

[19]中华人民共和国专利局

[51]Int.Cl<sup>6</sup>



# [12] 发明专利申请公开说明书

C12N 15/56  
C12N 9/28 C12N 1/21  
C12N 15/70

[21] 申请号 96194754.3

[43]公开日 1998年7月15日

[11] 公开号 CN 1187853A

[22]申请日 96.6.14

[30]优先权

[32]95.6.14 [33]JP[31]147257/95

[86]国际申请 PCT/JP96/01641 96.6.14

[87]国际公布 WO97/00324 英 97.1.3

[85]进入国家阶段日期 97.12.12

[71]申请人 花王株式会社

地址 日本东京

[72]发明人 秦田勇二 尾崎克也 荒胜俊

川合修次 伊藤进

[74]专利代理机构 中国国际贸易促进委员会专利商标  
事务所  
代理人 陈文平

权利要求书 1 页 说明书 18 页 附图页数 4 页

[54]发明名称 编码碱性液化 $\alpha$ -淀粉酶的基因

[57]摘要

本发明提供一种编码碱性液化 $\alpha$ -淀粉酶的DNA片段、包括该DNA片段的重组DNA、含有该重组DNA的转化的微生物和一种利用转化子生产碱性液化 $\alpha$ -淀粉酶的方法。本发明的方法能够大规模生产作为一种洗涤剂的成分有用的碱性液化 $\alpha$ -淀粉酶。

## 权 利 要 求 书

---

1. 一种编码碱性液化 $\alpha$ -淀粉酶活性的 DNA 分子。
2. 一种在权利要求 1 中所定义的 DNA 分子，它编码在序列号 1 中所描述的氨基酸序列或其功能片段。
3. 一种 DNA 分子，它编码一种表现碱性液化 $\alpha$ -淀粉酶活性的蛋白质并具有一个在序列号 1 中所描述的其中一个或多个氨基酸被置换、增加、缺失、倒位或插入的氨基酸序列。
4. 一种在权利要求 1 到 3 中的任何一项中所定义的 DNA 分子，进一步包括一个调节基因表达的核苷酸序列。
5. 一种重组 DNA，它含有权利要求 1 到 4 中任何一项的 DNA 分子。
6. 一种含有权利要求 5 的重组 DNA 的转化的微生物。
7. 一种用于生产碱性液化 $\alpha$ -淀粉酶的方法，包括培养权利要求 6 的转化的微生物和分离由该微生物产生的碱性液化 $\alpha$ -淀粉酶。
8. 一种 DNA 分子，它与同序列号 2 的核酸序列互补的一种 DNA 序列杂交。
9. 一种由权利要求 9 的 DNA 分子编码的蛋白质。
10. 一种 DNA 分子，它与同序列号 2 中的核酸序列互补的一种 DNA 序列杂交，其中所说的 DNA 分子编码一种具有碱性液化 $\alpha$ -淀粉酶活性的蛋白质。
11. 一种由权利要求 11 的 DNA 分子编码的蛋白质。
12. 重组 DNA 质粒 pAML100。
13. 重组大肠杆菌菌株 HB101 ( pAML100 )。

# 说明书

## 编码碱性液化 $\alpha$ -淀粉酶的基因

本发明涉及编码碱性液化 $\alpha$ -淀粉酶的基因和其片段，还涉及重组 DNA 和带有该基因或该基因的片段的转化子。

长久以来 $\alpha$ -淀粉酶一直被用于多个领域。例如，在发酵业中它一直被用于谷物和马铃薯的糖化，在纺织业中用作淀粉糊去除剂，在制药业中用作助消化药，而在食品业中用于制造粘稠的麦芽糖浆。 $\alpha$ -淀粉酶是一种作用于淀粉相关多糖如直链淀粉和支链淀粉，单一地水解多糖分子中 $\alpha$ -1,4-糖苷键的酶。自从 1833 年当时 Payen 和 Persoz 首次发现这种酶开始，已经从包括细菌、真菌、植物种子和动物消化腺等许多不同的来源中获得了 $\alpha$ -淀粉酶的晶体样品或电泳上均质的样品。

本发明的发明者们最近发现当将 $\alpha$ -淀粉酶和支链淀粉酶都掺入洗碗剂和衣物洗涤剂中时，洗碗剂和衣物洗涤剂的效能可被大大提高，特别是对淀粉污渍（日本专利申请公开（kokai）2-132192号）。然而，以前在自然界中发现的大部分 $\alpha$ -淀粉酶在中性到酸性 pH 范围内表现出最大且稳定的酶切活性，但在 pH9-10 的碱性溶液中却很少发挥作用。已知只有少数的淀粉酶在碱性的 pH 范围内表现最大的活性（所谓的碱性 $\alpha$ -淀粉酶和耐碱 $\alpha$ -淀粉酶）。这些碱性 $\alpha$ -淀粉酶和耐碱 $\alpha$ -淀粉酶包括由芽孢杆菌属 sp. A-40-2 产生的一种酶〔Horikoshi, K.等,《农业与生物化学》(Agric. Biol. Chem.), 35, 1783 (1971)〕, 由芽孢杆菌属 sp. NRRL B-3881 产生的一种酶〔Boyer, E.等,《细菌学杂志》(J. Bacteriol.), 110, 992 (1972)〕, 由链霉菌属 sp.KSM-9 产生的一种酶（日本专利申请公开（kokai）61-209528号）, 由芽孢杆菌属 sp. H-167 产生的一种酶（日本专利申请公开（kokai）62-208278号）, 由嗜热碱性芽孢杆菌 sp. A3-8 产生的一种酶（日本专利申请公开（kokai）2-49584号）以及由嗜盐碱球菌属 sp. AH-36 产生的一种酶（日本专利申请公开（kokai）4-211369号）。

