

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200580033745.7

[51] Int. Cl.

C12N 15/82 (2006.01)

A01H 5/00 (2006.01)

C07H 21/04 (2006.01)

[43] 公开日 2007年9月12日

[11] 公开号 CN 101035900A

[22] 申请日 2005.8.4

[21] 申请号 200580033745.7

[30] 优先权

[32] 2004.8.4 [33] US [31] 60/598,671

[86] 国际申请 PCT/US2005/027729 2005.8.4

[87] 国际公布 WO2006/015376 英 2006.2.9

[85] 进入国家阶段日期 2007.4.3

[71] 申请人 巴斯福植物科学有限公司

地址 德国路德维希港

[72] 发明人 R·A·阿申齐 G·布茨泽维斯基

G·凯克福达 B·K·辛格

[74] 专利代理机构 北京市中咨律师事务所

代理人 黄革生 林柏楠

权利要求书 11 页 说明书 47 页 序列表 12 页  
附图 4 页

[54] 发明名称

单子叶植物 AHASS 序列和使用方法

[57] 摘要

描述了编码乙酰羟酸合酶小亚基(AHASS)多肽的分离的多核苷酸以及由这些多核苷酸编码的氨基酸序列。本发明提供了包含编码 AHASS 多肽或 AHASS 融合多肽的多核苷酸的表达盒和植物表达载体。还提供的是用本发明的多核苷酸、表达盒或表达载体转化的植物、种子和宿主细胞。本发明还提供了使用本发明多核苷酸来增强 AHAS 活性并且增强植物对除草剂的耐受性的方法。

1. 包含选自下列的核苷酸序列的分离的多核苷酸:

(a) 在 SEQ ID NO:1、SEQ ID NO:3 中定义的多核苷酸; SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565;

(b) 多核苷酸, 其与在 SEQ ID NO:1、SEQ ID NO:3 中定义的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565 具有至少 80% 的序列同一性, 其中多核苷酸编码具有乙酰羟酸合酶小亚基(AHASS)活性的多肽;

(c) 多核苷酸, 其在严格条件下与在 SEQ ID NO:1、SEQ ID NO:3 中定义的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565 杂交, 其中多核苷酸编码具有 AHASS 活性的多肽;

(d) 编码多肽的多核苷酸, 所述的多肽如在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481 或 SEQ ID NO:5 的连续氨基酸 64-471 中所定义的;

(e) 编码与下列具有至少 81% 序列同一性的多肽的多核苷酸: 如在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5 中所述的氨基酸序列、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481, 或 SEQ ID NO:5 的连续氨基酸 64-471, 其中多核苷酸编码具有 AHASS 活性的多肽;

(f) 编码与 SEQ ID NO:5 的连续氨基酸 64-471 具有至少 77% 序列同一性的多肽的多核苷酸, 其中多核苷酸编码具有 AHASS 活性的多肽;

(g) 在 SEQ ID NO:10 所定义的多核苷酸; 以及

(h) 在 SEQ ID NO:11 所定义的多核苷酸。

2. 权利要求 1 的分离的多核苷酸, 其中多核苷酸包含在 SEQ ID NO:1、SEQ ID NO:3 中定义的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565。

3. 权利要求 1 的分离的多核苷酸, 其中多核苷酸包含与下列具有至

少 90%序列同一性的核苷酸序列：如在 SEQ ID NO:1、SEQ ID NO:3 中所述的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495，或 SEQ ID NO:3 的连续核苷酸 342-1565，其中多核苷酸编码具有 AHASS 活性的多肽。

4. 权利要求 1 的分离的多核苷酸，其中多核苷酸包含编码多肽的核苷酸序列，所述的多肽如在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481、或 SEQ ID NO:5 的连续氨基酸 64-471 中所定义的。

5. 权利要求 1 的分离的多核苷酸，其中多核苷酸包含编码与下列具有至少 90%序列同一性的多肽的多核苷酸：如在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5 中所述的氨基酸序列、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481、或 SEQ ID NO:5 的连续氨基酸 64-471，其中多核苷酸编码具有 AHASS 活性的多肽。

6. 权利要求 1 的分离的多核苷酸，其中多核苷酸包含如在 SEQ ID NO:10 或 SEQ ID NO:11 中所示的多核苷酸。

7. 权利要求 1 的(a)到(f)的任意一项的分离的多核苷酸，其中多核苷酸在包含有效连接到多核苷酸的启动子的表达盒中。

8. 权利要求 7 的分离的多核苷酸，其中多核苷酸在植物表达载体中。

9. 权利要求 7 的分离的多核苷酸，其中表达盒还包含编码有效连接到多核苷酸上的叶绿体转运肽的核苷酸序列。

10. 权利要求 7 的分离的多核苷酸，其中启动子能在选自下列的宿主细胞中推动多核苷酸的表达：细菌、真菌细胞、动物细胞和植物细胞。

11. 权利要求 7 的分离的多核苷酸，其中表达盒存在于选自下列的宿主细胞中：细菌、真菌细胞、动物细胞和植物细胞。

12. 权利要求 7 的分离的多核苷酸，其中表达盒在植物中。

13. 权利要求 8 的分离的多核苷酸，其中植物表达载体还包含含有有效连接到编码真核 AHASL 多肽的第二核苷酸序列上的第二启动子的第二多核苷酸构建体，其中两种启动子都能够在植物细胞中推动基因表达。

14. 权利要求 13 的分离的多核苷酸，其中多核苷酸选自：SEQ ID

NO:1、SEQ ID NO:3 中所示的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565; 以及编码如在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481、或 SEQ ID NO:5 的连续氨基酸 64-471 中所示多肽的核苷酸序列。

15. 权利要求 13 的分离的多核苷酸, 其中真核 AHASL 多肽是植物 AHASL 多肽。

16. 权利要求 13 的分离的多核苷酸, 其中真核 AHASL 多肽是除草剂耐受的 AHASL 多肽。

17. 权利要求 13 的分离的多核苷酸, 其中表达载体在植物细胞中。

18. 权利要求 12 的植物表达载体, 其中表达载体在植物中。

19. 包含选自下列的氨基酸序列的分离的多肽:

(a) 如在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481、或 SEQ ID NO:5 的连续氨基酸 64-471 中所述的多肽;

(b) 与下列具有至少 81% 序列同一性的多肽: 如在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5 中所述的氨基酸序列、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481、或 SEQ ID NO:5 的连续氨基酸 64-471, 其中多肽具有乙酰羟酸合酶小亚基(AHASS)活性;

(c) 与 SEQ ID NO:5 的连续氨基酸 64-471 具有至少 77% 序列同一性的多肽, 其中多肽具有 AHASS 活性;

(d) 由多核苷酸所编码的多肽, 所述的多核苷酸与在 SEQ ID NO:1、SEQ ID NO:3 中所述的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565 具有至少 80% 序列同一性, 其中多肽具有 AHASS 活性; 以及

(e) 由多核苷酸所编码的多肽, 所述的多核苷酸在严格条件下杂交在 SEQ ID NO:1、SEQ ID NO:3 中定义的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565, 其中多肽具

有 AHASS 活性。

20. 权利要求 19 的分离的多肽, 其中多肽是在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481, 或 SEQ ID NO:5 的连续氨基酸 64-471 中所述的。

21. 包含含有选自下列的核苷酸序列的多核苷酸构建体的转基因植物细胞:

(a) 在 SEQ ID NO:1、SEQ ID NO:3 中定义的多核苷酸、SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565;

(b) 多核苷酸, 其与在 SEQ ID NO:1、SEQ ID NO:3 中定义的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565 具有至少 80% 的序列同一性, 其中多核苷酸编码具有乙酰羟酸合酶小亚基(AHASS)活性的多肽;

(c) 多核苷酸, 其在严格条件下与在 SEQ ID NO:1、SEQ ID NO:3 中定义的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565 杂交, 其中多核苷酸编码具有 AHASS 活性的多肽;

(d) 编码多肽的多核苷酸, 所述的多肽如在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5 中、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481、或 SEQ ID NO:5 的连续氨基酸 64-471 中所定义的;

(e) 编码与下列具有至少 81% 序列同一性的多肽的多核苷酸: 在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5 中所述的氨基酸序列、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481, 或 SEQ ID NO:5 的连续氨基酸 64-471, 其中多核苷酸编码具有 AHASS 活性的多肽; 以及

(f) 编码与 SEQ ID NO:5 的连续氨基酸 64-471 具有至少 77% 序列同一性的多肽的多核苷酸, 其中多核苷酸编码具有 AHASS 活性的多肽。

22. 权利要求 21 的转基因植物细胞, 其中多核苷酸构建体有效连接到选自下列的启动子: 组成型启动子和组织优选的启动子。

23. 权利要求 21 的转基因植物细胞, 其中多核苷酸构建体还包含有

效连接到第一核苷酸序列的编码叶绿体转运肽的第二核苷酸序列。

24. 权利要求 21 的转基因植物细胞, 其中与野生型品种的植物细胞相比, 转基因植物细胞的 AHAS 活性增加。

25. 权利要求 21 的转基因植物细胞, 其中与野生型品种的植物细胞相比, 转基因植物细胞对至少一种除草剂的耐受性增加。

26. 权利要求 21 的转基因植物细胞, 其中转基因植物细胞是选自下列的单子叶植物细胞: 玉米、小麦、稻、大麦、黑麦、燕麦、黑小麦、粟和高粱。

27. 权利要求 21 的转基因植物细胞, 其中转基因植物细胞是来自选自下列的双子叶植物细胞: 大豆、棉、芸苔属物种、烟草、马铃薯、甜菜、苜蓿、向日葵、红花和落花生。

28. 权利要求 21 的转基因植物细胞, 其中转基因植物细胞在植物中。

29. 权利要求 21 的转基因植物细胞, 其中转基因植物细胞在种子中。

30. 在植物中增强 AHAS 活性的方法, 其包括向植物细胞中引入多核苷酸构建体并且从植物细胞生成比野生型品种的植物具有增加的 AHAS 活性的转基因植物, 其中多核苷酸构建体包含选自下列的核苷酸序列:

(a) 在 SEQ ID NO:1、SEQ ID NO:3 中定义的多核苷酸、SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565;

(b) 多核苷酸, 其与在 SEQ ID NO:1、SEQ ID NO:3 中定义的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565 具有至少 80% 的序列同一性, 其中多核苷酸编码具有乙酰羟酸合酶小亚基(AHASS)活性的多肽;

(c) 多核苷酸, 其在严格条件下与在 SEQ ID NO:1、SEQ ID NO:3 中定义的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565 杂交, 其中多核苷酸编码具有 AHASS 活性的多肽;

(d) 编码多肽的多核苷酸, 所述的多肽如在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481, 或 SEQ ID NO:5 的连续氨基酸 64-471 中所定义的;

(e)编码与下列具有至少 81%序列同一性的多肽的多核苷酸：在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5 中所述的氨基酸序列、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481，或 SEQ ID NO:5 的连续氨基酸 64-471，其中多核苷酸编码具有 AHASS 活性的多肽；以及

(f)编码与 SEQ ID NO:5 的连续氨基酸 64-471 具有至少 77%序列同一性的多肽的多核苷酸，其中多核苷酸编码具有 AHASS 活性的多肽。

31. 权利要求 30 的方法，其中与野生型品种的植物相比，转基因植物对除草剂具有增加的耐受性。

32. 权利要求 31 的方法，其中与野生型品种的植物相比，转基因植物对咪唑啉酮除草剂具有增加的耐受性。

33. 权利要求 30 的方法，其中植物是除草剂耐受的植物。

34. 权利要求 33 的方法，其中植物是咪唑啉酮耐受的植物。

35. 权利要求 33 的方法，其中植物包含除草剂耐受的乙酰羟酸合酶大亚基(AHASL)多肽。

36. 权利要求 30 的方法，其中多核苷酸构建体还包含有效连接到核苷酸序列的启动子，并且其中启动子选自组成型启动子和组织优选的启动子。

37. 权利要求 30 的方法，其中多核苷酸构建体还包含编码除草剂耐受的乙酰羟酸合酶大亚基(AHASL)多肽的多核苷酸序列。

38. 通过这样的方法产生的比野生型品种的植物具有增加的 AHAS 活性的转基因植物，所述的方法包括将多核苷酸构建体引入植物细胞中并且从植物细胞生成转基因植物，其中多核苷酸构建体包含选自下列的核苷酸序列：

(a) 在 SEQ ID NO:1、SEQ ID NO:3 中定义的多核苷酸、SEQ ID NO:1 的连续核苷酸 275-1495，或 SEQ ID NO:3 的连续核苷酸 342-1565；

(b)多核苷酸，其与在 SEQ ID NO:1、SEQ ID NO:3 中定义的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495，或 SEQ ID NO:3 的连续核

核苷酸 342-1565 具有至少 80%的序列同一性, 其中多核苷酸编码具有乙酰羟酸合酶小亚基(AHASS)活性的多肽;

(c)多核苷酸, 其在严格条件下与在 SEQ ID NO:1、SEQ ID NO:3 中定义的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565 杂交, 其中多核苷酸编码具有 AHASS 活性的多肽;

(d)编码多肽的多核苷酸序列, 所述的多肽如在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5 中、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481, 或 SEQ ID NO:5 的连续氨基酸 64-471 中所定义的;

(e)编码与下列具有至少 81%序列同一性的多肽的多核苷酸: 在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5 中所述的氨基酸序列、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481, 或 SEQ ID NO:5 的连续氨基酸 64-471, 其中多核苷酸编码包含 AHASS 活性的多肽; 以及

(f)编码与 SEQ ID NO:5 的连续氨基酸 64-471 具有至少 77%序列同一性的多肽的多核苷酸, 其中多核苷酸编码包含 AHASS 活性的多肽。

39. 权利要求 38 的转基因植物, 其中多核苷酸构建体包含在 SEQ ID NO:1、SEQ ID NO:3 中定义的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565。

40. 权利要求 38 的转基因植物, 其中多核苷酸构建体包含编码多肽的核苷酸序列, 所述的多肽如在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481 或 SEQ ID NO:5 的连续氨基酸 64-471 中所定义的。

41. 在植物附近控制杂草的方法, 其包括将咪唑啉酮除草剂应用到杂草和植物上, 其中与野生型品种的植物相比, 所述植物对咪唑啉酮除草剂具有增加的耐受性, 并且其中所述植物包含含有选自下列的核苷酸序列的多核苷酸构建体:

(a) 在 SEQ ID NO:1、SEQ ID NO:3 中定义的多核苷酸、SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565;



(b)多核苷酸，其与在 SEQ ID NO:1、SEQ ID NO:3 中定义的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495，或 SEQ ID NO:3 的连续核苷酸 342-1565 具有至少 80%的序列同一性，其中多核苷酸编码具有乙酰羟酸合酶小亚基(AHASS)活性的多肽；

(c)多核苷酸，其在严格条件下与在 SEQ ID NO:1、SEQ ID NO:3 中定义的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495，或 SEQ ID NO:3 的连续核苷酸 342-1565 杂交，其中多核苷酸编码具有 AHASS 活性的多肽；

(d)编码多肽的多核苷酸序列，所述的多肽如在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5 中、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481，或 SEQ ID NO:5 的连续氨基酸 64-471 中所定义的；

(e)编码与下列具有至少 81%序列同一性的多肽的多核苷酸：在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5 中所述的氨基酸序列、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481，或 SEQ ID NO:5 的连续氨基酸 64-471，其中多核苷酸编码包含 AHASS 活性的多肽；以及

(f)编码与 SEQ ID NO:5 的连续氨基酸 64-471 具有至少 77%序列同一性的多肽的多核苷酸，其中多核苷酸编码包含 AHASS 活性的多肽。

42. 包含有效连接到乙酰羟酸合酶小亚基(AHASS)结构域的乙酰羟酸合酶大亚基(AHASL)结构域的融合多肽；其中融合多肽包含 AHAS 活性，其中 AHASL 结构域包含成熟真核 AHASL 多肽的氨基酸序列，并且其中 AHASS 结构域包含选自下列的 AHASS 多肽的氨基酸序列：

(a)如在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481、或 SEQ ID NO:5 的连续氨基酸 64-471 中所述的多肽；

(b)与下列具有至少 81%序列同一性的多肽：如在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5 中所述的氨基酸序列、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481、或 SEQ ID NO:5 的连续氨基酸 64-471，其中多肽具有乙酰羟酸合酶小亚基(AHASS)活性；

(c)与 SEQ ID NO:5 的连续氨基酸 64-471 具有至少 77%序列同一性的多肽, 其中多肽具有 AHASS 活性;

(d)由多核苷酸所编码的多肽, 所述的多核苷酸与在 SEQ ID NO:1、SEQ ID NO:3 中所述的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565 具有至少 80%序列同一性, 其中多肽具有 AHASS 活性; 以及

(e)由多核苷酸所编码的多肽, 所述的多核苷酸在严格条件下杂交在 SEQ ID NO:1、SEQ ID NO:3 中定义的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565, 其中多肽具有 AHASS 活性。

43. 权利要求 42 的融合多肽, 其中真核 AHASL 多肽是植物 AHASL 多肽。

44. 权利要求 42 的融合多肽, 其还包含有效连接在 AHASL 结构域和 AHASS 结构域之间的连接区。

45. 权利要求 42 的融合多肽, 其中 AHASL 多肽和 AHASS 多肽来自不同物种。

46. 分离的多核苷酸, 其中多核苷酸编码乙酰羟酸合酶大亚基(AHASL)-乙酰羟酸合酶小亚基(AHASS)融合多肽, 其中 AHASL 是真核 AHASL 多肽, 并且其中 AHASS 包含选自下列的氨基酸序列:

(a)如在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481、或 SEQ ID NO:5 的连续氨基酸 64-471 中所述的多肽;

(b)与下列具有至少 81%序列同一性的多肽: 如在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5 中所述的氨基酸序列、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481、或 SEQ ID NO:5 的连续氨基酸 64-471, 其中多肽具有乙酰羟酸合酶小亚基(AHASS)活性;

(c)与 SEQ ID NO:5 的连续氨基酸 64-471 具有至少 77%序列同一性的多肽, 其中多肽具有 AHASS 活性;

(d)由多核苷酸所编码的多肽,所述的多核苷酸与在 SEQ ID NO:1、SEQ ID NO:3 中所述的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495,或 SEQ ID NO:3 的连续核苷酸 342-1565 具有至少 80%序列同一性,其中多肽具有 AHASS 活性;以及

(e)由多核苷酸所编码的多肽,所述的多核苷酸在严格条件下杂交在 SEQ ID NO:1、SEQ ID NO:3 中定义的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495,或 SEQ ID NO:3 的连续核苷酸 342-1565,其中多肽具有 AHASS 活性。

47. 权利要求 46 的分离的多核苷酸,其中 AHASS 包含在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5 中所述的氨基酸序列、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481,或 SEQ ID NO:5 的连续氨基酸 64-471。

48. 权利要求 46 的分离的多核苷酸,其中多核苷酸还包含有效连接的编码连接区的第三核苷酸序列。

49. 权利要求 46 的分离的多核苷酸,其中多核苷酸还包含有效连接到多核苷酸的叶绿体靶向序列。

50. 权利要求 46 的分离的多核苷酸,其中真核 AHASL 多肽是植物 AHASL 多肽。

51. 权利要求 46 的分离的多核苷酸,其中真核 AHASL 多肽是除草剂耐受的 AHASL 多肽。

52. 权利要求 46 的分离的多核苷酸,其中多核苷酸在植物表达载体中。

53. 权利要求 46 的分离的多核苷酸,其中多核苷酸在植物细胞中。

54. 权利要求 46 的分离的多核苷酸,其中多核苷酸在种子中。

55. 产生具有增加的 AHAS 活性的转基因植物的方法,其包括将多核苷酸构建体引入植物细胞中并且从转基因植物细胞生成比野生型品种的植物具有增加的 AHAS 活性的转基因植物,其中多核苷酸构建体编码乙酰羟酸合酶大亚基(AHASL)-乙酰羟酸合酶小亚基(AHASS)融合多肽,其中

AHASL 是真核 AHASL 多肽, 并且其中 AHASS 包含选自下列的氨基酸序列:

(a) 如在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481、或 SEQ ID NO:5 的连续氨基酸 64-471 中所述的多肽;

(b) 与下列具有至少 81% 序列同一性的多肽: 如在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5 中所述的氨基酸序列、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481、或 SEQ ID NO:5 的连续氨基酸 64-471, 其中多肽具有乙酰羟酸合酶小亚基(AHASS)活性;

(c) 与 SEQ ID NO:5 的连续氨基酸 64-471 具有至少 77% 序列同一性的多肽, 其中多肽具有 AHASS 活性;

(d) 由多核苷酸所编码的多肽, 所述的多核苷酸与在 SEQ ID NO:1、SEQ ID NO:3 中所述的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565 具有至少 80% 序列同一性, 其中多肽具有 AHASS 活性; 以及

(e) 由多核苷酸所编码的多肽, 所述的多核苷酸在严格条件下杂交在 SEQ ID NO:1、SEQ ID NO:3 中定义的核苷酸序列、SEQ ID NO:1 的连续核苷酸 275-1495, 或 SEQ ID NO:3 的连续核苷酸 342-1565, 其中多肽具有 AHASS 活性。

56. 权利要求 55 的方法, 其中 AHASS 多肽在 SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481 或 SEQ ID NO:5 的连续氨基酸 64-471 中定义。

