

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12N 5/10

C12N 15/63 C12N 15/79

C12N 15/85 C12P 21/06



[12] 发明专利申请公开说明书

[21] 申请号 02814382.5

[43] 公开日 2004年9月29日

[11] 公开号 CN 1533432A

[22] 申请日 2002.6.4 [21] 申请号 02814382.5

[30] 优先权

[32] 2001.6.4 [33] US [31] 60/295,961

[32] 2001.11.26 [33] US [31] 60/333,620

[32] 2002.1.29 [33] US [31] 60/352,404

[86] 国际申请 PCT/US2002/017763 2002.6.4

[87] 国际公布 WO2002/099089 英 2002.12.12

[85] 进入国家阶段日期 2004.1.16

[71] 申请人 ML 实验室公开有限公司

地址 英国伦敦

[72] 发明人 特里什·本顿

克里斯多佛·罗伯特·贝宾顿

卡尔拉·安·亨宁 戴维·J·金

罗伯特·克龙比 向 少

[74] 专利代理机构 中原信达知识产权代理有限责
任公司

代理人 丁业平 王维玉

权利要求书 6 页 说明书 49 页 序列表 60 页
附图 22 页

[54] 发明名称 高水平、大规模生产重组蛋白的组
合体和方法

[57] 摘要

本发明公开用于高水平、大规模生产重组蛋白的组合物和方法。示范的组合物包含一种或多种能够高水平表达重组蛋白和/或多肽的表达载体，和能够在无血清、悬浮培养条件下生长的无限增殖化宿主细胞系。双向 UCOE 载体允许在单个基于 UCOE 的质粒载体上同时高水平表达两种或多种重组蛋白和/或多肽。

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

-
1. 一种获得高水平、大规模蛋白和/或多肽表达的组合物，所述组合物包括：
- 5 (a) 能够在培养基中持续生长的无限增殖化宿主细胞系，其中所述宿主细胞系能够在无血清悬浮培养基中生长；以及
- (b) 持续过表达重组蛋白和/或多肽的载体，
- 其中所述宿主细胞系用所述载体转染。
- 10 2. 权利要求 1 的组合物，其中所述无限增殖化宿主细胞系倍增时间不大于 16 小时。
3. 权利要求 2 的组合物，其中所述倍增时间不大于 12 小时。
- 15 4. 权利要求 1 的组合物，其转染效率至少 70%。
5. 权利要求 4 的组合物，其中所述转染效率至少 75%。
6. 权利要求 4 的组合物，其中所述转染效率至少 85%。
- 20 7. 权利要求 4 的组合物，其中所述转染效率至少 95%。
8. 权利要求 1 的组合物，其中所述宿主细胞系易受选自潮霉素，G418 和嘌呤霉素的筛选剂的影响。
- 25 9. 权利要求 1 的组合物，其中所述宿主细胞系的特征在于所述重组蛋白和/或多肽没有半乳糖-半乳糖糖基化形式。
10. 权利要求 1 的组合物，其中所述宿主细胞选自 CHO-S，293-F，293-H，COS-7L，D.Mel-2，Sf21 和 Sf9。
- 30

11. 权利要求 1 的组合体, 其中所述载体进一步具有下述特征: (a) 含有一个或多个元件有助于在无限增殖化宿主细胞系中高水平、大规模表达; 以及(b)对重组蛋白和/或多肽的阻遏作用有抗性。

5

12. 权利要求 1 的组合体, 其中所述载体进一步包括一个或多个通用染色质开放元件(UCOE)。

13. 权利要求 1 的组合体, 其中所述组合体的特征在于每升培养基中能够得到至少 50mg 重组蛋白和/或多肽的表达水平。

10

14. 权利要求 13 的组合体, 其中所述组合体的特征在于每升培养基中能够得到至少 100mg 重组蛋白和/或多肽的表达水平。

15. 权利要求 13 的组合体, 其中所述组合体的特征在于每升培养基中能够得到至少 200mg 重组蛋白和/或多肽的表达水平。

15

16. 权利要求 1 的组合体, 其中所述组合体能够扩大培养至至少 100 升规模, 并且其中所述组合体能够生产至少 1 克蛋白和/或多肽。

20

17. 权利要求 16 的组合体, 其中所述组合体能够产生至少 10 克蛋白和/或多肽。

18. 权利要求 16 的组合体, 其中所述组合体能够产生至少 20 克蛋白和/或多肽。

25

19. 一种高水平、大规模生产蛋白和/或多肽的方法, 其中所述方法包括下列步骤:

(a) 获得能够以悬浮方式生长的无限增殖化宿主细胞系;

30

(b) 使所述无限增殖化宿主细胞系适合在无血清培养基中生

长；

(c) 用适用于高水平表达重组蛋白和/或多肽的载体转染所述适合无血清生长的无限增殖化细胞系。

5 20. 权利要求 19 的方法，其中所述无限增殖化宿主细胞系倍增时间不超过 16 小时。

21. 权利要求 19 的方法，其中所述倍增时间不超过 12 小时。

10 22. 权利要求 19 的方法，其转染效率至少 70%。

23. 权利要求 22 的方法，其中所述转染效率至少 75%。

24. 权利要求 22 的方法，其中所述转染效率至少 85%。

15

25. 权利要求 22 的方法，其中所述转染效率至少 95%。

26. 权利要求 19 的方法，其中所述宿主细胞系易受选自潮霉素，G418 和嘌呤霉素的筛选剂的影响。

20

27. 权利要求 19 的方法，其中所述宿主细胞系的特征在于所述重组蛋白和/或多肽没有半乳糖-半乳糖糖基化形式。

28. 权利要求 19 的方法，其中所述宿主细胞系选自 CHO-S，
25 293-F，293-H，COS-7L，D.Mel-2，Sf21 和 Sf9。

29. 权利要求 19 的方法，其中所述载体进一步包括下列特性：(a) 具有一个或多个元件有助于在无限增殖化宿主细胞中高水平、大规模表达以及(b)对重组蛋白和/或多肽的阻遏作用有抗性。

30

30. 权利要求 19 的方法，其中所述载体进一步包括一个或多个通用染色质开放元件(UCOE)。

5 31. 权利要求 19 的方法，其中所述方法的特征在于每升培养物能够获得至少 50mg 重组蛋白和/或多肽的表达水平。

32. 权利要求 31 的方法，其中所述方法的特征在于每升培养物能够获得至少 100mg 重组蛋白和/或多肽的表达水平。

10 33. 权利要求 31 的方法，其中所述方法的特征在于每升培养物能够获得至少 200mg 重组蛋白和/或多肽的表达水平。

34. 权利要求 19 的方法，其中所述方法能够扩大培养至至少 100 升规模，并且其中所述方法能够生产至少 1 克蛋白和/或多肽。

15

35. 权利要求 34 的方法，其中所述方法能够产生至少 10 克蛋白和/或多肽。

20 36. 权利要求 34 的方法，其中所述方法能够产生至少 20 克蛋白和/或多肽。

37. 一种用于高水平、大规模表达多亚基蛋白和/或多肽的双向载体，所述组合体包括：

- 25 (a) 至少一个 UCOE 元件；及
(b) 第一个转录启动子；及
(c) 第二个转录启动子；

其中所述 UCOE 元件与所述第一和第二个转录启动子可操作连接，以及其中所述第一个转录启动子与所述第二个转录启动子方向相反。

30

38. 权利要求 37 的双向载体，其中所述 UCOE 元件是 RNP UCOE。

39. 权利要求 37 的双向载体，其中所述第一个转录启动子选自人 CMV 启动子，鼠 CMV 启动子和人 β -肌动蛋白启动子。

40. 一种获得高水平、大规模蛋白和/或多肽表达的组合物，所述组合物包括：

(a) 无限增殖化宿主细胞系，其能够在培养基中持续生长，其中所述宿主细胞系能够在无血清悬浮培养基中生长，以及

(b) 权利要求 37 的双向载体，
其中所述宿主细胞系用所述载体转染。

41. 一种用于高水平、大规模生产蛋白和/或多肽的方法，所述方法包括下列步骤：

(a) 获得能够持续生长的宿主细胞系；

(b) 使所述宿主细胞系适应在无血清培养基中生长以产生能够在无血清培养基中持续生长的细胞系；

(c) 用权利要求 37 的载体转染能够在无血清培养基中持续生长的所述细胞系。

42. 权利要求 41 的方法，其中所述宿主细胞系能够持续生长，也能够以悬浮方式生长。

43. 权利要求 42 的方法，其中能够以悬浮方式持续生长的所述宿主细胞系是 CHO-S 细胞系。

44. 一种用于高水平、大规模表达多亚基蛋白和/或多肽的载体，其中所述组合物包括：

(a) 至少一个 UCOE 元件；及

(b) 转录启动子;

所述载体进一步含有如表 4 和图 14 所述的在 RNP UCOE 区域内选自 Δ BS, Δ EcoNI, Δ EM, Δ MluI, 和 Δ RV 中的的一个或多个缺失。

5 45. 权利要求 44 的载体, 其中所述缺失为在 RNP UCOE 区域内如表 4 和图 14 中所述的 Δ BS。

46. 权利要求 44 的载体, 其中所述缺失至少 100bp。

10 47. 权利要求 44 的载体, 其中所述缺失至少 1,000bp。

48. 权利要求 44 的载体, 其中所述缺失至少 4,000bp。

高水平、大规模生产重组蛋白的组合体和方法

5 发明背景

发明领域

本发明总的涉及基因表达和蛋白生产，更具体涉及过量表达重组蛋白的组合体(composition)和方法。这种组合体和方法在重组蛋白的高水平、大规模生产中十分有用。

10

相关技术描述

生物技术产业的主要目标是发展用于大规模表达重组蛋白，例如重组抗体的基于稳定细胞系的系统。开发适合的重组宿主细胞系的标准方法耗时费力。常规情况下，在含有胎牛血清的条件下培养细胞，
15 例如 CHO-K1 或 CHO DUX，将目的表达载体转染到细胞内。接着整个细胞群经过筛选去除未转入表达载体的细胞。所有含载体的细胞一般再经过亚克隆和高表达株筛选过程。扩增这样获得的每一株高表达株，并使之逐渐适应无血清悬浮培养，但这种适应通常导致重组蛋白和/或多肽表达的丧失。

20

除了这些常见的重组蛋白表达中的限制因素，有效的表达功能性多亚基蛋白，例如抗体，要求能够适当平衡的表达两条亚基链。例如，传统表达抗体重链和轻链的方法需要共转染分别带有重链和轻链编码区的质粒，但这仍难以使重链和轻链保持相同的拷贝数，而且可能由于载体彼此之间在基因组中整合的位置相近导致基因间的转录干扰。
25

因此，尽管有大量的研究，本领域仍需要改良的组合体和方法用于高水平大规模表达包括抗体重链和轻链在内的重组蛋白和/或多肽。通过使用预先适应了无血清悬浮培养的宿主细胞系以及适当的重组蛋白表达载体，本发明满足了这一需要，并进一步提供了其它相关有利
30

之处。本文还提供了一种双向 UCOE 载体，使得能够在基于单个 UCOE 质粒载体上同时高水平表达两种或多种重组蛋白和/或多肽。

发明概述

5 本法明旨在提供快速有效培养适于高水平表达、大规模生产重组蛋白和/或多肽的重组细胞系的组合体和方法。

10 一方面，本发明提供的组合体包括：(a)无限增殖化宿主细胞系，能够持续在培养基中生长，这种宿主细胞系能够在无血清悬浮培养条件下生长，以及(b)能够持续过表达重组蛋白和/或多肽的载体，例如本文描述的基于 UCOE 的载体。

15 本发明另一方面提供高水平大规模生产多肽的方法。特指的方法包括下述步骤：(a)获得能够以悬浮方式生长的无限增殖化宿主细胞系；(b)使宿主细胞系适应无血清培养；(c)用适于过表达重组蛋白和/或多肽的载体转染获得的能够在无血清培养基中悬浮生长的无限增殖化宿主细胞系。

20 根据本发明的组合体和方法，适当的无限增殖化宿主细胞系可以具有下述一种或多种特性：(a)倍增时间不超过 16 小时，优选的介于 12 和 16 小时之间；(b)转染效率至少 70%，优选至少 75%，80%，85%，90%或 95%；(c)易受常用的筛选试剂，例如，潮霉素，G418，嘌呤霉素的影响；(d)缺失半乳糖-半乳糖糖基化的重组蛋白和/或多肽。

25 在本发明权限范围中可被采用的示范性无限增殖化宿主细胞系包括，但不局限于下述可购买获得的宿主细胞系：(a)CHO-S(中国仓鼠卵巢细胞系)；(b)293-F(人宿主细胞系)；(c)293-H(人宿主细胞系)；(d)COS-7L(猴宿主细胞系)；(e)D.Mel-2(昆虫宿主细胞系)；(f)Sf21(昆虫宿主细胞系)；以及(g)Sf9(昆虫宿主细胞系)。或者，可以根据本文描述的
30 方法通过常规试验获得适当的宿主细胞系。

5 本发明的组合体和方法中适用的过表达重组蛋白和/或多肽的载体具有下述一种或多种特性：(a)含有一种或多种有助于在无限增殖化宿主细胞系中高水平大规模表达的元件，及(b)对重组蛋白和/或多肽的阻遏有抗性。

10 在特定的实施方案中，本发明的载体可以进一步包括一个或多个如下文中定义的通用染色质开放元件(UCOE)。而且或者选择性地，本文所述的载体可以包括一个或多个转录启动子，例如，CMV 启动子。

15 本发明优选的组合体和方法能够获得表达水平是每升培养物至少产生 50mg 重组蛋白和/或多肽，更优选的每升培养物至少 100mg 重组蛋白和/或多肽，更优选的至少每升 200mg 重组蛋白和/或多肽。

20 本发明进一步提供的组合体和方法能够将规模扩大至至少 100 升，产量(每 100 升培养物)至少 1 克蛋白和/或多肽，更优选的至少 5 克蛋白和/或多肽，更为优选的至少 10 克蛋白和/或多肽，而最优选的至少 20 克蛋白和/或多肽。

25 本发明进一步提供的组合体和方法采用能够在基于单个 UCOE 的质粒载体上高水平表达两种或多种重组蛋白的双向载体系统。示范性的双向载体系统可以包括一个或多个选自鼠 CMV 启动子，人 CMV 启动子和人 β -肌动蛋白启动子中的转录启动子。

30 本发明还提供用于改良表达一种或多种重组蛋白的组合体和方法，它们含有基于 RNP UCOE 的质粒载体，例如 CET720GFP，或者在 8kb 的 RNP UCOE 部分内含有一个或多个缺失。例证性的 UCOE 缺失结构优选保留显著的 UCOE 活性，例如，相对本文所述的 8kb RNP UCOE 元件有至少约 50%，优选至少 75%，而且更优选的至少 90%或更高的 UCOE 活性。有代表的缺失可以，或者选择性的，包括在 RNP

UCOE 区域内选自 Δ BS, Δ EcoNI, Δ EM, Δ mluI, 和 Δ RV 中的缺失, 如表 4 和图 14 所示。本发明范围内的缺失优选至少 100bp, 更优选至少 250bp, 更为优选至少 1000bp, 更优选的至少 2500bp, 更优选甚至至少 4000bp。因此本发明中特指的例证性 UCOE 载体最低包括至少一个或多个 UCOE 部分, 其中 UCOE 部分保留所需水平的 UCOE 活性。在一个说明性实施方案中, 采用至少约 4.1kb UCOE 部分, 对应于 CET720GFP(SEQ ID NO:9)的 5152-9254 位核苷酸残基。这种 UCOE 部分, 例如本文中所述的, 与位于 CET720GFP(SEQ ID NO:9)的 2225-10525 位核苷酸残基的 8kb 全长 UCOE 元件观察到的活性相比仍具有一定的活性。根据本文所述的路线和本领域承认的技术, 这些 UCOE 及其它 UCOE 部分易于鉴定, 其活性也易于测定。

参考下文详细的说明和附图, 本发明的各个方面均明白易懂。本文描述的所有参考文献均以其全文引作参考, 如同每一参考文献独立地纳入本文。

附图与序列标识符简述

图 1 图示了基于 UCOE 的的抗体表达盒。

图 2A 和 2B 是可用于表达重组人抗体的载体的质粒图谱。图 2A 显示表达重组人 Ig 重链的质粒。图 2B 显示表达重组人 Ig kappa 轻链的质粒。

图 3 描述了用与不用 UCOE 转染的 CHO 细胞的抗体表达水平。

图 4 显示在摇瓶和 2 升生物反应器中, 用表达抗体 Ab1 重链和轻链的载体转染 CHO-S 细胞系放大培养的结果。左图显示了 ELISA 方法检测的抗体滴度。右图显示细胞的生长。

图 5 描述了鼠杂交瘤、CHO-K1 和 CHO-S 细胞表面上 Gal-Gal 残基的水平。

图 6 图示了双向 UCOE 质粒载体 pBDUneo100。

图 7 图示了双向 UCOE 质粒载体 pBDUneo200。

图 8 图示了双向 UCOE 质粒载体 pBDUpuro300。

图 9 图示了双向 UCOE 质粒载体 pBDUpuro400。

图 10 图示了双向 UCOE 质粒载体 pBDUneo500。

图 11 图示了双向 UCOE 质粒载体 pBDUneo600。

图 12 图示了双向 UCOE 质粒载体 pBDUpuro700。

5 图 13 图示了双向 UCOE 质粒载体 pBDUpuro800。

图 14 图示了删除了 CET720GFP 中 8kb 的 RNP UCOE。

图 15 图示了双向 UCOE 质粒载体 pBDUpuro350。

图 16 图示了双向 UCOE 质粒载体 pBDUpuro450。

图 17 图示了双向 UCOE 质粒载体 pBDUneo1200。

10 图 18 图示了双向 UCOE 质粒载体 pBDUpuro1450。

图 19 图示了双向 UCOE 质粒载体 pBDUneo1600。

图 20 图示了双向 UCOE 质粒载体 pBDUpuro1800。

图 21 描述了含双向 UCOE 质粒载体的示例性细胞系产生抗体的速率。

15

序列标识简述

SEQ ID NO:1 是 pBDUneo100 的多核苷酸序列。

SEQ ID NO:2 是 pBDUneo200 的多核苷酸序列。

SEQ ID NO:3 是 pBDUpuro300 的多核苷酸序列。

20 SEQ ID NO:4 是 pBDUpuro400 的多核苷酸序列。

SEQ ID NO:5 是 pBDUneo500 的多核苷酸序列。

SEQ ID NO:6 是 pBDUneo600 的多核苷酸序列。

SEQ ID NO:7 是 pBDUpuro700 的多核苷酸序列。

SEQ ID NO:8 是 pBDUpuro800 的多核苷酸序列。

25 SEQ ID NO:9 是载体 CET720GFP 的多核苷酸序列。

SEQ ID NOs:10-26 代表了根据本发明为了制备改良 UCOE 载体，在实施例 4 中使用的示例性引物序列。

SEQ ID NO:27 是 pBDUpuro350 的多核苷酸序列。

SEQ ID NO:28 是 pBDUpuro450 的多核苷酸序列。

30 SEQ ID NO:29 是 pBDUneo1200 的多核苷酸序列。

SEQ ID NO:30 是 pBDUpuro1450 的多核苷酸序列。

SEQ ID NO:31 是 pBDUneo1600 的多核苷酸序列。

SEQ ID NO:32 是 pBDUpuro1800 的多核苷酸序列。

5 发明详述

本发明总的涉及用于高水平大规模生产重组蛋白和/或多肽的组份和方法。如下详述，本发明示例的组合体包括无限增殖化、无血清悬浮宿主细胞系连同—个或多个适于高水平大规模表达重组蛋白和/或多肽的表达载体，但并不局限于此。

10

除非特别指出，本发明使用的操作均是本领域熟悉的传统病毒学、免疫学、微生物学、分子生物学和重组 DNA 技术的方法，下面描述了其中的许多方法进行说明。这些技术在一些文献中有充分的说明。参考，例如，Sambrook 等，分子克隆：实验室手册(第二版，1989)；Maniatis 等，分子克隆：实验室手册(1982)；DNA 克隆：实用方法，(DNA Cloning: A Practical Approach)第一、二卷(D.Glover 编)；寡聚核苷合成(N.Gai 编，1984)；核酸杂交(B.Hames & S.Higgins 编，1985)；转录和翻译(B.Hames & S.Higgins 编，1984)；动物细胞培养(R.Frshney 编，1986)；Perbal，分子克隆应用指南(A Practical Guide to Molecular Cloning)(1984)。

15

20

本文引用的所用著作，专利和专利申请，无论是上文提及的还是下文将要提到的，均以其全文在此引作参考。

25

在本说明及所附权利要求中使用的单数形式“一个”及“这个”包括复数含义，除非内容中有清楚的其它说明。

制备并筛选无血清，悬浮培养的宿主细胞系

30

理想的适用于本发明的组合体及方法的宿主细胞系可以具有下述一种或多种特征：(a)能够永生，持续在培养基中生长；(b)适应悬浮生

长；(c)生长快速，倍增时间优选为 12-16 小时；(d)高转染效率，优选至少 70%；(e)易于通过常用的筛选试剂，优选潮霉素，G418 或嘌呤霉素加以筛选；(f)蛋白糖基化模式与用于人体治疗一致，优选缺失半乳糖-半乳糖糖基化形式；(g)适于在无血清培养基中生长，优选在无间接动物来源组分的化学组成的无蛋白的培养基中生长。

具有上述一种或多种特征的宿主细胞系可用于开发快速培养重组宿主细胞系的系统，这种重组细胞系较现有的方法更省时省力，能够转化成重组蛋白和/或多肽高水平、大规模生产的研发和制造中。

10

长期高产量生产重组蛋白，通常优选稳定表达。例如，稳定表达目的多核苷酸的细胞系可以用同时含有内源性表达元件和筛选标记基因的表达载体或表达元件和筛选基因独立存在的表达载体进行转染。细胞中引入载体后，在丰富培养基中生长 1-2 天换用筛选培养基。筛选标记可以使细胞对筛选产生抗性，其存在使得成功表达引入序列的细胞生长和回收。利用与细胞类型相适应的组织培养技术可增殖稳定的具有抗性的转化细胞株。

任何筛选系统都可用于获得转化细胞系。包括，但不局限于，单纯疱疹病毒胸苷激酶(Wigler, M.等(1977)细胞 11:223-32)基因和腺嘌呤磷酸核糖基转移酶(Lowy, L.等(1990)细胞 22:817-23)基因可分别用于筛选 tk.sup.-或 aprt.sup-细胞。同样，抗代谢药、抗生素或除草剂抗性也可以用于选择；例如，dhfr可以使细胞对氨甲蝶呤产生抗性(Wigler, M.等(1980)美国国家科学院院报(Proc. Natl. Acad. Sci.)77:3567-70)；谷氨酰胺合成酶(GS)可以使细胞不依赖谷氨酰胺生长并对 Sulphoximine 产生抗性 (Bebbingtong 等(1992)生物技术 10(2):169-75；和 Cockett 等(1991)核酸研究 25；19(2):319-25；npt 可以使细胞对氨基葡萄糖苷、新霉素和 G-418 产生抗性(Colbere-Garapin, F.等(1981)分子生物学杂志 150:1-14)；以及 als 或 pat 可以使细胞分别对绿磺隆(chlorsulfuron)和 phosphinotricin 乙酰基转移酶产生抗性(Murry, 同上)。还描述了其他

5 可选择的基因，例如，*trpB* 使细胞可以利用色氨酸上的吲哚基团，*hisD* 使细胞可以利用组氨酸上的 *histinol* 基团(Hartman, S.C.和 R.C. Mulligan(1998) 美国国家科学院院报 85:8047-51)。利用可视标记物已获得广泛欢迎，例如，花色素苷、 β -葡萄糖苷酸酶和它的底物 GUS，以及荧光素酶和它的底物荧光素不仅广泛用于鉴定转化子还用于对特定载体系统进行瞬时或稳定蛋白表达的定量分析(Rhodes, C.A.等(1995) 分子生物学方法 55:121-131)。

10 虽然标记基因表达存在/缺乏说明目的基因存在，但是目的基因的存在与表达仍需鉴定。例如，如果编码多肽的序列插入标记基因内，含有此序列的重组细胞可以通过标记基因功能的缺失得到确认。或者，标记基因与多肽编码序列串联排列并受同一个启动子调控。标记基因经诱导或在筛选条件下表达通常也说明与其串联的基因表达。

15 或者，含有并表达所需多核苷酸序列的宿主细胞可以通过本领域技术人员熟知的多种程序得到鉴定。这些程序包括，但不局限于，DNA-DNA 或 DNA-RNA 杂交和蛋白生物分析或免疫分析技术，这些技术包括，例如，检测和/或定量核酸或蛋白的基于膜、溶液或芯片的技术。

20

使用对多核苷酸编码的产物特定的多克隆或单克隆抗体检测和测量所述产物的多种方法是本领域技术人员熟知的。例子包括酶联免疫吸附分析(ELISA)，放免分析(RIA)，及荧光活化细胞分类(FACS)。在一些应用中优选使用针对特定多肽上的两个独立表位的单克隆抗体进行两点单抗免疫分析，但是也可以采用竞争性结合分析。这些及其他分析方法在其它著作中也有描述，如 Hampto, R 等(1990; 血清学方法，实验室手册，APS 出版社，St. Paul. Minn.)及 Maddox, D.E.等(1983, 实验医学杂志 158:1211-1216)。

30

广泛使用的多种标记和偶联技术是本领域熟练技术人员熟知的，

- 可以用于各种核酸和氨基酸分析。制备检测相关多核苷酸的序列的标记杂交或 PCR 探针的方法包括寡聚标记，缺刻翻译，末尾标记或用标记核苷酸的 PCR 扩增。或者，将序列或其任何区段克隆到载体上制备 mRNA 探针。这些载体为本领域熟知，可购买获得，在加入适当的 RNA 聚合酶如 T7, T3 或 SP6 及标记核苷酸后能够用于体外合成 RNA 探针。可以购买各种试剂盒进行这些操作。能够使用的适当的报告分子或标记包括放射核素，酶，荧光，化学发光，或发色剂以及底物，辅助因子，抑制因子，磁性颗粒等等。
- 5
- 10 转化有目的多核苷酸序列的宿主细胞可在适宜于表达蛋白和从细胞培养物中回收蛋白的条件下培养。根据所用的序列和/或表达载体，重组细胞所产生的蛋白分泌出来或存在于细胞内。熟悉本领域的技术人员能够知道，含有本发明多核苷酸的表达载体可以包括信号序列，这段序列可以指引编码的多肽分泌通过原核细胞或真核细胞膜。还可以将编码多肽区的核苷酸序列与编码目的多肽的序列连接，这样的重组构建有助于可溶性蛋白的纯化。这种利于纯化的区域包括，但不局限于，金属螯合肽例如组氨酸-色氨酸模块可以在固定金属上纯化，蛋白 A 区在固定的免疫球蛋白上纯化，以及利用 FLAGS 扩展/亲和纯化系统(Immunex 公司，Seattle,WA)的结构域。可切割的接头序列包括如
- 15
- 20 因子 XA 或肠激酶(Invitrogen)的特异识别序列，其位于纯化区和编码多肽之间，可便于纯化。一种这样的表达载体涉及含有目的多肽和编码 6 个组氨酸残基的融合蛋白的表达，6 个组氨酸残基之后有硫氧还蛋白或肠激酶的切割位点。如 Porath, J 等人(1992, 蛋白表达纯化 3:263-281)所述，组氨酸残基有助于在 IMIAC(固定金属离子亲和色谱)
- 25
- 30 上纯化，而肠激酶切割位点可以把目的多肽从融合蛋白中纯化出来。Kroll, D.J.等人对含有融合蛋白的载体进行过讨论(1993; DNA 细胞生物学 12:441-453)。

- 无血清无限增殖化宿主细胞系易于从各种公开的和/或商业途径获得，例如，美国典型培养物保藏中心(ATCC; Manassas, VA), Celox(St.

Paul, MN); Invitrogen(Carlsban, CA); 欧洲和日本细胞库(分别为 ECACC, Salisbury, Wiltshire(UK)和 JCRB, Shinjuky, Japan)。

5 适宜的宿主细胞系可以通过筛选具有一种或多种上述特点的现有宿主细胞系获得，并改造和/或选择宿主细胞系的变异株以获得其它特性。使用预先改造的宿主细胞系确保细胞能够在转染和重组蛋白表达前获得想要的条件。如下所示，这种细胞系适合与含有 UCOE 的表达载体共同使用，因为这些载体系统具有长期稳定高水平表达蛋白的特点。

10

根据本发明的组合体和方法，可以用于改良和/或改造的适宜的示范性宿主细胞系包括，但不局限于，(a)293-F，人宿主细胞系；(b)293-H，人宿主细胞系；(c)COS-7L，猴宿主细胞系；(d)D.MEL-2，昆虫宿主细胞系；(e)SF-21，昆虫宿主细胞系；(f)SF9，昆虫宿主细胞系；(g)CHO-S，中国仓鼠卵巢宿主细胞系。

15

例如，中国仓鼠卵巢亚克隆(CHO-S; Invitrogen/Gibco)已经能够获得商品化化学组成的无蛋白培养基，适合在本发明的组合体和方法中使用。见，D Anna 等，放射研究 148:260-271(1997); D Anna 等，20 细胞科学方法(Methods in Cell Science) 18:115-125(1996); Deaven 等，染色体 41:129-144(1973); Gorfein 等，动物细胞技术：基础和应用 9:247-252(Kluwer Academic Publishers, Netherlands, 1998)。CHO-S 宿主细胞系摇瓶培养 12 到 16 小时生长一倍，细胞密度达到峰值 $9-11 \times 10^6$ 细胞/ml。它们对浓度为 400ug/ml 的潮霉素和浓度为 600ug/ml 的新霉素(G418)敏感。即使静止培养，这些细胞也是悬浮单细胞生长。

25

临床上使用的重组蛋白存在 Gal α 1 \rightarrow 3Gal β 1 \rightarrow 4GlcNAc-R(Gal-Gal)碳水化合物残基，这与蛋白被快速从血清中清除有关。虽然在人的糖蛋白中没有发现 Gal-Gal 残基，但啮齿动物细胞分泌的糖蛋白碳水化合物结构是 Gal-Gal 二糖末端。因此，制备没有这种特定碳水化

30

合物结构的重组蛋白是有利的。

如下所述，CHO-S 宿主细胞系是优选的，它与含有一个或多个 UCOE 元件的表达载体共同使用。这种宿主细胞系具有良好的生长特性，并且在它表面的糖蛋白中未检测到 Gal-Gal 碳水化合物。因此，CHO-S 宿主细胞系适合表达临床使用的重组蛋白和/或多肽。

表达载体的制备和选择

根据本发明，适合表达重组蛋白和/或多肽的载体系统包括一种或多种如下特点：(a)易于操作；(b)使高水平表达整合位点独立的元件；(c)使表达能抗沉默/阻遏的元件，从而获得持久稳定的长期表达；以及(d)在不同细胞型和不同物种中高水平表达的元件。

为了表达目的蛋白和/或多肽，将编码多肽的核苷酸序列或功能等效物插入到适宜的表达载体，即载体含有对插入的编码序列的转录和翻译所必须的元件。构建含有编码目的多肽的序列和适当转录和翻译控制元件的表达载体，所使用的方法是本领域的技术人员所熟知的。这些方法包括体外重组 DNA 技术、合成技术和体内遗传重组。这些技术参见如 Sambrook, J 等(1989)分子克隆：实验室手册，冷泉港出版社，Plainview, N.Y.和 Ausubel, F.M.等(1989)当代分子生物学指南 (Current Protocols in Molecular Biology), John Wiley&Sons, New York. N.Y.。

含有和表达多核苷酸序列的各种表达载体/宿主系统均可使用。这些系统包括但不限于质粒或粘粒 DNA 表达载体；用病毒表达载体(如杆状病毒)感染的昆虫细胞系统；病毒表达载体(菜花花叶病毒，CaMV；烟草花叶病毒，TMV)转化的植物细胞系统；或动物细胞系统。

表达载体中的“控制元件”或“调控序列”是载体中的非翻译区——增强子，启动子，5'和 3'非翻译区——这些非翻译区与宿主细胞

蛋白相互作用以调控转录和翻译。这些元件的长度和特异性可有变异。根据使用的载体系统和宿主细胞可以选用任何适当的组成型和诱导型的转录和表达元件。在哺乳细胞系统中，哺乳细胞基因或哺乳细胞病毒基因的启动子都是优选的。如果必须获得含有多拷贝编码多肽序列的细胞系，含有 GS 或 DHFR 选择标记的载体或带有适合选择标记的基于 SV40 或 EBV 的载体优先使用。

昆虫表达系统也可以用于表达目的多肽。例如，在秋粘虫 (*Spodoptera frugiperda*) 细胞或粉纹夜蛾 (*Trichoplusia*) 幼虫中，使用 Autographa 加利福尼亚核多角体病毒 (AcNPV) 做载体表达外源基因。编码多肽的序列可以克隆到病毒的非必需区，如多角体蛋白基因，并且置于多角体蛋白启动子的控制下。成功插入编码多肽的序列使多角体蛋白基因失活并且产生缺乏衣壳蛋白的重组蛋白。然后，重组病毒可用于感染，如秋粘虫细胞或粉纹夜蛾幼虫，从而目的蛋白在其中得以表达 (Engelhard, E.K. 等 (1994) 美国国家科学院院报 92:3224-3227)。

在哺乳动物宿主细胞，通常可以获得来源于病毒的多种表达系统。例如，如果使用腺病毒作为表达载体，可以将编码目的多肽的序列连接到腺病毒转录/翻译复合物中，该复合物由晚启动子和三分式先导序列组成。在病毒基因组的非必需区 E1 或 E3 区插入则可以获得有活性的病毒，其在感染宿主细胞中能够表达多肽 (Logan, J. 和 Shenk, T (1984) 美国科学院院报 81:3655-3659)。另外，转录增强子，例如劳氏肉瘤病毒 (RSV) 增强子，能够用于提高哺乳动物宿主细胞的表达水平。

还可以使用特异起始信号获得更有效的目的多肽编码序列的翻译。这种信号包括 ATG 起始密码子和邻近序列。如果多肽编码序列的起始密码子和上游序列都插入到适当的表达载体中，就不需要其它的转录或翻译调控信号。然而，如果只有编码序列或其一部分插入载体中，则需提供外源转录调控信号，包括 ATG 起始密码子。而且，

起始密码子必须在正确的阅读框架内以确保整个插入序列的翻译。外源翻译元件和起始密码子可以有不同的来源，天然的及合成的。表达效率可以通过加入增强子得以提高，增强子适用于特定的细胞系统，例如一些著作中描述的那些增强子(Scharf, D.等(1994)Results Probl. Cell Differ., 20:125-162)。

实施例优选的元件适于使高表达水平与插入位点无关，包括，通用染色质开放元件(UCOE)。UCOE 是保持染色质处于“开放”构型的多核苷酸序列。参照，例如，Crombie 等，PCT 专利申请号 WO0005393(2000)。在表达载体启动子上游含有 UCOE 能够产生高表达水平，这与插入位点无关并对基因沉寂具有抗性。与通常含有非 UCOE 的载体相比，单拷贝插入基因位点即可获得有效的表达，这使得在筛选的细胞群体中表达标记基因的细胞百分率提高。再结合使用适于无血清悬浮培养的亲代细胞系就可在短时间内快速生产大量蛋白。使用 UCOE 载体能够提高效率，显著减少了需经筛选才能获得高产亚克隆的转化子数目。

在悬浮培养的宿主细胞系中，利用含有一个或多个 UCOE 的载体，可以快速研发和大规模生产蛋白和/或多肽，例如，抗体或其片段。由于 UCOE 的存在，只需筛选少量亚克隆即可在无血清条件下于 5 周内获得产量至少为 50mg/L 蛋白和/或多肽，优选至少为 100mg/L 蛋白和/或多肽，更优选至少为 200mg/L 蛋白和/或多肽的克隆。

优选的，本发明的成分和方法中使用的表达载体系统每升悬浮培养物的表达产量能够超过 1g 蛋白和/或多肽。更优选的，能够用于稳定的宿主细胞系中的表达载体至少每个细胞每天产生 20pg 蛋白和/或多肽。

如下文中详细讨论的，本发明特定实施方案中，蛋白和/或多肽可以含有一个或多个亚基，例如，抗体重链和轻链及其片段。正如本

领域熟知的，有效的功能性抗体需要适当平衡的表达重链和轻链。转染分别含有重链或轻链的不同质粒难以保持它们相同的拷贝数，而且如果两种载体在基因组中插入位点彼此邻近可能产生基因间的转录干扰。因此，可以采用能够在同一载体上共表达两种基因的双向载体。

5 如本文下述实施例中详述的，作为范例的双向基于 UCOE 的载体系统，在本发明范围内，可以根据“杂交”RNP/ β -肌动蛋白 UCOE(Cobra Therapeutics)任选加以构建。载体可以含有一个或多个抗生素抗性标记，例如，新霉素或嘌呤霉素抗性标记，和/或可含有一个或多个哺乳动物启动子，例如，鼠 CMV 启动子(mCMV)，人 CMV 启动子(hCMV)，
10 或人肌动蛋白启动子，以驱动轻链或重链的表达。

用本发明的表达载体转染宿主细胞系

转染常用的宿主细胞系，事先已适应了在大规模培养装置中生长，能够快速进行细胞系培养从而加速了从研究到开发应用的转换速度。相反，传统方法使用转染后逐步适应无血清悬浮培养的亲代细胞系，更需要筛选大量亚克隆，这是由于许多亚克隆不能在大规模生产蛋白的条件下生长。使用预先适应的细胞系能够将研发细胞系的时间从几个月缩短至几周。预先适应了化学组成、无蛋白培养基的细胞系能够在摇瓶或生物反应器中快速生长至高细胞密度。

20

适用的转染程序很容易为本领域熟练技术人员掌握或获得。作为范例的转染程序适用于获得高水平、大规模的转染，如 Invitrogen/Gibco 推荐的 CHO-S 宿主细胞系的转染。通常，筛选阳性转化细胞可以使用的试剂，如，潮霉素，G418，及嘌呤霉素。典型的转染效率至少是
25 70%，更优选至少 75%，80%，85%，90%或 95%。转染和筛选后的所有克隆可以任选进一步被亚克隆以鉴定出蛋白表达水平最高的单个克隆。

细胞培养条件的选择

30 专业人员根据本发明通过常规试验即可挑选和测试适于无限增殖

化悬浮细胞培养的无血清培养基。对于上文描述的 CHO-S 细胞，适用 CD-CHO 培养基(例如可获自 Invitrogen 或 Gibco)。

适于高水平大规模生产的示例性蛋白和/或多肽

5 本文使用的术语“蛋白”和“多肽”是常见的含义，即是氨基酸序列。多肽并不限制为特定长度的产物；因此，肽，寡肽，和蛋白均在多肽的范畴内，除非特别说明这些术语可以替换使用。此术语也并非特指或排除表达后修饰的多肽，例如，糖基化，乙酰基化，磷酸化等，以及本领域熟悉的其它修饰，无论天然存在的还是非天然存在的。
10 如上文所述，无论如何本发明优选的蛋白和/或多肽缺失半乳糖-半乳糖糖基化。多肽可以是完整的蛋白，或是其部分序列。本发明上下文中特定的目的多肽是包括表位的部分氨基酸序列，所述表位即是基本上使多肽产生免疫原性并引发免疫反应的抗原决定簇。

15 特定优选的实施例中，本发明产生和/或使用的多肽有免疫原性，即，在免疫分析中(如 ELISA 或 T 细胞刺激分析)能够检测到与抗血清和/或癌症病人的 T 细胞反应。可以使用专业人员熟知的技术筛选免疫原活性。例如，使用 Harlow 和 Lane 在《抗体：实验室指南(Antibodies: A Laboratory Manual)，冷泉港实验室，1988》中描述的方法进行筛选。
20 在一个说明性实施例中，多肽固定于固相支持物上，与病人的血清接触以使血清中的抗体结合到固定化的多肽上。去除未结合的血清，检测结合的抗体，例如使用 ^{125}I 标记的蛋白 A 检测。

 根据本文提供的说明制备的多肽免疫原部分可能被专业人员识别，因此这些免疫原部分也属于本发明范畴。本文使用的“免疫原部分”是本发明中免疫原性多肽的片段，本身可与识别多肽的 B 细胞和/或 T 细胞表面抗原受体产生免疫反应(例如特异性结合)。免疫原部分通常可以经本领域熟知的技术进行鉴定，如 Paul 在《基础免疫学，第三版，243-247(Raven 出版社，1993)》中概述的以及本文引用的其它
25 文献。这些技术包括筛选能够与抗原特异性抗体，抗血清和/或 T 细胞
30

系或克隆反应的多肽。如本文中应用的抗血清和抗体，如果可以特异地与抗原结合则是“抗原特异性”的(即在 ELISA 或其它免疫分析中与蛋白反应，但不能检测到与其它不相关的蛋白反应)。这种抗血清和抗体可以根据本文描述及熟知的技术进行制备。

5

在一个优选实施方案中，本发明多肽的免疫原部分是能够与抗血清和/或 T 细胞进行反应，反应活性基本上不低于全长多肽(如 ELISA 和/或 T 细胞反应分析)。优选的，免疫原部分免疫原性活性至少为全长多肽的 50%，优选至少 70%，最优选的大于约 90%。在一些情况下，
10 优选免疫原部分的免疫原性活性与对应的全长多肽相比更强，例如活性强于约 100%或 150%或者更高的免疫原性活性。

在其它一些实施方案中，示例的免疫原部分可能包括 N-末端先导序列和/或跨膜区序列缺失的肽。其它示例的免疫原部分相对成熟蛋白可以含有少量 N-和/或 C-末端缺失(例如 1-30 个氨基酸，优选 5-15
15 个氨基酸)。

另一实施例中，本发明制备和/或使用的蛋白和/或多肽可以包括一种或多种多肽，它们可以与针对本发明多肽所产生的 T 细胞和/或抗体发生免疫反应，尤其是含有本文描述的氨基酸序列的多肽，或其免疫原性片段或变体。
20

本文使用的术语多肽“变体”，与本文具体描述的多肽有明显差别，有单个或多个替换，缺失，添加和/或插入。这种变体是天然存在的或合成的，例如，修饰上述本发明的一种或多种多肽序列并用本文
25 描述的方法和/或其他本领域熟知的技术检测其活性。本发明中示例的变体序列与本文提供的 8kb RNP UCOE 序列，或其部分序列相关、有一定的同源性，并保留所需的 UCOE 活性。

30 在一实施方案中，例如，特别是本发明示例的变异序列包含的多

核苷酸序列与此处特别涉及的 UCOD 多核苷酸至少有 70%、75%、80%、85%、90%、95%或 99%或更高的同一性。当与本发明涉及的 8kb RNP UCOE 元件所示的 UCOE 活性比较时，这些优选的变体显示了至少 70%、75%、80%、85%、90%、95%或 100%或更高的 UCOE 活性。

5

在许多情况中，变体会含有保守取代。“保守取代”是一个氨基酸取代另一个氨基酸，而特性相似，熟悉肽化学领域的技术人员希望多肽的二级结构和水疗性没有根本改变。如上所述，对本发明的多核苷酸和多肽的结构可进行修饰，还得到了编码变异或衍生多肽的功能性分子，它们具有所要的特征，比如具有免疫原性特征。当想要改变多肽的氨基酸序列产生一个与本发明多肽等效或甚至改良的变体或部分，熟悉本领域的技术人员根据表 1 有代表性的改变一个或多个编码 DNA 序列的密码子。

10

15

例如，在蛋白结构中某个氨基酸被另一个氨基酸取代，而与结构的相互作用结合能力没有明显降低，比如与抗体的抗原结合区或底物分子上的结合位点的结合能力。因为蛋白的相互作用能力和性质限定了蛋白的生物学功能活性，所以可以在蛋白序列中进行某个氨基酸取代，当然，DNA 编码序列先发生改变，仍然可以得到特性相近的蛋白。因此，预计本发明公开的组合体的肽序列和编码该肽的相应 DNA 序列可以进行各种改变，而它们的生物利用度或活性不会显著降低。

20

表 1

氨基酸			密码子					
丙氨酸	Ala	A	GCA	GCC	GCG	GCU		
半胱氨酸	Cys	C	UGC	UGU				
天冬氨酸	Asp	D	GAC	GAU				
谷氨酸	Glu	E	GAA	GAG				
苯丙氨酸	Phe	F	UUC	UUU				
甘氨酸	Gly	G	GGA	GGC	GGG	GGU		
组氨酸	His	H	CAC	CAU				
异亮氨酸	Ile	I	AUA	AUC	AUU			
赖氨酸	Lys	K	AAA	AAG				
亮氨酸	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU
甲硫氨酸	Met	M	AUG					
天冬酰胺	Asn	N	AAC	AAU				
脯氨酸	Pro	P	CCA	CCC	CCG	CCU		
谷氨酰胺	Gln	Q	CAA	CAG				
精氨酸	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
丝氨酸	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
苏氨酸	Thr	T	ACA	ACC	ACG	ACU		
缬氨酸	Val	V	GUA	GUC	GUG	GUU		
色氨酸	Trp	W	UGG					
酪氨酸	Tyr	Y	UAC	UAU				

在进行这些变化时，应考虑氨基酸的水疗指数。在本领域，通常可以理解水疗氨基酸指数对蛋白相互作用的生物功能很重要(Kyte 和 Doolittle, 1982, 此处引作参考)。认为氨基酸的相对水疗特性影响蛋白的二级结构，从而规定了蛋白与其它分子的相互作用，例如，酶、底物、受体、DNA、抗体、抗原等。每个氨基酸根据它的疏水性和电荷都有一个水疗指数(Kyte 和 Doolittle,1982)。这些值为：异亮氨酸(+4.5)；缬氨酸(+4.2)；亮氨酸(+3.8)；苯丙氨酸(+2.8)；半胱氨酸/胱氨酸(+2.5)；甲硫氨酸(+1.9)；丙氨酸(+1.8)；甘氨酸(-0.4)；苏氨酸(-0.7)；丝氨酸(-0.8)；色氨酸(-0.9)；酪氨酸(-1.3)；脯氨酸(-1.6)；组氨酸(-3.2)；谷氨酸(-3.5)；谷氨酰胺(-3.5)；天冬氨酸(-3.5)；天冬酰胺(-3.5)；赖氨酸(-3.9)和精氨酸(-4.5)。

本领域已知某些氨基酸可被其它具有相似水疗指数或分数的氨基酸取代，所得到的蛋白仍具有相似的生物学活性，即仍得到生物学

功能等同的蛋白。在进行这些改变时，取代氨基酸的水疗指数在 ± 2 以内为优选， ± 1 为特别优选， ± 0.5 为更优选。根据亲水性有效地替代相似的氨基酸也是本领域熟知的。美国专利 4,554,101(以其全文具体在此引作参考)，说明蛋白局部平均最大的亲水性，由其邻近氨基酸决定，与蛋白的生物特性有关。

如美国专利 4,554,101 所述，氨基酸残基的亲水值如下，精氨酸(+3.0);赖氨酸(+3.0);天冬氨酸(+3.0 \pm 1);谷氨酸(+3.0 \pm 1);丝氨酸(+0.3);天冬酰胺(+0.2);谷氨酰胺(+0.2);甘氨酸(0);苏氨酸(-0.4);脯氨酸(-0.5 \pm 1);丙氨酸(-0.5);组氨酸(-0.5);半胱氨酸(-1.0);甲硫氨酸(-1.3);缬氨酸(-1.5);亮氨酸(-1.8);异亮氨酸(-1.8);酪氨酸(-2.3);苯丙氨酸(-2.5);色氨酸(-3.4)。已知氨基酸可被其它具有相近亲水值的氨基酸取代，所得到的蛋白仍具有等同的生物学活性，特别是免疫学上等同的蛋白。在进行这些改变时，取代氨基酸的亲水值在 ± 2 以内为优选， ± 1 为特别优选，而 ± 0.5 为更优选。

综上所述，氨基酸取代通常基于氨基酸侧链取代基的相对相似性，例如，它们的疏水性、亲水性、电荷、大小等。熟悉本领域的技术人员考虑到上述各种特征已知可以举例的取代包括：精氨酸和赖氨酸；丝氨酸和苏氨酸；谷氨酰胺和天冬酰胺；缬氨酸、亮氨酸和异亮氨酸。

另外，可以进一步修饰多核苷酸以增加它在体内的稳定性。可能的修饰包括，不但局限于，增加 5'和/或 3'末端的侧翼序列；在骨架中使用硫代磷酸酯或 2'O-甲基而不用磷酸二酯酶连接；和/或含有非传统的碱基如次黄嘌呤核苷、Q 核苷(queosine)和 wybutosine，以及腺嘌呤、胞嘧啶、鸟嘌呤、胸腺嘧啶和尿嘧啶的乙酰、甲基、硫代和其它修饰形式。

进行氨基酸取代还基于残基的极性、电荷、可溶性、疏水性、亲

水性和/或两性分子的本性中的相似性。例如，带负电荷氨基酸包括天冬氨酸和谷氨酸；带正电荷的氨基酸包括赖氨酸和精氨酸；不带电荷的极性首基的氨基酸具有相似的亲水值，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；以及丝氨酸、苏氨酸、苯丙氨酸和酪氨酸。其他组的氨基酸可能意味保守改变，包括(1)ala, pro, gly, glu, asp, gln, asn, ser, thr; (2)cys, ser, tyr, thr; (3)val, ile, leu, met, ala, phe; (4)lys, arg, his; 和(5)phe, tyr, trp, his。变体可能还含有非保守的改变。在一个优选的实施方案中，通过取代、删除或增加 5 个或 5 个以下的氨基酸，得到的变异多肽与天然序列不同。变体还可能被修饰，如删除或增加对多肽的免疫原性、二级结构和水溶性有极微影响的氨基酸。

如上所述，在蛋白的 N 末端，多肽可能含有一个信号(或引导)序列，共同翻译或翻译后引导蛋白转移。多肽还可以与接头和其它序列结合，使多肽易于合成、纯化或鉴定(如聚组氨酸接头)，或增加多肽与固相支持物的结合。例如，多肽可以与免疫球蛋白的 Fc 区结合。

如下所述，当比较两个多肽序列时，如果氨基酸序列完全一致，则说明这两个序列是“同一性”的。两个序列的比对通常是比较比对窗内的序列，鉴定和比较序列局部区域的近似性。本文使用的“比对窗”是指至少有 20 个邻近位点的区段，通常 30-约 75，40-约 50，其中一个序列可以与具有同样数目邻近位点的参考序列在这两个序列最佳比对后进行比较。

使用生物信息学软件(DNASTAR, Inc., Madison, WI)Lasergene 组中的 Megalign 程序，使用默认参数可以进行序列的最佳比对。此程序包含几个同源性比较方案，描述见下面参考文献：Dayhoff, M.O.(1978) 蛋白进化模型——检测远源关系的矩阵。蛋白序列和结构图集 Dayhoff, M.O.(编)，国家生物医学研究基金会，Washington DC Vol.5, Suppl.3, 第 345-358 页；Hein J.(1990)同源性比较和系统发育统一标准化方法，第

626-645 页, 酶学方法, 183 卷, Academic Press, Inc., San Diego, CA; Higgins, D.G.和 Sharp, P.M.(1989)CABIOS 5:151-153; Myers, E.W.和 Muller W.(1988)CABIOS 4:11-17; Robinson, E.D.(1971)Comb. Theor 11:105; Saitou, N. Nei, M.(1987)分子生物学进展。4:406-425; Sneath, P.H.A.和 Sokal, R.R.(1973)数字分类学——数字分类学原理和应用, Freeman 出版社, San Francisco, CA; Wilbur, W.J.和 Lipman, D.J.(1983)美国国家科学院院报(Proc. Natl. Acad., Sci. USA)80:726-730。

另外, 最佳的同源性序列比较可以运用 Smith 和 Waterman 的局部同一性算法(1981), 加法应用数学 2:4 82, 运用 Needleman 和 Wunsch 的同源性比较算法(1970), 分子生物杂志 48:443, 运用 Pearson 和 Lipman 的相似性方法检索(1988)美国国家科学院院报 85:2444, 运用计算机执行这些算法(Wisconsin 遗传学软件包中的 GAP, BESTFIT, BLAST, FASTA 和 TFASTA, 遗传学计算小组(GCG), 575 Science Dr., Madison, WI), 或通过检查的方法进行比对。

算法的一个优选实施例, 适于判定序列的同一性和相似性的比例是 BLAST 和 BLAST 2.0 算法, 分别见 Altschul 等(1977)核酸研究, 25:3389-3402 和 Altschul 等(1990)分子生物学杂志 215:403-410。例如, BLAST 和 BLAST 2.0 适用于此处所描述的参数, 以判定本发明多核苷酸和多肽序列同一性的比例。国家生物技术信息中心公开提供 BLAST 分析软件。对于氨基酸序列, 使用得分矩阵计算累计分数。当碰到下列情况时, 各方向上的字击打的延伸停止: 通过 X 累计比对得分从它的最大值下降; 由于累加一个或多个负分的残基比对值, 累加得分为零或更低; 或者抵达序列的末端。BLAST 算法参数 W, T 和 X 确定比较序列的敏感性和速度。

在一个优选的方法中, 通过在一个至少 20 个位点的对比窗中比较两个理想的对比序列来确定“序列同一性百分数”, 其中相对用于两条最佳对比序列的参考序列(不含碱基添加或缺失), 比对窗中的多

肽序列部分可以含有 20%或更少，通常 5-15%，或 10-12%的碱基添加或缺失。通过确定相同氨基酸残基同时出现在两条序列中产生的匹配位点数目计算百分数，即用参考序列中的位点总数(即，对比窗大小) 除匹配位点数，再将结果乘以 100 得出序列同一性的百分数。

5

在另一个说明性实施方案中，根据本发明产生的和/或采用的多肽可以是异种多肽，它与作为参考序列的人类多肽(也称为自体同源抗原)基本上具有如上所述的序列同一性，但是来源于不同的、非人类的物种。本领域专业技术人员会发现“自身”抗原并非好的 CD8+和 CD4+T 淋巴细胞应答反应刺激物，因此有效的针对肿瘤多肽的免疫治疗策略需要研发一种能够克服特定自身肿瘤多肽免疫耐受性的方法。例如，用异种(非人类的)前列腺蛋白(Prostate Protein)免疫的人能够产生抗相应的人类蛋白的免疫反应，例如存在于人肿瘤细胞的人前列腺(Prostate)肿瘤蛋白。因此，本发明的一个方面还提供了本文描述的蛋白和/或多肽的异种变体。

10

15

更特别的，本发明中用作上文提及的人类多肽的异种形式是指小鼠、大鼠、猴、猪和其它非人类多肽。

20

在其它说明性实施方案中，本发明采用和/或产生的融合多肽包括如本文所描述的多种多肽和/或多肽亚基，或包括至少一种本文描述的多肽和其它不相关的序列。融合对象，例如，可以帮助提供 T 辅助表位(一种免疫学融合对象)，优选可被人类识别的 T 辅助表位，或帮助表达蛋白(表达增强子)使产量高于天然重组蛋白。某些优选的融合对象既是免疫学也是表达增强的融合对象。也可以选用其它融合对象提高多肽可溶性，或使多肽可以靶向目的胞内小室。而且融合对象还可以包括亲和标记物，有助于多肽的纯化。

25

30

通常使用常规技术制备融合多肽，包括化学偶联。优选的，采用本发明的组合体和方法表达出重组多肽形式的融合多肽，并使表达系

5 统的产量提高。简言之，例如编码多肽组分的 DNA 序列可以分别组装，并连接入适当的表达载体。编码一个多肽组分的 DNA 序列 3'端，通过肽接头或者直接与编码另一个多肽组分的 DNA 序列 5'端连接以保持序列的读码框一致。这样可以翻译出一条融合多肽并保持两个多肽组分的生物学活性。

10 使用多肽接头序列可以使第一和第二两个多肽组分分开足够的距离以确保每个组分能够折叠形成其二级和三级结构。这种多肽接头序列可以通过本领域熟知的常规技术整合到融合多肽中。可根据下述因素选择适当的肽接头序列：(1)具有柔韧性良好的构型；(2)不会产生二级结构与第一和第二多肽组分上的功能表位相互作用；及(3)没有可能会与多肽功能表位反应的疏水性或带电残基。优选的肽接头序列包括甘氨酸，天冬氨酸和丝氨酸残基。其它接近中性的氨基酸，例如苏氨酸和丙氨酸也可用于接头序列。可以用作接头的氨基酸序列包括
15 Maratea 等在《基因》40:39-46，1985 中；Murphy 等在《美国科学院院报》83:8258-8262，1986 中；美国专利号 4,935,233 和美国专利号 4,751,180 中描述的。接头序列长度通常为 1 至约 50 个氨基酸。当第一和第二多肽的非必需 N-端氨基酸区可以用于分离功能结构域和防止空间位阻影响时不需要接头序列。

20

连在一起的 DNA 序列可以与适当的转录或翻译调控元件可操作连接。负责 DNA 表达的调控元件只能位于编码第一多肽的 DNA 序列的 5'端。同样，用于终止翻译的终止密码子和转录终止信号只能位于编码第二多肽的 3'端。

25

融合多肽可以包括本发明中生产和/或描述的多肽和不相关的蛋白，例如能够引发记忆反应的免疫原性蛋白。这种蛋白的实例有破伤风、肺结核和肝炎蛋白(参照，例如，Stoute 等，新英格兰医学杂志(New Engl. J. Med.)，336:86-91，1997)。

30

在一个优选实施例中，免疫学融合对象衍生于分支杆菌属，例如结核分支杆菌衍生的 Ra12 片段。Ra12 用于增强表达和/或增强异源多核苷酸/多肽序列的免疫原性的组合体和方法在美国专利申请 60/158,585 中有说明，其公开内容以全文引入本文作为参考。简言之，

5 Ra12 是一段多核苷酸序列，是结核分支杆菌 MTB32A 核酸的亚序列。MTB32A 是结核分支杆菌毒株和无毒株的基因编码的分子量为 32kD 的丝氨酸蛋白酶。MTB32A 的核苷酸序列和氨基酸序列已有说明(例如，美国专利申请 60/158,585；或参照，Skeiky 等，感染和免疫 (1999)67:3998-4007，本文引作参考)。MTB32A 的 C-末端片段编码序列能够高水平表达，纯化全过程维持为可溶性多肽。而且，Ra12 可以增强与其融合的异源免疫原性多肽的免疫原性。一种优选的 Ra12 融合多肽包括对应于 MTB32A 的第 192-323 位氨基酸残基的 14KD C-端片段。其它优选的 Ra12 多核苷酸通常包括连续至少约 15 个，至少约 30 个，至少约 60 个，至少约 100 个，至少约 200 个，或至少约 300

10 个编码 Ra12 多肽部分的核苷酸。Ra12 多核苷酸可以包括天然序列(即编码 Ra12 多肽或其部分的内源序列)或包括这种序列的变体。Ra12 多核苷酸变体可包括一个或多个替换，碱基添加，缺失和/或插入，与包含天然 Ra12 多肽的融合多肽相比，这种编码的融合多肽生物学活性基本上没有下降。与编码天然 Ra12 多肽或其部分的多核苷酸序列相比，变体优选具有至少约 70%同一性，更优选至少约 80%同一性，最优选的至少约 90%同一性。

15

20

其它优选实施方案中，免疫学融合对象衍生于蛋白 D，它是革兰氏阴性细菌流感嗜血菌 B 的表面蛋白(WO 91/18926)。优选的，蛋白 D 衍生物包括蛋白的大约三分之一(例如，N-端前 100-110 个氨基酸)，蛋白 D 衍生物可以被脂质化。在某些优选的实施方案中，脂蛋白 D 融合对象的前 109 个残基定位于 N-端可以使多肽额外具有外源性 T-细胞表位并提高在大肠杆菌中的表达水平(从而功能相当于表达增强子)。尾部的脂类确保更好的将抗原递呈于抗原递呈细胞。其它融合对象包括流感病毒的非结构蛋白，NS1(红血球凝集素)。尽管可以使用

25

30

包括 T-辅助表位的不同片段，但是通常使用 N-端 81 个氨基酸。

5 另一实施方案中，免疫学融合对象是已知的 LYTA 蛋白，或其部分（优选 C-端部分）。LYTA 衍生于肺炎链球菌 (*Streptococcus pneumoniae*)，其合成的 N-乙酰化-L-丙氨酸酰胺酶即是已知的酰胺酶 LYTA(LytA 基因编码；基因 43:265-292，1986)。LYTA 是自溶素，特异性降解肽聚糖骨架中特定的键。LYTA 蛋白 C-端结构域负责与胆碱或胆碱类似物如 DEAE 的亲合反应。此特性可以用于能够表达融合蛋白的大肠杆菌 C-LYTA 表达质粒的开发。氨基端含有 C-LYTA 片段的杂合蛋白的纯化已有描述(参照生物技术 10:795-798，1992)。在一个
10 优选实施方案中，融合多肽中可整入 LYTA 的重复部分。重复部分位于 C-末端，起始于第 178 位残基。特别优选的重复部分整入第 188-305 位残基。

15 然而另一个说明性实施方案涉及融合多肽，和其编码的多核苷酸，其中融合对象包括导向信号能够将多肽定位于内吞小泡/溶酶体小泡中，如美国专利号 5,633,234 中所述。本发明的免疫原性多肽与这种导向信号融合能够更有效的与 MHC II 类分子结合，因此能够增强体内这种多肽特异的 CD4+T 细胞刺激。

20 通常，本发明的蛋白和/多肽(包括融合多肽)是分离的。“分离的”多肽是指与其原始环境分离。例如，天然存在的蛋白或多肽如果与一些或所有的在天然环境下和其共存的物质分离，就是分离的。优选的这些多肽也经过纯化，例如，至少有约 90%的纯度，更优选的至少约
25 95%的纯度，最优选的至少约 99%的纯度。

30 用本发明中的方法生产的特别优选的多肽包括结合剂，例如抗体和其抗原结合片段，能够与目的靶多肽发生免疫结合，例如与特定病态关联的多肽，或其部分、变体或衍生物。抗体或其抗原结合片段被认为能够与本发明的多肽“特异性结合”，“免疫结合”，和/或对其

具有“免疫反应性”，条件是其与多肽在可检测的水平(在如 ELISA 反应中)发生反应，并且不能检测到与不相关的多肽在相似条件下的反应。

5 本文中使用的免疫结合，通常是指在免疫球蛋白分子和其特异的抗原之间发生的非共价反应类型。免疫结合反应的强度，或亲和力可以用反应的解离常数(k_d)这个术语表示，其中 k_d 越小表示亲和力越大。选用多肽的免疫结合特性能够通过本领域熟知的方法定量。一种方法必须测定抗原结合位点/抗原复合物形成和解离的速率，其中速率依赖于复合物对象的浓度，相互作用的亲和力，和均等影响两个方向速率的几何参数。这样，“结合速度常数”(K_{on})和“解离速度常数”(K_{off})可以通过计算浓度和实际的结合和解离速度确定。 K_{off}/K_{on} 的比率能够消除所有与亲和力无关的参数，所以与解离常数 K_d 相等。通常参照，
10 Davies 等(1990)生物化学年报 59:439-473。

15 抗体的“抗原结合位点”，或“结合部分”是指免疫球蛋白分子参与抗原结合的部分。抗原结合位点是由重链(“H”)和轻链(“L”)N-末端可变区(“V”)的氨基酸残基形成。重链和轻链 V 区的三个高度多样化区段是指“高变区”，两侧为更保守的区段称为“框架区”，
20 或“FR”。术语“FR”是指天然存在于免疫球蛋白分子的高变区之间或与其临近的氨基酸序列。抗体分子中，轻链的三个高变区和重链的三个高变区在三维空间中处于相互对应的位置能够形成抗原结合表面。抗原结合表面与结合抗原的三维空间表面互补，重链和轻链各自的三个高变区也称作“互补决定区”，或“CDR”。

25 使用本文提供和本领域已知的具有代表性的分析方法，例如那些特异于肿瘤相关蛋白的某些结合剂，还能够用于区分患有肿瘤和未患肿瘤的病人。例如，与肿瘤蛋白结合的抗体或结合剂优选产生一种信号，说明至少约 20%的病人患有肿瘤，更优选至少约 30%病人。或者，
30 或另外，抗体可以产生负信号说明至少约 90%的个体没有肿瘤。为了

测定结合剂是否满足这种要求，可以用本文描述的方法对患有肿瘤和未患肿瘤的病人(用常规临床检测法确定)的生物样品(例如，血液，血清，痰液，尿和/或肿瘤活检)加以检测从而确定是否存在与结合剂结合的多肽。优选的，分析的患有癌症和未患癌症的样品数目具有统计学意义。每种结合剂都应符合上述标准；然而本领域普通技术人员应该认识到结合剂可以联合使用提高灵敏度。其它根据本发明生产的结合剂亦可能由于其对肿瘤相关多肽序列的特异性而具有治疗价值。

