



(12) 发明专利申请

(10) 申请公布号 CN 112585265 A

(43) 申请公布日 2021.03.30

(21) 申请号 201980055617.4

(22) 申请日 2019.06.21

(30) 优先权数据

18180087.1 2018.06.27 EP

18209806.1 2018.12.03 EP

(85) PCT国际申请进入国家阶段日

2021.02.24

(86) PCT国际申请的申请数据

PCT/EP2019/066480 2019.06.21

(87) PCT国际申请的公布数据

W02020/002152 EN 2020.01.02

(71) 申请人 巴斯夫欧洲公司

地址 德国莱茵河畔路德维希港

(72) 发明人 A·斯卡法尔 A·加勒 J·范里

L·登布尔 K·达伦

(74) 专利代理机构 北京市中咨律师事务所

11247

代理人 张莉 黄革生

(51) Int.Cl.

C12N 9/88 (2006.01)

C12N 15/82 (2006.01)

权利要求书9页 说明书72页

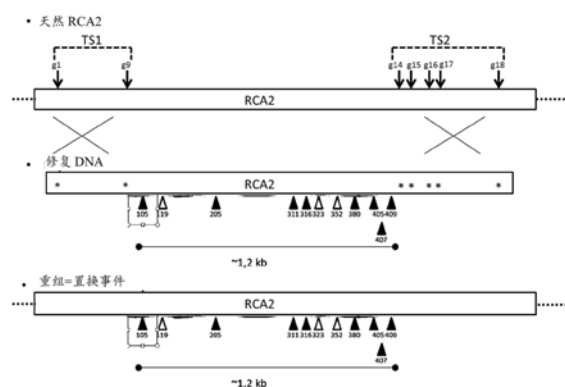
序列表68页 附图6页

(54) 发明名称

热稳定性RUBISCO活化酶及其用途

(57) 摘要

本发明涉及农业领域。具体地，本发明提供热稳定性Rca蛋白、重组基因、包含重组基因的植物、和改善胁迫条件下禾谷类植物耐热性的方法。



1. 一种用于增加禾谷类植物中热稳定性Rca (Rubisco活化酶) 蛋白比率的方法, 包括:
 - a.
 - i. 向禾谷类植物的细胞提供重组基因, 所述重组基因包含以下有效连接的元件:
 1. 启动子, 优选地在植物中可表达;
 2. 核酸, 编码选自以下的热稳定性Rca蛋白:
 - a. Rca 1 β 蛋白及其变体; 和
 - b. 热稳定性Rca 2蛋白变体
 - 和, 任选地
 3. 转录终止和多腺苷酸化区, 优选地在植物中有功能的转录终止和多腺苷酸化区; 和
 - ii. 减少所述禾谷类植物细胞中内源热不稳定性Rca 2蛋白的表达, 其中与不包含所述重组基因的对照禾谷类植物细胞相比, 所述比率增加; 或
 - b. 向禾谷类植物的细胞引入至少一个热稳定性Rca 2等位基因, 其中所述热稳定性Rca 2等位基因编码包含以下的氨基酸:
 - i. SEQ ID NOs:32或35的氨基酸序列或
 - ii. 与SEQ ID NOs:32或35具有90%同一性并且包含选自以下的至少一个氨基酸的氨基酸序列:
 1. 与SEQ ID NO:32或35的位置59相对应的位置处的异亮氨酸;
 2. 与SEQ ID NO:32或35的位置73相对应的位置处的天冬氨酸;
 3. 与SEQ ID NO:32或35的位置160相对应的位置处的异亮氨酸;
 4. 与SEQ ID NO:32或35的位置265相对应的位置处的精氨酸;
 5. 与SEQ ID NO:32或35的位置270相对应的位置处的脯氨酸;
 6. 与SEQ ID NO:32或35的位置277相对应的位置处的亮氨酸;
 7. 与SEQ ID NO:32或35的位置307相对应的位置处的谷氨酸;
 8. 与SEQ ID NO:32或35的位置334相对应的位置处的异亮氨酸;
 9. 与SEQ ID NO:32或35的位置359相对应的位置处的赖氨酸;
 10. 与SEQ ID NO:32或35的位置361相对应的位置处的亮氨酸; 和
 11. 与SEQ ID NO:32或35的位置363相对应的位置处的谷氨酸,
- 其中与不包含所述热稳定性Rca 2等位基因的对照禾谷类植物细胞相比, 所述比率增加。
2. 根据权利要求1所述的方法, 其中所述Rca 1 β 蛋白及其变体包含选自以下的氨基酸序列:
 - a. SEQ ID NO:8的氨基酸序列;
 - b. 与SEQ ID NO:8的氨基酸序列具有至少90%同一性并且包含选自以下的至少一个氨基酸的氨基酸序列:
 - i. 与SEQ ID NO:8的位置109相对应的位置处的异亮氨酸;
 - ii. 与SEQ ID NO:8的位置123相对应的位置处的天冬氨酸;
 - iii. 与SEQ ID NO:8的位置210相对应的位置处的异亮氨酸;
 - iv. 与SEQ ID NO:8的位置315相对应的位置处的精氨酸;
 - v. 与SEQ ID NO:8的位置320相对应的位置处的脯氨酸;

- vi. 与SEQ ID NO:8的位置327相对应的位置处的亮氨酸;
- vii. 与SEQ ID NO:8的位置357相对应的位置处的谷氨酸;
- viii. 与SEQ ID NO:8的位置384相对应的位置处的异亮氨酸;
- ix. 与SEQ ID NO:8的位置409相对应的位置处的赖氨酸;
- x. 与SEQ ID NO:8的位置411相对应的位置处的亮氨酸;和
- xi. 与SEQ ID NO:8的位置413相对应的位置处的谷氨酸。

3. 根据权利要求1或2所述的方法, 其中所述编码Rca 1 β 蛋白及其变体的核酸包含选自以下的编码性核酸序列:

- a. SEQ ID NO:7的核酸或其互补物;
- b. 与SEQ ID NO:7的核酸或其互补物具有60%同一性的核酸。

4. 根据权利要求1所述的方法, 其中所述热稳定性Rca 2蛋白变体包含选自以下的氨基酸序列:

- a. SEQ ID NO:30或33的氨基酸序列;
- b. 与SEQ ID NO:30或33的氨基酸序列具有至少90%同一性并且包含选自以下的至少一个氨基酸的氨基酸序列:

- i. 与SEQ ID NO:30或33的位置105相对应的位置处的异亮氨酸;
- ii. 与SEQ ID NO:30或33的位置119相对应的位置处的天冬氨酸;
- iii. 与SEQ ID NO:30或33的位置206相对应的位置处的异亮氨酸;
- iv. 与SEQ ID NO:30或33的位置311相对应的位置处的精氨酸;
- v. 与SEQ ID NO:30或33的位置316相对应的位置处的脯氨酸;
- vi. 与SEQ ID NO:30或33的位置323相对应的位置处的亮氨酸;
- vii. 与SEQ ID NO:30或33的位置353相对应的位置处的谷氨酸;
- viii. 与SEQ ID NO:30或33的位置380相对应的位置处的异亮氨酸;
- ix. 与SEQ ID NO:30或33的位置405相对应的位置处的赖氨酸;
- x. 与SEQ ID NO:30或33的位置407相对应的位置处的亮氨酸;和
- xi. 与SEQ ID NO:30或33的位置409相对应的位置处的谷氨酸;

5. 根据权利要求1所述的方法, 其中所述热稳定性Rca 2蛋白变体包含选自以下的氨基酸序列:

- a. SEQ ID NO:32或35的氨基酸序列, 并且还包含叶绿体靶向肽;
- b. 与SEQ ID NO:32或35的氨基酸序列具有至少90%同一性的氨基酸序列, 还包含叶绿体靶向肽, 并且包含选自以下的至少一个氨基酸:

- i. 与SEQ ID NO:32或35的位置59相对应的位置处的异亮氨酸;
- ii. 与SEQ ID NO:32或35的位置73相对应的位置处的天冬氨酸;
- iii. 与SEQ ID NO:32或35的位置160相对应的位置处的异亮氨酸;
- iv. 与SEQ ID NO:32或35的位置265相对应的位置处的精氨酸;
- v. 与SEQ ID NO:32或35的位置270相对应的位置处的脯氨酸;
- vi. 与SEQ ID NO:32或35的位置277相对应的位置处的亮氨酸;
- vii. 与SEQ ID NO:32或35的位置307相对应的位置处的谷氨酸;
- viii. 与SEQ ID NO:32或35的位置334相对应的位置处的异亮氨酸;

- ix. 与SEQ ID NO:32或35的位置359相对应的位置处的赖氨酸;
 - x. 与SEQ ID NO:32或35的位置361相对应的位置处的亮氨酸;和
 - xi. 与SEQ ID NO:32或35的位置363相对应的位置处的谷氨酸。
6. 根据权利要求5所述的方法,其中所述叶绿体靶向肽包含选自以下的氨基酸序列:
- a. SEQ ID NO:30从位置1至位置46的氨基酸序列或SEQ ID NO:8从位置1至位置47的氨基酸序列;和
 - b. 与SEQ ID NO:30从位置1至位置46的氨基酸序列或SEQ ID NO:8从位置1至位置47的氨基酸序列具有至少80%同一性的氨基酸序列。
7. 根据权利要求1、4至6中任一项所述的方法,其中编码热稳定性Rca 2蛋白变体的所述核酸包含选自以下的编码性核苷酸序列:
- a. SEQ ID Nos:31、34、36或37的核苷酸序列或其互补物;
 - b. 与SEQ ID Nos:31、34、36或37的核苷酸序列或其互补物具有至少60%同一性的核苷酸序列。
8. 根据权利要求1所述的方法,其中减少内源热不稳定性Rca 2蛋白的表达包括:
- a. 向所述禾谷类植物的细胞引入至少一个敲除突变Rca 2等位基因;或
 - b. 向所述禾谷类植物的细胞提供能够特异性阻抑内源热不稳定性Rca 2基因表达的第二重组基因。
9. 根据权利要求8所述的方法,其中所述敲除突变Rca 2等位基因是来自小麦亚基因组B、A或D的Rca 2 β 基因或来自小麦亚基因组B、A或D的Rca 2 α 基因的敲除突变等位基因。
10. 根据权利要求8所述的方法,其中所述能够特异性阻抑内源热不稳定性Rca 2基因表达的第二重组基因包含以下有效连接的元件:
- a. 启动子,优选地在植物中可表达;
 - b. 核酸,当转录时所述核酸产生对编码热不稳定性Rca蛋白的内源Rca 2基因具有抑制、但不对编码热稳定性Rca蛋白的基因具有抑制的RNA分子;和,任选地
 - c. 转录终止和多腺苷酸化区,优选地在植物中有功能的转录终止和多腺苷酸化区。
11. 根据权利要求8或10所述的方法,其中内源热不稳定性Rca 2基因包含SEQ ID NO:1的编码性核苷酸序列或与SEQ ID NO:1的核苷酸序列具有至少60%同一性并且不编码选自以下的氨基酸的编码性核苷酸序列:
- a. 与SEQ ID NO:4的位置59相对应的位置处的异亮氨酸;
 - b. 与SEQ ID NO:4的位置73相对应的位置处的天冬氨酸;
 - c. 与SEQ ID NO:4的位置160相对应的位置处的异亮氨酸;
 - d. 与SEQ ID NO:4的位置265相对应的位置处的精氨酸;
 - e. 与SEQ ID NO:4的位置270相对应的位置处的脯氨酸;
 - f. 与SEQ ID NO:4的位置277相对应的位置处的亮氨酸;
 - g. 与SEQ ID NO:4的位置307相对应的位置处的谷氨酸;
 - h. 与SEQ ID NO:4的位置334相对应的位置处的异亮氨酸;
 - i. 与SEQ ID NO:4的位置359相对应的位置处的赖氨酸;
 - j. 与SEQ ID NO:4的位置361相对应的位置处的亮氨酸;和
 - k. 与SEQ ID NO:4的位置363相对应的位置处的谷氨酸。

12. 根据权利要求1至11中任一项所述的方法, 其中所述启动子是组成型启动子、组织特异性启动子或诱导型启动子。

13. 根据权利要求1所述的方法, 其中所述热稳定性突变Rca 2等位基因包含SEQ ID NOs:31、34、36或37的编码性核苷酸序列或与SEQ ID NOs:31、34、36或37具有至少60%同一性并且编码蛋白质的编码性核苷酸序列, 其中所述蛋白质包含选自以下的至少一个氨基酸:

- a. 与SEQ ID NO:32或35的位置59相对应的位置处的异亮氨酸;
- b. 与SEQ ID NO:32或35的位置73相对应的位置处的天冬氨酸;
- c. 与SEQ ID NO:32或35的位置160相对应的位置处的异亮氨酸;
- d. 与SEQ ID NO:32或35的位置265相对应的位置处的精氨酸;
- e. 与SEQ ID NO:32或35的位置270相对应的位置处的脯氨酸;
- f. 与SEQ ID NO:32或35的位置277相对应的位置处的亮氨酸;
- g. 与SEQ ID NO:32或35的位置307相对应的位置处的谷氨酸;
- h. 与SEQ ID NO:32或35的位置334相对应的位置处的异亮氨酸;
- i. 与SEQ ID NO:32或35的位置359相对应的位置处的赖氨酸;
- j. 与SEQ ID NO:32或35的位置361相对应的位置处的亮氨酸; 和
- k. 与SEQ ID NO:32或35的位置363相对应的位置处的谷氨酸。

14. 增加禾谷类植物耐热性的方法, 包括:

- a. 根据权利要求1至13中任一项, 增加热稳定性Rca蛋白的比率; 以及
- b. 再生所述禾谷类植物,

其中与不包含所述增加比率的热稳定性Rca蛋白的禾谷类植物相比, 所述耐热性增加。

15. 增加热胁迫条件下禾谷类植物产量的方法, 包括:

- a. 根据权利要求1至13中任一项, 增加热稳定性Rca的比率; 和
- b. 再生所述禾谷类植物,

其中与热稳定性Rca的比率未增加的禾谷类植物的产量相比, 实现所述产量增加。

16. 根据权利要求15所述的方法, 其中所述产量是种子产量。

17. 根据权利要求15所述的方法, 其中所述产量是千粒种子重。

18. 产生耐热性增加的禾谷类植物的方法, 包括

- a. 根据权利要求1至12中任一项, 增加热稳定性Rca的比率; 和
- b. 再生所述禾谷类植物。

19. 热稳定性Rca 2蛋白变体, 其包含选自以下的氨基酸序列:

- a. SEQ ID NO:30或33的氨基酸序列;
- b. 与SEQ ID NO:30或33的氨基酸序列具有至少90%同一性并且包含选自以下的至少一个氨基酸的氨基酸序列:

- i. 与SEQ ID NO:30或33的位置105相对应的位置处的异亮氨酸;
- ii. 与SEQ ID NO:30或33的位置119相对应的位置处的天冬氨酸;
- iii. 与SEQ ID NO:30或33的位置206相对应的位置处的异亮氨酸;
- iv. 与SEQ ID NO:30或33的位置311相对应的位置处的精氨酸;
- v. 与SEQ ID NO:30或33的位置316相对应的位置处的脯氨酸;

vi. 与SEQ ID NO:30或33的位置323相对应的位置处的亮氨酸;
vii. 与SEQ ID NO:30或33的位置353相对应的位置处的谷氨酸;
viii. 与SEQ ID NO:30或33的位置380相对应的位置处的异亮氨酸;
ix. 与SEQ ID NO:30或33的位置405相对应的位置处的赖氨酸;
x. 与SEQ ID NO:30或33的位置407相对应的位置处的亮氨酸;和
xi. 与SEQ ID NO:30或33的位置409相对应的位置处的谷氨酸;
c. SEQ ID NO:32或35的氨基酸序列, 并且任选地还包含叶绿体靶向肽;
d. 与SEQ ID NO:32或35的氨基酸序列具有至少90%同一性的氨基酸序列, 任选地还包含叶绿体靶向肽, 并且包含选自以下的至少一个氨基酸:

- i. 与SEQ ID NO:32或35的位置59相对应的位置处的异亮氨酸;
- ii. 与SEQ ID NO:32或35的位置73相对应的位置处的天冬氨酸;
- iii. 与SEQ ID NO:32或35的位置160相对应的位置处的异亮氨酸;
- iv. 与SEQ ID NO:32或35的位置265相对应的位置处的精氨酸;
- v. 与SEQ ID NO:32或35的位置270相对应的位置处的脯氨酸;
- vi. 与SEQ ID NO:32或35的位置277相对应的位置处的亮氨酸;
- vii. 与SEQ ID NO:32或35的位置307相对应的位置处的谷氨酸;
- viii. 与SEQ ID NO:32或35的位置334相对应的位置处的异亮氨酸;
- ix. 与SEQ ID NO:32或35的位置359相对应的位置处的赖氨酸;
- x. 与SEQ ID NO:32或35的位置361相对应的位置处的亮氨酸;和
- xi. 与SEQ ID NO:32或35的位置363相对应的位置处的谷氨酸。

20. 根据权利要求19所述的热稳定性Rca 2蛋白变体, 其中所述叶绿体靶向肽包含选自以下的氨基酸序列:

a. SEQ ID NO:30从位置1至位置46的氨基酸序列或SEQ ID NO:8从位置1至位置47的氨基酸序列;和

b. 与SEQ ID NO:30从位置1至位置46的氨基酸序列或SEQ ID NO:8从位置1至位置47的氨基酸序列具有至少80%同一性的氨基酸序列。

21. 编码根据权利要求19或20所述的热稳定性Rca 2蛋白变体的核酸, 包含选自以下的编码性核苷酸序列:

a. SEQ ID Nos:31、34、36或37的核苷酸序列或其互补物;

b. 与SEQ ID Nos:31、34、36或37的核苷酸序列或其互补物具有至少60%同一性的核苷酸序列。

22. 重组基因, 包含以下有效连接的元件:

a. 启动子, 优选地在植物中可表达;

b. 核酸, 编码选自以下的Rca蛋白:

a. Rca 1 β 蛋白及其变体, 和

b. 热稳定性Rca 2蛋白变体;

和, 任选地

c. 转录终止和多腺苷酸化区, 优选地在植物中有功能的转录终止和多腺苷酸化区。

23. 根据权利要求22所述的重组基因, 其中所述启动子是组成型启动子、组织特异性启

动子或诱导型启动子。

24. 根据权利要求22或23所述的重组基因,其中所述Rca 1 β 蛋白及其变体包含选自以下的氨基酸序列:

- a. SEQ ID NO:8的氨基酸序列;
- b. 与SEQ ID NO:8的氨基酸序列具有至少90%同一性并且包含选自以下的至少一个氨基酸的氨基酸序列:
 - i. 与SEQ ID NO:8的位置109相对应的位置处的异亮氨酸;
 - ii. 与SEQ ID NO:8的位置123相对应的位置处的天冬氨酸;
 - iii. 与SEQ ID NO:8的位置210相对应的位置处的异亮氨酸;
 - iv. 与SEQ ID NO:8的位置315相对应的位置处的精氨酸;
 - v. 与SEQ ID NO:8的位置320相对应的位置处的脯氨酸;
 - vi. 与SEQ ID NO:8的位置327相对应的位置处的亮氨酸;
 - vii. 与SEQ ID NO:8的位置357相对应的位置处的谷氨酸;
 - viii. 与SEQ ID NO:8的位置384相对应的位置处的异亮氨酸;
 - ix. 与SEQ ID NO:8的位置409相对应的位置处的赖氨酸;
 - x. 与SEQ ID NO:8的位置411相对应的位置处的亮氨酸;和
 - xi. 与SEQ ID NO:8的位置413相对应的位置处的谷氨酸。

25. 根据权利要求22至24中任一项所述的重组基因,其中所述编码Rca 1 β 蛋白及其变体的核酸包含选自以下的编码性核酸序列:

- a. SEQ ID NO:7的核酸或其互补物;
- b. 与SEQ ID NO:7的核酸或其互补物具有60%同一性的核酸。

26. 根据权利要求22或23所述的重组基因,其中所述热稳定性Rca 2蛋白变体包含选自以下的氨基酸序列:

- a. SEQ ID NO:30或33的氨基酸序列;
- b. 与SEQ ID NO:30或33的氨基酸序列具有至少90%同一性并且包含选自以下的至少一个氨基酸的氨基酸序列:
 - i. 与SEQ ID NO:30或33的位置105相对应的位置处的异亮氨酸;
 - ii. 与SEQ ID NO:30或33的位置119相对应的位置处的天冬氨酸;
 - iii. 与SEQ ID NO:30或33的位置206相对应的位置处的异亮氨酸;
 - iv. 与SEQ ID NO:30或33的位置311相对应的位置处的精氨酸;
 - v. 与SEQ ID NO:30或33的位置316相对应的位置处的脯氨酸;
 - vi. 与SEQ ID NO:30或33的位置323相对应的位置处的亮氨酸;
 - vii. 与SEQ ID NO:30或33的位置353相对应的位置处的谷氨酸;
 - viii. 与SEQ ID NO:30或33的位置380相对应的位置处的异亮氨酸;
 - ix. 与SEQ ID NO:30或33的位置405相对应的位置处的赖氨酸;
 - x. 与SEQ ID NO:30或33的位置407相对应的位置处的亮氨酸;和
 - xi. 与SEQ ID NO:30或33的位置409相对应的位置处的谷氨酸;

27. 根据权利要求22或23所述的重组基因,其中所述热稳定性Rca 2蛋白变体包含选自以下的氨基酸序列

- a. SEQ ID NO:32或35的氨基酸序列,任选地还包含叶绿体靶向肽;
 - b. 与SEQ ID NO:32或35的氨基酸序列具有至少90%同一性的氨基酸序列,任选地还包含叶绿体靶向肽,并且包含选自以下的至少一个氨基酸:
 - i. 与SEQ ID NO:32或35的位置59相对应的位置处的异亮氨酸;
 - ii. 与SEQ ID NO:32或35的位置73相对应的位置处的天冬氨酸;
 - iii. 与SEQ ID NO:32或35的位置160相对应的位置处的异亮氨酸;
 - iv. 与SEQ ID NO:32或35的位置265相对应的位置处的精氨酸;
 - v. 与SEQ ID NO:32或35的位置270相对应的位置处的脯氨酸;
 - vi. 与SEQ ID NO:32或35的位置277相对应的位置处的亮氨酸;
 - vii. 与SEQ ID NO:32或35的位置307相对应的位置处的谷氨酸;
 - viii. 与SEQ ID NO:32或35的位置334相对应的位置处的异亮氨酸;
 - ix. 与SEQ ID NO:32或35的位置359相对应的位置处的赖氨酸;
 - x. 与SEQ ID NO:32或35的位置361相对应的位置处的亮氨酸;和
 - xi. 与SEQ ID NO:32或35的位置363相对应的位置处的谷氨酸。
28. 根据权利要求27所述的重组基因,其中所述叶绿体靶向肽包含选自以下的氨基酸序列:
- a. SEQ ID NO:30从位置1至位置46的氨基酸序列或SEQ ID NO:8从位置1至位置47的氨基酸序列;和
 - b. 与SEQ ID NO:30从位置1至位置46的氨基酸序列或SEQ ID NO:8从位置1至位置47的氨基酸序列具有至少80%同一性的氨基酸序列。
29. 根据权利要求22、23、26至28中任一项所述的重组基因,其中编码热稳定性Rca 2蛋白变体的所述核酸包含选自以下的编码性核苷酸序列:
- a. SEQ ID Nos:31、34、36或37的核苷酸序列或其互补物;
 - b. 与SEQ ID Nos:31、34、36或37的核苷酸序列或其互补物具有至少60%同一性的核苷酸序列。
30. 载体,包含根据权利要求22至29中任一项所述的重组基因。
31. 宿主细胞,包含根据权利要求22至29中任一项所述的重组基因或根据权利要求30所述的载体。
32. 根据权利要求31所述的细胞,其是植物细胞。
33. Rca 2基因的敲除等位基因。
34. 根据权利要求33所述的敲除等位基因,其中Rca 2基因是来自小麦亚基因组B、A或D的Rca 2 β 基因或来自小麦亚基因组B、A或D的Rca 2 α 基因。
35. 能够特异性阻抑内源Rca 2基因表达的重组基因,包含以下有效连接的元件:
- a. 启动子,优选地在植物中可表达;
 - b. 核酸,当转录时所述核酸产生对编码热不稳定性Rca蛋白的内源Rca 2基因具有抑制、但对编码热稳定性Rca蛋白的基因不具有抑制的RNA分子;和,任选地
 - c. 转录终止和多腺苷酸化区,优选地在植物中有功能的转录终止和多腺苷酸化区。
36. 根据权利要求35所述的重组基因,其中内源Rca 2基因包含SEQ ID NO:1的编码性核苷酸序列或与SEQ ID NO:1的核苷酸序列具有至少60%同一性的编码性核苷酸序列。

37. 根据权利要求35或36所述的重组基因,其中所述启动子是组成型启动子、组织特异性启动子或诱导型启动子。

38. 载体,包含根据权利要求35至37中任一项所述的重组基因。

39. 宿主细胞,包含根据权利要求35至37中任一项所述的重组基因或根据权利要求38所述的载体。

40. 根据权利要求39所述的细胞,其是植物细胞。

41. Rca 2基因的热稳定性等位基因。

42. 根据权利要求41所述的热稳定性等位基因,其包含

a. SEQ ID NOs:31、34、36或37的编码性核苷酸序列,或
b. 与SEQ ID NO:31、34、36或37具有至少60%同一性并且编码蛋白质的编码性核苷酸序列,其中所述蛋白质包含选自以下的至少一个氨基酸:

- i. 与SEQ ID NO:32或35的位置59相对应的位置处的异亮氨酸;
- ii. 与SEQ ID NO:32或35的位置73相对应的位置处的天冬氨酸;
- iii. 与SEQ ID NO:32或35的位置160相对应的位置处的异亮氨酸;
- iv. 与SEQ ID NO:32或35的位置265相对应的位置处的精氨酸;
- v. 与SEQ ID NO:32或35的位置270相对应的位置处的脯氨酸;
- vi. 与SEQ ID NO:32或35的位置277相对应的位置处的亮氨酸;
- vii. 与SEQ ID NO:32或35的位置307相对应的位置处的谷氨酸;
- viii. 与SEQ ID NO:32或35的位置334相对应的位置处的异亮氨酸;
- ix. 与SEQ ID NO:32或35的位置359相对应的位置处的赖氨酸;
- x. 与SEQ ID NO:32或35的位置361相对应的位置处的亮氨酸;和
- xi. 与SEQ ID NO:32或35的位置363相对应的位置处的谷氨酸。

43. 根据权利要求41或42所述的Rca 2的热稳定性等位基因,其中Rca 2基因是来自亚基因组B、A或D的小麦Rca 2 β 基因或来自亚基因组B、A或D的小麦Rca 2 α 基因。

44. 根据权利要求32所述的植物细胞,还包含:

a. 根据权利要求33或34所述的Rca 2基因的至少一个敲除等位基因;或
b. 根据权利要求35至37中任一项所述的能够特异性阻抑内源Rca 2基因表达的重组基因、或根据权利要求38所述的载体。

45. 禾谷类植物细胞,其包含至少一个根据权利要求41至43中任一项所述的Rca 2基因的热稳定性突变等位基因。

46. 细胞,包含根据权利要求19或20所述的热稳定性Rca 2蛋白变体。

47. 植物、植物部分或种子,基本上由根据权利要求32、40和44至46中任一项所述的植物细胞组成。

48. 根据权利要求47所述的植物、植物部分或种子,其是禾谷类植物、禾谷类植物部分或禾谷类植物种子。

49. 增加禾谷类植物耐热性的方法,包括:

- a. 增加热稳定性Rca蛋白的比率;和
- b. 再生所述禾谷类植物。

50. 根据权利要求49所述的方法,其中所述热稳定性Rca蛋白是Rca 1 β 蛋白或其变体。

51. 根据权利要求49所述的方法, 其中所述热稳定性Rca蛋白是根据权利要求书19或20所述的热稳定性Rca 2蛋白变体。

52. 根据权利要求所述的19或20的热稳定性Rca 2蛋白变体、根据权利要求21所述的编码热稳定性Rca 2蛋白变体的核酸、根据权利要求22至29中任一项所述的重组基因、根据权利要求30所述的载体、根据权利要求35至37中任一项所述的重组基因、根据权利要求38所述的载体、或根据权利要求41至43中任一项所述的Rca 2基因的热稳定性等位基因的用途, 用于增加禾谷植物中热稳定性Rca蛋白的比率、增加禾谷类植物耐热性、增加热胁迫条件下禾谷类植物的产量、或产生耐热性增加的禾谷类植物。

53. 产生食物、饲料或工业产品的方法, 包括

- a. 获得根据权利要求47或48所述的植物、其部分或种子; 和
- b. 从所述植物、其部分或种子制备食物、饲料或工业产品。

54. 根据权利要求53所述的方法, 其中

- a. 食物或饲料是饼粕、粮食、淀粉、面粉或蛋白质; 或
- b. 工业产品是生物燃料、纤维、工业化学品、药物或营养药。

55. 增加Rca 2蛋白的热稳定性的方法, 包括向所述Rca 2蛋白的氨基酸序列引入至少一个氨基酸置换, 其中所述氨基酸置换选自:

- a. 与SEQ ID NO:4的位置59相对应的位置处异亮氨酸置换缬氨酸;
- b. 与SEQ ID NO:4的位置73相对应的位置处天冬氨酸置换甘氨酸;
- c. 与SEQ ID NO:4的位置160相对应的位置处异亮氨酸置换甲硫氨酸;
- d. 与SEQ ID NO:4的位置265相对应的位置处精氨酸置换谷氨酰胺;
- e. 与SEQ ID NO:4的位置270相对应的位置处脯氨酸置换丝氨酸;
- f. 与SEQ ID NO:4的位置277相对应的位置处亮氨酸置换异亮氨酸;
- g. 与SEQ ID NO:4的位置307相对应的位置处谷氨酸置换丝氨酸;
- h. 与SEQ ID NO:4的位置334相对应的位置处异亮氨酸置换缬氨酸;
- i. 与SEQ ID NO:4的位置359相对应的位置处赖氨酸置换苏氨酸;
- j. 与SEQ ID NO:4的位置361相对应的位置处亮氨酸置换甲硫氨酸;
- k. 与SEQ ID NO:4的位置363相对应的位置处谷氨酸置换谷氨酰胺。

56. 根据权利要求55所述的方法, 其中增加热稳定性约7℃。

57. 产生热稳定性Rca 2蛋白变体的方法, 包括培养根据权利要求31所述的宿主细胞并且分离产生的蛋白质。

58. 禾谷类植物, 包含根据权利要求19或20所述的热稳定性Rca 2蛋白变体。

热稳定性RUBISCO活化酶及其用途

技术领域

[0001] 本发明涉及增加禾谷植物中热稳定性Rubisco活化酶(Rca)蛋白的比率和改善禾谷类植物的热胁迫耐受性的方法和手段。具体地说,本发明提供与天然Rca 2蛋白相比增加了热稳定性的Rca 2蛋白变体。

背景技术

[0002] 核酮糖-1,5-二磷酸羧化酶/加氧酶(Rubisco)是使来自大气的惰性CO₂气体转化成糖的光合作用核心酶。Rubisco受严密调节并且Rubisco的活性部位易遭抑制,甚至受其糖底物核酮糖-1,5-二磷酸(RuBP)抑制¹。借助Rubisco的伴侣蛋白酶Rubisco活化酶(Rca),可以对Rubisco进行调节并且从Rubisco活性部位移除抑制物²。Rca是AAA⁺酶家族的成员并且利用ATP从Rubisco活性部位机械地移除紧密结合的抑制物³。Rca的定义性特征之一是,它为热不稳定性蛋白质,并且甚至在施加中等量的热时,在许多植物物种中也会解离、变性并从溶液中聚集,至远高于Rubisco的程度⁴⁻⁶。从而,认为Rca是暴露于高于最适温度的植物缺少光合功能的主导原因之一⁷⁻¹⁰。因此认为改进Rca的热稳定性是在暴露于热胁迫的有害影响的作物中,改进其光合作用并且因此可能改进其产量的最有前景方式之一^{11,12}。

[0003] Rca的热稳定性依赖于物种并且与物种演化所经历的气候相关,温带物种具有相对于热带物种为热不稳定的Rca^{13,14}。甚至来自不同环境的近缘物种,如水稻(*Oryza sativa*)和北澳大利亚炎热环境下特有的近亲澳洲野生稻(*Oryza australiensis*),也具有趋异的热稳定性Rca变体¹⁵。目前利用物种之间Rca热稳定性的遗传多样性来改进更脆弱的Rca变体的稳定性。例如,最近一项研究尝试,基于更为热稳定的来自CAM荒漠植物太匮龙舌兰(*Agave tequilana*)的Rca,来改进稻Rca的热稳定性¹⁶。

[0004] 禾谷植物如小麦(普通小麦(*Triticum aestivum*))是重要的粮食作物。小麦是温带草本植物,小麦的光合作用在刚低于40℃的温度即已经受损¹⁷。因此仍需要改善禾谷植物Rca蛋白(如小麦Rca蛋白)的热稳定性。

发明概述

[0005] 在一个方面,本发明提供一种用于增加禾谷类植物如小麦中热稳定性Rca(Rubisco活化酶)蛋白比率的方法,所述方法包括(a)向禾谷类植物的细胞提供基因,如重组基因,所述基因包含启动子作为有效连接的元件(其中所述启动子优选地在植物中可表达)、编码Rca 1β蛋白及其变体或编码热稳定性Rca2蛋白变体的核酸、和任选地转录终止和多腺苷酸化区,优选地在植物中有功能的转录终止和多腺苷酸化区;并且减少所述禾谷类植物细胞中内源热不稳定性Rca 2蛋白的表达,其中与不包含所述重组基因的对照禾谷类植物细胞相比,所述比率升高;或(b)向禾谷类植物的细胞引入至少一个热稳定性Rca 2等位基因,其中所述热稳定性Rca 2等位基因编码氨基酸序列,所述氨基酸序列包含SEQ ID NOs:32或35的氨基酸序列、或与SEQ ID NOs:32或35具有90%同一性的氨基酸序列,且包含至少一个选自以下的氨基酸:i)与SEQ ID NO:32或35的位置59相对应的位置处的异亮氨酸;ii)与SEQ ID NO:32或35的位置73相对应的位置处的天冬氨酸;iii)与SEQ ID NO:32或

35的位置160相对应的位置处的异亮氨酸;iv)与SEQ ID NO:32或35的位置265相对应的位置处的精氨酸;v)与SEQ ID NO:32或35的位置270相对应的位置处的脯氨酸;vi)与SEQ ID NO:32或35的位置277相对应的位置处的亮氨酸;vii)与SEQ ID NO:32或35的位置307相对应的位置处的谷氨酸;viii)与SEQ ID NO:32或35的位置334相对应的位置处的异亮氨酸;ix)与SEQ ID NO:32或35的位置359相对应的位置处的赖氨酸;x)与SEQ ID NO:32或35的位置361相对应的位置处的亮氨酸;和xi)与SEQ ID NO:32或35的位置363相对应的位置处的谷氨酸,其中与不包含所述热稳定性Rca 2等位基因的对照禾谷类植物细胞相比,所述比率升高。

[0006] 在又一个实施方案中,Rca 1 β 蛋白质及其变体包含氨基酸序列,所述氨基酸序列选自SEQ ID NO:8的氨基酸序列或与SEQ ID NO:8的氨基酸序列具有至少90%同一性且包含至少一个选自以下的氨基酸的氨基酸序列:(i)与SEQ ID NO:8的位置109相对应的位置处的异亮氨酸;(ii)与SEQ ID NO:8的位置123相对应的位置处的天冬氨酸;(iii)与SEQ ID NO:8的位置210相对应的位置处的异亮氨酸;(iv)与SEQ ID NO:8的位置315相对应的位置处的精氨酸;(v)与SEQ ID NO:8的位置320相对应的位置处的脯氨酸;(vi)与SEQ ID NO:8的位置327相对应的位置处的亮氨酸;(vii)与SEQ ID NO:8的位置357相对应的位置处的谷氨酸;(viii)与SEQ ID NO:8的位置384相对应的位置处的异亮氨酸;(ix)与SEQ ID NO:8的位置409相对应的位置处的赖氨酸;(x)与SEQ ID NO:8的位置411相对应的位置处的亮氨酸;和(xi)与SEQ ID NO:8的位置413相对应的位置处的谷氨酸。

[0007] 另外,编码Rca 1 β 蛋白及其变体的核酸可以包含选自以下的编码性核酸序列:SEQ ID NO:7的核酸或其互补物、和与SEQ ID NO:7的核酸或其互补物具有至少60%同一性的核酸。

[0008] 在又一个实施方案中,热稳定性Rca 2蛋白变体包含氨基酸序列,所述氨基酸序列选自SEQ ID NO:30或33的氨基酸序列、和与SEQ ID NO:30或33的氨基酸序列具有至少90%同一性且包含至少一个选自以下的氨基酸的氨基酸序列:(i)与SEQ ID NO:30或33的位置105相对应的位置处的异亮氨酸,(ii)与SEQ ID NO:30或33的位置119相对应的位置处的天冬氨酸,(iii)与SEQ ID NO:30或33的位置206相对应的位置处的异亮氨酸,(iv)与SEQ ID NO:30或33的位置311相对应的位置处的精氨酸,(v)与SEQ ID NO:30或33的位置316相对应的位置处的脯氨酸,(vi)与SEQ ID NO:30或33的位置323相对应的位置处的亮氨酸,(vii)与SEQ ID NO:30或33的位置353相对应的位置处的谷氨酸,(viii)与SEQ ID NO:30或33的位置380相对应的位置处的异亮氨酸,(ix)与SEQ ID NO:30或33的位置405相对应的位置处的赖氨酸,(x)与SEQ ID NO:30或33的位置407相对应的位置处的亮氨酸和(xi)与SEQ ID NO:30或33的位置409相对应的位置处的谷氨酸。

[0009] 另外,热稳定性Rca 2蛋白变体可以包含氨基酸序列,所述氨基酸序列选自:SEQ ID NO:32或35且还包含叶绿体靶向肽的氨基酸序列,或与SEQ ID NO:32或35具有90%同一性、还包含叶绿体靶向肽且包含至少一个选自以下的氨基酸的氨基酸序列:(i)与SEQ ID NO:32或35的位置59相对应的位置处的异亮氨酸,(ii)与SEQ ID NO:32或35的位置73相对应的位置处的天冬氨酸,(iii)与SEQ ID NO:32或35的位置160相对应的位置处的异亮氨酸,(iv)与SEQ ID NO:32或35的位置265相对应的位置处的精氨酸,(v)与SEQ ID NO:32或35的位置270相对应的位置处的脯氨酸,(vi)与SEQ ID NO:32或35的位置277相对应的位置

处的亮氨酸, (vii) 与SEQ ID NO:32或35的位置307相对应的位置处的谷氨酸, (viii) 与SEQ ID NO:32或35的位置334相对应的位置处的异亮氨酸, (ix) 与SEQ ID NO:32或35的位置359相对应的位置处的赖氨酸, (x) 与SEQ ID NO:32或35的位置361相对应的位置处的亮氨酸和 (xi) 与SEQ ID NO:32或35的位置363相对应的位置处的谷氨酸。此外, 编码热稳定性Rca 2蛋白质变体的核酸可以包含选自以下的编码性核苷酸序列: (a) SEQ ID Nos:31、34、36或37的核苷酸序列或其互补物, (b) 与SEQ ID Nos:31、34、36或37的核苷酸序列或其互补物具有至少60%同一性的核苷酸序列。

[0010] 在另一个实施方案中, 减少内源热不稳定性Rca 2蛋白的表达包括: 向所述禾谷类植物的细胞引入至少一个敲除突变Rca 2等位基因, 或向所述禾谷类植物的细胞提供能够特异性阻抑内源热不稳定性Rca 2基因表达的第二重组基因。

[0011] 在又一个实施方案中, 能够特异性阻抑内源热不稳定性Rca 2基因表达的第二重组基因包含以下有效相连的元件: (a) 启动子, 其优选地在植物中可表达, (b) 核酸, 当转录时所述核酸产生对编码热不稳定性Rca蛋白的内源Rca 2基因具有抑制、但对编码热稳定性Rca蛋白的基因不具有抑制的RNA分子; 和, 任选地 (c) 转录终止和多腺苷酸化区域, 优选地在植物中有功能的转录终止和多腺苷酸化区。

[0012] 在又一个实施方案中, 热稳定性突变体Rca 2等位基因包含SEQ ID NOs:31、34、36或37的编码性核苷酸序列或与SEQ ID NOs:31、34、36或37具有至少60%同一性且编码蛋白质的编码性核苷酸序列, 所述蛋白质包含选自以下氨基酸中的至少一者: (i) 与SEQ ID NO:32或35的位置59相对应的位置处的异亮氨酸, (ii) 与SEQ ID NO:32或35的位置73相对应的位置处的天冬氨酸, (iii) 与SEQ ID NO:32或35的位置160相对应的位置处的异亮氨酸, (iv) 与SEQ ID NO:32或35的位置265相对应的位置处的精氨酸, (v) 与SEQ ID NO:32或35的位置270相对应的位置处的脯氨酸, (vi) 与SEQ ID NO:32或35的位置277相对应的位置处的亮氨酸, (vii) 与SEQ ID NO:32或35的位置307相对应的位置处的谷氨酸, (viii) 与SEQ ID NO:32或35的位置334相对应的位置处的异亮氨酸, (ix) 与SEQ ID NO:32或35的位置359相对应的位置处的赖氨酸, (x) 与SEQ ID NO:32或35的位置361相对应的位置处的亮氨酸, 和 (xi) 与SEQ ID NO:32或35的位置363相对应的位置处的谷氨酸。

[0013] 还提供用于增加禾谷类植物的耐热性、用于增加热胁迫条件下禾谷类植物的产量、和用于产生耐热性增加的禾谷类植物的方法。这些方法包括根据本发明增加热稳定性Rca蛋白的比率并且再生禾谷类植物。

[0014] 在另一个方面, 本发明提供一种热稳定性Rca蛋白变体和编码它的核酸, 其中所述热稳定性Rca2蛋白变体包含选自以下的氨基酸序列: (a) SEQ ID NO:30或33的氨基酸序列; 和 (b) 氨基酸序列, 其与SEQ ID NO:30或33的氨基酸序列具有90%同一性, 并且包含至少一个选自以下的氨基酸: (i) 与SEQ ID NO:30或33的位置105相对应的位置处的异亮氨酸, (ii) 与SEQ ID NO:30或33的位置119相对应的位置处的天冬氨酸, (iii) 与SEQ ID NO:30或33的位置206相对应的位置处的异亮氨酸, (iv) 与SEQ ID NO:30或33的位置311相对应的位置处的精氨酸, (v) 与SEQ ID NO:30或33的位置316相对应的位置处的脯氨酸, (vi) 与SEQ ID NO:30或33的位置323相对应的位置处的亮氨酸, (vii) 与SEQ ID NO:30或33的位置353相对应的位置处的谷氨酸, (viii) 与SEQ ID NO:30或33的位置380相对应的位置处的异亮氨酸, (ix) 与SEQ ID NO:30或33的位置405相对应的位置处的赖氨酸, (x) 与SEQ ID NO:30

或33的位置407相对应的位置处的亮氨酸,和(xi)与SEQ ID NO:30或33的位置409相对应的位置处的谷氨酸;(c)SEQ ID NO:32或35的氨基酸序列,并且其任选地还包含叶绿体靶向肽;(d)氨基酸序列,其与SEQ ID NO:32或35的氨基酸序列具有90%同一性,任选地还包含叶绿体靶向肽,并且包含至少一个选自以下的氨基酸:(i)与SEQ ID NO:32或35的位置59相对应的位置处的异亮氨酸,(ii)与SEQ ID NO:32或35的位置73相对应的位置处的天冬氨酸,(iii)与SEQ ID NO:32或35的位置160相对应的位置处的异亮氨酸,(iv)与SEQ ID NO:32或35的位置265相对应的位置处的精氨酸,(v)与SEQ ID NO:32或35的位置270相对应的位置处的脯氨酸,(vi)与SEQ ID NO:32或35的位置277相对应的位置处的亮氨酸,(vii)与SEQ ID NO:32或35的位置307相对应的位置处的谷氨酸,(viii)与SEQ ID NO:32或35的位置334相对应的位置处的异亮氨酸,(ix)与SEQ ID NO:32或35的位置359相对应的位置处的赖氨酸,(x)与SEQ ID NO:32或35的位置361相对应的位置处的亮氨酸;和(xi)与SEQ ID NO:32或35的位置363相对应的位置处的谷氨酸。还提供包含本发明的耐热性Rca 2蛋白变体的禾谷类植物。

[0015] 在又一个方面,基因,如重组基因,包含以下有效连接的元件:(a)启动子,优选地在植物中可表达,(b)编码选自以下的Rca蛋白的核酸:(i)Rca 1 β 蛋白及其变体;和(ii)热稳定性Rca 2蛋白变体和,任选地(c)转录终止和多腺苷酸化区,优选地在植物中有功能的转录终止和多腺苷酸化区。在其他实施方案中,Rca 1 β 蛋白及其变体包含如上文所述的氨基酸序列并且热稳定性Rca 2蛋白变体包含根据本发明的氨基酸序列。

[0016] 进一步提供Rca 2基因的敲除等位基因和能够特异性阻抑内源Rca 2基因表达的重组基因,所述重组基因包含以下有效连接的元件:(a)启动子,其优选地在植物中可表达,(b)核酸,当转录时所述核酸产生对编码热不稳定性Rca蛋白的内源Rca 2基因具有抑制、但对编码热稳定性Rca蛋白的基因不具有抑制的RNA分子;和,任选地(c)转录终止和多腺苷酸化区域,优选地在植物中有功能的转录终止和多腺苷酸化区。

[0017] 在另一个实施方案中,提供Rca 2基因的热稳定性等位基因。本发明的热稳定性等位基因可以包含(a)SEQ ID NOs:31、34、36或37的编码性核苷酸序列,或(b)与SEQ ID NO:31、34、36或37具有60%同一性并且编码蛋白质的编码性核苷酸序列,其中所述蛋白质包含选自以下氨基酸中至少一者:(i)与SEQ ID NO:32或35的位置59相对应的位置处的异亮氨酸,(ii)与SEQ ID NO:32或35的位置73相对应的位置处的天冬氨酸,(iii)与SEQ ID NO:32或35的位置160相对应的位置处的异亮氨酸,(iv)与SEQ ID NO:32或35的位置265相对应的位置处的精氨酸,(v)与SEQ ID NO:32或35的位置270相对应的位置处的脯氨酸,(vi)与SEQ ID NO:32或35的位置277相对应的位置处的亮氨酸,(vii)与SEQ ID NO:32或35的位置307相对应的位置处的谷氨酸,(viii)与SEQ ID NO:32或35的位置334相对应的位置处的异亮氨酸,(ix)与SEQ ID NO:32或35的位置359相对应的位置处的赖氨酸,(x)与SEQ ID NO:32或35的位置361相对应的位置处的亮氨酸和(xi)与SEQ ID NO:32或35的位置363相对应的位置处的谷氨酸。

[0018] 在另一个方面,提供细胞,其包含本发明的(重组)基因、如本文所述的Rca 2基因的至少一个敲除等位基因、特异性阻抑内源Rca 2基因表达的重组基因、根据本发明的至少一个Rca 2基因热稳定性等位基因和/或本文所述的热稳定性Rca 2蛋白变体。还提供基本上由本发明细胞组成的禾谷类植物、植物部分或种子。这些禾谷类植物、植物部分或种子可

以是小麦植物、小麦植物部分或小麦种子。

[0019] 还提供本发明的热稳定性Rca 2蛋白变体、编码本发明热稳定性Rca 2蛋白变体的核酸、本发明的重组基因、本文所述的能够特异性阻抑内源Rca 2基因表达的重组基因、或本文提供的Rca 2基因的热稳定性等位基因的用途,用于增加禾谷植物中热稳定性Rca蛋白的比率、增加禾谷类植物耐热性、增加热胁迫条件下禾谷类植物的产量、或产生耐热性增加的禾谷类植物。

附图简述

[0020] 图1.比较在对照(22℃昼/夜)或热处理(38/22℃,两个昼夜周期)后从小麦叶提取的Rca的基因表达和温度依赖性Rubisco活化速度。(A)采用对小麦Rca1 β (TaRca1- β)基因和Rca2基因 α (TaRca2- α)和Rca2 β 剪接变体(TaRca2- β)特异的引物,依据qPCR的基因表达。(B)Rca活化Rubisco的绝对初始速度,其中所述Rca从对照和热处理的小麦叶提取出来,在0.2mM ATP存在下于所示温度温育10分钟,之后在标准25℃进行分析。(C)小麦叶Rca的Rubisco活化速度vs.温育温度,其中所述速度相对于实现的最快速度进行了归一化。值是三个或更多个生物学重复的均值 $\pm$ SD。曲线是可变斜率模型的普通最小二乘法拟合(方程1)。

[0021] 图2.对于重组小麦Rca2 α (TaRca2- α)、 β (TaRca2- β)、Rca1 β (TaRca1- β)和稻Rca β (OsRca- β)同种型,Rca在25℃的Rubisco活化速度(A)、Rca的温度依赖性Rubisco活化(B)、和差示扫描荧光测定(C)。Rubisco活化实验为将 $1.4 \pm 0.2 \mu\text{M}$ 的Rca原体(protoomer)添加至 $0.2 \pm 0.05 \mu\text{M}$ 受RuBP(ER)抑制的Rubisco活性部位。对于温度响应曲线,将Rca在0.2mM ATP存在下于所示温度温育10分钟,之后在标准25℃进行分析,将值相对于实现的最快速度进行归一化。对于DSF,将样品按1℃/分钟加热并且将荧光信号相对于记录的最大值进行归一化。值是四个实验重复的均值 $\pm$ SD。曲线是玻尔兹曼Sigmoidal方程的普通最小二乘法拟合(方程1和2)。

[0022] 图3.小麦Rca2 β (TaRca2- β)、Rca1 β (TaRca1- β)、稻Rca β (OsRca- β)、暖适应物种和冷适应物种的共有序列、有11个氨基酸变化(TaRca2- β -11AA)和八氨基酸变化(TaRca2- β -8AA)的小麦Rca2 β 突变体的序列比对。共有序列分别从温暖环境和寒冷环境特有的八个物种和九个物种的比对生成。基于TaRca1- β 序列选择在空心三角形和实心三角形所指示的位置对TaRca2- β -11AA所做的突变,标准为TaRca1- β 匹配暖物种共有序列并且异于冷物种共有序列。基于上述标准和以下额外标准,选择在实心三角形指示的位置对TaRca2- β -8AA所做的突变,其中所述额外的标准为OsRca- β 不能匹配TaRca2- β 或冷物种共有序列。突出显示序列之间的残基差异。下划线显示未纳入分析的叶绿体信号肽。

[0023] 图4.对于图2中所呈现的重组小麦Rca2 β (TaRca2- β)、Rca1 β (TaRca1- β)和稻Rca β (OsRca- β)同种型以及TaRca2- β 11氨基酸突变体(TaRca2- β -11AA)和8氨基酸突变体(TaRca2- β -8AA),Rca在25℃的Rubisco活化速度(A)、Rca的温度依赖性Rubisco活化(B)和差示扫描荧光测定(C)。值是四个或更多个实验重复的均值 $\pm$ SD。曲线是玻尔兹曼Sigmoidal方程的普通最小二乘法拟合(方程1和2)。

[0024] 图5.实施例6中遵循的基因替换策略的示意图。TS1:靶位点1、TS2:靶部位2、箭头标记的g1、g9、g14至18代表向导RNA在基因组DNA中引起双链断裂的位置。以空心三角形和实心三角形指示TaRca2- β -11AA中的氨基酸置换连同位置,而以实心三角形指示TaRca2- β -

8AA中的氨基酸置换连同位置。星号代表跨切割位点的沉默突变。

发明详述

[0025] 本发明基于令人惊讶的发现：小麦Rca 1 β 蛋白是热稳定性Rca同种型，并且热不稳定性Rca 2蛋白上有限数目的氨基酸变化可以导致体外Rubisco活化测定法中Rca 2蛋白的热稳定性改进至少7°C。

定义

[0026] 如本文对蛋白质序列所用，术语“序列同一性百分数”指在最佳对齐的多肽的窗口的二个区段之间保守氨基酸的百分数。本领域技术人员熟知用于对齐比较窗口的序列最佳比对并且可以通过矩阵如BLOSUM(模块替代矩阵)和PAM(点接受的突变)(Henikoff和Henikoff1992,PNAS 89(22):10915-10919)计算保守百分数。测试序列和参考序列的已比对区段的“同一性分数”是两个已比对序列共有的相同或保守残基的数目除以参考序列区段(即完整的参考序列或参考序列的较小限定的部分)中残基的总数目。序列同一性百分数表述为同一性分数乘以100。一个或多个蛋白质序列的比较可以是相对于全长蛋白质序列或其部分，或相对于更长的蛋白质序列进行。

[0027] 在本文中，术语“蛋白质”可与术语“多肽”互换使用，描述一组由多于30个氨基酸组成的分子，而术语“肽”描述由最多30个氨基酸组成的分子。蛋白质和肽可以进一步形成二聚体、三聚物 and 高级低聚物，即由不止一个(多)肽分子组成。形成这类二聚体、三聚物等的蛋白质或肽分子可以是相同或不相同的。相应的更高阶结构因此称作同型二聚体或异二聚体、同型三聚体或异三聚体等。术语“蛋白质”和“肽”也指天然修饰的蛋白质或肽，其中例如通过糖基化、乙酰化、磷酸化等实现所述修饰。这类修饰是本领域熟知的。

[0028] 相对于本发明的核苷酸序列而言，术语“变体”意指基本上相似的序列。可以使用熟知的分子生物学技术，例如采用如本文先前描述的聚合酶链反应(PCR)和杂交技术，鉴定天然存在的等位基因变体(如这些变体)。变体核苷酸序列还包括合成方式衍生的核苷酸序列，例如通过利用对SEQ ID NOs:7、46或48中任一者的位点定向诱变所生成的那些。通常，本发明的核苷酸序列变体将与天然(野生型或内源)的核苷酸序列具有至少40%、至少50%、至少60%，到至少70%，例如，优选地至少71%、至少72%、至少73%、至少74%、至少75%、至少76%、至少77%、至少78%，到至少79%、通常至少80%，例如，至少81%到至少84%、至少85%，例如，至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%，到至少98%和至少99%核苷酸序列同一性。本文公开的DNA分子的衍生物可以包括但不限于，序列缺失、单点或多点突变、在特定限制性酶位点处改变、添加功能性元件、或其他的分子修饰手段。用于获得这类衍生物的技术是本领域熟知(参见，例如，J.F.Sambrook,D.W.Russell和N.Irwin(2000)Molecular Cloning:A Laboratory Manual,第3版,第1、2和3卷,Cold Spring Harbor Laboratory Press)。本领域技术人员熟悉标准来源资料，所述标准来源资料描述构建、操作和分离大分子(例如，DNA分子、质粒等)以及产生重组生物和筛选及分离DNA分子的具体条件和方法。

[0029] 如本文对核苷酸序列所用，术语“序列同一性百分数”指在最佳对齐的DNA的窗口的二个区段之间相同核苷酸的百分数。用于对齐比较窗口的序列最佳比对是本领域技术人员熟知的并且可以由以下工具实施：如Smith和Waterman的局部同源性算法(Waterman,

M.S.Introduction to Computational Biology:Maps,sequences and genomes.Chapman&Hall.London(1995)、Needleman和Wunsch的同源性比对算法(J.Mol.Biol.,48:443-453(1970)、Pearson和Lipman的相似性检索方法(Proc.Natl.Acad.Sci.,85:2444(1988)),并且优选地使用这些算法的计算机化执行,如作为GCG(注册商标),Wisconsin Package(来自加利福尼亚州圣迭戈Accelrys Inc.的注册商标)的部分可获得的GAP、BESTFIT、FASTA和TFASTA。测试序列和参考序列的已比对区段的“同一性分数”是两个已比对序列共有的相同组分的数目除以参考序列区段(即完整的参考序列或参考序列的较小限定的部分)中组分的总数目。序列同一性百分数表述为同一性分数乘以100。一个或多个DNA序列的比较可以是相对于全长DNA序列或其部分,或相对于更长的DNA序列进行。

[0030] 术语“重组基因”指任何人工基因,其含有a) DNA序列,包括在自然界中并非一起存在的调节序列和编码性序列,或b) 编码并不天然毗连的蛋白质部分的序列,或c) 并不天然毗连的启动子部分。因此,重组基因可以包含衍生自不同来源的调节序列和编码性序列,即异源序列,或包含衍生自相同来源、但按照与自然界中存在不同的方式排列的调节序列和编码性序列。

[0031] 术语“异源”指衍生自不同来源的两个或更多个核酸序列或蛋白质序列之间的关系。例如,启动子相对于有效连接的DNA区域(如编码性序列)是异源,如果正常情况下这种组合不存在于自然界中。此外,特定序列可以相对于插入该序列的细胞或生物是“异源”的(即不天然存在于这种特定细胞或生物中)。例如,本文公开的重组基因是异源核酸。

[0032] 术语“内源”指源自植物或细胞内部。内源基因因此是最初存在于给定植物或细胞中的基因。

[0033] “分离的核酸”,与如本文所用的“分离的DNA”可互换地使用,指不在其天然基因组背景下出现的核酸,无论其长度和序列是什么。分离的DNA可以例如指与其基因组背景物理分离的DNA,如基因组DNA的片段。分离的DNA也可以是人工产生的DNA,如化学合成的DNA,或如借助扩增反应(如本领域熟知的聚合酶链反应(PCR))产生的DNA。分离的DNA还可以指在非其天然存在的DNA背景下存在的DNA。例如,分离的DNA可以指质粒中存在的一段DNA。另外,分离的DNA可以指在与它天然存在的背景相异的另一个染色体背景下(例如在基因组中异于天然位置的另一个位置、在与它天然存在的物种相异的另一个物种的基因组中、或在人工染色体中)存在的一段DNA。

[0034] 当两个核酸分子在适宜条件下彼此退火时,杂交出现。核酸杂交是DNA操作领域技术人员熟知的技术。给定的核酸对的杂交特性是其相似性或同一性的指标。两个核酸序列基本上相同的另一指标是这两个分子在严格条件下彼此杂交。短语“特异性杂交”指,某分子当存在于复杂混合物(例如,总细胞)DNA或RNA中时,在严格条件下仅与特定核苷酸序列结合、形成双链体或杂交。“基本上结合”指探针核酸和靶核酸之间的互补性杂交,并且包括可以通过降低杂交介质的严格性以实现所需的靶核酸序列检测而容许的少量错配。在核酸杂交实验如DNA印迹和RNA印迹杂交的背景下,“严格杂交条件”和“严格杂交洗涤条件”是序列依赖性的,并且在不同的环境参数下不同。高度严格的洗涤条件的一个例子是0.15M NaCl在72℃持续约15分钟。严格洗涤条件的一个例子是在65℃时0.2X SSC洗涤液持续15分钟。也可以通过添加去稳定剂如甲酰胺实现严格性条件。通常,在特定杂交测定法中,与使用不相关探针所观察到的相比,2X(或更高)信噪比表示检测到特异性杂交。在严格条件下

彼此不杂交的核酸如果编码的蛋白质基本上相同,则这些核酸仍是基本上相同的。例如,当使用遗传密码所允许的最大密码子简并性产生核酸副本时,这种情况出现。

[0035] 短语“DNA”、“DNA序列”、“核酸序列”、“核酸分子”、“核苷酸序列”和“核酸”指包含核苷酸的有序排列的物理结构。术语“序列”和“分子”可互换地使用。DNA序列或核苷酸序列可以含于更大的核苷酸分子、载体等内部。此外,这些序列中核酸的有序排列可以用序列列表、图、表、电子介质等形式描述。

[0036] 如本文所用,“包含”应解释为特指存在如所提到的特征、整数、步骤或组分,但是不排除存在或附加一种或多种特征、整数、步骤或组分或它们的群组。因此,例如,包含一个核苷酸或氨基酸序列的核酸或蛋白质可以包含比实际所引述更多的核苷酸或氨基酸,即,包含于更大的核酸或蛋白质中。包含以结构方式或功能方式定义的核酸的重组基因可以包含额外的DNA区域等。然而,在本公开的背景下,术语“包含”也包括“由……组成”。

[0037] 可以理解当提及单数形式的某个词汇(例如一株植物或一个等位基因)时,本文中还包括复数形式(例如多株植物、多个等位基因)。因而,借助非限定冠词“a(一个)”或“an(一种)”对某要素的提及不排除存在多于一个该要素的可能性,除非上下文清楚地要求应当存在一个要素且仅存在一个所述要素。非限定冠词“a”或“an”因而通常意指“至少一个”。

[0038] 如本文中所用,术语“等位基因”意指在特定基因座位置的基因的一个或多个可变形式中的任意一个形式。在生物的二倍体(或双二倍体)细胞中,给定基因的等位基因位于染色体上的特定位置或基因座(复数基因座(loci))。成对同源染色体的每条染色体上各存在一个等位基因。

[0039] 如本文所用,术语“基因座”(复数形式loci)意指染色体上其中存在例如基因或遗传标记的一个特定位置或多个位置或一个位点。例如,“Rca 2A基因座”指在A基因组的染色体上可以找到Rca 2A基因(和两个Rca 2A等位基因)的位置,而“Rca 2B基因座”指在B基因组的染色体上可以找到Rca 2B基因(和两个Rca 2B等位基因)的位置,并且“Rca 2D基因座”指在D基因组的染色体上可以找到Rca 2D基因(和两个Rca 2D等位基因)的位置。