57. 权利要求 55 的方法, 其中与野生型品种植物相比, 转基因植物对咪唑啉酮除草剂具有增加的耐受性。

## 单子叶植物 AHAS 序列和使用方法

### 发明领域

本发明涉及编码乙酰羟酸合酶小亚基并且可以用来增强作物植物的乙酰羟酸合酶活性和除草剂耐受性的新多核苷酸。

### 发明背景

乙酰羟酸合酶(AHAS; EC 4.1.3.18, 也已知为乙酰乳酸合酶或 ALS)是催化支链氨基酸缬氨酸、亮氨酸和异亮氨酸的生物化学合成的第一个酶(Singh, 1999, "Biosynthesis of valine, leucine and isoleucine," in *Plant Amino Acids*, Singh, ed., Marcel Dekker Inc. New York, New York, pp. 227-247)。AHAS 是四个结构不同的除草剂家族作用位点, 所述除草剂家族包括磺酰脲(LaRossa 和 Falco, 1984, *Trends Biotechnol.* 2:158-161)、咪唑啉酮(Shaver 等人, 1984, *Plant Physiol.* 76:545-546)、三唑并嘧啶(Subramanian 和 Gerwick, 1989, "Inhibition of acetolactate synthase by triazolopyrimidines," 于 *Biocatalysis in Agricultural Biotechnology* 一书中, Whitaker 和 Sonnet 编著, ACS Symposium Series, American Chemical Society, Washington, D.C., pp. 277-288)和嘧啶基氧基苯甲酸酯类(pyrimidyloxybenzoates) (Subramanian 等人, 1990, *Plant Physiol.* 94:239-244)。由于咪唑啉酮和磺酰脲除草剂在极低应用率下的效力和在动物中的相对无毒性, 它们在现代农业中被广泛应用。通过抑制 AHAS 活性, 这些除草剂家族阻止包括许多杂草物种在内的易感植物的生长和发育。可商购的咪唑啉酮除草剂的一些实例是 PURSUIT® (咪草烟)、SCEPTER® (灭草啶)和 ARSENAL® (灭草烟)。磺酰脲除草剂的实例是绿黄隆、甲黄隆、嘧黄隆、氟嘧黄隆、噻黄隆、苯黄隆、禾草芊混剂、烟嘧黄隆、胺苯黄隆、玉嘧黄隆、氟胺黄隆、醚苯黄隆、氟嘧黄隆、醚黄隆、磺氨黄隆

(amidosulfuron)、fluzasulfuron、啶咪黄隆、吡嘧黄隆和吡氯黄隆。

由于咪唑啉酮和磺酰脲的高效力和低毒性，可以通过喷在广泛区域植被顶部来促成应用它们。将除草剂喷到广泛区域植被顶部的能力减少与耕地建设和保持相关的费用，并且减少了在使用此类化学制剂前现场准备的需要。在所需耐受物种顶部喷射也使得由于竞争物种不存在而能够实现所需物种最大产量的潜力。但是，使用此类在上喷射技术的能力取决于在喷射区域中所需植被的抗咪唑啉酮物种的存在。

在主要的农业作物中，一些豆科物种如大豆对咪唑啉酮除草剂有天然抗性，这是由于它们能够迅速代谢除草剂化合物 (Shaner 和 Robson, 1985, *Weed Sci.* 33:469-471)。其它作物如玉米(Newhouse 等人, 1992, *Plant Physiol.* 100:882886)和稻(Barrette 等人, 1989, *Crop Safeners for Herbicides*, Academic Press, New York, pp. 195-220)对咪唑啉酮除草剂有一定的易感。对咪唑啉酮除草剂的差别易感性取决于特定除草剂的化学性质和在每种植物中化合物从有毒到无毒形式的差别代谢(Shaner 等人, 1984, *Plant Physiol.* 76:545-546; Brown *et al.*, 1987, *Pestic. Biochem. Physiol.* 27:24-29)。其它植物生理学差异如吸收和运输也在敏感性中起到重要作用(Shaner 和 Robson, 1985, *Weed Sci.* 33:469-471)。

对咪唑啉酮、磺酰脲和三唑并嘧啶有抗性的作物栽培种已经通过在玉米(*Zea mays*)、拟南芥(*Arabidopsis thaliana*)、油菜(*Brassica napus*)、大豆(*Glycine max*)和烟草(*Nicotiana tabacum*)中使用种子、小孢子、花粉和愈伤组织诱变成功产生(Sebastian 等人, 1989, *Crop Sci.* 29:1403-1408; Swanson 等人, 1989, *Theor. Appl. Genet.* 78:525-530; Newhouse 等人, 1991, *Theor. Appl. Genet.* 83:65-70; Sathasivan 等人, 1991, *Plant Physiol.* 97:1044-1050; Mourand 等人, 1993, *J. Heredity* 84:91-96)。在所有的情形中，单一的、部分显性的核基因赋予抗性。在普通小麦(*Triticum aestivum* L. cv. Fidel)的种子诱变后也预先分离出 4 种抗咪唑啉酮的小麦植物(Newhouse 等人, 1992, *Plant Physiol.* 100:882-886)。遗传研究证明单一的、部分显性的基因赋予抗性。基于等位研究，作者得出结论，在 4 个经鉴定品系中的突变位于相同

的基因座。Fidel 栽培种抗性基因之一被称作 FS-4(Newhouse 等人, 1992, *Plant Physiol.* 100:882-886)。

对咪唑啉酮除草剂有抗性的植物也在一些专利中进行了报道。美国专利号 4,761,373、5,331,107、5,304,732、6,211,438、6,211,439 和 6,222,100 总体上描述了改变的 AHAS 基因在植物中引起除草剂抗性的用途, 并且具体公开了某些抗咪唑啉酮的玉米品系。美国专利号 5,013,659 公开了由于在一个或多个保守性区域中在至少一个氨基酸中的突变, 而表现出除草剂抗性的植物。在本文描述的突变编码对咪唑啉酮和磺酰脲的交叉抗性或对磺酰脲特异性的抗性, 但是没有描述对咪唑啉酮特异性的抗性。另外, 美国专利号 5,731,180 和美国专利号 5,767,361 讨论了在野生型单子叶植物 AHAS 氨基酸序列中具有单氨基酸替代(产生咪唑啉酮特异性抗性)的分离的基因。另外, 对于干扰乙酰羟酸合酶的除草剂有抗性的稻植物可以通过突变育种并且还通过其它培养来发育(见美国专利号 5,545,822、5,736,629、5,773,703、5,773,704、5,952,553 和 6,274,796)。

在植物中, AHAS 酶由两个亚基组成: 大亚基(催化作用)和小亚基(调节作用)(Duggleby 和 Pang, 2000, *J. Biochem. Mol. Biol.* 33:1-36)。AHAS 大亚基蛋白质(命名为 AHASL)在拟南芥和稻中是由单一基因编码的, 或者在玉米、油菜和棉中是由多个基因家族成员编码的。AHASL 中特异、单一的核苷酸替代使得酶对一类或多类除草剂具有一定程度的不敏感性(Chang 和 Duggleby, 1998, *Biochem J.* 333:765-777)。

也已经合理设计了抗除草剂的 AHASL 基因。WO 96/33270、美国专利号 5,853,973 和 5,928,937 公开了用于 AHAS 变体制备的基于结构的建模方法, 所述的 AHAS 变体包括对除草剂如咪唑啉酮和 AHAS 抑制性除草剂表现出选择性增加的抗性那些。AHAS 抑制剂复合物三维构象的基于计算机的建模预测了在所提出的抑制剂结合袋中的几个氨基酸作为所引入的突变可能赋予对咪唑啉酮的选择性抗性的位点(Ott 等人, 1996, *J. Mol. Biol.* 263:359-368)。使用在 AHAS 酶的所提出结合位点中这些合理设计的突变中的一些突变产生的小麦植物事实上表现出对单一类除草剂的特异性抗性

(Ott 等人, 1996, *J. Mol. Biol.* 263:359-368)。

从原核系统中的研究了解了关于 AHAS 酶功能的大量方面。这些研究使 AHAS 小亚基(AHASS)蛋白质的作用清楚明白的显示出来。原核 AHAS 酶以两种不同但是物理相关的蛋白质亚基存在。在原核生物中, 两种多肽“大亚基”和“小亚基”自分别的基因表达。已经在肠细菌中鉴定出了三种主要的 AHAS 酶, 称作 I、II 和 III, 都具有大亚基和小亚基。在原核生物中, 已经表明在分支的氨基酸生物合成途径中 AHAS 酶是调节酶(Miflin, 1971, *Arch. Biochem. Biophys.* 146:542-550), 并且已经观察到只有大亚基具有催化活性。从微生物系统的 AHAS 酶的研究中, 已经描述了小亚基的两种作用。一种作用是当异亮氨酸、亮氨酸、或缬氨酸或其组合存在时, 对催化的大亚基的变构反馈抑制。另一作用是在异亮氨酸、亮氨酸或缬氨酸不存在时对大亚基催化活性的增强。也已经表明小亚基增加大亚基活性构象的稳定性(Weinstock 等人, 1992, *J. Bacteriol.* 174:5560-5566)。如对于来自大肠杆菌(*E. coli*)的 AHAS I 所见, 小亚基的表达还可以增加大亚基的表达(Weinstock 等人, 1992, *J. Bacteriol.* 174:5560-5566)。

体外研究已经证明了在小亚基不存在时原核大亚基表现出 AHAS 活性的基本水平并且该活性不会被氨基酸异亮氨酸、亮氨酸或缬氨酸反馈抑制。当将小亚基添加到含有大亚基的同一反应混合物中时, 大亚基的特异性活性增加。

虽然还已知 AHAS 的小亚基存在于植物中, 但是关于其体内功能知之甚少。WO 98/37206 公开了编码来自野生烟草(*Nicotiana plumbaginifolia*)的 AHASS cDNA 序列的核苷酸序列以及此序列在筛选抑制 AHAS 全酶活性的除草剂中的用途。另外, WO 98/37206 公开了玉米 AHASS 蛋白质的部分长度 cDNA 序列。美国专利号 6,348,643 公开了来自拟南芥的全长 AHASS 蛋白质的核苷酸和氨基酸序列。该专利另外公开了通过添加拟南芥 AHASS 蛋白质对拟南芥 AHASL 蛋白质的野生型和抗除草剂形式的激活。激活是通过公开拟南芥 AHASS 蛋白质在增加 AHASL 蛋白质野生型和抗除草剂形式的特异性 AHAS 活性上的能力来证明的。更近期, 美国



专利公开 No. 2001/0044939 报道了用非物种特异性的 AHASL 蛋白质重组天然植物 AHASS 蛋白质的有利效果, 如通过野生烟草 AHASS 蛋白质增加来自其它双子叶植物(拟南芥)的 AHASL 蛋白质特异性活性的能力所表明的。

## 发明概述

本发明提供了编码玉米、稻和小麦乙酰羟酸合酶小亚基(AHASS)多肽的分离的多核苷酸, 所述 AHASS 多肽在本文中分别称作玉米 AHAS 小亚基亚型 1 种内同源物 a(ZmAHASS1a)、稻(*Oryza sativa*)AHAS 小亚基亚型 1(OsAHASS1), 和普通小麦 AHAS 小亚基亚型 1(TaAHASS1X)。本发明的多核苷酸包含选自下列的核苷酸序列: 在 SEQ ID NO:1 和 3 中所示的核苷酸序列, 和编码在 SEQ ID NO:2、4 和 5 中所示的氨基酸序列的核苷酸序列, 以及编码含有 AHASS 活性的多肽的核苷酸序列的片段和变体。

在一个实施方案中。本发明的多核苷酸包含 SEQ ID NO:1 的连续核苷酸 275 - 1495 或 SEQ ID NO:3 的连续核苷酸 342 - 1565。在另一个实施方案中, 本发明的多核苷酸与在 SEQ ID NO:1 或 SEQ ID NO:3 中所示的核苷酸序列, 或与 SEQ ID NO:1 的连续核苷酸 275 - 1495 或 SEQ ID NO:3 的连续核苷酸 342 - 1565 具有至少 80% 的序列同一性, 其中此类多核苷酸编码具有 AHASS 活性的多肽。本发明分离的多核苷酸还包括编码本发明 AHASS 多肽成熟形式的多核苷酸。AHASS 多肽的此类成熟形式没有位于 N 末端的叶绿体转运肽。

本发明还提供了包含稻 AHASS 启动子的多核苷酸序列。本领域的技术人员将认识到此多核苷酸包含稻 AHASS 基因的转录起始位点上游的稻基因组区域, 人们可以操作所述的区域以生成在植物中仍然起作用的最小长度的启动子。包含此启动子的稻基因组片段在 SEQ ID NO:10 中示出。

本发明另外提供了包含稻 AHASS 终止子的多核苷酸序列。本领域的技术人员将认识到此多核苷酸包含稻 AHASS 基因的翻译终止密码子下游的稻基因组区域, 人们可以操作所述的区域以生成在植物中仍然起作用的

最小长度的终止子。包含此终止子的稻基因组片段在 SEQ ID NO:11 中示出。

本发明还提供了用来在植物、植物细胞和其它非人宿主细胞(包括但不限于细菌、真菌细胞和动物细胞)中表达本发明多核苷酸的表达盒。表达盒包含在目的植物、植物细胞或其它宿主细胞中可表达的启动子,其有效连接至编码全长 AHASS 多肽(即包括叶绿体转运肽)或成熟 AHASS 多肽(即无叶绿体转运肽)的本发明多核苷酸上。如果在植物或植物细胞的质体中需要表达,表达盒可以另外包含有效连接的编码叶绿体转运肽的叶绿体靶向序列。

本发明另外提供了用于在目的植物或宿主细胞中表达真核 AHASL 多肽和 AHASS 多肽的植物表达载体。在一个实施方案中,植物表达载体包含第一多核苷酸构建体和第二多核苷酸构建体,其中第一多核苷酸构建体包含有效连接了编码真核 AHASL 多肽的核苷酸序列的第一启动子,其中第二多核苷酸构建体包含有效连接了编码 AHASS 多肽的核苷酸序列的第二启动子,并且其中第一和第二启动子能够在目的植物或宿主细胞中推动基因表达。在一个实施方案中,第一和第二多核苷酸构建体另外包含有效连接的叶绿体靶向序列。在另一个实施方案中,真核 AHASL 多肽是植物 AHASL 多肽,并且在一些情况中是除草剂耐受的 AHASL 多肽。

本发明提供了包含 AHASS 多肽的分离的多肽。分离的多肽包含选自下列的氨基酸序列:在 SEQ ID NO:2、4 和 5 中所示的氨基酸序列,由在 SEQ ID NO:1 和 3 中所示的核苷酸序列编码的氨基酸序列,以及编码包含 AHASS 活性的多肽的氨基酸序列的片段和变体。此类片段包括但不限于本发明 AHASS 多肽的成熟形式,尤其是选自下列的氨基酸序列:在 SEQ ID NO:2 中所示的氨基酸序列的氨基酸 77-483、在 SEQ ID NO:4 中所示的氨基酸序列的氨基酸 74-481、在 SEQ ID NO:5 中所示的氨基酸序列的氨基酸 64-471、由在 SEQ ID NO:1 中所示的核苷酸序列的核苷酸 275-1495 编码的氨基酸序列,以及由在 SEQ ID NO:3 中所示的核苷酸序列的核苷酸 342-1565 所编码的氨基酸序列。本发明还提供了与在 SEQ ID NO:2、4 或

5 中所示的氨基酸序列具有至少 81%序列同一性的多肽，或与 SEQ ID NO:5 的连续氨基酸 64-471 具有至少 77%序列同一性的多肽，其中此类多肽包含 AHASS 活性。

本发明另外提供了包含本发明 AHASS 多核苷酸的转基因植物、种子和转基因植物细胞。在一个实施方案中，AHASS 多核苷酸有效连接至植物细胞中推动其表达的启动子。在另一个实施方案中，启动子是组成型启动子或组织优选的启动子。在另一个实施方案中，多核苷酸构建体另外包含有效连接 AHASS 多核苷酸的叶绿体靶向序列。在一个实施方案中，转基因植物是选自下列的单子叶植物：玉米、小麦、稻、大麦、黑麦、燕麦、黑小麦、粟和高粱。在另一个实施方案中，转基因植物是选自下列的双子叶植物：大豆、棉、芸苔属物种(*Brassica spp.*)、烟草、马铃薯、甜菜、苜蓿、向日葵、红花和落花生。优选的，包含本发明 AHASS 多核苷酸的这些转基因植物、种子和植物细胞具有 AHAS 活性和/或对至少一种除草剂的抗性，所述抗性与野生型品种植物相比增加。

本发明提供了在植物中增强 AHAS 活性的方法，包括用本发明的 AHASS 多核苷酸转化植物。在一个实施方案中，AHASS 多核苷酸在包含启动子的表达盒中，所述的启动子有效连接至 AHASS 核苷酸序列，能够在植物细胞中推动基因表达。在另一个实施方案中，启动子是组成型启动子或组织优选的启动子。在另一个实施方案中，植物包含除草剂耐受的乙酰羟酸合酶大亚基(AHASL)多肽。本发明的方法可以用来增强或增加植物对至少一种干扰 AHAS 酶催化活性的除草剂的抗性。还提供了由这些方法产生的转基因植物，其中与野生型品种的植物相比，在此类转基因植物中 AHAS 活性增加。

本发明还提供了用于在除草剂耐受的植物中增强除草剂耐受性的方法，包括用本发明的 AHASS 多核苷酸转化植物。在一个实施方案中，AHASS 多核苷酸在包含有启动子的表达盒中，所述的启动子有效连接至 AHASS 核苷酸序列，能够在植物细胞中推动基因表达。在另一个实施方案中，启动子是组成型启动子或组织优选的启动子。在一个实施方案中，

AHASS 多核苷酸构建体另外包含编码除草剂耐受的 AHASL 多肽的核苷酸序列。在另一个实施方案中，除草剂耐受的植物包含 AHASL 多肽。在另一个实施方案中，对除草剂耐受的植物进行或没进行遗传改造以表达除草剂耐受的 AHASL 多肽。在另一个实施方案中，除草剂耐受的植物是咪唑啉酮耐受的植物。还提供了由这些方法产生的转基因植物，其中与野生型品种的植物相比，在此类转基因植物中 AHAS 活性增加。本发明还提供了在植物的附近控制杂草的方法，包括对杂草和植物应用咪唑啉酮除草剂，其中与野生型品种的植物相比，植物对咪唑啉酮除草剂具有增加的耐受性，并且其中植物包含含有本发明 AHASS 核苷酸序列的多核苷酸构建体。在一个实施方案中，AHASS 核苷酸序列在 SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:1 的连续核苷酸 275-1495，或 SEQ ID NO:3 的连续核苷酸 342-1565 中定义。在另一个实施方案中，AHASS 核苷酸包含编码如在下列中所定义多肽的多核苷酸：SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:2 的连续氨基酸 77-483、SEQ ID NO:4 的连续氨基酸 74-481，或 SEQ ID NO:5 的连续氨基酸 64-471。

本发明另外提供了包含有效连接至 AHASS 结构域的 AHASL 结构域的融合多肽，其中融合多肽包含 AHAS 活性。AHASL 结构域包含成熟真核 AHASL 多肽的氨基酸序列。AHASS 结构域包含选自下列的氨基酸序列：在 SEQ ID NO:2、4 和 5 中所示的氨基酸序列；由 SEQ ID NO:1 和 3 中所示的核苷酸序列编码的氨基酸序列；以及编码含有 AHASS 活性的多肽的氨基酸序列的片段和变体。此类片段包括但不限于本发明 AHASS 多肽的成熟形式，尤其是选自下列的氨基酸序列：在 SEQ ID NO:2 中所示的氨基酸序列的氨基酸 77-483、在 SEQ ID NO:4 中所示的氨基酸序列的氨基酸 74-481、在 SEQ ID NO:5 中所示的氨基酸序列的氨基酸 64-471、由在 SEQ ID NO:1 中所示的核苷酸序列的核苷酸 275-1495 和在 SEQ ID NO:3 中所示的核苷酸序列的核苷酸 342-1565 所编码的氨基酸序列。在一个实施方案中，真核 AHASL 多肽是植物 AHASL 多肽。在另一个实施方案中，真核 AHASL 多肽是除草剂耐受的植物 AHASL 多肽。在另一个实施方案

中，融合多肽另外包含在 AHASL 结构域和 AHASS 结构域之间有效连接的连接区。优选的，AHASL 多肽和 AHASS 多肽来自不同的物种。

本发明还提供了用于在目的植物或宿主细胞中表达 AHASL-AHASS 融合多肽的表达载体。表达载体包含有效连接至编码 AHASL-AHASS 融合多肽的多核苷酸的启动子。多核苷酸包含有效连接了第二核苷酸序列的第一核苷酸序列，其中第一核苷酸序列编码包含真核成熟 AHASL 多肽的氨基酸序列并且第二核苷酸序列编码包含本发明成熟 AHASS 多肽的氨基酸序列。多核苷酸可以另外包含编码位于融合多肽的 AHASL 和 AHASS 结构域之间的连接区的有效连接的第三核苷酸序列。在另一个实施方案中，融合多肽的真核 AHASL 结构域是植物 AHASL 多肽。在另一个实施方案中，真核 AHASL 多肽是除草剂耐受的植物 AHASL 多肽。

本发明另外提供了包含编码 AHASL-AHASS 融合多肽的多核苷酸的转基因植物、种子和植物细胞。还提供的是用于产生除草剂耐受植物的方法，包括用包含有效连接至编码 AHASL-AHASS 融合多肽的多核苷酸的启动子的表达载体转化植物细胞，并且从转基因植物细胞生成转基因植物，其中与野生型品种的植物相比，包含 AHASL-AHASS 融合多肽的转基因植物对至少一种除草剂具有增加的耐受性。

## 附图简述

图 1 是下列本发明成熟 AHASS 多肽的氨基酸序列比对，所述的多肽：**ZmAHASS1a** (SEQ ID NO:2 的残基 77-483)、**OsAHASS1** (SEQ ID NO:4 的残基 74-481)，和 **TaAHASS1X** (SEQ ID NO:5 的残基 64-471)。使用 Clustal X 版本 1.81，多比对模式 (Multiple Alignment Mode) 对上述推论氨基酸序列(减去预测的可变叶绿体转运肽)进行比对。使用默认参数来重复进行完全比对(至少 3 次)。“\*” 指示了氨基酸序列在所有序列中是一致的。“:” 和 “.” 是递减的保守性替代。保守的结构域 1 和结构域 2 区域以粗体指示。结构域 1 在 N 端，并且结构域 2 在 C 端。存在位于结构域 1 和 2 之间的间插可变连接区。

