgcatatttctataaggaatatagagacactaatcttcagcggaagattatcgggcccctatta tggaacttgccctaaagcctaaagagactgcttcagagacttcagaggaataatgaaataatcggagagtttatttcaagtcattggat gattatcttttgaaaaagtatgatgacagttacctgttaaatgaactgtctaaacttaaggactcttcaaggagactgttggaagacga taacttaagcgaatttaaggaaatcttcatattgacagatctattcaagaagattttttaaaggcattgaacagagccactagtttta attcatctcatctgattatgctttgaggtagtttgagatgggaatcacactgtatgcataatttcaaaatgaaagaaacccctgaactc tttaaatcaatttgctatccattatgatgctacagagagttttgatcctgaaaagaatgcaatagacactttagcctcagttattggaacc tttcaatgacaataatcttaacaattcgacagaaaacttatttagccattaaactgggtgtattgaacaatttccctggaatcatctt atgctaataagagattacactaaactcaagtttaattatagaagagcgaacataatttgaatctaatgaggttcagataacatttatgga gataaggttagttttgtcacttttcagtgattacaacatgtttgaaatgaatgatgataaaatcacaactacacatcatcaaaagtacat ctcttccctatttaataagataaaccataaggagatacaaacctttttatgttgcatatctcaaggataaaggactctcactttatca atcctataatctataattatgagatgctgatggatgagaggtccagaagacaactcattgattatttaatacaagatttttcataaaatac agaaaaatcatttcaactagagacttatgaaactttattatgagataatagttcctccggaattcctgaaaagtggagatttggacaa tatcaatgatttcatggactattcattgcctaatttactctcgatataccagaaaggtcagatttatgaaagctatgcaatgaattgg atccgacattgcatcgtaatgaatcttttagataagttcataatagacttgaacattaatgacgatactgaaaagataattaaatcggttat tttgatttcactcgtttaatttccctgaagagatggagagatctggttgatttttagggagtttaacaattctgaaaaaggaaggtt caccataattctaaatagggttgagattctatggaagagcattcaagaataatttcaaggataaaggtttatttaaatattttaa aatatctgtatgcctataatacacagctctcataaaggattacaaaatactgttcactgaagtttaaggagcgtatttttaattggaaaggt tcctataagaagaataactatatgtattgacactttggatttcatttaaggtgataaaaaatttaaacatttaaacctctgtgtgataaatt	137

图 4A-7

GAMOLA ORF	核苷酸序列/描述	SEQ ID NO:
	tgaataaagttttattagatggtctcttttaggaaatagatttaaaacagataataaaattttttcagttgaatctaaacaaaaga aaataaccttttaaatgttgatttttcattatatacaatacaataatgaaattatataatggttttaagcctaataatcagataaagacgat ttaataaatttttagatgaatttattataaattatttagatgaagatacagatgatgatttatatgtaattagtttagaacttatgaaga atttttatttgagtcgaatgattttggcacttttgaaatttaagaggggttaa	
2062	atggatttttcagaaaaattataatattattataaaacaaatgacttgtgatgataataaa aggaaatttaatacatcagatttaataaaattctccattgttaccttttaaaactaataact cctaaaaaagctaattttgaaaatggttttgatataatattgggtgaattatccagaatc ttattaaataaaactattgaaaaaaattttaaattagataataattgtttcaaatttaatt gataataatattgaaattgagatggaactaaggagatatacaaaaaattgctaaatgaa tatttattcgatgaaaaaaatgatttaaaaaatctcatccgaatttgtatttatatt cctctttcaaaataaagttcttaattggagaacaggaagtagcattgtttttaagagat attttttgtgaagaataatcaaaatcttataaacttttttgaaatcattatgattcaaatcac attatcttaaatttaattttaaaaaaactcctaatttaccatcataaaataactgaaact aagtatgtaattcattttgaggaaattgcaaatttatttaataagaatatcaattatgca atactttataagaattttttatggaacacattggttaatttttgcattactatttc ttttatactctcaatttaattcaaaaatttctaaggctttaatgataataaataatgaaaaattgtactatttattggactggga atctgcaagtaaaaatcgtaaatcattaaatagttatagtttactaaaaacatcattcaaacattatatgcaaaaatggcagtcataag atcaaataaatacacatattagtagcaaaataaatttattggaataagatatattcagaatattttaataatttgacatacaattctaaaaat aattttttacattttcttaaaaaatgggtttcagattataggtatgtaaggaaattttgagataaagaattaccagataaattttattaga attaaactgaaattctttttgaaagttaaaaaatgaaaaactgggttggtgagctgacaaatcagagataatgttgcttttaactggaag atatagctaaaaaataattgttaagcgtgagaggttcgtatggttattgatttaaaatattatagagataatgttgcttttaactgca ttatggtttaagataaagaaattaaagttaaatcaactatttattgaatttgagaaaaagggagtttattttgacaaaatattcaaaaga agaagtagtaaaattttttaactaaattaaatttaattgacaaagaaagtgatagtgagatgctcaaatatgttaaacccagtttttatgat atgttaaacaccagtttttatgatattatctataataattactaaattattttgatgataactaagatttttaagtggtgaaaaatttttcat tagttttgatgaagatgaccaaataatgcttttttataatagtttaagaagatttgcaaaaactaaatttttctgttctgaattcatat atgtccatactatttcaggaaaggaataataactatttccataaacattaatggagtcgaattgtcatttctgaaagttaacaatt aacgtagatttcttagttactttaagaaatcaagttactctcaagaggggtgtttgggaagatactgcattattagttatctgtaaatga agcaatagatagatttggaaggtatgagaaatttacagaaaggaatgccattaaatgtaaaatctatttcaaaaaatttagagg atgaaattaacgattctcaaatcttaactatccgataaagcaaatcgtaagttttcttgaatattcaagaagaagaattgtttcaa acaacattatgggattatgaaactatcttataataatcaacaagggtttgttagtgatgaagatttaagagagttgaaatcttttttaa agatgaccaactaaatacaaatctcctcaaaaagatttaaaaaagatttaaaaaaattatgacacatttaataatgaagtataaaatttt cacaatatggagataaaaaaagaaacagcttaaaaaatgttttactgattctggagtttctattttaagtaaaagatgactggtataaagca gaatggaaaaatggttaaaaaatcaaggatgactttatttaatacaaaaaatcctttaaatcaaatgaaatcttgaaaaaatcacaga	138
2063		139

