

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12N 15/82

A01H 5/10 A01H 5/00

C12N 15/29

[12] 发明专利申请公开说明书

[21] 申请号 00813447.2

[43] 公开日 2002 年 10 月 23 日

[11] 公开号 CN 1376204A

[22] 申请日 2000.8.25 [21] 申请号 00813447.2

[30] 优先权

[32]1999.8.27 [33]US [31]60/151,044

[32]1999.10.27 [33]US [31]60/161,722

[32]2000.5.30 [33]CA [31]2,310,304

[86] 国际申请 PCT/CA00/00988 2000.8.25

[87] 国际公布 WO01/16340 英 2001.3.8

[85] 进入国家阶段日期 2002.3.27

[71] 申请人 塞姆柏奥希斯遗传学公司

地址 加拿大艾伯塔

共同申请人 联邦科学和工业研究组织

[72] 发明人 S·考德瑞 G·范罗金

M·M·莫洛尼

S·辛

[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 李 瑛

权利要求书 4 页 说明书 39 页 附图页数 24 页

[54] 发明名称 亚麻种子特异性启动子

[57] 摘要

本发明提供了在亚麻种子和其它植物种类种子中表达非天然基因的新方法。这些方法涉及从亚麻得到的种子特异性启动子的应用。同时提供的有新的亚麻种子特异性启动子,含有新的亚麻种子特异性启动子的嵌合核酸构建体,含有新的亚麻种子特异性启动子的转基因植物细胞,转基因植物和转基因植物的种子。这些启动子和方法可用于例如改变亚麻种子或其它植物种子中的种子油和蛋白质组成。

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

权 利 要 求 书

1. 在亚麻种子中表达需要的核酸序列的方法，包括步骤：

(a) 制备在 5' 到 3' 转录方向含有下列可操作连接的成分的嵌合核酸构建体：

(1) 从亚麻得到的种子特异性启动子；和

(2) 所述的需要的核酸序列，其中所述的需要的核酸对所述的种子特异性启动子是非天然的；

(b) 将所述的嵌合核酸构建体导入亚麻植物细胞；和

(c) 使所述的亚麻植物细胞生长成能够结种子的成熟的亚麻植物，其中所述的需要的核酸序列是在所述的种子特异性启动子的控制下在种子中表达的。

2. 根据权利要求 1 所述的方法，其中所述的种子特异性启动子给予它的天然核酸序列的至少一个表达特征是给予所述的非天然核酸序列的。

3. 根据权利要求 2 所述的方法，其中所述的表达特征是表达的计时，表达的水平，对光照条件变化的反应，对温度变化的反应，对化学试剂浓度变化的反应。

4. 根据权利要求 1 所述的方法，其中所述的亚麻种子特异性启动子是从包括油质蛋白启动子，2S 贮藏蛋白启动子和类豆球蛋白种子贮藏蛋白启动子的启动子组中选择的。

5. 根据权利要求 1 所述的方法，其中所述的亚麻种子特异性启动子，包括：

(a) 如图 1 (SEQ. ID. NO. :1)，图 2 (SEQ. ID. NO. :4)，图 3 (SEQ. ID. NO. :6) 或图 4 (SEQ. ID. NO. :8) 所示的核酸序列，其中 T 也可以是 U；

(b) 当 (a) 的核酸序列互补的核酸序列；

(c) 具有与 (a) 或 (b) 的核酸序列同源的基本序列核酸序列；

(d) 是 (a)，(b) 或 (c) 的核酸序列的类似物的核酸序列；或

(e) 在严格杂交条件下，与 (a)，(b)，(c) 或 (d) 的核酸序列杂交的

核酸序列。

6. 根据权利要求 1 所述的方法，其中需要的所述核酸序列的表达导致所述的种子中蛋白质或脂肪酸组成的改变。

7. 根据下列方法制备的转基因亚麻种子，包括：

(a) 制备在 5' 到 3' 转录方向含有下列可操作地连接的成分的嵌合核酸构建体：

(1) 从亚麻得到的种子特异性启动子；和

(2) 需要的核酸序列，其中所述的需要的核酸对所述的种子特异性启动子是非天然的；

(b) 将所述的嵌合核酸构建体导入亚麻植物细胞；和

(c) 使所述的亚麻植物细胞生长成能够长种子的成熟的亚麻植物，其中所述的需要的核酸序列是在所述的种子特异性启动子的控制下，在种子中表达的。

8. 根据权利要求 7 所述的转基因亚麻种子，其中所述的种子特异性启动子给予它的天然核酸序列的至少一个表达特征是给予所述的非天然核酸序列。

9. 根据权利要求 8 所述的方法，其中所述的表达特征是表达的时间或表达的水平。

10. 根据权利要求 8 所述的转基因亚麻种子，其中所述的种子特异性启动子是种子贮藏蛋白启动子，油质蛋白启动子，2S 贮藏蛋白启动子或类豆球蛋白种子贮藏蛋白启动子。

11. 根据权利要求 8 所述的转基因亚麻种子，其中所述的种子特异性启动子包括：

(a) 如图 1 (SEQ. ID. NO. :1), 图 2 (SEQ. ID. NO. :4), 图 3 (SEQ. ID. NO. :6) 或图 4 (SEQ. ID. NO. :8) 所示的核酸序列，其中 T 也可以是 U；

(b) 与 (a) 的核酸序列互补的核酸序列；

(c) 与 (a) 或 (b) 的核酸序列有基本序列同源性的核酸序列；

(d) 是 (a), (b) 或 (c) 的核酸序列的类似物的核酸序列；或

(e) 在严格杂交条件下，与 (a), (b), (c) 或 (d) 的核酸序列杂交的

核酸序列。

12. 根据权利要求 8 所述的转基因亚麻种子，其中所述的需要的非天然基因的表达导致种子蛋白质或脂肪酸组成的改变。

13. 通过下列方法制备的能够发育产生种子的转基因亚麻植物，包括

(a) 制备在 5' 到 3' 转录方向上含有下列可操作地连接的成分的嵌合核酸构建体：

(1) 从亚麻得到的种子特异性启动子；和

(2) 需要的核酸序列，其中需要的所述核酸对所述的种子特异性启动子是非天然的；

(b) 将所述的嵌合核酸构建体导入亚麻植物细胞；和

(c) 使所述的亚麻植物细胞生长成能够结种子的成熟亚麻植物，其中所述的需要的核酸序列是在所述的种子特异性启动子的控制下，在种子中表达的。

14. 在植物中能够指导种子特异性表达的分离的核酸序列，包括：

(a) 如图 1 (SEQ. ID. NO. :1)，图 2 (SEQ. ID. NO. :4)，图 3 (SEQ. ID. NO. :6) 或图 4 (SEQ. ID. NO. :8) 所示的核酸序列，其中 T 也可以是 U；

(b) 与 (a) 所述的核酸序列互补的核酸序列；

(c) 与 (a) 或 (b) 的核酸序列具有基本序列同源性的核酸序列；或

(d) 是 (a)，(b) 或 (c) 的核酸序列的类似物的核酸序列；或

(e) 与 (a)，(b)，(c) 或 (d) 的核酸序列，在严格的杂交条件下杂交的核酸序列。

15. 分离的嵌合核酸序列，包括：

(a) 含有从亚麻得到的种子特异性启动子的第一个核酸序列，它包括：

(1) 如图 1 (SEQ. ID. NO. :11)，图 2 (SEQ. ID. NO. :4)，图 3 (SEQ. ID. NO. :6) 或图 4 (SEQ. ID. NO. :8) 所示的核酸序列，其中 T 也可以是 U；