图 2 提供了来自成熟 AHASS 多肽配对比校的百分比氨基酸序列同一性。比较包括所有公开已知的植物 AHASS 序列和 ZmAHASS1a (SEQ ID NO:2)、OsAHASS1 (SEQ ID NO:4)和 TaAHASS1X (SEQ ID NO:5)的本发明氨基酸序列。使用 ClustalW 算法对来自所有公开基因编码序列的推论氨基酸序列和其它推定全长序列进行比对。基于此算法计算配对差异。数据以两个序列间的百分比序列同一性的形式来显示。术语：“GmAHASS1”指的是大豆 AHAS 小亚基亚型 1(美国专利申请公开号 2001/00044039A1 的 SEQ ID NO:18); “NpAHASS1”指的是野生烟草 AHAS 小亚基亚型 1(登录号 AJ234901.1); “ZmAHASS2”指的是玉米 AHAS 小亚基亚型 2(美国专利申请公开号 2001/00044039A1 的 SEQ ID NO:10); “OsAHASS2”指的是稻 AHAS 小亚基亚型 2(美国专利申请公开号 2001/00044039A1 的 SEQ ID NO:16); “AtAHASS1”指的是拟南芥 AHAS 小亚基亚型 1 (NM\_179843.1); 并且“AtAHASS2”指的是拟南芥 AHAS 小亚基亚型 2 (NM\_121634.2)。

图 3 提供了来自 AHASS 多肽结构域 1 的配对比校的百分比氨基酸序列同一性。该比较包括来自所有公开已知植物 AHASS 序列的结构域 1 和来自 ZmAHASS1a (SEQ ID NO:2)、OsAHASS1 (SEQ ID NO:4)和 TaAHASS1X (SEQ ID NO:5)的本发明氨基酸序列的结构域 1。氨基酸序列的术语和百分比氨基酸序列同一性如在上面对图 2 所描述的，只是只有对应于结构域 1 的氨基酸序列用在确定百分比序列同一性中。

图 4 提供了来自 AHASS 多肽的结构域 2 配对比校的百分比氨基酸序列同一性。比较包括来自所有公开已知植物 AHASS 序列的结构域 2 和来自 ZmAHASS1a (SEQ ID NO:2)、OsAHASS1 (SEQ ID NO:4)和 TaAHASS1X (SEQ ID NO:5)的本发明氨基酸序列的结构域 2。氨基酸序列的术语和百分比氨基酸序列同一性如在上面对图 2 所描述的，只是只有对应于结构域 2 的氨基酸序列用在确定百分比序列同一性中。结构域 1 和 2 是根据经验从已知 AHASS 多肽的氨基酸序列中确定的。每个结构域包含 ACT 结构域。已知的植物 AHASS 多肽具有类细菌 AHASS 多肽的两个重复序列。尽管作为祖代复制结果是可能的，现在“重复序列”完全不同于

彼此并且在本文中称为结构域 1 和 2。

图 5 提供了 *OsAHASS1* (SEQ ID NO:4)的氨基酸序列与可从 The Institute for Genomic Research (TIGR)获得的 *OsAHASS1* 基因组 DNA 注释的翻译(SEQ ID NO:12)比对。在两个氨基酸序列中对应位置上一致的氨基酸是加上阴影的。还提供了共有序列。

图 6 描述了两个 EST 的重叠序列和一个专有重叠群的比对和区域,该重叠群用来构建全长 *OsAHASS1* 核苷酸序列(SEQ ID NO:3)。

### 发明详述

本发明涉及分离的多核苷酸分子,其包含编码乙酰羧酸合酶小亚基(AHASS)多肽的核苷酸序列。具体地,本发明涉及分离的多核苷酸分子,其编码来自玉米、稻和小麦的单子叶植物 AHASS 多肽,在本文中它们分别称为 *ZmAHASS1a*、*OsAHASS1* 和 *TaAHASS1X*。更具体的,本发明涉及分离的多核苷酸分子,其包含选自下列的多核苷酸序列:在 SEQ ID NO:1 或 SEQ ID NO:3 中所述的核苷酸序列、编码在 SEQ ID NO:2、4 和 5 中所述的 AHASS 多肽的核苷酸序列,以及编码功能性 AHASS 多肽的此类核苷酸序列的片段和变体。

另外,本发明提供了编码成熟 *ZmAHASS1a*、*OsAHASS1* 或 *TaAHASS1X* 多肽的分离的多核苷酸。本发明成熟 AHASS 多肽缺乏在 *ZmAHASS1a*、*OsAHASS1* 和 *TaAHASS1X* 多肽各自的 N 末端发现的叶绿体转运肽,但是保留了 AHASS 活性。具体地,本发明的多核苷酸包含选自下列的核苷酸序列:在 SEQ ID NO:1 中所示核苷酸序列的核苷酸 275-1495、在 SEQ ID NO:3 中所示核苷酸序列的核苷酸 342-1565、编码在 SEQ ID NO:2 中所示氨基酸序列的氨基酸 77-483 的核苷酸序列、编码在 SEQ ID NO:4 中所示氨基酸序列的氨基酸 64-471 的核苷酸序列、编码在 SEQ ID NO:5 中所示氨基酸序列的氨基酸 74-481 的核苷酸序列,以及编码包含 AHASS 活性的成熟 AHASS 多肽的这些核苷酸序列的片段和变体。

如本文所用到的,除非另外说明,“AHASS 活性”指的是 AHASS 多

肽的生物学活性，借此，与 AHASS 多肽不存在时 AHASL 多肽的 AHAS 活性相比，当该 AHASS 和 AHASL 多肽彼此存在时，AHASS 多肽增加了至少一种 AHASL 多肽的 AHAS 活性。

本发明分离的 AHASS 多核苷酸分子可以用来转化作物植物以增强作物植物对除草剂的耐受性，所述的除草剂尤其是已知会抑制 AHAS 活性的除草剂，并且具体的是咪唑啉酮和磺酰脲除草剂。此类 AHASS 多核苷酸可以用在表达盒、表达载体、转化载体、质粒等中。在用此类多核苷酸构建体转化后得到的转基因植物对 AHAS 抑制性除草剂如咪唑啉酮和磺酰脲除草剂表现出增加的耐受性。如本文用到的，术语“耐受性”和“抗性”可以互换使用并且指的是植物抵抗这种水平除草剂效应的能力，在所述的水平上通常杀死野生型品种的植物或抑制其生长。如本文用到的，植物的“野生型品种”指的是作为对照植物分析用于对比目的的植物组，其中除了野生型品种的植物没有用 AHASS 多核苷酸转化和/或在野生型品种植物中 AHASS 多核苷酸的表达没有改变之外，野生型品种的植物与测试植物(用 AHASS 多核苷酸转化的植物或其中 AHASS 多核苷酸表达被改变的植物)一致。因此，使用术语“野生型品种”植物不用于暗指植物在其基因组中没有重组 DNA。

本发明的组合物包括编码 AHASS 多肽的核苷酸序列。具体地，本发明提供了分离的多核苷酸分子(在本文中也称为“核酸分子”)，其包含编码如在 SEQ ID NO:2、4 和 5 中所示氨基酸序列的核苷酸序列。另外提供了多肽，其具有由本文所描述的多核苷酸分子(例如在 SEQ ID NO:1 和 3 中所示的那些)编码的氨基酸序列，以及其片段和变体。

本发明包括分离的或基本纯化的核酸或多核苷酸组合物。“分离的”或“纯化的”多核苷酸分子或多肽，或其生物学活性部分基本上没有通常与其天然存在环境中发现的多核苷酸分子或多肽伴行或相互作用的成分。因此，分离的或纯化的多核苷酸分子或多肽当通过重组技术产生时基本上没有其它细胞材料或培养基，或者当化学合成时基本上没有化学前体或其它化学品。优选的，“分离的”核酸没有在此核酸所来源生物的基因组中的



核酸(即位于核酸的5'和3'端的序列)天然侧翼的序列(优选的是多肽编码序列)。例如,在多个实施方案中,分离的多核苷酸分子可以包含少于大约5kb、4kb、3kb、2kb、1kb、0.5kb或0.1kb的核苷酸序列,该核苷酸序列是来自此核酸所来源细胞基因组DNA中多核苷酸分子的天然侧翼。基本上没有细胞材料的多肽包括具有大约少于30%、20%、10%、5%或1%(干重)污染多肽的多肽制品。当本发明的多肽或其生物学活性部分是重组产生的,培养基优选的少于大约30%、20%、10%、5%或1%(干重)的化学前体或非目的多肽的化学品。

本发明提供了包含AHASS多肽(ZmAHASS1a、OsAHASS1和TaAHASS1X)的分离的多肽。如本文所用到的,术语“蛋白质”和“多肽”可以互换使用,指的是由肽键连接的至少4个氨基酸的链。链可以是线性的、分支的、环状的或其组合。分离的多肽可以包含选自下列的氨基酸序列:在SEQ ID NO:2、4和5中所示的氨基酸序列;由在SEQ ID NO:1和3中所示的核苷酸序列编码的氨基酸序列;编码包含AHASS活性的AHASS多肽的氨基酸序列的功能性片段和变体。术语“功能性片段和变体”指的是包含AHASS活性的例示性多肽的片段和变体。

另外提供的是包含本发明AHASS多肽成熟形式的分离的多肽。此类分离的多肽包含选自下列的氨基酸序列:在SEQ ID NO:2中所示氨基酸序列的氨基酸77-483、在SEQ ID NO:4中所示氨基酸序列的氨基酸74-481、在SEQ ID NO:5中所示氨基酸序列的氨基酸64-471、由在SEQ ID NO:1中所示核苷酸序列的核苷酸275-1495编码的氨基酸序列、由在SEQ ID NO:3中所示核苷酸序列的核苷酸342-1565所编码的氨基酸序列,以及编码包含AHASS活性的成熟AHASS多肽的氨基酸序列的片段和变体。

在本发明某些实施方案中,方法涉及到使用除草剂耐受的或抗除草剂的植物。“除草剂耐受的”或“抗除草剂”植物指的是植物对至少一种在这种水平上的除草剂耐受或有抗性,在所述的水平上通常能杀死正常或野生型品种的植物或抑制其生长。优选的,本发明除草剂耐受的植物包含除草剂耐受的或抗除草剂的AHASL蛋白质。术语“除草剂耐受的AHASL蛋

白质”或“抗除草剂的 AHASL”蛋白质指的是这种 AHASL 蛋白质，即当在已知会干扰 AHAS 活性并且处于已知会抑制野生型 AHASL 蛋白质的 AHAS 活性的浓度或水平上的除草剂存在时，与野生型 AHASL 蛋白质的 AHAS 活性相比，所述的 AHASL 蛋白质表现出更高的 AHAS 活性。

另外，认识到可以通过用编码除草剂耐受的或抗除草剂的 AHASL 蛋白质的核苷酸序列转化植物或其祖代，将除草剂耐受的或抗除草剂的 AHASL 蛋白质引入至植物中。此类除草剂耐受的或抗除草剂的 AHASL 蛋白质由除草剂耐受的或抗除草剂的 AHASL 多核苷酸编码。可选的，在植物中除草剂耐受的或抗除草剂的 AHASL 蛋白质可以作为天然存在的或在植物或其祖代基因组中内源 AHASL 基因中诱导突变的结果而存在。

本发明提供了对至少一种除草剂具有增加的抗性或耐受性的转化的植物、转化的植物组织、转化的植物细胞和转化的宿主细胞。除草剂的优选量或浓度是“有效量”或“有效浓度”。术语“有效量”或“有效浓度”指的是这种量或浓度，即它足够杀死下列或抑制下列生长：类似的、未转化的植物、植物组织、植物细胞或宿主细胞，但是该量不会杀死下列或如此严重的抑制下列的生长：转化的植物、转化的植物细胞或转化的宿主细胞。术语“类似的、未转化的植物、植物细胞或宿主细胞”分别指的是没有特定的本发明多核苷酸的植物、植物组织、植物细胞或宿主细胞，所述的本发明多核苷酸用来制造本发明的转化的植物、转化的植物细胞或转化的宿主细胞。因此使用术语“未转化的”不用于暗指植物、植物组织、植物细胞或其它宿主细胞在其基因组中没有重组 DNA。

本发明提供了用来增强植物、植物组织、植物细胞或其它宿主细胞对至少一种干扰 AHAS 酶活性的除草剂的耐受性或抗性的方法。优选的，此类除草剂是咪唑啉酮或磺酰脲除草剂。对于本发明，咪唑啉酮除草剂包括但不限于 PURSUIT® (咪草烟)、CADRE® (甲咪唑烟酸)、RAPTOR® (咪草啶酸)、SCEPTER® (灭草嗪)、ASSERT® (imazethabenz)、ARSENAL® (灭草烟)、任意前述除草剂的衍生物，或两种或多种前述除草剂的混合物，例如灭草烟/咪草啶酸 (ODYSSEY®)。更具体的，咪唑啉酮除草剂可以选

自但不限于 2-(4-异丙基-4-甲基-5-氧代-2-咪唑啉-2-基)-烟酸、[2-(4-异丙基)-4-甲基-5-氧代-2-咪唑啉-2-基]-3-喹啉羧酸、[5-乙基-2-(4-异丙基)-4-甲基-5-氧代-2-咪唑啉-2-基]-烟酸、2-(4-异丙基-4-甲基-5-氧代-2-咪唑啉-2-基)-5-(甲氧甲基)-烟酸、[2-(4-异丙基-4-甲基-5-氧代-2-咪唑啉-2-基)]-5-甲基烟酸，以及甲基 [6-(4-异丙基-4-甲基-5-氧代-2-咪唑啉-2-基)]-间-甲苯甲酸酯和甲基[2-(4-异丙基-4-甲基-5-氧代-2-咪唑啉-2-基)]-对甲苯甲酸酯的混合物。使用 5-乙基-2-(4-异丙基-4-甲基-5-氧代-2-咪唑啉-2-基)-烟酸和 [2-(4-异丙基-4-甲基-5-氧代-2-咪唑啉-2-基)]-5-(甲氧甲基)-烟酸是优选的。使用[2-(4-异丙基-4-甲基-5-氧代-2-咪唑啉-2-基)]-5-(甲氧甲基)-烟酸是尤其优选的。对于本发明，磺酰脲除草剂包括但不限于绿黄隆、甲黄隆、噻黄隆、氟噻黄隆、噻黄隆、苯黄隆、禾草芊混剂、烟噻黄隆、胺苯黄隆、玉噻黄隆、氟胺黄隆、醚苯黄隆、氟噻黄隆、醚黄隆、磺氨黄隆(amidosulfuron)、fluzasulfuron、啉咪黄隆、吡噻黄隆和吡氟黄隆。

本发明提供了在植物中增强 AHAS 活性的方法，包括用 AHASS 多核苷酸构建体转化植物。如本文用到的，术语“AHASS 多核苷酸”指的是包含 AHASS 核苷酸序列的多核苷酸。方法包括向至少一个植物细胞中引入本发明多核苷酸构建体并且从其生成转化的植物。在一个实施方案中，AHASS 多核苷酸构建体包含有效连接至 AHASS 核苷酸序列的启动子，其中该启动子在植物细胞中能够推动基因表达。优选的，该启动子是组成型启动子或组织优选的启动子。方法可以用来增强或增加植物对至少一种干扰 AHAS 酶催化活性的除草剂的耐受性。

本发明还提供了在除草剂耐受的植物中增强除草剂耐受性的方法，包括用 AHASS 多核苷酸构建体转化植物。这些方法包括向至少一个植物细胞中引入本发明 AHASS 多核苷酸构建体并且从其再生转化的植物。在一个实施方案中，除草剂耐受的植物包含除草剂耐受的 AHASL 蛋白质，它使得植物对至少一种已知会干扰 AHAS 酶活性的除草剂具有耐受性。在另一个实施方案中，AHASS 多核苷酸构建体包含有效连接到 AHASS 核苷酸序列的启动子，其中该启动子在植物细胞中能够推动基因表达。方法还可

以用来增加除草剂耐受的植物对至少一种干扰 AHAS 酶活性的除草剂的耐受性。因此，方法允许对除草剂耐受的植物应用更高水平的除草剂，而不会杀死或显著伤害除草剂耐受的植物。

本发明提供了用于在植物、植物组织、植物细胞和其它宿主细胞中表达本发明 AHASS 多核苷酸的表达盒。表达盒包含在目的植物、植物组织、植物细胞和其它宿主细胞中可表达的有效连接本发明多核苷酸的启动子，所述的多核苷酸编码全长 AHASS 多肽(即包括叶绿体转运肽)或成熟 AHASS 多肽(即没有叶绿体转运肽)。如果希望在植物或植物细胞的质体中表达，表达盒可以包含有效连接的编码叶绿体转运肽的叶绿体靶向序列。

本发明的表达盒可以用在增强植物或宿主细胞的除草剂耐受性的方法中。方法包括用本发明的表达盒转化植物或宿主细胞，其中表达盒包含可在目的植物或宿主细胞中表达的启动子并且其中启动子有效连接本发明的 AHASS 多核苷酸。

本发明还提供了用于在目的植物或宿主细胞中表达真核 AHASL 多肽和本发明 AHASS 多肽的表达载体。在一个实施方案中，植物表达载体包含第一多核苷酸构建体和第二多核苷酸构建体，其中第一多核苷酸构建体包含有效连接编码真核 AHASL 蛋白质的核苷酸序列的第一启动子，其中第二多核苷酸构建体包含有效连接编码 AHASS 蛋白质的核苷酸序列的第二启动子，并且其中第一和第二启动子在目的植物或宿主细胞中推动基因的表达。在一个实施方案中，第一和第二多核苷酸构建体还包含有效连接的叶绿体靶向序列。在另一个实施方案中，真核 AHASL 蛋白质是植物 AHASL 蛋白质，并且在一些情形中是除草剂耐受的 AHASL 蛋白质。对于在植物和植物细胞中的表达，表达载体在本文中指的是植物表达载体。植物表达载体的第一和第二启动子能够在植物细胞中推动基因表达。

本文中使用的术语“多核苷酸构建体”不用于将本发明限制在包含 DNA 的多核苷酸构建体上。本领域的普通技术人员将认识到，由核糖核苷酸以及核糖核苷酸和脱氧核糖核苷酸的组合组成的多核苷酸构建体，尤其是多核苷酸和寡核苷酸，也可以用在本文公开的方法中。因此，本发明的

多核苷酸构建体包括可以用在本发明方法中用于转化植物的所有多核苷酸构建体，包括但不限于由脱氧核糖核苷酸、核糖核苷酸和其组合组成的那些。此类脱氧核糖核苷酸和核糖核苷酸包括天然存在的分子和合成类似物。本发明的多核苷酸构建体还包括所有类型的多核苷酸构建体，它们包括但不限于单链形式、双链形式、发夹结构、茎和环结构等。另外，本领域的普通技术人员理解在本文公开的每种核酸序列还包括所例示的核苷酸序列的互补序列。

另外，认识到本发明的方法可以使用这样的多核苷酸构建体，即它在转化的植物中能够指导至少一种蛋白质，或至少一种 RNA(例如与 mRNA 至少一部分互补的反义 RNA)的表达。通常此类多核苷酸构建体由蛋白质的编码序列或有效连接至 5'和 3'转录调节区的 RNA 组成。可选的，还认识到本发明的方法可以使用不能在转化的植物中指导蛋白质或 RNA 表达的多核苷酸构建体。

本发明提供了包含有效连接 AHASS 结构域的真核 AHASL 结构域的融合蛋白质，其中 AHASL 结构域包含成熟真核 AHASL 蛋白质的氨基酸序列，并且其中 AHASS 结构域包含本发明 AHASS 蛋白质的氨基酸序列。AHASL 结构域可以包含成熟真核 AHASL 蛋白质的氨基酸序列，它与 AHASS 结构域的 AHASS 蛋白质的氨基酸序列来自相同或不同的真核物种。因此，AHASL 结构域包含来自任何真核生物的成熟真核 AHASL 蛋白质的任何已知氨基酸序列，所述的真核生物包括但不限于单子叶植物、双子叶植物、藻类、动物或真菌。本发明还提供了编码此类融合蛋白质的核苷酸序列。

本发明提供了用于在植物或目的宿主细胞中表达 AHASL-AHASS 融合多肽的表达载体。表达载体包含这样的启动子，即它能够在目的植物或宿主细胞中推动基因表达，有效连接到编码 AHASL-AHASS 融合多肽的多核苷酸上。多核苷酸包含第一核苷酸序列，它编码包含真核成熟 AHASL 多肽的氨基酸序列，并且有效连接到编码包含本发明成熟 AHASS 多肽的氨基酸序列的第二核苷酸序列上。在特定的实施方案中，多核苷酸另外包

含有效连接的第三核苷酸序列，它编码位于第一和第二核苷酸序列之间的连接区。

当在植物或宿主细胞中表达时，本发明的 AHASL-AHASS 融合多肽包含 AHAS 活性。优选的，AHASL-AHASS 融合多肽包含这种水平的 AHAS 活性，即所述水平的活性高于当相应的 AHASS 多肽不存在时对应的 AHASL 多肽的活性。

本发明提供了用于产生除草剂耐受植物的方法，包括用包含有效连接到编码 AHASL-AHASS 融合多肽的多核苷酸的启动子的植物表达载体转化植物细胞，并且从转基因植物细胞中生成转基因植物。方法可以用来产生对至少一种干扰 AHAS 酶的除草剂具有增加的耐受性的作物植物。

