

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12Q 1/68

C12N 1/14 C12N 15/00

C07H 21/02

[12] 发明专利申请公开说明书

[21] 申请号 01801513.1

[43] 公开日 2002 年 11 月 20 日

[11] 公开号 CN 1380905A

[22] 申请日 2001.4.13 [21] 申请号 01801513.1

[30] 优先权

[32]2000.4.13 [33]WO [31]PCT/US00/10199

[86] 国际申请 PCT/US01/12335 2001.4.13

[87] 国际公布 WO01/79558 英 2001.10.25

[85] 进入国家阶段日期 2002.1.29

[71] 申请人 马克·阿龙·埃马尔法尔布

地址 美国佛罗里达州

[72] 发明人 彼得·扬·蓬特 科妮莉亚·范塞尔

科尔内留斯·范登洪德

[74] 专利代理机构 中科专利商标代理有限责任公司

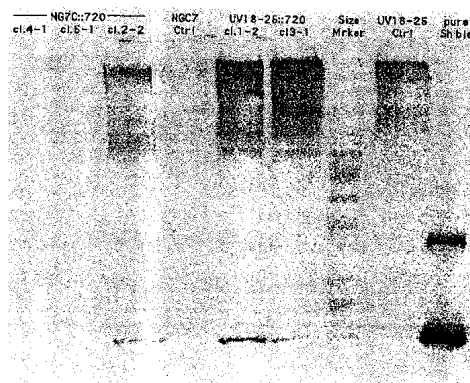
代理人 李悦

权利要求书 6 页 说明书 58 页 附图 20 页

[54] 发明名称 丝状真菌中 DNA 表达文库的高通量筛选

[57] 摘要

本发明提供了一种在丝状真菌中表达外源基因 DNA 文库的方法。该真菌可以处理有内含子的真核基因,并且也可以进行诸如糖基化、蛋白质折叠的翻译后加工。本发明提供了形态学改变的真菌的用途,其允许高通量筛选和对表达蛋白直接改造。所述转化的真菌可以用于生产大量的蛋白质,用来进行蛋白质的分离、测定和应用测试,也许还适合于蛋白质的商业化生产。



ISSN 1008-4274

1. 一种表达由一组 DNA 载体库编码的多种蛋白质的方法，其中载体库包括多种不同的载体，每种载体都包含一种不同的编码蛋白质的核酸序列，所述核酸序列可操作地连接于一种表达调控区和任选连接一种分泌信号编码序列，所述方法包括以下步骤：

- (a) 提供一种具有悬浮生长表型特征的丝状真菌株，并且具有在悬浮条件下产生可转移性复制单元的特征；
- 10 (b) 使用上述 DNA 载体库稳定转化上述丝状真菌株，使每个单一菌株都至少含有一种编码外源蛋白的核酸序列；
- (c) 将转化的丝状真菌突变株在有利于形成可转移性复制单元的悬浮培养条件下培养；
- (d) 将各种不同的可转移性复制单元相互分离；和
- 15 (e) 在有利于外源蛋白编码核酸序列表达外源蛋白的条件下，培养成单个可转移性复制单元的单克隆培养物或单克隆。

2. 一种在由 DNA 载体库编码的多种蛋白质中筛选具有所需活性或特性蛋白质的方法，包括以下步骤：

- (a) 按照权利要求 1 的方法在单克隆丝状真菌培养物或单克隆丝状真菌克隆中表达多种蛋白质；和
- 20 (b) 筛选单个克隆培养物或克隆菌落的所需活性或特性。

3. 一种得到编码具有所需活性或特性的蛋白质的 DNA 分子的方法，包括以下步骤：

- (a) 按照权利要求 1 的方法，在单克隆丝状真菌培养物或单克隆丝状真菌克隆中表达多种蛋白；
- 25 (b) 在单个克隆培养物或克隆菌落中筛选具有所需特性或活性的蛋白；和
- (c) 从表现出所需特性或活性的克隆培养物或克隆菌落中分离 DNA 分子。

30 4. 一种得到编码具有所需特性或活性蛋白的 DNA 分子的核苷酸序

列的方法，包括以下步骤：

(a) 按照权利要求 3 的方法从表现出所需特性或活性的克隆培养物或克隆菌落中分离 DNA 分子；和

(b) 将上述 DNA 测序。

5 5. 一种得到具有所需特性或活性蛋白的氨基酸序列的方法，包括以下步骤：

(a) 按照权利要求 4 的方法，得到编码具有所需特性或活性蛋白的 DNA 序列；和

(b) 将上述 DNA 序列转化成氨基酸序列。

10 6. 一种从多种单克隆丝状真菌培养物或单克隆丝状真菌克隆中筛选具有所需活性或特性的代谢物的方法，包括以下步骤：

(a) 按照权利要求 1 的方法，在单克隆丝状真菌培养物或单克隆丝状真菌克隆中表达多种蛋白；和

(b) 在每个单一克隆培养物或克隆菌落中筛选所需特性或活性。

15 7. 一种优化所需蛋白活性或特性的方法，包含如下步骤：

(a) 提供一种包含编码蛋白质突变型的 DNA 序列的载体库；

(b) 提供一种表型特征为可悬浮生长和在悬浮条件下产生可转移复制单元的丝状真菌株；

20 (c) 使用上述 DNA 载体库稳定转化上述丝状真菌株，使每个单一菌株都至少含有一种编码外源蛋白的核酸序列；

(d) 在有利于形成可转移复制单元的条件下培养转化的丝状真菌株；

(e) 将各种不同的可转移性复制单元相互分离；

25 (f) 在有利于由外源蛋白编码核酸序列编码的外源蛋白表达的条件下，培养成单一可转移复制单元的单克隆培养物或克隆菌落；

(g) 在各生物体、克隆培养物或克隆菌落中筛选具有所需特性或活性的表达蛋白；

(h) 分离一个或多个表达具有所需活性或特性蛋白的单一生物体、克隆培养物或克隆菌落；

30 (i) 从编码表现所需活性或特性蛋白的分离的单一生物体、克隆培养

物或克隆菌落中分离 DNA，并对其进行突变。

(j) 提供包含(i)步骤中得到的突变 DNA 序列的载体库；和

(k) 重复步骤(b)到(g)直到所需活性或特性达到满意的水平或不再提高。

5 8. 权利要求 7 的方法，在步骤(h)和(i)间进一步包含如下步骤：培养一个或多个在步骤(h)中分离到的单一生物体，克隆培养物或克隆菌落；分离表现出所需特性或活性的表达蛋白；对所分离蛋白的特定活性进行评价。

9. 权利要求 2 的方法，其中筛选步骤为高通量筛选。

10 10. 权利要求 3 的方法，其中筛选步骤为高通量筛选。

11. 权利要求 4 的方法，其中筛选步骤为高通量筛选。

12. 权利要求 5 的方法，其中筛选步骤为高通量筛选。

13. 权利要求 6 的方法，其中筛选步骤为高通量筛选。

14. 权利要求 7 的方法，其中筛选步骤为高通量筛选。

15 15. 权利要求 8 的方法，其中筛选步骤为高通量筛选。

16. 权利要求 1-15 中任一权利要求的方法，其中真菌表型特征为：在优化或近优化条件，有充分营养物的悬浮培养时，发酵终止时培养物粘度小于 200cP。

17. 权利要求 1-15 中任一权利要求的方法，其中真菌表型特征为：
20 在优化或近优化条件，有充分营养物的悬浮培养时，发酵终止时培养物粘度小于 100cP。

18. 权利要求 1-15 中任一权利要求的方法，其中真菌表型特征为：在优化或近优化条件，有充分营养物的悬浮培养时，发酵终止时培养物粘度小于 60cP。

25 19. 权利要求 1-15 中任一权利要求的方法，其中真菌表型特征为：在优化或近优化条件，有充分营养物的悬浮培养时，发酵终止时培养物粘度小于 10cP。

20. 权利要求 1-15 中任一权利要求的方法，其中载体包含一种真菌信号序列。

30 21. 权利要求 20 的方法，其中真菌信号序列是编码下列蛋白的真

菌基因的信号序列：纤维素酶， β -半乳糖苷酶，木聚糖酶，果胶酶，酯酶，蛋白酶，淀粉酶，多聚半乳糖醛酸酶和 hydrophobin。

22. 权利要求 1-15 中任一权利要求的方法，其中载体包含一种编码选择性标记的核苷酸序列。

5 23. 权利要求 1-15 中任一权利要求的方法，其中载体包含一种可操作地与编码蛋白的核酸序列连接的表达调控区。

24. 权利要求 23 的方法，其中表达调控区包含一种可诱导的启动子。

10 25. 权利要求 1-15 中任一权利要求的方法，其中真菌为 *Eucomycetes* 类。

26. 权利要求 25 的方法，其中真菌属于 *Onygenales* 目。

27. 权利要求 25 的方法，其中真菌属于散囊菌目

28. 权利要求 1-15 中任一权利要求的方法，其中真菌属于子囊菌类，限定条件是它们不属于酵母目。

15 29. 权利要求 1-15 中任一权利要求的方法，其中真菌属于选自曲霉属，木霉属，金孢属，链孢霉属，*Rhizomucor*，*Hansenula*，腐质霉属，毛霉属，*Tolypocladium*，镰刀霉属，青霉属，*Talaromyces*，*Emericella* 和 *Hypocrea* 之一的属。

20 30. 权利要求 29 的方法，其中真菌属于选自曲霉属，镰刀霉属，金孢属和木霉属之一的属。

31. 权利要求 30 的方法，其中真菌是金孢属菌株 UV-25，保藏号是 VKM F-3631D。

32. 权利要求 30 的方法，其中真菌是 *Trochoderma longibranchiatum* 菌株 X-252。

25 33. 权利要求 30 的方法，其中真菌是 *Aspergillus sojae* 菌株 pclA。

34. 权利要求 30 的方法，其中真菌是 *Aspergillus niger* 菌株 pclA。

35. 权利要求 1-15 中任一权利要求的方法，其中表达的蛋白与生物量之比至少是 1:1。

36. 权利要求 35 的方法，其中表达的蛋白与生物量之比至少是 2:1。

30 37. 权利要求 36 的方法，其中表达的蛋白与生物量之比至少是 6:1。

38. 权利要求 37 的方法，其中表达的蛋白与生物量之比至少是 8:1。
39. 权利要求 1-15 中任一权利要求的方法，其中可转移的复制单元是单一的真菌细胞。
40. 权利要求 1-15 中任一权利要求的方法，其中可转移的复制单元是孢子。
41. 权利要求 1-15 中任一权利要求的方法，其中可转移的复制单元是菌丝片段。
42. 权利要求 1-15 中任一权利要求的方法，其中可转移的复制单元是菌丝体小块。
43. 权利要求 1-15 中任一权利要求的方法，其中可转移的复制单元是原生质体。
44. 一种获得具有所需活性或特征的蛋白质的方法，包括以下步骤：
(a) 按照权利要求 2 的方法，从一个由 DNA 载体库编码的众多蛋白中筛选具有所需活性或特征的蛋白；
(b) 在有助于表达由编码外源蛋白质的核酸序列所编码的外源蛋白的条件下，培养适当规模的表达所需活性或特征的单克隆培养物或单克隆菌落；和
(c) 分离表达的蛋白。
45. 一种获得具有所需活性或特征的蛋白质的方法，包括用权利要求 7 或权利要求 8 的方法优化所需活性或特征，以适当的规模培养最后一步(h)中分离的单一生物体，克隆培养物，或克隆菌落，并从培养物中分离表达的蛋白。
46. 一种构建转化的丝状真菌库的方法，包括以下步骤：
(a) 提供一种具有如下表型特征的丝状真菌，在培养液中悬浮生长并在悬浮培养时产生可转移的复制单元；和
(b) 用一种 DNA 载体库稳定地转化所述丝状真菌，以使多种单一真菌的每一种都至少转入一个外源蛋白编码核酸序列；
其中 DNA 载体库包括大量不同的载体，各不同载体包含一种不同的蛋白质编码核酸序列，所述核酸序列可操作地连接于一个表达调控区，并任选连接一个分泌信号编码序列。

47. 按照权利要求 43 的方法制备的转化的丝状真菌库。

48. 一种获得表达所需活性或特征的蛋白质的转化丝状真菌宿主的方法，包括以下步骤：

- (a) 用权利要求 2 的方法筛选大量由 DNA 载体库编码的蛋白质以获得
5 所需活性或特征的蛋白；和
- (b) 分离表达所需活性或特征的单克隆培养物或单克隆菌落。

丝状真菌中 DNA 表达文库的高通量筛选

5

发明概述

本发明提供了一种在丝状真菌宿主中表达和后续筛选 DNA 文库，尤其是合成文库、基因组文库和 cDNA 文库的方法。该系统采用转化或转染的丝状真菌，这些真菌在深层培养中通过诸如有效的孢子形成等途径产生可转移的复制单元。这些真菌优选表现这样一种形态，即很少或不形成纠缠的菌丝体。其中更为优选的菌株还能表达可分离量的外源蛋白，用于评测。本发明的突变的真菌菌株由于产生可转移的复制单元、可高水平的表达外源基因、以及非常低的培养粘度，因此它们非常适合高通量的筛选技术。

15

发明背景

天然微生物种群表现广泛的生化和代谢多样性。由于许多微生物的分离和培养非常困难，这些微生物中大量有潜在价值的蛋白质和多肽不能被发现。事实上，据估计目前全世界只有不到 1% 的微生物被分离和培养。发明新的方法从未被认知乃至已被分离的微生物中分离蛋白质、多肽和代谢物仍然是一个迫切的需求。（术语“蛋白质”在后面的文章中应被理解为包括蛋白质和多肽。）鉴定和分离编码这些蛋白质的基因，从而能修饰和/或生产这些蛋白质也是人们所需要的。

Short 在美国专利 5938672、6001574、6030779 和 6057103（它们的内容通过参考合并到本发明中）中阐述了解决这一问题的一个途径。在该方法中，他们直接从环境样品（如土样）中制备了基因组 DNA 文库，而没有试图从中分离或培养任何可能存在的生物。该文库在大肠杆菌中被表达，并从获得资源甚至纯粹兴趣的角度筛选表达的蛋白。Short 暗示，但没有明确论述或肯定在该方法中使用了真菌宿主细胞。

上述方法有许多重大缺陷，其中之一就是大肠杆菌不能有效地表达有

内含子的基因。土壤中大约 90% 的物种是真核生物（主要是真菌），因此其中大部分生物的基因组 DNA 中有内含子。虽然已经发现的真菌有大约 10 万种，但是估计大约有 100 万种真菌仍未被发现（B. Kendrick, 《第五界》，Mycologue 出版社，1999），真菌基因组编码蛋白质和代谢物的潜在多样性远比已知的真菌高，但是由于内含子的存在，使得绝大部分真菌特有蛋白和代谢物不能在细菌表达系统中有效表达。不仅许多种酶类（例如分泌形真菌木质素过氧化物酶和锰依赖型过氧化物酶）是真菌特有的，而且有许多真菌蛋白，包括酶类（例如木质素过氧化物酶和黑曲霉蔗糖酶）是糖基化的；但是如果这些蛋白在大肠杆菌中表达就不能被糖基化。由于真菌基因组的数量大，长度和复杂度非常高，许多蛋白是真菌特有的，并且许多真菌蛋白都要进行糖基化，所有这些都表明，如果将给定的环境微生物样品的基因组 DNA 在细菌中进行表达，能被细菌识别并正确表达的微生物蛋白和代谢物的丰度不超过 10%。

部分由于爱滋病的蔓延和接受器官移植的人数量的增加，越来越多的人受到免疫损害或免疫抑制，导致真菌感染的数量和种类稳步增长（《传染病医学杂志》 16: 380-382,385-386(1999)）。因此正在进行的研制新抗真菌药的过程中需要分离和鉴定来自致病真菌的蛋白质，而这又需要有筛选来源于真菌基因组的 DNA 文库的能力。由于真菌基因组中内含子的存在，使得在目前已知的绝大多数细菌宿主细胞中表达这些文库变得非常困难。同样，由于越来越多的细菌感染对抗生素产生抗药性，也需要对有抗生素活性的真菌代谢物进行高通量筛选。

真核生物的基因组太复杂，以至于不能在细菌中完全表达其 DNA 文库。如果算上所有的真核生物，细菌只代表所有已知物种的 0.3%(E.O.Wilson, "生物多样性的现状", 《生物多样性》，国家学术出版社，华盛顿特区，1988，第一章)；因而能在细菌中表达的全世界基因多样性是极其有限的。

为了避免内含子的问题，可以构建 cDNA 文库在细菌中表达。但是这种方法必须有转录产物 RNA 的存在，而任何当时未被活跃转录的基因都不会出现在该文库中。而许多我们需要的蛋白只在特定条件下才表达（例如致病真菌中的毒性因子），并且不一定在收获 mRNA 时存在。而

且，为了获得足够量的 RNA 构建 DNA 文库，需要培养大量的菌体。对于环境样品中不易培养的生物，或者新的适宜培养条件还未找到的微生物而言，获得大量的 RNA 是困难或不可靠的。相比之下，通过很少量的单个细胞，采用随机引物或适合某一类基因的引物进行 PCR 扩增，就能获得足够的基因组 DNA。最后，在某个生物体中表达量高的基因也会导致其 mRNA 的超量合成，而在基因文库中，它的超量合成是以掩盖一些微量表达的基因为代价的。如果是构建 cDNA 文库而不是基因组文库，为了包括尽可能多的现存 mRNA 种类，必须要筛选多得多的克隆，因为后者更能均衡地代表现存的基因种类。显然，如果可能的话，人们更愿意构建基因组 DNA 文库。

并且，大肠杆菌不能将许多蛋白质分泌出来，这对于用于筛选目的，而筛选依赖于宿主细胞将外源基因产物分泌到胞外的宿主细胞是一个缺陷。对于大肠杆菌和所有细菌宿主细胞来说，另一个缺陷是，许多真核生物蛋白需要多种翻译后加工以获得活性，而细菌不能对它们进行翻译后加工。除了糖基化，亚基切分、二硫键形成和蛋白的正确折叠是生成有活性蛋白所常需的翻译后加工的范例。

为了保证这些翻译后加工，有时可以利用哺乳动物细胞，但是这些细胞很难培养，需要昂贵的培养基，并且不总是能高效地转化外源基因。虽然科学家们努力尝试用哺乳动物细胞做 cDNA 文库筛选的宿主 (Schouten *et al.*, WO 99/64582)，但这种转化系统显然不适合蛋白质的高通量筛选。一种在文库筛选前将转化的原生质体与哺乳动物细胞融合的方法已被报道 (美国专利 5,989,814)，但是细菌或酵母细胞中蛋白质文库的表达发生在细胞融合之前。有人试图用酶在宿主细胞表达外源蛋白后对蛋白的糖基化方式进行改造 (Meynial-salles 和 Combes, *J. Biotechnol.*, **46**: 1-14(1996))，但是这种方法必须针对不同的蛋白而改变，并且不适用于从 DNA 文库中表达蛋白。在更近一些时候，Maras *et al.*, *Eur. J. Biochem.*, **249**:701-707(1997)(又见美国专利 5,834,251) 报道一株里氏木霉 (*Trichoderma reesei*) 基因工程菌表达人类乙酰糖苷 (GlcNAc) 转移酶 I。该酶将所表达的其它外源基因的 N-乙酰葡萄糖苷转化为甘露糖残基，这是向有天然活性的哺乳动物蛋白迈进的第一步。

用酵母作为宿主细胞解决了上述的一些问题，但却引入了新的问题。酵母倾向于过度糖基化外源蛋白质（Bretthauer 和 Castellino, 1999, *Biotechnol. Appl. Biochem.* **30**: 193-200），而由于糖基化方式的改变，常使得表达的哺乳动物蛋白质具有高抗原性（C. Ballou, *Molecular Biology of the yeast saccharomyces*(酵母的分子生物学), J. Strathern *et al.*, eds., Cold Spring Harbor Laboratory Press, NY, 1982, 335-360)。虽然酵母可以应付少数一些内含子，但是它们基本上不能处理从高等物种，比如脊椎动物的复杂基因。即使来自丝状真菌的基因对于酵母来说也太复杂以至不能有效地转录，再加上酵母和丝状真菌的表达和基因拼接序列不同，使这一问题变得更加复杂（参见例如：M. Innis *et al.*, *Science* 1985, **228**: 21-26）。尽管有这些缺点，酵母转化和表达体系还是被广泛地开发，主要用作 cDNA 文库的表达。人们开发酵母表达系统，用来筛选源自哺乳动物的自然分泌蛋白和膜蛋白（Klein, *et al.*, *Proc. Natl. Acad. Sci. USA* 1996 **93**: 7108-7113; Treco, 美国专利 5,783,385），异源的真菌蛋白（Dalboge 和 Heldt-Hansen, *Mol. Gen. Genet.* **243**: 253-260(1994)），和哺乳动物蛋白（Tekamp-Olson 和 Meryweather, 美国专利 6,017,731）。

术语“酵母”在本文酵母表达系统部分中是指分类上属于酵母目的生物，例如酿酒酵母和巴斯德毕赤氏酵母。为了明确上述定义，术语“真菌”和“真菌的”应被理解为担子菌纲、结合菌纲、卵菌纲、壶菌纲和隶属于真子囊菌纲的子囊菌，而不包括酵母目。丝状真菌和酵母可以通过它们在没有性繁殖时菌丝体的延长和需要空气进行碳代谢来区别（酵母的无性繁殖是单个细胞的出芽生殖，而且酵母可以进行发酵生长代谢）。

真菌宿主细胞能最有效地进行真菌基因产物正确的内含子的拼接、糖基化、折叠和其它翻译后加工，这使得丝状真菌成为最好的进行土壤样品基因组 DNA 筛选的宿主菌。同时丝状真菌也成为生产有商业价值的真菌酶类，如蛋白酶、纤维素酶和淀粉酶的最好宿主菌。人们同时发现，丝状真菌能够转录、翻译、加工和分泌其它真核生物基因产物，包括哺乳动物基因。该特性使得丝状真菌成为生产有生物医药价值的蛋白质的合适宿主菌。与酵母相比，丝状真菌引入的糖基化模式更接近于哺乳动物蛋白的糖基化模式。由于上述原因，科学家投入了巨大的努力来开发

用来表达外源性蛋白的真菌宿主系统，并且已经开发出了许多真菌表达系统。欲了解该领域的工作，请参见 Maras *et al.*, *glycoconjugate J.*, **16**: 99-107(1999); PEBERDY, *Acta Microbiol. Immunol. Hung.* **46**: 165-174(1999); Kruszewska, *Acta Biochim. Pol.* **46**: 181-195(1999); Archer *et al.*,
5 *Crit. Rev. Biotechnol.* **17**: 273-306(1997); 和 Jeenes *et al.*, *Biotech. Genet. Eng. Rev.* **9**: 327-367(1991)。

从真菌基因组构建的 DNA 文库的高通量表达和鉴定也可以被用来确定许多现在功能未知的哺乳动物基因的功能。例如，一旦发现了一个真菌蛋白质具有一种令人感兴趣的功能特性，可以将编码该蛋白的基因序
10 列与人类基因组序列相比较来寻找同源基因。

Yelton 等在美国专利第 4,816,405 号中阐述了改造丝状子囊菌以使其生产并分泌异源蛋白质。Buxton 等在美国专利第 4,885,249 号，以及 Baxton 和 Radford, *Mol. Gen. Genet.* **196**: 339-344(1994)中通过一个带有可以整合到宿主细胞中的选择标记的 DNA 载体对黑曲霉进行了转化。McKnight
15 等在美国专利第 4,935,349 号,以及 Boel 在美国专利第 5,536,661 号阐述了利用可以引导异源基因在根霉和其它丝状真菌中表达外源基因的启动子在曲霉中表达真核基因的方法。Royer 等在美国专利第 5,837,847 号中，以及 Berka 等在 WO 00/56900 中，介绍了一种采用天然或突变过的镰刀菌启动子的有毒镰刀菌表达系统。Conneely 等在美国专利第 5,955,613 号
20 中介绍了所构建的适合在曲霉中表达和生产乳铁蛋白的质粒。枝孢菌的葡萄糖氧化酶已经在曲霉中被表达（美国专利第 5,879,921 号）。

相似的技术也被应用在链孢霉中。Lambowitzr 在美国专利第 4,486,533 号中，介绍了一种可以在丝状真菌中自动复制的 DNA 载体及其在链孢酶内引入和表达外源基因中的应用。Stuart 等在美国专利第 5,695,965 号中
25 介绍了厚壁链孢菌(*Neurospora crassa*)原生质球的哺乳动物基因和内源性转录调节因子的共转化，并在美国专利第 5,776,730 号中描述了一株经改良后减少了胞外蛋白酶的链孢菌。美国专利第 5,834,191 号中公开了用于链孢菌转化的载体。Taraki 等在美国专利第 5,436,158 号中介绍了一种用于 *Rhizopus* 中的转化系统。Sisniege-Barroso 等在 WO 99/51736 中介绍了
30 用于丝状真菌的转化系统，该系统采用了来自于 *Aspergillus awamori* 的

谷氨酸脱氢酶启动子。Dantas-Barbosa 等在 *FERMS Microbiol. Lett.* 1998 **169**: 185-190 中,介绍了通过醋酸锂法或电激法转化灰色腐殖菌(*Humicola grisea*)的 *Thermoidea* 变种,使它获得了对潮霉素 B 的抗性。

更加成功的一些真菌表达系统是曲霉属和木霉菌属,相关例子请参见
5 Berka 等在美国专利第 5,578,463 号中的说明,在 Devchand 和 Gwynne 的
文章 *J. Biotechnol.* **17**: 3-9(1991),以及 Gouka 等的文章 *Appl. Microbiol.*
Biotechnol. **47**: 1-11(1997)中也有论述。菌种 *Myceliophthora thermophila*,
Acremonium alabamense, 陆生梭菌(*Thielavia terrestris*)和 *Sporotrichum*
celluliphilum 转化株的例子分别在 WO 96/02563, 美国专利第 5,602,004、
10 5,604,129、5,695,985 号中有报道。这些报道称曲霉和木霉系统有一定的
缺陷,暗示其它的真菌也许更适于大规模蛋白质生产。

除子囊菌门外其它门的转化方法也有文章报道。请参见例子:
Munoz-Rivas 等, *Mol. Gen. Genet.* 1986 **205**: 103-106 *Schizophyllum commune*;
van de Rhee 等, *Mol. Gen. Genet.* 1996 **250**: 252-258, 双孢伞菌(*Agaricus*
15 *bisporus*); Arnau 等, *Mol. Gen. Genet.* 1991 **225**: 193-198, 拳卷毛霉(*Mucor*
circinnelloides); Liou 等, *Biosci. Biotechnol. Biochem.* 1992 **56**: 1503-1504
光亮根霉(*Rhizopus niveus*); Judelson 等, *Mol. Plant Microbe Interact.* 1991 **4**:
602-607, 致病疫霉(*Phytophthora infestans*); 以及 de Groot 等, *Nature*
Biotechnol. 1998 **16**: 839-842, 双孢伞菌(*Agaricus bisporus*)。