[0040] 如本文中所用,“野生型”(也书写为“野生型(wildtype)”或“野生型(wild-type)”)指植物或基因最常出现在自然界中的典型形式。“野生型植物”指一种植物,其具有在天然群体中此类植物的最常见表型。“野生型等位基因”指产生野生型蛋白和野生型表型所需要的基因的等位基因。相反,“突变体植物”指一种植物,该植物具有此类植物在天然群体中的不同罕有表型、或通过人类干预,例如通过诱变或基因编辑产生,并且“突变等位基因”指产生突变体蛋白和/或突变体表型所需要的基因的等位基因。

[0041] 如本文所用的“突变体”指植物或基因的形式,所述形式与天然群体中的这种植物或基因不同并且通过人类干预(例如通过诱变或基因编辑)产生,并且“突变等位基因”指一种等位基因,其在天然群体或繁育种群的植物中不存在,而是通过人类干预如诱变或基因编辑产生。

[0042] 如本文中所用,术语“野生型等位基因”(例如野生型Rca 2B等位基因、野生型Rca 2A等位基因或野生型Rca 2D等位基因)意指在禾谷类植物、尤其小麦植物内找到的天然存在性等位基因,其编码有功能的非耐热蛋白质(例如有功能的非耐热Rca 2A、Rca 2B或Rca 2D蛋白)。相反,如本文所用,术语“突变等位基因”(例如突变Rca 2A等位基因,突变Rca 2B等位基因或突变Rca 2D等位基因)指一种等位基因,其不编码有功能的非耐热蛋白质,即编

码无功能的Rca 2蛋白(例如无功能的Rca 2A或Rca 2B或Rca 2D)的Rca 2等位基因、或编码有功能的耐热蛋白质(例如有功能的耐热Rca 2A或Rca 2B或Rca 2D)的Rca 2等位基因。如本文所用,编码无功能的Rca 2蛋白的突变Rca 2等位基因指,与相应的野生型有功能Rca 2蛋白相比,没有生物学活性或生物学活性显著降低的Rca 2蛋白,或根本不编码Rca 2蛋白。如本文所用,编码有功能的耐热蛋白质的Rca 2等位基因指如下文描述的耐热Rca 2蛋白变体。敲除Rca 2等位基因是编码无功能Rca 2蛋白的突变Rca 2等位基因的等同术语。耐热Rca 2等位基因是编码有功能耐热蛋白质的Rca 2等位基因的等同术语。

[0043] 如本文中所用,“诱变”指这样的过程,其中对植物细胞(例如,多个禾谷类种子或其他部分,如花粉等)施与可以在所述细胞的DNA中引起突变的技术,如接触诱变剂如化学物质(如甲磺酸乙酯(EMS)、乙基亚硝基脲(ENU)等)或电离辐射(中子(如在快中子诱变法等中)、 α 射线、伽玛射线(如钴60源所提供)、X-射线、紫外辐射等),或T-DNA插入诱变(Azpiroz-Leehan等人(1997) Trends Genet 13:152-156)、转座子诱变(McKenzie等人(2002) Theor Appl Genet 105:23-33)或组织培养诱变(诱导体细胞克隆变异)、或这些技术中两者或更多者的组合。因此,可以通过使用上述方法之一完成一个或多个Rca 2等位基因的所需诱变。尽管通过辐射产生的突变经常是大的缺失或其他严重损伤如易位或复杂重排,然而通过化学诱变产生的突变经常是更离散的损伤如点突变。例如,EMS使鸟嘌呤碱基烷基化,这导致碱基错配:烷基化的鸟嘌呤将与胸腺嘧啶碱基配对,从而主要产生G/C至A/T转换。诱变后,使用已知技术从处理的细胞再生禾谷植物。例如,可以根据常规培育方法种植所得的谷物种子,并且在自我授粉后在该植物上形成种子。可以收获因这种自我授粉在当前或后续世代中形成的额外种子并对其筛选突变Rca 2等位基因的存在。已知几项技术可以筛选特定突变等位基因,例如,DeleteageneTM(Delete-a-gene;Li等人,2001,Plant J 27:235-242)使用聚合酶链反应(PCR)测定法来筛选通过快中子诱变法产生的缺失突变体,TILLING(基因组中定向诱导的局部损伤;McCallum等人,2000,Nat Biotechnol 18:455-457)鉴定EMS诱导的点突变等。

[0044] 如本文所用,基因编辑指使用序列特异性酶(如核酸内切酶、切刻酶、碱基转化酶)和/或供体核酸(例如dsDNA、寡聚物),定向修饰基因组DNA以在DNA中引入所需的改变。可以接受编程以识别特定DNA序列的序列特异性核酸酶包括大范围核酸酶(meganucleases)(MGN)、锌指核酸酶(ZFN)、TAL效应子核酸酶(TALEN)和RNA指导或DNA指导的核酸酶如Cas9,Cpf1,CasX,CasY,C2c1,C2c3、某些基于Argonaut的系统(参见,例如Osakabe和Osakabe,Plant Cell Physiol.2015Mar;56(3):389-400;Ma等人,Mol Plant.2016Jul 6;9(7):961-74;Bortessie等人,Plant Biotech J,2016,14;Murovec等人,Plant Biotechnol J.2017Apr 1;Nakade等人,Bioengineered 8-3,2017;Burststein等人,Nature 542,37-241;Komor等人,Nature 533,420-424,2016;所有文献均通过引用的方式并入本文)。供体核酸可以用作修复序列特异性核酸酶所致DNA断裂的模板,但也可以就此用于基因打靶(不引起DNA断裂)以将所需的变化引入基因组DNA。也可以在没有供体核酸的情况下使用序列特异性核酸酶,由此借助非同源末端接合修复机制允许插入突变或缺失突变。基因编辑可以用来创建突变Rca 2等位基因。

[0045] 突变核酸分子或突变等位基因可以包含一个或多个突变或修饰,如:

- a. “错义突变”,其是核酸序列中导致一个氨基酸置换另一个氨基酸的改变;

b. “无义突变”或“终止密码子突变”，其是在核酸序列中导致引入提前终止密码子和由此翻译终止(产生截短的蛋白质)的改变；植物基因含有翻译终止密码子“TGA”(RNA中是UGA)、“TAA”(RNA中是UAA)和“TAG”(RNA中是UAG)；因而导致这些密码子之一出现于正在翻译的成熟mRNA中(可读框中)的任何核苷酸置换、插入、缺失都将终止翻译；

c. 的一个或多个氨基酸的“插入突变”，其由在核酸的编码性序列中已经添加一个或多个密码子而引起；

d. 一个或多个氨基酸的“缺失突变”，其由在核酸的编码性序列中已经缺失一个或多个密码子而引起；

e. “移码突变”，其导致核酸序列在该突变的下游在不同的读框中翻译。移码突变可具有多种原因，如插入、缺失或重复一个或多个核苷酸；

f. 突变的剪接位点，其导致改变的剪接，从而导致改变的mRNA加工并因此导致编码的蛋白质改变，所述蛋白质可以含有各种长度的缺失、置换或插入，并可能与提前翻译终止组合。

Rca蛋白的热稳定性变体中的保守残基

[0046] 本文已经鉴定十一个在来自多个植物物种的Rca蛋白的热稳定性变体中保守的氨基酸残基(参见实施例3)。这些氨基酸是：(a)与SEQ ID NO:4的位置59或SEQ ID NO:8的位置109相对应的位置处的异亮氨酸，本文中称作“AA1”，(b)与SEQ ID NO:4的位置73或SEQ ID NO:8的位置123相对应的位置处的天冬氨酸，本文中称作“AA2”，(c)与SEQ ID NO:4的位置160或SEQ ID NO:8的位置210相对应的位置处的异亮氨酸，本文中称作“AA3”，(d)与SEQ ID NO:4的位置265或SEQ ID NO:8的位置315相对应的位置处的精氨酸，本文中称作“AA4”，(e)与SEQ ID NO:4的位置270或SEQ ID NO:8的位置320相对应的位置处的脯氨酸，本文中称作“AA5”，(f)与SEQ ID NO:4的位置277或SEQ ID NO:8的位置327相对应的位置处的亮氨酸，本文中称作“AA6”，(g)与SEQ ID NO:4的位置307或SEQ ID NO:8的位置357相对应的位置处的谷氨酸，本文中称作“AA7”，(h)与SEQ ID NO:4的位置334或SEQ ID NO:8的位置384相对应的位置处的异亮氨酸，本文中称作“AA8”，(i)与SEQ ID NO:4的位置359或SEQ ID NO:8的位置409相对应的位置处的赖氨酸，本文中称作“AA9”，(j)与SEQ ID NO:4的位置361或SEQ ID NO:8的位置411相对应的位置处的亮氨酸，本文中称作“AA10”，和(k)与SEQ ID NO:4的位置363或SEQ ID NO:8的位置413相对应的位置处的谷氨酸，本文中称作“AA11”。

[0047] 可以理解，本文描述的热稳定性蛋白质可以包含氨基酸序列，所述氨基酸序列与SEQ ID NOs:8、47、49、2、4、6、30、33、39、41、43或45的氨基酸序列具有90%同一性并且包含至少一个、至少两个、至少三个、至少四个、至少五个、至少六个、至少七个、至少八个、至少九个、至少十个或全部十一个在来自各植物物种的Rca蛋白的热稳定性变体中保守的氨基酸残基。

[0048] 所述至少两个氨基酸残基可以是(a) AA1和AA2、(b) AA1和AA3、(c) AA1和AA4、(d) AA1和AA5、(e) AA1和AA6、(f) AA1和AA7、(g) AA1和AA8、(h) AA1和AA9、(i) AA1和AA10、或(j) AA1和AA11。所述至少两个氨基酸残基还可以是(a) AA2和AA3、(b) AA2和AA4、(c) AA2和AA5、(d) AA2和AA6、(e) AA2和AA7、(f) AA2和AA8、(g) AA2和AA9、(h) AA2和AA10或(i) AA2和AA11。所述至少两个氨基酸残基还可以是(a) AA3和AA4、(b) AA3和AA5、(c) AA3和AA6、(d) AA3和AA7、(e) AA3和AA8、(f) AA3和AA9、(g) AA3和AA10或(h) AA3和AA11。所述至少两个氨基酸残基还可

以是 (a) AA4和AA5、(b) AA4和AA6、(c) AA4和AA7、(d) AA4和AA8、(e) AA4和AA9、(f) AA4和AA10或 (g) AA4和AA11。作为一种备选,至少两个氨基酸残基可以是 (a) AA5和AA6、(b) AA5和AA7、(c) AA5和AA8、(d) AA5和AA9、(e) AA5和AA10或 (f) AA5和AA11。作为另一种备选,至少两个氨基酸残基可以是 (a) AA6和AA7、(b) AA6和AA8、(c) AA6和AA9、(d) AA6和AA10或 (e) AA6和AA11。作为又一种备选,至少两个氨基酸残基可以是 (a) AA7和AA8、(b) AA7和AA9、(c) AA7和AA10或 (d) AA7和AA11。所述至少两个氨基酸残基还可以是 (a) AA8和AA9、(b) AA8和AA10或 (c) AA8和AA11。所述至少两个氨基酸残基还进一步可以是 (a) AA9和AA10或 (b) AA9和AA11。所述至少两个氨基酸残基还可以是AA10和AA11。

[0049] 所述至少三个氨基酸残基可以是 (a) AA1、AA2和AA3、(b) AA1、AA2和AA4、(c) AA1、AA2和AA5、(d) AA1、AA2和AA6、(e) AA1、AA2和AA7、(f) AA1、AA2和AA8、(g) AA1、AA2和AA9、(h) AA1、AA2和AA10或 (i) AA1、AA2和AA11。所述至少三个氨基酸残基还可以是 (a) AA1、AA3和AA4、(b) AA1、AA3和AA5、(c) AA1、AA3和AA6、(d) AA1、AA3和AA7、(e) AA1、AA3和AA8、(f) AA1、AA3和AA9、(g) AA1、AA3和AA10或 (h) AA1、AA3和AA11。所述至少三个氨基酸残基还可以是 (a) AA1、AA4和AA5、(b) AA1、AA4和AA6、(c) AA1、AA4和AA7、(d) AA1、AA4和AA8、(e) AA1、AA4和AA9、(f) AA1、AA4和AA10或 (g) AA1、AA4和AA11。作为一种备选,所述至少三个氨基酸残基可以是 (a) AA1、AA5和AA6、(b) AA1、AA5和AA7、(c) AA1、AA5和AA8、(d) AA1、AA5和AA9、(e) AA1、AA5和AA10或 (f) AA1、AA5和AA11。作为另一种备选,所述至少三个氨基酸残基可以是 (a) AA1、AA6和AA7、(b) AA1、AA6和AA8、(c) AA1、AA6和AA9、(d) AA1、AA6和AA10或 (e) AA1、AA6和AA11。作为又一种备选,所述至少三个氨基酸残基可以是 (a) AA1、AA7和AA8、(b) AA1、AA7和AA9、(c) AA1、AA7和AA10或 (d) AA1、AA7和AA11。所述至少三个氨基酸残基还可以是 (a) AA1、AA8和AA9、(b) AA1、AA8和AA10或 (c) AA1、AA8和AA11。所述至少三个氨基酸残基还进一步可以是 (a) AA1、AA9和AA10或 (b) AA1、AA9和AA11。所述至少三个氨基酸残基还可以是AA1、AA10和AA11。

[0050] 所述至少三个氨基酸残基还可以是 (a) AA2、AA3和AA4、(b) AA2、AA3和AA5、(c) AA2、AA3和AA6、(d) AA2、AA3和AA7、(e) AA2、AA3和AA8、(f) AA2、AA3和AA9、(g) AA2、AA3和AA10或 (h) AA2、AA3和AA11。所述至少三个氨基酸残基还可以是 (a) AA2、AA4和AA5、(b) AA2、AA4和AA6、(c) AA2、AA4和AA7、(d) AA2、AA4和AA8、(e) AA2、AA4和AA9、(f) AA2、AA4和AA10或 (g) AA2、AA4和AA11。作为一种备选,所述至少三个氨基酸残基可以是 (a) AA2、AA5和AA6、(b) AA2、AA5和AA7、(c) AA2、AA5和AA8、(d) AA2、AA5和AA9、(e) AA2、AA5和AA10或 (f) AA2、AA5和AA11。作为另一种备选,所述至少三个氨基酸残基可以是 (a) AA2、AA6和AA7、(b) AA2、AA6和AA8、(c) AA2、AA6和AA9、(d) AA2、AA6和AA10或 (e) AA2、AA6和AA11。作为又一种备选,所述至少三个氨基酸残基可以是 (a) AA2、AA7和AA8、(b) AA2、AA7和AA9、(c) AA2、AA7和AA10或 (d) AA2、AA7和AA11。所述至少三个氨基酸残基还可以是 (a) AA2、AA8和AA9、(b) AA2、AA8和AA10或 (c) AA2、AA8和AA11。所述至少三个氨基酸残基还进一步可以是 (a) AA2、AA9和AA10或 (b) AA2、AA9和AA11。所述至少三个氨基酸残基还可以是AA2、AA10和AA11。

[0051] 所述至少三个氨基酸残基可以是 (a) AA3、AA4和AA5、(b) AA3、AA4和AA6、(c) AA3、AA4和AA7、(d) AA3、AA4和AA8、(e) AA3、AA4和AA9、(f) AA3、AA4和AA10或 (g) AA3、AA4和AA11。作为一种备选,所述至少三个氨基酸残基可以是 (a) AA3、AA5和AA6、(b) AA3、AA5和AA7、(c) AA3、AA5和AA8、(d) AA3、AA5和AA9、(e) AA3、AA5和AA10或 (f) AA3、AA5和AA11。作为另一种备

选,所述至少三个氨基酸残基可以是(a) AA3、AA6和AA7、(b) AA3、AA6和AA8、(c) AA3、AA6和AA9、(d) AA3、AA6和AA10或(e) AA3、AA6和AA11。作为又一种备选,所述至少三个氨基酸残基可以是(a) AA3、AA7和AA8、(b) AA3、AA7和AA9、(c) AA3、AA7和AA10或(d) AA3、AA7和AA11。所述至少三个氨基酸残基还可以是(a) AA3、AA8和AA9、(b) AA3、AA8和AA10或(c) AA3、AA8和AA11。所述至少三个氨基酸残基还进一步可以是(a) AA3、AA9和AA10或(b) AA3、AA9和AA11。所述至少三个氨基酸残基还可以是AA3、AA10和AA11。

[0052] 所述至少三个氨基酸残基可以是(a) AA4、AA5和AA6、(b) AA4、AA5和AA7、(c) AA4、AA5和AA8、(d) AA4、AA5和AA9、(e) AA4、AA5和AA10或(f) AA4、AA5和AA11。作为另一种备选,所述至少三个氨基酸残基可以是(a) AA4、AA6和AA7、(b) AA4、AA6和AA8、(c) AA4、AA6和AA9、(d) AA4、AA6和AA10或(e) AA4、AA6和AA11。作为又一种备选,所述至少三个氨基酸残基可以是(a) AA4、AA7和AA8、(b) AA4、AA7和AA9、(c) AA4、AA7和AA10或(d) AA4、AA7和AA11。所述至少三个氨基酸残基还可以是(a) AA4、AA8和AA9、(b) AA4、AA8和AA10或(c) AA4、AA8和AA11。所述至少三个氨基酸残基还进一步可以是(a) AA4、AA9和AA10或(b) AA4、AA9和AA11。所述至少三个氨基酸残基还可以是AA4、AA10和AA11。

[0053] 所述至少三个氨基酸残基可以是(a) AA5、AA6和AA7、(b) AA5、AA6和AA8、(c) AA5、AA6和AA9、(d) AA5、AA6和AA10或(e) AA5、AA6和AA11。作为又一种备选,所述至少三个氨基酸残基可以是(a) AA5、AA7和AA8、(b) AA5、AA7和AA9、(c) AA5、AA7和AA10或(d) AA5、AA7和AA11。所述至少三个氨基酸残基还可以是(a) AA5、AA8和AA9、(b) AA5、AA8和AA10或(c) AA5、AA8和AA11。所述至少三个氨基酸残基还进一步可以是(a) AA5、AA9和AA10或(b) AA5、AA9和AA11。所述至少三个氨基酸残基还可以是AA5、AA10和AA11。

[0054] 所述至少三个氨基酸残基可以是(a) AA6、AA7和AA8、(b) AA6、AA7和AA9、(c) AA6、AA7和AA10或(d) AA6、AA7和AA11。所述至少三个氨基酸残基还可以是(a) AA6、AA8和AA9、(b) AA6、AA8和AA10或(c) AA6、AA8和AA11。所述至少三个氨基酸残基还进一步可以是(a) AA6、AA9和AA10或(b) AA6、AA9和AA11。所述至少三个氨基酸残基还可以是AA6、AA10和AA11。所述至少三个氨基酸残基可以是(a) AA7、AA8和AA9、(b) AA7、AA8和AA10或(c) AA7、AA8和AA11。所述至少三个氨基酸残基还进一步可以是(a) AA7、AA9和AA10或(b) AA7、AA9和AA11。所述至少三个氨基酸残基还可以是AA7、AA10和AA11。所述至少三个氨基酸残基还可以是(a) AA8、AA9和AA10或(b) AA8、AA9和AA11。所述至少三个氨基酸残基还可以是AA8、AA10和AA11。所述至少三个氨基酸残基还可以是AA9、AA10和AA11。

[0055] 此类所述至少四个氨基酸残基可以是(a) AA1、AA2、AA3和AA4、(b) AA1、AA2、AA3和AA5、(c) AA1、AA2、AA3和AA6、(d) AA1、AA2、AA3和AA7、(e) AA1、AA2、AA3和AA8、(f) AA1、AA2、AA3和AA9、(g) AA1、AA2、AA3和AA10或(h) AA1、AA2、AA3和AA11。所述至少四个氨基酸残基还可以是(a) AA1、AA2、AA4和AA5、(b) AA1、AA2、AA4和AA6、(c) AA1、AA2、AA4和AA7、(d) AA1、AA2、AA4和AA8、(e) AA1、AA2、AA4和AA9、(f) AA1、AA2、AA4和AA10或(g) AA1、AA2、AA4和AA11。作为一种备选,所述至少四个氨基酸残基可以是(a) AA1、AA2、AA5和AA6、(b) AA1、AA2、AA5和AA7、(c) AA1、AA2、AA5和AA8、(d) AA1、AA2、AA5和AA9、(e) AA1、AA2、AA5和AA10或(f) AA1、AA2、AA5和AA11。作为另一种备选,所述至少四个氨基酸残基可以是(a) AA1、AA2、AA6和AA7、(b) AA1、AA2、AA6和AA8、(c) AA1、AA2、AA6和AA9、(d) AA1、AA2、AA6和AA10或(e) AA1、AA2、AA6和AA11。作为又一种备选,所述至少四个氨基酸残基可以是(a) AA1、AA2、AA7和AA8、(b) AA1、AA2、AA7

和AA9、(c) AA1、AA2、AA7和AA10或(d) AA1、AA2、AA7和AA11。所述至少四个氨基酸残基还可以是(a) AA1、AA2、AA8和AA9、(b) AA1、AA2、AA8和AA10或(c) AA1、AA2、AA8和AA11。所述至少四个氨基酸残基还进一步可以是(a) AA1、AA2、AA9和AA10或(b) AA1、AA2、AA9和AA11。所述至少四个氨基酸残基还可以是AA1、AA2、AA10和AA11。此类所述至少四个氨基酸残基可以是(a) AA1、AA3、AA4和AA5、(b) AA1、AA3、AA4和AA6、(c) AA1、AA3、AA4和AA7、(d) AA1、AA3、AA4和AA8、(e) AA1、AA3、AA4和AA9、(f) AA1、AA3、AA4和AA10或(g) AA1、AA3、AA4和AA11。作为一种备选,所述至少四个氨基酸残基可以是(a) AA1、AA3、AA5和AA6、(b) AA1、AA3、AA5和AA7、(c) AA1、AA3、AA5和AA8、(d) AA1、AA3、AA5和AA9、(e) AA1、AA3、AA5和AA10或(f) AA1、AA3、AA5和AA11。作为另一种备选,所述至少四个氨基酸残基可以是(a) AA1、AA3、AA6和AA7、(b) AA1、AA3、AA6和AA8、(c) AA1、AA3、AA6和AA9、(d) AA1、AA3、AA6和AA10或(e) AA1、AA3、AA6和AA11。作为又一种备选,所述至少四个氨基酸残基可以是(a) AA1、AA3、AA7和AA8、(b) AA1、AA3、AA7和AA9、(c) AA1、AA3、AA7和AA10或(d) AA1、AA3、AA7和AA11。所述至少四个氨基酸残基还可以是(a) AA1、AA3、AA8和AA9、(b) AA1、AA3、AA8和AA10或(c) AA1、AA3、AA8和AA11。所述至少四个氨基酸残基还进一步可以是(a) AA1、AA3、AA9和AA10或(b) AA1、AA3、AA9和AA11。所述至少四个氨基酸残基还可以是AA1、AA3、AA10和AA11。此类所述至少四个氨基酸残基可以是(a) AA1、AA4、AA5和AA6、(b) AA1、AA4、AA5和AA7、(c) AA1、AA4、AA5和AA8、(d) AA1、AA4、AA5和AA9、(e) AA1、AA4、AA5和AA10或(f) AA1、AA4、AA5和AA11。作为另一种备选,所述至少四个氨基酸残基可以是(a) AA1、AA4、AA6和AA7、(b) AA1、AA4、AA6和AA8、(c) AA1、AA4、AA6和AA9、(d) AA1、AA4、AA6和AA10或(e) AA1、AA4、AA6和AA11。作为又一种备选,所述至少四个氨基酸残基可以是(a) AA1、AA4、AA7和AA8、(b) AA1、AA4、AA7和AA9、(c) AA1、AA4、AA7和AA10或(d) AA1、AA4、AA7和AA11。所述至少四个氨基酸残基还可以是(a) AA1、AA4、AA8和AA9、(b) AA1、AA4、AA8和AA10或(c) AA1、AA4、AA8和AA11。所述至少四个氨基酸残基还进一步可以是(a) AA1、AA4、AA9和AA10或(b) AA1、AA4、AA9和AA11。所述至少四个氨基酸残基还可以是AA1、AA4、AA10和AA11。此类所述至少四个氨基酸残基可以是(a) AA1、AA5、AA6和AA7、(b) AA1、AA5、AA6和AA8、(c) AA1、AA5、AA6和AA9、(d) AA1、AA5、AA6和AA10或(e) AA1、AA5、AA6和AA11。作为又一种备选,所述至少四个氨基酸残基可以是(a) AA1、AA5、AA7和AA8、(b) AA1、AA5、AA7和AA9、(c) AA1、AA5、AA7和AA10或(d) AA1、AA5、AA7和AA11。所述至少四个氨基酸残基还可以是(a) AA1、AA5、AA8和AA9、(b) AA1、AA5、AA8和AA10或(c) AA1、AA5、AA8和AA11。所述至少四个氨基酸残基还进一步可以是(a) AA1、AA5、AA9和AA10或(b) AA1、AA5、AA9和AA11。所述至少四个氨基酸残基还可以是AA1、AA5、AA10和AA11。此类所述至少四个氨基酸残基可以是(a) AA1、AA6、AA7和AA8、(b) AA1、AA6、AA7和AA9、(c) AA1、AA6、AA7和AA10或(d) AA1、AA6、AA7和AA11。所述至少四个氨基酸残基还可以是(a) AA1、AA6、AA8和AA9、(b) AA1、AA6、AA8和AA10或(c) AA1、AA6、AA8和AA11。所述至少四个氨基酸残基还进一步可以是(a) AA1、AA6、AA9和AA10或(b) AA1、AA6、AA9和AA11。所述至少四个氨基酸残基还可以是AA1、AA6、AA10和AA11。此类所述至少四个氨基酸残基可以是(a) AA1、AA7、AA8和AA9、(b) AA1、AA7、AA8和AA10或(c) AA1、AA7、AA8和AA11。所述至少四个氨基酸残基还进一步可以是(a) AA1、AA7、AA9和AA10或(b) AA1、AA7、AA9和AA11。所述至少四个氨基酸残基还可以是AA1、AA7、AA10和AA11。此类所述至少四个氨基酸残基可以是(a) AA1、AA8、AA9和AA10或(b) AA1、AA8、AA9和AA11。所述至少四个氨基酸残基还可以是AA1、AA8、AA10和AA11。所述至少四个氨基酸残基还可以是AA1、AA9、AA10和AA11。

[0056] 此类所述至少四个氨基酸残基可以是 (a) AA2、AA3、AA4和AA5、(b) AA2、AA3、AA4和AA6、(c) AA2、AA3、AA4和AA7、(d) AA2、AA3、AA4和AA8、(e) AA2、AA3、AA4和AA9、(f) AA2、AA3、AA4和AA10或 (g) AA2、AA3、AA4和AA11。作为一种备选,所述至少四个氨基酸残基可以是 (a) AA2、AA3、AA5和AA6、(b) AA2、AA3、AA5和AA7、(c) AA2、AA3、AA5和AA8、(d) AA2、AA3、AA5和AA9、(e) AA2、AA3、AA5和AA10或 (f) AA2、AA3、AA5和AA11。作为另一种备选,所述至少四个氨基酸残基可以是 (a) AA2、AA3、AA6和AA7、(b) AA2、AA3、AA6和AA8、(c) AA2、AA3、AA6和AA9、(d) AA2、AA3、AA6和AA10或 (e) AA2、AA3、AA6和AA11。作为又一种备选,所述至少四个氨基酸残基可以是 (a) AA2、AA3、AA7和AA8、(b) AA2、AA3、AA7和AA9、(c) AA2、AA3、AA7和AA10或 (d) AA2、AA3、AA7和AA11。所述至少四个氨基酸残基还可以是 (a) AA2、AA3、AA8和AA9、(b) AA2、AA3、AA8和AA10或 (c) AA2、AA3、AA8和AA11。所述至少四个氨基酸残基还进一步可以是 (a) AA2、AA3、AA9和AA10或 (b) AA2、AA3、AA9和AA11。所述至少四个氨基酸残基还可以是AA2、AA3、AA10和AA11。此类所述至少四个氨基酸残基可以是 (a) AA2、AA4、AA5和AA6、(b) AA2、AA4、AA5和AA7、(c) AA2、AA4、AA5和AA8、(d) AA2、AA4、AA5和AA9、(e) AA2、AA4、AA5和AA10或 (f) AA2、AA4、AA5和AA11。作为另一种备选,所述至少四个氨基酸残基可以是 (a) AA2、AA4、AA6和AA7、(b) AA2、AA4、AA6和AA8、(c) AA2、AA4、AA6和AA9、(d) AA2、AA4、AA6和AA10或 (e) AA2、AA4、AA6和AA11。作为又一种备选,所述至少四个氨基酸残基可以是 (a) AA2、AA4、AA7和AA8、(b) AA2、AA4、AA7和AA9、(c) AA2、AA4、AA7和AA10或 (d) AA2、AA4、AA7和AA11。所述至少四个氨基酸残基还可以是 (a) AA2、AA4、AA8和AA9、(b) AA2、AA4、AA8和AA10或 (c) AA2、AA4、AA8和AA11。所述至少四个氨基酸残基还进一步可以是 (a) AA2、AA4、AA9和AA10或 (b) AA2、AA4、AA9和AA11。所述至少四个氨基酸残基还可以是AA2、AA4、AA10和AA11。此类所述至少四个氨基酸残基可以是 (a) AA2、AA5、AA6和AA7、(b) AA2、AA5、AA6和AA8、(c) AA2、AA5、AA6和AA9、(d) AA2、AA5、AA6和AA10或 (e) AA2、AA5、AA6和AA11。作为又一种备选,所述至少四个氨基酸残基可以是 (a) AA2、AA5、AA7和AA8、(b) AA2、AA5、AA7和AA9、(c) AA2、AA5、AA7和AA10或 (d) AA2、AA5、AA7和AA11。所述至少四个氨基酸残基还可以是 (a) AA2、AA5、AA8和AA9、(b) AA2、AA5、AA8和AA10或 (c) AA2、AA5、AA8和AA11。所述至少四个氨基酸残基还进一步可以是 (a) AA2、AA5、AA9和AA10或 (b) AA2、AA5、AA9和AA11。所述至少四个氨基酸残基还可以是AA2、AA5、AA10和AA11。此类所述至少四个氨基酸残基可以是 (a) AA2、AA6、AA7和AA8、(b) AA2、AA6、AA7和AA9、(c) AA2、AA6、AA7和AA10或 (d) AA2、AA6、AA7和AA11。所述至少四个氨基酸残基还可以是 (a) AA2、AA6、AA8和AA9、(b) AA2、AA6、AA8和AA10或 (c) AA2、AA6、AA8和AA11。所述至少四个氨基酸残基还进一步可以是 (a) AA2、AA6、AA9和AA10或 (b) AA2、AA6、AA9和AA11。所述至少四个氨基酸残基还可以是AA2、AA6、AA10和AA11。此类所述至少四个氨基酸残基可以是 (a) AA2、AA7、AA8和AA9、(b) AA2、AA7、AA8和AA10或 (c) AA2、AA7、AA8和AA11。所述至少四个氨基酸残基还进一步可以是 (a) AA2、AA7、AA9和AA10或 (b) AA2、AA7、AA9和AA11。所述至少四个氨基酸残基还可以是AA2、AA7、AA10和AA11。此类所述至少四个氨基酸残基可以是 (a) AA2、AA8、AA9和AA10或 (b) AA2、AA8、AA9和AA11。所述至少四个氨基酸残基还可以是AA2、AA8、AA10和AA11。所述至少四个氨基酸残基还可以是AA2、AA9、AA10和AA11。

[0057] 此类所述至少四个氨基酸残基可以是 (a) AA3、AA4、AA5和AA6、(b) AA3、AA4、AA5和AA7、(c) AA3、AA4、AA5和AA8、(d) AA3、AA4、AA5和AA9、(e) AA3、AA4、AA5和AA10或 (f) AA3、AA4、AA5和AA11。作为另一种备选,所述至少四个氨基酸残基可以是 (a) AA3、AA4、AA6和AA7、(b)

AA3、AA4、AA6和AA8、(c) AA3、AA4、AA6和AA9、(d) AA3、AA4、AA6和AA10或(e) AA3、AA4、AA6和AA11。作为又一种备选,所述至少四个氨基酸残基可以是(a) AA3、AA4、AA7和AA8、(b) AA3、AA4、AA7和AA9、(c) AA3、AA4、AA7和AA10或(d) AA3、AA4、AA7和AA11。所述至少四个氨基酸残基还可以是(a) AA3、AA4、AA8和AA9、(b) AA3、AA4、AA8和AA10或(c) AA3、AA4、AA8和AA11。所述至少四个氨基酸残基还进一步可以是(a) AA3、AA4、AA9和AA10或(b) AA3、AA4、AA9和AA11。所述至少四个氨基酸残基还可以是AA3、AA4、AA10和AA11。此类所述至少四个氨基酸残基可以是(a) AA3、AA5、AA6和AA7、(b) AA3、AA5、AA6和AA8、(c) AA3、AA5、AA6和AA9、(d) AA3、AA5、AA6和AA10或(e) AA3、AA5、AA6和AA11。作为又一种备选,所述至少四个氨基酸残基可以是(a) AA3、AA5、AA7和AA8、(b) AA3、AA5、AA7和AA9、(c) AA3、AA5、AA7和AA10或(d) AA3、AA5、AA7和AA11。所述至少四个氨基酸残基还可以是(a) AA3、AA5、AA8和AA9、(b) AA3、AA5、AA8和AA10或(c) AA3、AA5、AA8和AA11。所述至少四个氨基酸残基还进一步可以是(a) AA3、AA5、AA9和AA10或(b) AA3、AA5、AA9和AA11。所述至少四个氨基酸残基还可以是AA3、AA5、AA10和AA11。此类所述至少四个氨基酸残基可以是(a) AA3、AA6、AA7和AA8、(b) AA3、AA6、AA7和AA9、(c) AA3、AA6、AA7和AA10或(d) AA3、AA6、AA7和AA11。所述至少四个氨基酸残基还可以是(a) AA3、AA6、AA8和AA9、(b) AA3、AA6、AA8和AA10或(c) AA3、AA6、AA8和AA11。所述至少四个氨基酸残基还进一步可以是(a) AA3、AA6、AA9和AA10或(b) AA3、AA6、AA9和AA11。所述至少四个氨基酸残基还可以是AA3、AA6、AA10和AA11。此类所述至少四个氨基酸残基可以是(a) AA3、AA7、AA8和AA9、(b) AA3、AA7、AA8和AA10或(c) AA3、AA7、AA8和AA11。所述至少四个氨基酸残基还进一步可以是(a) AA3、AA7、AA9和AA10或(b) AA3、AA7、AA9和AA11。所述至少四个氨基酸残基还可以是AA3、AA7、AA10和AA11。此类所述至少四个氨基酸残基可以是(a) AA3、AA8、AA9和AA10或(b) AA3、AA8、AA9和AA11。所述至少四个氨基酸残基还可以是AA3、AA8、AA10和AA11。所述至少四个氨基酸残基还可以是AA3、AA9、AA10和AA11。

[0058] 此类所述至少四个氨基酸残基可以是(a) AA4、AA5、AA6和AA7、(b) AA4、AA5、AA6和AA8、(c) AA4、AA5、AA6和AA9、(d) AA4、AA5、AA6和AA10或(e) AA4、AA5、AA6和AA11。作为又一种备选,所述至少四个氨基酸残基可以是(a) AA4、AA5、AA7和AA8、(b) AA4、AA5、AA7和AA9、(c) AA4、AA5、AA7和AA10或(d) AA4、AA5、AA7和AA11。所述至少四个氨基酸残基还可以是(a) AA4、AA5、AA8和AA9、(b) AA4、AA5、AA8和AA10或(c) AA4、AA5、AA8和AA11。所述至少四个氨基酸残基还进一步可以是(a) AA4、AA5、AA9和AA10或(b) AA4、AA5、AA9和AA11。所述至少四个氨基酸残基还可以是AA4、AA5、AA10和AA11。此类所述至少四个氨基酸残基可以是(a) AA4、AA6、AA7和AA8、(b) AA4、AA6、AA7和AA9、(c) AA4、AA6、AA7和AA10或(d) AA4、AA6、AA7和AA11。所述至少四个氨基酸残基还可以是(a) AA4、AA6、AA8和AA9、(b) AA4、AA6、AA8和AA10或(c) AA4、AA6、AA8和AA11。所述至少四个氨基酸残基还进一步可以是(a) AA4、AA6、AA9和AA10或(b) AA4、AA6、AA9和AA11。所述至少四个氨基酸残基还可以是AA4、AA6、AA10和AA11。此类所述至少四个氨基酸残基可以是(a) AA4、AA7、AA8和AA9、(b) AA4、AA7、AA8和AA10或(c) AA4、AA7、AA8和AA11。所述至少四个氨基酸残基还进一步可以是(a) AA4、AA7、AA9和AA10或(b) AA4、AA7、AA9和AA11。所述至少四个氨基酸残基还可以是AA4、AA7、AA10和AA11。此类所述至少四个氨基酸残基可以是(a) AA4、AA8、AA9和AA10或(b) AA4、AA8、AA9和AA11。所述至少四个氨基酸残基还可以是AA4、AA8、AA10和AA11。所述至少四个氨基酸残基还可以是AA4、AA9、AA10和AA11。

[0059] 此类所述至少四个氨基酸残基可以是(a) AA5、AA6、AA7和AA8、(b) AA5、AA6、AA7和

AA9、(c) AA5、AA6、AA7和AA10或(d) AA5、AA6、AA7和AA11。所述至少四个氨基酸残基还可以是(a) AA5、AA6、AA8和AA9、(b) AA5、AA6、AA8和AA10或(c) AA5、AA6、AA8和AA11。所述至少四个氨基酸残基还进一步可以是(a) AA5、AA6、AA9和AA10或(b) AA5、AA6、AA9和AA11。所述至少四个氨基酸残基还可以是AA5、AA6、AA10和AA11。此类所述至少四个氨基酸残基可以是(a) AA5、AA7、AA8和AA9、(b) AA5、AA7、AA8和AA10或(c) AA5、AA7、AA8和AA11。所述至少四个氨基酸残基还进一步可以是(a) AA5、AA7、AA9和AA10或(b) AA5、AA7、AA9和AA11。所述至少四个氨基酸残基还可以是AA5、AA7、AA10和AA11。此类所述至少四个氨基酸残基可以是(a) AA5、AA8、AA9和AA10或(b) AA5、AA8、AA9和AA11。所述至少四个氨基酸残基还可以是AA5、AA8、AA10和AA11。所述至少四个氨基酸残基还可以是AA5、AA9、AA10和AA11。

[0060] 此类所述至少四个氨基酸残基可以是(a) AA6、AA7、AA8和AA9、(b) AA6、AA7、AA8和AA10或(c) AA6、AA7、AA8和AA11。所述至少四个氨基酸残基还进一步可以是(a) AA6、AA7、AA9和AA10或(b) AA6、AA7、AA9和AA11。所述至少四个氨基酸残基还可以是AA6、AA7、AA10和AA11。此类所述至少四个氨基酸残基可以是(a) AA6、AA8、AA9和AA10或(b) AA6、AA8、AA9和AA11。所述至少四个氨基酸残基还可以是AA6、AA8、AA10和AA11。所述至少四个氨基酸残基还可以是AA6、AA9、AA10和AA11。此类所述至少四个氨基酸残基可以是(a) AA7、AA8、AA9和AA10或(b) AA7、AA8、AA9和AA11。所述至少四个氨基酸残基还可以是AA7、AA8、AA10和AA11。所述至少四个氨基酸残基还可以是AA7、AA9、AA10和AA11。所述至少四个氨基酸残基还可以是AA8、AA9、AA10和AA11。

[0061] 此类所述至少五个氨基酸残基可以是(a) AA1、AA2、AA3、AA4和AA5、(b) AA1、AA2、AA3、AA4和AA6、(c) AA1、AA2、AA3、AA4和AA7、(d) AA1、AA2、AA3、AA4和AA8、(e) AA1、AA2、AA3、AA4和AA9、(f) AA1、AA2、AA3、AA4和AA10或(g) AA1、AA2、AA3、AA4和AA11。作为一种备选,所述至少五个氨基酸残基可以是(a) AA1、AA2、AA3、AA5和AA6、(b) AA1、AA2、AA3、AA5和AA7、(c) AA1、AA2、AA3、AA5和AA8、(d) AA1、AA2、AA3、AA5和AA9、(e) AA1、AA2、AA3、AA5和AA10或(f) AA1、AA2、AA3、AA5和AA11。作为另一种备选,所述至少五个氨基酸残基可以是(a) AA1、AA2、AA3、AA6和AA7、(b) AA1、AA2、AA3、AA6和AA8、(c) AA1、AA2、AA3、AA6和AA9、(d) AA1、AA2、AA3、AA6和AA10或(e) AA1、AA2、AA3、AA6和AA11。作为又一种备选,所述至少五个氨基酸残基可以是(a) AA1、AA2、AA3、AA7和AA8、(b) AA1、AA2、AA3、AA7和AA9、(c) AA1、AA2、AA3、AA7和AA10或(d) AA1、AA2、AA3、AA7和AA11。所述至少五个氨基酸残基还可以是(a) AA1、AA2、AA3、AA8和AA9、(b) AA1、AA2、AA3、AA8和AA10或(c) AA1、AA2、AA3、AA8和AA11。所述至少五个氨基酸残基还进一步可以是(a) AA1、AA2、AA3、AA9和AA10或(b) AA1、AA2、AA3、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA2、AA3、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA1、AA2、AA4、AA5和AA6、(b) AA1、AA2、AA4、AA5和AA7、(c) AA1、AA2、AA4、AA5和AA8、(d) AA1、AA2、AA4、AA5和AA9、(e) AA1、AA2、AA4、AA5和AA10或(f) AA1、AA2、AA4、AA5和AA11。作为另一种备选,所述至少五个氨基酸残基可以是(a) AA1、AA2、AA4、AA6和AA7、(b) AA1、AA2、AA4、AA6和AA8、(c) AA1、AA2、AA4、AA6和AA9、(d) AA1、AA2、AA4、AA6和AA10或(e) AA1、AA2、AA4、AA6和AA11。作为又一种备选,所述至少五个氨基酸残基可以是(a) AA1、AA2、AA4、AA7和AA8、(b) AA1、AA2、AA4、AA7和AA9、(c) AA1、AA2、AA4、AA7和AA10或(d) AA1、AA2、AA4、AA7和AA11。所述至少五个氨基酸残基还可以是(a) AA1、AA2、AA4、AA8和AA9、(b) AA1、AA2、AA4、AA8和AA10或(c) AA1、AA2、AA4、AA8和AA11。所述至少五个氨基酸残基还进一步可以是(a) AA1、AA2、AA4、

AA9和AA10或(b) AA1、AA2、AA4、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA2、AA4、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA1、AA2、AA5、AA6和AA7、(b) AA1、AA2、AA5、AA6和AA8、(c) AA1、AA2、AA5、AA6和AA9、(d) AA1、AA2、AA5、AA6和AA10或(e) AA1、AA2、AA5、AA6和AA11。作为又一种备选,所述至少五个氨基酸残基可以是(a) AA1、AA2、AA5、AA7和AA8、(b) AA1、AA2、AA5、AA7和AA9、(c) AA1、AA2、AA5、AA7和AA10或(d) AA1、AA2、AA5、AA7和AA11。所述至少五个氨基酸残基还可以是(a) AA1、AA2、AA5、AA8和AA9、(b) AA1、AA2、AA5、AA8和AA10或(c) AA1、AA2、AA5、AA8和AA11。所述至少五个氨基酸残基还可以进一步可以是(a) AA1、AA2、AA5、AA9和AA10或(b) AA1、AA2、AA5、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA2、AA5、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA1、AA2、AA6、AA7和AA8、(b) AA1、AA2、AA6、AA7和AA9、(c) AA1、AA2、AA6、AA7和AA10或(d) AA1、AA2、AA6、AA7和AA11。所述至少五个氨基酸残基还可以是(a) AA1、AA2、AA6、AA8和AA9、(b) AA1、AA2、AA6、AA8和AA10或(c) AA1、AA2、AA6、AA8和AA11。所述至少五个氨基酸残基还可以进一步可以是(a) AA1、AA2、AA6、AA9和AA10或(b) AA1、AA2、AA6、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA2、AA6、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA1、AA2、AA7、AA8和AA9、(b) AA1、AA2、AA7、AA8和AA10或(c) AA1、AA2、AA7、AA8和AA11。所述至少五个氨基酸残基还可以进一步可以是(a) AA1、AA2、AA7、AA9和AA10或(b) AA1、AA2、AA7、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA2、AA7、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA1、AA2、AA8、AA9和AA10或(b) AA1、AA2、AA8、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA2、AA8、AA10和AA11。所述至少五个氨基酸残基还可以是AA1、AA2、AA9、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA1、AA3、AA4、AA5和AA6、(b) AA1、AA3、AA4、AA5和AA7、(c) AA1、AA3、AA4、AA5和AA8、(d) AA1、AA3、AA4、AA5和AA9、(e) AA1、AA3、AA4、AA5和AA10或(f) AA1、AA3、AA4、AA5和AA11。作为另一种备选,所述至少五个氨基酸残基可以是(a) AA1、AA3、AA4、AA6和AA7、(b) AA1、AA3、AA4、AA6和AA8、(c) AA1、AA3、AA4、AA6和AA9、(d) AA1、AA3、AA4、AA6和AA10或(e) AA1、AA3、AA4、AA6和AA11。作为又一种备选,所述至少五个氨基酸残基可以是(a) AA1、AA3、AA4、AA7和AA8、(b) AA1、AA3、AA4、AA7和AA9、(c) AA1、AA3、AA4、AA7和AA10或(d) AA1、AA3、AA4、AA7和AA11。所述至少五个氨基酸残基还可以是(a) AA1、AA3、AA4、AA8和AA9、(b) AA1、AA3、AA4、AA8和AA10或(c) AA1、AA3、AA4、AA8和AA11。所述至少五个氨基酸残基还可以进一步可以是(a) AA1、AA3、AA4、AA9和AA10或(b) AA1、AA3、AA4、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA3、AA4、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA1、AA3、AA5、AA6和AA7、(b) AA1、AA3、AA5、AA6和AA8、(c) AA1、AA3、AA5、AA6和AA9、(d) AA1、AA3、AA5、AA6和AA10或(e) AA1、AA3、AA5、AA6和AA11。作为又一种备选,所述至少五个氨基酸残基可以是(a) AA1、AA3、AA5、AA7和AA8、(b) AA1、AA3、AA5、AA7和AA9、(c) AA1、AA3、AA5、AA7和AA10或(d) AA1、AA3、AA5、AA7和AA11。所述至少五个氨基酸残基还可以是(a) AA1、AA3、AA5、AA8和AA9、(b) AA1、AA3、AA5、AA8和AA10或(c) AA1、AA3、AA5、AA8和AA11。所述至少五个氨基酸残基还可以进一步可以是(a) AA1、AA3、AA5、AA9和AA10或(b) AA1、AA3、AA5、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA3、AA5、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA1、AA3、AA6、AA7和AA8、(b) AA1、AA3、AA6、AA7和AA9、(c) AA1、AA3、AA6、AA7和AA10或(d) AA1、AA3、AA6、AA7和AA11。所述至少五个氨基酸残基还可以是(a) AA1、AA3、AA6、AA8和AA9、(b) AA1、AA3、AA6、AA8和AA10或(c) AA1、AA3、AA6、AA8和AA11。

所述至少五个氨基酸残基还进一步可以是 (a) AA1、AA3、AA6、AA9和AA10或 (b) AA1、AA3、AA6、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA3、AA6、AA10和AA11。此类所述至少五个氨基酸残基可以是 (a) AA1、AA3、AA7、AA8和AA9、(b) AA1、AA3、AA7、AA8和AA10或 (c) AA1、AA3、AA7、AA8和AA11。所述至少五个氨基酸残基还进一步可以是 (a) AA1、AA3、AA7、AA9和AA10或 (b) AA1、AA3、AA7、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA3、AA7、AA10和AA11。此类所述至少五个氨基酸残基可以是 (a) AA1、AA3、AA8、AA9和AA10或 (b) AA1、AA3、AA8、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA3、AA8、AA10和AA11。所述至少五个氨基酸残基还可以是AA1、AA3、AA9、AA10和AA11。此类所述至少五个氨基酸残基可以是 (a) AA1、AA4、AA5、AA6和AA7、(b) AA1、AA4、AA5、AA6和AA8、(c) AA1、AA4、AA5、AA6和AA9、(d) AA1、AA4、AA5、AA6和AA10或 (e) AA1、AA4、AA5、AA6和AA11。作为又一种备选,所述至少五个氨基酸残基可以是 (a) AA1、AA4、AA5、AA7和AA8、(b) AA1、AA4、AA5、AA7和AA9、(c) AA1、AA4、AA5、AA7和AA10或 (d) AA1、AA4、AA5、AA7和AA11。所述至少五个氨基酸残基还可以是 (a) AA1、AA4、AA5、AA8和AA9、(b) AA1、AA4、AA5、AA8和AA10或 (c) AA1、AA4、AA5、AA8和AA11。所述至少五个氨基酸残基还进一步可以是 (a) AA1、AA4、AA5、AA9和AA10或 (b) AA1、AA4、AA5、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA4、AA5、AA10和AA11。此类所述至少五个氨基酸残基可以是 (a) AA1、AA4、AA6、AA7和AA8、(b) AA1、AA4、AA6、AA7和AA9、(c) AA1、AA4、AA6、AA7和AA10或 (d) AA1、AA4、AA6、AA7和AA11。所述至少五个氨基酸残基还可以是 (a) AA1、AA4、AA6、AA8和AA9、(b) AA1、AA4、AA6、AA8和AA10或 (c) AA1、AA4、AA6、AA8和AA11。所述至少五个氨基酸残基还进一步可以是 (a) AA1、AA4、AA6、AA9和AA10或 (b) AA1、AA4、AA6、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA4、AA6、AA10和AA11。此类所述至少五个氨基酸残基可以是 (a) AA1、AA4、AA7、AA8和AA9、(b) AA1、AA4、AA7、AA8和AA10或 (c) AA1、AA4、AA7、AA8和AA11。所述至少五个氨基酸残基还进一步可以是 (a) AA1、AA4、AA7、AA9和AA10或 (b) AA1、AA4、AA7、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA4、AA7、AA10和AA11。此类所述至少五个氨基酸残基可以是 (a) AA1、AA4、AA8、AA9和AA10或 (b) AA1、AA4、AA8、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA4、AA8、AA10和AA11。所述至少五个氨基酸残基还可以是AA1、AA4、AA9、AA10和AA11。此类所述至少五个氨基酸残基可以是 (a) AA1、AA5、AA6、AA7和AA8、(b) AA1、AA5、AA6、AA7和AA9、(c) AA1、AA5、AA6、AA7和AA10或 (d) AA1、AA5、AA6、AA7和AA11。所述至少五个氨基酸残基还可以是 (a) AA1、AA5、AA6、AA8和AA9、(b) AA1、AA5、AA6、AA8和AA10或 (c) AA1、AA5、AA6、AA8和AA11。所述至少五个氨基酸残基还进一步可以是 (a) AA1、AA5、AA6、AA9和AA10或 (b) AA1、AA5、AA6、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA5、AA6、AA10和AA11。此类所述至少五个氨基酸残基可以是 (a) AA1、AA5、AA7、AA8和AA9、(b) AA1、AA5、AA7、AA8和AA10或 (c) AA1、AA5、AA7、AA8和AA11。所述至少五个氨基酸残基还进一步可以是 (a) AA1、AA5、AA7、AA9和AA10或 (b) AA1、AA5、AA7、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA5、AA7、AA10和AA11。此类所述至少五个氨基酸残基可以是 (a) AA1、AA5、AA8、AA9和AA10或 (b) AA1、AA5、AA8、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA5、AA8、AA10和AA11。所述至少五个氨基酸残基还可以是AA1、AA5、AA9、AA10和AA11。此类所述至少五个氨基酸残基可以是 (a) AA1、AA6、AA7、AA8和AA9、(b) AA1、AA6、AA7、AA8和AA10或 (c) AA1、AA6、AA7、AA8和AA11。所述至少五个氨基酸残基还进一步可以是 (a) AA1、AA6、AA7、AA9和AA10或 (b) AA1、AA6、AA7、AA9和AA11。所述至少五个氨基酸残基还可以

是AA1、AA6、AA7、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA1、AA6、AA8、AA9和AA10或(b) AA1、AA6、AA8、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA6、AA8、AA10和AA11。所述至少五个氨基酸残基还可以是AA1、AA6、AA9、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA1、AA7、AA8、AA9和AA10或(b) AA1、AA7、AA8、AA9和AA11。所述至少五个氨基酸残基还可以是AA1、AA7、AA8、AA10和AA11。所述至少五个氨基酸残基还可以是AA1、AA7、AA9、AA10和AA11。所述至少五个氨基酸残基还可以是AA1、AA8、AA9、AA10和AA11。

[0062] 此类所述至少五个氨基酸残基可以是(a) AA2、AA3、AA4、AA5和AA6、(b) AA2、AA3、AA4、AA5和AA7、(c) AA2、AA3、AA4、AA5和AA8、(d) AA2、AA3、AA4、AA5和AA9、(e) AA2、AA3、AA4、AA5和AA10或(f) AA2、AA3、AA4、AA5和AA11。作为另一种备选,所述至少五个氨基酸残基可以是(a) AA2、AA3、AA4、AA6和AA7、(b) AA2、AA3、AA4、AA6和AA8、(c) AA2、AA3、AA4、AA6和AA9、(d) AA2、AA3、AA4、AA6和AA10或(e) AA2、AA3、AA4、AA6和AA11。作为又一种备选,所述至少五个氨基酸残基可以是(a) AA2、AA3、AA4、AA7和AA8、(b) AA2、AA3、AA4、AA7和AA9、(c) AA2、AA3、AA4、AA7和AA10或(d) AA2、AA3、AA4、AA7和AA11。所述至少五个氨基酸残基还可以是(a) AA2、AA3、AA4、AA8和AA9、(b) AA2、AA3、AA4、AA8和AA10或(c) AA2、AA3、AA4、AA8和AA11。所述至少五个氨基酸残基还进一步可以是(a) AA2、AA3、AA4、AA9和AA10或(b) AA2、AA3、AA4、AA9和AA11。所述至少五个氨基酸残基还可以是AA2、AA3、AA4、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA2、AA3、AA5、AA6和AA7、(b) AA2、AA3、AA5、AA6和AA8、(c) AA2、AA3、AA5、AA6和AA9、(d) AA2、AA3、AA5、AA6和AA10或(e) AA2、AA3、AA5、AA6和AA11。作为又一种备选,所述至少五个氨基酸残基可以是(a) AA2、AA3、AA5、AA7和AA8、(b) AA2、AA3、AA5、AA7和AA9、(c) AA2、AA3、AA5、AA7和AA10或(d) AA2、AA3、AA5、AA7和AA11。所述至少五个氨基酸残基还可以是(a) AA2、AA3、AA5、AA8和AA9、(b) AA2、AA3、AA5、AA8和AA10或(c) AA2、AA3、AA5、AA8和AA11。所述至少五个氨基酸残基还进一步可以是(a) AA2、AA3、AA5、AA9和AA10或(b) AA2、AA3、AA5、AA9和AA11。所述至少五个氨基酸残基还可以是AA2、AA3、AA5、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA2、AA3、AA6、AA7和AA8、(b) AA2、AA3、AA6、AA7和AA9、(c) AA2、AA3、AA6、AA7和AA10或(d) AA2、AA3、AA6、AA7和AA11。所述至少五个氨基酸残基还可以是(a) AA2、AA3、AA6、AA8和AA9、(b) AA2、AA3、AA6、AA8和AA10或(c) AA2、AA3、AA6、AA8和AA11。所述至少五个氨基酸残基还进一步可以是(a) AA2、AA3、AA6、AA9和AA10或(b) AA2、AA3、AA6、AA9和AA11。所述至少五个氨基酸残基还可以是AA2、AA3、AA6、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA2、AA3、AA7、AA8和AA9、(b) AA2、AA3、AA7、AA8和AA10或(c) AA2、AA3、AA7、AA8和AA11。所述至少五个氨基酸残基还进一步可以是(a) AA2、AA3、AA7、AA9和AA10或(b) AA2、AA3、AA7、AA9和AA11。所述至少五个氨基酸残基还可以是AA2、AA3、AA7、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA2、AA3、AA8、AA9和AA10或(b) AA2、AA3、AA8、AA9和AA11。所述至少五个氨基酸残基还可以是AA2、AA3、AA8、AA10和AA11。所述至少五个氨基酸残基还可以是AA2、AA3、AA9、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA2、AA4、AA5、AA6和AA7、(b) AA2、AA4、AA5、AA6和AA8、(c) AA2、AA4、AA5、AA6和AA9、(d) AA2、AA4、AA5、AA6和AA10或(e) AA2、AA4、AA5、AA6和AA11。作为又一种备选,所述至少五个氨基酸残基可以是(a) AA2、AA4、AA5、AA7和AA8、(b) AA2、AA4、AA5、AA7和AA9、(c) AA2、AA4、AA5、AA7和AA10或(d) AA2、AA4、AA5、AA7和AA11。所述至少五个氨基酸残基还可以是(a) AA2、AA4、AA5、AA8和AA9、(b) AA2、AA4、AA5、AA8和AA10或(c) AA2、AA4、AA5、AA8和AA11。所述至少五个

氨基酸残基还进一步可以是 (a) AA2、AA4、AA5、AA9 和 AA10 或 (b) AA2、AA4、AA5、AA9 和 AA11。所述至少五个氨基酸残基还可以是 AA2、AA4、AA5、AA10 和 AA11。此类所述至少五个氨基酸残基可以是 (a) AA2、AA4、AA6、AA7 和 AA8、(b) AA2、AA4、AA6、AA7 和 AA9、(c) AA2、AA4、AA6、AA7 和 AA10 或 (d) AA2、AA4、AA6、AA7 和 AA11。所述至少五个氨基酸残基还可以是 (a) AA2、AA4、AA6、AA8 和 AA9、(b) AA2、AA4、AA6、AA8 和 AA10 或 (c) AA2、AA4、AA6、AA8 和 AA11。所述至少五个氨基酸残基还进一步可以是 (a) AA2、AA4、AA6、AA9 和 AA10 或 (b) AA2、AA4、AA6、AA9 和 AA11。所述至少五个氨基酸残基还可以是 AA2、AA4、AA6、AA10 和 AA11。此类所述至少五个氨基酸残基可以是 (a) AA2、AA4、AA7、AA8 和 AA9、(b) AA2、AA4、AA7、AA8 和 AA10 或 (c) AA2、AA4、AA7、AA8 和 AA11。所述至少五个氨基酸残基还进一步可以是 (a) AA2、AA4、AA7、AA9 和 AA10 或 (b) AA2、AA4、AA7、AA9 和 AA11。所述至少五个氨基酸残基还可以是 AA2、AA4、AA7、AA10 和 AA11。此类所述至少五个氨基酸残基可以是 (a) AA2、AA4、AA8、AA9 和 AA10 或 (b) AA2、AA4、AA8、AA9 和 AA11。所述至少五个氨基酸残基还可以是 AA2、AA4、AA8、AA10 和 AA11。所述至少五个氨基酸残基还可以是 AA2、AA4、AA9、AA10 和 AA11。此类所述至少五个氨基酸残基可以是 (a) AA2、AA5、AA6、AA7 和 AA8、(b) AA2、AA5、AA6、AA7 和 AA9、(c) AA2、AA5、AA6、AA7 和 AA10 或 (d) AA2、AA5、AA6、AA7 和 AA11。所述至少五个氨基酸残基还可以是 (a) AA2、AA5、AA6、AA8 和 AA9、(b) AA2、AA5、AA6、AA8 和 AA10 或 (c) AA2、AA5、AA6、AA8 和 AA11。所述至少五个氨基酸残基还进一步可以是 (a) AA2、AA5、AA6、AA9 和 AA10 或 (b) AA2、AA5、AA6、AA9 和 AA11。所述至少五个氨基酸残基还可以是 AA2、AA5、AA6、AA10 和 AA11。此类所述至少五个氨基酸残基可以是 (a) AA2、AA5、AA7、AA8 和 AA9、(b) AA2、AA5、AA7、AA8 和 AA10 或 (c) AA2、AA5、AA7、AA8 和 AA11。所述至少五个氨基酸残基还进一步可以是 (a) AA2、AA5、AA7、AA9 和 AA10 或 (b) AA2、AA5、AA7、AA9 和 AA11。所述至少五个氨基酸残基还可以是 AA2、AA5、AA7、AA10 和 AA11。此类所述至少五个氨基酸残基可以是 (a) AA2、AA5、AA8、AA9 和 AA10 或 (b) AA2、AA5、AA8、AA9 和 AA11。所述至少五个氨基酸残基还可以是 AA2、AA5、AA8、AA10 和 AA11。所述至少五个氨基酸残基还可以是 AA2、AA5、AA9、AA10 和 AA11。此类所述至少五个氨基酸残基可以是 (a) AA2、AA6、AA7、AA8 和 AA9、(b) AA2、AA6、AA7、AA8 和 AA10 或 (c) AA2、AA6、AA7、AA8 和 AA11。所述至少五个氨基酸残基还进一步可以是 (a) AA2、AA6、AA7、AA9 和 AA10 或 (b) AA2、AA6、AA7、AA9 和 AA11。所述至少五个氨基酸残基还可以是 AA2、AA6、AA7、AA10 和 AA11。此类所述至少五个氨基酸残基可以是 (a) AA2、AA6、AA8、AA9 和 AA10 或 (b) AA2、AA6、AA8、AA9 和 AA11。所述至少五个氨基酸残基还可以是 AA2、AA6、AA8、AA10 和 AA11。所述至少五个氨基酸残基还可以是 AA2、AA6、AA9、AA10 和 AA11。此类所述至少五个氨基酸残基可以是 (a) AA2、AA7、AA8、AA9 和 AA10 或 (b) AA2、AA7、AA8、AA9 和 AA11。所述至少五个氨基酸残基还可以是 AA2、AA7、AA8、AA10 和 AA11。所述至少五个氨基酸残基还可以是 AA2、AA7、AA9、AA10 和 AA11。所述至少五个氨基酸残基还可以是 AA2、AA8、AA9、AA10 和 AA11。