图 4A-8

GAMOLA ORF	核苷酸序列/描述	SEQ ID NO:
	aaatggcctaaattattgggaatttcctaactcattactaaaaactggaaaaagaaacgaaatatcatctgtttttaatccaaatcatt ctcatgaagttagtttaaagttcagtttgatcagattcttagtaattcattccttaatacaaaattcaaaaaattcactattgcaaga ggcaaatcattaatagtagcaatttactctggattctagtgaaaccaattttcaaaactataaaatataaaacataaaaaatgaaaatatctc tgaatttacttttaataattgtgtgtttggaattttgagcctgaaaatttttaattctataaaatctcgttttagtgtcaatgtcaaatcta aacagattattgtaacaaatgatgaggatagtttttgataatttctgaacaattctctgcttggaatgatggaataataataaagaaaat ggtgaaaaagtatatctttatgatgatgaatcttttaataatttctgaacaattctctgcttggaatgatggaataataataaagaaaat atataaagacaataattatgcccatttttgattaaagaaaaactaagaaaaacactgccagtaaaattcttatgtaaatttgaatttga aaagaagaaaatatgaaaaactttatttttaattgagtaaaaagctgttcaagatgtaaacagtttctatctcgtagaagaatttaaagaa tttctaaaaatggaacgagaaaatcattaaacaggacatatatttatgctaaaaaggaatttgatggttctttagaaaaaaattggaagtttc attcagtaacgagtttagaaaactgcttacatggacataatttaatttatataagacttttgatgattctccagaagaacaattctccaagtt tagtttatttaaatgatgatttaaaagagtatatataaagaaattataactatcttcaataaggaataatcagaataatgaagagaattct atttgtctgattttaatataaaaagaatttgcttaagttagtgcgtattgaaactgatacaaaaaataatgtattctccattatctcc actaaatatcgcttatcaattagaagtttctaaacaatcggaataaggatttatccgtaaaatatattagaacggttagttccaata acttaattccttatattgttcagatgatggaacaaaatgtatttatttctaattgtgttactgaaaaacttaatacatttgcacaacatttta aaaagtgaagaagtatcaattggaacaaactcaccattaaaaattaaacttaatacaataaaagatgataaagaagttgtgaaaggagattta ttacttatttagtttaacaactcctcacaataaaactaaaaattaaacttaatacaataaaagatgataaagaagttgtgaaaggagattta attttattagatcagattgcctgataaaactaaaaacaaaaaagttattcctgttgaaatcaatatttacaacgacgctgaaaaagagt tcttttgacaatttatttgattgtcagctgaaattcagttactggaagaatttggaattaaaaaattaaaaatcagatatattttgatcc tatagatattatccatgatgtcaaaaataattttcattatataaaacatccatttaaaaaagaagaatatgaatatgctcatctttctt tctataaagtcacaatacaataaacaatatgtcaaatgacaatatggatgaaactggaattatcttttaaacggtttatatcttca gttacatcaactactaaacttcagaatatagaaacggtttcgggtacaaaacaatttttaaatatgagtaattcctcttataaaaaacagt gattaactgaaacgaactgttgagaatagcaagaactttggaaaaaaactatttcaaaaaataaaatcagttatcactactgttgaat tagaagaagacaatatgagaattgtacgacaaatctcatgggttaacatttattgaaacctacatttgggtatcgaatactttgacagt tctgtagtaaatgtgatcattattcattacagcgatcaatatagtttctctagcaagatgatacattactgaactaataaatacaac acaatatgaggaaattataagagatttctccaactcaaatatgtaaaagttacagatgaagaattgtatgatgtaataaagaagtgtta attcaataatggtgaattggttacttagagtaattttctaattccggccatttatgatagggaaaaaataaagattatttctgctattaaa tattgcttattccatttggaccataaagatatgtttggattccagtttctatggaagaaactttaagaatttcagagaaattgttaaat agataaaaaataaaggaattttgactctaagttataaaaaaggaatcatagtgatgatctgctattcattgggtgaaaaatcaatgaag ataatagaatcgaagttatattctaccaatcgaagttaaaataggtttgaataatgcttcaactataaaaaagggttaaaagtcaatta gataaacttataaaactcttaaaactcaactacaaaatttagtaactaatcatatttggaatgagaatactgaaattagaataaattctt catccaaattattgtccaatgagcaaaaaattagtaactaatcatatttggaatgaaaaaggattggatagaattgaagaattcaag cagaactattaaatgatgaatatgatattttatgtgttagaagagtagacattggtaaaaggttcttgggttcttataaaaaagaatct catcacatctctatatggattgtgataagcaagtaattgagttaccagaagattttgcttattacgggttagccactctataag agaaatccatgatgagattcagtcagacaatacacagatatcttgcagaacaactgttatcacatgtagatataaagtgaatttagagcaa	

图 4A-9

GAMOLA ORF	核苷酸序列/描述	SEQ ID NO:
	agaataatgatataatcggaattcaaatgaggatattgtcaatagatgatgatttttgataaatttaagtgaatttgaggatagttttattgaa gaagaatcagaaaattttctgaagagcctgatgaagaattaaactgatgcaacttcaagcagtgataaatgaatcaattgagaatattggaga atctccttctaaaatttctaatgtcagagcatttaattggaactcagaagggtataatcataaagtttattgggaatttggacatcctt ctttagcaaatagycataatgttaattcaaggaaaactcggacaaggttaaacatatttcatccaacgaattgtaaaagagttgtctatt caaggcattcctagcataattattgattatactgattcgaagccttcaatgaacccaattttaaagactcctttaggagataa aataagccaattattttgttaaaagaaaatttccaataaaatccatttaacgaacactatcatgattgataaggacataattttattg aagaagacaacagcactattgagccgatttaaatctatcatttaattctgtttattggctaggtattcaacagtcataactttatat caaacagttttagattgcttggaataagtgatgataaattttgatttgaataatcctaaagaaataatctaaagacgaatacaaatag tgcacaaacagctttgaataaaattaaatgagttattggataaaaatccgttcgcatccactgatttcgactggtctgtcttagataata aagacggtaaaagtttacattatccaattaaactgcactttcctaaggataacagacaataataactgaattcactcctttgggatttatgg aattacaaaattacaatggtagtgaagataatcctttcattgttttttagatgaagtcataaatcttgatttttccaatgactctcc atgtagtcaaaattcttaaaagaaaggcagaaaatttggtggtctggtggtttgcaacacaaatctgttaaaggatctatgaaaattgatg aaatagctaaaatttagagaatgcagatgagaagataatctccatccacggatgtttcaacaaatagctaaagatttgtcaaaagacaat gaagataaaaaaataatgaaaaagagttatctcaattgacaaaaggatattgtattgttcaagggtctgcaatagattcaagtggcaa tttgtatcagccaaatcctgtaaactgtaaaaattggaggagataaagttttgtatgataaaattaa	140
2064	atggaaagttaagacagaatagagaataactattttcaatgataaaagaatggatgcttat atagacaaaattttgaagagtttactctgagttctgaacaagaagagcactgaatatt tggatagataaattaaataacgatcaattaaacagtgaaaagggaattatcataatttt tttgaaattattcttgagagtttacttggtataaacgttctgtattgaaaagatggaagattat aatattggtgatgaaggccatcctgtagagtttgtattgaaaagatggaagattat gtgattatagaactaaaaggacaactataaagacccctcactaaaagagaagacctggacag caatcaccagtagagcaagctacaaaattatgctagtgtctaaaagaaaactgaatgggct acagtttcaaatatgatgaatttagatttttttaattccaacagcagagagataattatatt tcattcaagtttaggcaattgaaagatttggaatcctttaaanaatttttattagttatc agtaaaatttcacttatcgatgaagacataacctaaaataacttaataatgaaacaaaagtc attgaaagagaattagaaaatgaattttatcaattgtatagtgacactagattaatgatt attaagggaattggaatattctcagaagataattaaaggattgaagcaataaaaattatct caaataatattaaatagatttatttcccttgttttgagaagatttagcgcttatggag gaagagacaactgctgagtattatttaacacccctttaaagcatgaaaatttaactcggaat acaatgtggaataggttaaatgaattatttttttgcaatcaagggaataaaacatagg agaatacctgcatttaattggaggattatttgaggatgatttataatttgaaaaataagg gatgaaattgaagataggctcttcttcttgaaaattggaattttaaagaagatttttgaagac aaatatgaggatattgctaaactaatggagtttataaagatactcttaataccaatttc	140