20 丝状真菌除了例如原生质体融合等常规转化方法外, Chakraborty 和
Kapoor, 在 *Nucleic. Acids Res.* **18**: 6737(1990)中介绍了利用电激法转化丝
状真菌。De Groot 等在 *Nature Biotechnol.* **16**:839-842(1998)中采用根癌农
杆菌(*Agrobacterium tumefaciens*)介导的几种丝状真菌的转化。科学家也开
始利用生物介质将外源 DNA 转入真菌;例子有: Christiansen 等, *Curr. Genet.*
25 **29**: 100-102(1995); Durand 等, *Curr. Genet.* **31**: 158-161(1997); 以及
Barcellos 等, *Can. J. Microbiol.* **44**: 1137-1141(1998)。以磁粒作为“磁性
生物介质”转化细胞的报道参见美国专利第 5,516,670 号和 5,753,477 号,
科学家希望这种方法适用于丝状真菌。

显然人们作了许多努力开发以真菌作为宿主菌的表达系统。然而,普
30 通的真菌宿主都是丝状真菌,在不加搅拌的培养液中会形成纠缠的菌丝

体，而在加以搅拌的生物反应器中培养液粘度又会很大。丝状真菌的这些特性也会在以之为宿主菌的工业化酶生产中导致问题。例如，高粘稠度和/或在局部形成的密集菌丝体堆积，将会对搅拌、通气和营养物分散造成困难。一般来说，由于培养液的粘度，用微量移液器将丝状真菌悬浮培养物吸到微量滴定板中有一定困难。而且，由于典型丝状真菌培养物的菌丝体互相纠缠，用它来表达 DNA 文库时，大规模地将纠缠的菌丝体分离为单个克隆并不容易，这将使检测单个基因型这一高通量检测系统所需的步骤遇到困难。

在不进行持续的搅拌时，典型的丝状真菌倾向于在液体培养基的表面生长并形成一层菌丝体，并在其上产气生孢子。而在培养基液面以下生长时又不总是产生孢子。这两个特性都对在微量滴定板中培养丝状真菌菌落，以及对有效地操作和利用这种培养物进行高通量的筛选制造了巨大的困难。悬浮的孢子和其它的可复制单元适于分离并分散到单个的微量滴定板孔中，但如果允许气生菌丝体的形成，气生孢子的产生会导致板孔间的交叉污染。在微量滴定板孔的培养基中进行搅拌以防止气生菌丝层的形成是不可行的。除了难以控制的气生孢子问题，气生菌丝层会干扰光的转移，使得许多检测（尤其是分光光度吸收检测）较难甚至无法进行。气生菌丝层还对诸如供氧、试剂和营养的添加以及取样等产生干扰。

真菌的形态特征对其培养物的物理特性的影响已经被认知，人们已经找到了有更好特性的自然产生的菌株，Jenssen 和 Boominathan 的美国专利第 5,695,985 号提供了一个例子。松散的菌丝体均一分布，有规律的分枝，都是所描述的有益的形态特征。Schuster 和 Royer 在国际专利申请 WO 97/26330 和美国专利 6,184,206 中提到了相似的方法，用来分离鉴定形态特征更适合于工业化生产外源蛋白质的真菌。该方法不是直接筛选野生型菌株，而是筛选亲本真菌细胞的突变株系，转化突变株，检测是否有转化子比亲本产生更多的外源蛋白。比亲本至少多出 10% 菌丝体分枝的突变株被特别注意。该方法被应用于木霉属、镰刀霉属和曲霉属的筛选，他们还暗示此方法可以应用于其它许多属。

Bock 等在 *Biotechnol. Bioeng.* **65**:638-648(1999)中检测了分枝频率对米

曲霉突变株培养物粘度的影响；结果发现分枝越多的突变株培养物粘度越低。而 Van Wezel 等在 PCT 应用 WO 00/00613 中，介绍了一种减少菌丝分枝和/或增加菌丝体片段的数量的方法，培养液的粘度也被降低了。该方法将灰色链霉菌的 *SsgA* 基因转入了微生物中。该方法采用的是放线菌目的丝状细菌，但宣称适用于丝状真菌。在 WO 00/56893 中，Dunn-Coleman 等描述了一株构巢曲霉的 *HbrA2* 突变株在 42°C 以上生长时分枝数大大增加，并且观察到菌丝分枝度与培养物粘度之间呈线性关系。

对丝状真菌表达系统的绝大多数前期研究都是为了分离鉴定适于工业化酶生产的菌株，因而人们的注意力主要集中在培养液粘度、转化的稳定性、单位体积的外源蛋白产量，以及产量占生物量的百分比这些方面。已经在真菌中表达了一些 DNA 文库，例如 Gems 和 Clutterbuck 在 *Curr. Genet.* 1993 24:520-524 这篇论文中，用构巢曲霉系统表达了构巢曲霉文库。Gems 等在 *Mol.Gen. Genet.* 1994 242: 467-471 这篇论文中，用曲霉系统表达了一个来源于青霉的基因组文库。但是这两个例子都没有透露或暗示对表达的蛋白进行了筛选；而是通过对宿主突变的等位基因的互补作用来证实来自 DNA 文库的基因已经获得了表达。这种互补的方法对每个想要检测活性的外源蛋白都需要一个特定的突变株做宿主菌，因而不能对一般的文库筛选提供一个工具。

Berges 等人在 *Curr. Genet.* 1993 24:53-59 这篇论文中，阐述了用在 *Trichoderma reesei* 中表达的方法克隆了一个黑曲霉蔗糖酶基因。具体方法是利用一个构建在有选择标记的粘粒上的黑曲霉基因组文库，并采用一株不能利用蔗糖的里氏木霉 (*T. reesei*) 作为宿主菌，用同胞选择的方法克隆了一个黑曲霉蔗糖酶基因。在这里，同样需要有特定性状的宿主菌来检测特定单个外源蛋白的表达，基因组文库的筛选没有被提及甚或不可能。

一个适于表达 DNA 文库的宿主菌的特点在许多方面与适合于工业化蛋白质生产的宿主菌不同。总的来说，一株适合于高通量筛选的真菌宿主菌应具备许多条件，其中包括：

- 宿主菌必须能进行高效转化。
- 宿主菌必须能处理含有内含子的基因，并能进行任何必要的拼接。

- 宿主菌必须能对表达的蛋白进行翻译后加工，以使蛋白变为有活性的形式。

- 当从文库中筛选一种蛋白质时，宿主菌必须能够生产足够多的该蛋白以供分析之用。

5 - 宿主菌必须能接受许多种表达调控因子，并能轻松利用各种调控因子。

- 宿主菌培养物中的菌丝体不能相互纠缠而对分离单个克隆产生困难，也不能相互纠缠而升高粘度到不能在小型化的高通量筛选中（如微量取样）有效转移和复制的程度。

10 - 宿主菌不能在培养基表面形成菌丝层，而应优先在培养基液面以下生长。

- 在高通量筛选所提供的生长条件下，宿主菌必须能在培养基液面以下有效地形成孢子或其它无性繁殖体。

15 当被筛选物是代谢产物时，如果宿主细胞能将代谢产物分泌到培养基是再好不过了。在培养基中代谢产物可以被很容易地检测和/或分析。在理想状态下，宿主菌应该只分泌外源蛋白。

20 当对蛋白质进行分析时，如果宿主菌也能表达足够用来分离和纯化的外源蛋白就非常有利了。如果有具备这种特点的宿主菌，就可以只通过培养宿主细胞，而不用烦琐费时的分子生物学操作将基因转入其他生物体，来进一步鉴定所有感兴趣的外源蛋白。宿主菌能将蛋白质分泌到胞外是一个好的特性，这样可以进行更可靠和更多样的检测。

如果宿主细胞能方便地分离外源 DNA 也将是非常有利的，这样进一步的研究以及改造基因本身也可以进行。

25 除了上述的特性，这种转化系统还应具备一些其它的特点。转化的效率应该足够高，以产生有意义的筛选所必须的足够转化子数。在理想状态下，外源蛋白的表达应由一种诱导物诱导，并且该诱导物通过一条途径作用在一个启动子上。

30 迄今为止，还没有任何宿主菌和转化系统的组合符合全部这些条件，甚至连能符合大部分的也没有。因此，人们仍然需要能有效表达 DNA 文库——尤其是基因组文库和/或真核基因组文库的基因产物的真菌宿主和转

化系统。

发明简述

本发明采用在深层培养时能产生“可转移的复制单元”的丝状真菌。

5 “可转移的复制单元”是指孢子、无性芽、菌丝片段、原生质体、皮层碎片，或其它符合下面两个特征的真菌单元：(1) 在培养基中能容易地与其它这类单元分离，和(2) 能够自身复制成一个单克隆培养物。优选真菌还表现出丝状形态不明显和/或生长紧密的优点，并且产生适于高通量 DNA 文库筛选中涉及的物理操作的低粘度培养物。尤其优选的是，即使
10 在不加搅拌的情况下仍然倾向于在培养基内部生长，而不是在培养基表面形成菌丝层的丝状真菌。

本发明利用了在国际专利申请 PCT/NL99/00618 和 PCT/EP99/202516 中公开的转化系统的特点。这些申请描述了一种对诸如 *Chrysosporium lucknowense* 、*Aspergillus sojae* 等丝状真菌宿主非常有效的转化系统。这些
15 申请还公开已经获得了保留野生型宿主细胞所有优良性状，但部分失去其丝状表型的突变株，这种突变株降低了培养物的粘度。

本发明所采用的优选真菌能够大量表达并分泌外源蛋白质，与已知丝状真菌宿主相比蛋白质/生物量之比获得了提高。本发明提供了一个表现高转化子产量的转化系统。本发明还提供了能有效表达外源 cDNA 插入
20 体，尤其是基因组 DNA 插入体的蛋白产物的真菌转化株库。本发明的另一方面，转化的真菌库可以用来筛选外源蛋白的活性或特性，或者筛选转化真菌由于外源蛋白的作用而产生的代谢物，也可以用来筛选外源 DNA 或其转录产物 RNA。应该理解本发明还能够高通量筛选具有上述性状的非转化菌株的代谢物。

25 本发明使用的术语“丝状真菌突变株”仅指在自然界找不到的真菌。导致例如紧密的生长状态、低粘度、低蛋白酶含量、在液面以下生长等可描述的表性特征的“突变”可以通过紫外线诱变和化学诱变等经典方法，或者通过如盒式突变等分子生物学手段随机产生，还可以通过缜密的基因工程方法来获得。如果一株天然真菌具有这些必要特性，它将理
30 所当然地可以被应用于本发明的方法中。

在本发明的另一方面，转化的真菌库可以用来筛选真菌自身的有用性质，例如，一种特定表达的蛋白或代谢物的高水平生产。本发明的这一方面可以通过对目标表达蛋白的定量分析来说明，其中可以检测到有最优良的蛋白生产、蛋白加工和蛋白分泌特性组合的特殊菌株。

- 5 在本发明的再一方面，转化真菌库可以用来筛选能够与目标核酸探针杂交的 DNA 序列的存在。

附图说明

- 图 1 是实施例中描述的 Western 印迹图。
- 10 图 2 是质粒 pUT720 的图谱。
图 3 是质粒 pUT970G 的图谱。
图 4 是质粒 pUT1064 的图谱。
图 5 是质粒 pUT1065 的图谱。
图 6 是质粒 pF6g 的图谱。
- 15 图 7 是质粒 pUT1150 的图谱。
图 8 是质粒 pUT1152 的图谱。
图 9 是质粒 pUT1155 的图谱。
图 10 是质粒 pUT1160 的图谱。
图 11 是质粒 pUT1162 的图谱。
- 20 图 12 是 *pclA* 蛋白的结构图。
图 13A 是黑曲霉(*Aspergillus niger*) 野生型的显微照相图。
图 13B 是黑曲霉 *pclA* 突变株的显微照相照片。
图 14A 是酱油曲霉 (*Aspergillus sojiae*) 野生型的显微照相照片。
图 14B 是酱油曲霉 *pclA* 突变株的显微照相照片。
- 25 图 15A-E 是 *pyrE* 基因的测序结果。带下划线的是氨基酸序列；由于一些核酸序列不确定导致氨基酸序列不连续。标出的氨基酸序列是最可能的序列，黑体字代表推测的/可能的内含子。

本发明的详细描述

- 30 从最广义的角度来说，本发明涉及在悬浮培养液中能产生可转移的复

制单元的转化丝状真菌；涉及这种真菌库；还涉及从这种库中筛选诸如与表达的外源蛋白质或相关代谢产物（即由内源性和/或外源性酶产生的小分子量产物）有关的生物化学活性或生物学活性等所需生物学特性的方法。这个低粘度的丝状真菌库包括含有核酸序列的真菌，每段核酸序列编码一个外源蛋白，每段所述核酸序列都可操作地连接于一个表达调控区，并可选择地连接分泌信号编码区和/或载体蛋白编码区。本发明中的转化菌株优选可以分泌外源蛋白。

本发明所使用的表达和筛选方法，和其中使用的真菌菌株，对生产真菌、蛋白质、代谢物和 DNA 分子这些在各方面有广泛用途的物质很有用。本发明的方法对获得核酸和蛋白质序列信息也很有用，这些信息本身也被看作本方法的有用产品之一。

本发明的优选丝状真菌特征在于培养基粘度低。典型的工业级丝状真菌的培养物粘度远大于 200 厘泊 (cP)，通常是 1,000 cP 以上，甚至能达到 10,000 cP；本发明的真菌在有充足营养及最适或接近最适生长条件下培养 48 小时或更长时间以后，培养物粘度在 200 cP 以下，优选 100 cP 以下，更优选 60 cP 以下，最优选 10 cP 以下。本发明的丝状真菌通常呈现短的、不连续、不相互纠缠的菌丝或小团的形态。这些小团是从一个单克隆产生的有轻微或无相互纠缠的菌丝集合体，与从多个相互缠绕的克隆产生的菌丝团是不一样的，后者要大得多。例如，*Chrysosporium lucknowense* 突变株 UV18-25（培养粘度<10 cP）和形态相似的 *Trichoderma longbanchiatum* 突变株 X-525（培养粘度<60cP），特征在于有短的、明显的、不相互纠缠的菌丝，长度在 100-200 微米之间；低粘度基因工程突变株 *Aspergillus sojae* pclA 的形态特征是有多分枝和短菌丝的紧密形态（参见图 14）。而 WO97/2630 中提到的低培养粘度真菌在形态上“分枝更多”，本发明中的真菌与非突变菌株相比菌丝分枝与之相当甚或更少。看起来菌丝的长度在控制培养物粘度上起关键作用。

特别优选的真菌菌株有高的外源分泌蛋白/生物量之比。这个比率大于 1:1，较有大于 2:1，更优选 6:1 或更高，最优选 8:1 或更高。这样高的比率在高通量筛选环境中是有利的，因为它会使外源蛋白的浓度更高，允许更灵敏和/或更快速的筛选检测。这在检测的体积减小时尤其有利，

例如从 96 孔微量滴定板到 384 孔板,再到 1536 孔板时。本发明的方法适用于所有这些型号的微量滴定板,也适用于绝大多数其它采用液体样品的、HTS 样式的滴定板。

我们期望用本发明所讲的突变方法,可以将所有丝状真菌转变为适合于在本发明中应用的突变株。优选的丝状真菌的属有:金孢属 (*Chrysosporium*),梭孢壳属 (*Thielavia*),链孢酶属 (*Neurospora*),金色担子菌属 (*Aureobasidium*),线黑粉菌属 (*Filobasidium*), *Piomyces*, *Crylococcus*, *Acrimonium*, *Tolypocladium*, *Scrytalidium*,裂褶菌属 (*Schizophyllum*),侧孢霉属 (*Sporotrichum*),青霉属 (*Penicillium*),赤霉菌属 (*Gibberella*),毁丝霉属 (*Myceliophthora*),毛霉属 (*Mucor*),曲霉属 (*Aspergillus*),镰刀霉属 (*Fusarium*),腐质霉属 (*Humicola*),和木霉属 (*Trichoderma*),和它们的无性型及远距离刺激变形型。其中优选金孢属,木霉属,曲霉属和镰刀霉属。更优选金孢属。真菌的种属可以根据与 Barnett 和 Hunter 所著的《不完美的真菌属的图解》第三版,1972, Burgess 出版公司——书中描绘的形态的一致性来定义。对金孢属真菌的分类提供详细原始资料的有: Van Oorschot, C.A.N.(1980) “对金孢属及其相关属的修订”, *Studies in Mycology No. 20* Centraal Bureau voor Schimmelcultures(CBS), Barm, The Netherlands, 第 1-36 页。根据这些资料,金孢属属于丝孢菌目的丛梗孢科(*Moniliaceae*)。

另一个提供真菌命名的方便资源是布达佩斯条约的保藏,尤其是那些提供在线数据库的(以下这些网址采用 HTTP 传输协议)。美国的 ATCC 在 www.atcc.org 提供信息, CBS(NE)在 www.cbs.knaw.nl 提供信息,俄罗斯的 VKM 在 www.bdt.org.br/bdt.msdn.vkm/general 提供信息。另一个资料来源是 NT.ars-grin.gov/fungaldatabases。所有这些机构都能提供区分真菌种的特征的教程。在 www.ncbi.nlm.nih.gov/htbonpost/taxonomy/wgetorg?mode=Undef&id=4890 网站可以找到 *Ascomycota* 的另一种分类。根据这另外一种分类,金孢属隶属于 *Ascomycota* 门、*Onygenales* 目、*Onygenaceae* 科。

金孢属的范围包括但不局限于这些菌株: *C.botryoides*, *C.carmichaelli*, *C.crassitunicatum*, *C.Europe*, *C.evolceannui*, *C.farinicola*, *C.fastidium*, *C.*

filiforme, *c. georgiae*, *C. globiferum*, *C. globiferum var articulatum*, *C. globiferum var niveum*, *C. hirundo*, *C. hispanicum*, *C. holmii*, *C. indicum*, *C. inops*, *C. keratinophilum*, *C. kreisei*, *C. kuzurovianum*, *C. lignorum*, *C. lobatum*, *C. lucknowense*, *C. lucknowense* Grag 27K, *C. medium*, *C. medium*
5 *var spissescens*, *C. mephiticum*, *C. merdarium*, *C. merdarium var roseum*, *C. minor*, *C. pannicola*, *C. parvum*, *C. parvum var crescens*, *C. pilosum*, *C. pseudomerdarium*, *C. periformis*, *C. queenslandium*, *C. sigleri*, *C. sulfureum*, *C. synchronum*, *C. tropicum*, *C. undulatum*, *C. vollenarense*, *C. vsepertilium*, *C. zonatum*.

10 *C. Lucknowense* 是金孢属的一个人们特别感兴趣的种，因为在自然条件下它的纤维素酶蛋白的产量比较高（国际专利申请 WO 98/15633, PCT/NL99/00618, 和美国专利 5,511,381 和 6,015,707）。国际保藏号为 ATCC 44006, CBS 251.72, CBS 143.77, CBS 272.77, 和 VKM F-3500D 的菌种都是 *Chrysosporium lucknowense* 菌株的例子。金孢属的定义范围还
15 包括从金孢属同一祖先衍生来的菌株，包括自发突变或诱变的突变株。本发明的方法，在一个具体方面，利用了一株通过辐射和化学诱变获得的金孢属突变株，该菌株在悬浮培养时倾向于产生可转移的复制单元，其形态特征是短的、不连续的、不相互纠缠的菌丝（紧密生长），还有一个表型特征是在悬浮培养时在液面以下生长，发酵培养基的粘度较低。
20 在另一个方面，本发明采用了形态相似的木霉属突变株。在另一个方面，采用了形态相似的酱油曲霉或黑根霉突变株。

例如，将 VKM F-3500D（C1 菌株）用紫外线照射，获得菌株 UV13-6。随后，该菌株用 N-甲基-N'-硝基-N-亚硝基鸟嘌呤进一步诱变，产生菌株 NG7C-19。后者接着被紫外线诱变，获得菌株 UV18-2(VKM F-
25 361D)。在突变过程中，该菌株在液体培养基或平板以及显微镜下的形态特征发生了一些变化。在每一次连续的突变后，在平板上培养的突变株的金孢属典型绒毛状和毡状的形态特征就会减少一些。野生型在某些培养基中生长时会产生棕色色素，而突变株的这种特征则不明显。在液体培养基中生长时，突变株 UV18-25 的粘度明显低于野生型菌株 C1 和突
30 变株 UV13-6 和 NG7C-19。虽然所有菌株都大致保持着金孢属的显微镜

下特征，但所有菌株每经过一代突变菌丝就变细一些，而 UV18-25 可以明显观察到菌丝呈片段状。这种菌丝片段可能是造成 UV18-25 的培养物粘度降低的原因之一。每进行一代突变，菌株产生气生孢子的能力就减弱一些。这些结果表明一个菌株可以总体上属于金孢属，而在形态上表现出与传统分类（形态上）定义有一些差异。

特别是金孢属的无性型，我们发现它特别适合本发明所述的筛选应用。无性型的代谢使其特别适合高效表达。有性型也应该适合，因为无性型和有性型的基因组成是一样的。无性型和有性型的区别在于一个是无性状态，另一个是有性状态；这两种状态在特定培养条件下表现出不同的形态。

另一个代表基因工程突变菌株的例子是酱油曲霉(*Aspergillus sojae*)。这些突变株中的一株的编码一个特定内切蛋白酶的基因被阻断了。这一突变导致了菌丝变短，分枝增加，菌丝生长紧凑，在培养基液面以下形成菌丝小片。并且，在本应用中提到的 *Aspergillus sojae* 在特定的液体培养基内生长可以被诱导表现出高效的生孢子能力，这使得它特别适合于高通量筛选系统。在这个例子中，诱导产生可转移的复制单元的条件仅仅是由一种含 0.6 克/毫升 EDTA 的合成培养基组成。诱导条件会随着不同的宿主而改变，但很明显的是如果一个宿主菌适合，这个条件应该早就为人所知了。

人们喜欢用不产毒素和非致病的真菌菌株，现有技术已知道很多这种菌，因为这会降低操作者的危险性，并能简化总体的筛选过程。在优选的实施方式中真菌还将是蛋白酶缺陷型，这样可以最大程度地降低外源蛋白的降解，和/或易于抑制蛋白酶的产生。将蛋白酶缺失的菌株用作表达宿主菌是众所周知的，例如 PCT 申请 WO 96/29391。蛋白酶缺失的菌株可以通过筛选突变株来产生，或者蛋白酶基因可以被“敲除”，或者可以用本文中的方法使其失活，例如 Christensen 和 Hynes 在美国专利 6,025,185 号中所描述的(*areA* 基因功能丧失的 *Aspergillus sojae*)。

人们发现可以处理金孢属突变株使其蛋白酶表达量下降，这就使它们更适合于生产蛋白类产品，尤其是当蛋白类产品对蛋白酶活性敏感时。因此本发明也可以使用一株金孢属突变株，该突变株比诸如 C.

lucknowense 菌株 C1(VKM F-3500 D)等非突变金孢属菌株产生的蛋白酶少。这些突变株的特别之处是其蛋白酶活性（除了所有的有选择的用来切割分泌的融合蛋白的蛋白酶）小于 C1 菌株产生的活性的一半，更好的情况下小于 30%，最好时小于 10%。被降低的蛋白酶的活性可以用已知的方法测定，例如测量脱脂牛奶平板上的抑菌圈或牛血清白蛋白降解。

已经发现可以构建蛋白酶表达水平降低的金孢属突变株，以使它们更适于生产蛋白产物，尤其是当蛋白产物对该蛋白酶活性敏感时。因此，本发明也可以利用比金孢属非突变株例如 *C. lucknowense* C1 株 (VKM F-3500 D) 产生更少蛋白酶的金孢属突变株。尤其是当该突变株的蛋白酶活性(而非裂解融合蛋白的蛋白酶)不及 C1 株的一半或少于 30%甚至 10%时更好。蛋白酶活性的减少用已知方法测定，如测定脱脂牛奶平板形成的抑菌圈或牛血清白蛋白的降解。

为了减少分析时宿主蛋白的干扰，去除宿主丝状真菌的其它基因如编码纤维素酶和其它分泌较多的蛋白的基因的活性，是再好不过了。可以删除或突变编码分泌蛋白的基因，或者对非目的蛋白的诱导系统及表达的途径进行改造以降低其表达。本发明（见下面）中载体使用内源性操纵子，是为了灭活同一个诱导子控制下的其它的基因的活性。蛋白酶分泌较易抑制的真菌的蛋白酶表达往往受环境调节因子如氨基酸浓度的控制，而通过这些环境因子可以减少蛋白酶的生产。

转化载体中优选应用能够在选择的宿主中高表达的同源表达调节区。来源于异源宿主如木霉属或曲霉属的高表达调节区，众所周知而且都在使用。举例但并不限于这些，已知大量表达并因此为本发明的应用提供合适表达调节序列的蛋白质的例子是 hydrophobin，蛋白酶、淀粉酶、木聚糖酶、果胶酶、酯酶、 β -半乳糖苷酶、纤维素酶 (endo-glucanase, cellobiohydrolase) 及多聚半乳糖醛酸酶。

表达调节区包含一个可控制的启动子序列，和要表达的蛋白的核酸序列连接在一起。启动子和要表达序列的起始密码子紧密相连，以达到表达的目的。启动子序列可以是组成型的，但诱导型的更好。使用诱导型启动子和适当的诱导介质更利于和启动子相连的序列的表达。我们曾考察过一些来源于同源物种或异源株系的、能够促使蛋白质表达的表

节序列。这种表达调节序列适合于真菌的表达调节区，例如子囊菌纲的调节区。合适的子囊菌纲的表达调节区是来自下列属任意一种的调节区：曲霉属，木霉属，金孢属，腐质霉属，链孢霉属，*Tolyocladium*，镰刀霉属，青霉属，*Talaromyces*，或它们的另一种性别型例如 *Emericella* 及镰刀霉属。

- 5 来自木霉属的 cellobiohydrolase 启动子；曲霉属的乙醇脱氢酶 A，乙醇脱氢酶 R，谷氨酸脱氢酶，RAKA 淀粉酶，甘油淀粉酶，甘油醛磷酸脱氢酶启动子，链孢霉属的磷酸甘油及旁路的启动子；*Rhizomucor miehei* 的脂酶、冬氨酸蛋白酶启动子；*penicillium canescens* 的 β -半乳糖苷酶的启动子；而 cellobiohydrolase，葡聚糖内切酶，木聚糖酶，甘油醛磷酸脱氢酶和
10 蛋白酶的启动子是典型代表。和宿主同属的表达调节序列是最佳选择，因为它更可能特别地适应宿主。