(2) 在严格杂交条件下，与 (a) 的核酸序列杂交的核酸序列；

(3) 与 (a) 的核酸序列互补的核酸序列；或

- (4) 与(a)的核酸序列有基本序列同源性的核酸序列；和
- (b) 对所述的亚麻种子特异性启动子是非天然的第二个核酸序列。

16. 在植物种子中表达需要的核酸序列的方法，包括步骤：

- (a) 将权利要求 15 所述的嵌合核酸序列导入植物细胞；和
- (b) 使所述的植物细胞生长成能够结种子的成熟植物，其中第二个核酸序列是在种子特异性启动子的控制下，在种子中表达的。

17. 根据权利要求 16 所述的方法，其中所述的植物细胞选自下列植物的组：大豆 (*Glycine max*)，油菜 (*Brassia napus*, *Brassica campestris*)，向日葵 (*Helianthus annuus*)，棉花 (*Gossypium hirsutum*)，玉米 (*Zea mays*)，烟草 (*Nicotiana tobacum*)，苜蓿 (*Medicago Sativa*)，小麦 (*Triticum sp.*)，大麦 (*Hordeum vulgare*)，燕麦 (*Avena Sativa L.*)，高粱 (*Sorgham bicolor*)，拟南芥 (*Arabidopsis*)，土豆 (*Solanum sp.*)，亚麻/亚麻子 (*Linum usitatissimum*)，红花 (*Carthamus tinctorius*)，油棕榈 (*Eleais guineeis*)，落花生 (*Arachis hypogaea*)，巴西坚果 (*Bertholletia excelsa*)，椰子 (*Cocus nucifera*)，蓖麻 (*Ricinus communis*)，芫荽 (*Coriandrum sativum*)，南瓜 (*Cucurbita maxima*)，灌木 (*Simmondsia chinensis*) 和水稻 (*Oryza Sativa*)。

- 18. 根据权利要求 16 所述的方法制备的植物。
- 19. 含有权利要求 15 所述的嵌合核酸序列的植物细胞。
- 20. 含有权利要求 15 所述的嵌合核酸序列的植物种子。
- 21. 从根据权利要求 16 所述的方法制备的植物中得到的植物种子。
- 22. 含有权利要求 14 所述的核酸序列的重组表达载体。
- 23. 含有权利要求 15 所述的核酸序列的重组表达载体。

说明书

亚麻种子特异性启动子

发明领域

本发明涉及用于改变植物种子成分的植物基因工程方法，更具体地说，本发明涉及已经从亚麻中得到，并且能够指导亚麻种子以及其它植物种子中非天然基因的表达的启动子。

发明背景

亚麻或亚麻子(*Linum usitatissimum*)是具有经济价值的重要油料种子作物。亚麻油和粉是从亚麻种子得到的有价值的原材料。另一种有经济价值的原材料，亚麻纤维，是可以从该植物的茎中得到的。亚麻油组分可用于非食用目的，例如用于制造清漆和油料，并且通过新培育称为 *Linola* 的新品种，最近已经可用于制造许多食用产品，如黄油和色拉油和调味品(Green (1986 年) 《Can. J. Plant Sci.》，66 卷: 499 - 503 页)。亚麻粉首先可用作反刍动物的饲料中的成分，而亚麻纤维可用于制造亚麻布纺织品。了解了亚麻作为原材料来源的经济重要性，考虑到农业上的性能，例如：种子产量，对病原体 and 低温的抗性，和适于后加工应用中的原材料的产量和质量，进一步改良和区分可利用的亚麻作物是需要的。虽然如开发 *Linola* 品种所证实的，通过传统的植物育种，得到改良的亚麻品种是可能的，但由于需要的时间长，开发优良的农业植物品系需要在植物育种中大量投资，植物遗传工程技术允许直接从不相关的种类中分离基因并将这些基因转移到良好的农学背景中，从而显著地减少开发新的品种所需要的时间。另外，植物遗传工程允许生产自然情况下不能从亚麻得到的产品如治疗剂。

为了通过植物遗传工程开发新的亚麻品种，控制导入的外源或非天然基因的表达是十分重要的。根据开发的植物品系的用途，需要的非天然基因的表达特征，如非天然基因的表达水中，非天然基因表达的特定

植物组织或器官，和植物生命周期中非天然基因表达的特定时间是会改变的。例如，改变种子的油组成可能需要种子发育早期中脂肪酸代谢中参与的酶低水平地种子特异性表达。（参见例如：美国专利 5,420,034）。另一方面，收获该植物的叶子，药用蛋白质的表达可能优选地需要高水平的叶子特异性表达。（参见例如，美国专利：5,929,304）。

为了控制非天然基因的表达特征，可以利用许多因子。一个因子是所利用的转录启动子的选择。目前有效的植物兼容启动子很多，详细文献记载的启动子中的一些包括组成性启动子如 35 - SCaMv 启动子（Rothstein 等人，（1987 年）《基因》53 卷：153 - 161 页），和泛素启动子（美国专利 5,614,399），组织特异性启动子如种子特异性启动子，如菜豆蛋白启动子（Sengupta - Gopalan 等人，（1985 年），PNAS USA 82 卷：3320 - 3324 页）和可诱导启动子，如热（Czarnecka 等人，（1989 年），《分子细胞生物学》，9(8)：3457 - 3464 页），紫外光，诱导物和创伤（Lois 等人，（1989 年）EMBO J. 8(6)：1641 - 1648 页），或化学物如内源激素（Skriver 等人（1991 年），美国国家科学院院报，88(16)卷：7266 - 7270 页）可诱导的那些启动子。为了控制植物中非天然基因的表达特征，可利用的其它因子包括转录修饰因子如：内含子，多聚腺苷酸化位点和转录终止位点。非天然基因的表达特征可以进一步利用影响翻译的因子如，核糖体结合位点和宿主显示的密码子偏倚性来掌握。另外，非天然基因本身可以影响转基因植物的存活，从而特定地限制可以达到的表达水平。在一些情况中，通过特异于组织的方式，如在叶子或种子中表达蛋白质，或由于存在或缺乏能够将蛋白质送到不同亚细胞部分的特异靶序列来限制蛋白质在不同的亚细胞部分如细胞质，内质网或液泡中的积累，可能能够克服这一问题。另一个影响表达特征的因子是构建体将自身插入宿主染色体中的座位。这一因子的效果可以提供关于为什么用同一重组构建体转化的不同植物会使重组蛋白质表达水平有波动。

就本发明人所知，亚麻种子中非天然基因的表达只在 PCT 专利申请 WO 98/18948 中有记载。这一申请公开了从亚麻得到的两个硬脂酰 - 酰基载体蛋白质去饱和酶 (SAD) 基因。相关的 SAD 启动子序列可用于修饰亚麻

和其它植物以表达内源或外源基因。但是 WO 98/18948 所教授的方法受到 SAD 启动子在亚麻中是非种子特异的，并且使其在叶子，茎，花和种子中表达的限制。所以非天然基因的表达可能在非种子组织中产生不需要的副作用。另外，利用 SAD 启动子允许有限地控制表达水平和表达时间。

所以，本领域需要进一步改进亚麻种子和其它植物种子中非天然基因的表达方法。

发明概述

本发明涉及改良的植物中非天然基因的种子特异性表达方法。本发明特别涉及了亚麻中非天然基因的种子特异性表达的已改良的方法。

所以，在一方面，本发明提供了亚麻种子中需要的核酸序列的表达方法，该方法包括步骤：

(a) 制备在 5' 到 3' 转录方向含有作为可操作地连接的下列成分的嵌合核酸构建体：

(1) 从亚麻中得到的种子特异性启动子；和

(2) 需要的核酸序列，其中所述的需要的核酸对于所述的亚麻种子特异性启动子是非天然的；

(b) 将所述的嵌合核酸构建体导入亚麻植物细胞；和

(c) 使所述的亚麻植物细胞生长成能够结种子的成熟的亚麻植物，其中所述的需要的核酸序列是在所述的亚麻种子特异性启动子控制下在种子中表达的。

在本发明的优选的实施方案中，启动子给予非天然核酸序列的至少一个表达特征，如植物生命周期中表达的计时是和给予天然核酸序列时的表达特征相似的。在其它优选的实施方案中，亚麻种子特异性启动子是油质蛋白启动子，2S 贮藏蛋白启动子或类豆球蛋白种子贮藏蛋白启动子。

