



(12)发明专利申请

(10)申请公布号 CN 108463552 A

(43)申请公布日 2018.08.28

(21)申请号 201680066362.8

(74)专利代理机构 北京市柳沈律师事务所
11105

(22)申请日 2016.11.15

代理人 张文辉 涂滔

(66)本国优先权数据

PCT/CN2015/094695 2015.11.16 CN

(51)Int.Cl.

G12N 9/46(2006.01)

(85)PCT国际申请进入国家阶段日

2018.05.14

D06M 16/00(2006.01)

G11D 3/386(2006.01)

(86)PCT国际申请的申请数据

PCT/CN2016/105927 2016.11.15

(87)PCT国际申请的公布数据

W02017/084560 EN 2017.05.26

(71)申请人 诺维信公司

地址 丹麦鲍斯韦

(72)发明人 赖伟坚 彭伟 刘晔 李明

权利要求书8页 说明书77页
序列表75页

(54)发明名称

纤维素酶变体和对其编码的多核苷酸

(57)摘要

提供了纤维素酶变体和编码这些变体的多核苷酸。并提供包含这些多核苷酸的核酸构建体、载体和宿主细胞。此外提供了使用这些变体的方法。

1. 一种纤维素酶变体,该纤维素酶变体在与SEQ ID NO:2的成熟多肽的位置292、274、266、265、255、246、237、224以及221相对应的一个或多个位置处包含改变,并且该变体具有纤维素酶活性。

2. 如权利要求1所述的变体,其中该纤维素酶是GH45纤维素酶。

3. 如权利要求1-2中任一项所述的变体,其中该改变是插入或缺失。

4. 如权利要求1-3中任一项所述的变体,其中该改变是取代。

5. 如权利要求1-4中任一项所述的变体,该变体是亲本纤维素酶的变体,包含催化结构域和纤维素酶结合结构域,其中该纤维素酶结合结构域与该催化结构域是异源的。

6. 如权利要求5所述的变体,其中接头在该催化结构域与该纤维素结合结构域之间。

7. 如权利要求6所述的变体,其中该接头与该纤维素结合结构域是同源的。

8. 如权利要求6所述的变体,其中该接头与该催化结构域是同源的。

9. 如权利要求1-8中任一项所述的变体,该变体是亲本纤维素酶的变体,该亲本纤维素酶选自下组,该组由以下各项组成:

a. 与SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:18的成熟多肽具有至少60%序列一致性的多肽;

b. 由在低严格条件下与以下项杂交的多核苷酸编码的多肽:(i) SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:17的成熟多肽编码序列,(ii) 其cDNA序列,或(iii) (i) 或(ii) 的全长互补体;

c. 由与SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:17的成熟多肽编码序列或其cDNA序列具有至少60%一致性的多核苷酸编码的多肽;以及

d. SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:18的成熟多肽的片段,该片段具有纤维素酶活性。

10. 如权利要求5所述的变体,其中该亲本纤维素酶与SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:18的成熟多肽具有至少60%,例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列一致性。

11. 如权利要求1-10中任一项所述的变体,其中该亲本纤维素酶由如下多核苷酸编码,该多核苷酸在低严格条件、中严格条件、中-高严格条件、高严格条件或非常高严格条件下与以下项杂交:(i) SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:17的成熟多肽编码序列,或者(ii) (i) 的全长互补体。

12. 如权利要求1-11中任一项所述的变体,其中该亲本纤维素酶由如下多核苷酸编码,该多核苷酸与SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:17的成熟多肽编码序列或其cDNA序列具有至少60%,例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列一致性。

13. 如权利要求1-12中任一项所述的变体,其中该亲本纤维素酶包含以下项或由以下项组成:SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:18的成熟多肽。

14. 如权利要求1-13中任一项所述的变体,其中该亲本纤维素酶是SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:18的成熟多肽的片段,其中该片段具有纤维素酶活性。

15. 如权利要求1-14中任一项所述的变体,该变体与该亲本纤维素酶的氨基酸序列具有至少60%,例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%一致性、至少96%、至少97%、至少98%、或至少99%但小于100%的序列一致性。

16. 如权利要求1-15中任一项所述的变体,其中该变体由200至315个,例如210至290、215至295、220至300个氨基酸组成。

17. 如权利要求1-16中任一项所述的变体,其中改变的数目是1-20个,例如1-10和1-5个,如1、2、3、4、5、6、7、8、9或10个改变。

18. 如权利要求1-17中任一项所述的变体,该变体在对应于位置292的位置处包含改变。

19. 如权利要求18所述的变体,其中该改变是被Tyr取代。

20. 如权利要求1-19中任一项所述的变体,该变体在对应于位置274的位置处包含改变。

21. 如权利要求20所述的变体,其中该改变是被Tyr取代。

22. 如权利要求1-19中任一项所述的变体,该变体在对应于位置266的位置处包含改变。

23. 如权利要求22所述的变体,其中该改变是被Tyr取代。

24. 如权利要求1-23中任一项所述的变体,该变体在对应于位置265的位置处包含改变。

25. 如权利要求24所述的变体,其中该改变是被Lys取代。

26. 如权利要求1-25中任一项所述的变体,该变体在对应于位置255的位置处包含改变。

27. 如权利要求26所述的变体,其中该改变是被Pro取代。

28. 如权利要求1-27中任一项所述的变体,该变体在对应于位置246的位置处包含改变。

29. 如权利要求28所述的变体,其中该改变是被Asn取代。

30. 如权利要求1-29中任一项所述的变体,该变体在对应于位置237的位置处包含改变。

31. 如权利要求30所述的变体,其中该改变是被Asn取代。

32. 如权利要求1-31中任一项所述的变体,该变体在对应于位置224的位置处包含改变。

33. 如权利要求32所述的变体,其中该改变是被Pro取代。

34. 如权利要求1-33中任一项所述的变体,该变体在对应于位置221的位置处包含改变。

35. 如权利要求34所述的变体,其中该改变是被Pro取代。

36. 如权利要求1-35中任一项所述的变体,该变体在对应于位置292、274、266、265、255、246、237、224和221中任何位置的两个位置处包含改变。

37. 如权利要求1-35中任一项所述的变体,该变体在对应于位置292、274、266、265、255、246、237、224和221中任何位置的三个位置处包含改变。

38. 如权利要求1-35中任一项所述的变体,该变体在对应于位置292、274、266、265、255、246、237、224和221的每一个位置处包含改变。

39. 如权利要求1-38中任一项所述的变体,该变体包含选自下组的一个或多个取代,该组由以下各项组成:W292Y、F274Y、W266Y、R265K、S255P、T246N、S224P、S224P和S221P。

40. 如权利要求1-39中任一项所述的变体,该变体包含W292Y和一个或多个选自下组的其他取代,该组由以下各项组成:S221P、S224P、S224P、T246N、S255P、R265K、W266Y和F274Y。

41. 如权利要求40所述的变体,该变体包含选自下组的一个或多个取代,该组由以下各项组成:

W292Y

W266Y+W292Y

R265K+W292Y

R265K+W266Y+W292Y

F274Y+W292Y

W266Y+F274Y+W292Y

S221P+S224P+W292Y

S224P+T246N+W292Y

S255P+W292Y

W266Y+W292Y。

42. 如权利要求1-42中任一项所述的变体,该变体在对应于位置202的位置处进一步包含改变。

43. 如权利要求1-42中任一项所述的变体,该变体包含以下项或由以下项组成:SEQ ID NO:6、SEQ ID NO:10、SEQ ID NO:16、SEQ ID NO:20、SEQ ID NO:22、SEQ ID NO:24、SEQ ID NO:26、SEQ ID NO:28、SEQ ID NO:30、SEQ ID NO:32、SEQ ID NO:34或SEQ ID NO:36或其成熟多肽。

44. 如权利要求1-43中任一项所述的变体,该变体相对于该亲本具有改进的特性,其中该改进的特性是改进的生物整理性能、降低的含纤维素纺织品的重量损失和/或降低的含纤维素纺织品强度损失。

45. 如权利要求44所述的变体,其中生物整理是生物抛光或牛仔布磨损。

46. 亲本GH45纤维素酶的变体,该变体包含催化结构域和纤维素结合结构域,其中该纤维素酶结合结构域与该催化结构域是异源的,并且其中该变体与该亲本GH45纤维素酶相比具有改进的生物整理活性。

47. 如权利要求46所述的变体,其中接头位于该催化结构域与该纤维素结合结构域之间。

48. 如权利要求47所述的变体,其中该接头与该纤维素结合结构域是同源的。

49. 如权利要求47所述的变体,其中该接头与该催化结构域是同源的。

50. 如权利要求46-49中任一项所述的变体,其中该催化结构域选自下组,该组由以下各项组成:

a. 与SEQ ID NO:12、SEQ ID NO:13或SEQ ID NO:14的催化结构域具有至少60%序列一致性的多肽；

b. 由在低严格条件下与以下项杂交的多核苷酸编码的多肽：(i) SEQ ID NO:37、SEQ ID NO:38或SEQ ID NO:39的催化结构域编码序列，(ii) 其cDNA序列，或(iii) (i) 或(ii) 的全长互补体；

c. 由与SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39的催化结构域编码序列或其cDNA序列具有至少60%一致性的多核苷酸编码的多肽；以及

d. SEQ ID NO:12、SEQ ID NO:13或SEQ ID NO:14的催化结构域的片段，该片段具有纤维素酶活性。

51. 如权利要求46-50中任一项所述的变体，其中该催化结构域与SEQ ID NO:12、SEQ ID NO:13或SEQ ID NO:14的催化结构域具有至少60%，例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列一致性。

52. 如权利要求46-51中任一项所述的变体，其中催化结构域由以下多核苷酸编码，该多核苷酸在低严格条件、中严格条件、中-高严格条件、高严格条件或非常高严格条件下与以下项杂交：(i) SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39的催化结构域编码序列，或(ii) (i) 的全长互补体。

53. 如权利要求46-52中任一项所述的变体，其中该催化结构域由如下多核苷酸编码，该多核苷酸与SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39的催化结构域编码序列或其cDNA序列具有至少60%，例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列一致性。

54. 如权利要求46-53中任一项所述的变体，其中该催化结构域包含以下项或由以下项组成：SEQ ID NO:12的氨基酸22-223、SEQ ID NO:13的氨基酸1-200或SEQ ID NO:14的氨基酸22-223。

55. 如权利要求46-54中任一项所述的变体，其中该纤维素结合结构域选自下组，该组由以下各项组成：

a. 与SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:8的纤维素结合结构域具有至少60%序列一致性的多肽；

b. 由在低严格条件下与以下项杂交的多核苷酸编码的多肽：(i) SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:7的纤维素结合结构域编码序列，(ii) 其cDNA序列，或(iii) (i) 或(ii) 的全长互补体；

c. 由与SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:7的纤维素结合结构域编码序列或其cDNA序列具有至少60%一致性的多核苷酸编码的多肽；以及

d. SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:8的纤维素结合结构域的片段，该片段具有纤维素结合活性。

56. 如权利要求46-55中任一项所述的变体，其中该纤维素结合结构域与SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:8的纤维素结合结构域具有至少60%，例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列一致性。

57. 如权利要求46-56中任一项所述的变体,其中该纤维素结合结构域由如下多核苷酸编码,该多核苷酸在低严格条件、中严格条件、中-高严格条件、高严格条件或非常高严格条件下与以下项杂交:(i) SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39、SEQ ID NO:7的纤维素结合结构域编码序列,或(ii) (i)的全长互补体。

58. 如权利要求46-57中任一项所述的变体,其中该纤维素结合结构域由如下多核苷酸编码,该多核苷酸与SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:7的纤维素结合结构域编码序列或其cDNA序列具有至少60%,例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列一致性。

59. 如权利要求46-58中任一项所述的变体,其中该纤维素结合结构域包含以下项或由以下项组成:SEQ ID NO:12的氨基酸262-299、SEQ ID NO:13的氨基酸258-295、SEQ ID NO:14的氨基酸257-293、或SEQ ID NO:8的氨基酸249-286。

60. 如权利要求47-59中任一项所述的变体,其中接头选自下组,该组由以下各项组成:

a. 与SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:8的接头具有至少60%序列一致性的多肽;

b. 由在低严格条件下与以下项杂交的多核苷酸编码的多肽:(i) SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:7的接头编码序列、(ii) 其cDNA序列、或(iii) (i) 或(ii)的全长互补体;

c. 由与SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:7的接头编码序列或其cDNA序列具有至少60%一致性的多核苷酸编码的多肽;以及

d. SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:8的接头的片段,该片段具有纤维素结合活性。

61. 如权利要求47-60中任一项所述的变体,其中该接头与SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:8的接头具有至少60%,例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列一致性。

62. 如权利要求47-61中任一项所述的变体,其中该接头由如下多核苷酸编码,该多核苷酸在低严格条件、中严格条件、中-高严格条件、高严格条件或非常高严格条件下与以下项杂交:(i) SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39、SEQ ID NO:7的接头编码序列,或(ii) (i)的全长互补体。

63. 如权利要求47-62中任一项所述的变体,其中该接头由如下多核苷酸编码,该多核苷酸与SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:7的接头编码序列或其cDNA序列具有至少60%,例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列一致性。

64. 如权利要求47-63中任一项所述的变体,其中该接头包含以下项或由以下项组成:SEQ ID NO:12的氨基酸224-261、SEQ ID NO:13的氨基酸201-257、SEQ ID NO:14的氨基酸224-256、或SEQ ID NO:8的氨基酸203-248。

65. 如权利要求46-64中任一项所述的变体,该变体与该亲本纤维素酶的氨基酸序列具有至少60%,例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%

一致性、至少96%、至少97%、至少98%、或至少99%但小于100%的序列一致性。

66. 如权利要求46-65中任一项所述的变体,其中该变体由200至315个,例如210至290、215至295、220至300个氨基酸组成。

67. 如权利要求46-66中任一项所述的变体,该变体在与SEQ ID NO:2的成熟多肽的位置292、274、266、265、255、246、237、224以及221相对应的一个或多个位置处包含改变。

68. 如权利要求46-67中任一项所述的变体,其中改变的数目是1-20个,例如1-10和1-5个,如1、2、3、4、5、6、7、8、9或10个改变。

69. 如权利要求46-68中任一项所述的变体,该变体在对应于位置292的位置处包含改变。

70. 如权利要求69所述的变体,其中该改变是被Tyr取代。

71. 如权利要求46-69中任一项所述的变体,该变体在对应于位置274的位置处包含改变。

72. 如权利要求71所述的变体,其中该改变是被Tyr取代。

73. 如权利要求46-72中任一项所述的变体,该变体在对应于位置266的位置处包含改变。

74. 如权利要求73所述的变体,其中该改变是被Tyr取代。

75. 如权利要求46-74中任一项所述的变体,该变体在对应于位置265的位置处包含改变。

76. 如权利要求75所述的变体,其中该改变是被Lys取代。

77. 如权利要求46-76中任一项所述的变体,该变体在对应于位置255的位置处包含改变。

78. 如权利要求77所述的变体,其中该改变是被Pro取代。

79. 如权利要求46-78中任一项所述的变体,该变体在对应于位置246的位置处包含改变。

80. 如权利要求79所述的变体,其中该改变是被Asn取代。

81. 如权利要求46-80中任一项所述的变体,该变体在对应于位置237的位置处包含改变。

82. 如权利要求81所述的变体,其中该改变是被Asn取代。

83. 如权利要求46-82中任一项所述的变体,该变体在对应于位置224的位置处包含改变。

84. 如权利要求83所述的变体,其中该改变是被Pro取代。

85. 如权利要求46-84中任一项所述的变体,该变体在对应于位置221的位置处包含改变。

86. 如权利要求85所述的变体,其中该改变是被Pro取代。

87. 如权利要求46-86中任一项所述的变体,该变体在对应于位置292、274、266、265、255、246、237、224和221中任何位置的两个位置处包含改变。

88. 如权利要求46-87中任一项所述的变体,该变体在对应于位置292、274、266、265、255、246、237、224和221中任何位置的三个位置处包含改变。

89. 如权利要求46-88中任一项所述的变体,该变体在对应于位置292、274、266、265、

255、246、237、224和221的每一个位置处包含改变。

90. 如权利要求46-89中任一项所述的变体,该变体包含选自下组的一个或多个取代,该组由以下各项组成:W292Y、F274Y、W266Y、R265K、S255P、T246N、S224P、S224P和S221P。

91. 如权利要求90所述的变体,该变体包含W292Y和一个或多个选自下组的其他取代,该组由以下各项组成:S221P、S224P、S224P、T246N、S255P、R265K、W266Y和F274Y。

92. 如权利要求91所述的变体,该变体包含选自下组的一个或多个取代,该组由以下各项组成:

W292Y

W266Y+W292Y

R265K+W292Y

R265K+W266Y+W292Y

F274Y+W292Y

W266Y+F274Y+W292Y

S221P+S224P+W292Y

S224P+T246N+W292Y

S255P+W292YW

266Y+W292Y。

93. 如权利要求46-91中任一项所述的变体,该变体在对应于位置202的位置处进一步包含改变。

94. 如权利要求46-93中任一项所述的变体,该变体包含以下项或由以下项组成:SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:18、SEQ ID NO:22、SEQ ID NO:24、SEQ ID NO:26、SEQ ID NO:28、SEQ ID NO:30、SEQ ID NO:32、SEQ ID NO:34或SEQ ID NO:36。

95. 如权利要求46-94中任一项所述的变体,该变体相对于该亲本GH45纤维素酶具有改进的生物整理特性、降低的含纤维素纺织品的重量损失和/或降低的含纤维素纺织品强度损失。

96. 如权利要求95所述的变体,其中生物整理是生物抛光或牛仔布磨损。

97. 一种编码如权利要求1-96中任一项所述的变体的多核苷酸。

98. 一种包含如权利要求97所述的多核苷酸的核酸构建体。

99. 一种包含如权利要求97所述的多核苷酸的表达载体。

100. 一种包含如权利要求97所述的多核苷酸的宿主细胞。

101. 一种产生纤维素酶变体的方法,该方法包括:

a. 在适于表达该变体的条件下培养如权利要求100所述的宿主细胞;并且

b. 回收该纤维素酶变体。

102. 一种用如权利要求97所述的多核苷酸转化的转基因植物、植物部分或植物细胞。

103. 一种产生如权利要求1-96中任一项所述的变体的方法,该方法包括:

a. 在有益于产生该变体的条件下培养包含编码该变体的多核苷酸的转基因植物或植物细胞;并且

b. 回收该变体。

104. 一种用于获得如权利要求1-45中任一项所述的变体的方法,该方法包括向亲本纤维素酶在对应于SEQ ID NO:2的成熟多肽的位置292、274、266、265、255、246、237、224和221的一个或多个位置处引入改变,其中每个改变独立地是取代、缺失或插入并且该变体具有纤维素酶活性;并且回收该变体。

105. 一种用于获得如权利要求46-96中任一项所述的变体的方法,该方法包括用异源纤维素结合结构域替代纤维素结合结构域;或将异源纤维素结合结构域添加至不包含纤维素结合结构域的亲本纤维素酶。

106. 一种用于生物整理的工艺,该工艺包括用如权利要求1-96中任一项所述的变体处理含纤维素纺织品。

107. 一种用于处理含纤维素纺织品的工艺,该工艺包括

a. 退浆;

b. 颜色修饰;

其中在步骤(a)和步骤(b)之前、期间或之后添加如权利要求1-96中任一项所述的变体。

108. 一种用于处理含纤维素纺织品的工艺,该工艺包括

a. 退浆;

b. 精练;

c. 漂白;

d. 染色;

其中在步骤(a)、(b)、(c)或(d)之前、期间或之后添加如权利要求1-96中任一项所述的变体。

109. 如权利要求106-109中任一项所述的工艺,其中该纤维素酶在低机械作用下被处理或添加。

110. 如权利要求1-96中任一项所述的变体用于含纤维素纺织品的生物整理的用途。

纤维素酶变体和对其编码的多核苷酸

[0001] 对序列表的引用

[0002] 本申请包含处于计算机可读形式的序列表,将其通过引用结合在此。

[0003] 发明背景

技术领域

[0004] 本发明涉及纤维素酶变体、编码这些变体的多核苷酸、产生这些变体的方法以及使用这些变体的方法。

背景技术

[0005] 纤维素酶广泛用于改进含纤维素织物的外观和柔软度。广泛应用纤维素酶来去除织物中或织物上的棉绒毛和松散的表面纤维。该工艺被称为“生物抛光”,并平滑了织物的表面,从而改进了织物的柔软度和外观。纤维素酶处理还有助于防止随后形成使服装呈现磨损的纤维起球。在去起球(depilling)过程中,期望由于纤维素酶的水解作用使该织物的强度损失最小化。

[0006] 纤维素酶的另一个工业应用是用于处理牛仔布织物,以便赋予它们“石洗”外观。此工艺在业内被称为“生物石磨”。因为习惯上使用浮石处理织物,所以采用术语“生物石磨”。然而,近年来,纤维素酶已经在很大程度上取代了浮石。生物石磨与去起球有很大的区别,因为生物石磨目的是从牛仔布中去除颜色并控制其在织物上的再沉积,而去起球目的仅仅是改进柔软度和外观。

[0007] 纤维素酶是一组糖苷水解酶,其催化纤维素聚合物中 β -1,4-糖苷键的水解,并且通常包含纤维素结合结构域(CBD)和催化结构域。这两个结构域之间的称为“接头”或“接头肽”的区域用作CBD和催化结构域之间的柔性间隔基。单个纤维素酶组分的催化结构域由酶学委员会(EC)和糖苷水解酶(GH)家族系统二者分类。酶学委员基于他们对内部 β -1,4键的切割(内切葡聚糖酶或“EG”,EC 3.2.1.4)或从纤维素聚合物的还原性或非还原性末端释放纤维二糖(纤维二糖水解酶或“CBH”,EC 3.2.1.91,有时也称为外切葡聚糖酶)的偏好来区分两类纤维素酶。相比之下,GH家族系统基于一级和二级结构的保守性以及催化反应的立体化学来区分纤维素酶组分的催化结构域。提供了所有已知的纤维素酶催化和结合结构域的GH家族名称并通过可在URL:cazy.org.获得的碳水化合物活性酶(CAZy)数据库(Cantarel等人,2009,Nucleic Acids Res[核酸研究]37:D233-238)进行持续更新。纤维素酶可见于许多GH家族,包括但不限于家族5、6、7、8、9、10、12、16、18、19、26、44、45、48、51、61和74。此外,可以将一些较大的GH家族中的纤维素酶分组为亚家族。

[0008] 工业上表现良好的内切- β -1,4-葡聚糖酶描述于例如WO 91/17243、WO 91/17244和WO 91/10732。具体的纤维素酶变体描述于WO 94/07998和WO 1998/012307。US 7981654披露了纤维素酶融合蛋白,其包含A.衍生自一个物种的纤维素酶核心的任选经修饰的第一氨基酸序列,和B.衍生自另一个物种的接头和/或纤维素结合结构域(CBD)的任选经修饰的第二氨基酸序列,其中在所述第一氨基酸序列和所述第二氨基酸序列之间已经引入了结合

区。

[0009] 尽管做了这些努力,但仍然需要改进纤维素酶及其组合物的组合,这些组合物在生物整理含纤维素纺织品中更有效。具体地,对更有效的纤维素酶组合物存在持续的需要以改进该工艺经济性。本发明目的是满足这些需要。

[0010] 本发明提供了与其亲本相比具有改进的特性的纤维素酶变体。

发明内容

[0011] 本发明涉及纤维素酶变体,这些变体在与SEQ ID NO:2的成熟多肽的位置292、274、266、265、255、246、237、224以及221相对应的一个或多个(例如,若干个)位置处包含改变,其中这些变体具有纤维素酶活性。

[0012] 本发明进一步涉及亲本GH45纤维素酶的变体,该变体包含催化结构域和纤维素结合结构域,其中该纤维素酶结合结构域与该催化结构域是异源的,并且其中该变体与亲本GH45纤维素酶相比具有改进的生物整理活性。

[0013] 本发明还涉及编码这些变体的多核苷酸;包含这些多核苷酸的核酸构建体、载体和宿主细胞;以及产生这些变体的方法。

[0014] 本发明还涉及处理含纤维素纺织品的方法。

[0015] 定义

[0016] 纤维素酶:术语“纤维素酶”是指催化纤维素聚合物中 β -1,4-糖苷键水解的一组糖苷水解酶。为了本发明的目的,根据实例中描述的程序确定纤维素酶活性。在一个方面中,本发明的变体具有SEQ ID NO:2的成熟多肽的至少20%,例如至少40%、至少50%、至少60%、至少70%、至少80%、至少90%、至少95%或至少100%的纤维素酶活性。

[0017] 等位基因变体:术语“等位基因变体”意指占用同一染色体位点的一种基因的两个或更多个替代形式中的任一者。等位基因变异通过突变天然产生,并且可以导致群体内的多态性。基因突变可以是沉默的(在所编码的多肽中没有改变)或可编码具有改变的氨基酸序列的多肽。多肽的等位基因变体是由基因的等位基因变体编码的多肽。

[0018] cDNA:术语“cDNA”意指可以通过从获得自真核或原核细胞的成熟的、剪接的mRNA分子进行反转录而制备的DNA分子。cDNA缺乏可存在于对应基因组DNA中的内含子序列。最初的初级RNA转录物是mRNA的前体,其在呈现为成熟的剪接的mRNA之前要通过一系列的步骤(包括剪接)进行加工。

[0019] 编码序列:术语“编码序列”意指直接指明变体的氨基酸序列的多核苷酸。编码序列的边界一般由开放阅读框决定,该开放阅读框以起始密码子(如ATG、GTG或TTG)开始并且以终止密码子(如TAA、TAG或TGA)结束。编码序列可为基因组DNA、cDNA、合成DNA或其组合。

[0020] 控制序列:术语“控制序列”意指对于表达编码本发明的变体的多核苷酸所必需的核酸序列。每个控制序列对于编码该变体的多核苷酸来说可以是天然的(即,来自相同基因)或外源的(即,来自不同基因),或相对于彼此是天然的或外源的。此类控制序列包括但不限于前导子、聚腺苷酸化序列、前肽序列、启动子、信号肽序列和转录终止子。最少,控制序列包括启动子、以及转录和翻译终止信号。出于引入促进控制序列与编码变体的多核苷酸的编码区域连接的特异性限制位点的目的,控制序列可以提供有接头。

[0021] 表达:术语“表达”包括涉及变体产生的任何步骤,包括但不限于,转录、转录后修

饰、翻译、翻译后修饰以及分泌。

[0022] 表达载体:术语“表达载体”意指线性或环状DNA分子,该分子包括编码变体的多核苷酸并且该多核苷酸可操作地与提供用于其表达的控制序列相连接。

[0023] 片段:术语“片段”意指从成熟多肽的氨基和/或羧基末端缺失一个或多个(例如,若干个)氨基酸的多肽;其中片段具有纤维素酶活性。在一个方面,一个片段包含至少200个氨基酸残基(例如,SEQ ID NO:2的氨基酸200至315,例如,210至290、215至295、220至300。

[0024] 高严格条件:术语“高严格条件”意指对于长度为至少100个核苷酸的探针而言,遵循标准DNA印迹程序,在42℃下在5X SSPE、0.3%SDS、200微克/ml剪切并变性的鲑鱼精子DNA和50%甲酰胺中预杂交和杂交12至24小时。最后在65℃下使用2X SSC、0.2%SDS将载体材料清洗三次,每次15分钟。

[0025] 宿主细胞:术语“宿主细胞”意指易于用包含本发明的多核苷酸的核酸构建体或表达载体转化、转染、转导等的任何细胞类型。术语“宿主细胞”涵盖由于复制过程中发生的突变而与亲本细胞不同的亲本细胞的任何后代。

[0026] 改进的特性:术语“改进的特性”意指与亲本相比改进的变体相关的特征。这种改进的特性包括但不限于改进的生物整理特性、降低的含纤维素纺织品的重量损失和降低的含纤维素纺织品强度损失。

[0027] 成熟多肽:术语“成熟多肽”意指在翻译和任何翻译后修饰如N-末端加工、C-末端截短、糖基化作用、磷酸化作用等之后处于其最终形式的多肽。在一个方面中,基于预测SEQ ID NO:2的氨基酸-21至-1是信号肽的SignalP程序(Nielsen等人,1997,Protein Engineering[蛋白质工程]10:1-6),成熟多肽是SEQ ID NO:2的氨基酸-21至-1。本领域已知,宿主细胞可产生由相同多核苷酸表达的两个或更多个不同成熟多肽(即,具有不同C-末端和/或N-末端氨基酸)的混合物。

[0028] 成熟多肽编码序列:术语“成熟多肽编码序列”意指编码具有纤维素酶活性的成熟多肽的多核苷酸。在一方面,基于预测SEQ ID NO:1的核苷酸1至63编码信号肽的SignalP(Nielsen等人,1997,同上)程序,成熟多肽编码序列是SEQ ID NO:1的核苷酸1至63或其cDNA序列。

[0029] 中严格条件:术语“中严格条件”意指对于长度为至少100个核苷酸的探针而言,遵循标准DNA印迹程序,在42℃下在5X SSPE、0.3%SDS、200微克/ml剪切并变性的鲑鱼精子DNA和35%甲酰胺中预杂交和杂交12至24小时。最后在55℃下使用2X SSC、0.2%SDS将载体材料清洗三次,每次15分钟。

[0030] 中-高严格条件:术语“中-高严格条件”意指对于长度为至少100个核苷酸的探针而言,遵循标准DNA印迹程序,在42℃下在5X SSPE、0.3%SDS、200微克/ml剪切并变性的鲑鱼精子DNA以及35%甲酰胺中预杂交和杂交12至24小时。最后在60℃下使用2X SSC、0.2%SDS将载体材料清洗三次,每次15分钟。

[0031] 突变体:术语“突变体”意指编码变体的多核苷酸。

[0032] 核酸构建体:术语“核酸构建体”意指单链-或双链的核酸分子,该核酸分子是从天然存在的基因中分离的,或以本来不存在于自然界中的方式被修饰成包含核酸的区段,或是合成的,该核酸分子包含一个或多个控制序列。

[0033] 可操作地连接:术语“可操作地连接”意指如下的构造,其中,控制序列相对于多核

苷酸的编码序列安置在适当位置,从而使得该控制序列指导该编码序列的表达。

[0034] 亲本或亲本纤维素酶:术语“亲本”或“亲本纤维素酶”意指具有纤维素酶活性的任何多肽,对其进行改变以产生本发明的酶变体。

[0035] 序列一致性:用参数“序列一致性”来描述两个氨基酸序列之间或两个核苷酸序列之间的相关性。

[0036] 出于本发明的目的,使用如在EMBOSS包(EMBOSS:欧洲分子生物学开放软件套件(The European Molecular Biology Open Software Suite),Rice等人,2000,Trends Genet.[遗传学趋势]16:276-277)(优选5.0.0版本或更新版本)的尼德尔(Needle)程序中所实施的尼德尔曼-翁施(Needleman-Wunsch)算法(Needleman和Wunsch,1970,J.Mol.Biol.[分子生物学杂志]48:443-453)来确定两个氨基酸序列之间的序列一致性。使用的参数是缺口开放罚分10、缺口扩展罚分0.5以及EBLOSUM62(BLOSUM62的EMBOSS版本)取代矩阵。将标记为“最长一致性”的Needle的输出(使用-nobrief选项获得)用作百分比一致性并且计算如下:

[0037] $(\text{一致的残基} \times 100) / (\text{比对长度} - \text{比对中的空位总数})$

[0038] 出于本发明的目的,使用如在EMBOSS包(EMBOSS:欧洲分子生物学开放软件套件,Rice等人,2000,同上)(优选5.0.0版或更新版本)的Needle程序中所实施的Needleman-Wunsch算法(Needleman和Wunsch,1970,同上)来确定两个脱氧核糖核苷酸序列之间的序列一致性。所使用的参数是空位开放罚分10,空位延伸罚分0.5和EDNAFULL(NCBI NUC4.4的EMBOSS版本)取代矩阵。将标记为“最长一致性”的Needle的输出(使用-nobrief选项获得)用作百分比一致性并且计算如下:

[0039] $(\text{相同脱氧核糖核苷酸} \times 100) / (\text{比对的长度} - \text{在比对中的空位总数})$ 。

[0040] 子序列:术语“子序列”意指使一个或多个(例如,若干个)核苷酸从成熟多肽编码序列的5'端和/或3'端缺失的多核苷酸;其中子序列编码具有纤维素酶活性的片段。在一个方面,子序列含有SEQ ID NO:1的核苷酸600至945(例如,SEQ ID NO:1的核苷酸645至885、核苷酸660至900)。

[0041] 变体:术语“变体”意指在一个或多个(例如,若干个)位置处包含改变(即,取代、插入和/或缺失)的具有纤维素酶活性的多肽。取代意指用不同的氨基酸替代占用某一位置的氨基酸;缺失意指去除占用某一位置的氨基酸;并且插入意指在邻接并且紧随占用某一位置的氨基酸之后添加氨基酸。术语“变体”还可以指“杂合多肽”,其中一种多肽的区域融合在另一种多肽的区域的N-末端或C-末端;或“融合多肽”,其中另一多肽融合在一种多肽的N-末端或C-末端。本发明的变体具有亲本纤维素酶的成熟多肽的至少20%,例如至少40%、至少50%、至少60%、至少70%、至少80%、至少90%、至少95%、或至少100%、至少110%、至少120%、至少150%、至少180%、至少200%的纤维素酶活性。

[0042] 非常高严格条件:术语“非常高严格条件”意指对于长度为至少100个核苷酸的探针而言,遵循标准DNA印迹程序,在42°C下在5X SSPE、0.3%SDS、200微克/ml剪切并变性的鲑精DNA和50%甲酰胺中预杂交和杂交12至24小时。最后在70°C下使用2X SSC、0.2%SDS将载体材料清洗三次,每次15分钟。

[0043] 非常低严格条件:术语“非常低严格条件”是指对于长度为至少100个核苷酸的探针而言,遵循标准DNA印迹程序,在42°C下在5X SSPE、0.3%SDS、200微克/ml剪切并变性的

鲑精DNA和25%甲酰胺中预杂交和杂交12至24小时。最后在45℃下使用2X SSC、0.2%SDS将载体材料洗涤三次,每次15分钟。

[0044] 野生型纤维素酶:术语“野生型”纤维素酶意指由天然存在的微生物(如在自然界中发现的细菌、酵母或丝状真菌)表达的纤维素酶。

[0045] 变体命名规则

[0046] 出于本发明的目的,将SEQ ID NO:2中披露的成熟多肽用以确定另一种纤维素酶中的相对应的氨基酸残基。将另一种纤维素酶的氨基酸序列与SEQ ID NO:2中披露的成熟多肽进行比对,并且基于该比对,使用如在EMBOSS包(EMBOSS:欧洲分子生物学开放软件套件,Rice等人,2000,Trends Genet.[遗传学趋势]16:276-277)(优选5.0.0版或更新版本)的尼德尔程序中所实施的尼德尔曼-翁施算法(Needleman和Wunsch,1970,J.Mol.Biol.[分子生物学杂志]48:443-453)来确定与SEQ ID NO:2中所披露的成熟多肽中的任何氨基酸残基相对应的氨基酸位置编号。使用的参数是缺口开放罚分10、缺口扩展罚分0.5以及EBLOSUM62(BLOSUM62的EMBOSS版本)取代矩阵。

[0047] 可以通过使用若干计算机程序,使用其对应默认参数比对多个多肽序列来确定在另一种纤维素酶中的对应氨基酸残基的鉴定,这些计算机程序包括但不限于MUSCLE(通过对数预期的多序列比较;版本3.5或更新版本;Edgar,2004,Nucleic Acids Research[核酸研究]32:1792-1797),MAFFT(版本6.857或更新版本;Kato和Kuma,2002,Nucleic Acids Research[核酸研究]30:3059-3066;Kato等人,2005,Nucleic Acids Research[核酸研究]33:511-518;Kato和Toh,2007,Bioinformatics[生物信息学]23:372-374;Kato等人,2009,Methods in Molecular Biology[分子生物学方法]537:39-64;Kato和Toh,2010,Bioinformatics[生物信息学]26:1899-1900),以及采用ClustalW(1.83或更新版本;Thompson等人,1994,Nucleic Acids Research[核酸研究]22:4673-4680)的EMBOSS EMMA。

[0048] 当其他酶与SEQ ID NO:2的成熟多肽相背离使得传统的基于序列的比较方法不能检测其相互关系时(Lindahl和Elofsson,2000,J.Mol.Biol.[分子生物学杂志]295:613-615),可以应用其他成对序列比较算法。在基于序列的搜索中较高的敏感度可使用搜索程序来获得,这些搜索程序采用多肽家族的概率表现(特征曲线)来搜索数据库。例如,PSI-BLAST程序通过迭代数据库搜索过程来产生多个谱图,并且能够检测远距离同源物(Atschul等人,1997,Nucleic Acids Res[核酸研究]25:3389-3402)。如果多肽的家族或超家族具有在蛋白结构数据库中的一个或多个代表,甚至可以实现更高的敏感度。程序如GenTHREADER(Jones,1999,J.Mol.Biol.[分子生物学杂志]287:797-815;McGuffin和Jones,2003,Bioinformatics[生物信息学]19:874-881)利用来自多种来源(PSI-BLAST、二级结构预测、结构比对谱和溶剂化势)的信息作为预测查询序列的结构折叠的神经网络的输入。类似地,Gough等人,2000,J.Mol.Biol.[分子生物学杂志]313:903-919的方法可以用于比对未知结构的序列与存在于SCOP数据库中的超家族模型。这些比对进而可以用于产生多肽的同源性模型,并且使用出于该目的而开发的多种工具可以评估此类模型的准确度。

[0049] 对于已知结构的蛋白质,若干工具和资源可用于检索并产生结构比对。例如,蛋白质的SCOP超家族已经在结构上进行比对,并且那些比对是可访问且可下载的。可以使用多种算法如距离比对矩阵(Holm和Sander,1998,Proteins[蛋白质]33:88-96)或组合延伸(Shindyalov和Bourne,1998,ProteinEngineering[蛋白质工程]11:739-747)比对两种或

更多种蛋白质结构,并且这些算法的实施可以另外用于查询具有感兴趣结构的结构数据库,以便发现可能的结构同源物(例如,Holm和Park,2000,Bioinformatics[生物信息学]16:566-567)。

[0050] 在描述本发明的变体中,以下所述的命名法适于方便参考。采用了已接受的IUPAC单字母和三字母的氨基酸缩写。

[0051] 取代。对于氨基酸取代,使用以下命名法:原始氨基酸、位置、取代氨基酸。因此,将在位置226处的苏氨酸被丙氨酸取代表示为“Thr226Ala”或“T226A”。多个突变由加号(“+”)分开,例如“Gly205Arg+Ser411Phe”或“G205R+S411F”表示在位置205和位置411处的甘氨酸(G)和丝氨酸(S)分别被精氨酸(R)和苯丙氨酸(F)取代。

[0052] 缺失。对于氨基酸缺失,使用以下命名法:初始氨基酸、位置、*。因此,将在位置195处的甘氨酸的缺失表示为“Gly195*”或“G195*”。多个缺失由加号(“+”)分开,例如“Gly195*+Ser411*”或“G195*+S411*”。

[0053] 插入。对于氨基酸插入,使用以下命名法:原始氨基酸、位置、原始氨基酸、插入氨基酸。因此,将在位置195处的甘氨酸之后插入赖氨酸表示为“Gly195GlyLys”或“G195GK”。将多个氨基酸的插入表示为[原始氨基酸、位置、原始氨基酸、插入氨基酸#1、插入氨基酸#2等]。例如,将在位置195处的甘氨酸之后插入赖氨酸和丙氨酸表示为“Gly195GlyLysAla”或“G195GKA”。

[0054] 在此类情况下,通过将小写字母添加至在所插入的氨基酸残基之前的氨基酸残基的位置编号中来对所插入的氨基酸残基进行编号。在以上实例中,该序列因此将是:

[0055]

亲本:	变体:
195	195 195a 195b
G	G-K-A

[0056] 多种改变。包含多种改变的变体由加号(“+”)分开,例如“Arg170Tyr+Gly195Glu”或“R170Y+G195E”表示在位置170和位置195处的精氨酸和甘氨酸分别被酪氨酸和谷氨酸取代。

[0057] 不同改变。在可以在一个位置处引入不同的改变的情况下,这些不同的改变由逗号分开,例如“Arg170Tyr,Glu”代表在位置170处的精氨酸被酪氨酸或谷氨酸取代。因此,“Tyr167Gly,Ala+Arg170Gly,Ala”表示以下变体:

[0058] “Tyr167Gly+Arg170Gly”、“Tyr167Gly+Arg170Ala”、“Tyr167Ala+Arg170Gly”和“Tyr167Ala+Arg170Ala”。

具体实施方式

[0059] 本发明涉及纤维素酶变体,这些变体在与SEQ ID NO:2的成熟多肽的位置292、274、266、265、255、246、237、224以及221相对应的一个或多个位置处包含改变,并且该变体具有纤维素酶活性。

[0060] 本发明涉及亲本GH45纤维素酶的变体,该变体包含催化结构域和纤维素结合结构域,其中该纤维素酶结合结构域与该催化结构域是异源的,并且其中该纤维素酶变体与亲本GH45纤维素酶相比具有改进的生物整理活性。

[0061] 变体

[0062] 本发明提供了纤维素酶变体,这些纤维素酶变体在与位置292、274、266、265、255、246、237、224和221相对应的一个或多个(例如,若干个)位置处包含改变,其中每个改变独立地是取代、插入或缺失并且变体具有纤维素酶活性。

[0063] 在一方面,该纤维素酶是GH45纤维素酶。

[0064] 在另一方面,该改变是取代。

[0065] 在另一方面,该变体是亲本纤维素酶的变体,其包含催化结构域和纤维素酶结合结构域,其中纤维素酶结合结构域与催化结构域是异源的。在一个优选的实施例中,接头位于催化结构域与纤维素结合结构域之间。在更优选的实施例中,接头与纤维素结合结构域是同源的。在更优选的实施例中,接头与催化结构域是同源的。

[0066] 在另一方面,该变体与该亲本纤维素酶的氨基酸序列具有至少60%、例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、或至少99%但小于100%的序列一致性。

[0067] 在一方面,在本发明的变体中的改变数目是1-20个,例如1-10和1-5个,例如1、2、3、4、5、6、7、8、9或10个改变。

[0068] 在另一方面,变体在对应于位置292、274、266、265、255、246、237、224和221的一个或多个(例如,若干个)位置处包含改变。在另一方面,变体在对应于位置292、274、266、265、255、246、237、224和221的两个位置处包含改变。在另一方面,变体在对应于位置292、274、266、265、255、246、237、224和221中的任何位置的三个位置处包含改变。在另一方面,变体在对应于位置292、274、266、265、255、246、237、224和221的每个位置处包含改变。

[0069] 在另一方面,该变体在与位置292相对应的位置处包含一个改变或由其组成。在另一方面,与位置292相对应的位置处的氨基酸被Ala、Arg、Asn、Asp、Cys、Gln、Glu、Gly、His、Ile、Leu、Lys、Met、Phe、Pro、Ser、Thr、Trp、Tyr、或Val取代,优选被Tyr取代。在另一方面,该变体包含SEQ ID NO:2的成熟多肽的取代W292Y或由其组成。

[0070] 在另一方面,该变体在与位置274相对应的位置处包含一个改变或由其组成。在另一方面,与位置274相对应的位置处的氨基酸被Ala、Arg、Asn、Asp、Cys、Gln、Glu、Gly、His、Ile、Leu、Lys、Met、Phe、Pro、Ser、Thr、Trp、Tyr、或Val取代,优选被Tyr取代。在另一方面,该变体包含SEQ ID NO:2的成熟多肽的取代F274Y或由其组成。

[0071] 在另一方面,该变体在与位置266相对应的位置处包含一个改变或由其组成。在另一方面,与位置266相对应的位置处的氨基酸被Ala、Arg、Asn、Asp、Cys、Gln、Glu、Gly、His、Ile、Leu、Lys、Met、Phe、Pro、Ser、Thr、Trp、Tyr、或Val取代,优选被Tyr取代。在另一方面,该变体包含SEQ ID NO:2的成熟多肽的取代W266Y或由其组成。

[0072] 在另一方面,该变体在与位置265相对应的位置处包含一个改变或由其组成。在另一方面,与位置265相对应的位置处的氨基酸被Ala、Arg、Asn、Asp、Cys、Gln、Glu、Gly、His、Ile、Leu、Lys、Met、Phe、Pro、Ser、Thr、Trp、Tyr、或Val取代,优选被Lys取代。在另一方面,该变体包含SEQ ID NO:2的成熟多肽的取代R265K或由其组成。

[0073] 在另一方面,该变体在与位置255相对应的位置处包含一个改变或由其组成。在另一方面,在与位置255相对应的位置处的氨基酸被Ala、Arg、Asn、Asp、Cys、Gln、Glu、Gly、

His、Ile、Leu、Lys、Met、Phe、Pro、Ser、Thr、Trp、Tyr、或Val取代,优选地被Pro取代。在另一方面,该变体包含SEQ ID NO:2的成熟多肽的取代S255P或由其组成。

[0074] 在另一方面,该变体在与位置246相对应的位置处包含一个改变或由其组成。在另一方面,对应于位置246的位置处的氨基酸被Ala、Arg、Asn、Asp、Cys、Gln、Glu、Gly、His、Ile、Leu、Lys、Met、Phe、Pro、Ser、Thr、Trp、Tyr、或Val取代,优选被Asn取代。在另一方面,该变体包含SEQ ID NO:2的成熟多肽的取代T246N或由其组成。

[0075] 在另一方面,该变体在与位置237相对应的位置处包含一个改变或由其组成。在另一方面,对应于位置237的位置处的氨基酸被Ala、Arg、Asn、Asp、Cys、Gln、Glu、Gly、His、Ile、Leu、Lys、Met、Phe、Pro、Ser、Thr、Trp、Tyr、或Val取代,优选被Asn取代。在另一方面,该变体包含SEQ ID NO:2的成熟多肽的取代T237N或由其组成。