来自表达巨大量蛋白质的金孢属菌株的天然表达调节序列特别优选。这些菌株的例子已经根据布达佩斯条约保藏于莫斯科的全俄菌种保藏机构(All Russian Collection, VKM)。野生型 C1 株保藏号为 VKM F-3500 D，
15 保藏日为 1996 年 8 月 29 日，C1 UV13-6 突变株保藏号 VKM F-3632 D，保藏日 1998 年 9 月 2 日，C1 NG7C-19 突变株保藏号 VKM F-3633 D，保藏日为 1998 年 9 月 2 日，C1 UV18-25 突变株，保藏号 VKM F-3631 D，保藏日为 1998 年 9 月 2 日。这些菌株也优选用来作为生产低-粘性突变株的菌源；事实上 VKM F-3631D 菌株已经表现必需的低粘度表型特征。
20 经过对 *Trichoderma longibrachiatum*18.2kk 的两轮辐射选择，已经得到木霉属的一个编号为 X-252 低粘度突变株，而 X-252 低粘度突变株是从 *Trichoderma longibrachiatum* 的 QM 9414 株的突变体 (ATCC 26921) 得到。本发明的其它实施方式利用了 *Aspergillus sojae* 和 *Aspergillus niger* 表现型近似的突变株。

- 25 当宿主是金孢属时，优选应用金孢属的启动子序列，以保证启动子被宿主正确识别。某些异源表达调节序列也和金孢属本身的表达调节序列一样在，金孢属中具有相同的调节效率。这样以来就使我们可以构建用于转化，金孢属的出色载体，并提供在此宿主中进行高效转化和表达的许多可能性。如，标准的曲霉属转化技术可参考 Christiansen 等在
30 *Bio/Technology*1998 6:1419-1422 一文的描述，其它有关转化载体的详细

描述包括：美国专利 4,816,405、5,198,345、5,503,991、5,364,770、5,705,358、5,728,547、5,578,463、EP-B-215.594(也适用于木霉属),它们的内容通过参考并入本文。因为在金孢属中发现纤维素的表达率非常高,纤维素基因的表达调节区很受欢迎。

- 5 本发明中载体允许插入编码酶(尤其对分泌出去的酶)的基因的启动子序列。这些酶包括糖降解酶(纤维素酶,木聚糖酶,甘露聚糖酶,甘露糖苷酶,果胶酶,淀粉酶,例如葡萄糖淀粉酶, α -淀粉酶, α 和 β 半乳糖苷酶, α 和 β 葡萄糖苷酶, β -葡聚糖酶,几丁质酶),蛋白酶(内源蛋白酶,氨基蛋白酶,氨基及羧基多肽酶),其他水解酶(脂肪酶,酯酶,肌醇六磷酸酶),和转移酶(糖基转移酶,谷氨酰胺转移酶,异构酶,转化酶)。来自 *Chrysosporium lucknowense* 的几个酶的例子见表 A。

表 A: 来自 *Chrysosporium lucknowense* 的部分酶的特性

样品	氨基酸数量	保留 >50% 活性的最高 PH			保留 >70% 活性的最高 PH			稳定性 20h, 50 °C pH 7.5/8
		CMC ase	RBB CMC ase	其它底物	CMC ase	RBB CMC ase	其它底物	保留的最高活性 %
30 Kd 碱性蛋白酶		-	-	12.5	-	-	12.0	-
30 kD Xyl (碱性)	333	-	-	10.0	-	-	8.5	80
51 kD Xyl		-	-	8.0	-	-	7.5	-
60 kD Xyl		-	-	9.5	-	-	9.0	85
30 kD endo (EG3)	247							
45 kD endo		7.0	8.0	-	6.5	7.0	-	75
55 kD endo	247	8.0	8.0	-	7.0	7.0	-	55
25 kD(21.8 kD)endo (EG5)	225	7.5	10.0	-	6.5	9.0	-	80
43 kD(39.6 kD)endo (EG6)	395	8.0	8.0	-	7.2	7.2	-	-
45 kD α , β -Gal/ β -Gluc		-	-	6.8	-	-	5.7	-
48 kD CBH		5.2	7.5	8.0	5.0	6.8	-	-
55 kD CBH1	526	8.0	9.0	-	7.4	8.5	-	70
65 kD PGU		-	-	8.0	-	-	7.3	-
90 kD 蛋白酶		-	-	9.0	-	-	9.0	-
100 kD 蛋白酶		-	-	9.0	-	-	9.0	-

注: *通过 MALDI 测定的分子量; 其它均为通过 SDS PAGE 测定的分子量

xyl = 木聚糖酶

30 endo = 葡聚糖内切酶

gal = 半乳糖苷酶

gluc = 葡糖苷酶

CBN = cellobiohydrolase

PGU = 半乳糖醛酸酶

5

核酸构建体优选包括来自金孢属、最好是来自 *Chrysosporium lucknowense* 或其衍生物的表达调控区，该调节区可操作地与一个编码待表达蛋白的核酸序列连接。特别优选的核酸构建体将包括来自表达纤维素酶或木聚糖酶的金孢属的表达调控区，优选包括来自表达
10 cellobiohydrolase, 更优选表达表 A 描述的 55 kDa cellobiohydrolase (CBH1) 的金孢属的表达调控区。再举些例子，金孢属的 hydrophobin、蛋白酶、淀粉酶、木聚糖酶、酯酶、果胶酶、 β -半乳糖苷酶、纤维素酶(葡聚糖内切酶, cellobiohydrolase)和多聚半乳糖醛酸酶的启动子序列也在本发明的范围内。

15 表 A 中公开的酶的所有启动子或表达调控区都可以恰当地得以使用。可以很容易地从金孢属得到这些启动子或表达调控区的核酸序列。确定启动子序列的方法多种多样而且众所周知。启动子序列通常紧靠在相关基因开始的起始密码子 ATG 的上游。举例来说，可以用去除相关基因上游的序列、重组 DNA 技术、及检验去除这些序列对基因表达的影响来
20 鉴别启动子序列。再例如，通过比较相关基因上游序列的同源启动子序列，也可以发现启动子序列。

比如，C1 葡聚糖内切酶的启动子序列就是通过克隆相应基因得来的（见 PCT/NL99/00618）。本发明倾向于使用来自金孢属 55 kDa cellobiohydrolase (CBH1)、甘油醛磷酸脱氢酶 A、30k Da 的木聚糖酶(XylIF)
25 的启动子，因为这些酶应用它们的启动子时表达水平较高。*Chrysosporium lucknowense* 尤其是 *C.lucknowense* GARG 27K 糖降解酶的启动子因可在其它真菌宿主中表达很多蛋白，很好使用。

本发明核酸序列的具体实施方式是已知的金孢属，曲霉属和木霉属的核酸序列。金孢属的启动子序列见 PCT/NL99/00618 的描述。现有技术提
30 供了大量可在曲霉属中应用的表达调控区，例如美国专利 4,935,349;

5,198,345; 5,252,726; 5,705,358; 5,965,384 和 PCT 申请 WO93/07277。在木霉属菌株中表达公开在美国专利 6,002,725 中。这些专利的内容通过参考其全文并入本发明。

Hydrophobin 基因是一个高表达的真菌基因。有人认为 Hydrophobin 基因的启动子序列、优选来自金孢属的启动子序列，适合用作本发明适当实施方式中的表达调节序列。*Trichoderma reesei* 和 *Trichoderma harzianum* 的 Hydrophobin 基因序列已经为现有技术所公开，并且 *Aspergillus fumigatus* 和 *Aspergillus nidulans* 的基因序列及相关序列信息被参考并入本发明（Nakari-Setala 等, *Eur. J. Biochem.* 1996, 235:248-255; Parta 等, *Infect. Immun.* 1994 62:4389-4395; Munoz 等, *Curr. Genet.* 1997, 32:225-230; Stringer 等, *Mol. Microbio.* 1995 16:33-44）。利用这些信息，本领域普通技术人员按照标准方法例如前面提到的技术，不必经过繁重实验即可得到金孢属 hydrophobin 基因的表达调控序列。根据本发明的一个重组金孢属菌株可以包含一个可操作地与编码外源蛋白的序列相连的 hydrophobin 调控区。

表达调控序列也可以另外包含一个增强子或沉默子。它们也是现有技术已知的，它们一般远离启动子。表达调控序列还包括具有激活物结合位点和阻抑物结合位点的启动子。在某些情况下，这些部位也可以得以修饰以去除这种类型的调控。例如，存在 creA 位点的丝状真菌启动子的已有报道。可以使 creA 位点突变，以确保由于存在的 creA 被去除而导致葡萄糖的抑制。利用这样的启动子可以产生在葡萄糖存在下启动子控制的核酸序列编码的蛋白库。这种技术的例子见 WO 94/13820 及 WO 97/09438。这些启动子在有后无其 creA 位点的情况下均能使用。creA 位点突变的突变体可用作本发明重组株的表达调控序列，而它所调控的核酸序列库可以在葡萄糖存在下得以表达。这类启动子如 WO 97/09438 中所描述的类似方式确保去抑制。CreA 位点的鉴别知识是现有技术已知的。另外，可以在抑制系统以外具有突变的宿主菌株中应用 CreA 结合位点未突变的启动子，从而该菌株尽管存在 CreA 结合位点，仍能够在葡萄糖存在下产生蛋白库。

终止序列也属于表达调控序列，它们可操作地连接在要表达的序列的

3'端。可能有许多已知的真菌终止子在本发明的宿主菌中起作用。如 *A. nidulans* trpC 的终止子, *A. niger* 的 α -葡糖苷酶终止子、*A. niger* 的葡糖淀粉酶终止子、*Mucor miehei* 的羧基蛋白酶终止子(参见 US 5578463), 以及 *Trichoderma reesei* 的 cellobiohydrolase 终止子。金孢属的终止子序
5 列, 例如 EG6 终止子, 当然在金孢属中发挥很好的得作用。

用于本发明的合适转化载体可以任选具有可操作地与信号序列的编码序列连接的待表达外源核酸序列。而信号序列为一段氨基酸序列, 当和待表达蛋白的氨基酸序列可操作地连在一起时, 能够使该蛋白从宿主生物体中分泌出去。这样的一种信号序列可以是与异源蛋白有关的或是宿
10 主本身的。编码信号肽的核酸序列必须定位在读码框中, 以使信号肽和异源蛋白一同翻译。本发明信号肽的位置(包括单一的、分泌外源蛋白)最好处于分子进化该分子所处位置。

一个载体上只连接一个信号序列来进行文库的表达, 不是最好选择, 因为这将减少表达目的蛋白的可能。随机剪断 DNA 而建立的基因组文库, 然后再克隆到一个载体中, 库中基因正好和信号序列连接在一个阅
15 读框架的可能性较小。而且, 即便是连接到一起, 选择的信号序列可能不会对所有基因起作用。鉴于此, 最好不要应用信号序列来筛选基因组库, 而要通过检测细胞内外源蛋白的存在来筛选基因组库。细胞内蛋白的存在和活性分析可以通过用酶将真菌转化为原生质体、然后再溶解后
20 得以实现。有关操作见 Zeyl 等, *J. Biotechnol.* 59:221-224(1997) 的描述。这种方法已经应用于生长在微量培养板上的金孢属转化子的 PCR 克隆检测。

讲解一下能使金孢属菌株分泌蛋白的信号序列。此类序列以真菌的信号序列为好, 但以 *Ascomycete* 的信号序列为最好。通常可以从真核细胞
25 得到合适的信号序列, 最好从酵母和以下各属的真菌得到: 曲霉属, 金孢属, 木霉属, *Pichia*, 链孢霉属, *Rhizomucor*, *Hansenula*, 腐质霉属, 毛霉属, *Tolypocladium*, 镰刀霉属, 青霉属, *Saccharomyces*, *Talaromyces*, 或不同生殖方式的 *Emericela*、*Hypocrea*。那些和 cellobiohydrolase、葡聚糖内切酶、 β -半乳糖苷酶、木聚糖酶、果胶酶、酯酶、蛋白酶、hydrophobin 或
30 淀粉酶生来就连在一起的信号肽特别有用。如曲霉属或腐质霉属的淀粉

酶或葡糖淀粉酶, *Aspergillus oryzae* TAKA 淀粉酶, *Aspergillus niger* 的 α -淀粉酶, 毛霉属的羧基肽酶 (US 5578463), *Rhizomucor miehei* 的脂肪酶或蛋白酶, 木霉属的 cellobiohydrolase, 来自金孢属 *Penicillium canescens* CBH1 株的 β -半乳糖苷酶, *Saccharomyces* 的 α -交配因子的信号序列。

另外信号序列也可以是杆状菌的淀粉酶或枯草杆菌蛋白的信号序列。从同一属的宿主菌株得到的信号序列尤其适合, 因为它最可能会特别适应特定宿主的要求; 因此当宿主是 *Chrysosporium lucknowense* 时, 信号序列最好是金孢属的信号序列, 金孢属菌株分泌的蛋白量异常地高, 当然来自这些株系的信号序列让人感兴趣。来自丝状真菌和酵母的信号序列, 也许和非真菌的信号序列一样有用。

按照本发明的任意实施方式, 转化的重组真菌宿主可以进一步携带一个选择标记。此选择标记可用来选择转化或转染细胞。选择标记通常编码一种基因产物, 进而导致宿主产生不同于非转化细胞的特定类型的抗性。通常可以是对重金属的抗性、抗生素或杀生剂抗性。原养型是非抗生素变种的一种有用选择型标记。自养型标记导致宿主产生营养缺陷, 而纠正这些缺陷的基因可以用来作为选择的标记。经常使用的抗性和自养选择标记是 amdS(己酰胺酶)、hph(潮霉素磷酸转移酶)、pyrG(乳清酸核苷-5'-磷酸脱羧酶)、pyr (乳清酸 P-核糖转移酶)、trpC (邻氨基苯甲酸合成酶)、argB(鸟氨酸氨甲酰基转移酶)、sC(硫酸腺苷转移酶)、bar(膦丝菌素乙酰转移酶)、niaD(硝酸还原酶)、Sh-ble (博来霉素-腐草霉素抗性)、突变的乙酰乳酸合成酶 (磺酰脲抗性)、新霉素磷酸转移酶 (氨基糖苷抗性)。金孢属菌株最好的选择标记是乳清酸 P-核糖转移酶。选择标记单独在一个载体上或者选择标记和要表达异源蛋白的核酸序列在同一片段上, 都能通过共转化进行选择。

可以通过使用 AMA1 复制子序列进一步提高转化效率, 这对 *Aspergillus niger*(Verdoes 等, *Gene* 146:159-165(1994))是有用。该序列导致许多丝状真菌的转化效率增加 10-100 倍。另外, 导入的 DNA 不整合到真菌的基因组中, 而是以多拷贝、自主的形式保留在真菌细胞里。这对本发明中的高通量筛选方法有利, 因为非整合状态减少不同转化子在基

因表达水平上的差异。此外，因为导入的 DNA 不重组到宿主 DNA 中，宿主基因组不会发生非目的突变。借助于载体如 AMA1 的自主复制或真菌端粒序列促进下的自主复制，外源基因得以均一表达。（参见 A. Alesenko 和 L.Ivaniva, *Mol. Gen. Genet.* 1998 260:159-164.）

5 本发明中使用的术语“异源蛋白”一词，是指不能在本发明的宿主菌株中正常表达或分泌的蛋白质或多肽。异源蛋白既可来自原核生物，也可来自真菌、植物、昆虫、或来自诸如哺乳动物的高等动物。对于药物筛选目的，则人蛋白是首选，因此一个优选的实施方式是 DNA 文库属于人源的宿主。因此这样一些实施方式也被认为是本发明的适当实施例。

10 用本发明的丝状真菌表达人的基因，即来源于人基因组文库的基因，有许多理由期望得到有效表达。现在知道人的基因平均大小为 3000-5000bp，人的内含子平均为 75-150bp（总范围分布在 40 至 >50000），丝状真菌的内含子为 40-75bp，但它们能处理碱基数高达 500bp 的内含子。人每个基因平均有 3-5 个内含子(M. Deutsch, M.Long, *Nucl. Acids Res.* 1999
15 27:3219-3228；见表 B)。已知人的信号序列在真菌中也能正常工作。基于上述理由，很可能大量人基因可用本发明的方法高水平表达和分泌。

表 B

生物体	每个基因内含子数	平均内含子大小 (nt) (范围)	内含子结构
20 动物 / 植物	3-5	75-150 (40->50000) 150nt 以下有 80%	GTnnGt.....CtxAC.....yAG
真菌	3	40-75 (40-500)	GTAnGy.....CtxAC.....yAG
25 酵母	0.01	50-60 (?-?)	GTATGT..TACTAAC..yAG

本发明的方法有望对来自原核和真核基因组的 DNA 文库的表达有所帮助。如上所述，该方法可以用来表达和检测分泌及细胞内蛋白，而且
30 适应于多种基因和蛋白的情况。

本发明的另一方面包括构建和筛选真菌突变体库以及通过本发明方法构建的真菌突变体库。可以应用本领域技术人员熟知的方法，利用任何整合或非整合方法，转化本发明的真菌宿主得到上述突变体库。通过优选宿主菌株形成的真菌库可以用微量和/或高通量模式筛选方法处理和
5 筛选外源蛋白的所需活性或特性。目的特性或活性意味着与文库中基因的外源蛋白有关的任何物理的、生理化学的、化学的、生物学的或催化特性，或所述特性的任何增加、减少及改进。该文库还可以筛选由于外源蛋白和/或内源蛋白的存在而产生的代谢物、或与代谢物有关的特性或活性。该文库还可以用来筛选蛋白或代谢产物的量增加或减少的真菌。

10 在本发明的再一方面，转化的真菌文库可以用来筛选具有所需特性的真菌代谢物的存在。例如，这些代谢产物包括聚酮化合物、生物碱、和萜类天然产物。多基因或基因簇（操纵子）预期能转导入本发明的宿主细胞中，而且不会因为编码的酶的活性而在宿主细胞中产生非蛋白产物。例如，已经公开编码生产 lovastatin 所必需的蛋白的 DNA 就能转入
15 *Aspergillus oryzae*(US 5362638；也可参见 US 5849541)。

在本发明的另一个实施方式中，真菌转化文库可以用来筛选与目标核酸探针杂交的 DNA 的存在。在这个实施方式中，外源蛋白的表达和/或分泌并非必需的，尽管这通常是期望的。其中蛋白表达不是必须的，可以理解载体上的调控序列也不是必须的。

20 在本发明的再一个实施方式中，真菌转化文库可以用来筛选真菌本身的某些期望特性，如真菌对物理、化学极端环境条件的耐受性或产生、修饰、降解或代谢目的物质的能力。此类期望特性也许归功于单个外源蛋白的存在、也许没有关系。当被用作定向进化过程的一部分时，该实施方式将会非常有用。

25 异源 DNA 可以是利用多种已知的方法从生物样品制备的基因组 DNA 或 cDNA。生物样品可以是环境中取样（如混合肥料、森林垃圾、土壤、海水或新鲜水），或通过提取、过滤或离心或其它富集方法得到的样品。从环境中得到的样品进行的微生物混合培养也能拿来应用。生物样品也可以来自单个物种的生物，如培养的微生物，或植物，昆虫，或
30 诸如哺乳动物的其它动物。另外，外源 DNA 可以是合成的或半合成的，

如随机 DNA 片段或经过重组的、突变过的或其它方式改变过的 DNA 片段。一个半合成核酸库的例子参见 Wagner 等, WO 00/0632。从环境中取
样(或其混合培养物)的 DNA 对发现新蛋白十分有利,而从单一物种得
来的 DNA 则有下列优点(1)更加明智地选择合适的载体;和(2)如果
5 已经鉴别了一种感兴趣的蛋白,操作者可以直接对相关或类似的物种进行深入的筛选。

和传统的真菌宿主相比较,呈现致密丝状形态的 UV 18-25 的金孢属
菌株的转化、表达和分泌率惊人地高。根据本发明,重组菌株最好为这种
形态。然而本发明也覆盖呈现这类特性的非重组菌株或通过其它方式
10 得到的工程菌。本发明的另一吸引人的方面是,使用在相应或等同培
养条件下比 NG7C-19 菌株、最好比 UV18-25 粘性低的金孢属菌株。我们测量
了 UV18-25 培养物的粘度低于 10cP,相比之下,已知 *Trichoderma reesei*
的粘度为 200-600cP,而传统的 *Aspergillus niger* 在最理想的培养条件下的
发酵中期和晚期,粘度为 1500-2000cP。因此,本发明可以使用具有低
15 粘性的任何工程或突变丝状真菌,如金孢属的 UV18-25 (VKM F-3631D)
菌株,木霉属 X252 菌株, *A. sojae* *pc1A* 菌株(源自 ATCC 11906)或 *A. niger*
pc1A。

丝状真菌培养物的流动性可以在从近乎半固体到自由流动的液体的较
宽范围内变化,粘性可以用 Brookfield 旋转粘度计、运动粘度试管、落
20 球式粘度仪或杯状粘度仪轻而易举地进行定量。培养液为非牛顿型发酵
液,总体粘性在某种程度上和收率相关(Goudar 等, *Appl. Microbiol.*
*Biotechnol.*1999 51:310-315),本发明应用的低粘性培养物这种效应非常
不明显。

应用这种低粘性的培养物对筛选本发明方法的表达文库十分有利。丝
25 状真菌中表达的 DNA 文库的筛选曾受相对低效和烦琐的方法所限制。通
常,真菌一旦转化(和任意选择的转化体),就必须制备孢子或分生孢子,
或机械地打断菌丝体,以便将转化的真菌分散为单个生物,进一步培养
成克隆化克隆或培养物。然后稀释孢子、分生孢子或菌丝体,在标准平
板上铺开,对单个克隆进行颜色、对底物变化的检查、验证要寻找的蛋
白特性或活性导致的其它可检测的指标。另外,可以将分泌蛋白印迹到
30

膜上，然后用对膜上因存在目的蛋白活性或特性而产生的标记进行搜索或检测。应用膜的方法对那些存在降解问题的外源蛋白十分有用 (Asgeirsdottir 等, *Appl. Environ. Microbiol.*1999,65:2250-252)。这些操作的劳动强度大，尚未证明易于实现自动化，因此，对传统丝状真菌中真菌表达蛋白的检测尚未实现高通量检测。本申请中，高通量筛选就是指用部分、或全部自动化的筛选方法，即所述筛选方法一天能评估 1,000 个或更多转化体所表达的蛋白，优选指一天能评估 5,000 个或更多转化子所表达蛋白的方法，更优选一天能评估 10,000 个或更多转化子所表达蛋白的方法。

10 根据本发明，对转化真菌文库的高通量自动化筛选，可以用许多已知方法进行。对细菌或酵母适用的方法通常也适用于本发明的低粘性真菌。低粘度真菌通过可转移的复制单元的作用，使要用的突变菌丝体变得相对不再缠绕。在自动高通量筛选的机械操作时，突变的真菌和/或它们的可转移的复制单元行为方面在本质上和单个细菌或酵母类似。这和野生型真菌、许多工业上驯化的真菌产生缠绕的菌丝、彼此之间不能分离的情况形成鲜明的对比。

例如，根据本发明，转化真菌的稀释悬浮液可以通过机械微量加样器加到 96 孔微量反应板上。能向 384 孔和 1536 孔反应板加样的液体装置也可以用来向微量反应板添加微生物。悬浮的微生物的浓度可以根据需要调整，以便使每孔微生物（或其它的可转移复制单元）控制在平均数量上。最好反应孔内有多个微生物，对孔内的物质进行稀释、再培养为单个克隆化克隆或培养后，然后鉴别孔内目的蛋白的活性和性质。这样以来，系统的总处理能力就会增加，但以需要对每孔的物质进行鉴定而增加的“干扰”为代价。

25 本发明另外还包括，在液体流通途径中，可以加进去一个细胞分选器，根据分拣室的检测结果将培养物流向微量反应板不同的孔内。这可以准确地将一个细胞分到一个孔内，对鉴别转化子十分有用，因为本方法允许荧光激活的细胞分选器自动选择转化子。

在另一个实施方式中，固体培养基上生长的克隆可以用自动菌落挑选器挑选菌落，并且将微生物转移到微量培养板。每孔只有单个的菌落

生长。

分散的微生物在微量反应板上长成克隆培养物。诱导剂、营养物等用自动液体分流系统按要求添加。这套系统也可用于添加检测目的蛋白的活性或特性所需要的试剂。例如，添加发光或荧光底物可用于分光光度检测或荧光检测酶的活性。微孔中培养物的低粘性和液面下生长的习性使试剂在培养物中扩散很快，大大地增加了分析的灵敏度和可靠性。氧和氮的扩散也得到了加强，保证了菌体的快速生长、外源蛋白表达和分泌的增加。某些分析方法，如闪烁亲近测定法，依赖于液体成分的扩散才能达到平衡状态；本发明中真菌的低粘性才能使高通量的检测成为可能。最后，对一个高度自动化系统，希望具有自动挑选菌落、供气、从微量反应板微孔中将目的克隆吸出的功能，而低粘性和在液面下生长习性才使这成为可能。就传统的丝状真菌培养物的粘性，尤其是培养物在微量培养板上长成搅不动、剪不开的一团的问题，使上述操作十分困难或者根本不可能。

在另一个实施方式中，单个细胞通过微量液体装置时，可就感兴趣的特性或活性进行光学检测（Wada 等,WO99/67639）。低粘性对微量液体设备是必须的，本发明的培养物的低粘度就是为了适合微量液体操作。Short 等在美国专利 6,174,673 中描述了如何利用荧光底物检测感兴趣的酶的活性，及如何利用宿主细胞表达的此活性来用荧光激活的细胞分选器分离细胞。本发明的方法和这种表达蛋白的鉴别方法是一致的。

本发明中，当转化子携带有荧光蛋白作为标记时，荧光可以定量，也可以作为基因表达或/和特定培养中表达蛋白的度量。在该实施方式中，即可以检测感兴趣的外源蛋白，也可以估计特定蛋白的活性，正如 Blyna 等在 WO 00/78997 中描述的那样。当本发明的筛选方法用作定向进化过程的一部分时，该实施方式将特别优选。

在较大粘性可以接受的情况下，培养真菌时用胶形成基质具有一定的优势，正如 Bochner 在美国 6,046,021 专利中描述的那样，在微量板上进行生化分析。

另一类筛选方法为光度测量分析，利用数字成像分光光度法，对生长在固体物质上的许多单个菌落进行筛选。参见 Youvan 等在 1994,*Meth.*

*Enzymol.*246:732-748 所述。在此方法中，特定试剂的总吸光值或发射光谱的变化代表了异源目的蛋白活性或性质的存在。本发明的变异真菌具有弱侧向生长趋势，在固体培养基上形成光滑的、致密且轮廓清晰的菌斑，对这种筛选系统有利。再者，和细菌相比，真菌的较好的表达和分泌特性可以提供较大数量的蛋白供光度分析。