这里所用的，术语“碱性 $\alpha$ -淀粉酶”指最适 pH 落在碱性 pH 范围中的 $\alpha$ -淀粉酶，而术语“耐碱 $\alpha$ -淀粉酶”指最适 pH 在中性至酸性范围内但其在碱性范围内的活性与在最适 pH 时的活性类似且在碱性范围内保持其稳定性的淀粉酶。术语“中性范围”意指低于 8 且不小于 6 的 pH 范围，而“碱性”表示高于“中性范围”的 pH。

这些碱性 $\alpha$ -淀粉酶和耐碱 $\alpha$ -淀粉酶的大部分是将淀粉或淀粉相关多糖分解为葡萄糖、麦芽糖或麦芽三糖的所谓的糖化 $\alpha$ -淀粉酶。据此，虽然它们被有效地用于制糖业中，但如果将它们用作洗涤剂的酶却会产生问题。因此，仍然需要对用于洗涤剂中的表面活性剂具有抗性并以一种高度随机的方式分解淀粉或淀粉相关多糖的所谓的碱性液化 $\alpha$ -淀粉酶。本发明者们持续深入地寻找产生适于作为洗涤剂成分的碱性 $\alpha$ -淀粉酶的微生物，并发现其生长最适 pH 在碱性范围内的嗜碱芽孢杆菌 sp. KSM-AP1378 菌株产生一种表现碱性液化 $\alpha$ -淀粉酶活性的酶。他们证明这种酶作为用于洗碗和厨房用具的洗涤剂组合物以及用于洗衣的洗涤剂组合物中的一种添加剂是有效的 (WO94/26881)。

通过改善用于产碱性 $\alpha$ -淀粉酶微生物芽孢杆菌 sp. KSM-AP1378 的培养方法或通过利用突变可以有效增加所产生的酶的量。然而，为了以一种工业化的规模有利地生产这种酶，必需采取其它方法。

利用一种基因工程的方法可以提高所生产的酶的量，另外，利用一种基因工程方法通过改变编码酶的基因提高酶的催化活性。然而，仍未获得编码一种碱性液化 $\alpha$ -淀粉酶的基因。

相应地，本发明的一个目的是提供编码碱性液化 $\alpha$ -淀粉酶的基因和其片段、一种含有包括该基因的重组 DNA 的转化子和一种用于利用该转化子生产碱性液化 $\alpha$ -淀粉酶的方法。

编码碱性液化 $\alpha$ -淀粉酶基因的 DNA 进一步可被用来制备用于从其它微生物中分离其它的同源碱性液化 $\alpha$ -淀粉酶基因的探针。因此，本发明的另一目的是提供一种筛选和分离其它碱性液化 $\alpha$ -淀粉酶的方法。

本发明者们试图从一种嗜碱芽孢杆菌菌株的染色体 DNA 中分离一种含

有编码碱性液化 $\alpha$ -淀粉酶基因的 DNA 片段，结果，他们成功分离了一个大约 1.8 kb 的编码一种碱性液化 $\alpha$ -淀粉酶的 DNA 片段。当他们利用被连至一合适载体的该 DNA 片段转化一种宿主微生物时，证实了所获的重组微生物产生一种碱性液化 $\alpha$ -淀粉酶。此外，发现所编码的碱性液化 $\alpha$ -淀粉酶的氨基酸序列与以前已知的淀粉酶不同。基于这个发现完成了本发明。

相应地，本发明提供编码一种碱性液化 $\alpha$ -淀粉酶的 DNA 片段。

本发明还提供一种含有上述的编码碱性液化 $\alpha$ -淀粉酶的 DNA 片段的重组 DNA。

本发明还提供一种包含上述的含有编码碱性液化 $\alpha$ -淀粉酶的 DNA 片段的重组 DNA 的转化的微生物。

本发明进一步提供一种用于通过培养上述的转化的微生物产生一种碱性液化 $\alpha$ -淀粉酶并收集酶的方法。

图 1 表现编码一种碱性液化淀粉酶的基因片段的限制性酶图谱。

图 2 是一个描述利用编码一种碱性液化淀粉酶的基因片段构建 pAML100 的图。

图 3 表现所用引物的核苷酸序列。

图 4 是一个由芽孢杆菌 sp. KSM-AP1378 产生的一种碱性液化 $\alpha$ -淀粉酶的 pH 曲线。

在本发明中，用作一种碱性液化 $\alpha$ -淀粉酶基因供体的一个有效的微生物例如可以是芽孢杆菌 sp. KSM-AP1378 (FERM BP-3048, 1989 年 7 月 24 日寄存于发酵研究所 (Fermentation Research Institute), 工业科学和技术代理处 (Agency of Industrial Science and Technology) 1-3, Higashi 1-chome, Tsukuba-shi, Ibaraki, 305 日本), 它是一种嗜碱芽孢杆菌菌株。这种菌株是由本发明者在日本 Tochigi 县的 Tochigi 城附近从土壤中分离到并鉴定为一种产生显著量的碱性液化 $\alpha$ -淀粉酶的菌株。该菌株在 1990 年 8 月 8 日以 FERM BP-3048 (最初于 1989 年 7 月 24 日保存为 P-10886) 保存在发酵研究所, 工业科学和技术代理处 (1-3,

Higashi 1-chome, Tsukuba-shi, Ibaraki-ken, 305, 日本)。

为了从供体微生物中获得染色体 DNA, 可利用由 Marmur, J. (《分子生物学杂志》(J. Mol. Biol.), 3, 208 (1961)) 建议的方法或由 Saito, H. 和 Miura, K. (《生物化学与生物物理学学报》(Biochem. Biophys. Acta), 72, 619 (1963)) 建议的方法。也可使用其它类似的方法。

通过利用限制性酶切割所获的染色体 DNA 制备含有碱性液化 $\alpha$ -淀粉酶基因的 DNA 片段。对于可使用的限制性酶只要不切割该基因则并无特殊限制。也可通过 PCR 获得碱性液化 $\alpha$ -淀粉酶的基因 (Mullis, K.B. 和 Faloona, F.A., 《酶学方法》(Methods Enzymol.), 155, 355 (1987); Saiki, R.K. 等, 《科学》(Science), 239, 487 (1988))。例如, 根据在序列号 2 中描述的核苷酸序列通过合成具有与那些在核心区域的 5'-末端的上游和 3'-末端的下游序列相应的序列的引物, 随后通过将产一种碱性液化 $\alpha$ -淀粉酶的微生物的染色体 DNA 用作模板进行 PCR 可以获得基因。另外, 可通过利用任何方法从产一种碱性液化 $\alpha$ -淀粉酶的微生物中首先获得碱性液化 $\alpha$ -淀粉酶基因的一个片段, 随后通过 PCR 扩增该片段基因的上游和下游区。