任何满足上述要求的作用物都可以是结合剂。例如，结合剂可以是核酶，含有或不含多肽组分，RNA 分子或多肽。在一个优选实施方案中，结合剂是抗体或其抗原结合片段。可以用本领域普通技术人员熟知的各种技术制备抗体。例如，参照，Harlow 和 Lane，抗体：实验室手册，冷泉港实验室，1988。除了本文作为范例所述的本发明方法，专业人员还可获得大量的抗体制备技术。例如，采用细胞培养技术制备抗体，包括生成本文所述的单克隆抗体，或将抗体基因转染到适当的细菌或哺乳细胞宿主中，以产生重组抗体。在一种技术中，首先将含有多肽的免疫原注射到多种多样的任一哺乳动物(例如，小鼠，大鼠，兔子，绵羊或山羊)中。在这一步中，本发明的多肽可以用作不加修饰的免疫原。或者，特别对于较短的多肽，如果多肽与载体蛋白，例如牛血清蛋白或匙孔血蓝蛋白相连，可以引发很强的免疫反应。免疫原注入宿主动物中，优选根据预先制定的计划引入一次或多次加强免疫接种，并分时段给动物取血。针对多肽的特异性多克隆抗体可以再进行纯化从抗血清中分离，通过例如，用偶联于适当的固相支持物上的多肽进行亲和层析。

25

制备针对目的抗原多肽的特异性单克隆抗体可以通过，例如 Kohler 和 Milstein 在《欧洲免疫学杂志》6:511-519，1976 中描述的技术或其改良技术进行。简言之，这些方法包括能够产生具有所需特异性的抗体(即与目的多肽反应)的无限增殖化细胞系的制备。这些细胞系可以产生于例如，按上述方法免疫动物的脾细胞。脾细胞接着被无

30

5 限增殖化，通过例如与骨髓瘤细胞融合对象融合，优选是一种免疫动物的同源细胞。可以采用各种融合技术。例如，用非离子去垢剂混合脾细胞和骨髓瘤细胞数分钟，再以较低浓度涂布于支持杂交细胞生长而不支持骨髓瘤细胞生长的选择性培养基上。优选筛选技术使用 HAT(次黄嘌呤，氨基蝶呤，胸苷)筛选。经过充足的时间，通常 1-2 周，可以观察到杂交株。挑选单个克隆，测定其培养上清液对多肽的结合活性。优选具有高活性和特异性的杂交瘤。

10 可以从培养杂交瘤克隆的上清液中分离出单克隆抗体。而且可采用各种技术提高产量，例如将杂交瘤细胞系注入适当的脊椎动物，如小鼠的宿主腹腔中。然后从腹水或血液中收获单克隆抗体。通过常规技术去除抗体中的污染物，例如色谱，凝胶过滤，沉淀，和抽提。本发明的多肽可以用于纯化过程，例如亲和层析步骤中。

15 本领域已知的许多治疗用分子含有抗原结合位点，能显示抗体分子的免疫结合特性。木瓜蛋白酶优选切割 IgG 分子产生若干片段，其中两个片段(“F(ab)”片段)各含有包括完整抗原结合位点的共价异二聚体。胃蛋白酶能切割 IgG 分子提供若干个片段，包括含有两个抗原结合位点的“F(ab)₂”片段。可以通过优选的蛋白水解酶切割 IgM
20 产生“Fv”片段，少数情况下切割 IgG 或 IgA 免疫球蛋白分子。然而，使用本领域熟知的重组技术生产 Fv 片段更常见。Fv 片段含有包括了抗原结合位点的非共价 V_H::V_L 异二聚体，此位点很大程度保留了天然抗体的抗原识别和结合能力。Inbar 等(1972)美国科学院院报 69:2659-2662; Hochman 等(1976)生物化学 15:2706-2710; 及 Ehrlich 等(1980)
25 生物化学 19:4091-4096。

30 单链 Fv(“sFv”)多肽是自通过多肽编码接头连接的包括 V_H 和 V_L 编码基因的融合基因表达而成的共价连接的 V_H::V_L 异二聚体。Huston 等(1988)美国科学院院报 85(16):5879-5883。已描述了许多分辨化学结构的方法可以将天然聚合的——但是化学分离的——抗体 V 区的重链

和轻链转变成 sFv 分子，它可以折叠成与抗原结合位点基本相似的三维结构。参照，例如授权于 Huston 等人的美国专利号 5,091,513 和 5,132,405 以及授权于 Ladner 等的美国专利号 4,946,778。

5 上述每种分子均包括重链和轻链 CDR 区，分别位于重链和轻链的 FR 区之间，FR 区对 CDR 区有支持作用而且界定了 CDR 相互之间的空间关系。如本文使用的术语“CDR 区”指重链或轻链 V 区中的高变区。从重链或轻链的 N-端起，这些区段分别命名为“CDR1”，“CDR2”，和“CDR3”。所以抗原结合位点包括六个 CDR，包括重链和轻链各自 V 区的 CDR 区。含有一个 CDR(例如 CDR1，CDR2 或
10 CDR3)的多肽在本文中是指“分子识别单位”。晶体学分析了大量的抗原-抗体复合物，结果显示 CDR 的氨基酸残基与结合的抗原形成了广泛的接触，其中接触最多的是重链的 CDR3。所以，分子识别单位是主要负责抗原结合位点的特异性。

15

 如本文使用的术语“FR 区”指重链或轻链 V 区的 CDR 区中构成 CDR 两侧的四条氨基酸序列。某些 FR 残基可与结合抗原接触；然而，FR 主要负责 V 区的折叠以形成抗原结合位点，特别是直接临近 CDR 的 FR 残基。FR 内一些氨基酸残基和结构特征是高度保守的。在这方
20 面，所有的 V 区序列含有大约 90 个氨基酸残基的内部二硫环。当 V 区折叠成结合位点，CDR 显示出突出的环基序形成抗原结合表面。通常认为影响 CDR 环折叠成某些“规范”结构的形状的是 FR 的保守结构区而不是 CDR 精确的氨基酸序列。而且，已知一些 FR 残基参与了结构域间的非共价接触稳定了抗体重链和轻链间的相互作用。

25

 已描述了大量“人源化的”抗体分子，它们含有衍生于非人的免疫球蛋白分子的抗原结合位点，包括嵌合抗体，它们含有啮齿类的 V 区及相连的 CDR 并与人的恒定区融合(Winter 等(1991)自然 349:293-299; Lobuglio 等(1989)美国科学院院报 86:4220-4224; Shaw 等(1987)
30 免疫学杂志 138:4534-4538; Brown 等(1987)癌症研究 47:3577-3583)，

啮齿类 CDR 在与适当的人抗体恒定区融合之前嫁接到人的有支持作用的 FR 上(Riechmann 等(1988)自然 332:323-327; Verhoeyen 等(1988)科学 239:1534-1536; Jones 等(1986)自然 321:522-525), 及啮齿类 CDR 由重组修饰的啮齿类 FR 支持(欧洲专利公开号 519,596, 出版日期 1992 年 12 月 23 日)。这些“人源化的”分子是为了减弱不必要的针对啮齿类抗人的抗体分子的免疫反应而设计的, 这种反应限制了这些部分在人类受体治疗应用时的持续时间和效果。

如本文中使用的术语“修饰的 FR”和“重组修饰的 FR”指选择性的替换 FR 残基, 例如, 用人 FR 残基替换啮齿类重链或轻链 V 区中的 FR 残基形成具有抗原结合位点的异源分子, 该位点基本上保留了所有天然 FR 多肽折叠结构。修饰技术的根据是抗原结合位点的配体结合特性主要由其结构和抗原结合表面中重链和轻链的 CDR 区的相对位置决定。Davis 等(1990)生物化学年报 59:439-473。所以抗原结合特性可以在人源化抗体中保留只要 CDR 结构、它们相互之间的相互作用及与其它 V 区结构域的相互作用能小心地得以维持即可。使用修饰技术, 容易为免疫系统识别的外部(如溶剂可接近的)FR 残基可以选择使用人类的残基替换, 形成具有较弱免疫原性或基本上无免疫原性的表面经修饰的杂合分子。

修饰过程利用可以得到的人类抗体可变区序列数据, Kabat 等汇编, 免疫学所需蛋白序列(Sequences of Proteins of Immunological Interest), 第 4 版(美国卫生与人类服务部, 美国政府印刷局, 1987), Kabat 更新数据库, 和其它可以得到的美国和外国数据库(核酸和蛋白)。V 区与溶剂可接触的氨基酸可以从已知的人和鼠抗体片段的三维结构推测获得。修饰鼠的抗原结合位点通常有两步。首先, 比较目的抗体分子可变区 FR 与对应的来自上述确定数据的人可变结构域 FR 序列。然后逐个氨基酸残基对比同源性最高的人 V 区与相应的鼠氨基酸。鼠 FR 中不同于人相应部分的氨基酸残基用本领域熟知的技术替换为人类部分中存在的残基。残基替换只能选用至少部分暴露的部分

进行(溶剂可以接触的),而且氨基酸替换时必须小心,因其会对V区结构域的三级结构产生显著影响,例如脯氨酸,甘氨酸和带电氨基酸。

5 这样,获得“修饰的”鼠抗原结合位点的设计要保留鼠CDR残基,与CDR基本上邻近的残基,经鉴定被包埋或大部分被包埋的残基(溶剂不能接触的),认为参与重链和轻链间的非共价(例如静电作用和疏水作用)相互作用的残基,以及认为影响CDR环的“规范”三级结构的FR保守结构区的残基。再根据这些设计标准制备连接了鼠重链和轻链CDR及人源FR的重组核苷酸序列,此序列可以转染哺乳动物细胞表达具有鼠抗体分子抗原特异性的重组人抗体。

10

本发明另一个实施方案中,根据本发明生产的抗体可以与一个或多个治疗剂偶联。这方面适用的试剂包括放射性核素,分化诱导物,药物,毒素,和它们的衍生物。优选的放射性核素包括 ^{90}Y , ^{123}I , ^{125}I , ^{131}I , ^{186}Re , ^{188}Re , ^{211}At , 和 ^{212}Bi 。优选的药物包括氨甲蝶呤,嘧啶和嘌呤类似物。优选的分化诱导物包括佛波酯和丁酸。优选的毒素包括蓖麻毒素,相思豆毒素,白喉毒素,霍乱毒素,核糖体失活蛋白(gelonin),假单孢菌外毒素,痢疾毒素,和美洲商陆抗病毒蛋白。

15

20 治疗剂可以直接或间接(例如通过接头基团)与适当的单克隆抗体偶联(例如共价连接)。如果试剂和抗体各具有能彼此相互作用的取代基,两者可以直接反应。例如,一种分子上的亲核基团,如氨基或巯基基团,能够与另一分子上的含羰基的基团,例如酮或酰卤,或含容易取代的基团(例如卤化物)的烷基反应。

25 或者,适合通过接头基团偶联治疗剂和抗体。接头基团用作抗体和试剂间的间隔子能将两者分开以避免对结合能力的影响。接头基团也可以用于提高在试剂或抗体上取代基的化学活性,从而提高偶联效率。化学活性的提高也便于使用试剂,或试剂上的官能团,否则不可能使用它们。

30

5 使用各种双官能团或多官能团试剂，功能相同和功能不同的(例如 Pierce 化学公司 Rockford, IL 目录中描述的那些)作为接头基团对本领域熟练技术人员是显而易见的。偶联可能会受到例如，氨基基团，羧基基团，巯基基团或氧化的碳水化合物残基的影响。有很多参考文献描述了这些方法，例如，授权于 Rodwell 等人的美国专利号 4,671,958。

10 如果本发明的免疫偶联物的治疗剂脱离抗体会更有使用价值，则使用能够在进入细胞的过程中或进入细胞后发生切割的接头基团更为理想。大量可被切割的不同接头基团已有描述。胞内从接头基团上释放偶联剂的机理包括还原二硫键的切割(例如，授权于 Spitler 的美国专利号 4,489,710)，照射光敏键(例如，授权于 Kohn 等的美国专利号 4,638,045)，衍生化氨基酸侧链的水解(例如，授权于 Kohn 等人的美国专利号 4,638,045)，血清补体介导的水解(例如，授权于 Rodwell 等的美国专利号 4,671,958)，以及酸催化水解(例如，授权于 Blattler 等的美国专利号 4,569,789)。

适合蛋白和/或多肽表达的多核苷酸

20 另一方面本发明提供了编码上文描述的重组蛋白和/或多肽的多核苷酸。术语“DNA”和“多核苷酸”在本文交换使用，指从特定物种总基因组 DNA 中分离出的游离 DNA 分子。本文用的“分离”意思是多核苷酸基本上与其它编码序列分离，而且该 DNA 分子不含大部分不相关的编码 DNA，例如大的染色体片段或其它功能基因或多肽编码区。当然，所指的 DNA 分子是分离出的原始片段，不排除后来人工添加到这些区段上的基因或编码区。

30 多核苷酸可以包括原始序列(即编码蛋白和/或多肽的内源序列，例如抗体，或其部分)或包括编码变体或衍生物的序列，优选这种序列的免疫原性变体或衍生物。在一些实施方案中，多核苷酸序列可以编

码如上文所述的免疫原性多肽。

5 通常，多核苷酸变体含有一个或多个替换，添加，缺失和/或插入，优选由变体多核苷酸编码的多肽的免疫原性与本文特意提及的多核苷酸序列编码的多肽相比基本上没有下降。对术语“变体”的理解应包括异种来源的同源基因。

10 本发明的多核苷酸或其片段，无论编码序列本身的长度，可以与其它 DNA 序列，例如启动子，多聚腺苷酸信号，附加内切酶位点，多克隆位点，其它编码区段等等联合使用，这样它们总长度差异较大。因此预期可以使用各种长度的核酸片段，优选总长度受制备的难易度和要使用的重组 DNA 技术限制。例如，作为范例的多核苷酸区段总长度约为 10,000，5000，3000，2000，1000，500，100，50 个碱基对等，(包括所有中间产物的长度)预期在本发明的许多方面是有用的。

15

根据本发明适用于高表达、大规模生产的多核苷酸可以使用许多已有的技术鉴定，制备和/或操作(参照，Sambrook 等，分子克隆：实验室手册，冷泉港实验室，冷泉港，纽约，1989，以及其它参考文献)。例如，通过 cDNA 微阵列筛选肿瘤相关表达的方法可鉴定多核苷酸。
20 这种筛选可以根据制造商的说明书(并基本上如 Schena 等，美国科学院院报 93:10614-10619，1996 及 Heller 等，美国科学院院报 94:2150-2155，1997)所述采用，例如 Affymetrix 公司的微阵列技术(Santa Clara，加利福尼亚)来进行。或者，可从表达本文描述的蛋白的细胞，例如肿瘤细胞，制备的 cDNA 扩增多核苷酸。

25

可以使用许多依赖模板的方法扩增样品中存在的目的靶序列。最著名的扩增方法是多聚酶链式反应(PCRTM)，在美国专利号 4,683,196，4,683,202 和 4,800,159 中有详述，分别以其全文在此引作参考。简言之，PCRTM中，制备两条与靶序列反相互补链的区域互补的引物序列。
30 将过量的脱氧三磷酸核苷与 DNA 聚合酶(例如 Taq 聚合酶)一起加入到

反应混合物中。如果样品中存在靶序列，引物就会与靶结合，聚合酶通过添加核苷酸使引物沿靶序列延伸。升高和降低反应混合物的温度，延伸的引物会与靶解离形成反应产物，多余的引物与靶及反应产物结合，延伸过程得以重复。采用优选的反转录和 PCRTM 扩增程序可
5 定量扩增的 mRNA 数量。聚合酶链式反应方法是本领域熟知的。

其他众多依赖模板的任何方法，其中许多都是 PCRTM 扩增技术的变化形式，为本领域熟知并可以获得的。作为范例的一些方法包括连接酶链式反应(称作 LCR)，在例如，欧洲专利申请公开号 320,308 和
10 美国专利号 4,883,750 中有描述；Qbeta 复制酶，在 PCT 国际专利申请公开号 PCT/US87/00880 中有描述；链替换扩增(SDA)和修复链式反应(RCR)。还有其它扩增方法在大英国专利申请号 2 202 328，和 PCT 国际专利申请公开号 PCT/US89/01025 中有描述。其它核酸扩增程序包括以转录为基础的扩增系统(TAS)(PCT 国际专利申请公开号
15 WO88/10315)，包括基于核酸序列的扩增(NASBA)和 3SR。欧洲专利申请公开号 329,822 描述了包括循环合成单链 RNA(“ssRNA”)，ssDNA 和双链 DNA(dsDNA)的核酸扩增程序。PCT 国际专利申请公开号 WO89/06700 描述了一种基于启动子/引物与靶单链 DNA(“ssDNA”)杂
交接着转录出该序列的许多 RNA 拷贝的核酸扩增方法。其它扩增方
20 法如“RACE”(Frohman, 1990)，和“一侧 PCR(one-sided PCR)”(Ohara, 1989)也是本领域熟练技术人员熟知的。

可以采用熟知的技术，用本发明多核苷酸的扩增部分从适当的文库(例如，肿瘤 cDNA 文库)中分离出全长基因。这些技术中，用适合于扩增的一种或多种多核苷酸探针或引物筛选文库(cDNA 或基因组)。
25 优选的，文库的大小经过筛选并包括较大的分子。随机引物文库优选用于鉴定基因的 5'端和上游区。基因组文库优选用于获得内含子和延长的 5'序列。或者，或另外，根据本发明在常规的亚克隆技术中基本上可使用任何扩增的多核苷酸以获得基于 UCOE 的载体。

30

对杂交技术而言，可以使用众所周知的技术标记部分序列(例如，使用 ^{32}P 进行缺刻翻译或末端标记)。通常用标记探针和含有变性细菌克隆(或长有噬菌斑的菌苔)的膜杂交筛选细菌或噬菌体文库(参照，Sambrook 等，分子克隆：实验室手册，冷泉港实验室，冷泉港，纽约，1989)。筛选并扩增杂交克隆或噬菌斑，分离 DNA 作进一步分析。例如，使用一条与部分序列相同的引物和一条与载体序列相同的引物进行 PCR，可以分析 cDNA 克隆从而确定附加序列的量。制作限制性内切酶图谱和部分序列鉴定一个或多个重叠克隆。通过常用技术，可能包括制备一系列缺失克隆，确定完整的序列。产生的重叠序列可以再组装成一条连续的序列。使用已知的技术，可以通过连接适当的片段获得全长 cDNA 分子。

或者，如上所述的扩增技术可用来从部分 cDNA 序列获得全长编码序列。一种这样的扩增技术是反向 PCR(参照，Triglia 等，核酸研究，16:8186，1988)，它采用限制性内切酶在基因的已知区域产生片段。再通过分子内连接环化该片段并以其为模板用已知区域的趋异引物进行 PCR。在一个替代方法中，部分序列的邻近序列的恢复可以通过使用一条接头序列的引物和一条特异于已知区域的引物扩增进行。扩增得到的序列通常在使用同一条接头引物和另一条特异于已知区域的引物进行第二轮扩增。这一方法的不同之处在 WO96/38591 中有描述，它采用两条引物从已知序列的不同方向进行初始延伸。另一技术称为“cDNA 末端快速扩增”或 RACE。这一技术要使用一条内引物和一条外引物，后者与多聚腺苷酸区或载体序列杂交，鉴定已知序列的 5'和 3'端。其它技术包括捕获 PCR(Lagerstrom 等，PCR 方法应用 1:111-19，1991)和行走 PCR(Parker 等，核酸研究 19:3055-60，1991)。也可使用其它扩增方法获得全长 cDNA 序列。

在一些情形下，可以通过分析表达序列标签(EST)数据库中提供的序列来获得全长 cDNA 序列，所述数据库诸如从 GenBank 中获得的数据库。寻找重叠 EST 通常可以通过使用众所周知的程序(例如 NCBI

BLAST 搜索)进行, 可以使用这些 EST 产生连续的全长序列。也可以通过分析基因组片段获得全长 DNA 序列。

5 本发明一些优选的实施方案中, 多核苷酸序列或其片段可以用于基于 UCOE 的载体的构建和/或使用中, 并编码一条或多条目的多肽, 例如抗体或融合蛋白或其功能等价物。由于遗传密码本身的简关性, 会产生其它基本上编码相同或功能相当的氨基酸序列的 DNA 序列, 这些序列可用于克隆和表达给定蛋白。

10 如本领域技术人员理解的, 在一些情形下产生具有非天然存在的密码子的多肽编码核苷酸序列是有利的。例如, 可以选用特定的原核或真核宿主优选的密码子提高蛋白表达速率或产生具有所需特性的重组 RNA 转录本, 例如半衰期比天然序列生成的转录本更长。

15 而且, 本发明的多核苷酸序列能够通过本领域广泛熟知的方法加以改造, 以适应多种需求改变多肽编码序列, 包括但不限于, 修饰克隆, 加工, 和/或基因产物的表达而造成的改变。例如, 可以采用随机片段化重组 DNA, PCR 重新组装基因片段及合成寡核苷酸来改造核苷酸序列。另外, 可以采用定点突变插入新的内切酶位点, 改变糖
20 基化模式, 改变密码子嗜好性, 形成剪切变体或引入突变等等。

新合成的多肽应基本上纯化, 例如, 采用制备高效液相色谱(例如, Creighton, T.(1983)蛋白, 结构和分子原理, WH Freeman 和 Co., 纽约市, 纽约州)或其它相应的本领域可以获得的技术。合成多肽的组
25 成可以用氨基酸分析或测序的方法进行鉴定(例如, Edman 降解法)。另外, 多肽或其部分的氨基酸序列可以经直接合成和/或用化学法与其它蛋白或其部分的序列组合而得到改变, 产生变体多肽。

提供下述实施例进行说明, 但本发明不局限于此。

30

实施例

实施例 1

用基于 UCOE 的载体系统表达重组抗体

5 本实施例比较了使用含有或不含有 UCOE 的载体表达重组抗体的水平。

10 用如图 1 所示的含有人 RNP UCOE 的载体表达改造的人抗体 Ab3。还构建了相同的但不含 UCOE 元件的载体。本实施例中 Ig 重链编码序列含有改造的人 V 区序列，此序列位于人 Ig γ -1 恒定区编码基因组 DNA 片段上游并保持读码框一致。Ig 轻链编码区包含改造的人 V 区序列，此序列位于人 Ig κ 恒定区编码 cDNA 片段上游并与之保持读码框一致。表达 Ig 重链的载体还含有一个 neo 筛选标记基因，表达 Ig 轻链的载体含有潮霉素筛选标记。参照图 2A。

15 根据厂商说明用 lipofectamine(Life Technodgies)对 CHO-K1 细胞进行轻链和重链载体的共转染。细胞筛选用潮霉素和 G418。培养转染的细胞群，在转染后不同时间用 ELISA 检测分泌到培养基中组装的免疫球蛋白水平。(图 3)。不含 RNP UCOE 的载体转染 48 小时抗体表达水平很低(大约 48ng/ml)，之后表达水平下降。相反，用含 RNP UCOE
20 的表达载体转染的细胞群中，随着转染细胞扩大培养抗体水平持续累积，转染后 15 天可达 3 毫克/ml。所以使用 UCOE 使得高水平表达重组免疫球蛋白的转染细胞群快速增殖。

实施例 2

25 在转染有基于 UCOE 的表达载体系统的 CHO 宿主细胞系中实现高水平大规模表达

30 用含有 UCOE 抗体表达盒的载体共转染 CHO-S 细胞(如图 1 所示)获得改造的人抗体 Ab1。Ig 重链编码序列含有改造的人 V 区序列，此序列位于人 Ig γ -4 恒定区编码基因组 DNA 片段上游并保持读码框一致。Ig 轻链编码区包含改造的人 V 区序列，此序列位于人 Ig κ 恒定区

编码 cDNA 片段上游并保持读码框一致。表达 Ig 重链的载体还含有一个 neo 筛选标记基因，表达 Ig 轻链的载体含有潮霉素筛选标记。参照图 2B。

- 5 根据厂商说明用 lipofectamine(Life Technologies)进行转染。在 CD-CHO 培养基(Life Technologies)中加入潮霉素和 G418 筛选细胞，挑选亚克隆。此过程大约需 5 周。将一个亚克隆扩大培养，在 2L 生物反应器中进行最终参数优化，再扩大到 100L 的生物反应器中。采用此方法表达重组抗体的大部分转化子典型的生产率大约为 5pg/细胞/天。细胞培养悬液中一种抗体的产量大约可达 200mg/l。参照图 4。共转染到 CHO-S 细胞中的两种表达载体含有 UCOE，它使得可以快速分离转化子克隆并立即悬浮培养于组成培养基中。
- 10

实施例 3

- 15 CHO-K1 和 CHO-S 宿主细胞系的低水平 Gal-Gal 残基

 如上文所述，含有 Gal α 1 \rightarrow 3Gal β 1 \rightarrow 4GlcNAc-R(Gal-Gal)糖残基的抗体在用于人体治疗时，已经与蛋白在血清中被迅速清除有关。所以，生产不含这种糖残基的重组蛋白是有利的。参照，例如，Borrebaeck 等，当代免疫学 14:477-479(1993)和 Kagawa 等生物化学杂志 263:17508-17515(1988)。使用 FITC 标记的 IB₄凝集素和流式细胞仪可见 CHO-S 细胞表面没有 Gal-Gal 残基。参照图 5；方法描述参见 Cho 等，生物化学杂志 272: 13622-13628(1997)和 Gorelik 等，癌症研究 55:4185-4173(1995)。在这方面，CHO-S 与另一种广泛使用并经测试的 CHO 细胞系，CHO-K1 相似。相反，小鼠杂交瘤细胞系在此试验测试中显示细胞表面有高水平的 Gal-Gal 糖基。上述细胞系表达的重组蛋白纯化后质谱分析显示无 Gal-Gal 糖基(数据未显示)。

20

25

实施例 4

改良的多亚基重组蛋白表达水平的双向 UCOE 载体

- 30 此实施例公开了双向 UCOE 载体系统上重组抗体重链和轻链蛋白

表达水平的提高。

引入接头分子将 pORT1(Cobra Therapeutics)的两个 Sfi I 位点变为 Mfe I 位点，接头分子含有退火的寡核苷酸 Mfe.F，5'-AACAATTGGC
5 GGC(SEQ ID NO:10)和 Mfe.R，5'-GCCAATTGTTGCC(SEQ ID
NO:11)。使用引物 TK.F，5'-ACGCGTCGACGGAAGGAGACAA
TACCGGAAG(SEQ ID NO:12)和 TK.R，5'-CCGCTCGAGTTGGGGTGG
GGAAAAGGAA(SEQ ID NO:13)从 pVgRXR(Invitrogen)上扩增 HSV TK
多聚腺苷酸位点，在 Sal I 位点插入 Sal I-Xho I 的片段。接着，从雄性
10 BALB/c 基因组 DNA(Clontech)扩增鼠 PKG 多聚腺苷酸位点，使用引
物 mPGK.F，5'-CGGGATCCGCCTGAGAAAGGAAGTGAGCTG(SEQ
ID NO:14)和 mPGK.R，5'-GAAGATCTG GAGGAATGAGCTGG
CCCTTA(SEQ ID NO:15)，将 BamH I-Bgl II 片段克隆到 BamH I 位点。
用 T4 DNA 聚合酶处理含有 neo 表达盒的 pcDNA3.1 的 Ase I-Sal I 片
15 段，与 Spe I 接头连接(5'-GACTAGTC; SEQ ID NO:16)连接，接着将 Spe
I 片段克隆到 Spe I 位点中产生 pORTneoF；或者用 T4 DNA 聚合酶处
理携带嘌呤霉素抗性表达盒的 CET700(Cobra Therapeutics)的 EcoR I-
Not I 片段，与 Xba I 接头连接，再将 Xba I 片段克隆到 Xba I 位点中
产生 pORTpuroF。将来自 pCMVEGFPN-1(Cobra)的 Hind III-BamH I
20 鼠 CMV 启动子片段亚克隆到 BKS+(Cobra)的杂合 UCOE 的 Hind III-
BamH I 位点。从质粒 pIRESneo(Clontech)上扩增人 CMV 启动子，使
用引物 hCMVF，5'-CTCGAGTTATTAATAGTAATCAATTACGGGG
TCAT(SEQ ID NO:17)和 hCMVR，5'-GTCGACGATCTGACGGTT
CACTAAA CCAGCTCT(SEQ ID NO:18)，将 Xho I-Sal I 片段克隆到 Sal
25 I 位点。再将 BamH I-Sal I 片段克隆到 pORTneoF 的 BamH I-Sal I 位点
产生 pBDUneo100，或者克隆到 pORTpuroF 产生 pBDUpuro300。通过
定点突变改变 BKS+中杂合 UCOE Sal I 克隆位点上游的两个 ATG 密
码子，再将 BamH I-Sal I 片段克隆到 pORTneoF 的 BamH I-Sal I 位点
产生 pBDUneo200，或者克隆到 pORpuroF 产生 pBDUpuro400。

30

将人抗体轻链克隆到四种双向 UCOE 载体 (pBDUneo100, pBDUneo200, pBDUpuro300 和 pBDUpuro400; 分别见图 6-9 及 SEQ ID NO:1-4) 的任一 BamH I 或 Sal I 位点, 接着将重链插入留下的 BamH I 或 Sal I 克隆位点产生 pBDUneo112, pBDUneo121, pBDUneo212, pBDUneo221, pBDUpuro112, pBDUpuro121, pBDUpuro212 和 pBDUpuro221。

其它适于共表达两种或多种重组蛋白的双向 UCOE 载体如图 10-13 所示 (SEQ ID NO:5-8), 分别称为 pBDUneo500, pBDUneo600, pBDUpuro700 和 pBDUpuro800。可以使用这些载体, 例如优化杂合 UCOE 表达抗体的方向, 并提供替代启动子的组合进行优化。

用 Xba I (部分) 和 NsiI 消化质粒 pORTpuroF 以去除牛生长激素多聚腺苷酸位点, 再与 SV40 早多聚腺苷酸位点相连, 此位点用引物 14506, 5'-CCAATGCATAGGTTGGGCTTCGGGAATCGT (SEQ ID NO:19) 和 14507, 5'-GCTCTAGATCTCGACGGTATAACAGACATGAT (SEQ ID NO:20) 扩增而来并经 XbaI 和 NsiI 消化, 产生质粒 pORTpuroF2。杂合 UCOE 载体含有位于人 RNP UCOE 下游的鼠 CMV 启动子和两个突变的位于肌动蛋白启动子和 Sal I 位点之间的 ATG 密码子, 用 BamHI 和 HindIII 消化此载体去除鼠 CMV 启动子, 然后与人 CMV 启动子相连, 此启动子用引物 14425, 5'-CCCAAGCTTATTAATAGTAATCAATTACGGGGTCAT (SEQ ID NO:21) 和 14426, 5'-CAAGGATCCGATCTGACGGTTCCTAAACCAGCTCT (SEQ ID NO:22) 扩增而来并经 BamHI 和 HindIII 消化。将含有退火的寡核苷酸 14466, 5'-TCGAGTCGTTTAAACTCTAG (SEQ ID NO:23) 和 14465, 5'-TCGACTAGAGTTTAAACGAC (SEQ ID NO:24) 的接头插入 SalI 位点, 用 PmeI 和 SalI 消化, 再与鼠 CMV 启动子相连, 此启动子已由引物 14435, 5'-GAATTCGAGCTCGCCCAACTCCGCCCCTTTTAT (SEQ ID NO:25) 和 14436, 5'-ATTTGTCGACTCTAGACCCGGGCTGCAGCGAGGAGCTCT (SEQ ID NO:26) 扩增而来并经 SalI 消化。用 BamHI

和 *Sa*II 消化含有或不含鼠 CMV 启动子的质粒，与 *Bam*HI 和 *Sa*II 消化的 pORTneoF 相连产生质粒 pBDUneo500 和 pBDUneo600；或与 *Bam*HI 和 *Sa*II 消化的质粒 pORTpuroF2 相连分别产生质粒 pBDUpuro700 和 pBDUpuro800。

5

根据厂商说明用 lipofectamine 或 DMRIE-C(Invitrogen)分别将表达重链和轻链的 G418 或嘌呤霉素抗性的双向 UCOE 载体转染到 CHO-K1 或 CHO-S 细胞中，用 500ug/ml G418(neo 载体)或 12.5ug/ml 嘌呤霉素(puro 载体)筛选。挑选细胞群并比较不同构建体之间的抗体生产率以确定 CHO 细胞表达抗体联合使用的优化的启动子和筛选标记。

10

CHO-S 悬浮细胞表达研究结果归纳于表 2。这些数据显示含有轻链的载体用鼠 CMV 启动子表达，获得最佳的抗体表达水平。使用的载体含有嘌呤霉素或 G418 抗性筛选标记。另外，将两种双向载体，一种含有嘌呤霉素抗性标记，一种含有 G418 抗性标记，进行共转染。挑选细胞群，确定抗体生产率。只转染 G418 或只转染嘌呤霉素抗性的转化子生产率相似，但是共转染的细胞群生产率显著增高。这说明有可能通过用不同的筛选标记保持双拷贝的抗体表达载体来提高生产率。用高浓度嘌呤霉素(25-50 μ g/ml 相对 12.5 μ g/ml)挑选出的细胞群与增高的产量不相关。

15

20

从携带 pBDUpuro421 的嘌呤霉素抗性细胞群中分离克隆细胞系。二十二个克隆细胞系中有十五个表达的抗体量能够检测到。测量的初始生产率显示细胞系分泌蛋白的速率最大为 16pg/细胞/天(表 3)。Southern 印迹法分析鉴定出至少一个克隆的生产率为 13pg/细胞/天，而且大致含有一个拷贝的载体 DNA(克隆 S421.7)。从细胞群中分离出生产率为 3-18pg/细胞/天的克隆。选用表达量约为 5pg/细胞/天的克隆进行初始发酵试验。

25

30

表 2

双向 UCOE 载体表达 hAb1(IgG4)

载体	H3 启动子	K1 启动子	生产率 (pg/细胞/天)
PBDUneo112	鼠 CMV	人 CMV	0.3
PBDUneo121	人 CMV	鼠 CMV	1.5
PBDUneo212	鼠 CMV	人 β -肌动蛋白	0.06
PBDUneo221	人 β -肌动蛋白	鼠 CMV	1.3
PBDUpuro312	鼠 CMV	人 CMV	0.5
PBDUpuro321	人 CMV	鼠 CMV	1.4
PBDUpuro412	鼠 CMV	人 β -肌动蛋白	0.05
PBDUpuro421	人 β -肌动蛋白	鼠 CMV	2.3
共转染**	人 CMV	人 CMV	0.7
PBDUneo221	人 β -肌动蛋白	鼠 CMV	1.3
PBDUpuro421	人 β -肌动蛋白	鼠 CMV	1
PBDUneo221+	人 β -肌动蛋白	鼠 CMV	5
PBDUpuro421			

** 以前使用均由 4kb UCOE CMV 载体(潮霉素和新霉素筛选)驱动
的相同抗体基因实施共转染

5

表 3

转染有 pBDUpuro421 的克隆 CHO-S 细胞系表达 hAb1

潮霉素抗性细胞系	生产率 (pg/细胞/天)
S421.2	5.4
S421.3	0.5
S421.4	0.5
S421.7	13.4
S421.8	5.4
S421.9	0.04
S421.12	1.4
S421.14	6.7
S421.15	0.3
S421.16	7.2
S421.17	5
S421.18	0.8
S421.20	1.2
S421.21	0.3
S421.22	16

实施例 5

RNP UCOE 的缺失分析

本实施例描述了在 RNP UCOE 质粒载体内缺失多核苷酸来提高重组蛋白的表达水平。简言之，在 8kb RNP UCOE 内制备一系列缺失，以鉴定在不影响 UCOE 功能的情况下可以被删除的重要的功能元件和区域。将绿色荧光蛋白基因 (GFP) 克隆到质粒 CET720(Cobra Therapeutics)，再将缺失的序列导入 UCOE 区域(图 14)。将第一套缺失序列转染 CHO-S 细胞，并检测是否可以表达 GFP。瞬时转染分析(转染两天后)中，荧光显微技术检测所有的质粒都可以表达 GFP。筛选稳定的携带不同构建体的细胞群，用 FACS 分析确定 GFP 的表达。转染后一个月，相对不含 UCOE 的对照质粒，所有的缺失都显示出更高的阳性细胞百分数(>50%对不含 UCOE 的 10%)，阳性群体的平均荧光高于不含 UCOE 的对照质粒(表 4)。

这些数据更精确地界定了完全活性所需的人 RNP UCOE 的区域，并鉴定出更短(大约 7kb)的完全活性的 UCOE 元件。通过缺失 Δ RV 形成这段新的 7kb UCOE 元件并从第 2225 位延伸至第 9254 位核苷酸，如图 14。

表 4

在 8kb RNP UCOE 内含有缺失的质粒的 GFP 表达

质粒	缺失区	阳性百分数	阳性细胞群的平均 荧光性
CET720GFP(8kb UCOE)	无	68	516
CET700GFP(无 UCOE)	核苷酸 2225-10525	10	136
Δ BS(4 kb UCOE)	核苷酸 2225-6341	61	370
Δ EcoNI	核苷酸 3875-6916	65	439
Δ EX2	核苷酸 6916-7053	53	384
Δ EM	核苷酸 6916-7209	66	423
Δ MX	核苷酸 7053-7209	66	464
Δ MluI	核苷酸 7209-8293	58	448
Δ RV	核苷酸 9254-10342	72	548

用 EcoRV, MluI, EcoNI, 或 BamHI 加上 SalI 消化载体 CET720GFP(如 SEQ ID NO:9 所示, 含有 8kb 人 RNP UCOE), 用 T4 DNA 聚合酶补平末端, 再使载体重新连接分别产生载体 δ RV, δ MluI, δ EcoNI 和 δ BS。用 PflMI 消化 CET720, T4 DNA 聚合酶补平末端, 再用 BamHI 切割。将平末端到 BamHI 粘端的片段克隆到 pBluescript II SK(+)的 EcoRV-BamHI 位点产生 pPB720。用 EcoNI 和 MluI, MluI 和 XhoI(部分), 或 EcoNI 和 XhoI(部分)消化 pPB720, 用 T4 DNA 聚合酶处理末端后再环化。将每一个这样产生的载体的 PshAI 片段克隆到 CET720GFP 的 PshAI 位点分别产生范例载体 δ EM, δ EX 和 δ MX。

实施例 6

RNP UCOE 其它缺失分析

上文所述的实施例通过缺失分析已鉴定出载体 CET720GFP(SEQ ID NO:9)的第 2225-6916 位和第 9254-10342 位核苷酸的 UCOE 区域可以被去除而不损失 UCOE 活性(参照上文实施例 5)。本实施例中, 进一步确定了最小的对活性很重要的 8kb UCOE 区域。重要的是, 此分析更精确地界定了保留全部活性的 4.1kb 的范例性人 RNP UCOE。

简言之, 补平 8kb RNP UCOE 片段, 与 HindIII 接头连接(New England Biolabs; 目录编号 S1098S), 用 HindIII 消化并与 HindIII 消化及小牛肠碱性磷酸酶处理的载体 CET700GFP 连接。用 DMRIE-C(Invitrogen)将载体转染到 CHO-S 细胞中, 其中瞬时转染分析显示所有的构建体都可以表达 GFP(数据未显示)。嘌呤霉素筛选两周后, 用 FACS 确定阳性细胞群的荧光几何平均数, 并以对照(CET720GFP)的百分数表达, 结果总结于下文表 5 中。载体 700FRV 含有 4.1 kb 的 MfeI-EcoRV 的 RNP UCOE 片段, 相应于 CET720GFP 的第 5152-9254 位核苷酸残基, 相对于 CET720 GFP 的第 2225-10525 位核苷酸残基的 8kb UCOE 区域保留了全部 UCOE 活性。所以, 此 4.1kb UCOE 片段代表一种新型最小的 UCOE 元件, 它保留了与全长 8kb UCOE 元件相

当的活性。

表 5

质粒	存在的 UCOE 区域	对照百分数
CET720GFP(8kb UCOE)	核苷酸 2225-10525	100
CET700(无 UCOE)	无	10
δ RV	核苷酸 2225-9254 核苷酸 10342-10525	99
700HRV.F	核苷酸 2240-9254	121
700FRV.F	核苷酸 5152-9254	122
700BRV.F	核苷酸 6341-9254	73

- 5 还测定了三个包含于 700HRV.R, 700FRV.R 和 700BRV.R 中的 UCOE 片段的活性, UCOE 片段以相反方向插入分别产生质粒 700HRV.F, 700FRV.F 和 700BRV.F。同样的, 瞬时转染分析中所有的质粒都可以表达 GFP。嘌呤霉素筛选三周后, 用 FACS 确定阳性细胞群的荧光几何平均数, 并以对照(CET720GFP)百分数表示, 结果总结于下文表 6 中。然而含有相反方向 UCOE 的质粒观察到的活性水平
- 10 较低, 但是所有片段仍然都保留了 UCOE 活性。

表 6

质粒	存在的 UCOE 区域	对照百分数
CET720GFP(8kb UCOE)	核苷酸 2225-10525	100
CET700(无 UCOE)	无	6
700HRV.F	核苷酸 2240-9254	59
700FRV.F	核苷酸 5152-9254	43
700BRV.F	核苷酸 6341-9254	30

15 实施例 7

制备其它范例性双向 UCOE 载体

上述实施例描述了各种范例性 UCOE 载体的制备和评估。本实施例中, 构建了其它 UCOE 载体。例如, 制备载体 pBDpuro350(SEQ ID NO:27) 和 pBDUpuro450(SEQ ID NO:28) 使与上文描述的载体 pBDpuro300 和 pBDpuro450 相当, 除了用 SV40 多聚腺苷酸位点替代

20

嘌呤霉素抗性基因之后的多聚腺苷酸位点(参照图 15 和 16)。其它几个载体中用上文缺失分析鉴定含有全部 UCOE 活性的 4.1kb MfeI-EcoRV 片段替代 8kb RNP UCOE 元件。为了改变 pBDUpuro 系列载体嘌呤霉素抗性盒的多聚腺苷酸位点,从 pBSneo.23 上用聚合酶链式反应扩增 SV40 多聚腺苷酸位点,反应产物用 NsiI 和 XbaI 消化并插入 pORTpuroF 的 NsiI-XbaI 位点以替代 BGH 多聚腺苷酸位点。再用 BamHI 和 Sall 消化这个新载体 pORTpuroF', 并克隆到 HUCMV(与鼠 CMV 启动子杂合的 UCOE)的 BamHI-Sall 位点产生质粒 pBDUpuro350(SEQ ID NO:27; 仍参照图 15), 或克隆到 pUCOEact3 的 BamHI 位点(与 ATG 密码子定点突变的肌动蛋白启动子杂合的 UCOE)产生 pBDUpuro450(SEQ ID NO:28; 参照图 16)。其它 UCOE 载体的构建是在质粒 pUCOEact3 和 pUCOEact3hCMV 的人 β -肌动蛋白和 RNP UCOE 片段之间边缘的 KpnI 位点处插入 HindIII 位点。然后去除携带 RNP UCOE 的 4kb HindIII 片段, 替换为 700FRV.R 的 4.1kb RNP UCOE 片段。将 Sall-BamHI(部分)片段克隆到 pORTneoF 和 pORTpuroF' 的 Sall-BamHI 位点产生 pBDUpuro1200(SEQ ID NO:29; 参照图 17), pBDUpuro1450(SEQ ID NO:30; 参照图 18), pBDUneo1600(SEQ ID NO:31; 参照图 19)和 pBDUpuro1800(SEQ ID NO:32; 参照图 20)。

20 实施例 8

对双向 UCOE 活性重要的载体特征评定

1. 双向 UCOE 载体拷贝数对 CHO-S 细胞中抗体生产率的影响:

CHO-S 细胞系 S421.7 显示含有单拷贝的表达 hAb1(IgG4)的载体 pBDUpuro421。为了确定增加载体拷贝数是否能提高抗体表达水平, 用也表达 hAb1 但携带不同筛选标记(G418)抗性的载体 pBDUneo221 重新转染 S421.7。分离克隆细胞并分析生产率(图 21)。许多细胞系显示出比亲代细胞系 S421.7 更高的生产率, 说明增加载体拷贝数可以提高产量。初步拷贝数分析显示细胞系 S7.16, S7.20 和 S7.23 含有 1-2 个拷贝的 pBDUneo221 载体(未显示数据)。

30

2. 杂合 UCOE 方向和启动子的选择对 CHO-S 细胞中抗体生产率的影响

分析稳定的含有表达 hAb1(IgG4)的不同的双向 UCOE 载体的 CHO-S 细胞群，以确定杂合 UCOE 相对于抗体基因的方向的影响，及不同启动子对抗体表达速率的影响。用一系列的表达 hAb1(IgG4)的双向 UCOE 载体转染 CHO-S 细胞，用 12.5 μ g/ml 嘌呤霉素或 500 μ g/ml G418 筛选出稳定的细胞群。重链(H)和轻链(K) 相对于杂合 UCOE 元件(肌动蛋白末端对 RNP 末端)和使用的启动子的定位如下文表 7 所示。用 ELISA 测定抗体生产率，western 印迹分析确定轻链和重链相对细胞裂解物在上清液中的分布。杂合 UCOE 的方向对抗体表达水平仅显示极小的作用，然而联合使用的启动子的选择对生产率具有一定的影响。这些试验中获得的最高生产率是使用人 β -肌动蛋白启动子表达重链，及使用鼠 CMV 或人 CMV 表达轻链的范例性载体(例如，pBDUpuro454 和 pBDUpuro804)。

15

表 7

载体	肌动蛋白末端	RNP 末端	重链 (上清液)	重链 (细胞裂解物)	κ 链 (上清液)	κ 链 (细胞裂解物)	生产率 (pg/细胞/天)
pBDUpuro352	hCMV-K	mCMV-H	+	++	+	-	0.159
PBDUpuro354	hCMV-H	mCMV-K	+	+	+++	+	0.256
PBDUpuro452	肌动蛋白-K	mCMV-H	+/-	++	+/-	-	0.0056
PBDUpuro454	肌动蛋白-H	mCMV-K	++	+	+++	++	0.657
PBDUpuro702	hCMV-K	mCMV-H	++	++	++	+	0.391
PBDUpuro704	hCMV-H	mCMV-K	++	++	++	+/-	0.170
PBDUpuro802	肌动蛋白-K	mCMV-H	+/-	+++	+/-	-	0.020
PBDUpuro804	肌动蛋白-H	mCMV-K	+++	+++	+++	++	0.608

3. CHO-S 细胞中的转录对生产率

从携带 pBDUpuro452, pBDUpuro454 和 pBDUpuro804 的嘌呤霉素抗性细胞群中分离出克隆细胞系。大约三分之二的携带 pBDUpuro454 和 pBDUpuro804 的克隆细胞系的抗体生产率可以测出

20

为 1-10pg/细胞/天，与上文用 pBDUpuro421 载体获得的结果相似(结果未显示)。基因组样品进行的 TaqMan 分析显示克隆细胞系 S452.3, S454.5 和 S804.4 分别携带单拷贝的双向 UCOE 载体 pBDUpuro452, pBDUpuro454 和 pBDUpuro804。细胞系 S421.7, 上文中 Southern 分析显示具有单拷贝的 pBDUpuro421(重链由人肌动蛋白启动子表达, 轻链由鼠 CMV 启动子表达的 pBDUpuro400), 用作对照。为了研究生产率和抗体链转录之间的相关性, 对这些细胞系进行 TaqMan RT-PCR 分析, 其结果总结于下表 8 中。S452.3 细胞系中的重链和轻链 RNA 水平均明显低于已显示抗体表达水平高的对照细胞系 D6 和 S421.7 中观察到的结果。然而, 细胞系 S454.5 和 S804.4 的 RNA 水平和产量与阳性对照细胞系相似。加上 western 印迹分析(结果未显示), 这些结果说明所述细胞系中观察到的抗体重链和轻链的 RNA 水平与观察到的生产率具有相关性。

表 8

细胞系	生产率 (pg/细胞/天)	轻链 (Ct)	重链 (Ct)
CHO-S	0	40	40
D6	5.5	20.39	22.86
S421.7	4.57	21.91	23.90
S454.5	3.52	22.12	23.96
S804.4	3.62	22.40	24.11
S452.3	0.07	29.62	26.47

Ct, 循环数阈值; CHO-S, 亲代细胞系; D6, 克隆细胞系携带表达 hAb1 的轻链的载体和 4-6 拷贝由 hCMV 启动子表达重链的载体; S421.7, 克隆细胞系携带单个拷贝的 pBDUpuro421; S454.5, 克隆细胞系携带单个拷贝的 pBDUpuro454; S804.4, 克隆细胞系携带单个拷贝的 pBDUpuro804; S452.3, 克隆细胞系携带单个拷贝的 pBDUpuro452。

在本说明书特别提及和/或在申请数据页列出的美国专利, 美国专利申请出版物, 美国专利申请, 外国专利, 外国专利申请以及非专

利著作均以其全文在此引作参考。

5 根据上文可以理解，尽管以说明为目的对本发明的特定实施方案予以描述，仍可以在不偏离本发明主旨和范围的前提下进行各种修改。因此，本发明除所附权利要求外不受其它限制。

<110> ML 实验室公开有限公司

<120> 高水平、大规模生产重组蛋白的组合体和方法
(COMPOSITIONS AND METHODS FOR HIGH-LEVEL, LARGE-SCALE
PRODUCTION OF RECOMBINANT PROTEINS)

<130> SCT033644-47

<140> PCT/US02/17763

<141> 2002-06-04

<160> 32

<210> 1

<211> 12701

<212> DNA

<213> 载体

<220>

<221> misc_特征

<222> 9602

<223> n = A, T, C 或 G

<400> 1

```

acgttgtaaa acgacggcca gtgaattgta atacgactca ctatagggcg aattgggtac 60
cgggccccc ctcgaggtcg agttggggtg gggaaaagga agaaacgcgg gcgtattggc 120
cccaatgggg tctcggtagg gtatcgacag agtgccagcc ctgggaccga accccgcgtt 180
tatgaacaaa cgaccacaaca cccgtgcgtt ttattctgtc tttttattgc cgtcatagcg 240
egggttcctt cgggtattgt ctccctccgt cgacgatctg acggttcact aaaccagctc 300
tgcttatata gacctccac cgtacacgcc taccgcccac ttgctcaat ggggaggagt 360
tgttacgaca ttttgaaaag tcccgttgat tttggtgcca aaacaaactc ccattgacgt 420
caatggggtg gagacttggg aatccccgtg agtcaaaccg ctatccacgc ccattgatgt 480
actgcaaaaa ccgcatcacc atggtaatag cgatgactaa tacgtagatg tactgccaag 540
taggaaagtc ccataaggtc atgtactggg cataatgcca ggccggccat ttaccgcat 600
tgacgtcaat agggggcgta cttggcatat gatacacttg atgtactgcc aagtggggcag 660
ttaccgtaa atactccacc cattgacgtc aatggaaagt ccctattggc gttactatgg 720
gaacatacgt cattattgac gtcaatgggc gggggtcgtt gggcggtcag ccaggcgggc 780
catttaccgt aagttatgta acgcggaact ccataatagg gctatgaact aatgaccccc 840
taattgatta ctattaataa ctcgacggta tcatggtagc gaccggcatg gtgagctgcg 900
agaatagccg ggcgcgctgt gagccgaagt cgcgcccgcc ctggccactt ccggcgcgcc 960
gagtccttag gccgccaggg ggcgcggcgc cgcgcccaga ttggggacaa aggaagccgg 1020
gccggccgcg ttattaccat aaaaggcaaa cactggtcgg aggcgtcccc gcggcgcgcg 1080
gcaggaagcc aggcccaaac cccctcccaa ccggggcgcca gccccgcctc cggccgggtc 1140
aaacagcgac cgggtcgcgc gcgcgcacgc agcggccaca ccctcggggc ccagcggctc 1200
gggcaggaag tggcgcaagc gcccgggccc cagaacgcac gcgcgattag cgccattgag 1260
tcccagcgcg cacgcgcaat tagcgccaat tcccagcgcg cacgcagtta gcgccccaaag 1320
gaccagcgcg cacgcgcatg gcgccccagc ccccaccggg cctgacgggg gctacgccc 1380
gccaccgtg cgatccccat tggcaagagc cgggctcaga caaagacccc gccggttggc 1440
cccgccccga gagcggcacc cccggagcgc gcccgcccga gcgcggcctc gcgcctgcca 1500
actggcgtgg ggtgtcccc atctccggag gcccgggggc ttctcccgcg cccccacgg 1560
cggtcgggtt ccgccccatg cgcgcccggc tgcggcccag acggcggtc tgcacgggcg 1620
aaggcccgcg gccgcagcc ccggtcggct ggccggggtt acctggcggc ggggtgtggac 1680
gggcggcgga tcggcaaagg cgaggctctg tgctcgcggg cggacgcggg ctcgccgggtg 1740
gtggcgcgct gcgcccgtgg gttttatagg gcgcccgcgc ggccgctcga gccataaaag 1800

```

gcaactttcg	gaacggcgca	cgctgattgg	ccccgcgccc	ctcactcacc	ggcttcgccc	1860
cacagtgcag	cattttttta	ccccctctcc	cctccttttg	cgaaaaaaaa	aaagagcgag	1920
agcgagattg	aggaagagga	ggagggagag	ttttggcggt	ggccgccttg	gggtgctggg	1980
cccgggggct	gggggcgccc	gccgtggccc	ccgcgcccc	cgctgggcag	tgcccggttc	2040
ggccccgcat	ggccaggcct	gcccccgccc	tgcccgtctc	tcgggcccc	caccaccgpc	2100
gggacatcct	aggtgtggac	atctcttggg	cactgagcgc	ccaggtgggg	tgggcccagg	2160
tctgcacggg	tgccagggcc	ctgggttctg	tacgtccttg	cagaaggagc	tcttggaggg	2220
catggagtgg	ccaggcagtc	actccccctt	gccgacttca	gagcaactgc	cctgaaagca	2280
gggacctgag	acctctggct	gtggggctca	gctagctaaa	tgtgctgggt	gggtcactag	2340
ggagagacct	ggccttgaga	ggtagagtgt	ggtgttgggg	gagtcaggtg	gcttgcggcc	2400
attagagtgc	caggaccaca	ctccccagga	cagggcaggg	gccagcggtc	cagtggctgg	2460
aggtggcccc	tgatgaaggc	tacaaacct	cccagccgca	gccctgggaa	ggaagtgggc	2520
tctacagggc	agggcacctt	ttaccctgga	gctgcctgct	tttgagggta	acagtacgpc	2580
ccagccaaga	ccaggcctgg	ggcgtagtg	ggtgacctag	gcactgcggg	gcgggggggc	2640
tggttctaca	cagcctgggt	ctgggcccac	cgctccgttg	atgtctgcta	tgcgagacca	2700
cagctgaact	gccctcccag	accatctgga	ggccgctggg	ggactctggg	gaccaagact	2760
ccatgtgcca	cagaggattg	ggggcggggc	ggtgctagga	actcaaagcc	agcctgggaa	2820
gaccctgtcc	ttgtcacctt	ttcttgcctt	gggtctgtcc	actgagtagc	acacaagacc	2880
gggtgggag	ggtccgttct	gctccgggaa	tcacagactg	tgtgtacca	ggtggtgggc	2940
atgcagcgat	cagtggcggt	ggaccacaga	gggggcccgc	ggtacctaaa	acagcttcac	3000
atggcttaaa	ataggggacc	aatgtctttt	ccaatctaag	tcccatttat	aataaagtc	3060
atgttccatt	tttaaaggac	aatccttttc	gtttaaaacc	aggcacgatt	acccaaaca	3120
ctcacaacct	taaagcactg	tgaatcttct	ctgttctgca	atcccaactt	ggttctgtct	3180
cagaaacctt	ccctctttcc	aatcgttaat	taaataacaa	aaggaaaaaa	cttaagatgc	3240
ttcaaccccc	ttctgtgaca	ctttgaaaaa	agaatcacct	cttgcaaaaa	cccgcctccc	3300
acccccgccc	ctgaagcccc	gcgtccagag	gcctaagcgc	gggtgcccgc	ccccacccgg	3360
gagcgcgggc	ctcgtgggtc	gcgcatccgc	ggggagaaa	aaaggccgpc	gcacgggggc	3420
tcaagggcac	tgccgacac	cgcacgcgpc	tacccccgpc	cggccacggt	aactggcggg	3480
cgccgcagcc	tcgggacagc	cggccgcgpc	ccgcccaggc	cgcggacgpc	ggaccacgpc	3540
ccgcccctcc	ggaggcccaa	gtctcgacct	agccccgpc	ggcgcctggg	gagggggcpc	3600
ctccgcccga	acgcgggtgg	gggaggggag	ggggaaatgc	gctttgtctc	gaaatggggc	3660
aaccgtcgcc	acagctccct	acccccctga	gggcagagca	gtccccccac	taactaccgg	3720
gctggccgpc	gcaccaggcca	gcccgcaggc	caccgcccga	ccctccactc	cttcccgcag	3780
ctcccggcpc	ggggtccggc	gagaagggga	ggggagggga	gaggagaacc	gggccccggg	3840
gacgcgtgtg	gcatctgaag	caccaccagc	gagcgagagc	tagagagaag	gaaagccacc	3900
gacttcaccg	cctccgagct	gctccgggtc	gcgggtctgc	agcgtctccg	gccctccgpc	3960
cctacagctc	aagccacatc	cgaagggggc	gggagccggg	agctgcgpc	ggggccgpc	4020
gggggagggg	tgccaccgcc	cacgcccggc	ggccacgaag	ggcggggcag	cgggcgpc	4080
cgccgcccgg	ggagggggcc	gcgcccggcc	cgctgggaa	tggggcccta	gggggagggc	4140
ggagggcgg	acgaccgpc	cacttaccgt	tcgcccgtg	gcgcccggg	gtccccagg	4200
ggaggggaag	gggaggggg	gcgaggacag	tgaccggagt	ctcctcagpc	gtggcttttc	4260
tgcttggcag	cctcagcggc	tggcgccaaa	accggactcc	gcccacttcc	tcgcccggpc	4320
gtgcgaggg	gtggaatcct	ccagacgctg	ggggaggggg	agttgggagc	ttaaaaacta	4380
gtaccctttt	gggaccactt	tcagcagcga	actctcctgt	acaccagggg	tcagttccac	4440
agacgcgggc	caggggtggg	tcattgcggc	gtgaacaata	atgtgactag	aagttgatcc	4500
gggtgtttcc	ggaagggggc	gagtcaatcc	gccgagttgg	ggcacggaaa	acaaaaagg	4560
aaggctacta	agatttttct	ggcgggggtt	atcattggcg	taactgcagg	gaccacctcc	4620
cgggttgagg	gggctggatc	tccaggctgc	ggattaagcc	cctcccgtcg	gcgttaattt	4680
caaaactgpc	gacgtttctc	acctgccttc	gccaaggcag	gggcccggac	cctattccaa	4740
gaggtagtaa	ctagcaggac	tctagccttc	cgcaattcat	tgagcgcatt	tacggaagta	4800
acgtcgggta	ctgtctctgg	ccgcaagggt	gggaggagta	cgcatttggc	gtaaggtggg	4860
gcgtagagcc	ttcccgcct	tggcggcgga	tagggcgttt	acgcgacggc	ctgacgtagc	4920
ggaagacgpc	ttagtggggg	ggaaggttct	agaaaagcgg	cggcagcggc	tctagcggca	4980
gtagcagcag	cgccgggtcc	cgtgcggagg	tgctcctcgc	agagttgttt	ctcagacagc	5040
ggcagttctc	actacagcpc	caggacgagt	ccggttcgtg	ttcgtccgpc	gagatctctc	5100
tcactctgct	cggctgcggg	aaatcgggct	gaagcagctg	agtcgcgcat	ggaggtaacg	5160
ggtttgaaat	caatgagtta	ttgaaaagg	catggcgagg	ccgttggcpc	ctcagtggaa	5220

gtcggccagc	cgctccgtg	ggagagaggg	aggaaatcgg	accaattcag	tagcagtggg	5280
gcttaaggtt	tatgaacggg	gtcttgagcg	gaggcctgag	cgtacaaaca	gcttccccac	5340
cctcagcctc	ccggcgccat	ttcccttcac	tgggggtggg	ggatggggag	ctttcacatg	5400
gcgagcgctg	ccccgctggg	gtgaaagtgg	ggcgcggagg	cggaattctt	tattcccttt	5460
ctaaagcacg	ctgcttcggg	ggccacggcg	tctcctcggc	gagcgtttcg	gcgggcagca	5520
ggctctcgtg	agcagggctg	cggagcttcc	cctccccctc	tctccccgga	accgatttgg	5580
cggccgccat	tttcatggct	cgcttctctc	tcagcgtttt	ccttataact	cttttatttt	5640
cttagtgtgc	tttctctatc	aagaagtaga	agtggttaac	tatttttttt	ttcttctcgg	5700
gtctgtttca	tatcgtttcc	agggtgattt	ggagtgtttt	gtgagcttgg	atcttttagag	5760
tcttgccgac	ctcattaaag	gcgctcagcc	ttcccctcga	tgaaatggcg	ccattgcggt	5820
cggaagccac	accgaagagc	ggggaggggg	ggtgctccgg	gtttgcgggc	ccggtttcag	5880
agaagatatic	accaccagg	gcgctggggc	gggttcaatg	cgagccgtag	gacaaagaaa	5940
ccattttatg	tttttctgt	cttttttttc	ctttgagtaa	cggttttatc	tgggtctgca	6000
gtcagtaaaa	cgacagatga	accgcgga	aataaacata	aattggaagc	catcggccac	6060
gaggggcagg	gacgaaggtg	gttttctggg	cgggggaggg	atattcgcgt	cagaatcctt	6120
tactgttctt	aaggattccg	tttaagttgt	agagctgact	cattttaagt	aatggtgtta	6180
ctgagaagtt	taacccttac	gggacagatc	catggacctt	tatagatgat	tacgaggaaa	6240
gtgaaataac	gattttgtcc	ttagttatac	ttcgattaaa	acatggcttc	agaggctcct	6300
tctgtaatg	cgtatggatt	gatgtgcaaa	actgttttgg	gcctgggccc	ctctgtattt	6360
gaactttggt	acttttctca	ttttgtttgc	aatcttgggt	gaacattaca	ttgataagca	6420
taaggtctca	agcgaagggg	gtctacctgg	ttatttttct	ttgaccctaa	gcacgtttat	6480
aaaataacat	tgtttaaaat	cgatagtggg	catcgggtaa	gtttggataa	attgtgaggt	6540
aagtaatgag	tttttgcttt	ttgttagtga	ttgttaaaac	ttgttataaa	tgtacattat	6600
ccgtaatttc	agtttagaga	taacctatgt	gctgacgaca	attaagaata	aaaactagct	6660
gaaaaaatga	aaataactat	cgtgacaagt	aaccatttca	aaagactgct	ttgtgtctca	6720
taggagctag	tttgatcatt	tcagttaatt	ttttctttaa	tttttacgag	tcatgaaaac	6780
tacaggaaaa	aaaatctgaa	ctgggtttta	ccactacttt	ttaggagttg	ggagcatgcg	6840
aatggagggg	gagctccgta	gaactgggat	gagagcagca	attaatgctg	cttgctagga	6900
acaaaaaata	attgattgaa	aattacgtgt	gacttttttag	tttgattat	gcgtttgtag	6960
cagttggctc	tggatatcac	tttctctcgt	ttgaggtttt	ttaacctagt	taacttttaa	7020
gacaggtttc	cttaacattc	ataagtgcc	agaatacagc	tgtgtagtac	agcatataaa	7080
gatttcagct	ctgaggtttt	tcctattgac	ttggaaaatt	gttttgtgcc	tgtcgcttgc	7140
cacatggcca	atcaagtaag	cttcgaattc	gagctcgccc	aactccgccc	gttttatgac	7200
tagaaccaat	agtttttaat	gccaaatgca	ctgaaatccc	ctaatttgca	aagccaaacg	7260
ccccctatgt	gagtaatacg	gggacttttt	acccaatttc	ccaagcggaa	agccccctaa	7320
tacactcata	tggcatatga	atcagcacgg	tcatgcactc	taatggcggc	ccatagggga	7380
ttccacata	gggggcgctt	accatttccc	agcagagggg	tggtgactca	atggccttta	7440
ccaagtaca	ttgggtcaat	gggaggtaa	ccaatgggtt	tttccatta	ctggcaagca	7500
cactgagtca	aatgggactt	tccactgggt	tttgcccagg	tacattgggt	caatgggagg	7560
tgagccaatg	ggaaaaacc	attgctgcca	agtacactga	ctcaataggg	actttccaat	7620
gggtttttcc	attgttgcca	agcatataag	gtcaatgtgg	gtgagtcaat	agggactttc	7680
cattgtattc	tgcccagtac	ataaggtcaa	taggggtgga	atcaacagga	aagtcccatt	7740
ggagccaagt	acactgcgtc	aatagggact	ttccattggg	ttttgcccag	tacataaggt	7800
caatagggga	tgagtcaatg	ggaaaaacc	attggagcca	agtacactga	ctcaataggg	7860
actttccatt	gggttttgcc	cagtacataa	ggtcaatagg	gggtgagtca	acaggaaagt	7920
cccattggag	ccaagtacat	tgagtcaata	gggactttcc	aatgggtttt	gccaggtaca	7980
taaggtcaat	gggaggtaa	ccaatgggtt	tttccatta	ctggcacgta	tactgagtca	8040
ttagggactt	tccaatgggt	tttgcccagt	acataaggct	aataggggtg	aatcaacagg	8100
aaagtcccat	tggagccaag	tacactgagt	caatagggac	tttccattgg	gttttgccca	8160
gtacaaaagg	tcaatagggg	gtgagtcaat	gggtttttcc	cattattggc	acgtacataa	8220
ggtcaatagg	ggtgagtcat	tgggtttttc	cagccaattt	aattaaaacg	ccatgtactt	8280
tcccaccatt	gacgtcaatg	ggctattgaa	actaatgcaa	cgtgaccttt	aaacgggtact	8340
ttcccatagc	tgattaatgg	gaaagtaccg	ttctcgagcc	aatacacgtc	aatgggaagt	8400
gaaagggcag	ccaaaacgta	acaccgcccc	ggttttcccc	tggaaattcc	atattggcac	8460
gcattctatt	ggctgagctg	cgttctacgt	gggtataaga	ggcgcgacca	gcgtcggtac	8520
cgtcgcagtc	ttcggctgta	ccaccgtaga	acgcagagct	cctcgctgca	gcccgggtct	8580
agaggatccg	cctgagaaag	gaagtgagct	gtaaaggctg	agctctctct	ctgacgtatg	8640

tagcctctgg	ttagcttctg	cactcactgt	tcttgactca	gcatggcaat	ctgatgaaat	8700
cccagctgta	agtctgcaga	aattgatgat	ctattaacaa	ataaagatgt	ccactaaaat	8760
ggaagttttt	cctgtcatac	tttgtaaga	agggtgagaa	cagagtacct	acattttgaa	8820
tggaaggatt	ggagctacgg	gggtgggggt	gggggaggat	tagataaatg	cctgctcttt	8880
actgaaggct	ctttactatt	gctttatgat	aatgtttcat	agttggatat	cataatttaa	8940
acaagcaaaa	ccaaattaag	ggccagctca	ttcctccaga	tccactagta	attctgtgga	9000
atgtgtgtca	gttagggtgt	ggaaagtccc	caggtccccc	agcaggcaga	agtatgcaaa	9060
gcatgcatct	caattagtca	gcaaccaggt	gtggaaagtc	cccaggctcc	ccagcaggca	9120
gaagtatgca	aagcatgcat	ctcaattagt	cagcaaccat	agtcccggcc	ctaactccgc	9180
ccatcccggc	cctaactccg	cccagttccg	cccattctcc	gccccatggc	tgactaattt	9240
tttttattta	tgagaggcc	gaggccgcct	ctgcctctga	gctattccag	aagtagtgag	9300
gaggcttttt	tggaggccta	ggcttttgca	aaaagotccc	gggagcttgt	atatccattt	9360
tcgatctga	tcaagagaca	ggatgaggat	cgtttcgcac	gattgaacaa	gatggattgc	9420
acgcaggttc	tccggccgct	tgggtggaga	ggctattcgg	ctatgactgg	gcacaacaga	9480
caatcggctg	ctctgatgcc	gccgtgttcc	ggctgtcagc	gcagggggcg	ccggttcttt	9540
ttgtcaagac	cgacctgtcc	ggtgccctga	atgaactgca	ggacgaggca	gcgcggttat	9600
cstggctggc	cacgacgggc	gttccttgcg	cagctgtgct	cgacgttgtc	actgaagcgg	9660
gaagggactg	gctgctattg	ggcgaagtgc	cggggcagga	tctcctgtca	tctcaccttg	9720
ctcctgccga	gaaagtatcc	atcatggctg	atgcaatgcg	gcggtgcat	acgcttgatc	9780
ggctacctg	cccattcgac	caccaagcga	aacatcgcat	cgagcgagca	cgtactcgga	9840
tggaaagcgg	tcttctcgat	caggatgac	tggacgaaga	gcatcagggg	ctcgcgccag	9900
ccgaactggt	cgccaggctc	aaggcgcgca	tgcccagcgg	cgaggatctc	gtcgtgaccc	9960
atggcgatgc	ctgcttggcg	aatatcatgg	tggaaaatgg	ccgcttttct	ggattcatcg	10020
actgtggccg	gctgggtgtg	gcggaccgct	atcaggacat	agcgttggct	accctgata	10080
ttgctgaaga	gcttggcggc	gaatgggctg	accgcttctc	cgtgctttac	ggtatcgccg	10140
ctcccgattc	gcagcgcac	gccttctatc	gccttcttga	cgagttcttc	tgagcgggac	10200
tctggggttc	gaaatgaccg	accaagcgac	gcccacctg	ccatcacgag	atttcgattc	10260
caccgccgcc	ttctatgaaa	ggttgggctt	cggaatcggt	ttccgggacg	ccggctggat	10320
gatcctccag	cgccgggatc	tcatgctgga	gttcttcgcc	cacccaact	tgtttattgc	10380
agcttataat	ggttacaaat	aaagcaatag	catcacaat	ttcacaata	aagcattttt	10440
ttcactgcat	tctagttgtg	gtttgtccaa	actcatcaat	gtatcttatc	atgtctgat	10500
accgtcgaga	ctagttctag	agcggccgcc	accgcgggtg	agctccagct	tttgttccct	10560
ttagtgaggg	ttaatttcga	gcttggcgta	atcatggtca	tagctgtttc	ctgtgtgaaa	10620
ttgttatccg	ctcacaattc	cacacaacat	acgagccgga	agcataaagt	gtaaagcctg	10680
gggtgcctaa	tgagttagct	aactcacatt	aattgcgttg	cgctcactgc	ccgctttcca	10740
gtcgggaaac	ctgtcgtgcc	aggggggtacc	taggcggggc	aacaattggc	ggccggccgc	10800
acttttcggg	gaaatgtgcg	cggaaccctt	atltgtttat	ttttctaaat	acattcaaat	10860
atgtatccgc	tcatgagaca	ataaccctga	taaatgcttc	aataatattg	aaaaaggaag	10920
agtatgagta	ttcaacattt	ccgtgtcgcc	cttattccct	tttttgccgc	atlttgctt	10980
cctgtttttg	ctcaccaga	aacgctgggt	aaagtaaaag	atgtgaaga	tcagttgggt	11040
gcacgagtgg	gttacatcga	actggatctc	aacagcggta	agatccttga	gagttttcgc	11100
cccgaagaac	gttttccaat	gatgagcact	tttaaagttc	tgctatgtgg	cgcggtatta	11160
tcccgtattg	acgccgggca	agagcaactc	ggtcgccgca	tacactattc	tcagaatgac	11220
ttggttgagt	actcaccagt	cacagaaaag	catcttacgg	atggcatgac	agtaagagaa	11280
ttatgcagtg	ctgccataac	catgagtgat	aacactgccc	ccaacttact	tctgacaacg	11340
atcggaggac	cgaaggagct	aaccgctttt	ttgcacaaca	tgggggatca	tgtaactcgc	11400
cttgatcgtt	gggaaccgga	gctgaatgaa	gccataccaa	acgacgagcg	tgacaccacg	11460
atgcctgtag	caatggcaac	aacgttgccg	aaactattaa	ctggcgaact	acttactcta	11520
gcttcccggc	aacaattaat	agactggatg	gaggcggata	aagttgcagg	accacttctg	11580
cgctcggccc	ttccggctgg	ctggtttatt	gctgataaat	ctggagccgg	tgagcgtggg	11640
tctcgggta	tcattgcagc	actggggcca	gatggtaagc	cctcccgtat	cgtagttatc	11700
tacacgacgg	ggagtccagg	aactatggat	gaacgaaata	gacagatcgc	tgagataggt	11760
gcctcactga	ttaagcattg	gtaactgtca	gaccctagcc	cgggcaacaa	ttggcggccg	11820
gccctgcatt	aatgaatcgg	ccaacgcgcg	gggagaggcg	gtttgcgtat	tgggcgctct	11880
tccgcttctc	cgctcactga	ctcgcctgcg	tccgctcgtc	ggctgcggcg	agcggtatca	11940
gctcactcaa	agggcgtaat	acggttatcc	acagaatcag	gggataacgc	aggaagaac	12000
atgtgagcaa	aaggccagca	aaaggccagg	aaccgtaaaa	aggccgcgct	gctggcgttt	12060

```

ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 12120
cgaaaccgga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgccc 12180
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cttttctccc ttcggaagc 12240
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cgggtgtaggt cgttcgctcc 12300
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcccctt atccggtaac 12360
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggg 12420
aacaggatta gcagagcgag gtatgtaggg ggtgctacag agttctttaa gtgggtggcct 12480
aactacggct aactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 12540
ttcggaaaaa gagttggtag ctcttgatcc ggcaaaaaa ccaccgctgg tagcgggtgg 12600
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 12660
atcttttcta cggggtctga cgctcagtg aacgaaaact c 12701