[0076] 在另一方面,该变体在与位置224相对应的位置处包含一个改变或由其组成。在另一方面,在与位置224相对应的位置处的氨基酸被Ala、Arg、Asn、Asp、Cys、Gln、Glu、Gly、His、Ile、Leu、Lys、Met、Phe、Pro、Ser、Thr、Trp、Tyr、或Val取代,优选地被Pro取代。在另一方面,该变体包含SEQ ID NO:2的成熟多肽的取代S224P或由其组成。

[0077] 在另一方面,该变体在与位置221相对应的位置处包含一个改变或由其组成。在另一方面,在与位置221相对应的位置处的氨基酸被Ala、Arg、Asn、Asp、Cys、Gln、Glu、Gly、His、Ile、Leu、Lys、Met、Phe、Pro、Ser、Thr、Trp、Tyr、或Val取代,优选地被Pro取代。在另一方面,该变体包含SEQ ID NO:2的成熟多肽的取代S221P或由其组成。

[0078] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292和274、位置292和266、位置292和265、位置292和265、位置292和246、位置292和237、位置292和224、位置292和221的位置处的改变,如上文描述的那些。

[0079] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置274和266、位置274和265、位置274和255、位置274和246、位置274和237、位置274和224、位置274和221的位置处的改变,如上文描述的那些。

[0080] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置266和265、位置266和255、位置266和246、位置266和237、位置266和224、位置266和221的位置处的改变,如上文描述的那些。

[0081] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置255和246、位置255和237、位置255和224、位置255和221的位置处的改变,如上文描述的那些。

[0082] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置246和237、位置246和224、位置246和221的位置处的改变,如上文描述的那些。

[0083] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置237和224、位置237和221的位置处的改变,如上文描述的那些。

[0084] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置224和221的位置处的改变,如上文描述的那些。

[0085] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、274和266,位置292、274和265,位置292、274和255,位置292、274和246,位置292、274和237,位置292、274和224,位置292、274和221的位置处的改变,如上文描述的那些。

[0086] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、266和265,位

置292、266和255,位置292、266和246,位置292、266和237,位置292、266和224,位置292、266和221的位置处的改变,如上文描述的那些。

[0087] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、265和255,位置292、265和246,位置292、265和237,位置292、265和224,位置292、265和221的位置处的改变,如上文描述的那些。

[0088] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、255和246,位置292、255和237,位置292、255和224,位置292、255和221的位置处的改变,如上文描述的那些。

[0089] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、246和237,位置292、246和224,位置292、246和221的位置处的改变,如上文描述的那些。

[0090] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、237和224,位置292、237和221的位置处的改变,如上文描述的那些。

[0091] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、224和221的位置处的改变,如上文描述的那些。

[0092] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置274、266和265,位置274、266和255,位置274、266和246,位置274、266和237,位置274、266和224,位置274、266和221的位置处的改变,如上文描述的那些。

[0093] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置274、265和255,位置274、265和246,位置274、265和237,位置274、265和224,位置274、265和221的位置处的改变,如上文描述的那些。

[0094] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置274、255和246,位置274、255和237,位置274、255和224,位置274、255和221的位置处的改变,如上文描述的那些。

[0095] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置274、246和237,位置274、246和224,位置274、246和221的位置处的改变,如上文描述的那些。

[0096] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置274、237和224,位置274、237和221的位置处的改变,如上文描述的那些。

[0097] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置274、224和221的位置处的改变,如上文描述的那些。

[0098] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置266、265和255,位置266、265和246,位置266、265和237,位置266、265和224,位置266、265和221的位置处的改变,如上文描述的那些。

[0099] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置266、255和246,位置266、255和237,位置266、255和224,位置266、255和221的位置处的改变,如上文描述的那些。

[0100] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置266、246和237,位置266、246和224,位置266、246和221的位置处的改变,如上文描述的那些。

[0101] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置266、237和224,位置266、237和221的位置处的改变,如上文描述的那些。

[0102] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置266、224和221的位置处的改变,如上文描述的那些。

[0103] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置265、255和246,位置265、255和237,位置265、255和224,位置265、255和221的位置处的改变,如上文描述的那些。

[0104] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置265、246和237,位置265、246和224,位置265、246和221的位置处的改变,如上文描述的那些。

[0105] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置265、237和224,位置265、237和221的位置处的改变,如上文描述的那些。

[0106] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置265、224和221的位置处的改变,如上文描述的那些。

[0107] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置255、246和237,位置255、246和224,位置255、246和221的位置处的改变,如上文描述的那些。

[0108] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置255、237和224,位置255、237和221的位置处的改变,如上文描述的那些。

[0109] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置255、224和221的位置处的改变,如上文描述的那些。

[0110] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置246、237和224、位置246、237和221的位置处的改变,如上文描述的那些。

[0111] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置237、224和221的位置处的改变,如上文描述的那些。

[0112] 在另一方面,该变体包含以下项或由以下项组成:对应于位置292、274、266和265,位置292、274、266和255,位置292、274、266和246,位置292、274、266和237,位置292、274、266和224,位置292、274、266和221的位置处的改变,如上文描述的那些。

[0113] 在另一方面,该变体包含以下项或由以下项组成:对应于位置292、266、265和255,位置292、266、265和255,位置292、266、265和246,位置292、266、265和237,位置292、266、265和224,位置292、266、265和221的位置处的改变,如上文描述的那些。

[0114] 在另一方面,该变体包含以下项或由以下项组成:对应于位置292、265、255和246,位置292、265、255和237,位置292、265、255和224,位置292、265、255和221的位置处的改变,如上文描述的那些。

[0115] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、255、246和237,位置292、255、246和224,位置292、255、246和221的位置处的改变,如上文描述的那些。

[0116] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、246、237和224,位置292、246、237和221的位置处的改变,如上文描述的那些。

[0117] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、237、224和221的位置处的改变,如上文描述的那些。

[0118] 在另一方面,该变体包含以下项或由以下项组成:对应于位置274、266、265和255,位置274、266、265和246,位置274、266、265和237,位置274、266、265和224,位置274、266、265和221的位置处的改变,如上文描述的那些。

[0119] 在另一方面,该变体包含以下项或由以下项组成:对应于位置274、265、255和246,位置274、265、255和237,位置274、265、255和224,位置274、265、255和221的位置处的改变,如上文描述的那些。

[0120] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置274、255、246和237,位置274、255、246和224,位置274、255、246和221的位置处的改变,如上文描述的那些。

[0121] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置274、246、237和224,位置274、246、237和221的位置处的改变,如上文描述的那些。

[0122] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置274、237、224及221的位置处的改变,如上文描述的那些。

[0123] 在另一方面,该变体包含以下项或由以下项组成:对应于位置266、265、255和246,位置266、265、255和237,位置266、265、255和224,位置266、265、255和221的位置处的改变,如上文描述的那些。

[0124] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置266、255、246和237,位置266、255、246和224,位置266、255、246和221的位置处的改变,如上文描述的那些。

[0125] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置266、246、237和224,位置266、246、237和221的位置处的改变,如上文描述的那些。

[0126] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置266、237、224和221的位置处的改变,如上文描述的那些。

[0127] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置265、255、246和237,位置265、255、246和224,位置265、255、246和221的位置处的改变,如上文描述的那些。

[0128] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置265、246、237和224,位置265、246、237和221的位置处的改变,如上文描述的那些。

[0129] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置265、237、224和221的位置处的改变,如上文描述的那些。

[0130] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置255、246、237和224,位置255、246、237和221的位置处的改变,如上文描述的那些。

[0131] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置255、237、224和221的位置处的改变,如上文描述的那些。

[0132] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置246、237、224及221的位置处的改变,如上文描述的那些。

[0133] 在另一方面,该变体包含以下项或由以下项组成:对应于位置292、274、266、265和255,位置292、274、266、265和246,位置292、274、266、265和237,位置292、274、266、265和224,位置292、274、266、265和221的位置处的改变,如上文描述的那些。

[0134] 在另一方面,该变体包含以下项或由以下项组成:对应于位置292、266、265、255和246,位置292、266、265、255和237,位置292、266、265、255和224,位置292、266、265、255和221的位置处的改变,如上文描述的那些。

[0135] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、265、255、246和237,位置292、265、255、246和224,位置292、265、255、246和221的位置处的改变,如上文描述的那些。

[0136] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、255、246、237和224,位置292、255、246、237和221的位置处的改变,如上文描述的那些。

[0137] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、246、237、224和221的位置处的改变,如上文描述的那些。

[0138] 在另一方面,该变体包含以下项或由以下项组成:对应于位置274、266、265、255和246,位置274、266、265、255和237,位置274、266、265、255和224,位置274、266、265、255和221的位置处的改变,如上文描述的那些。

[0139] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置274、265、255、246和237,位置274、265、255、246和224,位置274、265、255、246和221的位置处的改变,如上文描述的那些。

[0140] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置274、255、246、237和224,位置274、255、246、237和221的位置处的改变,如上文描述的那些。

[0141] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置274、246、237、224和221的位置处的改变,如上文描述的那些。

[0142] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置266、265、255、246和237,位置266、265、255、246和224,位置266、265、255、246和221的位置处的改变,如上文描述的那些。

[0143] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置266、255、246、237和224,位置266、255、246、237和221的位置处的改变,如上文描述的那些。

[0144] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置266、246、237、224和221的位置处的改变,如上文描述的那些。

[0145] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置265、255、246、237和224,位置265、255、246、237和221的位置处的改变,如上文描述的那些。

[0146] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置265、246、237、224和221的位置处的改变,如上文描述的那些。

[0147] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置255、246、237、224和221的位置处的改变,如上文描述的那些。

[0148] 在另一方面,该变体包含以下项或由以下项组成:对应于位置292、274、266、265、255和246,位置292、274、266、265、255和237,位置292、274、266、265、255和224,位置292、274、266、265、255和221的位置处的改变,如上文描述的那些。

[0149] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、266、265、255、246和237,位置292、266、265、255、246和224,位置292、266、265、255、246和221的位置处的改变,如上文描述的那些。

[0150] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、265、255、246、237和224,位置292、265、255、246、237和221的位置处的改变,如上文描述的那些。

[0151] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、255、246、237、224和221的位置处的改变,如上文描述的那些。

[0152] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置274、266、265、255、246和237,位置274、266、265、255、246和224,位置274、266、265、255、246和221的位置

处的改变,如上文描述的那些。

[0153] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置274、265、255、246、237和224,位置274、265、255、246、237和221的位置处的改变,如上文描述的那些。

[0154] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置274、255、246、237、224和221的位置处的改变,如上文描述的那些。

[0155] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置266、265、255、246、237和224,位置266、265、255、246、237和221的位置处的改变,如上文描述的那些。

[0156] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置266、255、246、237、224和221的位置处的改变,如上文描述的那些。

[0157] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置265、255、246、237、224和221的位置处的改变,如上文描述的那些。

[0158] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、274、266、265、255、246和237,位置292、274、266、265、255、246和224,位置292、274、266、265、255、246和221的位置处的改变,如上文描述的那些。

[0159] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、266、265、255、246、237和224,位置292、266、265、255、246、237和221的位置处的改变,如上文描述的那些。

[0160] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、265、255、246、237、224和221的位置处的改变,如上文描述的那些。

[0161] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置274、266、265、255、246、237和224,位置274、266、265、255、246、237和221的位置处的改变,如上文描述的那些。

[0162] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置274、265、255、246、237、224和221的位置处的改变,如上文描述的那些。

[0163] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置266、265、255、246、237、224和221的位置处的改变,如上文描述的那些。

[0164] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、274、266、265、255、246、237和224,位置292、274、266、265、255、246、237和221的位置处的改变,如上文描述的那些。

[0165] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、266、265、255、246、237、224和221的位置处的改变,如上文描述的那些。

[0166] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置274、266、265、255、246、237、224和221的位置处的改变,如上文描述的那些。

[0167] 在另一方面,该变体包含以下项或由以下项组成:在对应于位置292、274、266、265、255、246、237、224和221的位置处的改变,如上文描述的那些。

[0168] 在另一方面,该变体包含以下项或由以下项组成:选自下组的一个或多个(例如,若干个)取代,该组由W292Y、F274Y、W266Y、R265K、S255P、T246N、T237N、S224P和S221P组成。

[0169] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少

90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+F274Y、W292Y+W266Y、W292Y+R265K、W292Y+S255P、W292Y+T246N、W292Y+T237N、W292Y+S224P、W292Y+S221P。

[0170] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+W266Y、F274Y+R265K、F274Y+S255P、F274Y+T246N、F274Y+T237N、F274Y+S224P、F274Y+S221P。

[0171] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W266Y+R265K、W266Y+S255P、W266Y+T246N、W266Y+T237N、W266Y+S224P、W266Y+S221P。

[0172] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代R265K+S255P、R265K+T246N、R265K+T237N、R265K+S224P、R265K+S221P。

[0173] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代S255P+T246N、S255P+T237N、S255P+S224P、S255P+S221P。

[0174] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代T246N+T237N、T246N+S224P、T246N+S221P。

[0175] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代T237N+S224P、T237N+S221P。

[0176] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代S224P+S221P。

[0177] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、

至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+F274Y+W266Y、W292Y+F274Y+R265K、W292Y+F274Y+S255P、W292Y+F274Y+T246N、W292Y+F274Y+T237N、W292Y+F274Y+S224P、W292Y+F274Y+S221P。

[0178] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+W266Y+R265K、W292Y+W266Y+S255P、W292Y+W266Y+T246N、W292Y+W266Y+T237N、W292Y+W266Y+S224P、W292Y+W266Y+S221P。

[0179] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+R265K+S255P、W292Y+R265K+T246N、W292Y+R265K+T237N、W292Y+R265K+S224P、W292Y+R265K+S221P。

[0180] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+S255P+T246N、W292Y+S255P+T237N、W292Y+S255P+S224P、W292Y+S255P+S221P。

[0181] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+T246N+T237N、W292Y+T246N+S224P、W292Y+T246N+S221P。

[0182] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+T237N+S224P、W292Y+T237N+S221P。

[0183] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+S224P+S221P。

[0184] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+W266Y+R265K、F274Y+W266Y+S255P、F274Y+W266Y+T246N、F274Y+W266Y+T237N、F274Y+W266Y+S224P、F274Y+W266Y+S221P。

[0185] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+R265K+S255P、F274Y+R265K+T246N、F274Y+R265K+T237N、F274Y+R265K+S224P、F274Y+R265K+S221P。

[0186] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+S255P+T246N、F274Y+S255P+T237N、F274Y+S255P+S224P、F274Y+S255P+S221P。

[0187] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+T246N+T237N、F274Y+T246N+S224P、F274Y+T246N+S221P。

[0188] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+T237N+S224P、F274Y+T237N+S221P。

[0189] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+S224P+S221P。

[0190] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W266Y+R265K+S255P、W266Y+R265K+T246N、W266Y+R265K+T237N、W266Y+R265K+S224P、W266Y+R265K+S221P。

[0191] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W266Y+S255P+T246N、W266Y+S255P+T237N、W266Y+S255P+S224P、W266Y+S255P+S221P。

[0192] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W266Y+T246N+T237N、W266Y+T246N+S224P、W266Y+T246N+S221P。

[0193] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W266Y+T237N+S224P、W266Y+T237N+S221P。

[0194] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W266Y+S224P+S221P。

[0195] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代R265K+S255P+T246N、R265K+S255P+T237N、R265K+S255P+S224P、R265K+S255P+S221P。

[0196] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代R265K+T246N+T237N、R265K+T246N+S224P、R265K+T246N+S221P。

[0197] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代R265K+T237N+S224P、R265K+T237N+S221P。

[0198] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代R265K+S224P+S221P。

[0199] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代S255P+T246N+T237N、S255P+T246N+S224P、S255P+T246N+S221P。

[0200] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代S255P+T237N+S224P、S255P+T237N+S221P。

[0201] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与

SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代S255P+S224P+S221P。

[0202] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代T246N+T237N+S224P、T246N+T237N+S221P。

[0203] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代T246N+S224P+S221P。

[0204] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代T237N+S224P+S221P。

[0205] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+F274Y+W266Y+R265K、W292Y+F274Y+W266Y+S255P、W292Y+F274Y+W266Y+T246N、W292Y+F274Y+W266Y+T237N、W292Y+F274Y+W266Y+S224P、W292Y+F274Y+W266Y+S221P。

[0206] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+W266Y+R265K+S255P、W292Y+W266Y+R265K+T246N、W292Y+W266Y+R265K+T237N、W292Y+W266Y+R265K+S224P、W292Y+W266Y+R265K+S221P。

[0207] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+R265K+S255P+T246N、W292Y+R265K+S255P+T237N、W292Y+R265K+S255P+S224P、W292Y+R265K+S255P+S221P。

[0208] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+S255P+T246N+T237N、W292Y+S255P+T246N+S224P、W292Y+S255P+T246N+S221P。

[0209] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与

SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+T246N+T237N+S224P、W292Y+T246N+T237N+S221P。

[0210] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+T237N+S224P+S221P。

[0211] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+W266Y+R265K+S255P、F274Y+W266Y+R265K+T246N、F274Y+W266Y+R265K+T237N、F274Y+W266Y+R265K+S224P、F274Y+W266Y+R265K+S221P。

[0212] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+R265K+S255P+T246N、F274Y+R265K+S255P+T237N、F274Y+R265K+S255P+S224P、F274Y+R265K+S255P+S221P。

[0213] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+S255P+T246N+T237N、F274Y+S255P+T246N+S224P、F274Y+S255P+T246N+S221P。

[0214] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+T246N+T237N+S224P、F274Y+T246N+T237N+S221P。

[0215] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+T237N+S224P+S221P。

[0216] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W266Y+R265K+S255P+T246N、W266Y+R265K+S255P+T237N、W266Y+R265K+S255P+S224P、W266Y+R265K+S255P+S221P。

[0217] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与

SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W266Y+S255P+T246N+T237N、W266Y+S255P+T246N+S224P、W266Y+S255P+T246N+S221P。

[0218] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W266Y+T246N+T237N+S224P、W266Y+T246N+T237N+S221P。

[0219] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W266Y+T237N+S224P+S221P。

[0220] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代R265K+S255P+T246N+T237N、R265K+S255P+T246N+S224P、R265K+S255P+T246N+S221P。

[0221] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代R265K+T246N+T237N+S224P、R265K+T246N+T237N+S221P。

[0222] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代R265K+T237N+S224P+S221P。

[0223] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代S255P+T246N+T237N+S224P、S255P+T246N+T237N+S221P。

[0224] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代S255P+T237N+S224P+S221P。

[0225] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、

至少99%一致性的具有纤维素酶活性的多肽的取代T246N+T237N+S224P+S221P。

[0226] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+F274Y+W266Y+R265K+S255P、W292Y+F274Y+W266Y+R265K+T246N、W292Y+F274Y+W266Y+R265K+T237N、W292Y+F274Y+W266Y+R265K+S224P、W292Y+F274Y+W266Y+R265K+S221P。

[0227] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+W266Y+R265K+S255P+T246N、W292Y+W266Y+R265K+S255P+T237N、W292Y+W266Y+R265K+S255P+S224P、W292Y+W266Y+R265K+S255P+S221P。

[0228] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+R265K+S255P+T246N+T237N、W292Y+R265K+S255P+T246N+S224P、W292Y+R265K+S255P+T246N+S221P。

[0229] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+S255P+T246N+T237N+S224P、W292Y+S255P+T246N+T237N+S221P。

[0230] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+T246N+T237N+S224P+S221P。

[0231] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+W266Y+R265K+S255P+T246N、F274Y+W266Y+R265K+S255P+T237N、F274Y+W266Y+R265K+S255P+S224P、F274Y+W266Y+R265K+S255P+S221P。

[0232] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+R265K+S255P+T246N+T237N、F274Y+R265K+S255P+T246N+S224P、F274Y+R265K+S255P+T246N+S221P。

[0233] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与

SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+S255P+T246N+T237N+S224P、F274Y+S255P+T246N+T237N+S221P。

[0234] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+T246N+T237N+S224P+S221P。

[0235] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W266Y+R265K+S255P+T246N+T237N、W266Y+R265K+S255P+T246N+S224P、W266Y+R265K+S255P+T246N+S221P。

[0236] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W266Y+S255P+T246N+T237N+S224P、W266Y+S255P+T246N+T237N+S221P。

[0237] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W266Y+T246N+T237N+S224P+S221P。

[0238] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代R265K+S255P+T246N+T237N+S224P、R265K+S255P+T246N+T237N+S221P。

[0239] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代R265K+T246N+T237N+S224P、R265K+T246N+T237N+S221P。

[0240] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代S255P+T246N+T237N+S224P+S221P。

[0241] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、

至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+F274Y+W266Y+R265K+S255P+T246N、W292Y+F274Y+W266Y+R265K+S255P+T237N、W292Y+F274Y+W266Y+R265K+S255P+S224P、W292Y+F274Y+W266Y+R265K+S255P+S221P。

[0242] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的W292Y+W266Y+R265K+S255P+T246N+T237N、W292Y+W266Y+R265K+S255P+T246N+S224P、W292Y+W266Y+R265K+S255P+T246N+S221P。

[0243] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+R265K+S255P+T246N+T237N+S224P、W292Y+R265K+S255P+T246N+T237N+S221P。

[0244] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+S255P+T246N+T237N+S224P+S221P。

[0245] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+W266Y+R265K+S255P+T246N+T237N、F274Y+W266Y+R265K+S255P+T246N+S224P、F274Y+W266Y+R265K+S255P+T246N+S221P。

[0246] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+R265K+S255P+T246N+T237N+S224P、F274Y+R265K+S255P+T246N+T237N+S221P。

[0247] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+S255P+T246N+T237N+S224P+S221P。

[0248] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W266Y+R265K+S255P+T246N+T237N+S224P、W266Y+R265K+S255P+T246N+T237N+S221P。

[0249] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W266Y+S255P+T246N+T237N+S224P+S221P。

[0250] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代R265K+S255P+T246N+T237N+S224P+S221P。

[0251] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+F274Y+W266Y+R265K+S255P+T246N+T237N、W292Y+F274Y+W266Y+R265K+S255P+T246N+S224P、W292Y+F274Y+W266Y+R265K+S255P+T246N+S221P。

[0252] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的W292Y+W266Y+R265K+S255P+T246N+T237N+S224P、W292Y+W266Y+R265K+S255P+T246N+T237N+S221P。

[0253] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+R265K+S255P+T246N+T237N+S224P+S221P。

[0254] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+W266Y+R265K+S255P+T246N+T237N+S224P、F274Y+W266Y+R265K+S255P+T246N+T237N+S221P。

[0255] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+R265K+S255P+T246N+T237N+S224P+S221P。

[0256] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、

至少99%一致性的具有纤维素酶活性的多肽的取代W266Y+R265K+S255P+T246N+T237N+S224P+S221P。

[0257] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+F274Y+W266Y+R265K+S255P+T246N+T237N+S224P、W292Y+F274Y+W266Y+R265K+S255P+T246N+T237N+S221P。

[0258] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的W292Y+W266Y+R265K+S255P+T246N+T237N+S224P+S221P。

[0259] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代F274Y+W266Y+R265K+S255P+T246N+T237N+S224P+S221P。

[0260] 在另一方面,该变体包含以下项或由以下项组成:SEQ ID NO:2的成熟多肽的或与SEQ ID NO:2的成熟多肽具有至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%一致性的具有纤维素酶活性的多肽的取代W292Y+F274Y+W266Y+R265K+S255P+T246N+T237N+S224P+S221P。

[0261] 在另一方面,该变体包含选自下组的取代或由选自下组的取代组成:SEQ ID NO:2的成熟多肽的

[0262] W292Y

[0263] W266Y+W292Y

[0264] R265K+W292Y

[0265] R265K+W266Y+W292Y

[0266] F274Y+W292Y

[0267] W266Y+F274Y+W292Y

[0268] S221P+S224P+W292Y

[0269] S224P+T246N+W292Y

[0270] S255P+W292Y

[0271] W266Y+W292Y。

[0272] 这些变体在一个或多个(例如,若干个)其他位置上可进一步包含一个或多个另外的改变。

[0273] 这些氨基酸变化可以具有微小性质,即,不会显著地影响蛋白质的折叠和/或活性的保守氨基酸取代或插入;典型地1-30个氨基酸的小缺失;小的氨基-末端或羧基末端延伸,例如氨基末端的甲硫氨酸残基;多至20-25个残基的小接头肽;或通过改变净电荷或另

外的功能促进纯化的小的延伸,如聚组氨酸段、抗原表位或结合结构域。

[0274] 保守取代的实例是在下组之内:碱性氨基酸(精氨酸、赖氨酸及组氨酸)、酸性氨基酸(谷氨酸和天冬氨酸)、极性氨基酸(谷氨酰胺和天冬酰胺)、疏水性氨基酸(亮氨酸、异亮氨酸及缬氨酸)、芳香族氨基酸(苯丙氨酸、色氨酸及酪氨酸)及小氨基酸(甘氨酸、丙氨酸、丝氨酸、苏氨酸及甲硫氨酸)。一般不会改变比活性的氨基酸取代是本领域已知的并且例如由H.Neurath和R.L.Hill,1979,于The Proteins[蛋白质],Academic Press[学术出版社],New York[纽约]中描述。常见取代为Ala/Ser、Val/Ile、Asp/Glu、Thr/Ser、Ala/Gly、Ala/Thr、Ser/Asn、Ala/Val、Ser/Gly、Tyr/Phe、Ala/Pro、Lys/Arg、Asp/Asn、Leu/Ile、Leu/Val、Ala/Glu和Asp/Gly。

[0275] 在一方面,该变体还包含在对应于位置202的位置处的改变。

[0276] 可替代地,这些氨基酸改变具有这样一种性质使得多肽的物理化学性质发生改变。

[0277] 例如,该变体可以在一个或多个以下位置处包含取代、插入和/或缺失:4、5、6、7、8、9、10、11、12、13、14、15、16、18、19、20、21、21a、42、44、45、47、48、49、49a、49b、74、82、95j、110、111、112、113、114、115、116、119、121、123、127、128、129、130、131、132、132a、133、145、146、147、148、149、150b、178和/或179(WO 1998/012307的纤维素酶编号,将该文件通过引用结合在此)。

[0278] 可以根据本领域中已知的程序,如定点诱变或丙氨酸扫描诱变(Cunningham和Wells,1989,Science[科学]244:1081-1085)来鉴定多肽中的必需氨基酸。在后一技术中,在分子中的每个残基处引入单一的丙氨酸突变,并对得到的突变体分子检验纤维素酶活性从而鉴定对分子活性至关重要的氨基酸残基。还参见Hilton等人,1996,J.Biol.Chem.[生物化学杂志]271:4699-4708。酶或其他生物学相互作用的活性位点还可以通过对结构的物理分析来确定,如通过这样的技术确定:如核磁共振、晶体学、电子衍射或光亲和标记,连同对推定的接触位点氨基酸进行突变。参见,例如de Vos等人,1992,Science[科学]255:306-312;Smith等人,1992,J.Mol.Biol.[分子生物学杂志]224:899-904;Wlodaver等人,1992,FEBS Lett.[欧洲生物化学学会联盟通讯]309:59-64。还可以从与相关多肽的比对来推断必需氨基酸的一致性。

[0279] 这些变体可以由200至315个,例如,210至290个、215至295个、220至300个氨基酸组成。

[0280] 在一方面,该变体相对于亲本具有改进的特性,其中改进的特性是改进的生物整理性能、降低的含纤维素纺织品的重量损失和/或降低的含纤维素纺织品强度损失。

[0281] 在另一方面,生物整理是生物抛光或牛仔布磨损。

[0282] 亲本纤维素酶

[0283] 在一方面,亲本是GH45纤维素酶。

[0284] GH家族45纤维素酶(以前称为家族K)以转化异头构型方式起作用,以产生作为产物的寡糖的 α -D异头体。已经阐明,在活性部位中,一种天冬氨酸氨基酸充当广义酸并且另一种充当广义碱。

[0285] 家族45酶的三维结构已被阐明(参见,例如,以下文献中的特异腐质霉的结构: Davies等人,1996,Acta Crystallographica Section D-Biological Crystallography[结

晶学学报D辑-生物晶体学]52:7-17第1部分)。这些酶含有附着第七条链的六条链的 β -桶状结构。该结构含有平行和反平行的 β 链。活性中心位于开放的底物结合槽中。

[0286] 如在此使用的,术语“GH45纤维素酶”、“家族45纤维素酶”或“Cel45”意指含有根据EC 3.2.1.4分类的糖苷水解酶家族45催化结构域的碳水化合物活性纤维素酶。该术语涵盖使用转化机制水解纤维素和纤维寡糖的碳水化合物活性酶,并且在催化性天冬氨酸氨基酸附近具有以下两个特征序列之一:(i) A/S/T-T-R/N/T-Y/F/T-X-D-X-X-X-X-X-C/A-A/G/S-W/C的第一保守特征序列和H/Q/D/N-F/L-D-I/L/F的第二保守特征序列;或(ii)具有H/Q/D/N-F/L-D-I/L/F的第二保守特征序列,但缺少所述第一保守序列。在一个实施例中,第二保守特征序列是H-F-D-I。

[0287] 家族45纤维素酶已被分为至少两个亚家族(称为“A”和“B”)(Igarashi等人, Applied and Environmental Microbiology [应用与环境微生物学], 2008, 74(18):5628-5634)。根据本发明的一个实施例,家族45纤维素酶是亚家族A成员。根据本发明的另一个实施例,家族45纤维素酶是亚家族B成员。这包括但不限于下表中列出的那些亚家族A和亚家族B酶。

[0288] 家族45纤维素酶亚家族B成员:

[0289]

生物	缩写名称	基因库登录号
里氏木霉	TrCel45A	CAA83846.1
绿色木霉 (<i>Trichoderma viride</i>)	TvEGV	AAQ21385.1
斜卧青霉	PdCel 45A	ACF33814.1
构巢曲霉	AnAN6786.2	EAA58604.1

[0290]

盘鲍 (<i>Hadiotis discus discus</i>)	HddEG1	ABO26608.1
福寿螺	AcEG27I	ABR92637.1
福寿螺	AcEG27II	ABR92638.1
贻贝	MeEG	CAC59695.1
黄孢原毛平革菌	PcCel45A	BAG68300.1

[0291] 家族45纤维素酶亚家族A成员:

[0292]

生物	缩写名称	基因库登录号和 WO 公开号
特异腐质霉	HiCel45A	AAE16508.1
灰腐质霉高温变种 (<i>Humicola grisea</i> var. <i>thermoidea</i>)	HgEgl3	BAA74956.1
黑腐质霉 (<i>Humicola nigrescens</i>)	HnCel45A	CAB42308.1
毡状地丝霉 (<i>Geomyces pannorum</i>)	Gp RF6293 Cel45A	在 WO 2010/076388 中的 SEQ ID NO: 13
毡状地丝霉 (<i>Geomyces pannorum</i>)	Gp RF6293 Cel45B	在 WO 2010/076388 中的 SEQ ID NO: 15
木贼镰孢菌相似种 (<i>Fusarium cf. equiseti</i>)	Fe RF6318 Cel45B	在 WO 2010/076388 中的 SEQ ID NO: 17
毡状地丝霉 (<i>Geomyces pannorum</i>)	Gp RF6546 Cel45A	在 WO 2010/076388 中的 SEQ ID NO: 19
毡状地丝霉 (<i>Geomyces pannorum</i>)	Gp RF6608 Cel45A	在 WO 2010/076388 中的 SEQ ID NO: 21
毡状地丝霉 (<i>Geomyces pannorum</i>)	Gp RF6608 Cel45B	在 WO 2010/076388 中的 SEQ ID NO: 23
大孢圆孢霉	ScSTCE1	BAG69187.1
大孢圆孢霉	ScSTCE1	在 WO 2005/054475 中的 SEQ ID NO: 3
粪生粪壳菌	Sfcel45	在 WO 2014/026630 中的 SEQ ID NO: 2
白毛黑果菌	MaCel45A	CAD56665.1

[0293]

(<i>Melanocarpus albomyces</i>)		
柄孢霉	PaCel45A	CAP61565.1
嗜热枝顶孢霉 (<i>Acremonium thermophilum</i>)	AtSEQ6	ACE10216.1
土生梭孢壳霉	TtCel45A	在 WO 2012/089024 中的 SEQ ID NO: 4
粉红聚端孢菌	TroCel45A	CAB42312.1
嗜热枝顶孢霉 (<i>Acremonium thermophilum</i>)	AtSEQ2	ABW41463.1
<i>Fusarium anguoides</i>	FaCel45A	CAB42310.1
粉红粘帚霉(<i>Clonostachys rosea</i> f. <i>catenulata</i>)	CrCel45A	CAB42311.1
粗糙脉孢菌	NcCEI45A	CAD70529.1
<i>Volutella colletotrichoides</i>	VcSEQ22	AA00854.1
玉米赤霉	GzCel45A	AAR02399.1
尖孢镰孢菌	FoCel45A	AAA65589.1
枝顶孢霉属物种	AsSEQ10	AA00848.1
枝顶孢霉属物种	AsSEQ8	AA00847.1
卢克诺文思金孢子菌 (<i>Chrysosporium lucknowense</i>)	ClCel45A	AAQ38150.1
异宗梭孢壳霉 (<i>Thielavia heterothallica</i>)	ThSEQ2	AA00844.1
卷枝毛霉	McMce1	BAD95808.1
栖北散白蚁	RshpCel45A	BAA98037.1
松材线虫	BxEng1	BAD34546.1
灰葡萄孢霉	BfCel45A	XP_JX11547700.1
嗜热枝顶孢霉 (<i>Acremonium</i>)	AtSEQ4	ABW41464.1

[0294]

<i>thermophilum</i>)		
短柄帚霉	SbEgl	Q7M4T4*
总状共头霉	SrCBHI	ABU49185.2
米根霉	RoRce1	BAC53956.1
柄毛皮伞 (<i>Crinipellis scabella</i>)	CsSEQ16	AAY00851.1
菜豆壳球孢菌	MpSEQ14	AAY00850.1
柄孢霉	PaCel45B	CAP69443.1
米根霉	RoRce3	BAC53988.1
松材线虫	BxEng2	BAD34544.1
松材线虫	BxEng3	BAD34548.1
灰腐质霉高温变种	HgEgl4	BAA74957.1
闪光须霉	PnPcel	BAD77808.1
米根霉	RoRce2	BAC53987.1
达尔文澳白蚁后肠共生物种	MdhsFm4	CAD39200.1
稻瘟病菌	MgCel45A	XP_363402.1
达尔文澳白蚁后肠共生物种	MdhsFm3	CAD39199.1
达尔文澳白蚁后肠共生物种	MdhsFm1	CAD39197.1
达尔文澳白蚁后肠共生物种	MdhsFm2	CAD39198.1
四孢脉孢菌	Ntcel45	在 WO 2015/058700 中的 SEQ ID NO: 2
毕赤酵母 GS115	PpCel45A	CAY71902.1
马梨囊鞭菌 (<i>Piromyces equi</i>)	PeCel45A	CAB92325.1
桑天牛	AgCelI	AAN78326.1
桑天牛	AgCelIII	AAR22385.1
互隔交链孢霉	AaK1	AAF05700.1

[0295]

辣根猿叶甲 (<i>Phaedon cochleariae</i>)	PcEg	CAA76931.1
埃默森篮状菌	TeCel45A	CAJ75963.1
玉蜀黍黑粉菌	UmEgII	AAB36147.1

[0296] *Uniprot条目

[0297] 在一方面,亲本纤维素酶包含催化结构域和纤维素酶结合结构域,其中纤维素酶结合结构域与催化结构域是异源或同源的。

[0298] 在另一方面,接头位于催化结构域与纤维素结合结构域之间。

[0299] 在另一方面,接头与纤维素结合结构域是异源或同源的。

[0300] 在另一方面,接头与催化结构域是异源或同源的。

[0301] 在另一个实施例中,亲本纤维素酶选自下组,该组由以下各项组成:

[0302] a.与SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:18的成熟多肽具有至少60%序列一致性的多肽;

[0303] 由在低严格条件下与以下项杂交的多核苷酸编码的多肽:(i) SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:17的成熟多肽编码序列,(ii) 其cDNA序列,或(iii) (i) 或(ii) 的全长互补体;

[0304] 由与SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:17的成熟多肽编码序列或其cDNA序列具有至少60%一致性的多核苷酸编码的多肽;以及

[0305] SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:18的成熟多肽的片段,该片段具有纤维素酶活性。

[0306] 在另一方面,该亲本纤维素酶与SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:18的成熟多肽具有至少60%,例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列一致性。

[0307] 在另一实施例中,该亲本纤维素酶由如下多核苷酸编码,该多核苷酸在低严格条件、中严格条件、中-高严格条件、高严格条件或非常高严格条件下与以下项杂交:(i) SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:17的成熟多肽编码序列,或者(ii) (i) 的全长互补体。

[0308] 在另一方面,该亲本纤维素酶由如下多核苷酸编码,该多核苷酸与SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:17的成熟多肽编码序列或其cDNA序列具有至少60%,例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列一致性。

[0309] 在另一方面,该亲本纤维素酶包含以下项或由以下项组成:SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:18的成熟多肽。

[0310] 在另一方面,该亲本纤维素酶是SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:18的成熟多肽的片段,其中该片段具有纤维素酶活性。

[0311] 在另一方面,该亲本是SEQ ID NO:2的成熟多肽的片段,该片段包含200个至315个,例如210个至290个、215个至295个、220个至300个氨基酸残基。

[0312] 在另一个实施例中,该亲本是SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:18的成熟多肽的等位基因变体。

[0313] SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:17的多核苷酸或其子序列,连同SEQ ID NO:4、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:18的多肽或其片段可以用于设计核酸探针,以根据本领域熟知的方法从不同属或种的菌株鉴定并克隆编码亲本的DNA。具体而言,此类探针可以遵循标准DNA印迹程序用于与感兴趣的细胞的基因组DNA或cDNA杂交,以便鉴定和分离其中相应的基因。此类探针可以明显短于完整序列,但是长度应为至少15,例如至少25、至少35、或至少70个核苷酸。优选地,该核酸探针的长度为至少100个核苷酸,例如长度为至少200个核苷酸、至少300个核苷酸、至少400个核苷酸、至少500个核苷酸、至少600个核苷酸、至少700个核苷酸、至少800个核苷酸、或至少900个核苷酸。可以使用DNA和RNA探针两者。典型地将探针进行标记(例如,用³²P、³H、³⁵S、生物素、或抗生物素蛋白),以检测相应的基因。本发明涵盖此类探针。

[0314] 可以筛选由这样的其他菌株制备的基因组DNA或cDNA文库中与上文所述的探针杂交并编码亲本的DNA。来自此类其他菌株的基因组DNA或其他DNA可以通过琼脂糖或聚丙烯酰胺凝胶电泳或其他分离技术来分离。可以将来自文库的DNA或分离的DNA转移到并固定在硝化纤维素或其他适合的载体材料上。为了鉴定与SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:17或其子序列杂交的克隆或DNA,将载体材料用于DNA印迹中。

[0315] 出于本发明的目的,杂交表示该多核苷酸与对应于以下各项的经标记的核酸探针杂交:(i) SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:17;(ii) SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:17的成熟多肽编码序列;(iii) 其cDNA序列;(iv) 其全长互补体;或(v) 其子序列;杂交是在非常低的至非常高的严格条件下进行。可以使用例如X-射线胶片或本领域已知的任何其他检测手段来检测在这些条件下核酸探针杂交的分子。

[0316] 在一个方面中,核酸探针是SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39、或SEQ ID NO:17的成熟多肽编码序列。在另一方面,核酸探针是SEQ ID NO:1的核苷酸64至1101。在另一方面,核酸探针是编码以下项的多核苷酸:SEQ ID NO:2的多肽;其成熟多肽;或其片段。在另一方面,核酸探针是SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:17或其cDNA序列。

[0317] 该亲本可以是一种杂合多肽,其中一种多肽的一个区域在另一种多肽的一个区域的N-末端或C-末端处融合。

[0318] 该亲本可以是融合多肽或可切割的融合多肽,其中另一个多肽的一个区域在一个多肽的N-末端或C-末端处融合。融合多肽是通过将编码另一种多肽的多核苷酸融合至本发明的多核苷酸而产生,其中一种多肽的区域(例如,纤维素酶的纤维素酶结合结构域)被添加或融合到另一多肽的区域的N-末端或C-末端,例如另一个纤维素酶的催化结构域和接头。用于产生融合多肽的技术在本领域是已知的,并且包括连接编码多肽的编码序列,这样使得它们在框内并且使得融合多肽的表达处于一个或多个相同启动子和终止子的控制下。还可以使用内含肽技术构建融合多肽,其中在翻译后产生融合多肽(Cooper等人,1993,EMBO J. [欧洲分子生物学学会杂志]12:2575-2583;Dawson等人,1994,Science [科学]266:776-779)。

[0319] 融合多肽可进一步包含两个多肽之间的切割位点。在融合蛋白分泌之时,该位点被切割而释放这两个多肽。切割位点的实例包括但不限于在以下文献中公开的位点:Martin等,2003,J. Ind. Microbiol. Biotechnol. 3:568-576;Svetina等人,2000,J. Biotechnol. [生物技术杂志]76:245-251;Rasmussen-Wilson等人,1997,Appl. Environ. Microbiol. [应用与环境微生物学],63:3488-3493;Ward等人,1995,Biotechnology [生物技术]13:498-503;以及Contreras等人,1991,Biotechnology [生物技术]9:378-381;Eaton等人,1986,Biochemistry [生物化学]25:505-512;Collins-Racie等人,1995,Biotechnology [生物技术]13:982-987;Carter等人,1989,Proteins: Structure, Function, and Genetics [蛋白质:结构、功能与遗传学]6:240-248;以及Stevens,2003,Drug Discovery World [药物发现世界]4:35-48。

[0320] 该亲本可以从任何属类的微生物中获得。出于本发明的目的,如在此结合一种给定的来源使用的术语“从……中获得”应意指由多核苷酸编码的亲本是由该来源或者由已经插入来自该来源的多核苷酸的菌株产生的。在一方面,该亲本是胞外分泌的。

[0321] 在另一方面,该亲本是梭孢壳属纤维素酶、圆孢霉属纤维素酶、脉孢菌属纤维素酶或梭孢壳属纤维素。优选地,该亲本是大孢圆孢霉纤维素酶、四孢脉孢菌纤维素酶、土生梭孢壳霉或赫坎梭孢壳霉纤维素酶。

[0322] 应理解的是对于前述的物种,本发明涵盖完全和不完全阶段(perfect and imperfect states),和其他分类学的等同物(equivalent),例如无性型(anamorph),而与它们已知的种名无关。本领域的技术人员将容易地识别适当等效物的身份。

[0323] 这些物种的菌株可容易地在许多培养物保藏中心为公众所获得,如美国典型培养物保藏中心(ATCC)、德国微生物和细胞培养物保藏中心(Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH,DSMZ)、荷兰菌种保藏中心(Centraalbureau Voor Schimmelcultures,CBS)以及美国农业研究服务专利培养物保藏中心北方地区研究中心(Agricultural Research Service Patent Culture Collection,Northern Regional Research Center,NRRL)。

[0324] 可以使用上文提及的探针从其他来源,包括从自然界(例如,土壤、堆肥、水等)分离的微生物或直接从自然材料(例如,土壤、堆肥、水等)获得的DNA样品鉴定和获得该亲本。用于直接地从自然生活环境分离微生物和DNA的技术是本领域中熟知的。然后可通过类似地筛选另一种微生物或混合DNA样品的基因组DNA或cDNA文库来获得编码亲本的多核苷酸。一旦已经用一种或多种探针检测到编码亲本的多核苷酸,就可以通过使用本领域普通技术

人员已知的技术分离或克隆该多核苷酸(参见,例如,Sambrook等人,1989,见上文)。

[0325] 本发明进一步涉及亲本GH45纤维素酶的变体,该变体包含催化结构域和纤维素结合结构域,其中该纤维素酶结合结构域与该催化结构域是异源的,并且其中该变体与亲本GH45纤维素酶相比具有改进的生物整理活性。

[0326] 在一方面,接头位于催化结构域与纤维素结合结构域之间。

[0327] 在一个实施例中,接头与纤维素结合结构域是同源的。

[0328] 在一个实施例中,接头与催化结构域是同源的。

[0329] 在一个实施例中,催化结构域选自下组,该组由以下各项组成:

[0330] a.与SEQ ID NO:12、SEQ ID NO:13或SEQ ID NO:14的催化结构域具有至少60%序列一致性的多肽;

[0331] b.由在低严格条件下与以下项杂交的多核苷酸编码的多肽:(i) SEQ ID NO:37、SEQ ID NO:38或SEQ ID NO:39的催化结构域编码序列,(ii) 其cDNA序列,或(iii) (i) 或(ii)的全长互补体;

[0332] c.由与SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39的催化结构域编码序列或其cDNA序列具有至少60%一致性的多核苷酸编码的多肽;以及

[0333] d. SEQ ID NO:12、SEQ ID NO:13或SEQ ID NO:14的催化结构域的片段,该片段具有纤维素酶活性。

[0334] 在一个实施例中,催化结构域与SEQ ID NO:12、SEQ ID NO:13或SEQ ID NO:14的催化结构域具有至少60%,例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列一致性。

[0335] 在一个实施例中,催化结构域由以下多核苷酸编码,该多核苷酸在低严格条件、中严格条件、中-高严格条件、高严格条件或非常高严格条件下与以下项杂交:(i) SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39的催化结构域编码序列,或(ii) (i)的全长互补体。

[0336] 在一个实施例中,催化结构域由多核苷酸编码,该多核苷酸与SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39的催化结构域编码序列或其cDNA序列具有至少60%,例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列一致性。

[0337] 在一个实施例中,催化结构域包含SEQ ID NO:12的氨基酸22-223、SEQ ID NO:13的氨基酸1-200或SEQ ID NO:14的氨基酸22-223或由其组成。

