



(12)发明专利申请

(10)申请公布号 CN 107920489 A

(43)申请公布日 2018.04.17

(21)申请号 201680050837.4

(22)申请日 2016.09.06

(30)优先权数据

15184528.6 2015.09.09 EP

(85)PCT国际申请进入国家阶段日

2018.03.02

(86)PCT国际申请的申请数据

PCT/EP2016/070972 2016.09.06

(87)PCT国际申请的公布数据

W02017/042162 EN 2017.03.16

(71)申请人 菲利普莫里斯生产公司

地址 瑞士纳沙泰尔

(72)发明人 L·博维特

(74)专利代理机构 中国国际贸易促进委员会专利商标事务所 11038

代理人 傅宇昌

(51)Int.Cl.

A01H 5/12(2018.01)

C12N 9/00(2006.01)

C12N 15/82(2006.01)

权利要求书3页 说明书86页

序列表45页 附图12页

(54)发明名称

天冬酰胺含量减少的植物

(57)摘要

本文描述了具有降低的天冬酰胺合成酶表达或活性的突变型、非天然存在的或转基因植物或其部分,所述天冬酰胺合成酶包含以下、由以下组成或基本上由以下组成:(i)聚核苷酸序列,包含以下、由以下组成或基本上由以下组成:与SEQ ID NO:1具有至少90%序列一致性或与SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7具有至少72%序列一致性的序列;或(ii)由(i)中所示的聚核苷酸编码的多肽;或(iii)与SEQ ID NO:2或SEQ ID NO:4或SEQ ID NO:6或SEQ ID NO:8具有至少78%序列一致性的多肽;其中(i)、(ii)或(iii)中所示的天冬酰胺合成酶的表达或活性与对照植物相比降低。

1. 一种天冬酰胺合成酶的表达或活性降低的突变型、非天然存在的或转基因植物或其部分,所述天冬酰胺合成酶包含以下、由以下组成或基本上由以下组成:

(i) 包含以下、由以下组成或基本上由以下组成的聚核苷酸序列:与SEQ ID NO:1具有至少90%序列一致性或与SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7具有至少72%序列一致性的序列;或

(ii) 由(i)中所示的聚核苷酸编码的多肽;或

(iii) 与SEQ ID NO:2或SEQ ID NO:4或SEQ ID NO:6或SEQ ID NO:8具有至少78%序列一致性的多肽;

其中(i)、(ii)或(iii)中所示的所述天冬酰胺合成酶的表达或活性与对照植物相比降低。

2. 根据权利要求1所述的突变型、非天然存在的或转基因植物或其部分,其中所述植物在编码所述天冬酰胺合成酶的聚核苷酸的调节区或编码序列中包含至少一个基因改变。

3. 根据权利要求1或权利要求2所述的突变型、非天然存在的或转基因植物或其部分,其中所述天冬酰胺合成酶的降低的表达或活性赋予源自所述突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶中天冬酰胺的水平与源自所述对照植物的经烘烤或干燥的叶中天冬酰胺的水平相比降低,合适地,其中所述天冬酰胺的水平与所述对照植物相比降低了至少约17%;和/或

其中所述天冬酰胺合成酶的降低的表达或活性赋予在加热或燃烧源自所述突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶时所产生的气溶胶中丙烯酰胺的水平与源自所述对照植物的经烘烤或干燥的叶相比降低,合适地,其中所述气溶胶中丙烯酰胺的水平与所述对照植物的经烘烤或干燥的叶相比降低了至少20%;和/或

其中与所述对照植物的经烘烤或干燥的叶和气溶胶相比,源自所述突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶中天冬酰胺的水平降低至少约22%并且在加热或燃烧源自所述突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶时所产生的气溶胶中丙烯酰胺的量降低至少24%;合适地,其中与所述对照植物的经烘烤或干燥的叶和气溶胶相比,源自所述突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶中天冬酰胺的水平降低至少约44%并且在加热或燃烧源自所述突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶时所产生的气溶胶中丙烯酰胺的量降低至少66%;合适地,其中与所述对照植物的经烘烤或干燥的叶和气溶胶相比,源自所述突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶中天冬酰胺的水平降低至少约70%并且在加热或燃烧源自所述突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶时所产生的气溶胶中丙烯酰胺的量降低至少约88%。

4. 根据前述权利要求中任一项所述的突变型、非天然存在的或转基因植物或其部分,其中来自所述突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶中尼古丁的水平与所述对照植物的经烘烤或干燥的叶中尼古丁的水平基本上相同;和/或

其中源自所述突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶中谷氨酰胺、天冬氨酸和谷氨酸的形成与所述对照植物的经烘烤或干燥的叶相比增加;和/或

其中所述植物选自由烟草或茶组成的群组,优选地,

其中所述烟草植物是普通烟草植物,优选地,白肋型烟草植物。

5. 根据前述权利要求中任一项所述的突变型、非天然存在的或转基因植物或其部分, 其中所述植物在(i)、(ii)或(iii)中所示的编码所述天冬酰胺合成酶的聚核苷酸序列的每个拷贝中包含至少一个突变, 合适地, 干扰编码所述天冬酰胺合成酶的RNA转录物的翻译的终止突变和/或基因片段。

6. 根据前述权利要求中任一项所述的突变型、非天然存在的或转基因植物或其部分, 其中所述天冬酰胺合成酶核苷酸序列包含在与SEQ ID No. 19、21和23中编码终止突变的核苷酸序列的位置等同的位置处编码所述终止突变的核苷酸序列, 或其中所述天冬酰胺合成酶包含在与SEQ ID No. 20、22和24中终止密码子的位置等同的位置处编码所述终止密码子的核苷酸序列。

7. 根据前述权利要求中任一项所述的突变型、非天然存在的或转基因植物或其部分, 其中所述天冬酰胺合成酶包含选自SEQ ID No. 19、21和23组成的群组的核苷酸序列, 或由选自SEQ ID No. 19、21和23组成的群组的所述核苷酸序列编码, 或包含以下、由以下组成或基本上由以下组成: SEQ ID No. 20、22和24中所示的多肽序列; 和/或

其中来自所述突变型、非天然存在的或转基因植物或其部分植物的叶生物质的量与来自所述对照植物的叶生物质的量基本上相同。

8. 根据权利要求3至7中任一项所述的突变型、非天然存在的或转基因植物或其部分, 其中所述叶是经晾制的, 合适地, 其中所述经晾制的叶是经日光烘烤或明火烘烤的; 或其中所述叶是经风干的, 合适地, 其中所述经风干的叶是经日光干燥或经明火干燥的。

9. 一种植物材料或经烘烤或干燥的植物材料, 其来自根据权利要求1至8中任一项所述的植物。

10. 一种植物产品, 其包含根据权利要求1至8中任一项所述的植物的至少一部分或根据权利要求9所述的植物材料。

11. 一种制备植物材料的方法, 源自所述植物材料的气溶胶中具有与来自对照植物的植物材料相比降低的天冬酰胺水平和降低的丙烯酰胺水平, 所述方法包含以下步骤:

(a) 提供包含如下聚核苷酸的植物或其部分, 所述聚核苷酸包含以下、由以下组成或基本上由以下组成: 编码天冬酰胺合成酶并且与SEQ ID NO: 1具有至少90%序列一致性或与SEQ ID NO: 3或SEQ ID NO: 5或SEQ ID NO: 7具有至少72%序列一致性的序列;

(b) 降低所述植物或其部分中所述聚核苷酸的表达或由其编码的蛋白质的活性;

(c) 从所述植物或其部分收获植物材料;

(d) 干燥或烘烤所述植物材料;

(e) 任选地, 测量所述植物或其部分中天冬酰胺的水平和/或测量源自所述植物或其部分的气溶胶中丙烯酰胺的水平; 和

(f) 获得经烘烤或干燥的植物材料, 源自所述植物材料的气溶胶中具有降低的天冬酰胺水平和降低的丙烯酰胺水平, 合适地, 其中尼古丁的水平与所述对照植物中的尼古丁水平基本上相同; 合适地, 其中谷氨酰胺、天冬氨酸和谷氨酸的形成与所述对照植物的经烘烤或干燥的叶相比增加。

12. 根据权利要求11所述的方法, 其中所述植物材料在收获后烘烤或至少约3天; 和/或

其中所述植物材料是经晾制的, 合适地, 其中所述经晾制的叶是经日光烘烤的或经明火烘烤的, 或其中所述植物材料是经风干的或经日光干燥的或经明火干燥的。

13. 一种植物材料,其是通过或能够通过根据权利要求11或权利要求12所述的方法获得的。

14. 一种用于产生植物材料的方法,源自经烘烤或干燥的植物材料的气溶胶中具有与对照植物相比降低的天冬酰胺水平和降低的丙烯酰胺水平,所述方法包含以下步骤:

(a) 提供根据权利要求1至8中任一项所述的植物或根据权利要求9所述的植物材料;

(b) 从所述植物收获植物材料;

(c) 烘烤或干燥所述植物材料一段时间;和

(d) 获得经烘烤或干燥的植物材料,源自所述经烘烤或干燥的植物材料的气溶胶中具有降低的天冬酰胺水平和降低的丙烯酰胺水平,合适地,其中所述尼古丁的水平与所述对照植物中尼古丁的水平基本上相同;合适地,其中所述谷氨酰胺、天冬氨酸和谷氨酸的形成与所述对照植物相比增加。

15. 一种烟草产品或一种吸烟制品,其包含根据权利要求9或权利要求13所述的植物材料。

天冬酰胺含量减少的植物

技术领域

[0001] 本发明公开了编码来自烟草 (*Nicotiana tabacum*) 的天冬酰胺合成酶的基因的聚核苷酸序列及其变体、同源物和片段。还公开了由其编码的多肽序列及其变体、同源物和片段。公开了一种或多种基因的表达或由其编码的蛋白质活性的修饰,以调节植物中天冬酰胺的水平。

背景技术

[0002] 丙烯酰胺是式 C_3H_5NO (IUPAC名称是丙-2-烯酰胺)的化合物,关于其潜在毒性已经引起关注。包含烟草的吸烟制品的气溶胶中的丙烯酰胺的来源可能是氨基酸天冬酰胺,其可能存在于用于制造所述吸烟制品的烟草材料中。因此,可能需要降低植物中某些氨基酸(如天冬酰胺)的水平,特别是在用于烟草生产的那些植物中。

[0003] US2013/0068240描述了经遗传修饰的烟草植物和烟草产品,其表现出至少一种氨基酸的减少,使得在加热和/或燃烧烟草时,与未修饰的亲本烟草植物相比,产生水平降低的衍生自所述至少一种氨基酸的化合物。描述了用甲磺酸乙酯 (EMS) 诱变种子以获得香烟烟雾中具有降低的天冬酰胺水平和降低的丙烯酰胺水平的突变烟草品系。多种EMS突变体描述于US2013/0068240中。在烤烟上部叶中,三种EMS突变体,即FC Up 10NH-5、FC Up 10NH-18和FC-Up 10NH-23显示:(i) 天冬酰胺分别降低71%、30%和62%;(ii) 烟草烟雾中的丙烯酰胺分别降低47%、42%和44%;和(iii) 烤烟上部烟草中的尼古丁分别降低38%、增加111%和降低19%。因此,突变体FC Up 10NH-5分别使天冬酰胺和丙烯酰胺的最高水平降低了71%和47%。然而,在这个突变体中,与对照相比,尼古丁水平显示出38%的降低。在烤烟中部叶中,三种EMS突变体,即FC Mid10NH-5、FC Up 10NH-18和FC-Up 10NH-23显示:(i) 天冬酰胺分别降低6%、增加531%和降低29%;(ii) 烟草烟雾中的丙烯酰胺分别降低19%、增加140%和增加200%;和(iii) 烤烟中部烟草中的尼古丁分别降低36%、增加113%和降低19%,因此,FC Up 10NH-5使丙烯酰胺的最高水平降低了19%,但是天冬酰胺的水平仅降低了6%并且尼古丁的水平降低了36%。

[0004] 在FC Up 10NH-5的烤烟上部叶中发现天冬酰胺和丙烯酰胺的最高水平分别降低了71%和47%。然而,在这个突变体中,与对照相比,尼古丁水平显示出38%的降低。在FC Up 10NH-5的烤烟中部叶中,与对照相比,丙烯酰胺的水平仅降低了19%,天冬酰胺的水平仅降低了6%并且尼古丁水平同样降低(36%)。

[0005] 尼古丁水平的降低是不可取的,因为烟草产品消费者将需要消耗更多的烟草才能得到相同水平的尼古丁。从US2013/0068240中呈现的结果也清楚地表明,各种EMS突变体之间的天冬酰胺、丙烯酰胺和尼古丁的水平变化很大和随机。

[0006] 希望找到降低植物中天冬酰胺水平的替代方式,并且由此使由植物的加热或燃烧部分产生的气溶胶中丙烯酰胺的量最小化。还希望找到将丙烯酰胺的量进一步降低到比US2013/0068240中所公开的量更低的水平的方法。还希望维持植物中的尼古丁水平。开发能够始终产生具有这些特性的植物的方法也是合乎需要的。本发明试图解决这一需要。

发明内容

[0007] 已经在烟草中鉴定了六种全长天冬酰胺合成酶基因,称为NtASN1-S (SEQ ID NO: 1)、NtASN1-T (SEQ ID NO:3)、NtASN3-S (SEQ ID NO:9)、NtASN3-T (SEQ ID NO:11)、NtASN5-S (SEQ ID NO:5)和NtASN5-T (SEQ ID NO:7)。发现ASN3-S和ASN3-T基因在烘烤期间仅不良地表达,而出人意料地发现NtASN1-S、NtASN1-T、NtASN5-S和NtASN5-T在烘烤期间高度表达,例如在前10、20、30、40、50、60、70、80、90或100小时的烘烤期间。ASN3-S或ASN3-T基因或由其编码的蛋白质不适合用于本公开中。本文显示降低NtASN1和/或NtASN5的表达或由其编码的蛋白质的活性有助于降低烘烤叶(如烘烤的中下部叶)中天冬酰胺的水平。这又使得在加热或燃烧烘烤叶时产生的气溶胶中丙烯酰胺的水平降低。与US2013/0068240中获得的结果相反,降低NtASN1或NtASN5的表达不仅导致叶中天冬酰胺的减少,而且始终导致气溶胶中丙烯酰胺的减少。这种天冬酰胺和丙烯酰胺水平的持续降低伴随着烘烤叶中尼古丁的水平即使不高于对照植物也大致相同。有利地,在本公开的某些实施例中,可以获得较低的约89%的天冬酰胺减少值和约70%的丙烯酰胺减少值,而对尼古丁含量具有基本上最小的影响或没有影响。

[0008] 不希望受理论所束缚,已经进一步确定,在烟草气溶胶中产生的丙烯酰胺依赖于在叶烘烤期间产生的天冬酰胺库。因此,通过调节(例如降低)在烘烤或干燥的早期阶段期间有活性的天冬酰胺合成酶的表达或活性,可以通过减少在烘烤或干燥期间产生的天冬酰胺的库来调节(例如降低)气溶胶中的丙烯酰胺水平。描述了具有降低的天冬酰胺含量并表现出天冬酰胺合成酶表达或功能降低的植物。在实施例中,这是通过使用本文所述的在烘烤或干燥期间表达的天冬酰胺合成酶编码序列的RNA干扰或突变来实现的。降低天冬酰胺水平减少气溶胶中存在的丙烯酰胺的量。当应用于烟草时,这使得吸烟制品(包括加热或燃烧的烟草产品)的消费者吸入的气溶胶中存在的丙烯酰胺的量降低。本文中描述了本发明的其它优势。

[0009] 本发明的方面和实施例

[0010] 本发明的方面和实施例在所附权利要求中阐述。

[0011] 在一个方面,提供一种具有降低的天冬酰胺合成酶表达或活性的突变型、非天然存在的或转基因植物或其部分,所述天冬酰胺合成酶包含以下、由以下组成或基本上由以下组成:(i)聚核苷酸序列,包含以下、由以下组成或基本上由以下组成:与SEQ ID NO:1具有至少72%或至少90%序列一致性或与SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7具有至少72%序列一致性的序列;或(ii)由(i)中所示的聚核苷酸编码的多肽;或(iii)与SEQ ID NO:2或SEQ ID NO:4或SEQ ID NO:6或SEQ ID NO:8具有至少78%序列一致性的多肽;其中(i)、(ii)或(iii)中所示的天冬酰胺合成酶的表达或活性与对照植物相比降低。

[0012] 合适地,植物在编码天冬酰胺合成酶的聚核苷酸的调节区或编码序列中包含至少一个基因改变。

[0013] 合适地,降低的天冬酰胺合成酶的表达或活性赋予源自突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶中天冬酰胺的水平与源自对照植物的经烘烤或干燥的叶中天冬酰胺的水平相比降低,合适地,其中天冬酰胺的水平与对照植物相比降低了至少约17%。

[0014] 合适地,降低的天冬酰胺合成酶的表达或活性赋予在加热或燃烧源自突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶时所产生的气溶胶中丙烯酰胺的水平与源自对照植物的经烘烤或干燥的叶相比降低,合适地,其中气溶胶中丙烯酰胺的水平与对照植物的经烘烤或干燥的叶相比降低了至少20%。

[0015] 合适地,与对照植物的经烘烤或干燥的叶和气溶胶相比,源自突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶中天冬酰胺的水平降低至少约22%并且在加热或燃烧源自突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶时所产生的气溶胶中丙烯酰胺的量降低至少24%;合适地,其中与对照植物的经烘烤或干燥的叶和气溶胶相比,源自突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶中天冬酰胺的水平降低至少约44%并且在加热或燃烧源自突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶时所产生的气溶胶中丙烯酰胺的量降低至少66%;合适地,其中与对照植物的经烘烤或干燥的叶和气溶胶相比,源自突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶中天冬酰胺的水平降低至少约70%并且在加热或燃烧源自突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶时所产生的气溶胶中丙烯酰胺的量降低至少约88%。

[0016] 合适地,来自突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶中尼古丁的水平与对照植物的经烘烤或干燥的叶中尼古丁的水平基本上相同。

[0017] 合适地,与对照植物的经烘烤或干燥的叶相比,源自突变型、非天然存在的或转基因植物或其部分的经烘烤或干燥的叶中谷氨酰胺、天冬氨酸和谷氨酸的形成增加。

[0018] 合适地,植物选自烟草或茶树 (*Camellia sinensis*) 组成的群组。

[0019] 合适地,烟草植物是烟草植物,优选地白肋型烟草植物。

[0020] 合适地,植物在(i)、(ii)或(iii)中所示的编码天冬酰胺合成酶的聚核苷酸序列的每个拷贝中包含至少一个突变,合适地是终止突变和/或干扰编码天冬酰胺合成酶的RNA转录物翻译的基因片段。

[0021] 合适地,天冬酰胺合成酶核苷酸序列包含在与SEQ ID No.19、21和23中编码终止突变的核苷酸序列的位置等同的位置处编码终止突变的核苷酸序列,或其中天冬酰胺合成酶包含在与SEQ ID No.20、22和24中终止密码子的位置等同的位置处编码终止密码子的核苷酸序列。

[0022] 合适地,天冬酰胺合成酶包含选自SEQ ID No.19、21和23组成的群组的核苷酸序列,或由选自SEQ ID No.19、21和23组成的群组的核苷酸序列编码,或包含以下、由以下组成或基本上由以下组成:SEQ ID No.20、22和24中所示的多肽序列。

[0023] 合适地,来自突变型、非天然存在的或转基因植物或其部分的叶生物质的量与来自对照植物的叶生物质的量基本上相同。

[0024] 合适地,叶是晾制的,合适地,其中晾制叶是晒制或明火烘烤的。

[0025] 合适地,叶是风干的,合适地,其中风干叶是晒干或明火烘干的。

[0026] 在另一方面,提供来自本公开的植物的植物材料或经烘烤或干燥的植物材料。

[0027] 在另一方面,提供包含至少一部分本公开的植物或植物材料的植物产品。

[0028] 在另一方面,提供一种制备植物材料的方法,所述植物材料与来自对照植物的植物材料相比天冬酰胺的水平降低并且由所述植物材料获得的气溶胶中丙烯酰胺的水平降

低,所述方法包含以下步骤:(a)提供具有包含以下、由以下组成或基本上由以下组成的聚核苷酸的植物或其部分:编码天冬酰胺合成酶并且与SEQ ID NO:1具有至少72%或至少90%序列一致性或与SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7具有至少72%序列一致性的序列;(b)降低所述植物或其部分中所述聚核苷酸的表达或由其编码的蛋白质的活性;(c)从所述植物或其部分收集植物材料;(d)干燥或烘烤所述植物材料;(e)任选地,测量所述植物或其部分中天冬酰胺的水平 and/或测量由所述植物或其部分获得的气溶胶中丙烯酰胺的水平;和(f)获得经烘烤或干燥的植物材料,所述植物材料的天冬酰胺水平降低并且由所述植物材料获得的气溶胶中丙烯酰胺的水平降低,合适地,其中尼古丁的水平与对照植物中的尼古丁水平基本上相同;合适地,其中谷氨酰胺、天冬氨酸和谷氨酸的形成与对照植物的经烘烤或干燥的叶相比增加。

[0029] 合适地,植物材料在收集后烘烤或干燥至少约3天。

[0030] 合适地,植物材料是晾制的,合适地,其中晾制叶是晒制或明火烘烤的。

[0031] 合适地,植物材料是风干或晒干或明火烘干的。

[0032] 还公开了通过这一方法获得或能够通过这一方法获得的植物材料。

[0033] 在另一方面,提供一种用于鉴定植物中的一个或多个基因改变的方法,所述基因改变与来自所述植物的经烘烤或干燥的植物材料中天冬酰胺的水平 and 由所述经烘烤或干燥的植物或其部分获得的气溶胶中丙烯酰胺的水平与不包含所述一个或多个基因改变的对照植物相比降低相关,所述方法包含以下步骤:(a)鉴定经烘烤或干燥的植物,来自所述植物的植物材料中天冬酰胺的水平 and 由所述经烘烤或干燥的植物材料获得的气溶胶中丙烯酰胺的水平与来自对照植物的植物材料相比降低;(b)提供来自步骤(a)中所鉴定的植物的核酸样品;和(c)在来自步骤(b)的核酸样品中鉴定编码天冬酰胺合成酶并且与SEQ ID NO:1具有至少72%或至少90%序列一致性或与SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7具有至少72%序列一致性的聚核苷酸序列中的一个或多个基因改变。

[0034] 在另一方面,提供一种用于产生植物材料的方法,所述植物材料与对照植物相比天冬酰胺的水平降低并且由经烘烤或干燥的植物材料获得的气溶胶中丙烯酰胺的水平降低,所述方法包含以下步骤:(a)提供根据本公开的植物或植物材料;(b)收集来自所述植物的植物材料;(c)烘烤或干燥所述植物材料一段时间;和(d)获得经烘烤或干燥的植物材料,所述植物材料具有降低的天冬酰胺水平并且由所述经烘烤或干燥的植物材料获得的气溶胶中丙烯酰胺的水平降低,合适地,其中尼古丁的水平与对照植物的尼古丁水平基本上相同;合适地其中谷氨酰胺、天冬氨酸和谷氨酸的形成与对照植物相比增加。

[0035] 在另一方面,提供一种经分离的聚核苷酸序列,其包含以下、由以下组成或基本上由以下组成:编码天冬酰胺合成酶并且与SEQ ID NO:1具有至少72%或至少90%序列一致性或与SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7具有至少72%序列一致性的序列,或包含以下、由以下组成或基本上由以下组成:编码天冬酰胺合成酶并且与SEQ ID No. 19、21和23具有至少72%序列一致性的序列,条件是与SEQ ID No. 19、21和23具有至少72%序列一致性的序列在与SEQ ID No. 19、21和23中终止密码子的位置等同的位置处包含编码终止密码子的核苷酸序列,优选地,其中所述经分离的聚核苷酸是合成聚核苷酸或cDNA。

[0036] 在另一方面,提供一种由本文所述的聚核苷酸编码的经分离的多肽。

[0037] 在另一方面,提供一种经分离的多肽,其包含以下、由以下组成或基本上由以下组

成:编码天冬酰胺合成酶并且与SEQ ID NO:2或SEQ ID NO:4或SEQ ID NO:6或SEQ ID NO:8或SEQ ID NO:20或SEQ ID NO:22或SEQ ID NO:24具有至少78%序列一致性的序列,或包含以下、由以下组成或基本上由以下组成:编码天冬酰胺合成酶并且与SEQ ID NO:2或SEQ ID NO:4或SEQ ID NO:6或SEQ ID NO:8或SEQ ID NO:20或SEQ ID NO:22或SEQ ID NO:24具有至少78%序列一致性的序列,其条件是与SEQ ID NO:20或SEQ ID NO:22或SEQ ID NO:24具有至少78%序列一致性的序列在与SEQ ID NO:20或SEQ ID NO:22或SEQ ID NO:24中终止密码子的位置等同的位置处包含终止密码子优选地,其中所述经分离的聚核苷酸是合成聚核苷酸或cDNA。

[0038] 在另一方面,提供一种包含本文所述的经分离的聚核苷酸的构建体、载体或表达载体。

[0039] 在另一方面,提供一种来自本文所述的植物的突变型、非天然存在的或转基因植物细胞。

[0040] 在另一方面,提供包含本文所述的细胞的经烘烤或干燥的植物材料。

[0041] 在另一方面,提供一种包含本文所述的植物材料的烟草产品或吸烟制品。

[0042] 还公开了所阐述的一个或多个实施例的组合。

[0043] 一些优势

[0044] 本文显示天冬酰胺在烘烤期间主动产生。已显示,天冬酰胺的这种累积导致在加热或燃烧经烘烤的植物材料时所产生的气溶胶中形成丙烯酰胺。有利地,NtASN1-S、NtASN1-T、NtASN5-S和NtASN5-T在烘烤期间,特别是从烘烤开始就高度表达。降低这些基因中的一种或多种的表达可以导致气溶胶中丙烯酰胺的水平降低,因为在整个烘烤过程中天冬酰胺的水平可以降低。

[0045] 本文所述的经修饰的植物中对尼古丁水平的影响有限,这在打算将经修饰的植物用于生产烟草植物时是合乎需要的。

[0046] 本文所述的方法始终产生在烘烤或干燥期间天冬酰胺的累积减少并且气溶胶中丙烯酰胺的形成也减少的植物。

[0047] 当将本文所述的经修饰的植物用于商业目的时,对叶生物质和表型的影响有限,这是有利的。

[0048] 本公开允许创建未经遗传修饰的植物,这可能对于消费者来说更能接受。

[0049] 本公开不限于使用如US2013/0068240中的EMS突变植物。EMS突变植物在育种后可能不具有改善作物特性的潜力。一旦开始育种,由于不同的原因,EMS突变植物的理想特征可能会丢失。举例来说,可能需要多个突变,突变可以是显性或隐性的,并且可能难以在基因靶标中鉴定点突变。相反,本公开利用了可以特异性操作和/或敲除的某些基因的用途,以产生具有所需表型的植物。本公开可以应用于其它烟草品种或作物,由此提供期望的解决方案以至少减少丙烯酰胺。

附图说明

[0050] 图1显示在同一田间(重复平均5个地块)培养并且同时烘烤的3种不同白肋烟草中的天冬氨酸(Asp)、谷氨酰胺(Glu)、天冬酰胺(Asn)和总游离氨基酸(aa)含量。三种白肋烟是TN90(A)、Banket A1(B)和Kentucky 14(C)。

[0051] 图2显示在白肋烟叶烘烤期间天冬酰胺、谷氨酰胺和天冬氨酸进化的时程。

[0052] 图3显示天冬酰胺合成酶活性途径。

[0053] 图4显示在单品级烟草弗吉尼亚 (FC)、东方 (OR)、蜘蛛兰 (KS) 或白肋 (BU) 和包含弗吉尼亚 (FC)、蜘蛛兰 (KS) 和东方 (OR) 的混合物的掺合物 (称为RRP) 中天冬酰胺的存在之间存在的相关性。

[0054] 图5显示当瑞士白肋Stella在瑞士生长并且在晾房中烘烤10周 (传统的农业做法, BU-AC) 或恰如瑞士弗吉尼亚烟草一样在烟道烘烤炉中烘烤 (BU-FC) 时所获得的结果。在烘烤后, 研磨叶片并且进行天冬酰胺测定 (A)。制成均化的烟草材料粉末 (小于或等于200微米)。基本上如W02013098405中所述, 将均化的烟草材料用于气溶胶生成制品中, 从而与气溶胶生成装置一起使用。使用加拿大卫生部 (Health-Canada) 吸烟制度抽吸气溶胶生成制品以测定气溶胶中丙烯酰胺的含量 (B)。

[0055] 图6显示由对应于NtASN1-S、NtASN1-T、NtASN3-S、NtASN3-T、NtASN5-S、NtASN5-T的烟草基因推导的氨基酸序列与拟南芥蛋白质AtASN1 (At3g47340)、AtASN2 (At5g65010) 和AtASN3 (At5g10240) 以及番茄蛋白质Soly06g007180和Soly04g055200相比的系统发生树。

[0056] 图7显示2010年在瑞士生长的瑞士白肋烟草 (Stella) 在烘烤期间ASN基因的表达。ARN是从田间 (-4小时) 和晾房中直到烘烤96小时以时程方式收集的叶样 (中部秆位) 分离的。通过使用Affymetrix Tobarray芯片的微阵列分析来确定特异性探针上的基因表达。

[0057] 图8显示瑞士白肋 (Stella, A和C) 和瑞士弗吉尼亚 (亮色, ITB683, B和B) 烟草在48小时烘烤后的ASN基因的表达。

[0058] 图9显示2013年在贝耶纳 (Payerne) 田间生长的4种不同的白肋烟叶 (下部秆位) 在60小时晾制后ASN1、ASN5 (A) 和ASN3 (B) 的收获后表达。

[0059] 图10显示从田间生长的成熟白肋烟草收集的8种组织 (未成熟的花 (Im_flower; 下部_LS、中部_MS和上部_US秆位处的叶; 花瓣、根、萼片和茎) 中ASN1、ASN5 (左图) 和ASN3 (右图) 的表达。。

[0060] 图11显示在MMV或E4启动子的控制下在温室中生长的ASN1-RNAi植物中的天冬酰胺含量和在烘烤3天后测定的ASN1转录物水平。在每个转化事件的15株植物的中下部秆叶位置测量天冬酰胺的量 (A)。通过qPCR使用在每个转化事件的15株植物烘烤3天后收集的RNA估计ASN1转录物的相对表达 (B)。

[0061] 图12显示在MMV或E4启动子的控制下在温室中生长的ASN1-RNAi植物中的谷氨酰胺、天冬氨酸、谷氨酸和尼古丁含量。在每个转化事件的15株植物的中下部秆叶位置测量Gln (A)、Asp (B)、Glu (C) 和尼古丁 (D) 的量。

[0062] 图13显示在MMV或E4启动子的控制下在温室中生长的ASN5-RNAi植物中的天冬酰胺含量和在烘烤3天后测定的ASN5转录物水平。在每个转化事件的15株植物的中下部秆叶位置测量天冬酰胺的量 (A)。通过qPCR使用在每个转化事件的15株植物烘烤3天后收集的RNA估计ASN5转录物的相对表达 (B)。

[0063] 图14显示在MMV或E4启动子的控制下在温室中生长的ASN5-RNAi植物中的谷氨酰胺、天冬氨酸、谷氨酸和尼古丁含量。在每个转化事件的15株植物的中下部秆叶位置测量Gln (A)、Asp (B)、Glu (C) 和尼古丁 (D) 的量。

[0064] 图15显示ASN1-和ASN5-RNAi品系的叶生物质。从每个转化事件的4株植物的中部秆位收集四片叶子,随机选择并称重。显示重量平均值和标准偏差。

[0065] 图16显示在瑞士白肋(Stella)晾制的早期步骤期间CYP82E4的表达。使用Tobarray Affymetrix芯片基于针对细胞色素P450单加氧酶CYP82E4的特异性探针NtPM1a1g2e2_st确定转录物。

[0066] 图17显示ASN变体中天冬酰胺减少的百分比。相对于每个变体品系的WT外分离体计算百分比值。

[0067] 图18显示在中下部秆位收获的烘烤叶(4)的生物质。

[0068] 图19显示如W02013/098405中所述,通过加热由ASN1-RNAi、ASN5-RNAi和它们各自的对照叶制成的气溶胶形成制品生成的气溶胶中丙烯酰胺的测定。在与50%弗吉尼亚烟草掺合后,测定ASN1-RNAi、ASN5-RNAi和对照叶(参见图11和图13)的细粉中的天冬酰胺(A、C)。将烟草掺合物转化成铸叶并在气溶胶形成制品内加热。随后测定气溶胶中的丙烯酰胺(B、D)。

[0069] 图20显示由ASN1-RNAi和它们各自的对照中部秆叶片的切丝填料制成的可燃香烟的烟雾中丙烯酰胺的测定。测定由ASN1-RNAi和对照叶制备的切丝填料中的天冬酰胺(图A)。抽吸可燃香烟并且分析烟雾中的丙烯酰胺。

[0070] 图21显示来自ASN1-S_stop突变体和对照(WT外分离体)叶制成的气溶胶生成制品的气溶胶中丙烯酰胺的测定。在与50%弗吉尼亚烟草掺合后,测定ASN1-S_stop突变体和对照叶的细粉中的天冬酰胺(A)。将烟草掺合物转化成铸叶和由气溶胶形成制品生成的气溶胶。随后测定气溶胶中的丙烯酰胺(B)。

[0071] 图22显示用于生成ASN-RNAi植物的载体的实例。

具体实施方式

[0072] 定义

[0073] 在本申请的范围内使用的技术术语和表述通常被赋予在植物和分子生物学的相关领域中经常应用于它们的含义。所有以下术语定义适用于本申请的整个内容。词语“包含”不排除其它元件或步骤,且不定冠词“一(a)”或“一(an)”不排除多个/多种。单个步骤可以完成在权利要求中列出的几个特征的功能。在给定数值或范围背景下,术语“约”、“基本上”和“大约”指在给定值或范围20%内,10%内或者5%、4%、3%、2%或1%内的值或范围。

[0074] 术语“经分离的”指取自其天然环境的任何实体,但所述术语不意味着任何纯化程度。

[0075] “表达载体”是包含允许核酸表达的核酸组分组合的核酸媒介物。合适表达载体包括能够染色体外复制的游离基因,如环状双链核酸质粒;经线性化的双链核酸质粒;以及任何来源的其它功能等效的表达载体。如下文所定义,表达载体包含位于核酸、核酸构建体或核酸结合物上游并且与核酸、核酸构建体或核酸结合物可操作地连接的至少一个启动子。

[0076] 术语“构建体”指包含一种或多种聚核苷酸的双链重组核酸片段。构建体包含与互补“有义链或编码链”碱基配对的“模板链”。给定构建体可以在两个可能定向上插入载体内,所述两个可能定向是就置于载体(如表达载体)内的启动子定向而言的相同(或有义)定向或相反(或反义)定向。

[0077] “载体”指核酸媒介物,其包含允许核酸、核酸构建体和核酸结合物等等转运的核酸组分组合。合适载体包括能够染色体外复制的游离基因,如环状双链核酸质粒;经线性化的双链核酸质粒;以及任何来源的其它载体。

[0078] “启动子”指通常位于双链DNA片段上游并且与双链DNA片段可操作地连接的核酸元件/序列。启动子可以完全源自接近天然目的基因的区域,或可以由源自不同天然启动子的不同元件或合成DNA区段构成。

[0079] 术语“同源性、一致性或相似性”指通过序列比对比较的两个多肽或两个核酸分子之间的序列相似性程度。所比较的两个不连续核酸序列之间的同源性程度随着可比较位置处的一致或匹配核苷酸数目变化。一致性百分比可通过目视检查和数学计算进行确定。或者,可以通过使用计算机程序(如ClustalW、BLAST、FASTA或Smith-Waterman)比较序列信息,确定两个核酸序列的一致性百分比。两个序列的一致性百分比可能取不同的值,具体取决于:(i)用于比对序列的方法,例如ClustalW、BLAST、FASTA、Smith-Waterman(在不同的程序中实现),或者来自3D比较的结构比对;和(ii)比对方法使用的参数,例如局部相对于全局比对,使用的配对得分矩阵(例如BLOSUM62、PAM250、Gonnet等)和缺口罚分,例如功能形式和常量。在进行比对之后,有不同的方式计算两个序列之间的一致性百分比。举例来说,可以将一致性数目除以:(i)最短序列的长度;(ii)比对的长度;(iii)序列的平均长度;(iv)非缺口位置的数目;或(iv)不包括突出端的等价位置的数目。此外,应了解一致性百分比也具有强烈的长度依赖性。因此,一对序列越短,可以预期偶然出现的序列一致性越高。流行的多重比对程序ClustalW(Nucleic Acids Research(1994) 22,4673-4680;Nucleic Acids Research(1997),24,4876-4882)是用于产生多肽或聚核苷酸的多重比对的合适方式。ClustalW的合适参数可能如下:对于聚核苷酸比对:缺口开放罚分=15.0,缺口延伸罚分=6.66,并且矩阵=一致性。对于多肽比对:缺口开放罚分=10.0,缺口延伸罚分=0.2,并且矩阵=Gonnet。对于DNA和蛋白质比对:ENDGAP=-1,并且GAPDIST=4。本领域的技术人员将会意识到,可能有必要改变这些和其它参数以达到最佳序列比对。然后,由这样的比对合适地以(N/T)计算一致性百分比,其中N是序列共享一致残基的位置数,T是比较的位置总数,包括缺口但不包括突出端。

[0080] “变体”是指基本上相似的序列。变体可以具有与野生型序列相似的功能或基本上相似的功能。对于天冬酰胺合成酶,相似的功能是在相同条件下将天冬氨酸转化为天冬酰胺的野生型酶功能的至少约50%、60%、70%、80%或90%。对于天冬酰胺合成酶,基本上相似的功能是在相同条件下将天冬氨酸转化为天冬酰胺的野生型酶功能的至少约90%、95%、96%、97%、98%或99%。与野生型多肽相比,变体可以具有一个或多个有利的突变,导致所述酶具有降低的天冬酰胺合成酶活性水平。变体可以具有一个或多个有利的突变,导致其天冬酰胺合成酶活性被敲除(即,100%抑制,因此是非功能性多肽)。野生型ASN1-S的示例性变体是ASN1-S_W156*(SEQ ID No 19和20),其具有与野生型ASN1-S多肽相比导致天冬酰胺减少约17%的有利终止突变。野生型ASN1-T的示例性变体是ASN1-T_W156*(SEQ ID No 21和22),其具有与野生型ASN1-S多肽相比导致天冬酰胺减少约83%的有利终止突变。野生型ASN5-S的示例性变体是ASN5-S_Q66*(SEQ ID No 23和24),其具有与野生型ASN1-S多肽相比导致天冬酰胺减少约44%的有利终止突变。

[0081] 术语“植物”指处于其生命周期或发育的任何阶段的任何植物或植物的部分以及

其后代。在一个实施例中,植物是“烟草植物”,这指属于烟草属(*Nicotiana*)的植物。本文中描述了优选的烟草植物物种。合适地,植物是与对照植物相比一种或多种基因的表达或一种或多种蛋白质的活性被调节的突变型、非天然存在的或转基因植物。合适地,使植物成为突变型、非天然存在的或转基因植物的改变导致调节一种或多种基因的表达或调节一种或多种蛋白质的活性。在某些实施例中,改变是基因改变或基因修饰。

[0082] “植物部分”包括植物细胞、植物原生质体、可再生完整植物的植物细胞组织培养物、植物愈伤组织、植物团块以及植物或植物的部分中完整的植物细胞,所述植物的部分如胚、花粉、花药、胚珠、种子、叶、花、茎、枝、果实、根、根尖等。再生植物的后代、变体和突变体也包括在本公开的范围,限制条件是其包含本文所述的引入的聚核苷酸。植物的叶特别优选地用于本公开中。

[0083] “植物细胞”指植物的结构和生理单位。植物细胞可以采取不含细胞壁的原生质体、经分离的单细胞或培养细胞的形式,或作为更高等组构单位的一部分,如但不限于植物组织、植物器官或全植物。

[0084] 术语“植物材料”指可得自植物的任何固体、液体或气体组合物或其组合,包括生物质、叶、茎、根、花或花的部分、果实、花粉、卵细胞、合子、种子、插条、分泌物、提取物、细胞或组织培养物、或任何其它植物部分或产物。在一个实施例中,植物材料包含生物质、茎、种子或叶,或者由生物质、茎、种子或叶组成。在另一个实施例中,植物材料包含叶或由叶组成。

[0085] 术语“品种”指共享恒定特征的植物群体,所述恒定特征使其与相同物种的其它植物分开。尽管具有一种或多种独特性状,但品种的特征进一步在于所述品种内个体之间的极小整体变化。品种通常市场上有出售。

[0086] 如本文所用的术语“品系”或“育种品系”指示在植物育种期间使用的一群植物。品系可与品种区分,因为它展示对于一种或多种目的性状在个体间的很少变化,尽管对于其它性状在个体间可能存在一些变化。

[0087] 如本文所用的术语“非天然存在的”描述并非天然形成或在自然界中不存在的实体(例如聚核苷酸、基因突变、多肽、植物、植物细胞和植物材料)。可以通过本文描述或本领域已知的方法,制备、合成、起始、修饰、干预或操纵这类非天然存在的实体或人工实体。可以由人制备、合成、起始、修饰、干预或操纵这类非天然存在的实体或人工实体。因此,例如,非天然存在的植物、非天然存在的植物细胞或非天然存在的植物材料可以使用遗传操纵技术(如反义RNA、干扰RNA、大范围核酸酶等等)进行制备。进一步举例来说,可以通过第一植物或植物细胞基因渗入第二植物或植物细胞(其自身可以是天然存在的)内,或通过将一个或多个基因突变(例如一种或多种多态性)从第一植物或植物细胞转移到第二植物或植物细胞内来制备非天然存在的植物、非天然存在的植物细胞或非天然存在的植物材料,使得所得到的植物、植物细胞或植物材料或其后代包含并非天然形成或在自然界中不存在的基因组成(例如基因组、染色体或其区段)。所得到的植物、植物细胞或植物材料因此是人工的或非天然存在的。相应地,可以通过修饰第一天然存在的植物或植物细胞中的基因序列来制备人工的或非天然存在的植物或植物细胞,即使所得到的基因序列在第二植物或植物细胞中天然存在,所述第二植物或植物细胞包含与第一植物或植物细胞不同的遗传背景。

[0088] 术语“调节”可以指降低、抑制、增加或以其它方式影响多肽的表达或活性。所述术

语还可以指降低、抑制、增加或以其它方式影响编码多肽的基因的活性,可以包括但不限于调节转录活性。

[0089] 如本文所用,术语“降低”或“降低的”指数量或活性约10%至约99%的降低,或者至少10%、至少20%、至少25%、至少30%、至少40%、至少50%、至少60%、至少70%、至少75%、至少80%、至少90%、至少95%、至少98%、至少99%或至少100%或更多的降低,所述活性如但不限于多肽活性、转录活性和蛋白质表达。

[0090] 如本文所用,术语“抑制”或“抑制的”指数量或活性约98%至约100%的降低,或者至少98%、至少99%但特别是100%的降低,所述活性如但不限于多肽活性、转录活性和蛋白质表达。

[0091] 细胞的转化可以是稳定的或瞬时的。术语“瞬时转化”或“瞬时转化的”或其变化形式是指在外源聚核苷酸不整合到宿主细胞基因组中的情况下将一种或多种外源聚核苷酸引入细胞中。相反,术语“稳定转化”或“稳定转化的”是指将一个或多个外源聚核苷酸引入和整合到细胞基因组中。术语“稳定转化体”是指将一个或多个外源聚核苷酸稳定地整合到基因组或细胞器DNA中的细胞。应理解,用本公开的核酸、构建体和/或载体转化的生物体或其细胞可以瞬时地以及稳定地转化。在某些实施例中,稳定转化是优选的。

[0092] 如本文所用,术语“增加”或“增加的”指数量或活性约5%至约99%的增加,或者至少5%、至少10%、至少20%、至少25%、至少30%、至少40%、至少50%、至少60%、至少70%、至少75%、至少80%、至少90%、至少95%、至少98%、至少99%或至少100%或更多的增加,所述活性如但不限于多肽活性、转录活性和蛋白质表达。

[0093] 如本文所用并且当在量的背景下使用时,术语“基本上”意指所述量是与其相比的量的至少约10%、至少约9%、至少约8%、至少约7%、至少约6%、至少约5%、至少约4%、至少约3%、至少约2%、至少约1%或至少约0.1%。

[0094] 在对照植物或对照植物细胞等背景下的术语“对照”意指目的基因或蛋白质的表达或活性未被调节的植物或植物细胞,因此它可以提供与其中酶的表达或活性已被修饰的植物或植物细胞的比较或参考。因此,在本发明的上下文中,对照将不包括降低天冬酰胺合成酶表达或活性的至少一种基因改变。对照植物或植物细胞可以包含空白载体。对照植物或植物细胞可以对应于野生型植物或野生型植物细胞等等。在所有这类情况中,出于比较目的,使用相同的方案来培养和收获主题植物和对照植物。本文所述的基因或多肽的水平、比率、活性或分布的变化,或植物表型的变化,特别是天冬酰胺累积减少和/或丙烯酰胺累积减少和/或谷氨酰胺累积增加,和/或天冬氨酸累积增加和/或谷氨酸累积增加可以通过将主题植物与对照植物进行比较来测量,合适地,其中主题植物和对照植物已经使用相同的方案进行培养和/或收获。对照植物可以提供用于测量主题植物的表型变化的参照点。对表型变化的测量可以在植物的任何时间测量,包含在植物发育、衰老期间或在烘烤后。对表型变化的测量可以在任何条件下生长的植物中测量,包含在生长室、温室或在野外生长的植物。表型变化可以通过测量天冬酰胺含量和/或谷氨酰胺含量和/或天冬氨酸含量和/或谷氨酸含量(如Moldoveanu (2005) 中所述) 和/或丙烯酰胺含量(如Papousek等人, 2014和/或Onoa等人., 2003和/或美国环境保护局方法8032A-丙烯酰胺气相色谱法,修订1, 1996年12月和/或尼古丁含量(使用烟草科学研究合作中心推荐的(CORESTA) 方法编号62,通过气相色谱分析(2005年2月) 测定烟草和烟草产品中的尼古丁) 可以在烘烤或干燥之前和/或期

间和/或之后使用本领域中熟知的方法进行监测。使用由Moldoveanu (2005) 改编的方法进行游离氨基酸分析。

具体实施方式

[0095]

[0096] 在一个实施例中,提供一种经分离的聚核苷酸,其包含以下、由以下组成或基本上由以下组成:与本文所述的序列中的任一种,包括序列表中所示的聚核苷酸中的任一种具有至少60%序列一致性的聚核苷酸序列。合适地,经分离的聚核苷酸包含以下、由以下组成或基本上由以下组成:与其具有至少60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、78%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列一致性的序列。更合适地,经分离的聚核苷酸包含以下、由以下组成或基本上由以下组成:与其具有至少72%、73%、74%、75%、78%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列一致性的序列。在表1中示出本文所述序列的比对。

[0097] 在另一个实施例中,提供一种经分离的聚核苷酸,其包含以下、由以下组成或基本上由以下组成:与SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11具有至少60%序列一致性的聚核苷酸序列。合适地,经分离的聚核苷酸SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11包含以下、由以下组成或基本上由以下组成:与SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11具有至少约60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、75%、78%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%或100%序列一致性的序列。更合适地,经分离的聚核苷酸包含以下、由以下组成或基本上由以下组成:与SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11具有至少72%、73%、74%、75%、78%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列一致性的序列。

[0098] 更合适地,经分离的聚核苷酸包含以下、由以下组成或基本上由以下组成:与SEQ ID NO:1具有至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列一致性的序列。

[0099] 更合适地,经分离的聚核苷酸包含以下、由以下组成或基本上由以下组成:与SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11具有至少72%、73%、74%、75%、78%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列一致性的序列。

[0100] 在另一个实施例中,提供一种经分离的聚核苷酸,其包含以下、由以下组成或基本上由以下组成:与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23具有至少60%序列一致性的聚核苷酸序列,条件是所述经分离的聚核苷酸在与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23中所示的位置等同的位置处包含编码终止密码子的突变。

[0101] 合适地,经分离的聚核苷酸包含以下、由以下组成或基本上由以下组成:与SEQ ID

NO:19或SEQ ID NO:21或SEQ ID NO:23具有至少约60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、75%、78%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%或100%序列一致性的序列。更合适地,经分离的聚核苷酸包含以下、由以下组成或基本上由以下组成:与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23具有至少72%、73%、74%、75%、78%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列一致性的序列。

[0102] 在另一个实施例中,提供一种聚核苷酸,其包含以下、由以下组成或基本上由以下组成:与SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11具有相当大源性(即,序列相似性)或相当大一致性的聚核苷酸。

[0103] 在另一个实施例中,提供一种聚核苷酸,其包含以下、由以下组成或基本上由以下组成:与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23具有相当大源性(即,序列相似性)或相当大一致性的聚核苷酸,条件是所述经分离的聚核苷酸在与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23中所示的位置等同的位置处包含编码终止密码子的突变。

[0104] 在另一个实施例中,提供聚核苷酸变体,其与SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11的序列具有至少约60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、75%、78%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%或99.9%序列一致性。合适地,提供聚核苷酸变体,其与SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11的序列具有至少约72%、73%、74%、75%、78%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%或99.9%序列一致性。

[0105] 合适地,提供聚核苷酸变体,其与SEQ ID NO:1的序列具有至少约90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%或99.9%序列一致性。

[0106] 合适地,提供聚核苷酸变体,其与SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11的序列具有至少约72%、73%、74%、75%、78%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%或99.9%序列一致性。

[0107] 在另一个实施例中,提供聚核苷酸变体,其与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23的序列具有至少约60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、75%、78%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%或99.9%序列一致性,条件是所述聚核苷酸变体在与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23中所示的位置等同的位置处包含编码终止密码子的突变。合适地,提供聚核苷酸变体,其与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23的序列具有至少约72%、73%、74%、75%、78%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%或

99.9%序列一致性,条件是所述聚核苷酸变体在与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23中所示的位置等同的位置处包含编码终止密码子的突变。在另一个实施例中,提供与SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11具有相当大同源性(即,序列相似性)或相当大一致性的SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11的片段,所述片段与SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11的相应片段具有至少约60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、78%、75%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%或100%序列一致性。

[0108] 在另一个实施例中,提供与SEQ ID NO:1具有相当大同源性(即,序列相似性)或相当大一致性的SEQ ID NO:1的片段,所述片段与SEQ ID NO:1的相应片段具有至少约90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%或100%序列一致性。

[0109] 在另一个实施例中,提供与SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11具有相当大同源性(即,序列相似性)或相当大一致性的SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11的片段,所述片段与SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11的相应片段具有至少约60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、78%、75%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%或100%序列一致性。

[0110] 合适地,提供与SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11具有相当大同源性(即,序列相似性)或相当大一致性的SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11的聚核苷酸变体的片段,所述片段与SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11的相应片段具有至少约72%、73%、74%、75%、78%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%或99.9%序列一致性。

[0111] 合适地,提供与SEQ ID NO:1具有相当大同源性(即,序列相似性)或相当大一致性的SEQ ID NO:1的聚核苷酸变体的片段,所述片段与SEQ ID NO:1的相应片段具有至少约72%、73%、74%、75%、78%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%或99.9%序列一致性。

[0112] 合适地,提供与SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11具有相当大同源性(即,序列相似性)或相当大一致性的SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11的聚核苷酸变体的片段,所述片段与SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11的相应片段具有至少约72%、73%、74%、75%、78%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、

99.7%、99.8%或99.9%序列一致性。

[0113] 在另一个实施例中,提供与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23具有相当大同性(即,序列相似性)或相当大一致性的SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23的聚核苷酸片段,所述片段与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23的相应片段具有至少约60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、78%、75%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%或100%序列一致性,条件是所述片段在与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23中所示的位置等同的位置处包含编码终止密码子的突变。合适地,提供与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23具有相当大同性(即,序列相似性)或相当大一致性的SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23的聚核苷酸变体的片段,所述片段与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23的相应片段具有至少约72%、73%、74%、75%、78%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%或99.9%序列一致性,条件是所述片段在与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23中所示的位置等同的位置处包含编码终止密码子的突变。

[0114] 在另一个实施例中,提供包含与SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7或SEQ ID NO:9或SEQ ID NO:11的足够或相当大程度的一致性 or 相似性的聚核苷酸,其编码充当天冬酰胺合成酶的多肽。合适地,本文所述的聚核苷酸编码具有天冬酰胺合成酶活性的蛋白质,所述活性是SEQ ID NO:2或SEQ ID NO:4或SEQ ID NO:6或SEQ ID NO:8或SEQ ID NO:10或SEQ ID NO:12中所示蛋白质的活性的至少约60%、70%、80%、90%、95%、96%、97%、98%、99%或100%或更多。为了确定多肽是否是功能性天冬酰胺合成酶,可以使用由Romagni和Dayan(2000) J Agric Food Chem. 5月;48(5):1692-6所描述的分析。在另一个实施例中,提供包含与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23的足够或相当大程度的一致性 or 相似性的聚核苷酸,其编码充当天冬酰胺合成酶的多肽,条件是所述聚核苷酸在与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23中所示的位置等同的位置处包含编码终止密码子的突变。合适地,本文所述的聚核苷酸编码具有天冬酰胺合成酶活性的蛋白质,所述活性是SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23中所示蛋白质的活性的至少约60%、70%、80%、90%、95%、96%、97%、98%、99%或100%或更多。

[0115] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的聚核苷酸的表达:编码天冬酰胺合成酶并且与SEQ ID NO:1和SEQ ID NO:3和SEQ ID NO:5和SEQ ID NO:7具有至少72%序列一致性的序列。

[0116] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的聚核苷酸的表达:编码天冬酰胺合成酶并且与SEQ ID NO:1具有至少90%序列一致性的序列。

[0117] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的聚核苷酸的表达:编码天冬酰胺合成酶并且与SEQ ID NO:3和SEQ ID NO:5和SEQ ID NO:7具有至少72%序列一致性的序列。

[0118] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的聚核苷酸的表达:编码天冬酰胺合成酶并且与SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:5或SEQ ID

NO:7具有至少72%序列一致性的序列。

[0119] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的聚核苷酸的表达:编码天冬酰胺合成酶并且与SEQ ID NO:1具有至少90%序列一致性的序列。

[0120] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的聚核苷酸的表达:编码天冬酰胺合成酶并且与SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7具有至少72%序列一致性的序列。

[0121] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的聚核苷酸的表达:编码天冬酰胺合成酶并且与SEQ ID NO:1和SEQ ID NO:3具有至少72%序列一致性的序列。在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的聚核苷酸的表达:编码天冬酰胺合成酶并且与SEQ ID NO:1具有至少90%序列一致性且与SEQ ID NO:3具有至少72%序列一致性的序列。在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的聚核苷酸的表达:编码天冬酰胺合成酶并且与SEQ ID NO:5和SEQ ID NO:7具有至少72%序列一致性的序列。在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的聚核苷酸的表达:编码天冬酰胺合成酶并且与SEQ ID NO:1和SEQ ID NO:3和SEQ ID NO:5和/或SEQ ID NO:7具有至少72%序列一致性的序列。在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的聚核苷酸的表达:编码天冬酰胺合成酶并且与SEQ ID NO:1具有至少90%序列一致性且与SEQ ID NO:3和SEQ ID NO:5和/或SEQ ID NO:7具有至少72%序列一致性的序列。在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的聚核苷酸的表达:编码天冬酰胺合成酶并且与SEQ ID NO:1和/或SEQ ID NO:3和SEQ ID NO:5和SEQ ID NO:7具有至少72%序列一致性的序列。在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的聚核苷酸的表达:编码天冬酰胺合成酶并且与SEQ ID NO:1具有至少90%序列一致性和/或与SEQ ID NO:3和SEQ ID NO:5和SEQ ID NO:7具有至少72%序列一致性的序列。

[0122] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的聚核苷酸的表达:编码天冬酰胺合成酶并且与SEQ ID NO:19和SEQ ID NO:21和SEQ ID NO:23具有至少72%序列一致性的序列,条件是所述聚核苷酸在与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23中所示的位置等同的位置处包含编码终止密码子的突变。

[0123] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的聚核苷酸的表达:编码天冬酰胺合成酶并且与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23具有至少72%序列一致性的序列,条件是所述聚核苷酸在与SEQ ID NO:19或SEQ ID NO:21或SEQ ID NO:23中所示的位置等同的位置处包含编码终止密码子的突变。

[0124] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的聚核苷酸的表达:编码天冬酰胺合成酶并且与SEQ ID NO:19和SEQ ID NO:21具有至少72%序列一致性的序列,条件是所述聚核苷酸在与SEQ ID NO:19和SEQ ID NO:21中所示的位置等同的位置处包含编码终止密码子的突变。

[0125] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的聚核苷酸的表达:编码天冬酰胺合成酶并且与SEQ ID NO:21和SEQ ID NO:23具有至少72%序列一致性的序列,条件是所述聚核苷酸在与SEQ ID NO:21和SEQ ID NO:23中所示的位置等同的位置处包含编码终止密码子的突变。

[0126] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的聚核苷酸的表达:编码天冬酰胺合成酶并且与SEQ ID NO:19和SEQ ID NO:21和/或SEQ ID NO:23具有至少72%序列一致性的序列,条件是所述聚核苷酸在与SEQ ID NO:19、SEQ ID NO:21和SEQ ID NO:23中所示的位置等同的位置处包含编码终止密码子的突变。

[0127] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的聚核苷酸的表达:编码天冬酰胺合成酶并且与SEQ ID NO:19和/或SEQ ID NO:21和SEQ ID NO:23具有至少72%序列一致性的序列,条件是所述聚核苷酸在与SEQ ID NO:19、SEQ ID NO:21和SEQ ID NO:23中所示的位置等同的位置处包含编码终止密码子的突变。

[0128] 如本文所述的聚核苷酸可以包括核苷酸聚合物,其可以是未经修饰或经修饰的脱氧核糖核酸(DNA)或核糖核酸(RNA)。相应地,聚核苷酸可以是但不限于基因组DNA、互补DNA(cDNA)、mRNA或反义RNA或其片段。此外,聚核苷酸可以是单链或双链DNA、单链和双链区混合的DNA、包含DNA和RNA的杂交分子或具有单链和双链区的混合物的杂交分子或其片段。另外,聚核苷酸可以由包含DNA、RNA或两者的三链区或者其片段构成。聚核苷酸可以含有一个或多个经修饰的碱基,如硫代磷酸酯,并且可以是肽核酸。一般来说,聚核苷酸可以由经分离或克隆的cDNA片段、基因组DNA、寡核苷酸或个别核苷酸或前述的组合装配。尽管本文所述的聚核苷酸序列显示为DNA序列,但序列包括其相应的RNA序列以及其互补(例如完全互补)DNA或RNA序列,包括其逆向互补序列。

[0129] 如本文所述的聚核苷酸一般将会含有磷酸二酯键,尽管在一些情况下,包括聚核苷酸类似物,其可能具有替代主链,包含例如氨基磷酸酯、硫代磷酸酯、二硫代磷酸酯或O-甲基亚磷酰胺键;以及肽聚核苷酸主链和键。其它类似聚核苷酸包括具有阳性主链;非离子主链和非核糖主链的那些。核糖-磷酸主链的修饰可以出于多种原因而完成,例如增加这种分子在生理环境中的稳定性和半衰期,或作为生物芯片上的探针。可以制备天然存在的聚核苷酸和类似物的混合物;或者,可以制备不同聚核苷酸类似物的混合物,以及天然存在的聚核苷酸和类似物的混合物。

[0130] 多种聚核苷酸类似物是已知的,包括例如氨基磷酸酯、硫代磷酸酯、二硫代磷酸酯、O-甲基亚磷酰胺键以及肽聚核苷酸主链和键。其它类似聚核苷酸包括具有阳性主链、非离子主链和非核糖主链的那些。还包括含一种或多种碳环糖的聚核苷酸。

[0131] 其它类似物包括作为肽聚核苷酸类似物的肽聚核苷酸。这些主链在中性条件下是基本上非离子的,与天然存在的聚核苷酸的高度荷电的磷酸二酯主链形成对比。这可以产生优势。首先,肽聚核苷酸主链可以显示出改善的杂交动力学。对于错配碱基对相对于完全匹配的碱基对,肽聚核苷酸在解链温度中具有更大变化。对于内部错配,DNA和RNA通常显示出在解链温度中的2-4℃下降。在非离子肽聚核苷酸主链的情况下,下降接近于7-9℃。类似地,由于其非离子性质,连接至这些主链的碱基的杂交对盐浓度相对不敏感。另外,肽聚核苷酸可以不被细胞酶降解或被细胞酶降解至更少程度,并且因此可以是更稳定的。

[0132] 在所公开的聚核苷酸及其片段的用途中有片段作为核酸杂交分析中的探针或用于核酸扩增分析的引物的用途。这类片段一般包含DNA序列的至少约10、11、12、13、14、15、16、17、18、19或20个或更多个邻接核苷酸。在其它实施例中,DNA片段包含DNA序列的至少约10、15、20、30、40、50或60个或更多个邻接核苷酸。因此,在一个方面,还提供一种用于检测编码具有天冬酰胺合成酶活性的蛋白质或编码天冬酰胺合成酶的聚核苷酸的方法,所述方

法包含使用探针或引物或两者。示例性引物在SEQ ID NO:13至16中示出。因此,另一方面涉及一种寡核苷酸引物,其包含以下、由以下组成或基本上由以下组成:与SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:15或SEQ ID NO:16具有至少60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、75%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%或100%序列一致性的核苷酸序列。还公开了使用这些引物检测编码具有天冬酰胺合成酶活性的蛋白质的聚核苷酸的方法及其用途。

[0133] 影响聚核苷酸杂交条件选择的基本参数和设计合适条件的指导由Sambrook,J., E.F.Fritsch和T.Manias(1989,Molecular Cloning:A Laboratory Manual,Cold Spring Harbor Laboratory Press,Cold Spring Harbor,N.Y.)描述。使用遗传密码的知识与本文所述的氨基酸序列组合,可以制备简并寡核苷酸组。这类寡核苷酸可用作例如聚合酶链反应(PCR)中的引物,由此分离且扩增DNA片段。在某些实施例中,简并引物可以用作基因文库的探针。这类文库将包括但不限于cDNA文库、基因组文库、以及甚至电子表达序列标签或DNA文库。通过这种方法鉴定的同源序列随后用作探针,以鉴定本文鉴定的序列的同源物。