[0063] 此类所述至少五个氨基酸残基可以是 (a) AA3、AA4、AA5、AA6 和 AA7、(b) AA3、AA4、AA5、AA6 和 AA8、(c) AA3、AA4、AA5、AA6 和 AA9、(d) AA3、AA4、AA5、AA6 和 AA10 或 (e) AA3、AA4、AA5、AA6 和 AA11。作为又一种备选,所述至少五个氨基酸残基可以是 (a) AA3、AA4、AA5、AA7 和 AA8、(b) AA3、AA4、AA5、AA7 和 AA9、(c) AA3、AA4、AA5、AA7 和 AA10 或 (d) AA3、AA4、AA5、AA7 和 AA11。所述至少五个氨基酸残基还可以是 (a) AA3、AA4、AA5、AA8 和 AA9、(b) AA3、AA4、AA5、AA8 和 AA10 或 (c) AA3、AA4、AA5、AA8 和 AA11。所述至少五个氨基酸残基还进一步可以是 (a) AA3、AA4、AA5、AA9 和 AA10 或 (b) AA3、AA4、AA5、AA9 和 AA11。所述至少五个氨基酸残基还可以是

[illegible]

[0064] 此类所述至少五个氨基酸残基可以是 (a) AA4、AA5、AA6、AA7和AA8、(b) AA4、AA5、AA6、AA7和AA9、(c) AA4、AA5、AA6、AA7和AA10或 (d) AA4、AA5、AA6、AA7和AA11。所述至少五个氨基酸残基还可以是 (a) AA4、AA5、AA6、AA8和AA9、(b) AA4、AA5、AA6、AA8和AA10或 (c) AA4、AA5、AA6、AA8和AA11。所述至少五个氨基酸残基还进一步可以是 (a) AA4、AA5、AA6、AA9和AA10或 (b) AA4、AA5、AA6、AA9和AA11。所述至少五个氨基酸残基还可以是AA4、AA5、AA6、AA10和AA11。此类所述至少五个氨基酸残基可以是 (a) AA4、AA5、AA7、AA8和AA9、(b) AA4、AA5、AA7、AA8和AA10或 (c) AA4、AA5、AA7、AA8和AA11。所述至少五个氨基酸残基还进一步可以是 (a) AA4、AA5、AA7、AA9和AA10或 (b) AA4、AA5、AA7、AA9和AA11。所述至少五个氨基酸残基还可

以是AA4、AA5、AA7、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA4、AA5、AA8、AA9和AA10或(b) AA4、AA5、AA8、AA9和AA11。所述至少五个氨基酸残基还可以是AA4、AA5、AA8、AA10和AA11。所述至少五个氨基酸残基还可以是AA4、AA5、AA9、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA4、AA6、AA7、AA8和AA9、(b) AA4、AA6、AA7、AA8和AA10或(c) AA4、AA6、AA7、AA8和AA11。所述至少五个氨基酸残基还进一步可以是(a) AA4、AA6、AA7、AA9和AA10或(b) AA4、AA6、AA7、AA9和AA11。所述至少五个氨基酸残基还可以是AA4、AA6、AA7、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA4、AA6、AA8、AA9和AA10或(b) AA4、AA6、AA8、AA9和AA11。所述至少五个氨基酸残基还可以是AA4、AA6、AA8、AA10和AA11。所述至少五个氨基酸残基还可以是AA4、AA6、AA9、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA4、AA7、AA8、AA9和AA10或(b) AA4、AA7、AA8、AA9和AA11。所述至少五个氨基酸残基还可以是AA4、AA7、AA8、AA10和AA11。所述至少五个氨基酸残基还可以是AA4、AA7、AA9、AA10和AA11。所述至少五个氨基酸残基还可以是AA4、AA8、AA9、AA10和AA11。

[0065] 此类所述至少五个氨基酸残基可以是(a) AA5、AA6、AA7、AA8和AA9、(b) AA5、AA6、AA7、AA8和AA10或(c) AA5、AA6、AA7、AA8和AA11。所述至少五个氨基酸残基还进一步可以是(a) AA5、AA6、AA7、AA9和AA10或(b) AA5、AA6、AA7、AA9和AA11。所述至少五个氨基酸残基还可以是AA5、AA6、AA7、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA5、AA6、AA8、AA9和AA10或(b) AA5、AA6、AA8、AA9和AA11。所述至少五个氨基酸残基还可以是AA5、AA6、AA8、AA10和AA11。所述至少五个氨基酸残基还可以是AA5、AA6、AA9、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA5、AA7、AA8、AA9和AA10或(b) AA5、AA7、AA8、AA9和AA11。所述至少五个氨基酸残基还可以是AA5、AA7、AA8、AA10和AA11。所述至少五个氨基酸残基还可以是AA5、AA7、AA9、AA10和AA11。所述至少五个氨基酸残基还可以是AA5、AA8、AA9、AA10和AA11。此类所述至少五个氨基酸残基可以是(a) AA6、AA7、AA8、AA9和AA10或(b) AA6、AA7、AA8、AA9和AA11。所述至少五个氨基酸残基还可以是AA6、AA7、AA8、AA10和AA11。所述至少五个氨基酸残基还可以是AA6、AA7、AA9、AA10和AA11。所述至少五个氨基酸残基还可以是AA6、AA8、AA9、AA10和AA11。所述至少五个氨基酸残基还可以是AA7、AA8、AA9、AA10和AA11。

[0066] 此类所述至少六个氨基酸残基可以是(a) AA1、AA2、AA3、AA4、AA5和AA6、(b) AA1、AA2、AA3、AA4、AA5和AA7、(c) AA1、AA2、AA3、AA4、AA5和AA8、(d) AA1、AA2、AA3、AA4、AA5和AA9、(e) AA1、AA2、AA3、AA4、AA5和AA10或(f) AA1、AA2、AA3、AA4、AA5和AA11。作为另一种备选,所述至少六个氨基酸残基可以是(a) AA1、AA2、AA3、AA4、AA6和AA7、(b) AA1、AA2、AA3、AA4、AA6和AA8、(c) AA1、AA2、AA3、AA4、AA6和AA9、(d) AA1、AA2、AA3、AA4、AA6和AA10或(e) AA1、AA2、AA3、AA4、AA6和AA11。作为又一种备选,所述至少六个氨基酸残基可以是(a) AA1、AA2、AA3、AA4、AA7和AA8、(b) AA1、AA2、AA3、AA4、AA7和AA9、(c) AA1、AA2、AA3、AA4、AA7和AA10或(d) AA1、AA2、AA3、AA4、AA7和AA11。所述至少六个氨基酸残基还可以是(a) AA1、AA2、AA3、AA4、AA8和AA9、(b) AA1、AA2、AA3、AA4、AA8和AA10或(c) AA1、AA2、AA3、AA4、AA8和AA11。所述至少六个氨基酸残基还进一步可以是(a) AA1、AA2、AA3、AA4、AA9和AA10或(b) AA1、AA2、AA3、AA4、AA9和AA11。所述至少六个氨基酸残基还可以是AA1、AA2、AA3、AA4、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA1、AA2、AA3、AA5、AA6和AA7、(b) AA1、AA2、AA3、AA5、AA6和AA8、(c) AA1、AA2、AA3、AA5、AA6和AA9、(d) AA1、AA2、AA3、AA5、AA6和AA10或(e) AA1、AA2、AA3、AA5、AA6和AA11。作为又一种备选,所述至少六个氨基酸残基可以是(a) AA1、AA2、AA3、AA5、

[illegible]

AA1、AA2、AA5、AA6、AA8和AA10或(c) AA1、AA2、AA5、AA6、AA8和AA11。所述至少六个氨基酸残基还进一步可以是(a) AA1、AA2、AA5、AA6、AA9和AA10或(b) AA1、AA2、AA5、AA6、AA9和AA11。所述至少六个氨基酸残基还可以是AA1、AA2、AA5、AA6、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA1、AA2、AA5、AA7、AA8和AA9、(b) AA1、AA2、AA5、AA7、AA8和AA10或(c) AA1、AA2、AA5、AA7、AA8和AA11。所述至少六个氨基酸残基还进一步可以是(a) AA1、AA2、AA5、AA7、AA9和AA10或(b) AA1、AA2、AA5、AA7、AA9和AA11。所述至少六个氨基酸残基还可以是AA1、AA2、AA5、AA7、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA1、AA2、AA5、AA8、AA9和AA10或(b) AA1、AA2、AA5、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA1、AA2、AA5、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA1、AA2、AA5、AA9、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA1、AA2、AA6、AA7、AA8和AA9、(b) AA1、AA2、AA6、AA7、AA8和AA10或(c) AA1、AA2、AA6、AA7、AA8和AA11。所述至少六个氨基酸残基还进一步可以是(a) AA1、AA2、AA6、AA7、AA9和AA10或(b) AA1、AA2、AA6、AA7、AA9和AA11。所述至少六个氨基酸残基还可以是AA1、AA2、AA6、AA7、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA1、AA2、AA6、AA8、AA9和AA10或(b) AA1、AA2、AA6、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA1、AA2、AA6、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA1、AA2、AA6、AA9、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA1、AA2、AA7、AA8、AA9和AA10或(b) AA1、AA2、AA7、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA1、AA2、AA7、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA1、AA2、AA7、AA9、AA10和AA11。所述至少六个氨基酸残基还可以是AA1、AA2、AA8、AA9、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA1、AA3、AA4、AA5、AA6和AA7、(b) AA1、AA3、AA4、AA5、AA6和AA8、(c) AA1、AA3、AA4、AA5、AA6和AA9、(d) AA1、AA3、AA4、AA5、AA6和AA10或(e) AA1、AA3、AA4、AA5、AA6和AA11。作为又一种备选,所述至少六个氨基酸残基可以是(a) AA1、AA3、AA4、AA5、AA7和AA8、(b) AA1、AA3、AA4、AA5、AA7和AA9、(c) AA1、AA3、AA4、AA5、AA7和AA10或(d) AA1、AA3、AA4、AA5、AA7和AA11。所述至少六个氨基酸残基还可以是(a) AA1、AA3、AA4、AA5、AA8和AA9、(b) AA1、AA3、AA4、AA5、AA8和AA10或(c) AA1、AA3、AA4、AA5、AA8和AA11。所述至少六个氨基酸残基还进一步可以是(a) AA1、AA3、AA4、AA5、AA9和AA10或(b) AA1、AA3、AA4、AA5、AA9和AA11。所述至少六个氨基酸残基还可以是AA1、AA3、AA4、AA5、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA1、AA3、AA4、AA6、AA7和AA8、(b) AA1、AA3、AA4、AA6、AA7和AA9、(c) AA1、AA3、AA4、AA6、AA7和AA10或(d) AA1、AA3、AA4、AA6、AA7和AA11。所述至少六个氨基酸残基还可以是(a) AA1、AA3、AA4、AA6、AA8和AA9、(b) AA1、AA3、AA4、AA6、AA8和AA10或(c) AA1、AA3、AA4、AA6、AA8和AA11。所述至少六个氨基酸残基还进一步可以是(a) AA1、AA3、AA4、AA6、AA9和AA10或(b) AA1、AA3、AA4、AA6、AA9和AA11。所述至少六个氨基酸残基还可以是AA1、AA3、AA4、AA6、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA1、AA3、AA4、AA7、AA8和AA9、(b) AA1、AA3、AA4、AA7、AA8和AA10或(c) AA1、AA3、AA4、AA7、AA8和AA11。所述至少六个氨基酸残基还进一步可以是(a) AA1、AA3、AA4、AA7、AA9和AA10或(b) AA1、AA3、AA4、AA7、AA9和AA11。所述至少六个氨基酸残基还可以是AA1、AA3、AA4、AA7、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA1、AA3、AA4、AA8、AA9和AA10或(b) AA1、AA3、AA4、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA1、AA3、AA4、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA1、AA3、AA4、AA9、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA1、AA3、AA5、

[illegible]

至少六个氨基酸残基可以是 (a) AA1、AA4、AA7、AA8、AA9和AA10或 (b) AA1、AA4、AA7、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA1、AA4、AA7、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA1、AA4、AA7、AA9、AA10和AA11。所述至少六个氨基酸残基还可以是AA1、AA4、AA8、AA9、AA10和AA11。此类所述至少六个氨基酸残基可以是 (a) AA1、AA5、AA6、AA7、AA8和AA9、(b) AA1、AA5、AA6、AA7、AA8和AA10或 (c) AA1、AA5、AA6、AA7、AA8和AA11。所述至少六个氨基酸残基还进一步可以是 (a) AA1、AA5、AA6、AA7、AA9和AA10或 (b) AA1、AA5、AA6、AA7、AA9和AA11。所述至少六个氨基酸残基还可以是AA1、AA5、AA6、AA7、AA10和AA11。此类所述至少六个氨基酸残基可以是 (a) AA1、AA5、AA6、AA8、AA9和AA10或 (b) AA1、AA5、AA6、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA1、AA5、AA6、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA1、AA5、AA6、AA9、AA10和AA11。此类所述至少六个氨基酸残基可以是 (a) AA1、AA5、AA7、AA8、AA9和AA10或 (b) AA1、AA5、AA7、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA1、AA5、AA7、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA1、AA5、AA7、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA1、AA5、AA8、AA9、AA10和AA11。此类所述至少六个氨基酸残基可以是 (a) AA1、AA6、AA7、AA8、AA9和AA10或 (b) AA1、AA6、AA7、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA1、AA6、AA7、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA1、AA6、AA7、AA9、AA10和AA11。所述至少六个氨基酸残基还可以是AA1、AA6、AA8、AA9、AA10和AA11。所述至少六个氨基酸残基还可以是AA1、AA7、AA8、AA9、AA10和AA11。

[0067] 此类所述至少六个氨基酸残基可以是 (a) AA2、AA3、AA4、AA5、AA6和AA7、(b) AA2、AA3、AA4、AA5、AA6和AA8、(c) AA2、AA3、AA4、AA5、AA6和AA9、(d) AA2、AA3、AA4、AA5、AA6和AA10或 (e) AA2、AA3、AA4、AA5、AA6和AA11。作为又一种备选,所述至少六个氨基酸残基可以是 (a) AA2、AA3、AA4、AA5、AA7和AA8、(b) AA2、AA3、AA4、AA5、AA7和AA9、(c) AA2、AA3、AA4、AA5、AA7和AA10或 (d) AA2、AA3、AA4、AA5、AA7和AA11。所述至少六个氨基酸残基还可以是 (a) AA2、AA3、AA4、AA5、AA8和AA9、(b) AA2、AA3、AA4、AA5、AA8和AA10或 (c) AA2、AA3、AA4、AA5、AA8和AA11。所述至少六个氨基酸残基还进一步可以是 (a) AA2、AA3、AA4、AA5、AA9和AA10或 (b) AA2、AA3、AA4、AA5、AA9和AA11。所述至少六个氨基酸残基还可以是AA2、AA3、AA4、AA5、AA10和AA11。此类所述至少六个氨基酸残基可以是 (a) AA2、AA3、AA4、AA6、AA7和AA8、(b) AA2、AA3、AA4、AA6、AA7和AA9、(c) AA2、AA3、AA4、AA6、AA7和AA10或 (d) AA2、AA3、AA4、AA6、AA7和AA11。所述至少六个氨基酸残基还可以是 (a) AA2、AA3、AA4、AA6、AA8和AA9、(b) AA2、AA3、AA4、AA6、AA8和AA10或 (c) AA2、AA3、AA4、AA6、AA8和AA11。所述至少六个氨基酸残基还进一步可以是 (a) AA2、AA3、AA4、AA6、AA9和AA10或 (b) AA2、AA3、AA4、AA6、AA9和AA11。所述至少六个氨基酸残基还可以是AA2、AA3、AA4、AA6、AA10和AA11。此类所述至少六个氨基酸残基可以是 (a) AA2、AA3、AA4、AA7、AA8和AA9、(b) AA2、AA3、AA4、AA7、AA8和AA10或 (c) AA2、AA3、AA4、AA7、AA8和AA11。所述至少六个氨基酸残基还进一步可以是 (a) AA2、AA3、AA4、AA7、AA9和AA10或 (b) AA2、AA3、AA4、AA7、AA9和AA11。所述至少六个氨基酸残基还可以是AA2、AA3、AA4、AA7、AA10和AA11。此类所述至少六个氨基酸残基可以是 (a) AA2、AA3、AA4、AA8、AA9和AA10或 (b) AA2、AA3、AA4、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA2、AA3、AA4、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA2、AA3、AA4、AA9、AA10和AA11。此类所述至少六个氨基酸残基可以是 (a) AA2、AA3、AA5、AA6、AA7和AA8、(b) AA2、AA3、AA5、AA6、AA7和AA9、(c)

[illegible]

AA10或(b) AA2、AA4、AA7、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA2、AA4、AA7、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA2、AA4、AA7、AA9、AA10和AA11。所述至少六个氨基酸残基还可以是AA2、AA4、AA8、AA9、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA2、AA5、AA6、AA7、AA8和AA9、(b) AA2、AA5、AA6、AA7、AA8和AA10或(c) AA2、AA5、AA6、AA7、AA8和AA11。所述至少六个氨基酸残基还进一步可以是(a) AA2、AA5、AA6、AA7、AA9和AA10或(b) AA2、AA5、AA6、AA7、AA9和AA11。所述至少六个氨基酸残基还可以是AA2、AA5、AA6、AA7、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA2、AA5、AA6、AA8、AA9和AA10或(b) AA2、AA5、AA6、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA2、AA5、AA6、AA9、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA2、AA5、AA7、AA8、AA9和AA10或(b) AA2、AA5、AA7、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA2、AA5、AA7、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA2、AA5、AA7、AA9、AA10和AA11。所述至少六个氨基酸残基还可以是AA2、AA5、AA8、AA9、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA2、AA6、AA7、AA8、AA9和AA10或(b) AA2、AA6、AA7、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA2、AA6、AA7、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA2、AA6、AA7、AA9、AA10和AA11。所述至少六个氨基酸残基还可以是AA2、AA6、AA8、AA9、AA10和AA11。所述至少六个氨基酸残基还可以是AA2、AA7、AA8、AA9、AA10和AA11。

[0068] 此类所述至少六个氨基酸残基可以是(a) AA3、AA4、AA5、AA6、AA7和AA8、(b) AA3、AA4、AA5、AA6、AA7和AA9、(c) AA3、AA4、AA5、AA6、AA7和AA10或(d) AA3、AA4、AA5、AA6、AA7和AA11。所述至少六个氨基酸残基还可以是(a) AA3、AA4、AA5、AA6、AA8和AA9、(b) AA3、AA4、AA5、AA6、AA8和AA10或(c) AA3、AA4、AA5、AA6、AA8和AA11。所述至少六个氨基酸残基还进一步可以是(a) AA3、AA4、AA5、AA6、AA9和AA10或(b) AA3、AA4、AA5、AA6、AA9和AA11。所述至少六个氨基酸残基还可以是AA3、AA4、AA5、AA6、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA3、AA4、AA5、AA7、AA8和AA9、(b) AA3、AA4、AA5、AA7、AA8和AA10或(c) AA3、AA4、AA5、AA7、AA8和AA11。所述至少六个氨基酸残基还进一步可以是(a) AA3、AA4、AA5、AA7、AA9和AA10或(b) AA3、AA4、AA5、AA7、AA9和AA11。所述至少六个氨基酸残基还可以是AA3、AA4、AA5、AA7、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA3、AA4、AA5、AA8、AA9和AA10或(b) AA3、AA4、AA5、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA3、AA4、AA5、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA3、AA4、AA5、AA9、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA3、AA4、AA6、AA7、AA8和AA9、(b) AA3、AA4、AA6、AA7、AA8和AA10或(c) AA3、AA4、AA6、AA7、AA8和AA11。所述至少六个氨基酸残基还进一步可以是(a) AA3、AA4、AA6、AA7、AA9和AA10或(b) AA3、AA4、AA6、AA7、AA9和AA11。所述至少六个氨基酸残基还可以是AA3、AA4、AA6、AA7、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA3、AA4、AA6、AA8、AA9和AA10或(b) AA3、AA4、AA6、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA3、AA4、AA6、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA3、AA4、AA6、AA9、AA10和AA11。此类所述至少六个氨基酸残基可以是(a) AA3、AA4、AA7、AA8、AA9和AA10或(b) AA3、AA4、AA7、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA3、AA4、AA7、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA3、AA4、AA7、AA9、AA10和AA11。所述至少六个氨基酸残基还可以是AA3、AA4、AA8、AA9、AA10和AA11。此类所述至少六个氨基酸残基可

以是 (a) AA3、AA5、AA6、AA7、AA8和AA9、(b) AA3、AA5、AA6、AA7、AA8和AA10或 (c) AA3、AA5、AA6、AA7、AA8和AA11。所述至少六个氨基酸残基还进一步可以是 (a) AA3、AA5、AA6、AA7、AA9和AA10或 (b) AA3、AA5、AA6、AA7、AA9和AA11。所述至少六个氨基酸残基还可以是AA3、AA5、AA6、AA7、AA10和AA11。此类所述至少六个氨基酸残基可以是 (a) AA3、AA5、AA6、AA8、AA9和AA10或 (b) AA3、AA5、AA6、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA3、AA5、AA6、AA8、AA10和AA11。此类所述至少六个氨基酸残基可以是 (a) AA3、AA5、AA7、AA8、AA9和AA10或 (b) AA3、AA5、AA7、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA3、AA5、AA7、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA3、AA5、AA7、AA9、AA10和AA11。所述至少六个氨基酸残基还可以是AA3、AA5、AA8、AA9、AA10和AA11。此类所述至少六个氨基酸残基可以是 (a) AA3、AA6、AA7、AA8、AA9和AA10或 (b) AA3、AA6、AA7、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA3、AA6、AA7、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA3、AA6、AA7、AA9、AA10和AA11。所述至少六个氨基酸残基还可以是AA3、AA6、AA8、AA9、AA10和AA11。所述至少六个氨基酸残基还可以是AA3、AA7、AA8、AA9、AA10和AA11。

[0069] 此类所述至少六个氨基酸残基可以是 (a) AA4、AA5、AA6、AA7、AA8和AA9、(b) AA4、AA5、AA6、AA7、AA8和AA10或 (c) AA4、AA5、AA6、AA7、AA8和AA11。所述至少六个氨基酸残基还进一步可以是 (a) AA4、AA5、AA6、AA7、AA9和AA10或 (b) AA4、AA5、AA6、AA7、AA9和AA11。所述至少六个氨基酸残基还可以是AA4、AA5、AA6、AA7、AA10和AA11。此类所述至少六个氨基酸残基可以是 (a) AA4、AA5、AA6、AA8、AA9和AA10或 (b) AA4、AA5、AA6、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA4、AA5、AA6、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA4、AA5、AA6、AA9、AA10和AA11。此类所述至少六个氨基酸残基可以是 (a) AA4、AA5、AA7、AA8、AA9和AA10或 (b) AA4、AA5、AA7、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA4、AA5、AA7、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA4、AA5、AA7、AA9、AA10和AA11。所述至少六个氨基酸残基还可以是AA4、AA5、AA8、AA9、AA10和AA11。此类所述至少六个氨基酸残基可以是 (a) AA4、AA6、AA7、AA8、AA9和AA10或 (b) AA4、AA6、AA7、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA4、AA6、AA7、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA4、AA6、AA7、AA9、AA10和AA11。所述至少六个氨基酸残基还可以是AA4、AA6、AA8、AA9、AA10和AA11。所述至少六个氨基酸残基还可以是AA4、AA7、AA8、AA9、AA10和AA11。此类所述至少六个氨基酸残基可以是 (a) AA5、AA6、AA7、AA8、AA9和AA10或 (b) AA5、AA6、AA7、AA8、AA9和AA11。所述至少六个氨基酸残基还可以是AA5、AA6、AA7、AA8、AA10和AA11。所述至少六个氨基酸残基还可以是AA5、AA6、AA8、AA9、AA10和AA11。所述至少六个氨基酸残基还可以是AA5、AA7、AA8、AA9、AA10和AA11。所述至少六个氨基酸残基还可以是AA6、AA7、AA8、AA9、AA10和AA11。

[0070] 此类所述至少七个氨基酸残基可以是 (a) AA1、AA2、AA3、AA4、AA5、AA6和AA7、(b) AA1、AA2、AA3、AA4、AA5、AA6和AA8、(c) AA1、AA2、AA3、AA4、AA5、AA6和AA9、(d) AA1、AA2、AA3、AA4、AA5、AA6和AA10或 (e) AA1、AA2、AA3、AA4、AA5、AA6和AA11。作为又一种备选,所述至少七个氨基酸残基可以是 (a) AA1、AA2、AA3、AA4、AA5、AA7和AA8、(b) AA1、AA2、AA3、AA4、AA5、AA7和AA9、(c) AA1、AA2、AA3、AA4、AA5、AA7和AA10或 (d) AA1、AA2、AA3、AA4、AA5、AA7和AA11。所述

[illegible]

[illegible]

[illegible]

AA6、AA7、AA8和AA9、(b) AA1、AA4、AA5、AA6、AA7、AA8和AA10或(c) AA1、AA4、AA5、AA6、AA7、AA8和AA11。所述至少七个氨基酸残基还可以进一步可以是(a) AA1、AA4、AA5、AA6、AA7、AA9和AA10或(b) AA1、AA4、AA5、AA6、AA7、AA9和AA11。所述至少七个氨基酸残基还可以是AA1、AA4、AA5、AA6、AA7、AA10和AA11。此类所述至少七个氨基酸残基可以是(a) AA1、AA4、AA5、AA6、AA8、AA9和AA10或(b) AA1、AA4、AA5、AA6、AA8、AA9和AA11。所述至少七个氨基酸残基还可以是AA1、AA4、AA5、AA6、AA8、AA10和AA11。所述至少七个氨基酸残基还可以是AA1、AA4、AA5、AA6、AA9、AA10和AA11。此类所述至少七个氨基酸残基可以是(a) AA1、AA4、AA5、AA7、AA8、AA9和AA10或(b) AA1、AA4、AA5、AA7、AA8、AA9和AA11。所述至少七个氨基酸残基还可以是AA1、AA4、AA5、AA7、AA9、AA10和AA11。所述至少七个氨基酸残基还可以是AA1、AA4、AA5、AA8、AA9、AA10和AA11。此类所述至少七个氨基酸残基可以是(a) AA1、AA4、AA6、AA7、AA8、AA9和AA10或(b) AA1、AA4、AA6、AA7、AA8、AA9和AA11。所述至少七个氨基酸残基还可以是AA1、AA4、AA6、AA7、AA8、AA10和AA11。所述至少七个氨基酸残基还可以是AA1、AA4、AA6、AA7、AA9、AA10和AA11。所述至少七个氨基酸残基还可以是AA1、AA4、AA6、AA8、AA9、AA10和AA11。所述至少七个氨基酸残基还可以是AA1、AA4、AA6、AA7、AA8、AA9、AA10和AA11。此类所述至少七个氨基酸残基可以是(a) AA1、AA5、AA6、AA7、AA8、AA9和AA10或(b) AA1、AA5、AA6、AA7、AA8、AA9和AA11。所述至少七个氨基酸残基还可以是AA1、AA5、AA6、AA7、AA8、AA10和AA11。所述至少七个氨基酸残基还可以是AA1、AA5、AA6、AA7、AA9、AA10和AA11。所述至少七个氨基酸残基还可以是AA1、AA5、AA6、AA8、AA9、AA10和AA11。所述至少七个氨基酸残基还可以是AA1、AA5、AA7、AA8、AA9、AA10和AA11。所述至少七个氨基酸残基还可以是AA1、AA6、AA7、AA8、AA9、AA10和AA11。

[0071] 此类所述至少七个氨基酸残基可以是 (a) AA2、AA3、AA4、AA5、AA6、AA7和AA8、(b) AA2、AA3、AA4、AA5、AA6、AA7和AA9、(c) AA2、AA3、AA4、AA5、AA6、AA7和AA10或 (d) AA2、AA3、AA4、AA5、AA6、AA7和AA11。所述至少七个氨基酸残基还可以是 (a) AA2、AA3、AA4、AA5、AA6、AA8和AA9、(b) AA2、AA3、AA4、AA5、AA6、AA8和AA10或 (c) AA2、AA3、AA4、AA5、AA6、AA8和AA11。所述至少七个氨基酸残基还进一步可以是 (a) AA2、AA3、AA4、AA5、AA6、AA9和AA10或 (b) AA2、AA3、AA4、AA5、AA6、AA9和AA11。所述至少七个氨基酸残基还可以是AA2、AA3、AA4、AA5、AA6、AA10和AA11。此类所述至少七个氨基酸残基可以是 (a) AA2、AA3、AA4、AA5、AA7、AA8和AA9、(b) AA2、AA3、AA4、AA5、AA7、AA8和AA10或 (c) AA2、AA3、AA4、AA5、AA7、AA8和AA11。所述至少七个氨基酸残基还进一步可以是 (a) AA2、AA3、AA4、AA5、AA7、AA9和AA10或 (b) AA2、AA3、AA4、AA5、AA7、AA9和AA11。所述至少七个氨基酸残基还可以是AA2、AA3、AA4、AA5、AA7、AA10和AA11。此类所述至少七个氨基酸残基可以是 (a) AA2、AA3、AA4、AA5、AA8、AA9和AA10或 (b) AA2、AA3、AA4、AA5、AA8、AA9和AA11。所述至少七个氨基酸残基还可以是AA2、AA3、AA4、AA5、AA8、AA10和AA11。所述至少七个氨基酸残基还可以是AA2、AA3、AA4、AA5、AA9、AA10和AA11。此类所述至少七个氨基酸残基可以是 (a) AA2、AA3、AA4、AA6、AA7、AA8和AA9、(b) AA2、AA3、AA4、AA6、AA7、AA8和AA10或 (c) AA2、AA3、AA4、AA6、AA7、AA8和AA11。所述至少七个氨基酸残基还进一步可以是 (a) AA2、AA3、AA4、AA6、AA7、AA9和AA10或 (b) AA2、AA3、AA4、AA6、AA7、AA9和AA11。所述至少七个氨基酸残基还可以是AA2、AA3、AA4、AA6、AA7、AA10和AA11。此类所述至少七个氨基酸残基可以是 (a) AA2、AA3、AA4、AA6、AA8、AA9和AA10或 (b) AA2、AA3、AA4、AA6、AA8、AA9和AA11。所述至少七个氨基酸残基还可以是AA2、AA3、AA4、AA6、AA8、AA10和AA11。所

[illegible]

AA11。所述至少七个氨基酸残基还可以是AA2、AA6、AA7、AA8、AA9、AA10和AA11。

[0072] 此类所述至少七个氨基酸残基可以是 (a) AA3、AA4、AA5、AA6、AA7、AA8和AA9、(b) AA3、AA4、AA5、AA6、AA7、AA8和AA10或 (c) AA3、AA4、AA5、AA6、AA7、AA8和AA11。所述至少七个氨基酸残基还进一步可以是 (a) AA3、AA4、AA5、AA6、AA7、AA9和AA10或 (b) AA3、AA4、AA5、AA6、AA7、AA9和AA11。所述至少七个氨基酸残基还可以是AA3、AA4、AA5、AA6、AA7、AA10和AA11。此类所述至少七个氨基酸残基可以是 (a) AA3、AA4、AA5、AA6、AA8、AA9和AA10或 (b) AA3、AA4、AA5、AA6、AA8、AA9和AA11。所述至少七个氨基酸残基还可以是AA3、AA4、AA5、AA6、AA8、AA10和AA11。所述至少七个氨基酸残基还可以是AA3、AA4、AA5、AA6、AA9、AA10和AA11。此类所述至少七个氨基酸残基可以是 (a) AA3、AA4、AA5、AA7、AA8、AA9和AA10或 (b) AA3、AA4、AA5、AA7、AA8、AA9和AA11。所述至少七个氨基酸残基还可以是AA3、AA4、AA5、AA7、AA8、AA10和AA11。所述至少七个氨基酸残基还可以是AA3、AA4、AA5、AA7、AA9、AA10和AA11。所述至少七个氨基酸残基还可以是AA3、AA4、AA5、AA8、AA9、AA10和AA11。此类所述至少七个氨基酸残基可以是 (a) AA3、AA4、AA6、AA7、AA8、AA9和AA10或 (b) AA3、AA4、AA6、AA7、AA8、AA9和AA11。所述至少七个氨基酸残基还可以是AA3、AA4、AA6、AA7、AA8、AA10和AA11。所述至少七个氨基酸残基还可以是AA3、AA4、AA6、AA7、AA9、AA10和AA11。所述至少七个氨基酸残基还可以是AA3、AA4、AA6、AA8、AA9、AA10和AA11。所述至少七个氨基酸残基还可以是AA3、AA4、AA7、AA8、AA9、AA10和AA11。此类所述至少七个氨基酸残基可以是 (a) AA3、AA5、AA6、AA7、AA8、AA9和AA10或 (b) AA3、AA5、AA6、AA7、AA8、AA9和AA11。所述至少七个氨基酸残基还可以是AA3、AA5、AA6、AA7、AA9、AA10和AA11。所述至少七个氨基酸残基还可以是AA3、AA5、AA6、AA8、AA9、AA10和AA11。所述至少七个氨基酸残基还可以是AA3、AA5、AA7、AA8、AA9、AA10和AA11。所述至少七个氨基酸残基还可以是AA3、AA5、AA7、AA8、AA9、AA10和AA11。所述至少七个氨基酸残基还可以是AA3、AA5、AA6、AA7、AA9、AA10和AA11。所述至少七个氨基酸残基还可以是AA3、AA5、AA6、AA8、AA9、AA10和AA11。所述至少七个氨基酸残基还可以是AA4、AA5、AA6、AA7、AA8、AA9和AA10或 (b) AA4、AA5、AA6、AA7、AA8、AA9和AA11。所述至少七个氨基酸残基还可以是AA4、AA5、AA6、AA7、AA8、AA10和AA11。所述至少七个氨基酸残基还可以是AA4、AA5、AA6、AA7、AA9、AA10和AA11。所述至少七个氨基酸残基还可以是AA4、AA5、AA6、AA8、AA9、AA10和AA11。所述至少七个氨基酸残基还可以是AA4、AA5、AA7、AA8、AA9、AA10和AA11。所述至少七个氨基酸残基还可以是AA4、AA6、AA7、AA8、AA9、AA10和AA11。所述至少七个氨基酸残基还可以是AA5、AA6、AA7、AA8、AA9、AA10和AA11。

[0073] 此类所述至少八个氨基酸残基可以是 (a) AA1、AA2、AA3、AA4、AA5、AA6、AA7和AA8、(b) AA1、AA2、AA3、AA4、AA5、AA6、AA7和AA9、(c) AA1、AA2、AA3、AA4、AA5、AA6、AA7和AA10或 (d) AA1、AA2、AA3、AA4、AA5、AA6、AA7和AA11。所述至少八个氨基酸残基还可以是 (a) AA1、AA2、AA3、AA4、AA5、AA6、AA8和AA9、(b) AA1、AA2、AA3、AA4、AA5、AA6、AA8和AA10或 (c) AA1、AA2、AA3、AA4、AA5、AA6、AA8和AA11。所述至少八个氨基酸残基还进一步可以是 (a) AA1、AA2、AA3、AA4、AA5、AA6、AA9和AA10或 (b) AA1、AA2、AA3、AA4、AA5、AA6、AA9和AA11。所述至少八个氨基酸残基还可以是AA1、AA2、AA3、AA4、AA5、AA6、AA10和AA11。此类所述至少八个氨基酸残基可以是 (a) AA1、AA2、AA3、AA4、AA5、AA7、AA8和AA9、(b) AA1、AA2、AA3、AA4、AA5、AA7、AA8和AA10或 (c) AA1、AA2、AA3、AA4、AA5、AA7、AA8和AA11。所述至少八个氨基酸残基还进一步可以是 (a) AA1、AA2、AA3、AA4、AA5、AA7、AA9和AA10或 (b) AA1、AA2、AA3、AA4、AA5、AA7、AA9和AA11。所述至少八个氨基酸残基还可以是AA1、AA2、AA3、AA4、AA5、AA7、AA10和AA11。此类所述至少八

[illegible]

[illegible]

至少八个氨基酸残基还可以是AA1、AA4、AA5、AA7、AA8、AA9、AA10和AA11。所述至少八个氨基酸残基还可以是AA1、AA4、AA6、AA7、AA8、AA9、AA10和AA11。所述至少八个氨基酸残基还可以是AA1、AA5、AA6、AA7、AA8、AA9、AA10和AA11。

[illegible]

[0075] 此类所述至少八个氨基酸残基可以是 (a) AA3、AA4、AA5、AA6、AA7、AA8、AA9和AA10或 (b) AA3、AA4、AA5、AA6、AA7、AA8、AA9和AA11。所述至少八个氨基酸残基还可以是AA3、AA4、AA5、AA6、AA7、AA8、AA10和AA11。所述至少八个氨基酸残基还可以是AA3、AA4、AA5、AA6、AA7、AA9、AA10和AA11。所述至少八个氨基酸残基还可以是AA3、AA4、AA5、AA6、AA8、AA9、AA10和AA11。所述至少八个氨基酸残基还可以是AA3、AA4、AA5、AA7、AA8、AA9、AA10和AA11。所述至少八个氨基酸残基还可以是AA3、AA4、AA6、AA7、AA8、AA9、AA10和AA11。所述至少八个氨基酸残基还可以是AA3、AA5、AA6、AA7、AA8、AA9、AA10和AA11。所述至少八个氨基酸残基还可以是

AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11。

[illegible]

AA11。

[0077] 此类所述至少九个氨基酸残基可以是(a) AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9和AA10或(b) AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9和AA11。所述至少九个氨基酸残基还可以是AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA10和AA11。所述至少九个氨基酸残基还可以是AA2、AA3、AA4、AA5、AA6、AA7、AA9、AA10和AA11。所述至少九个氨基酸残基还可以是AA2、AA3、AA4、AA5、AA6、AA8、AA9、AA10和AA11。所述至少九个氨基酸残基还可以是AA2、AA3、AA4、AA5、AA7、AA8、AA9、AA10和AA11。所述至少九个氨基酸残基还可以是AA2、AA3、AA4、AA6、AA7、AA8、AA9、AA10和AA11。所述至少九个氨基酸残基还可以是AA2、AA3、AA5、AA6、AA7、AA8、AA9、AA10和AA11。所述至少九个氨基酸残基还可以是AA2、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11。所述至少九个氨基酸残基还可以是AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11。

[0078] 此类所述至少十个氨基酸残基可以进一步是(a) AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9和AA10、(b) AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9和AA11,或(c) AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA10和AA11。此类所述至少十个氨基酸残基甚至可以是AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA9、AA10和AA11。此类所述至少十个氨基酸残基还可以是(a) AA1、AA2、AA3、AA4、AA5、AA6、AA8、AA9、AA10和AA11、(b) AA1、AA2、AA3、AA4、AA5、AA7、AA8、AA9、AA10和AA11、(c) AA1、AA2、AA3、AA4、AA6、AA7、AA8、AA9、AA10和AA11,或(d) AA1、AA2、AA3、AA5、AA6、AA7、AA8、AA9、AA10和AA11。此类所述至少十个氨基酸残基还可以是(a) AA1、AA2、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11、(b) AA1、AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11,或(c) AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11。

Rca蛋白和核酸

[0079] 在许多物种包括稻中,存在基于前体mRNA可变剪接成较大 α 同种型和较短 β 同种型而编码两个蛋白质同种型的单一Rca基因^{18,19}。类似地,小麦存在 α 和 β 同种型。此外,小麦具有在小麦第四染色体上与第一Rca基因紧邻的独立Rca基因,其编码 β 同种型的第二变体。后者称作形式1并且可变剪接形式称作形式2Rca¹¹。因此,在小麦中——除了三个小麦染色体亚基因组上各自几乎相同副本外——存在三种可以表达的不同Rca基因产物:形式1 β (TaRca1- β)、形式2 α (TaRca2- α) 和形式2 β (TaRca2- β)。

[0080] Rca蛋白是可以形成六聚体蛋白质复合体并与Rubisco相互作用的AAA+伴侣蛋白。有功能的Rca蛋白包含中央ATP酶结构域(AAA+模块)和涉及Rubisco-Rca和Rca-Rca相互作用的C末端结构域³³⁻³⁵。AAA+模块位于从等同于SEQ ID NO:2上位置57的氨基酸位置至等同于SEQ ID NO:2上位置345的氨基酸位置,ATP酶核心位于从等同于SEQ ID NO:2上位置182的氨基酸位置至等同于SEQ ID NO:2上位置282的氨基酸位置,并且C末端结构域位于从等同于SEQ ID NO:2上位置346的氨基酸位置至等同于SEQ ID NO:2上位置427的氨基酸位置。更具体地,有功能的Rca蛋白进一步包含在等同SEQ ID NO:2上位置123和124的氨基酸位置处的N-接头(IA)、在等同SEQ ID NO:2上位置155至161的氨基酸位置处的Walker A基序(GxxxxGK)、在等同SEQ ID NO:2上位置215至219的氨基酸位置处的Walker B基序(LxxxD)、在等同SEQ ID NO:2上位置253至265的氨基酸位置处的Rubisco相互作用环(Shivhare等人,2017)、在等同SEQ ID NO:2上位置339至347的氨基酸位置处的Rca-Rca界面(Stotz等人2011,Nature Structural and Molecular Biology 18:1366-1370)和在等同SEQ ID NO:2上位置406的氨基酸位置处的酪氨酸(Y)。

[0081] 向包含Rubisco活化酶和Rubisco蛋白的蛋白质复合体赋予热稳定性意味着升高中点温度,在所述中点温度Rca对Rubisco的活化速度被降低一半(T_m ,参见实施例2)。所述蛋白质复合体被赋予的热稳定性可以是中点温度升高约1℃、约2℃、约3℃、约4℃、约5℃、约6℃、约7℃、约8℃、约9℃、或约10℃。它也可以是中点温度升高至少约1℃、至少约2℃、至少约3℃、至少约4℃、至少约5℃、至少约6℃、至少约7℃、至少约8℃、至少约9℃、或至少约10℃。它也可以是中点温度升高约1℃至4℃、约1℃至5℃、约1℃至6℃、约1℃至7℃、约1℃至8℃、约1℃至9℃、约1℃至10℃、约2℃至5℃、约2℃至6℃、约2℃至7℃、约2℃至8℃、约2℃至9℃、约2℃至10℃、约3℃至6℃、约3℃至7℃、约3℃至8℃、约3℃至9℃、约3℃至10℃、约4℃至7℃、约4℃至8℃、约4℃至9℃、约4℃至10℃、约5℃至8℃、约5℃至9℃、约5℃至10℃、约6℃至9℃、约6℃至10℃或约7℃至10℃。

[0082] 这种热稳定性复合体的中点温度(T_m ,即Rca对Rubisco的活化速度被降低一半时的温度)可以是至少约或约37℃、至少约或约38℃、至少约或约39℃、至少约或约40℃、至少约或约41℃、至少约或约42℃、至少约或约43℃、至少约或约44℃、或至少约或约45℃。它也可以是约37℃至约40℃、约37℃至约42℃、约37℃至约45℃、约39℃至约41℃、约39℃至约43℃、或约39℃至约45℃。

[0083] 这种热稳定性复合体可以在热胁迫条件下具有酶活性。

[0084] 包含Rubisco活化酶和Rubisco蛋白的复合体可以在体外或体内形成。例如,可以通过使本发明的热稳定性Rca 2或如本文所述的小麦Rca1- β 与来自小麦的存在于叶提取物中的Rubisco或已经从小麦纯化的Rubisco接触,在体外形成复合体。备选地,可以通过在植物中表达本发明的Rca多肽,由此与植物的内源Rubisco形成复合体,而在体内形成所述复合体。

[0085] 本领域已知的任何方法均可以用来分析包含Rca多肽的复合体的活性(包括但不限于,Rubisco活化和ATP水解,参见例如Chakrabarti等人2002J.Biochem.Biophys.Methods 52:179-187及McC Lilley和Portis 1997Plant Physiol.114:605-613)。

[0086] “热稳定性Rca蛋白”是能够向包含Rubisco活化酶和Rubisco蛋白的蛋白质复合体赋予热稳定性的Rca蛋白。热稳定性Rca蛋白也可以是Rca解折叠的中点温度较高的Rca蛋白。

[0087] 热稳定性Rca蛋白解折叠的中点温度可以高于热不稳定性Rca蛋白解折叠的中点温度至少约或约1℃、至少约或约2℃、至少约或约3℃、至少约或约4℃、至少约或约5℃、至少约或约6℃、至少约或约7℃、至少约或约8℃、至少约或约9℃、或至少约或约10℃。它也可以比热不稳定性Rca蛋白解折叠的中点温度高约1℃至4℃、约1℃至5℃、约1℃至6℃、约1℃至7℃、约1℃至8℃、约1℃至9℃、约1℃至10℃、约2℃至5℃、约2℃至6℃、约2℃至7℃、约2℃至8℃、约2℃至9℃、约2℃至10℃、约3℃至6℃、约3℃至7℃、约3℃至8℃、约3℃至9℃、约3℃至10℃、约4℃至7℃、约4℃至8℃、约4℃至9℃、约4℃至10℃、约5℃至8℃、约5℃至9℃、约5℃至10℃、约6℃至9℃、约6℃至10℃、或约7℃至10℃。

[0088] 热稳定性蛋白质解折叠的中点温度可以是至少约或约35℃、至少约或约36℃、至少约或约37℃、至少约或约38℃、至少约或约39℃、至少约或约40℃、至少约或约41℃、至少约或约42℃、或至少约或约43℃。它也可以是约35℃与约38℃之间、约35℃与约40℃之间、

约35℃与约43℃之间、约37℃与约39℃之间、约37℃与约41℃之间、或约37℃与约43℃之间。

Rca1-β

[0089] 本文中显示Rca1-β蛋白向包含这种Rubisco活化酶和Rubisco蛋白的蛋白质复合物赋予热稳定性。

[0090] 本文描述并使用的Rca 1β蛋白及其变体包含氨基酸序列,所述氨基酸序列选自(a) SEQ ID N0s:8、47或49的氨基酸序列;和(b)与SEQ ID N0s:8、47或49的氨基酸序列具有90%同一性且包含选自以下的至少一个氨基酸的氨基酸序列:(i)与SEQ ID N0:8的位置109相对应的位置处的异亮氨酸,(ii)与SEQ ID N0:8的位置123相对应的位置处的天冬氨酸,(iii)与SEQ ID N0:8的位置210相对应的位置处的异亮氨酸,(iv)与SEQ ID N0:8的位置315相对应的位置处的精氨酸,(v)与SEQ ID N0:8的位置320相对应的位置处的脯氨酸,(vi)与SEQ ID N0:8的位置327相对应的位置处的亮氨酸,(vii)与SEQ ID N0:8的位置357相对应的位置处的谷氨酸,(viii)与SEQ ID N0:8的位置384相对应的位置处的异亮氨酸,(ix)与SEQ ID N0:8的位置409相对应的位置处的赖氨酸,(x)与SEQ ID N0:8的位置411相对应的位置处的亮氨酸和(xi)与SEQ ID N0:8的位置413相对应的位置处的谷氨酸。

[0091] SEQ ID N0:8代表来自小麦亚基因组B的Rca 1β蛋白的氨基酸序列,SEQ ID N0:47代表来自小麦亚基因组A的Rca 1β蛋白的氨基酸序列并且SEQ ID N0:49代表来自小麦亚基因组D的Rca 1β蛋白的氨基酸序列。

[0092] 以下Rca 1β蛋白的氨基酸序列适用于本发明,所述氨基酸序列包含与本文中所述的蛋白质具有至少80%、或至少85%、或至少90%、或至少91%、或至少92%、或至少93%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少100%序列同一性的氨基酸序列,并且也称作变体。相对于本发明的SEQ ID N0:8氨基酸序列而言,术语“变体”意指基本上相似的序列。SEQ ID N0:47和SEQ ID N0:49的氨基酸序列是SEQ ID N0:8的氨基酸序列的变体。

[0093] 另外,显而易见,其中已经缺失、置换或插入一个或多个氨基酸残基的Rca 1β蛋白变体也可以在本发明的方法中使用以到达相同的效果,前提是中央ATP酶结构域(AAA+模块)、C末端结构域、N-接头、Walker A、Walker B基序、Rubisco相互作用环、Rca-Rca界面、和在等同SEQ ID N0:2上位置406的氨基酸位置处的酪氨酸(Y)不受所述氨基酸缺失、置换或插入影响。

[0094] 这种Rca 1β蛋白变体还可以包含至少一个、或至少两个、或至少三个、或至少四个、或至少五个、或至少六个、或至少七个、或至少八个、或至少九个、或至少十个、或全部十个鉴定为在来自各个植物物种的Rca蛋白的热稳定性变体中保守的上文所列氨基酸残基。上文列出了至少两个、或至少三个、或至少四个、或至少五个、或至少六个、或至少七个、或至少八个、或至少九个、或至少十个这类氨基酸残基的可能组合。

[0095] 另外,编码这种Rca 1β蛋白及其变体的核酸可以包含选自以下的编码性核酸序列:(a) SEQ ID N0s:7、46或48的核酸或其互补物和(b)与SEQ ID N0s:7、46或48的核酸或其互补物具有60%同一性的核酸。

[0096] SEQ ID N0:7代表来自亚基因组B的小麦Rca 1β基因的编码性核苷酸序列,SEQ ID N0:46代表来自亚基因组A的小麦Rca 1β基因的编码性核苷酸序列,并且SEQ ID N0:48代表

来自亚基因组D的小麦Rca 1 β 基因的编码性核苷酸序列。

[0097] 编码Rca 1 β 蛋白的如下核酸适于本发明,所述核酸包含与本文中所述的基因具有至少60%、或至少70%、或至少80%、或至少85%、或至少90%、或至少95%、或至少98%序列同一性的核苷酸序列并且也称作变体。

[0098] 包含与SEQ ID N0s:7、46或48具有60%序列同一性的核苷酸序列的核酸因此可以是包含与SEQ ID N0s:7、46或48分别具有至少60%、或至少70%、或至少80%、或至少85%、或至少90%、或至少95%、或至少98%、或100%序列同一性的核苷酸序列的核酸。

内源热不稳定性Rca 2

[0099] 与Rca 1 β 相反,已知Rca 2内源蛋白是热不稳定的。如本文所述的,内源热不稳定性Rca 2蛋白质可以包含SEQ ID N0s:2、6、39、41、43或45的氨基酸序列。内源热不稳定性Rca 2蛋白还可以包含氨基酸序列,所述氨基酸序列与SEQ ID N0s:2、6、39、41、43或45的氨基酸序列具有至少90%序列同一性且不编码选自以下的氨基酸:(a)与SEQ ID N0:4的位置59相对应的位置处的异亮氨酸,(b)与SEQ ID N0:4的位置73相对应的位置处的天冬氨酸,(c)与SEQ ID N0:4的位置160相对应的位置处的异亮氨酸,(d)与SEQ ID N0:4的位置265相对应的位置处的精氨酸,(e)与SEQ ID N0:4的位置270相对应的位置处的脯氨酸,(f)与SEQ ID N0:4的位置277相对应的位置处的亮氨酸,(g)与SEQ ID N0:4的位置307相对应的位置处的谷氨酸,(h)与SEQ ID N0:4的位置334相对应的位置处的异亮氨酸,(i)与SEQ ID N0:4的位置359相对应的位置处的赖氨酸,(j)与SEQ ID N0:4的位置361相对应的位置处的亮氨酸和(k)与SEQ ID N0:4的位置363相对应的位置处的谷氨酸。

[0100] 热不稳定性内源Rca 2蛋白因此还可以包含与SEQ ID N0s:2、6、39、41、43或45的氨基酸序列具有至少90%序列同一性并且不包含氨基酸AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11的氨基酸序列。

[0101] 热不稳定性内源Rca 2蛋白可以包含与本文中所述的蛋白质具有至少80%、或至少85%、或至少90%、或至少91%、或至少92%、或至少93%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%或至少100%序列同一性的氨基酸序列,所述氨基酸序列不包含氨基酸AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11。

[0102] SEQ ID N0s:2和6分别代表来自亚基因组B的小麦Rca 2 β 和2 α 蛋白的氨基酸序列。SEQ ID N0s:39和43分别代表来自亚基因组A的小麦Rca 2 β 和2 α 蛋白的氨基酸序列。SEQ ID N0s:41和45分别代表来自亚基因组D的小麦Rca 2 β 和2 α 蛋白的氨基酸序列。SEQ ID N0:4代表来自亚基因组B的小麦Rca 2 β 的氨基酸酸序列,不包括叶绿体靶向肽。

[0103] 另外,编码所述热不稳定性Rca 2蛋白的内源热不稳定性Rca 2基因可以包含SEQ ID N0s:1、5、38、40、42或44的编码性核苷酸序列。

[0104] SEQ ID N0s:1和5分别代表来自亚基因组B的小麦Rca 2 β 和2 α 基因的编码性核苷酸序列。SEQ ID N0s:38和42分别代表来自亚基因组A的小麦Rca 2 β 和2 α 基因的编码性核苷酸序列。SEQ ID N0s:40和44分别代表来自亚基因组D的小麦Rca 2 β 和2 α 基因的编码性核苷酸序列。

[0105] 编码所述热不稳定性Rca 2蛋白的内源热不稳定性Rca 2基因可以包含SEQ ID N0s:1、5、38、40、42或44的编码性核苷酸序列或这样的编码性核苷酸序列,其与SEQ ID N0:1、5、38、40、42或44的核苷酸序列具有至少60%同一性并且不编码选自以下的氨基酸:(a)

与SEQ ID NO:4的位置59相对应的位置处的异亮氨酸, (b) 与SEQ ID NO:4的位置73相对应的位置处的天冬氨酸, (c) 与SEQ ID NO:4的位置160相对应的位置处的异亮氨酸, (d) 与SEQ ID NO:4的位置265相对应的位置处的精氨酸, (e) 与SEQ ID NO:4的位置270相对应的位置处的脯氨酸, (f) 与SEQ ID NO:4的位置277相对应的位置处的亮氨酸, (g) 与SEQ ID NO:4的位置307相对应的位置处的谷氨酸, (h) 与SEQ ID NO:4的位置334相对应的位置处的异亮氨酸, (i) 与SEQ ID NO:4的位置359相对应的位置处的赖氨酸, (j) 与SEQ ID NO:4的位置361相对应的位置处的亮氨酸和 (k) 与SEQ ID NO:4的位置363相对应的位置处的谷氨酸。

[0106] 编码所述热不稳定性Rca 2蛋白的内源热不稳定性Rca 2基因因此还可以包含与SEQ ID NOs:1、5、38、40、42或44的核苷酸序列具有至少60%序列同一性并且不编码氨基酸AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11的核苷酸序列。

[0107] 包含与SEQ ID NOs:1、5、38、40、42或44具有60%序列同一性并且不编码氨基酸AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11的核苷酸序列的核酸可以是包含与SEQ ID NOs:1、5、38、40、42或44分别具有至少60%、或至少70%、或至少80%、或至少85%、或至少90%、或至少95%、或至少98%、或100%序列同一性并且不编码氨基酸AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11的核苷酸序列的核酸。

Rca 2敲除等位基因

[0108] 在另一个实施方案中,提供Rca 2基因的敲除等位基因。在又一个实施方案中,Rca 2基因是来自小麦亚基因组B、A或D的Rca 2 β 基因或来自小麦亚基因组B、A或D的Rca 2 α 基因。

[0109] Rca 2基因的敲除等位基因可以是完全敲除等位基因或部分敲除等位基因。

[0110] 如本文所用,“完全敲除”或“无效”等位基因指一种突变等位基因,其编码与相应的野生型有功能蛋白相比没有生物学活性的蛋白质,或根本不编码蛋白。这种“完全敲除突变等位基因”例如可以是在其核酸序列中包含一个或多个突变的野生型等位基因,所述突变为例如,一个或多个无义、错义、插入、缺失、移码或突变的剪接位点突变。特别地,这种完全敲除突变Rca 2等位基因可以是包含突变的野生型Rca 2等位基因,其中所述突变优选地导致产生Rca 2蛋白,所述蛋白缺少至少一个功能结构域或基序,如中央ATP酶结构域(AAA+模块)、C末端结构域、N-接头、Walker A、Walker B基序、Rubisco相互作用环、Rca-Rca界面,或缺少至少一个对其功能关键的氨基酸,如在等同SEQ ID NO:2上位置406的氨基酸位置处的酪氨酸(Y),从而Rca 2蛋白的生物学活性彻底消除,或其中所述(诸)修饰优选地导致不产生Rca 2蛋白。

[0111] 如本文所用,“部分敲除”突变等位基因指一种突变等位基因,其编码与相应的野生型有功能蛋白相比生物学活性显著降低的蛋白质。这种“部分敲除突变等位基因”例如可以是在其核酸序列中包含一个或多个突变的野生型等位基因,例如,一个或多个错义突变。尤其,这种部分敲除突变等位基因可以是包含突变的野生型等位基因,其中所述突变优选地导致产生蛋白质,在所述蛋白质中至少一个保守和/或功能性氨基酸被置换成另一氨基酸,从而生物学活性显著降低但是未彻底消除。

[0112] 如本文所用,Rca 2等位基因中的错义突变可以是Rca 2等位基因中的如下任何突变(缺失、插入或置换),借助所述突变,一个或多个改变的密码子被并入相应野生型Rca 2等位基因的编码性DNA和相应mRNA序列中,导致野生型Rca 2蛋白中的一个或多个氨基酸置

换成突变Rca 2蛋白中的一个或多个其他氨基酸。包含错义突变的突变Rca 2等位基因是其中置换一个氨基酸的Rca 2等位基因。

[0113] 如本文中所用,Rca 2等位基因中的无义突变是Rca 2等位基因中这样的突变,借助所述突变,一个或多个翻译终止密码子被引入相应野生型Rca 2等位基因的编码性DNA和相应的mRNA序列。翻译终止密码子是TGA (mRNA中是UGA)、TAA (UAA) 和TAG (UAG)。因而,在编码性序列中导致读框内终止密码子产生的任何突变(缺失、插入或替换)将导致翻译终止和氨基酸链的截短。截短的蛋白质缺少由该突变下游(即Rca 2蛋白的C端部分)的编码性DNA编码的氨基酸并保持由该突变上游(即Rca 2蛋白的N端部分)的编码性DNA编码的氨基酸。与野生型蛋白Rca 2相比,突变体Rca 2蛋白截短得越多,截短越可能导致Rca 2蛋白活性显著降低。认为,为了突变体Rca 2蛋白丧失一定的生物学活性,其应当至少不再包含在等同SEQ ID NO:2上位置406的氨基酸位置处的酪氨酸(Y)。

表1:导致读框内终止密码子生成的置换突变实例。

SEQ ID NO: 2 上的置换位置	置换之前的密码子	所得的终止密码子
79	CAG	TAG
182	TGG	TAG
183	TGG	TGA
214	CAG	TAG
217	CAG	TAG
259	CAG	TAG
319	CAG	TAG
461	TGG	TAG
462	TGG	TGA
475	CAA	TAA
490	CAG	TAG
592	CAG	TAG

694	CAG	TAG
712	CAG	TAG
763	CAG	TAG
884	TGG	TAG
885	TGG	TGA
931	CAG	TAG
988	CAA	TAA
1049	TGG	TAG
1050	TGG	TGA
1123	CAG	TAG
1174	CAG	TAG
1180	CAG	TAG
1201	CAG	TAG

[0114] 如本文所用,Rca 2等位基因中的移码突变是Rca 2等位基因中的如下突变(缺失、插入、复制等),所述突变导致在该突变下游,核酸序列在不同读框中翻译。

[0115] 如本文所用,Rca 2等位基因中的剪接位点突变是Rca 2等位基因中的如下突变(缺失、插入、置换、复制等),其中借助所述突变,剪接供体位点或剪接受体位点遭突变,导致mRNA的加工改变并因此导致编码的蛋白质改变,所述改变的蛋白质可以具有各种长度的插入、缺失、置换或可以被截短。

[0116] 如本文所用,Rca 2等位基因中的缺失突变是Rca 2等位基因的如下突变,所述突变导致产生Rca 2蛋白,所述蛋白缺少由缺失的编码性DNA编码的氨基酸并且维持由该缺失上游(即Rca 2蛋白的N端部分)的编码性DNA编码的氨基酸和由该缺失下游(即C端部分Rca 2蛋白)的编码性DNA编码的氨基酸。