在公开号为 11-304666 的日本专利申请中，描述了一种自动处理微生物的装置。该装置能转移含单个细胞的微滴，鉴于本发明真菌的形态，有望用该装置对单个菌落进行微量操作。

Beydon 等在 *J.Biomol.Screening* 5:13-21(2000)一文中描述了一种微生物自动高通量筛选系统。该自动系统可转移 400 nl 的液滴到琼脂平板上，每小时处理 10000 个筛选点，而且已经用于对酵母的双杂合子筛选。本发明的真菌宿主和酵母一样，预期可以应用该类型系统进行高通量筛选。

作为微量反应板的替代，转化子也可以以微菌斑的形式生长在平板上，按 WO 00/78997 描述的方法进行光学分析。

Jayawickreme 和 Kost 在 *Curr. Opin. Biotechnol.*8:629-634(1997)文章中探讨了高通量筛选技术的进展情况。而对很少有人谈到的差异表达基因的高通量筛选技术见 von Stein 在 *Nucleic Acids Res.*35:2598-2602 中的论述。

金孢属 UV18-25 菌株和木霉属 X252 菌株和本发明的要求相当吻合。然而本发明也可以使用其它的突变或利用其它方法得到的、可在悬浮状态下产生可转移性复制单元及培养粘性低的工程丝状真菌。真菌的特定形态并不关键，本发明人曾发现，这两种菌株具有短的、不互相缠绕的菌丝体，然而呈现如致密、菌丝分枝较多的形态，也能降低粘性。根据本发明，中性 pH 和 25-43℃ 的温度是真菌生长的理想条件。这种筛选条件对保持外源蛋白、尤其是对那些易降解和被酸性 pH 灭活的蛋白的活性有利。大部分哺乳动物蛋白，尤其是人类蛋白，进化成在生理 pH 和温度下起作用，筛选人的酶的正常活性最好在這些条件下进行，这将使筛选条件最有利。金孢属菌株严格遵循这一特性，在中性 pH 和 35-40℃ 生长良好，而其它常用的真菌宿主在酸性 pH 生长状态最好，也许因为这

个原因它们不是十分合适。

本发明方法的另一应用体现在“定向进化”过程中，其中产生新的编码蛋白的 DNA 序列、在宿主细胞中表达被编码的蛋白、选择表现所需特征的编码蛋白的序列、突变、和再表达。这个过程经过多次重复，直到得到具有期望特性的蛋白质。此过程包括基因重组、蛋白质工程、易错 PCR、定点突变、组合及随机突变，通过该过程可以得到编码外源蛋白的新 DNA 序列。美国专利 5223409、5780279、5770356 提供了这方面的指导。也可参见以下文献：Kuchner 和 Arnold, *Trends in Biotechnology*, 15:523-530(1997); Schmidt-Dannert 和 Arnold, *Trends in Biotech.*, 17:135-136(1999); Arnold 和 Volkov, *Curr. Opin. Chem. Biol.*, 3:54-59(1999); Zhao 等, *Manual of Industrial Microbiology and Biotechnology*, 2nd Ed., (Demain 和 Davies, eds.) pp.597-604, ASM Press, Washington DC, 1999; Arnold 和 Wintrode, *生物处理技术百科全书: 发酵, 生物催化和生物分离*, (Flickinger 和 Drew, eds.) pp.971-987, John Wiley & Sons, New York, 1999; 以及 Minshull 与 Stemmer, *生物化学当代观点 (Curr. Opin. Chem. Biol)* 3:284-290。

组合突变的应用见 Hu 等, *生物化学(Biochemistry)*. 1998 37:10006-10015。美国专利 5763192 描述了通过随机产生合成的序列、再导入宿主、从中选择具有期望特性的宿主细胞的方法获得新的蛋白编码 DNA 序列的方法。影响人造基因重组的方法包括随机引物重组 (Z. Shao 等, *Nucleic Acid Res.*, 26:681-683(1998)), 推迟延伸过程 (H. Zhao 等, *Nature Biotech.*, 16:258-262(1998)), 异源双链重组 (H. Volkov 等, *Nucleic Acids Res.*, 27:e18(1999))。易错 PCR 是另一种途径 (Song 和 Rhee, *Appl. Environ. Microbiol.* 66:890-894(2000))。

定向进化过程选择方面，有两种广泛使用的方法。一种方法是，使目的蛋白的活性在某种程度上对宿主细胞的生存成为必需。例如，如果必需的活性是在 pH 8 时有活性的纤维素酶，可以将纤维素酶基因突变，然后再导入宿主细胞。以纤维素为唯一碳源培养转化子，逐渐提高 pH 直到仅有少数细胞存活。存活细胞编码纤维素的突变基因，假设在相对高 pH 下编码纤维素酶，经受又一轮的突变，该过程多次重复，直到获得以纤

5 纤维素为碳源在 pH 8 的条件生长良好的转化子为止。酶热稳定突变子也可以通过类似的方法获得，即通过多轮基因突变和高温培养宿主细胞（Liao 等, *Prroc.Natl. Acad.Sci.USA* 1986 83:576-580; Giver 等,*Proc.Natl.Acad.Sci.USA*. 1998 95:12809-12813)。为了达到应用的目的，
5 编码外源蛋白的突变 DNA 序列的获得，可以借助于定向进化中使用的多种方法如基因重组、体内重组或盒突变。

这种方法的主要优点是选择步骤中“最适者生存”的整体平行性质。数以亿计的突变可以不经单独评估、考虑而被剔除，但不太可能总能够把酶活性和宿主的存活联系起来。如期望蛋白的特性是和目的靶向选择
10 地结合，则把结合特性和存活联系起来就比较困难。诸如高温或极端 pH 窘迫下的生存，可能取决于多个因子，选择一种期望的突变是不可能的，而且当宿主因为和突变蛋白性质无关的原因而不能生存时，这种突变还会丢失。

另一种整体平行的“最适者生存”的研究是连续筛选。在该方法中，
15 使用传统的方法如观察生长在指示培养基上的菌落周围的透明带或染色带、比色或荧光测定酶法、免疫分析、结合鉴定法等筛选单个转化体。参见例如 Joo 等的文章 *Nature* 399:670-673(1999)，其中细胞色素 P450 单氧化酶经过多轮突变和筛选进化为不以 NADH 为辅助因子；May 等，*Nature Biotech.* 18:317-320(2000)，乙内酰脲酶的反向立体选择性的选择也
20 是通过类似的方式；Miyazaki 等，在 *J Mol Biol* 297:1015-1026(2000)文章中讲述了如何是得到热稳定枯草杆菌素。

这种筛选途径比单一的“存活筛选”有明显优势，尤其是如果它能够以接近整体平行“存活筛选”技术的高通量模式筛选时，则更显出其优势。例如，通过应用以下测量引入平行度数，如转化生物体的数字图像(Joo
25 等,*Chemistry & Biology*,6:699-706(1999))或菌落的数字光度分析（Youvan 等,1994,*Meth.Enzymol.* 246: 732-748)。一种高通量筛选的成熟方法涉及用商业上可获得的读板机，自动评价微量滴定板上的表达蛋白，而本发明方法非常适用于应用该高通量筛选模式筛选“定向进化”。

在本发明的这一实施方式中，编码目的蛋白的基因通过所有已知的方法
30 产生多种突变子，编码突变蛋白的 DNA 通过合适的表达载体转入到低

粘性丝状真菌中,可随意地筛选和培养转化子。宿主细胞如前所述散布到微量滴定板的孔中,或空间上分布到可处理的位置,以便产生单个的单克隆培养(或含有少于100个不同克隆的多克隆)。细胞最好分布在微量滴定板的孔中。突变DNA编码的蛋白质最好分泌到微量滴定板的加样孔中。对每个培养物进行目的蛋白的检测,对那些呈现需要特性的培养物进行选择。选择的培养物编码目的蛋白的基因再行突变,把突变的DNA导入低粘性的丝状真菌宿主,再对转化子进行筛选。突变和再筛选过程重复多次,直至感兴趣的特性到达期望的水平为止。

再一个实施方式中,可通过突变和目的基因在另外一种生物中如在大肠杆菌扩增而进行“定向进化”,然后根据本发明把突变的基因转入丝状真菌,再进行筛选。

本领域技术人员容易理解,一种基于筛选方法似乎是人们感兴趣的蛋白质,不一定具有商业应用所需的所有其它特性。例如,具有酶的活性,尽管特异活性较高,不一定预示其具有必备的热或pH稳定性,或变性剂或蛋白酶抗性,或无免疫原性,或其它商业方面有活力的产品期望应有的或必需的特性。需要简单的方法来判断一种已鉴别的蛋白是否具有商业方面有用的特性。

现有技术的筛选方法不能对上述问题提出解决方案,因为宿主生物(细菌或酵母)对生产可分离量的蛋白不适应。因此有必要把有潜在用途的基因从一种生物转向另一种生物,如把一个基因进行文库建立、基因表达、筛选、批量生产基因产物的研究、直至寻找工业生产合适的高效表达菌株。另一方面,本发明的突变的丝状真菌,是外源蛋白出色的高效表达和分泌菌株,尤其是使用本发明的载体时。分离充分的蛋白质目的不仅在于鉴定,而且还在于评价应用的实验。事实上,本发明筛选方法中应用的菌株也适于工业生产,因为它们拥有所需的生产特性例如低粘度、高表达率、和非常高的蛋白/生物量比率。

因此,在本发明的一个优选实施方式中,该方法进一步包括:在外源蛋白库(或其前体)能够表达和分泌、并且能够提取随后产生的蛋白以获得目的蛋白的条件下,培养按照本发明方法鉴定的克隆菌落或培养物。文库蛋白的表达和分泌可通过被克隆基因和异源蛋白的基因表达调节序

列如和其相应的信号序列的框内融合、或和第三个蛋白的信号序列而得到加强。通过这种途径可产生在文库蛋白的上游与外源蛋白相融合的融合蛋白。然后，可以用已知的纯化技术来分离、回收融合蛋白前体物。当载体被基因工程操作，使蛋白酶酶切位点连接一种分泌良好的蛋白载体和目的蛋白后，可以用 Kex-2、Kex-2 类蛋白酶或其它选择性蛋白酶进行该酶切步骤。

用从前的方法直接筛选宿主生物，是不能获得突变蛋白的。本发明提供的便利在于，突变的目的蛋白可以使用高通量筛选技术进行大量的分离（毫克级），以便供进一步特性分析或做应用实验时使用。本发明的特点是允许工作人员根据蛋白期望特性而不是高通量筛选出的一个特性，筛选突变蛋白为下一轮的定向进化做准备。可能由本发明制定的较为严格的筛选标准，将导致更有效和花费较少的“定向进化”。

本发明的重组突变丝状真菌的生产方法包括，将具有表达调控的序列和控制一套编码异源蛋白的 DNA 序列连接起来后转入本发明的低粘性突变丝状真菌内。DNA 序列转导可以采用目前已知的转化丝状真菌本身的所有方法。这些成熟的方法如 CaCl_2 聚乙烯乙二醇刺激真菌原生质体的 DNA 吸收 (Johnstone *et al.*, *EMBO J.*, 1985, 4:1307-1311)。原生质体转化方法在实施例中有描述。另外已知的原生质体或球形体的转化方法可以使用上述已介绍的其他丝状真菌的转化方法。多拷贝异源 DNA 整合到真菌基因组的适合载体众所周知；参见 Giuseppin *et al.*, WO 91/00920 的例子。很早以来就知道自主复制质粒是有效的真菌转化工具 (Gems 等, *gene* 1991 98:61-67; Verdoes 等, *Gene* 1994 146:159-165; Aleksenko 和 Clutterbuck, *Fungal Genetics Biol.* 1997 21:373-387; Aleksenko 等, *Mol. Gen. Genet.* 1996 253:242-246)。这些方法的详细介绍可在引证的参考文献中查找，它们引入本发明作为参考。

本发明的方法实施例包括使用低粘性金孢属或 *A. sojae* 突变株作为携带异源 DNA 的载体转入的起始物，这些方法实施例将在下面描述。

实施例

30 A. 致密生长形态突变株的研制

许多专利申请中讲述到可以用不同的筛选方法分离形态突变子。WO 96/02653、O97/26330 描述了展现致密形态的非局限性突变子。发现 *A. sojae* 的一种蛋白原加工异常的突变株具有以外的异常生长表现型（高度分枝），但没有对蛋白产生损害效应。该株的培养试验发现，其呈现十分致密小球状的生长表型。这种观察到的特性不仅存在 *A. sojae* 而且也存在于 *A. niger* 等其它的突变真菌中。

(1) *A. niger* 蛋白原加工突变子的构建

从 *A. niger* 中克隆蛋白原转换酶基因，使用了 PCR 方法。通过对多种酵母株及真核生物的不同蛋白原转换酶基因的比对，设计了不同的简并性 PCR 引物，简并性分别为 4, 2, 2, 512, 1152, 4602, 2048 和 49152。使用 PE4 和 PE6 PCR 引物，得到了两个克隆，编码的蛋白序列和 *S. cerevisiae* KEX2 的序列高度同源，这些克隆用于进一步的实验。

根据其它克隆的 PCR 片段的蛋白原转换酶基因观察到的同源性，*A. niger* 的相应基因标记为 *pclA* (原蛋白转换酶样)。Southern 印迹分析表明 *pclA* 基因是单拷贝基因，在 *A. niger* 基因组中没有密切相关的基因，即使在不同的杂交条件下（50℃，60×SSC 条件下洗涤）也没有明显的额外杂交信号。对 *A. niger* N401 (van Hartingsvelt 等, *Mol. Gen. Genet.* 1987 206:71-75) 的 EMBL3 基因组文库的第一轮筛选，尽管筛选了约 10-20 个基因组等同物，也没有产生任何阳性的杂交斑。对 *A. niger* N401 的第二轮筛选，分离到了 *A. niger* 基因的基因组全长拷贝 (Goosen 等, *Curr. Genet.* 11:499-503(1987))。

通过对大小相当于 5-10 个基因组的第一轮筛选，8 个阳性杂交斑中，6 个仍为阳性。正如基因组 DNA 一样，PCR 产物的 EcoRV 的 3 和 4kb 杂交片段 (PCR 片段中包括 EcoRV 限制性位点)，所有 6 个克隆很可能都携带有全长的 *pclA* 基因，通过和其它的蛋白原转化酶的大小相比较，这些片段包含 *pclA* 全长基因的 3' 5' 两侧的序列。两个 EcoRV 片段和相交迭的 5kb 进行亚克隆以进行进一步的特性分析。

根据限制性酶谱，EcoRI 和 EcorRV 亚克隆测定了 *A. niger pclA* 完整基因 DNA 序列。序列分析表明一个阅读框架和 *S. cerevisiae* KEX2 基因及其它蛋白原转换酶具有相当的类似性。根据比对，在编码区发现两个

可能的内含子序列。用内含子两侧的引物，对 pEMBLyex 建立的 *A. niger* cDNA 文库随后进行 PCR 分析，只有两个序列的 5' 最末端代表真正的内含子序列，编码的 *pclA* 蛋白的一般结构很明显和其它蛋白原转化酶类似。*pclA* 蛋白和其它蛋白原转化酶的总类似性为约 50%。

- 5 为了证实克隆的 *pclA* 基因是一个编码功能蛋白的功能基因，尝试构建了没有 *pclA* 基因的菌株。构建了 *pclA* 缺失载体 pPCL1A，其中大部分 *pclA* 编码区域被 *oryzae pyrG* 选择标记所取代。其后，用此载体的 EcoR I 的 5kb 插入片段用来转化不同的 *A. niger* 菌株。

- 10 从这些转化（基于 *pyrG* 选择）中，获得了许多转化子。有趣的是，部分转化子（变化范围是 1-50%）出现了一种十分明显的变异表型（表 13）。对野生型和变异型的转化子 Southern 分析表明，表现为十分严重的限制（致密）变异生长表型突变子，丢失了 *pclA* 基因。所有表现为正常野生型生长的菌株，都携带一个替代的片段并整合到野生型 *pclA* 基因附近或非同源位置上。

15 (2) 构建 *A. sojae* 蛋白原加工突变子

- 为了构建 *A. sojae* 的相应的突变子，用 *sojae* ATCC 11906 粘粒文库转化 *A. sojae pclA* 突变子，进行了低粘性突变子 *A. sojae* 的功能互补分析。从产生的 *A. sojae* 互补突变子中，分离了基因组粘粒文库，包含 *A. sojae* 的蛋白加工蛋白酶 *pclA*。分离到的序列部分序列分析证实克隆到了
- 20 *A. sojae pclA* 基因。根据克隆的 *A. sojae pclA* 基因序列，用我们在其他实施例中反复描述过的方法，使用 WO 01/09352 描述的可再次使用的 *pyrG* 选择标记，构建了基因替代载体。另外还构建了携带 *pyrG* 选择标记和 *A. sojae pcl* 基因 5', 3' 端不完整片段的基因阻断载体，利用基因替代载体和基因阻断载体用于构建 ATCC11906 及 ATCC11906 衍生物 突变子。
- 25 利用部分产生的突变子作培养实验，发现了改进的形态特征，尤其是致密生长和微沉淀方面。

(3) 分离其他 *A. sojae* 的致密生长突变株

- 可用携带真菌筛选标记的线状 DNA 片段进行 *A. sojae* ATCC 11906 及其衍生物的转化，如果没有特异的复制序列，用此方法得到的转化子宿主细胞基因组中整合有导入的 DNA 片段。如导入的选择标记是异源的
- 30

(*A. sojae*), 只发生同源重组, 导致基因组不同位置携带有标记的 DNA 的多种转化子。这种整合倾向于对内源 *A. sojae* 序列产生损坏, 因而可产生一系列 *A. sojae* 突变株。这种情况通过对携带 *A. sojae* *pyrG* 选择标记的 *A. sojae* ATCC 11906 *alpA pyrG* 的分析得到了证明。总共分析了数千个转化子, 其中 5-10 个表现为形态异常突变子, 部分具有和 *pclA* 类同的表性, 和克隆 *A. sojae* *pclA* 基因类似, 可以通过形态表性的互补的方法, 从 *A. sojae* 基因文库中分离和此突变有关的基因。根据已克隆的基因, 可以产生相应的基因阻断/缺失突变子。

(4) 分离金孢属致密生长突变子

10 利用一种类似基于 PCR 的克隆形式 (已在 *Aspergillus niger* *pclA* 基因克隆中描述), 从金孢属 BLUESTAR(TM) 基因文库中克隆了金孢属蛋白原加工基因 *pcl1*。携带有完整基因组基因拷贝的基因片段在 BLUESTAR(TM) 克隆中得到了亚克隆。根据得到的亚克隆, 象构建 *A. sojae* 载体那样, 也获得了一个基因阻断载体。利用重复的侧翼拷贝 *A. sojae* 基因代替 *pyrG* 来作为金孢属的选择标记。金孢属的基因阻断转化导致产生

15 了致密生长表现型的菌株。

B. 粘度测定

下列操作参数数据范围是利用 5 种不同的真菌微生物用作真菌发酵而得来的。这五种真菌为 *Aspergillus niger*, *Trichoderma longibrachiatum* 18.2KK(*T. reesei*), *Trichoderma longibrachiatum* X252, *Chrysosporium lucknowense* UV18-25 和 *Aspergillus sojae* *pclA*。真菌培养物的粘性在发酵过程中是变化的, 而且随营养物质浓度变化而变化。现在报道的测量数据, 培养基中的碳水化合物碳源含量为 20-100g/l (即纤维素酶、乳糖、蔗糖、木糖和葡萄糖等), 和真菌一起培养, 允许培养基通过 200 rpm 震荡培养, 开始用发酵容器为 500-1000rpm 的搅拌器搅拌, 在生长期结束或接近生长期结束时, 一般来说粘性最高。此时培养基补充碳源, 补充速度以碳源浓度不高于 0.5g/l 为准。典型的补充速度为 1-3g/l/hr。

25

在 Brookfield LVF 粘性仪上用微量进样器和 31 号转子 37°C 条件下测量粘性。取 10ml 新鲜发酵液样品, 放置到微量样品转子上。转子速度

30

调到粘性读数为 0-80 范围内。4 分钟后从粘性仪刻度上读取数据。数据乘以下列参数，得出粘性 (centipoise,cp) 水化合物

	转子速度	乘数因子
	6	50
5	12	25
	30	10
	60	5

发酵结束时测定终粘性:

	菌株	终粘性,cp (均数±标准差)
10	<i>T. longibrachiatum</i> 18.2KK	(297± 173)
	<i>A. niger</i>	1500-2000
	<i>T. longibrachiatum</i> X252	≤60
	<i>C. lucknowense</i> UV18-25	≤10
	<i>A. sojae</i> pclA	n.d.

15

C. 金孢属、木霉属和 *Tolypocladium* 的转化

使用的转化培养基如下:

	Mandels 基础培养基:	MnP 培养基
	KH ₂ PO ₄ 2.0 g/l	Mandels 基础培养基添加
20	(NH ₄) ₂ SO ₄ 1.4 g/l	蛋白胨 1 g/l
	MgSO ₄ · 7H ₂ O 0.3 g/l	MES 2 g/l
	CaCl ₂ 0.3 g/l	蔗糖 100 g/l
	痕量元素 1.0 ml/l	调整到 pH 5
25	MnR	MnP Ca ²⁺ :
	MnP +蔗糖 130 g/l	MnP 培养基+
	酵母提取物 2.5 g/l	CaCl ₂ · 2H ₂ O 50 mM
	葡萄糖 2.5 g/l	调整到 pH 6.5
	琼脂 15 g/l	

30

MnR 软培养基: MnR 仅加 7.5 g/l 琼脂

MPC:

	CaCl ₂	50 Mm	pH 5.8
5	MOPS	10 mM	
	PEG	40%	

用于筛选和培养的培养基:

GS:

10	葡萄糖	10 g/l	
	生物大豆酶降解物	5 g/l	
	琼脂	15 g/l	pH 应当为 6.8

PDA:

15	土豆葡聚糖琼脂 (Difco)	39 g/l	pH 应当为 5.5
----	-----------------	--------	------------

MPG:

Mandels 基础培养基添加

	邻苯二甲酸酯 K	5 g/l	
20	葡萄糖	30 g/l	
	酵母提取物	5 g/l	

IC1

	0.5 g/L KH ₂ PO ₄	
25	0.15 g/L MgSO ₄ · 7H ₂ O	pH 7.0
	0.05 g/L KCl	
	0.007 g/L FeSO ₄ · 7H ₂ O	
	1 g/L 酵母提取物 (ohly KAT)	
	10 g/L 蛋白胍或 Pharmamedia	
30	10 g/L 乳糖	
	10 g/L 葡萄糖	

再生培养基 (MnR) 中添加 $50 \mu\text{g/ml}$ 腐草霉素或 $100\text{-}150 \mu\text{g/ml}$ 潮霉素用以筛选转化子。GS 培养基, 再加入 $5 \mu\text{g/ml}$ 腐草霉素用以确认抗生素抗性。

PDA 是快速生长和良好生孢的完全培养基。液体培养基中加入 1/20 的孢子悬浮液 (所有孢子来自 90mm 平板, 体积为 5ml (0.1%Tween))。此类培养基在 27°C 条件下震荡培养 (200rpm)。用两种未转化的金孢属 C1 株和一种 *Trichoderma reesei* 参照菌株检验两种培养基 (GS pH6.8, Pridham 琼脂糖, PA, pH6.8)。为了检验抗生素抗性水平, 从培养 7 天的 PDA 平板上收集孢子。选择的平板在 32°C 条件下培养, 在 2、4、5 天进行分值评价。C-1 株 NG7C-19 和 UV18-25 很明显对腐草霉素和潮霉素具有低的基本抗性, 和参照菌株 *T. reesei* 实验室株水平相当。这清楚地表明这些标准真菌筛选标记可用于金孢属菌株。未发现其他标准真菌筛选标记存在的问题。

在 $50 \mu\text{g/ml}$ 条件下, 成功地筛选出了转化金孢属株的腐草霉素抗性突变子 (Sh-ble)。这也是 *T. reesei* 的筛选水平。由此看来, 对金孢属很容易进行差异筛选。这对 $150 \mu\text{g/ml}$ 水平的潮霉素抗性突变株来说也是一样。

基于最常使用的真菌转化技术, 用原生质体转化技术对金孢属进行了转化。所有 90mm 平板的孢子用 8ml ICl 收集, 转移到 50ml ICl 培养基摇瓶中, 35°C 200rpm 温育 15 分钟。培养物离心后, 沉淀物用 10ml MnP 洗涤, 用 10ml MnP 及 10mg/ml Caylase C_3 悬浮, 35°C 温育 30 分钟并搅拌 (150rpm)。

溶液过滤, 滤过物 3500rpm 离心 10 分钟, 沉淀物用 10ml MnP Ca^{2+} 洗涤。 25°C 离心 10 分钟, 加入 50ml 冷 MPC。混合物置于冰上 30 分钟, 加入 2.5ml PMC, 室温下放置 15 分钟, 加入 500ml 处理过的原生质体, 和 3ml MnR 混合后, 立即铺于含腐草霉素或潮霉素筛选试剂的平板上。 30°C 培养 5 天后对转化子进行分析 (48 小时后, 克隆变的可见)。 $30 \mu\text{g/ml}$ 参照 pAN8-1 质粒来测定转化效率, 结果见表 C。

表 C: 转化效率 (10 μ g 参照质粒 pAN8-1)

	<i>T. reesei</i>	NG7C-19	UV18-25
存活率	10 ⁶ /200 μ l	5 \times 10 ⁶ /200 μ l	5 \times 10 ⁶ /200 μ l
每 200 μ l 中含转化子	2500	10 ⁴	10 ⁴
每 10 ⁶ 个存活细胞中的转化子	2500	2000	2000

结果显示, 金孢属转化子的存活率高于金孢属菌株。菌株的转化率是
5 可比较的, 因此一个实验中金孢属得到的转化子数目比 *T. reesei* 得到的
数目高 4 倍。因此, 金孢属转化系统不是仅仅等同于, 而是优于通常使
用的 *T. reesei* 系统。这种改进对比 pAN8-1 载体转化效率低的载体尤其有
用。用同源的金孢属蛋白编码序列及金孢属转化实验中所用的异源蛋白
编码序列, 构建了许多转化和表达质粒。载体图谱见图 6-11。

10 待表达的同源蛋白选自由金孢属产生的一组纤维素酶, 由家族 6 的葡
聚糖内切酶组成(MW 43kDa), 异源蛋白由属于 *pennicillium* 家族 12(MW
25 kDa) 的葡聚糖内切酶 3 组成。

pF6g 包括金孢属葡聚糖内切酶 6 启动子片段, 该片段和葡聚糖内切
酶 6 开放阅读框架的信号序列相连, 后接葡聚糖内切酶 6 终止子序列。
15 转化子选择采用选择的载体进行共转化。

pUT1150 包括 *Trichoderma reesei* 的细胞生物水解酶启动子, 和葡聚
糖内切酶 6 开放阅读框架的信号序列相连, 后接 *T. reesei* 细胞生物水解
酶终止子序列。另外, 该载体带有一个含选择标记, 即腐草霉素抗性基
因 (Sh-ble 基因) 的第二表达盒。

20 pUT1152 包括 *Aspergillus nidulans* 甘油醛-3-磷酸脱氢酶 A 启动子,
和 *Trichoderma reesei* 的细胞生物水解酶信号肽相连, 后接 *A. nidulans* 氨
基苯甲酸酯合成酶 (*trpC*) 终止子序列。另外, 该载体带有一个含选择标
记, 即腐草霉素抗性基因 (Sh-ble 基因) 的第二表达盒。

25 pUT1155 包括 *A. nidulans* 甘油醛-3-磷酸脱氢酶 A 启动子, 和木霉属
的细胞生物水解酶信号肽相连, 后者和载体蛋白在同一阅读框架内, 后

接 *A. nidulans* traC 终止子序列。该载体应用载体蛋白与目的蛋白融合技术，已知该技术极大改进目的蛋白的分泌。

pUT1160 包括 *A. nidulans* 甘油醛-3-磷酸脱氢酶 A 启动子，并和 *Trichoderma reesei* 的细胞生物水解酶信号肽相连，然后和青霉属葡聚糖内切酶 3 的开放阅读框架、最后和 *A. nidulans* traC 终止子序列相连。

pUT1162 包括 *A. nidulans* 甘油醛-3-磷酸脱氢酶 A 启动子，并和 *Trichoderma reesei* 的细胞生物水解酶信号肽相连，然后 *A. nidulans* traC 终止子序列相连。此外，该载体携带一个含有潮霉素抗性基因（Sh-ble 基因）asa 选择标记的第二个表达盒。

基因组或 cDNA 样品可以很容易地被切割、消化为蛋白编码片段，以及被连接到载体中的片段，例如本发明说明的那些片段，从而产生表达载体库，这对本领域的技术人员来说是显而易见的。采用共转染的方法的可应用性，以及自主复制载体或整合载体可以用上述载体库来转染丝状真菌同样也是显而易见的。

15

表 D: 转化子的比较

载体	菌株	转化	转化数量
pUT1150	UV18-25	在腐草霉素上筛选	285
	<i>T. geodes</i>	在腐草霉素上筛选	144
PUT1152	UV18-25	共转化 pAN8.1	398
	<i>T. geodes</i>	共转化 pAN8.1	45
pF6G	UV18-25	共转化 pAN8.1	252
	<i>T. geodes</i>	共转化 pAN8.1	127
pUT1162	UV18-25	在腐草霉素上筛选	>400
	<i>T. geodes</i>	(n.d.)	