[0134] 潜在用途还有聚核苷酸和寡核苷酸(例如引物或探针),其在降低的严格条件下(通常为中等严格条件,且通常为高度严格条件下)与如本文所述的聚核苷酸杂交。影响杂交条件选择的基本参数和用于设计合适条件的指导可以由本领域普通技术人员基于例如聚核苷酸的长度或碱基组成容易地确定。一种实现中等严格条件的方式涉及使用预洗涤溶液(所述预洗涤溶液含有5×标准柠檬酸钠、0.5%十二烷基硫酸钠、1.0mM乙二醇四乙酸(pH 8.0)),具有约50%甲酰胺、6×标准柠檬酸钠的杂交缓冲液,和约55℃的杂交温度(或其它相似的杂交溶液,如含有约50%甲酰胺的杂交溶液,伴随约42℃的杂交温度),以及在0.5×标准柠檬酸钠、0.1%十二烷基硫酸钠中,约60℃的洗涤条件。一般地,高度严格条件定义为如上的杂交条件,但使用在大约68℃下的洗涤、0.2×标准柠檬酸钠、0.1%十二烷基硫酸钠。SSPE(1×SSPE是0.15M氯化钠、10mM磷酸钠和1.25mM乙二醇四乙酸,pH 7.4)可以替代杂交和洗涤缓冲液中的标准柠檬酸钠(1×标准柠檬酸钠是0.15M氯化钠和15mM柠檬酸钠);洗涤在杂交完成后执行15分钟。应理解,通过应用控制杂交反应和双链体稳定性的基本原则,可以根据需要调整洗涤温度和洗涤盐浓度,以实现所需严格性程度,如本领域技术人员已知的和下文进一步描述的(参见例如,Sambrook,J.,E.F.Fritsch和T.Manias(1989,Molecular Cloning:A Laboratory Manual,Cold Spring Harbor Laboratory Press,Cold Spring Harbor,N.Y.)。当使聚核苷酸与未知序列的靶聚核苷酸杂交时,杂交物长度假定为杂交聚核苷酸的长度。当杂交已知序列的聚核苷酸时,可以通过比对聚核苷酸的序列并且鉴定一个或多个最佳序列互补性区域,来确定杂交物长度。预期长度小于50个碱基对的杂交物的杂交温度应比杂交物的解链温度低5至10℃,其中解链温度根据下述等式确定。对于长度小于18个碱基对的杂交物,解链温度(℃)=2(A+T碱基数目)+4(G+C碱基数目)。对于长度在18个碱基对以上的杂交物,解链温度(℃)=81.5+16.6(log₁₀[Na⁺])+0.41(%G+C)-(600/N),其中N是杂交物中的碱基数目,并且[Na⁺]是杂交缓冲液中的钠离子浓度(1×标准柠檬酸钠的[Na⁺]=0.165M)。通常,每种这类杂交聚核苷酸的长度是与它杂交的聚核苷酸的长度的至少25%(通常至少50%、60%或70%,并且最常至少80%),并且同与它杂交

的聚核苷酸具有至少60%序列一致性(例如至少70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%)。

[0135] 如本领域技术人员应理解,线性DNA具有两个可能定向:5'-至-3'方向和3'-至-5'方向。举例来说,如果参考序列以5'-至-3'方向放置,并且如果第二序列以5'-至-3'方向置于相同聚核苷酸分子/链中,那么参考序列和第二序列以相同方向定向,或具有相同定向。通常,启动子序列和处于给定启动子调节下的目的基因以相同定向放置。然而,就以5'-至-3'方向放置的参考序列而言,如果第二序列以3'-至-5'方向置于相同聚核苷酸分子/链中,那么参考序列和第二序列以反义方向定向,或具有反义定向。如果参考序列(5'-至-3'方向)和参考序列的反向互补序列(以5'-至-3'放置的参考序列)置于相同聚核苷酸分子/链中,那么就彼此而言具有反义定向的两个序列可以替代地描述为具有相同定向。本文所示的序列以5'-至-3'方向显示。

[0136] 本文提供的重组构建体可以用于转化植物或植物细胞,以便调节蛋白质表达和/或活性水平。重组聚核苷酸构建体可以包含编码如本文所述的一种或多种聚核苷酸的聚核苷酸,所述一种或多种聚核苷酸可操作地连接于适合表达多肽的调节区。因此,聚核苷酸可以包含编码如本文所述的多肽的编码序列。蛋白质表达和/或活性水平经调节的植物或植物细胞可以包括突变型、非天然存在的、转基因的、人造的或经基因工程改造的植物或植物细胞。合适地,植物或植物细胞包含已通过重组DNA的稳定整合而改变的基因组。重组DNA包括已在细胞外部经基因工程改造和构建的DNA,并且包括含有天然存在的DNA或cDNA或合成DNA的DNA。植物可以包括由最初转化的植物细胞再生的植物和来自经转化的植物的以后世代或杂交的后代植物。合适地,与对照植物相比较,修饰改变本文所述的聚核苷酸或多肽的表达或活性。

[0137] 由重组聚核苷酸编码的多肽可以是天然多肽,或对于细胞可以是异源的。在一些情况下,重组构建体含有可操作地连接于调节区的调节表达的聚核苷酸。在本文中描述了合适调节区的实例。

[0138] 还提供含有重组聚核苷酸构建体的载体,例如本文所述的那些。合适的载体主链包括例如本领域常规使用的那些,如质粒、病毒、人工染色体、细菌人工染色体、酵母人工染色体或噬菌体人工染色体。合适的表达载体包括但不限于源自例如噬菌体、杆状病毒和逆转录病毒的质粒和病毒载体。众多载体和表达系统是商购可得的。载体可以包括例如复制起点、支架附着区或标记。标记基因可以赋予植物细胞可选择表型。举例来说,标记可以赋予杀生物剂抗性,如对抗生素(例如卡那霉素(kanamycin)、G418、博来霉素(bleomycin)或潮霉素(hygromycin))或除草剂(例如草甘膦(glyphosate)、氯磺隆(chlorsulfuron)或草胺膦(phosphinothricin))的抗性。另外,表达载体可以包括设计为促进所表达多肽的操纵或检测(例如纯化或定位)的标签序列。标签序列,如荧光素酶、 β -葡糖醛酸酶、绿色荧光蛋白、谷胱甘肽S-转移酶、聚组氨酸、c-myc或血凝素序列,通常表达为与所编码多肽的融合物。这类标签可以插入多肽内的任何地方,包括在羧基或氨基末端处。

[0139] 植物或植物细胞可以通过具有重组聚核苷酸整合到其基因组进行转化,以变得稳定转化。本文所述的植物或植物细胞可以是稳定转化的。稳定转化的细胞通常伴随每次细胞分裂保留所引入的聚核苷酸。植物或植物细胞可以进行瞬时转化,使得重组聚核苷酸不整合到其基因组内。瞬时转化的细胞通常伴随每次细胞分裂失去所引入的重组聚核苷酸的

全部或一部分,使得在足够数目的细胞分裂后,所引入的重组聚核苷酸无法在子细胞中检测到。同样设想使用基因组编辑。

[0140] 用于转化植物细胞的许多方法是本领域可用的,所述方法全部涵盖在本文中,包括生物射弹、基因枪技术、土壤杆菌属 (*Agrobacterium*) 介导的转化、病毒载体介导的转化和电穿孔。用于将外源DNA整合到植物染色体内的土壤杆菌属系统已被广泛研究、修改和开发用于植物基因工程改造。通过常规方法将包含以有义或反义定向可操作地连接于调控序列的对应于主题纯化蛋白质的DNA序列的裸重组DNA分子连接至适当的T-DNA序列。通过聚乙二醇技术或电穿孔技术将这些引入原生质体内,所述两种技术都是标准的。或者,将包含编码主题纯化蛋白质的重组DNA分子的这类载体引入活土壤杆菌属细胞中,所述活土壤杆菌属细胞随后将DNA转移到植物细胞中。通过裸DNA而无伴随T-DNA载体序列的转化可以经由原生质体与含DNA脂质体的融合或经由电穿孔来完成。不伴随T-DNA载体序列的裸DNA也可以用于经由惰性、高速度微弹转化细胞。

[0141] 如果细胞或培养的组织用作转化的受体组织,那么需要时,通过本领域技术人员已知的技术,可以由经转化的培养物再生植物。

[0142] 有待包括在重组构建体中的调节区的选择取决于几个因素,包括但不限于效率、可选择性、可诱导性、所需表达水平和细胞或组织优先表达。通过适当选择调节区且相对于编码序列放置调节区,调节编码序列的表达对于本领域技术人员是常规工作。聚核苷酸的转录可以相似方式进行调节。一些合适的调节区仅或占优势地在某些细胞类型中起始转录。用于鉴定且表征植物基因组DNA中的调节区的方法是本领域已知的。

[0143] 启动子的实例包括由存在于不同组织或细胞类型中的组织特异性因子识别的组织特异性启动子(例如根特异性启动子、芽特异性启动子、木质部特异性启动子),或存在于不同发育阶段期间或响应不同环境条件存在。启动子的实例包括组成型启动子,其可以在大多数细胞类型中活化而不需要特定的诱导剂。用于控制RNAi多肽产生的启动子的实例包括花椰菜花叶病毒35S (CaMV/35S)、SSU、OCS、1b4、usp、STLS1、B33、nos或泛素启动子或菜豆蛋白启动子。本领域技术人员能够产生重组启动子的多种变体。

[0144] 组织特异性启动子是仅在植物发育期间的特定时间,在特定细胞或组织中(例如在营养组织或生殖组织中)活跃的转录控制元件。例如,当聚核苷酸在某些组织中的表达是优选的时,组织特异性表达可以是有利的。在发育控制下的组织特异性启动子的实例包括可以仅(或主要仅)在某些组织中起始转录的启动子,所述组织如营养组织,例如根或叶,或生殖组织,如果实、胚珠、种子、花粉、雌蕊 (pistil)、花或任何胚组织。生殖组织特异性启动子可以是例如花药特异性、胚珠特异性、胚特异性、胚乳特异性、珠被特异性、种子和种皮特异性、花粉特异性、花瓣特异性、萼片特异性或其组合。

[0145] 叶特异性启动子的实例包括来自C4植物(玉蜀黍)的丙酮酸正磷酸双激酶 (PPDK) 启动子、来自玉蜀黍的cab-m1Ca²⁺启动子、拟南芥 (*Arabidopsis thaliana*) myb相关基因启动子 (Atmyb5)、二磷酸核酮糖羧化酶 (RBCS) 启动子(例如,在叶和光生长幼苗中表达的番茄RBCS 1、RBCS2和RBCS3A基因,在发育中的番茄果实中表达的RBCS1和RBCS2,或几乎完全以高水平在叶片和叶鞘的叶肉细胞中表达的二磷酸核酮糖羧化酶启动子)。

[0146] 衰老特异性启动子的实例包括在果实催熟、叶枯萎和脱落期间活跃的番茄启动子、编码半胱氨酸蛋白酶的基因的玉蜀黍启动子、82E4的启动子和SAG基因的启动子。

[0147] 在某些实施例中,衰老特异性启动子是优选的。合适地,启动子可以是E4启动子。E4启动子已作为使用GUS作为报告基因的烟草的合适衰老启动子报告在Plant Mol Biol. (2008) 3月;66 (4) :415-27中。另一种示例性启动子是SAG12启动子,如Plant Cell. (1999) 6月;11 (6) :1073-80中所述。

[0148] 花药特异性启动子是其它实例。可以选择本领域技术人员已知的根优先的启动子。种子优先的启动子包括种子特异性启动子(在种子发育期间活跃的那些启动子,如种子贮藏蛋白质的启动子)和种子发芽启动子(在种子发芽期间活跃的那些启动子)。这类种子优先的启动子包括但不限于Cim1(细胞分裂素诱导的信息);cZ19B1(玉蜀黍19kDa玉米醇溶蛋白);milps(肌醇-1-磷酸合成酶);mZE40-2,也称为Zm-40;nuc1c;和cel1A(纤维素合成酶)。γ-玉米醇溶蛋白是胚乳特异性启动子。Glob-1是胚特异性启动子。对于双子叶植物,种子特异性启动子包括但不限于豆β-菜豆蛋白、油菜籽蛋白、ββ-伴大豆球蛋白、大豆凝集素、十字花科蛋白(cruciferin)等等。对于单子叶植物,种子特异性启动子包括但不限于玉蜀黍15kDa玉米醇溶蛋白启动子、22kDa玉米醇溶蛋白启动子、27kDa玉米醇溶蛋白启动子、g-玉米醇溶蛋白启动子、27kDa γ-玉米醇溶蛋白启动子(如gzw64A启动子,参见Genbank登记号S78780)、蜡质启动子、shrunk 1启动子、shrunk 2启动子、球蛋白1启动子(参见Genbank登记号L22344)、ltp2启动子、cim1启动子、玉蜀黍end1和end2启动子、nuc1启动子、Zm40启动子、eep1和eep2;lec1、硫氧还蛋白H启动子;m1ip15启动子、PCNA2启动子;和shrunk-2启动子。

[0149] 诱导型启动子的实例包括响应病原体攻击、厌氧条件、高温、光、干旱、寒冷温度或高盐浓度的启动子。病原体诱导型启动子包括来自发病机制相关蛋白质(PR蛋白质)的那些,所述发病机制相关蛋白质在通过病原体感染后诱导(例如PR蛋白质、SAR蛋白质、β-1,3-葡聚糖酶、壳多糖酶)。

[0150] 除植物启动子之外,其它合适的启动子可源自细菌来源(例如章鱼碱合成酶启动子、胭脂碱合成酶启动子和其它源自Ti质粒的启动子),或可源自病毒启动子(例如花椰菜花叶病毒(CaMV)的35S和19S RNA启动子、烟草花叶病毒的组成型启动子、花椰菜花叶病毒(CaMV) 19S和35S启动子或玄参花叶病毒35S启动子)。在某些实施例中,紫茉莉花叶病毒(MMV)启动子是优选的。在某些实施例中,35S启动子是优选的。

[0151] 在另一方面,提供一种经分离的多肽,其包含以下、由以下组成或基本上由以下组成:与本文所述的多肽序列中的任一种,包括序列表中所示的多肽中的任一种具有至少60%序列一致性的多肽序列。合适地,经分离的多肽包含以下、由以下组成或基本上由以下组成:与其具有至少60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、75%、76%、77%、78%、79%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%或100%序列一致性的序列。

[0152] 在一个实施例中,经分离的多肽包含以下、由以下组成或基本上由以下组成:与SEQ ID NO:2或SEQ ID NO:4或SEQ ID NO:6或SEQ ID NO:8或SEQ ID NO:10或SEQ ID NO:12具有至少60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、75%、76%、77%、78%、79%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%

或100%序列一致性的序列。在另一个实施例中,经分离的多肽包含以下、由以下组成或基本上由以下组成:与SEQ ID NO:2或SEQ ID NO:4或SEQ ID NO:6或SEQ ID NO:8或SEQ ID NO:10或SEQ ID NO:12具有至少78%、79%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%或100%序列一致性的序列。

[0153] 在一个实施例中,经分离的多肽包含以下、由以下组成或基本上由以下组成:与SEQ ID NO:20或SEQ ID NO:22或SEQ ID NO:24具有至少60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、75%、76%、77%、78%、79%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%或100%序列一致性的序列,条件是所述多肽在与SEQ ID NO:20或SEQ ID NO:22或SEQ ID NO:24中所示的位置等同的位置处包含终止密码子突变。

[0154] 在另一个实施例中,经分离的多肽包含以下、由以下组成或基本上由以下组成:与SEQ ID NO:20或SEQ ID NO:22或SEQ ID NO:24具有至少78%、79%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%或100%序列一致性的序列,条件是所述多肽在与SEQ ID NO:20或SEQ ID NO:22或SEQ ID NO:24中所示的位置等同的位置处包含终止密码子突变。在另一个实施例中,提供一种多肽变体,其包含以下、由以下组成或基本上由以下组成:由与SEQ ID NO:2或SEQ ID NO:4或SEQ ID NO:6或SEQ ID NO:8或SEQ ID NO:10或SEQ ID NO:12具有至少约60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、75%、76%、77%、78%、79%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%序列一致性的聚核苷酸变体编码的氨基酸序列。

[0155] 在另一个实施例中,提供一种多肽变体,其包含以下、由以下组成或基本上由以下组成:由与SEQ ID NO:20或SEQ ID NO:22或SEQ ID NO:24具有至少约60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、75%、76%、77%、78%、79%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%序列一致性的聚核苷酸变体编码的氨基酸序列,条件是所述多肽在与SEQ ID NO:20或SEQ ID NO:22或SEQ ID NO:24中所示的位置等同的位置处包含终止密码子突变。在另一个实施例中,提供SEQ ID NO:2或SEQ ID NO:4或SEQ ID NO:6或SEQ ID NO:8或SEQ ID NO:10或SEQ ID NO:12的多肽的片段或SEQ ID NO:2或SEQ ID NO:4或SEQ ID NO:6或SEQ ID NO:8或SEQ ID NO:10或SEQ ID NO:12的片段,其与SEQ ID NO:2或SEQ ID NO:4或SEQ ID NO:6或SEQ ID NO:8或SEQ ID NO:10或SEQ ID NO:12的相应片段具有至少约60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、75%、76%、77%、78%、79%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%或100%序列一致性。

[0156] 在另一个实施例中,提供SEQ ID NO:20或SEQ ID NO:22或SEQ ID NO:24的多肽的片段或SEQ ID NO:20或SEQ ID NO:22或SEQ ID NO:24的片段,其与SEQ ID NO:20或SEQ ID

NO:22或SEQ ID NO:24的相应片段具有至少约60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、75%、76%、77%、78%、79%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%或100%序列一致性,条件是所述多肽在与SEQ ID NO:20或SEQ ID NO:22或SEQ ID NO:24中所示的位置等同的位置处包含终止密码子突变。在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的多肽的活性:编码天冬酰胺合成酶并且与SEQ ID NO:2或SEQ ID NO:4或SEQ ID NO:6或SEQ ID NO:8具有至少78%序列一致性的序列。在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的多肽的活性:编码天冬酰胺合成酶并且与SEQ ID NO:2和SEQ ID NO:4和SEQ ID NO:6和SEQ ID NO:8具有至少78%序列一致性的序列。在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的多肽的活性:编码天冬酰胺合成酶并且与SEQ ID NO:2和SEQ ID NO:4具有至少78%序列一致性的序列。在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的多肽的活性:编码天冬酰胺合成酶并且与SEQ ID NO:6和SEQ ID NO:8具有至少78%序列一致性的序列。在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的多肽的活性:编码天冬酰胺合成酶并且与SEQ ID NO:2和SEQ ID NO:4和SEQ ID NO:6和/或SEQ ID NO:8具有至少78%序列一致性的序列。在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的多肽的活性:编码天冬酰胺合成酶并且与SEQ ID NO:2和/或SEQ ID NO:4和SEQ ID NO:6和SEQ ID NO:8具有至少78%序列一致性的序列。

[0157] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的多肽的活性:编码天冬酰胺合成酶并且与SEQ ID NO:20或SEQ ID NO:22或SEQ ID NO:24具有至少78%序列一致性的序列,条件是所述多肽在与SEQ ID NO:20或SEQ ID NO:22或SEQ ID NO:24中所示的位置等同的位置处包含终止密码子突变。

[0158] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的多肽的活性:编码天冬酰胺合成酶并且与SEQ ID NO:20和SEQ ID NO:22和SEQ ID NO:24具有至少78%序列一致性的序列,条件是所述多肽在与SEQ ID NO:20和SEQ ID NO:22和SEQ ID NO:24中所示的位置等同的位置处包含终止密码子突变。

[0159] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的多肽的活性:编码天冬酰胺合成酶并且与SEQ ID NO:20和SEQ ID NO:22具有至少78%序列一致性的序列,条件是所述多肽在与SEQ ID NO:20和SEQ ID NO:22中所示的位置等同的位置处包含终止密码子突变。

[0160] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的多肽的活性:编码天冬酰胺合成酶并且与SEQ ID NO:22和SEQ ID NO:24具有至少78%序列一致性的序列,条件是所述多肽在与SEQ ID NO:22和SEQ ID NO:24中所示的位置等同的位置处包含终止密码子突变。

[0161] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的多肽的活性:编码天冬酰胺合成酶并且与SEQ ID NO:20和SEQ ID NO:24具有至少78%序列一致性的序列,条件是所述多肽在与SEQ ID NO:20和SEQ ID NO:24中所示的位置等同的位置处包含终止密码子突变。

[0162] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的多肽的活性:

编码天冬酰胺合成酶并且与SEQ ID NO:20和SEQ ID NO:22和/或SEQ ID NO:24具有至少78%序列一致性的序列,条件是所述多肽在与SEQ ID NO:20和SEQ ID NO:22和/或SEQ ID NO:24中所示的位置等同的位置处包含终止密码子突变。

[0163] 在某些实施例中,调节包含以下、由以下组成或基本上由以下组成的多肽的活性:编码天冬酰胺合成酶并且与SEQ ID NO:20和/或SEQ ID NO:22和SEQ ID NO:24具有至少78%序列一致性的序列,条件是所述多肽在与SEQ ID NO:20和/或SEQ ID NO:22和SEQ ID NO:24中所示的位置等同的位置处包含终止密码子突变。所述多肽可以包括包含足够或相当大程度的一致性 or 相似性的序列的片段以充当天冬酰胺合成酶。多肽的片段通常保留全长序列的活性的一些或全部,如至少约95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%或100%的活性。

[0164] 如本文所讨论,多肽还包括通过引入任何类型的改变(例如,氨基酸的插入、缺失或置换;糖基化状态的变化;影响重折叠或异构化、三维结构或自缔合状态的变化)而产生的突变体,条件是它们仍具有其作为天冬酰胺合成酶的一些或全部功能或活性。合适地,调节、降低或抑制作为天冬酰胺合成酶的功能或活性。合适地,作为天冬酰胺合成酶的功能或活性被抑制,使得天冬酰胺合成酶活性是不可检测的。

[0165] 多肽包括通过引入任何类型的改变(例如,氨基酸的插入、缺失或置换;糖基化状态的变化;影响重折叠或异构化、三维结构或自缔合状态的变化)而产生的变体,所述变体可以是有意改造或天然分离的。变体可以具有产生沉默变化且导致功能等效蛋白质的改变。有意的氨基酸置换可以在残基的极性、电荷、溶解性、疏水性、亲水性和两亲性质中的相似性基础上作出,只要物质的二级结合活性被保留。例如,带负电的氨基酸包括天冬氨酸和谷氨酸;带正电的氨基酸包括赖氨酸和精氨酸;并且具有相似亲水性值含不带电极性首基的氨基酸包括亮氨酸、异亮氨酸、缬氨酸、甘氨酸、丙氨酸、天冬酰胺、谷氨酰胺、丝氨酸、苏氨酸、苯丙氨酸和酪氨酸。保守置换可以例如根据下表进行。第二列中的相同块和优选第三列中的相同行中的氨基酸可以彼此置换:

[0166]

脂肪族	非极性	Gly Ala Pro Ile Leu Val
	极性 - 不带电	Cys Ser Thr Met Asn Gly
	极性 - 带电	Asp Glu Lys Arg
芳香族		His Phe Trp Tyr

[0167] 多肽可以是成熟蛋白质或不成熟蛋白质或源自不成熟蛋白质的蛋白质。多肽可以采取线性形式或使用已知方法环化。多肽通常包含至少10个、至少20个、至少30个、或至少40个、或至少50个、或至少100个、或至少200个、或至少300个、或至少400个、或至少500个、或至少600个邻接氨基酸。

[0168] 还公开了由SEQ ID NO:2或SEQ ID NO:4或SEQ ID NO:6或SEQ ID NO:8或SEQ ID NO:10或SEQ ID NO:12或SEQ ID NO:20或SEQ ID NO:22或SEQ ID NO:24编码的与其具有100%序列一致性的多肽,或包含以下、由以下组成或基本上由以下组成的多肽:SEQ ID

NO:2或SEQ ID NO:4或SEQ ID NO:6或SEQ ID NO:8或SEQ ID NO:10或SEQ ID NO:12或SEQ ID NO:20或SEQ ID NO:22或SEQ ID NO:24中所示的序列。

[0169] 多肽可以通过在适合表达多肽的培养条件下培养经转化的或重组宿主细胞进行制备。所得到的表达多肽随后可以使用已知的纯化过程由这类培养物进行纯化。多肽的纯化可以包括含有与多肽结合的试剂的亲合柱；一个或多个在这类亲和树脂上的柱步骤；一个或多个涉及疏水作用色谱的步骤；或免疫亲和色谱。或者，多肽还可以促进纯化的形式表达。举例来说，它可以表达为融合多肽，如麦芽糖结合多肽、谷胱甘肽-5-转移酶、组氨酸标签或硫氧还蛋白的那些。用于融合多肽的表达和纯化的试剂盒是商购可得的。多肽可以用表位加标签，并且随后通过使用针对这类表位的特异性抗体纯化。一个或多个液相色谱步骤-如反相高效液相色谱可以用于进一步纯化多肽。前述纯化步骤中的一些或全部以各种组合可用于提供基本上均质的重组多肽。由此纯化的多肽可以基本上不含其它多肽，并且在本文中定义为“基本上纯化的多肽”；这类纯化的多肽包括多肽、片段、变体等等。多肽和片段的表达、分离和纯化可以通过任何合适技术实现，所述合适技术包括但不限于本文所述的方法。

[0170] 也可以使用亲和柱(如针对多肽产生的单克隆抗体)亲和纯化所表达的多肽。这些多肽可以使用常规技术，例如在高盐洗脱缓冲液中，从亲和柱中取出，且随后透析到更低盐的缓冲液内用于使用，或取决于利用的亲和基质，通过改变pH或其它组分，或使用天然存在的亲和部分的底物竞争性取出。

[0171] 公开了分离的或基本上纯化的聚核苷酸或蛋白质组合物。“分离的”或“纯化的”聚核苷酸或蛋白质或其生物学上的活性部分基本上或大体上不含通常伴随聚核苷酸或蛋白质或者与其相互作用(如在其天然存在的环境中发现的那样)的组分。因此，分离的或纯化的聚核苷酸或蛋白质在通过重组技术产生时，基本上不含其它细胞物质或培养基，或当以化学方式合成时，基本上不含化学前体或其它化学品。最佳地，“分离的”聚核苷酸不含所述聚核苷酸的来源生物体的基因组DNA中天然地侧接于所述聚核苷酸的序列(最佳地，蛋白质编码序列)(例如，位于聚核苷酸的5'和3'末端的序列)。举例来说，在各种实施例中，分离的聚核苷酸可以含有少于约5kb、4kb、3kb、2kb、1kb、0.5kb或0.1kb的核苷酸序列，所述核苷酸序列在所述聚核苷酸的来源细胞的基因组DNA中天然地侧接于所述聚核苷酸。基本上不含细胞物质的蛋白质包括具有小于约30%、20%、10%、5%或1%(以干重计)污染性蛋白质的蛋白质制备物。

[0172] 多肽还可以通过已知的常规化学合成来产生。通过合成方式用于构建多肽或其片段的方法是本领域技术人员已知的。由于与天然多肽共享一级、二级或三级结构或构象特征，合成构建的多肽序列可以具有与之共同的生物特性，包括生物活性。

[0173] 遗传背景中的差异可以通过表型差异或本领域已知的分子生物学技术进行检测，所述分子生物学技术如核酸测序、遗传标记(例如微卫星RNA标记)的存在或不存在。

[0174] 还提供了与本文所述的多肽具有免疫反应性的抗体。如本文所述的多肽、片段、变体、融合多肽等等可以用作产生与其免疫反应的抗体的“免疫原”。这类抗体可以经由抗体的抗原结合位点与多肽特异性结合。特异性结合抗体是特异性识别且结合多肽、同源物和变体，而不是其它分子的那些。在一个实施例中，抗体对于具有与如本文所示的氨基酸序列的多肽是特异性的，并且不与其它多肽交叉反应。

[0175] 更具体来说,多肽、片段、变体、融合多肽等等含有引发抗体形成的抗原决定簇或表位。这些抗原决定簇或表位可以是线性的或构象的(不连续的)。线性表位由多肽的单个氨基酸区段组成,而构象或不连续表位由来自多肽链的不同区域的氨基酸区段组成,所述氨基酸区段在多肽折叠后达到紧密接近。表位可以通过本领域已知的方法中的任一种进行鉴定。另外,来自多肽的表位可以用作分析中的研究试剂,且纯化来自如多克隆血清或来自培养杂交瘤的上清液等物质的特异性结合抗体。这类表位或其变体可以使用本领域已知技术如固相合成、多肽的化学或酶促切割或使用重组DNA技术产生。

[0176] 针对多肽的多克隆抗体和单克隆抗体都可以通过常规技术进行制备。本文还涵盖了产生特异性针对多肽的单株抗体的杂交瘤细胞系。这类杂交瘤可以通过常规技术进行生产且鉴定。对于抗体生产,多种宿主动物可以通过用多肽、其片段、变体或突变体注射进行免疫接种。仅举几个例子,这类宿主动物可以包括但不限于兔、小鼠和大鼠。多种佐剂可以用于增加免疫应答。取决于宿主物种,这类佐剂包括但不限于弗氏(完全和不完全)、矿物凝胶如氢氧化铝,表面活性物质如溶血卵磷脂、普鲁兰尼克多元醇、聚阴离子、肽、油乳剂、钥孔血蓝蛋白、二硝基苯酚,以及潜在有用的人类佐剂如BCG(卡介苗)和短小棒状杆菌(*Corynebacterium parvum*)。单克隆抗体可以通过常规技术进行回收。这类单克隆抗体可以属于任何免疫球蛋白类别,包括IgG、IgM、IgE、IgA、IgD及其任何亚类。

[0177] 抗体还可以用于分析中,以在体外或体内检测多肽或片段的的存在。抗体还可以用于通过免疫亲和色谱纯化多肽或片段。

[0178] 还公开了聚核苷酸和由其编码的多肽的片段。聚核苷酸的片段可以编码保留天然蛋白质的生物活性的蛋白质片段。或者,适用作杂交探针或PCR引物的聚核苷酸的片段通常不编码保留生物活性的片段蛋白质。此外,所公开的核苷酸序列的片段包括可以在本文所论述的重组构建体内组装的那些。聚核苷酸序列的片段可以在至少约25个核苷酸、约50个核苷酸、约75个核苷酸、约100个核苷酸、约150个核苷酸、约200个核苷酸、约250个核苷酸、约300个核苷酸、约400个核苷酸、约500个核苷酸、约600个核苷酸、约700个核苷酸、约800个核苷酸、约900个核苷酸、约1000个核苷酸、约1100个核苷酸、约1200个核苷酸、约1300个核苷酸、或约1500个核苷酸、约2000个核苷酸、约3000个核苷酸、约4000个核苷酸、约5000个核苷酸、约6000个核苷酸、约7000个核苷酸、约8000个核苷酸、约9000个核苷酸、

[0179] 约10000个核苷酸、约15000个核苷酸、约20000个核苷酸并且至多编码本文所述的多肽的全长聚核苷酸的范围内。多肽序列的片段可以在至少约25个氨基酸、约50个氨基酸、约75个氨基酸、约100个氨基酸、约150个氨基酸、约200个氨基酸、约250个氨基酸、约300个氨基酸、约400个氨基酸、约500个氨基酸、约600个氨基酸或至多本文所述的全长多肽的范围内。

[0180] 应理解,可以通过多种方式来实现降低本文所述的天冬酰胺合成酶的表达或活性。根据某些实施例,降低天冬酰胺合成酶的表达可以使用干扰转录和/或翻译的多种分子在基因组和/或转录物水平进行,所述分子包括但不限于反义、siRNA、核酶或脱氧核酶分子。也可以使用向至少一个基因插入一个或多个突变,包括缺失、插入、位点特异性突变、锌指核酸酶等。根据其它实施例,可以使用拮抗剂或裂解多肽等的酶在蛋白质水平抑制表达。

[0181] 在一个方面,描述了一种突变植物或其部分,包含至少一个突变和(i)包含以下、由以下组成或基本上由以下组成的聚核苷酸:编码天冬酰胺合成酶并且与SEQ ID NO:1或

SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7具有至少72%序列一致性的序列;(ii)由(i)中所示的聚核苷酸编码的多肽;(iii)包含以下、由以下组成或基本上由以下组成的多肽:编码天冬酰胺合成酶并且与SEQ ID NO:2或SEQ ID NO:4或SEQ ID NO:6或SEQ ID NO:8具有至少78%序列一致性的序列;或(iv)包含(i)中所示的经分离的聚核苷酸的构建体、载体或表达载体,其中与不包含所述至少一个突变的对照植物相比,所述至少一个突变降低天冬酰胺合成酶的表达或活性。在另一方面,描述了一种突变植物或其部分,包含至少一个突变和(i)包含以下、由以下组成或基本上由以下组成的聚核苷酸:编码天冬酰胺合成酶并且与SEQ ID NO:1具有至少90%序列一致性或与SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7具有至少72%序列一致性的序列;(ii)由(i)中所示的聚核苷酸编码的多肽;(iii)包含以下、由以下组成或基本上由以下组成的多肽:编码天冬酰胺合成酶并且与SEQ ID NO:2或SEQ ID NO:4或SEQ ID NO:6或SEQ ID NO:8具有至少78%序列一致性的序列;或(iv)包含(i)中所示的经分离的聚核苷酸的构建体、载体或表达载体,其中与不包含所述至少一个突变的对照植物相比,所述至少一个突变降低天冬酰胺合成酶的表达或活性。

[0182] 植物或植物细胞因此可以包含NtASN1-S(SEQ ID NO:1)和/或NtASN1-T(SEQ ID NO:3)和/或NtASN3-S(SEQ ID NO:9)和/或NtASN3-T(SEQ ID NO:11)和/或NtASN5-S(SEQ ID NO:5)和/或NtASN5-T(SEQ ID NO:7)中的一个或多个突变,其中所述突变使得所述基因的表达降低或由其编码的蛋白质的功能降低。植物或植物细胞可以如所述包含NtASN1-S(SEQ ID NO:1)或NtASN1-T(SEQ ID NO:3)或NtASN5-S(SEQ ID NO:5)或NtASN5-T(SEQ ID NO:7)中的一个或多个突变,其中所述突变使得所述基因的表达降低或由其编码的蛋白质的功能降低。植物或植物细胞可以如所述包含NtASN1-S(SEQ ID NO:1)和NtASN1-T(SEQ ID NO:3)或NtASN5-S(SEQ ID NO:5)和NtASN5-T(SEQ ID NO:7)中的一个或多个突变(例如,如本文所述的无义突变等),其中所述突变使得所述基因的表达降低或由其编码的蛋白质的功能降低。可以调节、抑制或降低突变体的表达或功能。除本文所述的一个或多个突变以外,突变植物或植物细胞可以在一个或多个其它基因或多肽中具有一个或多个其它突变。在某些实施例中,突变体可以在一个或多个其它基因或多肽具有一个或多个其它突变。本文中描述了NtASN1-S、NtASN1-T、NtASN5-S和ASN5-T的突变体。在某些实施例中,一个或多个突变在谷氨酰胺酶结构域中。

[0183] 在ASN1-S(ASN1-S_W156*-SEQ ID No 19和20)、ASN1-T(ASN1-T_W156*-SEQ ID No 21和22)和ASN5-S(ASN5-S_Q66*-SEQ ID No 23和24)拷贝中鉴定出三种终止突变,但在ASN5-T拷贝中没有。由于在ASN5-T中没有鉴定出终止突变,因此将无终止突变鉴定为潜在候选者。ASN1-S_W156*、ASN1-T_W156*和ASN5-S_Q66*与其相应的WT分离体相比存在明显的天冬酰胺减少倾向,但ASN5-T_G59D不存在。ASN1-T_W156*(83%)观察到最强减少,然后是ASN5-S_Q66*(44%)并且最后是ASN1-S_W156*(17%)。品系ASN5-T_G59D显示天冬酰胺略微增加。ASN1-S和ASN5-S中的终止突变对植物生物质和尺寸没有影响。根据Canales等人., 2012,所有Nt-ASN基因中的突变位于从谷氨酰胺释放氨的谷氨酰胺酶结构域、定位在Nt-ASN中从氨基酸R211至D451共同区域的合成酶结构域中。

[0184] 所述突变植物或植物细胞的突变可以是杂合或纯合的。所述突变植物或植物细胞的至少一个突变可以是杂合的,并且至少一个不同突变是纯合的。合适地,突变植物或植物细胞的突变是纯合的。

[0185] 突变体可以在与SEQ ID NO:2的氨基酸位置156等同的位置处具有突变(例如,产生终止密码子的无义突变)。这类突变体的实例是如本文所述的ASN1-S_W156*。

[0186] 突变体可以在与SEQ ID NO:4的氨基酸位置156等同的位置处具有突变(例如,产生终止密码子的无义突变)。这类突变体的实例是如本文所述的ASN1-T_W156*。

[0187] 突变体可以在与SEQ ID NO:6的氨基酸位置66等同的位置处具有突变(例如,产生终止密码子的无义突变)。这类突变体的实例是如本文所述的ASN5-S_Q66*。

[0188] 还涵盖了这些突变的组合。所述组合可以包括ASN1-S_W156*和ASN1-T_W156*的组合;或ASN1-S_W156*和ASN5-S_Q66*的组合;或ASN1-T_W156*和ASN5-S_Q66*的组合;或ASN1-S_W156*、ASN1-T_W156*和ASN5-S_Q66*的组合。

[0189] 在另一方面,提供一种用于降低植物或源自植物的植物材料中天冬酰胺的水平的方法,所述方法包含将一个或多个降低至少一个天冬酰胺合成酶基因表达的突变引入到所述植物的基因组中,其中所述至少一个天冬酰胺合成酶基因编码NtASN1-S、NtASN1-T、NtASN5-S或NtASN5-T。合适地,除NtASN1-S、NtASN1-T、NtASN5-S或NtASN5-T中的突变之外,还可以将一个或多个突变引入到至少一个、两个或三个或更多个其它天冬酰胺合成酶基因的至少一个等位基因中。还涵盖了NtASN1-S和NtASN1-T或NtASN5-S和NtASN5-T中的突变。

[0190] 还提供一种用于鉴定天冬酰胺水平降低的植物的方法,所述方法包含从目的植物筛选存在NtASN1-S、NtASN1-T、NtASN5-S或ASN5-T中的一个或多个突变的核酸样品。合适地,所述方法另外包含从所述目的植物筛选所述核酸样品或存在NtASN1-S中的突变、存在NtASN1-T中的突变、或存在NtASN5-S中的突变、或存在NtASN5-T中的突变的其它核酸样品。

[0191] 还公开了一种植物或植物细胞,其编码NtASN1-S、NtASN1-T、NtASN5-S或NtASN5-T的基因中的一个或多个突变是杂合或纯合的,其中所述突变使得所述基因的表达降低或由其编码的蛋白质的功能降低。

[0192] 在一些实施例中,使用诱变法将有利突变引入植物或植物细胞中,并且使用本领域技术人员已知的方法(如DNA印迹分析法、DNA测序、PCR分析法或表型分析法)鉴定或选择所引入的突变。可以使用本领域熟知的方法来确定影响基因表达或干扰所编码的蛋白质的功能的突变。基因外显子中的插入突变通常导致空突变。保守残基中的突变可以在抑制或降低所编码的蛋白质的代谢功能中特别有效。

[0193] 还公开了用于获得突变型聚核苷酸和多肽的方法。任何目的植物,包括植物细胞或植物材料,可以通过多种已知诱导诱变的方法进行遗传修饰,所述方法包括定点诱变、寡核苷酸指导的诱变、化学诱导的诱变、辐射诱导的诱变、利用经修饰的碱基的诱变、利用缺口双链体DNA的诱变、双链断裂诱变、利用修复缺陷型宿主株的诱变、通过全基因合成的诱变、DNA改组和其它等效方法。

[0194] 突变型多肽变体可以用于制备包含一种或多种突变型多肽变体的突变型、非天然存在的或转基因植物(例如,突变型、非天然存在的、转基因、人造或基因工程改造的植物)。合适地,突变型多肽变体保留未突变多肽的活性。与未突变的多肽相比较,突变型多肽变体的活性可更高、更低或大约相同。

[0195] 本文所述的核苷酸序列和多肽中的突变可以包括人造突变或合成突变或基因工程改造的突变。本文所述的核苷酸序列和多肽中的突变可以通过包括体外或体内操纵步

骤的过程获得或可获得的突变。本文所述的核苷酸序列和多肽中的突变可以通过包括人为干预的过程获得或可获得的突变。作为实例,所述过程可以包括使用外源添加的化学品的诱变,如致突变、致畸形或致癌有机化合物,例如甲磺酸乙酯(EMS),所述化学品在遗传物质中产生随机突变。进一步举例来说,所述过程可以包括一个或多个基因工程改造步骤-如本文所述的基因工程改造步骤中的一个或多个或其组合。进一步举例来说,所述过程可以包括一个或多个植物杂交步骤。

[0196] 如果转化活性在统计学上低于未经修饰以抑制天冬酰胺合成酶多肽的转化活性且已使用相同方案培养和收获的植物中的相同天冬酰胺合成酶多肽的转化活性,那么根据本公开,植物中的一个或多个天冬酰胺合成酶多肽的活性被降低或抑制。当无法通过本文所述的分析方法检测时,考虑消除植物中天冬酰胺合成酶多肽的活性。本文中描述了测定天冬酰胺合成酶多肽的活性的方法。

[0197] 除诱变以外,可以调节本文所述的一种或多种聚核苷酸或多肽的表达或活性的组合物包括但不限于:可以干扰一种或多种内源基因的转录的序列特异性聚核苷酸;可以干扰RNA转录物翻译的序列特异性聚核苷酸(例如双链RNA、siRNA、核酶);可以干扰一种或多种蛋白质的稳定性的序列特异性多肽;可以干扰一种或多种蛋白质的酶促活性或者一种或多种蛋白质相对于底物或调节蛋白质的结合活性的序列特异性聚核苷酸;显示出对于一种或多种蛋白质的特异性的抗体;可以干扰一种或多种蛋白质的稳定性、或者一种或多种蛋白质的酶促活性、或者一种或多种蛋白质的结合活性的小分子化合物;结合一种或多种聚核苷酸的锌指蛋白;以及具有针对一种或多种聚核苷酸的活性的大范围核酸酶。基因编辑技术、遗传编辑技术和基因组编辑技术是本领域众所周知的。

[0198] 一种基因编辑方法涉及转录激活因子样效应物核酸酶(transcription activator-like effector nuclease,TALEN)的使用,其诱导细胞可以修复机制响应的双链断裂。非同源末端连接使来自双链断裂任一侧的DNA再连接,其中存在很少的序列重叠或无序列重叠用于退火。这一修复机制经由插入或缺失、或染色体重排诱导基因组中的错误。任何这类误差可以致使在该位置处编码的基因产物无功能。另一种基因编辑方法涉及细菌CRISPR/Cas系统的使用。细菌和古细菌显示出称为规律成簇间隔短回文重复(clustered regularly interspaced short palindromic repeat,CRISPR)的染色体元件,它是适应性免疫系统的一部分,防止侵入病毒和质粒DNA。在II型CRISPR系统中,CRISPR RNA(crRNA)与反式激活crRNA(tracrRNA)和CRISPR相关(Cas)蛋白质起作用,以在靶DNA中引入双链断裂。通过Cas9的靶切割要求在crRNA和tracrRNA之间的碱基配对,以及在crRNA和靶DNA之间的碱基配对。靶识别通过称为原型间隔序列毗邻基序(proto-spacer-adjacent motif,PAM)的短基序的存在得到促进,所述PAM符合序列NGG。这一系统可以用于基因组编辑。Cas9通常通过双重RNA按程序工作,所述双重RNA由crRNA和tracrRNA组成。然而,这些RNA的核心组分可以组合成单一杂交物‘引导RNA’用于Cas9靶向。对靶DNA使用非编码RNA引导用于位点特异性切割有希望比现有技术(如TALEN)明显更直截了当。使用CRISPR/Cas策略,再靶向核酸酶复合物仅要求引入新RNA序列,并且不需要再改造蛋白质转录因子的特异性。

[0199] 反义技术是可以用于调节多肽表达的另一熟知方法。将待阻遏基因的聚核苷酸克隆且可操作地连接于调节区和转录终止序列,使得RNA的反义链被转录。重组构建体随后转化到植物细胞内,并且产生RNA的反义链。聚核苷酸无需是待阻遏基因的整个序列,但通常

与待阻遏基因的有义链的至少一部分基本上互补。

[0200] 聚核苷酸可以转录成核酶,或催化RNA,其影响mRNA的表达。核酶可以设计为与几乎任何靶RNA特异性配对,且切割在特定位置处的磷酸二酯主链,由此使靶RNA功能失活。异源聚核苷酸可以编码设计为切割特定mRNA转录物的核酶,从而阻止多肽的表达。锤头状核酶可用于破坏特定mRNA,尽管可以使用在位点特异性识别序列处切割mRNA的多种核酶。锤头状核酶在由侧翼区指示的位置处切割mRNA,所述侧翼区与靶mRNA形成互补碱基对。唯一要求是靶RNA含有5'-UG-3'核苷酸序列。锤头状核酶的构建和产生是本领域已知的。锤头状核酶序列可以嵌入稳定RNA如转移RNA(tRNA)内,以增加体内切割效率。

[0201] 在一个实施例中,可以干扰RNA转录物翻译的序列特异性聚核苷酸是干扰RNA。RNA干扰或RNA沉默是进化上保守的过程,特异性mRNA通过其可被靶向用于酶促降解。一种双链RNA(双链RNA)通过细胞(例如双链RNA病毒、或干扰RNA聚核苷酸)引入或产生,以起始干扰RNA途径。双链RNA可以通过RNA酶III而转化成21-23bp长度的多个小干扰RNA双链体,所述RNA酶III是双链RNA特异性核酸内切酶。小干扰RNA随后可以由RNA诱导沉默复合物识别,所述RNA诱导沉默复合物促进通过ATP依赖性过程的小干扰RNA的解旋。小干扰RNA的解旋的反义链将活化RNA诱导沉默复合物导向靶向mRNA,所述靶向mRNA包含与小干扰RNA反义链互补的序列。靶向mRNA和反义链可以形成A形螺旋,并且A形螺旋的大沟可以由活化RNA诱导沉默复合物识别。可以在由小干扰RNA链的5'-端的结合位点限定的单个位点处,通过活化RNA诱导沉默复合物切割靶mRNA。活化RNA诱导沉默复合物可以再循环,以催化另一切割事件。来自NtASN1-S可以用于使NtASN1-S和NtASN1-T拷贝沉默的编码序列的实例显示在SEQ ID NO:17中。因此,另一方面涉及SEQ ID NO:17中所示的核苷酸序列或与其具有至少72%、73%、74%、75%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%序列一致性的序列。来自NtASN5-T可以用于使NtASN1-S和NtASN1-T拷贝沉默的编码序列的实例显示在SEQ ID NO:18中。因此,另一方面涉及SEQ ID NO:17中所示的序列。因此,另一方面涉及SEQ ID NO:18中所示的核苷酸序列或与其具有至少72%、73%、74%、75%、80%、85%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%序列一致性的序列。还涵盖了使用这些编码序列的基因沉默方法及其用途。

[0202] 干扰RNA表达载体可以包含编码干扰RNA聚核苷酸的干扰RNA构建体,其通过降低mRNA、mRNA前体或相关RNA变体的表达水平,显示出RNA干扰活性。图22中示出了示例性构建体。表达载体可以包含置于干扰RNA构建体上游且可操作地连接于干扰RNA构建体的启动子,如本文进一步描述的。干扰RNA表达载体可以包含合适的最小核心启动子、目的干扰RNA构建体、上游(5')调节区、下游(3')调节区,包括转录终止和聚腺苷酸化信号,以及本领域技术人员已知的其它序列,如多种选择标记。

[0203] 聚核苷酸可以多种形式产生,包括作为双链结构(即,包含反义链和互补有义链的双链RNA分子)、双链发夹样结构、或单链结构(即,仅包含反义链的ssRNA分子)。结构可以包含具有自互补的有义链和反义链的双链体、不对称双链体、发夹或不对称的发夹二级结构。双链干扰RNA可以酶促转化为双链小干扰RNA。小干扰RNA双链体的链之一可对靶mRNA和相关RNA变体内的互补序列退火。小干扰RNA/mRNA双链体由RNA诱导沉默复合物识别,所述RNA诱导沉默复合物可以序列依赖性方式在多个位点处切割RNA,从而导致靶mRNA和相关RNA变体的降解。

[0204] 双链RNA分子可以包括以茎环结构由单个寡核苷酸装配的小干扰RNA分子,其中小干扰RNA分子的自互补的有义区和反义区借助于基于聚核苷酸或基于非聚核苷酸的接头以及具有两个或更多个环结构和包含自互补的有义链和反义链的茎的环状单链RNA连接,其中所述环状RNA可以在体内或体外加工,以生成能够介导干扰RNA的活性小干扰RNA分子。

[0205] 还涵盖了小发夹RNA分子的使用。除了反向互补(有义)序列外,它们还包含特定的反义序列,通常由间隔物或环序列分开。间隔物或环的切割提供了单链RNA分子及其反向互补序列,使得它们可以退火以形成双链RNA分子(任选具有另外的加工步骤,可以导致来自任一或两条链的3'端或5'端的一个、两个、三个或更多个核苷酸的添加或删除)。间隔物可以具有足够长度,以在间隔物切割(和任选地,可以导致来自任一或两条链的3'端或5'端的一个、两个、三个、四个或更多个核苷酸的添加或删除的后续加工步骤)之前,允许反义和有义序列退火且形成双链结构(或茎)。间隔序列通常是位于两个互补的核苷酸序列区域之间的无关核苷酸序列,当退火成双链聚核苷酸时,所述两个互补的核苷酸序列区域包含小发夹RNA。间隔序列一般包含约3至约100个核苷酸。

[0206] 可以通过选择用于产生发夹双链体的合适序列组成、环大小和茎长,产生任何目的RNA聚核苷酸。用于设计发夹双链体的茎长的合适范围包括至少约10、11、12、13、14、15、16、17、18、19或20个核苷酸的茎长,如约14-30个核苷酸、约30-50个核苷酸、约50-100个核苷酸、约100-150个核苷酸、约150-200个核苷酸、约200-300个核苷酸、约300-400个核苷酸、约400-500个核苷酸、约500-600个核苷酸以及约600-700个核苷酸。用于设计发夹双链体的环长的合适范围包括约4-25个核苷酸、约25-50个核苷酸或如果发夹双链体的茎长相当大,那么更长的环长。在某些实施例中,双链RNA或ssRNA分子长度在约15和约40个核苷酸之间。在另一个实施例中,小干扰RNA分子是长度在约15和约35个核苷酸之间的双链RNA或ssRNA分子。在另一个实施例中,小干扰RNA分子是长度在约17和约30个核苷酸之间的双链RNA或ssRNA分子。在另一个实施例中,小干扰RNA分子是长度在约19和约25个核苷酸之间的双链RNA或ssRNA分子。在另一个实施例中,小干扰RNA分子是长度在约21至约23个核苷酸之间的双链RNA或ssRNA分子。在某些实施例中,具有长于21个核苷酸的双链区的发夹结构可以促进有效的小干扰RNA指导的沉默,与环序列和长度无关。用于RNA干扰的示例性编码序列在SEQ ID NO:17和18中示出。

[0207] 靶mRNA序列通常长度为约14至约50个核苷酸。因此,可以对靶mRNA进行约14至约50个核苷酸长度的区域的扫描,所述区域优选地满足靶序列的一个或多个以下标准:A+T/G+C比在约2:1和约1:2之间;在靶序列的5'端处的AA二核苷酸或CA二核苷酸;对靶mRNA独特的至少10个连续核苷酸的序列(即,序列不存在于来自相同植物的其它mRNA序列中);以及不超过三个连续鸟嘌呤(G)核苷酸或超过三个连续胞嘧啶(C)核苷酸的“运行”。这些标准可以使用本领域已知的多种技术进行评价,例如计算机程序如BLAST可以用于搜索可公开获得的数据库,以确定所选择的靶序列是否对于靶mRNA是独特的。或者,使用商购可得的计算机软件(例如商购可得的OligoEngine、Target Finder和小干扰RNA设计工具),可以选择靶序列(和设计的小干扰RNA序列)。

[0208] 在一个实施例中,选择满足上述标准中的一个或多个、长度在约14和约30个核苷酸之间的靶mRNA序列。在另一个实施例中,选择满足上述标准中的一个或多个、长度在约16和约30个核苷酸之间的靶序列。在另一个实施例中,选择满足上述标准中的一个或多个、长

度在约19和约30个核苷酸之间的靶序列。在另一个实施例中,选择满足上述标准中的一个或多个、长度在约19和约25个核苷酸之间的靶序列。

[0209] 在一个示例性实施例中,小干扰RNA分子包含特异性反义序列,其与本文所述聚核苷酸序列中的任何一个的至少12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30个或更多个邻接核苷酸互补。

[0210] 由小干扰RNA分子包含的特异性反义序列可以与靶序列的互补序列一致或基本上一致。在一个实施例中,由小干扰RNA分子包含的特异性反义序列与靶mRNA序列的互补序列至少约75%、80%、85%、90%、95%、96%、97%、98%、99%或100%一致。确定序列一致性的方法是本领域已知的,并且可以例如通过使用University of Wisconsin Computer Group (GCG) 软件的或NCBI网站上提供的BLASTN程序进行确定。

[0211] 通过在一个、两个、三个、四个或更多个核苷酸处与靶mRNA的序列不同(例如,通过核苷酸置换,包括转换或颠换),小干扰RNA分子的特异性反义序列可以显示出变化性。当这类核苷酸置换存在于双链RNA分子的反义链中时,在有义链中置换核苷酸通常与之形成氢键碱基配对的互补核苷酸可以相应地置换或不置换。还涵盖了双链RNA分子,其中一种或多种核苷酸置换在有义序列而不是反义链中发生。如上所述,当小干扰RNA分子的反义序列包含在小干扰RNA的核苷酸序列和靶核苷酸序列之间的一个或多个错配时,错配可以在反义序列的3'端、5'端或中心部分发现。

[0212] 在另一个实施例中,小干扰RNA分子包含特异性反义序列,在严格条件下,所述特异性反义序列能够与天然存在的靶基因或靶mRNA的一部分选择性杂交。如本领域普通技术人员已知的,可以通过改变用于杂交和洗涤步骤的时间、温度或溶液浓度,来实现杂交条件的严格性中的变化。合适条件还可以部分取决于使用的具体核苷酸序列,例如靶mRNA或基因的序列。

[0213] 一种用于在植物中诱导双链RNA沉默的方法是用产生发夹RNA的基因构建体转化(参见Smith等人.(2000) *Nature*, 407, 319-320)。这类构建体包含由适当间隔物分开的靶基因序列的反向区。由于内含子剪接的发夹RNA的生成,功能性植物内含子区作为间隔物片段的插入另外增加基因沉默诱导的效率(Wesley等人.(2001) *Plant J.*, 27, 581-590)。合适地,茎长为约50个核苷酸至约1千个碱基长度。用于产生内含子剪接的发夹RNA的方法在本领域中充分描述(参见例如Bioscience, Biotechnology, and Biochemistry (2008) 72, 2, 615-617)。

[0214] 具有双链体或双链结构的干扰RNA分子,例如双链RNA或小发夹RNA可具有平端,或可具有3'或5'突出端。如本文所用,“突出端”指当一个RNA链的3'端延伸超出另一条链的5'端(3'突出端),或反之亦然(5'突出端)时,由双链体结构突出的一个或多个不成对的核苷酸。包含突出端的核苷酸可以是核糖核苷酸、脱氧核糖核苷酸或其修饰形式。在一个实施例中,干扰RNA分子的至少一条链具有长度约1至约6个核苷酸的3'突出端。在其它实施例中,3'突出端长度为约1至约5个核苷酸、约1至约3个核苷酸以及约2至约4个核苷酸。

[0215] 当干扰RNA分子在分子的一端包含3'突出端时,另一端可以是平端的或也具有突出端(5'或3')。当干扰RNA分子在分子的两端包含突出端时,突出端的长度可以是相同的或不同的。在一个实施例中,干扰RNA分子在分子的两端包含约1至约3个核苷酸的3'突出端。在另一个实施例中,干扰RNA分子是在分子的两端具有2个核苷酸的3'突出端的双链RNA。在

另一个实施例中,构成干扰RNA的突出端的核苷酸是TT二核苷酸或UU二核苷酸。

[0216] 当确定包含一个或多个突出端的干扰RNA分子与靶mRNA序列的一致性百分比时,一个或多个突出端可加以考虑或不加以考虑。例如,来自3'突出端的核苷酸以及来自双链的5'或3'端的至多2个核苷酸可以进行修饰,而小干扰RNA分子活性无明显丧失。

[0217] 干扰RNA分子可以包含一个或多个5'或3'帽状结构。干扰RNA分子可以包含在干扰RNA分子的有义链、反义链、或有义链和反义链两者的3'端处;或者在有义链、反义链、或有义链和反义链两者的5'端处的帽状结构。或者,干扰RNA分子可以包含在干扰RNA分子的3'端和5'端两者处的帽状结构。术语“帽状结构”指在寡核苷酸的任一端处并入的化学修饰,这使分子免于核酸外切酶降解,并且还可以促进在细胞内的递送或定位。

[0218] 另一种可应用于干扰RNA分子的修饰是干扰RNA分子与一个或多个部分或结合物的化学连接,这增强干扰RNA分子的活性、细胞分布、细胞摄取、生物利用度或稳定性。聚核苷酸可以通过本领域充分确定的方法进行合成或修饰。化学修饰可以包括但不限于2'修饰、非天然碱基的引入、与配体的共价连接以及用硫代磷酸酯键替换磷酸酯键。在这个实施例中,双链体结构的完整性通过至少一个,且通常为两个化学键得到加强。化学连接可以通过多种众所周知技术中的任一种实现,例如通过引入共价键、离子键或氢键;疏水作用、范德华力或堆积相互作用;借助于金属离子配位,或通过使用嘌呤类似物。

[0219] 可以修饰两条单链中的一条或两条的核苷酸以调节细胞酶的活化,例如但不限于某些核酸酶。用于降低或抑制细胞酶活化的技术是本领域已知的,包括但不限于2'-氨基修饰、2'-氟修饰、2'-烷基修饰、不带电的主链修饰、吗啉代修饰、2'-O-甲基修饰和氨基磷酸酯。因此,双链RNA上的核苷酸的至少一个2'-羟基替换为化学基团。另外,至少一个核苷酸可以进行修饰,以形成锁核苷酸。这类锁核苷酸含有连接核糖的2'-氧与核糖的4'-碳的亚甲基或亚乙基桥。锁核苷酸引入寡核苷酸内改善互补序列的亲合力,并且使解链温度升高几度。

[0220] 配体可以结合至干扰RNA分子,例如以增强其细胞吸收。在某些实施例中,疏水性配体结合至分子,以促进细胞膜的直接渗透。这些方法已经用于促进反义寡核苷酸的细胞渗透。在某些情况下,阳离子配体与寡核苷酸的结合通常使得对核酸酶的抗性改善。阳离子配体的代表性实例包括丙铵和二甲基丙铵。当阳离子配体分布遍及寡核苷酸时,反义寡核苷酸可以保留其与mRNA的高结合亲和力。

[0221] 本文所述的分子和聚核苷酸可以使用众所周知的固相合成技术进行制备。可以另外或替代地采用本领域已知的用于这类合成的任何其它方式。

[0222] 多个实施例涉及包含本文所述的聚核苷酸中的一种或多种或者一种或多种干扰RNA构建体的表达载体。图21中示出了示例性构建体。

[0223] 多个实施例涉及包含一种或多种聚核苷酸或一种或多种干扰RNA构建体的表达载体,所述干扰RNA构建体编码本文所述的一种或多种干扰RNA聚核苷酸,其能够自我退火以形成发夹结构,其中所述构建体包含(a)本文所述的聚核苷酸中的一种或多种;(b)编码间隔元件的第二序列,所述间隔元件形成发夹结构的环;和(c)置于与第一序列相同的定向、包含第一序列的反向互补序列的第三序列,其中第二序列置于第一序列和第三序列之间,并且第二序列可操作地连接于第一序列和第三序列。

[0224] 所公开的序列可以用于构建不形成发夹结构的多种聚核苷酸。举例来说,可以通

过(1)通过可操作地连接于第一启动子来转录DNA的第一条链,和(2)通过可操作地连接于第二启动子来转录DNA片段的第一条链的反向互补序列来形成双链RNA。聚核苷酸的每条链可以由相同表达载体或不同表达载体转录。具有RNA干扰活性的RNA双链体可以酶促转化为小干扰RNA,以调节RNA水平。

[0225] 因此,多个实施例涉及包含一种或多种本文所述的聚核苷酸或编码能够自我退火的干扰RNA聚核苷酸的干扰RNA构建体的表达载体,其中所述构建体包含(a)本文所述的聚核苷酸中的一种或多种;和(b)置于与第一序列相同的定向、包含第一序列的互补(例如反向互补)序列的第二序列。

[0226] 提供了通过促进基因表达的共抑制,用于调节本文所述多肽中的一种或多种(或如本文所述的其任何组合)的内源表达水平的多种组合物和方法。由于将转基因的多个拷贝引入植物细胞宿主内,发生共抑制现象。转基因的多个拷贝的整合可以导致转基因和靶向内源基因的调节表达。共抑制程度取决于转基因和靶向内源基因之间的序列一致性程度。内源基因和转基因两者的沉默都可以通过沉默基因座(即内源启动子和内源目的基因)的广泛甲基化而发生,所述广泛甲基化可以阻碍转录。或者,在一些情况下,内源基因和转基因的共抑制可以通过转录后基因沉默而发生,其中可以产生转录物,但增强的降解速率阻碍转录物的累积。通过转录后基因沉默的共抑制机制被认为类似RNA干扰,因为RNA看起来是这些过程中的重要起始子和靶,并且可以至少部分地通过相同分子机制、可能通过RNA指导的mRNA降解进行介导。

[0227] 可以通过将核酸或其片段的多个拷贝作为转基因整合到目的植物的基因组内来实现核酸的共抑制。宿主植物可以用表达载体进行转化,所述表达载体包含可操作地连接于核酸或其片段的启动子。多个实施例涉及用于促进内源基因的共抑制的表达载体,所述内源基因包含可操作地连接于聚核苷酸的启动子。

[0228] 多个实施例涉及通过将聚核苷酸的多个拷贝整合到植物基因组内,用于调节本文所述的聚核苷酸中的一种或多种(或如本文所述的其任何组合)的表达水平的方法,其包含:用表达载体转化植物细胞宿主,所述表达载体包含可操作地连接于聚核苷酸的启动子。

[0229] 提供了通过调节mRNA的翻译用于调节内源基因表达水平的多种组合物和方法。宿主植物细胞可以用表达载体转化,所述表达载体包含:可操作地连接于聚核苷酸的启动子,所述聚核苷酸以就启动子而言的反义定向放置,以使得能够表达具有与mRNA的一部分互补的序列的RNA聚核苷酸。

