[19]中华人民共和国国家知识产权局

[51]Int. Cl[7]
C12N 9/16
C12N 15/55 A23K 1/165

[12]发明专利申请公开说明书

[21]申请号 99806375.4

[43]公开日 2001年8月22日

[11]公开号 CN 1309699A

[22]申请日 1999.3.22 [21]申请号 99806375.4
[30]优先权
[32]1998.3.23 [33]DK [31]0407/1998
[32]1998.6.19 [33]DK [31]PA199800806
[32]1998.9.18 [33]DK [31]PA199801176
[32]1999.1.22 [33]DK [31]PA199900091
[86]国际申请 PCT/DK99/00153 1999.3.22
[87]国际公布 WO99/49022 英 1999.9.30
[85]进入国家阶段日期 2000.11.20
[71]申请人 诺维信公司
地址 丹麦鲍斯韦
[72]发明人 阿伦·斯文德森 德克·科斯特鲁瓦
马丁·莱曼 路易斯·帕萨蒙蒂斯
安德鲁·汤姆斯奇 阿道弗斯·范隆
库尔特·沃格尔 马库斯·怀斯
索伦·F·拉森
[74]专利代理机构 柳沈知识产权律师事务所
代理人 巫肖南

权利要求书8页 说明书50页 附图页数51页

[54]发明名称 植酸酶变体

[57]摘要

植酸酶变体、其制备及应用，当按照图1进行对齐排列时，此类植酸酶变体与模式植酸酶相比在许多位点中的至少一处被改变。优选的模式植酸酶为担子菌纲植酸酶和子囊菌纲植酸酶，例如隔孢伏革菌属植酸酶和曲霉属植酸酶。优选的植酸酶变体表现出改变的活性特征，例如改善的特异性活性和/或改善的热稳定性。

ISSN1008-4274

权 利 要 求 书

1、一种植酸酶变体，当按照图1进行对齐排列时，利用与P_lycii相应的位点编号，与模式植酸酶相比，其在至少下列位点之一处被改变：

24; 27; 31; 33; 39; 40; 41; 42; 43; 44; 45; 46; 47; 49; 51; 56; 58; 59; 61; 62; 68; 69; 70; 71; 72; 73; 74; 75; 76; 77; 78; 79; 80; 81; 82; 83; 84; 88; 90; 102; 115; 116; 117; 118; 119; 120; 121; 122; 123; 124; 125; 126; 127; 128; 132; 143; 148; 149; 150; 151; 152; 153; 154; 155; 156; 157; 158; 159; 160; 161; 162; 163; 170f; 170g; 171; 172; 173; 184; 185; 186; 187; 187a; 190; 191; 192; 193; 194; 195; 198; 199; 200; 201; 201a; 201b; 201c; 201d; 201e; 201f; 202; 203; 203a; 204; 205; 211; 215; 220; 223; 228; 232; 233; 234; 235; 236; 237; 238; 239; 242; 243; 244; 246; 251e; 253; 256; 260; 264; 265; 267; 270; 271; 272; 273; 274; 275; 276; 277; 278; 279; 280; 283; 285; 287; 288; 292; 293; 302; 304; 332; 333; 334; 335; 336; 337; 338; 339; 340; 341; 342; 343; 348; 349; 352; 360; 362; 364; 365; 366; 367; 368; 369; 370; 371; 372; 373; 374; 375; 376; 383k; 387; 393; 394; 396; 404; 409; 411; 412; 413; 417; 421; 431.

2、一种植酸酶变体，当按照图1进行对齐排列时，利用与P_lycii相应的位点编号，与模式植酸酶相比，其包括至少下列改变之一：

24C; 27P; 31Y; 33C; 39H,S,Q; 40L,N; 42S,G; 43A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y; 44N; 45D,S; 47Y,F; 49P; 51E,A,R; 56P; 58D,K,A; 59G; 61R; 62V,I; 69Q; 75W,F; 78D,S; 79G; 80K,A; 81A,G,Q,E; 82T; 83A,I,K,R,Q; 84I,Y,Q,V; 88I; 90R,A; 102Y; 115N; 116S; 118V,L; 119E; 120L; 122A; 123N,Q,T; 125M,S; 126H,S,V; 127Q,E,N; 128A,S,T; 132F,I,L; 143N; 148V,I; 151A,S; 152G; 153D,Y; 154D,Q,S,G; 157V; 158D,A; 159T; 160A,S; 161T,N; 162N; 163W; 170fH; 170gA; 171N; 172P; 173Q,S; 184Q,S,P; 185S; 186A,E,P; 187A; 187aS; 190A,P; 193S; 194S,T; 195T,V,L; 198A,N,V; 200G,V; 201D,E; 201a(); 201b(); 201c(); 201d(); 201e(); 201f(); 201eT; 202S,A; 203R,K,S; 203aV,T; 204Q,E,S,A,V; 205E; 211L,V;

215A,P; 220L,N; 223H,D; 228N; 232T; 233E; 235Y,L,T; 236Y,N; 237F; 238L,M; 242P,S; 244D; 246V; 251eE,Q; 253P; 256D; 260A,H; 264R,I; 265A,Q; 267D; 270Y,A,L,G; 271D,N; 273D,K; 275F,Y; 278T,H; 280A,P; 283P; 287A,T; 288L,I,F; 292F,Y; 293A,V; 302R,H; 304P,A; 332F; 336S; 337T,G,Q,S; 338I; 339V,I; 340P,A; 343A,S,F,I,L; 348Y; 349P; 352K; 360R; 362P; 364W,F; 365V,L,A,S; 366D,S,V; 367A,K; 368K; 369I,L; 370V; 373A,S; 374S,A; 375H; 376M; 383kQ,E; 387P; 393V; 396R; 404A,G; 409R; 411K,T; 412R; 417E,R; 421F,Y; 431E.

3、权利要求1或2中任一项的植酸酶变体，其衍生自子囊菌纲植酸酶。

4、权利要求3的植酸酶变体，其衍生自曲霉属植酸酶。

5、权利要求4的植酸酶变体，其中所述模式植酸酶为黑曲霉、无花果曲霉、构巢曲霉、烟曲霉、土曲霉的菌株。

6、权利要求5的植酸酶变体，其中所述模式植酸酶为构巢曲霉 DSM 9743；或者土曲霉以下菌株之任一：CBS 116.46、DSM 9076、CBS220.95。

7、权利要求6的植酸酶变体，其中所述模式植酸酶为示于图10中的构巢曲霉植酸酶序列；或者示于图12中的土曲霉植酸酶序列。

8、权利要求3的植酸酶变体，其中所述模式植酸酶为 Thermomyces lanuginosus、嗜热踝节菌或嗜热毁丝霉的菌株。

9、权利要求8的植酸酶变体，其中所述模式植酸酶为 Thermomyces lanuginosus CBS 586.94；或者嗜热踝节菌以下菌株之任一：ATCC 20186、ATCC 74338；或者嗜热毁丝霉以下菌株之任一：ATCC 34625、ATCC 74340。

10、权利要求9的植酸酶变体，其中所述模式植酸酶为示于图14中的 Thermomyces lanuginosus 植酸酶序列；或者示于图13中的嗜热踝节菌植酸酶序列；或者示于图7中的嗜热毁丝霉植酸酶序列。

11、权利要求3的植酸酶变体，其中所述模式植酸酶为子囊菌纲共有植酸酶序列。

12、权利要求1或2中任一项的植酸酶变体，其衍生自担子菌纲植酸酶。

13、权利要求 12 的植酸酶变体，其中所述模式植酸酶为 Paxillus involutus、绒毛栓菌、平田头菇或 Peniophora lycii 的菌株。

14、权利要求 13 的植酸酶变体，其中所述模式植酸酶为绒毛栓菌 CBS 100232 或 Paxillus involutus CBS 100231。

15、权利要求 14 的植酸酶变体，其中所述模式植酸酶为图 4 中的绒毛栓菌植酸酶序列或者图 2 或 3 中的 Paxillus involutus 植酸酶序列。

16、权利要求 1 或 2 中任一项的植酸酶变体，其包括至少下述改变之一：

R24C; V27P; H39Q,S; L40N; G42S;

Q43A,C,D,E,F,G,H,I,K,L,M,N,P,R,S,T,V,W,Y; Y44N; A45D,S; F47Y; S49P; A51E,R; V56P; A58D,K; V62I; S69Q; Y75W,F; D78S; S79G; K80A; G81A,Q,E; K82T; K83A,I,R,Q; Y84Q,I,V; E90R,A; D115N; D116S; T118V,L; P119E; F120L; E122A; Q123N,T; L125S,M ; V126H,S; N127Q,E; S128A,T; F132I,L; I148V; S151A; S153D,Y; S154Q,D,G; I157V; A158D; S159T: G160A,S; K161T,N; K162N; F163W; R170fH; Q171N; G173Q,S; S184P,Q; E185S; A186E,P; S187A; T190P,A; P193S; G194S,T; T195V,L; V198A,N; E200G,V; D201E; S201d(); E201e(),T;

L201f()；优选所有三个缺失；A202S；D203R,K,S；

D203aV,T; V204Q,E,S,A; T211L,V; S215AP; L220N; D223H; T228N; T235Y,L; Y236N; L237F; M238L; S242P; I246V; K251eE,Q; H260A; I264R; N265Q,A; Q270Y,A,L,G; S271D,N; K273D; Y275F; H278T; A280P; T287A; Q288L,I,F; Y292F; A293V; H302R; P304A; N336S; G337S,T,Q; I339V; S340P,A; F343A,S,F,I,L; N349P; N360R; T362P; F364W; S365V,L,A; S366D,V; A367K; W368K; T369I,L; A373S: S374A; R375H; L376M; Q383kE; P404A,G; T411K; R417E; F421Y; A431E.

17、权利要求 16 的植酸酶变体，其模式植酸酶为衍生自曲霉的植酸酶，优选衍生自无花果曲霉或黑曲霉。

18、权利要求 17 的植酸酶变体，其模式植酸酶为衍生自无花果(黑)曲霉 NRRL 3135、黑曲霉 ATCC 9142 或黑曲霉 ATCC 74337 之任一的植酸酶。

19、权利要求 18 的植酸酶变体，其模式植酸酶为图 11 的无花果曲霉植酸酶序列。

20、权利要求1或2中任一项的植酸酶变体，其植酸酶变体包括至少下述改变之一：

A24C; V27P; H39,S,Q; L40N; G42S; Q43C,D,E,F,H,K,M,P,R,S,W,Y; Y44N; S45D; F47Y; S49P; E51A,R; L56P; K58D,A; D59G; I62V; S69Q; Y75W,F; S78D; S79G; K80A; S81A,G,Q,E; K82T; K83A,I,Q,R; Y84Q,V,I; V88K; A90R; F102Y; D115N; D116S; T118V,L; P119E; F120L; E122A; Q123N,T; L125S,M; V126H,S; N127Q,E; S128A,T; F132,I,L; S143N; I148V; S151A; S153D,Y; D154Q,S,G; I157V; A158D; S159T; G160A,S; E161T,N; K162N; F163W; G170fH; ()171N; N173Q,S; T172P; P184Q,S; E185S; S186A,E,P; E187A; T187aS; T190P,A; G194S,T; V195L,T; K198A,N,V; E200G,V; A201D,E; S201d(); Q201e(),T；L201f()；优选所有三个缺失；G202S,A；D203R,K,S; E203aV,T; V204Q,E,S,A; A205E; L211V; A220L,N; H223D; T228N; E232T; D233E; V235Y,L,T; V236Y,N; L237F; M238L; C242P,S; T246V; Q251eE,Q; Q256D; H260A; K264R,I; K265Q,A; N267D; Q270Y,A,L,G; S271D,N; G273D,K; Y275F; Y278T,H; A280P; A287T; Q288L,I,F; F292Y; T293A,V; R302H; P304A; F332F; N336S; S337T,G,Q; M338I; V339I; S340P,A; F343A,S,I,L; N349P; E352K; S360R; K362P; Y364W,F; S365V,L,A; A366D,V,S; S367A,K; W368K; V369I,L; G373S,A; R375H; A376M; K383kQ,E; D404A,G; K411T; I393V; L412R; K417E,R; W421F,Y; G431E.

21、权利要求20的植酸酶变体，其衍生自曲霉属植酸酶，优选利用衍生自烟曲霉的模式植酸酶。

22、权利要求21的植酸酶变体，其模式植酸酶为衍生自烟曲霉以下菌株之任一的植酸酶：ATCC 13073、ATCC 32722、ATCC 58128、ATCC 26906或ATCC 32239。

23、权利要求22的植酸酶变体，其模式植酸酶为图8的烟曲霉植酸酶序列。

24、权利要求1或2中任一项的植酸酶变体，其植酸酶变体包括至少下述改变之一：

G24C; V27P; H39S,Q; L40N; G42S;
Q43A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y; Y44N; S45D; Y47F;
S49P; E51A,R; V56P; D58K,A; D59G; V62I; S69Q; Y75W,F; S78D;
S79G; K80A; S81A,G,Q,E; K82T; A83I,Q,K,R; Y84,Q,I,V; A90R;
D115N; D116S; T118V,L; F119E; P120L; E122A; N123Q,T; M125S;
V126H,S; N127Q,E; S128A,T; Y132F,I,L; K143N; I148V; S151A;
S153D,Y; D154Q,S,G; I157V; A158D; S159T; A160S; E161T,N; K162N;
F163W; G170fH; S170gA; Q171N; H173Q,S; P184Q,S; E185S;
G186A,E,P; S187A; G187aS; T190P,A; H193S; G194S,T; T195V,L;
A198N,V ; E200G,V ; D201E ; S201d() ; E201e(),T ; L201f() ; 优选所有三个; G202S,A ; D203R,K,S ; D203aV,T ; V204Q,S,A,E ; L211V ;
A215P; L220N; D223H, T228N; E232T; D233E; V235Y,L,T; Y236N;
L237F; M238L; P242S; E244D; E251e,Q; A256D; H260A; R264I; Q265A;
Q270Y,A,L,G; S271D,N; G273D,K; Y275F; Y278T,H; A280P; A287T;
Q288L,I,F; F292Y; A293V; R302H; P304A; N336S; S337T,Q,G; M338I;
I339V; S340P,A; F343A,S,I,L; N349P; A352K; S360R; E362P;
Y364W,F; S365V,L,A; A366D,V,S; S367K,A; W368K; T369I,L; G373S,A;
A374S; R375H; A376M; Q383kE; A404G; K411T; E417R; F421Y; A431E.

25、权利要求24的植酸酶变体，其模式植酸酶为一种子囊菌纲共有植酸酶。

26、权利要求25的植酸酶变体，其模式植酸酶为图9的子囊菌纲共有序列“conphys”。

27、权利要求1或2中任一项的植酸酶变体，其植酸酶变体包括至少下述改变之一：
V24C; F27P; ()31Y; F33C; D39H,S,Q; S40L,N; A42S,G;
A43C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y; Y44N; T45D,S; Y47F;
Q51E,A,R; K58D,A; K61R; I62V; F75W; S78D; A80K; G81A,Q,E;
R83A,I,Q,K; I84Y,Q,V; V88I; K90R,A; L102Y; D115N; D116S; V118L;
P119E; F120L; L123N,T,Q; S125M; S126H,V; Q127E,N; A128S,T;
T132F,I,L; E143N; V148I; S151A; S152G; S153D,Y; N154D,Q,S,G;
D158A; S159T; A160S; T161N; ()170fH; ()170gA; ()171N; H173Q,S;
H172P; S184Q,P; E185S; S186A,E,P; L187A; ()187aS; T190P,A;

D193S; A194S,T; M195T,V,L; N198A,V; G200V; S201D,E ()201eT; S202A; D203R,K,S; P203aV,T; Q204E,S,A,V; T205E; I211L,V; P215A; L220N; Q223D,H; A232T; D233E; S235Y,L,T; N236Y; L237F; I238L,M; A242P,S; E244D; I246V; ()251eE,Q; N256D; P260A,H; A264R,I; Q265A; E267D; G270Y,A,L; L332F; D271N; D273K; F275Y; T278H; Y280A,P; Y283P; V287A,T; Q288L,I,F; Y292F; I293A,V; E302R,H; P304A; L332F; N336S; Q337T,S,G; M338I; I339V; A340P; S343A,F,I,L; F348Y; N349P; S352K; P360R; R362P; W364F; V365L,A,S; T366D,V,S; S367K,A; R368K; L369I; T370V; S373A; A374S; R375H; S383kQ,E; T387P; A396R; G404A; L409R; T411K; L412R; E417R; Y421F.

28、权利要求27的植酸酶变体，其模式植酸酶为衍生自平田头菇的植酸酶。

29、权利要求27的植酸酶变体，其模式植酸酶为衍生自平田头菇CBS 900.96的植酸酶。

30、权利要求29的植酸酶变体，其模式植酸酶为图5的平田头菇植酸酶序列。

31、权利要求1-2中任一项的植酸酶变体，其植酸酶变体包括至少下述改变之一：

F24C; V27P; L31Y; I33C; S39H,Q; N40L; G42S; P43A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y; Y44N; D45S; F47Y; E51A,R; E58D,K,A; T61R; V62I; W75F; S78D; A80K; R81Q,E,G,A; S82T; R83A,I,Q,K; Q84Y,V,I; V88I; K90R,A; A115N; D116S; L118V; P119E; F120L; N123T,Q; S125M; H126S,V; Q127E,N; T128A,S; M132F,I,L; G143N; V148I; A151S; D153Y; Q154D,S,G; D158A; S159T; S160A; T161N; ()170fH; ()170gA; S171NG172P; E173Q,S; Q184S,P; E185S; E186A,P; G187A; ()187aS; T190P,A; N193S; N194S,T; M195T,V,L; N198A,V; V200G; D201E; ()201eT; G202S,A; D203R,K,S; ()203aV,T; E204Q,S,A,V; S205E; V211L; N215A,P; L220N; A223D,H; S232T; D233E; L235Y,T; T236Y,N; L237F; M238L; P242S; L246V; ()251eE,Q; A260H; V264R,I; S265Q,A; E267D; Y270A,L,G; D271N;

D273K; F275Y; G278T,H; P280A; A283P; T287A; Q288L,I,F; Y292F; V293A; G302R,H; A304P; N336S; T337Q,S,G; M338I; V339I; P340A; A343S,F,I,L; F348Y; N349P; A352K; E360R; R362P; W364F; V365L,A,S; D366V,S; S367K,A; L369I; S373A; G374A,S; ()383kQ,E; E387P; A396R; G404A; V409R; E411K,T; L412R; E417R; Y421F; A431E.

32、权利要求31的植酸酶变体，其模式植酸酶为衍生自 Peniophora lycii 的植酸酶。

33、权利要求32的植酸酶变体，其模式植酸酶为衍生自 Peniophora lyciiCBS 686.96 的植酸酶。

34、权利要求33的植酸酶变体，其模式植酸酶为图6中的 Peniophora lycii 植酸酶序列。

35、一种植酸酶多肽，其包括根据上述权利要求中任一项的植酸酶变体。

36、一 DNA 构建体，其包括编码根据权利要求1-34中任一项的植酸酶变体的 DNA 序列。

37、一种重组表达载体，其包括根据权利要求36的 DNA 构建体。

38、一种宿主细胞，其由根据权利要求36的 DNA 构建体或根据权利要求37的载体转化。

39、一种用于制备植酸酶变体的方法，该方法包括在允许产生植酸酶变体的条件下培养根据权利要求38的宿主细胞，以及从培养液中回收植酸酶。

40、一种饲料或食物，其包括至少一种根据权利要求1-34中任一项的植酸酶变体。

41、一种用于制备根据权利要求40的饲料或食物的方法，其中将上述至少一种植酸酶变体添加至食物或饲料组分中。

42、一种组合物，其包括至少一种根据权利要求1-34中任一项的植酸酶变体。

43、权利要求42的组合物，其适用于制备食物或饲料。

44、权利要求42-43中任一项的组合物，其为一种动物饲料添加剂。

45、一种用于降低动物粪尿中植酸盐水平的方法，其包括用有效量的根据权利要求40的饲料或根据权利要求41可获得的饲料来喂养动物。

46、权利要求 1-34 中任一项的植酸酶变体；或者权利要求 42-43 中任一项的组合物用于从植酸酶底物释放三价磷的应用。

47、一种转基因植物或植物部分，其能够表达根据权利要求 1-34 中任一项的植酸酶变体。

说　明　书

植酸酶变体

发明领域

本发明涉及植酸酶的变体，特别是子囊菌纲植酸酶变体和担子菌纲植酸酶变体、相应的克隆的 DNA 序列、制备此类植酸酶变体的方法以及其在多种工业用途上的应用。

发明背景

植酸或肌醇-1,2,3,4,5,6-六二氢磷酸盐(或缩写为肌醇六磷酸盐)是植物种子中肌醇的主要来源和磷酸盐的基本储存形式。植酸钙镁是一种混合的肌醇钾、镁和钙盐。

植酸的磷酸盐部分螯合二价和三价阳离子，例如金属离子，即营养所必需的离子钙、铁、锌和镁以及痕量矿物质锰、铜和钼。

植酸及其盐，植酸盐，通常不被代谢，即其三价磷和螯合的金属离子均非营养上可利用的。

相应地，食物和饲料的制备需要补充无机磷酸盐，而且通常还必需补充营养所必需的离子，例如铁和钙。

此外，植酸盐的三价磷穿过这些动物的胃肠道并随粪尿排泄，引起不希望的环境磷酸盐污染，导致例如水环境的富养化和藻类的大量生长。

植酸或植酸盐，这些在本申请中除非另外指明，作为同义词使用或随机使用的术语，均通过植酸酶降解。

植物以及微生物的植酸酶产生已有报道。在微生物中，产植酸酶细菌以及产植酸酶真菌是已知的。

有一些对属于真菌门子囊菌纲的产植酸酶丝状真菌的描述。特别是，有几篇产植酸酶的子囊菌纲曲霉属，例如土曲霉的文献(Yamada 等， 1986，农业生物化学(Agric. Biol. Chem. 322:1275-1282))。此外，黑曲霉变种泡盛曲霉的植酸酶基因的克隆和表达也已经被描述(Piddington 等， 1993 ，基因(Gene)133:55-62)。 EP 0420358 描述了无花果(黑)曲霉的植酸酶的克隆和表达。 EP 0684313 描述了子囊菌纲黑曲霉、嗜热毁丝霉、土曲霉的植酸酶的克

隆和表达。此外还给出了构巢曲霉、嗜热踝节菌、烟曲霉以及另一株土曲霉的植酸酶的一些部分序列。

在 WO 97/35017 中描述了 Thermomyces lanuginosus 植酸酶的克隆和表达。

目前需要具有改变的性质或特性的植酸酶，例如热稳定性增加的、最佳 pH 改变的(体外处理需要一高的最佳 pH ，体内胃肠道中的处理需要一低的最佳 pH)和/或较高特异性活性的植酸酶。

发明概述

在一方面，本发明提供植酸酶变体，其特性与所谓的模式植酸酶相比是改变的。

任何模式植酸酶，其与本文特别公开的十三种模式植酸酶具有某种相似性，都可以用作此类变体的模式。

在另一方面，本发明涉及衍生自 Cladorrhinum foecundissimum 的一种新型植酸酶。

在另一方面，本发明还提供编码这些植酸酶变体和该植酸酶的 DNA 序列，以及它们的制备方法。

最后，本发明大体上还涉及植酸酶和植酸酶变体从任何植酸酶底物中释放三价磷，特别是从植酸盐或植酸中释放无机磷酸盐的应用。

附图简述

在下面本发明的详细描述中，对图进行了说明，其中

图 1 是十三种特定植酸酶序列的对齐排列(多序列对齐排列，按照程序 PileUp ；空隙权(GapWeight)： 3.000; 空隙长度权(GapLengthWeight)：0.100)；

图 2 该图显示一种植酸酶(“ P_involtus-Al ”)的氨基酸和 DNA 序列，其衍生自 97 年 11 月 28 日保藏的 Paxillus involutus CBS 100231 株；包含其全长 cDNA 序列的表达质粒 pYES2.0 被转化到 97 年 11 月 12 日保藏的大肠杆菌 DSM 11842 株(见 WO 98/28409)中；

图 3 该图显示另一种植酸酶(“ P_involtus-A2 ”)的氨基酸和 DNA 序列，其衍生自 97 年 11 月 28 日保藏的 Paxillus involutus CBS 100231 株；包含其

全长 cDNA 序列的表达质粒 pYES2.0 被转化到 97 年 11 月 12 日保藏的大肠杆菌 DSM 11843 株(见 WO 98/28409)中；

图 4 该图显示一种植酸酶(“绒毛栓菌(T_pubescens)”)的氨基酸和 DNA 序列，其衍生自 97 年 11 月 28 日保藏的绒毛栓菌 CBS 100232 株；包含其全长 cDNA 序列的表达质粒 pYES2.0 被转化到 97 年 11 月 12 日保藏的大肠杆菌 DSM 11844 株(见 WO 98/28409)中；

图 5 该图显示一种植酸酶(“平田头菇(A_pediades)”)的氨基酸和 DNA 序列，其衍生自 96 年 12 月 4 日保藏的平田头菇 CBS 900.96 株；包含其全长 cDNA 序列的表达质粒 pYES2.0 被转化到 96 年 12 月 2 日保藏的大肠杆菌 DSM 11313 株(见 WO 98/28409)中；

图 6 该图显示一种植酸酶(“P_lycii”)的氨基酸和 DNA 序列，其衍生自 96 年 12 月 4 日保藏的 Peniophora lycii CBS 686.96 株；包含其全长 cDNA 序列的表达质粒 pYES2.0 被转化到 96 年 12 月 2 日保藏的大肠杆菌 DSM 11312 株(见 WO 98/28409)中；

图 7 该图等同于 EP 0684313 的图 2，显示一种植酸酶(“嗜热毁丝霉(M_thermophila)”)的氨基酸和 DNA 序列，其衍生自 97 年 3 月 14 日重新保藏的嗜热毁丝霉 ATCC 48102 株(= ATCC 74340)；

图 8 该图显示一种植酸酶(“烟曲霉(A_fumigatus)”)的氨基酸和 DNA 序列，其衍生自烟曲霉 ATCC 13073 株(见 EP 0897985)；

图 9 该图显示一种子囊菌属共有植酸酶(在本申请中称为“共有植酸酶 A(consphyA)”)的氨基酸(“Conphys”)和 DNA 序列 (见 EP 0897985)；

图 10 该图显示一种植酸酶(“构巢曲霉(A_nidulans)”)的氨基酸和 DNA 序列，其衍生自构巢曲霉 DSM 9743 株(见 EP 0897985)；

图 11 该图等同于 EP 0420358 的图 8，显示一种植酸酶(“无花果曲霉(A_ficuum)”)的氨基酸和 DNA 序列，其衍生自无花果曲霉 NRRL-3135 株；

图 12 该图等同于 EP 0684313 的图 1，显示一种植酸酶(“土曲霉(A_terreus)”)的氨基酸和 DNA 序列，其衍生自土曲霉 CBS 220.95 株；

图 13 该图显示一种植酸酶(“嗜热踝节菌(T_thermo)”)的氨基酸和 DNA 序列，其衍生自 97 年 3 月 14 日重新保藏的嗜热踝节菌 ATCC 20186(= ATCC 74338)株(见 EP 0897985)；

图 14 该图等同于 WO 97/35017 的图 2，显示一种植酸酶

(“T_lanuginosa”)的氨基酸和DNA序列，其衍生自Thermomyces lanuginosus CBS 586.94株；将包含其全长cDNA序列的质粒转化到96年2月23日保藏于NRRL的大肠杆菌DH5 α(pMWR46) 株B-21527中；

图15该图显示一种植酸酶(“C_foecundissimum”)的氨基酸和DNA序列，其衍生自1997年1月23日保藏的Cladorrhinum foecundissimum CBS 427.97株；将包含其全长cDNA序列的表达质粒pYES2.0转化到1999年3月17日保藏的大肠杆菌DSM 12742株中；

图16该图显示C_foecundissimum植酸酶与作为模式植酸酶的嗜热毁丝霉的对齐排列，利用程序GAP gcg(空隙权3.000; 长度权0.100)；以及

图17显示如何能将C_foecundissimum植酸酶粘贴于图1的对齐排列中。

发明详述

植酸酶

在本申请中，植酸酶是一种催化植酸盐(肌醇六磷酸盐)水解为(1)肌醇和/或(2)其单、二、三、四和/或五磷酸盐以及(3)无机磷酸盐的酶。为了简短，在后文中上述化合物有时分别由IP6、I、IP1、IP2、IP3、IP4、IP5和P指代。这意味着通过植酸酶的作用，IP6被降解成P加一个或多个IP5、IP4、IP3、IP2、IP1和I组分。或者，将共携有n个磷酸基团于p、q、r、……位点处的肌醇表示为Ins(p,q,r,..)Pn。为了方便，Ins(1,2,3,4,5,6)P6(植酸)缩写为PA。