[0338] 在一个实施例中,纤维素结合结构域选自下组,该组由以下各项组成:

[0339] a.与SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:8的纤维素结合结构域具有至少60%序列一致性的多肽;

[0340] b.由在低严格条件下与以下项杂交的多核苷酸编码的多肽:(i) SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:7的纤维素结合结构域编码序列,(ii) 其cDNA序列,或(iii) (i) 或(ii)的全长互补体;

[0341] c.由与SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:7的纤维素结合结构域编码序列或其cDNA序列具有至少60%一致性的多核苷酸编码的多肽;以及

[0342] d. SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:8的纤维素结合结构域的片段,该片段具有纤维素结合活性。

[0343] 在一个实施例中,纤维素结合结构域与SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:8的纤维素结合结构域具有至少60%,例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列一致性。

[0344] 在一个实施例中,纤维素结合结构域由以下多核苷酸编码,该多核苷酸在低严格条件、中严格条件、中-高严格条件、高严格条件或非常高严格条件下与以下项杂交:(i) SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39、SEQ ID NO:7的纤维素结合结构域编码序列,或(ii) (i)的全长互补体。

[0345] 在一个实施例中,纤维素结合结构域由以下多核苷酸编码,该多核苷酸与SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:7的纤维素结合结构域编码序列或其cDNA序列具有至少60%,例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列一致性。

[0346] 在一个实施例中,纤维素结合结构域包含以下项或由以下项组成:SEQ ID NO:12的氨基酸262-299、SEQ ID NO:13的氨基酸258-295、SEQ ID NO:14的氨基酸257-293、或SEQ ID NO:8的氨基酸249-286。

[0347] 在一个实施例中,接头选自下组,该组由以下各项组成:

[0348] a.与SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:8的接头具有至少60%序列一致性的多肽;

[0349] b.由在低严格条件下与以下项杂交的多核苷酸编码的多肽:(i) SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:7的接头编码序列,(ii) 其cDNA序列,或(iii) (i)或(ii)的全长互补体;

[0350] c.由与SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:7的接头编码序列或其cDNA序列具有至少60%一致性的多核苷酸编码的多肽;以及

[0351] d.SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:8的接头的片段,该片段具有纤维素结合活性。

[0352] 在一个实施例中,接头与SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:8的接头具有至少60%,例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列一致性。

[0353] 在一个实施例中,接头由以下多核苷酸编码,该多核苷酸在低严格条件、中严格条件、中-高严格条件、高严格条件或非常高严格条件下与以下项杂交:(i) SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39、SEQ ID NO:7的接头编码序列,或(ii) (i)的全长互补体。

[0354] 在一个实施例中,接头由多核苷酸编码,该多核苷酸与SEQ ID NO:37、SEQ ID NO:38、SEQ ID NO:39或SEQ ID NO:7的接头编码序列或其cDNA序列具有至少60%,例如至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列一致性。

[0355] 在一个实施例中,接头包含以下项或由以下项组成:SEQ ID NO:12的氨基酸224-261、SEQ ID NO:13的氨基酸201-257、SEQ ID NO:14的氨基酸224-256、或SEQ ID NO:8的氨基酸203-248。

[0356] 在一个实施例中,该变体与该亲本纤维素酶的氨基酸序列具有至少60%,例如至

少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%一致性、至少96%、至少97%、至少98%、或至少99%、但小于100%的序列一致性。

[0357] 在一个实施例中,该变体由200至315,例如210至290、215至295、220至300个氨基酸组成。

[0358] 在一个实施例中,该变体在与SEQ ID NO:2的成熟多肽的位置292、274、266、265、255、246、237、224以及221相对应的一个或多个位置处包含改变。

[0359] 在一个实施例中,改变的数目是1-20个,例如,1-10和1-5个,如1、2、3、4、5、6、7、8、9或10个改变。

[0360] 变体的制备

[0361] 本发明还涉及用于获得具有纤维素酶活性的变体的方法,这些方法包括:(a)在与SEQ ID NO:2的成熟多肽的位置292、274、266、265、255、246、237、224和221对应的一个或多个(例如,若干个)位置处向亲本纤维素酶引入改变,其中该变体具有纤维素酶活性;并且(b)回收该变体。

[0362] 可以使用本领域已知的任何诱变程序来制备变体,如定点诱变、合成基因构建、半合成基因构建、随机诱变、改组等。

[0363] 定点诱变是在编码该亲本的多核苷酸中的一个或多个限定位点处引入一个或多个(例如,若干个)突变的技术。

[0364] 通过涉及使用含有所希望的突变的寡核苷酸引物的PCR可以体外实现定点诱变。也可以通过盒式诱变进行体外定点诱变,该盒式诱变涉及由限制酶在包含编码亲本的多核苷酸的质粒中的位点处切割并且随后将含有突变的寡核苷酸连接在多核苷酸中。通常,消化该质粒与该寡核苷酸的限制酶是相同的,从而允许该质粒的粘性末端以及插入片段彼此连接。参见例如,Scherer和Davis,1979,Proc.Natl.Acad.Sci.USA 76:4949-4955[Scherer和Davis,1979,美国国家科学院院刊,76:4949-4955];和Barton等人,1990,Nucleic Acids Res.18:7349-4966[Barton等人,1990,核酸研究,18:7349-4966]。

[0365] 还可以通过本领域已知的方法在体内实现定点诱变。参见例如,美国专利申请公开号2004/0171154;Storici等人,2001,Nature Biotechnol.19:773-776[Storici等人,2001,自然生物技术,19:773-776];Kren等人,1998,Nat.Med.4:285-290[Kren等人,1998,自然医学,4:285-290];以及Calissano and Macino,1996,Fungal Genet.Newslett.43:15-16[Calissano和Macino,1996,真菌遗传学通讯,43:15-16]。

[0366] 在本发明中可以使用任何定点诱变程序。存在许多可以用于制备变体的可商购的试剂盒。

[0367] 合成基因构建需要体外合成设计的多核苷酸分子以编码感兴趣的多肽。基因合成可以利用多种技术来进行,如由Tian等人(2004,Nature[自然]432:1050-1054)所述的基于多路微芯片的技术、以及在光可编程的微流芯片上合成并组装寡核苷酸的类似技术。

[0368] 使用已知的诱变、重组和/或改组方法、随后进行一个相关的筛选程序可以做出单一或多种氨基酸取代、缺失和/或插入并对其进行测试,这些相关的筛选程序例如由Reidhaar-Olson和Sauer,1988,Science[科学]241:53-57;Bowie和Sauer,1989,Proc.Natl.Acad.Sci.USA[美国国家科学院院刊]86:2152-2156;WO 95/17413;或WO 95/22625。其他可以使用的方法包括易错PCR、噬菌体展示(例如Lowman等人,1991,

Biochemistry[生物化学]30:10832-10837;美国专利号5,223,409;WO 92/06204)以及区域定向诱变(Derbyshire等人,1986,Gene[基因]46:145;Ner等人,1988,DNA 7:127)。

[0369] 诱变/改组方法可以与高通量自动化筛选方法组合以检测由宿主细胞表达的克隆的诱变多肽的活性(Ness等人,1999,Nature Biotechnology[自然生物技术]17:893-896)。可从宿主细胞回收编码活性多肽的诱变的DNA分子,并使用本领域的标准方法快速测序。这些方法允许迅速确定多肽中个体氨基酸残基的重要性。

[0370] 通过组合合成基因构建、和/或定点诱变、和/或随机诱变、和/或改组的多方面来实现半合成基因的构建。半合成构建典型地是利用合成的多核苷酸片段的过程结合PCR技术。因此,基因的限定区域可以从头合成,而其他区域可以使用位点特异性诱变引物来扩增,然而还有其他区域可以进行易错PCR或非易错PCR扩增。然后可以对多核苷酸子序列进行改组。

[0371] 本发明进一步涉及用于获得变体的方法,其包括用异源纤维素结合结构域替代亲本纤维素酶的纤维素结合结构域;或将异源纤维素结合结构域添加至不包含纤维素结合结构域的亲本纤维素酶。用于产生变体的技术在本领域是已知的,并且包括连接编码序列,这样使得它们在框内并且使得变体的表达处于一个或多个相同启动子和终止子的控制下。还可以使用内含肽技术构建变体,其中在翻译后产生变体(Cooper等人,1993,EMBO J.[欧洲分子生物学学会杂志]12:2575-2583;Dawson等人,1994,Science[科学]266:776-779)。

[0372] 多核苷酸

[0373] 本发明还涉及编码本发明的变体的多核苷酸。

[0374] 核酸构建体

[0375] 本发明还涉及包含编码本发明的变体的、可操作地连接至一个或多个控制序列上的多核苷酸的核酸构建体,该一个或多个控制序列在与控制序列相容的条件下指导编码序列在适合的宿主细胞中表达。

[0376] 可以按多种方式来操纵该多核苷酸以提供变体的表达。取决于表达载体,在多核苷酸插入载体之前对其进行操纵可以是令人希望的或必需的。用于利用重组DNA方法修饰多核苷酸的技术在本领域是熟知的。

[0377] 控制序列可以是启动子,即由宿主细胞识别用于表达该多核苷酸的多核苷酸。启动子含有介导变体的表达的转录控制序列。启动子可以在宿主细胞中显示出转录活性的任何多核苷酸,包括突变型、截短型和杂合型启动子,并且可以从编码与宿主细胞同源或异源的细胞外或细胞内多肽的基因获得。

[0378] 用于在细菌宿主细胞中指导本发明核酸构建体的转录的适合启动子的实例是从以下基因中获得的启动子:解淀粉芽孢杆菌 α -淀粉酶基因(amyQ)、地衣芽孢杆菌 α -淀粉酶基因(amyL)、地衣芽孢杆菌青霉素酶基因(penP)、嗜热脂肪芽孢杆菌产麦芽糖淀粉酶基因(amyM)、枯草芽孢杆菌果聚糖蔗糖酶基因(sacB)、枯草芽孢杆菌xy1A和xy1B基因、苏云金芽孢杆菌cryIIIA基因(Agaisse和Lereclus,1994,Molecular Microbiology[分子微生物学]13:97-107)、大肠杆菌lac操纵子、大肠杆菌trc启动子(Egon等人,1988,Gene[基因]69:301-315)、天蓝链霉菌琼脂水解酶基因(dagA)和原核 β -内酰胺酶基因(Villa-Kamaroff等人,1978,Proc.Natl.Acad.Sci.USA[美国国家科学院院刊]75:3727-3731)以及tac启动子(DeBoer等人,1983,Proc.Natl.Acad.Sci.USA[美国国家科学院院刊]80:21-25)。其他启动

子描述于以下文献中：“Useful proteins from recombinant bacteria[来自重组细菌的有用蛋白质]”，Gilbert等人，1980，Scientific American[科学美国人]242:74-94；以及 Sambrook等人，1989，见上文。串联启动子的实例披露在WO 99/43835中。

[0379] 用于指导本发明的核酸构建体在丝状真菌宿主细胞中的转录的适合启动子的实例是从以下各项的基因获得的启动子：构巢曲霉乙酰胺酶、黑曲霉中性 α -淀粉酶、黑曲霉稳定性 α -淀粉酶、黑曲霉或泡盛曲霉葡糖淀粉酶(glaA)、米曲霉TAKA淀粉酶、米曲霉碱性蛋白酶、米曲霉丙糖磷酸异构酶、尖镰孢胰蛋白酶样蛋白酶(WO 96/00787)、镶片镰孢淀粉葡糖苷酶(WO 00/56900)、镶片镰孢Daria(*Fusarium venenatum* Daria)(WO 00/56900)、镶片镰孢Quinn(*Fusarium venenatum* Quinn)(WO 00/56900)、米黑根毛霉(*Rhizomucor miehei*)脂肪酶、米黑根毛霉天冬氨酸蛋白酶、里氏木霉 β -葡糖苷酶、里氏木霉纤维二糖水解酶I、里氏木霉纤维二糖水解酶II、里氏木霉内切葡聚糖酶I、里氏木霉内切葡聚糖酶II、里氏木霉内切葡聚糖酶III、里氏木霉内切葡聚糖酶IV、里氏木霉内切葡聚糖酶V、里氏木霉木聚糖酶I、里氏木霉木聚糖酶II、里氏木霉 β -木糖苷酶，以及NA2-tpi启动子(一种修饰的启动子，其来自曲霉属中性 α -淀粉酶基因，其中未翻译的前导子由曲霉属丙糖磷酸异构酶基因的未翻译的前导子替代；非限制性实例包括来自编码中性 α -淀粉酶的黑曲霉基因的修饰的启动子，其中已经用来自编码丙糖磷酸异构酶的构巢曲霉或米曲霉基因的未翻译的前导子替换未翻译的前导子)；及其突变型、截短型及杂合型启动子。

[0380] 在酵母宿主中，有用的启动子从以下各项的基因获得：酿酒酵母烯醇酶(ENO-1)、酿酒酵母半乳糖激酶(GAL1)、酿酒酵母醇脱氢酶/甘油醛-3-磷酸脱氢酶(ADH1、ADH2/GAP)、酿酒酵母丙糖磷酸异构酶(TPI)、酿酒酵母金属硫蛋白(CUP1)、以及酿酒酵母3-磷酸甘油酸激酶。由Romanos等人，1992，Yeast[酵母]8:423-488描述了酵母宿主细胞的其他有用启动子。

[0381] 控制序列也可为由宿主细胞识别以终止转录的转录终止子。终止子序列被可操作地连接至编码变体的多核苷酸的3'-末端。可以使用在宿主细胞中具有功能的任何终止子。

[0382] 细菌宿主细胞的优选终止子从以下各项的基因获得：克劳氏芽孢杆菌碱性蛋白酶(aprH)、地衣芽孢杆菌 α -淀粉酶(amyL)和大肠杆菌核糖体RNA(rrnB)。

[0383] 丝状真菌宿主细胞的优选终止子是从以下各项的基因中获得的：构巢曲霉邻氨基苯甲酸合酶、黑曲霉葡糖淀粉酶、黑曲霉 α -葡糖苷酶、米曲霉TAKA淀粉酶以及尖镰孢胰蛋白酶样蛋白酶。

[0384] 用于酵母宿主细胞的优选终止子从以下各项的基因获得：酿酒酵母烯醇酶、酿酒酵母细胞色素C(CYC1)、以及酿酒酵母甘油醛-3-磷酸脱氢酶。由Romanos等人，1992(见上文)描述了酵母宿主细胞的其他有用终止子。

[0385] 控制序列还可以是启动子下游和基因的编码序列上游的mRNA稳定子区域，其增加该基因的表达。

[0386] 适合的mRNA稳定子区域的实例是从以下基因获得：苏云金芽孢杆菌cryIIIA基因(WO 94/25612)和枯草芽孢杆菌SP82基因(Hue等人，1995，Journal of Bacteriology[细菌学杂志]177:3465-3471)。

[0387] 该控制序列也可以是前导序列，一种对宿主细胞翻译很重要的非翻译mRNA区域。前导子序列可操作地连接至编码该变体的多核苷酸的5'-末端。在宿主细胞中起作用的任

何前导子都可以使用。

[0388] 用于丝状真菌宿主细胞的优选前导子是从米曲霉TAKA淀粉酶和构巢曲霉丙糖磷酸异构酶的基因获得。

[0389] 对于酵母宿主细胞合适的前导子从以下酶的基因获得：酿酒酵母烯醇化酶(ENO-1)、酿酒酵母3-磷酸甘油酸激酶、酿酒酵母 α 因子、和酿酒酵母醇脱氢酶/甘油醛-3-磷酸脱氢酶(ADH2/GAP)。

[0390] 该控制序列还可以是一种多腺苷酸化序列，即被可操作地连接至该变体编码序列的3'-末端并且当转录时由宿主细胞识别成将多腺苷酸残基添加到所转录的mRNA上的一个信号的一种序列。可以使用在宿主细胞中起作用的任何多腺苷酸化序列。

[0391] 用于丝状真菌宿主细胞的优选聚腺苷酸化序列是从以下各项的基因获得：构巢曲霉邻氨基苯甲酸合酶、黑曲霉葡糖淀粉酶、黑曲霉 α -葡糖苷酶、米曲霉TAKA淀粉酶以及尖镰孢胰蛋白酶样蛋白酶。

[0392] 对于酵母宿主细胞有用的多聚腺苷酸化序列由Guo和Sherman, 1995, Mol. Cellular Biol. [分子细胞生物学] 15: 5983-5990描述。

[0393] 控制序列还可以是信号肽编码区，其编码与变体的N-末端连接的信号肽，并且指导变体进入细胞的分泌通路。多核苷酸的编码序列的5'-端可以固有地含有信号肽编码序列，该信号肽编码序列在翻译阅读框中与编码变体的编码序列的区段天然地连接在一起。可替代地，编码序列的5'-端可以含有对于编码序列是外源的信号肽编码序列。在编码序列天然地不含有信号肽编码序列的情况下，可能需要外源信号肽编码序列。可替代地，外源信号肽编码序列可以简单地替代天然信号肽编码序列，以便增强变体的分泌。然而，可以使用指导表达的变体进入宿主细胞的分泌通路的任何信号肽编码序列。

[0394] 用于细菌宿主细胞的有效信号肽编码序列是从以下各项的基因获得的信号肽编码序列：芽孢杆菌属NCIB 11837产麦芽糖淀粉酶、地衣芽孢杆菌枯草杆菌蛋白酶、地衣芽孢杆菌 β -内酰胺酶、嗜热脂肪芽孢杆菌 α -淀粉酶、嗜热脂肪芽孢杆菌中性蛋白酶(nprT、nprS、nprM)和枯草芽孢杆菌prsA。另外的信号肽由Simonen和Palva, 1993, Microbiological Reviews [微生物评论] 57: 109-137描述。

[0395] 用于丝状真菌宿主细胞的有效信号肽编码序列是从以下酶的基因获得的信号肽编码序列：黑曲霉中性淀粉酶、黑曲霉葡糖淀粉酶、米曲霉TAKA淀粉酶、特异腐质霉(*Humicola insolens*)纤维素酶、特异腐质霉内切葡聚糖酶V、疏棉状腐质霉(*Humicola lanuginosa*)脂肪酶和米黑根毛霉天冬氨酸蛋白酶。

[0396] 用于酵母宿主细胞的有用的信号肽从酿酒酵母 α -因子和酿酒酵母转化酶的基因获得。Romanos等人(1992, 见上文)描述了其他有用的信号肽编码序列。

[0397] 该控制序列还可以是编码位于变体的N-末端的前肽的前肽编码序列。生成的多肽被称为前体酶(proenzyme)或多肽原(或在一些情况下被称为酶原(zymogen))。多肽原通常是无活性的并且可以通过催化切割或自身催化切割来自多肽原的前肽而转化为活性多肽。可以从枯草芽孢杆菌碱性蛋白酶(aprE)、枯草芽孢杆菌中性蛋白酶(nprT)、嗜热毁丝霉漆酶(WO 95/33836)、米黑根毛霉天冬氨酸蛋白酶和酿酒酵母 α -因子的基因获得前肽编码序列。

[0398] 在信号肽序列和前肽序列二者都存在的情况下，前肽序列定位成紧邻变体的N-末

端并且信号肽序列定位成紧邻前肽序列的N-末端。

[0399] 还令人希望的可以是添加相对于宿主细胞的生长来调节变体的表达的调节序列。调节系统的实例是响应于化学或物理刺激而引起基因的表达开启或关闭的那些,包括调控化合物的存在。原核系统中的调节系统包括lac、tac和trp操纵子系统。在酵母中,可以使用ADH2系统或GAL1系统。在丝状真菌中,可以使用黑曲霉葡萄糖淀粉酶启动子、米曲霉TAKA α -淀粉酶启动子、以及米曲霉葡萄糖淀粉酶启动子。调控序列的其他实例是允许基因扩增的那些。在真核系统中,这些调控序列包括在甲氨蝶呤存在下被扩增的二氢叶酸还原酶基因以及用重金属扩增的金属硫蛋白基因。在这些情况下,编码该变体的多核苷酸将与该调节序列可操作地连接。

[0400] 表达载体

[0401] 本发明还涉及包含编码本发明的变体的多核苷酸、启动子以及转录和翻译终止信号的重组表达载体。各种核苷酸和控制序列可以连接在一起以产生重组表达载体,该重组表达载体可以包括一个或多个合宜的限制位点以允许在这样的位点处插入或取代编码变体的多核苷酸。可替代地,可以通过将多核苷酸或包含该多核苷酸的核酸构建体插入用于表达的适当载体中来表达该多核苷酸。在产生表达载体时,编码序列位于载体中,这样使得编码序列与用于表达的适当控制序列可操作地连接。

[0402] 重组表达载体可以是可方便地经受重组DNA程序并且可引起多核苷酸表达的任何载体(例如,质粒或病毒)。载体的选择将典型地取决于载体与待引入载体的宿主细胞的相容性。载体可以是线状或闭合的环状质粒。

[0403] 载体可以是自主复制载体,即作为染色体外实体存在的载体,其复制独立于染色体复制,例如质粒、染色体外元件、微染色体或人工染色体。载体可以含有用于确保自我复制的任何装置。可替代地,载体可以是这样的载体,当它引入宿主细胞中时整合入基因组中并与其中已整合了它的染色体一起复制。此外,可以使用单独的载体或质粒或两个或更多个载体或质粒,其共同含有待引入宿主细胞基因组的总DNA,或可以使用转座子。

[0404] 载体优选地含有一个或多个选择性标记,这些选择性标记容许容易地选择转化细胞、转染细胞、转导细胞等。选择性标记是这样一种基因,该基因的产物提供了杀生物剂抗性或病毒抗性、重金属抗性、营养缺陷型的原养型等。

[0405] 细菌选择性标记的实例是地衣芽孢杆菌或枯草芽孢杆菌da1基因,或赋予抗生素抗性(如氨苄青霉素、氯霉素、卡那霉素、新霉素、大观霉素或四环素抗性)的标记。用于酵母宿主细胞的适合的标记包括但不限于ADE2、HIS3、LEU2、LYS2、MET3、TRP1和URA3。用于在丝状真菌宿主细胞中使用的选择性标记包括但不限于amdS(乙酰胺酶)、argB(鸟氨酸氨甲酰基转移酶)、bar(草胺磷乙酰转移酶)、hph(潮霉素磷酸转移酶)、niaD(硝酸还原酶)、pyrG(乳清苷-5'-磷酸脱羧酶)、sC(硫酸腺苷基转移酶)、以及trpC(邻氨基苯甲酸合酶),连同其等效物。优选在曲霉属细胞中使用的是构巢曲霉或米曲霉amdS和pyrG基因以及吸水链霉菌bar基因。

[0406] 载体优选含有允许载体整合到宿主细胞的基因组中或载体在细胞中独立于基因组自主复制的元件。

[0407] 对于整合到该宿主细胞基因组中,该载体可以依靠编码该变体的多核苷酸序列或用于通过同源或非同源重组整合到该基因组中的载体的任何其他元件。可替代地,载体可

以含有用于指导通过同源重组而整合入宿主细胞基因组的染色体中的精确位置的另外的多核苷酸。为了增加在精确位置整合的可能性,整合的元件应含有足够数量的核酸,如100至10,000个碱基对、400至10,000个碱基对、以及800至10,000个碱基对,其与相应的靶序列具有高度的序列一致性以增强同源重组的可能性。这些整合元件可为与宿主细胞基因组中的靶序列同源的任何序列。此外,这些整合元件可为非编码多核苷酸或编码多核苷酸。另一方面,载体可通过非同源重组整合入宿主细胞的基因组中。

[0408] 对于自主复制,载体可以进一步含有使载体能够在所讨论的宿主细胞中自主地进行复制的复制起点。复制起点可为在细胞中有功能的介导自主复制的任何质粒复制子。术语“复制起点”或“质粒复制子”意指使质粒或载体能够在体内复制的多核苷酸。

[0409] 细菌复制起点的实例是允许在大肠杆菌中复制的质粒pBR322、pUC19、pACYC177、以及pACYC184的复制起点,以及允许在芽孢杆菌属中复制的质粒pUB110、pE194、pTA1060、以及pAMB1的复制起点。

[0410] 用于酵母宿主细胞中的复制起点的实例是2微米复制起点、ARS1、ARS4、ARS1与CEN3的组合、及ARS4与CEN6的组合。

[0411] 用于丝状真菌细胞中的复制起点的实例是AMA1和ANS1 (Gems等人,1991, Gene [基因] 98:61-67; Cullen等人,1987, Nucleic Acids Res. [核酸研究] 15:9163-9175; WO 00/24883)。可根据WO 00/24883中公开的方法完成AMA1基因的分离和包含该基因的质粒或载体的构建。

[0412] 可以将多于一个拷贝的本发明的多核苷酸插入宿主细胞中以增加变体的产生。通过将序列的至少一个另外的拷贝整合到宿主细胞基因组中或通过包括一个与该多核苷酸一起的可扩增的选择性标记基因可以获得多核苷酸的增加的拷贝数目,其中通过在适当的选择性试剂的存在下培养细胞可以选择包含选择性标记基因的经扩增的拷贝的细胞、以及由此该多核苷酸的另外的拷贝。

[0413] 用于连接以上所述的元件以构建本发明的重组表达载体的程序是本领域的普通技术人员熟知的(参见,例如,Sambrook等人,1989(见上文))。

[0414] 宿主细胞

[0415] 本发明还涉及重组宿主细胞,这些重组宿主细胞包含编码本发明的变体的、可操作地连接至一个或多个控制序列的多核苷酸,该一个或多个控制序列指导本发明的变体的产生。将包含多核苷酸的构建体或载体引入宿主细胞中,这样使得该构建体或载体被维持作为染色体整合体或作为自主复制的染色体外载体,如早前所述。术语“宿主细胞”涵盖由于复制过程中发生的突变而与亲本细胞不同的亲本细胞的任何后代。宿主细胞的选择在很大程度上将取决于编码变体的基因及其来源。

[0416] 宿主细胞可以是在变体的重组产生中有用的任何细胞,例如原核细胞或真核细胞。

[0417] 原核宿主细胞可以是任何革兰氏阳性或革兰氏阴性细菌。革兰氏阳性细菌包括但不限于芽孢杆菌属、梭菌属、肠球菌属、土芽孢杆菌属、乳杆菌属、乳球菌属、海洋芽孢杆菌属、葡萄球菌属、链球菌属和链霉菌属。革兰氏阴性细菌包括但不限于:弯曲杆菌属、大肠杆菌、黄杆菌属、梭杆菌属、螺杆菌属、泥杆菌属、奈瑟氏菌属、假单胞菌属、沙门氏菌属、以及脲原体属。

[0418] 细菌宿主细胞可以是任何芽孢杆菌属细胞,包括但不限于嗜碱芽孢杆菌、解淀粉芽孢杆菌、短芽孢杆菌、环状芽孢杆菌、克劳氏芽孢杆菌、凝结芽孢杆菌、坚硬芽孢杆菌、灿烂芽孢杆菌、迟缓芽孢杆菌、地衣芽孢杆菌、巨大芽孢杆菌、短小芽孢杆菌、嗜热脂肪芽孢杆菌、枯草芽孢杆菌以及苏云金芽孢杆菌细胞。

[0419] 细菌宿主细胞还可为任何链球菌属细胞,包括但不限于:似马链球菌、酿脓链球菌、乳房链球菌和马链球菌兽瘟亚种细胞。

[0420] 细菌宿主细胞还可以是任何链霉菌属细胞,包括但不限于:不产色链霉菌、除虫链霉菌、天蓝色链霉菌、灰色链霉菌、以及变铅青链霉菌细胞。

[0421] 将DNA引入芽孢杆菌属细胞中可以通过以下来实现:原生质体转化(参见例如,Chang和Cohen,1979,Mol.Gen.Genet.[分子遗传学与基因组学]168:111-115)、感受态细胞转化(参见例如,Young和Spizizen,1961,J.Bacteriol.[细菌学杂志]81:823-829,或Dubnau和Davidoff-Abelson,1971,J.Mol.Biol.[分子生物学杂志]56:209-221)、电穿孔(参见例如,Shigekawa和Dower,1988,Biotechniques[生物技术]6:742-751)或接合(参见例如,Koehler和Thorne,1987,J.Bacteriol.[细菌学杂志]169:5271-5278)。将DNA引入大肠杆菌细胞中可以通过以下来实现:原生质体转化(参见例如,Hanahan,1983,J.Mol.Biol.[分子生物学杂志]166:557-580)或电穿孔(参见例如,Dower等人,1988,Nucleic Acids Res.[核酸研究]16:6127-6145)。将DNA引入链霉菌属细胞中可以通过以下来实现:原生质体转化、电穿孔(参见例如,Gong等人,2004,Folia Microbiol.(Praha)[叶线形微生物学(布拉格)]49:399-405)、接合(参见例如,Mazodier等人,1989,J.Bacteriol.[细菌学杂志]171:3583-3585)或转导(参见例如,Burke等人,2001,Proc.Natl.Acad.Sci.USA[美国国家科学院院刊]98:6289-6294)。将DNA引入假单胞菌属细胞中可以通过以下来实现:电穿孔(参见例如,Choi等人,2006,J.Microbiol.Methods[微生物学方法杂志]64:391-397)或接合(参见例如,Pinedo和Smets,2005,Appl.Environ.Microbiol.[应用与环境微生物学]71:51-57)。将DNA引入链球菌属细胞中可以通过以下来实现:天然感受态(参见例如,Perry和Kuramitsu,1981,Infect.Immun.[感染与免疫]32:1295-1297)、原生质体转化(参见例如,Catt和Jollick,1991,Microbios[微生物学]68:189-207)、电穿孔(参见例如,Buckley等人,1999,Appl.Environ.Microbiol.[应用与环境微生物学]65:3800-3804)或接合(参见例如,Clewell,1981,Microbiol.Rev.[微生物学评论]45:409-436)。然而,可使用本领域已知的将DNA引入宿主细胞的任何方法。

[0422] 宿主细胞还可为真核生物,如哺乳动物、昆虫、植物或真菌细胞。

[0423] 宿主细胞可为真菌细胞。如在此使用的“真菌”包括子囊菌门(Ascomycota)、担子菌门(Basidiomycota)、壶菌门(Chytridiomycota)和接合菌门(Zygomycota)以及卵菌门(Oomycota)和所有有丝分裂孢子真菌(如由Hawksworth等人所定义的,在:Ainsworth and Bisby's Dictionary of The Fungi[Ainsworth和Bisby的真菌大词典],第8版,1995,国际CAB,大学出版社,剑桥,英国)。

[0424] 真菌宿主细胞可以是酵母细胞。如在此所用的“酵母”包括产子囊酵母(ascosporogenous yeast)(内孢霉目(Endomycetales))、产担子酵母(basidiosporogenous yeast)和属于半知菌类(Fungi Imperfecti)(芽孢纲(Blastomycetes))的酵母。由于酵母的分类可在将来变化,出于本发明的目的,酵母应当如

Biology and Activities of Yeast [酵母的生物学与活性] (Skinner, Passmore 和 Davenport 编, Soc. App. Bacteriol. Symposium Series No. 9 [应用细菌学学会专题论文集系列9], 1980) 中所描述的那样定义。

[0425] 酵母宿主细胞可以是假丝酵母属细胞、汉逊酵母属细胞、克鲁维酵母属细胞、毕赤酵母属细胞、酵母菌属细胞、裂殖酵母或耶罗维亚酵母属细胞、如乳酸克鲁维酵母细胞、卡氏酵母细胞、酿酒酵母细胞、糖化酵母细胞、道格拉斯酵母 (*Saccharomyces douglasii*) 细胞、克鲁弗酵母细胞、诺地酵母细胞、卵形酵母细胞或解脂耶罗维亚酵母细胞。

[0426] 真菌宿主细胞可为丝状真菌细胞。“丝状真菌”包括真菌门 (Eumycota) 和卵菌门 (Oomycota) 的亚门的所有丝状形式 (如由 Hawksworth 等人, 1995, 见上文)。丝状真菌通常的特征在于由几丁质、纤维素、葡聚糖、壳多糖、甘露聚糖、以及其他复杂多糖构成的菌丝体壁。营养生长是通过菌丝延长, 而碳分解代谢是专性需氧的。相反, 酵母 (如酿酒酵母) 的营养生长是通过单细胞菌体的出芽 (budding), 而碳分解代谢可以是发酵性的。

[0427] 丝状真菌宿主细胞可以是枝顶孢属、曲霉属、短梗霉属、烟管霉属 (*Bjerkandera*)、拟腊菌属、金孢子菌属、鬼伞属、革盖菌属 (*Coriolus*)、隐球菌属、线黑粉菌科 (*Filibasidium*)、镰孢属、腐质霉属、梨孢菌属、毛霉属、毁丝霉属、新美鞭菌属、链孢菌属、拟青霉属、青霉属、平革菌属、射脉菌属 (*Phlebia*)、瘤胃壶菌属、侧耳属 (*Pleurotus*)、裂褶菌属、篮状菌属、嗜热子囊菌属、梭孢壳属、弯颈霉属、栓菌属 (*Trametes*) 或木霉属细胞。

[0428] 例如, 丝状真菌宿主细胞可以是泡盛曲霉、臭曲霉、烟曲霉、日本曲霉、构巢曲霉、黑曲霉、米曲霉、黑刺烟管菌 (*Bjerkandera adusta*)、干拟蜡菌 (*Ceriporiopsis aneirina*)、卡内基拟蜡菌 (*Ceriporiopsis caregiea*)、浅黄拟蜡孔菌 (*Ceriporiopsis gilvescens*)、潘诺希塔拟蜡菌 (*Ceriporiopsis pannocinta*)、环带拟蜡菌 (*Ceriporiopsis rivulosa*)、微红拟蜡菌 (*Ceriporiopsis subrufa*)、虫拟蜡菌 (*Ceriporiopsis subvermispora*)、狭边金孢子菌 (*Chrysosporium inops*)、嗜角质金孢子菌、卢克诺文思金孢子菌 (*Chrysosporium lucknowense*)、粪状金孢子菌 (*Chrysosporium merdarium*)、租金孢子菌、女王杜香金孢子菌 (*Chrysosporium queenslandicum*)、热带金孢子菌、褐薄金孢子菌 (*Chrysosporium zonatum*)、灰盖鬼伞 (*Coprinus cinereus*)、毛革盖菌 (*Coriolus hirsutus*)、杆孢状镰孢、谷类镰孢、库威镰孢、大刀镰孢、禾谷镰孢、禾赤镰孢、异孢镰孢、合欢木镰孢、尖镰孢、多枝镰孢、粉红镰孢、接骨木镰孢、肤色镰孢、拟分枝孢镰孢、硫色镰孢、圆镰孢、拟丝孢镰孢、镶片镰孢、特异腐质霉、柔毛腐质霉、米黑毛霉、嗜热毁丝霉、粗糙链孢菌、产紫青霉、黄孢平革菌 (*Phanerochaete chrysosporium*)、射脉菌 (*Phlebia radiata*)、刺芹侧耳 (*Pleurotus eryngii*)、土生梭孢壳霉、长域毛栓菌 (*Trametes villosa*)、变色栓菌 (*Trametes versicolor*)、哈茨木霉、康宁木霉、长枝木霉、里氏木霉、或绿色木霉细胞。

[0429] 真菌细胞可以通过下述过程转化, 该过程涉及原生质体形成、原生质体的转化、以及以本身已知的方式的细胞壁的再生。用于转化曲霉属和木霉属宿主细胞的适合程序描述于以下文献中: EP 238023, Yelton 等人, 1984, Proc. Natl. Acad. Sci. USA [美国国家科学院院刊] 81:1470-1474 以及 Christensen 等人, 1988, Bio/Technology [生物/技术] 6:1419-1422。用于转化镰孢菌属物种的适合方法由 Malardier 等人 1989, Gene [基因] 78:147-156 和 WO 96/00787 描述。可以使用由如以下文献描述的程序转化酵母: Becker 和 Guarente, 在阿贝尔森, J.N. 和 Simon, M.I. 编辑, Guide to Yeast Genetics and Molecular Biology [酵

母遗传学与分子生物学指南], *Methods in Enzymology* [酶学方法], 第194卷, 第182-187页, 学术出版社有限公司, 纽约; Ito等人1983, *J. Bacteriol.* [细菌学杂志] 153:163; 以及 Hinnen等人, 1978, *Proc. Natl. Acad. Sci. USA* [美国国家科学院院刊] 75:1920。

[0430] 产生方法

[0431] 本发明还涉及产生变体的方法, 这些方法包括: (a) 在适于表达该变体的条件下培养本发明的宿主细胞; 并且 (b) 回收该变体。

[0432] 使用本领域已知的方法在适合于产生变体的营养培养基中培养宿主细胞。例如, 可以通过摇瓶培养, 或者在适合的培养基中并在允许变体表达和/或分离的条件下在实验室或工业发酵罐中进行小规模或大规模发酵 (包括连续发酵、分批发酵、分批给料发酵或固态发酵) 来培养细胞。培养是使用本领域已知的程序, 在适合营养培养基中发生, 该培养基包含碳和氮来源及无机盐。适合的培养基可以从商业供应商获得或可以根据公开的组成 (例如, 在美国典型培养物保藏中心 (American Type Culture Collection) 的目录中) 制备。如果变体被分泌到营养培养基中, 则变体可以直接从培养基中回收。如果变体没有分泌, 则它可以从细胞裂解液中回收。

[0433] 可以使用本领域已知的对这些变体特异的方法检测这些变体。这些检测方法包括但不限于, 特异性抗体的使用、酶产物的形成或酶底物的消失。例如, 可以使用酶测定来确定该变体的活性。

[0434] 可以使用本领域已知的方法来回收该变体。例如, 可以通过多种常规程序从该营养培养基中回收该变体, 这些常规程序包括但不限于: 收集、离心、过滤、提取、喷雾干燥、蒸发、或沉淀。

[0435] 可以通过本领域已知的多种程序来纯化变体以获得基本上纯的变体, 这些程序包括但不限于: 色谱法 (例如, 离子交换色谱、亲和色谱、疏水作用色谱、色谱聚焦、以及尺寸排阻色谱)、电泳程序 (例如, 制备型等电点聚焦)、差别溶解度 (例如, 硫酸铵沉淀)、SDS-PAGE、或提取 (参见例如, *Protein Purification*, Janson and Ryden, editors, VCH Publishers, New York, 1989 [蛋白质纯化, Janson和Ryden编辑, VCH出版社, 纽约, 1989])。

[0436] 在可替代的方面, 没有回收该变体, 而是将表达该变体的本发明的宿主细胞用作该变体的来源。

[0437] 纺织品

[0438] 如在此使用的, 术语“纺织品”是指纤维、纱线、织物、服装、和无纺布。该术语涵盖由天然、合成的 (例如, 制造的)、和各种天然和合成的共混物而制成的纺织品。纺织品可以是未加工的或加工的纤维、纱线、机织或针织织物、无纺布、和服装, 并且可以是使用各种材料制成的, 这些材料中的一些在此提及了。

[0439] 最有利地, 将本发明的工艺应用于含纤维素纺织品, 如棉、粘胶纤维、人造丝、苧麻、亚麻、天丝、或其混合物、或任何这些纤维的混合物、或任何这些纤维连同合成纤维的混合物, 如棉和氨纶 (弹力牛仔布) 的混合物。具体地, 该织物是染色的织物。在实施例中, 该织物是牛仔布。可以将该牛仔布织物用以下还原染料进行染色: 如靛蓝或靛蓝相关染料, 如硫靛蓝。

[0440] 在本发明的工艺的实施例中, 含纤维素纺织品是含棉纺织品或含人造纤维素的纺织品。

[0441] 纤维素酶变体的生物整理活性的测量

[0442] 为了确定纤维素酶变体的生物整理活性,通常使用已知技术对其进行纯化。

[0443] 如在此使用的术语“生物整理”是指使用纤维素酶处理纺织品,并且包括但不限于生物抛光和生物石磨。

[0444] 如在此使用的“生物整理活性”,特别是“生物抛光活性”如实例中所述进行确定。GH45纤维素酶的生物抛光效果可以通过从织物去除绒毛或绒毛小球(称为起球)的活性来测量。去起球可以表示为每单位蛋白质的去起球活性(即,特定的去起球活性)。

[0445] 根据本发明的一个实施例,测量生物整理活性的测定是针对生物抛光活性的起球记录测试。相对于亲本GH45纤维素酶的生物整理效果,本发明的GH45纤维素酶变体提供了对含棉纺织品的增强的生物整理。例如,GH45纤维素酶变体提供了比亲本GH45纤维素酶强约1%、约5%、约10%、约15%、约20%、约30%、约50%的生物抛光效果。在牛仔布磨损方面,GH45纤维素酶变体提供比亲本GH45纤维素酶强约1%、约5%、约10%、约15%、约20%、约30%、约50%的生物石磨效果。在起球记录测试中,GH45纤维素酶变体提供比亲本GH45纤维素酶的起球记录多出约0.1、约0.2、约0.5、约0.8、约1.0的起球记录。在改进工艺经济性方面,这种GH45纤维素酶变体的利用可能是一个进步。

[0446] 纺织品制造工艺

[0447] 将织物,如纤维材料的织物加工成准备用于服装制造的材料涉及若干步骤:将纤维纺成纱线;从纱线制造机织(woven)或针织(knit)织物;及随后的准备过程,染色/印染和修整(finishing)操作。在例如染色/印染和整理之前,对于去除来自纤维的天然的和人工诱导的杂质和用于改进其美学外观和可加工性而言准备过程是必需的。常见的准备过程包含退浆(用于机织物)、精练和漂白,这些过程产生适于染色或整理的织物。

[0448] 机织物是通过在织机上沿纵轴向方向伸展的经纱之间机织“纬线(filling)”或者“纬(weft)”纱而编制的。为了润滑和防止经纱在机织过程中纬纱高速插入时的打磨,经纱在机织之前必须上浆。常用的浆液试剂是淀粉(或淀粉衍生物以及改性淀粉)、聚(乙烯醇)、羧甲基纤维素(即CMC),其中淀粉占优势。浆液混合液中通常包含石蜡、丙烯酸粘合剂以及多种润滑剂。纬纱可以穿过经纱以“上一个-下一个(over one-under the next)”的方式纺织(平织),或以“上一个-下两个(over one-under two)”(斜纹织)或任何其它无数种排列。一般来说,裙装、衬衫、裤子、被单、毛巾、帷幔等都是由机织物制成的。当织物制成之后,必须再次去除织物上的浆料(即退浆)。

[0449] 针织法通过将连锁的纱线圈连接在一起而形成织物。与由两类纱线织成并含有许多“线头”的机织物相反,针织物是由单股连续的纱线织成的。从织法上讲,有许多种不同的方式将纱线圈结起来,而最终织物的特性则同时取决于纱线和针织的类型。内衣、针织衫(sweater)、袜子、运动衫、汗衫(sweat shirt)等衍生自针织织物。

[0450] 退浆

[0451] 退浆是将浆料从机织物中的经纱降解和/或去除。通常,通过酶退浆工序去除淀粉。此外,有时使用氧化退浆以及用酸或碱的化学退浆。

[0452] 在一些实施例中,该退浆酶是一种淀粉分解酶,如 α -淀粉酶、 β -淀粉酶、甘露聚糖酶、葡糖淀粉酶、或其组合。

[0453] 适合的 α 以及 β -淀粉酶包括细菌或真菌来源的那些,连同这类淀粉酶的化学或基

因修饰的突变体以及变体。适合的 α -淀粉酶包括从芽孢杆菌属物种可获得的 α -淀粉酶。合适的可商购的淀粉酶包括但不限于,OPTISIZE[®]NEXT、OPTISIZE[®]FLEX和OPTISIZE[®]COOL(全部来自于杰能科国际公司(Genencor International Inc.))、以及DURAMYL[™]、ERMAMYL[™]、FUNGAMYL[™]、TERMAMYL[™]、AQUAZYME[™]以及BAN[™](均可从诺维信公司,巴格斯瓦德(Bagsvaerd),丹麦获得)。

[0454] 其他合适的淀粉分解酶包括CGTase(环糊精葡聚糖转移酶,EC2.4.1.19),例如,从芽孢杆菌属、热厌氧杆菌属或嗜热厌氧杆菌属物种获得的那些。

[0455] 精练

[0456] 精练用于从纤维中去除杂质、用于膨润纤维并且用于去除棉籽壳。这是最关键的步骤之一。精练的主要目的是a)均匀清洁织物,b)软化籽屑及其他废料,c)改进织物吸水性,d)使脂肪、油以及蜡皂化并且溶解,以及e)使不成熟棉减到最少。在约沸腾温度下的氢氧化钠精练对于100%棉而言是被接受的处理,而氢氧化钙和碳酸钠使用频率较低。合成纤维在更温和的条件下进行精练。表面活性剂和螯合剂是碱性精练所必需的。已经引入酶精练,其中据报道纤维素酶、半纤维素酶、果胶酶、脂肪酶、以及蛋白酶均具有精练作用。

[0457] 漂白

[0458] 漂白是破坏有颜色的颜料和/或有色的杂质,以及棉籽壳屑去除。通过利用氧化或还原化学进行漂白。氧化剂可以进一步细分为使用或产生以下的那些:a)次氯酸盐(OCl⁻),b)二氧化氯(ClO₂),c)高锰酸盐(MnO₄⁻),d)臭氧,以及氢过氧化物的那些种类(OOH⁻和/或OOH)。还原剂是典型的二氧化硫、亚硫酸氢盐等。已经报道了使用葡萄糖氧化酶或过氧化物酶的酶促漂白(例如,参见WO 2013/040991)。传统地,在该工艺中使用过氧化氢。

[0459] 印染和染色

[0460] 通过任何适合用于将染料结合至纺织品中纤维的方法,通过将颜料施用至纺织品进行纺织品的印染和染色。纺织品的染色可以例如通过使织物通过颜料的浓溶液,随后将湿的织物存储在不漏气的外壳中以允许扩散时间,以及在冲洗掉未反应的颜料之前将颜料与织物基材反应而进行。可替代地,在冲洗之前,颜料可以通过随后对纺织品蒸汽蒸馏进行固定。这些染料包括合成的和天然的染料。典型的染料是具有阴离子官能团的那些(例如,酸性染料、直接染料、媒染染料和反应性染料),具有阳离子基团的那些(例如,碱性染料),在应用前需要化学反应的那些(例如,还原染料、硫黄染料和偶氮染料),分散染料和溶剂染料。

