



(12) 发明专利申请

(10) 申请公布号 CN 101809155 A

(43) 申请公布日 2010.08.18

(21) 申请号 200880101619.4

C12N 15/53 (2006.01)

(22) 申请日 2008.08.01

C12N 15/29 (2006.01)

(30) 优先权数据

C07K 14/415 (2006.01)

60/953,562 2007.08.02 US

A01H 5/00 (2006.01)

(85) PCT申请进入国家阶段日

2010.02.02

(86) PCT申请的申请数据

PCT/EP2008/060106 2008.08.01

(87) PCT申请的公布数据

W02009/016249 EN 2009.02.05

(71) 申请人 巴斯夫植物科学有限公司

地址 德国路德维希港

(72) 发明人 A·舍利

(74) 专利代理机构 北京市中咨律师事务所

11247

代理人 黄革生 林柏楠

(51) Int. Cl.

C12N 15/82 (2006.01)

C12N 9/02 (2006.01)

权利要求书 1 页 说明书 35 页 序列表 50 页
附图 4 页

(54) 发明名称

具有增加的胁迫耐受性和产量的转基因植物

(57) 摘要

本发明公布了多聚核苷酸,其能够使转化后含有这种多聚核苷酸的植株在水有限的条件下增加生长、产量,和/或增加对环境胁迫的耐受性。本发明还提供了使用这种多聚核苷酸的方法,以及包含这样的多聚核苷酸作为转基因的转基因植株和农产品,包括种子。

1. 用表达盒转化的转基因植株,所述表达盒包含分离的多聚核苷酸,其编码的多肽的序列选自 SEQ ID NO :2, SEQ ID NO :4, SEQ ID NO :6, SEQ ID NO :8 和 SEQ ID NO :10。

2. 用表达盒转化的转基因植株,所述表达盒包含分离的多聚核苷酸,其编码的多肽的序列选自 SEQ ID NO :12, SEQ ID NO :14 和 SEQ ID NO :16。

3. 用表达盒转化的转基因植株,所述表达盒包含分离的多聚核苷酸,其编码的多肽的序列选自 SEQ ID NO :18, SEQ ID NO :20, SEQ ID NO :22, SEQ ID NO :24, SEQ ID NO :26, SEQ ID NO :28, SEQ ID NO :30, SEQ ID NO :32, SEQ ID NO :34, SEQ ID NO :36, SEQ ID NO :38, SEQ ID NO :40, SEQ ID NO :42 和 SEQ ID NO :44。

4. 分离的多聚核苷酸,其序列选自表 1 所列的多聚核苷酸序列。

5. 分离的多肽,其序列选自表 1 所列的多肽序列。

6. 生产包含表 1 所列的至少一种多聚核苷酸的转基因植株的方法,其中多聚核苷酸在植株中的表达导致了植株与野生型的多种植株相比,在正常或者水有限的条件下增加了生长和 / 或产量,和 / 或增加了对环境胁迫的耐受性,所述方法包括下列步骤:

(a) 将包含表 1 所列的至少一种多聚核苷酸的表达载体引入植物细胞,

(b) 由该植物细胞产生表达多聚核苷酸的转基因植株,其中多聚核苷酸在转基因植株中的表达使得植株与野生型的多种植株相比,在正常或者水有限的条件下增加生长和 / 或产量,和 / 或增加对环境胁迫的耐受性。

7. 增加植株在正常或者水有限的条件下的生长和 / 或产量、和 / 或增加植株对环境胁迫的耐受性的方法,所述方法包含在植株中增加表 1 所列的至少一种多聚核苷酸的表达的步骤。

具有增加的胁迫耐受性和产量的转基因植物

[0001] 本申请要求提交于 2007 年 8 月 2 日的美国临时专利申请序列号 60/953562 的优先权权益,该申请的内容通过引用并入本文。

发明领域

[0002] 一般而言,本发明涉及过量表达编码下述多肽的核酸序列的转基因植物,所述多肽能在正常或非生物胁迫条件下赋予增加的胁迫耐受性以及由此增加的植物生长和作物产量。此外,本发明涉及新颖的经分离的编码下述多肽的核酸序列,所述多肽向植物赋予:非生物胁迫条件下,增加的耐受性,和 / 或,正常或非生物胁迫条件下,增加的植物生长和 / 或增加的产量。

[0003] 发明背景

[0004] 非生物环境胁迫,例如干旱、盐度、热和冷,是植物生长和作物产量的主要限制性因素。作物产量在本文中被定义为每英亩收获的相关农业产品(例如谷粒(grain)、草料(forage)或种子)的蒲式耳数量。这些胁迫导致的多种作物(例如大豆、稻、玉米(玉蜀黍)、棉花和小麦)的作物损失和作物产量损失是显著的经济和政治因素,它们在很多非发达国家导致食品短缺。

[0005] 水可利用性是非生物胁迫及其对植物生长的影响的重要方面。持续暴露给干旱条件导致植物代谢的重大改变,这最终引起细胞死亡以及由此导致产量损失。因为一些土壤中的高盐含量导致细胞可摄入的水较少,高盐浓度对植物产生类似于干旱对植物的影响。此外,在冷冻的温度下,由于植物内形成冰,植物细胞丢失水。因此,干旱、热、盐度和冷胁迫导致的作物损伤主要都是由于脱水导致的。

[0006] 因为植物在其生命周期期间一般暴露给水可利用性降低的环境,大多数植物都进化出了针对非生物胁迫导致的干燥的保护性机制。但是,如果干燥条件的严重性太高以及持续时间太长的话,对大多数作物植物的发育、生长、植物尺寸和产量会有很大影响。因此,开发有效使用水的植物是可能能在世界范围内显著改进人类生活的策略。

[0007] 传统植物育种策略相对较慢,并且需要非生物胁迫耐受的基础品系与其它种质杂交,以开发新的非生物胁迫抗性品系。对于此类基础品系来说有限的种质来源以及远族相关的植物物种之间杂交的不相容性代表了传统育种中遇到的显著问题。针对耐受性的育种还远未成功。

[0008] 很多农业生物技术公司已试图鉴定能赋予对非生物胁迫应答的耐受性的基因,努力开发转基因的非生物胁迫耐受性作物植物。虽然涉及植物中胁迫应答或水分利用效率的一些基因已被分析,但是,对赋予胁迫耐受性和 / 或水分利用效率的植物基因的分析 and 克隆仍相当不完全,还仅是碎片式的。迄今为止,在开发转基因非生物胁迫耐受性作物植物上取得的成功是有限的,还没有此类植物被商品化。

[0009] 为了开发转基因非生物胁迫耐受性的作物植物,必须要在模式植物系统、作物植物的温室研究和田间试验中检验多个参数。例如,水分利用效率(WUE)是通常与干旱耐受性相关的参数。对植物对干燥、渗透压震扰和极端温度的应答的研究也被用于测定植物对非

生物胁迫的耐受性或抗性。当针对转基因的存在对植物的胁迫耐受性的影响加以测试时，对土壤性质、温度、水和养分利用率以及光强度进行标准化的能力是温室或植物生长室环境较之田间的固有优势。

[0010] WUE 已通过多种途径被定义和测量。一种手段是计算整株植物干重与植物生命期间消耗的水的重量之比。另一变化是当测量生物量积累和水分利用时，采用较短的时间间隔。还一手段使用来自植物的局限部份的测量，例如，仅测量地上部分的生长和水分利用。WUE 还被定义为来自叶或叶的部分的 CO_2 摄取与水蒸气损失的比例，这通常在非常短的时间期间上测量（例如数秒 / 数分钟）。用同位素比质谱仪测量的植物组织中固定的 $^{13}\text{C}/^{12}\text{C}$ 的比例也被用于评估利用 C_3 光合作用的植物中的 WUE。

[0011] WUE 的增加指示生长和水消耗相对提高效率，但是该信息单独采用时并不能表明这两种过程中的一种改变或者两者都改变。在选择用于改善作物的性状时，由于水分利用减少导致的 WUE 增加（生长不改变）将在水输入费用较高的灌溉农业系统中具有特别的益处。主要由于生长增加驱动的 WUE 增加（而没有相应的水分利用跃增）将可用于所有农业系统。在水供应不受限的很多农业系统中，生长的增加，即使其随着水分利用的增加花费（即，WUE 无改变），也可增加产量。因此，需要增加 WUE 和生物量积累两者的新方法改善农业生产。

[0012] 伴随对与非生物胁迫耐受性相关的参数的测量，还测量指示转基因对作物产量的潜在影响的参数。对于叶类作物，例如，苜蓿、青贮玉米和干草而言，植物生物量与总产量相关。但是对于谷粒作物而言，已用其它参数来估计产量，例如植物尺寸（通过总的植物干重测量）、地上部分干重、地上部分鲜重、叶面积、茎体积、植物高度、莲座直径、叶长度、根长度、根量、分蘖数量和叶数量。通常，早期发育阶段的植物尺寸将与发育晚期的植物尺寸相关。通常，具有较大叶面积的较大植物可比较小的植物吸收更多的光和二氧化碳，因此在相同期间可能获得更多的重量。除了可能的微环境持续或遗传优点之外，植物必须最初就获得较大的尺寸。植物尺寸和生长速率存在强烈的遗传因素，因此对于一系列多种基因型而言，一种环境条件下的植物尺寸可能与另一种条件下的尺寸相关。以这种方式，用标准环境来近似田间作物在不同位置和时间遇到的多样的和动态的环境。

[0013] 收获指数——种子产量与地上部分干重之比，在很多环境条件下都相对稳定，因此植物尺寸和谷粒产量之间可能强相关。植物尺寸和谷粒产量内在联系，因为谷粒生物量中的大部分都依赖于植物茎叶流动或贮藏的光合作用生产力。因此，已经通过选择植物尺寸（甚至是在发育早期阶段）来筛选出暴露给田间测试时可能展示出增加的产量的植物。至于非生物胁迫耐受性，生长室或温室中标准条件下对早期发育时的植物尺寸的测量是测量转基因的存在所赋予的可能的产量优点的标准实践。

[0014] 因此，人们需要鉴定出在胁迫耐受性植物和 / 或在使用水中有有效的植物中表达的其它基因，它们能向宿主植物和其它植物物种赋予胁迫耐受性和 / 或增加的水分利用效率。新产生的胁迫耐受性植物和 / 或具有增加的水分利用效率的植物将具有很多优点，例如，可种植作物植物的范围增加，例如通过减少植物物种的水需求而实现。其它想要的优点包括增加的对枝条或茎应答于风、雨、害虫或疾病的倒伏、弯曲的抗性。

[0015] 发明概述

[0016] 本发明已经公开：用某些多核苷酸转化植物，当所述多核苷酸作为转基因存在于

植物中时,导致植物生长和对环境胁迫的应答增强,因而增加了植物的农业产品的产量。能介导此类增强的多核苷酸已从欧洲油菜 (*Brassica napus*)、稻 (*Oryza sativa*)、大豆 (*Glycine max*)、小麦 (*Triticum aestivum*)、大麦 (*Hordeum vulgare*)、玉米 (*Zea mays*) 和亚麻 (*Linum usitatissimum*) 中分离,并列在表 1 中,其序列在表 1 所示的序列表中示出。

[0017] 表 1

[0018]

基因 ID	生物体	多聚核苷酸 SEQ ID	氨基酸 SEQ ID
		NO	NO
BN51364980	欧洲油菜	1	2
OS34096188	稻 (<i>O. sativa</i>)	3	4
OS32583643	稻	5	6
GM53626178	大豆 (<i>G. max</i>)	7	8
TA56540264	小麦 (<i>T. aestivum</i>)	9	10
BN45206322	欧洲油菜	11	12
GM48923793	大豆	13	14
TA55969932	小麦	15	16
BN47310186	欧洲油菜	17	18
基因 ID	生物体	多聚核苷酸 SEQ ID	氨基酸 SEQ ID
		NO	NO
BN51359456	欧洲油菜	19	20
HV62552639	大麦 (<i>H. vulgare</i>)	21	22
ZM61995511	玉米 (<i>Z. mays</i>)	23	24
LU61567101	亚麻 (<i>L. usitatissimum</i>)	25	26
LU61893412	亚麻	27	28
OS39781852	稻	29	30

基因 ID	生物体	多聚核苷酸 SEQ ID NO	氨基酸 SEQ ID NO
OS34701560	稻	31	32
OS36821256	稻	33	34
GM51659494	大豆	35	36
GM49780101	大豆	37	38
GM59637305	大豆	39	40
TA55974113	小麦	41	42

[0019]

[0020] 在一种实施方式中,本发明提供了经包含经分离的下述多核苷酸的表达盒转化的转基因植物,所述多核苷酸编码甲硫氨酸亚砷还原酶家族蛋白质,其选自 SEQ ID NO :2、SEQ ID NO :4、SEQ ID NO :6、SEQ ID NO :8 和 SEQ ID NO :10,或者编码 Se1R 蛋白质区域。

[0021] 在另一实施方式中,本发明提供了经包含经分离的下述多核苷酸的表达盒转化的转基因植物,所述多核苷酸编码同源异型域-亮氨酸拉链蛋白质,其具有 SEQ ID NO :12、SEQ ID NO :14 或者 SEQ ID NO :16 所示序列。

[0022] 在另一实施方式中,本发明提供了经包含经分离的下述多核苷酸的表达盒转化的转基因植物,所述多核苷酸编码含有 A20 区域与 AN1 区域的锌指蛋白质,选自 SEQ ID NO :18、SEQ ID NO :20、SEQ ID NO :22、SEQ ID NO :24、SEQ ID NO :26、SEQ ID NO :28、SEQ ID NO :30、SEQ ID NO :32、SEQ ID NO :34、SEQ ID NO :36、SEQ ID NO :38、SEQ ID NO :40、SEQ ID NO :42 和 SHQ ID NO :44,或者编码 AN1- 样和 A20- 样的锌指蛋白质结构域。

[0023] 在又一实施方式中,本发明涉及到本发明中的转基因植株产生的种子,其中种子是包含上述多聚核苷酸的转基因的纯种。来自本发明的种子的植株与野生型的各种植株相比,在正常或者胁迫的条件下,对环境胁迫压力的耐受性增加,和 / 或植株生长增加,和 / 或产量增加。

[0024] 在另一方面,本发明涉及由本发明的转基因植株、其植株的部分、或其种子生产或由其制成的产品,比如食品、饲料、食品补剂、饲料补剂、化妆品或药物。

[0025] 本发明还提供了下面的表 1 中所鉴定的分离的多聚核苷酸和表 1 中所鉴定的多肽。本发明还用包含本发明的分离的多聚核苷酸的重组载体具体实施。

[0026] 在另一实施方式中,本发明涉及到制造上述的转基因植株的方法,其中该方法包含用含有本发明分离的多聚核苷酸的表达载体转化植株细胞,从这个植株细胞产生表达该多聚核苷酸编码的多肽的转基因植株。与野生型的多种植株相比,这个多肽在植株中的表达使植株在正常和 / 或者有胁迫的条件下,增加了对环境胁迫压力的耐受性、和 / 或者植株生长、和 / 或者产量。

[0027] 在另一实施方式中,本发明提供了增加植物对环境胁迫的耐受性和/或生长和/或产量的方法。所述方法包括下述步骤:用包含本发明的经分离的多核苷酸的表达盒转化植物细胞,以及从植物细胞产生转基因植物,其中所述转基因植物包含所述多核苷酸。

[0028] 附图概述

[0029] 图 1:BN51364980、OS34096188、OS32583643、GM53626178、TA56540264 与已知的甲硫氨酸亚砷还原酶家族蛋白质的比对。

[0030] 图 2:BN45206322、GM48923793 和 TA55969932 与已知的同源异型域-亮氨酸拉链蛋白质的比对。

[0031] 图 3:BN47310186、BN51359456、HV62552639、ZM61995511、LU61567101、LU61893412、OS39781852、OS34701560、OS36821256、GM51659494、GM49780101、GM59637305 和 TA55974113 与已知的含 A20 和 AN1 结构域的锌指蛋白质的比对。

[0032] 优选实施方式详述

[0033] 在本申请中,提到了多篇公开文献。所有这些公开文献及这些公开文献中提到的那些参考文献都通过引用整体并入本申请,以更完整地解释本发明所述领域的状态。本文中使用的术语仅仅是为了描述特定实施方式的目的,而非意欲加以限制。在本文中使用,“一个/种”(“a”或“an”)可表示一个或多个/种,这取决于其所使用的上下文。因此,例如,提到“细胞”可表示可使用至少一个细胞。

[0034] 在一种实施方式中,本发明提供了过量表达表 1 中鉴定的经分离的多核苷酸或其同源物的转基因植物。本发明的转基因植物展示出较之植物的野生型品种而言对环境胁迫的增加的耐受性。任选地,较之植物的野生型品种而言,在正常或胁迫条件下,此类经分离的核酸在植物中的过量表达可导致植物生长的增加或相关农业产品的产量增加。并没希望被任何理论所束缚,本发明的转基因植株对环境胁迫的耐受力的增加,生长的增加和/或者产量的增加相信是由植株水使用效率(water use efficiency)的增加引起的。

[0035] 如本文中的定义,“转基因植物”是这样的植物,已使用重组 DNA 技术对其进行了改造,使得其含有经分离的、本来不存在于植物中的核酸。在本文中使用,术语“植物”包括整株植物、植物细胞和植物部分。植物部分包括但不限于,茎、根、胚珠、雄蕊、叶、胚、分生组织区域、愈伤组织、配子体、孢子体、花粉、小孢子等。本发明的转基因植物可以是雄性不育或雄性能育的,并且还可包括并不是包含本文所述的经分离多核苷酸的转基因的转基因。

[0036] 在本文中使用,术语“品种”指物种内共享恒定特征的植物组,所述特征将它们与该物种内的典型形式和其它可能品种分开。虽然具有至少一种区别性的性状,但是品种也通过品种内的各个体间的一些变异来表征,这主要基于以后世代的后代间孟德尔性状分离。如果对于某特定性状而言遗传上是纯合的——到这样的程度:纯育的品种被自花授粉,在后代中观察不到显著量的性状的自由分离,品种就被认为是特定性状的“纯育”。在本发明中,性状是从引入植物品种的一种或多种经分离的多核苷酸的转基因表达产生的。如还在本文中所使用的,术语“野生型品种”指这样作为对照植物针对比较目的对其加以分析的一组植物,其中,野生型品种植物与转基因植物(经根据本发明的经分离的多核苷酸转化的植物)相同,但是,野生型品种植物没有经本发明的经分离的多核苷酸转化。

[0037] 如本文中的定义,术语“核酸”和“多核苷酸”可相互替换,它们表示直链或带支链的、单链或双链的 RNA 或 DNA 或其杂交体。该术语还包括 RNA/DNA 杂交体。“经分离的”核

酸分子是这样的：其基本上与所述核酸天然来源中存在的其它核酸分子（即，编码其它多肽的序列）分开。例如，经克隆的核酸被认为是经分离的。如果核酸已被人为干涉所改变，或者被放在并非其天然位点的基因座或位置上，或者如果其通过转化被引入细胞，核酸就被认为是经分离的。此外，经分离的核酸分子，例如 cDNA 分子，当通过重组技术生产时可不合其天然相关的一些其它细胞材料，当通过化学合成时，不含化学前体或其它化学品。虽然其任选包括位于基因编码区域 3' 和 5' 末端的非翻译序列，但是其优选去除了其天然存在的复制子中编码区域天然侧翼的序列。

[0038] 在本文中使用时，术语“环境胁迫”指与盐度、干旱、氮、温度、金属、化学品、病原性或氧化胁迫或其任何组合相关的次最优条件。术语“水分利用效率”和“WUE”指植物产生的有机质的量除以植物生产有机质所使用的水的量，即，植物的干重相对于植物的水分利用。在本文中使用时，术语“干重”指植物中水之外的所有物质，其包括，例如，碳水化合物、蛋白、油和矿物质营养物。

[0039] 任何植物物种可被转化，以产生根据本发明的转基因植物。本发明的转基因植物可以是双子叶植物或单子叶植物。例如，并且并非限制性地，本发明的转基因植物可源于下述双子叶植物科中的任何：豆科 (Leguminosae)，包括豌豆、苜蓿和大豆等植物；伞形科 (Umbelliferae)，包括胡萝卜和芹菜等植物；茄科 (Solanaceae)，包括番茄、马铃薯、茄子、烟草和胡椒等植物；十字花科 (Cruciferae)，特别是芸苔属 (Brassica)，其包括油菜 (oilseed rape)、甜菜、甘蓝、花椰菜和椰菜等植物；以及拟南芥 (*A. thaliana*)；菊科 (Compositae)，其包括莴苣等植物；锦葵科 (Malvaceae)，其包括棉花；豆科 (Fabaceae)，其包括花生等植物，等等。本发明的转基因植物可源于单子叶植物，例如，小麦、大麦、高粱、粟、黑麦、黑小麦、玉米、稻、燕麦和甘蔗。本发明的转基因植物还包括树，例如苹果、梨、椴树、李、樱桃、桃、油桃、杏、番木瓜、芒果和其它木本物种，包括松柏类和落叶类的树，例如杨树、松、红杉、雪松、橡树等。尤其优选的是拟南芥、烟草 (*Nicotiana tabacum*)、油菜、大豆、玉米（玉蜀黍）、小麦、亚麻、马铃薯和万寿菊。

[0040] 如表 1 所示，本发明的一种实施方式是经包含经分离的下述多核苷酸的表达盒转化的转基因植物，所述多核苷酸编码蛋氨酸亚砷还原酶家族蛋白质。蛋氨酸亚砷还原酶 (MSR) 催化硫氧还蛋白质依赖的蛋氨酸亚砷 (MetSO) 还原为正确的蛋氨酸残基。蛋氨酸很容易受到氧化损伤，蛋氨酸氧化导致许多蛋白质活动和构象上的修饰。

[0041] 有两种 MSR，A 型及 B 型；但是，这两种类型在序列和结构上均不相关。MSRB 酶选择性地催化还原蛋氨酸亚砷 R 对映体。MSRB 型蛋白质在两个 CxxC 基序中含有四个保守的半胱氨酸残基，其中 x 可以是任何氨基酸。这些 CxxC 基序有可能与锌固定有关。

[0042] 已描述有来自几种植物物种的例子，其中环境胁迫性的条件导致了增加的活性氧族 (ROS) 的水平，且由此产生的氧化损伤导致了 MSR 基因表达的修饰。MSR 本身是直接抗氧化剂的优良候选，因为蛋氨酸残基循环的氧化和还原可以充当消除细胞中活性氧族的有效途径。在真核生物中，衰老和多种疾病都是由蛋氨酸氧化造成的蛋白质结构与功能的破坏所引发的。MSR 蛋白质的底物很多还是未知之数。迄今为止，第一个被鉴定的植株 MSR 底物是小的质体 (plastidic) 热休克蛋白质 Hsp21。Hsp21 包含保守的 N 端区域，其高度富含蛋氨酸残基。此蛋氨酸区域必须保持还原形式，以保持 Hsp21 的伴侣样活性。

[0043] 此实施方式中的转基因植株可能包括任何编码蛋氨酸亚砷还原酶家族蛋白质的

多聚核苷酸。此实施方式中的转基因植株优选包括编码 Se1R 结构域的多聚核苷酸,该结构域含有的序列包含 SEQ ID NO:2 中的氨基酸 77 到 199;SEQ ID NO:4 中的氨基酸 79 到 200;SEQ ID NO:6 中的氨基酸 91 到 213;SEQ ID NO:8 中的氨基酸 79 到 200;SEQ ID NO:10 中的氨基酸 80 到 201。此实施方式中的转基因植株更优选包括编码蛋氨酸亚砷还原酶家族蛋白质的多聚核苷酸,所述蛋氨酸亚砷还原酶家族蛋白质的序列包含 SEQ ID NO:2 中的氨基酸 1 到 205;SEQ ID NO:4 中的氨基酸 1 到 204;SEQ ID NO:6 中的氨基酸 1 到 214;SEQ ID NO:8 中的氨基酸 1 到 202;SEQ ID NO:10 中的氨基酸 1 到 206。

[0044] 在另一个实施方式中,本发明提供了用含有编码同源结构域-亮氨酸拉链蛋白质的分离的多聚核苷酸的表达盒转化的转基因植株。同源结构域-亮氨酸拉链(HDZip)蛋白质属于转录因子家族,其作为二聚体通过亮氨酸拉链结构域相互作用,并通过其同源结构域以序列特异性的方式结合 DNA。基于序列,HDZip 家族蛋白质分为四类。人们认为第 I 类 HDZip 蛋白质调节植株对 ABA 的反应,并可能具有与 ABA 的信号传导相关的调控性作用。第 I 类 HDZip 蛋白质的成员可以在体外形成异源二聚体,因此,这一类可以构成蛋白质相互作用网络,其介导对环境刺激的反应和/或整合信号以调节相似的靶基因集合。