[0117] “有功能的Rca 2蛋白的量明显减少”(例如有功能的Rca 2A、Rca 2B或Rca 2C蛋白)指,与不包含突变Rca 2等位基因的细胞所产生的有功能Rca 2蛋白的量相比,包含该突变Rca 2等位基因的细胞产生的有功能蛋白的量减少至少30%、40%、50%、60%、70%或80%。然而不消除有功能Rca 2蛋白的产生。该定义包括产生在体内具有降低的生物学活性的“无功能”Rca 2蛋白(例如,截短的Rca 2蛋白)、减少有功能的Rca 2蛋白的绝对量(例如,因Rca 2基因中的突变导致不产生有功能的Rca 2蛋白)、产生与有功能的野生型Rca 2蛋白的活性相比具有明显降低生物学活性的Rca 2蛋白(例如,其中对编码的Rca2蛋白的生物学活性重要的一个或多个氨基酸残基被置换成另一氨基酸残基的Rca 2蛋白)。

热稳定性Rca 2变体

[0118] 本发明的一个目的是,提供热稳定性Rca 2蛋白变体。这种热稳定性Rca 2蛋白变

体可以包含选自以下的氨基酸序列:SEQ ID NO:30或33的氨基酸序列、与SEQ ID NO:30或33的氨基酸序列具有90%同一性并且包含鉴定为在来自各个植物物种的Rca蛋白的热稳定性变体中保守的上文所列氨基酸残基中至少一者的氨基酸序列。

[0119] SEQ ID NO:30代表这样的氨基酸序列,所述氨基酸序列基于来自亚基因组B的小麦Rca 2 β 之一(即SEQ ID NO:2),但包含全部十一个鉴定为在来自各个植物物种的Rca蛋白的热稳定性变体中保守的氨基酸残基。SEQ ID NO:33代表这样的氨基酸序列,所述氨基酸序列基于来自亚基因组B的小麦Rca2 β 之一(SEQ ID NO:2),但包含十一个鉴定为在来自各个植物物种的Rca蛋白的热稳定性变体中保守的氨基酸残基中的八个。这八个氨基酸残基是(a) AA1、(b) AA3、(d) AA4、(e) AA5、(f) AA8、(g) AA9、(h) AA10和(i) AA11。

[0120] 热稳定性Rca 2蛋白变体还可以包含氨基酸序列,所述氨基酸序列选自SEQ ID NO:32或35的氨基酸序列且还包含叶绿体靶向肽,或者与SEQ ID NO:32或35的氨基酸序列具有90%同一性、还包含叶绿体靶向肽且包含选自以下的至少一个氨基酸:(i) 与SEQ ID NO:32或35的位置59相对应的位置处的异亮氨酸,(ii) 与SEQ ID NO:32或35的位置73相对应的位置处的天冬氨酸,(iii) 与SEQ ID NO:32或35的位置160相对应的位置处的异亮氨酸,(iv) 与SEQ ID NO:32或35的位置265相对应的位置处的精氨酸,(v) 与SEQ ID NO:32或35的位置270相对应的位置处的脯氨酸,(vi) 与SEQ ID NO:32或35的位置277相对应的位置处的亮氨酸,(vii) 与SEQ ID NO:32或35的位置307相对应的位置处的谷氨酸,(viii) 与SEQ ID NO:32或35的位置334相对应的位置处的异亮氨酸,(ix) 与SEQ ID NO:32或35的位置359相对应的位置处的赖氨酸,(x) 与SEQ ID NO:32或35的位置361相对应的位置处的亮氨酸和(xi) 与SEQ ID NO:32或35的位置363相对应的位置处的谷氨酸。

[0121] 可以理解,本发明的热稳定性Rca 2蛋白变体可以包含氨基酸序列,所述氨基酸序列与SEQ ID NOs:30、32、33或35的氨基酸序列具有至少90%同一性并且包含至少一个、或至少两个、或至少三个、或至少四个、或至少五个、或至少六个、或至少七个、或至少八个、或至少九个、或至少十个、或全部十一个鉴定为在来自各个植物物种的Rca蛋白的热稳定性变体中保守的本文所列氨基酸残基,AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11。

[0122] 热稳定性Rca 2蛋白变体可以包含与SEQ ID NOs:30、32、33或35的氨基酸序列具有至少80%、或至少85%、或至少90%、或至少91%、或至少92%、或至少93%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少100%序列同一性的氨基酸序列,并且包含鉴定为在来自各个植物物种的Rca蛋白的热稳定性变体中保守的且在本文中列示为AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11的氨基酸残基中的至少一者。

[0123] 热稳定性Rca 2蛋白变体可以包含与SEQ ID NOs:30、32、33或35的氨基酸序列具有至少80%、或至少85%、或至少90%、或至少91%、或至少92%、或至少93%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少100%序列同一性的氨基酸序列,并且包含鉴定为在来自各个植物物种的Rca蛋白的热稳定性变体中保守的且在本文中列示为AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11的氨基酸残基中的至少二者。

[0124] 热稳定性Rca 2蛋白变体可以包含与SEQ ID NOs:30、32、33或35的氨基酸序列具有至少80%、或至少85%、或至少90%、或至少91%、或至少92%、或至少93%、或至少94%、

或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少100%序列同一性的氨基酸序列,并且包含鉴定为在来自各个植物物种的Rca蛋白的热稳定性变体中保守的且在本文中列示为AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11的氨基酸残基中的至少三者。

[0125] 热稳定性Rca 2蛋白变体可以包含与SEQ ID N0s:30、32、33或35的氨基酸序列具有至少80%、或至少85%、或至少90%、或至少91%、或至少92%、或至少93%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少100%序列同一性的氨基酸序列,并且包含鉴定为在来自各个植物物种的Rca蛋白的热稳定性变体中保守的且在本文中列示为AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11的氨基酸残基中的至少四者。

[0126] 热稳定性Rca 2蛋白变体可以包含与SEQ ID N0s:30、32、33或35的氨基酸序列具有至少80%、或至少85%、或至少90%、或至少91%、或至少92%、或至少93%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少100%序列同一性的氨基酸序列,并且包含鉴定为在来自各个植物物种的Rca蛋白的热稳定性变体中保守的且在本文中列示为AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11的氨基酸残基中的至少五者。

[0127] 热稳定性Rca 2蛋白变体可以包含与SEQ ID N0s:30、32、33或35的氨基酸序列具有至少80%、或至少85%、或至少90%、或至少91%、或至少92%、或至少93%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少100%序列同一性的氨基酸序列,并且包含鉴定为在来自各个植物物种的Rca蛋白的热稳定性变体中保守的且在本文中列示为AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11的氨基酸残基中的至少六者。

[0128] 热稳定性Rca 2蛋白变体可以包含与SEQ ID N0s:30、32、33或35的氨基酸序列具有至少80%、或至少85%、或至少90%、或至少91%、或至少92%、或至少93%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少100%序列同一性的氨基酸序列,并且包含鉴定为在来自各个植物物种的Rca蛋白的热稳定性变体中保守的且在本文中列示为AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11的氨基酸残基中的至少七者。

[0129] 热稳定性Rca 2蛋白变体可以包含与SEQ ID N0s:30、32、33或35的氨基酸序列具有至少80%、或至少85%、或至少90%、或至少91%、或至少92%、或至少93%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少100%序列同一性的氨基酸序列,并且包含鉴定为在来自各个植物物种的Rca蛋白的热稳定性变体中保守的且在本文中列示为AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11的氨基酸残基中的至少八者。

[0130] 热稳定性Rca 2蛋白变体可以包含与SEQ ID N0s:30、32、33或35的氨基酸序列具有至少80%、或至少85%、或至少90%、或至少91%、或至少92%、或至少93%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少100%序列同一性的氨基酸序列,并且包含鉴定为在来自各个植物物种的Rca蛋白的热稳定性变体中保守的且在本文中列示为AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11的氨基酸残基中

的至少九者。

[0131] 热稳定性Rca 2蛋白变体可以包含与SEQ ID NOs:30、32、33或35的氨基酸序列具有至少80%、或至少85%、或至少90%、或至少91%、或至少92%、或至少93%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少100%序列同一性的氨基酸序列,并且包含鉴定为在来自各个植物物种的Rca蛋白的热稳定性变体中保守的且在本文中列示为AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11的氨基酸残基中的至少十者。

[0132] 此外,显而易见,其中已经缺失、置换或插入一个或多个氨基酸残基的Rca 2蛋白的热稳定性变体也可以在本发明的方法中用于相同的效果,前提是中央ATP酶结构域(AAA+模块)、C末端结构域、N-接头、Walker A、Walker B基序、Rubisco相互作用环、Rca-Rca界面、和在等同SEQ ID NO:2上位置406的氨基酸位置处的酪氨酸(Y)不受所述氨基酸缺失、置换或插入影响。

[0133] 能够实现本发明Rca蛋白亚细胞靶向定位至叶绿体的叶绿体靶向肽适于本发明。叶绿体转运肽、叶绿体靶向序列和基质靶向转运肽是等同术语。叶绿体靶向肽基于以下三个结构域的存在可识别:始于甲硫氨酸、后继丙氨酸并且止于甘氨酸或脯氨酸的约10个残基的不带电荷N端结构域、缺少酸性残基但富含丝氨酸和苏氨酸的中央结构域、和富含精氨酸并形成两亲性 β 链的C端结构域(Bruce,2000,trends in cell biology,Vol 10,440-447)。

[0134] 这类叶绿体靶向肽在本文中鉴定为SEQ ID NOs:8、47或49的位置1至位置47的氨基酸序列、SEQ ID NOs:2、6、30、33、39、41、43或45的位置1至位置46的氨基酸序列、SEQ ID NO:11的位置1至位置46的氨基酸序列、SEQ ID NO:12的位置1至位置46的氨基酸序列、SEQ ID NO:13的位置1至位置55的氨基酸序列、SEQ ID NO:14的位置1至位置56的氨基酸序列、SEQ ID NO:15的位置1至位置57的氨基酸序列、SEQ ID NO:16的位置1至位置50的氨基酸序列、SEQ ID NO:17的位置1至位置57的氨基酸序列、SEQ ID NO:20的位置1至位置57的氨基酸序列、SEQ ID NO:21的位置1至位置58的氨基酸序列、SEQ ID NO:22的位置1至位置59的氨基酸序列、SEQ ID NO:23的位置1至位置56的氨基酸序列、SEQ ID NO:24的位置1至位置51的氨基酸序列、SEQ ID NO:25的位置1至位置66的氨基酸序列、SEQ ID NO:26的位置1至位置57的氨基酸序列、SEQ ID NO:27的位置1至位置55的氨基酸序列、SEQ ID NO:28的位置1至位置74的氨基酸序列。

[0135] 还适于本发明的是具有如下氨基酸序列的叶绿体靶向肽,所述氨基酸序列与SEQ ID NO:30从位置1至位置46的氨基酸序列或SEQ ID NO:8从位置1至位置47的氨基酸序列具有至少70%、至少80%、至少85%、至少90%、至少92%、至少95%、至少98%、至少99%同一性。具有与SEQ ID NO:30从位置1至位置46的氨基酸序列或SEQ ID NO:8从位置1至位置47的氨基酸序列至少80%序列同一性的氨基酸序列的叶绿体靶向肽,因此可以是具有如下氨基酸序列的叶绿体靶向肽,所述氨基酸序列与SEQ ID NO:30从位置1至位置46的氨基酸序列或SEQ ID NO:8从位置1至位置47的氨基酸序列具有至少80%、或至少85%、或至少90%、至少92%、至少95%、至少98%、至少99%或甚至100%序列同一性。

[0136] 此外,编码热稳定性Rca 2蛋白质变体的核酸可以包含选自以下的编码性核酸序列:(a) SEQ ID Nos:31、34、36或37的核苷酸序列或其互补物,(b) 与SEQ ID NOs:31、34、36

或37的核苷酸序列或其互补物具有至少60%同一性的核苷酸序列。

[0137] SEQ ID NOs:31和36代表编码SEQ ID NO:30的氨基酸序列的核苷酸序列。SEQ ID NOs:31和36因此代表热稳定性Rca 2 β 基因变体的编码性核苷酸序列,其编码全部十一个鉴定为在来自各植物物种的Rca蛋白的热稳定性变体中保守的氨基酸残基。SEQ ID NOs:34和37代表编码SEQ ID NO:33的氨基酸序列的核苷酸序列。SEQ ID NOs:34和37因此代表热稳定性Rca 2 β 基因变体的编码性核苷酸序列,其编码十一个鉴定为在来自各植物物种的Rca蛋白的热稳定性变体中保守的氨基酸残基中的八者。这八个氨基酸残基是(a) AA1、(b) AA3、(d) AA4、(e) AA5、(f) AA8、(g) AA9、(h) AA10和(i) AA11。

[0138] 编码Rca 2蛋白的如下核酸适用于本发明,所述核酸包含与本文中所述的基因具有至少40%、至少50%、或至少60%、或至少70%、或至少80%、或至少85%、或至少90%、或至少95%、或至少98%序列同一性的核苷酸序列并且也称作变体。

[0139] 包含与SEQ ID NOs:31、34、36或37具有至少60%序列同一性的核苷酸序列的核酸,因此可以是包含与SEQ ID NOs:31、34、36或37分别具有至少60%、或至少70%、或至少80%、或至少85%、或至少90%、或至少95%、或至少98%、或100%序列同一性的核苷酸序列的核酸。

热稳定性Rca 2等位基因

[0140] 热稳定性Rca 2蛋白变体可以由Rca 2基因的热稳定性等位基因编码。

[0141] 在一个实施方案中,所述热稳定性等位基因可以包含(a) SEQ ID NOs:31、34、36或37的编码性核苷酸序列,或(b)与SEQ ID NO:31、34、36或37具有至少60%同一性并且编码蛋白质的编码性核苷酸序列,所述蛋白质包含选自以下氨基酸中至少一者:(i)与SEQ ID NO:32或35的位置59相对应的位置处的异亮氨酸,(ii)与SEQ ID NO:32或35的位置73相对应的位置处的天冬氨酸,(iii)与SEQ ID NO:32或35的位置160相对应的位置处的异亮氨酸,(iv)与SEQ ID NO:32或35的位置265相对应的位置处的精氨酸,(v)与SEQ ID NO:32或35的位置270相对应的位置处的脯氨酸,(vi)与SEQ ID NO:32或35的位置277相对应的位置处的亮氨酸,(vii)与SEQ ID NO:32或35的位置307相对应的位置处的谷氨酸,(viii)与SEQ ID NO:32或35的位置334相对应的位置处的异亮氨酸,(ix)与SEQ ID NO:32或35的位置359相对应的位置处的赖氨酸,(x)与SEQ ID NO:32或35的位置361相对应的位置处的亮氨酸,和(xi)与SEQ ID NO:32或35的位置363相对应的位置处的谷氨酸。

[0142] 还提供小麦Rca 2基因的热稳定性等位基因,其中小麦Rca 2基因是来自亚基因组B、A或D的Rca 2 β 基因或来自亚基因组B、A或D的Rca 2 α 基因。

[0143] 使用基因编辑技术,可以通过以下方式,将植物中编码热不稳定性Rca 2蛋白的内源等位基因转化成热稳定性Rca 2等位基因:对现有的Rca 2基因做出所需变化(错义突变),或将一个或多个编码热不稳定性Rca 2蛋白的内源序列替换为编码热稳定性Rca 2蛋白(例如本文所述的热稳定性Rca 2蛋白)的序列(缺失和插入突变体)。

[0144] 还可以通过使用诱变,对现有的Rca 2基因做出所需变化(错义突变),以将禾谷类植物(如小麦)中编码热不稳定性Rca 2蛋白的内源等位基因转化成热稳定性Rca 2等位基因。

重组基因和载体

[0145] 在又一个方面,重组基因包含以下有效连接的元件:(a)启动子,优选地在植物中

可表达, (b) 编码选自以下的Rca蛋白的核酸: (i) Rca 1 β 蛋白及其变体; 和(ii) 热稳定性Rca 2蛋白变体和, 任选地(c) 转录终止和多腺苷酸化区域, 优选地在植物中有功能的转录终止和多腺苷酸化区。在其他实施方案中, Rca 1 β 蛋白和核酸及其变体包含如上文所述的氨基酸序列和核苷酸序列, 并且热稳定性Rca 2蛋白和核酸变体包含本发明的氨基酸序列和核苷酸序列。在又一个实施方案中, 所述启动子是组成型启动子、组织特异性启动子或诱导型启动子。在又一个实施方案中, 启动子可以是绿色组织特异性启动子、叶肉特异性启动子、光诱导型启动子或温度诱导启动子。

[0146] 另外, 也提供能够特异性阻抑内源Rca 2基因表达的第二重组基因, 其包含以下有效连接的元件: (a) 启动子, 其优选地在植物中可表达; (b) 核酸, 当转录时所述核酸产生对编码热不稳定性Rca 2蛋白的内源Rca 2基因具有抑制性、但对编码热稳定性Rca蛋白变体的基因不具有抑制性的RNA分子; 和, 任选地(c) 转录终止和多腺苷酸化区域, 优选地在植物中有功能的转录终止和多腺苷酸化区。在一个后续实施方案中, 所述内源Rca 2基因包含SEQ ID NO:1的编码性核苷酸序列或与SEQ ID NO:1的核苷酸序列具有至少60%同一性的编码性核苷酸序列。

[0147] 这种抑制性RNA分子可以减少基因的表达, 例如通过RNA介导的基因沉默机制减少。它可以是下调靶基因表达的沉默性RNA。如本文所用, “沉默性RNA”或“沉默性RNA分子”指当引入植物细胞时减少靶基因表达的任何RNA分子。这种沉默性RNA可以例如是所谓的“反义RNA”, 因而该RNA分子包含与靶核酸的序列、优选靶基因的编码性序列的互补物具有95%序列同一性的至少20个连续核苷酸的序列。然而, 反义RNA也可以针对靶基因的调节序列, 包括启动子序列和转录终止信号和多腺苷酸化信号。沉默性RNA还包括所谓“有义RNA”, 其中该RNA分子包含与靶核酸的序列具有95%序列同一性的至少20个连续核苷酸的序列。其他沉默性RNA可以是“未多腺苷酸化的RNA”, 其包含与靶核酸的互补序列具有95%序列同一性的至少20个连续核苷酸, 如W001/12824或US6423885中描述(两份文件均通过引用方式并入本文)。又一个类型的沉默性RNA是如W003/076619(通过引用方式并入本文)中所述的RNA分子, 其包含与靶核酸或其互补物的序列具有95%序列同一性的至少20个连续核苷酸, 并且还包含大部分呈双链的区域, 如W003/076619中所述(包括这样的大部分呈双链的区域, 其包含来自马铃薯纺锤形块茎类病毒属类病毒的核定位信号或包含CUG三核苷酸重复序列)。沉默性RNA也可以是包含如本文定义的有义链和反义链的双链RNA, 其中有义链和反义链能够彼此碱基配对以形成双链RNA区域(优选地, 有义和反义RNA的所述至少20个连续核苷酸彼此互补)。有义区和反义区也可以在一个RNA分子内部存在, 从而当有义区和反义区形成双链RNA区域时, 可以形成发夹RNA(hpRNA)。本领域范围内熟知hpRNA(参见, 例如通过引用方式并入本文的W099/53050)。hpRNA可以分为长hpRNA, 其具有可以基本上互补、但不必要完全互补的长的有义区和反义区(一般大于约200bp, 范围在200-1000bp之间)。hpRNA也可以相当小, 在尺寸上是约30至约42bp, 但不比94bp长得多(参见W004/073390, 其通过引用方式并入本文)。沉默性RNA也可以是在例如W02005/052170、W02005/047505或US 2005/0144667中所述的人工micro-RNA分子或如W02006/074400中所述的ta-siRNA(全部文件均通过引用方式并入本文)。所述能够调节基因表达的RNA也可以是RNA核酶。

[0148] 短语“有效地连接”指两个或更多个核酸区域或核酸序列的功能性空间排列。例如, 启动子区可以相对于核酸序列如此安置, 从而核酸序列的转录受启动子区指导。由此,

启动子区“有效连接于”核酸序列。“功能性连接”是等同术语。

[0149] 如本文所用的“转录终止和多腺苷酸化区”是植物细胞中有功能的序列,所述序列控制新生RNA的切割,此后在所得的RNA 3'末端添加poly(A)尾。植物细胞中有功能的转录终止和多腺苷酸化信号包括但不限于3' nos、3' 35S、3' his和3' g7。

[0150] 如本文中使用时,术语“植物可表达启动子”意指能够在植物细胞中控制(启动)转录的DNA序列。这包括任何的植物源启动子,还包括能够在植物细胞中指导转录的任何非植物源启动子,即,某些病毒源启动子或细菌源启动子,例如CaMV35S(Harpster等人(1988) Mol Gen Genet. 212(1):182-90,地三叶草病毒启动子第4号或第7号(W09606932)或T-DNA基因启动子,还包括组织特异性或器官特异性启动子,包括但不限于种子特异性启动子(例如,W089/03887)、器官原基特异性启动子(An等人(1996) Plant Cell 8(1):15-30)、茎特异性启动子(Keller等人(1988) EMBO J. 7(12):3625-3633)、叶特异性启动子(Hudspeth等人(1989) Plant Mol Biol. 12:579-589)、叶肉特异性启动子(如光诱导型Rubisco启动子)、根特异性启动子(Keller等人(1989) Genes Dev. 3:1639-1646)、块茎特异性启动子(Keil等人(1989) EMBO J. 8(5):1323-1330)、维管组织特异性启动子(Peleman等人(1989) Gene 84:359-369)、雄蕊选择性启动子(W089/10396、W092/13956)、开裂区特异性启动子(W097/13865)等。

[0151] 合适的本发明启动子是组成型植物可表达启动子。组成型植物可表达启动子是本领域熟知的并且包括CaMV35S启动子(Harpster等人(1988) Mol Gen Genet. 212(1):182-90)、肌动蛋白启动子,例如,来自稻肌动蛋白基因的启动子(McElroy等人,1990, Plant Cell 2:163)、木薯叶脉花叶病毒启动子(Verdaguer等人,1996 Plant Mol. Biol. 31:1129)、GOS启动子(de Pater等人,1992, Plant J. 2:837)、组蛋白H3启动子(Chaubet等人1986, Plant Mol Biol 6:253)、根癌农杆菌(*Agrobacterium tumefaciens*)胭脂碱合酶(Nos)启动子(Depicker等人,1982, J. Mol. Appl. Genet. 1:561),或遍在蛋白启动子,例如,玉米遍在蛋白1基因启动子(Christensen等人,1992, Plant Mol. Biol. 18:675)。

[0152] 适于本发明的又一启动子是驱动编码Rca蛋白的基因表达的内源启动子。

[0153] 可以在重组载体中提供上文描述的任何核酸序列。重组载体一般按照5'至3'方向包含:指导核酸序列转录的启动子和核酸序列。如需要,重组载体可以进一步包含3'转录终止子、3'多腺苷酸化信号、其他非翻译核酸序列、转运和靶向核酸序列,选择标记、增强子和操纵基因。措辞“5'UTR”指基因编码区上游或5'的DNA非翻译区,“3'UTR”指基因编码区下游或3'的DNA非翻译区。用于制备重组载体的手段是本领域熟知的。US4971908、US4940835、US4769061和US4757011中描述了用于产生特别适用于植物转化的重组载体的方法。可用于高等植物中表达核酸的常见载体是本领域熟知的,并且包括衍生自根癌农杆菌(*Agrobacterium tumefaciens*)肿瘤诱导型(Ti)质粒的载体。也可以在重组载体中提供一个或多个额外的启动子。这些启动子可以有效连接至例如但不限于上文描述的任何核酸序列。备选地,启动子可以有效连接至其他核酸序列,如编码转运肽、选择标记蛋白质或反义序列的核酸序列等等。这些额外的启动子可以基于将要插入载体的细胞类型选择。另外,细菌、酵母和植物中发挥作用的启动子均也在本领域中有充分教导。额外的启动子也可以基于其调节特征选择。这类特征的实例包括转录活性增强作用、诱导性、组织特异性和发育阶段特异性。

[0154] 重组载体还可以含有一个或多个额外的核酸序列。这些额外的核酸序列通常可以是适用于重组载体中的任何序列。此类核酸序列包括但不限于上文描述的任何核酸序列及其修饰形式。额外的结构性核酸序列也可以有效连接至前述的任一启动子。一个或多个结构性核酸序列可以各自有效连接至独立的启动子。备选地,这些结构性核酸序列可以有效连接至单一启动子(即,单个操纵子)。

方法和用途

[0155] 在一个方面,本发明提供一种用于增加禾谷类植物中热稳定性Rca (Rubisco活化酶) 蛋白比率的方法,所述方法包括(a) 向禾谷类植物的细胞提供重组基因,所述重组基因包含如下有效连接的元件:启动子,所述启动子优选地在植物中可表达;编码Rca 1 β 蛋白及其变体或编码热稳定性Rca 2蛋白变体的核酸和,任选地转录终止和多腺苷酸化区,优选地在植物中有功能的转录终止和多腺苷酸化区;以及减少所述禾谷类植物细胞中内源热不稳定性Rca 2蛋白的表达,其中与不包含所述重组基因的对照禾谷类植物细胞相比,所述比率升高;或(b) 向禾谷类植物的细胞引入本发明的至少一个热稳定性Rca 2等位基因,其中与不包含所述热稳定性Rca 2等位基因的对照禾谷类植物细胞相比,所述比率升高。在又一个实施方案中,禾谷类植物是小麦植物。

[0156] 在又一个方面,Rca 1 β 蛋白及其变体包含如上文所述的氨基酸序列并且由包含上述编码性核酸序列的核酸编码。

[0157] 在另一个实施方案中,所述热稳定性Rca 2蛋白变体包含如上文所述的氨基酸序列并且由包含上述编码性核酸序列的核酸编码。

[0158] 如本文所用的“增加热稳定性Rca比率”意指,相对于(热稳定性和热不稳定性)Rca蛋白的总体丰度,增加热稳定性Rca蛋白的相对丰度。这可以通过增加热稳定性Rca蛋白的丰度、通过降低热不稳定性Rca蛋白的丰度、和/或通过增加热稳定性Rca蛋白的丰度且降低热不稳定性Rca蛋白的丰度来实现。增加的热稳定性Rca比率可以是至少约或约15%、至少约或约30%、至少约或约45%、至少约或约60%、至少约或约75%、至少约或约90%、或至少约或约100%。增加的热稳定性Rca比率可以在约15%和约30%之间、在约15%和约45%之间、在约15%和约60%之间、在约15%和约75%之间、在约15%和约90%之间、在约15%和约100%之间、在约30%和约45%之间、在约30%和约60%之间、在约30%和约75%之间、在约30%和约90%之间、在约30%和约100%之间、在约45%和约60%之间、在约45%和约75%之间、在约45%和约90%之间、在约45%和约100%之间、在约60%和约75%之间、在约60%和约90%之间、在约60%和约100%之间、在约75%和约90%之间、在约75%和约100%之间、在约90%和约100%之间。

[0159] 另外通过增加热稳定性Rca蛋白的比率,包含Rca蛋白和Rubisco的复合体的热稳定性按比例增加(Shivhare和Mueller-Cajar,2017,Plant Physiol DOI 10.1104/pp17.00554)。

[0160] 与本申请相关的“引入”涉及通过人工手段将遗传信息放入植物细胞或植物中。这可以通过本领域已知的向植物细胞、原生质体、愈伤组织、根、块茎、种子、茎、叶、籽苗、胚、花粉和小孢子、其他植物组织或完整植株引入RNA或DNA的任何方法实现。“引入”还包括稳定整合入植物的基因组。可以通过转化、或通过与转化获得的植物或其后代杂交(也称作“渐渗”(introgression)),引入重组基因。还可以通过诱变或通过基因编辑引入等位基因。

[0161] 术语“提供”可以指通过转化将外源DNA分子引入植物细胞,任选地后续从转化的植物细胞再生出植物。该术语还可以指通过以下方式引入重组DNA分子:使包含重组DNA分子的转基因植物与另一植物杂交并且选择已经继承重组DNA分子或转基因的后代植物。又一个备选的提供手段指,通过如原生质体融合以及任选地后续从融合的原生质体再生出植物等技术,引入重组DNA分子。

[0162] 可以通过本领域熟知的方法向植物细胞提供重组基因。

[0163] 本文中术语“转化”指将核酸引入(或转移入)受体宿主,如植物或任何植物部分或组织,包括植物细胞、原生质体、愈伤组织、根、块茎、种子、茎、叶、纤维、籽苗、胚和花粉。含有转化的核酸序列的植物称作“转基因植物”。“转化的”、“转基因”和“重组”指已经向其中引入异源核酸分子(例如表达盒或重组载体)的宿主生物如植物。核酸可以稳定整合入植物的基因组。

[0164] 如本文所用,短语“转基因植物”指,核酸已经稳定地整合入植物基因组(例如,细胞核或质体基因组)中的植物。换句话说,含有转化的核酸序列的植物称作“转基因植物”并且包括直接从转化获得的植物及它们的后代(T_x代)。转基因和重组指已经向其中引入异源核酸分子(例如,如本文所述的启动子、重组基因或载体)的宿主生物如植物。核酸可以稳定整合入植物的基因组。

[0165] 明显地,所用的转化方法与本发明不甚相关。植物转化现在是常规技术。有利地,几种转化方法中的任一方法可以用来将目的核酸/基因引入合适的祖先细胞。转化方法包括使用脂质体、电穿孔法、增加游离DNA摄入的化学品、DNA直接注射至植物、粒子枪轰击法、使用病毒或花粉的转化法和显微投射法(microprojection)。转化方法可以选自用于原生质体的钙/聚乙二醇法(Krens等人(1982) *Nature* 296:72-74;Negrutiu等人(1987) *Plant.Mol.Biol.*8:363-373);原生质体电穿孔法(Shillito等人(1985) *Bio/Technol.*3:1099-1102);微量注入植物材料法(Crossway等人(1986) *Mol.Gen.Genet.*202:179-185);DNA涂布粒子或RNA涂布粒子轰击法(Klein等人(1987) *Nature* 327:70)、(非整合性)病毒感染法等。

[0166] 可以建立用于多种禾谷植物的不同转化系统:组织电穿孔法、原生质体转化法、和可再生组织和细胞中采用粒子轰击的DNA转移法(关于综述,参见Jane, *Euphytica* 85(1995),35-44)。已经在文献中数次描述小麦的转化(关于综述,参见Maheshwari, *Critical Reviews in Plant Science* 14(2)(1995),149-178;Nehra等人, *Plant J.*5(1994),285-297)。

[0167] 本发明的重组DNA分子可以使用本领域熟知的方法,稳定地或暂时地提供给植物。可以将重组基因引入植物,或可以如EP 1339859中所述在植物细胞内部产生重组基因。

[0168] 如本文所用的“对照植物”指,在遗传上类似于受测试植物,但不携带重组基因的植物(如野生型植物或无效分离子植物),或不携带突变等位基因的植物(如野生型植物或野生型分离子植物)。

[0169] 通过本文所述方法获得的转化的植物细胞和植物可以进一步用于本领域熟知的育种程序中,如杂交、自交和回交。育种程序可以涉及杂交以产生F₁(第一子)代,随后是几个自交世代(产生F₂、F₃等)。育种程序还可以涉及回交(BC)步骤,其中子代与亲本品系之一(称作轮回亲本)回交。

[0170] 通过本文公开的方法获得的转化的植物细胞和植物还可以进一步用于后续转化程序中,例如以引入再一重组基因。

[0171] 在又一个实施方案中,如上文所述,减少内源热不稳定性Rca 2蛋白的表达包括,向禾谷类植物的细胞引入本发明的至少一个敲除突变Rca 2等位基因,或向所述禾谷类植物的细胞提供能够特异性阻抑内源热不稳定性Rca 2基因表达的第二重组基因。

[0172] 适于本发明的是增加禾谷植物中热稳定性Rca蛋白比率的方法,所述方法包括向禾谷类植物的细胞引入本发明的至少两个、至少三个、至少四个、至少五个或甚至全部六个敲除突变Rca 2等位基因。这类至少两个敲除突变Rca 2等位基因可以是来自亚基因组B的两个敲除突变Rca 2等位基因、来自亚基因组D的两个敲除突变Rca 2等位基因、来自亚基因组A的两个敲除突变Rca 2等位基因、来自亚基因组B的一个敲除突变Rca 2等位基因和来自亚基因组D的一个敲除突变Rca 2等位基因、来自亚基因组B的一个敲除突变Rca 2等位基因和来自亚基因组A的一个敲除突变Rca 2等位基因、或来自亚基因组D的一个敲除突变Rca 2等位基因和来自亚基因组A的一个敲除突变Rca 2等位基因。这类至少三个敲除突变Rca 2等位基因可以是来自亚基因组B的两个敲除突变Rca 2等位基因和来自亚基因组A的一个敲除突变Rca 2等位基因、来自亚基因组B的两个敲除突变Rca 2等位基因和来自亚基因组D的一个敲除突变Rca 2等位基因、来自亚基因组D的两个敲除突变Rca 2等位基因和来自亚基因组B的一个敲除突变Rca 2等位基因、来自亚基因组D的两个敲除突变Rca 2等位基因和来自亚基因组A的一个敲除突变Rca 2等位基因、来自亚基因组A的两个敲除突变Rca 2等位基因和来自亚基因组B的一个敲除突变Rca 2等位基因、来自亚基因组A的两个敲除突变Rca 2等位基因和来自亚基因组D的一个敲除突变Rca 2等位基因、或来自亚基因组B的一个敲除突变Rca 2等位基因、来自亚基因组A的一个敲除突变Rca 2等位基因和来自亚基因组D的一个敲除突变Rca 2等位基因。

[0173] 这类至少四个敲除突变Rca 2等位基因可以是来自亚基因组B的两个敲除突变Rca 2等位基因和来自亚基因组A的两个敲除突变Rca 2等位基因;来自亚基因组B的两个敲除突变Rca 2等位基因和来自亚基因组D的两个敲除突变Rca 2等位基因;或来自亚基因组D的两个敲除突变Rca 2等位基因和来自亚基因组A的两个敲除突变Rca 2等位基因。这类至少四个敲除突变Rca 2等位基因也可以是来自亚基因组B的两个敲除突变Rca 2等位基因、来自亚基因组A的一个敲除突变Rca 2等位基因和来自亚基因组D的一个敲除突变Rca 2等位基因;或来自亚基因组D的两个敲除突变Rca 2等位基因、来自亚基因组A的一个敲除突变Rca 2等位基因和来自亚基因组B的一个敲除突变Rca 2等位基因;或来自亚基因组A的两个敲除突变Rca 2等位基因、来自亚基因组B的一个敲除突变Rca 2等位基因和来自亚基因组D的一个敲除突变Rca 2等位基因。这类至少五个敲除突变Rca 2等位基因可以是来自亚基因组B的两个敲除突变Rca 2等位基因、来自亚基因组A的两个敲除突变Rca 2等位基因和来自亚基因组D的一个敲除突变Rca 2等位基因;或来自亚基因组B的两个敲除突变Rca 2等位基因、来自亚基因组D的两个敲除突变Rca 2等位基因和来自亚基因组A的一个敲除突变Rca 2等位基因;或来自亚基因组D的两个敲除突变Rca 2等位基因、来自亚基因组A的两个敲除突变Rca 2等位基因和来自亚基因组B的一个敲除突变Rca 2等位基因。

[0174] “减少禾谷类植物细胞中内源热不稳定性Rca 2蛋白的表达”指,与不包含本发明至少一个敲除突变Rca 2等位基因或能够特异性阻抑内源热不稳定性Rca 2基因表达的第

二重组基因(如上文所述)的细胞产生的有功能热不稳定性Rca 2蛋白的量相比,由包含本发明该至少一个敲除突变Rca 2等位基因或能够特异性阻抑内源热不稳定性Rca 2基因表达的第二重组基因(如上文所述)的细胞产生的有功能热不稳定性Rca 2蛋白的量减少至少30%、40%、50%、60%、70%、80%、90%或甚至100%。

[0175] 还适于本发明的是增加禾谷植物(如小麦)中热稳定性Rca蛋白比率的方法,所述方法包括向禾谷类植物的细胞引入本发明的至少两个、至少三个、至少四个、至少五个或甚至全部六个热稳定性Rca 2等位基因。这类至少两个热稳定性突变Rca 2等位基因可以是来自亚基因组B的两个热稳定性Rca 2等位基因、来自亚基因组D的两个热稳定性Rca 2等位基因、来自亚基因组A的两个热稳定性Rca 2等位基因、来自亚基因组B的一个热稳定性Rca 2等位基因和来自亚基因组D的一个热稳定性Rca 2等位基因、来自亚基因组B的一个热稳定性Rca 2等位基因和来自亚基因组A的一个热稳定性Rca 2等位基因、或来自亚基因组D的一个热稳定性Rca 2等位基因和来自亚基因组A的一个热稳定性Rca 2等位基因。这类至少三个热稳定性Rca 2等位基因可以是来自亚基因组B的两个热稳定性Rca 2等位基因和来自亚基因组A的一个热稳定性Rca 2等位基因、来自亚基因组B的两个热稳定性Rca 2等位基因和来自亚基因组D的一个热稳定性Rca 2等位基因、来自亚基因组D的两个热稳定性Rca 2等位基因和来自亚基因组B的一个热稳定性Rca 2等位基因、来自亚基因组D的两个热稳定性Rca 2等位基因和来自亚基因组A的一个热稳定性Rca 2等位基因、来自亚基因组A的两个热稳定性Rca 2等位基因和来自亚基因组B的一个热稳定性Rca 2等位基因、来自亚基因组A的两个热稳定性Rca 2等位基因和来自亚基因组D的一个热稳定性Rca 2等位基因、或来自亚基因组B的一个热稳定性Rca 2等位基因、来自亚基因组A的一个热稳定性Rca 2等位基因和来自亚基因组D的一个热稳定性Rca 2等位基因。

[0176] 这类至少四个热稳定性Rca 2等位基因可以是来自亚基因组B的两个热稳定性Rca 2等位基因和来自亚基因组A的两个热稳定性Rca 2等位基因;来自亚基因组B的两个热稳定性Rca 2等位基因和来自亚基因组D的两个热稳定性Rca 2等位基因;或来自亚基因组D的两个热稳定性Rca 2等位基因和来自亚基因组A的两个热稳定性Rca 2等位基因。这类至少四个热稳定性Rca 2等位基因也可以是来自亚基因组B的两个热稳定性Rca 2等位基因、来自亚基因组A的一个热稳定性Rca 2等位基因和来自亚基因组D的一个热稳定性Rca 2等位基因;或来自亚基因组D的两个热稳定性Rca 2等位基因、来自亚基因组A的一个热稳定性Rca 2等位基因和来自亚基因组B的一个热稳定性Rca 2等位基因;或来自亚基因组A的两个热稳定性Rca 2等位基因、来自亚基因组B的一个热稳定性Rca 2等位基因和来自亚基因组D的一个热稳定性Rca 2等位基因。这类至少五个热稳定性Rca 2等位基因可以是来自亚基因组B的两个热稳定性Rca 2等位基因、来自亚基因组A的两个热稳定性Rca 2等位基因和来自亚基因组D的一个热稳定性Rca 2等位基因;或来自亚基因组B的两个热稳定性Rca 2等位基因、来自亚基因组D的两个热稳定性Rca 2等位基因和来自亚基因组A的一个热稳定性Rca 2等位基因;或来自亚基因组D的两个热稳定性Rca 2等位基因、来自亚基因组A的两个热稳定性Rca 2等位基因和来自亚基因组B的一个热稳定性Rca 2等位基因。

[0177] 在另一个方面,提供一种增加禾谷类植物耐热性的方法,所述方法包括增加热稳定性Rca蛋白的比率并且再生所述植物,其中与不包含所述增加比率的热稳定性Rca蛋白的禾谷类植物相比,耐热性增加。在又一个实施方案中,根据本文所述的用于增加禾谷植物中

热稳定性Rca (Rubisco活化酶) 蛋白比率的方法, 增加热稳定性Rca蛋白的比率。在其他实施方案中, 所述热稳定性Rca蛋白是如本文所述的Rca1 β 蛋白或其变体, 或者所述热稳定性Rca蛋白是本发明的热稳定性Rca 2蛋白变体。

[0178] 在热胁迫下, 不耐热的Rca蛋白与Rubisco解离。Rubisco酶因此活性更小或甚至无活性并且光合作用减少或停止。本文公开的热稳定性Rca蛋白向包含Rca蛋白和Rubisco的复合体赋予热稳定性。这种热稳定性复合体允许Rubisco在热胁迫下保持活性并且因此允许维持光合活性。因此可以通过测量植物的光合活性, 测量植物耐热性。本领域熟知测量植物中光合活性的方法 (参见例如Kalaji等人2012Photosynth Res 114:69-96)。

[0179] 在本发明的又一个方面, 提供一种增加热胁迫条件下禾谷类植物 (如小麦植物) 产量的方法, 所述方法包括增加热稳定性Rca蛋白的比率并且再生所述植物, 其中与不包含所述增加比率的热稳定性Rca蛋白的禾谷类植物相比, 产量增加。在又一个实施方案中, 根据本文所述的用于增加禾谷植物中热稳定性Rca (Rubisco活化酶) 蛋白比率的方法, 增加热稳定性Rca蛋白的比率。在其他实施方案中, 所述热稳定性Rca蛋白是如本文所述的Rca 1 β 蛋白或其变体, 或者所述热稳定性Rca蛋白是本发明的热稳定性Rca 2蛋白变体。增加的产量可以是种子产量或千粒种子重量。

[0180] 进一步提供一种产生耐热性增加的禾谷类植物 (如小麦植物) 的方法, 所述方法包括如本文中公开的增加热稳定性Rca蛋白的比率并且再生所述植物。

[0181] 还提供本发明的热稳定性Rca 2蛋白变体、编码本发明热稳定性Rca 2蛋白变体的核酸、本发明的重组基因、能够特异性阻抑本文所述的内源Rca 2基因表达的重组基因、或本文提供的Rca 2基因的热稳定性等位基因的用途, 用于增加禾谷植物中热稳定性Rca蛋白的比率、增加禾谷类植物耐热性、增加热胁迫条件下禾谷类植物的产量、或产生耐热性增加的禾谷类植物。这种禾谷类植物可以是小麦植物。

[0182] 如本文所用的“产量”可以包括收获的植物或植物部分 (如生物质或种子) 的产量, 包括种子蛋白质含量、种子重量 (计量为千粒种子重量)、种子数目。增加的产量可以是增加的每穗产量、增加的每分蘖产量、增加的每植株产量、和增加的每表面单位栽培用地的产量, 如每公顷的产量。可以通过增加例如对非生物胁迫条件的耐受性, 增加产量。

[0183] 当产量是种子产量时, 与热稳定性Rca蛋白的比率未增加的植物相比, 用本文所述方法实现的产量增加可以是至少约5%、至少约6%、至少约7%、至少约8%、至少约9%或至少约10%。当产量是种子重量时, 与热稳定性Rca蛋白的比率未增加的植物相比, 用本文所述方法实现的产量增加可以是至少约5%、至少约6%、至少约7%、至少约8%、至少约9%或至少约10%。

[0184] “胁迫”指非最佳环境条件如非生物胁迫。非生物胁迫可以包括环境胁迫因素如旱、涝、极端温度 (高温或低温)、土壤盐度或重金属、低氧 (hypoxia)、缺氧 (anoxia)、渗透胁迫、氧化胁迫、低养分水平如氮或磷。

[0185] 如本文所用的“热胁迫”指植物暴露于高温持续规定的时间。这种高温可以仅每日持续数小时并且可以出现至少或不超过2天、至少或不超过3天、至少或不超过4天、至少或不超过5天、至少或不超过6天、至少或不超过7天、至少或不超过8天、至少或不超过9天、至少或不超过10天、至少或不超过15天、或至少或不超过20天。它也可以持续更长时间如至少或不超过3周、至少或不超过4周、至少或不超过5周、至少或不超过6周、至少或不超过2个

月、或至少或不超过3个月。针对禾谷植物如小麦的高温意指超过禾谷植物如小麦的最佳温度范围的温度,并且可以是至少约28摄氏度(°C)、至少约29°C、至少约30°C、至少约31°C、至少约32°C、至少约33°C、至少约34°C、至少约35°C、至少约36°C、至少约37°C、至少约38°C、至少约39°C、至少约40°C、至少约41°C、至少约42°C、至少约43°C、至少约44°C、至少约45°C。针对禾谷植物如小麦的高温也可以是约28°C与约30°C之间、约28°C与约32°C之间、约28°C与约34°C之间、约28°C与约36°C之间、约28°C与约38°C之间、约28°C与约40°C之间、约28°C与约42°C之间、约28°C与约45°C之间、约30°C与约32°C之间、约30°C与约34°C之间、约30°C与约36°C之间、约30°C与约38°C之间、约30°C与约40°C之间、约30°C与约42°C之间、约30°C与约45°C之间、约32°C与约34°C之间、约32°C与约36°C之间、约32°C与约38°C之间、约32°C与约40°C之间、约32°C与约42°C之间、约32°C与约45°C之间、约34°C与约36°C之间、约34°C与约38°C之间、约34°C与约40°C之间、约34°C与约42°C之间、约34°C与约45°C之间、约36°C与约38°C之间、约36°C与约40°C之间、约36°C与约42°C之间、约36°C与约45°C之间、约38°C与约40°C之间、约38°C与约42°C之间、约38°C与约45°C之间、约40°C与约42°C之间、约40°C与约45°C之间、或约42°C与45°C之间的温度。

[0186] 本发明的另一个方面提供一种产生食物或饲料如饼粕、粮食、淀粉、面粉或蛋白质、或工业产品如生物燃料、纤维、工业化学品、药物或营养药的方法,所述方法包括获得本发明的植物或其部份,并且从该植物或其部份制备食物、饲料或工业产品。

[0187] 在小麦植物或其他禾谷类植物的情况下,食物产品的实例包括面粉、淀粉、发酵面包或无酵面包、面食、面条、动物饲料、早餐谷物食品、点心食品、蛋糕、麦芽、馅饼、面筋、和含有面粉基酱的食物。

[0188] 本领域熟知从小麦产生这种食物、饲料或工业产品的方法。例如,通过在面粉厂细磨籽实,产生面粉(参见例如www.madehow.com/Volume-3/Flour.html),和从小麦秆或小麦秆和小麦粗粉的混合物产生生物燃料(参见例如Erdei等人,Biotechnology for Biofuels,2010,3:16)。

[0189] 在又一个实施方案中,提供一种增加Rca 2蛋白的热稳定性的方法,所述方法包括向所述Rca 2蛋白的氨基酸序列引入至少一个氨基酸置换,其中所述氨基酸置换选自(a)将SEQ ID NO:4的位置59相对应的位置处的缬氨酸置换或替换为异亮氨酸、(b)将SEQ ID NO:4的位置73相对应的位置处的甘氨酸置换或替换为天冬氨酸、(c)将SEQ ID NO:4的位置160相对应的位置处的甲硫氨酸置换或替换为异亮氨酸、(d)将SEQ ID NO:4的位置265相对应的位置处的谷氨酰胺置换或替换为精氨酸、(e)将SEQ ID NO:4的位置270相对应的位置处的丝氨酸置换或替换为脯氨酸、(f)将SEQ ID NO:4的位置277相对应的位置处的异亮氨酸置换或替换为亮氨酸、(g)将SEQ ID NO:4的位置307相对应的位置处的丝氨酸置换或替换为谷氨酸、(h)将SEQ ID NO:4的位置334相对应的位置处的缬氨酸置换或替换为异亮氨酸、(i)将SEQ ID NO:4的位置359相对应的位置处的苏氨酸置换或替换为赖氨酸、(j)将SEQ ID NO:4的位置361相对应的位置处的甲硫氨酸置换或替换为亮氨酸,和(k)将SEQ ID NO:4的位置363相对应的位置处的谷氨酰胺置换或替换为谷氨酸,其中与不包含任何所列氨基酸置换的Rca 2蛋白相比,所述Rca 2蛋白的热稳定性升高。在又一个实施方案中,增加Rca 2蛋白热稳定性约7°C。

[0190] 适于本发明的是包含所述氨基酸置换的Rca 2蛋白的热稳定性增加至少约或约3

℃、至少约或约4℃、至少约或约5℃、至少约或约6℃、至少约或约7℃、至少约或约8℃、至少约或约9℃、至少约或约10℃。

[0191] 可以理解,本文所述的方法可以包括引入至少两个、至少三个、至少四个、至少五个、至少六个、至少七个、至少八个、至少九个、至少十个或甚至全部十一个氨基酸置换。由于上文所列的氨基酸置换导致引入本文中鉴定为与Rca蛋白热稳定性相关的一个或多个氨基酸(AA1、AA2、AA3、AA4、AA5、AA6、AA7、AA8、AA9、AA10和AA11),因此清楚的,至少两个、至少三个、至少四个、至少五个、至少六个、至少七个、至少八个、至少九个或至少十个氨基酸置换的不同组合对应于鉴定为与Rca蛋白热稳定性相关的且列于本文中的至少两个、至少三个、至少四个、至少五个、至少六个、至少七个、至少八个、至少九个或至少十个氨基酸残基的组合。

[0192] 此外,本文中提供一种用于产生热稳定性Rca蛋白或热稳定性Rca蛋白变体的方法,所述方法包括培养包含重组基因的宿主细胞和分离产生的蛋白质,所述重组基因包含编码如上文所述的热稳定性Rca蛋白或热稳定性Rca蛋白变体的核酸。

[0193] 所述宿主细胞表达或过量表达本发明的热稳定性Rca蛋白或热稳定性Rca蛋白变体。因此,在宿主细胞中产生并从其中分离本发明的所述蛋白质。如若宿主细胞产生本发明的蛋白质并将它分泌到周围的培养基(例如,由在蛋白质上连接的合适信号肽导致),则分离指将包含蛋白质的培养基与宿主细胞分离。所述培养基随后可以经历其他纯化步骤(参见下文)。

[0194] 本领域技术人员熟知用于培养原核或真核宿主的合适条件。例如,培养细菌的合适条件是在Luria Bertani (LB) 培养基中通气培养。为了增加表达产物的产量和溶解度,可以用已知增强或促进这两者的合适添加物进行缓冲或补充培养基。可以从4℃至约37℃培养大肠杆菌,确切温度或温度顺序取决于待过量表达的分子。通常而言,技术人员也知晓,可能需要针对宿主需求和所表达多肽的要求,适应性调整这些条件。如若诱导型启动子在宿主细胞中存在的载体上控制本发明核酸,则可以通过添加适宜的诱导剂,诱导多肽表达。技术人员已知合适的表达方案和策略。

[0195] 用于真核细胞的合适表达方案是技术人员熟知的并且可以取自例如Sambrook, 2001。

[0196] 用于昆虫细胞培养的合适培养基例如是TNM+10%FCS或SF900培养基。通常将昆虫细胞在27℃作为贴壁或悬浮培养物培养。

[0197] 分离已产生多肽的方法是本领域熟知,并且包括而不限于方法步骤如硫酸铵析出、离子交换色谱、凝胶过滤色谱(大小排阻色谱)、亲和色谱、高压液相色谱(HPLC)、反相HPLC、圆盘凝胶电泳或免疫沉淀法,例如参见Sambrook, 2001。

细胞和植物

[0198] 其他实施方案提供宿主细胞,如大肠杆菌细胞、农杆菌细胞、酵母细胞或植物细胞,所述宿主细胞包含:(a) 包含编码本发明热稳定性Rca蛋白的核酸的重组基因或包含这种重组基因的载体,(b) 能够特异性阻抑如本文所述内源热不稳定性Rca 2基因表达的重组基因或包含这种重组基因的载体,或(c) 本发明的热稳定性Rca 2蛋白变体。

[0199] 其他实施方案提供一种植物细胞,其包含(a) 至少一个如本文所述的敲除Rca 2等位基因和/或(b) 本发明的至少一个热稳定性Rca 2等位基因。在又一个实施方案中,包含了

含有编码本发明热稳定性Rca蛋白的核酸的重组基因或含有这种重组基因的载体的植物细胞,可以进一步包含能够特异性阻抑如本文所述内源热不稳定性Rca 2基因表达的重组基因或包含这种重组基因的载体或至少一个如本文所述的敲除Rca 2等位基因。植物细胞可以是禾谷类植物细胞或小麦植物细胞。

[0200] 在又一个实施方案中,提供表达本发明耐热Rca 2蛋白变体的植物。所述植物可以是禾谷类植物或小麦植物。

[0201] 也可以向宿主细胞引入其他核酸序列连同本文所述的重组基因(例如,也与本发明载体连接)。这些其他序列可以包括3'转录终止子、3'多腺苷酸化信号、其他非翻译的核酸序列、转运或靶向序列、选择标记、增强子和操纵基因。上文描述了本发明的优选核酸序列,包括重组载体、结构性核酸序列、启动子和其他调节元件。

[0202] 在其他实施方案中,提供包含任何本发明重组基因和等位基因的植物。又一个实施方案提供从本发明植物可获得的植物部分和种子。这些植物部分和种子包含上文描述的重组基因或等位基因。在另一个实施方案中,本发明的植物、植物部分或种子是小麦植物、植物部分或种子。

[0203] 包含任何本发明重组基因的植物细胞或植物可以是包含如下重组基因的植物细胞或植物,所述重组基因的启动子或与所述启动子有效连接的异源核酸序列相对于植物细胞是异源的。这类植物细胞或植物可以是其中借助转化引入重组基因的转基因植物。备选地,植物的植物细胞可以包含源自相同物种的本发明启动子与有效连接的也源自相同物种的核酸,即启动子和有效连接的核酸相对于植物细胞均不是异源的,但是启动子有效连接至自然界中不与它连接的核酸。重组基因可以借助转化引入在植物或植物细胞中,从而启动子和有效连接的核苷酸均在基因组中它们不天然存在的位置处。备选地,本发明启动子可以按照定向方式整合在植物或植物细胞的基因组中位于编码目的表达产物的内源核酸上游,即,旨在调节内源性基因的表达模式。按照定向方式整合在内源核酸上游的启动子可以是整合在最初衍生它的植物物种的细胞中或异源植物物种的细胞中。备选地,异源核酸可以按照定向方式整合在植物或植物细胞的基因组中位于本发明启动子下游,从而所述异源核酸按根偏好方式表达并具有胁迫诱导性。所述异源核酸可以是相对于启动子为异源的核酸,即启动子与所述异源核酸的组合通常在自然界中不存在。所述异源核酸可以是对其中插入它的所述植物物种为异源的核酸,但是,它也可以在所述植物物种中天然存在但位于植物基因组中的不同位置处。可以通过靶向的序列插入,例如使用如W02005/049842中所述的方法,将所述启动子或所述异源核酸定向地整合在植物基因组中。

[0204] “植物”涵盖“单子叶植物”。“单子叶植物”,也称作“单子叶植物(monocot plant)”或“单子叶植物(monocot)”,是本领域熟知的并且是种子一般具有一片子叶的植物。单子叶植物的实例是草本植物,如草地早熟禾(蓝草、早熟禾属(Poa)),饲用草如羊茅、黑麦草、温带牧草、如剪股颖属(Agrostis),和禾谷植物,例如,小麦、燕麦、黑麦、大麦、稻、黑小麦、斯佩耳特小麦、单粒小麦、二粒小麦、硬粒小麦、kamut、高粱和玉蜀黍(玉米)。

[0205] 本发明的植物可以是禾谷类植物。本发明的禾谷类植物可以是小麦植物。

[0206] 如本文所用的“小麦”或“小麦植物”可以是可用于生长小麦的任何品种。小麦的实例是但不限于普通小麦(*Triticum aestivum*)、埃塞俄比亚小麦(*Triticum aethiopicum*)、密穗小麦(*Triticum Compactum*)、野生二粒小麦(*Triticum dicoccoides*)、二粒小麦

(*Triticum dicoccon*)、硬粒小麦(*Triticum durum*)、单粒小麦(*Triticum monococcum*)、斯佩耳特小麦(*Triticum spelta*)、圆锥小麦(*Triticum turgidum*)。“小麦”进一步涵盖春小麦品种和冬小麦品种,其中冬小麦品种依据开花需要春化来定义,而春小麦品种不需要春化以开花。

[0207] 如本文所用的“植物部分”是植物的部分,所述部分可以是细胞、组织或器官如种子、分离的部分如根、叶、花、花粉等。

[0208] 本发明的植物可以额外地含有赋予除草剂抗性的内源基因或转基因,如赋予草铵膦(glufosinate ammonium) (Liberty®、Basta®或Ignite®) 抗性的bar或pat基因[EP 0 242 236和EP 0 242 246,通过引用的方式并入];或者任何修饰的EPSPS基因,如来自玉米的2mEPSPS基因[EP 0 508 909和EP 0 507 698,通过引用方式并入],或者赋予草甘膦(RoundupReady®)抗性的草甘膦乙酰转移酶或草甘膦氧化还原酶,或者赋予溴苯腈抗性的溴苯腈腈水解酶,或赋予对磺酰脲类、咪唑啉酮类、磺酰基氨基羰基三唑啉酮类、三唑并嘧啶类或嘧啶基(氧代/硫代)苯甲酸类的抗性的任何修饰的AHAS基因。

[0209] 本发明的植物或植物种子可以用化学化合物进一步处理,如选自以下的化学化合物:除草剂:烯草酮(Clethodim)、二氯吡啶酸(Clopyralid)、氯甲草(Diclofop)、胺苯磺隆(Ethametsulfuron)、吡氟禾草灵(Fluazifop)、草铵膦(Glufosinate)、草甘膦(Glyphosate)、吡草胺(Metazachlor)、喹草酸(Quinmerac)、喹禾灵(Quizalofop)、酞肟草(Tepaloxymid)、氟乐灵(Trifluralin)。杀真菌剂/PGR:腈嘧菌酯(Azoxystrobin)、N-[9-(二氯亚甲基)-1,2,3,4-四氢-1,4-亚甲基萘-5-基]-3-(二氟甲基)-1-甲基-1H-吡唑-4-甲酰胺(Benzovindiflupyr、Benzodiflupyr)、Bixafen、啶酰菌胺(Boscalid)、多菌灵(Carbendazim)、萎锈灵(Carboxin)、氯化矮壮素(Chlormequat-chloride)、盾壳霉(Coniothyrium minitans)、环唑醇(Cyproconazole)、环丙嘧啶(Cyprodinil)、噁醚唑(Difenoconazole)、烯酰吗啉(Dimethomorph)、醚菌胺(Dimoxystrobin)、氧唑菌(Epoxiconazole)、噁唑酮菌(Famoxadone)、氟啶胺(Fluazinam)、咯菌腈(Fludioxonil)、氟吡菌胺(Fluopicolide)、氟吡菌酰胺(Fluopyram)、氟嘧菌酯(Fluoxastrobin)、喹唑菌酮(Fluquinconazole)、氟硅唑(Flusilazole)、Fluthianil、粉唑醇(Flutriafol)、氟唑菌酰胺(Fluxapyroxad)、异丙定(Iprodione)、吡唑萘菌胺(Isopyrazam)、精甲霜灵(Mefenoxam)、氯化缩节胺(Mepiquat-chloride)、甲霜灵(Metalaxyl)、环戊唑菌(Metconazole)、叉氨苯酰胺(Metominostrobin)、多效唑(Paclobutrazole)、戊苯吡菌胺(Penflufen)、吡噻菌胺(Penthiopyrad)、啶氧菌酯(Picoxystrobin)、丙氯灵(Prochloraz)、丙硫菌唑(Prothioconazole)、唑菌胺酯(Pyraclostrobin)、氟唑环菌胺(Sedaxane)、戊唑醇(Tebuconazole)、氟醚唑(Tetraconazole)、甲基托布津(Thiophanate-methyl)、福美双(Thiram)、唑菌醇(Triadimenol)、肟菌酯(Trifloxystrobin)、坚强芽孢杆菌(*Bacillus firmus*)、坚强芽孢杆菌(*Bacillus firmus*) 菌株I-1582、枯草芽孢杆菌(*Bacillus subtilis*)、枯草芽孢杆菌菌株GB03、枯草芽孢杆菌菌株QST 713、短小芽孢杆菌(*Bacillus pumilis*)、短小芽孢杆菌菌株GB34。杀昆虫剂:啉虫脒(Acetamiprid)、涕灭威(Aldicarb)、印楝素(Azadirachtin)、克百威(carbofuran)、氯虫苯甲酰胺(Chlorantraniliprole) (Rynaxypyr)、可尼丁(clothianidin)、氰虫酰胺(Cyantraniliprole) (Cyazypyr)、(β-) 氟氯氰菊酯((beta-) Cyfluthrin)、γ-氯氟氰菊酯