图 4A-10

GAMOLA ORF	核苷酸序列/描述	SEQ ID NO:
	attaatttatttaataatttctacatatgactttgattcagaacttgatgtaaatatctta ggacatatatttttgaatacagattattagcgatatattgaagagttaaaaatgataatcaagag caaagaaaaaagatggagtgattacactccagaatatattacagattatattttgtaga aatacaattattccatatatttaagtatttctggttaagccagcacagttcatgaattatta tacgaatatgaatcatctaattcattggatgttttagattcaaaattaactaacatcaaa gtcttagatcctgcctggtagtggagcatgttaataaatcagtagatatattattt gaaattcatgaagcattacatgcaagtaaatatgctggagattcatcttttagatagattt tttgatagtttagaaaaacgaaaggaaattataagtaataataatttatgggttgatttg aacgaggagtccttgaaattactaaattatccttattcctaaaattagcaactactgtt ggacttaagaagggtttcaacttcctagtttagataaacatatataaatgtgggattcca ttagtgatgatgagtcattgcccgaataaaagcatttaactgggtatgaatcattttca gaagttttgaagtgccggtttgatattattgtaggaaatcctccgtatgtggatatt aaagaaatggatgaaaaaacagcaaatcacatatattgtaggaaatcctccgtatgtggatatt aggataaatattatctacatttggtaaaaaagttattatttgcataaaaaatgaagg atttttcoatttatcatgcctaatccatttatttaattctacttactctaaaaatcaga gaactaattcttaataatacatccatatataaatattgtagaacttctgatgatgttttt aaagatgctaaagtggaacctattatttttaattctcaaaaaaggatatgatgaaggaaat aagactaagatactataaaaaaagatgatatggatgaaattccaataaataattatttca gaacatttcttacacaagaagaagatggtttgaaaaataattcaataaataatttttttca gatgattttacttttgattttattaaagaaaaattgatggaataaatgaaagatttaattgat tattgtgatttttagtttaggtttaactccttatgataaataataaaggatgagtggagat attattaaaaatagaaaaattccattcaaaaataaaaactgacgatacttttaagaatta ttggatgggtccgatattacaagatataatgtaaaatggcgaaaaaagatatattaaa tatgggactgttaggagcacctcgggaagaaaaattcttaaaaatcctagaaatttta ataagacagattttatctatagctcctaaaggaatctcgaaaaaggatttttgcgcgttac acagaagaagaattatataacgctcaaatgtcttttaatttagtattgaaagaaggcttt gatgataaaaaatttattgaaatatttttaggaataaataattctaaaaatttgattgg tattatgaagaagaatttatggataaaaaataaaagaattttgcaaaaattttgattgaa aatgctaaaaatctccagtcattatttaattcaaattttttagatgagatagtatctaat gtggattcaataatagagftaaataaagagtttttagtgaagaatgcatttcaaaaca tggcttaaaaatagaatttgaaattgaaaaactcttaagaaactagaaaaactattatgat ttaaaactttgaagaattcttaaaagagataaaaaagaaaaagtagtcattagaccaaaat caaatacaggactgtccgaattatttaataagagtttaggaaaaatagaatatctgcaa agagagatttaagaagcagacgacgaaaaaaatttaacctactggtctatgaattatatggttta	

图 4A-11

GAMOLA ORF	核苷酸序列/描述	SEQ ID NO:
2065	aatcatgaagaaatagaaattatagaaaatagctttaatgattaa atggacgatgaaacattaatcctaattgaatacatcagaaacgcgccaacaagagaaaatggtgcttaagtcatttgaaggagtggaactt tataagaccacaatccaaatttctaggaacacaggcattcaccacaatgtcagcaagaaaattaaaggatcttcgcgaacatgaaactgg tctatgtcataaatacctgagtaccatgttcctaactttacaggcttactgaaaagggtaaaaatatgcttcaattctttag	141
	噬菌体整合位点, 基因组整合基因座	
2069	atgaataaaaaattatcttatccctcttttagtattatagtagctatttctgtctctgcagttgcagcagcagatgctgatgtcac atatataaacgatgctgcagatgtagacgatgttgacgacgaaaaagtgtgctcctcttacagctagtgtgacacaagacatccaaa ctaagcttgataatgctaaacctggagacacaaatgaattagaaaaacaagacatatgacgcttgatacaacatttaatt	142

图 4A-12

[illegible]