[0045] 此实施方式中的转基因植株可能包括任何编码同源结构域-亮氨酸拉链蛋白质的多聚核苷酸。此实施方式中的转基因植株优选包括编码同源域的多聚核苷酸,所述同源域的序列包含 SEQ ID NO:12 中的氨基酸 62 到 116;SEQ ID NO:14 中的氨基酸 83 到 147;SEQ ID NO:16 中的氨基酸 76 到 130;或者多聚核苷酸编码与亮氨酸拉链结构域相关的同源域,其序列包括 SEQ ID NO:12 中的氨基酸 117 到 161;SEQ ID NO:14 中的氨基酸 138 到 182;SEQ ID NO:16 中的氨基酸 131 到 175。此实施方式中的转基因植株更优选包括编码同源结构域-亮氨酸拉链蛋白的多聚核苷酸,所述同源结构域-亮氨酸拉链蛋白的序列包含 SEQ ID NO:12 中的氨基酸 1 到 310;SEQ ID NO:14 中的氨基酸 1 到 331;SEQ ID NO:16 中的氨基酸 1 到 340。

[0046] 在另一实施方式中,本发明提供了表达盒转化的转基因植株,该表达盒所含的分离的多聚核苷酸编码含有 A20 和 AN1 结构域的锌指蛋白质。含有 A20 和 AN1 结构域的锌指蛋白质存在于所有真核生物。这些蛋白质的特点是存在含有多个半胱氨酸 2/半胱氨酸 2 指基序的 A20 锌指结构域和 AN1 锌指结构域。AN1 结构域通常是在包含 A20 锌指结构域的蛋白质中。这些蛋白质在动物系统中的功能已经被很好地表征,但是对这些蛋白质在植株中的功能知之甚少。

[0047] 大米 OsiSAP1 蛋白质是第一个被鉴定的拥有 A20 和 AN1 锌指结构域的植物蛋白质。这种蛋白质被发现与多种胁迫性有关。该 OsiSAP1 基因在应对环境胁迫诸如冷、盐度、干旱、水淹、受伤和重金属时被诱导。也已知来自于豆类的直系同源被诱发剂(eliciter)处理和应对受伤时可被诱导。当在烟草中过表达时,OsiSAP1 产生非生物胁迫耐受性。OsiSAP1 没有典型的核定位信号,因此被认为是通过对锌指结构域来行使蛋白质的相互作用的功能的。

[0048] 此实施方式中的转基因植株包含任何编码内含 A20 域和 AN1 结构域的锌指蛋白质的多聚核苷酸。此实施方式中的转基因植株优选包括多聚核苷酸,其编码的 A20 样锌指所具有的序列包含 SEQ ID NO:18 中的氨基酸 15 到 39;SEQ ID NO:20 中的氨基酸 13 到 37;SEQ ID NO:22 中的氨基酸 15 到 39;SEQ ID NO:24 中的氨基酸 14 到 38;SEQ ID NO:26 中

的氨基酸 14 到 38 ;SEQ ID NO :28 中的氨基酸 40 到 64 ;SEQ ID NO :30 中的氨基酸 15 到 39 ;SEQ ID NO :32 中的氨基酸 19 到 43 ;SEQ ID NO :34 中的氨基酸 13 到 37 ;SEQ ID NO :36 中的氨基酸 19 到 43 ;SEQ ID NO :38 中的氨基酸 18 到 42 ;SEQ ID NO :40 中的氨基酸 15 到 39 ;SEQ ID NO :42 中的氨基酸 15 到 39 ;SEQ ID NO :44 中的氨基酸 19 到 43 ,且该多聚核苷酸编码 AN1 样锌指结构域,其序列包含 SEQ ID NO :18 中的氨基酸 118 到 158 ;SEQ ID NO :20 中的氨基酸 128 到 168 ;SEQ ID NO :22 中的氨基酸 95 到 135 ;SEQ ID NO :24 中的氨基酸 112 到 152 ;SEQ ID NO :26 中的氨基酸 115 到 155 ;SEQ ID NO :28 中的氨基酸 179 到 219 ;SEQ ID NO :30 中的氨基酸 110 到 150 ;SEQ ID NO :32 中的氨基酸 105 到 145 ;SEQ ID NO :34 中的氨基酸 105 到 145 ;SEQ ID NO :36 中的氨基酸 111 到 151 ;SEQ ID NO :38 中的氨基酸 102 到 142 ;SEQ ID NO :40 中的氨基酸 113 到 153 ;SEQ ID NO :42 中的氨基酸 115 到 155 ;SEQ ID NO :44 中的氨基酸 106 到 146。此实施方式中的转基因植株更优选包括多聚核苷酸,其编码的含有 A20 和 AN1 结构域的锌指蛋白质的序列包含 SEQ ID NO :18 中的氨基酸 1 到 177 ;SEQ ID NO :20 中的氨基酸 1 到 187 ;SEQ ID NO :22 中的氨基酸 1 到 154 ;SEQ ID NO :24 中的氨基酸 1 到 171 ;SEQ ID NO :26 中的氨基酸 1 到 174 ;SEQ ID NO :28 中的氨基酸 1 到 239 ;SEQ ID NO :30 中的氨基酸 1 到 169 ;SEQ ID NO :32 中的氨基酸 1 到 164 ;SEQ ID NO :34 中的氨基酸 1 到 164 ;SEQ ID NO :36 中的氨基酸 1 到 170 ;SEQ ID NO :38 中的氨基酸 1 到 161 ;SEQ ID NO :40 中的氨基酸 1 到 172 ;SEQ ID NO :42 中的氨基酸 1 到 174 ;SEQ ID NO :44 中的氨基酸 1 到 165。

[0049] 本发明还提供了通过表达表 1 所列出的多核苷酸的转基因植物产生的种子,其中,所述种子含有所述多核苷酸,并且其中,所述植物较之植物的野生型品种而言,在正常或胁迫条件下增加的生长和 / 或产量和 / 或对环境胁迫的增加的耐受性方面是纯育的。本发明还提供了通过或从表达所述多核苷酸的转基因植物、其植物部分或其种子生产的产品。可使用本领域公知的多种方法获得产品。在本文中使用时,词语“产品”包括但不限于食物、饲料、食物补充物、饲料补充物、化妆品或药物。食物被认为是用于营养或用于补充营养的组合物。动物饲料和动物饲料补充物特别地被认为是食物。本发明还提供了通过转基因植物、植物部分和植物种子中的任何生产的农业产品。农业产品包括但不限于:植物提取物、蛋白、氨基酸、碳水化合物、脂肪、油、聚合物、维生素等。

[0050] 在一种优选的实施方式中,本发明的经分离的多核苷酸包含下述多核苷酸,其具有选自表 1 列出的多核苷酸序列的序列。这些多核苷酸可包含编码区域的序列以及 5' 非翻译序列和 3' 非翻译序列。

[0051] 本发明的多聚核苷酸可以使用标准的分子生物学技术和这里提供的序列信息分离。用于聚合酶链反应扩增的合成的寡聚核苷酸引物可基于表 1 所示的核苷酸序列设计。本发明的核酸分子可用 cDNA 或者可选地,用基因组 DNA 为模板和适当的寡聚核苷酸引物,按标准 PCR 扩增技术扩增。这样扩增的核酸分子可以克隆到适当的载体中,用 DNA 序列分析表征。此外,与列于表 1 中的核苷酸序列相应的寡聚核苷酸可以用标准的合成技术制备,例如,使用自动 DNA 合成仪。

[0052] “同源物”在本文中被定义为分别具有相似或基本相同的核苷酸或氨基酸序列的两条核酸或多肽。同源物包括等位变体、类似物和直向同源物,它们如下文所定义。在本文中使用,术语“类似物”指具有相同或相似功能但是在不相关的生物中分别进化的两条核

酸。在本文中使用时,术语“直向同源物”指来自不同物种但是通过物种形成从共同的祖先基因进化来的两条核酸。术语同源物还包括由于遗传密码子简并性而与表 1 所示的核苷酸序列之一有所不同但却编码同样多肽的核酸分子。在本文中使用时,“天然存在的”核酸分子指具有自然界中存在的核苷酸序列的 RNA 或 DNA 分子(即编码天然多肽)。

[0053] 为测定两条氨基酸序列(例如,表 1 的多肽序列之一及其同源物)的百分比序列同一性,就最优比较目的而言,将序列对齐(例如,可在一条多肽的序列中引入缺口,以与另外的多肽或核酸最优对齐)。然后比较相应氨基酸位置上的氨基酸残基。当一条序列中的位置被另一序列中相应位置上的同样的氨基酸残基占据时,分子在该位置就是相同的。可在两条核酸序列之间进行同样类型的比较。

[0054] 优选地,本发明的多肽的经分离的氨基酸同源物、类似物和直向同源物与表 1 所示的整条氨基酸序列至少大约 50-60%,优选至少大约 60-70%,更优选至少大约 70-75%、75-80%、80-85%、85-90%或 90-95%,最优选至少大约 96%、97%、98%、99%相同或者更相同。在另一优选的实施方式中,本发明的经分离的核酸同源物包含下述核苷酸序列,所述核苷酸序列与表 1 所示的核苷酸序列至少大约 40-60%,优选至少大约 60-70%,更优选至少大约 70-75%、75-80%、80-85%、85-90%或 90-95%,进一步优选至少大约 95%、96%、97%、98%、99%相同或者更相同。

[0055] 可以例如使用 Vector NTI 9.0(PC) 软件包(Invitrogen,1600 FaradayAve., Carlsbad,CA92008)来测定两条核酸或多肽序列之间的百分比序列同一性。当使用 Vector NTI 时,为测定两条核酸的百分比同一性,可使用的缺口开放罚分为 15,缺口延伸罚分为 6.66。为测定两条多肽的百分比同一性,可使用的缺口开放罚分为 10,缺口延伸罚分为 0.1。所有其它 VectorNTI 参数都可按照默认设置来设置。就使用 Vector NTI 的多重比对的目的而言(Clustal W 算法),缺口开放罚分为 10,缺口延伸罚分为 0.05,使用 blosum 62 矩阵。可选地,可以使用 Align 2.0(Myers 和 Miller(1989)CABIOS 4,11-17),所有参数设置为默认设置。应当理解,就测定序列同一性的目的而言,当比较 DNA 序列和 RNA 序列时,胸腺嘧啶核苷酸等同于尿嘧啶核苷酸。

[0056] 对应于表 1 所列出的多肽的同源物、类似物和直向同源物的核酸分子可基于其与所述多肽的同一性来分离,其中使用编码各多肽的多核苷酸或基于其的引物作为杂交探针,根据标准杂交技术在严紧杂交条件下来进行。本文中关于 DNA 到 DNA 的印迹杂交而言,术语“严紧条件”指在 10xDenhart' s 溶液、6X SSC、0.5% SDS 和 100 μ g/ml 变性的鲑鱼精 DNA 中于 60℃进行杂交。在 62℃,3X SSC/0.1% SDS 接着 1X SSC/0.1% SDS 以及最后 0.1X SSC/0.1% SDS 中对印迹依序洗 30 分钟(每次)。此外,在本文中使用时,在一种优选的实施方式中,术语“严紧条件”指在 6X SSC 溶液中于 65℃进行杂交。在另一实施方式中,“高度严紧条件”指在 10XDenhart' s 溶液、6X SSC、0.5% SDS 和 100 μ g/ml 变性的鲑鱼精 DNA 中于 65℃杂交过夜。在 65℃,3X SSC/0.1% SDS 接着 1X SSC/0.1% SDS 以及最后 0.1X SSC/0.1% SDS 中对印迹依序洗 30 分钟(每次)。核酸杂交的方法在文献 Meinkoth 和 Wahl,1984,Anal. Biochem. 138 :267-284 中有描述;是本领域公知的(见,比如,Current Protocols in Molecular Biology, Chapter 2, Ausubel 等,编, Greene Publishing 和 Wiley-Interscience, New York,1995; 和 Tijssen,1993, Laboratory Techniques in Biochemistry and Molecular Biology:Hybridization with Nucleic Acid Probes,第 I

部分,第2章,Elsevier,New York,1993)。优选地,在严紧或高度严紧条件下与表1列出的核苷酸序列杂交的本发明的经分离的核酸分子对应于天然存在的核酸分子。

[0057] 有多种方法可用于从简并寡核苷酸序列生产可能的同源物的文库。可在自动DNA合成仪中进行简并基因序列的化学合成,然后将合成的基因连接进合适的表达载体。使用简并基因组允许在一种混合物中提供编码想要的可能序列的组的所有序列。用于合成简并寡核苷酸的方法是本领域已知的(见,例如,Narang,1983, Tetrahedron 39 :3 ;Itakura等,1984, Annu. Rev. Biochem. 53 :323 ;Itakura等,1984, Science 198 :1056 ;Ike等,1983, Nucleic Acid Res. 11 :477)。

[0058] 此外,可制造经优化的核酸。优选地,经优化的核酸编码下述多肽,所述多肽具有与表1所列出的那些多肽相似的功能和/或当其在植物中过量表达时能在正常和/或水受限条件下调节植物生长和/或产量和/或对环境胁迫的耐受性,以及更优选地,在正常和/或水受限条件下增加植物生长和/或产量和/或对环境胁迫的耐受性。在本文中使用,“经优化的”指经遗传改造以增加其在给定的植物或动物中的表达的核酸。为提供经针对植物进行优化的核酸,可将基因的DNA序列修饰为:1) 包含高度表达的植物基因偏好的密码子;2) 包含植物中基本发现的核苷酸碱基组成中的A+T含量;3) 形成植物起始序列;4) 以消除导致RNA不稳定、不适当聚腺苷酸化、降解和终止的序列,或者消除形成二级结构发卡或RNA剪接位点的序列;或5) 消除反义可读框。可通过利用一般植物或特定植物中的密码子选择分布频率,来实现核酸在植物中的增加表达。用于优化植物中核酸表达的方法可在EPA 0359472、EPA 0385962、PCT申请No. WO 91/16432、美国专利No. 5,380,831、美国专利No. 5,436,391、Perlack等人,1991, Proc. Natl. Acad. Sci. USA 88 :3324-3328和Murray等人,1989, Nucleic Acids Res. 17 :477-498中找到。

[0059] 可对本发明的经分离的多核苷酸加以优化,使得其密码子选择分布频率与高度表达的植物基因偏差优选不超过25%,更优选不超过大约10%。此外,要考虑到简并第三碱基的百分比G+C含量(单子叶植物看起来在该位置偏好G+C,而双子叶植物则不)。还要认识到,XCG(其中X是A、T、C或G)核苷酸在双子叶植物中是最不优选的密码子,而XTA密码子在单子叶和双子叶植物中都应当避免。本发明的经优化的核酸优选还具有与所选用的宿主植物十分接近CG和TA双联体避免指数。更优选地,这些指数与宿主的偏差不超过大约10-15%。

[0060] 本发明还提供了经分离的重组表达载体,其包含上文所述的多核苷酸,其中,所述载体在宿主细胞中的表达导致:较之宿主细胞的野生型品种而言,植物在正常或水受限条件下增加生长和/或产量,和/或增加对环境胁迫的耐受性。本发明的重组表达载体包含以适于在宿主细胞中表达核酸的形式存在的本发明的核酸,这意味着,重组表达载体包括一条或多条调控序列,其是基于待被使用用于表达的宿主细胞选择的,其与待表达的核酸序列有效地相连。在本文中关于重组表达载体使用时,“有效地相连”意欲表示感兴趣的核苷酸序列与调控序列以允许核苷酸序列表达的方式相连(例如,当载体被引入宿主细胞中时,在细菌或植物宿主细胞中)。术语“调控序列”意欲包括启动子、增强子和其它表达控制元件(例如聚腺苷化信号)。此类调控序列是本领域公知的。调控序列包括指导核苷酸序列在多种类型的宿主细胞中组成型表达的那些以及仅在某些宿主细胞或某些条件下指导核苷酸序列表达的那些。本领域技术人员理解,对表达载体的设计可取决于下述因素,例

如,对待转化的宿主细胞的选择、想要的多肽的表达水平等。本发明的表达载体可引入宿主细胞,以由此生产本文所述的核酸编码的多肽。

[0061] 植物基因表达应当与合适的启动子有效地相连,所述启动子赋予适时地、细胞特异性或组织特异性方式的基因表达。可用于本发明的表达盒的启动子包括能在植物细胞中启动转录的任何启动子。此类启动子包括但不限于:可从植物、植物病毒和细菌(含有在植物中表达的基因,例如农杆菌属(*Agrobacterium*)和根瘤菌属(*Rhizobium*))获得的那些

[0062] 启动子可以是组成型、诱导型、发育阶段优先、细胞类型优先、组织优先或器官优先的。组成型启动子在大多数条件下都有活性。组成型启动子的例子包括 CaMV 19S 和 35S 启动子(Ode11 等,1985, *Nature* 313:810-812), sX CaMV 35S 启动子(Kay 等,1987, *Science* 236:1299-1302), Sep1 启动子,稻肌动蛋白启动子(McElroy 等,1990, *Plant Cell* 2:163-171),拟南芥肌动蛋白启动子,泛素(ubiquitin)启动子(Christensen 等,1989, *Plant Molec. Biol.* 18:675-689), pEmu(Last 等,1991, *Theor. Appl. Genet.* 81:581-588), 玄参花叶病毒 35S 启动子, Sma 启动子(Velten 等,1984, *EMBO J* 3:2723-2730), 超级启动子(美国专利 No. 5,955,646), GRP1-8 启动子,肉桂醇脱氢酶启动子(美国专利 No. 5,683,439),来自农杆菌属的 T-DNA 的启动子,例如,甘露碱合酶、胭脂碱合酶和章鱼碱合酶,二磷酸核酮糖羧化酶小亚基(ssu-RUBISCO)启动子等。

[0063] 诱导型启动子偏好于某些环境条件下具有活性,例如,存在或不存在营养物或代谢产物时、热或冷、光照、病原体攻击、厌氧条件等。例如,来自芸苔属(*Brassica*)的 hsp80 启动子是热激诱导的;PPDK 启动子由光诱导;来自烟草、拟南芥和玉蜀黍的 PR-1 启动子可通过病原体的感染诱导;Adh1 启动子由缺氧和冷胁迫诱导。还可通过诱导型启动子来协助植物基因表达(综述参见 Gatz,1997, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 48:89-108)。如果想要基因表达以时间特异性方式发生,可化学诱导的启动子是尤其合适的。此类启动子的例子是水杨酸诱导型启动子(PCT 申请 No. WO 95/19443)、四环素诱导型启动子(Gatz 等人,1992, *Plant J.* 2:397-404)和乙醇诱导型启动子(PCT 申请 No. WO 93/21334)。

[0064] 在本发明的一种优选的实施方式中,诱导型启动子是胁迫诱导型启动子。就本发明的目的而言,胁迫诱导型启动子偏好于在一种或多种下述胁迫下具有活性:与盐度、干旱、氮、温度、金属、化学品、病原体和氧化胁迫相关的次最优条件。胁迫诱导型启动子包括但不限于 Cor78(Chak 等人,2000, *Planta* 210:875-883;Hovath 等人,1993, *Plant Physiol.* 103:1047-1053)、Cor15a(Artus 等人,1996, *PNAS* 93(23):13404-09)、Rci2A(Medina 等人,2001, *Plant Physiol.* 125:1655-66;Nylander 等人,2001, *Plant Mol. Biol.* 45:341-52;Navarre 和 Goffeau,2000, *EMBO J.* 19:2515-24;Capel 等人,1997, *Plant Physiol.* 115:569-76)、Rd22(Xiong 等人,2001, *Plant Cell* 13:2063-83;Abe 等人,1997, *Plant Cell* 9:1859-68;Iwasaki 等人,1995, *Mol. Gen. Genet.* 247:391-8)、cDet6(Lang 和 Palve,1992, *Plant Mol. Biol.* 20:951-62)、ADH1(Hoeren 等人,1998, *Genetics* 149:479-90)、KAT1(Nakamura 等人,1995, *Plant Physiol.* 109:371-4)、KST1(Muller-Röber 等人,1995, *EMBO J.* 14:2409-16)、Rha1(Terryn 等人,1993, *Plant Cell* 5:1761-9;Terryn 等人,1992, *FEBS Lett.* 299(3):287-90)、ARSK1(Atkinson 等人,1997, GenBank 登录号 L22302 和 PCT 申请 No. WO 97/20057)、PtxA(Plesch 等人, GenBank 登录号 X67427)、SbHRGP3(Ahn 等

人,1996, Plant Cell 18 :1477-90)、GH3(Liu 等人,1994, Plant Cell 6 :645-57)、病原体诱导型 PRP1 基因启动子(Ward 等人,1993, Plant. Mol. Biol. 22 :361-366)、来自番茄的热诱导型 hsp80- 启动子(美国专利 No. 5187267)、来自马铃薯的冷诱导型 α -淀粉酶启动子(PCT 申请 No. WO 96/12814)或创伤诱导型 pinII 启动子(欧洲专利 No. 375091)。关于干旱、冷和盐诱导型的启动子的其它例子,例如 RD29A 启动子,参见 Yamaguchi-Shinozaki 等人,1993, Mol. Gen. Genet. 236 :331-340。

[0065] 发育阶段优先的启动子偏好在某些发育阶段表达。组织和器官优先的启动子包括偏好在某些组织或器官(例如叶、根、种子或木质部)中表达的那些。组织优先和器官优先启动子的例子包括但不限于:果实优先、胚珠优先、雄性组织优先、种子优先、珠被优先、块茎优先、柄(stalk)优先、果皮优先、叶优先、柱头优先、花粉优先、花药优先、花瓣优先、萼片优先、花梗优先、长角果优先、茎优先、根优先的启动子等。种子优先的启动子偏好在种子发育和/或萌发期间表达。例如,种子优先的启动子可以是胚优先、胚乳优先和种皮优先的(见 Thompson 等人,1989, BioEssays 10 :108)。种子优先的启动子的例子包括但不限于纤维素合酶(celA)、Cim1、 γ -玉米醇溶蛋白、球蛋白-1、玉米 19kD 玉米醇溶蛋白(cZ19B1)等。

[0066] 其它合适的组织优先或器官优先启动子包括来自油菜(rapeseed)的 napin 基因启动子(美国专利 No. 5,608,152)、来自蚕豆(Vicia faba)的 USP 启动子(Baeumlein 等人,1991, Mol. Gen. Genet. 225(3):459-67)、来自拟南芥属(Arabidopsis)的油质蛋白启动子(PCT 申请 No. W098/45461)、来自菜豆(Phaseolus vulgaris)的菜豆蛋白启动子(美国专利 No. 5,504,200)、来自芸苔属的 Bce4 启动子(PCT 申请 No. W091/13980)或豆球蛋白 B4 启动子(LeB4;Baeumlein 等人,1992, Plant Journal, 2(2):233-9)以及在单子叶植物(例如玉米、大麦、小麦、黑麦、稻等)中赋予种子特异性表达的启动子。值得注意的合适启动子是来自大麦的 Ipt2 或 Ipt1 基因启动子(PCR 申请 No. WO 95/15389 和 PCT 申请 No. WO 95/23230)或者 PCT 申请 No. WO 99/16890 中描述的那些启动子(来自大麦的大麦醇溶蛋白基因、稻谷蛋白基因、稻水稻素基因、稻谷醇溶蛋白基因、小麦的麦醇溶蛋白基因、小麦谷蛋白基因、燕麦谷蛋白基因、高粱 kasirin 基因和黑麦的裸麦醇溶蛋白基因的启动子)。

[0067] 可用于本发明的表达盒的其它启动子包括但不限于主要叶绿体 a/b 结合蛋白启动子,组蛋白启动子, Ap3 启动子, β -conglycin 启动子,油卡诺拉油菜蛋白启动子,大豆凝集素启动子,玉米 15kD 玉米醇溶蛋白启动子、22kD 玉米醇溶蛋白启动子、27kD 玉米醇溶蛋白启动子、 γ -玉米醇溶蛋白启动子、waxy、shrunk 1、shrunk 2 和 bronze 启动子和 Zm13 启动子(美国专利 No. 5,086,169)、玉米多聚半乳糖醛酸酶启动子(PG)(美国专利 Nos. 5,412,085 和 5,545,546)和 SGB6 启动子(美国专利 No. 5,470,359)以及合成的或其它天然启动子。

[0068] 控制异源基因在植物中表达的其它机动性可通过使用来自异源来源的 DNA 结合结构域和应答元件(即,来自非植物来源的 DNA 结合结构域)获得。此类异源 DNA 结合结构域的例子是 LexA DNA 结合结构域(Brent 和 Ptashne,1985, Cell 43 :729-736)。

[0069] 在本发明的一种优选的实施方式中,表 1 列出的多核苷酸在来自高等植物(例如种子植物,例如作物植物)的植物细胞中表达。多核苷酸可通过任何手段“引入”植物细胞,包括转染、转化或转导、电穿孔、颗粒轰击、农杆菌感染等。用于转化或转染植物