根据酶命名法数据库ExPASy〔一个酶命名法相关信息的收集处，其主要基于国际生物化学和分子生物学联合会(IUBMB)命名委员会为描述每种类型鉴定过并为其提供了一个EC(酶学委员会(Enzyme Commission))编号的酶所提出的建议〕，已知有两种不同类型的植酸酶：一种被称为的3-植酸酶(肌醇六磷酸盐3-磷酸水解酶，EC 3.1.3.8)和一种被称为的6-植酸酶(肌醇六磷酸盐6-磷酸水解酶，EC 3.1.3.26)。3-植酸酶首先水解D-3位的酯键，而6-植酸酶首先水解D-6或L-6位的酯键。

“植酸酶”或“表现植酸酶活性的多肽或酶”包括任何能够影响无机磷酸盐或三价磷从各种肌醇磷酸盐中释放的酶。此类肌醇磷酸盐(植酸酶底物)的实例是植酸及其任意盐，例如，植酸钠或植酸钾或混合盐。肌醇的单、二、

三、四或五磷酸盐的任何立体异构体都可以用作植酸酶底物。优选的植酸酶底物是植酸及其盐。

根据上述的定义，可以利用其中采用了这些底物之一的任何测试来确定植酸酶的活性。在本申请中(除非另外指定)植酸酶活性是以 FYT 单位来确定的，一个FYT是在下述条件下： pH5.5 ;温度37 ℃ ;底物是浓度为0.0050mol/l的植酸钠($C_6H_6O_{24}P_6Na_{12}$)，每分钟释放 1 μ mol 无机正磷酸盐的酶的量。在试验部分描述了一种适当的植酸酶检测法。

本发明提供遗传工程植酸酶，如在所附的权利要求中所述。

遗传工程植酸酶是非自然存在的植酸酶，其与模式植酸酶，例如野生型植酸酶不同。遗传工程植酸酶包括，但不限于，通过定点诱变、基因重组、随机诱变等制备的植酸酶。

本发明还提供了 DNA 构建体、载体、宿主细胞以及制备这些遗传工程植酸酶和植酸酶变体的方法及其应用。

植酸酶变体是一种多肽或酶或其一个片段，其表现出植酸酶活性而且与模式植酸酶相比其被改变。

改变意指通过一个或多个氨基酸或肽的置换、缺失、插入和/或增加的方式（在每种情况下，一个或多个氨基酸发生改变）。此类置换、缺失、插入、加入可以通过本领域所知的任何方法来实现，例如基因重组、随机诱变、定点诱变等。

可衍生出植酸酶变体的模式或亲代植酸酶可以是任何植酸酶，例如，一种野生型植酸酶或其一种衍生物、突变体或变体，包括等位变体和物种变体，及其基因工程变体，它们可以通过例如定点诱变、随机诱变、重组等来制备。

在模式植酸酶概念中还包括任何杂交或嵌合植酸酶，即组合了至少两种植酸酶的部分氨基酸序列的植酸酶。

杂交植酸酶可能组合了至少两种子囊菌纲植酸酶、至少两种担子菌纲植酸酶或至少一种子囊菌纲植酸酶和至少一种担子菌纲植酸酶的部分氨基酸序列。这些可衍生出部分氨基酸序列的子囊菌纲和担子菌纲植酸酶可以是例如本文所提及的那些特定植酸酶中的任何一种。

在本申请中，单纯衍生自子囊菌纲植酸酶的杂交的、重组的、随机诱变的、定点诱变的或其它遗传工程改造的植酸酶仍然是子囊菌纲植酸酶；而单

纯衍生自作为模式的担子菌纲植酸酶的杂交的、重组的、随机诱变的、定点诱变的或其它遗传工程改造的植酸酶仍然是担子菌纲植酸酶。将任何衍生自至少一种子囊菌纲植酸酶以及至少一种担子菌纲植酸酶的杂交体称为混合的子囊菌纲/担子菌纲植酸酶，而且此种植酸酶在本申请中也是一种模式植酸酶。

类似地，衍生自一种或多种曲霉类植酸酶的杂交的、重组的、随机诱变的、定点诱变的或其它遗传工程改造的植酸酶仍然是曲霉植酸酶；而衍生自本文提到的任何其它分类亚群的杂交的、重组的、随机诱变的、定点诱变的或其它遗传工程改造的植酸酶也将被指定为该分类亚群的植酸酶。

进而，在本申请中“衍生自”意指由所指定生物的一个株产生的或可产生的植酸酶，以及由分离自该株的DNA序列编码并在由该DNA序列转化的宿主生物中所产生的植酸酶。最后，该术语意指一种植酸酶，其由一合成的和/或cDNA来源的DNA序列编码而且具有所指定植酸酶的鉴定特征。

模式植酸酶优选为一种可以按下述与图1中十三种植酸酶(均为特别优选的模式植酸酶)的任何一个进行对齐排列的植酸酶。

优选的野生型植酸酶(即，未经重组、或重排或其它遗传工程改造的植酸酶)与这十三种植酸酶中任一种的38 - 403号(隔孢伏革菌的编号)氨基酸序列具有至少40 %，更优选至少50 %，还更优选至少60 %，特别是至少70 %，尤其是至少80 %的相似性或同源性，优选等同性，而在一最优选的实施方案中是至少90 %的相似度。

优选的重组的或重排的或者其它遗传工程改造的模式植酸酶与这十三种植酸酶中任何一种的38 - 49、63 - 77、274 - 291、281 - 300以及389 - 403号(隔孢伏革菌的编号)部分序列具有至少60 %，更优选至少70 %，还更优选至少80 %，特别是至少90 %的相似性或同源性，优选等同性。

在一优选的实施方案中，相似性程度是基于与十三种植酸酶中任一种的完整氨基酸序列的比较。

相似性或同源性，或者等同性的程度可以利用任何本领域已知的对齐排列程序来确定。一个优选的对齐排列程序是在GCG版本8程序包(Wisconsin Package程序手册，版本8，1994年8月，遗传学计算机组，575 Science Drive，Madison，Wisconsin，USA 53711，还见Needleman, S.B.和Wunsch, C.D.，(1970)，分子生物学杂志(Journal of Molecular Biology)，48，443 -

453)中提供的 GAP 。利用 GAP 按照下述设置进行多肽序列比较： 3.000 的空隙权和 0.100 的空隙长度权。

还优选的是一种野生型模式植酸酶，其包括一段由下述 DNA 序列编码的氨基酸序列，该 DNA 序列与编码图 1 中十三种特定植酸酶序列中任一种的 38 - 403 位(隔孢伏革菌的编号)氨基酸序列的 DNA 序列杂交。

更优选的模式植酸酶是一种遗传工程植酸酶，其包括一段由下述 DNA 序列编码的氨基酸序列，该 DNA 序列与编码图 1 中十三种特定植酸酶序列中任一种的 38 - 49 位、63 - 77 位、274 - 291 位、281 - 300 位以及 389 - 403 位的氨基酸序列(隔孢伏革菌的编号)的 DNA 序列杂交。

在一优选的实施方案中，杂交是与十三种植酸酶中任何一种的完整植酸酶编码部分的杂交。

用于确定一给定的DNA或RNA序列是否与一特定的核苷酸或寡核苷酸探针“杂交”的适当试验条件包括在 5 × SSC(氯化钠/柠檬酸钠)中将含有待测 DNA 片段或 RNA 的滤膜预先浸泡 10 分钟，(J. Sambrook， E.F. Fritsch，和 T. Maniatis ， 1989 ，分子克隆，实验室手册(Molecular Cloning ， A Laboratory Manual)，第二版，冷泉港，纽约)，并在 5 × SSC 、 5 × Denhardt 液(Sambrook 等， 1989)、 0.5% SDS 和 100 μ g/ml 变性的超声处理鲑鱼精 DNA(Sambrook 等， 1989)的溶液中预杂交，随后在大约 45 ℃于含有随机引物引发的（ Feinberg, A.P.和 Vogelstein, B.(1983) 分析生物化学(Anal. Biochem.)132 ： 6 - 13 ） ^{32}P-dCTP 标记的（特异活性>1 × 10^9cpm/μg ）探针 10ng/ml 的相同溶液中杂交 12 小时。

然后在至少 55 ℃(低严谨度)、至少 60 ℃(中严谨度)、至少 65 ℃(中/高严谨度)、至少 70 ℃(高严谨度)或至少 75 ℃(极高严谨度)将滤膜在 2 × SSC 、 0.5 % SDS 中洗涤两次 30 分钟。

利用 X 线胶片可检测到在这些条件下寡核苷酸探针所杂交的分子。

需要注明的是，某种特定的植酸酶变体实际上不需要由特定的模式植酸酶来制备，因为在本申请中该模式植酸酶用作一种“样板植酸酶”。在后来与模式植酸酶比较时变体表现出至少一种本文所指的改变便足够了。

图 1 的对齐排列是利用程序 PileUp （ Wisconsin Package 程序手册，版本 8， 1994 年 8 月，遗传学计算机小组， 575 Science Drive， Madison， Wisconsin， USA 53711 ）按照 3.000 的空隙权和 0.100 的空隙长度权作出。

当排列一种新型模式植酸酶或一种新型植酸酶变体时，与新型植酸酶(变体)一起的可以包括所有十三个序列，或者与新型植酸酶(变体)一起的至少包括图 1 的十三个序列之一。

对图 1 的新型模式植酸酶(或植酸酶变体)进行对齐排列的优选程序如下：将新型模式植酸酶与图 1 的十三条序列中与新型模式植酸酶同源性程度最高的特定序列进行对齐排列。为了对两个序列进行同源性程度的计算以及“根据图 1 进行对齐排列”，优选利用下面提到的程序 GAP 。首先对齐排列两条序列，利用第一个对齐排列的结果(按照对齐排列所规定的，将相同和同源的氨基酸残基彼此对齐)，将新型模式植酸酶(或植酸酶变体)添加(粘贴)到图 1 的对齐排列中，随后相应位点便很容易判断。

实施例 7 显示一个如何向图 1 的对齐排列中增加一种新型模式植酸酶并推断其相应植酸酶变体的实例。

同实施例 7 类似，可以对齐排列其他模式植酸酶并推断变体。特别是对于下述模式植酸酶更是如此：黑曲霉变种泡盛曲霉的植酸酶(美国专利号 5,830,733)； WO 98/06858 的芽孢杆菌植酸酶； WO 98/20139 的大豆植酸酶； WO 98/05785 的玉米植酸酶； WO 97/38096 的曲霉植酸酶； WO 98/13480 的亚克红曲霉植酸酶； EP 0699762 的西方许旺氏酵母植酸酶等。

当利用本文所描述的对齐排列来比较一种模式植酸酶和一种推测的植酸酶变体时，可以鉴定相应的氨基酸位点，即模式植酸酶的模式位点和变体的变异位点——在对齐排列中只是简单地将相应的模式位点和变异位点一个置于另一个上方。如果模式位点的模式氨基酸同变异位点的变异氨基酸不同，则称在给定的位点发生了改变。这些位点优选的改变表现为氨基酸置换、缺失或增加。

在至少一个位点被改变意味着在一个或多个位点，即在一个、两个、三个、四个、五个、六个、七个、八个、九个、十个、十一个、十二个等，高达所有列出的 N 个位点处被改变。该定义包括其任何可能的亚组合，例如任何有两个置换的组、任何有三个置换的组、任何有四个置换的组等，——直至任何有(N-1)个位点置换的组。

在本申请中，不论模式植酸酶为何，均利用同植酸酶 P_lycii 相应的编号对所有序列，以及包括图 1 的十三条序列进行编号。这些“隔孢伏革菌编号”同“对齐排列编号”一起在图 1 中标出。P_lycii 的编号起始于 M1 ，终

止于 E439 处。

如上面所解释的，对齐排列揭示了在除 P_lycii 外的各种植酸酶序列中哪些位点同给出的 P_lycii 位点等价或相应。

本文将氨基酸的置换标注为如“ 3S ”，其表示在位点 3 处可以用氨基酸 S 置换“原始”或模式位点 3 处的氨基酸，不论其为何种氨基酸。因此，该置换可以在相应的变异位点产生一个 S 。现在鉴于图 1 上的对齐排列，对于所示各种植酸酶的类似如“ 3S ”的置换将做如下解释(第一个标明的氨基酸是在“隔孢伏革菌位点” 3 处的“原始”或模式氨基酸)：

P_involtus_A1:	F3S(3 号 F 由 S 置换)
P_involtus_A2:	L3S
绒毛栓菌：	M1S
平田头菇：	M1S
P_lycii ：	多余的(已经为一 S)
烟曲霉：	T5S
共有植酸酶 A ：	V5S
构巢曲霉：	T5S
无花果曲霉_NRRL3135 ：	A5S
土曲霉：	A5S
嗜热踝节菌：	L5S
T_lanuginosa ：	V11S
嗜热毁丝霉：	G5S

然而，下文对上述的特定置换将做如下定义(始终采用隔孢伏革菌编号)：

P_involtus_A1:	F3S
P_involtus_A2:	L3S
绒毛栓菌：	M3S
平田头菇：	M3S
P_lycii ：	多余的(已经为一 S)
烟曲霉：	T3S
共有植酸酶 A ：	V3S
构巢曲霉：	T3S

无花果曲霉_NRRL3135： A3S
土曲霉： A3S
嗜热踝节菌： L3S
T_lanuginosa： V3S
嗜热毁丝霉： G3S

此外，类似如“ 3S,F,G ”的符号意指将正讨论的模式植酸酶位点 3(隔孢伏革菌编号)处的氨基酸由 S 、 F 或 G 的任何一个来置换，即，例如标示“ 3S,F,G ”可以被认为完全等同于标示“ 3S,3F,3G ”。

类似()3S 的符号意指在该序列中与隔孢伏革菌 3 号相应的位点处增加了氨基酸 S(在实际序列中的空隙处)，反之为缺失(S3())。

在隔孢伏革菌植酸酶序列比图 1 中某些其他植酸酶存在更大缺失的区域的情况下，例如在位点 201 和 202(隔孢伏革菌编号)间的区域中，中间的位点(在其他序列中的氨基酸残基)通过在最后一个隔孢伏革菌位点编号处添加 a 、 b 、 c 、 d 等小写字母来编号，例如对于嗜热毁丝霉植酸酶： E201 ； G201a ； P201b ； Y201c ； S201d ； T201e ； I201f ； G202 ； D203 等。

在本申请书的一个优先权申请中，有两个位点编号的小错误；根据上述定义，在第一个优先权申请中所提及的 204 和 205 位点(隔孢伏革菌编号)为错误指定；它们本应分别编号为 203a 和 204 。因此，在本申请中 204 由 203a 、 205 由 204 代替。

本发明的一个优选的植酸酶变体包括以下氨基酸序列，其包括，优选含有，一个或多个下述氨基酸置换：

24C; 27P; 31Y; 33C; 39H,S,Q; 40L,N; 42S,G; 43A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y; 44N; 45D,S; 47Y,F; 49P; 51E,A,R; 56P; 58D,K,A; 59G; 61R; 62V,I; 69Q; 75W,F; 78D,S; 79G; 80K,A; 81A,G,Q,E; 82T; 83A,I,K,R,Q; 84I,Y,Q,V; 88I; 90R,A; 102Y; 115N; 116S; 118V,L; 119E; 120L; 122A; 123N,Q,T; 125M,S; 126H,S,V; 127Q,E,N; 128A,S,T; 132F,I,L; 143N; 148V,I; 151A,S; 152G; 153D,Y; 154D,Q,S,G; 157V; 158D,A; 159T; 160A,S; 161T,N; 162N; 163W; 170fH; 170gA; 171N; 172P; 173Q,S; 184Q,S,P; 185S; 186A,E,P; 187A; 187aS; 190A,P; 193S; 194S,T; 195T,V,L; 198A,N,V; 200G,V ； 201D,E ； 201a 、 201b 、 201c 、 201d 、 201e 、 201f 中的至少一个，优选全部的缺失； 201eT ； 202S,A ； 203R,K,S ；

203aV,T; 204Q,E,S,A,V; 205E; 211L,V; 215A,P; 220L,N; 223H,D; 228N; 232T; 233E; 235Y,L,T; 236Y,N; 237F; 238L,M; 242P,S; 244D; 246V; 251eE,Q; 253P; 256D; 260A,H; 264R,I; 265A,Q; 267D; 270Y,A,L,G; 271D,N; 273D,K; 275F,Y; 278T,H; 280A,P; 283P; 287A,T; 288L,I,F; 292F,Y; 293A,V; 302R,H; 304P,A; 332F; 336S; 337T,G,Q,S; 338I; 339V,I; 340P,A; 343A,S,F,I,L; 348Y; 349P; 352K; 360R; 362P; 364W,F; 365V,L,A,S; 366D,S,V; 367A,K; 368K; 369I,L; 370V; 373A,S; 374S,A; 375H; 376M; 383kQ,E; 387P; 393V; 396R; 404A,G; 409R; 411K,T; 412R; 417E,R; 421F,Y; 431E.

在一优选的实施方案中，这是以模式植酸酶在所指定的位点处并未包括上面提到的氨基酸置换或增加或缺失为条件的。或者，以每个位点的模式氨基酸尚不是此处提出的变异氨基酸为条件的。但由于“改变”的表述，可以说这些附加条件事实上已经为上述措辞所固有。

权利要求16-34的各种优选植酸酶变体包括（优选含有或具有）以下氨基酸序列，它们包括或含有在各权利要求中所列的一个或多个氨基酸的置换，增加或缺失。

在一优选的实施方案中，各种植酸酶变体包括这些置换中的1、2、3、4、5、6、7、8、9个或者甚至10个；或者10-15、15-20、20-30个或者甚至30-50个的大量置换；最终高达60、70、80或90个置换。

在另一优选的实施方案中，各种植酸酶变体的氨基酸序列包括本文下面所列置换亚群的一个或多个置换；或者归类至两个或多个亚群的置换的组合。

通常，由于被本文所描述的一个或多个置换所修饰，优选变体的氨基酸序列并非“包括”、“含有”或“具有”，而是“基本上由”或者“由”图1中特定模式植酸酶“组成”。

在本发明中，担子菌意指担子菌纲的微生物。此担子菌纲与诸如子囊菌纲一起，都属于真菌界。

分类学问题可以通过查阅下面所列的参考或通过查阅互联网的下述地址：http;//www3.ncbi.nlm.nih.gov/Taxonomy/tax.html所提供的真菌分类数据库(NIH数据库(Entrez))来明确。

对于担子菌纲的定义，以Julich，1981，担子菌纲高级分类法(Higher Taxa of Basidiomycetes)；Ainsworth和Bisby's(编)真菌字典(Dictionary of the

Fungi)， 1995 ， Hawksworth ， D.L.， P.M. Kirk ， B.C. Sutton 和 D.N. Pegler ；或者 Hansen 和 Knudsen(编)， Nordic Macromycetes ，第 2 卷(1992)和第 3 卷(1997)作为参考。优选参考 Hansen 和 Knudsen 。

对于子囊菌纲的定义，以上面引用的 Ainsworth 和 Brisby 或者由 Eriksson ， O.E.和 D.L. Hawksworth ，第 16 卷， 1998 编辑的系统子囊菌纲(Systema Ascomycetum)作为参考。优选参考 Eriksson 等。

大体上，归类于以上所列参考（包括数据库）的子囊菌纲/担子菌纲的微生物是本申请中的子囊菌/担子菌。

一些曲霉菌株很难分类，因为它们是渐变的，因此它们应被归类为半知菌。然而，一旦发现有性配对部分，便按照分类法将其重新分类。例如，构巢曲霉的有性型是构巢皮胞囊果属(子囊菌门的不整囊菌纲，散囊菌目，发菌科)。优选子囊菌门的这些亚群以及子囊菌门的核菌纲，粪壳目，Lasiosphaeriaceae 科。

本文所用的“子囊菌纲”和类似物包括曲霉属、 Thermomyces 、毁丝霉属、踝节菌属的渐变的，并因此应被归类于半知菌的菌株。

优选的担子菌纲植酸酶是在 WO 98/28409 中以“发明详述”为标题的章节的最开始处所列举的那些。

按照在附图简述下所列各文献的指导，可以制备编码这十三种特别列举的模式植酸酶以及其他模式植酸酶的 DNA 序列。

利用本领域众所周知的各种方法可以从产生目的植酸酶的任何细胞或微生物中分离编码模式植酸酶的 DNA 序列。首先，应该利用来自产该植酸酶生物的染色体 DNA 或信使 RNA 构建一基因组 DNA 和/或 cDNA 文库。然后，如果已知该植酸酶的氨基酸序列，可以合成同源的标记寡核苷酸探针，并用来从由该生物制备的基因组文库中鉴定植酸酶编码克隆。或者，利用低严谨度的杂交和洗涤条件，可以将含有同一已知植酸酶基因同源的序列的标记寡核苷酸探针用作探针，来鉴定植酸酶编码克隆。

另一种鉴定植酸酶编码克隆的方法还可以包括将基因组 DNA 片段插入到一表达载体，例如质粒中，利用所获的基因组 DNA 文库转化植酸酶阴性的细菌，然后将转化的细菌涂平板至含有植酸酶底物的琼脂上，由此可以鉴定表达植酸酶的克隆。

或者，可以通过已建立的标准方法来合成编码该酶的 DNA 序列，例如，

由 S.L. Beaucage 和 M.H. Caruthers(1981)所描述的磷酰胺方法或者由 Matthes 等(1984)所描述方法。在磷酰胺方法中，对寡核苷酸进行合成，例如在自动 DNA 合成仪中，将其纯化、退火、连接并克隆到适当的载体中。

最后，按照标准的技术，通过连接合成的、基因组的或 cDNA 来源的片段(若适当，这些片段对应于整个 DNA 序列的各个部分)所制备的 DNA 序列可以是基因组和合成混合来源的、合成和 cDNA 混合来源的或者基因组和 cDNA 混合来源的。也可以利用特定的引物通过聚合酶链式反应(PCR)来制备 DNA 序列如在美国专利 4,683,202 或 R.K. Saiki 等(1988)中所描述。

编码本发明中植酸酶变体的 DNA 可以通过本领域所知的方法来制备，例如定点诱变。一旦分离出编码目的模式植酸酶的 DNA 序列，并确定了预期的突变位点，就可以利用合成的寡核苷酸引入突变。这些寡核苷酸含有预期突变位点的侧翼核苷酸序列；在寡核苷酸合成过程中插入突变的核苷酸。在一特定的方法中，在一携带植酸酶编码基因的载体中建立一个桥接植酸酶编码序列的 DNA 单链空隙。随后将带有预期突变的合成核苷酸与单链 DNA 的同源部分退火。然后用 DNA 聚合酶 I(Klenow 片段)补平其余的空隙，并用 T4 连接酶连接该构建体。该方法的一个特定的实例描述于 Morinaga 等(1984)中。US 4,760,025 公开了通过进行基因盒的局部改动对编码多个突变的寡核苷酸的引入。然而，因为可以引入各种长度的多种寡核苷酸，所以利用 Morinaga 方法可以在任何时候引入更加多样的突变。

在 Nelson 和 Long(1989)中描述了向编码一预期模式植酸酶的 DNA 序列中引入突变的另一种方法。其涉及一个 3 步产生的含有预期突变的 PCR 片段，这些突变是通过利用一化学合成的 DNA 链作为 PCR 反应中的引物之一来引入的。由 PCR 生成的片段，可以通过限制性核酸内切酶切割来分离携带该突变的 DNA 片段并将其重新插入到一表达质粒中。

突变编码一种模式植酸酶的 DNA 序列的另一方法是随机诱变。在翻译成所示氨基酸序列的基因的至少三个部分或在整个基因内适当地进行局部随机诱变或特定区域随机诱变。

利用本领域已知的任何方法可以方便地对编码模式植酸酶的 DNA 序列进行随机诱变。

同上面的相关，本发明进一步还涉及产生模式植酸酶的变体的方法，其中该变体优选表现出下述的改变的特征，该方法包括：

(a)使编码该模式植酸酶的DNA序列进行定点诱变，或用Nelson和Long PCR诱变方法诱变或者随机诱变，

(b)在宿主细胞中表达步骤(a)中获得的突变DNA序列，并

(c)筛选表达相对于该模式植酸酶性质改变的植酸酶变体的宿主细胞。

当利用随机诱变时，本发明上述方法的步骤(a)优选利用添加的引物进行。

例如，可以通过利用一种适宜的物理或化学诱变剂、利用一种适宜的寡核苷酸、或者通过将该DNA序列进行PCR生成的诱变来进行随机诱变。此外，可以通过利用这些诱变剂的任意组合来进行随机诱变。例如，该诱变剂可诱导转换、颠换、倒位、scrambling、缺失、和/或插入。

适于本目的的物理或化学诱变剂的实例包括紫外线(UV)辐射、羟胺、N-甲基-N'-硝基-N-亚硝基胍(MNNG)、wycb甲基羟胺、亚硝酸、乙基甲磺酸盐(EMS)、亚硫酸氢钠、甲酸以及核苷酸类似物。利用此类诱变剂时，通常通过在适于诱变发生的条件下于所选诱变剂存在下温育待诱变的编码亲代酶的DNA序列，并筛选具有预期性质的突变DNA来进行诱变。

当利用寡核苷酸来进行诱变时，在寡核苷酸合成过程中于该寡核苷酸待改变的位点处应添加或增强标定三种非亲代核苷酸。实施添加或增强标定或以避开无用氨基酸的密码子。通过任何公开的技术，利用认为适当的如PCR、LCR或者任何DNA聚合酶和连接酶，可以将添加或增强标定的寡核苷酸整合到编码植酸酶的DNA中。

优选利用“持续随机添加”进行添加，其中在每个位点上的野生型和突变的百分比是预先确定的。此外，根据对某些核苷酸引入的优选以及由此的对一个或多个特定氨基酸残基的优选可以指导添加。例如，添加可在每个位点引入90%野生型和10%突变。在选择添加设计时还应根据遗传学以及蛋白质结构限制加以考虑。添加设计可以通过利用DOPE程序来进行，其特别保证防止引入终止密码子。

在利用PCR生成的诱变时，在增加核苷酸错误增强标定的条件下对编码模式植酸酶的化学处理或非处理的基因进行PCR(Deshler 1992；Leung等，技术(Technique)，第1卷，1989，11-15页)。

通过例如向一诱变株中转化一含有亲代糖基化酶的质粒，培养转有质粒的诱变株并从诱变株中分离突变的质粒，大肠杆菌(Fowler等，分子普通遗传

学(Molec. Gen. Genet.)， 133， 1974， 179-191 页)、酿酒酵母或任何其他微生物的诱变株可以用于对编码模式植酸酶 DNA 进行随机诱变，即通过例如向一诱变株中转化一含有亲代糖基化酶的质粒，培养转有质粒的诱变株并从诱变株中分离突变的质粒。随后可以将突变的质粒转化到表达生物体中。

待诱变的 DNA 序列可以方便地存在于由表达该模式植酸酶的生物体所制备的基因组或 cDNA 文库中。或者，该 DNA 序列可以存在于一适当的载体上，例如质粒或噬菌体，其本身可以同诱变剂进行温育或者暴露于诱变剂。待诱变的 DNA 还可以通过整合到细胞的基因组中或者通过存在于驻留细胞中的一种载体上而存在于该宿主细胞中。最后，待诱变的 DNA 可以处于分离的形式。当然待进行随机诱变的 DNA 序列优选为一 cDNA 或一基因组 DNA 序列。

有些情况下，在进行表达步骤 b)或筛选步骤 c)之前扩增突变 DNA 序列可能是方便的。此类扩增可以按照本领域已知的方法进行，目前优选的方法是利用基于亲代酶的 DNA 或氨基酸序列制备的寡核苷酸引物进行 PCR 生成的扩增。

在同诱变剂进行温育或暴露于诱变剂后，通过在便于表达的条件下培养携带该 DNA 序列的适当宿主细胞来表达该突变 DNA 。用于此目的的宿主细胞可以是由选择性地存在一载体上的突变 DNA 序列转化的那种，或者是在诱变处理过程中携带编码亲代酶 DNA 序列的那种。适宜的宿主细胞的实例如下：革兰氏阳性细菌，例如枯草芽孢杆菌、地衣芽孢杆菌、迟缓芽孢杆菌、短芽孢杆菌、嗜热脂肪芽孢杆菌、嗜碱芽孢杆菌、解淀粉芽孢杆菌、凝结芽孢杆菌、环状芽孢杆菌、灿烂芽孢杆菌、巨大芽孢杆菌、苏云金芽孢杆菌，浅青紫链霉菌或鼠灰链霉菌；以及革兰氏阴性细菌，例如大肠杆菌。