本发明包括用本文描述的多核苷酸转化的宿主细胞，所述的多核苷酸包括但不限于 AHASS 核苷酸序列、编码 AHASL-AHASS 融合多肽的核苷酸序列、多核苷酸构建体、表达盒和表达载体。本发明的宿主细胞包括原核和真核细胞，包括但不限于植物细胞、动物细胞、细菌细胞、酵母细胞和其它真菌细胞。优选的，本发明的宿主细胞是非人宿主细胞。更优选的，宿主细胞是植物细胞、细菌细胞和酵母细胞。最优选的，宿主细胞是植物细胞。

另外，认识到对于在目的宿主细胞中本发明多核苷酸的表达，多核苷酸可以有效连接到能够在目的宿主细胞中推动基因表达的启动子上。用于在宿主细胞中表达多核苷酸的本发明方法不依赖于特定的启动子。方法包括使用在本领域已知并且能够在目的宿主细胞中推动基因表达的任何启动子。

本发明包括 AHASS 多核苷酸分子以及其片段和变体。作为这些核苷酸序列片段的多核苷酸也包括在本发明中。术语“片段”指的是编码本发明 AHASS 多肽的核苷酸序列的部分。本发明的 AHASS 核苷酸序列的片段可以编码 AHASS 多肽的生物学活性部分，或者它可以是用作使用下列公开方法的杂交探针或 PCR 引物的片段。AHASS 多肽的生物学活性部分可以通过下列来制备：分离一种本发明 AHASS 核苷酸序列的一部分，表

达 AHASS 多肽的所述编码部分(例如通过体外重组表达), 并且评估 AHASS 多肽的所述编码部分的活性。作为 AHASS 核苷酸序列片段的多核苷酸分子至少包含大约 15、20、50、75、100、200、300、350、400、450、500、550、600、650、700、750、800、850、900、950、1000、1050、1100、1150、1200、1250、1300、1350、1400、1500、1600、1700 或 1800 个核苷酸, 或者多达在本文公开的全长核苷酸序列中存在的核苷酸数量(例如, 对于 SEQ ID NO:1 和 3 分别为 1726 和 1861 个核苷酸), 这取决于目的用途。

要理解的是分离的片段包括本发明之前没有公开的任何邻接序列以及基本相同但是没有公开的序列。因此, 如果分离的片段在本发明之前公开, 则该片段就不用于包括在本发明中。当序列没有在本发明之前公开时, 分离的核酸片段至少是大约 12、15、20、25 或 30 个邻接核苷酸。核苷酸序列的其它区域可以包含多种大小的片段, 这取决于与之前公开的序列的潜在同源性。

编码本发明 AHASS 多肽生物学活性部分的 AHASS 核苷酸序列的片段将编码至少大约 15、25、30、50、75、100、125、150、175、200、250、300、350、400 或 450 个邻接氨基酸, 或多达在本发明全长 AHASS 多肽中存在的氨基酸总数(例如对于 SEQ ID NO:2、4 和 5 分别是 483、481 和 471 个氨基酸)。可用作 PCR 引物杂交探针的 AHASS 核苷酸序列的片段不需要编码 AHASS 多肽的生物学活性部分。

作为在本文中公开的核苷酸序列变体的多核苷酸分子也包括在本发明中。本发明的 AHASS 核苷酸序列的“变体”包括编码本文公开的 AHASS 多肽但是由于遗传密码的简并性而适当不同的那些序列。这些天然存在的等位基因变体可以使用熟知的分子生物学技术来鉴定, 如下面说明的聚合酶链反应(PCR)和杂交技术。变体核苷酸序列还包括合成的衍生核苷酸序列, 例如, 它们是通过使用位点定向诱变产生的但按照下面讨论仍编码本发明所公开的 AHASS 多肽。通常, 本发明的核苷酸序列变体将与本文公开的特定核苷酸序列具有至少约 75%、80%、85%、90%、91%、92%、

93%、94%、95%、96%、97%、98%或99%的序列同一性。变体 AHASS 核苷酸序列将分别编码 AHASS 多肽，它们具有与本文公开的 AHASS 多肽的氨基酸序列有至少约 75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的同一性的氨基酸序列。

另外，技术人员还了解，可以通过突变将改变引入到本发明的核苷酸序列中从而使得所编码 AHASS 多肽的氨基酸序列改变而不改变 AHASS 多肽的生物学活性。因此，通过将一个或多个核苷酸替代、添加或缺失引入到本文公开的相应核苷酸序列中，从而将一个或多个氨基酸替代、添加或缺失引入到所编码的多肽中，来产生编码具有分别不同于 SEQ ID NO:2、4 或 5 所示序列的 AHASS 多肽的分离的多核苷酸分子。可以通过标准技术如位点定向诱变和 PCR 介导的诱变来引入突变。此类变体核苷酸序列也包括在本发明中。

例如，优选的，可以在一个或多个预测的，优选非必需的氨基酸残基处产生保守性氨基酸替代。“非必需”氨基酸残基是可以从 AHASS 多肽的野生型序列改变(例如分别是 SEQ ID NO:2、4 或 5 的序列)但是不改变生物学活性的残基，而“必需”氨基酸残基是生物学活性所需要的。“保守性氨基酸残基替代”是其中氨基酸残基用具有类似侧链的氨基酸残基置换的这种。具有类似侧链的氨基酸残基家族已经在本领域中定义。这些家族包括具有碱性侧链的氨基酸(例如赖氨酸、精氨酸、组氨酸)、具有酸性侧链的氨基酸(例如天冬氨酸、谷氨酸)、具有不带电的极性侧链的氨基酸(例如甘氨酸、天冬酰胺、谷氨酰胺、丝氨酸、苏氨酸、酪氨酸、半胱氨酸)、具有非极性侧链的氨基酸(例如丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、苯丙氨酸、甲硫氨酸、色氨酸)、具有  $\beta$ -分支的侧链的氨基酸(例如苏氨酸、缬氨酸、异亮氨酸)，以及具有芳香族侧链的氨基酸(例如酪氨酸、苯丙氨酸、色氨酸、组氨酸)。对于保守的氨基酸残基或对于保守基序中存在的氨基酸残基不进行此类替代。

本发明的蛋白质可以以多种方式进行改变，包括氨基酸替代、缺失、截短和插入。用于此类操作的方法通常是本领域已知的。例如可以通过在



DNA 中产生突变来制备 AHASS 多肽的氨基酸序列变体。诱变和核苷酸序列变更的方法是本领域已知的。见例如, Kunkel, 1985, *Proc. Natl. Acad. Sci. USA* 82:488-492; Kunkel 等人, 1987, *Methods in Enzymol.* 154:367-382; 美国专利号 4,873,192; Walker 和 Gaastra 编著, 1983, *Techniques in Molecular Biology*, MacMillan Publishing Company, New York, 以及在其中引用的文献。不影响目的蛋白质生物学活性的适当氨基酸替代的指导可以于在此引用作为参考的 Dayhoff 等人, 1978, *Atlas of Protein Sequence and Structure*, Natl. Biomed. Res. Found., Washington, D.C. 的模型中找到。保守性替代, 如一个氨基酸用具有类似性质的另一个氨基酸来替换, 可以是优选的。

可选的, 可以通过在沿着所有或部分的 AHASS 编码序列中随机引入突变(如通过饱和诱变)来产生变体 AHASS 核苷酸序列, 并且可以筛选所得突变体的 AHASS 生物学活性来鉴定保留了活性的突变体。在诱变后, 可以重组表达编码的多肽, 并且多肽的活性可以使用标准测定技术来确定。

因此, 本发明的核苷酸序列包括在本文公开的序列以及其片段和变体。本发明的 AHASS 核苷酸序列, 以及其片段和变体可以用作探针和/或引物以在其它植物中鉴别和/或克隆 AHASS 同源物。此类探针可以用来检测编码相同或一致多肽的转录物或基因组序列。

以这种方式, 方法如 PCR、杂交等可以用来鉴定与本发明序列具有相当大同一性的此类序列。见例如 Sambrook 等人, 1989, *Molecular Cloning: Laboratory Manual*, 第二版, Cold Spring Harbor Laboratory Press, Plainview, NY, 和 Innis 等人, 1990, *PCR Protocols: A Guide to Methods and Applications*, Academic Press, NY。本发明包括这样的 AHASS 核苷酸序列, 所述 AHASS 核苷酸序列是基于它们的序列与本文所示 AHASS 核苷酸序列或者其片段或变体的同一性来分离的。

在杂交方法中, 所有或部分的已知 AHASS 核苷酸序列可以用来筛选 cDNA 或基因组文库。构建此类 cDNA 和基因组文库的方法通常是本领域中已知的并且在 Sambrook 等人, 1989, *Molecular Cloning: A Laboratory*

*Manual*, 第二版, Cold Spring Harbor Laboratory Press, Plainview, NY 中公开。所称的杂交探针可以是基因组 DNA 片段、cDNA 片段、RNA 片段或其它寡核苷酸, 并且可以用可检测基团来标记, 如  $^{32}\text{P}$  或任何其它可检测标记物如其它放射性同位素、荧光化合物、酶或酶的辅因子。杂交探针可以通过标记基于本文公开的已知 AHASS 核苷酸序列的合成寡核苷酸来制备。可以另外使用基于在已知 AHASS 核苷酸序列或所编码氨基酸序列中的保守性核苷酸或氨基酸残基来设计的简并引物。探针通常包含在严格条件下杂交到本发明 AHASS 核苷酸序列或其片段或变体的至少约 12 个、优选约 25 个, 更优选约 50、75、100、125、150、175、200、250、300、350、400、500、600、700、800、900、1000、1200、1400、1600 或 1800 个连续核苷酸的核苷酸序列的区域。杂交探针制备方法通常是本领域已知的并且在在本文引用作为参考的 Sambrook 等人, 1989, *Molecular Cloning: A Laboratory Manual*, 第二版, Cold Spring Harbor Laboratory Press, Plainview, NY 中公开。

例如, 在本文公开的全长 AHASS 序列或其一部分或多部分, 可以用作能够特异性杂交到对应的 AHASS 序列和信使 RNA 上的探针。杂交技术包括平板 DNA 文库(噬斑或菌落, 见例如 Sambrook 等人, 1989, *Molecular Cloning: A Laboratory Manual*, 第二版, Cold Spring Harbor Laboratory Press, Plainview, NY)的杂交筛选。

此类序列的杂交可以在严格条件下进行。术语“严格条件”或“严格杂交条件”指的是这种条件, 即在该条件下, 相比杂交到其它序列上, 探针将以可检测的更高程度杂交到它的靶标序列上(例如至少 2 倍于背景)。严格条件是序列依赖的并且在不同的环境中不同。

通常, 严格条件是那些条件, 即其中在 pH 为 7.0 到 8.3 下盐浓度少于大约 1.5M Na 离子, 通常为大约 0.01 到 1.0M Na 离子浓度(或其它盐), 并且对于短探针(例如 10 到 50 个核苷酸)温度至少是大约 30°C, 并且对于长探针(例如大于 50 个核苷酸)至少约 60°C。严格条件可以通过添加去稳定剂如甲酰胺来实现。例示性的低严格条件包括在 37°C 用 30 到 35% 甲酰胺、

1M NaCl、1% SDS(十二烷基硫酸钠)的缓冲液杂交,并且在 50 到 55°C 在  $1 \times$  到  $2 \times$  SSC( $20 \times$  SSC=3.0 M NaCl/0.3 M 柠檬酸三钠)中洗涤。例示性的中等严格条件包括在 37°C 在 40 到 45% 甲酰胺、1.0 M NaCl、1% SDS 中杂交,并且在 55 到 60°C 在  $0.5 \times$  到  $1 \times$  SSC 中洗涤。例示性的高严格条件包括在 37°C 在 50% 甲酰胺、1 M NaCl、1% SDS 中杂交,并且在 60 到 65°C 在  $0.1 \times$  SSC 中洗涤。任选的,洗涤缓冲液可以包含大约 0.1% 到大约 1% 的 SDS。杂交的持续时间通常少于约 24 小时,通常是大约 4 到 12 小时。

特异性通常与杂交后洗涤相关,关键因素是最终洗涤溶液的离子强度和温度。对于 DNA-DNA 杂交体, $T_m$  可以从 Meinkoth 和 Wahl (1984) *Anal. Biochem.* 138:267-284 的公式估计:  $T_m = 81.5^\circ\text{C} + 16.6 (\log M) + 0.41 (\%GC) - 0.61 (\% \text{ 甲酰胺}) - 500/L$ ; 其中  $M$  是一价阳离子的体积摩尔浓度, %GC 是 DNA 中鸟苷和胞嘧啶核苷酸的百分比, %甲酰胺是杂交溶液中甲酰胺的百分比, 并且  $L$  是杂交体碱基对的长度。 $T_m$  是当 50% 的互补靶标序列杂交到完全匹配的的探针上时的温度(在确定的离子强度和 pH 下)。对于每 1% 的错配,  $T_m$  降低了大约  $1^\circ\text{C}$ ; 因此, 可以调整  $T_m$ 、杂交和/或洗涤条件以杂交到所需同一性的序列上。例如, 如果搜索  $\geq 90\%$  同一性的序列,  $T_m$  可以下降  $10^\circ\text{C}$ 。通常选择杂交条件是要大约比在确定的离子强度和 pH 下特异性序列和其互补序列的热解链温度低约  $5^\circ\text{C}$ 。但是, 特别严格杂交条件可以在比热解链温度( $T_m$ )低 1、2、3 或  $4^\circ\text{C}$  下进行杂交和/或洗涤; 中等严格条件可以在比热解链温度( $T_m$ )低 6、7、8、9 或  $10^\circ\text{C}$  下进行杂交和/或洗涤; 低严格条件可以在比热解链温度( $T_m$ )低 11、12、13、14、15 或  $20^\circ\text{C}$  下进行杂交和/或洗涤。使用此公式、杂交和洗涤成分以及所需  $T_m$ , 本领域的技术人员将理解在杂交严格性和/或洗涤溶液中的变化是原来描述的。如果所需程度的错配使得  $T_m$  低于  $45^\circ\text{C}$  (水溶液)或  $32^\circ\text{C}$  (甲酰胺溶液), 优选的增加 SSC 浓度使得可以使用更高的温度。对核酸杂交的深入指导可以在下列中找到: Tijssen, 1993, *Laboratory Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Acid Probes*, Part I, Chapter 2, Elsevier, NY; 以及 Ausubel 等人编著, 1995,

*Current Protocols in Molecular Biology*, Chapter 2, Greene Publishing and Wiley-Interscience, NY. 也见 Sambrook 等人, 1989, *Molecular Cloning: A Laboratory Manual*, 第二版, Cold Spring Harbor Laboratory Press, Plainview, NY.

认识到本发明的多核苷酸分子和多肽包括含有这种核苷酸或氨基酸序列的多核苷酸分子和多肽, 所述的核苷酸或氨基酸序列与 SEQ ID NO:1 或 SEQ ID NO:3 的核苷酸序列, 或与 SEQ ID NO:2、4 或 5 的氨基酸序列充分相同。术语“充分相同”在本文中用来指的是包含足够或最小量的与第二氨基酸或核苷酸序列一样或相当(例如具有类似的侧链)的氨基酸残基或核苷酸的第一氨基酸或核苷酸序列, 从而第一和第二氨基酸或核苷酸序列具有共同的结构域和/或共同的功能活性。例如, 在本文中 will 含有共同结构域的氨基酸或核苷酸序列(具有至少约 45%、55% 或 65% 同一性, 优选为 75% 同一性, 更优选为 77%、80%、81%、85%、95% 或 98% 同一性) 定义为充分相同。

为了确定两个氨基酸序列或两个核酸的百分比同一性, 比对序列用于最佳比较目的。两个序列之间的百分比同一性是序列共有的一致位置数的函数(即, 百分比同一性 = 一致位置数/位置总数(例如重叠位置) × 100)。在一个实施方案中, 两个序列是一样长的。可以使用与下列描述的那些技术类似的技术(有或没有允许空位)确定两个序列之间的百分比同一性。在计算百分比同一性中, 通常计算准确匹配。对于本发明, 序列同一性/相似性值优选来自本发明的全长核苷酸或全长氨基酸序列对第二核苷酸或氨基酸序列的无空位比对。

使用数学算法可以实现两个序列之间的百分比同一性的确定。用于比较两个序列的数学算法优选的非限制性实例是 Karlin 和 Altschul, 1990, *Proc. Natl. Acad. Sci. USA* 87:2264 的算法, 在 Karlin 和 Altschul, 1993, *Proc. Natl. Acad. Sci. USA* 90:5873-5877 中得到改进。将此类算法引入到 Altschul 等人, 1990, *J. Mol. Biol.* 215:403 的 NBLAST 和 XBLAST 程序。可以用 NBLAST 程序、分值=100、字长=12 进行 BLAST 核苷酸搜索以得到与本

发明多核苷酸分子同源的核苷酸序列。可以使用 XBLAST 程序、分值 = 50、字长 = 3 进行 BLAST 蛋白质搜索以得到与本发明蛋白质分子同源的氨基酸序列。为了得到用于比较目的的空位比对，可以按照在 Altschul 等人, 1997, *Nucleic Acids Res.* 25:3389 中描述的使用空位 BLAST。备选的，可以使用 PSIBlast 以进行检测分子间远缘关系的重复搜索。见 Altschul 等人, 1997, 见上。当使用 BLAST、空位 BLAST 和 PSI-Blast 程序时，可以使用各自程序（例如 XBLAST 和 NBLAST）的默认参数。见 <http://www.ncbi.nlm.nih.gov>。另外优选的、非限制的用于序列比较的数学算法实例是 Myers 和 Miller, 1988, *CABIOS* 4:11-17 的算法。将此类算法引入到 ALIGN 程序(版本 2.0)中，该程序是 GCG 序列比对软件包的一部分。当使用 ALIGN 程序用于比较氨基酸序列时，可以使用 PAM120 权重残基表、空位长度罚分为 12，以及空位罚分为 4。也可以通过目测手动进行比对。

除非另外在本文中说明，配对百分比序列同一性从比对两个核苷酸或两个氨基酸序列生成，所述的比对使用 ClustalX 版本 1.81 和使用简单 p 距离模型的 MEGA (分子进化遗传分析)版本 2.1。术语“相当的程序”指的是这种任意的序列比较程序，即对于任意两个所讨论的序列，当与 ClustalX 版本 1.81 所生成的对应比对和通过使用简单 p 距离模型的 MEGA(分子进化遗传分子)版本 2.1 所计算的百分比同一性对比，所述的程序生成具有一致的核苷酸或氨基酸残基匹配和一致的百分比序列同一性比对。

本发明的 AHASS 核苷酸序列包括天然存在的序列以及突变形式。同样的，本发明的多肽包括天然存在的多肽以及其变体和修饰形式。此类变体将继续具有所需要的 AHASS 活性。明显的，在编码变体的 DNA 中产生的突变不会将序列放在阅读框外并且优选的将不会制造产生二级 mRNA 结构的互补区(见例如，EP 专利申请公开号 75,444)。

不希望本文所包括的蛋白质序列缺失、插入和替代引起蛋白质特征的根本变化。但是，当在进行此类任务之前很难预测替代、缺失或插入的精

确作用时,本领域的技术人员将了解此作用可以通过常规的筛选测定来评估。那就是,可以通过 AHAS 活性测定来评估此活性。见,例如 Singh 等人,1988, *Anal. Biochem.* 171:173-179, 在本文引用作为参考。

变体核苷酸序列和蛋白质还包括来自诱变方法和重组方法(如 DNA 混编)的序列和蛋白质。用此类方法,可以操作一种或多种不同的 AHASS 编码序列来产生具有所需性质的新 AHASS 蛋白质。以此方式,重组多核苷酸文库可以从相关序列多核苷酸总体生成,其包含具有基本的序列同一性并且可以被体外或体内同源重组的序列区。例如,使用此方法,编码目的结构域的序列基序可以在本发明 AHASS 基因和其它已知 AHASS 基因之间被改组以得到编码具有改进目的性质(如在酶的情况下,增加的  $K_m$ )的蛋白质的新基因。本领域已知此类 DNA 改组的方法。见,例如 Stemmer, 1994, *Proc. Natl. Acad. Sci. USA* 91:10747-10751; Stemmer, 1994, *Nature* 370:389-391; Cramer 等人, 1997, *Nature Biotech.* 15:436-438; Moore 等人, 1997, *J. Mol. Biol.* 272:336-347; Zhang 等人, 1997, *Proc. Natl. Acad. Sci. USA* 94:4504-4509; Cramer 等人, 1998, *Nature* 391:288-291; 和美国专利号 5,605,793 和 5,837,458。

本发明的核苷酸序列可以用来从其它生物,特别是从其它植物,更特别是其它单子叶植物中分离相应的序列。以此方式,方法如 PCR、杂交等可以用来鉴定此类序列,这是基于它们的序列与本文所示序列的同一性。本发明包括基于其序列与本文所示的全长 AHASS 序列或其片段的同一性来分离的序列。因此,本发明包括编码 AHASS 蛋白质并且在严格条件下杂交到本文所公开的序列或其片段上的分离的序列。

在 PCR 方法中,可以设计寡核苷酸引物用在 PCR 反应中,所述的 PCR 反应从任何目的植物所提取的 cDNA 或基因组 DNA 中扩增相应的 DNA 序列。设计 PCR 引物和 PCR 克隆的方法是本领域熟知的并且在: Sambrook 等人, 1989, *Molecular Cloning: A Laboratory Manual*, 第二版, Cold Spring Harbor Laboratory Press, Plainview, NY 中公开。也见 Innis 等人编著, 1990, *PCR Protocols: A Guide to Methods and Applications*, Academic Press, NY;

Innis 和 Gelfand 编著, 1995, *PCR Strategies*, Academic Press, NY; 以及 Innis 和 Gelfand 编著, 1999, *PCR Methods Manual*, Academic Press, NY. PCR 已知方法包括但不限于, 使用下列引物的方法: 配对引物、嵌套引物、单特异性引物、简并引物、基因特异性引物、载体特异性引物、部分错配引物等。