细胞的合适方法已被公开,例如,使用美国专利 Nos. 4, 945, 050、5, 036, 006、5, 100, 792、5, 302, 523、5, 464, 765、5, 120, 657、6, 084, 154 等所示的颗粒轰击。更优选地,可使用农杆菌转化来制造本发明的转基因玉米种子,如美国专利 Nos. 5, 591, 616、5, 731, 179、5, 981, 840、5, 990, 387、6, 162, 965、6, 420, 630, 美国专利公开号 2002/0104132 等所述。可使用例如欧洲专利 No. EP0424047、U. S. 专利 No. 5, 322, 783、欧洲专利 No. EP 0397 687、美国专利 No. 5, 376, 543 或美国专利 No. 5, 169, 770 所述的技术来进行对大豆的转化。小麦转化的一个特别例子可于 PCT 申请 No. WO 93/07256 中发现。可使用美国专利 Nos. 5, 004, 863、5, 159, 135、5, 846, 797 等中公开的方法来转化棉花。可使用美国专利 Nos. 4, 666, 844、5, 350, 688、6, 153, 813、6, 333, 449、6, 288, 312、6, 365, 807、6, 329, 571 等中公开的方法来转化稻。其它植物转化方法公开于,例如,美国专利 Nos. 5, 932, 782、6, 153, 811、6, 140, 553、5, 969, 213、6, 020, 539 等中。适于将转基因插入特定植物的任何植物转化方法都可根据本发明使用。

[0070] 根据本发明,引入的多核苷酸如果并入非染色体自主复制子或整合进植物染色体中,则其可稳定保持于植物细胞中。或者,引入的多核苷酸可存在于染色体外非复制载体上且可瞬时转化或具有瞬时活性。

[0071] 本发明的另一方面涉及经分离的多肽,其具有选自表 1 列出的多肽序列的序列。“经分离的”或“经纯化的”多肽不含(通过重组 DNA 技术生产时)一些细胞材料或(化学合成时)化学品前体或其它化学品。“基本不含细胞材料”这种说法包括这样的多肽制备物,其中,将多肽与天然或重组生产其的细胞中的一些细胞组分分开。在一种实施方式中,“基本不含细胞材料”这种说法包括这样的本发明的多肽制备物:其具有少于大约 30% (以干重)污染性多肽,更优选少于大约 20% 的污染性多肽,进一步更优选少于大约 10% 的污染性多肽,最优选少于大约 5% 的污染性多肽。

[0072] 对酶活性和动力学参数的测定在本领域中已很成熟。用于测定任何给定的经改造的酶的活性的实验必须针对野生型酶的特定活性加以定制,这是本领域技术人员能力范围内的。关于酶的综述是一般性的,关于测定很多酶活性的结构、动力学、原理、方法、应用和例子详细描述有很多,它们是本领域技术人员公知的。

[0073] 本发明还涉及生产包含表 1 列出的至少一种多核苷酸的转基因植物的方法,其中,所述多核苷酸在植物中的表达导致:较之植物的野生型品种而言,植物在正常或水受限条件下生长和/或产量增加,和/或对环境胁迫的耐受性增加,所述方法包括下述步骤:(a)向植物细胞中引入包含表 1 所列出的至少一种多核苷酸的表达载体,以及(b)从植物细胞产生表达所述多核苷酸的转基因植物,其中,多核苷酸在转基因植物中的表达导致:较之植物的野生型品种而言,植物在正常或水受限条件下生长和/或产量增加,和/或对环境胁迫的耐受性增加。植物可以是,但不限于,原生质体、配子生产细胞和再生为整株植物的细胞。在本文中使用时,“转基因”指含有表 1 列出的至少一种重组多核苷酸的任何植物、植物细胞、愈伤组织、植物组织或植物部分。在很多情况中,重组多核苷酸被稳定整合进染色体或稳定的染色体外元件,使得其传到后续世代。

[0074] 本发明还提供了在正常或水受限条件下增加植物生长和/或产量和/或增加植物对环境胁迫的耐受性的方法,所述方法包括下述步骤:在植物中增加表 1 所列出的至少一种多核苷酸的表达。可通过本领域技术人员已知的任何方法增加蛋白的表达。

[0075] 可通过在正常和 / 或较不合适的条件下培养经修饰的植物, 然后分析植物的生长特征和 / 或代谢, 来评估遗传修饰对植物生长和 / 或产量和 / 或胁迫耐受性的影响。此类分析技术是本领域技术人员公知的, 其包括干重、湿重、多肽合成、碳水化合物合成、脂类合成、蒸发蒸腾速率、一般植物和 / 或作物产量、开花、繁殖、结籽 (seed setting)、根生长、呼吸速率、光合速率、代谢组成等, 其中使用生物技术领域技术人员已知的方法来进行。

[0076] 通过下述实施例来进一步阐述本发明, 所述实施例不应被以任何方式理解为对本发明的范围施加限制。

[0077] 实施例 1

[0078] 全长 cDNA 的克隆

[0079] 小立碗藓 (*Physcomitrella patens*) EST65 蛋氨酸亚砷还原酶家族蛋白的全长 DNA 序列 (SEQ ID NO :43) 在 e 值为 e^{-10} 的水平上与卡诺拉油菜 (canola), 大豆, 水稻, 玉米, 亚麻子, 向日葵, 大麦和小麦 cDNA 专利数据库比对 (Altschul 等, 1997, *Nucleic Acids Res.* 25 :3389-3402)。对于假定的全长序列分析所有的重叠群击中, 且代表假定的全长重叠群的最长的克隆被完全测序。一个卡诺拉油菜来源的同系物, 两个水稻来源的同系物, 一个大豆来源的同系物和一个小麦来源地同系物被鉴定。这些序列与各自最近的已知的公共序列的氨基酸同一性和相似性的程度列于表 2 至 6 (Align 2.0)。

[0080] 表 2

[0081] BN51364980 (SEQ ID NO :2) 与已知的蛋氨酸亚砷还原酶的比较

[0082]

公共数据库登录号	物种	序列同源性 (%)
NP_564640	拟南芥	82.00%
AAM65202	拟南芥	66.70%
BAD35399	稻	56.10%
NP_001057620	稻	55.80%
ZP_01592095	泥菌 (<i>Geobacter lovleyi</i> SZ)	42.40%

[0083] 表 3

[0084] OS34096188 (SEQ ID NO :4) 与已知的蛋氨酸亚砷还原酶的比较

[0085]

公共数据库登录号	物种	序列同源性 (%)
ABE84787	蒺藜苜蓿 (<i>Medicago truncatula</i>)	72.60%
NP_567639	拟南芥	61.90%

公共数据库登录号	物种	序列同源性 (%)
AAM62876	拟南芥	61.00%
NP_567271	拟南芥	58.70%
EAY98001	稻	57.80%

[0086] 表 4

[0087] OS32583643 (SEQ ID NO :6) 与已知的蛋氨酸亚砷还原酶的比较

[0088]

公共数据库登录号	物种	序列同源性 (%)
NP_001057620	稻	99.10%
BAD35399	稻	74.80%
NP_564640	拟南芥	59.60%
AAM65202	拟南芥	53.70%
YP_846684	Syntrophobacter fumaroxidans MPOB	44.20%

[0089] 表 5

[0090] GM53626178 (SEQ ID NO :8) 与已知的蛋氨酸亚砷还原酶的比较

[0091]

公共数据库登录号	物种	序列同源性 (%)
ABE84787	蒺藜苜蓿	72.80%
NP_567639	拟南芥	60.90%
AAM62876	拟南芥	60.90%
EAY98001	稻	59.10%
AA072582	稻	57.60%

[0092] 表 6

[0093] TA56540264 (SEQ ID NO :10) 与已知的蛋氨酸亚砷还原酶的比较

[0094]

公共数据库登录号	物种	序列同源性 (%)
EAY98001	稻	72.90%
NP_001055501	稻	69.50%
ABE84787	蒺藜苜蓿	65.50%
NP_567639	拟南芥	60.60%
AAM62876	拟南芥	59.60%

[0095] 小立碗藓 EST12 同源结构域亮氨酸拉链蛋白质的全长 DNA 序列 (SEQ ID NO :44) 在 e 值为 e^{-10} 的水平上与卡诺拉油菜, 大豆, 水稻, 玉米, 亚麻子, 向日葵, 大麦和小麦 cDNA 专利数据库比对 (Altschul 等, 1997, Nucleic Acids Res. 25 :3389-3402)。对于假定的全长序列分析所有的重叠群击中, 且代表假定的全长重叠群的最长的克隆被完全测序。一个卡诺拉油菜来源的同系物, 一个大豆来源的同系物和一个小麦来源的同系物被鉴定。这些序列与各自最近的已知的公共序列的氨基酸同一性和相似性的程度列于表 7 至 9 (Align 2.0)。

[0096] 表 7

[0097] BN45206322 (SEQ ID NO :12) 与已知的同源结构域亮氨酸拉链蛋白质的比较

[0098]

公共数据库登录号	物种	序列同源性 (%)
AAR04932	欧洲油菜	95.20%
AAF73482	芜菁 (Brassica rapa)	91.30%
AAD41726	拟南芥	81.20%
NP_195716	拟南芥	69.90%
AAK96762	拟南芥	69.60%

[0099] 表 8

[0100] GM48923793 (SEQ ID NO :14) 与已知的同源结构域亮氨酸拉链蛋白质的比较

[0101]

公共数据库登录号	物种	序列同源性 (%)
AAX98670	大豆	56.00%
AAK84886	菜豆 (Phaseolus vulgaris)	51.60%

公共数据库登录号	物种	序列同源性 (%)
CAA64417	番茄 (<i>Solanum lycopersicum</i>)	47.40%
BAA05624	胡萝卜 (<i>Daucus carota</i>)	46.70%
AAF01765	大豆	44.00%

[0102] 表 9

[0103] TA55969932 (SEQ ID NO :16) 与已知的同源结构域亮氨酸拉链蛋白质的比较

[0104]

公共数据库登录号	物种	序列同源性 (%)
NP_001048008	稻	74.00%
EAY87390	稻	73.80%
NP_001061807	稻	49.90%
AAD37698	稻	49.60%
AAS83417	稻	46.30%

[0105] 小立碗藓 EST307 含 A20 和 AN1 结构域的锌指蛋白质的全长 DNA 序列 (SEQ ID NO : 45) 在 e 值为 e^{-10} 的水平上与卡诺拉油菜, 大豆, 水稻, 玉米, 亚麻子, 向日葵, 大麦和小麦 cDNA 专利数据库比对 (Altschul 等, 1997, *Nucleic Acids Res.* 25 :3389-3402)。对于假定的全长序列分析所有的重叠群击中, 且代表假定的全长重叠群的最长的克隆被完全测序。两个卡诺拉油菜来源的同系物, 一个大麦来源的同系物, 两个玉米来源的同系物, 两个亚麻子来源的同系物, 三个大豆来源的同系物, 三个水稻来源的同系物和一个小麦来源的同系物被鉴定。这些序列与各自最近的已知的公共序列的氨基酸同一性和相似性的程度列于表 10 至 22 (Align 2.0)。

[0106] 表 10

[0107] BN47310186 (SEQ ID NO :18) 与已知含 A20 和 AN1 结构域的锌指蛋白质的比较

[0108]

公共数据库登录号	物种	序列同源性 (%)
NP_564585	拟南芥	88.70%
AAN71995	拟南芥	88.10%
ABL67658	佛手 (<i>Citrus cv. Shiranuhi</i>)	59.40%

公共数据库登录号	物种	序列同源性 (%)
AAQ84334	稻	56.00%
AAD38146	杏 (<i>Prunus armeniaca</i>)	55.20%

[0109] 表 11

[0110] BN51359456 (SEQ ID NO :20) 与已知含 A20 和 AN1 结构域的锌指蛋白质的比较

[0111]

公共数据库登录号	物种	序列同源性 (%)
NP_190848	拟南芥	71.60%
AAK68811	拟南芥	71.10%
NP_565844	拟南芥	66.00%
ABE93196	蒺藜苜蓿	51.10%
AAN71995	拟南芥	47.10%

[0112]

[0113] 表 12

[0114] HV62552639 (SEQ ID NO :22) 与已知含 A20 和 AN1 结构域的锌指蛋白质的比较

[0115]

公共数据库登录号	物种	序列同源性 (%)
NP_001055132	稻	70.10%
AAR96005	小果野芭蕉 (<i>Musa acuminata</i>)	51.80%
AAA33773	菜豆	42.60%
EAZ09556	稻	40.80%
EAZ45178	稻	39.40%

[0116] 表 13

[0117] ZM61995511 (SEQ ID NO :24) 与已知含 A20 和 AN1 结构域的锌指蛋白质的比较

[0118]

公共数据库登录号	物种	序列同源性 (%)
AAQ84334	稻	79.50%

公共数据库登录号	物种	序列同源性 (%)
AAX14637	玉米	77.00%
EAZ01657	稻	71.70%
ABL67658	佛手	69.50%
NP_001046186	稻	65.90%

[0119] 表 14

[0120] LU61567101 (SEQ ID NO :26) 与已知含 A20 和 AN1 结构域的锌指蛋白质的比较

[0121]

公共数据库登录号	物种	序列同源性 (%)
CAE73100	线虫 (Caenorhabditis briggsae)	29.30%
NP_190848	拟南芥	29.20%
XP_001357850	果蝇 (Drosophila pseudoobscura)	29.20%
EAY92150	稻	29.10%
ABL97956	芜菁	28.90%

[0122]

[0123] 表 15

[0124] LU61893412 (SEQ ID NO :28) 与已知含 A20 和 AN1 结构域的锌指蛋白质的比较

[0125]

公共数据库登录号	物种	序列同源性 (%)
ABL67658	佛手	67.40%
AAD38146	杏	64.60%
AAQ84334	稻	61.70%
ABN08135	蒺藜苜蓿	61.50%
AAN71995	拟南芥	61.00%

[0126] 表 16

[0127] OS39781852 (SEQ ID NO :30) 与已知含 A20 和 AN1 结构域的锌指蛋白质的比较

[0128]

公共数据库登录号	物种	序列同源性 (%)
EAZ45178	稻	99.40%
EAZ09556	稻	99.40%
NP_001063521	稻	65.90%
ABI23728	案头菊 (Chrysanthemum x morifolium)	58.80%
AAA33773	菜豆	47.30%

[0129] 表 17

[0130] OS34701560 (SEQ ID NO :32) 与已知含 A20 和 AN1 结构域的锌指蛋白质的比较

[0131]

公共数据库登录号	物种	序列同源性 (%)
NP_565844	拟南芥	58.80%
ABE93196	蒺藜苜蓿	56.40%
NP_190848	拟南芥	55.60%
AAK68811	拟南芥	55.60%
ABL67658	佛手	48.30%

[0132] 表 18

[0133] OS36821256 (SEQ ID NO :34) 与已知含 A20 和 AN1 结构域的锌指蛋白质的比较

[0134]

公共数据库登录号	物种	序列同源性 (%)
EAZ45178	稻	63.00%
EAZ09556	稻	63.00%
ABI23728	案头菊	51.10%
AAA33773	菜豆	43.40%
AAX14637	玉米	43.30%

[0135] 表 19

[0136] GM51659494 (SEQ ID NO :36) 与已知含 A20 和 AN1 结构域的锌指蛋白质的比较
[0137]

公共数据库登录号	物种	序列同源性 (%)
AAA33773	菜豆	58.00%
EAZ09556	稻	54.20%
EAZ45178	稻	54.20%
NP_566429	拟南芥	51.50%
ABI23728	案头菊	44.40%

[0138]

[0139] 表 20

[0140] GM49780101 (SEQ ID NO :38) 与已知含 A20 和 AN1 结构域的锌指蛋白质的比较
[0141]

公共数据库登录号	物种	序列同源性 (%)
ABN08135	蒺藜苜蓿	76.70%
ABL67658	佛手	66.30%
AAD38146	杏	64.60%
AAQ84334	稻	64.40%
AAX14637	玉米	61.90%

[0142] 表 21

[0143] GM59637305 (SEQ ID NO :40) 与已知含 A20 和 AN1 结构域的锌指蛋白质的比较
[0144]

公共数据库登录号	物种	序列同源性 (%)
AAD38146	杏	69.30%
ABL67658	佛手	68.60%
AAQ84334	稻	65.90%
AAX14637	玉米	64.00%
NP_564585	拟南芥	63.70%

[0145] 表 22

[0146] TA55974113 (SEQ ID NO :42) 与已知含 A20 和 AN1 结构域的锌指蛋白质的比较

[0147]

公共数据库登录号	物种	序列同源性 (%)
EAZ09556	稻	73.20%
EAZ45178	稻	72.60%
ABI23728	案头菊	54.20%
NP_001063521	稻	49.00%
AAA33773	菜豆	46.10%

[0148]

[0149] 实施例 2

[0150] 水胁迫耐受的拟南芥植株

[0151] 表 1 中的多聚核苷酸连接到含有选择性标记的二元载体中。这样形成的重组载体包含在组成性的启动子控制下的相应的有义方向的基因。该重组载体根据标准的条件转化到根癌农杆菌系中。拟南芥生态型 Col-0 或者 C24 根据标准条件种植和转化。T1 和 T2 植株筛选具有选择性的标记基因赋予的对选择性的试剂的抗性的。T3 种子用在温室或者生长室的实验中。

[0152] 在种植前大约 3-5 天,将种子冷藏用以分层。然后将种子种植,施肥,且用透明的圆屋顶保持湿度。对于“生物量”的测定,植株在 22℃种植在温室中,光周期为 16 小时光照 8 小时黑暗,一周浇水两次。对于“周期性干旱”的测定,植株种植在温度为 22℃相对湿度为 55%的生长室中,光周期设定为 16 小时光照 8 小时黑暗,在播种后 0 天、18 天和 25 天浇水。

[0153] 在第 19 和 22 天,每株植株的植株面积,叶子面积,生物量,颜色分布,颜色强度和生长速率都可以用市售的成像系统来测量。生物量的计算是最后一个测量时间点的植株叶子的总面积。生长速率的计算是最后一个测量的时间点的植株叶子的面积减去第一个测量的时间点的植株叶子的面积后再除以第一个测量的时间点的植株叶子的面积。健康指数的计算是深绿色叶子的面积除以植株叶子的总面积。表 23 呈现了过表达了 SEQ ID NO : 5, 7 和 13 的多聚核苷酸的转基因植株的独立转化事件 (株系) 的生物量、生长速率和健康指数。计算一个系相对于混合的野生型对照的变化百分比,用 t 检验来计算显著性的值 (p 值)。如果这个百分比变化大于 0 且 p 值小于 0.1,这个系被认为是阳性的,如果 p 值大于 0.1,不管百分比的变化是多少,则被认为没有显著性 (NS),如果百分比变化小于 0 且 p 值小于 0.1,则被认为是阴性的。

[0154] 表 23

[0155]

基因编码	结果分类	每个测量中每个结果分类中事件的数目		
		生物量	生长速率	健康指数
OS32583643 SEQ ID NO: 5	阳性	1	2	2
	没有显著性	8	5	6
	阴性	1	3	2
GM53626178 SEQ ID NO: 7	阳性	6	2	1
	没有显著性	3	6	7
	阴性	1	2	2
GM48923793 SEQ ID NO: 13	阳性	3	1	4
	没有显著性	2	2	4
	阴性	5	7	2

[0156] 实施例 3: 氮胁迫耐受性的拟南芥属植物

[0157] 将表 1 的多核苷酸连接进含有可选择标记的双元载体。得到的重组载体含有处于组成型启动子下的正义向的相应基因。按照标准条件将重组载体转化进根癌农杆菌菌株。按照标准条件培养和转化拟南芥生态型 Col-0 或 C24。针对对可选择标记基因赋予的对选择剂的抗性, 来筛选 T1 和 T2 植物。

[0158] 使用不含有机组分的基底 (substrate), 将植物生长于平台 (flat) 上。每块平台用水弄湿, 然后将对选择剂有抗性的幼苗移种到基底上。在设置到 22°C、相对湿度为 55%、光照周期设置为 16 小时光照 / 8 小时黑暗的生长室中培养植物。在第 12、15、22 和 29 天向供水中加入受控的低或高氮营养溶液。在第 18、25 和 32 天给予没有营养溶液的供水。使用商业可获得的成像系统, 在第 26、30 和 33 天, 采集托盘 (tray) 上所有植物的图像。在每个成像时间点, 测量针对每株植物的生物量和植物表型, 包括植物面积、叶面积、生物量、颜色分布、颜色强度和生长速率。

[0159] 实施例 4

[0160] 耐逆性的卡诺拉油菜植株

[0161] 用 4 天龄卡诺拉油菜幼苗的子叶叶柄作为外植体, 用于组织培养, 并按照 EP1566443 进行转化。商业栽培种 Westar (Agriculture Canada) 是用于转化的标准品种, 但是也可使用其它品种。含有双元载体的根癌农杆菌 GV3101 : pMP90RK 被用于卡诺拉油菜转化。用于转化的标准双元载体是 pSUN (W002/00900), 但是还描述过用于植物转化的很多不同的双元载体系统 (例如, An, G., 在 Agrobacterium Protocols, Methods in Molecular Biology 卷 44, 第 47-62 页, Gartland KMA 和 MR Davey 编著, Humana Press, Totowa, New Jersey)。使用的植株基因表达盒包含一个选择性标记基因与调控编码多聚核苷酸的 cDNA 转录的植株启动子。可使用多种选择标记基因, 包括美国专利 Nos. 5, 767, 366 和 6, 225, 105 公开的突变的乙酰羟酸合酶 (AHAS) 基因。用合适的启动子调控性状基因, 以提供对基因转录的组成型、发育、组织或环境调控。

[0162] 将卡诺拉油菜种子在 70%乙醇中表面灭菌 2 分钟,在 55℃温自来水中孵育 15 分钟,然后在 1.5%的次氯酸钠中孵育 10 分钟,接着用灭菌的蒸馏水漂洗三次。然后将种子置于没有激素的 MS 培养基上,该培养基含有 Gamborg B5 维生素、3%蔗糖和 0.8% Oxoidagar。在 24℃低光照 ($< 50 \mu\text{Mol/m}^2\text{s}$, 16 小时光照) 下令种子发芽 4 天。从体外幼苗上切下连有子叶的子叶柄外植体,通过将叶柄外植体的切口端浸入细菌悬浮液来用农杆菌属接种。然后在 24℃, 16 小时光照下,在包括维生素的 MS 培养基上对外植体培养 3 天,培养基中含有 3.75mg/l 的 BAP、3%蔗糖、0.5g/l MES (pH 5.2)、0.5mg/l GA3、0.8% Oxoidagar。与农杆菌共培养 3 天之后,将叶柄外植体转到含有 3.75mg/l BAP、0.5mg/l GA3、0.5g/l MES (pH 5.2)、300mg/l 特美汀和选择剂的再生培养基上,直到枝条再生。一旦外植体开始发育枝条,就将其转移到枝条延长培养基 (A6, 含有完全强度的 MS 培养基,其包括维生素、2%蔗糖、0.5% Oxoidagar、100mg/l 肌醇、40mg/l 硫酸腺嘌呤、0.5g/l MES (pH 5.8)、0.0025mg/l BAP、0.1mg/l IBA、300mg/l 特美汀和选择剂) 上。

[0163] 使用 TaqMan 探针,通过 qPCR,分析来自初代转基因植物 (T0) 的体外和温室材料的样品,以验证 T-DNA 的存在,以及测定 T-DNA 整合的数量。

[0164] 通过自花授粉从初代转基因植物产生种子。将第二代植物培养于温室条件下,并自花授粉。使用 TaqMan 探针通过 qPCR 分析植物,以验证 T-DNA 的存在,以及测定 T-DNA 整合的数量。比较纯合转基因植物、杂合转基因植物和单性 (azygous) (失效转基因) 植物的胁迫耐受性,例如,在实施例 2 和 3 所述相似的检验中,还在温室和田间研究两者中比较它们的产量。

[0165] 实施例 5:筛选胁迫耐受性稻植物

[0166] 使用已知的方法来产生包含表 1 的多核苷酸的转基因稻植物。产生大约 15 至 20 个独立转化子 (T0)。将初代转化子从组织培养室转移到温室,用于培养和收获 T1 种子。针对转基因的存在 / 不存在以 3 : 1 分离的 T1 后代的 5 个事件被保留下来。对于这些事件中的每个而言,通过视觉标记筛选,选择含有转基因 (杂合和纯合子) 的 10 株 T1 幼苗和缺乏转基因 (失效合子 (nullizygotes)) 的 10 株 T1 幼苗。将选出的 T1 植物转到温室。每株植物获得独特的条形码标记,以将表型数据与相应植物明确联系起来。在下述环境设置下,将选出的 T1 植物培养于 10cm 直径盆中的土壤上:光周期 = 11.5 小时,日光强度 = 30,000lux 或更高,日间温度 = 28℃或更高,夜间温度 = 22℃,相对湿度 = 60-70%。转基因植物和相应的失效合子以随机位置并排种植。从播种阶段起,到成熟阶段为止,令植物数次通过数字成像箱。每个时间点,以至少 6 个不同角度采集每株植物的影像 (2048x1536 像素,1600 万色)。