然后将所制备的基因片段进行克隆。只要宿主细菌表达本发明的碱性液化 $\alpha$ -淀粉酶基因, 重组 DNA 分子在宿主菌中可被复制并且整合的基因可被稳定地保留, 则对可利用的宿主/载体系统并无特殊限制。例如, 宿主是大肠杆菌 K-12 的 EK 系统的成员和宿主为 Marburg 枯草芽孢杆菌的 BS 系统的成员都可以使用。利用包含许多种载体并且遗传学上被深入地研究过的 EK 系统可获得良好的结果, 因而是优选的。宿主细菌的特定的例子包括 EK 系统的 HB101、C600 和 JM109, 以及 BS 系统的 BD170、MI112 和 ISW1214。载体的特定的例子包括用于 EK 系统 pBR322 和 pUC18, 以及用于 BS 系统的 pUB110 和 pHY300PLK。

通过用一限制性酶切割一种载体随后与上述的染色体或 PCR 扩增的 DNA 片段连接产生一种重组质粒 DNA 分子。例如通过利用一种 DNA 连接酶可以实现连接。

对利用一重组 DNA 分子转化宿主细菌的方法并无特殊的限制。例如，在 EK 系统宿主的情况下可用氯化钙法（Mandel, M.和 Higa, A., 《分子生物学杂志》，53, 159 (1970)），而在 BS 系统宿主的情况下可用原生质体法（Chang, C.和 Cohen, S.N., 《分子遗传学与普通遗传学》（Mol. Gen. Genet.），168, 111 (1978)）。按照下面的方法进行重组子微生物的筛选。首先，利用一种未由外源染色体 DNA 片段的插入而失活的特性作为一种指标如载体 DNA 上编码的对抗生素的抗性，来筛选用含有载体衍生的 DNA 片段转化的微生物。例如，在一利用 EK 系统的 pBR322 作为载体并将染色体 DNA 的一个 HindIII 片段插入 pBR322 的 HindIII 切割位点中的特定例子中，四环素抗性基因失活，所以可以通过在氨苄青霉素基因中不带有 HindIII 切割位点而赋予氨苄青霉素抗性的转化子的生长进行初次筛选。随后，将所选的转化子利用如影印方法转移至含有淀粉的琼脂平板上，然后进行培养以便形成菌落。因为靶重组子微生物在集落周围分解淀粉，所以通过利用一种含碘溶液对在含有淀粉的琼脂平板中所含的淀粉染色可以对其进行筛选。

利用制备质粒或噬菌体 DNA 的标准程序可以提取由所获的重组微生物包含的重组 DNA 分子（Maniatis, T.等, 《分子克隆》（Molecular Cloning），冷泉港实验室，纽约（1982））。在利用电泳分析通过各种限制性酶所获得的切割模式时，证实重组 DNA 分子是载体 DNA 分子和一个含有碱性液化 $\alpha$ -淀粉酶的 DNA 片段的连接产物。

在图 1 的限制性酶切图谱中所表现的一个大约 2.1kb 的 DNA 片段内含有编码一种碱性液化 $\alpha$ -淀粉酶的基因，并存在于由白色带表示的大约 1.6kb 的片段中。该基因具有一个显示在序列号 2 中的核苷酸序列。在此序列中，5'末端和 3'末端分别对应于显示在序列 2 中的大约 2.1kb 片段的左手侧和右手侧。在此序列中观察到一个在第 145 个核苷酸 ATG 处开始并编码一个由序列号 1 中所描述的 516 个氨基酸残基组成的序列的开放阅读框架（ORF）。ORF 上游 13 个碱基（13b）处存在一个与枯草芽孢杆菌 16S 核糖体 RNA 的 3'末端序列高度互补的 AAGGAG 序列（McLaughlin, J.R.等,《生物学和化学杂志》（J. Biol. Chem.），256,

11283 (1981))。在核苷酸由 9 延伸至 36 的更加上游的区域中, 存在一个与一种  $\sigma^A$  - 型启动子的共有序列具有高度同源性的 TTGAAA ..... 16b ..... TATGGT 序列 (Gitt, M.A.等, 《生物学与化学杂志》, 260, 7178 (1985))。类似地, 在从 95 到 125 的核苷酸中发现另一个  $\sigma^A$  - 型启动子序列。从芽孢杆菌 sp. KSM-AP1378 培养物纯化的碱性液化  $\alpha$  - 淀粉酶中氨基末端侧的 10 个氨基酸残基的氨基酸序列与由该 DNA 片段的核苷酸序列所推导的第 37 个氨基酸延伸的序列 (序列号 2 中的 37 - 46 号氨基酸) 一致。

当将本发明中基因的核苷酸序列和所推导的氨基酸序列与那些目前已知的  $\alpha$  - 淀粉酶相比较时, 证实该基因包括一种新型核苷酸序列, 所推导的由该基因编码的氨基酸序列与那些其它  $\alpha$  - 淀粉酶如由解淀粉芽孢杆菌产生的一种液化  $\alpha$  - 淀粉酶 (Takinen, K.等, 《生物学与化学杂志》, 258, 1007 (1983))、由嗜热脂肪芽孢杆菌产生的一种液化  $\alpha$  - 淀粉酶 (Nakajima, R.等, 《细菌学杂志》 (J. Bacteriol.), 163, 401 (1985))、由地衣芽孢杆菌产生的一种液化  $\alpha$  - 淀粉酶 (Yuuki, T.等, 《生物化学杂志》 (J. Biochem.), 98, 1147 (1985)) 或由芽孢杆菌 sp. 707 产生的一种液化  $\alpha$  - 淀粉酶 (Tsukamoto, A.等, 《生物化学与生物物理学研究通讯》 (Biochem. Biophys. Res. Commun.), 151, 25 (1988)) 不同。