在用于在目的植物中表达的表达盒中也提供了本发明的 AHASS 序列。该盒包括有效连接本发明 AHASS 核苷酸序列的 5' 调节序列和 3' 调节序列。在此用到的术语“有效连接”指的是在启动子和第二序列之间的功能性连接, 其中启动子序列起始和介导了对应于第二序列的 DNA 序列的转录。通常, 有效连接意思是核酸或氨基酸序列是连接的从而使得序列都能满足所使用序列所具有的功能和活性。在一个实施方案中, 有效连接意思是连接的核酸序列是邻接的, 并且对于以邻接和相同的阅读框来连接两个蛋白质编码区是必需的。盒可以另外包括待共转化到生物中的至少一种额外基因。备选的, 可以在多个表达盒上提供额外的基因。

提供了具有多个限制性位点的此类表达盒, 所述限制性位点用于受到调节区转录调控的 AHASS 序列的插入。表达盒可以另外包含选择性标记基因。

在转录的 5' - 3' 方向, 表达盒可以包括在植物中有功能的转录和翻译起始区(即启动子)、本发明的 AHASS 序列、以及转录和翻译终止区(即终止区)。启动子可以是天然的或类似的, 或对于植物宿主和/或对于本发明的 AHASS 序列是外源的或异源的。另外, 启动子可以是天然序列或备选是合成序列。在启动子对于植物宿主是“外源的”或“异源的”时, 它指的是在该启动子要引入的天然植物中为不存在的启动子。在启动子对于本发明的 AHASS 序列是“外源的”或“异源的”时, 它指的是启动子对于有效连接的本发明 AHASS 序列不是本身的或天然存在的。如本文用到的, 嵌合基因包含有效连接到与编码序列异源的转录起始区的编码序列。

尽管使用异源启动子表达序列可以是优选的, 也可以使用本身的 AHASS 或 AHASL 启动子序列。此类构建体将改变 AHASS 蛋白质在植物

或植物细胞中的表达水平。因此，植物或植物细胞的表型也改变。

终止区可以是与转录起始区来源相同的，可以是与有效连接的目的 *AHASS* 序列来源相同的，可以与植物宿主来源相同的，或可以来自其它来源(即对于启动子、目的 *AHASS* 序列、植物宿主或其任何组合是外源的或异源的)。常规终止区可以从根瘤农杆菌(*A. tumefaciens*)的 Ti 质粒得到，如章鱼碱合酶和胭脂氨酸合成酶终止区(也见 Guerineau 等人, 1991, *Mol. Gen. Genet.* 262:141-144; Proudfoot, 1991, *Cell* 64:671-674; Sanfacon 等人, 1991, *Genes Dev.* 5:141-149; Mogen 等人 1990, *Plant Cell* 2:1261-1272; Munroe 等人, 1990, *Gene* 91:151-158; Ballas 等人, 1989, *Nucleic Acids Res.* 17:7891-7903; 和 Joshi 等人, 1987, *Nucleic Acid Res.* 15:9627-9639)。

根据需要，可以优化基因用于在转化的植物中增加的表达。那就是，可以使用植物优选的密码子来合成基因用于改进的表达(见，例如用于宿主优选密码子选择讨论的 Campbell 和 Gowri, 1990, *Plant Physiol.* 92:1-11)。方法可自合成植物优选基因的领域中获得(见，例如美国专利号 5,380,831 和 5,436,391，以及 Murray 等人, 1989, *Nucleic Acids Res.* 17:477-498，在本文引用作为参考)。

已知额外的序列修饰来增强细胞宿主中的基因表达。这些包括消除编码假加多聚 A 信号、外显子-内含子剪接位点信号、转座子样重复的序列和对于基因表达有害的其它此类明确表征的信号。可以将序列的 G-C 含量调整到所给细胞宿主的平均水平，所述的平均水平是参考在宿主细胞中表达的已知基因来计算的。可能的话，将序列修饰以避免预测的发夹二级 mRNA 结构。

也可以将增强基因表达的核苷酸序列用在植物表达载体中。这些包括玉米 *AdhI* 的内含子，内含子 1 基因(Callis 等人, 1987, *Genes and Development* 1:1183-1200)，以及来自烟草花叶病毒(TMV)、玉米褪绿矮缩病毒和苜蓿花叶病毒的前导序列(W-序列)(Gallie 等人, 1987, *Nucleic Acid Res.* 15:8693-8711 和 Skuzeski 等人, 1990, *Plant Molec. Biol.* 15:65-79)。已经表明来自玉米的 *shrunkent-1* 基因座的第一内含子增加在嵌合基因构建



体中基因的表达。美国专利号 5,424,412 和 5,593,874 公开了在基因表达构建体中特定内含子的用途, 并且 Gallie 等人, 1994, *Plant Physiol.* 106:929-939 也表明内含子可用于在组织特异性基础上调节基因表达。为了进一步增强或优化 AHASS 小亚基基因表达, 本发明的植物表达载体还可以包含含有基质附着区(MAR)的 DNA 序列。之后, 用此类修饰的表达系统转化的植物细胞可以表现出本发明核苷酸序列的过量表达或组成型表达。

表达盒可以额外在表达盒构建体中包含 5'前导序列。此类前导序列可以作用为增强翻译。翻译前导序列在本领域已知并且包括: 小 RNA 病毒前导序列, 例如 EMCV 前导序列(脑心肌炎 5'非编码区) (Elroy-Stein 等人, 1989, *Proc. Natl. Acad. Sci. USA* 86:6126-6130); 马铃薯 Y 病毒前导序列, 例如, TEV 前导序列 (烟草蚀斑病毒) (Gallie 等人, 1995, *Gene* 165(2):233-238)、MDMV 前导序列(玉米矮花叶病毒) (*Virology* 154:9-20), 和人免疫球蛋白重链结合蛋白 (BiP) (Macejak 等人, 1991, *Nature* 353:90-94); 来自苜蓿花叶病毒的外被蛋白 mRNA(AMV RNA 4)的非翻译前导序列 (Jobling 等人, 1987, *Nature* 325:622-625); 烟草花叶病毒(TMV)前导序列(Gallie 等人, 1989, 于 *Molecular Biology of RNA* 一书中, Cech, Liss 编著, New York, pp. 237-256); 玉米褪绿斑驳病毒前导序列(MCMV) (Lommel 等人, 1991, *Virology* 81:382-385)。也见 Della-Cioppa 等人, 1987, *Plant Physiol.* 84:965-968。增强翻译的其它已知方法也可以使用, 例如内含子等。

在制备表达盒中, 可以操作多种 DNA 片段, 从而提供在合适方向中并且根据需要在恰当的阅读框中的 DNA 序列。为此目的, 可以使用衔接头或连接体来连接 DNA 片段或可以包括其它操作来提供方便的限制性位点、移除多余的 DNA、移除限制性位点等。为此目的, 可以涉及体外诱变、引物修复、限制、复性和替代, 例如转换和颠换。

可以在本发明的实践中使用一些启动子。启动子基于所需要的结果选择。核酸可以与组成型、组织优选的或其它启动子组合用于在植物中表达。

例如, 此类组成型启动子包括 Rsyn7 启动子的核心启动子和在 WO 99/43838 和美国专利号 6,072,050 中公开的其它组成型启动子; 核心 CaMV 35S 启动子(Odell 等人, 1985, *Nature* 313:810-812); 稻肌动蛋白(McElroy 等人, 1990, *Plant Cell* 2:163-171); 遍在蛋白(Christensen 等人, 1989, *Plant Mol. Biol.* 12:619-632 和 Christensen 等人, 1992, *Plant Mol. Biol.* 18:675-689); pEMU (Last 等人, 1991, *Theor. Appl. Genet.* 81:581-588); MAS (Velten 等人, 1984, *EMBO J.* 3:2723-2730); ALS 启动子 (美国专利号 5,659,026) 等。其它组成型启动子包括例如美国专利号 5,608,149、5,608,144、5,604,121、5,569,597、5,466,785、5,399,680、5,268,463、5,608,142 和 6,177,611。

可以使用组织优选启动子在特定的植物组织中得到增强的 AHASS 表达。此类组织优选的启动子包括但不限于, 叶优选的启动子、根优选的启动子、种子优选的启动子和茎优选的启动子。组织优选的启动子包括 Yamamoto 等人, 1997, *Plant J.* 12(2):255-265; Kawamata 等人, 1997, *Plant Cell Physiol.* 38(7):792-803; Hansen 等人, 1997, *Mol. Gen. Genet.* 254(3):337-343; Russell 等人, 1997, *Transgenic Res.* 6(2):157-168; Rinehart 等人, 1996, *Plant Physiol.* 112(3):1331-1341; Van Camp 等人, 1996, *Plant Physiol.* 112(2):525-535; Canevascini 等人, 1996, *Plant Physiol.* 112(2):513-524; Yamamoto 等人, 1994, *Plant Cell Physiol.* 35(5):773-778; Lam, 1994, *Results Probl. Cell Differ.* 20:181-196; Orozco 等人, 1993, *Plant Mol Biol.* 23(6):1129-1138; Matsuoka 等人, 1993, *Proc Natl. Acad. Sci. USA* 90(20):9586-9590; 和 Guevara-Garcia 等人, 1993, *Plant J.* 4(3):495-505 所公开的启动子。如果需要可以修饰此类启动子用于弱表达。

在一个实施方案中, 将目的核酸靶向到叶绿体用于表达。以这种方式, 在目的核酸不是直接插入至叶绿体中时, 表达盒会额外包含叶绿体靶向序列, 该靶向序列包含编码指引目的基因产物到叶绿体中的叶绿体转运肽的核苷酸序列。此类转运肽是本领域已知的。见, 例如, Von Heijne 等人, 1991, *Plant Mol. Biol. Rep.* 9:104-126; Clark 等人, 1989, *J. Biol. Chem.*

264:17544-17550; Della-Cioppa 等人, 1987, *Plant Physiol.* 84:965-968; Romer 等人, 1993, *Biochem. Biophys. Res. Commun.* 196:1414-1421; 和 Shah 等人, 1986, *Science* 233:478-481。尽管本发明的 AHASS 多肽包括天然叶绿体转运肽, 可以通过将叶绿体靶向序列有效连接到编码本发明成熟 AHASS 多肽的核苷酸序列的 5'端, 将本领域已知的任何叶绿体转运肽融合到本发明成熟 AHASS 多肽的氨基酸序列上。

叶绿体靶向序列是本领域已知的并且包括核酮糖-1,5-二磷酸羧化酶 (Rubisco) 的叶绿体小亚基 (de Castro Silva Filho 等人, 1996, *Plant Mol. Biol.* 30:769-780; Schnell 等人, 1991, *J. Biol. Chem.* 266(5):3335-3342); 5-(烯醇式丙酮酰)莽草酸-3-磷酸合酶 (EPSPS) (Archer 等人, 1990, *J. Bioenerg. Biomemb.* 22(6):789-810); 色氨酸合酶 (Zhao 等人, 1995, *J. Biol. Chem.* 270(11):6081-6087); 质体蓝素 (Lawrence 等人, 1997, *J. Biol. Chem.* 272(33):20357-20363); 分支酸合成酶 (Schmidt 等人, 1993, *J. Biol. Chem.* 268(36):27447-27457); 和捕获光能叶绿素 a/b 结合蛋白 (LHBP) (Lamppa 等人, 1988, *J. Biol. Chem.* 263:14996-14999)。也见 Von Heijne 等人, 1991, *Plant Mol. Biol. Rep.* 9:104-126; Clark 等人, 1989, *J. Biol. Chem.* 264:17544-17550; Della-Cioppa 等人, 1987, *Plant Physiol.* 84:965-968; Romer 等人, 1993, *Biochem. Biophys. Res. Commun.* 196:1414-1421; 和 Shah 等人, 1986, *Science* 233:478-481。

叶绿体转化的方法是本领域已知的 (见, 例如, Svab 等人, 1990, *Proc. Natl. Acad. Sci. USA* 87:8526-8530; Svab 和 Maliga, 1993, *Proc. Natl. Acad. Sci. USA* 90:913-917; Svab 和 Maliga, 1993, *EMBO J.* 12:601-606)。方法依靠含有选择性标记的 DNA 的基因枪传递和通过同源重组将 DNA 靶向到质体基因组。另外, 可以通过核编码和质体导向 RNA 聚合酶的组织优选表达, 通过沉默质体产生转基因的反式激活作用来实现质体转化。此类系统在 McBride 等人, 1994, *Proc. Natl. Acad. Sci. USA* 91:7301-7305 中进行了报道。

可以优化待靶向到叶绿体的目的核酸用于在叶绿体中表达以说明在植

物细胞核和此细胞器之间密码子选择的差异。以这种方式，可以使用叶绿体优选的密码子合成目的核酸(见，例如，美国专利号 5,380,831，在本文引用作为参考)。

如本文所公开的，本发明的 *AHASS* 核苷酸序列可以用来增强包含编码除草剂耐受 *AHASL* 多肽的基因的植物的除草剂耐受性。此类 *AHASL* 基因可以掺入到植物基因组中并且可以是内源基因或转基因。另外，在某些实施方案中，可以将本发明的核酸序列与任何目的核苷酸序列的组合堆积在一起以产生具有所需表型的植物。例如，将本发明的多核苷酸与编码具有杀虫和/或杀昆虫(insecticidal)活性的多肽的任何其它多核苷酸堆积，所述的多肽例如苏云金芽孢杆菌(*Bacillus thuringiensis*)毒蛋白(在美国专利号 5,366,892; 5,747,450; 5,737,514; 5,723,756; 5,593,881; 和 Geiser 等人, 1986, *Gene* 48:109 中描述)。生成的组合还可以包括任何一种目的多核苷酸的多个拷贝。

认识到用这些核苷酸序列，可以构建与 *AHASS* 序列的信使 RNA(mRNA)至少一部分互补的反义构建体。构建反义核苷酸以与相应的 mRNA 杂交。可以对反义序列进行修饰，只要反义序列杂交对应的 mRNA 并且干扰其表达即可。以这种方式，可以使用与对应的反义序列具有 70%、优选 80%、更优选 85% 的序列同一性的反义构建体。另外，可以使用反义核苷酸的部分来破坏靶基因的表达。通常，可以使用至少 50 个核苷酸、100 个核苷酸、200 个核苷酸或更长的序列。

还可以使用有义方向的本发明核苷酸序列以抑制植物中内源基因的表达。本领域已知使用有义方向的核苷酸序列抑制植物中基因表达的方法。方法通常包括用包含启动子的 DNA 构建体转化植物，所述的启动子在植物中推动表达，有效连接到对应内源基因转录物的核苷酸序列的至少一部分。通常，此类核苷酸序列与内源基因转录物的序列具有相当大的序列同一性，优选的，大于约 65% 序列同一性，更优选的大于约 85% 序列同一性，最优选的大于约 95% 序列同一性(见美国专利号 5,283,184 和 5,034,323; 在本文引用作为参考)。

通常，表达盒将包含用于转化细胞选择的选择性标记基因。选择性标记基因是用于转化的细胞或组织的选择。标记基因包括编码赋予了抗生素抗性的多肽的基因，如编码新霉素磷酸转移酶 II(NEO)和潮霉素磷酸转移酶(HPT)的那些基因，以及编码赋予针对除草剂化合物抗性的多肽的基因，如草铵膦、溴苯腈、咪唑啉酮和 2,4-二氯苯氧乙酸(2,4-D)。总体上见 Yarranton, 1992, *Curr. Opin. Biotech.* 3:506-511; Christopherson 等人, 1992, *Proc. Natl. Acad. Sci. USA* 89:6314-6318; Yao 等人, 1992, *Cell* 71:63-72; Reznikoff, 1992, *Mol. Microbiol.* 6:2419-2422; Barkley 等人, 1980, in *The Operon*, pp. 177-220; Hu 等人, 1987, *Cell* 48:555-566; Brown 等人, 1987, *Cell* 49:603-612; Figge 等人, 1988, *Cell* 52:713-722; Deuschle 等人, 1989, *Proc. Natl. Acad. Sci. USA* 86:5400-5404; Fuerst 等人, 1989, *Proc. Natl. Acad. Sci. USA* 86:2549-2553; Deuschle 等人, 1990, *Science* 248:480-483; Gossen, 1993, Ph.D. Thesis, University of Heidelberg; Reines 等人, 1993, *Proc. Natl. Acad. Sci. USA* 90:1917-1921; Labow 等人, 1990, *Mol. Cell. Biol.* 10:3343-3356; Zambretti 等人, 1992, *Proc. Natl. Acad. Sci. USA* 89:3952-3956; Baim 等人, 1991, *Proc. Natl. Acad. Sci. USA* 88:5072-5076; Wyborski 等人, 1991, *Nucleic Acids Res.* 19:4647-4653; Hillenand-Wissman, 1989, *Topics Mol. Struc. Biol.* 10:143-162; Degenkolb 等人, 1991, *Antimicrob. Agents Chemother.* 35:1591-1595; Kleinschmidt 等人, 1988, *Biochemistry* 27:1094-1104; Bonin, 1993, Ph.D. Thesis, University of Heidelberg; Gossen 等人, 1992, *Proc. Natl. Acad. Sci. USA* 89:5547-5551; Oliva 等人, 1992, *Antimicrob. Agents Chemother.* 36:913-919; Hlavka 等人, 1985, *Handbook of Experimental Pharmacology*, Vol. 78, Springer-Verlag, Berlin; Gill 等人, 1988, *Nature* 334:721-724。此类公开在本文引用作为参考。

选择性标记基因的上述列表不是为了限制。在本发明中可以使用任何选择性标记基因。

编码 AHASS 多肽的分离的多核苷酸分子可以在载体中使用来转化植物，所以产生的植物对除草剂尤其是咪唑啉酮除草剂具有增强的耐受性。可

以将编码 AHASS 多肽的分离的多核苷酸分子单独用在载体中或与编码 AHAS 酶大亚基的核苷酸序列组合用在载体中以及在植物中赋予除草剂抗性(见美国专利号 6,348,643, 其在本文中以其全文引用作为参考)。

本发明的 AHASS 核苷酸序列还可以与 AHASL 核苷酸序列组合使用, 作为标记用于选择转化的植物细胞、植物组织和植物。可以将任何目的基因掺入到包含编码 AHASS 和 AHASL 多肽的核苷酸序列的载体中。可以将载体引入到对 AHAS-抑制性除草剂敏感的植物细胞或组织中。可以在除草剂存在时, 使用本领域已知的标准技术选择含有这些载体的转化的植物、植物组织和植物细胞。

本发明还提供了包含这种启动子的植物表达载体, 所述的启动子在植物中推动表达, 有效连接本发明分离的 AHASS 多核苷酸分子。分离的多核苷酸分子包含编码单子叶植物 AHASS 多肽的核苷酸序列, 所述的 AHASS 多肽尤其是包含在 SEQ ID NO:2、SEQ ID NO:4 或 SEQ ID NO:5 中所示的氨基酸序列或其功能性片段或变体的 AHASS 多肽。本发明的植物表达载体不依赖于特定的启动子, 只是此类启动子要能够在植物细胞中推动基因表达。优选的启动子包括但不限于组成型启动子和组织优选的启动子。

在另一个实施方案中, 植物表达载体包含: 有效连接到编码 AHASL 多肽的核苷酸序列的真核 AHASL 基因启动子, 能够在植物细胞中推动表达、有效连接本发明 AHASS 核苷酸序列的启动子, 其中 AHASS 核苷酸序列选自: 在 SEQ ID NO:1 和 SEQ ID NO:3 中所示的核苷酸序列、编码在 SEQ ID NO:2、SEQ ID NO:4 和 SEQ ID NO:5 中所示的氨基酸序列的核苷酸序列, 以及编码包含 AHASS 活性的成熟 AHASS 多肽的其片段和变体。

与在 AHASS 多肽不存在时 AHASL 多肽的 AHAS 活性相比, 当此类 AHASS 和 AHASL 多肽彼此存在时, 此类成熟 AHASS 多肽能够增加至少一种 AHASL 多肽的 AHAS 活性。

在另一个实施方案中, 在植物中表达异源 AHAS 基因的植物表达载体包含有效连接编码融合多肽的核苷酸序列的植物启动子, 所述的融合多肽包含融合到 AHASS 多肽氨基酸序列的成熟 AHASL 多肽氨基酸序列。此类多

核苷酸构建体包含有效连接本发明 AHASS 核苷酸序列的、编码成熟 AHASL 多肽的核苷酸序列。

如本文用到的，在编码融合多肽的此类多核苷酸的上下文中术语“有效连接”指的是编码第一氨基酸序列的第一核苷酸序列以这种方式连接或融合到编码第二氨基酸序列的第二核苷酸序列，即以所述的方式由融合核苷酸序列编码的融合氨基酸序列包含第一和第二氨基酸序列。认识到编码本发明融合多肽的多核苷酸构建体还可以包含额外的核苷酸序列并且此类额外的核苷酸序列可以位于第一编码序列的 5'、第二编码序列的 3'，或在第一和第二编码序列之间。另外认识到在本发明的某些实施方案中，在编码融合多肽的此类融合核苷酸序列中需要包含编码连接体氨基酸序列的额外的核苷酸序列。在所得融合多肽中，连接体氨基酸序列将位于第一和第二氨基酸序列之间。认识到需要具有此类连接体氨基酸序列以允许在对应于第一和第二氨基酸序列的融合多肽部分之间有最佳相互作用。此类连接体氨基酸序列可以包含 1、2、3、4、5、6、7、8、9、10、15、20、25、30、40、50、75、100 或更多个氨基酸。