图 4B-1

GAMOLA ORF	氨基酸序列/描述	SEQ ID NO:
1996	mieciqeeqdfdtwvpssssdikiyivsvddegndlfscpdpfyrkrsmphsnpsyckhirqvleednrlqml	13
1998	mdninkttkslakfeeffstvykdevmevlekyypeertlvvdyenlemfdpdladlliekpdeviaasqkaiknidplmkdpklidkfk nvsncidfvnadskylgklisfeakvmeakepplidlavycrgcmslreipqtinsleplscpecggrsfrllqdeseflesqlli vssddtskslkvllrdecfsdlysmgqevritgilkfsfssnygyeyflecnielindsedseydeygnrnspeyr	14
1999	mankirvnltdpnlwqlakdklpcsrseffengqlkmfigieddeseiikdiqtkenainaialrdklchvrksqklklesnksmekamas lnrmhkygkigenginqirnlahvhkvdfddllkceqdcnmcniifefaevpkhdsvm	15
2000	mtigslndfsgsdgenmpedfcdsvffemyskalfeildvevgsfaelldvyknvemdylkrhalkqkeilywfntdwkeelgkekpt ekdkekwiirgkigydsfvveglevklkhiirmyetalkhsfeaik	16
2001	mnvktvmdnliglskefegveyeiesknsiifysfpkymkegivilkysaiydlhtilkgmdgiivdilevednpgdekdkllyvqiev ke	17
2002	misdeweeeyykvneeleqvevrfsskvdrlvfaqsysssfsfeeaeilcdricdiltndlgnakyylggkdel	18
2003	mnyrewvasqfelrmtedgrvcfktpcsyytfskedfeilremfnfes	19
2004	mdrikelnvcgtckhshlipdingeiaavnicrigseavnkdgltycvdtwtprrkls	20
2005	mlskkeaiqmtldnekhypvkckycgkpfckshnrqmycdscrnalreqkaryqakrrlkikqkvlivdeykyglgsygttsanghr knfsheymaiaqkemkriglr	21
2006	mylakfcpncgnkveendkfclycgnklrvliipekkvkrssnsindektskyvevidgmlrykyvfpslplvkiykvnygttsdeikn ilengnyynikinihyflqnkklyfrspknpnmfkkfhndrfneemrkdnkeiishissyarvhrpsltkifnfnqektfillankhfeeni rktlrl	22
2007	miydkatvtsliavaiilplmsmlgigeltqnyilaiivsgmialvwyynekhnssdlvsgttkdcocelcyggddeal	23
2008	mkhlyeiiipyrtrwtitgflktvtssamittgvvilfnstehpyfnewdeigivlgivsitiacliyamidrwkerrkkeeldtiety inrkaeeianmkvlrkleee	24
2010	mieistikitdikpaeynprimsglehtklrnsmetfgvvdpiilinknnhilgghqryevlldksmmednefikelhirlrgdvvgwafp esdlevedddhekalnlalnniegewdlpklepiltldkdvgfdieltgfsdieltelnennlvfaeeefepdeseevdlediydepv kemlqcpacdhvdvkrfrvdsqg	25
2012	mdvdnllgydkvlelrsllleevtdk vipvwhknrgikdkqmcqdynfvsi sgwrnedvkdqfihfvrhahngcrihglgltrrkv ldrvpfdsvdssswlqtilyarlgkqlskfaterrgdlavlsyikwmktqeeyykwrhyhd	26
2013	minerirpynlfnfndkkvifvtlffvsnlislalavkfnlgfwtcdgnllfplgylmadvitevygertarrvillglfanill ivattltvmpypsywtggayaymfgftprivlagfiaylvvgfyfynarlmvlikkwtnskylfmrtigtstlggelcdscicssiaayyg ivpnsgillfilmgyvkvwtwvwmqpltyksiaawakdg	27
	DNA包装	
2014	mpepwerqrdengklepikafeyfteyltmdkprsmrvlcerlgkkgdvirqlhaysstwnwveraeydehiilkkrlrkekfydelv	28

图 4B-2

GAMOLA ORF	氨基酸序列/描述	SEQ ID NO:
	eselpnlrkrleyynknmdietdmttkptskahaydnkskahsttlnellmigkpteiketsleadiesdnkidlenrvevditsde fmeselzymrkmiiek	
2016	mkdivnhyygylspykptirsdskaknkfklnepyrqgmlsagagsgimgyagllivddpiknvaeeaskvrqaklkdwgggtiksrvg rrsnqlppikiviahqlhldhgiiketeptipandafrilrnngsidpntwdfnlpaicdseddilgrkigevlweeqrdyewlma ekrsngsylfnslyqgqpvdergeifkrewfqdevnhkltclidpkdipqlrlywdfgasgdagdttsailtsy	29
2017	MSKKQEmmrierlkhyayqtgllipifknhdlikkiengkitntdeikiyeenekqikrrrefisiydnckyfkydsicykkliskvn nfeiksleelmneiesekkngrfnieekelikenenkgrlkeyvriirreyglftsrvklklkidnnkirskeelneeieekkkk elrriyvdsdlndlwleleskirhneittkeelikerrriiefiniidydnrdfkldyvtikqlvskigsneittkeelikem	30
2018	msknakadafvttedgsydivdadvleryaiksesdetgsklktgdweyddtllleplydplqlcelleintyhenecvdivdarsagi gydiivptgekekelnkpkltnflenielpninellyqmnydratgygalelirkdkskepnvlshissytlrtsdgrvkrvgrvgtk tvwfiygnknydkegnlcvhsetgefhpynslskeeranellwtmeyttksyvgplkiygaipaiysdisrskyn	31
噬菌体衣壳模块		
2019	mmalkvrhskrqiqnikreyrrrllieeqcsreiadffrrlerkihkvmdewheselglfhlknkvsdiidqsrqeyydilfkycdkdSYM kgreateerrfnrkienismkadvnitrledlfpdpptirynlnnkvfqasahtmdrvdnrimenitqsyddglgideakdrltveyngl ksweaqriarteinsagndgafdygelgveyhqwtqadervretpqadhrhelhgkivkvngsfnglqypgdr	32
2021	miieipgndrteeinlpgnqfvilitylqendmmslpdgkyvcpfmiqlfteeggeliceleenphyntrfyntvfehiddegieyr	33
2022	mlrvvertyyqqeiktldcrireagvntyslarqgavdyptyesynevreerikeakekyesyhyhwrdevetyfyylgffsfdeei ekylertvtykphreelkeamerldkrfeevinefwysleeyeditdvleqlkegdctvhceflikefkfvnivicriikqneinhke f	34
2023	mdainvinqnkilvdvlyrgtvnldivigdalvydnptvvvkcyttldlafatteideivleneefellityedgefsillkshnlgel gyiewvl	35
2024	mvfvdliceneiadmvesfyrngdgsrilltnravgeivehycsfevdgrittnlrdfllysvvlydtigeavddgvnleevfiedrnc ytgksqvili99	36
2025	mvnwlkaigdnfsvdylllalfssgdililvaivlnsygvvispenrelvidyisyrvkvdifwrhlrrprmsfedyvldnfeemetgelt reqvvefvsrqrkgltfcneifiaivplkkgskddivelwneyfedykenwleqhenlgwndkklkkekivengggddfgifrnhl dcvlmey	37
2026	mtmkvtfedengertvefgddvdfvliessddgnieiregdweidgdsdddweeyddwdeef	38
	mdefvetlfdtywkvnnengeymstldeqdfyiakvapcvrnwsiwiecncfcfhckefvyhengailleigmeisslylsqmeikdkli midc	39
2028	mflegifeqdgenvreqviywrksnqnhwfvvnaqgdgedncqphsvsreqleelrdlcraavlndkkaeellptrpgfffgaidydew yyydlqytkidevlkddrylyfeycsww	40
2029	mqlvvegenmecpcdncemlrenepikvidwksnspfgngixihfwrerlqdvgyng	41