```

<210> 2

<211> 12109

<212> DNA

<213> 载体

<220>

<221> misc_特征

<222> 9010

<223> n = A, T, C 或 G

<400> 2

```

acgttgtaaa acgacggcca gtgaattgta atacgactca ctatagggcg aattgggtac 60
ggggcccccc ctcgaggtcg agttgggggtg gggaaaagga agaaacgcgg gcgtattggc 120
cccaatgggg tctcggtggg gtatcgacag agtgccagcc ctgggaccga accccgcgtt 180
tatgaacaaa cgacccaaca cccgtgctgtt ttattctgtc tttttattgc cgtcatagcg 240
cgggttccct cgggtattgt ctccctccgt cgacggtatc aagggtggcg cgggaatggg 300
gagctgagag aatagccggg cgcgctgtga gccgaagtcg cccccgccct ggccacttcc 360
ggcgcgccga gtccttaggc cgccaggggg cgccggcgcg cgccagatt ggggacaaaag 420
gaagccgggg cggccgcggt attaccataa aaggcaaaaca ctggtcggag gcgtccccgc 480
ggcgcgccgg aggaagccag gccccaaccc cctcccaacc gggcgccagc ccgcctccg 540
cccgttcaa acagcgaccg ggtcgcgcg cgcaacgcag cggccacacc ctcgggcgcc 600
agcggctcgg gcaggaagtg gcgcaagcgc ccgggcccc gaacgcacgc gcgattagcg 660
ccattgagtc ccagcgcgca cgcgcaatta gcgccaattc ccagcgcgca cgcagttagc 720
gccccaaagga ccagcgcgca cgcgcatggc gccccagccc ccaccgggccc tgacggggggc 780
tacgccgcgc ccaccgtgcg atccccattg gcaagagccc ggctcagaca aagacccccg 840
cggttgcccc cgccccgaga gcggcaccgc cggagcgcg cggcccagc gcggcctcgc 900
gcctgcgaac tggcgtgggg tgtcccccat ctccggaggc ccaggggctt ctcccgcgcc 960
ccccacggcg gtccggttcc gccccatgcg ccccccgctg cggcccagac ggcggtctctg 1020
cacgggcgaa gggccgcggc cgcatgcccc ggtcggctgg ccgggcttac ctggcgggcg 1080
gtgtggacgg gcggcggatc ggcaaaaggc aggctctgtg ctcgcgggcg gacgcggtct 1140
cggcggtggg ggcgctcgc gccgctgggt ttatagggc gccgcccgg ccgctcgagc 1200
cataaaaggc aactttcgg aacggcgacg ctgattggcc ccgcccggct cactcacggg 1260
cttcgcccga cagtgcagca ttttttacc ccctctcccc tcttttgcg aaaaaaaaaa 1320
agagcgagag cgagattgag gaagaggagg agggagagtt ttggcgttgg ccgccttggg 1380
gtgctgggccc cgggggctgg gggcgcgcg cgtggcccc gcgccccacg ctgggcaagt 1440
cccggttcgg ccccgcattg ccaggcctgc ccccggcctg ccggtctctc gggcccccca 1500
cccaccgcgg gacatcctag gtgtggacat ctcttgggca ctgagcggcc aggtgggggtg 1560
ggccagggtc tgcacgggtg ccaggccctt gggttctgta cgctcctgca gaaggagctc 1620
ttggagggca tggagtggcc aggcagtcac tcccccttgc cgacttcaga gcaactgccc 1680
tgaaagcagg gcctgaggac ctctggctgt ggggctcagc tagctaaatg tgctgggtgg 1740
gtcactaggg agagacctgg gcttgagagg tagagtgtgg tgttggggga gtcagggtgg 1800
ttgcggccat tagagtgcga ggaccacact ccccaggaca gggcagggggc cagcgggtcca 1860
gtggctggag gtggcccgtg atgaaggcta caaacctacc cagccgcagc cctgggaagg 1920
aagtgggctc tacagggcag ggcacctttt accctggagc tgctgcttt tgagggtaac 1980

```

agtcacgccc	agccaagacc	aggcctgggg	cgttagtggg	tgacctaggg	actgcggggg	2040
gggggggctg	ggtctacaca	gcctgggtct	gggcccaccg	tccgttgtat	gtctgtctatg	2100
cgagccaca	gctgaactgc	cctcccagac	catctggagg	ccgctggggg	actctggggg	2160
ccaagactcc	atgtgccaca	gaggattggg	ggcggggcgg	tgctaggaac	tcaaagccag	2220
cctgggaaga	ccctgtcctt	gtcacccttt	cttgccttgg	gtctgtccac	tgagtagcac	2280
acaagaccgg	gtgggcaggg	tccgttctgc	tccgggaatc	acagactgtg	tgtaccagg	2340
tggtgggcat	gcagcgatca	gtggcgtggg	accacagagg	gggcccgcgg	tacctaaaac	2400
agcttcacat	ggcttaaaat	aggggaccaa	tgtcttttcc	aatctaagtc	ccatttataa	2460
taaagttccat	gttccattht	taaaggacaa	tcccttcggg	ttaaaaccag	gcagattac	2520
ccaaacaact	cacaacggta	aagcactgtg	aatcttctct	gttctgcaat	cccaacttgg	2580
tttctgtca	gaaaccctcc	ctctttccaa	tccgtaatta	aataacaaaa	ggaaaaaact	2640
taagatgctt	caaccccggt	tctgtgacct	ttgaaaaaag	aatcacctct	tgcaaacacc	2700
cgctcccagc	ccccgcgct	gaagcccggc	gtccagaggc	ctaagcgcgg	gtgcccggcc	2760
ccacccggga	gcgcgggcct	cggtggtcagc	gcatccgcgg	ggagaaacaa	aggccgcggc	2820
acgggggctc	aagggcactg	cgccacaccg	cacgcgccta	ccccgcgcg	gccacgttaa	2880
ctggcggctc	ccgcagcctc	gggacagccg	gccgcgcgcc	gccaggctcg	cggacgcggg	2940
accacgcgcc	gcccctccggg	aggcccaagt	ctcgaccag	ccccgcgtgg	cgctggggga	3000
gggggcgcct	ccgccggaac	gcgggtgggg	gaggggaggg	ggaaatgcgc	tttgtctcga	3060
aatggggcaa	ccgtcgccac	agctccctac	cccctcgagg	gcagagcagt	ccccccacta	3120
actaccgggc	tggccgcgcg	ccaggccagc	cgcgaggcca	ccgcccagcc	ctccactcct	3180
tcccgcagct	cccggcgcgg	ggtccggcga	gaaggggagg	ggaggggagc	ggagaaccgg	3240
gcccccgga	cgctgtggc	atctgaagca	ccaccagcga	gcgagagcta	gagagaagga	3300
aagccaccga	cttaccgcgc	tccgagctgc	tcgggtcgc	gggtctgcag	cgtctccggc	3360
cctccgcgcc	tacagctcaa	gccacatccg	aaggggaggg	gagccgggag	ctgcccgcgg	3420
ggccgcggg	gggaggggtg	gcaccgcccc	cgccgggcgg	ccacgaaggg	cggggcagcg	3480
ggcgcgcgcg	cggcgggggg	aggggcccgc	gccgcgcccg	ctgggaattg	gggcccagg	3540
gggagggcgg	aggcgcggac	gaccgcggca	cttaccgttc	gcggcgtggc	gcccgggtgg	3600
ccccaaaggg	agggaaaggg	gaggcggggc	gaggacagtg	accggagtct	cctcagcggg	3660
ggcttttctg	cttggcagcc	tcagcggctg	gcgcaaaaac	cggactccgc	ccacttctct	3720
gcccgcgggt	gcgaggggtg	ggaatcctcc	agacgctggg	ggagggggag	ttgggagctt	3780
aaaaactagt	acccctttgg	gaccactttc	agcagcgaac	tctcctgtac	accaggggtc	3840
agttccacag	acgcgggcca	ggggtgggtc	attgcgcgct	gaacaataat	ttgactagaa	3900
gttgattcgg	gtgtttccgg	aaggggcccga	gtcaatccgc	cgagttgggg	cacggaaaac	3960
aaaaagggaa	ggctactaag	atthttctgg	cgggggttat	cattggcgta	actgcagggg	4020
ccactctccc	ggttgagggg	gctggatctc	caggtcgcgg	attaagcccc	tcccgtcggc	4080
gtaaatctca	aactgcccga	cgthttctac	tgcccttcgc	caaggcaggg	gcccggacc	4140
tattccaaga	ggtagtaact	agcaggactc	tagccttccg	caattcattg	agcgcattta	4200
cggaagtaac	gtcgggtact	gtctctggcc	gcaaggggtg	gaggagtacg	catttggcgt	4260
aaggtggggc	gtagagcctt	cccgccattg	gcggcggata	gggcgtttac	gcgacggcct	4320
gacgtagcgg	aagacgcggt	agtggggggg	aaggttctag	aaaagcggcg	gcagcggctc	4380
tagcggcagc	agcagcagcg	ccgggtcccg	tgccgaggtg	ctcctcgcag	agttgtttct	4440
cgagcagcgg	cagttctcac	tacagcgcga	ggacgagctc	ggttcgtggt	cgtccgcgga	4500
gatctctctc	atctcgctcg	gctgcgggaa	atcgggctga	agcagactgag	tccgcgatgg	4560
aggtaacggg	tttgaatca	atgagttatt	gaaaagggca	tggcgaggcc	gttggcgcc	4620
cagtggaagt	cggccagccg	cctccgtggg	agagagggcag	gaaatcggac	caattcagta	4680
gcagtggggc	ttaaggttta	tgaaacgggg	cttgagcggg	ggcctgagcg	tacaaacagc	4740
ttccccacc	tcagcctccc	ggcgccattt	cccttcaactg	ggggtggggg	atggggagct	4800
ttcacatggc	ggacgctgcc	ccgctggggg	gaaagtgggg	cgcgaggcgg	ggaattctta	4860
ttccctttct	aaagcacgct	gcttcggggg	ccacggcgtc	tccctcggcg	gcgtttcggc	4920
gggcagcagg	tctcgtgag	cgaggctcgc	gagcttcccc	tccccctctc	tcccgggaac	4980
cgatttgcg	gccgccattt	tcattggctcg	ccttcctctc	agcgttttcc	ttataactct	5040
ttattttct	tagtgtgctt	tctctatcaa	gaagtagaag	tggttaacta	tttttttttt	5100
cttctcgggc	tgthttcata	tcgtttcag	gtggatttgg	agtgttttgt	gagcttggat	5160
ctttagagtc	ctgcgcacct	cattaaaggc	gctcagcctt	cccctcgatg	aaatggcgcc	5220
attgcggtcg	gaagccacac	cgaagagcgg	ggaggggggg	tgctccgggt	ttgcggggcc	5280
ggtttcagag	aagatatcac	caccagggc	gtcgggcccg	gthcaatgcg	agccgtagga	5340
caaagaaacc	atthtatgth	tttctgtct	tttttttct	ttgagtaacg	gthttatctg	5400

ggtctgcagt	cagtaaaacg	acagatgaac	cgcggaacaa	taaacataaa	ttggaagcca	5460
tcggccacga	ggggcagggg	cgaaggtggt	tttctgggcg	ggggagggat	attcgcgtca	5520
gaatccttta	ctgttcttaa	ggattccggt	taagttgtag	agctgactca	ttttaagtaa	5580
tgttgttact	gagaagttta	acccttacgg	gacagatcca	tggaccttta	tagatgatta	5640
cgaggaaagt	gaaataacga	ttttgtcctt	agttataact	cgattaaaac	atggcctcag	5700
aggctccttc	ctgtaatgcg	tatggattga	tgtgcaaaac	tgttttgggc	ctgggcccgt	5760
ctgtatttga	actttgttac	ttttctcatt	ttgtttgcaa	tcttggttga	acattacatt	5820
gataagcata	aggctctcaag	cgaagggggg	ctacctgggt	atTTTTcttt	gaccctaagc	5880
acgtttataa	aataacattg	tttaaaatcg	atagtgagca	tcgggtaagt	ttggataaat	5940
tgtgaggtaa	gtaatgagtt	tttgcttttt	gttagtgatt	tgtaaaactt	gtataaatg	6000
tacattatcc	gtaatttcag	tttagagata	acctatgtgc	tgacgacaat	taagaataaa	6060
aactagctga	aaaaatgaaa	ataactatcg	tgacaagtaa	ccatttcaaa	agactgcttt	6120
gtgtctcata	ggagctagtt	tgatcatttc	agttaatttt	ttctttaatt	tttacgagtc	6180
atgaaaacta	caggaaaaaa	aatctgaact	gggttttacc	actacttttt	aggagtggg	6240
agcatgcgaa	tggagggaga	gctccgtaga	actgggatga	gagcagcaat	taatgctgct	6300
tgctaggaac	aaaaaataat	tgattgaaaa	ttacgtgtga	ctttttagtt	tgctattatg	6360
gtttgtagca	gttggctctg	gatatcactt	tctctcgttt	gaggtttttt	aacctagtta	6420
acttttaaga	caggtttcct	taacattcat	aagtgccag	aatacagctg	tgtagtacag	6480
catataaaga	tttcagctct	gaggtttttc	ctattgactt	ggaaaattgt	tttgtgcctg	6540
tcgcttgcca	catggccaat	caagtaagct	tcgaattcga	gctcgcccaa	ctccgcccgt	6600
tttatgacta	gaaccaatag	tttttaatgc	caaatgcact	gaaatcccct	aatttgcaaa	6660
gccaaacgcc	ccctatgtga	gtaatacggg	gactttttac	ccaatttccc	aagcggaaaag	6720
cccctaata	caactcatatg	gcatatgaat	cagcacgggc	atgcactcta	atggcggccc	6780
atagggactt	tccacatagg	gggcgttcac	catttcccag	cataggggtg	gtgactcaat	6840
ggcctttacc	caagtaacatt	gggtcaatgg	gaggtaaagc	aatgggtttt	tcccattact	6900
ggcaagcaca	ctgagtcaaa	tgggactttc	cactgggttt	tgcccaagta	cattgggtca	6960
atgggaggtg	agccaatggg	aaaaacccat	tgctgccaag	tacactgact	caatagggac	7020
tttccaatgg	gtttttccat	tgttggcaag	catataaggt	caatgtgggt	gagtcaatag	7080
ggactttcca	ttgtattctg	cccagtaac	aaggtcaata	gggggtgaat	caacaggaaa	7140
gtcccattgg	agccaagtac	actgcgtcaa	tagggacttt	ccattgggtt	ttgcccagta	7200
cataaggtca	ataggggatg	agtcaatggg	aaaaacccat	tggagccaag	tacactgact	7260
caatagggac	tttcattggg	gttttgccca	gtacataaag	tcaatagggg	gtgagtcaac	7320
aggaaagtcc	cattggagcc	aagtacattg	agtcaatagg	gactttccaa	tgggttttgc	7380
ccagtacata	aggccaatgg	gaggtaaagc	aatgggtttt	tcccattact	ggcacgtata	7440
ctgagtcatt	agggactttc	caatgggttt	tgcccagtac	ataaggtcaa	taggggtgaa	7500
tcaacaggaa	agtcccattg	gagccaagta	caactgagta	atagggactt	tccattgggt	7560
tttgcccagt	acaaaaggtc	aatagggggg	gagtcaatgg	gtttttccca	ttattgggac	7620
gtacataaag	tcaatagggg	tgagtcattg	ggtttttcca	gccaatttaa	ttaaaacgcc	7680
atgtactttc	ccaccattga	cgtaaatggg	ctattgaaac	taatgcaacg	tgaccttaa	7740
acggactttt	cccatagctg	attaatggga	aagtaccggt	ctcgagccaa	tacacgtcaa	7800
tgggaagtga	aagggcagcc	aaaacgtaac	accgccccgg	ttttcccctg	gaaattccat	7860
attggcacgc	attctattgg	ctgagctgcg	ttctacgtgg	gtataagagg	cgcgaccagc	7920
gtcggtagcc	tcgcagctct	cggctctgacc	accgtagaac	gcagagctcc	tcgctgcagc	7980
ccgggtctag	aggatccgcc	tgagaaagga	agtgagctgt	aaaggctgag	ctctctctct	8040
gacgatgta	gcctctgggt	agcttcgtca	ctcactgttc	ttgactcagc	atggcaatct	8100
gatgaaatcc	cagctgtaag	tctgcagaaa	ttgatgatct	attaacaat	aaagatgtcc	8160
actaaaatgg	aagtttttcc	tgtcataact	tgtaagaag	ggtgagaaca	gagtacctac	8220
atTTTgaatg	gaaggattgg	agctacgggg	gtgggggtgg	ggtgggatta	gataaatgcc	8280
tgctctttac	tgaaggctct	ttactattgc	tttatgataa	tgtttcatag	ttggatatca	8340
taatttaaac	aagcaaaaacc	aaattaaggg	ccagctcatt	cctccagatc	cactagtaat	8400
tctgtggaat	gtgtgtcagt	taggggtgtg	aaagccccca	ggctccccag	caggcagaag	8460
atgcaaagc	atgcatctca	attagtcact	aaccaggtgt	ggaaagtccc	caggctcccc	8520
agcaggcaga	agtatgcaaa	gcatgcatct	caattagtca	gcaacctatg	tcccggcccct	8580
aactccgccc	atcccggccc	taactccgcc	cagttccgcc	cattctccgc	ccccgggctg	8640
actaatTTTT	tttatttatg	cagaggccga	ggccgcctct	gcctctgagc	tattccagaa	8700
gtagtgagga	ggctTTTTTg	gaggcctagg	cttttgcaaa	aagctcccgg	gagcttgtat	8760
atccattttc	ggatctgatc	aagagacagg	atgaggatcg	tttcgcatga	ttgaacaaga	8820

tggattgcac	gcaggttctc	cgcccgcttg	ggtggagagg	ctattcggct	atgactgggc	8880
acaacagaca	atcggctgct	ctgatgccgc	cgtgttccgg	ctgtcagcgc	aggggcgccc	8940
ggttcttttt	gtcaagaccg	acctgtccgg	tgccctgaat	gaactgcagg	acgaggcagc	9000
gcggtatcs	tggctggcca	cgacgggcgt	tccttgcgca	gctgtgctcg	acgttgtcac	9060
tgaagcggga	agggactggc	tgctattggg	cgaagtgccg	gggcaggatc	tcctgtcatc	9120
tcaccttgct	cctgccgaga	aagtatccat	catggctgat	gcaatgcggc	ggctgcatac	9180
gcttgatccg	gctacctgcc	cattcgacca	ccaagcga	catcgcatcg	agcgagcagc	9240
tactcggatg	gaagccggtc	ttgtcgatca	ggatgatctg	gacgaagagc	atcaggggct	9300
cgcgccagcc	gaactgttcc	ccaggctcaa	ggcgcgcgatg	cccgcggcgg	aggatctcgt	9360
cgtagcccat	ggcgatgcct	gcttgcgaa	tatcatgggtg	gaaaatggcc	gcttttctgg	9420
attcatcgac	tgtggccggc	tgggtgtggc	ggaccgctat	caggacatag	cgttggctac	9480
ccgtgatatt	gctgaagagc	ttggcggcga	atgggctgac	cgcttcctcg	tgctttacgg	9540
tatgcccgct	cccgattcgc	agcgcacgc	cttctatcgc	cttcttgacg	agttcttctg	9600
agcgggactc	tggggttcga	aatgaccgac	caagcgcgc	ccaacctgcc	atcacgagat	9660
ttcgattcca	ccgccgcctt	ctatgaaagg	ttgggcttcg	gaatcgtttt	ccgggacgcc	9720
ggctggatga	tcctccagcg	cggggatctc	atgctggagt	tcttcgccc	ccccaaactg	9780
tttattgcag	cttataatgg	ttacaaataa	agcaatagca	tcacaaat	cacaaataaa	9840
gcattttttt	cactgcattc	tagttgtggt	ttgtccaaac	tcatcaatgt	atcttatcat	9900
gtctgtatac	cgtcgagact	agttctagag	cgcccgccac	cgcggtggag	ctccagcttt	9960
tgttcccttt	agtgagggtt	aatttcgagc	ttggcgtaat	catggtcata	gctgtttcct	10020
gtgtgaaatt	gttatccgct	cacaattcca	cacaacatac	gagccggaag	cataaagtgt	10080
aaagcctggg	gtgcctaattg	agtgagctaa	ctcacattaa	ttgcgttgcg	ctcactgccc	10140
gctttccagt	cgggaaacct	gtcgtgccag	gggtacctta	ggccgggcaa	caattggcgg	10200
ccggccgcac	ttttcgggga	aatgtgcgcg	gaaccctat	ttgtttattt	ttctaaatac	10260
attcaaatat	gtatccgctc	atgagacaat	aacctgata	aatgcttcaa	taatattgaa	10320
aaaggaagag	tatgagtatt	caacatttcc	gtgtcgccct	tattcccttt	tttgcggcat	10380
tttgccttcc	tgtttttgct	caccagaaa	cgctggtgaa	agtaaaagat	gctgaagatc	10440
agttgggtgc	acgagtgggt	tacatcgaac	tggatctcaa	cagcggtaag	atccttgaga	10500
gttttcgccc	cgaagaacgt	tttccaatga	tgagcacttt	taaagtctg	ctatgtggcg	10560
cggtattatc	ccgtattgac	gccgggcaag	agcaactcgg	tcgccgcata	cactattctc	10620
agaatgactt	ggttgagtac	tcaccagtca	cagaaaagca	tcttacggat	ggcatgacag	10680
taagagaatt	atgcagtgct	gccataacca	tgagtgataa	cactgcggcc	aacttacttc	10740
tgacaacgat	cggaggaccg	aaggagctaa	ccgctttttt	gcacaacatg	ggggatcatg	10800
taactgcctt	tgatcgttgg	gaaccggagc	tgaatgaagc	cataccaaac	gacgagcgtg	10860
acaccacgat	gctgtagca	atggcaacaa	cgttgcgcaa	actattaact	ggcgaactac	10920
ttactctagc	ttcccggcaa	caattaatag	actggatgga	ggcggataaa	gttgcaggac	10980
cacttctgcg	ctcggccctt	ccggctggct	ggtttattgc	tgataaatct	ggagccgggtg	11040
agcgtgggtc	tcgcggtatc	attgcagcac	tggggccaga	tggttaagccc	ttccgtatcg	11100
tagttatcta	cacgagggg	agtcaggcaa	ctatggatga	acgaaataga	cagatcgctg	11160
agataggtgc	ctcactgatt	aagcattggt	aactgtcaga	ccctaggccg	ggcaacaatt	11220
ggcggccggc	cctgcattaa	tgaatcggcc	aacgcgcggg	gagaggcggg	ttgcgtattg	11280
ggcgcctctc	cgcttcctcg	ctcactgact	cgctgcgctc	ggtcgttcgg	ctgcggcgag	11340
cggtatcagc	tcactcaaag	gcggtataac	ggttatccac	agaatcaggg	gataacgcag	11400
gaaagaacat	gtgagcaaaa	ggccagcaaa	aggccaggaa	ccgtaaaaag	gccgcggttc	11460
tggcgttttt	ccataggctc	cgccccctg	acgagcatca	caaaaatcga	cgctcaagtc	11520
agaggtggcg	aaacccgaca	ggactataaa	gataccaggc	gtttccccct	ggaagctccc	11580
tcgtgcgctc	tctgttccg	accctgccgc	ttaccggata	cctgtccgcc	tttctccctt	11640
cgggaagcgt	ggcgttttct	catagctcac	gctgtaggta	tctcagttcg	gtgtaggtcg	11700
ttcgctcaa	gctgggctgt	gtgcacgaac	ccccggttca	gcccgaccgc	tgcgccttat	11760
ccgtaacta	tcgtcttgag	tccaacccgg	taagacacga	cttatcgcca	ctggcagcag	11820
ccactggtaa	caggattagc	agagcgagg	atgtaggcgg	tgctacagag	ttcttgaagt	11880
ggtggcctaa	ctacggctac	actagaagga	cagtatttgg	tatctgcgct	ctgctgaagc	11940
cagttacctt	cggaaaaaga	gttggtagct	cttgatccgg	caaacaaacc	accgctggta	12000
gcggtggttt	ttttgtttgc	aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	12060
atcctttgat	cttttctacg	gggtctgacg	ctcagtgga	cgaaaactc		12109

<210> 3

<211> 12680

<212> DNA

<213> 载体

<220>

<221> misc_特征

<222> 10382

<223> n = A, T, C 或 G

<400> 3

```

acgttgtaaa acgacggcca gtgaattgta atacgactca ctatagggcg aattgggtac 60
cgggccccc ctcgaggtcg agttgggggtg gggaaaagga agaaacgcgg gcgtattggc 120
cccaatgggg tctcggtggtg gtatcgacag agtgccagcc ctgggaccga accccgcgtt 180
tatgaacaaa cgaccacaaca cccgtgcggt ttattctgtc tttttattgc cgtcatagcg 240
cgggttcctt cgggtattgt ctcttccgt cgacgatctg acggttact aaaccagctc 300
tgcttatata gacctccac cgtacacgcc taccgcccac ttgctcaat gggcgaggat 360
tgttacgaca ttttgaaaag tcccgttgat tttggtgcca aaacaaactc ccattgatgt 420
caatgggggtg gagacttggg aatccccgtg agtcaaaccg ctatccacgc ccattgatgt 480
actgccaata ccgcatcacc atggtaatag cgatgactaa tacgtagatg tactgccaag 540
taggaaagtc ccataaggtc atgtactggg cataatgcca ggcgggccat ttaccgtcat 600
tgacgtcaat agggggcgta cttggcatat gatacacttg atgtactgcc aagtgggcag 660
tttaccgtaa atactccacc cattgacgtc aatggaaagt ccctattggc gttactatgg 720
gaacatacgt cattattgac gtcaatgggc gggggctcgtt gggcggtcag ccaggcgggc 780
catttaccgt aagttatgta acgcggaact ccatatatgg gctatgaact aatgaccccg 840
taattgatta ctattaataa ctcgacggta tcatggtggc gaccggcatg gtgagctgcg 900
agaatagccg ggcgcgctgt gagccgaagt cgccccgcc ctggccactt ccggcgcgcc 960
gagtccttag gccgccaggg ggcgccggcg cgcgccaga ttggggacaa aggaagccgg 1020
gccggccgcg ttattaccat aaaaggcaaa cactggtcgg aggcgtccc gcggcgcgcg 1080
gcaggaagcc agggcccaac cccctccaa cggggcgcca gccccgctc cgcccgggtc 1140
aaacagcgac cgggtcgcgc gcgcgcacgc agcggccaca ccctcgggcg ccagcggctc 1200
gggcaggaag tggcgcaagc gcccgggcc cagaacgcac gcgcgattag cgccattgag 1260
tcccagcgcg cacgcgcaat tagcgcctc tcccagcgcg cagcagttta gcgccaaaag 1320
gaccagcgcg cagcgcctag gcgcccagc ccccaccggg cctgacgggg gctacgcgcg 1380
gcccaccgtg cgatccccat tggcaagagc ccggctcaga caaagacccc gccggttgc 1440
cccggccgga gagcggcacc cccggagcgc gcccggccga gcgcggcctc gcgcctgcga 1500
actggcgtgg ggtgtcccc atctccggag gccaggggc ttctcccgcg cccccacgg 1560
cggtcgggtt ccgccccatg cccccccgc tgcggcccag acggcggctc tgcacgggcg 1620
aagggccgcg gccgcatgcc ccggtcggct ggccgggctt acctggcggc ggggtgtggac 1680
gggcggcgga tcggcaaagg cgaggctctg tgctcgcggg cggacgcggt ctcggcggtg 1740
gtggcgcgct gcgcccgtgg gttttatagg gcgcccgcg ggccgctcga gccataaaag 1800
gcaactttcg gaacggcgca cgctgattgg cccgcgcgcg ctactcacc ggcttcgccg 1860
cacagtgcag cttttttta cccctctcc cctcctttt cgaaaaaaaa aaagagcgag 1920
agcgagattg aggaagagga ggagggagag ttttggcgtt ggccgccttg ggggtctggg 1980
cccgggggct gggggcgcgc gccgtggccc ccgcgcccc cgctgggcag tgcccgggtc 2040
ggccccgcat ggccaggcct gccccggcc tgcgcctctc tcgggcccc caccaccgc 2100
gggacatcct aggtgtggac atctcttggg cactgagcgc ccagggtggg tgggcccagg 2160
tctgcacggg ctggagggcc ctgggttctg tacgctcctg cagaaggagc tcttggaggg 2220
catggagtgg ccaggcagtc actccccctt gccgacttca gagcaactgc cctgaaagca 2280
gggcctgagg acctctggct gtggggctca gctagctaaa tgtgctgggt gggctactag 2340
ggagagacct gggcttgaga ggtagagtgt ggtgttgggg gagtcagggt gcttgcggcc 2400
attagagtcg caggaccaca ctccccagga cagggcaggg gccagcggtc cagtggctgg 2460
aggtggcccg tgatgaaggc taaaaccta cccagccgca gccctgggaa ggaagtgggc 2520
tctacagggc agggcacctt ttaccctgga gctgcctgct tttgagggta acagtcacgc 2580
ccagccaaga ccaggcctgg ggcgttagtg ggtgacctag gcaactcggg gcgggggggc 2640
tggttctaca cagcctgggt ctgggcccac cgctcgttgt atgtctgcta tgcgcagcca 2700
cagctgaact gccctccag accatctgga ggccgctggg ggactctggg gaccaagact 2760

```

ccatgtgcca	cagaggattg	ggggcggggc	ggtgctagga	actcaaagcc	agcctgggaa	2820
gaccctgtcc	ttgtcaccct	ttcttgcttt	gggtctgtcc	actgagtagc	acacaagacc	2880
gggtgggcag	ggtccgttct	gctccgggaa	tcacagactg	tgtgtaccca	ggtggtgggc	2940
atgcagcgat	cagtggcggtg	ggaccacaga	gggggcccgc	ggtacctaata	acagcttcac	3000
atggcttaaa	ataggggacc	aatgtctttt	ccaatctaag	tcccatttat	aataaagtc	3060
atgttccatt	tttaaaggac	aatcctttcg	gtttaaaacc	aggcacgatt	acccaaacaa	3120
ctcacaacgg	taaagcactg	tgaatcttct	ctgttctgca	atcccaactt	ggtttctgtc	3180
cagaaacctt	ccctctttcc	aatcggtaat	taaataacaa	aaggaaaaaa	cttaagatgc	3240
ttcaaccccg	tttctgtgaca	ctttgaaaaa	agaatcacct	cttgcaaca	cccgtcccgc	3300
acccccgcgc	ctgaagcccgc	gcgtccagag	gcctaagcgc	gggtgcccgc	ccccaccccgc	3360
gagcgcgggc	ctcgtgggtca	gcgcatccgc	gggagaaaac	aaaggcccgc	gcacgggggc	3420
tcaaggacac	tcgcccacac	cgacgcgcgc	tacccccgcg	cggccacggt	aactggcggg	3480
cgccgcagcc	tcgggacagc	cggccgcgcg	ccgccaggct	cgcgacgcgc	ggaccacgcg	3540
ccgccctccg	ggaggcccaa	gtctcgacct	agccccgcgt	ggcgctgggg	gagggggcgc	3600
ctccgcggga	acgccccgtg	gggaggggag	ggggaaatgc	gctttgtctc	gaaatggggc	3660
aaccgtcgcc	acagctccct	acccccctga	gggcagagca	gtccccccac	taactaccgc	3720
gctggccgcg	cgccaggcca	gccgcgaggc	caccgcccga	ccctccactc	cttcccgcag	3780
ctcccggcgc	ggggtccggc	gagaagggga	ggggagggga	gcggagaacc	gggccccccg	3840
gacgcgtgtg	gcatctgaag	caccaccagc	gagcgcgagc	tagagagaag	gaaagccacc	3900
gacttcaccg	cctccgagct	gctccgggtc	gcgggtctgc	agcgtctccg	gccctccgcg	3960
cctacagctc	aagccacatc	cgaaggggga	gggagccggg	agctgcgcgc	ggggccgcgc	4020
gggggagggg	tggcaccgcc	cacgcccggc	ggccacgaag	ggcggggcag	cgggcgcgcg	4080
cgcgccgggg	ggaggggccc	gcgcccgcgc	cgctgggaat	tggggcccta	gggggagggc	4140
ggagggcgcg	acgaccgcgc	cacttaccgt	tcgcccgcgt	gcgcccgggt	gtccccagg	4200
ggaggggaag	gggagggcgg	gcgaggacag	tgaccggagt	ctcctcagcg	gtggcttttc	4260
tgcttggcag	cctcagcggc	tggcgccaaa	accggactcc	gcccacttcc	tcgcccgcgc	4320
gtcgcagggg	gtggaatcct	ccagacgctg	ggggaggggg	agttgggagc	ttaaaaacta	4380
gtaccctttt	gggaccactt	tcagcagcga	actctcctgt	acaccagggg	tcagttccac	4440
agacgcgggc	caggggtggg	tcattgcggc	gtgaacaata	atltgactag	aagttgattc	4500
gggtgtttcc	ggaagggggc	gagtcfaatc	gccgagttgg	ggcacggaaa	acaaaaagg	4560
aaggctacta	agatttttct	ggcggggggt	atcattggcg	taactgcagg	gaccacctcc	4620
cggtttgagg	gggctggatc	tccaggctgc	ggattaagcc	cctcccgtcg	gcgttaattt	4680
caaactgcgc	gacgttttct	acctgccttc	gccaaaggcag	gggcccggac	cctattccaa	4740
gaggtagtaa	ctagcaggac	tctagccttc	cgcaattcat	tgagcgcatt	tacggaagta	4800
acgtcgggta	ctgtctctgg	ccgcaagggt	gggaggagta	cgcatltggc	gtaaggtggg	4860
gcgtagagcc	ttcccgccat	tggcggcgga	tagggcgttt	acgcgacggc	ctgagctagc	4920
ggaagacgcg	ttagtggggg	ggaaggttct	agaaaagcgg	cggcagcggc	tctagcggca	4980
gtagcagcag	cctccgggtc	cgtgcggagg	tgctcctcgc	agagttgttt	ctcagcagc	5040
ggcagttctc	actacagcgc	caggacgagt	ccggttcgtg	ttcgtccgcg	gagatctctc	5100
tcactctcgt	cggctgcggg	aaatcgggct	gaagcgcact	agtcgcgcgt	ggaggtaacg	5160
ggtttgaaat	caatgagtta	ttgaaaagg	catggcgagg	ccgttggcgc	ctcagtgga	5220
gtcggccagc	cgctccgtg	ggagagaggc	aggaaatcgg	accaattcag	tagcagtggg	5280
gcttaagggt	tatgaacggg	gtcttgagcg	gaggcctgag	cgtacaaaca	gcttccccac	5340
cctcagcctc	ccggcgccat	ttcccttcac	tgggggtggg	ggatggggag	ctttcacatg	5400
gcggacgctg	ccccgcctgg	gtgaaagtgg	ggcgcggagg	cgggaattct	tattcccttt	5460
ctaaagcacg	ctgcttcggg	ggccacggcg	tctcctcggc	gagcgtttcg	gcgggcagca	5520
ggctcctcgt	agcagggctg	cggagcttcc	cctccccctc	tctcccggga	accgatttgg	5580
cgcccgccat	tttcatggct	cgcttctctc	tcagcgtttt	ccttataact	cttttatttt	5640
cttagtgtgc	tttctctatc	agaagtaga	agtggttaac	tatttttttt	ttcttctcgg	5700
gctgttttca	tatcgtttcg	aggtggattt	ggagtgtttt	gtgagcttgg	atctttagag	5760
tccctgcgcac	ctcattaaag	gcgctcagcc	ttccccctga	tgaaatggcg	ccattgcggt	5820
cggaagccac	accgaagagc	ggggaggggg	ggtgctccgg	gtttgcgggc	ccggtttcag	5880
agaagatata	accaccaggg	gcgtcgggcc	gggttcaatg	cgagccgtag	gacaaagaaa	5940
ccatlttatg	tttttctctg	cttttttttc	ctttgagtaa	cggttttatc	tgggtctgca	6000
gtcagtaaaa	cgacagatga	accgcggcaa	aataaacata	aattggaagc	catcggccac	6060
gaggggcagg	gacgaagggt	gttttctggg	cgggggaggg	atattcgcgt	cagaatcctt	6120
tactgttctt	aaggattccg	tttaagttgt	agagctgact	catttttaagt	aatgttgtaa	6180

ctgagaagtt	taacccttac	gggacagatc	catggacctt	tatagatgat	tacgaggaaa	6240
gtgaaataac	gattttgtcc	ttagttatac	ttcgattaa	acatggcttc	agaggctcct	6300
tcctgtaatg	cgtatggatt	gatgtgcaaa	actgttttgg	gcctgggccg	ctctgtattt	6360
gaactttggt	acttttctca	ttttgtttgc	aatcttgggt	gaacattaca	ttgataagca	6420
taaggctca	agcgaagggg	gtctacctgg	ttatttttct	ttgaccctaa	gcacgtttat	6480
aaaataacat	tgtttaaaat	cgatagtgga	catcgggtaa	gtttgataaa	attgtgaggt	6540
aagtaatgag	tttttgcttt	ttgttagtga	tttgtaaaac	ttgttataaa	tgtacattat	6600
ccgtaatttc	agtttagaga	taacctatgt	gctgacgaca	attaagaata	aaaactagct	6660
gaaaaaatat	aaataactat	cgtgacaagt	aaccatttca	aaagactgct	ttgtgtctca	6720
taggagctag	tttgatcatt	tcagttaatt	ttttctttaa	tttttacgag	tcatgaaaac	6780
tacaggaaaa	aaaatctgaa	ctgggtttta	ccactacttt	ttaggagtgt	ggagcatgcg	6840
aatggagggg	gagctccgta	gaactgggat	gagagcagca	attaatgctg	cttgctagga	6900
acaaaaaata	attgattgaa	aattacgtgt	gactttttag	tttgattat	gcgtttgtag	6960
cagttggctc	tggatatcac	tttctctcgt	ttgaggtttt	ttaacctagt	taacttttaa	7020
gacaggtttc	cttaacattc	ataagtgcc	agaatacagc	tgtgtagtac	agcatataaa	7080
gatttcagct	ctgaggtttt	tcctattgac	ttggaaaatt	gttttgtgcc	tgtcgcttgc	7140
cacatggcca	atcaagtaag	cttcgaattc	gagctcgccc	aactccgccc	gttttatgac	7200
tagaaccaat	agtttttaat	gccaaatgca	ctgaaatccc	ctaatttgca	aagccaaacg	7260
ccccctatgt	gagtaatacg	gggacttttt	acccaatttc	ccaagcggaa	agccccctaa	7320
tacactcata	tggcatatga	atcagcacgg	tcatgcactc	taatggcggc	ccatagggag	7380
tttccacata	gggggcgttc	accatttccc	agcatagggg	tggtgactca	atggccttta	7440
ccaagtaca	ttgggtcaat	gggaggtgag	ccaatgggtt	tttccatta	ctggcaagca	7500
cactgagctca	aatgggactt	tccactgggt	tttgcccaag	tacattgggt	caatgggagg	7560
tgagccaatg	ggaaaaaccc	attgctgcc	agtacactga	ctcaataggg	actttccaat	7620
gggtttttcc	attgttggca	agcatataag	gtcaatgtgg	gtgagtcaat	agggactttc	7680
cattgtattc	tgcccagtac	ataaggtcaa	taggggtgga	atcaacagga	aagtcccatt	7740
ggagccaagt	acactgcgtc	aatagggact	ttccattggg	ttttgccag	tacataaggt	7800
caatagggga	tgagtcaatg	ggaaaaaccc	attggagcca	agtacactga	ctcaataggg	7860
actttccatt	gggttttgcc	cagtacataa	gggtcaatag	gggtgagtca	acaggaaagt	7920
cccattggag	ccaagtacat	tgagtcaata	gggactttcc	aatgggtttt	gccagtagca	7980
taaggtcaat	gggaggtgag	ccaatgggtt	tttccatta	ctggcacgta	tactgagtca	8040
ttagggactt	tccaatgggt	tttgcccagt	acataaggtc	aataggggtg	aatcaacagg	8100
aaagtcccat	tggagccaag	tacactgagt	caatagggac	tttccattgg	gttttgccca	8160
gtacaaaagg	tcaatagggg	gtgagtcaat	gggtttttcc	cattattggc	acgtacataa	8220
ggtcaatagg	ggtgagctcat	tgggtttttc	cagcaattt	aattaaaacg	ccatgtactt	8280
tcccaccatt	gacgtcaatg	ggctattgaa	actaatgcaa	cgtgaccttt	aaacggtact	8340
ttcccatagc	tgattaatgg	gaaagtaccg	ttctcgagcc	aatacacgct	aatgggaagt	8400
gaaagggcag	ccaaaacgta	acaccgcccc	ggttttcccc	tggaaattcc	atattggcac	8460
gcattctatt	ggctgagctg	cgttctacgt	gggtataaga	ggcgcgacca	gcgtcggtag	8520
cgctcgagtc	ttcggctctga	ccaccgtaga	acgcagagct	cctcgtctga	gcccgggtct	8580
agaggatccg	cctgagaaaag	gaagtgagct	gtaaaggctg	agctctctct	ctgacgtatg	8640
tagcctctgg	ttagcttctg	cactcactgt	tcttgactca	gcatggcaat	ctgatgaaat	8700
cccagctgta	agtctgcaga	aattgatgat	ctattaaaca	ataaagatgt	ccactaaaat	8760
ggaagttttt	cctgtcatac	tttgtttaaga	agggtgagaa	cagagtacct	acattttgaa	8820
tggaaggatt	ggagctacgg	gggtgggggt	gggtgaggat	tagataaatg	cctgctcttt	8880
actgaaggct	ctttactatt	gctttatgat	aatgtttcat	agttggatat	cataatttaa	8940
acaagcaaaa	ccaaattaag	ggccagctca	ttcctccaga	tccactagtt	ctagagcaaa	9000
ttctaccggg	taggggaggc	gcttttccca	aggcagctctg	gagcatgcgc	tttagcagcc	9060
ccgctgggca	cttggcgeta	cacaagtggc	ctctggcctc	gcacacattc	cacatccacc	9120
ggtaggcgcc	aaccggctcc	gttctttggg	ggccccttcg	cgccaccttc	tactcctccc	9180
ctagtcagga	agttcccccc	cgccccgag	ctcgcgtcgt	gcaggacgtg	acaaatggaa	9240
gtagcacgtc	tcactagtct	cgtgcagatg	gacagcaccg	ctgagcaatg	gaagcgggta	9300
ggcctttggg	gcagcggcca	atagcagctt	tgctccttcg	ctttctgggc	tcagaggctg	9360
ggaaggggtg	ggtccggggg	cgggctcagg	ggcgggctca	ggggcggggc	gggcccga	9420
aggtcctccg	gaggcccggc	attctgcacg	cttcaaaagc	gcacgtctgc	cgcgctgttc	9480
tcctcttctt	catctccggg	cctttcgacc	agcttaccat	gaccgagtag	aagcccacgg	9540
tgcgctctcg	cacccgcgac	gacgtcccca	gggcccgtacg	caccctcgcc	gccgcgttcg	9600

ccgactaccc	cgccacgcgc	cacaccgtcg	atccggaccg	ccacatcgag	cgggtcaccg	9660
agctgcaaga	actcttcctc	acgcgcgtcg	ggctcgacat	cggcaagggtg	tgggtcgcg	9720
acgacggcgc	cgcggtggcg	gtctggacca	cgccggagag	cgtcgaagcg	ggggcgggtg	9780
tcgcccagat	cggcccgcgc	atggccgagt	tgagcggttc	ccggctggcc	gcgcagcaac	9840
agatggaagg	cctcctggcg	ccgcaccggc	ccaaggagcc	cgcgtggttc	ctggccaccg	9900
tcggcgtctc	gcccgaccac	cagggcaagg	gtctgggcag	cgccgtcgtg	ctccccggag	9960
tggaggcggc	cgagcgcgcc	ggggtgcccc	ccttcctgga	gacctccgcg	ccccgcaacc	10020
tccccctcta	cgagcggctc	ggcttcaccg	tcaccgcca	cgtcgagggtg	cccgaaggac	10080
cgcgcacctg	gtgcatgacc	cgcaagcccc	gtgcctgacg	cccgcaccac	gaccgcagc	10140
gcccgaccga	aaggagcgc	cgacccccatg	catcgtagag	ctcgtctgatc	agcctcgact	10200
gtgccttcta	gttgccagcc	atctgttgtt	tgccccctcc	ccgtgccttc	cttgaccctg	10260
gaagggtgcca	ctcccactgt	cctttcctaa	taaaatgagg	aaattgcatc	gcattgtctg	10320
agtaggtgtc	attctattct	gggggggtggg	gtggggcagg	acagcaaggg	gggggattgg	10380
gragacaata	gcaggcatgc	tgggggggcg	gtgggggcta	tggcttctga	ggcggaaaga	10440
accagctggg	gctcgagatc	cactagtctt	agcctcgagg	ctagagcggc	ctgctctaga	10500
gcggccgcca	ccgcggtgga	gctccagctt	ttgttcctt	tagtgagggt	taatttcgag	10560
cttggcgtaa	tcatggtcat	agctgtttcc	tgtgtgaaat	tgttatccgc	tcacaattcc	10620
acacaacata	cgagccggaa	gcataaagt	taaagcctgg	ggtgccta	gagtgagcta	10680
actcacatta	attgcgttgc	gctcactgcc	cgctttccag	tcgggaaacc	tgtcgtgcca	10740
gggggtacct	aggccgggca	acaattggcg	gccggccgca	cttttcgggg	aaatgtgccc	10800
ggaaccctga	tttgtttatt	tttctaaata	cattcaaata	tgtatccgct	catgagacaa	10860
taacctctat	aaatgcttca	ataatattga	aaaaggaaaga	gtatgagtat	tcaacatttc	10920
cgtgtcgccc	ttattccctt	ttttgcggca	ttttgccttc	ctgtttttgc	tcaccagaaa	10980
acgctggtga	aagtaaaaga	tgctgaagat	cagttgggtg	cacgagtggg	ttacatcgaa	11040
ctggatctca	acagcggtaa	gatccttgag	agttttcgcc	ccgaagaacg	ttttccaatg	11100
atgagcactt	ttaaagtctt	gctatgtggc	gcggtattat	cccgtattga	cgccgggcaa	11160
gagcaactcg	gtcgcgcgat	acactattct	cagaatgact	tggttgagta	ctcaccagtc	11220
acagaaaagc	atcttacgga	tggcatgaca	gtaagagaat	tatgcagtgc	tgccataacc	11280
atgagtgata	acactgcggc	caacttactt	ctgacaacga	tcggaggacc	gaaggagcta	11340
accgcttttt	tgcaaacat	gggggatcat	gtaactcgcc	ttgatcgttg	ggaaccggag	11400
ctgaatgaag	ccataccaaa	cgacgagcgt	gacaccacga	tgctgtagc	aatggcaaca	11460
acgttgcgca	aactattaac	tggcgaacta	cttactctag	cttcccggca	acaattaata	11520
gactggatgg	aggcggataa	agttgcagga	ccacttctgc	gctcggccct	tccggctggc	11580
tggtttattg	ctgataaatc	tggagccgg	gagcgtgggt	ctcgcggtat	cattgcagca	11640
ctggggccag	atggtaagcc	ctcccgtatc	gtagttatct	acacgacggg	gagtcaggca	11700
actatggtag	aacgaaatag	acagatcgct	gagatagggtg	cctcactgat	taagcattgg	11760
taactgtcag	accctaggcc	gggcaacaat	tggcggccgg	ccctgcatta	atgaatcggc	11820
caacgcgcgg	ggagaggcgg	tttgcgtatt	gggcgctctt	ccgcttctc	gctcactgac	11880
tcgctgcgct	cggtcgcttcg	gctgcggcga	gcggtatcag	ctcactcaa	ggcggtaata	11940
cggttatcca	cagaatcagg	ggataacgca	ggaaagaaca	tgtgagcaaa	aggccagcaa	12000
aaggccagga	accgtaaaaa	ggccgcgttg	ctggcgtttt	tccataggct	ccgccccct	12060
gacgagcatc	acaaaaatcg	acgctcaagt	cagaggtggc	gaaacccgac	aggactataa	12120
agataccagg	cgtttcccc	tggaagctcc	ctcgtgcgct	ctcctgttcc	gaccctgccg	12180
cttaccggat	acctgtccgc	ctttctccct	tcgggaagcg	tggcgtttc	tcatagctca	12240
cgctgtaggt	atctcagttc	ggtgtaggtc	gttcgctcca	agctgggctg	tgtgcacgaa	12300
cccccgcttc	agcccagaccg	ctgcgcctta	tccggtaact	atcgtcttga	gtccaaccgg	12360
gtaagacacg	acttatcgcc	actggcagca	gccactggta	acaggattag	cagagcgagg	12420
tatgtaggcg	gtgctacaga	gttcttgaag	tggtggccta	actacggcta	cactagaagg	12480
acagtatttg	gtatctgcgc	tctgtgaag	ccagttacct	tcggaaaaag	agttggtagc	12540
tcttgatccg	gcaaaaaaac	caccgctgg	agcggtggtt	tttttgtttg	caagcagcag	12600
attacgcgca	gaaaaaaagg	atctcaagaa	gatcctttga	tcttttctac	ggggtctgac	12660
gctcagtgga	acgaaaactc					12680

<210> 4
 <211> 12088
 <212> DNA
 <213> 载体

<220>

<221> misc_特征

<222> 9790

<223> n = A, T, C 或 G

<400> 4

```

acgttgtaaa acgacggcca gtgaattgta atacgactca ctatagggcg aattgggtac 60
cgggccccc ctcgaggtcg agttgggggtg gggaaaagga agaaacgcgg gcgattggc 120
cccaatgggg tctcgggtggg gtatcgacag agtgccagcc ctgggaccga acccgcggt 180
tatgaacaaa cgacccaaca cccgtgcggtt ttattctgtc tttttattgc cgtcatagcg 240
cgggttcctt cgggtattgt ctcttccgt cgacggtatc aagggtggcga ccggaatgg 300
gagctgcgag aatagccggg cgcgctgtga gccgaagtcg cccccccct ggccacttcc 360
ggcgcgccga gtccttaggc cgccaggggg cgccggcgcg cgccagatt ggggacaaaag 420
gaagccgggc cggccgcggt attaccataa aaggcaaaaca ctggtcggag gcgccccgc 480
ggcgcgcggc aggaagccag gcccacacc cctcccacc gggcgccagc cccgcctccg 540
cccggttcaa acagcgaccg ggtcgcgcgc gcgcacgcag cggccacacc ctcgggcgcc 600
agcggctcgg gcaggaagtg gcgcaagcgc ccgggcccc gaacgcacgc gcgattagcg 660
ccattgagtc ccagcgcgca cgcgcaatta gcgccaattc ccagcgcgca cgcagttagc 720
gccc aaagga ccagcgcgca cgcgatggc gccacagccc ccaccgggccc tgacgggggc 780
tacccgcgcg caccgtgcg atccccattg gcaagagccc ggctcagaca aagaccccgc 840
cggttgcccc cgccccgaga gcggcacccc cggagcgcgc ccgcccagac gcgccctcgc 900
gcctgcgaac tggcgtgggg tgtccccat ctccggaggc ccaggggctt ctccgcgcc 960
ccccacggcg gtccggttcc gcccacatgc cccccgcgtg cggcccagac ggcggctctg 1020
cacgggcaaa gggccgcggc cgcacgccc ggtcggctgg ccgggcttac ctggcgccg 1080
gtgtggacgg gcggcgatc ggcaaaggcg aggtctctgt ctcgcgggcg gacgcggtct 1140
cggcgggtgg ggcgcgtcgc gccgctgggt tttatagggc gccgcccggc ccgctcagac 1200
cataaaaggc aactttcgga acggcgcacg ctgattggcc ccgcccgcct cactaccgg 1260
cttcgccgca cagtgcagca ttttttacc ccctctccc tccttttgcg aaaaaaaaa 1320
agagcgagag cgagattgag gaagaggagg agggagagtt ttggcgttgg ccgccttggg 1380
gtgctgggccc cgggggctgg gggcgcgcgc cgtggcccc gcgccccacg ctgggacgtg 1440
cccggttcgg ccccgcatgg ccaggcctgc ccccgccctg cccgtctctc gggccccca 1500
cccaccgcgg gacatcctag gtgtggacat ctcttgggca ctgagcgcgc aggtgggggtg 1560
ggcagggtc tgcacgggtg ccaggccctt gggttctgta cgctcctgca gaaggagctc 1620
ttggaggcca tggagtggcc aggcagtcac tcccccttgc cgacttcaga gcaactgcc 1680
tgaaagcagg gcctgaggac ctctggctgt ggggctcagc tagctaaatg tgctgggtgg 1740
gtcactaggg agagacctgg gcttgagagg tagagtgtgg tgttggggga gtcaggtggc 1800
ttgcggccat tagagtcgca ggaccacact ccccaggaca gggcaggggc cagcggcca 1860
gtggctggag gtggcccgtg atgaaggcta caaacctacc cagccgcagc cctgggaagg 1920
aagtgggctc tacagggcag ggcaccttt accctggagc tgctgcttt tgagggtaac 1980
agtcacgccc agccaagacc aggcctgggg cgttagtggg tgacctaggc actgcggggc 2040
gggggggctg ggtctacaca gcctgggtct gggcccaccg tccgttgtat gtctgctatg 2100
cgcagccaca gctgaactgc cctcccagac catctggagg ccgctggggg actctgggga 2160
ccaagactcc atgtgccaca gaggattggg ggcggggcgg tgctaggaac tcaaagccag 2220
cctgggaaga cctgtcctt gtcaccctt cttgccttgg gtctgtccac tgagtagcac 2280
acaagaccgg gtgggcaggg tccgttctgc tccgggaatc acagactgtg tgtaccagg 2340
tgggtggcat gcagcgtca gtggcgtggg accacagagg gggcccgcgg tacctaaaac 2400
agcttcacat ggcttaaaat aggggaccaa tgccttttcc aatctaagtc ccatttataa 2460
taaagtccat gttccatttt taaaggacaa tcctttcggg ttaaaaaccag gcacgattac 2520
ccaacaact cacaacggta aagcactgtg aatcttctct gttctgcaat cccaacttgg 2580
tttctgctca gaaaccctcc ctctttccaa tcggttaatta aataacaaaa ggaaaaaact 2640
taagatgctt caaccccggt tcgtgacact ttgaaaaaag aatcacctct tgcaaacacc 2700
cgctcccagc ccccgccgct gaagcccggc gtccagaggc ctaagcgcgg gtgcccgcgc 2760
ccaccgggga gcgcgggccc cgtggtcagc gcacccgcgg ggagaaacaa aggcgcggc 2820
acgggggctc aagggcactg cgccacaccg cacgcgccta ccccgcgcg gccacgttaa 2880
ctggcggctc ccgagcctc gggacagccg gccgcgcgcc gccaggctcg cggacgcggg 2940
accacgcgcc gccctccggg aggcccaagt ctcgaccag cccgcgctgg cgctggggga 3000