[0167] 在第二次实验中,用 T2 植物验证用 T1 植物进行的第一次实验中获得的数据。具有正确表达模式的品系被选用来进行进一步的分析。通过监测标记表达,筛选来自 T1 中阳性植物 (杂合子和纯合子) 的种子批次。对于选择的每一事件而言,再保留杂合子的种子批次,用于 T2 评估。在每批种子批次内,在温室中培养同样数量的阳性和阴性植物以用于评估。

[0168] 针对其改善的生长和 / 或产量和 / 或胁迫耐受性,筛选转基因植物,例如,使用实施例 2 和 3 描述的相似检验,对产量而言,温室和田间研究两者中均进行。

[0169] 实施例 6:胁迫耐受性的大豆植物

[0170] 使用共有的共同未决国际申请号 WO 2005/121345 中描述的方法,将表 1 的多核苷酸转化进大豆,该文献的内容通过引用并入本文。

[0171] 然后针对其在水受限条件下改善的生长和 / 或干旱、盐和 / 或冷耐受性,筛选转基因植物,例如,使用实施例 2 和 3 描述的相似检验,并且在温室和田间两者研究中针对产量来筛选。

[0172] 实施例 7 :胁迫耐受性的小麦植物

[0173] 使用 Ishida 等人,1996, Nature Biotech. 14745-50 所述的方法将表 1 的多核苷酸转入小麦。将不成熟的胚与携带有“超级二元”载体的根癌农杆菌共培养,通过器官发生回收转基因植物。该流程提供了 2.5% 至 20% 的转化效率。然后针对其在水受限条件下改善的生长和 / 或产量和 / 或胁迫耐受性来筛选转基因植物,例如,使用实施例 2 和 3 描述的相似检验,对产量而言,温室和田间研究两者中均进行。

[0174] 实施例 8 :胁迫耐受性的玉米植物

[0175] 使用农杆菌属将表 1 的多核苷酸转化进玉米的不成熟胚。在吸涨 (imbibition) 后,将胚转到没有选择剂的培养基。7 至 10 天后,将胚转到含有选择剂的培养基上,培养 4 周 (2 周转移一次),以获得经转化的愈伤组织细胞。通过将抗性愈伤组织转到补充有选择剂的培养基上来启动植物再生并在 25-27°C 光照下培养二至三周。然后将再生的枝条转到具有含选择剂的培养基的生根盒上。将具有根的小植物转到温室中小盆中的盆栽混合物中,适应新环境之后将其移种到更大的盆中,并在温室中保持,直到成熟。

[0176] 对这些植物中的每株都独特标签、取样并针对转基因拷贝数加以分析。标记出转基因阳性和阴性植物,将相同大小的配对,用于一起移种到大盆中。这为转基因阳性和阴性植物提供了统一的并有竞争性的环境。对大盆浇水,达到一定的土壤田间持水量百分比,这取决于需要的水胁迫的严重性。通过隔天浇水来保持土壤水水平。在生长期测量植物生长和生理性状,例如高度、茎直径、叶卷曲、植物萎蔫、叶延伸速率、叶水状态、叶绿素含量和光合速率。生长期之后,收获植物的地上部分,对每株植物称鲜重和干重。然后对转基因阳性和阴性植物之间的干旱耐受性表型加以比较。

[0177] 对盆加盖,所述盖允许幼苗长起来但使得水的损失最小化。定期对每只盆加以称重,加水以保持最初的水含量。在实验末尾,测量每株植物的鲜重和干重,计算每株植物消耗的水,计算每株植物的 WUE。在实验期间测量植物生长和生理性状,例如,WUE、高度、茎直径、叶卷曲、植物萎蔫、叶延伸速率、叶水状态、叶绿素含量和光合速率。然后对转基因阳性和阴性植物之间的 WUE 表型加以比较。

[0178] 将这些盆保持于温室中有统一环境条件的区域,并最优培养。对这些植物中的每株都独特标签、取样并针对转基因拷贝数加以分析。令植物在这些条件下生长,直到它们达到预定的生长阶段。然后停止给水。随着胁迫强度的增加,测量植物生长和生理性状,例如,高度、茎直径、叶卷曲、植物萎蔫、叶延伸速率、叶水状态、叶绿素含量和光合速率。然后对转基因阳性和阴性植物之间的干燥耐受性表型加以比较。

[0179] 在小盆中种植针对转化事件而言分离的转基因玉米种子,以在循环性干旱检验中进行测试。将这些盆保持于温室中有统一环境条件的区域,并最优培养。对这些植物中的每株都独特标签、取样并针对转基因拷贝数加以分析。令植物在这些条件下生长,直到它们达到预定的生长阶段。然后以固定的时间间隔,对植物重复浇水至饱和。在实验期间重复

该水 / 干旱循环。在生长阶段期间测量植物生长和生理性状,例如,高度、茎直径、叶卷曲、叶延伸速率、叶水状态、叶绿素含量和光合速率。在实验末尾,收获植物,称地上部分的鲜重和干重。然后对转基因阳性和阴性植物之间的循环性干旱耐受性表型加以比较。

[0180] 为测试分离的转基因玉米在无雨条件下的干旱耐受性,使用单个位置或多个位置处的受管理干旱胁迫。通过滴水带 (dip tape) 和顶部灌溉来控制作物水可利用性,这在下述位置进行,该位置在平均 5 个月的季节内具有小于 10cm 的降水以及超过 5℃ 的最低温度,或者是这样的位置,其具有期待的被自动“防雨篷 (rain out shelter)”拦截的应季降雨,防雨篷在不需要的时候撤回以提供开放的田间条件。遵循区域中的标准农业实践,进行土壤准备、种植、施肥和害虫控制。在每小块上播种针对单个转基因插入事件的存在而分离的种子。在叶样品上使用 Taqman 转基因拷贝数检验,以将转基因与失效分离对照植物区分开。还针对与干旱耐受性、生长和产量相关的一系列表型,对以这种方式基因分型的植物加以计分。这些表型包括植物高度、每植物的谷粒重量、每植物的谷粒数、每植物的穗数、地上干重、叶对水蒸汽的传导率、叶 CO_2 摄取、叶的叶绿素含量、光合作用相关的叶绿素荧光参数、水分利用效率、叶的水势、叶相对水含量、茎液流速率、茎导水率、叶温度、叶反射率、叶的光吸光度、叶面积、到开花所需天数、开花 - 抽丝间隔、籽粒灌浆持续时间、渗透势、渗透调节、根尺寸、叶延伸速率、叶角度、叶卷曲和存活。所有测量都是使用厂商提供的标准方案,用于田间生理学的商业可获得的设备来进行的。各植物被用作作为每种事件的重复单位。

[0181] 为测试无雨条件下非分离转基因玉米的干旱耐受性,使用单个位置或多个位置处的受管理干旱胁迫。通过滴水带和顶部灌溉来控制作物水利用率,这在下述位置进行,该位置在平均 5 个月的季节内具有小于 10cm 的降水量以及超过 5℃ 的最低温度,或者是这样的位置,其具有期待的被自动“防雨篷”拦截的应季降水量,防雨篷在不需要的时候撤回以提供开放的田间条件。遵循区域中的标准农业实践,进行土壤准备、种植、施肥和害虫控制。设计试验方案,使得含有非分离转基因事件的小块土地与失效分离对照的邻接小块土地成对。失效分离是由于孟德尔分离不含转基因的转基因植物的后代 (或者源于后代的品系)。在试验周围分布有针对特定事件的额外重复成对的小块土地。在成对的小块土地中对与干旱耐受性、生长和产量相关的一系列表型加以计分,并在小块土地的水平上对其加以评估。当测量技术仅能应用于个体植物时,在小块土地内每次随机选择植物。这些表型包括植物高度、每植物的谷粒重量、每植物的谷粒数、每植物的穗数、地上干重、叶对水蒸汽的传导率、叶 CO_2 摄取、叶的叶绿素含量、光合作用相关的叶绿素荧光参数、水分利用效率、叶的水势、叶相对水含量、茎液流速率、茎导水率、叶温度、叶反射率、叶的光吸光度、叶面积、到开花所需天数、开花 - 抽丝间隔、籽粒灌浆持续时间、渗透势、渗透调节、根尺寸、叶延伸速率、叶角度、叶卷曲和存活。所有测量都是使用厂商提供的标准方案,用于田间生理学的商业可获得的设备来进行的。各小块土地被用作作为每种事件的重复单位。

[0182] 为针对干旱耐受性和产量对转基因玉米进行多位置测试,选择 5 至 20 个包括主要玉米种植区的位置。将其广泛分布,以提供一系列预计的作物水利用率 (基于平均温度、湿度、降水量和土壤类型)。不对作物水利用率加以标准农业实践外的改变。设计试验方案,使得含有非分离转基因事件的小块土地与失效分离对照的邻接小块土地成对。在成对的小块土地中对与干旱耐受性、生长和产量相关的一系列表型加以计分,并在小块土地

的水平上对其加以评估。当测量技术仅能应用于个体植物时,在小块土地内每次随机选择植物。这些表型包括植物高度、每植物的谷粒重量、每植物的谷粒数、每植物的穗数、地上干重、叶对水蒸汽的传导率、叶 CO₂ 摄取、叶的叶绿素含量、光合作用相关的叶绿素荧光参数、水分利用效率、叶的水势、叶相对水含量、茎液流速率、茎导水率、叶温度、叶反射率、叶的光吸光度、叶面积、到开花所需天数、开花 - 抽丝间隔、籽粒灌浆持续时间、渗透势、渗透调节、根尺寸、叶延伸速率、叶角度、叶卷曲和存活。所有测量都是使用厂商提供的标准方案,用于田间生理学的商业可获得的设备来进行的。各小块土地被用作为每种事件的重复单位。

[0183] 附录

[0184] 源于卡诺拉油菜的 BN51364980 的 cDNA 序列 (SEQ ID NO :1) :

[0185] atggcttcttcttagttgtttcaccattcagtcacgtttcgtctcagcgagaaacaaagctcgattcaat
ctccaaaccgagtcctctccgattcgcttgcgttctcttataaaaccagaaacttgaatctctctgttcttctt
cggtgttccatgggttcccttaactcttctcagaaatcagacaacgtccaagaagctgcaaagagtgacttttgctt
caataagtgaaggtgagtggaagaaacggctaaccaccagaacagtattacatcaccagacagaagggaacagagag
agcttttacttggtgagtattggaatacaaagaccccaggagtatacaaatgtatctgttgcgacacgccactgttt
gactcatcaacaaagcttgatagtgaaccgggtggccatcgattaccaacctattggaacaatgtgaagtcaa
agctggacctctctatcatcttcatgcctagacaagaagttatctgtgctgtttgtaacgcccatcttggtcatgt
cttcgatgacgggtccacgaccaaccggaaaacgatattgcctcaacagtgtctgctctgaaacttgagtcattggag
agaacaagagaatga

[0186] BN51364980 cDNA 翻译为下列氨基酸序列 (SEQ ID NO :2) :

[0187] massscftiqsrfsartkldsiskspslsgfacrsltkprnlslvllrcsmgsfnssqksdnvqeaak
sdfasisegekkrltpeqyitrqkgteraftgeywnktkpgvykiccdtpldfsstkldsgtgwpsyyqpignn
vksklldsiifmprqevicavcnahlghvfdgprptgkryclnsaalkleslertre

[0188] 源于水稻的 OS34096188 的 cDNA 序列 (SEQ ID NO :3) :

[0189] atgggcttcaatattctgagaaccacttccatctccactcctatctcttccctccaaatccaaacccat
tttctcaactcttcttctcgttcttctcttccaccattttccccccaaagtcggttactcccaccactcttttctgtt
tctgccaccccccttcttactctccatcccaagcttggttttctgtggtgggattgtggccatggccgcacctggc
tctctccgcaaattccgaggaagagtggcgcgcaattctctcccctgaacagtttcggatcctcaggcaaaagggc
accgagttccctggaacaggagagtatgacaagttctatgaagaggagtttacaactgtgctggttgtgggact
ccactctacaggtccataacaaaattcaattctggttgtggctggccagccttctatgaggggattcccgagacc
ataaatcgcaatccggatcctgatgggatgaggacagaaataacgtgtgctgcttgtgggggacatctaggtcac
gtcttttaaaggagaaggattttccaacaccactaacgaacgccatttgtgtcaatagcatttcgctgaaatttgcg
ccagccaattcttattcttaa

[0190] OS34096188 cDNA 翻译成下列氨基酸序列 (SEQ ID NO :4) :

[0191] mgfnilrttsistpiisskskipfstllrsspstifppksvtpttlfvsatpfftlhpklgfrggivam
aapgsrlrkeeewrailspeqfrilrqkgtefpgtgeydkfyeegvncagcgtplyrsitkfnsgcgwpafyegip
gainrnpdpdgmrtetcaacgghlghvfkgegfptptnerhcvnsislkfapansys

[0192] 源于水稻的 OS32583643 cDNA 序列 (SEQ ID NO :5) :

[0193] atggccatgcggcaatacgcggctgctaccgctgcctcctccagtttcagagcacgtccacgggcgcgc
ccctcctgcctcccagccgcccgccctggcgcccttgctgtggtgtggcggtggagccgtgctagctacaggcg

[0204] 源于卡诺拉油菜的 BN45206322 cDNA 序列 (SEQ ID NO:11) :

[0205] atgatgaagagattaagcagttcagattcagtggtggtctcatctctttatgtcccactacttccacagatcagccgaatccaagaagatgcgggagagaatttcagtcgatgctcgaaggttacgaggaggaagaagaagaagccataaccgaggaagaggacaaaccggttttagccgagaagaagagacggttaaaccattaaccaagttaaagccttggagaaaaatttcaggttagagaacaagcttgagcctgagaggaaagtgaagttagctcaagaacttggctctccaacctcgtcaagtagctgttttggtttcagaaccgccgtgcgcggtggaagacaaaacagcttgagaaagattacgggtgttctcaaaacgcaatacagattctctccgccataactttgattccctccgccgtgaaaatgaatctcttcttcaagagatcgttaaactaaaagctaagcttaacggagaagaagaaggagatgatgttgatgaagaagagaacaacttggcgacgatggagagtgatgtttccgtcaaggaagaagaagtttcgttgccggagcagatcacagagccgccgtcttctcctcccgagcttctagagcattccgacagtttcaattaccggagtttaccgacctccgcgaccttcttccgttaaaggccgcggttctctccgtcgcgcgctggatcgctcggacagtagcgattcgagcgccgtgttgaacgaggaaagtagctctaactttacggcggtcccggcacgggttcccgcggcagtttcttgagtttgtgaaaatggagcagacggaggatcacgacgactttctgagtgagagaagaagcgtgcgggtttttctccgatgaacagccaccgtctctgcactgggtattccaccgttgatcagtggaactga

[0206] BN45206322 cDNA 翻译成下列氨基酸序列 (SEQ ID NO:12) :

[0207] mmkrlsssdsvglislcpttstdqpnprrcgrefqsmlegyeeeeeeaiteergqtglakkrlnlnqvkaleknfelenkleperkvklaqelglqprqvavwfqnrrarwkqllekdygvlktqydslrhnfdsllrenesllqeigklkaklngeeegddvdeennlatmesdvsvkeeevslpeqiteppssppqllehsdsfnysftdlrdllplkaaassvaaagssdssdsavlneesssnvtaapatvppgsfllqfvkmeqtedhddflsgeeacgffsdeqppsllhwystvdqwn

[0208] 源于大豆的 GM48923793 cDNA 序列 (SEQ ID NO:13) :

[0209] atggcgggtagtggaagtgccttttccaacatcactagctttcttcgcaccaacaaccctcttctcaacctctcgattcttctctcttctcctctctgcaccttctctgctcttctcctcggttcgagatccatgatgagttttgattggagaaggaggaaggggtgtaacggctccttcttccgcgctttgacatggacgacaatggggatgagtgcattgacgagtactttcatcaaccgcgagaagaagcgacgtctctctgcgagccaggttcagtttctagagaagagcttcgaggagagaacaagcttgaaccgcgagagaaagaccaaactagccaaagaccttggtttgcagccacggcaagttgctatttggttccagaaccgtagagctcggtggaagaacaacagctggagaaggattacgagactctgcatgcaagttttgagagtctcaagtccaactatgactgtcttctcaaggagaaagacaagttaaaagctgaggtggcgagcctcactgagaaggtgcttgcaagagggaaacaagaggggcacatgaagcaggctgaaagtgaagtgaaagaacaaaaggattattgcatttgcaggaacaggaaccacccagaggcttttactgcaatcagtttcggagggagaaggatccaaagtctcttctgtcgtttgggggttgtaaacaggaagatatcagttcagcaaggagtgcatttttgattcagatagtcacatttaccgatggagttcactctgcgctgctagagcatggtgattcttcttatgtgtttgagcctgatcaatcagatatgtcaagatgaagaagataacctcagcaagagtctctacccttcgtaccttttccaaacttgaagaagatgtggattactccgaccacctgaaagtcttctgtaattttggatttccctgaggaagatcatgtcctttggacctgggcttactactaa

[0210] GM48923793 cDNA 翻译成下列氨基酸序列 (SEQ ID NO:14) :

[0211] magsgsafsnitsflrtqqpssqpldsslflsapssapflgsrsmmsfdgeggkgcngsfrrafdmddngdecmdyfhqpekrrlsasvqfleksfeeenkleperktklakdlglqprqvaiwfqnrrarwknkqlekdyetlhasfeslksnydcllkekdklkaevasltekvlargkqeghmkqaeseseetkgllhlqeqppqrrlllqsvsege

gskvssvvggckqedissarsdildsdsphytdgvhsallehgdssyvfepdqsdmsqdeednlkslypsylfpl
eedvdysdpesscnfgfpeedhvlwtwayy

[0212] 源于小麦的 TA55969932 cDNA 序列 (SEQ ID NO :15) :

[0213] atggagccccggcggtcatcttcaacacgtcgggctccggcaacggacagatgctcttcatggactgc
ggcgcgggcgggcatcgccggcgcgggccggcatgttccatcgaggggtgagaccggtcctcgcgggcatggaagaagg
gcgcgggcgtgaageggcccttcttcacctcgccggatgacatgctggaggaggagtactacgacgagcagctcccg
agaagaagcgggcgctcacgccggagcaggtccacctgctggagaggagcttcgaggaggagaacaagctggagccg
gagaggaagacggagctggcccgcaagctcgggctgcagccacggcaggtggcgtctgtgttccagaaccgccgcgc
ccggtggaagacaaagacgtggagcgcgacttcgaccgcctcaaggcgtccttcgacgccctccgcgccgaccacg
acgcgtcctccaggacaaccaccgggtccgggtcacaggtggtaacgttgaccgagaagatgcaagataaggaggcg
ccggaaggcagcttcggtgcagccgccgacgcctcggagccggagcagggcgcgggaggcgaaggcttccttggc
cgacgccgaggagcagggcgcgggcagcgaggcggttcgaggtggtgcagcagcagctgcacgtgaaggacgaggaga
ggctgagcccggggagcgggcgggagcgcggtgctggacgcgagggacgcgctgctcgggagcggtatcgccctcgcc
ggcgtggtggacagcagcgtggactcgactgcttccccggggggcgccggcgggcagcagtagccacgagtgctgtgt
gggccccgtggcgggcgccatccagtcggaggaggacgacggcgcgggcagcgacgagggtgcagctactaccccg
acgacgcccgctcttcttcgccgccgcgaagggcacggccaccatcgacggacgacgacgatcagcaggacgac
ggccagatcagctacttgatgtggaactag

[0214] TA55969932 cDNA 翻译成下列氨基酸序列 (SEQ ID NO :16) :

[0215] mepgrlifntsgsgngqmlfmdcgaggiagaagmfhrgrvpvlggmeegrvgkrpffftspddmleeeyy
deqlpekrrrltpeqvhlrsfreenkleperktelarklglqprqvavwfqnrrarwkktlerdfdrllkasfda
lradhdallqdnhrllrsqvvtltekmdkeapegsfgaaadasepeqaaaaeakasladaeeqaaaaeafenvqqqlh
vkdeerlspgsggsavldardallgsgcglagvvdssvdsycfpggaggdeyhecvgpvaggiqseeddgagsdeg
csyypddaavffaaaqghghrtddddqddgqisywmwn

[0216] 源于卡诺拉油菜的 BN47310186 cDNA 序列 (SEQ ID NO :17) :

[0217] atggaccacgacaaaacaggatgccaaagcccacctgaaggtcccaagctatgcatcaacaactgcggt
ttcttcggaagcgtgccacaatgaacatgtgttccaagtgtcacaaggctatcctgtttcaacaggaaacagggggc
taggtttgcatctgcagtgcttggtggtacatcatcatccagcaacatcttaaaggaaacctttgctgctaccgcgc
tggttgatgctgaaacaaatccgttgagccgggtggtgctctgtgtacagccatcttctgtccaagttgccgcagag
gtagtagctccagaagccgctgcagcaaaactaaaggaaggaccaagccgatgtgctacttgcaataaacgggttgg
tctgactggattcaaatgtcgctgtggtgacctcttctcggggacgcaccgttatgcagacatacacaactgtcct
tcaattaccatgccgctgcgcaagaagctatagctaaagcaaaccgggttgtaaggcagagaagcttgacaaaatc
tga

[0218] BN47310186 cDNA 翻译成下列氨基酸序列 (SEQ ID NO :18) :

[0219] mdhdktgcspppegpklcinncgffgsaatmnmcskchkailfqeqgarfasavsggtssssnilket
faatalvdaetksvepvavsvqpssvqvaaevvapeaaaaaklkegpsrcatenkrvgltgfkrcgdlfcgthryad
ihncsfnyhaaaqaeaiakanpvvkaekldki

[0220] 源于卡诺拉油菜的 BN51359456 cDNA 序列 (SEQ ID NO :19)

[0221] atggcggaagagcatcgatgccagacgccggaaggccaccgtctctgtgctaacaactgcggcttcctc
ggcagctccgccaccatgaatctatgctccaactgctacggcgatctctgcttaagcaacagcaagcttccatgaa

atccaccgtcgaatcctctctctccgccgtatctcctccgtcgtcagagatcggctctatgcaatccaccgttgaat
cctctctctccgacgtatctcctccatcaccggagaccatttccatctcctctccaatgatccagcctctcgttcga
aaccatcagctgaattggaggtaacggcgacgaagacgggtgactccgccgccggagcagcagcagaaacggccgaa
tcggtgcacgacgtgtaggaacgggtcggttgaccgggttcaagtgccggtgcgggacgactttttgcggggctc
acaggtaccggaggtccatggatgcaccttcgatttcaaatcggccggtcgcgaagagatcgccaaggcgaacca
ctcgtcaaagcggcgaagcttcagaagatttga

[0222] BN51359456 cDNA 翻译成下列氨基酸序列 (SEQ ID NO :20) :

[0223] maehrcqtpeghrlcanncgflgssatmnlcsncygdllkqqqasmkstvesslsavspssseigsm
qstvesslsdvspsspetisisppmiqplvrnpsaelevtatktvtpppeqqkrpnrttcrkrvglgtgfkrcrgt
tfcgahrypevhgctfdksagreeiakanplvkaaklqki

[0224] 源于大麦的 HV62552639 cDNA 序列 (SEQ ID NO :21) :

[0225] atggcccaggagagttgtgatctcaacaaggacgaggccgagatcctgaagccatcctcctccacacc
ttgcctccttcgccagccacaccaccaccaaccgctcaataaccagaaccacaacctccacactcaccacca
caaccaccggcagctcaattcttgtccaggccctgcgaggttggttcccatagagacttccaaaagaggaaacat
gctgatgcggtgtcaatggccattgtggttgagccattgtcgtctgtgctgttcgttaaccgttgcaacgtgtgc
cgcaagagagttggtttgaccgggttcggttgccggtgtgagaagctcttttgcgcgccaccggcattcagaa
agccacgactgctcatttgattataaaactgtgggtcgggaggagattgcccgggcaaaccctctgatcagggt
gccaagatcattagatatga

[0226] HV62552639 cDNA 翻译成下列氨基酸序列 (SEQ ID NO :22) :

[0227] maqescdlnkdeaeilkpssstpsppspatpppptaqi pepqpphsppqpaaqlsrpcevvpjets
kkrkhadavsmavveplssvlfnrcnvrkrvglgtgfrercklfcprhrhseshdcsfdyktvgreeiaranp
liraakiiri

[0228] 源于玉米的 ZM61995511 cDNA 序列 (SEQ ID NO :23) :