[0230] 用于调节mRNA翻译的多种表达载体可以包含:可操作地连接于聚核苷酸的启动子,其中序列以就启动子而言的反义定向放置。反义RNA聚核苷酸的长度可以改变,并且可以是约15-20个核苷酸、约20-30个核苷酸、约30-50个核苷酸、约50-75个核苷酸、约75-100个核苷酸、约100-150个核苷酸、约150-200个核苷酸以及约200-300个核苷酸。

[0231] 还可以通过将转座子(例如,IS元件)引入目的植物的基因组内来靶向基因以灭活。这些活动基因元件可以通过伴性异花受精而引入,并且可以就蛋白质活性的丧失筛选插入突变体。通过例如伴性异花受精,通过使亲本植物与未实施转座子诱导诱变的植物杂交,可将亲本植物中的破坏基因引入其它植物内。可以利用本领域技术人员已知的任何标准育种技术。在一个实施例中,可以通过插入一个或多个转座子灭活一种或多种基因。突变可能导致一种或多种基因的纯合破坏、一种或多种基因的杂合破坏,或如果破坏超过一种

基因,那么可能导致纯合和杂合破坏两者的组合。合适的转座元件包括反转录转座子、反转座子和SINE样元件。这类方法是本领域技术人员已知的。

[0232] 或者,可以通过将源自许多小环状RNA的核酶引入植物中来靶向基因以灭活,所述小环状RNA能够自切割和复制。这些RNA可以单独复制(类病毒RNA)或伴随辅助病毒(卫星RNA)而复制。合适RNA的实例包括源自鳄梨日斑病类病毒的那些,以及源自烟草环斑病毒、苜蓿短暂条纹病毒、绒毛烟草斑驳病毒、莴苣斑驳病毒和地下三叶草斑驳病毒的卫星RNA。多种靶RNA特异性核酶是本领域技术人员已知的。

[0233] 如本文所讨论,可以通过非转基因方式来调节一种或多种多肽的表达,如本文所讨论的在一个或多个基因中产生一个或多个突变。在基因序列中随机引入突变的方法可以包括化学诱变、EMS诱变和辐射诱变。将一个或多个靶向突变引入细胞内的方法包括但不限于基因组编辑技术(确切地说锌指核酸酶介导的诱变和靶向诱导的基因组局部损伤(TILLING))、同源重组、寡核苷酸引导的诱变和大范围核酸酶介导的诱变。在一个实施例中,使用TILLING。这是可以用于生成和/或鉴定编码具有修饰的表达和/或活性的多肽的聚核苷酸的诱变技术。TILLING还允许选择携带这类突变体的植物。TILLING组合高密度诱变与高通量筛选方法。用于TILLING的方法是本领域众所周知的(参见McCallum等人,(2000)Nat Biotechnol 18:455-457和Stemple(2004)Nat Rev Genet 5(2):145-50)。

[0234] 突变的一些非限制性实例是至少一个核苷酸的缺失、插入和错义突变,单核苷酸多态性和简单序列重复。在突变后,可以执行筛选,以鉴定产生提前终止密码子或以其它方式无功能的基因的突变。在突变后,可以执行筛选以鉴定产生功能基因的突变,所述功能基因能够以升高水平表达。可以通过测序或使用特异性针对基因或蛋白质的一个或多个探针或引物进行突变体的筛选。还可以产生聚核苷酸中的特异性突变,其可以导致调节的基因表达、mRNA的稳定性调节或蛋白质的稳定性调节。这类植物在本文中被称作“非天然存在的”或“突变型”植物。通常,突变型或非天然存在的植物包括外源或合成或人造核酸(例如DNA或RNA)的至少一部分,其在被操纵前不存在于植物中。外源核酸可以是单个核苷酸、两个或更多个核苷酸、两个或更多个邻接核苷酸、或者两个或更多个不邻接核苷酸-如至少10、20、30、40、50、100、200、300、400、500、600、700、800、900、1000、1100、1200、1300、1400或1500个或更多个邻接或不邻接核苷酸。

[0235] 突变型或非天然存在的植物或植物细胞可以具有一个或多个基因中的一个或多个突变的任何组合,所述一个或多个突变使得蛋白质水平得以调节。举例来说,突变型或非天然存在的植物或植物细胞可以具有单个基因中的单个突变;单个基因中的多个突变;两个或更多个或者三个或更多个或者四个或更多个基因中的单个突变;或者两个或更多个或者三个或更多个或者四个或更多个基因中的多个突变。这些突变的实例描述于本文中。进一步举例来说,突变型或非天然存在的植物或植物细胞可以具有在基因的特定部分中的一个或多个突变,所述特定部分如在基因中编码蛋白质的活性位点或其一部分的区域中。进一步举例来说,突变型或非天然存在的植物或植物细胞可以具有在一个或多个基因外部区域中的一个或多个突变,所述区域如在其调节的基因的上游或下游区域,只要其调节基因的活性或表达。上游元件可以包括启动子、增强子或转录因子。一些元件如增强子可以置于它调节的基因的上游或下游。元件无需定位接近于它调节的基因,因为一些元件已发现位于它调节的基因上游或下游几百上千个碱基对处。突变型或非天然存在的植物或植物细胞

可以具有位于基因的前100个核苷酸内、基因的前200个核苷酸内、基因的前300个核苷酸内、基因的前400个核苷酸内、基因的前500个核苷酸内、基因的前600个核苷酸内、基因的前700个核苷酸内、基因的前800个核苷酸内、基因的前900个核苷酸内、基因的前1000个核苷酸内、基因的前1100个核苷酸内、基因的前1200个核苷酸内、基因的前1300个核苷酸内、基因的前1400个核苷酸内、或基因的前1500个核苷酸内的一个或多个突变。突变型或非天然存在的植物或植物细胞可以具有位于基因的100个核苷酸的第一、第二、第三、第四、第五、第六、第七、第八、第九、第十、第十一、第十二、第十三、第十四或第十五集合或其组合内的一个或多个突变。公开了包含突变型多肽变体的突变型或非天然存在的植物或植物细胞(例如,如本文所述的突变型、非天然存在的或转基因植物或植物细胞等)。

[0236] 在一个实施例中,使来自植物的种子诱变且随后生长成第一代突变型植物。随后使第一代植物自花授粉,并且使来自第一代植物的种子生长成第二代植物,所述第二代植物随后就其基因座中的突变进行筛选。尽管诱变的植物材料可以针对突变进行筛选,但筛选第二代植物的优点在于所有体细胞突变都对应于生殖系突变。本领域技术人员应理解,多种植物材料包括但不限于种子、花粉、植物组织或植物细胞可以进行诱变,以便产生突变植物。然而,当植物核酸针对突变进行筛选时,可以影响诱变的植物材料类型。举例来说,当在非诱变植物授粉之前对花粉实施诱变时,使授粉获得的种子生长成第一代植物。第一代植物的每一个细胞将含有在花粉中产生的突变;因此这些第一代植物随后可针对突变进行筛选,而不是等到第二代进行。

[0237] 主要产生点突变和短缺失、插入、颠换和或转换的诱变剂包括化学诱变剂或辐射,可用于产生突变。诱变剂包括但不限于甲磺酸乙酯、甲磺酸甲酯、N-乙基-N-亚硝基脒、三乙基蜜胺、N-甲基-N-亚硝基脒、丙卡巴肼、苯丁酸氮芥、环磷酰胺、硫酸二乙酯、丙烯酰胺单体、美法仑、氮芥、长春新碱、二甲基亚硝胺、N-甲基-N'-硝基-亚硝基胍、亚硝基胍、2-氨基嘌呤、7,12-二甲基苯并蒽、环氧乙烷、六甲基磷酰胺、白消安、二环氧烷烃(二环氧辛烷、二环氧丁烷等)、2-甲氧基-6-氯-9[3-(乙基-2-氯乙基)氨基丙基氨基]吡啶二盐酸盐和甲醛。

[0238] 也考虑了可能不是由诱变剂直接引起的基因座中的自发突变,只要它们导致期望的表型。合适的诱变试剂还可以包括例如电离辐射,如X射线、 γ 射线、快中子照射和UV辐射。本领域技术人员已知的植物核酸制备的任何方法均可以用于制备用于突变筛选的植物核酸。

[0239] 可以任选合并由个别植物、植物细胞或植物材料制备的核酸,以便加速源于诱变的植物组织、细胞或材料的植物群体中的突变筛选。可以筛选植物、植物细胞或植物材料的一个或多个后续世代。任选合并的组的大小取决于使用的筛选方法的灵敏度。

[0240] 在核酸样品任选合并后,可以对它们实施聚核苷酸特异性扩增技术,如聚合酶链反应。特异性针对基因或与基因紧邻的序列的任何一个或多个引物或探针可以用于扩增任选合并的核酸样品内的序列。寡核苷酸引物的实例在SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:15和SEQ ID NO:16中示出。合适地,SEQ ID NO:13和SEQ ID NO:14组合使用以检测NtAsn-1。合适地,SEQ ID NO:15和SEQ ID NO:16组合使用以检测NtAsn-5。合适地,一个或多个引物或探针设计为扩增最可能出现有用突变的基因座的区域。最优选地,引物设计为检测聚核苷酸区域内的突变。另外,引物和探针优选避免已知的多态性位点,以便容易筛选点突变。为了便于扩增产物的检测,一个或多个引物或探针可以使用任何常规标记方法进行

行标记。使用本领域充分理解的方法,可以基于本文所述的序列来设计引物或探针。

[0241] 为了便于扩增产物的检测,可以使用任何常规标记方法来标记引物或探针。使用本领域充分理解的方法,可以基于本文所述的序列来设计这些。多态性可以通过本领域已知的方法进行鉴定,并且一些已在文献中得到描述。

[0242] 在另一方面,提供了制备突变型植物的方法。所述方法涉及提供包含编码本文所述的功能性聚核苷酸(或如本文所述的其任何组合)的基因的植物的至少一个细胞。接下来,植物的至少一个细胞在有效调节本文所述的聚核苷酸的活性的条件下进行处理。至少一个突变型植物细胞随后繁殖成突变型植物,其中与对照植物相比较,所述突变型植物具有调节水平的所述多肽(或如本文所述的其任何组合)。在这一制备突变型植物的方法的一个实施例中,处理步骤涉及使至少一个细胞经受如上所述的化学诱变剂,且在有效获得至少一个突变型植物细胞的条件下。在这一方法的另一个实施例中,处理步骤涉及在有效获得至少一个突变型植物细胞的条件下,使至少一个细胞经受辐射源。术语“突变型植物”包括这样的突变型植物,与对照植物相比较,所述突变型植物中的基因型适当地通过除了基因工程改造或基因修饰之外的方式进行修饰。

[0243] 在某些实施例中,突变型植物、突变型植物细胞或突变型植物材料可以包含一个或多个突变,所述一个或多个突变在另种植物、植物细胞或植物材料中天然存在,且赋予所需性状。可以将这一突变并入(例如基因渗入)到另种植物、植物细胞或植物材料(例如,具有与得到突变的植物不同的遗传背景的植物、植物细胞或植物材料)中以产生在该植物中非天然存在的突变并赋予其性状。因此,例如,在第一植物中天然存在的突变可以引入到第二植物中,如与第一植物具有不同遗传背景的第二植物。技术人员因此能够搜索且鉴定在基因组中天然携带本文所述基因的一种或多种突变等位基因的植物,所述基因赋予所需性状。可以通过多种方法包括育种、回交和基因渗入,将天然存在的突变等位基因转移至第二植物,以产生在本文所述基因中具有一个或多个突变的品系、品种或杂交物。可以在突变型植物的库中筛选显示所需性状的植物。合适地,利用如本文所述的核苷酸序列的知识进行选择。因此,能够与对照相比较筛选遗传性状。这类筛选方法可以涉及如本文讨论的常规核酸扩增和/或杂交技术的应用。因此,另一方面涉及用于鉴定突变型植物的方法,其包含以下步骤:(a)提供来自植物的包含核酸的样品;和(b)测定聚核苷酸的核酸序列,其中与对照植物的聚核苷酸序列相比较,在聚核苷酸序列中的差异指示所述植物是突变型植物。在另一方面,提供了用于鉴定突变型植物的方法,与对照植物相比较,所述突变型植物累积降低水平的天冬酰胺,所述方法包含以下步骤:(a)提供来自待筛选的植物的样品;(b)测定所述样品是否包含在本文所述的聚核苷酸中的一种或多种中的一个或多个突变;和(c)测定所述植物的至少天冬酰胺含量。

[0244] 在另一方面,提供了用于制备突变型植物的方法,与对照植物相比,所述突变型植物具有降低水平的天冬酰胺,所述方法包含以下步骤:(a)提供来自第一植物的样品;(b)确定所述样品是否包含在本文所述的聚核苷酸中的一种或多种中的一个或多个突变,所述一个或多个突变导致天冬酰胺水平降低;和(c)将一个或多个突变转移到第二植物内。可以使用本领域已知的多种方法,如通过基因工程改造、基因操纵、基因渗入、植物育种、回交等等,将突变转移到第二植物内。在一个实施例中,第一植物是天然存在的植物。在一个实施例中,第二植物具有与第一植物不同的遗传背景。

[0245] 在另一方面,提供了用于制备突变型植物的方法,与对照植物相比,所述突变型植物具有降低水平的天冬酰胺,所述方法包含以下步骤:(a) 提供来自第一植物的样品;(b) 确定所述样品是否包含在本文所述的聚核苷酸中的一种或多种中的一个或多个突变,所述一个或多个突变导致天冬酰胺水平降低;和(c) 将来自第一植物的一个或多个突变基因渗入到第二植物内。在一个实施例中,基因渗入步骤包含植物育种,任选包括回交等等。在一个实施例中,第一植物是天然存在的植物。在一个实施例中,第二植物具有与第一植物不同的遗传背景。在一个实施例中,第一植物不是栽培品种或优良栽培品种。在一个实施例中,第二植物是栽培品种或优良栽培品种。

[0246] 另一方面涉及通过本文所述的方法获得或可获得的突变型植物(包括栽培品种或优良栽培品种突变型植物)。在某些实施例中,“突变型植物”可以具有仅定位于植物的特定区域,如在本文所述的一种或多种聚核苷酸的序列内的一个或多个突变。根据这一实施例,突变型植物的剩余基因组序列将与诱变前的植物相同或基本上相同。

[0247] 在某些实施例中,突变型植物可以具有定位于植物的超过一个区域,如在本文所述的聚核苷酸中的一种或多种的序列内以及基因组的一个或多个其它区域中的一个或多个突变。根据这一实施例,突变型植物的剩余基因组序列将与诱变前的植物不同或基本上不同。在某些实施例中,突变型植物可能不具有本文所述的聚核苷酸的一个或多个、两个或更多个、三个或更多个、四个或更多个、或者五个或更多个外显子中的一个或多个突变;或可能不具有本文所述的聚核苷酸的一个或多个、两个或更多个、三个或更多个、四个或更多个、或者五个或更多个内含子中的一个或多个突变;或可能不具有本文所述的聚核苷酸的启动子中的一个或多个突变;或可能不具有本文所述的聚核苷酸的3'非翻译区中的一个或多个突变;或可能不具有本文所述的聚核苷酸的5'非翻译区中的一个或多个突变;或可能不具有本文所述的聚核苷酸的编码区中的一个或多个突变;或可能不具有本文所述的聚核苷酸的非编码区中的一个或多个突变;或其部分中的其两个或更多个、三个或更多个、四个或更多个、五个或更多个;或者六个或更多个的任何组合。

[0248] 在另一方面,提供了鉴定在基因中包含突变的植物、植物细胞或植物材料的方法,所述基因编码本文所述的聚核苷酸,所述方法包含:(a) 对植物、植物细胞或植物材料实施诱变;(b) 由所述植物、植物细胞或植物材料或其后代获得核酸样品;和(c) 测定编码本文所述的聚核苷酸的基因或其变体或片段的核酸序列,其中所述序列中的差异指示其中的一个或多个突变。

[0249] 锌指蛋白也可以用于调节本文所述的聚核苷酸中的一种或多种的表达或活性。在多个实施例中,通过锌指核酸酶介导的诱变修饰包含聚核苷酸编码序列的一部分或全部的基因组DNA序列。在基因组DNA序列中搜索锌指蛋白结合的独特位点。或者,在基因组DNA序列中搜索锌指蛋白结合的两个独特位点,其中两个位点在相反链上且紧靠在一起,例如相隔1、2、3、4、5、6个或更多个碱基对。相应地,提供了与聚核苷酸结合的锌指蛋白。

[0250] 锌指蛋白可以进行工程改造,以识别基因中的所选靶位点。通过截短或扩增或选择方法偶联的定点诱变过程(所述选择方法例如但不限于噬菌体展示选择、细菌双杂交选择或细菌单杂交选择),锌指蛋白可以包含源自天然锌指DNA结合结构域和非天然锌指DNA结合结构域的任何基序组合。术语“非天然锌指DNA结合结构域”指锌指DNA结合结构域,其结合靶核酸内的三碱基对序列,并且在包含待修饰的核酸的细胞或生物中不存在。用于

设计结合特异性核苷酸序列的锌指蛋白的方法是本领域已知的,所述特异性核苷酸序列对于靶基因是独特的。

[0251] 可以通过制备编码结合聚核苷酸的锌指蛋白的第一聚核苷酸以及编码非特异性核酸内切酶的第二聚核苷酸(例如但不限于IIS型核酸内切酶的那些)的融合物来构建锌指核酸酶。在锌指蛋白和核酸酶之间的融合蛋白可以包含由两个碱基对组成的间隔物,或者,间隔物可以由三个、四个、五个、六个、七个或更多个碱基对组成。在多个实施例中,锌指核酸酶在聚核苷酸的基因组DNA序列的调节区、编码区或非编码区中引入双链断裂,且导致聚核苷酸的表达水平降低,或由其编码的蛋白质活性降低。通过锌指核酸酶的切割通常导致在通过非同源末端连接的DNA修复后在切割位点处的DNA缺失。

[0252] 在其它实施例中,可以选择锌指蛋白以结合于聚核苷酸的调节序列。更具体来说,调节序列可以包含转录起始位点、起始密码子、外显子区、外显子-内含子边界、终止子或终止密码子。相应地,本公开提供了在本文所述的一种或多种聚核苷酸附近或其内通过锌指核酸酶介导的诱变产生的突变型、非天然存在的或转基因植物或植物细胞,以及通过锌指核酸酶介导的诱变用于制备这类植物或植物细胞的方法。用于将锌指蛋白和锌指核酸酶递送至植物的方法类似于下文对于大范围核酸酶递送描述的那些。

[0253] 在另一方面,描述了使用大范围核酸酶如I-CreI,用于产生突变型、非天然存在的或转基因或以其它方式遗传修饰的植物的方法。天然存在的大范围核酸酶以及重组大范围核酸酶可以用于特异性引起在植物基因组DNA的单个位点或相对少数位点处的双链断裂,以允许破坏本文所述的一种或多种聚核苷酸。大范围核酸酶可以是具有改变的DNA识别特性的经工程改造的大范围核酸酶。大范围核酸酶蛋白质可以通过本领域已知的多种不同机制递送到植物细胞内。

[0254] 本公开还涵盖使用大范围核酸酶使植物细胞或植物中的本文所述的聚核苷酸(或如本文所述的其任何组合)失活。具体来说,本公开提供使用大范围核酸酶使植物中的聚核苷酸失活的方法,其包含:a)提供包含如本文所述的聚核苷酸的植物细胞;b)将大范围核酸酶或编码大范围核酸酶的构建体引入所述植物细胞内;和(c)允许大范围核酸酶使聚核苷酸基本上失活

[0255] 大范围核酸酶可以用于切割在聚核苷酸的编码区内的大范围核酸酶识别位点。这类切割通常导致在通过非同源末端连接的诱变DNA修复后,在大范围核酸酶识别位点处的DNA缺失。基因编码序列中的这类突变通常足以使基因失活。这种修饰植物细胞的方法首先涉及使用合适的转化方法将大范围核酸酶表达盒递送至植物细胞。为了最高效率,期望将大范围核酸酶表达盒连接至可选标记,且在选择剂的存在下选择成功转化的细胞。这种方法使得大范围核酸酶表达盒整合到基因组内,然而,如果植物可能需要监管机构批准,那么这可能是不理想的。在这类情况下,使用常规育种技术,大范围核酸酶表达盒(和连接的可选标记基因)可以在后续植物世代中分离。或者,植物细胞可以最初用缺乏可选标记的大范围核酸酶表达盒进行转化,并且可以在缺乏选择剂的培养基上生长。在这类条件下,经处理的细胞的一部分获得大范围核酸酶表达盒,并且瞬时表达经工程改造的大范围核酸酶,而不将大范围核酸酶表达盒整合到基因组内。因为它不负责转化效率,所以后面这种转化程序需要筛选更多数目的经处理的细胞,以获得所需基因组修饰。当利用锌指蛋白或锌指核酸酶时,上述方法还可以应用于修饰植物细胞。

[0256] 在大范围核酸酶表达盒递送后,植物细胞最初在对于使用的具体转化程序典型的条件下生长。这可能意味着在低于26℃的温度下,通常在黑暗中,使经转化的细胞在培养基上生长。这类标准条件可以使用一段时间,优选1-4天,以允许植物细胞从转化过程恢复。在此最初恢复期后的任何点,生长温度可以上升,以刺激经工程改造的大范围核酸酶的活性,从而切割且突变大范围核酸酶识别位点。

[0257] 对于某些应用,可能期望精确地从植物的基因组中去除聚核苷酸。这类应用可能使用一对经工程改造的大范围核酸酶,所述一对大范围核酸酶各自切割在预期缺失的任一侧上的大范围核酸酶识别位点。还可以使用TAL效应物核酸酶(TALEN),其能够识别且结合基因,并且将双链断裂引入基因组内。因此,在另一方面,涵盖了使用TAL效应物核酸酶,用于生产如本文所述的突变型、非天然存在的或转基因或以其它方式遗传修饰的植物的方法。

[0258] 在烟草的白肋品种中已经鉴定为在烘烤期间参与天冬酰胺合成的基因可能适用于其它植物和其它植物品种。因此,本公开可能在其它植物中是可再现的并且适用于用变异品系进行育种。

[0259] 适用于本公开的植物包括但不限于单子叶和双子叶植物和植物细胞系统,包括来自以下各科之一的物种:爵床科(Acanthaceae)、葱科(Alliaceae)、六出花科(Alstroemeriaceae)、石蒜科(Amaryllidaceae)、夹竹桃科(Apocynaceae)、棕榈科(Arecaceae)、菊科(Asteraceae)、小檗科(Berberidaceae)、红木科(Bixaceae)、十字花科(Brassicaceae)、凤梨科(Bromeliaceae)、大麻科(Cannabaceae)、石竹科(Caryophyllaceae)、三尖杉科(Cephalotaxaceae)、藜科(Chenopodiaceae)、秋水仙科(Colchicaceae)、葫芦科(Cucurbitaceae)、薯蓣科(Dioscoreaceae)、麻黄科(Ephedraceae)、古柯科(Erythroxylaceae)、大戟科(Euphorbiaceae)、豆科(Fabaceae)、唇形科(Lamiaceae)、亚麻科(Linaceae)、石松科(Lycopodiaceae)、锦葵科(Malvaceae)、黑药花科(Melanthiaceae)、芭蕉科(Musaceae)、桃金娘科(Myrtaceae)、蓝果树科(Nyssaceae)、罂粟科(Papaveraceae)、松科(Pinaceae)、车前草科(Plantaginaceae)、禾本科(Poaceae)、蔷薇科(Rosaceae)、茜草科(Rubiaceae)、杨柳科(Salicaceae)、无患子科(Sapindaceae)、茄科(Solanaceae)、红豆杉科(Taxaceae)、山茶科(Theaceae)或葡萄科(Vitaceae)。

[0260] 合适物种可以包括以下各属的成员:黄葵属(Abelmoschus)、冷杉属(Abies)、槭属(Acer)、剪股颖属(Agrostis)、葱属(Allium)、六出花属(Alstroemeria)、凤梨属(Ananas)、穿心莲属(Andrographis)、须芒草属(Andropogon)、蒿属(Artemisia)、芦竹属(Arundo)、颠茄属(Atropa)、小檗属(Berberis)、甜菜属(Beta)、红木属(Bixa)、芸苔属(Brassica)、金盏菊属(Calendula)、山茶属(Camellia)、喜树属(Camptotheca)、大麻属(Cannabis)、辣椒属(Capsicum)、红花属(Carthamus)、长春花属(Catharanthus)、三尖杉属(Cephalotaxus)、菊属(Chrysanthemum)、金鸡纳属(Cinchona)、西瓜属(Citrullus)、咖啡属(Coffea)、秋水仙属(Colchicum)、鞘蕊花属(Coleus)、甜瓜属(Cucumis)、南瓜属(Cucurbita)、狗牙根属(Cynodon)、曼陀罗属(Datura)、石竹属(Dianthus)、洋地黄属(Digitalis)、薯蓣属(Dioscorea)、油棕属(Elaeis)、麻黄属(Ephedra)、蔗茅属(Erianthus)、古柯属(Erythroxylum)、桉树属(Eucalyptus)、羊茅属(Festuca)、草莓属(Fragaria)、雪花莲属(Galanthus)、大豆属(Glycine)、棉属(Gossypium)、向日葵属(Helianthus)、橡胶树属

(Hevea)、大麦属(*Hordeum*)、天仙子属(*Hyoscyamus*)、麻风树属(*Jatropha*)、莴苣属(*Lactuca*)、亚麻属(*Linum*)、黑麦草属(*Lolium*)、羽扇豆属(*Lupinus*)、番茄属(*Lycopersicon*)、石松属(*Lycopodium*)、木薯属(*Manihot*)、苜蓿属(*Medicago*)、薄荷属(*Mentha*)、芒属(*Miscanthus*)、芭蕉属(*Musa*)、烟草属、稻属(*Oryza*)、黍属(*Panicum*)、罂粟属(*Papaver*)、银胶菊属(*Parthenium*)、狼尾草属(*Pennisetum*)、矮牵牛属(*Petunia*)、藨草属(*Phalaris*)、梯牧草属(*Phleum*)、松属(*Pinus*)、早熟禾属(*Poa*)、一品红属(*Poinsettia*)、杨属(*Populus*)、萝芙木属(*Rauwolfia*)、蓖麻属(*Ricinus*)、蔷薇属(*Rosa*)、甘蔗属(*Saccharum*)、柳属(*Salix*)、血根草属(*Sanguinaria*)、赛茛苳属(*Scopolia*)、黑麦属(*Secale*)、茄属(*Solanum*)、高粱属(*Sorghum*)、米草属(*Spartina*)、菠菜属(*Spinacea*)、菊蒿属(*Tanacetum*)、红豆杉属(*Taxus*)、可可属(*Theobroma*)、小黑麦属(*Triticosecale*)、小麦属(*Triticum*)、北美穗草属(*Uniola*)、藜芦属(*Veratrum*)、长春花属(*Vinca*)、葡萄属(*Vitis*)和玉蜀黍属(*Zea*)。

[0261] 合适物种可以包括黍属物种(*Panicum* spp.)、高粱属物种(*Sorghum* spp.)、芒属物种(*Miscanthus* spp.)、甘蔗属物种(*Saccharum* spp.)、蔗茅属物种(*Erianthus* spp.)、杨属物种(*Populus* spp.)、大须芒草(*Andropogon gerardii*) (大蓝秆草)、象草(*Pennisetum purpureum*) (象草)、藨草(*Phalaris arundinacea*) (草芦)、狗牙根(*Cynodon dactylon*) (狗牙根)、高羊茅(*Festuca arundinacea*) (高羊茅)、互花米草(*Spartina pectinata*) (草原索草)、苜蓿(*Medicago sativa*) (苜蓿)、芦竹(*Arundo donax*) (芦竹)、黑麦(*Secale cereale*) (黑麦)、柳属物种(*Salix* spp.) (柳树)、桉树属物种(*Eucalyptus* spp.) (桉树)、小黑麦(*Triticosecale*) (小麦杂交黑麦)、竹、向日葵(*Helianthus annuus*) (向日葵)、红花(*Carthamus tinctorius*) (红花)、桐油树(*Jatropha curcas*) (麻风树)、蓖麻(*Ricinus communis*) (蓖麻)、油棕(*Elaeis guineensis*) (棕榈)、亚麻(*Linum usitatissimum*) (亚麻)、芥菜(*Brassica juncea*)、甜菜(*Beta vulgaris*) (甜菜)、木薯(*Manihot esculenta*) (木薯)、番茄(*Lycopersicon esculentum*) (番茄)、莴苣(*Lactuca sativa*) (生菜)、香蕉(*Musyclise alca*) (香蕉)、马铃薯(*Solanum tuberosum*) (土豆)、甘蓝(*Brassica oleracea*) (绿花椰菜、花椰菜、孢子甘蓝(*Brussels sprouts*))、山茶(*Camellia sinensis*) (茶)、草莓(*Fragaria ananassa*) (草莓)、可可(*Theobroma cacao*) (可可)、咖啡(*Coffeycliseca*) (咖啡)、葡萄(*Vitis vinifera*) (葡萄)、菠萝(*Ananas comosus*) (菠萝)、辣椒(*Capsicum annum*) (辣椒和甜椒)、洋葱(*Allium cepa*) (洋葱)、甜瓜(*Cucumis melo*) (甜瓜)、黄瓜(*Cucumis sativus*) (黄瓜)、笋瓜(*Cucurbita maxima*) (南瓜)、南瓜(*Cucurbita moschata*) (南瓜)、菠菜(*Spinacea oleracea*) (菠菜)、西瓜(*Citrullus lanatus*) (西瓜)、秋葵(*Abelmoschus esculentus*) (秋葵)、茄子(*Solanum melongena*) (茄子)、蔷薇属物种(*Rosa* spp.) (玫瑰)、香石竹(*Dianthus caryophyllus*) (康乃馨)、矮牵牛属物种(*Petunia* spp.) (矮牵牛)、一品红(*Poinsettia pulcherrima*) (一品红)、白羽扇豆(*Lupinus albus*) (羽扇豆)、燕麦(*Uniola paniculata*) (燕麦)、翦股颖(翦股颖属物种(*Agrostis* spp.))、美洲山杨(*Populus tremuloides*) (白杨)、松属(*Pinus* spp.) (松树)、冷杉属物种(*Abies* spp.) (冷杉)、槭属物种(*Acer* spp.) (枫木)、大麦(*Hordeum vulgare*) (大麦)、草地早熟禾(*Poa pratensis*) (早熟禾)、黑麦草属物种(*Lolium* spp.) (黑麦草)和猫尾草(*Phleum pratense*) (梯牧草)、柳枝稷(*Panicum virgatum*) (柳枝稷)、高粱

(*Sorghum* spp.) 或(高粱、苏丹草)、芒草 (*Miscanthus giganteus*) (芒草)、甘蔗属物种 (*Saccharum* sp.) (能源蔗)、香脂杨 (*Populus balsamifera*) (白杨)、玉蜀黍 (*Zea mays*) (玉米)、大豆 (*Glycine max*) (黄豆)、油菜 (*Brassica napus*) (芸苔)、小麦 (*Triticum aestivum*) (小麦)、陆地棉 (*Gossypium hirsutum*) (棉花)、水稻 (*Oryza sativa*) (稻)、向日葵 (*Helianthus annuus*) (向日葵)、苜蓿 (*Medicago sativa*) (苜蓿)、甜菜 (*Beta vulgaris*) (甜菜) 或珍珠粟 (*Pennisetum glaucum*) (珍珠粟)。

[0262] 多个实施例涉及突变型、非天然存在的或转基因植物或植物细胞,其经修饰以调节基因表达水平,由此产生与对照相比较,其中多肽的表达水平在目的组织中经调节的植物或植物细胞(如烟草植物或烟草植物细胞)。所公开的组合物和方法可以应用于烟草属的任何物种,包括黄花烟草 (*N. rustica*) 和普通烟草(例如 LA B21、LN KY171、TI 1406、Basma、Galpao、Perique、Beinhart 1000-1 和 Petico)。其它物种包括无茎烟草 (*N. acaulis*)、尖叶烟草 (*N. acuminata*)、非洲烟草 (*N. africana*)、花叶烟草 (*N. alata*)、阿米基诺氏烟草 (*N. ameghinoi*)、抱茎烟草 (*N. amplexicaulis*)、阿伦兹氏烟草 (*N. arentsii*)、渐狭叶烟草 (*N. attenuata*)、阿姆布吉烟草 (*N. azambujae*)、贝纳莫特氏烟草 (*N. benavidesii*)、本赛姆氏烟草 (*N. benthamiana*)、印度烟草 (*N. bigelovii*)、博内里烟草 (*N. bonariensis*)、洞生烟草 (*N. cavicola*)、克利夫兰氏烟草 (*N. clevelandii*)、心叶烟草 (*N. cordifolia*)、伞床烟草 (*N. corymbosa*)、迪伯纳氏烟草 (*N. debneyi*)、木丝烟草 (*N. excelsior*)、福尔吉特氏烟草 (*N. forgetiana*)、香烟草 (*N. fragrans*)、粉蓝烟草 (*N. glauca*)、粘烟草 (*N. glutinosa*)、古特斯比氏烟草 (*N. goodspeedii*)、哥西氏烟草 (*N. gossei*)、杂交烟草 (*N. hybrid*)、因古儿巴烟草 (*N. ingulba*)、卡瓦卡米氏烟草 (*N. kawakamii*)、奈特氏烟草 (*N. knightiana*)、郎氏烟草 (*N. langsdorffii*)、渐尖叶烟草 (*N. linearis*)、长花烟草 (*N. longiflora*)、海滨烟草 (*N. maritima*)、特大管烟草 (*N. megalosiphon*)、摩西氏烟草 (*N. miersii*)、夜花烟草 (*N. noctiflora*)、裸茎烟草 (*N. nudicaulis*)、欧布斯特烟草 (*N. obtusifolia*)、西方烟草 (*N. occidentalis*)、西方亚种香芥烟草 (*N. occidentalis* subsp. *hesperis*)、耳状烟草 (*N. otophora*)、圆锥烟草 (*N. paniculata*)、少花烟草 (*N. pauciflora*)、矮牵牛状烟草 (*N. petunioides*)、蓝茉莉叶烟草 (*N. plumbaginifolia*)、夸德瑞伍氏烟草 (*N. quadrivalvis*)、雷蒙德氏烟草 (*N. raimondii*)、波缘烟草 (*N. repanda*)、莲座烟草 (*N. rosulata*)、莲座亚种因古儿巴烟草 (*N. rosulata* subsp. *ingulba*)、圆叶烟草 (*N. rotundifolia*)、赛特氏烟草 (*N. setchellii*)、拟似烟草 (*N. simulans*)、茄叶烟草 (*N. solanifolia*)、斯佩格茨氏烟草 (*N. spegazzinii*)、斯托可通氏烟草 (*N. stocktonii*)、香甜烟草 (*N. suaveolens*)、美花烟草 (*N. sylvestris*)、拟穗状烟草 (*N. thyrsiflora*)、绒毛烟草 (*N. tomentosa*)、绒毛状烟草 (*N. tomentosiformis*)、三角叶烟草 (*N. trigonophylla*)、荫生烟草 (*N. umbratica*)、波叶烟草 (*N. undulata*)、颤毛烟草 (*N. velutina*)、序叶烟草 (*N. wigandioides*) 和花烟草 (*N. xsanderae*)。

[0263] 本文还涵盖使用烟草栽培品种和优良烟草栽培品种。因此,转基因、非天然存在的或突变型植物可以是烟草品种或优良烟草栽培品种,其包含一种或多种转基因、或者一个或多个基因突变或其组合。基因突变(例如,一种或多种多态性)可以是非天然存在于个别烟草品种或烟草栽培品种(例如,优良烟草栽培品种)的突变,或可以是的天然存在的基因突变,条件是所述突变并非天然存在于个别烟草品种或烟草栽培品种(例如,优良烟草栽培品种)。

培品种)中。

[0264] 特别有用的烟草品种包括白肋烟型、深型、烟道烘烤型和东方型烟草。品种或栽培品种的非限制性实例是:BD 64、CC 101、CC 200、CC 27、CC 301、CC 400、CC 500、CC 600、CC 700、CC 800、CC 900、Coker 176、Coker 319、Coker 371Gold、Coker 48、CD 263、DF911、DT 538LC Galpao烟草、GL 26H、GL 350、GL 600、GL 737、GL 939、GL 973、HB 04P、HB 04P LC、HB3307PLC、杂交403LC、杂交404LC、杂交501LC、K 149、K 326、K 346、K 358、K394、K 399、K 730、KDH 959、KT 200、KT204LC、KY10、KY14、KY 160、KY 17、KY 171、KY 907、KY907LC、KY14xL8LC、Little Crittenden、McNair 373、McNair 944、msKY 14xL8、窄叶Madole、窄叶Madole LC、NBH 98、N-126、N-777LC、N-7371LC、NC 100、NC 102、NC 2000、NC 291、NC 297、NC 299、NC 3、NC 4、NC 5、NC 6、NC7、NC 606、NC 71、NC 72、NC 810、NC BH 129、NC 2002、Neal Smith Madole、OXFORD 207、PD 7302LC、PD 7309LC、PD 7312LC、'Perique'烟草、PVH03、PVH09、PVH19、PVH50、PVH51、R 610、R 630、R 7-11、R 7-12、RG 17、RG 81、RG H51、RGH 4、RGH 51、RS 1410、Speight 168、Speight 172、Speight 179、Speight 210、Speight 220、Speight 225、Speight 227、Speight 234、Speight G-28、Speight G-70、Speight H-6、Speight H20、Speight NF3、TI 1406、TI 1269、TN 86、TN86LC、TN 90、TN 97、TN97LC、TN D94、TN D950、TR (Tom Rosson) Madole、VA 309、VA359、AA 37-1、B13P、Xanthi (Mitchell-Mor)、Bel-W3、79-615、Samsun Holmes NN、KTRDC2号杂交49、白肋21、KY8959、KY9、MD 609、PG01、PG04、P01、P02、P03、RG11、RG 8、VA509、AS44、Banket A1、Basma Drama B84/31、Basma I Zichna ZP4/B、Basma Xanthi BX 2A、Batek、Besuki Jember、C104、Coker 347、Criollo Misionero、Delcrest、Djebel 81、DVH 405、Galpão Comum、HB04P、希克斯阔叶、Kabakulak Ellassona、Kutsage E1、LA BU 21、NC 2326、NC 297、PVH 2110、Red Russian、Samsun、Saplak、Simmaba、Talgat 28、Wislica、Yayaldag、Prilep HC-72、Prilep P23、Prilep PB 156/1、Prilep P12-2/1、Yaka JK-48、Yaka JB 125/3、TI-1068、KDH-960、TI-1070、TW136、Basma、TKF 4028、L8、TKF 2002、GR141、Basma xanthi、GR149、GR153、Petit Havana。即使本文未特别指明,也涵盖上述的低转化亚变种。

[0265] 在一个实施例中,使用烟草的白肋型品种。

[0266] 适用于本公开的其它植物包括但不限于茶(茶树)。

[0267] 实施例还涉及用于产生突变型植物、非天然存在的植物、杂交植物或转基因植物的组合物和方法,所述植物已进行修饰,以调节本文所述的聚核苷酸(或如本文所述的其任何组合)的表达或活性。有利地,所获得的突变型植物、非天然存在的植物、杂交植物或转基因植物可以在整体外观上与对照植物相似或基本上相同。多种表型特征,如成熟程度、每一植物叶数、秆高、叶插入角度、叶大小(宽度和长度)、节间距离以及叶片-中脉比可以通过田间观察进行评价。

[0268] 一个方面涉及本文所述的突变型植物、非天然存在的植物、杂交植物或转基因植物的种子。优选地,所述种子是烟草种子。另一方面涉及本文所述的突变型植物、非天然存在的植物、杂交植物或转基因植物的花粉或胚珠。另外,提供了如本文所述的突变型植物、非天然存在的植物、杂交植物或转基因植物,其还含有赋予雄性不育的核酸。

[0269] 还提供了如本文所述的突变型植物、非天然存在的植物、杂交植物或转基因植物或其一部分的可再生细胞的组织培养物,其中培养物再生能够表达亲本的所有形态和生理

特征的植物。可再生细胞包括但不限于来自以下的细胞：叶、花粉、胚、子叶、下胚轴、根、根尖、花药、花及其一部分、胚珠、枝条、茎、秆、髓和荚膜或源自其的愈伤组织或原生质体。

[0270] 另一方面涉及源自或可源自突变型、非天然存在的或转基因植物或细胞的烘烤的植物材料，如烘烤的叶或烘烤的烟草，其中本文所述聚核苷酸中的一或多种的表达或者由其编码的蛋白质的活性是降低的并且其导致其中天冬酰胺的水平降低。

[0271] 合适地，所述植物（例如，叶）的视觉外观与对照植物基本上相同。合适地，所述植物是烟草植物。

[0272] 实施例还涉及用于产生突变型、非天然存在的或转基因植物或植物细胞的组合物和方法，所述植物或植物细胞已进行修饰，以调节本文所述聚核苷酸或多肽中的一种或多种的表达或活性，所述修饰可以产生天冬酰胺水平降低的植物或植物组分（例如叶，如经烘烤或干燥的叶）或植物细胞。

[0273] 根据本文所述方法获得的突变型、非天然存在的或转基因植物在视觉外观上与对照植物相似或基本上相同。在一个实施例中，突变型、非天然存在的或转基因植物的叶重与对照植物基本上相同。在一个实施例中，突变型、非天然存在的或转基因植物的叶数目与对照植物基本上相同。在一个实施例中，突变型、非天然存在的或转基因植物的叶重和叶数目与对照植物基本上相同。在一个实施例中，例如在田间移植后一、二或三或更多个月或者在打顶后10、20、30或36或更多天，突变型、非天然存在的或转基因植物的秆高与对照植物基本上相同。例如，突变型、非天然存在的或转基因植物的秆高不低于对照植物的秆高。在另一个实施例中，突变型、非天然存在的或转基因植物的叶绿素含量与对照植物基本上相同。在另一个实施例中，突变型、非天然存在的或转基因植物的秆高与对照植物基本上相同，并且突变型、非天然存在的或转基因植物的叶绿素含量与对照植物基本上相同。在其它实施例中，突变型、非天然存在的或转基因植物的叶的大小、或形状、或数目、或着色与对照植物基本上相同。合适地，所述植物是烟草植物。

[0274] 在另一方面，提供了用于调节（例如，降低）植物的至少一部分（例如叶，如经烘烤或干燥的叶，或者烟草）中的天冬酰胺的量的方法，其包含以下步骤：(i) 调节（例如，降低）本文所述多肽中的一种或多种的表达或活性，合适地，其中多肽由本文所述的相应聚核苷酸序列编码；(ii) 测量在步骤(i)中获得的突变型、非天然存在的或转基因植物的至少一部分（例如叶，如经烘烤的叶）中的天冬酰胺含量；和(iii) 鉴定与对照植物相比较，其中的天冬酰胺含量已降低的突变型、非天然存在的或转基因植物。合适地，所述突变型、非天然存在的或转基因植物的视觉外观与对照植物基本上相同。合适地，所述植物是烟草植物。

[0275] 在另一方面，提供了用于降低经烘烤或干燥的植物材料的至少一部分（如经烘烤或干燥的叶）中天冬酰胺的量的方法，其包含以下步骤：(i) 降低本文所述多肽中的一种或多种的表达或活性，合适地，其中多肽由本文所述的相应聚核苷酸序列编码；(ii) 收获植物材料（如叶中的一或多种），且烘烤或干燥一段时间；(iii) 测量在步骤(ii)中获得的经烘烤或干燥的植物材料的至少一部分中的天冬酰胺含量；和(iv) 鉴定与对照植物相比较，其中的天冬酰胺含量已降低的经烘烤或干燥的植物材料。

[0276] 与对照相比较，表达中的降低可为约5%至约100%，或者至少10%、至少20%、至少25%、至少30%、至少40%、至少50%、至少60%、至少70%、至少75%、至少80%、至少90%、至少95%、至少98%或100%的降低，这包括转录活性或聚核苷酸表达或多肽表达或

其组合中的降低。

[0277] 与对照相比较,活性中的降低可为约5%至约100%,或者至少10%、至少20%、至少25%、至少30%、至少40%、至少50%、至少60%、至少70%、至少75%、至少80%、至少90%、至少95%、至少98%或100%的降低。

[0278] 本文所述的聚核苷酸和重组构建体可以用于调节本文所述的酶在目的植物物种(合适地,烟草)中的表达。

[0279] 许多基于聚核苷酸的方法可以用于增加例如在植物和植物细胞中的基因表达。作为实例,可以制备与待转化的植物相容的构建体、载体或表达载体,其包含目的基因连同能够在植物或植物细胞中过表达所述基因的上游启动子。示例性启动子在本文中描述。在转化后且当生长在合适条件下时,所述启动子可以驱动表达,以便调节(例如,降低)这种酶在植物或在其特定组织中的水平。在一个示例性实施例中,生成携带本文所述的一种或多种聚核苷酸(或如本文所述的其任何组合)的载体,以在植物或植物细胞中过表达所述基因。所述载体携带位于转基因上游的合适启动子(如花椰菜花叶病毒CaMV 35S启动子),从而驱动所述转基因在植物的所有组织中的组成型表达。所述载体还携带抗生素抗性基因,以便对经转化的愈伤组织和细胞系赋予选择。

[0280] 多个实施例涉及通过将聚核苷酸的多个拷贝整合到植物基因组内,用于降低本文所述的一种或多种聚核苷酸的表达水平的方法,其包含:用包含可操作地连接于本文所述的一种或多种聚核苷酸的启动子的表达载体转化植物细胞宿主。由重组聚核苷酸编码的多肽可以是天然多肽,或对于细胞可以是异源的。

[0281] 携带本文所述的一种或多种聚核苷酸(或如本文所述的其任何组合)的突变等位基因的植物可以用于植物育种计划,以产生有用的品系、品种和杂种。特别地,可以使所述突变等位基因基因渗入上述商业上重要的品种内。因此,提供了用于植物育种的方法,其包含将如本文所述的突变型植物、非天然存在的植物或转基因植物与含有不同遗传一致性的植物进行杂交。所述方法可以进一步包括将后代植物与另一植物杂交,且任选重复杂交直到获得具有期望的遗传性状或遗传背景的后代。此类育种方法发挥的一个目的是将期望的遗传性状引入其它品种、育种品系、杂种或栽培品种,尤其是具有商业利益的那些。另一个目的是便于在单一植物品种、品系、杂种或栽培品种中叠加不同基因的遗传修饰。考虑种内和种间交配。起于此类杂交的后代植物也称为育种品系,是本公开的非天然存在的植物的实例。

[0282] 在一个实施例中,提供了用于产生非天然存在的植物的方法,其包括:(a)将突变型或转基因植物与第二植物杂交,以获得后代烟草种子;(b)在植物生长条件下,使后代种子生长,以获得非天然存在的植物。该方法还可以包括:(c)将非天然存在的植物的上一代与自身或另一植物杂交,以获得后代种子;(d)在植物生长条件下,使步骤(c)的后代种子生长,以获得另外的非天然存在的植物;和(e)将(c)和(d)的杂交和生长步骤重复多次,以获得非天然存在的植物的更多世代。所述方法可以任选包括在步骤(a)之前提供亲本植物的步骤,所述亲本植物包含得到表征且不同于突变型或转基因植物的遗传一致性。在一些实施例中,取决于育种计划,杂交和生长步骤重复0至2次、0至3次、0至4次、0至5次、0至6次、0至7次、0至8次,0至9次或0至10次,以便生成非天然存在的植物的世代。回交是此类方法的实例,其中后代与其亲本之一或与其亲本遗传相似的另一植物进行杂交,以便获得在下一

代中具有更接近于亲本之一的遗传一致性的后代植物。用于植物育种、特别是植物育种的技术，是众所周知的且可用于本公开的方法中。本公开还提供了通过这些方法产生的非天然存在的植物。某些实施例排除选择植物的步骤。

[0283] 在本文所述方法的一些实施例中，使用标准田间程序，在田间评估起因于育种和筛选变体基因的品系。包括原始未诱变亲本的对照基因型包括在内，并且按随机化的完全区组设计或其它适当的田间设计，将入选者(entry)排列于田间。对于烟草，使用标准的农学实践，例如将烟草收获、称量且取样，用于在烘烤或干燥之前和期间的化学及其它常见测试。执行数据的统计分析，以确认所选择品系与亲本品系之间的相似性。任选执行所选植物的细胞遗传学分析，以确认染色体组和染色体配对关系。

[0284] DNA指纹鉴定、单核苷酸多态性、微卫星标记或类似技术可用在标记辅助选择(MAS)的育种计划中，以如本文所述的，将基因的突变等位基因转移或培育到其它植物内。例如，育种者可由含有突变等位基因的基因型与农学期望的基因型的杂交而产生分离的群体。可使用本文所列出的技术之一，使用从基因组序列或其片段所开发的标记来筛选F₂中的植物或回交世代。鉴定为具有突变等位基因的植物可以回交或自花授粉，以产生待筛选的第二群体。取决于预期遗传模式或所用MAS技术，有必要在每轮回交之前对所选择的植物进行自花授粉，以帮助鉴定期望的个体植物。可重复进行回交或其它育种操作，直到恢复轮回亲本的所需表型。

[0285] 在育种计划中，成功的杂交获得能育的F₁植物。所选择的F₁植物可与亲本之一杂交，并且第一回交世代植物进行自花授粉，以产生再次筛选变体基因表达(例如，基因的无效版本)的群体。将回交、自花授粉和筛选的过程重复例如至少4次，直到最终筛选产生能育且与轮回亲本相当相似的植物。如果需要的话，这种植物进行自花授粉，并且随后再次筛选后代，以确认植物显示出变体基因表达。在一些实施例中，在F₂代的植物群体中筛选变体基因表达，例如，根据标准方法如通过使用具有引物的PCR方法，鉴定由于基因缺失而未能表达多肽的植物，所述引物基于本文所述的聚核苷酸(或如本文所述的其任何组合)的核苷酸序列信息。

[0286] 杂交品种可通过以下方式产生：阻止第一品种的雌性亲本植物(即，种子亲本)的自花授粉，允许来自第二品种的雄性亲本植物的花粉使雌性亲本植物受精，且允许F₁杂种种子在雌性植物上形成。可通过在花发育早期阶段将花朵去雄，来阻止雌性植物的自花授粉。或者，可使用雄性不育的形式阻止在雌性亲本植物上的花粉形成。例如，通过细胞质雄性不育(CMS)或转基因雄性不育可产生雄性不育，其中转基因抑制小孢子和/或花粉形成、或自交不相容。含有CMS的雌性亲本植物是特别有用的。在雌性亲本植物是CMS的实施例中，从雄性能育植物收获花粉且人工施用于CMS雌性亲本植物的柱头，并且收获所得到的F₁种子。

[0287] 本文所述品种和品系可用于形成单杂交F₁杂种。在此类实施例中，亲本品系的植物可生长为基本上同质的相邻群体，以便于雄性亲本植物与雌性亲本植物的天然异花授粉。通过常规方式选择性收获在雌性亲本植物上形成的F₁种子。还可大批种植两个亲本品系品种，且收获由于自花授粉在雌性亲本上形成的F₁杂种种子和在雄性亲本上形成的种子的掺和物。或者，可进行三系杂交，其中单杂交F₁杂种用作雌性亲本，并且与不同的雄性亲本杂交。作为另一选择，可产生双杂交杂种，其中两个不同单杂交的F₁后代进行自身杂交。

[0288] 可在突变型、非天然存在的或转基因植物群体中,筛选或选择具有所需性状或表型的那些群体成员。例如,可在单一转化事件的后代群体中,筛选具有由其编码的多肽的所需表达或活性水平的那些植物。物理和生物化学方法可用于鉴定表达或活性水平。这些包括用于检测聚核苷酸的DNA分析或PCR扩增;用于检测RNA转录物的RNA印迹、S1RNA酶保护、引物延伸或RT-PCR扩增;用于检测多肽和聚核苷酸的酶或核酶活性的酶促分析;以及检测多肽的蛋白质凝胶电泳、蛋白质印迹、免疫沉淀和酶联免疫分析。其它技术如原位杂交、酶染色和免疫染色以及酶分析也可用于检测多肽或聚核苷酸的存在或表达或活性。

[0289] 如本文所述的突变型、非天然存在的或转基因植物细胞和植物包含一种或多种重组聚核苷酸、一种或多种聚核苷酸构建体、一种或多种双链RNA、一种或多种结合物或者一种或多种载体/表达载体。

[0290] 无限制地,在表达或活性已根据本公开进行调节之前或之后,本文所述的植物可出于其它目的进行修饰。下列遗传修饰中的一种或多种可存在于突变型、非天然存在的或转基因植物中。在一个实施例中,修饰涉及含氮代谢中间产物转化的一种或多种基因,使得植物(如叶)在烘烤时产生比对照植物更低水平的至少一种烟草特异性亚硝胺。如本文所述,可进行修饰的基因的非限制性实例包含编码天冬酰胺合成酶的基因,如CYP82E4、CYP82E5和CYP82E10,其参与烟碱向降烟碱的转化,并且描述在W02006091194、W02008070274、W02009064771和PCT/US2011/021088中且如本文所述。在另一个实施例中,修饰涉及重金属摄取或重金属转运的一种或多种基因,使得植物或植物部分(如叶)与不含修饰的对照植物或其部分相比具有更低重金属含量。非限制性实例包括多药抗性相关蛋白质家族、阳离子扩散协助蛋白(CDF)家族、Zrt-,Irt样蛋白(ZIP)家族、阳离子交换剂(CAX)家族、铜转运蛋白(COPT)家族、重金属P型ATP酶家族(例如HMA,如W02009074325中所述)、天然抗性相关巨噬细胞蛋白(NRAMP)同系物家族、以及参与重金属如镉转运的ATP结合盒(ABC)转运蛋白家族(例如MRP,如W02012/028309中所述)中的基因。如本文所用,术语重金属包括过渡金属。其它修饰的实例包括除草剂耐受性,例如,草甘膦是许多广谱除草剂的活性成分。通过转移aroA基因(来自鼠伤寒沙门氏菌(*Salmonella typhimurium*)和大肠杆菌(*E.coli*)的草甘膦EPSP合成酶),已开发草甘膦抗性转基因植物。通过转化来自拟南芥属(*Arabidopsis*)的突变型ALS(乙酰乳酸合成酶)基因,已产生磺脲抗性植物。来自突变型绿穗苋(*Amaranthus hybridus*)的光系统II的OB蛋白已转移到植物内,以产生莠去津抗性转基因植物;并且通过掺入来自细菌肺炎克雷伯氏菌(*Klebsiella pneumoniae*)的bxn基因,已产生溴苯腈抗性转基因植物。另一示例性修饰导致对昆虫具有抗性的植物。苏云金芽孢杆菌(*Bacillus thuringiensis*) (Bt)毒素可提供延迟Bt抗性害虫出现的有效途径,如最近在绿花椰菜中所示,其中金字塔式的cry1Ac和cry1C Bt基因控制小菜蛾(diamondback moth)对任一单一蛋白质的抗性,且明显延迟抗性昆虫的进化。另一示例性修饰产生对由病原体(例如病毒、细菌、真菌)引起的疾病具有抗性的植物。已工程改造了表达Xa21基因(对白叶枯病的抗性)的植物,以及表达Bt融合基因和壳多糖酶基因(对三化螟的抗性和对鞘(sheath)的耐受性)两者的植物。另一示例性修饰导致改变的生殖能力,例如雄性不育。另一示例性修饰导致耐受非生物胁迫(例如,干旱、温度、盐度)的植物,并且通过转移来自拟南芥属的酰基甘油磷酸酶,已产生耐受的转基因植物;编码甘露醇脱氢酶和山梨糖醇脱氢酶的基因改善抗旱性,所述甘露醇脱氢酶和山梨糖醇脱氢酶涉及甘露醇和山梨糖醇合成。

其它示例性修饰可导致具有改善的贮藏蛋白和油的植物、具有增强的光合效率的植物、具有延长的贮存期限的植物、具有增强的碳水化合物含量的植物以及对真菌具有抗性的植物；编码涉及生物碱生物合成的酶的植物。也可涵盖这样的转基因植物，其中S-腺苷-L-甲硫氨酸 (SAM) 和/或胱硫醚 γ -合酶 (CGS) 的表达已被调节。

[0291] 一种或多种此类性状可基因渗入来自另一栽培品种的突变型、非天然存在的或转基因植物，或可直接转化到其内。性状基因渗入到突变型、非天然存在的或转基因植物中可通过本领域中已知的任何植物育种方法实现，例如谱系育种、回交、双单倍体育种等（参见 Wernsman, E. A 和 Rufty, R. C. 1987. 第十七章. Tobacco. 第669-698页于: Cultivar Development. Crop Species. W. H. Fehr (编), MacMillan Publishing Co, Inc., New York, N. Y 第761页）。上述基于分子生物学的技术特别是RFLP和微卫星标记，可用于此类回交中，以鉴定与轮回亲本具有最高遗传一致性程度的后代。这允许加速生产与轮回亲本具有至少90%、优选至少95%、更优选至少99%遗传一致性、再更优选与轮回亲本遗传一致的品种，且还包括由供体亲本基因渗入的性状。此类遗传一致性的确定可基于本领域已知的分子标记。

[0292] 最后的回交世代可进行自交，以给出对于被转移的核酸纯的育种后代。除转移的性状（例如，一种或多种单基因性状）之外，所得到的植物一般具有突变型、非天然存在的或转基因烟草植物的基本上所有形态和生理特征。确切的回交方案将取决于被改变的性状，以确定合适的测试方案。虽然当被转移的性状是显性等位基因时，回交方法是简化的，但也可转移隐性等位基因。在这种情况下，可能有必要引入对后代的测试，以确定是否已成功转移所需性状。

[0293] 多个实施例提供了突变型植物、非天然存在的植物或转基因植物以及生物质，其中聚核苷酸（或如本文所述的其任何组合）的表达水平被降低，以降低其中的天冬酰胺含量。

[0294] 此类植物特别是烟草植物的一部分，以及更特别地烟草植物的叶片和中脉，可掺入或用于制备多种消费品，包括但不限于气溶胶形成材料、气溶胶形成装置、吸烟制品、可抽吸制品、无烟产品和烟草产品。气溶胶形成材料的实例包括但不限于烟草组合物、烟草、烟草提取物、烟丝、切丝填料、经烘烤或干燥的烟草、膨胀烟草、均质烟草、再造烟草和烟斗烟草。吸烟制品和可抽吸制品是气溶胶形成装置的类型。吸烟制品或可抽吸制品的实例包括但不限于香烟、小雪茄和雪茄。无烟产品的实例包括嚼烟和鼻烟。在某些气溶胶形成装置而不是燃烧中，烟草组合物或另一气溶胶形成材料被例如一个或多个电加热元件或碳热源进行加热，以产生气溶胶。通常在此类加热式吸烟制品中，通过从热源到物理上分开的气溶胶形成基质或材料的热转移而生成气溶胶，所述气溶胶形成基质或材料可位于热源内、热源周围或热源下游。在抽吸期间，挥发性化合物通过来自热源的热传递从气溶胶形成基质释放，并且夹带在被抽吸穿过吸烟制品的空气中。随着所释放的化合物冷却，化合物冷凝以形成由使用者吸入的气溶胶。此类装置包括例如电加热的气溶胶生成装置，其中通过从气溶胶生成装置到加热式吸烟制品的气溶胶形成基质的热转移而生成气溶胶。合适地，在气溶胶形成基质的加热期间，不发生烟草的燃烧。合适的气溶胶形成制品描述在W02013/098405中并且包含气溶胶形成基质，用于在被气溶胶生成装置的内部加热元件加热时生成可吸入的气溶胶。它可以包含具有内部加热元件的电加热气溶胶生成装置。它还可以线性

顺序排列包含气溶胶形成基质、位于气溶胶形成基质紧下游的支撑元件、位于支撑元件下游的气溶胶冷却元件以及围绕气溶胶形成基质、支撑元件和气溶胶冷却元件的外包装。支撑元件可以邻接气溶胶形成基质。气溶胶形成基质可被气溶胶生成装置的加热元件穿透。

[0295] 如本文所用,术语“燃烧”是指氧化还原化学反应,其中反应物分子(即,燃料和氧化剂)混合并重新排列成同时释放热量的产物分子。燃烧可以通过在气态产物中存在相关量的氮氧化物而肯定地指示出来,氮氧化物不是由存在于原始反应物基质中的硝酸盐的分解所形成,并且明确证明同时发生总体放热过程。燃烧的气体产物中氮氧化物的相关量的演变可以通过比较在相关条件中(例如在空气中)形成的氮氧化物与在相同条件下但在没有的氧气(例如,在纯氮气或氦气气氛中)形成的氮氧化物的总量来确定。在加热气溶胶形成装置的另一类型中,通过将热量从可燃性燃料元件或热源转移到物理上分开的气溶胶形成材料来产生气溶胶,所述气溶胶形成材料可位于热源内、热源周围或热源下游。无烟烟草产品和多种含烟草的气溶胶形成材料可含有任何形式的烟草,包括沉积在其它成分上、混合于其它成分中、由其它成分包围或以其它方式与其它成分组合的干燥颗粒、碎片、小颗粒、粉末或浆料,所述其它成分采取任何形式,如絮片、膜、卡(tab)、泡沫或珠。如本文所用,术语‘烟’用于描述一类气溶胶,其由吸烟制品如可燃香烟、或通过燃烧气溶胶形成材料而产生。

[0296] 在某些实施例中,加热而不燃烧植物材料是优选的。

[0297] 在一个实施例中,还提供了来自本文所述的突变型、转基因和非天然存在的烟草植物的经烘烤的植物材料。烘烤绿色烟叶的工艺是本领域普通技术人员已知的,并且包括但不限于空气烘烤、明火烘烤、烟道烘烤和日光烘烤。烘烤绿色烟叶的工艺取决于收获的烟草类型。例如,白肋烟和某些深色品系通常是晾制的,而烟斗烟草、嚼烟和鼻烟通常是明火烘烤的。

[0298] 在另一个实施例中,还提供了来自本文所述的突变型、转基因和非天然存在的植物的经干燥的植物材料。干燥叶的工艺是本领域普通技术人员已知的,包括但不限于风干和晒干。干燥叶的确切工艺取决于收获的植物类型。合适地,植物材料在收获之后干燥。因此,本文预期使用干燥材料和收获后的干燥材料。干燥过程可以激活一个或多个衰老相关基因。可以在烘烤或干燥期间监测本文所述的基因和蛋白质的活性的表达。同样,可以在烘烤或干燥期间监测天冬酰胺、尼古丁、谷氨酰胺、天冬氨酸和谷氨酸的水平。举例来说,可以在烘烤或干燥0、10、20、30、40、50、60、70、80、90或100小时后进行测量。进一步举例来说,可以在0天、1天、2天、3天、4天、5天或10天或更长时间之后进行测量。