含有碱性液化  $\alpha$  - 淀粉酶基因完整区域的一个优选的重组 DNA 分子的例子是质粒 pAML100 (图 2)。这个重组质粒大小为 4.4kb 并形成含有包括碱性液化  $\alpha$  - 淀粉酶基因和 pUC19 的一个 1.8kb 片段的片段。包含该重组 DNA 分子的一个优选的重组微生物是一种大肠杆菌 HB101 (pAML100) 菌株。通过用重组质粒 pAML100 以一种标准的转化方法转化大肠杆菌 HB101 获得该菌株。当用一种常规用来培养大肠杆菌的培养基培养该菌株时, 它产生一种碱性液化  $\alpha$ -淀粉酶。所产生的酶的最适反应 pH 为 pH 8-9。这与对由该基因的供体细菌菌株芽孢杆菌 sp. KSM-AP1378 产生的碱性液化  $\alpha$ -淀粉酶所确定的活性-pH 关系曲线 (图 4) 非常一致。

本发明的 DNA 片段, 只要它们编码一种表现所需酶切活性的蛋白质,



则不一定仅限于那些编码在下述的序列表中所示氨基酸序列的 DNA，这包括编码一种其中一个或多个氨基酸被置换、增加、缺失、倒位或插入的氨基酸序列的 DNA 片段。这种 DNA 的一个例子是一个编码一种等价于在序列号 1 中所描述的氨基酸序列而 N-末端有多至 32 个氨基酸已缺失的氨基酸序列的 DNA。

为了利用本发明中被转化的微生物生产一种碱性液化 $\alpha$ -淀粉酶，将一种被转化的含有本发明中前面提到的 DNA 片段的微生物进行培养。另外，可将该 DNA 片段整合入许多表达载体来获得具有提高了表达能力的被转化的微生物，随后培养所获的转化子。此外，根据微生物的特性可在不同的条件下培养被转化的微生物。因此，可以使用适于宿主的培养条件。为了从所获的培养物中收集一种碱性液化 $\alpha$ -淀粉酶，可使用一种常规的方法（如在 WO94/26881 中所描述的方法）。

可以进一步将本发明中的 DNA 片段用作探针用于从其它生物中分离同源的碱性液化 $\alpha$ -淀粉酶基因。

## 实施例

下面将要通过实施例更详细地描述本发明，这些实施例不应被认为是将本发明限定于此。在实施例中的浓度都是以重量的%为基础。

### 实施例 1：

将产生一种碱性液化 $\alpha$ -淀粉酶的芽孢杆菌 sp. KSM-AP1378 接种在 5ml 培养基 A（表 1）中并在 30℃ 进行振荡培养 24 小时。将 1ml 培养物接种于 100ml 同样的培养基中，然后在 30℃ 再次振荡培养 12 小时。随后，通过离心收集细胞并按照由 Saito 和 Miura 推荐的方法（Saito, H. 和 Miura, K., 《生物化学与生物物理学学报》（Biochem. Biophys. Acta），72, 619（1963））获得大约 1mg 的染色体 DNA。

表 1

培养基 A 的组成

可溶性淀粉	1.0%
多聚蛋白胨	1.0%
酵母提取物	0.5%
$\text{KH}_2\text{PO}_4$	0.1%
$\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$	0.25%
$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	0.02%
$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$	0.02%
$\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$	0.001%
$\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$	0.0001%
$\text{Na}_2\text{CO}_3$	1.0% (单独灭菌)

### 实施例 2:

已知淀粉酶家族的许多成员具有氨基酸序列高度保守的 I - IV 区 ( Nakajima, R.等, 《应用微生物学与生物技术》 ( Appl. Microbiol. Biotechnol. ), 23, 355 ( 1986 ) )。因而, 根据在已知的碱性液化 $\alpha$ -淀粉酶的 I 到 IV 区中非常保守的 II 区和 IV 区的氨基酸序列合成与 II 区和 IV 区相应的引物 1 和 2 (图 1 和 3)。利用所合成的引物和 KSM-AP1378 的染色体 DNA (用作模板) 进行 PCR ( 1 个循环 =  $94\text{ }^\circ\text{C} \times 1\text{min} + 42\text{ }^\circ\text{C} \times 1\text{min} + 60\text{ }^\circ\text{C} \times 2\text{min}$ , 30 个循环)。获得了在图 1 中所示的一个大约 0.3kb 的基因片段 (片段 A), 并测定了该片段的核苷酸序列。结果, 发现该片段编码一种表现出与已知液化淀粉酶的从 II 区延伸到 IV 区的氨基酸序列具有一种不可忽视水平同源性的氨基酸序列。

### 实施例 3:

利用片段 A 作为一个探针, 对经 XbaI 消化的 KSM-AP1378 的染色体 DNA 进行 Southern 杂交。结果, 证实在大约 1.0kb 位置处杂交有一条带。利用从片段 A 的末端序列合成的引物 (在 II 区侧: 引物 3; 在 IV 区侧:

引物 4) 和通过分子内连接 XbaI 消化的 KSM-AP1378 的染色体 DNA (图 1) 而获得的 DNA 作为模板, 通过一种反向 PCR 方法 (Triglia, T.等, 《核酸研究》 (Nucleic Acids Res.), 16, 81 (1988)) 获得一个大约 0.7kb 的扩增片段 (片段 B)。对片段 B 测定其核苷酸序列, 其序列揭示该片段含有一段 IV 区下游的大约 0.6kb 区域的序列。该片段含有一个 ORF 的终止密码子, 推测其归因于碱性液化 $\alpha$ -淀粉酶。