表 D 显示了金孢属 UV18-25 和 *Tolyocladium geodes* 转化的结果。转化操作过程在异源转化章节中叙述。

20

D.金孢属转化子中异源及同源表达

C1 株(NG7C-19 及 UV18-25)用来检验他们分泌的各种异源蛋白的能力: 细菌蛋白 (*Streptoalloteichus hinustanus* 腐草霉素抗性蛋白 Sh-ble)、真菌蛋白 (*Trichoderma reesei* 木聚糖酶 II, XYN2) 和人蛋白 (人溶菌酶, HLZ)。详细方法描述如下:

5 (1) *Streptoalloteichus hinustanus* 腐草霉素抗性蛋白 (Sh-ble) 的 C1 分泌株 C1 株

C1 菌株 NG7C-19 及 UV18-25 用质粒 pUT720 转化 (参考文献 1)。该载体具有以下真菌表达盒:

10 -*Aspergillus nidulans* 甘油醛-3-磷酸脱氢酶 (*gpdA*) 启动子 (参考文献 2)。

-合成 *Trichoderma reesei* 细胞色素水解酶 (*chb1*) 信号序列 (参考文献 1、3)。

-*Streptoalloteichus hinustanus* 腐草霉素抗性蛋白 Sh-ble (参考文献 4)。

-*Aspergillus nidulans* 色氨酸合成酶 (*trpC*) 终止子 (参考文献 5)。

15 该载体也携带 β -内酰胺酶基因 (*bla*) 和 pUT18 (参考文献 6) *E coli* 复制起始点。质粒图谱详见图 2。

C1 原生质体根据 Durand 等 (文献 7) 进行转化以适应 C1 株: 所有来自 90 mm PDA 的孢子收集在 8 ml ICl 中, 转入摇瓶中, 50ml ICl 培养基 35°C 150rpm 搅拌摇动培养。然后将培养物离心, 用 MnP 洗涤沉淀, 20 用 10ml MnP 加 10mg/ml Caylase C₃ 重新溶解, 35°C 150rpm 搅拌温育 30 分钟。过滤溶液, 滤过物 3500rpm 离心 10 分钟。沉淀用 10ml MnP 洗涤, 然后 3500rpm 离心 10 分钟, 沉淀溶于 1ml MnPCa²⁺ 中。向 200 μ l 原生质体溶液中加入 pUT 720 DNA, 室温 (20°C 左右) 温育 10 分钟。再加入 50 μ l 冷 MPC。混合物置于冰上 30 分钟, 同时加入 2.5ml PMC。室温下放置 25 15 分钟后, 500 μ l 处理过的原生质体和 3ml 的 MnR 软培养基混合。马上铺于含腐草霉素选择剂的平板上。30°C 培养 5 天后, 对转化子进行分析 (48 小时后可看到菌落形成)。

C1 转化子 (腐草霉素抗性菌落) 产生 Sh-ble 的分析如下: 最初的转化子用牙签挑出涂布于 GS+腐草霉素 (5 μ g/ml) 平板, 32°C 培养 5 天用于抗性确认。每个真实的抗性菌落亚克隆于 GS 平板, 每转化子的两个亚

30

克隆用 PDA 平板培养，用于收集孢子来进行液体培养。含 ICI 的液体培养 27°C 培养 5 天（200 rpm 振荡）。然后把培养物离心（5000 rpm，10 分钟）。收集 500 μ l 上清。这些样品用 TCA 方法沉淀，用 Western 上样缓冲液重新悬浮于 4 mg/ml 总蛋白（Lowry 法，参考文献 8）。取 10 μ l（约 40 μ g 总蛋白）上样至 12% 丙烯酰胺/SDS 胶，进行电泳（Mini Trans-Blot™ 系统，BioRad Laboratories）。Western 电泳根据 BioRad 的说明进行（Schleicher & Schuler 0.2 μ m 膜），用兔抗 Sh-ble 抗血清（socite Cayla, Toulouse FR, catalog#ANTI-0010）作为主要抗体。结果列于图 1 和表 E。

10

表 E: C1 转化子中 Sh-ble 的估计产量

	Western 印迹估计的 Sh-ble 产量	生产培养基中 Sh-ble 的 估计浓度
未转化 NG7C-19	不能检测	
NG7C-19::720 克隆 4-1	25 ng	0.25 mg/L
NG7C-19::720 克隆 5-1	25 ng	0.25 mg/L
NG7C-19::720 克隆 2-2	250 ng	2.5 mg/L
未转化 UV18-25	不能检测	
UV18-25::720 克隆 1-2	500 ng	5.0 mg/L
UV18-25::720 克隆 3-1	250 ng	2.5 mg/L

这些数据显示：

- 1) 来自 pUT 720 的异源转录/翻译信号在金孢属中是有功能的。
- 15 2) pUT 720 的异源信号序列在金孢属中是有功能的。
- 3) 金孢属可作为宿主用来分泌异源细菌蛋白。

(2) 人溶菌酶 (HLZ) 的 C1 株

用质粒 pUT970G (参考文献 9) 转化 C1 株 NG7C-19 和 UV18-25。
该载体具有下列真菌表达盒:

5 --*Aspergillus nidulans* 甘油醛-3-磷酸脱氢酶 (*gpdA*) 启动子 (参考文献 2)。

--合成 *Trichoderma reesei* 细胞色素水解酶 I (*chb1*) 信号序列 (参考文献 1、3)。

10 --*Streptoalloteichus hinustanus* 腐草霉素抗性基因 *Sh-ble4* 作为载体蛋白 (文献 10)。克隆自质粒 pAN56-2 的 *Aspergillus nidulans* 葡糖淀粉酶 (*glaA2*) 铰链区 (文献 11、12)。

--具 KEX-2 样蛋白酶裂解部位的连接肽 (LGERK)。

--合成的人溶菌酶基因 (*hlz*) (文献 10)。

--*Aspergillus nidulans* 色氨酸合成酶 (*trpC*) 终止子 (文献 5)。

15 该载体也携带有 β -内酰胺酶基因 (*bla*) 及来自 pUC186 *E coli* 复制起始区。详细质粒图谱见图 3。

用实施例 1 中所描述的 pUT970G 质粒来转化 C1 原生质体, 融合蛋白(*sh-ble::GAM hinge::HLZ*)于腐草霉素抗性来说是有功能的, 易于选择转化子。C1 转化子(腐草霉素抗性菌落)的 HLZ 产量用溶菌酶活性分析如下: 初始转化子用牙签涂布于 GS+腐草霉素 (5 μ g/ml) 平板上(抗性确认)及 LYSO 平板(用可见清除区进行 HLZ 活性检测)(参见文献 1, 20 10), 平板在 32°C 条件下培养 5 天。将每个确认的克隆的亚克隆用 PDA 平板培养, 以收集液体培养用孢子。含 ICl 的液体培养基 27°C 条件下 180rpm 震荡培养 5 天后, 将培养物离心 5000g, 10 分钟, 根据 Morsky 等 (文献 13)的方法测定溶菌酶活性。

表 F: C1 株中活性 HLZ 生产水平

	培养基中活性 HLZ 的浓度
未转化 NG7C-19	0 mg/L
NG7C-19::970G 克隆 4	4 mg/L
NG7C-19::970G 克隆 5	11 mg/L
未转化 UV18-25	0 mg/L
UV18-25::970G 克隆 1	8 mg/L
UV18-25::970G 克隆 2	4 mg/L
UV18-25:: 970G 克隆 3	2 mg/L
UV18-25::970G 克隆 2	2.5 mg/L

这些数据显示:

- 5 1)、实施例 1 的 1, 2 两点均得到了确认。
- 2)、Sh-ble 作为抗性标记在金孢属是有功能的。
- 3)、Sh-ble 作为载体蛋白在金孢属是有功能的。
- 4)、KEX-2 样蛋白酶酶解位点在金孢属是有功能的(否则 HLZ 将无活性)。
- 5)、金孢属作为宿主用来分泌异源哺乳动物蛋白。

10 (3) C1 株分泌 *Trichoderma reesei* 木聚糖酶 II (XYN2)

用质粒 pUT1064 和 pUT1065 转化 C1 株 UV18-25。pUT1064 具有下列两个真菌表达盒:

第一个表达盒能够筛选腐草霉素抗性转化子:

- 15 --*Neurospora crassa* 交叉途径控制基因 (cpc-1) 启动子 (参考文献 14)。
- Streptoalloteichus hinustanus* 腐草霉素抗性基因 Sh-ble(参考文献 4)。
- Aspergillus nidulans* 色氨酸合成酶 (trpC) 终止子 (文献 5)。

第二个表达盒是木聚糖酶生产盒:

- T. reesei* 菌株 TR2 cbh1 启动子 (参考文献 15)。
- T. reesei* 菌株 TR2 xyn2 基因 (包括它的信号肽) (参考文献 16)。
- 20 --*T. reesei* 菌株 TR2 cbh1 终止子 (参考文献 15)。

该载体还携带来自质粒 pUC19(文献 6)的 *E. coli* 复制起始点。质粒图

谱详见图 4。

pUT1065 具有以下真菌表达盒：

- Aspergillus nidulans* 甘油醛-3-磷酸脱氢酶 (gpdA) 启动子 (参考文献 2)。
- 合成 *T. reesei* 细胞生物脱氢酶 I (cbh1) 信号肽序列 (参考文献 1、3)。
- 5 -*Streptoalloteichus hinustanus* 腐草霉素抗性基因 Sh-ble4 用作载体蛋白(参考文献 10)。
- 以 KEX-2 样蛋白酶裂解部位为特点的连接载体(SGERK) (参考文献 1)
- T. reesei* 株 TR2xyn2 基因 (无信号序列) (参考文献 16)
- Aspergillus nidulans* 色氨酸合成酶 (trpC) 终止子 (参考文献 5)

10 该载体还携带 β -内酰胺基因 (bla) 和质粒 p18 的 *E.coli* 复制起始点 (参考文献 6)。质粒详细图谱见图 5。

C1 原生质体用质粒 pUT1064 或 pUT1065 按照例 1 的步骤进行转化。质粒 pUT1065 融合蛋白 (Sh-ble::XYN2) 在腐草霉素抗性方面是有功能的，可以借此容易地对 C1 转化子进行筛选。另外，腐草霉素抗性和 xyn2 15 表达水平大体相关。在 pUT1064，xyn2 自己本身的信号序列一起被克隆。

按下列木聚糖酶活性分析方法对 C1 转化子的木聚糖生产进行分析：原初转化子用牙签挑到 GS+腐草霉素 (5 μ g/ml) 平板上 (抗性确认) 及 XYLAN 平板 (参见文献 17)，用木聚糖酶活性的亮区来检验。将每个确认的克隆的两个亚克隆用 PDA 平板培养，以收集液体培养用孢子。含 20 ICl+5g/l K⁺ Phtalate 的液体培养基 27°C 条件下 180rpm 震荡培养 5 天后，将培养物离心 5000g, 10 分钟，根据 Miller 等(文献 18)的 DNS 方法测定木聚糖酶活性。

表 G: (生产木聚糖酶最好的) C1 株中活性 XYN2 生产水平

	培养基中活性木聚糖酶 II 浓度	培养基中木聚糖酶 II 特异性活性
未转化的 UV18-25	3.9 U/ml	3.8 U/mg 总产量
UV18-25::1064 克隆 7-1	4.7 U/ml	4.7 U/mg 总产量
UV18-25:: 1064 克隆 7-2	4.4 U/ml	4.3 U/mg 总产量
UV18-25:: 1065 克隆 1-1	29.7 U/ml	25.6 U/mg 总产量
UV18-25:: 1065 克隆 1-2	30.8 U/ml	39.4 U/mg 总产量

这些数据显示:

- 5 1)、实施例 2 的 1 至 4 点均得到了确认。
- 2)、C1 作为宿主用来表达异源真菌蛋白的分泌。

(4) 总结

10 表 H 显示用质粒转化的 UV18-25 的结果。表中的结果是用外源表达的调控序列和信号序列表达葡聚糖内切酶和细胞色素脱氢酶的水平。也有用同源表达序列和信号序列的相应水平。各种质粒的详细情况可在其它描述和图表中见到。表达条件为 35℃、碱性环境。

表 H: UV18-25 转化菌株的表达数据
(相对于 UV18-25 亲本株的百分比)

培养物	总蛋白量 毫克/毫升	CMCase		β-葡聚糖酶		PH 值
		U/毫升	U/毫克	U/毫升	U/毫克	
UV18-25	100%	100%	100%	100%	100%	7.90
1150-23	94%	105%	111%	140%	149%	7.90
-30	96%	105%	110%	145%	151%	8.10
1152-3	94%	112%	120%	147%	156%	7.85
-4	100%	105%	105%	132%	132%	7.90
1160-2	69%	81%	118%	90%	131%	7.90
-4	73%	72%	98%	83%	114%	8.35
-1	92%	95%	103%	120%	130%	8.45
1162-1	102%	105%	103%	145%	142%	8.20
-11	112%	109%	98%	115%	103%	8.20
F6g-20	104%	102%	98%	130%	125%	7.90
-25	-	-	-	-	-	-

培养条件(摇瓶): 88 小时, 35°C, 230 转/分

5

E. *Aspergillus sojae* 基因文库的建立

(1) 载体库

采用了一种以前描述的方法(Punt, van den Hondel, *Methods Enzymol.* 1992 **216**:447-457), 获得 ATCC 11906 的原生质体, 制备 *A. sojae* 的基因组 DNA。经过分离以后, 用 Kolar 等, *Gene* 1988 **62**:127-34 所述的方法将 DNA 从原生质体中提取出来。随后将 DNA 用 MboI 部分消化, 获得平均大小为 30-50kb 的 DNA 片段。

用于构建基因文库的载体 pAOpyrGcosarp1 是通过将一段 pANsCos1 (Osiewacz, *Curr Genet.* 1994 **26**:97-90) 的 3kb BamHI-HindIII 片段和 pAO4.2(De Rooter-Jacobs 等, *Curr. Genet.* 1989 **16**:159-163) 的 3.2kb Acc65I-

HindIII 片段连接到经 Acc65I-BamHI 酶切的 pHELP1 载体(Gems 等,*Gene* 1991 **98**:61-67)上构建的。这个粘粒载体含有 *A.oryzae* 的 pyrG 选择标记并可以在丝状真菌中自我复制。

MboI 消化的基因组 DNA 被连接到 BamHI 消化的 pAOpyrGcosarp1, 并且该连接混合物被通过 Statagene Supercos1 载体试剂盒 (Statagene Inc. La Jolla CA) 包裹到噬菌体颗粒中。这就产生了总共 30,000 个单克隆, 代表了大约 30 倍的 *A.sojae* 基因组。所产生的克隆混合物的储存物(在 15% 的甘油中)贮存于 -80℃, 供以后使用。

(2) 高频率转化

10 一株 *A sojae* ATCC 11906 的 pyrG 突变株被作为 ATCC 11906 的抗花乳清酸酸的衍生株被选育出来。该菌株, *A sojae* ATCC 11906pyrG, 被两个带有黑曲霉 pyrG 基因的载体转化。一个载体 pAB4-1(van Hartingsveldt 等, *Mol.Gen. Genet.* **206**:71-75(1987))只携带有 pyrG 基因, 而 pAB4-arp1(Verdoes 等, *gene* **146**: 159-165(1994))携带有 pyrG 基因和 *A. nidulans* 的 AMA1 序列。ATCC 11906 pyrG 的转化产生了每毫克 pAB4-1 DNA 5-10 个转化子, 而用 pAB4-arp1 转化频率至少高 10-100 倍。转化子的表型分析显示 pAB4-arp1 化子的 pyrG 的表型只有在持续的选择下才能保持, 而 pAB4-1 转化子的 pyrG 表型无论有没有筛选都是稳定的。这些结果肯定了 pAB4-arp1 转化子中的被转化质粒能自主复制。用其它的带有 AMA1 15 20 序列的真菌转化载体及其衍生质粒如 pAOpyrGcosarp1 也获得了相似的结果。

(3) 真菌转化株库的建立

用一个构建在转化载体 pAOpyrGcosarp1 上的 *A.sojae* 文库转化 *A.sojae* ATCC 11906pyrG 或其相关突变株, 尤其是那些形态紧凑的突变株。这个载体能产生高频率的、载体能自由复制的转化子。将真菌原生质体用 Punt 和 den Hondel 在 *Methods Enzymol.* 1992 **216**:447-457 中所述的方法, 用携带有 *A.sojae* 或金孢属真菌基因组 DAN 粘粒文库中的 DNA 处理, 并将转化的原生质体系列稀释后在选择琼脂平板上涂平板来确定获得的转化频率。剩下小原生质体用选择培养基复壮后贮存在 4℃。根据 30 获得的转化频率结果 (依据的实验中的值可以达到几千个转化子/毫克粘

粒文库 DNA), 将复壮的原生质体的限量稀释物倒在 96,248 孔或其它型号的微孔检测板中, 使每孔有一个转化的原生质体, 将检测板在 35°C 孵育来产生真菌生物量, 获得的转化株库将在后继实验中使用。

5 相似的策略被用于构建一组携带突变的金孢属等位基因的真菌转化株。该策略也能用于构建从其它任意感兴趣的基因衍生的突变株的文库, 不管突变是通过诱变, 基因重组还是基因进化方法产生的。

F. 浸没发酵中孢子产生的介绍

10 许多真菌, 诸如 *Aspergillus sojiae*, 在浸没发酵时不产生孢子。这里我们介绍一种以前不知道的在这些条件下产生孢子的方法。A *sojiae* ATCC 11906 及其尤其是生长形态紧凑的突变株在添加了酵母提取物的合成生长培养基中生长。在这种条件下, 无论是静置培养还是搅拌培养生物量都会迅速积累。但是在培养液中没有孢子产生。在相似的培养基中加入 0.6 克/千克 EDTA 后, 35°C 培养 2-4 天, 会导致大量孢子的产生, 最多能达到每毫升培养物 10^9 个孢子。

合成培养基 (+/-EDTA)

克/千克培养基

	KH_2PO_4	2.5
	NH_4Cl	7.2
20	$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	0.7
	$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$	0.2
	酵母提取物	20
	$\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$	0.015
	$\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$	0.005
25	$\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$	0.016
	$\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$	0.040
	H_3BO_4	0.005
	KI	0.003
	$\text{MnCl}_2 \cdot 2\text{H}_2\text{O}$	0.012
30	$\text{Na}_2 \cdot \text{MoO}_4 \cdot 2\text{H}_2\text{O}$	0.003

EDTA (0.6 或 0.0)
pH 用 NaOH/ H₃PO₄ 调整到 5.5

G. 金孢属和曲霉属的转化系统

5 (1) *A.niger*(黑曲霉)乳清酸磷酸核糖基转移酶基因 *pyrE* 的克隆

许多多用途的丝状真菌转化系统是采用尿嘧啶营养缺陷型突变株为基础的。这些突变株或者是尿苷-5-磷酸去羧基酶缺陷型(OMPD), 或者乳清酸磷酸核糖基转移酶缺陷型(OPRT)。(T. Goosen 等, *Curr Genet.* 1987, 11:499-503; J. Begueret et al., *Gene*, 1984 32: 487-92.) 先前我们已经
10 分离出了黑曲霉 *OMPD* 基因 *pyrG* (W. van Hartingsveldt 等, *Mol. Gen. Genet.* 1987 206:71-5)。黑曲霉 *OPRT* 基因(*pyre*)的克隆是通过一株 *FOA* 抗性、尿嘧啶营养缺陷型、非 *pyrG* 突变株的互补进行的。为了互补采用了一个构建在 *pAOpyrGcosarp1* 载体上的黑曲霉粘粒文库。从互补的转化子中, 基因组粘粒克隆被分离出来, 这些克隆带有互补的黑曲霉基因, 现在起
15 名叫 *pyrE*。将一段带有 *pyrE* 基因的 5.5kb *SstII* 酶切片段克隆到 *Pbluescript*TM(Stratagene), 产生载体 *pBLUEpyrE*。将该载体的一段含有 *pyrE* 基因的 1.6kb 片段进行测序, 以确认 *OPRT* 基因的位点(参见表 15)。

(2) *Chrysosporium lucknowense* 的营养缺陷型转化系统

尿嘧啶缺陷型的 *Chrysosporium lucknowense* 菌株被作为 *C1* 和
20 *UV18-25* 的花乳清酸抗性突变株, 所采用的方法如 *PCT* 公布的 *WO* 01/09352 中所述。花乳清酸抗性突变株的筛选能导致两种尿嘧啶缺陷型突变株的分离, 例如, 或者是 *orotidine-5-磷酸脱羧酶(OMPD)* 突变株, 或者是 *orotate-p-ribosyl 转移酶 (OPRT)* 突变株(T. Goosen 等, *Curr Genet.* 1987,11:499-503)。为了检测金孢属的性状, 进行了转化实验, 所用的是
25 能得到的黑曲霉基因 *pyrG* (*OMPD*; *pAB4-1* 载体, W. van Hartingsveldt 等, *Mol. Gen. Genet.* 1987 206:71-5)和 *pyrE* (*pBLUE-pyrE*; *OPRT*)。如表 I 示, 只有对突变株的 *pyrE* 基因转化才产生原养型转化子, 暗示这些金孢属菌株是 *OPRT* 突变株。根据我们所采用的金孢属基因命名法, 这些突变株被命名为 *pyr5*。

表 I

基因	来源	载体 ¹⁻⁴	UV18FOA ^R #4	C1#B
OMPD (PyrG/pyr4)	Aspergillus niger	pAB4-1	-	-
	Aspergillus oryzae	pAO4-2	-	-
	Neurospora crassa	pDJB3	-	-
OPRT (pyrE/pyr5)	Aspergillus niger	pBLUEpyrE	+	+

1.pAB4-1: W. van Hartingsveldt 等, Mol. Gen. Genet. 1984 206:71-5

2.pAO4-2: Y. de Rooter-Jacobs 等, Curr. Genet. 1989 16:159-63

5 3.pDJB3: D. Ballance, G. Turner, Mol. Gen. Genet. 1986 202:271-5

4.pBLUEpyrE: 见上

(3) 自主复制真菌转化载体的构建和使用。

基于 pBLUE 载体衍生出两个载体, 其带有的序列使载体在转入丝状
10 真菌时可以自主复制。带有 *Aspergillus nidulans* AMA1 序列(J. Verdoes 等,
Gene 1994 **146**:159-65)的 5.5kb HindIII 切片段, 插入到 pBLUEpyrE 的单
一 HindIII 位点, 得到 pBLUEpyrE-AMA。带有人类端粒序列(A.Aleksenko,
L. Ivanova, *Mol. Gen. Genet.* 1998 **260**: 159-64)的 2.1kb HindIII (部分)酶切
15 片段插入到 pBLUEpyrE 的单一 HindIII 位点, 得到 pBLUEpyrE-TEL。这
些载体转入到曲霉属和金孢属 OPRT 突变株, 得到原养型的转化子。得
到的几个转化子都表现出带有自由复制质粒的转化子所表现出的粗糙表
型特征(J. Verdoes 等, *Gene* 1994 **146**:159-65)。

(4) *Chrysosporium lecknowense* 的转化

此方法来源于用来转化曲霉的方法(P. Punt, C. van den Hondel,
20 *Methods in Enzymology* 1992 **216**:447-457)。在 1 升 Erlenmeyer 摇瓶中, 以