```

gggggcgct	ccgccggaac	gcgggtgggg	gaggggaggg	ggaaatgcmc	tttgtctcga	3060
aatggggcaa	ccgtcgccac	agctccctac	cccctcgagg	gcagagcagt	ccccccacta	3120
actaccgggc	tggccgcgcg	ccaggccagc	cgcgaggcca	ccgcccgacc	ctccactcct	3180
tcccgcagct	cccggcgcgg	ggtccggcga	gaaggggagg	ggaggggagc	ggagaaccgg	3240
gcccccgga	cgcgtgtggc	atctgaagca	ccaccagcga	gcgagagcta	gagagaagga	3300
aagccaccga	cttcaccgcc	tccgagctgc	tccgggtcgc	gggtctgcag	cgtctccggc	3360
cctccgcgcc	tacagctcaa	gccacatccg	aagggggagg	gagccgggag	ctgcgcgcgg	3420
ggccgcggg	gggaggggtg	gcaccgcccc	cgccgggcgg	ccacgaaggg	cggggcagcg	3480
ggcgcgcgcg	cggcgggggg	aggggcgggc	gcccgcggcg	ctgggaattg	ggccctagg	3540
gggagggcgg	aggcggcgac	gaccgcggca	cttaccgttc	gcggcgtggc	gcccgtggt	3600
ccccaaaggg	agggaagggg	gaggcggggc	gaggacagtg	accggagtct	cctcagcggg	3660
ggcttttctg	cttggcagcc	tcagcggctg	gcgccaaaac	cggactccgc	ccacttcttc	3720
gcccgcgggt	gcgaggggtg	ggaatcctcc	agacgcctgg	ggagggggag	ttgggagctt	3780
aaaaactagt	acccctttgg	gaccactttc	agcagcgaac	tctcctgtac	accaggggtc	3840
agttccacag	acgcgggcca	ggggtgggtc	attgcggcgt	gaacaataat	ttgactagaa	3900
gttgattcgg	gtgtttccgg	aaggggccga	gtcaatccgc	cgagttgggg	cacggaanaac	3960
aaaaagggaa	ggctactaag	atthttctgg	cgggggttat	cattggcgtg	actgcagggg	4020
ccacctcccg	ggttaggggg	gctggatctc	caggctgcgg	attaagcccc	tcccgtcggc	4080
gttaatttca	aactgcgcga	cgthttctac	ctgccttcgc	caaggcaggg	gccgggacct	4140
tattccaaga	ggtagtaact	agcaggactc	tagccttccg	caattcattg	agcgcattta	4200
cggaagtaac	gtcgggtact	gtctctggcc	gcaaggggtg	gaggagtacg	catttggcgt	4260
aaggtggggc	gtagagcctt	cccgccattg	gcggcggata	ggcggtttac	gcgacggcct	4320
gacgtagcgg	aagacgcggt	agtggggggg	aaggttctag	aaaagcggcg	gcagcggctc	4380
tagcggcag	agcagcagcg	ccgggtcccg	tgcgaggtg	ctcctcgcag	agttgtttct	4440
cgagcagcgg	gatttctcac	tacagcggca	ggacgagtcc	ggttcgtggt	cgccgcggga	4500
gatctctctc	atctcgctcg	gctgcgggaa	atcgggctga	agcgcactgag	tccgcgatgg	4560
aggtaacggg	tttgaatca	atgagttatt	gaaaagggca	tggcgaggcc	gttggcgcct	4620
cagtggaaat	cggccagccg	cctccgtggg	agagaggcag	gaaatcggac	caattcagta	4680
gcagtggggc	ttaaggttta	tgaacggggg	cttgagcggg	ggcctgagcg	tacaaacagc	4740
ttccccacct	tcagcctccc	ggcgcattt	cccttactg	ggggtggggg	atggggagct	4800
ttcacatggc	ggacgctgcc	ccgctggggg	gaaagtgggg	cgcgaggggc	ggaattctta	4860
ttccctttct	aaagcacgct	gcttcggggg	ccacggcgtc	tctcggcgca	gcgttccggc	4920
gggcagcagg	tcctcgtgag	cgaggctgcg	gagcttcccc	tccccctctc	tcccgggaac	4980
cgatttggcg	gccgccattt	tcattgctcg	ccttctctct	agcgttttcc	ttataactct	5040
tttattttct	tagtgtgctt	tctctatcaa	gaagtagaag	tggttaacta	ttttttttt	5100
cttctcgggc	tgthttcata	tcgthttcag	gtggatttgg	agtgthttgt	gagcttggat	5160
cttttagagtc	ctgcgcacct	cattaaaggc	gctcagcctt	cccctcgatg	aaatggcgcc	5220
attgcgthtcg	gaagccacac	cgaagagcgg	ggaggggggg	tgctccgggt	ttgcggggccc	5280
ggtthtcagag	aagatatcac	caccagggc	gtcgggcccg	gthcaatgcm	agccgtagga	5340
caaagaaacc	atthttatgtt	thttctgtct	thtttttctt	ttgagtaacg	gthttatctg	5400
ggtctgcag	cagtaaaacg	acagatgaac	cgcgcaaaa	taaacataaa	ttggaagcca	5460
tcggccacga	ggggcagggg	cgaaggtggt	thttctggcg	ggggagggat	attcgcgtca	5520
gaatccttta	ctgttcttaa	ggattccgth	taagthgtag	agctgactca	thttaaagtaa	5580
tgthgttact	gagaagthta	acccttacgg	gacagatcca	tggacctthta	tagatgatta	5640
cgaggaaagt	gaaataacga	thttgtcctt	agthtatactt	cgattaaaac	atggcttcag	5700
aggctccttc	ctgtaatgcm	tatggattga	tgtgcaaaa	tgthttgggc	ctgggcccgt	5760
ctgattthga	actthgttac	thttctcatt	ttgthttgcaa	tcttggthtga	acattacatt	5820
gataagcata	aggtctcaag	cgaagggggg	ctacctggth	atthttctth	gacctaaagc	5880
acgthttataa	aataacattg	thtaaaatcg	atagthggaca	tcgggtaagt	ttggataaat	5940
tgtgaggtaa	gtaatgagth	thttgctthtt	gthtagthatt	tgtaaaactt	gthataaatg	6000
tacattatcc	gtaatthtcag	thttagagata	acctatgthc	tgacgacaat	taagaataaa	6060
aactagctga	aaaaatgaaa	ataactatcg	tgacaagtaa	ccatttcaaa	agactgctth	6120
gtgtctcata	ggagctagth	tgatcattth	agthtaathth	thctthtaath	thttacgagth	6180
atgaaaacta	caggaaaaaa	aatctgaact	gggthttacc	actactthth	aggagthggg	6240
agcatgcmgaa	tggaggggaga	gctccgtaga	actgggtaga	gagcagcaat	taatgctgct	6300
tgctaggaac	aaaaataaat	tgattgaaaa	ttacgthgta	ctthtttagth	tgattatgcm	6360
gthttgtagca	gthggctcctg	gatatcactt	tctctcgtth	gagthththth	aacctagthta	6420

acttttaaga	caggtttct	taacattcat	aagtgccag	aatacagctg	tgtagtacag	6480
catataaaga	tttcagctct	gaggtttttc	ctattgactt	ggaaaattgt	tttgtgctg	6540
tcgcttgcca	catggccaat	caagtaagct	tcgaattcga	gctcgccaa	ctccgccgt	6600
tttatgacta	gaaccaatag	tttttaatgc	caaatgcact	gaaatcccct	aatttgcaaa	6660
gccaaacgcc	ccctatgtga	gtaatacggg	gactttttac	ccaatttccc	aagcggaaag	6720
ccccctaata	cactcatatg	gcatatgaat	cagcacggtc	atgcactcta	atggcgccc	6780
atagggactt	tccacatagg	gggcgttcac	catttcccag	cataggggtg	gtgactcaat	6840
ggcctttacc	caagtacatt	gggtcaatgg	gaggtaaagc	aatgggtttt	tcccattact	6900
ggcaagcaca	ctgagtcaaa	tgggactttc	cactgggttt	tgcccaagta	cattgggtca	6960
atggggagtg	agccaatggg	aaaaacccat	tgctgccaa	tacactgact	caatggggac	7020
tttccaatgg	gtttttccat	tggtggcaag	catataaggt	caatgtgggt	gagtcaatag	7080
ggactttcca	ttgtattctg	cccagtacat	aagggtcaata	gggggtgaat	caacaggaaa	7140
gtcccattgg	agccaagtac	actgctcaa	tagggacttt	ccattgggtt	ttgccagta	7200
cataaggtca	ataggggatg	agtcaatggg	aaaaacccat	tggagccaag	tacactgact	7260
caatagggac	tttcattggg	gttttgccca	gtacataagg	tcaatagggg	gtgagtcaac	7320
aggaaagtcc	cattggagcc	aagtacattg	agtcaatagg	gactttccaa	tgggttttgc	7380
ccagtacata	aggtcaatgg	gaggtaaagc	aatgggtttt	tcccattact	ggcacgtata	7440
ctgagtcatt	agggactttc	caatgggttt	tgcccagtac	ataaggtcaa	taggggtgaa	7500
tcaacaggaa	agtcccattg	gagccaagta	cactgagtca	atagggactt	tccattgggt	7560
tttgcccagt	acaaaaggtc	aataggggtt	gagtcaatgg	gtttttccca	ttattggcac	7620
gtacataaag	tcaatagggg	tgagtcattg	ggtttttcca	gccaatttaa	ttaaaacgcc	7680
atgtactttc	ccaccattga	cgtcaatggg	ctattgaaac	taatgcaacg	tgaccttaa	7740
acggtaactt	cccatagctg	attaatggga	aagtaccgtt	ctcgagccaa	tacacgtcaa	7800
tgggaagtga	aagggcagcc	aaaacgtaac	accgcccgg	ttttcccctg	gaaattccat	7860
attggcacgc	attctattgg	ctgagctgcg	ttctacgtgg	gtataagagg	cgcgaccagc	7920
gtcggtagcg	tcgcagtctt	cggctctgacc	accgtagaac	gcagagctcc	tcgctgcagc	7980
ccgggtctag	aggatccgcc	tgagaaagga	agtgagctgt	aaaggctgag	ctctctctct	8040
gacgtatgta	gcctctgggt	agcttcgtca	ctcactgttc	ttgactcagc	atggcaatct	8100
gatgaaatcc	cagctgtaag	tctgcagaaa	ttgatgatct	attaacaat	aaagatgtcc	8160
actaaaatgg	aagtttttcc	tgtcatactt	tgtaagaag	ggtgagaaca	gagtacctac	8220
atthtgaatg	gaaggattgg	agctacgggg	gtgggggtgg	ggtgggatta	gataaatgcc	8280
tgctctttac	tgaaggctct	ttactattgc	tttatgataa	tgtttcatag	ttggatatca	8340
taatttaaac	aagcaaaacc	aaattaaggg	ccagctcatt	cctccagatc	cactagttct	8400
agagcaaatt	ctaccgggta	ggggagggcg	ttttcccagg	gcagtctgga	gcatgcgctt	8460
tagcaccccc	gtgggcact	tggcgctaca	caagtggcct	ctggcctcgc	acacattcca	8520
catccaccct	taggcgcaa	ccggctccgt	tctttggtgg	ccccttcgcg	ccaccttcta	8580
ctcctcccct	agtcaggaag	ttcccccccg	ccccgagct	cgctcgtgc	aggacgtgac	8640
aaatggaagt	agcacgtctc	actagtctcg	tgcagatgga	cagcaccgct	gagcaatgga	8700
agcgggtagg	cctttggggc	agcggccaat	agcagctttg	ctccttcgct	ttctgggctc	8760
agaggctggg	aaggggtggg	tccggggcg	ggctcagggg	cgggctcagg	ggcggggcgg	8820
gcgcccgaag	gtcctccgga	ggcccggcat	tctgcacgct	tcaaaagcgc	acgtctgccg	8880
cgctgttctc	ctcttctca	tctccgggcc	tttcgaccag	cttaccatga	ccgagtacaa	8940
gcccacggtg	cgctcgcga	cccgcgacga	cgtcccaggg	gccgtacgca	ccctcgcgc	9000
cgcttcgccc	gactacccc	ccacgcgcca	caccgtcgat	ccggaccgcc	acatcgagcg	9060
ggtcaccgag	ctgcaagaac	tcttctcac	gcgctcggg	ctcgacatcg	gcaagggtgtg	9120
ggtcgaggac	gacggcgccg	cgggtggcgg	ctggaccacg	ccggagagcg	tcgaagcggg	9180
ggcgggtgtc	gccgagatcg	gcccgcgcat	ggccgagttg	agcggttccc	ggctggcccgc	9240
gcagcaacag	atggaaggcc	tcttggcgcc	gcaccggccc	aaggagcccg	cgtggttctt	9300
ggccaccgtc	ggcgtctcgc	ccgaccacca	gggcaagggg	ctgggcagcg	ccgtcgtgct	9360
ccccggagtg	gaggcggccg	agcgcgcccg	ggtgcccgcc	ttcctggaga	cctccgccc	9420
ccgcaacctc	cccttctacg	agcggctcgg	cttcaccgct	accgcccagc	tcgaggtgcc	9480
cgaaggaccg	cgcacctggt	gcatgaccgg	caagcccggg	gcctgacgcc	cgccccacga	9540
cccgcagcgc	ccgaccgaaa	ggagcgcacg	accccatgca	tcgtagagct	cgctgatcag	9600
cctcgactgt	gccttctagt	tgccagccat	ctgttgtttg	cccctcccc	gtgccttctt	9660
tgaccctgga	aggtgccact	cccactgtcc	tttctaata	aaatgaggaa	attgcatcgc	9720
attgtctgag	taggtgtcat	tctattctgg	gggggtgggg	ggggcaggac	agcaaggggg	9780
gggattgggr	agacaatagc	agggcatgctg	ggggggcggg	gggggctatg	gcttctgagg	9840

```

cggaaagaac cagctggggc tcgagatcca ctagttctag cctcgaggct agagcggcct 9900
gctctagagc ggccgccacc gcggtggagc tccagctttt gttcccttta gtgaggggta 9960
atctcgagct tggcgtaatc atggtcatag ctgtttcctg tgtgaaattg ttatccgctc 10020
acaattccac acaacatacg agccggaagc ataaagtgtg aagcctgggg tgcctaatag 10080
gtgagctaac tcacattaat tgcgttgccg tccactgccc ctttccagtc gggaaacctg 10140
tcgtgccagg gggtagctag gccgggcaac aattggcggc cggccgcact tttcggggaa 10200
atgtgcgcgg aaccctattt tgtttatttt tctaaataca ttcaaatatg tatccgctca 10260
tgagacaata accctgataa atgcttcaat aatattgaaa aaggaagagt atgagtattc 10320
aacatttccg tgcgcgccct attccccttt ttgcggcatt ttgccttctt gtttttgctc 10380
accagaaaac gctgggtgaaa gtaaaagatg ctgaagatca gttgggtgca cgagtgggtt 10440
acatcgaact ggatctcaac agcggtaaga tccttgagag ttttcgcccc gaagaacgct 10500
ttccaatgat gagcactttt aaagtctctg tatgtggcgc ggtattatcc cgtattgacg 10560
ccgggcaaga gcaactcggc cgccgcatac actattctca gaatgacttg gttgagtact 10620
caccagtcac agaaaagcat cttacggatg gcatgacagt aagagaatta tgcagtgctg 10680
ccataaccat gagtgataac actgcggcca acttacttct gacaacgatc ggaggaccga 10740
aggagctaac cgcttttttg cacaacatgg gggatcatgt aactcgctt gatcgttggg 10800
aaccggagct gaatgaagcc ataccaaaac acgagcgtga caccacgatg cctgtagcaa 10860
tggcaacaac gttgcgcaaa ctattaactg gcgaactact tactctagct tcccggcaac 10920
aatatataga ctggatggag gcggataaag ttgcaggacc acttctgcgc tcggcccttc 10980
cggctggctg gtttattgct gataaatctg gagccgggtg gcgtgggtct cgcggtatca 11040
ttgcagcact ggggcccagat ggtaagccct cccgatcgt agttatctac acgacgggga 11100
gtcaggcaac tatggatgaa cgaaatagac agatcgtgga gatagggtgcc tccactgatta 11160
agcattggta actgtcagac cctaggccgg gcaacaattg gcggccggcc ctgcattaat 11220
gaatcgcca acgcgcgggg agaggcgggt tgcgtattgg gcgctcttcc gcttctcgc 11280
tccactgact gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 11340
cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaag 11400
gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc 11460
gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgcagac 11520
gactataaag ataccaggcg tttcccctg gaagetccct cgtgcgctct cctgttccga 11580
ccctgccgct taccggatac ctgtccgctt ttctcccttc gggaaagcgtg gcgctttctc 11640
atagctcac ctgtaggat ctccagttcgg tgtaggctgt tcgctccaag ctgggctgtg 11700
tgcacgaacc ccccgctcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt 11760
ccaaccgggt aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca 11820
gagcgaggta tgtaggcggg gctacagagt tcttgaagtg gtggcctaac tacggctaca 11880
ctagaaggac agtatttggg atctgcgctc tgctgaagcc agttaccttc ggaaaaagag 11940
ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca 12000
agcagcagat tacgcgcaga aaaaaggat ctcaagaaga tcctttgatc ttttctacgg 12060
ggtctgacgc tcagtggaac gaaaactc                                     12088

```

<210> 5

<211> 12704

<212> DNA

<213> 载体

<220>

<221> misc_特征

<222> 9605

<223> n = A, T, C 或 G

<400> 5

```

acgttgtaaa acgacggcca gtgaattgta atacgactca ctatagggcg aattgggtac 60
cgggcccccc ctcgaggtcg agttgggggtg gggaaaagga agaaacgcgg gcgtattggc 120
cccaatgggg tctcgggtggg gtatcgacag agtgccagcc ctgggaccga accccgcggt 180
tatgaacaaa cgaccocaaca cccgtgcggt ttattctgtc tttttattgc cgtcatagcg 240
cgggttcctt ccggtattgt ctcttccgct cgactctaga cccgggctgc agcgaggagc 300
tctgcgctct acgggtggtca gaccgaagac tgcgacggta ccgacgctgg tcgcgctct 360

```

tatacccacg	tagaacgcag	ctcagccaat	agaatgcgtg	ccaatatgga	atthccaggg	420
gaaaaccggg	gcggtgttac	gttttggctg	ccctttcact	tcccattgac	gtgtattggc	480
tcgagaacgg	tactttccca	ttaatcagct	atgggaaagt	accgtttaa	ggtcacgttg	540
cattagtttc	aatagcccat	tgacgtcaat	ggtgggaaag	tacatggcgt	tttaattaaa	600
ttggctggaa	aaacccaatg	actcaccctt	attgacctta	tgtacgtgcc	aataatggga	660
aaaaccatt	gactcaccct	ctattgacct	tttgtactgg	gcaaaacca	atggaaagt	720
cctattgact	cagtgtactt	ggctccaatg	ggactttcct	gttgattcac	ccctattgac	780
cttatgtact	gggcaaaacc	cattggaaag	tccctaata	ctcagtatac	gtgccagtaa	840
tgggaaaaac	ccattggctt	acctcccatt	gaccttatgt	actgggcaaa	acccattgga	900
aagtccctat	tgactcaatg	tacttggctc	caatgggact	ttcctgttga	ctcaccctt	960
attgacctta	tgtactgggc	aaaacccaat	ggaaagtccc	tattgagtca	gtgtacttgg	1020
ctccaatggg	tttttcccat	tgactcatcc	cctattgacc	ttatgtactg	ggcaaaacc	1080
aatggaaagt	ccctattgac	gcagtgtact	tggctccaat	gggactttcc	tgttgattca	1140
ccccctattg	accttatgta	ctgggcagaa	tacaatggaa	agtccctatt	gactcaccca	1200
cattgacctt	atatgcttgc	caacaatgga	aaaaccatt	ggaaagtccc	tattgagtca	1260
gtgtacttgg	cagcaatggg	tttttcccat	tggctcacct	cccattgacc	caatgtactt	1320
gggcaaaacc	cagtggaaag	tcccatttga	ctcagtgtgc	ttgccagtaa	tgggaaaaac	1380
ccattggctt	acctcccatt	gacccaatgt	acttgggtaa	aggccattga	gtcaccaccc	1440
ctatgctggg	aaatggtgaa	cgccccctat	gtggaaagt	cctatgggcc	gccattagag	1500
tgcatgaccg	tgtctgattca	tatgccatat	gagtgtatta	gggggctttc	cgcttgggaa	1560
attgggtaaa	aagtccccgt	attactcaca	tagggggcgt	ttggctttgc	aaattagggg	1620
atthcagtgc	atthggcatt	aaaaactatt	ggttctagtc	ataaaacggg	cgagttggg	1680
cgagctcgaa	ttcaaacgac	tcgacgggat	caaggtggcg	accggaatgg	tgagctgcga	1740
gaatagccgg	gcgctgtgtg	agccgaagtc	gcccccgccc	tggccacttc	cggcgcccg	1800
agtccttagg	ccgcccagggg	gcgccggcgc	gcgcccagat	tggggacaaa	ggaagccggg	1860
ccggcccgct	tattaccata	aaaggcaaac	actggtcgga	ggcgtccccg	cggcgccgg	1920
caggaagcca	ggcccccaacc	ccctcccaac	cgggcgccag	ccccgcctcc	gccccgttca	1980
aacagcgacc	gggtcgccgcg	cgcgcacgca	gcgccacac	cctcggggcg	cagcggctcg	2040
ggcaggaagt	ggcgcaagcg	cccgggcccc	agaacgcacg	cgcgattagc	gccattgagt	2100
cccagcgcgc	acgcgcaatt	agcgcgaatt	cccagcgcgc	acgcagttag	cgcccaaagg	2160
accagcgcgc	acgcgcatgg	cgccccagcc	cccaccgggc	ctgacggggg	ctacgcccg	2220
cccaccgtgc	gatccccatt	ggcaagagcc	cggtcagac	aaagaccccg	ccggttgccc	2280
ccgccccgag	agcggcacc	ccggagcgcg	cccggcccag	cgcggcctcg	cgcttgcgaa	2340
ctggcgtggg	gtgtccccca	tctccggagg	cccaggggct	tctcccgcgc	ccccacggc	2400
ggtccggttc	cgccccatgc	gccccccgct	gcgccccaga	cggcggctct	gcacgggcca	2460
agggccgcgg	ccgcctgccc	cggctggctg	gcccggctta	cctggcggcg	ggtgtggagc	2520
ggcggcggat	cggcaaaagg	gaggtctgtg	gctcgcgggc	ggacgcggtc	tcggcggtag	2580
tggcgcgtcg	cgccgctggg	ttttataggg	cgccgcgcgc	gcccgtcgag	ccataaaagg	2640
caactttcgg	aacggcgcac	gctgattggc	cccgcgcgc	tactcaccg	gcttcgccgc	2700
acagtgcagc	atthtttttac	cccctctccc	ctccttttgc	gaaaaaaaa	aagagcgaga	2760
gagagattga	ggaagaggag	gagggagagt	tttggcgttg	gccccttgg	ggtgctgggc	2820
ccgggggctg	ggggcgcgcg	ccgtggcccc	cgcgccccac	gctgggcagt	gccccgttcc	2880
gccccgcgat	gccaggcctg	ccccggcct	gcccgtctct	cgggcccccc	accaccgcg	2940
ggacatccta	ggtgtggaca	tctcttgggc	actgagcgc	caggtgggg	gggcccagg	3000
ctgcacgggt	gccagggcc	tgggttctgt	acgtcctgc	agaaggagct	cttggagggc	3060
atggagtggc	caggcagtca	ctccccctt	ccgacttcag	agcaactgcc	ctgaaagcag	3120
ggcctgagga	cctctggctg	tggggctcag	ctagctaaat	gtgctgggtg	ggtcactagg	3180
gagagacctg	ggcttgagag	gtagagtgtg	gtgttggggg	agtcaggtgg	cttgcggcca	3240
ttagagtgcg	aggaccacac	tccccaggac	agggcagggg	ccagcggctc	agtggttggg	3300
ggtggcccg	gatgaaggct	acaaacctac	ccagccgcag	ccctgggaag	gaagtgggct	3360
ctacagggca	gggcaccttt	tacctggag	ctgcctgctt	ttgagggtaa	cagtcacgcc	3420
cagccaagac	caggcctggg	gcgttagtgg	gtgacctagg	cactgcgggg	cgggggggct	3480
gggtctacac	agcctgggtc	tgggcccacc	gtccgttgta	tgtctgctat	gcgcagccac	3540
agctgaactg	ccctcccaga	ccatctggag	gcccgtgggg	gactctgggg	accaagactc	3600
catgtgccac	agaggattgg	gggcccggcg	gtgctaggaa	ctcaaagcca	gcctgggaag	3660
accctgtcct	tgtcaccctt	tcttgccttg	ggtctgtcca	ctgagtagca	cacaagaccg	3720
ggtgggcagg	gtccgttctg	ctccgggaat	cacagactgt	gtgtaccag	gtgggtggga	3780

tgcagcgatc	agtggcgtgg	gaccacagag	ggggcccgcg	gtacctaaaa	cagcttcaca	3840
tggcttaaaa	taggggacca	atgtcttttc	caatctaagt	cccatttata	ataaagtcca	3900
tgttccattt	ttaaaggaca	atcctttcgg	tttaaaacca	ggcacgatta	cccaaacaac	3960
tcacaacggt	aaagcactgt	gaatctttctc	tgttctgcaa	tcccaacttg	gtttctgctc	4020
agaaaccctc	cctctttcca	atcggtaatt	aaataacaaa	aggaaaaaac	ttaagatgct	4080
tcaaccccgt	ttcgtgacac	tttgaaaaaa	gaatcacctc	ttgcaaacac	ccgctcccga	4140
ccccgcgcg	tgaagcccgg	cgtccagagg	cctaagcgcg	ggtgcccgcg	cccacccggg	4200
agcgcggggc	tctgtgtcag	cgcatccgcg	gggagaaaca	aaggccgcgg	cacgggggct	4260
caagggcact	gcgccacacc	gcacgcgcct	accccgcgcg	ggccacgta	actggcggct	4320
gccgcagcct	cgggacagcc	ggccgcgcgc	cgccaggctc	gcggacgcgg	gaccacgcgc	4380
cgccctccgg	gaggcccaag	tctcgaccca	gccccgcgtg	gcgctggggg	agggggcgcc	4440
tccgcgggaa	cgcggtggg	ggaggggagg	gggaaatgcg	ctttgtctcg	aaatggggca	4500
accgtcgcca	cagctcccta	ccccctcgag	ggcagagcag	tccccccact	aactaccggg	4560
ctggccgcgc	gccaggccag	ccgcgaggcc	accgcccgcg	cctccactcc	ttcccgcagc	4620
tcccggcgcg	gggtccggcg	agaaggggag	gggaggggag	cggagaaccg	ggcccccggg	4680
acgcgtgtgg	catctgaagc	accaccagcg	agcgagagct	agagagaagg	aaagccaccg	4740
acttcaccgc	ctccgagctg	ctccgggtcg	cgggtctgca	gcgtctccgg	ccctccgcgc	4800
ctacagctca	agccacatcc	gaagggggag	ggagccggga	gctgcgcgcg	gggcccgcgc	4860
ggggaggggt	ggcaccgccc	acgcccggcg	gccacgaagg	gcggggcagc	gggcccgcgc	4920
gcggcggggg	gaggggcccg	cgccgcgcgc	gctgggaatt	ggggccctag	ggggagggcg	4980
gaggcgccga	cgaccgcggc	acttacggtt	cgccggcgtg	cgcccgggtg	tcccccaagg	5040
gagggaaagg	ggaggcgggg	cgaggacagt	gaccggagtc	tcctcagcgg	tggtttttct	5100
gcttggcagc	ctcagcggct	ggcgccaaaa	ccggactccg	cccacttcct	cgcccgcggc	5160
tgcgagggtg	tggaatcctc	cagacgctgg	gggaggggga	gttgggagct	taaaaactag	5220
tacccttttg	ggaccacttt	cagcagcga	ctctcctgta	caccaggggt	cagttccaca	5280
gacgcggggc	aggggtgggt	cattgcggcg	tgaacaataa	tttgactaga	agttgattcg	5340
ggtgtttccg	gaaggggccc	agtcaatccg	ccgagttggg	gcacggaaaa	caaaaagggg	5400
aggctactaa	gattttttctg	gcgggggtta	tcattggcgt	aactgcaggg	accacctccc	5460
gggttgaggg	ggctggatct	ccaggctcgc	gattaagccc	ctcccgtcgg	cgttaatctc	5520
aaactgcgcg	acgtttctca	cctgccttcg	ccaaggcagg	ggccgggacc	ctattccaag	5580
aggtagtaac	tagcaggact	ctagccttcc	gcaattcatt	gagcgcattt	acggaagtaa	5640
cgtcgggtac	tgtctctggc	cgcaaggggtg	ggagggagtac	gcatttgccg	taaggtgggg	5700
cgtagagcct	tcccgccatt	ggcggcggat	agggcgttta	cgcgacggcc	tgacgtagcg	5760
gaagacgcgt	tagtgggggg	gaaggttcta	gaaaagcggc	ggcagcggct	ctagcggcag	5820
tagcagcagc	gccgggtccc	gtgcggaggt	gctcctcgca	gagttgtttc	tcgagcagcg	5880
gcagttctca	ctacagcggc	aggacgagtc	cggttcgtgt	tcgtccgcgg	agatctctct	5940
catctcgctc	ggctgcggga	aatcgggctg	aagcgactga	gtccgcgatg	gaggtaacgg	6000
gtttgaaatc	aatgagttat	tgaaaagggc	atggcagggc	cgttggcgcc	tcagtggaag	6060
tcggccagcc	gcctccgtgg	gagagaggca	ggaaaatcgg	ccaattcagt	agcagtgggg	6120
cttaaggttt	atgaacgggg	tcttgagcgg	aggcctgagc	gtacaaaacag	cttccccacc	6180
ctcagcctcc	cggcgccatt	tcccttcaact	gggggtgggg	gatggggagc	tttccatagg	6240
cggacgcctg	cccgtcgggg	tgaaagtggg	gcgcggaggc	gggaattctt	attccctttc	6300
taaagcacgc	tgcttcgggg	gccacggcgt	ctcctcggcg	agcgtttcgg	cgggcagcag	6360
gtcctcgtga	gcgaggctgc	ggagcttccc	ctccccctct	ctcccgggaa	ccgatttggc	6420
ggccgccatt	ttcatggctc	gccttcctct	cagcgttttc	cttataactc	ttttattttc	6480
ttagtgtgct	ttctctatca	agaagtagaa	gtgggttaact	atTTTTTTTT	tcttctcggg	6540
ctgttttcat	atcgtttcga	ggtggatttg	gagtgttttg	tgagcttggg	tcttttagagt	6600
cctgcgcacc	tcattaaagg	cgctcagcct	tcccctcgat	gaaatggcgc	cattgcgttc	6660
ggaagccaca	ccgaagagcg	gggagggggg	gtgctccggg	tttgccggcc	cggtttcaga	6720
gaagatatca	ccaccagggg	cgtcggggcg	ggttcaatgc	gagccgtagg	acaagaaac	6780
cattttatgt	ttttctgctc	ttttttttcc	tttgagtaac	ggttttatct	gggtctgcag	6840
tcagtaaaac	gacagatgaa	ccgcggcaaa	ataaacataa	attggaagcc	atcgccacag	6900
aggggcaggg	acgaaggtgg	ttttctgggc	gggggagggg	tattcgcgtc	agaatccttt	6960
actgttctta	aggattccgt	ttaagttgta	gagctgactc	atTTTaaagta	atgttggttac	7020
tgagaagttt	aacccttacg	ggacagatcc	atggaccttt	atagatgatt	acgaggaaaag	7080
tgaaataacg	atTTTgtcct	tagttatact	tcgattaaaa	catggcttca	gaggctcctt	7140
cctgtaatgc	gtatggattg	atgtgcaaaa	ctgttttggg	cctgggcccgc	tctgtatttg	7200

aactttgtta	cttttctcat	tttgtttgca	atcttgggtg	aacattacat	tgataagcat	7260
aaggtctcaa	gcgaaggggg	tctacctggt	tatTTTTctt	tgaccctaag	cacgtttata	7320
aaataacatt	gtttaaaatc	gatagtggac	atcgggtaag	tttggataaa	ttgtgaggta	7380
agtaatgagt	ttttgctttt	tgttagtgat	ttgtaaaact	tgttataaat	gtacattatc	7440
cgtaatttca	gtttagagat	aacctatgtg	ctgacgacaa	ttaagaataa	aaactagctg	7500
aaaaaatgaa	aataactatc	gtgacaagta	accatttcaa	aagactgctt	tgtgtctcat	7560
aggagctagt	ttgatcattt	cagttaattt	tttctttaat	ttttacgagt	catgaaaact	7620
acaggaaaaa	aaatctgaac	tgggttttac	cactactttt	taggagttgg	gagcatgcca	7680
atggagggag	agctccgtag	aactgggatg	agagcagcaa	ttaatgctgc	ttgttaggaa	7740
caaaaaataa	ttgattgaaa	attacgtgtg	actttttagt	ttgcattatg	cgttttagc	7800
agttggctct	ggatatcact	ttctctcgtt	tgaggttttt	taacctagt	aacttttaag	7860
acaggtttcc	ttaacattca	taagtgccca	gaatacagct	gtgtagtaca	gcatataaag	7920
atctcagctc	tgaggttttt	cctattgact	tggaaaattg	ttttgtgcct	gtcgttgcc	7980
acatggccaa	tcaagtaagc	ttattaatag	taatcaatta	cggggtcatt	agttcatagc	8040
ccatataatg	agttccgctg	tacataactt	acggtaaatg	gcccgcctgg	ctgaccgccc	8100
aacgaccccc	gcccattgac	gtcaataatg	acgtatgttc	ccatagtaac	gccaataggg	8160
actttccatt	gacgtcaatg	ggtggagtat	ttacggtaaa	ctgcccactt	ggcagtacat	8220
caagtgtatc	atatgccaa	tacgccccct	attgacgtca	atgacggtaa	atggcccggc	8280
tggcattatg	cccagtacat	gacettatgg	gactttccta	cttggcagta	catctacgta	8340
ttagtcacgc	ctattaccat	ggtgatgcgg	ttttggcagt	acatcaatgg	gctgtgatag	8400
cggtttgact	cacggggatt	tccaagtctc	cacccattg	acgtcaatgg	gagtttgttt	8460
tggcaccaaa	atcaacggga	ctttccaaaa	tgtcgttaaca	actccgcccc	attgacgcaa	8520
atgggcggta	ggcgtgtacg	gtgggaggtc	tatataagca	gagctggttt	agtgaaccgt	8580
cagatcggat	ccgcttgaga	aaggaagtga	gctgtaaagg	ctgagctctc	tctctgacgt	8640
atgtagcctc	tggttagctt	cgctactcac	tgttcttgac	tcagcatggc	aatctgatga	8700
aatcccagct	gtaagtctgc	agaaattgat	gatctattaa	acaataaaga	tgtccactaa	8760
aatggaagtt	tttctgtca	tactttgtta	agaaggggtg	gaacagagta	cctacatttt	8820
gaatggaagg	attggagcta	cgggggtggg	ggtgggtgg	gattagataa	atgcctgctc	8880
tttactgaag	gctctttact	attgctttat	gataatgttt	catagttgga	tatcataatt	8940
taacaagca	aaaccaaatt	aaggccagc	tcaattctcc	agatccacta	gtaattctgt	9000
ggaatgtgtg	tcagttaggg	tgtgaaaagt	ccccaggctc	cccagcaggc	agaagtatgc	9060
aaagcatgca	tctcaattag	tcagcaacca	ggtgtggaaa	gtccccaggc	tccccagcag	9120
gcagaagtat	gcaaagcatg	catctcaatt	agtcagcaac	catagtcccg	cccctaactc	9180
cgcccattccc	gcccctaact	ccgcccagtt	ccgcccattc	tccgcccact	ggctgactaa	9240
ttttttttat	ttatgcagag	gcccaggccg	cctctgcctc	tgagctattc	cagaagtatg	9300
gaggaggctt	ttttggaggc	ctaggctttt	gcaaaaagct	cccgggagct	tgtatatcca	9360
tttccggtatc	tgatcaagag	acaggatgag	gatcgtttcg	catgattgaa	caagatggat	9420
tgcacgcagg	ttctccggcc	gcttgggtgg	agaggctatt	cggctatgac	tgggcacaac	9480
agacaatcgg	ctgctctgat	gcccgcgtgt	tccggctgtc	agcgcagggg	cgcccggttc	9540
ttttgtcaa	gaccgacctg	tccggtgccc	tgaatgaact	gcaggacgag	gcagcgcggc	9600
tatcstggct	ggccacgacg	ggcgttctt	gcgcagctgt	gctcgacgtt	gtcactgaag	9660
cgggaaggga	ctggctgcta	ttgggcgaag	tgcggggca	ggatctctctg	tcatctcacc	9720
ttgctcctgc	cgagaaagta	tccatcatgg	ctgatgcaat	gcggcggtcg	catacgcttg	9780
atccggctac	ctgcccattc	gaccaccaag	cgaaacatcg	catcgagcga	gcacgtactc	9840
ggatggaagc	cgttcttgtc	gatcaggatg	atctggacga	agagcatcag	gggctcgcgc	9900
cagccgaact	gttcgccagg	ctcaaggcgc	gcatgccca	cggcgaggat	ctcgtcgtga	9960
cccatggcga	tgcttctgtg	ccgaatatca	tggtggaaaa	tggccgcttt	tctggattca	10020
tcgactgtgg	ccggtgggt	gtggcgacc	gctatcagga	catagcgttg	gctaccgctg	10080
atattgctga	agagcttggc	ggcgaatggg	ctgacgctt	cctcgtgctt	tacggtatcg	10140
ccgctcccga	ttcgcagcgc	atcgccttct	atcgccttct	tgacgagttc	ttctgagcgg	10200
gactctgggg	ttcgaaatga	ccgaccaagc	gacgccaac	ctgccatcac	gagatttcca	10260
ttccaccgcc	gccttctatg	aaaggttggg	cttcggaatc	gttttccggg	acgcccgtctg	10320
gatgatcctc	cagcgcgggg	atctcatgct	ggagttcttc	gcccaccca	acttgtttat	10380
tgcagcttat	aatggttaca	aataaagcaa	tagcatcaca	aatttcacaa	ataaagcatt	10440
tttttactg	cattctagtt	gtggtttgtc	caaactcatc	aatgtatctt	atcatgtctg	10500
tataccgtcg	agactagttc	tagagcggcc	gccaccgcg	tggagctcca	gcttttgttc	10560
cctttagtga	gggttaattt	cgagcttggc	gtaatcatgg	tcatagctgt	ttcctgtgtg	10620

```

aaattgttat cgcctcacia ttccacacia catacgagcc ggaagcataa agtgtaaagc 10680
ctgggggtgcc taatgagtga gctaactcac attaattgcg ttgcgctcac tgcccgcttt 10740
ccagtcggga aacctgtcgt gccagggggg acctaggccg ggcaacaatt ggcggccggc 10800
cgcacttttc ggggaaatgt gcgcggaacc cctatttggt tatttttcta aatacattca 10860
aatatgtatc cgctcatgag acaataaacc tgataaatgc ttcaataata ttgaaaaagg 10920
aagagtatga gtattcaaca tttccgtgtc gcccttattc ccttttttgc ggcattttgc 10980
cttcctgttt ttgctcacc agaaacgctg gtgaaagtaa aagatgctga agatcagttg 11040
gggtgcacgag tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt 11100
cgccccgaag aacgttttcc aatgatgagc acttttaaag ttctgctatg tggcgcggtg 11160
ttatcccgta ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat 11220
gacttggttg agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga 11280
gaattatgca gtgctgccat aacctgagt gataaacatg cggccaactt acttctgaca 11340
acgatcggag gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact 11400
cgccttgatc gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc 11460
acgatgcctg tagcaatggc aacaacgctg cgcaactat taactggcga actacttact 11520
ctagcttccc ggcaacaatt aatagactgg atggaggcgg ataaagtgc aggaccactt 11580
ctgcgctcgg cccttcggc tggctgggtt attgctgata aatctggagc cggtgagcgt 11640
gggtctcgcg gtatcattgc agcactgggg ccagatggta agccctccc tatcgtagtt 11700
atctacacga cggggagtca ggcaactat gatgaacgaa atagacagat cgctgagata 11760
gggtgcctcac tgattaagca ttggtaactg tcagacccta ggccgggcaa caattggcgg 11820
ccggccctgc attaatgaat cggccaacgc gcggggagag gcggtttgcg tattgggccc 11880
tcttcgctt cctcgetcac tgactcgetg cgctcggtcg ttcggtcgc gcgagcggta 11940
tcagctcact caaaggcggg aatacgggta tccacagaat caggggataa cgcaggaaaag 12000
aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg 12060
tttttcata ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg 12120
tggcgaaaacc ggacaggact ataaagatag caggcgtttc cccctggaag ctccctcgtg 12180
cgctctcctg ttccgaccct gccgcttacc ggatacctgt ccgcctttct cccttcggga 12240
agcgtggcgc tttctcatag ctcacgctgt aggtatctca gttcggtgta ggtcgttcgc 12300
tccaagctgg gctgtgtgca cgaaccccc gttcagccc accgctgccc cttatccggt 12360
aactatcgtc ttgagtccaa cccggtaaaga cagacttat cgccactggc agcagccact 12420
ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtgggtg 12480
cctaactacg gctacactag aaggacagta tttggatctt gcgctctgct gaagccagtt 12540
accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggg 12600
ggttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct 12660
ttgatcttt ctacggggtc tgacgctcag tggaaacgaaa actc 12704

```

<210> 6

<211> 11273

<212> DNA

<213> 载体

<220>

<221> misc_特征

<222> 8174

<223> n = A, T, C 或 G

<400> 6

```

acgttgtaaa acgacggcca gtgaattgta atacgactca ctatagggcg aattgggtac 60
cgggcccccc ctcgaggtcg agttggggtg gggaaaagga agaaacgagg gcgtattggc 120
cccaatgggg tctcggtggg gtatcgacag agtgccagcc ctgggaccga accccgctt 180
tatgaacaaa cgaccaaca cccgtgcggt ttattctgtc tttttattgc cgtcatagcg 240
cgggttcctt ccggtattgt ctccctccgt cgacgggtatc aagggtggcga ccggaatggg 300
gagctgcgag aatagccggg cgcgctgtga gccgaagtcg cccccgcctt ggccacttcc 360
ggcgcgccga gtccttaggc cgccaggggg cgccggcgcg cgcccagatt ggggacaaaag 420
gaagccgggc cggccgctt attaccataa aaggcaacaa ctggtcggag gcgtccccgc 480
ggcgcgcggc aggaagccag gccccaaacc cctcccaacc gggcgccagc cccgcctccg 540

```

```

ccccggttcaa acagcgaccg ggtcgcgcgc gcgcacgcag cggccacacc ctcgggcgcc 600
agcgggctcgg gcaggaagtg gcgcaagcgc cggggcccca gaacgcacgc gcgattagcg 660
ccattgagtc ccagcgcgca cgcgcaatta gcgccaattc ccagcgcgca cgcagttagc 720
gccccaaagga ccagcgcgca cgcgcatggc gccccagccc ccaccggggc tgacgggggc 780
tacgccgcgc ccaccgtgcg atccccattg gcaagagccc ggctcagaca aagacccccgc 840
cggttgcccc cgccccgaga gcggcacccc cggagcgcgc ccgccccgagc gcggcctcgc 900
gcctgcgaac tggcgtgggg tgtcccccat ctccggaggc ccaggggctt ctccccgcgc 960
ccccacggcg gtccggttcc gccccatgcy cccccgctg cggcccagac ggcggtctg 1020
cacgggcgaa gggccgcggc gcgatgcccc ggtcggctgg cggggcttac ctggcgggcg 1080
gtgtggacgg gcggcggatc ggcaaaggcg aggtctctgt ctcgcgggcg gacgcggtct 1140
cggcggtggt ggcgcgtcgc gccgctgggt tttatagggc gccgcccggc ccgctcgagc 1200
cataaaaggc aactttcggg acggcgcacg ctgattggcc ccgcccgcct cactcaccgg 1260
cttcgccgca cagtgcagca tttttttacc ccctctcccc tccttttgcg aaaaaaaaaa 1320
agagcgagag cgagattgag gaagaggagg agggagagtt ttggcgttgg ccgcttggg 1380
gtgctggggc cgggggctgg gggcgcgcgc cgtggcccc gcgccccacg ctgggagtg 1440
cccggttcgg ccccgcattg ccaggcctgc ccccggcctg cccgtctctc gggcccccca 1500
cccaccgchg gacatcctag gtgtggacat ctcttgggca ctgagcgccc aggtggggtg 1560
ggccagggtc tgacgggtg ccagggccct gggtctgta cgctcctgca gaaggagctc 1620
ttggagggca tggagtggcc aggcagtcac tcccccttgc cgacttcaga gcaactgcc 1680
tgaaagcagg gcctgaggac ctctggctgt ggggtcagc tagctaaatg tgctgggtgg 1740
gtcactagg agagacctg gcttgagagg tagagtgtg ttgtggggga gtcaggtggc 1800
ttcggccat tagagtgcga ggaccacact ccccaggaca gggcaggggc cagcgggcca 1860
gtggctggag gtggcccgtg atgaaggcta caaacctacc cagccgcagc cctgggaaag 1920
aagtgggctc tacagggcag ggcacctttt accctggagc tgcttcttt tgagggtaac 1980
agtcacgccc agccaagacc aggcctgggg cgtagtgagg tgacctaggc actgcggggc 2040
gggggggctg ggtctacaca gcctgggtct gggcccaccg tccgttgat gtctgctatg 2100
cgcagccaca gctgaactgc cctcccagac catctggagg ccgctggggg actctgggga 2160
ccaagactcc atgtgccaca gaggattggg ggcggggcgg tgctaggaac tcaaagccag 2220
cctgggaaga ccctgtcctt gtcaccctt cttgccttg gtctgtccac tgagtagcac 2280
acaagaccgg gtgggcaggg tccgttctgc tccgggaatc acagactgtg tgtaccagc 2340
tggtgggcat gcagcatca gtggcgtggg accacagagg gggcccgcgg tacctaaaac 2400
agcttccat ggcttaaaat aggggaccaa tgtcttttcc aatctaagtc ccatataa 2460
taaagtccat gttccathtt taaaggacaa tcccttccgt ttaaaccag gcacgattac 2520
ccaaacaact cacaacggta aagcactgtg aatcttctct gttctgcaat ccaacttg 2580
tttctgctca gaaaccctcc ctctttccaa tcggtaatta aatacaaaa ggaaaaaac 2640
taagtctct caaccccggt tcgtgacact ttgaaaaaag aatcacctct tgcaaacacc 2700
cgtcccgcac ccccgccgct gaagcccggc gtccagaggc ctaagcgcg gtgcccgc 2760
ccaccgggga gcgcgggcct cgtggctcagc gcaccccgcg ggagaaacaa aggcgcggc 2820
acgggggctc aagggcactg cgccacaccg cacgcgccta cccccgcg cgccagttaa 2880
ctggcggtcg ccgcaacctc gggacagccg gccgcgcgcc gccaggctcg cggacgcggg 2940
accacgcgcc gccctccggg aggcccaagt ctgacccag ccccgctgg cgctggggga 3000
ggggcgccct ccgccggaac gcgggtgggg gaggggaggg ggaaatgccc tttgtctcga 3060
aatggggcaa ccgtcgccac agctccctac cccctcgagg gcagagcagt cccccacta 3120
actaccgggc tggccgcgcy ccaggccagc cgcgaggcca ccgcccgacc ctccactcct 3180
tcccgcagct cccggcgcgg ggtccggcga gaaggggagg ggaggggagc ggagaaccgg 3240
gccccgggga cgcgtgtggc atctgaagca ccaccagcga gcgagagcta gagagaagga 3300
aagccaccga cttcaccgcc tccgagctgc tccgggtcgc gggctctgag cgtctccggc 3360
cctccgcgcc tacagctcaa gccacatccg aagggggagg gagccgggag ctgcccggc 3420
ggccgcggcg gggaggggtg gcaccgcccc cgcggggcgg ccacgaaggg cggggcagcg 3480
ggcgcgcgcy cggcgggggg aggggcccgc gccgcccgc ctgggaattg gggccctagg 3540
gggagggcgg aggcgcccga gaccgcccga cttaccgttc gcggcgtggc gcccggtgg 3600
ccccaagggg aggggaaggg gaggcggggc gaggacagtg accggagtct cctcagcgg 3660
ggcttttctg cttggcagcc tcagcggtg gcgcaaaaac cggactccgc ccacttctc 3720
gcccgcgggt gcgaggtgt ggaatcctcc agacgctggg ggagggggag ttgggagctt 3780
aaaaactagt acccctttg gaccacttcc agcagcgaac tctcctgtac accaggggtc 3840
agttccacag acgcgggcca ggggtgggtc attgcccgt gaacaataat ttgactagaa 3900
gttgattcgg gtgtttccgg aaggggccga gtcaatccgc cgagttgggg cacggaaaac 3960

```

aaaaagggaa	ggctactaag	atTTTTctgg	cggggggtat	cattggegta	actgcagggg	4020
ccacctccc	ggttgagggg	gctggatctc	caggctgcgg	attaagcccc	tcccgtcggc	4080
gttaatttca	aactgcgcga	cgTTTTctcac	ctgccttcgc	caaggcaggg	gccgggaccc	4140
tattccaaga	ggtagtaact	agcaggactc	tagccttccg	caattcattg	agcgattta	4200
cggaagtaac	gtcgggtact	gtctctggcc	gcaaggggtg	gaggagtacg	catttggcgt	4260
aaggtggggc	gtagagcctt	cccgccattg	gcggcggata	gggcgtttac	gcgacggcct	4320
gacgtagcgg	aagacgcggt	agtggggggg	aaggttctag	aaaagcggcg	gcagcggctc	4380
tagcggcagt	agcagcagcg	ccgggtcccc	tgcggaggtg	ctcctcgcag	agttgtttct	4440
cgagcagcgg	cagttctcac	tacagcgcca	ggacgagtc	ggttcgtggt	cgtccgcgga	4500
gatctctctc	atctcgcctc	gctgcgggaa	atcgggctga	agcgactgag	tccgcgatgg	4560
aggtaacggg	tttgaaatca	atgagttatt	gaaaagggca	tggcgaggcc	gttggcgctc	4620
cagtgggaag	cggccagccg	cctccgtggg	agagaggcag	gaaatcggac	caattcagta	4680
gcagtggggc	ttaaggttta	tgaacggggg	cctgagcggg	ggcctgagcg	tacaaacagc	4740
ttccccacc	tcagcctccc	ggcgccattt	cccttactg	ggggtggggg	atggggagct	4800
ttcacatggc	ggacgctgcc	ccgctggggg	gaaagtgggg	cgcgaggcgg	ggaattctta	4860
ttccctttct	aaagcacgct	gcttcggggg	ccacggcgtc	tcctcggcga	gcgtttcggc	4920
gggcagcagg	tcctcgtgag	cgaggctgcg	gagcttcccc	tccccctctc	tcccgggaac	4980
cgatttggcg	gccgccattt	tcatggctcg	ccttctctc	agcgttttcc	ttataactct	5040
tttattttct	tagtgtgctt	tctctatcaa	gaagtagaag	tggttaacta	ttttttttt	5100
cttctcgggc	tgttttcata	tcgtttcgag	gtggatttgg	agtgttttgt	gagcttggat	5160
ctttagagtc	ctgcgcacct	cattaaaggc	gctcagcctt	cccctcgatg	aaatggcgcc	5220
attgcgttcg	gaagccacac	cgaagagcgg	ggaggggggg	tgctccgggt	ttgcgggccc	5280
ggtttcagag	aagatatcac	caccagggc	gtcgggccgg	gttcaatgcg	agccgtagga	5340
caaagaaacc	atTTtatggt	tttctgtct	tttttttct	ttgagtaacg	gttttatctg	5400
ggtctgcagt	cagtaaaacg	acagatgaac	cgcggcaaaa	taaacataaa	ttggaagcca	5460
tcggccacga	ggggcagggg	cgaaggtggt	tttctggcgg	ggggagggat	attcgcgcca	5520
gaatccttta	ctgttcttaa	ggattccggt	taagtgttag	agctgactca	ttttaagtaa	5580
tgttgttact	gagaagttta	acccttacgg	gacagatcca	tggaccttta	tagatgatta	5640
cgaggaaagt	gaaataacga	ttttgtcctt	agttatactt	cgattaaaac	atggcttcag	5700
aggctccttc	ctgtaatgcg	tatggattga	tgtgcaaaac	tgttttgggc	ctgggcccgt	5760
ctgtatttga	actttgttac	ttttctcatt	ttgtttgcaa	tcttggttga	acattacatt	5820
gataagcata	aggctcctca	cgaaaggggt	ctacctgggt	atTTTcttt	gaccctaagc	5880
acgtttataa	aataacattg	tttaaaatcg	atagtgagca	tcgggtaagt	ttggataaat	5940
tgtagggtaa	gtaatgagtt	tttgcttttt	gttagtgatt	tgtaaaactt	gttataaatg	6000
tacattatcc	gtaatttcag	tttagagata	acctatgtgc	tgacgacaat	taagaataaa	6060
aactagctga	aaaaatgaaa	ataactatcg	tgacaagtaa	ccatttcaaa	agactgcttt	6120
gtgtctcata	ggagctagtt	tgatcatttc	agttaatTTT	ttctttaatt	tttacgagtc	6180
atgaaaacta	caggaaaaaa	aatctgaact	gggttttacc	actacttttt	aggagtggg	6240
agcatgcgaa	tggagggaga	gctccgtaga	actgggtaga	ggcagcaat	taagtctgct	6300
tgtaggaac	aaaaaataat	tgattgaaaa	ttacgtgtga	ctttttagtt	tgattatgct	6360
gtttgtagca	gttggtcctg	gatatcactt	tctctcgttt	gaggtttttt	aacctagtta	6420
acttttaaga	caggtttctt	taacattcat	aagtgcccg	aatacagctg	tgtagtacag	6480
catataaaga	tttcagctct	gaggtttttc	ctattgactt	ggaaaattgt	tttgtgcctg	6540
tcgcttgcca	catggccaat	caagtaagct	tattaatagt	aatcaattac	ggggtcatta	6600
gttcatagcc	catatatgga	gttccgcggt	acataactta	cggtaaatgg	cccgcctggc	6660
tgaccgcca	acgacccccg	cccattgacg	tcaataatga	cgtaggttcc	catagtaacg	6720
ccaatagggg	ctttccattg	acgtcaatgg	gtggagtatt	tacggtaaac	tgcccacttg	6780
gcagtacatc	aagtgtatca	tatgccaaat	acgcccccta	ttgacgtcaa	tgacggtaaa	6840
tggccccgct	ggcattatgc	ccagtacatg	accttatggg	actttcctac	ttggcagtac	6900
atctacgtat	tagtcatcgc	tattaccatg	gtgatgcggg	tttggcagta	catcaatggg	6960
cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	acccattgga	cgtcaatggg	7020
agtttgTTTT	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtaacaa	ctccgcccc	7080
ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	agctggTTTT	7140
gtgaaccgtc	agatcgatc	gcctgagaaa	aggaagtgag	ctgtaaaggc	tgagctctct	7200
ctctgacgta	tgtagcctct	ggttagcttc	gtcactcact	gttcttgact	cagcatggca	7260
atctgatgaa	atcccagctg	taagtctgca	gaaattgatg	atctattaaa	caataaagat	7320
gtccactaaa	atggaagttt	ttcctgtcat	actttgttaa	gaagggtag	aacagagtac	7380

ctacatTTTTg	aatggaagga	ttggagctac	gggggtgggg	gtgggggtggg	attagataaa	7440
tgctgtctct	ttactgaagg	ctctttacta	ttgctttatg	ataatgtttc	atagttggat	7500
atcataatTT	aaacaagcaa	aaccaaatta	agggccagct	cattcctcca	gatccactag	7560
taattctgtg	gaatgtgtgt	cagttagggg	gtggaaagtc	cccaggctcc	ccagcaggca	7620
gaagtatgca	aagcatgcat	ctcaattagt	cagcaaccag	gtgtggaaag	tccccaggct	7680
ccccagcagg	cagaagtatg	caaagcatgc	atctcaatta	gtcagcaacc	atagtccccg	7740
ccctaactcc	gccccatccc	cccctaactc	cgcccagttc	cgcccattct	ccgccccatg	7800
gctgactaat	TTTTTTtatt	tatgcagagg	ccgaggccgc	ctctgcctct	gagctattcc	7860
agaagtatgt	aggaggcttt	tttggaggcc	taggcttttg	caaaaagctc	ccgggagctt	7920
gtatatccat	tttcggatct	gatcaagaga	caggatgagg	atcgtttcgc	atgattgaac	7980
aagatggatt	gcacgcaggt	tctccggccg	cttgggtgga	gaggctattc	ggctatgact	8040
gggcacaaca	gacaatcggc	tgctctgatg	ccgccgtgtt	ccggctgtca	gcgcaggggc	8100
gccccgttct	ttttgtcaag	accgacctgt	ccggtgccct	gaatgaactg	caggacgagg	8160
cagcgcggct	atcstggctg	gccacgacgg	gcgttccttg	cgcagctgtg	ctcgcagctt	8220
tactgaagc	gggaaggac	tggtctgat	tgggcgaagt	gccggggcag	gatctcctgt	8280
catctcacct	tgctcctgcc	gagaaagtat	ccatcatggc	tgatgcaatg	cggcggctgc	8340
atacgcttga	tccggctacc	tgcccattcg	accaccaagc	gaaacatcgc	atcgagcgag	8400
cacgtactcg	gatggaagcc	ggtcctgtcg	atcaggatga	tctggacgaa	gagcatcagg	8460
ggctcgcgcc	agccgaactg	ttcgccaggc	tcaaggccgc	catgcccgac	ggcgaggatc	8520
tcgtcgtgac	ccatggcgat	gcctgcttgc	cgaatatcat	ggtggaaaaa	ggccgctttt	8580
ctggattcat	cgactgtggc	cggtgggtg	tgccggaccg	ctatcaggac	atagcgttgg	8640
ctaccctgta	tattgctgaa	gagcttggcg	gcgaatgggc	tgaccgcttc	ctcgtgcttt	8700
acggtatcgc	cgctcccgat	tcgcagcgca	tcgccttcta	tcgccttctt	gacgagtctt	8760
tctgagcggg	actctggggg	tcgaaatgac	cgaccaagcg	acgcccacc	tgccatcacg	8820
agatttcgat	tccaccggcg	ccttctatga	aaggttgggc	ttcggaatcg	ttttccggga	8880
cgccggctgg	atgatcctcc	agcgcgggga	tctcatgctg	gagttcttcg	cccaccccaa	8940
cttgtttatt	gcagcttata	atggttacia	ataaagcaat	agcatcacia	atctcaciaa	9000
taaagcattt	ttttcactgc	attctagtgt	tggtttgtcc	aaactcatca	atgtatctta	9060
tcatgtctgt	ataccgtcga	gactagtctt	agagcggccg	ccaccgcggt	ggagctccag	9120
ctttgtttcc	ctttagttag	ggttaatttc	gagcttggcg	taatcatggt	catagctgtt	9180
tcctgtgtga	aattgttatc	cgctcacaat	tccacacaac	atacgagccg	gaagcataaa	9240
gtgtaaagcc	tgggggtgcct	aatgagtgag	ctaactcaca	ttaattgctg	tgcgctcact	9300
gcccgccttc	cagtcgggaa	acctgtcgtg	ccagggggta	cctaggcccg	gcaacaattg	9360
gcggccggcc	gcacttttcg	gggaaatgtg	cgcggaacc	ctatttgttt	atttttctaa	9420
atacattcaa	atatgtatcc	gctcatgaga	caataaccct	gataaatgct	tcaataatat	9480
tgaaaaagga	agagtatgag	tattcaacat	ttccgtgtcg	cccttattcc	cttttttgcg	9540
gcattttgcc	ttcctgtttt	tgctcaccca	gaaacgctgg	tgaaagtaaa	agatgctgaa	9600
gatcagttgg	gtgcacgagt	gggttacatc	gaactggatc	tcaacagcgg	taagatcctt	9660
gagagttttc	gccccgaaga	acgttttcca	atgatgagca	cttttaaagt	tctgctatgt	9720
ggcgcggtat	tatcccgtat	tgacgccggg	caagagcaac	tcggtcgcgg	catacactat	9780
tctcagaatg	acttggttga	gtactacca	gtcacagaaa	agcatcttac	ggatggcatg	9840
acagtaagag	aattatgcag	tgctgccata	accatgagtg	ataaactgc	ggccaactta	9900
cttctgacaa	cgatcggagg	accgaaggag	ctaaccgctt	ttttgcacia	catgggggat	9960
catgtaactc	gccttgatcg	ttgggaaccg	gagctgaatg	aagccatacc	aaacgacgag	10020
cgtgacacca	cgatgcctgt	agcaatggca	acaacgttgc	gcaaactatt	aactggcgaa	10080
ctacttactc	tagcttcccg	gcaacaatta	atagactgga	tggaggcgga	taaagttgca	10140
ggaccacttc	tgcgctcggc	ccttccggct	ggctggttta	ttgctgataa	atctggagcc	10200
ggtgagcgtg	ggtctcgcgg	tatcattgca	gcactggggc	cagatggtaa	gccctcccgt	10260
atcgtagtta	tctacacgac	ggggagtcag	gcaactatgg	atgaacgaaa	tagacagatc	10320
gctgagatag	gtgcctcact	gattaagcat	tggttaactgt	cagaccctag	gcccggcaac	10380
aattggcggc	cggccctgca	ttaatgaatc	ggccaacgcg	cggggagagg	cggtttgctg	10440
attgggcgct	cttccgcttc	ctcgtcact	gactcgtctg	gctcggctgt	tcggctcggg	10500
cgagcgggat	cagctcactc	aaaggcggta	atacggttat	ccacagaatc	aggggataac	10560
gcaggaaaaga	acatgtgagc	aaaaggccag	caaaaggcca	ggaaccgtaa	aaaggccgcg	10620
ttgctggcgt	ttttccatag	gctccgcccc	cctgacgagc	atcacaaaaa	tcgacgctca	10680
agtcagaggt	ggcgaaacc	gacaggacta	taaagatacc	aggcgtttcc	ccctggaagc	10740
tcctcgtgct	gctctcctgt	tccgaccctg	ccgcttaccg	gatacctgtc	cgcttttctc	10800

```

ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag ttcgggtgtag 10860
gtcgttcgct ccaagctggg ctgtgtgcac gaaccccccg ttcagcccga ccgctgcgcc 10920
ttatccggtg actatcgtct tgagtccaac ccggtaaagac acgacttatc gccactggca 10980
gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagtctctg 11040
aagtgggtggc ctaactacgg ctacactaga aggacagtat ttggtatctg cgctctgctg 11100
aagccagtta ccttcggaaa aagagttggt agctcttgat ccggcaaaaa aaccaccgct 11160
ggtagcggtg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa 11220
gaagatcctt tgatcttttc tacggggtct gacgctcagt ggaacgaaaa ctc 11273

```

```

<210> 7
<211> 12591
<212> DNA
<213> 载体

```

```

<400> 7
acgttgtaaa acgacggcca gtgaattgta atacgactca ctatagggcg aattgggtac 60
cgggcccccc ctcgaggctg agttgggggtg gggaaaagga agaaacgcgg gcgtattggc 120
cccaatgggg tctcgggtggg gtatcgacag agtgccagcc ctgggaccga accccgcggt 180
tatgaacaaa cgaccaaca cccgtgcggt ttattctgtc tttttattgc cgtcatagcg 240
cgggttcctt ccggtattgt ctcttccgt cgactctaga cccgggctgc agcgaggagc 300
tctgcgttct acggtggtca gaccgaagac tgcgacggta ccgacgctgg tcgcgctct 360
tatacccacg tagaacgcag ctacgccaat agaatgcggtg ccaatatgga atttccaggg 420
gaaaaccggg gcggtgttac gttttggctg ccctttcact tcccattgac gtgtattggc 480
tcgagaacgg tactttccca ttaatcagct atgggaaagt accgtttaa gggtcacgtt 540
cattagtttc aatagcccat tgacgtcaat ggtgggaaag tacatggcgt ttttaattaa 600
ttggctggaa aaaccaatg actcaccctc attgacctta tgtacgtgcc aataatggga 660
aaaaccatt gactcaccct ctattgacct tttgtactgg gcaaaaacca atggaaagtc 720
cctattgact cagtgtactt ggctccaatg ggactttcct gttgattcac ccctattgac 780
cttatgtact gggcaaaaacc cattggaaag tccctaataga ctcagtatac gtgccagtaa 840
tgggaaaaac ccatggctt acctccatt gaccttatgt actgggcaaa acccattgga 900
aagtccctat tgactcaatg tacttggtc caatgggact ttctgttga ctcaccctc 960
attgacctta tgtactgggc aaaaccaat ggaaagtccc tattgagtca gtgtattgg 1020
ctccaatggg tttttcccat tgactcatcc cctattgacc ttatgtactg ggcaaaacc 1080
aatggaaagt ccctattgac gcagtgtact tggctccaat gggactttcc tgttgattca 1140
ccccctattg acctatgta ctgggcagaa tacaatggaa agtccctatt gactcaccca 1200
cattgacctt atatgcttgc caacaatgga aaaaccatt ggaaagtccc tattgagtca 1260
gtgtacttgg cagcaatggg tttttcccat tggctcacct cccattgacc caatgtactt 1320
gggcaaaacc cagtggaaag tcccatttga ctcagtgtgc ttgccagtaa tgggaaaaac 1380
ccattggctt acctccatt gaccaatgt acttgggtaa aggccattga gtcaccacc 1440
ctatgctggg aaatggtgaa cgccccctat gtggaaagtc cctatgggcc gccattagag 1500
tgcgatgacc tgctgattca tatgccatat gagtgtatta gggggcttcc cgcttgggaa 1560
attgggtaaa aagtccccgt attactcaca tagggggcgt ttggctttgc aaattagggg 1620
atctcagtgc atttggcatt aaaaactatt ggttctagtc ataaaacggg cggagttggg 1680
cgagctcgaa ttcaaacgac tcgacgggat caaggtggcg accggaatgg tgagctgcga 1740
gaatagccgg gcgcgctgtg agccgaagtc gccccgcgcc tggccacttc cggcgcgccg 1800
agtccctagg ccgcccgggg gcgcccggcg gcgcccagat tggggacaaa ggaagccggg 1860
cggcccgctg tattaccata aaaggcaaac actggtcgga ggcgtccccg cggcgcgccg 1920
caggaagcca ggccccaac ccctccaac cgggcgccag cccgcctcc gcccggttca 1980
aacagcgacc gggtcgcgcg cgcgcacgca gcggccacac cctcgggcgc cagcggctcg 2040
ggcaggaagt ggcgcaagcg cccgggcccc agaacgcacg cgcgattagc gccattgagt 2100
cccagcgcgc acgcgcaatt agcgcgaatt cccagcgcgc acgcagttag cgcctaaagg 2160
accagcgcgc acgcgcatgg cgcgccagcc cccaccgggc ctgacggggg ctacgcccgc 2220
cccaccgtgc gatccccatt ggcaagagcc cggctcagac aaagaccccg ccggttgccc 2280
ccgccccgag agcggcacc cggagcgcgc cccgcccag cgcggcctcg gcctgcgaa 2340
ctggcggtgg gtgtcccca tctccggagg cccaggggct tctccgcgc ccccacggc 2400
ggtccggttc cgccccatgc gcccccgct gcggcccaga cggcggctct gcacgggcga 2460
agggccgcgg ccgcatgccc cggtcggctg gccgggctta cctggcggcg ggtgtggacg 2520