[0299] 在另一个实施例中,描述了包括含有烟草的气溶胶形成材料的烟草产品,所述气溶胶形成材料包含来自本文所述的突变型烟草植物、转基因烟草植物或非天然存在的烟草植物的植物材料,例如叶,优选经烘烤或干燥的叶。本文所述的烟草产品可以是混合的烟草产品,其还可包含未修饰的烟草。

[0300] 在烘烤或干燥(例如烘烤或干燥3天)后,本文所述的植物、植物部分、植物材料、植物产品或烟草产品中天冬酰胺的量比对照植物低约20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%和100%,如低约200%或300%。合适地,在烘烤或干燥(例如烘烤或干燥3天)后,本文所述的植物、植物部分、植物材料、植物产品或烟草产品中天冬酰胺的量比对照低至少约60%、65%、70%、

75%、80%、85%、90%、95%、96%、97%、98%、99%和100%，如低约200%或300%，更合适地，比对照低至少约80%或低至少约90%。

[0301] 在烘烤或干燥(例如烘烤或干燥3天)后，本文所述的植物、植物部分、植物材料、植物产品或烟草产品中天冬酰胺的量为约40mg/g干重叶或更少、更合适地约35mg/g干重叶或更少、更合适地约30mg/g干重叶或更少、更合适地约25mg/g干重叶或更少、更合适地约20mg/g干重叶或更少、更合适地约15mg/g干重或更少、更合适地约12mg/g干重、更合适地约10mg/g干重干重叶或更少、或更合适地约5mg/g干重叶或更少。合适地，从植物的中下部秆位收获叶。

[0302] 在烘烤或干燥后源自植物、植物部分、植物材料、植物产品或烟草产品并且通过加热获得的气溶胶中丙烯酰胺的量比对照低至少约70%。举例来说，在烘烤或干燥后源自烟草产品并且通过加热获得的气溶胶中丙烯酰胺的量为每根香烟约1.5 μ g。

[0303] 在烘烤或干燥后源自植物、植物部分、植物材料、植物产品或烟草产品并且通过燃烧获得的气溶胶中丙烯酰胺的量比对照低至少约55%。举例来说，在烘烤或干燥后源自烟草产品并且通过燃烧获得的气溶胶中丙烯酰胺的量为每根香烟约2.5 μ g。

[0304] 合适地，在烘烤或干燥后源自植物、植物部分、植物材料、植物产品或烟草产品并且通过燃烧或加热获得的气溶胶中丙烯酰胺的量比对照低至少约55%至约70%。举例来说，在烘烤或干燥后源自烟草产品并且通过燃烧或加热获得的气溶胶中丙烯酰胺的量在每根香烟约1.5 μ g和每根香烟2.5 μ g之间。

[0305] 植物、植物部分、植物材料、植物产品、烟草产品或气溶胶中尼古丁的量可以与来自对照植物的植物、植物部分、植物材料、可抽吸制品、无烟产品的尼古丁的量基本上相同。举例来说，尼古丁的量可以在对照植物中所存在的尼古丁的量的约20%、15%、10%、9%、8%、7%、6%、5%、4%、3%、2%或1%内或更少。合适地，尼古丁的量在对照植物中所存在的尼古丁的量的约10%、9%、8%、7%、6%、5%、4%、3%、2%或1%内或更少。尼古丁的量通常在约15和20mg/g叶的干重生物物质之间。

[0306] 在一个实施例中，与对照相比，烘烤后的植物、植物部分、植物材料、植物产品或烟草产品(例如经烘烤的铸叶)中天冬酰胺的量低至少约20%(例如低至少约22%)，并且源自烘烤后的植物、植物部分、植物材料、植物产品或烟草产品(例如经烘烤的铸叶)的气溶胶中丙烯酰胺的量低至少约25%(例如低至少约24%或更低)。在另一个实施例中，烘烤后的植物、植物部分、植物材料、植物产品或烟草产品(例如经烘烤的铸叶)中天冬酰胺的量为约12mg/g干重或更少，并且源自烘烤后的植物、植物部分、植物材料、植物产品或烟草产品(例如经烘烤的铸叶)的气溶胶中丙烯酰胺的量为约2.5 μ g/香烟或更少。合适地，尼古丁水平与对照中的尼古丁水平基本上相同。

[0307] 在另一个实施例中，与对照相比，烘烤后的植物、植物部分、植物材料、植物产品或烟草产品(例如经烘烤的铸叶)中天冬酰胺的量低至少约70%(例如低至少约66%)，并且源自烘烤后的植物、植物部分、植物材料、植物产品或烟草产品的气溶胶中丙烯酰胺的量低至少约45%(例如低至少约44%)。在另一个实施例中，烘烤后的植物、植物部分、植物材料、植物产品或烟草产品(例如经烘烤的铸叶)中天冬酰胺的量为约12mg/g干重或更少，并且源自烘烤后的植物、植物部分、植物材料、植物产品或烟草产品的气溶胶中丙烯酰胺的量为约3 μ g/香烟或更少。合适地，尼古丁水平与对照中的尼古丁水平基本上相同。

[0308] 在另一个实施例中,与对照相比,烘烤后的植物、植物部分、植物材料、植物产品或烟草产品(例如经烘烤的铸叶)中天冬酰胺的量低至少约90%(例如至少约88%或更低),并且源自烘烤后的植物、植物部分、植物材料、植物产品或烟草产品的气溶胶中丙烯酰胺的量低至少约70%。在另一个实施例中,烘烤后的植物、植物部分、植物材料、植物产品或烟草产品(例如经烘烤的铸叶)中天冬酰胺的量为约5mg/g干重或更少,并且源自烘烤后的植物、植物部分、植物材料、植物产品或烟草产品的气溶胶中丙烯酰胺的量为约1.5μg/香烟或更少。合适地,尼古丁水平与对照中的尼古丁水平基本上相同。

[0309] 在另一个实施例中,经由RNAi调节NtASN1的表达可以:(i)使烘烤或干燥后的植物、植物部分、植物材料、植物产品或烟草产品(例如经烘烤的铸叶)中天冬酰胺的量降低24%至约89%(例如达到约35mg/g至约5mg/g);(ii)使尼古丁的水平维持在对照植物中所存在的尼古丁的量的约20%、15%、10%、9%、8%、7%、6%、5%、4%、3%、2%或1%或更少(例如约15和20mg/g叶的干重生物物质);和(iii)使气溶胶中丙烯酰胺的水平降低至少约70%达到1.5 3μg/香烟或更少。叶生物物质的量一般在约300-400g之间。

[0310] 在另一个实施例中,经由RNAi调节NtASN5的表达可以:(i)使烘烤或干燥后的植物、植物部分、植物材料、植物产品或烟草产品(例如经烘烤的铸叶)中天冬酰胺的量降低约35%至约66%(例如达到约31mg/g至约18mg/g);(ii)使尼古丁的水平维持在对照植物中所存在的尼古丁的量的约20%、15%、10%、9%、8%、7%、6%、5%、4%、3%、2%或1%或更少(例如约15和20mg/g叶的干重生物物质);和(iii)使气溶胶中丙烯酰胺的水平降低至少约44%达到3μg/香烟或更少。叶生物物质的量一般在约300-400g之间。

[0311] 在另一个实施例中,经由诱变调节NtASN1的表达可以:(i)使烘烤或干燥后的植物、植物部分、植物材料、植物产品或烟草产品(例如经烘烤的铸叶)中天冬酰胺的量降低约17%至约83%(例如达到约19.4mg/g至约4.7mg/g);和(iii)使气溶胶中丙烯酰胺的水平降低至少约24%达到约2.5μg/香烟或更少。

[0312] 在另一个实施例中,经由诱变调节NtASN5的表达可以使烘烤或干燥后的植物、植物部分、植物材料、植物产品或烟草产品(例如经烘烤的铸叶)中天冬酰胺的量降低约44%(例如达到约14.5mg/g)。突变型、非天然存在的或转基因植物可以在例如农业中具有其它用途。例如,本文所述的突变型、非天然存在的或转基因植物可用于制备动物饲料和人类食物产品。

[0313] 本公开还提供了用于产生种子的方法,其包括培养本文所述的突变型植物、非天然存在的植物或转基因植物,并且从栽培的植物收集种子。来自本文所述植物的种子可通过本领域中已知的方式进行条件处理,且包装在包装材料中,以形成制造物品。包装材料如纸和布是本领域众所周知的。种子的包装可带有描述其中种子的性质的标记,例如固定至包装材料的标签或标记、印刷在包装上的标记。

[0314] 用于植物基因分型以鉴定、选择或育种的组合物、方法和试剂盒可包括检测聚核苷酸样品中的聚核苷酸(或如本文所述的其任何组合)存在的手段。相应地,描述了组合物,其包含用于特异性扩增聚核苷酸中的一种或多种的至少一部分的一个或多个引物,以及用于进行扩增或检测的任选的一个或多个探针和任选的一种或多种试剂。

[0315] 相应地,公开了基因特异性的寡核苷酸引物或探针,其包含对应于本文所述的聚核苷酸的约10个或更多个邻接聚核苷酸。所述引物或探针可包含以下或由以下组成:约15、

20、25、30、40、45或50个或更多个邻接聚核苷酸,所述引物或探针与本文所述的聚核苷酸杂交(例如,特异性杂交)。在一些实施例中,所述引物或探针可包含以下或由以下组成:约10至50个邻接核苷酸、约10至40个邻接核苷酸、约10至30个邻接核苷酸、或约15至30个邻接核苷酸,其可用于基因鉴定(例如,DNA杂交)、或分离(例如,细菌菌落或噬菌体噬菌斑的原位杂交)、或基因检测(例如,作为核酸扩增或检测中的一个或多个扩增引物)的序列依赖性方法。可设计一个或多个特异性引物或探针,且用于扩增或检测聚核苷酸的部分或全部。作为具体实例,两个引物可用于聚合酶链反应方案中,以扩增编码核酸如DNA或RNA的核酸片段。还可使用源自核酸序列的一个引物和第二引物(其与核酸序列上游或下游的序列杂交,所述序列如启动子序列、mRNA前体的3'端或源自载体的序列)执行聚合酶链反应。可用于聚核苷酸体外扩增的热和等温技术的实例是本领域众所周知的。样品可以是或可源自植物、植物细胞或植物材料,或者由如本文所述的植物、植物细胞或植物材料制备或衍生的烟草产品。

[0316] 在另一方面,还提供了检测样品中本文所述的聚核苷酸(或如本文所述的其任何组合)的方法,其包括以下步骤:(a)提供包含或疑似包含聚核苷酸的样品;(b)使所述样品与多个引物之一或一个或多个探针接触,用于特异性检测聚核苷酸的至少一部分;和(c)检测扩增产物的存在,其中扩增产物的存在指示样品中存在聚核苷酸。在另一方面,还提供了一个或多个引物或探针用于特异性检测聚核苷酸的至少一部分的用途。还提供了用于检测聚核苷酸的至少一部分的试剂盒,其包含用于特异性检测聚核苷酸的至少一部分的多个引物或探针之一。试剂盒可包含用于聚核苷酸扩增(如PCR)的试剂,或用于探针杂交检测技术(如DNA印迹、RNA印迹、原位杂交或微阵列)的试剂。试剂盒可包含用于抗体结合检测技术(如蛋白质印迹、ELISA、SELDI质谱法或测试条)的试剂。试剂盒可包含用于DNA测序的试剂。试剂盒可包含用于至少测定植物材料、经烘烤或干燥的植物材料或经烘烤或干燥的叶中的天冬酰胺含量和/或丙烯酰胺气溶胶含量的试剂和说明书。

[0317] 在一些实施例中,试剂盒可包含用于所述方法中的一种或多种的说明书。所述试剂盒可用于遗传一致性确定、系统发生研究、基因分型、单倍体分型、谱系分析或植物育种,特别是共显性评分。

[0318] 本公开还提供了对包含如本文所述的聚核苷酸的植物、植物细胞或植物材料进行基因分型的方法。基因分型提供了区分染色体对的同源物的手段,并且可用于区分植物群体中的分离体。分子标记方法可用于系统发生研究、表征作物品种之间的遗传关系、鉴定杂交或体细胞杂种、定位影响单基因性状的染色体区段、图位克隆和定量遗传研究。基因分型的具体方法可采用任意数目的分子标记分析技术,包括扩增片段长度多态性(AFLP)。AFLP是由核苷酸序列变异性引起的扩增片段之间的等位基因差异的产物。因此,本公开还提供了使用如AFLP分析这样的技术来跟踪一种或多种基因或核酸以及遗传连锁至这些基因或核酸的染色体序列的分离的手段。

[0319] 在一个实施例中,还提供了来自本文所述的突变型、转基因和非天然存在的植物的经烘烤或干燥的植物材料。例如,烘烤或干燥烟叶的工艺是本领域技术人员已知的,并且包括但不限于空气烘烤、明火烘烤、烟道烘烤和日光烘烤。烘烤绿色烟叶的工艺取决于如本文所述收获的烟草类型。

[0320] 在另一个实施例中,描述了烟草产品,其包括包含来自本文所述的突变型、转基因

和非天然存在的植物的植物材料(如叶,经适当烘烤的植物材料,如经烘烤或干燥的叶)的烟草产品,或通过本文所述的方法生产的烟草产品。本文所述的烟草产品还可包含未修饰的烟草。

[0321] 在另一个实施例中,描述了烟草产品,其包含来自本文所述的突变型、转基因和非天然存在的植物的植物材料,优选叶,如经烘烤或干燥的叶。例如,所述植物材料可加入烟草产品内部或外部,并且因此在燃烧后释放期望的香气。根据此实施例的烟草产品甚至可以是未修饰的烟草或经修饰的烟草。根据此实施例的烟草产品甚至可源自突变型、转基因或非天然存在的植物,其在除了本文公开的基因之外的一种或多种基因中具有修饰。

[0322] 本发明还在下文实例中描述,提供所述实例以更详细地描述本发明。这些实例阐述目前考虑用于进行本发明的优选方式,意欲说明而不是限制本发明。

[0323] 实例

[0324] 实例1-经烘烤的烟叶中天冬酰胺的合成

[0325] 在烟叶并且特别是白肋烟栽培品种的叶烘烤期间,主动产生天冬酰胺,占经烘烤的叶中所存在的总氨基酸的约50%。在图1中,显示了在同一田间和晾房中平行生长和晾制的三种白肋烟草栽培品种的经烘烤的叶中天冬酰胺和总游离氨基酸的增加,主要由天冬酰胺增加引起。令人惊讶的是,天冬酰胺是在烘烤期间明显积累的唯一氨基酸。考虑到天冬酰胺是在早期烘烤阶段合成的,因此反应依赖于使用谷氨酰胺和天冬氨酸作为底物的天冬酰胺合成酶活性。然而,在图1中,与天冬酰胺相比,Asp和Gln仅显示很小的增加,因此表明在长时间的晾制工艺(>2个月)期间Gln和Asp的库可能变化。没有观察到谷氨酸库的变化(数据未显示)。

[0326] 图2表明Gln在烘烤的前4天(96小时)增加,然后在天冬酰胺积累开始时降低。相比之下,天冬氨酸库基本保持平坦,尽管可以观察到一些波动。不希望受任何具体的理论所束缚,所观察到的谷氨酰胺库可用作天冬酰胺合成酶在烘烤的前4-5天期间产生天冬酰胺的底物(参见图3)。

[0327] 如上所述,在烘烤期间观察到的天冬酰胺累积可能由衰老诱导的天冬酰胺合成酶的特异性活性引起。一旦加热或燃烧,天冬酰胺和还原糖就有可能通过美拉德反应(Maillard reaction)在气溶胶中生成丙烯酰胺。经烘烤的烟叶或烟草混合物中积累的天冬酰胺与气溶胶中测量的丙烯酰胺之间存在良好的相关性(参见图4)。不希望受任何具体的理论所束缚,这表明气溶胶中丙烯酰胺的增加严格取决于天冬酰胺的存在。另外,叶组织中天冬酰胺的存在也与所用的烘烤工艺有关,如图5所示。

[0328] 实例2-烟草中的天冬酰胺合成酶基因

[0329] 在烟草中鉴定到6种全长天冬酰胺合成酶基因,并且命名为NtASN1-S、NtASN1-T、NtASN3-S、NtASN3-T、NtASN5-S和NtASN5-T。有趣的是,拟南芥有3个与天冬酰胺合成酶相关的基因AtASN1、AtASN2和AtASN3,番茄只有2个全长基因(Solyc06g007180和Solyc04g055200)。使用从烟草基因组中提取的基因组序列(Sierro等人,2014)得到的假定蛋白质序列进行的系统发生分析表明,两组天冬酰胺合成酶在进化过程中发散(参见图6)。实际上,包含Solyc04g055200,AtASN2和3的一组接近于NtASN3-S和NtASN3-T。属于这一组的AtASN2的基因产物定位于韧皮部中,并且对于绿色组织中的氮同化、分布和再活化显然是必需的(Gaufichon等人,2013Plant Cell Environ. 2月;36(2):328-42.)。这一组属于II

类的双子叶植物子类。包括Soly06g007180、AtASN1、NtASN1-S、NtASN1-T、NtASN5-S和NtASN5-T的第二组属于I类的双子叶植物子类 (Gaufichon等人, 2010 Ann Bot. 2010年6月; 105 (7) : 1141-57.)。AtASN1的过表达有助于增加花和发育中的长角果中的天冬酰胺含量, 并提高种子蛋白质含量, 从而证明天冬酰胺从来源到沉积组织的运输增强。这表明AtASN1组更加致力于氮源向种子的转运, 以允许适当的种子萌发和幼苗对限氮底物的高耐受性 (Lam等人., 2003 Plant Physiol. 2003年6月; 132 (2) : 926-35)。

[0330] 基因产物之间的一致性百分比显示NtASN3-S和NtASN3-T是一致的, 并且非常接近于Soly04g055200和AtASN3 (85%一致性), 表明这一基因簇在进化过程中相对保守, 可能保持营养组织中的氮同化 (见上文)。有趣的是, 每个美花烟草和绒毛状烟草都推导出NtASN1-S/NtASN1-T和NtASN5-S/NtASN5-T的蛋白质拷贝均享有接近于98%的高一致性百分比。这表明ASN1和ASN5拷贝的基因功能在每个烟草祖先之间可能是保守的, 并且可能源自番茄和拟南芥中发现的单个共同祖先。此外, NtASN5-S与Soly06g007180共享95%一致性, 表明具有相似的功能。

[0331] 已知天冬酰胺在白肋烟的晾制过程中迅速增加, 在烘烤的早期阶段 (参见图1和图2) 天冬酰胺的产生可能不是蛋白质或多肽的蛋白水解降解直接产生的, 而是来自烘烤期间的从头合成。如果是这样, 并如上所述, 天冬酰胺合成只能源自活性天冬酰胺合成酶。

[0332] 为了确定在烘烤期间天冬酰胺合成酶基因的表达, 分析使用Tobarray Affymetrix基因芯片 (Martin等人 BMC Genomics 2012, 13:674) 从微阵列分析收集的数据。叶RNA在烘烤的前96小时期间以时程方式分离。鉴定出与六个ASN编码序列匹配的特异性探针 (100%一致性) (NtPMI1g47375e1_st针对ASN1-T; NtPMI1g31395e1_st针对ASN1-S; NtPMI1g25255e1_st针对ASN5-T; NtPMI1g57337e1_st针对ASN5-S; NtPMI1g152348e1_st针对ASN3-T以及NtPMI1g121582e1_st针对ASN3-S)。对于RNA制备, 直接在田间 (-4h) 在Stella (CH-Burley) 的中部秆位收集叶, 时间0是在叶悬挂在晾房中的时刻。ASN3-S和ASN3-T仅在晾制期间在白肋烟叶中弱表达。相比之下, ASN1-S、ASN1-T和ASN5-S在烘烤 (衰老, 变黄阶段) 开始时高度表达, 而ASN5-T仅在收获后上调, 然后在烘烤的前96小时期间不断降低 (图7)。

[0333] 为了检查微阵列分析中可能发生的任何交叉杂交, 特别是对于具有高度序列同源性的基因, 分析了使用RNAseq技术的ASN1-S/T和ASN5-S/T的表达。数据呈现在图8中。在瑞士白肋烟“Stella”中 (图8A), RNA-seq数据证实ASN5-T在烘烤期间比其它三个ASN1/5拷贝诱导的更少 (参见图7), ASN1-T是在前两天晾制后最集中表达的基因。这种ASN基因活性在烘箱中传统烘烤的瑞士亮色弗吉尼亚“ITB683”中不会出现类似的强度。在这种烟草中, ASN1-S/T和ASN5-S/T的表达在48小时烘烤后仍然很低。在两种烟草中, ASN3S/T表达不是通过烘烤诱导的 (图8C和D), 因此证实了图7中呈现的Affymetrix数据。

[0334] 为了测试ASN基因诱导在白肋烟烘烤早期阶段的稳固性, 在晾房中烘烤60小时之前和之后从同时在瑞士田间 (2013年) 生长的四个不同白肋烟种质 (BU1-4: Saplack、BanketA1、Stella和TN90) 分离叶RNA。分离RNA的时间点延长到60小时, 因为TN90在瑞士需要更多的时间开始变黄。表达数据 (图9A) 显示, 在所有白肋烟叶中, ASN1-T和ASN5-T同样分别是在晾制期间表达最多和最少的ASN基因 (除BU2中的ASN1-T外), 从而证实了图8中呈现的数据。转录物ASN1-S和ASN5-S在不同白肋烟种质叶的早期烘烤过程中不断并且牢固地表

达。图9B证实来自绒毛状烟草和美花烟草的两个ASN3拷贝在烘烤期间均未被诱导,由此间接证实这两个更接近于AtASN2的拷贝(参见图6)可能在植物营养组织中具有更多的基础功能用于维持天冬酰胺生理库。

[0335] 从田间开花期和田间生长的TN90植物中收集的8种组织中的相应转录物的分析表明,ASN5-T在花中比在叶中更活跃,ASN5-T RNA拷贝尤其在未成熟的花、花瓣和萼片中被发现(参见图10,左图)。除了在叶中表达之外,ASN1-T和ASN1-S也在可能同化硝酸盐的花瓣中强烈表达,这是已经在拟南芥中关于直系同源基因AtASN1描述的功能(Lam等人,2003Plant Physiol.2003年6月;132(2):926-35)(参见图6)。ASN3-S/T在所有这8种组织中都具有低的基础表达(参见图10,右图)。

[0336] 实例3-ASN1和ASN5RNAi TN90e3品系

[0337] 对白肋烟草植物中的ASN1-S/ASN1-T和ASN5-S/ASN5-T拷贝的沉默进行了调查,以确定其是否有助于减少白肋烟叶中的天冬酰胺。将特定的DNA片段克隆到强组成型MMV(紫茉莉花叶病毒)启动子或衰老启动子E4(天冬酰胺合成酶)之间。ASN基因片段侧接在MMV或E4与根癌土壤杆菌*Agrobacterium tumefaciens*胭脂碱合成酶基因的3'端终止子序列之间(Cheng等人.,2003In Vitro Cellular&Developmental Biology-Plant,39 595 604)。图22中显示了本研究中使用的构建体的图。白肋烟草品系TN90e4e5e10(Zyvert)使用标准土壤杆菌介导的转化方案分别用五种构建体中的每一种转化。TN90e4e5e10(Zyvert)代表甲磺酸乙酯(EMS)诱变白肋烟种群的选择,其在烟草的主要天冬酰胺合成酶基因CYP82E4、CYP82E5v2和CYP82E10中含有敲除突变(Lewis等人,2010)预防降烟碱产生。使用这样的背景品系可以避免解释未来TSNA数据时可能出现的复杂情况。

[0338] 为了能够选择高水平表达相应转基因的植物,通过qPCR分析每个构建体的至少10个独立的T0植物。从每个构建体的三个T0系收获种子,其显示出ASN沉默事件的最高水平。为了测试遗传了RNAi基因片段的T1后代是否执行有效的基因沉默,从4个选择的独立转化事件植物中提取RNA以获得用于将来田间测试(2016年)的T2种子。TN90e4e5e10(Zyvert)对照品系用作背景来比较转化的效率。

[0339] 根据ASN1和ASN5转录物的水平,在T0ASN1-RNAi和ASN5-RNAi植物中选择三个独立的转化事件。产生ASN1-RNAi植物以在MMV启动子或E4启动子的控制下使ASN1-T和ASN1-S拷贝沉默,并产生ASN5-RNAi植物以在MMV启动子或E4启动子的控制下使ASN5-T和ASN5-S拷贝沉默。对卡那霉素选择15个植物用于每个转化事件并在温室中生长。成熟时,从每个植物收集四个位于中下部的叶子,并进行7周的晾晒。

[0340] 氨基酸和尼古丁的浓度通过HPLC-MS确定:在50℃下将等份(~30mg)磨碎的烟草用乙醇-水(1:1;6mL)提取45分钟。将离心过的提取物稀释10倍。在45℃下在酰胺柱(Waters Acquity BEH amide,2.1×150mm,1.7μm)上用2mM甲酸铵水溶液+0.25%甲酸(洗脱液A)和乙腈+0.1%甲酸(洗脱液B)的梯度洗脱,应用洗脱梯度(0min-10%A,0.5min-10%A,4min-60%A,4.5min-60%A,4.6min-10%A,6.8min-10%A;流速0.5ml/min)进行液相色谱分离。对于质谱检测氨基酸,使用Q Exactive仪器(Thermo Scientific)以正电喷雾模式获取全扫描质谱。提取物中的尼古丁浓度由光电二极管UV/VIS检测器的260nm轨迹中的峰面积计算。

[0341] 天冬酰胺是在经烘烤的叶片中测量的。在黄化阶段烘烤3天后,提取叶片的冲孔,

从每个植物中分离RNA并进行qPCR。在图11A中,显示在MMV和E4启动子控制下的ASN1RNAi基因片段影响经烘烤的叶样品中天冬酰胺的水平,天冬酰胺的量在不同的转化事件中变化。与平行生长的背景植物TN90e4e5e10 (Zyvert) 相比,E284-4显示24%的天冬酰胺减少,而E284-6显示89%的天冬酰胺减少。对于启动子MMV或E4,天冬酰胺含量与烘烤3天后测量的ASN1RNA的水平相关(图11B),证实天冬酰胺是在烘烤期间合成的,ASN1-T和/或ASN1-S拷贝涉及在烘烤早期的天冬酰胺产生。组成型启动子(如MMV)和衰老启动子(如E4)都具有使ASN1沉默的活性。E4启动子的效率证明ASN1基因主要在衰老(变黄)期间有活性,从而有助于经烘烤的叶中的天冬酰胺产生,这可能由于从源叶到种子的氮的再分配。在烘烤三天后,与转录分析同时进行,还测量了天冬酰胺,已经显示6个测试品系的类似天冬酰胺还原曲线,虽然在此烘烤期后仅合成约三分之一的这一氨基酸(对照TN90e4e5e10 (Zyvert) 中的16.4mg/g,数据未显示),在叶收获后9-10天的晾晒后这一天冬酰胺的峰值是最大的(参见图2)。

[0342] 天冬酰胺合成酶从谷氨酰胺(Gln)转移氨基到天冬氨酸(Asp),谷氨酰胺是以从叶中储存的硝酸盐同化的氮形式固定氮的第一个氨基酸。由这一使用ATP的活性,它使得形成天冬酰胺(天冬酰胺)和谷氨酸(Glu,也参见图3)。因此,我们可以推测沉默ASN1将导致Gln和Asp的增加以及Glu的减少。

[0343] 因此,在相同的样品中还分析了Gln、Asp、Glu和尼古丁的量。在显示天冬酰胺强烈减少的样品,即E281-6和E282-1中可见Gln伴随增加,相反在显示低天冬酰胺的样品,即E282-7中Gln含量较低(图12A)。对于Gln和天冬酰胺,由ASN1基因的沉默导致的氨基酸平衡在经烘烤的叶内以相反的方式被修饰。对于Asp和Glu也观察到相似的效果,但是比Gln更不明显。在所有显示出相似Gln概况的转化品系(除了E282-7中的Asp外),Asp和Glu都显示增加($\sim 1.5-3\times$)。观察天冬酰胺合成酶活性(参见图3),预计Glu不会增加。

[0344] 尼古丁含量在MMV构建体和E4构建体中均不受ASN1沉默影响,证实ASN1-S和/或ASN1-T的活性与尼古丁合成途径无关,而限于在早衰程序期间的收获后叶成熟。

[0345] 使用RNAi转基因系进行类似的实验来使ASN5-S和ASN5-T沉默。数据显示,当ASN5RNAi片段在组成型启动子MMV的控制下表达时,天冬酰胺被有效地减少,特别是在两个品系中,但在E4启动子的控制下未减少或仅略微减少(图13A)。有趣的是,数据仅与烘烤3天后通过qPCR估计的ASN5mRNA的水平部分相关(图13B)。首先,在E4启动子控制下的转化体事件显示比在MMV控制下更充足的ASN5mRNA,证实E4启动子对于沉默ASN5而言不如MMV启动子有效,可能表明ASN5蛋白质合成程序比叶收获后ASN1蛋白质的产生更快和更早地发生(图7)。在某些品系中,天冬酰胺的百分比与mRNA水平不相关,比较例如E303-14与E303-16或E305-5与E305-15。这可能是由于ASN5-T拷贝参与了不存在或仅略微存在于经烘烤的叶(图9)中而是存在于开花组织(图10)内的活性。在烘烤3天后,还测量了天冬酰胺,对于ASN1转基因品系(参见上文),显示了6个品系(数据未显示)的类似天冬酰胺减少概况。

[0346] 在ASN5-RNAi品系中也测量了Gln、Asp、Glu和尼古丁的量(图14)。在ASN1-RNAi品系中,在显示天冬酰胺的强烈减少的样品(即E303-1和E303-16)中看到Gln的伴随增加,相反在显示更高天冬酰胺含量的样品(即E303-14和E-305-16)中的Gln量降低。与对照和姐妹系E-305-15和E305-16相比,E305-5显示小于10%的天冬酰胺减少和增加的Gln、Asp和Glu,表明在某些情况下,E4启动子可以部分地使ASN5拷贝沉默。与对照相比,品系E303-1显示出

最高的mRNA减少(图13B)、强的天冬酰胺减少(66%,图13A)以及Gln(14×)、Asp(3×)和Glu(3×)的相关增加(图14A、B和C)。因此,对于ASN1-RNAi植物,沉默ASN5导致天冬酰胺明显减少以及Gln、Asp和Glu增加。对于ASN1-RNAi品系,观察到ASN5-RNAi植物中的尼古丁水平没有变化(图14D),表明反应与尼古丁途径断开。

[0347] 在每个转化事件和对照品系中,称量中部秆位的叶。数据显示当ASN1-RNAi和ASN5-RNAi植物在受控环境如温室中生长时,叶物质没有大的变化(图15)。

[0348] 总之,ASN1和ASN5基因家族都在相应的MMV RNAi品系中沉默,从而表明ASN1和ASN5基因拷贝在白肋烟烘烤期间都参与了天冬酰胺的合成。

[0349] 通过交叉qPCR分析来确定ASN1和ASN5转录物家族是否在相应的MMV启动子控制下的ASN1-RNAi和ASN5-RNAi植物中被沉默。在表2中,数据显示在ASN5-RNAi植物中,ASN1和ASN5转录物都是沉默的,但在ASN1-RNAi植物中,特别是对于ASN1-1和ASN1-2,这还不太清楚。这一观测结果可能是由于ASN5和ASN1(S和T)基因非常相似,属于同一个基因簇(图6)。然而,这也可能表明,只要灭活ASN1基因簇以改变经烘烤的叶中的天冬酰胺合成就足够了。

[0350] 在白肋烟烘烤期间CYP82E4的表达分析表明E4启动子只有在烘烤48小时后才变得完全活跃,由此表明所有在E4启动子控制下的转基因构建体或基因片段在早期烘烤阶段较晚表达(>48小时)。因此,基于图11和图13中呈现的数据,预测ASN1-T和-S稍后在烘烤过程的第一阶段期间被激活,因为ASN1沉默可以由MMV和E4启动子同等地驱动。E4启动子微弱控制ASN5的沉默,从而表明ASN1基因由与诱导的叶衰老程序相关的一系列事件控制。

[0351] 实例4-ASN变体(EMS突变体)品系

[0352] 鉴定两个ASN1-S和-T拷贝以及两个ASN5-S和-T拷贝的ASN EMS突变体。在预测的外显子序列中的基因组DNA上筛选特定的突变。在ASN1-S(ASN1-S_W156*)、ASN1-T(ASN1-T_W156*)和ASN5-S(ASN5-S_Q66*)拷贝中鉴定出三个终止突变,但是在ASN5-T拷贝中未鉴定到。由于在ASN5-T中没有鉴定到终止突变,因此选择一个显示SIFT得分为0.0004的非终止突变作为可以替代终止密码子的潜在候选物(ASN5-T_G59D)。

[0353] 从M2亲本中,在选择基因(WT外分离体)中携带突变(变体)或未突变的后代种子被基因分型,鉴定和在田间种植。在各个地块,每个选择的变体的10株植物在其相应的WT外分离体的10株植物的一侧生长以防止田间变化的任何影响(即非均匀的硝酸盐肥和土壤变化)。在收获时收集四个中下部秆叶,并在晾房中挂在绳上2个月。叶子被剥离,叶片磨成细粉状。在每个叶库中测定天冬酰胺(mg/g)。尽管可以观察到一些变化,由于地块间的变化性,与相应的WT分离体相比,ASN1-S_W156*、ASN1-T_W156*和ASN5-S_Q66*中天冬酰胺减少的明显倾向是可见的,但是ASN5-T_G59D没有这种倾向(参见表3)。

[0354] 天冬酰胺减少的百分比显示在图17中。ASN1-T_W156*(83%)观察到最强减少,然后是ASN5-S_Q66*(44%)并且最后是ASN1-S_W156*(17%)。品系ASN5-T_G59D显示天冬酰胺略微增加。ASN变体中的天冬酰胺含量与在烘烤期间的基因活化水平相关。实际上,ASN1-T在烘烤期间比ASN5-S和ASN1-S在烘烤48小时和60小时后更多地表达(参见图8和图9)。然而,这一观察结果仍应谨慎使用,因为ASN1-T_W156*的数据仅来自田间的一个地块,而且ASN1-T_W156*和此品系的WT外分离体植物都在此田间实验中显示出强烈的生物物质减少(参见图18)。

[0355] ASN1-T_W156*和其相应的WT外分离体ASN1-T_WT*中的生物物质的减少可能是由两

种变体品系中仍然存在的其它突变造成的。实际上,用ASN1/5-RNAi品系进行的转基因实验倾向于表明ASN1-T和任何其它ASN1和ASN5拷贝对植物高度和生物质表型没有影响(参见图15)。与背景品种AA37杂交和随后的植物自交将允许我们总结关于ASN1-T终止突变对植物大小的任何影响。图18也显示ASN1-S和ASN5-S中的终止突变对植物生物质没有影响。这对于ASN5-T中的G59D点突变也是有效的,这有可能根据表达数据(图7和图8)优先涉及其它组织,如花(参见图10)。

[0356] 用ASN1-T(包括10株外分离体植物和10株终止突变植物的6个地块)、ASN1-S(包括10株外分离体植物和10株终止突变植物的6个地块)和ASN5-S(包括10株外分离体植物和10株终止突变植物的9个地块)的类似终止突变体进行第二个田间试验。在叶烘烤(中部秆位的6片叶)后,ASN1-T终止突变植物表现出与外分离体相比叶片中的天冬酰胺减少19%,ASN1-S终止突变植物表现出与外分离体相比叶片中的天冬酰胺减少23%,最后,ASN5-S终止突变植物表现出与外分离体相比叶片中的天冬酰胺减少约10%。这些结果证实了表3中所示的先前结果,证明选定的终止突变体能有效减少经烘烤的叶中的天冬酰胺。当与来自ASN-RNAi品系的天冬酰胺数据相比时,三种终止突变体ASN1-T_W156*、ASN1-S_W156*和ASN5-S_Q66*的组合应提供在经烘烤叶的叶片中类似的天冬酰胺减少。

[0357] 我们从这些数据得出结论,ASN1-S、ASN1-T和ASN5-S参与白肋烟叶晾晒过程中天冬酰胺的形成。

[0358] 实例5-ASN转基因品系中的天冬酰胺减少对丙烯酰胺形成的影响

[0359] 将来自转基因品系MMV-ASN1-RNAi(品系E281-6及其各自的对照,参见图11)和MMV-ASN5-RNAi(品系E303-1及其相应的对照,参见图13)的叶的叶片粉末与50%弗吉尼亚烟草粉末根据制造规格掺合制成浆料和铸叶材料。生产气溶胶形成制品,并对通过将其加热产生的气溶胶进行分析(加拿大卫生部)以测量丙烯酰胺。数据显示在图19中。与相应的烟草对照相比,ASN1-RNAi品系中表现出88%减少的天冬酰胺转化为70%丙烯酰胺的减少。用第二种转基因品系得到类似的数据,与烟草对照相比,在气溶胶中丙烯酰胺减少40%的ASN5-RNAi品系的叶中天冬酰胺表现出66%的减少。

[0360] 丙烯酰胺是吸烟者的生物标记(Naufal等人.,2011)。也分析了可燃香烟烟雾中丙烯酰胺的产生(图20)。使用以下程序从ASN1-RNAi植物的中部秆烘烤叶制备可燃香烟。烟叶在22℃和60%环境湿度下调理。切丝填料是在去除中脉后由叶片制备的。香烟是手工制作的,使用标准香烟管填充电动香烟注射机“Powermatic 2”。使插入香烟的切丝填料的重量变化(约800mg)以获得对应于120至140mm水位的抽吸阻力。然后用手控制每根香烟的坚实度(硬度)。一旦确定了一组切丝填料的参数,就将它们维持用于每个制成的香烟棒系列(每个样品200个)。

[0361] 在受控植物中,烟草切丝填料中天冬酰胺的含量(图20)与用于铸叶的掺合粉末中测量的天冬酰胺的含量(参见图19)相当,与图11中测量的量相比低约10%。这种差异是由于收获时的叶秆位,在白肋烟烘烤后,下部秆叶中天冬酰胺比上部秆叶中的天冬酰胺更丰富。然而,与野生型植物相比,ASN-RNAi中天冬酰胺的减少与图11和图19中所示的数据类似。有趣的是,可燃性香烟烟雾中丙烯酰胺的减少(57%)不如气溶胶生成制品气溶胶(70%)的有效,这可能表明导致形成天冬酰胺的美拉德反应在加热时比在燃烧的烟草产品更有效。

[0362] 由ASN1-S终止突变体的叶片制成的气溶胶形成制品的气溶胶中丙烯酰胺的减少也是有效的(图21)。然而,在这种情况下,这种减少是有限的,因为与例如ASN1-T相比,单独的ASN1-S作为烘烤叶中天冬酰胺合成的贡献者的影响是有限的(基于图17中所示的数据)。数据表明,在ASN-RNAi转基因品系和ASN突变品系中,经烘烤的叶片中天冬酰胺的减少伴随随后通过加热产生的可燃香烟烟雾和气溶胶中丙烯酰胺的减少。换句话说,敲除ASN1和ASN5基因拷贝导致烟草基质内天冬酰胺的减少和加热产生的可燃香烟烟雾和气溶胶中丙烯酰胺形成的减少。

[0363] 参考文献

[0364] Canales J,Rueda-López M,Craven-Bartle B,Avila C,Cánovas FM. (2012). Novel insights into regulation of asparagine synthetase in conifers. *Front Plant Sci.* 3:100

[0365] Lewis RS,Bowen SW,Keogh MR and Dewey RE. (2010) Three nicotine demethylase genes mediate nornicotine biosynthesis in *Nicotiana tabacum* L.: functional characterization of the CYP82E10 gene. *Phytochemistry* 71:17-18.

[0366] Moldoveanu SC (2005) Analysis of Protein Amino Acids in Tobacco Using Microwave Digestion of Plant Material. *Contributions to Tobacco Research* 21: 451-465

[0367] Naufal ZS,Marano KM,Kathman SJ,Wilson CL. (2011) Differential exposure biomarker levels among cigarette smokers and smokeless tobacco consumers in the National Health and Nutrition Examination Survey 1999-2008. *Biomarkers* 16: 222-235

[0368] Onoa H,Chudab Y,Ohnishi-Kameyama M,H.Yadaa,M.Ishizakac,H.Kobayashia& M.Yoshidaa. 2003. Analysis of acrylamide by LC-MS/MS and GC-MS in processed Japanese foods. *Food Additives&Contaminants*, 20:215-220.

[0369] Papousek et al., (2014) Determination of Acrylamide and Acrolein in Smoke from Tobacco and E-Cigarettes, *Chromatographia* 77:1145-1151

[0370] Sierro N,Batley JN,Ouadi S,Bakaher N,Bovet L,Willig A,Goepfert S, Peitsch MC and Ivanov NV. (2014) The tobacco genome sequence and its comparison with those of tomato and potato. *Nat Commun.* 5:3833.

[0371] 另一方面涉及天冬酰胺合成酶的表达或活性降低的突变型、非天然存在的或转基因植物或其部分,所述天冬酰胺合成酶包含以下、由以下组成或基本上由以下组成:(i) 包含以下、由以下组成或基本上由以下组成的聚核苷酸序列:与SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7具有至少72%序列一致性的序列;或(ii) 由(i)中所示的聚核苷酸编码的多肽;或(iii) 与SEQ ID NO:2或SEQ ID NO:4或SEQ ID NO:6或SEQ ID NO:8具有至少78%序列一致性的多肽;其中(i)、(ii)或(iii)中所示的天冬酰胺合成酶的表达或活性与对照植物相比降低。

[0372] 另一方面涉及一种制备植物材料的方法,源自所述植物材料的气溶胶中具有与来自对照植物的植物材料相比降低的天冬酰胺水平和降低的丙烯酰胺水平,所述方法包含以下步骤:(a) 提供包含如下聚核苷酸的植物或其部分,所述聚核苷酸包含以下、由以下组成

或基本上由以下组成：编码天冬酰胺合成酶并且与SEQ ID NO:1或SEQ ID NO:3或SEQ ID NO:5或SEQ ID NO:7具有至少72%序列一致性的序列；(b)降低所述植物或其部分中所述聚核苷酸的表达或由其编码的蛋白质的活性；(c)从所述植物或其部分收获植物材料；(d)干燥或烘烤所述植物材料；(e)任选地，测量所述植物或其部分中天冬酰胺的水平和/或测量源自所述植物或其部分的气溶胶中丙烯酰胺的水平；和(f)获得经烘烤或干燥的植物材料，源自所述植物材料的气溶胶中具有降低的天冬酰胺水平和降低的丙烯酰胺水平，合适地，其中尼古丁的水平与对照植物中尼古丁的水平基本上相同；合适地，其中谷氨酰胺、天冬氨酸和谷氨酸的形成与对照植物的经烘烤或干燥的叶相比增加。

[0373] 在本文中引用的或描述的任何出版物提供了在本申请的提交日期之前公开的有关信息。本文中的陈述不应解释为承认：发明人丧失先于这样的公开的资格。在上面的说明书中提及的所有出版物通过引用并入本文。在不脱离本发明的范围和精神的情况下，本发明的各种修改和变化对于本领域的技术人员将是显而易见的。尽管已结合特定优选实施例描述本发明，但应理解，如所要求的本发明不应不恰当地限于此类特定实施例。实际上，细胞生物学、分子生物学和植物生物学或有关领域的技术人员显而易见的用于实现本发明的所描述模式的不同改进意图在以下权利要求的范围内。

[0374]

表 1

在烟草、番茄和拟南芥天冬酰胺合成酶（ASN）基因产物之间的无缺口比对区域中的一致残基的百分比。

	AtASN1	AtASN2	AtASN3	NtASN5-T	NtASN5-S	NtASN1-T	NtASN1-S	NtASN3-T	NtASN3-S	Solyc06g007180	Solyc04g055200
AtASN1	100	76.91	77.26	83.01	83.73	85.1	85.27	76.17	76.17	83.9	77.44
AtASN2	76.91	100	91.33	77.22	76.99	78.03	78.03	84.64	84.64	76.82	85.46
AtASN3	77.26	91.33	100	77.4	76.99	78.37	78.37	85.14	85.14	76.99	85.96
NtASN5-T	83.01	77.22	77.4	100	97.88	90.27	89.91	77.16	77.16	93.81	80.2
NtASN5-S	83.73	76.99	76.99	97.88	100	90.85	90.51	77.28	77.28	95.08	79.7
NtASN1-T	85.1	78.03	78.37	90.27	90.85	100	98.31	77.62	77.62	90.34	80.45
NtASN1-S	85.27	78.03	78.37	89.91	90.51	98.31	100	77.45	77.45	89.83	80.2
NtASN3-T	76.17	84.64	85.14	77.16	77.28	77.62	77.45	100	100	75.93	64.73
NtASN3-S	76.17	84.64	85.14	77.16	77.28	77.62	77.45	100	100	75.93	64.73
Solyc06g007180	83.9	76.82	76.99	93.81	95.08	90.34	89.83	75.93	75.93	100	79.95
Solyc04g055200	77.44	85.46	85.96	80.2	79.7	80.45	80.2	64.73	64.73	79.95	100

[0375]

表2

[0376] qPCR分析(相对基因表达单位) 使用特异性引物ASN1p和ASN5p以确定ASN1-RNAi

(ASN1-1、ASN1-2和ASN1-3) 和ASN5-RNAi (ASN5-1、ASN5-2和ASN5-3) 转基因品系中ASN1和ASN5转录物的沉默程度。

[0377]

	ASN1p	ASN5p
对照	1	1
ASN1-1	0.35	0.88
ASN1-2	0.54	0.74
ASN1-3	0.10	0.40
ASN5-1	0.31	0.24
ASN5-2	0.50	0.34
ASN5-3	0.16	0.35

[0378] 表3

[0379] 在以下四种ASN变体中在中部/下部秆位收获的经晾晒的叶中天冬酰胺 (mg/g) 的分析:ASN1-S_W156*和其对应的WT外分离体ASN1-S_WT*、ASN1-T_W156*和其对应的WT外分离体ASN1-T_WT*、ASN5-S_Q66*和其WT外分离体ASN5-S_WT*、ASN5-T_G59D和其对应的WT外分离体ASN5-T_WT*。每个变体和WT外分离体在个别地块中并排种植,以防止任何不希望的田间效应(即硝酸盐和土壤变化)。

[0380]

地块	ASN1-S_WT#	ASN1-S_W156*	ASN1-T_WT#	ASN1-T_W156*	ASN5-S_WT#	ASN5-S_Q66*	ASN5-T_WT#	ASN5-T_G59D
1	24.2	20.5	27.5	4.7	20.9	12.0	25.0	26.4
2	26.0	20.6			22.3	16.2	18.5	27.9
3	22.9	20.5			22.0	15.2	17.7	28.0
4	25.7	21.3					28.3	28.7
5	30.9	22.0					25.4	31.9
6	17.4	14.1						
7	18.4	17.3						
%减少	17+/-7		83		34 +/- 8		否	

[0381] 序列

[0382] SEQ ID NO:1-NtASN1-S,基因组DNA序列

[0383]

atgtgcgggatcttggctgttttgggttggtctgatgattctcaggccaaaagggttcgtgttctcgagctctctcg
caggtaaattatatttcccttcttttctgtattgttaattgttattgtgttgcggtattatattgattctttttatga
gaaagtagtttctgaacgtttgcattttagaataaaaagaagactaaagcttaagatagattatatttttagatatt
cttgacgttgacctatatcttagattcacatataattttctcggaccaactgttttagcatgacttgtaaattgatg
ttgatatgaaaagctgtgatttaattagtaggagtaattcatttccatgttacaattacttatctttctcctccttg
tgttcttgccactgctagaggcagatctataattcctacttagagtattgaacacattgtacttttgaaattgtaag
tgtgcagtttaataatttgtaaattttgaaggtttttttaattacatgtatatcgtgacgttatatagcctctcta
ccgcataaggtaggggtaattgtctgcgtacaaataggetgatttcactctgtgttcagttgaatcaatatatgcgg
cateccacctgtttgcaaatgcatgggaatggaaaagtcacgtcactgtcacacatgatttttattttattt
atttattatatatgattcttttaaattctatttagtgggtactgaattcttgctaactgatcgattgatgaaacaggttga
agcatcgtggaccagattggagtggtgtatcaacatggggactgttacttggcacatcagcgtctagctattgtt

gatcctgcttccggtgatcaacctctgtttaacgaagataagacgattgttggttacggtgggtgcttttaatttcat
gtttactctcatttcttctgtcagtttactgatgaaaacgaaactaatgctgataactatcaaccaatcaaccacata
caagtaggggtcgactatatgattcattatatacctttcctctattcaaatagttgatgattaatttcatgccccaaa
tttcaggtaaatggagagatctacaatcacgagcaacttcgtaagcaaatgcctaatacagaattccggactggcag
tgactgtgatgtcattgcacacctagtgagtactcatcttccaacttaggaaacatgtatgtatatgtaaaattatt
ccctgattatattatgtacattaatcagtatgaagaacatggagaagattttgtggacatgctggatgggatccttcg
cttttgtgttattggatactcgagataacagctttcttgttgcgtgatgccattggaattacttccctttatatt
ggttggggacttgatggtaagattccatatgattttctcacttagattttgttattgcaaatgaatatgttgaatcc
ttatatgtactaattgttgaattctgttccctctcagggtctgtatggataatcatctgagcttaagggttggaatg
atgactgcgaacattttgaagttttccaccagggtcacttgtagcttagcaagaatggcggttttaggaggtggtac
aatcctccttgggttctctgagggcattccttccactcggttatgatcccttagttctcaggcgtgcctttgaaaatgt
gagttaatattgttgttgttgaatggatccaacatgttggttttaaggtatataatcattttgcaggctgttatca
aaaggttgatgactgatgtccctttgggtgttctctcctcgggggactcgattcattccttgggtgcttcgattact
gctcgtacttggctggtacaaaggtgccagcagtggggagcacagcttcattccttctgtgttggccttgagggt
cagataatcttgagacttgtgtgattacaacattaatacgacaacacctcattcacaagctagtaaggctcagctat
atgaaatctgataccgatttgccttgccttccaccgaacagggtcaccggatctcaaggctgcaagagaagttgctga
ctacttgggaaccgttcaccacgagtttacttccaggttcagggttttgatggagcaactctattttcacttgatct
atcactcctcttcgtttctcctttgtttacctgatggcaaatgacatgtgttgatttgcaggatggaattgatgcaa
ttgaagatgttattttaccatattgagacatacgatgtaacgacaatcagagcaagcactcctatgttcccttatgtcg
cgtaagattaagtcactaggagtgaagatggatcatactcggggaaggatctgatgaagtgttgggtggctacttgta
ctttcacaaggctcccaacaaggaagagttccacaaggaaacatgtcgcaaggtataagaaagtgttccagttcct
agtttaaataggacaaggagttgagaaaatgttgatttttccctatgcagattaaagcgttcaccaatatgactg
cttaagagcaataagtcacatctgcatgggtttagaagctagagtcctttccttagataaggagttcatcaatg
ttgccatgagtattgatccagagtgggaagttggtaagtagttgacccccaattgttgatcgaataaacaatcaactg
tccctaaaactatgtattttattcatttcagattaaaccagagcaaaggaggattgaaaagtgggctctaaggagggc
ctttgatgatgaggagcatccttatctcccaaagggtgtactaaatcgccattttcaaactgatttgcaaatgatttt
gaatgttatattattaatcaaaactaacaatttactcattacattccttgcagcacatcctgtataggcaaaaagaac
aattcagtgatggcgtgggctatagttggatagatggactcaaagcacatgctgaacaacatgtgcgtttcaatgac
taatcaacgactttctactgctcttattttatcttccaatgcttgtttgttttaaagctgaccgaggaaacaaat
ttccttttaataggtgaccaataggatgatgtttaatgcttcacatatattccctcacaacacaccattacaaagga
agcatactactataggatgattttcgagcgccttttccacaggttaattgttcagataatagtgagaatttatgat
aaatggactaattttacgtttgttgcgattttgcaaccttccaacagtgtagcaagctcacataatcggaattttt
tgttcaaataataatgcattggaaatcataatgctaaattgagattttacttttaacctgttccctctaccatttttg
gttctgcagaattcagctgggctaaccgttccctggaggagcaagtgtagcgtgttagcacagctaaagctgtagagt
ggatgcttcttgggtcaaagaaccttgatccttcaggcagggtgctatttggtgtacataactcggttatgagaatc
atgtacctgctatggctaattgggaatttgacaaaaaaatcattgggtcgtgtgccttctatggtagaagttgggtgct
gctcccgagctcacaataaagagttag

[0384] SEQ ID NO:2-NtASN1-S,推导的氨基酸序列

[0385] MCGILAVLGCSDDSQAKRVRVLELSRRLKHRGPDWSGLYQHGDYLAHQRLAIVDPASGDQPLFNEDKT

67

caattgaagatgttattttaccatattgagacatacgatgtaacgacaatcagagcaagcactcctatgttccctatg
tcgcgtaagattaagtcactaggagtgaagatgggtatatactggggaaggctctgatgaagtgtttgggtggctactt
gtactttcacaaggctcccaacaaggaagagttccacaaggaaacatgtcgcaaggtataaaaaagtgtttccagtt
ccttgtttaaaaccttattttaaataggataagcaacaacaacaacaaccaggtataatcccactagtggggctct
gaggagggtattgtgtatgcagaccttaccctaccctggggtagagaggctgtttcagatagaaactcgactccct
ccctccaaaaaccttattttaaataggacaaggagttgagaaaaatgttgatcttttccctccatctctatgcagatta
aagcacttcaccaatatgactgtttaagagcaaaataagtcaacatctgcatggggcttagaagctagagtgcccttc
ctagataaggagttcatcaatgttgccatgagtattgatccagagtgggaagttggtaagtagttgaaccccaattgt
tgatcgaataaacgatcaactgtccctaaaaccttgatttctatccatttcagattaaaccagagcaaaggaggatt
gagaagtgggctcctaaggagggcctttgatgatgaggagcatccctatctcccaaaggtggtactaaatcgcccttt
tcaaactgatttgcaagtgatttcgaatgttatattgttaaccaaactaacaatttnaaggctctgatgaagtgttt
gggtggctacttgtactttcacaaggctcccaacaagaagagttccacaaggaaacatgtcgcaaggtataaaaaag
tgtttccagttcccttgtttaaaaccttattttaaataggataagcaacaacaacaacaaccaggtataatcccaac
tagtggggctctgaggagggtattgtgtatgcagaccttaccctaccctggggtagagaggctgtttcagatagaaa
ctcgactccctccctccaaaaaccttattttaaataggacaaggagttgagaaaaatgttgatcttttccctccatctc
tatgcagattaaagcacttcaccaatatgactgtttaagagcaaaataagtcaacatctgcatggggcttagaagcta
gagtgcctttccctagataaggagttcatcaatgttgccatgagtattgatccagagtgggaagttggtaagtagttga
accccaattgttgatcgaataaacgatcaactgtccctaaaaccttgatttctatccatttcagattaaaccagagc
aaaggaggattgagaagtgggctcctaaggagggcctttgatgatgaggagcatccctatctcccaaaggtggtacta
aatcgcccttttcaaactgatttgcaagtgatttcgaatgttatattgttaaccaaactaacaatttactcattaca
tttccctgcagcacatccctatacaggcagaaagaacaattcagtgatggcgtaggctatagttggatagatggactca
aagcacatgctgaacaacatgtgcgtttcaatgactaatcaacaactttctaatgttcttatttatcttccaatgtt
tgtttgttttaagctgacttggaaattttttgtttcccttaaataggtgaccaataggatgatgcttaatgcttcac
atatattccctcataacacaccgattacaaaggaagcatactattataggatgattttcgagcgctttttccacag
gtaatttgttcatgtattgtgaggattcatgataaattagctaaacttatgctgggttttcaaccttgcaaccttccga
taagtcacattgttggagtttgttgttcaaatgcgagatttacttttaacatgttccctctgggtatttttggttctg
cagaattcagctgggctaaccgttccctggaggagcgagtggtggcgtgtagcacagctaaagctgtagagtgggatgc
ttcttgggtcaaagaaccttgatccctcaggaagggtgctattgggtgtacataactcagcttatgagaatcatgaac
ctgctatggctaattgggaatttggccacaaaaatcattggccgtgcgcgtctatggtagaagttgggtgctgctcat
gagctcacaataaggagtttag

[0388] SEQ ID NO:4-NtASN1-T, 推导的氨基酸序列

[0389] MCGILAVLGCSDDSQAKRVRVLELSRRLKHRGPDWSGLYQHGDYLAHQRLAIVDPASGDQPLFNEDKT
IVVTVNGEIYNHEQLRKQMPDHFRTGSDCDVIAHLYEEHGDFVMDLGI FAFVLLDTRDNSFLVARDAIGITSLY
IGWGLDGSVWISSELKGLNDDCEHFEVFPFGHLYSSKNGGFRRWYNPLWFSEAIPSTPYDPLVLRRAFENAVIKRLM
TDVPFGVLLSGGLDSSLVASITARYLAGTKAAKQWGAQLHSFCVGLGSPDLKAAREVADYLGTVHHEFHFTVQDGI
DAIEDVIYHIETYDVTITRASTPMFLMSRKIKSLGVKMVISGEGSDEVFGGYLYFHKAPNKEEFHKETCRKIKALHQ
YDCLRANKSTSAWGLEARVPFLDKEFINVAMSIDPEWKLIKPEQRRIEKWALRRAFDDEEHPYLPKHILYRQKEQFS
DVGYSWIDGLKAHAEEHVNTNRMLNASHIFPHNTPITKEAYYYRMIFERFFPQNSAGLTPGGASVACSTAKAVEW
DASWSKNLDPSGRAAIGVHNSAYENHEPAMANGNLATKIIGRAPSMVEVGAAHELTIRS

[0390] SEQ ID NO:5-NtASN5-S,基因组DNA序列

[0391]

atgtgtggaatcttggctttgttgggttgttcagatgattctcaggccaaaagggttcgagttcttgagctctctcg
caggcaatttccccctttcttttctcttttattttattactccatcatataattattctactttcacctttttgttt
tttaatttcgatttgtttgctttttgctttcttcatttaagtgtcttgcaattgctctagctcctatacgaataatag
gagctaaagtaatagagtgaaggattatgttgagcaacggtaaagtattttgtatagttcacaggttgagcg
taaaagcagttacttagtgtttgcattagggttaagtgtctatcaggaaagatatttgtagcaattgaccacccttg
gcaaaaaattatgcgtaaaaatattactacttgttatgtattcaaaaaatattgactttgaacatccttgtatagccta
gtggcaaagggtgttcaaaaattttgaacacctttattgaattttctggtttcgtcgttgcgtctacaacacactc
tgaagtgcggccttccctcaaaccctacgtgaatatggaacgcctagtgtgccgggatgcccttttaggctgagtaa
tagagtactatgttgatatacctaaaatacgggaactggcacgagagataaattcagaatttatagcttaattattta
aaatttatcatttgatttatttcaaccagagttcattctagtagagtttatgttgcatatttgtctgatttgcacca
tttgttttaagtgtcttgcattactctaattttattatgaataataagctaaactaattgagtttatgttgatgtaa
ctgacctgagttggcagtgatgtttggctacttgaatttaaaacaagattaagacttgtgagttggcagtaattgtt
tggtacttgaatttagaacaagaataaaacaagattaattaaggatacattattcctaaccggaccctaattctta
tagattgattattaaattttcttggaccaaccgtcacctgacctgttcattgatgggtgatggcaaaagccataata
cacttgatttctgttttcaatttttaagggttcttattattatacttattatatactttgaaaattatgggtcaaaat
tttaattttgttatttccctccgtatgtttgactcaacttgagttcatgaaaaaaaaaaaaatttttttaacctt
gtggtcttaacaagtcacgatataatgtgtggttgtgaatcgccctcgtaggataaagtagaaatttttaaggttaa
actatttcgaaataaaaaagtatgatagtgttttgggacatactaaaaaggaaataacacataaaatggaacaaag
aagtataatttttagtgatttttcacatatataatcagttatttgtatcaaaaatgctgggttcagttgaactcattgaac
tggtgtctccatccacctccgagttactgcaacatgtgaccggtggcagatgcaacccttacgtaccagggtcatctg
aacctagtataaaatataatttgtgaaaaatccctaaaatttcaataattagtagattttaaacccttaattttaaaact
acaatgagttcagttttaaaatctttatatgtcaaacctatcaagattaaatcctggatccaccttttgcattgtgac
agtacaatagccttttgtaattgaaacaaatatttgaattttggtaactatttgatcacataggttgaagcatcgtg
gaccagattggagtggtatatacaacatggtgatttttacttagcacatcaacgttttagcaattatcgatcctgct
tctggtgatcagcctctgttttaataagataagacgattgttgttacagtgagtgccctctaattttacctctttttt
tttcttttttttttttacgatcgcttgttttagattatttatggatgaaaatttgtatatttgattattgattgtgtg
ctaaattttgcagggtcaatggagagatttacaatcatgagaaacttcgtaatcttatgcctaatacagaagttcagaa
ctggaagtgattgtgatgttattgcacatctgtgagttagtttacttttctcattgaccaacttaggccagtgaa
tatattaggaatttcgggtgtttggatatggtcaaattagcttataagcacttattatcaattttaacatttttate
catacacgtaactgttcatcataaagtcatttttagcacttgatgtcatcaacctatgttgttgggactcttcgaa
aatgtcagcccgatagcatttcggagagtcggagcaacatagcaatcaactacttttaatcagttaaacgaacatgc
tcatagtataacctctggaaaatatgttatagtaaatagaaaaatttatacgaatgtctcactatcggtatgtatt
gatcagtatgaagaatatggagaaaattttgtggacatgttggatgggtgttctcttttgtattgttggatacgcg
cgataacagctttcttgcgtcgtgatgcaattggaattactcccctataatttggttggggacttgatggtaaga
ttttctatacaattttccagttagaaattaaaactgaagctattcctatttactattgaagaatcttgatacatgtt
atttgggtatttctgctgtgttattgtctcaggctctgtgtgatttcatctgagctaaaaggattaaatggtgact
gtgaacattttgaagttttccctcccggtcacttgtactcgagcaagaatggcggtttaggagatggtacaatcct

caatggttctctgaggetattccatcaaatecttacgaccccttagttttgagacgtgccttcgaaaatgtgaggct
tgtaaatgaatctagtgtaacaatcttgatgtttttctctcgactttggttttagcactcaaaaagtgtgtgtaca
aacgatcttggtataaggagtggtggttaattttgcaggctgttatcaaacgattgatgaccgatgtccctttgg
tgttctgctctccgggggacttgattcgctctttggttgcttctgtcactgctcgctacttggctggaacaaaagctg
ctaagcaatggggagcacagcttcattccttctgtgttggtctcgaggtcagactatcaatcctgagcggagacata
tttggggcggttctgtgggttcaactgaacccattgcttttggttcgaaccatatatacaaatccaagaaaaata
agatgtgtacatatataatgatgacactcattctgagcctactgacgctagatactggactcgcttttggtcttgagact
ttataatagaaaaatagatatattcaataccaatttctgtttttcttttctgactgacagggtcaccagatctcaaggc
tgcaagagaagttgtgactatttgggaaccgttcaccacgagttcaccttcacagttcagggttgtaaatcgtgtt
gctactccgatttattcatgtgagtccttctgttttctccttagaaagcaacaagtgttcatgttttgcaggatgg
aattgatgctattgaagatgttatttaccatatcgagacatacgatgtaacaacgatcagagcaagcactcctatgt
tcttatgtcgcgtaagattaaatcactgggagttaaagtgttcataatcaggggaaggctcagatgaactgtttggc
ggctatttgtacttccacaaggctccgaacaaggaagaattccatgtggagacatgtcacaaggtaataaaacacgt
cggtcccttggcaagctctgattctccttgtcattgtttgtctcggttggaataaattgctataattcgagaagaa
atgttaatectgctctttcactcgcatgcagataaaagcgcttcaccaatacgactgtttgagagcaaaataaggcaa
catcagcatggggcttagaagctagagtaccatttctggataaaagagttcatcaacgttgctatgagtatcgatcct
gaatggaagatggtaacaaaactgagctccaattgtcaatgagatgagaagcataactgaatatgcagttctctaa
aacttgtttttcatccgtttcagattaaacacgatcatggtaggatcgagaagtgggttcttaggaaggcttttgat
gatgaggagcaaccctatctcccaaagggtccacaaagtcattgttttcttactgatttgccgatttagtatggatcaa
aactaacatgttttgttcttttatgtactctgcagcatattctgtaccggcagaaagaacaattcagtgtatggcgta
ggctatagttggatcgatggactcaaagcacatgctgaacaacatgtgagcttctaaataaatccagaactattaca
tcgttatcgctccaattttcgtttgttttgcgtgcagctgatcttgaaacaattcttttctctcaggtgactgatag
gatgatgcttaatgtgcacatatcttccctcacaacactccaactacaaaggaagcatactattacaggatgattt
tcgagagggttcttccacaggtacttatagcgcgagactagaatcatatatctcttaaatggaaacatgaaaaataa
aggaggtgttttttctccttggaaaaagattttactactgattaaagcaccttcattcaccatcttgcacttccg
gtactttactaagttgtggttctaaaaaagaagatttagcttggtatattccatcacgaggaagccacaaaaagatt
ctacttataagaagttggatacttttaattggtgtgagaccttttgggaaaaaccatgcgggcttggcccaaagccga
caatatcacaccatgttaagagtatctttgcttcgtttttagccaaacagtacaataagtcaatttttaaacatgtt
gttactattgttggatttgcagaattcagcaaggctaactgttcttgaggaccgagtatagcttgcagcacagcta
aagctattgagtgggatgcttcgtggtcgaacaaccttgatccttccggtagggtgcaatcggtgtacataactcg
gcttatgacgatcatctcccgatgttggtaattgggaatttggacacaacgatcatcgataatgtgccgaggatggt
aggagtgggtgctgctgcagagctcacaataaggagctag