(γ -Cyhalothrin)、 λ -氟氯氰菊酯(λ -Cyhalothrin)、氯氰菊酯(Cypermethrin)、溴氰菊酯(Deltamethrin)、乐果(Dimethoate)、呋虫胺(Dinotofuran)、乙虫腈(Ethiprole)、氟啶虫酰胺(Flonicamid)、氟虫双酰胺(Flubendiamide)、氟噻虫砒(Fluensulfone)、氟吡菌酰胺(Fluopyram)、Flupyradifurone、氟胺氰菊酯(tau-fluvalinate)、Imicyafos、吡虫啉(imidacloprid)、氟氟虫脞(Metaflumizone)、甲硫威(Methiocarb)、吡蚜酮(Pymetrozine)、吡氟啶脞(Pyrifluquinazon)、乙基多杀菌素(Spinetoram)、艾克敌(Spinosad)、螺虫乙酯(Spirotetramate)、氟啶虫胺腈(Sulfoxaflo)、噻虫啉(Thiacloprid)、噻虫嗪(Thiamethoxam)、1-(3-氯吡啶-2-基)-N-[4-氰基-2-甲基-6-(甲基氨基甲酰基)苯基]-3-[[5-(三氟甲基)-2H-四唑-2-基]甲基]-1H-吡唑-5-甲酰胺、1-(3-氯吡啶-2-基)-N-[4-氰基-2-甲基-6-(甲基氨基甲酰基)苯基]-3-[[5-(三氟甲基)-1H-四唑-1-基]甲基]-1H-吡唑-5-甲酰胺、1-{2-氟-4-甲基-5-[(2,2,2-三氟乙基)亚磺酰]苯基}-3-(三氟甲基)-1H-1,2,4-三唑-5-胺、(1E)-N-[(6-氯吡啶-3-基)甲基]-N'-氰基-N-(2,2-二氟乙基)乙脞、坚强芽孢杆菌、坚强芽孢杆菌菌株I-1582、枯草芽孢杆菌、枯草芽孢杆菌菌株GB03、枯草芽孢杆菌菌株QST 713、金龟子绿僵菌(Metarhizium anisopliae)F52。

[0210] 在提到本发明的“植物”或“多个植物”时,除非另外说明,否则应当理解其中也涵盖植物部分(细胞、组织或器官、种荚、种子、分离的部分如根、叶、花、花粉等)、仍保留亲本的区别特征的植物后代,如通过自交或杂交获得的种子,例如(通过两个近交亲本系杂交获得的)杂交种子、杂交植物和源自其的植物部分。

[0211] 在一些实施方案中,本发明的植物细胞以及根据本发明的方法产生的植物细胞可以是非繁殖性细胞。

[0212] 根据本发明获得的植物可以在常规育种方案中用于产生具有相同特征的其他植物,或用于在相同植物物种的其它品种或相关植物物种中或在杂交植物中引入相同特征。获得的植物还可以用于产生繁殖材料。本发明的植物还可以用来产生通过本发明方法所获得的植物的配子、种子、胚(合子胚或体细胞胚)、后代或杂种。本发明还涵盖从本发明植物获得的种子。

[0213] 如本文所用,“产生繁殖材料”指本领域已知的用于产生其他植物、植物部分或种子的任何手段,并且尤其包括营养性繁殖方法(例如空中压条法或地面压条法、分株、嫁(芽)接、微繁殖、匍匐茎或纤匐枝、贮藏器官如鳞茎、球茎、块茎和根状茎、插条(striking)或扦插、双鳞片繁殖)、有性繁殖(与另一株植物杂交)和无性繁殖(例如无融合生殖、体细胞杂交)。

[0214] 在某些专利管辖区域,专门通过主要是生物学的过程获得的本发明植物可能不能被授予专利,其中如果植物产生过程完全由天然现象如杂交或选择组成,则认定该过程是主要生物学过程。因此,本发明的植物也涵盖并非专门通过主要生物学过程而获得的那些植物。

[0215] 包含在178千字节(Microsoft Windows®中测量的大小)的文件“BCS18-2016_ST25.txt”中的序列表包含69个序列:SEQ ID NO:1至SEQ ID NO:69,在此以电子提交方式提交并且通过引用的方式并入本文。

[0216] 在说明书和实施例中,述及以下序列:

序列

SEQ ID NO:1:来自亚基因组B的TaRca 2b的核苷酸序列
SEQ ID NO:2:来自亚基因组B的TaRca 2b的氨基酸序列
SEQ ID NO:3:来自亚基因组B的TaRca 2b的核苷酸序列,减去信号肽
SEQ ID NO:4:来自亚基因组B的TaRca 2b的氨基酸序列,减去信号肽
SEQ ID NO:5:来自亚基因组B的TaRca 2a的核苷酸序列
SEQ ID NO:6:来自亚基因组B的TaRca 2a的氨基酸序列
SEQ ID NO:7:来自亚基因组B的TaRca 1b的核苷酸序列
SEQ ID NO:8:来自亚基因组B的TaRca 1b的氨基酸序列
SEQ ID NO:9:来自亚基因组B的TaRca 1b的核苷酸序列,减去信号肽
SEQ ID NO:10:来自亚基因组B的TaRca 1b的氨基酸序列,减去信号肽
SEQ ID NO:11:来自稻 (*Oryza sativa*) 的Rca的氨基酸序列BAA97584.1
SEQ ID NO:12:来自澳洲野生稻 (*Oryza australiensis*) 的Rca的氨基酸序列
ANH11447.1
SEQ ID NO:13:来自 *Larrea tridentate* 的Rca的氨基酸序列Q7X999.1
SEQ ID NO:14:来自小果野蕉 (*Musa acuminata*) 的Rca的氨基酸序列XP_
009419709.1
SEQ ID NO:15:来自头序野麻 (*Datisca glomerata*) 的Rca的氨基酸序列
AAC62207.1
SEQ ID NO:16:来自可可 (*Theobroma cacao*) 的Rca的氨基酸序列E0Y07450.1
SEQ ID NO:17:来自普通烟草 (*Nicotiana tabacum*) 的Rca的氨基酸序列
AAA78277.1
SEQ ID NO:18:来自陆地棉 (*Gossypium hirsutum*) 的Rca的氨基酸序列XP_
016753736.1
SEQ ID NO:19:来自“暖”物种的Rca的氨基酸共有序列
SEQ ID NO:20:来自拟南芥 (*Arabidopsis thaliana*) 的Rca的氨基酸序列NP_
850321.1
SEQ ID NO:21:来自甘蓝 (*Brassica oleracea*) 的Rca的氨基酸序列AFH35543.1
SEQ ID NO:22:来自北美云杉 (*Picea sitchensis*) 的Rca的氨基酸序列
ABK25255.1
SEQ ID NO:23:来自菠菜 (*Spinacia oleracea*) 的Rca的氨基酸序列AAA34038.1
SEQ ID NO:24:来自野草莓 (*Fragaria vesca*) 的Rca的氨基酸序列XP_
004305457.1
SEQ ID NO:25:来自蔓花生 (*Arachis duranensis*) 的Rca的氨基酸序列XP_
015938754.1
SEQ ID NO:26:来自二穗短柄草 (*Brachypodium distachyon*) 的Rca的氨基酸序列
XP_003580722.1
SEQ ID NO:27:来自小花南芥 (*Arabis alpine*) 的Rca的氨基酸序列KFK41750.1
SEQ ID NO:28:来自冰叶日中花 (*Mesembryanthemum crystallinum*) 的Rca的氨基
酸序列AAZ41846.1

SEQ ID NO:29:来自“冷”物种的Rca的氨基酸共有序列

SEQ ID NO:30:来自亚基因组B的TaRca 2b的氨基酸序列,带11个氨基酸置换

SEQ ID NO:31:来自亚基因组B的TaRca 2b的核苷酸序列,带11个氨基酸置换,减去信号肽,密码子优化用于大肠杆菌中表达

SEQ ID NO:32:来自亚基因组B的TaRca 2b的氨基酸序列,带11个氨基酸置换,减去信号肽

SEQ ID NO:33:来自亚基因组B的TaRca 2b的氨基酸序列,带8个氨基酸置换

SEQ ID NO:34:来自亚基因组B的TaRca 2b的核苷酸序列,带8个氨基酸置换,减去信号肽,密码子优化用于大肠杆菌中表达

SEQ ID NO:35:来自亚基因组B的TaRca 2b的氨基酸序列,带8个氨基酸置换,减去信号肽

SEQ ID NO:36:T-DNA Prca0m::Rca 1β的核苷酸序列

SEQ ID NO:37:T-DNA Prbcs::Rca 1β的核苷酸序列

SEQ ID NO:38:来自亚基因组A的TaRca 2b的核苷酸序列

SEQ ID NO:39:来自亚基因组A的TaRca 2b的氨基酸序列

SEQ ID NO:40:来自亚基因组D的TaRca 2b的核苷酸序列

SEQ ID NO:41:来自亚基因组D的TaRca 2b的氨基酸序列

SEQ ID NO:42:来自亚基因组A的TaRca 2a的核苷酸序列

SEQ ID NO:43:来自亚基因组A的TaRca 2a的氨基酸序列

SEQ ID NO:44:来自亚基因组D的TaRca 2a的核苷酸序列

SEQ ID NO:45:来自亚基因组D的TaRca 2a的氨基酸序列

SEQ ID NO:46:来自亚基因组A的TaRca 1b的核苷酸序列

SEQ ID NO:47:来自亚基因组A的TaRca 1b的氨基酸序列

SEQ ID NO:48:来自亚基因组D的TaRca 1b的核苷酸序列

SEQ ID NO:49:来自亚基因组D的TaRca 1b的氨基酸序列

SEQ ID NO:50:TaRca-1b正向引物

SEQ ID NO:51:TaRca-1b反向引物

SEQ ID NO:52:TaRca-2a正向引物

SEQ ID NO:53:TaRca-2a反向引物

SEQ ID NO:54:TaRca-2b正向引物

SEQ ID NO:55:TaRca-2b反向引物

SEQ ID NO:56:Ta54227正向引物

SEQ ID NO:57:Ta54227反向引物

SEQ ID NO:58:Ta54238正向引物

SEQ ID NO:59:Ta54238反向引物

SEQ ID NO:60:T-DNA PubiZm::hpRca2的核苷酸序列

SEQ ID NO:61:向导RNA g1的核苷酸序列

SEQ ID NO:62:向导RNA g2的核苷酸序列

SEQ ID NO:63:向导RNA g13的核苷酸序列

SEQ ID NO:64:向导RNA g9的核苷酸序列
SEQ ID NO:65:向导RNA g14的核苷酸序列
SEQ ID NO:66:向导RNA g15的核苷酸序列
SEQ ID NO:67:向导RNA g16的核苷酸序列
SEQ ID NO:68:向导RNA g17的核苷酸序列
SEQ ID NO:69:向导RNA g18的核苷酸序列

实施例

[0217] 除非在实施例中另外指出,否则全部重组DNA技术按照Sambrook和Russell (2001) *Molecular Cloning: A Laboratory Manual*,第3版,Cold Spring Harbor Laboratory Press, NY中、Ausubel等人(1994) *Current Protocols in Molecular Biology*, *Current Protocols*, USA第1和第2卷中和Brown(1998) *Molecular Biology LabFax*,第2版,Academic Press (UK) 的第I和第II卷中所述的标准方案实施。在BIOS Scientific Publications Ltd (UK) 和Blackwell Scientific Publications, UK联合出版的 *Plant Molecular Biology Labfax* (1993) (R.D.D.Croy) 中描述了用于植物分子操作的标准材料和方法。用于聚合酶链反应的标准材料和方法可以在Dieffenbach和Dveksler (1995) *PCR Primer: A Laboratory Manual*, Cold Spring Harbor Laboratory Press中并在McPherson等人 (2000) *PCR-Basics: From Background to Bench*,第1版,德国Springer Verlag中找到。

实施例1-确定3种小麦Rca同种型的表达谱

原位加热实验

[0218] 将普通小麦栽培品种Fielder (小麦) 种子润湿并在4℃分层置于发芽纸上放置1周,随后在17cm直径带盆栽土壤的花盆中播种。将花盆置于生长箱中,使用12小时光周期,恒定22℃和 $300\mu\text{mol m}^{-2}\text{s}^{-1}$ 光合有效辐射。播种后35天,将一半花盆重新置于38℃光周期温度的毗邻生长箱。在播种后38天,将全部健康叶材料在光周期的中间收获并且在液N₂中急冻,之后储存在-80℃直至使用前。为了测定Rca基因表达,使用SEQ ID NOs:50至59的引物和以下PCR循环:95℃、10分钟;随后95℃15秒、60℃60秒的40个循环;和解链曲线:95℃15秒、60℃60秒、以增量0.3℃从60℃至95℃,以及95℃15秒,对叶材料进行qRT-PCR,并且用参比基因Ta54227和Ta54238归一化。

基因表达结果

[0219] 当小麦暴露于38℃昼间温度历经两个日间循环后,Rca同种型表达存在与22℃对照的偏移。尽管TaRca2- α 同种型的表达保持恒定,但是施加热处理时,TaRca2- β 同种型的表达下降(图1A)。更有意义的是,由独立等位基因编码并且具有与其他剪接同种型高度序列趋异的TaRca1- β 的表达,从对照温度时表达水平不可检出,上升到在热处理下的明显检出。即使如此,TaRca1- β 表达仍明显低于TaRca2剪接变体的表达。然而,似乎热诱导可以TaRca1- β 基因的表达,提示TaRca1- β 蛋白参与小麦的热反应。Rca和Rubisco叶提取和纯化

[0220] 从上文描述的标准生理条件下长出的叶,提取蛋白质。使用液N₂和研钵和研杵,将冷冻的叶组织研磨成细粉。在冰上时,在提取缓冲液中添加叶粉并反复涡旋混合,所述提取缓冲液由100mM Tris pH 8.0、1mM EDTA、7.5mM MgCl₂、2mM DTT、1mM ATP、2%W/V PVPP和蛋白酶抑制剂混合物组成,之后,过单层Miricloth和Lingette Gaze以移除固形物。将样品

在24,000g在4℃离心20分钟并保留上清液。添加35%V/V饱和硫酸铵并且将样品在冰上保持30分钟,之后再次离心。对于Rca纯化,将沉淀物重悬于无PVPP的叶提取缓冲液中,在相同缓冲液中使用Sephadex PD-10脱盐柱脱盐,并采用1ml/分钟流量加载于用脱盐缓冲液平衡的1ml HiTrap Q FF柱上。使用1ml/分钟流速经20ml含有从0至100%的0.5mM KCl的脱盐缓冲液梯度来洗脱Rca,并且合并通过Bradford测定法确定含有蛋白质的级分,使用10kDa截断值Amicon浓缩器(Merck)浓缩到浓度 2.2 ± 0.5 mg/ml,储存在-80℃直至使用前。对于Rubisco纯化,在上文给出的35%V/V饱和硫酸铵析出步骤后,向上清液逐滴添加60%V/V饱和硫酸铵并在4℃缓慢搅拌30分钟,之后再离心。将所得的沉淀物悬浮于100mM Tricine pH 8.0、0.5mM EDTA的样品缓冲液中,并且使用PD-10脱盐柱脱盐入相同的缓冲液。添加20%甘油,并且将样品等分并急冻,储存在-80℃直至使用前。

热处理的小麦Rcas的热稳定性增加

[0221] 为确立观测到的由热引起的Rca表达谱变化是否影响Rca全酶的活性和热稳定性,从叶中提取并分离蛋白质,并且在一系列温度温育10分钟后,在25℃测量以确定热中点(T_m),即一半Rca速度被损害时的温度。就25℃测量的Rca对Rubisco的活化速度(V_{25})而言,相对于对照叶,从热处理的叶提取的Rca具有较低的绝对速率(图1B),提示热处理引起的Rca同种型组成变化降低了Rca的潜在最大速度。但是,对于热处理的植物,Rca T_m 存在小而显著的1℃增加(表2),当Rca速度归一化时温度反应的迁移明确地可观测(图1C)。

实施例2-表征三个小麦Rca同种型每者的热稳定性

重组蛋白生成

[0222] 从头合成全部Rca目的基因(GENEWIZ, South Plainfield, NJ, USA),其中相应于信号肽的N端46个氨基酸缺失,且6氨基酸His标签连接在C末端。将基因连接入Novagen pET-23d+载体(Merck KGaA, 达姆施塔特, 德国),之后遵循标准程序转化入BL21 (DE3) Star大肠杆菌(Escherichia coli)菌株。如Scafaro等人(2016)描述15,进行加His标签的Rca表达和纯化。将最终的Rca蛋白以 2.2 ± 0.5 mg/ml的浓度脱盐至含有20mM Tris pH 8、0.2mM EDTA、7.5mM $MgCl_2$ 、1mM DTT和50mM KCl的缓冲液中,急冻并储存在-80℃直至使用为止。采用牛血清白蛋白(BSA)标准品,使用Protein Assay Dye Reagent Concentrate(Bio-Rad)测定Rca蛋白浓度,并且分别使用针对小麦 α 同种型和 β 同种型的分子量50,954Da和47,110Da及针对稻 β 同种型的分子量47,930Da,计算摩尔浓度。使用IMAGEJ软件(美国国家卫生研究院, Bethesda, MA, USA)通过凝胶图像分析计算Rca浓度时,还考虑分离的Rca样品中的杂质。Rca预温育和Rubisco活化测定

[0223] 在测量Rubisco活化速度之前,添加或不添加0.2mM ATP的5 μ l 2.2 ± 0.5 mg/ml Rca样品置于4Tititude PCR管(BIOKE, Leiden, 荷兰)中,并且在Labcycler(SensoQuest GmBH, 哥廷根, 德国)上加热,其中取决于Rca变体,Labcycler设定在38℃或等同温度,8℃的温度范围跨12个位点。全部管均起初在25℃加热2分钟,随后跨8℃温度范围加热各管10分钟。随后将管以4000g离心5分钟,之后测定酶促速度。遵循Scales等人的ADP非敏感偶联酶分光光度法23,配合以下的修改细节,测量Rca活化Rubisco的能力。全部试剂购自Merck KGaA,除了如先前描述24表达并纯化d-2,3磷酸甘油酸变位酶。测定法按比例缩减到100 μ l反应并且在使用Eppendorf热混合仪(Eppendorf, Hamburg, 德国)加热到25℃的Coster 96孔平底聚苯乙烯平板(Corning, NY, USA)中测量。在一组孔中,添加终体积80 μ l的反应液,所

述反应液组成如下: N₂鼓吹过的MiliQ H₂O、5%W/V PEG-4000、100mM Tricine pH 8、10mM MgCl₂、10mM NaHCO₃、5mM DTT、2.4U/ml烯醇化酶、3.75U/ml磷酸烯醇式丙酮酸羧化酶、6U/ml苹果酸酯脱氢酶、0.2mM 2,3双-磷酸甘油酸、4U/ml d-2,3磷酸甘油酸变位酶、10U/ml碳酸酐酶、4mM磷酸肌酸、20U/ml肌酸磷酸激酶、2mM ATP和0.8mM NADH。在另一组孔中, 20μl终体积由0.25±0.05μM Rubisco活性部位组成, 添加至; 1.活化溶液(N₂鼓吹过的MiliQ H₂O、20mM Tricine pH 8、20mM NaHCO₃和10mM MgCl₂)以测定Rubisco总氨甲酰化活性(ECM), 或2.4mM核酮糖-1,5-二磷酸(RuBP; 纯度99%)用于Rubisco底物抑制(ER)。测量前两分钟, 将4μl预温育的Rca作为与Rubisco溶液分开的液滴, 添加至ER孔。当测量自发基线活性时, 不添加Rca至ER样品。添加Rubisco后10分钟, 通过多通道移液器, 将反应液孔的内容物添加至含有Rubisco的孔, 并且在Infinite M200 Pro平板读数仪(TECAN, Männedorf, 瑞士)上立即每15秒测量波长340nm处吸光度值, 历时8分钟。同时测定多达10份样品。通过Loganathan等人2016²³所述的方法, 计算每分钟由Rca引起的ECM再生反应的定量(mol ECM min⁻¹ × 10⁻³ mol⁻¹ Rca)。通过与前60秒从ECM样品生成的3-磷酸甘油酸(3PG)产物相对应的数据点和25℃小麦Rubisco反应速率常数(K_{cat}) 2.1²⁵的因数, 从线性回归的斜率确定向测定法添加的Rubisco活性部位的量。

温度曲线分析

[0224] 为了确定将Rca的Rubisco活化速度(V)降低一半的中点温度(T_m), 对V相对于温育前温度(T)的曲线, 使用可变斜率模型进行普通最小二乘法拟合:

$$v = \frac{V_{min} + \frac{V_{最大值}}{V_{min}}}{1 + 10^{(T_m - T) \cdot Hillslope}} \quad \text{方程 1}$$

[0225] 其中V_{min}是记录的最慢Rubisco活化速度, V_{max}是最快活化速度, T是以℃计的温育前温度, Hillslope是V下降的陡度。T_m取作V在V_{min}和V_{max}之间中点处的T, 并且Hillslope取作Sign(T_{max}处V-T_{min}处V)。使用Graphpad Prism 5.0软件(GraphPad Prism Software Inc., San Diego, CA, USA)或R编程语言(R Core Team 2017; www.R-project.org)实施全部数据与统计分析。全部实验均重复3-6次, 呈现的所有值和误差线均为均值和标准差(SD)。

小麦Rca 1β同种型比Rca 2同种型更为热稳定

[0226] 重组表达和纯化的Rca同种型用来检查在热处理的叶中观察到的TaRca1-β表达诱导和相关的Rca热稳定性增加是否实际上归因于TaRca1-β是比TaRca2-β或TaRca2-α更为热稳定的Rca变体。使用稻Rca的β同种型(OsRca-β)作为热稳定性的参比, 因为不同于温带小麦, 稻是热带物种。两个TaRca2α和β同种型的T_m无显著差异, 彼此在1℃范围内, 并且一定程度如预期, OsRca-β具有分别高于TaRca2-α剪接变体和TaRca2-β剪接变体7.2℃和7.7℃的T_m(图2, 表2)。有意义地, TaRca1-β同种型的确是小麦的Rca热稳定变体, 并达到这样的程度: 其T_m是42℃, 也高于两种TaRca2剪接型变体7℃并且在稻OsRca-β同种型的1℃范围内。然而TaRca1-β的V₂₅显著小于具有相似V₂₅的其他小麦和稻Rca同种型的速度。

实施例3-热稳定性小麦Rca 2变体的设计

序列比对

[0227] 利用分析软件CLC Main Workbench V.7.8.1(QIAGEN Aarhus A/S)比对全部蛋白质序列。通过比对稻(Oryza sativa) (BAA97584.1)、澳洲野生稻(Oryza australiensis) (ANH11447.1)、Larrea tridentate (Q7X999.1)、小果野蕉(Musa acuminata) (XP_

009419709.1)、头序野麻(*Datisca glomerata*) (AAC62207.1)、可可(*Theobroma cacao*) (E0Y07450.1)、普通烟草(*Nicotiana tabacum*) (AAA78277.1)和陆地棉(*Gossypium hirsutum*) (XP_016753736.1)的Rca蛋白序列,生成暖物种共有序列。从拟南芥(*Arabidopsis thaliana*) (NP_850321.1)、甘蓝(*Brassica oleracea*) (AFH35543.1)、北美云杉(*Picea sitchensis*) (ABK25255.1)、菠菜(*Spinacia oleracea*) (AAA34038.1)、野草莓(*Fragaria vesca*) (XP_004305457.1)、蔓花生(*Arachis duranensis*) (XP_015938754.1)、二穗短柄草(*Brachypodium distachyon*) (XP_003580722.1)、小花南芥(*Arabis alpine*) (KFK41750.1)和冰叶日中花(*Mesembryanthemum crystallinum*) (AAZ41846.1),产生冷物种共有序列。

鉴定与Rcas热稳定性有关的氨基酸

[0228] 使用序列比对作为工具以确定在Rca热稳定性变体中保守的残基,以便可以对热敏感性TaRca2-β进行突变,以人工升高其热稳定性。通过TaRca2-β与热稳定性TaRca1-β和OsRca-β以及冷物种共有序列和暖物种共有序列的蛋白质比对,生成了突变(图3)。通过比对我们分类为冷环境或暖环境特有的物种的Rca序列,生成冷共有序列和暖共有序列。基于TaRca2-β序列和TaRca1-β序列之间的差异,使用标准:TaRca1-β匹配暖物种共有序列并且异于冷物种共有序列,生成的首个TaRca2-β突变体具有11个氨基酸置换(TaRca2-β-11AA)。为了潜在地减少为赋予热稳定性所需要的突变的数目,基于TaRca2-β-11AA,但采用以下额外标准生成了第二突变体:在所述11个氨基酸突变位点处,OsRca-β不能匹配TaRca2-β或冷物种共有序列。这使突变数目减少了三个,导致一个8氨基酸的突变体(TaRca2-β-8AA)。

评价包含产生的Rca 2蛋白变体和Rubisco蛋白的复合体的热稳定性

[0229] Rca活化Rubisco的温度响应曲线显示,TaRca2-β-11AA和TaRca2-β-8AA突变体的确具有比热敏感性TaRca2-β更大的热稳定性(图4,表2)。实际上,TaRca2-β-11AA具有类似于TaRca1-β和OsRca-β热稳定性同种型的温度响应和42.4℃的 T_m ,再次比TaRca2-β高7℃。尽管TaRca2-β-8AA具有改进的热稳定性, T_m 为40℃,但它居于TaRca2-β和其他热稳定性同种型之间,使TaRca2-β T_m 增加5.2℃。这表明,TaRca2-β-11AA突变体的该额外三个残基置换在小麦中在促进Rca的热稳定性方面提供进一步有意义的作用。TaRca2-β-11AA和TaRca2-β-8AA突变体二者均具有与TaRca2-β相似的 V_{25} 值,这表明热稳定性改进并不以突变体酶的动力学活性为代价。

评价产生的Rca 2蛋白变体的热稳定性

[0230] 通过如Niesen等人2007,Nat Protoc.2,2212-2221中所述的示差扫描荧光法(DSF),确定Rca发生解折叠的点。 $10 \pm 0.2 \mu\text{M}$ 的Rca原体在由1:5000稀释的Sypro Orange染料(Invitrogen)、Rca脱盐缓冲液和0.2mM ATP构成的20μl终体积中。将样品在CFX384 qPCR仪(Bio-Rad)上按1℃/分钟从20℃加热至50℃,并且在每个温度在490nm激发荧光并于610nm记录发射(FU)。使用玻尔兹曼S型方程,迭代地拟合普通最小二乘法模型:

$$v = \frac{V[FU]_{\min} + \frac{V[FU]_{\max}}{V[FU]_{\min}}}{1 + \exp(T_m - T) * \text{斜率}} \quad \text{方程 2}$$

[0231] 其中 $V[FU]_{\min}$ 是记录的最小荧光单位, $V[FU]_{\max}$ 是记录的最大荧光单位,T是以℃计的温度并且斜率是上升/下降的陡度。 T_m 取作 $V[FU]$ 处于 $V[FU]_{\min}$ 和 $V[FU]_{\max}$ 之间中点处的T,

并且斜率设在初值2。使用Graphpad Prism 5.0软件(GraphPad Prism Software Inc., San Diego, CA, USA) 或R编程语言(R Core Team 2017; <https://www.R-project.org/>) 实施全部数据与统计分析。全部DSF实验均重复20次或更多次。展示的全部值和误差线均是均值和标准差(SD)。

[0232] 使用差示扫描荧光法(DSF) 确定半数Rca结构完整性受损(即解折叠)的温度(T_m)，产生的结果与确定使Rca速度减半的温度时所获得的结果可比(表2)，但是来自DSF的 T_m 总是比从速度测定法获得的值低2至4℃。对于Fielder对照样品， T_m 无显著差异。使用DSF，热处理的Fielder具有比对照处理显著更高的 T_m 。

[0233] 这些结果还确认，TaRca2- β -11AA和TaRca2- β -8AA突变体的确具有比热敏感性TaRca2- β 野生型更大的热稳定性(图4，表2)。实际上，TaRca2- β -11AA具有类似于TaRca1- β 和OsRca- β 热稳定性同种型的温度响应和40.2℃的 T_m ，再次比TaRca2- β 高7℃。尽管TaRca2- β -8AA具有改进的热稳定性， T_m 为36.1℃，但它居于TaRca2- β 和其他热稳定性同种型之间，增加TaRca2- β T_m 达5.2℃。这表明，TaRca2- β -11AA突变体的该额外三个残基置换在小麦中促进Rca的热稳定性方面提供进一步有意义的作用。

[0234] 表2. 从Rca活化测定法(V) 或差示扫描荧光法(DSF) 计算的热中点，即半数Rca速度或结构稳定性丧失的点(T_m)。值是四个或更多个实验重复的均值±SD。对于选择的变体，在0.2mM ATP存在下进行Rca的温育前加热。上标字母指使用单因素方差分析(one-way ANOVA) 和Tukey多重比较检验，分别分析叶提取物和重组蛋白的情况下，变体之间的显著差异， $p \leq 0.05$ 。

变体	T_m (°C) (V)	T_m (°C) (DSF)
TaRca-leaf-25°C	38.6±0.2 ^a	31.4±0.4 ^a
TaRca-leaf-38°C	39.5±0.5 ^b	32.6±0.4 ^b
TaRca2- α	35.3±0.2 ^c	31.8±0.5 ^a
TaRca2- β	34.8±0.5 ^c	31.8±0.6 ^a
TaRca1- β	42.0±0.2 ^d	40.9±0.7 ^c
OsRca- β	42.5±0.4 ^d	39.8±0.5 ^d
TaRca2- β -11AA	42.4±0.5 ^d	40.2±0.5 ^d
TaRca2- β -8AA	40±0.4 ^b	36.1±0.6 ^c

实施例4-生成内源Rca2蛋白水平降低的小麦植物
生成特异性沉默内源Rca 2基因的构建体

[0235] 使用标准重组DNA技术，在含有bar选择标记盒(SEQ ID NO:60的位置3856至5520)的载体中组装玉蜀黍(Zea mays) 遍在蛋白基因的组成型启动子区(根据SEQ ID NO:60的核苷酸位置157至2153的序列)、靶向来自小麦亚基因组A、B和D的Rca 2基因的发夹DNA片段(根据SEQ ID NO:60的核苷酸位置2162至3543的序列)、和花椰菜花叶病毒35S转录物基因的3'非翻译序列(根据SEQ ID NO:60的核苷酸位置3547至3771的序列)，以产生T-DNA PubiZm::hpRca2(SEQ ID NO:60)。

生成包含上文提到的沉默构建体的转基因小麦植物

[0236] 包含表达盒PubiZm::hpRca2的重组载体用来以Yuji Ishida等人2015, Methods

in Molecular Biology, 1223:189-198中所述的方法稳定转化小麦。选择纯合和无效分离子植物。

生成敲除Rca 2突变体小麦植物

借助诱变

[0237] 通过EMS诱变构建了诱变的小麦群体。基于对Rca 2基因周围的区域测序,鉴定了在来自B亚基因组、来自A亚基因组或来自D亚基因组的Rca 2基因中带敲除突变的突变体植物。取回纯合突变体植物及其野生型分离子。

[0238] 使这类突变体植物杂交,以产生在来自亚基因组A和B、或来自亚基因组A和D或来自亚基因组B和D的Rca 2基因中带敲除突变的双突变体植物。使这类所得的双突变体植物进一步杂交,以产生在来自全部三个亚基因组(即A、B和D)的Rca 2基因中带敲除突变的突变体植物。

借助定向敲除

[0239] 通过使用例如CAS-finder工具,设计了用于CRISPR介导的基因编辑的向导RNA,靶向来自D亚基因组的Rca 2基因的mRNA编码序列、优选地蛋白质编码序列;靶向来自D和A亚基因组的Rca 2基因的mRNA编码序列、优选地蛋白质编码序列;靶向来自A和B亚基因组的Rca 2基因的mRNA编码序列、优选地蛋白质编码序列;或靶向来自A、B和D亚基因组的Rca 2基因的mRNA编码序列、优选地靶向蛋白质编码序列。通过PEG介导向含有Rca 2基因的小麦品系的原生质体瞬时共递送在适宜启动子控制下的gRNA表达载体连同相应核酸酶(例如Cas9或Cpf1)的表达载体,测试了向导RNA的靶向效率。在递送向导RNA和核酸酶载体后,从原生质体提取基因组DNA。在PCR扩增后,通过测序评估靶向的Rca 2基因序列的完整性。

[0240] 最有效的向导RNA用于小麦中的稳定基因编辑。选定的向导RNA是靶向亚基因组A、B和D的g1 (SEQ ID NO:61);靶向亚基因组A和D的g2 (SEQ ID NO:62);靶向亚基因组A和B的g13 (SEQ ID NO:63);和靶向亚基因组D的g9 (SEQ ID NO:64)。为此目的,例如使用粒子枪轰击法,将选定的向导RNA表达载体,与核酸酶表达模块和选择标记基因一起,引入小麦胚胎。使用本领域技术人员已知的方法,再生显示选择剂抗性的转基因植物。通过PCR扩增和测序,鉴定了至少13个含有基因打靶事件,优选地导致无功能Rca 2基因的小缺失或插入,的转基因T0植物。表3中显示获得的敲除突变体的实例。

[0241] 使含有至少一个Rca 2基因的敲除突变(优选地处于纯合态、但备选地处于杂合态)的转基因T0植物杂交,以产生在来自亚基因组A和B、或来自亚基因组A和D或来自亚基因组B和D的Rca 2基因中带敲除突变的植物。使所得的植物进一步杂交,以产生在来自全部三个亚基因组(即A、B和D)的Rca 2基因中带敲除突变的突变体植物。

[0242] 表3

突变系	产生的突变	AA 突变位置	突变蛋白长度
1	在 SEQ ID NO: 38 的位置 268 的核苷酸后插入 C	与 SEQ ID NO: 39 相比, 自氨基酸位置 90 起修饰的序列	165
2	在 SEQ ID NO: 38 的位置 268 的核苷酸后插入 A	与 SEQ ID NO: 39 相比, 自氨基酸位置 268 起修饰的序列	274
3	缺失 SEQ ID NO: 1 核苷酸 97 至 101	与 SEQ ID NO: 2 相比, 自氨基酸位置 33 起修饰的序列	199
4	缺失 SEQ ID NO: 1 核苷酸 100 和 101	与 SEQ ID NO: 2 相比, 自氨基酸位置 34 起修饰的序列	200
5	缺失 SEQ ID NO: 1 核苷酸 102 至 110	与 SEQ ID NO: 2 相比, 自氨基酸位置 34 起修饰的序列	33
6	在 SEQ ID NO: 1 的位置 101 的核苷酸后插入 A	与 SEQ ID NO: 2 相比, 自氨基酸位置 34 起修饰的序列	33
7	在 SEQ ID NO: 1 的位置 800 的核苷酸后插入 A	与 SEQ ID NO: 2 相比, 自氨基酸位置 268 起修饰的序列	274
8	缺失 SEQ ID NO: 1 核苷酸 803 和 804	与 SEQ ID NO: 2 相比, 自氨基酸位置 268 起修饰的序列	273
9	缺失 SEQ ID NO: 40 核苷酸 100 和 101	与 SEQ ID NO: 41 相比, 自氨基酸位置 34 起修饰的序列	164
10	缺失 SEQ ID NO: 40 核苷酸 270 至 285	与 SEQ ID NO: 41 相比, 自氨基酸位置 91 起修饰的序列	134
11	缺失 SEQ ID NO: 40 的核苷酸 267	与 SEQ ID NO: 41 相比, 自氨基酸位置 90 起修饰的序列	139
12	在 SEQ ID NO: 40 的位置 299 的核苷酸后插入 A	与 SEQ ID NO: 41 相比, 自氨基酸位置 100 起修饰的序列	165
13	缺失 SEQ ID NO: 40 的核苷酸 299	与 SEQ ID NO: 41 相比, 自氨基酸位置 100 起修饰的序列	139

实施例5-生成耐热性增加的小麦植物

生成具有小麦热稳定性Rca的表达构建体

[0243] 使用标准重组DNA技术,在含有bar选择标记盒(SEQ ID NO:36的位置2537至4201)的载体中,组装南方野生稻(*Oryza meridionalis*) Rubisco活化酶基因的启动子区和5' UTR (根据SEQ ID NO:36的核苷酸位置81至880的序列)、编码小麦Rca 1 β 的DNA片段(根据SEQ ID NO:36的核苷酸位置890至2188的序列),和南方野生稻Rubisco活化酶基因的3'非翻译序列(根据SEQ ID NO:36的核苷酸位置2203至2462的序列),以产生T-DNA Prca0m::Rca 1 β (SEQ ID NO:36)。

[0244] 使用标准重组DNA技术,在含有bar选择标记盒(SEQ ID NO:37的位置4971至6635)的载体中,组装稻(*Oryza sativa*) Rubisco小亚基基因的启动子区(根据SEQ ID NO:37的核苷酸位置75至2821的序列)、稻肌动蛋白1基因的第一内含子(根据SEQ ID NO:37的核苷酸位置2825至3286的序列)、编码小麦Rca 1 β 的DNA片段(根据SEQ ID NO:37的核苷酸位置3294至4592的序列)、和根癌农杆菌(*Agrobacterium tumefaciens*) 胭脂碱合酶基因的3'非翻译序列(根据SEQ ID NO:37的核苷酸位置4630至4890的序列),以产生T-DNA Prbcs::Rca 1 β (SEQ ID NO:37)。

[0245] 从头合成和设计DNA分子i)以编码根据SEQ ID 30或33的多肽,ii)以优化核苷酸序列以便在小麦植物细胞中表达和iii)以避免它被上文描述的发夹构建体靶向沉默。为此目的,考虑诸多因素如密码子选择、mRNA二级结构、AT含量、隐匿性剪接位点或限制性位点。使用标准重组DNA技术,在载体中用植物可表达启动子和转录终止子组装这些DNA分子每

者。

包含含有小麦热稳定性Rca的表达构建体的转基因小麦植物的生成

[0246] 包含表达盒Prca0m::Rca 1 β 和Prbcs::Rca 1 β 的重组载体用来以Yuji Ishida等人2015,Methods in Molecular Biology,1223:189-198中所述的方法稳定转化小麦。已经选择纯合和无效分离子植物。

[0247] 包含带Rca 2 β 变体的表达盒的重组载体用来以Yuji Ishida等人2015,Methods in Molecular Biology,1223:189-198中所述的方法稳定转化小麦。选择纯合和无效分离子植物。

[0248] 获得的转基因小麦植物随后与其中沉默或敲除了内源Rca 2基因的上述小麦系(参见实施例4)杂交,并且选出表达纯合和无效分离子的株系。

实施例6-生成包含热稳定性Rca 2等位基因的小麦植物

[0249] 在Rca2中跨约1.2kb区域引入11个AA置换的方法基于如下进行:在紧邻第一氨基酸置换(AA1)的靶位点1(TS1)处以及在紧邻最末氨基酸置换(AA2)的靶位点2(TS2)处同时诱导DBS,并通过同源重组,将约1.2kb天然Rca2序列替换为含有11AA的约1.2kb Rca2突变序列。因而,已经开发一种修复DNA,其包含旁侧有与TS1紧上游和TS2紧下游的序列同源的区域的该约1.2kb Rca2突变序列,具有在gRNA靶位点上的沉默突变以防止修复DNA被切割(图6)。

[0250] 为了鉴定可以同时受切割的两个靶位点(TS1和TS2),通过PEG介导的转染,向小麦原生质体共递送Cas9核酸酶和靶向TS1和TS2的sgRNA对。同时切割TS1和TS2可以导致缺失该2个靶位点之间的约1.2kb区域。通过使用分别在TS1上游和TS2下游的引物进行PCR扩增,观察到与缺失了该约1.2kb区域时预期大小相应的较小PCR片段。表3显示与Cas9核酸酶共递送时同时切割其相应靶位点TS1和TS2并由此导致缺失约1.2kb Rca2区域的gRNA对。

表4

gRNA对	亚基因组	TS1序列	TS2序列
g1-g14	A,B,D	SEQ ID NO:61	SEQ ID NO:65
g1-g15	A,B,D	SEQ ID NO:61	SEQ ID NO:66
g1-g16	A,B,D	SEQ ID NO:61	SEQ ID NO:67
g1-g17	A,B,D	SEQ ID NO:61	SEQ ID NO:68
g9-g18	D	SEQ ID NO:64	SEQ ID NO:69

[0251] 已经通过粒子轰击法将Cas9基因、用于靶位点TS1和TS2处切割的2个sgRNA和修复DNA引入不成熟胚胎的盾片细胞。已经使用无选择过程的组织培养方法或通过基于选择的组织培养方法,通过共递送Cas9、2个gRNA和修复DNA以及选择标记基因(例如epsp^s、bar),从轰击过的胚胎获得再生的植物。可以通过后代分离,从含有Rca2 11AA置换的替换事件中移除选择标记基因。

序列表

<110> 巴斯夫欧洲公司 (BASF SE)	
<120> 热稳定性 RUBISCO 活化酶及其用途	
<130> BCS18-2016	
<160> 69	
<170> PatentIn version 3.5	
<210> 1	
<211> 1284	
<212> DNA	
<213> 普通小麦 (Triticum aestivum)	
<400> 1	
atggetgctg ctttctctc caccgtcggg gccccggctt ctacgccgac caacttcctt	60
gggaagaagc tcaagaagca ggtgacctcg gccgtgaact accatggcaa gagctccaag	120
gccaacaggt tcacagtcac ggcagcggaa aacctcgacg agaagaggaa cacggacaag	180
tggaagggtc ttgcgtacga tatctccgac gaccagcagg acatcaccag aggggaaggc	240
atcgtggact cgctcttcca ggcgcccacg ggcgacggca cccacgaggc cgtcctcagc	300
tcctacagat acgtcagcca gggactcaag aagtacgact tcgacaacac catgggaggc	360
ttctacatcg ctcttgcttt catggacaag ctgtttgtcc atctctccaa gaacttcattg	420
accctgcccc acatcaagat cccactcacc ttgggtatct ggggaggcaa ggggtcaagga	480
aaatcattcc agtgcgagct tgtcttcgcc aagatgggca tcaaccaat catgatgagt	540
[0001] gccggagagc tggagagtgg caacgccgga gagccagcca agctcatcag gcagcggtag	600
cgtgaggctg cagacatgat caagaagggt aagatgtgct gcctcttcat caacgatctt	660
gacgtgggtg cgggtcggat gggcgggacc acacagtaca ccgtcaacaa ccagatgggtg	720
aacgccaccc tcatgaacat cgccgatgcc cccaccaacg tgcagctccc aggcattgtac	780
aacaaggagg agaaccctcg tgtgcccacc gtcgtcactg gtaacgattt ctgcagttg	840
tacgcccctc tgatccgtga tggccgtatg gagaagttct actgggtccc caccgcgac	900
gaccgtatcg gtgtctgcaa gggatcttcc cagaccgaca atgtcagcga cgagtccgtc	960
gtcaagatcg tcgacacctt cccaggacaa tccatcgact ttttcgggtg tctgcgtgct	1020
cgggtgtacg acgacgaggt ggcgaagtgg gtgacctcta ccggtatcga gaacattggc	1080
aagaggtggt tgaactcgcg ggacggaccc gtgacctttg agcagccaaa gatgacagt	1140
gagaagctgc tagagtacgg gcacatgctt gtccaggagc aggacaatgt caagcgtgtg	1200
cagcttgctg acacctacat gagccaggca gctctgggtg atgctaacca ggatgcgatg	1260
aagactgggt ctttctacgg ttag	1284
<210> 2	
<211> 427	
<212> PRT	
<213> 普通小麦 (Triticum aestivum)	
<400> 2	
Met Ala Ala Ala Phe Ser Ser Thr Val Gly Ala Pro Ala Ser Thr Pro	
1 5 10 15	
Thr Asn Phe Leu Gly Lys Lys Leu Lys Lys Gln Val Thr Ser Ala Val	
20 25 30	

Asn Tyr His Gly Lys Ser Ser Lys Ala Asn Arg Phe Thr Val Met Ala		
35	40	45
Ala Glu Asn Leu Asp Glu Lys Arg Asn Thr Asp Lys Trp Lys Gly Leu		
50	55	60
Ala Tyr Asp Ile Ser Asp Asp Gln Gln Asp Ile Thr Arg Gly Lys Gly		
65	70	75
Ile Val Asp Ser Leu Phe Gln Ala Pro Thr Gly Asp Gly Thr His Glu		
85	90	95
Ala Val Leu Ser Ser Tyr Glu Tyr Val Ser Gln Gly Leu Lys Lys Tyr		
100	105	110
Asp Phe Asp Asn Thr Met Gly Gly Phe Tyr Ile Ala Pro Ala Phe Met		
115	120	125
Asp Lys Leu Val Val His Leu Ser Lys Asn Phe Met Thr Leu Pro Asn		
130	135	140
Ile Lys Ile Pro Leu Ile Leu Gly Ile Trp Gly Gly Lys Gly Gln Gly		
145	150	155
Lys Ser Phe Gln Cys Glu Leu Val Phe Ala Lys Met Gly Ile Asn Pro		
165	170	175
Ile Met Met Ser Ala Gly Glu Leu Glu Ser Gly Asn Ala Gly Glu Pro		
180	185	190
Ala Lys Leu Ile Arg Gln Arg Tyr Arg Glu Ala Ala Asp Met Ile Lys		
195	200	205
[0002] Lys Gly Lys Met Cys Cys Leu Phe Ile Asn Asp Leu Asp Ala Gly Ala		
210	215	220
Gly Arg Met Gly Gly Thr Thr Gln Tyr Thr Val Asn Asn Gln Met Val		
225	230	235
Asn Ala Thr Leu Met Asn Ile Ala Asp Ala Pro Thr Asn Val Gln Leu		
245	250	255
Pro Gly Met Tyr Asn Lys Glu Glu Asn Pro Arg Val Pro Ile Val Val		
260	265	270
Thr Gly Asn Asp Phe Ser Thr Leu Tyr Ala Pro Leu Ile Arg Asp Gly		
275	280	285
Arg Met Glu Lys Phe Tyr Trp Ala Pro Thr Arg Asp Asp Arg Ile Gly		
290	295	300
Val Cys Lys Gly Ile Phe Gln Thr Asp Asn Val Ser Asp Glu Ser Val		
305	310	315
Val Lys Ile Val Asp Thr Phe Pro Gly Gln Ser Ile Asp Phe Phe Gly		
325	330	335
Ala Leu Arg Ala Arg Val Tyr Asp Asp Glu Val Arg Lys Trp Val Thr		
340	345	350
Ser Thr Gly Ile Glu Asn Ile Gly Lys Arg Leu Val Asn Ser Arg Asp		
355	360	365
Gly Pro Val Thr Phe Glu Gln Pro Lys Met Thr Val Glu Lys Leu Leu		
370	375	380
Glu Tyr Gly His Met Leu Val Gln Glu Gln Asp Asn Val Lys Arg Val		

	385	390	395	400
	Gln Leu Ala Asp Thr Tyr Met Ser Gln Ala Ala Leu Gly Asp Ala Asn			
		405	410	415
	Gln Asp Ala Met Lys Thr Gly Ser Phe Tyr Gly			
	420	425		
	<210> 3			
	<211> 1284			
	<212> DNA			
	<213> 普通小麦(Triticum aestivum)			
	<400> 3			
	atggetgctg ctttctctc caccgtcggt gccccggtt ctacgccgac caacttcctt			60
	gggaagaagc tcaagaagca ggtgacctcg gccgtgaact accatggcaa gagtccaag			120
	gccaacaggt tcacagtcac ggcagcggaa aacctcgacg agaagaggaa cacggacaag			180
	tggaagggtc ttgcgtacga tatctccgac gaccagcagg acatcaccag agggaagggc			240
	atcgtggact cgctcttcca ggcgcccacg ggcgacggca cccacgaggc cgtcctcagc			300
	tcctacgagt acgtcagcca gggactcaag aagtacgact tcgacaacac catgggaggc			360
	ttctacatcg ctctctgttt catggacaag cttgttgctc atctctccaa gaacttcattg			420
	acctgccca acatcaagat cccactcatc ttgggtatct ggggaggcaa gggtaagga			480
	aaatcattcc agtgcgagct tgtcttcgcc aagatgggca tcaacccaat catgatgagt			540
	gccggagagc tggagagtgg caacgccgga gagccagcca agctcatcag gcagcggtag			600
	cgtgaggctg cagacatgat caagaagggt aagatgtgct gcctcttcat caacgatctt			660
	gacgtggtg cgggtcggat gggcgggacc acacagtaca ccgtcaacaa ccagatggtg			720
[0003]	aacgccaccc tcatgaacat cgccgatgcc cccaccaacg tgcagctccc aggcattgac			780
	aacaaggagg agaaccctcg tgtgcccacg gtcgtcactg gtaacgattt ctcgacgttg			840
	tacccccctc tgatccgtga tggccgtatg gagaagttct actgggctcc caccgcgcac			900
	gaccgtatcg gtgtctgcaa gggatcttcc cagaccgaca atgtcagcga cgagtccgtc			960
	gtcaagatcg tcgacacctt cccaggacaa tccatcgact ttttcgggtg tctgcgtgct			1020
	cgggtgtacg acgacgaggt gcgcaagtgg gtgacctcta ccggtatcga gaacattggc			1080
	aagaggctgg tgaactcgcg ggacggaccc gtgacctttg agcagccaaa gatgacagtg			1140
	gagaagctgc tagagtacgg gcacatgctt gtccaggagc aggacaatgt caagcgtgtg			1200
	cagcttgctg acacctacat gagccaggca gctctgggtg atgctaacca ggatgcgatg			1260
	aagactggtt cttctacgg ttag			1284
	<210> 4			
	<211> 381			
	<212> PRT			
	<213> 普通小麦(Triticum aestivum)			
	<400> 4			
	Met Ala Ala Glu Asn Leu Asp Glu Lys Arg Asn Thr Asp Lys Trp Lys			
	1 5 10 15			
	Gly Leu Ala Tyr Asp Ile Ser Asp Asp Gln Gln Asp Ile Thr Arg Gly			
	20 25 30			
	Lys Gly Ile Val Asp Ser Leu Phe Gln Ala Pro Thr Gly Asp Gly Thr			
	35 40 45			
	His Glu Ala Val Leu Ser Ser Tyr Glu Tyr Val Ser Gln Gly Leu Lys			
	50 55 60			

86

	atggctgctg cttctcctc caccgtcggt gccccggctt ctacgccgac caacttcctt	60
	gggaagaagc tcaagaagca ggtgacctcg gccgtgaact accatggcaa gagctccaag	120
	gccaacaggt tcacagtcac ggcagcggaa aacctcgacg agaagaggaa cacggacaag	180
	tggaagggtc ttgcgtacga tatctccgac gaccagcagg acatcaccag aggggaagggc	240
	atcgtggact cgctcttcca ggcgccacg ggcgacggca cccacgaggc cgctctcagc	300
	tcctacgagt acgtcagcca gggactcaag aagtagcact tcgacaacac catgggaggc	360
	ttctacatcg ctctctgttt catggacaag cttgttgctc atctctccaa gaacttcag	420
	acctgccca acatcaagat cccactcatc ttgggtatct ggggaggcaa gggtaagga	480
	aaatcattcc agtgcgagct tgtcttcgcc aagatgggca tcaaccaat catgatgagt	540
	gccggagagc tggagagtgg caacgccgga gagccagcca agctcatcag gcagcggtag	600
	cgtgaggctg cagacatgat caagaagggt aagatgtgct gcctcttcat caacgatctt	660
	gacgtggtg cgggtcggat gggcgggacc acacagtaca ccgtcaacaa ccagatggtg	720
	aagccaccc tcatgaacat cgccgatgcc cccaccaacg tgcagctccc aggcagtac	780
	aacaaggagg agaacctcg tgtgccatc gtctcactg gtaacgattt ctgcagttg	840
	tacgccctc tgatccgtga tggccgtatg gagaagtctt actgggctcc caccgcgac	900
	gaccgtatcg gtgtctgcaa gggatatctt cagaccgaca atgtcagcga cgagtcgctc	960
	gtcaagatcg tcgacacctt cccaggacaa tccatcgact ttttcgggtc tctgcgtgct	1020
	cgggtgtacg acgacgaggt gcgcaagtgg gtgacctcta ccggtatcga gaacattggc	1080
	aagaggctgg tgaactcgcg ggacggaccc gtgacctttg agcagccaaa gatgacagt	1140
	gagaagctgc tagagtacgg gcacatgctt gtccaggagc aggacaatgt caagcgtgtg	1200
	cagcttgctg acacctacat gagccaggca gctctgggtg atgctaacca ggatgcgatg	1260
[0005]	aagactggtt cttctacgg taaaggggca cagcaagta ctttgctgt accggcagga	1320
	tgcaccgacc agactgcaa gaacttcgac ccaacggcga ggagtgcga cgggagctgc	1380
	ctttacacct tttaa	1395
	<210> 6	
	<211> 464	
	<212> PRT	
	<213> 普通小麦(Triticum aestivum)	
	<400> 6	
	Met Ala Ala Ala Phe Ser Ser Thr Val Gly Ala Pro Ala Ser Thr Pro	
	1 5 10 15	
	Thr Asn Phe Leu Gly Lys Lys Leu Lys Lys Gln Val Thr Ser Ala Val	
	20 25 30	
	Asn Tyr His Gly Lys Ser Ser Lys Ala Asn Arg Phe Thr Val Met Ala	
	35 40 45	
	Ala Glu Asn Leu Asp Glu Lys Arg Asn Thr Asp Lys Trp Lys Gly Leu	
	50 55 60	
	Ala Tyr Asp Ile Ser Asp Asp Gln Gln Asp Ile Thr Arg Gly Lys Gly	
	65 70 75 80	
	Ile Val Asp Ser Leu Phe Gln Ala Pro Thr Gly Asp Gly Thr His Glu	
	85 90 95	
	Ala Val Leu Ser Ser Tyr Glu Tyr Val Ser Gln Gly Leu Lys Lys Tyr	
	100 105 110	
	Asp Phe Asp Asn Thr Met Gly Gly Phe Tyr Ile Ala Pro Ala Phe Met	

115	120	125
Asp Lys Leu Val Val His Leu Ser Lys Asn Phe Met Thr Leu Pro Asn		
130	135	140
Ile Lys Ile Pro Leu Ile Leu Gly Ile Trp Gly Gly Lys Gly Gln Gly		
145	150	155
Lys Ser Phe Gln Cys Glu Leu Val Phe Ala Lys Met Gly Ile Asn Pro		
165	170	175
Ile Met Met Ser Ala Gly Glu Leu Glu Ser Gly Asn Ala Gly Glu Pro		
180	185	190
Ala Lys Leu Ile Arg Gln Arg Tyr Arg Glu Ala Ala Asp Met Ile Lys		
195	200	205
Lys Gly Lys Met Cys Cys Leu Phe Ile Asn Asp Leu Asp Ala Gly Ala		
210	215	220
Gly Arg Met Gly Gly Thr Thr Gln Tyr Thr Val Asn Asn Gln Met Val		
225	230	235
Asn Ala Thr Leu Met Asn Ile Ala Asp Ala Pro Thr Asn Val Gln Leu		
245	250	255
Pro Gly Met Tyr Asn Lys Glu Glu Asn Pro Arg Val Pro Ile Val Val		
260	265	270
Thr Gly Asn Asp Phe Ser Thr Leu Tyr Ala Pro Leu Ile Arg Asp Gly		
275	280	285
Arg Met Glu Lys Phe Tyr Trp Ala Pro Thr Arg Asp Asp Arg Ile Gly		
290	295	300
Val Cys Lys Gly Ile Phe Gln Thr Asp Asn Val Ser Asp Glu Ser Val		
305	310	315
Val Lys Ile Val Asp Thr Phe Pro Gly Gln Ser Ile Asp Phe Phe Gly		
325	330	335
Ala Leu Arg Ala Arg Val Tyr Asp Asp Glu Val Arg Lys Trp Val Thr		
340	345	350
Ser Thr Gly Ile Glu Asn Ile Gly Lys Arg Leu Val Asn Ser Arg Asp		
355	360	365
Gly Pro Val Thr Phe Glu Gln Pro Lys Met Thr Val Glu Lys Leu Leu		
370	375	380
Glu Tyr Gly His Met Leu Val Gln Glu Gln Asp Asn Val Lys Arg Val		
385	390	395
Gln Leu Ala Asp Thr Tyr Met Ser Gln Ala Ala Leu Gly Asp Ala Asn		
405	410	415
Gln Asp Ala Met Lys Thr Gly Ser Phe Tyr Gly Lys Gly Ala Gln Gln		
420	425	430
Gly Thr Leu Pro Val Pro Ala Gly Cys Thr Asp Gln Thr Ala Lys Asn		
435	440	445
Phe Asp Pro Thr Ala Arg Ser Asp Asp Gly Ser Cys Leu Tyr Thr Phe		
450	455	460
<210> 7		
<211> 1299		

	<212> DNA	
	<213> 普通小麦(<i>Triticum aestivum</i>)	
	<400> 7	
	atggtctctg ctttctcgtc caccgttga gctccggcgt caaccccgac caccttctc	60
	gggaagaagg tgaagaagca ggccggtgcg ttgaactact accatggtgg caacaagatc	120
	aagagtaagg tggtcagggc catggcggcc aaaaaggaac ttgacgaggg caagcagacc	180
	aatgccgata ggtggaaggg tctcgttac gacatctccg acgaccagca ggacatcacg	240
	agtgggaaag gtatcgtaga ctccctgttc caggccccc tgggcgacgg caccacgag	300
	gccatcctga gctcctacga gtacatcagc cagggcctgc ggaagtacga cttcgacaac	360
	accatggacg gcctctacat cgccccggca ttcatggaca agctcatcgt ccacctcgcc	420
	aagaacttca tgacactccc caacatcaag gttcctctca tcctgggcat ctggggaggc	480
	aagggacagg gcaagtcgtt ccagtgcgag ctggtgttcg ccaagatggg catcaacccc	540
	atcatgatga gcgccggtga gctggagagc ggcaacgccg gcgagccggc gaagctgac	600
	cggcagaggt accgcgaggc tgccgacatt atcaacaagg gcaagatgtg ctgcctcttc	660
	atcaacgacc tggacgccg cgcgggggcg atgggcggga cgacgcagta cacggtgaac	720
	aaccagatgg tgaacgccac cctgatgaac atcgcgagc cgcccaccaa cgtgcagctc	780
	cccggtgatgt acaacaagga ggagaacccc cgctgcccc tcatcgtcac gggcaacgac	840
	ttctcaacgc tgtacgcgcc cctgatccgc gatggccgca tggagaagtt ctactgggcg	900
	cccaccgcg aggaccgcat cggcgtgtgc aaggcctct tccgcaccga caacgtcccc	960
	gacgaggccg tggtaggct ggtggacacc ttcccgggc agtccatcga cttcttcggc	1020
	gcgctgcggg cgcggtgtga cgacgacgag gtgcgcaagt gggtcggcga gatcgcgctg	1080
	gagaacatct ccaagcggct cgtcaactcc agggaggggc cgccgacgtt cgaccagccc	1140
[0007]	aagatgacca tcgagaagct catggagtac ggccacatgc tgggtccagga gcaggagaac	1200
	gtgaagcgcg tgcagctcgc cgacaagtac ctcagcgagg cggcgctcgg ccaagccaac	1260
	gacgacgcca tgaagaccgg cgccttctac ggcaagtag	1299
	<210> 8	
	<211> 432	
	<212> PRT	
	<213> 普通小麦(<i>Triticum aestivum</i>)	
	<400> 8	
	Met Ala Ser Ala Phe Ser Ser Thr Val Gly Ala Pro Ala Ser Thr Pro	
	1 5 10 15	
	Thr Thr Phe Leu Gly Lys Lys Val Lys Lys Gln Ala Gly Ala Leu Asn	
	20 25 30	
	Tyr Tyr His Gly Gly Asn Lys Ile Lys Ser Lys Val Val Arg Ala Met	
	35 40 45	
	Ala Ala Lys Lys Glu Leu Asp Glu Gly Lys Gln Thr Asn Ala Asp Arg	
	50 55 60	
	Trp Lys Gly Leu Ala Tyr Asp Ile Ser Asp Asp Gln Gln Asp Ile Thr	
	65 70 75 80	
	Ser Gly Lys Gly Ile Val Asp Ser Leu Phe Gln Ala Pro Met Gly Asp	
	85 90 95	
	Gly Thr His Glu Ala Ile Leu Ser Ser Tyr Glu Tyr Ile Ser Gln Gly	
	100 105 110	
	Leu Arg Lys Tyr Asp Phe Asp Asn Thr Met Asp Gly Leu Tyr Ile Ala	

115 120 125
 Pro Ala Phe Met Asp Lys Leu Ile Val His Leu Ala Lys Asn Phe Met
 130 135 140
 Thr Leu Pro Asn Ile Lys Val Pro Leu Ile Leu Gly Ile Trp Gly Gly
 145 150 155 160
 Lys Gly Gln Gly Lys Ser Phe Gln Cys Glu Leu Val Phe Ala Lys Met
 165 170 175
 Gly Ile Asn Pro Ile Met Met Ser Ala Gly Glu Leu Glu Ser Gly Asn
 180 185 190
 Ala Gly Glu Pro Ala Lys Leu Ile Arg Gln Arg Tyr Arg Glu Ala Ala
 195 200 205
 Asp Ile Ile Asn Lys Gly Lys Met Cys Cys Leu Phe Ile Asn Asp Leu
 210 215 220
 Asp Ala Gly Ala Gly Arg Met Gly Gly Thr Thr Gln Tyr Thr Val Asn
 225 230 235 240
 Asn Gln Met Val Asn Ala Thr Leu Met Asn Ile Ala Asp Ala Pro Thr
 245 250 255
 Asn Val Gln Leu Pro Gly Met Tyr Asn Lys Glu Glu Asn Pro Arg Val
 260 265 270
 Pro Ile Ile Val Thr Gly Asn Asp Phe Ser Thr Leu Tyr Ala Pro Leu
 275 280 285
 Ile Arg Asp Gly Arg Met Glu Lys Phe Tyr Trp Ala Pro Thr Arg Glu
 [0008] 290 295 300
 Asp Arg Ile Gly Val Cys Lys Gly Ile Phe Arg Thr Asp Asn Val Pro
 305 310 315 320
 Asp Glu Ala Val Val Arg Leu Val Asp Thr Phe Pro Gly Gln Ser Ile
 325 330 335
 Asp Phe Phe Gly Ala Leu Arg Ala Arg Val Tyr Asp Asp Glu Val Arg
 340 345 350
 Lys Trp Val Gly Glu Ile Gly Val Glu Asn Ile Ser Lys Arg Leu Val
 355 360 365
 Asn Ser Arg Glu Gly Pro Pro Thr Phe Asp Gln Pro Lys Met Thr Ile
 370 375 380
 Glu Lys Leu Met Glu Tyr Gly His Met Leu Val Gln Glu Gln Glu Asn
 385 390 395 400
 Val Lys Arg Val Gln Leu Ala Asp Lys Tyr Leu Ser Glu Ala Ala Leu
 405 410 415
 Gly Gln Ala Asn Asp Asp Ala Met Lys Thr Gly Ala Phe Tyr Gly Lys
 420 425 430
 <210> 9
 <211> 1158
 <212> DNA
 <213> 普通小麦(*Triticum aestivum*)
 <400> 9
 atggcggcca aaaaggaact tgacgagggc aagcagacca atgccgatag gtggaagggt