图 4B-3

GAMOLA ORF	氨基酸序列/描述	SEQ ID NO:
2031	melmtremegklksfpysqdgkgddaiavmkffnpyglgtwyvleakqgngdylffgyvespitpfeineygyfslselenlkipiki ngltvsygrierdlyfervrigdiagn	42
2032	mfgqkfevkmyhvgdvvelvmdaqaappsgtrgeilfvddigqihvrwengsglaliygedrfkvverkge	43
2033	mdllylddltarrevydvsglsfvkykfssrkeaqdfalkygaelliee	44
2034	mierrrlgmkydiftildeisrklddgeldeqvdflqlqmeilveegtidegaqdvmmgdy	45
2035	mmklsikelgeeiemilaeggltdqvdyllyletciadgsiteeqkrelicrdf	46
2036	mvregerlimyleidevkcenidrieddvameivltdekvyerikrwlksneidydciredryfanlieyviritww	47
2037	mtvrvdvgftieerffltaqeleyseveehesvidraiallytklrrdfeftdeerelledafvivsdq	48
2038	makenvidykierqndntwsylyvtergrgniaassfgelrqvlkrglpwndisnlfkksdsdnvrrresvkniddesvladvakkss dsnvleavrkisdnyvlidivknasdydvrrreavrkindsvlediaknnndenvrleavrnindesvlenisknasdskvrieaikk indetiiiklaknndedvrieavrkinkdvtviidfaknasdskvrrreavrkindsavlayvlnkdpwlvrieavr	49
2039	mvkpcncfsprvskcedtnikwqcdkckfngahgafinaemekveqitiekierertereqferaikeakeqfererteqfererke rlerekrekeriekiererkerlernrikieerereriknerirrenernriksdkreerekseerikrnerirkanernsikrekr ererernvmtideyyrsigygstgskvwsaiilpillviciililmfygggm	50
2040	mteggirqiaheylanyslvdknheffetrevigvpvesyitnepislgldgtvneypkgtwiatttritdeemekalngeygtysit tvskkfadqigqlprvlmkdikpvgftislvrkpcvrgakfcmkedengdvsvseniddkleetkgfvgsikgfnkeddedkn pededielidikaivdevtkdfvntddfetfknelekalsdkfetlgaelfkslkslekdkaeakksgr	51
2041	mgneatlnqlvneqekavfksmrtdmetgkavlneqglgyflreatldntilrdadfkmlksfkklhlnrvqingrvltnyngvndgetdp eipaaadvdfganeldvkklkamceieddedekedmtqaqfeqlqmmgerigedleywalfadsevarsddplntndgwllkkanhis srapiapsngmfiedgpeamfdami kalpprfrknrmkfyvvpfevedayrnilnrgtglgdsaqigfnalsykgipiehcstldde dgrgmngnrvcsn	52
2042	mrrrclnspehngmishiiivllcfiglveailmalvdwedlaisvrksprklynvlkdelgelpewnelsvlierrsmkkrayavirdsfp elppweelsvidrrshkrlykklksvydgyddpslegppaaavgpqkeipleaeayp	53
2043	mtwigtedviefgvkpgqtrfekgdtssletllekwilqaegliisyncnydfndleeippavvnvclrltanmvalaqarkdtpviqv kewnvgtvssnifsnldlkrdltpfvherksykddeidffvitgddds	54
2044	mVklqidveelkpleprfkkvaktvvtanelqnlkklspvdhrglqgswifqtelertvksakyaifvndgtglygplghkir pkngkflaftpnkgkfgklvvvptwrgqkqprfversmenterrvqefmiramemnds	55
2045	mrfvntaslvpqtvkaylereiceggllledvetlipsvnsdvpdpapaiwlvqpttrwsgsqpnlsnkiamsvpfefvcveysddlee aeilgislasrvsgslmknfnkvkvdssmpnrfhkletletlypvgevtvvgkseripatsiifnfyvvdwlkcnrryd	56
2046	mgirvvgmkeearygvaesapdfhqevskakasinstpntkssgrsmkkkaragykptaniegevdllkrihyllkafldnyhftdgggs nnpnvhefwggennklssftlwvtfdi fektivgslldnfkmevsdeymkftadvkykteesdeieneliykvklldgdwalmfydvsvve idenappgi vssfsfdgkninnvdkti glsgsrgpqrkaaaqgrdisisfvstleretleliqkaeygevgtpepseck	57

图 4B-4

GAMOLA ORF	氨基酸序列/描述	SEQ ID NO:
2047	mvvkkssdilkvkkiekviealdgdemylrplsqaieinevdeieakamgiffetnetahrgrrrqkpkvsvveskginlelqqkaqhqa ktaiflsldneknvgeeaeseteiegmphklfeelfnhvkrslgsiieldedddvdfh	58
2049	mpssnmvniivkaedmassvaqkvensfrklngtldstfttslntkfnqelstfgtdldkvtqrlkqvngvngssfnqltnaerrtle klsefdpvsavqlqhlslrigitgqqtfnqisvseqslmmmkstaqqleevnslrvvgigaqaanmlnqmkldpsvgsnldraklkv semgysldstkgkilvlgtaigtsglnkwdsktkvqttatnirtalgnaltsvkskvqnlgnafsglgiissaig	59
2052	masvtkypsnvsgttgkfvsfnlani knnadgahavssvlikskkspnprstvsckgfgslpegaeptkitvtyrhrknagsdys sknktchicniggtisllqvsgfssksgccttmthtkafsvhgklraqvnsanfgvklodyptnsntnyngymrisvrvrtveyitsq ysvsvkhvsgtyededyvvsllgisnknltsynptctltvpagftykgvtgaatgtvtkvnnrtfswnpqlqgragrqislafepnvtf pegttsy	60
2053	mgiaivvmadneenflqldpdlctinetieelqlrtlefnykfdyvedrdlfrignkiwisnsqslcdcliyintpvensvyqenyfa ceieevlaelyyaplfstqltsangftlrrtngqetvvdwnalnywfglyfnigvqecclgtyanritvngtmnrlnlrsieeqtg nrfvtryekdlldntihryldflnpvnvskwnkieydfiyeddgeyceaytsdgnpiseiyddieeeddivdfpp	61
裂解		
2055	mvekitvspqevrgygnvdekeledygysrcdvsessevikgveerifsvsgvpapalsianvtedtrrrgrcahisasfedgedgdld dkaislksgdvlatittgsgenvfdvlydsaqlyavfdgdyppavseaitvnpakslwdvefildeeevsgdtailsgtvgtiiv deivdgetvrrqmeanvtltvtdlgrircstnangefvlqvpniqngwrvviaatsthlvfnnglndvphdysl	62
2058	mvrfsrdmlqdgaakrmfkwlrkgeglpnylimydmrnrkeyklvpkeyaglyesnrifwikngrepnyvtltsvarnplvmidyqntny ccptslslasqmlhykssesecakalgtsgsgtspaqlianapklgfkiiipikrdskevkkylkkgfpvichwqvnqsrnckgdytgn fghygliwdmtsthyvvdapkgvnrkykfsclndnankgyrqnyvvcpa	63
溶原转变		
2061	mkkhcfyflqdsfadi cneamfcekhlvegnylsdiiragkaseiitvnicelegqdgglissgqkkrlemlykgliisdyiykrlnhir kirnkavhghlsdiednanilhaylylicayfykeyrdtnfsaedytgpmidiaspkpetasetnednigefissplddylfkydd syllnelssklkdskeavednlsefkeylhidsiqedflkalnratsfnashlimlsgsvdgkshlianlkknpelfnqfaihyd atesfdpeknaidt lasvlepfnndnlnnsteklilainlgvlhnnflessyanedytklklieeanifesnevsnidygdksvftfs dynmfelnddensnytskysisslfnkitqkednfpvyaylkkdshfinpiinyemlmdeevqktildylifikyrkiistrdl lnfiyeiivppeflksedlnndfmdyslpnlalfgyppersdlklcnelptlhnrnesdkfiidlninddtekilnryfdtfrfnfl eeygeylvdfrefnnsekekvtnilirfavfygksii knnfkdvyl nylkylayntqshkdykylftevkdai fnwksyknntici dtlds fkvynklklkpsvdkfekslldglfngfrktdikiyfsvesnkkkiplnvdfslsyqymklyngfkpnqskddliildefin nllddetddilyvisletyeeffesndfgtfeckrg	64
2062	mdfsennyillkqmtcdvnrklqhinqnspplpfknttpkkanfengfdilgelsrillnktieknfkldnivsnlndniedg tkeyitklilneylfdekndlkishpnlylyiplsnkssngeqevalfrdificnnqnlinfesydnsnhiilnlilknptnlhkhkit	65