[0392] SEQ ID NO:6-NtASN5-s, 推导的氨基酸序列

[0393] MCGILALLGCSDDSAKRVRLVLELSRRLKHRGPDWSGIYQHGFYLAHQRLAIDPASGDQPLFNQDKT
IVTVNGEIIYNHEKLRNLMPNHKFRTGSDCDVIAHLYEEYGENFVMDLDGVFSFVLLDTRDNSFLAARDAIGITPLY
IGWGLDGSVWISSSELKGLNGDCEHFVFPFGHLYSSKNGGFRRWYNPQWFSEAI PSNPYDPLVLRRAFENAVIKRLM
TDVPFGVLLSGGLDSSLVASVTARYLAGTKAAKQWGAQLHSFCVGLGSPDLKAAREVADYLGTVHHEFTFTVQDGI
DAIEDVIYHIETYDVTTIRASTPMFLMSRKIKSLGVKMWISGEGSDELFGGYLYFHKAPNKEEFHVETCHKIKALHQ
YDCLRANKATSAWGLEARVPFLDKEFINVAMSIDPEWKMIKHDHGRIEKWVLRKAFDDEEQPYLPKHILYRQKEQFS

DGVGYSWIDGLKAHAEQHVTDMMMLNAAHIFPHNTPTTKEAYYYRMIFERFFPQNSARLTVPGGPSIACSTAKAIEW
DASWSNNLDPSGRAAIIGVHNSAYDDHLPDVGNGNLDTTIIDNVPRMVGVGAAAELTIRS

[0394] SEQ ID NO:7-NtASN5-T, 基因组DNA序列

[0395]

atgacgtagtttagcttgttcgtccacggcataaaaaacgcctataaattcggtgggtcattttgggttcatttacagaa
tcccacgatcatcttttgtgtattatcattttgccttctttacgtgttctctctcccaaaatcctgcaaattcttta
actatcttcccatatttctctccctctttctctttctttttgttatctactttactacttctatttagtagagcttta
tccaattaaagtcattctatatattggatatagttttactactactttatcttgtttgggaaagaaaagaataaaaaat
tagagtgtctatatattaagtataaaaaaggtgaaaaataaaaaagtagtgatcaatcatgtgtgggatcttggctttgt
tgggttgtccagatgattctcaagccaaaaggggtcgagttcttgagctctctcgcaggcaattccccctttcttttc
tcttttattttattactccatcatattattcaactataaatttcaccttttttgtttttcaattaataatttgtctgct
tttgcctttcttcatttaaatgttcttgcattgctctagttccctttatgaataataggctaaagtaatagagtgaagga
gtagagtcggagcaacgggtgaagctgtctttacatgatctatagctcacgggttcaccatagaagcagccaatggcg
aagctagaaaattcaataagggtgttctttcctttgccagtgggctatgcaagggtgttcaaagctctatttttaatc
aataacaagtaataattttatcttatagtgaaagcaaccactaatacttgcatttaggtaattagctgtctatcaagg
gaagtatatgtgtggttaagtagtggcggagccagaaatttcgttaaggggtgtcaaaatatataaaagtaaacatacc
aggaaattaaggggagtcatacatagtatatacatataaatttatttttttacctagctatacagtataatttttt
ccgcgaagtgacaccccttctataaggtggctccgccactgggtggtaaggggtgggtcaattgaacaccccttcgccga
aaaattatactctggatgggtgtaaaatattactacttgttactgattcaataaataaattttgaaaaccccttgcata
gtctagtagcaaagagtgttcaaaattttgaacacatttatgaattctctggtttcgccactgctgtctacatcac
atcccttaaggtgcgatcttttttcgaatcatgtgtgaatacgggatgccttgtgttccaggctgcccccttaggct
aaagtaatagagtacgtcttaattcttatgttgatcatataacctgaaactcggaactggcacgagagatataaattc
agaatttatagcaaatatatcattatgttgcatttgttttaagtgttcttgcattactctaatttttttatgaataat
aagctaaactaattgagtttatgttgatgtaactgacctgagtttggctacttgaaagttaaacaagattaagact
tgtgagttggcagcgattgtttggatacttgaacacgtttaaaacaagattaagacattatgatggatacattatc
ctaacctgggtgggacctaattctatggattcattattaaaatttcttggacaaaccatcacatgatctgttcatg
atgttttatggcaaaagccataatacactttatctgtatcacaaactactttatacacatgttatagagttttaaaa
ttattttaatatattgttatctttagtctcctgataattgactccgatgaaattcaagaaaaaaaaagaagatt
tttaaaatttatagtcgaaaacaagccatcgatatagtgtggttgttaataacctcgtaggcgtaaagtagaaat
tttaaagttaaactatcttaataagaaaatatgatatgtttttgcatagactaataaggaaagtatgacaagta
aaataaaacaaagaagtataattttaatgatttttccgtgtcaaaaatgtcagtttagttgaacttgaactcattga
aactgggtgaccagtgggtggatgcaaaccttaggtacggagggtcatccgaacctagtatagttatatatttgtgaaaaat
cactaaaatttcaataattactagatatgaaaccttaattttaaaactacaatgagtgcagttttaaaatctttaaa
tgtcaaacccgatcaagattaaattatggatccacctttaaagtcaaacccgacagtacaatagccttttgcaattga
aacaatatataacattacatttattaaactagttttaattttgggttactatttgattacataggttgaagcatcgtg
gaccagattggagtgggatataatcaacatgggtgatttttacttagcacatcaacgttagcaattatcgatccctact
tctggtgatcagcctctgttttaataagataagactattgttgttacagtgagtgccctctaatttaccctttttttt
gtggacgaatgcgttgttttagattactttaattatggctgaaaatttgtatatatttgattattgattgtgtgccaat
gttgcagggtcaatggagaaatttacaatcatgagaaacttcgtaatcttatgcctaatacacaagttcagaaccggaa

gtgattgtgatgttattgcacatcttgtgagttagtttactttactcatcaaccaacttaggctagtgaattcaaag
gcggagctagaattttgagtatatgagttctggaaattaggacaagataagttactgggttcggatagattatntag
acatattaagtagatttcttaacacaaatacacggcccgagccgaagctattgggttctgccgaaccgtagctaga
cttgtagctccgccattgagtgaaattatataaggaatttcgggtgtttgggttaaacttataagctgggtcaaattagc
ttataagcacttattatcaatttcaacatttctatccaaacacgtaactattcattcgatttttagcacttgatgctt
atcgactacttttaatacagttaaacgaacaggctcatagtgacaaactctagaaaatatgggtattgtaaattagtag
aagaatttatacgatgtctcactatcgatatgtattgatcagtatgaagaatatggagaaaattttgtggacatggt
ggatgggggtgttctcttttgtattgttggatacgcgcgataacagcttcttgcctgctcgatgcgattggaatta
ctccccctctatatattgggttggggacttgatggtaagattttctatacaattttccaagtagaaaattaaaactgaagcc
attccaatttgctattgtagaatcttgatatatgttatttgggtatttcaggctctgtgtggatttcactctgagttaa
agggtctaaatgatgattgtgaacattttgaagtttccctcccgggcacttgctactcgagcaagaatggcgggttt
aggagatgggtacaatcctcaatgggttctctgaggctgttccatcaaatccttatgaccccttagttctgaggcgctgc
cttcgaaaatgtgagaaatgaatctaggagggttccctctcgacttggacttagcactctaaaaatttgtctgtac
aaacgatcttggctataaggagtggttgttaatttcgcaggtgttattaaacggttgatgaccgatgtaccctttg
gtgttctgctctccgggggacttgattcgtctttgggtgcttctgtcactgctcgctacttggctggaacaaaagct
gctaagcaatggggagcgcagcttcattccttctgtgttggctcgcaggtcagtcacatcaatcttgagtagaaaaa
tttagggtcggttctgtgggttaacttaactcattgcttttggctcgaaaccatatatgcaaatccaaaagattta
agatgtatacatataataacactcattctgagcctactgacgctagatactggactgtgcctttggctcttgagac
ttttgtaatagaaaattgatattcaataccaatttctgttcttttgtgactaacagggtccaccagatctcaaggct
gcaagagaagttgctgactatttgggaaccgttcaccacgagttcaccttcacagttcaggtttgtaaatcgcgatg
ctacttcaatttattcatgtgaggctcttctgttttctactagaaaagcaacatgtgtgcatgttttggcaggacgg
aattgatgctattgaagatgttatttaccatatcgagacgtatgatgtaacaacgatcagagcaagcaccctatgt
tccttatgtcgcgtaagattaaatcacttggagtgaaagtgtcatatcaggggaaggctcagatgaactgtttgggt
ggctacttgtacttccacaaggctcccaacaaggaagaattccacacggagacatgtcacaaggtaataaaacacat
cggcaagtctgattctccttgtcattgtttgtctcgggttggataaaattgctataattcgagatatggaatgagtt
aagaaatgttaatcatgtttttccattcgcatgcagataaaaagcgttcaccaatacgactgtttgagagcaaaata
ggcaacatcagcatggggcttagaagctagagtaccatttctggataaagagttcatcaatattgctatgagtatcg
atcctgaatggaagatggtaagaaaactgagctccaattgtcaatgagacgagaagcataactgaatatgcgcgttc
tctaaaacctgtttttcatccgtttcagattaaacacgatcaaggtaggatcgagaagtgggttcttaggaaggctt
ttgatgatgaggagcaccctatctcccaaaggctccacaaactcataattttctactgatttggcaataattatgg
atcaaaaactaacatgtttgttctgttatgtgctctgcagcataatttgtaccggcgagaaagaacaattcagcgatgg
tgtaggctatagttggatcgatgggctcaaagcacatgtcgaacaacatgtgagcttcccttaataaatccagaacta
atgcagcgttatcgctccaatttttgtttgttttactatcaactaattttagtcatgttagtatgtttaccttaggta
ctatgtttttaaggagtttttagtctgtgtcaagaatttatcgttggtaaaaatagattgaacttccagtattttc
tacttttattatgcataatgattggactgagatgaggttaaatggaacaacatggttgaggattaataataaccaatc
caacttgggttggaaattgaggcgctcgtagttgttgttagatctcagaaacaattcttttctctcaggtgactgtagg
atgatgcttaatgcttcacatatcttccctcacaacactccaactacaaaggaagcatactattacaggatgatttt
tgagaggttctttccacaggtaatatattagtgagagattaaaatctatagattttgtttatggctgaggtttaaatca
tagatttcttaagctcaaaaactaaaaccatgtccctgctactgttggatttgcagaattcagcaaggctaactgtt

cttcccttgcataatagacttcacatttccagcgtttctgataagcggttcttgccttattcctgtccttctacttggtt
gattctcctctcacttagttcactatcttcttgcacgcattttcaaattgttttatcaattctcttgccttcttttc
atctgaccccaaaaaatacctcccaaaagcctttctgcttttccgagctgaaattatctgcttattactgccttct
aagtattctgatgcgttgactgttttttcccttgtagtaatgcatgggtcgtgtccaatctctattttgtttttt
gtgaaccactattctttgttgccttctgtttcgtcatcctcagttcccatcttctcagtttgattttcctcatttaa
aaagcggtgtgcccccttttcttgcctacccctctaaaaaatcagttgctgtcatctcgtttgatctgtcatgatccca
gttgatgccacatgttttagttctttacctctggacaacttttacttctgagaaccacgattttcttttccctgaagaa
attcttagaaattgggtatgttaaaggaaattcccagcaatcaccatttggcctctacctggccttgcttttcaactta
ccgacacgtgttcatecccccttttaacagatgctgaccattatcttctgcttgttggggcatacttcttacctatggg
gttaacctgatctttcgtatcacatttttcaactcttaattctcacatgtaaagtgcctatctttgcctcttttagt
tgatcttgggttgggctgttaatccaaattctctgggcttgccttgaagtatgataagcggtagctttactacagt
accatattctacattgctacagacgtgaaattatcctgatttgccttgccttgcctatcaacggttggcaatagtag
ccacagaggacctgattggagtggttgcatagccatgatgactgttatcttgcctcatcaacggttggcaatagtag
accaacttctggagatcagccactttacaacgaggacaagaccattgttgcggtatggaaatatttattatat
tttttcttagatatattaccggtacatggactttggacaatatgttgcgttttgatcttctatagtacaaactttct
cgtaggataaaatttctctactttgaatggatcttaacaggaaaaatcggaataaacaatttgcgttttgcgtttt
tagttcttacatatattggaacatatgcggtaaagtaactaggcgattgactaggtgaatacggagaatttctctac
tgttatcaatatattactcagctcatttgaaagatctctgggtcgtctagcagtacagctatttaggcaacttaacca
aatcccccttcccatgtgtttcactttcttgccttattttatatttgacattttgcaatgccgaaaaataaattgttt
ttccattttatgacctttaaaatgttgagggtgaagggttaggcgttactttaattgacttatcagcaacaatgtatt
agatacagagcacaaacttgtttgggtgcctaatttgggtatttttggagaacgatggagtaaatggaaatatgaatt
agcaaaatattgaaatagtagatctgaaaaatgaatgactgaatttagaggaggaccttttcttcttcttgcctga
agtgggtataagagaattatgtgaagaacatgtaagctttagggtaggttacaccataacagtcatttccctc
aactattagtgtagctgatgaatgccttcagaaatgtggctgttttttgcctgcgtattttatgtctttcatcatct
caaaattttatgtgttttaacactgcttgggaaggcctagagcttccgaaaatagtaggttgagaaaacatgagtaa
gcaaaatttcttctgcctggcaattaatcgttttatatgagggtgaagccacacaatctaacatgggtcttagacataaa
aatattatttttgacatagagcagacagaaatcttgggtctagttgctatctcccttccctaataagatgcctctat
cctcattgctatatacccaaacagattcacactagaatgaataaatctctccttgcagccacattttatattaataag
tagtgtaacaccaaacatttgccttctgcaaagctgttggaagtggacgattaagtaataaattgtgtgagtttg
actgaagagaaagtaataaattgggtgagcttctcacatttttcaggataatgtcaattcatctgattgaacttttt
gaaggacatctccttttctaaaagtttttgccttccctgaaaatgggttgccttcccaattttattctcctgtatt
ttccttgccttttttttttaccaggatgactctggccttcccttttgcaaatagtgacttcattaccttttaaca
aaagaccaaatagcctttgtatgcttgaaatcgaaagcagaatttacggctcaatttctctagagatgtgtgacagt
aacataaaacgtgggttaaagcaacttgtaagcactgccactgtccaatttgacatgactctattactatttagagtg
tcaaatagtggttcttttgactaccaattttcaatttctttttataaccttttcttgcctgtgacttataagta
tattttatgtagtttcagaatatatatatttagtttttaaagaattgatgtctgaaagtacagccaatttagatgggtt
caccctcgtacttttaaatcttataagttgtcttaggttaaaggaaataatgcatgattacattgtttggattgata
ttcagaagatcaaatgtgccggacaaaagcttcgagaagataaattctgtcattttgttgttttagataaaactacce
ttgtcaaataacctaaaactaattcacttttcaggatccttcttgattggactactctgatcttcttccatatatt

gatatacctattctttttgcttttaggttaacggggagatatataaccatagagacttacgggaaaagctgaagtccc
accagtttcgaactggcagtgactgtgaagttattgcacatcttgtaagtactttcagaatatatttagataactat
acatgacttctcagtttctgtgcgcagaatgatttcaaaataccatgactcaaattttcatcaaaagcgaaaaagg
tggtattgtagtaggagaatgagatTTTTGTcataatttctttcttatgggtacttgaactccgcaatctacctatcg
cagtatgaagactatggagaagacttcgttcacatgttggttggaatgttctcctttgtgcttcttgacacccgtga
taaaagtttcatgtctgcgcgggatgcaattggcattacacccctttatatgggatggggcttggatgggtatgtaatt
tgttccatttcataatttttggtatatgttccttttaataaatagcccgtgagctatatgaagttcctttgtttgttt
ccttgTTTTTatttgactgatgtatctttgtttgtttgcctactattaggggtgtcaatgggttcgggttcggccggt
tattttataaaatttgtaccataccaatttttaggttattctattatgtataaccaaaattagactttttgaaaccg
tcccaatcatgtcggtttctcttcgggtatcggtacgggttcgggttaatttttcgggtatttttaatatcatgtaaaatt
caccagtagaagtagaattgcaataacatacatctcttttataggacttagcaaaactctctagatatttttactgtt
taaagggtgatgaattaaaaaaaagaaaagaaagatggctagaatatagatccatcaactattatacaacagcgta
aaagaaatcaaacaaagacaaagaaaatatattaacacacgagttgaaagatataccaaggtaggactcaagaaataa
gtctatagaagattaaatatcaaaaagataaatctaaattatatgaaaggaaacataattcaattcattgtagtttg
ctactcataatcgctagaatactttgtgtcttgctaataaagatacttgaaataatttagtttaagtagaagtagca
taataggttttatgaattagtttttgagtttaattacttgttggttgtaataattttcataattccaaggctcaa
agaaaatttattgcattattatttttaaaacttactaaataaataataattttctacatgtaaaatttatcgggtatggt
tctgtattttttcggtttatttttataaaataaaaaacctaccctaattatcggtgcggttgtagatttatataaaa
acctacggtttcttaaaaagaaagctaaaaatcggttcgggtgcggtacgggttcgggtcggttttagtcgattttcgaat
atccattgacacccctacctactataatgctggctcatgctgttggttacaacaaaatgatgtataatagtgcatag
tacatgttttcttgccgctgtagatcaacaatgtaccaagaacatctgtcttagctagttcaacctaatcttcagtt
atctatcattccagtgaataacacatatagagttaggccgacattacctaaaggccaactatacaacattgtgagggtg
cgactaaggaactctgaacactgatcgggagaaaaagcaacttatccgtgtttgttttgaaagtactatgaggaggca
gcaaccttgctggatttatcacttgacagtaaatgtcaatggttgaggaaacatttttaatgaccggaggagaaacag
agaagagaagaaataattgattttaaaaaagcagaatagaaaattttgtgaagtcgtatgttaatttcaatgaagag
gatcaacagggtgcaatatttgacttggcacctagcatgtagaagtttctggttggggttgagcatggtaggcactgt
tgctttttacaactgtcattagagtcggaaacatatgtattcttgaggagcggttgctatggataaaaaaggcctat
ctactacttgatatattgggcattcagaactagttgtctaggaggaagaaaaggatttaaaatggtgagtattgttagg
ctttctttgccgtccaggatgcctgggctagttccctgaattgggatactggtggaggatttccctgaaattgatcat
ggccctgatatgttaggggataagaacattgaaggatcgaaggaaagtgccggcgctaagttttgctagaaatgaac
agttccattatggcttttgtcaaagatgggaatggaattggccccaacttctaggeatctccatccatcgtcagggtg
aactaaagtaaggttacactggtagaatttttcatatttcatattaaagtaagggtgaagccgtcaacaatcattgcg
tgagaaggaaccaggaacaaaatctagaagatgtctactaaagtgttaatagataatctgcaaaatttagcacggta
aggctagtgtagatctggttcaaatttttaaatgggtccaaaatgcgttaataaatgtcgggcctcgttttggteccc
cgaaagggttgctaaataaagacttgacaaagaaacagggtgaaggagaagcccaaacatttgaacaaactacacctag
cctaaacaaagtgtattgtttgcgaggaggaaaagcgagcatgcggagagagcctcaaacacaaaaaatggcctagtc
ttgtaatgagtcctagccaaatcatgtaagctccgactttcaagagagatcaacagaaaacttcgtcaatccggataa
atcaagaaaattatgataatggaaaggaggctggacattctactgccgacgcaaccctggttttggagtcagtgag
acaagtggaaactaaagaagaatatggatcaaaagagtttatccagttgacacctgatagactcactacatcagtttt

cagttcaagaacatttcaattcttgcctgataaagatagttgccgataaagataggcaaactgttacctaccaatcaa
aggaggttggtttcaattgagtttagaggtatcaaaggaggagccagtatatatagagcaagtttgatgaaggaggtt
gcctctatttgagaccaccaggtactcaagtgcctcccgtagttcctaggcattttatgatggatatgatagagaaa
tgtcttattgacaagaatatactagggctacagggataatagggttgggaaggtaggacttctttggttgctgtca
ggaacttggaacaaaggagaaaggggttggaatcaaaggttttgatttcgtttttgatttataaggaaataatgg
gtagtggatggagtcaggataggaaagaggtttaccagttgcaccattgaatatgtaattagagacactatcctct
tgttagggaacacgactatcacagcacgtaagggtagtgaaagctccgacatggacgtgggtttattcgggtatgt
tggcacgaagccttcgtgtctaggtttggtgttgggttaagtgcacctccttaggaaggcaaataaacttgggacag
atttgatagttgtcggaagttgcacaatgatgatattttctttatccaaacgctacgaatgattgcacatatagtaa
ctagagttatttctcatttgagcttttatatgtaagaatcaaaataaaatgattgatgaaatgcatgaaagctcaaat
gagaaataacttagttgtgagaatgcacgggggaaaaataattatattgattaaagtgttgtacaaccctatttaga
tacagtaattacataataataggtatctacttcccgatgtgggacactagacatgactaactacttaacaatcccc
tcaagccggtgcataataatcatatgtatcgagctgtttacagatgtaactaatacagaaccagtaagagacttag
tgaaaatatctgctagctaatcattcgactttacaaactttgtaacaatatctcttgagagtattttttctctaaca
aagtgcagtcgatctcaatgtgttttagtcctttcatagaacaccggatttgacgcaatatgaagagcagcttgatt
atcacacaccagttccatcttgcctgatttctccgaacttcaactccttgagcaactgcttgatccaaactagaatca
cacgttgccacaaccatggcccgatattcggtctcggcgctagatcgagcaactacattctgtttcttgccttcca
agagactaaattacctctactagaacacaatatccagacgtagaacgtcaatcagcaggtaatcctgcccatacag
catctgtgtacceacaatctgctcgtggcctcgatcctcaaatcgtaaccctttgcctggagctgactttatatac
cgaagaatgcgaacaactgcacccagtgactatcacaggagagaaaccataaactaacttgcaacactcacggaaa
agaaatattaggtctagtcactgtgtggtaatttaatttgccaaccaacctctctatctcgttaggatctctaagat
gctccccctgtccatgcagaagcttagcattcggtatccatagaagtgtcaactggctctacaacccatcattccagtc
tctcaagaatgtctaaagcatatttccgctgtgaaataacaatacctgagctagactgagcgacctcaatacc tag
aaaatacttcaatctgcccagatccttagtctggaagtgtgaaagagatgctgcttcagattagtaataccatctt
gatcattgccaataataataatcatcaacataccactcaactttggcatttctatctcaatttggctggaatatc
ctggacaatgcccataaacactgcgtctcttctcagtagctggaacaatggaggggcaaataattaacagaagaaat
ggtggaagtagttcctgcttgatctggtgggtcatgtggaaggaaagaatgcagagtttttgagaataggccta
gctcttagcctcttaccagaaaattaagcttaattgtttttgttatttcttttgggtgcaaagagagctataggg
aggatgtacattctttactagatttgctagaagaaatgtaatcgattagaataggcaaggcagttgcagatcttttg
gggtctgattgtggcttacactatctcctgtaaatatttttaatttttatataatttgttaccattctaaaaaaa
acataaaccaccagataaataacagattaggagcagaatgtcgataaagcacagagtgatccgcttactacgagt
catgccgaactcctgaataattatgtctgaacttaccaaaccaagctcgagggaactgtttcaaccataatagtgacc
tgcgcaatcggcatacaagaccactagacttcccttaagcaacaaaaccaggtgattgctccatataaatttcttcc
tcaagatcacctggaggaaagcattcttagtgcttaactgataaagaggtcattgacgtacaacaacctgaacaa
gaagagacaaccgatgctacttttagccacgggagagagtgtcactataatcaagcccaaaaatctgagtgtatcctt
ttgcaacaagacgagccttaagacgatcaacttgcccatccgggtgactttgactgcataaaccacagataacca
acaatagacttacttgaaggaagaggaacaagcttccaagtgcactcacatgtaagcagacatctcctcaatcat
agtttgcgccatcctggatgagatagtgccctcactgtagactttgggatagaacaattgacaaagatgatataa
aagcataatgaggtgatgcagacgatgataactaaatcgacataatgaggattaagattaagagtggatcataca

cccttccgaagtgcatacgggtgtacttaggaagagacaagtccgcagtaggagcagggtcaagtgtaggatgtgaatc
agttgggacctgatgctgggtgcggacgacgatgatatgtcaagagtgggtgttccgggtggcgggggaatc taggaggag
ccacactagactccccaacggttggaatgggtaagacttatgtggcgggaaggtgaaggaggagctatagtaagctct
ttaaaggctcggatataggtaaacctcagatatatacaaggtggtcagaagaggtaaagaaagggttagactcaaaaaa
tgtgacgtcagatgacataaagtaattacgaaaatcaagtgagtaacaacgatatcccttttgaacacgagaaataac
caaggaagacatacttgagagtacgaggagctaacttatcttttcttaggggctaaattatgaacgaaacaagagctc
cccaaaacacgaggaggaacagagtataaggggtgactgggaaacaatactgcatacggaaatctgattctggatggg
gatgaaggcattcgattaaacaaataacaagctgtgagaactgtatcgccataaaaaacgcaacggaacatgagattc
aaagagaagtgtgcgagcagtttcaatgatgtgccattctttctctctgcaacccccattttgctgaggggtaaaag
gacaagaggtctgatgaataattccttgagaagtcataaagtgcataaaattgagaggataactattctaaggcatta
tcactgcgaaaagtgcgaatagaaacaccaaattgatttttaatttcatacaaaaaattctggaatatagaaaacaa
ctcagaacgatctttcattaaagaaaatccaagtacatcttgatgatcatcaatgaaactatcaaac taacgaaatc
ccaaggttgagttgactctactatgacctatatacagaatgaaataaagaaaaaacagactatgcatgactctca
atactacgaggaaaggaggctcaggtatgtttcccgagctgacatgactcacactctaatgtagataaaactaggcac
tatcttctgaagcttggaatgtccataacatctgtgaattaggtccggaggatctgtaactagacatgccttgaggga
attgagtgggttaaggtagtaaaggccttctaattcaagtcctgttccaatcgctctgtccgtactgcggtcctgcat
aataaaagaatcattataaaaaatataaccacaatggagggcacaagtcaaacgactaacagatgcaagattaaaag
gacaaccagggaacataaagaacggaatctagagtgcagagggtaggggattcgcttggtccaactccttttgcttca
gtttgagaccggttggtctaaagtaacagtggaaagagactgtgaatacgcataatttgacaaaagtgatttattacc
agagatatgatcagaagcggttgagtccacaacccattgtccaagagtacttagactgggaaacacaagcaaaagaat
taccagcaacaaaagcatcagcttgagcaacagaggctacttggtggagatgtctgcttacttgctcgatactgaagg
aacttattataactccccctccgataaagaaaaataccggttacctatagctctcggtctgagcaacataagcattttt
gggtggacgaccatgtaaagaatagcacacgtcacgagtgtgtccaagtttatgacaataagagcacttgggcttag
atcttccaaaacgacctctctctgctctattctccatagtttgagatgcccgatgtttcacttactgggatacga
acagatgagtcaggctctgtgatgagctcactgggtgacttggtgctgcagcaaggcgaagtaatcgagagaataa
ttcatcaactgtggggacagtcggactagccaaaaatcggtcacgtactgaatcaaggctcattagggagtccatcaa
gtgtaagaactagaaacatcttctgtcgttgctcttggtgctttttaatactagcagaaactggcatcaacgtctca
aattcttccatgattgccgtgacttgctccctagtaagtagacatatccaattcctgtttcttcaagcttgctatc
gcgatattacatcatagaaacgagatatgtcatttagtgtataaaattacgagccttttcccaaacataaatacatgtc
tggaatggacggaacaagggtcatcaacttggaatcaatagatcgccacaggatactacataactgagcatcgacctt
cttcaaaagtgttttgcccttttcatcaccttcgctagccttttttggttaaattgatcttgaactccttgacctttac
aacacaactcgacagacgaagcccaagctaagtagtttgaacttctcattaaagggttctgaggtaatcataatagca
gaacttcagaactcgtgttttttagacccaaaatcacatccactcccaaagacattattggattgaaaagagatctagc
aaattagcaccaataaaacaaagaatcaactgtgggttgcccaaaaactgcccggaaaactactgtagttgcccggaaa
attttcaaagtgtctcggaatcaaaaaataaaaaatagggcaggctcggaatggtagagcgatcagactaacctaaag
gagtttttctgaaaaaattgacggaacgggctccatgcgctgggtgctgaggtagatctcgccggagaagactgtct
ccgatcggcgctggcggtgctgaggcggttatgacggagggtgttcgctgggggtttggctcgccggaagtggggg
accttggtggtggtgtttggtttttgcacaacaccgatggaattggttttgacgaaaaatagccctaaaagggtcacggg
aattgaagcacggcgacggctggttttttttccccggatgttttctcactgcccgtctgataccatgtgagaatg

cacgggagaaaaatatattatattgattaaagtgtgtgtacaaccctatttatatacagtaattacataataataggt
atctacttccccgatgtgggacactaaacatgactaactacttaacactagtaacctcatgggtctcatcctacttca
ttgaccacttggccacatccttgggtgcctgcactgtagttgtagctgtaatgaagatagtttttttctttcttg
atactcttatattgtgacaatttttctctatttgcctgtaattggatgaggttatgtgactctccggtggtgagaatg
ttaatgaacatcaagctcacatgtcatttctatttctgattgatcttgccttctactctttacgagtttaattgtgtt
agtccttccaagttctgtccatctatctttctgggtgtgcttaacccatgtatattggaccacactttgtattcatgc
tgattgcaacaaagcctgactgatcctgaactcctcctttatgccttaggtagaaattttgtgtgttttagtgttaca
tgtttcgcttcaaaggaaaaaggagaaaacttaaggtagaatcctgctccttgacctttggatccttgacattgaa
gtggacagacatgtttgctttgctttgatgagttgttttagtagtcggtctaaacaactgctgatgtattctgcttt
attgtaagtgtctataatttttatctttccacaatttgccttcagctctttctgcatgtagagaaaaataacagaagc
ttctatttgatcatgtaatatgtgactgcgcaagttcagttttaacaactttggctatgcataatgcaagcaatgacc
tgaaataaatattattttttatgcagggtctgtatggttttcttcagagatgaaagccttaagtgtgactgtgaac
gatttgttagtttcttccccgggtcacatttattcaagcaaaaaatggtagcccatcgaagctcatttctgagcttg
tgcttctatttattgttgatagttgatcatgtttcctatacgcacagctacttttgcctatttatttcagatttacttt
aggacctgtcggaacagctcttgaaacccttattgatgggtgtaattgttatctcttgatcaaccaaggagtaatt
acatgtaattattacaaatttgaataaaaataacttttaggataaccagtgagcttaataatcttttaggatatt
tgggatgctagagtgtatgtacttttgataatgtcggttaataataaaaatcctctgactttgtactgatgtactggctg
agcagattttatattactttctagtcctctgcaataaaaaaaattgtgagactgcaatataacaccattttttaat
gcatttacaatgttttaacttgatgtagcagtgagaacatgtcgaccttaaacagaggttctatttgtcctagata
aaggaacatcagcgtaatgaaacattacaggagagaagctctattgagttttgttaatctagaggagcccgtaaa
tagaacaccaattcatctgttatatgacttcttttctatttggaaacgttttaaaagtttgacaacattttttgatc
tccgtacactaaactttttgatgtattaaactaactatttccactactttaataattttcttcttaacaactcatg
ttagcatcttaactgtgtgtgttcaataccctcataaaaaattaaacgagtatttttttatatttgccttatgccct
tggagattgggatggggaagttcatcatttgagggtgaaaggtaaattcttttggatggacaaaaagtgaacctgc
tttggcatggagggactatttaaatcgatcatcattaccaagtcacctttcaaatgtttactttcaagaactattc
atgctatttctcaagtatattatgggtgcaggaggactcagaagatgggtacaaccaccatgggtactcagaaacat
tccttctactccatatgatcaccttgtcttacggaaagcttttgagaaggctattctattttggatgtcgcttgttt
ctgagttgtacataatatattaggtgccactttgttatattttgaacttcatcaaaagcctttggttttttgaggaca
catatcccccttgactgactgggcgtgaatttcttgatgtattgtgcttattttgattgtcctcttctttgtaata
ctcttgcctataaataaaccttctttatcaaaagaaaaatgataaaatgttaataaacactttgtgatataatgttgaag
gggagctcttggcataactggtaaagttgctgcatgtgaccaggaggtcacgggttcaagccgtggaaacagcctct
tgcagaaatgtagggtaaggtgcatacaatcgaccttgttggtccgaccttccccggaccccgcgcataacggga
gcttagtgacccgggtgtccttttttttgcatagactttacatgatattttatttcaatctttatcaattctttt
ttaccttcacatgcaattatactaccgtatgtgaaaatgcaatagtttttcccgcaattttcatctttcatgtac
ctctctctactatccttgtattactactgatttgggtgcattagatgttaattttatttcccaacctttactttatgca
ggctgtagtcaagcgacttatgacggatgtaccatttgggtgtgcttctctcaggtggactagattcttcaacttgttg
ctgcagtggccaaccgttatttggctgatactgaagctgcgcgacaatggggatcacagttgcataccttttgcgta
ggcttgaagggtggtgatgtcctttttgtaaacttgcctctgattaggaattgcttagctgtatcatccaattcaca
tggaagttctagaccattattgctgtttcttttccatatatagggttttctgaaaaatagtaaaaaggaaatggaat

79

agatgaaagtagcagagatgaggatgctgaggtggatgtgcgggcatacaaggatggacaagattaggaatgaagtt
attcgagagaaggtntttagccgaggggtctcctggaaacagcctctctgccccctcggggtaggggtaaggtctgcgt
acatactaccctccccagacccccattgggtgggattatactgggatgttgttgttgttgttgttgttgttgtacatat
aagggtattatatgtacacactcctccacacctttcacttataattacgtaaaacacatgcatgtgtaaaacctat
atatgtgtgtgtgatatataataaaaatgatgtttgaaaagacattatgatcaagtagttgatcttcatctgcttt
ccgttttttagcatatcttgtataggcagaaggaacagttcagtgtggagttggctacagttggattgatggtttga
aggatcacgcaagcagactgggtgagcttttttagttgttttctcttccctcctcttccattctttatcccttccttg
tggtttttgttgaacacatctcgtatatttggccaggtttctgattctatgttagcgaatgcaagtttcgtttaccc
gcataacacacccacgacaaaaggaaggatactactatagaactatttttgagcgatatttcccccaaggtgggtctta
gatagtcttcttttcttgttattcatgttcaatttttgatattttgtcccaacaatgttttcttttcttcccttc
tgttaacaaggctcgttctatcttcatatttgttgataaaggtaatcacattgataactattatttggagcagaatg
ctgcgagggaaacagttccaggtgggtccaagtgtggcatgcagcactgcaaaagcagtagaatgggacgcagcttgg
tccaagaatctagatccatctggacgagctgcactcgggtgttcatgcagctgcttatgaggatgcatcagaggttaa
aaccacagtcgcaacagataccccctcagaaacttgaagttgataaagctgcagtagctgtttga

[0400] SEQ ID NO:10-NtASN3-S, 推导的氨基酸序列

[0401] MCGILAVFGCIDNSQAKRSRI IELSRRLRHRGPDWSGLHSHDDCYLAHQRLA IVDPTSGDQPLYNEDKT
IVVAVNGEIYNHRDLREKLKSHQFRTGSDCEVIAHLYEDYGEDFVHMLDGMFSFVLLDTRDKSFIAARDAIGITPLY
MGWGLDGSVWFSSEMKA LSDDCERFVSFLPGHIYSSKNGGLRRWYNPPWYSETIPSTPYDHLVLRKA FEKAVVKRLM
TDVPFGVLLSGGLDSSLVA AVANRYLADTEAARQWGSQ LHTFCVGLKGSPDLKAAREVADYLGRHHEFHFTVQEG I
DALDEVIYHVETYDVTTIRASTPMFLMSRKIKSLGVK MVLSGEGSDEIFGGYLYFHKAPNKEEFHQETCRKIKALHL
YDCLRANKSTSAWGVEARVPFLDKEFVNVAMNMDPESKMIRPDLGRIEKWVLRNAFDDDQNPYLPKHILYRQKEQFS
DGVGYSWIDGLKDHASRLVSDSMLANASFVYPHNTPTTKEGYYYRTIFERYFPKNAARETVPGGPSVACSTAKAVEV
DAAWSKNLDPSGRAALGVHAAAYEDASEVKTTVATDTPQKLEVDKAAVAV

[0402] SEQ ID NO:11-NtASN3-T, 基因组DNA序列

atgtgtggaataactagcagttttcggttgcaattgataattctcaggccaagcgttccccgaatcatcgaactttctag
aaggtagaatttttttaaaaaaaaattgcattccaatttttctacagcattcttattctttggggtgacaatggta
tgcatgtgttttggattttgctattcgtggacttgctcgttttctcgtggctcttcgattttgtttttgttaatttataa
tcagtgtttttatatggattttctttttgaatctctcaaatatatgacatttgatttccacgagaaaaattggcatt
tttatggttcttatgatatgcggttaacccttcgccttgcatcgttcagcgttttaaaagacgaaatgatggctat
tagtaatactgctagagtggttattgaattacattaaataacagattgcttagtagatcaaattcagaatgacagtga
aatactcttgaaaacttgaatccccgttcaattcgctggtaaaacctggcaaaaaatagctttatctcaaaggaga
gtatgtcactgtgaaattgggtggatgtgttgtttaagctttaaggcatgatatgctgctcttaggtgatgtcagga
tgataataccatcctttcacactagactcagctttgaagcactttcatctcagttaatcgggtggaggaaaatctag
tttgcctttgggcataatataccaaatgtcttcttttgttttctagtgtttgttttctcaagaattgaaaaaccaacct
acaatatagtgatataattaactagttagctcttctgtttttctgaaaatttccaataattgagaagcgtttacag
gtcaattagtaaagggagaaacggaacatttttttctctaaagaacttctctttttccatacagaatgactttctat
acaaggcctaagctctgcttttagtgtctgcgattaagcatcttctgggaaacttcacacttgaaagaagtagaactag
gagggggcgaggtctagcaagataattgcttcagtcatttctcataaaaaaacattgtttcattcattatgtgcattc
atatagagttgggttaagctgctaagctgagctaatagtatgatagctatgaatctcgacaaagtaggttgcttctta

81

tccttgcagccacatatatattaataagtagtgtaacactaaaacatttgtgcttctgcaaagtcgttgaagtggg
cgattaagtaataaattgtgtgagtttgactgaagagaaagtaataaattgtgtgagcttctcacattttcatgata
atgtcaattcatctgattgaactttttgaaggacatctccttttctaaaagttttgttcacttcttgaaaatggtt
gctttcccaaatgtattctcctgtattttccttgtgctttatgttgaccaggatgactctggccttcccttttcaa
atgggtgacttcattaccttttaacaaaagatcaaatagcctttgtatgcttgaaatcgaagcagaatttagggctca
atttctctaggagatgtgtagcagtaacataacacgtgggttaaagcaacttttaagcactgccactgtcaaatgtga
tatgacgctattactatttagagtgtcaaatagtatttcttttgactaccaatttttcaatatctttttatacctt
tttctttgactgtgacttatagtatatatttatgtagttcagaatatatatatttagtttttaaagaattgatgtct
gaaagtacagccaattagatggtttcacccctgtacttcaaatcttataagttgttttaggttaaaggaataatatgc
atgattacattgtttggattgaaattcagaagatcaaatgtgcaggcaaaaagcttgcggaagatataattctgtca
ttttgtttgttttagattaacttccccgtcaaatagctaaaactaattcatactttctcaggatccttcttgattgga
ctactctgatcttcttccatatattgatatatcatattccttttgccttttaggttaacggggagatatataaccata
gagaattacgggaaaagctgaagtcaccagtttgcgaactggcagtgactgtgaagttattgcacatcttgaagt
actttcagaatatattcaaataaatatacatgacttctcagtttctgtgcgaagaatgatttcaaaaataccatgact
caaatttctcatcaaaaagagaaaaaggtggattgtagtaagagaatgagattttttgtcataatttctttcttatgg
tatttgaactccacgatctacctatcgcagtatgaagactttggagaagacttcgttcacatgttgatggaaatgtt
ctcctttgtgcttcttaacacccgtgataaaaagtttcattgctgctcggtatgcaattggcattacgcccccttata
tgggatggggctcttgatgggtatgtaattttccatttcataatttttggtatattgcccttttaataaatagcctgt
gagttatgtgaagttcctttgtttgtttccttgtttttatttgctgatgcatctttgtttgtttgcctactataat
gctggtttcatgctgttgttacaaacaaatgatgtataatagtgaaacagtacatgttttcttgccgctgtagatcaa
caatgtaccaagaacatctgtcttaagctagttcaacctaatcttcagttatctatcattccagcgaataacacatg
tagagttaaccggacattacctaaaggccagctatacaacattgtgaggtgtgactaaggaactctgaacactgtttg
gagaaatggaaagcaacatatccgtgtttgttttgaagtactataaggaggcagcaaccttgctggatttatcactt
gacagcaaatgtcaatggttgaggaacattttttatggctggaggagaaacagagaagagaaaaataattgatttt
aaaaaacagaggagaaaaattttgtgaagtcgtatgttaattcaatggctggaggtgcaatatattgacttggcaact
agcgtgtagaagtttctggttgggggttgacatggaagacactgttgcttttacaactgtcattagagtgcgaaaca
tattgtattcttgagggagcgttgcaatggataaaaaagtcctacctactacttgatattttgggcattcagaactgg
ttgtctaaagctgagaaggaagaaaaaggatttaaaatgggtgagtattgttaggttttcttttctgtccagggtgcctg
gactagttccctgaattgggatactgggtggaggatttctgaaattgatcatggccaaagagatatgtcaggggata
agaacattgaaggatcaaaggaaggtgctgcgctaagttttgctagaaatgaacagttccactatggctttttgtcaa
agatgggaattgaattggccccaaacttctgggcattccatccatcgtcaggtgaacaaaagtaaggttacaaatgggt
agaatttttctatttttcaacaagtaagggtgaagtcgtcagcaatcattgcctgagaaggaaccaggaacaaaatc
tagaagatgtctgctaaagtgctaataagataatctgccaaatttagcacggtaaggctagtgcagatctggttcaaa
ttttaaattgggtccaaaatgctttaataaatgtcggacctcgttttgggtcccccaaaaggttgctatataaagactt
gacaaagaaacaggtgaaggagaagcccaacaattgaacaaactacacctagcctaaagaaagtggttgatagttt
ggctcgaggagggaagagcgagcctgcggagagagcctcaaacataaaaaatgacctagctcttgtaatgagtctggcc
aaatcaagtaagctctgactttcaagagagatcaacagaaaacttcgtcagcttagataaatcaagaaaattatgat
aatgaaaaggaggctggacattctactgccgaccaacctgggttttgggtgctcagtgagacaagtggaactaaagaag
aatatggatcaaaaagagtttatccggttgacacctgataaactcactacatcagttttcagttcaaaaacatttcaa

83

84

taacataaatcctctgactttgtactgatgtactggctgagcagattttatattacttttctaateccctctacaaata
aaaaaattgtgagactgcaatataacaccatttttttaatgcatttacaatgtttaaacttgatttagcaacgtaga
acatgtcgaccttaaacagaggttctattttgtcctagataaaggaagatcagcataatgaaacattacaggagagga
gctctattgagtttttgttaatctagagggagcccgtaaatagtacaccaattcatctcttatatgacttcttttct
attttggaaacgttttaaaagtttgactacattttttgatctccgtacactaaactttttgatgtattaaactaacta
tttccactacttttaatattttcttcccttaacaactcctgttagcatcttaaacctctgtgtgttcaataccctcata
taaaattatacagagtatttttttatatttgccttatgtccttggagtttgggatggggaagctcatcatttgaggcg
aaaggtaaattcttttgaatggacaaaaagtgaaacctcggtctttggcatggagggactatttaagtcgccatcatt
accaaagtcatectttcaagtgtttgctttcaagaactattcatgtctatgtctcaagtatattatgtgtgcaggagg
actcagaagatggtacaaccacccatggtactcagaaaccattccttctactccatatgatcaccttgtcttacgga
aagcttttgagaaggctattctattttggaagtcgcttgtttctgagttgtatataatatttaggtgccactttgtt
atattttaaaacttcatcaaaagcctttggattctttgaggacacatatcccttgactgactgggagctgaatttcc
ttgatgtatagtgttattttgattgtcctctcctttgtaatactcttgcctataaataaaccttcccttatcaaaag
aaaatgataaaatgttaaaaatactttgtgataagtggtgactatagactttaaatgatatatttatctcaatcttta
tcaattcttttttaccttcacatgcaaattatactacagtatgtgaaatatgcaatagtttgtcctgcaattttcat
ctttcatgtacctctctctactatccgtgtattactactgatttgatgcattaaatgctaattttattcccaacctt
tactttatgcaggctgttagtcaagcgacttatgacggatgtaccatttgggtgtgcttctctcaggcgagtagattc
ttcacttgttgcctgcagtgcccaaccgctatttggctgatactgaagctgcgcgacaatggggatcacagttgcata
ccttttgcgtaggcttgaaggtggtgatgtcctttttgtaaacttgctctgatttgggaattgctagctgtatcat
cccaatttacatggaagttctagaccattattgctgtttcttttccatatatagggattttctgaaaaatagtaaaaa
ggaaatggaatttctctttattaaagataccgaatgcctttttgggaaatgatcttacagttctgaatcgaaaattt
ccttttgacattgtgattaattctgatatgaatttgggtccatcctggctttccttctcttgattgtatagattttcctt
ctcttgttacaattatatgcctgtctttttgtcttactctctgagacaggcaatgtaactccagttgcaaggtttgg
taatatgacttgtttattttctgttatcacatcctcctgaaaaagtgctactagcttgtattcatgcgggaatttga
aattttcaagtgtatagctttacaaaatgatcagaactgatcacagtcattgctctaaataaaattttctaaaaagc
aggtattgttgtgatataatgtgacacttaaaataaatttaaacctacatcctttcgaaatctgtaatagctccacta
atatttcttgtgtctaaactattccaacacctcttatcgatgccaattttgacagaccttagtgcacttaaaagtgtt
tagtgttctttatgcagtaggtatgttattgtccaaaatgacagaaataaagttagagatttgttgaaagccacttg
gtgtgattgagcagttcggagatgagtgatttatccattttatatacttgttagtgttttgttttcttttgcctttgt
cattgttaattaaggttgtcttttgtttgacattagggttctcctgatctgaaagctgccagagaggtggcagactac
cttggaaacctgcacatgagtttctacatttacagtgacaggttaactttccacaaagacatcctaccatgtgaccagga
ggtcacgggtttgagctgtggaaacaaccttttgcagaaatgcagcgtaaggttgcgtacaatagacccttgtggctc
cgcccttccccgaccccgacatagcgggagcttagtgaccgggtgcccacctaagctgattttccttatagg
agtataaccgcttaagtcaaaattggctatctacttctttctattgtctacatcaggaaggaattgatgcactagatga
agtcatttatcatgtttgaaacatatgatgtgaccactatcagagccagtcaccaatgtttctcatgtctcgggaaga
taaagcttttgggtgtgaaaatgggtctatctgggtgaaggttctgatgaaatttttggcggttatttatatttccac
aagggaaccaacaagaggagtttccaagaacactttagaaaaggtggatctcattgtcattttcttcagtcattca
ggattattggaagtaactgtatttttactagaatacttctttgaactattttctagctctgatggacgctttgatgtt
aatccagattaaagcacttcatctttatgattgcttgagggccacaaaatctacttcagcttggggtgttgaagctc

gtgtacctttcttggataaagaatttatcaatgttgcaatgaacattgatccagagtggaaaatggtaacctacagt
gcttccgtctcattttctccacccccaccccccaaaaaaagaacagaaaagaattgaaaagagagattcttttct
tccttcaaggaagaagggggaacaaaagaatagctgcaagttagtcagaaaaaatagctcttccatgtcatgggaga
ccagggttcaaactctgattgttaaaaaaacactagggtgaactcttctaactctgtttaagccttggtggacatagtt
aactagtatatgagctagtttagagctgactcgaaaccaccattattcaaaaaattaagaaagcttcaatgttgttc
caccgtgaattcatgatgtgtcgactttcatgtgtttatctctggatagatgaagagtttttttgatcttgattaa
ttctgttttgcccccttgaggttctgcacgctatttctagactcaaacttgctgcgacctatgttaggtctagagaa
tatctgcctaataaaacaatggaagttacattcaagtattttacctccatattacaagtcagtgtagttaaatggg
tctgatataatgggatagaaaatggcttcatacccataacctcgtaaataatgttgctgatgtgctgatctcatctgta
catctagtataatcctaattacaaattaatgcattcttttgcggtgactatagtatgctgaagaacctgtcttgaaa
cattgcagatcagacctgatctcggaagaatagaaaagtggttctacgcaatgcttttgacgatgatcagaatcct
tatctgcaaagggtttgtttgttttcttgatcagggtccaaataatcataatctttgtttattcatttgcataaaagg
cagtagactcccaatacattgtgtacgcttatgcacaaatgtatgagatgtgtattatatgtgtcacatataagggt
attatatgtacacactcctccaccttccacatatattatgtgtgtgtgtgtgtgtgtataatgaaatgatgtttga
aaagacattatgatcaagtagttgatcctcatctgctttctgttttttagcatactctgtataggcagaaggaacagt
tcagtgatggagttggctacagttggattgatggcttgaaggatcacgcaagcagctctggtgagctcttttagttgtt
ttctcttccctcctcttccattctttatcccttcttttgtggttttgttgaaacacatctcatatacttggtccagg
tttctgattctatgttagcgaatgcaagttttgtttacccgcataacacaccacgacaaaggaaggatactattat
agaactatttttgagcgatatttccccaagggtgggtcttacatagctcttcttttcttctgttattcatgttcagttta
tgatattttgtcccaacaatgttttcttttcttcttttctgttaacaagatcgcttctatcttctgtattttgtt
gataaagttaatcacattgttaactattatttggagcagaatgctgcgagggaacagttccagggtggtccaagtgt
ggcatgcagcactgcaaaagcagtagaatgggacgcagcttggccaagaatctagatccatctggtcgagctgcac
tcggtgttcatgcagctgcttatgaggatgcatcagagggttaaaaccacagtcgacgacagatactgctcagaacctt
gacgttgataaagctgcagtagctgtttga

[0403] SEQ ID NO:12-NtASN3-T,推导的氨基酸序列

[0404] MCGILAVFGCIDNSQAKRSRI IELSRRLRHRGPDWGLSHDDCYLAHQRLAIVDPTSGDQPLYNEDKT
IVVAVNGEIYNHRDLREKLKSHQFRTGSDCEVIAHLYEDYGEDFVHMLDGMFSFVLLDTRDKSFIAARDAIGITPLY
MGWGLDGSVWFSSEMKA LSDDCERFVSFLPGHIYSSKNGGLRRWYNPPWYSETIPSTPYDHLVLRKAFKAVVKRLM
TDVPFGVLLSGGLDSSLVA AVANRYLADTEAARQWGSQ LHTFCVGLKGS PDLKAAREVADYLGTRHHEFHFTVQEGI
DALDEVIYHVETVDVTTIRASTPMFLMSRKIKSLGVKMVLSGEGSDEIFGGYLYFHKAPNKEEFHQETCRKIKALHL
YDCLRANKSTSAWGVEARVPFLDKEFVNVAMNMDPESKMIRPDLGRIEKWVLRNADFDDQNPYLPKHILYRQKEQFS
DGVGYSWIDGLKDHASRLVSDSMLANASFVYPHNTPTTKEGYYYRTIFERYFPKNAARETVPGGPSVACSTAKAVEW
DAAWSKNLDPSGRAALGVHAAAYEDASEVKTTVATDTPQKLEVDKAAVAV

[0405] SEQ ID NO:13-NtASN1 引物ASN1-f的DNA序列

[0406] tcacatatattccctcataacacac

[0407] SEQ ID NO:14-NtASN1 引物ASN1-r的DNA序列

[0408] aagaagcatccactctacagctt

[0409] SEQ ID NO:15-NtASN5 引物ASN5-f的DNA序列

[0410] atatcttccctcacaacactccasct

[0411] SEQ ID NO:16-NtASN5引物ASN5-r的DNA序列

[0412] Ctaccggaaggatcaaggttgt

[0413] SEQ ID NO:17-来自NtASN1-S用于使NtASN1-S和NtASN1-T拷贝沉默的DNA编码序列

[0414]

tcacatatattccctcataacacacccattacaaaggaagcatactactataggtgattttcgagcgctttttccc
acagaattcagctgggctaaccgttcctggaggagcaagtgtggcgtgtagcacagctaaagctgtagagtgggatg
cttcttgggtcaaagaaccttgatccttcaggcagggtgctattgggtgtacataactcggttatgagaatcatgta
cctgctatggctaattgggaatttgacaaaaaaatcattggtcgtgtgccttctatggtagaagttgggtgctgctcc
cgagctcacaataaagagtttag

[0415] SEQ ID NO:18-来自NtASN5-T用于使NtASN5-S和NtASN5-T拷贝沉默的DNA编码序列

[0416]

gaccagattggagtgggatatatcaacatggtgatttttacttagcacatcaacgttttagcaattatcgatcctact
tctggtgatcagcctctgtttaatcaagataagactattgttgttacagtcaatggagaaatttacaatcatgagaa
acttcgtaatcttatgcctaatacacaagttcagaaccggaagtgttgtgtgttattgcacatctttatgaagaat
atggagaaaaattttgtggacatgttggatgggggttctcttttgtattgttggatacgcgcgataacagctttctt
gctgctcgtgatgcgatttgaattactccccctctatatattggttggggacttgatgg

[0417] SEQ ID NO:19-ASN1-S终止突变体的DNA序列:ASN1-S_W156 *

[0418]

atgtgcgggatcttggctgttttgggttggtctgatgattctcaggccaaaagggttcgtgttctcgagctctctcg
caggtaggttgaagcatcgtggaccagattggagtgggctgtatcaacatggggactgttacttggcacatcagcgt
ctagctattgttgatcctgcttccggtgatcaacctctgtttaacgaagataagacgattgttgttacggtaggtaa
atggagagatctacaatcacgagcaacttcgtaagcaaatgccaatcataagtccggactggcagtgactgtgat
gtcattgcacacctagtagtatgaagaacatggagaagattttgtggacatgctggatgggatcttcgcttttgtgt
tattggatactcgagataacagctttcttgttgcctgctgatgccattggaattacttccctttatatattggttgggga
cttgatggtagggctctgtatgaatatcatctgagcttaagggttgatgatgactgcgaacattttgaagttttcc
caccagggcacttgtactctagcaagaatggcggctttaggaggtggtacaatcctccttgggttctctgaggccatt
ccttccactcgttatgateccttagttctcaggcgtgcctttgaaaatgtaggctgttatcaaaaggttgatgactg
atgtcccccttgggtgttctcctctccgggggactcgattcatccttgggttgccttcgattactgctcgtacttggct
ggtacaaaggctgccaagcagtggggagcacagcttcattccttctgtgttggccttgaggtagggtcaccggatc
tcaaggctgcaagagaagttgctgactacttgggaaccgttcaccacgagtttacttcaccgttcaggtaggttg
aattgatgcaattgaagatgttatttaccatattgagacatacgatgtaacgacaatcagagcaagcactcctatgt
tccttatgtcgcgtaagattaagtcactaggagtgaagatggtcatactggggaaggatctgatgaagtgtttggt
ggctacttgtactttcacaaggctcccaacaagggaagagttccacaaggaaacatgtcgcaaggtagattaaagcgc
ttaccaatatgactgcttaagagcaaataagtcacatctgcatggggtttagaagctagagtccttttccctagat
aaggagttcatcaatgttggcatgagtattgatccagagtgggaagttgagattaaaccagagcaaaggaggattgaa
aagtgggctctaaggagggttggatgatgaggagcatccttatctcccaaaggtagcacatcctgtataggcaaaa
aagaacaattcagtgatggcgtgggcatagttggatagatggactcaaagcacatgctgaacaacatgtaggtgacc

aataggatgatgtttaatgcttcacatatattccctcataacacacccattacaaaggaagcatactactataggat
gatttttcgagcgctttttccacaggtagaattcagctgggctaaccgttcctggaggagcaagtgtggcgtgtacg
acagctaaagctgtagagtgggatgcttcttggtaaagaaccttgatccttcaggcagggctgctatttggtgtaca
taactcggcttatgagaatcatgtacctgctatggctaattgggaatttgacaaaaaaatcattggctcgtgtgcctt
ctatggtagaagttgggtgctgctcccgagctcacaataaagagt

[0419] SEQ ID NO:20-ASN1-S终止突变体的氨基酸序列:ASN1-S_W156*

[0420]

MetCysGlyIleLeuAlaValLeuGlyCysSerAspAspSerGlnAlaLysArgValArgValLeuGluLeuSerAr
gArgLeuLysHisArgGlyProAspTrpSerGlyLeuTyrGlnHisGlyAspCysTyrLeuAlaHisGlnArgLeuA
laIleValAspProAlaSerGlyAspGlnProLeuPheAsnGluAspLysThrIleValValThrValAsnGlyGlu
IleTyrAsnHisGluGlnLeuArgLysGlnMetProAsnHisLysPheArgThrGlySerAspCysAspValIleAl
aHisLeuTyrGluGluHisGlyGluAspPheValAspMetLeuAspGlyIlePheAlaPheValLeuLeuAspThrA
rgAspAsnSerPheLeuValAlaArgAspAlaIleGlyIleThrSerLeuTyrIleGlyTrpGlyLeuAspGlySer
ValSTOP

[0421] SEQ ID NO:21-ASN1-T终止突变体的DNA序列:ASN1-T_W156*

[0422]

atgtgcgggatcttggctgttttgggttggtctgatgattctcaggccaaaagggttcgtgttctcgagctctctcg
caggtaggttgaagcatcgtaggaccagattggagtgggctgtatcaacatggggactgttacttggcacatcagcgt
ctagctattgttgatcctgcttccgggtgatcaacctctgtttaacgaagataagacgattgttgttacggtaggttaa
atggagagatctacaatcacgagcaacttcgcaagcaaatgcctgatcataagtccggactggaagtgactgtgat
gtcattgcacacctagtagtatgaagaacatggagaagattttgtggacatgctggatgggatcttcgcttttgtgt
tactggatactcgagataacagcttttcttgttgcctgctgatgccattggaattacttcccttttatattggttgggga
cttgatggtagggctctgtatgaatatcatctgagcttaagggttgaaatgatgactgcgaacattttgaagttttcc
caccaggacacttgtagcttagcaagaatggcggtttaggaggtggtacaatcctcttttggttctctgaggetatt
ccttccactccttatgatcccttagttctcaggcgcgctttgaaaatgtaggctgttatcaaaagggttgatgactg
atgtcccttttgggtgttctgctctccgggggactcgattcatccttggttgcttcgattactgcccgtacttggct
ggcacaagggtgccaagcagtgaggagcacagcttcattccttctgtgttgcccttgaggtagggatcacggatc
tcaaggctgcaagagaagttgctgactacttgggaaccgttcaccacgagtttcacttcaccgttcaggtaggatgg
aattgatgcaattgaagatgttatttaccatattgagacatacgaatgaacgacaatcagagcaagcactcctatgt
tccttatgtcgcgtaagattaagtcactaggagtgaagatgggttatcttggggaaggctctgatgaagtgtttggt
ggctacttgtactttcacaaggctcccaacaaggaagagttccacaaggaaacatgtcgaaggtagattaaagcac
ttaccaatatgactgtttaagagcaaataagtcacatctgcatggggcttagaagctagagtgcctttccttagat
aaggagttcatcaatgttgccatgagtattgatccagagtgggaagttggttagattaaaccagagcaaaggaggattg
agaagtgggctctaaggagggcctttgatgatgaggagcatccctatctcccaaaggtagcacatcctatacaggca
gaaagaacaattcagtgatggcgtaggctatagttggatagatggactcaaagcacatgctgaacaacatgtaggtg
accaataggatgatgcttaatgcttcacatatattccctcataacacaccgattacaaaggaagcatactattatag
gatgattttcgagcgctttttccacaggtagaattcagctgggctaaccgttcctggaggagcgagtggtggcgtgt
agcacagctaaagctgtagagtgggatgcttcttggtaaagaaccttgatccttcaggaagggtgctatttggtgt
acataactcagcttatgagaatcatgaacctgctatggctaattgggaatttgccacaaaaatcattggccgtgcgc

cgtctatggtagaagttggtgctgctcatgagctcacaataaggagt

[0423] SEQ ID NO:22-ASN1-T终止突变体的氨基酸序列:ASN1-T_W156 *

[0424]

MetCysGlyIleLeuAlaValLeuGlyCysSerAspAspSerGlnAlaLysArgValArgValLeuGluLeuSerArgArgLeuLysHisArgGlyProAspTrpSerGlyLeuTyrGlnHisGlyAspCysTyrLeuAlaHisGlnArgLeuAlaIleValAspProAlaSerGlyAspGlnProLeuPheAsnGluAspLysThrIleValValThrValAsnGlyGluIleTyrAsnHisGluGlnLeuArgLysGlnMetProAspHisLysPheArgThrGlySerAspCysAspValIleAlaHisLeuTyrGluGluHisGlyGluAspPheValAspMetLeuAspGlyIlePheAlaPheValLeuLeuAspThrArgAspAsnSerPheLeuValAlaArgAspAlaIleGlyIleThrSerLeuTyrIleGlyTrpGlyLeuAspGlySerValSTOP

[0425] SEQ ID NO:23-ASN5-S终止突变体的DNA序列:ASN5-S_Q66 *

[0426]

atgtgtggaatcttggctttgttgggttgttcagatgattctcaggccaaaagggttcagattcttgagctctctcg
caggcaggttgaagcatcgtaggaccagattggagtgggataatacaacatggtgatttttacttagcacatcaacgt
ttagcaattatcgatcctgcttctggatgacgctctgtttaattaaagataagacgattgttgttacagtaggtca
atggagagatttacaatcatgagaaacttcgtaatcttatgcctaatacacaagttcagaactggaagtgtgtgat
gttattgcacatctttagtagtatgaagaatatggagaaaaatttgtggacatgttggatgggggtgttctcttttgtat
tgtttggatacgcgcgataacagctttcttctgctgctcgatgcaattggaattactcccctatatattggttgggga
cttgatggtaggctctgtgtggatttcatctgagctaaaaggattaaatggtgactgtgaacattttgaagttttcc
ctcccggtcacttgtactcgagcaagaatggcgggttaggagatggtacaatcctcaatggttctctgaggctatt
ccatcaaatccttacgaccccttagttttgagacgtgccttcgaaaatgtaggctgttatcaaacgattgatgaccg
atgtcccccttgggtgttctgctctccgggggacttgattcgtctttggttgcctctgtcactgctcgtacttggct
ggaacaaaagctgctaagcaatggggagcacagcttcattccttctgtgttggctctcgaggtagggtccaccagatc
taaggtgcaagagaagttgctgactatttgggaaccgttcaccacgagttcaccttcacagttcaggtaggtatgg
aattgatgctattgaagatgttatttaccatatcgagacatacgatgtaacaacgatcagagcaagcactcctatgt
tccttatgtcgcgtaagattaaatcactgggagttaaagtgtcatatcagggggaaggctcagatgaactgtttggc
ggctatttgtacttccacaaggtccgaacaaggaagaattccatgtggagacatgtcacaaggtagataaaaagcgc
ttaccaatacagactgtttgagagcaaataaggcaacatcagcatggggcttagaagctagagtaccatttctggat
aaagagttcatcaacgttgctatgagtatcgatcctgaatggaagatggtagattaaacacgatcatggtaggatcg
agaagtgggttcttaggaaggcttttgatgatgaggagcaaccctatctcccaaaggtagcatattctgtaccggca
gaaagaacaattcagtgatggcgtaggctatagttggatcgatggactcaaagcacatgctgaacaacatgtagggtg
actgataggatgatgcttaatgctgcacatatcttccctcacaacactccaactacaaaggaagcatactattacag
gatgattttcgagaggttcttcccacaggtagaattcagcaaggctaactgttcttgaggaccgagtatagcttgc
agcacagctaaagctattgagtgggatgcttcgtgggtcgaacaaccttgatccttccggtagggctgcaatcggtgt
acataactcggttatgacgatcatctccccgatgttggttaattgggaatttggacacaacgatcatcgataatgtgc
cgaggatggttaggagtgggtgctgctgcagagctcacaataaggagc

[0427] SEQ ID NO:24-ASN5-S终止突变体的氨基酸序列:ASN5-S_Q66 *

[0428]

MetCysGlyIleLeuAlaLeuLeuGlyCysSerAspAspSerGlnAlaLysArgValArgValLeuGluLeuSerAr

gArgLeuLysHisArgGlyProAspTrpSerGlyIleTyrGlnHisGlyAspPheTyrLeuAlaHisGlnArgLeuA
laIleIleAspProAlaSerGlyAspGlnProLeuPheAsnSTOP

序列表

<110> Philip Morris Products S.A.