[0461] 在染色之后必须去除未与纤维结合的过量的可溶染料,以确保染色的纺织品的坚牢度并且防止消费者在纺织品的洗涤过程中发生不期望的染料转移。一般,需要大量的水用于彻底去除多余的染料。在常规工艺中,首先将印染或染色的纺织品用冷水冲洗,然后在高温下添加适合的添加剂洗涤来减少回染,像聚(乙烯吡咯烷酮)(PVP)。

[0462] 在WO 99/34054中披露了用洗涤液从染色的织物中移除过量颜料的酶工艺,该洗涤液包括至少一种过氧化物酶、氧化酶试剂以及至少一种介质,如包含过氧化物酶、过氧化氢酶以及像1-羟基-苯并三唑的介质的液体。

[0463] 生物抛光

[0464] 不应用整理组分的情况下,大多数棉织品和混纺棉织物具有相当硬和坚硬的手感问题。织物表面同样不光滑,因为小的有绒毛的微纤维从其中突出。另外,在相对短期的穿

着之后,起球发生于该织物表面上,从而给予它无吸引力的磨损的样子。

[0465] 生物抛光是 在纤维素织物制造过程中用酶(例如纤维素酶)对其进行处理的方法,该方法就“减少起球”而言改进了织物品质。生物抛光的最重要的效果可以由以下项表征:较少的起毛和起球,增加光彩/光泽(gloss/luster),改进织物手感,增加持久柔软性和/或改进吸水性。通常在制造针织和机织织物或服装的湿润过程中进行生物抛光。湿润过程包括例如以下步骤:退浆、精练、漂白、洗涤、染色/印染以及整理。可以在湿润步骤的任一步骤之后将生物抛光作为分开步骤而进行或可以与那些湿润步骤的任一步骤组合进行。

[0466] 在本发明中,生物整理步骤在退浆、漂白或印染和染色步骤之前、期间或之后进行。

[0467] 牛仔布织物的制造

[0468] 一些染色的织物(例如牛仔布织物)需要在机织之前将纱线染色。对于牛仔布织物,在机织之前,将经线例如用靛蓝染色并上浆。优选地,牛仔纱的染色是环染。本发明的一个优选实施例是用还原染料(例如靛蓝)、或靛蓝相关的染料(例如硫靛蓝)、或硫化染料、或直接染料、或活性染料或萘酚环染纱线。还可以用多于一种染料为纱线染色,例如首先用硫化染料并且然后用还原染料,或反之亦然。

[0469] 优选地,在将纱线染色之前,使其经历精练和/或漂白,以便使牛仔布织物达到较高的品质。通常,在机织为染色的织物(例如牛仔布)之后,染色的织物或服装进入退浆阶段,优选地接着是石磨或打磨步骤和/或颜色修饰步骤。

[0470] 如在此使用的退浆工艺与如正文中以上提及的工艺相同。

[0471] 在退浆后,染色的织物经历了生物石磨步骤。可以用酶类或浮石或两者进行生物石磨步骤。如在此使用的,术语“生物石磨”、“石洗”和“打磨”是可互换的,这意味着在含有一种机械磨擦剂(如浮石、研磨纤维素酶或这些的组合)的水性介质中搅动牛仔布,以便提供“石洗”外观。在所有情况下,需要机械作用以去除颜料,并且该处理通常在洗涤机(像鼓式洗涤机、圆形洗涤机(belly washer))中进行。由于颜料去除不均匀,染色的区域和已经去除颜料的区域之间存在对比度。用纤维素酶处理可以完全代替用浮石处理。然而,当希望产生重度打磨的抛光时,纤维素酶处理还可以与浮石处理结合。对于牛仔布制造,“生物整理”包括“生物石磨”。

[0472] 优选地,打磨后是颜色修饰步骤。如在此使用的,术语“颜色修饰”或“颜色调节”用于无差别地指由破坏、修饰、或去除与纺织品有关的着色剂造成的纺织品颜色的任何变化。不局限于理论,提出了颜色修饰起因于与纺织品材料相关的生色团的修饰,由此改变其外貌。这些生色团可以是天然与用于制造纺织品的材料相关(例如,棉的白色)或与特殊整理,如染色或印染相关。颜色修饰涵盖对生色团的化学修饰连同对生色团所附着的材料的化学修饰。

[0473] 在纺织品特别是牛仔布的某些地方,获得褪色或漂白的 外观在纺织品制造中是重要的一环。这通常通过在打磨步骤之后将 KMnO_4 (或 $\text{KMnO}_4/\text{H}_3\text{PO}_4$)溶液(通过刷涂、摩擦或喷雾)涂到干燥的牛仔布上来实现。染色区域在干燥并用 $\text{Na}_2\text{S}_2\text{O}_5$ 溶液洗涤后将被漂白。在这个过程中,靛蓝/硫化染料被 KMnO_4 通过氧化破坏,并且然后用 $\text{Na}_2\text{S}_2\text{O}_5$ 洗涤以去除由氧化产物引起的棕色。这种处理将造成局部颜色修饰,即牛仔布上的特定漂白图案以满足客户的需要。在本发明中,生物整理步骤在退浆或颜色修饰步骤之前、期间或之后进行。

[0474] 本发明还涉及一种用于生物整理的工艺,该工艺包括用本发明的变体处理含纤维素纺织品。

[0475] 本发明进一步涉及一种用于处理含纤维素纺织品的工艺,该工艺包括

[0476] (a) 退浆;

[0477] (b) 颜色修饰;

[0478] 其中在步骤(a)和步骤(b)之前、期间或之后添加本发明的变体。

[0479] 本发明进一步涉及一种用于处理含纤维素纺织品的工艺,该工艺包括

[0480] (a) 退浆;

[0481] (b) 精练;

[0482] (c) 漂白;

[0483] (d) 染色;

[0484] 其中在步骤(a)、(b)、(c)或(d)之前、期间或之后添加本发明的变体。

[0485] 在一个实施例中,纤维素酶在低机械作用下被处理或添加。低机械作用可通过将低搅动速度或低机械物质(例如橡胶球)施加到包含纤维素酶和含纤维素纺织品的系统上来进行。

[0486] 本发明涉及本发明的变体用于对含纤维素纺织品进行生物整理的用途。

[0487] 植物

[0488] 本发明还涉及植物,例如转基因植物、植物部分或植物细胞,其包括本发明的多核苷酸,以便以可回收的量表达和产生该变体。该变体可以从植物或植物部分回收。可替代地,包含该变体的植物或植物部分可以按原样用于改进食品或进料的质量,例如,改进营养价值、可口性以及流变特性,或用以破坏抗营养因子。

[0489] 转基因植物可以是双子叶的(双子叶植物)或单子叶的(单子叶植物)。单子叶植物的实例是草,如草甸草(蓝草,早熟禾属);饲草,如羊茅属(*Festuca*)、黑麦草属(*Lolium*);温带草,如翦股颖属(*Agrostis*);以及谷类,例如小麦、燕麦、黑麦、大麦、稻、高粱、以及玉蜀黍(玉米)。

[0490] 双子叶植物的实例是烟草、豆类(如羽扇豆(*lupins*)、马铃薯、糖甜菜(*sugar beet*)、豌豆、豆(*bean*)和大豆(*soybean*))、以及十字花科植物(十字花科(*family Brassicaceae*))(如花椰菜、油菜籽、以及紧密相关的模式生物拟南芥)。

[0491] 植物部分的实例是茎、愈伤组织、叶、根、果实、种子和块茎、以及包括这些部分的独立组织,例如,表皮、叶肉、薄壁组织(*parenchyme*)、维管组织、分生组织。特定植物细胞区室,如叶绿体、质外体(*apoplast*)、线粒体、液泡、过氧化物酶体以及细胞质也被认为是植物部分。此外,任何植物细胞,无论是何种组织来源,都被认为是植物部分。同样地,植物部分,如分离以有助于本发明的利用的特定组织和细胞也被认为是植物部分,例如胚、胚乳、糊粉和种皮。

[0492] 同样包括于本发明范围内的是此类植物、植物部分以及植物细胞的后代。

[0493] 表达变体的转基因植物或植物细胞可以根据本领域已知的方法来构建。简而言之,通过如下方法构建该植物或植物细胞:将编码变体的一个或多个表达构建体并入到植物宿主基因组或叶绿体基因组中,并且使所得的修饰植物或植物细胞繁殖为转基因植物或植物细胞。

[0494] 表达构建体宜为包括编码变体的多核苷酸的核酸构建体,该多核苷酸与在选择的植物或植物部分中表达该多核苷酸所需的适当的调节序列可操作地连接。而且,表达构建体可以包括用于鉴定整合了此表达构建体的植物细胞的选择性标记,和将此构建体引入所讨论的植物所必需的DNA序列(后者取决于所用的引入DNA的方法)。

[0495] 例如,基于希望在何时、何处、以及如何表达该变体来确定对调节序列如启动子和终止子序列和任选的信号或转运序列的选择。例如,编码变体的基因的表达可以是组成型的或诱导型的,或可以是发育、阶段或组织特异性的,并且可以使基因产物靶向特定组织或植物部分,如种子或叶。调节序列由例如塔格(Tague)等人,1988,植物生理学(Plant Physiology) 86:506描述。

[0496] 对于组成型表达,可以使用35S-CaMV、玉蜀黍泛素1、或稻肌动蛋白1启动子(Franck等人,1980,Cell[细胞]21:285-294;Christensen等人,1992,Plant Mol.Biol.[植物分子生物学]18:675-689;Zhang等人,1991,Plant Cell[植物细胞]3:1155-1165)。器官特异性启动子可以是以下各项的启动子,例如来自贮藏库组织(如种子、马铃薯块茎、和果实)(Edwards和Coruzzi,1990,Ann.Rev.Genet.[遗传学年鉴]24:275-303),或来自代谢库组织(如分生组织)(Ito等人,1994,Plant Mol.Biol.[植物分子生物学]24:863-878)的启动子,种子特异性启动子,如来自稻的谷蛋白、醇溶谷蛋白、球蛋白或白蛋白启动子(Wu等人,1998,Plant Cell Physiol.[植物与细胞生理学]39:885-889),来自豆球蛋白B4的蚕豆启动子和来自蚕豆的未知种子蛋白基因(Conrad等人,1998,J.Plant Physiol.[植物生理学杂志]152:708-711),来自种子油体蛋白的启动子(Chen等人,1998,Plant Cell Physiol.[植物与细胞生理学]39:935-941),来自欧洲油菜的贮藏蛋白napA启动子,或本领域已知的任何其他种子特异性启动子,例如,如在WO 91/14772中所描述的。此外,启动子可以是叶特异性启动子,如来自稻或番茄的rbcS启动子(Kyozuka等人,1993,Plant Physiol.[植物生理学]102:991-1000)、小球藻病毒腺嘌呤甲基转移酶基因启动子(Mitra和Higgins,1994,植物分子生物学[Plant Mol.Biol.]26:85-93)、来自稻的aldP基因启动子(Kagaya等人,1995,分子遗传学与基因组学[Mol.Gen.Genet.]248:668-674)、或伤口诱导型启动子(如马铃薯pin2启动子)(Xu等人,1993,植物分子生物学[Plant Mol.Biol.]22:573-588)。同样,该启动子可以通过如温度、干旱或盐度变化等非生物处理来诱导,或通过外源地施加使该启动子活化的物质(例如乙醇;雌激素;植物激素,如乙烯、脱落酸及赤霉素;及重金属)来诱导。

[0497] 启动子增强子元件也可以用于实现变体在植物中的较高表达。例如,启动子增强子元件可以是置于启动子与编码变体的多核苷酸之间的内含子。例如,Xu等人,1993,见上文,披露了使用稻肌动蛋白1基因的第一内含子以增强表达。

[0498] 该选择性标记基因及该表达构建体的任何其他部分可以选自本领域中可用的那些。

[0499] 可以根据本领域中已知的常规技术将核酸构建体结合到植物基因组中,这些常规技术包括农杆菌介导的转化、病毒介导的转化、微注射、粒子轰击、生物射弹转化、以及电穿孔(Gasser等人,1990,Science[科学]244:1293;Potrykus,1990,Bio/Technology[生物/技术]8:535;Shimamoto等人,1989,Nature[自然]338:274)。

[0500] 目前根癌土壤杆菌介导的基因转移是用于产生转基因双子叶植物(关于综述,请

参见Hooykas和Schilperoort,1992,Plant Mol.Biol.[植物分子生物学]19:15-38)并且用于转化单子叶植物的方法,尽管对于这些植物可以使用其他转化方法。用于产生转基因单子叶植物的方法是粒子(涂覆有转化DNA的微观金或钨粒子)轰击胚愈伤组织或发育中的胚(Christou,1992,植物杂志[Plant J.]2:275-281;Shimamoto,1994,生物技术当前述评[Curr.Opin.Biotechnol.]5:158-162;Vasil等人,1992,生物/技术[Bio/Technology]10:667-674)。用于转化单子叶植物的替代方法是基于天然质体转化,如由Omirulleh等人,1993,Plant Mol.Biol.[植物分子生物学]21:415-428所描述。另外的转化方法包括美国专利号6,395,966和7,151,204(两者都通过引用以其全文结合在此)中所描述的那些。

[0501] 在转化后,根据本领域熟知的方法选出已并入了表达构建体的转化体,并使其再生成为完整植物。通常设计转化程序用于通过如下方法在再生期间或在后续世代中选择性消除选择基因:例如,使用带有两个独立的T-DNA构建体的共转化或利用特异性重组酶位点特异性地切除选择基因。

[0502] 除用本发明的构建体直接转化具体植物基因型之外,还可以通过使具有构建体的植物与缺乏该构建体的第二植物杂交来产生转基因植物。例如,可以通过杂交将编码变体的构建体引入特定植物品种中,无需总是直接地转化该给定品种的植物。因此,本发明不仅涵盖了从根据本发明已经转化的细胞直接再生的植物,而且还涵盖了此类植物的后代。如在此所用的,后代可以是指根据本发明制备的亲本植物的任何代的后代。此类后代可以包括根据本发明制备的DNA构建体。杂交导致通过将起始系用供体植物系交叉授粉,将转基因引入植物系。此类步骤的非限制性实例描述于美国专利号7,151,204中。

[0503] 植物可以通过回交转化工艺生成。例如,植物包括被称为回交转化的基因型、种系、近交体、或杂交体的植物。

[0504] 可以使用遗传标记以协助本发明的一种或多种转基因从一个遗传背景渗入到另一个。标记协助的选择提供了相对于常规育种的优势,在于其可以用于避免由表型变异引起的错误。另外,遗传标记可以在具体杂交的个别后代中提供有关良种种质相对程度的数据。例如,当具有所希望性状并且另外具有非农艺学所希望的遗传背景的植物与良种亲本杂交时,可以使用遗传标记来选择不仅具有感兴趣的性状,还具有相对较大比例所希望种质的后代。以此方式,使一种或多种性状渗入特定遗传背景所需的世代数得以最小化。

[0505] 本发明亦涉及产生本发明的变体的方法,其包括:(a)在有助于产生该变体的条件下培养转基因植物或植物细胞,该植物或植物细胞包含编码具有变体的多核苷酸;和(b)回收该变体。

[0506] 通过以下实例进一步描述本发明,所述实例不应理解为对本发明范围的限制。

[0507] 实例

[0508] 材料

[0509] 用作缓冲液和底物的化学品是至少试剂级的商业产品。

[0510] 介质

[0511] 1)具有50mM乙酸盐的pH 5.0缓冲液:将2.873g乙酸钠和0.901g乙酸溶解于1L去离子水中;

[0512] 2)具有50mM磷酸盐的pH 6.5缓冲液:将5.642g磷酸氢二钠十二水合物($\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$)和5.344g磷酸二氢钠二水合物($\text{NaH}_2\text{PO}_4 \cdot 2\text{H}_2\text{O}$)溶解于1L去离子水中;

[0513] 3) 具有50mM磷酸盐的pH 7.5缓冲液:将15.045g磷酸氢二钠十二水合物 ($\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$) 和1.248g磷酸二氢钠二水合物 ($\text{NaH}_2\text{PO}_4 \cdot 2\text{H}_2\text{O}$) 溶解于1L去离子水中。

[0514] 酶

[0515]

酶	核心	接头	CBM	对 应 于 O42EB3 成熟 多肽的突变	SEQ ID NO
组 1					
O42EB3	Tt 纤维素酶	Sc 纤维素酶	Sc 纤维素酶		SEQ ID NO: 1 和 2
O32JRC	Tt 纤维素酶	Sc 纤维素酶	Tt 纤维素酶		SEQ ID NO: 3 和 4
O62FA6	Tt 纤维素酶	Sc 纤维素酶	Sc 纤维素酶	W292Y	SEQ ID NO: 5 和 6
O62FA8	Tt 纤维素酶	Th 纤维素酶	Th 纤维素酶	R202A;	SEQ ID NO: 7 和 8
O62FAA	Tt 纤维素酶	Th 纤维素酶	Th 纤维素酶	R202A ; Y292W	SEQ ID NO: 9 和 10
O52EWC	土生梭孢壳霉 GH45 纤维素酶 (Tt 纤维素酶),				SEQ ID NO: 37 和 12

[0516]

	参见 WO 2012/089024 中的 SEQ ID NO: 3 和 4, 通过引用特此结合				
组 2					
O82AH2	大孢圆孢霉 GH45 纤维素酶(Sc 纤维素酶), 参见在 WO 2005/054475 中的 SEQ ID NO 39 和 3, 通过引用特此结合				SEQ ID NO: 38 和 13
O82DT2	Sc 纤维素酶			W292Y	SEQ ID NO 15 和 16
O82DT4	Sc 纤维素酶	Tt 纤维素酶	Tt 纤维素酶	R202A	SEQ ID NO 17 和 18
组 3					
P24YEZ	四孢脉孢菌 GH45 纤维素酶 (Nt 纤维素酶) 参见 WO 2015/058700 中的 SEQ ID NO: 1 和 SEQ ID NO: 2, 通过引用特此结合				SEQ ID NO 39 和 14
P446XD	Nt 纤维素酶			Y292W	SEQ ID NO 19 和 20
组 4					
O14G83	Tt 纤维素酶	Sc 纤维素酶	Sc 纤维素酶	R265K W266Y W292Y	+SEQ ID NO 21 和 +22
O141FN	Tt 纤维素酶	Sc 纤维素酶	Sc 纤维素酶	W266Y F274Y W292Y	+SEQ ID NO 23 和 +24
O141FS	Tt 纤维素酶	Sc 纤维素酶	Sc 纤维素酶	T237N T246N W292Y	+SEQ ID NO 25 和 +26
O141FK	Tt 纤维素酶	Sc 纤维素酶	Sc 纤维素酶	F274Y W292Y	+SEQ ID NO 27 和 28
O141FU	Tt 纤维素酶	Sc 纤维素酶	Sc 纤维素酶	S255P W292Y	+SEQ ID NO 29 和 30
O141FF	Tt 纤维素酶	Sc 纤维素酶	Sc 纤维素酶	W266Y W292Y	+SEQ ID NO 31 和 32
O141FH	Tt 纤维素酶	Sc 纤维素酶	Sc 纤维素酶	R265K W292Y	+SEQ ID NO 33 和 34
O141FQ	Tt 纤维素酶	Sc 纤维素酶	Sc 纤维素酶	S221P S224P W292Y	+SEQ ID NO 35 和 +36

[0517] 织物

[0518] 双棉布(Cotton interlock):40S,漂白的,HM-A0008,可获得自中国广州的HM棉业公司(HM Cotton)。

[0519] 单面棉针织布(Cotton single jersey):40S/1AC+20D,漂白的,中国四川Migen spandex plain公司。

[0520] 基本和竹类牛仔织物和蓝色/棕色针织织物从中国广东购买。

[0521] 用活性染料染色的海军蓝织物购自广东溢达集团[Esquel]。

[0522] 方法

[0523] 重量损失确定

[0524] 将小块布样在编号之前置于控制室(65%±5%湿度,20℃±1℃)中24小时,通过分析天平(用于100g以下的样品)或精密天平(用于100g以上的样品)称重并记录。在处理之后,将所有样品滚动干燥1小时并在如以上提及的控制室中调节24小时。对于每份样品,重量损失定义如下:

[0525]

$$\frac{\text{重量损失}}{\text{处理之前的重量}} = \frac{\text{处理之前的重量} - \text{处理之后的重量}}{\text{处理之前的重量}} \times 100$$

[0526] 起球记录(pilling notes)测试

[0527] 将已经在标准气候(65%湿度,21℃)下预调节至少24小时的织物(包括处理的和未处理的)针对起球记录用Nu-马丁代尔(Martindale)测试仪(詹姆斯(James)H.希尔有限公司(Heal Co.Ltd),英格兰)进行测试,用相同类型的未处理的织物作为磨损的织物。在2000回转之后进行标准的起球测试(瑞士标准(SN)198525),通过从具有如下的定义的含义的1-5标记,其中1显示不良的抗起球而5显示优异的抗起球特性。因此,马丁代尔起球记录得分越高,该内切葡聚糖酶生物抛光处理越有效。

[0528] 记录5:没有起球

[0529] 记录4:轻微起球

[0530] 记录3:中度起球

[0531] 记录2:明显起球

[0532] 记录1:严重起球

[0533] 允许1/2、1/4记录

[0534] 为了使结果更可靠,对每个样品由不同的人进行3个分开读数,并且这3个读数的平均值被采用作为起球记录的最终结果。

[0535] 强度测定

[0536] 按照GB/T19976-2005用YGB031PC电子大理石破裂强度试验机测试织物强度。包括五个平行测试,并且平均值是以每次处理的织物强度计算的。

[0537] 用于牛仔布的量色法

[0538] 通过用预校准的DataColor SF450X(可替代地,可以使用等效设备)测量反射比而确定牛仔布样品的磨损水平和返染水平。每个样品取四个读数,并使用这些读数的平均值。用指数CIE L*在样品的蓝面(正面)上评估磨损水平,并且用指数CIE b*在样品的背面上评估返染水平。

[0539] L*指示从0至100等级的白/黑变化,并且L*的降低意味着黑色的增加(白色的减

少),并且L*的增加意味着白色的增加(黑色的减少)。 δL^* 单位=用某一纤维素酶处理的小块布样的L*-在纤维素酶处理之前的小块布样的L*。 δL^* 单位越大,牛仔布磨损水平越高,例如4的 δL^* 单位具有比3的 δL^* 单位高的磨损水平。

[0540] b*指示蓝/黄变化,并且b*的降低意味着蓝色的增加(黄色的减少),并且b*的增加意味着黄色的增加(蓝色的减少)。 δb^* 单位=用某一纤维素酶处理的小块布样的b*-在纤维素酶处理之前的小块布样的b*。 δb^* 单位越大,对应的返染水平越低,例如-1.5的 δb^* 单位具有比-2.5的 δb^* 单位低的返染水平。

[0541] 蛋白质含量

[0542] 根据产品手册,用BCA™蛋白测定试剂盒(BCA™ Protein Assay Kit)(产品编号23225,可商购自赛默飞世尔科技公司(Thermo Fisher Scientific Inc.))测量酶产品中的酶蛋白。

[0543] 纤维素酶活性测定(CNUR/g)

[0544] 将底物羧甲基纤维素(CMC)在pH 7.5,50℃下用纤维素酶水解30分钟。反应由含有PAHBAH和铋的碱性试剂终止,该碱性试剂与还原糖形成复合物。复合物形成导致颜色产生,其可以通过分光光度计在405nm处读取。所产生的颜色与纤维素酶活性成比例。在Konelab分析仪中自动进行酶反应和吸光度测量。相对于诺维信酶标准品确定纤维素酶活性。该测定法的详细描述以及Renozyme™标准品的样品可从丹麦DK-2880巴格斯瓦德(Bagsvaerd)克罗格休约尔(Krogshoejvej)36的诺维信公司(Novozymes A/S)(EB-SM-0787.02-D)索取获得。

[0545] 实例1:通过将不同接头/CBM附接至Tt_Ce145a(CBM+)核心来构建纤维素酶杂合基因

[0546] 纤维素酶杂合基因RenoCorenew_ScLNK_ScCBM(SEQ ID NO:11&2)是通过以下方式产生:将来自大孢圆孢霉(*Staphylotrichum coccosporum*)GH45 SWISSPROT:B5BNY1的接头和CBM附接至Tt_Ce145a(CBM+)的核心区域的C-末端(WO 2012/089024,通过引用特此结合)。使用WO 2012/089024中描述的pRenoCBD质粒作为扩增杂合基因核心区域的模板。合成了大孢圆孢霉GH45SWISSPROT:B5BNY1的接头和CBM区域,并将其由金唯智公司(GENEWIZ Inc.)克隆在ScLNKCBM_SynA0(SEQ ID NO:40)中。

[0547] 以下引物由英杰公司(Invitrogen)合成。

[0548]

引物 1	ACACA ACTGGGGATC <u>CACC</u> atgcgctctactcccgttcttc	SEQ ID NO: 50
引物 2	<i>gctgcaagcgc aacgaecgatggtaacttcctgtgttcacc</i>	SEQ ID NO: 41
引物 3	<i>ggfgaacacaggggaagttaccatc gtcgttgcgcttgcagc</i>	SEQ ID NO: 42
引物 4	CCCTCTAGATCTCGAG <u>tca</u> <i>gagacactgggagtaccagtgcttc</i>	SEQ ID NO: 43

[0549] 引物1中的大写字母表示与pCaHj505(WO 2013/029496,通过引用特此结合)的插入位点同源的区域。带下划线的字母表示作为翻译过程起始物的Kozak序列。小写字母表示Tt_Ce145a(CBM+)的编码序列。在引物2中,常规字母表示Tt_Ce145a(CBM+)的核心序列,而斜体字母表示大孢圆孢霉GH45接头的序列。引物3含有引物2的互补序列。引物4中的大写字

母表示与pCaHj505的插入位点同源的区域。小写字母表示大孢圆孢霉GH45 CBM的互补序列。粗体字母和下划线字母是终止密码子。

[0550] 进行剪接重叠延伸PCR以产生杂合基因。在第一轮PCR中,核心区域和接头/CBM区域被独立扩增,详述如下。

[0551] 针对核心区域扩增,将各20pmol的引物1和3用于PCR反应,该PCR反应由以下项构成:在50 μ l的终体积中2 μ l的pRenoCBD质粒DNA,10 μ l的5X**Phusion**[®]HF缓冲液(Finnzymes Oy公司,艾斯堡,芬兰),1.5 μ l的DMSO,各1.5 μ l的2.5mM dATP、dTTP、dGTP和dCTP,和0.6单位的PHUSION[™]高保真度DNA聚合酶(Finnzymes Oy公司,艾斯堡,芬兰)。扩增使用Peltier Thermal Cycler(MJ Research Inc.,South San Francisco,CA,USA)进行,其程序为:在98 $^{\circ}$ C变性1分钟;10个循环,每循环在98 $^{\circ}$ C变性30秒,在65 $^{\circ}$ C退火30秒,其中每个循环降低1 $^{\circ}$ C,并在72 $^{\circ}$ C延伸70秒;25个循环,各自为在98 $^{\circ}$ C下持续30秒,在60 $^{\circ}$ C下持续30秒,并且在72 $^{\circ}$ C下持续70秒;和在72 $^{\circ}$ C最终延伸7分钟。加热块然后进入4 $^{\circ}$ C浸泡循环。通过1.0%的琼脂糖凝胶电泳,使用90mM Tris-硼酸盐及1mM EDTA(TBE)缓冲液来分离PCR产物,其中在UV光下观测来自反应的约0.8kb的单一产物条带。然后,根据制造商的说明书,通过使用**illustra**[™]**GFX**[™]PCR DNA和凝胶条带纯化试剂盒(GE医疗(GE Healthcare),白金汉郡,英国)从溶液中纯化出PCR产物。其被指定为片段1。

[0552] 对于接头和CBM区域扩增,将各20pmol的引物2和4用于PCR反应中,该PCR反应由以下项构成:在50 μ l的终体积中2 μ l的ScLNKCBM_SynA0,10 μ l的5X**Phusion**[®]HF缓冲液,1.5 μ l的DMSO,各1.5 μ l的2.5mM dATP、dTTP、dGTP、和dCTP,和0.6单位的PHUSION[™]高保真度DNA聚合酶。使用了与片段1扩增相同的程序。纯化的PCR产物约为0.3kb,并且被指定为片段2。

[0553] 在第二轮PCR中,将20pmol的引物1和4用于PCR反应中,该PCR反应由以下项构成:在50 μ l的终体积中1 μ l的片段1和1 μ l的片段2,10 μ l的5X**Phusion**[®]HF缓冲液,1.5 μ l的DMSO,各1.5 μ l的2.5mM的dATP、dTTP、dGTP和dCTP,和0.6单位的PHUSION[™]高保真度DNA聚合酶。使用了与片段1扩增相同的程序。获得约1.1kb的PCR产物。将纯化的PCR产物指定为片段3。

[0554] 将质粒pCaHj505用Bam HI和Xho I消化,使用TBE缓冲液通过1.0%琼脂糖凝胶电泳来分离,并且根据制造商的说明书,使用**ILLUSTRATE**[™]**GFX**[™]PCR DNA和凝胶条带纯化试剂盒进行纯化。使用**In-Fusion**[®]HD克隆试剂盒(克罗泰克实验有限公司(Clontech Laboratories, Inc.),山景城,加利福尼亚,美国)将片段3直接克隆至表达载体pCaHj505中,无需限制性消化。

[0555] 详细来说,将用Bam HI和Xho I消化的1 μ l的50ng/ μ l的pCaHj505和3 μ l的ddH₂O添加至In-Fusion Dry-Down Mix以获得溶解的球粒。然后将2 μ l的In-Fusion Dry-Down Mix溶液添加至新试管中并且添加3 μ l的片段3。将该反应在50 $^{\circ}$ C处孵育15分钟。使用连接反应物转化大肠杆菌TOP10感受态细胞(天根生物科技有限公司(TIANGEN Biotech Co.Ltd.),北京,中国)。获得三种转化体并发走用于序列确认。将插入在p505-RenoCorenew_ScLNK_ScCBM中的纤维素酶杂合基因编码序列RenoCorenew_ScLNK_ScCBM通过使用3730XL DNA分析仪(应用生物系统公司(Applied Biosystems Inc),福斯特市(Foster City),加利福尼亚州,美国)进行DNA测序来确认。因此,纤维素酶杂合基因的转录处于米曲霉 α -淀粉酶基因启动子的控制之下。使用**QIAPREP**[®]Spin Miniprep试剂盒(凯杰公司(QIAGEN GmbH),希

尔登,德国)来制备p505-RenoCorenew_ScLNK_ScCBM的质粒。

[0556] 第二纤维素酶杂合基因,RenoCorenew_ScLNK_RenoCBM (SEQ ID NO:3和4) 通过用衍生自Tt_Cel45a (CBM+) 的CBM替代RenoCorenew_ScLNK_ScCBM的CBM区域来产生。使用上述p505-RenoCorenew_ScLNK_ScCBM的质粒作为用于扩增核心和接头区域的模板。质粒pRenoCBD被用作扩增CBM区域的模板。

[0557] 以下引物由英杰公司 (Invitrogen) 合成。

[0558]

引物5	cgctcgtccaccggagga ggetgcacgtctcagaagtg	SEQ ID NO:44
引物6	cacttctgagacgtgcagcc tctcctcggtggacgacg	SEQ ID NO:45
引物7	CCCTCTAGATCTCGAG ggaatcaaccagcagtcgc	SEQ ID NO:46

[0559] 在引物5中,常规字母表示RenoCorenew_ScLNK_ScCBM的接头区域的序列,并且斜体字母表示Tt_Cel45a (CBM+) 的CBM区域的序列。引物6含有引物5的互补序列。引物7中的大写字母表示与pCaHj505的插入位点同源的区域。斜体小写字母与Tt_Cel45a (CBM+) CDS的下游序列互补。

[0560] 进行剪接重叠延伸PCR以产生杂合基因。

[0561] 对于核心和接头区域扩增,引物1和6被用作引物对,并且p505-RenoCorenew_ScLNK_ScCBM用作模板以获得PCR产物片段4 (约1kb)。PCR程序如下:在98°C变性持续1分钟;8个循环,每循环在98°C变性30秒,在67°C退火30秒,其中每循环降低1°C,并在72°C延伸60秒;25个循环,各自为在98°C下持续30秒,在60°C下持续30秒,并且在72°C下持续60秒;和在72°C最终延伸7分钟。

[0562] 对于CBM区域,使用引物5和7作为引物对,并使用pRenoCBD作为模板以获得PCR产物片段5 (约0.3kb)。使用了与片段1扩增相同的程序。

[0563] 然后通过使用引物1和7作为引物对并且使用片段4和5作为模板,制备与RenoCorenew_ScLNK_ScCBM的第二轮PCR中描述的类似的反应。使用与片段4扩增相同的程序。得到的PCR片段约为1.3kb,并且被指定为片段6。

[0564] 用In-Fusion[®] HD克隆试剂盒将片段6连接到载体pCaHj505上。大肠杆菌TOP10中的转化导致几种转化体。含有靶基因的转化体通过DNA测序确认并被指定为p505-RenoCorenew_ScLNK_RenoCBM)。在实例2中制备质粒用于米曲霉转化。

[0565] RenoCorenew_ScLNK_ScCBM的CBM区域的编码序列进一步被优化为RenoCorenew_ScLNK_ScCBM_2 (SEQ ID NO:1) 用于米曲霉表达。ScCBM_2 (SEQ ID NO:79) 在金唯智公司中合成。

[0566]

引物8	cgctcgtccaccggaggaggctgcgcagcgcagcgcg	SEQ ID NO:47
引物9	tctcctcggtggacgacg	SEQ ID NO:48
引物10	CCCTCTAGATCTCGAG tcagaggcactgcgagtaccagtcg	SEQ ID NO:49

[0567] 引物8中的常规字母表示RenoCorenew_ScLNK_ScCBM的接头区域,并且斜体字母表示ScCBM_2的5'末端。引物9含有接头区域的互补序列。引物10中的大写字母表示与pCaHj505的插入位点同源的区域,而斜体字母表示ScCBM_2的3'末端。

[0568] 使用p505-RenoCorenew_ScLNK_RenoCBM的质粒DNA以通过使用引物1和9作为引物

对来扩增核心和接头区域。通过使用引物8和10作为引物对,使用ScCBM_2来扩增CBM区域。通过使用引物1和10作为引物对,将所得PCR片段、片段7和片段8用于RenoCorenew_ScLNK_ScCBM_2扩增。将所得PCR产物片段9连接至pCaHj505。通过DNA测序确认含有靶基因的大肠杆菌转化体。在实例2中制备质粒DNA p505-RenoCorenew_ScLNK_ScCBM_2.用于米曲霉转化。

[0569] 实例2:纤维素酶杂合基因在米曲霉中的表达

[0570] 使用米曲霉菌株MT3568以异源表达编码纤维素酶杂合基因RenoCorenew_ScLNK_ScCBM_2和RenoCorenew_ScLNK_RenoCBM的基因。米曲霉MT3568是米曲霉JaL355的amdS(乙酰胺酶)破坏的基因衍生物(WO 02/40694),其中通过用pyrG基因破坏米曲霉乙酰胺酶(amdS)基因来恢复pyrG营养缺陷型。

[0571] 根据如US 20140179588 A1的实例2中“曲霉表达宿主的转化”所描述的方法制备原生质体。使用三μg质粒DNA转化米曲霉MT3568。

[0572] 用p505-RenoCorenew_ScLNK_ScCBM_2转化米曲霉MT3568或将米曲霉MT3568转化为p505-RenoCorenew_ScLNK_RenoCBM每个产生约20个转化体。分离每种转化的3-4个转化体以进行铺板重新分离,并且然后分别接种到24孔板中的3ml YPM培养基中,并在30°C、150rpm孵育。3天的孵育之后,根据制造商的说明书,在具有50mM MES的NUPAGE®NOVEX®4%-12%Bis-Tris凝胶(英杰公司,卡尔斯巴德,加利福尼亚州,美国)上对来自每一培养物的20μl上清液进行分析。将所得的凝胶用INSTANTBLUE™(艾本德有限公司(Expedeon Ltd.),巴布拉汉,剑桥,英国)染色。培养物的SDS-PAGE图谱显示大多数转化体针对两种基因具有大约55kDa的主要条带。因此,选择RenoCorenew_ScLNK_ScCBM_2的转化体#2和RenoCorenew_ScLNK_RenoCBM的转化体#1用于摇瓶发酵并分别指定为042EB3和032JRC。

[0573] 实例3:米曲霉表达菌株042EB3和032JRC的发酵

[0574] 将表达菌株042EB3的斜面用10ml YPM洗涤并接种到四个各含有400ml YPM培养基(1%酵母提取物、2%蛋白胨和2%麦芽糖)的2升烧瓶中,在30°C,80rpm振荡。在第3天时采集该培养物,并且使用0.45μmDURAPORE膜(美国马萨诸塞州贝德福德的密理博公司(Millipore, Bedford, MA, USA))进行过滤。

[0575] 将表达菌株032JRC的斜面用10ml YPM洗涤并接种到四个各含有400ml YPM培养基的2升烧瓶中,在30°C、80rpm振荡。在第3天时采集该培养物,并且使用0.45μm DURAPORE膜进行过滤。

[0576] 实例4:来自米曲霉042EB3和032JRC的纤维素酶杂合物的纯化和活性测定

[0577] 将1600ml体积的米曲霉042EB3(实例3)的经过滤的上清液用硫酸铵(80%饱和)沉淀,再溶解于50ml的20mM Tris-HCl(pH 7.5)中,用同一缓冲液透析,并且通过0.45μm过滤器过滤。最终体积是60ml。将该溶液施加至用20mM Tris-HCl(pH 7.5)平衡的40ml Q SEPHAROSE®快流柱(GE医疗,白金汉郡,英国)上。使用0-0.25M NaCl的线性梯度NaCl对这些蛋白质进行洗脱。收集未结合的蛋白质并通过使用用20mM Tris-HCl(pH 7.5)平衡的MonoQ HR16/10柱(GE医疗,白金汉郡,英国)进一步纯化。使用0-0.15M的线性梯度NaCl对蛋白质进行洗脱。收集未结合的蛋白质,并通过使用用20mM NaAc(pH 4.5)平衡的40ml SP SEPHAROSE®快流柱(GE医疗,白金汉郡,英国)进一步纯化。使用0-0.25M NaCl的线性

梯度NaCl对这些蛋白质进行洗脱。使用具有50mM MES的NUPAGE® NOVEX®4%-12% Bis-Tris凝胶通过SDS-PAGE分析各部分。将所得凝胶用INSTANTBLUE™染色。合并含有大约55kDa的条带的部分。然后将合并的溶液通过超滤进行浓缩。

[0578] 将1600ml体积的米曲霉032JRC的经过滤的上清液用硫酸铵(80%饱和)沉淀,再溶解于50ml的20mM Tris-HCl(pH 7.5)中,用同一缓冲液透析,并且通过0.45μm过滤器过滤。最终体积是75ml。将该溶液施加到用20mM Tris-HCl(pH 7.5)平衡的40ml Avicel pH-101柱上。将蛋白质用1%三乙胺洗脱。收集洗脱的级分,并使用用20mM Tris-HCl(pH 7.5)平衡的40ml QSEPHAROSE®快流柱(GE医疗,白金汉郡,英国)进一步纯化。使用0-0.5M NaCl的线性梯度NaCl洗脱这些蛋白质。使用具有50mM MES的NUPAGE® NOVEX®4%-12% Bis-Tris凝胶通过SDS-PAGE分析各部分。将所得凝胶用INSTANTBLUE™染色。合并含有大约55kDa的条带的部分。然后将合并的溶液通过超滤进行浓缩。

[0579] 通过蓝色底物检测纤维素酶活性。将AZCL-HE-纤维素(麦格酶公司(Megazyme))的蓝色底物溶于pH 6.0的100mM Bis-Tris中,终浓度为0.2% (w/v)。将20μl酶与100μl蓝色底物混合,并在50℃用Thermomixer(艾本德公司)以600rpm孵育30分钟,然后将60μl的混合物上清液转移到新孔中以测量595nm处的吸光度。纤维素酶杂合物示出良好的纤维素酶活性。

[0580] 实例5:062FA6 GH45纤维素酶的克隆

[0581] 选择GH45纤维素酶基因,062FA6(DNA序列为SEQ ID NO:5以及推导的氨基酸序列为SEQ ID NO:6)用于表达克隆。

[0582] 设计下文所示的寡核苷酸引物以从质粒p505-RenoCorenew_ScLNK_ScCBM_2扩增全长GH45纤维素酶基因,其对应于042EB3成熟多肽具有单个取代W292Y。这些引物是由中国北京的英杰公司(Invitrogen)合成的。

[0583] 正向引物D24H9NF(基因特异性):

[0584] 5' ACACAAGTGGGGATC CACC atgcgctctactcccgttcttc 3' (SEQ ID NO:50)

[0585] 反向引物D24H9NR(致突变的):

[0586] 5' CCCTCTAGATCTCGAG tcagaggcactgcgagtagtagtc 3' (SEQ ID NO:51)

[0587] 正向引物中的小写字符字母表示基因的编码区并且反向引物中的小写字符字母表示基因的侧翼区域,而粗体字符字母表示与pCaHj505(WO 2013/029496)的插入位置同源的区域。正向引物D24H9NF中的编码序列之前的4个字母表示作为翻译过程起始物的Kozak序列。斜体字母表示突变的氨基酸的核苷酸。

[0588] 将二十皮摩尔的上述正向引物和反向引物用于PCR反应中,该PCR反应由以下项构成:在50μl的终体积中0.2μl的p505-RenoCorenew_ScLNK_ScCBM_2质粒DNA,10μl的5X Phusion® HF缓冲液(Finnzymes Oy公司,艾斯堡,芬兰),1.5μl的DMSO,各1.5μl的2.5mM的dATP、dTTP、dGTP和dCTP,和0.6单位的Phusion®高保真度DNA聚合酶(Finnzymes Oy公司,艾斯堡,芬兰)。扩增使用佩尔特尔热循环仪(Peltier Thermal Cycler)(MJ研究有限公司(MJ Research Inc.),南旧金山,加利福尼亚州,美国)进行,其程序为:在98℃变性1分钟;10个循环,每循环在98℃变性30秒,在65℃退火30秒,其中每循环降低1℃,以及在72℃延伸90秒;25个循环,各自为在98℃下持续30秒,在56℃下持续30秒,并且在72℃下持续90秒;和在72℃最终延伸7分钟。加热块然后进入4℃浸泡循环。

[0589] 通过使用90mM Tris-硼酸盐和1mM EDTA(TBE)缓冲液进行1.0%琼脂糖凝胶电泳

分离PCR产物,其中在UV光下观察到反应中约1.1kb的单一产物条带,并且然后使用ILLUSTRATMTMGFXTMPCR DNA和凝胶条带纯化试剂盒(GE医疗,白金汉郡,英国)根据制造商的说明书进行纯化。

[0590] 将质粒pCaHj505用Bam HI和Xho I消化,使用TBE缓冲液通过1.0%琼脂糖凝胶电泳来分离,并且根据制造商的说明书,使用ILLUSTRATMTMGFXTMPCR DNA和凝胶条带纯化试剂盒(GE医疗,白金汉郡,英国)进行纯化。

[0591] 使用In-Fusion HD克隆试剂盒(克罗泰克实验有限公司(Clontech Laboratories, Inc.), 山景城,加利福尼亚州,美国)将PCR片段直接克隆至表达载体pCaHj505中,无需限制性消化和连接。根据制造商的说明书,使用In-Fusion HD克隆试剂盒将纯化的PCR产物和消化的载体连接在一起,得到质粒,其中O62FA6GH45纤维素酶多肽编码序列的转录处于米曲霉 α -淀粉酶基因启动子的控制之下。简言之,将1 μ l用Bam HI和Xho I消化的30ng/ μ l的pCaHj505和3 μ l的含约60ng O62FA6GH45纤维素酶片段的纯化PCR产物添加到1 μ l的5X In-Fusion HD酶预混物中。将反应在50 $^{\circ}$ C孵育15分钟,并且然后使用5 μ l的连接反应物转化大肠杆菌TOP10感受态细胞(天根生物科技有限公司,北京,中国)。通过菌落PCR检测含有表达构建体的大肠杆菌转化体。菌落PCR是一种用于直接从大肠杆菌菌落快速筛选质粒插入片段的方法。简言之,将一个单一菌落转移至PCR管中的预混合的PCR溶液中,该PCR溶液包括PCR缓冲液、MgCl₂、dNTPs、以及从其中产生PCR片段的引物对。筛选若干个菌落。在PCR后,将反应通过1.0%琼脂糖凝胶电泳使用TBE缓冲液进行分析。使用QIAPREP[®] Spin Miniprep试剂盒(凯杰公司(QIAGEN GmbH), 希尔登,德国)从示出具有预期大小的插入物的菌落制备质粒DNA。插入到pCaHj505中的O62FA6GH45纤维素酶编码序列通过使用3730XL DNA分析仪(应用生物系统公司(Applied Biosystems Inc), 福斯特市,加利福尼亚州,美国)进行DNA测序来确认。

[0592] O62FA6突变体多肽编码序列的DNA序列和推导的氨基酸序列分别在SEQ ID NO:5和SEQ ID NO:6中示出。编码序列是1104bp(包括终止密码子),它被78bp(核苷酸90至167)和69bp(核苷酸417至485)的2个内含子间断。编码蛋白是具有21个残基的信号肽的318个氨基酸。成熟蛋白包含297个氨基酸,具有30kDa的预测分子量和5.0的预测等电点。