```

ggcggcggat	cggcaaaggc	gaggctctgt	gctcgcgggc	ggacgcggtc	tcggcgggtg	2580
tggcgcgtcg	cgccgctggg	ttttataggg	cgccgcccgc	gccgctcgag	ccataaaagg	2640
caactttcgg	aacggcgcac	gctgattggc	cccgcgccgc	tcactcaccg	gcttcgccgc	2700
acagtgcagc	atTTTTTTac	ccccctctcc	ctccttttgc	gaaaaaaaaa	aagagcgaga	2760
gcgagattga	ggaagaggag	gagggagagt	tttggcgttg	gccgccttgg	ggtgctgggc	2820
ccgggggctg	ggggcgcgcg	ccgtggcccc	cgcgccccac	gctgggcagt	gccccggttcg	2880
gccccgcattg	gccaggcctg	ccccggcctt	gcccgtctct	cgggcccccc	accaccgcgc	2940
ggacatccta	ggtgtggaca	tctcttgggc	actgagcgcc	caggtggggg	gggcagggtg	3000
ctgcacgggt	gccaggggcc	tgggttctgt	acgctcctgc	agaaggagct	cttggagggt	3060
atggagtggc	caggcagtca	ctcccccttg	ccgacttcag	agcaactgcc	ctgaaagcag	3120
ggcctgagga	cctctggctg	tggggctcag	ctagctaaat	gtgctgggtg	ggtcactagg	3180
gagagacctg	ggcttgagag	gtagagtgtg	gtgttggggg	agttaggtgg	cttgcggcca	3240
ttagagtctc	aggaccaçac	tccccaggac	agggcagggg	ccagcgggtc	agtggctgga	3300
ggtggcccg	gatgaaggct	acaaacctac	ccagccgcag	ccctgggaag	gaagtgggct	3360
ctacagggca	gggcaccttt	taccctggag	ctgcctgctt	ttgagggtaa	cagtcacgcc	3420
cagccaagac	caggcctggg	gcgttagtgg	gtgacctagg	cactgcgggg	cgggggggct	3480
gggtctacac	agcctgggtc	tgggcccacc	gtccgttgta	tgtctgctat	gcgagccac	3540
agctgaactg	ccctcccaga	ccatctggag	gccgctgggg	gactctgggg	accaagactc	3600
catgtgccac	agaggattgg	gggcggggcg	gtgctaggaa	ctcaaagcca	gcctgggaag	3660
accctgtcct	tgtcaccett	tettgccttg	ggtctgtcca	ctgagtagca	cacaagaccg	3720
ggtgggcagc	gtccgttctg	ctccgggaat	cacagactgt	gtgtaccag	gtggtgggca	3780
tgacgcgatg	agtggcgtag	gaccacagag	ggggcccgcg	gtacctaaaa	cagcttcaca	3840
tggcttaaaa	taggggacca	atgtcttttc	caatctaagt	cccatttata	ataaagtcca	3900
tgttcattt	ttaaaggaca	atcctttcgg	tttaaaacca	ggcacgatta	cccaaacaac	3960
tcacaacggt	aaagcactgt	gaatcttctc	tgttctgcaa	tcccaacttg	gtttctgctc	4020
agaaaccctc	cctctttcca	atcggtaatt	aaataacaaa	aggaaaaaac	ttaagatgct	4080
tcaaccccg	ttcgtgacac	tttgaaaaaa	gaatcacctc	ttgcaaacac	ccgctcccga	4140
ccccgcgcg	tgaagcccgg	cgccagagg	cctaagcgcg	ggtgcccggc	cccacccggg	4200
agcgcgggcc	tcgtggtcag	cgcatccgcg	gggagaaaca	aaggccgcgg	cacgggggct	4260
caagggcact	gcgccacacc	gcacgcgcct	accccgcgcg	ggccacgtta	actggcggtc	4320
gccgcagcct	cgggacagcc	ggccgcgcgc	cgccaggctc	gccgacgcgg	gaccacgcgc	4380
cgccctccgg	gaggcccaag	tctcgaccca	gccccgcgtg	gcgctggggg	agggggcgcc	4440
tccgccggaa	cgcggtggg	ggaggggagg	gggaaatgcg	ctttgtctcg	aaatggggca	4500
accgtcgcca	cagctcccta	ccccctcgag	ggcagagcag	tcccccaact	aactaccggg	4560
ctggccgcgc	gccaggccag	ccgcgaggcc	accgcccagc	cctccactcc	ttccccgagc	4620
tcccgcgcgc	gggtccggcg	agaaggggag	gggaggggag	cgggagaaccg	ggcccccgcc	4680
acgcgtgtgg	catctgaagc	accaccagcg	agcgagagct	agagagaagg	aaagccaccg	4740
acttaccgc	ctccgagctg	ctccgggtcg	cgggtctgca	gcgtctccgg	ccctccgcgc	4800
ctacagctca	agccaatcc	gaagggggag	ggagccggga	gctgcgcgcg	gggccgcgcg	4860
ggggaggggt	ggcaccgccc	acgccgggcg	gccacgaagg	gcggggcagc	gggcgcgcgc	4920
gcggcggggg	gaggggcccg	cgccgcgccc	gctgggaatt	ggggccctag	ggggagggcg	4980
gaggcgccga	cgaccgcggc	acttaccgtt	cgcggcgtgg	cgcccggtgg	tccccagggg	5040
gagggaaagg	ggaggcgggg	cgaggacagt	gaccggagtc	tcctcagcgg	tggcttttct	5100
gcttggcagc	ctcagcggct	ggcgccaaaa	ccggactccg	cccacttctc	cgcccccgcg	5160
tgcgaggggtg	tggaatcctc	cagacgctgg	gggaggggga	ggtgggagct	taaaaactag	5220
taccctttg	ggaccacttt	cagcagcgaa	ctctcctgta	caccaggggt	cagttccaca	5280
gacgcggggc	aggggtgggt	cattgcggcg	tgaacaataa	tttgactaga	agttgattcg	5340
ggtgtttccg	gaaggggccc	agtcaatccg	ccgagttggg	gcacggaaaa	caaaaaggga	5400
aggctactaa	gatttttctg	gcgggggtta	tcattggcgt	aactgcaggg	accacctccc	5460
gggttgaggg	ggctggatct	ccaggctgcg	gattaagccc	ctcccgtcgg	cgttaatttc	5520
aaactgcgcg	acgtttctca	cctgccttcg	ccaaggcagg	ggccgggacc	ctattccaag	5580
aggtagtaac	tagcaggact	ctagccttcc	gcaattcatt	gagcgcattt	acggaagtaa	5640
cgtcgggtac	tgtctctggc	cgcaagggtg	ggaggagtac	gcatttggcg	taagggtggg	5700
cgtagagcct	tcccgccatt	ggcggcggat	agggcgttta	cgcgacggcc	tgacgtagcg	5760
gaagacgcgt	tagtgggggg	gaaggttcta	gaaaagcggc	ggcagcggct	ctagcggcag	5820
tagcagcagc	gccgggtccc	gtgcggaggt	gctcctcgca	gagttgtttc	tcgagcagcg	5880
gcagttctca	ctacagcgcc	aggacgagtc	cggttcgtgt	tcgtccgcgg	agatctctct	5940

catctcgctc	ggctgcgggga	aatcgggctg	aagcgactga	gtccgcgatg	gaggtaacgg	6000
gtttgaaatc	aatgagttat	tgaaaagggc	atggcgaggc	cgttggcgcc	tcagtgggaag	6060
tcggccagcc	gcctccgtgg	gagagaggca	ggaaatcggg	ccaattcagt	agcagtgggg	6120
cttaagggtt	atgaacgggg	tcttgagcgg	aggcctgagc	gtacaaacag	cttccccacc	6180
ctcagcctcc	cggcgccatt	tcccttcact	gggggtgggg	gatggggagc	tttcacatgg	6240
cggacgctgc	cccgtgsggg	tgaaagtggg	gcgcggaggc	gggaattcct	attccctttc	6300
taaagcacgc	tgcttcgggg	gccacggcgt	ctcctcggcg	agcgtttcgg	cgggcagcag	6360
gtcctcgtga	gcgaggctgc	ggagcttccc	ctccccctct	ctcccgggaa	ccgatttggc	6420
ggccgccatt	ttcatggctc	gccttcctct	cagcgttttc	cttataactc	ttttattttc	6480
ttagtgtgct	ttctctatca	agaagtagaa	gtggttaact	atTTTTTTTT	tcttctcggg	6540
ctgttttcat	atcgtttcca	ggtagatttg	gagtgttttg	tgagcttggg	tcttttagagt	6600
cctgcgacc	tcattaaagg	cgctcagcct	tcccctcgat	gaaatggcgc	cattgcgctc	6660
ggaagccaca	ccgaagagcg	gggagggggg	gtgctccggg	tttgccggcc	cggtttcaga	6720
gaagatatca	ccaccaggg	cgctgggccc	ggttcaatgc	gagccgtagg	acaaagaaac	6780
cattttatgt	ttttcctgtc	ttttttttcc	tttgagtaac	ggttttatct	gggtctgcag	6840
tcagtaaaac	gacagatgaa	ccgcggcaaa	ataaacataa	attggaagcc	atcggccacg	6900
aggggcaggg	acgaaggtgg	ttttctgggc	gggggagggg	tattcgcgctc	agaatccttt	6960
actgttctta	aggattccgt	ttaagttgta	gagctgactc	atTTTAAGTA	atgttgttac	7020
tgagaagttt	aacccttacg	ggacagatcc	atggaccttt	atagatgatt	acgaggaaaag	7080
tgaaataacg	atTTTgtcct	tagttatact	tcgattaaaa	catggcttca	gaggctcctt	7140
cctgtaatgc	gtatggattg	atgtgcaaaa	ctgttttggg	cctgggcccgc	tctgtatttg	7200
aactttgtta	cttttctcat	tttgtttgca	atccttggtg	aacattacat	tgataagcat	7260
aaggtctcaa	gcgaaggggg	tctacctggg	tatttttctt	tgaccctaag	caggtttata	7320
aaataacatt	gtttaaaatc	gatagtggac	atcgggtaag	tttgataaaa	ttgtgaggtg	7380
agtaatgagt	ttttgctttt	tgtagtgat	ttgtaaaact	tgttataaat	gtacattatc	7440
cgtaatttca	gtttagagat	aacctatgtg	ctgacgacaa	ttaagaataa	aaactagctg	7500
aaaaaatgaa	aataactatc	gtgacaagta	accatttcaa	aagactgctt	tgtgtctcat	7560
aggagctagt	ttgatcattt	cagttaattt	tttctttaat	ttttacgagt	catgaaaact	7620
acaggaaaaa	aatctgaac	tgggttttac	cactactttt	taggagttgg	gagcatgcga	7680
atggagggag	agctccgtag	aactgggatg	agagcagcaa	ttaatgctgc	ttgctaggaa	7740
caaaaaataa	ttgattgaaa	attacgtgtg	actttttagt	ttgcattatg	cgtttgtagc	7800
agttggtcct	ggatcact	ttctctcgtt	tgaggttttt	taacctagtt	aacttttaag	7860
acaggtttcc	ttaacattca	taagtgccca	gaatacacgct	gtgtagtaca	gcatataaag	7920
atTtcagctc	tgaggttttt	cctattgact	tggaaaattg	ttttgtgcct	gtcgcctgcc	7980
acatggccaa	tcaagtaagc	ttattaatag	taatcaatta	cggggctcatt	agttcatagc	8040
ccatataatg	agttccgctg	tacataactt	acggtaaattg	gcccgcctgg	ctgaccgcc	8100
aacgaccccc	gccattgac	gtcaataatg	acgtatgttc	ccatagtaac	gccaataggg	8160
actttccatt	gacgtcaatg	ggtggagtag	ttacggtaaa	ctgcccactt	ggcagtagac	8220
caagtgtatc	atatgccaag	tacgccccct	attgacgtca	atgacggtaa	atggccccgc	8280
tggcattatg	cccagtagat	gaccttatgg	gactttccta	cttggcagta	catctacgta	8340
ttagtcatcg	ctattaccat	ggtgatgagg	ttttggcagt	acatcaatgg	gcgtggatag	8400
cggtttgact	cacggggatt	tccaagtctc	caccccattg	acgtcaatgg	gagtttgttt	8460
tggcaccaaa	atcaacggga	ctttccaaaa	tgctgtaaca	actccgcccc	attgacgcaa	8520
atgggcggta	ggcgtgtacg	gtgggaggtc	tatataagca	gagctggttt	agtgaaccgt	8580
cagatcggat	ccgctgaga	aaggaagtga	gctgtaaagg	ctgagctctc	tctctgacgt	8640
atgtagcttc	tggttagctt	cgtcactcac	tgttcttgac	tcagcatggc	aatctgatga	8700
aatcccagct	gtaagtctgc	agaaattgat	gatctattaa	acaataaaga	tgtccactaa	8760
aatggaagtt	tttctgtca	tactttgtta	agaaggggtg	gaacagagta	cctacatttt	8820
gaatggaagg	attggagcta	cgggggtggg	ggtgggggtg	gattagataa	atgcctgctc	8880
tttactgaag	gctctttact	attgctttat	gataatgttt	catagttgga	tatcataatt	8940
taacaagca	aaaccaaat	aagggccagc	tcattcctcc	agatccacta	gttcttagagc	9000
aaattctacc	gggtagggga	ggcgcttttc	ccaagggcag	ctggagcatg	cgcttagca	9060
gcccgtctgg	gcacttgccg	ctacacaagt	ggcctctggc	ctcgcacaca	ttccacatcc	9120
accggtaggc	gccaaaccggc	tccgttcttt	ggtggcccc	tcgcccacc	ttctactcct	9180
cccctagtea	ggaagttccc	ccccgcccc	cagctcgcgt	cgtgcaggac	gtgacaaatg	9240
gaagtagcac	gtctcactag	tctcgtgcag	atggacagca	ccgctgagca	atggaagcgg	9300
gtaggccttt	ggggcagcgg	ccaatagcag	ctttgctcct	tcgctttctg	ggctcagagg	9360

ctgggaagg	gtgggtccg	gggcgggctc	aggggagggc	tcaggggagg	ggcggggcgc	9420
cgaaggtcct	ccggaggccc	ggcattctgc	acgcttcaaa	agcgcacgtc	tgcgcgcgctg	9480
ttctcctctt	cctcatctcc	gggcctttcg	accagcttac	catgaccgag	tacaagccca	9540
cggtgcgcct	cgccacccgc	gacgacgctc	ccagggccgt	acgcaccctc	gcccgcgcgt	9600
tcgcccacta	ccccgccacg	cgccacaccg	tcgatccgga	ccgccacatc	gagcgggtca	9660
ccgagctgca	agaactcttc	ctcacgcgcg	tcgggctcga	catcggcaag	gtgtgggtcg	9720
cggacgacgg	cgccgcggtg	gcgggtctgga	ccacgcggga	gagcgtcgaa	gccccggcgg	9780
tgttcgccga	gatcggcccc	cgcatggccg	agttgagcgg	ttcccggctg	gcccgcgcaga	9840
acagatggaa	ggcctcctgg	cgccgcaccg	gcccgaaggag	cccgcgtggt	tcctggccac	9900
gctcgcgtct	cgcccagcca	ccagggcaag	ggtctgggca	gcgcgcgtcgt	gctccccgga	9960
gtggaggcgg	ccgagcgcgc	cggggtgccc	gccttcctgg	agacctccgc	gccccgaac	10020
ctccccctct	acgagcggct	cggttcacc	gtcaccgcg	acgtcgaggt	ccccgaagga	10080
ccgcgcacct	ggtgcatgac	ccgcaagccc	ggtgcctgac	gcccgcacca	cgaccgcgag	10140
cgcccagacc	aaaggagcgc	acgaccccat	gcataggttg	ggcttcggaa	tcgttttccg	10200
ggacgcgggc	tggatgatcc	tccagcgcgg	ggatctcatg	ctggagttct	tcgcccacc	10260
caacttgttt	attgcagctt	ataatggtta	caaataaagc	aatagcatca	caaatttcac	10320
aaataaagca	tttttttcac	tgcattctag	ttgtggtttg	tccaaactca	tcaatgtatc	10380
ttatcatgtc	tgtataccgt	cgagatctag	agcggccgcc	accgcggtgg	agctccagct	10440
tttgttccct	ttagtgaggg	ttaatttcga	gcttggcgta	atcatggtca	tagctgtttc	10500
ctgtgtgaaa	ttgttatccg	ctcacaatte	cacacaacat	acgagccgga	agcataaagt	10560
gtaaagcctg	gggtgcctaa	tgagtgagct	aactcacatt	aattgcggtg	cgctcactgc	10620
ccgctttcca	gtcgggaaac	ctgtcgtgcc	agggggtacc	taggccgggc	aacaattggc	10680
ggccggccgc	acttttcggg	gaaatgtgcy	cggaaaccct	atgtgtttat	ttttctaaat	10740
acattcaaat	atgtatccgc	tcatgagaca	ataaccctga	taaagtcttc	aataatattg	10800
aaaaaggaag	agtatgagta	ttcaacattt	ccgtgtcgcc	cttattccct	ttttgcygc	10860
attttgctt	cctgtttttg	ctcaccaga	aacgctggty	aaagtaaaag	atgctgaaga	10920
tcagttgggt	gacagagtg	gttacatcga	actggatctc	aacagcggta	agatccttga	10980
gagttttcgc	cccgaagaac	gttttccaat	gatgagcact	tttaaagttc	tgctatgtgg	11040
cgcggtatta	tcccgtattg	acgcccggca	agagcaactc	ggtcgcccga	tacactattc	11100
tcagaatgac	ttggttgagt	actcaccagt	cacagaaaag	catcttacgg	atggcatgac	11160
agtaagagaa	ttatgcagtg	ctgccataac	catgagtgat	aacactgcgg	ccaacttact	11220
tctgacaacg	atcggaggac	cgaaggagct	aaccgctttt	ttgcacaaca	tgggggatca	11280
tgtaactcgc	cttgatcggt	gggaaccgga	gctgaatgaa	gccataccaa	acgacgagcg	11340
tgacaccacg	atgcctgtag	caatggcaac	aacgttgcgc	aaactattaa	ctggcgaaact	11400
acttactcta	gcttcccggc	aacaattaat	agactggatg	gaggcggata	aagttgcagg	11460
accacttctg	cgctcggccc	ttccggctgg	ctggtttatt	gctgataaat	ctggagccgg	11520
tgagcgtggg	tctcgcggta	tcattgcagc	actggggcca	gatggtaagc	cctcccgtat	11580
cgtagttatc	tacacgacgg	ggagtcaggc	aactatggat	gaacgaaata	gacagatcgc	11640
tgagataggt	gctcactga	ttaagcattg	gtaactgtca	gaccctaggc	gggcaacaa	11700
ttggcggcct	gcccctgcat	aatgaatcgg	ccaacgcgcg	gggagaggcg	gtttgcgtat	11760
tgggcgctct	tccgcttcc	cgctcactga	ctcgtgcgc	tcggtcgttc	ggctgcggcg	11820
agcggtatca	gctcactcaa	aggcggtaat	acggttatcc	acagaatcag	gggataacgc	11880
aggaaagaac	atgtgagcaa	aaggccagca	aaaggccagg	aaccgtaaaa	aggccgcggt	11940
gctggcggtt	ttccataggc	tccgcccc	tgacgagcat	cacaaaaatc	gacgctcaag	12000
tcagaggtgg	cgaaccgca	caggactata	aagataccag	gcgtttcccc	ctggaagctc	12060
cctcgtgcgc	tctcctgttc	cgaccctgcc	gcttaccgga	tacctgtccg	cctttctccc	12120
ttcgggaagc	gtggcgcttt	ctcatagctc	acgctgtagg	tatctcagtt	cggtgtaggt	12180
cgttcgctcc	aagctgggct	gtgtgcacga	acccccggt	cagcccagacc	gctgcgcctt	12240
atccggtaac	tatcgtcttg	agtccaacc	ggtaagacac	gacttatcgc	cactggcagc	12300
agccactggg	aacaggatta	gcagagcgag	gtatgtaggc	ggtgctacag	agttcctgaa	12360
gtgggtggct	aactacggct	acactagaag	gacagtattt	ggtatctgcy	ctctgctgaa	12420
gccagttacc	ttcggaaaaa	gagttggtag	ctcttgatcc	ggcaaaaaaa	ccaccgctgg	12480
tagcgggtgg	ttttttgtt	gcaagcagca	gattacgcgc	agaaaaaaag	gatctcaaga	12540
agatcctttg	atcttttcta	cggggtctga	cgctcagtyg	aacgaaaaact	c	12591

<210> 8

<211> 11160

<212> DNA

<213> 载体

<400> 8

```

acgttgtaaa acgacggcca gtgaattgta atacgactca ctatagggcg aattgggtac 60
cgggccccc ctcgaggtcg agttgggggtg gggaaaagga agaaacgcgg gcgtattggc 120
cccaatgggg tctcggttgg gtatcgacag agtgccagcc ctgggaccga acccccgctt 180
tatgaacaaa cgaccaaca cccgtgcgtt ttattctgtc tttttattgc cgtcatagcg 240
cgggttcctt ccggtattgt ctccttccgt cgacggtatc aaggtggcga ccggaatggg 300
gagctgcgag aatagccggg cgcgctgtga gccgaagtcg cccccgccct ggccacttcc 360
ggcgcgccga gtccttaggc cgccaggggg cgccggcgcg cgcccagatt ggggacaaaag 420
gaagccgggg cggccgcgctt attaccataa aaggcaaaaca ctggctcggag gcgtccccgc 480
ggcgcgcggc aggaagccag gcccacaacc cctcccacc gggcgccagc cccgcctccg 540
cccggttcaa acagcgaccg ggtcgcgcgc gcgcacgcag cggccacacc ctccgggcgcc 600
agcggctcgg gcaggaagtg gcgcaagcgc cggggcccca gaacgcacgc gcgattagcg 660
ccattgagtc ccagcgcgca cgcgcaatta gcgccaattc ccagcgcgca cgcagttagc 720
gccc aaagga ccagcgcgca cgcgcatggc gccccagccc ccaccgggccc tgacggggggc 780
tacgccgcgc ccaccgtgcg atccccattg gcaagagccc ggctcagaca aagacccccgc 840
cggttgcccc cgccccgaga gcggcacccc cggagcgcgc ccgcccgagc gcggcctcgc 900
gcctgcgaac tggcgtgggg tgtcccccat ctccggaggc ccaggggctt ctcccgcgcc 960
ccccacgggg gtccgggttcc gccccatgcy ccccccgctg cggcccagac ggcggtctcg 1020
cacgggcgaa gggcgcgcggc cgcattgcccc ggtcggctgg ccgggcttac ctggcggcgg 1080
gtgtggacgg gcggcgatc ggcaaaggcg aggtctgtg ctcgcgggcg gacgcggtct 1140
cggcggtggt ggcgcgtcgc gccgctgggt tttatagggc gccgcccggc ccgctcgagc 1200
cataaaaggc aactttcgga acggcgcacg ctgattggcc ccgcgccgct cactcacccg 1260
cttcgccgca cagtgcagca ttttttacc cctctcccc tccttttgcg aaaaaaaaaa 1320
agagcgagag cgagattgag gaagaggagg agggagagtt ttggcgttgg ccgccttggg 1380
gtgctggggc cgggggctgg gggcgcgcgc cgtggcccc gcgccccacg ctgggcagtg 1440
cccggttcgg ccccgcatgg ccaggcctgc cccggcctg ccgctctctc gggcccccca 1500
cccaccgcgg gacatcctag gtgtggacat ctcttgggca ctgagcgcgc aggtgggggtg 1560
ggccagggtc tgcacgggtg ccagggccct gggttctgta cgctcctgca gaaggagctc 1620
ttggagggca tggagtggcc aggcagtcac tcccccttgc cgacttcaga gcaactgccc 1680
tgaaagcagg gcctgaggac ctctggctgt ggggctcagc tagctaaatg tgctgggtgg 1740
gtcactaggg agagacctgg gcttgagagg tagagtgtgg tgttggggga gtcagggtggc 1800
ttgcggccat tagagtgcga ggaccacact cccagggaca gggcaggggc cagcggtaaca 1860
gtggctggag ttggcccgtg atgaaggcta caaacctacc cagccgcagc cctgggaagg 1920
aagtgggctc tacagggcag ggcacctttt accctggagc tgcctgcttt tgagggtaac 1980
agtacgcccc agccaagacc aggcctgggg cgttagtggg tgacctaggc actgcggggc 2040
gggggggctg ggtctacaca gcctgggtct gggcccaccg tccgttgtat gtctgctatg 2100
cgcagccaca gctgaactgc cctcccagac catctggagg ccgctggggg actctggggg 2160
ccaagactcc atgtgccaca gaggattggg ggcggggcgg tgctaggaac tcaaagccag 2220
cctgggaaga cctgtcctt gtcacccttt ctgcccctgg gtctgtccac tgagtagcac 2280
acaagaccgg gtgggcaggg tccgttctgc tccgggaatc acagactgtg tgtaccagc 2340
tggtgggcat gcagcgatca gtggcgtggg accacagagg gggcccgcgg tacctaaaac 2400
agcttcacat ggcttaaaat aggggaccaa tgtcttttcc aatctaagtc ccatttataa 2460
taaagtccat gttccathtt taaaggacaa tcctttcggg ttaaaaccag gcacgattac 2520
ccaaacaact cacaacggta aagcactgtg aatcttctct gttctgcaat cccaacttgg 2580
tttctgctca gaaacctccc ctctttccaa tcggtaatta aataacaaaa ggaaaaaact 2640
taagatgctt caaccccggt tctgacact ttgaaaaaag aatcacctct tgcaaacacc 2700
cgctcccgac ccccgccgct gaagcccggc gtcagaggc ctaagcgcgg gtgcccggcc 2760
ccaccgggga gcgcgggcct cgtggctcagc gcatccgcgg ggagaaacaa aggcgcggc 2820
acgggggctc aagggcactg cgccacaccg caocgcgcta cccccgcgcg gccacgttaa 2880
ctggcggtcg ccgcagcctc gggacagccg gccgcgcgcc gccaggctcg cggacgcggg 2940
accacgcgcc gccctccggg aggcccaagt ctgcacccag ccccgctgg cgctggggga 3000
gggggcgcct ccgcccgaac gcgggtgggg gaggggaggg ggaaatgcgc tttgtctcga 3060
aatggggcaa ccgtcggcac agctccctac cccctcgagg gcagagcagt cccccacta 3120
actaccgggc tggccgcgcg ccaggccagc cgcgaggcca ccgcccgacc ctccactcct 3180

```

tcccgcagct	cccggcgcgg	ggtccggcga	gaaggggagg	ggagggggagc	ggagaaccgg	3240
gcccccgga	cgcgtgtggc	atctgaagca	ccaccagcga	gcgagagcta	gagagaagga	3300
aagccaccga	cttcaccgcc	tccgagctgc	tccgggtcgc	gggtctgcag	cgctccggc	3360
cctccgcgcc	tacagctcaa	gccacatccg	aagggggagg	gagccgggag	ctgcgcgcgg	3420
ggccgcggg	gggaggggtg	gcaccgccca	cgccgggcgg	ccacgaaggg	cggggcagcg	3480
ggcgcgcgcg	cggcgggggg	aggggcgggc	gccgcgcccg	ctgggaattg	gggccctagg	3540
gggagggcgg	agggcggcag	gaccgcggca	cttaccgttc	gcggcgtggc	gcccgggtgt	3600
ccccaaaggg	agggaaaggg	gaggcggggc	gaggacagtg	accggagtct	ccctcagcgt	3660
ggctttctg	cttggcagcc	tcagcggctg	gcgccaaaac	cggactccgc	ccacttcctc	3720
gcccgcgggt	gcgaggggtg	ggaatcctcc	agacgctggg	ggagggggag	ttgggagctt	3780
aaaaactagt	acccttttgg	gaccactttc	agcagcgaac	tctcctgtac	accaggggtc	3840
agttccacag	acgcggggcca	gggggtgggtc	attgcccgtt	gaacaataat	ttgactagaa	3900
gttgattcgg	gtgtttccgg	aaggggccga	gtcaatccgc	cgagttgggg	cacggaaaac	3960
aaaaagggaa	ggctactaag	atTTTTctgg	cgggggttat	cattggcgta	actgcagggg	4020
ccacctcccg	ggttgagggg	gctggatctc	caggctgcgg	attaagcccc	tcccgtcggc	4080
gtaatttca	aactgcgcga	cgtttctcac	ctgccttcgc	caaggcaggg	gccgggacc	4140
tattccaaga	ggtagtaact	agcaggactc	tagccttccg	caattcattg	agcgcattta	4200
cggaagtaac	gtcgggtact	gtctctggcc	gcaaggggtg	gaggagtacg	catttggcgt	4260
aaggtggggc	gtagagcctt	cccgccattg	gcggcggata	gggcgtttac	gcgacggcct	4320
gacgtagcgg	aagacgcgtt	agtggggggg	aaggttctag	aaaagcggcg	gcagcggctc	4380
tagcggcagc	agcagcagcg	ccgggtcccc	tgccgaggtg	ctcctcgcag	agttgtttct	4440
cgagcagcgg	cagttctcac	tacagcggca	ggacgagtcc	ggttcgtgtt	cgtccgcgga	4500
gatctctctc	atctcgtcgc	gctgcgggaa	atcgggtcga	agcgaactgag	tccgcgatgg	4560
aggtaacggg	tttgaatca	atgagttatt	gaaaagggca	tggcgaggcc	gttggcgcct	4620
cagtggaagt	cggccagccg	cctccgtggg	agagaggcag	gaaatcggac	caattcagta	4680
gcagtggggc	ttaaggttta	tgaacggggg	cttgagcggg	ggcctgagcg	tacaaacagc	4740
ttccccaccc	tcagcctccc	ggcgccattt	cccttcaactg	ggggtggggg	atggggagct	4800
ttcacatggc	ggacgctgcc	ccgctggggg	gaaagtgggg	cgccgagggc	ggaattctta	4860
ttccctttct	aaagcacgct	gcttcggggg	ccacggcgtc	tctcggcga	gcgtttcggc	4920
gggcagcagg	tcctcgtgag	cgaggctgcg	gagcttcccc	tccccctctc	tcccgggaac	4980
cgatttggcg	gcccattt	tcatggctcg	ccttctctc	agcgttttcc	ttataactct	5040
tttattttct	tagtgtgctt	tctctatcaa	gaagtagaag	tggttaacta	ttttttttt	5100
cttctcgggc	tgttttcata	tcgtttcgag	gtggatttgg	agtgttttgt	gagcttggat	5160
ctttagagtc	ctgcgcacct	cattaaaggc	gctcagcctt	cccctcgatg	aaatggcgcc	5220
attgcgttcg	gaagccacac	cgaagagcgg	ggaggggggg	tgctccgggt	ttgcggggcc	5280
ggtttcagag	aagatcac	caccagggc	gtcggggcgg	gttcaatgcg	agccgtagga	5340
caaagaaacc	attttatggt	tttctgtct	tttttttct	ttgagtaacg	gttttatctg	5400
ggtctgcagt	cagtaaaacg	acagatgaac	cgcgcaaaaa	taaacataaa	ttggaagcca	5460
tcggccacga	ggggcagggg	cgaaggtggt	tttctggggc	ggggagggat	attcgcgtca	5520
gaatccttta	ctgttcttaa	ggattccggt	taagttgtag	agctgactca	ttttaagtaa	5580
tgttgttact	gagaagttta	acccttacgg	gacagatcca	tggaccttta	tagatgatta	5640
cgaggaaagt	gaaataacga	ttttgtcctt	agttataact	cgattaaaac	atggcttcag	5700
aggctccttc	ctgtaatgcg	tatggattga	tgtgcaaaaac	tgttttgggc	ctgggccgct	5760
ctgtatttga	actttgttac	ttttctcatt	ttgtttgcaa	tcttgggttga	acattacatt	5820
gataagcata	aggtctcaag	cgaagggggg	ctacctgggt	atttttcttt	gaccctaagc	5880
acgtttataa	aataacattg	tttaaaatcg	atagtggaac	tcgggtaagt	ttggataaat	5940
tgtgaggtaa	gtaatgagtt	tttgcttttt	gttagtgtt	tgtaaaactt	gttataaatg	6000
tacattatcc	gtaatttcag	tttagagata	acctatgtgc	tgacgacaat	taagaataaa	6060
aactagctga	aaaaatgaaa	ataactatcg	tgacaagtaa	ccatttcaaa	agactgcctt	6120
gtgtctcata	ggagctagtt	tgatcatttc	agttaatttt	ttctttaatt	tttacgagtc	6180
atgaaaacta	caggaaaaaa	aatctgaact	gggttttacc	actacttttt	aggagttggg	6240
agcatgcgaa	tggaggggaga	gctccgtaga	actgggatga	gagcagcaat	taatgctgct	6300
tgctaggaac	aaaaaataat	tgattgaaaa	ttacgtgtga	cttttttagtt	tgcatatgct	6360
gtttgtagca	gttggctctg	gatatactt	tctctogttt	gaggtttttt	aacctagtta	6420
acttttaaga	caggtttctt	taacattcat	aagtgccag	aatacagctg	tgtagtacag	6480
catataaaga	tttcagctct	gaggtttttc	ctattgactt	ggaaaattgt	tttgtgcctg	6540
tcgcttgcca	catggccaat	caagtaagct	tattaatagt	aatcaattac	ggggtcatta	6600

gttcatagcc	catatatgga	gttccgcggt	acataactta	cggtaaatgg	cccgcctggc	6660
tgaccgcccc	acgacccccg	cccattgacg	tcaataatga	cgatatgttc	catagtaacg	6720
ccaatagggg	ctttccattg	acgtcaatgg	gtggagtatt	tacggtaaac	tgcccacttg	6780
gcagtacatc	aagtgtatca	tatgccaaat	acgcccccta	ttgacgtcaa	tgacggtaaa	6840
tgccccgctt	ggcattatgc	ccagtacatg	accttatggg	actttcctac	ttggcagtac	6900
atctacgtat	tagtcatcgc	tattaccatg	gtgatgcggg	tttggcagta	catcaatggg	6960
cgtagtatgc	ggtttgactc	acggggattt	ccaagtctcc	accccattga	cgtagatggg	7020
agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	gtcgtaacaa	ctccgcccc	7080
ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	atataagcag	agctggttta	7140
gtgaaccgct	agatcggatc	cgcttgagaa	aggaagtggg	ctgtaaaggc	tgagctctct	7200
ctctgacgta	tgtagcctct	ggtttagctt	gtcactcact	gttcttgact	cagcatggca	7260
atctgatgaa	atcccagctg	taagtctgca	gaaattgatg	atctattaa	caataaagat	7320
gtccactaaa	atggaagttt	ttcctgtcat	actttgttaa	gaaggggtgag	aacagagtac	7380
ctacattttg	aatggaagga	ttggagctac	gggggtgggg	gtgggggtggg	attagataaa	7440
tgccctgctt	ttactgaagg	ctctttacta	ttgctttatg	ataatgtttc	atagttggat	7500
atcataattt	aaacaagcaa	aaccaaatta	agggccagct	cattcctcca	gatccactag	7560
ttctagagca	aattctaccg	ggtagggggg	gcgcttttcc	caaggcagtc	tggagcatgc	7620
gcttttagcag	ccccgctggg	cacttggcgc	tacacaagtg	gcctctggcc	tgcacacat	7680
tccacatcca	ccggtagggc	ccaaccggct	ccgttctttg	gtggccccctt	cgcgccacct	7740
tctactectc	ccctagtcag	gaagttcccc	cccgcctggc	agctcgcgct	gtgcaggacg	7800
tgacaaatgg	aagtagcacg	tctcactagt	ctcgtgcaga	tggacagcac	cgctgagcaa	7860
tggaagcggg	taggcctttg	gggcagcggc	caatagcagc	tttgctcctt	cgctttctgg	7920
gctcagagcg	tgggaagggg	tgggtccggg	ggcgggctca	ggggcgggct	caggttcggg	7980
gcgggcgccc	gaaggtcctc	cggaggcccc	gcattctgca	cgcttcaaaa	gcgcacgtct	8040
gcccgcgtgt	tctcctcttc	ctcatctccg	ggcctttcga	ccagcttacc	atgaccgagt	8100
acaagcccac	ggtgcgcttc	gccaccgcg	acgacgtccc	cagggccgta	cgcacctctg	8160
ccgcccgcgt	cgccgactac	cccgccacgc	gccacaccgt	cgatccggac	cgccacatcg	8220
agcgggtcac	cgagctgcaa	gaactcttcc	tcacgcgctg	cgggctcgac	atcggcaagg	8280
tgtgggtcgc	ggacgacggc	gcccgcgtgg	cggtctggac	cacgccggag	agcgtcgaag	8340
cgggggctgg	gttcgcccag	atcggccccg	gcattggcca	gttgagcggg	tcccggctgg	8400
ccgcccagaa	cagatggaag	gcctcctggc	gccgcaccgg	cccaaggagc	ccgctggttt	8460
cctggccacc	gtcgcgtctc	gcccgaccac	cagggcaagg	gtctgggagc	cgccgtcgtg	8520
ctccccggag	tggaggcggc	cgagcgcgcc	gggggtgccc	ccttctgga	gacctccgcg	8580
ccccgcaacc	tcccccttca	cgagcggctc	ggcttcaccg	tcaccgcca	cgctgaggtg	8640
cccgaaggag	cgccgacctg	gtgcatgacc	cgcaagcccc	gtgcctgacg	cccgcctcag	8700
gaccgcgacg	gcccgaaccg	aaggagcgca	cgacccccat	cataggttgg	gcttcggaat	8760
cgttttccgg	gacgcccggc	ggatgatcct	ccagcggcgg	gatctcatgc	tggagtctct	8820
cgcccacccc	aacttgttta	ttgcagctta	taatggttac	aaataaagca	atagcatcac	8880
aaatttcaca	aataaagcat	tttttctact	gcattctagt	tgtggtttgt	ccaaactcat	8940
caatgtatct	tatcatgtct	gtataccgtc	gagatctaga	gcggccgcca	ccgcggtgga	9000
gctccagctt	ttgttccctt	tagtgagggt	taatttcgag	cttggcgtaa	tcatggtcat	9060
agctgtttcc	tgtgtgaaat	tgttatccgc	tcacaattcc	acacaacata	cgagccggaa	9120
gcataaagtg	taaagcctgg	ggtgccta	gagtgagcta	actcacatta	attgctgtgc	9180
gctcactgcc	cgctttccag	tccggaaacc	tgtcgtgcca	gggggtacct	aggccgggca	9240
acaattggcg	gccggccgca	cttttcgggg	aaatgtgctc	ggaaccctca	tttgtttatt	9300
tttctaata	cattcaata	tgtatccgct	catgagacaa	taaccctgat	aatgcttca	9360
ataatattga	aaaaggaaga	gtatgagtat	tcaacatttc	cggtgcgcc	ttattccctt	9420
ttttgcggca	ttttgccttc	ctgtttttgc	tcaccagaaa	acgctgggtg	aagtaaaaga	9480
tgetgaagat	cagttgggtg	cacgagtggg	ttacatcgaa	ctggatctca	acagcggtaa	9540
gatccttgag	agttttcgcc	ccgaagaacg	ttttccaatg	atgagcactt	ttaaagtctt	9600
gctatgtggc	gcggtattat	cccgtattga	cgccgggcaa	gagcaactcg	gtcgcctcat	9660
acactattct	cagaatgact	tggttgagta	ctcaccagtc	acagaaaagc	atcttacgga	9720
tggcatgaca	gtaagagaat	tatgcagtgc	tgccataacc	atgagtgata	acactgcggc	9780
caacttactt	ctgacaacga	tccggaggacc	gaaggagcta	accgcttttt	tgcacaacat	9840
gggggatcat	gtaactcgcc	ttgatcgttg	ggaaccggag	ctgaatgaag	ccataccaaa	9900
cgacgagcgt	gacaccacga	tgctgtagc	aatggcaaca	acgttgcgca	aactattaac	9960
tggcgaacta	cttactctag	cttcccggca	acaattaata	gactggatgg	aggcggataa	10020

```

agttgcagga ccacttctgc gctcggccct tccggctggc tggtttattg ctgataaatc 10080
tggagccggg gagcgtgggt ctcgcggtat cattgcagca ctggggccag atggtaagcc 10140
ctcccgtatc gtagttatct acacgacggg gagtcaggca actatggatg aacgaaatag 10200
acagatcgct gagatagggt cctcactgat taagcattgg taactgtcag accctaggcc 10260
gggcaacaat tggcggccgg ccctgcatta atgaatcggc caacgcgcgg ggagaggcgg 10320
tttgcgtatt gggcgctctt ccgcttccct gctcactgac tcgctgcgct cggtcgctcg 10380
gctgcgggca gcggtatcag ctcaactcaa ggcggtaata cggttatcca cagaatcagg 10440
ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa 10500
ggcgcgcttg ctggcgcttt tccataggct ccgccccct gacgagcatc acaaaaatcg 10560
acgctcaagt cagagggtggc gaaaccgcag aggactataa agataccagg cgtttcccc 10620
tggaaagctc ctcgctgcgt ctctgttcc gaccctgccg cttaccggat acctgtccgc 10680
ctttctccct tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc 10740
ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa cccccgctc agcccagacc 10800
ctgcgctta tccggtaact atcgtcttga gtccaaccgg gtaagacacg acttatcgcc 10860
actggcagca gccactggta acaggattag cagagcgagg tatgtaggag gtgctacaga 10920
gttcttgaag tggtagccta actacggcta cactagaagg acagtatttg gtatctgcgc 10980
tctgctgaag ccagttacct tcggaaaag agttggtagc tcttgatccg gcaaacaaac 11040
caccgctggt agcgggtggt tttttgtttg caagcagcag attacgcgca gaaaaaaagg 11100
atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc 11160

```

<210> 9

<211> 14262

<212> DNA

<213> 载体

<220>

<221> misc_特征

<222> 13483

<223> n = A,T,C 或 G

<400> 9

```

ggtagcactt ttcggggaaa tgtgcgcgga acccctattt gtttattttt ctaaatacat 60
tcaaatatgt atccgctcat gagacaataa ccctgataaa tgcttcaata atattgaaaa 120
aggaagagta tgagtattca acatttccgt gtcgccctta tccccttttt tgcggcattt 180
tgccctcctg tttttgctca ccagaaaacg ctgggtgaaag taaaagatgc tgaagatcag 240
ttgggtgcac gagtgggtta catcgaactg gatctcaaca gcggtgaagat ccttgagagt 300
tttgcggcgg aagaacgttt tccaatgatg agcactttta aagttctgct atgtggcgcg 360
gtattatccc gtattgacgc cgggcaagag caactcggtc gccgcataca ctattctcag 420
aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta 480
agagaattat gcagtgctgc cataaccatg agtgataaca ctgcccga cttacttctg 540
acaacgatcg gaggaccgaa ggagctaacc gcttttttgc acaacatggg ggatcatgta 600
actcgccttg atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac 660
accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt 720
actctagctt cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca 780
cttctgcgct cggcccttcc ggctggctgg tttattgctg ataaatctgg agccgggtgag 840
cgtgggtctc gcggtatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta 900
gttatctaca cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgctgag 960
ataggtgcct cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt 1020
tagattgatt taaaacttca tttttaattt aaaaggatct aggtgaagat cttttttgat 1080
aatctcatga ccaaaatccc ttaacgtgag ttttcgcttc actgagcgtc agaccccgta 1140
gaaaagatca aaggatcttc ttgagatcct tttttctg cgcgtaatctg ctgcttgcaa 1200
acaaaaaac caccgctacc agcgggtggt tgtttgccgg atcaagagct accaactctt 1260
tttccgaagg taactggctt cagcagagcg cagataccaa atactgtcct tctagtgtag 1320
ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta 1380
atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca 1440
agacgatagt taccggataa ggcgcagcgg tcgggctgaa cgggggggtc gtgcacacag 1500

```

cccagcttgg	agcgaacgac	ctacaccgaa	ctgagatacc	tacagcgtga	gctatgagaa	1560
agcgccacgc	ttcccgaagg	gagaaaaggcg	gacaggtatc	cggtaagcgg	cagggtcgga	1620
acaggagagc	gcacgagggg	gcttccaggg	gaaaacgcct	ggtatcttta	tagtcctgtc	1680
gggtttcgcc	acctctgact	tgagcgtcga	tttttgtgat	gctcgtcagg	ggggcggagc	1740
ctatggaaaa	acgccagcaa	cgcggccttt	ttacgggttc	tggccttttg	ctggcctttt	1800
gctcacatgt	tctttcctgc	gttatcccct	gattctgtgg	ataaccglat	taccgccttt	1860
gagtgagctg	ataccgctcg	ccgcagccga	acgaccgagc	gcagcgagtc	agtgagcgag	1920
gaagcggaag	agcgcccaat	acgcaaaccg	cctctccccg	cgcgttggcc	gattcattaa	1980
tgagctggc	acgacaggtt	tcccgactgg	aaagcgggca	gtgagcgcaa	cgcaattaat	2040
gtgagttagc	tcactcatta	ggcaccocag	gctttacact	ttatgcttcc	ggctcgtatg	2100
ttgtgtgaa	ttgtgagcgg	ataacaattt	cacacaggaa	acagctatga	ccatgattac	2160
gccaaagcgc	caattaacc	tcactaaagg	gaacaaaagc	tgggtaccgg	gccccccctc	2220
gaggtcgacg	gtatcgataa	gcttcaatgt	ttttagcacc	ctctgtgtgg	aggaaaataa	2280
tgacagattat	tctaattagt	gtaatatcta	accacattaa	aatatattac	atagtaaact	2340
acactccata	attttataaa	tttgactccc	cagggtaata	aactagtctc	tagtctgctc	2400
accttcaact	gtacaataaa	gtcttggttc	ttttgaaata	gacctcaaat	gagacaccta	2460
aaattcaaag	tgtctttaca	tttaaagaca	cctacaggaa	agcaggtaaa	agagccaggt	2520
taaaaacaaa	ttctaaaacc	acttagctgc	agttaaacat	atagtaaaga	tgactaaag	2580
tttcttactc	tgtaaatccc	ttccacttca	ggaaatattc	cactttccca	ttcactacac	2640
gtcgatctag	tactttttcc	acgacaaatt	cttcaggctc	tgctcttca	acttttttac	2700
tctttccatt	ctgttttttt	cccatttttt	gctaaaataa	aacaaaagag	aaattaagaa	2760
atattcctct	tgaattttga	gcacattttc	aaggctcaat	tgcttatatt	attatcacat	2820
tcgacataaa	tttttacttc	tatatcccct	ggcagacacc	ttctggaag	attaaaagtc	2880
aacagacaat	aaaataaaag	aatgctttat	cttgctcatt	tagttcaaac	ttacaacca	2940
ccaccaaaaat	aatacaataa	aaaaacacta	tctggaaaca	gttatttttt	tccagtcttt	3000
ttttttgaga	cagggctca	cactcttctc	gcccaggctg	gagtgagctg	gcgtgatctc	3060
agctcactgc	aacctccgcc	tcccagggtt	caagcagttc	tcatgcctca	gcctccagag	3120
tagctgggat	tatagcgga	tgccaccatg	ccgggcta	tttttttg	tttttattag	3180
aaacaggggt	tcaccatggt	gaccaggctg	gtctcaaact	cctgacctga	agtgattcac	3240
cagcctgggc	ctcccaaagt	gctggcatta	cagggctgag	ccactgcgcc	cggccctgta	3300
gtcttaaaaag	accaagttta	ctaattttca	ctcattttta	caacactgca	acaacaact	3360
atgcaggaag	tacctaaagg	gtgatccaga	gaagcaagta	gtagtgcag	gtcttaggtg	3420
aacctatgac	agaccttgta	tccaccccc	gatggtaaaa	gccccagccc	ccttctcaat	3480
tcaaatatta	atgtcaaaaag	catcaatgat	acagagaaaa	gataaatgca	gaatgaaaac	3540
atggttcaaa	atcctgatac	caactgcagg	gtcaactata	gagaccacta	ggaggttcaa	3600
ttaaaggaca	agattatttt	tccataatct	ctgtagataa	tatttcctac	cacttagaac	3660
aaaactataa	agctatcact	tcaagagacc	aacattacaa	atttatttta	attccctaag	3720
gtgaaaaaaa	tccttctctc	ctggtttctc	aagagaaagt	ctatactggt	aaccaaatc	3780
actttaaaca	ggcattttct	ttggatgac	actatttaag	agaagcagga	aaccaacgtg	3840
aaccagctct	ttccaatggc	tcaagatttc	ctatgagagg	actaaaaatg	gggaaaattt	3900
ttatgagagg	attaaaaatg	ggggaaaaaa	aaccctgaaa	tggttaatca	gaagatccta	3960
tgggctgaga	aggaatccat	cttaacattt	catcttaag	caaatgctat	tgccgggggc	4020
agtggctcat	gcctgtaatc	ccagcacttt	gggaggccga	ggtgggcaga	tcatctgagg	4080
tcaggagttt	gagaccagcc	tgaccaacat	ggagaaacc	cgtttctact	aaaaatacaa	4140
aattagccag	gcatagtggg	gcatgcctgt	aatcccagct	acttgggagg	ctgaggcagg	4200
agaactgctt	gaaccagga	ggcttaagtt	gcggtgagcc	aagatcacgc	cattgcactc	4260
tagcctggac	aacaagagaa	aaactctgtc	tcaaaaaaac	acaaaaaca	aaaacccaaa	4320
tactatttaa	aaaagataaa	ccttaattgc	tcaatcatta	aagccatccc	acaagtaaag	4380
cagcaagcag	aaaaaagtta	agaacacctc	aaggctacag	aaggacattt	caagctatgc	4440
aggcatatga	agtgtgcaga	cagatgtgta	agaaagccct	caagactgca	aaagggcatt	4500
tcaagctatg	caagcatata	ggtaacacat	acacacacac	aaaataaaat	cccctgaaat	4560
acaaaaacat	gcagcaaaca	cctgacgttt	ttggatacca	tttctaagtc	aggtgttatg	4620
attctcatta	gtcaagatac	ttgagtactg	ggcccaaa	gctttctgcc	actgtacagt	4680
acaagaaggt	aggaataatg	gtgggaggag	caaagacaaa	ctgtaataga	cagaagtgtg	4740
tcagatacct	atactacatg	aaaaacaaaa	cagctactgc	cacaaagga	gaaggctaac	4800
aaaataaagt	caacaataaa	tacagaaaat	gaaaaggata	cacactaagg	tttacaacaaa	4860
aaaaaaggca	gacaaaatgc	catacagtat	tcattcacta	ctatggcatt	cataagctag	4920

tttcaaatgc	tcactatthtt	cttttatagt	atatatttgc	cttaaccag	cacttttttc	4980
caaaagtggg	tgagtcaaaa	taaatttccc	attattttaag	tgaaattaac	agcacacata	5040
tctcacaaca	ctaataaatt	tttaaaatgg	aaagttaaga	acttttaaaag	tggccaacct	5100
gtgatccttc	acaaaataaa	ctaaatacaa	taacagacc	caaaggctat	caattgctgt	5160
caaaaacaac	ttctgttttc	cagggtaaac	agaatcta	gcagaatcta	atgcagggta	5220
aacagactta	atgcagaatc	taatgatggc	acaaatthaa	aatcactaac	gtgccctttt	5280
tagtgtgaaa	cccagagaga	gcacatacaa	gccaaaaaca	aatgctttat	tttacctagg	5340
agacattaac	attcaccttt	acgtgtttta	gattaatgca	atgttaaata	ttgtgaaaac	5400
tgtaactttg	aatttcatga	tttttatgtg	aatattccag	ggtttaaaaa	aacttghtaac	5460
atgacatggc	tgataaagat	aaaaaaaaaa	tctagccttt	tctcccttct	ggctcatatt	5520
tgcgatttcg	atcattttgt	ttaaaaaaca	aaacactgca	atgaatthaa	cttaatatct	5580
ttctatgttt	tagagtaagt	taaaacaaga	taaagtgacc	aaagtaattt	gaaagattca	5640
atgacttttg	ctccaacct	gggtcacaa	gtacctgtt	ctthaaattg	ggctthaaatg	5700
aaaatacttc	tccagaattc	tggggattta	agaaaaatta	tgccaacca	caagggcttt	5760
accattttat	gtaacatttt	tcaacgctgc	aaaaatgtgt	gtatttctat	ttgaagataa	5820
aaatcctcag	caaaatccac	attgcaactg	ccttcaaaga	ttagccttct	ttgaactagt	5880
taagacacta	ttaagccaag	ccagtatctc	cctgtaatga	attcgttttt	ctcttaattt	5940
tcccctgtaa	tttactactg	gagagctggg	aaatatgtgg	atgthaaattt	ctcagccaca	6000
gagatgcaaa	gttatactgt	ggggaaaaaa	aacttgagtt	aaatccttac	atatttttagg	6060
ttttcattaa	cttaccaatg	tagttttgtt	ggaggccatt	ttttttattg	cagacttgaa	6120
gagctattac	tagaaaaatg	catgacagtt	aaghtaagtt	tgcatgacac	aaaaaaggta	6180
actaaataca	aattctgttt	ggattccaac	ccccaaatag	agagcgcaca	ctttcaaacg	6240
tgaatacaaaa	tccagagtag	atctgctctc	ctacctacat	tgcttatgat	gtactthaaatg	6300
acgtgtccta	accatgtgag	tctagaaaaga	ctttactggg	gatcctggta	cctaaaaacag	6360
ctttcacatgg	ctthaaatag	gggaccaatg	tcttttccaa	tctaagtccc	atttataata	6420
aagtccatgt	tccattttta	aaggacaatc	ctttcggttt	aaaaccaggc	acgattacc	6480
aaacaactca	caacggtaaa	gcaactgtgaa	tcttctctgt	tctgcaatcc	caacttggtt	6540
tctgctcaga	aacctctcct	ctttccaatc	ggtaattthaa	taacaaaagg	aaaaaactta	6600
agatgcttca	accccgtttc	gtgacacttt	gaaaaaaaga	tcacctcttg	caaacacccg	6660
ctcccagacc	ccgcccgtga	agcccggcgt	ccagaggcct	aagcgcgggt	gcccggcccc	6720
acccgggagc	gcgggcctcg	tggtcagcgc	atccgcgggg	agaaacaaag	gccgcggcac	6780
gggggctcaa	gggcaactgc	ccacaccgca	cgccctacc	cccgcgcggc	cacgttaact	6840
ggcggctcgc	gcagcctcgg	gacagccggc	cgccgcggc	caggctcgcg	gacgcgggac	6900
cacgcgcgcg	cctccgggag	gcccagttct	cgaccagcc	ccgcgtggcg	ctgggggagg	6960
gggcgcctcc	gccggaacgc	gggtggggga	ggggaggggg	aaatgcgctt	tgtctcgaaa	7020
tggggcaacc	gtcggccacag	ctccctacc	cctcgagggc	agagcagttc	ccccactaac	7080
taccgggctg	gccgcgcgcg	aggccagcc	cgaggccacc	gcccgaacct	ccactcttc	7140
cccgagctcc	cgtgcggggg	tccgycgaga	aggggagggg	aggggagcgg	agaaccgggc	7200
ccccgggagc	cggtgtggcat	ctgaagcacc	accagcggag	gagagctaga	gagaaggaaa	7260
gccaccgact	tcaccgcctc	cgagctgctc	cgggctcgcg	gtctgcagcg	tctccggccc	7320
tccgcgccta	cagctcaagc	cacatccgaa	gggggagggg	gccgggagct	gcgcgcgggg	7380
ccgcccgggg	gaggggtggc	accgcccacg	ccgggcggcc	acgaaggcgc	gggcagcggg	7440
cgcgcgcgcg	gcggggggag	ggccggcgc	cgcccccgt	gggaattggg	gccctagggg	7500
gagggcggag	gcgcccagca	ccgcccact	taccgttcgc	ggcgtggcgc	ccgggtggcc	7560
ccaaggggag	ggaaggggga	ggcggggcga	ggacagtgac	cggagttctc	tcagcgggtg	7620
ctttctgct	tggcagcctc	agcggctggc	gccaaaaacc	gactccgccc	acttccctcg	7680
ccgcccgtgc	gaggggtgtg	aatcctccag	acgctggggg	agggggagtt	gggagctthaa	7740
aaactagtag	ccctttggga	ccactttcag	cagcgaactc	tcctgtacac	caggggtcag	7800
ttccacagac	gcgggcccag	ggtgggtcat	tgccgcgtga	acaataattt	gactagaagt	7860
tgattcgggt	gtttccggaa	ggggccgag	caatccgccg	agttggggca	cggaaaaaca	7920
aaagggaaag	ctactaagat	ttttctggcg	ggggttatca	ttggcgtaac	tgcagggacc	7980
acctcccggg	ttgagggggc	tggatctcca	ggctgcggat	taagccctc	ccgtcggcgt	8040
taatttcaaa	ctgcgcgacg	tttctcacct	gccttcgcca	aggcaggggc	cgggacccta	8100
ttccaagagg	tagtaactag	caggactcta	gccttccgca	attcattgag	cgcatttacg	8160
gaagtaacgt	cgggtactgt	ctctggccgc	aagggtgggg	ggagtaacgca	tttggcgtaa	8220
ggtggggcgt	agagccttcc	cgccattggc	ggcggatagg	gcgtttacgc	gacggcctga	8280
cgtagcggaa	gacgcgttag	tgggggggaa	ggttctagaa	aagcggcggc	agcggctcta	8340

gcggcagtag	cagcagcgcc	gggtcccgtg	cggaggtgct	cctcgcagag	ttgtttctcg	8400
agcagcggca	gttctcacta	cagcgccagg	acgagtcg	ttcgtgttcg	tccgcggaga	8460
tctctctcat	ctcgcctcggc	tgcgggaaat	cgggctgaag	cgactgagtc	cgcgatggag	8520
gtaacgggtt	tgaatcaat	gagttattga	aaagggcatg	gcgaggccgt	tggcgcctca	8580
gtggaagtcg	gccagccgcc	tccgtgggag	agaggcagga	aatcggacca	attcagtagc	8640
agtggggcct	aaggtttatg	aacgggggtc	tgagcggagg	cctgagcgta	caaacagctt	8700
ccccaccctc	agcctcccgg	cgccatttcc	cttcaactggg	ggtgggggat	ggggagcttt	8760
ccatggcgg	acgctgcccc	gctgggggtg	aagtggggcg	cggaggcggg	aattcttatt	8820
ccctttctaa	agcacgctgc	ttcgggggcc	acggcgcttc	ctcggcgagc	gtttcggcgg	8880
gcagcaggtc	ctcgtgagcg	aggctgcgga	gcttcccctc	cccctctctc	cggggaaccg	8940
atgtggcggc	cgccatttcc	atggctcgcc	ttcctctcag	cgttttcctt	ataactcttt	9000
tatgtttctt	gtgtgctttc	tctatcaaga	agtagaagtg	gttaactatt	ttttttttct	9060
tctcgggctg	ttttcatatc	gtttcgagg	ggatttggag	tgttttgtga	gcttggatct	9120
ttagagtcct	gcgcacctca	ttaaaggcgc	tcagccttcc	cctcgatgaa	atggcgccat	9180
tgcgttcggg	agccacaccg	aagagcgggg	aggggggggtg	ctccgggttt	gcgggcccgg	9240
tttcagagaa	gatatcacca	cccagggcgt	cgggcccgggt	tcaatgcgag	ccgtaggaca	9300
aagaaacat	tttatgtttt	tccctgtctt	tttttccttt	gagtaacggt	tttatctggg	9360
tctgcagtca	gtaaaacgac	agatgaaccg	cggcaaaata	aacataaatt	ggaagccatc	9420
ggccacgagg	ggcagggagc	aagggtggtt	tctgggcggg	ggagggatat	tccggtcaga	9480
atcctttact	gttcttaagg	attccgttta	agttgtagag	ctgactcatt	ttaagtaatg	9540
ttgttactga	gaagttaaac	ccttacggga	cagatccatg	gacctttata	gatgattacg	9600
aggaaagtga	aataacgatt	ttgtccttag	ttataactcg	attaaaacat	ggcttcagag	9660
gctccttctc	gtaatgcgta	tggattgatg	tgcaaaactg	ttttggcctc	gggcccgtct	9720
gtatgtgaac	ttgtttactt	ttctcatttt	gtttgcaatc	ttggttgaac	attacattga	9780
taagcataag	gtctcaagcg	aaggggtctc	acctggttat	ttttctttga	ccctaagcac	9840
gtttataaaa	taacattggt	taaaatcgat	agtggacatc	gggtaagttt	ggataaattg	9900
tgaggtaagt	aatgagtttt	tgctttttgt	tagtgatttg	taaaacttgt	tataaatgta	9960
cattatccgt	aatttcagtt	tagagataac	ctatgtgctg	acgacaatta	agaataaaaa	10020
ctagctgaaa	aatgaaaat	aactatcgtg	acaagtaacc	atttcaaaag	actgctttgt	10080
gtctcatagg	agctagtttg	atcatttcag	ttaatttttt	ctttaatttt	tacgagtcac	10140
gaaaactaca	ggaaaaaaa	tctgaactgg	gttttaccac	tacttttttag	gagttgggag	10200
catgccaatg	gagggagagc	tccgtagaac	tgggatgaga	gcagcaatta	atgctgcttg	10260
ctaggaacaa	aaaataattg	attgaaaatt	acgtgtgact	tttttagttg	cattatgcgt	10320
ttgtagcagt	tggtcctgga	tatcactttc	tctcgtttga	ggttttttaa	cctagttaac	10380
tttaagaca	ggtttcctta	acattcataa	gtgccagaa	tacagctgtg	tagtacagca	10440
tataaagatt	tcagctctga	ggtttttcct	attgacttgg	aaaattggtt	tgtcctgtc	10500
gcttgccaca	tggccaatca	agtaagcttg	attaatagta	atcaattacg	gggtcattag	10560
ttcatagccc	atatatggag	ttccgcgtta	cataacttac	ggtaaatggc	ccgcttggct	10620
gaccgccc	cgacccccgc	ccattgacgt	caataatgac	gtatgttccc	atagtaaacgc	10680
caatagggac	tttccattga	cgtcaatggg	tggagtattt	acggtaaaact	gcccacttgg	10740
cagtacatca	agtgtatcat	atgccaagta	cgccccctat	tgacgtcaat	gacggtaaat	10800
ggcccgcctg	gcattatgcc	cagtacatga	ccttatggga	ctttcctact	tggcagtaca	10860
tctacgtatt	agtcatcgct	attaccatgg	tgatgcgggt	ttggcagtac	atcaatgggc	10920
gtggatagcg	gtttgactca	cggggatttc	caagtctcca	ccccattgac	gtcaatggga	10980
gtttgttttg	gcacaaaaat	caacgggact	ttccaaaaatg	tcgtaacaac	tccgccccat	11040
tgacgcaaat	ggcggttagg	cgtgtacggg	gggaggtcta	tataagcaga	gctgggttag	11100
tgaaccgtca	gatccgctag	ccggtcgcca	ccatgggtgag	caagggcgag	gagctgttca	11160
ccgggggtgt	gcccacctc	gtcagctg	acggcgacgt	aaacggccac	aagttcagcg	11220
tgccggcga	ggcgagggc	gatgccacct	acggcaagct	gaccctgaag	ttcatctgca	11280
ccaccggcaa	gctgcccgtg	ccctggccca	ccctcgtgac	caccctgacc	tacggcgtgc	11340
agtgttcag	ccgctacccc	gaccacatga	agcagcacga	cttcttcaag	tccgccatgc	11400
ccgaaggcta	cgtccaggag	cgcacatct	tcttcaagga	cgacggcaac	tacaagacc	11460
gcgccgaggt	gaagtccgag	ggcgacaccc	tgggtgaaccg	catcgagctg	aagggcatcg	11520
acttcaagga	ggacggcaac	atcctggggc	acaagctgga	gtacaactac	aacagccaca	11580
acgtctatat	catggccgac	aagcagaaga	acggcatcaa	ggtgaacttc	aagatccgcc	11640
acaacatcga	ggacggcagc	gtgcagctcg	ccgaccacta	ccagcagaac	acccccatcg	11700
gcgacggccc	cgtgctgctg	cccgacaacc	actacctgag	caccagtc	gccctgagca	11760

```

aagaccccaa cgagaagcgc gatcacatgg tcctgctgga gttcgtgacc gccgcccggga 11820
tcactctcgg catggacgag ctgtacaagt aaagcggccg cgactctaga tcataatcag 11880
ccataccaca tttgtagagg ttttacttgc tttaaaaaac ctcccacacc tccccctgaa 11940
cctgaaacat aaaatgaatg caattgttgt tgtaacttg tttattgcag cttataatgg 12000
ttacaaataa agcaatagca tcacaaatth cacaaataaa gcattttttt cactgcattc 12060
tagttgtggt ttgtcctaac tcatcaatgt atcttaaadc gaattctacc gggtagggga 12120
ggcgcttttc ccaaggcagt ctggagcatg cgcttagca gccccgctgg gcaactggcg 12180
ctacacaagt ggcctctggc ctgcacaca ttccacatcc accggtaggc gccaaccggc 12240
tccgttcttt ggtggcccc tgcgccacc ttctactcct cccctagtca ggaagtccc 12300
ccccgccccg cagctcgcgt cgtgcaggac gtgacaaatg gaagtagcac gtctcactag 12360
tctcgtgcag atggacagca ccgctgagca atggaagcgg gtaggccttt ggggcagcgg 12420
ccaatagcag ctttgcctct tcgctttctg ggctcagagg ctgggaaggg gtgggtccgg 12480
gggcgggctc aggggcgggc tcaggggcgg ggcgggcgcc cgaaggtcct ccggaggccc 12540
ggcattctgc acgcttcaaa agcgcacgct tgcgcgctg ttctcctctt cctcatctcc 12600
gggcctttcg accagcttac catgaccgag tacaagcca cgggtgcgct ccgcccccgc 12660
gacgacgtcc ccagggcctg acgcaccctc gccgcccgtc tcgccgacta cccccccagc 12720
cgccacaccg tcgatccgga ccgccacatc gagcgggtca ccgagctgca agaactcttc 12780
ctcacgcgcg tggggctcga catcggcaag gtgtgggtcg cggacgacgg cgccgcggtg 12840
gcggtctgga ccacgccgga gagcgtcga gcgggggcgg tgttcgccga gatcgccccg 12900
cgcatggccg agttgagcgg ttcccggctg gccgcgcaga acagatggaa ggcctcctgg 12960
cgccgcaccg gccaaggag cccgcgtggt tctggccac cgtcgcgtct cgcccagca 13020
ccagggcaag ggtctgggca gcgcgtcgt gctccccgga gtggaggcgg ccgagcgcgc 13080
cggggtgcc gccttcttgg agacctccgc gccccgcaac ctccccttct acgagcggct 13140
cggcttcacc gtcaccgccc acgtcagagt gccccaagga ccgcgcacct ggtgcatgac 13200
ccgcaagccc ggtgcctgac gcccgcccc cgacccgcag cgcccgaccg aaaggagcgc 13260
acgaccccat gcacgtaga gctcgtgat cagcctcgc tgtgccttct agttgccagc 13320
catctgttgt ttgcccctcc cccgtgcctt ccttgaccct ggaagggtgc actcccactg 13380
tcctttccta ataaaatgag gaaattgcat cgcattgtct gagtaggtgt cattctattc 13440
tggggggtgg ggtggggcag gacagcaagg ggggggattg ggragacaat agcaggcatg 13500
ctgggggggc ggtgggggct atggcttctg aggcggaaaag aaccagctgg ggctcgagat 13560
ccactagttc tagcctcgag gctagagcgg ccgccaccgc ggtggagctc caattcgccc 13620
tatagtgagt cgtattacgc gcgctcactg gccgctggtt tacaacgtcg tgactgggaa 13680
aaccctggcg ttacccaact taatcgctt gcagcacatc cccctttcgc cagctggcgt 13740
aatagcgaag agggccgcac cgatcgccct tcccaacagt tgcgcagcct gaatggcgaa 13800
tggaattgt aagcgttaat attttgtaa aattcgcgtt aaattttgt taaatcagct 13860
catttttaa ccaatagcc gaaatcggca aaatccctta taaatcaaaa gaatagaccg 13920
agataggtgt gagtgtgtt ccagtttggg acaagagtc actattaaag aacgtggact 13980
ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg cccactacgt gaaccatcac 14040
cctaatcaag ttttttgggg tcgaggtgcc gtaaagcact aaatcggaac cctaaagggg 14100
gccccgatt tagagcttga cggggaaagc cggcgaaagt ggcgagaaag gaaggggaga 14160
aagcgaag agcgggcgct agggcgctgg caagtgtagc ggtcacgctg cgcgtaacca 14220
ccacaccg cgcgcttaat gcgcccgtac agggcgcgct ag 14262