[0229] atggaacacaaggaggcgggtgccagcagccgaggggccaatcctatgcatcaataactgcggctt
cttcggcagtgctgcgacgatgaacatgtgctccaagtgcacaaaggagatgataacgaagcaggagcaggccag
ctggctgcctcccccatcgatagcattgtcaatggcgtgacggcggaaggacctgtaattgctgcatctgta
aatgtggcagttcctcaagttgagcagaagactattgttgcagccatgcttgtagctgaaaccagcagggt
gctgctgtaatcccaaggccaaggaaggcccagaccgggtgcgcggcctgcaggaagcgtgttgggctgacggga
tttagctgccgatgcgggaacatgtactgttcggtgcaccgctactccgacaaacatgactgtcagttcgactat
cggactgcagcaaggagcgcgat tgccaaggccaatcctgtggtgaggcgaggagaagctcgacaagatctga

[0230] ZM61995511 cDNA 翻译成下列氨基酸序列 (SEQ ID NO :24) :

[0231] mehkeagcqqpcgpilcinncgffgsaatmnmcskchkemi tkqeqaqlaaspidsivngdgkgpv
iaasvnavvpqveqktivvqpmivaetseaaavipkakegpdraacrkrgltgfscregnmycsvhrysdkhdcq
fdyrtaardaiakanpvvraekldki

[0232] 源于亚麻子的 LU61567101 cDNA 序列 (SEQ ID NO :25) :

[0233] atggetccttcaccttgctccacgggtgcacggccaattgccccgctgccactcttacggacacccc
atcttcgggaactcagatctcgccgctggcggcagcgatacgtccacgtcggtgtttggaaaagtaggatccgtcgt
gattcagtcgctgcgaagaatcacgcgttcggccaagcttgtggccccgttttccctcgagctcctcccccttcc
gccgatcaagttcgccccaaagatggcgaggggaaaggaccgctgaagccgatcgagaagcagccgtcgaagaag

cgctccgtttctgcttctctccccgacgagacgattgacgcgacggttcctccgtccaccaaacggttcggttcggtccg
ttccgtctgtgtcacggacgccgacgaggccaggttgaaggcgaaccgcgagttcttcgctccggtatccccgaaac
gtggcttcgatccgactgacatgaccttcggtaacgccgccgctgcggctaatagcgaggaggaagcgaagaag
tggtgcggcagttgcaagaagecgctggggctgttaggggtcaagtgcaggtgtacgaagtcttctgttggaagca
tcggtatcctgaggagcatggttgtagcttcgatcatgtggcggttcgggaggcggattatcgagaaacagaatcctg
ttctcgagaccgacaagctggtggacagaatctga

[0234] LU61567101 cDNA 翻译成下列氨基酸序列 (SEQ ID NO :26) :

[0235] mapspcvhgtancprchsyghpifgnsdlaaggsdtstsvfkgvgsvviqspaknhafgqacgpvfps
ssspfrrikfpgkdgegkgplkpiekqpskkrfcfspdetidatvppstkpfgsfrsvcvtdadearlkanreffa
pvsrkrghfdptdmtfgnaaaaaaanareeakkwcgsckkrvgllgfkcrctkffcgkhrypeehgctfdhvafgrrii
ekqnpvletdklvdri

[0236] 源于亚麻子的 LU61893412 cDNA 序列 (SEQ ID NO :27) :

[0237] atggaccatgacgaggcaggtgccaggtccttcgatcatcctattctgtgcgttaacaattgcggc
ttcttcggaagtgtgccaccatgaacatgtgctcaaagtgccacaaggatacgtatgctaaaccaagagcaatccaa
gcttgctgcttcacgagcgaagtatcctcaacggatcgtcgatgagcctcggaagggaactcgttattgctgcta
agaccaattcggtagaaccaagaccatctccgtccaaccatcttctgcttcaagtgtgaagagagtatcgaaatg
aagctgccaaaagaagggccagtaggtgcaacacttgcaacaaacgtgtcggtttgaccggattcaaatgtcggtg
cgagaacatgtttctgcgcaaaccatcgctactcggacaagcacaattgcccccttgattaccgcactgttgccgtg
aagctatctcaaaggccaatccttttggtgaaggcggagaagctcgacaaaatctga

[0238] LU61893412 cDNA 翻译成下列氨基酸序列 (SEQ ID NO :28)

[0239] mdhdeagcqapsdhpilcvnncgffgsaatmmmskchkdtmlnqeqsklaassaasilngssmslgr
elviaaktnsvepktisvqpssassaeesiemklpkempsrentcnkrvgltgfkrcenmfcanhrysdhncp
dyrtagreaiskanplvkaekldki

[0240] 源于水稻的 OS39781852 cDNA 序列 (SEQ ID NO :29) :

[0241] atggcgcagcgcgacaagaaggatcaggagccgacggagctcagggcgccggagatcacgctgtgcgcc
aacagctgcggattccccgggaacccggccacgcagaacctctgccagaactgcttcttgggcgccacggcgctccac
ctcgctgcggtcttctttgtcgctaccgggtgctcgacaagcagccgccgaggccggcgccgctggttgagcctc
aggctcctctcccaccgcctgtggaggagatggcctccgcgctcgcgacggcgccggcgccggtcgccaagacgtcg
gcggtgaaccgggtgctccaggtgccggaagcgtgtcggcctcaccgggttcgggtgccggtgcggccacctgttctg
cgcgagcacccggtactccgaccgccacgggtgcagctacgactacaagtcggcgggcgaggagcccatcgccaggg
acaaccgggtggtgcgcgcggccaagatcggttaggttctga

[0242] OS39781852 cDNA 翻译成下列氨基酸序列 (SEQ ID NO :30) :

[0243] maqrddkqdptelrapeitlcanscgfpnpatqnlcqnclaatstsspslsspvldkqpprpa
aplvepqaplpppveemasalatapavaktsavnrcsrerkrvgltgfrcrcghlfcgehrysdhrhgsydyksa
ardaiardnpvvraakivrf

[0244] 源于水稻的 OS34701560 cDNA 序列 (SEQ ID NO :31) :

[0245] atggccgaagaacaccgatgccaagctcccgaaggtcacagactctgctccaacaactgcggtttcttt
ggtagccccgccaccatgaatctctgttccaaatgctacagagacatccgtttgaaggaagaagaacaagccaaaac
caaateccacaategaaaccgctctttcaggatcttctccgccaccgtcaccgcaaccgctcgttgctcctccg

tggaatcccccttcggcgccggttgaatccctccctcaaccaccggtgctgatttcgccgatatagccgcaccggtt
caggcgaaccggtgcgggcgctgttaggaagcgcggtggggttgacagggttcaagtgcaggtgcggaacaacgttttg
tgggagccacaggtaccccagaaacacgcgtgtggcttcgatttcaaggcggtggggagagaggagatagcacggg
cgaatcccgatgataaaggcgagaagctacggaggatttaa

[0246] OS34701560 cDNA 翻译成下列氨基酸序列 (SEQ ID NO :32) :

[0247] maehrcqapeghrlcsnncgffgspatmnlcskcyrdirlkeeeqaktkstietalsgsssatvtat
avvassvespsapveslpqppvlispdiaapvqanrcgacrkrvgtlgtfkcrcgttfcgshrypekhaacgfdkav
greeiaranpvikgeklrri

[0248] 源于水稻的 OS36821256 cDNA 序列 (SEQ ID NO :33) :

[0249] atggcgagagggagaagaaggtggaggagccgacggagctgagggcgccgagatgacgctctgcgcc
aacagctgcgggttcccgggcaaccggcgaccaacaacctctgccagaactgcttcttggtgcctcggcgctcttc
ttcttcttcttcgcgcgtgcctcgccgtcgacgacgtcgttgccggtgtttccggtggtggagaagccgaggcagg
ccgtacagtcgtcggcgggcgggcggtggcgctggtggttgagcgggcgacggcggggcccgggtggagtcgtcgtcg
aaggcgtcgaggtcgtcgtcggtcaaccgatgccacagctgccggaggcggtgggacctgaccgggttccggtgccg
ctgcggcgagctctactgcggcgcgacccggtactccgaccgccacgactgcagcttcgactacaagtcggcgggcga
gggacgccatcgccagggagaaccccgctcgtcgcgcgcccaagatcggttaggttctaa

[0250] OS36821256 cDNA 翻译成下列氨基酸序列 (SEQ ID NO :34) :

[0251] maqrekkveptelrapemtlcansefgpnpatnnlcqncflaasassssssaaaspsttspvfpv
vekprqavqssaaaavalvverptagpvessskasrsvnrhscrrrvgtlgtfrcrcgelycgahrysdrhdcs
fdyksaardaiarenpvvraakivrf

[0252] 源于大豆的 GM51659494 cDNA 序列 (SEQ ID NO :35) :

[0253] atggctcagaaaaccgagaaagaagaaccgacttcaaagttccggaaacgattacgctttgcgtcaac
aactgcggcgctaccggaaacctgccacgaataacatgtgccagaagtgttactgcctctaccgccaccacttc
cggcgcgggaggtgccggaatagcttctccggcgaccagatccggcgctctccgcgcgtcctcagaagagatcttttc
ctgaagagccctcgcgggtggcggtatcctccttcttcggaccagacgacgccgtcggaggcgaagcgcggtggtcaac
cgctgctccggatgccggcggaaggtcggactcaccggattccggtgccggtgcggcgagctcttctgcgccgagca
ccggtactccgaccgccacgactgcagctatgactacaaagccgccggaagagaagccatcgcgaggggagaatccgg
tgatcagagctgcgaagatcgtaaaagtctga

[0254] GM51659494 cDNA 翻译成下列氨基酸序列 (SEQ ID NO :36) :

[0255] maqktekeetdfkvpetitlcvnncgvtgnpatnnmcqkftastattsgaggagiaspatrsgvsarp
qkrsfpeepspvadppssdqttpseakrvvnrcsgcrrkvgl tgfrcrcgelfcaehrysdrhdcsydykaagrea
iarenpviraakivkv

[0256] 源于大豆的 GM49780101 cDNA 序列 (SEQ ID NO :37) :

[0257] atggagcctcatgatgagactggatgccaggtcctgaacgccccattctttgcattaataattgtggc
ttcttttgaagagcagctaccatgaacatgtgttccaagtgttacaaggacatgctgttgaagcaggagcaggacaa
atttgcagcatcatccgttgaaaacattgtgaatggcagttccaatggcaatggaaagcaggctgtggctactgggtg
ctgttctgtgtacaagttgaagctgtggaggtcaagattgtctgtgtcagagttctgtggattcgtcctccggtgat
agtttggagatgaaagccaagactgggtcccagtagatgtgctacatgccggaaacgtgttggtttaactggtttcag
ctgcaaatgtggcaacctcttctgtgcaatgcatcgctattctgataaacatgattgcccttttgattataggactg

[0272] EST307 氨基酸序列 (SEQ ID NO :45)

[0273] matervsqettsqapegpvmcknlcgffgsqatmglskcyretvmqakmtalaeqatqaaqatsataa
avqppapvhetkl tcevertmivphqsssyqqdlvtpaaaapqavkssiaapsrpepnrcgscrkrvgltgfkcreg
nlycalhrysdkhtctydykaagqeaia kanplvvaekvvkf

序列表 <110> 巴斯夫植物科学有限公司

<120> 具有增加的胁迫耐受性和产量的转基因植物

<130> PF57972

<160> 45

<170> PatentIn version 3.4

<210> 1

<211> 615

<212> DNA

<213> 欧洲油菜

<220>

<221> CDS

<222> (1).. (615)

<223> 甲硫氨酸亚砷还原酶家族蛋白质 (BN51364980)

<400> 1

atg gct tct tct agt tgt ttc acc att cag tca cgt ttc gtc tca gcg	48
Met Ala Ser Ser Ser Cys Phe Thr Ile Gln Ser Arg Phe Val Ser Ala	
1 5 10 15	
aga aca aag ctc gat tca atc tcc aaa ccg agt ctc tcc gga ttc gct	96
Arg Thr Lys Leu Asp Ser Ile Ser Lys Pro Ser Leu Ser Gly Phe Ala	
20 25 30	
tgt cgt tct ctt aca aaa ccc aga aac ttg aat ctc tct gtt ctt ctt	144
Cys Arg Ser Leu Thr Lys Pro Arg Asn Leu Asn Leu Ser Val Leu Leu	
35 40 45	
cgg tgt tcc atg ggt tcc ttt aac tct tct cag aaa tca gac aac gtc	192
Arg Cys Ser Met Gly Ser Phe Asn Ser Ser Gln Lys Ser Asp Ash Val	
50 55 60	
caa gaa gct gca aag agt gac ttt gct tca ata agt gaa ggt gag tgg	240
Gln Glu Ala Ala Lys Ser Asp Phe Ala Ser Ile Ser Glu Gly Glu Trp	
65 70 75 80	
aag aaa cgg cta aca cca gaa cag tat tac atc acc aga cag aag gga	288
Lys Lys Arg Leu Thr Pro Glu Gln Tyr Tyr Ile Thr Arg Gln Lys Gly	

85	90	95	
aca gag aga gct ttc act ggt gag tat tgg aat aca aag acc cca gga			336
Thr Glu Arg Ala Phe Thr Gly Glu Tyr Trp Asn Thr Lys Thr Pro Gly			
100	105	110	
gta tac aaa tgt atc tgt tgc gac acg cca ctg ttt gac tca tca aca			384
Val Tyr Lys Cys Ile Cys Cys Asp Thr Pro Leu Phe Asp Ser Ser Thr			
115	120	125	
aag ctt gat agt gga acc ggg tgg cca tcg tat tac caa cct att gga			432
Lys Leu Asp Ser Gly Thr Gly Trp Pro Ser Tyr Tyr Gln Pro Ile Gly			
130	135	140	
aac aat gtg aag tca aag ctg gac ctc tct atc atc ttc atg cct aga			480
Asn Asn Val Lys Ser Lys Leu Asp Leu Ser Ile Ile Phe Met Pro Arg			
145	150	155	160
caa gaa gtt atc tgt gct gtt tgt aac gcc cat ctt ggt cat gtc ttc			528
Gln Glu Val Ile Cys Ala Val Cys Asn Ala His Leu Gly His Val Phe			
165	170	175	
gat gac ggt cca cga cca acc gga aaa cga tat tgc ctc aac agt gct			576
Asp Asp Gly Pro Arg Pro Thr Gly Lys Arg Tyr Cys Leu Asn Ser Ala			
180	185	190	
gct ctg aaa ctt gag tca ttg gag aga aca aga gaa tga			615
Ala Leu Lys Leu Glu Ser Leu Glu Arg Thr Arg Glu			
195	200		

<210>2

<211>204

<212>PRT

<213> 欧洲油菜

<400>2

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

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

<210>3

<211>615

<212>DNA

<213> 稻

<220>

<221>CDS

<222>(1).. (615)

<223> 甲硫氨酸亚砷还原酶家族蛋白质 (OS34096188)

<400>3

atg ggc ttc aat att ctg aga acc act tcc atc tcc act cct atc tct	48
Met Gly Phe Asn Ile Leu Arg Thr Thr Ser Ile Ser Thr Pro Ile Ser	
1 5 10 15	
tcc tcc aaa tcc aaa ccc att ttc tca act ctt ctt cgt tct tct cct	96
Ser Ser Lys Ser Lys Pro Ile Phe Ser Thr Leu Leu Arg Ser Ser Pro	
20 25 30	
tcc acc att ttc ccc cca aag tcc gtt act ccc acc act ctt ttc gtt	144
Ser Thr Ile Phe Pro Pro Lys Ser Val Thr Pro Thr Thr Leu Phe Val	
35 40 45	
tct gcc acc ccc ttc ttc act ctc cat ccc aag ctt ggt ttt cgt ggt	192
Ser Ala Thr Pro Phe Phe Thr Leu His Pro Lys Leu Gly Phe Arg Gly	

50	55	60	
ggg att gtg gcc atg gcc gca cct ggc tct ctc cgc aaa tcc gag gaa			240
Gly Ile Val Ala Met Ala Ala Pro Gly Ser Leu Arg Lys Ser Glu Glu			
65	70	75	80
gag tgg cgc gca att ctc tcc cct gaa cag ttt cgg atc ctc agg caa			288
Glu Trp Arg Ala Ile Leu Ser Pro Glu Gln Phe Arg Ile Leu Arg Gln			
85	90	95	
aag ggc acc gag ttc cct gga aca gga gag tat gac aag ttc tat gaa			336
Lys Gly Thr Glu Phe Pro Gly Thr Gly Glu Tyr Asp Lys Phe Tyr Glu			
100	105	110	
gag gga gtt tac aac tgt gct ggt tgt ggg act cca ctc tac agg tcc			384
Glu Gly Val Tyr Asn Cys Ala Gly Cys Gly Thr Pro Leu Tyr Arg Ser			
115	120	125	
ata aca aaa ttc aat tct ggt tgt ggc tgg cca gcc ttc tat gag ggg			432
Ile Thr Lys Phe Asn Ser Gly Cys Gly Trp Pro Ala Phe Tyr Glu Gly			
130	135	140	
att ccc gga gcc ata aat cgc aat ccg gat cct gat ggg atg agg aca			480
Ile Pro Gly Ala Ile Asn Arg Asn Pro Asp Pro Asp Gly Met Arg Thr			
145	150	155	160
gaa ata acg tgt gct gct tgt ggg gga cat cta ggt cac gtc ttt aaa			528
Glu Ile Thr Cys Ala Ala Cys Gly Gly His Leu Gly His Val Phe Lys			
165	170	175	
gga gaa gga ttt cca aca ccc act aac gaa cgc cat tgt gtc aat agc			576
Gly Glu Gly Phe Pro Thr Pro Thr Asn Glu Arg His Cys Val Asn Ser			
180	185	190	
att tcg ctg aaa ttt gcg cca gcc aat tct tat tct taa			615
Ile Ser Leu Lys Phe Ala Pro Ala Asn Ser Tyr Ser			
195	200		

<210>4

<211>204

<212>PRT

<213> 稻

<400>4

Met Gly Phe Asn Ile Leu Arg Thr Thr Ser Ile Ser Thr Pro Ile Ser
1 5 10 15
Ser Ser Lys Ser Lys Pro Ile Phe Ser Thr Leu Leu Arg Ser Ser Pro

atg gcc atg cgg caa tac gcg gct gct acc gct gcc tcc tcc agt ttc	48
Met Ala Met Arg Gln Tyr Ala Ala Ala Thr Ala Ala Ser Ser Ser Phe	
1 5 10 15	
aga gca cgt cca cgg gcg cgc ccc tcc tgc crc cca gcc gcc gcc ctg	96

Arg	Ala	Arg	Pro	Arg	Ala	Arg	Pro	Ser	Cys	Leu	Pro	Ala	Ala	Ala	Leu	
20				25				30								
ccc	ttg	gcg	cct	tgc	tgt	ggt	gtg	gcg	tgg	agc	cgt	gct	agc	tac	agg	144
Pro	Leu	Ala	Pro	Cys	Cys	Gly	Val	Ala	Trp	Ser	Arg	Ala	Ser	Tyr	Arg	
35				40				45								
cga	gcc	tcc	gtt	cgt	gcc	atg	ggt	gcc	gct	tca	tcg	tct	tcg	tcg	tcg	192
Arg	Ala	Ser	Val	Arg	Ala	Met	Gly	Ala	Ala	Ser	Ser	Ser	Ser	Ser	Ser	
50				55				60								
tcg	tcg	tcg	tct	ccg	tcg	ccg	cag	ggt	caa	gcc	caa	gcc	caa	gcc	caa	240
Ser	Ser	Ser	Ser	Pro	Ser	Pro	Gln	Gly	Gln	Ala	Gln	Ala	Gln	Ala	Gln	
65				70				75				80				
ggt	aaa	ccg	aac	tac	agt	aca	tct	ctg	act	gat	gag	gag	tgg	agg	aag	288
Gly	Lys	Pro	Asn	Tyr	Ser	Thr	Ser	Leu	Thr	Asp	Glu	Glu	Trp	Arg	Lys	
85				90				95								
cgc	ctg	aca	aaa	gat	cag	tat	tac	att	act	cgg	cag	aag	ggc	aca	gaa	336
Arg	Leu	Thr	Lys	Asp	Gln	Tyr	Tyr	Ile	Thr	Arg	Gln	Lys	Gly	Thr	Glu	
100				105				110								
aga	gca	ttt	act	ggg	gaa	tac	tgg	aac	acc	aaa	acc	ccg	ggc	atc	tac	384
Arg	Ala	Phe	Thr	Gly	Glu	Tyr	Trp	Asn	Thr	Lys	Thr	Pro	Gly	Ile	Tyr	
115				120				125								
cat	tgt	gtc	tgc	tgt	gac	acc	cct	ctt	ttt	gag	tca	tcg	acc	aaa	ttt	432
His	Cys	Val	Cys	Cys	Asp	Thr	Pro	Leu	Phe	Glu	Ser	Ser	Thr	Lys	Phe	
130				135				140								
gat	agt	ggt	act	ggg	tgg	ccg	tca	tat	tat	caa	ccc	att	gga	gat	aat	480
Asp	Ser	Gly	Thr	Gly	Trp	Pro	Ser	Tyr	Tyr	Gln	Pro	Ile	Gly	Asp	Asn	
145				150				155				160				
gta	aag	tgc	aag	ctt	gat	atg	tcc	atc	ata	ttc	atg	cct	cgg	act	gag	528
Val	Lys	Cys	Lys	Leu	Asp	Met	Ser	Ile	Ile	Phe	Met	Pro	Arg	Thr	Glu	
165				170				175								
gtg	ctg	tgt	gct	gtc	tgt	gac	gct	cat	ctg	ggg	cac	gtg	ttt	gat	gat	576
Val	Leu	Cys	Ala	Val	Cys	Asp	Ala	His	Leu	Gly	His	Val	Phe	Asp	Asp	
180				185				190								
ggg	cca	cga	cca	aca	ggg	aaa	aga	tac	tgt	atc	aat	agc	gca	tct	ctc	624
Gly	Pro	Arg	Pro	Thr	Gly	Lys	Arg	Tyr	Cys	Ile	Asn	Ser	Ala	Ser	Leu	
195				200				205								
aag	ctg	aag	aag	acc	cag	tag										645
Lys	Leu	Lys	Lys	Thr	Gln											
210																

<210>6

<211>214

<212>PRT

<213> 稻

<400>6

```
Met Ala Met Arg Gln Tyr Ala Ala Ala Thr Ala Ala Ser Ser Ser Phe
1           5           10           15
Arg Ala Arg Pro Arg Ala Arg Pro Ser Cys Leu Pro Ala Ala Ala Leu
          20           25           30
Pro Leu Ala Pro Cys Cys Gly Val Ala Trp Ser Arg Ala Ser Tyr Arg
          35           40           45
Arg Ala Ser Val Arg Ala Met Gly Ala Ala Ser Ser Ser Ser Ser Ser
          50           55           60
Ser Ser Ser Ser Pro Ser Pro Gln Gly Gln Ala Gln Ala Gln Ala Gln
65           70           75           80
Gly Lys Pro Asn Tyr Ser Thr Ser Leu Thr Asp Glu Glu Trp Arg Lys
          85           90           95
Arg Leu Thr Lys Asp Gln Tyr Tyr Ile Thr Arg Gln Lys Gly Thr Glu
          100          105          110
Arg Ala Phe Thr Gly Glu Tyr Trp Asn Thr Lys Thr Pro Gly Ile Tyr
          115          120          125
His Cys Val Cys Cys Asp Thr Pro Leu Phe Glu Ser Ser Thr Lys Phe
          130          135          140
Asp Ser Gly Thr Gly Trp Pro Ser Tyr Tyr Gln Pro Ile Gly Asp Asn
145          150          155          160
Val Lys Cys Lys Leu Asp Met Ser Ile Ile Phe Met Pro Arg Thr Glu
          165          170          175
Val Leu Cys Ala Val Cys Asp Ala His Leu Gly His Val Phe Asp Asp
          180          185          190
Gly Pro Arg Pro Thr Gly Lys Arg Tyr Cys Ile Asn Ser Ala Ser Leu
          195          200          205
Lys Leu Lys Lys Thr Gln
          210
```

<210>7

<211>609

<212>DNA

<213> 大豆

gga gaa gga ttt cca acg ccc act aac gaa cgc cat tgt gtc aat agc 576
 Gly Glu Gly Phe Pro Thr Pro Thr Asn Glu Arg His Cys Val Asn Ser
 180 185 190
 att tca ctg aaa ttt gcg cca gcc aat tct taa 609
 Ile Ser Leu Lys Phe Ala Pro Ala Asn Ser
 195 200