在其它方面，本发明提供了包括下面步骤的方法的转基因亚麻种子：

(a) 制备在 5' 到 3' 转录方向含有作为可操作地连接的下列成分的

嵌合核酸构建体:

- (1) 从亚麻得到的种子特异性启动子; 和
- (2) 需要的核酸序列, 其中需要的所述核酸对于所述的种子特异性启动子是非天然的;
- (b) 在亚麻植物细胞中导入所述的嵌合核酸构建体; 和
- (c) 使所述的亚麻植物细胞生长成能够长种子的成熟的亚麻植物, 其中所述的需要的核酸序列是在所述的种子特异性启动子的控制下, 在种子中表达的。

在本发明的优选的实施方案中, 至少一种启动子给予非天然核酸序列的表达特征, 例如: 在植物的生命周期中, 表达的计时是与给予天然核酸序列时的表达特征相似的。在其它的实施方案中, 亚麻种子特异性启动子是油质蛋白启动子, 2S 贮藏蛋白启动子或类豆球蛋白种子贮藏蛋白启动子。

在另一个方面, 本发明提供根据下面方法制备的转基因亚麻种子:

(a) 制备在 5' 到 3' 转录方向含有下面可操作地连接的成分的嵌合核酸构建体:

- (1) 从亚麻得到的种子特异性启动子; 和
- (2) 需要的核酸序列, 其中所述的需要的核酸对所述的种子特异性启动子是非天然的;
- (b) 将所述的嵌合核酸构建体导入亚麻植物细胞; 和
- (c) 使所述的亚麻植物细胞生长成能够结种子的成熟亚麻植物, 其中所述的需要的核酸序列是在所述的种子特异性启动子的控制下, 在种子中表达的。

在另一方面, 本发明提供了由下列方法制备的能够结种子的亚麻植物:

(a) 制备在 5' 到 3' 转录方向含有下面可操作连接的成分的嵌合核酸构建体:

- (1) 从亚麻得到的种子特异性启动子; 和
- (2) 需要的核酸序列, 其中所述的需要的核酸对所述的种子特异性

启动子是非天然的;

(b) 将所述的嵌合核酸构建体导入亚麻植物细胞; 和

(c) 使所述的亚麻植物细胞生长成能够结种子的成熟亚麻植物, 其中所述的需要的核酸序列在所述的种子特异性启动子的控制下, 在种子中表达。

在另一方面, 本发明提供了用于亚麻种子和其它植物种类的种子中非天然基因的表达, 例如用于种子的蛋白质或油组成的修饰的新的亚麻种子特异性启动子。

在优选的实施方案中, 种子特异性启动子包含:

(a) 如图 1 (SEQ. ID. NO. :1), 图 2 (SEQ. ID. NO. :4), 图 3 (SEQ. ID. NO. :6) 或图 4 (SEQ. ID. NO. :8) 所示的核酸序列, 其中 T 也可以是 U;

(b) 与 (a) 的核酸序列互补的核酸序列;

(c) 具有与 (a) 或 (b) 的核酸序列的基本序列同源性的核酸序列;

(d) 是 (a), (b) 或 (c) 的核酸序列类似物的核酸序列;

(e) 与 (a), (b), (c) 或 (d) 的核酸序列在严格杂交条件下杂交的核酸序列。

在另一方面, 本发明提供了含有与对所述的第一个核酸序列是非天然的第二个核酸序列可操作地连接的从亚麻得到的第一个核酸序列的嵌合核酸序列, 其中所述的第一个核酸序列含有新的亚麻种子特异性启动子。

本发明的其它特征和优点从下面的详细说明中可以容易地了解。但是, 应该理解, 因为本详细说明书的领域中的技术人员将能够了解本发明的精神和范围内的各种变化和修改, 所以当表示本发明的优选实施方案时, 只通过说明的方式给出了详述和具体的实施例。

附图的简要说明

下面根据附图描述本发明, 其中:

图 1 表示了编码 16.0 kDa 油质蛋白 (SEQ. ID. NO. :2 和 3) 的亚麻基因组克隆的 DNA 序列, (SEQ. ID. NO. :1)。

图 2 表示编码 18.6 kDa 油质蛋白 (SEQ. ID. NO. :5) 的亚麻基因组克隆的 DNA 序列 (SEQ. ID. NO. :4)。

图 3 表示编码 2S 贮藏蛋白 (SEQ. ID. NO. :7) 的亚麻基因组克隆的 DNA 序列 (SEQ. ID. NO. :6)。

图 4 表示编码 54.5kDa 类豆球蛋白贮藏蛋白 (SEQ. ID. NOS: 9-12) 的亚麻基因组克隆的 DNA 序列 (SEQ. ID. NO. :8)。

图 5 表示了用亚麻油质蛋白 DNA 序列探测的亚麻基因组 DNA 的 Southern 印迹分析结果。

图 6 表示了对亚麻油质蛋白的种子特异性表达的 Northern 印迹分析的结果。

图 7 表示了对种子发育过程中亚麻油质蛋白的发育表达的 Northern 印迹分析的结果。

图 8 表示了用亚麻油质蛋白启动子 - GUS - 亚麻终止子基因构建体轰击的亚麻胚的 GUS 活性。

图 9 表示了用 2S 蛋白质基因启动子 GUS 融合体转化的植物的发育中的亚麻胚和拟南芥属种子中的 GUS 表达。

图 10 表示用核丝启动子 - GUS - 核丝终止子基因构建体转化的转基因亚麻植物中 GUS 的组织特异性表达。

图 11 表示了用核丝启动子 - GUS - 核丝终止子基因构建体转化的转基因亚麻植物中 GUS 的暂时表达。

图 12 表示在用核丝启动子 - GUS - 核丝终止子基因构建体转化的转基因油菜植物 (L1 到 L9) 中 GUS 的表达。

图 13 表示了用核丝启动子 - GUS - 核丝终止子基因构建体转化的转基因拟南芥属植物在种子发育的不同时期中 GUS 的表达。

本发明的详细描述

如上文提到, 本发明涉及植物中, 特别是亚麻中非天然基因的表达的改良方法。本发明提供了允许在亚麻中非天然基因的种子特异性表达的方法。本发明的方法的优势是, 得到了对亚麻种子中非天然基因的表

达的更好的控制。非天然基因的表达局限于种子，因此，限制了在其它植物器官或组织中表达产生的潜在的不尽人意的影响。另外，提供的方法允许更好地控制表达特征，如非天然基因的表达水平和植物发育周期中非天然基因的表达计时。本发明的方法特别有用，因为根据本发明，与有价值的原材料如油，蛋白质和多糖相关的种子组成是可以在量和性上发生改变的。