当谈及组合两种氨基酸序列以形成融合蛋白质时所使用的“有效连接”是指两种氨基酸序列融合或连接以形成单一连续的氨基酸序列，从而两种序列满足所使用序列所具有的功能或活性。在一个实施方案中，此类融合蛋白质是包含有效连接到第二核苷酸序列的第一核苷酸序列的单一连续核苷酸序列的翻译产物。第一核苷酸序列编码第一氨基酸序列，并且第二核苷酸序列编码第二氨基酸序列。之后融合多肽作为单一连续核苷酸序列的翻译产物产生。

在另一个实施方案中，植物表达载体包含能够在植物细胞中推动基因表达、有效连接编码融合多肽的多核苷酸的启动子，所述的融合多肽包含成熟 AHASL 多肽的氨基酸序列以及本发明的成熟 AHASS 多肽的氨基酸序列。因此，融合多肽由两个结构域 AHASL 结构域和 AHASS 结构域组成。此类融合多肽可以包含：从 N 末端，AHASL 结构域接着是 AHASS 结构域，或备选的，AHASS 结构域接着是 AHASL 结构域。另外，融合

多肽可以另外包含连接区的氨基酸序列。在此类融合多肽中，连接区位于 AHASL 和 AHASS 结构域之间。如果需要，对于叶绿体表达，编码融合多肽的多核苷酸另外包含编码叶绿体转运肽的叶绿体靶向序列。此类叶绿体转运肽可以选自：来自融合多肽的天然 AHASS 或 AHASL 多肽的叶绿体转运肽，或本领域已知的任何其它叶绿体转运肽。认识到此类叶绿体转运肽通常在蛋白质的 N 末端。

本发明的 AHASS 核苷酸序列可以用来产生粘连的 (tethered) AHAS 酶，其包含本发明的 AHASL-AHASS 融合多肽。例如，在本发明的实施方案中，通过编码连接区(或连接体多肽)的连接体核苷酸序列，如聚甘氨酸(polyGly)，编码本发明 AHASS 多肽的第一多核苷酸分子与编码真核 AHASL 蛋白质的氨基酸序列的第二多核苷酸分子翻译偶联。那就是，连接体核苷酸序列有效连接到第一核苷酸序列的 3'端和第二核苷酸序列的 5'端，从而编码连续包含下列的多肽：AHASS 多肽的氨基酸序列、连接区的氨基酸序列和 AHASL 多肽的氨基酸序列。备选的配置包括用小亚基转运序列调换连接区转录物周围的大亚基和小亚基的成熟编码序列。本发明不依赖于具有特定数量氨基酸的连接区，只是融合多肽要具有 AHAS 活性，优选是比在对应的 AHASS 多肽不存在时相应的 AHASL 多肽更高水平的 AHAS 活性。

认识到通过保持大 AHAS 亚基结构域和小 AHAS 亚基结构域彼此接近，粘连的 AHAS 酶可以用来增强除草剂耐受性。用大肠杆菌 AHAS 酶已经表明，大亚基和小亚基之间的联系是宽松的。在大肠杆菌中据估计在每个亚基浓度为  $10^{-7}$ M 时，作为  $\alpha_2\beta_2$  活性全酶，大亚基仅仅是半连接的 (Sella 等人, 1993, *J. Bacteriology* 175:5339-5343)。

认识到当相对于 AHASL 蛋白质的摩尔浓度，存在摩尔过量的 AHASS 蛋白质时，就达到最高活性。由于已经确定了当两个亚基相连时 AHAS 酶是最稳定和活性的 (Weinstock 等人, 1992, *J. Bacteriology*, 174:5560-5566, Sella 等人, 1993, *J. Bacteriology* 175:5339-5343)，可以通过将两个亚基融合在单一多肽中来产生高活性和稳定性的酶。已经表明粘连多肽起到恰当的



作用。Gilbert 等人表示在大肠杆菌中两种粘连寡糖合成酶产生有功能的、体外稳定的，并且可溶的酶(Gilbert 等人, 1998, *Nature Biotechnology* 16: 769-772)。

表达作为来自单一核苷酸构建体的单一多肽的 AHAS 的大亚基和小亚基还具有产生转基因除草剂耐受作物的优势。使用单一基因转化并且将植物培育成原种作物品种比使用两个或更多的基因更容易和更有利。

可以构建包含两个多核苷酸构建体(一个构建体编码 AHASL 多肽并且另一个编码 AHASS 多肽)的植物表达载体。以这种方式，两个基因作为单一基因座分离，这使得除草剂耐受植物的育种更加容易。备选地，可以将大亚基和小亚基融合为表达自单一启动子的单一基因。融合多肽可以具有提高水平的 AHAS 活性和除草剂耐受性。AHAS 大亚基可以是野生型序列的(如果在独立或融合的小亚基存在时赋予了抗性)，或者可以是其自身会对除草剂具有一定水平抗性的突变大亚基。小亚基的存在可以增强大亚基的活性，增强大亚基的除草剂耐受性，增加酶在体内表达时的稳定性，和/或增加对大亚基降解的抗性。以这种方式小亚基可以提高植物/作物对咪唑啉酮或其它除草剂的耐受性。在对转化的植物没有植物毒性下提高的耐受性允许应用除草剂和/或增加杂草控制率的安全性。理想的，赋予的耐受性会提高对已知会干扰 AHAS 的除草剂(例如咪唑啉酮和磺酰脲除草剂)的耐受性。

大亚基和小亚基的关联在原核生物中似乎是高特异性的。例如大肠杆菌具有 3 种 AHASL 同工酶和 3 种 AHASS 同工酶。每种 AHASL 同工酶特异性地与仅仅一种 AHASS 同工酶相关联，尽管所有的亚基在相同的生物中表达(Weinstock 等人, 1992, *J. Bacteriology*, 174:5560-5566)。但是关于来自相同或不同物种或来自相同物种的不同同工酶对的真核 AHASL 和 AHASS 蛋白质之间相互作用特异性知之甚少。

本发明的 AHASS 多肽可以从例如玉米、稻和小麦植物中纯化出来并且可以用在组合物中。同样的，编码本发明 AHASS 蛋白质的分离的多核苷酸分子可以用来在微生物如大肠杆菌中表达本发明 AHASS 多肽。用本

领域普通技术人员已知的任何方法，可以将表达的 AHASS 多肽从大肠杆菌的提取物中纯化出来。

本发明还涉及产生对除草剂有抗性的转基因植物的方法。此类方法包括用包含启动子的植物表达载体转化植物，所述的启动子在植物中推动表达，有效连接至本发明分离的多核苷酸分子上。分离的多核苷酸分子包含编码单子叶植物 AHASS 多肽的核苷酸序列，尤其是包含选自下列的氨基酸序列的 AHASS 多肽：在 SEQ ID NO:2、SEQ ID NO:4 或 SEQ ID NO:5 中所示的氨基酸序列；在 SEQ ID NO:2 中所示的氨基酸序列的氨基酸 77-483、在 SEQ ID NO:4 中所示的氨基酸序列的氨基酸 64-471，和在 SEQ ID NO:5 中所示的氨基酸序列的氨基酸 74-481；或其功能性片段或变体。

本发明还涉及对植物细胞赋予除草剂耐受性的方法。方法包括用下列共转化植物细胞：包含有效连接至编码 AHASL 多肽的核苷酸序列的第一植物可表达启动子的第一植物表达载体，以及包含有效连接至编码本发明 AHASS 多肽的核苷酸序列的第二植物可表达启动子的第二植物表达载体。优选的，编码 AHASL 多肽的核苷酸序列编码真核 AHASL 多肽。在一个实施方案中，编码 AHASL 多肽的核苷酸序列编码植物 AHASL 多肽。在另一个实施方案中，编码 AHASL 蛋白质的核苷酸序列编码单子叶植物 AHASL 多肽。在另一个实施方案中，编码 AHASL 多肽的核苷酸序列编码 AHASL 多肽，对于它来说，已知 AHAS 活性通过本发明 AHASS 多肽来增强。

本发明另外涉及增强表达编码 AHASL 多肽或其突变体或变体的转基因植物的除草剂耐受性的方法。此类方法包括用本发明的 AHASS 多核苷酸分子转化转基因植物。优选的，多核苷酸分子有效连接至能在植物或其至少一种细胞中推动基因表达的启动子。

本发明还提供了在抗除草剂植物的后代植物中增强除草剂抗性的方法。方法包括无性(somatically)或有性地将其基因互补序列包含编码抗除草剂真核 AHASL 多肽的核苷酸序列的植物与用编码本发明 AHASS 多肽的多核苷酸分子转化的植物杂交，以及选择表现出增强的除草剂抗性的后代

植物。在一个实施方案中，所选择的后代包含稳定掺入到其基因组中的编码本发明 AHASS 多肽的多核苷酸分子。当与野生型品种植物的除草剂抗性相比，此类后代植物对至少一种除草剂包含增强的抗性。

本发明还提供了本发明方法产生的转基因植物和后代植物，该植物对干扰 AHAS 酶的除草剂表现出增强的抗性。本发明的组合物和方法可以用来增强植物或宿主细胞对任意种类 AHAS 抑制剂的抗性，所述的 AHAS 抑制剂包括但不限于咪唑啉酮和磺酰脲：triazolopyrimidines(chloransulam-methyl、双氟磺草胺、唑啉磺胺、唑草磺胺、唑啉磺草胺)；嘧啶基(硫代)苯甲酸盐(肟啉草、嘧啶苯甲酸钠、环酯草醚、嘧啶草肟、双嘧啶苯甲酸钠)；以及磺酰氨基-羰基-三唑啉酮(flucarbenzone-Na、prooxycarbazone-Na)。优选的，本发明的除草剂是那些在农业中有用的除草剂，例如咪唑啉酮、磺酰脲、chloransulam-methyl 和双氟磺草胺。在本发明的一个实施方案中，除草剂是包含咪唑啉酮除草剂的可商购除草剂产品，包括但不限于 Backdraft™、Beyond™ 除草剂、Cadre®、Extreme®、Lightning®除草剂、Pursuit®、Raptor®和 Sceptor®。

本发明的一些实施方案涉及使用编码 AHASL 多肽的核苷酸序列。此类核苷酸序列是本领域已知的。本发明不依赖于编码特定 AHASL 多肽的特定核苷酸序列，只是此类 AHASL 多肽的活性要能够通过本发明 AHASS 多肽来增强或增加。优选的，核苷酸序列编码真核 AHASL 多肽。更优选的，核苷酸序列编码植物 AHASL 多肽。编码 AHASL 多肽的核苷酸序列包括在下列登录号中所示的那些：登录号 AAR06607.1 (小果亚麻荠 (*Camelina microcarpa*))、AAK68759.1 (拟南芥(*Arabidopsis thaliana*))、AAK50821.1 (*Amaranthus powellii*)、CAA87083.1 (陆地棉 (*Gossypium hirsutum*))、CAA87084.1 (陆地棉)、CAA18088.1 (虞美人(*Papaver rhoeas*))、BAB20812.1 (稻(*Oryza sativa*))、AAG40279.1 (降低抗性龙葵(*Solanum ptycanthum*))、AAG53548.1 (普通小麦)、AAG53550.1 (普通小麦)、AAM03119.1 (早雀麦(*Bromus tectorum*))，和 AAC14572.1 (大麦(*Hordeum vulgare*))。

本发明的 AHASS 多核苷酸可以用在增强除草剂耐受植物耐受性的方法中。具体地, 此类除草剂耐受植物包含除草剂耐受的或抗除草剂的 AHASL 多肽。此类除草剂耐受植物包括用除草剂耐受的 AHASL 核苷酸序列转化的植物和在其基因组中包含编码除草剂耐受的 AHASL 多肽的内源基因的植物。编码除草剂耐受的 AHASL 多肽的核苷酸序列和包含编码除草剂耐受的 AHASL 多肽的内源基因的除草剂耐受植物是本领域已知的。见例如美国专利号 5,013,659、5,731,180、5,767,361、5,545,822、5,736,629、5,773,703、5,773,704、5,952,553 和 6,274,796; 所有的这些以其全文在此引用作为参考。

多种植物转化载体和转化植物的方法是可用的。见例如, An 等人, 1986, *Plant Physiol.*, 81:301-305; Fry 等人, 1987, *Plant Cell Rep.* 6:321-325; Block, 1988, *Theor. Appl Genet.* 76:767-774; Hinchee 等人, 1990, *Stadler. Genet. Symp.* 203-212; Cousins 等人, 1991, *Aust. J. Plant Physiol.* 18:481-494; Chee 和 Slightom, 1992, *Gene.* 118:255-260; Christou 等人, 1992, *Trends. Biotechnol.* 10:239-246; D'Halluin 等人, 1992, *Bio/Technol.* 10:309-314; Dhir 等人, 1992, *Plant Physiol.* 99:81-88; Casas 等人, 1993, *Proc. Nat. Acad Sci. USA* 90:11212-11216; Christou, 1993, *In Vitro Cell. Dev. Biol.-Plant*; 29P:119-124; Davies 等人, 1993, *Plant Cell Rep.* 12:180-183; Dong 和 Mchughen, 1993, *Plant Sci.* 91:139-148; Franklin 和 Trieu, 1993, *Plant Physiol.* 102:167; Golovkin 等人, 1993, *Plant Sci.* 90:41-52; *Guo Chin Sci Bull.* 38:2072-2078; Asano 等人, 1994, *Plant Cell Rep.* 13; Ayeres 和 Park, 1994, *Crit. Rev. Plant. Sci.* 13:219-239; Barcelo 等人, 1994, *Plant. J.* 5:583-592; Becker 等人, 1994, *Plant. J.* 5:299-307; Borkowska 等人, 1994, *Acta. Physiol Plant.* 16:225-230; Christou, 1994, *Agro. Food. Ind. Hi Tech.* 5: 17-27; Eapen 等人 (1994) *Plant Cell Rep.* 13:582-586; Hartman, 等人 (1994) *Bio-Technology* 12: 919923; Ritala 等人, 1994, *Plant. Mol. Biol.* 24:317-325; 以及 Wan 和 Lemaux, 1994, *Plant Physiol.* 104:3748.

本发明的某些方法包括将多核苷酸构建体引入植物中。如本文用到的,

术语“引入”指的是将多核苷酸构建体以这种方式呈递至植物中，即以所述的方式构建体可以进入到植物细胞的内部。本发明的方法不依赖于将多核苷酸构建体引入植物中的特定方法，只是要多核苷酸构建体进入到植物至少一个细胞的内部。向植物中引入多核苷酸构建体的方法是本领域已知的，包括但不限于稳定转化方法、瞬时转化方法和病毒介导的方法。

如本文用到的，术语“稳定转化”指的是这种转化方法，其中引入到植物中的多核苷酸构建体整合到植物的基因组中并且能够通过其后代遗传。如本文用到的，术语“瞬时转染”指的是这种转化方法，其中引入到植物中的多核苷酸构建体没有整合到植物的基因组中。

对于植物和植物细胞的转化，使用标准技术将本发明的核苷酸序列插入到植物或植物细胞中适合于核苷酸序列表达的本领域已知的任何载体。载体的选择取决于优选的转化技术和待转化的靶标植物物种。在优选的实施方案中，*AHASS* 核苷酸序列有效连接已知用于植物细胞中高水平表达的植物启动子，并且之后将此构建体引入到在其基因组中包含抗除草剂 *AHASL* 等位基因的植物中。此类抗除草剂的 *AHASL* 等位基因可以是天然的或对植物基因组是内源的，或可以通过本领域已知的任何植物转化方法被引入到植物基因组中。以这种方式，抗除草剂的基因(*AHASL*)的效力可以通过大亚基蛋白质的稳定或激活来增强。可以将此方法应用到任何植物物种中；但是当应用到作物植物(尤其是通常在除草剂存在时生长的作物植物)中时是最有益的。

构建植物表达盒和将外源核酸引入植物中的方法通常是本领域已知的并且已经在前面描述过。例如，可以使用瘤诱导(Ti)质粒载体将外源 DNA 引入到植物中。用于外源 DNA 递送的其它方法包括使用 PEG 介导的原生质体转化、电穿孔、显微注射须(*whisker*)，和用于直接 DNA 摄取的生物射弹轰击或微粒轰击(见例如 Vasil 等人的美国专利号 5,405,765; Bilang 等人, 1991, *Gene* 100: 247-250; Scheid 等人, 1991, *Mol. Gen. Genet.*, 228: 104-112; Guerche 等人, 1987, *Plant Science* 52: 111-116; Neuhauser 等人, 1987, *Theor. Appl. Genet.* 75: 30-36; Klein 等人, 1987, *Nature* 327: 70-73;

Howell 等人, 1980, *Science* 208:1265; Horsch 等人, 1985, *Science* 227: 1229-1231; DeBlock 等人, 1989, *Plant Physiology* 91: 694-701; Weissbach 和 Weissbach, eds., 1988, *Methods for Plant Molecular Biology*, Academic Press, Inc. 以及 Schuler 和 Zielinski, eds., 1989, *Methods in Plant Molecular Biology*, Academic Press, Inc.)。转化方法取决于待转化的植物细胞、所用载体的稳定性、基因产物的表达水平和其它参数。

将核苷酸序列引入植物细胞并且随后插入到植物基因组中的其它适当方法包括如 Crossway 等人 (1986, *Biotechniques* 4:320-334) 所描述的显微注射, 如 Riggs 等人 (1986, *Proc. Natl. Acad. Sci. USA* 83:5602-5606) 所描述的电穿孔; 如 Townsend 等人 (美国专利号 5,563,055) 和 Zhao 等人 (美国专利号 5,981,840) 所描述的农杆菌介导的转化; 如 Paszkowski 等人 (1984, *EMBO J.* 3:2717-2722) 所描述的直接基因转移; 在例如 Sanford 等人 (美国专利号 4,945,050)、Tomes 等人 (美国专利号 5,879,918)、Tomes 等人 (美国专利号 5,886,244)、Bidney 等人 (美国专利号 5,932,782)、Tomes 等人 (1995, "Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment," in *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg and Phillips, Springer-Verlag, Berlin; McCabe 等人, 1988, *Biotechnology* 6:923-926) 中所描述的弹击粒子加速; 以及 Lec1 转化(WO 00/28058)。也见, Weissinger 等人, 1988, *Ann. Rev. Genet.* 22:421-477; Sanford 等人, 1987, *Particulate Science and Technology* 5:27-37 (洋葱); Christou 等人, 1988, *Plant Physiol.* 87:671-674 (大豆); McCabe 等人, 1988, *Bio/Technology* 6:923-926 (大豆); Finer 和 McMullen, 1991, *In Vitro Cell Dev. Biol.* 27P:175-182 (大豆); Singh 等人, 1998, *Theor. Appl. Genet.* 96:319-324 (大豆); Datta 等人, 1990, *Biotechnology* 8:736-740 (rice); Klein 等人, 1988, *Proc. Natl. Acad. Sci. USA* 85:4305-4309 (玉米); Klein 等人, 1988, *Biotechnology* 6:559-563 (玉米); Tomes, 美国专利号 5,240,855; Buising 等人, 美国专利号 5,322,783 和 5,324,646; Tomes 等人, 1995, "Direct DNA Transfer into Intact Plant Cells via Microprojectile

Bombardment,” in *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg, Springer-Verlag, Berlin, (玉米); Klein 等人, 1988, *Plant Physiol.* 91:440-444 (玉米); Fromm 等人, 1990, *Biotechnology* 8:833-839 (玉米); Hooykaas-Van Slogteren 等人, 1984, *Nature (London)* 311:763-764; Bowen 等人, 美国专利号 5,736,369 (谷类); Bytebier 等人, 1987, *Proc. Natl. Acad. Sci. USA* 84:5345-5349 (百合科(Liliaceae)); De Wet 等人, 1985, in *The Experimental Manipulation of Ovule Tissues*, ed. Chapman 等人, Longman, NY, pp. 197-209 (花粉); Kaeppler 等人, 1990, *Plant Cell Reports* 9:415-418 和 Kaeppler 等人, 1992, *Theor. Appl. Genet.* 84:560-566 (须(whisker)介导的转化); D'Halluin 等人, 1992, *Plant Cell* 4:1495-1505 (电穿孔); Li 等人, 1993, *Plant Cell Reports* 12:250-255 以及 Christou 和 Ford, 1995, *Annals of Botany* 75:407-413 (稻); Osjoda 等人, 1996, *Nature Biotechnology* 14:745-750 (通过根瘤农杆菌转化的玉米); 它们所有在此引用作为参考。

还可以通过让植物接触病毒或病毒核酸将本发明的多核苷酸引入至植物中。通常, 此类方法包括将本发明多核苷酸构建体掺入病毒 DNA 或 RNA 分子内部。认识到本发明的 AHASS 多肽最初可以作为病毒多蛋白的一部分被合成的, 随后通过体内或体外的蛋白水解来加工以产生所需的重组 AHASS 多肽。另外, 认识到本发明的启动子还包括用于通过病毒 RNA 聚合酶的转录中的启动子。将多核苷酸构建体(包括病毒 DNA 或 RNA 分子)引入到植物中并且在其中表达所编码蛋白质的方法是本领域已知的(见例如美国专利号 5,889,191、5,889,190、5,866,785、5,589,367 和 5,316,931; 在此引用作为参考)。

根据常规方法, 使在其中引入了 AHASS 多核苷酸的本发明细胞生长为植物(见, 例如 McCormick 等人, 1986, *Plant Cell Reports* 5:81-84)。之后这些植物生长, 并且用相同的转化品系或不同的品系对它授粉, 并且可以鉴定所得的具有所需表型特征组成型表达的杂种。可以培育 2 代或多代以确保所需表型特征的表达稳定维持并遗传, 并且随后可以收获种子以确保

实现了所需表型特征的表达。以这种方式，本发明提供了具有本发明 *AHASS* 多核苷酸构建体的转化的种子(也指“转基因种子”)。在一个实施方案中，本发明的 *AHASS* 多核苷酸稳定的掺入到植物基因组中。