[0593] 实例6:O62FA8和O62FAA GH45纤维素酶的克隆

[0594] 选择GH45纤维素酶基因O62FA8(DNA序列为SEQ ID NO:7以及推导的氨基酸序列为SEQ ID NO:8)和O62FAA(DNA序列为SEQ ID NO:9以及推导出的氨基酸序列为SEQ ID NO:10)用于表达克隆。

[0595] 对于O62FA8,设计下文所示的寡核苷酸引物以分开从质粒p505-RenoCorenew_ScLNK_ScCBM_2扩增GH45纤维素酶核心,并从质粒pGH45_Thihy3331(WO 2014/101753,特此通过引用结合)扩增接头/CBM结构域;对应于O42EB3成熟多肽具有单个取代R202A。这些引物是由中国北京的英杰公司(Invitrogen)合成的。

[0596] 正向引物D24H9NF:5' ACACAAGTGGGGATC CACC atgcgctctactcccgttcttc 3' (SEQ ID NO:50)

[0597] 反向引物D24H9TR:5' agctggagtcgctcgtt cgccttgcagccgg 3' (SEQ ID NO:52)

[0598] 正向引物D24H9TlinkerF:5' ctccggctgcaaggcg aacgacgactccagcttccc 3' (SEQ ID NO:53)

[0599] 反向引物D24H9TlinkerR:5' CCCTCTAGATCTCGAG ttagaggcactgcgagtagt agtcg 3' (SEQ ID NO:54)

[0600] 大写字母是与pCaHj505(WO 2013/029496)的插入位点同源的区域。小写字母表示土生梭孢壳霉GH45纤维素酶(Tt纤维素酶)核心的编码区域,并且粗体字符是赫坎梭孢壳霉GH45内切葡聚糖酶(Th纤维素酶)接头/CBM结构域的编码区域。正向引物D24H9NF中的编码序列之前的4个字母表示作为翻译过程起始物的Kozak序列。

[0601] 使用正向引物D24H9NF和反向引物D24H9TR来扩增来自042EB3的GH45纤维素酶核心。将二十皮摩尔的上述正向引物和反向引物用于PCR反应中,该PCR反应由以下项构成:在50 μ l的终体积中0.2 μ l的p505-RenoCorenew_ScLNK_ScCBM_2质粒DNA,10 μ l的5X **Phusion**[®] HF缓冲液(Finnzymes Oy公司,艾斯堡,芬兰),1.5 μ l的DMSO,各1.5 μ l的2.5mM的dATP、dTTP、dGTP和dCTP,和0.6单位的**Phusion**[®]高保真度DNA聚合酶(Finnzymes Oy公司,艾斯堡,芬兰)。扩增使用佩尔特热循环仪(Peltier Thermal Cycler)(MJ研究有限公司(MJ Research Inc.),南旧金山,加利福尼亚州,美国)进行,其程序为:在98 $^{\circ}$ C变性1分钟;10个循环,每循环在98 $^{\circ}$ C变性30秒,在65 $^{\circ}$ C退火30秒,其中每循环降低1 $^{\circ}$ C,以及在72 $^{\circ}$ C延伸90秒;25个循环,各自为在98 $^{\circ}$ C下持续30秒,在56 $^{\circ}$ C下持续30秒,并且在72 $^{\circ}$ C下持续90秒;和在72 $^{\circ}$ C最终延伸7分钟。加热块然后进入4 $^{\circ}$ C浸泡循环。

[0602] 通过使用90mM Tris-硼酸盐和1mM EDTA(TBE)缓冲液的1.0%琼脂糖凝胶电泳分离PCR产物,其中在UV光下观察到反应中约816bp的单一产物条带,并且然后使用ILLUSTRA[™]GFX[™]PCR DNA和凝胶条带纯化试剂盒(GE医疗,白金汉郡,英国)根据制造商的说明书进行纯化。PCR产物条带被指定为F3。

[0603] 使用正向引物D24H9T接头F和反向引物D24H9T接头R从Th纤维素酶扩增接头/CBM结构域。将二十皮摩尔的上述正向引物和反向引物用于PCR反应中,该PCR反应由以下项构成:在50 μ l的终体积中0.2 μ l的pGH45_Thihy3331质粒DNA,10 μ l的5X **Phusion**[®] HF缓冲液(Finnzymes Oy公司,艾斯堡,芬兰),1.5 μ l的DMSO,各1.5 μ l的2.5mM的dATP、dTTP、dGTP和dCTP,和0.6单位的**Phusion**[®]高保真度DNA聚合酶(Finnzymes Oy公司,艾斯堡,芬兰)。扩增使用佩尔特热循环仪(Peltier Thermal Cycler)(MJ研究有限公司(MJ Research Inc.),南旧金山,加利福尼亚州,美国)进行,其程序为:在98 $^{\circ}$ C变性1分钟;10个循环,每循环在98 $^{\circ}$ C变性30秒,在65 $^{\circ}$ C退火30秒,其中每循环降低1 $^{\circ}$ C,以及在72 $^{\circ}$ C延伸90秒;25个循环,各自为在98 $^{\circ}$ C下持续30秒,在56 $^{\circ}$ C下持续30秒,并且在72 $^{\circ}$ C下持续90秒;和在72 $^{\circ}$ C最终延伸7分钟。加热块然后进入4 $^{\circ}$ C浸泡循环。

[0604] 通过使用90mM Tris-硼酸盐和1mM EDTA(TBE)缓冲液的1.0%琼脂糖凝胶电泳分离PCR产物,其中在UV光下观察到反应中约255bp的单一产物条带,并且然后使用ILLUSTRA[™]GFX[™]PCR DNA和凝胶条带纯化试剂盒(GE医疗,白金汉郡,英国)根据制造商的说明书进行纯化。PCR产物条带被指定为F4。

[0605] 将质粒pCaHj505用Bam HI和Xho I消化,使用TBE缓冲液通过1.0%琼脂糖凝胶电泳来分离,并且根据制造商的说明书,使用ILLUSTRA[™]GFX[™]PCR DNA和凝胶条带纯化试剂盒(GE医疗,白金汉郡,英国)进行纯化。

[0606] 使用In-Fusion HD克隆试剂盒(克罗泰克实验有限公司(Clontech

Laboratories, Inc.), 山景城, 加利福尼亚州, 美国) 将PCR片段直接克隆至表达载体pCaHj505中, 无需限制性消化和连接。根据制造商的说明书, 使用In-Fusion HD克隆试剂盒将纯化的PCR产物和消化的载体连接在一起, 得到质粒, 其中062FA8GH45纤维素酶多肽编码序列的转录处于米曲霉 α -淀粉酶基因启动子的控制之下。简言之, 将0.5 μ l用Bam HI和Xho I消化的30ng/ μ l的pCaHj505、1.5 μ l的F3 PCR产物和0.6 μ l的F4 PCR产物添加到1 μ l的5X In-Fusion HD酶预混物, 最终体积为5 μ l。PCR片段的摩尔比为1:1, 并且载体:插入率为1:5。将反应在50 $^{\circ}$ C孵育15分钟, 并且然后使用5 μ l的连接反应物转化大肠杆菌TOP10感受态细胞(天根生物科技有限公司, 北京, 中国)。通过菌落PCR检测含有表达构建体的大肠杆菌转化体。菌落PCR是一种用于直接从大肠杆菌菌落快速筛选质粒插入片段的方法。简言之, 将一个单一菌落转移至PCR管中的预混合的PCR溶液中, 该PCR溶液包括PCR缓冲液、MgCl₂、dNTPs、以及从其中产生PCR片段的引物对。筛选若干个菌落。在PCR后, 将反应通过1.0%琼脂糖凝胶电泳使用TBE缓冲液进行分析。使用QIAPREP[®] Spin Miniprep试剂盒(凯杰公司(QIAGEN GmbH), 希尔登, 德国)从示出具有预期大小的插入物的菌落制备质粒DNA。插入到pCaHj505中的062FA8 GH45纤维素酶编码序列通过使用3730XL DNA分析仪(应用生物系统公司(Applied Biosystems Inc), 福斯特市, 加利福尼亚州, 美国)进行DNA测序来确认。

[0607] 062FA8突变体多肽编码序列的DNA序列和推导的氨基酸序列分别在SEQ ID NO:7和SEQ ID NO:8中示出。编码序列是1071bp(包括终止密码子), 它被78bp(核苷酸90至167)和69bp(核苷酸417至485)的2个内含子间断。编码蛋白是具有21个残基的信号肽的307个氨基酸。成熟蛋白包含286个氨基酸, 具有29kDa的预测分子量和5.8的预测等电点。

[0608] 对于062FAA, 设计如下所示的寡核苷酸引物以分开从质粒p505-RenoCorenew_ScLNK_ScCBM_2扩增GH45纤维素酶核心, 并从质粒pGH45_Thihy3331扩增接头/CBM结构域; 对应于042EB3成熟多肽具有单个取代R202A和Y292W。这些引物是由中国北京的英杰公司(Invitrogen)合成的。

[0609] 正向引物D24H9NF: 5' ACACAAC TGGGGATC CACC atgcgctctactcccgttcttc 3' (SEQ ID NO:50)

[0610] 反向引物D24H9TR: 5' agctggagtcgctcgtt cgccttgcagccgg 3' (SEQ ID NO:52)

[0611] 正向引物D24H9TlinkerF: 5' ctccggctgcaaggcg aacgacgactccagcttccc 3' (SEQ ID NO:53)

[0612] 反向引物D24H9X1linkerR: 5' CCCTCTAGATCTCGAG

[0613] ttagaggcactgcgagtaccagtc 3' (SEQ ID NO:55)

[0614] 大写字符是与pCaHj505(WO 2013/029496)的插入位点同源的区域。小写字符代表Tt纤维素酶核心的编码区域, 并且粗体字符是Th纤维素酶接头/CBM结构域的编码区域。正向引物D24H9NF中的编码序列之前的4个字母表示作为翻译过程起始物的Kozak序列。

[0615] 单独扩增和纯化816bp和255bp的两个PCR片段, 并且然后如针对062FA8所述直接克隆到表达载体pCaHj505中。062FAA突变体多肽编码序列的DNA序列和推导的氨基酸序列分别在SEQ ID NO:9和SEQ ID NO:10中示出。编码序列是1071bp(包括终止密码子), 它被78bp(核苷酸90至167)和69bp(核苷酸417至485)的2个内含子间断。编码蛋白是具有21个残基的信号肽的307个氨基酸。成熟蛋白包含286个氨基酸, 具有29kDa的预测分子量和5.8的预测等电点。

[0616] 实例7:082AH2、082DT2和082DT4GH45纤维素酶的克隆

[0617] 选择GH45纤维素酶基因,082AH2 (DNA序列为SEQ ID NO:38以及推导的氨基酸序列为SEQ ID NO:13)、082DT2 (DNA序列为SEQ ID NO:15以及推导的氨基酸序列为SEQ ID NO:16)和082DT4 (DNA序列为SEQ ID NO:17以及推导的氨基酸序列为SEQ ID NO:18)用于表达克隆。

[0618] 对于082AH2,基于从在WO 2005/054475中的SEQ ID NO:3获得的序列信息,设计下文所示的寡核苷酸引物以从合成的(金唯智公司,苏州,中国)的质粒D24YEA (SEQ ID NO:38)扩增全长GH45纤维素酶基因。这些引物是由中国北京的英杰公司 (Invitrogen) 合成的。

[0619] 正向引物D24YEA F:

[0620] 5' ACACA AACTGGGGATC CACC atgcgctcctcccctg 3' (SEQ ID NO:56)

[0621] 反向引物D24YEA R:

[0622] 5' CCCTCTAGATCTCGAG ttagagacactgggagtaccagtcg 3' (SEQ ID NO:57)

[0623] 大写字母是与pCaHj505 (WO 2013/029496)的插入位点同源的区域,并且小写字母表示082AH2GH45纤维素酶的编码区域。正向引物D24YEA F中的编码序列之前的4个字母表示作为翻译过程起始物的Kozak序列。

[0624] 如实例5中针对062FA6所述,扩增、纯化951bp的PCR片段,并且然后将其直接克隆到表达载体pCaHj505中。082AH2多肽编码序列的DNA序列和推导的氨基酸序列分别在SEQ ID NO:38和SEQ ID NO:13中示出。编码序列是951bp,包括终止密码子,不含内含子。编码蛋白是具有21个残基的信号肽的316个氨基酸。成熟蛋白包含295个氨基酸,具有31kDa的预测分子量和7.3的预测等电点。

[0625] 对于082DT2,设计如下所示的寡核苷酸引物以从质粒D24YEA扩增全长GH45纤维素酶基因,其对应于042EB3成熟多肽具有单个取代W292Y。这些引物是由中国北京的英杰公司 (Invitrogen) 合成的。

[0626] 正向引物D24YEA F:

[0627] 5' ACACA AACTGGGGATC CACC atgcgctcctcccctg 3' (SEQ ID NO:56)

[0628] 反向引物D347FKR:

[0629] 5' CCCTCTAGATCTCGAG ttagagacactgggagtagtagtcgttc 3' (SEQ ID NO:58) (斜体字母表示突变的氨基酸的核苷酸)

[0630] 大写字母是与pCaHj505 (WO 2013/029496)的插入位点同源的区域,并且小写字母表示082DT2GH45纤维素酶的编码区域。正向引物D24YEA F中的编码序列之前的4个字母表示作为翻译过程起始物的Kozak序列。斜体字母表示突变的氨基酸的核苷酸。

[0631] 如实例1中针对062FA6所述,扩增、纯化951bp的PCR片段,并且然后将其直接克隆到表达载体pCaHj505中。082DT2突变体多肽编码序列的DNA序列和推导的氨基酸序列分别在SEQ ID NO:15和SEQ ID NO:16中示出。编码序列是951bp,包括终止密码子,不含内含子。编码蛋白是具有21个残基的信号肽的316个氨基酸。成熟蛋白包含295个氨基酸,具有31kDa的预测分子量和7.3的预测等电点。

[0632] 对于082DT4,设计下文所示的寡核苷酸引物以分开从质粒D24YEA扩增GH45纤维素酶核心,并从质粒1 (WO 2012/089024,通过引用特此结合) 扩增接头/CBM结构域;对应于042EB3成熟多肽具有单个取代R202A。这些引物是由中国北京的英杰公司 (Invitrogen) 合

成的。

[0633] 正向引物D24YEAF:5' ACACAACCTGGGGATC CACC atgcgctcctcccctg 3' (SEQ ID NO:56)

[0634] 反向引物D3379PlinkerR:5' tggagtcgtcgtt tgcgacagccggctcctg 3' (SEQ ID NO:59)

[0635] 正向引物D3379PlinkerF:5' gaccggctgtcgcgca aacgacgactccagcttccc 3' (SEQ ID NO:60)

[0636] 反向引物D24H9TlinkerR:5' CCCTCTAGATCTCGAG ttagaggcactgcgagtagtagtgcg 3' (SEQ ID NO:54)

[0637] 大写字母是与pCaHj505(WO 2013/029496)的插入位点同源的区域。小写字母代表大孢圆孢霉GH45纤维素酶(Sc纤维素酶)核心的编码区域,并且粗体字母是土生梭孢壳霉GH45纤维素酶(Tt纤维素酶)接头/CBM结构域的编码区域。正向引物D24YEAF中的编码序列之前的4个字母表示作为翻译过程起始物的Kozak序列。

[0638] 单独扩增和纯化663bp和231bp的两个PCR片段,并且然后如针对O62FA8所述直接克隆到表达载体pCaHj505中。O82DT4突变体多肽编码序列的DNA序列和推导的氨基酸序列分别在SEQ ID NO:17和SEQ ID NO:18中示出。编码序列是894bp,包括终止密码子,不含内含子。编码蛋白是具有21个残基的信号肽的297个氨基酸。成熟蛋白包含276个氨基酸,具有28kDa的预测分子量和5.2的预测等电点。

[0639] 实例8:P446XD GH45纤维素酶的克隆

[0640] 选择GH45纤维素酶基因,P446XD(DNA序列为SEQ ID NO:19以及推导的氨基酸序列为SEQ ID NO:20)用于表达克隆。

[0641] 设计以下所示的寡核苷酸引物以从质粒001-8#1扩增全长GH45纤维素酶基因(包含四孢脉孢菌纤维素酶,WO 2015/058700中的SEQ ID NO:2,通过引用特此结合),其对应于O42EB3成熟多肽具有单个取代Y292W。这些引物是由中国北京的英杰公司(Invitrogen)合成的。

[0642] 正向引物D43NXPF(基因特异性):

[0643] 5' ACACAACCTGGGGATC CACC atgcgctcctccactgttc 3' (SEQ ID NO:61)

[0644] 反向引物D43NXPR(致突变的):

[0645] 5' CCCTCTAGATCTCGAG ttaggcacactggtggtaccaatc 3' (SEQ ID NO:62)(斜体字母表示突变的氨基酸的核苷酸)

[0646] 大写字母是与pCaHj505(WO 2013/029496)的插入位点同源的区域,并且小写字母表示P446XD GH45纤维素酶的编码区域。正向引物D43NXPF中的编码序列之前的4个字母表示作为翻译过程起始物的Kozak序列。斜体字母表示突变的氨基酸的核苷酸。

[0647] 如针对O62FA6所述,扩增、纯化935bp的PCR片段,并且然后将其直接克隆到表达载体pCaHj505中。P446XD突变体多肽编码序列的DNA序列和推导的氨基酸序列分别在SEQ ID NO:19和SEQ ID NO:20中示出。编码序列是935bp(包括终止密码子),其被一个53bp(核苷酸339至391)的内含子间断。编码蛋白是具有21个残基的信号肽的293个氨基酸。成熟蛋白包含272个氨基酸,具有28kDa的预测分子量和6.9的预测等电点。

[0648] 实例9:O62FA6变体(O14G83、O141FN、O141FS、O141FK、O141FU、O141FF、O141FH和

0141FQ)的克隆

[0649] 选择062FA6的八个变体用于表达克隆,包括014G83 (DNA序列为SEQ ID NO:21以及推导的氨基酸序列为SEQ ID NO:22)、0141FN (DNA序列为SEQ ID NO:23以及推导的氨基酸序列为SEQ ID NO:24)、0141FS (DNA序列为SEQ ID NO:25以及推导的氨基酸序列为SEQ ID NO:26)、0141FK (DNA序列为SEQ ID NO:27以及推导的氨基酸序列为SEQ ID NO:28)、0141FU (DNA序列为SEQ ID NO:29以及推导的氨基酸序列为SEQ ID NO:30)、0141FF (DNA序列为SEQ ID NO:31以及推导的氨基酸序列为SEQ ID NO:32)、0141FH (DNA序列为SEQ ID NO:33以及推导的氨基酸序列为SEQ ID NO:34)和0141FQ (DNA序列为SEQ ID NO:35以及推导的氨基酸序列为SEQ ID NO:36)。

[0650] 对于014G83,设计下文所示的寡核苷酸引物以从用于062FA6 (SEQ ID NO:5)的质粒D24H9N扩增全长GH45纤维素酶基因,其对应于042EB3成熟多肽具有两个取代,R265K和W266Y。这些引物是由中国北京的英杰公司 (Invitrogen) 合成的。

[0651] 正向引物D24H9NF:5' ACACAAC TGGGATC CACC atgcgctctactcccgttcttc 3' (SEQ ID NO:50)

[0652] 反向引物D43NG3linkerRnew:5' cgcactgagcgtactt ctgcgctgcgcagc 3' (SEQ ID NO:63)

[0653] 正向引物D43NG3linkerFnew:5' aagtacgctcagtgccggtgg 3' (SEQ ID NO:64)

[0654] 反向引物D24H9NR:5' CCCTCTAGATCTCGAG tcagaggcactgcgagtagtagtc 3' (SEQ ID NO:51)

[0655] 大写字母是与pCaHj505 (WO 2013/029496)的插入位点同源的区域,并且小写字母表示014G83GH45纤维素酶的编码区域。正向引物D24H9NF中的编码序列之前的4个字母表示作为翻译过程起始物的Kozak序列。斜体字母表示突变的氨基酸的核苷酸。

[0656] 单独扩增和纯化1018bp和102bp的两个PCR片段,并且然后如针对062FA8所述直接克隆到表达载体pCaHj505中。014G83突变体多肽编码序列的DNA序列和推导的氨基酸序列分别在SEQ ID NO:21和SEQ ID NO:22中示出。编码序列是1104bp (包括终止密码子),它被78bp (核苷酸90至167)和69bp (核苷酸417至485)的2个内含子间断。编码蛋白是具有21个残基的信号肽的318个氨基酸。成熟蛋白包含297个氨基酸,具有30kDa的预测分子量和5.8的预测等电点。

[0657] 对于0141FN,设计下文所示的寡核苷酸引物以从用于062FA6 (SEQ ID NO:5)的质粒D24H9N扩增全长GH45纤维素酶基因,其对应于042EB3成熟多肽具有两个取代,W265Y和F274Y。这些引物是由中国北京的英杰公司 (Invitrogen) 合成的。

[0658] 正向引物D24H9NF:5' ACACAAC TGGGATC CACC atgcgctctactcccgttcttc 3' (SEQ ID NO:50)

[0659] 反向引物D43NJ3linkerR:5' gctgtagccgatgccca ccgactgagcgtaacgc 3' (SEQ ID NO:65)

[0660] 正向引物D43NJ3linkerF:5' gcgttacgctcagtgcc ggtggcactcggctacagc 3' (SEQ ID NO:66)

[0661] 反向引物D24H9NR:5' CCCTCTAGATCTCGAG tcagaggcactgcgagtagtagtc 3' (SEQ ID NO:51)

[0662] 大写字母是与pCaHj505 (WO 2013/029496)的插入位点同源的区域,并且小写字母表示0141FN GH45纤维素酶的编码区域。正向引物D24H9NF中的编码序列之前的4个字母表示作为翻译过程起始物的Kozak序列。斜体字母表示突变的氨基酸的核苷酸。

[0663] 单独扩增和纯化1035bp和103bp的两个PCR片段,并且然后如针对062FA8所述直接克隆到表达载体pCaHj505中。0141FN突变体多肽编码序列的DNA序列和推导的氨基酸序列分别在SEQ ID NO:23和SEQ ID NO:24中示出。编码序列是1104bp(包括终止密码子),它被78bp(核苷酸90至167)和69bp(核苷酸417至485)的2个内含子间断。编码蛋白是具有21个残基的信号肽的318个氨基酸。成熟蛋白包含297个氨基酸,具有30kDa的预测分子量和5.8的预测等电点。

[0664] 对于0141FS,设计下文所示的寡核苷酸引物以从用于062FA6 (SEQ ID NO:5)的质粒D24H9N扩增全长GH45纤维素酶基因,其对应于042EB3成熟多肽具有两个取代,T237N和T246N。这些引物是由中国北京的英杰公司(Invitrogen)合成的。

[0665] 正向引物D24H9NF:5' ACACAAC TGGGATC CACC atgcgctctactcccgtttcttc 3' (SEQ ID NO:50)

[0666] 反向引物D43PNTlinkerR:5' ggagtttgtggaggtt gctttggtcgaagtgttgac 3' (SEQ ID NO:67)

[0667] 正向引物D43PNTlinkerF:5' cgtccaacacttcgac caagcaacctccacaaactcc 3' (SEQ ID NO:68)

[0668] 反向引物D24H9NR:5' CCCTCTAGATCTCGAG tcagaggcactgcgagtagtagtc 3' (SEQ ID NO:51)

[0669] 大写字母是与pCaHj505 (WO 2013/029496)的插入位点同源的区域,并且小写字母表示0141FS GH45纤维素酶的编码区域。正向引物D24H9NF中的编码序列之前的4个字母表示作为翻译过程起始物的Kozak序列。斜体字母表示突变的氨基酸的核苷酸。

[0670] 单独扩增和纯化951bp和191bp的两个PCR片段,并且然后如针对062FA8所述直接克隆到表达载体pCaHj505中。0141FS突变体多肽编码序列的DNA序列和推导的氨基酸序列分别在SEQ ID NO:25和SEQ ID NO:26中示出。编码序列是1104bp(包括终止密码子),它被78bp(核苷酸90至167)和69bp(核苷酸417至485)的2个内含子间断。编码蛋白是具有21个残基的信号肽的318个氨基酸。成熟蛋白包含297个氨基酸,具有30kDa的预测分子量和5.8的预测等电点。

[0671] 对于0141FK,设计下文所示的寡核苷酸引物以从质粒D24H9N (SEQ ID NO:5)扩增全长GH45纤维素酶基因,其具有单个取代F274Y。这些引物是由中国北京的英杰公司(Invitrogen)合成的。

[0672] 正向引物D24H9NF:5' ACACAAC TGGGATC CACC atgcgctctactcccgtttcttc 3' (SEQ ID NO:50)

[0673] 反向引物D43NHUlinkerR:5' gctgtagccgatgcc a cgcactgagcccaac 3' (SEQ ID NO:69)

[0674] 正向引物D43NHUlinkerF:5' gcgttgggctcagtgc ggtggcatcggtacagc 3' (SEQ ID NO:70)

[0675] 反向引物D24H9NR:5' CCCTCTAGATCTCGAG tcagaggcactgcgagtagtagtc 3' (SEQ ID

NO:51)

[0676] 大写字母是与pCaHj505 (WO 2013/029496) 的插入位点同源的区域,并且小写字母表示0141FK GH45纤维素酶的编码区域。正向引物D24H9NF中的编码序列之前的4个字母表示作为翻译过程起始物的Kozak序列。斜体字母表示突变的氨基酸的核苷酸。

[0677] 单独扩增和纯化1035bp和103bp的两个PCR片段,并且然后如针对062FA8所述直接克隆到表达载体pCaHj505中。0141FK突变体多肽编码序列的DNA序列和推导的氨基酸序列分别在SEQ ID NO:27和SEQ ID NO:28中示出。编码序列是1104bp(包括终止密码子),它被78bp(核苷酸90至167)和69bp(核苷酸417至485)的2个内含子间断。编码蛋白是具有21个残基的信号肽的318个氨基酸。成熟蛋白包含297个氨基酸,具有30kDa的预测分子量和5.8的预测等电点。

[0678] 对于0141FU,设计下文所示的寡核苷酸引物以从用于062FA6 (SEQ ID NO:5) 的质粒D24H9N扩增全长GH45纤维素酶基因,其对应于042EB3成熟多肽具有单个取代S255P。这些引物是由中国北京的英杰公司 (Invitrogen) 合成的。

[0679] 正向引物D24H9NF:5' ACACAAC TGGGATC CACC atgcgctctactcccgttcttc 3' (SEQ ID NO:50)

[0680] 反向引物D43PNNlinkerR:5' ggacggcgacgtctgg gacgaggcgggtggaagttg 3' (SEQ ID NO:71)

[0681] 正向引物D43PNNlinkerF:5' cttccaccgcctcgtc ccagacgtcgccgtcc 3' (SEQ ID NO:72)

[0682] 反向引物D24H9NR:5' CCCTCTAGATCTCGAG tcagaggcactgcgagtagtagtc 3' (SEQ ID NO:51)

[0683] 大写字母是与pCaHj505 (WO 2013/029496) 的插入位点同源的区域,并且小写字母表示0141FU GH45纤维素酶的编码区域。正向引物D24H9NF中的编码序列之前的4个字母表示作为翻译过程起始物的Kozak序列。斜体字母表示突变的氨基酸的核苷酸。

[0684] 单独扩增和纯化978bp和158bp的两个PCR片段,并且然后如针对062FA8所述直接克隆到表达载体pCaHj505中。0141FU突变体多肽编码序列的DNA序列和推导的氨基酸序列分别在SEQ ID NO:29和SEQ ID NO:30中示出。编码序列是1104bp(包括终止密码子),它被78bp(核苷酸90至167)和69bp(核苷酸417至485)的2个内含子间断。编码蛋白是具有21个残基的信号肽的318个氨基酸。成熟蛋白包含297个氨基酸,具有30kDa的预测分子量和5.8的预测等电点。

[0685] 对于0141FF,设计下文所示的寡核苷酸引物以从用于062FA6 (SEQ ID NO:5) 的质粒D24H9N扩增全长GH45纤维素酶基因,其对应于042EB3成熟多肽具有单个取代W266Y。这些引物是由中国北京的英杰公司 (Invitrogen) 合成的。

[0686] 正向引物D24H9NF:5' ACACAAC TGGGATC CACC atgcgctctactcccgttcttc 3' (SEQ ID NO:50)

[0687] 反向引物D43NFYlinkerR:5' agcgtaacgctgcgct ggcgagcctcctccg 3' (SEQ ID NO:73)

[0688] 正向引物D43NFYlinkerF:5' ccgaggaggctgcgc agcgcagcgttacgctc 3' (SEQ ID NO:74)

[0689] 反向引物D24H9NR:5' CCCTCTAGATCTCGAG *tcagaggcactg*cgagtagtagtc 3' (SEQ ID NO:51)

[0690] 大写字母是与pCaHj505 (WO 2013/029496) 的插入位点同源的区域,并且小写字母表示0141FF GH45纤维素酶的编码区域。正向引物D24H9NF中的编码序列之前的4个字母表示作为翻译过程起始物的Kozak序列。斜体字母表示突变的氨基酸的核苷酸。

[0691] 单独扩增和纯化1011bp和125bp的两个PCR片段,并且然后如针对062FA8所述直接克隆到表达载体pCaHj505中。0141FF突变体多肽编码序列的DNA序列和推导的氨基酸序列分别在SEQ ID NO:31和SEQ ID NO:32中示出。编码序列是1104bp (包括终止密码子),它被78bp (核苷酸90至167) 和69bp (核苷酸417至485) 的2个内含子间断。编码蛋白是具有21个残基的信号肽的318个氨基酸。成熟蛋白包含297个氨基酸,具有30kDa的预测分子量和5.8的预测等电点。

[0692] 对于0141FH,设计下文所示的寡核苷酸引物以从用于062FA6 (SEQ ID NO:5) 的质粒D24H9N扩增全长GH45纤维素酶基因,其对应于042EB3成熟多肽具有单个取代R265K。这些引物是由中国北京的英杰公司 (Invitrogen) 合成的。

[0693] 正向引物D24H9NF:5' ACACAACTGGGGATC CACC *atgcgctctactcccgttcttc* 3' (SEQ ID NO:50)

[0694] 反向引物D43NG1linkerR:5' *ccacttctgcgctg*cg cagcctcctccggtgg 3' (SEQ ID NO:75)

[0695] 正向引物D43NG1linkerF:5' *ccaccggaggaggctg* cgcagcgcagaagtgg 3' (SEQ ID NO:76)

[0696] 反向引物D24H9NR:5' CCCTCTAGATCTCGAG *tcagaggcactg*cgagtagtagtc 3' (SEQ ID NO:51)

[0697] 大写字母是与pCaHj505 (WO 2013/029496) 的插入位点同源的区域,并且小写字母表示0141FH GH45纤维素酶的编码区域。正向引物D24H9NF中的编码序列之前的4个字母表示作为翻译过程起始物的Kozak序列。斜体字母表示突变的氨基酸的核苷酸。

[0698] 单独扩增和纯化1008bp和128bp的两个PCR片段,并且然后如针对062FA8所述直接克隆到表达载体pCaHj505中。0141FH突变体多肽编码序列的DNA序列和推导的氨基酸序列分别在SEQ ID NO:33和SEQ ID NO:34中示出。编码序列是1104bp (包括终止密码子),它被78bp (核苷酸90至167) 和69bp (核苷酸417至485) 的2个内含子间断。编码蛋白是具有21个残基的信号肽的318个氨基酸。成熟蛋白包含297个氨基酸,具有30kDa的预测分子量和5.8的预测等电点。

[0699] 对于0141FQ,设计下文所示的寡核苷酸引物以从用于062FA6 (SEQ ID NO:5) 的质粒D24H9N扩增全长GH45纤维素酶基因,其对应于042EB3成熟多肽具有两个取代S221P和S224P。这些引物是由中国北京的英杰公司 (Invitrogen) 合成的。

[0700] 正向引物D24H9NF:5' ACACAACTGGGGATC CACC *atgcgctctactcccgttcttc* 3' (SEQ ID NO:50)

[0701] 反向引物D43NXRlinkerR:5' *ggacggggaggacggg* gaggactggcctcccg 3' (SEQ ID NO:77)

[0702] 正向引物D43NXRlinkerF:5' *gggaggccagtcctcc* ccgtcctcccgtcc 3' (SEQ ID

NO:78)

[0703] 反向引物D24H9NR:5' CCCTCTAGATCTCGAG tcagaggcactgcgagtagtagtc 3' (SEQ ID NO:51)

[0704] 大写字母是与pCaHj505 (WO 2013/029496) 的插入位点同源的区域,并且小写字母表示0141FQ GH45纤维素酶的编码区域。正向引物D24H9NF中的编码序列之前的4个字母表示作为翻译过程起始物的Kozak序列。斜体字母表示突变的氨基酸的核苷酸。

[0705] 单独扩增和纯化885bp和250bp的两个PCR片段,并且然后如针对062FA8所述直接克隆到表达载体pCaHj505中。0141FQ突变体多肽编码序列的DNA序列和推导的氨基酸序列分别在SEQ ID NO:35和SEQ ID NO:36中示出。编码序列是1104bp (包括终止密码子),它被78bp (核苷酸90至167) 和69bp (核苷酸417至485) 的2个内含子间断。编码蛋白是具有21个残基的信号肽的318个氨基酸。成熟蛋白包含297个氨基酸,具有30kDa的预测分子量和5.8的预测等电点。

[0706] 实例10:GH45纤维素酶变体在米曲霉中的表达

[0707] 米曲霉菌株MT3568用于GH45纤维素酶变体基因的异源表达。米曲霉MT3568是米曲霉JaL355的amdS (乙酰胺酶) 破坏的基因衍生物 (WO 02/40694), 其中通过用pyrG基因破坏米曲霉乙酰胺酶 (amdS) 基因恢复pyrG营养缺陷型。

[0708] 根据Christensen等人,1988,Bio/Technology [生物/技术]6:1419-1422中所描述的方法制备原生质体。使用三μg实例5-9中所述的每种质粒转化米曲霉MT3568。

[0709] 用每种质粒转化米曲霉MT3568产生约10个转化体。将四个转化体分离到各个基本培养基平板中,并且然后分别接种到24孔板中的3ml YPM培养基 (1%酵母提取物、2%蛋白胨和2%麦芽糖) 中,并在30°C、150rpm孵育。3天的孵育之后,根据制造商的说明书,在具有MES的NUPAGE®NOVEX®4%-12%Bis-Tris凝胶 (英杰公司,卡尔斯巴德,加利福尼亚州,美国) 上对来自每一培养物的21μl上清液进行分析。将所得的凝胶用INSTANTBLUE™ (艾本德有限公司 (Expedeon Ltd.), 巴布拉汉,剑桥,英国) 染色。培养物的SDS-PAGE概貌显示大多数转化体具有大约30kDa的主要条带。选择每一种转化体中的一个作为表达菌株。

[0710] 实例11:表达菌株的发酵

[0711] 将实例10中所述的表达菌株的斜面用10ml YPM洗涤,并接种到各含有400ml YPM培养基的2升烧瓶中以产生用于纯化和表征该酶的培养液。在第3天采集培养物并使用0.22 μm 1升瓶顶部真空过滤器 (康宁公司,康宁,纽约州,美国) 过滤。

[0712] 实例12:GH45纤维素酶变体的纯化和活性测定

[0713] 062FA6

[0714] 首先用硫酸铵 (80%饱和度) 沉淀062FA6的培养上清液,然后用pH 4.5的20mM NaAc透析。将溶液用0.45μm过滤器过滤,并且然后加载到用pH4.5的20mM NaAc平衡的Q SEPHAROSE®快流柱 (GE医疗) 中。使用NaCl浓度梯度作为从0至1M的洗脱缓冲液,并且然后收集洗脱级分和贯流级分以检测纤维素酶活性。尝试使用具有纤维素酶活性的贯流级分加载Mono Q柱 (GE医疗),但活性物仍然在贯流级分中。然后将活性物加载到用pH 7.0的20mM Tris-HCl (添加了0.3M NaCl) 平衡的SUPERDEX™ 75柱 (GE医疗) 中。最后,通过在pH 7.0的20mM Tris-HCl中的终浓度为1.2M的硫酸铵添加来自SUPERDEX™ 75的活性物,并且然后将活性物加载至具有苯基琼脂糖6快流的色谱柱 (GE医疗) 中。活性物通过盐浓度的梯度

降低而被洗脱。具有活性物的级分通过SDS-PAGE分析,并且然后浓缩用于进一步评估。

[0715] 062FA8和062FAA具有与062FA6类似的PI,并且纯化程序与062FA6相似,只是没有Mono Q步骤。

[0716] 082DT2

[0717] 首先用硫酸铵(80%饱和度)沉淀082DT2的培养上清液,然后用pH 9.0的20mM Tris-HCl透析。将溶液用0.45 μ m过滤器过滤,并且然后加载到用pH 9.0的20mM Tris-HCl平衡的QSEPHAROSE®快流柱(GE医疗)中。使用NaCl浓度梯度作为从0至1M的洗脱缓冲液,然后分析洗脱级分和贯流样品的纤维素酶活性。贯流样品是靶纯化蛋白。

[0718] 082DT4

[0719] 通过具有1.2M最终浓度的硫酸铵添加082DT4的培养上清液并将该上清液加载到用添加了1.2M硫酸铵的20mM PBS(pH 6.0)平衡的苯基琼脂糖6快流柱(GE医疗)中。建立硫酸铵浓度从1.2M到0的梯度降低,并且分析并合并具有纤维素酶活性的级分。

[0720] 将082AH2、0141FS、0141FK、0141FU、0141FQ、0141FH、0141FF、P446XD、014G83通过与082DT4相似的程序纯化。靶蛋白可以从贯流级分或洗脱级分获得。通过蓝色底物检测纤维素酶活性。将AZCL-HE-纤维素(麦格酶公司(Megazyme))的蓝色底物溶于pH 6.0的100mM Bis-Tris中,终浓度为0.2%(w/v)。将20 μ l酶与100 μ l蓝色底物混合,并在50 $^{\circ}$ C用Thermomixer(艾本德公司)以600rpm孵育30分钟,然后将60 μ l的混合物上清液转移到新孔中以测量595nm处的吸光度。变体示出良好的纤维素酶活性。

[0721] 实例13:在耐洗牢度试验仪(Laundry-0-meter)中采用042EB3、052EWC的生物抛光

[0722] 在本实例中使用由实例4纯化的野生型蛋白052EWC(SEQ ID NO:12)和融合蛋白042EB3(SEQ ID NO:2)进行生物抛光。

[0723] 将棉布小块布样剪成约16cm*16cm(各自约5克)。将这些小块布样在编号之前置于控制室(65%相对湿度,21 $^{\circ}$ C)中24小时,通过分析天平称重并记录。用耐洗牢度试验仪进行生物抛光。在每个烧杯中放置两个经调节的小块布样和20个大钢球(总重量220克)或4个橡胶球,以分别提供高和低水平的机械辅助。将烧杯中填充根据表1的酶和如介质部分所述而制备的缓冲液,至总体积为100ml左右,这样可以获得约10:1(v/w)的液体与织物之比率。

[0724] 在选择所需程序后启动耐洗牢度试验仪(LOM)机器,并且当温度达到预设温度55 $^{\circ}$ C时,将其保持。每个烧杯装配一种内衬有2个neoprin垫片的盖子并且用金属夹紧装置紧密闭合。在竖直位,在4个鼓位的每一个中放置5个烧杯。在55 $^{\circ}$ C用下表中指定的剂量和机械辅助处理1小时后,将小块布样从烧杯中取出并转移到含有2g/L碳酸钠的失活溶液中并保持于85 $^{\circ}$ C持续10分钟。然后,将小块布样在热水中漂洗2次并在冷水中漂洗2次。并且,在评估重量损失和起球记录之前,将它们滚动干燥1小时,在65%相对湿度、21 $^{\circ}$ C下调节24小时。

[0725] 如表1所总结的,在每个烧杯中在采用20个钢球的高机械作用下,基于蛋白质,融合蛋白042EB3比052EWC稍微提供更佳生物抛光性能;而在每个烧杯中在采用4个橡胶球的低机械作用下,基于蛋白质,042EB3比052EWC更佳,在于它比052EWC提供高0.5-0.7的起球记录。

[0726] 表1.在55 $^{\circ}$ C、pH 6.5下采用042EB3和052EWC的LOM生物磨损

[0727]

酶	剂量 (mg/g 织物)	每个烧杯中的球	起球记录
O42EB3	0.008	每个烧杯中有 20 个钢球	2.9 ± 0.2
	0.016		3.8 ± 0.0
	0.024		3.8 ± 0.0
O52EWC	0.008		3.0 ± 0.0
	0.016		3.4 ± 0.2
	0.024		3.8 ± 0.0
O42EB3	0.008	每个烧杯中有 4 个橡胶球	2.6 ± 0.2
	0.016		3.6 ± 0.2
	0.024		3.4 ± 0.2
O52EWC	0.008		1.9 ± 0.2
	0.016		2.9 ± 0.2

[0728]

	0.024		2.9 ± 0.2
--	-------	--	-----------

[0729] 实例14:在耐洗牢度试验仪中采用O82AH2、O82DT2和O82DT4的生物抛光

[0730] 在本实例中使用从实例12纯化的野生型蛋白O82AH2 (SEQ ID NO:13) 和变体O82DT2 (W292Y, SEQ ID NO:16) 和融合蛋白O82DT4 (SEQ ID NO:18) 用于生物抛光。

[0731] 织物的制备和试验操作与实例13相似。如表2所总结的,在每个烧杯中采用20个钢球的高机械作用下,基于蛋白质,O82DT2和O82DT4比野生型纤维素酶O82AH2稍微提供更好的生物抛光性能;而在每个烧杯中采用4个橡胶球的低机械作用下,基于蛋白质,O82DT2和O82DT4比O82AH2更佳,在于它比O82AH2提供高1.0和0.5的起球记录。

[0732] 表2.在55°C、pH 6.5下采用O82AH2、O82DT2和O82DT4的LOM生物磨损

[0733]

酶	剂量 (mg/g)	每个烧杯中的球	起球记录
O82DT4	0.008	20 个钢球	3.8 ± 0.1
O82AH2			3.7 ± 0.4
O82DT2			3.9 ± 0.2
O82DT4	0.024	4 个橡胶球	3.3 ± 0.1
O82AH2			2.8 ± 0.3
O82DT2			3.8 ± 0.1

[0734] 实例15:在耐洗牢度试验仪中采用O42EB3、O62FA6、O62FAA、O62FA8的生物抛光

[0735] 在本实例中使用从实例4和14中纯化出的在组内仅具有一个氨基酸差异的两组变体O42EB3/O62FA6、O62FAA/O62FA8进行生物抛光。

[0736] 织物的制备和试验操作与实例13相似。如表3所总结,在每个烧杯中采用20个钢球的高机械作用下,随着主链O42EB3中的W292Y的突变,得到的酶O62FA6在具有相同蛋白质剂量的起球记录方面提供高得多的生物抛光性能,并且超出期望的是,为达到等同的起球记录水平,与O42EB3相比,O62FA6的重量损失减少;相似地,所得到的酶蛋白O62FA8在起球记录的更佳生物抛光性能上示出显著改进,并且在等同的起球记录水平上示出减少的重量损失。

[0737] 表3. 在55℃、pH 6.5下采用O42EB3、O62FA6、O62FAA、O62FA8的LOM生物磨损

[0738]

酶	剂量 (mg/g)	重量损失 (%)	起球记录
O42EB3	0.004	0.1 ± 0.1	1.8 ± 0.0
	0.008	0.1 ± 0.1	1.8 ± 0.0
	0.016	0.7 ± 0.2	3.5 ± 0.0
O62FA6	0.004	0.2 ± 0.3	2.5 ± 0.0
	0.008	0.7 ± 0.1	3.8 ± 0.0
	0.016	1.5 ± 0.0	4.0 ± 0.0
O62FAA	0.004	-0.1 ± 0.2	1.8 ± 0.2
	0.008	0.0 ± 0.2	1.6 ± 0.2
	0.016	0.3 ± 0.0	2.9 ± 0.2
O62FA8	0.004	-0.1 ± 0.1	2.4 ± 0.2
	0.008	0.0 ± 0.2	2.5 ± 0.0
	0.016	0.3 ± 0.1	3.8 ± 0.0

[0739] 实例16:在wascator中采用O62FA6和O52EWC的牛仔布磨损

[0740] 在本实例中使用两种蛋白质-从实例12纯化的O62FA6、以及O52EWC,进行牛仔布磨损试验。

[0741] 切割并缝合两种粗牛仔布,形成尺寸为35cm×45cm并且重量为180克的筒形物。根据产品应用指南,将粗牛仔布筒形物用淀粉酶退浆。将这些筒形物置于控制室(65%相对湿度,21℃)中24小时,之后将它们进行编号,通过天平称重并记录。将具有少量填充物以达到1000g的5个经调节的筒形物(3个1型筒形物和2个2型筒形物)加载到wascator (Electrolux Wascator, FOM71CLS) 中。将牛仔布筒形物用10L淡水在室温漂洗10分钟,并且然后用0.03mg纤维素酶蛋白质/g织物在55℃在10L淡水中处理60分钟,并用乙酸和/或氢氧化钠将pH调节至6-6.5,然后用淡水冲洗10分钟两次,之后将它们离心并滚筒干燥(AEG, LAVATHERM 37700,德国)1小时并调节。

[0742] 通过用预校准DataColor SF450X测量纤维素酶处理之前和之后的反射比而确定牛仔布样品的磨损和返染水平。对于L*和b*两者而言,每个织物都取四个读数并且使用这四个读数的平均值。用样品蓝面的指数CIE L*评估磨损水平,并且用样品背面的指数CIE b*评估返染水平。

[0743] 如表4所示,在55℃、pH 6-6.5下,蛋白质剂量相同的情况下,与052EWC相比,分别在1型和2型牛仔布上062FA6导致的牛仔布磨损高得多和略高;并且在等同的磨损水平上,两种纤维素酶都引起类似的返染水平。