因此，在一个方面，本发明提供了在亚麻种子中表达需要的核酸序列的方法，方法包括步骤：

(a) 制备在 5' 到 3' 转录方向含有可操作地连接的成分的嵌合核酸构建体；

这些成分有：

(1) 从亚麻得到的种子特异性启动子；和

(2) 需要的核酸序列，其中所述的需要的核酸对于所述的亚麻种子特异性启动子是非天然的；

(b) 在亚麻植物细胞中导入所述的嵌合核酸构建体；和

(c) 使所述的亚麻植物细胞生长成能够结种子的成熟植物，其中所述的需要的核酸序列是在所述的亚麻种子特异性启动子的控制下在种子中表达的。

如本文所用，术语“非天然”指任何核酸序列，包括任何 RNA 或 DNA 序列，该序列与种子特异性启动子正常情况不相结合。此核酸序列包括从不同于启动子的植物种类得到的异源核酸序列，以及从相同于启动子的植物种类得到的同源核酸序列，但这些核酸序列与野生型（非转基因）植物的启动子无关。

与从亚麻得到的种子特异性启动子连接时，非天然核酸序列产生了嵌合构建体。将嵌合构建体导入亚麻植物细胞中产生转基因亚麻植物细胞，从而产生了与非转基因亚麻植物细胞或从中生长成的亚麻植物相比明显不同表现型的亚麻植物细胞或从中生长成的亚麻植物。在非转化亚麻植物细胞或从中生长成的亚麻植物中不存在与嵌合构建体的核酸序列相同的邻接核酸序列。在这一方面，嵌合核酸序列包括的那些序列含有

一个亚麻启动子，与从另一个植物种类得到的核酸序列连接，或与来自亚麻，但与该启动子的连接是不正常的核酸序列连接。如本文所用的嵌合核酸序列进一步包括的序列含有亚麻启动子和与该启动子正常连接但还含有非天然核酸序列的核酸序列。例如，如果启动子是亚麻种子特异性油质蛋白启动子，对亚麻油质蛋白启动子是“非天然”的序列也包括含有与油质启动子天然连接的亚麻油质基因和与该启动子非天然连接的需要的编码序列之间的融合体的序列。术语非天然也指含有上文中的融合基因，该基因另外含有分隔与启动子序列正常连接的核酸序列和编码需要的蛋白质的基因的切割序列。

术语“核酸序列”指含有天然存在的碱基，糖和糖之间(支链)的键的核苷酸或核苷单体的序列。该术语也包括含有发挥相似功能的非天然存在的单体或其部分的修饰的或取代的序列。本发明的核酸序列可以是核糖核酸(RNA)或脱氧核糖核酸(DNA)，并且可以含有天然存在的碱基，包括腺嘌呤，鸟嘌呤，胞嘧啶，胸腺嘧啶和尿嘧啶。序列也可以含有修饰碱基，如黄嘌呤，次黄嘌呤，2-氨基腺嘌呤，6-甲基，2-丙基，和其它烷基腺嘌呤，5-卤素尿嘧啶，5-卤素胞嘧啶，6-氨基尿嘧啶，6-氨基胞嘧啶和6-氨基胸腺嘧啶，假尿嘧啶，4-硫尿嘧啶，8-卤素腺嘌呤，8-氨基腺嘌呤，8-硫醇腺嘌呤，8-硫烷基腺嘌呤，8-羟基腺嘌呤和其它8-取代基腺嘌呤，8-卤素鸟嘌呤，8-氨基鸟嘌呤，8-硫醇鸟嘌呤，8-硫烷基鸟嘌呤，8-羟基鸟嘌呤和其它8-取代基鸟嘌呤，其它氮和脱氮尿嘧啶，胸腺嘧啶，胞嘧啶，腺嘌呤，或鸟嘌呤，5-三氟甲基尿嘧啶和5-三氟胞嘧啶。

术语“种子特异性启动子”，指在启动子控制下表达的基因在有或没有基本表达的植物种子中优先表达，而在其它植物组织中通常少于整个表达水平的5%。

在其它方面，本发明提供了在亚麻种子和其它植物种类的种子中用于表达非天然基因的新的亚麻种子特异性启动子。这些启动子可以用于修饰例如种子的蛋白质，油或多糖组成。在优选的实施方案中，种子特异性启动子包括：

- (a) 如图 1 (SEQ. ID. NO. :1), 图 2 (SEQ. ID. NO. :4), 图 3 (SEQ. ID. NO. :6) 或图 4 (SEQ. ID. NO. :8) 中所示的核酸序列, 其中 T 也可以是 U;
- (b) 与 (a) 的核酸序列互补的核酸序列;
- (c) 与 (a) 或 (b) 的核酸序列具有基本序列同源性的所属核酸序列;
- (d) 是 (a), (b) 或 (c) 的核酸序列的类似物的核酸序列; 或
- (e) 与 (a), (b), (c) 或 (d) 的核酸序列在严格杂交条件下杂交的核酸序列。

术语“具有基本序列同源性的序列”指与 (a) 或 (b) 中的序列相比有轻微的或无关紧要的序列变化的那些核酸序列, 这些序列以基本相同的方式发挥作用并且能够驱动非天然核酸序列的种子特异性表达。这些变化可能是由于局部的突变或结构上的修饰。具有基本同源性的核酸序列包括与图 1 (SEQ. ID. NO. :1), 图 2 (SEQ. ID. NO. :4), 图 3 (SEQ. ID. NO. :6) 或图 4 (SEQ. ID. NO. :8) 所示的核酸序列至少有 65%, 更优选地至少 85%, 最优选地 90 - 95% 的同一性的核酸序列。

术语“杂交的序列”指可以在严格杂交条件下与 (a), (b), (c) 或 (d) 的序列杂交的核酸序列。促进 DNA 杂交的适当的“严格杂交条件”是本领域技术人员已知的, 或者可以在《分子生物学》, John Wiley 和 Sons, 纽约 (1989 年), 6.3.1 - 6.3.6 中的通用方案中找到。例如, 可以利用下面的条件: $6.0 \times$ 氯化钠/柠檬酸钠 (SSC), 约 45°C , 接着于 50°C 用 $2.0 \times$ SSC 洗。根据洗涤步骤利用的条件可以选择严格度。例如, 可以从高严格度: 约 $0.2 \times$ SSC, 在 50°C 下选择洗涤步骤中的盐浓度。另外, 洗涤步骤中的温度可以是在约 65°C 的高严格度条件下。

术语“作为类似物的核酸序列”指与 (a), (b) 或 (c) 的序列相比, 已经修饰的核酸序列, 其中的修饰不改变如上所述的序列 (即作为种子特异性启动子) 的用途。修饰的序列或类似物可能具有比 (a), (b) 或 (c) 所示的序列更好的特性。修饰制备类似物的一个例子是用修饰的碱基如黄嘌呤, 次黄嘌呤, 2-氨基腺嘌呤, 6-甲基, 2-丙基和其它烷基腺嘌呤, 5-卤素尿嘧啶, 5-卤素胞嘧啶, 6-氮尿嘧啶, 6-氮胞嘧啶和 6-氮胸腺嘧啶, 假尿嘧啶, 4-硫尿嘧啶, 8-卤素腺嘌呤, 8-氨基腺嘌呤, 8

- 硫醇腺嘌呤, 8- 硫醇烷基腺嘌呤, 8- 羟基腺嘌呤和其它 8- 取代基腺嘌呤, 8- 卤素鸟嘌呤, 8- 氨基鸟嘌呤, 8- 硫醇鸟嘌呤, 8- 硫醇烷基鸟嘌呤, 8- 羟基鸟嘌呤和其它 8- 取代基鸟嘌呤, 其它氮和脱氮尿嘧啶, 胸腺嘧啶, 胞嘧啶, 腺嘌呤, 或鸟嘌呤, 5- 三氟甲基尿嘧啶和 5- 三氟胞嘧啶取代图 1, 图 2, 图 3 或图 4 中所示的序列的一个天然存在的碱基(即, 腺嘌呤, 鸟嘌呤, 胞嘧啶或胸腺嘧啶)。