<210>8

<211>202

<212>PRT

<213> 大豆

<400>8

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

195	200	
<210>9		
<211>621		
<212>DNA		
<213> 小麦		
<220>		
<221>CDS		
<222>(1).. (621)		
<223> 甲硫氨酸亚砷还原酶家族蛋白质 (TA6540264)		
<400>9		
atg gcg tcg ccc cac gcc cac ccg gcc acg cgg ccc ctc tca tcg ctc	48	
Met Ala Ser Pro His Ala His Pro Ala Thr Arg Pro Leu Ser Ser Leu		
1 5 10 15		
ccg tcc ctc ctc ctc gcc cgc tcc tcc tcc gcc gcc acc gcc gcc gcg	96	
Pro Ser Leu Leu Leu Ala Arg Ser Ser Ser Ala Ala Thr Ala Ala Ala		
20 25 30		
tcg tcc gcc cgc ccc gcc tcc ctc tcc ctc tcg tgc tcg cgg tcg cgg	144	
Ser Ser Ala Arg Pro Ala Ser Leu Ser Leu Ser Cys Ser Arg Ser Arg		
35 40 45		
gcg cgg gcc tac tgc cca gcc gga cga cgg ttg ccg ggc gcc gtg gtg	192	
Ala Arg Ala Tyr Cys Pro Ala Gly Arg Arg Leu Pro Gly Ala Val Val		
50 55 60		
gct atg tcg tcg gcg gcg ccc acg ccg ggg ccc gtg cag aag tcg gag	240	
Ala Met Ser Ser Ala Ala Pro Thr Pro Gly Pro Val Gln Lys Ser Glu		
65 70 75 80		
gag gag tgg gag gcc gtc ctc acg ccg gag cag ttc cgc atc ctc cgc	288	
Glu Glu Trp Glu Ala Val Leu Thr Pro Glu Gln Phe Arg Ile Leu Arg		
85 90 95		
cgc aag ggc acc gag tat cct gga aca ggt gaa tat gac aag ttc ttc	336	
Arg Lys Gly Thr Glu Tyr Pro Gly Thr Gly Glu Tyr Asp Lys Phe Phe		
100 105 110		
agt gag ggt att tac gga tgt gct ggc tgt gga acc ccc ttg tac aaa	384	
Ser Glu Gly Ile Tyr Gly Cys Ala Gly Cys Gly Thr Pro Leu Tyr Lys		
115 120 125		
tca tct acg aag ttc aac tca ggg tgt ggt tgg cca gca ttc tat gaa	432	
Ser Ser Thr Lys Phe Asn Ser Gly Cys Gly Trp Pro Ala Phe Tyr Glu		

130	135	140	
gga ttt cct gga gcc ata aaa cgg acg gcg gat cct gat ggg agg cga			480
Gly Phe Pro Gly Ala Ile Lys Arg Thr Ala Asp Pro Asp Gly Arg Arg			
145	150	155	160
att gag atc aca tgt gct gct tgt gaa gga cat ctg ggg cat gtg ttc			528
Ile Glu Ile Thr Cys Ala Ala Cys Glu Gly His Leu Gly His Val Phe			
165	170	175	
aaa ggg gag ggg ttc aac acg ccg act gat gag cga cac tgc gtc aac			576
Lys Gly Glu Gly Phe Asn Thr Pro Thr Asp Glu Arg His Cys Val Asn			
180	185	190	
agt atc tca ctc aag ttc gtt ccg gcc tct gaa gag gct agt tga			621
Ser Ile Ser Leu Lys Phe Val Pro Ala Ser Glu Glu Ala Ser			
195	200	205	

<210>10

<211>206

<212>PRT

<213> 小麦

<400>10

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

Gly Phe Pro Gly Ala Ile Lys Arg Thr Ala Asp Pro Asp Gly Arg Arg
 145 150 155 160
 Ile Glu Ile Thr Cys Ala Ala Cys Glu Gly His Leu Gly His Val Phe
 165 170 175
 Lys Gly Glu Gly Phe Asn Thr Pro Thr Asp Glu Arg His Cys Val Asn
 180 185 190
 Ser Ile Ser Leu Lys Phe Val Pro Ala Ser Glu Glu Ala Ser
 195 200 205

<210>11

<211>933

<212>DNA

<213> 欧洲油菜

<220>

<221>CDS

<222>(1)..(933)

<223> 同源异型域 - 亮氨酸拉链蛋白质 (BN45206322)

<400>11

atg atg aag aga tta agc agt tca gat tca gtg ggt ggt ctc atc tct	48
Met Met Lys Arg Leu Ser Ser Ser Asp Ser Val Gly Gly Leu Ile Ser	
1 5 10 15	
tta tgt ccc act act tcc aca gat cag ccg aat cca aga aga tgc ggg	96
Leu Cys Pro Thr Thr Ser Thr Asp Gln Pro Asn Pro Arg Arg Cys Gly	
20 25 30	
aga gaa ttt cag tcg atg ctc gaa ggt tac gag gag gaa gaa gaa gaa	144
Arg Glu Phe Gln Ser Met Leu Glu Gly Tyr Glu Glu Glu Glu Glu Glu	
35 40 45	
gcc ata acc gag gaa aga gga caa acc ggt tta gcc gag aag aag aga	192
Ala Ile Thr Glu Glu Arg Gly Gln Thr Gly Leu Ala Glu Lys Lys Arg	
50 55 60	
cgg tta aac att aac caa gtt aaa gcc ttg gag aaa aat ttc gag tta	240
Arg Leu Asn Ile Asn Gln Val Lys Ala Leu Glu Lys Asn Phe Glu Leu	
65 70 75 80	
gag aac aag ctt gag cct gag agg aaa gtg aag tta gct caa gaa ctt	288
Glu Asn Lys Leu Glu Pro Glu Arg Lys Val Lys Leu Ala Gln Glu Leu	
85 90 95	
ggt ctc caa cct cgt caa gta gct gtt tgg ttt cag aac cgc cgt gcg	336
Gly Leu Gln Pro Arg Gln Val Ala Val Trp Phe Gln Asn Arg Arg Ala	

100	105	110	
cgg tgg aag aca aaa cag ctt gag aaa gat tac ggt gtt ctc aaa acg			384
Arg Trp Lys Thr Lys Gln Leu Glu Lys Asp Tyr Gly Val Leu Lys Thr			
115	120	125	
caa tac gat tct ctc cgc cat aac ttt gat tcc ctc cgc cgt gaa aat			432
Gln Tyr Asp Ser Leu Arg His Asn Phe Asp Ser Leu Arg Arg Glu Asn			
130	135	140	
gaa tct ctt ctt caa gag atc ggt aaa cta aaa gct aag ctt aac gga			480
Glu Ser Leu Leu Gln Glu Ile Gly Lys Leu Lys Ala Lys Leu Asn Gly			
145	150	155	160
gaa gaa gaa gga gat gat gtt gat gaa gaa gag aac aac ttg gcg acg			528
Glu Glu Glu Gly Asp Asp Val Asp Glu Glu Glu Asn Asn Leu Ala Thr			
165	170	175	
atg gag agt gat gtt tcc gtc aag gaa gaa gaa gtt tcg ttg ccg gag			576
Met Glu Ser Asp Val Ser Val Lys Glu Glu Glu Val Ser Leu Pro Glu			
180	185	190	
cag atc aca gag ccg ccg tct tct cct ccg cag ctt cta gag cat tcc			624
Gln Ile Thr Glu Pro Pro Ser Ser Pro Pro Gln Leu Leu Glu His Ser			
195	200	205	
gac agt ttc aat tac cgg agt ttc acc gac ctc cgc gac ctt ctt ccg			672
Asp Ser Phe Asn Tyr Arg Ser Phe Thr Asp Leu Arg Asp Leu Leu Pro			
210	215	220	
tta aag gcc gcg gct tcc tcc gtc gcc gcc gct gga tcg tcg gac agt			720
Leu Lys Ala Ala Ala Ser Ser Val Ala Ala Ala Gly Ser Ser Asp Ser			
225	230	235	240
agc gat tcg agc gcc gtg ttg aac gag gaa agt agc tct aac gtt acg			768
Ser Asp Ser Ser Ala Val Leu Asn Glu Glu Ser Ser Ser Asn Val Thr			
245	250	255	
gcg gct ccg gcg acg gtt ccc ggc ggc agt ttc ttg cag ttt gtg aaa			816
Ala Ala Pro Ala Thr Val Pro Gly Gly Ser Phe Leu Gln Phe Val Lys			
260	265	270	
atg gag cag acg gag gat cac gac gac ttt ctg agt gga gaa gaa gcg			864
Met Glu Gln Thr Glu Asp His Asp Asp Phe Leu Ser Gly Glu Glu Ala			
275	280	285	
tgc ggg ttt ttc tcc gat gaa cag cca ccg tct ctg cac tgg tat tcc			912
Cys Gly Phe Phe Ser Asp Glu Gln Pro Pro Ser Leu His Trp Tyr Ser			
290	295	300	
acc gtt gat cag tgg aac tga			933
Thr Val Asp Gln Trp Asn			

305

310

<210>12

<211>310

<212>PRT

<213> 欧洲油菜

<400>12

```

Met Met Lys Arg Leu Ser Ser Ser Asp Ser Val Gly Gly Leu Ile Ser
1           5           10           15
Leu Cys Pro Thr Thr Ser Thr Asp Gln Pro Asn Pro Arg Arg Cys Gly
           20           25           30
Arg Glu Phe Gln Ser Met Leu Glu Gly Tyr Glu Glu Glu Glu Glu Glu
           35           40           45
Ala Ile Thr Glu Glu Arg Gly Gln Thr Gly Leu Ala Glu Lys Lys Arg
           50           55           60
Arg Leu Asn Ile Asn Gln Val Lys Ala Leu Glu Lys Asn Phe Glu Leu
65           70           75           80
Glu Asn Lys Leu Glu Pro Glu Arg Lys Val Lys Leu Ala Gln Glu Leu
           85           90           95
Gly Leu Gln Pro Arg Gln Val Ala Val Trp Phe Gln Asn Arg Arg Ala
           100          105          110
Arg Trp Lys Thr Lys Gln Leu Glu Lys Asp Tyr Gly Val Leu Lys Thr
           115          120          125
Gln Tyr Asp Ser Leu Arg His Asn Phe Asp Ser Leu Arg Arg Glu Asn
           130          135          140
Glu Ser Leu Leu Gln Glu Ile Gly Lys Leu Lys Ala Lys Leu Asn Gly
145           150           155           160
Glu Glu Glu Gly Asp Asp Val Asp Glu Glu Glu Asn Asn Leu Ala Thr
           165           170           175
Met Glu Ser Asp Val Ser Val Lys Glu Glu Glu Val Ser Leu Pro Glu
           180          185          190
Gln Ile Thr Glu Pro Pro Ser Ser Pro Pro Gln Leu Leu Glu His Ser
           195          200          205
Asp Ser Phe Asn Tyr Arg Ser Phe Thr Asp Leu Arg Asp Leu Leu Pro
           210          215          220
Leu Lys Ala Ala Ala Ser Ser Val Ala Ala Ala Gly Ser Ser Asp Ser
225           230           235           240

```

Ser Asp Ser Ser Ala Val Leu Asn Glu Glu Ser Ser Ser Asn Val Thr
 245 250 255
 Ala Ala Pro Ala Thr Val Pro Gly Gly Ser Phe Leu Gln Phe Val Lys
 260 265 270
 Met Glu Gln Thr Glu Asp His Asp Asp Phe Leu Ser Gly Glu Glu Ala
 275 280 285
 Cys Gly Phe Phe Ser Asp Glu Gln Pro Pro Ser Leu His Trp Tyr Ser
 290 295 300
 Thr Val Asp Gln Trp Asn
 305 310

<210>13

<211>996

<212>DNA

<213> 大豆

<220>

<221>CDS

<222>(1).. (996)

<223> 同源异型域 - 亮氨酸拉链蛋白质 (GM48923793)

<400>13

atg gcg ggt agt gga agt gcc ttt tcc aac atc act agc ttt ctt cgc 48
 Met Ala Gly Ser Gly Ser Ala Phe Ser Asn Ile Thr Ser Phe Leu Arg
 1 5 10 15
 acc caa caa ccc tct tct caa cct ctc gat tct tct ctc ttc ctc tct 96
 Thr Gln Gln Pro Ser Ser Gln Pro Leu Asp Ser Ser Leu Phe Leu Ser
 20 25 30
 gca cct tcc tct gct cct ttc ctc ggt tcg aga tcc atg atg agt ttt 144
 Ala Pro Ser Ser Ala Pro Phe Leu Gly Ser Arg Ser Met Met Ser Phe
 35 40 45
 gat gga gaa gga ggg aag ggg tgt aac ggc tcc ttc ttc cgc gcg ttt 192
 Asp Gly Glu Gly Gly Lys Gly Cys Asn Gly Ser Phe Phe Arg Ala Phe
 50 55 60
 gac atg gac gac aat ggg gat gag tgc atg gac gag tac ttt cat caa 240
 Asp Met Asp Asp Asn Gly Asp Glu Cys Met Asp Glu Tyr Phe His Gln
 65 70 75 80
 ccc gag aag aag cga cgt ctc tct gcg agc cag gtt cag ttt cta gag 288
 Pro Glu Lys Lys Arg Arg Leu Ser Ala Ser Gln Val Gln Phe Leu Glu

85	90	95	
aag agc ttc gag gag gag aac aag ctt gaa ccc gag aga aag acc aaa			336
Lys Ser Phe Glu Glu Glu Asn Lys Leu Glu Pro Glu Arg Lys Thr Lys			
100	105	110	
cta gcc aaa gac ctt ggt ttg cag cca cgg caa gtt gct att tgg ttc			384
Leu Ala Lys Asp Leu Gly Leu Gln Pro Arg Gln Val Ala Ile Trp Phe			
115	120	125	
cag aac cgt aga gct cgg tgg aag aac aaa cag ctg gag aag gat tac			432
Gln Asn Arg Arg Ala Arg Trp Lys Asn Lys Gln Leu Glu Lys Asp Tyr			
130	135	140	
gag act ctg cat gca agt ttt gag agt ctc aag tcc aac tat gac tgt			480
Glu Thr Leu His Ala Ser Phe Glu Ser Leu Lys Ser Asn Tyr Asp Cys			
145	150	155	160
ctt ctc aag gag aaa gac aag tta aaa gct gag gtg gcg agc ctc act			528
Leu Leu Lys Glu Lys Asp Lys Leu Lys Ala Glu Val Ala Ser Leu Thr			
165	170	175	
gag aag gtg ctt gca aga ggg aaa caa gag ggg cac atg aag cag gct			576
Glu Lys Val Leu Ala Arg Gly Lys Gln Glu Gly His Met Lys Gln Ala			
180	185	190	
gaa agt gaa agt gaa gaa aca aaa gga tta ttg cat ttg cag gaa cag			624
Glu Ser Glu Ser Glu Glu Thr Lys Gly Leu Leu His Leu Gln Glu Gln			
195	200	205	
gaa cca ccc cag agg ctt tta ctg caa tca gtt tcg gag gga gaa gga			672
Glu Pro Pro Gln Arg Leu Leu Leu Gln Ser Val Ser Glu Gly Glu Gly			
210	215	220	
tcc aaa gtc tct tct gtc gtt ggg ggt tgt aaa cag gaa gat atc agt			720
Ser Lys Val Ser Ser Val Val Gly Gly Cys Lys Gln Glu Asp Ile Ser			
225	230	235	240
tca gca agg agt gac att ttg gat tca gat agt cca cat tac acc gat			768
Ser Ala Arg Ser Asp Ile Leu Asp Ser Asp Ser Pro His Tyr Thr Asp			
245	250	255	
gga gtt cac tct gcg ctg cta gag cat ggt gat tct tct tat gtg ttt			816
Gly Val His Ser Ala Leu Leu Glu His Gly Asp Ser Ser Tyr Val Phe			
260	265	270	
gag cct gat caa tca gat atg tca caa gat gaa gaa gat aac ctc agc			864
Glu Pro Asp Gln Ser Asp Met Ser Gln Asp Glu Glu Asp Asn Leu Ser			
275	280	285	
aag agt ctc tac cct tcg tac ctc ttt ccc aaa ctt gaa gaa gat gtg			912
Lys Ser Leu Tyr Pro Ser Tyr Leu Phe Pro Lys Leu Glu Glu Asp Val			

54

195					200					205						
Glu	Pro	Pro	Gln	Arg	Leu	Leu	Leu	Gln	Ser	Val	Ser	Glu	Gly	Glu	Gly	
210					215					220						
Ser	Lys	Val	Ser	Ser	Val	Val	Gly	Gly	Cys	Lys	Gln	Glu	Asp	Ile	Ser	
225					230					235					240	
Ser	Ala	Arg	Ser	Asp	Ile	Leu	Asp	Ser	Asp	Ser	Pro	His	Tyr	Thr	Asp	
245					250					255						
Gly	Val	His	Ser	Ala	Leu	Leu	Glu	His	Gly	Asp	Ser	Ser	Tyr	Val	Phe	
260					265					270						
Glu	Pro	Asp	Gln	Ser	Asp	Met	Ser	Gln	Asp	Glu	Glu	Asp	Asn	Leu	Ser	
275					280					285						
Lys	Ser	Leu	Tyr	Pro	Ser	Tyr	Leu	Phe	Pro	Lys	Leu	Glu	Glu	Asp	Val	
290					295					300						
Asp	Tyr	Ser	Asp	Pro	Pro	Glu	Ser	Ser	Cys	Asn	Phe	Gly	Phe	Pro	Glu	
305					310					315					320	
Glu	Asp	His	Val	Leu	Trp	Thr	Trp	Ala	Tyr	Tyr						
325					330											

 $\langle 210 \rangle_{15}$ $\langle 211 \rangle 1023$

<212>DNA

〈213〉 小麦

 $\langle 220 \rangle$ $\langle 221 \rangle$ CDS $\langle 222 \rangle (1) \dots (1023)$

〈223〉 同源异型域-亮氨酸拉链蛋白质 (TA55969932)

 $\langle 400 \rangle_{15}$

atg	gag	ccc	ggc	cgg	ctc	atc	ttc	aac	acg	tcg	ggc	tcc	ggc	aac	gga	48
Met	Glu	Pro	Gly	Arg	Leu	Ile	Phe	Asn	Thr	Ser	Gly	Ser	Gly	Asn	Gly	
1			5					10						15		
cag	atg	ctc	ttc	atg	gac	tgc	ggc	gcg	ggc	ggc	atc	gcc	ggc	gcg	gcc	96
Gln	Met	Leu	Phe	Met	Asp	Cys	Gly	Ala	Gly	Gly	Ile	Ala	Gly	Ala	Ala	
			20					25						30		
ggc	atg	ttc	cat	cga	ggg	gtg	aga	ccg	gtc	ctc	ggc	ggc	atg	gaa	gaa	144
Gly	Met	Phe	His	Arg	Gly	Val	Arg	Pro	Val	Leu	Gly	Gly	Met	Glu	Glu	
			35					40						45		
ggg	cgc	ggc	gtg	aag	cgg	ccc	ttc	ttc	acc	tcg	ccg	gat	gac	atg	ctg	192

Gly	Arg	Gly	Val	Lys	Arg	Pro	Phe	Phe	Thr	Ser	Pro	Asp	Asp	Met	Leu	
50				55				60								
gag	gag	gag	tac	tac	gac	gag	cag	ctc	ccg	gag	aag	aag	cgg	cgc	ctc	240
Glu	Glu	Glu	Tyr	Tyr	Asp	Glu	Gln	Leu	Pro	Glu	Lys	Lys	Arg	Arg	Leu	
65				70				75				80				
acg	ccg	gag	cag	gtc	cac	ctg	ctg	gag	agg	agc	ttc	gag	gag	gag	aac	288
Thr	Pro	Glu	Gln	Val	His	Leu	Leu	Glu	Arg	Ser	Phe	Glu	Glu	Glu	Asn	
85				90				95								
aag	ctg	gag	ccg	gag	agg	aag	acg	gag	ctg	gcc	cgc	aag	ctc	ggg	ctg	336
Lys	Leu	Glu	Pro	Glu	Arg	Lys	Thr	Glu	Leu	Ala	Arg	Lys	Leu	Gly	Leu	
100				105				110								
cag	cca	cgg	cag	gtg	gcc	gtc	tgg	ttc	cag	aac	cgc	cgc	gcc	cgg	tgg	384
Gln	Pro	Arg	Gln	Val	Ala	Val	Trp	Phe	Gln	Asn	Arg	Arg	Ala	Arg	Trp	
115				120				125								
aag	aca	aag	acg	ctg	gag	cgc	gac	ttc	gac	cgc	ctc	aag	gcg	tcc	ttc	432
Lys	Thr	Lys	Thr	Leu	Glu	Arg	Asp	Phe	Asp	Arg	Leu	Lys	Ala	Ser	Phe	
130				135				140								
gac	gcc	ctc	cgc	gcc	gac	cac	gac	gcg	ctc	ctc	cag	gac	aac	cac	cgg	480
Asp	Ala	Leu	Arg	Ala	Asp	His	Asp	Ala	Leu	Leu	Gln	Asp	Asn	His	Arg	
145				150				155				160				
ctc	cgg	tca	cag	gtg	gta	acg	ttg	acc	gag	aag	atg	caa	gat	aag	gag	528
Leu	Arg	Ser	Gln	Val	Val	Thr	Leu	Thr	Glu	Lys	Met	Gln	Asp	Lys	Glu	
165				170				175								
gcg	ccg	gaa	ggc	agc	ttc	ggt	gca	gcc	gcc	gac	gcc	tcg	gag	ccg	gag	576
Ala	Pro	Glu	Gly	Ser	Phe	Gly	Ala	Ala	Ala	Asp	Ala	Ser	Glu	Pro	Glu	
180				185				190								
cag	gcg	gcg	gcg	gag	gcg	aag	gct	tcc	ttg	gcc	gac	gcc	gag	gag	cag	624
Gln	Ala	Ala	Ala	Glu	Ala	Lys	Ala	Ser	Leu	Ala	Asp	Ala	Glu	Glu	Gln	
195				200				205								
gcc	gcg	gca	gcg	gag	gcg	ttc	gag	gtg	gtg	cag	cag	cag	ctg	cac	gtg	672
Ala	Ala	Ala	Ala	Glu	Ala	Phe	Glu	Val	Val	Gln	Gln	Gln	Leu	His	Val	
210				215				220								
aag	gac	gag	gag	agg	ctg	agc	ccg	ggg	agc	ggc	ggg	agc	gcg	gtg	ctg	720
Lys	Asp	Glu	Glu	Arg	Leu	Ser	Pro	Gly	Ser	Gly	Gly	Ser	Ala	Val	Leu	
225				230				235				240				
gac	gcg	agg	gac	gcg	ctg	ctc	ggg	agc	gga	tgc	ggc	ctc	gcc	ggc	gtg	768
Asp	Ala	Arg	Asp	Ala	Leu	Leu	Gly	Ser	Gly	Cys	Gly	Leu	Ala	Gly	Val	
245				250				255								
gtg	gac	agc	agc	gtg	gac	tcg	tac	tgc	ttc	ccg	ggg	ggc	gcc	ggc	ggc	816

Val Asp Ser Ser Val Asp Ser Tyr Cys Phe Pro Gly Gly Ala Gly Gly	
260	265
gac gag tac cac gag tgc gtg gtg ggc ccc gtg gcg ggc ggc atc cag	864
Asp Glu Tyr His Glu Cys Val Val Gly Pro Val Ala Gly Gly Ile Gln	
275	280
tgc gag gag gac gac ggc gcg ggc agc gac gag ggc tgc agc tac tac	912
Ser Glu Glu Asp Asp Gly Ala Gly Ser Asp Glu Gly Cys Ser Tyr Tyr	
290	295
ccc gac gac gcc gcc gtc ttc ttc gcc gcc gcg caa ggg cac ggc cac	960
Pro Asp Asp Ala Ala Val Phe Phe Ala Ala Ala Gln Gly Hi s Gly His	
305	310
cat cgc acg gac gac gac gat cag cag gac gac ggc cag atc agc tac	1008
His Arg Thr Asp Asp Asp Asp Gln Gln Asp Asp Gly Gln Ile Ser Tyr	
325	330
tgg atg tgg aac tag	1023
Trp Met Trp Asn	
340	

<210>16

<211>340

<212>PRT

<213> 小麦

<400>16

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

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

<210>17

<211>534

<212>DNA

<213> 欧洲油菜

<220>

<221>CDS

<222>(1).. (534)

$\langle 400 \rangle_{17}$

59

<210>18

<211>177

<212>PRT

<213> 欧洲油菜

<400>18

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

<210>19

<211>564

<212>DNA

<213> 欧洲油菜

<220>

<221>CDS

<222>(1).. (564)

<400>19

61

<210>20

<211>187

<212>PRT

<213> 欧洲油菜

<400>20

```

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

```

<210>21

<211>465

<212>DNA

<213> 大麦

<220>

<221>CDS

<222>(1).. (465)

<223> 含有 A20 结构域与 AN1 结构域的锌指蛋白质 (HV62552639)