[0744] 表4.在wascator中,在55℃、1小时下,062FA6和052EWC导致的牛仔布磨损
[0745]

酶	剂量 (mg/g 织物)	牛仔布类型	前面的 L*	背面的 b*
O62FA6	0.03	1	23.9 ± 0.2	-11.3 ± 0.3
		2	23.4 ± 0.0	-10.8 ± 0.3
O52EWC	0.03	1	23.2 ± 0.0	-10.9 ± 0.2
		2	23.3 ± 0.0	-10.9 ± 0.6

[0746] 实例17.在wascator中采用062FA6和052EWC的生物抛光

[0747] 含有实例16中提及的两种蛋白质的发酵液也在wascator中测试了不同织物的生物抛光。基于活性,测量发酵液的CNUR活性并加载到本研究中。

[0748] 在每个wascator生物抛光试验中,包括5种不同的棉针织物:两种漂白的双棉布,以及3种分别为蓝色、棕色和海军蓝的单面棉针织布。如上所述对织物进行预调节并称重。在wascator中的每次生物抛光试验中,五种织物总计约1kg。用与实例16相似的程序处理织物,只是在本实例中,在55℃和pH 6-6.5的主洗涤中加载18CNUR/g织物;并且在两次漂洗步骤之前,在80℃用1g/L碳酸钠进行另外一次灭活步骤持续10分钟。然后将织物离心并滚筒干燥(AEG, LAVATHERM 37700, 德国)1小时,并进行调节。

[0749] 如表5所总结,在55℃、pH 6-6.5下,在相同的活性剂量下,与052EWC相比,062FA6导致在起球记录上的相似的生物抛光性能,以及在所有3种类型的染色织物上具有高0.5-1的起球记录的更佳生物抛光性能;在对于等同或甚至更高的起球记录上的重量损失方面,062FA6在3/5的织物上提供比052EWC更佳的性能,其等同的或甚至更高的起球记录上有更少的重量损失,而在其他的2/5的织物上,062FA6导致与更高的起球记录相一致的略微更高的重量损失。对于强度来说,对于用062FA6处理的织物来说,与052EWC相比,在等同或甚至更高的起球记录下2/5织物上保留了更高的强度,在等同的起球记录下1/5织物上保留较低的强度,最后2/5具有高0.5-1的起球记录和更低的强度损失。总而言之,在相同的活性剂量下,062FA6提供更佳生物抛光性能、更少的重量和强度损失。

[0750] 表5.062FA6和052EWC在wascator中在55℃、1小时下的生物抛光

[0751]

酶	剂量 (CNUR/g 织物)	织物	起球记录	重量损失 (%)	强度 (N)
O52EWC	18	蓝色	3.7 ± 0.2	2.0	650 ± 35
		棕色	1.3 ± 0.3	-0.9	573 ± 37
		白色	3.9 ± 0.3	2.0	885 ± 44
		普白	4.2 ± 0.3	2.4	912 ± 67
		海军蓝	1.6 ± 0.5	-0.4	663 ± 28
O62FA6	18	蓝色	4.2 ± 0.3	2.3	606 ± 40
		棕色	2.3 ± 0.4	-0.4	609 ± 17
		白色	3.9 ± 0.3	1.9	807 ± 55
		普白	4.1 ± 0.2	1.9	955 ± 53
		海军蓝	2.6 ± 0.6	-0.3	618 ± 49

[0752] 实例18.在喷射式染色机中采用O62FA6和O52EWC的生物抛光

[0753] 还在喷射式染色机 (Allfit 10, Fong's) 中针对不同织物的生物抛光来测试实例17中提及的发酵液。

[0754] 在每个喷射式染色机生物抛光试验中,包括2种不同的棉针织物:约2kg双棉布和约1kg棉/氨纶单面棉针织布。如上所述对织物进行预调节并称重。在喷射式染色机的每个生物抛光试验中,两种类型的棉织物和5kg棉填充物合计高达约8kg。将它们用80L淡水在室温漂洗10分钟,并且在主洗中在30CNUR/g织物中用O62FA6或O52EWC于55℃在80L淡水中处理90分钟,并用乙酸和/或氢氧化钠将pH调节至6-6.5,纤维素酶分别加载75分钟和90分钟后收集织物小块布样;用纤维素酶主洗后,排水并用80L淡水中的1g/L碳酸钠在80℃灭活10分钟,用淡水漂洗10分钟两次。在喷射式染色机的整个过程中,绞车速度设定为40m/min,并且设定合适的泵压力以确保织物在机器中平稳运行。然后将处理过的织物离心并滚筒干燥1小时,并进行调节。

[0755] 如表6所总结,在55℃、pH 6-6.5下,在相同的活性剂量下,与O52EWC相比,O62FA6导致显著更佳生物抛光性能:在90min,与O52EWC相比,O62FA6导致在两种织物上的起球记录分别高0.8和0.6;或者为达到O52EWC在90min的起球记录水平,O62FA6仅需要75min。除了更高的起球记录水平外,O62FA6还保留了比O52EWC更高的强度。

[0756] 表6.在喷射式染色机中在55℃、90min由O62FA6和O52EWC导致的生物抛光

[0757]

酶	剂量(CNUR/g 织物)	织物	75 min	90 min	
			起球记录	起球记录	强度(N)
O62FA6	30	双棉布	2.6 ± 0.1	3.6 ± 0.1	979 ± 51
		棉/氨纶单面棉针织布	1.8 ± 0.3	3.5 ± 0.1	433 ± 19
O52EWC	30	双棉布	1.8 ± 0.3	2.8 ± 0.4	930 ± 65
		棉/氨纶单面棉针织布	1.5 ± 0.4	1.9 ± 0.4	429 ± 22

[0758] 实例19:在耐洗牢度试验仪中采用O14G83、O52EWC的生物抛光

[0759] 在本实例中使用来自实例12的纯化蛋白O14G83和实例17中提及的发酵液O52EWC进行生物抛光试验。

[0760] 织物的制备和试验操作与实例13相似。如表7所总结,在每个烧杯中采用20个钢球的高机械作用下,在加载相同的蛋白时,与O52EWC相比,融合蛋白O14G83提供更佳的生物抛光性能和更高的起球记录,而重量损失相似。

[0761] 表7. 在55℃、pH 6.5下采用O14G83和O52EWC的LOM生物磨损

[0762]

酶	剂量(mg/g 织物)	重量损失(%)	起球记录
O14G83	0.004	0.6 ± 0.0	3.0 ± 0.0
	0.016	1.4 ± 0.1	4.0 ± 0.3
O52EWC	0.004	0.5 ± 0.1	2.7 ± 0.3
	0.016	1.4 ± 0.1	3.5 ± 0.0

[0763] 实例20:在耐洗牢度试验仪中采用O141FN、O141FS、O141FK、O141FU和O52EWC的生物抛光

[0764] 在本实例中使用来自实例12中纯化的蛋白O141FN、O141FS、O141FK、O141FU和实例16中提及的发酵液O52EWC进行生物抛光试验。

[0765] 织物的制备和试验操作与实例13相似。如表8所总结,在每个烧杯中采用20个钢球的高机械作用下,当加载相同的蛋白质时,在这里测试的所有4种融合蛋白都提供比O52EWC更佳的生物磨损性能和更高的起球记录。

[0766] 表8. 在55℃、pH 6.5下采用O141FN、O141FS、O141FK、O141FU和O52EWC的LOM生物磨损

[0767]

酶	剂量 (mg/g 织物)	重量损失 (%)	起球记录
O141FN	0.002	1.1 ± 0.0	3.1 ± 0.2
	0.008	2.1 ± 0.1	3.9 ± 0.2
O141FS	0.002	0.5 ± 0.2	3.3 ± 0.0
	0.008	1.7 ± 0.1	3.9 ± 0.2
O141FK	0.002	0.7 ± 0.2	2.8 ± 0.0
	0.008	1.9 ± 0.0	3.9 ± 0.2
O141FU	0.002	1.0 ± 0.2	2.6 ± 0.2
	0.008	2.0 ± 0.1	4.1 ± 0.2
O52EWC	0.002	0.5 ± 0.1	2.6 ± 0.2
	0.008	1.2 ± 0.2	3.6 ± 0.2

[0768] 实例21:在耐洗牢度试验仪中采用O141FF、O141FH、O141FQ和O52EWC的生物抛光

[0769] 在本实例中使用来自实例12中纯化的蛋白O141FF、O141FH、O141FQ和实例16中提及的发酵液O52EWC进行生物抛光试验。

[0770] 织物的制备和试验操作与实例13相似。如表9所总结,在每个烧杯中采用20个钢球的高机械作用下,当加载相同的蛋白质时,在这里测试的所有3种融合蛋白都提供比O52EWC更佳的生物磨损性能和更高的起球记录;在每个烧杯中采用4个橡胶球的低机械作用下,O141FH比O52EWC和其他融合蛋白提供显著更高的起球记录。

[0771] 表9.在55°C、pH 6.5下采用O141FF、O141FH、O141FQ和O52EWC的LOM生物磨损

[0772]

酶	剂量 (mg/g)	每个烧杯中的球	重量损失 (%)	起球记录
O141FQ	0.002	20 个钢球	0.5 ± 0.1	3.4 ± 0.2
O141FH			0.1 ± 0.2	3.6 ± 0.2
O141FF			0.2 ± 0.1	3.6 ± 0.2
O52EWC			0.0 ± 0.1	2.4 ± 0.2
O141FQ	0.024	4 个橡胶球	1.5 ± 0.1	3.1 ± 0.5
O141FH			1.5 ± 0.1	3.7 ± 0.1
O141FF			1.4 ± 0.1	3.0 ± 0.2
O52EWC			1.1 ± 0.1	3.1 ± 0.3

[0773] 实例22:在耐洗牢度试验仪中采用P24YEZ、P446XD的生物抛光

[0774] 在本实例中使用两种蛋白质-P24YEZ (SEQ ID NO:14) 和从实例12纯化的P446XD,进行生物抛光试验。

[0775] 织物的制备和试验操作与实例13相似。如表10中所总结,在每个烧杯中采用20个

钢球的高机械作用下,在主链P24YEZ中具有Y292W突变时,得到的酶P446XD在具有相同蛋白质剂量的情况下在起球记录方面提供相似的生物抛光性能,并且超出预期的是,为达到等同的起球记录水平,与P24YEZ相比,来自P446XD的重量损失大大增加。例如,为在pH 7.5时达到起球记录水平3.8,P446XD和P24YEZ导致的重量损失分别为1.5%和0.8%。

[0776] 表10.在35℃,pH 7.5下采用P24YEZ、P446XD的LOM生物磨损

[0777]

酶	PH	剂量 (Mg/g 织物)	重量损失 (%)	起球记录
P24YEZ	6.5	0.008	0.2 ± 0.0	3.4 ± 0.2
		0.032	1.2 ± 0.1	3.6 ± 0.2
	7.5	0.008	0.1 ± 0.1	3.3 ± 0.0
		0.032	0.8 ± 0.1	3.8 ± 0.0
P446XD	6.5	0.008	0.4 ± 0.2	3.5 ± 0.0
		0.032	1.4 ± 0.3	3.8 ± 0.0
	7.5	0.008	0.5 ± 0.0	3.1 ± 0.2

[0778]

		0.032	1.5 ± 0.2	3.8 ± 0.0
--	--	-------	-----------	-----------

[0779] 实例23:在耐洗牢度试验仪中采用032JRC、052EWC的生物抛光

[0780] 在本实例中使用两种蛋白质-从实例4纯化的032JRC、以及052EWC,进行生物抛光试验。

[0781] 织物制备和试验操作与实例13相似,不同之处在于,在当前实例中,测试了采用20个钢球的高机械作用和采用2个橡胶球的低机械作用,并且分别在纤维素酶处理30分钟、60分钟和90分钟后收集了样品。如表11所总结,在从30分钟到90分钟的整个过程中,032JRC在高和低机械作用下均比052EWC提供更佳的生物抛光性能,其比052EWC有更高的起球记录。

[0782] 表11.在55℃,pH 6.5下采用032JRC、052EWC的LOM生物磨损

[0783]

酶	剂量 (mg/g)	每个烧杯中的球	时间 (min)	起球记录
空白	0	2 个橡胶球	90	1.5 ± 0.0
O32JRC	0.016	20 个钢球	30	1.9 ± 0.2
			60	3.8 ± 0.0
			90	3.9 ± 0.2
	0.064	2 个橡胶球	30	1.9 ± 0.2
			60	3.4 ± 0.2
			90	3.6 ± 0.2
O52EWC	0.016	20 个钢球	30	1.6 ± 0.2
			60	3.4 ± 0.2
			90	3.3 ± 0.0
	0.064	2 个橡胶球	30	1.6 ± 0.2
			60	3.1 ± 0.2
			90	2.9 ± 0.2

[0784] 在此描述并且要求保护的本发明不局限于在此披露的特定方面的范围,因为这些方面旨在作为本发明若干方面的说明。预期任何等效方面都处于本发明的范围内。实际上,除本申请中所示和描述的那些之外,对于本领域的技术人员而言本发明的多种修改将从前述的描述变得显而易见。这样的修改也旨在落入所附权利要求书的范围内。在有冲突的情况下,以包括定义的本披露为准。

序列表

- <110> 诺维信公司(Novozymes A/S)
- <120> 纤维素酶变体和对其编码的多核苷酸
- <130> 13134-W0-PCT[2]
- <160> 79
- <170> PatentIn 3.5版本
- <210> 1
- <211> 1104
- <212> DNA
- <213> 人工序列
- <220>
- <223> 合成构建体
- <220>
- <221> 外显子
- <222> (1) .. (89)
- <220>
- <221> 信号肽
- <222> (1) .. (63)
- <220>
- <221> CDS
- <222> (1) .. (89)
- <220>
- <221> 成熟肽
- <222> (64) .. (1101)
- <220>
- <221> 内含子
- <222> (90) .. (167)
- <220>
- <221> 外显子
- <222> (168) .. (416)
- <220>
- <221> CDS
- <222> (168) .. (416)
- <220>
- <221> 内含子
- <222> (417) .. (485)
- <220>
- <221> 外显子

<222> (486)..(1101)
 <220>
 <221> CDS
 <222> (486)..(1101)
 <400> 1
 atg cgc tct act ccc gtt ctt cgc aca acc ctg gcc gct gca ctt cct 48
 Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro
 -20 -15 -10
 ctg gtc gcc tcc gcg gcc agt ggc agt ggc cag tcc acg ag 89
 Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg
 -5 -1 1 5
 gtatgctcc ctcaatgcca atgcctcacc aaacagcta ggtgtccagg acgccgccca 149
 tgctgactcg ttcccag a tac tgg gac tgc tgc aag ccg tgc tgc gct tgg 201
 Tyr Trp Asp Cys Cys Lys Pro Ser Cys Ala Trp
 10 15 20
 ccc ggg aag gcc gcc gtc agc caa ccg gtc tac gcg tgc gat gcc aac 249
 Pro Gly Lys Ala Ala Val Ser Gln Pro Val Tyr Ala Cys Asp Ala Asn
 25 30 35
 ttc cag cgc ctg tcc gac ttc aat gtc cag tgc ggc tgc aac ggc ggc 297
 Phe Gln Arg Leu Ser Asp Phe Asn Val Gln Ser Gly Cys Asn Gly Gly
 40 45 50
 tgc gcc tac tcc tgc gcc gac cag act ccc tgg gcg gtg aac gac aat 345
 Ser Ala Tyr Ser Cys Ala Asp Gln Thr Pro Trp Ala Val Asn Asp Asn
 55 60 65
 ctc gcc tac ggc ttc gcc gcg acg agc atc gcc ggc ggc tcc gaa tcc 393
 Leu Ala Tyr Gly Phe Ala Ala Thr Ser Ile Ala Gly Gly Ser Glu Ser
 70 75 80
 tgc tgg tgc tgc gcc tgc tac gc gtaagtctc tctgccagct accaggaagt 446
 Ser Trp Cys Cys Ala Cys Tyr Ala
 85 90
 tggattagcg cgagctaacc tcaactccate aactccag g etc acc ttc act tcc 501
 Leu Thr Phe Thr Ser
 95
 ggt ccc gtc gcc ggc aag aca atg gtg gtg cag tca acg agc act ggc 549
 Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr Ser Thr Gly
 100 105 110
 ggc gac ctg gga agt aac cat ttc gat atc gcc atg ccc ggc ggc ggc 597
 Gly Asp Leu Gly Ser Asn His Phe Asp Ile Ala Met Pro Gly Gly Gly
 115 120 125

gtg ggc atc ttc aac ggc tgc agc tcg cag ttc ggc ggc ctc ccc ggc	645
Val Gly Ile Phe Asn Gly Cys Ser Ser Gln Phe Gly Gly Leu Pro Gly	
130 135 140 145	
gct caa tac ggc ggc att tcg tcg cgc gac cag tgc gat tcc ttc ccc	693
Ala Gln Tyr Gly Gly Ile Ser Ser Arg Asp Gln Cys Asp Ser Phe Pro	
150 155 160	
gcg ccg ctc aag ccc ggc tgc cag tgg cgg ttt gac tgg ttc cag aac	741
Ala Pro Leu Lys Pro Gly Cys Gln Trp Arg Phe Asp Trp Phe Gln Asn	
165 170 175	
gcc gac aac ccg acg ttc acg ttc cag cag gtg cag tgc ccc gcc gag	789
Ala Asp Asn Pro Thr Phe Thr Phe Gln Gln Val Gln Cys Pro Ala Glu	
180 185 190	
atc gtt gcc cgc tcc ggc tgc aag cgc aac gac gat ggt aac ttc cct	837
Ile Val Ala Arg Ser Gly Cys Lys Arg Asn Asp Asp Gly Asn Phe Pro	
195 200 205	
gtg ttc acc cct cct tcg gga ggc cag tcc tcc tcg tcc tcc tcg tcc	885
Val Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Ser Ser Ser Ser	
210 215 220 225	
tcg tcg gcc aag ccg aca tcg acg tcg acg tcc act act tcg acc aaa	933
Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Thr Thr Ser Thr Lys	
230 235 240	
gca acc tcc aca act tcc acc gcc tcg tcc cag acg tcg tcg tcc acc	981
Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser Gln Thr Ser Ser Ser Thr	
245 250 255	
gga gga ggc tgc gca gcg cag cgt tgg gct cag tgc ggt ggc atc ggc	1029
Gly Gly Gly Cys Ala Ala Gln Arg Trp Ala Gln Cys Gly Gly Ile Gly	
260 265 270	
ttc agc gga tgc acc acc tgt gtc tct ggc acc acc tgc aac aag cag	1077
Phe Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys Asn Lys Gln	
275 280 285	
aac gac tgg tac tcg cag tgc ctc tga	1104
Asn Asp Trp Tyr Ser Gln Cys Leu	
290 295	
<210> 2	
<211> 318	
<212> PRT	
<213> 人工序列	
<220>	
<223> 合成构建体	

<400> 2

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

Thr	Thr	Cys	Asn	Lys	Gln	Asn	Asp	Trp	Tyr	Ser	Gln	Cys	Leu
	285					290					295		

<210> 3
 <211> 1297
 <212> DNA
 <213> 人工序列
 <220>
 <223> 合成构建体
 <220>
 <221> 信号肽
 <222> (1) .. (63)
 <220>
 <221> 外显子
 <222> (1) .. (89)
 <220>
 <221> CDS
 <222> (1) .. (89)
 <220>
 <221> 成熟肽
 <222> (64) .. (1101)
 <220>
 <221> 内含子
 <222> (90) .. (167)
 <220>
 <221> 外显子
 <222> (168) .. (416)
 <220>
 <221> CDS
 <222> (168) .. (416)
 <220>
 <221> 内含子
 <222> (417) .. (485)
 <220>
 <221> 外显子
 <222> (486) .. (1101)
 <220>
 <221> CDS
 <222> (486) .. (1101)
 <400> 3

atg cgc tct act ccc gtt ctt cgc aca acc ctg gcc gct gca ctt cct 48
 Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro
 -20 -15 -10
 ctg gtc gcc tcc gcg gcc agt ggc agt ggc cag tcc acg ag 89
 Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg
 -5 -1 1 5
 gtatgcgtcc ctcaatgcga atgcctcacc aaacgagcta ggtgtccagg acgccgccca 149
 tgctgactcg tttcccag a tac tgg gac tgc tgc aag ccg tcg tgc gct tgg 201
 Tyr Trp Asp Cys Cys Lys Pro Ser Cys Ala Trp
 10 15 20
 ccc ggg aag gcc gcc gtc agc caa ccg gtc tac gcg tgc gat gcc aac 249
 Pro Gly Lys Ala Ala Val Ser Gln Pro Val Tyr Ala Cys Asp Ala Asn
 25 30 35
 ttc cag cgc ctg tcc gac ttc aat gtc cag tcg ggc tgc aac ggc ggc 297
 Phe Gln Arg Leu Ser Asp Phe Asn Val Gln Ser Gly Cys Asn Gly Gly
 40 45 50
 tcg gcc tac tcc tgc gcc gac cag act ccc tgg gcg gtg aac gac aat 345
 Ser Ala Tyr Ser Cys Ala Asp Gln Thr Pro Trp Ala Val Asn Asp Asn
 55 60 65
 ctc gcc tac ggc ttc gcc gcg acg agc atc gcc ggc ggg tcc gaa tcc 393
 Leu Ala Tyr Gly Phe Ala Ala Thr Ser Ile Ala Gly Gly Ser Glu Ser
 70 75 80
 tcg tgg tgc tgc gcc tgc tac gc gtaagtctc tctgccagct accaggaagt 446
 Ser Trp Cys Cys Ala Cys Tyr Ala
 85 90
 tggattagcg cgagctaacc tcaactccatc aactccag g ctc acc ttc act tcc 501
 Leu Thr Phe Thr Ser
 95
 ggt ccc gtc gcc ggc aag aca atg gtg gtg cag tca acg agc act ggc 549
 Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr Ser Thr Gly
 100 105 110
 ggc gac ctg gga agt aac cat ttc gat atc gcc atg ccc ggc ggc ggc 597
 Gly Asp Leu Gly Ser Asn His Phe Asp Ile Ala Met Pro Gly Gly Gly
 115 120 125
 gtg ggc atc ttc aac ggc tgc agc tcg cag ttc ggc ggc ctc ccc ggc 645
 Val Gly Ile Phe Asn Gly Cys Ser Ser Gln Phe Gly Gly Leu Pro Gly
 130 135 140 145
 gct caa tac ggc ggc att tcg tcg cgc gac cag tgc gat tcc ttc ccc 693
 Ala Gln Tyr Gly Gly Ile Ser Ser Arg Asp Gln Cys Asp Ser Phe Pro

150	155	160	
gcg ccg ctc aag ccc ggc tgc cag tgg cgg ttt gac tgg ttc cag aac			741
Ala Pro Leu Lys Pro Gly Cys Gln Trp Arg Phe Asp Trp Phe Gln Asn			
165	170	175	
gcc gac aac ccg acg ttc acg ttc cag cag gtg cag tgc ccc gcc gag			789
Ala Asp Asn Pro Thr Phe Thr Phe Gln Gln Val Gln Cys Pro Ala Glu			
180	185	190	
atc gtt gcc cgc tcc ggc tgc aag cgc aac gac gat ggt aac ttc cct			837
Ile Val Ala Arg Ser Gly Cys Lys Arg Asn Asp Asp Gly Asn Phe Pro			
195	200	205	
gtg ttc acc cct cct tcg gga ggc cag tcc tcc tcg tcc tcc tcg tcc			885
Val Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Ser Ser Ser Ser Ser			
210	215	220	225
tcg tcg gcc aag ccg aca tcg acg tcg acg tcc act act tcg acc aaa			933
Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Thr Thr Ser Thr Lys			
230	235	240	
gca acc tcc aca act tcc acc gcc tcg tcc cag acg tcg tcg tcc acc			981
Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser Gln Thr Ser Ser Ser Thr			
245	250	255	
gga gga ggc tgc acg tct cag aag tgg gct cag tgc ggt gcc atc gcc			1029
Gly Gly Gly Cys Thr Ser Gln Lys Trp Ala Gln Cys Gly Gly Ile Gly			
260	265	270	
ttc agc gga tgc acc acc tgt gtc tct ggc acc acc tgc cag aag ttg			1077
Phe Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys Gln Lys Leu			
275	280	285	
aac gac tac tac tcg cag tgc ctc taaacagctt ttcgcacgag gtggcgggac			1131
Asn Asp Tyr Tyr Ser Gln Cys Leu			
290	295		
ggagcaagga gaccgtcaac ttcgtcatgc atatTTTTTg agcgetcaat acatacataa			1191
ccttcgattc ttgtacatag cacgccgta cacatctcac accgactttg gggcggaat			1251
caggccccgtt ttacccccgag ttcgetgcca ctgetggttg atttcc			1297
<210> 4			
<211> 318			
<212> PRT			
<213> 人工序列			
<220>			
<223> 合成构建体			
<400> 4			
Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro			

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

<210> 5
 <211> 1104
 <212> DNA
 <213> 人工序列
 <220>
 <223> 合成构建体
 <220>
 <221> 信号肽
 <222> (1) .. (63)
 <220>
 <221> 外显子
 <222> (1) .. (89)
 <220>
 <221> CDS
 <222> (1) .. (89)
 <220>
 <221> 成熟肽
 <222> (64) .. (1101)
 <220>
 <221> 内含子
 <222> (90) .. (167)
 <220>
 <221> 外显子
 <222> (168) .. (416)
 <220>
 <221> CDS
 <222> (168) .. (416)
 <220>
 <221> 内含子
 <222> (417) .. (485)
 <220>
 <221> 外显子
 <222> (486) .. (1101)
 <220>
 <221> CDS
 <222> (486) .. (1101)

<400> 5

atg cgc tct act ccc gtt ctt cgc aca acc ctg gcc gct gca ctt cct 48
 Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro

-20	-15	-10	
ctg gtc gcc tcc gcg gcc agt ggc agt ggc cag tcc acg ag			89
Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg			
-5	-1 1	5	
gtatgcgtcc ctcaatgcga atgcctcacc aaacgagcta ggtgtccagg acgccgccca			149
tgetgactcg tttcccag a tac tgg gac tgc tgc aag ccg tcg tgc gct tgg			201
	Tyr Trp Asp Cys Cys Lys Pro Ser Cys Ala Trp		
	10	15	20
ccc ggg aag gcc gcc gtc age caa ccg gtc tac gcg tgc gat gcc aac			249
Pro Gly Lys Ala Ala Val Ser Gln Pro Val Tyr Ala Cys Asp Ala Asn			
	25	30	35
ttc cag cgc ctg tcc gac ttc aat gtc cag tgg ggc tgc aac ggc ggc			297
Phe Gln Arg Leu Ser Asp Phe Asn Val Gln Ser Gly Cys Asn Gly Gly			
	40	45	50
tcg gcc tac tcc tgc gcc gac cag act ccc tgg gcg gtg aac gac aat			345
Ser Ala Tyr Ser Cys Ala Asp Gln Thr Pro Trp Ala Val Asn Asp Asn			
	55	60	65
ctc gcc tac ggc ttc gcc gcg acg age atc gcc ggc ggg tcc gaa tcc			393
Leu Ala Tyr Gly Phe Ala Ala Thr Ser Ile Ala Gly Gly Ser Glu Ser			
	70	75	80
tcg tgg tgc tgc gcc tgc tac gc gtaagtctc tctgccagct accaggaagt			446
Ser Trp Cys Cys Ala Cys Tyr Ala			
85	90		
tggattagcg cgagctaacc tcaactccatc aactccag g ctc acc ttc act tcc			501
		Leu Thr Phe Thr Ser	
		95	
ggt ccc gtc gcc ggc aag aca atg gtg gtg cag tca acg age act ggc			549
Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr Ser Thr Gly			
	100	105	110
ggc gac ctg gga agt aac cat ttc gat atc gcc atg ccc ggc ggc ggc			597
Gly Asp Leu Gly Ser Asn His Phe Asp Ile Ala Met Pro Gly Gly Gly			
	115	120	125
gtg ggc atc ttc aac ggc tgc age tgg cag ttc ggc ggc ctc ccc ggc			645
Val Gly Ile Phe Asn Gly Cys Ser Ser Gln Phe Gly Gly Leu Pro Gly			
	130	135	140
gct caa tac ggc ggc att tgg tgg cgc gac cag tgc gat tcc ttc ccc			693
Ala Gln Tyr Gly Gly Ile Ser Ser Arg Asp Gln Cys Asp Ser Phe Pro			
	150	155	160
gcg ccg ctc aag ccc ggc tgc cag tgg cgg ttt gac tgg ttc cag aac			741

Ala Pro Leu Lys Pro Gly Cys Gln Trp Arg Phe Asp Trp Phe Gln Asn	
165	170
175	
gcc gac aac ccg acg ttc acg ttc cag cag gtg cag tgc ccc gcc gag	789
Ala Asp Asn Pro Thr Phe Thr Phe Gln Gln Val Gln Cys Pro Ala Glu	
180	185
190	
atc gtt gcc cgc tcc ggc tgc aag cgc aac gac gat ggt aac ttc cct	837
Ile Val Ala Arg Ser Gly Cys Lys Arg Asn Asp Asp Gly Asn Phe Pro	
195	200
205	
gtg ttc acc cct cct tcg gga ggc cag tcc tcc tcg tcc tcc tcg tcc	885
Val Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Ser Ser Ser Ser Ser	
210	215
220	225
tcg tcg gcc aag ccg aca tcg acg tcg acg tcc act act tcg acc aaa	933
Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Thr Thr Ser Thr Lys	
230	235
240	
gca acc tcc aca act tcc acc gcc tcg tcc cag acg tcg tcg tcc acc	981
Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser Gln Thr Ser Ser Ser Thr	
245	250
255	
gga gga ggc tgc gca gcg cag cgt tgg gct cag tgc ggt ggc atc ggc	1029
Gly Gly Gly Cys Ala Ala Gln Arg Trp Ala Gln Cys Gly Gly Ile Gly	
260	265
270	
ttc agc gga tgc acc acc tgt gtc tct ggc acc acc tgc aac aag cag	1077
Phe Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys Asn Lys Gln	
275	280
285	
aac gac tac tac tcg cag tgc ctc tga	1104
Asn Asp Tyr Tyr Ser Gln Cys Leu	
290	295
<210> 6	
<211> 318	
<212> PRT	
<213> 人工序列	
<220>	
<223> 合成构建体	
<400> 6	
Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro	
-20	-15
-10	
Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg Tyr Trp	
-5	-1 1
5	10
Asp Cys Cys Lys Pro Ser Cys Ala Trp Pro Gly Lys Ala Ala Val Ser	
15	20
25	

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

<210> 7

<211> 1071

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<220>

<221> 信号肽

<222> (1) .. (63)

<220>

<221> 外显子

<222> (1) .. (89)

<220>

<221> CDS

<222> (1) .. (89)

<220>

<221> 成熟肽

<222> (64) .. (1068)

<220>

<221> 内含子

<222> (90) .. (167)

<220>

<221> 外显子

<222> (168) .. (416)

<220>

<221> CDS

<222> (168) .. (416)

<220>

<221> 内含子

<222> (417) .. (485)

<220>

<221> 外显子

<222> (486) .. (1068)

<220>

<221> CDS

<222> (486) .. (1068)

<400> 7

atg cgc tct act ccc gtt ctt cgc aca acc ctg gcc gct gca ctt cct 48

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

-20

-15

-10

ctg gtc gcc tcc gcg gcc agt ggc agt ggc cag tcc acg ag 89

Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg

-5

-1 1

5

gtatgcgtcc ctcaatgcga atgcctcacc aaacgagcta ggtgtccagg acgccgccca 149

tgctgactcg tttcccag a tac tgg gac tgc tgc aag ccg tgc tgc gct tgg	201
Tyr Trp Asp Cys Cys Lys Pro Ser Cys Ala Trp	
10 15 20	
ccc ggg aag gcc gcc gtc agc caa ccg gtc tac gcg tgc gat gcc aac	249
Pro Gly Lys Ala Ala Val Ser Gln Pro Val Tyr Ala Cys Asp Ala Asn	
25 30 35	
ttc cag cgc ctg tcc gac ttc aat gtc cag tgc ggc tgc aac ggc ggc	297
Phe Gln Arg Leu Ser Asp Phe Asn Val Gln Ser Gly Cys Asn Gly Gly	
40 45 50	
tgc gcc tac tcc tgc gcc gac cag act ccc tgg gcg gtg aac gac aat	345
Ser Ala Tyr Ser Cys Ala Asp Gln Thr Pro Trp Ala Val Asn Asp Asn	
55 60 65	
ctc gcc tac ggc ttc gcc gcg acg agc atc gcc ggc ggc tcc gaa tcc	393
Leu Ala Tyr Gly Phe Ala Ala Thr Ser Ile Ala Gly Gly Ser Glu Ser	
70 75 80	
tgc tgg tgc tgc gcc tgc tac gc gtaagtctc tctgccagct accaggaagt	446
Ser Trp Cys Cys Ala Cys Tyr Ala	
85 90	
tggattagcg cgagctaacc tcaactccatc aactccag g ctc acc ttc act tcc	501
Leu Thr Phe Thr Ser	
95	
ggt ccc gtc gcc ggc aag aca atg gtg gtg cag tca acg agc act ggc	549
Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr Ser Thr Gly	
100 105 110	
ggc gac ctg gga agt aac cat ttc gat atc gcc atg ccc ggc ggc ggc	597
Gly Asp Leu Gly Ser Asn His Phe Asp Ile Ala Met Pro Gly Gly Gly	
115 120 125	
gtg ggc atc ttc aac ggc tgc agc tgc cag ttc ggc ggc ctc ccc ggc	645
Val Gly Ile Phe Asn Gly Cys Ser Ser Gln Phe Gly Gly Leu Pro Gly	
130 135 140 145	
gct caa tac ggc ggc att tgc tgc cgc gac cag tgc gat tcc ttc ccc	693
Ala Gln Tyr Gly Gly Ile Ser Ser Arg Asp Gln Cys Asp Ser Phe Pro	
150 155 160	
gcg ccg ctc aag ccc ggc tgc cag tgg cgg ttt gac tgg ttc cag aac	741
Ala Pro Leu Lys Pro Gly Cys Gln Trp Arg Phe Asp Trp Phe Gln Asn	
165 170 175	
gcc gac aac ccg acg ttc acg ttc cag cag gtg cag tgc ccc gcc gag	789
Ala Asp Asn Pro Thr Phe Thr Phe Gln Gln Val Gln Cys Pro Ala Glu	
180 185 190	

atc gtt gcc cgc tcc ggc tgc aag gcg aac gac gac tcc agc ttc ccc 837
 Ile Val Ala Arg Ser Gly Cys Lys Ala Asn Asp Asp Ser Ser Phe Pro
 195 200 205
 gtc ttc acc ccg ggc acc tcg ggc tcg tcc tcg acc gcc aag ccc gct 885
 Val Phe Thr Pro Gly Thr Ser Gly Ser Ser Ser Thr Ala Lys Pro Ala
 210 215 220 225
 tcc tct tcg acc cgg gcc acg tcc acg aag acc tcg gct cct gcg acc 933
 Ser Ser Ser Thr Arg Ala Thr Ser Thr Lys Thr Ser Ala Pro Ala Thr
 230 235 240
 cag acg tcg tcg act gga ggc ggc tgc gtc gcg cag aag tgg gcg caa 981
 Gln Thr Ser Ser Thr Gly Gly Gly Cys Val Ala Gln Lys Trp Ala Gln
 245 250 255
 tgt ggt ggt agc gga ttt tet ggc tgc acc acc tgc gct gcc ggg tcg 1029
 Cys Gly Gly Ser Gly Phe Ser Gly Cys Thr Thr Cys Ala Ala Gly Ser
 260 265 270
 acg tgc acg aag cag aac gac tac tac tcg cag tgc ctc taa 1071
 Thr Cys Thr Lys Gln Asn Asp Tyr Tyr Ser Gln Cys Leu
 275 280 285
 <210> 8
 <211> 307
 <212> PRT
 <213> 人工序列
 <220>
 <223> 合成构建体
 <400> 8
 Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro
 -20 -15 -10
 Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg Tyr Trp
 -5 -1 1 5 10
 Asp Cys Cys Lys Pro Ser Cys Ala Trp Pro Gly Lys Ala Ala Val Ser
 15 20 25
 Gln Pro Val Tyr Ala Cys Asp Ala Asn Phe Gln Arg Leu Ser Asp Phe
 30 35 40
 Asn Val Gln Ser Gly Cys Asn Gly Gly Ser Ala Tyr Ser Cys Ala Asp
 45 50 55
 Gln Thr Pro Trp Ala Val Asn Asp Asn Leu Ala Tyr Gly Phe Ala Ala
 60 65 70 75
 Thr Ser Ile Ala Gly Gly Ser Glu Ser Ser Trp Cys Cys Ala Cys Tyr
 80 85 90

Ala Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Thr Met Val Val
95 100 105

Gln Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe Asp Ile
110 115 120

Ala Met Pro Gly Gly Gly Val Gly Ile Phe Asn Gly Cys Ser Ser Gln
125 130 135

Phe Gly Gly Leu Pro Gly Ala Gln Tyr Gly Gly Ile Ser Ser Arg Asp
140 145 150 155

Gln Cys Asp Ser Phe Pro Ala Pro Leu Lys Pro Gly Cys Gln Trp Arg
160 165 170

Phe Asp Trp Phe Gln Asn Ala Asp Asn Pro Thr Phe Thr Phe Gln Gln
175 180 185

Val Gln Cys Pro Ala Glu Ile Val Ala Arg Ser Gly Cys Lys Ala Asn
190 195 200

Asp Asp Ser Ser Phe Pro Val Phe Thr Pro Gly Thr Ser Gly Ser Ser
205 210 215

Ser Thr Ala Lys Pro Ala Ser Ser Ser Thr Arg Ala Thr Ser Thr Lys
220 225 230 235

Thr Ser Ala Pro Ala Thr Gln Thr Ser Ser Thr Gly Gly Gly Cys Val
240 245 250

Ala Gln Lys Trp Ala Gln Cys Gly Gly Ser Gly Phe Ser Gly Cys Thr
255 260 265

Thr Cys Ala Ala Gly Ser Thr Cys Thr Lys Gln Asn Asp Tyr Tyr Ser
270 275 280

Gln Cys Leu
285

<210> 9
<211> 1071
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<220>
<221> 信号肽
<222> (1) .. (63)
<220>
<221> 外显子
<222> (1) .. (89)
<220>

<221> CDS
 <222> (1) .. (89)
 <220>
 <221> 成熟肽
 <222> (64) .. (1068)
 <220>
 <221> 内含子
 <222> (90) .. (167)
 <220>
 <221> 外显子
 <222> (168) .. (416)
 <220>
 <221> CDS
 <222> (168) .. (416)
 <220>
 <221> 内含子
 <222> (417) .. (485)
 <220>
 <221> 外显子
 <222> (486) .. (1068)
 <220>
 <221> CDS
 <222> (486) .. (1068)
 <400> 9
 atg cgc tct act ccc gtt ctt cgc aca acc ctg gcc gct gca ctt cct 48
 Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro
 -20 -15 -10
 ctg gtc gcc tcc gcg gcc agt ggc agt ggc cag tcc acg ag 89
 Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg
 -5 -1 1 5
 gtatgctgcc ctcaatgcga atgectcacc aaacgageta ggtgtccagg acgccgccca 149
 tgctgactcg ttteccag a tac tgg gac tgc tgc aag ceg teg tgc gct tgg 201
 Tyr Trp Asp Cys Cys Lys Pro Ser Cys Ala Trp
 10 15 20
 ccc ggg aag gcc gcc gtc agc caa ceg gtc tac gcg tgc gat gcc aac 249
 Pro Gly Lys Ala Ala Val Ser Gln Pro Val Tyr Ala Cys Asp Ala Asn
 25 30 35
 ttc cag cgc ctg tcc gac ttc aat gtc cag teg ggc tgc aac ggc ggc 297
 Phe Gln Arg Leu Ser Asp Phe Asn Val Gln Ser Gly Cys Asn Gly Gly

tcg gcc tac tcc tgc gcc gac cag act ccc tgg gcg gtg aac gac aat	345		
Ser Ala Tyr Ser Cys Ala Asp Gln Thr Pro Trp Ala Val Asn Asp Asn			
55	60		
ctc gcc tac ggc ttc gcc gcg acg agc atc gcc ggc ggg tcc gaa tcc	393		
Leu Ala Tyr Gly Phe Ala Ala Thr Ser Ile Ala Gly Gly Ser Glu Ser			
70	75	80	
tcg tgg tgc tgc gcc tgc tac gc gtaagtctc tctgccagct accaggaagt	446		
Ser Trp Cys Cys Ala Cys Tyr Ala			
85	90		
tggattagcg cgagctaacc tcaactccate acaactccag g ctc acc ttc act tcc	501		
Leu Thr Phe Thr Ser			
95			
ggc ccc gtc gcc ggc aag aca atg gtg gtg cag tca acg agc act ggc	549		
Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr Ser Thr Gly			
100	105	110	
ggc gac ctg gga agt aac cat ttc gat atc gcc atg ccc ggc ggc ggc	597		
Gly Asp Leu Gly Ser Asn His Phe Asp Ile Ala Met Pro Gly Gly Gly			
115	120	125	
gtg ggc atc ttc aac ggc tgc agc tcg cag ttc ggc ggc ctc ccc ggc	645		
Val Gly Ile Phe Asn Gly Cys Ser Ser Gln Phe Gly Gly Leu Pro Gly			
130	135	140	145
gct caa tac ggc ggc att tcg tcg cgc gac cag tgc gat tcc ttc ccc	693		
Ala Gln Tyr Gly Gly Ile Ser Ser Arg Asp Gln Cys Asp Ser Phe Pro			
150	155	160	
gcg ccg ctc aag ccc ggc tgc cag tgg cgg ttt gac tgg ttc cag aac	741		
Ala Pro Leu Lys Pro Gly Cys Gln Trp Arg Phe Asp Trp Phe Gln Asn			
165	170	175	
gcc gac aac ccg acg ttc acg ttc cag cag gtg cag tgc ccc gcc gag	789		
Ala Asp Asn Pro Thr Phe Thr Phe Gln Gln Val Gln Cys Pro Ala Glu			
180	185	190	
atc gtt gcc cgc tcc ggc tgc aag gcg aac gac gac tcc agc ttc ccc	837		
Ile Val Ala Arg Ser Gly Cys Lys Ala Asn Asp Asp Ser Ser Phe Pro			
195	200	205	
gtc ttc acc ccg ggc acc tcg ggc tcg tcc tcg acc gcc aag ccc gct	885		
Val Phe Thr Pro Gly Thr Ser Gly Ser Ser Ser Thr Ala Lys Pro Ala			
210	215	220	225
tcc tct tcg acc cgg gcc acg tcc acg aag acc tcg gct cct gcg acc	933		
Ser Ser Ser Thr Arg Ala Thr Ser Thr Lys Thr Ser Ala Pro Ala Thr			

	230	235	240	
	cag acg tcg tcg act gga ggc ggc tgc gtc gcg cag aag tgg gcg caa			981
	Gln Thr Ser Ser Thr Gly Gly Gly Cys Val Ala Gln Lys Trp Ala Gln			
	245	250	255	
	tgt ggt ggt agc gga ttt tct ggc tgc acc acc tgc gct gcc ggg tcg			1029
	Cys Gly Gly Ser Gly Phe Ser Gly Cys Thr Thr Cys Ala Ala Gly Ser			
	260	265	270	
	acg tgc acg aag cag aac gac tgg tac tcg cag tgc ctc taa			1071
	Thr Cys Thr Lys Gln Asn Asp Trp Tyr Ser Gln Cys Leu			
	275	280	285	
	<210> 10			
	<211> 307			
	<212> PRT			
	<213> 人工序列			
	<220>			
	<223> 合成构建体			
	<400> 10			
	Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro			
	-20	-15	-10	
	Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg Tyr Trp			
	-5	-1 1	5	10
	Asp Cys Cys Lys Pro Ser Cys Ala Trp Pro Gly Lys Ala Ala Val Ser			
	15	20	25	
	Gln Pro Val Tyr Ala Cys Asp Ala Asn Phe Gln Arg Leu Ser Asp Phe			
	30	35	40	
	Asn Val Gln Ser Gly Cys Asn Gly Gly Ser Ala Tyr Ser Cys Ala Asp			
	45	50	55	
	Gln Thr Pro Trp Ala Val Asn Asp Asn Leu Ala Tyr Gly Phe Ala Ala			
	60	65	70	75
	Thr Ser Ile Ala Gly Gly Ser Glu Ser Ser Trp Cys Cys Ala Cys Tyr			
	80	85	90	
	Ala Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Thr Met Val Val			
	95	100	105	
	Gln Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe Asp Ile			
	110	115	120	
	Ala Met Pro Gly Gly Gly Val Gly Ile Phe Asn Gly Cys Ser Ser Gln			
	125	130	135	
	Phe Gly Gly Leu Pro Gly Ala Gln Tyr Gly Gly Ile Ser Ser Arg Asp			
	140	145	150	155

Gln Cys Asp Ser Phe Pro Ala Pro Leu Lys Pro Gly Cys Gln Trp Arg
 160 165 170
 Phe Asp Trp Phe Gln Asn Ala Asp Asn Pro Thr Phe Thr Phe Gln Gln
 175 180 185
 Val Gln Cys Pro Ala Glu Ile Val Ala Arg Ser Gly Cys Lys Ala Asn
 190 195 200
 Asp Asp Ser Ser Phe Pro Val Phe Thr Pro Gly Thr Ser Gly Ser Ser
 205 210 215
 Ser Thr Ala Lys Pro Ala Ser Ser Ser Thr Arg Ala Thr Ser Thr Lys
 220 225 230 235
 Thr Ser Ala Pro Ala Thr Gln Thr Ser Ser Thr Gly Gly Gly Cys Val
 240 245 250
 Ala Gln Lys Trp Ala Gln Cys Gly Gly Ser Gly Phe Ser Gly Cys Thr
 255 260 265
 Thr Cys Ala Ala Gly Ser Thr Cys Thr Lys Gln Asn Asp Trp Tyr Ser
 270 275 280