突变的 DNA 序列进而可以包括编码允许该突变 DNA 序列表达的功能的一段 DNA 序列。

利用局部随机诱变可以将随机诱变便利地定位于模式植酸酶的一部分。例如，当已经确定了酶的某些区域对于该酶的一已知特性特别重要时，以及当预计修饰会产生具有改变特性的变体时，这可能是有益的。在亲代酶的三级结构已被阐明而且同酶的功能相关时通常可以确定此类区域。

通过利用上面所描述的 PCR 生成的诱变技术或本领域所知的任何其他适当的技术可以方便地进行局部的或者区域特定的随机诱变。或者，编码待

修饰的DNA序列中一部分的DNA序列可以被分离，例如通过插入到一适当的载体中，随后通过利用上面讨论的任何诱变方法可以对该部分进行诱变。

为了改变如模式植酸酶的特异性活性而进行区域特定的随机诱变时，可以适当地针对同图1中所列氨基酸序列的下述氨基酸残基相应的密码子位点而进行：

残基：41-47、68-80、83-84、115-118、120-126、128、149-163、184-185、191-193、198-201e、202-203、205、235-236、238-239、242-243、270-279、285、288、332-343、364-367、369-375、394。

区域：41-47、68-80、120-128、149-163、270-279、332-343、364-375。

随机诱变可以通过下面的步骤进行：

1. 在亲代酶中选择待修饰的目的区域

2. 在选定的区域中确定突变位点和非突变位点

3. 例如根据待构建变体的预期的稳定性和/或特性，确定应进行哪类突变

4. 选择结构上合理的突变

5. 根据步骤4调整通过步骤3选定的残基

6. 利用一适当的添加算法分析核苷酸分布

7. 如果必要，根据遗传密码的实践调整所需的残基，例如考虑到由遗传密码产生的限制，如为了避免终止密码子的引入；技术人员将清楚一些密码子的组合在实践中不能使用并将需要调整。

8. 制备引物。

9. 利用这些引物进行随机诱变

10. 通过筛选预期的改善的性质来选择所获的植酸酶变体

用于步骤6中的适宜的添加算法是本领域众所周知的。Tomandl，D.等，1997，计算机辅助分子设计杂志(Journal of Computer-Aided Molecular Design)11：29-38描述了一种此类算法。另一种算法是DOPE(Jensen，LJ，Andersen，KV，Svendsen，A和Kretzschmar，T(1998)核酸研究(Nucleic Acids Research)26：697-702)。

编码本发明的模式植酸酶或植酸酶变体的DNA序列可以利用表达载体，重组表达载体来表达，该表达载体通常包括编码启动子、操纵子、核糖

体结合位点、翻译启始信号的控制序列以及选择性地一个抑制子基因或各种激活子基因。

重组表达载体应该是可以方便地进行重组 DNA 操作的任何载体，而且载体的选择通常将有赖于其将被引入的宿主细胞。因此，载体可以是一种自主复制载体，例如质粒、噬菌体或者染色体外元件。或者，载体可以是当引入宿主细胞时被整合到宿主细胞基因组中并同其整合于其中的染色体一起复制的那种。

在载体中， DNA 序列应该是可操作地同一适当的启动子序列相连。启动子可以是在选定的宿主细胞中表现出转录活性的任何 DNA 序列，而且可以来自编码宿主细胞同源或者异源性蛋白质的基因。用于指导编码本发明中植酸酶变体的 DNA 序列转录，特别是在一细菌宿主中转录的适当启动子的一个实例是大肠杆菌的 lac 乳糖操纵子。对于在真菌宿主中的转录，有效启动子的实例是来自编码米曲霉 TAKA 淀粉酶基因的那些。

本发明的表达载体还可以包括一个适当的转录终止子以及在真核细胞中可操纵地连接至编码本发明中植酸酶变体 DNA 序列的的多聚腺苷酸化序列。终止和多聚腺苷酸化序列可以适当地衍生自同启动子相同的来源。

载体进一步可以包括使载体能够在所选宿主细胞中复制的 DNA 序列。此类序列的实例是质粒 pUC19 、 pACYC177 、 pUB110 、 pE194 、 pAMB1 和 pIJ702 的复制起点。

载体还可能包括一可筛选的标志，例如，其产物在宿主细胞中弥补一缺陷的基因，如枯草芽孢杆菌或地衣芽孢杆菌的 dal 基因，或者提供抗生素抗性的基因，如氨苄青霉素抗性。进而，载体可以包括曲霉筛选标记，如 amdS 、 argB 、 niaD 和 sC ，或者筛选可以通过共转化来完成，例如按照在 WO 91/17243 所描述的。

用于分别连接本发明中编码植酸酶变体的 DNA 构建体、启动子、终止子和其他元件并将它们插入到含有复制所需信息的适当载体中的方法，对于本领域的技术人员是众所周知的(参阅，例如， Sambrook 等，(1989))。

本发明的细胞，其包括按照上面所确定的其包括本发明的 DNA 构建体或表达载体，在本发明的植酸酶变体的重组制备中便利地用作宿主细胞。通过将本发明中编码变体的 DNA 构建体整合到宿主染色体中(一个或多个拷贝)，可以方便地用该 DNA 构建体转化该细胞。这种整合通常被认为具有优

势，因为 DNA 序列更有可能稳定地保留在细胞中。可以按照传统的方法，例如通过同源或异源重组，将 DNA 构建体整合到宿主染色体中。或者，可以用如上所述的与不同的宿主细胞类型相关表达载体来转化该细胞。

一种分离的 DNA 分子或者一段“克隆的 DNA 序列”、“一个 DNA 构建体”、“一段 DNA 片段”或“一段分离的 DNA 序列”指一种 DNA 分子或序列，其可以按照在遗传工程中用来将 DNA 片段由其天然位置重新定位至其将被复制的一不同位点的标准的克隆方法来进行克隆。该术语通常指基本上没有其他核酸序列的一段核酸序列，例如通过琼脂糖凝胶电泳确定的；至少大约 20%的纯度，优选至少 40%的纯度，更优选大约 60%的纯度，还要更优选 80%的纯度，最优选大约 90%的纯度，更加更优选大约 95%的纯度。克隆的步骤可能包括切除和分离含有编码该多肽的核酸序列的预期核酸片段，将该片段插入一载体分子，以及将重组载体整合到一种宿主细胞中，在该宿主细胞中该核酸序列的多个拷贝或克隆将被复制。核酸序列可以是基因组、cDNA 、RNA 、半合成、合成来源的，或者其任意的组合。

术语“载体”规定为包括象“核酸构建体”、“ DNA 构建体”、“表达载体”或“重组载体”那样的术语/对象。

核酸构建体包括本发明中可操作地连接至一个或多个调控序列的核酸序列，这些调控序列能够指导编码序列在同调控序列相容的条件下于一种合适的宿主细胞中表达。

本文“核酸构建体”定义为一段核酸分子，单链或双链，其分离自一个天然存在的基因或者其经修饰含有以自然界不会存在的方式结合或并置的核酸片段。

在核酸构建体含有本发明中编码序列表达所需的所有调控序列时，术语核酸构建体与术语表达基因盒应该是同义的。

本文所定义的术语“编码序列”首先包括一段序列，当将其置于上述调控序列的调控下时，其被转录成 mRNA 并翻译成本发明的一种多肽。编码序列的边界一般是通过 5’-末端的翻译启始密码子 ATG 和 3’-末端的翻译终止密码子来确定的。编码序列可以包括，但不限于， DNA 、 cDNA 以及重组核酸序列。

术语“调控序列”本文定义为包括所有对于核酸序列的编码序列表达所必需或有利的所有成分。每个调控序列对于编码多肽的核酸序列可以是原来

或外来的。此类调控序列包括，但不限于，一个先导序列、一个多聚腺苷酸化序列、一个前肽序列、一个启动子、一个信号序列以及一个转录终止子。至少，调控序列包括一个启动子以及转录和翻译终止信号。为了引入便于连接调控序列和编码多肽的核酸序列编码区的特定限制性酶切位点，调控序列可以带有接头。

“宿主细胞”或“重组宿主细胞”包括亲代细胞的任何子代，其因在复制过程中发生的突变而与亲代细胞不同。

细胞优选被包括本发明核酸序列的载体转化，随后该载体整合到该宿主染色体中。

“转化”意指在宿主细胞中引入包括本发明核酸序列的载体，以便该载体作为染色体整合子或自我复制的染色体外载体来保持。整合通常被认为是一种优势，因为核酸序列更可有能在细胞中稳定保持。载体在宿主染色体中的整合可以通过如上所述的同源或非同源重组来进行。

宿主细胞可以是单细胞微生物，例如原核生物，或者非单细胞微生物，例如真核生物。真核生物细胞的实例是哺乳动物细胞、昆虫细胞、植物细胞或真菌细胞。有用的哺乳动物细胞包括中国仓鼠卵巢(CHO)细胞、 HeLa 细胞、幼仓鼠肾(BHK)细胞、 COS 细胞或者例如可由美国典型培养物保藏中心所提供的许多其他的永生化细胞系。

在一优选的实施方案中，宿主细胞是真菌细胞。

真菌细胞可以通过涉及原生质体形成、原生质体转化以及细胞壁再生的方法，按照其原本的形式进行转化。

本发明还涉及转基因植物、植物的部分例如植物的种子或者植物细胞，其已经由编码本发明中植酸酶的 DNA 序列所转化以便表达或产生该酶。根据本发明的使用和食物/饲料权利要求，此类植物或植物部分的组合物和应用也属于本发明的范围，特别是其用做饲料和食物或者其添加剂。

转基因植物可以是双子叶或单子叶的，简称为双的子叶或单子叶植物。主要的兴趣是，可作为潜在的食物或种子组分，而且含有植酸的此类植物。种子组分的正常植酸水平为 0.1-100g/kg，或者更通常为 0.5-50g/kg，最通常为 0.5-20g/kg 。单子叶植物的实例是禾本科植物，例如草地早熟禾(早熟禾属)，禾本科牧草，如狐茅属、黑麦草属，温带禾本，如，剪股颖属，以及禾谷类，如小麦属、燕麦属、黑麦属、大麦属、稻属、高粱属和玉米属(玉米)。

双子叶植物的实例是豆科植物，例如羽扇豆属、豌豆属、菜豆属和大豆，以及十字花科植物(十字花科)，例如花椰菜、油籽油菜以及密切相关的模式生物拟南芥。

此类转基因植物等能够降解其自身的植酸，相应地减低了对于向含有这类植物的食物或饲料中添加此类酶的需要。优选将植物或植物部分，例如种子，磨碎或粉碎，并且可能在添加至食物或饲料前或者在应用，如摄入前还要浸泡，以便使酶促降解的速度适应实际应用。

如果需要，还可以从植物中回收植物所产生的酶。在某些情况下，优选从植物中回收，以便在随后可能的沉淀步骤中保证制剂的热稳定性。

植物部分的实例是茎、胼胝、叶、根、果实、种子、块茎等。但是任何植物组织也都包括在本定义中。

任何植物细胞，不论组织来源，都包括在上面植物细胞的定义中。

在本发明的范围内还包括这些植物、植物部分和植物细胞的子代。

技术人员将知道如何构建一个 DNA 表达构建体来插入指定植物，特别是考虑到是否该酶应以一种组织特异性的方式来分泌。同这个评价有关的是在植物的表达室中植酸酶的稳定性(pH 稳定性、由内源性蛋白酶的可降解性等)。如果需要的话，他还将能够选择适当的调节序列，如启动子和终止子序列以及信号或转位序列(Tague 等，植物，生理，(Plant， Phys.)， 86， 506，1988)。

利用任何已知的方法，可以用该 DNA 构建体转化植物、植物部分等。这种方法的一个实例是通过病毒或细菌载体的转化，例如经遗传学设计而包括编码本发明中植酸酶的基因的农杆菌属的细菌种。直接将植酸酶 DNA 引入植物细胞或植物组织的方法在本领域也是已知的，例如显微注射和电穿孔(Gasser 等，科学(Science)， 244 ， 1293 ； Potrykus ，生物/技术(Bio/Techn.)8， 535， 1990 ； Shimamoto 等，自然(Nature)， 338， 274，1989)。

在转化后，利用技术人员已知的任何方法筛选转化子，随后将它们再生成为完整的植物。

然后，这些植物等以及它们的子代携带植酸酶编码 DNA 作为它们遗传学部件的一部分。

大体上，以 WO 9114782A 和 WO 9114772A 作为参考。

根癌农杆菌介导的基因转移是选择来产生转基因双子叶植物的方法(综述见 Hooykas 和 Schilperoort， 1992，植物分子生物学(Plant Mol. Biol.)19：15-38)，然而，其还可以用于转化单子叶植物。由于宿主范围的限制，在根癌农杆菌辅助下转化单子叶植物通常是不可能。此处需使用其他的方法。选择来产生转基因单子叶植物的方法是对胚胎胼胝或发育中胚胎的微粒轰击(包被有转化 DNA 的显微金或钨颗粒)(Christou， 1992. 植物杂志(Plant J.)2：275-281；Shimatoto， 1994. 生物技术的当前观点(Curr. Opin. Biotechnol.)5：158-162；Vasil 等， 1992. 生物技术(Bio/Technology)10：667-674)。

用于将游离 DNA 转移入这些植物的其他系统也是优选的方法，包括病毒载体(Joshi 和 Joshi， 1991. FEBS Lett. 281：1-8)、通过聚乙二醇或电穿孔的原生质体转化(综述见 Potyrkus， 1991. 植物生理学和植物分子生物学年鉴(Annu. Rev. Plant Physiol. Plant Mol. Biol.)42：205-225)、向叶肉原生质体中的 DNA 显微注射(Crossway 等， 1986. Mol. Gen. Genet. 202：79-85)以及向禾谷类植物嫩的花分檗中的 DNA 显微注射(de la Pena 等， 1987. 自然 325：274-276)。

通常将编码本发明植酸酶变体的 cDNA 或基因置于由在目标植物中有活性的一适当启动子和一适当终止子(转录的终止)组成的一个表达基因盒中(例如 Pietrzak 等， 1986. 核酸研究(Nuleic Acids Res.)14：5857-5868)。单子叶植物通过微粒轰击将该基因盒(当然包括合适的筛选标志，见下面)转化入该植物。如果是双子叶植物，首先将表达基因盒置于一提供 T-DNA 边界以及一个合适的筛选标志的适宜载体中，该载体又被转化入根癌农杆菌。双子叶植物将通过载有由 T-DNA 为侧翼的表达基因盒以及筛选标志的农杆菌按照标准的方法(例如 Akama 等， 1992. 植物细胞报道(Plant Cell Reports)12：7-11)来转化。最近对 T-DNA 由农杆菌向植物细胞的转移已经进行了综述(Zupan 和 Zambryski， 1995. 植物生理学(Plant Physiol.)107：1041-1047)。用于通过农杆菌进行植物转化的载体是商品化的或者可以从许多构建此类载体的实验室获得(例如， Deblaere 等， 1985. 核酸研究 13：4777-4788；综述见 Klee 等， 1987. 植物生理学年鉴(Annu. Rev. Plant Physiol.)38：467-486)。

可利用的植物启动子：根据操作的步骤，可能需要器官和/或细胞特异性

的表达以及适当的发育控制和环境控制。例如，理想的是在玉米胚乳等中表达植酸酶 cDNA 。最常用的启动子是组成型的 35S-CaMV 启动子(Franck 等，1980. 细胞(Cell)21 ： 285-294)。表达在植物各部位中将或多或少是等同的。该启动子已经成功地用于改造除草剂和病原体抗性植物(综述见 Stitt 和 Sonnewald ， 1995. 植物生理学和植物分子生物学年鉴(Annu. Rev. Plant Physiol. Plant Mol. Biol.)46 ： 341-368)。器官特异性启动子已经被报道用于储存性沉降组织(sink tissue)，例如种子、马铃薯块茎和果实(Edwards 和 Coruzzi ， 1990. 遗传学年鉴(Annu. Rev. Genet.)24 ： 275-303)，以及用于代谢性沉降组织，例如分生组织(Ito 等， 1994. 植物分子生物学(Plant Mol. Biol.)24 ： 863-878)。

用于培养转化宿主细胞的培养基可以是适于所选宿主细胞生长的任何常规培养基。所表达的植酸酶可以方便地分泌到培养基中，并且可以通过众所周知的方法从中回收，包括通过离心或过滤将细胞同培养基分离，用盐，如硫酸铵沉淀培养基的蛋白成分，随后进行层析，如离子交换层析、亲和层析等。

优选的宿主细胞是镰孢属、汉逊氏酵母属、木霉属或曲霉属的菌株，特别是禾谷镰孢、毒性镰孢(Fusarium venenatum)、谷类镰孢(Fusarium cerealis)、具有镰孢 ATCC20334 株的鉴别特征的进一步描述于 PCT/US/95/07743 中的各种镰孢菌种、多形汉逊氏酵母、 harzianum 木霉(Trichoderma harzianum)或 reesei 木霉(Trichoderma reesei)、黑曲霉或米曲霉中的一株。

在多形汉逊氏酵母中表达的参考： Gellissen ， G.、 Piontek ， M.、 Dahlems ， U.、 Jenzelewski ， V.、 Gavagan ， J.E.、 DiCosimo ， R.、 Anton ， D.I.和 Janowicz ， Z.A.(1996)。重组的多形汉逊氏酵母作为生物催化剂：菠菜乙醇酸酯氧化酶(GO)以及酿酒酵母催化酶 T(CTT1)基因的共表达。应用微生物学生物技术(Appl. Microbiol. Biotechnol.)46 ， 46-54 。

根据本发明植酸酶变体的一些更特定的应用登载于 PCT/DK97/00568 ，发明详述部分的最后数页。

在一优选的实施方案中，本发明的植酸酶变体基本上不含其他非植酸酶多肽，例如，通过 SDS-PAGE 确定的至少大约 20%的纯度、优选至少大约 40%的纯度、更优选大约 60%的纯度、还要更优选大约 80%的纯度、最优选

大约90%的纯度而且更加更优选大约95%的纯度。有时此类多肽或者指一种“纯化的”和/或“分离的”植酸酶。

含有本发明植酸酶变体的植酸酶多肽包括融合的多肽或者可切割的融合多肽，其中另一多肽融合于该多肽或其片段的N-末端或C-末端。融合的多肽是通过将编码另一多肽的核酸序列(或其一部分)融合至编码本发明一种植酸酶变体的核酸序列(或其一部分)来制备的。用于制备融合多肽的技术在本领域是已知的，而且包括连接编码多肽的编码序列，以便它们在同一阅读框架中而且融合多肽的表达处于相同启动子和终止子的控制下。

“饲料”和“食物”分别指预期或适合分别被动物或人食用、摄取、消化的任何天然的或人工的饮食、膳食或诸如此类或者此类膳食的组分。

本发明的植酸酶变体可以在体外或体内发挥其效应，即分别在摄入前或在机体的胃中。也可能是联合作用。

根据本发明的植酸酶组合物通常包括至少一种本发明的植酸酶。

通常，植酸酶组合物是液态的或干燥的。

液态组合物不需要含有除植酸酶外的任何组分，优选处于一种高度纯化的形式。然而，通常也加入一种稳定剂如甘油、山梨醇或单丙二醇。液态组合物还可以包括其他添加剂，例如盐、糖、防腐剂、pH调节剂、蛋白质、植酸盐(植酸酶底物)。典型的液态组合物是基于水或油的粘合液。液态组合物在其选择性沉淀后可以加至食物或饲料中。

干燥的组合物可以是喷雾干燥的组合物，在这种情况下组合物不需要含有除干燥形式的酶以外的任何组分。然而，通常干燥的组分就是所谓的颗粒可能易于同例如食物或饲料组分混合，或者更优选形成预混合料的组分。酶颗粒的颗粒大小优选同混合物中其他组分的大小一致。这提供了将酶掺进例如动物饲料中的一种安全而且方便的方式。

在一高剪切力的搅拌机(例如Lodige)中利用团聚技术制备团聚颗粒，在此过程中一种填充剂材料同酶团聚成颗粒。通过使载体材料的核心吸收酶/被酶包被来制备吸收颗粒。

典型的填充剂材料是盐，例如硫酸钠。其他填充剂为陶土、滑石、硅酸镁铝以及纤维素纤维。选择性地，粘合剂，例如糊精也包括在团聚颗粒中。

典型的载体材料是淀粉，例如木薯、玉米、马铃薯、大米和小麦。盐也可以使用。

选择性地，颗粒由包被混合物来包裹。此类混合物包括包被剂，优选疏水包被剂，例如氢化棕榈油和牛油，以及必要时加其他添加剂，例如碳酸钙或陶土。

另外，植酸酶组合物可以含有其他替代物，例如生色剂、芳香化合物、稳定剂、维生素、矿物质、其他饲料或食物强化酶，即提高饲料/食物的营养性质的酶，等。特别是对于所谓的预混合料就是如此。

“食物或饲料添加剂”是用来或适于添加到食物或饲料中的基本上纯的化合物或多组分组合物。特别是其是一种物质，其预期的用途是成为一种食物或饲料产品的组分或影响一种食物或饲料产品的任何特征。其由如上面对植酸酶组合物所说明的所组成。典型的添加剂通常包括一种或多种化合物，例如维生素、矿物质或饲料强化酶以及适当的载体和/或赋形剂。

在一优选的实施方案中，本发明的植酸酶组合物额外包括有效数量的一种或多种饲料强化酶，特别是选自由α-半乳糖苷酶、β-半乳糖苷酶，特别是乳糖酶，其他植酸酶、β-葡聚糖酶，特别是内-β-1,4-葡聚糖酶和内-β-1,3(4)葡聚糖酶，纤维素酶、木糖苷酶、半乳聚糖酶，特别是阿拉伯半乳聚糖内-1,4-β-半乳糖苷酶和阿拉伯半乳聚糖内-1,3-β-半乳糖苷酶，内葡聚糖酶，特别是内-1,2-β-葡聚糖酶、内-1,3-α-葡聚糖酶和内-1,3-β-葡聚糖酶，果胶降解酶，特别是果胶酶、果胶酯酶、果胶裂解酶、多聚半乳糖醛酸酶、阿拉伯聚糖酶、鼠李糖半乳糖醛酸酶、鼠李糖半乳糖醛酸乙酰酯酶、鼠李糖半乳糖醛酸-α-鼠李糖苷酶、果胶酸盐裂解酶和α-半乳糖醛酸苷酶，甘露聚糖酶、β-甘露糖苷酶、甘露聚糖乙酰酯酶、木聚糖乙酰酯酶、蛋白酶、木聚糖酶、阿拉伯糖基木聚糖酶以及脂解酶例如脂肪酶、磷脂酶和角质酶所组成组合的饲料强化酶。

本发明的动物饲料添加剂是在进食前或进食同时添加给单胃动物。本发明的动物饲料添加剂优选在进食的同时添加给单胃动物。在一更优选的实施方案中，动物饲料添加剂以颗粒或稳定液体的形式添加到膳食中。

在食物或饲料中植酸酶的有效数量是从大约 10 到 20,000；优选从大约 10 到 15,000，更优选从大约 10 到 10,000，特别是从大约 100 到 5,000，尤其是从大约 100 到大约 2,000 FYT/kg 饲料或食物。

本发明植酸酶的其他特定用途的实例是在大豆加工或在肌醇或其衍生物的生产中。

本发明还涉及用于降低动物粪尿中植酸盐水平的方法，其中给动物喂养含有有效数量的本发明植酸酶的饲料。

本发明还包括在食物或饲料制品或添加剂的制备过程中对本发明植酸酶的应用，即植酸酶仅在生产过程中发挥其植酸酶活性而在最终食物或饲料产品中没有活性。该方面与例如调面和烘烤相关。

本发明涉及一种植酸酶变体，当利用同 P_lycii 相应的位点编号按照图 1 对齐排列时，同一模式植酸酶相比，其在至少下述位点之一处是改变的：
24; 27; 31; 33; 39; 40; 41; 42; 43; 44; 45; 46; 47; 49; 51; 56; 58; 59; 61; 62; 68; 69; 70; 71; 72; 73; 74; 75; 76; 77; 78; 79; 80; 81; 82; 83; 84; 88; 90; 102; 115; 116; 117; 118; 119; 120; 121; 122; 123; 124; 125; 126; 127; 128; 132; 143; 148; 149; 150; 151; 152; 153; 154; 155; 156; 157; 158; 159; 160; 161; 162; 163; 170f; 170g; 171; 172; 173; 184; 185; 186; 187; 187a; 190; 191; 192; 193; 194; 195; 198; 199; 200; 201; 201a; 201b; 201c; 201d; 201e; 201f; 202; 203; 203a; 204; 205; 211; 215; 220; 223; 228; 232; 233; 234; 235; 236; 237; 238; 239; 242; 243; 244; 246; 251e; 253; 256; 260; 264; 265; 267; 270; 271; 272; 273; 274; 275; 276; 277; 278; 279; 280; 283; 285; 287; 288; 292; 293; 302; 304; 332; 333; 334; 335; 336; 337; 338; 339; 340; 341; 342; 343; 348; 349; 352; 360; 362; 364; 365; 366; 367; 368; 369; 370; 371; 372; 373; 374; 375; 376; 383k; 387; 393; 394; 396; 404; 409; 411; 412; 413; 417; 421; 431.

从这些变体，我们预期改变的特征，优选改变的活性特征。事实上，几种变体已经表现出此类改变的特征(见实验部分)。同上面相似，“改变”意指同模式植酸酶相比。“改变的活性特征”意指在至少一个植酸酶活性相关方面被改变，例如(非排他性列举)：pH 稳定性、温度稳定性、pH 适应性、温度适应性、特异性活性(特别是同 pH 和温度相关的)、底物特异性、底物切割模式、底物结合、位点特异性、磷酸盐从玉米释放的速度和水平、反应率、植酸盐降解率、释放的磷酸盐所达到的最终水平。

优选的改变活性特征是改变的特异性活性，优选提高的活性，而且优选在 pH 3、4、5 或 6 下提高的活性；改变的 pH 或温度适应性；和/或改变

的，优选提高的热稳定性，例如如 DSC 所测定的提高的熔解温度。

优选的植酸酶变体是：当利用同 P_lycii 相应的位点编号根据图 1 对齐排列时，同一模式植酸酶相比，其在至少下述位点之一处是改变的植酸酶变体：

43; 44; 47; 51; 58; 62; 78; 80; 83; 88; 90; 102; 143; 148; 153; 154; 186; 187a; 195; 198; 201e; 204; 205; 211; 215; 220; 242; 244; 251e; 260; 264; 265; 267; 270; 273; 278; 302; 336; 337; 339; 352; 365; 373; 383k; 404; 417.

下述烟曲霉的变体组成一个亚群：

Q43L; Q270L; G273D,K; N336S; A205E; Y278H; Q43L+Q270L; Q43L+Q270L+G273D; Q43L+Q270L+G273D+N336S; G273K+A205E; G273K+A205E+Y278H (see EP 0897010).