#### 实施例 4:

基于来自 KSM-AP1378 菌株的碱性液化 $\alpha$ -淀粉酶的 N-末端氨基酸序列 (7 个氨基酸) (图 3) 设计并合成一个引物。利用所得的引物 (引物 5) 联合前面所述的引物 3 (图 3) 并以 KSM-AP1378 的染色体 DNA 作为模板, 进行 PCR 获得一个大约 0.7kb 的片段 (片段 C, 图 1), 然后测定其核苷酸序列。

#### 实施例 5:

合成一个含有 21 个碱基、在编码所纯化的酶 N-末端氨基酸序列的核苷酸序列的下游直接延伸的引物 (引物 6)。利用引物 6 和 7 (图 1 和 3) 并以通过分子内连接 HindIII 消化的 KSM-AP1378 染色体 DNA 所获的 DNA (图 1) 为模板, 进行反向 PCR, 获得一个其中在 HindIII 位点将一个上游的 0.8kb 片段 (片段 D) 和一个下游的 PstI-HindIII 0.4kb 片段连接起来的 1.2kb 片段。对片段 D 区的核苷酸序列进行测定, 揭示了一个由 31 个氨基酸组成的信号序列 MKLHNRIISVLLTLLAVAVLFPYMTEPAQA (从序列号 2 的第 1 到第 31 位)、一个推导的由 AAGGAG 组成的 SD 序列 (核苷酸 127 - 132; McLaughlin, J.R.等, 《生物学与化学杂志》, 260, 7178 (1985)) 和两种推导的启动子序列 (-35 序列, TTGAAA; -10 序列, TATGGT 和 -35 序列, TTGACT; -10 序列, TAAATT) 的存在。

#### 实施例 6:

利用定位于启动子序列上游大约 0.1kb 处的引物 A、定位于中止密码子下游 79b 处的引物 B 并以 KSM-AP1378 的染色体 DNA 作为模板, 通过 PCR 扩增在引物间的大约 1.8kb 的片段。将所获的扩增片段插入 pUC19 的 SmaI 位点, 然后引入大肠杆菌 HB101。转化子可以在含有 0.4% 的淀粉天青和 15ug/ml 氨基青霉素的 LB 琼脂培养基上生长。分离在周围形成透明晕环的集落作为产生液化 $\alpha$ -淀粉酶的大肠杆菌菌株。从该转化子中制备重组质粒, 并制备质粒的限制性酶切图谱。在图谱中, 证实含有一个在图 1 中所示的大约 1.8kb 的 DNA 片段(片段 E)。这个重组质粒被定名为质粒 pAML100 (图 2)。

#### 实施例 7:

将在实施例 6 中获得的重组大肠杆菌于 5ml 含有 50ug/ml 氨基青霉素的 LB 液体培养基中振荡培养 12 小时。将 1ml 培养物接种于 100ml LB 培养基(含有氨基青霉素)中, 然后在 37℃ 振荡培养 24 小时。将通过离心分离收集的细胞悬于 Tris-HCl 缓冲液(pH8.0)中, 并通过超声处理破坏细胞。细胞被超声处理后, 通过离心分离清除细胞碎片, 并将所获的上清用作一种无细胞提取物。作为对照, 以一种相似的方式分别制备 HB101 (pUC19) 菌株的无细胞提取物。首先通过在一含有 50mM 甘氨酸-NaCl-NaOH 缓冲液(pH10)和可溶性淀粉的反应混合物中于 50℃ 反应 15 分钟, 然后通过利用 3,5-二硝基水杨酸方法定量测定所产生的还原糖来检测这些提取物中的 $\alpha$ -淀粉酶活性(WO94/26881)。将 1 个单位的酶活性定义为每分钟产生还原糖的量等价于 1 $\mu$ mol 葡萄糖的蛋白质的量。因而, 在菌株 HB101 (pAML100) 的无细胞提取物中检测 $\alpha$ -淀粉酶的活性。发现 $\alpha$ -淀粉酶的最适工作 pH 落在 8 到 9 之间 pH 范围中。这个结果与由芽孢杆菌 sp. KSM-AP1378 产生的液化 $\alpha$ -淀粉酶的最适 pH 非常一致(图 4)。对于酶活性的测定, 利用下面表 2 中所示的缓冲液(各为 40mM)。

表 2

---

pH 3.5-5.5 :	醋酸盐缓冲液
pH 5.5-8.5 :	Tris-马来酸缓冲液
pH 8.5-10.5 :	甘氨酸-NaCl- NaOH 缓冲液
pH10.5-11.0 :	Na <sub>2</sub> CO <sub>3</sub> -NaHCO <sub>3</sub> 缓冲液

---

根据本发明, 获得一个编码在碱性 pH 范围内表现最大活性的碱性液化 α - 淀粉酶的基因和含有这种基因的微生物是可能的。利用它们便于碱性液化 α - 淀粉酶的大量生产。

序列表

序列号 1 的资料:

(i) 序列特征:

(A) 长度: 516 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑结构: 线性

(ii) 分子类型: 肽

(xi) 序列描述: 序列号 1:

```

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

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

Asp Val Pro Leu His Tyr Asn Leu Tyr Asn Ala Ser Asn Ser Gly Gly  
                   325                                  330                                  335  
 Tyr Phe Asp Met Arg Asn Ile Leu Asn Gly Ser Val Val Gln Lys His  
                   340                                  345                                  350  
 Pro Ile His Ala Val Thr Phe Val Asp Asn His Asp Ser Gln Pro Gly  
                   355                                  360                                  365  
 Glu Ala Leu Glu Ser Phe Val Gln Ser Trp Phe Lys Pro Leu Ala Tyr  
                   370                                  375                                  380  
 Ala Leu Ile Leu Thr Arg Glu Gln Gly Tyr Pro Ser Val Phe Tyr Gly  
 385                                  390                                  395                                  400  
 Asp Tyr Tyr Gly Ile Pro Thr His Gly Val Pro Ser Met Lys Ser Lys  
                   405                                  410                                  415  
 Ile Asp Pro Leu Leu Gln Ala Arg Gln Thr Tyr Ala Tyr Gly Thr Gln  
                   420                                  425                                  430  
 His Asp Tyr Phe Asp His His Asp Ile Ile Gly Trp Thr Arg Glu Gly  
                   435                                  440                                  445  
 Asp Ser Ser His Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asp Gly  
                   450                                  455                                  460  
 Pro Gly Gly Asn Lys Trp Met Tyr Val Gly Lys His Lys Ala Gly Gln  
 465                                  470                                  475                                  480  
 Val Trp Arg Asp Ile Thr Gly Asn Arg Ser Gly Thr Val Thr Ile Asn  
                   485                                  490                                  495  
 Ala Asp Gly Trp Gly Asn Phe Thr Val Asn Gly Gly Ala Val Ser Val  
                   500                                  505                                  510  
 Trp Val Lys Gln  
                   516

**序列号 2 的资料:**

**(i) 序列特征:**

**(A) 长度: 1776 个碱基对**

**(B) 类型: 核酸**

**(C) 成链类型: 双链**

**(D) 拓扑结构: 线性**

**(ii) 分子类型: DNA (基因组)**

**(vi) 来源:**

**(A) 生物体: 芽孢杆菌属**

**(B) 菌株: KSM-AP1378**

**(xi) 序列描述: 序列号 2:**



ATATAAAATTT GAAATGAAACA CCTATGAAAA TATGGTAGCG ATTGCGCGAC GAGAAAAAAC 60  
 TTGGGAGTTA GGAAGTGATA TTAAGGATT TTTTGTGACT TCTTGTGAAA ACGCTTGCAT 120  
 AAATTGAAGG AGAGGGTGCT TTTT ATG AAA CTT CAT AAC CGT ATA ATT AGC GTA 174  