修饰的另一个例子是包括在图 1, 图 2, 图 3 或图 4 中所示的核酸分子中的磷酸主链, 短链烷基或环烷基糖间键或短链杂原子或杂环糖间键中的修饰的磷或氧杂原子。例如, 核酸序列可以含有硫代磷酸酯, 磷酸三酯, 甲基磷酸酯, 和二硫代磷酸酯。

本发明的核酸分子类似物的另一个例子是肽核酸(PNA), 其中用与肽中发现的相似的聚酰胺主链代替DNA(或RNA)中的脱氧核糖(或核糖)磷酸主链(P.E.Nielsen, 等人, 《科学》1991 年, 254 卷, 1997)。PNA 类似物已经显示对酶的降解具有抗性和在体内和体外具有延长的生命。由于在 PNA 链和 DNA 链之间缺乏电荷排斥, PNA 与互补的 DNA 序列的结合也更强。其它核酸类似物可以含有具有聚合物主链, 环状主链, 或无环主链的核苷酸。例如, 核苷酸可以具有吗啉主链结构(美国专利号, 5, 034, 506)。类似物也可以含有基团, 如报道基团, 一个改进核酸序列的药物动力学或药物动态学特性的基团。

在另一个方面, 本发明提供了嵌合核酸序列, 含有与第二个核酸序列可操作地连接的从亚麻得到的第一个核酸序列, 其中第二个核酸序列对所述第一个核酸序列是非天然的, 其中所述的第一个核酸序列含有新的亚麻种子特异性启动子。优选地, 启动子是从图 1, 图 2, 图 3 和图 4 的启动子的组中选择的或是在严格条件下与它们杂交的核酸序列。

根据本发明, 可以将嵌合核酸序列以已知的方法掺入能保证在种子细胞中有良好的表达的重组表达载体。因此, 本发明包括含有适于在种子细胞中表达的本发明的嵌合核酸序列的重组表达载体。

术语“适于在种子细胞中表达”指重组表达载体含有本发明的嵌合核酸序列, 调节区和终止区, 是根据用于表达的种子细胞选择的, 它与

编码需要的氨基酸组成的多肽的核酸序列操作性地连接。操作性地连接指编码多肽的嵌合核酸序列与调节序列和允许在种子细胞中表达的终止区连接。典型的构建体在5'到3'方向由用能够指导植物中的表达的启动子完成的调节区，多肽编码区和在植物细胞中发挥功能的转录终止区组成。这些构建体可以根据分子生物学领域的技术人员已知的方法来制备(参见例如：Sambrook等人，(1990年)，《分子克隆》第二版，冷泉港出版社)。构建体的制备中可以涉及到如限制消化，连接，凝胶电泳，DNA测序和PCR等技术。许多克隆载体是进行必要的克隆步骤能够采用的。特别适用于本目的是带有在大肠杆菌中发挥功能的复制系统的克隆载体如pBR322，pUC系列，M13mp系列，pACYC184，pBluescript等。可以将核酸序列导入这些载体，并且这些载体可以用于转化大肠杆菌，大肠杆菌可以在适当的培养基中生长。在收获和溶解细胞后，可以从细胞中回收质粒。最后的构建体可以导入适合于整合进入植物的植物载体，如Ti和Ri质粒。

根据本发明在亚麻种子中表达非天然基因的方法可以利用任何亚麻种子特异性启动子实施并且不受选择的具体的亚麻种子特异性启动子的限制。在本发明的优选的实施方案中，亚麻种子特异性启动子给予非天然核酸序列至少一个与天然启动子给予天然核酸序列的表达特征相似或相同的表达特征。如本文所用，术语“表达特征”指亚麻种子特异性启动子给予与亚麻种子特异性启动子可操作地连接的核酸序列的任何可检测的特性或效果。所以，在优选的实施方案中，非天然核酸序列在植物生命周期中表达的计时是与天然核酸序列中的表达计时相似或相同的。在其它优选的实施方案中，异源核酸序列的表达水平与天然核酸序列的表达水平相似或相同。在其它特定的实施方案中，非天然基因对光照条件的改变，波长或光强度的改变，例如温度的改变，组织创伤，化学试剂例如植物激素和杀虫剂的浓度变化的反应，是与天然核酸序列对这些刺激的反应相似的。亚麻种子特异性启动子给予的其它需要的表达特征是能够被本领域技术人员认识到的，并且因此可以选择亚麻种子特异性启动子。

可以根据本发明使用的亚麻种子特异性启动子包括与种子贮藏蛋白如所有清蛋白和球蛋白，包括豌豆球蛋白和类豆球蛋白相关的启动子，以及与种子贮藏蛋白，如油质蛋白不相关的种子特异性启动子。其它特别需要的是与脂肪酸代谢，如酰基载体蛋白 (ACP)，饱和酶，去饱和酶，延伸酶等等相关的启动子。

在本发明的优选的实施方案中，所用的种子特异性启动子是油质蛋白启动子，类豆球蛋白种子贮藏蛋白启动子或 2S 贮藏蛋白启动子。在特别优选的实施方案中，种子特异性启动子具有图 1，图 2，图 3 或图 4 所示的序列或可从亚麻得到并在严格条件下与这四种核酸序列的任一种杂交的任何核酸序列。

根据本发明可以利用其它亚麻种子特异性启动子。这些启动子可以用许多方法得到。分离了亚麻种子蛋白质后，可以部分地测序，从而设计核酸探针用于鉴定种子特异性 RNA。为了进一步增加与种子特异性相关的 RNA，可以从种子细胞制备 cDNA，cDNA 可以减掉来自非种子细胞的 mRNA 或 cDNA。剩余的种子 cDNA 然后可以用于为互补序列探测基因组 DNA 文库。与 cDNA 杂交的序列随后就可以得到，相关的启动子区也可以分离。利用来自其它植物种类的已知种子特异性基因筛选从亚麻种子组织制备的基因组 DNA 文库并随后分离它们的相关启动子也是可能的。由于种子中种子贮藏蛋白相对富裕，通过随机测序亚麻种子 cDNA 文库来获得序列信息也是可能的。与已知的种子贮藏蛋白的序列相匹配的 cDNA 序列可以用于鉴定相关的启动子。对已知的种子特异性蛋白质和/或启动子可以寻找含有来自大规模的例如拟南芥属和玉米的测序得到的序列信息的数据库，该信息可以用于鉴定具有序列相似性的亚麻中的启动子序列。分离其它亚麻种子特异性启动子的备选方法可以利用，本领域技术人员可以发现新的亚麻种子特异性启动子并且根据本发明使用。

与启动子连接的需要的核酸序列可以是任何需要的核酸序列，包括编码需要的肽或蛋白质，例如酶的任何 RNA 或 DNA 序列，或与基因组序列互补的序列，其中基因组序列可以是至少一种可读框，内含子，非编码前导序列，或其中的互补序列将抑制转录，信使 RNA 加工，例如拼接

或翻译的任何序列。该需要的核酸序列可以是合成的，自然起源的或它们的结合体。另外，需要的核酸序列可以是天然序列的片段，例如只包括催化区或特别重要的结构。根据需要的核酸序列的特性，合成具有植物优选密码子的序列可能是需要的。植物优选密码子可以从在需要的特定植物种类中以最大量表达的蛋白质中频度最高的密码子中确定。

需要的核酸序列可以编码任何种类的重组蛋白质，通过本发明的方法可以表达的重组蛋白质的例子包括具有有利的催化功能的蛋白质或将积累到高水平并需要时提取的有价值蛋白质。具有催化功能的蛋白质包括，但不限于给予结种子新的生化表现型的蛋白质。新表现型可以包括如下的修饰：如改变种子蛋白质或种子油组成或种子多糖组成，提高预存在的需要产物或特性的生产，和利用反义，核酶或共抑制技术减少或甚至抑制不需要的基因产物(Izant 和 Weintraub(1984 年)《细胞》26 卷：1007 - 1015 页，反义；Hazelhoff 和 Gerlach(1988 年)《自然》334 卷：585 - 591 页，核酶；Napoli 等人(1990 年)《植物细胞》2 卷：279 - 289 页，共抑制)。