通常，如下可以推断或鉴定出本发明的变体：观察根据图 1 的对齐排列，比较两个序列，其中一个为一种有改变性质的模式植酸酶，确定在相关位点/区域的氨基酸差异并将在相关位点处的氨基酸从模式植酸酶转移到(置换)另一植酸酶序列。

本发明还涉及制备植酸酶变体的方法，其包括上面的方法，而且进一步包括相应 DNA 序列的推断和合成、宿主细胞的转化、宿主细胞的培养以及植酸酶变体的回收。

相关的位点/区域包括下面提到的同重要的植酸酶活性特征，如特异性活性、热稳定性、pH 活性/稳定性相关的那些。

本发明还涉及植酸酶变体(本文所定义的各种模式植酸酶)，其是通过上面概述的步骤可获得的，优选获得的，并且其预期表现改变的特征/性质，优选表现例如改善的特异性活性等此类改变的特征。

至少担子菌纲模式植酸酶 P_lycii 和绒毛栓菌表现出一种高度特异性的活性(如本文实施例 2 的方法所确定的)。

这是预期性质的一个实例，通过那种例如上面提到的推断步骤，可以将其转移至其他植酸酶，例如在图 1 中所列举的其他植酸酶，特别是 A_pediades 和子囊菌纲植酸酶例如烟曲霉、无花果曲霉、共有植酸酶 A 。

因此，变体预期具有改变的特异性活性，特别是改善的特异性活性，特别是在低 pH 和/或高温下，其中变体已经在相关区域被改变，即(i)在指向活

性位点裂口的氨基酸残基；或者(ii)在紧邻这些活性位点残基处的氨基酸残基。优选，紧邻意指在距离活性位点残基 10Å以内。

由 pdb 文件 1IHP(布鲁克海文数据库(Brookhaven Database)条目 18.03.98 re 1IHP，磷酸单酯酶的结构， D. Kostrewa ；或者如在自然结构生物学(Nature Structural Biology)4， 1997， 185-190 页中所发表的)可以鉴定活性位点区域，利用来自分子模拟 MSI，圣迭哥，加利弗尼亚的程序 INSIGHTII 并利用子命令，可以确定一个“活性位点外框”，其包括那些位置靠近催化残基的氨基酸残基，在无花果曲霉植酸酶中确定为 H59 、 D339 和 R58(分别同隔孢伏革菌编号 H71 、 D335 和 R70 相对应)。一个“活性位点外框(10Å)”包括那些位于距离上述催化残基 10Å以内的残基。

距离 H71 和 D335 在 10Å以内的残基如下(利用隔孢伏革菌编号)： 41-47 、 68-77 、 115-118 、 120-126 、 128 、 149-163 、 185 、 191-193 、 199 、 243 、 270-271 、 273-275 、 277-279 、 288 、 332-343 、 364-367 、 369-375 、 394(“活性位点外框(10Å)”)。

优选，还可以确定一个“底物结合外框”，其包括那些紧邻底物结合位点因而预期可能同底物接触的残基。

如上面所描述可以通过在活性位点裂口(形成活性位点的表面的残基)中接入一种植酸钙镁的糖类似物来推断这些信息。如果一种没有任何磷酸基团的糖例如，α-D-葡萄糖(椅式构象，由 INSIGHTII 程序提供的结构)接入活性位点裂口中，利用如下所示的一个固定距离，利用子命令以及距底物类似物 10Å的距离可以得出指向活性位点裂口的残基。或者，化合物肌醇-1,4,5-三磷酸酯(布鲁克海文数据库文件 1djx ，肌醇-1,4,5-三磷酸酯)可以接入活性位点裂口中。然而，该化合物和葡萄糖或多或少是可叠加的。

以埃(Å)计的距离为：从α-D-葡萄糖 6 号位处的氧原子到

H59 的原子 ND1 ： 5.84

R58 的原子 NH2 ： 6.77

R142 的原子 NH2 ： 5.09

N340 的原子 ND2 ： 3.00

H59 的原子 ND1 ： 7.76

R58 的原子 NH2 ： 8.58 。

(上述残基的隔孢伏革菌编号分别为： H71 、 R70 、 R155 、 N336 、

H71 和 R70 。)

这样，将与底物接触的残基鉴定如下(隔孢伏革菌编号)： 43-44 ； 70-80 ； 83-84 ； 115 ； 153 ； 155-156 ； 184 ； 191-192 ； 198-202 ； 205 ； 235 ； 238 ； 242 ； 270 ； 272-273 ； 275-277 ； 332-336 ； 338 ； 369 ； 371(“底物结合外框(10Å)”)。

在一个或多个(1)活性位点外框或(2)底物结合外框处被改变的变体被强烈预期具有改变的特异性活性。这引起下面位点的联合分组(仍为隔孢伏革菌编号和 10Å的外框)：

41-47, 68-80, 83-84, 115-118, 120-126, 128, 149-163, 184-185, 191-193, 198-201e, 202-203, 205, 235-236, 238-239, 242-243, 270-279, 285, 288, 332-343, 364-367, 369-375, 394.

优选，如上所述利用图 1 的担子菌纲模式植酸酶(隔孢伏革菌植酸酶为一种优选的模式)来确定活性位点外框和底物结合外框。利用图 1 的对齐排列可推断其他模式植酸酶的相应变体。

在一优选的实施方案中，在子命令中利用了一个 5Å的距离，因此确定具有一更有限大小的活性位点和底物结合外框，例如包括残基 43-44 、 69-74 、 117 、 125 、 155-156 、 159 、 274 、 332-340 、 370-374(距离 H71 和 D335 5Å)的一个活性位点外框，“活性位点外框(5Å)”。

通常活性位点外框和底物结合外框区域构成用于选择随机诱变区域的基础。优选的随机诱变区域的实例为

区域 69-74 、 332-340 、 370-374 ，待加入添加剂(一个 5Å的距离)；以及

区域 57-62 、 142-146 、 337-343 ，待加入添加剂(一个 10Å的距离)。

目前预计在这些位点的任意一个中的任何改变将产生一种具有改变的特征，例如改变的特异性活性的植酸酶。

上面的描述“在这些位点的任意一个中的任何改变”被认为完全等价于列举的每个位点以及每个置换，例如对于上面的亚群 41-47 如下：

41A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y;

42A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y;

43A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y;

44A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y;

45A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y;

46A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y;

47A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y.

在一优选的实施方案中，改变的特异性活性预期来自下面的变体：

42S,G; 43A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y; 45D,S; 47Y,F; 51E,A; 75W,F; 78S,D; 79G; 80K,A; 83I,Q; 84Q,V; 116S; 118V,L; 119E; 120L; 122A; 123N,T; 125S; 126H,S; 127Q,E; 128A,T; 151A,S; 152G; 153D,Y; 154Q,D,G; 157V; 158D,A; 159T; 160A,S; 161T,N; 162N; 163W; 184Q,S; 186A,E; 198A,N; 200G,V; 201D;

201a、201b、201c、201d、201e、201f中一个或多个——优选全部的缺失；

202S; 205Q,E; 235Y,L; 238L,M; 242P; 270Y,A,L; 271D; 273D,K; 275F,Y; 278T,H; 332F; 336S; 337T,Q; 339V; 340P,A; 343A,S; 364W,F; 365V,L; 366D,V; 367K; 368K; 369I,L; 370V; 373S; 374A; 375H; 376M; 393V.

特别优选的变体如下：78S；79G；80A；83I,Q；84Q,V；198A,N；200G,V；201D；201a、201b、201c、201d、201e、201f中一个或多个——优选全部的缺失；202S；205Q,E；235Y,L；238L,M；242P；273D；275F,Y。

其他特别优选的变体如下：43A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y ;特别是43M,P；75W,F；80K；153D；184Q,S；270Y,A；332F；369I,L。

下面的变体是非常优选的：43L,G,N,V,A,I,T；78D；153Y；154G；270L；273D,K。双重或三重变体(43L/270L)；(43L/270L/273D)；(43L/78D)以及(43L/153Y/154G)也是非常优选的。其他优选的变体是205E；278H；336S。

这些非常优选的单重、双重和三重变体优选为可以在图1对齐排列的模式植酸酶的变体，特别是在图1中所列举的特定模式植酸酶的变体。

至少已知共有植酸酶A具有一高的热稳定性。进一步P_lycii的热稳定

性也是相当高的。

这是一个能通过诸如上述的推断步骤转移到其他植酸酶，例如在图 1 中所列举的其他植酸酶，特别是担子菌纲植酸酶，例如 P_lycii 和 A_pediades 中的所需特性的实例。

根据这个背景，下述变体预期具有改变的热稳定性，特别是提高的热稳定性：

39H,S; 40L,N; 43P; 47Y,F; 49P; 51E,A; 56P; 58D; 61R; 62V; 80K; 83A; 84Y; 172P; 184P; 195T; 198A; 204V; 211L; 223D; 236Y; 242P; 246V; 253P; 264R; 265Q; 280A,P; 283P; 287A; 292F,Y; 293A; 302R; 304P; 337S; 348Y; 387P; 396R; 409R; 411K; 412R; 417E; 421F,Y.

特别优选下述具有改变的热稳定性的变体： 39S ； 40N ； 47Y,F ； 51A ； 83A ； 195T ； 204V ； 211L ； 242P ； 265A 。

具有改变的热稳定性的变体还有如下：

42G; 43T,L,G; 44N; 58K,A; 59G; 62I; 69Q; 75F; 78D; 79G; 80A; 81A,G; 82T; 83K,R; 84I; 88I; 90R,A; 102Y; 115N; 118V; 122A; 123Q,N; 125M,S; 126V,S; 127N,Q; 128S,A; 143N,K; 148V,I; 154S; 158D; 170fH; 170gA; 171T,N; 172N; 173W, 184S; 186A; 187A; 187aS; 193S; 195V,L; 198V; 201E; 201eT; 202A, 203aT; 204A; 211V, 215P,A; 220L,N; 223H; 228N; 232T; 322E; 235T; 236N; 242S; 244D; 251eQ,E; 256D; 264I; 260A,H; 265A; 267D; 270G; 271D; 273K,D; 278T,H; 287T; 293V; 302H; 337T,G; 338I; 339V,I; 340A; 352K; 365A,S; 366S; 367A; 369L; 373S,A; 374S; 376M; 383kE,Q; 404G,A; 411T; 417R; 431E.

预期能使植酸酶的热稳定性提高的本发明的其它构想如下－参照前文引用的 1IHP 结构并通过本文所概述的按图 1 之对齐排列进行的转移：

(A)在间隙位点引入脯氨酸，这些位点是能充满脯氨酸特异性双面夹角，和不阻碍氢原子结合网络，而且未发现空间冲突的位点。

(B)填补孔洞：通过置换内部空腔中较大的残基通常可以使稳定性提高。

(C)半胱氨酸桥：半胱氨酸桥通常将使蛋白质具有更大的刚性，而且增

加伸展能。

根据(A)至(C)的这些构想预期具有改变的热稳定性变体进一步还有：27P、31Y、132F、132I、132L、184P、186P、190P、280P、343F、343I、343L、349P、362P以及(33C和24C)。

构想(A)：27P、184P、186P、190P、349P、362P。

构想(B)：343F,I,L、31Y、132F,I,L、273F。

构想(C)：33C/24C。

另一个预期的性质是pH活性或稳定性的改变，优选稳定性的改变，特别是在低pH下的改变，而且特别是提高，其可以通过按图1进行对齐排列以及从pH适应性提高的模式植酸酶传递至其他植酸酶来被传递－如上述。

预期能使植酸酶在低pH时的稳定性提高的本发明其它构想如下－参照前文引用的1IHP结构并通过本文所概述的按图1之对齐排列进行的转移：

(D)表面电荷：在低pH下的更好的分布来避免负或正电荷的聚集，并避免带有相同电荷的残基过于靠近。

(E)脱酰氨基作用：暴露Q或N的表面同带有负电荷的残基紧密接触。

在低pH下pH稳定性/活性提高的植酸酶变体预期为：39H；39Q；80A；203R；271N；51R；154S；185S；194S；194T；288L；288I；288F；360R；173Q,S；204Q,S；303K,S；81Q,E。

构想(D)：203R、271N、51R、185S；360R；173Q,S；204Q,S；303K,S；81Q,E。

构想(E)：154S；194S,T；288L,I,F。

对于(D)和(E)这些构想的优选的模式植酸酶为P_lycii。

实验证实最佳pH值降低的是：子囊菌纲植酸酶变体80A，特别是烟曲霉的植酸酶和共有植酸酶A。

非常优选的单重、双重和三重变体为43L；(43L/270L)和(43L/270L/273D)。这些变体具有一改变的pH适应性。它们优选为在图1中所列举的特定模式植酸酶的变体。

对于上面所列举的所有优选变体：

稳定性优选在高温下被改变，即在50-100℃，特别是60-90℃的温度范围内，更优选在70-90℃的范围内。

活性优选在适用于动物胃肠道系统的相应温度范围内被改变，例如30-

40 ℃，更优选 32-38 ℃，最优选在 35-38 ℃的范围内。

稳定性优选在低的 pH 下被改变，即在 pH 1.5-7，优选 2-6，更优选 3-5 的范围内。

活性优选在 pH 1.5-5.5，更优选在 pH 2.5-4.5，还更优选 3-5 的范围内被改变。

对于如上提到的改变的植酸酶特征的测试法是本领域众所周知的，而且任何此类测试都可以用来比较植酸酶变体和模式植酸酶的性能。

在实施例 2 中给出了特异性活性的优选测试法。在实施例 3 中给出了 pH 和温度活性以及稳定性的优选测试法。热稳定性的更优选测试法为实施例 4 的 DSC 方法。

WO 98/28409 也公开了各种其他参数如位点特异性的测试。WO 98/28409 的所有测试法都为优选的测试法。

当然总体上所有这些测试法都可以在预期的 pH 值和温度下进行。

在附属的权利要求中，详细说明了基于本文明确公开的十三种模式植酸酶中五种的一些优选的植酸酶变体。

用类似的方式可以很容易推断出基于剩余八种明确公开的模式植酸酶的其他优选变体，方法是对图 1 中每一相应序列作上述改变。这些优选的变体明确地包括在本发明中，而且它们很容易推断，即如下：

衍生自 Paxillus，优选 Paxillus involutus，其中又优选 CBS 100231 株的模式植酸酶变体，优选 P_involtus-Al 的植酸酶变体，其序列示于图 2 中，上述的变体包括至少下列改变之一：

()24C; T27P; F31Y; I33C; R39H,S,Q; N40L; S42G; P43A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y; Y44N; S45D; Y47F; A51E,R; A58D;K; Q61R; I62V; F75W; S78D; A80K; T81Q,E,G,A; R83A,I,Q,K; I84Y,Q,V; L88I; K90R,A; F102Y; S115N; D116S; V118L; P119E; F120L; A123N,T,Q; S125M; F126H,S,V; D127Q,E,N; A128T,S; A132F,I,L; I148V; D151A,S; S153D,Y; D154Q,S,G; D158A; S159T; A160S; T161N; ()170fH; ()170gA; S171N; H172P; N173Q,S; P184Q,S; Q185S; T186A,E,P; G187A; ()187aS; T190P,A; D193S; N194S,T; M195T,V,L; A198N,V; G200V; D201E; ()201eT; S202A; D203R,K,S; P203aV,T; Q204E,S,A,V; V205E; V211L; S215A,I; L220N; A223D,H; D233E; F235Y,L,T; N236Y; L237F; V238L,M; A242P,S; M244D;

()251eE,Q; D253P; T256D; P260A,H; E264R,I; A265Q; A267D; G270Y,A,L; D271N; D273K; F275Y; T278H; Y280A,P; E283P; V287A,T; Q288L,I,F; Y292F; V293A; N302R,H; A304P; N336S; L337T,Q,S,G; M338I; V339I; A340P; S343A,F,I,L; F348Y; R349P; A352K; P360R; R362P; W364F; R365V,L,A,S; T366D,V,S; S367K,A; S368K; L369I; S373A; G374A,S; R375H; ()383kQ,E; T387P; Q396R; G404A; L409R; T411K; L412R; E417R; F421Y.

衍生自 Paxillus 属，优选 Paxillus involutus 种，其中又优选 CBS 100231 株的模式植酸酶变体，优选 P_involtus-A2 的植酸酶变体，其序列示于图 3 中，上述的变体包括至少下列改变之一：

P24C; I27P; F31Y; I33C; R39H,S,Q; N40L; S42G; P43A,C,D,E,F,G,H,I,K,L,M,N,Q,R,S,T,V,W,Y; Y44N; S45D; Y47F; A51E,R; A58D,K; E61R; I62V; F75W; S78D; A80K; A81Q,E,G; R83A,I,Q,R,K; I84Y,Q,V; L88I; K90R,A; F102Y; S115N; D116S; V118L; P119E; F120L; A123N,T,Q; S125M; F126H,S,V; D127Q,E,N; A128T,S; V132F,I,L; D143N; I148V; D151A,S; S153D,Y; D154Q,S,G; D158A; A160S; T161N; ()170fH; ()170gA; S171N; R172P; N173Q,S; P184Q,S; Q185S; T186A,E,P; G187A; ()187aS; T190P,A; D193S; N194S,T; M195T,V,L; A198N,V; G200V; E201D; ()201eT; S202A; D203R,K,S; P203aV,T; Q204E,S,A,V; V205E; S211L,V; S215A,P; L220N; A223D,H; A232T; F235Y,L,T; N236Y; L237F; V238L,M; P242S; M244D; ()251eE,Q; D253P; T256D; P260A,H; E264R,I; A265Q; A267D; G270Y,A,L; D271N; D273K; F275Y; T278H; Y280A,P; A283P; V287A,T; Q288L,I,F; Y292F; I293A,V; N302R,H; A304P; N336S; L337T,Q,S,G; M338I; V339I; 340P,A; A343S,F,I,L; F348Y; R349P; A352K; P360R; R362P; W364F; L365V,A,S; T366D,V,S; S367K,A; S368K; V369I,L; S373A; R375H; ()383kQ,E; T387P; Q396R; G404A; L409R; A411K,T; L412R; E417R; Y421F.

衍生自栓菌属的一个种，优选绒毛栓菌种，其中又优选 CBS 100232 株

的模式植酸酶变体，优选绒毛栓菌的植酸酶变体，其序列示于图 4 中，上述的变体包括至少下列改变之一：

R24C; T27P; L31Y; V33C; Q39H,S; S40L,N; S42G;
M43A,C,D,E,F,G,H,I,K,L,N,P,Q,R,S,T,V,W,Y; Y44N; S45D;
Y47F; A51E,R; A58D,K; S59G; Q61R; I62V; F75W; S78D; A80K;
A81Q,E,G; R83A,I,Q,K; I84Y,Q,V; V88I; K90R,A; L102Y; D115N;
V118L; T123N,Q; S125M; S126H,V; E127Q,N; A128T,S; A132F,I,L;
D143N; V148I; S151A; S153D,Y; D154Q,S,G; A158D; A160S; N161T;
()170fH; ()170gA; S171N; S172P; N173Q,S; S184Q,P; E185S;
A186E,P; G187A; ()187aS; T190P,A; N194S,T; M195T,V,L; A198N,V;
G200V; ()201eT; S202A; D203R,K,S; P203aV,T; Q204E,S,A,V; V205E;
Q211L,V; P215A; L220N; G223D,H; D233E; Y235L,T; N236Y; L237F;
L238M; P242S; E244D; ()251eE,Q; E253P; Q260A,H; D264R,I; A265Q;
A267D; A270Y,L,G; D271N; D273K; F275Y; T278H; Y280A,P; V287A,T;
Q288L,I,F; Y292F; I293A,V; A302R,H; N304P,A; N336S; Q337T,S,G;
M338I; V339I; A340P; S343A,F,I,L; F348Y; N349P; A352K; P360R;
R362P; F364W; L365V,A,S; V366D,S; K367A; I369L; A373S; A374S;
R375H; ()383kQ,E; Q387P; A396R; G404A; V409R; T411K; L412R;
E417R; Y421F.

衍生自曲霉属的一个种，优选构巢曲霉种，其中又优选 DSM 9743 株的模式植酸酶变体，优选构巢曲霉的植酸酶变体，其序列示于图 10 中，上述的变体包括至少下列改变之一：

V24C; A27P; H39S,Q; V40L,N; G42S;
Q43A,C,D,E,F,G,H,I,K,L,M,N,P,R,S,T,V,W,Y; Y44N; S45D;
Y47F; S49P; E51A,R; V56P; H58D,K,A; E61R; V62I; S69Q; Y75W,F;
E78D,S; S79G; K80A; S81Q,E,A,G; K82T; A83I,Q,K,R; Y84Q,V,I;
A90R; D115N; D116S; T118V,L; I119E; F120L; E122A; N123T,Q;
M125S; V126H,S; D127Q,E,N; S128A,T; F132I,L; K143N; I148V;
S151A; S153D,Y; D154Q,S,G; A158D; S159T; A160S; E161T,N; K162N;
F163W; G170fH; S170gA; ()171N; ()172P; K173Q,S; P184Q,S; E185S;
I186A,E,P; D187A; G187aS; T190P,A; H193S; S194T; S198A,N,V;

R201f()(缺失至少 201d 、 201e 、 201f 之一，优选全部)； A202S ；
D203R,K,S； E203aV,T； I204Q,E,S,A,V； I211L,V； P215A； L220N；
D223H； K228N； E232T； N233E； I235Y,L,T； Y236N； L237F； M238L；
S242P； M246V； E251eQ； A256D； E260A,H； L264R,I； Q270Y,A,L,G；
S271D,N； S273D,K； Y275F； G278T,H； A280P； A287T； Q288L,I,F；
F292Y； T293A,V； Q302R,H； P304A； N336S； S337T,Q,G； M338I； I339V；
S340P,A； F343A,S,I,L； N349P； Q352K； S360R； Q362P； Y364W,F；
A365V,L,S； A366D,V,S； S367K,A； W368K； T369I,L； G373S,A； A374S；
R375H； A376M； E383kQ； A404G； T411K； L412R； E417R； F421Y； K431E.

衍生自曲霉属的一个种，优选土曲霉种，其中又优选 CBS 220.95 株的模式植酸酶的变体，优选土曲霉的植酸酶变体，其序列示于图 12 中，上述的变体包括至少下列改变之一：

G24C； V27P； H39S,Q； K40L,N； G42S；
L43A,C,D,E,F,G,H,I,K,M,N,P,Q,R,S,T,V,W,Y； Y44N； A45D,S；
Y47F； S49P； Q51E,A,R； V56P； P58D,K,A； D59G； H61R； I62V； A69Q；
S75W,F； H78D,S； S79G； K80A； T81Q,E,A,G； A83I,Q,K,R； Y84Q,V,I；
A90R； E115N； E116S； T118V,L； P119E； F120L； R122A； N123T,Q；
L125S,H； R126H,S,V； D127Q,E,N； L128A,T,S； F132I,L； H143N； V148I；
T151A,S； D152G； A153D,Y； S154D,Q,G； H157V； E158D,A； S159T；
A160S； E161T,N； K162N； F163W； H173Q,S； P184Q,S； E185S；
G186A,E,P； S187A； A187aS； T190P,A； H193S； S194T； L195T,V；
A198N,V； E200G,V； S201D,E； S201d()； T201e()； V201f()； G202S,A；
D203R,K,S； D203aV,T； A204Q,E,S,V； V205E； V211L； A215P； L220N；
D223H； Q228N； D232T； D233E； V235Y,L,T； N236Y； L237F； M238L；
P242S； E244E； T251eE,Q； A260H； T264R,I； Q265A； N267D； L270Y,A,G；
S271D,N； K273D； Y275F； H278T； G280A,P； V287A,T； Q288L,I,F；
W292F,Y； A293V； Q302H； P304A； N337T,Q,S,G； L338I； V339I；
S340P,A； W343A,S,F,I,L； N349P； A352K； S360R； S362P； Y364W,F；
A365V,L,S； A366D,V,S； A367K； W368K； T369I,L； A373S； A374S；
R375H； A376M； R383kQ,E； P404A,G； K411T； A417E,R； F421Y； A431E.

衍生自踝节菌属的一个种，优选嗜热踝节菌种，其中又优选 ATCC 20186 或 ATCC 74338 株的模式植酸酶的变体，优选嗜热踝节菌的植酸酶变体，其序列示于图 13 中，上述的变体包括至少下列改变之一：

H24C; V27P; H39S,Q; S40L,N; G42S;
Q43A,C,D,E,F,G,H,I,K,L,M,N,P,R,S,T,V,W,Y; Y44N; S45D;
F47Y; S49P; A51E,R; V56P; Q58D,K,A; N59G; K61R; I62V; Y75W,F;
S78D; S79G; K80A; T81Q,E,A,G; E82T; L83A,I,Q,R,K; Y84Q,V,I;
R90A; D116S; T118V,L; P119E; F120L; E122A; N123T,Q; M125S;
I126H,S,V; Q127E,N; L128A,T,S; F132I,L; V148I; S151A; S153D,Y;
D154Q,S,G; I157V; A158D; S159T; G160A,S; R161T,N; L162N; F163W;
S170gA; D171N; K172P; H173Q,S; E184Q,S,P; E185S; G186A,E,P;
D187A; T190P,A; T193S; G194S,T; S195T,V,L; V198A,N; E200G,V;
D201E; S201d(); S201e(),T; S201f(); G202S,A; H203R,K,S;
D203aV,T; A204Q,E,S,V; Q205E; Q211L,V; A215P; I220N,L; H223D;
D228N; S232T; D233E; P235Y,L,T; Y236N; M237F; D238L,M; P242S;
E244D; L246V; ()251eE,Q; A256D; Q260A,H; Q264R,I; A265Q;
Q270Y,A,L,G; S271D,N; G273D,K; Y275F; N278T,H; G280A,P; A287T;
Q288L,I,F; F292Y; V293A; H302R; P304A; N336S; T337Q,S,G; M338I;
T339V,I; S340P,A; A343S,F,I,L; N349P; A352K; S360R; E362P;
Y364W,F; S365V,L,A; A366D,V,S; A367K; W368K; T369I,L; G373S,A;
G374A,S; R375H; A376M; D383kQ,E; E404A; K411T; R417E; F421Y.

衍生自 Thermomyces 属的一个种，优选 Thermomyces lanuginosus 种，其中又优选 DBS 586.94 株的模式植酸酶的变体，优选 T-lanuginosa 的植酸酶变体，其序列示于图 14 中，上述的变体包括至少下列改变之一：

K24C; ()27P; ()31Y; ()33C; R39H,S,Q; H40L,N; G42S;
Q43A,C,D,E,F,G,H,I,K,L,M,N,P,R,S,T,V,W,Y; Y44N; S45D;
F47Y; S49P; A51E,R; V56P; K58D,A; V62I; S69Q; Y75W,F; A78D,S;
H79G; K80A; S81Q,E,A,G; E82T; V83A,I,Q,K,R; Y84Q,V,I; L88I;
R90A; F102Y; D115N; N116S; T118V,L; R119E; F120L; E122A;
E123N,T,Q; M125S; M126H,S,V; E127Q,N; S128A,T; F132I,L; E143N;
V148I; A151S; S153D,Y; A154D,Q,S,G; I157V; A158D; S159T; A160S;
E161T,N; F162N; F163W; R170fH; S170gA; K172P; D173Q,S; S184Q,P;

E185S; E186A,P; T187A; G187aS; T190P,A; G193S; L194S,T; T195V,L; A198N,V; E200G,V; E201D; A201d(); P201e(),T; D202S,A; P203R,K,S; T203aV; Q204E,S,A,V; P205E; V211L; R215A,P; I220L,N; H223D; E232T; D233E; P235Y,L,T; L236Y,N; M238L; P242S; Q251eE; H256D; Q260H; M264R,I; A265Q; Y270A,L,G; T271D,N; D273K; Y275F; H278T; G280A,P; A283P; S287A; R288L,I,F; F292Y; V293A; G302R,H; P304A; N336S; T337Q,S,G; M338I; T339V,I; G340P,A; S343A,F,I,L; N349P; P360R; T362P; Y364W,F; A365V,L,S; A366D,V,S; S367K,A; W368K; T369I,L; A373S; A374S; R375H; A376M; E383kQ; R404A,G; R411K,T; K417E,R; F421Y; D431E.

衍生自毁丝菌属的一个种，优选嗜热毁丝菌种，其中又优选 ATCC 48102 或 ATCC 74340 株的模式植酸酶的变体，优选嗜热毁丝菌的植酸酶变体，其序列示于图 7 中，上述的变体包括至少下列改变之一：

S24C; F31Y; H39S,Q; F40L,N; G42S;

Q43A,C,D,E,F,G,H,I,K,L,M,N,P,R,S,T,V,W,Y; Y44N; S45D; Y47F; S49P; P51E,A,R; I56P; D58K,A; D59G; E61R; V62I; S69Q; A75W,F; L78D,S; K79G; R80K,A; A81Q,E,G; A82T; S83A,I,Q,K,R; Y84Q,V,I; R90A; D115N; E116S; T118V,L; R119E; T120L; Q122A; Q123N,T; M125S; V126H,S; N127Q,E; S128A,T; F132I,L; K143N; V148I; A151S; Q153D,Y; D154Q,S,G; H158D,A; S159T; A160S; E161T,N; G170fH; S170gA; T171N; F163W; V172P; R173Q,S; P184Q,S; E185S; T186A,E,P; G187aS; T190P,A; N193S; D194S,T; L195T,V; A198N,V; E200G,V; E201D; G201a(); P201b(); Y201c(); S201d(); T201e(); I201f(); G202S,A; D203R,K,S; D203aV,T; A204Q,E,S,V; Q205E; T211L,V; P215A; V220N,L; N223D,H; A232T; D233E; V235Y,L,T; A236Y,N; L237F; M238L; P242S; E244D; A251eE,Q; R256D; E260A,H; R264I; A265Q; Q270Y,A,L,G; S271D,N; K273D; Y275F; Y278T,H; P280A; T287A; Q288L,I,F; F292Y; V293A; ()302R,H; P304A; N336S; D337T,Q,S,G; M338I; M339V,I; G340P,A; G343A,S,F,I,L; D349P; P352K; D360R; E362P; Y364W,F; A365V,L,S; A366D,V,S;

S367K,A; W368K; A369I,L; A373S; A374S; R375H; I376M; E383kQ; E387P; G404A; M409R; T411K; L412R; E417R; F421Y; D431E.