<400>21

atg gcc cag gag agt tgt gat ctc aac aag gac gag gcc gag atc ctg	48
Met Ala Gln Glu Ser Cys Asp Leu Asn Lys Asp Glu Ala Glu Ile Leu	
1 5 10 15	
aag cca tcc tcc tcc aca cct tgc cct cct tgc cca gcc aca cca cca	96
Lys Pro Ser Ser Ser Thr Pro Ser Pro Pro Ser Pro Ala Thr Pro Pro	
20 25 30	
cca cca acc gct caa ata cca gaa cca caa cct cca cac tca cca cca	144
Pro Pro Thr Ala Gln Ile Pro Glu Pro Gln Pro Pro His Ser Pro Pro	
35 40 45	
caa cca ccg gca gct caa ttc ttg tcc agg ccc tgc gag gtt gtt ccc	192
Gln Pro Pro Ala Ala Gln Phe Leu Ser Arg Pro Cys Glu Val Val Pro	
50 55 60	
ata gag act tcc aaa aag agg aaa cat gct gat gcg gtg tca atg gcc	240
Ile Glu Thr Ser Lys Lys Arg Lys His Ala Asp Ala Val Ser Met Ala	
65 70 75 80	
att gtg gtt gag cca ttg tgc tct gtg ctg ttc gtt aac cgt tgc aac	288
Ile Val Val Glu Pro Leu Ser Ser Val Leu Phe Val Asn Arg Cys Asn	
85 90 95	
gtg tgc cgc aag aga gtt ggt ttg acc ggg ttc cgt tgc cgg tgt gag	336
Val Cys Arg Lys Arg Val Gly Leu Thr Gly Phe Arg Cys Arg Cys Glu	
100 105 110	
aag ctc ttt tgt ccg cgc cac cgg cat tca gaa agc cac gac tgc tca	384
Lys Leu Phe Cys Pro Arg His Arg His Ser Glu Ser His Asp Cys Ser	
115 120 125	
ttt gat tat aaa act gtg ggt cgg gag gag att gcc cgg gca aac cct	432
Phe Asp Tyr Lys Thr Val Gly Arg Glu Glu Ile Ala Arg Ala Asn Pro	
130 135 140	
ctg atc agg gct gcc aag atc att agg ata tga	465
Leu Ile Arg Ala Ala Lys Ile Ile Arg Ile	
145 150	

<210>22

<211>154

<212>PRT

<213> 大麦

<400>22

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

<210>23

<211>516

<212>DNA

<213> 玉米

<220>

<221>CDS

<222>(1).. (516)

<223> 含有 A20 结构域与 AN1 结构域的锌指蛋白质 (ZM61995511)

<400>23

atg gaa cac aag gag gcg ggc tgc cag cag ccg gag ggc cca atc cta 48
 Met Glu His Lys Glu Ala Gly Cys Gln Gln Pro Glu Gly Pro Ile Leu
 1 5 10 15
 tgc atc aat aac tgc ggc ttc ttc ggc agt gct gcg acg atg aac atg 96

Cys Ile Asn Asn Cys Gly Phe Phe Gly Ser Ala Ala Thr Met Asn Met	
20 25 30	
tgc tcc aag tgc cac aag gag atg ata acg aag cag gag cag gcc cag	144
Cys Ser Lys Cys His Lys Glu Met Ile Thr Lys Gln Glu Gln Ala Gln	
35 40 45	
ctg gct gcc tcc ccc atc gat agc att gtc aat ggc ggt gac ggc ggg	192
Leu Ala Ala Ser Pro Ile Asp Ser Ile Val Asn Gly Gly Asp Gly Gly	
50 55 60	
aaa gga cct gta att gct gca tct gta aat gtg gca gtt cct caa gtt	240
Lys Gly Pro Val Ile Ala Ala Ser Val Asn Val Ala Val Pro Gln Val	
65 70 75 80	
gag cag aag act att gtt gtg cag ccc atg ctt gta gct gaa acc agc	288
Glu Gln Lys Thr Ile Val Val Gln Pro Met Leu Val Ala Glu Thr Ser	
85 90 95	
gag gct gct gct gta atc ccc aag gcc aag gaa ggc cca gac cgg tgc	336
Glu Ala Ala Ala Val Ile Pro Lys Ala Lys Glu Gly Pro Asp Arg Cys	
100 105 110	
gcg gcc tgc agg aag cgt gtt ggg ctg acg gga ttt agc tgc cga tgc	384
Ala Ala Cys Arg Lys Arg Val Gly Leu Thr Gly Phe Ser Cys Arg Cys	
115 120 125	
ggg aac atg tac tgt tcg gtg cac cgc tac tcc gac aaa cat gac tgt	432
Gly Asn Met Tyr Cys Ser Val His Arg Tyr Ser Asp Lys His Asp Cys	
130 135 140	
cag ttc gac tat cgg act gca gca agg gac gcg att gcc aag gcc aat	480
Gln Phe Asp Tyr Arg Thr Ala Ala Arg Asp Ala Ile Ala Lys Ala Asn	
145 150 155 160	
cct gtg gtg agg gcg gag aag ctc gac aag atc tga	516
Pro Val Val Arg Ala Glu Lys Leu Asp Lys Ile	
165 170	

<210>24

<211>171

<212>PRT

<213> 玉米

<400>24

Met Glu His Lys Glu Ala Gly Cys Gln Gln Pro Glu Gly Pro Ile Leu
1 5 10 15

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

<210>25

<211>720

<212>DNA

<213> 亚麻

<220>

<221>CDS

<222>(1).. (720)

<223> 含有 A20 结构域与 AN1 结构域的锌指蛋白质 (LU61567101)

<400>25

atg gct cct tca cct tgc gtc cac ggc tgc acg gcc aat tgc ccc cgc	48
Met Ala Pro Ser Pro Cys Val His Gly Cys Thr Ala Asn Cys Pro Arg	
1 5 10 15	
tgc cac tct tac gga cac ccc atc ttc ggg aac tca gat ctc gcc gct	96
Cys His Ser Tyr Gly His Pro Ile Phe Gly Asn Ser Asp Leu Ala Ala	
20 25 30	
ggc ggc agc gat acg tcc acg tcg gtg ttt gga aaa gta gga tcc gtc	144

Gly	Gly	Ser	Asp	Thr	Ser	Thr	Ser	Val	Phe	Gly	Lys	Val	Gly	Ser	Val	
35				40				45								
gtg	att	cag	tgc	cct	gcg	aag	aat	cac	gcg	ttc	ggc	caa	gct	tgt	ggc	192
Val	Ile	Gln	Ser	Pro	Ala	Lys	Asn	His	Ala	Phe	Gly	Gln	Ala	Cys	Gly	
50				55				60								
ccg	gtt	ttt	ccc	tgc	agc	tcc	tcc	cct	ttc	cgc	cgc	atc	aag	ttc	ggc	240
Pro	Val	Phe	Pro	Ser	Ser	Ser	Ser	Pro	Phe	Arg	Arg	Ile	Lys	Phe	Gly	
65				70				75				80				
ccc	aaa	gat	ggc	gag	ggg	aaa	gga	ccg	ctg	aag	ccg	atc	gag	aag	cag	288
Pro	Lys	Asp	Gly	Glu	Gly	Lys	Gly	Pro	Leu	Lys	Pro	Ile	Glu	Lys	Gln	
85				90				95								
ccg	tgc	aag	aag	cgt	ccg	ttc	tgc	ttc	tct	ccc	gac	gag	acg	att	gac	336
Pro	Ser	Lys	Lys	Arg	Pro	Phe	Cys	Phe	Ser	Pro	Asp	Glu	Thr	Ile	Asp	
100				105				110								
gcg	acg	gtt	cct	ccg	tcc	acc	aaa	ccg	ttc	ggt	tgc	ttc	cgt	tcc	gtc	384
Ala	Thr	Val	Pro	Pro	Ser	Thr	Lys	Pro	Phe	Gly	Ser	Phe	Arg	Ser	Val	
115				120				125								
tgt	gtc	acg	gac	gcc	gac	gag	gcc	agg	ttg	aag	gcg	aac	cgc	gag	ttc	432
Cys	Val	Thr	Asp	Ala	Asp	Glu	Ala	Arg	Leu	Lys	Ala	Asn	Arg	Glu	Phe	
130				135				140								
ttc	gct	ccg	gta	tcc	cgc	aaa	cgt	ggc	ttc	gat	ccg	act	gac	atg	acc	480
Phe	Ala	Pro	Val	Ser	Arg	Lys	Arg	Gly	Phe	Asp	Pro	Thr	Asp	Met	Thr	
145				150				155				160				
ttc	ggt	aac	gcc	gcc	gcc	gct	gcg	gct	aat	gcg	agg	gag	gaa	gcg	aag	528
Phe	Gly	Asn	Ala	Ala	Ala	Ala	Ala	Ala	Asn	Ala	Arg	Glu	Glu	Ala	Lys	
165				170				175								
aag	tgg	tgc	ggc	agt	tgc	aag	aag	cgc	gtg	ggg	ctg	tta	ggg	ttc	aag	576
Lys	Trp	Cys	Gly	Ser	Cys	Lys	Lys	Arg	Val	Gly	Leu	Leu	Gly	Phe	Lys	
180				185				190								
tgc	agg	tgt	acg	aag	ttc	ttc	tgt	ggg	aag	cat	cgg	tat	cct	gag	gag	624
Cys	Arg	Cys	Thr	Lys	Phe	Phe	Cys	Gly	Lys	His	Arg	Tyr	Pro	Glu	Glu	
195				200				205								
cat	ggt	tgt	acg	ttc	gat	cat	gtg	gcg	ttc	ggg	agg	cgg	att	atc	gag	672
His	Gly	Cys	Thr	Phe	Asp	His	Val	Ala	Phe	Gly	Arg	Arg	Ile	Ile	Glu	
210				215				220								
aaa	cag	aat	cct	gtt	ctc	gag	acc	gac	aag	ctg	gtg	gac	aga	atc	tga	720
Lys	Gln	Asn	Pro	Val	Leu	Glu	Thr	Asp	Lys	Leu	Val	Asp	Arg	Ile		
225				230				235								
<210>26																

<211>239

<212>PRT

<213> 亚麻

<400>26

```

Met Ala Pro Ser Pro Cys Val His Gly Cys Thr Ala Asn Cys Pro Arg
1           5           10           15
Cys His Ser Tyr Gly His Pro Ile Phe Gly Asn Ser Asp Leu Ala Ala
          20           25           30
Gly Gly Ser Asp Thr Ser Thr Ser Val Phe Gly Lys Val Gly Ser Val
          35           40           45
Val Ile Gln Ser Pro Ala Lys Asn His Ala Phe Gly Gln Ala Cys Gly
          50           55           60
Pro Val Phe Pro Ser Ser Ser Ser Pro Phe Arg Arg Ile Lys Phe Gly
65           70           75           80
Pro Lys Asp Gly Glu Gly Lys Gly Pro Leu Lys Pro Ile Glu Lys Gln
          85           90           95
Pro Ser Lys Lys Arg Pro Phe Cys Phe Ser Pro Asp Glu Thr Ile Asp
          100          105          110
Ala Thr Val Pro Pro Ser Thr Lys Pro Phe Gly Ser Phe Arg Ser Val
          115          120          125
Cys Val Thr Asp Ala Asp Glu Ala Arg Leu Lys Ala Asn Arg Glu Phe
          130          135          140
Phe Ala Pro Val Ser Arg Lys Arg Gly Phe Asp Pro Thr Asp Met Thr
145          150          155          160
Phe Gly Asn Ala Ala Ala Ala Ala Asn Ala Arg Glu Glu Ala Lys
          165          170          175
Lys Trp Cys Gly Ser Cys Lys Lys Arg Val Gly Leu Leu Gly Phe Lys
          180          185          190
Cys Arg Cys Thr Lys Phe Phe Cys Gly Lys His Arg Tyr Pro Glu Glu
          195          200          205
His Gly Cys Thr Phe Asp His Val Ala Phe Gly Arg Arg Ile Ile Glu
          210          215          220
Lys Gln Asn Pro Val Leu Glu Thr Asp Lys Leu Val Asp Arg Ile
225          230          235

```

<210>27

<211>510

gtg aag gcg gag aag ctc gac aaa atc tga

510

Val Lys Ala Glu Lys Leu Asp Lys Ile

165

<210>28

<211>169

<212>PRT

<213> 亚麻

<400>28

Met Asp His Asp Glu Ala Gly Cys Gln Ala Pro Ser Asp His Pro Ile

1 5 10 15

Leu Cys Val Asn Asn Cys Gly Phe Phe Gly Ser Ala Ala Thr Met Asn

20 25 30

Met Cys Ser Lys Cys His Lys Asp Thr Met Leu Asn Gln Glu Gln Ser

35 40 45

Lys Leu Ala Ala Ser Ser Ala Ala Ser Ile Leu Asn Gly Ser Ser Met

50 55 60

Ser Leu Gly Arg Glu Leu Val Ile Ala Ala Lys Thr Asn Ser Val Glu

65 70 75 80

Pro Lys Thr Ile Ser Val Gln Pro Ser Ser Ala Ser Ser Ala Glu Glu

85 90 95

Ser Ile Glu Met Lys Leu Pro Lys Glu Gly Pro Ser Arg Cys Asn Thr

100 105 110

Cys Asn Lys Arg Val Gly Leu Thr Gly Phe Lys Cys Arg Cys Glu Asn

115 120 125

Met Phe Cys Ala Asn His Arg Tyr Ser Asp Lys His Asn Cys Pro Phe

130 135 140

Asp Tyr Arg Thr Ala Gly Arg Glu Ala Ile Ser Lys Ala Asn Pro Leu

145 150 155 160

Val Lys Ala Glu Lys Leu Asp Lys Ile

165

<210>29

<211>495

<212>DNA

<213> 稻

〈223〉 含有 A20 结构域与 AN1 结构域的锌指蛋白质 (OS39781852)

atc gtt agg ttc tga 495
Ile Val Arg Phe

<210>30

<211>164

<212>PRT

<213> 稻

<400>30

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

<210>31

<211>495

<212>DNA

<213> 稻

<220>

<221>CDS

<222>(1)..(495)

<223> 含有 A20 结构域与 AN1 结构域的锌指蛋白质 (OS34701560)

<400>31															
atg gcc gaa gaa cac cga tgc caa gct ccc gaa ggt cac aga ctc tgc															
Met Ala Glu Glu His Arg Cys Gln Ala Pro Glu Gly His Arg Leu Cys															
1					5					10					15
tcc aac aac tgc ggt ttc ttt ggt agc ccc gcc acc atg aat ctc tgt															
Ser Asn Asn Cys Gly Phe Phe Gly Ser Pro Ala Thr Met Asn Leu Cys															
					20					25					30
tcc aaa tgc tac aga gac atc cgt ttg aag gaa gaa gaa caa gcc aaa															
Ser Lys Cys Tyr Arg Asp Ile Arg Leu Lys Glu Glu Glu Gln Ala Lys															
					35					40					45
acc aaa tcc aca atc gaa acc gct ctt tca gga tct tcc tcc gcc acc															
Thr Lys Ser Thr Ile Glu Thr Ala Leu Ser Gly Ser Ser Ser Ala Thr															
					50					55					60
gtc acc gca acc gcc gtc gtt gcc tcc tcc gtg gaa tcc cct tcg gcg															
Val Thr Ala Thr Ala Val Val Ala Ser Ser Val Glu Ser Pro Ser Ala															
65					70					75					80
ccg gtt gaa tcc ctc cct caa cca ccg gtg ctg att tcg ccg gat ata															
Pro Val Glu Ser Leu Pro Gln Pro Pro Val Leu Ile Ser Pro Asp Ile															
					85					90					95
gcc gca ccg gtt cag gcg aac cgg tgc ggc gcg tgt agg aag cgc gtg															
Ala Ala Pro Val Gln Ala Asn Arg Cys Gly Ala Cys Arg Lys Arg Val															
					100					105					110
ggg ttg aca ggg ttc aag tgc agg tgc gga aca acg ttt tgt ggg agc															
Gly Leu Thr Gly Phe Lys Cys Arg Cys Gly Thr Thr Phe Cys Gly Ser															
					115					120					125
cac agg tac ccc gag aaa cac gcg tgt ggc ttc gat ttc aag gcg gtg															
His Arg Tyr Pro Glu Lys His Ala Cys Gly Phe Asp Phe Lys Ala Val															
					130					135					140
ggg aga gag gag ata gca cgg gcg aat ccc gtg atc aaa ggc gag aag															
Gly Arg Glu Glu Ile Ala Arg Ala Asn Pro Val Ile Lys Gly Glu Lys															
145					150					155					160
cta cgg agg att taa															
Leu Arg Arg Ile															

<210>32

<211>164

<212>PRT

<213> 稻

<400>32

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

<210>33

<211>513

<212>DNA

<213> 稻

<220>

<221>CDS

<222>(1)..(513)

<223> 含有 A20 结构域与 AN1 结构域的锌指蛋白质 (OS36821256)

<400>33

atg gcg cag agg gag aag aag gtg gag gag ccg acg gag ctg agg gcg 48
 Met Ala Gln Arg Glu Lys Lys Val Glu Glu Pro Thr Glu Leu Arg Ala
 1 5 10 15
 ccg gag atg acg ctc tgc gcc aac agc tgc ggg ttc ccg ggc aac ccg 96

Pro Glu Met Thr Leu Cys Ala Asn Ser Cys Gly Phe Pro Gly Asn Pro	
20 25 30	
gcg acc aac aac ctc tgc cag aac tgc ttc ttg gct gcc tcg gcg tct	144
Ala Thr Asn Asn Leu Cys Gln Asn Cys Phe Leu Ala Ala Ser Ala Ser	
35 40 45	
tct tct tct tct tcc gcc gct gcc tcg ccg tcg acg acg tcg ttg ccg	192
Ser Ser Ser Ser Ser Ala Ala Ala Ser Pro Ser Thr Thr Ser Leu Pro	
50 55 60	
gtg ttt ccg gtg gtg gag aag ccg agg cag gcc gta cag tcg tcg gcg	240
Val Phe Pro Val Val Glu Lys Pro Arg Gln Ala Val Gln Ser Ser Ala	
65 70 75 80	
gcg gcg gcg gtg gcg ctg gtg gtt gag cgg ccg acg gcg ggg ccg gtg	288
Ala Ala Ala Val Ala Leu Val Val Glu Arg Pro Thr Ala Gly Pro Val	
85 90 95	
gag tcg tcg tcg aag gcg tcg agg tcg tcg tcg gtc aac cga tgc cac	336
Glu Ser Ser Ser Lys Ala Ser Arg Ser Ser Ser Val Asn Arg Cys His	
100 105 110	
agc tgc cgg agg cgg gtg ggc ctg acc ggg ttc cgg tgc cgc tgc ggc	384
Ser Cys Arg Arg Arg Val Gly Leu Thr Gly Phe Arg Cys Arg Cys Gly	
115 120 125	
gag ctc tac tgc ggc gcg cac cgg tac tcc gac cgc cac gac tgc agc	432
Glu Leu Tyr Cys Gly Ala His Arg Tyr Ser Asp Arg His Asp Cys Ser	
130 135 140	
ttc gac tac aag tcg gcg gcg agg gac gcc atc gcc agg gag aac ccc	480
Phe Asp Tyr Lys Ser Ala Ala Arg Asp Ala Ile Ala Arg Glu Asn Pro	
145 150 155 160	
gtc gtc cgc gcc gcc aag atc gtt agg ttc taa	513
Val Val Arg Ala Ala Lys Ile Val Arg Phe	
165 170	

<210>34

<211>170

<212>PRT

<213> 稻

<400>34

Met Ala Gln Arg Glu Lys Lys Val Glu Glu Pro Thr Glu Leu Arg Ala
1 5 10 15

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

<210>35

<211>486

<212>DNA

<213> 大豆

<220>

<221>CDS

<222>(1)..(486)

<223> 含有 A20 结构域与 AN1 结构域的锌指蛋白质 (GM51659494)

<400>35

atg gct cag aaa acc gag aas gaa gaa acc gac ttc aaa gtt ccg gaa 48
 Met Ala Gln Lys Thr Glu Lys Glu Glu Thr Asp Phe Lys Val Pro Glu
 1 5 10 15
 acg att acg ctt tgc gtc aac aac tgc ggc gtc acc gga aac cct gcc 96
 Thr Ile Thr Leu Cys Val Asn Asn Cys Gly Val Thr Gly Asn Pro Ala
 20 25 30
 acg aat aac atg tgc cag aag tgc ttc act gcc tct acc gcc acc act 144

Thr Asn Asn Met Cys Gln Lys Cys Phe Thr Ala Ser Thr Ala Thr Thr	
35 40 45	
tcc ggc gcc gga ggt gcc gga ata gct tct ccg gcg acc aga tcc ggc	192
Ser Gly Ala Gly Gly Ala Gly Ile Ala Ser Pro Ala Thr Arg Ser Gly	
50 55 60	
gtc tcc gcg cgt cct cag aag aga tct ttt cct gaa gag ccc tcg ccg	240
Val Ser Ala Arg Pro Gln Lys Arg Ser Phe Pro Glu Glu Pro Ser Pro	
65 70 75 80	
gtg gcg gat cct cct tct tcg gac cag acg acg ccg tcg gag gcg aag	288
Val Ala Asp Pro Pro Ser Ser Asp Gln Thr Thr Pro Ser Glu Ala Lys	
85 90 95	
cgc gtg gtc aac cgc tgc tcc gga tgc cgg cgg aag gtc gga ctc acc	336
Arg Val Val Asn Arg Cys Ser Gly Cys Arg Arg Lys Val Gly Leu Thr	
100 105 110	
gga ttc cgg tgc cgg tgc ggc gag ctc ttc tgc gcc gag cac cgg tac	384
Gly Phe Arg Cys Arg Cys Gly Glu Leu Phe Cys Ala Glu His Arg Tyr	
115 120 125	
tcc gac cgc cac gac tgc agc tat gac tac aaa gcc gcc gga aga gaa	432
Ser Asp Arg His Asp Cys Ser Tyr Asp Tyr Lys Ala Ala Gly Arg Glu	
130 135 140	
gcc atc gcg agg gag aat ccg gtg atc aga gct gcg aag atc gtc aaa	480
Ala Ile Ala Arg Glu Asn Pro Val Ile Arg Ala Ala Lys Ile Val Lys	
145 150 155 160	
gtc tga	486
Val	

<210>36

<211>161

<212>PRT

<213> 大豆

<400>36

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

Ser Gly Ala Gly Gly Ala Gly Ile Ala Ser Pro Ala Thr Arg Ser Gly
 50 55 60
 Val Ser Ala Arg Pro Gln Lys Arg Ser Phe Pro Glu Glu Pro Ser Pro
 65 70 75 80
 Val Ala Asp Pro Pro Ser Ser Asp Gln Thr Thr Pro Ser Glu Ala Lys
 85 90 95
 Arg Val Val Asn Arg Cys Ser Gly Cys Arg Arg Lys Val Gly Leu Thr
 100 105 110
 Gly Phe Arg Cys Arg Cys Gly Glu Leu Phe Cys Ala Glu His Arg Tyr
 115 120 125
 Ser Asp Arg His Asp Cys Ser Tyr Asp Tyr Lys Ala Ala Gly Arg Glu
 130 135 140
 Ala Ile Ala Arg Glu Asn Pro Val Ile Arg Ala Ala Lys Ile Val Lys
 145 150 155 160
 Val

<210>37

<211>519

<212>DNA

<213> 大豆

<220>

<221>CDS

<222>(1).. (519)

<223> 含有 A20 结构域与 AN1 结构域的锌指蛋白质 (GM49780101)

<400>37

atg gag cct cat gat gag act gga tgc cag gct cct gaa cgc ccc att 48
 Met Glu Pro His Asp Glu Thr Gly Cys Gln Ala Pro Glu Arg Pro Ile
 1 5 10 15
 ctt tgc att aat aat tgt ggc ttc ttt gga aga gca gct acc atg aac 96
 Leu Cys Ile Asn Asn Cys Gly Phe Phe Gly Arg Ala Ala Thr Met Asn
 20 25 30
 atg tgt tcc aag tgt tac aag gac atg ctg ttg aag cag gag cag gac 144
 Met Cys Ser Lys Cys Tyr Lys Asp Met Leu Leu Lys Gln Glu Gln Asp
 35 40 45
 aaa ttt gca gca tca tcc gtt gaa aac att gtg aat ggc agt tcc aat 192
 Lys Phe Ala Ala Ser Ser Val Glu Asn Ile Val Asn Gly Ser Ser Asn
 50 55 60

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

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

<210>39

<211>525

<212>DNA

<213> 大豆

<220>

<221>CDS

<222>(1).. (525)

<223> 含有 A20 结构域与 AN1 结构域的锌指蛋白质 (GM59637305)