<120> 天冬酰胺含量减少的植物

<130> P10251W0

<140>

<141> 2015-09-09

<160> 24

<170> PatentIn version 3.5

<210> 1

<211> 3646

<212> DNA

<213> 烟草

<400> 1

```
atgtgcggga tcttggctgt tttgggttgt tctgatgatt ctcaggccaa aagggttcgt 60
gttctcgagc tctctcgagc gtaaattata tttcccttct tttctgtatt gttaatgtta 120
ttgtgttgcg ttattatatt gattcttttt atgagaaagt agtttctga acgtttgcat 180
tttagaataa aagaagacta aagcttaaga tagattatat atttagatat tcttgacgtg 240
gacctatata ttagattcac aatataattt tctcggacca actgttttag catgacttgt 300
aaatgatggt gatatgaaaa gctgtgattt aattagtagg agtaattcat ttccatgtta 360
caattactta tctttctcct ccttgtgttc ttgccactgc tagaggcaga tctataattc 420
ctacttagag tattgaacac attgtacttt tgaaattgta agtgtgcagt ttaatatgtg 480
ttaaaatttt gaagggtttt ttaattacat gtatatcgtg acgttatata gcctctctac 540
ccgcataagg taggggtaat gtctgcgtac aaataggctg atttactct gtgttcagtt 600
gaatcaatat atgcggcatc caccactgtt tgcaaatgca tgggaatgga aaagtcatgc 660
ctcatcactg tcacacatga tttttattta tttatttatt atatatgatt ctttaaactc 720
attagtgggt actgaatctt gctaactgat cgattgatga aacaggttga agcatcgtgg 780
accagattgg agtgggctgt atcaacatgg ggactgttac ttggcacatc agcgtctagc 840
tattgttgat cctgcttccg gtgatcaacc tctgtttaac gaagataaga cgattgttgt 900
tacggtgggt gcttttaatt tcatgtttac tctcatttct tgctcagttt actgatgaaa 960
acgaaactaa tgctgataac tatcaaccaa tcaaccacat acaagtaggg gtcgactata 1020
tgattcatta tatcctttcc tctattcaaa tagttgatga ttaatttcat gccaaaaatt 1080
tcaggtaaat ggagagatct acaatcacga gcaacttcgt aagcaaatgc ctaatcataa 1140
gttccggact ggcagtgact gtgatgtcat tgcacaceta gtgagtactc atcttccaac 1200
ttaggaaaca tgtatgtata tgtaaaatta ttccctgatt atattatgta cattaatcag 1260
tatgaagaac atggagaaga ttttgtggac atgctggatg ggatcttcgc ttttgtgtta 1320
ttggatactc gagataacag ctttcttgtt gctcgtgatg ccattggaat tacttccctt 1380
tatattgggt ggggacttga tggtaagatt ccatatgatt ttctcactta gattttgtta 1440
ttgcaaatga atatgttgaa tccttatatg tactaattgt tgtaattctg ttccttctca 1500
```

gggtctgtat ggatatcatc tgagcttaag ggcttgaatg atgactgcga acattttgaa 1560
gttttccac cagggcacctt gtactctagc aagaatggcg gctttaggag gtggtacaat 1620
cctccttgggt tctctgaggc cattccttcc actcgttatg atcccttagt tctcaggcgt 1680
gcctttgaaa atgtgagtta atattgttgt tgttgaaatg gatccaacat gttggtttta 1740
aggctatata tcattttgca ggctgttatc aaaagggtga tgactgatgt cccctttgggt 1800
gttctcctct cggggggact cgattcatcc ttgggtgctt cgattactgc tcgctacttg 1860
gctggtacaa aggctgccaa gcagtgggga gcacagcttc attccttctg tgttggcctt 1920
gaggtcagat aatcttgaga cttgtgtgat tacaacatta ataacgacaa cacctcattc 1980
acaagctagt aaggtcagct atatgaaatc tgataccgat ttgcttgctt tcaccgaaca 2040
gggctcaccg gatctcaagg ctgcaagaga agttgctgac tacttgggaa ccgttcacca 2100
cgagtttcac ttcaccgttc aggttttgat ggagcaactc tattttcact tgatctatca 2160
tcctcttcgt ttctcctttg tttacctgat ggcaaatgac atgtgttgta tttgcaggat 2220
ggaattgatg caattgaaga tgttatttac catattgaga catacgatgt aacgacaatc 2280
agagcaagca ctctatgtt ccttatgtcg cgtaagatta agtcactagg agtgaagatg 2340
gtcatatctg gggaaggatc tgatgaagtg tttggtggct acttgactt tcacaaggct 2400
cccaacaagg aagagttcca caaggaaaca tgtcgcaagg tataagaaag tgcttccagt 2460
tcctagttta aataggacaa ggagttgaga aaatgttgat tttttcccta tgcagattaa 2520
agcgcttcac caatatgact gcttaagagc aaataagtca acatctgcat ggggtttaga 2580
agctagagtc cctttcctag ataaggagtt catcaatgtt gccatgagta ttgatccaga 2640
gtggaagttg gtaagtagtt gaccccaat tgttgatcga ataaacaatc aactgtccct 2700
aaaactatgt atttattcat ttcagattaa accagagcaa aggaggattg aaaagtgggc 2760
tctaaggagg gcctttgatg atgaggagca tccttatctc ccaaagggtg actaaatcgc 2820
cattttcaaa ctgatttgca aatgattttg aatgttatat tattaatcaa actaacaatt 2880
tactcattac attccttgca gcacatcctg tataggcaaa aagaacaatt cagtgatggc 2940
gtgggctata gttggataga tggactcaaa gcacatgctg aacaacatgt gcgtttcaat 3000
gactaatcaa cgactttcta ctgctcttat ttatcttcca atgcttgttt gttttaaaagc 3060
tgaccgagga aacaaatttt ttcctttaat aggtgaccaa taggatgatg tttaatgctt 3120
cacatatatt ccctcataac acaccatta caaaggaagc atactactat aggatgattt 3180
tcgagcgctt tttccacag gtaatttggt cagataatag tgagaattta tgataaatgg 3240
actaatttac gtttggtgct gattttgcaa cttccaaca gtgtgagcaa gctcacataa 3300
tcggaatttt ttgttcaaat aataatgcat tggaaatcat aatgctaaat tgagatttac 3360
ttttaaccat gttcctctac catttttgggt tctgcagaat tcagctgggc taaccgttcc 3420
tgaggagca agtggtggcg ttagcacagc taaagctgta gagtgggatg cttcttggtc 3480
aaagaacctt gatccttcag gcagggtgc tattggtgta cataactcgg cttatgagaa 3540
tcatgtacct gctatggcta atgggaattt gacaaaaaa atcattggtc gtgtgccttc 3600
tatggtagaa gttggtgctg ctcccgagct cacaataaag agttag 3646

<210> 2

<211> 590

<212> PRT

<213> 烟草

<400> 2

```

Met Cys Gly Ile Leu Ala Val Leu Gly Cys Ser Asp Asp Ser Gln Ala
1           5           10           15
Lys Arg Val Arg Val Leu Glu Leu Ser Arg Arg Leu Lys His Arg Gly
          20           25           30
Pro Asp Trp Ser Gly Leu Tyr Gln His Gly Asp Cys Tyr Leu Ala His
          35           40           45
Gln Arg Leu Ala Ile Val Asp Pro Ala Ser Gly Asp Gln Pro Leu Phe
          50           55           60
Asn Glu Asp Lys Thr Ile Val Val Thr Val Asn Gly Glu Ile Tyr Asn
65           70           75           80
His Glu Gln Leu Arg Lys Gln Met Pro Asn His Lys Phe Arg Thr Gly
          85           90           95
Ser Asp Cys Asp Val Ile Ala His Leu Tyr Glu Glu His Gly Glu Asp
          100          105          110
Phe Val Asp Met Leu Asp Gly Ile Phe Ala Phe Val Leu Leu Asp Thr
          115          120          125
Arg Asp Asn Ser Phe Leu Val Ala Arg Asp Ala Ile Gly Ile Thr Ser
          130          135          140
Leu Tyr Ile Gly Trp Gly Leu Asp Gly Ser Val Trp Ile Ser Ser Glu
145          150          155          160
Leu Lys Gly Leu Asn Asp Asp Cys Glu His Phe Glu Val Phe Pro Pro
          165          170          175
Gly His Leu Tyr Ser Ser Lys Asn Gly Gly Phe Arg Arg Trp Tyr Asn
          180          185          190
Pro Pro Trp Phe Ser Glu Ala Ile Pro Ser Thr Arg Tyr Asp Pro Leu
          195          200          205
Val Leu Arg Arg Ala Phe Glu Asn Ala Val Ile Lys Arg Leu Met Thr
          210          215          220
Asp Val Pro Phe Gly Val Leu Leu Ser Gly Gly Leu Asp Ser Ser Leu
225          230          235          240
Val Ala Ser Ile Thr Ala Arg Tyr Leu Ala Gly Thr Lys Ala Ala Lys
          245          250          255
Gln Trp Gly Ala Gln Leu His Ser Phe Cys Val Gly Leu Glu Gly Ser
          260          265          270
Pro Asp Leu Lys Ala Ala Arg Glu Val Ala Asp Tyr Leu Gly Thr Val
          275          280          285
His His Glu Phe His Phe Thr Val Gln Asp Gly Ile Asp Ala Ile Glu

```

290	295	300
Asp Val Ile Tyr His Ile Glu Thr Tyr Asp Val Thr Thr Ile Arg Ala		
305	310	315
Ser Thr Pro Met Phe Leu Met Ser Arg Lys Ile Lys Ser Leu Gly Val		320
	325	330
Lys Met Val Ile Ser Gly Glu Gly Ser Asp Glu Val Phe Gly Gly Tyr		335
	340	345
Leu Tyr Phe His Lys Ala Pro Asn Lys Glu Glu Phe His Lys Glu Thr		350
	355	360
Cys Arg Lys Ile Lys Ala Leu His Gln Tyr Asp Cys Leu Arg Ala Asn		365
	370	380
Lys Ser Thr Ser Ala Trp Gly Leu Glu Ala Arg Val Pro Phe Leu Asp		
385	390	395
Lys Glu Phe Ile Asn Val Ala Met Ser Ile Asp Pro Glu Trp Lys Leu		400
	405	410
Ile Lys Pro Glu Gln Arg Arg Ile Glu Lys Trp Ala Leu Arg Arg Ala		415
	420	425
Phe Asp Asp Glu Glu His Pro Tyr Leu Pro Lys His Ile Leu Tyr Arg		430
	435	440
Gln Lys Glu Gln Phe Ser Asp Gly Val Gly Tyr Ser Trp Ile Asp Gly		445
	450	460
Leu Lys Ala His Ala Glu Gln His Val Thr Asn Arg Met Met Phe Asn		
465	470	475
Ala Ser His Ile Phe Pro His Asn Thr Pro Ile Thr Lys Glu Ala Tyr		480
	485	490
Tyr Tyr Arg Met Ile Phe Glu Arg Phe Phe Pro Gln Asn Ser Ala Gly		495
	500	505
Leu Thr Val Pro Gly Gly Ala Ser Val Ala Cys Ser Thr Ala Lys Ala		510
	515	520
Val Glu Trp Asp Ala Ser Trp Ser Lys Asn Leu Asp Pro Ser Gly Arg		525
	530	535
Ala Ala Ile Gly Val His Asn Ser Ala Tyr Glu Asn His Val Pro Ala		540
545	550	555
Met Ala Asn Gly Asn Leu Thr Lys Lys Ile Ile Gly Arg Val Pro Ser		560
	565	570
Met Val Glu Val Gly Ala Ala Pro Glu Leu Thr Ile Lys Ser		575
	580	585
		590

<210> 3

<211> 4564

<212> DNA

<213> 烟草

<220>

<221> misc_feature

<222> (3137) .. (3137)

<223> n是a、c、g或t

<400> 3

```

atgtgcggga tcttggctgt tttgggttgt tctgatgatt ctcaggccaa aagggttcgt 60
gttctcgcagc tctctcgcag gttaaattata ttcccttctt ttctctattg ttaatgttat 120
tgtgttgcat gcgttattaa ttctttttat gagaaagtag tttcctgagc gtttggagct 180
gctactgatt gcttgcattt tagaataaaa cttaagatgg attatatatt tagatattct 240
tgacgtggac ctatatctta gattcacaat tttctcggac caactgtttt agcatgactt 300
gtatatgatg ttgatatgaa aagctgtggg ttagtaattc atttccatgt tacaattact 360
aatctttctc ctatctgtgt tcttgccact gctagaggca gatctataat tcgtaattag 420
agtattgaac acattgtact tttcaaaatt ataagtgcgc aatttaatat ttgttaaaat 480
tttgaaggct tttttaatta catgtatata gtgtctcgta tagcctctct acccatgtaa 540
ggctctgcgta cacattaccc gccccatgcc ctcccgtaaa attttactgg ggtttgctgt 600
tgtttgttgg atcgtgtgtc gcataactt tgttctgttg aatcaatata tgcgtcatct 660
gccactgttt gcaaattgcat ggaatggaaa agtcatgcct catcactgtc acacaggatt 720
tttattttatt tattttattt ttatatatga ttctttttct tgatgtgaaa tctatttagtg 780
ggtactgaat ctggctaact gactaatgaa acaggttgaa gcatcgtgga ccagattgga 840
gtgggctgta tcaacatggg gactgttact tggcacatca gcgtctagct attgttgatc 900
ctgcttccgg tgatcaacct ctgtttaacg aagataagac gattgttggt acgggtgggtg 960
cttttaattt catgtttact ctcatctctt gctcagttta ctgatgaaca tgaaactaat 1020
gactatcaat cagccaatca accacgttta attccataca agtaggggtc gactatatga 1080
ttcattatat cctttcctct attccgggtt ctttttattc aaatagttga tgattaattt 1140
catgccaaaa atttcaggta aatggagaga tctacaatca cgagcaactt cgcaagcaaa 1200
tgccatgatca taagttccgg actggaagtg actgtgatgt cattgcacac ctagtgagta 1260
atctagatta ctcatcttcc aacttaggaa acatgtatgt attcccctga ttatttcate 1320
atctttatgt acattaatca gtatgaagaa catggagaag attttgtgga catgctggat 1380
gggatcttcg cttttgtgtt actggatact cgagataaca gctttcttgt tgctcgtgat 1440
gccattggaa ttacttcctt ttatatgttg tggggacttg atggttaagat tccatattat 1500
tttctcactt agattttgtt attgcaaatt aatatgttga atctttatat gtagtaactg 1560
ttgtaatttg gttccttctc agggctctgta tggatatcat ctgagcttaa gggcttgaat 1620
gatgactgcg aacattttga agttttccca ccaggacact tgtactctag caagaatggc 1680
ggcttttagga ggtggtacaa tcctcttttg ttctctgagg ctattccttc cactccttat 1740
gatcccttag ttctcaggcg cgcctttgaa aatgtgagtt attattgttg ttgttgaaat 1800
ggatccaaca tgttggttct aaggctggat atcattgtgc aggctgttat caaaaggttg 1860
atgactgatg tcccttttgg tgttctgctc tccgggggac tcgattcate cttggttgc 1920

```

```

tcgattactg cccgctactt ggctggcaca aaggctgcc aagcagtgagg agcacagctt 1980
cattccttct gtgttggcct tgaggtcaga taatcttgag atttctgtga ttacaacatt 2040
aataacgacc acacctcatt cacaagctag taaggtcagc tatatgaatc tgataccgat 2100
ttgcttgctt tcaaacgaac agggatcacc ggatctcaag gctgcaagag aagttgctga 2160
ctacttggga accgttcacc acgagtttca cttcaccgtt caggtattga tggagcaact 2220
ctattttgac ttgatctatc atcctctgcg ttcccttttg tttacctgat ggcaaatgac 2280
atgtgtcgta tttgcaggat ggaattgatg caattgaaga tggtatttac catattgaga 2340
catacgatgt aacgacaatc agagcaagca ctctatgtt ccttatgtcg cgtaagatta 2400
agtcactagg agtgaagatg gttatatctg gggaaggctc tgatgaagtg tttggtggct 2460
acttgtactt tcacaaggct cccaacaagg aagagttcca caaggaaaca gtgcgcaagg 2520
tataaaaaag tgtttccagt tccttgttta aaaccttatt taaataggat aagcaacaac 2580
aacaacaaca acccagtata atcccactag tggggtctga ggagggtatt gtgtatgcag 2640
accttaccct taccctgggg tagagagget gtttcagata gaaactcgac tccctccctc 2700
caaaaacctt atttaaatag gacaaggagt tgagaaaatg ttgatctttt tcttccatct 2760
ctatgcagat taaagcactt caccaatatg actgtttaag agcaaataag tcaacatctg 2820
catggggctt agaagctaga gtgcctttcc tagataagga gttcatcaat gttgcatga 2880
gtattgatcc agagtggag ttggttaagta gttgaacccc aattgttgat cgaataaacg 2940
atcaactgtc cctaaaacct tgtattctat ccatttcaga ttaaaccaga gcaaaggagg 3000
attgagaagt gggctctaag gagggccttt gatgatgagg agcatcccta tctcccaaag 3060
gtggtactaa atcgcccttt tcaaaactgat ttgcaagtga tttcgaatgt tatattgtta 3120
accaaaactaa caatttnaag gctctgatga agtgtttggg ggctacttgt actttcacia 3180
ggctcccaac aaagaagagt tccacaagga aacatgtcgc aaggtataaa aaagtgtttc 3240
cagttccttg tttaaaacct tatttaaata ggataagcaa caacaacaac aacaaccag 3300
tataatccca ctagtgggggt ctgaggagggt tattgtgtat gcagacctta cccctaccct 3360
ggggtagaga ggctgtttca gatagaaact cgactccctc cctccaaaaa cttatttta 3420
ataggacaag gaggttgagaa aatgttgatc ttttccctc atctctatgc agattaaagc 3480
acttcaccaa tatgactgtt taagagcaaa taagtcaaca tctgcatggg gcttagaagc 3540
tagagtgcct ttcctagata aggagttcat caatgttgcc atgagtattg atccagagtg 3600
gaagttggta agtagttgaa cccaattgt tgatcgaata aacgatcaac tgtccctaaa 3660
accttgtatt ctatccattt cagattaaac cagagcaaag gaggattgag aagtgggctc 3720
taaggagggc ctttgatgat gaggagcatc cctatctccc aaaggtggta ctaaategcc 3780
cttttcaaac tgatttgcaa gtgatttcga atgttatatt gttaaccaa ctaacaattt 3840
actcattaca tttcctgcag cacatcctat acaggcagaa agaacaattc agtgatggcg 3900
taggctatag ttggatagat ggactcaaag cacatgtcga acaacatgtg cgtttcaatg 3960
actaatcaac aactttctaa tgttcttatt tatcttccaa tgtttgtttg ttttaaagct 4020
gacttggaa tttttgttt ctttaaatag gtgaccaata ggatgatgct taatgcttca 4080
catatatcc ctcataacac accgattaca aaggaagcat actattatag gatgattttc 4140
gagcgctttt tcccacaggt aatttgttca tgtattgtga ggattcatga taaattagct 4200
aacttatgct ggttttcaac cttgcaacct tccgataagt cacattgttg gaggttgttg 4260

```

ttcaaatgcg agatttactt ttaaccatgt tcctctggta tttttggttc tgcagaattc 4320
 agctgggcta accgttcctg gaggagcgag tgtggcgtgt agcacagcta aagctgtaga 4380
 gtgggatgct tcttgggtcaa agaaccttga tccttcagga agggctgcta ttggtgtaca 4440
 taactcagct tatgagaatc atgaacctgc tatggctaata gggaatttgg ccacaaaaat 4500
 cattggccgt gcgccgtcta tggtagaagt tggtgctgct catgagctca caataaggag 4560
 ttag 4564

<210> 4

<211> 590

<212> PRT

<213> 烟草

<400> 4

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

Asp Val Pro Phe Gly Val Leu Leu Ser Gly Gly Leu Asp Ser Ser Leu			
225	230	235	240
Val Ala Ser Ile Thr Ala Arg Tyr Leu Ala Gly Thr Lys Ala Ala Lys			
	245	250	255
Gln Trp Gly Ala Gln Leu His Ser Phe Cys Val Gly Leu Glu Gly Ser			
	260	265	270
Pro Asp Leu Lys Ala Ala Arg Glu Val Ala Asp Tyr Leu Gly Thr Val			
	275	280	285
His His Glu Phe His Phe Thr Val Gln Asp Gly Ile Asp Ala Ile Glu			
	290	295	300
Asp Val Ile Tyr His Ile Glu Thr Tyr Asp Val Thr Thr Ile Arg Ala			
305	310	315	320
Ser Thr Pro Met Phe Leu Met Ser Arg Lys Ile Lys Ser Leu Gly Val			
	325	330	335
Lys Met Val Ile Ser Gly Glu Gly Ser Asp Glu Val Phe Gly Gly Tyr			
	340	345	350
Leu Tyr Phe His Lys Ala Pro Asn Lys Glu Glu Phe His Lys Glu Thr			
	355	360	365
Cys Arg Lys Ile Lys Ala Leu His Gln Tyr Asp Cys Leu Arg Ala Asn			
	370	375	380
Lys Ser Thr Ser Ala Trp Gly Leu Glu Ala Arg Val Pro Phe Leu Asp			
385	390	395	400
Lys Glu Phe Ile Asn Val Ala Met Ser Ile Asp Pro Glu Trp Lys Leu			
	405	410	415
Ile Lys Pro Glu Gln Arg Arg Ile Glu Lys Trp Ala Leu Arg Arg Ala			
	420	425	430
Phe Asp Asp Glu Glu His Pro Tyr Leu Pro Lys His Ile Leu Tyr Arg			
	435	440	445
Gln Lys Glu Gln Phe Ser Asp Gly Val Gly Tyr Ser Trp Ile Asp Gly			
	450	455	460
Leu Lys Ala His Ala Glu Gln His Val Thr Asn Arg Met Met Leu Asn			
465	470	475	480
Ala Ser His Ile Phe Pro His Asn Thr Pro Ile Thr Lys Glu Ala Tyr			
	485	490	495
Tyr Tyr Arg Met Ile Phe Glu Arg Phe Phe Pro Gln Asn Ser Ala Gly			
	500	505	510
Leu Thr Val Pro Gly Gly Ala Ser Val Ala Cys Ser Thr Ala Lys Ala			
	515	520	525
Val Glu Trp Asp Ala Ser Trp Ser Lys Asn Leu Asp Pro Ser Gly Arg			

530	535	540
Ala Ala Ile Gly Val His Asn Ser Ala Tyr Glu Asn His Glu Pro Ala		
545	550	555
Met Ala Asn Gly Asn Leu Ala Thr Lys Ile Ile Gly Arg Ala Pro Ser		
565	570	575
Met Val Glu Val Gly Ala Ala His Glu Leu Thr Ile Arg Ser		
580	585	590
<210> 5		
<211> 5276		
<212> DNA		
<213> 烟草		
<400> 5		
atgtgtggaa tcttggcttt gttgggttgt tcagatgatt ctcaggccaa aagggttcga 60		
gttccttgagc tctctcgcag gcaatttccc ctttcttttc tcttttattt tattactcca 120		
tcatataatt attctacttt cacctttttt gttttttaat ttcgatttgt ttgctttttg 180		
ctttcttcat ttaatgttct tgcattgctc tagctcctat acgaataata ggagctaaag 240		
taatagagtg aaggattatt ttggagcaac ggtaaagtta tttttgtata gttcacaggt 300		
tggagccgta aaagcagtta ctagtgtttg cattagggtta agttgtctat caggaaagat 360		
attttagtga attgaccacc ctttggcaaa aaattatgcg taaaatatta ctacttggtta 420		
ttgattcaaa aaatatgact ttgaacatcc ttgtatagcc tagtggcaaa ggggtgttcaa 480		
aatttttgaa caccttttatt gaattttctg gtttcgtcgt tgctgtctac aacacactct 540		
cgaagtgcgg ccttccctca aaccctacgt gaatatggaa cgcctagtgt gccgggatgc 600		
cctttaggct gagtaataga gtactatggt gatataccta aaatacggaa ctggcacgag 660		
agataaattc agaatttata gcttaattat ttaaaattta tcatttgatt tatttcaacc 720		
agagttcatt ctagtagagt ttatgttgca tatttgtctg atttgcacca tttgttttaa 780		
tgttcttgca ttactctaatt ttatttatga ataataagct aaactaattg agtttatggt 840		
gatgtaactg acctgagttg gcagtgatgt ttggctactt gaatttaaaa caagattaag 900		
acttgtgagt tggcagtaat tgtttgata cttgaattta gaacaagaat aaaacaagat 960		
taattaagga tacattatc ctaaccggga ccctaattct atagattgat tattaaattt 1020		
ttcttggacc aaccgtcacc gtgacctgtt catgatgggt atggcaaaag ccataatata 1080		
cttgattttct gttttcaatt tttaagggtt ttattattat acttattata tactttgaaa 1140		
attatgggtc aaaaattttta ttttgttat tcttccgta tgtttgactc aacttggagt 1200		
tcatgaaaaa aaaaaaaaaat ttttttaaac ttgtggtcct aaacaagtca tcgatatatg 1260		
tgtggttgtg aatcgcctcg ttaggataaa gtagaaattt taaagttaaa ctatttcgaa 1320		
ataaaaaagt atgatatgtt ttttgggaca tactaaaaag gaaataacac ataaaaatgga 1380		
acaagaagt atatttttagt gatttttcac atatatatca gtattttgta tcaaaaatgc 1440		
tggttcagtt gaactcattg aactgggtgc tccatccacc tccgagtact gcaacatgtg 1500		
accggtggca gatgcaaccc ttacgtacca gggtcactct aacctagtat aaatatattt 1560		
gtgaaaaatc ctaaaaattt caataattag tagattttaa accttaattt taaaactaca 1620		

```

atgagttcag ttttaaaatc tttatatgtc aaacctatca agattaaatc ctggatccac 1680
cttttgcattg tgacagtaca atagcctttt gtaattgaaa caaatatttg aatttttggt 1740
actatttgat cacatagggt gaagcatcgt ggaccagatt ggagtgggat atatcaacat 1800
ggatgattttt acttagcaca tcaacgttta gcaattatcg atcctgcttc tggatgatcag 1860
cctctgttta atcaagataa gacgattgtt gttacagtga gtgcctctaa tttacctctt 1920
ttttttttct tttttttttt acgatcgcgt tgtttagatt atttatggat gaaaatttgt 1980
atatttgatt attgattgtg tgctaaattt tgcagggtcaa tggagagatt tacaatcatg 2040
agaaacttcg taatcttatg cctaatacaca agttcagaac tggaagtgat tgtgatgtta 2100
ttgcacatct tgtgagttag tttacttttc tcattgacca acttaggcca gtgaattata 2160
ttaggaattt cgggtgtttg gatattgtca aattagctta taagcactta ttatcaattt 2220
taacattttt atccatacac gtaactgttc attcataaag tcatttttagc acttgatgct 2280
catcaacctt tgttgttggg actcttcgaa aatgtcagcc cgatagcatt tcggagagtc 2340
cgagcaacat agcaatcaac tacttttaat cagttaaacg aacatgctca tagtgataac 2400
ctctggaaaa tatgttatag taaattagaa aaatttatac gatgtctcac tateggtatg 2460
tattgatcag tatgaagaat atggagaaaa ttttgtggac atgttggatg ggggtgttctc 2520
ttttgtattg ttggatacgc gcgataacag ctttcttgcg gctcgtgatg caattggaat 2580
tactccccta tatattggtt ggggacttga tggtaagatt ttctatacaa tttccagtt 2640
agaaattaaa actgaagcta ttcctattta ctattgaaga atcttgatac atgttatttg 2700
gtatttcgtc gtgttgttat gtctcaggct ctgtgtggat ttcatctgag ctaaaaggat 2760
taaattggtg ctgtgaacat tttgaagttt tccctcccg tcaattgtac tcgagcaaga 2820
atggcgggtt taggatggg tacaatcctc aatggttctc tgaggctatt ccatcaaate 2880
cttacgaccc cttagttttg agacgtgcct tcgaaaatgt gaggtttgtt aatgaatcta 2940
gtgtacaate ttggatgttt ttctctcgac tttggtttta gcactcaaaa agttgtctgt 3000
acaaacgatac ttggctataa ggagtgggtg ttaattttgc aggctgttat caaacgattg 3060
atgaccgatg tcccctttgg tgttctgctc tccggggggc ttgattcgtc tttggttgc 3120
tctgtcactg ctcgctactt ggctggaaca aaagctgcta agcaatgggg agcacagctt 3180
cattccttct gtgttgttct cgaggtcaga ctatcaatcc tgagcggaga catatttggg 3240
gcgggttctg tgggttcaac tgaaccatt gcttttggtt cgaaccatat atacaaatcc 3300
aagaaaaatat aagatgtgta catataatat gacactcatt ctgagcctac tgacgctaga 3360
tactggactc gcctttggte ttgagacttt ataatagaaa atagatatcc aataccaatt 3420
tctgtttttc ttttctgact gacagggtc accagatctc aaggctgcaa gagaagttgc 3480
tgactatttg ggaaccgtc accacgagtt caccttcaca gttcaggttt gtaaatecgtg 3540
ttgctactcc gatttattca tgtgagtctc ttctgttttc ctctagaaa gcaacaagtg 3600
ttcatgtttt gcaggatgga attgatgcta ttgaagatgt tatttaccat atcgagacat 3660
acgatgtaac aacgatcaga gcaagcactc ctatgttctt tatgtcgcgt aagattaaat 3720
cactgggagt taagatggte atatcagggg aaggtcaga tgaactgttt ggcggtctatt 3780
tgtacttcca caaggctccg aacaaggaag aattccatgt ggagacatgt cacaaggtaa 3840
taaaacacgt cggtccctt ggcaagtctg attctccttg tcattgtttg tctcggttg 3900
gaataaattg ctataattcg agaagaaatg ttaatcctgc tctttcactc gcatgcagat 3960

```

aaaagcgctt caccaatac agctgtttgag agcaaataag gcaacatcag catggggctt 4020
 agaagctaga gtaccatttc tggataaaga gtatcatcaac gttgctatga gtatcgatcc 4080
 tgaatggaag atggtaacaa aactgagctc caattgtcaa tgagatgaga agcataactg 4140
 aatatgcacg ttctctaaaa ctgtgttttc atccgtttca gattaaacac gatcatggta 4200
 ggatcgagaa gtgggttctt aggaaggctt ttgatgatga ggagcaaccc tatctcccaa 4260
 aggtccacaa agtcatgttt ttctactga tttgccgatt agtatggatc aaaactaaca 4320
 tgttttgttc ttttatgtac tctgcagcat attctgtacc ggcagaaaga acaattcagt 4380
 gatggcgtag gctatagttg gatcgatgga ctcaaagcac atgctgaaca acatgtgagc 4440
 ttctaaataa atccagaact attacatcgt tategtccaa ttttcgtttg ttttgctgtc 4500
 agctgatctt ggaaacaatt cttttctctc aggtgactga taggatgatg cttaatgctg 4560
 cacatatctt ccctcacac actccaacta caaaggaagc atactattac aggatgattt 4620
 tcgagagggt cttccacacag gtacttatag cgcgagacta gaatcatata tctcttaaat 4680
 ggaaacatga aaaaataaagg aggtgttttt tctccttgga aaaagtattt actactgatt 4740
 aaagcacctt cattcaccca tcttgetact tcgcgtactt tactaagttg tggttctaaa 4800
 aaagaagatt tagcttggtg tattccatca cgaggaagcc acaaaaagat tctacttata 4860
 agaagttgga tacttttaat ggtgtgagac cttttgggaa aaaccatgcg ggcttgcccc 4920
 aaagccgaca atatcacacc atgttaagag tatctttgct tcgttttagc caaacgtac 4980
 aataagtcaa tttttaaac atgttggtac tattgttgga tttgcagaat tcagcaaggc 5040
 taactgttcc tggaggaccg agtatagctt gcagcacagc taaagctatt gagggggatg 5100
 ctctgtggtc gaacaacctt gatccttcg gtagggctgc aatcggtgta cataactcgg 5160
 cttatgacga tcactcccc gatgttggtg atgggaattt ggacacaacg atcatcgata 5220
 atgtgccgag gatggtagga gtgggtgctg ctgcagagct cacaataagg agctag 5276

<210> 6

<211> 590

<212> PRT

<213> 烟草

<400> 6

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

Ser Asp Cys Asp Val Ile Ala His Leu Tyr Glu Glu Tyr Gly Glu Asn		
100	105	110
Phe Val Asp Met Leu Asp Gly Val Phe Ser Phe Val Leu Leu Asp Thr		
115	120	125
Arg Asp Asn Ser Phe Leu Ala Ala Arg Asp Ala Ile Gly Ile Thr Pro		
130	135	140
Leu Tyr Ile Gly Trp Gly Leu Asp Gly Ser Val Trp Ile Ser Ser Glu		
145	150	155
Leu Lys Gly Leu Asn Gly Asp Cys Glu His Phe Glu Val Phe Pro Pro		
165	170	175
Gly His Leu Tyr Ser Ser Lys Asn Gly Gly Phe Arg Arg Trp Tyr Asn		
180	185	190
Pro Gln Trp Phe Ser Glu Ala Ile Pro Ser Asn Pro Tyr Asp Pro Leu		
195	200	205
Val Leu Arg Arg Ala Phe Glu Asn Ala Val Ile Lys Arg Leu Met Thr		
210	215	220
Asp Val Pro Phe Gly Val Leu Leu Ser Gly Gly Leu Asp Ser Ser Leu		
225	230	235
Val Ala Ser Val Thr Ala Arg Tyr Leu Ala Gly Thr Lys Ala Ala Lys		
245	250	255
Gln Trp Gly Ala Gln Leu His Ser Phe Cys Val Gly Leu Glu Gly Ser		
260	265	270
Pro Asp Leu Lys Ala Ala Arg Glu Val Ala Asp Tyr Leu Gly Thr Val		
275	280	285
His His Glu Phe Thr Phe Thr Val Gln Asp Gly Ile Asp Ala Ile Glu		
290	295	300
Asp Val Ile Tyr His Ile Glu Thr Tyr Asp Val Thr Thr Ile Arg Ala		
305	310	315
Ser Thr Pro Met Phe Leu Met Ser Arg Lys Ile Lys Ser Leu Gly Val		
325	330	335
Lys Met Val Ile Ser Gly Glu Gly Ser Asp Glu Leu Phe Gly Gly Tyr		
340	345	350
Leu Tyr Phe His Lys Ala Pro Asn Lys Glu Glu Phe His Val Glu Thr		
355	360	365
Cys His Lys Ile Lys Ala Leu His Gln Tyr Asp Cys Leu Arg Ala Asn		
370	375	380
Lys Ala Thr Ser Ala Trp Gly Leu Glu Ala Arg Val Pro Phe Leu Asp		
385	390	395
Lys Glu Phe Ile Asn Val Ala Met Ser Ile Asp Pro Glu Trp Lys Met		

405	410	415
Ile Lys His Asp His Gly Arg Ile Glu Lys Trp Val Leu Arg Lys Ala		
420	425	430
Phe Asp Asp Glu Glu Gln Pro Tyr Leu Pro Lys His Ile Leu Tyr Arg		
435	440	445
Gln Lys Glu Gln Phe Ser Asp Gly Val Gly Tyr Ser Trp Ile Asp Gly		
450	455	460
Leu Lys Ala His Ala Glu Gln His Val Thr Asp Arg Met Met Leu Asn		
465	470	475
Ala Ala His Ile Phe Pro His Asn Thr Pro Thr Thr Lys Glu Ala Tyr		
485	490	495
Tyr Tyr Arg Met Ile Phe Glu Arg Phe Phe Pro Gln Asn Ser Ala Arg		
500	505	510
Leu Thr Val Pro Gly Gly Pro Ser Ile Ala Cys Ser Thr Ala Lys Ala		
515	520	525
Ile Glu Trp Asp Ala Ser Trp Ser Asn Asn Leu Asp Pro Ser Gly Arg		
530	535	540
Ala Ala Ile Gly Val His Asn Ser Ala Tyr Asp Asp His Leu Pro Asp		
545	550	555
Val Gly Asn Gly Asn Leu Asp Thr Thr Ile Ile Asp Asn Val Pro Arg		
565	570	575
Met Val Gly Val Gly Ala Ala Ala Glu Leu Thr Ile Arg Ser		
580	585	590

<210> 7

<211> 5926

<212> DNA

<213> 烟草

<400> 7

```

atgacgtagt ttagcttggt cgtccacggc ataaaaacgc ctataaatc ggtggtcatt 60
ttggttcatt tacagaatcc cagcatcatc ttttgtgtat tateattttg ctttctttac 120
gtgttcttct ctcccaaaat cctgcaaate tttaactate tteccatatt tctctccctc 180
tttctctttc tttttgttat ctactttact acttctatta gtagagcttt atccaattaa 240
agtcacatcat attggatata gtttttacta ctactttate ttgtttggga aagaaaagaa 300
taaaaaatta gagtgtctat attaagtata aaaaggtgaa aaataaaaaa gtagtgatca 360
atcatgtgtg ggatcttggc tttgttgggt tgtccagatg attctcaagc caaaagggtt 420
cgagttcttg agctctctcg caggcaatc ccctttcttt tctcttttat tttattactc 480
catcatatta ttcaactata atttcacett ttttgTTTTT caattaatat ttgtctgctt 540
ttgctttctt catttaatgt tcttgcatg ctctagtccc tttatgaata ataggctaaa 600
gtaatagagt gaaggagtag agtcggagca acggtgaagc tgtctttaca tgatctatag 660

```

```

ctcacgggtt caccatagaa gcagccaatg gcgaagctag aaaattcaat aagggtgttc 720
tttcctttgc cagtgggcta tgcaagggtg ttcaaagtct atttttaatc aataacaagt 780
aatattttat cttatagtga aagcaaccac taatacttgc atttaggtaa ttagctgtct 840
atcaagggaa gttatgtgtg gtaagtagtg gcggagccag aattttcggt aaggggtgtc 900
aaaatatata aaagtaaaca taccaggaaa ttaaggggag tcaatacata gtatatatac 960
atataattta tttttttacc tagctataca gtatatTTTT tccgcgaagt gacacccctt 1020
cctataaggt ggctccgcca ctggtggtaa ggggtggtcaa ttgaacaccc ttcgccgaaa 1080
aattatactc tggatggtgt aaaatattac tacttggtac tgattcaata aataaatttt 1140
gaaaaccctt gcatagtcta gtagcaaaga gtgttcaaaa ttttgaacac atttattgaa 1200
ttctctggtt tcgccactgc tgtctacatc acatccctta aggtgcgatc ttttttcgaa 1260
tcatgtgtga atacgggatg ccttgtgttc caggtgccc ctttaggcta aagtaataga 1320
gtacgtctta attcttatgt tgatcatata cctgaaactc ggaactggca cgagagatat 1380
aaattcagaa tttatagcaa atttatcatt atgttgcat tgttttaatg ttcttgcatt 1440
actctaattt ttttatgaat aataagctaa actaattgag tttatgttga tgtaactgac 1500
ctgagtttgg ctacttgaag gttaaaacaa gattaagact tgtgagttgg cagcgattgt 1560
ttggatactt gaacacgttt aaaacaagat taagacatta tgatggatac attattccta 1620
acctgggtgg gacctaatc tatggattca ttattaaaat ttcttggaca aaccatcacc 1680
atgatctgtt catgatgttt atggcaaaaag ccataatata ctttatttct gtatcacaa 1740
tactttatac acatgttata gagttttaaa attattttta ttttgttat tccttgagtc 1800
tcttgatatt tgactccgca tgaaattcaa gaaaaaaaaa agaagatttt taaaatttat 1860
agtcgaaaac aagccatcga tatagtgtgt gttgttaata acctcgtag gcgtaaagta 1920
gaaattttta agttaaacta tttctaaata agaaaatatg atattgtttt tgcatagact 1980
aataaggaaa gtatgacaag taaaataaaa caaagaagta ttttttaatg attttttccg 2040
tgtcaaaaat gtcagtttag ttgaacttga actcattgaa actggtgacc agtggtggat 2100
gcaaacctta ggtacggagg tcatccgaac ctagtatagt tatatttgtg aaaaatcact 2160
aaaatttcaa taattactag atatgaaacc ttaattttta aactacaatg agtgcagttt 2220
taaaatcttt aaatgtcaaa ccgatcaaga ttaaattatg gatccacctt taaatgtcaa 2280
accgacagta caatagcctt ttgcaattga aacaaatata aacattacat ttattaaact 2340
agtttttaatt ttggttacta ttgattaca taggttgaag catcgtggac cagattggag 2400
tggtatatat caacatggtg atttttactt agcacatcaa cgttttagcaa ttatcgatcc 2460
tacttctggt gatcagcctc tgtttaatca agataagact attgttggtt cagtgaagtgc 2520
ctctaattta cccTTTTTTT tgtggacgaa tgcgttggtt agattacttt aattatggct 2580
gaaaatttgt atatttgatt attgattgtg tgccaaatgt tgcaggtcaa tggagaaatt 2640
tacaatcatg agaaacttcg taatcttatg cctaatacaca agttcagaac cggaagtgat 2700
tgtgatgtta ttgcacatct tgtgagttag tttactttac tcatcaacca acttaggcta 2760
gtgaattcaa aggcggagct agaattttga gtatatgagt tctggaattt aggacaagat 2820
aagttactgg gttcggatag attatttaga catattaagt agatttctta acacaaatac 2880
acggccccgag ccgaagctat tgggttctgc cgaaccgta gctagacttg tagctccgcc 2940
attgagtga tttatataagg aatttcgggt gtttgggtaa acttataagc tgggtcaaatt 3000

```

```

agcttataag cacttattat caatttcaac atttctatcc aaacacgtaa ctattcattc 3060
gatttttagca cttgatgcct atcgactact tttaatcagt taaacgaaca ggctcatagt 3120
gacaaactct agaaaatatg gtattgtaaa ttagtagaag aatttatacg atgtctcact 3180
atcgatatgt attgatcagt atgaagaata tggagaaaat tttgtggaca tgttgatgg 3240
gggtgttctct tttgtattgt tggatacgcg cgataacagc tttcttgctg ctcgtgatgc 3300
gattggaatt actccccctct atattggttg gggacttgat ggtaagattt tctatacaat 3360
tttccaagta gaaattaaaa ctgaagccat tccaatttgc tattgtagaa tcttgatata 3420
tgttattttgg tatttcaggc tctgtgtgga tttcatctga gttaaagggc ttaaattgatg 3480
attgtgaaca ttttgaagtt ttccctcccg ggcacttgta ctcgagcaag aatggcgggt 3540
ttaggagatg gtacaatcct caatggttct ctgaggtgt tccatcaaat ccttatgacc 3600
ccttagttct gaggcgtgcc ttcgaaaatg tgagaaatga atctaggagg tttccctctc 3660
gactttggac ttagcactct aaaaatttgt ctgtacaaac gatcttggct ataaggagtg 3720
gttggttaatt tcgcaggctg ttattaaacg gttgatgacc gatgtaccct ttgggtgttct 3780
gctctccggg ggacttgatt cgtcttttgt tgettctgtc actgctcgt acttggtgg 3840
aacaaaagct gctaagcaat ggggagcgca gtttcattcc ttctgtgttg gtctcgaggt 3900
cagtcaatca atcttgagta gaaaaaattt agggtcggtt ctgtgggttt aacttaactc 3960
attgcttttg gctcgaacca tatatgcaaa tccaaaaaga tttaatagatg atacatataa 4020
tataacactc attctgagcc tactgacgct agatactgga ctgtgccttt ggtcttgaga 4080
cttttgtaat agaaaattga tattcaatac caatttctgt tcttttgtaga ctaacagggc 4140
tcaccagatc tcaaggctgc aagagaagtt gctgactatt tgggaaccgt tcaccacgag 4200
ttcaccttca cagttcaggt ttgtaaatcg cgatgctact tcaatttatt catgtgaggc 4260
tcttctgttt tctactaga aagcaacatg tgtgcatgtt ttggcaggac ggaattgatg 4320
ctattgaaga tgtattttac catatcgaga cgtatgatgt aacaacgatc agagcaagca 4380
cccctatgtt ccttatgtcg cgtaagatta aatcacttgg agtgaagatg gtcatatcag 4440
gggaaggctc agatgaactg tttggtggct acttgtactt ccacaaggct cccaacaagg 4500
aagaattcca cacggagaca tgtcacaagg taataaaaca catcggcaag tctgattctc 4560
cttgtcattg tttgtctcgg gttggaataa attgctataa ttcgagatat ggaatgagtt 4620
aagaaatggt aatcatgttt ttccattcgc atgcagataa aagcgcttca ccaatacgac 4680
tgtttgagag caaataaggc aacatcagca tggggcttag aagctagagt accatttctg 4740
gataaagagt tcatcaatat tgctatgagt atcgatcctg aatggaagat ggtaagaaaa 4800
ctgagctcca attgtcaatg agacgagaag cataactgaa tatgcgcgtt ctctaaaacc 4860
tgtttttcat ccgtttcaga ttaaacacga tcaaggtagg atcgagaagt gggttcttag 4920
gaaggctttt gatgatgagg agcaccctta tctcccaaag gtccacaaac tcatattttt 4980
cctactgatt tgccaataat tatggatcaa aactaacatg tttgttctgt tatgtgctct 5040
gcagcatatt ttgtaccggc agaaagaaca attcagcgat ggtgtaggct atagttagg 5100
cgatgggctc aaagcacatg ctgaacaaca tgtgagcttc cttaataaat ccagaactaa 5160
tgcagcgta tcgtccaatt tttgtttgtt ttactatcaa ctaattttag tcatgttagt 5220
atgtttacct taggtactat gtttttaagg agtttttagt cgtgtcaaga atttattcgc 5280
tggtaaaaaa tagattgaac ttccagtatt ttctactttt attatgcata tgattggact 5340

```

gagatgaggt taaatggaac aacatggttg aggattaata taaccaattc caacttggtt 5400
 ggaattgagg cgtcgtagtt gttgttagat ctcagaaaca attcttttct ctcaggtgac 5460
 tgataggatg atgcttaatg cttcacatat cttccctcac aacactccaa ctacaaagga 5520
 agcatactat tacaggatga tttttgagag gttctttcca caggtaatat tagtgagaga 5580
 ttaaaatcta tagattttgt ttatggctga ggtttaaatc atagatttct taagctcaaa 5640
 aactaaaacc atgtccctgc tactgttgga ttgcagaat tcagcaaggc taactgttcc 5700
 tggaggaccg agtatagctt gcagcacggc taaagctatt gaggggacg cttcttggtc 5760
 gaacaacctt gatccttccg gtagggctgc tatcggtgta cataactcag cttatgacga 5820
 tcactacccc gatgttagta atgggaattt ggacacaacg atcatcgata atgtgccaag 5880
 gatggtagga gtgggtgctt ctgcagagct cacaataagg agctag 5926

<210> 8

<211> 565

<212> PRT

<213> 烟草

<400> 8

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

Val Ile Lys Arg Leu Met Thr Asp Val Pro Phe Gly Val Leu Leu Ser		
195	200	205
Gly Gly Leu Asp Ser Ser Leu Val Ala Ser Val Thr Ala Arg Tyr Leu		
210	215	220
Ala Gly Thr Lys Ala Ala Lys Gln Trp Gly Ala Gln Leu His Ser Phe		
225	230	235
Cys Val Gly Leu Glu Gly Ser Pro Asp Leu Lys Ala Ala Arg Glu Val		
	245	250
Ala Asp Tyr Leu Gly Thr Val His His Glu Phe Thr Phe Thr Val Gln		
	260	265
Asp Gly Ile Asp Ala Ile Glu Asp Val Ile Tyr His Ile Glu Thr Tyr		
	275	280
Asp Val Thr Thr Ile Arg Ala Ser Thr Pro Met Phe Leu Met Ser Arg		
	290	295
Lys Ile Lys Ser Leu Gly Val Lys Met Val Ile Ser Gly Glu Gly Ser		
305	310	315
Asp Glu Leu Phe Gly Gly Tyr Leu Tyr Phe His Lys Ala Pro Asn Lys		
	325	330
Glu Glu Phe His Thr Glu Thr Cys His Lys Ile Lys Ala Leu His Gln		
	340	345
Tyr Asp Cys Leu Arg Ala Asn Lys Ala Thr Ser Ala Trp Gly Leu Glu		
	355	360
Ala Arg Val Pro Phe Leu Asp Lys Glu Phe Ile Asn Ile Ala Met Ser		
	370	375
Ile Asp Pro Glu Trp Lys Met Ile Lys His Asp Gln Gly Arg Ile Glu		
385	390	395
Lys Trp Val Leu Arg Lys Ala Phe Asp Asp Glu Glu His Pro Tyr Leu		
	405	410
Pro Lys His Ile Leu Tyr Arg Gln Lys Glu Gln Phe Ser Asp Gly Val		
	420	425
Gly Tyr Ser Trp Ile Asp Gly Leu Lys Ala His Ala Glu Gln His Val		
	435	440
Thr Asp Arg Met Met Leu Asn Ala Ser His Ile Phe Pro His Asn Thr		
	450	455
Pro Thr Thr Lys Glu Ala Tyr Tyr Tyr Arg Met Ile Phe Glu Arg Phe		
465	470	475
Phe Pro Gln Asn Ser Ala Arg Leu Thr Val Pro Gly Gly Pro Ser Ile		
	485	490
Ala Cys Ser Thr Ala Lys Ala Ile Glu Trp Asp Ala Ser Trp Ser Asn		

500	505	510
Asn Leu Asp Pro Ser Gly Arg Ala Ala Ile Gly Val His Asn Ser Ala		
515	520	525
Tyr Asp Asp His Leu Pro Asp Val Ser Asn Gly Asn Leu Asp Thr Thr		
530	535	540
Ile Ile Asp Asn Val Pro Arg Met Val Gly Val Gly Ala Ser Ala Glu		
545	550	555
Leu Thr Ile Arg Ser		
565		

<210> 9

<211> 21008

<212> DNA

<213> 烟草

<220>

<221> misc_feature

<222> (18072) .. (18072)

<223> n是a、c、g或t

<220>

<221> misc_feature

<222> (19355) .. (19462)

<223> n是a、c、g或t

<220>

<221> misc_feature

<222> (19541) .. (19541)

<223> n是a、c、g或t

<220>

<221> misc_feature

<222> (19714) .. (19786)

<223> n是a、c、g或t

<220>

<221> misc_feature

<222> (20035) .. (20035)

<223> n是a、c、g或t

<400> 9

```

atgtgtggaa tactagctgt ttttggttgc attgataact ctcaggccaa gcgttcccga 60
atcatcgaac tttccagaag gttttcttctt tttattgett tccgattttc tacagcattc 120
ttattcttctt ggggtgtcaa tggatatgcatt gtgttttggga tttgctattc gtgaacttgt 180
tcgctttctc gtggtcttcg atttattttt gtcaatttat aatcagcggt ttttgatatgg 240
atttctcttt tgaatctctc aaatatatga catttgattt cctcgagaaa atttggcatt 300

```

```

tttatgggttc ttatgatatg tggttaaccc ctgctctgc atcgcttcag cgctttaaaa 360
cacgaagtga tggctattag taatactgct agagtgggta ttgaattaca ttaaataacg 420
ggttgctagt agatcaaatt cagaatgaca gtgaaatata tctggaaact tgaatccccg 480
ttcgattcgc tggtaaaacc tggctaaaaa atagctttat ctcaaaggag agtatgtcac 540
tgtgaaattg gtggatatgt ttgtttaagg catgatatgc tgctcttagg tgatgtcaag 600
atgataatac catccttttc acactagact cagctttgta gcacttacat ctgagttaat 660
cgatggagga aaatctagtc tgctttgggg catattatac caaatgtctt cttttgttt 720
tctagtgttg tttctcaaga attgaaaaac caatctacaa tatagtata taattaacta 780
gttaactctt cctgtgtttt ctgaaaattt ccaataattg agaagtgttt actggtcaat 840
tagtaaaggg agaacggaat atttttttcc tctaaagaac ttctcttttt ccatacagaa 900
tgactttcta tacaaggcct aagtctaatt tagtgtctgc gattaagcat cttctgggaa 960
cttcacactt gaaagaagta gaactaggag ggggcaggtc tagaaagata attgcttcag 1020
tcattttctc taaaaaacat tgcttcagtc attatgtgca ttcataataga gttggtcaag 1080
ctgctaagct gagctaatag tatgatagct gtgactctcg acaaagtagg ttgcctctta 1140
gggctttatc ctttggtccc gtaaagaatt tctttgggag ggctccagag tgggtaggat 1200
tgttctcatg gtaagctgta acaccaacga aacaagcctt tgtagctgtg gaagtccgag 1260
gggggtgggtg tgggtgggtg ggtgtgagag ttagagtggg ggaggagagg ggaggagtga 1320
agagaagttt atattttaaa gaatagggcc ttgagaactt ttagtagtga gcatgcatta 1380
ttattattgt tattgttatt attattatta ttattattat tattattata ttattattat 1440
tattatatta ttgttatgta attacttagc taatgagggt tttccaactg tttccaaatt 1500
aggaaatgtg ttgaaagcat agttaagtat ataaaagcgt gttctttcta atatccttag 1560
cttttagagc agataatcac aaaatttatc atgataatat agcagacaat ttttcctagt 1620
tcaaatctca tgccaccctt cataaaaaat ttccacgtgc ttggctggca tgtagggga 1680
cgtcttggca taattaaaca aatagaattg tccctttcta acagcttaag ctttttagatg 1740
aggcgggtcaa caattcaaca atttgcattt gtttgataag aagatattaa aagaacctta 1800
aaagggtgca gtatgtgaac aaactcaagg ggtagactaa ctagaagtca atttaaattt 1860
tctgcttgtc tgtgttttac ttcttttgca ctatagaaat tgcatttcca gcattccgaa 1920
agctgcttcc tttgcatata gacttcacat ttccagcgtt tctgataagc gttcttgctt 1980
attcctgtcc ttctacttgg ttgattctcc tctcacttag ttcactatct tcttgcacgc 2040
attttcaaat tgttttatca attctcttgc tttcttttca tctgaccca caaaatacct 2100
cccaaaagcc tttctgcttt tccgagtctg aaattatctg cttattactg cttttcaagt 2160
attctgatgc gttgactgtt tttttccctt gtagtaatgc atggttcgct gtccaatctc 2220
tattttgttt tttgtgaacc actattcttt gttgcttctg tttcgtcatc ctgagtccc 2280
catcttctca gtttgatttt cctcatttaa aaagcgtgtg ccccttttct tgtgtaccc 2340
tctaaaaaat cagttgctgt catctcgttt gatctgtcat gatcccagtt gatgccacat 2400
gtttagttct ttacctctgg acaactttta ctccctgaga accacgattt cttttcctga 2460
agaaattctt agaaattggg tatgtaaagg aaattcccag caatcaccat ttggcctcta 2520
cctggcttgc ttttcaactt accgacacgt gttcatcccc ttttaacaga tgctgacat 2580
tatcttctgc ttgttggggc atacttctta cccatggtgt taaccctgat ctttcgtatc 2640