Gln Cys Leu
 285

<210> 11

<211> 1104

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<400> 11

atgcgctcta ctcccgttct tcgcacaacc ctggccgctg cacttcctct ggtcgcctcc 60
 gcggccagtg gcagtgcca gtccacgagg tatgcgtccc tcaatgcgaa tgcctcacca 120
 aacgagctag gtgtccagga cgccgcccat getgactcgt ttcccagata ctgggactgc 180
 tgcaagccgt cgtgcgcttg gcccggaag gccgcccga gccaaccggt ctacgcgtgc 240
 gatgccaact tccagcgcct gtccgacttc aatgtccagt cgggctgcaa cggcggtcgc 300
 gcctactcct gcgccacca gactccctgg gcggtgaacg acaatctcgc ctacggcttc 360
 gccgcgacga gcatcgccgg cgggtccgaa tctcgtggt getgcgctg ctacgcgtaa 420
 gtccctctctg ccagctacca ggaagttgga ttagecgcag ctaacctac tccatcacac 480
 tccaggetca cttcacttc cgggtcccgc gccggcaaga caatggtggt gcagtcaacg 540
 agcactggcg gcgacctggg aagtaacct ttgatatcg ccatgcccg cggcggcgtg 600
 ggcattctca acggtgcag ctgcgagttc ggcgccctcc ccggcgcga ataccggcgc 660
 atttcgtcgc gcgaccagtg cgattccttc cccgcgccgc tcaagcccgg ctgccagtgg 720
 cggtttgact ggttccagaa cgccgacaac ccgacgttca cgttccagca ggtgcagtgc 780
 cccgccgaga tcgttcccc ctccggtcgc aagcgcacg acgatggtaa ctccctgtg 840

ttcaccctc cttcgggagg ccagtcctcc tcgtcctcct cgtcctcgtc ggccaagccg 900
 acatcgacgt cgacgtccac tacttgcacc aaagcaacct ccacaacttc caccgcctcg 960
 tcccagacgt cgtcgtccac cggaggaggc tgtgcagcgc agcgttgggc acagtgtggt 1020
 ggtatcggat tctcgggatg tactacatgt gtgtcgggca cgacgtgtaa caagcagaac 1080
 gactgggtact cccagtgtct ctga 1104

<210> 12

<211> 299

<212> PRT

<213> 土生梭孢壳霉

<400> 12

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

225 230 235 240
 Gly Thr Gly Thr Pro Thr Ser Thr Ala Pro Gly Ser Gly Gln Thr Ser
 245 250 255
 Pro Gly Gly Gly Ser Gly Cys Thr Ser Gln Lys Trp Ala Gln Cys Gly
 260 265 270
 Gly Ile Gly Phe Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys
 275 280 285
 Gln Lys Leu Asn Asp Tyr Tyr Ser Gln Cys Leu
 290 295
 <210> 13
 <211> 295
 <212> PRT
 <213> 大孢圆孢霉
 <400> 13
 Ala Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys Cys Lys Pro Ser Cys
 1 5 10 15
 Ser Trp Pro Gly Lys Ala Ser Val Asn Gln Pro Val Phe Ala Cys Ser
 20 25 30
 Ala Asn Phe Gln Arg Ile Ser Asp Pro Asn Val Lys Ser Gly Cys Asp
 35 40 45
 Gly Gly Ser Ala Tyr Ala Cys Ala Asp Gln Thr Pro Trp Ala Val Asn
 50 55 60
 Asp Asn Phe Ser Tyr Gly Phe Ala Ala Thr Ser Ile Ser Gly Gly Asn
 65 70 75 80
 Glu Ala Ser Trp Cys Cys Gly Cys Tyr Glu Leu Thr Phe Thr Ser Gly
 85 90 95
 Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr Ser Thr Gly Gly
 100 105 110
 Asp Leu Gly Thr Asn His Phe Asp Leu Ala Met Pro Gly Gly Gly Val
 115 120 125
 Gly Ile Phe Asp Gly Cys Ser Pro Gln Phe Gly Gly Leu Ala Gly Asp
 130 135 140
 Arg Tyr Gly Gly Val Ser Ser Arg Ser Gln Cys Asp Ser Phe Pro Ala
 145 150 155 160
 Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe Asp Trp Phe Lys Asn Ala
 165 170 175
 Asp Asn Pro Thr Phe Thr Phe Arg Gln Val Gln Cys Pro Ser Glu Leu
 180 185 190
 Val Ala Arg Thr Gly Cys Arg Arg Asn Asp Asp Gly Asn Phe Pro Val

195	200	205
Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Ser Ser Ser Ser Ser Ser		
210	215	220
Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Thr Thr Ser Thr Lys Ala		
225	230	235
Thr Ser Thr Thr Ser Thr Ala Ser Ser Gln Thr Ser Ser Ser Thr Gly		
	245	250
Gly Gly Cys Ala Ala Gln Arg Trp Ala Gln Cys Gly Gly Ile Gly Phe		
	260	265
Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys Asn Lys Gln Asn		
	275	280
Asp Trp Tyr Ser Gln Cys Leu		285
	290	295
<210> 14		
<211> 293		
<212> PRT		
<213> 四孢脉孢菌		
<400> 14		
Met Arg Ser Ser Thr Val Leu Gln Thr Gly Leu Val Ala Ala Leu Pro		
1	5	10
Phe Ala Val Gln Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg Tyr Trp		
	20	25
Asp Cys Cys Lys Pro Ser Cys Ser Trp Ser Gly Lys Ala Ser Val Asn		
	35	40
Arg Pro Val Leu Ala Cys Asp Ala Asn Asn Asn Pro Leu Ser Asp Ala		
	50	55
Ser Val Lys Ser Gly Cys Asp Gly Gly Ser Ala Tyr Thr Cys Ala Asn		
65	70	75
Asn Ser Pro Trp Ala Val Asn Asp Gln Leu Ser Tyr Gly Phe Ala Ala		
	85	90
Thr Lys Leu Ser Gly Gly Thr Glu Ser Ser Trp Cys Cys Ala Cys Tyr		
	100	105
Ala Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Thr Met Val Val		
	115	120
Gln Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe Asp Ile		
	130	135
Asn Met Pro Gly Gly Gly Val Gly Leu Phe Asp Gly Cys Thr Arg Gln		
145	150	155
Phe Gly Gly Leu Pro Gly Ala Gln Tyr Gly Gly Ile Ser Ser Arg Ser		

	165		170		175
Gln Cys Asp Ser Phe Pro Ala Ala Leu Lys Pro Gly Cys Gln Trp Arg					
	180		185		190
Phe Asp Trp Phe Gln Asn Ala Asp Asn Pro Asn Phe Thr Phe Lys Gln					
	195		200		205
Val Gln Cys Pro Ser Glu Leu Thr Ser Arg Thr Gly Cys Lys Arg Asn					
	210		215		220
Asp Asp Ser Gln Phe Pro Val Phe Thr Pro Pro Ser Gly Gly Gly Thr					
225		230		235	240
Asn Pro Ser Thr Pro Thr Thr Pro Pro Ser Ser Gly Gly Gly Ser Gly					
	245		250		255
Cys Thr Ala Asp Lys Tyr Ala Gln Cys Gly Gly Ser Gly Trp Ser Gly					
	260		265		270
Cys Thr Asn Cys Pro Ser Gly Ser Thr Cys Lys Thr Ile Asn Asp Tyr					
	275		280		285
Tyr His Gln Cys Ala					
	290				

<210> 15

<211> 951

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<220>

<221> 信号肽

<222> (1) .. (63)

<220>

<221> 外显子

<222> (1) .. (948)

<220>

<221> CDS

<222> (1) .. (948)

<220>

<221> 成熟肽

<222> (64) .. (948)

<400> 15

atg cgc tcc tcc cct gtc etc agg aca gcg etc gca gca gcc etc ccc 48

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

-20

-15

-10

ctc gca gcg ttg gca gca gat gga aag tcg aca cgc tat tgg gac tgt 96
 Leu Ala Ala Leu Ala Ala Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys
 -5 -1 1 5 10
 tgt aag cct tcg tgt tcg tgg cct ggc aag gca tcc gtg aac cag ccg 144
 Cys Lys Pro Ser Cys Ser Trp Pro Gly Lys Ala Ser Val Asn Gln Pro
 15 20 25
 gtc ttc gca tgt tcc gca aac ttc cag cgg atc tcc gat ccg aac gtc 192
 Val Phe Ala Cys Ser Ala Asn Phe Gln Arg Ile Ser Asp Pro Asn Val
 30 35 40
 aaa tcc ggt tgt gat gga ggc tcc gcc tac gcg tgt gcc gac cag aca 240
 Lys Ser Gly Cys Asp Gly Gly Ser Ala Tyr Ala Cys Ala Asp Gln Thr
 45 50 55
 cct tgg gca gtc aac gac aac ttc tcc tac ggc ttc gca gca act tcg 288
 Pro Trp Ala Val Asn Asp Asn Phe Ser Tyr Gly Phe Ala Ala Thr Ser
 60 65 70 75
 atc tcg ggt ggc aac gag gcc tcg tgg tgt tgt gga tgt tac gaa ctc 336
 Ile Ser Gly Gly Asn Glu Ala Ser Trp Cys Cys Gly Cys Tyr Glu Leu
 80 85 90
 acc ttc act tcg gga cct gtg gca ggt aag acc atg gtc gtc cag tcg 384
 Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser
 95 100 105
 acg tcg aca gga ggc gac ctc ggt acc aac cac ttc gat ttg gcg atg 432
 Thr Ser Thr Gly Gly Asp Leu Gly Thr Asn His Phe Asp Leu Ala Met
 110 115 120
 cct gga gga ggc gtg ggc atc ttc gac ggc tgt tcg cct cag ttc gga 480
 Pro Gly Gly Gly Val Gly Ile Phe Asp Gly Cys Ser Pro Gln Phe Gly
 125 130 135
 ggc ttg gca ggc gac cgg tat ggc ggt gtg tcc tcc cgg tcg cag tgt 528
 Gly Leu Ala Gly Asp Arg Tyr Gly Gly Val Ser Ser Arg Ser Gln Cys
 140 145 150 155
 gat tcc ttc cct gca gcc ctc aaa ccc ggt tgt tac tgg cga ttc gac 576
 Asp Ser Phe Pro Ala Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe Asp
 160 165 170
 tgg ttc aag aac gcc gac aac cct acc ttc acg ttc cga cag gtg cag 624
 Trp Phe Lys Asn Ala Asp Asn Pro Thr Phe Thr Phe Arg Gln Val Gln
 175 180 185
 tgt ccg tcg gag ttg gtc gcc agg acc ggc tgt cgc agg aac gat gat 672
 Cys Pro Ser Glu Leu Val Ala Arg Thr Gly Cys Arg Arg Asn Asp Asp
 190 195 200

ggt aac ttc cct gtg ttc acc cct cct tcg gga ggc cag tcc tcc tcg 720
 Gly Asn Phe Pro Val Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Ser
 205 210 215
 tcc tcc tcg tcc tcg tcg gcc aag ccg aca tcg acg tcg acg tcc act 768
 Ser Ser Ser Ser Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Thr
 220 225 230 235
 act tcg acc aaa gca acc tcc aca act tcc acc gcc tcg tcc cag acg 816
 Thr Ser Thr Lys Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser Gln Thr
 240 245 250
 tcg tcg tcc acc gga gga ggc tgt gca gcg cag cgt tgg gca cag tgt 864
 Ser Ser Ser Thr Gly Gly Gly Cys Ala Ala Gln Arg Trp Ala Gln Cys
 255 260 265
 ggt ggt atc gga ttc tcg gga tgt act aca tgt gtg tcg ggc acg acg 912
 Gly Gly Ile Gly Phe Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr
 270 275 280
 tgt aac aag cag aac gac tac tac tcc cag tgt ctc taa 951
 Cys Asn Lys Gln Asn Asp Tyr Tyr Ser Gln Cys Leu
 285 290 295
 <210> 16
 <211> 316
 <212> PRT
 <213> 人工序列
 <220>
 <223> 合成构建体
 <400> 16
 Met Arg Ser Ser Pro Val Leu Arg Thr Ala Leu Ala Ala Ala Leu Pro
 -20 -15 -10
 Leu Ala Ala Leu Ala Ala Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys
 -5 -1 1 5 10
 Cys Lys Pro Ser Cys Ser Trp Pro Gly Lys Ala Ser Val Asn Gln Pro
 15 20 25
 Val Phe Ala Cys Ser Ala Asn Phe Gln Arg Ile Ser Asp Pro Asn Val
 30 35 40
 Lys Ser Gly Cys Asp Gly Gly Ser Ala Tyr Ala Cys Ala Asp Gln Thr
 45 50 55
 Pro Trp Ala Val Asn Asp Asn Phe Ser Tyr Gly Phe Ala Ala Thr Ser
 60 65 70 75
 Ile Ser Gly Gly Asn Glu Ala Ser Trp Cys Cys Gly Cys Tyr Glu Leu
 80 85 90

Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser
 95 100 105
 Thr Ser Thr Gly Gly Asp Leu Gly Thr Asn His Phe Asp Leu Ala Met
 110 115 120
 Pro Gly Gly Gly Val Gly Ile Phe Asp Gly Cys Ser Pro Gln Phe Gly
 125 130 135
 Gly Leu Ala Gly Asp Arg Tyr Gly Gly Val Ser Ser Arg Ser Gln Cys
 140 145 150 155
 Asp Ser Phe Pro Ala Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe Asp
 160 165 170
 Trp Phe Lys Asn Ala Asp Asn Pro Thr Phe Thr Phe Arg Gln Val Gln
 175 180 185
 Cys Pro Ser Glu Leu Val Ala Arg Thr Gly Cys Arg Arg Asn Asp Asp
 190 195 200
 Gly Asn Phe Pro Val Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Ser
 205 210 215
 Ser Ser Ser Ser Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Thr
 220 225 230 235
 Thr Ser Thr Lys Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser Gln Thr
 240 245 250
 Ser Ser Ser Thr Gly Gly Gly Cys Ala Ala Gln Arg Trp Ala Gln Cys
 255 260 265
 Gly Gly Ile Gly Phe Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr
 270 275 280
 Cys Asn Lys Gln Asn Asp Tyr Tyr Ser Gln Cys Leu
 285 290 295

<210> 17

<211> 894

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<220>

<221> 信号肽

<222> (1) .. (63)

<220>

<221> 外显子

<222> (1) .. (891)

<220>

<221> CDS
 <222> (1) .. (891)
 <220>
 <221> 成熟肽
 <222> (64) .. (891)
 <400> 17
 atg cgc tcc tcc cct gtc ctc agg aca gcg ctc gca gca gcc ctc ccc 48
 Met Arg Ser Ser Pro Val Leu Arg Thr Ala Leu Ala Ala Ala Leu Pro
 -20 -15 -10
 ctc gca gcg ttg gca gca gat gga aag tcg aca cgc tat tgg gac tgt 96
 Leu Ala Ala Leu Ala Ala Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys
 -5 -1 1 5 10
 tgt aag cct tcg tgt tcg tgg cct ggc aag gca tcc gtg aac cag ccg 144
 Cys Lys Pro Ser Cys Ser Trp Pro Gly Lys Ala Ser Val Asn Gln Pro
 15 20 25
 gtc ttc gca tgt tcc gca aac ttc cag cgg atc tcc gat ccg aac gtc 192
 Val Phe Ala Cys Ser Ala Asn Phe Gln Arg Ile Ser Asp Pro Asn Val
 30 35 40
 aaa tcc ggt tgt gat gga ggc tcc gcc tac gcg tgt gcc gac cag aca 240
 Lys Ser Gly Cys Asp Gly Gly Ser Ala Tyr Ala Cys Ala Asp Gln Thr
 45 50 55
 cct tgg gca gtc aac gac aac ttc tcc tac ggc ttc gca gca act tcg 288
 Pro Trp Ala Val Asn Asp Asn Phe Ser Tyr Gly Phe Ala Ala Thr Ser
 60 65 70 75
 atc tcg ggt ggc aac gag gcc tcg tgg tgt tgt gga tgt tac gaa ctc 336
 Ile Ser Gly Gly Asn Glu Ala Ser Trp Cys Cys Gly Cys Tyr Glu Leu
 80 85 90
 acc ttc act tcg gga cct gtg gca ggt aag acc atg gtc gtc cag tcg 384
 Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser
 95 100 105
 acg tcg aca gga ggc gac ctc ggt acc aac cac ttc gat ttg gcg atg 432
 Thr Ser Thr Gly Gly Asp Leu Gly Thr Asn His Phe Asp Leu Ala Met
 110 115 120
 cct gga gga ggc gtg ggc atc ttc gac ggc tgt tcg cct cag ttc gga 480
 Pro Gly Gly Gly Val Gly Ile Phe Asp Gly Cys Ser Pro Gln Phe Gly
 125 130 135
 ggc ttg gca ggc gac cgg tat ggc ggt gtg tcc tcc cgg tcg cag tgt 528
 Gly Leu Ala Gly Asp Arg Tyr Gly Gly Val Ser Ser Arg Ser Gln Cys
 140 145 150 155

gat tcc ttc cct gca gcc ctc aaa ccc ggt tgt tac tgg cga ttc gac 576
 Asp Ser Phe Pro Ala Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe Asp
 160 165 170

tgg ttc aag aac gcc gac aac cct acc ttc acg ttc cga cag gtg cag 624
 Trp Phe Lys Asn Ala Asp Asn Pro Thr Phe Thr Phe Arg Gln Val Gln
 175 180 185

tgt ccg tcg gag ttg gtc gcc agg acc ggc tgt cgc gca aac gac gac 672
 Cys Pro Ser Glu Leu Val Ala Arg Thr Gly Cys Arg Ala Asn Asp Asp
 190 195 200

tcc agc ttc ccc gtc ttc acc ccc cca age ggt ggc aac ggt ggc acc 720
 Ser Ser Phe Pro Val Phe Thr Pro Pro Ser Gly Gly Asn Gly Gly Thr
 205 210 215

ggg acg ccc acg tcg act gcg cct ggg tcg ggc cag acg tct ccc ggc 768
 Gly Thr Pro Thr Ser Thr Ala Pro Gly Ser Gly Gln Thr Ser Pro Gly
 220 225 230 235

ggc ggc agt ggc tgc acg tct cag aag tgg gct cag tgc ggt ggc atc 816
 Gly Gly Ser Gly Cys Thr Ser Gln Lys Trp Ala Gln Cys Gly Gly Ile
 240 245 250

ggc ttc agc gga tgc acc acc tgt gtc tct ggc acc acc tgc cag aag 864
 Gly Phe Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys Gln Lys
 255 260 265

ttg aac gac tac tac tcg cag tgc ctc taa 894
 Leu Asn Asp Tyr Tyr Ser Gln Cys Leu
 270 275

<210> 18

<211> 297

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<400> 18

Met Arg Ser Ser Pro Val Leu Arg Thr Ala Leu Ala Ala Ala Leu Pro
 -20 -15 -10

Leu Ala Ala Leu Ala Ala Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys
 -5 -1 1 5 10

Cys Lys Pro Ser Cys Ser Trp Pro Gly Lys Ala Ser Val Asn Gln Pro
 15 20 25

Val Phe Ala Cys Ser Ala Asn Phe Gln Arg Ile Ser Asp Pro Asn Val
 30 35 40

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

<210> 19

<211> 935

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<220>

<221> 信号肽

<222> (1) .. (63)

<220>
 <221> 外显子
 <222> (1) .. (338)
 <220>
 <221> CDS
 <222> (1) .. (338)
 <220>
 <221> 成熟肽
 <222> (64) .. (932)
 <220>
 <221> 内含子
 <222> (339) .. (391)
 <220>
 <221> 外显子
 <222> (392) .. (932)
 <220>
 <221> CDS
 <222> (392) .. (932)
 <400> 19
 atg cgc tcc tcc act gtt ctg caa acc ggg cta gtg gcc get ctt ccc 48
 Met Arg Ser Ser Thr Val Leu Gln Thr Gly Leu Val Ala Ala Leu Pro
 -20 -15 -10
 ttc gcc gtt cag gcc gcg tcc gga tcc ggc cag tcc acc aga tat tgg 96
 Phe Ala Val Gln Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg Tyr Trp
 -5 -1 1 5 10
 gac tgc tgc aaa cca tct tgc tcc tgg tcc ggc aag gct tct gtc aac 144
 Asp Cys Cys Lys Pro Ser Cys Ser Trp Ser Gly Lys Ala Ser Val Asn
 15 20 25
 cga ccc gtc ctc gct tgc gat gea aac aac aac ccc ctg agc gac gcc 192
 Arg Pro Val Leu Ala Cys Asp Ala Asn Asn Asn Pro Leu Ser Asp Ala
 30 35 40
 agt gtc aag tct gga tgt gac ggc ggt tct gea tac acc tgt gcc aac 240
 Ser Val Lys Ser Gly Cys Asp Gly Gly Ser Ala Tyr Thr Cys Ala Asn
 45 50 55
 aac tca cca tgg gcg gtg aac gac cag ctc tcc tac ggc ttt get gcc 288
 Asn Ser Pro Trp Ala Val Asn Asp Gln Leu Ser Tyr Gly Phe Ala Ala
 60 65 70 75
 acg aaa ctc agt ggt gga acc gag tcg tct tgg tgc tgt gcc tgt tat 336
 Thr Lys Leu Ser Gly Gly Thr Glu Ser Ser Trp Cys Cys Ala Cys Tyr

80	85	90	
gc gtagtttcg	cgtaacagc	tggtatcac	ttgatggcta atttatactg cag c 392
Ala			
ctt acc ttc act tcg ggt cct gtt gct ggc aag acc atg gtc gtt cag			440
Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Thr Met Val Val Gln			
95	100	105	
tct acc agt acc ggc ggt gat ctc ggc tcc aac cac ttc gat atc aac			488
Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe Asp Ile Asn			
110	115	120	
atg ccc ggc ggc ggc gtc ggc ctg ttt gat ggc tgt aca cga cag ttt			536
Met Pro Gly Gly Gly Val Gly Leu Phe Asp Gly Cys Thr Arg Gln Phe			
125	130	135	140
ggc ggt ctc ccc ggc gct caa tat ggc ggc atc agc tcc cgc agc cag			584
Gly Gly Leu Pro Gly Ala Gln Tyr Gly Gly Ile Ser Ser Arg Ser Gln			
145	150	155	
tgc gat tca ttc cct gcc gcg ctc aag ccc ggt tgc cag tgg cgc ttc			632
Cys Asp Ser Phe Pro Ala Ala Leu Lys Pro Gly Cys Gln Trp Arg Phe			
160	165	170	
gac tgg ttc cag aac gcc gac aac ccc aac ttc acc ttc aag cag gtc			680
Asp Trp Phe Gln Asn Ala Asp Asn Pro Asn Phe Thr Phe Lys Gln Val			
175	180	185	
caa tgc cca tcc gag ctc acc tcc cgc acc ggc tgc aag cga aac gac			728
Gln Cys Pro Ser Glu Leu Thr Ser Arg Thr Gly Cys Lys Arg Asn Asp			
190	195	200	
gac tct caa ttc cct gtc ttc act ccg ccc tct ggt gga ggc acc aac			776
Asp Ser Gln Phe Pro Val Phe Thr Pro Pro Ser Gly Gly Gly Thr Asn			
205	210	215	220
ccc tct act ccg aca acc cct ccc tct tca ggc ggc ggt tcc gga tgt			824
Pro Ser Thr Pro Thr Thr Pro Pro Ser Ser Gly Gly Gly Ser Gly Cys			
225	230	235	
acg gcg gat aaa tac gct cag tgt ggc ggc tcg ggg tgg tct ggc tgc			872
Thr Ala Asp Lys Tyr Ala Gln Cys Gly Gly Ser Gly Trp Ser Gly Cys			
240	245	250	
acc aac tgc ccg tct gga tcg acc tgc aag act atc aac gat tgg tac			920
Thr Asn Cys Pro Ser Gly Ser Thr Cys Lys Thr Ile Asn Asp Trp Tyr			
255	260	265	
cac cag tgt gcc taa			935
His Gln Cys Ala			
270			

<210> 20
 <211> 293
 <212> PRT
 <213> 人工序列
 <220>
 <223> 合成构建体
 <400> 20
 Met Arg Ser Ser Thr Val Leu Gln Thr Gly Leu Val Ala Ala Leu Pro
 -20 -15 -10
 Phe Ala Val Gln Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg Tyr Trp
 -5 -1 1 5 10
 Asp Cys Cys Lys Pro Ser Cys Ser Trp Ser Gly Lys Ala Ser Val Asn
 15 20 25
 Arg Pro Val Leu Ala Cys Asp Ala Asn Asn Asn Pro Leu Ser Asp Ala
 30 35 40
 Ser Val Lys Ser Gly Cys Asp Gly Gly Ser Ala Tyr Thr Cys Ala Asn
 45 50 55
 Asn Ser Pro Trp Ala Val Asn Asp Gln Leu Ser Tyr Gly Phe Ala Ala
 60 65 70 75
 Thr Lys Leu Ser Gly Gly Thr Glu Ser Ser Trp Cys Cys Ala Cys Tyr
 80 85 90
 Ala Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Thr Met Val Val
 95 100 105
 Gln Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe Asp Ile
 110 115 120
 Asn Met Pro Gly Gly Gly Val Gly Leu Phe Asp Gly Cys Thr Arg Gln
 125 130 135
 Phe Gly Gly Leu Pro Gly Ala Gln Tyr Gly Gly Ile Ser Ser Arg Ser
 140 145 150 155
 Gln Cys Asp Ser Phe Pro Ala Ala Leu Lys Pro Gly Cys Gln Trp Arg
 160 165 170
 Phe Asp Trp Phe Gln Asn Ala Asp Asn Pro Asn Phe Thr Phe Lys Gln
 175 180 185
 Val Gln Cys Pro Ser Glu Leu Thr Ser Arg Thr Gly Cys Lys Arg Asn
 190 195 200
 Asp Asp Ser Gln Phe Pro Val Phe Thr Pro Pro Ser Gly Gly Gly Thr
 205 210 215
 Asn Pro Ser Thr Pro Thr Thr Pro Pro Ser Ser Gly Gly Gly Ser Gly
 220 225 230 235

<220>
 <221> CDS
 <222> (486)..(1101)
 <400> 21
 atg cgc tct act ccc gtt ctt cgc aca acc ctg gcc gct gca ctt cct 48
 Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro
 -20 -15 -10
 ctg gtc gcc tcc gcg gcc agt ggc agt ggc cag tcc acg ag 89
 Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg
 -5 -1 1 5
 gtatgcgtcc ctcaatgcca atgctcacc aaacgagcta ggtgtccagg acgccgccca 149
 tgctgactcg tttcccag a tac tgg gac tgc tgc aag ceg teg tgc gct tgg 201
 Tyr Trp Asp Cys Cys Lys Pro Ser Cys Ala Trp
 10 15 20
 ccc ggg aag gcc gcc gtc age caa ccg gtc tac gcg tgc gat gcc aac 249
 Pro Gly Lys Ala Ala Val Ser Gln Pro Val Tyr Ala Cys Asp Ala Asn
 25 30 35
 ttc cag cgc ctg tcc gac ttc aat gtc cag tgc ggc tgc aac ggc ggc 297
 Phe Gln Arg Leu Ser Asp Phe Asn Val Gln Ser Gly Cys Asn Gly Gly
 40 45 50
 teg gcc tac tcc tgc gcc gac cag act ccc tgg gcg gtg aac gac aat 345
 Ser Ala Tyr Ser Cys Ala Asp Gln Thr Pro Trp Ala Val Asn Asp Asn
 55 60 65
 ctc gcc tac ggc ttc gcc gcg acg agc atc gcc ggc ggc tcc gaa tcc 393
 Leu Ala Tyr Gly Phe Ala Ala Thr Ser Ile Ala Gly Gly Ser Glu Ser
 70 75 80
 tcg tgg tgc tgc gcc tgc tac gc gtaagtctc tctgccagct accaggaagt 446
 Ser Trp Cys Cys Ala Cys Tyr Ala
 85 90
 tggattagcg cgagctaacc tcaactccatc aactccag g etc acc ttc act tcc 501
 Leu Thr Phe Thr Ser
 95
 ggt ccc gtc gcc ggc aag aca atg gtg gtg cag tca acg agc act ggc 549
 Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr Ser Thr Gly
 100 105 110
 ggc gac ctg gga agt aac cat ttc gat atc gcc atg ccc ggc ggc ggc 597
 Gly Asp Leu Gly Ser Asn His Phe Asp Ile Ala Met Pro Gly Gly Gly
 115 120 125
 gtg ggc atc ttc aac ggc tgc age tcg cag ttc ggc ggc etc ccc ggc 645

Val Gly Ile Phe Asn Gly Cys Ser Ser Gln Phe Gly Gly Leu Pro Gly	
130	135 140 145
gct caa tac ggc ggc att tcg tcg cgc gac cag tgc gat tcc ttc ccc	693
Ala Gln Tyr Gly Gly Ile Ser Ser Arg Asp Gln Cys Asp Ser Phe Pro	
150	155 160
gcg ccg ctc aag ccc ggc tgc cag tgg cgg ttt gac tgg ttc cag aac	741
Ala Pro Leu Lys Pro Gly Cys Gln Trp Arg Phe Asp Trp Phe Gln Asn	
165	170 175
gcc gac aac ccg acg ttc acg ttc cag cag gtg cag tgc ccc gcc gag	789
Ala Asp Asn Pro Thr Phe Thr Phe Gln Gln Val Gln Cys Pro Ala Glu	
180	185 190
atc gtt gcc cgc tcc ggc tgc aag cgc aac gac gat ggt aac ttc cct	837
Ile Val Ala Arg Ser Gly Cys Lys Arg Asn Asp Asp Gly Asn Phe Pro	
195	200 205
gtg ttc acc cct cct tcg gga ggc cag tcc tcc tcg tcc tcc tcg tcc	885
Val Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Ser Ser Ser Ser Ser	
210	215 220 225
tcg tcg gcc aag ccg aca tcg acg tcg acg tcc act act tcg acc aaa	933
Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Thr Thr Ser Thr Lys	
230	235 240
gca acc tcc aca act tcc acc gcc tcg tcc cag acg tcg tcg tcc acc	981
Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser Gln Thr Ser Ser Ser Thr	
245	250 255
gga gga ggc tgc gca gcg cag aag tac gct cag tgc ggt ggc atc ggc	1029
Gly Gly Gly Cys Ala Ala Gln Lys Tyr Ala Gln Cys Gly Gly Ile Gly	
260	265 270
ttc agc gga tgc acc acc tgt gtc tct ggc acc acc tgc aac aag cag	1077
Phe Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys Asn Lys Gln	
275	280 285
aac gac tac tac tcg cag tgc ctc tga	1104
Asn Asp Tyr Tyr Ser Gln Cys Leu	
290	295
<210> 22	
<211> 318	
<212> PRT	
<213> 人工序列	
<220>	
<223> 合成构建体	
<400> 22	

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

285	290	295
<210> 23		
<211> 1104		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成构建体		
<220>		
<221> 信号肽		
<222> (1) .. (63)		
<220>		
<221> 外显子		
<222> (1) .. (89)		
<220>		
<221> CDS		
<222> (1) .. (89)		
<220>		
<221> 成熟肽		
<222> (64) .. (1101)		
<220>		
<221> 内含子		
<222> (90) .. (167)		
<220>		
<221> 外显子		
<222> (168) .. (416)		
<220>		
<221> CDS		
<222> (168) .. (416)		
<220>		
<221> 内含子		
<222> (417) .. (485)		
<220>		
<221> 外显子		
<222> (486) .. (1101)		
<220>		
<221> CDS		
<222> (486) .. (1101)		
<400> 23		
atg cgc tct act ccc gtt ctt cgc aca acc ctg gcc gct gca ctt cct 48		

```

Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro
-20 -15 -10
ctg gtc gcc tcc gcg gcc agt ggc agt ggc cag tcc acg ag 89
Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg
-5 -1 1 5
gtatgcgtcc ctcaatgcga atgcctcacc aaacgagcta ggtgtccagg acgccgccca 149
tgctgactcg tttcccag a tac tgg gac tgc tgc aag ccg tgc tgc gct tgg 201
Tyr Trp Asp Cys Cys Lys Pro Ser Cys Ala Trp
10 15 20
ccc ggg aag gcc gcc gtc age caa ccg gtc tac gcg tgc gat gcc aac 249
Pro Gly Lys Ala Ala Val Ser Gln Pro Val Tyr Ala Cys Asp Ala Asn
25 30 35
ttc cag cgc ctg tcc gac ttc aat gtc cag tgc ggc tgc aac ggc ggc 297
Phe Gln Arg Leu Ser Asp Phe Asn Val Gln Ser Gly Cys Asn Gly Gly
40 45 50
tcg gcc tac tcc tgc gcc gac cag act ccc tgg gcg gtg aac gac aat 345
Ser Ala Tyr Ser Cys Ala Asp Gln Thr Pro Trp Ala Val Asn Asp Asn
55 60 65
ctc gcc tac ggc ttc gcc gcg acg age atc gcc ggc ggc tcc gaa tcc 393
Leu Ala Tyr Gly Phe Ala Ala Thr Ser Ile Ala Gly Gly Ser Glu Ser
70 75 80
tcg tgg tgc tgc gcc tgc tac gc gtaagtctc tctgccagct accaggaagt 446
Ser Trp Cys Cys Ala Cys Tyr Ala
85 90
tggattagcg cgagctaacc tcaactccatc aactccag g ctc acc ttc act tcc 501
Leu Thr Phe Thr Ser
95
ggc ccc gtc gcc ggc aag aca atg gtg gtg cag tca acg age act ggc 549
Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr Ser Thr Gly
100 105 110
ggc gac ctg gga agt aac cat ttc gat atc gcc atg ccc ggc ggc ggc 597
Gly Asp Leu Gly Ser Asn His Phe Asp Ile Ala Met Pro Gly Gly Gly
115 120 125
gtg ggc atc ttc aac ggc tgc age tgc cag ttc ggc ggc ctc ccc ggc 645
Val Gly Ile Phe Asn Gly Cys Ser Ser Gln Phe Gly Gly Leu Pro Gly
130 135 140 145
gct caa tac ggc ggc att tgc tgc cgc gac cag tgc gat tcc ttc ccc 693
Ala Gln Tyr Gly Gly Ile Ser Ser Arg Asp Gln Cys Asp Ser Phe Pro
150 155 160

```

gcg ccg ctc aag ccc ggc tgc cag tgg cgg ttt gac tgg ttc cag aac	741
Ala Pro Leu Lys Pro Gly Cys Gln Trp Arg Phe Asp Trp Phe Gln Asn	
165 170 175	
gcc gac aac ccg acg ttc acg ttc cag cag gtg cag tgc ccc gcc gag	789
Ala Asp Asn Pro Thr Phe Thr Phe Gln Gln Val Gln Cys Pro Ala Glu	
180 185 190	
atc gtt gcc cgc tcc ggc tgc aag cgc aac gac gat ggt aac ttc cct	837
Ile Val Ala Arg Ser Gly Cys Lys Arg Asn Asp Asp Gly Asn Phe Pro	
195 200 205	
gtg ttc acc cct cct tcg gga ggc cag tcc tcc tcg tcc tcc tcg tcc	885
Val Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Ser Ser Ser Ser	
210 215 220 225	
tcg tcg gcc aag ccg aca tcg acg tcg acg tcc act act tcg acc aaa	933
Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Thr Thr Ser Thr Lys	
230 235 240	
gca acc tcc aca act tcc acc gcc tcg tcc cag acg tcg tcg tcc acc	981
Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser Gln Thr Ser Ser Ser Thr	
245 250 255	
gga gga ggc tgc gca gcg cag cgt tac gct cag tgc ggt ggc atc ggc	1029
Gly Gly Gly Cys Ala Ala Gln Arg Tyr Ala Gln Cys Gly Gly Ile Gly	
260 265 270	
tac agc gga tgc acc acc tgt gtc tct ggc acc acc tgc aac aag cag	1077
Tyr Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys Asn Lys Gln	
275 280 285	
aac gac tac tac tcg cag tgc ctc tga	1104
Asn Asp Tyr Tyr Ser Gln Cys Leu	
290 295	
<210> 24	
<211> 318	
<212> PRT	
<213> 人工序列	
<220>	
<223> 合成构建体	
<400> 24	
Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro	
-20 -15 -10	
Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg Tyr Trp	
-5 -1 1 5 10	
Asp Cys Cys Lys Pro Ser Cys Ala Trp Pro Gly Lys Ala Ala Val Ser	

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

<210> 25

<211> 1104

<212> DNA

<213> 人工序列

<220>
 <223> 合成构建体
 <220>
 <221> 信号肽
 <222> (1) .. (63)
 <220>
 <221> 外显子
 <222> (1) .. (89)
 <220>
 <221> CDS
 <222> (1) .. (89)
 <220>
 <221> 成熟肽
 <222> (64) .. (1101)
 <220>
 <221> 内含子
 <222> (90) .. (167)
 <220>
 <221> 外显子
 <222> (168) .. (416)
 <220>
 <221> CDS
 <222> (168) .. (416)
 <220>
 <221> 内含子
 <222> (417) .. (485)
 <220>
 <221> 外显子
 <222> (486) .. (1101)
 <220>
 <221> CDS
 <222> (486) .. (1101)
 <400> 25

```

atg cgc tct act ccc gtt ctt cgc aca acc ctg gcc gct gea ctt cct 48
Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro
      -20                -15                -10
ctg gtc gcc tcc gcg gcc agt ggc agt ggc cag tcc acg ag 89
Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg
-5                -1 1                5
  
```


gtatgcgtcc ctcaatgcga atgcctcacc aaacgageta ggtgtccagg acgccgccca 149
 tgctgactcg tttcccag a tac tgg gac tgc tgc aag ccg tcg tgc gct tgg 201
 Tyr Trp Asp Cys Cys Lys Pro Ser Cys Ala Trp
 10 15 20
 ccc ggg aag gcc gcc gtc agc caa ccg gtc tac gcg tgc gat gcc aac 249
 Pro Gly Lys Ala Ala Val Ser Gln Pro Val Tyr Ala Cys Asp Ala Asn
 25 30 35
 ttc cag cgc ctg tcc gac ttc aat gtc cag tcg ggc tgc aac ggc ggc 297
 Phe Gln Arg Leu Ser Asp Phe Asn Val Gln Ser Gly Cys Asn Gly Gly
 40 45 50
 tcg gcc tac tcc tgc gcc gac cag act ccc tgg gcg gtg aac gac aat 345
 Ser Ala Tyr Ser Cys Ala Asp Gln Thr Pro Trp Ala Val Asn Asp Asn
 55 60 65
 ctc gcc tac ggc ttc gcc gcg acg agc atc gcc ggc ggg tcc gaa tcc 393
 Leu Ala Tyr Gly Phe Ala Ala Thr Ser Ile Ala Gly Gly Ser Glu Ser
 70 75 80
 tcg tgg tgc tgc gcc tgc tac gc gtaagtctc tctgccagct accaggaagt 446
 Ser Trp Cys Cys Ala Cys Tyr Ala
 85 90
 tggattagcg cgagctaacc tcaactccatc aactccag g ctc acc ttc act tcc 501
 Leu Thr Phe Thr Ser
 95
 ggt ccc gtc gcc ggc aag aca atg gtg gtg cag tca acg agc act ggc 549
 Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr Ser Thr Gly
 100 105 110
 ggc gac ctg gga agt aac cat ttc gat atc gcc atg ccc ggc ggc ggc 597
 Gly Asp Leu Gly Ser Asn His Phe Asp Ile Ala Met Pro Gly Gly Gly
 115 120 125
 gtg ggc atc ttc aac ggc tgc agc tcg cag ttc ggc ggc ctc ccc ggc 645
 Val Gly Ile Phe Asn Gly Cys Ser Ser Gln Phe Gly Gly Leu Pro Gly
 130 135 140 145
 gct caa tac ggc ggc att tcg tcg cgc gac cag tgc gat tcc ttc ccc 693
 Ala Gln Tyr Gly Gly Ile Ser Ser Arg Asp Gln Cys Asp Ser Phe Pro
 150 155 160
 gcg ccg ctc aag ccc ggc tgc cag tgg cgg ttt gac tgg ttc cag aac 741
 Ala Pro Leu Lys Pro Gly Cys Gln Trp Arg Phe Asp Trp Phe Gln Asn
 165 170 175
 gcc gac aac ccg acg ttc acg ttc cag cag gtg cag tgc ccc gcc gag 789
 Ala Asp Asn Pro Thr Phe Thr Phe Gln Gln Val Gln Cys Pro Ala Glu

180	185	190	
atc gtt gcc cgc tcc ggc tgc aag cgc aac gac gat ggt aac ttc cct			837
Ile Val Ala Arg Ser Gly Cys Lys Arg Asn Asp Asp Gly Asn Phe Pro			
195	200	205	
gtg ttc acc cct cct tcg gga ggc cag tcc tcc tcg tcc tcc tcg tcc			885
Val Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Ser Ser Ser Ser Ser			
210	215	220	225
tcg tcg gcc aag ccg aca tcg acg tcg acg tcc aac act tcg acc aaa			933
Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Asn Thr Ser Thr Lys			
	230	235	240
gca acc tcc aca aac tcc acc gcc tcg tcc cag acg tcg tcg tcc acc			981
Ala Thr Ser Thr Asn Ser Thr Ala Ser Ser Gln Thr Ser Ser Ser Thr			
	245	250	255
gga gga ggc tgc gca gcg cag cgt tgg get cag tgc ggt ggc atc ggc			1029
Gly Gly Gly Cys Ala Ala Gln Arg Trp Ala Gln Cys Gly Gly Ile Gly			
	260	265	270
ttc agc gga tgc acc acc tgt gtc tct ggc acc acc tgc aac aag cag			1077
Phe Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys Asn Lys Gln			
	275	280	285
aac gac tac tac tcg cag tgc ctc tga			1104
Asn Asp Tyr Tyr Ser Gln Cys Leu			
290	295		
<210> 26			
<211> 318			
<212> PRT			
<213> 人工序列			
<220>			
<223> 合成构建体			
<400> 26			
Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro			
-20	-15	-10	
Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg Tyr Trp			
-5	-1 1	5	10
Asp Cys Cys Lys Pro Ser Cys Ala Trp Pro Gly Lys Ala Ala Val Ser			
	15	20	25
Gln Pro Val Tyr Ala Cys Asp Ala Asn Phe Gln Arg Leu Ser Asp Phe			
	30	35	40
Asn Val Gln Ser Gly Cys Asn Gly Gly Ser Ala Tyr Ser Cys Ala Asp			
	45	50	55

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

<210> 27

<211> 1104

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<220>

<221> 信号肽

<222> (1) .. (63)

<220>
 <221> 外显子
 <222> (1) .. (89)
 <220>
 <221> CDS
 <222> (1) .. (89)
 <220>
 <221> 成熟肽
 <222> (64) .. (1101)
 <220>
 <221> 内含子
 <222> (90) .. (167)
 <220>
 <221> 外显子
 <222> (168) .. (416)
 <220>
 <221> CDS
 <222> (168) .. (416)
 <220>
 <221> 内含子
 <222> (417) .. (485)
 <220>
 <221> 外显子
 <222> (486) .. (1101)
 <220>
 <221> CDS
 <222> (486) .. (1101)
 <400> 27
 atg cgc tct act ccc gtt ctt cgc aca acc ctg gcc gct gca ctt cct 48
 Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro
 -20 -15 -10
 ctg gtc gcc tcc gcg gcc agt ggc agt ggc cag tcc acg ag 89
 Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg
 -5 -1 1 5
 gtatgcgtcc ctcaatgcga atgcctcacc aaacgagcta ggtgtccagg acgccgccca 149
 tgetgactcg tttcccag a tac tgg gac tgc tgc aag ccg tgc tgc gct tgg 201
 Tyr Trp Asp Cys Cys Lys Pro Ser Cys Ala Trp
 10 15 20
 ccc ggg aag gcc gcc gtc agc caa ccg gtc tac gcg tgc gat gcc aac 249

Pro Gly Lys Ala Ala Val Ser Gln Pro Val Tyr Ala Cys Asp Ala Asn	
25	30
ttc cag cgc ctg tcc gac ttc aat gtc cag tcg ggc tgc aac ggc ggc	297
Phe Gln Arg Leu Ser Asp Phe Asn Val Gln Ser Gly Cys Asn Gly Gly	
40	45
tcg gcc tac tcc tgc gcc gac cag act ccc tgg gcg gtg aac gac aat	345
Ser Ala Tyr Ser Cys Ala Asp Gln Thr Pro Trp Ala Val Asn Asp Asn	
55	60
ctc gcc tac ggc ttc gcc gcg acg agc atc gcc ggc ggg tcc gaa tcc	393
Leu Ala Tyr Gly Phe Ala Ala Thr Ser Ile Ala Gly Gly Ser Glu Ser	
70	75
tcg tgg tgc tgc gcc tgc tac gc gtaagtctc tctgccagct accaggaagt	446
Ser Trp Cys Cys Ala Cys Tyr Ala	
85	90
tggattagcg cgagctaac tcactccatc aactccag g ctc acc ttc act tcc	501
	Leu Thr Phe Thr Ser
	95
ggt ccc gtc gcc ggc aag aca atg gtg gtg cag tca acg agc act ggc	549
Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr Ser Thr Gly	
100	105
ggc gac ctg gga agt aac cat ttc gat atc gcc atg ccc ggc ggc ggc	597
Gly Asp Leu Gly Ser Asn His Phe Asp Ile Ala Met Pro Gly Gly Gly	
115	120
gtg ggc atc ttc aac ggc tgc agc tcg cag ttc ggc ggc ctc ccc ggc	645
Val Gly Ile Phe Asn Gly Cys Ser Ser Gln Phe Gly Gly Leu Pro Gly	
130	135
gct caa tac ggc ggc att tcg tcg cgc gac cag tgc gat tcc ttc ccc	693
Ala Gln Tyr Gly Gly Ile Ser Ser Arg Asp Gln Cys Asp Ser Phe Pro	
	150
	155
	160
gcg ccg ctc aag ccc ggc tgc cag tgg cgg ttt gac tgg ttc cag aac	741
Ala Pro Leu Lys Pro Gly Cys Gln Trp Arg Phe Asp Trp Phe Gln Asn	
165	170
	175
gcc gac aac ccg acg ttc acg ttc cag cag gtg cag tgc ccc gcc gag	789
Ala Asp Asn Pro Thr Phe Thr Phe Gln Gln Val Gln Cys Pro Ala Glu	
180	185
	190
atc gtt gcc cgc tcc ggc tgc aag cgc aac gac gat ggt aac ttc cct	837
Ile Val Ala Arg Ser Gly Cys Lys Arg Asn Asp Asp Gly Asn Phe Pro	
195	200
	205
gtg ttc acc cct cct tcg gga ggc cag tcc tcc tcg tcc tcc tcg tcc	885