本发明可以用于任何植物物种的转化，包括但不限于单子叶植物和双子叶植物。目的植物物种的实例包括但不限于玉米 (*Zea mays*)，芸苔属物种(*Brassica* sp.) (例如欧洲油菜(*B. napus*)、芜菁(*B. rapa*)、芥菜(*B. juncea*))，尤其是可作为种子油来源的芸苔属物种，苜蓿(*Medicago sativa*)，稻 (*Oryza sativa*)、黑麦 (*Secale cereale*)、高粱 (两色蜀黍(*Sorghum bicolor*)、高粱 (*Sorghum vulgare*))、粟 (例如，御谷(*Pennisetum glaucum*)、稷(*Panicum miliaceum*)、谷子 (*Setaria italica*)、龙爪稷(*Eleusine coracana*))、向日葵 (*Helianthus annuus*)、红花(*Carthamus tinctorius*)、小麦 (普通小麦(*Triticum aestivum*)、硬粒小麦(*T. Turgidum ssp. durum*))，大豆(*Glycine max*)、烟草 (*Nicotiana tabacum*)、马铃薯 (*Solanum tuberosum*)、落花生 (*Arachis hypogaea*)、棉 (海岛棉(*Gossypium barbadense*)、陆地棉(*Gossypium hirsutum*))、苕薯(*Ipomoea batatas*)、木薯 (*Manihot esculenta*)、咖啡属物种 (*Coffea* spp.)、椰子 (*Cocos nucifera*)、凤梨(*Ananas comosus*)、柑桔属物种 (*Citrus* spp.)、可可树 (*Theobroma cacao*)、茶(*Camellia sinensis*)、芭蕉属物种(*Musa* spp.)、鳄梨(*Persea americana*)、无花果(*Ficus casica*)、番石榴 (*Psidium guajava*)、芒果(*Mangifera indica*)、油橄榄(*Olea europaea*)、番木瓜 (*Carica papaya*)、腰果(*Anacardium occidentale*)、全缘叶澳洲坚果(*Macadamia integrifolia*)、扁桃 (*Prunus amygdalus*)、甜菜(*Beta vulgaris*)、甘蔗属物种 (*Saccharum* spp.)、燕麦、大麦、蔬菜、观赏植物、松柏类植物。在一个实施方案中，本发明的植物是作物植物(例如，玉米、稻、小麦、甜菜、苜蓿、向日葵、芸苔、大豆、棉、红花、落花生、高粱、粟、烟草等)，优选是谷类植物(例如玉米、稻、小麦、大麦、高粱、黑麦、黑小麦等)，更优选的是玉米、稻和小麦植物。

还应当了解前述涉及本发明优选的实施方案并且在不偏离本发明范围内可以在其中产生多种变化。本发明通过下列实施例进一步说明，所述实施



例不以任何方式解释来限制其范围。相反的，要明确理解的是方法可以适用于多种其它实施方案、修饰和其等价物，并且在阅读本文说明书后，在不偏离本发明实质和/或附属权利要求范围下，它们是本领域的技术人员将联想到的。

## 实施例

### 实施例 1

来自玉米、稻和小麦的全长 *AHASS* 核苷酸序列的鉴定

使用 Concert™植物 RNA 提取液(Invitrogen Corp., Carlsbad, CA, USA), 将总 RNA 从玉米叶组织(栽培种 3394)中提取出来。此总 RNA 库用作使用 Invitrogen's Gene Racer (RLM-RACE) 试剂盒产生第一链 cDNA 文库的来源 RNA。设计用于 cDNA 末端快速扩增(RACE)的引物, 它靶向到少于全长的公开结构域 cDNA 序列(ZmAHASS1a: 登录号 AY105043)的 5'区。此局部序列具有破坏了 cDNA 的 ORF 的序列错误。用局部 ZmAHASS1a cDNA 序列与翻译的 AHASS1 cDNA 进行序列比较指示出可能引起移码的碱基。但是, 在得到实验来源的全长 cDNA 之前, 移码碱基的鉴定是不肯定的。用于 ZmAHASS1a 的 5'RACE 解析的引物是下列:

TTCACAAGGATGGAGAGAAGTATGCGAGCGA	gjb 17 (SEQ ID NO:6)
ACATCACCCCCAGCATTGGATGGTTGA	gjb 18 (SEQ ID NO:7)
AAGCAGCAGAAAATCGCCAGAAACGGG	gjb 42 (SEQ ID NO:8)
AACGCCTCTATCAGGTCTGGGTAAG	gjb 43 (SEQ ID NO:9).

使用 Promega's pGem T-easy 克隆试剂盒对 5'RACE 产物进行 TA 克隆。对 4 个分离的质粒克隆进行测序, 并且确定的核苷酸序列解析了 cDNA ZmAHASS1a 的实验起始密码子。使用与指定了终止密码子的公开局部序列偶联的实验来源起始密码子, 设计全长 cDNA 扩增的引物。进行 PCR, 从来自上述植物组织的第一链 cDNA 文库扩增 ZmAHASS1a cDNA。测序

和分析来自四个独立 cDNA 反应库的 23 个独立克隆, 确定了 ZmAHAS1a 的实验 cDNA 序列并且确定了在公开的局部 cDNA 序列 AY105043 中移码测序错误的身份。

在基于对已知 AHASS 核苷酸序列同源性的专有 EST 数据库中, 鉴定分别对应于 SEQ ID NO:1 和 3 的玉米和小麦 AHAS 核苷酸序列的表达序列标记(EST)。随后使用 cDNA 末端快速扩增法(RACE), 尤其是 5'-RACE 法, 得到全长玉米 cDNA 克隆(Frohman 等人, 1988, *Proc. Natl. Acad. Sci. USA* 85:8998-9002)。对所得 cDNA 测序以得到在 SEQ ID NO:1 中所示的核苷酸序列。

小麦氨基酸序列(SEQ ID NO:5)来自几个重叠简并 EST(核苷酸序列未显示)的核苷酸序列的预测氨基酸序列。重叠群 c5532171 组装自 6 个专有小麦 EST 和 2 个局部序列的基因文库呈递(gi2139744 和 gi21319X)。重叠群表达(Informax, Inc., North Bethesda, MD, USA)用来重复来自原始专有 EST 的上述组装。所得组装跨越完整的基因但是包含多个多态性。这些可能代表了在小麦的 3 个同源基因之间的变异。因此, 在重叠中不存在 100% 的同一性。之后将预测的氨基酸序列(代表共有区)与来自 ZmAHASS1a 和 OsAHASS1 的那些序列以及其它公开序列比对, 并且通过检查用于共有序列的原始核苷酸序列来检测每个“未预料到的”氨基酸。

全长稻 AHASS cDNA 组装自两个公开的 EST(登录号: AU064546 和 AU166867)和一个专有重叠群。稻 AHASS 的核苷酸序列在 SEQ ID NO:3 中示出。推论氨基酸序列在 SEQ ID NO:4 中示出。将本发明的稻 AHASS 核苷酸和氨基酸序列与可从 TIGR(The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850; 在线 [www.tigr.org](http://www.tigr.org))得到的 OsAHASS1 基因组 DNA 注释比较。OsAHASS1 基因组 DNA 注释的 TIGR 参考号是 TIGR 基因临时 id: 8351.t03738 和 8351.t03738。但是, 这些注释不相同于本发明的全长稻 AHASS 核苷酸(SEQ ID NO:3)和氨基酸(SEQ ID NO:4)序列。在核苷酸水平, 在注释和 SEQ ID NO:3 之间存在 2 个外显子差异。在氨基酸水平的差异在图 5 展示的比对中描述。

## 实施例 2

### 来自玉米、稻和小麦的 AHASS 蛋白质

本发明的玉米、稻和小麦 AHASS 蛋白质的氨基酸序列分别在 SEQ ID NO:2、4 和 5 中示出。从本发明氨基酸序列对其它已知植物 AHASS 氨基酸序列的比较中，对每种本发明的氨基酸序列确定叶绿体转运肽的位置。对于玉米 AHASS 蛋白质，叶绿体转运肽对应氨基酸 1-76，并且成熟蛋白质对应 SEQ ID NO:2 的氨基酸 77-483。对于稻 AHASS 蛋白质，叶绿体转运肽对应氨基酸 1-73，并且成熟蛋白质对应 SEQ ID NO:4 的氨基酸 74-481。对于小麦 AHASS 蛋白质，叶绿体转运肽对应氨基酸 1-63，并且成熟蛋白质对应 SEQ ID NO:5 的氨基酸 64-471。

在图 1 中提供了本发明成熟 AHASS 蛋白质氨基酸序列的比对。本发明的 3 个 AHAS 蛋白质每种包含 2 个保守结构域，称为结构域 1 和结构域 2，它们通过可变连接区分开。

将本发明氨基酸序列与其它已知植物 AHASS 氨基酸序列对比。图 2 提供了来自本发明成熟 AHASS 蛋白质与来自植物的已知成熟 AHASS 蛋白质配对比较的氨基酸序列同一性。图 3 和 4 分别提供了结构域 1 和 2 的相似比较的结果。

在说明书中提到的所有公开和专利申请表明了本发明所属领域的技术人员的水平。所有的公开和专利申请根据与每个单独公开或专利申请特定和独立的指示为在本文引用作为参考的相同程度，在本文引用作为参考。尽管通过用于明确理解目的的说明和实施例，已经有些详细的描述了前述发明，要明确的是在附属权利要求范围内可以实行一些改变和修饰。

<110> 巴斯福植物科学有限公司

<120> 单子叶植物 AHASS 序列和使用方法

<130> 15069

<140>

<141>

<160> 12

<170> PatentIn 版本 3.2

<210> 1

<211> 1726

<212> DNA

<213> 玉米 (Zea mays)

<400> 1

```

agcccaccac cagaccccg ccaacctcgc gctccacggc agcggcatgt ccgtcgcctc 60
ctcccaccgc ctgcgaccgt cgcctcggg cccggcgcgg cgcgccggat ccgcggcgcc 120
gcgcgtagcg ttctggccag ccggtgcctc gcccccgat cgcggccgct gcaccgtcgt 180
ggcggcggcg gctgccgtg gcgcggcgg cgaggcatct tcggccccg tggctgtgtc 240
ggctgtcgc cccgcggcg ctgccagga caaggtgcgg cgcacacga tttcggtgtt 300
cgtcggggac gagagcggca tgatcaaccg gatcgcagg gtgttcgcca ggagaggcta 360
caacatcgag tcgctcggc tggggctcaa caaggacaag gccctcttca ccattgtcgt 420
ctccgggacc gacagggtgc tcaaccaagt catcgagcag ctcaataagc ttgtcaacgt 480
cctgagtgtg gaagatctat ctaaagaacc tcaggtgaa agagagctga tgcttataaa 540
actaaatgtt gaacctgatc agcgtcctga ggtcatggtt ttagttgata ttttcagagc 600
aaaagtgtt gatatactg agaaaacact taccatggag gtagctggag atcctggcaa 660
aattgctgca gtgcagagga atctatgaa atttggaatc aaagaaattt gcaggacagg 720
aaaaattgct ttgagacgtg aaaggattg tgcaacagcc cgtttctgac gattttctgc 780
tgcttcttac ccagacctga tagaggcgtt gccgaaaaat ccaactacat ctgtaaatag 840
gacagttaat ggcagttttg gtcaaccatc caatgctggg ggtgatgtt atcctgtgga 900
atcttacgag agcttatcag tgaaccatgt acttgatgct cattgggggtg ttctggatga 960
tgatgatgag actggacttc gctgcatac tctctccatc cttgtgaatg actgtcctgg 1020
tgtcctcaac attgtaacag gagtctttgc tcgcaggggc tacaatatac agagccttgc 1080
tggtggccca gctgagaagg aaggcatttc gcggattaca acagttgttc ctggactgt 1140
tgaatccatt gagaagttag ttcagcagct ttacaagctt attgatgtgc acgaagttca 1200
tgacattacc cactcacctt ttgctgaaag ggaacttatg cttattaagg tttctgtcaa 1260
cactgctgct cggaaaggaaa tcctagatat tgctgaaatc ttccgagcaa aacctgttga 1320
tgtttctgac cacacagtaa cgcttcagct tactggagat cttgacaaga tggttgact 1380
acaaaggtta ttggagccat atggcatctg tgaggtcgcc agaactggac gagtagcact 1440
ggtcccgcaa tcgaaggtgg actccaagta cctccgcggc tactctcttc cactgtagcc 1500
ttgccagteg tgatgatgat ggcacatga agagctcggc tggttgttta tagagcagac 1560
catctcgcgc tgggttgtgt tgtgcgatta caggacttgt ttctttcatg tcgtgaactc 1620
ccccggcctg cgtgtccaat aacgcgtgtt gaggctgcat gtgtattcca acgactgtca 1680
gaataagatg catatcgtat ctgtttaaaa aaaaaaaaaa aaaaaa 1726

```

&lt;210&gt; 2

&lt;211&gt; 483

&lt;212&gt; PRT

&lt;213&gt; 玉米

&lt;400&gt; 2

Met Ser Val Ala Ser Ser His Arg Leu Arg Pro Ser Pro Ser Gly Pro  
 1                    5                    10                    15

Ala Arg Arg Pro Gly Ser Ala Ala Pro Arg Val Ala Phe Trp Pro Ala  
                   20                    25                    30

Gly Ala Ser Pro Pro Tyr Arg Gly Arg Cys Thr Val Val Ala Ala Ala  
                   35                    40                    45

Ala Ala Ala Gly Ala Gly Gly Glu Ala Ser Ser Ala Pro Val Ala Val  
                   50                    55                    60

Ser Ala Val Ala Pro Ala Gly Ala Ala Arg Asp Lys Val Arg Arg His  
 65                    70                    75                    80

Thr Ile Ser Val Phe Val Gly Asp Glu Ser Gly Met Ile Asn Arg Ile  
                   85                    90                    95

Ala Gly Val Phe Ala Arg Arg Gly Tyr Asn Ile Glu Ser Leu Ala Val  
                   100                    105                    110

Gly Leu Asn Lys Asp Lys Ala Leu Phe Thr Ile Val Val Ser Gly Thr  
                   115                    120                    125

Asp Arg Val Leu Asn Gln Val Ile Glu Gln Leu Asn Lys Leu Val Asn  
 130                    135                    140

Val Leu Ser Val Glu Asp Leu Ser Lys Glu Pro Gln Val Glu Arg Glu  
 145                    150                    155                    160

Leu Met Leu Ile Lys Leu Asn Val Glu Pro Asp Gln Arg Pro Glu Val  
                   165                    170                    175

Met Val Leu Val Asp Ile Phe Arg Ala Lys Val Val Asp Ile Ser Glu  
                   180                    185                    190

Lys Thr Leu Thr Met Glu Val Ala Gly Asp Pro Gly Lys Ile Ala Ala  
                   195                    200                    205

Val Gln Arg Asn Leu Trp Lys Phe Gly Ile Lys Glu Ile Cys Arg Thr  
 210                    215                    220

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

<210> 3  
 <211> 1861  
 <212> DNA  
 <213> 稻 (Oryza sativa)

<400> 3  
 ccgcataaac cctactagtg ctaccatggc ttaacccaaa aactaaaccg agtgccccgt 60  
 gccactgtc acacaataca gaccagcccc ccgtcacctt cgagctcgag ccaagccaaa 120  
 cgatgtccgt cgcctccccg caccatctcc gccctcggc gctggcgccg gcatgccgtg 180  
 ctggcggcgt ccccgcgcg ccccgggcgg cgcaccggcc atggtgcccc cgcgtccgcc 240  
 gggccgtcgc cgcggcctcc tccggcggcg gcggcggcga ggcggtgacg gcggtgagcg 300  
 cggcggcggg gggggcggcc gcgagcggcg cgagggacac ggtgcggcgc cacacgatct 360  
 ctgtgttcgt cggcgacgag agcgggatga taaaccggat cggcggggtg ttcgccagga 420  
 gaggtataca catcgagtgc ctcgccgtgg ggctcaacaa ggacaaggcc atgttcacca 480  
 ttgctgtctc cggcacggac aggggtgtca accaagtcac cgagcagctc aacaagcttg 540  
 tcaacgtctt gaatgtggaa gatctatcta aggagccaca ggttgaaaga gagctgatgc 600  
 ttataaaaaa taatgttgaa ccagatcagc gtctgaggt catggtttta gttgatattt 660  
 tccgagcgaa agttgttgat atttcggaga acacccttac catcgaggta actggagatc 720  
 ctggcaaaat tgttgctgtg caaaggaacc tcagcaaat tgggataaaa gaaatttgta 780  
 gaacgggaaa aattgctttg agacgtgaaa aaattggagc aactgcccgc ttctggggat 840  
 tttctgtcgc ttcttaccga gatctcatag aggcatggc caaaaattct cttcttactt 900  
 ctgtaataaa gacagtcaat ggaagtttg atcaaccatc caatgctggg ggcgatgtct 960  
 atcctgtgga accttatgag gtttcatcca tgaaccaagt acttgatgct cactggggcg 1020  
 tccttgatga tgaagattca agtggacttc gatcacatac tctatccatc cttgtcaatg 1080  
 attgccctgg tgttctcaac attgttacag gggcttttgc tcgagaggc tacaatatac 1140  
 agagtcttgc ttagggccca gctgaaaagt caggccttcc gcgtattaca acagttgctc 1200  
 ctggaacaga tgaatccatt gagaagttag ttcagcagct taacaaactt gttgatgtgc 1260  
 atgaggttca agatataact cacttgcctt ttgctgaaag agaacttatg cttatcaagg 1320  
 tttctgtgaa cactgctgct cggagagaca tactagatat tgctgaaatc tccggggcaa 1380  
 aatctgttga tgttctgat cacactgta cgttacagct tactggggat ctgacaaga 1440  
 tggttgcatc acaaaggctg ttggagcctt atggcatctg tgaggtcgcc agaacagggc 1500  
 gagtggcgct ggtcccgcaa tccggtgtcg attccaagta ccttctgggc tactcctttc 1560  
 cgttgtaatc ccaggtcttg tgagaagaaa ggacagtaat aaaatgcttg gtcggttggt 1620  
 tgctacctgt tacagcagag tgtttagta gagtgttga gtcagattcc gttcgttcag 1680  
 ttatgttggt tgttatgatg ctgttctttt gttgttgttt acctcctct ctgtaatgtg 1740  
 ccaatccgct ggcttcttgg ccagtaaaga tcatgatgca agagttgagc ctatgttttc 1800  
 tactgacagg caaaccaaac gtgcactcag ccacttacca tgttctggaa taaaattga 1860  
 a 1861

<210> 4  
 <211> 481  
 <212> PRT  
 <213> 稻

<400> 4  
 Met Ser Val Ala Ser Pro His His Leu Arg Pro Ser Pro Leu Ala Pro  
 1 5 10 15

Ala Cys Arg Ala Gly Gly Val Pro Ala Arg Ala Ala Ala Ala His Arg  
                   20                                  25                                  30

Pro Trp Cys Pro Arg Val Arg Arg Ala Val Ala Ala Ala Ser Ser Gly  
                   35                                  40                                  45

Gly Gly Gly Gly Glu Ala Val Thr Ala Val Ser Ala Ala Ala Val Gly  
           50                                  55                                  60

Ala Pro Ala Ser Ala Ala Arg Asp Thr Val Arg Arg His Thr Ile Ser  
   65                                  70                                  75                                  80

Val Phe Val Gly Asp Glu Ser Gly Met Ile Asn Arg Ile Ala Gly Val  
                                   85                                  90                                  95

Phe Ala Arg Arg Gly Tyr Asn Ile Glu Ser Leu Ala Val Gly Leu Asn  
                   100                                  105                                  110

Lys Asp Lys Ala Met Phe Thr Ile Val Val Ser Gly Thr Asp Arg Val  
           115                                  120                                  125

Leu Asn Gln Val Ile Glu Gln Leu Asn Lys Leu Val Asn Val Leu Asn  
   130                                  135                                  140

Val Glu Asp Leu Ser Lys Glu Pro Gln Val Glu Arg Glu Leu Met Leu  
   145                                  150                                  155                                  160

Ile Lys Ile Asn Val Glu Pro Asp Gln Arg Pro Glu Val Met Val Leu  
                                   165                                  170                                  175

Val Asp Ile Phe Arg Ala Lys Val Val Asp Ile Ser Glu Asn Thr Leu  
                   180                                  185                                  190

Thr Ile Glu Val Thr Gly Asp Pro Gly Lys Ile Val Ala Val Gln Arg  
           195                                  200                                  205

Asn Leu Ser Lys Phe Gly Ile Lys Glu Ile Cys Arg Thr Gly Lys Ile  
   210                                  215                                  220

Ala Leu Arg Arg Glu Lys Ile Gly Ala Thr Ala Arg Phe Trp Gly Phe  
   225                                  230                                  235                                  240

Ser Ala Ala Ser Tyr Pro Asp Leu Ile Glu Ala Leu Pro Lys Asn Ser  
                   245                                  250                                  255

Leu Leu Thr Ser Val Asn Lys Thr Val Asn Gly Ser Phe Asp Gln Pro  
           260                                  265                                  270