<400>39

atg gac cat gac aag act ggg tgc caa gct cct cct gaa ggt cct ata	48
Met Asp His Asp Lys Thr Gly Cys Gln Ala Pro Pro Glu Gly Pro Ile	
1 5 10 15	
ttg tgc atc aac aac tgt ggg ttt ttt gga agt gca gct acc atg aac	96
Leu Cys Ile Asn Asn Cys Gly Phe Phe Gly Ser Ala Ala Thr Met Asn	
20 25 30	
atg tgt tct aaa tgc cac aaa gac ata ttg ctg aaa cag gag cag gcc	144
Met Cys Ser Lys Cys His Lys Asp Ile Leu Leu Lys Gln Glu Gln Ala	
35 40 45	
aag ctt gca gca tca tcc att ggg aat att atg aat ggg tca tca agc	192
Lys Leu Ala Ala Ser Ser Ile Gly Asn Ile Met Asn Gly Ser Ser Ser	
50 55 60	
agc act gaa aag gaa cct gtt gtt gct gct gct gct aat att gat atc	240
Ser Thr Glu Lys Glu Pro Val Val Ala Ala Ala Ala Asn Ile Asp Ile	
65 70 75 80	

cca gtt att cca gta gag cct aaa act gtc tct gtg caa cct tta ttt	288
Pro Val Ile Pro Val Glu Pro Lys Thr Val Ser Val Gln Pro Leu Phe	
85 90 95	
ggt tca ggt cca gag ggg agt gtt gag gca aag ccg aag gat gga cca	336
Gly Ser Gly Pro Glu Gly Ser Val Glu Ala Lys Pro Lys Asp Gly Pro	
100 105 110	
aaa cgt tgc agc agc tgc aac aag cga gtt ggt ttg aca ggg ttt aat	384
Lys Arg Cys Ser Ser Cys Asn Lys Arg Val Gly Leu Thr Gly Phe Asn	
115 120 125	
tgt cga tgt ggt gac ctt ttt ttg tgc tgt aca tcg cta ctc gac aag	432
Cys Arg Cys Gly Asp Leu Phe Leu Cys Cys Thr Ser Leu Leu Asp Lys	
130 135 140	
cat aat tgc cca ttt gat tac cgc act gcc gct caa gat gct ata gct	480
His Asn Cys Pro Phe Asp Tyr Arg Thr Ala Ala Gln Asp Ala Ile Ala	
145 150 155 160	
aaa gca aac cca gtt gtc aag gct gaa aag ctt gat aag atc taa	525
Lys Ala Asn Pro Val Val Lys Ala Glu Lys Leu Asp Lys Ile	
165 170	

<210>40

<211>174

<212>PRT

<213> 大豆

<400>40

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

100	105	110
Lys Arg Cys Ser Ser Cys Asn Lys Arg Val Gly Leu Thr Gly Phe Asn		
115	120	125
Cys Arg Cys Gly Asp Leu Phe Leu Cys Cys Thr Ser Leu Leu Asp Lys		
130	135	140
His Asn Cys Pro Phe Asp Tyr Arg Thr Ala Ala Gln Asp Ala Ile Ala		
145	150	155
Lys Ala Asn Pro Val Val Lys Ala Glu Lys Leu Asp Lys Ile		
165	170	

<210>41

<211>498

<212>DNA

<213> 小麦

<220>

<221>CDS

<222>(1)..(498)

<223> 含有 A20 结构域与 AN1 结构域的锌指蛋白质 (TA55974113)

<400>41

atg gcg cag cgg gat cac aag cag gag gag ccc acg gag ctg cgg gcg	48
Met Ala Gln Arg Asp His Lys Gln Glu Glu Pro Thr Glu Leu Arg Ala	
1 5 10 15	
ccg gag atc acg ctc tgc gcc aac agc tgc ggc ttc ccg ggc aac ccg	96
Pro Glu Ile Thr Leu Cys Ala Asn Ser Cys Gly Phe Pro Gly Asn Pro	
20 25 30	
gcc acg cag aac ctc tgc cag aac tgc ttc ttg gcc ggc ccg gcg tcc	144
Ala Thr Gln Asn Leu Cys Gln Asn Cys Phe Leu Ala Gly Pro Ala Ser	
35 40 45	
acg tcg ccg tct tcc tcc tcc tcc tcc tcc tct tct ctg ccg ggc gtg	192
Thr Ser Pro Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Leu Pro Gly Val	
50 55 60	
tcc gcg ccg acc ccc gtc atc gac agg ccg agg ccg gcg ccg ttg gag	240
Ser Ala Pro Thr Pro Val Ile Asp Arg Pro Arg Pro Ala Pro Leu Glu	
65 70 75 80	
gcg gag ctg gca cgc ccc gcc gtc gac ctt gct ccg gcg acg gag gcg	288
Ala Glu Leu Ala Arg Pro Ala Val Asp Leu Ala Pro Ala Thr Glu Ala	
85 90 95	

aag ccg gcg agg acg tcg gtg aac cgg tgc tcc agc tgc cgg aag cgc	336
Lys Pro Ala Arg Thr Ser Val Asn Arg Cys Ser Ser Cys Arg Lys Arg	
100 105 110	
gtg ggg ctg acg ggg ttc cgg tgc cgg tgc ggc gac atg ttc tgc ggc	384
Val Gly Leu Thr Gly Phe Arg Cys Arg Cys Gly Asp Met Phe Cys Gly	
115 120 125	
gag cac cgg tac tcg gac cgg cac ggg tgc agc tac gac tac aag gcc	432
Glu His Arg Tyr Ser Asp Arg His Gly Cys Ser Tyr Asp Tyr Lys Ala	
130 135 140	
gcc gcc agg gac gcc atc gcc agg gac aac ccc gtc gtg cgc gcc gcc	480
Ala Ala Arg Asp Ala Ile Ala Arg Asp Asn Pro Val Val Arg Ala Ala	
145 150 155 160	
aag atc gtc agg ttc tga	498
Lys Ile Val Arg Phe	
165	

<210>42

<211>165

<212>PRT

<213> 小麦

<400>42

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

Glu His Arg Tyr Ser Asp Arg His Gly Cys Ser Tyr Asp Tyr Lys Ala
 130 135 140
 Ala Ala Arg Asp Ala Ile Ala Arg Asp Asn Pro Val Val Arg Ala Ala
 145 150 155 160
 Lys Ile Val Arg Phe
 165

<210>43

<211>207

<212>PRT

<213> 小立碗藓

<220>

<221>MISC_FEATURE

<222>(1).. (207)

<223> 甲硫氨酸亚砷还原酶家族蛋白质 (EST65)

<400>43

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

	165	170	175
Val Phe Asp Asp Gly Pro Arg Pro Thr Gly Lys Arg Tyr Cys Ile Asn			
	180	185	190
Ser Ala Ala Ile Asp Leu Lys Ala Glu Lys Gln Glu Glu Arg Asn			
	195	200	205

<210>44

<211>339

<212>PRT

<213> 小立碗藓

<220>

<221>MISC_FEATURE

<222>(1).. (339)

<223> 同源异型域 - 亮氨酸拉链蛋白质 (EST12)

<400>44

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

	165		170		175										
Leu	Lys	Ala	Glu	Val	Ser	Arg	Leu	Lys	Gly	Ile	Ser	Asn	Asp	Asp	Val
	180		185		190										
Lys	Pro	Ala	Glu	Phe	Val	Gln	Gly	Lys	Cys	Asp	Thr	Thr	Ser	His	Pro
	195		200		205										
Ala	Ser	Pro	Ala	Gln	Ser	Glu	Arg	Ser	Asp	Ile	Val	Ser	Ser	Arg	Asn
	210		215		220										
Arg	Thr	Thr	Pro	Thr	Ile	His	Val	Asp	Pro	Val	Ala	Pro	Glu	Glu	Ala
225			230		235										
Gly	Ala	His	Leu	Thr	Met	Ser	Ser	Asp	Ser	Asn	Ser	Ser	Glu	Val	Met
	245		250		255										
Asp	Ala	Asp	Ser	Pro	Arg	Thr	Ser	His	Thr	Ser	Ala	Ser	Arg	Ser	Thr
	260		265		270										
Leu	Ser	Thr	Ser	Val	Val	Gln	Pro	Asp	Glu	Gly	Leu	Gly	Val	Ala	Gln
	275		280		285										
Tyr	Pro	His	Phe	Ser	Pro	Glu	Asn	Phe	Val	Gly	Pro	Asn	Met	Pro	Glu
	290		295		300										
Ile	Cys	Ala	Asp	Gln	Ser	Leu	Ala	Ser	Gln	Val	Lys	Leu	Glu	Glu	Ile
305			310		315										
His	Ser	Phe	Asn	Pro	Asp	Gln	Thr	Phe	Leu	Leu	Leu	Pro	Asn	Trp	Trp
	325		330		335										
Asp	Trp	Ala													

<210>45

<211>188

<212>PRT

<213> 小立碗藓

<220>

<221>MISC_FEATURE

<222>(1).. (188)

<223> 含有 A20 结构域与 AN1 结构域的锌指蛋白质 (EST307)

<400>45

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

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

SEQ ID NO:2	(1)	-----MASSSCFTIQSRFVSARTKLDISIKP-----SLSGFACRSLTKPRNLNLSVLL-----RCSMGSFNSSQ
SEQ ID NO:43	(1)	-----MVAESVLVCRSSVVGAGLQSFVGEAKRESAGPGRSVFLGAQVQKMGAGMSA-----RSDVRPAAVPK
SEQ ID NO:8	(1)	-----MGLSILRSTISITPISSSKSKPIFSTLVRSSFASISPTKCVTPITL FVSATP-----FFTASPKRGFR
SEQ ID NO:6	(1)	MAMRQYAAATAASSFRARPRARPSCLPAAALPLAPCCGVASRASRYRRASVRAMGAASS-----SSSSSSSPSPQ
SEQ ID NO:4	(1)	-----MGFNILRTSISITPISSSKSKPIFSTLRRSSPSTIFPPKSVTPITL FVSATP-----FFTLPKLGFR
SEQ ID NO:10	(1)	-----MASPHAHPATRPLSLPS-----LLLARSSSAATAAASSARPPASLSLSCSRRARAYCPAGRRLP
SEQ ID NO:2	(60)	KSDNVQEAAS-D-FASISEGEWKKRLTPEQYIITRQKGTERTFTGEYWNTPGVYKICCCDTPLFDSSTKLDS
SEQ ID NO:43	(64)	ASGDVSEQTD-----YKTFSDDEWKKRLSQOQFYVARKKGTERTFTGEYWNTPKTAGTYLCVCCKTPLESSKTKFDS
SEQ ID NO:8	(64)	GGIVAMAAAGSLR-----KSEEWRAVLSPQFRILRQKGTETFPGTGEYDKFFDEGVYNCAGCGTPLYRSLTKFNS
SEQ ID NO:6	(73)	GQAQAOAQGKPNY-STSLTDEEWKRLTKDQYIITRQKGTERTFTGEYWNTPGIYHCVCCDTPLFESSTKFDS
SEQ ID NO:4	(64)	GGIVAMAAPGSLR-----KSEEWRAILSPQFRILRQKGTETFPGTGEYDKFYEEGVYNCAGCGTPLYRSITKFNS
SEQ ID NO:10	(61)	GAVVAMSSAAPTGPVQKSEEEWEAVLTPEQFRILRRKGTETPGTGEYDKFFSEGIYGCAGCGTPLYKSSTKFNS
SEQ ID NO:2	(133)	GTGWPSYYQPIGNVKSCLDLSIFMPRQEVICAVCNAHLGHVFDDG--PRPTGKRYCINSAALKLESLETRTRE--
SEQ ID NO:43	(135)	GTGWPSYYDTIGDNVKSMDWSIPFMPRTTEVVCACVDAHLGHVFDDG--PRPTGKRYCINSAAILDKAEKQEEERN
SEQ ID NO:8	(135)	GCGWPAFYEGIPGAINRNPDP---DGMRTETCAACGGHLGHVFKGEGFPTPTNERHCVNSISLKFAPANS-----
SEQ ID NO:6	(147)	GTGWPSYYQPIGDNVCKCLDMSIFMPRTEVLCAVCDAHLGHVFDDG--PRPTGKRYCINSAALKKKTQ-----
SEQ ID NO:4	(135)	GCGWPAFYEGIPGAINRNPDP---DGMRTETCAACGGHLGHVFKGEGFPTPTNERHCVNSISLKFAPANSYS--
SEQ ID NO:10	(136)	GCGWPAFYEGFPGAIKRTADP---DGRRIETCAACEGHLGHVEKGEFNTPTDERHCVNSISLKFVPASEEAS-

图 1

SEQ ID NO:12	(1)	-----MMKRLSSDSVGGGLISLCPTTSTDQPNPR-----RCGREFQSMLEGY
SEQ ID NO:44	(1)	MVVPSPAFGGQNAMLRNIDNNTDTLISLQSCSPRVSMQQVPRSESLENMMG---ACQKLPYFSSFDGPS
SEQ ID NO:14	(1)	-----MAGSGSAFNSITS---FLRTQPPSSQPLDSSLFLSAPSSAPFLGSRSMMSFDGEGKGCNGSGFFRA
SEQ ID NO:16	(1)	-----MEPGRLLIFNTSGSGNGQMLFMDCCGAGGIAGAAGMFHRGVRPVLGME-----E-GRGVKRPF
SEQ ID NO:12	(43)	EEEEEAITEERGQTGLAEKKRRRLNINQVKALEKNFELENKLEPERKVKLAQELGLQPRQVAVWFQNRARWTK
SEQ ID NO:44	(73)	VEEQEDVDEGIDEFAHVEKKRRRLSLEQVRSLEKNFEVENKLEPERKMQLAKELGLRPRQVAVWFQNRARWTK
SEQ ID NO:14	(64)	FDMDNGDECMDEYFHQPEKKRRRLSASQVQFLEKSFEEENKLEPERKTKLAKDLGLQPRQVAIWQNRARWKNK
SEQ ID NO:16	(57)	FTSPDDMLEEYYDEQLPEKKRRRLTPEQVHLLERSFEEENKLEPERKTELRKLGQPRQVAVWFQNRARWTK
SEQ ID NO:12	(118)	QLEKDYGVLTQYDSLRHNEDSLRRENESLLOEIGKLLKAKLNGE--EEGDDVDEEENNLATMESDVSVKEEE---V
SEQ ID NO:44	(148)	QLEHDIETLKKAYDRLKADFEAVTLDTNALKAEVSRLLKGISNDDVKPAEFFVQCKCDTTSHPASPAQSERSD---I
SEQ ID NO:14	(139)	QLEKDIETLHASFESILKSNYDCLLKEKDKLKAEVASLTEKVLARGKQEGHMKQAESESEETKGLLHLQEQEPPQR
SEQ ID NO:16	(132)	TLERDFDRLKASFDALRADHDALLQDNHRLRSQVVTLTETKMQDKEAPEGSFGAAADASEPEQAAAAEAKASLADAE
SEQ ID NO:12	(189)	SLPEQITEPPSPPPQLLEHSDSFNYSFTDLRLDILLPLKAAASSVAAAGSSDSSSAVLNEESSNNVTAAATVP
SEQ ID NO:44	(220)	VSSRNRTPTTIHVDPPVAPEEAGAHLTMSDDSNSEVMADSPRTSHTSASRSTLSTSVVQPDGLGVAQYPHFSP
SEQ ID NO:14	(214)	LLLQSVSEGEKSVSSVVGCKQEDISSARSDILDSDPHYTDGVHSALLEHGDSSYVFEPDQSDMSQDEEDNLS
SEQ ID NO:16	(207)	EQAAAAEAFEVVQQQLHVKDEERLSPGSGGSAVLDAARDALLGSGCGLAGVVDSSVDSYCFPGGAGGDEYHECVVG
SEQ ID NO:12	(264)	--GGSELFQVKMEQTEHDHDFLSGEEACGFFSDEQPPSLHWYSTVDQWN-----
SEQ ID NO:44	(295)	---ENFVGPNMPEICADQSLASQVKLEEIHSFNPDQTFLLLPNWWDA-----
SEQ ID NO:14	(289)	--KSLYPSYLFPKLEEDVDYSDPPSESCNFGFPEEDHVLWTWAY-----
SEQ ID NO:16	(282)	PVAGGIQSEEDDGAGSDEGCSYYPDDAAVFFAAQAQGHGHRTDDDDQDDGQISYMMWN-

图 2

SEQ ID NO:18	-----MDHDKTGCSPPGPKLCINNCGFFGSAATMNMCSKCHKAILFQEQGARFA--	(1)
SEQ ID NO:20	-----MAEEHRCQTP-EGHRLCANNCGFLGSSATMNLCSNCGDLCLKQ-QASMK--	(1)
SEQ ID NO:34	-----MAEEHRCQAP-EGHRLCSNCGFFGSPATMNLCSKCYRDIRLKEEQAKTK--	(1)
SEQ ID NO:38	-----MAQTEKEETDFKVP-ETITLCVNNCGVTGNPATNMCKQCFSTASTATSGAGAG--	(1)
SEQ ID NO:36	-----MAQREKKVEEPTLRAPEITLCA NSCGFPNATNLQNLCONCFLAASASS--SSS--	(1)
SEQ ID NO:32	-----MAQRDKDQEPTELRAPEITLCA NSCGFPNATNLQNLCONCFLAATASTSSPSSL--	(1)
SEQ ID NO:44	-----MAQRDHQEEPTLRAPEITLCA NSCGFPNATNLQNLCONCFLAGPASTS-PSSSS--	(1)
SEQ ID NO:22	-----MAQESCDLNKDEAEILKPSSTPSPPSPATPPPTA--	(1)
SEQ ID NO:28	MAPSPCVHGCTANCPCHSYGHP-IFGNSDLAAGSDTSTVFGKVGSVIQSPAKNHAFCQACGPFVFPSSSPFR	(1)
SEQ ID NO:45	-----MATERSVQETTSQAP-EGPVMCKNLGFFGSAATMGLCSKCYRETVMQAKMTALAE--	(1)
SEQ ID NO:42	-----MDHDKTGCOAPP-EGPILCINNCGFFGSAATMNMCSKCHKIDILLKQEQAKLAA--	(1)
SEQ ID NO:30	-----MDHDEAGCQAPSDHPILCVNCGFFGSAATMNMCSKCHKIDMLLNOEQSKLAA--	(1)
SEQ ID NO:40	-----MEPHDETGCQAP-ERPILCINNCGFFGSAATMNMCSKCHKIDMLLNOEQDKFAA--	(1)
SEQ ID NO:26	-----MEHKEAGCQAP-EGPILCINNCGFFGSAATMNMCSKCHKEMITKQDQAKLAA--	(1)
SEQ ID NO:24	-----MEHKEAGCQAP-EGPILCINNCGFFGSAATMNMCSKCHKEMITKQEQAKLAA--	(1)
SEQ ID NO:18	-----SAVSGGTS-----SSSNILKE-----TFAATALVDAETKS-----VEPVAHSV	(53)
SEQ ID NO:20	-----STVESLSAV-----SPSSEIGSMQSTVESLSDVSPSPETISISSPMIQPLV	(50)
SEQ ID NO:34	-----STIETALSGS-----S-----SATVTATAVAVASSVESAP-----VESLP	(51)
SEQ ID NO:38	-----IASPATRSG-----VSARPKRSFP-----EESPVPAD	(56)
SEQ ID NO:36	-----AAASPSTS-----LP-----VFPVVEKPRQ-----AVQSSAAA	(54)
SEQ ID NO:32	-----SPVLDKQPP-----RP-----AAPLVEPQAP-----LPPVVEEM	(57)
SEQ ID NO:44	-----SSSSSLPGV-----SA-----PTPVIDRPRP-----APLEAELA	(56)
SEQ ID NO:22	-----QIPEQPP-----HSPFPQAPAAQFLSRPCEVPIET	(37)
SEQ ID NO:28	RIKFGPKDGEKGLPIEQPSKKRPFCSFDETIDATVPPSTKPFSGFRSVCTDADAEARKANREFFVPSR	(76)
SEQ ID NO:45	-----QATQAAQATS-----ATAAAVQPPAPVHETKLTCEVERTMIV-----PHOSSSYQ	(56)
SEQ ID NO:42	-----SSIGNIMNGS-----SSSTEKEP-----VWAAANIDIPVIP-----VEPKTVSV	(53)
SEQ ID NO:30	-----SSAASIILNGS-----SMSLIGREL-----VIAAKTNS-----VEPKTISV	(53)
SEQ ID NO:40	-----SSVENIVNGS-----SNGNGKQ-----AVATGAVAVQVEA-----VEVKIVCA	(53)
SEQ ID NO:26	-----SSIDSIVNG-----SDAVME-----PVVAGSNTVWVAQ-----VELOTMNV	(52)
SEQ ID NO:24	-----SPIDSIVNG-----GDGKG-----PVIAAS-VNVAVPQ-----VEQKTI-V	(52)
SEQ ID NO:18	QPSSVQV-AAEVVAPEAAAKLKEGSPRCATCNKRVLGTGFKCRCGLFCGTHRYADIHNCSFNHYHAAQAEIAK	(91)
SEQ ID NO:20	RNPSAELEVTATKTVTPEEQOQKRNRCCTCRKRVGLTGFKCRCGTTFCGAHRYPEVHGCTDFKSAAGREEIAK	(100)
SEQ ID NO:34	-----VLISPDIAAPVQANRCGACRKRVLGTGFKCRCGTTFCGSHRYPEKHACGDFDKAVGREEIAR	(87)
SEQ ID NO:38	-----SDQTPSEAKRVNRCSCGRRKVLGTGFKCRCGELFCAEHRYSDRHDCSYDYKKAAGREAIAR	(84)
SEQ ID NO:36	-----AVALVERPTAGPVESSSKASRSSVNRCSSCRKRVLGTGFKCRCGELFCEGEHRYSDRHDCSYDYKKAAGREAIAR	(83)
SEQ ID NO:32	-----ASALATAPAPVAKTS-----AVNRCSSCRKRVLGTGFKCRCGELFCEGEHRYSDRHDCSYDYKKAAGREAIAR	(86)
SEQ ID NO:44	-----RPAVDLAPATEAKPART-----SVNRCSSCRKRVLGTGFKCRCGELFCEGEHRYSDRHDCSYDYKKAAGREAIAR	(85)
SEQ ID NO:22	-----SKKR-KHADAVSMAIVEPLSSVLFVNRCNVCRKRVLGTGFKRCCEKLFCEPRHSHSEHDCSFDYKTVGREEIAR	(88)
SEQ ID NO:28	-----KRGFDPDTMTFGNAAAANAAREAAKWCSCKKRVGLLGFKRCCTKFFCGKHRYPEEHGCTFDHVAFGRRRIEK	(151)
SEQ ID NO:45	-----QDLVTPAAAAPQAVKSSIAAPSRPEPNRCGSKRVGLTGFKCRCGNLYCALHRYSDKHTTCTDYKKAAGQAEIAK	(101)
SEQ ID NO:42	-----QPL-----FGSGPEGSEAKP-KDGPKRCSNCRKRVGLTGFCNRCGDLFLCCTSLLDKHNCFDYRTAAQDAIAK	(93)

图 3

SEQ ID NO: 30	(87)	QPS-----SASSAEESIEMKLPKEGSPSRCNTCNKRVGLTGFKRCENMFCAHRYSDKHNCPEFYRTAGREAIISK
SEQ ID NO: 40	(91)	QSS-----VDSSSGDSLEMKAKTGPSRCATCRKRVGLTGFSCKCGNLFCAMHRYSDKHDCPEFYRTVGQDAIAK
SEQ ID NO: 26	(89)	QQPADVAGPSEGVAAI SKGGK--VGNRC SACRKRVLGTGFNCRCGNLYCALHRYSDKHDCKFYRTAARDAIAK
SEQ ID NO: 24	(87)	VQPM LVAETSEAAAVI PKAK---EGPDRCAACRKRVLGTGFSCRCGNMYCSVHRYSDKHDCQFDYRTAARDAIAK
SEQ ID NO: 18	(165)	ANPVVKA EKLDKI---
SEQ ID NO: 20	(175)	ANPLVKA AKLQKI---
SEQ ID NO: 34	(152)	ANPVIKGEKL RRI---
SEQ ID NO: 38	(149)	ENPVIRAAKIVKV---
SEQ ID NO: 36	(158)	ENPVVRAAKIVRF---
SEQ ID NO: 32	(152)	DNPVVRAAKIVRF---
SEQ ID NO: 44	(153)	DNPVVRAAKIVRF---
SEQ ID NO: 22	(142)	ANPLIRAAKII RI---
SEQ ID NO: 28	(226)	QNPVLETDKLV DRI---
SEQ ID NO: 45	(176)	ANPLVVAEKVVKF---
SEQ ID NO: 42	(162)	ANPVVKA EKLDKI---
SEQ ID NO: 30	(157)	ANPLVKA EKLDKI---
SEQ ID NO: 40	(160)	ANPIIKADKLDKI---
SEQ ID NO: 26	(162)	ANPVVKADKLDKI---
SEQ ID NO: 24	(159)	ANPVVRA EKLDKI---

图 3 续