图 4B-5

GAMOLA ORF	氨基酸序列/描述	SEQ ID NO:
	etkyviahfeeanlfnedinyailykffmenignifayyyffiyqiliskiskgndnnefeklyyllldwesasknrkslnsyllkh hskplyakmavidqintllgttnlllekdiseyfnldinsknflhflkwsdyrvrnfddkelpdnllleilteifelskneklgvd gavqsyalnleadiakkyllkrrgsygyvlninrdmllvltalcvkdkklklnglfiiefekrgvyfdkyskeevnfltklnldkksd sgdaqyvkpvl	
2063	mlnqfydylsnkllnyfddtkilsgkffisfdeeddqimsfynslrsiaetnscsefiyvhvhtisgkeyntysiningvkvfviseslti nvdfvltrnqvtsqegvwdtallvneaidsigkmrnlqkegmrlnvksisknledeindsqilnysdkqiaakfslniqeeelfq ttlwdyetilsiinkgfvdsedlrelnlfkddqlnqnsqkmlkrkenydtfnvknfsqygdkkeqlknmftdsgvslskddwya ewkmvkskddfinqgnplnynenlekitenglnyweipnsftktgkrkrniivfnphshesvlskfsfdqilsnsflntnskkftiar gksliivnftldssepifktikykhkneniseftnfvlnfepeifnsksrfsnvksqilvntndedsfdi vfgtgskeiekliken geklylydesliiseqspawndgklsfklykdnnyapflikekssktlpvnsyviwnlkrnmenfifngvkaqvdnsvfylveefke flkmereliikqdi fya krnidgskleievsfsneletaymdifnyyktfddspednlpslvylnddlkelykktififnkeiseieens ilsdfkykknllklgrietdkimysplniayqlevskqcnedlsvnilerlvnnlipycsddgkelfrpiyqeeahewliye kseevisigttnvfisnvtelknqfvkhfnylfsfnnsipikinlinikddekvkvgnfirslpdkttkktkkipveiniyndaeks sfdnlfcdqseiqlleefgiklksdifiidpdiidihmiqnnsiyykhpfkfkeeyeahlsfykvkshnniandnmkdi etglsngllss vtsttkhseyrtgftnnilnmsnpliktvinlnelvensknfgntysknsvittveelednieelydkshwvtfieptfgieyfd sdsnlii ihydsqyssskydtitvntkstqyeiir dflqskvkvtddeelydvikmfsingewllrvinsghydrcklsisaiak yclsilhdhkdviwpvsmeeilriagnvklndknkgifdsklkgnhsddlifigvknednrievifypievkiglnnastikkgsq dntyklktqlqninvensefrnkfrnffiqill sneqlvtnhiwdekgl dreefkaellndeydilygleeyigkgslvskkes hhisiymdvdkqvielpedfayyglatp ireihdei qsdntdilaetllshvdi sei raknndicdsnedmsidddfdnlsefedsfie eeseiseepdeeltatsssdnesienigespskisinvr aligtqgynhkvweyefghpslanrhmligksggktyfiqrmkelsi qgipsi i idytdgfkpsql epnfkds lgdkisqyfvvkenfpinpfkrntimidkdi fieednstiasrfsiinsvylgiqqstly qtvldclidkyddnfdnlilkeeilkdesnaqtvl nklne lldknpfastdfdwsvldnkdgvyyi iqltalskdiqtiitefilwldw nykltngsednpi fivvldeahnl d fndspcskilkegrkfgwsgwfatqsvksgmki deia klenadekiyfhptdsvstia kds kdn edkkiye kelsqltkgycivqgsaidssgnlyqnpvptvkieeisfden	66
2064	meskdrientifndkrmdayidkyfeeflssqeaealnwidklnndqitsekgnynhnffeiilledllykrsdvkheenigdeghpv efvlekdgdgyviielkgttykdltkrpgqgspveqatnyasakketawatvsnnydeffrnfptardnyisfkrqlkdleifkklfl vfskfslidedipkllnetkvierelene fyqlysdtrlmii keleysednri eai ksqilnrflicfaedlameeettadv lltplkhrnligntmwrlne lfi fanqgnkhrripafngglfeddlnl kirdeiedrsffenwnlkedfedkyedia kl igvykdtl npifinll iistyd fdseldvnl ghifensi dieelkndnqekrkdgvytytpeyitdyicrntiipyisigkastvhelllyeyes snsldvlsklt n i kvl dpacgsgsm l nksvdl f e i h e a l h a s k y a g d s l d r f d s l e k r k e i i s n n i y g v d l n e e s v e i t k l s f l klattvqlkegfqlpsldkhi kcgds l v d d e s i a g n k a f n w y e s f e v f e s g g f d i i v g n p p y v d i k e m d e k t a k y i f d n y e t s f n r i n lystfveksyylknegifsfimpnsilfnstyskirelinlntsilnrvtsddvfkakvepiilifkkydegntk kilikddmd eipinnysehftqerwfennsiinifsd d f t d l l k k i d g n n e r l i d y c d f s l g l t p y d k y g m s e d i l k n r k f h s k i k l d d t f k e l l dgsdlt r y n v k w e g e y i k y g d w l g a p r e e k f f k n p r i l i r q i l s i a p k e s z k r i f a a y e e e l y n a q i a f n l v l k e g f d d k n l l k y f l	67