60

ctcgtttagc acatctccga cgaccagcag gacatcacga gtgggaaagg tatcgtagac	120
tccctgttcc aggcccccat gggcgacggc acccacgagg ccatcctgag ctctacgag	180
tacatcagcc agggcctgcg gaagtacgac ttcgacaaca ccatggacgg cctctacatc	240
gccccggcat tcatggacaa gctcatcgtc cacctcgcca agaacttcat gacactcccc	300
aacatcaagg ttctctctcat cctgggcatc tggggaggca agggacaggg caagtcgttc	360
cagtgcgagc tgggtgttcgc caagatgggc atcaacccca tcatgatgag cgccggtgag	420
ctggagagcg gcaacgccgg cgagccggcg aagctgatcc ggcagaggta ccgcgagget	480
gccgacatta tcaacaaggg caagatgtgc tgcctcttca tcaacgacct ggacgccggc	540
gcggggcgga tgggcgggac gacgcagtac acggtgaaca accagatggt gaacgccacc	600
ctgatgaaca tcgcggacgc gccccaac gtgcagctcc ccgggatgta caacaaggag	660
gagaaccccc gcgtgcccat catcgtcacg ggcaacgact tctcaacgct gtacgcgcc	720
ctgatccgcg atggccgcat ggagaagttc tactgggcgc ccaccgcga ggaccgcatc	780
ggcgtgtgca agggcatctt ccgcaccgac aacgtccccg acgaggccgt ggtgaggetg	840
gtggacacct tcccggggca gtccatcgac ttcttcggcg cgctgcgggc gcgggtgtac	900
gacgacgagg tgcgcaagtg ggtcggcgag atcggcgtgg agaacatctc caagcggtc	960
gtcaactcca gggaggggccc gccgacgttc gaccagccca agatgacat cgagaagctc	1020
atggagtacg gccacatgct ggtccaggag caggagaacg tgaagcgcgt gcagctcgcc	1080
gacaagtacc tcagcgaggc ggcgctcggc caagccaacg acgacgcat gaagaccggc	1140
gccttctacg gcaagtag	1158
<210> 10	
<211> 385	
<212> PRT	
[0009] <213> 普通小麦(Triticum aestivum)	
<400> 10	
Met Ala Ala Lys Lys Glu Leu Asp Glu Gly Lys Gln Thr Asn Ala Asp	
1 5 10 15	
Arg Trp Lys Gly Leu Ala Tyr Asp Ile Ser Asp Asp Gln Gln Asp Ile	
20 25 30	
Thr Ser Gly Lys Gly Ile Val Asp Ser Leu Phe Gln Ala Pro Met Gly	
35 40 45	
Asp Gly Thr His Glu Ala Ile Leu Ser Ser Tyr Glu Tyr Ile Ser Gln	
50 55 60	
Gly Leu Arg Lys Tyr Asp Phe Asp Asn Thr Met Asp Gly Leu Tyr Ile	
65 70 75 80	
Ala Pro Ala Phe Met Asp Lys Leu Ile Val His Leu Ala Lys Asn Phe	
85 90 95	
Met Thr Leu Pro Asn Ile Lys Val Pro Leu Ile Leu Gly Ile Trp Gly	
100 105 110	
Gly Lys Gly Gln Gly Lys Ser Phe Gln Cys Glu Leu Val Phe Ala Lys	
115 120 125	
Met Gly Ile Asn Pro Ile Met Met Ser Ala Gly Glu Leu Glu Ser Gly	
130 135 140	
Asn Ala Gly Glu Pro Ala Lys Leu Ile Arg Gln Arg Tyr Arg Glu Ala	
145 150 155 160	

	Ala Asp Ile Ile Asn Lys Gly Lys Met Cys Cys Leu Phe Ile Asn Asp	
	165	170 175
	Leu Asp Ala Gly Ala Gly Arg Met Gly Gly Thr Thr Gln Tyr Thr Val	
	180	185 190
	Asn Asn Gln Met Val Asn Ala Thr Leu Met Asn Ile Ala Asp Ala Pro	
	195	200 205
	Thr Asn Val Gln Leu Pro Gly Met Tyr Asn Lys Glu Glu Asn Pro Arg	
	210	215 220
	Val Pro Ile Ile Val Thr Gly Asn Asp Phe Ser Thr Leu Tyr Ala Pro	
	225	230 235 240
	Leu Ile Arg Asp Gly Arg Met Glu Lys Phe Tyr Trp Ala Pro Thr Arg	
	245	250 255
	Glu Asp Arg Ile Gly Val Cys Lys Gly Ile Phe Arg Thr Asp Asn Val	
	260	265 270
	Pro Asp Glu Ala Val Val Arg Leu Val Asp Thr Phe Pro Gly Gln Ser	
	275	280 285
	Ile Asp Phe Phe Gly Ala Leu Arg Ala Arg Val Tyr Asp Asp Glu Val	
	290	295 300
	Arg Lys Trp Val Gly Glu Ile Gly Val Glu Asn Ile Ser Lys Arg Leu	
	305	310 315 320
	Val Asn Ser Arg Glu Gly Pro Pro Thr Phe Asp Gln Pro Lys Met Thr	
	325	330 335
[0010]	Ile Glu Lys Leu Met Glu Tyr Gly His Met Leu Val Gln Glu Gln Glu	
	340	345 350
	Asn Val Lys Arg Val Gln Leu Ala Asp Lys Tyr Leu Ser Glu Ala Ala	
	355	360 365
	Leu Gly Gln Ala Asn Asp Asp Ala Met Lys Thr Gly Ala Phe Tyr Gly	
	370	375 380
	Lys	
	385	
	<210> 11	
	<211> 433	
	<212> PRT	
	<213> 稻(Oryza sativa)	
	<400> 11	
	Met Ala Ala Ala Phe Ser Ser Thr Val Gly Ala Pro Ala Ser Thr Pro	
	1	5 10 15
	Thr Asn Phe Leu Gly Lys Lys Leu Lys Lys Gln Val Thr Ser Ala Val	
	20	25 30
	Asn Tyr His Gly Lys Ser Ser Asn Ile Asn Arg Phe Lys Val Met Ala	
	35	40 45
	Lys Glu Leu Asp Glu Gly Lys Gln Thr Asp Gln Asp Arg Trp Lys Gly	
	50	55 60
	Leu Ala Tyr Asp Ile Ser Asp Asp Gln Gln Asp Ile Thr Arg Gly Lys	
	65	70 75 80

	Gly	Phe	Val	Asp	Ser	Leu	Phe	Gln	Ala	Pro	Thr	Gly	Asp	Gly	Thr	His
				85						90						95
	Glu	Ala	Val	Leu	Ser	Ser	Tyr	Glu	Tyr	Leu	Ser	Gln	Gly	Leu	Arg	Thr
			100						105					110		
	Tyr	Asp	Phe	Asp	Asn	Thr	Met	Gly	Gly	Phe	Tyr	Ile	Ala	Pro	Ala	Phe
			115					120					125			
	Met	Asp	Lys	Leu	Val	Val	His	Ile	Ser	Lys	Asn	Phe	Met	Thr	Leu	Pro
		130					135					140				
	Asn	Ile	Lys	Val	Pro	Leu	Ile	Leu	Gly	Ile	Trp	Gly	Gly	Lys	Gly	Gln
	145					150					155					160
	Gly	Lys	Ser	Phe	Gln	Cys	Glu	Leu	Val	Phe	Ala	Lys	Met	Gly	Ile	Asn
				165						170					175	
	Pro	Ile	Met	Met	Ser	Ala	Gly	Glu	Leu	Glu	Ser	Gly	Asn	Ala	Gly	Glu
			180						185					190		
	Pro	Ala	Lys	Leu	Ile	Arg	Gln	Arg	Tyr	Arg	Glu	Ala	Ala	Asp	Ile	Ile
			195				200						205			
	Lys	Lys	Gly	Lys	Met	Cys	Cys	Leu	Phe	Ile	Asn	Asp	Leu	Asp	Ala	Gly
		210				215						220				
	Ala	Gly	Arg	Met	Gly	Gly	Thr	Thr	Gln	Tyr	Thr	Val	Asn	Asn	Gln	Met
	225					230					235					240
	Val	Asn	Ala	Thr	Leu	Met	Asn	Ile	Ala	Asp	Asn	Pro	Thr	Asn	Val	Gln
				245						250					255	
[0011]	Leu	Pro	Gly	Met	Tyr	Asn	Lys	Glu	Asp	Asn	Pro	Arg	Val	Pro	Ile	Ile
			260						265					270		
	Val	Thr	Gly	Asn	Asp	Phe	Ser	Thr	Leu	Tyr	Ala	Pro	Leu	Ile	Arg	Asp
			275					280					285			
	Gly	Arg	Met	Glu	Lys	Phe	Tyr	Trp	Ala	Pro	Thr	Arg	Asp	Asp	Arg	Val
		290				295						300				
	Gly	Val	Cys	Lys	Gly	Ile	Phe	Arg	Thr	Asp	Asn	Val	Pro	Asp	Glu	Asp
	305					310					315					320
	Ile	Val	Lys	Ile	Val	Asp	Ser	Phe	Pro	Gly	Gln	Ser	Ile	Asp	Phe	Phe
				325					330						335	
	Gly	Ala	Leu	Arg	Ala	Arg	Val	Tyr	Asp	Asp	Glu	Val	Arg	Lys	Trp	Val
			340						345					350		
	Ser	Asp	Thr	Gly	Val	Glu	Asn	Ile	Gly	Lys	Arg	Leu	Val	Asn	Ser	Arg
		355						360					365			
	Glu	Gly	Pro	Pro	Glu	Phe	Glu	Gln	Pro	Lys	Met	Thr	Ile	Glu	Lys	Leu
		370				375						380				
	Met	Glu	Tyr	Gly	Tyr	Met	Leu	Val	Lys	Glu	Gln	Glu	Asn	Val	Lys	Arg
	385					390					395					4

Ser

<210> 12
 <211> 431
 <212> PRT
 <213> *Oryza australiensis*
 <400> 12

Met Ala Ala Ala Phe Ser Ser Thr Val Gly Ala Pro Ala Ser Thr Pro
 1 5 10 15
 Thr Asn Phe Leu Gly Lys Lys Leu Lys Lys Gln Val Thr Ser Ala Val
 20 25 30
 Asn Tyr His Gly Lys Ser Ser Asn Ile Asn Arg Phe Lys Val Met Ala
 35 40 45
 Lys Glu Leu Asp Glu Asp Lys Gln Thr Asp Gln Asp Lys Trp Lys Gly
 50 55 60
 Leu Ala Tyr Asp Ile Ser Asp Asp Gln Gln Asp Ile Thr Arg Gly Lys
 65 70 75 80
 Gly Leu Val Asp Ser Leu Phe Gln Ala Pro Met Gly Asp Gly Thr His
 85 90 95
 Glu Ala Val Leu Ser Ser Tyr Glu Tyr Leu Ser Gln Gly Leu Lys Met
 100 105 110
 Leu Asp Asn Thr Met Gly Gly Phe Tyr Ile Ala Pro Ala Phe Met Asp
 [0012] 115 120 125
 Lys Leu Val Val His Ile Ser Lys Asn Phe Met Ala Leu Pro Asn Ile
 130 135 140
 Lys Val Pro Leu Ile Leu Gly Ile Trp Gly Gly Lys Gly Gln Gly Lys
 145 150 155 160

Ser Phe Gln Cys Glu Leu Val Phe Ser Lys Met Gly Ile Asn Pro Ile
 165 170 175
 Met Met Ser Ala Gly Glu Leu Glu Ser Gly Asn Ala Gly Glu Pro Ala
 180 185 190
 Lys Leu Ile Arg Gln Arg Tyr Arg Glu Ala Ala Asp Ile Ile Lys Lys
 195 200 205
 Gly Lys Met Cys Cys Leu Phe Ile Asn Asp Leu Asp Ala Gly Ala Gly
 210 215 220
 Arg Met Gly Gly Thr Thr Gln Tyr Thr Val Asn Asn Gln Met Val Asn
 225 230 235 240
 Ala Thr Leu Met Asn Ile Ala Asp Asn Pro Thr Asn Val Gln Leu Pro
 245 250 255
 Gly Met Tyr Asn Lys Glu Asp Asn Pro Arg Val Pro Ile Ile Val Thr
 260 265 270
 Gly Asn Asp Phe Ser Thr Leu Tyr Ala Pro Leu Ile Arg Asp Gly Arg
 275 280 285
 Met Glu Lys Phe Tyr Trp Ala Pro Thr Arg Asp Asp Arg Val Gly Val

290	295	300
Cys Lys Gly Ile Phe Arg Thr Asp Asn Val Pro Asp Glu Asp Ile Val		
305	310	315
Lys Ile Val Asp Ser Phe Pro Gly Gln Ser Ile Asp Phe Phe Gly Ala		
	325	330
		335
Leu Arg Ala Arg Val Tyr Asp Asp Glu Val Arg Lys Trp Val Ser Asp		
	340	345
		350
Thr Gly Val Glu Asn Ile Gly Lys Lys Leu Val Asn Ser Arg Glu Gly		
	355	360
		365
Pro Pro Glu Phe Glu Gln Pro Lys Met Thr Ile Asp Lys Leu Met Glu		
	370	375
		380
Tyr Gly His Met Leu Val Arg Glu Gln Glu Asn Val Lys Arg Val Gln		
385	390	395
		400
Leu Ala Asp Lys Tyr Leu Ser Glu Ala Ala Leu Gly Asp Ala Asn Ser		
	405	410
		415
Asp Ala Met Lys Thr Gly Ser Phe Tyr Gly Ala Ala Pro Ser Ser		
	420	425
		430

<210> 13

<211> 435

<212> PRT

<213> Larrea tridentate

[0013]

<400> 13

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

	Leu	Val	Phe	Ala	Lys	Met	Gly	Ile	Asn	Pro	Ile	Met	Met	Ser	Ala	Gly	
							180				185					190	
	Glu	Leu	Glu	Ser	Gly	Asn	Ala	Gly	Glu	Pro	Ala	Lys	Leu	Ile	Arg	Gln	
							195				200					205	
	Arg	Tyr	Arg	Glu	Ala	Ala	Asp	Ile	Ile	Lys	Lys	Gly	Lys	Met	Cys	Cys	
							210				215					220	
	Leu	Phe	Ile	Asn	Asp	Leu	Asp	Ala	Gly	Ala	Gly	Arg	Met	Gly	Gly	Thr	
							225				230					235	
	Thr	Gln	Tyr	Thr	Val	Asn	Asn	Gln	Met	Val	Asn	Ala	Thr	Leu	Met	Asn	
							245				250					255	
	Ile	Ala	Asp	Asn	Pro	Thr	Asn	Val	Gln	Leu	Pro	Gly	Met	Tyr	Asn	Lys	
							260				265					270	
	Glu	Glu	Asn	Pro	Arg	Val	Pro	Ile	Ile	Val	Thr	Gly	Asn	Asp	Phe	Ser	
							275				280					285	
	Thr	Leu	Tyr	Ala	Pro	Leu	Ile	Arg	Asp	Gly	Arg	Met	Glu	Lys	Phe	Tyr	
							290				295					300	
	Trp	Ala	Pro	Thr	Arg	Glu	Asp	Arg	Ile	Gly	Val	Cys	Lys	Gly	Ile	Phe	
							305				310					315	
	Arg	Thr	Asp	Asn	Val	Pro	Glu	Glu	Asp	Ile	Val	Lys	Val	Val	Asp	Gln	
							325				330					335	
	Phe	Pro	Gly	Gln	Ser	Ile	Asp	Phe	Phe	Gly	Ala	Leu	Arg	Ala	Arg	Val	
							340				345					350	
[0014]	Tyr	Asp	Asp	Glu	Val	Arg	Lys	Trp	Val	Ser	Glu	Val	Gly	Val	Asp	Thr	
							355				360					365	
	Ile	Gly	Lys	Lys	Leu	Val	Asn	Ser	Lys	Glu	Gly	Pro	Pro	Thr	Phe	Glu	
							370				375					380	
	Gln	Pro	Lys	Met	Thr	Ile	Asp	Lys	Leu	Leu	Gln	Tyr	Gly	Asn	Met	Leu	
							385				390					395	
	Val	Glu	Glu	Gln	Glu	Asn	Val	Lys	Arg	Val	Gln	Leu	Ala	Asp	Lys	Tyr	
							405				410					415	
	Met	Ser	Glu	Ala	Ala	Leu	Gly	Asp	Ala	Asn	Gln	Asp	Ala	Ile	Lys	Arg	
							420				425					430	
	Gly	Thr	Phe														
							435										
	<210>	14															
	<211>	438															
	<212>	PRT															
	<213>	Musa acuminata															
	<400>	14															
	Met	Ala	Ala	Ala	Val	Ser	Thr	Val	Gly	Ala	Val	Asn	Arg	Val	Pro	Leu	
	1				5						10				15		
	Ser	Leu	His	Gly	Ser	Ser	Ser	Gly	Ala	Ser	Val	Pro	Ser	Ser	Ala	Phe	
					20						25				30		
	Phe	Gly	Ser	Ala	Leu	Lys	Lys	Ala	Asn	Ser	Gly	Leu	Ser	His	Gly	Arg	

	35		40		45
	Val Ser Thr Ala Ser Phe Lys	Val Leu Ala Ala Asp Leu Asp Glu Ser			
	50		55		60
	Lys Gln Thr Ser Ser Asp Arg Trp Ala Gly Leu Ala Ser Asp Val Ser				
	65		70		80
	Asp Asp Gln Gln Asp Ile Thr Arg Gly Lys Gly Leu Val Asp Thr Leu				
		85		90	95
	Phe Gln Ala Pro Met Gly Asp Gly Thr His Ile Pro Val Met Ser Ser				
	100		105		110
	Tyr Glu Tyr Ile Ser Gln Gly Leu Arg Gln Tyr Asn Phe Asp Asn Thr				
	115		120		125
	Met Asp Gly Tyr Tyr Ile Ala Pro Ala Phe Met Asp Lys Leu Val Val				
	130		135		140
	His Ile Ala Lys Asn Phe Met Asn Leu Pro Asn Ile Lys Val Pro Leu				
	145		150		155
	Ile Leu Gly Ile Trp Gly Gly Lys Gly Gln Gly Lys Ser Phe Gln Cys				
		165		170	175
	Glu Leu Val Phe Ala Lys Met Gly Ile Asn Pro Ile Val Met Ser Ala				
	180		185		190
	Gly Glu Leu Glu Ser Gly Asn Ala Gly Glu Pro Ala Lys Leu Ile Arg				
	195		200		205
	Gln Arg Tyr Arg Glu Ala Ala Asp Ile Ile Lys Lys Gly Lys Met Cys				
[0015]	210		215		220
	Cys Leu Phe Ile Asn Asp Leu Asp Ala Gly Ala Gly Arg Leu Gly Gly				
	225		230		235
	Thr Thr Gln Tyr Thr Val Asn Asn Gln Met Val Asn Ala Thr Leu Met				
		245		250	255
	Asn Ile Ala Asp Asn Pro Thr Asn Val Gln Leu Pro Gly Met Tyr Asn				
	260		265		270
	Lys Gln Glu Asn Ala Arg Val Pro Ile Ile Val Thr Gly Asn Asp Phe				
	275		280		285
	Ser Thr Leu Tyr Ala Pro Leu Ile Arg Asp Gly Arg Met Glu Lys Phe				
	290		295		300
	Tyr Trp Ala Pro Thr Arg Glu Asp Arg Ile Gly Val Cys Leu Gly Ile				
	305		310		315
	Phe Arg Thr Asp Asn Val Pro Met Glu Asp Ile Val Lys Leu Val Asp				
		325		330	335
	Ala Phe Pro Gly Gln Ser Ile Asp Phe Phe Gly Ala Leu Arg Ala Arg				
	340		345		350
	Val Tyr Asp Asp Glu Val Arg Lys Trp Val Ala Glu Ile Gly Val Glu				
	355		360		365
	Lys Val Gly Lys Lys Leu Val Asn Ser Leu Glu Gly Pro Pro Thr Phe				
	370		375		380
	Glu Gln Pro Lys Met Thr Leu Asp Lys Leu Thr Glu Tyr Gly Asn Met				
	385		390		395
					400

Leu Val Arg Glu Gln Glu Asn Val Lys Arg Val Gln Leu Ala Asp Lys
 405 410 415
 Tyr Leu Ser Glu Ala Ala Leu Gly Asp Ala Asn Ala Asp Ala Ile Lys
 420 425 430
 Thr Gly Ser Phe Tyr Gly
 435
 <210> 15
 <211> 373
 <212> PRT
 <213> *Datisca glomerata*
 <400> 15
 Met Ala Ser Ser Ala Val Ser Thr Ile Gly Ala Val Asn Arg Ser Pro
 1 5 10 15
 Leu Asn Leu Asn Asn Asn Gly Thr Gly Gly Leu Leu Pro Asn Thr Ala
 20 25 30
 Phe Phe Gly Ser Ser Leu Lys Lys Met Asn Ser Arg Leu Thr Asn Pro
 35 40 45
 Arg Ile Ala Ala Gly Asn Ile Lys Ala Val Ala Asp Asp Asp Glu Glu
 50 55 60
 Lys Gln Thr Ser Lys Asp Arg Trp Gly Gly Leu Ala Phe Asp Thr Ser
 65 70 75 80
 Asp Asp Gln Gln Asp Ile Thr Arg Gly Lys Gly Met Val Asp Ser Leu
 [0016] 85 90 95
 Phe Gln Ala Pro Met Gln Thr Gly Thr His Tyr Ala Val Met Ser Ser
 100 105 110
 Tyr Glu Tyr Leu Ser Thr Gly Leu Arg Gln Tyr Leu Asp Asn Asn Met
 115 120 125
 Asp Gly Phe Tyr Ile Ala Pro Ala Phe Met Asp Lys Leu Val Val His
 130 135 140
 Ile Thr Lys Asn Phe Met Thr Leu Pro Asn Ile Lys Val Pro Leu Ile
 145 150 155 160
 Leu Gly Ile Trp Gly Gly Lys Gly Gln Gly Lys Ser Phe Gln Cys Glu
 165 170 175
 Leu Val Phe Ala Lys Met Gly Ile Asn Pro Ile Met Met Ser Ala Gly
 180 185 190
 Glu Leu Glu Ser Gly Asn Ala Gly Glu Pro Ala Lys Leu Ile Arg Gln
 195 200 205
 Arg Cys Arg Glu Ala Ala Asp Ile Ile Lys Lys Gly Lys Met Ser Cys
 210 215 220
 Leu Phe Ile Asn Asp Leu Asp Ala Gly Ala Gly Arg Leu Gly Gly Thr
 225 230 235 240
 Thr Gln Tyr Thr Val Asn Asn Gln Met Val Asn Ala Thr Leu Met Asn
 245 250 255
 Ile Ala Asp Asn Pro Thr Asn Val Gln Leu Pro Gly Met Tyr Asn Lys
 260 265 270

[0017]

Glu Glu Asn Pro Arg Val Pro Ile Ile Val Thr Gly Asn Asp Phe Ser
 275 280 285
 Thr Leu Tyr Ala Pro Leu Ile Arg Asp Gly Arg Met Glu Lys Phe Tyr
 290 295 300
 Trp Ala Pro Thr Arg Glu Asp Arg Ile Gly Val Cys Thr Gly Ile Phe
 305 310 315 320
 Arg Ser Asp Asn Val Ala Lys Glu Asp Ile Val Lys Leu Val Asp Thr
 325 330 335
 Phe Pro Gly Gln Ser Ile Asp Phe Phe Gly Ala Leu Arg Ala Arg Val
 340 345 350
 Tyr Asp Asp Glu Val Arg Lys Trp Ile Ser Gly Val Gly Val Gln Asp
 355 360 365
 Val Gly Lys Ser Leu
 370
 <210> 16
 <211> 431
 <212> PRT
 <213> Theobroma cacao
 <400> 16
 Met Ser Leu Ser Arg Glu Leu Gln Trp Asn Leu Asn Gly Ser Lys Pro
 1 5 10 15
 Ala Ala Ala Ser Leu Pro Ser Ser Ala Phe Leu Gly Asn Thr Leu Lys
 20 25 30
 Lys Val Thr Ser Lys Ile Ser His Arg Lys Val Ser Pro Lys Asn Phe
 35 40 45
 Lys Val Val Ala Glu Tyr Asp Glu Ser Lys Lys Thr Ser Lys Asp Arg
 50 55 60
 Trp Lys Gly Leu Ala Tyr Asp Glu Ser Asp Asp Gln Gln Asp Ile Thr
 65 70 75 80
 Arg Gly Lys Gly Met Val Asp Thr Leu Phe Gln Ala Pro Met Gly Thr
 85 90 95
 Gly Thr His His Ala Ile Met Ser Ser Tyr Asp Tyr Ile Ser Gln Gly
 100 105 110
 Leu Arg Met Tyr Asn Leu Asp Asn Thr Met Asp Gly Phe Tyr Ile Ala
 115 120 125
 Pro Ala Phe Met Asp Lys Val Val Val His Ile Thr Lys Asn Phe Leu
 130 135 140
 Ser Leu Pro Asn Ile Lys Val Pro Leu Ile Leu Gly Ile Trp Gly Gly
 145 150 155 160
 Lys Gly Gln Gly Lys Ser Phe Gln Cys Glu Leu Val Phe Ala Lys Met
 165 170 175
 Gly Ile Asn Ser Ile Met Met Ser Ala Gly Glu Leu Glu Ser Gly Asn
 180 185 190
 Ala Gly Glu Pro Ala Lys Leu Ile Arg Gln Arg Tyr Arg Glu Ala Ala
 195 200 205

	Asp Ile Ile Arg Lys Gly Lys Met Cys Cys Leu Phe Ile Asn Asp Leu	
	210	215 220
	Asp Ala Gly Ala Gly Arg Met Gly Gly Thr Thr Gln Tyr Thr Val Asn	
	225	230 235 240
	Asn Gln Met Val Asn Ala Thr Leu Met Asn Ile Ala Asp Asn Pro Thr	
	245	250 255
	Asn Val Gln Leu Pro Gly Met Tyr Asn Lys Glu Glu Asn Pro Arg Val	
	260	265 270
	Pro Ile Ile Val Thr Gly Asn Asp Phe Ser Thr Leu Tyr Ala Pro Leu	
	275	280 285
	Ile Arg Asp Gly Arg Met Glu Lys Phe Tyr Trp Ala Pro Thr Arg Glu	
	290	295 300
	Asp Arg Ile Gly Val Cys Ser Gly Ile Phe Lys Thr Asp Arg Val Pro	
	305	310 315 320
	His Glu Asp Ser Val Lys Leu Val Asp Thr Phe Pro Gly Gln Ser Ile	
	325	330 335
	Asp Phe Phe Ser Ala Leu Arg Ala Arg Val Tyr Asp Asp Glu Val Arg	
	340	345 350
	Glu Trp Val Ser Glu Val Gly Val Gln Asn Val Gly Lys Arg Leu Val	
	355	360 365
	Asn Ser Lys Glu Gly Pro Pro Thr Phe Glu Gln Pro Lys Met Thr Leu	
	370	375 380
[0018]	Glu Lys Leu Leu Glu Tyr Gly Asn Met Leu Val Gln Glu Gln Glu Asn	
	385	390 395 400
	Val Lys Arg Val Gln Leu Ala Asp Lys Tyr Leu Lys Glu Ala Ala Leu	
	405	410 415
	Gly Glu Ala Asn Asp Asp Ala Ile Lys Asn Glu Ser Phe Tyr Gly	
	420	425 430
	<210> 17	
	<211> 442	
	<212> PRT	
	<213> 烟草(Nicotiana tabacum)	
	<400> 17	
	Met Ala Thr Ser Val Ser Thr Ile Gly Ala Val Asn Lys Thr Pro Leu	
	1	5 10 15
	Ser Leu Asn Asn Ser Val Ala Gly Thr Ser Val Pro Ser Thr Ala Phe	
	20	25 30
	Phe Gly Lys Thr Leu Lys Lys Val Tyr Gly Lys Gly Val Ser Ser Pro	
	35	40 45
	Lys Val Thr Asn Lys Ser Leu Arg Ile Val Ala Glu Gln Ile Asp Val	
	50	55 60
	Asp Pro Lys Lys Gln Thr Asp Ser Asp Arg Trp Lys Gly Leu Val Gln	
	65	70 75 80
	Asp Phe Ser Asp Asp Gln Gln Asp Ile Thr Arg Gly Lys Gly Met Val	
	85	90 95

Asp Ser Leu Phe Gln Ala Pro Thr Gly Thr Gly Thr His His Ala Val		
100	105	110
Leu Gln Ser Tyr Glu Tyr Val Ser Gln Gly Leu Arg Gln Tyr Asn Leu		
115	120	125
Asp Asn Lys Leu Asp Gly Phe Tyr Ile Ala Pro Ala Phe Met Asp Lys		
130	135	140
Leu Val Val His Ile Thr Lys Asn Phe Leu Lys Leu Pro Asn Ile Lys		
145	150	155
Val Pro Leu Ile Leu Gly Ile Trp Gly Gly Lys Gly Gln Gly Lys Ser		
165	170	175
Phe Gln Cys Glu Leu Val Phe Arg Lys Met Gly Ile Asn Pro Ile Met		
180	185	190
Met Ser Ala Gly Glu Leu Glu Ser Gly Asn Ala Gly Glu Pro Ala Lys		
195	200	205
Leu Ile Arg Gln Arg Tyr Arg Glu Ala Ala Glu Ile Ile Arg Lys Gly		
210	215	220
Asn Met Cys Cys Leu Phe Ile Asn Asp Leu Asp Ala Gly Ala Gly Arg		
225	230	235
Met Gly Gly Thr Thr Gln Tyr Thr Val Asn Asn Gln Met Val Asn Ala		
245	250	255
Thr Leu Met Asn Ile Ala Asp Asn Pro Thr Asn Val Gln Leu Pro Gly		
260	265	270
[0019] Met Tyr Asn Lys Gln Glu Asn Ala Arg Val Pro Ile Ile Val Thr Gly		
275	280	285
Asn Asp Phe Ser Thr Leu Tyr Ala Pro Leu Ile Arg Asp Gly Arg Met		
290	295	300
Glu Lys Phe Tyr Trp Ala Pro Thr Arg Glu Asp Arg Ile Gly Val Cys		
305	310	315
Thr Gly Ile Phe Arg Thr Asp Asn Val Pro Ala Glu Asp Val Val Lys		
325	330	335
Ile Val Asp Asn Phe Pro Gly Gln Ser Ile Asp Phe Phe Gly Ala Leu		
340	345	350
Arg Ala Arg Val Tyr Asp Asp Glu Val Arg Lys Trp Val Ser Gly Thr		
355	360	365
Gly Ile Glu Lys Ile Gly Asp Lys Leu Leu Asn Ser Phe Asp Gly Pro		
370	375	380
Pro Thr Phe Glu Gln Pro Lys Met Thr Ile Glu Lys Leu Leu Glu Tyr		
385	390	395
Gly Asn Met Leu Val Gln Glu Gln Glu Asn Val Lys Arg Val Gln Leu		
405	410	415
Ala Asp Lys Tyr Leu Lys Glu Ala Ala Leu Gly Asp Ala Asn Ala Asp		
420	425	430
Ala Ile Asn Asn Gly Ser Phe Phe Ala Ser		
435	440	
<210> 18		

<211> 421
 <212> PRT
 <213> 陆地棉(*Gossypium hirsutum*)
 <400> 18
 Met Ala Ala Glu Lys Glu Ile Asp Glu Glu Thr Gln Thr Glu Lys Asp
 1 5 10 15
 Arg Trp Lys Gly Leu Ala Tyr Asp Ile Ser Asp Asp Gln Gln Asp Ile
 20 25 30
 Thr Arg Gly Lys Gly Met Val Asp Ser Leu Phe Gln Ala Pro Met Asn
 35 40 45
 Asp Gly Thr His Tyr Ala Val Met Ser Ser Tyr Glu Tyr Ile Ser Gln
 50 55 60
 Gly Leu Lys Thr Tyr Asn Leu Asp Asn Asn Met Asp Gly Phe Tyr Ile
 65 70 75 80
 Ala Pro Ala Phe Met Asp Lys Leu Val Val His Ile Ser Lys Asn Phe
 85 90 95
 Met Ser Leu Pro Asn Ile Lys Val Pro Leu Ile Leu Gly Ile Trp Gly
 100 105 110
 Gly Lys Gly Gln Gly Lys Ser Phe Gln Cys Glu Leu Val Phe Ala Lys
 115 120 125
 Met Gly Ile Asn Pro Ile Met Met Ser Ala Gly Glu Leu Glu Ser Gly
 130 135 140
 [0020] Asn Ala Gly Glu Pro Ala Lys Leu Ile Arg Gln Arg Tyr Arg Glu Ala
 145 150 155 160
 Ala Asp Ile Ile Lys Lys Gly Lys Met Cys Ala Leu Phe Ile Asn Asp
 165 170 175
 Leu Asp Ala Gly Ala Gly Arg Met Gly Gly Thr Thr Gln Tyr Thr Val
 180 185 190
 Asn Asn Gln Met Val Asn Ala Thr Leu Met Asn Ile Ala Asp Asn Pro
 195 200 205
 Thr Asn Val Gln Leu Pro Gly Met Tyr Asn Lys Glu Glu Asn Pro Arg
 210 215 220
 Val Pro Ile Ile Val Thr Gly Asn Asp Phe Ser Thr Leu Tyr Ala Pro
 225 230 235 240
 Leu Ile Arg Asp Gly Arg Met Glu Lys Phe Tyr Trp Ala Pro Thr Arg
 245 250 255
 Asp Asp Arg Ile Gly Val Cys Lys Gly Ile Phe Arg Thr Asp Gly Val
 260 265 270
 Arg Asp Glu Asp Ile Val Lys Leu Val Asp Thr Phe Pro Gly Gln Ser
 275 280 285
 Ile Asp Phe Phe Gly Ala Leu Arg Ala Arg Val Tyr Asp Asp Glu Val
 290 295 300
 Arg Lys Trp Ile Ser Glu Val Gly Val Ala Ser Val Gly Lys Lys Leu
 305 310 315 320
 Val Asn Ser Arg Glu Gly Pro Pro Thr Phe Glu Gln Pro Lys Met Thr

	325	330	335
	Ile Glu Lys Leu Leu Glu Tyr Gly Asn Met Leu Val Ala Glu Gln Glu		
	340	345	350
	Asn Val Lys Arg Val Gln Leu Ala Asp Lys Tyr Leu Ser Glu Ala Ala		
	355	360	365
	Leu Gly Glu Ala Asn Glu Asp Ser Ile Asn Arg Gly Thr Phe Tyr Gly		
	370	375	380
	Lys Ala Ala Gln Gln Val Gly Val Pro Val Pro Glu Gly Cys Thr Asp		
	385	390	395
	Pro Asn Ala Asp Asn Phe Asp Pro Thr Ala Arg Ser Asp Asp Gly Thr		
	405	410	415
	Cys Thr Tyr Lys Phe		
	420		
	<210> 19		
	<211> 430		
	<212> PRT		
	<213> 人工序列		
	<220>		
	<223> 来自“暖”物种的氨基酸序列		
	<220>		
	<221> misc_feature		
	<222> (30).. (31)		
[0021]	<223> Xaa 可以是任何天然氨基酸		
	<220>		
	<221> misc_feature		
	<222> (35).. (36)		
	<223> Xaa 可以是任何天然氨基酸		
	<220>		
	<221> misc_feature		
	<222> (38).. (39)		
	<223> Xaa 可以是任何天然氨基酸		
	<220>		
	<221> misc_feature		
	<222> (44).. (44)		
	<223> Xaa 可以是任何天然氨基酸		
	<220>		
	<221> misc_feature		
	<222> (53).. (53)		
	<223> Xaa 可以是任何天然氨基酸		
	<220>		
	<221> misc_feature		
	<222> (362).. (362)		
	<223> Xaa 可以是任何天然氨基酸		
	<220>		
	<221> misc_feature		

<222> (395)..(395)
 <223> Xaa 可以是任何天然氨基酸
 <220>
 <221> misc_feature
 <222> (420)..(420)
 <223> Xaa 可以是任何天然氨基酸
 <400> 19
 Met Ala Ala Ser Ser Thr Val Gly Ala Val Asn Arg Thr Pro Leu Ser
 1 5 10 15
 Leu Asn Gly Ser Gly Val Pro Ser Thr Ala Phe Phe Gly Xaa Xaa Leu
 20 25 30
 Lys Lys Xaa Xaa Asn Xaa Xaa Lys Val Ser Ile Xaa Asn Phe Lys Val
 35 40 45
 Val Ala Lys Glu Xaa Asp Glu Asp Lys Gln Thr Asp Lys Asp Arg Trp
 50 55 60
 Lys Gly Leu Ala Tyr Asp Ile Ser Asp Asp Gln Gln Asp Ile Thr Arg
 65 70 75 80
 Gly Lys Gly Met Val Asp Ser Leu Phe Gln Ala Pro Met Gly Asp Gly
 85 90 95
 Thr His Tyr Ala Val Met Ser Ser Tyr Glu Tyr Ile Ser Gln Gly Leu
 100 105 110
 Arg Gln Tyr Asn Leu Asp Asn Thr Met Asp Gly Phe Tyr Ile Ala Pro
 [0022] 115 120 125
 Ala Phe Met Asp Lys Leu Val Val His Ile Thr Lys Asn Phe Met Ser
 130 135 140
 Leu Pro Asn Ile Lys Val Pro Leu Ile Leu Gly Ile Trp Gly Gly Lys
 145 150 155 160
 Gly Gln Gly Lys Ser Phe Gln Cys Glu Leu Val Phe Ala Lys Met Gly
 165 170 175
 Ile Asn Pro Ile Met Met Ser Ala Gly Glu Leu Glu Ser Gly Asn Ala
 180 185 190
 Gly Glu Pro Ala Lys Leu Ile Arg Gln Arg Tyr Arg Glu Ala Ala Asp
 195 200 205
 Ile Ile Lys Lys Gly Lys Met Cys Cys Leu Phe Ile Asn Asp Leu Asp
 210 215 220
 Ala Gly Ala Gly Arg Met Gly Gly Thr Thr Gln Tyr Thr Val Asn Asn
 225 230 235 240
 Gln Met Val Asn Ala Thr Leu Met Asn Ile Ala Asp Asn Pro Thr Asn
 245 250 255
 Val Gln Leu Pro Gly Met Tyr Asn Lys Glu Glu Asn Pro Arg Val Pro
 260 265 270
 Ile Ile Val Thr Gly Asn Asp Phe Ser Thr Leu Tyr Ala Pro Leu Ile
 275 280 285
 Arg Asp Gly Arg Met Glu Lys Phe Tyr Trp Ala Pro Thr Arg Glu Asp
 290 295 300

```

Arg Ile Gly Val Cys Lys Gly Ile Phe Arg Thr Asp Asn Val Pro Asp
305                310                315                320
Glu Asp Ile Val Lys Leu Val Asp Thr Phe Pro Gly Gln Ser Ile Asp
                325                330                335
Phe Phe Gly Ala Leu Arg Ala Arg Val Tyr Asp Asp Glu Val Arg Lys
                340                345                350
Trp Val Ser Glu Val Gly Val Glu Asn Xaa Gly Lys Lys Leu Val Asn
                355                360                365
Ser Arg Glu Gly Pro Pro Thr Phe Glu Gln Pro Lys Met Thr Ile Glu
                370                375                380
Lys Leu Leu Glu Tyr Gly Asn Met Leu Val Xaa Glu Gln Glu Asn Val
385                390                395                400
Lys Arg Val Gln Leu Ala Asp Lys Tyr Leu Ser Glu Ala Ala Leu Gly
                405                410                415
Asp Ala Asn Xaa Asp Ala Ile Lys Thr Gly Ser Phe Tyr Gly
                420                425                430
<210> 20
<211> 441
<212> PRT
<213> 拟南芥(Arabidopsis thaliana)
<400> 20
Met Ala Ala Ala Val Ser Thr Val Gly Ala Ile Asn Arg Ala Pro Leu
[0023] 1                5                10                15
Ser Leu Asn Gly Ser Gly Ser Gly Ala Val Ser Ala Pro Ala Ser Thr
                20                25                30
Phe Leu Gly Lys Lys Val Val Thr Val Ser Arg Phe Ala Gln Ser Asn
                35                40                45
Lys Lys Ser Asn Gly Ser Phe Lys Val Leu Ala Val Lys Glu Asp Lys
                50                55                60
Gln Thr Asp Gly Asp Arg Trp Arg Gly Leu Ala Tyr Asp Thr Ser Asp
65                70                75                80
Asp Gln Gln Asp Ile Thr Arg Gly Lys Gly Met Val Asp Ser Val Phe
                85                90                95
Gln Ala Pro Met Gly Thr Gly Thr His His Ala Val Leu Ser Ser Tyr
                100               105               110
Glu Tyr Val Ser Gln Gly Leu Arg Gln Tyr Asn Leu Asp Asn Met Met
                115               120               125
Asp Gly Phe Tyr Ile Ala Pro Ala Phe Met Asp Lys Leu Val Val His
                130               135               140
Ile Thr Lys Asn Phe Leu Thr Leu Pro Asn Ile Lys Val Pro Leu Ile
145                150                155                160
Leu Gly Ile Trp Gly Gly Lys Gly Gln Gly Lys Ser Phe Gln Cys Glu
                165                170                175
Leu Val Met Ala Lys Met Gly Ile Asn Pro Ile Met Met Ser Ala Gly
                180                185                190

```

	Glu	Leu	Glu	Ser	Gly	Asn	Ala	Gly	Glu	Pro	Ala	Lys	Leu	Ile	Arg	Gln	
	195							200					205				
	Arg	Tyr	Arg	Glu	Ala	Ala	Asp	Leu	Ile	Lys	Lys	Gly	Lys	Met	Cys	Cys	
	210						215					220					
	Leu	Phe	Ile	Asn	Asp	Leu	Asp	Ala	Gly	Ala	Gly	Arg	Met	Gly	Gly	Thr	
	225					230					235					240	
	Thr	Gln	Tyr	Thr	Val	Asn	Asn	Gln	Met	Val	Asn	Ala	Thr	Leu	Met	Asn	
					245				250					255			
	Ile	Ala	Asp	Asn	Pro	Thr	Asn	Val	Gln	Leu	Pro	Gly	Met	Tyr	Asn	Lys	
					260				265					270			
	Glu	Glu	Asn	Ala	Arg	Val	Pro	Ile	Ile	Cys	Thr	Gly	Asn	Asp	Phe	Ser	
					275				280					285			
	Thr	Leu	Tyr	Ala	Pro	Leu	Ile	Arg	Asp	Gly	Arg	Met	Glu	Lys	Phe	Tyr	
	290					295						300					
	Trp	Ala	Pro	Thr	Arg	Glu	Asp	Arg	Ile	Gly	Val	Cys	Lys	Gly	Ile	Phe	
	305					310					315					320	
	Arg	Thr	Asp	Lys	Ile	Lys	Asp	Glu	Asp	Ile	Val	Thr	Leu	Val	Asp	Gln	
					325						330				335		
	Phe	Pro	Gly	Gln	Ser	Ile	Asp	Phe	Phe	Gly	Ala	Leu	Arg	Ala	Arg	Val	
					340					345					350		
	Tyr	Asp	Asp	Glu	Val	Arg	Lys	Phe	Val	Glu	Ser	Leu	Gly	Val	Glu	Lys	
					355				360					365			
[0024]	Ile	Gly	Lys	Arg	Leu	Val	Asn	Ser	Arg	Glu	Gly	Pro	Pro	Val	Phe	Glu	
					370				375					380			
	Gln	Pro	Glu	Met	Thr	Tyr	Glu	Lys	Leu	Met	Glu	Tyr	Gly	Asn	Met	Leu	
	385					390					395					400	
	Val	Met	Glu	Gln	Glu	Asn	Val	Lys	Arg	Val	Gln	Leu	Ala	Glu	Thr	Tyr	
					405						410					415	
	Leu	Ser	Gln	Ala	Ala	Leu	Gly	Asp	Ala	Asn	Ala	Asp	Ala	Ile	Gly	Arg	
					420				425						430		
	Gly	Thr	Phe	Tyr	Gly	Lys	Thr	Glu	Val								
					435				440								
	<210>	21															
	<211>	438															
	<212>	PRT															
	<213>	甘蓝	(Brassica	oleracea)													
	<400>	21															
	Met	Ala	Ala	Ala	Val	Ser	Thr	Val	Gly	Ala	Ile	Asn	Arg	Ala	Pro	Leu	
	1				5					10					15		
	Ser	Leu	Asn	Gly	Ser	Gly	Ala	Gly	Ala	Ala	Ser	Val	Pro	Ala	Thr	Thr	
					20					25					30		
	Phe	Leu	Gly	Lys	Lys	Ala	Val	Thr	Ala	Ser	Arg	Phe	Thr	Gln	Ser	Asn	
					35					40					45		
	Asn	Lys	Lys	Ser	Asn	Gly	Ser	Phe	Lys	Val	Val	Ala	Val	Lys	Glu	Asp	
					50				55						60		

	Lys	Gln	Thr	Asp	Gly	Asp	Arg	Trp	Lys	Gly	Leu	Ala	Tyr	Asp	Thr	Ser
65						70				75						80
	Asp	Asp	Gln	Gln	Asp	Ile	Thr	Arg	Gly	Lys	Gly	Met	Val	Asp	Ser	Val
					85					90						95
	Phe	Gln	Ala	Pro	Met	Gly	Thr	Gly	Thr	His	Asn	Ala	Val	Leu	Ser	Ser
					100					105						110
	Tyr	Glu	Tyr	Ile	Ser	Gln	Gly	Leu	Lys	Gln	Tyr	Asn	Leu	Asp	Asn	Met
					115					120						125
	Met	Asp	Gly	Leu	Phe	Ile	Ala	Pro	Ala	Phe	Met	Asp	Lys	Leu	Val	Val
					130					135						140
	His	Ile	Thr	Lys	Asn	Phe	Leu	Thr	Leu	Pro	Asn	Ile	Lys	Val	Pro	Leu
145						150						155				160
	Ile	Leu	Gly	Ile	Trp	Gly	Gly	Lys	Gly	Gln	Gly	Lys	Ser	Phe	Gln	Cys
					165					170						175
	Glu	Leu	Val	Met	Ala	Lys	Met	Gly	Ile	Asn	Pro	Ile	Met	Met	Ser	Ala
					180					185						190
	Gly	Glu	Leu	Glu	Ser	Gly	Asn	Ala	Gly	Glu	Pro	Ala	Lys	Leu	Ile	Arg
					195					200						205
	Gln	Arg	Tyr	Arg	Glu	Ala	Ala	Asp	Met	Ile	Lys	Lys	Gly	Lys	Met	Cys
					210					215						220
	Cys	Leu	Phe	Ile	Asn	Asp	Leu	Asp	Ala	Gly	Ala	Gly	Arg	Met	Gly	Gly
225						230						235				240
[0025]	Thr	Thr	Gln	Tyr	Thr	Val	Asn	Asn	Gln	Met	Val	Asn	Ala	Thr	Leu	Met
					245					250						255
	Asn	Ile	Ala	Asp	Asn	Pro	Thr	Asn	Val	Gln	Leu	Pro	Gly	Met	Tyr	Asn
					260					265						270
	Lys	Glu	Glu	Asn	Ala	Arg	Val	Pro	Ile	Ile	Val	Thr	Gly	Asn	Asp	Phe
					275					280						285
	Ser	Thr	Leu	Tyr	Ala	Pro	Leu	Ile	Arg	Asp	Gly	Arg	Met	Glu	Lys	Phe
					290					295						300
	Tyr	Trp	Ala	Pro	Thr	Arg	Glu	Asp	Arg	Ile	Gly	Val	Cys	Lys	Gly	Ile
305						310						315				320
	Phe	Arg	Thr	Asp	Asn	Val	Lys	Asp	Glu	Asp	Ile	Val	Thr	Leu	Val	Asp
					325					330						335
	Gln	Phe	Pro	Gly	Gln	Ser	Ile	Asp	Phe	Phe	Gly	Ala	Leu	Arg	Ala	Arg
					340					345						350
	Val	Tyr	Asp	Asp	Glu	Val	Arg	Lys	Phe	Val	Glu	Gly	Leu	Gly	Val	Glu
					355					360						365
	Lys	Ile	Gly	Lys	Arg	Leu	Val	Asn	Ser	Arg	Glu	Gly	Pro	Pro	Val	Phe
					370					375						380
	Glu	Gln	Pro	Ala	Met	Thr	Leu	Glu	Lys	Leu	Met	Glu	Tyr	Gly	Asn	Met
385						390						395				400
	Leu	Val	Met	Glu	Gln	Glu	Asn	Val	Lys	Arg	Val	Gln	Leu	Ala	Asp	Gln
					405					410						415
	Tyr	Leu	Asn	Glu	Ala	Ala	Leu	Gly	Asp	Ala	Asn	Ala	Asp	Ala	Ile	Gly

	420	425	430
Arg Gly Thr Phe Tyr Gly			
435			
<210> 22			
<211> 440			
<212> PRT			
<213> Picea sitchensis			
<400> 22			
Met Ala Thr Ala Ala Ala Val Gly Ser Ile Thr Val Ser Ser Asn			
1 5 10 15			
Leu Pro Leu Ser Ser Gly Gln Ala Ser Ile Leu Lys Thr Val Pro Ser			
20 25 30			
Thr Gly Phe Phe Gly Cys Gly Leu Lys Lys Thr Ser Phe Asn Ala Ser			
35 40 45			
His Thr Phe Thr Lys Ala Asn His Phe Lys Val Val Ala Glu Ile Asp			
50 55 60			
Glu Gly Lys Gln Thr Asp Lys Asp Lys Trp Lys Gly Leu Ala Phe Asp			
65 70 75 80			
Glu Ser Asp Asp Gln Met Asp Ile Arg Arg Gly Lys Gly Lys Val Asp			
85 90 95			
Ser Leu Phe Gln Ala Pro Met Gly Ser Gly Thr His Asn Val Val Met			
100 105 110			

[0026]

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

[0027]

Asp Phe Ser Thr Leu Tyr Ala Pro Leu Ile Arg Asp Gly Arg Met Glu
 290 295 300
 Lys Phe Tyr Trp Ala Pro Thr Arg Asp Asp Arg Ile Gly Val Cys Gln
 305 310 315 320
 Gly Ile Phe Arg Ala Asp Asn Val His Pro Asp Asp Val Val Arg Leu
 325 330 335
 Val Asp Thr Phe Pro Gly Gln Ser Ile Asp Phe Phe Gly Ala Leu Arg
 340 345 350
 Ala Arg Val Tyr Asp Asp Glu Val Arg Lys Trp Val Ala Gly Thr Gly
 355 360 365
 Ile Gln Asn Ile Gly Lys Lys Leu Ile Asn Ser Lys Glu Gly Pro Pro
 370 375 380
 Thr Phe Glu Lys Pro Ala Met Thr Ile Glu Lys Leu Leu Glu Tyr Gly
 385 390 395 400
 Asn Met Leu Val Gly Glu Gln Asp Asn Val Lys Arg Val Gln Leu Ala
 405 410 415
 Asp Lys Tyr Met Ser Glu Ala Ala Leu Gly Asp Ala Asn Glu Asp Ser
 420 425 430
 Ile Lys Arg Gly Thr Phe Tyr Gly
 435 440
 <210> 23
 <211> 472
 <212> PRT
 <213> 菠菜(Spinacia oleracea)
 <400> 23
 Met Ala Thr Ala Val Ser Thr Val Gly Ala Ala Thr Arg Ala Pro Leu
 1 5 10 15
 Asn Leu Asn Gly Ser Ser Ala Gly Ala Ser Val Pro Thr Ser Gly Phe
 20 25 30
 Leu Gly Ser Ser Leu Lys Lys His Thr Asn Val Arg Phe Pro Ser Ser
 35 40 45
 Ser Arg Thr Thr Ser Met Thr Val Lys Ala Ala Glu Asn Glu Glu Lys
 50 55 60
 Asn Thr Asp Lys Trp Ala His Leu Ala Lys Asp Phe Ser Asp Asp Gln
 65 70 75 80
 Leu Asp Ile Arg Arg Gly Lys Gly Met Val Asp Ser Leu Phe Gln Ala
 85 90 95
 Pro Ala Asp Ala Gly Thr His Val Pro Ile Gln Ser Ser Phe Glu Tyr
 100 105 110
 Glu Ser Gln Gly Leu Arg Lys Tyr Asp Ile Asp Asn Met Leu Gly Asp
 115 120 125
 Leu Tyr Ile Ala Pro Ala Phe Met Asp Lys Leu Val Val His Ile Thr
 130 135 140
 Lys Asn Phe Leu Asn Leu Pro Asn Ile Lys Ile Pro Leu Ile Leu Gly

145 150 155 160
 Val Trp Gly Gly Lys Gly Gln Gly Lys Ser Phe Gln Cys Glu Leu Val
 165 170 175
 Phe Ala Lys Leu Gly Ile Asn Pro Ile Met Met Ser Ala Gly Glu Leu
 180 185 190
 Glu Ser Gly Asn Ala Gly Glu Pro Ala Lys Leu Ile Arg Gln Arg Tyr
 195 200 205
 Arg Glu Ala Ala Asp Leu Ile Ala Lys Gly Lys Met Cys Ala Leu Phe
 210 215 220
 Ile Asn Asp Leu Glu Pro Gly Ala Gly Arg Met Gly Gly Thr Thr Gln
 225 230 235 240
 Tyr Thr Val Asn Asn Gln Met Val Asn Ala Thr Leu Leu Asn Ile Ala
 245 250 255
 Asp Asn Pro Thr Asn Val Gln Leu Pro Gly Met Tyr Asn Lys Gln Asp
 260 265 270
 Asn Ala Arg Val Pro Ile Ile Val Thr Gly Asn Asp Phe Ser Thr Leu
 275 280 285
 Tyr Ala Pro Leu Ile Arg Asp Gly Arg Met Glu Lys Phe Tyr Trp Ala
 290 295 300
 Pro Thr Arg Glu Asp Arg Ile Gly Val Cys Thr Gly Ile Phe Lys Thr
 305 310 315 320
 Asp Lys Val Pro Ala Glu His Val Val Lys Leu Val Asp Ala Phe Pro
 [0028] 325 330 335
 Gly Gln Ser Ile Asp Phe Phe Gly Ala Leu Arg Ala Arg Val Tyr Asp
 340 345 350
 Asp Glu Val Arg Lys Trp Val Asn Ser Val Gly Val Asp Asn Val Gly
 355 360 365
 Lys Lys Leu Val Asn Ser Lys Asp Gly Pro Pro Val Phe Glu Gln Pro
 370 375 380
 Glu Met Thr Leu Gln Lys Leu Met Glu Tyr Gly Asn Met Leu Val Gln
 385 390 395 400
 Glu Gln Glu Asn Val Lys Arg Val Gln Leu Ala Asp Gln Tyr Met Ser
 405 410 415
 Ser Ala Ala Leu Gly Asp Ala Asn Lys Asp Ala Ile Asp Arg Gly Thr
 420 425 430
 Phe Phe Gly Lys Ala Ala Gln Gln Val Ser Leu Pro Val Ala Gln Gly
 435 440 445
 Cys Thr Asp Pro Glu Ala Lys Asn Tyr Asp Pro Thr Ala Arg Ser Asp
 450 455 460
 Asp Gly Ser Cys Thr Tyr Asn Leu
 465 470
 <210> 24
 <211> 428
 <212> PRT
 <213> Fragaria vesca

<400> 24
 Met Thr Leu Cys Asn Ser Phe Leu Leu Lys Thr Pro Ile Phe Pro Pro
 1 5 10 15
 His Ser His Tyr Phe Pro Pro Gln Phe Pro Thr Lys Pro Ile Phe Ser
 20 25 30
 Leu Ser Cys Ser Ala Asn Pro Ser Ser Ser Ser Asp Asp Asn Ala Thr
 35 40 45
 Asn Gly Gly Gly Glu Pro Lys Thr Lys Lys Lys Leu Ser Asp Gln Ser
 50 55 60
 Ser Trp Glu Ala Glu Asp Ser Ser Gly Glu Asp Tyr Leu Tyr Arg Leu
 65 70 75 80
 Gly Lys Glu Ala Asp Asn Met Asn Ile Ala Val Gly Ala Arg Ala Gly
 85 90 95
 Val Val Asp Asp Leu Phe Val Gly Gln Phe Leu Gly Arg Asp Ser Asp
 100 105 110
 Ile Val Phe Asp Tyr Arg Gln Lys Val Thr Arg Ser Phe Gln Tyr Leu
 115 120 125
 Gln Gly Asp Tyr Tyr Ile Ala Pro Leu Phe Met Asp Lys Val Val Cys
 130 135 140
 His Met Thr Lys Asn Tyr Leu Ala Pro Ile Leu Asn Thr Lys Val Pro
 145 150 155 160
 Leu Ile Leu Gly Ile Trp Gly Gly Lys Gly Gln Gly Lys Ser Phe Gln
 [0029] 165 170 175
 Thr Glu Leu Ile Phe Gln Ala Met Gly Val Glu Pro Val Ile Met Ser
 180 185 190
 Ala Gly Glu Leu Glu Ser Glu Arg Ala Gly Glu Pro Gly Lys Leu Ile
 195 200 205
 Arg Glu Arg Tyr Arg Thr Ala Ser Gln Val Val Gln Asn Gln Gly Lys
 210 215 220
 Met Ser Cys Leu Met Ile Asn Asp Ile Asp Ala Gly Leu Gly Arg Phe
 225 230 235 240
 Gly Asn Thr Gln Met Thr Val Asn Asn Gln Ile Val Val Gly Thr Leu
 245 250 255
 Met Asn Leu Cys Asp Asn Pro Thr Arg Val Ser Ile Gly Gln Glu Trp
 260 265 270
 Arg Asp Ser Asp Ile Thr Asn Arg Val Pro Val Ile Val Thr Gly Asn
 275 280 285
 Asp Phe Ser Lys Ile Tyr Ala Pro Leu Ile Arg Asp Gly Arg Met Glu
 290 295 300
 Lys Phe Tyr Trp Gln Pro Asn Arg Glu Asp Ile Ile Asn Ile Val Asn
 305 310 315 320
 Arg Met Tyr Glu Lys Asp Gly Ile Ser Arg Asp Glu Val Val Ser Ile
 325 330 335
 Val Asp Thr Phe Pro Asn Gln Ala Leu Asp Phe Tyr Gly Ala Leu Arg
 340 345 350

Ser Arg Thr Tyr Asp Arg Ser Ile Leu Lys Trp Val Asp Asp Ile Gly
 355 360 365
 Gly Val Glu Ser Leu Gly Asn Gln Leu Leu Lys Arg Arg Lys Asp Asp
 370 375 380
 Asn Leu Pro Val Tyr Thr Ala Pro Lys Gln Thr Val Glu Ala Leu Leu
 385 390 395 400
 Glu Ser Gly Tyr Ser Leu Leu Lys Glu Gln Gln Leu Val Met Glu Ser
 405 410 415
 Lys Leu Ser Lys Glu Tyr Met Lys Asn Met Asp Asp
 420 425
 <210> 25
 <211> 442
 <212> PRT
 <213> arachis duranensis
 <400> 25
 Met Pro Arg Gly Ile Glu Glu Leu Leu Val Gln Asn Glu Met Asn Glu
 1 5 10 15
 Ser Pro Ile Met Cys Lys Cys Ala Met Ala Ala Ile Pro Ser Gly Trp
 20 25 30
 Cys Ala Ala His Thr Ser Ile Arg Ile Glu Lys Lys Arg Leu Pro Phe
 35 40 45
 Leu Arg Val Gly Ala Lys Leu Cys Val Arg Cys Cys Ser Asp Gly Ser
 50 55 60
 Glu Lys Glu Glu Glu Lys Lys Lys Arg Leu Ser Glu Gln Ser Ser Trp
 65 70 75 80
 Glu Ala Lys Asp Ala Gln Gly Arg Asp Tyr Leu Tyr Arg Leu Gly Lys
 85 90 95
 Glu Ala Asp Asn Met Asn Ile Ala Val Gly Gln Arg Ala Gly Val Ile
 100 105 110
 Asp Ser Leu Phe Val Gly Asp Phe Leu Gly Lys Asp Ser Asp Ile Val
 115 120 125
 Phe Asp Tyr Arg Gln Lys Ala Thr Arg Ser Phe Glu Tyr Leu Gln Gly
 130 135 140
 Asp Tyr Tyr Ile Ala Pro Leu Phe Met Asp Lys Val Val Cys His Ile
 145 150 155 160
 Val Lys Asn Tyr Leu Ala His Leu Leu Asn Thr Lys Val Pro Leu Ile
 165 170 175
 Leu Gly Ile Trp Gly Gly Lys Gly Gln Gly Lys Ser Phe Gln Thr Glu
 180 185 190
 Leu Ile Phe Gln Ala Met Gly Ile Glu Pro Val Ile Met Ser Ala Gly
 195 200 205
 Glu Leu Glu Ser Glu Arg Ala Gly Glu Pro Gly Lys Leu Ile Arg Glu
 210 215 220
 Arg Tyr Arg Thr Ala Ser Gln Val Val Gln Asn Gln Gly Lys Met Ser
 225 230 235 240