本发明还提供了一种新型植酸酶，其衍生自 Cladorrhinum 的一个株，即 C. Foecundissimum 。相应地，本发明还涉及一种具有植酸酶活性的多肽，而且其包括 SEQ ID NO ： 2 或其成熟部分(16-495 位氨基酸)；或者一种与之至少 70，更优选 75 、 80 、 85 、 90 、 95%同源的多肽；同源性意指相似性，优选同一性，是通过利用程序 GAP 和本文上面所确定的设置来确定的。本发明涉及一种 DNA 构建体，其编码一种具有植酸酶活性的多肽，该 DNA 构建体包括一种 DNA 分子，其包括 SEQ ID NO ： 1 或其 20-70 和 207-1560 号核苷酸；或者其 20-70 和 207-1563 号核苷酸；或者其 65-70 和 207-1560 号核苷酸；或者其 65-70 和 207-1563 号核苷酸；或者一种 DNA 构建体或分子，其同这些核苷酸序列中任何一种至少 70 、 75 、 80 、 85 、 90 、 95%同源；同源性意指相似性，优选同一性，是通过利用本领域已知的计算机程序，例如在 GCG 程序包中提供的 GAP(威斯康星包的程序手册，版本 8， 1996 年 8 月，遗传学计算机组， 575 Science Drive ， Madison ， Wisconsin ，美国 53711)(Needleman, S.B.和 Wunsch, C.D.，(1970)，分子生物学杂志(Journal of Molecular Biology)， 48 ， 443 - 453)来确定的。利用 GAP 按照下列设置进行 DNA 序列比较： 5.0 的 GAP 形成补偿和 0.3 的 GAP 延伸补偿。本发明还涉及一种 DNA 构建体，其在本文上面提到的条件下同上述 DNA 序列中的任何一个杂交。

实施例

实施例 1

植酸酶活性检测(FYT)

利用下面的检测可以测定植酸酶活性：

将 10 μ l 稀释的酶样品(稀释于 0.1M 乙酸钠、 0.01%吐温 20 ， pH5.5)加入到在 0.1M 乙酸钠、 0.01%吐温 20 、 pH5.5 中的 5mM 植酸钠(Sigma)250 μ l(植酸钠溶解后调节 pH ；预热底物)中并于 37 ℃温育 30 分钟。通过加入 250 μ l 10%TCA 来终止反应，并通过加入在 100ml 钼酸盐试剂(将在 8ml H_2SO_4 中溶解的 2.5g $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$ 稀释至 250ml)中溶解的 7.3g

$FeSO_4$ 500 μl 来测定游离磷酸根。在一 96 孔微量滴定板中测定 200 μl 样品在 750nm 处的吸光度。同时测定的有底物和酶的空白对照。还包括一个磷酸根的标准曲线(0-2mM 磷酸根)。 1FYT 等于在给定的条件下每分钟释放 1 μ mol 磷酸根的酶的量。

实施例 2

对特异性活性的测试

按照下述方法可以确定特异性活性：

利用一种高度纯化的植酸酶样品(预先在一 SDS 聚丙烯酰胺凝胶上检测纯度，表现出仅有一种组分存在)。

按照下述方法通过氨基酸分析确定在植酸酶样品中的蛋白质浓度：在一抽真空的玻璃试管中于 110 ℃将一份植酸酶样品在 6N HCl 、 0.1%苯酚中水解 16 小时。利用 Applied Biosystems 420A 氨基酸分析系统按照制造商的说明进行操作对所获的氨基酸定量。根据氨基酸的量可以计算出水解样品中蛋白质的总质量-从而还有浓度。

以 FYT 为单位确定活性。一个 FYT 等于在 pH5.5 、 37 ℃下每分钟从植酸盐(5mM 植酸盐)释放 1 微摩尔无机磷酸根的酶的量；测试法描述于例如实施例 1 中。

特异性活性是 FYT/mg 酶蛋白的值。

实施例 3

对温度和 pH 活性及稳定性的测试

温度和 pH 活性及稳定性可以按照下述的方法确定：

温度适应性(即温度活性关系)通过在各种温度下进行实施例 1 的 FYT 测试(预热底物)而定。

温度稳定性通过在检测残存活性前在各种温度下于 0.1M 磷酸钠、pH5.5 中预温育植酸酶而定。

pH 稳定性通过在检测残存活性前在 40 ℃于 pH3(25mM 盐酸甘油)、pH4-5(25mM 乙酸钠)、 pH6(25mM MES)、 pH7-9(25mM Tris-HCl)下温育酶 1 小时而定。

pH 适应性(即 pH 活性关系)通过利用相同的缓冲液系统(50mM 、在溶解

底物时重新调节的 pH)在各种 pH 下进行测试。

实施例 4

DSC 作为对热稳定性的一个优选的测试

热稳定性或熔解温度， Tm ，可以按照如下的方法确定：

在 DSC 中，测定样品池相对于参考池保持恒速升温所消耗的热量。保持恒定的加热速率(例如 90 ℃/小时)。将吸热过程(热消耗过程-例如酶/蛋白质的展开)看作是为了保持恒速升温而传递至该池的热量的增加。

利用 MicroCal 的 MC2-仪可以进行 DSC 。在以 90 ℃/h 的扫描速率扫描至 90 ℃前将池在 20 ℃平衡 20 分钟。装载例如在 0.1M 乙酸钠， pH5.5 中的大约 2.5mg/ml 植酸酶样品。

实施例 5

具有改变的活性特征的植酸酶变体

按照在 EP 98104858.0(EP-A-0897010)实施例 2-3 和 5 中所述制备烟曲霉的模式植酸酶(衍生自 ATCC 13073 株的一种野生型植酸酶)的变体，而且按照其实施例 7 中所述测定植酸酶活性。 pH 和温度最佳值以及熔点按照在 EP 98113176.6(EP-A-0897985)实施例 9 和 10 中所描述的来测定。

在表 1 中，列举了在 pH5.0 下具有改善的特异性活性的变体。表 2 列举了在 pH3.0 下具有改善的相对活性的变体，而表 3 列举了具有改善的热稳定性的变体(通过例如 DSC 测定的最佳温度)。

表 1

改变的位点编号	置换为	在 pH5.0 下的特异性活性 (U/mg)
43	43L	83.4
	43N	45.5
	43T	106.9
	43I	91.2
	43V	35.0
	43A	27.3

	43G	59.6
43 和 270	43L 、 270L	88.7
43 和 270 以及 273	43L 、 270L 、 273D	92.3
43 和 78	43L 、 78D	118.5
43 和 153 以及 154	43L 、 153Y 、 154G	193.0
烟曲霉野生型植酸酶	-	26.5

表 2

改变的位点编号	置换为	在 pH3.0 下的相对植酸酶活性
205	205E	41%
273	273K	61%
278	278H	75%
273 和 205	273K 、 205E	65%
273 和 278	273K 、 278H	100%
273 和 205 以及 278	273K 、 205E 、 278H	96%
烟曲霉野生型植酸酶	-	32%

表 3

改变的位点编号	置换为	最佳温度(℃)	Tm(℃) (DSC)
43 、 47 、 88 、 102 、 220 、 242 以及 267	43T 、 47Y 、 88I 、 102Y 、 220L 、 242P 、 267D	60	67
同上再加 51 、 302 、 337 、 373 以及 115	同上再加 51A 、 302H 、 337T 、 373A 、 115N	63	-
烟曲霉野生型植酸酶	-	55	62.5

实施例 6

具有改变的活性特征的其它植酸酶变体

按照在 EP 98113176.6(EP-A-0897985)，实施例 4-8 中所描述的来制备图 9 中子囊菌纲共有序列“共有植酸酶”的变体。植酸酶活性，包括最佳 pH

和最佳温度以及熔点，分别按照在其实施例 9 和 10 中所描述的来测定。

下面的表列举出具有改变的活性特征的变体，即

表 4 在 pH6.0 下具有改善的特异性活性的变体；

表 5 具有改变的最佳 pH 的变体(所指的最佳 pH 是一个近似值，确定为在此 pH 下获得最大植酸酶活性的那个 pH 值(选自由 pH4.0 、 4.5 、 5.0 、 5.5 、 6.0 、 6.5 和 7.0 组成的一组))；

表 6 具有改善的热稳定性的变体(通过由差异扫描测热法(DSC)所测定的熔点来表示)；以及

表 7 具有改变的热稳定性(最佳温度)的变体；“+”或“-”分别表示对高达 1 ℃的最佳温度的正或负影响；而“++”和“--”分别指对在 1 和 3 ℃之间的最佳温度的正或负影响。

表 4

改变的位点编号	置换为	在 pH6.0 下的特异性活性 (U/mg)
43	43T	130
	43L	205
共有植酸酶	-	62

表 5

改变的位点编号	置换为	最佳 pH 大约
43	43T	6.0
	43L	5.5
	43G	6.5
43 和 44	43L 、 44N	6.0
	43T 、 44N	5.5
共有植酸酶	-	6.0

表 6

改变的位点编号	置换为	Tm(℃)
43	43T	78.9

共有植酸酶	-	78.1

表 7

改变的位点编号	置换为	最佳温度的改变
51	A	+
58	K	+
220	N	+
195	L	++
201e	T	++
244	D	+
264	I	+
302	H	+
337	T	++
352	K	+
373	A	++
47	F	-
62	I	-
83	K	-
90	R	-
143	N	-
148	V	--
186	A	--
187a	S	-
198	V	-
204	A	--
211	V	-
215	P	--
251e	Q	-
260	A	-
265	A	-

339	V	-
365	A	--
383k	E	-
404	G	--
417	R	--
共有植酸酶	-	0

表 8

改变的位点编号	置换为	Tm(℃) (DSC)	在 pH5.0 下的特异性活性(U/mg)
43、51、220、244、264、302、337、352 以及 373	51A、220N、244D、264I、302H、337T、352K、373A、43T	84.7	105
同上再加 80	同上再加 80A	85.7	180
共有植酸酶	-	78.1	30

实施例 7

Cladorrhinum foecundissimum 植酸酶的克隆

基本上按照在 WO 98/28409 中所描述的对编码 Cladorrhinum foecundissimum CBS 427.97 的植酸酶的 DNA 进行了克隆，并对该酶进行了分离和纯化。

图 15 显示在 pA2phy8 中 HindIII/XbaI 克隆的 PCR 产物的 DNA 序列。所克隆的 PCR 产物扩增自编码 Cladorrhinum foecundissimum CBS 427.97 中 phyA 基因的基因组区域。推测的内含子通过符合 GT-AG 构想的剪切-连接位点处的双下划线标明(R. Brethnach 等，美国国家科学院进展(Proc. Natl. Acad. Sci. USA)75(1978)4853-4857 页)。用于克隆的限制性位点标有下划线。

根据 SignalP V1.1 的预测(Henrik Nielsen、Jacob Engelbrecht、Stren Brunak 和 Gunnar von Heijne："原核和真核信号肽的鉴定和它们的切割位点的预测"，蛋白质工程(Protein Engineering)10， 1-6(1997))，该酶的信号肽部分对应于 1-15 位氨基酸，相应地成熟的酶是 16-495 位氨基酸。

该酶表现出在 pH6 左右有最佳 pH 而在低 pH(pH3)下没有活性，但较高活性保持至 pH7.5 ；因此，同无花果曲霉植酸酶相比其是一种更碱性的植酸酶。

在 pH5.5 下发现一个 60 ℃左右的最佳温度。因此，该植酸酶比无花果曲霉植酸酶热稳定性更好。

实施例 8

根据图 1 一种新型模式植酸酶的对齐排列

利用上面引用的 GAP 版本 8，按照 3.000 的空隙权和 0.100 的空隙长度权，将在实施例 7 中所公开的 Cladorrhinum foecundissimum 的植酸酶序列同图 1 中的 13 种模式植酸酶进行比较。比较了完整的氨基酸序列。嗜热毁丝霉植酸酶序列被证明是最同源的序列，表现出同 C. foecundissimum 序列 70.86%的相似度。

仍利用 GAP 程序以及上面提到的参数，现在将植酸酶序列“C_foecundissimum”同“嗜热毁丝霉”植酸酶进行对齐排列——见图 16。平均匹配为 0.540 ；平均错配为-0.396 ；质量(quanlity)445.2，长度 505，比率 0.914，空隙 9 个，相似百分比 70.860，同一性百分比 53.878 。

在下一个步骤中，见图 17，将 C_foecundissimum 粘贴(或者其可以简单地写)到图 1 的对齐排列上作为最底行，保证那些根据图 16 的对齐排列为相同(通过一个垂直线标明)或相似(通过一个或两个点标明)的氨基酸残基一一置于其上。在沿着序列的 5 个位置处 C_foecundissimum 序列包括“多余”的氨基酸残基，图 1 的对齐排列没有为其提供位置。在图 17 中，将这些多余的残基转移至下一行(但是它们能包括在多序列对齐排列中，而且按照前面在位点编号相关段落中所描述的进行编号(利用符号 a 、 b 、 c 等))。

然后，基于图 17 很容易地推断出 C_foecundissimum 植酸酶相应的变体。一些实例：在 C_foecundissimum 中通常表示为“ 80K,A ”和“ 43T ”的变体分别对应于“ K80A ”和“ Q43T ”。

序列表

<212> DNA
<213> Cladorrhinum foecundissimum

<220>
<221> 内含子
<222> (71)..(126)

<220>
<221> CDS
<222> (20)..(70)

<220>
<221> CDS
<222> (127)..(1563)

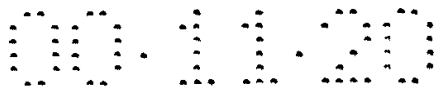
<220>
<221> 信号肽
<222> (20)..(64)

<400> 1

```
aagcttgggc aaactcatc atg ctc atc ttg atg att cca ctg ttc agc tac  52
                       Met Leu Ile Leu Met Ile Pro Leu Phe Ser Tyr
                         1               5                  10

ctg gct gct gct tct ctg tgggttcatc ctttgcccct gtctcgatgt          100
Leu Ala Ala Ala Ser Leu
             15

taaaatacta aacatatttc accaga cgt gta ctc tcc cct cag cca gtg tcc  153
                            Arg Val Leu Ser Pro Gln Pro Val Ser
                                20                  25

tgt gac agc ccg gag ctt ggt tac caa tgc gac cag cag aca acg cac   201
Cys Asp Ser Pro Glu Leu Gly Tyr Gln Cys Asp Gln Gln Thr Thr His
             30                  35                  40

acc tgg ggt caa tac tca ccc ttc ttc tct gtc ccg tca gag atc tcc   249
Thr Trp Gly Gln Tyr Ser Pro Phe Phe Ser Val Pro Ser Glu Ile Ser
         45                  50                  55

cct tcc gtt cct gat ggc tgc cgc ctc acc ttc gcc caa gtt ctc tcc   297
Pro Ser Val Pro Asp Gly Cys Arg Leu Thr Phe Ala Gln Val Leu Ser
     60                  65                  70

cgc cac ggc gcc cgc ttc cca acc ccg ggt aaa gcc gcc gcc atc tcc   345
Arg His Gly Ala Arg Phe Pro Thr Pro Gly Lys Ala Ala Ala Ile Ser
 75                  80                  85                  90

gct gtc ctc acc aaa atc aaa acc tct gcc acc tgg tac ggt tcc gac   393
Ala Val Leu Thr Lys Ile Lys Thr Ser Ala Thr Trp Tyr Gly Ser Asp
                 95                 100                 105

ttt cag ttc atc aag aac tac gac tat gta ctt ggc gta gac cac ctg   441
Phe Gln Phe Ile Lys Asn Tyr Asp Tyr Val Leu Gly Val Asp His Leu
             110                 115                 120
```

```
acc gcg ttc ggc gag caa gaa atg gtc aac tcc ggc atc aag ttc tac   489
Thr Ala Phe Gly Glu Gln Glu Met Val Asn Ser Gly Ile Lys Phe Tyr
        125                 130                 135

cag cgc tac tcc tcc ctc atc cag aca gaa gac tcg gat acg ctc ccc   537
Gln Arg Tyr Ser Ser Leu Ile Gln Thr Glu Asp Ser Asp Thr Leu Pro
    140                 145                 150

ttc gtc cgc gcc tct ggc cag gaa cgc gtc atc gcc tcc gcc gag aac   585
Phe Val Arg Ala Ser Gly Gln Glu Arg Val Ile Ala Ser Ala Glu Asn
155                 160                 165                 170

ttc acc acc ggc ttc tac tcg gcc ctc tca gcc gac aag aac cct cct   633
Phe Thr Thr Gly Phe Tyr Ser Ala Leu Ser Ala Asp Lys Asn Pro Pro
                175                 180                 185

tcc tcc tta cca aga cca gaa atg gtc atc att tct gag gag cca aca   681
Ser Ser Leu Pro Arg Pro Glu Met Val Ile Ile Ser Glu Glu Pro Thr
            190                 195                 200

gcc aac aac acc atg cac cac ggc ctc tgc cgc tcc ttt gaa gat tcc   729
Ala Asn Asn Thr Met His His Gly Leu Cys Arg Ser Phe Glu Asp Ser
        205                 210                 215

acc acc ggc gac caa gcc caa gcg gaa ttc atc gcc gcc acc ttc cca   777
Thr Thr Gly Asp Gln Ala Gln Ala Glu Phe Ile Ala Ala Thr Phe Pro
    220                 225                 230

ccc atc acc gcc cgt ctc aac gcc caa ggt ttc aaa ggc gtc acc ctc   825
Pro Ile Thr Ala Arg Leu Asn Ala Gln Gly Phe Lys Gly Val Thr Leu
235                 240                 245                 250

tcc aac acc gac gtc cta tca cta atg gac ctc tgc ccc ttt gac acc   873
Ser Asn Thr Asp Val Leu Ser Leu Met Asp Leu Cys Pro Phe Asp Thr
                255                 260                 265

gtc gcc tac ccc ctt tcc tcc ctc acc acc acc tct tcc gtt tct gga   921
Val Ala Tyr Pro Leu Ser Ser Leu Thr Thr Thr Ser Ser Val Ser Gly
            270                 275                 280

ggc ggc aag tta tcc ccc ttc tgc tct ctt ttc act gcc agc gac tgg   969
Gly Gly Lys Leu Ser Pro Phe Cys Ser Leu Phe Thr Ala Ser Asp Trp
        285                 290                 295

aca atc tac gat tac ctc cag tcc cta ggg aaa tac tac ggt ttc ggc   1017
Thr Ile Tyr Asp Tyr Leu Gln Ser Leu Gly Lys Tyr Tyr Gly Phe Gly
    300                 305                 310

ccc ggt aat tcc cta gct gcc acc cag ggg gta ggg tac gtc aac gag   1065
Pro Gly Asn Ser Leu Ala Ala Thr Gln Gly Val Gly Tyr Val Asn Glu
315                 320                 325                 330

ctt atc gcc cgc ttg atc cgt gct ccc gtc gta gat cac acg acg acc   1113
Leu Ile Ala Arg Leu Ile Arg Ala Pro Val Val Asp His Thr Thr Thr
                335                 340                 345
```

```
   aac tct act ctt gat ggc gac gaa aaa acg ttt ccg ttg aac aga acg   1161
   Asn Ser Thr Leu Asp Gly Asp Glu Lys Thr Phe Pro Leu Asn Arg Thr
                350                 355                 360
5
   gtg tat gcg gat ttt tcc cat gat aat gat atg atg aat atc ctg act   1209
   Val Tyr Ala Asp Phe Ser His Asp Asn Asp Met Met Asn Ile Leu Thr
           365                 370                 375

10 gct ttg cgg ata ttc gag cat atc agt ccg atg gat aac acc act atc   1257
   Ala Leu Arg Ile Phe Glu His Ile Ser Pro Met Asp Asn Thr Thr Ile
       380                 385                 390

   ccg acc aac tat ggc cag aca gga gat gac ggg gtg aag gaa agg gat   1305
15 Pro Thr Asn Tyr Gly Gln Thr Gly Asp Asp Gly Val Lys Glu Arg Asp
   395                 400                 405                 410

   ttg ttc aag gtt agt tgg gcg gtg ccc ttt gct ggg agg gtg tac ttt   1353
   Leu Phe Lys Val Ser Trp Ala Val Pro Phe Ala Gly Arg Val Tyr Phe
20                 415                 420                 425

   gag aaa atg gtt tgt gat gcg gat ggg gat ggc aag att gat agt gat   1401
   Glu Lys Met Val Cys Asp Ala Asp Gly Asp Gly Lys Ile Asp Ser Asp
               430                 435                 440
25
   gag gct cag aaa gag ttg gtg agg att ttg gtt aat gat cgg gtg atg   1449
   Glu Ala Gln Lys Glu Leu Val Arg Ile Leu Val Asn Asp Arg Val Met
           445                 450                 455

30 aga ttg aat ggg tgt gat gct gat gaa cag ggt agg tgt gga ttg gag   1497
   Arg Leu Asn Gly Cys Asp Ala Asp Glu Gln Gly Arg Cys Gly Leu Glu
       460                 465                 470

   aag ttt gtg gag agt atg gag ttt gcg agg aga ggg ggg gag tgg gag   1545
35 Lys Phe Val Glu Ser Met Glu Phe Ala Arg Arg Gly Gly Glu Trp Glu
   475                 480                 485                 490

   gag agg tgt ttt gtt tag ctctaga                                  1570
   Glu Arg Cys Phe Val
40                 495


   <210> 2
   <211> 495
45 <212> PRT
   <213> Cladorrhinum foecundissimum

   <400> 2
   Met Leu Ile Leu Met Ile Pro Leu Phe Ser Tyr Leu Ala Ala Ala Ser
50   1               5                  10                  15

   Leu Arg Val Leu Ser Pro Gln Pro Val Ser Cys Asp Ser Pro Glu Leu
               20                  25                  30

55 Gly Tyr Gln Cys Asp Gln Gln Thr Thr His Thr Trp Gly Gln Tyr Ser
           35                  40                  45
```



Pro Phe Phe Ser Val Pro Ser Glu Ile Ser Pro Ser Val Pro Asp Gly
50 55 60

Cys Arg Leu Thr Phe Ala Gln Val Leu Ser Arg His Gly Ala Arg Phe
65 70 75 80

Pro Thr Pro Gly Lys Ala Ala Ala Ile Ser Ala Val Leu Thr Lys Ile
85 90 95

Lys Thr Ser Ala Thr Trp Tyr Gly Ser Asp Phe Gln Phe Ile Lys Asn
100 105 110

Tyr Asp Tyr Val Leu Gly Val Asp His Leu Thr Ala Phe Gly Glu Gln
115 120 125

Glu Met Val Asn Ser Gly Ile Lys Phe Tyr Gln Arg Tyr Ser Ser Leu
130 135 140

Ile Gln Thr Glu Asp Ser Asp Thr Leu Pro Phe Val Arg Ala Ser Gly
145 150 155 160

Gln Glu Arg Val Ile Ala Ser Ala Glu Asn Phe Thr Thr Gly Phe Tyr
165 170 175

Ser Ala Leu Ser Ala Asp Lys Asn Pro Pro Ser Ser Leu Pro Arg Pro
180 185 190

Glu Met Val Ile Ile Ser Glu Glu Pro Thr Ala Asn Asn Thr Met His
195 200 205

His Gly Leu Cys Arg Ser Phe Glu Asp Ser Thr Thr Gly Asp Gln Ala
210 215 220

Gln Ala Glu Phe Ile Ala Ala Thr Phe Pro Pro Ile Thr Ala Arg Leu
225 230 235 240

Asn Ala Gln Gly Phe Lys Gly Val Thr Leu Ser Asn Thr Asp Val Leu
245 250 255

Ser Leu Met Asp Leu Cys Pro Phe Asp Thr Val Ala Tyr Pro Leu Ser
260 265 270

Ser Leu Thr Thr Thr Ser Ser Val Ser Gly Gly Gly Lys Leu Ser Pro
275 280 285

Phe Cys Ser Leu Phe Thr Ala Ser Asp Trp Thr Ile Tyr Asp Tyr Leu
290 295 300

Gln Ser Leu Gly Lys Tyr Tyr Gly Phe Gly Pro Gly Asn Ser Leu Ala
305 310 315 320

Ala Thr Gln Gly Val Gly Tyr Val Asn Glu Leu Ile Ala Arg Leu Ile
325 330 335

Arg Ala Pro Val Val Asp His Thr Thr Thr Asn Ser Thr Leu Asp Gly

00.11.20

```
                340                 345                 350

Asp Glu Lys Thr Phe Pro Leu Asn Arg Thr Val Tyr Ala Asp Phe Ser
        355                 360                 365

His Asp Asn Asp Met Met Asn Ile Leu Thr Ala Leu Arg Ile Phe Glu
    370                 375                 380

His Ile Ser Pro Met Asp Asn Thr Thr Ile Pro Thr Asn Tyr Gly Gln
385                 390                 395                 400

Thr Gly Asp Asp Gly Val Lys Glu Arg Asp Leu Phe Lys Val Ser Trp
                405                 410                 415

Ala Val Pro Phe Ala Gly Arg Val Tyr Phe Glu Lys Met Val Cys Asp
            420                 425                 430

Ala Asp Gly Asp Gly Lys Ile Asp Ser Asp Glu Ala Gln Lys Glu Leu
        435                 440                 445

Val Arg Ile Leu Val Asn Asp Arg Val Met Arg Leu Asn Gly Cys Asp
    450                 455                 460

Ala Asp Glu Gln Gly Arg Cys Gly Leu Glu Lys Phe Val Glu Ser Met
465                 470                 475                 480

Glu Phe Ala Arg Arg Gly Gly Glu Trp Glu Glu Arg Cys Phe Val
                485                 490                 495
```

说 明 书 附 图

隔孢伏革菌编号	1					37
对齐排列编号	1					50
P_involtus_A1	.....ML	FGFVALACLL	SLSEVLATSV	P.....KNT	APTFPIPESE
P_involtus_A2	.....MH	LGFVTLACLI	HLSEVFAASV	P.....RNI	APKFSIPESE
绒毛栓菌	.....	MAFSILASLL	FVCYAYARAV	PRAHIPLRDT	SACLDVTRDV
平田头菇	.....	MSLFIGGCLL	VFLQASAYGG	VVQATFVQPF	.....FPPQI
P_lycii	.....MV	SSAFAPSILL	SLMSSLALST	QFSF.....V	AAQLPIPAQN
烟曲霉	.....MVTL	TFLLSAAYLL	.SGRVSAAPS	SAGSKSCDTV	DLGYQCSPAT
共有植酸酶A	.....MGVF	VVLLSIATLF	GSTSGTALGP	RGNSHSCDTV	DGGYQCFPEI
构巢曲霉	.....MAFF	TVALSLYYLL	..SRVSAQAP	VVQNHSCNTA	DGGYQCFPNV
无花果曲霉NRRL3135	.....MGVS	AVLLPLYLLS	GVTSGLAVPA	SRNQSSCDTV	DQGYQCFSET
土曲霉	.....MGFL	AIVLSVALLF	RSTSGTPLGP	RGKHSDCNSV	DHGYQCFPEL
嗜热踝节菌	.....MSLL	LLVLSGGLVA	LYVS...RNP	HVDSHSCNTV	EGGYQCRPEI
T_lanuginosa	MAGIGLGSFL	VLLLQFSALL	TASPAIPPFW	RKKHPNVD..	I
嗜热毁丝霉	.....MTGL	GVMVVMVGFL	AIASL.....	QSESRPCDTP	DLGFQCGTAI
	38				83
	51				100
P_involtus_A1	QRNWSPYSPY	FPLAEYKA..	..PPAGCQIN	QVNIIQRHGA	RFPTSGATTR
P_involtus_A2	QRNWSPYSPY	FPLAEYKA..	..PPAGCEIN	QVNIIQRHGA	RFPTSGAATR
绒毛栓菌	QQSWSMYSPY	FPAATYVA..	..PPASCQIN	QVHIIQRHGA	RFPTSGAAKR
平田头菇	QDSWAAYTPY	YPVQAYTP..	..PPKDCKIT	QVNIIQRHGA	RFPTSGAGTR
P_lycii	TSNWGPYDPF	FPVEPYAA..	..PPEGCTVT	QVNLIQRHGA	RWPTSGARSR
烟曲霉	SHLWGQYSPF	FSLEDELSVS	SKLPKDCRIT	LVQVLSRHGA	RYPTSSKSKK
共有植酸酶A	SHLWGQYSPY	FSLEDESAIS	PDVPDDCRVT	FVQVLSRHGA	RYPTSSKSKA
构巢曲霉	SHVWGQYSPY	FSIEQESAIS	EDVPHGCEVT	FVQVLSRHGA	RYPTESKSKA
无花果曲霉NRRL3135	SHLWGQYAPF	FSLANESVIS	PEVPAGCRVT	FAQVLSRHGA	RYPTDSKGKK
土曲霉	SHKWGLYAPY	FSLQDESPFP	LDVPEDCHIT	FVQVLARHGA	RSPTHSKTKA
嗜热踝节菌	SHSWGQYSPF	FSLADQSEIS	PDVPQNCKIT	FVQLLSRHGA	RYPTSSKTEL
T_lanuginosa	ARHWGQYSPF	FSLAEVSEIS	PAVPKGCRVE	FVQVLSRHGA	RYPTAHKSEV
嗜热毁丝霉	SHFWGQYSPY	FSVP..SELD	ASIPDDCEVT	FAQVLSRHGA	RAPTLKRAAS
	84				133
	101				150
P_involtus_A1	IKAGLTKLQG	VQNFTDAKFN	FIKSFKYDLG	NSDLVPFGAA	QSFDAGQEAF
P_involtus_A2	IKAGLSKLQS	VQNFTDPKFD	FIKSFTYDLG	TSDLVPFGAA	QSFDAGLEVF
绒毛栓菌	IQTAVAKLKA	ASNYTDPLLA	FVTNYTYSLG	QDSLVELGAT	QSSEAGQEAF
平田头菇	IQAAVKKLQS	AKTYTDPRLD	FLTNYTYTLG	HDDLVPFGAL	QSSQAGEETF
P_lycii	QVAAVAKIQM	ARPFTDPKYE	FLNDFVYKFG	VADLLPFGAN	QSHQTGTDMY
烟曲霉	YKKLVTAIQA	NATDFKGKFA	FLKTYNYTLG	ADDLTPFGEQ	QLVNSGIKFY
共有植酸酶A	YSALIEAIQK	NATAFKGKYA	FLKTYNYTLG	ADDLTPFGEN	QMVNSGIKFY
构巢曲霉	YSGLIEAIQK	NATSFWGQYA	FLESYNYTLG	ADDLTIFGEN	QMVDSGAKFY
无花果曲霉NRRL3135	YSALIEEIQQ	NATTFDGKYA	FLKTYNYSLG	ADDLTPFGEQ	ELVNSGIKFY
土曲霉	YAATIAAIQK	SATAFPGKYA	FLQSYNYSLD	SEELTPFGRN	QLRDLGAQFY
嗜热踝节菌	YSQLISRIQK	TATAYKGYYA	FLKDYRYQLG	ANDLTPFGEN	QMIQLGIKFY
T_lanuginosa	YAELLQRIQD	TATEFKGDFA	FLRDYAYHLG	ADNLTRFGEE	QMMESGRQFY
嗜热毁丝霉	YVDLIDRIHH	GAISYGPGYE	FLRTYDYTLG	ADELTRTGQQ	QMVNSGIKFY
	134				176
	151				200
P_involtus_A1	ARYSKLVSKN	NLPFIRADGS	DRVVDSATNW	TAGFASA...	....SHNTVQ
P_involtus_A2	ARYSKLVSSD	NLPFIRSDGS	DRVVDTATNW	TAGFASA...	....SRNAIQ