Val Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Ser Ser Ser Ser Ser
 210 215 220 225
 tcg tcg gcc aag ccg aca tcg acg tcg acg tcc act act tcg acc aaa 933
 Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Thr Thr Ser Thr Lys
 230 235 240
 gca acc tcc aca act tcc acc gcc tcg tcc cag acg tcg tcg tcc acc 981
 Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser Gln Thr Ser Ser Ser Thr
 245 250 255
 gga gga ggc tgc gca gcg cag cgt tgg gct cag tgc ggt gcc atc gcc 1029
 Gly Gly Gly Cys Ala Ala Gln Arg Trp Ala Gln Cys Gly Gly Ile Gly
 260 265 270
 tac agc gga tgc acc acc tgt gtc tct gcc acc acc tgc aac aag cag 1077
 Tyr Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys Asn Lys Gln
 275 280 285
 aac gac tac tac tcg cag tgc etc tga 1104
 Asn Asp Tyr Tyr Ser Gln Cys Leu
 290 295
 <210> 28
 <211> 318
 <212> PRT
 <213> 人工序列
 <220>
 <223> 合成构建体
 <400> 28
 Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro
 -20 -15 -10
 Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg Tyr Trp
 -5 -1 1 5 10
 Asp Cys Cys Lys Pro Ser Cys Ala Trp Pro Gly Lys Ala Ala Val Ser
 15 20 25
 Gln Pro Val Tyr Ala Cys Asp Ala Asn Phe Gln Arg Leu Ser Asp Phe
 30 35 40
 Asn Val Gln Ser Gly Cys Asn Gly Gly Ser Ala Tyr Ser Cys Ala Asp
 45 50 55
 Gln Thr Pro Trp Ala Val Asn Asp Asn Leu Ala Tyr Gly Phe Ala Ala
 60 65 70 75
 Thr Ser Ile Ala Gly Gly Ser Glu Ser Ser Trp Cys Cys Ala Cys Tyr
 80 85 90
 Ala Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Thr Met Val Val

95	100	105
Gln Ser Thr Ser Thr Gly Gly Asp	Leu Gly Ser Asn His Phe Asp Ile	
110	115	120
Ala Met Pro Gly Gly Gly Val Gly Ile Phe Asn Gly Cys Ser Ser Gln		
125	130	135
Phe Gly Gly Leu Pro Gly Ala Gln Tyr Gly Gly Ile Ser Ser Arg Asp		
140	145	150
Gln Cys Asp Ser Phe Pro Ala Pro Leu Lys Pro Gly Cys Gln Trp Arg		
160	165	170
Phe Asp Trp Phe Gln Asn Ala Asp Asn Pro Thr Phe Thr Phe Gln Gln		
175	180	185
Val Gln Cys Pro Ala Glu Ile Val Ala Arg Ser Gly Cys Lys Arg Asn		
190	195	200
Asp Asp Gly Asn Phe Pro Val Phe Thr Pro Pro Ser Gly Gly Gln Ser		
205	210	215
Ser Ser Ser Ser Ser Ser Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr		
220	225	230
Ser Thr Thr Ser Thr Lys Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser		
240	245	250
Gln Thr Ser Ser Ser Thr Gly Gly Gly Cys Ala Ala Gln Arg Trp Ala		
255	260	265
Gln Cys Gly Gly Ile Gly Tyr Ser Gly Cys Thr Thr Cys Val Ser Gly		
270	275	280
Thr Thr Cys Asn Lys Gln Asn Asp Tyr Tyr Ser Gln Cys Leu		
285	290	295

<210> 29

<211> 1104

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<220>

<221> 信号肽

<222> (1) .. (63)

<220>

<221> 外显子

<222> (1) .. (89)

<220>

<221> CDS

<222> (1) .. (89)
 <220>
 <221> 成熟肽
 <222> (64) .. (1101)
 <220>
 <221> 内含子
 <222> (90) .. (167)
 <220>
 <221> 外显子
 <222> (168) .. (416)
 <220>
 <221> CDS
 <222> (168) .. (416)
 <220>
 <221> 内含子
 <222> (417) .. (485)
 <220>
 <221> 外显子
 <222> (486) .. (1101)
 <220>
 <221> CDS
 <222> (486) .. (1101)
 <400> 29
 atg cgc tct act ccc gtt ctt cgc aca acc ctg gcc gct gca ctt cct 48
 Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro
 -20 -15 -10
 ctg gtc gcc tcc gcg gcc agt ggc agt ggc cag tcc acg ag 89
 Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg
 -5 -1 1 5
 gtatgcgtcc ctcaatgcga atgectcacc aaacgagcta ggtgtccagg acgccgccca 149
 tgctgactcg tttcccag a tac tgg gac tgc tgc aag ceg teg tgc gct tgg 201
 Tyr Trp Asp Cys Cys Lys Pro Ser Cys Ala Trp
 10 15 20
 ccc ggg aag gcc gcc gtc age caa ceg gtc tac gcg tgc gat gcc aac 249
 Pro Gly Lys Ala Ala Val Ser Gln Pro Val Tyr Ala Cys Asp Ala Asn
 25 30 35
 ttc cag cgc ctg tcc gac ttc aat gtc cag teg ggc tgc aac ggc ggc 297
 Phe Gln Arg Leu Ser Asp Phe Asn Val Gln Ser Gly Cys Asn Gly Gly
 40 45 50

tcg gcc tac tcc tgc gcc gac cag act ccc tgg gcg gtg aac gac aat	345
Ser Ala Tyr Ser Cys Ala Asp Gln Thr Pro Trp Ala Val Asn Asp Asn	
55 60 65	
ctc gcc tac ggc ttc gcc gcg acg agc atc gcc ggc ggg tcc gaa tcc	393
Leu Ala Tyr Gly Phe Ala Ala Thr Ser Ile Ala Gly Gly Ser Glu Ser	
70 75 80	
tcg tgg tgc tgc gcc tgc tac gc gtaagtctc tctgccagct accaggaagt	446
Ser Trp Cys Cys Ala Cys Tyr Ala	
85 90	
tggattagcg cgagctaacc tcaactccatc acaactccag g ctc acc ttc act tcc	501
Leu Thr Phe Thr Ser	
95	
ggt ccc gtc gcc ggc aag aca atg gtg gtg cag tca acg agc act ggc	549
Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr Ser Thr Gly	
100 105 110	
ggc gac ctg gga agt aac cat ttc gat atc gcc atg ccc ggc ggc ggc	597
Gly Asp Leu Gly Ser Asn His Phe Asp Ile Ala Met Pro Gly Gly Gly	
115 120 125	
gtg ggc atc ttc aac ggc tgc agc tcg cag ttc ggc ggc ctc ccc ggc	645
Val Gly Ile Phe Asn Gly Cys Ser Ser Gln Phe Gly Gly Leu Pro Gly	
130 135 140 145	
gct caa tac ggc ggc att tcg tcg cgc gac cag tgc gat tcc ttc ccc	693
Ala Gln Tyr Gly Gly Ile Ser Ser Arg Asp Gln Cys Asp Ser Phe Pro	
150 155 160	
gcg ccg ctc aag ccc ggc tgc cag tgg cgg ttt gac tgg ttc cag aac	741
Ala Pro Leu Lys Pro Gly Cys Gln Trp Arg Phe Asp Trp Phe Gln Asn	
165 170 175	
gcc gac aac ccg acg ttc acg ttc cag cag gtg cag tgc ccc gcc gag	789
Ala Asp Asn Pro Thr Phe Thr Phe Gln Gln Val Gln Cys Pro Ala Glu	
180 185 190	
atc gtt gcc cgc tcc ggc tgc aag cgc aac gac gat ggt aac ttc cct	837
Ile Val Ala Arg Ser Gly Cys Lys Arg Asn Asp Asp Gly Asn Phe Pro	
195 200 205	
gtg ttc acc cct cct tcg gga ggc cag tcc tcc tcg tcc tcc tcg tcc	885
Val Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Ser Ser Ser Ser Ser	
210 215 220 225	
tcg tcg gcc aag ccg aca tcg acg tcg acg tcc act act tcg acc aaa	933
Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Thr Thr Ser Thr Lys	
230 235 240	

gca acc tcc aca act tcc acc gcc tcg tcc cag acg tcg ccg tcc acc 981
 Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser Gln Thr Ser Pro Ser Thr
 245 250 255
 gga gga ggc tgc gca gcg cag cgt tgg gct cag tgc ggt ggc atc ggc 1029
 Gly Gly Gly Cys Ala Ala Gln Arg Trp Ala Gln Cys Gly Gly Ile Gly
 260 265 270
 ttc agc gga tgc acc acc tgt gtc tct ggc acc acc tgc aac aag cag 1077
 Phe Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys Asn Lys Gln
 275 280 285
 aac gac tac tac tcg cag tgc ctc tga 1104
 Asn Asp Tyr Tyr Ser Gln Cys Leu
 290 295
 <210> 30
 <211> 318
 <212> PRT
 <213> 人工序列
 <220>
 <223> 合成构建体
 <400> 30
 Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro
 -20 -15 -10
 Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg Tyr Trp
 -5 -1 1 5 10
 Asp Cys Cys Lys Pro Ser Cys Ala Trp Pro Gly Lys Ala Ala Val Ser
 15 20 25
 Gln Pro Val Tyr Ala Cys Asp Ala Asn Phe Gln Arg Leu Ser Asp Phe
 30 35 40
 Asn Val Gln Ser Gly Cys Asn Gly Gly Ser Ala Tyr Ser Cys Ala Asp
 45 50 55
 Gln Thr Pro Trp Ala Val Asn Asp Asn Leu Ala Tyr Gly Phe Ala Ala
 60 65 70 75
 Thr Ser Ile Ala Gly Gly Ser Glu Ser Ser Trp Cys Cys Ala Cys Tyr
 80 85 90
 Ala Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Thr Met Val Val
 95 100 105
 Gln Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe Asp Ile
 110 115 120
 Ala Met Pro Gly Gly Gly Val Gly Ile Phe Asn Gly Cys Ser Ser Gln
 125 130 135

Phe Gly Gly Leu Pro Gly Ala Gln Tyr Gly Gly Ile Ser Ser Arg Asp
 140 145 150 155
 Gln Cys Asp Ser Phe Pro Ala Pro Leu Lys Pro Gly Cys Gln Trp Arg
 160 165 170
 Phe Asp Trp Phe Gln Asn Ala Asp Asn Pro Thr Phe Thr Phe Gln Gln
 175 180 185
 Val Gln Cys Pro Ala Glu Ile Val Ala Arg Ser Gly Cys Lys Arg Asn
 190 195 200
 Asp Asp Gly Asn Phe Pro Val Phe Thr Pro Pro Ser Gly Gly Gln Ser
 205 210 215
 Ser Ser Ser Ser Ser Ser Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr
 220 225 230 235
 Ser Thr Thr Ser Thr Lys Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser
 240 245 250
 Gln Thr Ser Pro Ser Thr Gly Gly Gly Cys Ala Ala Gln Arg Trp Ala
 255 260 265
 Gln Cys Gly Gly Ile Gly Phe Ser Gly Cys Thr Thr Cys Val Ser Gly
 270 275 280
 Thr Thr Cys Asn Lys Gln Asn Asp Tyr Tyr Ser Gln Cys Leu
 285 290 295

<210> 31

<211> 1104

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<220>

<221> 信号肽

<222> (1) .. (63)

<220>

<221> 外显子

<222> (1) .. (89)

<220>

<221> CDS

<222> (1) .. (89)

<220>

<221> 成熟肽

<222> (64) .. (1101)

<220>

<221> 内含子
 <222> (90)..(167)
 <220>
 <221> 外显子
 <222> (168)..(416)
 <220>
 <221> CDS
 <222> (168)..(416)
 <220>
 <221> 内含子
 <222> (417)..(485)
 <220>
 <221> 外显子
 <222> (486)..(1101)
 <220>
 <221> CDS
 <222> (486)..(1101)
 <400> 31
 atg cgc tct act ccc gtt ctt cgc aca acc ctg gcc get gca ctt cct 48
 Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro
 -20 -15 -10
 ctg gtc gcc tcc gcg gcc agt ggc agt ggc cag tcc acg ag 89
 Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg
 -5 -1 1 5
 gtatgctgccc ctcaatgcga atgcctcacc aaacgagcta ggtgtccagg acgccgccca 149
 tgctgactcg tttcccag a tac tgg gac tgc tgc aag ccg tgc tgc gct tgg 201
 Tyr Trp Asp Cys Cys Lys Pro Ser Cys Ala Trp
 10 15 20
 ccc ggg aag gcc gcc gtc agc caa ccg gtc tac gcg tgc gat gcc aac 249
 Pro Gly Lys Ala Ala Val Ser Gln Pro Val Tyr Ala Cys Asp Ala Asn
 25 30 35
 ttc cag cgc ctg tcc gac ttc aat gtc cag tgc ggc tgc aac ggc ggc 297
 Phe Gln Arg Leu Ser Asp Phe Asn Val Gln Ser Gly Cys Asn Gly Gly
 40 45 50
 tcg gcc tac tcc tgc gcc gac cag act ccc tgg gcg gtg aac gac aat 345
 Ser Ala Tyr Ser Cys Ala Asp Gln Thr Pro Trp Ala Val Asn Asp Asn
 55 60 65
 ctc gcc tac ggc ttc gcc gcg acg agc atc gcc ggc ggc tcc gaa tcc 393
 Leu Ala Tyr Gly Phe Ala Ala Thr Ser Ile Ala Gly Gly Ser Glu Ser

70	75	80	
tcg tgg tgc tgc gcc tgc tac gc gtaagtcctc tctgccagct accaggaagt			446
Ser Trp Cys Cys Ala Cys Tyr Ala			
85	90		
tggattagcg cgagctaacc tcaactccatc acactccag g ctc acc ttc act tcc			501
		Leu Thr Phe Thr Ser	
		95	
ggt ccc gtc gcc ggc aag aca atg gtg gtg cag tca acg agc act ggc			549
Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr Ser Thr Gly			
100	105	110	
ggc gac ctg gga agt aac cat ttc gat atc gcc atg ccc ggc ggc ggc			597
Gly Asp Leu Gly Ser Asn His Phe Asp Ile Ala Met Pro Gly Gly Gly			
115	120	125	
gtg ggc atc ttc aac ggc tgc agc tcg cag ttc ggc ggc ctc ccc ggc			645
Val Gly Ile Phe Asn Gly Cys Ser Ser Gln Phe Gly Gly Leu Pro Gly			
130	135	140	145
gct caa tac ggc ggc att tcg tcg cgc gac cag tgc gat tcc ttc ccc			693
Ala Gln Tyr Gly Gly Ile Ser Ser Arg Asp Gln Cys Asp Ser Phe Pro			
	150	155	160
gcg ccg ctc aag ccc ggc tgc cag tgg cgg ttt gac tgg ttc cag aac			741
Ala Pro Leu Lys Pro Gly Cys Gln Trp Arg Phe Asp Trp Phe Gln Asn			
	165	170	175
gcc gac aac ccg acg ttc acg ttc cag cag gtg cag tgc ccc gcc gag			789
Ala Asp Asn Pro Thr Phe Thr Phe Gln Gln Val Gln Cys Pro Ala Glu			
180	185	190	
atc gtt gcc cgc tcc ggc tgc aag cgc aac gac gat ggt aac ttc cct			837
Ile Val Ala Arg Ser Gly Cys Lys Arg Asn Asp Asp Gly Asn Phe Pro			
195	200	205	
gtg ttc acc cct cct tcg gga ggc cag tcc tcc tcg tcc tcc tcg tcc			885
Val Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Ser Ser Ser Ser Ser			
210	215	220	225
tcg tcg gcc aag ccg aca tcg acg tcg acg tcc act act tcg acc aaa			933
Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Thr Thr Ser Thr Lys			
	230	235	240
gca acc tcc aca act tcc acc gcc tcg tcc cag acg tcg tcg tcc acc			981
Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser Gln Thr Ser Ser Ser Thr			
	245	250	255
gga gga ggc tgc gca gcg cag cgt tac gct cag tgc ggt ggc atc ggc			1029
Gly Gly Gly Cys Ala Ala Gln Arg Tyr Ala Gln Cys Gly Gly Ile Gly			

260	265	270	
ttc agc gga tgc acc acc tgt gtc tct ggc acc acc tgc aac aag cag			1077
Phe Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys Asn Lys Gln			
275	280	285	
aac gac tac tac tcg cag tgc ctc tga			1104
Asn Asp Tyr Tyr Ser Gln Cys Leu			
290	295		
<210> 32			
<211> 318			
<212> PRT			
<213> 人工序列			
<220>			
<223> 合成构建体			
<400> 32			
Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro			
-20	-15	-10	
Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg Tyr Trp			
-5	-1 1	5	10
Asp Cys Cys Lys Pro Ser Cys Ala Trp Pro Gly Lys Ala Ala Val Ser			
15	20	25	
Gln Pro Val Tyr Ala Cys Asp Ala Asn Phe Gln Arg Leu Ser Asp Phe			
30	35	40	
Asn Val Gln Ser Gly Cys Asn Gly Gly Ser Ala Tyr Ser Cys Ala Asp			
45	50	55	
Gln Thr Pro Trp Ala Val Asn Asp Asn Leu Ala Tyr Gly Phe Ala Ala			
60	65	70	75
Thr Ser Ile Ala Gly Gly Ser Glu Ser Ser Trp Cys Cys Ala Cys Tyr			
80	85	90	
Ala Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Thr Met Val Val			
95	100	105	
Gln Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe Asp Ile			
110	115	120	
Ala Met Pro Gly Gly Gly Val Gly Ile Phe Asn Gly Cys Ser Ser Gln			
125	130	135	
Phe Gly Gly Leu Pro Gly Ala Gln Tyr Gly Gly Ile Ser Ser Arg Asp			
140	145	150	155
Gln Cys Asp Ser Phe Pro Ala Pro Leu Lys Pro Gly Cys Gln Trp Arg			
160	165	170	
Phe Asp Trp Phe Gln Asn Ala Asp Asn Pro Thr Phe Thr Phe Gln Gln			

175	180	185
Val Gln Cys Pro Ala Glu Ile	Val Ala Arg Ser Gly Cys Lys Arg Asn	
190	195	200
Asp Asp Gly Asn Phe Pro Val Phe Thr Pro Pro Ser Gly Gly Gln Ser		
205	210	215
Ser Ser Ser Ser Ser Ser Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr		
220	225	230
Ser Thr Thr Ser Thr Lys Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser		
240	245	250
Gln Thr Ser Ser Ser Thr Gly Gly Gly Cys Ala Ala Gln Arg Tyr Ala		
255	260	265
Gln Cys Gly Gly Ile Gly Phe Ser Gly Cys Thr Thr Cys Val Ser Gly		
270	275	280
Thr Thr Cys Asn Lys Gln Asn Asp Tyr Tyr Ser Gln Cys Leu		
285	290	295

<210> 33

<211> 1104

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<220>

<221> 信号肽

<222> (1) .. (63)

<220>

<221> 外显子

<222> (1) .. (89)

<220>

<221> CDS

<222> (1) .. (89)

<220>

<221> 成熟肽

<222> (64) .. (1101)

<220>

<221> 内含子

<222> (90) .. (167)

<220>

<221> 外显子

<222> (168) .. (416)

<220>
 <221> CDS
 <222> (168)..(416)
 <220>
 <221> 内含子
 <222> (417)..(485)
 <220>
 <221> 外显子
 <222> (486)..(1101)
 <220>
 <221> CDS
 <222> (486)..(1101)
 <400> 33
 atg cgc tct act ccc gtt ctt cgc aca acc ctg gcc gct gca ctt cct 48
 Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro
 -20 -15 -10
 ctg gtc gcc tcc gcg gcc agt ggc agt ggc cag tcc acg ag 89
 Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg
 -5 -1 1 5
 gtatgcgtcc ctcaatgcga atgcctcacc aaacgagcta ggtgtccagg acgccgccca 149
 tgctgactcg tttcccag a tac tgg gac tgc tgc aag ccg tgc tgc gct tgg 201
 Tyr Trp Asp Cys Cys Lys Pro Ser Cys Ala Trp
 10 15 20
 ccc ggg aag gcc gcc gtc agc caa ccg gtc tac gcg tgc gat gcc aac 249
 Pro Gly Lys Ala Ala Val Ser Gln Pro Val Tyr Ala Cys Asp Ala Asn
 25 30 35
 ttc cag cgc ctg tcc gac ttc aat gtc cag tgc ggc tgc aac ggc ggc 297
 Phe Gln Arg Leu Ser Asp Phe Asn Val Gln Ser Gly Cys Asn Gly Gly
 40 45 50
 tgc gcc tac tcc tgc gcc gac cag act ccc tgg gcg gtg aac gac aat 345
 Ser Ala Tyr Ser Cys Ala Asp Gln Thr Pro Trp Ala Val Asn Asp Asn
 55 60 65
 ctc gcc tac ggc ttc gcc gcg acg agc atc gcc ggc ggg tcc gaa tcc 393
 Leu Ala Tyr Gly Phe Ala Ala Thr Ser Ile Ala Gly Gly Ser Glu Ser
 70 75 80
 tgc tgg tgc tgc gcc tgc tac gc gtaagtcctc tctgccagct accaggaagt 446
 Ser Trp Cys Cys Ala Cys Tyr Ala
 85 90
 tggattagcg cgagctaacc tcaactccatc aactccag g etc acc ttc act tcc 501

														Leu	Thr	Phe	Thr	Ser	
														95					
ggt	ccc	gtc	gcc	ggc	aag	aca	atg	gtg	gtg	cag	tca	acg	agc	act	ggc	549			
Gly	Pro	Val	Ala	Gly	Lys	Thr	Met	Val	Val	Gln	Ser	Thr	Ser	Thr	Gly				
				100				105				110							
ggc	gac	ctg	gga	agt	aac	cat	ttc	gat	atc	gcc	atg	ccc	ggc	ggc	ggc	597			
Gly	Asp	Leu	Gly	Ser	Asn	His	Phe	Asp	Ile	Ala	Met	Pro	Gly	Gly	Gly				
				115				120				125							
gtg	ggc	atc	ttc	aac	ggc	tgc	agc	tcg	cag	ttc	ggc	ggc	ctc	ccc	ggc	645			
Val	Gly	Ile	Phe	Asn	Gly	Cys	Ser	Ser	Gln	Phe	Gly	Gly	Leu	Pro	Gly				
				130				135				140				145			
gct	caa	tac	ggc	ggc	att	tcg	tcg	cgc	gac	cag	tgc	gat	tcc	ttc	ccc	693			
Ala	Gln	Tyr	Gly	Gly	Ile	Ser	Ser	Arg	Asp	Gln	Cys	Asp	Ser	Phe	Pro				
				150				155				160							
gcg	ccg	ctc	aag	ccc	ggc	tgc	cag	tgg	cgg	ttt	gac	tgg	ttc	cag	aac	741			
Ala	Pro	Leu	Lys	Pro	Gly	Cys	Gln	Trp	Arg	Phe	Asp	Trp	Phe	Gln	Asn				
				165				170				175							
gcc	gac	aac	ccg	acg	ttc	acg	ttc	cag	cag	gtg	cag	tgc	ccc	gcc	gag	789			
Ala	Asp	Asn	Pro	Thr	Phe	Thr	Phe	Gln	Gln	Val	Gln	Cys	Pro	Ala	Glu				
				180				185				190							
atc	gtt	gcc	cgc	tcc	ggc	tgc	aag	cgc	aac	gac	gat	ggt	aac	ttc	cct	837			
Ile	Val	Ala	Arg	Ser	Gly	Cys	Lys	Arg	Asn	Asp	Asp	Gly	Asn	Phe	Pro				
				195				200				205							
gtg	ttc	acc	cct	cct	tcg	gga	ggc	cag	tcc	tcc	tcg	tcc	tcc	tcg	tcc	885			
Val	Phe	Thr	Pro	Pro	Ser	Gly	Gly	Gln	Ser	Ser	Ser	Ser	Ser	Ser	Ser				
				210				215				220				225			
tcg	tcg	gcc	aag	ccg	aca	tcg	acg	tcg	acg	tcc	act	act	tcg	acc	aaa	933			
Ser	Ser	Ala	Lys	Pro	Thr	Ser	Thr	Ser	Thr	Ser	Thr	Thr	Ser	Thr	Lys				
				230				235				240							
gca	acc	tcc	aca	act	tcc	acc	gcc	tcg	tcc	cag	acg	tcg	tcg	tcc	acc	981			
Ala	Thr	Ser	Thr	Thr	Ser	Thr	Ala	Ser	Ser	Gln	Thr	Ser	Ser	Ser	Thr				
				245				250				255							
gga	gga	ggc	tgc	gca	gcg	cag	aag	tgg	gct	cag	tgc	ggt	ggc	atc	ggc	1029			
Gly	Gly	Gly	Cys	Ala	Ala	Gln	Lys	Trp	Ala	Gln	Cys	Gly	Gly	Ile	Gly				
				260				265				270							
ttc	agc	gga	tgc	acc	acc	tgt	gtc	tct	ggc	acc	acc	tgc	aac	aag	cag	1077			
Phe	Ser	Gly	Cys	Thr	Thr	Cys	Val	Ser	Gly	Thr	Thr	Cys	Asn	Lys	Gln				
				275				280				285							
aac	gac	tac	tac	tcg	cag	tgc	ctc	tga								1104			

Asn Asp Tyr Tyr Ser Gln Cys Leu
 290 295
 <210> 34
 <211> 318
 <212> PRT
 <213> 人工序列
 <220>
 <223> 合成构建体
 <400> 34
 Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro
 -20 -15 -10
 Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg Tyr Trp
 -5 -1 1 5 10
 Asp Cys Cys Lys Pro Ser Cys Ala Trp Pro Gly Lys Ala Ala Val Ser
 15 20 25
 Gln Pro Val Tyr Ala Cys Asp Ala Asn Phe Gln Arg Leu Ser Asp Phe
 30 35 40
 Asn Val Gln Ser Gly Cys Asn Gly Gly Ser Ala Tyr Ser Cys Ala Asp
 45 50 55
 Gln Thr Pro Trp Ala Val Asn Asp Asn Leu Ala Tyr Gly Phe Ala Ala
 60 65 70 75
 Thr Ser Ile Ala Gly Gly Ser Glu Ser Ser Trp Cys Cys Ala Cys Tyr
 80 85 90
 Ala Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Thr Met Val Val
 95 100 105
 Gln Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe Asp Ile
 110 115 120
 Ala Met Pro Gly Gly Gly Val Gly Ile Phe Asn Gly Cys Ser Ser Gln
 125 130 135
 Phe Gly Gly Leu Pro Gly Ala Gln Tyr Gly Gly Ile Ser Ser Arg Asp
 140 145 150 155
 Gln Cys Asp Ser Phe Pro Ala Pro Leu Lys Pro Gly Cys Gln Trp Arg
 160 165 170
 Phe Asp Trp Phe Gln Asn Ala Asp Asn Pro Thr Phe Thr Phe Gln Gln
 175 180 185
 Val Gln Cys Pro Ala Glu Ile Val Ala Arg Ser Gly Cys Lys Arg Asn
 190 195 200
 Asp Asp Gly Asn Phe Pro Val Phe Thr Pro Pro Ser Gly Gly Gln Ser
 205 210 215

Ser Ser Ser Ser Ser Ser Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr
 220 225 230 235
 Ser Thr Thr Ser Thr Lys Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser
 240 245 250
 Gln Thr Ser Ser Ser Thr Gly Gly Gly Cys Ala Ala Gln Lys Trp Ala
 255 260 265
 Gln Cys Gly Gly Ile Gly Phe Ser Gly Cys Thr Thr Cys Val Ser Gly
 270 275 280
 Thr Thr Cys Asn Lys Gln Asn Asp Tyr Tyr Ser Gln Cys Leu
 285 290 295

<210> 35

<211> 1104

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<220>

<221> 信号肽

<222> (1) .. (63)

<220>

<221> 外显子

<222> (1) .. (89)

<220>

<221> CDS

<222> (1) .. (89)

<220>

<221> 成熟肽

<222> (64) .. (1101)

<220>

<221> 内含子

<222> (90) .. (167)

<220>

<221> 外显子

<222> (168) .. (416)

<220>

<221> CDS

<222> (168) .. (416)

<220>

<221> 内含子

<222> (417)..(485)
 <220>
 <221> 外显子
 <222> (486)..(1101)
 <220>
 <221> CDS
 <222> (486)..(1101)
 <400> 35
 atg cgc tct act ccc gtt ctt cgc aca acc ctg gcc gct gea ctt cct 48
 Met Arg Ser Thr Pro Val Leu Arg Thr Thr Leu Ala Ala Ala Leu Pro
 -20 -15 -10
 ctg gtc gcc tcc gcg gcc agt ggc agt ggc cag tcc acg ag 89
 Leu Val Ala Ser Ala Ala Ser Gly Ser Gly Gln Ser Thr Arg
 -5 -1 1 5
 gtatgctcc ctcaatgcga atgcctcacc aaacgagcta ggtgtccagg acgccgccca 149
 tgctgactcg tttcccag a tac tgg gac tgc tgc aag ccg tgc tgc gct tgg 201
 Tyr Trp Asp Cys Cys Lys Pro Ser Cys Ala Trp
 10 15 20
 ccc ggg aag gcc gcc gtc agc caa ccg gtc tac gcg tgc gat gcc aac 249
 Pro Gly Lys Ala Ala Val Ser Gln Pro Val Tyr Ala Cys Asp Ala Asn
 25 30 35
 ttc cag cgc ctg tcc gac ttc aat gtc cag tgc ggc tgc aac ggc ggc 297
 Phe Gln Arg Leu Ser Asp Phe Asn Val Gln Ser Gly Cys Asn Gly Gly
 40 45 50
 tcg gcc tac tcc tgc gcc gac cag act ccc tgg gcg gtg aac gac aat 345
 Ser Ala Tyr Ser Cys Ala Asp Gln Thr Pro Trp Ala Val Asn Asp Asn
 55 60 65
 ctc gcc tac ggc ttc gcc gcg acg agc atc gcc ggc ggg tcc gaa tcc 393
 Leu Ala Tyr Gly Phe Ala Ala Thr Ser Ile Ala Gly Gly Ser Glu Ser
 70 75 80
 tcg tgg tgc tgc gcc tgc tac gc gtaagtctc tetgcccagct accaggaagt 446
 Ser Trp Cys Cys Ala Cys Tyr Ala
 85 90
 tggattagcg cgagctaacc tcactccate acactccag g etc acc ttc act tcc 501
 Leu Thr Phe Thr Ser
 95
 ggt ccc gtc gcc ggc aag aca atg gtg gtg cag tca acg agc act ggc 549
 Gly Pro Val Ala Gly Lys Thr Met Val Val Gln Ser Thr Ser Thr Gly
 100 105 110

ggc gac ctg gga agt aac cat ttc gat atc gcc atg ccc ggc ggc ggc	597
Gly Asp Leu Gly Ser Asn His Phe Asp Ile Ala Met Pro Gly Gly Gly	
115 120 125	
gtg ggc atc ttc aac ggc tgc agc tcg cag ttc ggc ggc ctc ccc ggc	645
Val Gly Ile Phe Asn Gly Cys Ser Ser Gln Phe Gly Gly Leu Pro Gly	
130 135 140 145	
gct caa tac ggc ggc att tcg tcg cgc gac cag tgc gat tcc ttc ccc	693
Ala Gln Tyr Gly Gly Ile Ser Ser Arg Asp Gln Cys Asp Ser Phe Pro	
150 155 160	
gcg ccg ctc aag ccc ggc tgc cag tgg cgg ttt gac tgg ttc cag aac	741
Ala Pro Leu Lys Pro Gly Cys Gln Trp Arg Phe Asp Trp Phe Gln Asn	
165 170 175	
gcc gac aac ccg acg ttc acg ttc cag cag gtg cag tgc ccc gcc gag	789
Ala Asp Asn Pro Thr Phe Thr Phe Gln Gln Val Gln Cys Pro Ala Glu	
180 185 190	
atc gtt gcc cgc tcc ggc tgc aag cgc aac gac gat ggt aac ttc cct	837
Ile Val Ala Arg Ser Gly Cys Lys Arg Asn Asp Asp Gly Asn Phe Pro	
195 200 205	
gtg ttc acc cct cct tcg gga ggc cag tcc tcc ccg tcc tcc ccg tcc	885
Val Phe Thr Pro Pro Ser Gly Gly Gln Ser Ser Pro Ser Ser Pro Ser	
210 215 220 225	
tcg tcg gcc aag ccg aca tcg acg tcg acg tcc act act tcg acc aaa	933
Ser Ser Ala Lys Pro Thr Ser Thr Ser Thr Ser Thr Thr Ser Thr Lys	
230 235 240	
gca acc tcc aca act tcc acc gcc tcg tcc cag acg tcg tcg tcc acc	981
Ala Thr Ser Thr Thr Ser Thr Ala Ser Ser Gln Thr Ser Ser Ser Thr	
245 250 255	
gga gga ggc tgc gca gcg cag cgt tgg gct cag tgc ggt ggc atc gcc	1029
Gly Gly Gly Cys Ala Ala Gln Arg Trp Ala Gln Cys Gly Gly Ile Gly	
260 265 270	
ttc agc gga tgc acc acc tgt gtc tet ggc acc acc tgc aac aag cag	1077
Phe Ser Gly Cys Thr Thr Cys Val Ser Gly Thr Thr Cys Asn Lys Gln	
275 280 285	
aac gac tac tac tcg cag tgc etc tga	1104
Asn Asp Tyr Tyr Ser Gln Cys Leu	
290 295	
<210> 36	
<211> 318	
<212> PRT	

<213> 人工序列

<220>

<223> 合成构建体

<400> 36

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

	255		260		265
Gln Cys Gly Gly Ile Gly Phe Ser Gly Cys Thr Thr Cys Val Ser Gly					
	270		275		280
Thr Thr Cys Asn Lys Gln Asn Asp Tyr Tyr Ser Gln Cys Leu					
	285		290		295

<210> 37

<211> 1044

<212> DNA

<213> 土生梭孢壳霉

<220>

<221> 内含子

<222> (90) .. (167)

<220>

<221> 内含子

<222> (417) .. (485)

<400> 37

```

atgcgctcta ctcccgttct tcgcacaacc ctggccgctg cacttctctt ggtcgcctcc 60
gcggccagtg gcagtgcca gtccacgagg tatgcgtccc tcaatgcgaa tgctcacca 120
aacgagctag gtgtccagga cgccgcccat gctgactcgt ttcccagata ctgggactgc 180
tgcaagccgt cgtgcgcttg gcccgggaag gccgccgta gccaaccggt ctacgcgtgc 240
gatgccaact tccagcgcct gtccgacttc aatgtccagt cgggctgcaa cggcggctcg 300
gcctactcct gcgccacca gactccctgg gcggtgaacg acaatctcgc ctacggcttc 360
gccgcgacga gcatcgccgg cgggtccgaa tctctgtggt gctgcgctg ctacgcgtaa 420
gtcctctctg ccagctacca ggaagttgga ttagecgcgag ctaacctcac tccatcacac 480
tccaggctca cttcacttc cgtccccgtc gccggcaaga caatggtggt gcagtcaacg 540
agcactggcg gcgacctggg aagtaaccat ttcgatatcg ccatgcccgg cggcggcgtg 600
ggcatcttca acggctgcag ctcgcagttc ggcggcctcc ccggcgctca atacggcggc 660
atttcgtcgc gcgaccagtg cgattccttc cccgcgcgc tcaagcccgg ctgccagtgg 720
cggtttgact ggttccagaa cgccgacaac ccgacgttca cgttccagca ggtgcagtgc 780
cccgccgaga tcgttgcccg ctccggctgc aagegcaacg acgactccag cttccccgtc 840
ttcaccccc caagcgggtgg caacgggtggc accgggacgc ccacgtcgac tgcgcctggg 900
tcgggccaga cgtctcccgg cggcggcagt ggctgcacgt ctcagaagtg ggctcagtgc 960
ggtggcatcg gcttcagcgg atgcaccacc tgtgtctctg gcaccacctg ccagaagttg 1020
aacgactact actcgcagtg cctc 1044

```

<210> 38

<211> 951

<212> DNA

<213> 大孢圆孢霉

<400> 38

atgcgctcct cccctgtcct caggacagcg ctgcagcag ccctcccct cgcagcgttg 60
 gcagcagatg gaaagtcgac acgctattgg gactgttgta agccttcgtg ttcgtggcct 120
 ggcaaggcat ccgtgaacca gccggtcttc gcatgttccg caaactteca gcggatctcc 180
 gatccgaacg tcaaatccgg ttgtgatgga ggctccgct acgcgtgtgc cgaccagaca 240
 ccttgggcag tcaacgacaa cttctcctac ggcttcgcag caacttcgat ctcggtggc 300
 aacgaggcct cgtggtgttg tggatgttac gaactcaact tcaactcggg acctgtggca 360
 ggtaagacca tggctgtcca gtcgacgtcg acaggagcg acctcggtac caaccacttc 420
 gatttggcga tgccctggagg aggcgtgggc atcttcgacg gctgttcgcc tcagttcgga 480
 ggcttggcag gcgaccggtg tggcgggtgtg tctcccggg cgcagtgatg ttccttcct 540
 gcagccctca aacccggttg ttactggcga ttcgactggt tcaagaacgc cgacaacct 600
 accttcacgt tccgacaggt gcagtgctcg tcggagttgg tcgccaggac cggctgtcgc 660
 aggaacgatg atggtaactt cctgtgttc accctcctt cgggaggcca gtctctctcg 720
 tctcctcgt cctcgtcggc caagecgaca tcgacgtcga cgtccactac ttcgaccaa 780
 gcaacctca caactccac cgctcgtcc cagacgtcgt cgtccaccgg aggagctgt 840
 gcagcgcagc gttgggcaca gtgtggtggt atcggattct cgggatgtac tacatgtgtg 900
 tcgggcacga cgtgtaacaa gcagaacgac tggactccc agtgtctcta a 951

<210> 39

<211> 935

<212> DNA

<213> 四孢脉孢菌

<400> 39

atgcgctcct ccaactgttct gaaaccggg ctagtggcgg ctcttcctt cgccgttcag 60
 gcccgctccg gatccggcca gtccaccaga tattgggact gctgcaaacc atcttgctcc 120
 tggctccggca aggcttctgt caaccgacc gtctcgtctt gcgatgcaaa caacaacccc 180
 ctgagcgacg ccagtgtaa gtctggatgt gacggcgggt ctgcatacac ctgtgccaac 240
 aactcaccat gggcggtgaa cgaccagctc tctacggct ttgctgccac gaaactcagt 300
 ggtggaaccg agtcgtcttg gtgctgtgcc tgttatgcgt gagtttcgcg tcaacagctg 360
 ggtatcactt gatggctaatt ttatactgca gccttacctt cacttcgggt cctgttgctg 420
 gcaagacat ggtcgttcag tctaccagta ccggcgggta tctcggctcc aaccacttcg 480
 atatcaacat gcccggcggc ggcgtcggcc tgtttgatgg ctgtacacga cagtttggcg 540
 gtctccccgg cgctcaatat ggcggcatca gctcccagc ccagtgcat tcattccctg 600
 ccgcgctcaa gcccggttgc cagtggcgtc tcgactggtt ccagaacgcc gacaacccca 660
 acttcacctt caagcaggtc caatgcccac ccgagctcac ctcccgcacc ggctgcaage 720
 gaaacgacga ctctcaattc cctgttetca ctccgcctc tgggtggagc accaacccct 780
 ctactccgac aaccctccc tcttcaggcg gcggttcgg atgtacggcg gataaatacg 840
 ctcagtgtag cggtcggggg tggctggtg geaccaactg cccgtctgga tcgacctgca 900
 agactatcaa cgattattac caccagtggt cctaa 935

<210> 40

<211> 279

<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 40
gatggtaact tccctgtggt caccctcct tcgggaggec agtcctcctc gtcctcctcg 60
tcctcgtcgg ccaagccgac atcgacgtcg acgtccaacta cttegaccaa agcaacctcc 120
acaacttcca ccgcctcgtc ccagacgtcg tcgtccaccg gaggaggctg tgcagcgcag 180
cgttgggcac agtgtggtgg tatecggatc tcgggatgta ctacatgtgt gtcgggcacg 240
acgtgtaaca agcagaacga ctggtactcc cagtgtctc 279
<210> 41
<211> 41
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 41
gctgcaagcg caacgacgat ggtaacttcc ctgtgttcac c 41
<210> 42
<211> 41
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 42
ggtgaacaca gggaagttac catcgtcgtt gcgcttgacg c 41
<210> 43
<211> 44
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 43
ccctctagat ctcgagtcag agacaactggg agtaccagtc gttc 44
<210> 44
<211> 37
<212> DNA
<213> 人工序列
<220>

<223> 合成构建体
<400> 44
cgctgtccac cggaggaggc tgcacgtctc agaagtg 37
<210> 45
<211> 37
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 45
cacttctgag acgtgcagcc tcttccggtg gacgacg 37
<210> 46
<211> 36
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 46
ccctctagat ctcgaggaa atcaaccagc agtcgc 36
<210> 47
<211> 34
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 47
cgctgtccac cggaggaggc tgcgcagcgc agcg 34
<210> 48
<211> 17
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 48
tcttccggtg gacgacg 17
<210> 49
<211> 41
<212> DNA
<213> 人工序列

<220>
<223> 合成构建体
<400> 49
ccctctagat ctcgagtcag aggcactgcg agtaccagtc g 41
<210> 50
<211> 41
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 50
acacaactgg ggatccacca tgcgetetac tcecgttett c 41
<210> 51
<211> 40
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 51
ccctctagat ctcgagtcag aggcactgcg agtagtagtc 40
<210> 52
<211> 30
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 52
agctggagtc gtcgttcgcc ttgcagccgg 30
<210> 53
<211> 36
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 53
ctccggctgc aaggcgaacg acgactccag cttecc 36
<210> 54
<211> 41
<212> DNA

<213> 人工序列
<220>
<223> 合成构建体
<400> 54
ccctctagat ctcgagttag aggcactgcg agtagtagtc g 41
<210> 55
<211> 40
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 55
ccctctagat ctcgagttag aggcactgcg agtaccagtc 40
<210> 56
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 56
acacaactgg ggatccacca tgcgctcctc ccctg 35
<210> 57
<211> 41
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 57
ccctctagat ctcgagttag agacactggg agtaccagtc g 41
<210> 58
<211> 44
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 58
ccctctagat ctcgagttag agacactggg agtagtagtc gttc 44
<210> 59
<211> 32

<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 59
tggagtcgtc gtttgcgcga cagccggtcc tg 32
<210> 60
<211> 36
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 60
gaccggctgt cgcgcaaacg acgaactccag cttecc 36
<210> 61
<211> 38
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 61
acacaactgg ggatccacca tgcgctcctc cactgttc 38
<210> 62
<211> 40
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 62
ccctctagat ctcgagttag gcacactggt ggtaccaatc 40
<210> 63
<211> 30
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 63
cgcactgagc gtacttctgc gctgcgcagc 30
<210> 64

<211> 20
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 64
aagtacgctc agtgcggtgg 20
<210> 65
<211> 34
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 65
gctgtagccg atgccaccgc actgagcgta acgc 34
<210> 66
<211> 34
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 66
gcgttacgct cagtgcggtg gcatcggeta cage 34
<210> 67
<211> 37
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 67
ggagtttgtg gaggttgctt tggtcgaagt gttggac 37
<210> 68
<211> 38
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 68
cgtccaacac ttcgacaaa gcaacctcca caaactcc 38

<210> 69
<211> 32
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 69
gctgtagccg atgccaccgc actgagccca ac 32
<210> 70
<211> 34
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 70
gcgttgggct cagtgcggtg gcatcggcta cagc 34
<210> 71
<211> 35
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 71
ggacggcgac gtctgggacg aggcggtgga agttg 35
<210> 72
<211> 32
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 72
cttccaccgc ctcgtcccag acgtcgccgt cc 32
<210> 73
<211> 31
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 73

agcgtaacgc tgcgctgcgc agcctcctcc g 31
<210> 74
<211> 33
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 74

ccggaggagg ctgcgcagcg cagcgttacg etc 33
<210> 75
<211> 32
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 75

ccacttctgc gctgcgcagc ctctccggt gg 32
<210> 76
<211> 32
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 76

ccaccggagg aggctgcgca gcgcagaagt gg 32
<210> 77
<211> 32
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体
<400> 77

ggacggggag gacggggagg actggcctcc cg 32
<210> 78
<211> 31
<212> DNA
<213> 人工序列
<220>
<223> 合成构建体

<400> 78

gggaggccag tcctccccgt cctcccctc c 31

<210> 79

<211> 114

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<400> 79

ggctgcgcag cgcagcgttg ggetcagtgc ggtggcatcg gettcagegg atgcaccacc 60

tgtgtctctg gcaccacctg caacaagcag aacgactggt actcgcagtg cctc 114