10⁶孢子每毫升的密度使用完全培养基(250毫升),培养 *pyr5Chrysporium* 的突变株(同上),35℃空气摇床(330rpm)中培养24-48小时,通过无菌的 Miracloth™膜滤去菌丝体,并用100毫升1700 mosmol 的 NaCl/CaCl₂(0.27M CaCl₂/0.6M NaCl)清洗,将菌丝体称重,置于冰上,加入 Caylase™(Cayla)(20毫克每克菌丝体)和1700mosmol NaCl/CaCl₂(3.3毫克每克菌丝体),悬浮后置于空气摇床中33℃培养,在显微镜下观察原生质体。培养1-3个小时之后,大部分菌丝体都被消化,在显微镜下几乎都为原生质体。使用无菌的 Miracloth™膜过滤原生质体,使用等体积的冷STC1700(1.2M山梨醇/10mM Tris HCl pH7.5/50mM CaCl₂/35mM NaCl)清洗滤膜。4℃2500rpm离心10分钟,沉淀原生质体。使用STC1700重悬沉淀,再次离心。使用STC1700重悬沉淀后,将原生质体计数。加入STC1700至最终浓度为2×10⁸个原生质体每毫升。

将载体DNA(pAB4-1或pBLUE-pyrE,1-10微克)移入无菌试管底部,加入1μl 1M ATA和10μl原生质体(2×10⁷),实验同时包括负链阴性对照。混匀后室温培养原生质体25分钟,按如下步骤分次加入 PEG6000:250μl,混匀,250μl,混匀,850μl,混匀。溶液置于室温20分钟,在管内加入8ml STC1700,混匀,4℃2500rpm离心10分钟,沉淀重悬于250μl STC1700。样品的溶液用来在选择性培养基上铺板。对pyr⁺选择,平板中含有1.5% Daishin琼脂,1.2M山梨醇,含有硝酸盐的1×AspA,2mM MgSO₄·7H₂O,1×痕量元素,0.1%的酪素氨基酸和1%葡萄糖。如用来选择amdS(和pyr⁺),平板中应含有1.5% Oxoid琼脂,1.2M山梨醇,2mM MgSO₄·7H₂O,1×痕量元素,1%葡萄糖,无硝酸盐的1×AspA,15mM CsCl,10mM 乙酰胺或丙烯酰胺。平板置于30或35℃培养。

在PEG6000处理之前和之后,将孢子和可见的原生质体的STC1700溶液在含有硝酸盐和含有或不含有山梨醇的基础培养基上倒平板,计数。将100μl 10⁻¹,10⁻²,10⁻³的溶液铺在不含山梨醇的平板上,孢子计数。将100μl 10⁻²,10⁻³,10⁻⁴,10⁻⁵的溶液铺在不含山梨醇的平板上,将可见的原生质体计数,转化结果见表I。

H. 蛋白/生物量比率

对于产生纤维素酶和淀粉酶的金孢属、木霉属和曲霉属，其干重是通过下列方法测得：将一定量的完全肉汤培养基通过预称重的滤膜，将滤膜用去离子水清洗，然后在 60°C 烘干过夜，在 100°C 再烘干 1 小时；在分离器中冷却后，生物量通过从干滤膜加滤下的沉淀块的总重减去滤膜的重量再除以培养液的体积来测定。对于木霉属和曲霉属菌株，总的生物量被认为等于总干物质的量，因为在进行测定的时候除了生物量以外几乎没有其它不溶物。对于产生纤维素的金孢属菌株，在培养基中有相当量的纤维素，因此测定生物量要用总干物质的减去纤维素的量纤维素用以下方法测定：

将一定体积的完全肉汤培养基离心以去除固体物质，并弃去上清。沉淀用与初使肉汤培养基等体积的 0.1N NaOH 重悬浮，加上 1/10 体积的 0.5N NaOH。混合物在 65°C 孵育 4 小时。经过这样的处理除纤维素外的所有物质都已溶解。将此碱性混合物冷却理性，弃去上清。将剩下的沉淀用去离子水重悬并离心两次进行洗涤。将洗过的沉淀重悬于去离子水，转移到按上面的方法称重并干燥的盘中。纤维素的浓度通过干重除以所测定的液体体积来检测。

蛋白质的测定是用 Bradford 染色—结合过程(M. Bradford, 1976, *Anal. Biochem.* 72:248)并采用免疫球蛋白标准来测定。在各种丝状真菌中选出的表达蛋白的蛋白质/生物量之比如表 J 所示。

表 J

酶	菌株	克蛋白/克生物量
中性纤维素酶	<i>Chrysosporium lucknowense</i> UV18-25	8.2
中性纤维素酶	<i>Chrysosporium lucknowense</i> UV26-2	6.0
α - 淀粉酶	<i>Aspergillus oryzae</i> 108-318	0.89
葡萄糖淀粉酶	<i>Aspergillus niger</i>	0.78
葡萄糖淀粉酶	<i>Aspergillus niger</i>	1.11
酸性纤维素酶	<i>Trichoderma reesei</i> A-34	0.89
酸性纤维素酶	<i>Trichoderma reesei</i> A-1391	0.65
木聚糖酶	<i>Trichoderma reesei</i> X-252	2.4

I. 绿色荧光蛋白在 *A. sojae* 和 *C. lucknowense* 中的表达和分泌

- 5 作为一个多用和易筛选的报告蛋白的范例, 来自水母 *Aequoria victoria* 的绿色荧光蛋白(GFP)已经在 *A. sojae* 和 *C. lucknowense* 中表达。带有 GFP(*A. Santerre Henrisken* 等, *Microbiology*, 1999, **145**:729-734)和葡萄糖淀粉酶-GFP 融合基因的载体通过将 *glaA* 启动子置换为组成型表达的 *A. nidulans* *gpdA* 启动子来进行改造。采用 *pyrG* 或 *amdS* 选择标记, 用共转化的方法
- 10 将该载体引入 *A. sojae* 中。采用 *pyrG* 或 *amdS* 选择标记, 用共转化的方法将 pGPDGFP 载体及其衍生物转入金孢属菌株中。表达的结果在 *A. sojae* 和金孢属中都产生了强烈的荧光, 证实 GFP 被两个载体都表达了。在表达葡萄糖淀粉酶-GFP 融合蛋白的培养基的上清液中有荧光产生, 表明有荧光活性的蛋白被分泌到胞外。在从各种表达非分泌型细胞质类的
- 15 荧光蛋白的转化子中获得的孢子(或孢子样的无性芽)中也观察到了荧光蛋白的表达。

J. 真菌生长单元的转移

或者通过手工用多道加样枪的方法, 或者用自动滴定板操作系统, 在 96 孔微量滴定板的板孔中装入适当的培养基。体积太大会增加交叉感

染的机会，而为了避免蒸发带来的问题体积也不应太小。例如，如果采用 COSTER™3799 平底滴定板，150 μ l 是一个比较好的工作体积。用牙签将平皿上生长的菌落的孢子转移到滴定板中进行接种。也可以用微量移液器吸取少量的孢子，原生质体或菌丝片段的悬液对滴定板进行接种。

5 这些悬液可以从分离的孢子/原生质体溶液或从小平板生孢子培养基的培养物获得。也可以用单针或 96 针工具从微量滴定板接种。

随后滴定板在 35℃ 进行孵育。为了降低蒸发，可以采用带盖的滴定板，或者可以用允许 O₂, H₂O 和 CO₂ 交换并粘贴于板的表面的膜封住。为了进一步降低蒸发，可以采用可控制空气的孵育器。

10 经过 4 到 5 天的孵育，获得的生物量足够有效地转移到新的盛有新鲜培养基的微量滴定板中。为了准备一个影印滴定板，用了一个 96 针的工具。具有不同菌株排列方式的子滴定板可以通过手工或自动微量移液或用针转移来制备。为了保证在转移针上有可转移的复制单元存在，将带针的工具浸没在微量滴定板的培养基中并振荡 20 秒。随后该工具被小心地从初始滴定板移出，随后在一个新的微量滴定板中作了影印。一个同样有效的转移过程也可以用多道加样枪，转移大约 1 μ l 的父滴定板的培养液来获得。在这两种情况下，由于可转移的复制单元—例如孢子，孢子样的无性芽，原生质体或菌丝或菌丝体的片段的的存在，都能获得有效转移。可以用降解细胞壁的酶进行处理在微量板的板孔中获得原生质体，随后将这些原生质体转移。在微量平板上产生原生质体的方法在 C. van Zeiji 等, *J. Biotechnol.* 1997 59:221-224 中有描述。

15

20

通过将微量滴定板培养物在微量滴定板摇床上 35℃ 孵育，对转移进行了改进。这样增加了培养液中可转移的复制单元的数量。为了储存微量滴定板培养物，加入了终体积 15% 的甘油，并将滴定板储存于 -80℃。

25 在进行后继转移实验时，将滴定板解冻并按上述方法进行转移。*A. niger* 和 *A. sojae* 的野生型或商业菌株的有效转移在这里所采用的条件下不可行，因为这些菌株在培养一天以后会表现出很强的表面生长并产生气生孢子。气生孢子的产生会导致转移时的大范围污染，表面生长会盖住微孔继而使大部分已知的检测方法无法进行。

30

K.用于发现基因的真菌表达文库的构建

基于真菌表达载体 pAN52-1NOT(EMBL 分类号 Z32534)或其一个衍生质粒,构建了一个载体,其 *A. nidulans* *gpdA* 基因(P. Punt 等, *J Biotechnol.* 1991 17:19-33)之组成型表达广真菌宿主启动子的临近下游有一个唯一的 BamHI 克隆位点。由于该载体的构建方式,带有翻译起始密码子(ATG)的基因组 DNA 片段可以在其中表达。为了给这个载体提供一个选择标记,将来自 pBLUEyprE 的 NotI-BamHI 片段被连接到用 NotI-BglII 消化的被称为 pAN52-BamHI 的表达载体上,获得了载体 pAN52-pyrE。将金孢属的基因组 DNA 用 Sau3A 部分消化获得了大小 3-6 kb 的片段。将这些片段连接到 BamHI 消化的 pAN52-pyrE 载体上以后,获得了足以覆盖金孢属基因组几次的重组克隆。大量的这些克隆被混合以覆盖相当于至少 5-10 倍的真菌基因组。制备这些混合物的 DNA 用于转化金孢属 *pyr5* 或 *Aspergillus* *pyrE* 突变株。按上文所述的方法以微量平板的方式制备转化子的集合,用于以后的功能/活性筛选。或者,也可以如 PCT/NL99/00618 中所述,用特殊调控的金孢属启动子构建表达文库。

例子中引用的参考文献

(以下内容,以及上述引用的所有专利和参考文献,一并在这里以参考文献的方式列出)

1. Calmels T. P., Martin F., Durand H., and Tiraby G. (1991), 在处理真菌分泌蛋白质时的蛋白水解现象, *J. Biotechnol.* 17 (1): 51-66.
2. Punt P. J., Dingemans M. A., Jacobs-Meijsing B. J., Pouwels P. H., and van den Hondel C. A. (1988) *A. nidulans* 甘油醛-3-磷酸脱氢酶基因的分离和鉴定, *Gene* 69 (1): 49-57.
3. Shoemaker S., Schweickart V., Ladner M., Gelfand D., Kwok S., Myambo K., and Innis M. (1983) 来源于 *T. reesei* 的外-纤维二糖水解酶 I 的分子克隆, *Bio/Technology* Oct.: 691-696.
4. Drocourt D., Calmels T., Reynes J. P., Baron M., and Tiraby G. (1990) 用

- 于转化低等和高等真核生物以产生腐草霉素抗性的 *S. hindustanus* ble 基因组件, *Nucleic Acids Res.* 18 (13): 4009.
5. Mullaney E. J., Hamer J. E., Roberti K. A., Yelton M. M., and Timberlake W. E. (1985) 来源于 *A. nidulans* 的 *trpC* 基因的一级结构, *Mol. Gen. Genet.* 199 (1): 37-45.
6. Yanisch-Perron C., Vieira J., and Messing J. (1987) 改进的 M13 噬菌体克隆载体和宿主菌株: M13mp18 和 pUC19 载体的核苷酸序列, *Gene* 33: 103-119.
7. Durand H., Baron M., Calmels T., and Tiraby G. (1988) 应用于 *T. reesei* 以筛选纤维素分解能力提高的工业菌株的经典及分子遗传学, J. P. Aubert, Editor. Academic Press. pp. 135-151.
8. Lowry O. H., Rosebrough N. J., Farr A. L., and Randall R. J. (1951) 用福林-酚试剂进行蛋白质检测, *J. Biol. Chem* 193, 265-275.
9. Parriche M., Bousson J. C., Baron M., and Tiraby G. 丝状真菌外源蛋白质分泌系统的开发, 在第三届欧洲真菌遗传学研讨会上, 1996. Munster, Germany.
10. Baron M., Tiraby G., Calmels T., Parriche M., and Durand H. (1992) 真菌 *Tolytlocladium geodes* 中人溶菌酶与 sh-ble 抗腐草霉素融合蛋白的高效分泌, *J. Biotechnol.* 24 (3): 253-266.
11. Jeenes D. J., Marczinke B., MacKenzie D. A., and Archer D. B. (1993) 一个用于从黑曲霉 (*Aspergillus niger*) 中分泌外源蛋白的截短的葡萄糖淀粉酶融合蛋白, *FEMS Microbiol. Lett.* 107 (2-3) : 267-271.
12. Stone P. J., Makoff A. J., Parish J. H., and Radford A. (1993) *neurospora-crassa* 的葡萄糖淀粉酶基因的克隆和序列分析, *Current Genetics* 24 (3): 205-211.
13. Morsky P. (1983) 用 *Micrococcus lysodeikicus* 细胞通过比浊法测定溶菌酶: 反应条件的再检测, *Analytical Biochem.* 128: 77-85.

14. Paluh J. L., Orbach M. J., Legerton T. L., and Yanofsky C. (1988) *Neurospora crassa* 的交叉途径控制基因 *cpc-1*, 编码一个类似于酵母的 GCN4 基因和癌基因 *v-jun* 所编码蛋白的 DNA 结合区的蛋白, Proc. Natl. Acad. Sci. USA 85 (11): 3728-32.
- 5 15. Nakari T., Onnela M. L., Ilmen M., Nevalainen K., and Penttil M. (1994) 在有葡萄糖存在的情况下真菌的启动子是活跃的, WO 94/04673.
16. Torronen A., Mach R. L., Messner R., Gonzalez R., Kalkkinen N., Harkki A., and Kubicek C. P. (1992) 来源于 *Trichoderma reesei* 的两种主要木聚糖酶: 酶及其基因的特性, Biotechnology 10 (11): 1461-5.
- 10 17. Farkas V. (1985) 检测产生纤维素酶和木聚糖酶的微生物的新型培养基, FEMS Microbiol. Letters 28: 137-140.
18. Miller G. L. (1959) 用二硝基水杨酸试剂测定还原糖, Anal. Chem. 31: 426-428.
- 15 19. Punt P. J., Mattern I. E., van den Hondel C. A. M. J. J. (1988) 一个用于曲霉转化并赋予其腐草霉素抗性的载体, Fungal Genetics Newsletter 35, 25-30.