	Cys	Leu	Met	Ile	Asn	Asp	Ile	Asp	Ala	Gly	Leu	Gly	Arg	Phe	Gly	Asn
					245					250					255	
	Thr	Gln	Met	Thr	Val	Asn	Asn	Gln	Ile	Val	Val	Gly	Thr	Leu	Met	Asn
					260				265						270	
	Leu	Ser	Asp	Asn	Pro	Thr	Arg	Val	Ser	Ile	Gly	Gln	Asp	Trp	Arg	Glu
					275				280					285		
	Ser	Asp	Val	Thr	Asn	Arg	Ile	Pro	Ile	Ile	Val	Thr	Gly	Asn	Asp	Phe
					290				295					300		
	Ser	Thr	Ile	Tyr	Ala	Pro	Leu	Ile	Arg	Asp	Gly	Arg	Met	Asp	Lys	Phe
	305					310					315				320	
	Tyr	Trp	Gln	Pro	Asn	His	Glu	Asp	Ile	Val	Asn	Ile	Val	His	Arg	Met
					325						330				335	
	Tyr	Glu	Lys	Asp	Gly	Ile	Pro	Lys	Asp	Glu	Val	Ile	Arg	Ile	Val	Asn
					340					345				350		
	Thr	Phe	Pro	Asn	Gln	Ala	Leu	Asp	Phe	Tyr	Gly	Ala	Leu	Arg	Ser	Arg
					355				360					365		
	Thr	Tyr	Asp	Arg	Ser	Ile	Leu	Lys	Trp	Ile	Glu	Asp	Ile	Gly	Gly	Ile
					370				375					380		
	Glu	Asn	Phe	Gly	Lys	Lys	Phe	Leu	Lys	Arg	Lys	Lys	Asp	Gln	Ser	Leu
	385					390					395				400	
[0031]	Pro	Val	Phe	Ile	Pro	Pro	Glu	Gln	Thr	Met	Glu	Ala	Leu	Val	Glu	Ser
					405					410					415	
	Gly	Tyr	Ser	Leu	Ile	Lys	Glu	Gln	Gln	Leu	Ile	Met	Glu	Thr	Lys	Leu
					420					425					430	
	Ser	Lys	Glu	Tyr	Met	Lys	Asn	Ile	Gln	Glu						
					435				440							
	<210>	26														
	<211>	437														
	<212>	PRT														
	<213>	Brachypodium distachyon														
	<400>	26														
	Met	Ala	Thr	Thr	Ser	Leu	Ala	Ala	Ala	Ser	His	Arg	His	Ile	Leu	Phe
	1				5					10				15		
	Pro	Ser	Arg	Ser	Asn	Pro	Asn	Ser	Asn	Leu	Arg	Thr	Pro	Asn	Arg	Gly
					20					25				30		
	Arg	Leu	Cys	Pro	Thr	Phe	Pro	Pro	Ile	Leu	Cys	Ser	Ser	Ser	Ser	Ser
					35				40					45		
	Ala	Ser	Pro	Gln	Pro	Thr	Ala	Gly	Gly	Asp	Glu	Glu	Glu	Glu	Ala	Gly
					50				55					60		
	Arg	Gly	Arg	Arg	Leu	Ser	Lys	Gln	Ser	Ser	Trp	Glu	Ala	Thr	Asp	Ala
	65					70					75				80	
	Glu	Gly	Asp	Asp	Tyr	Leu	Tyr	Arg	Leu	Gly	Lys	Glu	Ala	Asp	Asn	Met
					85						90				95	
	Asn	Ile	Ala	Val	Gly	Ala	Arg	Ser	Gly	Ile	Ile	Asp	Asp	Leu	Phe	Val

	100	105	110
	Gly Lys Phe Leu Gly Arg Asp Ser Asp Ile Val Phe Asp Tyr Arg Gln		
	115	120	125
	Lys Val Thr Arg Lys Phe Glu Tyr Leu Gln Gly Asp Tyr Tyr Ile Ala		
	130	135	140
	Pro Ser Phe Leu Asp Lys Val Ala Cys His Ile Val Lys Asn Tyr Leu		
	145	150	155
	Ala Asn Asn Leu Asn Ile Lys Val Pro Leu Ile Leu Gly Val Trp Gly		
	165	170	175
	Gly Lys Gly Gln Gly Lys Thr Phe Gln Thr Glu Leu Ile Phe Lys Ala		
	180	185	190
	Met Gly Val Glu Pro Val Ile Met Ser Ala Gly Glu Leu Glu Ser Glu		
	195	200	205
	Arg Ala Gly Glu Pro Gly Arg Leu Ile Arg Asp Arg Tyr Arg Thr Ala		
	210	215	220
	Ser Gln Val Ile Gln Asn Gln Gly Lys Met Ser Cys Leu Met Ile Asn		
	225	230	235
	Asp Leu Asp Ala Gly Val Gly Arg Phe Gly Asn Thr Gln Met Thr Val		
	245	250	255
	Asn Asn Gln Ile Val Val Gly Thr Leu Met Asn Leu Ala Asp Asn Pro		
	260	265	270
	Thr Arg Val Ser Ile Gly Gln Lys Trp Arg Glu Ser Asp Ile Thr His		
[0032]	275	280	285
	Arg Val Pro Ile Ile Ala Thr Gly Asn Asp Phe Ser Thr Leu Tyr Ala		
	290	295	300
	Pro Leu Ile Arg Asp Gly Arg Met Glu Lys Phe Tyr Trp Gln Pro Asp		
	305	310	315
	Arg Glu Asp Ile Ile Asn Ile Val His Arg Met Tyr Met Lys Asp Gly		
	325	330	335
	Leu Ser Phe Glu Glu Val Ser Arg Ile Val Asp Thr Phe Pro Asn Gln		
	340	345	350
	Ala Leu Asp Phe Tyr Gly Ala Leu Arg Ser Arg Thr Tyr Asp Gln Ala		
	355	360	365
	Ile Leu Gln Trp Val Asn Asp Ile Gly Gly Tyr Glu Gln Leu Gly Glu		
	370	375	380
	Lys Leu Leu Lys Lys Lys Lys Arg Glu Lys Leu Pro Thr Phe Ile Pro		
	385	390	395
			400
	Pro Lys Pro Thr Leu Asp Ala Leu Ile Glu Ser Gly Asp Ser Leu Val		
	405	410	415
	Lys Glu Gln Glu Leu Ile Met Asn Ser Lys Leu Ser Lys Glu Tyr Met		
	420	425	430
	Lys Asn Leu Glu Gly		
	435		
	<210>	27	

<211> 436
 <212> PRT
 <213> Arabis alpina
 <400> 27
 Met Ala Leu Ala Asn Ile Ser Leu Arg Phe Lys Phe Pro Val Pro Gln
 1 5 10 15
 Leu Gln Ser Leu Ser Ser Ser Thr Phe His Thr Leu Ile Ser Ile Arg
 20 25 30
 Lys Pro Ser Ser Ile Ala Cys Ser Lys Pro Ser Ser Ser Ser Ser Leu
 35 40 45
 Ser Ala Gly Asp Glu Val Ser Lys Glu Ala Asn Asn Gly Gly Glu Lys
 50 55 60
 Arg Arg Lys Leu Ser Glu Gln Ser Ser Trp Glu Val Lys Asp Ser Glu
 65 70 75 80
 Gly Lys Asp Tyr Leu Tyr Arg Leu Gly Ala Glu Ser Asp Asn Val Asn
 85 90 95
 Ile Ala Val Gly Ala Arg Ser Gly Met Ile Asp Asp Val Phe Ile Gly
 100 105 110
 Asp Phe Leu Gly Lys Asp Ser Asp Ile Val Phe Asp Tyr Arg Gln Lys
 115 120 125
 Ala Thr Arg Ser Phe Glu His Leu Gln Gly Asp Tyr Tyr Ile Ala Pro
 130 135 140
 [0033] Thr Phe Leu Asp Lys Val Ala Val His Ile Val Lys Asn Tyr Leu Ala
 145 150 155 160
 Ser Ser Leu Asn Ile Lys Ile Pro Leu Ile Leu Gly Ile Trp Gly Gly
 165 170 175
 Lys Gly Gln Gly Lys Thr Phe Gln Thr Glu Leu Ile Phe Lys Thr Met
 180 185 190
 Gly Val Glu Pro Val Ile Met Ser Ala Gly Glu Leu Glu Ser Asp Arg
 195 200 205
 Ala Gly Glu Pro Gly Arg Leu Ile Arg Asp Arg Tyr Arg Thr Ala Ser
 210 215 220
 Gln Val Ile Gln Asn Gln Gly Lys Met Ser Val Leu Met Ile Asn Asp
 225 230 235 240
 Ile Asp Ala Gly Leu Gly Arg Phe Gly Glu Thr Gln Met Thr Val Asn
 245 250 255
 Asn Gln Ile Val Val Gly Thr Leu Met Asn Leu Ala Asp Asn Pro Thr
 260 265 270
 Arg Val Ser Val Gly Gln Asp Trp Arg Glu Ala Asp Ile Val Asn Arg
 275 280 285
 Val Pro Leu Ile Val Thr Gly Asn Asp Phe Ser Arg Leu Tyr Ala Pro
 290 295 300
 Leu Ile Arg Glu Gly Arg Met Glu Lys Phe Tyr Trp Gln Pro Thr Arg
 305 310 315 320
 Glu Asp Ile Val Asn Ile Val Ser Arg Met Tyr Glu Lys Asp Gly Ile

	325	330	335
Ser Arg Lys Glu Val Ala Ser Ile Val Asp Lys Phe Pro Asn Gln Ala			
340	345	350	
Leu Asp Phe Tyr Gly Ala Leu Arg Ser Arg Thr Tyr Asp Arg Ser Ile			
355	360	365	
Leu Lys Trp Val Asp Asp Ala Gly Gly Met Glu Thr Leu Gly Lys Thr			
370	375	380	
Leu Leu Arg Arg Lys Lys Thr Lys Glu Val Pro Gln Phe Ile Pro Pro			
385	390	395	400
Glu Gln Thr Met Glu Ala Leu Leu Glu Ser Gly Tyr Ser Leu Ile Arg			
405	410	415	
Glu Gln Lys Leu Ile Met Glu Thr Lys Leu Ser Arg Glu Tyr Met Lys			
420	425	430	
Asn Ile Asp Asp			
435			
<210> 28			
<211> 456			
<212> PRT			
<213> Mesembryanthemum crystallinum			
<400> 28			
Met Ala Leu Leu Ile Asn Pro Thr Asn Pro Thr Gln Phe Ser Asn Leu			
1	5	10	15
[0034] His Ser Leu Ser Ser Leu Thr Ser Phe Pro His Lys Leu Ser Leu Ser			
20	25	30	
Lys Pro Arg Leu Lys Leu Ser Gln Ile Arg Gln Phe Ser Arg Arg Lys			
35	40	45	
Asn Phe Ser Val Phe Val Glu Leu Lys Asn Gly Ala Gln Ile Glu Asn			
50	55	60	
Gly Asp Asp Lys Pro Ser Ser Ser Val Ala Glu Asn Thr Pro Val Arg			
65	70	75	80
Gly Thr Glu Arg Lys Lys Arg Leu Ser Glu Gln Ser Ser Trp Glu Ala			
85	90	95	
Lys Asp Ala Leu Gly Asn Asp Tyr Leu Tyr Arg Leu Gly Ala Glu Ala			
100	105	110	
Asp Asn Thr Asn Ile Thr Val Gly Ala Arg Pro Gly Val Ile Asp Asp			
115	120	125	
Leu Phe Thr Gly Asn Phe Leu Gly Arg Asp Ser Asp Ile Val Phe Asp			
130	135	140	
Tyr Arg Gln Lys Val Thr Arg Ser Phe Glu Tyr Leu Arg Gly Asp Tyr			
145	150	155	160
Tyr Ile Ala Pro Val Phe Leu Asp Gln Val Val Cys His Ile Val Lys			
165	170	175	
Asn Phe Met Ala His Leu Leu Asn Val Lys Val Pro Leu Ile Leu Gly			
180	185	190	
Ile Trp Gly Gly Lys Gly Gln Gly Lys Thr Phe Gln Thr Glu Leu Ile			

	195	200	205
	Phe Arg Ala Met Gly Val Glu Pro Val Ile Met Ser Ala Gly Glu Met		
	210	215	220
	Glu Ser Glu Lys Ala Gly Glu Pro Gly Arg Leu Ile Arg Glu Arg Tyr		
	225	230	235
	Arg Ala Ala Ser Gln Val Val Gln Asn His Gly Lys Leu Ser Cys Leu		240
		245	250
	Met Ile Asn Asp Leu Asp Ala Gly Ile Gly Arg Phe Gly Asn Thr Gln		255
		260	265
	Val Thr Val Asn Asn Gln Thr Ala Ala Gly Thr Leu Met Asn Leu Ala		270
		275	280
	Asp Asn Pro Thr Arg Val Ser Ile Gly Gln Lys Trp Arg Glu Asn Asp		285
	290	295	300
	Thr Thr His Arg Ile Pro Val Ile Val Thr Gly Asn Asp Phe Ser Thr		
	305	310	315
			320
	Ile Tyr Ala Pro Leu Ile Arg Asp Gly Arg Met Asp Lys Phe Tyr Trp		
		325	330
	Gln Pro Thr His Asp Asp Ile Val Asn Ile Val His Arg Met Tyr Glu		335
		340	345
	Lys Asp Gly Ile Ser Trp Asp Glu Val Val Ser Ile Val Asn Thr Phe		350
	355	360	365
[0035]	Pro Asn Gln Ala Leu Asp Phe Tyr Gly Ala Met Arg Ser Arg Thr Tyr		
	370	375	380
	Asp Arg Ser Val Leu Lys Trp Val Asp Asp Ile Gly Gly Ile Asp Asn		
	385	390	395
	Leu Gly Ala Lys Leu Leu Gln Leu Arg Lys Gly Asp Glu Leu Pro Val		400
		405	410
	Phe Val Pro Pro Glu Gln Asn Val Glu Ala Leu Leu Glu Ser Gly Tyr		415
		420	425
	Ser Leu Leu Arg Glu Gln Gln Leu Ile Asn Lys Thr Lys Leu Ala Lys		430
		435	440
	Glu Tyr Met Lys Asn Met Glu Asp		445
	450	455	
	<210> 29		
	<211> 429		
	<212> PRT		
	<213> 人工序列		
	<220>		
	<223> 来自“冷”RCA序列的共有序列		
	<220>		
	<221> misc_feature		
	<222> (9)..(9)		
	<223> Xaa 可以是任何天然氨基酸		
	<220>		

	<221> misc_feature
	<222> (17).. (17)
	<223> Xaa 可以是任何天然氨基酸
	<220>
	<221> misc_feature
	<222> (20).. (20)
	<223> Xaa 可以是任何天然氨基酸
	<220>
	<221> misc_feature
	<222> (23).. (23)
	<223> Xaa 可以是任何天然氨基酸
	<220>
	<221> misc_feature
	<222> (28).. (28)
	<223> Xaa 可以是任何天然氨基酸
	<220>
	<221> misc_feature
	<222> (30).. (30)
	<223> Xaa 可以是任何天然氨基酸
	<220>
	<221> misc_feature
	<222> (38).. (38)
[0036]	<223> Xaa 可以是任何天然氨基酸
	<220>
	<221> misc_feature
	<222> (41).. (41)
	<223> Xaa 可以是任何天然氨基酸
	<220>
	<221> misc_feature
	<222> (43).. (43)
	<223> Xaa 可以是任何天然氨基酸
	<220>
	<221> misc_feature
	<222> (46).. (46)
	<223> Xaa 可以是任何天然氨基酸
	<220>
	<221> misc_feature
	<222> (49).. (50)
	<223> Xaa 可以是任何天然氨基酸
	<220>
	<221> misc_feature
	<222> (57).. (58)
	<223> Xaa 可以是任何天然氨基酸
	<220>
	<221> misc_feature

<222> (63).. (63)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (97).. (97)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (124).. (124)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (150).. (150)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (186).. (186)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (202).. (202)
<223> Xaa 可以是任何天然氨基酸
[0037] <220>
<221> misc_feature
<222> (213).. (213)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (217).. (217)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (245).. (245)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (248).. (248)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (254).. (254)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (256).. (256)

<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (271).. (271)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (277).. (277)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (280).. (281)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (332).. (333)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (337).. (337)
<223> Xaa 可以是任何天然氨基酸
<220>
[0038] <221> misc_feature
<222> (360).. (361)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (367).. (367)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (387).. (387)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (395).. (395)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (405).. (405)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (415).. (415)
<223> Xaa 可以是任何天然氨基酸

```

<220>
<221> misc_feature
<222> (427)..(427)
<223> Xaa 可以是任何天然氨基酸
<220>
<221> misc_feature
<222> (429)..(429)
<223> Xaa 可以是任何天然氨基酸
<400> 29
Met Ala Ala Ala Val Val Gly Ala Xaa Asn Arg Ala Pro Leu Ser Leu
1          5          10          15
Xaa Gly Ser Xaa Ser Gly Xaa Ser Ser Val Pro Xaa Pro Xaa Phe Leu
          20          25          30
Gly Lys Lys Leu Thr Xaa Ser Cys Xaa Phe Xaa Gln Ser Xaa Ser Ser
          35          40          45
Xaa Xaa Asn Gly Ser Phe Lys Val Xaa Xaa Glu Val Gly Glu Xaa Lys
          50          55          60
Leu Ser Asp Gln Ser Ser Trp Glu Ala Leu Asp Ala Asp Gly Ser Asp
65          70          75          80
Tyr Leu Tyr Arg Leu Gly Glu Ala Asp Asn Met Asn Ile Ala Val Gly
          85          90          95
Xaa Arg Gly Met Val Asp Ser Leu Phe Gln Gly Pro Phe Leu Gly Thr
[0039]          100          105          110
Asp Ser Asp Ile Val Phe Asp Tyr Arg Gln Lys Xaa Thr Arg Ser Phe
          115          120          125
Glu Tyr Leu Gln Gly Asp Tyr Tyr Ile Ala Pro Ala Phe Met Asp Lys
          130          135          140
Val Val Val His Ile Xaa Lys Asn Phe Leu Ala Asn Leu Leu Asn Ile
          145          150          155          160
Lys Val Pro Leu Ile Leu Gly Ile Trp Gly Gly Lys Gly Gln Gly Lys
          165          170          175
Ser Phe Gln Thr Glu Leu Ile Phe Ala Xaa Met Gly Ile Glu Pro Val
          180          185          190
Ile Met Ser Ala Gly Glu Leu Glu Ser Xaa Arg Ala Gly Glu Pro Gly
          195          200          205
Lys Leu Ile Arg Xaa Arg Tyr Arg Xaa Ala Ser Gln Val Ile Gln Asn
          210          215          220
Gln Gly Lys Met Ser Cys Leu Met Ile Asn Asp Leu Asp Ala Gly Ala
          225          230          235          240
Gly Arg Phe Gly Xaa Thr Gln Xaa Thr Val Asn Asn Gln Xaa Val Xaa
          245          250          255
Gly Thr Leu Met Asn Leu Ala Asp Asn Pro Thr Arg Val Ser Xaa Gly
          260          265          270
Gln Met Trp Arg Xaa Ser Asp Xaa Xaa Arg Val Pro Ile Ile Val Thr
          275          280          285

```

Gly Asn Asp Phe Ser Thr Leu Tyr Ala Pro Leu Ile Arg Asp Gly Arg
 290 295 300
 Met Glu Lys Phe Tyr Trp Gln Pro Thr Arg Glu Asp Ile Ile Asn Ile
 305 310 315 320
 Val His Arg Met Tyr Glu Lys Asp Gly Ile Ser Xaa Xaa Glu Val Val
 325 330 335
 Xaa Ile Val Asp Thr Phe Pro Asn Gln Ala Leu Asp Phe Tyr Gly Ala
 340 345 350
 Leu Arg Ser Arg Thr Tyr Asp Xaa Xaa Val Leu Lys Trp Val Xaa Asp
 355 360 365
 Ile Gly Gly Val Glu Asn Leu Gly Lys Lys Leu Leu Asn Ser Lys Lys
 370 375 380
 Glu Gly Xaa Pro Val Phe Glu Pro Pro Glu Xaa Thr Leu Glu Ala Leu
 385 390 395 400
 Leu Glu Ser Gly Xaa Ser Leu Val Lys Glu Gln Glu Leu Val Xaa Arg
 405 410 415
 Val Lys Leu Ala Lys Glu Tyr Met Lys Asn Xaa Asp Xaa
 420 425

<210> 30

<211> 427

<212> PRT

<213> 人工序列

[0040] <220>

<223> 来自亚基因组 B 的 TaRCA 2b 的修饰序列

<400> 30

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

145 150 155 160
 Lys Ser Phe Gln Cys Glu Leu Val Phe Ala Lys Met Gly Ile Asn Pro
 165 170 175
 Ile Met Met Ser Ala Gly Glu Leu Glu Ser Gly Asn Ala Gly Glu Pro
 180 185 190
 Ala Lys Leu Ile Arg Gln Arg Tyr Arg Glu Ala Ala Asp Ile Ile Lys
 195 200 205
 Lys Gly Lys Met Cys Cys Leu Phe Ile Asn Asp Leu Asp Ala Gly Ala
 210 215 220
 Gly Arg Met Gly Gly Thr Thr Gln Tyr Thr Val Asn Asn Gln Met Val
 225 230 235 240
 Asn Ala Thr Leu Met Asn Ile Ala Asp Ala Pro Thr Asn Val Gln Leu
 245 250 255
 Pro Gly Met Tyr Asn Lys Glu Glu Asn Pro Arg Val Pro Ile Val Val
 260 265 270
 Thr Gly Asn Asp Phe Ser Thr Leu Tyr Ala Pro Leu Ile Arg Asp Gly
 275 280 285
 Arg Met Glu Lys Phe Tyr Trp Ala Pro Thr Arg Asp Asp Arg Ile Gly
 290 295 300
 Val Cys Lys Gly Ile Phe Arg Thr Asp Asn Val Pro Asp Glu Ser Val
 305 310 315 320
 Val Lys Leu Val Asp Thr Phe Pro Gly Gln Ser Ile Asp Phe Phe Gly
 [0041] 325 330 335
 Ala Leu Arg Ala Arg Val Tyr Asp Asp Glu Val Arg Lys Trp Val Thr
 340 345 350
 Glu Thr Gly Ile Glu Asn Ile Gly Lys Arg Leu Val Asn Ser Arg Asp
 355 360 365
 Gly Pro Val Thr Phe Glu Gln Pro Lys Met Thr Ile Glu Lys Leu Leu
 370 375 380
 Glu Tyr Gly His Met Leu Val Gln Glu Gln Asp Asn Val Lys Arg Val
 385 390 395 400
 Gln Leu Ala Asp Lys Tyr Leu Ser Glu Ala Ala Leu Gly Asp Ala Asn
 405 410 415
 Gln Asp Ala Met Lys Thr Gly Ser Phe Tyr Gly
 420 425

<210> 31

<211> 1143

<212> DNA

<213> 人工序列

<220>

<223> 修饰的 TaRCA 2b 序列, 来自亚基因组 B, 具有 11 个氨基酸置换

<400> 31

atggccgcgcg aaaatctaga tgaaaaacgt aacaccgata aatggaaagg tttagcatat 60

gacatctcag atgaccagca ggatataacc agaggtaaag gcattgtaga ttctcttttc 120

caagcaccta cgggggatgg aaccacagag gctgttctta gttcgtatga atatatctct 180

cagggattaa aaaagtatga tttcgataat accatggatg gtttctacat cgcgcccga	240
tttatggaca aactgtagt tcacctctt aaaaacttca tgacgtccc taacataaaa	300
atacctctga tattaggcat ctggggagga aaaggccaag gtaagagctt tcaatgtgaa	360
ctggtgttcg cgaagatggg cataaacct atcatgatgt ccgctggaga gttagaaagt	420
ggtaatgcgg gtgaaccagc taaattgatt cgtcaaaggt atagagagc tgctgacatc	480
atcaaaaagg gaaaaatgtg ttgcttggtc attaatgacc tcgacgccg cgcgggccgc	540
atgggtggga cgactcaata cacagtgaac aaccagatgg tgaacgtac tttaatgaac	600
atcgtgacg caccgactaa tgtacagcta cccggcatgt ataataaaga ggaaaaccca	660
cgtgtcccga tcgttggtac gggcaatgac tttccactc tatatgcacc gttaattcgc	720
gatgtagga tggaaaaatt ctattgggcc ccgactcgcg atgatcgaat tggcgtctgt	780
aagggtatat ttagaacaga taatgtgcca gatgaatcag tagtgaaact ttagatacc	840
tttccgggtc agtccattga cttcttggg gcgttaagag caagggttta tgacgatgag	900
gtacgcaaat gggttaccga aacgggcatt gagaatattg gtaaacgtct tgttaactca	960
agagacggac cagtaacctt tgaacagcca aaaatgacca tcgaaaaact ttggaatat	1020
ggcatatgt tagtccagga gcaagataac gtcaagcgtg tccaactggc agataataac	1080
ttaagtgaag ctgccctggg agatgcaaac caggatgcga tgaagacagg tagcttttat	1140
ggg	1143
<210> 32	
<211> 381	
<212> PRT	
<213> 人工序列	
<220>	
[0042] <223> 修饰的 TaRCA 2b , 具有 11 个氨基酸置换	
<400> 32	
Met Ala Ala Glu Asn Leu Asp Glu Lys Arg Asn Thr Asp Lys Trp Lys	
1 5 10 15	
Gly Leu Ala Tyr Asp Ile Ser Asp Asp Gln Gln Asp Ile Thr Arg Gly	
20 25 30	
Lys Gly Ile Val Asp Ser Leu Phe Gln Ala Pro Thr Gly Asp Gly Thr	
35 40 45	
His Glu Ala Val Leu Ser Ser Tyr Glu Tyr Ile Ser Gln Gly Leu Lys	
50 55 60	
Lys Tyr Asp Phe Asp Asn Thr Met Asp Gly Phe Tyr Ile Ala Pro Ala	
65 70 75 80	
Phe Met Asp Lys Leu Val Val His Leu Ser Lys Asn Phe Met Thr Leu	
85 90 95	
Pro Asn Ile Lys Ile Pro Leu Ile Leu Gly Ile Trp Gly Gly Lys Gly	
100 105 110	
Gln Gly Lys Ser Phe Gln Cys Glu Leu Val Phe Ala Lys Met Gly Ile	
115 120 125	
Asn Pro Ile Met Met Ser Ala Gly Glu Leu Glu Ser Gly Asn Ala Gly	
130 135 140	
Glu Pro Ala Lys Leu Ile Arg Gln Arg Tyr Arg Glu Ala Ala Asp Ile	
145 150 155 160	
Ile Lys Lys Gly Lys Met Cys Cys Leu Phe Ile Asn Asp Leu Asp Ala	

					165					170						175
	Gly	Ala	Gly	Arg	Met	Gly	Gly	Thr	Thr	Gln	Tyr	Thr	Val	Asn	Asn	Gln
					180					185						190
	Met	Val	Asn	Ala	Thr	Leu	Met	Asn	Ile	Ala	Asp	Ala	Pro	Thr	Asn	Val
					195					200						205
	Gln	Leu	Pro	Gly	Met	Tyr	Asn	Lys	Glu	Glu	Asn	Pro	Arg	Val	Pro	Ile
					210					215						220
	Val	Val	Thr	Gly	Asn	Asp	Phe	Ser	Thr	Leu	Tyr	Ala	Pro	Leu	Ile	Arg
	225						230					235				240
	Asp	Gly	Arg	Met	Glu	Lys	Phe	Tyr	Trp	Ala	Pro	Thr	Arg	Asp	Asp	Arg
							245					250				255
	Ile	Gly	Val	Cys	Lys	Gly	Ile	Phe	Arg	Thr	Asp	Asn	Val	Pro	Asp	Glu
							260					265				270
	Ser	Val	Val	Lys	Leu	Val	Asp	Thr	Phe	Pro	Gly	Gln	Ser	Ile	Asp	Phe
							275									285
	Phe	Gly	Ala	Leu	Arg	Ala	Arg	Val	Tyr	Asp	Asp	Glu	Val	Arg	Lys	Trp

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 修饰的 TaRCA 2b, 来自亚基因组 B, 具有 8 个氨基酸置换	
	<400> 34	
	atggccgcgg aaaatctgga cgagaaacgc aataccgata aatggaaagg tctggcctat	60
	gatatctctg atgacagca ggatattaca cgtggtaaag gcattgtgga ttcactgttt	120
	caggcaccga ctggtgatgg aaccacagaa gcagtattgt ctagctacga atatatctct	180
	caggggctta aaaagtatga ttttgacaac accatgggcg gtttttatat cgcgccggca	240
	ttcatggata agcttgttgt ccacctgtcg aaaaacttta tgacgtctcc taatattaaa	300
	attccattga ttctcggtat ttgggtggt aaagccagg gtaagtcttt ccagtgtgag	360
	ttagtttttg ccaagatggg cattaatccg atcatgatga gtgccgggga gctagaaagt	420
	ggtaatgccg gtgaacctgc taaacttatt cgccaacgct accgcgaagc cgcagacatc	480
	atcaaaaaag gtaagatgtg ttgcttattc atcaatgatc tcgatgctgg agctgggcgc	540
	atgggtggga caaccagta tactgttaac aatcagatgg tcaatgcaac gctgatgaac	600
	atcgtgatg cccaaccaa tgtgcaatta ccaggcatgt acaataagga agaaaatccc	660
	cgtgttccca tcgtagtac tggaaatgac tttagcacac tctatgcgcc cttattctgt	720
	gacgtagaa tggaaaaatt ttattgggcc ccaactaggg acgatcgtat cggcgtctgt	780
	aagggtatat tcaggacgga caatgtcccc gatgagagtg ttgttaagat cgttgataca	840
	ttccaggtc agtctataga tttcttcggt gcactgcggg cgcgggtcta cgacgatgag	900
	gtccgaaat gggtgacttc taccggaatt gaaaatateg gaaaacgtct cgtcaatagt	960
	agggacggcc cagtgcctt tgagcaaccc aaaatgacaa tagaaaagct tttggagtat	1020
[0045]	ggtcatatgc tcgttcaaga gcaagataac gttaagcgcg ttcagctggc agataaatat	1080
	ctgagcgagg cggccttggg ggatgcta atcaggacgcaa tgaaaaccgg gtccttctac	1140
	gga	1143
	<210> 35	
	<211> 381	
	<212> PRT	
	<213> 人工序列	
	<220>	
	<223> 来自亚基因组 B 的修饰 TaRCA 2b, 具有 8 个氨基酸置换	
	<400> 35	
	Met Ala Ala Glu Asn Leu Asp Glu Lys Arg Asn Thr Asp Lys Trp Lys	
	1 5 10 15	
	Gly Leu Ala Tyr Asp Ile Ser Asp Asp Gln Gln Asp Ile Thr Arg Gly	
	20 25 30	
	Lys Gly Ile Val Asp Ser Leu Phe Gln Ala Pro Thr Gly Asp Gly Thr	
	35 40 45	
	His Glu Ala Val Leu Ser Ser Tyr Glu Tyr Ile Ser Gln Gly Leu Lys	
	50 55 60	
	Lys Tyr Asp Phe Asp Asn Thr Met Gly Gly Phe Tyr Ile Ala Pro Ala	
	65 70 75 80	
	Phe Met Asp Lys Leu Val Val His Leu Ser Lys Asn Phe Met Thr Leu	
	85 90 95	
	Pro Asn Ile Lys Ile Pro Leu Ile Leu Gly Ile Trp Gly Gly Lys Gly	

	100	105	110
	Gln Gly Lys Ser Phe Gln Cys Glu Leu Val Phe Ala Lys Met Gly Ile		
	115	120	125
	Asn Pro Ile Met Met Ser Ala Gly Glu Leu Glu Ser Gly Asn Ala Gly		
	130	135	140
	Glu Pro Ala Lys Leu Ile Arg Gln Arg Tyr Arg Glu Ala Ala Asp Ile		
	145	150	155
	Ile Lys Lys Gly Lys Met Cys Cys Leu Phe Ile Asn Asp Leu Asp Ala		
	165	170	175
	Gly Ala Gly Arg Met Gly Gly Thr Thr Gln Tyr Thr Val Asn Asn Gln		
	180	185	190
	Met Val Asn Ala Thr Leu Met Asn Ile Ala Asp Ala Pro Thr Asn Val		
	195	200	205
	Gln Leu Pro Gly Met Tyr Asn Lys Glu Glu Asn Pro Arg Val Pro Ile		
	210	215	220
	Val Val Thr Gly Asn Asp Phe Ser Thr Leu Tyr Ala Pro Leu Ile Arg		
	225	230	235
	Asp Gly Arg Met Glu Lys Phe Tyr Trp Ala Pro Thr Arg Asp Asp Arg		
	245	250	255
	Ile Gly Val Cys Lys Gly Ile Phe Arg Thr Asp Asn Val Pro Asp Glu		
	260	265	270
	Ser Val Val Lys Ile Val Asp Thr Phe Pro Gly Gln Ser Ile Asp Phe		
[0046]	275	280	285
	Phe Gly Ala Leu Arg Ala Arg Val Tyr Asp Asp Glu Val Arg Lys Trp		
	290	295	300
	Val Thr Ser Thr Gly Ile Glu Asn Ile Gly Lys Arg Leu Val Asn Ser		
	305	310	315
	Arg Asp Gly Pro Val Thr Phe Glu Gln Pro Lys Met Thr Ile Glu Lys		
	325	330	335
	Leu Leu Glu Tyr Gly His Met Leu Val Gln Glu Gln Asp Asn Val Lys		
	340	345	350
	Arg Val Gln Leu Ala Asp Lys Tyr Leu Ser Glu Ala Ala Leu Gly Asp		
	355	360	365
	Ala Asn Gln Asp Ala Met Lys Thr Gly Ser Phe Tyr Gly		
	370	375	380

<210> 36

<211> 4326

<212> DNA

<213> 人工序列

<220>

<223> T-DNA

<400> 36

aattacaacg gtatatatcc tgccagtact gggccccctc gagggcgatc gctacgtacc 60

tgcaggcccg ggtaattaa gctccgatt gatgcttcac caaaaaacaa tatcaacagc 120

agtgcaaaat tagaatTTTT gtatttttGT ggtaacgcaa accggccatc aaaggggaaa 180

	aacgtacaat gcttatgttg tatgttaaga gaagtttgtg tgggtgcaaaa tgacagtcct	240
	agcctgacgg ttatcgagaa agcagaatat gtgcaggtgg cagagcaaaa tttttgtgt	300
	agtccaacta gaatacaatt tgcatgccat gcctcatcca agaagccggg caacgagagg	360
	cagcaaaagg cttttctgtg gtgatgcaaa atgaagaggt tatgtagtag ctgagctgat	420
	gaagcaactg gtcgctagct gccggccggg agacgaatgt gaggcaagga aagaaaagaa	480
	aaaacagaga gaaagagttg atcagaaatg ggtgaattct gtggtgagga aaggtcaagg	540
	aactgaagcc aagagatcct tctacctac actaatacaa tactcctaac tcgctcacag	600
	actccgatcc aggtccaagt catgctatgc tgtggatcgg ccggccgaga ttgcgccacg	660
	tgtgcagaac ccaatcttca gcgtgtggcc tgtgggggat ctggaagctg atccacaggg	720
	agggaggagt gtgtgcctct cacagcttcc aacttccatg gcgacgtcca atgctattgt	780
	attatttaag gcctaccgca gctcggcctc tacactttga gcagcagcgg ctggccatca	840
	tcagtgatcc tctacaatca tcgactttca gcaaattaag tccaaaacca tggcatcagc	900
	attttcgtcc actgtcggag ctctcgcac tactcctaca accttcttgg gcaaaaaggt	960
	gaagaaacaa gctggcgcgc taaactatta ccacggcgga aacaagatca acaatcgct	1020
	cgtacgggca atggctgcaa agaaagaact ggacgaaggg aagcagacag atgcagatcg	1080
	ctggaaaggc cttgcctatg acataagcga cgatcagcaa gacataaccc gaggaaaggg	1140
	gatagtcgat agcctgttcc aagctccgat gggatgatgg acacatgagg cgatcctatc	1200
	tagctacgag tacatctccc aaggactccg gaaatacgac ttcgacaaca caatggacgg	1260
	cttttacatc gctcctgcat tcatggacaa gctgacgtc cacttggcta agaactttat	1320
	gacgtgccc aacatcaaag tcccgtaat cttgggaatt tggggaggga aaggccaagg	1380
	aaagagcttt caatgcgagc tagtgttcgc gaaaaaggg atcaaccgga taatgatgag	1440
	cgctggagaa cttgagagtg gaaacgcggg agaaccagct aagctcatac gacagaggta	1500
[0047]	cagagaggca gcggacatca taaaaaaggg gaagatgtgc tgcttattca tcaacgactt	1560
	ggatgctgga gcaggaagga tggggggaac gactcagtag acggtcaata accagatggt	1620
	caacgcgacc ctaatgaaca tcgctgatgc tcccactaac gtccaattcc cggggatgta	1680
	caacaaggaa gagaatccgc gtgttccgat cattgtgacg gggaacgact tcagcacctt	1740
	atatgctcct ctcatacggg atggacggat ggagaaattc tattgggctc ctacgaggga	1800
	ggatagaatt ggcgtctgca aagggatctt ccgaactgac aacgttccgg atgaggtgt	1860
	cgttcaacta gttgacacct tccctgggtc gagcattgat ttctttggcg ccttaagagc	1920
	tcgctatac gacgatgaag tccgaaaatg ggtcggggag ataggagtcg agaacatctc	1980
	taagcgccta gttactcac gagaaggacc tcccacattc gaccaacca agatgacgat	2040
	cgagaagctc atggagtacg gccatatgct agtccaagag caagagaacg tcaagcgagt	2100
	tcaactcgcg gacaaatata tgagtgaagc tgcgttggga caagcgaacg atgacgcaat	2160
	ggctactgga gctttctacg ggaagtgatt gctagcacgc gtgcaggccg gcttgacgct	2220
	tgctattaat tatttctcct ttttatgttt tctctttgtt ctgtgtatca gatcgcgcc	2280
	aagccatagc tgggcatgac aagtttttgt taataataat ataataataa gatatgggct	2340
	atactataca tgagtatatt tcatttgggt atcatttgtc aatttctgtg tccatccagt	2400
	gttttcttag aacaatactg taagacggtt tactgtcagc caaatgcta aattaagcaa	2460
	tggatatcat taccctgtta tccctaaagc ttattaatat aacttcgtat agcatacatt	2520
	atacgaagtt atgtttccta cgcagcaggt ctcatcaaga cgatctacce gagtaacaat	2580
	ctccaggaga tcaaatacct tcccagaag gttaaagatg cagtcaaaag attcaggact	2640
	aattgcatca agaacacaga gaaagacata tttctcaaga tcagaagtac tattccagta	2700
	tggacgattc aaggcttgct tcataaacca aggcaagtaa tagagattgg agtctctaaa	2760
	aaggtagttc ctactgaatc taaggccatg catggagtct aagattcaaa tcgaggatct	2820
	aacagaactc gccgtgaaga ctggcgaaca gttcatacag agtcttttac gactcaatga	2880

	caagaagaaa atcttcgtca acatggtgga gcacgacact ctggtctact ccaaaaatgt	2940
	caaagataca gtctcagaag accaaagggc tattgagact ttccaacaaa ggataatttc	3000
	gggaacctc ctcggattcc attgccagc tatctgtcac ttcacgaaa ggacagtaga	3060
	aaaggaaggt ggctcctaca aatgccatca ttgcgataaa ggaaaggcta tcattcaaga	3120
	tgctctgcc gacagtggte ccaaagatgg acccccaccc acgaggagca tcgtggaaaa	3180
	agaagacgtt ccaaccacgt cttcaaagca agtggattga tgtgacatct ccactgacgt	3240
	aagggatgac gcacaatccc actatccttc gcaagaccct tcctctatat aaggaagttc	3300
	atttcatttg gagaggacac gctgaaatca ccagtccttc tctataaatc tatctctctc	3360
	tctataacaa tggaccacaga acgacgccc gccgacatcc gccgtgccac cgaggcgac	3420
	atgccggcgg tctgcacat cgtcaaccac tacatcgaga caagcacggt caacttccgt	3480
	accgagccgc aggaaccgca ggagtggacg gacgacctgc tccgtctgcg ggagcgtat	3540
	ccctggctcg tcgccgaggt ggacggcgag gtcccgcca tcgcctacgc ggcccctgg	3600
	aaggcacgca acgcctacga ctggacggcc gagtcgaccg tgtacgtctc cccccgccac	3660
	cagcggacgg gactgggtc cagctctac acccacctgc tgaagtcct ggaggcacag	3720
	ggcttcaaga gcgtgggtcg tgtcatcggt ctgcccaacg acccgagcgt gcgcatgcac	3780
	gagcgctcg gatagcccc ccgcggcatg ctgcggcgcg ccggcttcaa gcacgggaac	3840
	tggcatgacg tgggtttctg gcagctggac ttcagcctgc cggtagccgc ccgtccggtc	3900
	ctgcccgtca ccgagatctg agatcacccg ttctaggatc cgaagcagat cgttcaaaca	3960
	tttgcaata aagtttctta agattgaatc ctgttgccgg tcttgcatg attatcatat	4020
	aatttctgtt gaattacgtt aagcatgtaa taattaacat gtaatcatg acgttattta	4080
	tgagatgggt ttttatgatt agagtccgc aattatacat ttaatacgcg atagaaaaca	4140
	aaatatagcg cgaaactag gataaattat cgcgcgcggt gtcatctatg ttactagatc	4200
[0048]	gaaacataac ttcgtatagc atacattata cgaagtata tggatctcga ggcatcagg	4260
	cattacggca ctgcgaggg tcccaattcg agcatggagc catttacaat tgaatatatc	4320
	ctgccg	4326
	<210> 37	
	<211> 6760	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> T-DNA	
	<400> 37	
	aattacaacg gtatatatcc tgccagtact gggccccctc gagggcgatc gctacgtacc	60
	tgcaggcgcg ccgcaagctt ttggtggtag gaatgtagtt ttctgacaaa gtcaattact	120
	gaatataaaa aaaatctgca cagctctgcg tcaacagttg tccaagggat gcctcaaaaa	180
	tctgtgcaga ttatcagtcg tcacgcagaa gcagaacatc atggtgtgct aggtcagctt	240
	cttgcatggt gccatgaatc cggttggttg ttaatctctc ctctcttatt ctcttatatt	300
	aagatgcata actcttttat gtagtctaaa aaaaaatcca gtggatcgga tagtagtacg	360
	tcatggtgcc attaggtacc gttgaacctc acagatattt atgcatgtgt atatatatag	420
	ctatatagac aaaattggtg ccgattatag acccaaaagc aataggtata tataatataa	480
	tacagaccac accaccaaac taagaatcga tcaaatagac aaggcatgtc tccaaattgt	540
	cttaaaactat ttccgtaggt tcagccgttc aggagtcgaa tcagcctctg ccggcgtttt	600
	ctttgcacgt acgacggaca cacatgggca taccatatag ctggtccatg acattaggag	660
	agagaacgta cgtgttgacc tgtagctgag atataacaag gttgattata atatcaccaa	720
	acatgaaatc atccaaggat gaccataac tatcactact atagtactgc atctggtaaa	780

	agaaattgta tagactctat ttcgagcact accacataac gcctgcaatg tgacacccta	840
	cctattcact aatgtgecte ttcccacacg ctttccaccc gtactgtcga cagctttaag	900
	aaccagaaca aatgagtaat attagtgtcg gttcatggct aaaaccagca ctgatgtaca	960
	tgaccacata tgtcaaatgc tgcttctagg catgaccgcg tcttactaat acctactcat	1020
	cgctagaaga attttcggct gataaatfff caatttaagc aagagttatc cgcgttggtt	1080
	cataactcaa actgatggcc ccaaccatat tagtgcaaat ttcacatatg atcataacct	1140
	tttcatatga aatcggatcg tgatgaactt tatataaaca ttgtagctgt cgatgatacc	1200
	tacaatttta tagttcacia cttttttatt tcaagtcatt taaatgccc aatagggtgt	1260
	tcaaacttca gatagaaatg ttcaaaagta aaaaaggtcc ctatcataac ataattgata	1320
	tgtaagttag ttggaaaatg ataagtacgt gtgagagaga tcggagatca aattctggtg	1380
	taataatgta tgtatttcag tcataaaaat tggtagcagt agttggggct ctgtatatat	1440
	accggaagg atgggatggt agtagaataa ttctttttt gtttttagtt ttttctggtc	1500
	caaaatttca aatttgatc ctttacttgt accaactaat attaatgagt gttgaggcca	1560
	gtagaggtgc aactttacca taatccctct gtttcagggt ataagacgtt ttgactttaa	1620
	atttgactaa gtttatgcgc aaatatagta atatttataa tactaaatta gtttcattaa	1680
	ataaataatt gaatatattt tcataataaa ttgtgttga gttcaaaata ttattaattt	1740
	tttctacaaa cttgggtcaa cttaaagcag ttgactttg accaaagtca aaacgtctta	1800
	taacttgaag cggatggatt actttttttg tggggacaag tttaaatgt ttaataaaag	1860
	cacaatccat cttaatgttt tcaagctgaa tattgtaaaa ttcatggata aaccagcttc	1920
	taaatgttta accgggaaaa tgcgaacga caaattaata tttttaagt atggggagta	1980
	ttaattaagg agtgacaact caactttcaa tatcgtacta aactgtggga tttattttct	2040
	aaaattttat accctgcaa ttacagtggt gtagatcttt ttttttact aaccgacacc	2100
[0049]	aggtatatca attttattga atatagcagc aaaaagaatg tgttgtactt gtaaacaaaa	2160
	agcaactgt acataaaaa aaatgcactc ctatataatt aagctcataa agatgctttg	2220
	cttcgtgagg gcccaagttt tgatgacctt ttgcttgatc tcgaaattaa aatttaagta	2280
	ctgttaaggg agttcacacc accatcaatt ttccagctga agaaacagtt aaacaacgac	2340
	cccgatgacc agtctactgc tctccacata ctagtgcatt tattgatcac aaaacaaac	2400
	aaaacgaaat aaaaatcagc agcgagagtg tgcagagaga gacaaaggtg atctggcgtg	2460
	gatatctccc catccatcct caccgcgct gccatcact cgcgcgcga tactccatca	2520
	tgtggagaga ggaagacgag gaccacagcc agagcccggg tcgagatgcc accacggcca	2580
	caaccacga gcccggcgcg acaccaccgc gcgcgcgtga gccagccaca aacgccgcg	2640
	gataggcgcg cgcacgccgg ccaatcctac cacatcccc gcctccgcgg ctgcgagcg	2700
	ccgtgccat ccgatccgt gagttttggc tttttatac tactcgcgga gcctgtgtgc	2760
	agagcagtgc atctcaagaa gtactcgagc aaagaaggag agagcttggg gagctgcaga	2820
	gcagtaacc accccgcccc tctcctcttt ctttctcgt ttttttttc cgtctcggtc	2880
	tcgacttttg gccttggtag ttgggtggg cgagaggcgg cttcgtgcgc gccagatcg	2940
	gtgcgcggga gggcgggat ctgcggctg gggctctgc cggcgtggat ccggcccga	3000
	tctcggggg aatggggctc tcggatgtag atctcgatc cgcgttggt gggggagatg	3060
	atggggggtt taaaatttcc gccatgctaa acaagatcag gaagaggga aaaggcact	3120
	atggtttata tttttatata tttctgtgc ttctcagc ttagatgtgc tagatcttc	3180
	tttctcttt ttgtggtag aatttgaatc cctcagcatt gttcatcggt agttttctt	3240
	ttcatgattt gtgacaaatg cagcctcgtg cggagctttt ttgtaggtag accatggcat	3300
	cagcattttc gtccactgtc ggagctcctg catctactcc tacaaccttc ctgggcaaaa	3360
	aggtgaagaa acaagctggc gcgctaaact attaccacgg cggaacaag atcaacaatc	3420
	gcgtcgtacg ggcaatggct gcaagaaaag aactggacga agggaagcag acagatgcag	3480

	atcgtctggaa aggccttgcc tatgacataa gcgacgatca gcaagacata acccgaggaa	3540
	aggggatagt cगतगcctg ttccaagctc cगतगgggtga tggcacacat gaggcgatcc	3600
	tatctagcta cगतगtacatc tccaagggac tccggaaata cgacttcgac aacacaatgg	3660
	acggccttta catcgtcctt gcattcatgg acaagctgat cgtccacttg gctaagaact	3720
	ttatgacgct gcccaacatc aaagtccgcg taatcttggg aatttgggga gggaaaggcc	3780
	aaggaaagag ctttcaatgc gagctagtgt tcgcgaaaat ggggatcaac ccgataatga	3840
	tgagcgctgg agaacttgag agtggaaacg cggggaacc agctaagctc atacgacaga	3900
	ggtacagaga ggcagcggac atcataaaaa aggggaagat gtgctgctta ttcataacg	3960
	acttgatgc tggagcagga aggatggggg gaacgactca gtacacggtc aataaccaga	4020
	tggtcaacgc gacctaatg aacatcgctg atgtccccc taacgtccaa tccccgggga	4080
	tgtacaacaa ggaagagaat ccgctgttgc cगतगcattgt gacggggaac gacttcagca	4140
	ccttatatgc tctctcata cgggatggac ggatggagaa attctattgg gctctacga	4200
	gggaggatag aattggcgtc tgcaaaggga tcttccgaac tgacaacgtt ccgcatgagg	4260
	ctgtcgttca actagttgac accttccctg gtcagagcat tgatttcttt ggcgcttaa	4320
	gagctcgctg atacgacgat gaagtccgaa aatgggtcgg ggagatagga gtcgagaaca	4380
	tctctaagcg cctagttaac tcacgagaag gacctccac attcgacca cccaagatga	4440
	cगतगcagaa gctcatggag tacggccata tgctagtcca agagcaagag aacgtcaagc	4500
	gagttcaact cgcggacaaa tatctgagtg aagctgcgtt gggacaagcg aacgatgacg	4560
	caatggctac tggagctttc tacgggaagt gattgctagc acgcgtttaa ttaaatttaa	4620
	atggcgcgcc gaagcagatc gttcaaacat ttggcaataa agtttcttaa gattgaatcc	4680
	tggtgcgggt cttgcgatga ttatcatata atttctgttg aattacgtta agcatgtaat	4740
	aattaacatg taatgcatac cgttatattat gagatgggtt tttatgatta gagtcccga	4800
[0050]	attatacatt taatacgcga tagaaaacaa aatatagcgc gcaactagg ataaattatc	4860
	gcgcgggtg tcctctatgt tactagatcg gaattcgata tcattaccct gttatcccta	4920
	aagcttatta atataacttc gtatagcata cattatacga agttatgttt cctacgcagc	4980
	aggtctcatc aagacgatct acccgagtaa caatctccag gagatcaaat accttccaa	5040
	gaaggttaaa gatgcagtca aaagattcag gactaattgc atcaagaaca cagagaagaa	5100
	catatttctc aagatcagaa gtactattcc agtatggacg attcaaggct tgcttcataa	5160
	accaaggcaa gtaatagaga ttggagtctc taaaaaggta gttcctactg aatctaaggc	5220
	catgcatgga gtctaagatt caaatcgagg atctaacaga actcgcctg aagactggcg	5280
	aacagttcat acagagtctt ttacgactca atgacaagaa gaaaatcttc gtcaacatgg	5340
	tggagcacga cactctggtc tactccaaaa atgtcaaaga tacagtctca gaagacaaa	5400
	gggctattga gacttttcaa caaaggataa tttcgggaaa cctctcggga ttccattgcc	5460
	cagctatctg tcacttcacg gaaaggacag tagaaaagga aggtggctcc tacaaatgcc	5520
	atcattgcga taaaggaaag gctatcattc aagatgcctc tgccgacagt ggtcccaag	5580
	atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc acgtcttcaa	5640
	agcaagtgga ttgatgtgac atctccactg acgtaaggga tgacgcacaa tcccactatc	5700
	cttcgcaaga ccttctctct atataaggaa gttcatttca ttggagagg acacgtgaa	5760
	atcaccagtc tctctctata aatctatctc tctctctata acaatggacc cagaacgacg	5820
	cccggccgac atccgccgtg ccaccgaggc ggacatgccg gcggtctgca ccatcgtaa	5880
	ccactacatc gagacaagca cggtaactt cgtaccgag ccgcaggaac cgcaggagtg	5940
	gacggacgac ctctctcgtc tgcgggagcg ctatccctgg ctctctcggc aggtggacgg	6000
	cgaggtcgcc ggcatcgcct acgcgggccc ctggaaggca cgcaacgcct acgactggac	6060
	ggccgagtcg accgtgtacg tctcccccg ccaccagcgg acgggactgg gctccacgt	6120
	ctacaccac ctgctgaagt ccctggaggc acagggttc aagagcgtgg tcgtgtcat	6180

133

	Thr	Asn	Phe	Leu	Gly	Lys	Lys	Leu	Lys	Lys	Gln	Val	Thr	Ser	Ala	Val
				20				25					30			
	Asn	Tyr	His	Gly	Lys	Ser	Ser	Lys	Ala	Asn	Arg	Phe	Thr	Val	Met	Ala
			35					40					45			
	Ala	Glu	Asn	Ile	Asp	Glu	Lys	Arg	Asn	Thr	Asp	Lys	Trp	Lys	Gly	Leu
			50					55					60			
	Ala	Tyr	Asp	Ile	Ser	Asp	Asp	Gln	Gln	Asp	Ile	Thr	Arg	Gly	Lys	Gly
	65							70					75			80
	Ile	Val	Asp	Ser	Leu	Phe	Gln	Ala	Pro	Thr	Gly	Asp	Gly	Thr	His	Glu
								85					90			95
	Ala	Val	Leu	Ser	Ser	Tyr	Glu	Tyr	Val	Ser	Gln	Gly	Leu	Lys	Lys	Tyr
								100					105			110
	Asp	Phe	Asp	Asn	Thr	Met	Gly	Gly	Phe	Tyr	Ile	Ala	Pro	Ala	Phe	Met
								115					120			125
	Asp	Lys	Leu	Val	Val	His	Leu	Ser	Lys	Asn	Phe	Met	Thr	Leu	Pro	Asn
								130					135			140
	Ile	Lys	Ile	Pro	Leu	Ile	Leu	Gly	Ile	Trp	Gly	Gly	Lys	Gly	Gln	Gly
	145							150					155			160
	Lys	Ser	Phe	Gln	Cys	Glu	Leu	Val	Phe	Ala	Lys	Met	Gly	Ile	Asn	Pro
								165					170			175
	Ile	Met	Met	Ser	Ala	Gly	Glu	Leu	Glu	Ser	Gly	Asn	Ala	Gly	Glu	Pro
								180					185			190
[0052]	Ala	Lys	Leu	Ile	Arg	Gln	Arg	Tyr	Arg	Glu	Ala	Ala	Asp	Met	Ile	Lys
								195					200			205
	Lys	Gly	Lys	Met	Cys	Cys	Leu	Phe	Ile	Asn	Asp	Leu	Asp	Ala	Gly	Ala
								210					215			220
	Gly	Arg	Met	Gly	Gly	Thr	Thr	Gln	Tyr	Thr	Val	Asn	Asn	Gln	Met	Val
	225							230					235			240
	Asn	Ala	Thr	Leu	Met	Asn	Ile	Ala	Asp	Ala	Pro	Thr	Asn	Val	Gln	Leu
								245					250			255
	Pro	Gly	Met	Tyr	Asn	Lys	Glu	Glu	Asn	Pro	Arg	Val	Pro	Ile	Val	Val
								260					265			270
	Thr	Gly	Asn	Asp	Phe	Ser	Thr	Leu	Tyr	Ala	Pro	Leu	Ile	Arg	Asp	Gly
								275					280			285
	Arg	Met	Glu	Lys	Phe	Tyr	Trp	Ala	Pro	Thr	Arg	Asp	Asp	Arg	Ile	Gly
								290					295			300
	Val	Cys	Lys	Gly	Ile	Phe	Gln	Thr	Asp	Asn	Val	Ser	Asp	Glu	Ser	Val
	305							310					315			320
	Val	Lys	Ile	Val	Asp	Thr	Phe	Pro	Gly	Gln	Ser	Ile	Asp	Phe	Phe	Gly
								325					330			335
	Ala	Leu	Arg	Ala	Arg	Val	Tyr	Asp	Asp	Glu	Val	Arg	Lys	Trp	Val	Thr
								340					345			350
	Ser	Thr	Gly	Ile	Glu	Asn	Ile	Gly	Lys	Lys	Leu	Val	Asn	Ser	Arg	Asp
								355					360			365
	Gly	Pro	Val	Thr	Phe	Glu	Gln	Pro	Lys	Met	Thr	Val	Glu	Lys	Leu	Leu

	370	375	380	
	Glu Tyr Gly His Met Leu Val Gln Glu Gln Asp Asn Val Lys Arg Val			
	385	390	395	400
	Gln Leu Ala Asp Thr Tyr Met Ser Gln Ala Ala Leu Gly Asp Ala Asn			
	405	410	415	
	Gln Asp Ala Met Lys Thr Gly Ser Phe Tyr Gly			
	420	425		
	<210> 40			
	<211> 1284			
	<212> DNA			
	<213> 普通小麦(Triticum aestivum)			
	<400> 40			
	atggctgctg cttctctctc caccgtcggt gccccggett ctacgccgac caacttcctc	60		
	gggaagaagc tcaagaagca ggtgacctcg gccgtgaact accatggcaa gagctccaag	120		
	gccaacaggt tcacggtcat ggcagcggaa aacatcgacg agaagaggaa caccgacaag	180		
	tggaagggtc ttgcgtacga tatctccgac gaccagcagg acatcaccag agggaagggc	240		
	atcgtggact cctcttcca ggcgcccact ggcgacggca cccacgagge cgtctcagc	300		
	tcatacagat acgtcagcca gggactcaag aagtacgact tcgacaacac catgggagge	360		
	ttctacatcg ctctgtcttt catggacaag cttgttgtcc atctctccaa gaacttcattg	420		
	accctgccca acatcaagat cccactcatc ttgggtatct ggggaggcaa gggtaagga	480		
	aaatccttcc agtgtgagct tgtcttcgcc aagatgggca tcaaccaat catgatgagt	540		
[0053]	gccggagagc tggagagtgg caacgccgga gagccagcca agctcatcag gcagcgtac	600		
	cgtgaggtcg cagacatgat caagaagggt aagatgtgct gcctcttcat caacgatctt	660		
	gacgccgtg cgggtaggat gggcgggacc acacagtaca ccgtcaacaa ccagatggtg	720		
	aacgccaccc tcatgaacat cgccgatgcc cccaccaatg tgcagctccc aggcattgac	780		
	aacaaggagg agaaccacg tgtgcccatc gtcgtcactg gtaacgattt ctcgacgttg	840		
	tacgccctc tgatccgtga tggtcgtatg gagaagttct actgggctcc caccgcgac	900		
	gaccgtatcg gtgtctgcaa gggatatctt cagaccgaca atgtcagcga cgagtccgtc	960		
	gtcaagatcg tcgacacctt cccaggacaa tccatcgact ttttcgggtc tctgcgtgct	1020		
	cgggtgtacg acgacgaggt gcgcaagtgg gtgacctcta ccggtatcga gaacattggc	1080		
	aagaagctgg tgaactcgcg ggacggacca gtgacctttg agcagccaaa gatgacagtg	1140		
	gagaagctgc tagagtacgg gcacatgctc gtccaggagc aggacaatgt caagcgtgtg	1200		
	cagcttgctg acacctacat gagccaggca gctctgggtg atgctaacca ggatgcgatg	1260		
	aagactggtg cttctacgg ttag	1284		
	<210> 41			
	<211> 427			
	<212> PRT			
	<213> 普通小麦(Triticum aestivum)			
	<400> 41			
	Met Ala Ala Ala Phe Ser Ser Thr Val Gly Ala Pro Ala Ser Thr Pro			
	1 5 10 15			
	Thr Asn Phe Leu Gly Lys Lys Leu Lys Lys Gln Val Thr Ser Ala Val			
	20 25 30			
	Asn Tyr His Gly Lys Ser Ser Lys Ala Asn Arg Phe Thr Val Met Ala			