图 4B-6

GAMOLA ORF	氨基酸序列/描述	SEQ ID NO:
	giinskmmtwyyeerfmdknknkfakillienaknlpviinsnfldeivsnvdsiieelnkefysvrnafqtwlkiefieklskklenyy dlnfeflkeiikkkkvirpngiqdlselfneslgkieylqreikeadekinllvyelyglnheeieilensfnd	
2065	mddetlilieyirnaptrremvlsfegvdfirpiqisrktgilhpnnvskklldrehelvyvinpeyhvpklyrltekgktmlql	68
	噬菌体整合位点, 基因组整合基因座	
2069	mnkkiilslllavlsvsavaaadadvtyindaadvddvadekvapltasadaqdiqtkl dna kpgdtielenktydvdttfnvtkq vtikgqdtvfkasgasqgsgalfianeagtafegitfintdghknkygeqvgyaiqlaiengtvdncfkfidwssgvykgasfcsitn syfngsseqvtnggkkeygtkainlmgshditvtgctfegqvl da isia s nsgn imtdntfidncya i yfggastq	69

图 4B-7

SEQ ID NO:70	(101)	101	110	120	130	140	150
SEQ ID NO:71	PeP (98)	SRYNRFEKNGCRPPRVN	YS C--CGPSVSL	ETFKDMCKRAYNQF	EEENRR		
SEQ ID NO:72	PeW (83)	RRYEDFVRENGCRPNL	LEGC--KSDHVS	LCGTFKDMLEKYKDF	RIINGR		
SEQ ID NO:73	MRU ORF2058 (74)	-----VARNPLVMDYQNT	NYTCCPTSL	SLASONLYHYKSES	-----		
	共有 (101)	RY F INCRP VIFI NG	VSL TPKDMLKRYK FL	N R	区段 4		
	(151)	151	160	170	180	190	200
	PeP (146)	EPRIYVYTPPEPPVPR	REVRER	RLVLCERK	TAIQLYTLVS	RRCKYRFFYN	
	PeW (131)	EPNYLSQPPQLKCHWT	TKVIRKCTPH	-DATSLYRVRKTKYKYYN			
	MRU ORF2058 (110)	-----NCARALCTSR	CGTSPAQL	EEANAPRLGFKII			
	共有 (151)	EP I I P	EM K LGTRK SATSLY	LVAK CRYKFFYN	区段 5		
	(201)	201	210	220	230	240	250
	PeP (195)	DQTPNREAKKNTD	CINCTDACQL	PKPV	LCYSV	LIHVVV	CNDNK
	PeW (180)	DQVPHVAVHNT	TSCINCTDACQL	PSKVL	EEGCEYEVKI	HHVVKCNDCK	
	MRU ORF2058 (141)	PIKRDSEKVKKY	KKCGFPVICHQ	QNNQSRNCKDY	ICNF	-----	
	共有 (201)	DQ PN AVKRLT	CINCTDACQL	VIE LGYSVKIE	HHVVKCND K	区段 6	
	(251)	251	260	270	280	290	300
	PeP (245)	WYCHYFLRVAGKHL	ASVSLPSERWT	WDYVSA	TKTGRPL	CAPCCB	RCIQH
	PeW (230)	WYCHYLLRVLCF	ELKDC-----T	WDYVSA	TKTGRPL	CVPC	ACFQH
	MRU ORF2058 (180)	---CHYGLINDHT	STHYV-----VADPAK	GVNRKYK	FSCLDN	NANKGYR	
	共有 (251)	WYCHY LRVAG EL V	TVWDYVSA	TKTGRPL	C PCCS G QH	区段 7	
	(301)	301	312				
	PeP (295)	LCWGI	VSPKHD-				
	PeW (273)	LCWGI	VGPVMDK				
	MRU ORF2058 (220)	QNNYV	CCPA----				
	共有 (301)	LCWGI	V P HD				