Ser Asn Ala Gly Gly Asp Val Tyr Pro Val Glu Pro Tyr Glu Gly Ser



275	280	285
Ser Met Asn Gln Val Leu Asp Ala His Trp Gly Val Leu Asp Asp Glu 290	295	300
Asp Ser Ser Gly Leu Arg Ser His Thr Leu Ser Ile Leu Val Asn Asp 305	310	315 320
Cys Pro Gly Val Leu Asn Ile Val Thr Gly Val Phe Ala Arg Arg Gly 325	330	335
Tyr Asn Ile Gln Ser Leu Ala Val Gly Pro Ala Glu Lys Ser Gly Leu 340	345	350
Ser Arg Ile Thr Thr Val Ala Pro Gly Thr Asp Glu Ser Ile Glu Lys 355	360	365
Leu Val Gln Gln Leu Asn Lys Leu Val Asp Val His Glu Val Gln Asp 370	375	380
Ile Thr His Leu Pro Phe Ala Glu Arg Glu Leu Met Leu Ile Lys Val 385	390	395 400
Ser Val Asn Thr Ala Ala Arg Arg Asp Ile Leu Asp Ile Ala Glu Ile 405	410	415
Phe Arg Ala Lys Ser Val Asp Val Ser Asp His Thr Val Thr Leu Gln 420	425	430
Leu Thr Gly Asp Leu Asp Lys Met Val Ala Leu Gln Arg Leu Leu Glu 435	440	445
Pro Tyr Gly Ile Cys Glu Val Ala Arg Thr Gly Arg Val Ala Leu Val 450	455	460
Arg Glu Ser Gly Val Asp Ser Lys Tyr Leu Arg Gly Tyr Ser Phe Pro 465	470	475 480
Leu		

&lt;210&gt; 5

&lt;211&gt; 471

&lt;212&gt; PRT

&lt;213&gt; 普通小麦

&lt;400&gt; 5

Met Ser Val Ala Thr Ala His His Leu Arg Pro Ser Pro Pro Ala Ala

1

5

10

15

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

275	280	285
Val Leu Asp Asp Glu Asp Ser Asn Gly Leu Arg Ser His Thr Leu Ser		
290	295	300
Ile Leu Val Asn Asp Cys Pro Gly Val Leu Asn Ile Ile Thr Gly Val		
305	310	315
Phe Ala Arg Arg Gly Tyr Ser Ile Gln Ser Leu Ala Val Gly Pro Ala		
325	330	335
Glu Lys Glu Gly Ile Ser Arg Ile Thr Thr Val Val Pro Gly Thr Asp		
340	345	350
Glu Ser Ile Glu Lys Leu Val Gln Gln Leu Tyr Lys Leu Ile Asp Val		
355	360	365
His Lys Val Asp Asp Leu Thr Asp Leu Pro Phe Ala Glu Arg Glu Leu		
370	375	380
Met Val Ile Lys Val Ser Gly Asn Thr Ala Ala Arg Arg Glu Ile Leu		
385	390	395
Asp Ile Gly Asn Ile Phe Arg Ala Glu Cys Val Asp Leu Ser Asp His		
405	410	415
Thr Val Thr Leu Gln Leu Thr Gly Asp Leu Asp Lys Met Val Ala Leu		
420	425	430
Gln Arg Leu Leu Glu Pro Tyr Gly Ile Cys Glu Val Ala Arg Thr Gly		
435	440	445
Arg Ala Ala Leu Ile Arg Glu Ser Arg Val Gly Leu Gln Ser Thr Ser		
450	455	460
Ala Gly Tyr Ser Leu Pro Leu		
465	470	

&lt;210&gt; 6

&lt;211&gt; 31

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 人工序列的描述: 合成引物

&lt;400&gt; 6

ttcacaagga tggagagaag tatgcgagcg a

31

- <210> 7  
 <211> 27  
 <212> DNA  
 <213> 人工序列
- <220>  
 <223> 人工序列的描述: 合成引物
- <400> 7  
 acatcacccc cagcattgga tggttga 27
- <210> 8  
 <211> 27  
 <212> DNA  
 <213> 人工序列
- <220>  
 <223> 人工序列的描述: 合成引物
- <400> 8  
 aagcagcaga aaatcgccag aaacggg 27
- <210> 9  
 <211> 25  
 <212> DNA  
 <213> 人工序列
- <220>  
 <223> 人工序列的描述: 合成引物
- <400> 9  
 aacgcctcta tcaggtctgg gtaag 25
- <210> 10  
 <211> 1502  
 <212> DNA  
 <213> 稻
- <400> 10  
 ggtaccaccg ttgcaggcaa tcaatacctt tccggcccag ttaggcagat ataatgggcc 60  
 tcgtacgtcg cgatattggt tcgacccaaa tgcaaggggc tacaatgggc ctcaagtatg 120  
 tcgcgataac gtttcgttcc aaatacaatg ggctgaatgt ccgttgcaat cgctcgtgct 180  
 caccggggt cttttccctt ctttcctccg gaacgactcg attggttccg atcagatata 240  
 agggatataca tgtgtgcgtg ttaataacta ctacatgttt cccgtgcatt aagatatag 300  
 tctccaatct aaattctaac gtgtgttaga aacacgcgtg tgggtgcgtg tgttgtgtgt 360  
 gatgtgagcg tgggtgtgtg acgtatacgt cctaatttgc aacttaaaaa aagatccggt 420

gatatctcta ttttctgttt ttttagagtc gcgtggaggg agctcttggc atatggtctc 480  
gggccatctc tcggtttcat ctaaaaaatg gcgtgctacg atttacgaga ggaatcgatg 540  
acaacagtac tcacatgctc cacgatctca tgcttaattt ttttcctcac aaaatagtcg 600  
ctaaacacac gagaaacgtg acagctccca cgcacgtaaa caaaaacaaa atcgtcccaa 660  
ccacttggat ggagtaaagc aaacgccgcc tcattgcacc acgtaacgaa tcgattgacg 720  
tgcgcgcccc agaaaaacaa aactacaaaa agcaactccg aaacacgtgg catcatcttt 780  
tggccgtacc gctggcacct ggcatcatca catgtatata tcgacccttc gcacatgcgt 840  
gatactacac tttcacctac aaacccgagc ggctcgggtga cacgtacacg gtacaccctc 900  
caaaccctta agaaaagaaa acctcgcaaa ttcagccatt tcgacgcatc atcatgagtt 960  
ggatggggtg ggtcttttta tttttcgaag ataacctatg tgacctatag agagcgtaga 1020  
cacggagccg aattgccgaa agcctccaaa gcaaaagcag gcagctgtgg gagccgcagc 1080  
cgccggggcc gtcgacgtca gaccaagata ccgaatatcg gtcgggtccc cacgccgcc 1140  
agccgctgcc gctgcccggc caagtagtgc cccaacgcga acgcagggcc acccgtgacc 1200  
catcgcgagc ggatcgcgcc gcggcgcgc gcggcggtg gtgtcacgct cgcactcgca 1260  
ctcgcacacg ccgcacacgc cgtctcccc caaagccaag cggcgcgcgg ccgcgcgggg 1320  
gccagcccag taattttcca ccagccgcct tcgcccctcc actcgcgata aaccctacta 1380  
gtgctaccat ggcttaacct aaaaactaaa ccgagtgcc cgtgcccact gtcacacaat 1440  
acagaccagc cccccgtcac cttcgagctc gagccaagcc aaacgatgtc cgtcgtctca 1500  
ga 1502

<210> 11

<211> 1360

<212> DNA

<213> 稻

<400> 11

gaattccggt gtcgattcca agtaccttcg tggctactcc ttttcgttgt aatcccaggt 60  
cttgtgagaa gaaaggacag taataaaatg cttggtcggt tggttgctac ctgttacagc 120  
agagtgttgt agtcagattc cgttcgttca gttatgttgt ttgttatgat gctgttcttt 180  
tgttgttgtt taccttctc tctgtaatgt gccaatccgc tggcttcttg tccaataaag 240  
atcatgatgc aagagttgag cctatgtttt ctactgacag gcaaaccaaa cgtgactca 300  
gccactcacc atgttctgga ataaaaattg aatccagcgc tttgcaccgt gtaaaatgtt 360  
cgtaataatac catagaaaaa tgtttaaaaa atcatattaa tttattttac aggttttttt 420  
aaactaatac tccacctaatac taatcatgct ctaatgggcc gatagccttg tgtactactg 480  
ggcctgcagg tccaacaagg ctgtcggccc agtggcccac ggcatgaga tgggaagcgg 540  
ggaatagggga gaaggggaaga agaatccgga ttctgttgcg gtggttgttg gaagtttggga 600  
gctctcgggt cgggatcgat cgattgatcg gagaggagg ggaagaactc cggaggggga 660  
ggaggaggag gagcaccgat gacgcggggg aagcagaaga tcgatgcgca gcggcggaac 720  
gcggagaaga accagaagtc gaaggggtcc cagctcgagg cccgcgccgt cggcctcaag 780  
gtcatctgcc ccatctgcaa ggtgacgaat gaatccgtct cctcttccct tctgtcttgc 840  
taatccaacc agtagcttgt agcctgccgt attcgggtatt cctccccag ccctaaggct 900  
agaccgctag attggatgga ttcattggagg attcaggggg atggctggat ctttagcttc 960  
ttgttatcgt ctcatccacg atccaccag ccctatgctc ctatccatg cgatctgaaa 1020  
atctcgagcc gagtgatcga ttttgtccgt cttcgtctat agtccccgga tgcacctcaa 1080  
ttacagggta aggagagaca gggagttgat gtgaaatth cataggggat aatthtctat 1140  
tattgtaatt gttctaactt cgtctctcta agcactgcat ctaccctgcc tgccaagggt 1200  
ggaactagta tgattacatt accctagtga tttcttttg ggagatattg ctgctgttat 1260  
attgtttgtc cattaccatt tctatgataa tcaccgatca agggcgatc cagaaacaaa 1320  
aagttggggg gactcaataa ccgagtacac cgatgtacc 1360

<210> 12  
 <211> 487  
 <212> PRT  
 <213> 稻

<400> 12  
 Met Ser Val Ala Ser Pro His His Leu Arg Pro Ser Pro Leu Ala Pro  
 1 5 10 15  
 Ala Cys Arg Ala Gly Gly Val Pro Ala Arg Ala Ala Ala Ala His Arg  
 20 25 30  
 Pro Trp Cys Pro Arg Val Arg Arg Ala Val Ala Ala Ala Ser Ser Gly  
 35 40 45  
 Gly Gly Gly Gly Glu Ala Val Thr Ala Val Ser Ala Ala Ala Val Gly  
 50 55 60  
 Ala Pro Ala Ser Ala Ala Arg Asp Thr Arg Gly Tyr Asn Ile Glu Ser  
 65 70 75 80  
 Leu Ala Val Gly Leu Asn Lys Asp Lys Ala Met Phe Thr Ile Val Val  
 85 90 95  
 Ser Gly Thr Asp Arg Val Leu Asn Gln Val Ile Glu Gln Leu Asn Lys  
 100 105 110  
 Leu Val Asn Val Leu Asn Leu Glu Ile Pro Arg Ala Glu Thr Phe Tyr  
 115 120 125  
 Pro Asn Leu Val Lys Thr Leu Ala Leu Phe Ile Glu Asn Leu Leu Arg  
 130 135 140  
 Val Glu Asp Leu Ser Lys Glu Pro Gln Val Glu Arg Glu Leu Met Leu  
 145 150 155 160  
 Ile Lys Ile Asn Val Glu Pro Asp Gln Arg Pro Glu Val Met Val Leu  
 165 170 175  
 Val Asp Ile Phe Arg Ala Lys Val Val Asp Ile Ser Glu Asn Thr Leu  
 180 185 190  
 Thr Ile Glu Val Thr Gly Asp Pro Gly Lys Ile Val Ala Val Gln Arg  
 195 200 205  
 Asn Leu Ser Lys Phe Gly Ile Lys Glu Ile Cys Arg Thr Gly Lys Ile  
 210 215 220

Ala Leu Arg Arg Glu Lys Ile Gly Ala Thr Ala Arg Phe Trp Gly Phe  
 225 230 235 240

Ser Ala Ala Ser Tyr Pro Asp Leu Ile Glu Ala Leu Pro Lys Asn Ser  
 245 250 255

Leu Leu Thr Ser Val Asn Lys Thr Val Asn Gly Ser Phe Asp Gln Pro  
 260 265 270

Ser Asn Ala Gly Gly Asp Val Tyr Pro Val Glu Pro Tyr Glu Gly Ser  
 275 280 285

Ser Met Asn Gln Val Leu Asp Ala His Trp Gly Val Leu Asp Asp Glu  
 290 295 300

Asp Ser Ser Gly Leu Arg Ser His Thr Leu Ser Ile Leu Val Asn Asp  
 305 310 315 320

Cys Pro Gly Val Leu Asn Ile Val Thr Gly Val Phe Ala Arg Arg Gly  
 325 330 335

Tyr Asn Ile Gln Ser Leu Ala Val Gly Pro Ala Glu Lys Ser Gly Leu  
 340 345 350

Ser Arg Ile Thr Thr Val Ala Pro Gly Thr Asp Glu Ser Ile Glu Lys  
 355 360 365

Leu Val Gln Gln Leu Asn Lys Leu Val Asp Val His Glu Val Gln Asp  
 370 375 380

Ile Thr His Leu Pro Phe Ala Glu Arg Glu Leu Met Leu Ile Lys Val  
 385 390 395 400

Ser Val Asn Thr Ala Ala Arg Arg Asp Ile Leu Asp Ile Ala Glu Ile  
 405 410 415

Phe Arg Ala Lys Ser Val Asp Val Ser Asp His Thr Val Thr Leu Gln  
 420 425 430

Leu Thr Gly Asp Leu Asp Lys Met Val Ala Leu Gln Arg Leu Leu Glu  
 435 440 445

Pro Tyr Gly Ile Cys Glu Ile Ile Cys Ala Asn Thr Val Thr Val Leu  
 450 455 460

Thr His Ile Val Ser Lys Thr His Ser Pro His Val Met Asp Thr Phe  
 465 470 475 480

Ala Glu Val Asn Thr Cys Phe  
 485





<i>GmAHASS1</i>									
<i>NpAHASS1</i>	80.1								
<i>AtAHASS1</i>	79.8	76.2							
<i>ZmAHASS2</i>	72.3	70.5	72						
<i>OsAHASS2</i>	74.4	74.9	73.8	85					
<i>ZmAHASS1a</i>	76.7	73.1	76.2	74.1	80.1				
<i>OsAHASS1</i>	79.5	75.4	78.5	74.9	80.6	91.5			
<i>TaAHASS1X</i>	74.4	71.8	73.6	71.8	76.9	87.3	88.1		
<i>AtAHASS2</i>	73.1	71	72.3	72.3	72.8	71.5	73.6	69.9	

图 2

<i>ZmAHASS1</i>									
<i>OsAHASS1</i>	93.7								
<i>TaAHASS1X</i>	87.5	90							
<i>AtAHASS1</i>	78.8	79.4	76.3						
<i>GmAHASS1</i>	78.8	78.8	76.3	80					
<i>NpAHASS1</i>	80.6	83.1	80.6	83.1	85.6				
<i>AtAHASS2</i>	78.8	79.4	76.3	78.8	81.2	85			
<i>ZmAHASS2</i>	77.5	79.4	73.7	76.3	75.6	80.6	80		
<i>OsAHASS2</i>	85.6	86.9	81.9	77.5	79.4	86.2	82.5	86.9	

图 3

<i>ZmAHASS2</i>									
<i>OsAHASS2</i>	86.8								
<i>AtAHASS2</i>	72.5	72.5							
<i>ZmAHASS1a</i>	76.7	83.6	74.1						
<i>TaAHASS1X</i>	75.1	80.4	71.4	88.4					
<i>OsAHASS1</i>	76.7	83.1	77.2	91.5	87.3				
<i>GmAHASS1</i>	76.7	78.8	75.7	83.6	79.9	86.8			
<i>NpAHASSX</i>	68.3	70.9	66.7	75.1	72.5	77.2	85.2		
<i>AtAHASS1</i>	75.7	79.9	73.5	83.1	79.9	86.2	88.4	80.4	

图 4

		1	50
OsAHASS1 (SEQ ID NO: 4)	(1)	MSVASPHHLRPSFLAPACRAGGVPARAAAAHRPWC PRVRRVAAAASSGGG	
OsAHASS1 的翻译 (TIGR)	(1)	MSVASPHHLRPSFLAPACRAGGV PARAAAAHRPWC PRVRRVAAAASSGGG	
共有序列	(1)	MSVASPHHLRPSFLAPACRAGGV PARAAAAHRPWC PRVRRVAAAASSGGG	
		51	100
OsAHASS1 (SEQ ID NO: 4)	(51)	GGEAVTAVSAAAVGAPXSAARDTVRRHTISV FVGDESGMINRIAGVFARR	
OsAHASS1 的翻译 (TIGR)	(51)	GGEAVTAVSAAAVGAPASAARDT-----	
共有序列	(51)	GGEAVTAVSAAAVGAP SAARDT	
		101	150
OsAHASS1 (SEQ ID NO: 4)	(101)	GYNIESLAVGLNKDKAMFTI VVSGTDRVLNQVIEQLNKLNVNLN-----	
OsAHASS1 的翻译 (TIGR)	(75)	GYNIESLAVGLNKDKAMFTI VVSGTDRVLNQVIEQLNKLNVNLNLEIPRA	
共有序列	(101)	GYNIESLAVGLNKDKAMFTI VVSGTDRVLNQVIEQLNKLNVNLN	
		151	200
OsAHASS1 (SEQ ID NO: 4)	(145)	-----VEDLSKEPQVERELMLIKINVEPDQRPEVM	
OsAHASS1 的翻译 (TIGR)	(125)	ETFYPNLVKTALFNIENLLRVEDLSKEPQVERELMLIKINVEPDQRPEVM	
共有序列	(151)	VEDLSKEPQVERELMLIKINVEPDQRPEVM	
		201	250
OsAHASS1 (SEQ ID NO: 4)	(175)	VLVDI FRAKVVDISENTLTIEVTGDPGKIVAVQRNLSKFGIKEICRTGKI	
OsAHASS1 的翻译 (TIGR)	(175)	VLVDI FRAKVVDISENTLTIEVTGDPGKIVAVQRNLSKFGIKEICRTGKI	
共有序列	(201)	VLVDI FRAKVVDISENTLTIEVTGDPGKIVAVQRNLSKFGIKEICRTGKI	
		251	300
OsAHASS1 (SEQ ID NO: 4)	(225)	ALRREKIGATARFWGFSAAASYPDLIEALPKNSLLTSVNKTVNGSFDQPSN	
OsAHASS1 的翻译 (TIGR)	(225)	ALRREKIGATARFWGFSAAASYPDLIEALPKNSLLTSVNKTVNGSFDQPSN	
共有序列	(251)	ALRREKIGATARFWGFSAAASYPDLIEALPKNSLLTSVNKTVNGSFDQPSN	
		301	350
OsAHASS1 (SEQ ID NO: 4)	(275)	AGGDVYPVEPEYEGSSMNQVLDAHWGVLDDEDSGLRSHTLSILVNDCPGV	
OsAHASS1 的翻译 (TIGR)	(275)	AGGDVYPVEPEYEGSSMNQVLDAHWGVLDDEDSGLRSHTLSILVNDCPGV	
共有序列	(301)	AGGDVYPVEPEYEGSSMNQVLDAHWGVLDDEDSGLRSHTLSILVNDCPGV	
		351	400
OsAHASS1 (SEQ ID NO: 4)	(325)	LNIVTGVFARRGYNIQSLAVGPAEKSGLSRITTVAPGTDESIEKLVQQLN	
OsAHASS1 的翻译 (TIGR)	(325)	LNIVTGVFARRGYNIQSLAVGPAEKSGLSRITTVAPGTDESIEKLVQQLN	
共有序列	(351)	LNIVTGVFARRGYNIQSLAVGPAEKSGLSRITTVAPGTDESIEKLVQQLN	
		401	450
OsAHASS1 (SEQ ID NO: 4)	(375)	KLVDVHEVQDITHL PFAERE LMLIKVSVNTAARRDILDIAE IFRAKSVDV	
OsAHASS1 的翻译 (TIGR)	(375)	KLVDVHEVQDITHL PFAERE LMLIKVSVNTAARRDILDIAE IFRAKSVDV	
共有序列	(401)	KLVDVHEVQDITHL PFAERE LMLIKVSVNTAARRDILDIAE IFRAKSVDV	
		451	500
OsAHASS1 (SEQ ID NO: 4)	(425)	SDHTVTLQLTGDLDKMVALQRLLPEPYGICEVARTGRVALVR--ESGVDSK	
OsAHASS1 的翻译 (TIGR)	(425)	SDHTVTLQLTGDLDKMVALQRLLPEPYGICEI ICANTVTVLTHIVSKTHSP	
共有序列	(451)	SDHTVTLQLTGDLDKMVALQRLLPEPYGICEI V LL S S	
		501	514
OsAHASS1 (SEQ ID NO: 4)	(473)	YLRGYSFPL-----	
OsAHASS1 的翻译 (TIGR)	(475)	HVMDTFAEVNTCF-	
共有序列	(501)	HL L	

图 5

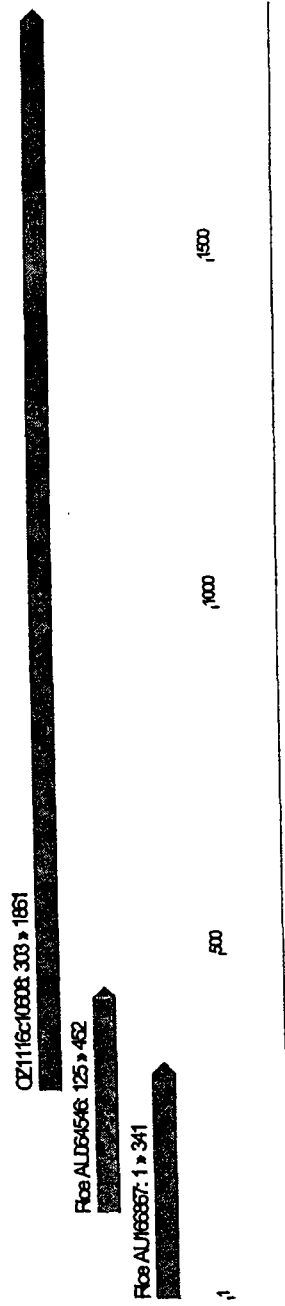


图 6