	35	40	45
	Ala Glu Asn Ile Asp Glu Lys Arg Asn Thr Asp Lys Trp Lys Gly Leu		
	50	55	60
	Ala Tyr Asp Ile Ser Asp Asp Gln Gln Asp Ile Thr Arg Gly Lys Gly		
	65	70	75
	Ile Val Asp Ser Leu Phe Gln Ala Pro Thr Gly Asp Gly Thr His Glu		
	85	90	95
	Ala Val Leu Ser Ser Tyr Glu Tyr Val Ser Gln Gly Leu Lys Lys Tyr		
	100	105	110
	Asp Phe Asp Asn Thr Met Gly Gly Phe Tyr Ile Ala Pro Ala Phe Met		
	115	120	125
	Asp Lys Leu Val Val His Leu Ser Lys Asn Phe Met Thr Leu Pro Asn		
	130	135	140
	Ile Lys Ile Pro Leu Ile Leu Gly Ile Trp Gly Gly Lys Gly Gln Gly		
	145	150	155
	Lys Ser Phe Gln Cys Glu Leu Val Phe Ala Lys Met Gly Ile Asn Pro		
	165	170	175
	Ile Met Met Ser Ala Gly Glu Leu Glu Ser Gly Asn Ala Gly Glu Pro		
	180	185	190
	Ala Lys Leu Ile Arg Gln Arg Tyr Arg Glu Ala Ala Asp Met Ile Lys		
	195	200	205
	Lys Gly Lys Met Cys Cys Leu Phe Ile Asn Asp Leu Asp Ala Gly Ala		
[0054]	210	215	220
	Gly Arg Met Gly Gly Thr Thr Gln Tyr Thr Val Asn Asn Gln Met Val		
	225	230	235
	Asn Ala Thr Leu Met Asn Ile Ala Asp Ala Pro Thr Asn Val Gln Leu		
	245	250	255
	Pro Gly Met Tyr Asn Lys Glu Glu Asn Pro Arg Val Pro Ile Val Val		
	260	265	270
	Thr Gly Asn Asp Phe Ser Thr Leu Tyr Ala Pro Leu Ile Arg Asp Gly		
	275	280	285
	Arg Met Glu Lys Phe Tyr Trp Ala Pro Thr Arg Asp Asp Arg Ile Gly		
	290	295	300
	Val Cys Lys Gly Ile Phe Gln Thr Asp Asn Val Ser Asp Glu Ser Val		
	305	310	315
	Val Lys Ile Val Asp Thr Phe Pro Gly Gln Ser Ile Asp Phe Phe Gly		
	325	330	335
	Ala Leu Arg Ala Arg Val Tyr Asp Asp Glu Val Arg Lys Trp Val Thr		
	340	345	350
	Ser Thr Gly Ile Glu Asn Ile Gly Lys Lys Leu Val Asn Ser Arg Asp		
	355	360	365
	Gly Pro Val Thr Phe Glu Gln Pro Lys Met Thr Val Glu Lys Leu Leu		
	370	375	380
	Glu Tyr Gly His Met Leu Val Gln Glu Gln Asp Asn Val Lys Arg Val		
	385	390	395
			400

[0055]	Gln Leu Ala Asp Thr Tyr Met Ser Gln Ala Ala Leu Gly Asp Ala Asn	
	405	410 415
	Gln Asp Ala Met Lys Thr Gly Thr Phe Tyr Gly	
	420	425
	<210> 42	
	<211> 1395	
	<212> DNA	
	<213> 普通小麦(Triticum aestivum)	
	<400> 42	
	atggtctgtg ctttctcctc caccgtcggt gccccggett ctacgccgac caacttcctc	60
	gggaagaagc tcaagaagca ggtgacctcg gccgtgaact accatggcaa gagctccaag	120
	gccaacaggt tcacagtcac ggcagcggaa aacatcgacg agaagaggaa cacagacaag	180
	tggaagggtc ttgcctacga tatctccgac gaccagcagg acatcaccag agggaagggc	240
	atcgtggact ccttcttcca ggcgcccacg ggcgacggca cccacgaggc cgtcctcagc	300
	tcctacgagt acgtcagcca gggactcaag aagtacgact tcgacaacac catgggaggc	360
	ttctacateg ctctctgttt catggacaag cttgttgtcc atctctccaa gaacttcattg	420
	acctgcccc acatcaagat ccactcctc ttgggtatct ggggaggcaa ggtcaagga	480
	aaatccttcc agtgtgagct tgtcttcgcc aagatgggca tcaaccaat catgatgagt	540
	gccggagagc tggagagtgg caacgccgga gagccagcca agctcatcag gcagcggtag	600
	cgtgaggctg cagacatgat caagaagggt aagatgtgct gcctcttcat caacgatctt	660
	gacgccggtg cgggtcggat ggcggggacc acacagtaca ccgtcaacaa ccagatgggtg	720
	aagccaccc tcatgaacat cgcgatgcc cccaccaacg tgcagctccc aggcattgac	780
	aacaaggagg agaaccctcg tgtgcccacg gtcgtcactg gtaacgattt ctcaacgttg	840
	tacgccctc tcatccgtga tggctgtatg gagaagttct actgggctcc caccgcgac	900
	gaccgtatcg gtgtctgcaa gggatcttc cagaccgaca atgtcagcga cgagtcgctc	960
	gtcaagatcg tcgacacctt cccaggacaa tccatcgact ttttcgggtg tctgcgtgct	1020
	cgggtgtacg acgacaggtg gcgcaagtgg gtgacctcta ccggtatcga gaacattggc	1080
	aagaagcttg tgaactcgcg ggacggacca gtgacctttg agcagccaaa gatgacagtc	1140
	gagaagctgc tagagtacgg gcacatgtc gtccaggagc aggacaatgt caagcgtgtg	1200
	cagcttgctg acacctacat gagccaggca gctctgggtg atgctaacca ggatgcgatg	1260
	aagactggtt ctttctacgg taaaggggca cagaaggta ctttgctgt gccgcagga	1320
	tgcaccgacc agactgcca gaacttcgac ccaacggcga ggagtgcga cggcagctgc	1380
	ctttacacct tttaa	1395
	<210> 43	
	<211> 464	
	<212> PRT	
	<213> 普通小麦(Triticum aestivum)	
	<400> 43	
	Met Ala Ala Ala Phe Ser Ser Thr Val Gly Ala Pro Ala Ser Thr Pro	
	1 5 10 15	
	Thr Asn Phe Leu Gly Lys Lys Leu Lys Lys Gln Val Thr Ser Ala Val	
	20 25 30	
	Asn Tyr His Gly Lys Ser Ser Lys Ala Asn Arg Phe Thr Val Met Ala	
	35 40 45	

	Ala	Glu	Asn	Ile	Asp	Glu	Lys	Arg	Asn	Thr	Asp	Lys	Trp	Lys	Gly	Leu
	50						55					60				
	Ala	Tyr	Asp	Ile	Ser	Asp	Asp	Gln	Gln	Asp	Ile	Thr	Arg	Gly	Lys	Gly
	65					70					75				80	
	Ile	Val	Asp	Ser	Leu	Phe	Gln	Ala	Pro	Thr	Gly	Asp	Gly	Thr	His	Glu
					85					90					95	
	Ala	Val	Leu	Ser	Ser	Tyr	Glu	Tyr	Val	Ser	Gln	Gly	Leu	Lys	Lys	Tyr
					100					105					110	
	Asp	Phe	Asp	Asn	Thr	Met	Gly	Gly	Phe	Tyr	Ile	Ala	Pro	Ala	Phe	Met
					115					120					125	
	Asp	Lys	Leu	Val	Val	His	Leu	Ser	Lys	Asn	Phe	Met	Thr	Leu	Pro	Asn
					130					135					140	
	Ile	Lys	Ile	Pro	Leu	Ile	Leu	Gly	Ile	Trp	Gly	Gly	Lys	Gly	Gln	Gly
	145					150					155				160	
	Lys	Ser	Phe	Gln	Cys	Glu	Leu	Val	Phe	Ala	Lys	Met	Gly	Ile	Asn	Pro
					165						170				175	
	Ile	Met	Met	Ser	Ala	Gly	Glu	Leu	Glu	Ser	Gly	Asn	Ala	Gly	Glu	Pro
					180						185				190	
	Ala	Lys	Leu	Ile	Arg	Gln	Arg	Tyr	Arg	Glu	Ala	Ala	Asp	Met	Ile	Lys
					195						200				205	
	Lys	Gly	Lys	Met	Cys	Cys	Leu	Phe	Ile	Asn	Asp	Leu	Asp	Ala	Gly	Ala
					210						215				220	
[0056]	Gly	Arg	Met	Gly	Gly	Thr	Thr	Gln	Tyr	Thr	Val	Asn	Asn	Gln	Met	Val
	225					230					235				240	
	Asn	Ala	Thr	Leu	Met	Asn	Ile	Ala	Asp	Ala	Pro	Thr	Asn	Val	Gln	Leu
					245						250				255	
	Pro	Gly	Met	Tyr	Asn	Lys	Glu	Glu	Asn	Pro	Arg	Val	Pro	Ile	Val	Val
					260						265				270	
	Thr	Gly	Asn	Asp	Phe	Ser	Thr	Leu	Tyr	Ala	Pro	Leu	Ile	Arg	Asp	Gly
					275						280				285	
	Arg	Met	Glu	Lys	Phe	Tyr	Trp	Ala	Pro	Thr	Arg	Asp	Asp	Arg	Ile	Gly
					290						295				300	
	Val	Cys	Lys	Gly	Ile	Phe	Gln	Thr	Asp	Asn	Val	Ser	Asp	Glu	Ser	Val
	305					310					315				320	
	Val	Lys	Ile	Val	Asp	Thr	Phe	Pro	Gly	Gln	Ser	Ile	Asp	Phe	Phe	Gly
					325						330				335	
	Ala	Leu	Arg	Ala	Arg	Val	Tyr	Asp	Asp	Glu	Val	Arg	Lys	Trp	Val	Thr
					340						345				350	
	Ser	Thr	Gly	Ile	Glu	Asn	Ile	Gly	Lys	Lys	Leu	Val	Asn	Ser	Arg	Asp
					355						360				365	
	Gly	Pro	Val	Thr	Phe	Glu	Gln	Pro	Lys	Met	Thr	Val	Glu	Lys	Leu	Leu
					370						375				380	
	Glu	Tyr	Gly	His	Met	Leu	Val	Gln	Glu	Gln	Asp	Asn	Val	Lys	Arg	Val
					385						390				395	
	Gln	Leu	Ala	Asp	Thr	Tyr	Met	Ser	Gln	Ala	Ala	Leu	Gly	Asp	Ala	Asn

	405	410	415
	Gln Asp Ala Met Lys Thr Gly Ser Phe Tyr Gly Lys Gly Ala Gln Gln		
	420	425	430
	Gly Thr Leu Pro Val Pro Ala Gly Cys Thr Asp Gln Thr Ala Lys Asn		
	435	440	445
	Phe Asp Pro Thr Ala Arg Ser Asp Asp Gly Ser Cys Leu Tyr Thr Phe		
	450	455	460
	<210> 44		
	<211> 1395		
	<212> DNA		
	<213> 普通小麦(Triticum aestivum)		
	<400> 44		
	atggctgctg ctttctcctc caccgtcgggt gccccggctt ctacgccgac caacttcctc	60	
	gggaagaagc tcaagaagca ggtgacctcg gccgtgaact accatggcaa gagctccaag	120	
	gccaacaggt tcacgggtcat ggcagcggaa aacatcgacg agaagaggaa caccgacaag	180	
	tggaagggtc ttgcgtacga tatctccgac gaccagcagg acatcaccag agggaagggc	240	
	atcgtggact cctcttccca ggcgcccact ggcgacggca cccacgaggc cgtcctcagc	300	
	tcatacagat acgtcagcca gggactcaag aagtacgact tcgacaacac catgggaggc	360	
	ttctacatcg ctctgtcttt catggacaag ctgtgtgtcc atctctccaa gaacttcattg	420	
	accctgccc aacatcaagat cccactcctc ttgggtatct ggggaggcaa gggtaagga	480	
	aaatccttcc agtgtgagct tgtcttcgcc aagatgggca tcaaccaat catgatgagt	540	
	gccggagagc tggagagtgg caacgccgga gagccagcca agctcatcag gcagcggtag	600	
[0057]	cgtgaggctg cagacatgat caagaagggt aagatgtgct gcctcttcat caacgatctt	660	
	gacgccggtg cgggtaggat gggcgggacc acacagtaca ccgtcaacaa ccagatggtg	720	
	aacgccacc tcataaacat cgcgatgcc cccaccaatg tgcagctccc aggcattgac	780	
	aacaaggagg agaaccacg tgtgcccacg gtcgtcactg gtaacgattt ctcgacgttg	840	
	tacgccctc tgatccgtga tggctgtatg gagaagttct actgggctcc caccgcgac	900	
	gaccgtatcg gtgtctgcaa gggatatctc cagaccgaca atgtcagcga cgagtccgtc	960	
	gtcaagatcg tcgacacctt cccaggacaa tccatcgact ttttcgggtg tctgcgtgct	1020	
	cgggtgtacg acgacgaggt gcgcaagtgg gtgacctcta ccggtatcga gaacattggc	1080	
	aagaagctgg tgaactcgcg ggacggacca gtgacctttg agcagccaaa gatgacagtg	1140	
	gagaagctgc tagagtacgg gcacatgctc gtccaggagc aggacaatgt caagcgtgtg	1200	
	cagcttgctg acacctacat gagccaggca gctctgggtg atgctaacca ggatgcgatg	1260	
	aagactggta cttctacgg taaaggggca cagcaaggta ctttgctgt accggcagga	1320	
	tgcaccgacc agactgcca gaacttcgac ccaacggcga ggagtacga cggcagctgc	1380	
	ctttacacct tttaa	1395	
	<210> 45		
	<211> 464		
	<212> PRT		
	<213> 普通小麦(Triticum aestivum)		
	<400> 45		
	Met Ala Ala Ala Phe Ser Ser Thr Val Gly Ala Pro Ala Ser Thr Pro		
	1 5 10 15		
	Thr Asn Phe Leu Gly Lys Lys Leu Lys Lys Gln Val Thr Ser Ala Val		
	20 25 30		

	Asn	Tyr	His	Gly	Lys	Ser	Ser	Lys	Ala	Asn	Arg	Phe	Thr	Val	Met	Ala	
	Ala	Glu	Asn	Ile	Asp	Glu	Lys	Arg	Asn	Thr	Asp	Lys	Trp	Lys	Gly	Leu	
	Ala	Tyr	Asp	Ile	Ser	Asp	Asp	Gln	Gln	Asp	Ile	Thr	Arg	Gly	Lys	Gly	
	Ile	Val	Asp	Ser	Leu	Phe	Gln	Ala	Pro	Thr	Gly	Asp	Gly	Thr	His	Glu	
	Ala	Val	Leu	Ser	Ser	Tyr	Glu	Tyr	Val	Ser	Gln	Gly	Leu	Lys	Lys	Tyr	
	Asp	Phe	Asp	Asn	Thr	Met	Gly	Gly	Phe	Tyr	Ile	Ala	Pro	Ala	Phe	Met	
	Asp	Lys	Leu	Val	Val	His	Leu	Ser	Lys	Asn	Phe	Met	Thr	Leu	Pro	Asn	
	Ile	Lys	Ile	Pro	Leu	Ile	Leu	Gly	Ile	Trp	Gly	Gly	Lys	Gly	Gln	Gly	
	Lys	Ser	Phe	Gln	Cys	Glu	Leu	Val	Phe	Ala	Lys	Met	Gly	Ile	Asn	Pro	
	Ile	Met	Met	Ser	Ala	Gly	Glu	Leu	Glu	Ser	Gly	Asn	Ala	Gly	Glu	Pro	
	Ala	Lys	Leu	Ile	Arg	Gln	Arg	Tyr	Arg	Glu	Ala	Ala	Asp	Met	Ile	Lys	
[0058]	Lys	Gly	Lys	Met	Cys	Cys	Leu	Phe	Ile	Asn	Asp	Leu	Asp	Ala	Gly	Ala	
	Gly	Arg	Met	Gly	Gly	Thr	Thr	Gln	Tyr	Thr	Val	Asn	Asn	Gln	Met	Val	
	Asn	Ala	Thr	Leu	Met	Asn	Ile	Ala	Asp	Ala	Pro	Thr	Asn	Val	Gln	Leu	
	Pro	Gly	Met	Tyr	Asn	Lys	Glu	Glu	Asn	Pro	Arg	Val	Pro	Ile	Val	Val	
	Thr	Gly	Asn	Asp	Phe	Ser	Thr	Leu	Tyr	Ala	Pro	Leu	Ile	Arg	Asp	Gly	
	Arg	Met	Glu	Lys	Phe	Tyr	Trp	Ala	Pro	Thr	Arg	Asp	Asp	Arg	Ile	Gly	
	Val	Cys	Lys	Gly	Ile	Phe	Gln	Thr	Asp	Asn	Val	Ser	Asp	Glu	Ser	Val	
	Val	Lys	Ile	Val	Asp	Thr	Phe	Pro	Gly	Gln	Ser	Ile	Asp	Phe	Phe	Gly	
	Ala	Leu	Arg	Ala	Arg	Val	Tyr	Asp	Asp	Glu	Val	Arg	Lys	Trp	Val	Thr	
	Ser	Thr	Gly	Ile	Glu	Asn	Ile	Gly	Lys	Lys	Leu	Val	Asn	Ser	Arg	Asp	
	Gly	Pro	Val	Thr	Phe	Glu	Gln	Pro	Lys	Met	Thr	Val	Glu	Lys	Leu	Leu	
	Glu	Tyr	Gly	His	Met	Leu	Val	Gln	Glu	Gln	Asp	Asn	Val	Lys	Arg	Val	

	385	390	395	400
	Gln Leu Ala Asp Thr Tyr Met Ser Gln Ala Ala Leu Gly Asp Ala Asn			
		405	410	415
	Gln Asp Ala Met Lys Thr Gly Thr Phe Tyr Gly Lys Gly Ala Gln Gln			
		420	425	430
	Gly Thr Leu Pro Val Pro Ala Gly Cys Thr Asp Gln Thr Ala Lys Asn			
		435	440	445
	Phe Asp Pro Thr Ala Arg Ser Asp Asp Gly Ser Cys Leu Tyr Thr Phe			
	450	455	460	
	<210> 46			
	<211> 1299			
	<212> DNA			
	<213> 普通小麦(Triticum aestivum)			
	<400> 46			
	atggcttctg ctttctcgtc caccgttga gctccggcgt cgaccccgac caccttcctc			60
	gggaagaagg tgaagaagca ggccggtgcg ttgaactact accatggtgg caacaagatc			120
	aacaataggg tggtcagggc catggcggcc aaaaaggaac ttgacgaggg caagcagacc			180
	gatgccgatc ggtggaagg tctcgttac gacatctccg atgaccagca ggacatcacg			240
	agggggaaag gcatcgtgga ctccctgttc caggccccc tgggcgacgg caccacgag			300
	gccatcctga gctcctacga gtacatcagc cagggcctgc gcaagtacga cttcgacaac			360
	accatggacg ggctgtacat cgccccggcg ttcatggaca agctcatcgt ccacctcgcc			420
	aagaacttca tgacactccc caacatcaag gtccctctca tcctgggtat ctggggaggc			480
[0059]	aaggacagg gcaagtcgtt ccagtgcgag ctggtgttcg ccaagatggg catcaacccc			540
	atcatgatga gcgccggaga gctggagagc ggcaacgccg gcgagccggc caagctgac			600
	cggcagaggt accgcgagge tgccgacatt atcaagaagg gcaagatgtg ctgcctcttc			660
	atcaacgacc tggacgccgg cgccggggcg atggcgggga cgacgcagta cacggtgaac			720
	aaccagatgg tgaacgccac cctgatgaac atcgcggacg cgcccaccaa cgtgcagctc			780
	ccggggatgt acaacaagga ggagaaccca cgcgtgccca tcatcgtcac cggcaacgac			840
	ttctcgacgc tgtacgcgcc cctcatccgg gacggccgca tggagaagtt ctactgggcg			900
	cccaccggg aggaccgcat cggcgtgtgc aagggcattc tccgcaccga caacgtcccc			960
	gacgaggccg tggtaggct ggtggacacc tccccggggc agtccatcga cttcttcggc			1020
	gcgctgcggg cgcgggtgta cgacgacgag gtgcgcaagt gggtcggcga gatcggcgtc			1080
	gagaacatct ccaagcggct cgtcaactcc agggaggggc cgccgacgtt cgaccagccc			1140
	aagatgacca tcgagaagct catggagtac ggccacatgc tggccagga gcaggagaac			1200
	gtgaagcgcg tgcagctcgc cgacaagtac ctacgcgagg cggcgctcgg ccaagccaac			1260
	gacgacgcca tggcgaccgg cgccttctac ggcaagtag			1299
	<210> 47			
	<211> 432			
	<212> PRT			
	<213> 普通小麦(Triticum aestivum)			
	<400> 47			
	Met Ala Ser Ala Phe Ser Ser Thr Val Gly Ala Pro Ala Ser Thr Pro			
	1 5 10 15			
	Thr Thr Phe Leu Gly Lys Lys Val Lys Lys Gln Ala Gly Ala Leu Asn			
	20 25 30			

	Tyr	Tyr	His	Gly	Gly	Asn	Lys	Ile	Asn	Asn	Arg	Val	Val	Arg	Ala	Met
	35							40					45			
	Ala	Ala	Lys	Lys	Glu	Leu	Asp	Glu	Gly	Lys	Gln	Thr	Asp	Ala	Asp	Arg
	50						55					60				
	Trp	Lys	Gly	Leu	Ala	Tyr	Asp	Ile	Ser	Asp	Asp	Gln	Gln	Asp	Ile	Thr
	65					70					75				80	
	Arg	Gly	Lys	Gly	Ile	Val	Asp	Ser	Leu	Phe	Gln	Ala	Pro	Met	Gly	Asp
					85					90					95	
	Gly	Thr	His	Glu	Ala	Ile	Leu	Ser	Ser	Tyr	Glu	Tyr	Ile	Ser	Gln	Gly
					100					105					110	
	Leu	Arg	Lys	Tyr	Asp	Phe	Asp	Asn	Thr	Met	Asp	Gly	Leu	Tyr	Ile	Ala
	115						120					125				
	Pro	Ala	Phe	Met	Asp	Lys	Leu	Ile	Val	His	Leu	Ala	Lys	Asn	Phe	Met
	130						135					140				
	Thr	Leu	Pro	Asn	Ile	Lys	Val	Pro	Leu	Ile	Leu	Gly	Ile	Trp	Gly	Gly
	145					150					155				160	
	Lys	Gly	Gln	Gly	Lys	Ser	Phe	Gln	Cys	Glu	Leu	Val	Phe	Ala	Lys	Met
					165					170					175	
	Gly	Ile	Asn	Pro	Ile	Met	Met	Ser	Ala	Gly	Glu	Leu	Glu	Ser	Gly	Asn
					180					185					190	
	Ala	Gly	Glu	Pro	Ala	Lys	Leu	Ile	Arg	Gln	Arg	Tyr	Arg	Glu	Ala	Ala
	195						200					205				
[0060]	Asp	Ile	Ile	Lys	Lys	Gly	Lys	Met	Cys	Cys	Leu	Phe	Ile	Asn	Asp	Leu
	210						215					220				
	Asp	Ala	Gly	Ala	Gly	Arg	Met	Gly	Gly	Thr	Thr	Gln	Tyr	Thr	Val	Asn
	225					230					235				240	
	Asn	Gln	Met	Val	Asn	Ala	Thr	Leu	Met	Asn	Ile	Ala	Asp	Ala	Pro	Thr
					245					250					255	
	Asn	Val	Gln	Leu	Pro	Gly	Met	Tyr	Asn	Lys	Glu	Glu	Asn	Pro	Arg	Val
					260					265					270	
	Pro	Ile	Ile	Val	Thr	Gly	Asn	Asp	Phe	Ser	Thr	Leu	Tyr	Ala	Pro	Leu
	275						280					285				
	Ile	Arg	Asp	Gly	Arg	Met	Glu	Lys	Phe	Tyr	Trp	Ala	Pro	Thr	Arg	Glu
	290					295						300				
	Asp	Arg	Ile	Gly	Val	Cys	Lys	Gly	Ile	Phe	Arg	Thr	Asp	Asn	Val	Pro
	305					310				315					320	
	Asp	Glu	Ala	Val	Val	Arg	Leu	Val	Asp	Thr	Phe	Pro	Gly	Gln	Ser	Ile
					325					330					335	
	Asp	Phe	Phe	Gly	Ala	Leu	Arg	Ala	Arg	Val	Tyr	Asp	Asp	Glu	Val	Arg
					340					345					350	
	Lys	Trp	Val	Gly	Glu	Ile	Gly	Val	Glu	Asn	Ile	Ser	Lys	Arg	Leu	Val
					355					360					365	
	Asn	Ser	Arg	Glu	Gly	Pro	Pro	Thr	Phe	Asp	Gln	Pro	Lys	Met	Thr	Ile
	370					375						380				
	Glu	Lys	Leu	Met	Glu	Tyr	Gly	His	Met	Leu	Val	Gln	Glu	Gln	Glu	Asn

	385	390	395	400
	Val Lys Arg Val Gln Leu Ala Asp Lys Tyr Leu Ser Glu Ala Ala Leu			
		405	410	415
	Gly Gln Ala Asn Asp Asp Ala Met Ala Thr Gly Ala Phe Tyr Gly Lys			
		420	425	430
	<210> 48			
	<211> 1299			
	<212> DNA			
	<213> 普通小麦(Triticum aestivum)			
	<400> 48			
	atggetttctg ctttctcatt caccgttggg gctccggcgt cgaccccgac caccttcctc			60
	gggaagaagg tgaagaagca ggccggtacg ttgaactact accatggtgg caacaagatc			120
	aacagtaggg tggtcagggc catggcggcc aaaaaggaac ttgacgaggg caagcagacc			180
	gatgccgatc ggtggaaggg tctcgtttac gacatctccg acgaccagca ggacatcacc			240
	agggggaaag gtatcgttga ctccctgttc caggccccc tgggcgacgg caccacagag			300
	gccatcctga gtcctacga gtacatcagc cagggcctgc ggaagtacga cttcgacaac			360
	accatggacg ggctctacat cgccccggca ttcattggaca agctcatcgt ccacctcgcc			420
	aagaacttca tgacgtccc caacatcaag gtccctctca tcttgggcat ctggggaggc			480
	aaggacaggg gcaagtctgt ccagtgcgag ctggtgttcg ccaagatggg catcaacccc			540
	atcatgatga gcgccggaga gctggagagc ggcaacgccg gcgagccggc caagctgac			600
	cggcagaggt accgcgaggg ggccgacatt atcaagaagg gcaagatgtg ctgcctcttc			660
	atcaacgacc tggacgccgg cgccggggcg atgggcggga cgacgcagta cacggtgaac			720
[0061]	aaccagatgg tgaacgccac cttgatgaac atcgccgacg cgcccaccaa cgtgcagctc			780
	ccggggatgt acaacaagga ggagaaccca cgcggtccca tcatcgtcac cggcaacgac			840
	ttctcgacgc tgtacgcgcc cttcatccgc gacggccgca tggagaagtt ctactgggcg			900
	cccaccgcg aggaccgcat cggggtgtgc aagggcattt tccgcaccga caacgtccc			960
	gacgaggtct tggtaggct ggtggacacc tccccgggc aatccatcga cttcttcggc			1020
	gcgtgcggg cgccgggtgt cgacgacgag gtgcgcaagt gggtcggcga gatcggcgtc			1080
	gagaacatct ccaagcgcct cgtcaactcc agggaggggc cgccgacgtt cgaccagccc			1140
	aagatgacca tcgagaagct catggagtac ggccacatgc tggccagga gcaggagaac			1200
	gtgaagcgtg tgcagctcgc cgacaagtac ctacgaggg cgccgctcgg ccaagccaac			1260
	gacgacgcca tggcgaccgg cgcttctac ggcaagtag			1299
	<210> 49			
	<211> 432			
	<212> PRT			
	<213> 普通小麦(Triticum aestivum)			
	<400> 49			
	Met Ala Ser Ala Phe Ser Ser Thr Val Gly Ala Pro Ala Ser Thr Pro			
	1 5 10 15			
	Thr Thr Phe Leu Gly Lys Lys Val Lys Lys Gln Ala Gly Thr Leu Asn			
	20 25 30			
	Tyr Tyr His Gly Gly Asn Lys Ile Asn Ser Arg Val Val Arg Ala Met			
	35 40 45			
	Ala Ala Lys Lys Glu Leu Asp Glu Gly Lys Gln Thr Asp Ala Asp Arg			
	50 55 60			

	Trp	Lys	Gly	Leu	Ala	Tyr	Asp	Ile	Ser	Asp	Asp	Gln	Gln	Asp	Ile	Thr
	65					70				75					80	
	Arg	Gly	Lys	Gly	Ile	Val	Asp	Ser	Leu	Phe	Gln	Ala	Pro	Met	Gly	Asp
					85					90					95	
	Gly	Thr	His	Glu	Ala	Ile	Leu	Ser	Ser	Tyr	Glu	Tyr	Ile	Ser	Gln	Gly
					100					105					110	
	Leu	Arg	Lys	Tyr	Asp	Phe	Asp	Asn	Thr	Met	Asp	Gly	Leu	Tyr	Ile	Ala
					115					120					125	
	Pro	Ala	Phe	Met	Asp	Lys	Leu	Ile	Val	His	Leu	Ala	Lys	Asn	Phe	Met
					130					135					140	
	Thr	Leu	Pro	Asn	Ile	Lys	Val	Pro	Leu	Ile	Leu	Gly	Ile	Trp	Gly	Gly
	145					150						155				160
	Lys	Gly	Gln	Gly	Lys	Ser	Phe	Gln	Cys	Glu	Leu	Val	Phe	Ala	Lys	Met
					165					170						175
	Gly	Ile	Asn	Pro	Ile	Met	Met	Ser	Ala	Gly	Glu	Leu	Glu	Ser	Gly	Asn
					180					185						190
	Ala	Gly	Glu	Pro	Ala	Lys	Leu	Ile	Arg	Gln	Arg	Tyr	Arg	Glu	Ala	Ala
					195					200						205
	Asp	Ile	Ile	Lys	Lys	Gly	Lys	Met	Cys	Cys	Leu	Phe	Ile	Asn	Asp	Leu
					210					215						220
	Asp	Ala	Gly	Ala	Gly	Arg	Met	Gly	Gly	Thr	Thr	Gln	Tyr	Thr	Val	Asn
	225					230						235				240
[0062]	Asn	Gln	Met	Val	Asn	Ala	Thr	Leu	Met	Asn	Ile	Ala	Asp	Ala	Pro	Thr
					245					250						255
	Asn	Val	Gln	Leu	Pro	Gly	Met	Tyr	Asn	Lys	Glu	Glu	Asn	Pro	Arg	Val
					260					265						270
	Pro	Ile	Ile	Val	Thr	Gly	Asn	Asp	Phe	Ser	Thr	Leu	Tyr	Ala	Pro	Leu
					275					280						285
	Ile	Arg	Asp	Gly	Arg	Met	Glu	Lys	Phe	Tyr	Trp	Ala	Pro	Thr	Arg	Glu
					290					295						300
	Asp	Arg	Ile	Gly	Val	Cys	Lys	Gly	Ile	Phe	Arg	Thr	Asp	Asn	Val	Pro
	305					310						315				320
	Asp	Glu	Ala	Val	Val	Arg	Leu	Val	Asp	Thr	Phe	Pro	Gly	Gln	Ser	Ile
					325					330						335
	Asp	Phe	Phe	Gly	Ala	Leu	Arg	Ala	Arg	Val	Tyr	Asp	Asp	Glu	Val	Arg
					340					345						350
	Lys	Trp	Val	Gly	Glu	Ile	Gly	Val	Glu	Asn	Ile	Ser	Lys	Arg	Leu	Val
					355					360						365
	Asn	Ser	Arg	Glu	Gly	Pro	Pro	Thr	Phe	Asp	Gln	Pro	Lys	Met	Thr	Ile
					370					375						380
	Glu	Lys	Leu	Met	Glu	Tyr	Gly	His	Met	Leu	Val	Gln	Glu	Gln	Glu	Asn
	385					390						395				400
	Val	Lys	Arg	Val	Gln	Leu	Ala	Asp	Lys	Tyr	Leu	Ser	Glu	Ala	Ala	Leu
					405					410						415
	Gly	Gln	Ala	Asn	Asp	Asp	Ala	Met	Ala	Thr	Gly	Ala	Phe	Tyr	Gly	Lys

	420	425	430
	<210> 50		
	<211> 19		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 引物		
	<400> 50		
	gggtcggcga gatcgcgct		19
	<210> 51		
	<211> 23		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 引物		
	<400> 51		
	ccagcatgtg gccgtactcc atg		23
	<210> 52		
	<211> 23		
	<212> DNA		
	<213> 人工序列		
[0063]	<220>		
	<223> 引物		
	<400> 52		
	ccttctacgg taaagggca cag		23
	<210> 53		
	<211> 21		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 引物		
	<400> 53		
	tgtaaaggca gctcccgtcg t		21
	<210> 54		
	<211> 24		
	<212> DNA		
	<213> 人工序列		
	<220>		
	<223> 引物		
	<400> 54		
	ccatacacac ccaccatctc ttgc		24
	<210> 55		
	<211> 21		
	<212> DNA		

	<213> 人工序列	
	<220>	
	<223> 引物	
	<400> 55	
	tgtaaaggca gctcccgtcg t	21
	<210> 56	
	<211> 25	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 引物	
	<400> 56	
	caaatacgcc atcagggaga acatc	25
	<210> 57	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 引物	
	<400> 57	
	cgctgccgaa accacgagac	20
	<210> 58	
[0064]	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 引物	
	<400> 58	
	ttcttttctc acaaccaac gac	23
	<210> 59	
	<211> 21	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 引物	
	<400> 59	
	gcctcccgac attgcatct g	21
	<210> 60	
	<211> 5739	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> T-DNA	
	<400> 60	
	aattacaacg gtatatatcc tgccagtact cggccgtcga ccgcggtacc ccggaatitt	60

	gtggcgctct atcatagcta taaacctatt cagcacaata ttagttagct agggggcccc	120
	tcgagggcga tcgtacgta cctgcagggc ggccgcctgc agtgcagcgt gacccggtcg	180
	tgcccccttc tagagataat gagcattgca tgtctaaagt ataaaaaatt accacatatt	240
	ttttttgtca cacttgtttg aagtgcagtt tatctatctt tatacatata tttaaacttt	300
	actctacgaa taatataatc tatagtacta caataatata agtgttttag agaatcatat	360
	aatgaacag ttagacatgg tctaaaggac aattgagtat ttgacaaca ggactctaca	420
	gttttatctt tttagtgtgc atgtgttctc cttttttttt gcaaatagct tcacctatat	480
	aatacttcat ccattttatt agtacatcca tttagggttt aggggttaatg gtttttatag	540
	actaatTTTT ttagtacatc tattttatct tattttagcc tctaaattaa gaaaactaaa	600
	actctatttt agttttttta tttaataatt tagatataaa atagaataaa ataaagtac	660
	taaaaattaa acaaatatcc tttaagaaat taaaaaaact aaggaaacat tttcttgtt	720
	tcgagtagat aatgccagcc tgttaaagcgc cgtcgatcga cgagtctaac ggacaccaac	780
	cagcgaacca gcagcgtcgc gtcgggcca gcaagcaga cggcacggca tctctgtcgc	840
	tgctctgga cccctctcga gagttccgct ccaccgttgg acttgcctcg ctgtcgcat	900
	ccagaaattg cgtggcggag cggcagacgt gagccggcac ggagggcggc ctctctctcc	960
	tctcacggca cggcagcta cgggggattc ctttccacc gctccttgc tttcccttcc	1020
	tcgcccgcg taataaatag acacccctc cacacccctt tcccccaacc tcgtgttgtt	1080
	cggagcgac acacacaca ccagatctcc cccaaatcca cccgtcggca cctccgttc	1140
	aagtgacgc gctcgtctc ccccccccc cctctctacc tctcttagat cggcgttccg	1200
	gtccatgctt agggcccgtt agttctactt ctgtccatgt ttgtgttaga tccgtgtttg	1260
	tgtagatcc gtgtacttag cgttcgtaca cggatgcgac ctgtacgtca gacacgttct	1320
	gattgctaac ttgccagtgt ttctctttgg ggaatcctgg gatggctcta gccgttccgc	1380
[0065]	agacgggatc gatttcatga tttttttgt ttcgttgcac agggtttggg ttgccctttt	1440
	cttttatttc aatatatgcc gtgcacttgt ttgtcgggtc atcttttcat gctttttttt	1500
	gtcttggttg tgatgatgtg gtctggttgg gcggtcgttc tagatcggag tagaattctg	1560
	tttcaaacta cctggtggat ttattaattt tggatctgta tgtgtgtgcc atacatattc	1620
	atagttacga attgaagatg atggatggaa atatcgatct aggataggta tacatgttga	1680
	tcgggtttt actgatgcat atacagagat gctttttgtt cgcttggttg tgatgatgtg	1740
	gtgtggttgg gcggtcgttc attcgttcta gatcggagta gaatactgtt tcaaactacc	1800
	tggtgtattt attaatTTT gaactgtatg tgtgtgtcat acatcttcat agttacgagt	1860
	ttaagatgga tggaaatata gatctaggat aggtatacat gttgatgtgg gttttactga	1920
	tgcatataca tgatggcata tgcagcatct attcatatgc tctaaccttg agtacctatc	1980
	tattataata aacaagtatg ttttataatt attttgatct tgatatactt ggatgatggc	2040
	atatgcagca gctatatgtg gattttttta gccctgcctt catacgtat ttatttgctt	2100
	ggtactgttt cttttgtcga tgcaccctt gttgtttggg gttacttctg caggctcagc	2160
	tcgttgggtc gaagtcttgg gcagcttggg cgggtcatcc tgccggtaca ggcaaagtac	2220
	cttgcgtgac cctttaccg tagaaggtag cagcttctat cgcaccttgg ttagcatcac	2280
	ccagagctgc ctggtcatg taggtgtcag caagctgcac acgcttgaca ttgtctgtct	2340
	cctggacgag catgtgcccc tactctagca gcttctccac tgtcatcttt ggctgtcaa	2400
	aggtcactgg tccgtcccg gagttacca gcttcttgc aatgttctcg ataccgtag	2460
	aggtcaccca cttgcgcacc tcgtcgtcgt acaccgagc acgcagagca ccgaaaagt	2520
	cgatggattg tcttgggaag gtgtcgacaa tcttgacgac ggactcgtcg ctgacattgt	2580
	cgggtccgcc ttgtttctcc tctgtctctt gatctgacta atcttgggtt atgattcgtt	2640
	gagtaatttt ggggaaagct agcttcgtcc acagtttttt ttctgatgaa cagtgcgcga	2700
	gtggcgctga tcttgatgc tatcctgcaa tcgtggtgaa cttatttctt ttatacctt	2760

	cactcccatg aaaaggctag taatctttct cgatgtaaca tcgtccagca ctgctattac	2820
	cgtgtggtcc atccgacagt ctggctgaac acatcatacg atattgagca aagatctatc	2880
	ttccctgttc tttaatgaaa gacgtcattt tcatcagtat gatctaagaa tgttgcaact	2940
	tgcaaggagg cgtttctttc tttgaattta actaactcgt tgagtggccc tgtttctcgg	3000
	acgtaaggcc tttctgtctc cacacatgtc cattcgaatt ttaccgtgtt tagcaagggc	3060
	gaaaagtttg catcttgatg atttagcttg actatgcgat tgctttcctg gacccgtgca	3120
	gccgacaatg tcagcgacga gtccgtcgtc aagattgtcg acaccttccc aggacaatcc	3180
	atcgactttt tcggtgctct gcgtgctcgg gtgtacgacg acgagggtcg caagtgggtg	3240
	acctctaccg gtatcgagaa cattggcaag aagctgggtg actcgcggga cggaccagtg	3300
	acctttgagc agccaaagat gacagtggag aagctgctag agtacgggca catgctcgtc	3360
	caggagcagg acaatgtcaa gcgtgtgcag cttgtcgaca cctacatgag ccaggcagct	3420
	ctgggtgatg ctaaccagga tgcgatgaag actggtacct tctacggtaa aggggcacag	3480
	caaggtactt tgctgtgacc ggcaggatgc accgaccaga ctgccaagaa cttcgaccca	3540
	acgcgttggc cacgtgaaa tcaccagtct ctctctacaa atctatctct ctctattttc	3600
	tccataataa tgtgtgagta gttcccagat aagggaatta gggttcctat agggtttcgc	3660
	tcatgtgttg agcatataag aaacccttag tatgtatttg tatttgtaaa atacttctat	3720
	caataaaatt tctaattcct aaaacaaaaa tccagtacta aaatccagat cactagaatt	3780
	cgatatcatt accctgttat ccctaaagct tattaatata acttcgtata gcatacatta	3840
	tacgaagtta tgtttcttac gcagcaggtc tcatcaagac gatctaccg agtaacaatc	3900
	tccaggagat caaatacctt cccaagaagg ttaaagatgc agtcaaaaga ttcaggacta	3960
	attgcatcaa gaacacagag aaagacatat ttctcaagat cagaagtact attccagtat	4020
	ggacgattca aggcttgctt cataaaccaa ggcaagtaat agagattgga gtctctaaaa	4080
[0066]	aggtagttcc tactgaatct aaggccatgc atggagtcta agattcaaatt cgaggatcta	4140
	acagaactcg ccgtgaagac tggcgaacag ttcatacaga gtcttttacg actcaatgac	4200
	aagaagaaaa tcttcgtcaa catggtggag cacgacactc tgggtctactc caaaaatgtc	4260
	aaagatacag tctcagaaga ccaaagggtc attgagactt ttcaacaaag gataatttcg	4320
	ggaaacctcc tcggattcca ttgccagct atctgtcact tcatcgaaag gacagtagaa	4380
	aaggaaggtg gtcctacaa atgccatcat tgcgataaag gaaaggctat cattcaagat	4440
	gcctctgccg acagtgtgct caaagatgga cccccacca cgaggagcat cgtggaaaaa	4500
	gaagacgttc caaccacgtc ttcaaagcaa gtggattgat gtgacatctc cactgacgta	4560
	agggatgacg cacaatccca ctatccttcg caagaccctt cctctatata aggaagtcca	4620
	tttcatttgg agaggacacg ctgaaatcac cagtctctct ctataaatct atctctctct	4680
	ctataacaat ggaccagaa cgacgcccg cgcacatccg ccgtgccacc gaggcggaca	4740
	tgccggcggt ctgcaccatc gtcaaccact acatcgagac aagcacggtc aacttccgta	4800
	ccgagccgca ggaaccgcag gagtggacgg acgacctcgt ccgtctgcgg gagcgctatc	4860
	cctggctcgt cgccgaggtg gacggcgagg tcgccggcat cgcctacgcg ggccccctgga	4920
	aggcacgcaa cgcctacgac tggacggccg agtcgacctg gtacgtctcc ccccgccacc	4980
	agcggacggg actgggctcc acgtcttaca ccacctgct gaagtccctg gaggcacagg	5040
	gcttcaagag cgtggctcgt gtcacgggc tgcccaacga cccgagcgtg cgcattgcacg	5100
	aggcgctcgg atatgcccc cgcggcatgc tgcggcggcg cggttcaag cacgggaact	5160
	ggcatgacgt gggtttctgg cagctggact tcagcctgcc ggtaccgccc cgtccggtcc	5220
	tgcccgtcac cgagatctga gatcacccgt tctaggatcc gaagcagatc gttcaaacat	5280
	ttggcaataa agtttcttaa gattgaatcc tgttgccggt cttgcgatga ttatcatata	5340
	atttctgttg aattacgtta agcatgtaat aattaacatg taatgcatga cgttatttat	5400
	gagatgggtt tttatgatta gagtcccgca attatacatt taatacgcca tagaaaaaa	5460

	aatatagcgc gcaaactagg ataaattatc gcgcgcggtg tcattctatgt tactagatcg	5520
	aaacataact tcgtatagca tacattatac gaagttatca aaacgtcgtg agacagtttg	5580
	gttaactata acggtcctaa ggtagcgatt agttagctag cgagcggcga actaataact	5640
	ccgctctacc gaaagttact cgaggcatta cggcattacg gcactcgcga gggtcaccaat	5700
	tegagcatgg agccatttac aattgaatat atcctgccg	5739
	<210> 61	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 向导 RNA	
	<400> 61	
	acctcggccg tgaactacca tgg	23
	<210> 62	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 向导 RNA	
	<400> 62	
	tcctcttcc aggcgccac ggg	23
[0067]	<210> 63	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 向导 RNA	
	<400> 63	
	tgacgacgat gggcacacga ggg	23
	<210> 64	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 向导 RNA	
	<400> 64	
	tgacgtactc gtatgagctg agg	23
	<210> 65	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 向导 RNA	
	<400> 65	
	ctacatgagc caggcagctc tgg	23

	<210> 66	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 向导 RNA	
	<400> 66	
	aaccaggatg cgatgaagac tgg	23
	<210> 67	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 向导 RNA	
	<400> 67	
[0068]	cccaccatct cttgctgcat agg	23
	<210> 68	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 向导 RNA	
	<400> 68	
	accatctctt gctgcatagg agg	23
	<210> 69	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 向导 RNA	
	<400> 69	
	attcttctta cctgttcttc tgg	23

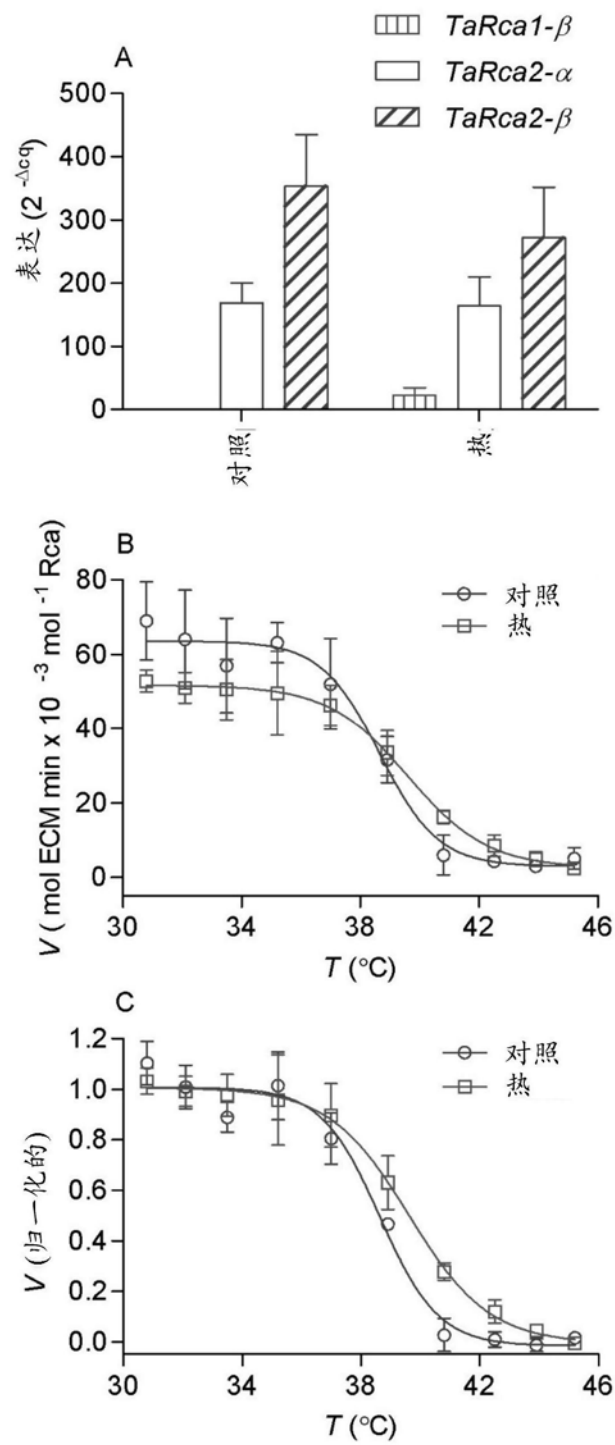


图1

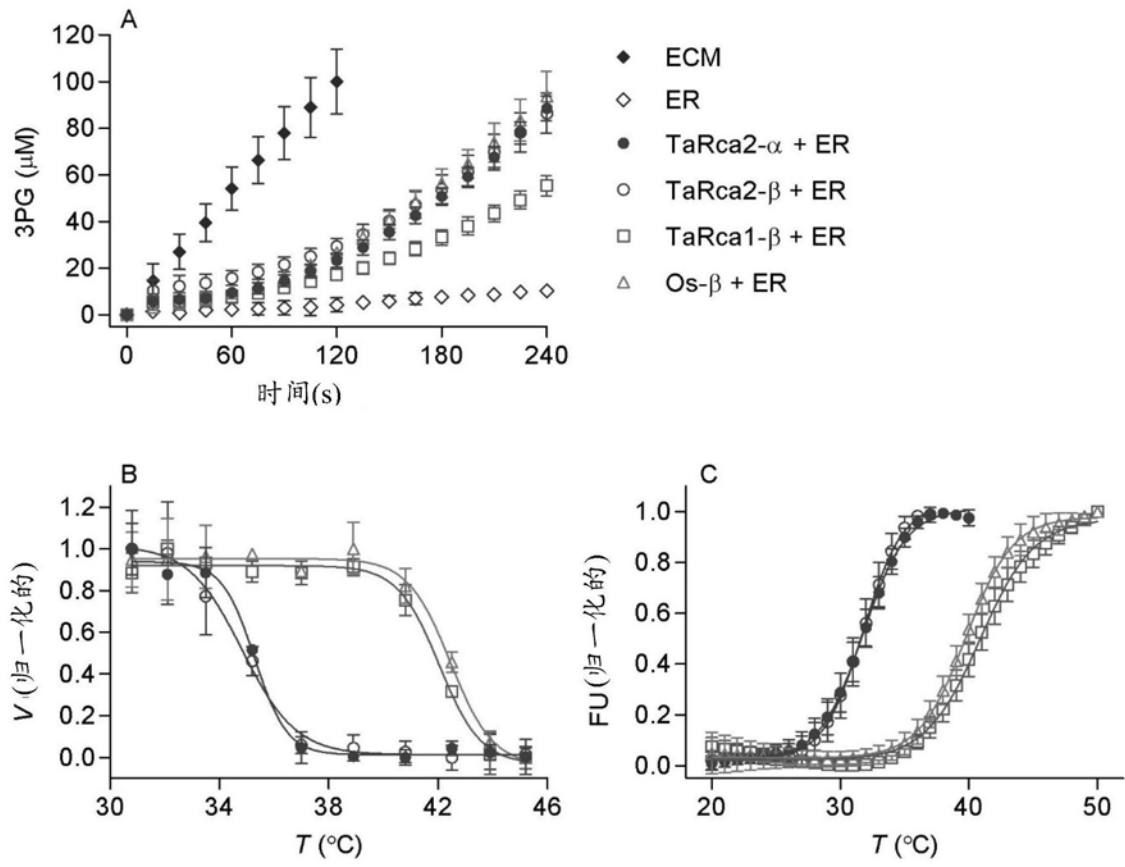


图2

```

TaRCA2b_CTP      1 maaafsstvgapastp-----tnflgkklkkqvt savnyhgkss-kanrftvmaeen-ld
TaRCA1b_CTP      1 maaafsstvgapastp-----tflgkklkkqvt savnyhgkss-kanrftvmaeen-ld
OsRCA_small      1 maaafsstvgapastp-----tnflgkklkkqvt savnyhgkss-ninrfkvm--keld
warm_consensus   1 maa--sstvgavnrtpslngsg-----vpstafgxxlkkxx-----nxxkvsixnfkvakexd
Cold_consensus   1 maaa---vvgaxnrtpslxgsxsgssvpxpflgkkltxscxf-----qssxxngsfkvxxevg
TaRCA2b_11AA_CTP 1 maaafsstvgapastp-----tnflgkklkkqvt savnyhgkss-kanrftvmaeen-ld
TaRCA2b_8AA_CTP 1 maaafsstvgapastp-----tnflgkklkkqvt savnyhgkss-kanrftvmaeen-ld

TaRCA2b_CTP      54 ekr--ntdkwkglaydisd-----dqdditr-gkivdslfqaptgdgtheavlsseyysqglkky
TaRCA1b_CTP      56 egkqtnadrdwkglaydisd-----dqdditr-gkivdslfqaptgdgtheavlsseyysqglrky
OsRCA_small      53 egkqtdqdrwkglaydisd-----dqdditr-gkivdslfqaptgdgtheavlsseyysqglrty
warm_consensus   55 edkqtdkdrwkglaydisd-----dqdditr-gkivdslfqaptgdgtheavlsseyysqglrty
Cold_consensus   62 exklsdgswealdadgsdylrlylgeadnmniavgxrgmvsifqgpf1-gtdsdiv--fdyrqkxtirsf
TaRCA2b_11AA_CTP 54 ekr--ntdkwkglaydisd-----dqdditr-gkivdslfqaptgdgtheavlsseyysqglkky
TaRCA2b_8AA_CTP 54 ekr--ntdkwkglaydisd-----dqdditr-gkivdslfqaptgdgtheavlsseyysqglkky

TaRCA2b_CTP      113 dfdntmggyfiapafmdklvvhlsknfnt-lpniklplilgiwggkqgkgsfqcelfvakmginpimmsa
TaRCA1b_CTP      117 dfdntmdglyi apafmdklvvhlsknfnt-lpnikvplilgiwggkqgkgsfqcelfvakmginpimmsa
OsRCA_small      114 dfdntmggyfiapafmdklvvhlsknfnt-lpnikvplilgiwggkqgkgsfqcelfvakmginpimmsa
warm_consensus   116 nldntmdgyfiapafmdklvvhlsknfnt-lpnikvplilgiwggkqgkgsfqcelfvakmginpimmsa
Cold_consensus   129 ey--lqgdyfiapafmdkvvhixknflnllnikvplilgiwggkqgkgsfqtellifaxmgiepvimsa
TaRCA2b_11AA_CTP 113 dfdntmdgyfiapafmdklvvhlsknfnt-lpniklplilgiwggkqgkgsfqcelfvakmginpimmsa
TaRCA2b_8AA_CTP 113 dfdntmggyfiapafmdklvvhlsknfnt-lpniklplilgiwggkqgkgsfqcelfvakmginpimmsa

TaRCA2b_CTP      182 gelesnagapaklirryreaadmi-kkgkmcclfindldagagrnggttqyvtvnnqmvnatlmniada
TaRCA1b_CTP      186 gelesnagapaklirryreaadii-nkgkmcclfindldagagrnggttqyvtvnnqmvnatlmniada
OsRCA_small      183 gelesnagapaklirryreaadii-kkgkmcclfindldagagrnggttqyvtvnnqmvnatlmniadn
warm_consensus   185 gelesnagapaklirryreaadii-kkgkmcclfindldagagrnggttqyvtvnnqmvnatlmniadn
Cold_consensus   197 gelesxragepgklirxyrxaqviqnqgkmsclmindldagagrfgxt-qxtvnnqvxxtlmnladn
TaRCA2b_11AA_CTP 182 gelesnagapaklirryreaadii-kkgkmcclfindldagagrnggttqyvtvnnqmvnatlmniada
TaRCA2b_8AA_CTP 182 gelesnagapaklirryreaadii-kkgkmcclfindldagagrnggttqyvtvnnqmvnatlmniada

```

图3

```

TaRCA2b_CTP      251 ptnvqlpgmynkeenprvpiivtgnfstlyaplirdgrmekfywaptrrdrigvckgifqtdnvsdesv
TaRCA1b_CTP      255 ptnvqlpgmynkeenprvpiivtgnfstlyaplirdgrmekfywaptredrigvckgifrtndnvpdeav
OsRCA_small      252 ptnvqlpgmynkeenprvpiivtgnfstlyaplirdgrmekfywaptreddrigvckgifrtndnvpdedi
warm_consensus   254 ptnvqlpgmynkeenprvpiivtgnfstlyaplirdgrmekfywaptredrigvckgifrtndnvpdedi
Cold_consensus   266 ptfrvsxgqmrxsdxxrvpiivtgnfstlyaplirdgrmekfywqptrediiinivhrmyekdgiisxxev
TaRCA2b_11AA_CTP 251 ptnvqlpgmynkeenprvpiivtgnfstlyaplirdgrmekfywaptreddrigvckgifrtndnvpdesv
TaRCA2b_8AA_CTP  251 ptnvqlpgmynkeenprvpiivtgnfstlyaplirdgrmekfywaptreddrigvckgifrtndnvpdesv
                ▲▲
TaRCA2b_CTP      321 vkivdtfpqgsidffgalravyddevrkwtstg-ienigkrlvnsr-dgpvtfeqpkmtveklleygh
TaRCA1b_CTP      325 vrlvdtfpqgsidffgalravyddevrkwtstg-ienigkrlvnsr-egpptfdqpkmtiekimeygh
OsRCA_small      322 vkivdsfpqgsidffgalravyddevrkwtstg-venigkrlvnsr-egppefeqpkmtiekimeygy
warm_consensus   324 vkivdtfpqgsidffgalravyddevrkwtstg-venigkrlvnsr-egppteqpkmtiekimeygn
Cold_consensus   336 vxivdtfpqaldfygalrgttdxxvllkwxvdiaggvenlkkllnskkekgyfeppeptleallesgx
TaRCA2b_11AA_CTP 321 vkivdtfpqgsidffgalravyddevrkwtstg-ienigkrlvnsr-dgpvtfeqpkmtiekimeygh
TaRCA2b_8AA_CTP  321 vkivdtfpqgsidffgalravyddevrkwtstg-ienigkrlvnsr-dgpvtfeqpkmtiekimeygh
                ▲▲
TaRCA2b_CTP      389 mlvqe qdnvkrvqladtymsgaalgdangdamktgsfyg-----
TaRCA1b_CTP      393 mlvqe qenvkrvqladkyliseaalganddamktgfaygk-----
OsRCA_small      390 mlvke qenvkrvqlaeqyliseaalganddamktgsfygsapss
warm_consensus   392 mlvxe qenvkrvqladkyliseaalganddamktgsfyg-----
Cold_consensus   406 slvke qelvxrvklakeynk-----nxdx-----
TaRCA2b_11AA_CTP 389 mlvqe qdnvkrvqladkyliseaalganddamktgsfyg-----
TaRCA2b_8AA_CTP  389 mlvqe qdnvkrvqladkyliseaalganddamktgsfyg-----
                ▲▲▲

```

图3 (续)

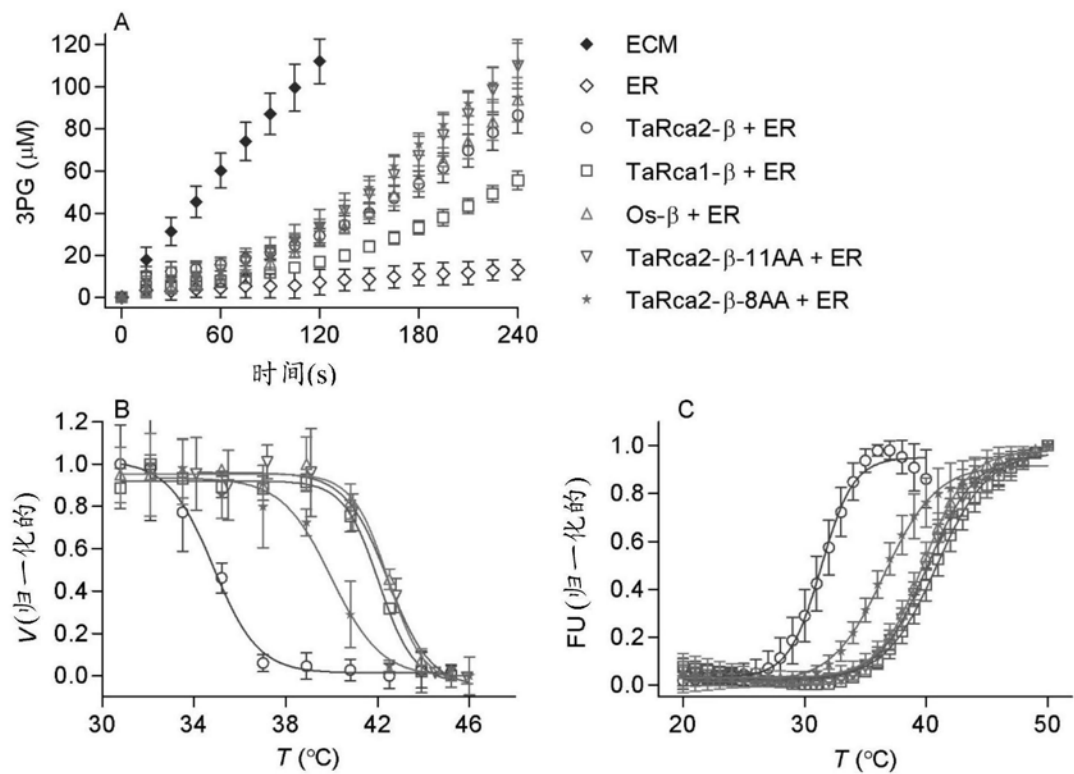


图4