尽管在不同的胚组织如胚轴和子叶中可以检测到改变的细胞表达，但仍然期望需要的蛋白质在所有的胚组织中表达。需要的核酸序列可以在种子发育的任何时期表达。表达的计时可能依据本发明的特定用途而不同。参与油修饰的酶的表达可能在种子发育的早期就需要，例如在种子贮藏蛋白的累积之前。

除了需要的启动子区和核酸序列，能够终止转录的核酸序列通常也包括在表达载体中。转录终止子优选地约为 200 到 1,000 个核苷酸碱基对，可以含有在植物中发挥功能的任何这样的序列，如胭脂碱合酶终止区(Bevan 等人，(1983 年)《核酸研究》11 卷：369 - 385 页)，菜豆蛋白终止子(Van der Geest 等人，(1994 年)《植物杂志》6(3)：413 - 423 页)，根癌土壤杆菌的章鱼氨酸合酶基因的终止子或其它相似地发挥功能的元件。这些转录终止区可以如 An(1987 年)，《酶学方法》153 卷：292 页所述得到，或已经存在于从商业源如 ClonTech, Palo Alto, 加利福尼亚得到的质粒中。合适的终止区的选择可能对转录速率有影响。

嵌合构建体可以进一步含有如 AMV 前导区 (Jobling 和 Gehrke (1987 年), 《自然》 32 卷: 622 - 625 页) 的增强子或内含子。应该理解的是表达载体的设计可根据如植物种类选择和/或待表达的多肽的类型这样的因素而不同。

表达载体正常情况下也将含有标记基因。标记基因包含能够区别已转化植物细胞与非转化细胞的所有基因, 包括可选择并可筛选的标记基因。为了方便, 标记可以是针对除草剂如甘氨酸磷或磷丝菌素, 或针对抗生素如给予通过化学方法可以选择的特性的卡那霉素, G418, 博来霉素, 潮霉素, 氯霉素等等的抗性标记。可筛选的标记可用于观察鉴定转化体。它们包括但不限于 β - 葡糖醛酸糖苷酶或 uidA 基因, β - 内酰胺酶基因或绿荧光蛋白质 (Niedz 等人 (1995 年) Plant Cell Rep., 14 卷: 403 页)。

为了在植物细胞中导入核酸序列, 通常, 技术人员可以利用各种技术。土壤杆菌介导的亚麻植物细胞转化已经有报道, 虽然其它各种技术 (参见下面) 也可以用于需要在亚麻细胞中导入嵌合 DNA 构建体, 但根据 Dong 和 McHughen (1993 年) 《植物科学》 88 卷: 61 - 77 页教授的方法可以得到亚麻转化体。

根据本领域的技术人员已知的常规农业实践培养的转化亚麻植物是允许结种子的。从而可从成熟的亚麻植物得到亚麻种子, 并分析其需要的相对于野生型种子已改变的特性。

可以使两代或更多代植物生长并且杂交或自交, 从而能鉴定出具有需要的表现型特征的植物和品系包括产生重组多肽。保证植物的纯合性以保证重组特性能够连续地遗传是需要的。选择纯合的植物的方法是植物育种领域的技术人员已知的, 包括轮回自花授粉和选择和花药及小孢子培养。通过转化单倍体细胞或组织, 接着繁殖单倍体小植物, 随后通过任何已知的各种方法 (例如, 用秋水仙素或其它微管裂解试剂处理) 转化成二倍体植物也可以得到纯合植物。

本发明也可以包括根据下列方法制备的转基因亚麻种子, 所述方法包括:

(a) 制备 5' 到 3' 转录方向上含有下列可操作地连接的成分的嵌合

核酸构建体:

(1) 从亚麻得到的种子特异性启动子; 和

(2) 需要的核酸序列, 其中所述的需要的核酸对所述的种子特异性启动子是非天然的;

(b) 在亚麻植物细胞中导入所述的嵌合核酸构建体; 和

(c) 使所述的亚麻植物细胞生长成能够结种子的成熟亚麻植物,

其中所述的需要的核酸序列是在所述的种子特异性启动子的控制下在种子中表达的。

在本发明的优选的实施方案中, 种子特异性启动子是从亚麻种子特异性启动子的组, 包括 2S 贮藏蛋白启动子, 球蛋白启动子, 油质蛋白启动子, 和类豆球蛋白种子贮藏蛋白启动子中选择的。图 1 (SEQ. ID. NO. :1), 图 2 (SEQ. ID. NO. :4), 图 3 (SEQ. ID. NO. :6) 和图 4 (SEQ. ID. NO. :8) 中表示了特定的启动子序列。

本发明进一步提供能够用下列方法制备的能够结种子的亚麻植物:

(a) 制备在 5' 到 3' 转录方向含有下列可操作连接成分的嵌合核酸构建体:

(1) 从亚麻得到的种子特异性启动子; 和

(2) 需要的核酸序列, 其中需要的所述核酸对所述的种子特异性启动子是非天然的;

(b) 在亚麻植物细胞中导入所述嵌合核酸构建体; 和

(c) 使所述的亚麻植物细胞生长成能结种子的成熟亚麻植物,

其中所述的需要的核酸序列是在所述的种子特异性启动子的控制下在种子中表达的。

本发明进一步提供了图 1 (SEQ. ID. NO. :1), 图 2 (SEQ. ID. NO. :4), 图 3 (SEQ. ID. NO. :6) 和图 4 (SEQ. ID. NO. :8) 中所示的新启动子在亚麻以外的植物种类中的使用方法。因此, 本发明也包括含有从图 1, 图 2, 图 3 和图 4 所示的启动子的组中选择的启动子和需要的核酸序列的嵌合核酸构建体的制备, 和以种子特异性方法, 将需要的核酸序列在亚麻以外的植物种类中并且在亚麻启动子的控制下进行表达。

在本发明的另一方面，提供了在植物种子中表达需要的核酸序列的方法，该方法包括：

(a) 制备在 5' 到 3' 转录方向含有下列可操作地连接的成分的嵌合核酸构建体：

(1) 从由下面成分组成的种子特异性启动子的组中选择的种子特异性启动子：

(i) 图 1 (SEQ. ID. NO. :1), 图 2 (SEQ. ID. NO. :4), 图 3 (SEQ. ID. NO. :6) 或图 4 (SEQ. ID. NO. :8) 所示的核酸序列，其中 T 也可以是 U；

(ii) 与 (i) 的核酸序列互补的核酸序列；

(iii) 与 (i) 或 (ii) 的核酸序列具有基本序列同源性的核酸序列；和

(iv) 是 (i), (ii) 或 (iii) 的核酸序列的类似物的核酸序列；

(v) 与 (i), (ii), (iii) 或 (iv) 的核酸序列，在严格条件下杂交的核酸序列；和

(2) 所述的需要的核酸；

(b) 将嵌合核酸构建体导入植物细胞；

(c) 使所述的植物细胞生长成能够结种子的成熟植物；其中所述的需要的核酸序列是在所述的种子特异性启动子的控制下在种子中表达的。

为了将核酸序列，特别是 DNA 导入植物宿主细胞，通常可以利用各种技术，例如，可以利用如 Moloney 等人 (1989 年)，Plant Cell Rep. ; 8 卷: 238 - 242 页或 Hinchee 等人 (1988 年)《生物技术》6 卷: 915 - 922 页所述的转化方案，或本领域技术人员已知的其它技术，利用标准土壤杆菌载体，可以将嵌合 DNA 构建体导入从双子叶植物，如烟草和产生油的种类，如欧洲油菜 (Brassica napus) 得到的宿主细胞。例如，利用 T - DNA 转化植物细胞已经进行了广泛的研究并且已经在 EP 0 120 516, Hoekema 等人, (1985 年)，第 V 章:《二元植物载体系统》Offset - drukkerij Kanters BV, Alblasterdam) 中; Knauf 等人, (1983 年)，《土壤杆菌的宿主表达的遗传学分析》，245 页，《细菌 - 植物的相互作用的分子遗传学》，Puhler, A 编辑, Springer - Verlag, 纽约中心; 和 An 等人, (1985 年)，《EMBO 杂志》，4 卷: 277 - 284 页) 中充分叙述。土壤杆