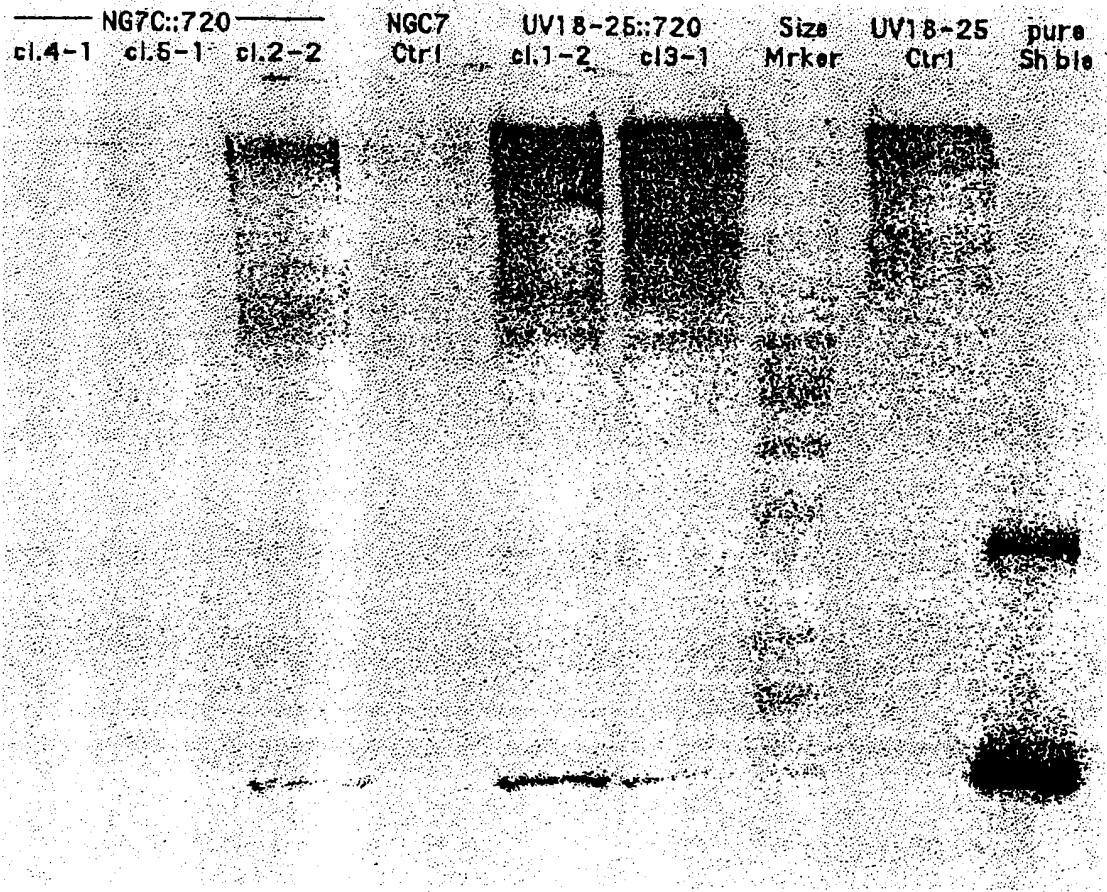


图 1

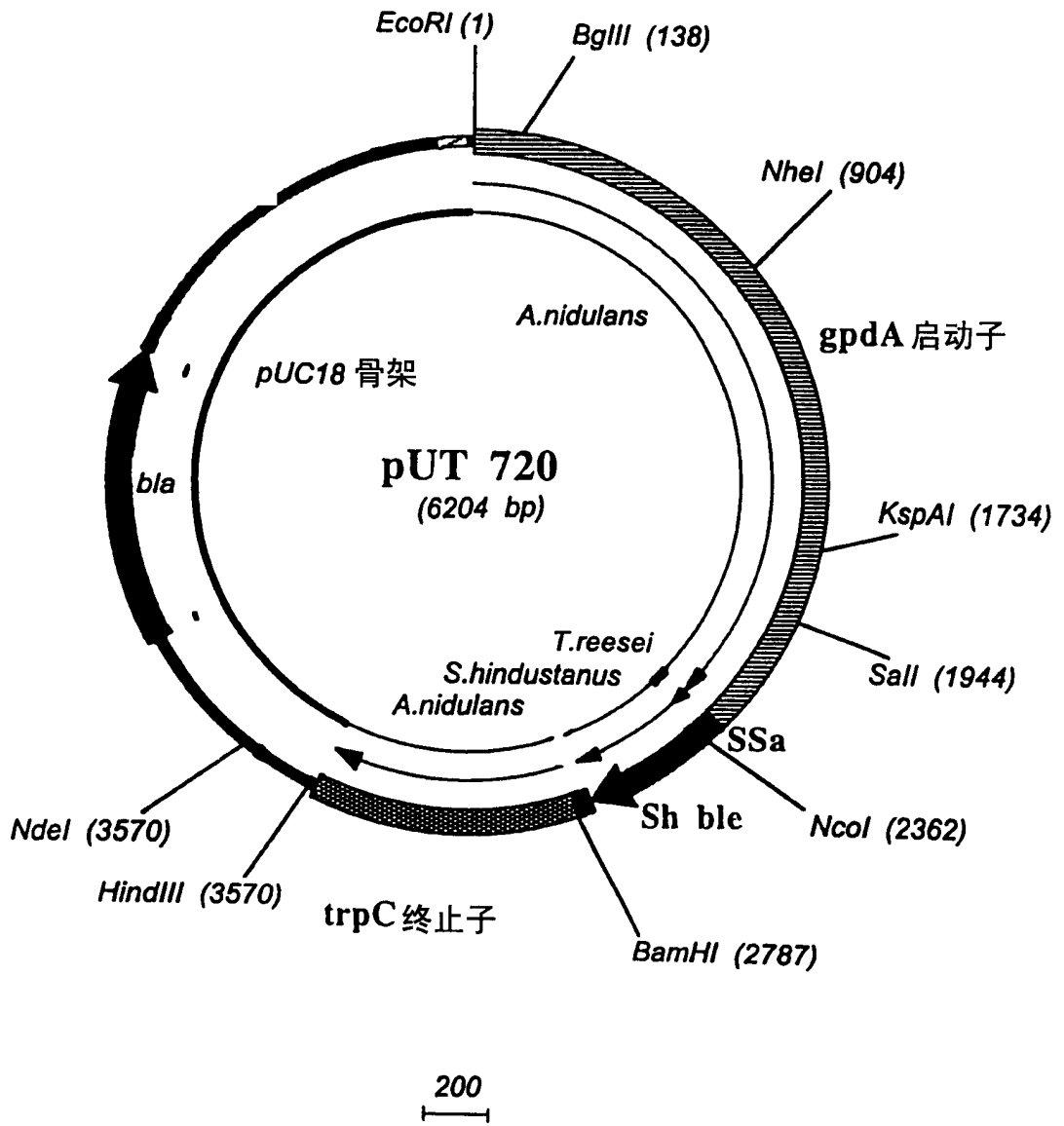


图 2

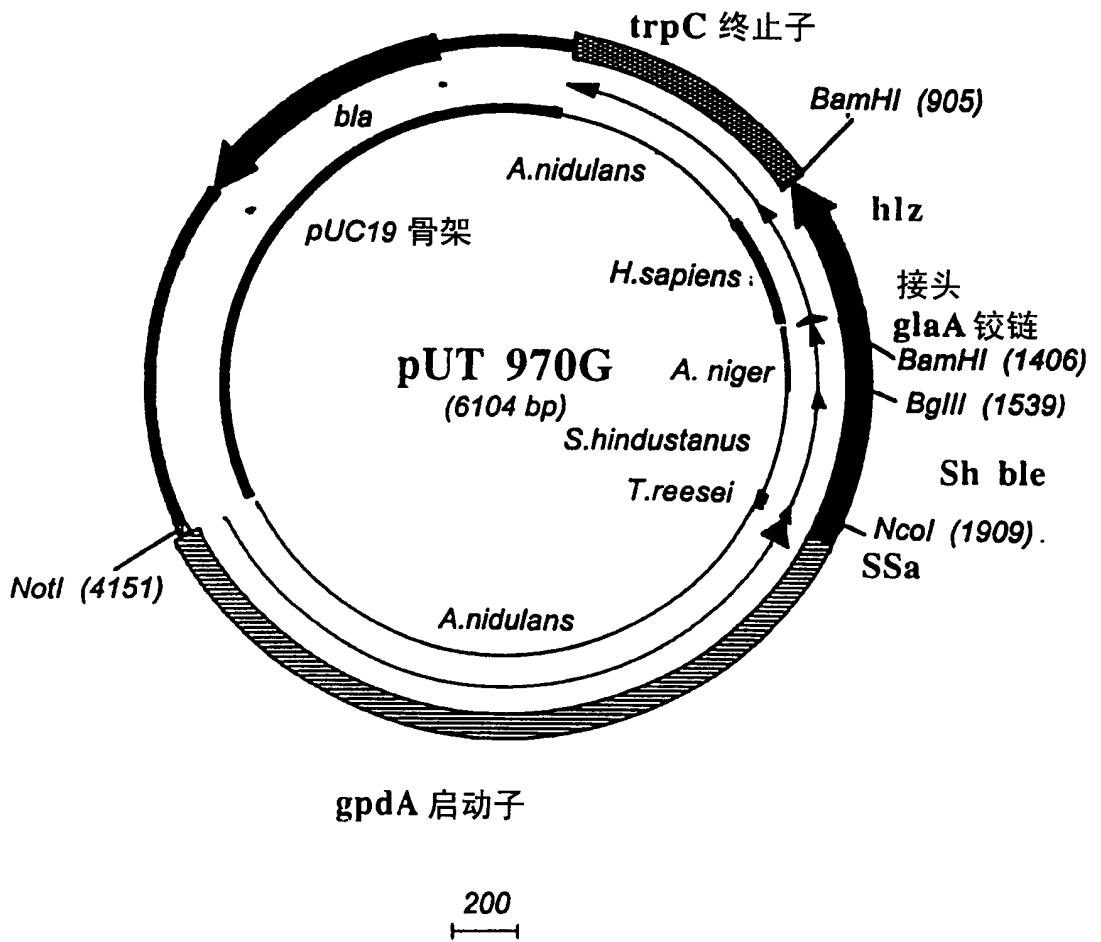


图 3

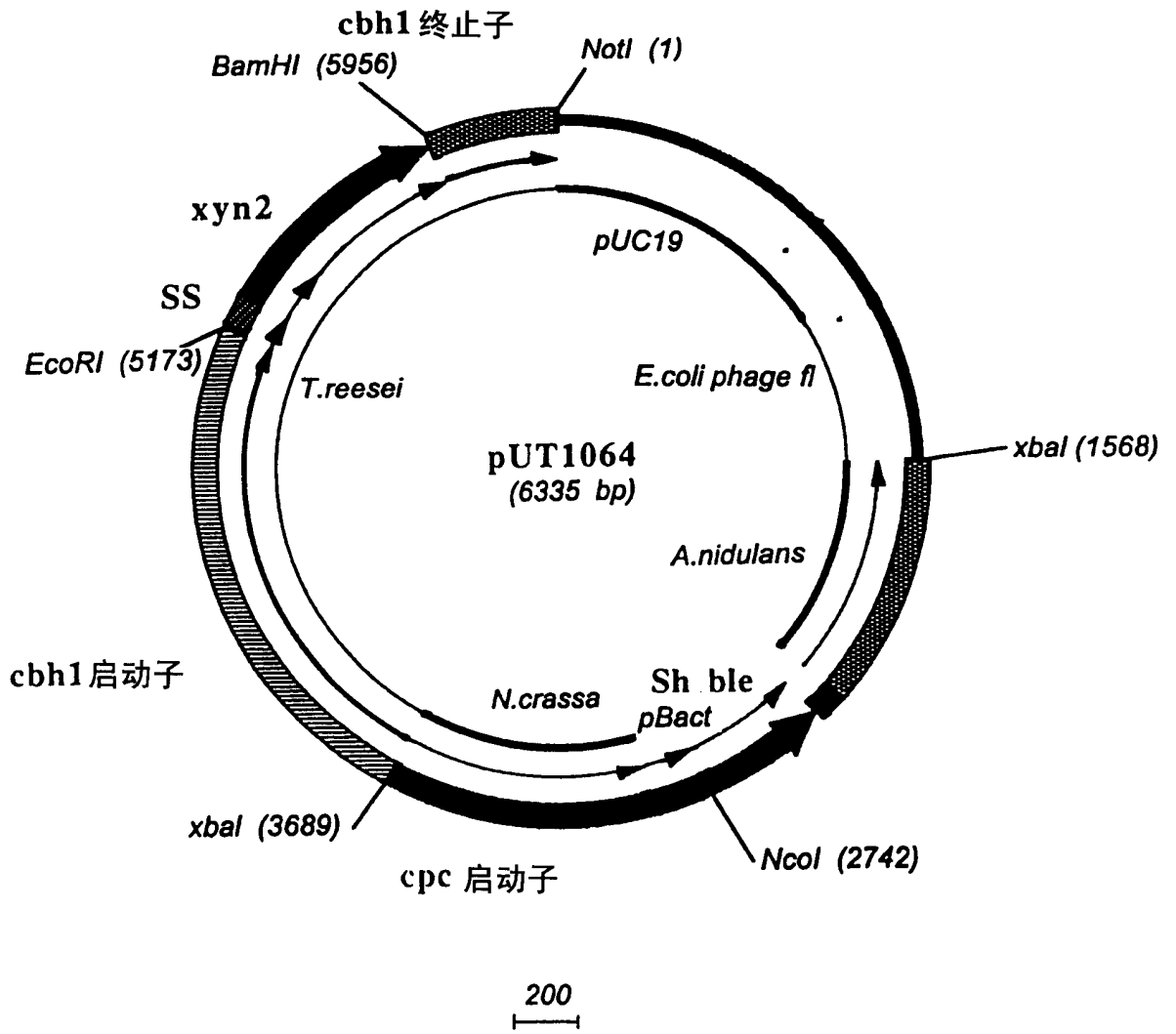


图 4

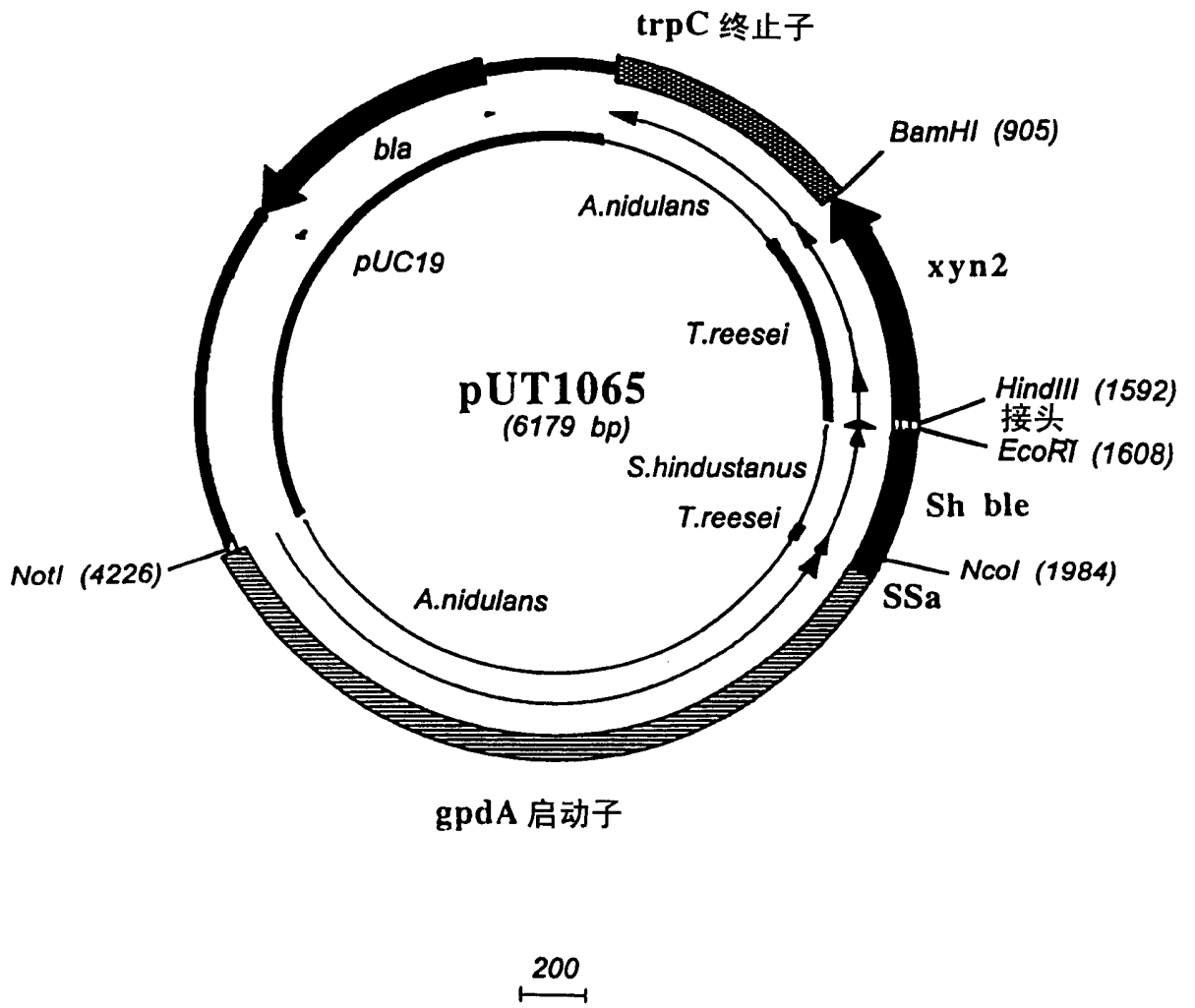


图 5

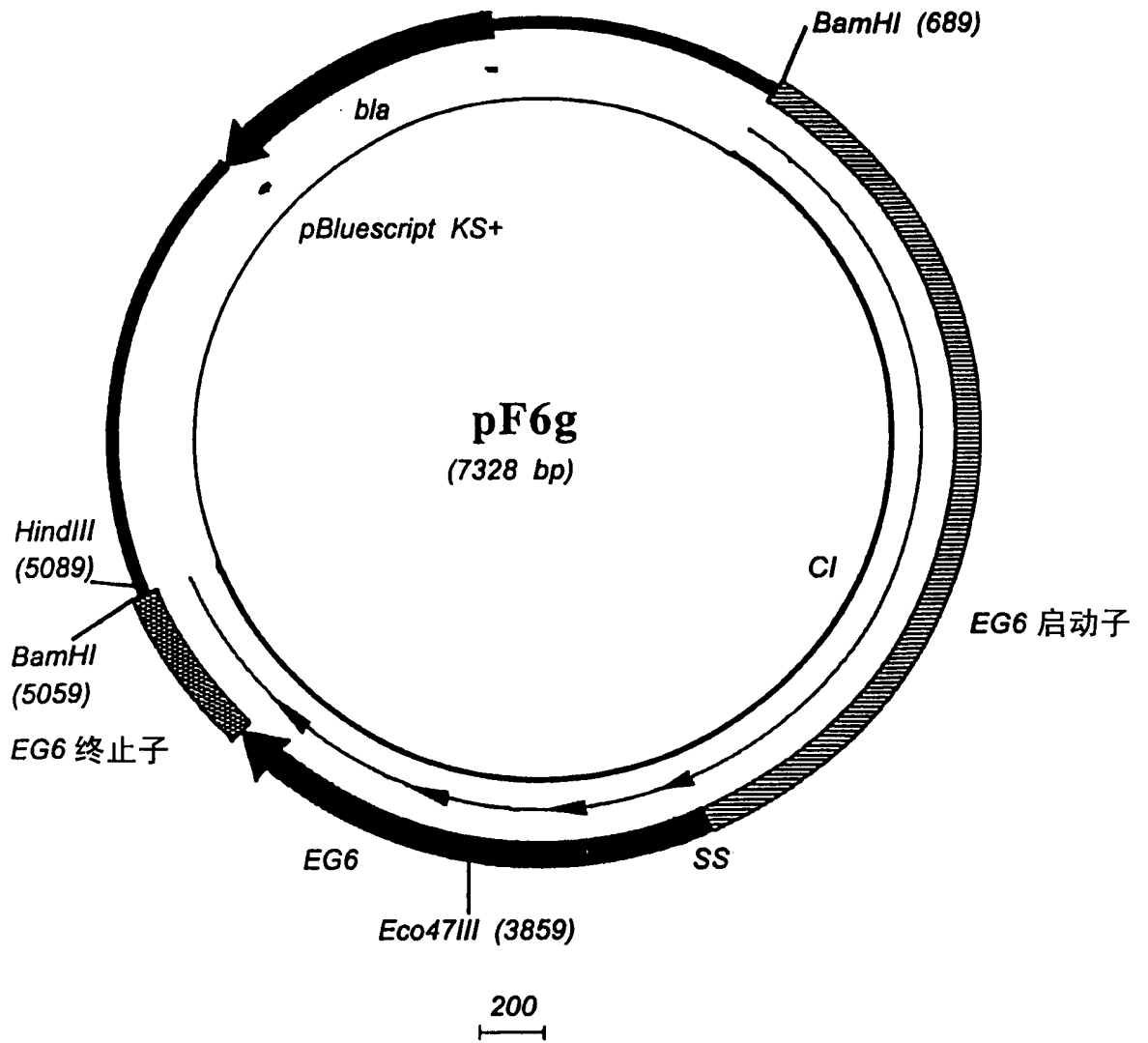


图 6

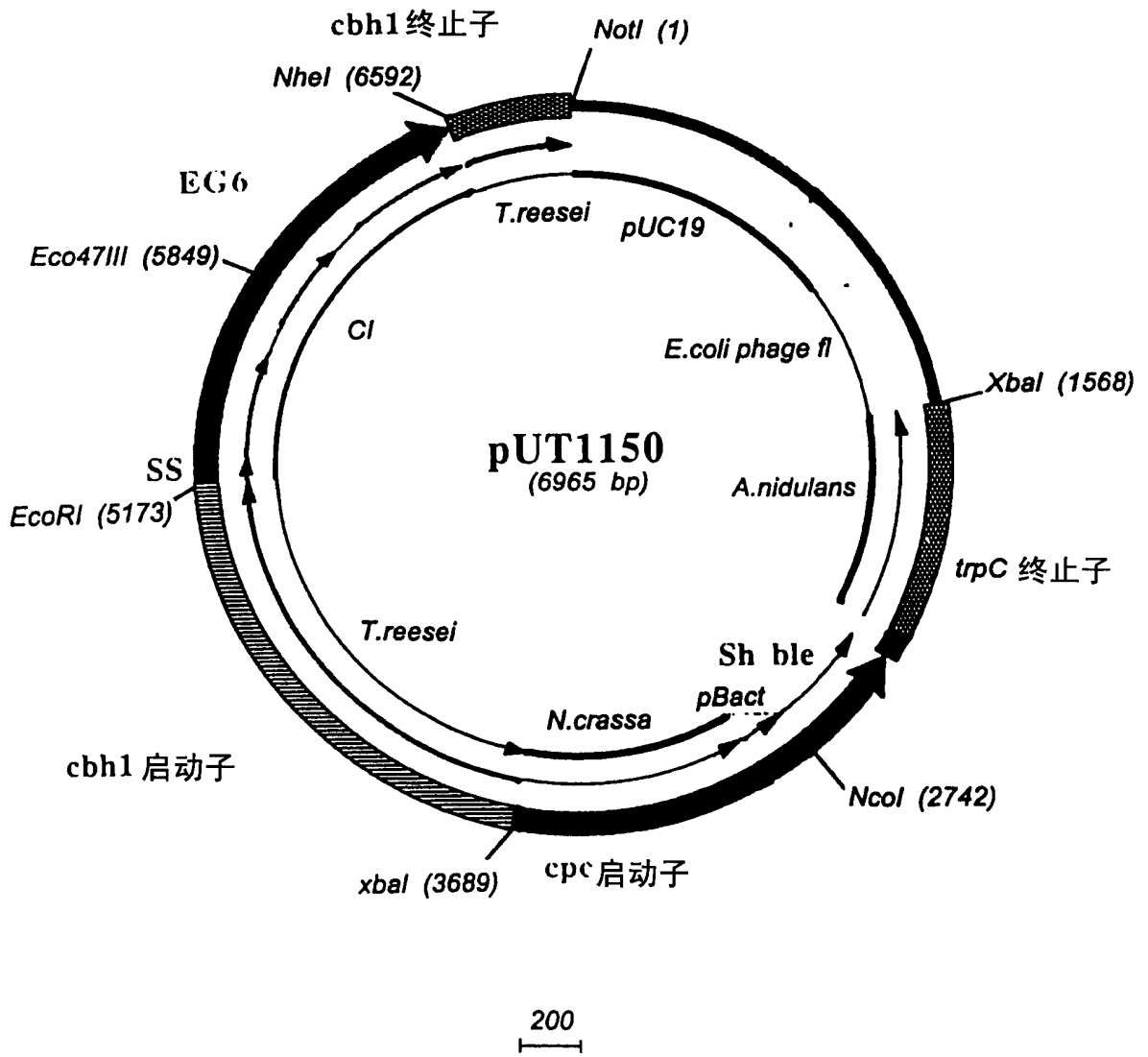


图 7

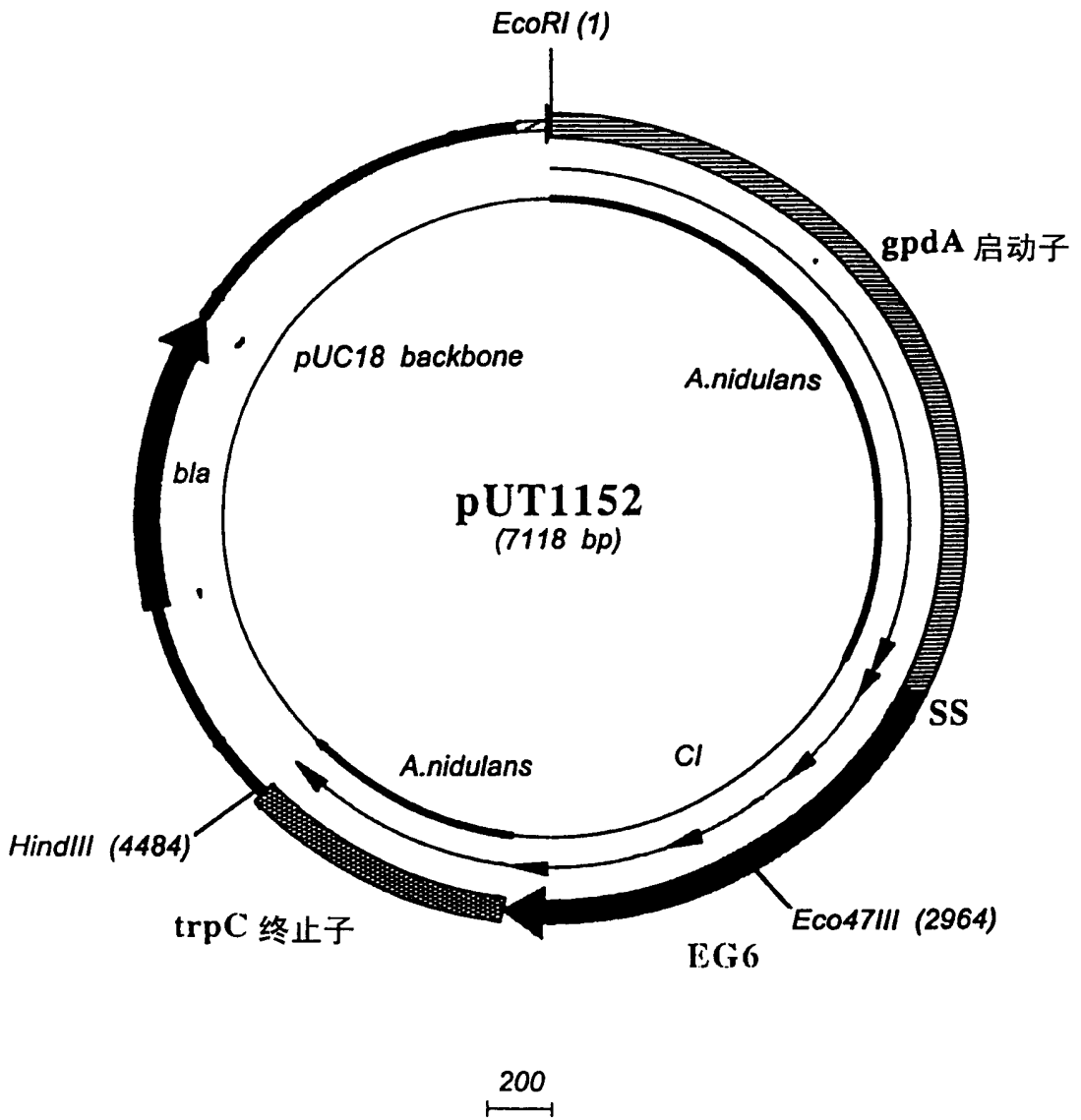
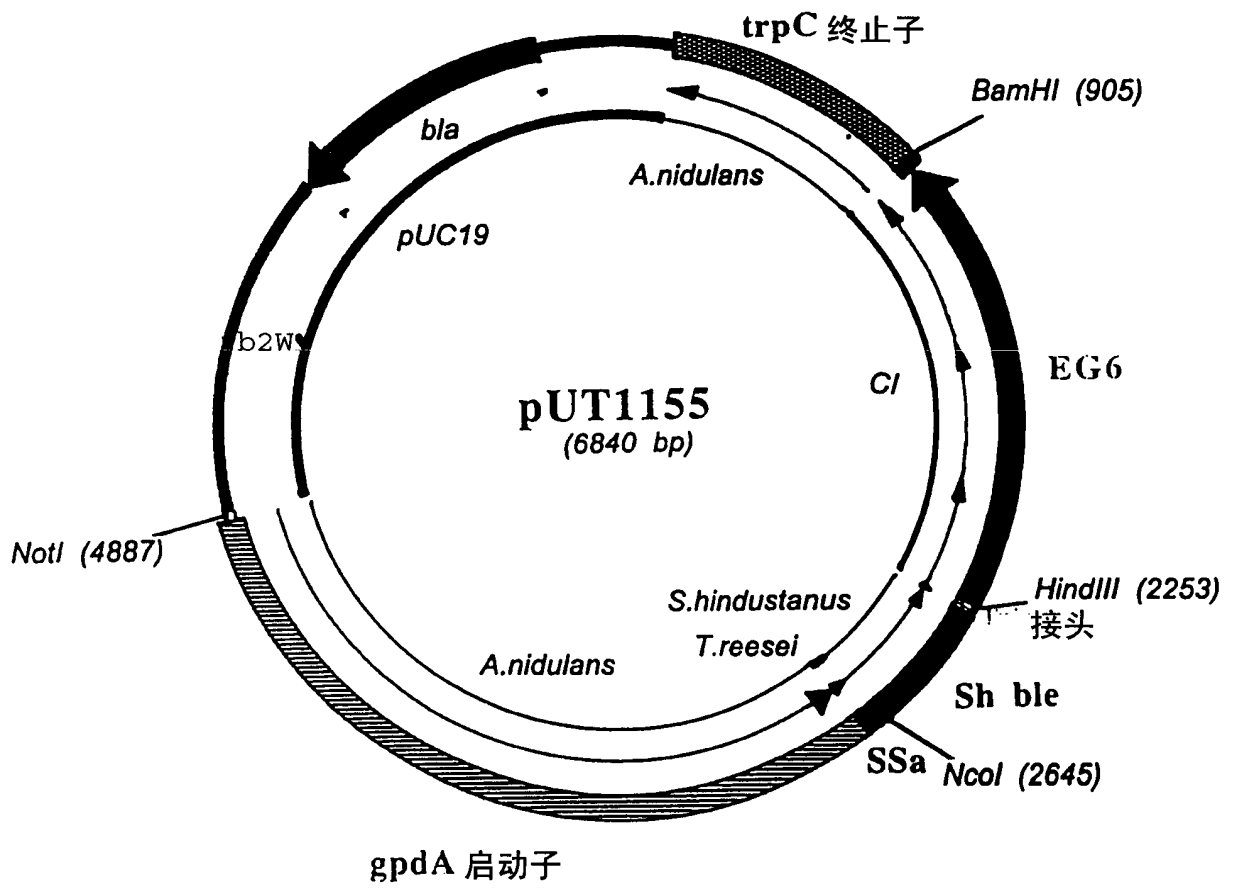