```

acatttttca actcttaatt ctacatgta aagtgcctat ctttgcctct ttagttgatc 2700
 ttgggttggg ctgttaatcc aaattctctg ggctttgctt tgaagtatga taagcggtag 2760
 ctttactaca gtaccatatt ctacattgct acagacgtga aattatcctg atttgtcttt 2820
 gatgtgatgt ctggcaagtg tagattacgc cacagaggac ctgattggag tggattgcat 2880
 agccatgatg actgttatct tgctcatcaa cggttggcaa tagtagacc aacttctgga 2940
 gatcagccac tttaacaacga ggacaagacc attgttgcg cggtatggaa atatttatta 3000
 tattttttct tagatattac cggtagatgg actttggaca atatttgcg gttttgatct 3060
 tctatagtac aaacttttct cgtaggataa atttcctcta ctttgaatgg atcttaacag 3120
 gaaaaatcgga ataaacaatt tgtgatttcg aaatttttag ttcttacata ttggaacat 3180
 atgcggtaaa gtaactaggc gattgactag gtgaatacgg agaatttcct ctactgttat 3240
 caatatattac tcagctcatt tgaaagatct ctgggtcgtc tagcagtaca gctatttagg 3300
 caacttaacc aaatccccct cccatgtgtt tcaatttctt gacttatttt atattttgac 3360
 attttgcaat gccgaaaata aattgttttt ccattttatg acctttaaaa tgttgaggtg 3420
 aagggttagg cggtagctta attgacttat cagcaacaat gtattagata cagagcacia 3480
 acttgtttgg tgccctaatt tgggtatttt tgagaacgat ggagtaaatg gaaatatgaa 3540
 ttagcaaaat attgaaatag tagatctgaa aaatgaatga ctgaatttag aggaggacct 3600
 tttcttcttt ccttgcctgaa gtgggtataa gagaattatg tgaagaacat gtaagcttta 3660
 gggtagtggt acaccataac agtcatgctt tccctcaact attagtgtag ctgatgaatg 3720
 ctttcagaaa tgtggctggt tttttgctgc gtattttatg tctttcatca tctcaaaatt 3780
 tatttgtgtt taacactgct tgggaaggcct agagtcttcc gaaaatagta ggttgagaaa 3840
 acatgagtaa gcaaatttct tcttgcctggc aattaatcgt tttatatgag gtgaagccac 3900
 acaatctaac atggctcttag acataaaaaat attatttttg acatagagca gacagaaatc 3960
 ttgggtctag ttgctatctc ccttcctaatt aagatgcctt ctatcctcat tgctatatcc 4020
 caaacagat tcacactaga atgaataaat ctctccttgc agccacattt atattaataa 4080
 gtagtgaac accaaaacat ttgtgcttct gcaaagtctg ttgaagtgga cgattaagta 4140
 ataaattgtg tgagtttgac tgaagagaaa gtaataaatt gggtagctt ctacattttt 4200
 tcaggataat gtcaattcat ctgattgaac ttttgaagg acatctcctt ttctaaaagt 4260
 ttttgttcac ttcctgaaaa tggttgcttt cccaaattta ttctcctgta tttccttgt 4320
 gctttttttt tttaccagga tgactctggc ctccctttt gcaaatagtg acttcattac 4380
 cttttaacaa aagaccacaa agcctttgta tgcttgaaat cgaagcagaa tttacggctc 4440
 aatttctcta ggagatgtgt agcagtaaca taaaacgtgg ttaaagcaac ttgtaagcac 4500
 tgccactgtc caatttgaca tgactctatt actatttaga gtgtcaaata gtgtttcttt 4560
 tgactacca tttttcaatt tcttttttat acctttttt tttgactgtg acttatagta 4620
 tattttatgt agtttcagaa tatatatatt agtttttaaa gaattgatgt ctgaaagtac 4680
 agccaattag atggtttcac cctcgtactt taaatcttat aagttgtctt aggttaaagg 4740
 aataatatgc atgattacat tgtttggatt gatattcaga agatcaaatg tgccggacaa 4800
 aagcttgacg aagatataat tctgtcattt tgttgtttag ataaactacc cttgtcaaat 4860
 acctaaaact aattcatact tttcaggatc cttcttgatt ggactactct gatcttcttc 4920
 catatatgta tatcctattc tttttgcttt taggttaacg gggagatata taaccataga 4980

gacttacggg aaaagctgaa gtcccaccag tttcgaactg gcagtgactg tgaagttatt 5040
 gcacatcttg taagtacttt cagaatatat ttagataact atacatgact tctcagtttc 5100
 tgtgcgcaga atgatttcaa aataccatga ctcaaatttt tcatcaaaag cgaaaaaggt 5160
 ggattgtagt aggagaatga gattttttgt cataatttct ttcttatggg acttgaactc 5220
 cgcaatctac ctatcgcagt atgaagacta tggagaagac ttcgttcaca tgttgatgg 5280
 aatgttctcc tttgtgcttc ttgacacccg tgataaaaagt ttcattgctg ctcgggatgc 5340
 aattggcatt acaccccttt atatgggatg gggctcttgat ggtatgtaat tgttccattt 5400
 cataattttt ggtatatgtt ccttttaata aatagcccgt gagctatatg aagttccttt 5460
 gttgtttcct tgttttttat ttgactgatg tatctttgtt gttttgccta ctattagggg 5520
 tgtcaatggg tcggttcggc cggttatttt ataaaatttg taccatacca attttttaggt 5580
 tattctatta tgtataacca aaattagact ttttgaaacc gtcccaatca tgtcggtttc 5640
 tcttcgggat cggtagcggt cggttaattt tcggtatttt ttaatatcat gtaaaattca 5700
 ccagtagaag tagaattgca ataacataca ttcttttata ggacttagca aaactctcta 5760
 gatattttta ctgtttaaag ggtgatgaat taaaaaaaaa gaaaagaaag atggctagaa 5820
 tatagatcca tcaactatta tacaacagcg taaaagaaat caaacaaga caaagaaat 5880
 attaatacaca cgagttgaaa gatataccaa ggtaggactc aagaataaag tctatagaag 5940
 attaaatatt caaaaagata aatctaaatt atatgaaagg aaacatattc aattcattgt 6000
 agtttgctac tcataatcgc tagaatactt tgtgtcttgc taataaagat acttgaaata 6060
 atttagttta agtagaagta gcataatagg ttttatgaat tagtattttg agtttaatta 6120
 cttgttggtt tgtaataatt ttcataattc caaggetcaa agaaaattta ttgcattatt 6180
 atttttaaac ttactaaata aatatatttt ctacatgtaa aatttatctg gtatggttct 6240
 gtattttttc ggtttatttt tataaaaataa aaaacctacc ctaattatcg gtgcggttgt 6300
 agatttatat aaaaacctac ggtttcttaa aaagaaagct aaaaatcggt tcggtgcggt 6360
 acggttcggt cggtttagtc gattttcgaa tatccattga caccctacc tactataatg 6420
 ctggctctcat gctgttgta caaacaagt atgtataata gtgcatagta catgttttct 6480
 tgccgctgta gatcaacaat gtaccaagaa catctgtctt agctagtcca acctaatctt 6540
 cagttatcta tcattccagt gaataacaca tatagagtta ggccgacatt acctaggcc 6600
 aactatacaa cattgtgagg tgcgactaag gaactctgaa cactgatcgg gagaaaaagc 6660
 aacttatccg tgtttgtttt gaagtactat gaggaggcag caaccttgct ggatttatca 6720
 cttgacagta aatgtcaatg gttgaggaac atttttaatg accggaggag aaacagagaa 6780
 gagaagaaat aattgatttt aaaaaagcag aatagaaaat tttgtgaagt cgtatgttaa 6840
 tttcaatgaa gaggatcaac aggtgcaata tttgacttgg cacctagcat gtagaagttt 6900
 ctggttgggg ttgagcatgg taggcactgt tgcttttaca actgtcatta gagtcgaaa 6960
 catattgtat tcttgaggga gcgttgctat ggataaaaaa ggcctatcta ctacttgata 7020
 tttgggcatt cagaactagt tgtctaggag gaagaaaagg atttaaatg gtgagtattg 7080
 taggttttct ttgccgtcca ggatgcctgg gctagtcccc tgaattggga tactgggtga 7140
 ggatttcctg aaattgatca tggccctgat atgttagggg ataagaacat tgaaggatcg 7200
 aaggaagggt ccggcgctaa gttttgctag aaatgaacag ttccattatg gcttttgta 7260
 aagatgggaa tggaattggc cccaacttct aggcactctc atccatcgtc aggtgaacta 7320

aagtaagggtt acactggttag aatttttcat tatttcatat taagtaagggt gaagccgtca 7380
acaatcattg cgtgagaagg aaccaggaac aaaatctaga agatgtctac taaagtgcta 7440
atagataatc tgcaaaattht agcacggtaa ggctagtgtg gatctggttc aaatthttaa 7500
ttgggtccaaa atgcgttaat aaatgtcggg cctcgttttg gtcccccgaa aggttgctaa 7560
ataaagactt gacaaagaaa cagggtgaagg agaagcccaa acatttgaac aaactacacc 7620
tagcctaaac aaagtgattg tttgcgagga ggaaaagcga gcatgcggag agagcctcaa 7680
acacaaaaaa tggcctagtc ttgtaatgag tctagccaaa tcatgtaagc tccgactttc 7740
aagagagatc aacagaaaaac ttcgtcaatc cggataaatc aagaaaatta tgataatgga 7800
aaggaggctg gacattctac tgccgacgca acccctgggtt ttggagtcag tgagacaagt 7860
ggaactaaaag aagaatatgg atcaaaaagag tttatccagt tgacacctga tagactcact 7920
acatcagttt tcagttcaag aacatttcaa ttcttctgta taaagatagt tgccgataaa 7980
gataggcaaa ctgttaccta ccaatcaaaag gaggttgttt caattgagtt agaggtatca 8040
aaggaggagc cagtatatat agagcaagtt ttgatgaagg gagttgcctc tatttgagac 8100
caccaggtag tcaagtgtt cccgtagttc ctaggcattt tatgatggat atgatagaga 8160
aatgtcttat tgacaagaat atactagggc tacagggata atagggttgg gaaggtgagg 8220
acttcttttg ttgctgtcag gaacttgga acaaaggag aaagggttg gaatcaaagg 8280
ttttgatttc gtttttgatt tataaggaaa taatgggtag tggatggagt cgaggatagg 8340
aaagaggttt accagttgca ccattgaata tgtaattaga gacactatcc tcttgttagg 8400
gaacacgact atcacaagca cgtaagggtg agtgaaagct ccgacatgga cgtggtttat 8460
tcgggtatgt tggcacgaag ccttcgtgtc taggtttggt gttgggttaa gtgcaactcc 8520
ttaggaaggc aaataaactt gggacagatt tgatagttgt cggaagttgc acaatgatga 8580
tattttcttt atccaaacgc tacgaatgat tgcacatata gtaactagag ttattttctca 8640
tttgagcttt tatatgtaag aatcaaataa atgattgat gaaatgcatg aaagctcaaa 8700
tgagaaataa ctagttgtga gaatgcacgg gggaaaaata tattatattg attaaagtgt 8760
tgtacaacc cttttagata cagtaattac ataataatag gtatctactt cccgatgtgg 8820
gacactagac atgactaact acttaacaat cccctcaag ccggtgcata taaatcatat 8880
gtatcgagct tgttacagat gtaactaata cgagaaccag taagagactt agtgaaaata 8940
tctgctagct aatcattcga ctttacaac tttgtaaca tatctcttga gagtattttt 9000
tctctaaca agtgacagtc gatctcaatg tgtttagtc tttcatagaa caccggattt 9060
gacgcaatat gaagagcagc ttgattatca cacaccagtt ccattctgct gatttctccg 9120
aacttcaact ccttgagcaa ctgcttgatc caaactagaa tcacacgttg ccacaaccat 9180
ggcccgatat tcggcttcgg cgctagatcg agcaactaca ttctgtttct tgctcttcca 9240
agagactaaa ttacctccta ctagaacaca atatccagac gtagaacgtc aatcagcagg 9300
taatcctgcc caatcagcat ctgtgtaccc aacaatctgc tcgtggcctc gatcctcaaa 9360
tcgtaaccct ttgcctggag ctgactttat ataccgaaga atgcgaacaa ctgcatccca 9420
gtgactatca caggagaaaa ccataaacta acttgcaaca ctaccggaa aagaaatatt 9480
aggtctagtc actgtgtggt aatttaattt gccaaccaac ctctctatc tcgtaggac 9540
tctaagatgc tccccctgc catgcagaag cttagcattc ggatccatag aagtgtcaac 9600
tggtctacaa cccatcattc cagtctctc aagaatgtct aaggcatatt tccgtgtga 9660

aataacaata cctgagctag actgagcgac ctcaatacct agaaaatact tcaatctgcc 9720
cagatcccta gtctggaagt gctgaaagag atgctgcttc agattagtaa taccatcttg 9780
atcattgcca ataataataa tatcatcaac ataccactca actttggcat ttctatctca 9840
atttggctgg aatatcctgg acaatgceca ataacactgc gtctcttctc agtagctgga 9900
acaatggagg ggcaaatatt aaacagaaga aatgggtggaa agtagttcct gcttgatatct 9960
gggtgggtcat gtggaaggaa aagaatgcag agtttttgaa gataggccta gctcttagcc 10020
tcttaccaga aaattaagct taattgtttt ttgttatctt atttttggtg caaagagagc 10080
tatagggagg atgtacattc ttactagat ttgctagaag aaatgtaatc gattagaata 10140
ggcaaggcag ttgcagatct tttgggggtct gattgtggct tacactatct cctgtaaata 10200
tttttaatct tttatataat ttgttaccat tctaaaaaaa aacataaacc accagataaa 10260
tatacagatt aggagcagaa tgtcgataaa gcacagagtg atccgcttca ctacgagtca 10320
tgccgaactc ctgaataatt atgctgaact taccaaacca agctcgaggg aactgtttca 10380
aaccatatag tgacctgcgc aatcggcata caagaccact agacttcct taagcaacaa 10440
aaccaggtga ttgctccata taaatttctt cctcaagatc accgtggagg aaagcattct 10500
tagtgtctaa ctgataaaga ggtcattgac gtacaacaac catgaacaag aagagacaac 10560
cgatgctact ttagccacgg gagagagtgt cactataatc aagcccaaaa atctgagtgt 10620
atccttttgc aacaagacga gccttaagac gatcaacttg gccatccggg ctgactttga 10680
ctgcataaac ccaacgataa ccaacaatag acttacttga aggaagagga acaagcttcc 10740
aagtgccact cacatgtaaa gcagacatct cctcaatcat agtttgctgc catcctggat 10800
gagatagtgc ctcacctgta gactttggga tagaaacaat tgacaaagat gatataaaag 10860
cataatgagg tgatgacaga cgatgataac ttaaatecgc ataatagga ttaagattaa 10920
gagtggatca tacaccttc cgaagtgcaa tcggtgtact aggaagagac aagtccgcag 10980
taggagcagg gtcaagtgtg ggatgtgaat cagttgggcc tgatgctggg tgcggacgac 11040
gatgatatgt caagagtggg gttccgggtg cggggaatct aggaggagcc acactagact 11100
ccccaacggg tggaatgggt aagacttatg tggcggaagg tgaaggagga gctatagtaa 11160
gctcttttaa ggctcggtata ggtaagacct cagatatatc aaggtggtca gaagaggtaa 11220
agaaagggtt agactcaaaa aatgtgacgt cagatgacat aaagtaatta cgaaaatcaa 11280
gtgagtaaca acgatatccc tttgaacac gagaataacc aaggaagaca tacttgagag 11340
tacgaggagc taacttatct ttctagggg ctaaattatg aacgaaacaa gagctcccca 11400
aaacacgagg aggaacagag tataagggtg actgggaaac aatactgcat acggaatctg 11460
attctggatg ggagatgaag gcattcgatt aaccaaataa caagctgtga gaactgtatc 11520
gcctaaaaaa cgcaacggaa catgagattc aaagagaagt gtgcgagcag tttcaatgat 11580
gtgcctattc tttctctctg caacccatt ttgctgaggg gtataaggac aagaggtctg 11640
atgaataatt ccttgagaag tcataaagtg ctaaaattga gaggataact attctaaggc 11700
attatcactg cgaaaagtgc gaatagaaac accaaattga tttttaattt catcacaaaa 11760
attctggaat atagaaaaca actcagaacg atctttcatt aagaaaatcc aagtacatct 11820
tgaatgatca tcaatgaaac tatcaaaacta acgaaatccc aaggttgagt tgactctact 11880
atgacctat atatcagaat gaactaaaga aaaaacagac tatgcatgac tctcaatact 11940
acgaggaaag gaggctcagg tatgtttccc gagctgacat gactcacact ctaatgtaga 12000

taaactaggc actatcttct gaagcttgga tgcctaaac atctgtgaat taggtccgga 12060
 ggatctgtaa ctagacatgc cttggaggaa ttgagtgggt taaggtagta aaggccttct 12120
 aattcaagtc ctgttccaat cgtctgtccg tactgcggtc ctgcataata aaagaatcat 12180
 taataaaaata tataccacaa tggagggcac aagtcaaacg actaacagat gcaagattaa 12240
 aaggacaacc agggacataa agaacggaat ctagagtgc agagggtagg ggattcgctt 12300
 gtccaactcc ttttgcttca gtttgagacc cgttggctaa agtaacagtg gaaagagact 12360
 gtgaatacgc aatatttgac aaaagtgatt tattaccaga gatatgatca gaagcggttg 12420
 agtccacaac ccattgtcca agagtactag actgggaaac acaagcaaaa gaattaccag 12480
 caacaaaagc atcagtctga gcaacagagg ctacttgttg agatgtctgc ttacttgctc 12540
 gatactgaag gaacttatta tactccctt ccgataaaga aaatacccg ttacctatag 12600
 tctcggctctg agcaacataa gcatttttgg gtggacgacc atgtaaagaa tagcacacgt 12660
 cacgagtgtg tccaagtta tgacaataag agcacttggg tctagatctt ccaaacgac 12720
 ctctctctcg tctattctcc atagtttgag atgcccagatt gttcacttac tgggatacga 12780
 gaacagatga gtcaggcttc tgtgatgagc tcaactgggtg acttggtgct gcagcaaggc 12840
 gaagtaatcg agagaataat tcatcaactg tggggacagt cggactagcc aaaatctggt 12900
 cacgtactga atcaaggta ttagggagtc catcaagtgt aagaactaga aacatcttct 12960
 gtcgttgctc ttgttgcttt ttaatactag cagaaactgg catcaacgct tcaaattctt 13020
 ccatgattgc ctgtacttgt cccctagtaa gtagacatat ccaattcctg tttcttcaag 13080
 cttgtcattc gcgatattac atcatagaaa cgagatatgt cattagtgt taaattacga 13140
 gccctttccc aaactaaata acatgtctgg aatggacgga acaaggcat caacttgga 13200
 tcaatagatc gccacaggat actacataac tgagcatcga cettettcaa aagtgttttg 13260
 gccctttcat caccttcgct agcctttttt gttaaagtat cttgaactcc ttgacctta 13320
 caacacaact cgacagacga agcccaagct aagtagtttg aacttctcat taaagttct 13380
 gaggtaatca taatagcaga acttcagaa ctctgttttt tagaccctaa tacatccact 13440
 cccaaagaca ttattggatt gaaaagagat ctagcaaatt agcacctaat aaaacaaaga 13500
 atcaactgtg gttgccccaa aactgccgga aaactactgt agttgccgga aaattttcaa 13560
 agtgctcggg atcaaaaaat aaaaatatgg gcaggctcgg aatggtagag cgatcagact 13620
 aacctaaagg agtttttctg aaaaaattga cggaacgggc tccatgcgct ggtgcgtgga 13680
 gtagatctcg ccggagaaga ctgtctccga tcggcgcgtg gcggtgcgtg aggcggctta 13740
 tgacggaggt gttcgcgtgg gtttggtcgc cggaagttgg gggacctgt ggtggtgttg 13800
 gtttttgac aacaccgatg gaattggttt tgacgaaaaa tagccctaaa aggtcaccgg 13860
 aattgaagca cggcgacggc tggttttttt tttcccggga tgttttctca ctgccgtct 13920
 gataccatgt gagaatgcac gggagaaaaa tatattatat tgattaaagt gttgtacaac 13980
 cctatttata tacagtaatt acataataat aggtatctac tccccgatgt gggacactaa 14040
 acatgactaa ctacttaaca ctagtaacct catggttctc atcctacttc attgaccact 14100
 tggccacatc cttgggtgcc tgcactgtag ttgtagctgt aatgaagata gtttttttc 14160
 ttttctggat actcttata tgtgacaatt tttctctatt tgctgttaat ggatgaggtt 14220
 atgtgactct ccggtgggta gaattgtaat gaacatcaag ctcacatgct atttcatttt 14280
 ctgattgate ttgctcttca ctctttacga gtttaagtgt ttagtccttc caagttctgt 14340

ccatctatct ttctgggtgtg cttaacccat gtatatgtga ccacactttg tattcatgct 14400
gattgcaaca aagcctgact gatcctgaac tcctccttta tgccttaggt agaaattttg 14460
tgtgttttagt gttacatgtt tcgcttcaaa ggaaaaaggg agaaacttaa ggtagaatcc 14520
tgctccttga cctttggatc ctgacacatt gaagtggaca gacatgtttg ctttgctttg 14580
atgagttgtt ttagtagtcg gtctaaacaa ctgctgatgt attctgcttt attgtaagtg 14640
tctataattt ttatctttcc acaattttgct ttcagctctt tctgcatgta gagaaaataa 14700
cagaagcttc tatttgatca tgtaatatgt actgcgcaag ttcagtttta aacaactttg 14760
gctatgcata tgcaagcaat gacctgaaat aaatattatt ttttatgcag ggtctgtatg 14820
gttttcttca gagatgaaag ccttaagtga tgactgtgaa cgatttgta gtttccttcc 14880
cggtcacatt tattcaagca aaaatggtat gccatacga agctcatttc tgagcttgtg 14940
cttcatttat gttgatatgt gatcatgttt cctatacga cagcgtactt ttgctattta 15000
tttcagattt actttaggac cctgtcggaa acagtcttga aaccttatt gatggtgtaa 15060
ttgttatctc ttgatcaacc aaaggagtaa ttacatgtaa ttattacaaa tttgcaataa 15120
aataactttt aggataacca gtggagctta atgaatcttt taggatattt gggatgctag 15180
agtgatgtac ttttgataat gtcggttaat ataaaatcct ctgactttgt actgatgtac 15240
tggtgagca gattttatat tactttctag tccctctgca aataaaaaaa ttgtgagact 15300
gcaatataac accatttttt aatgcattta caatgtttta acttgatgta gcagtgaaga 15360
acatgtcgac cttaaacaga ggttctattt gtccatgata aaggaacatc agcgtaatga 15420
aacattacag gagagaagct ctattgagtt tttgttaatc tagagggagc ccgtaaatag 15480
aacaccaatt catctgttat atgacttctt tttcattttg gaacgtttta aaagtttgac 15540
aacatttttt gatctccgta cactaaactt tttgatgtat taaactaact atttccacta 15600
ctttaatatt ttcttcctta aacaactcat gttagcatct taaactgtgt gtgttcaata 15660
ccctcataaa aaattaaacg agtatttttt tatatttgc t atgcccttg gagattggga 15720
tggggaagtt catcatttga ggtgaaaggt aaattccttt ggatggacaa aaagtgaacc 15780
tcgtcttttg catggaggga ctatttaaat cgtcatcatt accaaagtca tcctttcaaa 15840
tgtttacttt caagaactat tcatgctatt tctcaagtat attatgggtg caggaggact 15900
cagaagatgg tacaaccac catggtactc agaaaccatt ccttctactc catatgatca 15960
ccttgtctta cggaaagctt ttgagaaggt cattctattt tggatgtcgc ttgtttctga 16020
gttgtacata atatttaggt gccactttgt tatattttga acttcatcaa aagccttttg 16080
ttttttgagg acacatatcc ccttgactga ctgggcgctg aatttccttg atgtattgtg 16140
cttattttga ttgtcctctt ctttgtaata ctcttgecta taaataaacc ttctttatca 16200
aaagaaaatg ataaaatgtt aataaacactt tgtgatatat gttgaagggg agtcttggca 16260
taactggtaa agttgctgcc atgtgaccag gaggtcacgg gttcaagccg tggaaacagc 16320
ctcttcgaga aatgtagggg aaggetgcat acaatcgacc cttgttggtc gaccttccc 16380
cggacccgc gcataacggg agcttagtgc accgggctgt cctttttttt tgcatagact 16440
ttacatgata ttttatttca atctttatca attctttttt accttcacat gcaattata 16500
ctaccgatg tgaaatatgc aatagttttt cccgcaattt tcatctttca tgtacctctc 16560
tctactatcc ttgtattact actgatttgg tgcattagat gtttaatttta ttcccaacct 16620
ttactttatg caggctgtag tcaagcgact tatgacggat gtaccatttg gtgtgcttct 16680

ctcagggtgga ctagattctt cacttggtgc tgcagtggcc aaccgttatt tggctgatac 16740
 tgaagctgcg cgacaatggg gatcacagtt gcataccttt tgcgtaggct tgaaggtggt 16800
 gatgctccct tttgtaaact tgctctgatt agggaattgc tagctgtatc atcccaattc 16860
 acatggaagt tctagaccat tattgtctgtt tcttttccat atatagggat tttctgaaaa 16920
 tagtaaaaag gaaatggaat ttctcattat taaagatacc aaatgccttt ttgggaaatg 16980
 atcttacagt tgaatctgaa atttcctttt gacattgtga ttaattctga tatgaatttg 17040
 gtccatcctg gctttccttc tcttgattgt agattttttt tctctttttg caattatatg 17100
 cctttgtatt tcatactctt tattatgcta ttattttatt atgcattttt atggtactaa 17160
 tatatcggct cctgttgctt tttttgagcc gagattttgt ttattttctgc tatcacgtcc 17220
 tcctgaaaaa gtgctactat cttgtattca tgcgggagtt tgaaatttgc aagtgtataa 17280
 ctttacaaaa tgattcgaac tgatcacagt tcattgctaa taaactttta taaaaagcag 17340
 gtattgttgt gatataatgt gacacttaaa taaatttaaa cctacatcct ttcgaaatct 17400
 ctaatagctt cactaatatt tcttgtgcta aattcttcca acacctctta tcgatccaa 17460
 atttggcaga ccttagtgca attgaaagtg tttagtgttc tttatgcagt aggtatgtta 17520
 ttgtccaaaa tgacagaaat aaagttagag atttgttgaa aagccacttg gtgtgattga 17580
 gcagttcgga gatgagtgat ttatccatct ttttatactt ttagtgtttt gttttctttt 17640
 atattgtttt ctcttttgc tttgtcattg taattaaggt tgtcttttgt ttgactttag 17700
 ggttctcctg atctgaaagc tgccagagag gtggcagact accttggaac tcgtcaccat 17760
 gagtttctact ttactgtgca ggtaactttc ttcataaaga catcctatgc tgattttcct 17820
 tataggagta taactgctta agtcaaattg gctatctact tctttctatt gctacatcag 17880
 gaaggaattg atgcactaga tgaagtcatt tatcatgttg aaacatatga tgtgaccact 17940
 atcagagcca gtacaccaat gtttctcatg tctcggaaga taaagtcctt ggggtgtgaaa 18000
 atggttctat ctgggtgaagg ttctgatgaa atttttggcg gttatttata tttccacaag 18060
 gcaccaaca antttacaca aaaatgacac aagttatgca gaaaaaatag tcctcaatgt 18120
 catgggagac cagggttcaa atctgagggt aaaaaaata ctaggtgaac tcttctaate 18180
 cgtttaagcc ttggtagata gagttaaacta gtatatgagc tagttagagc tagcaggtac 18240
 ctggtagatt agtgcaggtg cacacgagct gactcaaaac caccattatt aaaaaataag 18300
 aaagttttca atgttgttcc accgtgaatt catgatgtgc cgactttcct gtgtttatct 18360
 ctggatagat gaagagtgtt tttggatctt gattaattct gtttggeccc tttgaggttc 18420
 tgcaagctat ttctagactc aaacttgctg cgatctatgt taggtctaga gaatatctgc 18480
 caaataaaac aatggaaggt aactcaagt attttatctc catattacaa gtcagtgtta 18540
 gttgaatggg tctgatatat gggacagaaa atggttcat acccataacc tcattaaaat 18600
 gttgctgatg tgctgatctc atctgtacat ctagtataat cctaattaca aattaatgca 18660
 tctctttgcg gtgactatag tatgtgaag aacttttctt gaaacattgc agatcagacc 18720
 tgatctcgga agaatagaaa agtgggttct acgcaatgct tttgacgatg atcagaatcc 18780
 ttatctgcca aaggtttggt tgttttcttg tatcaggtcc aaaatcaaaa tctttgttta 18840
 ttcatattgca taaacaggca gtagactccc aatacattgt gtacacttat gcacaaatgt 18900
 atgagatgtg tattatatgt gtcacatata agggattttc cacatggtgt ttattgatct 18960
 agagaaagcg tacgacaggg ttcctagga ggtcctatgg agatgcctag aggttaaagg 19020

```

ggtcccggtt gactgcatta ggggtgattaa agacatgtat gatggagcta agactcgaat 19080
taggacagta ggaggcaact ccgaccattt tccggttggt acgggggttg atcaagggtc 19140
tgcgttcaac cctttcctat ttgccttggt gatggatgct ataaccgcat catattcaag 19200
gggaggtgcc atggtgcatg ttatttgctg atgacatagt cctaattgac gagacacgac 19260
gcggcgtcaa cgagaggcta gaggtttgga gacatgccct tgagtctaaa ggtttcaggc 19320
tgagcaggac gaagacggaa tatctcgagt gcaannnnnn nnnnnnnnnn nnnnnnnnnn 19380
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 19440
nnnnnnnnnn nnnnnnnnnn nnatcgtata tgggtagggg ggatgaagtg gaggttagca 19500
acgggagttc tgtgtgacaa gaaagtgtca ccgttactaa ntctcgagtg caactttggg 19560
gccgagccga tggaagcatg agtggaagtg aggtctgact cacaagtcac ccctaagagg 19620
ggtagtttca agtaccttgg gtcagttatt caggggaccg gggagatcga cgaggatgtc 19680
acacatcgta taggggtggg gtggatgaag tggnnnnnnn nnnnnnnnnn nnnnnnnnnn 19740
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnatag gggtaggggtg 19800
gatgaagtgg aggttagcaa cgggagtctt gtgtgacaag aaagtgtcac cgttactaaa 19860
aggtaagttt tacagagcag tggttaggcc tgctatgttg catgggaccg agtgttggcc 19920
ggttaaaaat tcacacatct agaagatgaa agtagcagag atgaggatgc tgaggtggat 19980
gtgcgggcat acaaggatgg acaagattag gaatgaagtt attcgagaga aggtnttgag 20040
ccgagggctc cctggaaaca gcctctctgc ccctcgggtt aggggtaagg tctgcgtaca 20100
tactaccctc ccagacccc attggtggga ttatactggg atgttgttgt tgttgttgtt 20160
gttgtgttac atataagggt attatatgta cacactctc ccacettca cttatatatt 20220
cacgtaaaac acatgcatgt gtaaaaccta tatatgtgtg tgtgatatat ataataaaat 20280
gatgtttgaa aagacattat gatcaagtag ttgatcttca tctgctttcc gtttttagca 20340
tatcttgtat aggcagaagg aacagttcag tgatggagtt ggctacagtt ggattgatgg 20400
tttgaaggat cacgcaagca gactggtgag ctttttagtt gttttctctt ccctcctctt 20460
ccattcttta tcccttcctt tgtggttttg ttgaacacat ctcgtatatt tgggccagg 20520
ttctgattct atgttagcga atgcaagttt cgtttaccg cataacacac ccacgacaaa 20580
ggaaggatac tactatagaa ctatttttga gcgatatttc cccaagggtg gtcttagata 20640
gtcttctttt ccttgttatt catgttcaat tttgatatt ttgtcccaa caatgttttc 20700
tttttcttcc ttctgttaac aaggtcgtt ctatcttcat atttgttgat aaaggtaatc 20760
acattgataa ctattatttg gagcagaatg ctgcgaggga aacagttcca ggtggtccaa 20820
gtgtggcatg cagcactgca aaagcagtag aatgggacgc agcttgggtc agaactctag 20880
atccatctgg acgagctgca ctcggtgttc atgcagctgc ttatgaggat gcatcagagg 20940
ttaaaccac agtcgcaaca gataccctc agaaacttga agttgataaa gctgcagtag 21000
ctgtttga 21008

```

<210> 10

<211> 581

<212> PRT

<213> 烟草

<400> 10

Met	Cys	Gly	Ile	Leu	Ala	Val	Phe	Gly	Cys	Ile	Asp	Asn	Ser	Gln	Ala
1				5					10					15	
Lys	Arg	Ser	Arg	Ile	Ile	Glu	Leu	Ser	Arg	Arg	Leu	Arg	His	Arg	Gly
			20						25					30	
Pro	Asp	Trp	Ser	Gly	Leu	His	Ser	His	Asp	Asp	Cys	Tyr	Leu	Ala	His
			35						40					45	
Gln	Arg	Leu	Ala	Ile	Val	Asp	Pro	Thr	Ser	Gly	Asp	Gln	Pro	Leu	Tyr
			50						55					60	
Asn	Glu	Asp	Lys	Thr	Ile	Val	Val	Ala	Val	Asn	Gly	Glu	Ile	Tyr	Asn
65					70					75					80
His	Arg	Asp	Leu	Arg	Glu	Lys	Leu	Lys	Ser	His	Gln	Phe	Arg	Thr	Gly
					85					90					95
Ser	Asp	Cys	Glu	Val	Ile	Ala	His	Leu	Tyr	Glu	Asp	Tyr	Gly	Glu	Asp
			100						105					110	
Phe	Val	His	Met	Leu	Asp	Gly	Met	Phe	Ser	Phe	Val	Leu	Leu	Asp	Thr
			115						120					125	
Arg	Asp	Lys	Ser	Phe	Ile	Ala	Ala	Arg	Asp	Ala	Ile	Gly	Ile	Thr	Pro
			130						135					140	
Leu	Tyr	Met	Gly	Trp	Gly	Leu	Asp	Gly	Ser	Val	Trp	Phe	Ser	Ser	Glu
145					150					155					160
Met	Lys	Ala	Leu	Ser	Asp	Asp	Cys	Glu	Arg	Phe	Val	Ser	Phe	Leu	Pro
					165					170					175
Gly	His	Ile	Tyr	Ser	Ser	Lys	Asn	Gly	Gly	Leu	Arg	Arg	Trp	Tyr	Asn
			180						185					190	
Pro	Pro	Trp	Tyr	Ser	Glu	Thr	Ile	Pro	Ser	Thr	Pro	Tyr	Asp	His	Leu
			195						200					205	
Val	Leu	Arg	Lys	Ala	Phe	Glu	Lys	Ala	Val	Val	Lys	Arg	Leu	Met	Thr
			210						215					220	
Asp	Val	Pro	Phe	Gly	Val	Leu	Leu	Ser	Gly	Gly	Leu	Asp	Ser	Ser	Leu
225					230					235					240
Val	Ala	Ala	Val	Ala	Asn	Arg	Tyr	Leu	Ala	Asp	Thr	Glu	Ala	Ala	Arg
					245					250					255
Gln	Trp	Gly	Ser	Gln	Leu	His	Thr	Phe	Cys	Val	Gly	Leu	Lys	Gly	Ser
			260						265					270	
Pro	Asp	Leu	Lys	Ala	Ala	Arg	Glu	Val	Ala	Asp	Tyr	Leu	Gly	Thr	Arg
			275						280					285	
His	His	Glu	Phe	His	Phe	Thr	Val	Gln	Glu	Gly	Ile	Asp	Ala	Leu	Asp
			290						295					300	
Glu	Val	Ile	Tyr	His	Val	Glu	Thr	Tyr	Asp	Val	Thr	Thr	Ile	Arg	Ala

305	310	315	320
Ser Thr Pro Met Phe	Leu Met Ser Arg Lys	Ile Lys Ser Leu Gly	Val
	325	330	335
Lys Met Val Leu Ser	Gly Glu Gly Ser Asp	Glu Ile Phe Gly	Gly Tyr
	340	345	350
Leu Tyr Phe His Lys	Ala Pro Asn Lys Glu	Glu Phe His Gln	Glu Thr
	355	360	365
Cys Arg Lys Ile Lys	Ala Leu His Leu Tyr	Asp Cys Leu Arg	Ala Asn
	370	375	380
Lys Ser Thr Ser Ala	Trp Gly Val Glu Ala	Arg Val Pro Phe	Leu Asp
385	390	395	400
Lys Glu Phe Val Asn	Val Ala Met Asn Met	Asp Pro Glu Ser	Lys Met
	405	410	415
Ile Arg Pro Asp Leu	Gly Arg Ile Glu Lys	Trp Val Leu Arg	Asn Ala
	420	425	430
Phe Asp Asp Asp Gln	Asn Pro Tyr Leu Pro	Lys His Ile Leu	Tyr Arg
	435	440	445
Gln Lys Glu Gln Phe	Ser Asp Gly Val Gly	Tyr Ser Trp Ile	Asp Gly
	450	455	460
Leu Lys Asp His Ala	Ser Arg Leu Val Ser	Asp Ser Met Leu	Ala Asn
465	470	475	480
Ala Ser Phe Val Tyr	Pro His Asn Thr Pro	Thr Thr Lys Glu	Gly Tyr
	485	490	495
Tyr Tyr Arg Thr Ile	Phe Glu Arg Tyr Phe	Pro Lys Asn Ala	Ala Arg
	500	505	510
Glu Thr Val Pro Gly	Gly Pro Ser Val Ala	Cys Ser Thr Ala	Lys Ala
	515	520	525
Val Glu Trp Asp Ala	Ala Trp Ser Lys Asn	Leu Asp Pro Ser	Gly Arg
	530	535	540
Ala Ala Leu Gly Val	His Ala Ala Ala Tyr	Glu Asp Ala Ser	Glu Val
545	550	555	560
Lys Thr Thr Val Ala	Thr Asp Thr Pro Gln	Lys Leu Glu Val	Asp Lys
	565	570	575
Ala Ala Val Ala Val			
	580		

<210> 11

<211> 17971

<212> DNA

<213> 烟草

<220>

<221> misc_feature

<222> (1389) .. (1389)

<223> n是a、c、g或t

<400> 11

```

atgtgtggaa tactagcagt tttcggttgc attgataatt ctcaggccaa gcgttcccca 60
atcatcgaac tttctagaag gtagaatttt tttaaaaaaa aaattgcatt ccaatttttc 120
tacagcattc ttattctttg gggtgacaat ggtatgcatg tgttttggat ttgctattcg 180
tggacttggt cgctttctcg tggctcttca tttgtttttg ttaatttata atcagtgttt 240
ttatatggat ttcttttttg aatctctcaa atatatgaca ttgatttcc acgagaaaat 300
ttggcatttt tatggttctt atgatatgcg gttaaccctt cgcttgcac cgcttcagcg 360
ctttaaaaga cgaaatgatg gctattagta atactgctag agtggttatt gaattacatt 420
aaataacaga ttgctagtag atcaaattca gaatgacagt gaaatatctc tggaaacttg 480
aatccccgtt caattcgctg gtaaaacctg gcaaaaaaat agctttatct caaaggagag 540
tatgtcactg tgaaattggg ggtatgtgtt gtttaagctt taaggcatga tatgctgctc 600
ttaggtgatg tcaggatgat aataccatcc ttttcacact agactcagct ttgaagcact 660
ttcatctcag ttaatcgggt gaggaaaatc tagtttgctt tgggcataat ataccaaag 720
tcttcctttt gttttctagt gttgtttctc aagaattgaa aaaccaacct acaatatagt 780
gatataatta actagttagc tcttcctgtt tttctgaaaa tttccaataa ttgagaagcg 840
tttacaggtc aattagtaaa gggagaacgg aacatttttt tcctctaaag aacttctctt 900
tttcataca gaatgacttt ctatacaagg cctaagtctg ctttagtgct tgcgattaag 960
catcttctgg gaaacttcac acttgaaaga agtagaacta ggagggggca ggtctagcaa 1020
gataattgct tcagtcattt ctcataaaaa aacattgttt cattcattat gtgcattcat 1080
atagagttgg ttaagctgct aagctgagct aatagtatga tagctatgaa tctcgacaaa 1140
gtaggttgct tcttagggct ttattctttg gctccgtaaa gaatttcttt tggaagggtc 1200
tcagagtggg taggactggt ctgatgataa gctgtaacac caacgaaaca agcctttgta 1260
gctgtggaag tccgaggggg tgggtgtgtt ggtggtgtgt tgagagttag agtggtggag 1320
gagaggggag gagtgaagag aagtttatat tttaaagaat agggccttga gaacttttag 1380
tagtgagcng gggggggggg ggggcagagg ggaggagtga agagaagtat atattttaaa 1440
gaatagggcc ttgagaactt atagtagtaa gcatgcatta ttattattgt tgttattatt 1500
attattatta ttatttttat tattatgtaa gatgtaatta cttagctaat gaggtttttc 1560
caactatttc caaattgtga aatgtgttga aagcataatt aagtatataa aagcgtgttc 1620
tttctaatat ctttagcttt aagaggagat aatcacaaaa tttatcttga taatatagca 1680
gacaattttt ctagttcaa atctcacgtc acccttcata aaaaatttcc acgtgcttgg 1740
cttgcattgt aggggacgtc ttgacataat taaacaaata taattgtccc tctctaacag 1800
cttaagcttt tagatgagc ggtcaacaat tcaacaattt gcatttgttt gataagaagt 1860
tatattaaaa gaacacttaa aaggtgtcag tatgtgcaca aactcaaggt gtagactaac 1920
tagaagtcaa ttcaaatttt ctgcttgttc tctctgtttt acttcttttg cactatagaa 1980
attgcatttc cagcattccc gaaagctgct tcctttgcat atagacatca catttccagc 2040

```

```

gtgtctgata agcgttcttg cttattcctg tccctctact tggttgattt tcctctcact 2100
tagttcactg tcttcttgca cgcattatca aattgtgttc tcaattctct tgccttcttt 2160
tcactcgacc ccacaaaata cctcccaaaa agcctttagt ttctgctttt ccgagtcctga 2220
aattatctgc ttattactgc ctttcaagta ttctgatacg ttgaccattg ttttccccag 2280
tagtaatgca ttgttctctg tccaatctct attttttttt ttgtgaacca ctattctttg 2340
ttgcttctgt ttcgtcatcc tcagttcccc atcttctcaa tttgattttc ctcatttaaa 2400
aagcgtgtgc cccttttctt gtgctaccct ctcaaaaatc agttgctgtc atctcgtttg 2460
atctgtcatg aaccagttg atgccacatg tttagttctt tacctctgga caacttttac 2520
ttcctgagaa ccacaatttc ctttcttaa gaaattctta gaaattgggt atgtaactg 2580
taaagggaaa atttcccagc aatcaccatt tggectctac gtggcttgct tttcaactta 2640
ccgacacgtg ttcacccctt tttacaaaat gctgaccatt atcttctgct tgttggggca 2700
tgcttcttac ccatgggtgt aacctgate tttcgtatca ctttttcaa ctcttaattc 2760
tcactttata aagtgcata ccttgctctt ttagttgate ttgggttggg ctgttaatcc 2820
aaattccatg ggctttgctt tgaagtatga taagcattag ctttactaca gtagcatatt 2880
ctacgttgct acagacatga aattttcctg atttgtcttt gatgtgatgt ctggcaagtg 2940
tagattacgc cacagaggac ctgactggag tggattgcat agccatgatg actgttatct 3000
tgctcatcaa cggttggcaa tagtagacc aacttctgga gatcagccac tttacaatga 3060
ggacaagacc attgttgtcg cggtagggaa atatttatta tatcatttct tagatagtta 3120
agtattttac cagtacatgg actttggaca atatttgtgc gttttgatct tctatagcac 3180
aaacttttct cgtaggataa atttctcta ctttgagtga atcttaacag gaaaatcgga 3240
ataaacaatt tgtgatttcg aaatttttag ttcttacata tttgaaacat gtgcggtgaa 3300
gtaactagge gcttgactag gtgaatacgg agaatttct ctactgttat cgatattttc 3360
tcagttcatt tgaaagatct ctgggtcgtc tagcagtaca actatttagg caacttaatc 3420
aaatccccct cccatgagtt tcactttctt gacttgtttt atattttgac attttgcaat 3480
gccgaaaata aattgttttt ccattttatg accttaaaa tgttgagggtg aagggttagg 3540
cggctactta atcgacttat cggcaacaat gtattagata cagagcaca acttgtttgg 3600
gtgcctaaat tgatattttt gaaaagtcga tggagtaaat ggaaatatgg attagcaaaa 3660
atattgaaat agtagatcta aacaatggat gactgaattt agaggaggac cttttcttct 3720
ttccttggtg aagtgggtat aagagaatta ttggaagaaa atgtaagctt tagggtagtg 3780
ttgcaccata acagtcatgc tttccctcaa ttattagtgt agctgatgaa tgcttttggg 3840
aagggtggctg attttttgct gcgtatttta tgtctttcat catctcaaaa tttgttgtgt 3900
tttaacagtg cttggaagge cttagtctt ctgaaaaaag taggctgaga aaacatgagt 3960
aagcaaatth cttcctgctg gcaattaatc tttgtatatg aggtgaagcc acataatcta 4020
acatgggtatt agacataata atattatttt tgacatagag cagacagaaa tcttgggtct 4080
agttgctatc tcccttccta ataagatgcc ctcatgtcta tatcccaaaa cagattcacc 4140
ctagaacgaa taaatctctc cttgcagcca catatatatt aataagtagt gtaacactaa 4200
aacatttgtg cttctgcaaa gtctgttgaa gtgggcgatt aagtaataaa ttgtgtgagt 4260
ttgactgaag agaaagtaat aaattgtgtg agcttctcac attttcatga taatgtcaat 4320
tcactcgatt gaactttttg aaggacatct ctttttctaa aagtttttgt tcacttctg 4380

```

```

aaaatggttg ctttcccaaa tgtattctcc tgtattttcc ttgtgcttta tgttgaccag 4440
gatgactctg gccttccctt tttcaaatgg tgacttcatt accttttaac aaaagatcaa 4500
atagcctttg tatgcttgaa atcgaagcag aatttagggc tcaatttctc taggagatgt 4560
gtagcagtaa cataacacgt ggtaaagca acttttaagc actgccactg tcaaatttga 4620
tatgacgcta ttactattta gagtgtcaaa tagtatttct ttgactacc aatttttcaa 4680
tatctttttt ataccttttt tctttgactg tgacttatag tatattttat gtagtttcag 4740
aatatatatt ttagttttta aagaattgat gtctgaaagt acagccaatt agatggtttc 4800
acccttgtag ttcaaatctt ataagttggt taggttaaag gaataatatg catgattaca 4860
ttgtttggat tgaaattcag aagatcaaat gtgcaggcaa aaagcttgcg gaagatataa 4920
ttctgtcatt ttgttgttta gattaacttc cctgtcaaa tagctaaaac taattcatac 4980
tttctcagga tccttcttga ttggactact ctgatcttct tccatatatt gatatatcat 5040
attccttttg cttttagggt aacggggaga tatataacca tagagaatta cgggaaaagc 5100
tgaagtccca ccagtttcga actggcagtg actgtgaagt tattgcacat cttgtaagta 5160
ctttcagaat atattcaaat aaatatacat gacttctcag tttctgtgcg aagaatgatt 5220
tcaaaatacc atgactcaaa tttctcatca aaagagaaaa aggtggattg tagtaagaga 5280
atgagatttt ttgtcataat ttctttctta tggattttga actccacgat ctacctatcg 5340
cagtatgaag actttggaga agacttcggt cacatgttgg atggaatggt ctcctttgtg 5400
cttcttaaca cccgtgataa aagtttcatt gctgctcggg atgcaattgg cattacgccc 5460
ctttatatgg gatgggggtct tgatggtatg taatttttcc atttcataat ttttggtata 5520
ttgccctttt aataaatagc ctgtgagtta tgtgaagttc ctttgttggt tccttgtttt 5580
ttatttgcct gatgcatctt tgttgttttg cctactataa tgctggtttc atgctgttgt 5640
tacaacaaaa tgatgtataa tagtgaacag tacatgtttt cttgccgctg tagatcaaca 5700
atgtaccaag aacatctgtc ttaagctagt tcaacctaatt cttcagttat ctatcattcc 5760
agcgaataac acatgtagag ttaaccggac attacctaag gccagctata caacattgtg 5820
aggtgtgact aaggaactct gaacactggt tggagaaatg gaaagcaaca tatccgtggt 5880
tgttttgaag tactataagg aggcagcaac cttgctggat ttatcacttg acagcaaatg 5940
tcaatgggtg aggaacattt tttatggctg gaggagaaac agagaagaga aaaaataatt 6000
gatttttaaaa aacagaggag aaaattttgt gaagtcgtat gttaatttca atggctggag 6060
gtgcaatatt tgacttggca actagcgtgt agaagtttct ggttggggtt gagcatggaa 6120
gacactgttg cttttacaac tgtcattaga gtcggaaaca tattgtattc ttgagggagc 6180
gttgcaatgg ataaaaaagt cctacctact acttgatatt tgggcattca gaactgggtg 6240
tctaaagctg agaaggaaga aaaggattta aaatggtgag tattgtaggc tttcttttct 6300
gtccagggtg cctggactag ttccctgaat tgggatactg gtggaggatt tcctgaaatt 6360
gatcatggcc aaagagatat gtcaggggat aagaacattg aaggatcaaa ggaaggtgct 6420
gcgctaagtt ttgctagaaa tgaacagttc cactatggct tttgtcaaag atgggaattg 6480
aattggcccc aacttctggg catctccatc catcgtcagg tgaacaaaag taaggttaca 6540
atggtagaat ttttcattat ttcaaacaag taaggtgaag tcgtcagcaa tcattgctg 6600
agaaggaacc aggaacaaaa tctagaagat gtctgctaaa gtgctaatag ataatctgcc 6660
aaatttagca cggtaaaggct agtgcagatc tggttcaaat tttaaattgg tccaaaatgc 6720

```

```

tttaataaat gtcggacctc gttttgttcc cccaaaaggt tgctatataa agacttgaca 6780
aagaaacagg tgaaggagaa gcccaaacia ttgaacaaac tacacctagc ctaaagaaag 6840
tggttgatag tttggctcga ggaggaagag cgagcctgcg gagagagcct caaacataaa 6900
aaatgacctc gtcttgtaat gagtctggcc aaatcaagta agctctgact ttcaagagag 6960
atcaacagaa aacttcgtca gtctagataa atcaagaaaa ttatgataat gaaaaggagg 7020
ctggacattc tactgccgac caaccctggt tttggtgtca gtgagacaag tggaactaaa 7080
gaagaatatg gatcaaaaaga gtttatccgg ttgacacctg ataaactcac tacatcagtt 7140
ttcagttcaa aaacatttca attcttgcgt ataaagacag gcaaactgtc acctaccaat 7200
caaaggaggt tgtttcaatt gagtttagagg tatcaaagga ggcggcagta tatatggagc 7260
aagttttgat gaagggagtt gcttctattg gagaccacca ggtactcaag tgcttcccgt 7320
agttcctagg cattttataa ttgatatgac agagaaatgt cttattgaca agaatagact 7380
aggactacag ggataatagt gttgggaagg tgaggacttc tttggttgct gtcaggaaact 7440
tggaacaaaa gggagaaaagg ggttggaaac aaaggttttg atttcatttt tgatttatga 7500
ggaaataatg ggtagtggat ggagtcgagg atgggaaaga ggtttacca ttgcaccatt 7560
gaatatgtaa ttagagacac tatcctctta tttgggaaca cgactagggt gagggaacgc 7620
tccgacaggg atgtggttta ttcgggtatg ttggcacgaa gcttttgtgt ctaggtttgg 7680
tgttggctta agtgcaacac cttaggaaga caaacaact taggacagat ttgatagttg 7740
tcggaaattg cacagtgatg atattttctt cttccaaacg ttgggaatga ttgcacatgt 7800
agtaactaga gttatttctc atttaagctt ttatatgtaa gaatcaaata aaataattga 7860
tgaaatgcat gaaagcaact gtcttacgcc tgaatgcttt taaaaggcat gctttacaaa 7920
gtttttggta catctatctg acatgacaca ctaacattca ctcttatatg tgtacatatt 7980
taactaagtc caagcctaac tatctaactg gagatatcaa agttcttagc tctttaaagc 8040
tttcttaact ttaggaagat cctgcagtgt tgttttttta taaccgtggt atccaggcta 8100
gcttgcacac atctcgacta attccaccgg gtgcttgcta tctcccacca acacagatta 8160
cctggtaact catgggaaga aatcacctat tgttttgagt ctgctggcat ttgaagtttt 8220
ggactactaa cctcatggtt ctcacccac ttcatgacc acttgccac atccttggtt 8280
gcctgcactg tagctgtaac gaagatagtt ttttttccg ggatactctt atattgtgat 8340
aatttttttg tattcgctgt taatggatga ggttatgtga ctctccggtg gtgagagtgt 8400
taatgaacat caagctcata tgtcatttca tttcttgatt gatcttgctc ttcactcttt 8460
acgagtttaa tgtgttagtc cttcccagtt ttttttaaga gcgagaagcg caaaaaagcg 8520
acaagggctc gcctcgcttc aaaagcgaag cgcaaagcga agcgcacact ttattaaagt 8580
gaagcgcaat tcttaaaaaa cataatgtaa accttgcaaa gacacaatat aaaattataa 8640
ataatcaaaa agttcaaatt gtcaaaaatca aagctactaa attactagaa tcaacctctt 8700
attcttctac tcttcttggt cttcttcaag attgtcaaaa tcttgaaatc ctctattatc 8760
ttcatattgc tcgtcatctt cctcctcccc ctctcttctc tcttcatgat cactttcttc 8820
ttcatcaatt agggataggg atcgacttgt agtagccaca ctttttccct tctaatcga 8880
gcttgaactt gaagtattcc cccttaaacc ataaggattc tccccaactc cactagcaac 8940
cgcaacatca ccccaagtga aattagaatc gccttcaaat acttcttcat cttcacaatt 9000
ttcggggact ccggttagcc actcattagc ctcatcaata ttgtccaaaa gaattggatc 9060

```

aattagattg cgggtggttgt agcgacgcct caatgctcta ttgtatttta tgaacactag 9120
attatggagg cgcttcaagg ttagtcgatt cctcttcttt gtatgaatct gcaaacaaga 9180
tgtgtcaact aattagtagg agttgggaca attgagatat aatttacttt atctcaaaac 9240
acgaaataat tctttttatt tctcacgtgt tcaaaaacgc tccagttcct ttcacatccg 9300
gatgagctac aagttaaact tagaactctg atggcgaaag tctgtaaatt cggagtctct 9360
acaccatatt ggtcccacca ctcaactata gaagtgaana aaaaaattc aagagttaat 9420
aagtataaaa aatacattaa agttagagct ataaaaatata gaagcacttg gtcacctggc 9480
gatttcgtct tcttttgttt aatagcaagt cggagcttaa aaagtccctc agctgccttg 9540
taaatagcaa gctgatctac tatcttttct tgcatgtctt catctggggg caacttgata 9600
acaacctcat ggaatcctgt ccacacttct ctagecaatg aattattctc atgctgatca 9660
taaaagagtg acgggttcag aataagtcca gctgcatgca aaggtctatg aagttgctca 9720
ctccatcttg catcaatgat ctgaaagacc tttgcatatt tctgctcacc agtgaatgat 9780
gcttgaatag cctccttggc cctatccata gcttcataga ggtagcccat tgggtggtttt 9840
tgcctcccat ccaccaaagc gagtacttta accaaagggc caccaatttt aagagcatga 9900
aggacattat tcaaaaaaga ataagaaaga ataatgcgtg caacatcttt cctgcactt 9960
tcctttgcaa atttactctt gctccattcc tctgaagtga acaactttct caaattggat 10020
ttttgcaagt ggatactatg taaagtcaag aaagcagtgg cgaaccttgt cttgcccggt 10080
ttcaccaaat ttttttgtcc ggtgaatctt ctcatcatat tcaataacaa gggccgctga 10140
gaaatataag aatgtaccct aacgccctgg ccaaaaactg tagagaaggg ttttcccttg 10200
aaaatgtccc cgaagattaa gttgatacaa tgagccgcac atggagtcca atagacattc 10260
ttgtacgctc ctccaccat gccaccgct ttcacatttt cactcgcatt atcagtgacc 10320
acttgaacaa ctttgcttgg gccaatcttt tcaatggtgt tctgaaacaa ggtgaacatt 10380
ttgatgtggg cagtggatga gtcgctagca tcaatggact caagaaacaa acttcccttg 10440
ggagaattca ccaacacggt aataatcatt ttcccagttc ttgccgtcca cttatccatc 10500
ataatggagc agccatactt gttccacgca actttatgtt cctccacaat tttattagtc 10560
tcctccactt ccttatttag ataaggacct ctgatttcat gataagtggg aggcttcatt 10620
ccagggtccgt attgaccaac ggcctcaata aagtctccaa aagtgtcagt atagttgaca 10680
caattgaagg ggagcccagc atcatagacc catcgtgcaa aagctctaac tgcacgatcc 10740
ctcaaaatgt ccttagcaat tttttgtaca ttttttccac cgtcttttcc ttccggcttc 10800
tttgggaaat agcaatctat aggaccttta gtctactcg taccgctcga tccatggctt 10860
gaagagattg catctcttcc ccgttttgtt ggaagtgaca attcttcaat atcatcatca 10920
ccatcatcat caagattggg caccaatggg tgatgactca tttgattttt ttgctccttc 10980
ttcttctcaa caaaattctt tatttcatcc ctacctccg gtggacattt cggacaactt 11040
gtgacgtttc tatcgccacc aataagatgg aatttaaage gataaattcc acccgttgta 11100
atcttggtac aaaacttgca tacaatattt gtattcttct cattaactct atcgccatat 11160
cgccaagccg ggtctttatc tttcgatctt tgagacattt tgaaatccct attaagatat 11220
aagaaaatac ataatacaat aaactgaaaa tacatataat atatataata attaatttta 11280
taaaccgaaa tacacattaa aaaaatcaaa taaactaaaa tataacatat ataattttct 11340
aaattgaaat acatataaaa ttaatcaaat aaactgaaaa cataacatat atataataat 11400

```

taattttata tactgaaaaa tgcattaaga aaatcaaata aactgaaaaa tataacatat 11460
ataaataata attaatTTtT tatactgaaa taaaaatcaa ataaactaaa atataacata 11520
tatatataat aattaatttt ctaaaatgaa atacatataa aattaatcaa ataaactgaa 11580
aatataacat atataaataa taattaatat tatatactga aatatacatt aaaaaatcaa 11640
ataaactgaa aaatataaca tatataaata ataattaatt ttatatgctg aaatatacat 11700
taaaaaaatc aaataaactg aaaaatataa catatataaa taataattaa ttttataaac 11760
taaaatacat attaaaagta ataaataaga agaataaagg ggggaaaaaa acaaaatggg 11820
ttagagataa aaggaagaag aagacgctga agtctgaaca gtgcaaactg gactgtcaga 11880
gtctcagaga taaacgaaga agaagaagga agaaagaagg aaaaagaaag aaaaaaaagc 11940
agaatctgac gggaccaaga agaagaaaga aagaagaaag aaaaaaaga gaagaagaaa 12000
attacctgga ctggtcagca atgtcaatga agttgaacce taggctttgg tcgacctcga 12060
tcgaagaaga gagaagaaga aatggccatt ttttgtcaat ttagggctctg ttttgtataa 12120
tagaaaaaca gacccaaaagt tttttttaa aaaaaaaggg cctgcgcttt tttacaaaa 12180
gcgatcgctt cgtcgcttcc tccgttgaag cgtgcgcttc cctcggtcga gtcgcctcag 12240
gcagctagaa gcgacactga gtcgcgtcgc gtcgcttcgc gctttaagcg cgaagcgatc 12300
gcttttctaa acactggctc ttccaagttc tgtccatctt tctgggtgtgc ttaacccatg 12360
tatattggac cacactttgt attcatgctg attgcaacaa agcctgactg atcctgaact 12420
cctccttttt gccttgggta gaattatttt tgtgttttta gtgttacata tttcaattta 12480
aaggaaaaag ggagaagctt aaggtagaat cctgctcctt gacctttgga tccttacaca 12540
ttgaagtgga cagacatgtt tgctttgctt tgctttgatg agttgtttta gtagtcagtc 12600
taagcaactg ctgatgtatt ctgctttatt gtaagtgtct ataattttta tctttccaca 12660
atttgctttc agctctttct gcatataaga gaaaataaca gaagcttcta ttcgatcatg 12720
taatattgac tgcacaagtg catattttaa caactttggc tatgcgtatg caagcaataa 12780
taaataattac tttttatgca gggctctgtat ggttttcttc agagatgaaa gccttaagtg 12840
atgactgtga acgatttggt agtttccttc ccggtcatat ttattcaagc aaaaatggta 12900
tgcccatacg aagctcattt ctgagcttgc gcttcattta tgatcagcgt acttttccta 12960
tttatttcag atttacttta ggaccctgtt ggaaacagtc ttgaaccctt attgatgggtg 13020
taattattat ctcttgatca accaaaggag taattacatg taattattac aaatttgcaa 13080
caaaatcttt taggataatc agtggagctt catgaatctt ttaggatagt tgggatgcta 13140
gagtgatgta cttttgataa tgtcggttaa cataaatcct ctgactttgt actgatgtac 13200
tggtcgagca gattttatat tactttctaa tcctctaca aataaaaaaa ttgtgagact 13260
gcaatataac accatttttt taatgcattt acaatgttta aacttgattt agcaacgtag 13320
aacatgtcga ccttaaacag aggttctatt tgtcctagat aaaggaagat cagcataatg 13380
aaacattaca ggagaggagc tctattgagt ttttgttaat ctagagggag cccgtaaata 13440
gtacaccaat tcatctctta tatgaactct ttttcatttt ggaacgtttt aaaagtttga 13500
ctacattttt tgatctccgt aactaaact ttttgatgta ttaaactaac tatttccact 13560
actttaatat tttcttcctt aaacaactcc tgttagcatc ttaaactctg tgtgttcaat 13620
accctcatat aaaattatac gagtattttt tttatatttg cttatgtcct tggagtttgg 13680
gatggggaag ctcatcattt gaggcgaaag gtaaattcct ttgaatggac aaaaagtga 13740