图 1A

```
绒毛栓菌          TRYSSLVSAD ELPFVRASGS DRVVATANNW TAGFALA... ....SSNSIT
平田头菇          QRYSFLVSKE NLPFVRASSS NRVVDSATNW TEGFSAA... ....SHHVLN
P_lycii           TRYSTLFEGG DVPFVRAAGD QRVVDSSTNW TAGFGDA... ....SGETVL
烟曲霉            QRYKAL.ARS VVPFIRASGS DRVIASGEKF IEGFQQAKLA DPGA.TNRAA
共有植酸酶A       RRYKAL.ARK IVPFIRASGS DRVIASAEKF IEGFQSAKLA DPGSQPHQAS
构巢曲霉          RRYKNL.ARK NTPFIRASGS DRVVASAEKF INGFRKAQLH DHGS..KRAT
无花果曲霉NRRL3135 QRYESL.TRN IVPFIRSSGS SRVIASGKKF IEGFQSTKLK DPRAQPGQSS
土曲霉            ERYNAL.TRH INPFVRATDA SRVHESAEKF VEGFQTARQD DHHANPHQPS
嗜热踝节菌        NHYKSL.ARN AVPFVRCSGS DRVIASGRLF IEGFQSAKVL DPHSDKHDAP
T_lanuginosa      HRYREQ.ARE IVPFVRAAGS ARVIASAEFF NRGFQDAKDR DPRSNKDQAE
嗜热毁丝霉        RRYRAL.ARK SIPFVRTAGQ DRVVHSAENF TQGFHSALLA DRGSTVRPTL

                                                                    217
                  177                                               250
                  201
P_involtus_A1     PKLNLILPQT G..NDTLEDN MCPAAGD... ...SDPQVNA WLAVAFPSIT
P_involtus_A2     PKLDLILPQT G..NDTLEDN MCPAAGE... ...SDPQVDA WLASAFPSVT
绒毛栓菌          PVLSVIISEA G..NDTLDDN MCPAAGD... ...SDPQVNQ WLAQFAPPMT
平田头菇          PILFVILSES L..NDTLDDA MCPNAGS... ...SDPQTGI WTSIYGTPIA
P_lycii           PTLQVVLQEE G..NCTLCNN MCPNEVD... ...GD.ESTT WLGVFAPNIT
烟曲霉            PAISVIIPES ETFNNTLDHG VCTKFEA... SQLGDEVAAN FTALFAPDIR
共有植酸酶A       PVIDVIIPEG SGYNNTLDHG TCTAFED... SELGDDVEAN FTALFAPAIR
构巢曲霉          PVVNVIIPEI DGFNNTLDHS TCVSFEN... DERADEIEAN FTAIMGPPIR
无花果曲霉NRRL3135 PKIDVVISEA SSSNNTLDPG TCTVFED... SELADTVEAN FTATFVPSIR
土曲霉            PRVDVAIPEG SAYNNTLEHS LCTAFES... STVGDDAVAN FTAVFAPAIA
嗜热踝节菌        PTINVIIEEG PSYNNTLDTG SCPVFED... SSGGHDAQEK FAKQFAPAIL
T_lanuginosa      PVINVIISEE TGSNNTLDGL TCPAAEE... AP.DPTQPAE FLQVFGPRVL
嗜热毁丝霉        PYDMVVIPET AGANNTLHND LCTAFEEGPY STIGDDAQDT YLSTFAGPIT

                                                                    252
                  218                                               300
                  251
P_involtus_A1     ARLNAAAPSV NLTDTDAFNL VSLCAFLTVS KEKK..... .........S
P_involtus_A2     AQLNAAAPGA NLTDADAFNL VSLCPFMTVS KEQK..... .........S
绒毛栓菌          ARLNAGAPGA NLTDTDTYNL LTLCPFETVA TERR..... .........S
平田头菇          NRLNQQAPGA NITAADVSNL IPLCAFETIV KETP..... .........S
P_lycii           ARLNAAAPSA NLSDSDALTL MDMCPFDTLS SGNA..... .........S
烟曲霉            ARAEKHLPGV TLTDEDVVSL MDMCSFDTVA RTSD..ASQ. .........LS
共有植酸酶A       ARLEADLPGV TLTDEDVVYL MDMCPFETVA RTSD..ATE. .........LS
构巢曲霉          KRLENDLPGI KLTNENVIYL MDMCSFDTMA RTAH..GTE. .........LS
无花果曲霉NRRL3135 QRLENDLSGV TLTDTEVTYL MDMCSFDTIS TSTV..DTK. .........LS
土曲霉            QRLEADLPGV QLSTDDVVNL MAMCPFETVS LTDD..AHT. .........LS
嗜热踝节菌        EKIKDHLPGV DLAVSDVPYL MDLCPFETLA RNHT..DT.. .........LS
T_lanuginosa      KKITKHMPGV NLTLEDVPLF MDLCPFDTVG SDPVLFPRQ. .........LS
嗜热毁丝霉        ARVNANLPGA NLTDADTVAL MDLCPFETVA SSSSDPATAD AGGGNGRPLS

                                                                    300
                  253                                               350
                  301
P_involtus_A1     DFCTLFEGIP GSFEAFAYGG DLDKFYGTGY GQELGPVQGV GYVNELIARL
P_involtus_A2     DFCTLFEGIP GSFEAFAYAG DLDKFYGTGY GQALGPVQGV GYINELLARL
绒毛栓菌          EFCDIYEELQ AE.DAFAYNA DLDKFYGTGY GQPLGPVQGV GYINELLARL
平田头菇          PFCNLFT..P EEFAQFEYFG DLDKFYGTGY GQPLGPVQGV GYINELLARL
P_lycii           PFCDLFT..A EEYVSYEYYY DLDKYYGTGP GNALGPVQGV GYVNELLARL
烟曲霉            PFCQLFT..H NEWKKYNYLQ SLGKYYGYGA GNPLGPAQGI GFTNELIARL
共有植酸酶A       PFCALFT..H DEWRQYDYLQ SLGKYYGYGA GNPLGPAQGV GFANELIARL
构巢曲霉          PFCAIFT..E KEWLQYDYLQ SLSKYYGYGA GSPLGPAQGI GFTNELIARL
```

图 1B

```
无花果曲霉NRRL3135  PFCDLFT..H DEWINYDYLQ SLKKYYGHGA GNPLGPTQGV GYANELIARL
土曲霉              PFCDLFT..A TEWTQYNYLL SLDKYYGYGG GNPLGPVQGV GWANELMARL
嗜热踝节菌          PFCALST..Q EEWQAYDYYQ SLGKYYGNGG GNPLGPAQGV GFVNELIARM
T_lanuginosa        PFCHLFT..A DDWMAYDYYY TLDKYYSHGG GSAFGPSRGV GFVNELIARM
嗜热毁丝霉          PFCRLFS..E SEWRAYDYLQ SVGKWYGYGP GNPLGPTQGV GFVNELLARL

                    301                                                 349
                    351                                                 400
P_involtus_A1       TNS.AVRDNT QTNRTLDASP VTFPLNKTFY ADFSHDNLMV AVFSAMGLFR
P_involtus_A2       TNS.AVNDNT QTNRTLDAAP DTFPLNKTMY ADFSHDNLMV AVFSAMGLFR
绒毛栓菌            TAQ.NVSDHT QTNSTLDSSP ETFPLNRTLY ADFSHDNQMV AIFSAMGLFN
平田头菇            TEM.PVRDNT QTNRTLDSSP LTFPLDRSIY ADLSHDNQMI AIFSAMGLFN
P_lycii             TGQ.AVRDET QTNRTLDSDP ATFPLNRTFY ADFSHDNTMV PIFAALGLFN
烟曲霉              TRS.PVQDHT STNSTLVSNP ATFPLNATMY VDFSHDNSMV SIFFALGLYN
共有植酸酶A         TRS.PVQDHT STNHTLDSNP ATFPLNATLY ADFSHDNSMI SIFFALGLYN
构巢曲霉            TQS.PVQDNT STNHTLDSNP ATFPLDRKLY ADFSHDNSMI SIFFAMGLYN
无花果曲霉NRRL3135  THS.PVHDDT SSNHTLDSSP ATFPLKSTLY ADFSHDNGII SILFALGLYN
土曲霉              TRA.PVHDHT CVNNTLDASP ATFPLNATLY ADFSHDSNLV SIFWALGLYN
嗜热踝节菌          THS.PVQDYT TVNHTLDSNP ATFPLNATLY ADFSHDNTMT SIFAALGLYN
T_lanuginosa        TGNLPVKDHT TVNHTLDDNP ETFPLDAVLY ADFSHDNTMT GIFSAMGLYN
嗜热毁丝霉          A.GVPVRDGT STNRTLDGDP RTFPLGRPLY ADFSHDNDMM GVLGALGAYD

                    350                                         383
                    401                                                 450
P_involtus_A1       QPAPLSTSVP NPWR.....T WRTSSLVPFS GRMVVERLSC ..........
P_involtus_A2       QSAPLSTSTP DPNR.....T WLTSSVVPFS ARMAVERLSC ..........
绒毛栓菌            QSAPLDPTTP DPAR.....T FLVKKIVPFS ARMVVERLDC ..........
平田头菇            QSSPLDPSFP NPKR.....T WVTSRLTPFS ARMVTERLLC QRDGTGSGGP
P_lycii             ATA.LDPLKP DENR.....L WVDSKLVPFS GHMTVEKLAC ..........
烟曲霉              GTEPLSRTSV ESAKE..LDG YSASWVVPFG ARAYFETMQC ..........
共有植酸酶A         GTAPLSTTSV ESIEE..TDG YSASWTVPFG ARAYVEMMQC ..........
构巢曲霉            GTQPLSMDSV ESIQE..MDG YAASWTVPFG ARAYFELMQC ..........
无花果曲霉NRRL3135  GTKPLSTTTV ENITQ..TDG FSSAWTVPFA SRLYVEMMQC ..........
土曲霉              GTAPLSQTSV ESVSQ..TDG YAAAWTVPFA ARAYVEMMQC ..........
嗜热踝节菌          GTAKLSTTEI KSIEE..TDG YSAAWTVPFG GRAYIEMMQC ..........
T_lanuginosa        GTKPLSTSKI QPPTGAAADG YAASWTVPFA ARAYVELLRC ETETSSEEEE
嗜热毁丝霉          GVPPLDKTAR RDPEE..LGG YAASWAVPFA ARIYVEKMRC SGGGGGGGGG

                          384                                           425
                    451                                                 500
P_involtus_A1       .......FGT TKVRVLVQDQ VQPLEFCGGD RNGLCTLAKF VESQTFARSD
P_involtus_A2       .......AGT TKVRVLVQDQ VQPLEFCGGD QDGLCALDKF VESQAYARSG
绒毛栓菌            .......GGA QSVRLLVNDA VQPLAFCGAD TSGVCTLDAF VESQAYARND
平田头菇            SRIMRNGNVQ TFVRILVNDA LQPLKFCGGD MDSLCTLEAF VESQKYARED
P_lycii             .......SGK EAVRVLVNDA VQPLEFCGG. VDGVCELSAF VESQTYAREN
烟曲霉              K..S...EKE PLVRALINDR VVPLHGCDVD KLGRCKLNDF VKGLSWARSG
共有植酸酶A         Q..A...EKE PLVRVLVNDR VVPLHGCAVD KLGRCKRDDF VEGLSFARSG
构巢曲霉            E......KKE PLVRVLVNDR VVPLHGCAVD KFGRCTLDDW VEGLNFARSG
无花果曲霉NRRL3135  Q..A...EQA PLVRVLVNDR VVPLHGCPVD ALGRCTRDSF VRGLSFARSG
土曲霉              R..A...EKE PLVRVLVNDR VMPLHGCPTD KLGRCKRDAF VAGLSFAQAG
嗜热踝节菌          D..D...SDE PVVRVLVNDR VVPLHGCEVD SLGRCKRDDF VRGLSFARQG
T_lanuginosa        E..G...EDE PFVRVLVNDR VVPLHGCRVD RWGRCRRDEW IKGLTFARQG
嗜热毁丝霉          E..GRQEKDE EMVRVLVNDR VMTLKGCGAD ERGMCTLERF IESMAFARGN

                    426         439
```

图 1C

00.1.20

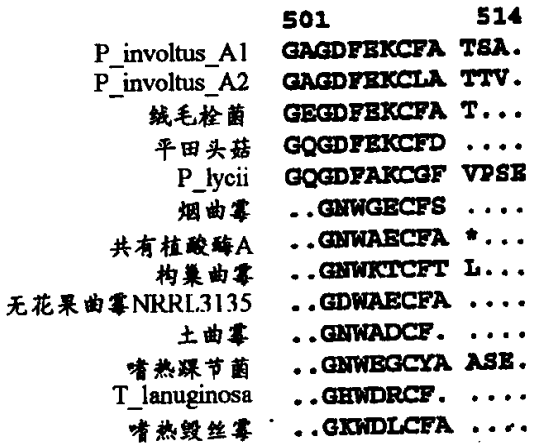

501 514
P_involtus_A1 GAGDFEKCFA T&A.
P_involtus_A2 GAGDFEKCLA TTV.
绒毛栓菌 GEGDFEKCFA T...
平田头菇 GQGDFEKCFD ....
P_lycii GQGDFAKCGF VPSE
烟曲霉 ..GNWGECFS ....
共有植酸酶A ..GNWAECFA *...
构巢曲霉 ..GNWKTCFT L...
无花果曲霉NRRL3135 ..GDWAECFA ....
土曲霉 ..GNWADCF. ....
嗜热琢节菌 ..GNWEGCYA ASE.
T_lanuginosa ..GHWDRCF. ....
嗜热毁丝霉 ..GKWDLCFA ....

图 1D

(2) SEQ ID NO:25的信息:

(i) 序列特征:
(A) 长度:1522碱基对
(B) 类型:核酸
(C) 链型:单链
(D) 拓扑学:线性

(ii) 分子类型: cDNA

(vi) 来源:
(A) 生物: Paxillus involutus
(B) 株: CBS 100231

(ix) 特征:
(A) 名称/关键词: CDS
(B) 位置: 58..1383

(ix) 特征:
(A) 名称/关键词: 成熟肽
(B) 位置: 115..1383

(ix) 特征:
(A) 名称/关键词: 信号肽
(B) 位置: 58..114

(xi) 序列描述: SEQ ID NO:25

```
GGATCCGAAT TCGGCACTCG TACGGTCCCC CGGTCTACCC TCTGCTCGCC TTGGAAG          57

ATG CTC TTC GGT TTC GTC GCC CTC GCC TGT CTC TTG TCC CTC TCC GAG         105
Met Leu Phe Gly Phe Val Ala Leu Ala Cys Leu Leu Ser Leu Ser Glu
-19             -15                 -10                  -5

GTC CTT GCG ACC TCC GTG CCC AAG AAC ACA GCG CCG ACC TTC CCC ATT         153
Val Leu Ala Thr Ser Val Pro Lys Asn Thr Ala Pro Thr Phe Pro Ile
                1               5                  10

CCG GAG AGT GAG CAG CGG AAC TGG TCC CCG TAC TCG CCC TAC TTC CCT         201
Pro Glu Ser Glu Gln Arg Asn Trp Ser Pro Tyr Ser Pro Tyr Phe Pro
    15                  20                  25

CTT GCC GAG TAC AAG GCT CCT CCG GCG GGC TGC CAG ATC AAC CAG GTC         249
Leu Ala Glu Tyr Lys Ala Pro Pro Ala Gly Cys Gln Ile Asn Gln Val
30                  35                  40                  45

AAC ATC ATC CAA AGA CAT GGT GCC CGG TTC CCG ACC TCT GGC GCG ACC         297
Asn Ile Ile Gln Arg His Gly Ala Arg Phe Pro Thr Ser Gly Ala Thr
                50                  55                  60

ACC CGT ATC AAG GCG GGT TTG ACC AAG TTG CAA GGC GTC CAG AAC TTT         345
Thr Arg Ile Lys Ala Gly Leu Thr Lys Leu Gln Gly Val Gln Asn Phe
            65                  70                  75

ACC GAC GCC AAA TTC AAC TTC ATC AAG TCG TTC AAG TAC GAT CTC GGT         393
Thr Asp Ala Lys Phe Asn Phe Ile Lys Ser Phe Lys Tyr Asp Leu Gly
        80                  85                  90

AAC TCG GAC CTC GTT CCG TTC GGT GCA GCA CAG TCC TTC GAC GCT GGT         441
Asn Ser Asp Leu Val Pro Phe Gly Ala Ala Gln Ser Phe Asp Ala Gly
    95                  100                 105
```

图 2A

```
CAG GAG GCC TTC GCC CGC TAC TCG AAG CTT GTC AGC AAG AAC AAC CTG      489
Gln Glu Ala Phe Ala Arg Tyr Ser Lys Leu Val Ser Lys Asn Asn Leu
110                 115                 120                 125

CCG TTC ATT CGT GCC GAT GGA AGT GAT CGT GTT GTG GAT TCT GCT ACA      537
Pro Phe Ile Arg Ala Asp Gly Ser Asp Arg Val Val Asp Ser Ala Thr
                130                 135                 140

AAC TGG ACT GCG GGT TTC GCT TCG GCA AGT CAC AAC ACG GTC CAG CCC      585
Asn Trp Thr Ala Gly Phe Ala Ser Ala Ser His Asn Thr Val Gln Pro
            145                 150                 155

AAG CTG AAC CTG ATT CTC CCG CAA ACT GGC AAT GAT ACC CTG GAA GAT      633
Lys Leu Asn Leu Ile Leu Pro Gln Thr Gly Asn Asp Thr Leu Glu Asp
        160                 165                 170

AAT ATG TGC CCT GCT GCT GGC GAT TCT GAC CCC CAG GTC AAC GCG TGG      681
Asn Met Cys Pro Ala Ala Gly Asp Ser Asp Pro Gln Val Asn Ala Trp
    175                 180                 185

TTG GCT GTT GCT TTC CCT TCC ATC ACT GCA CGG CTC AAC GCC GCC GCG      729
Leu Ala Val Ala Phe Pro Ser Ile Thr Ala Arg Leu Asn Ala Ala Ala
190                 195                 200                 205

CCC TCT GTC AAC CTC ACC GAC ACG GAC GCG TTC AAC CTC GTC AGT CTC      777
Pro Ser Val Asn Leu Thr Asp Thr Asp Ala Phe Asn Leu Val Ser Leu
                210                 215                 220

TGC GCT TTC TTG ACA GTC TCG AAG GAG AAG AAG AGT GAC TTC TGC ACC      825
Cys Ala Phe Leu Thr Val Ser Lys Glu Lys Lys Ser Asp Phe Cys Thr
            225                 230                 235

CTG TTC GAG GGC ATC CCT GGC TCT TTC GAG GCG TTC GCC TAT GGT GGC      873
Leu Phe Glu Gly Ile Pro Gly Ser Phe Glu Ala Phe Ala Tyr Gly Gly
        240                 245                 250

GAC CTT GAC AAG TTC TAC GGT ACC GGT TAC GGT CAG GAA CTC GGA CCC      921
Asp Leu Asp Lys Phe Tyr Gly Thr Gly Tyr Gly Gln Glu Leu Gly Pro
    255                 260                 265

GTT CAA GGC GTC GGC TAC GTC AAC GAG CTC ATC GCC CGC CTC ACC AAC      969
Val Gln Gly Val Gly Tyr Val Asn Glu Leu Ile Ala Arg Leu Thr Asn
270                 275                 280                 285

TCC GCC GTC CGC GAC AAC ACC CAG ACG AAC CGC ACA CTC GAC GCC TCG     1017
Ser Ala Val Arg Asp Asn Thr Gln Thr Asn Arg Thr Leu Asp Ala Ser
                290                 295                 300

CCC GTA ACC TTC CCG TTG AAC AAG ACG TTC TAC GCC GAT TTC TCC CAC     1065
Pro Val Thr Phe Pro Leu Asn Lys Thr Phe Tyr Ala Asp Phe Ser His
            305                 310                 315

GAC AAC CTC ATG GTC GCC GTC TTC TCC GCC ATG GGC CTC TTC CGC CAG     1113
Asp Asn Leu Met Val Ala Val Phe Ser Ala Met Gly Leu Phe Arg Gln
        320                 325                 330

CCC GCG CCG CTC AGC ACG TCC GTG CCG AAC CCA TGG CGC ACG TGG CGC     1161
Pro Ala Pro Leu Ser Thr Ser Val Pro Asn Pro Trp Arg Thr Trp Arg
    335                 340                 345
```

图 2B

```
ACG AGC TCC CTC GTC CCC TTC TCC GGA CGC ATG GTC GTG GAA CGC CTC      1209
Thr Ser Ser Leu Val Pro Phe Ser Gly Arg Met Val Val Glu Arg Leu
350                 355                 360                 365

AGC TGT TTC GGC ACG ACC AAG GTT CGC GTC CTC GTG CAG GAC CAG GTG      1257
Ser Cys Phe Gly Thr Thr Lys Val Arg Val Leu Val Gln Asp Gln Val
                370                 375                 380

CAG CCG CTC GAG TTC TGC GGG GGT GAT AGG AAC GGG CTG TGC ACG CTT      1305
Gln Pro Leu Glu Phe Cys Gly Gly Asp Arg Asn Gly Leu Cys Thr Leu
            385                 390                 395

GCT AAG TTT GTG GAG AGC CAG ACG TTT GCG AGG AGT GAT GGT GCG GGG      1353
Ala Lys Phe Val Glu Ser Gln Thr Phe Ala Arg Ser Asp Gly Ala Gly
        400                 405                 410

GAC TTT GAG AAG TGC TTC GCG ACC TCG GCG TGAGGATGGA CGAACAAAAT        1403
Asp Phe Glu Lys Cys Phe Ala Thr Ser Ala
    415                 420

TAAATTGGGG TATTTTATCG TATAATTATG GTGTGTGTAG AACATGGGCT CGGGGTCGAT    1463

GGTGAAAAGC AAAGGTTTAT CGTCTAAAAA AAAAAAAAAA AAAAAATTCC TGCGGCCGC     1522
```

图 2C

(2) SEQ ID NO:27的信息:

(i) 序列特征:
(A) 长度:1642碱基对
(B) 类型:核酸
(C) 链型:单链
(D) 拓扑学:线性

(ii) 分子类型: cDNA

(vi) 来源:
(A) 生物: Paxillus involutus
(B) 株: CBS 100231

(ix) 特征:
(A) 名称/关键词: CDS
(B) 位置: 48 . . 1373

(ix) 特征:
(A) 名称/关键词: 成熟肽
(B) 位置: 105 . . 1373

(ix) 特征:
(A) 名称/关键词: 信号肽
(B) 位置: 48 . . 104

(xi) 序列描述: SEQ ID NO:27

```
GGATCCGAAT TCCAGTCCCC AAGCTAATCC TCTGCTCGCC TTGGAAG ATG CAC CTC        56
                                                    Met His Leu
                                                    -19

GGC TTC GTC ACC CTC GCT TGT CTC ATA CAC CTC TCC GAG GTC TTC GCG       104
Gly Phe Val Thr Leu Ala Cys Leu Ile His Leu Ser Glu Val Phe Ala
    -15                 -10                 -5

GCA TCC GTG CCC CGG AAT ATT GCT CCG AAG TTC TCA ATT CCG GAA AGC       152
Ala Ser Val Pro Arg Asn Ile Ala Pro Lys Phe Ser Ile Pro Glu Ser
  1               5                  10                  15

GAG CAG CGA AAC TGG TCG CCT TAC TCT CCT TAC TTT CCC CTA GCC GAA       200
Glu Gln Arg Asn Trp Ser Pro Tyr Ser Pro Tyr Phe Pro Leu Ala Glu
            20                  25                  30

TAC AAG GCT CCT CCA GCA GGC TGC GAG ATT AAC CAA GTC AAT ATT ATC       248
Tyr Lys Ala Pro Pro Ala Gly Cys Glu Ile Asn Gln Val Asn Ile Ile
        35                  40                  45

CAA CGG CAT GGC GCA CGG TTC CCA ACC TCG GGT GCG GCC ACT CGC ATC       296
Gln Arg His Gly Ala Arg Phe Pro Thr Ser Gly Ala Ala Thr Arg Ile
    50                  55                  60

AAG GCT GGT TTA AGC AAG CTG CAA TCC GTC CAG AAT TTC ACC GAC CCC       344
Lys Ala Gly Leu Ser Lys Leu Gln Ser Val Gln Asn Phe Thr Asp Pro
65                  70                  75                  80

AAA TTC GAC TTC ATC AAG TCG TTC ACA TAC GAT CTT GGT ACT TCC GAC       392
Lys Phe Asp Phe Ile Lys Ser Phe Thr Tyr Asp Leu Gly Thr Ser Asp
                85                  90                  95
```