图 8



200

图 9

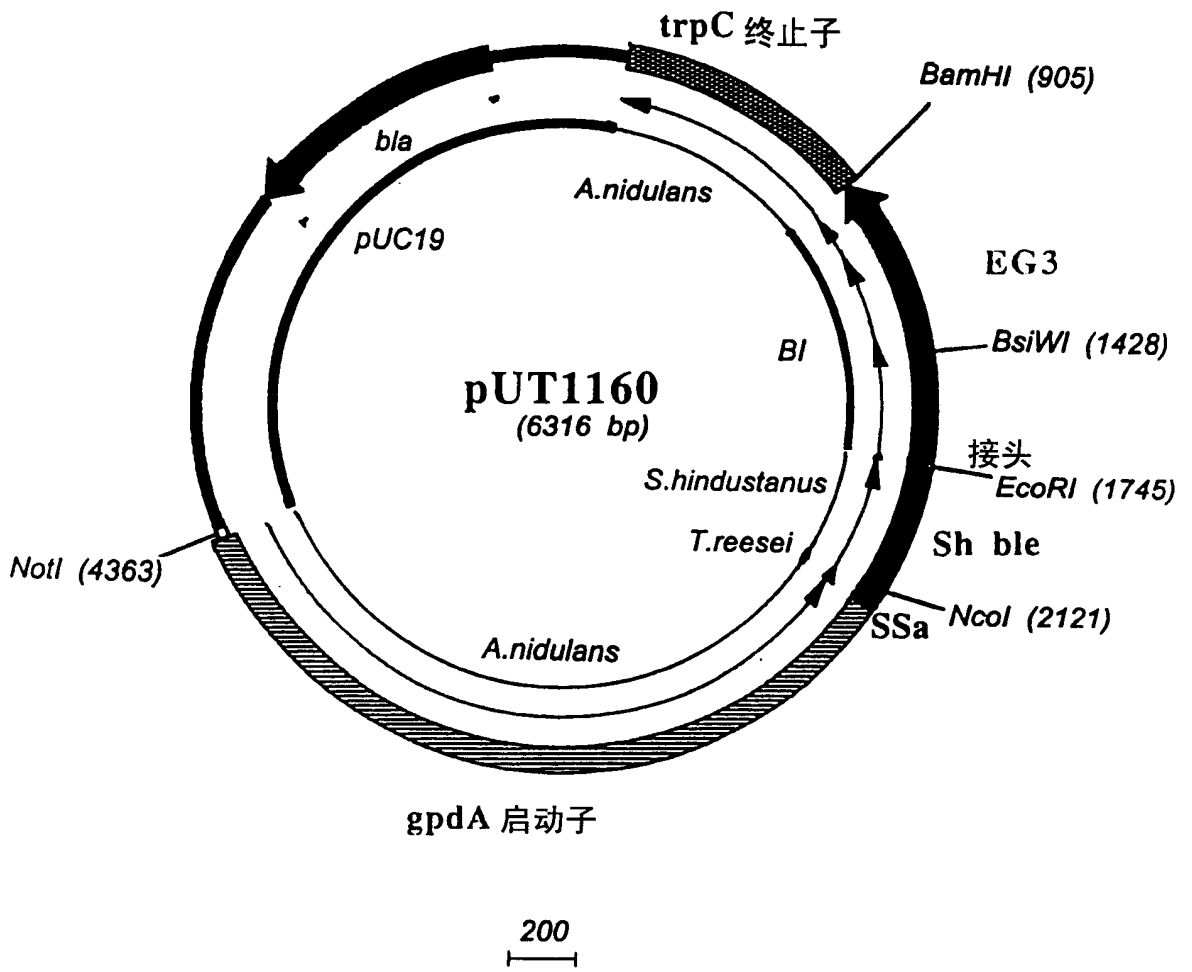


图 10

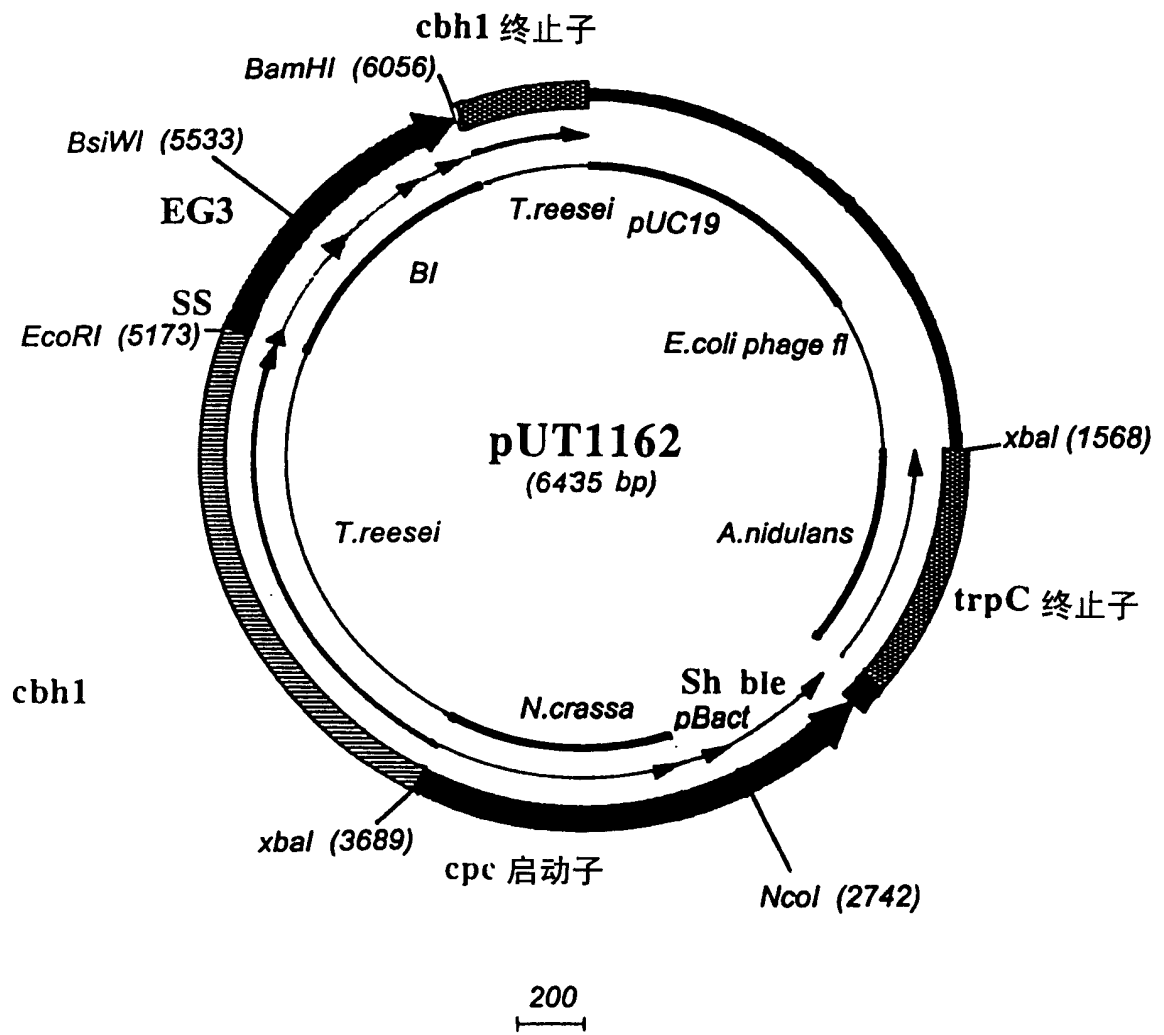


图 11

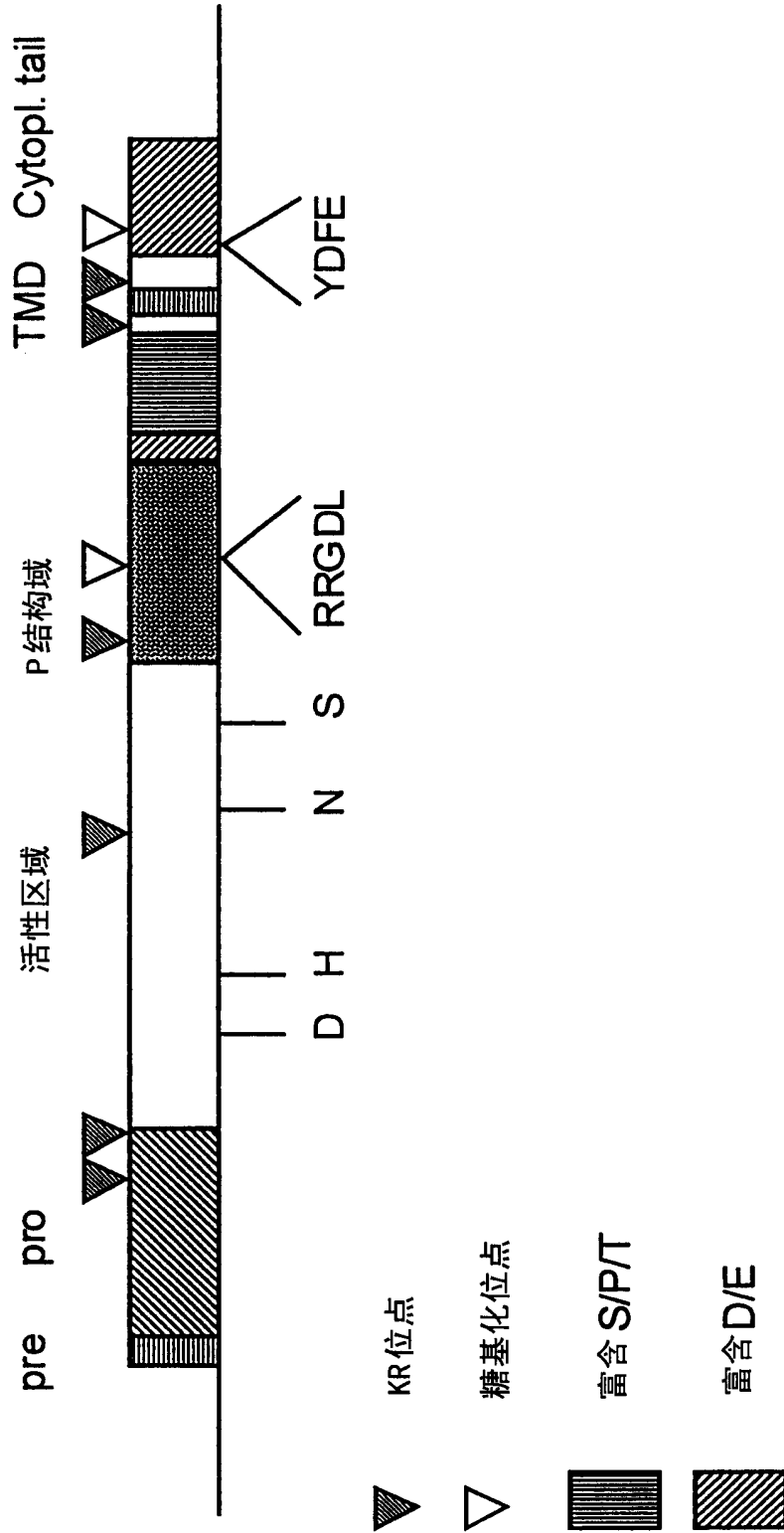


图 12

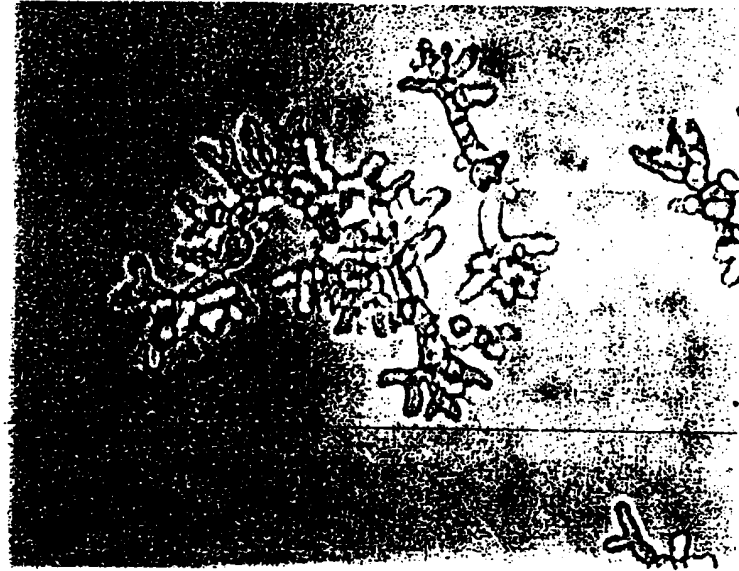


图 13B

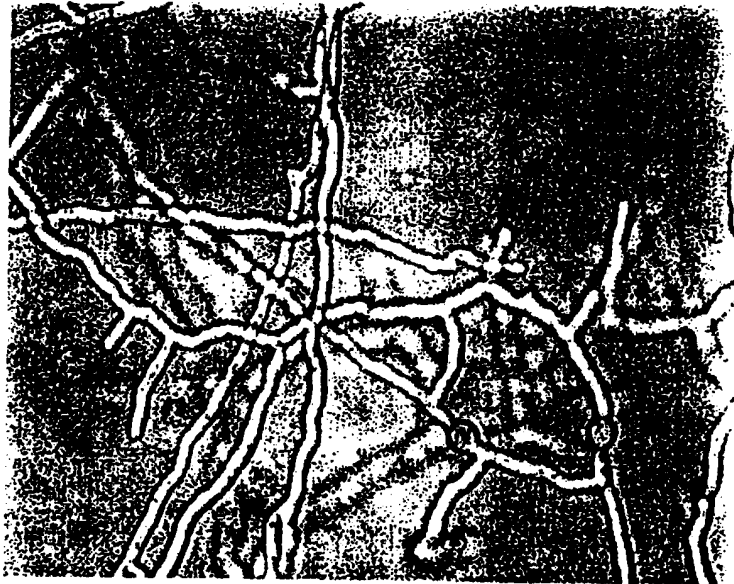


图 13A

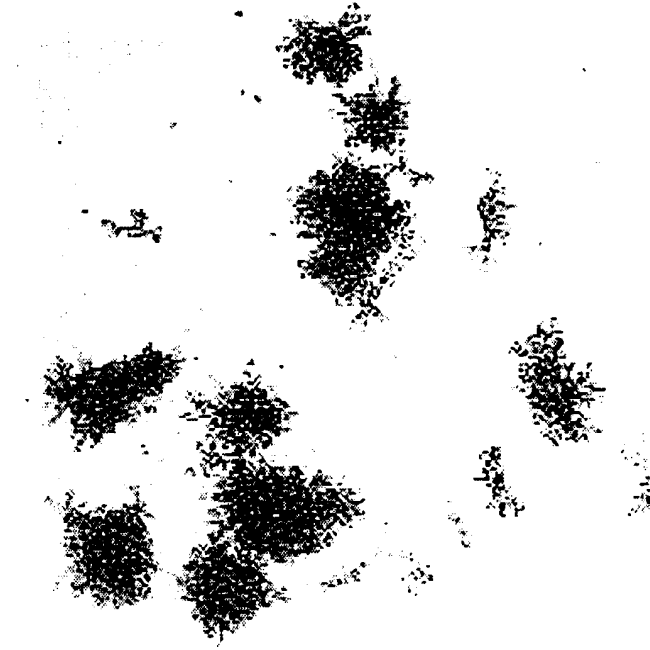


图 14B

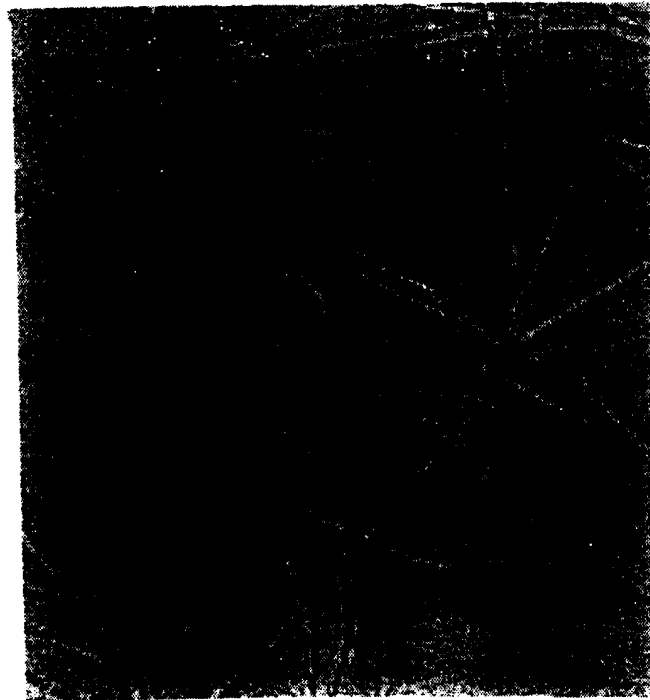


图 14A

PyrE 基因序列

SEQ ID NO:1

PyrE/niger Length: 1578 March 9, 2001 09:28 Type: N Check: 2282

```

1   GGGTTAATGT GAAGGCGTTA GTGGTAATGT ATATTAATGG TGAGATGGGC
51  TTTGATTGGG TTTAATTGGA ATCTGTATAT TTTCAGATGG AGTCAACTTT
101 TGAATGGCCA ATATATCCTC GGCGATACCG TCGGAGATAA GATAAGAATA
151 ATCGCACACT ATTCCCAAAG CATACTGGTA CATACTGCAT TCGGCTAGTG
201 CGGGGTGCTT ACCTCATCCA CCCGAATGAG CCCAACTTTT TTGTCTCAAT
251 CAATAATTGC ATCCAAATTC CCCC GCAACT TCCCCCTCCA ACCCCGTGTC
301 TATACCACTC CCTCCACACC CACACAATCA CAATGGCTCT CCCTGCCTAC
351 AAGACCGCCT TCCTGGAGTC TCTCGTCGGC CAACGTGCTG ACTTTCGGCA
401 CCTTCACCCT GAAGTCGGGT CGCCGTGCGT CACCCCTCCA ACACCGGCAT
451 TATCGCAATC GGAAGACTTA CCACTGTATA CAGACTCCCC CTACTTCTTC
501 AACGCCGGCA TCTTCAACAC CGCCTCTCTC CTCTCCGCC TCTCCACCAT
551 GGCCACACC ATCATCACCT TCCTCGCTGA GAACCCTTCC ATCCCCAAGC
601 CCGACGTCAT GCTTCGGGTA AAAAACCCCC TCTTCCCCA ATACCCCACT
651 TCCACTCAAC AACCCATAAA TAACTAACA AAACCCCTA AACAGCCCCG
701 CATAAAAGG CATCCCCCTC GCGTGCGCCA CCCTCCTTGA ACTCAACCGC
751 ATCGACCCCG CCACCTGGGG CAGCGTGTCC TACAGCTACA ACCGCAAAGA
801 AGCCAAGGAT CACGGCGAAG GCGGCAACAT TGTCGGCGCC GCTCTGAAGG
851 GCAAGACCGT GCTTGTGATC GACGATGTCA TCACGGCCGG TACCGCCATG
901 CGTGAGACCC TCAACCTGGT CGCCAAGGAG GGCGGCAAGG TCGTCGGATT
951 CACTGTTGCT CTGGACCGCT TGGAGAAGAT GCCCGGACCC AAGGACGAGA
1001 ACGGTGTCGA GGACGATAAG CCCAGAATGA GTGCTATGGG TCAGATCCGT
1051 AAGGAGTATG GTGTGCCAC GACGAGTATT GTTACTCTGG ATGATTTGAT
1101 CAAGTTGATG CAGGCGAAGG GCAATGAGGC CGATATGAAG CGGTTGGAGG
1151 AGTATAGGGC TAAGTATCAG GCTAGTGATT AGTCGGTTTC ATTGACCGAT

```

图 15A

1201 TGTTTGGGTG GGTGTGAGAG GTTAGGTTAG GTTGTGGGCG TAGGAATGAA
 1251 AAGCTGTATA CATAGGGGCC TGAAGAGGTG CGTAGAGACG GTCGTGAGAT
 1301 GTTTTATGTC AAAATCTTGA ACAAATGACA CCTTAAAAAA GACCCCTTGG
 1351 TTTCAGCTGA ATTAGCCCGG AAAGATGCTC GGCACGCCAT GAGTCTAGCC
 1401 CACTCAGTGG GCACCCGTTT CCCACATTG AAGTGGCCGA CGCTTATTG
 1451 GCTGAGGCTG TGGCCTGGAA AGGCACTATG GCGTGCTGCG GTACAAGGCC
 1501 GGGGCTGGCG TACGAACCAC GACGCCCGAA GGGAACTCTT CGGTCTTACT
 1551 ACTACTATGT CCCAGTTGA CCCCCGA

SEQ ID NO:2

PytE 的翻译(1-1578)

通用密码子

1 GGGTTAATGTGAAGGCGTTAGTGGTAATGTATATTAATGGTGAGATGGGCTTTGATTGGG
 CCAATTACACTTCCGCAATCACCATTACATATAATTACCACTCTACCCGAAACTAACC

 1 G L M * R R * W * C I L M V R W A L I G
 1 G * C E G V S G N V Y * W * D G L * L G
 1 V N V K A L V V M Y I N G E M G F D W V

 61 TTTAATTGGAATCTGTATATTTTCAGATGGAGTCAACTTTTGAATGGCCAATATATCCTC
 AAATTAACCTTAGACATATAAAAGTCTACCTCAGTTGAAAACCTTACCGTTATATAGGAG

 21 F N W N L Y I F R W S Q L L N G Q Y I L
 21 L I G I C I F S D G V N F * M A N I S S
 21 * L E S V Y F Q M E S T F E W P I Y P R

 121 GGCGATACCGTCGGAGATAAGATAAGAATAATCGCACACTATTCCCAAAGCATACTGGTA
 CCGCTATGGCAGCCTCTATTCTATTCTTATTAGCGTGTGATAAGGGTTTCGTATGACCAT

 41 G D T V G D K I R I I A H Y S Q S I L V
 41 A I P S E I R * E * S H T I P K A Y W Y
 41 R Y R R R * D K N N R T L F P K H T G T

 181 CATACTGCATTCCGGCTAGTGCGGGGTGCTTACCTCATCCACCCGAATGAGCCCAACTTTT
 GTATGACGTAAGCCGATCACGCCCCACGAATGGAGTAGGTGGGCTTACTCGGGTTGAAAA

 61 H T A F G * C G V L T S S T R M S P T F
 61 I L H S A S A G C L P H P P E * A Q L F
 61 Y C I R L V R G A Y L I H P N E P N F F

图 15B

241 TTGTCTCAATCAATAATTGCATCCAAATTCCTCCGCAACTTCCCCCTCCAACCCCGTGTCAACAGAGTTAGTTATTAACGTAGGTTTAAGGGGGCGTTGAAGGGGGAGGTTGGGGCACAG

81 L S Q S I I A S K F P R N F P L Q P R V
81 C L N Q * L H P N S P A T S P S N P V S
81 V S I N N C I Q I P P Q L P P P T P C L
?????????

301 TATACCACTCCCTCCACACCCACACAATCACAATGGCTCTCCCTGCCTACAAGACCGCCTATATGGTGAGGGAGGTGTGGGTGTGTTAGTGTACCGAGAGGGACGGATGTTCTGGCGGA

101 Y T T P S T P T Q S Q W L S L P T R P P
101 I P L P P H P H N H N G S P C L Q D R L
101 Y H S L H T H T I T M A L P A Y K T A F

361 TCCTGGAGTCTCTCGTCGGCCAACGTGCTGACTTTCGGCACCTTCACCCTGAAGTCGGGTAGGACCTCAGAGAGCAGCCGGTTGCACGACTGAAAGCCGTGGAAGTGGGACTTCAGCCCA
????????????????????????????

121 S W S L S S A N V L T F G T F T L K S G
121 P G V S R R P T C * L S A P S P * S R V
121 L E S L V G Q R A D F R H L H P E V G S

INTRON I

421 CGCCGTGCGTCAACCCCTCCAACACCGGCATTATCGCAATCGGAAGACTTACCCTGTATACGCGCACGCAGTGGGAGGTTGTGGCCGTAATAGCGTTAGCCTTCTGAATGGTGACATAT

141 R R A S P L Q H R H Y R N R K T Y H C I
141 A V R H P S N T G I I A I G R L T T V Y
141 P C V T P P T P A L S Q S E D L P L Y T

481 CAGACTCCCCCTACTTCTTCAACGCCGGCATCTTCAACACCGCCTCTCTCCTCTCCGCCCCTCTGAGGGGGATGAAGAAGTTGCGGCCGTTAGTAAGTTGTGGCGGAGAGAGGAGGGCGGG

161 Q T P P T S S T P A S S T P P L S S P P
161 R L P L L L Q R R H L Q H R L S P L R P
161 D S P Y F F N A G I F N T A S L L S A L

NcoI

541 TCTCCACCATGGCCACACCATCATCACCTTCTCGCTGAGAACCCTTCCATCCCCAAGCAGAGGTGGTACCGGGTGTGGTAGTAGTGAAGGAGCGACTCTTGGGAAGGTAGGGGTTCCG

181 S P P W P T P S S P S S L R T L P S P S
181 L H H G P H H H L P R * E P F H P Q A
181 S T M A H T I I T F L A E N P S I P K P
????????????? INTRON II

601 CCGACGTCATGCTTCGGGTAAAAAACCCCTCTTCCCAATACCCCACTTCCACTCAACGGCTGCAGTACGAAGCCATTTTTTGGGGGAGAAAGGGGTTATGGGGTGAAGGTGAGTTG

201 P T S C F G * K T P S F P N T P L P L N
201 R R H A S G K K P P L S P I P H F H S T
201 D V M L R V K N P L F P Q Y P T S T Q Q

图 15C

661 **AACCCATAAATAACTAACAAAAACCCCTAAACAGCCCCGCATACAAAGGCATCCCCCTC**
TTGGGTATTTATTGATTGTTTTTGGGGGATTTGTCGGGGCGTATGTTTCCGTAGGGGGAG

221 N P * I T N K N P L N S P A Y K G I P L
221 T H K * L T K T P * T A P H T K A S P S
221 P I N N * Q K P P K Q P R I Q R H P P R

721 GCGTGCGCCACCCTCCTTGAACCTCAACCGCATCGACCCCGCCACCTGGGGCAGCGTGTCC
CGCACGCGGTGGGAGGAACCTGAGTTGGCGTAGCTGGGGCGGTGGACCCCGTCGCACAGG

241 A C A T L L E L N R I D P A T W G S V S
241 R A P P S L N S T A S T P P P G A A C P
241 V R H P P * T Q P H R P R H L G Q R V L

781 TACAGCTACAACCGCAAAGAAGCCAAGGATCACGGCGAAGGCGGCAACATTGTCGGCGCC
ATGTCGATGTTGGCGTTTCTTCGGTTCCTAGTGCCGCTTCCGCCGTTGTAACAGCCGCGG

261 Y S Y N R K E A K D H G E G G N I V G A
261 T A T T A K K P R I T A K A A T L S A P
261 Q L Q P Q R S Q G S R R R R Q H C R R R

841 *KpnI*
GCTCTGAAGGGCAAGACCGTGCTTGTGATCGACGATGTCATCACGGCCGGTACCGCCATG
CGAGACTTCCCGTTCCTGGCAGCAACTAGCTGCTACAGTAGTGCCGGCCATGGCGGTAC

281 A L K G K T V L V I D D V I T A G T A M
281 L * R A R P C L * S T M S S R P V P P C
281 S E G Q D R A C D R R C H H G R Y R H A

901 CGTGAGACCCTCAACCTGGTCGCCAAGGAGGGCGGCAAGGTCGTCCGATTCACTGTTGCT
GCACTCTGGGAGTTGGACCAGCGGTTCTCCCGCCGTTCCAGCAGCCTAAGTGACAACGA

301 R E T L N L V A K E G G K V V G F T V A
301 V R P S T W S P R R A A R S S D S L L L
301 * D P Q P G R Q G G R Q G R R I H C C S

961 CTGGACCGCTTGGAGAAGATGCCCGGACCCAAGGACGAGAACGGTGTGCGAGGACGATAAG
GACCTGGCGAACCTCTTCTACGGGCCTGGGTTCCTGCTCTTGCCACAGCTCCTGCTATTC

321 L D R L E K M P G P K D E N G V E D D K
321 W T A W R R C P D P R T R T V S R T I S
321 G P L G E D A R T Q G R E R C R G R * A

1021 CCCAGAATGAGTGCTATGGGTCCAGATCCGTAAGGAGTATGGTGTGCCACGACGAGTATT
GGGTCTTACTCACGATACCCAGTCTAGGCATTCCTCATACCACACGGGTGCTGCTCATAA

341 P R M S A M G Q I R K E Y G V P T T S I
341 P E * V L W V R S V R S M V C P R R V L
341 Q N E C Y G S D P * G V W C A H D E Y C

图 15D

1081 GTTACTCTGGATGATTTGATCAAGTTGATGCAGGCGAAGGGCAATGAGGCCGATATGAAG
 CAATGAGACCTACTAACTAGTTCAACTACGTCCGCTTCCCGTTACTCCGGCTATACTTC

361 V T L D D L I K L M Q A K G N E A D M K
 361 L L W M I * S S * C R R R A M R P I * S
 361 Y S G * F D Q V D A G E G Q * G R Y E A

1141 CGGTTGGAGGAGTATAGGGCTAAGTATCAGGCTAGTGATTAGTCGGTTTTATTGACCGAT
 GCCAACCTCCTCATATCCCGATTCATAGTCCGATCACTAATCAGCCAAAGTAACTGGCTA

381 R L E E Y R A K Y Q A S D * S V S L T D
 381 G W R S I G L S I R L V I S R F H * P I
 381 V G G V * G * V S G * * L V G F I D R L

1201 TGTTTGGGTGGGTGTGAGAGGTTAGGTTAGGTTGTGGGCGTAGGAATGAAAAGCTGTATA
 ACAAACCCACCCACACTCTCCAATCCAATCCAACCCCGCATCCTTACTTTTCGACATAT

401 C L G G C E R L G * V V G V G M K S C I
 401 V W V G V R G * V R L W A * E * K A V Y
 401 F G W V * E V R L G C G R R N E K L Y T

1261 CATAGGGCCTGAAGAGGTGCGTAGAGACGGTCGTGAGATGTTTTATGTCAAATCTTGA
 GTATCCCCGGACTTCTCCACGCATCTCTGCCAGCACTCTACAAAATACAGTTTTAGAACT

421 H R G L K R C V E T V V R C F M S K S *
 421 I G A * R G A * R R S * D V L C Q N L E
 421 * G P E E V R R D G R E M F Y V K I L N

1321 ACAAATGACACCTTAAAAAGACCCCTTGGTTTCAGCTGAATTAGCCCGGAAAGATGCTC
 TGTTTACTGTGGAATTTTTCTGGGGAACCAAAGTCGACTTAATCGGGCCTTTCTACGAG

441 T N D T L K K T P W F Q L N * P G K M L
 441 Q M T P * K R P L G F S * I S P E R C S
 441 K * H L K K D P L V S A E L A R K D A R

1381 GGCACGCCATGAGTCTAGCCCACTCAGTGGGCACCCGTTTCCCACATTTGAAGTGGCCGA
 CCGTGCGGTACTCAGATCGGGTGAGTCACCCGTGGGCAAAGGGTGTAACCTTCACCGGCT

461 G T P * V * P T Q W A P V S H I * S G R
 461 A R H E S S P L S G H P F P T F E V A D
 461 H A M S L A H S V G T R F P H L K W P T

1441 CGCTTATTTGGCTGAGGCTGTGGCCTGGAAAGGCACTATGGCGTGCTGCGGTACAAGGCC
 GCGAATAAACCGACTCCGACACCGGACCTTCCGTGATACCGCAGCAGCCATGTTCCGG

481 R L F G * G C G L E R H Y G V L R Y K A
 481 A Y L A E A V A W K G T M A C C G T R P
 481 L I W L R L W P G K A L W R A A V Q G R

图 15E

```
1501      GGGGCTGGCGTACGAACCACGACGCCCGAAGGGA ACTCTTCGGTCTTACTACTACTATGT
          CCCC GACCGCATGCTTGCTGCTGCGGGCTTCCCTTGAGAAGCCAGAATGATGATGATACA

501      G A G V R T T T P E G N S S V L L L L C
501      G L A Y E P R R P K G T L R S Y Y Y Y V
501      G W R T N H D A R R E L F G L T T T M S

1561      CCCCAGTTGACCCCCCGA
          GGGGTCAACTGGGGGGCT

521      P Q L T P R
521      P S * P P
521      P V D P P
```