```

<210> 10

<211> 13

<212> DNA

<213> 人工序列

<220>

<223> PCR引物

<400> 10

aacaattggc ggc

13

<210> 11

<211> 13

<212> DNA
 <213> 人工序列

 <220>
 <223> PCR 引物

 <400> 11
 gcccaattggt gcc 13

 <210> 12
 <211> 31
 <212> DNA
 <213> 人工序列

 <220>
 <223> PCR 引物

 <400> 12
 acgcgtcgc ggaaggagac aataccggaa g 31

 <210> 13
 <211> 28
 <212> DNA
 <213> 人工序列

 <220>
 <223> PCR 引物

 <400> 13
 ccgctcgagt tggggtgggg aaaaggaa 28

 <210> 14
 <211> 30
 <212> DNA
 <213> 人工序列

 <220>
 <223> PCR 引物

 <400> 14
 cgggatccgc ctgagaaagg aagtgagctg 30

 <210> 15
 <211> 29
 <212> DNA
 <213> 人工序列

 <220>
 <223> PCR 引物

 <400> 15
 gaagatctgg aggaatgagc tggccctta 29

 <210> 16
 <211> 8

<212> DNA	
<213> 人工序列	
<220>	
<223> PCR 引物	
<400> 16	
gactagtc	8
<210> 17	
<211> 35	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR 引物	
<400> 17	
ctcgagttat taatagtaat caattacggg gtcac	35
<210> 18	
<211> 33	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR 引物	
<400> 18	
gtcgcagcgtc tgacgggtca ctaaaccagc tct	33
<210> 19	
<211> 30	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR 引物	
<400> 19	
ccaatgcata gggtgggctt cggaatcgt	30
<210> 20	
<211> 31	
<212> DNA	
<213> 人工序列	
<220>	
<223> PCR 引物	
<400> 20	
gctctagatc tcgacgggtat acagacatga t	31
<210> 21	
<211> 36	

<212> DNA
 <213> 人工序列

 <220>
 <223> PCR 引物

 <400> 21
 cccaagctta ttaatagtaa tcaattacgg ggtcat 36

 <210> 22
 <211> 36
 <212> DNA
 <213> 人工序列

 <220>
 <223> PCR 引物

 <400> 22
 caaggatccg atctgacggt tcaactaaacc agctct 36

 <210> 23
 <211> 20
 <212> DNA
 <213> 人工序列

 <220>
 <223> PCR 引物

 <400> 23
 tcgagtcggt taaactctag 20

 <210> 24
 <211> 20
 <212> DNA
 <213> 人工序列

 <220>
 <223> PCR 引物

 <400> 24
 tcgactagag tttaaacgac 20

 <210> 25
 <211> 33
 <212> DNA
 <213> 人工序列

 <220>
 <223> PCR 引物

 <400> 25
 gaattcgagc tcgccaact ccgcccgttt tat 33

 <210> 26
 <211> 39

<212> DNA
<213> 人工序列

<220>

<223> PCR 引物

<400> 26

atattgtcgac tctagacccg ggctgcagcg aggagctct

39

<210> 27

<211> 12588

<212> DNA

<213> 载体

<400> 27

```

acgttgtaaa acgacggcca gtgaattgta atacgactca ctatagggcg aattgggtac 60
cgggccccc ctcgaggctg agttggggtg gggaaaagga agaaacgcgg gcgtattggc 120
cccaatgggg tctcggtggg gtatcgacag agtgccagcc ctgggaccga accccgcgtt 180
tatgaacaaa cgacccaaca cccgtgcggt ttattctgtc tttttattgc cgtcatagcg 240
cgggttcctt ccggtattgt ctcttccgt cgacgatctg acggttcaact aaaccagctc 300
tgcttatata gacctccac cgtacacgcc taccgcccac ttgctgcaat ggggcggagt 360
tgttacgaca ttttgaaaag tcccgtgat ttgggtgcca aaacaaactc ccattgacgt 420
caatggggtg gagacttggg aatccccgtg agtcaaaccg ctatccacgc ccattgatgt 480
actgccaaaa ccgcatcacc atggtaatag cgatgactaa tacgtagatg tactgccaag 540
taggaaagtc ccataaggtc atgtactggg cataatgcca ggccgggcat ttaccgtcat 600
tgacgtcaat agggggcgta cttggcatat gatacacttg atgtactgcc aagtgggcag 660
tttaccgtaa atactccacc cattgacgtc aatggaaagt ccctattggc gttactatgg 720
gaacatacgt cattattgac gtcaatgggc gggggctcgtt gggcggtcag ccaggcgggc 780
catttaccgt aagttagta acgcggaact ccataatgg gctatgaact aatgaccccg 840
taattgatta ctattaataa ctcgacggta tcatgggtggc gaccggcatg gtgagctgcg 900
agaatagccg ggcgcgctgt gagccgaagt cgccccgcc ctggccactt ccggcgcgcc 960
gagtccttag gccgccaggg ggcgccggcg cgcgcccaga ttggggacaa aggaagccgg 1020
gccggccgcg ttattacccat aaaaggcaaa cactggctcg aggcgtccc gccggcgcg 1080
gcaggaagcc agggcccaac cccctccca cggggcgcca gccccgcctc cgcccggttc 1140
aaacagcgac cgggtcgcgc gcgcgcacgc agcggccaca ccctcgggcg ccagcggctc 1200
gggcaggaag tggcgcaagc gcccgggccc cagaacgcac gcgcgattag cgccattgag 1260
tcccagcgcg cacgcgcaat tagcgccaat tcccagcgcg cacgcagtta gcgccc aaaag 1320
gaccagcgcg cacgcgcatg gcgcccagc ccccaccggg cctgacgggg gctacgccgc 1380
gcccaccgtg cgatccccat tggcaagagc ccggctcaga caaagacccc gccggttgcc 1440
cccgcccgga gagcggcacc cccggagcgc gcccgcccga gcgcggcctc gcgctgcga 1500
actggcgtgg ggtgtcccc atctccggag gccaggggc ttctcccgcg cccccacgg 1560
cggtcgggtt ccgccccatg cgccccccgc tgcggcccag acggcggctc tgacggggcg 1620
aaggcccgcg gccgcatgcc ccggtcggct ggccgggctt acctggcggc ggggtgtggac 1680
ggcgcgcgga tcggcaaaag cgaggctctg tgctcgcggg cggacgcggg ctcggcgggtg 1740
gtggcgcgtc gcgcccgtgg gttttatagg gcgcccgcgc ggccgctcga gccataaaaag 1800
gcaactttcg gaacggcgca cgctgattgg ccccgcgccg ctcaactcacc ggcttcgccg 1860
cacagtgcag cattttttta ccccctctcc cctccttttg cgaaaaaaaa aaagagcgag 1920
agcgagattg aggaagagga ggagggagag ttttggcggt ggccgccttg gggtgctggg 1980
cccgggggct gggggcgcgc gccgtggccc ccgcgcccc cgctgggcag tgcccgggtc 2040
ggccccgcat ggccaggcct gccccggcc tgcccgtctc tcgggcccc caccaccgc 2100
gggacatcct aggtgtggac atctcttggg cactgagcgc ccaggtgggg tgggccaggg 2160
tctgcacggg tgccagggcc ctgggttctg tacgctcctg cagaaggagc tcttggaggg 2220
catggagtgg ccaggcagtc actccccctt gccgacttca gagcaactgc cctgaaagca 2280
gggcctgagg acctctggct gtggggctca gctagctaaa tgtgctgggt gggtcactag 2340
ggagagacct gggcttgaga ggtagagtgt ggtgttgggg gagtcaaggc gcttgcggcc 2400

```

attagagtcg	caggaccaca	ctccccagga	cagggcaggg	gccagcggtc	cagtggctgg	2460
aggtggcccc	tgatgaaggc	tacaaaccta	cccagccgca	gccctgggaa	ggaagtgggc	2520
tctacagggc	agggcacctt	ttaccctgga	gctgcctgct	tttgagggta	acagtcacgc	2580
ccagccaaga	ccaggcctgg	ggcgtagtg	ggtgacctag	gcaactgcggg	gcgggggggc	2640
tgggtctaca	cagcctgggt	ctgggcccac	cgctccgttg	atgtctgcta	tgcgcagcca	2700
cagctgaact	gccctcccag	accatctgga	ggcgcctggg	ggactctggg	gaccaagact	2760
ccatgtgcca	cagaggattg	ggggcggggc	ggtgctagga	actcaaagcc	agcctgggaa	2820
gaccctgtcc	ttgtcacctt	ttcttgccct	gggtctgtcc	actgagtagc	acacaagacc	2880
gggtgggcag	ggtccgttct	gctccgggaa	tcacagactg	tgtgtacca	ggtggtgggc	2940
atgcagcgat	cagtggcggtg	ggaccacaga	ggggcccgc	ggtacctaaa	acagcttcac	3000
atggcttaaa	ataggggacc	aatgtctttt	ccaatctaag	tcccatttat	aataaagtcc	3060
atgttccatt	tttaaaggac	aatcctttcg	gtttaaaacc	aggcacgatt	acccaaacia	3120
ctcacaacgg	taaagcactg	tgaatcttct	ctgttctgca	atcccaactt	ggtttctgct	3180
cagaaacctt	ccctctttcc	aatcggtaat	taaataacaa	aaggaaaaaa	cttaagatgc	3240
ttcaaccccc	tttcgtgaca	ctttgaaaaa	agaatcacct	cttgcaaaaa	cccgtctccg	3300
acccccgccg	ctgaagcccc	gcgtccagag	gcctaagcgc	gggtgccccg	ccccaccggg	3360
gagcgcgggc	ctcgtgggtca	gcgcatccgc	ggggagaaac	aaaggccgcg	gcacgggggc	3420
tcaagggcac	tgcgccacac	cgacgcgcc	tacccccgcg	cgccacggtt	aactggcggg	3480
cgccgcagcc	tcgggacagc	cgccgcgcgc	cgccagggct	cgccgacgcg	ggaccaagcg	3540
ccgccctccg	ggaggcccaa	gtctcgacct	agccccgcgt	ggcgcctggg	gagggggcgc	3600
ctccgccgga	acgccccgtg	gggaggggag	ggggaaatgc	gctttgtctc	gaaatggggc	3660
aaccgtcgcc	acagctccct	acccccctga	ggcagagca	gtccccccac	taactaccgg	3720
gctggccgcg	cgccaggcca	gccgcgaggc	caccgcccga	ccctccactc	cttccccag	3780
ctccccgcgc	gggttccggc	gagaagggga	ggggagggga	gcggagaacc	gggcccccg	3840
gacgcgtgtg	gcatctgaag	caccaccagc	gagcgcgagc	tagagagaag	gaaagccacc	3900
gacttcaccg	cctccgagct	gctccgggtc	gcgggtctgc	agcgtctccg	gccctccgcg	3960
cctacagctc	aagccacatc	cgaaggggga	gggagccggg	agctgcgcgc	ggggccgccc	4020
gggggagggg	tggcaccgcc	cacgccgggc	ggccacgaag	ggcggggcag	cgggcgcgcg	4080
cgcgccgggg	ggaggggccc	gcgccgcgcc	cgctgggaat	tggggcccta	gggggagggc	4140
ggagggcccg	acgaccgcgg	cacttaccgt	tcgcccgcgtg	gcgccgggtg	gtccccagg	4200
ggagggaaag	gggagggcgg	gcgaggacag	tgaccggagt	ctcctcagcg	gtggcttttc	4260
tgcttggcag	cctcagcggc	tggcgccaaa	accggactcc	gccccacttc	tcgccccccg	4320
gtgcgagggg	gtggaatcct	ccagacgctg	ggggaggggg	agttggggag	ttaaaaacta	4380
gtaccocctt	gggaccactt	tcagcagcga	actctcctgt	acaccagggg	tcagttccac	4440
agacgcgggc	caggggtggg	tcattgcggc	gtgaacaata	atttgactag	aagttagatt	4500
gggtgtttcc	ggaaggggccc	gagtcaatcc	gccgagttgg	ggcacggaaa	acaaaaaggg	4560
aaggctacta	agatttttct	ggcggggggt	atcattggcg	taactgcagc	gaccacctcc	4620
cggttggagg	ggcctggatc	tcaggctgc	ggattaagcc	cctcccgtcg	gcgttaattt	4680
caaaactgccc	gacgtttctc	acctgccttc	gccaaaggcag	gggcccgggac	cctattccaa	4740
gaggtagtaa	ctagcaggac	tctagccttc	cgcaattcat	tgagcgcatt	tacggaagta	4800
acgtcgggta	ctgtctctgg	ccgcaagggt	gggaggagta	cgcatttggc	gtaaggtggg	4860
gcgtagagcc	ttcccgccat	tggcggcgga	tagggcgttt	acgcgacggc	ctgacgtagc	4920
ggaagacgcg	ttagtggggg	ggaaggttct	agaaaagcgg	cggcagcggc	tctagcggca	4980
gtagcagcag	cgccgggtcc	cgtgcggagg	tgctcctcgc	agagttgttt	ctcagacagc	5040
ggcagttctc	actacagcgc	caggacgagt	ccggttcgtg	ttcgtccgcg	gagatctctc	5100
tcatctcgct	cggctgcggg	aaatcgggct	gaagcgcactg	agtccgcgat	ggaggtaacg	5160
ggtttgaaat	caatgagtta	ttgaaaaggg	catggcgagg	ccgttggcgc	ctcagtggaa	5220
gtcggccagc	cgctccgtg	ggagagaggg	aggaaatcgg	accaattcag	tagcagtggg	5280
gcttaagggt	tatgaacggg	gtcttgagcg	gaggcctgag	cgtacaaaca	gcttccccac	5340
cctcagcctc	ccggcgccat	ttcccttcac	tgggggtggg	ggatggggag	ctttcacatg	5400
gcggacgctg	ccccgctggg	gtgaaagtgg	ggcgcggagg	cggaattct	tattcccttt	5460
ctaaagcacg	ctcgttcggg	ggccacggcg	tctcctcggc	gagcgtttcg	gcgggcagca	5520
ggtcctcggtg	agcagggctg	cgagcttcc	cctccccctc	tctcccggga	accgatttgg	5580
cgcccgccat	tttcatggct	cgcttctctc	tcagcgtttt	ccttataact	cttttatttt	5640
cttagtggtc	tttctctatc	aagaagtaga	agtggttaac	tatttttttt	ttcttctcgg	5700
gctgttttca	tatcgtttcg	aggtggattt	ggagtgtttt	gtgagcttgg	atcttttagag	5760
tcctgcgcac	ctcattaaag	gcgctcagcc	ttccccctga	tgaaatggcg	ccattgcggt	5820

cggaagccac	accgaagagc	ggggaggggg	ggtgctccgg	gtttgcgggc	ccggtttcag	5880
agaagatatic	accaccagg	gcgtcgggcc	gggttcaatg	cgagccgtag	gacaaagaaa	5940
ccattttatg	tttttcctgt	cttttttttc	ctttgagtaa	cggttttatc	tgggtctgca	6000
gtcagtaaaa	cgacagatga	accgcggcaa	aataaacata	aattggaagc	catcggccac	6060
gaggggcagg	gacgaaggtg	gtttctctggg	cgggggaggg	atattcgcgt	cagaatcctt	6120
tactgttctt	aaggattccg	tttaagttgt	agagctgact	cattttaagt	aatgttgta	6180
ctgagaagtt	taacccttac	gggacagatc	catggacctt	tatagatgat	tacgaggaaa	6240
gtgaaataac	gattttgtcc	ttagttatac	ttcgattaaa	acatggcttc	agaggctcct	6300
tcctgtaatg	cgtatggatt	gatgtgcaaa	actgttttgg	gcctgggccc	ctctgtattt	6360
gaactttggt	acttttctca	ttttgtttgc	aatcttgggt	gaacattaca	ttgataagca	6420
taaggtctca	agcgaagggg	gtctacctgg	ttatttttct	ttgaccctaa	gcagtttat	6480
aaaataacat	tgtttaaaat	cgatagtggg	catcgggtaa	gtttggataa	attgtgaggt	6540
aagtaatgag	tttttgcttt	ttgttagtga	tttgtaaaac	ttgttataaa	tgtacattat	6600
ccgtaatttc	agtttagaga	taacctatgt	gctgacgaca	attaagaata	aaaactagct	6660
gaaaaaatga	aaataactat	ctgacaagt	aaccatttca	aaagactgct	tttgtctca	6720
taggagctag	tttgatcatt	tcagttaatt	ttttctttaa	tttttacgag	tcatagaaaac	6780
tacaggaaaa	aaaatctgaa	ctgggtttta	ccactacttt	ttaggagttg	ggagcatgcg	6840
aatggaggga	gagctccgta	gaactgggat	gagagcagca	attaatgctg	cttgctagga	6900
acaaaaata	attgattgaa	aattacgtgt	gactttttag	tttgattat	gcgtttgtag	6960
cagttggctc	tggatatcac	tttctctcgt	ttgaggtttt	ttaacctagt	taacttttaa	7020
gacaggtttc	cttaacattc	ataagtgcc	agaatacagc	tgtgtagtac	agcatataaa	7080
gatttcagct	ctgaggtttt	tcctattgac	ttggaaaatt	gttttgtgcc	tgtcgcttgc	7140
cacatggcca	atcaagtaag	cttcgaattc	gagctcgccc	aactccgccc	gttttatgac	7200
tagaaccaat	agtttttaat	gccaaatgca	ctgaaatccc	ctaatttgca	aagccaaacg	7260
cccctatgt	gagtaatacg	gggacttttt	acccaatttc	ccaagcggaa	agccccctaa	7320
tacactcata	tggcatatga	atcagcacgg	tcactgactc	taatggcggc	ccatagggac	7380
ttccacata	gggggcgctc	accatttccc	agcatagggg	tggtgactca	atggccttta	7440
cccaagtaca	ttgggtcaat	gggaggtaa	ccaatgggtt	tttcccatta	ctggcaagca	7500
caactgagca	aatgggactt	tccactgggt	tttgcccagg	tacattgggt	caatgggagg	7560
tgagccaatg	ggaaaaacc	attgctgcca	agtacactga	ctcaataggg	actttccaat	7620
gggtttttcc	attgttggca	agcatataag	gtcaatgtgg	gtgagtcaat	agggactttc	7680
cattgtattc	tgcccagtac	ataaggtcaa	taggggggtga	atcaacagga	aagtcccatt	7740
ggagccaagt	acactgcgtc	aatagggact	ttccattggg	ttttgcccag	tacataaggt	7800
caatagggga	tgagtcaatg	ggaaaaacc	attggagcca	agtacactga	ctcaataggg	7860
actttccatt	gggttttgcc	cagtacataa	ggtcaatagg	gggtgagtca	acaggaaagt	7920
cccattggag	ccaagtacat	tgagtcaata	gggactttcc	aatgggtttt	gccagtagca	7980
taaggtcaat	gggaggtaa	ccaatgggtt	ttcccatta	ctggcacgta	tactagtca	8040
ttagggactt	tccaatgggt	tttgcccagt	acataagggtc	aataggggtg	aatcaacagg	8100
aaagtcccat	tggagccaag	tacactgagt	caatagggac	tttccattgg	gttttgccca	8160
gtacaaaagg	tcaatagggg	gtgagtcaat	gggtttttcc	cattattggc	acgtacataa	8220
ggtcaatagg	ggtgagtcat	tgggtttttc	cagccaattt	aattaaaacg	ccatgtactt	8280
tcccaccatt	gacgtcaatg	ggctattgaa	actaatgcaa	cgtgaccttt	aaacgggtact	8340
ttcccatagc	tgattaatgg	gaaagtaccg	ttctcgagcc	aatacacgtc	aatgggaagt	8400
gaaagggcag	ccaaaacgta	acaccgcccc	ggttttcccc	tggaaattcc	atattggcac	8460
gcattctatt	ggctgagctg	cgttctacgt	gggtataaga	ggcgcgacca	gcgtcggtag	8520
cgtcgcagtc	ttcgggtctga	ccaccgtaga	acgcagagct	cctcgtgca	gcccgggtct	8580
agaggatccg	cctgagaaa	gaagtgagct	gtaaaggctg	agctctctct	ctgacgtatg	8640
tagcctctgg	ttagcttctg	cactcactgt	tcttgactca	gcatggcaat	ctgatgaaat	8700
cccagctgta	agtctgcaga	aattgatgat	ctattaaaca	ataaagatgt	ccactaaaat	8760
ggaagttttt	cctgtcatac	tttgtaaga	agggtgagaa	cagagtacct	acattttgaa	8820
tggaaggatt	ggagctacgg	gggtgggggt	gggggtgggat	tagataaatg	cctgctcttt	8880
actgaaggct	ctttactatt	gctttatgat	aatgtttcat	agttggatat	cataatttaa	8940
acaagcaaaa	ccaattaag	ggccagctca	ttcctccaga	tccactagtt	ctagagcaaa	9000
ttctaccggg	taggggaggc	gcttttccca	aggcagctct	gagcatgctc	tttagcagcc	9060
ccgctgggca	cttggcgcta	cacaagtggc	ctctggcctc	gcacacattc	cacatccacc	9120
ggtaggcgct	aaccggctcc	gttctttggg	ggccccttcg	cgccaccttc	tactcctccc	9180
ctagtcagga	agttcccccc	cgccccgcag	ctcgcgtcgt	gcaggacgtg	acaaatggaa	9240

gtagcacgtc	tactagtct	cgtgcagatg	gacagcaccg	ctgagcaatg	gaagcgggta	9300
ggcctttggg	gcagcggcca	atagcagctt	tgctccttcg	ctttctgggc	tcagaggctg	9360
ggaaggggtg	ggtccggggg	cgggctcagg	ggcgggctca	ggggcggggc	gggcgcccga	9420
aggtcctccg	gaggcccggc	attctgcacg	cttcaaaagc	gcacgtctgc	cgcgctgttc	9480
tcctcttcc	catctccggg	cctttcgacc	agcttaccat	gaccgagtac	aagcccacgg	9540
tgcgctcgc	caccgcgac	gacgtcccca	gggcegtacg	caccctcgcc	gccgcgttcg	9600
ccgactacc	cgccacgcgc	cacaccgtcg	atccggaccg	ccacatcgag	egggtcaccg	9660
agctgcaaga	actcttcttc	acgcgcgtcg	ggctcgacat	cggcaagggtg	tgggtcgcgg	9720
acgacggcgc	cgcggtggcg	gtctggacca	cgccggagag	cgtcgaagcg	ggggcgggtg	9780
tcgcccagat	cggcccgcgc	atggccgagt	tgagcggttc	ccggctggcc	gcgcagaaca	9840
gatggaaggc	ctcctggcgc	cgcaccggcc	caaggagccc	gcgtggttcc	tggccaccgt	9900
cgcgtctcgc	ccgaccacca	gggcaagggt	ctggcagcgc	ccgtcgtgct	cccggagtg	9960
gaggcggccg	agcgcgcgg	ggtgccggcc	ttcctggaga	cctccgcgcc	ccgcaacctc	10020
cccttctacg	agcggctcgg	cttcaccgtc	accgccgacg	tcgaggtgcc	cgaaggaccg	10080
cgcacctggt	gcatgaccgc	caagcccgg	gcctgacgcc	cgccccacga	cccgcagcgc	10140
ccgaccgaaa	ggagcgcacg	accccatgca	taggttgggc	ttcggaatcg	ttttccggga	10200
cgccggctgg	atgatcctcc	agcgcgggga	tctcatgctg	gagttcttcg	cccaccccaa	10260
cttgtttatt	gcagcttata	atggttacia	ataaagcaat	agcatcaaaa	atttcacaaa	10320
taaagcattt	ttttactg	attctagtgt	tggtttgtcc	aaactcatca	atgtatctta	10380
tcatgtctgt	ataccgtcga	gatctagagc	ggccgccacc	gcggtggagc	tccagctttt	10440
gttcccttta	gtgagggtta	atctcgagct	tggcgtaatc	atggtcatag	ctgtttcctg	10500
tgtgaaattg	ttatccgctc	acaattccac	acaacatacg	agccggaagc	ataaagtgtg	10560
aagcctgggg	tgccaatga	gtgagctaag	tcacattaat	tgctgtgcgc	tactgccccg	10620
ctttccagtc	gggaaacctg	tcgtgccagg	gggtacctag	gccgggcaac	aattggcggc	10680
cggccgcact	tttcggggaa	atgtgcgcgg	aaccctatt	tgtttatttt	tctaaataca	10740
ttcaaatatg	tatccgctca	tgagacaata	accctgataa	atgcttcaat	aatattgaaa	10800
aaggaagagt	atgagtattc	aacatttccg	tgctgccctt	attccctttt	ttgcggcatt	10860
ttgccttcc	gtttttgctc	acccagaaac	gctggtgaaa	gtaaaagatg	ctgaagatca	10920
gttgggtgca	cgagtgggtt	acatcgaact	ggatctcaac	agcggtaaga	tccttgagag	10980
ttttcgcccc	gaagaacggt	ttccaatgat	gagcactttt	aaagttctgc	tatgtggcgc	11040
ggtattatcc	cgtattgacg	ccgggcaaga	gcaactcggg	cgccgcatac	actattctca	11100
gaatgacttg	gttgagtact	caccagtcac	agaaaagcat	cttacggatg	gcatgacagt	11160
aagagaatta	tgcagtgctg	ccataaccat	gagtgataac	actgcgcca	acttacttct	11220
gacaacgac	ggaggaccga	aggagctaac	cgcttttttg	cacaacatgg	gggatcatgt	11280
aactcgctt	gatcgttggg	aaccggagct	gaatgaagcc	ataccaaacg	acgagcgtga	11340
caccacgatg	cctgtagcaa	tggcaacaac	gttgcgcaaa	ctattaactg	gcgaactact	11400
tactctagct	tcccggcaac	aattaataga	ctggatggag	gcggataaag	ttgcaggacc	11460
acttctgcgc	tcggcccttc	cggctggctg	gtttattgct	gataaatctg	gagccggtga	11520
gcgtgggtct	cgcggtatca	ttgcagcact	ggggccagat	ggtaagccct	cccgtatcgt	11580
agttatctac	acgacgggga	gtcaggcaac	tatggatgaa	cgaaatagac	agatcgctga	11640
gatagggtgcc	tactgatta	agcattggta	actgtcagac	cctaggccgg	gcaacaattg	11700
gcggccggcc	ctgcattaat	gaatcgcca	acgcgcgggg	agaggcgggt	tgcgtattgg	11760
gcgctcttcc	gcttctcgc	tactgactc	gctgcgctcg	gtcgttcggc	tgcggcgagc	11820
ggtatcagct	cactcaaagg	cggtaatagc	gttatccaca	gaatcagggg	ataacgcagg	11880
aaagaacatg	tgagcaaaaag	gccagcaaaa	ggccaggaac	cgtaaaaagg	ccgcggttgc	11940
ggcgtttttc	cataggctcc	gccccctga	cgagcatcac	aaaaatcgac	gctcaagtca	12000
gaggtggcga	aaccgcacag	gactataaag	ataccaggcg	tttccccctg	gaagctccct	12060
cgtgcgctct	cctgttccga	ccctgccgct	taccggatac	ctgtccgctt	ttctcccttc	12120
gggaagcgtg	gcgctttctc	atagctcaag	ctgtaggat	ctcagttcgg	tgtaggctct	12180
tcgctccaag	ctgggctgtg	tgcacgaacc	ccccgttcag	cccgaccgct	gcgccttate	12240
cggtaaactat	cgtcttgagt	ccaaccgggt	aagacacgac	ttatcgccac	tggcagcagc	12300
cactggtaac	aggattagca	gagcgaggta	tgtaggcggg	gctacagagt	tcttgaagtg	12360
gtggcctaac	tacggctaca	ctagaaggac	agtatttggg	atctgcgctc	tgcgtgaagc	12420
agttaccttc	ggaaaaagag	ttggtagctc	ttgatccggc	aaacaaacca	ccgctggtag	12480
cgggtggtttt	tttgtttgca	agcagcagat	tacgcgcaga	aaaaaaggat	ctcaagaaga	12540
tcctttgatc	ttttctacgg	ggtctgacgc	tcagtggaac	gaaaactc		12588

<210> 28
 <211> 11998
 <212> DNA
 <213> 载体

<400> 28

```

acgttgtaaa acgacggcca gtgaattgta atacgactca ctatagggcg aattgggtac 60
ggggccccc ctcgaggtcg agttgggggtg gggaaaagga agaaacgcgg gcgtattggc 120
cccaatgggg tctcggtggg gtatcgacag agtgccagcc ctgggaccga accccgcgtt 180
tatgaacaaa cgacccaaca cccgtgcggt ttattctgtc tttttattgc cgtcatagcg 240
ggggttcctt cgggtattgt ctcttccgt cgacggtatc aagggtggcg cgggaatggg 300
gagctgcgag aatagccggg cgcgctgtga gccgaagtgc cccccgccct ggccaactcc 360
ggcgcgccga gtccttaggc cgccaggggg cgccggcgcg cgcccagatt ggggacaaaag 420
gaagccgggg cggccgcggt attaccataa aaggcaaaaa ctggtcggag gcgtccccgc 480
ggcgcgccgg aggaagccag gcccacaacc cctcccaacc gggcgccagc cccgcctccg 540
cccggttcaa acagcgaccg ggtcgcgcgc gcgcacgcag cggccacacc ctcgggcgcc 600
agcggctcgg gcaggaagtg gcgcaagcgc cggggcccca gaacgcacgc gcgattagcg 660
ccattgagtc ccagcgcgca cgcgcaatta gcgccaatc ccagcgcgca cgcagttagc 720
gccc aaagga ccagcgcgca cgcgcatggc gccccagccc ccaccggggc tgacggggggc 780
tccgcccgcg ccaccgtgcg atccccattg gcaagagccc ggctcagaca aagaccccgc 840
cggttgcccc cgccccgaga gcggcacccc cggagcgcgc cgcgccgagc gcggcctcgc 900
gcctgcgaac tggcgtgggg tgtcccccat ctccggaggc ccaggggctt ctcccgcgcc 960
ccccacggcg gtccggttcc gcccacatgc cccccgcgtg cggcccagac ggcggtctctg 1020
cacgggcaaa gggcgcgggc cgcacgcccc ggtcggctgg ccgggcttac ctggcgggcg 1080
gtgtggacgg gcggcgatc ggcaaaggcg aggtctctgt ctgcggggcg gacgcggtct 1140
cggcggtggg ggcgcgtcgc gccgctgggt tttatagggc gccgcccggc ccgctcgagc 1200
cataaaaggc aactttcgga acggcgcacg ctgattggcc ccgcgccgct cactaccg 1260
cttcgccgca cagtgcagca ttttttacc ccctctcccc tccttttgcg aaaaaaaaaa 1320
agagcgagag cgagattgag gaagaggagg agggagagtt ttggcggtgg ccgccttggg 1380
gtgctggggc cgggggctgg gggcgcgcgc cgtggcccc gcgccccacg ctgggcaagt 1440
cccggttcgg ccccgcatgg ccaggcctgc ccccgccctg cccgtctctc gggcccccca 1500
cccaccgchg gacatcctag gtgtggacat ctcttgggca ctgagcggcc aggtgggggtg 1560
ggccagggtc tgcacgggtg ccagggccct gggttctgta cgctcctgca gaaggagctc 1620
ttggagggca tggagtggcc aggcagtcac tcccccttgc cgacttcaga gcaactgcc 1680
tgaaagcagg gcctgaggac ctctggctgt ggggctcagc tagctaaatg tgctgggtgg 1740
gtcactaggg agagacctgg gcttgagagg tagagtgtgg tgttggggga gtcagggtggc 1800
ttgcggccat tagagtcgca ggaccacact cccagggaca gggcaggggc cagcgggtcca 1860
gtggctggag gtggcccgtg atgaaggcta caaacctacc cagccgcagc cctgggaagg 1920
aagtgggctc tacagggcag ggcacctttt accctggagc tgcttcttt tgagggtaac 1980
agtcacgccc agccaagacc aggcctgggg cgttagtggg tgacctaggc actgcggggc 2040
gggggggctg ggtctacaca gcctgggtct gggcccaccg tccgttgtat gtctgctatg 2100
cgcagccaca gctgaactgc cctcccagac catctggagg ccgctggggg actctgggga 2160
ccaagactcc atgtgccaca gaggattggg ggcggggcgg tgctaggaac tcaaagccag 2220
cctgggaaga ccctgtcctt gtcacccttt cttgcccttg gtctgtccac tgagtagcac 2280
acaagaccgg gtgggcaggg tccgttctgc tccgggaatc acagactgtg tgtaccagc 2340
tgggtggcat gcagcgatca gtggcgtagg accacagagg gggcccgcgg tacctaaaac 2400
agcttcacat ggcttaaaat aggggaccaa tgtcttttcc aatctaagtc ccattataa 2460
taaagtccat gttccatttt taaaggacaa tcctttcggt ttaaaccag gcacgattac 2520
ccaaacaact cacaacggta aagcactgtg aatcttctct gttctgcaat cccaacttgg 2580
tttctgctca gaaacctcc ctctttccaa tcggtaatta aataacaaa ggaaaaaact 2640
taagatgctt caaccccgtt tcgtgacact ttgaaaaaag aatcacctct tgcaaacacc 2700
cgctcccga ccccgccgct gaagcccggc gtccagaggc ctaagcgcgg gtgcccggc 2760
ccaccgggga gcgcgggcct cgtggtcagc gcacccgcgg ggagaaacaa aggcgcggc 2820
acgggggctc aagggcactg cgccacaccg cacgcgccta cccccgcgcg gccacgttaa 2880
ctggcggtcg ccgcagcctc gggacagccg gccgcgcgcc gccaggctcg cggacgcggg 2940
accacgcgac gccctccggg aggcccaggt ctgcaccag ccccgcggtg cgctggggga 3000
gggggcgcct ccgcccgaac gcgggtgggg gaggggaggg ggaaatgcgc tttgtctcga 3060

```

aatggggcaa	ccgtcgccac	agctccctac	cccctcgagg	gcagagcagt	ccccccacta	3120
actaccgggc	tggccgcgcg	ccaggccagc	cgcgaggcca	ccgcccgacc	ctccactcct	3180
tcccgcagct	cccggcgcgg	ggtccggcga	gaaggggagg	ggaggggagc	ggagaaccgg	3240
gcccccgga	cgcggtgtggc	atctgaagca	ccaccagcga	gcgagagcta	gagagaagga	3300
aagccaccga	cttcaccgcc	tccgagctgc	tccgggtcgc	gggtctgcag	cgctccggc	3360
cctccgcgcc	tacagctcaa	gccacatccg	aagggggagg	gagccgggag	ctgcgcgcgg	3420
ggccgcgggg	gggaggggtg	gcaccgcccc	cgccggggcg	ccacgaaggg	cggggcagcg	3480
ggcgcgcgcg	cggcgggggg	aggggcccgc	gcccgcgccg	ctgggaattg	gggcccctagg	3540
gggagggggg	aggcgccgac	gaccgcggca	cttaccgttc	gcggcgtggc	gcccgggtgg	3600
ccccaaaggg	agggaaaggg	gaggcggggc	gaggacagtg	accggagtct	cctcagcggt	3660
ggcttttctg	cttggcagcc	tcagcggctg	gcgccaaaac	cggactccgc	ccacttcctc	3720
gcccgcgggt	gcgaggggtg	ggaatcctcc	agacgctggg	ggagggggag	ttgggagctt	3780
aaaaactagt	acccttttgg	gaccactttc	agcagcgaac	tctcctgtac	accaggggtc	3840
agttccacag	acgcgggcca	gggggtgggtc	attgcggcgt	gaacaataat	ttgactagaa	3900
gttgattcgg	gtgtttccgg	aaggggcccga	gtcaatccgc	cgagttgggg	cacggaaaac	3960
aaaaagggaa	ggctactaag	atttttctgg	cgggggttat	cattggcgta	actgcagggg	4020
ccacctcccg	ggttgagggg	gctggatctc	caggctgcgg	attaagcccc	tcccgtcggc	4080
gttaatttca	aactgcgcga	cgtttctcac	ctgccttcgc	caaggcaggg	gccgggacct	4140
tattccaaga	ggtagtaact	agcaggactc	tagccttcgg	caattcattg	agcgcattta	4200
cggaagtaac	gtcgggtact	gtctctggcc	gcaaggggtg	gaggagtacg	catttggcgt	4260
aaggtggggc	gtagagcctt	cccgccattg	gcggcggata	gggcgtttac	gcgacggcct	4320
gacgtagcgg	aagacgcggt	agtggggggg	aaggttctag	aaaagcggcg	gcagcggctc	4380
tagcggcagc	agcagcagcg	ccgggtcccg	tgccgaggtg	ctcctcgcag	agttgtttct	4440
cgagcagcgg	cagttctcac	tacagcgcca	ggacgagtcc	ggttcgtgtt	cgcccgcgga	4500
gatctctctc	atctcgctcg	gctgcgggaa	atcgggctga	agcgaactgag	tcccgatgg	4560
aggtaacggg	tttgaatca	atgagttatt	gaaaagggca	tggcgaggcc	gttggcgcct	4620
cagtggaaag	cggccagccg	cctccgtggg	agagaggcag	gaaatcggac	caattcagta	4680
gcagtggggc	ttaaggttta	tgaacggggg	cttgagcggg	ggcctgagcg	tacaaacagc	4740
ttccccacce	tcagcctccc	ggcgccattt	cccttactcg	gggggtggggg	atggggagct	4800
ttcacatggc	ggacgctgcc	ccgctggggg	gaaagtgggg	cgccggaggcg	ggaattctta	4860
ttccctttct	aaagcacgct	gcttcggggg	ccacggcgtc	tctcggcga	gcgtttcggc	4920
gggcagcagg	tcctcgtgag	cgaggctgcg	gagcttccc	tccccctctc	tcccgggaac	4980
cgatttggcg	gccgccattt	tcatggctcg	ccttcctctc	agcgttttcc	ttataactct	5040
tttattttct	tagtgtgctt	tctctatcaa	gaagtagaag	tggttaacta	tttttttttt	5100
cttctcgggc	tgttttcata	tcgtttcgag	gtggatttgg	agtgttttgt	gagcttggat	5160
ctttagagtc	ctgcgcacct	cattaaaggc	gctcagcctt	cccctcgatg	aaatggcgcc	5220
attgcgttcg	gaagccacac	cgaagagcgg	ggaggggggg	tgctccgggt	ttgctggggc	5280
ggtttcagag	aagatcacac	caccagggcg	gtcggggccg	gttcaatgcg	agccgtagga	5340
caaagaaacc	atttatgtt	tttctgtct	ttttttctct	ttgagtaacg	gttttatctg	5400
ggtctgcagt	cagtaaaacg	acagatgaac	cgccggcaaaa	taaacataaa	ttggaagcca	5460
tcggccacga	ggggcagggg	cgaaggtggt	tttctgggcg	ggggagggat	attcgcgtca	5520
gaatccttta	ctgttcttaa	ggattccggt	taagttgtag	agctgactca	ttttaagtaa	5580
tgttgttact	gagaagttaa	acccttacgg	gacagatcca	tggaccttta	tagatgatta	5640
cgaggaaagt	gaaataacga	ttttgtcctt	agttatactt	cgattaaaac	atggcttcag	5700
aggctccttc	ctgtaatgcg	tatggattga	tgtgcaaaa	tgttttgggc	ctggggccgct	5760
ctgtatttga	actttgttac	ttttctcatt	ttgtttgcaa	tcttgggtga	acattacatt	5820
gataagcata	aggtctcaag	cgaagggggg	ctacctgggt	atttttcttt	gaccctaagc	5880
acgtttataa	aataacattg	tttaaaatcg	atagtgagca	tcgggtaagt	ttggataaat	5940
tgtgaggtaa	gtaatgagtt	tttgcttttt	gttagtgatt	tgtaaaactt	gttataaatg	6000
tacattatcc	gtaatttcag	tttagagata	acctatgtgc	tgacgacaat	taagaataaa	6060
aactagctga	aaaaatgaaa	ataactatcg	tgacaagtaa	ccatttcaaa	agactgcttt	6120
gtgtctcata	ggagctagtt	tgatcatttc	aggttaatttt	ttctttaatt	tttacgagtc	6180
atgaaaacta	caggaaaaaaa	aatctgaact	gggtttttacc	actacttttt	aggagtgggg	6240
agcatgcgaa	tggaggggaga	gctccgtaga	actgggatga	gagcagcaat	taatgctgct	6300
tgctaggaac	aaaaaataat	tgattgaaaa	ttacgtgtga	ctttttagtt	tgctattatgc	6360
gtttgtagca	gttggtcctg	gatatactt	tctctcgttt	gagggttttt	aacctagtta	6420
acttttaaga	caggtttctt	taacattcat	aagtgcccg	aatacagctg	tgtagtacag	6480

catataaaga	tttcagctct	gaggtttttc	ctattgactt	ggaaaattgt	tttgtgcctg	6540
tcgcttgcca	catggccaat	caagtaagct	tcgaattcga	gctcgcccaa	ctccgcccgt	6600
tttatgacta	gaaccaatag	tttttaatgc	caaatgcact	gaaatcccct	aatttgcaaa	6660
gccaaacgcc	ccctatgtga	gtaatacggg	gactttttac	ccaatttccc	aagcggaaa	6720
ccccctaata	cactcatatg	gcatatgaat	cagcacggtc	atgcactcta	atggcggccc	6780
atagggactt	tccacatagg	gggcgttcac	catttcccag	cataggggtg	gtgactcaat	6840
ggcctttacc	caagtaacatt	gggtcaatgg	gaggtaagcc	aatgggtttt	tcccattact	6900
ggcaagcaca	ctgagtcaaa	tgggactttc	cactgggttt	tgcccaagta	cattgggtca	6960
atgggagggtg	agccaatggg	aaaaacccat	tgctgccaag	tacactgact	caatagggac	7020
tttccaatgg	gtttttccat	tgttggcaag	catataaggt	caatgtgggt	gagtcaatag	7080
ggactttcca	ttgtattctg	cccagtacat	aaggtaata	gggggtgaat	caacaggaaa	7140
gtcccattgg	agccaagtac	actgcgtcaa	tagggacttt	ccattgggtt	ttgccagta	7200
cataaggtca	ataggggatg	agtcaatggg	aaaaacccat	tggagccaag	tacactgact	7260
caatagggac	tttccattgg	gttttgccca	gtacataagg	tcaatagggg	gtgagtcaac	7320
aggaaagtcc	cattggagcc	aagtacattg	agtcaatagg	gactttccaa	tgggttttgc	7380
ccagtacata	aggtcaatgg	gaggtaaagc	aatgggtttt	tcccattact	ggcacgtata	7440
ctgagtcatt	agggactttc	caatgggttt	tgcccagtac	ataaggtcaa	taggggtgaa	7500
tcaacaggaa	agtcccattg	gagccaagta	cactgagtca	atagggactt	tccattgggt	7560
tttgcccagt	acaaaaggtc	aataggggtt	gagtcaatgg	gtttttccca	ttattggcac	7620
gtacataagg	tcaatagggg	tgagtcattg	ggtttttcca	gccaattdaa	ttaaaacgcc	7680
atgtactttc	ccaccattga	cgtcaatggg	ctattgaaac	taatgcaacg	tgaccttdaa	7740
acggactttt	cccatagctg	attaatggga	aagtaccgtt	ctcgagccaa	tacacgtcaa	7800
tgggaagtga	aagggcagcc	aaaacgtaac	accgccccgg	ttttcccctg	gaaattccat	7860
attggcacgc	attctattgg	ctgagctgcg	ttctactggt	gtataagagg	cgcgaccagc	7920
gtcggtagcg	tcgcagctct	cggctctgac	accgtagaac	gcagagctcc	tcgctgcagc	7980
ccgggtctag	aggatccgcc	tgagaaagga	agtgagctgt	aaaggctgag	ctctctctct	8040
gacgtatgta	gcctctgggt	agcttcgtca	ctcactgttc	ttgactcagc	atggcaatct	8100
gatgaaatcc	cagctgtaag	tctgcagaaa	ttgatgatct	attaacaat	aaagatgtcc	8160
actaaaatgg	aagtttttcc	tgtcatactt	tgtaagaag	ggtgagaaca	gagtacctac	8220
atthtgaatg	gaaggattgg	agctacgggg	gtgggggtgg	ggtgggatta	gataaatgcc	8280
tgctctttac	tgaaggctct	ttactattgc	ttatgataa	tgtttcatag	ttggatatca	8340
taattttaa	aagcaaaacc	aaattaaggg	ccagctcatt	cctccagatc	cactagtctt	8400
agagcaaat	ctaccgggta	ggggagggcg	ttttcccaag	gcagctcggg	gcatgcgctt	8460
tagcagcccc	gctgggcact	tggcgctaca	caagtggcct	ctggcctcgc	acacattcca	8520
catccaccgg	taggcgccaa	ccggctccgt	tcttttggtg	ccccttcgcg	ccaccttcta	8580
ctctccccct	agtcaggaag	ttccccccg	ccccgcagct	cgcgtcgtgc	aggacgtgac	8640
aaatggaagt	agcacgtctc	actagctctc	tgacagatga	cagcaccgct	gagcaatgga	8700
agcgggtagg	cctttggggc	agcggccaat	agcagctttg	ctccttcgct	ttctgggctc	8760
agaggctggg	aaggggtggg	tccggggggc	ggctcagggg	cgggctcagg	ggcggggcgg	8820
gcgcccgaag	gtcctccgga	ggcccggcat	tctgcacgct	tcaaaagcgc	acgtctgccg	8880
cgctgttctc	ctcttctca	tctccgggcc	tttcgaccag	cttaccatga	ccgagtacaa	8940
gcccacgggtg	cgcttcgcca	cccgcgacga	cgtccccagg	gcogtacgca	ccctcgcgcg	9000
cgcttcgccc	gactaccccg	ccacgcgcca	caccgtcgat	ccggaccgcc	acatcgagcg	9060
ggtcaccgag	ctgcaagaac	tcttctctac	gcgctcggg	ctcgacatcg	gcaaggtgtg	9120
ggtcgcggac	gacggcgccg	cggtggcggg	ctggaccacg	ccggagagcg	tccaagcggg	9180
ggcgggtgtc	gccgagatcg	gcccgcgcat	ggccgagttg	agcgggtccc	ggctggccgc	9240
gcagcaacag	atggaaggcc	tcctggcgcc	gcaccggccc	aaggagcccg	cgtggttcct	9300
ggccaccgtc	ggcgtctcgc	ccgaccacca	gggcaagggt	ctgggcagcg	ccgtcgtgct	9360
ccccggagtg	gaggcggccg	agcgcgcccg	ggtgcccgcc	ttcctggaga	cctccgcgcc	9420
ccgcaacctc	cccttctacg	agcggctcgg	cttcaccgct	accgcccagc	tcgaggtgcc	9480
cgaaggaccg	cgacactggg	gcatgaccgc	caagcccggg	gcctgacgcc	cgccccacga	9540
cccgcagcgc	ccgaccgaaa	ggagcgcacg	accccatgca	taggttgggc	ttcggaatcg	9600
ttttccggga	cgccggctgg	atgatcctcc	agcgcgggga	tctcatgctg	gagttcttcg	9660
cccaccccc	cttgtttatt	gcagcttata	atggttacia	ataaagcaat	agcatcacia	9720
atthtcacaa	taaagcattt	ttttcactgc	attctagttg	tggtttgtcc	aaactcatca	9780
atgtatctta	tcatgtctgt	ataccgtcga	gatctagagc	ggccgcccac	gcgggtggagc	9840
tccagctttt	gttcccttta	gtgaggggta	atthcagact	tggcgtaatc	atggctcatag	9900

```

ctgtttcctg tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc 9960
ataaagtgta aagcctgggg tgcctaata gtagactaac tcacattaat tgcgttgccg 10020
tcaactgccc ctttccagtc gggaaacctg tgcgtgccag ggttacctag gccgggcaac 10080
aattggcggc cggccgcact tttcggggaa atgtgcccgg aaccctatt tgtttatatt 10140
tctaaataca ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat 10200
aatattgaaa aaggaagagt atgagtattc aacatttccg tgtcgcctt attccctttt 10260
ttgcggcatt ttgccttctt gtttttgctc acccagaaac gctggtgaaa gtaaaagatg 10320
ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga 10380
tccttgagag ttttcgcccc gaagaacggt ttccaatgat gagcactttt aaagtctctg 10440
tatgtggcgc ggtattatcc cgtattgacg cgggcaaga gcaactcggg cgccgcatac 10500
actattctca gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg 10560
gcatgacagt aagagaatta tgcagtgctg ccataacatc gagtgataac actgcggcca 10620
acttacttct gacaacgata ggaggaccga aggagctaac cgcttttttg cacaacatgg 10680
gggatcatgt aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg 10740
acgagcgtga caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg 10800
gcgaactact tactctagct tcccggcaac aattaataga ctggatggag gcggataaag 10860
ttgcaggacc acttctgcgc tcggcccttc cggctggctg gtttattgct gataaatctg 10920
gagccggatg gcgtgggtct cgcggtatca ttgcagcact ggggcccagat ggtaagccct 10980
cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac 11040
agatcgcctg gataggtgcc tcaactgata agcattggta actgtcagac cctaggccgg 11100
gcaacaattg gcggccggcc ctgcattaat gaatcggcca acgcccgggg agaggcgggt 11160
tgcgtattgg gcgctcttcc gcttctctgc tcaactgactc gctgcgctcg gtcgttcggc 11220
tgccgagcgc ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg 11280
ataacgcagg aaagaacatg tgagcaaaa gcccagaaa ggccaggaac cgtaaaaagg 11340
ccgcgttgct ggcgtttttc cataggctcc gccccctga cgagcatcac aaaaatcgac 11400
gctcaagtca gaggtggcga aaccgacag gactataaag ataccaggcg tttccccctg 11460
gaagctccct cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct 11520
ttctcccttc gggaaagcgtg gcgctttctc atagctcacg ctgtaggtat ctgagttcgg 11580
tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct 11640
gcgcttctc cggttaactat cgtcttgagt ccaaccgggt aagacacgac ttatcgccac 11700
tggcagcagc cactggtaac aggattagca gagcaggtg ttagggcggg gctacagagt 11760
tcttgaagtg gtggcctaac tacggctaca ctagaaggac agtatttggg atctgcgctc 11820
tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca 11880
ccgctggtag cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat 11940
ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactc 11998

```

<210> 29

<211> 12052

<212> DNA

<213> 载体

<220>

<221> misc_特征

<222> 8953

<223> n = A, T, C 或 G

<400> 29

```

acgttgtaaa acgacggcca gtgaattgta atacgactca ctatagggcg aattgggtac 60
cgggcccccc ctgaggtcgc agttgggggtg gggaaaagga agaaacgcgg gcgtattggc 120
cccaatgggg tctcggtggt gtatcgacag agtgccagcc ctgggaccga accccgcggt 180
tatgaacaaa cgaccacaac cccgtgcggt ttattctgtc tttttattgc cgtcatagcg 240
cgggttctt cgggtattgt ctcttccgt cgacggatc aagggtggcga ccggaatggt 300
gagctgcgag aatagccggg cgcgctgtga gccgaagtcg cccccgcct ggccacttcc 360
ggcgcgccga gtccttaggc cgccaggggg cgccggcgcg cgcccagatt ggggacaaag 420
gaagccgggc cggccgcggt attacataa aaggcaaaac ctggtcggag gcgtccccgc 480
ggcgcgcggc aggaagccag gccccaaacc cctccaacc gggcgcgcag cccgcctccg 540

```

ccccggttcaa	acagcgaccg	ggtcgcgcg	gcgcacgcag	cggccacacc	ctcggggcgcc	600
agcgggctcgg	gcaggaagtg	gcgcaagcgc	ccgggccccca	gaacgcacgc	gcgattagcg	660
ccattgagtc	ccagcgcgca	cgcgcaatta	gcgccaattc	ccagcgcgca	cgcagttagc	720
gccccaaagga	ccagcgcgca	cgcgcatggc	gccccagccc	ccaccggggc	tgacgggggc	780
tacgccgcgc	ccaccgtgcg	atccccattg	gcaagagccc	ggctcagaca	aagacccccgc	840
cggttgcccc	cgccccgaga	gcgccacccc	eggagcgcgc	ccgcccgagc	gcgccctcgc	900
gcttgcgaac	tggcgtgggg	tgtcccccat	ctccggaggc	ccaggggctt	ctccccgcgc	960
ccccacggcg	gtccggttcc	gccccatgcy	ccccccgctg	cggcccagac	ggcggctctg	1020
cacggggcgaa	gggcccgggc	cgcatgcccc	ggtcggctgg	ccgggcttac	ctggcggcgc	1080
gtgtggacgg	gcgcgggatc	ggcaaaggcg	aggctctgtg	ctcgcggggc	gacgcggtct	1140
cggcggtggt	ggcgcgtcgc	gccgctgggt	tttatagggc	gccgcccggg	ccgctcgagc	1200
cataaaaggc	aactttcggg	acggcgcacg	ctgattggcc	ccgcgccgct	cactcaccgg	1260
cttcgccgca	cagtgcagca	tttttttacc	ccctctcccc	tccttttgcy	aaaaaaaaaaa	1320
agagcgcagag	cgagattgag	gaagaggagg	agggagagtt	ttggcgttgg	ccgccttggg	1380
gtgctggggc	cgggggctgg	gggcccgcgc	cgtggccccc	gccccccacg	ctgggcagtg	1440
ccccggttcgg	ccccgcattg	ccaggcctgc	ccccggcctg	cccgtctctc	gggcccccca	1500
cccaccgcgg	gacatcctag	gtgtggacat	ctcttgggca	ctgagcgcgc	aggtgggggtg	1560
ggccagggctc	tgcacgggtg	ccagggccct	gggttctgta	cgctcctgca	gaaggagctc	1620
ttggagggca	tggagtggcc	aggcagtcac	tcccccttgc	cgacttcaga	gcaactgcc	1680
tgaagcagg	gcctgaggac	ctctggctgt	ggggctcagc	tagctaaatg	tgtcgggtgg	1740
gtcactaggg	agagacctgg	gcttgagagg	tagagtgtgg	tgttggggga	gtcaggtggc	1800
ttgcggccat	tagagtcgca	ggaccacact	ccccaggaca	gggcaggggc	cagcgggtcca	1860
gtggctggag	gtggcccgtg	atgaaggcta	caaacctacc	cagccgcagc	cctgggaagg	1920
aagtgggctc	tacagggcag	ggcacctttt	accctggagc	tgcttgcctt	tgagggtaac	1980
agtcacgccc	agccaagacc	aggcctgggg	cgttagtggg	tgacctaggc	actgcggggc	2040
gggggggctg	ggtctacaca	gcctgggtct	gggcccaccg	tccgttgat	gtctgctatg	2100
cgcagccaca	gctgaactgc	cctcccagac	catctggagg	ccgctggggg	actctggggg	2160
ccaagactcc	atgtgccaca	gaggattggg	ggcggggcgg	tgctaggaac	tcaaagccag	2220
cctgggaaga	ccctgtcctt	gtcacccttt	cttgcccttg	gtctgtccac	tgagtagcac	2280
acaagaccgg	gtgggcaggg	tccgttctgc	tccgggaatc	acagactgtg	tgtaccagg	2340
tgggtgggcat	gcagcgatca	gtggcgtggg	accacagagg	gggcccgcgg	taccaagctt	2400
gggaattgcy	tgcaaaaaa	acttctgttt	tccagggtaa	acagaatcta	atgcagaatc	2460
taatgcaggg	taaacagact	taatgcagaa	tctaattgatg	gcacaaaatta	aaaatcacta	2520
acgtgccctt	tttagtgtga	aaccagaga	gagcacatac	aagccaaaaa	caaatgcttt	2580
atttaccta	ggagacatta	acattcacct	ttacgtgttt	aagattaatg	caatgttaaa	2640
tattgtgaaa	actgtaactt	tgaatttcat	gatttttatg	tgaatattcc	aggggttaaa	2700
aaaacttgta	acatgacatg	gctgaataag	ataaaaaaaa	aatctagcct	ttctccctt	2760
ctggctcata	tttgcgattt	cgatcatttt	gtttaaaaaa	caaaacactg	caatgaatta	2820
aacttaatat	tcttctatgt	tttagagtaa	gttaaaaaca	gataaagtga	ccaaagtaat	2880
ttgaaagatt	caatgacttt	tgctccaacc	taggtgcaca	aggtaccttg	ttctttaa	2940
tgggctttaa	tgaaaatact	tctccagaat	tctggggatt	taagaaaaat	tatgccaac	3000
aacaagggct	ttaccatttt	atgtaacatt	tttcaacgct	gcaaaaatgt	gtgtatttct	3060
atgtgaagat	aaaaatcctc	agcaaaatcc	acattgcact	gtccttcaaa	gattagcctt	3120
ctttgaacta	gttaagacac	tattaagcca	agccagtatc	tccctgtaat	gaattcgttt	3180
ttctcttaat	tttcccctgt	aatttacact	gggagagctg	ggaaatatgt	ggatgtaaat	3240
ttctcagcca	cagagatgca	aagttatact	gtggggaaaa	aaaacttgag	ttaaatcctt	3300
acataattta	ggttttcatt	aacttaccaa	tgtagttttg	ttggaggcca	ttttttttat	3360
tgcagacttg	aagagctatt	actagaaaaa	tgcatgacag	ttaaggtaag	tttgcagac	3420
acaaaaagg	taactaaata	caaattctgt	ttggattcca	acccccaggt	agagagcgca	3480
cactttcaaa	cgtgaataca	aatccagagt	agatctagcgc	tccctacctac	attgcttatg	3540
atgtacttaa	gtacgtgtcc	taacctgtg	agtctagaaa	gactttactg	gggatcctgg	3600
tacctaaaac	agcttcacat	ggcttaaaat	aggggaccaa	tgtcttttcc	aatctaagtc	3660
ccatttataa	taaagtccat	gttccatttt	taaaggacaa	tcctttcggg	ttaaaaccag	3720
gcacgattac	ccaaacaact	cacaacggta	aagcaactgtg	aatcttctct	gttctgcaat	3780
cccaacttgg	tttctgctca	gaaaccctcc	ctctttccaa	tcggtaatta	aataacaaaa	3840
ggaaaaaact	taagatgctt	caaccccggt	togtgacact	ttgaaaaaag	aatcacctct	3900
tgcaaacacc	cgctcccgc	ccccgcogct	gaagcccggc	gtccagaggc	ctaagcgcgg	3960