图 3A

```
CTC GTG CCA TTC GGC GCA GCA CAA TCA TTC GAT GCC GGC CTG GAG GTC      440
Leu Val Pro Phe Gly Ala Ala Gln Ser Phe Asp Ala Gly Leu Glu Val
            100                 105                 110

TTC GCT CGC TAT TCG AAG CTC GTC AGC TCG GAC AAC CTG CCT TTC ATT      488
Phe Ala Arg Tyr Ser Lys Leu Val Ser Ser Asp Asn Leu Pro Phe Ile
        115                 120                 125

CGC TCA GAT GGT AGC GAT CGT GTA GTC GAC ACT GCT ACG AAC TGG ACT      536
Arg Ser Asp Gly Ser Asp Arg Val Val Asp Thr Ala Thr Asn Trp Thr
    130                 135                 140

GCA GGT TTT GCT TCC GCG AGC CGC AAC GCG ATC CAA CCC AAG CTC GAC      584
Ala Gly Phe Ala Ser Ala Ser Arg Asn Ala Ile Gln Pro Lys Leu Asp
145                 150                 155                 160

TTG ATA CTT CCA CAA ACT GGC AAT GAC ACC CTC GAG GAC AAC ATG TGT      632
Leu Ile Leu Pro Gln Thr Gly Asn Asp Thr Leu Glu Asp Asn Met Cys
                165                 170                 175

CCA GCT GCT GGC GAA TCC GAC CCT CAG GTC GAT GCG TGG TTG GCG TCC      680
Pro Ala Ala Gly Glu Ser Asp Pro Gln Val Asp Ala Trp Leu Ala Ser
            180                 185                 190

GCC TTC CCA TCT GTC ACC GCG CAG CTC AAC GCT GCA GCG CCT GGT GCC      728
Ala Phe Pro Ser Val Thr Ala Gln Leu Asn Ala Ala Ala Pro Gly Ala
        195                 200                 205

AAT CTC ACA GAC GCC GAC GCC TTC AAC CTC GTC AGC CTG TGT CCC TTC      776
Asn Leu Thr Asp Ala Asp Ala Phe Asn Leu Val Ser Leu Cys Pro Phe
    210                 215                 220

ATG ACA GTT TCG AAG GAG CAG AAG AGC GAC TTC TGC ACG TTG TTC GAG      824
Met Thr Val Ser Lys Glu Gln Lys Ser Asp Phe Cys Thr Leu Phe Glu
225                 230                 235                 240

GGA ATC CCT GGA TCG TTC GAG GCG TTT GCC TAT GCC GGC GAC CTT GAC      872
Gly Ile Pro Gly Ser Phe Glu Ala Phe Ala Tyr Ala Gly Asp Leu Asp
                245                 250                 255

AAG TTC TAT GGG ACC GGC TAT GGC CAA GCC CTC GGA CCG GTC CAA GGC      920
Lys Phe Tyr Gly Thr Gly Tyr Gly Gln Ala Leu Gly Pro Val Gln Gly
            260                 265                 270

GTC GGC TAC ATC AAC GAG CTC CTT GCA CGC CTG ACC AAC TCC GCA GTG      968
Val Gly Tyr Ile Asn Glu Leu Leu Ala Arg Leu Thr Asn Ser Ala Val
        275                 280                 285

AAC GAC AAC ACA CAG ACG AAC CGC ACA CTC GAC GCC GCA CCA GAC ACG     1016
Asn Asp Asn Thr Gln Thr Asn Arg Thr Leu Asp Ala Ala Pro Asp Thr
    290                 295                 300

TTC CCG CTC AAC AAG ACC ATG TAC GCC GAT TTC TCA CAC GAC AAC CTC     1064
Phe Pro Leu Asn Lys Thr Met Tyr Ala Asp Phe Ser His Asp Asn Leu
305                 310                 315                 320

ATG GTC GCC GTG TTC TCC GCC ATG GGC CTC TTC CGC CAA TCC GCA CCG     1112
Met Val Ala Val Phe Ser Ala Met Gly Leu Phe Arg Gln Ser Ala Pro
                325                 330                 335
```

图 3B

```
CTC AGC ACG TCC ACA CCG GAT CCG AAC CGC ACG TGG CTC ACG AGC TCT      1160
Leu Ser Thr Ser Thr Pro Asp Pro Asn Arg Thr Trp Leu Thr Ser Ser
            340                 345                 350

GTC GTT CCG TTC TCC GCG CGC ATG GCC GTG GAA CGC CTC AGC TGT GCT      1208
Val Val Pro Phe Ser Ala Arg Met Ala Val Glu Arg Leu Ser Cys Ala
        355                 360                 365

GGT ACC ACG AAG GTG CGC GTC CTG GTG CAG GAC CAG GTC CAG CCA CTC      1256
Gly Thr Thr Lys Val Arg Val Leu Val Gln Asp Gln Val Gln Pro Leu
    370                 375                 380

GAG TTC TGC GGC GGC GAC CAG GAT GGG TTG TGC GCG CTA GAC AAG TTC      1304
Glu Phe Cys Gly Gly Asp Gln Asp Gly Leu Cys Ala Leu Asp Lys Phe
385                 390                 395                 400

GTC GAG AGC CAG GCG TAT GCA CGG AGT GGT GGC GCA GGT GAC TTT GAG      1352
Val Glu Ser Gln Ala Tyr Ala Arg Ser Gly Gly Ala Gly Asp Phe Glu
                405                 410                 415

AAG TGT CTT GCG ACG ACG GTG TGAGATGGGG TAATCTACGG TGAAGCAGCG         1403
Lys Cys Leu Ala Thr Thr Val
            420

GAGAGCCTCT CAACGAATGC AAAGGATAGG TTCGAGGCTT ACTTCATCAA CCTATATCAT    1463

CATAGGACAA GCCCCCCAAT AGCCAGACTC GTCGTTTGAC ATCGTGTATG AAAATAACCC    1523

ACCCACGCAC TCCGCTGCCA CTATTCGCGT GTATCGCATA CTAGGCGTTT TCGCCCAGTT    1583

GAACATGAGC CCATTCTGTC CCCAGTGAAA AAAAAAAAAA AAAAAATTCC TGCGGCCGC     1642
```

图 3C

(2) SEQ ID NO: 29的信息:

(i) 序列特征:
(A) 长度:1536碱基对
(B) 类型:核酸
(C) 链型:单链
(D) 拓扑学:线性

(ii) 分子类型: cDNA

(vi) 来源:
(A) 生物: 绒毛栓菌
(B) 株: CBS 100232

(ix) 特征:
(A) 名称/关键词: CDS
(B) 位置: 79 . . 1407

(ix) 特征:
(A) 名称/关键词: 成熟肽
(B) 位置: 130 . . 1407

(ix) 特征:
(A) 名称/关键词: 信号肽
(B) 位置: 79 . . 129

(xi) 序列描述: SEQ ID NO:29

```
GGATCCGAAT TCGCCCCCAC ATTCGTTCCA TCTTAGCAGC CGTCCGCGCC CAGGTCTTCG      60

ATAACCCCCC GCGTGACT ATG GCC TTC TCA ATC TTG GCC TCG CTG CTC TTC        111
                    Met Ala Phe Ser Ile Leu Ala Ser Leu Leu Phe
                    -17     -15                 -10

GTG TGT TAT GCA TAC GCC AGG GCT GTG CCC CGT GCA CAT ATC CCG CTC        159
Val Cys Tyr Ala Tyr Ala Arg Ala Val Pro Arg Ala His Ile Pro Leu
 -5                  1                   5                  10

CGC GAC ACC TCC GCG TGT CTA GAT GTA ACA CGC GAT GTG CAG CAG AGC        207
Arg Asp Thr Ser Ala Cys Leu Asp Val Thr Arg Asp Val Gln Gln Ser
                15                  20                  25

TGG TCC ATG TAC TCT CCC TAT TTC CCG GCA GCA ACT TAT GTG GCT CCG        255
Trp Ser Met Tyr Ser Pro Tyr Phe Pro Ala Ala Thr Tyr Val Ala Pro
            30                  35                  40

CCC GCG AGT TGC CAG ATC AAT CAG GTC CAC ATC ATC CAA CGT CAT GGT        303
Pro Ala Ser Cys Gln Ile Asn Gln Val His Ile Ile Gln Arg His Gly
        45                  50                  55

GCA CGC TTT CCC ACG TCT GGC GCA GCA AAG CGC ATC CAG ACA GCA GTA        351
Ala Arg Phe Pro Thr Ser Gly Ala Ala Lys Arg Ile Gln Thr Ala Val
    60                  65                  70

GCG AAG CTG AAG GCC GCG TCC AAC TAC ACC GAT CCC CTG CTC GCG TTC        399
Ala Lys Leu Lys Ala Ala Ser Asn Tyr Thr Asp Pro Leu Leu Ala Phe
75                  80                  85                  90

GTT ACG AAC TAC ACC TAC AGC TTA GGT CAG GAC AGC CTC GTT GAA CTC        447
Val Thr Asn Tyr Thr Tyr Ser Leu Gly Gln Asp Ser Leu Val Glu Leu
                95                  100                 105
```

图 4A

```
GGT GCG ACT CAG TCC TCC GAA GCG GGC CAG GAG GCA TTC ACG CGG TAC      495
Gly Ala Thr Gln Ser Ser Glu Ala Gly Gln Glu Ala Phe Thr Arg Tyr
            110                 115                 120

TCA TCC CTC GTG AGC GCG GAC GAG CTT CCC TTC GTT CGG GCG TCG GGC      543
Ser Ser Leu Val Ser Ala Asp Glu Leu Pro Phe Val Arg Ala Ser Gly
        125                 130                 135

TCA GAT CGC GTC GTT GCG ACT GCC AAC AAC TGG ACT GCA GGT TTC GCG      591
Ser Asp Arg Val Val Ala Thr Ala Asn Asn Trp Thr Ala Gly Phe Ala
    140                 145                 150

CTT GCG AGC TCA AAC AGC ATC ACG CCC GTG CTC TCA GTC ATC ATT TCC      639
Leu Ala Ser Ser Asn Ser Ile Thr Pro Val Leu Ser Val Ile Ile Ser
155                 160                 165                 170

GAA GCG GGC AAT GAC ACC CTC GAC GAC AAC ATG TGC CCC GCT GCA GGC      687
Glu Ala Gly Asn Asp Thr Leu Asp Asp Asn Met Cys Pro Ala Ala Gly
                175                 180                 185

GAT TCG GAT CCC CAG GTC AAT CAA TGG CTC GCG CAG TTC GCA CCG CCG      735
Asp Ser Asp Pro Gln Val Asn Gln Trp Leu Ala Gln Phe Ala Pro Pro
            190                 195                 200

ATG ACT GCT CGC CTC AAC GCA GGC GCG CCC GGC GCG AAC CTC ACG GAC      783
Met Thr Ala Arg Leu Asn Ala Gly Ala Pro Gly Ala Asn Leu Thr Asp
        205                 210                 215

ACG GAC ACC TAC AAC CTG CTC ACG CTA TGC CCG TTC GAG ACT GTA GCC      831
Thr Asp Thr Tyr Asn Leu Leu Thr Leu Cys Pro Phe Glu Thr Val Ala
    220                 225                 230

ACC GAG CGG CGT AGT GAA TTC TGC GAC ATC TAC GAG GAG CTG CAG GCG      879
Thr Glu Arg Arg Ser Glu Phe Cys Asp Ile Tyr Glu Glu Leu Gln Ala
235                 240                 245                 250

GAA GAC GCC TTC GCG TAC AAT GCC GAT CTC GAC AAG TTC TAC GGC ACT      927
Glu Asp Ala Phe Ala Tyr Asn Ala Asp Leu Asp Lys Phe Tyr Gly Thr
                255                 260                 265

GGA TAC GGC CAG CCC CTC GGA CCC GTG CAA GGC GTC GGG TAC ATC AAC      975
Gly Tyr Gly Gln Pro Leu Gly Pro Val Gln Gly Val Gly Tyr Ile Asn
            270                 275                 280

GAG CTC ATC GCG CGC CTC ACC GCG CAG AAC GTG TCC GAC CAC ACG CAG     1023
Glu Leu Ile Ala Arg Leu Thr Ala Gln Asn Val Ser Asp His Thr Gln
        285                 290                 295

ACG AAC AGC ACA CTC GAC TCC TCG CCC GAG ACG TTC CCG CTC AAC CGC     1071
Thr Asn Ser Thr Leu Asp Ser Ser Pro Glu Thr Phe Pro Leu Asn Arg
    300                 305                 310

ACG CTC TAC GCG GAC TTC TCG CAC GAC AAC CAG ATG GTC GCG ATC TTC     1119
Thr Leu Tyr Ala Asp Phe Ser His Asp Asn Gln Met Val Ala Ile Phe
315                 320                 325                 330

TCG GCC ATG GGT CTC TTC AAC CAG TCC GCG CCG CTC GAC CCG ACG ACG     1167
Ser Ala Met Gly Leu Phe Asn Gln Ser Ala Pro Leu Asp Pro Thr Thr
                335                 340                 345
```

图 4B

```
CCC GAC CCC GCG CGC ACG TTC CTC GTC AAG AAG ATC GTG CCG TTC TCC      1215
Pro Asp Pro Ala Arg Thr Phe Leu Val Lys Lys Ile Val Pro Phe Ser
            350                 355                 360

GCG CGC ATG GTC GTC GAG CGC CTC GAC TGC GGC GGT GCG CAG AGC GTG      1263
Ala Arg Met Val Val Glu Arg Leu Asp Cys Gly Gly Ala Gln Ser Val
        365                 370                 375

CGC CTG CTC GTG AAC GAC GCA GTG CAG CCG CTG GCG TTC TGC GGG GCG      1311
Arg Leu Leu Val Asn Asp Ala Val Gln Pro Leu Ala Phe Cys Gly Ala
    380                 385                 390

GAC ACG AGC GGG GTG TGC ACG CTG GAC GCG TTT GTC GAG AGC CAG GCG      1359
Asp Thr Ser Gly Val Cys Thr Leu Asp Ala Phe Val Glu Ser Gln Ala
395                 400                 405                 410

TAC GCG CGG AAC GAT GGC GAG GGC GAC TTC GAG AAG TGC TTC GCG ACA      1407
Tyr Ala Arg Asn Asp Gly Glu Gly Asp Phe Glu Lys Cys Phe Ala Thr
                415                 420                 425

TAGTTCCAGG TGTAGATACC CGGGGAAGAT GTACTCTCTA GACACCTCGC ATGTACTTAT    1467

CGATTAGAAA GAGACCCTGG CTGCTCTGCC CTCAAAAAAA AAAAAAAAAA AAAAAATTCC    1527

TGCGGCCGC                                                            1536
```

图 4C

(2) SEQ ID NO:21的信息:

(i) 序列特征:
(A) 长度:1501碱基对
(B) 类型:核酸
(C) 链型:单链
(D) 拓扑学:线性

(ii) 分子类型: cDNA

(vi) 来源:
(A) 生物: 平田头菇
(B) 株: CBS 900. 96

(ix) 特征:
(A) 名称/关键词: CDS
(B) 位置: 17 . . 1375

(ix) 特征:
(A) 名称/关键词: 信号肽
(B) 位置: 17 . . 94

(ix) 特征:
(A) 名称/关键词: 成熟肽
(B) 位置: 95 . . 1375

(xi) 序列描述: SEQ ID NO:21

```
GGATCCGAAT TCACTT ATG TCC CTC TTC ATC GGC GGC TGT TTG CTC GTG          49
                  Met Ser Leu Phe Ile Gly Gly Cys Leu Leu Val
                  -26 -25                 -20

TTT TTA CAG GCG AGC GCA TAC GGC GGC GTC GTG CAG GCC ACA TTC GTG        97
Phe Leu Gln Ala Ser Ala Tyr Gly Gly Val Val Gln Ala Thr Phe Val
-15                 -10                 -5                   1

CAG CCG TTT TTC CCT CCA CAG ATT CAG GAC TCT TGG GCA GCT TAT ACA       145
Gln Pro Phe Phe Pro Pro Gln Ile Gln Asp Ser Trp Ala Ala Tyr Thr
              5                  10                  15

CCA TAT TAT CCT GTT CAG GCG TAC ACG CCT CCC CCG AAG GAT TGC AAG       193
Pro Tyr Tyr Pro Val Gln Ala Tyr Thr Pro Pro Pro Lys Asp Cys Lys
         20                  25                  30

ATC ACA CAA GTT AAC ATT ATT CAA CGA CAT GGT GCC CGC TTT CCG ACA       241
Ile Thr Gln Val Asn Ile Ile Gln Arg His Gly Ala Arg Phe Pro Thr
     35                  40                  45

TCG GGG GCA GGC ACA AGG ATC CAA GCA GCT GTG AAG AAG CTT CAA TCA       289
Ser Gly Ala Gly Thr Arg Ile Gln Ala Ala Val Lys Lys Leu Gln Ser
 50                  55                  60                  65

GCT AAA ACC TAT ACG GAT CCT CGT CTC GAC TTT CTG ACC AAC TAT ACC       337
Ala Lys Thr Tyr Thr Asp Pro Arg Leu Asp Phe Leu Thr Asn Tyr Thr
                 70                  75                  80

TAT ACC CTT GGT CAC GAC GAT CTC GTA CCG TTT GGA GCG CTT CAA TCA       385
Tyr Thr Leu Gly His Asp Asp Leu Val Pro Phe Gly Ala Leu Gln Ser
             85                  90                  95
```

图 5A

```
TCA CAA GCT GGA GAG GAA ACG TTT CAA CGA TAC TCG TTT CTG GTG TCC      433
Ser Gln Ala Gly Glu Glu Thr Phe Gln Arg Tyr Ser Phe Leu Val Ser
        100                 105                 110

AAA GAG AAC TTA CCT TTT GTA AGA GCT TCG AGT TCC AAT CGA GTC GTC      481
Lys Glu Asn Leu Pro Phe Val Arg Ala Ser Ser Ser Asn Arg Val Val
    115                 120                 125

GAC TCA GCT ACC AAC TGG ACG GAA GGT TTT TCT GCG GCC AGT CAC CAC      529
Asp Ser Ala Thr Asn Trp Thr Glu Gly Phe Ser Ala Ala Ser His His
130                 135                 140                 145

GTC TTG AAT CCC ATT CTC TTT GTA ATC CTC TCA GAA AGT CTC AAT GAC      577
Val Leu Asn Pro Ile Leu Phe Val Ile Leu Ser Glu Ser Leu Asn Asp
                150                 155                 160

ACG CTT GAC GAT GCC ATG TGC CCT AAC GCG GGC TCC TCC GAC CCG CAG      625
Thr Leu Asp Asp Ala Met Cys Pro Asn Ala Gly Ser Ser Asp Pro Gln
            165                 170                 175

ACT GGT ATC TGG ACC TCG ATA TAC GGG ACG CCT ATT GCC AAC CGA CTA      673
Thr Gly Ile Trp Thr Ser Ile Tyr Gly Thr Pro Ile Ala Asn Arg Leu
        180                 185                 190

AAT CAG CAG GCT CCG GGT GCA AAT ATT ACA GCT GCC GAT GTG TCG AAC      721
Asn Gln Gln Ala Pro Gly Ala Asn Ile Thr Ala Ala Asp Val Ser Asn
    195                 200                 205

CTT ATA CCG CTT TGC GCA TTC GAG ACG ATA GTA AAG GAG ACG CCA AGT      769
Leu Ile Pro Leu Cys Ala Phe Glu Thr Ile Val Lys Glu Thr Pro Ser
210                 215                 220                 225

CCT TTC TGT AAT TTG TTC ACC CCC GAA GAG TTC GCA CAG TTT GAA TAT      817
Pro Phe Cys Asn Leu Phe Thr Pro Glu Glu Phe Ala Gln Phe Glu Tyr
                230                 235                 240

TTC GGT GAC CTG GAC AAG TTC TAT GGG ACA GGT TAT GGA CAA CCG TTA      865
Phe Gly Asp Leu Asp Lys Phe Tyr Gly Thr Gly Tyr Gly Gln Pro Leu
            245                 250                 255

GGA CCT GTG CAA GGT GTC GGC TAC ATC AAT GAA CTT CTT GCC CGA CTC      913
Gly Pro Val Gln Gly Val Gly Tyr Ile Asn Glu Leu Leu Ala Arg Leu
        260                 265                 270

ACA GAA ATG CCA GTT CGA GAT AAC ACC CAG ACG AAC AGG ACA CTC GAC      961
Thr Glu Met Pro Val Arg Asp Asn Thr Gln Thr Asn Arg Thr Leu Asp
    275                 280                 285

TCT TCT CCG CTT ACA TTT CCC CTC GAC CGC AGT ATC TAC GCT GAC CTC     1009
Ser Ser Pro Leu Thr Phe Pro Leu Asp Arg Ser Ile Tyr Ala Asp Leu
290                 295                 300                 305

TCG CAC GAT AAC CAA ATG ATC GCG ATA TTT TCA GCG ATG GGT CTT TTC     1057
Ser His Asp Asn Gln Met Ile Ala Ile Phe Ser Ala Met Gly Leu Phe
                310                 315                 320

AAC CAG AGT TCA CCT TTG GAT CCG TCC TTC CCC AAC CCC AAG CGT ACT     1105
Asn Gln Ser Ser Pro Leu Asp Pro Ser Phe Pro Asn Pro Lys Arg Thr
            325                 330                 335
```

图 5B

```
TGG GTC ACC AGT CGG CTT ACG CCT TTC AGC GCG AGA ATG GTC ACT GAG      1153
Trp Val Thr Ser Arg Leu Thr Pro Phe Ser Ala Arg Met Val Thr Glu
        340                 345                 350

CGG TTG CTG TGT CAA AGG GAT GGG ACA GGG AGC GGT GGA CCA TCC AGG      1201
Arg Leu Leu Cys Gln Arg Asp Gly Thr Gly Ser Gly Gly Pro Ser Arg
    355                 360                 365

ATC ATG CGG AAT GGA AAT GTG CAG ACG TTT GTG AGG ATT CTT GTC AAC      1249
Ile Met Arg Asn Gly Asn Val Gln Thr Phe Val Arg Ile Leu Val Asn
370                 375                 380                 385

GAT GCT TTA CAG CCT TTG AAG TTC TGC GGA GGG GAC ATG GAT AGT TTG      1297
Asp Ala Leu Gln Pro Leu Lys Phe Cys Gly Gly Asp Met Asp Ser Leu
                390                 395                 400

TGT ACT CTG GAA GCG TTC GTC GAG AGC CAG AAG TAT GCA CGA GAG GAT      1345
Cys Thr Leu Glu Ala Phe Val Glu Ser Gln Lys Tyr Ala Arg Glu Asp
            405                 410                 415

GGT CAA GGC GAT TTT GAA AAA TGT TTT GAT TAAATATTGC AGTATGCTCA        1395
Gly Gln Gly Asp Phe Glu Lys Cys Phe Asp
        420                 425

GTGAGTAGAC TACAGTGCAG GCCCTGTAAC TCTTGTATTG TGTTTCTGGA ATTCCTCGGA    1455

GCGTAGTTTG TAGCAAAAAA AAAAAAAAAA AAATTCCTGC GGCCGC                   1501
```

图 5C

(2) SEQ ID NO:23的信息:

(i) 序列特征:
(A) 长度:1593碱基对
(B) 类型:核酸
(C) 链型:单链
(D) 拓扑学:线性

(ii) 分子类型: cDNA

(vi) 来源:
(A) 生物: Peniophora Lycii
(B) 株: CBS 686. 96

(ix) 特征:
(A) 名称/关键词: 信号肽
(B) 位置: 123 . . 213

(ix) 特征:
(A) 名称/关键词: 成熟肽
(B) 位置: 213 . . 1439

(ix) 特征:
(A) 名称/关键词: CDS
(B) 位置: 123 . . 1439

(xi) 序列描述: SEQ ID NO:23

```
GGATCCGAAT TCCATCTTCT GCTCTGACCT CCATCTCGCT GAGCGGCCGA CGAGAACCTA      60

GGGGCTCTAA GTCCACGTAC TATCGCCGCG CCTGTGAAGG CCCCATACCA GCCCTTATCG     120

AT ATG GTT TCT TCG GCA TTC GCA CCT TCC ATC CTA CTT AGC TTG ATG         167
   Met Val Ser Ser Ala Phe Ala Pro Ser Ile Leu Leu Ser Leu Met
   -30                 -25                 -20

TCG AGT CTT GCT TTG AGC ACG CAG TTC AGC TTT GTT GCG GCG CAG CTA       215
Ser Ser Leu Ala Leu Ser Thr Gln Phe Ser Phe Val Ala Ala Gln Leu
-15                 -10                  -5                   1

CCT ATC CCC GCA CAA AAC ACA AGT AAT TGG GGG CCT TAC GAT CCC TTC       263
Pro Ile Pro Ala Gln Asn Thr Ser Asn Trp Gly Pro Tyr Asp Pro Phe
               5                  10                  15

TTT CCC GTC GAA CCG TAT GCA GCT CCG CCG GAA GGG TGC ACA GTG ACA       311
Phe Pro Val Glu Pro Tyr Ala Ala Pro Pro Glu Gly Cys Thr Val Thr
          20                  25                  30

CAG GTC AAC CTG ATT CAG AGG CAC GGC GCG CGT TGG CCC ACA TCC GGC       359
Gln Val Asn Leu Ile Gln Arg His Gly Ala Arg Trp Pro Thr Ser Gly
     35                  40                  45

GCG CGG TCG CGG CAG GTC GCC GCC GTA GCG AAG ATA CAA ATG GCG CGA       407
Ala Arg Ser Arg Gln Val Ala Ala Val Ala Lys Ile Gln Met Ala Arg
 50                  55                  60                  65

CCA TTC ACG GAT CCC AAG TAT GAG TTC CTC AAC GAC TTC GTG TAC AAG       455
Pro Phe Thr Asp Pro Lys Tyr Glu Phe Leu Asn Asp Phe Val Tyr Lys
                70                  75                  80
```

图 6A

```
TTC GGC GTC GCC GAT CTG CTA CCG TTC GGG GCT AAC CAA TCG CAC CAA          503
Phe Gly Val Ala Asp Leu Leu Pro Phe Gly Ala Asn Gln Ser His Gln
                85                  90                  95

ACC GGC ACC GAT ATG TAT ACG CGC TAC AGT ACA CTA TTT GAG GGC GGG          551
Thr Gly Thr Asp Met Tyr Thr Arg Tyr Ser Thr Leu Phe Glu Gly Gly
        100                 105                 110

GAT GTA CCC TTT GTG CGC GCG GCT GGT GAC CAA CGC GTC GTT GAC TCC          599
Asp Val Pro Phe Val Arg Ala Ala Gly Asp Gln Arg Val Val Asp Ser
    115                 120                 125

TCG ACG AAC TGG ACG GCA GGC TTT GGC GAT GCT TCT GGC GAG ACT GTT          647
Ser Thr Asn Trp Thr Ala Gly Phe Gly Asp Ala Ser Gly Glu Thr Val
130                 135                 140                 145

CTC CCG ACG CTC CAG GTT GTG CTT CAA GAA GAG GGG AAC TGC ACG CTC          695
Leu Pro Thr Leu Gln Val Val Leu Gln Glu Glu Gly Asn Cys Thr Leu
                150                 155                 160

TGC AAT AAT ATG TGC CCG AAT GAA GTG GAT GGT GAC GAA TCC ACA ACG          743
Cys Asn Asn Met Cys Pro Asn Glu Val Asp Gly Asp Glu Ser Thr Thr
            165                 170                 175

TGG CTG GGG GTC TTT GCG CCG AAC ATC ACC GCG CGA TTG AAC GCT GCT          791
Trp Leu Gly Val Phe Ala Pro Asn Ile Thr Ala Arg Leu Asn Ala Ala
        180                 185                 190

GCG CCG AGT GCC AAC CTC TCA GAC AGC GAC GCG CTC ACT CTC ATG GAT          839
Ala Pro Ser Ala Asn Leu Ser Asp Ser Asp Ala Leu Thr Leu Met Asp
    195                 200                 205

ATG TGC CCG TTC GAC ACT CTC AGC TCC GGG AAC GCC AGC CCC TTC TGT          887
Met Cys Pro Phe Asp Thr Leu Ser Ser Gly Asn Ala Ser Pro Phe Cys
210                 215                 220                 225

GAC CTA TTT ACC GCG GAG GAG TAT GTG TCG TAC GAG TAC TAC TAT GAC          935
Asp Leu Phe Thr Ala Glu Glu Tyr Val Ser Tyr Glu Tyr Tyr Tyr Asp
                230                 235                 240

CTC GAC AAG TAC TAT GGC ACG GGC CCC GGG AAC GCT CTC GGT CCT GTC          983
Leu Asp Lys Tyr Tyr Gly Thr Gly Pro Gly Asn Ala Leu Gly Pro Val
            245                 250                 255

CAG GGC GTC GGA TAC GTC AAT GAG CTG CTT GCA CGC TTG ACC GGC CAA         1031
Gln Gly Val Gly Tyr Val Asn Glu Leu Leu Ala Arg Leu Thr Gly Gln
        260                 265                 270

GCC GTT CGA GAC GAG ACG CAG ACG AAC CGC ACG CTC GAC AGC GAC CCT         1079
Ala Val Arg Asp Glu Thr Gln Thr Asn Arg Thr Leu Asp Ser Asp Pro
    275                 280                 285

GCA ACA TTC CCG CTG AAC CGT ACG TTC TAC GCC GAC TTC TCG CAT GAT         1127
Ala Thr Phe Pro Leu Asn Arg Thr Phe Tyr Ala Asp Phe Ser His Asp
290                 295                 300                 305

AAC ACC ATG GTG CCC ATC TTT GCG GCG CTC GGG CTC TTC AAC GCC ACC         1175
Asn Thr Met Val Pro Ile Phe Ala Ala Leu Gly Leu Phe Asn Ala Thr
                310                 315                 320
```