菌转化也可以用于转化单子叶植物种类(美国专利 5,591,616)。

为了方便,利用根癌土壤杆菌或发根病土壤杆菌培养外植体能够使转录构建体转移进植物宿主细胞。在利用土壤杆菌转化后,将植物细胞分散在适当的培养基中选择,然后回收愈伤组织,茎干和最终的植物。土壤杆菌宿主细胞将含有质粒,该质粒含有将 T-DNA 转移到植物细胞所必需的 vir 基因。为了注入和电穿孔(参见如下),可以将去臂的 Ti 质粒(缺乏肿瘤基因),特别是 T-DNA 区)可以导入植物细胞。

非土壤杆菌技术的利用允许利用本文所述的构建体获得在许多单子叶和双子叶植物种类中的转化和表达。这些技术特别可用于土壤杆菌转化系统中难处理的植物种类的转化。基因转移的其它技术包括粒子轰击(Sanford, (1988),《生物技术趋势》6 卷: 299-302 页,电穿孔(Fromm 等人, (1985 年),《美国国家科学院院报》,82 卷: 5824-5828 页; Riggs 和 Bates, (1986 年),《美国国家科学院院报》,83 卷: 5602-5606 页),PEG 介导的 DNA 吸收(Potrykus 等人, (1985 年),《分子基因遗传学》,199 卷: 169-177 页),显微注射(Reich 等人,《生物技术》(1986 年)4 卷: 1001-1004 页)和硅氧烷碳化物细丝(Kaeppler 等人(1990 年)Plant Cell Rep., 9 卷: 415-418 页)。

在其它特定的应用中,如在欧洲油菜中,如 Moloney 等人, (1989 年)Plant Cell Rep., 8 卷: 238-242 页所述,被靶击接受重组 DNA 构建体的宿主细胞通常将是来自子叶的叶柄得到的。利用有商业用途的油种子的其它例子包括在大豆外植体中转化子叶(Hinchee 等人, (1988 年)《生物技术》,6: 915-922 页)和棉花的茎转化(Umbeck 等人, (1987 年)《生物技术》,5 卷: 263-266 页)。

在转化后,细胞,例如叶片是在选择培养基上生长的。一旦茎干开始出现,就把它切下,放置在长根的培养基上。在形成足够的根后,将植物转移到土壤中。然后,测试推定的转化植物中标记的存在。利用适当的探针对基因组 DNA 进行 Southern 印迹分析,表明整合到宿主细胞的基因组中。

本发明提供的方法可以和大范围的植物种类结合使用。根据本发明

使用的特别优选的植物细胞包括来自下列植物的细胞：大豆(*Glycine max*)，油菜(*Brassica napus*, *Brassica campestris*)，向日葵(*Helianthus annuus*)，棉花(*Gossypium hirsutum*)，玉米(*Zea mays*)，烟草(*Nicotiana, tabacum*)，苜蓿(*Medicago Sativa*)，小麦(*Triticum sp.*)，大麦(*Hordeum vulgare*)，燕麦(*Avena Sativa L.*)，高粱(*Sorghum bicolor*)，拟南芥(*Arabidopsis thaliana*)，土豆(*Solanum sp.*)，亚麻/亚麻子(*Linum usitatissimum*)，红花(*Carthamus tinctorius*)，油棕榈(*Eleais guineeis*)，落花生(*Arachis hypogaea*)，巴西坚果(*Bertholletia excelsa*)，椰子(*Cocus nucifera*)，蓖麻(*Ricinus communis*)，芫荽(*Coriandrum sativum*)，南瓜(*Cucurbita maxima*)，灌木(*Simmondsia chinensis*)和水稻(*Oryza sativa*)。

本发明具有各种用途，包括通过积累已改变的多肽或新的重组肽或通过引入或去除代谢步骤，提高植物种子的内在价值。利用本发明可能导致提高蛋白质的质量(例如，增高必需的或稀有的氨基酸的浓度)，通过修饰脂肪组成提高液体质量，或改良或增多的碳水化合物组成。例子包括富硫蛋白质如在含硫氨基酸缺乏的种子中，在白羽扇豆或巴西坚果中发现的那些富硫蛋白质的表达。蛋白质质量的改良也可以通过必需氨基酸，包括赖氨酸，半胱氨酸，甲硫氨酸和色氨酸丰富的蛋白质或蛋白质片段的表达来实现。或者，可以表达脂肪酰基辅酶A，一种能够改变甘油三酯(贮藏脂)中脂肪酸的比例的转移酶。在种子中允许积累重组蛋白质的情况中，蛋白质也可以是具有药学或工业价值的肽。在这种情况下，可以从种子中提取肽，并且如果对预期用途合适，可以粗品或纯化的形式利用。同样，在种子中表达的多肽可以是天然蛋白质的片段或衍生物。可药用的蛋白质可以包括，但不限于抗凝剂，如水蛭素，抗体，包括单克隆抗体和抗体片段，疫苗，细胞因子或生长因子如牛生长因子，胆碱能分化因子(CDF)，睫状神经营养因子(CNTF)，成纤维细胞生长因子(FGF)，鱼生长因子，促性腺激素，粒细胞集落刺激因子(G-CSF)，粒细胞-巨噬细胞集落刺激因子(GM-CSF)，人生长激素，干扰素 α (IFN- α)，干扰素 β (IFN- β)，干扰素 γ (IFN- γ)，白介素1- α (IL1-

α), 白介素 1- β (IL1- β), 白介素-2 (IL-2), 白介素-3 (IL-3), 白介素-4 (IL-4), 白介素-5 (IL-5), 白介素-6 (IL-6), 白介素-10 (IL-10), 白血病抑制因子 (LIF), 硫氧还蛋白, 巨噬细胞集落刺激因子 (M-CSF), 骨髓单核细胞生长因子, 神经生长因子 (NGF), 制癌蛋白 M, 血小板起源生长因子 (PDGF), 促乳素, 转化生长因子 α (TGF- α), 转化生长因子 β 2 (TGF- β 2), 肿瘤坏死因子 α (TNF- α), 肿瘤坏死因子 β (TNF- β)。药物用途的蛋白质也可以包括哺乳动物蛋白质, 例如, 但不限于 α -1-抗胰蛋白酶, 抗肥胖蛋白, 血液蛋白, 胶原蛋白, 胶原酶, 弹性蛋白, 弹性蛋白酶, 肠肽酶, 纤维蛋白原, 血红蛋白, 人血清白蛋白, 胰岛素, 乳铁蛋白, 肌红蛋白和肺表面活性蛋白。