gtgcccgccc	ccacccggga	gcgcgggcct	ogtggtcagc	gcatccgcgg	ggagaaacaa	4020
aggccgcggc	acgggggctc	aagggcactg	cgccaacccg	cacgcgccta	ccccgcgcg	4080
gccacgttaa	ctggcggtcg	ccgcagcctc	gggacagccg	gccgcgcgcc	gccaggctcg	4140
cggacgcggg	accacgcgcc	gccctccggg	aggcccaagt	ctcgaccag	ccccgcgtgg	4200
cgctggggga	gggggcgct	ccgcccgaac	gcggggtggg	gaggggaggg	ggaaatgctc	4260
tttgtctcga	aatggggcaa	ccgtcgccac	agctccctac	cccctcgagg	gcagagcagt	4320
cccccaacta	actaccgggc	tggccgcgcg	ccaggccagc	cgcgaggcca	ccgcccagcc	4380
ctccactcct	tcccgcagct	cccggcgccg	ggtccggcga	gaaggggagg	ggaggggagc	4440
ggagaaccgg	gccccggga	cgctgtggc	atctgaagca	ccaccagcga	gcgagagcta	4500
gagagaaggga	aagccaccga	cttcaccgcc	tccgagctgc	tccgggtcgc	gggtctgcag	4560
cgtctccggc	cctccgcgcc	tacagctcaa	gccacatccg	aagggggagg	gagccgggag	4620
ctgcgcgcgg	ggccgcgggg	gggaggggtg	gcaccgcca	cgccggggcg	ccacgaaggg	4680
cggggcagcg	ggcgcgcgcg	cggcgggggg	aggggcccgc	gccgcgcccg	ctgggaattg	4740
gggcccaggg	gggagggcgg	aggcgccgac	gaccgcggca	cttaccgttc	gcggcgtggc	4800
gcccgggtgt	ccccaaaggg	agggaaaggg	gagggcgggc	gaggacagtg	accggagtct	4860
cctcagcggg	ggcttttctg	cttggcagcc	tcagcggctg	gcgccaaaac	cggactccgc	4920
ccacttcctc	gcccgcgggt	gcgaggggtg	ggaatcctcc	agacgctggg	ggagggggag	4980
ttgggagctt	aaaaactagt	acccttttgg	gaccactttc	agcagcgaac	tctcctgtac	5040
accaggggtc	agttccacag	acgcgggcca	gggggtgggtc	attgcccgt	gaacaataat	5100
ttgactagaa	gttgattcgg	gtgtttccgg	aaggggcccga	gtcaatccgc	cgagttgggg	5160
cacggaaaac	aaaaaggga	ggctactaag	atttttctgg	cgggggttat	cattggcgta	5220
actgcaggga	ccacctcccg	ggttgagggg	gctggatctc	caggctgcgg	attaagcccc	5280
tcccgtcggc	gttaatttca	aactgcgcga	cgtttctcac	ctgccttcgc	caaggcaggg	5340
gccgggaccc	tattccaaga	ggtagtaact	agcaggactc	tagccttcgc	caattcattg	5400
agcgcattta	cggaagtaac	gtcgggtact	gtctctggcc	gcaaggggtg	gaggagtacg	5460
catttgccgt	aaggtggggc	gtagagcctt	cccgccattg	gcggcggata	gggcttttac	5520
gcgacggcct	gacgtagcgg	aagacgcggt	agtggggggg	aaggttctag	aaaagcggcg	5580
gcagcgggctc	tagcggcagt	agcagcagcg	ccgggtcccg	tgcggagggtg	ctcctcgcag	5640
agttgtttct	cgagcagcgg	cagttctcac	tacagcgcga	ggacgagtc	ggttcgtgtt	5700
cgtccgcgga	gatctctctc	atctcgcctc	gctgcgggaa	atcgggctga	agcagctgag	5760
tccgcgatgg	aggtaacggg	tttgaatca	atgagttatt	gaaaagggca	tggcgaggcc	5820
gttggcgcct	cagtggaagt	cggccagccg	cctccgtggg	agagaggcag	gaaatcggac	5880
caattcagta	gcagtggggc	ttaaggttta	tgaacggggg	cttgagcggg	ggcctgagcg	5940
tacaaacagc	ttccccaccc	tcagcctccc	ggcgccattt	cccttcaactg	gggggtggggg	6000
atggggagct	ttcacatggc	ggacgctgcc	ccgctggggg	gaaagtgggg	cgcgaggcgc	6060
ggaattctta	ttccctttct	aaagcacgct	gcttcggggg	ccacggcgctc	tcccctcgcg	6120
gcgtttcggc	gggcagcagg	tcctcgtgag	cgaggctgcg	gagcttcccc	tccccctctc	6180
tcccgggaac	cgatttgcg	cccgccattt	tcattggctcg	ccttcctctc	agcgttttcc	6240
ttataactct	tttattttct	tagtgtgctt	tctctatcaa	gaagtagaag	tggttaacta	6300
tttttttttt	cttctcgggc	tgttttcata	tcgtttcgag	gtggatttgg	agtgttttgt	6360
gagcttggat	cttttagagtc	ctgcgcacct	cattaaaggc	gctcagcctt	cccctcgatg	6420
aaatggcgcc	attgcgttcg	gaagccacac	cgaagagcgg	ggaggggggg	tgctccgggt	6480
ttgcggggcc	ggtttcagag	aagatcccaa	gcttcgaatt	cgagctcgcc	caactccgcc	6540
cgttttatga	ctagaaccaa	tagtttttaa	tgccaaatgc	actgaaatcc	cctaatttgc	6600
aaagccaaac	gccccctatg	tgagtaatac	ggggactttt	tacccaattt	cccaagcggg	6660
aagcccccta	atacactcat	atggcatatg	aatcagcacg	gtcatgcact	ctaattggcg	6720
cccatagggg	ctttccacat	agggggcggt	caccatttcc	cagcataggg	gtggtgactc	6780
aatggccttt	acccaagtac	attgggtcaa	tgggaggtaa	gccaatgggt	ttttccatt	6840
actggcaagc	acactgagtc	aaatgggact	ttccactggg	ttttgcccga	gtacattggg	6900
tcaatgggag	gtgagccaat	gggaaaaacc	cattgctgcc	aagtacactg	actcaatagg	6960
gactttccaa	tgggtttttc	cattgttggc	aagcatataa	ggtcaatgtg	ggtgagctca	7020
tagggacttt	ccattgtatt	ctgcccagta	cataaggtca	atagggggtg	aatcaacagg	7080
aaagtcccat	tggagccaag	tacactgcgt	caatagggac	tttccattgg	gttttgccca	7140
gtacataagg	tcaatagggg	atgagtcaat	gggaaaaacc	cattggagcc	aagtacactg	7200
actcaatagg	gactttccat	tgggttttgc	ccagtacata	aggtcaatag	gggggtgagtc	7260
aacaggaaag	tcccattgga	gccaagtaca	ttgagtcaat	agggactttc	caatgggttt	7320
tgcccagtac	ataaggtcaa	tgggaggtaa	gccaatgggt	ttttccatt	actggcacgt	7380

atactgagtc	attagggact	ttccaatggg	ttttgccag	tacataaggt	caataggggt	7440
gaatcaacag	gaaagtccca	ttggagccaa	gtacactgag	tcaatagggg	ctttccattg	7500
ggttttgccc	agtacaaaag	gtcaataggg	ggtgagtcaa	tgggtttttc	ccattattgg	7560
cacgtacata	aggtcaatag	gggtgagtca	ttgggttttt	ccagccaatt	taattaaaac	7620
gccatgtact	ttcccacat	tgacgtcaat	gggctattga	aactaatgca	acgtgacctt	7680
taaacggtac	tttcccatag	ctgattaatg	ggaaagtacc	gttctcgagc	caatacacgt	7740
caatgggaag	tgaaagggca	gccaaaacgt	aacaccgccc	cggttttccc	ctggaaattc	7800
catattggca	cgcattctat	tggctgagct	gcgttctacg	tgggtataag	aggcgcgacc	7860
agcgtcggta	ccgtcgcagt	cttcgggtctg	accaccgtag	aacgcagagc	tcctcgtctc	7920
agcccgggtc	tagaggatcc	gcctgagaaa	ggaagtgagc	tgtaaaggct	gagctctctc	7980
tctgacgtat	gtagcctctg	gtagccttcg	tcactcactg	ttcttgactc	agcatggcaa	8040
tctgatgaaa	tcccagctgt	aagtctgcag	aaattgatga	tctattaaac	aataaagatg	8100
tccactaaaa	tggaaagttt	tcctgtcata	ctttgttaag	aagggtgaga	acagagtacc	8160
tacattttga	atggaaggat	tggagctacg	gggggtgggg	tgggggtggg	ttagataaat	8220
gcctgctctt	tactgaaggc	tccttactat	tgctttatga	taatgtttca	tagttggata	8280
tcataattta	aacaagcaa	accaaattaa	gggccagctc	attcctccag	atccactagt	8340
aattctgtgg	aatgtgtgtc	agttaggggtg	tggaaagtcc	ccaggctccc	cagcaggcag	8400
aagtatgcaa	agcatgcatc	tcaattagtc	agcaaccagg	tgtggaaagt	cccaggctc	8460
cccagcaggc	agaagtatgc	aaagcatgca	tctcaattag	tcagcaacca	tagtcccgcc	8520
cctaactccg	cccatcccgc	ccctaactcc	gccagttcc	gccattctc	cgccccatgg	8580
ctgactaatt	ttttttat	atgcagaggg	cgaggccgcc	tctgcctctg	agctattcca	8640
gaagtagtga	ggaggctttt	ttggaggcct	aggcttttgc	aaaaagctcc	cgggagcttg	8700
tatatccatt	ttcggatctg	atcaagagac	aggatgagga	tcgtttcgca	tgattgaaca	8760
agatggattg	cacgcaggtt	ctccggccgc	ttgggtggag	aggctattcg	gctatgactg	8820
ggcacaacag	acaatcggct	gctctgatgc	gcctgtgttc	cggctgtcag	cgcaggggcg	8880
cccggttcct	tttgtcaaga	ccgacctgtc	cggtgcctg	aatgaaactgc	aggacgaggc	8940
agcgcggcta	tcstggctgg	ccacgacggg	cgttccttgc	gcagctgtgc	tcgacgttgt	9000
caactgaagcg	ggaagggact	ggctgctatt	gggcgaagtg	ccggggcagg	atctcctgtc	9060
atctcacctt	gctcctgccg	agaaagtatc	catcatggct	gatgcaatgc	ggcggctgca	9120
tacgcttgat	ccggctacct	gccattcga	ccaccaagcg	aaacatcgca	tcgagcgagc	9180
acgtactcgg	atggaagccg	gtcttgcga	tcaggatgat	ctggacgaag	agcatcaggg	9240
gctcgcgcca	gccgaactgt	tcgccaggct	caaggcgcgc	atgcccgacg	gcgaggatct	9300
cgctcgtgacc	catggcgatg	cctgcttggc	gaatatcatg	gtggaaaatg	gccgcttttc	9360
tggattcatc	gactgtggcc	ggctgggtgt	ggcggaccgc	tatcaggaca	tagcgttggc	9420
taccctgat	attgctgaag	agcttggcgg	cgaatgggct	gaccgcttcc	tcgtgcttta	9480
cggtatcgcc	gctcccgat	cgcagcgcac	cgcttctat	cgcttcttg	acgagttctt	9540
ctgagcggga	ctctgggggt	cgaaatgacc	gaccaagcga	cgcccaacct	gccatcagca	9600
gatttcgatt	ccaccgccgc	cttctatgaa	agtttgggct	tcggaatcgt	ttccggggac	9660
gccggctgga	tgatcctcca	gcgcggggat	ctcatgctgg	agttcttcgc	ccaccccaac	9720
ttgtttat	cagcttataa	tggttacaaa	taaagcaata	gcatcacaaa	tttcacaaat	9780
aaagcatttt	tttactgca	ttctagttgt	ggtttgtcca	aactcatcaa	tgtatcttat	9840
catgtctgta	taccgtcgag	actagttcta	gagcggccgc	caccgcggtg	gagctccagc	9900
ttttgttccc	tttagtgagg	gttaatttcg	agcttggcgt	aatcatggtc	atagctgttt	9960
cctgtgtgaa	attgttatcc	gctcacaatt	ccacacaaca	tacgagccgg	aagcataaag	10020
tgtaaagcct	ggggtgccta	atgagtgagc	taactcacat	taattgcggt	gcgctcactg	10080
cccgtttcc	agtccggaaa	cctgtcgtgc	caggggggtac	ctaggccggg	caacaattgg	10140
cggccggccg	cacttttcgg	ggaaatgtgc	gcggaaacccc	tatttgttta	tttttctaaa	10200
tacattcaaa	tatgtatccg	ctcatgagac	aataaccctg	ataaatgctt	caataatatt	10260
gaaaaaggaa	gagtatgagt	attcaacatt	tcctgtcgc	ccttattccc	ttttttgcgg	10320
cattttgcct	tcctgttttt	gctcaccacg	aaacgctggt	gaaagtaaaa	gatgctgaa	10380
atcagttggg	tgacagagtg	ggttacatcg	aactggatct	caacagcggg	aagatccttg	10440
agagttttcg	ccccgaagaa	cgttttccaa	tgatgagcac	ttttaaagtt	ctgctatgtg	10500
gcgcggtatt	atcccgtatt	gacgcggggc	aagagcaact	cggctcggccg	atacactatt	10560
ctcagaatga	cttgggttgag	tactcaccag	tcacagaaaa	gcatcttacg	gatggcatga	10620
cagtaagaga	attatgcagt	gctgccataa	ccatgagtga	taacactcgc	gccaaacttac	10680
ttctgacaac	gatcggagga	ccgaaggagc	taaccgcttt	tttgacaaac	atgggggatc	10740
atgtaactcg	ccttgatcgt	tgggaaccgg	agctgaatga	agccatacca	aacgacgagc	10800

gtgacaccac	gatgcctgta	gcaatggcaa	caacgttgcg	caaactatta	actggcgaac	10860
tacttactct	agcttcccgg	caacaattaa	tagactggat	ggagggcggat	aaagttgcag	10920
gaccacttct	gcgctcggcc	cttccggctg	gctggtttat	tgctgataaa	tctggagccg	10980
gtgagcgtgg	gtctcgcggg	atcattgcag	cactggggcc	agatggtaag	ccctcccgtg	11040
tcgtagttat	ctacacgacg	gggagtcagg	caactatgga	tgaacgaaat	agacagatcg	11100
ctgagatagg	tgccctcactg	attaagcatt	ggtaactgtc	agaccctagg	ccgggcaaca	11160
attggcggcc	ggccctgcat	taatgaatcg	gccaacgcgc	ggggagaggc	ggttttgcgta	11220
ttgggcgctc	ttccgcttcc	tcgctcactg	actcgtcgcg	ctcggtcggt	cggctgcggc	11280
gagcggatc	agctcactca	aaggcggtaa	tacggttatc	cacagaatca	ggggataacg	11340
caggaaagaa	catgtgagca	aaaggccagc	aaaaggccag	gaaccgtaaa	aaggccgcgt	11400
tgctggcggt	tttccatagg	ctccgcccc	ctgacgagca	tcacaaaaat	cgacgctcaa	11460
gtcagagggtg	gcgaaaccgg	acaggactat	aaagatacca	ggcgtttccc	cctggaagct	11520
ccctcgtgcg	ctctcctggt	ccgaccctgc	cgcttaccgg	atacctgtcc	gcctttctcc	11580
cttcgggaag	cgtggcgctt	tctcatagct	cacgctgtag	gtatctcagt	tcgggtgtagg	11640
tcgttcgctc	caagctgggc	tgtgtgcacg	aacccccgt	tcagcccgcg	cgctgcgcct	11700
tatccggtaa	ctatcgtctt	gagtccaacc	cggttaagaca	cgacttatcg	ccactggcag	11760
cagccactgg	taacaggatt	agcagagcga	ggtagttagg	cggtgtctaca	gagttcttga	11820
agtgggtggc	taactacggc	tacactagaa	ggacagtatt	tggtatctgc	gctctctgta	11880
agccagttac	cttcggaaaa	agagttggta	gctcttgatc	cggcaaaaa	accaccgctg	11940
gtagcgggtg	tttttttggt	tgcaagcagc	agattacgcg	cagaaaaaaa	ggatctcaag	12000
aagatccttt	gatcttttct	acgggggtctg	acgctcagtg	gaacgaaaac	tc	12052

<210> 30

<211> 11941

<212> DNA

<213> 载体

<400> 30

acgttgtaaa	acgacggcca	gtgaattgta	atacgactca	ctatagggcg	aattgggtac	60
cgggcccccc	ctcgaggctg	agttgggggtg	gggaaaagga	agaaacgcgg	gcgtattggc	120
cccaatgggg	tctcggtggg	gtatcgacag	agtgccagcc	ctgggaccga	accccgcgtt	180
tatgaacaaa	cgaccecaaca	cccgtgcggt	ttattctgtc	tttttattgc	cgatcatagcg	240
cgggttcctt	ccggtattgt	ctccttccgt	cgacggtatc	aaggtagcga	ccggataggt	300
gagctgccc	aatagccggg	cgcgctgtga	gccgaagtcg	ccccgcct	ggcacttcc	360
ggcgcgccga	gtccttaggc	cgccaggggg	cgccggcgcg	cgcccagatt	ggggacaaaag	420
gaagccgggc	cggccgcggt	attaccataa	aaggcaaaaa	ctggtcggag	gcgtccccgc	480
ggcgcgcggc	aggaagccag	gccccaaccc	cctccccacc	gggcgccagc	cccgcctccg	540
cccggttcaa	acagcgaccg	ggtcgcgcgc	gcgcacgcag	cgccacacc	ctcgggcgcc	600
agcggctcgg	gcaggaagtg	gcgcaagcgc	ccgggcccc	gaacgcacgc	gcgattagcg	660
ccattgagtc	ccagcgcgca	cgcgcaatta	gcgccaattc	ccagcgcgca	cgcagttagc	720
gccccaaagga	ccagcgcgca	cgcgcatggc	gccccagccc	ccaccgggcc	tgacgggggc	780
tacgccgcgc	ccaccgtgcg	atccccattg	gcaagagccc	ggctcagaca	aagacccccg	840
cggttgcccc	cgccccgaga	gcggcacccc	cggagcgcgc	cgccccgagc	gcggcctcgc	900
gcctgcgaac	tggcgtgggg	tgtcccccat	ctccggaggc	ccaggggctt	ctccccgcgc	960
ccccacggcg	gtccggttcc	gccccatgcg	ccccccgctg	cgcccagac	ggcggctctg	1020
cacggggcga	gggcccgggc	cgcatgcccc	ggtcggctgg	ccgggcttac	ctggcggcgg	1080
gtgtggacgg	gcgccggatc	ggcaaaggcg	aggctctgtg	ctcgcggggc	gacgcggtct	1140
ggcgggtggt	ggcgcgtcgc	gccgctgggt	tttatagggc	gccgccgcgg	ccgctcagac	1200
cataaaaaggc	aactttcggg	acggcgcacg	ctgattggcc	ccgcgccgct	cactcaccgg	1260
cttcgcgcga	cagtgacgca	tttttttacc	ccctctcccc	tccttttgcg	aaaaaaaaaa	1320
agagcgagag	cgagattgag	gaagaggagg	agggagagtt	ttggcgttgg	ccgccttggg	1380
gtgctgggccc	cgggggctgg	gggcgcgcgc	cgtggccccc	gcgccccacg	ctgggcagtg	1440
cccggttcgg	ccccgcattg	ccaggcctgc	ccccggcctg	cccgtctctc	gggcccccca	1500
cccaccgcgg	gacatcctag	gtgtggacat	ctcttgggca	ctgagcgcgc	aggtgggggtg	1560
ggccagggctc	tgacgggtg	ccagggccct	gggttctgta	cgctcctgca	gaaggagctc	1620
ttggagggca	tggagtggcc	aggcagtcac	tcccccttgc	cgacttcaga	gcaactgccc	1680
tgaagcagg	gcctgaggac	ctctggctgt	ggggctcagc	tagctaaatg	tgctgggtgg	1740

gtcactaggg	agagacctgg	gcttgagagg	tagagtgtgg	tgttggggga	gtcaggtggc	1800
ttgcgccat	tagagtcgca	ggaccacact	cccaggaca	gggcaggggc	cagcggcca	1860
gtggctggag	gtggcccgtg	atgaaggcta	caaacctacc	cagccgcagc	cctgggaagg	1920
aagtgggctc	tacagggcag	ggcacctttt	accctggagc	tgccctgctt	tgagggtaac	1980
agtcacgccc	agccaagacc	aggcctgggg	cgtagtgagg	tgacctaggc	actgcggggc	2040
gggggggctg	ggtctacaca	gcctgggtct	gggccaccg	tccgttgat	gtctgctatg	2100
cgcagccaca	gctgaactgc	cctcccagac	catctggagg	ccgctggggg	actctggggg	2160
ccaagactcc	atgtgccaca	gaggattggg	ggcggggcgg	tgctaggaac	tcaaagccag	2220
cctgggaaga	ccctgtcctt	gtcacccttt	cttgcccttg	gtctgtccac	tgagtagcac	2280
acaagaccgg	gtgggcaggg	tccgttctgc	tccgggaatc	acagactgtg	tgtaccagg	2340
tggtagggcat	gcagcgatca	gtggcgtggg	accacagagg	gggccgcgg	taccaagctt	2400
gggaattgcg	tgcaaaaaca	acttctgttt	tccagggtaa	acagaatcta	atgcagaatc	2460
taatgcaggg	taaacagact	taatgcagaa	tctaattgatg	gcacaaatta	aaaatcacta	2520
acgtgccctt	tttagtgtga	aaccagaga	gagcacatac	aagccaaaaa	caaagtcttt	2580
atthttaccta	ggagacatta	acattcacct	ttacgtgttt	aagattaatg	caatgttaaa	2640
tattgtgaaa	actgtaactt	tgaatttcat	gatttttatg	tgaatattcc	agggtttaa	2700
aaaacttgta	acatgacatg	gctgaataag	ataaaaaaaaa	aatctagcct	tttctccctt	2760
ctggctcata	tttgcgattt	cgatcatttt	gtttaaaaaa	caaaactctg	caatgaatta	2820
aacttaatat	tcttctatgt	tttagagtaa	gttaaaaca	gataaagtga	ccaaagtaat	2880
ttgaaagatt	caatgacttt	tgctccaacc	taggtgcaca	aggtaccttg	ttctttaa	2940
tgggctttaa	tgaaaatact	tctccagaat	tctggggatt	taagaaaaat	tatgccaacc	3000
aacaagggat	ttaccatttt	atgtaacatt	tttcaacgct	gcaaaaatgt	gtgtatttct	3060
atttgaagat	aaaaatcctc	agcaaaatcc	acattgact	gtccttcaa	gattgcctt	3120
ctttgaacta	gttaagacac	tattaagcca	agccagtatc	tccctgta	gaattcgctt	3180
ttctctta	tttcccctgt	aatttacact	gggagagctg	ggaaatatgt	ggatgtaa	3240
ttctcagcca	cagagatgca	aagttatact	gtggggaaaa	aaaacttgag	ttaa	3300
acata	ggtttctatt	aacttacaa	tgtagt	ttggaggcca	tttttt	3360
tgca	aagagctatt	actagaaaa	tgcatgacag	ttaaggtaag	tttgc	3420
acaaaaa	taactaaata	caaattctgt	ttggattcca	acccccagt	agagagc	3480
cactttcaa	cgtgaataca	aatccagagt	agatctgcgc	tcctacctac	attgcttatg	3540
atgtacttaa	gtacgtgtcc	taaccatgtg	agtctagaaa	gactttactg	gggatcctgg	3600
tacctaaac	agcttcacat	ggcttaaaat	aggggaccaa	tgtcttttcc	aatctaagtc	3660
ccatttata	taaagtccat	gttccatttt	taaaggacaa	tcctttcggt	ttaa	3720
gcacgattac	ccaaacaact	cacaacggta	aagcactgtg	aatcttctct	gttctgca	3780
cccaactgg	tttctgtca	gaaaccctcc	ctctttccaa	tcggttaatta	aataca	3840
ggaaaaaact	taagatgctt	caaccccggt	tcgtgacact	ttgaaaaaag	aatcacctct	3900
tgcaaacacc	cgctcccagc	cccgcgct	gaagcccggc	gtccagaggc	ctaagcgcgg	3960
gtgcccgcc	ccaccggga	gcgcggcct	cgtagtcagc	gcatccgcgg	ggagaa	4020
aggccgcggc	acgggggctc	aagggcactg	cgccacaccg	cacgcgccta	ccccgcgcg	4080
gccacgttaa	ctggcggtcg	ccgcagcctc	gggacagccg	gccgcgcgcc	gccaggctcg	4140
cggacgcggg	accacgcgcc	gccctccggg	aggcccaagt	ctcgaccag	ccccgcgtgg	4200
cgctggggga	gggggcgcct	ccgcgggaac	gcgggtgggg	gaggggaggg	ggaaatgcgc	4260
tttgtctcga	aatggggcaa	ccgtcgccac	agctccctac	cccctcgagg	gcagagcagt	4320
cccccaacta	actaccgggc	tggccgcgcg	ccaggccagc	cgcgaggcca	ccgcccagacc	4380
ctccactcct	tcccgcagct	cccggcgcgg	ggtccggcga	gaaggggagg	ggaggggagc	4440
ggagaaaccg	gccccggga	cgctgtggc	atctgaagca	ccaccagcga	gcgagagcta	4500
gagagaagga	aagccaccga	cttaccgcc	tccgagctgc	tccgggtcgc	gggtctgcag	4560
cgtctccggc	cctccgcgcc	tacagctcaa	gccacatccg	aagggggagg	gagccgggag	4620
ctgcgcgcgg	ggccgcggg	gggaggggtg	gcaccgccca	cgccgggcgg	ccacgaagg	4680
cgggcagcg	ggcgcgcgcg	cggcggggg	agggccggc	gccgcgcgcc	ctgggaattg	4740
ggccctagg	gggagggcgg	aggcggcagc	gaccgcggca	cttaccgttc	gcggcgtggc	4800
gcccggtggt	ccccaaaggg	agggaaaggg	gaggcggggc	gaggacagtg	accggagtct	4860
cctcagcgg	ggcttttctg	cttggcagcc	tcagcggctg	gcgccaaaac	cggactccgc	4920
ccacttctc	gcccgcgggt	gcgaggggtg	ggaatcctcc	agacgctggg	ggagggggag	4980
ttgggagctt	aaaaactagt	accctttgg	gaccactttc	agcagcgaac	tctcctgtac	5040
accaggggtc	agttccacag	acgcgggcca	gggggtgggtc	attgcggcgt	gaacaataat	5100
ttgactagaa	gttgattcgg	gtgtttccgg	aaggggccga	gtcaatccgc	cgagttgggg	5160

cacggaaaac	aaaaagggaa	ggctactaag	atTTTTctgg	cgggggttat	cattggcgta	5220
actgcagggg	ccacctcccc	ggttgagggg	gctggatctc	caggctgcgg	attaagcccc	5280
tcccgtcggc	gttaattttca	aactgcgcga	cgtttctcac	ctgccttcgc	caaggcaggg	5340
gccgggaccc	tattccaaga	ggtagtaact	agcaggactc	tagccttccg	caattcattg	5400
agcgcattta	cggaagtaac	gtcgggtact	gtctctggcc	gcaaggggtg	gaggagtacg	5460
catttggcgt	aaggtggggc	gtagagcctt	cccgccattg	gcggcggata	gggcgtttac	5520
gcgacggcct	gacgtagcgg	aagacgcgtt	agtggggggg	aaggttctag	aaaagcggcg	5580
gcagcggctc	tagcggcagt	agcagcagcg	ccgggtcccg	tgccggaggtg	ctcctcgcag	5640
agttgtttct	cgagcagcgg	cagttctcac	tacagcgcca	ggacgagtcc	ggttcgtgtt	5700
cgtccgcgga	gatctctctc	atctcgcctc	gctgcgggaa	atcgggctga	agcactgag	5760
tccgcgatgg	aggtaacggg	tttgaatca	atgagttatt	gaaaagggca	tggcgaggcc	5820
gttggcgcct	cagtggaagt	cgccagccg	cctccgtggg	agagaggcag	gaaatcggac	5880
caattcagta	gcagtggggc	ttaaggttta	tgaacggggg	cttgagcggg	ggcctgagcg	5940
tacaaacagc	ttccccaccc	tcagcctccc	ggcgccattt	cccttcaactg	gggggtggggg	6000
atggggagct	ttcacatggc	ggacgctgcc	ccgctggggg	gaaagtgggg	cgccggagggc	6060
ggaattctta	ttccctttct	aaagcacgct	gcttcggggg	ccacggcgtc	tcctcggcga	6120
gcgtttcggc	gggcagcagg	tcctcgtgag	cgaggctgcg	gagcttcccc	tccccctctc	6180
tcccgggaac	cgatttggcg	gccgccattt	tcattggctcg	ccttcctctc	agcgttttcc	6240
ttataactct	tttattttct	tagtgtgctt	tctctatcaa	gaagtagaag	tggttaacta	6300
tttttttttt	cttctcgggc	tgttttcata	tcgtttcgag	gtggatttgg	agtgttttgt	6360
gagcttggat	ctttagagtc	ctgcgcacct	cattaaaggc	gctcagcctt	cccctcgatg	6420
aaatggcgcc	attgcgttcg	gaagccacac	cgaagagcgg	ggaggggggg	tgctccggctg	6480
ttcggggccc	ggtttcagag	aagatcccaa	gcttcgaatt	cgagctcgcc	caactccgcc	6540
cgttttatga	ctagaaccaa	tagtttttaa	tgccaaatgc	actgaaatcc	cctaatttgc	6600
aaagccaaac	gccccctatg	tgagtaatac	ggggactttt	tacccaattt	cccaagcggg	6660
aagcccccta	atacactcat	atggcatatg	aatcagcacg	gtcatgcact	ctaattggcgg	6720
cccatagggg	ctttccacat	agggggcgtt	caccatttcc	cagcataggg	gtggtgactc	6780
aatggccttt	acccaagtac	attgggtcaa	tgggaggtaa	gccaatgggt	ttttccatt	6840
actggcaagc	acactgagtc	aaatgggact	ttccactggg	ttttgcccaa	gtacattggg	6900
tcaatgggag	gtgagccaat	gggaaaaacc	cattgctgcc	aagtacactg	actcaatagg	6960
gactttccaa	tgggtttttc	cattgttggc	aagcatataa	gggtcaatgtg	ggtgagtcaa	7020
tagggacttt	ccattgtatt	ctgcccagta	cataaggcca	ataggggggtg	aatcaacagg	7080
aaagtcccat	tggagccaag	tacactgcgt	caatagggac	tttccattgg	gttttgccca	7140
gtacataaag	tcaatagggg	atgagtcaat	gggaaaaacc	cattggagcc	aagtacactg	7200
actcaatagg	gactttccat	tgggttttgc	ccagtacata	aggtcaatag	gggggtgagtc	7260
aacaggaag	tcccatttga	gccaagtaca	ttgagtcaat	agggactttc	caatgggttt	7320
tgcccagtac	ataagggtcaa	tgggaggtaa	gccaatgggt	ttttccatt	actggcacgt	7380
atactgagtc	attagggact	ttccaatggg	ttttgcccag	tacataaggt	caataggggt	7440
gaatcaacag	gaaagtccca	ttggagccaa	gtacactgag	tcaatagggg	ctttccattg	7500
ggttttgccc	agtacaaaag	gtcaataggg	ggtgagtcaa	tgggtttttc	ccattattgg	7560
cacgtacata	aggtcaatag	gggtgagtca	ttgggttttt	ccagccaatt	taattaaac	7620
gccatgtact	ttcccacat	tgacgtcaat	gggctattga	aactaatgca	acgtgacctt	7680
taaacggtac	tttcccatag	ctgattaatg	ggaaagtacc	gttctcgagc	caatacacgt	7740
caatgggaag	tgaaagggca	gccaaaacgt	aacaccgccc	cggttttccc	ctggaaattc	7800
catattggca	cgcattctat	tggctgagct	gcgttctacg	tgggtataag	agggcgcgacc	7860
agcgtcggta	ccgtcgcagt	cttcggctcg	accaccgtag	aacgcagagc	tcctcgcctg	7920
agccccggtc	tagaggatcc	gcctgagaaa	ggaagtgagc	tgtaaaggct	gagctctctc	7980
tctgacgtat	gtagcctctg	gttagcttcg	tcaactcactg	ttcttgactc	agcatggcaa	8040
tctgatgaaa	tcccagctgt	aagtctgcag	aaattgatga	tctattaaac	aataaagatg	8100
tccactaaaa	tggaaagttt	tcctgtcata	ctttgttaag	aagggtgaga	acagagtacc	8160
tacattttga	atggaaggat	tggagctacg	gggggtggggg	tgggggtggga	ttagataaat	8220
gcctgctctt	tactgaaggc	tctttactat	tgctttatga	taatgtttca	tagttggata	8280
tcataattta	aacaagcaaa	accaaattaa	gggccagctc	attcctccag	atccactagt	8340
tctagagcaa	attctaccgg	gtaggggagg	cgcttttccc	aaggcagctc	ggagcatgcg	8400
ctttagcagc	cccgtggggc	acttggcgct	acacaagtgg	cctctggcct	cgcacacatt	8460
ccacatccac	cggtaggcgc	caaccggctc	cgttctttgg	tggccccctc	gcgcacacct	8520
ctactcctcc	cctagtcagg	aagttcccc	ccgccccgca	gctcgcgctc	tgcaggacct	8580

gacaaatgga	agtagcacgt	ctcactagtc	tcgtgcagat	ggacagcacc	gctgagcaat	8640
ggaagcgggt	aggccttttg	ggcagcggcc	aatagcagct	ttgctccttc	gctttctggg	8700
ctcagaggct	gggaaggggt	gggtccgggg	gcgggctcag	gggcgggctc	aggggcgggg	8760
cgggcgcccg	aaggtcctcc	ggaggcccgg	cattctgcac	gcttcaaaag	cgcacgtctg	8820
ccgcgctggt	ctcctcttcc	tcatctccgg	gcctttcgac	cagcttacca	tgaccgagta	8880
caagcccacg	gtgcgctcgc	ccaccgcgga	cgacgtcccc	agggccgtac	gcaccctcgc	8940
cgccgcgttc	gccgactacc	ccgccacgcg	ccacaccgtc	gatccggacc	gccacatcga	9000
gcgggtcacc	gagctgcaag	aactcttctc	cacgcgcgtc	gggctcgaca	tcggcaaggt	9060
gtgggtcgcg	gacgacggcg	ccgcgggtgg	ggtctggacc	acgccggaga	gcgtcgaagc	9120
ggggcggggtg	ttcgccgaga	tcggcccgcg	catggccgag	ttgagcgggt	cccggctggc	9180
cgcgacgcaa	cagatggaag	gcctcctggc	gccgcaccgg	cccaaggagc	ccgctggtt	9240
cctggccacc	gtcggcgtct	cgccccacca	ccagggcaag	ggtctgggca	gcgccgtcgt	9300
gctccccgga	gtggaggcgg	ccgagcgcgc	cggggtgccc	gccttctctg	agacctccgc	9360
gccccgcaac	ctccccttct	acgagcggct	cggcttcacc	gtcaccgccg	acgtcagaggt	9420
gcccgaagga	ccgcgcacct	ggtgcatgac	ccgcaagccc	ggtgcctgac	gccccccca	9480
cgacccgcag	cgcccgaccg	aaaggagcgc	acgacccccat	gcataggttg	ggcttcggaa	9540
tcgttttccg	ggacgccggc	tggatgatcc	tccagcgcgg	ggatctcatg	ctggagtctt	9600
tcgcccaccc	caacttgttt	attgcagctt	ataatggtta	caaataaagc	aatagcatca	9660
caaatttcac	aaataaagca	tttttttcac	tgcattctag	ttgtggtttg	tccaaactca	9720
tcaatgtatc	ttatcatgtc	tgtataccgt	cgagatctag	agcggccgcc	accgcggtgg	9780
agctccagct	tttgttccct	ttagtggagg	ttaatttcga	gcttggcgta	atcatggtca	9840
tagctgtttc	ctgtgtgaaa	ttgttatccg	ctcacaattc	cacacaacat	acgagccgga	9900
agcataaagt	gtaaagcctg	gggtgcctaa	tgagtgagct	aactcacatt	aattgcgttg	9960
cgctcactgc	ccgctttcca	gtcgggaaac	ctgtcgtgcc	agggggtacc	taggcggggc	10020
aacaattggc	ggcggccgcg	acttttcggg	gaaatgtgcg	cggaaaccct	atgtgtttat	10080
ttttctaaat	acattcaaat	atgtatccgc	tcatgagaca	ataaccctga	taaagtcttc	10140
aataatattg	aaaaaggaag	agtatgagta	ttcaacattt	ccgtgtcgcc	cttattccct	10200
tttttgcggc	attttgctt	cctgtttttg	ctcaccacga	aacgctggtg	aaagtaaaag	10260
atgctgaaga	tcagttgggt	gcacgagtgg	gttacatcga	actggatctc	aacagcggtg	10320
agatccttga	gagttttcgc	cccgaagaac	gttttccaat	gatgagcact	tttaaagttc	10380
tgctatgtgg	cgcggtatta	tcccgtattg	acgccgggca	agagcaactc	ggtcgccgca	10440
tacactattc	tcagaatgac	ttggttgagt	actcaccagt	cacagaaaag	catcttacgg	10500
atggcatgac	agtaagagaa	ttatgcagtg	ctgccataac	catgagtgat	aacactgcgg	10560
ccaacttact	tctgacaacg	atcggaggac	cgaaggagct	aaccgctttt	ttgcacaaca	10620
tgggggatca	tgtaactcgc	cttgatcggt	gggaaccgga	gctgaatgaa	gccataccaa	10680
acgacgagcg	tgacaccacg	atgcctgtag	caatggcaac	aacgttgcgc	aaactattaa	10740
ctggcgaact	acttactcta	gcttcccggc	aacaattaat	agactggatg	gaggcggata	10800
aagttgcagg	accacttctg	cgctcggccc	ttccggctgg	ctggtttatt	gctgataaat	10860
ctggagccgg	tgagcgtggg	tctcgcggta	tcatgtcagc	actggggcca	gatggtaagc	10920
cctcccgtat	cgtagttatc	tacacgacgg	ggagtccagg	aactatggat	gaacgaaata	10980
gacagatcgc	tgagataggt	gcctcactga	ttaagcattg	gtaactgtca	gaccctaggc	11040
cgggcaacaa	ttggcggccg	gccctgcatt	aatgaatcgg	ccaacgcgcg	gggagaggcg	11100
gtttgcgcat	tgggcgctct	tccgcttctc	cgctcactga	ctcgcctgcg	tcggctgctt	11160
ggctgcggcg	agcggtatca	gctcactcaa	agggcgtaat	acggttatcc	acagaatcag	11220
gggataacgc	aggaaagaac	atgtgagcaa	aaggccagca	aaaggccagg	aaccgtaaaa	11280
aggccgcggt	gctggcggtt	ttccataggc	tccgcccccc	tgacgagcat	cacaaaaatc	11340
gacgctcaag	tcagaggtgg	cgaaaaccga	caggactata	aagataccag	gcgtttcccc	11400
ctggaagctc	cctcgtgcgc	tctcctgctc	cgaccctgcc	gcttaccgga	tacctgtccg	11460
cctttctccc	ttcgggaagc	gtggcgcttt	ctcatagctc	acgctgtagg	tatctcagtt	11520
cgggtgtaggt	cgttcgcctc	aagctgggct	gtgtgcacga	acccccgctt	cagcccagcc	11580
gctgcgcctt	atccggtaac	tatcgtcttg	agtccaaccc	ggtaagacac	gacttatcgc	11640
cactggcagc	agccactggg	aacaggatta	gcagagcgag	gtatgtaggc	gggtgtacag	11700
agttcctgaa	gtgggtggcct	aactacggct	acactagaag	gacagtattt	ggtatctgcg	11760
ctctgctgaa	gccagttacc	ttcggaaaaa	gagttggtag	ctcttgatcc	ggcaaaacaaa	11820
ccaccgctgg	tagcgggtgg	ttttttggtt	gcaagcagca	gattacgcgc	agaaaaaaag	11880
gatctcaaga	agatcctttg	atcttttcta	cggggtctga	cgctcagtg	aacgaaaact	11940
c						11941

```

<210> 31
<211> 11216
<212> DNA
<213> 载体

<220>
<221> misc_特征
<222> 8117
<223> n = A,T,C 或 G

<400> 31
acgttgtaaa acgacggcca gtgaattgta atacgactca ctatagggcg aattgggtac 60
ggggccccc ctcgaggtcg agttgggggtg gggaaaagga agaaacgcgg gcgtattggc 120
cccaatgggg tctcggtggg gtatcgacag agtgccagcc ctgggaccga accccgcgtt 180
tatgaacaaa cgacccaaca cccgtgcgtt ttattctgtc tttttattgc cgtcatagcg 240
cgggttcctt cgggtattgt ctccttcctg cgacggtatc aaggtggcga ccggaatggt 300
gagctgagag aatagccggg cgcgctgtga gccgaagtcg cccccgccct ggccacttcc 360
ggcgcgccga gtccttaggc cgccaggggg cgccggcgcg cgcccagatt ggggacaaaag 420
gaagccgggg cggccgcggt attaccataa aaggcaaaaca ctggtcggag gcgtccccgc 480
ggcgcgcggc aggaagccag gccccaaacc cctcccaacc gggcgccagc cccgcctccg 540
cccggttcaa acagcgaccg ggtcgcgcgc gcgcacgcag cggccacacc ctcgggcgcc 600
agcggctcgg gcaggaagtg gcgcaagcgc ccgggcccc gaacgcacgc gcgattagcg 660
ccattgagtc ccagcgcgca cgcgcaatta gcgcaattc ccagcgcgca cgcagttagc 720
gccccaaagg ccagcgcgca cgcgcatggc gccccagccc ccaccgggccc tgacggggggc 780
tacgccgcgc ccaccgtgcg atccccattg gcaagagccc ggctcagaca aagacccccg 840
cggttgcccc cgccccgaga gcggcaccccc cggagcgcgc ccgcccagac gcggcctcgc 900
gctgcggaac tggcgtgggg tgtcccccat ctccggaggc ccaggggctt ctcccgcgcc 960
cccacggcg gtcocggtcc gccccatgcg ccccccgctg cggcccagac ggcggtctcg 1020
cacgggcgaa gggccgcggc cgcattgcccc ggtcggctgg ccgggcttac ctggcgggcg 1080
gtgtggacgg gcggcgatc ggcaaaggcg aggcctctgtg ctcgcgggcg gacgcggtct 1140
cggcgggtgt ggcgctcgc gccgctgggt ttatagggc gccgcccggg ccgctcgagc 1200
cataaaagcc aactttcgga acggcgcacg ctgattggcc ccgcgcgct cactcaccgg 1260
cttcgccgca cagtgcagca tttttttacc cctctcccc tccttttgcg aaaaaaaaaa 1320
agagcgagag cgagattgag gaagaggagg agggagagtt ttggcgttgg ccgccttggg 1380
gtgctgggccc cgggggctgg gggcgcgcgc cgtggccccc gcgccccacg ctgggcagtg 1440
cccggttcgg ccccgcatgg ccaggcctgc ccccggcctg ccgctctctc gggcccccca 1500
cccaccgcgg gacatcctag gtgtggacat ctcttgggca ctgagcgcgc aggtgggggtg 1560
ggccagggtc tgcacgggtg ccaggggcct gggttctgta cgctcctgca gaaggagctc 1620
ttggagggca tggagtggcc aggcagtcac tcccccttgc cgacttcaga gcaactgcc 1680
tgaaagcagg gcctgaggac ctctggctgt ggggctcagc tagctaaatg tgctgggtgg 1740
gtcactaggg agagacctgg gcttgagagg tagagtgtgg tgttggggga gtcagggtgg 1800
ttgcggccat tagagtcgca ggaccacact cccaggaca gggcaggggc cagcgggtcca 1860
gtggctggag gtggcccgtg atgaaggcta caaacctacc cagccgcagc cctgggaagg 1920
aagtgggctc tacagggcag ggcacctttt accctggagc tgccctgctt tgagggtaac 1980
agtcacgccc agccaagacc aggcctgggg cgttagtggg tgacctaggc actgcggggc 2040
gggggggctg ggtctacaca gcctgggtct gggcccaccg tccgttgat gtctgctatg 2100
cgcagccaca gctgaactgc cctcccagac catctggagg ccgctggggg actctggggg 2160
ccaagactcc atgtgccaca gaggattggg ggcggggcgg tgctaggaac tcaaagccag 2220
cctgggaaga ccctgtcctt gtcacccttt ctctgccttg gtctgtccac tgagtagcac 2280
acaagaccgg gtgggcaggg tccgttctgc tccgggaate acagactgtg tgtaccagg 2340
tggtgggcat gcagcgatca gtggcgtggg accacagagg gggcccgcgg taccaagctt 2400
gggaattgcg tgcaaaaaa acttctgttt tccagggtaa acagaatcta atgcagaatc 2460
taatgcaggg taaacagact taatgcagaa tctaattgat gcacaaatta aaaatcacta 2520
acgtgccctt tttagtgtga aaccagaga gagcacatac aagccaaaaa caaatgcttt 2580
attttaccta ggagacatta acattcacct ttacgtgttt aagattaatg caatgttaaa 2640

```

tattgtgaaa	actgtaactt	tgaatttcat	gatttttatg	tgaatattcc	agggtttaaa	2700
aaaacttgta	acatgacatg	gctgaataag	ataaaaaaaaa	aatctagcct	tttctccctt	2760
ctggctcata	tttgcgattt	cgatcatttt	gttttaaaaa	caaaacactg	caatgaatta	2820
aacttaatat	tcttctatgt	tttagagtaa	gttaaaacaa	gataaagtga	ccaaagtaat	2880
ttgaaagatt	caatgacttt	tgctccaacc	taggtgcaca	aggtaccttg	ttctttaaat	2940
tgggctttta	tgaaaatact	tctccagaat	tctggggatt	taagaaaaat	tatgcccaacc	3000
aacaagggtc	ttaccatttt	atgtaacatt	tttcaacgct	gcaaaaatgt	gtgtatttct	3060
atltgaagat	aaaaatcctc	agcaaaatcc	acattgcact	gtccttcaaa	gattagcctt	3120
ctttgaacta	gttaagacac	tattaagcca	agccagtatc	tccctgtaat	gaattcgttt	3180
ttctcttaat	tttcccctgt	aatttacct	gggagagctg	ggaaatatgt	ggatgtaaat	3240
ttctcagcca	cagagatgca	aagtataact	gtggggaaaa	aaaacttgag	ttaaatcctt	3300
acataattta	ggttttcatt	aacttaccaa	tgtagttttg	ttggaggcca	ttttttttat	3360
tgacagactg	aagagctatt	actagaaaaa	tgcatgacag	ttaaggtaag	tttgcatgac	3420
acaaaaaagg	taactaaata	caaattctgt	ttggattcca	accccccaagt	agagagcgca	3480
cactttcaaa	cgtgaataca	aatccagagt	agatctgcgc	tcctacctac	attgcttatg	3540
atgtacttaa	gtacgtgtcc	taaccatgtg	agtctagaaa	gactttactg	gggatcctgg	3600
tacctaaaac	agcttcacat	ggcttaaaat	aggggaccaa	tgtcttttcc	aatctaagtc	3660
ccatttataa	taaagtccat	gttccatttt	taaaggacaa	tcctttcggg	ttaaaaccag	3720
gcacgattac	ccaacaact	cacaacggta	aagcactgtg	aatcttctct	gttctgcaat	3780
cccaacttgg	tttctgctca	gaaaccctcc	ctctttccaa	tcggttaatta	aatacaaaaa	3840
ggaaaaaact	taagatgctt	caaccccgtt	tcgtgacact	ttgaaaaaag	aatcacctct	3900
tgcaaacacc	cgctcccagc	ccccgccgct	gaagcccggc	gtccagaggc	ctaagcgcg	3960
gtgcccggcc	ccaccgggga	gcgcgggctt	cggtgtcagc	gcatccgcgg	ggagaaacaa	4020
agggcgggc	acgggggctc	aagggcactg	cgccacaccg	cacgcgccta	ccccgcgcg	4080
gccacgttaa	ctggcggctg	ccgcagcctc	gggacagccg	gccgcgcgcc	gccaggtctg	4140
cgacgcggg	accacgggcc	gccctccggg	aggcccaagt	ctcgaccag	ccccgctgg	4200
cgctggggga	gggggcccct	ccgcgggaac	gcgggtgggg	gaggggaggg	ggaaatgccc	4260
tttgtctcga	aatggggcaa	ccgtcgccac	agctccctac	cccctcgagg	gcagagcagt	4320
ccccccacta	actaccgggc	tgcccgcgcg	ccaggccagc	cgcgaggcca	ccgcccgacc	4380
ctccactcct	tcccgcagct	cccggcgcg	ggtccggcga	gaaggggagg	ggaggggagc	4440
ggagaaccgg	gcccccgga	cgctgtgtgg	atctgaagca	ccaccagcga	gcgagagcta	4500
gagagaagga	aagccaccga	cttaccggcc	tccgagctgc	tccgggtcgc	gggtctgcag	4560
cgctcgggc	cctccgcgcc	tacagctcaa	gccaatccg	aagggggagg	gagccgggag	4620
ctgcgcgcgg	ggcgcgcggg	gggaggggtg	gcaccgccc	cgccgggccc	ccacgaagg	4680
cggggcagcg	ggcgcgcgcg	cggcgggggg	aggggcccgc	gccgcgccc	ctgggaattg	4740
gggcccagg	gggagggcgg	aggcgcggac	gaccgcggca	cttaccgttc	gcggcgtggc	4800
gcccgtgggt	ccccaaaggg	agggaaaggg	gaggcggggc	gaggacagtg	accggagtct	4860
cctcagcgtg	ggcttttctg	cttggcagct	tcagcggctg	gcgccaaaac	cggactccgc	4920
ccacttcctc	gcccgcgggt	gcgaggggtg	ggaatcctcc	agacgctggg	ggaggggggag	4980
ttgggagctt	aaaaactagt	acccttttgg	gaccactttc	agcagcgaac	tctcctgtac	5040
accaggggtc	agttccacag	acgcgggcca	ggggtgggtc	attgcggcgt	gaacaataat	5100
ttgactagaa	gttgattcgg	gtgtttccgg	aaggggccga	gtcaatccgc	cgagttgggg	5160
cacggaaaac	aaaaagggaa	ggctactaag	atltttctgg	cgggggttat	cattggcgta	5220
actgcaggg	ccacctccc	ggttgagggg	gctggatctc	caggctgcgg	attaagcccc	5280
tcccgctcgg	gttaatttca	aactgcgcga	cgtttctcac	ctgccttcgc	caaggcaggg	5340
gccgggacct	tattccaaga	ggtagtaact	agcaggactc	tagccttccg	caattcattg	5400
agcgcattta	cggaagtaac	gtcgggtact	gtctctggcc	gcaaggggtg	gaggagtacg	5460
catttggcgt	aaggtggggc	gtagagcctt	cccgccattg	gcggcggata	gggcgtttac	5520
gcagcggcct	gacgtagcgg	aagacgcggt	agtggggggg	aaggttctag	aaaagcggcg	5580
gcagcggctc	tagcggcagt	agcagcagcg	ccgggtccc	tgccggaggtg	ctcctcgcag	5640
agttgtttct	cgagcagcgg	cagttctcac	tacagcgcca	ggacgagtc	ggtctcgtgt	5700
cgctccgcga	gatctctctc	atctcgctcg	gctgcgggaa	atcgggctga	agcgaactgag	5760
tccgcgatgg	aggtaacggg	tttgaaatca	atgagttatt	gaaaagggca	tgccgagggc	5820
gttggcgcct	cagtggaagt	cggccagccg	cctccgtggg	agagaggcag	gaaatcggac	5880
caattcagta	gcagtggggc	ttaaaggttta	tgaacggggg	cttgagcgg	ggcctgagcg	5940
tacaaaacag	ttccccacc	tcagcctccc	ggcgcattt	cccttcaactg	gggggtggggg	6000
atggggagct	ttcacatggc	ggacgctgcc	ccgctggggg	gaaagtgggg	cgcggaggcg	6060

ggaattotta	ttccctttct	aaagcacgct	gcttcggggg	ccacggcgct	tcctcggcga	6120
gcgtttcggc	gggcagcagg	tectcgtgag	cgaggctgcg	gagcttcccc	tccccctctc	6180
tcccgggaac	cgatttggcg	gccgccattt	tcatggctcg	ccttcctctc	agcgttttcc	6240
ttataactct	tttattttct	tagtgtgctt	tctctatcaa	gaagtagaag	tggttaacta	6300
tttttttttt	cttctcgggc	tgttttcata	tcgtttcgag	gtggatttgg	agtgttttgt	6360
gagcttggat	ctttagagtc	ctgcgcacct	cattaaaggc	gctcagcctt	cccctcgatg	6420
aaatggcgcc	attgcgttcg	gaagccacac	cgaagagcgg	ggaggggggg	tgctccgggt	6480
ttgcggggccc	ggtttcagag	aagatcccaa	gcttattaat	agtaatcaat	tacgggggtca	6540
ttagttcata	gccccatata	ggagttccgc	gttacataac	ttacggtaaa	tggccccgct	6600
ggctgaccgc	ccaacgaccc	ccgcccattg	acgtcaataa	tgacgtatgt	tccccatgta	6660
acgccaatag	ggactttcca	ttgacgtcaa	tgggtggagt	atttacggta	aactgcccac	6720
ttggcagtac	atcaagtgta	tcatatgcc	agtacgcccc	ctattgacgt	caatgacggt	6780
aaatggcccc	cctggcatta	tgcccagtac	atgaccttat	gggactttcc	tacttggcag	6840
tacatctacg	tattagtcac	cgctattacc	atgggtgatgc	ggttttggca	gtacatcaat	6900
gggcgtggat	agcggtttga	ctcacgggga	tttccaagtc	tccacccccat	tgacgtcaat	6960
gggagtttgt	tttggcacca	aaatcaacgg	gactttccaa	aatgtcgtaa	caactccgcc	7020
ccattgacgc	aaatgggcgg	taggcgtgta	cggtgggagg	tctatataag	cagagctggt	7080
ttagtgaacc	gtcagatcgg	atccgcctga	gaaaggaagt	gagctgtaaa	ggctgagctc	7140
tctctctgac	gtatgtagcc	tctggttagc	ttcgtcactc	actgttcttg	actcagcatg	7200
gcaatctgat	gaaatcccag	ctgtaagtct	gcagaaattg	atgatctatt	aaacaataaa	7260
gatgtccact	aaaatggaag	tttttctctg	catactttgt	taagaagggt	gagaacagag	7320
tacctacatt	ttgaatggaa	ggattggagc	tacgggggtg	ggggtggggg	gggattagat	7380
aaatgcctgc	tctttactga	aggctcttta	ctattgcttt	atgataatgt	ttcatagttg	7440
gatatcataa	tttaacaag	caaaacccaa	ttaagggcca	gctcattcct	ccagatccac	7500
tagtaattct	gtggaatgtg	tgtcagttag	ggtgtggaaa	gtccccaggc	tccccagcag	7560
gcagaagtat	gcaaagcatg	catctcaatt	agtacgcaac	caggtgtgga	aagtccccag	7620
gctccccagc	aggcagaagt	atgcaaagca	tgcatctcaa	ttagtcagca	accatagctc	7680
cgccccctaac	tccgccccatc	ccgccccctaa	ctccgcccag	ttccgccccat	tctccgcccc	7740
atggctgact	aatttttttt	atztatgcag	aggccgaggc	cgctctgccc	tctgagctat	7800
tccagaagta	gtgaggaggc	ttttttggag	gcctaggctt	ttgcaaaaag	ctcccggggag	7860
cttgatatac	catttttcgga	tctgatcaag	agacaggatg	aggatcgttt	cgcatgattg	7920
aacaagatgg	attgcacgca	ggttctccgg	ccgcttgggt	ggagaggcta	ttcggctatg	7980
actgggcaca	acagacaatc	ggctgctctg	atgccgccgt	gttccggctg	tcagcgcagg	8040
ggcgccccggt	tctttttgtc	aagaccgacc	tgtccggtgc	cctgaatgaa	ctgcaggacg	8100
aggcagcgcg	gctatcstgg	ctggccacga	cgggcgttcc	ttgctcagct	gtgctcgacg	8160
ttgtcactga	agcgggaagg	gactggctgc	tattgggcga	agtgccgggg	caggatctcc	8220
tgctcatcca	ccttgctcct	gcccagaaaag	tatccatcat	ggctgatgca	atgcgggggc	8280
tgcatagcgt	tgatccggct	acctgccccat	tgcaccaaca	agcgaaacat	cgcatcgagc	8340
gagcacgtac	tcggatggaa	gccggtcttg	tcgatcagga	tgatctggac	gaagagcatc	8400
aggggctcgc	gccagccgaa	ctgttcgcca	ggctcaaggc	gcgcatgccc	gacggcgagg	8460
atctcgtcgt	gacccatggc	gatgcctgct	tgcccgaatat	catgggtgga	aatggccgct	8520
tttctggatt	catcgactgt	ggccggctgg	gtgtggcgga	ccgctatcag	gacatagcgt	8580
tggctacccg	tgatattgct	gaagagcttg	gcggcgaatg	ggctgaccgc	ttcctcgtgc	8640
tttacgggat	cgccgctccc	gattcgcagc	gcctcgcctt	ctatcgcctt	cttgacgagt	8700
tcttctgagc	gggactctgg	ggttcgaaat	gaccgaccaa	gcgacgcccc	acctgccatc	8760
acgagatttc	gattccaccg	ccgcttcta	tgaaagggtg	ggcttcggaa	tcgttttccg	8820
ggacgcccgc	tggatgatcc	tccagcgcgg	ggatctcatg	ctggagttct	tcgcccaccc	8880
caacttgttt	attgcagctt	ataatggtta	caaataaagc	aatagcatca	caaatttcac	8940
aaataaagca	tttttttcac	tgcatcttag	ttgtggtttg	tccaaactca	tcaatgtatc	9000
ttatcatgtc	tgtataccgt	cgagactagt	tctagagcgg	ccgccaccgc	ggtggagctc	9060
cagcttttgt	tccctttagt	gagggttaat	ttcgagcttg	gcgtaatcat	ggtcatagct	9120
gtttcctgtg	tgaaattggt	atccgctcac	aattccacac	aacatacgag	ccggaagcat	9180
aaagtgtaaa	gcctggggtg	cctaattgagt	gagctaactc	acattaattg	cgttgcgctc	9240
actgcccgct	ttccagtcgg	gaaacctgtc	gtgccagggg	gtacctaggc	cgggcaacaa	9300
ttggcgccgc	gccgcacttt	tccgggaaat	gtgcgcggaa	cccctatttg	tttatttttc	9360
taaatacatt	caaatatgta	tccgctcatg	agacaataac	cctgataaat	gcttcaataa	9420
tattgaaaaa	ggaagagtat	gagtattcaa	catttccgtg	tcgcccttat	tccctttttt	9480

```

gcggcatttt gccttcctgt ttttgetcac ccagaaacgc tggtgaaagt aaaagatgct 9540
gaagatcagt tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaaagatc 9600
cttgagagtt ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta 9660
tgtggcgcgg tattatcccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac 9720
tattctcaga atgacttggg tgagtactca ccagtcacag aaaagcatct tacggatggc 9780
atgacagtaa gagaattatg cagtgtctgc ataaccatga gtgataacac tgcggccaac 9840
ttacttctga caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg 9900
gatcatgtaa ctgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac 9960
gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaact attaatggc 10020
gaactactta ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagt 10080
gcaggaccac ttctgcgctc ggcccttccg gctggctggg ttattgctga taaatctgga 10140
gccggtgagc gtgggtctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc 10200
cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag 10260
atcgtcgaga taggtgcctc actgattaag cattggtaac tgtcagacc taggccgggc 10320
aacaattggc ggccggccct gcattaatga atcggccaac gcgcggggag aggcggtttg 10380
cgtattgggc gctcttccgc ttctctgctc actgactcgc tgcgctcggg cgttcggctg 10440
cggcgagcgg tatcagctca ctcaaaggcg gtaatacggg tatccacaga atcaggggat 10500
aacgcaggaa agaacatgtg agcaaaaagg cagcaaaaagg ccaggaaccg taaaaaggcc 10560
gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcaca aaatcgacgc 10620
tcaagtacga ggtggcgaaa cccgacagga ctataaagat accaggcggt tccccctgga 10680
agctccctcg tgcgctctcc tgttccgacc ctgcccgtta ccggatacct gtccgccttt 10740
ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcgggtg 10800
taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgctcagcc cgaccgctc 10860
gccttatccg gtaactatcg tcttgagctc aaccgggtaa gacacgact atcgccactg 10920
gcagcagcca ctggtaacag gattagcaga gcgaggatg taggcgggtg tacagagttc 10980
ttgaagtggg ggcctaacta cggctacact agaaggacag tatttgggat ctgcgctctg 11040
ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaacacc 11100
gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct 11160
caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactc 11216

```

<210> 32

<211> 11105

<212> DNA

<213> 载体

<400> 32

```

acgttgtaaa acgacggcca gtgaattgta atacgactca ctatagggcg aattgggtac 60
cgggcccccc ctcgaggtcg agttgggggtg gggaaaagga agaaacgcgg gcgtattggc 120
cccaatgggg tctcggtggg gtatcgacag agtgccagcc ctgggaccca accccgcgtt 180
tatgaacaaa cgaccaca cccgtgcggt ttattctgtc tttttattgc cgtcatagcg 240
cgggttcctt ccgggtattgt ctcttccgct cgacgggtatc aagggtggcga ccggaatggg 300
gagctgcgag aatagccggg cgcgctgtga gccgaagtgc cccccgcctt ggccaacttc 360
ggcgcgccga gtcccttaggc cgccaggggg cgccggcgcg cgcccagatt ggggacaaag 420
gaagccgggc cggccgcggt attaccataa aaggcaacaa ctggtcggag gcgtccccgc 480
ggcgcgcggc aggaagccag gccccaaccc cctcccaacc gggcgcagc cccgcctccg 540
cccggttcaa acagcgcacc ggtcgcgcgc gcgcacgcag cggccacacc ctcgggcgcc 600
agcggctcgg gcaggaagtg ggcgaagcgc ccgggccccca gaacgcacgc gcgattagcg 660
ccattgagtc ccagcgcgca cgcgcaatta gcgccaattc ccagcgcgca cgcagttagc 720
gccccaaagg ccagcgcgca cgcgcatggc gccccagccc ccaccgggccc tgacggggggc 780
tacgccgcgc ccaccgtgcg atccccattg gcaagagccc ggctcagaca aagaccccgc 840
cggttgcccc cgccccgaga gggcaccccc cggagcgcgc ccgcccgagc gcggcctcgc 900
gcctgcgaac tggcgtgggg tgtcccccat ctccggaggc ccaggggctt ctccccgcgc 960
ccccacgggc gtccggttcc gcccatgctg cccccgctg cggcccagac ggcgctctg 1020
cacggggcga gggcgcggc cgcattgccc cctcggctgg cgggcttac ctggcggcgg 1080
gtgtggacgg gcggcgatc ggcaaggcg aggtctgtg ctgcggggcg gacgcggtct 1140
cggcgggtgg ggcgcgctgc gccgctgggt tttatagggc gccgcgcggc ccgctcagac 1200
cataaaaggc aactttcgga acggcgcacg ctgattggcc ccgcgcgct cactcaccgg 1260