```

cctcgggtcctt tggcatggag ggactattta agtcgccatc attaccaaag tcateccttc 13800
aagtgtttgc tttcaagaac tattcatgct atgtctcaag tatattatgt gtgcaggagg 13860
actcagaaga tggtaacaacc caccatggta ctcaaaaacc attccttcta ctccatatga 13920
tcaccttgct ttacggaaaag cttttgagaa ggtcattcta ttttggaagt cgcttgtttc 13980
tgagttgtat ataatattta ggtgccactt tgttatattt taaacttcat caaaagcctt 14040
tggaattcctt gaggacacat atccccttga ctgactggga gctgaatttc cttgatgtat 14100
agtgccttatt ttgattgtcc tctcctttgt aatactcttg cctataaata aaccttcctt 14160
atcaaaaagaa aatgataaaa tgttaaaaat actttgtgat agatgttgac tatagacttt 14220
aaatgatatt ttatctcaat ctttatcaat tcttttttac cttcacatgc aaattatact 14280
acagtatgtg aaatatgcaa tagtttgtcc tgcaattttc atctttcatg tacctctctc 14340
tactatccgt gtattactac tgatttgatg cattaaatgc taattttatt cccaaccttt 14400
actttatgca ggctgtagtc aagcgactta tgacggatgt accatttggg gtgcttctct 14460
caggcggact agattcttca cttgttgetg cagtggccaa ccgctatttg gctgatactg 14520
aagctgcgcg acaatgggga tcacagttgc ataccttttg cgtaggcttg aaggtggtga 14580
tgctcctttt tgtaaaactg ctctgatttg ggaattgcta gctgtatcat cccaatttac 14640
atggaagttc tagaccatta ttgctgtttc tttccatat atagggattt tctgaaaata 14700
gtaaaaagga aatggaattt ctctttatta aagataccga atgccttttt gggaaatgat 14760
cttacagttc tgaatcgaaa atttcctttt gacattgtga ttaattctga tatgaatttg 14820
gtccatcctg gctttccttc tcttgattgt agattttcct tctcttgta caattatatg 14880
cctgtccttt tgtcttactc tctgagacag gcaatgtaac tccagttgca aggtttggta 14940
atattgactt gtttatttct gttatcacat cctcctgaaa aagtgtact agcttgattt 15000
catgcgggaa tttgaaattt tcaagtgtat agctttacaa aatgatcaga actgatcaca 15060
gtcattgctc taaataaaat tttctaaaaa gcaggtattg ttgtgatata atgtgacact 15120
taaataaatt taaacctaca tcctttcgaa atctgtaata gctccactaa tatttcttgt 15180
gctaaactat tccaacacct cttatcgatg ccaattttga cagaccttag tgcacttaaa 15240
agtgttttagt gttctttatg cagtaggtat gttattgtcc aaaatgacag aaataaagtt 15300
agagatttgt tgaaagccac ttggtgtgat tgagcagttc ggagatgagt gatttatcca 15360
ttttatatac ttgtagtggt ttgttttctt ttgcttttgt cattgtaatt aaggttgtct 15420
tttgtttgac attaggggtc tctgatctg aaagctgcca gagaggtggc agactacctt 15480
ggaacccgtc accatgagtt tcactttaca gtgcaggtaa ctttccacaa agacatccta 15540
ccatgtgacc aggaggtcac gggtttgagc tgtggaaaca accttttgca gaaatgcagc 15600
gtaaggttgc gtacaataga cccttggtgt ccggcccttc cccgaccccg cacatagcgg 15660
gagcttagtg caccgggctg cccatcctaa gctgattttc cttataggag tataaccgct 15720
taagtcaaat tggctatcta cttctttcta ttgtacatc aggaaggaat tgatgcacta 15780
gatgaagtca tttatcatgt tgaaacatat gatgtgacca ctatcagagc cagtacacca 15840
atgtttctca tgtctcgga gataaagtct ttgggtgtga aaatggttct atctggtgaa 15900
ggttctgatg aaatttttgg cggttattta tatttccaca aggacccaa caaagaggag 15960
tttcaccaag aaactgttag aaaggtggat ctcatgtca tttcttcagt cattcaggat 16020
tattggaagt aactgtattt ttactagaat acttctttga actatttcta gctctgatgg 16080

acgctttgat gttaatccag attaaagcac ttcattcttta tgattgcttg agggccaaca 16140
 aatctacttc agcttgggggt gttgaagctc gtgtaccttt cttggataaa gaatttatca 16200
 atgttgcaat gaacattgat ccagagtgga aaatggtaac ctacagtgc tccgtctcat 16260
 ttctccaccc cccaccccc caaaaaaag aacagaaaag aattgaaaag agagattctt 16320
 ttcttccttc aaggaagaag ggggtaacaa aagaatagct gcaagttatg cagaaaaaat 16380
 agctctccat gtcattgggag accagggttc aaatctgatt gttaaaaaaa acactagggtg 16440
 aactcttcta atctgtttta gccttgggtg acatagttaa ctagtatatg agctagttag 16500
 agctgactcg aaaccacat tattcaaaaa attaagaaag tcttcaatgt tgttccaccg 16560
 tgaattcatg atgtgtcgac tttcatgtgt ttatctctgg atagatgaag agtttttttg 16620
 gatcttgatt aattctgttt ggcccccttg aggttctgca cgctatttct agactcaaac 16680
 ttgctgcgac ctatgttagg tctagagaat atctgcctaa taaaacaatg gaagttacat 16740
 tcaagtattt tacctccata ttacaagtca gtgttagtta aatgggtctg atatatggga 16800
 tagaaaatgg ctccataccc ataacctcgt taaaatgttg ctgatgtgct gatctcatct 16860
 gtacatctag tataatccta attacaaatt aatgcattct tttgcggtga ctatagtatg 16920
 ctgaagaacc tgtcttgaag cattgcagat cagacctgat ctcggaagaa tagaaaagt 16980
 ggttctacgc aatgcttttg acgatgatca gaatccttat ctgccaaagg tttgtttgtt 17040
 ttcttgtatc aggtccaata atcataatct ttgtttatct atttgcataa aaggcagtag 17100
 actccaata cattgtgtac gcttatgcac aaatgtatga gatgtgtatt atatgtgtca 17160
 catataaggg tattatatgt acacactcct cccaccttcc acatatatta tgtgtgtgtg 17220
 tgtgtgtgta taatgaaatg atgtttgaaa agacattatg atcaagtagt tgatcctcat 17280
 ctgctttctg tttttagcat atcttgtata ggcagaagga acagttcagt gatggagttg 17340
 gctacagttg gattgatggc ttgaaggatc acgcaagcag tctggtgagc tctttagttg 17400
 ttttctcttc cctcctcttc cattctttat cccttccttt tgtgggtttg ttgaacacat 17460
 ctcatatact tgggtccagg tttctgattct atgttagcga atgcaagttt tgtttaccgg 17520
 cataacacac ccacgacaaa ggaaggatac tattatagaa ctatttttga gcgatatttc 17580
 cccaagggtg gtcttacata gtcttctttt ccttgttatt catgttcagt ttatgatatt 17640
 ttgtcccaaa caatgttttc ttttcttcc ttttctgtt aacaagatcg cttctatctt 17700
 cgtatttggt gataaagtta atcacattgt taactattat ttggagcaga atgctgcgag 17760
 ggaaacagtt ccagggtggc caagtgtggc atgcagcact gcaaaagcag tagaatggga 17820
 cgcagcttgg tccaagaatc tagatccatc tggtcgagct gactcgggtg ttcattgcagc 17880
 tgcttatgag gatgcatcag aggttaaaac cacagtcgcg acagatactg ctcagaaact 17940
 tgacgttgat aaagctgcag tagctgtttg a 17971

<210> 12

<211> 581

<212> PRT

<213> 烟草

<400> 12

Met Cys Gly Ile Leu Ala Val Phe Gly Cys Ile Asp Asn Ser Gln Ala

1

5

10

15

Lys Arg Ser Arg Ile Ile Glu Leu Ser Arg Arg Leu Arg His Arg Gly		
20	25	30
Pro Asp Trp Ser Gly Leu His Ser His Asp Asp Cys Tyr Leu Ala His		
35	40	45
Gln Arg Leu Ala Ile Val Asp Pro Thr Ser Gly Asp Gln Pro Leu Tyr		
50	55	60
Asn Glu Asp Lys Thr Ile Val Val Ala Val Asn Gly Glu Ile Tyr Asn		
65	70	75
His Arg Asp Leu Arg Glu Lys Leu Lys Ser His Gln Phe Arg Thr Gly		
85	90	95
Ser Asp Cys Glu Val Ile Ala His Leu Tyr Glu Asp Tyr Gly Glu Asp		
100	105	110
Phe Val His Met Leu Asp Gly Met Phe Ser Phe Val Leu Leu Asp Thr		
115	120	125
Arg Asp Lys Ser Phe Ile Ala Ala Arg Asp Ala Ile Gly Ile Thr Pro		
130	135	140
Leu Tyr Met Gly Trp Gly Leu Asp Gly Ser Val Trp Phe Ser Ser Glu		
145	150	155
Met Lys Ala Leu Ser Asp Asp Cys Glu Arg Phe Val Ser Phe Leu Pro		
165	170	175
Gly His Ile Tyr Ser Ser Lys Asn Gly Gly Leu Arg Arg Trp Tyr Asn		
180	185	190
Pro Pro Trp Tyr Ser Glu Thr Ile Pro Ser Thr Pro Tyr Asp His Leu		
195	200	205
Val Leu Arg Lys Ala Phe Glu Lys Ala Val Val Lys Arg Leu Met Thr		
210	215	220
Asp Val Pro Phe Gly Val Leu Leu Ser Gly Gly Leu Asp Ser Ser Leu		
225	230	235
Val Ala Ala Val Ala Asn Arg Tyr Leu Ala Asp Thr Glu Ala Ala Arg		
245	250	255
Gln Trp Gly Ser Gln Leu His Thr Phe Cys Val Gly Leu Lys Gly Ser		
260	265	270
Pro Asp Leu Lys Ala Ala Arg Glu Val Ala Asp Tyr Leu Gly Thr Arg		
275	280	285
His His Glu Phe His Phe Thr Val Gln Glu Gly Ile Asp Ala Leu Asp		
290	295	300
Glu Val Ile Tyr His Val Glu Thr Tyr Asp Val Thr Thr Ile Arg Ala		
305	310	315
Ser Thr Pro Met Phe Leu Met Ser Arg Lys Ile Lys Ser Leu Gly Val		

	325	330	335
Lys Met Val Leu Ser Gly Glu Gly Ser Asp Glu Ile Phe Gly Gly Tyr			
	340	345	350
Leu Tyr Phe His Lys Ala Pro Asn Lys Glu Glu Phe His Gln Glu Thr			
	355	360	365
Cys Arg Lys Ile Lys Ala Leu His Leu Tyr Asp Cys Leu Arg Ala Asn			
	370	375	380
Lys Ser Thr Ser Ala Trp Gly Val Glu Ala Arg Val Pro Phe Leu Asp			
385	390	395	400
Lys Glu Phe Val Asn Val Ala Met Asn Met Asp Pro Glu Ser Lys Met			
	405	410	415
Ile Arg Pro Asp Leu Gly Arg Ile Glu Lys Trp Val Leu Arg Asn Ala			
	420	425	430
Phe Asp Asp Asp Gln Asn Pro Tyr Leu Pro Lys His Ile Leu Tyr Arg			
	435	440	445
Gln Lys Glu Gln Phe Ser Asp Gly Val Gly Tyr Ser Trp Ile Asp Gly			
	450	455	460
Leu Lys Asp His Ala Ser Arg Leu Val Ser Asp Ser Met Leu Ala Asn			
465	470	475	480
Ala Ser Phe Val Tyr Pro His Asn Thr Pro Thr Thr Lys Glu Gly Tyr			
	485	490	495
Tyr Tyr Arg Thr Ile Phe Glu Arg Tyr Phe Pro Lys Asn Ala Ala Arg			
	500	505	510
Glu Thr Val Pro Gly Gly Pro Ser Val Ala Cys Ser Thr Ala Lys Ala			
	515	520	525
Val Glu Trp Asp Ala Ala Trp Ser Lys Asn Leu Asp Pro Ser Gly Arg			
	530	535	540
Ala Ala Leu Gly Val His Ala Ala Ala Tyr Glu Asp Ala Ser Glu Val			
545	550	555	560
Lys Thr Thr Val Ala Thr Asp Thr Pro Gln Lys Leu Glu Val Asp Lys			
	565	570	575
Ala Ala Val Ala Val			
	580		

<210> 13

<211> 25

<212> DNA

<213> 人工序列

<220>

<223> 引物

<400> 13
tcacatatat tccctcataa cacac 25
<210> 14
<211> 24
<212> DNA
<213> 人工序列
<220>
<223> 引物
<400> 14
aagaagcatc ccactctaca gctt 24
<210> 15
<211> 26
<212> DNA
<213> 人工序列
<220>
<223> 引物
<400> 15
atatcttccc tcacaacact ccaact 26
<210> 16
<211> 22
<212> DNA
<213> 人工序列
<220>
<223> 引物
<400> 16
ctaccggaag gatcaaggtt gt 22
<210> 17
<211> 330
<212> DNA
<213> 烟草
<400> 17
tcacatatat tccctcataa cacacccatt acaaaggaag cataactacta taggatgatt 60
ttcgagcgct ttttcccaca gaattcagct gggctaaccg ttcttgagg agcaagtgtg 120
gcgtgtagca cagctaaagc tgtagagtgg gatgcttctt ggtcaaagaa ctttgatcct 180
tcaggcaggg ctgctattgg tgtacataac tcggettatg agaatcatgt acctgctatg 240
gctaattggga atttgaccaa aaaaatcatt ggtcgtgtgc cttctatggt agaagttggt 300
gctgctcccg agctcacaat aaagagttag 330
<210> 18
<211> 364

<212> DNA

<213> 烟草

<400> 18

gaccagattg gagtgggata tatcaacatg gtgattttta cttagcacat caacgttttag 60
 caattatcga tcctacttct ggtgatcagc ctctgtttta tcaagataag actattgttg 120
 ttacagtcaa tggagaaatt tacaatcatg agaaacttcg taatcttatg cctaatacaca 180
 agttcagaac cggaagtgat tgtgatgtta ttgcacatct ttatgaagaa tatggagaaa 240
 attttgtgga catgttggat ggggtgttct cttttgtatt gttggatacg cgcgataaca 300
 gctttcttgc tgctcgtgat gcgattggaa ttactccctt ctatattggt tggggacttg 360
 atgg 364

<210> 19

<211> 1815

<212> DNA

<213> 烟草

<400> 19

atgtgcggga tcttggctgt tttgggttgt tctgatgatt ctcaggccaa aagggttcgt 60
 gttctcgagc tctctcgag gtaggttgaa gcatcgtgga ccagattgga gtgggctgta 120
 tcaacatggg gactgttact tggcacatca gcgtctagct attgttgatc ctgcttccgg 180
 tgatcaacct ctgtttaacg aagataagac gattgttgtt acggtaggta aatggagaga 240
 tctacaatca cgagcaactt cgtaagcaaa tgcctaata taagttccgg actggcagtg 300
 actgtgatgt cattgcacac ctagtagtat gaagaacatg gagaagattt tgtggacatg 360
 ctggatggga tcttcgcttt tgtgttattg gatactcgag ataacagctt tcttgttgct 420
 cgtgatgcca ttggaattac ttccctttat attgggtggg gacttgatgg tagggctctgt 480
 atgaatatca tctgagctta agggcttgaa tgatgactgc gaacattttg aagttttccc 540
 accagggcac ttgtactcta gcaagaatgg cggctttagg aggtggtaca atcctccttg 600
 gttctctgag gccattcctt ccactcgta tgatccctta gttctcaggc gtgcctttga 660
 aaatgtaggc tgttatcaaa aggttgatga ctgatgtccc ctttgggtgtt ctctctccg 720
 ggggactcga ttcattccttg gttgcttcga ttactgctcg ctacttggct ggtacaaagg 780
 ctgccaaagca gtggggagca cagcttcatt cttctgtgt tggccttgag gtagggctca 840
 ccggatctca aggctgcaag agaagttgct gactacttgg gaaccgttca ccacagattt 900
 cacttcaccg ttcaggtagg atggaattga tgcaattgaa gatgttattt accatattga 960
 gacatacgat gtaacgacaa tcagagcaag cactcctatg ttctttatgt cgcgtaagat 1020
 taagtcacta ggagtgaaga tggcatatc tggggaagga tctgatgaag tgtttggtgg 1080
 ctacttgtag tttcacaagg ctcccaacaa ggaagagttc cacaaggaaa catgtcgcaa 1140
 ggtagattaa agcgcttcac caatatgact gcttaagagc aaataagtca acatctgcat 1200
 ggggtttaga agctagagtc ctttccctag ataaggagtt catcaatgtt gccatgagta 1260
 ttgatccaga gtggaagttg agattaaacc agagcaaagg aggattgaaa agtgggctct 1320
 aaggagggcc tttgatgatg aggagcatcc ttatctccca aaggtagcac atcctgtata 1380
 ggcaaaaaga acaattcagt gatggcgtgg gcatagttgg atagatggac tcaaagcaca 1440

tgctgaacaa catgtaggtg accaatagga tgatgtttaa tgcttcacat atattccctc 1500
 ataacacacc cattacaaag gaagcatact actataggat gatttttcgag cgctttttcc 1560
 cacaggtaga attcagctgg gctaaccgtt cctggaggag caagtgtggc gtgtagcaca 1620
 gctaaagctg tagagtggga tgcttcttgg tcaaagaacc ttgatacctc aggcaaggct 1680
 gctattgggtg tacataactc ggcttatgag aatcatgtac ctgctatggc taatgggaat 1740
 ttgacaaaaa aaatcattgg tcgtgtgcct tctatggtag aagttgggtgc tgctcccag 1800
 ctcacaataa agagt 1815

<210> 20

<211> 155

<212> PRT

<213> 烟草

<400> 20

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

<210> 21

<211> 1818

<212> DNA

<213> 烟草

<400> 21

atgtgcggga tcttggtgtg tttgggttgt tctgatgatt ctcaggccaa aagggttcgt 60
 gttctcgagc tctctcgcag gtaggttgaa gcatcgtgga ccagattgga gtgggctgta 120

tcaacatggg gactgttact tggcacatca gcgtctagct attgttgatc ctgcttccgg 180
 tgatcaacct ctgtttaacg aagataagac gattgttggt acggtaggta aatggagaga 240
 tctacaatca cgagcaactt cgcaagcaaa tgcctgatca taagttccgg actggaagtg 300
 actgtgatgt cattgcacac ctagtagtat gaagaacatg gagaagattt tgtggacatg 360
 ctggatggga tcttcgcttt tgtgttactg gatactcgag ataacagctt tcttggtgct 420
 cgtgatgcca ttggaattac ttccctttat attgggtggg gacttgatgg tagggtctgt 480
 atgaatatca tctgagctta agggcttgaa tgatgactgc gaacattttg aagttttccc 540
 accaggacac ttgtactcta gcaagaatgg cggttttagg aggtggtaca atcctctttg 600
 gttctctgag gctattcctt ccactcetta tgatecetta gttctcaggc gcgcctttga 660
 aaatgtaggc tgttatcaaa aggttgatga ctgatgtccc ttttgggtgt ctgctctccg 720
 ggggactcga ttcatccttg gttgcttcga ttactgccc ctacttggct ggcacaaagg 780
 ctgccaagca gtggggagca cagcttcatt ccttctgtgt tggccttgag gtagggatca 840
 ccgcatctca aggtgcaag agaagttgct gactacttgg gaaccgttca ccacgagttt 900
 cacttcaccg ttcaggtagg atggaattga tgcaattgaa gatgttattt accatattga 960
 gacatacgat gtaacgacaa tcagagcaag cactcctatg ttccttatgt cgcgtaagat 1020
 taagtcacta ggagtgaaga tggttatata tggggaaggc tctgatgaag tgtttggtgg 1080
 ctacttgtag tttcacaagg ctcccaacaa ggaagagttc cacaaggaaa catgtcgcaa 1140
 ggtagattaa agcacttcac caatatgact gtttaagagc aaataagtca acatctgcat 1200
 ggggcttaga agctagagtg cctttcctag ataaggagtt catcaatgtt gccatgagta 1260
 ttgatccaga gtggaagttg gtagattaaa ccagagcaaa ggaggattga gaagtgggct 1320
 ctaaggaggg cctttgatga tgaggagcat ccctatctcc caaaggtagc acatcctata 1380
 caggcagaaa gaacaattca gtgatggcgt aggctatagt tggatagatg gactcaaagc 1440
 acatgctgaa caacatgtag gtgaccaata ggatgatgct taatgcttca catatatcc 1500
 ctcataacac accgattaca aaggaagcat actattatag gatgattttc gagcgctttt 1560
 tcccacaggt agaattcagc tgggctaacc gttcctggag gagcgagtgt ggcgtgtagc 1620
 acagctaaag ctgtagagtg ggatgcttct tggtaaaaga accttgatcc ttcaggaagg 1680
 gctgctattg gtgtacataa ctgagcttat gagaatcatg aacctgctat ggctaattgg 1740
 aatttggcc caaaaatcat tggccgtgcg ccgtctatgg tagaagttgg tgctgctcat 1800
 gagctcacia taaggagt 1818

<210> 22

<211> 155

<212> PRT

<213> 烟草

<400> 22

Met Cys Gly Ile Leu Ala Val Leu Gly Cys Ser Asp Asp Ser Gln Ala

1 5 10 15

Lys Arg Val Arg Val Leu Glu Leu Ser Arg Arg Leu Lys His Arg Gly

20 25 30

Pro Asp Trp Ser Gly Leu Tyr Gln His Gly Asp Cys Tyr Leu Ala His

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

<210> 23

<211> 1818

<212> DNA

<213> 人工序列

<220>

<223> ASN5-S终止突变体ASN5-S_Q66*的DNA序列

<220>

<221> mutation

<222> (199) .. (201)

<400> 23

```

atgtgtggaa tcttggcttt gttgggttgt tcagatgatt ctcaggccaa aagggttcga 60
gttcttgagc tctctcgag gcaggttgaa gcatcgtgga ccagattgga gtgggatata 120
tcaacatggg gatttttact tagcacatca acgttttagca attatcgatc ctgcttctgg 180
tgatcagcct ctgtttaatt aagataagac gattgttgtt acagtaggtc aatggagaga 240
tttacaatca tgagaaactt cgtaatctta tgctaatca caagttcaga actggaagtg 300
attgtgatgt tattgcacat cttgtagtat gaagaatat gagaaaattt tgtggacatg 360
ttggatgggg tgttctcttt tgtattgttg gatacgcgcg ataacagctt tcttgctgct 420
cgtgatgcaa ttggaattac tccccatat attggttggg gacttgatgg taggctctgt 480
gtggatttca tctgagctaa aaggattaaa tggtgactgt gaacattttg aagttttccc 540
tcccggtcac ttgtactcga gcaagaatgg cgggtttagg agatgtgaca atcctcaatg 600
gttctctgag gctattccat caaatcetta cgaccetta gttttgagac gtgccttcga 660
aaatgtaggc tgttatcaaa cgattgatga ccgatgtccc ctttggtgtt ctgctctccg 720
ggggacttga ttcgtctttg gttgcttctg tcaactgctg ctacttggct ggaacaaaag 780
ctgctaagca atggggagca cagcttcatt ccttctgtgt tggctctcag gtagggctca 840

```

ccagatctca aggctgcaag agaagttgct gactatttgg gaaccgttca ccacgagttc 900
 accttcacag ttcaggtagg atggaattga tgctattgaa gatgttattt accatatcga 960
 gacatacgat gtaacaacga tcagagcaag cactcctatg ttccttatgt cgcgtaagat 1020
 taaatcactg ggagttaaga tggtcataatc aggggaaggc tcagatgaac tgtttggcgg 1080
 ctatttgtac ttccacaagg ctccgaacaa ggaagaattc catgtggaga catgtcacia 1140
 ggtagataaa agcgcttcac caatacgact gtttgagagc aaataaggca acatcagcat 1200
 ggggcttaga agctagagta ccatttctgg ataaagagtt catcaacggt gctatgagta 1260
 tcgatcctga atggaagatg gtagattaaa cacgatcatg gtaggatcga gaagtgggtt 1320
 cttaggaagg cttttgatga tgaggagcaa ccctatctcc caaaggtagc atattctgta 1380
 ccggcagaaa gaacaattca gtgatggcgt aggtatagtt tggatcgatg gactcaaagc 1440
 acatgctgaa caacatgtag gtgactgata ggatgatgct taatgctgca catatcttcc 1500
 ctcacaacac tccaactaca aaggaagcat actattacag gatgatttcc gagaggttct 1560
 tcccacaggt agaattcagc aaggetaact gttcctggag gaccgagtat agcttgcagc 1620
 acagctaaag ctattgagtg ggatgcttcg tggtcgaaca accttgatcc ttccggtagg 1680
 gctgcaatcg gtgtacataa ctcggcttat gacgatcctc tccccgatgt tggtaatggg 1740
 aatttggaca caacgatcat cgataatgtg ccgaggatgg taggagtggg tgctgctgca 1800
 gagctcacia taaggagc 1818

<210> 24

<211> 65

<212> PRT

<213> 人工序列

<220>

<223> ASN5-S终止突变体ASN5-S_Q66*的氨基酸序列

<400> 24

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

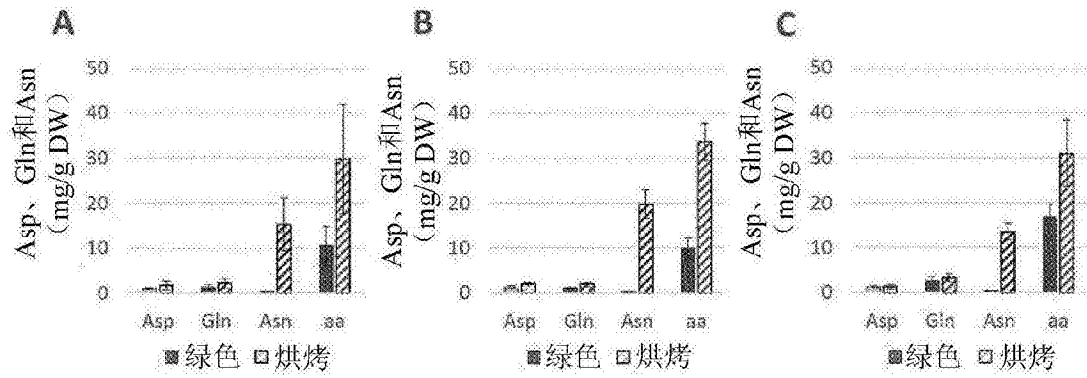


图1

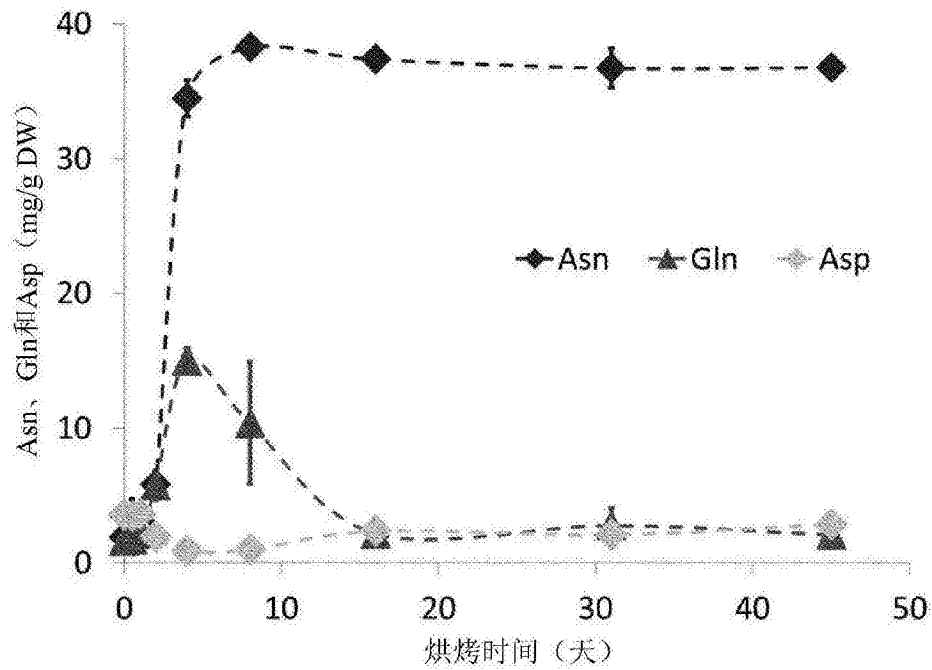


图2



图3

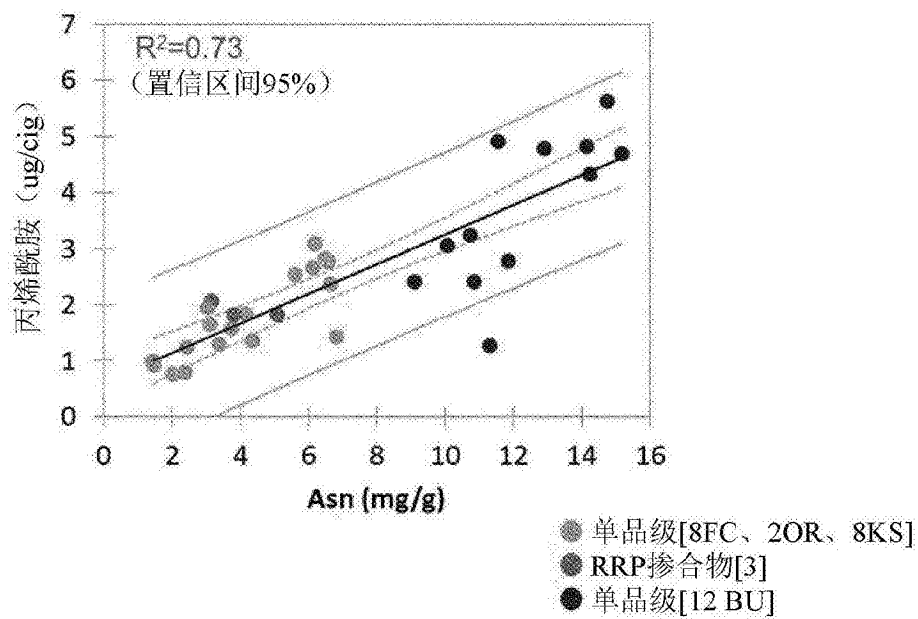


图4

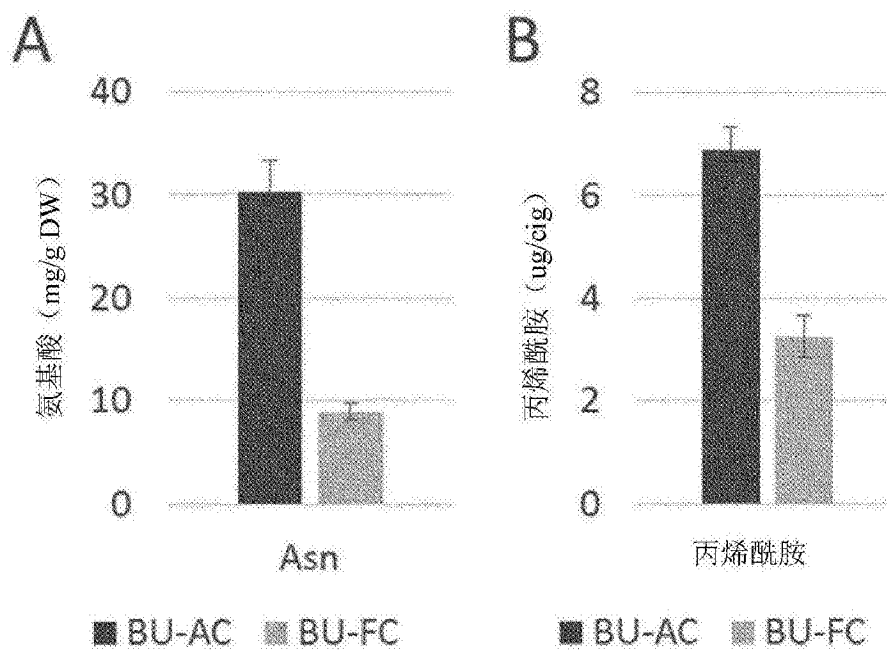


图5

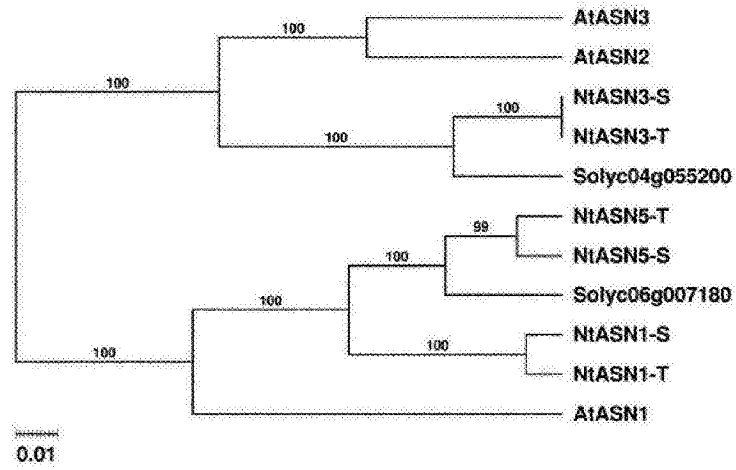


图6

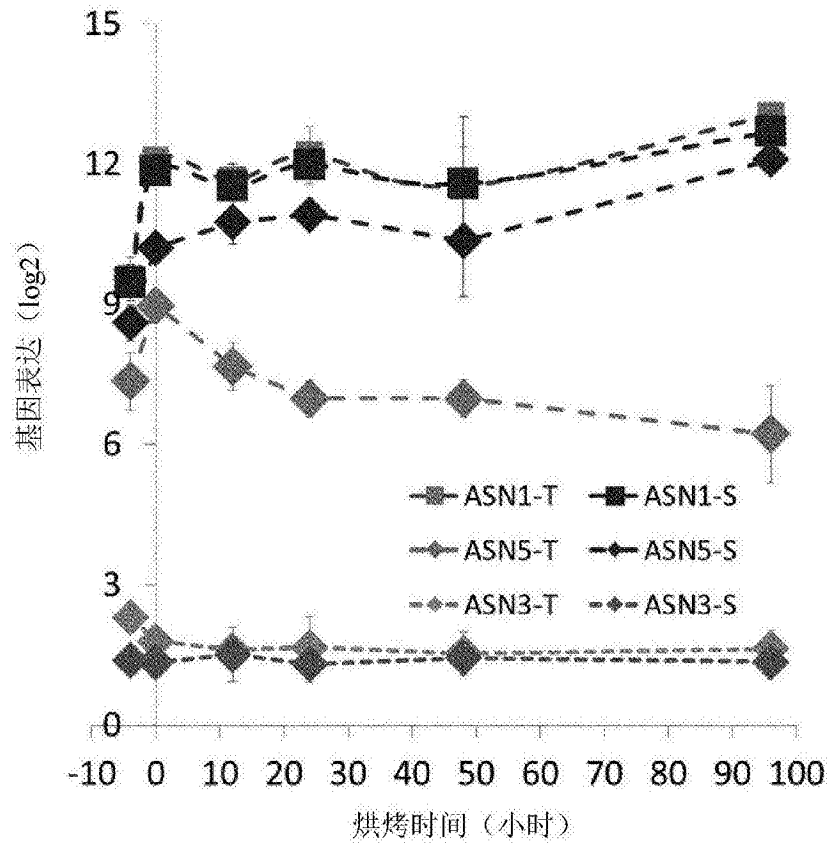


图7

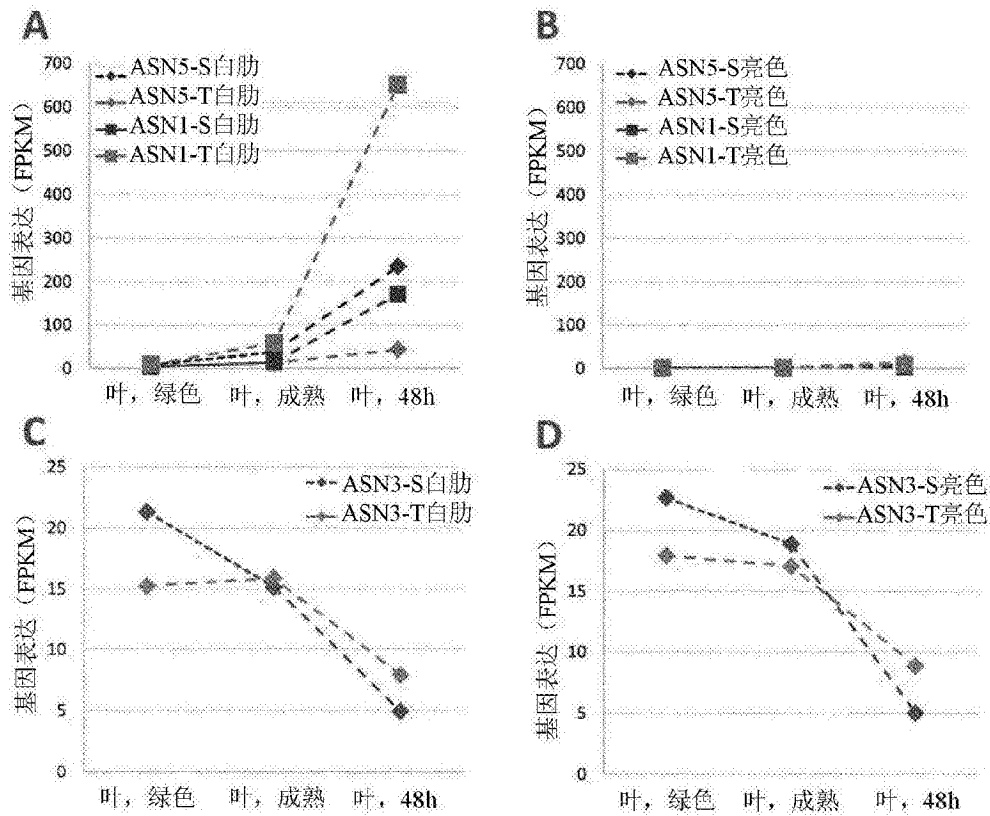


图8

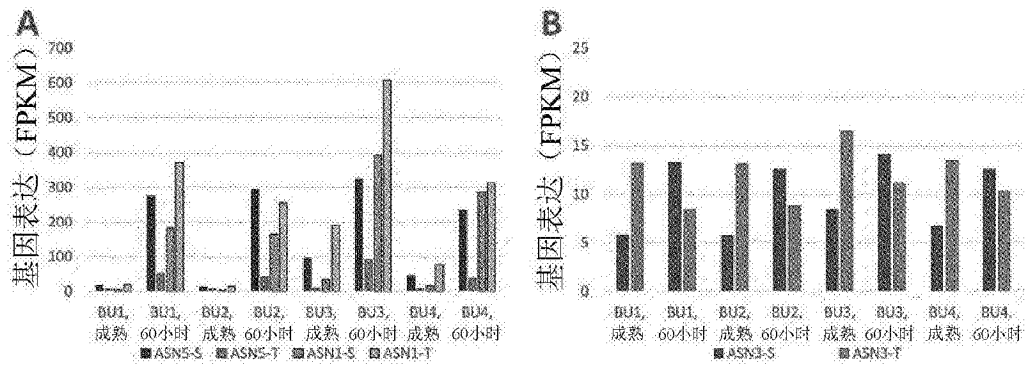


图9

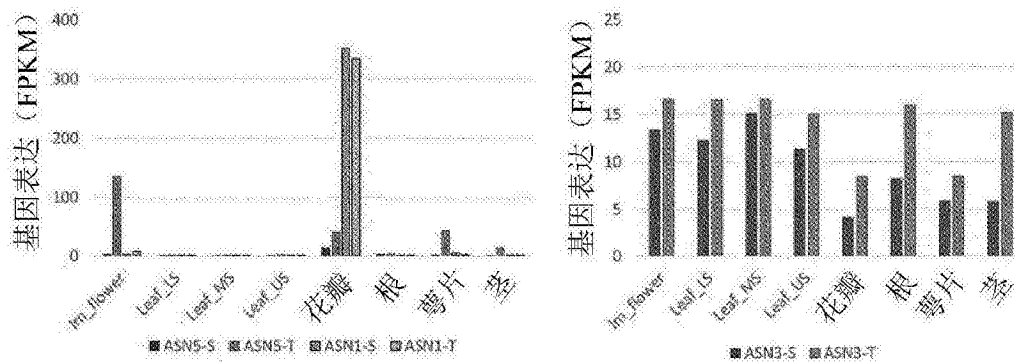


图10

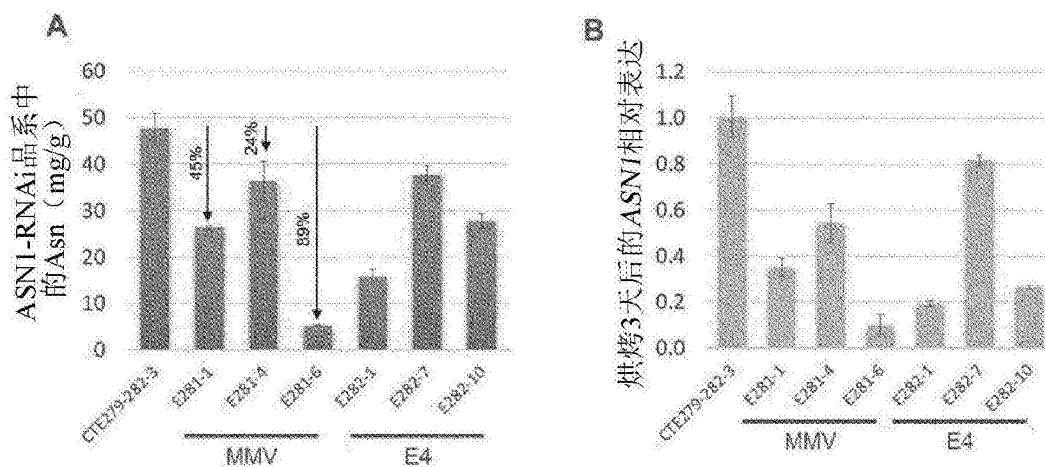


图11

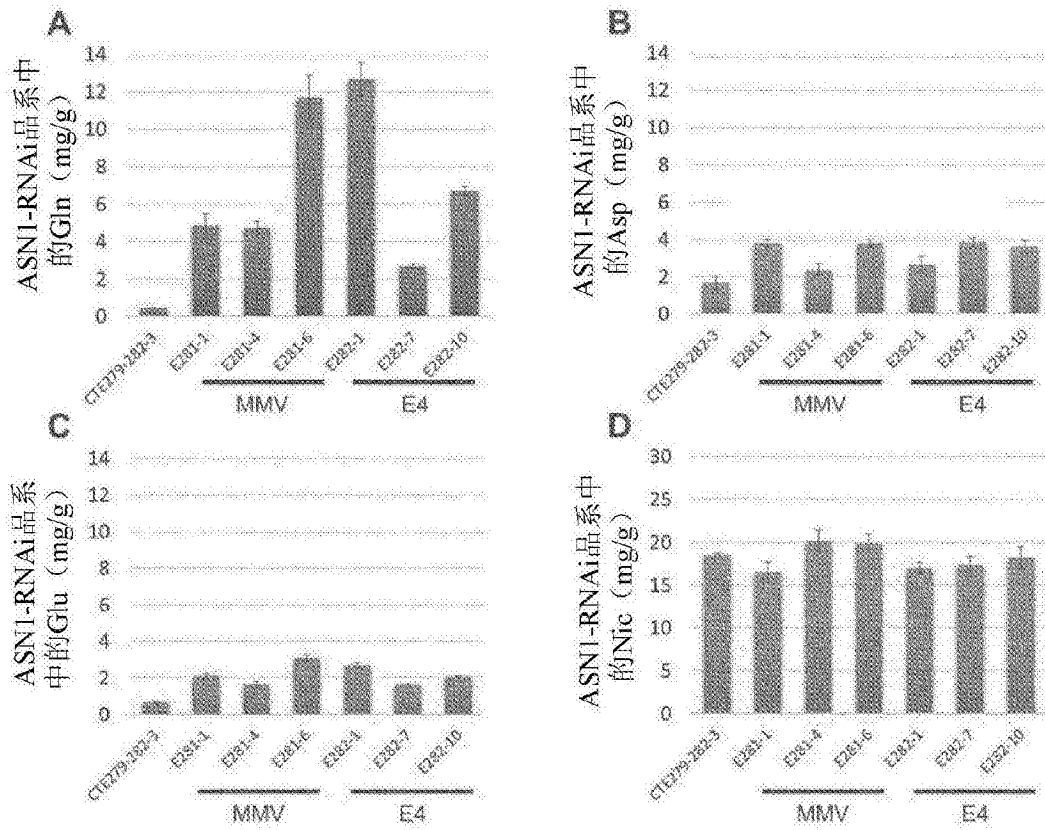


图12

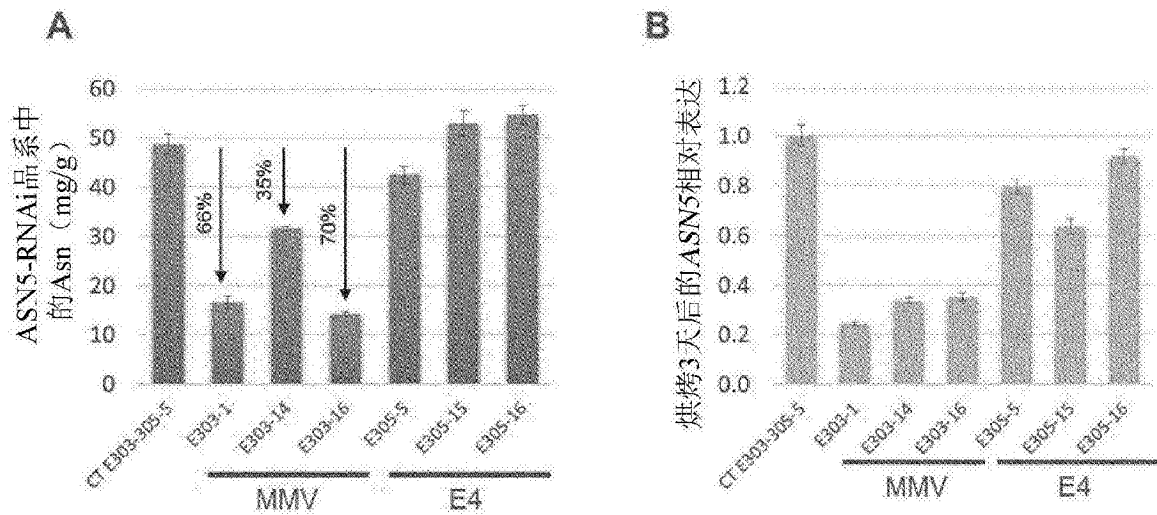


图13

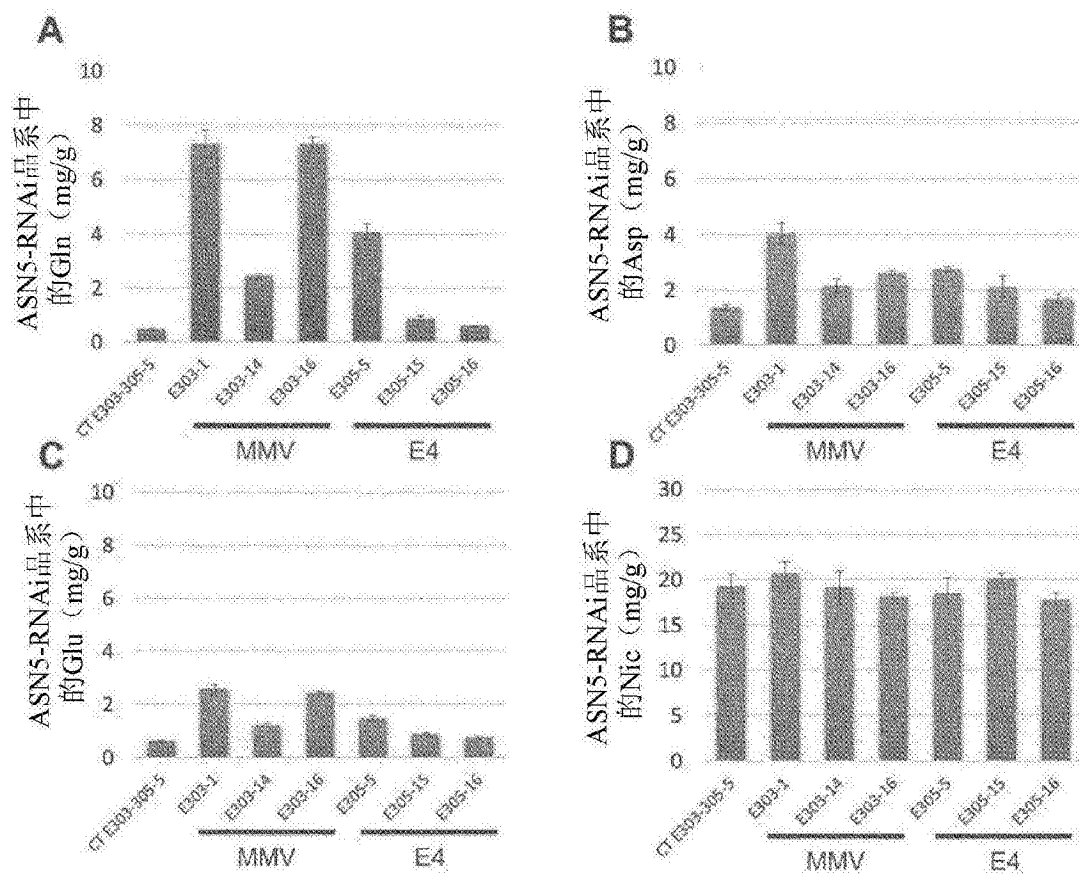


图14

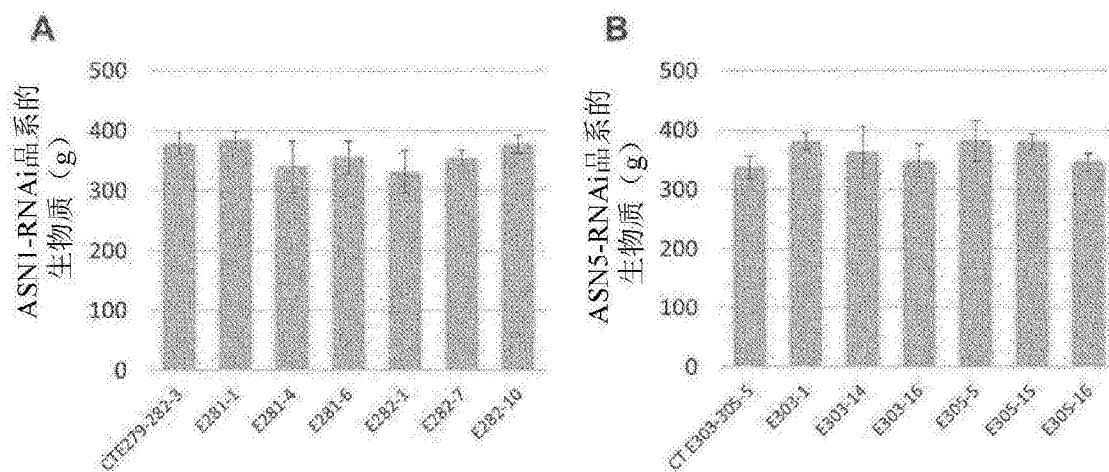


图15

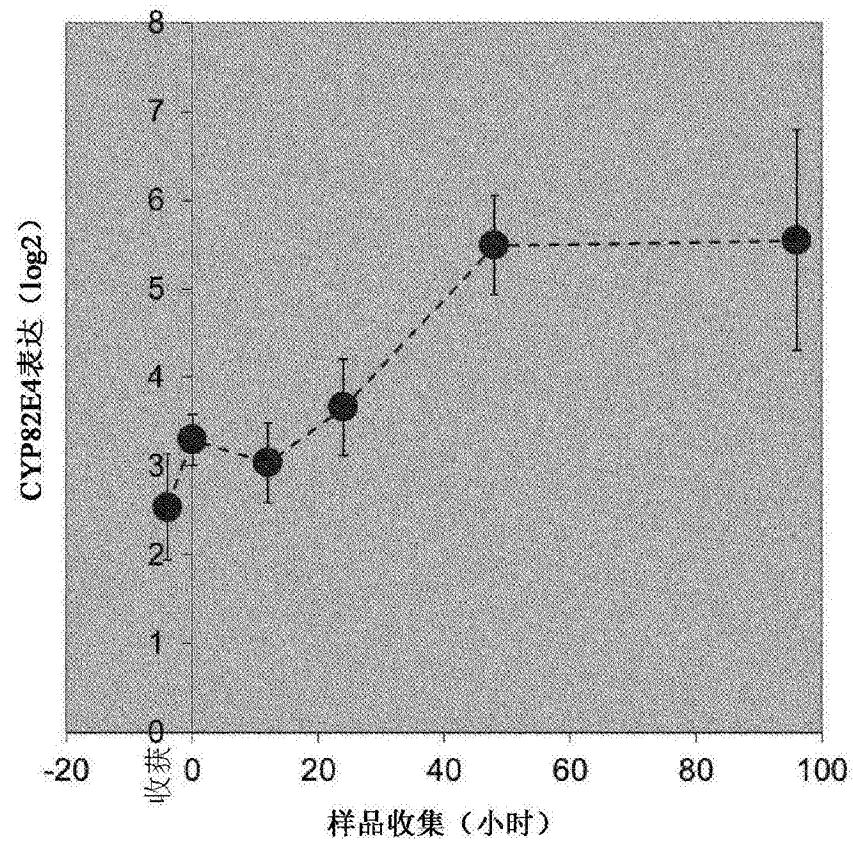


图16

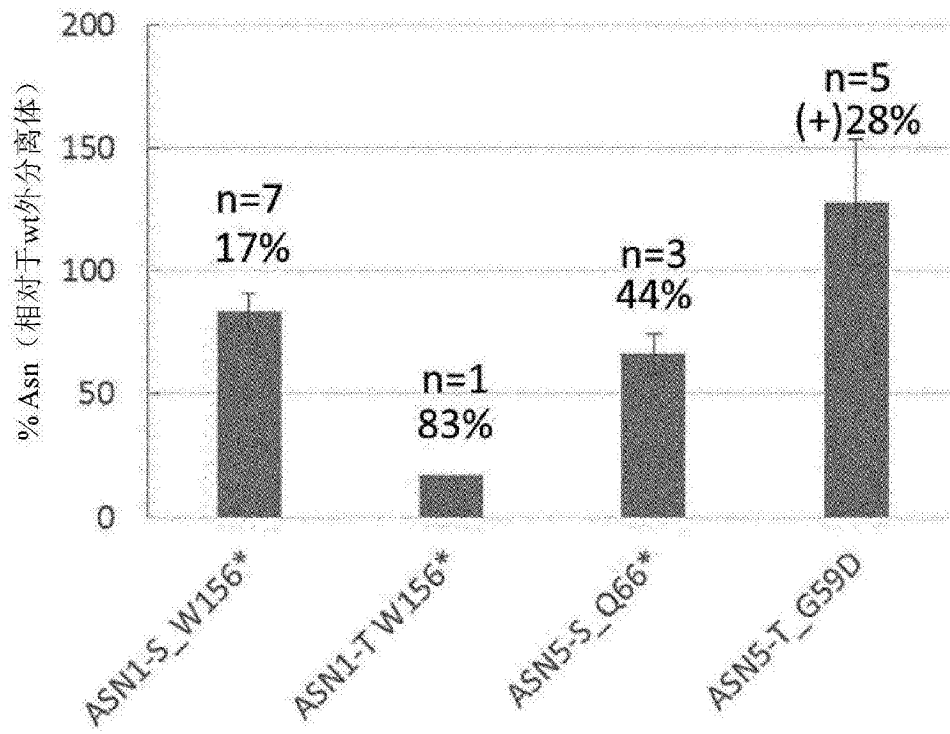


图17

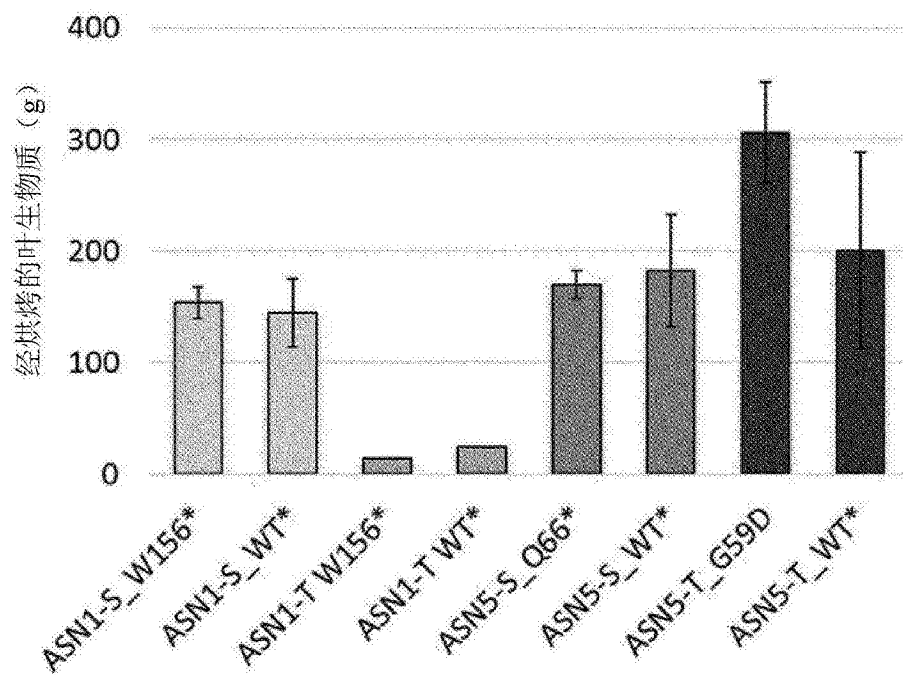


图18

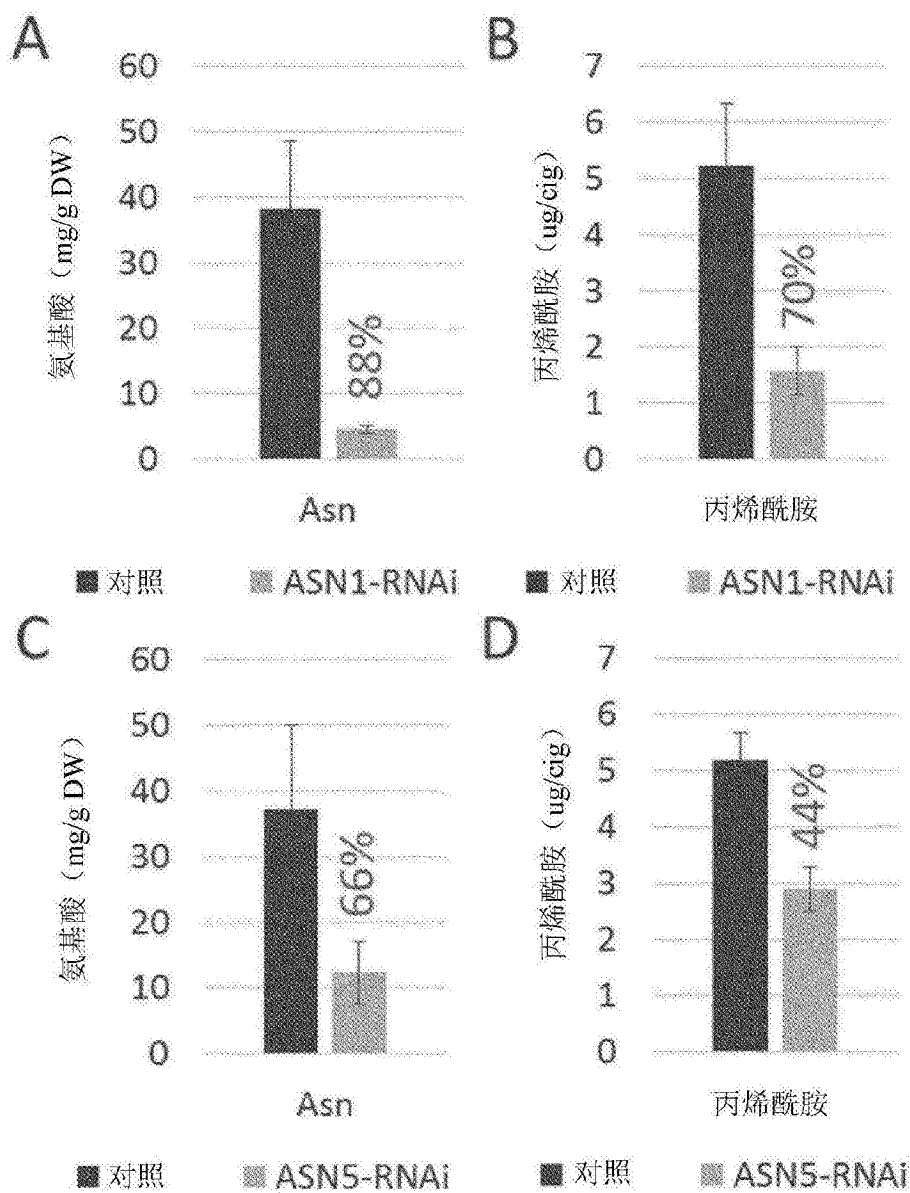


图19

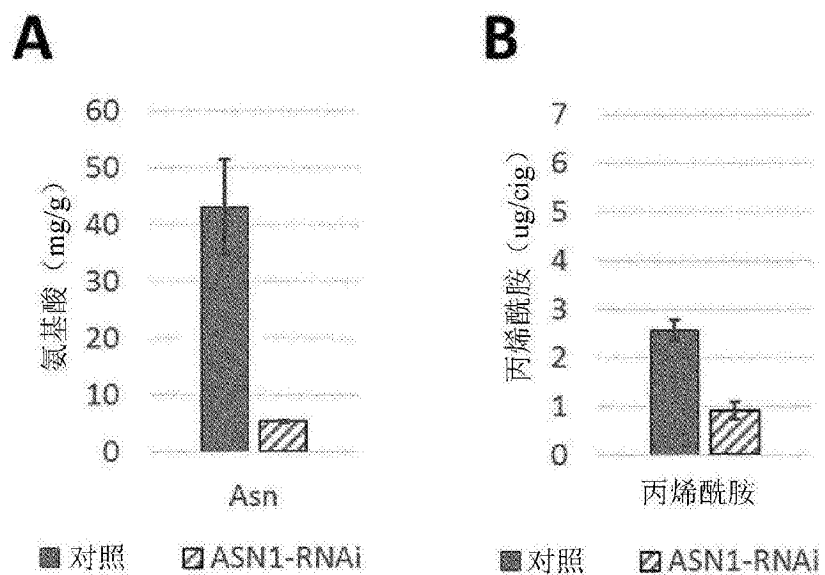


图20

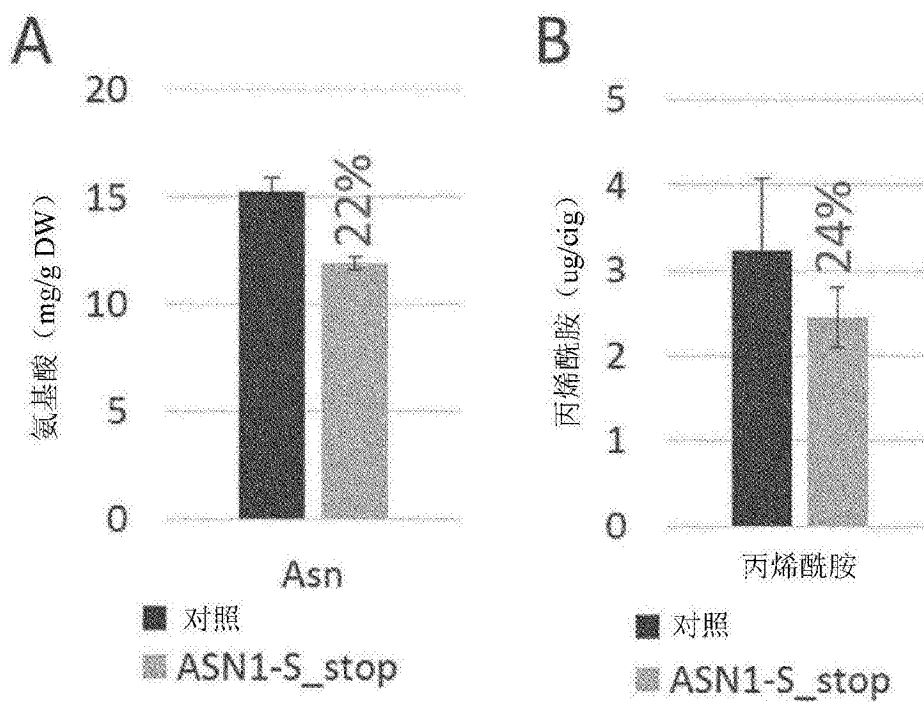


图21