 Met Lys Leu His Asn Arg Ile Ile Ser Val  
 1 5 10  
 CTA TTA ACA-CTA TTG TTA GCT GTA GCT GTT TTG TTT CCA TAT ATG ACG 222  
 Leu Leu Thr Leu Leu Leu Ala Val Ala Val Leu Phe Pro Tyr Met Thr  
 15 20 25  
 GAA CCA GCA CAA GCC CAT CAT AAT GGG ACG AAT GGG ACC ATG ATG CAG 270  
 Glu Pro Ala Gln Ala His His Asn Gly Thr Asn Gly Thr Met Met Gln  
 30 35 40  
 TAT TTT GAA TGG CAT TTG CCA AAT GAC GGG AAC CAC TGG AAC AGG TTA 318  
 Tyr Phe Glu Trp His Leu Pro Asn Asp Gly Asn His Trp Asn Arg Leu  
 45 50 55  
 CGA GAT GAC GCA GCT AAC TTA AAG AGT AAA GGG ATT ACC GCT GTT TGG 366  
 Arg Asp Asp Ala Ala Asn Leu Lys Ser Lys Gly Ile Thr Ala Val Trp  
 60 65 70  
 ATT CCT CCT GCA TGG AAG GGG ACT TCG CAA AAT GAT GTT GGG TAT GGT 414  
 Ile Pro Pro Ala Trp Lys Gly Thr Ser Gln Asn Asp Val Gly Tyr Gly  
 75 80 85 90  
 GCC TAT GAT TTG TAC GAT CTT GGT GAG TTT AAC CAA AAG GGA ACC GTC 462  
 Ala Tyr Asp Leu Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val  
 95 100 105  
 CGT ACA AAA TAT GGC ACA AGG AGT CAG TTG CAA GGT GCC GTG ACA TCT 510  
 Arg Thr Lys Tyr Gly Thr Arg Ser Gln Leu Gln Gly Ala Val Thr Ser  
 110 115 120  
 TTG AAA AAT AAC GGG ATT CAA GTT TAT GGG GAT GTC GTG ATG AAT CAT 558  
 Leu Lys Asn Asn Gly Ile Gln Val Tyr Gly Asp Val Val Met Asn His  
 125 130 135

AAA GGT GGA GCA GAC GGG ACA GAG ATG GTA AAT GCG GTG GAA GTG AAC	606
Lys Gly Gly Ala Asp Gly Thr Glu Met Val Asn Ala Val Glu Val Asn	
140 145 150	
CGA AGC AAC CGA AAC CAA GAA ATA TCA GGT GAA TAC ACC ATT GAA GCA	654
Arg Ser Asn Arg Asn Gln Glu Ile Ser Gly Glu Tyr Thr Ile Glu Ala	
155 160 165 170	
TGG ACG AAA TTT GAT TTC CCT GGA AGA GGA AAT ACC CAT TCC AAC TTT	702
Trp Thr Lys Phe Asp Phe Pro Gly Arg Gly Asn Thr His Ser Asn Phe	
175 180 185	
AAA TGG CGC TGG TAT CAT TTT GAT GGG ACA GAT TGG GAT CAG TCA CGT	750
Lys Trp Arg Trp Tyr His Phe Asp Gly Thr Asp Trp Asp Gln Ser Arg	
190 195 200	
CAG CTT CAG AAC AAA ATA TAT AAA TTC AGA GGT ACC GGA AAG GCA TGG	798
Gln Leu Gln Asn Lys Ile Tyr Lys Phe Arg Gly Thr Gly Lys Ala Trp	
205 210 215	
GAC TGG GAA GTA GAT ATA GAG AAC GGC AAC TAT GAT TAC CTT ATG TAT	846
Asp Trp Glu Val Asp Ile Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr	
220 225 230	
GCA GAC ATT GAT ATG GAT CAT CCA GAA GTA ATC AAT GAA CTT AGA AAT	894
Ala Asp Ile Asp Met Asp His Pro Glu Val Ile Asn Glu Leu Arg Asn	
235 240 245 250	
TGG GGA GTT TGG TAT ACA AAT ACA CTT AAT CTA GAT GGA TTT AGA ATC	942
Trp Gly Val Trp Tyr Thr Asn Thr Leu Asn Leu Asp Gly Phe Arg Ile	
255 260 265	

GAT GCT GTG AAA CAT ATT AAA TAC AGC TAT ACG AGA GAT TGG CTA ACA	990
Asp Ala Val Lys His Ile Lys Tyr Ser Tyr Thr Arg Asp Trp Leu Thr	
270 275 280	
CAT GTG CGT AAC ACC ACA GGT AAA CCA ATG TTT GCA GTT GCA GAA TTT	1038
His Val Arg Asn Thr Thr Gly Lys Pro Met Phe Ala Val Ala Glu Phe	
285 290 295	