```

cttcgcccga	cagtgcagca	tttttttacc	ccctctcccc	tccttttgcg	aaaaaaaaaa	1320
agagcgagag	cgagattgag	gaagaggagg	agggagagtt	ttggcgttgg	ccgccttggg	1380
gtgctggggcc	cgggggctgg	gggcgcgcg	cgtggccccc	gcgccccacg	ctgggcagtg	1440
cccggttcgg	ccccgcatgg	ccaggcctgc	ccccggcctg	cccgctctctc	gggcccccca	1500
cccaccgcg	gacatcctag	gtgtggacat	ctcttgggca	ctgagcgccc	aggtgggggtg	1560
ggccagggtc	tgacgggtg	ccagggccct	gggttctgta	cgctcctgca	gaaggagctc	1620
ttggagggca	tggagtggcc	aggcagtcac	tcccccttgc	cgacttcaga	gcaactgcc	1680
tgaaagcagg	gcctgaggac	ctctggctgt	ggggctcagc	tagctaaatg	tgctgggtgg	1740
gtcactagg	agagacctgg	gcttgagagg	tagagtgtgg	tggtggggga	gtcaggtggc	1800
ttcgggccat	tagatcgca	ggaccacact	ccccagaca	gggcaggggc	cagcgggtcca	1860
gtggctggag	gtggcccgtg	atgaaggcta	caaacctacc	cagccgcagc	cctgggaagg	1920
aagtgggctc	tacagggcag	ggcacctttt	accctggagc	tgctgtcttt	tgagggtaac	1980
agtcacgccc	agccaagacc	aggcctgggg	cgttagtggg	tgacctaggc	actgcggggc	2040
gggggggctg	ggtctacaca	gcctgggtct	gggcccaccg	tccgttgat	gtctgtctatg	2100
cgcagccaca	gctgaactgc	cctcccagac	catctggagg	ccgctggggg	actctggggga	2160
ccaagactcc	atgtgccaca	gaggattggg	ggcggggcgg	tgctaggaac	tcaaagccag	2220
cctgggaaga	ccctgtcctt	gtcacccttt	cttgccttgg	gtctgtccac	tgagtagcac	2280
acaagaccgg	gtgggcaggg	tccgttctgc	tccgggaatc	acagactgtg	tgtaccagg	2340
tggtgggcat	gcagcgatca	gtggcgtggg	accacagagg	gggcccgcgg	taccaagctt	2400
gggaattgcg	tgcaaaaaca	acttctgttt	tccagggtaa	acagaatcta	atgcagaatc	2460
taatgcaggg	taaacagact	taatgcagaa	tctaattgatg	gcacaaatta	aaaatcacta	2520
acgtgccctt	tttagtgtga	aaccagaga	gagcacatac	aagccaaaaa	caaatgcttt	2580
atttaccta	ggagacatta	acattcacct	ttacgtgttt	aagattaatg	caatgttaa	2640
tattgtgaaa	actgtaactt	tgaatttcat	gatttttatg	tgaatattcc	agggtttaa	2700
aaaacttgta	acatgacatg	gctgaataag	ataaaaaaaaa	aatctagcct	tttctccctt	2760
ctggctcata	tttgcgattt	cgatcatttt	gtttaaaaaa	caaaacactg	caatgaatta	2820
aacttaatat	tcttctatgt	tttagagtaa	gttaaaacaa	gataaagtga	ccaaagtaat	2880
ttgaaagatt	caatgacttt	tgctccaacc	taggtgcaca	aggtaccttg	ttctttaa	2940
tgggctttaa	tgaaaatact	tctccagaat	tctggggatt	taagaaaaat	tatgccaac	3000
aacaagggct	ttaccatttt	atgtaacatt	tttcaacgct	gcaaaaatgt	gtgtatttct	3060
atgtgaagat	aaaaatcctc	agcaaaatcc	acattgcact	gtccttcaaa	gattagcctt	3120
ctttgaaacta	gttaagacac	tattaagcca	agccagtatc	tccctgtaat	gaattcgttt	3180
ttctctta	tttcccctgt	aatttacact	gggagagctg	ggaaatatgt	ggatgtaaat	3240
ttctcagcca	cagagatgca	aagttatact	gtggggaaaa	aaaacttgag	ttaaatcctt	3300
acataatttta	ggttttcatt	aacttaccaa	tgtagttttg	ttggaggcca	ttttttttat	3360
tgagactttg	aagagctatt	actagaaaaa	tgcatgacag	ttaaggtaag	tttgcagc	3420
acaaaaagg	taactaaata	caaattctgt	ttggattcca	acccccagt	agagcgcga	3480
cactttcaaa	cgtgaataca	aatccagagt	agatctgcgc	tcctacctac	attgcttatg	3540
atgtacttaa	gtacgtgtcc	taaccatgtg	agtctagaaa	gactttactg	gggatcctgg	3600
tacctaaaac	agcttcacat	ggcttaaaat	aggggaccaa	tgtcttttcc	aatctaagtc	3660
ccatttataa	taaagtccat	gttccatttt	taaaggacaa	tcctttcggg	ttaaaaccag	3720
gcacgattac	ccaacaact	cacaacggtg	aagcactgtg	aatcttctct	gttctgcaat	3780
ccaacttgg	tttctgctca	gaaaccctcc	ctctttccaa	toggtaatta	aataacaaaa	3840
ggaaaaaact	taagatgctt	caaccccggt	togtgacact	ttgaaaaaag	aatcacctct	3900
tgcaaacacc	cgctcccagc	ccccgcggct	gaagcccggc	gtccagaggc	ctaagcggcg	3960
gtgcccggcc	ccacccggga	gcgcgggct	cgtggtcagc	gcatccgcgg	ggagaaacaa	4020
aggccgcggc	acgggggctc	aagggcactg	gcgcaacccg	cacgcgccta	ccccgcgcg	4080
gccacgttaa	ctggcggtcg	ccgcagcctc	gggacagccg	gcccgcgcgc	gccaggctcg	4140
cggacgcggg	accacgcgcc	gccctccggg	aggcccaagt	ctcgaccag	ccccgcgtgg	4200
cgctggggga	gggggcgcct	ccgcgggaac	gcgggtgggg	gaggggagg	ggaaatgcgc	4260
tttgtctcga	aatggggcaa	ccgtcgccac	agctccctac	cccctcgagg	gcagagcagt	4320
ccccccacta	actaccgggc	tggcgcgcgc	ccaggccagc	cgcgaggcca	ccgcccagcc	4380
ctccactcct	tcccgcagct	cccggcgcgg	ggtccggcga	gaaggggagg	ggaggggagc	4440
ggagaaccgg	gccccgggga	cgctgtggc	atctgaagca	ccaccagcga	gcgagagcta	4500
gagagaagga	aagccaccga	cttaccggcc	tccgagctgc	tccgggtcgc	gggtctgcag	4560
cgtctccggc	cctccgcgcc	tacagctcaa	gccacatccg	aagggggagg	gagccgggag	4620
ctgcgcgcgg	ggccgcggg	gggaggggtg	gcaccgccc	cgccggggcg	ccacgaagg	4680

cggggcagcg	ggcgcgcgcg	cggggggggg	aggggcccgc	gccgcgccc	ctgggaattg	4740
gggcccctagg	gggagggcgg	agggcggcac	gaccgcggca	cttaccgttc	gcggcgtggc	4800
gcccgggtggt	ccccaaaggg	agggaaaggg	gagggcgggc	gaggacagt	accggagtct	4860
cctcagcggg	ggcttttctg	cttggcagcc	tcagcggctg	gcgccaaaac	cggactccgc	4920
ccacttcctc	gcccgcgggt	gcgaggggtg	ggaatcctcc	agacgctggg	ggagggggag	4980
ttgggagctt	aaaaactagt	acccttttgg	gaccactttc	agcagcgaac	tctcctgtac	5040
accaggggtc	agttccacag	acgcggggcca	gggggtgggtc	attgcccgtt	gaacaataat	5100
ttgactagaa	gttgattcgg	gtgtttccgg	aaggggcccga	gtcaatccgc	cgagttgggg	5160
cacggaaaac	aaaaagggaa	ggctactaag	atttttctgg	cggggggtat	cattggcgta	5220
actgcagggg	ccacctcccg	ggttgagggg	gctggatctc	caggctgcgg	attaagcccc	5280
tcccgtcggc	gttaatttca	aactgcgcga	gctttctcac	ctgccttcgc	caagtcaggg	5340
gccgggaccc	tattccaaga	ggtagtaact	agcaggactc	tagccttcgc	caattcattg	5400
agcgcattta	cggaagtaac	gtcgggtact	gtctctggcc	gcaaggggtg	gaggagtacg	5460
catttggcgt	aaggtggggc	gtagagcctt	cccgccattg	gcggcgggata	gggcttttac	5520
gcgacggcct	gacgtagcgg	aagacgcgtt	agtggggggg	aaggttctag	aaaagcggcg	5580
gcagcggctc	tagcggcagt	agcagcagcg	ccgggtcccg	tgccggaggtg	ctcctcgag	5640
agttgtttct	cgagcagcgg	cagttctcac	tacagcggca	ggacgagtcc	ggttcgtgtt	5700
cgtccgcgga	gatctctctc	atctcgctcg	gctgcgggaa	atcgggctga	agcactgag	5760
tccgcgatgg	aggtaacggg	tttgaaatca	atgagttatt	gaaaagggca	tggcgaggcc	5820
gttggcgcct	cagtggaaat	cgccagcccg	cctccgtggg	agagaggcag	gaaatcggac	5880
caattcagta	gcagtggggc	ttaaggttta	tgaacggggg	cttgagcggg	ggcctgagcg	5940
tacaaacagc	ttccccacc	tcagcctccc	ggcgccattt	cccttactg	ggggtggggg	6000
atggggagct	ttcacatggc	ggacgctgcc	ccgctggggg	gaaagtgggg	cgcgaggcgg	6060
ggaattctta	ttccctttct	aaagcacgct	gcttcggggg	ccacggcgtc	tctcggcgca	6120
gcgtttcggc	gggcagcagg	tctcgtgag	cgaggctgcg	gagcttccc	tccccctctc	6180
tcccgggaac	cgatttggcg	gccgccattt	tcatggctcg	ccttctctct	agcgttttcc	6240
ttataactct	tttattttct	tagtgtgctt	tctctatcaa	gaagtagaag	tggttaacta	6300
tttttttttt	cttctcgggc	tgttttcata	tcgtttcgag	gtggatttgg	agtgttttgt	6360
gagcttggat	ctttagagtc	ctgcgcacct	cattaaaggc	gctcagcctt	cccctcgatg	6420
aaatggcgcc	attgcgttcg	gaagccacac	cgaagagcgg	ggaggggggg	tgctccgggt	6480
ttgcggggcc	ggtttcagag	aagatcccaa	gcttattaat	agtaatcaat	tacggggtca	6540
ttagttcata	gccatataat	ggagttccgc	gttacataac	ttacggtaaa	tggcccgcct	6600
ggctgaccgc	ccaacgacc	ccgcccattg	acgtcaataa	tgacgtatgt	tcccatagta	6660
acgccaatag	ggactttcca	ttgacgtcaa	tgggtggagt	atttacggta	aactgcccac	6720
ttggcagtac	atcaagtgtg	tcataatgcca	agtacgcccc	ctattgacgt	caatgacggg	6780
aaatggcccc	cctggcatta	tgcccagtac	atgaccttat	gggactttcc	tacttggcag	6840
tacatctacg	tattagtcat	cgctattacc	atggttagtc	ggttttggca	gtacatcaat	6900
ggcgctggat	agcggtttga	ctcacgggga	tttccaagtc	tccaccccat	tgacgtcaat	6960
gggagtttgt	tttggcacca	aaatcaacgg	gactttccaa	aatgtcgtaa	caactccgcc	7020
ccattgacgc	aaatgggcgg	taggcgtgta	cggtgggagg	tctatataag	cagagctggt	7080
ttagtgaacc	gtcagatcgg	atccgcctga	gaaaggaagt	gagctgtaaa	ggctgagctc	7140
tctctctgac	gtatgtagcc	tctggttagc	ttcgtcactc	actgttcttg	actcagcatg	7200
gcaatctgat	gaaatcccag	ctgtaagtct	gcagaaattg	atgatctatt	aaacaataaa	7260
gatgtccact	aaaatggaag	tttttctctg	catactttgt	taagaagggg	gagaacagag	7320
tacctacatt	ttgaatggaa	ggattggagc	tacgggggtg	gggggtgggg	gggattagat	7380
aaatgcctgc	tctttactga	aggctcttta	ctattgcttt	atgataatgt	ttcatagttg	7440
gatatacata	tttaaacaag	caaaaaccaa	ttaagggcca	gctcattcct	ccagatccac	7500
tagttctaga	gcaaatctta	ccgggtaggg	gagggcgttt	tcccaggca	gtctggagca	7560
tgcgctttag	cagccccgct	gggcacttgg	cgctacacaa	gtggcctctg	gcctcgcaca	7620
cattccacat	ccaccggtag	gcgccaaccg	gctccgttct	ttgggtggcc	cttcgcgcca	7680
ccttctactc	ctcccctagt	caggaagtcc	ccccccgccc	cgcagctcgc	gtcgtgcagg	7740
acgtgacaaa	tggaagtagc	acgtctcact	agtctcgtgc	agatggacag	caccgctgag	7800
caatggaagc	gggtaggcct	ttggggcagc	ggccaatagc	agctttgctc	cttcgctttc	7860
tgggctcaga	ggctgggaag	gggtgggtcc	gggggcgggc	tcagggggcg	gctcaggggc	7920
ggggcggggc	cccgaaggtc	ctccggaggc	ccggcattct	gcacgcttca	aaagcgcacg	7980
tctgccgcgc	tgttctcctc	ttctcatct	ccgggccttt	cgaccagctt	accatgaccg	8040
agtacaagcc	cacggtgcgc	ctcgccacc	gogacgacgt	ccccagggc	gtacgcacc	8100

tcgcccgcgc	gttcgcgcgac	taccccgcca	cgcgccacac	cgtcgatccg	gaccgcccaca	8160
tcgagcgggt	caccgagctg	caagaactct	tcctcacgcg	cgtcgggctc	gacatcggca	8220
aggtgtgggt	cgcgacgac	ggcgccgcgg	tggcggctctg	gaccacgccg	gagagcgtcg	8280
aagcgggggc	ggtgttcgcc	gagatcggcc	cgcgcatggc	cgagttgagc	ggttcccggc	8340
tgcccgcgca	gcaacagatg	gaaggcctcc	tggcgcgcga	ccggcccgaag	gagcccgcgt	8400
ggttcctggc	caccgtcggc	gtctcgcccc	accaccaggg	caagggctctg	ggcagcggcg	8460
tcgtgctccc	cggagtggag	gcggccgagc	gcgcccgggt	gcccgccttc	ctggagacct	8520
ccgcgccccg	caacctcccc	ttctacgagc	ggctcggctt	caccgtcacc	gccgacgtcg	8580
aggtgcccga	aggaccgcg	acctgggtga	tgacccgcaa	gcccggtgcc	tgacgcccgc	8640
cccacgacct	gcagcggccg	accgaaagga	gcgcacgacc	ccatgcatag	ggtgggcttc	8700
ggaatcgttt	tccgggacgc	cggctggatg	atcctccagc	gcggggatct	catgctggag	8760
ttcttcgccc	accccaactt	gtttattgca	gcttataatg	gttacaata	aagcaatagc	8820
atcacaat	tcacaataa	agcatttttt	tactgctatt	ctagttgtgg	tttgtccaaa	8880
ctcatcaatg	tatcttatca	tgtctgtata	ccgtcgagat	ctagagcggc	cgccaccgcg	8940
gtggagctcc	agcttttgtt	cccttttagt	agggttaatt	tcgagcttgg	cgtaatcatg	9000
gtcatagctg	tttctgtgt	gaaattgtta	tccgctcaca	attccacaca	acatacgagc	9060
cggaagcata	aagtgtaaag	cctgggggtgc	ctaattgagtg	agctaactca	cattaattgc	9120
gttgcgctca	ctgcccgcctt	tccagtcggg	aaacctgtcg	tgccaggggg	tacctaggcc	9180
gggcaacaat	tggcggcccg	ccgcactttt	cggggaaatg	tgcgcggaac	ccctatttgt	9240
ttatttttct	aaatacattc	aaatatgtat	ccgctcatga	gacaataacc	ctgataaatg	9300
cttcaataat	attgaaaaag	gaagagtatg	agtattcaac	atttccgtgt	cgcccttatt	9360
cccttttttg	cggcattttg	cettctgttt	tttgtcacc	cagaaacgct	ggtgaaagta	9420
aaagatgctg	aagatcagtt	gggtgcacga	gtgggttaca	tcgaactgga	tctcaacagc	9480
ggtaaagatcc	ttgagagttt	tcgccccgaa	gaacgttttc	caatgatgag	cacttttaaa	9540
gttctgctat	gtggcgcggt	attatcccgt	attgacgccc	ggcaagagca	actcggctcg	9600
cgcatacact	attctcagaa	tgacttggtt	gagtactcac	cagtacagaa	aaagcatctt	9660
acggatggca	tgacagtaag	agaattatgc	agtgtgcca	taaccatgag	tgataaacact	9720
gcggccaact	tactttctgac	aacgatcgga	ggaccgaagg	agctaaccgc	ttttttgcac	9780
aacatggggg	atcatgtaac	tcgccttgat	cgttgggaac	cggagctgaa	tgaagccata	9840
ccaacgacg	agcgtgacac	cacgatgcct	gtagcaatgg	caacaacggt	gcgcaacta	9900
ttaactggcg	aactacttac	tctagcttcc	cggcaacaat	taatagactg	gatggaggcg	9960
gataaagttg	caggaccact	tctgcgctcg	gcccttccgg	ctggctgggt	tattgctgat	10020
aaatctggag	ccggtgagcg	tgggtctcgc	ggtatcattg	cagcactggg	gccagatggg	10080
aagccctccc	gtatcgtagt	tatctacacg	acggggagtc	aggcaactat	ggatgaacga	10140
aatagacaga	tcgctgagat	aggtgcctca	ctgattaagc	attggtaact	gtcagaccct	10200
aggccgggca	acaattggcg	gccggccctg	cattaatgaa	tcggccaacg	cgcggggaga	10260
ggcggtttgc	gtattgggcg	ctcttccgct	tcctcgctca	ctgactcgct	gcgctcggtc	10320
gttcggctgc	ggcgagcggg	atcagctcac	tcaaaggcgg	taatacgggt	atccacagaa	10380
tcaggggata	acgcagga	gaacatgtga	gcaaaaaggcc	agcaaaaaggc	caggaaccgt	10440
aaaaaggcgc	cgttgctggc	gtttttccat	aggctccgcc	cccctgacga	gcatcacaaa	10500
aatcgacgct	caagtacagag	gtggcgaaac	ccgacaggac	tataaagata	ccaggcgttt	10560
ccccctggaa	gctccctcgt	gcgctctcct	gttccgacct	tgccgcttac	cggatacctg	10620
tccgccttcc	tcccttcggg	aagcgtggcg	ctttctcata	gctcacgctg	taggtatctc	10680
agttcgggtg	aggtcgttcg	ctccaagctg	ggctgtgtgc	acgaaccccc	cgttcagccc	10740
gaccgctgcg	ccttatccgg	taactatcgt	cttgagtcca	acccggtaag	acacgactta	10800
tcgccactgg	cagcagccac	tggtaacagg	attagcagag	cgaggatgtg	aggcgggtgct	10860
acagagttct	tgaagtgggtg	gcctaactac	ggctacacta	gaaggacagt	atctgggtatc	10920
tcgctctg	tgaagccagt	taccttcgga	aaaagagttg	gtagctcttg	atccggcaaaa	10980
caaaccaccg	ctggtagcgg	tggttttttt	gtttgcaagc	agcagattac	gcgcagaaaa	11040
aaagatctc	aagaagatcc	tttgatcttt	tctacggggg	ctgacgctca	gtggaacgaa	11100
aactc						11105

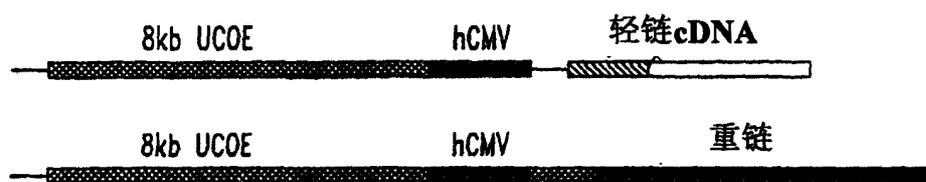


图1

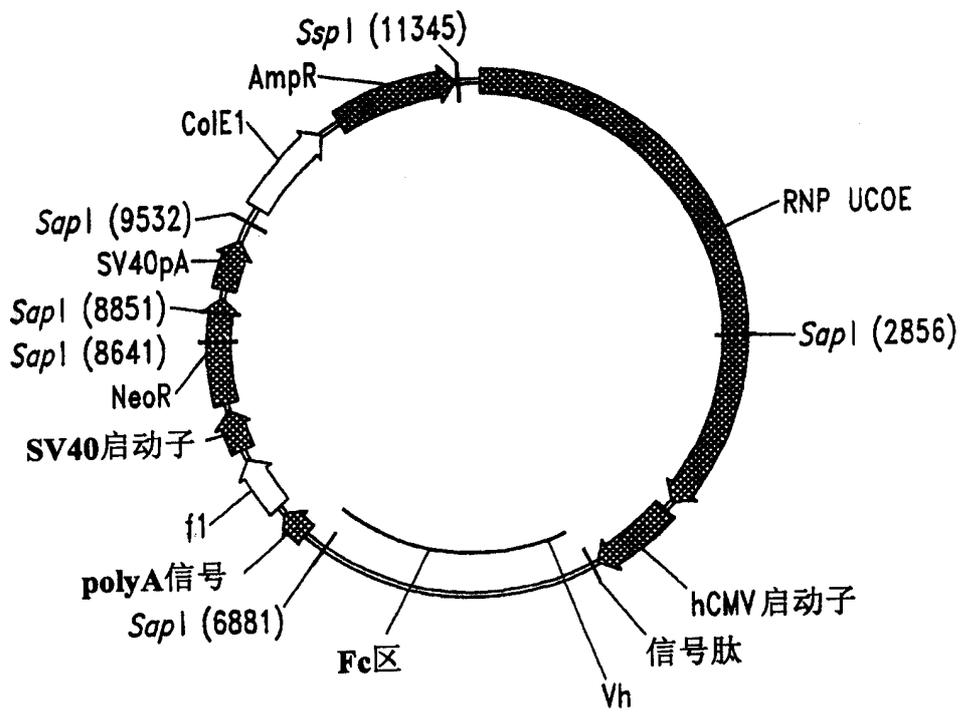


图2A

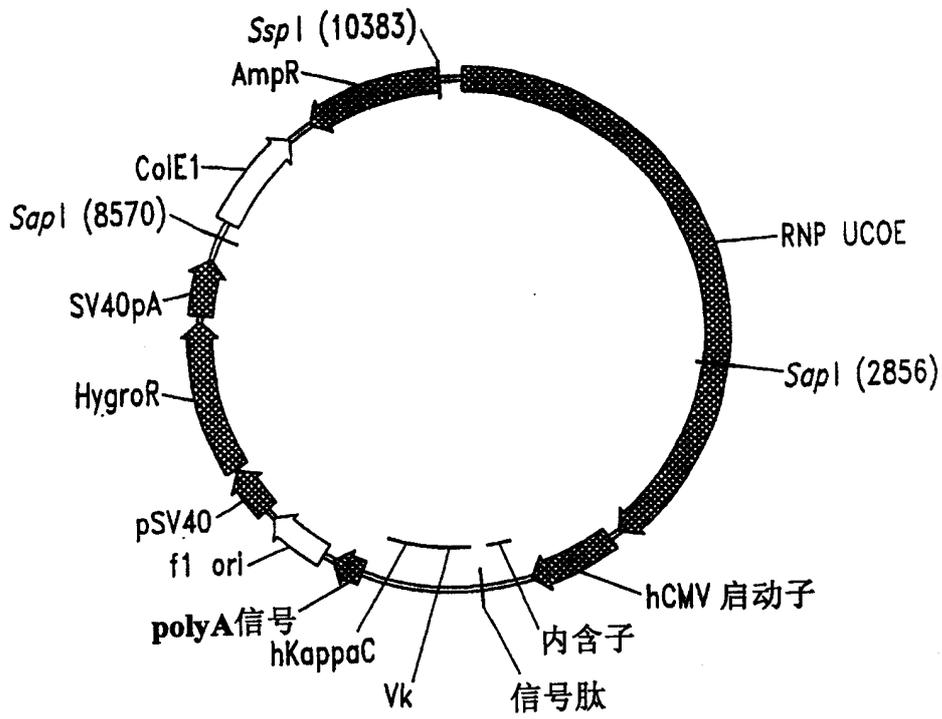


图2B

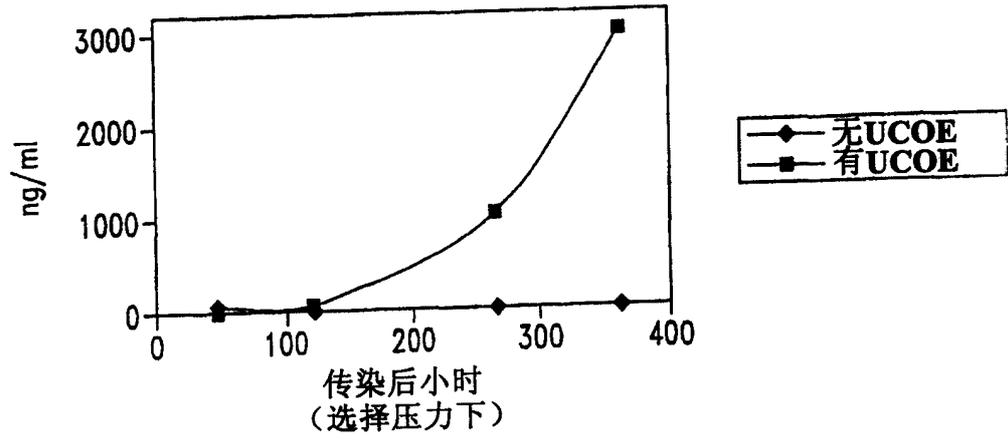


图3

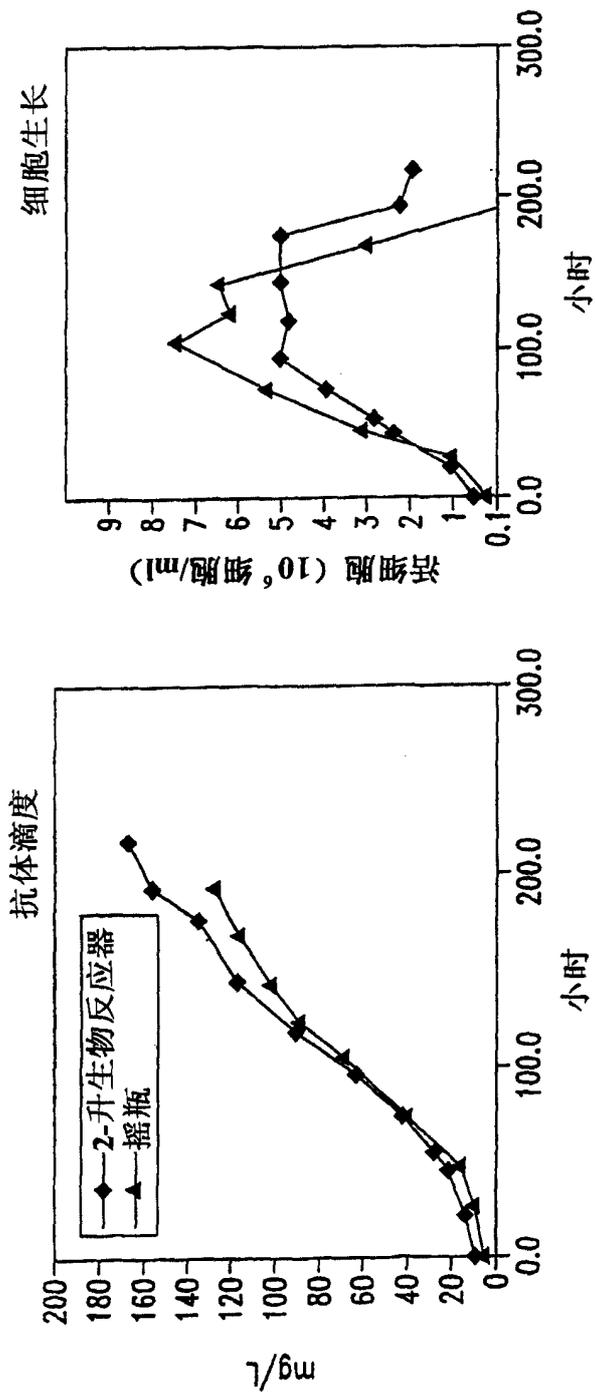


图4

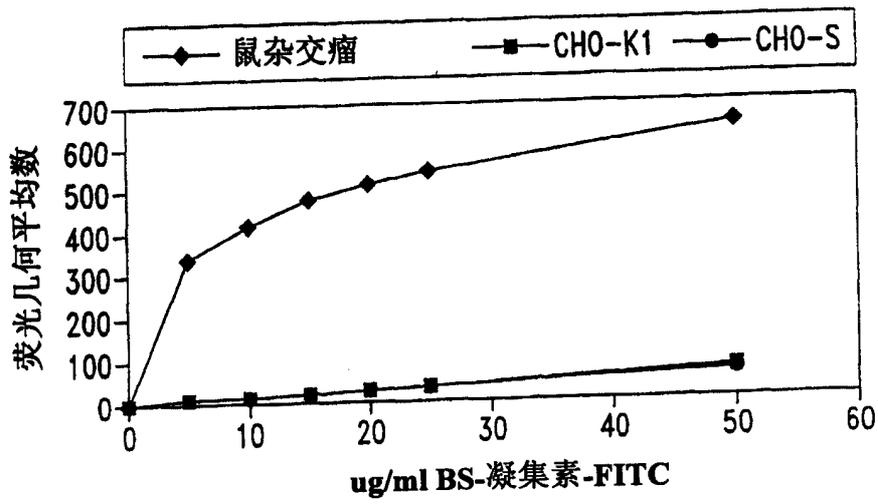


图5

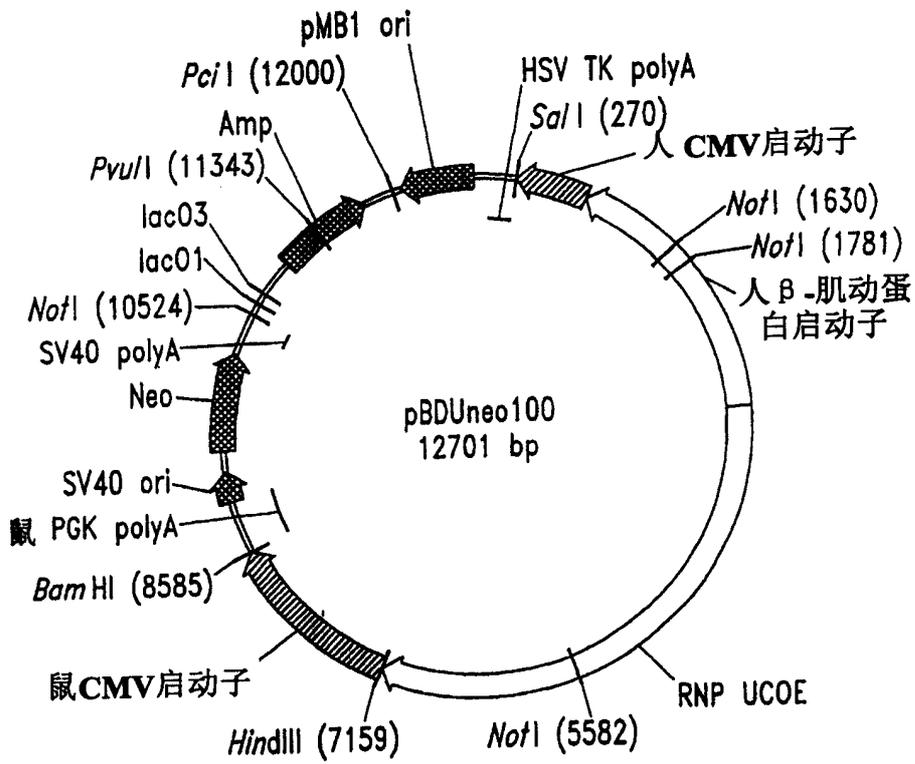


图6

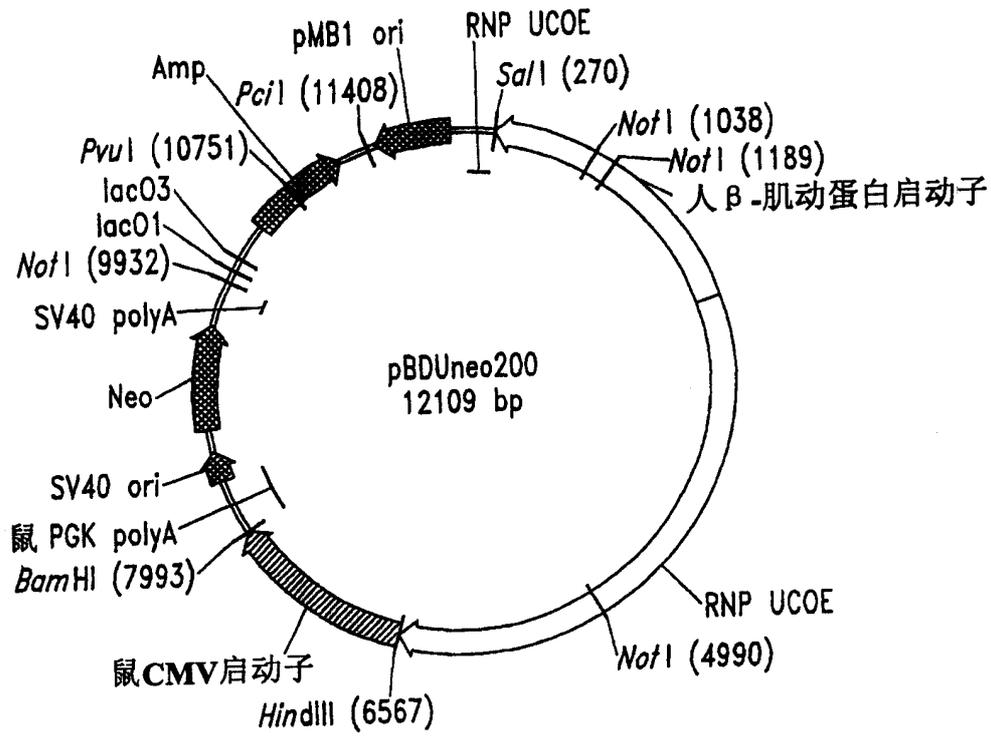


图7

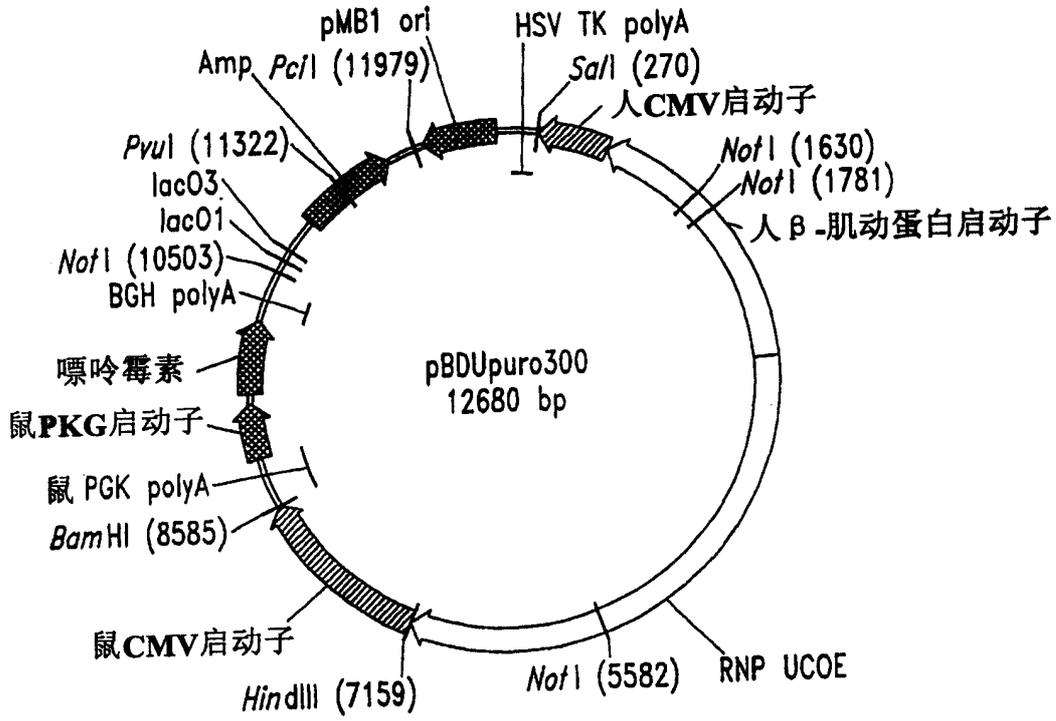


图8

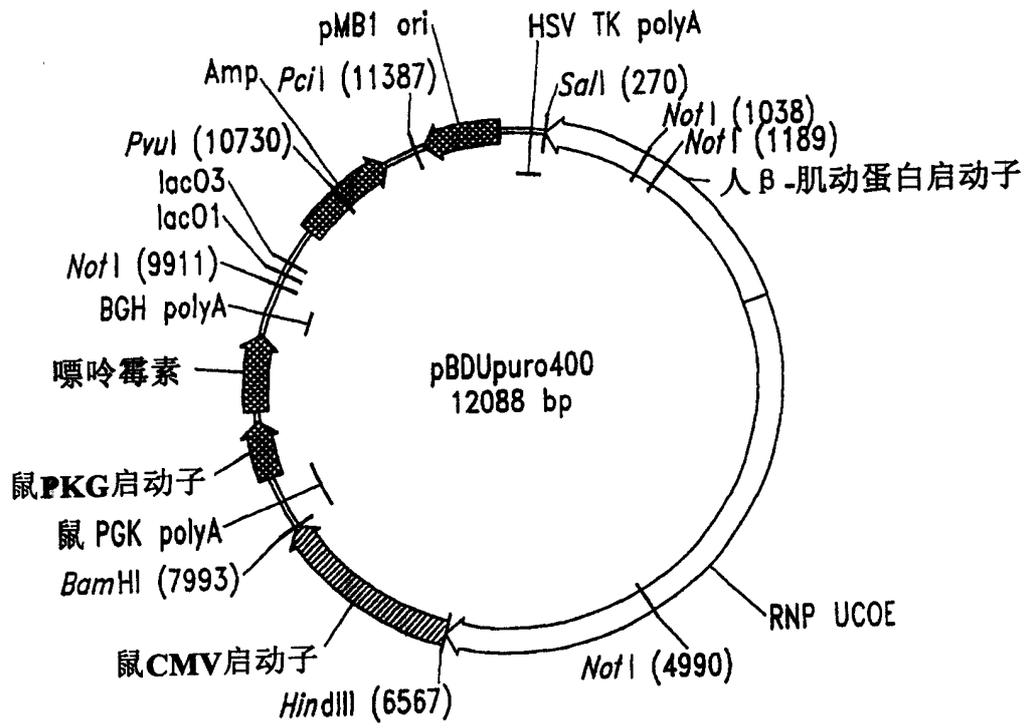


图9

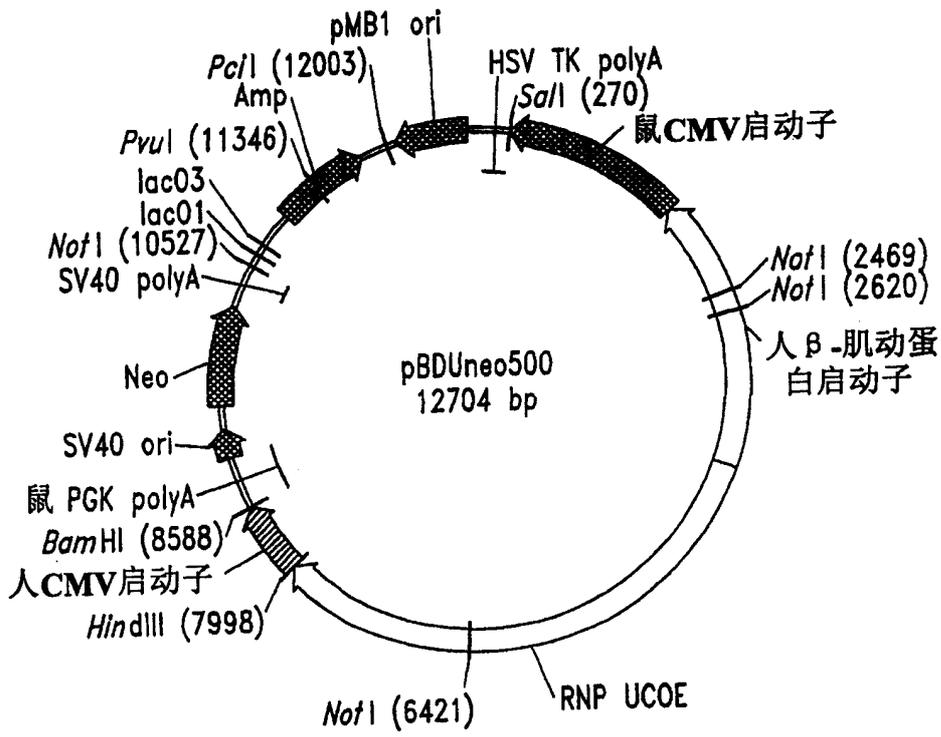


图10

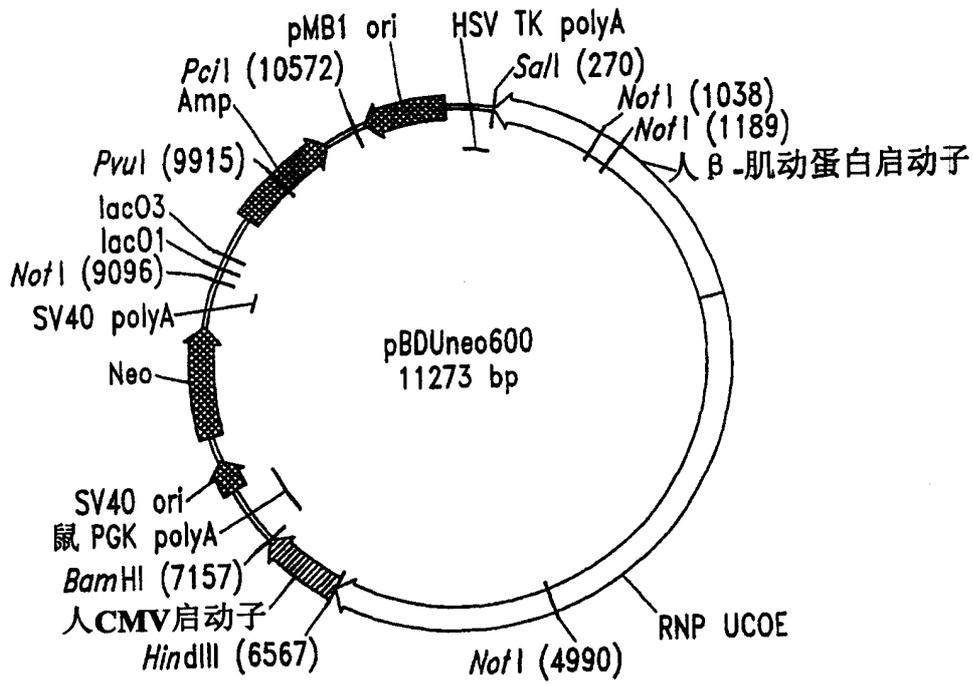


图11

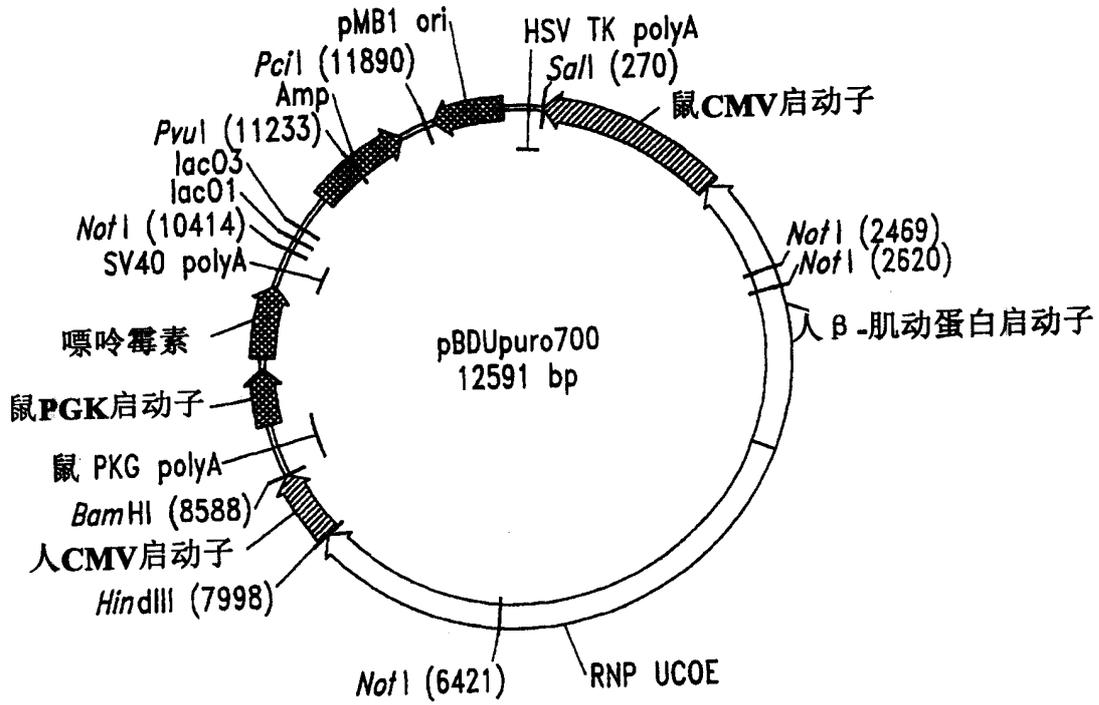


图12

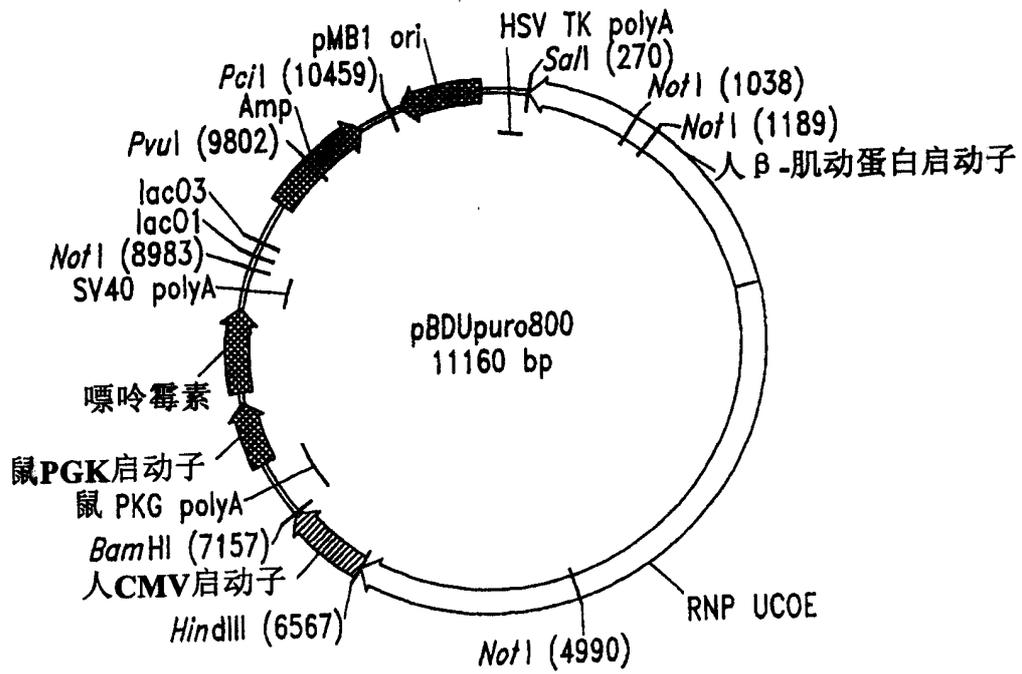


图13

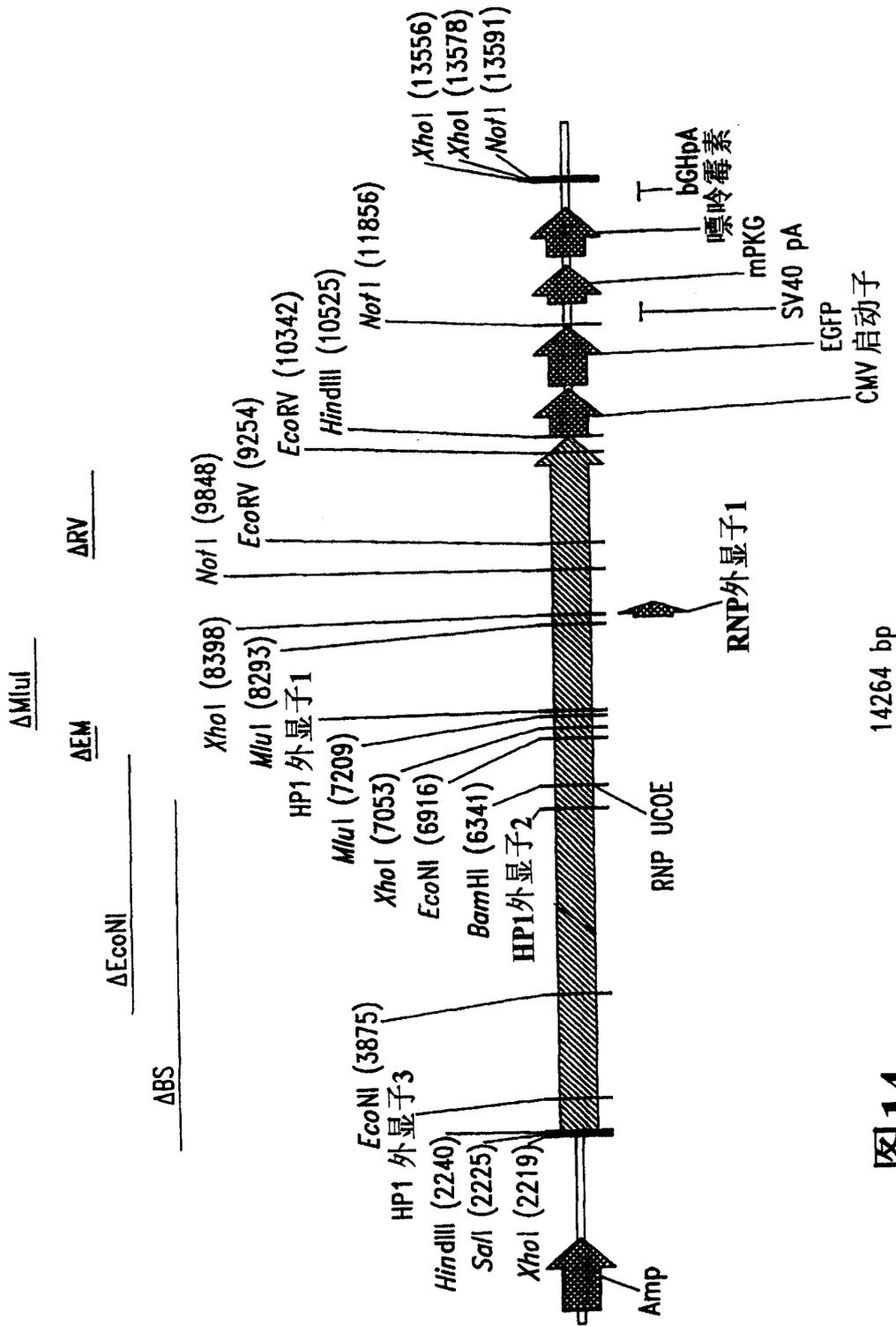


图14

图形表示载体pBDUpuro350

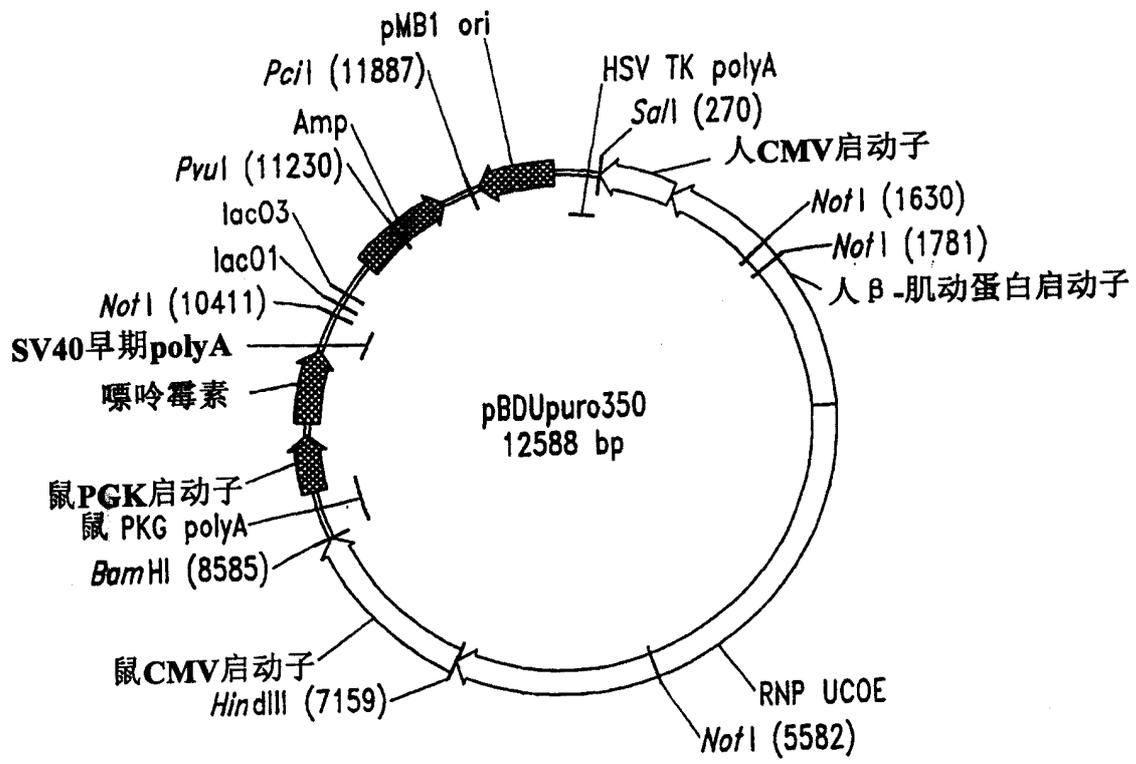


图15

图形表示载体pBDUpuro350

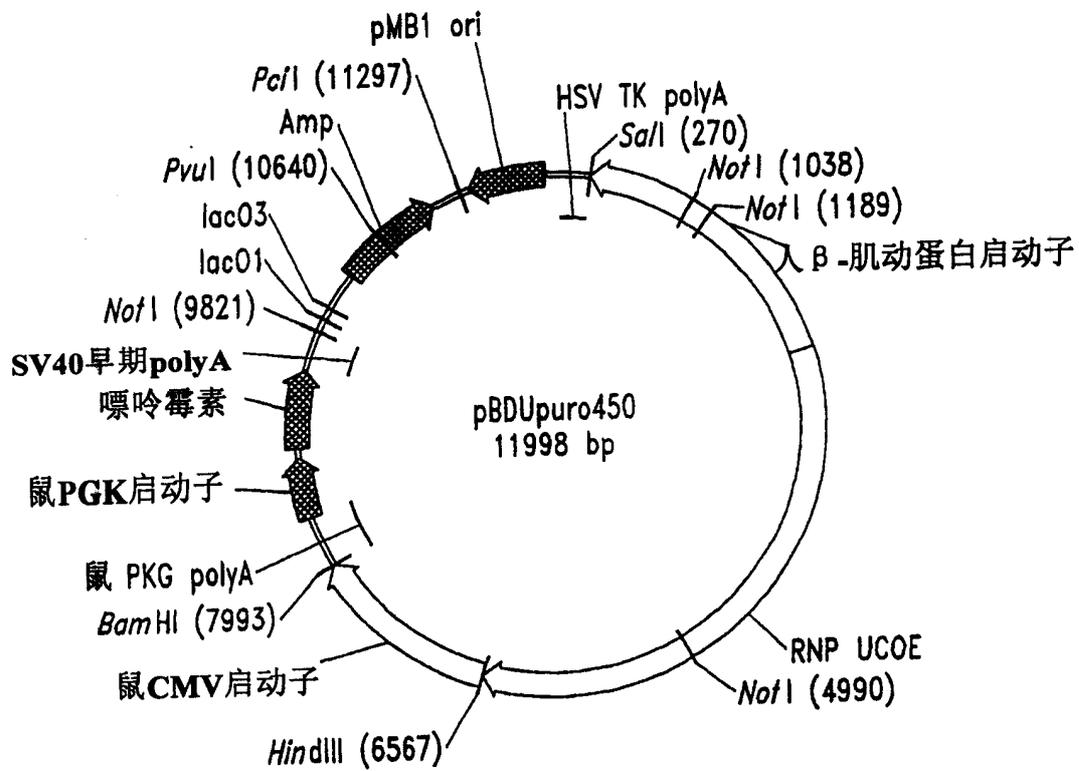


图16

图形表示载体pBDUneo1200

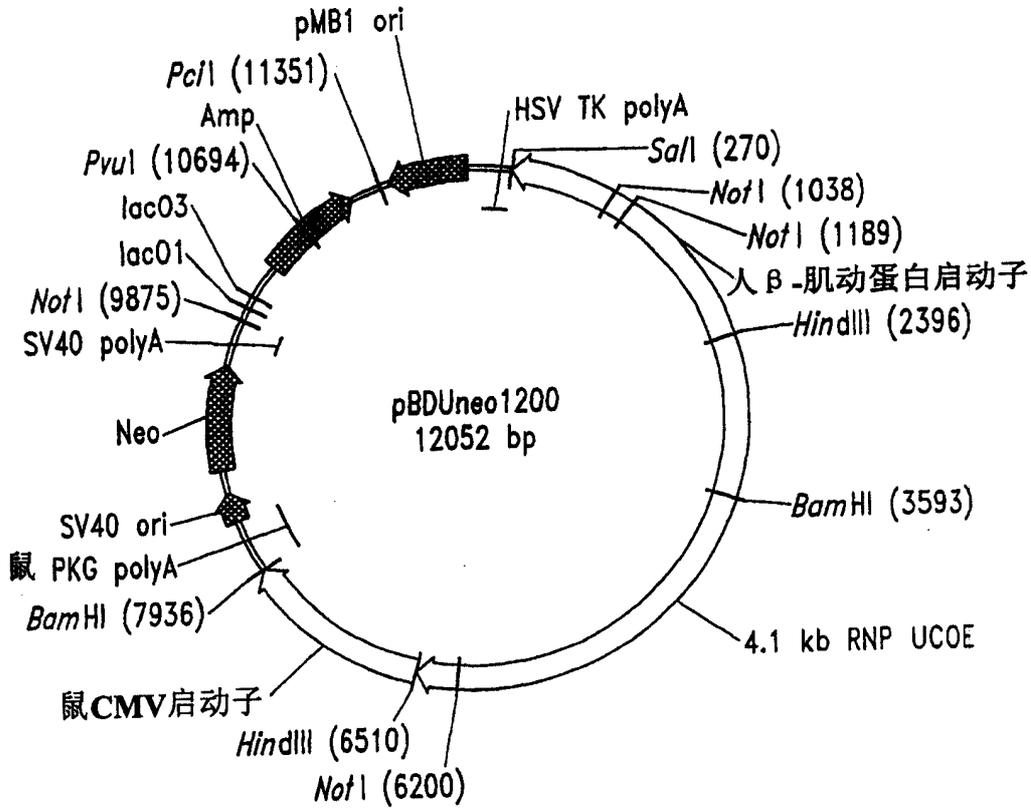


图17

图形表示载体pBDUpuro1450

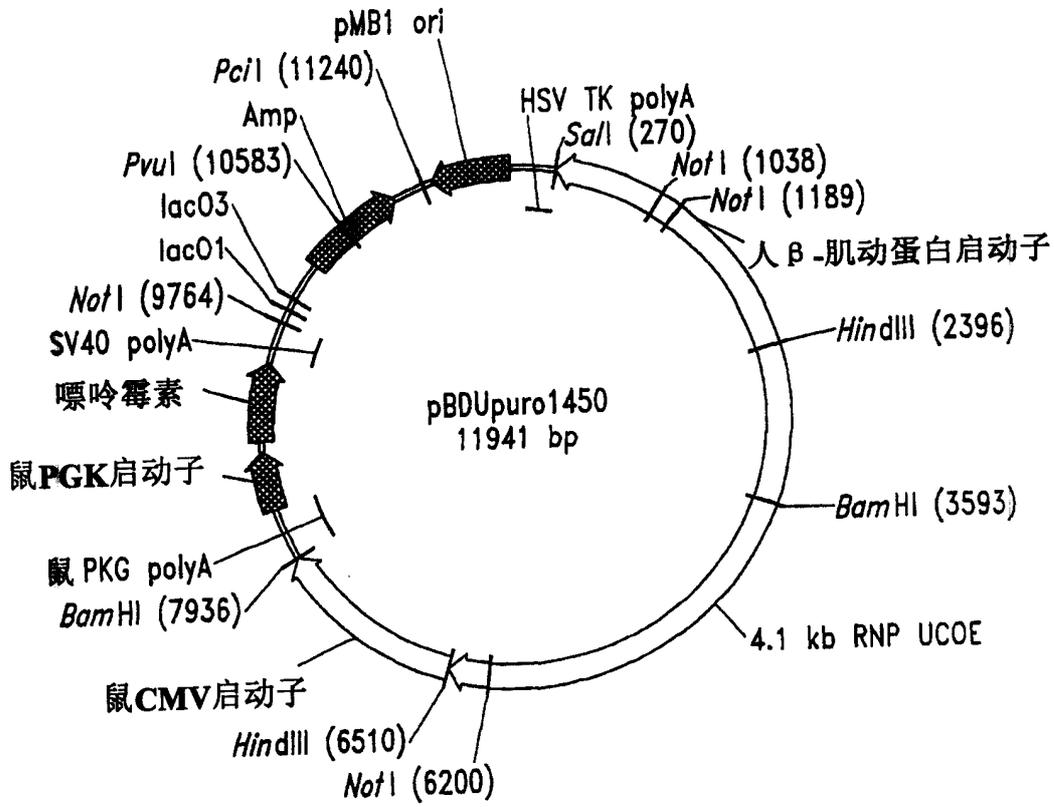


图18

图形表示载体pBDUpneo1600

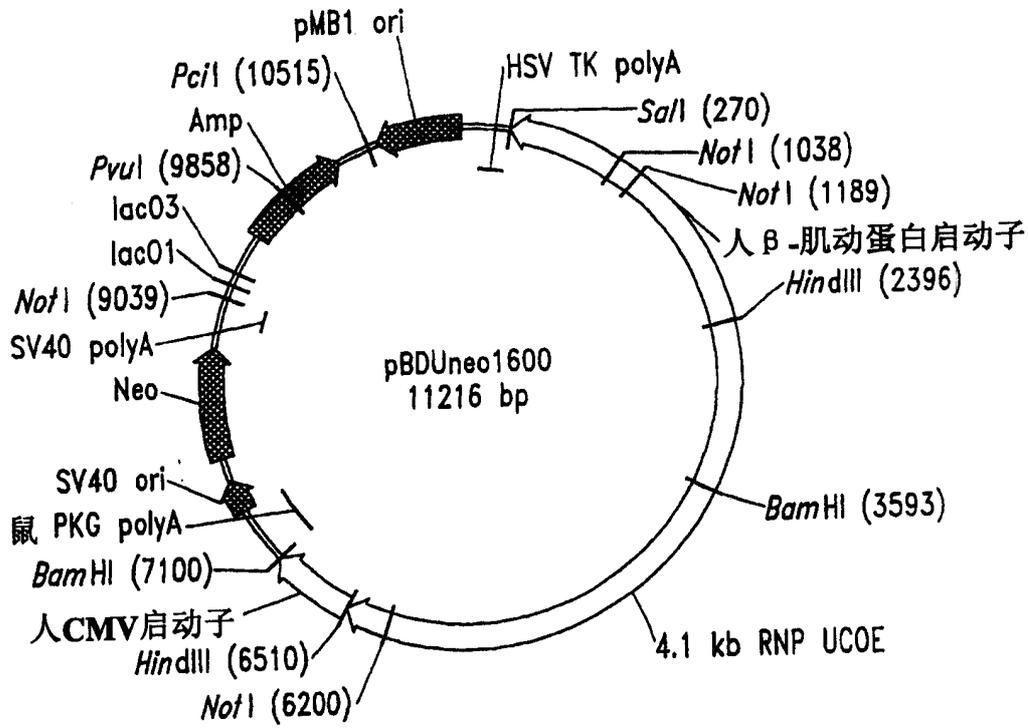


图19

图形表示载体pBDUpuro1800

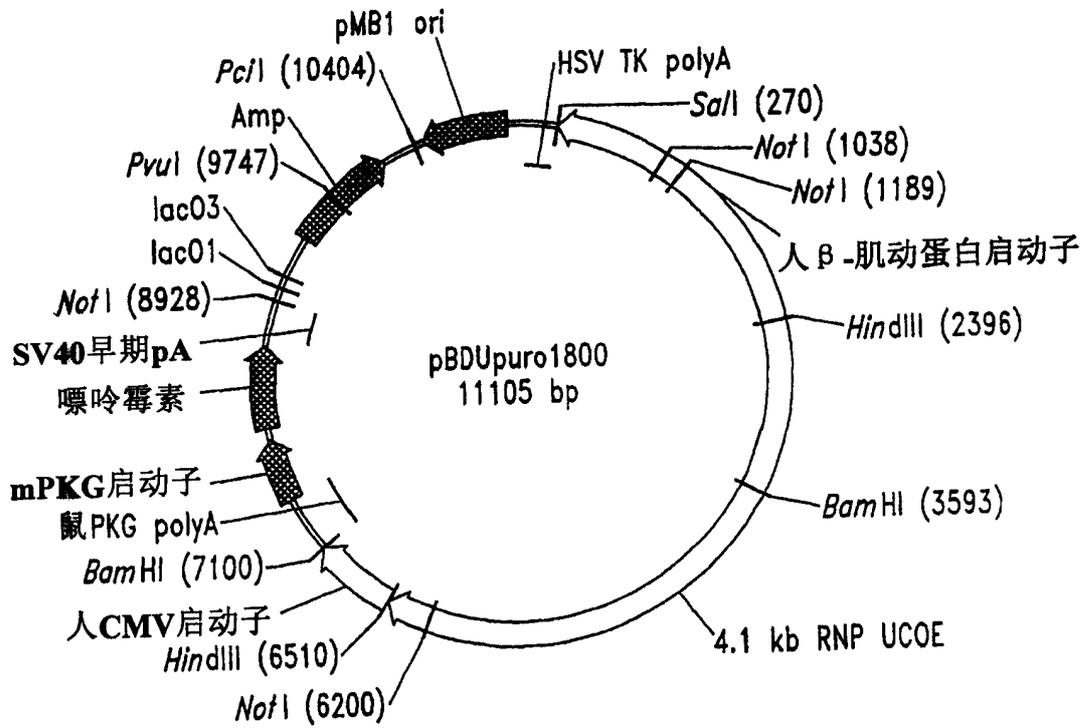


图20

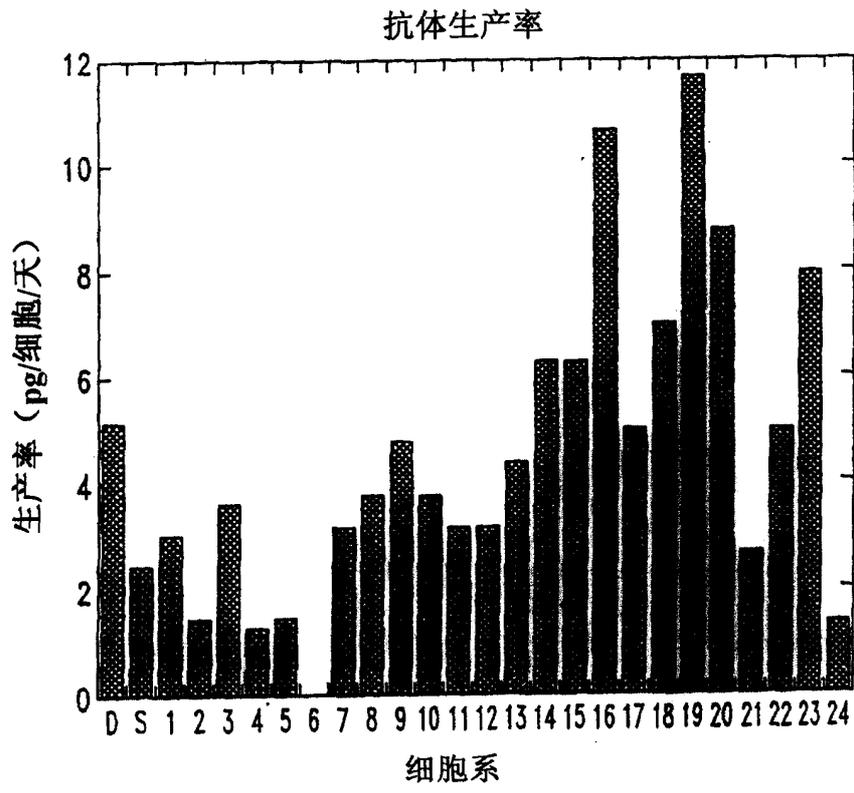


图21