有工业用途的肽可能包括, 但不限于 α -淀粉酶或其它淀粉酶, 淀粉葡萄糖苷酶, 阿拉伯聚糖酶, 过氧化氢酶, 纤维二糖水解酶, 纤维素酶, 几丁质酶, 胰凝乳蛋白酶, 脱氢酶, 内葡聚糖酶, 凝乳酶, 内半乳聚糖酶, 酯酶, β -半乳糖苷酶, α -半乳糖苷酶或其它半乳糖苷酶, 胃脂肪酶, 葡聚糖酶, 葡萄糖异构酶, 半纤维素酶, 水解酶, 异构酶, 木素酶, 脂酶, 裂解酶, 溶菌酶, 氧化酶, 氧化还原酶, 木瓜蛋白酶, 果胶酶, 果胶裂合酶, 过氧化物酶, 磷酸酶, 植酸酶, 蛋白酶, 支链淀粉酶, 还原酶, 丝氨酸蛋白酶, 硫氧还蛋白, 转移酶, 胰蛋白酶, 和木聚糖酶。

下面的非限制实施例是本发明的解释:

实施例

实施例 1

种子特异性亚麻启动子的分离

种子特异 cDNA 克隆是从亚麻种子特异性 cDNA 文库中分离到的。将这些 cDNA 克隆测序, 利用 Basic Local Alignment Search Tool (BLAST) 比较这些序列和公共数据库如基因库中的其它序列。这一比较发现, 几个分离的 cDNA 的推测的氨基酸序列与低的和高的分子量类别的油质蛋白, 2S-白蛋白和类豆球蛋白的贮藏蛋白具有高度的相似性。分别从编码油质蛋白, 2S 白蛋白和类豆球蛋白的贮藏蛋白的 cDNA 部分制备探针, 并利用这些探针筛选从亚麻品系 Forge 制备的基因组文库, 该品系对于四个

锈病抗性基因是纯合的(Anderson 等人, (1997 年), 《植物细胞》9 卷: 641 - 651 页)。在高度严格筛选后, 鉴定对于每个探针的几个阳性 λ 克隆。将插入片段亚克隆进质粒载体 pBluescript 并测序。序列信息揭露, 我们已经分离到了基因组中油质蛋白, 2S 白蛋白的相当部分和 cDNA 类豆球蛋白 cDNA。图 1 到 4 分别表示了含有编码高和低分子量油质蛋白同工型, 2S 白蛋白的序列的基因组克隆和类豆球蛋白基因的序列信息。

图 1 和 SEQ. ID. NO. :1 表示了编码 16.0kDa 油质蛋白(低分子量或 L - 同工型)的亚麻基因组克隆的 DNA 序列。推测的调节元件得到了鉴定并指示出。这些调节元件包括倒转重复(碱基对 805 到 813 和 821 到 829; 碱基对 1858 到 1866 和 1877 到 1885), 同向重复(碱基对 184 到 193 和 1102 和 1111; 碱基对 393 到 402 和 1701 到 1710; 碱基对 683 到 692 和 1546 到 1555; 碱基对 770 到 781 和 799 到 810; 碱基对 955 到 964 和 1936 到 1945; 碱基对 1483 到 1496 和 1513 到 1526), 抗坏血酸应答元件 (ABRE) (碱基对 1859 到 1866), CACA 盒子(碱基对 1933 到 1936), TATA 盒(碱基对 1925 到 1931)和 CAT 盒(碱基对 3020 到 3025)。可读框是用 1 个短内含子(已标记)中断的并且 2 个外显子翻译了, 并且 IUPAC 单字母氨基酸密码表示。

图 2 和 SEQ. ID. NO. :4 表示了编码 18.6kDa 的油质蛋白(高分子量或 H - 同工型)的亚麻基因组克隆的 DNA 序列。推测的调节元件鉴定并指示了。这些元件包括同向重复(碱基对 14 - 25 和 1427 - 1438; 碱基对 80 - 89 和 1242 - 1251; 碱基对 177 - 186 和 837 - 846; 碱基对 1281 - 1290 和 1242 - 1251; 碱基对 1591 - 1600 和 1678 - 1287)。这一可读框没有被内含子中断, 翻译了和用 IUPAC 单字母氨基酸密码表示了。

图 3 和 SEQ. ID. NO. :6 表示了编码 2S 贮藏蛋白的亚麻基因组克隆的 DNA 序列。对这一克隆的核苷酸序列分析表明它具有一个 174 个氨基酸的可读框, 这些氨基酸表现出与植物 2S 贮藏蛋白组有同源性。该序列编码与甘蓝 (*Brassica oleracea*) 2S 贮藏蛋白具有 38% 的整体相似性, 包括完全保守的一段富谷氨酸序列 QQQGQQGQQQ (SEQ. ID. NO. :13) 一个可读框。另外, 2S 贮藏蛋白基因启动子含有几个推测的启动子调节元件。这

些元件包括富 AT 的重复(碱基对 25 - 36, 97 - 108 和 167 - 190), RY 类似重复(碱基对 240 - 247), G- 盒 - 类似元件(碱基对 274 - 280), 种子特异性盒类似基序(碱基对 285 - 290)和 TATA 盒(碱基对 327 - 333)。

图 4 和 SEQ. ID. NO. :8 表示了编码 54.4kDa 的类亚麻豆球蛋白种子贮藏蛋白的亚麻基因组克隆的 DNA 序列。这一类豆球蛋白种子贮藏蛋白基因也可称为“核丝”。将核丝基因的推测的氨基酸序列与来自 *R. Communis* 的类豆球蛋白蛋白质, 来自 *M. salicifolia*, *Q. robur* 和 *G. hirsutum* 的豆球蛋白前体, 来自 *O. Sativa* 的谷蛋白前体和来自 *A. thaliana* 的 12S 种子贮藏蛋白比较。核丝基因表现出与来自 *R. Communis*, *M. salicifolia*, *Q. robur*, *G. hirsutum*, *O. Saliva* 和 *A. thaliana* 的相应的蛋白质分别具有 59/15, 47/16, 50/17, 45/17, 43/18 和 43/18 百分数的序列同一性/相似性。在启动子区中的推测的调节元件被鉴定和指明。这些元件包括倒转重复(碱基对 265 到 276 和 281 到 292; 碱基对 513 到 524 和 535 到 545), 重复(碱基对 1349 到 1360 和 1367 到 1378; 碱基对 1513 到 1529 和 1554 到 1572), 抗坏血酸应答元件(ABRE)(碱基对 1223 和 1231), 豆球蛋白盒(RY 重复)(在碱基对 1223 和 1231 之间), 可能的豌豆球蛋白盒区(碱基对 1887 到 1894), CAAT 盒(碱基对 1782 到 1785)和 TATA 盒(碱基对 1966 到 1970)。同样, 指示了用于 ER 膜导向的信号肽(碱基对 2034 - 2080)。可读框是用 3 个短内含子(被标记的)中断的, 并且 4 个外显子被翻译并用 IUPAC 单字母氨基酸代码表示。

图 5 表示了亚麻基因组 DNA 的 Southern 印迹分析。从叶子中分离了 60 μ g 亚麻基因组 DNA, 用 EcoRI(泳道 1), HindIII(泳道 2)和 BamHI(泳道 3)消化, 并加载于各自的泳道上。A) 用随机引导的 32 P- 标记的 3T cDNA(高分子量亚麻油质蛋白同工型)进行杂交。B) 用随机引导的 32 P- 标记的 7R cDNA(低分子量油质蛋白同工型)进行杂交。结果证明, 3T(高分子量油质蛋白同工型)和 7R(低分子量油质蛋白同工型)油质蛋白 cDNA 与亚麻基因组 DNA 杂交。更具体地说, 这些结果表明, 3T 似乎是代表亚麻中的 2- 拷贝基因, 正如每条消化泳道中的两条带所见到的。同样, 7R 似乎代表了亚麻中的多个基因家族, 因为每条消化泳道都检测到多个条