TGG AAA AAT GAC CTT GCT GCA ATC GAA AAC TAT TTA AAT AAA ACA AGT	1086
Trp Lys Asn Asp Leu Ala Ala Ile Glu Asn Tyr Leu Asn Lys Thr Ser	
300 305 310	
TGG AAT CAC TCC GTG TTC GAT GTT CCT CTT CAT TAT AAT TTG TAC AAT	1134
Trp Asn His Ser Val Phe Asp Val Pro Leu His Tyr Asn Leu Tyr Asn	
315 320 325 330	
GCA TCT AAT AGT GGT GGC TAT TTT GAT ATG AGA AAT ATT TTA AAT GGT	1182
Ala Ser Asn Ser Gly Gly Tyr Phe Asp Met Arg Asn Ile Leu Asn Gly	
335 340 345	
TCT GTC GTA CAA AAA CAC CCT ATA CAT GCA GTC ACA TTT GTT GAT AAC	1230
Ser Val Val Gln Lys His Pro Ile His Ala Val Thr Phe Val Asp Asn	
350 355 360	
CAT GAC TCT CAG CCA GGA GAA GCA TTG GAA TCC TTT GTT CAA TCG TGG	1278
His Asp Ser Gln Pro Gly Glu Ala Leu Glu Ser Phe Val Gln Ser Trp	
365 370 375	
TTC AAA CCA CTG GCA TAT GCA TTG ATT CTG ACA AGG GAG CAA GGT TAC	1326
Phe Lys Pro Leu Ala Tyr Ala Leu Ile Leu Thr Arg Glu Gln Gly Tyr	
380 385 390	
CCT TCC GTA TTT TAC GGT GAT TAC TAC GGT ATA CCA ACT CAT GGT GTT	1374
Pro Ser Val Phe Tyr Gly Asp Tyr Tyr Gly Ile Pro Thr His Gly Val	
395 400 405 410	
CCT TCG ATG AAA TCT AAA ATT GAT CCA CTT CTG CAG GCA CGT CAA ACG	1422
Pro Ser Met Lys Ser Lys Ile Asp Pro Leu Leu Gln Ala Arg Gln Thr	



415	420	425	
TAT GCC TAC GGA ACC CAA CAT GAT TAT TTT GAT CAT CAT GAT ATT ATC			1470
Tyr Ala Tyr Gly Thr Gln His Asp Tyr Phe Asp His His Asp Ile Ile			
430	435	440	
GGC TGG ACG AGA GAA GGG GAC AGC TCC CAC CCA AAT TCA GGA CTT GCA			1518
Gly Trp Thr Arg Glu Gly Asp Ser Ser His Pro Asn Ser Gly Leu Ala			
445	450	455	
ACT ATT ATG TCC GAT GGG CCA GGG GGT AAT AAA TGG ATG TAT GTC GGG			1566
Thr Ile Met Ser Asp Gly Pro Gly Gly Asn Lys Trp Met Tyr Val Gly			
460	465	470	
AAA CAT AAA GCT GGC CAA GTA TGG AGA GAT ATC ACC GGA AAT AGG TCT			1614
Lys His Lys Ala Gly Gln Val Trp Arg Asp Ile Thr Gly Asn Arg Ser			
475	480	485	490
GGT ACC GTC ACC ATT AAT GCA GAT GGT TGG GGG AAT TTC ACT GTA AAC			1662
Gly Thr Val Thr Ile Asn Ala Asp Gly Trp Gly Asn Phe Thr Val Asn			
495	500	505	
GGA GGG GCA GTT TCG GTT TGG GTG AAG CAA TAAATAAGGA ACAAGAGGCG			1712
Gly Gly Ala Val Ser Val Trp Val Lys Gln			
510	515		
AAAATTACTT TCCTACATGC AGAGCTTTCC GATCACTCAT ACACCCAATA TAAATTGGAA			1772
GCTT			1776

说明书附图

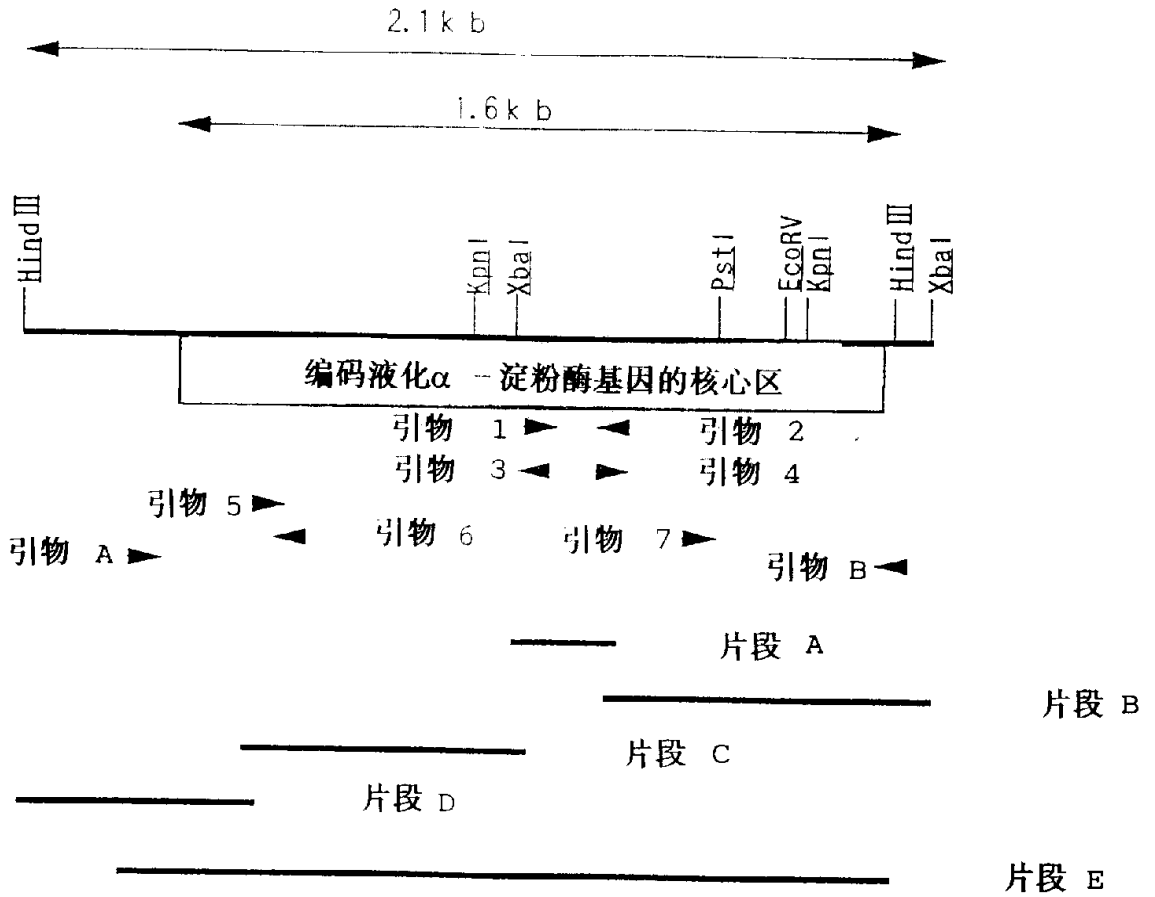


图 1

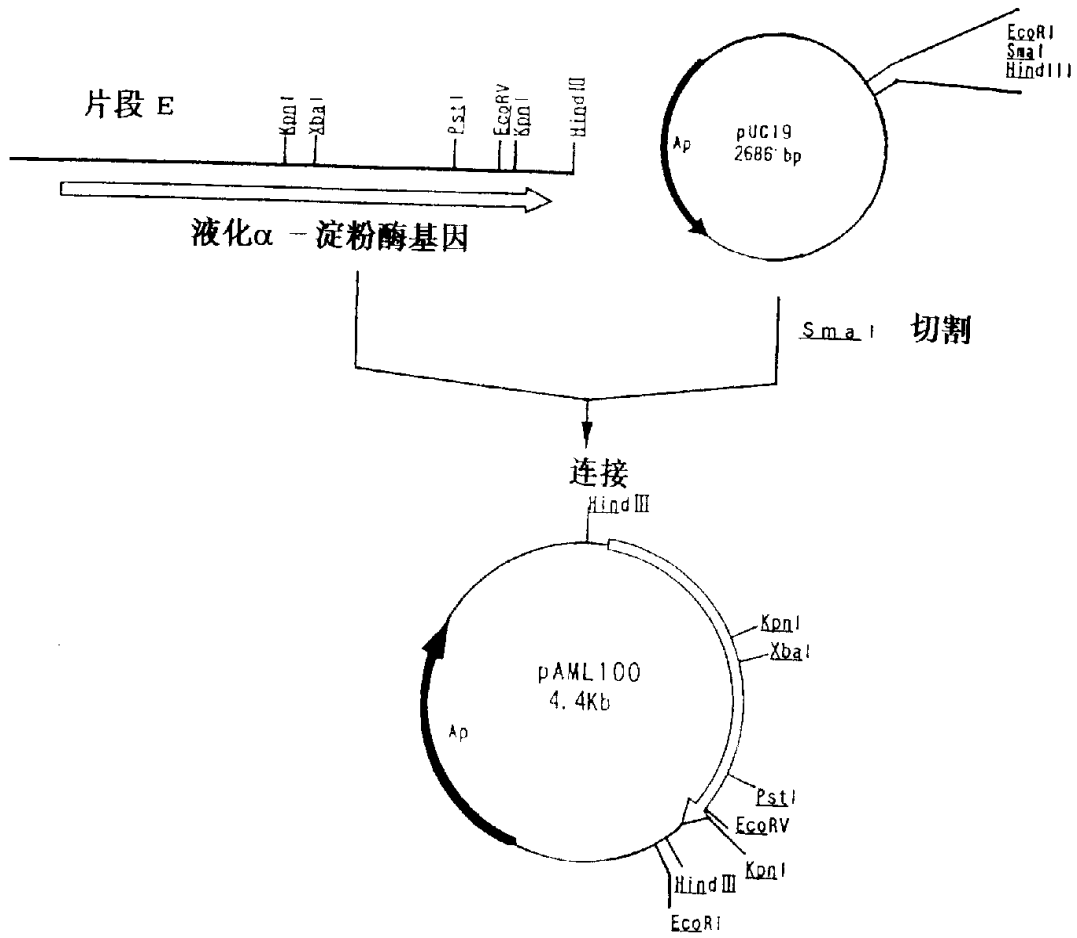


图 2

图 3

- 引物 1 5' TAGACGCAGTAAAACACATAAA 3'  
           C T C C G T C  
           G G G T  
           T T T
- 引物 2 3' CGACAATGAAAACAACACTATTAGTACT 5'  
           G G G G G G G G  
           C C C C  
           T T T T
- 引物 3 5' AGCCAATCTCTCGTATAGCTGTA 3'
- 引物 4 5' GTACAAAAACACCCTATACATG 3'
- 引物 5 5' AATGGAACAATGATGCAGTA 3'  
           T T T
- 引物 6 5' CATTGGCAAATGCCATTCAA 3'
- 引物 7 5' AAAATTGATCCACTTCTGCAG 3'
- 引物 A 5' CAGCGGTGATAATATAAATTTGAAT 3'
- 引物 B 5' AAGCTTCCAATTTATATTGGGTGTAT 3'

图 4