图 5

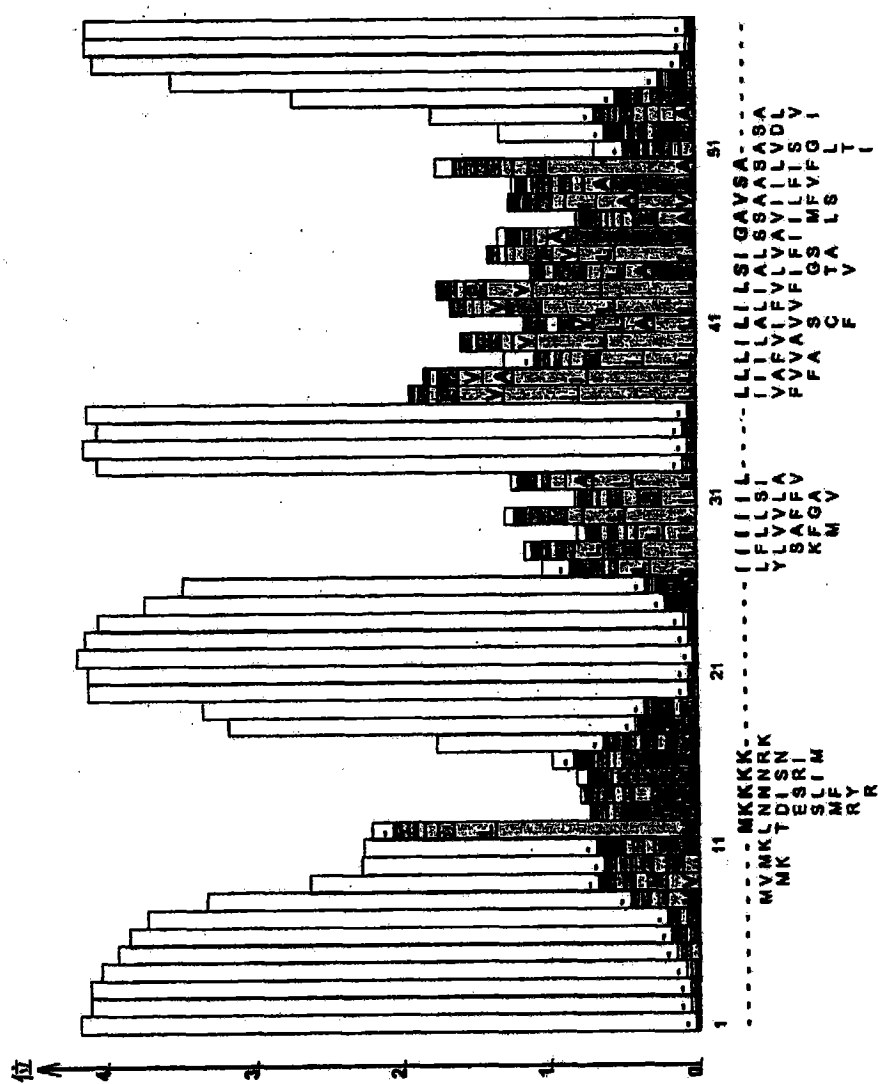


图 6

ORF 2058裂解酶对反刍甲烷短杆菌静息细胞的影响

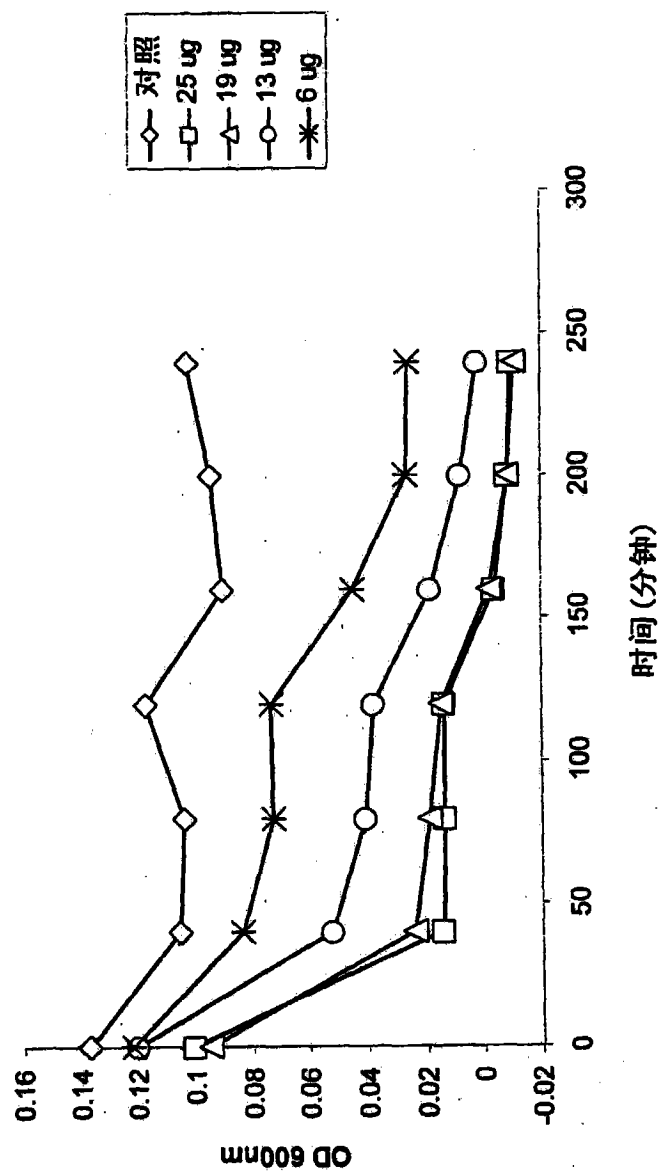


图 7

ORF 2058裂解酶对反刍甲烷短杆菌生长和甲烷形成的影响

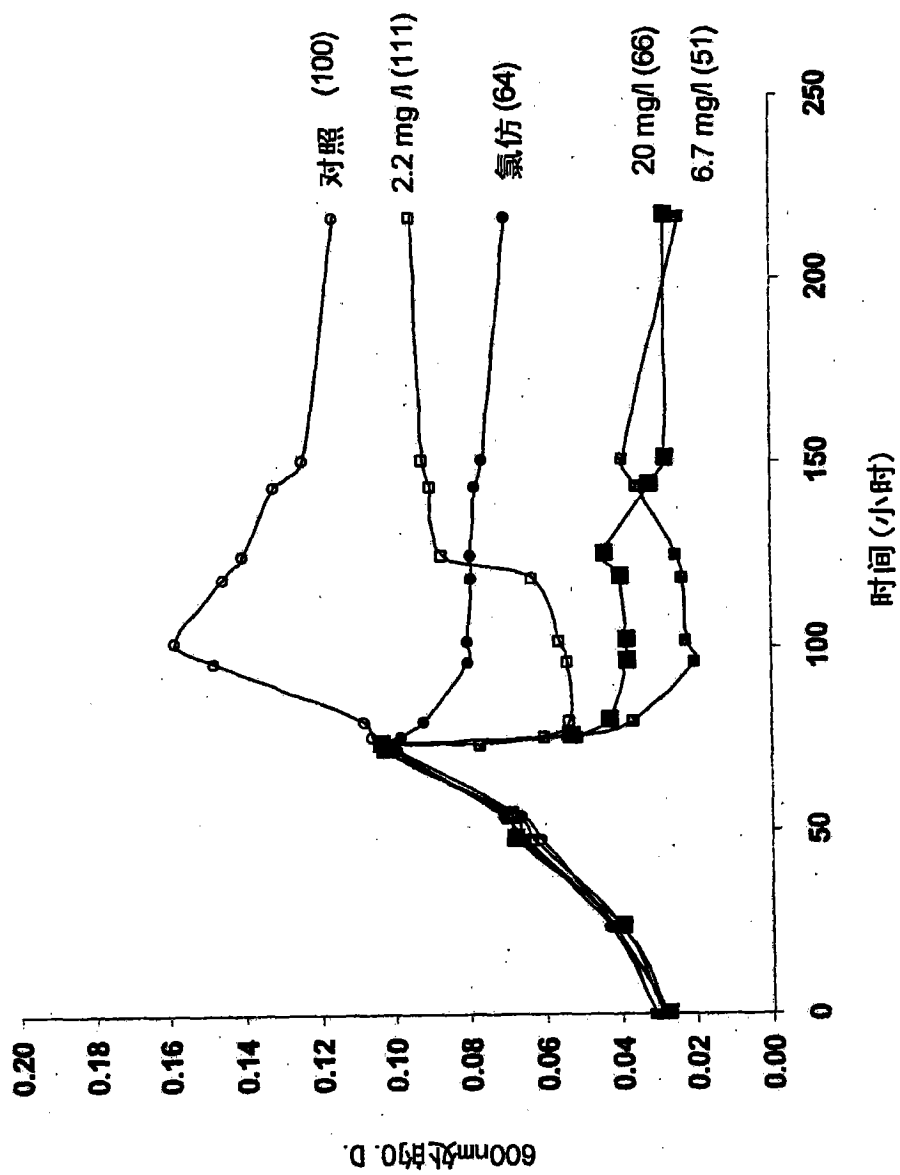


图 8