图 6B

```
GCC CTC GAC CCG CTG AAG CCC GAC GAG AAC AGG TTG TGG GTG GAC TCT      1223
Ala Leu Asp Pro Leu Lys Pro Asp Glu Asn Arg Leu Trp Val Asp Ser
            325                 330                 335

AAG CTG GTA CCG TTC TCT GGA CAT ATG ACG GTC GAG AAG CTG GCA TGT      1271
Lys Leu Val Pro Phe Ser Gly His Met Thr Val Glu Lys Leu Ala Cys
        340                 345                 350

TCT GGG AAG GAG GCG GTC AGG GTG CTC GTG AAC GAC GCG GTG CAG CCG      1319
Ser Gly Lys Glu Ala Val Arg Val Leu Val Asn Asp Ala Val Gln Pro
    355                 360                 365

CTG GAG TTC TGC GGA GGT GTT GAT GGG GTG TGC GAG CTT TCG GCT TTC      1367
Leu Glu Phe Cys Gly Gly Val Asp Gly Val Cys Glu Leu Ser Ala Phe
370                 375                 380                 385

GTA GAG AGC CAG ACG TAT GCG CGG GAG AAT GGG CAA GGC GAC TTC GCC      1415
Val Glu Ser Gln Thr Tyr Ala Arg Glu Asn Gly Gln Gly Asp Phe Ala
                390                 395                 400

AAG TGC GGC TTT GTT CCG TCG GAA TAGCGGGAGA CCGTCTATGC TACACAGTAA      1469
Lys Cys Gly Phe Val Pro Ser Glu
            405

TTGTGTACTC TATAGCACTG TAGCTGTACT TACAAGTCGT AGGGTACGAT CGTACTTACG      1529

CTCGTTTATT GATCCTTCCT TTAAAAAAAA AAAAAAAAAA AAAAAAAAAA ATTCCTGCGG      1589

CCGC                                                                  1593
```

图 6C

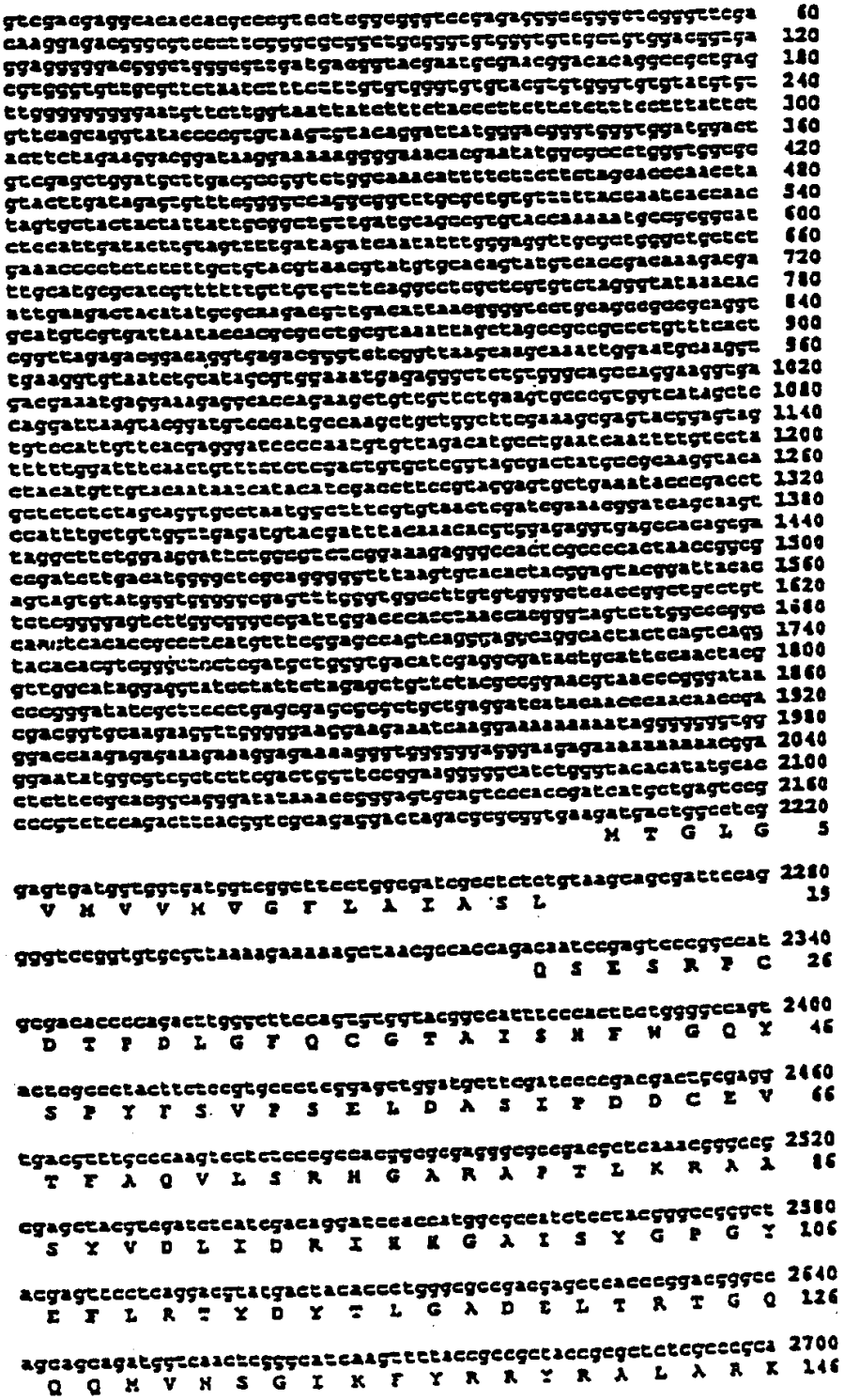


图 7A

```
agtcgatccccttcgtccgcaccgccggccaggaccgcgtcgtccactcggccgagaact 2760
 S  I  P  F  V  R  T  A  G  Q  D  R  V  V  H  S  A  E  N  F  166

tcacccagggcttccactctgccctgctcgccgaccgcgggtccaccgtccggcccaccc 2820
 T  Q  G  F  H  S  A  L  L  A  D  R  G  S  T  V  R  P  T  L  186

tcccctatgacatggtcgtcatcccggaaaccgccggcgccaacaacacgctccacaacg 2880
 P  Y  D  M  V  V  I  P  E  T  A  G  A  N  N  T  L  H  N  D  206

acctctgcaccgccttcgaggaaggcccgtactcgaccatcggcgacgacgcccaagaca 2940
 L  C  T  A  F  E  E  G  P  Y  S  T  I  G  D  D  A  Q  D  T  226

cctacctctccaccttcgccggacccatcaccgcccgggtcaacgccaacctgccgggcg 3000
 Y  L  S  T  F  A  G  P  I  T  A  R  V  N  A  N  L  P  G  A  246

ccaacctgaccgacgccgacacggtcgcgctgatggacctctgccccttcgagacggtcg 3060
 N  L  T  D  A  D  T  V  A  L  M  D  L  C  P  F  E  T  V  A  266

cctcctcctcctccgacccggcaacggcggacgcgggggcggcaacgggcggccgctgt 3120
 S  S  S  S  D  P  A  T  A  D  A  G  G  G  N  G  R  P  L  S  286

cgcccttctgccgcctgttcagcgagtccgagtggcgcgcgtacgactacctgcagtcgg 3180
 P  F  C  R  L  F  S  E  S  E  W  R  A  Y  D  Y  L  Q  S  V  306

tgggcaagtggtacgggtacgggccgggcaacccgctggggccgacgcaggggtcgggt 3240
 G  K  W  Y  G  Y  G  P  G  N  P  L  G  P  T  Q  G  V  G  F  326

tcgtcaacgagctgctggcgcggctggccggggtccccgtgcgcgacggcaccagcacca 3300
 V  N  E  L  L  A  R  L  A  G  V  P  V  R  D  G  T  S  T  N  346

accgcaccctcgacggcgacccgcgcaccttcccgctcggccggcccctctacgccgact 3360
 R  T  L  D  G  D  P  R  T  F  P  L  G  R  P  L  Y  A  D  F  366

tcagccacgacaacgacatgatgggcgtcctcggcgccctcggcgcctacgacggcgtcc 3420
 S  H  D  N  D  M  M  G  V  L  G  A  L  G  A  Y  D  G  V  P  386

cgcccctcgacaagaccgcccgccgcgacccggaagagctcggcgggtacgcggccagct 3480
 P  L  D  K  T  A  R  R  D  P  E  E  L  G  G  Y  A  A  S  W  406

 gggccgtcccgttcgccgccaggatctacgtcgagaagatgcggtgcagcggcggcggcg 3540
 A  V  P  F  A  A  R  I  Y  V  E  K  M  R  C  S  G  G  G  G  426

gcggcggcggcggcggcgaggggcggcaggagaaggatgaggagatggtcagggtgctgg 3600
 G  G  G  G  G  E  G  R  Q  E  K  D  E  E  M  V  R  V  L  V  446

tgaacgaccgggtgatgacgctgaaggggtgcggcgccgacgagaggggatgtgtacgc 3660
 N  D  R  V  M  T  L  K  G  C  G  A  D  E  R  G  M  C  T  L  466

tagaacggttcatcgaaagcatggcgtttgcgagggggaacggcaagtgggatctctgct 3720
 E  R  F  I  E  S  M  A  F  A  R  G  N  G  K  W  D  L  C  F  486

ttgcttgatatgcccacgcccgagattgaacagaacttgtgatgggggtagagtgtggta 3780
 A                                                            487

ttcgagatgatagttcacagttttcgggaatcaaaaatcggttagactggcgaaattcaa 3840
gtctggggcctgcggcgtctgcattctccgttccctgttgttaccttcttaatggttttt 3900
ttttatttttttatttttcttaaattttcacacaaaccttttattgtcttttttttcttctt 3960
tttcttcttctgcacatcggatgggaattgtcgac                           3995
```

图 7B

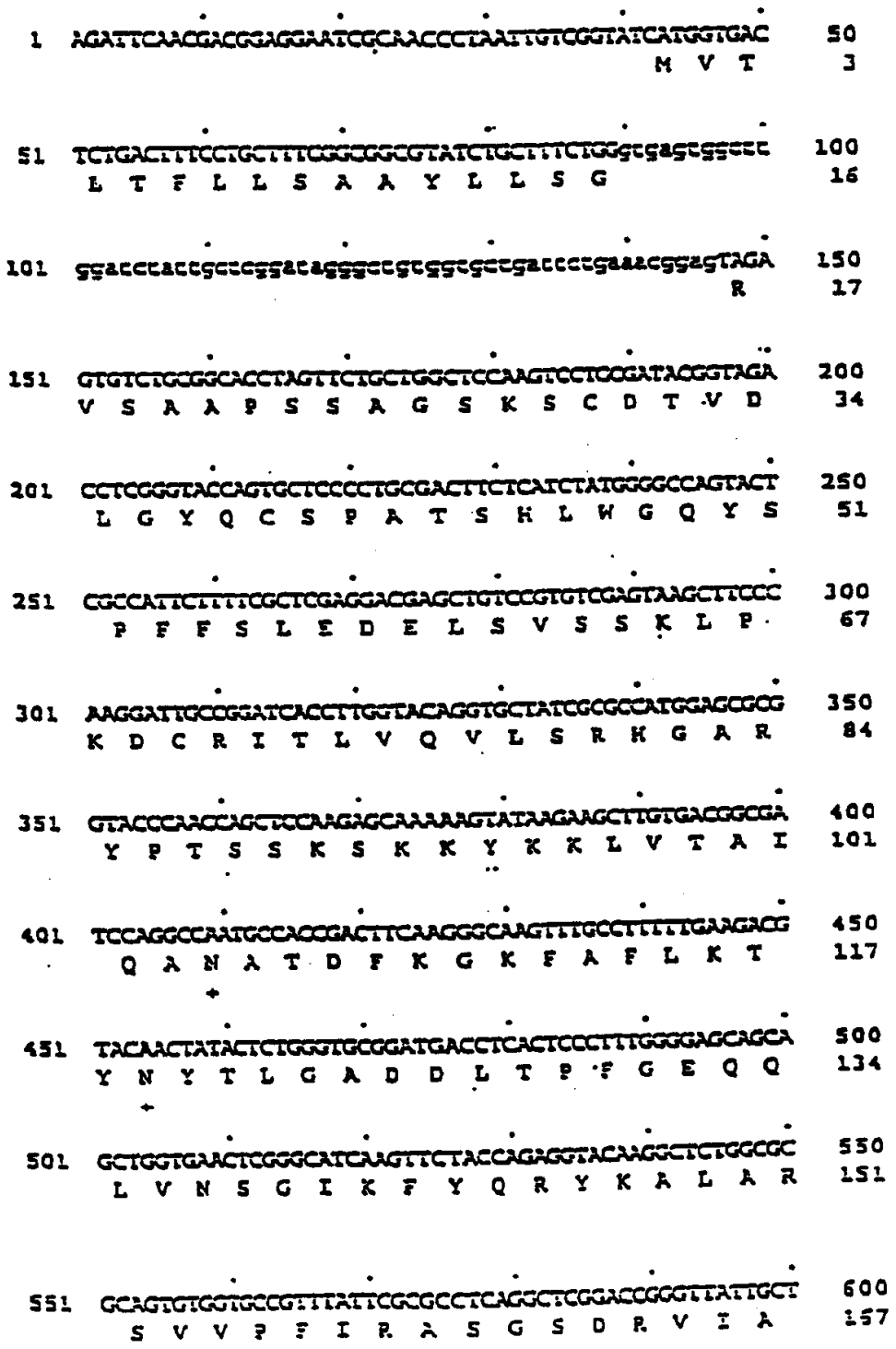

1 AGATTCAACGACGGAGGAATCGCAACCCTAATTGTCGGTATCATGGTGAC 50
M V T 3
51 TCTGACTTTCCTGCTTTCGGGCGGCGTATCTGCTTTCTGGgtgagtggcctc 100
L T F L L S A A Y L L S G 16
101 ggacctactgcctcggatagggcctgtggtgcctgatcctgaaacggagTAGA 150
R 17
151 GTGTCTGCGGCACCTAGTTCTGCTGGCTCCAAGTCCTCCGATACGGTAGA 200
V S A A P S S A G S K S C D T V D 34
201 CCTCGGGTACCAGTGCTCCCCTGCGACTTCTCATCTATGGGGCCAGTACT 250
L G Y Q C S P A T S H L W G Q Y S 51
251 CGCCATTCTTTTCGCTCGAGGACGAGCTGTCCGTGTCGAGTAAGCTTCCC 300
P F F S L E D E L S V S S K L P 67
301 AAGGATTGCCGGATCACCTTGGTACAGGTGCTATCGCGCCATGGAGCGCG 350
K D C R I T L V Q V L S R H G A R 84
351 GTACCCAACCAGCTCCAAGAGCAAAAAGTATAAGAAGCTTGTGACGGCGA 400
Y P T S S K S K K Y K K L V T A I 101
401 TCCAGGCCAATGCCACCGACTTCAAGGGCAAGTTTGCCTTTTTGAAGACG 450
Q A N A T D F K G K F A F L K T 117
451 TACAACTATACTCTGGGTGCGGATGACCTCACTCCCTTTGGGGAGCAGCA 500
Y N Y T L G A D D L T P F G E Q Q 134
501 GCTGGTGAACTCGGGCATCAAGTTCTACCAGAGGTACAAGGCTCTGGCGC 550
L V N S G I K F Y Q R Y K A L A R 151
551 GCAGTGTGGTGCCGTTTATTCGCGCCTCAGGCTCGGACCGGGTTATTGCT 600
S V V P F I R A S G S D R V I A 167

图 8A

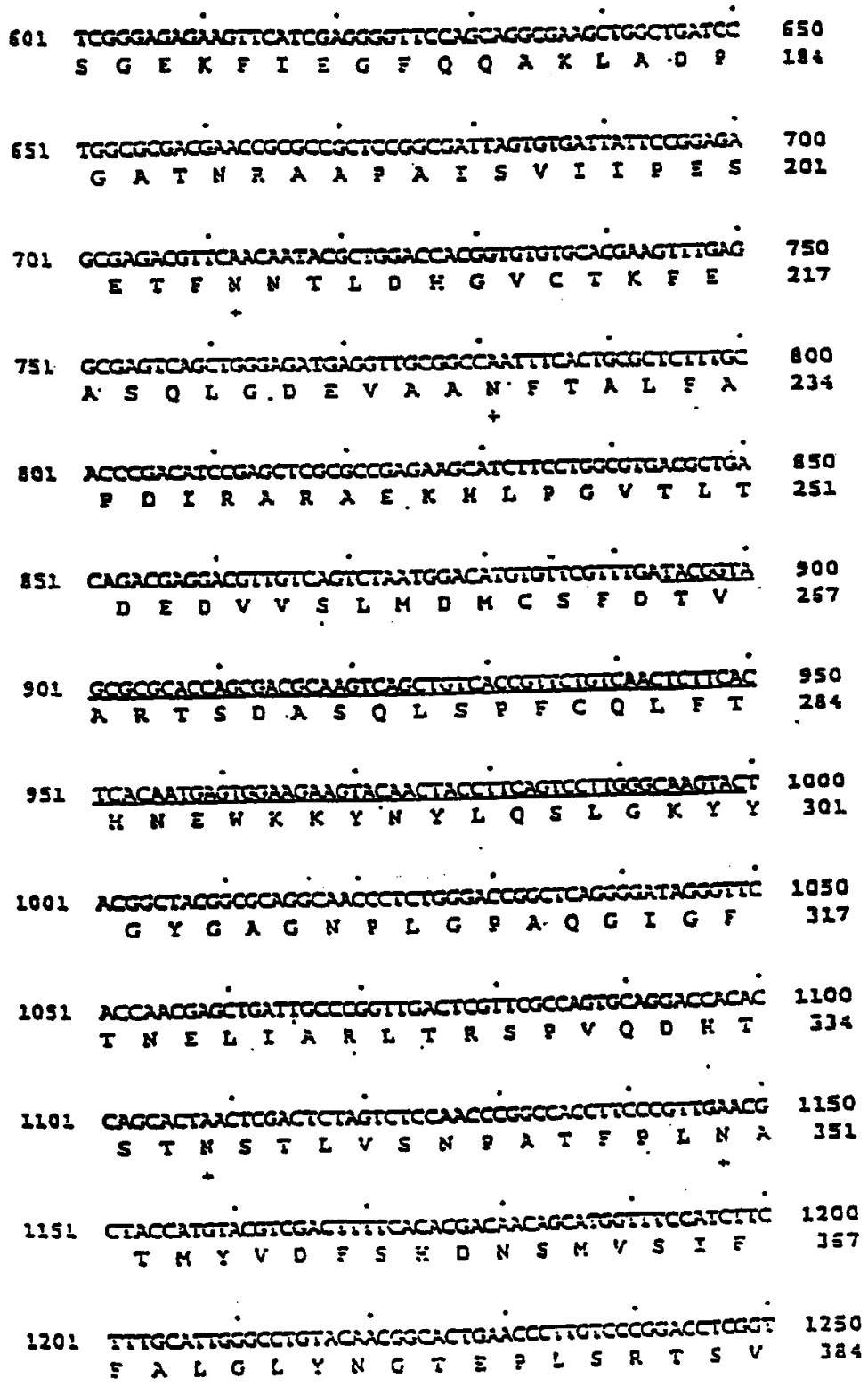

601 650
S G E K F I E G F Q Q A K L A D P 184
651 700
G A T N R A A P A I S V I I P E S 201
701 750
E T F N N T L D H G V C T K F E 217
751 800
A S Q L G D E V A A N F T A L F A 234
801 850
P D I R A R A E K H L P G V T L T 251
851 900
D E D V V S L M D M C S F D T V 267
901 950
A R T S D A S Q L S P F C Q L F T 284
951 1000
H N E W K K Y N Y L Q S L G K Y Y 301
1001 1050
G Y G A G N P L G P A Q G I G F 317
1051 1100
T N E L I A R L T R S P V Q D H T 334
1101 1150
S T N S T L V S N P A T F P L N A 351
1151 1200
T M Y V D F S H D N S M V S I F 367
1201 1250
F A L G L Y N G T E P L S R T S V 384

图 8B

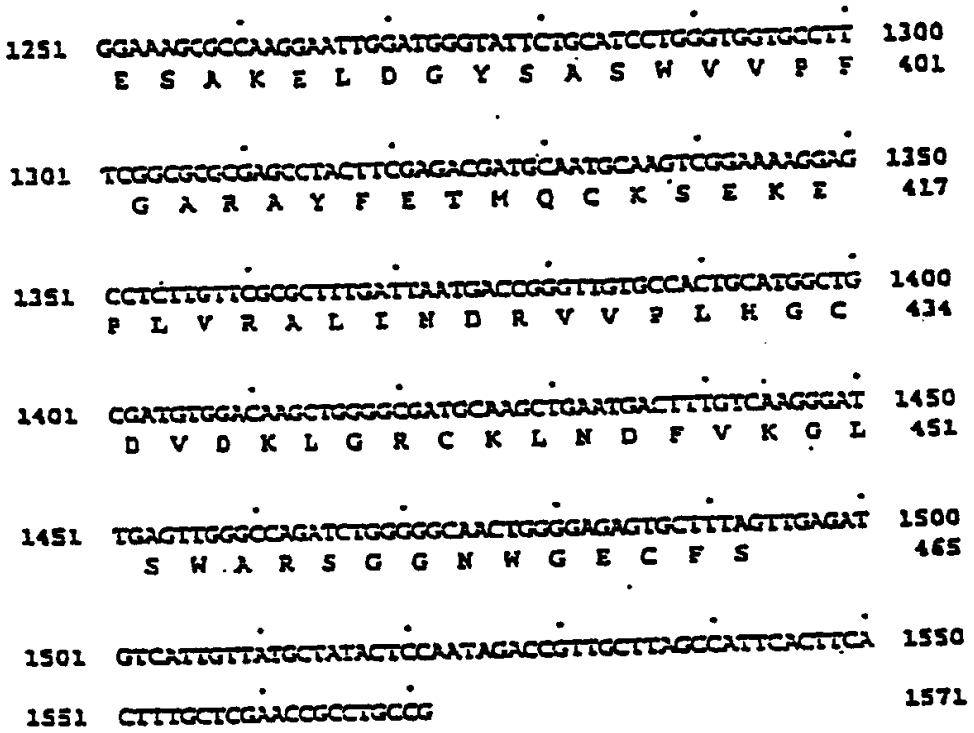

1251 GGAAAGCGCCAAGGAATTGGATGGGTATTCTGCATCCTGGGTGGTGCCTT 1300
E S A K E L D G Y S A S W V V P F 401
1301 TCGGCGCCGAGCCTACTTCGAGACGATGCAATGCAAGTCGGAAAGGAG 1350
G A R A Y F E T M Q C K S E K E 417
1351 CCTCTTGTTCGGCGTTTGATTAATGACCGGGTTGTGCCACTGCATGGCTG 1400
P L V R A L I N D R V V P L H G C 434
1401 CGATGTGGACAAGCTGGGGCGATGCAAGCTGAATGACTTTGTCAAGGGAT 1450
D V D K L G R C K L N D F V K G L 451
1451 TGAGTTGGGCCAGATCTGGGGGCAACTGGGGAGAGTGCTTTAGTTGAGAT 1500
S W A R S G G N W G E C F S 465
1501 GTCATTGTTATGCTATACTCCAATAGACCGTTGCTTAGCCATTCACTTCA 1550
1551 CTTTGCTCGAACCGCCTGCCG 1571

图 8C

```
                                                
```

```
                                                                                                  200
                                151
A. 土曲霉 9A-1                  NALTRhInPF VRATDASRVh ESAEKFVEGF QTARqDDHhA nPHQPSPrVd
A. 土曲霉 cbs                   DTLTRhInPF VRAADSSRVh ESAEKFVEGF QNARqGDPhA nPHQPSPrVd
A. 黑曲霉变种泡盛曲霉          ESLTRNIIPF IRSSGSSRVI ASGEKFIEGF QSTKLkDPrA qPgQSSPkId
A. 黑曲霉 T213                 ESLTRNIIPF IRSSGSSRVI ASGEKFIEGF QSTKLkDPrA qPgQSSPkId
A. 黑曲霉 NRRL3135             ESLTRNIVPF IRSSGSSRVI ASGKKFIEGF QSTKLkDPrA qPgQSSPkId
A. 烟曲霉 13073                KALARSVVPF IRASGSDRVI ASGEKFIEGF QqAKLADPGA .TNRAAPAIs
A. 烟曲霉 32722                KALARSVVPF IRASGSDRVI ASGEKFIEGF QqAKLADPGA .TNRAAPAIs
A. 烟曲霉 58128                KALARSVVPF IRASGSDRVI ASGEKFIEGF QqAKLADPGA .TNRAAPAIs
A. 烟曲霉 26906                KALARSVVPF IRASGSDRVI ASGEKFIEGF QqAKLADPGA .TNRAAPAIs
A. 烟曲霉 32239                KALAqSVVPF IRSSGSDRVI ASGEKFIEGF QqANVADPGA .TNRAAPVIs
A. 构巢曲霉                    KNLARKnTPF IRASGSDRVV ASAEKFINGF RKAQLhDHGS ..gQATPVVn
T. 嗜热踩节菌                  KSLARNaVPF VRCSGSDRVI ASGrlFIEGF QSAKVlDPhS dKHDAPPTIn
M. 嗜热毁丝菌                  RALARKsIPF VRTAGqDRVV hSAENFTQGF HSALLADRGS tVRPTLPydm

共有序列                        KALARKIVPF IRASGSDRVI ASAEKFIEGF QSAKLADPGS -PHQASPVI-
共有植酸酶                      KALARKIVPF IRASGSDRVI ASAEKFIEGF QSAKLADPGS QPHQASPVID

                                                                                                  250
                                201
A. 土曲霉 9A-1                  VaIPEGSAYN NTLEHSlCTA FES...StVG DDAvANFTAV FAPAIaQRLE
A. 土曲霉 cbs                   VVIPEGTAYN NTLEHSICTA FEA...StVG DAAaDNFTAV FAPAIakRLE
A. 黑曲霉变种泡盛曲霉          VVISEASSsN NTLDPGTCTV FED...SELA DTVEANFTAT FAPSIRQRLE
A. 黑曲霉 T213                 VVISEASSsN NTLDPGTCTV FED...SELA DTVEANFTAT FAPSIRQRLE
A. 黑曲霉 NRRL3135             VVISEASSsN NTLDPGTCTV FED...SELA DTVEANFTAT FVPSIRQRLE
A. 烟曲霉 13073                VIIPESETFN NTLDHGVCTk FEA...SQLG DEVaANFTAl FAPDIRARaE
A. 烟曲霉 32722                VIIPESETFN NTLDHGVCTk FEA...SQLG DEVaANFTAl FAPDIRARaE
A. 烟曲霉 58128                VIIPESETFN NTLDHGVCTk FEA...SQLG DEVaANFTAl FAPDIRARaK
A. 烟曲霉 26906                VIIPESETFN NTLDHGVCTk FEA...SQLG DEVaANFTAl FAPAIRARIE
A. 烟曲霉 32239                VIIPESETYN NTLDHSVCTN FEA...SELG DEVEANFTAl FAPAIRARIE
A. 构巢曲霉                    VIIPElDGFN NTLDHSTCVS FEN...DErA DEiEANFTAI MGPPIRkRLE
T. 嗜热踩节菌                  VIIeEGPSYN NTLDtGSCPV FED...SSqG HDAQEKFAkq FAPAILEKIK
M. 嗜热毁丝菌                  VVIPETAGaN NTLHNOlCTA FEEqpyStIG ODAQDTYlST FAGPItARVN

共有序列                        VIIPEGSGYN NTLDHGTCTA FED---SELG DDAEANFTAT FAPAIRARLE
共有植酸酶                      VIIPEGSGYN NTLDHGTCTA FED...SELG DDVEANFTAL FAPAIRARLE

                                                                                                  300
                                251
A. 土曲霉 9A-1                  ADLPGVqLST DDVVnLMAMC PFETVSlTD. .......... DAhTLSPFCD
A. 土曲霉 cbs                   ADLPGVqLSA DDVVnLMAMC PFETVSlTD. .......... DAhTLSPFCD
A. 黑曲霉变种泡盛曲霉          NDLSGVTLTD TEVTyLMDMC SFDTIStST. .......... vDTKLSPFCD
A. 黑曲霉 T213                 NDLSGVTLTD TEVTyLMDMC SFDTIStST. .......... vDTKLSPFCD
A. 黑曲霉 NRRL3135             NDLSGVTLTD TEVTyLMDMC SFDTIStST. .......... vDTKLSPFCD
A. 烟曲霉 13073                kHLPGVTLTD EDVVsLMDMC SFDTVARTS. .......... DASQLSPFCQ
A. 烟曲霉 32722                kHLPGVTLTD EDVVsLMDMC SFDTVARTS. .......... DASQLSPFCQ
A. 烟曲霉 58128                kHLPGVTLTD EDVVsLMDMC SFDTVARTS. .......... DASQLSPFCQ
A. 烟曲霉 26906                kHLPGVTLTD EDVVsLMDMC SFDTVARTS. .......... DASQLSPFCQ
A. 烟曲霉 32239                kHLPGVqLTD DDVVsLMDMC SFDTVARTA. .......... DASELSPFCA
A. 构巢曲霉                    NDLPGIKLTN ENVIyLMDMC SFDTMARTA. .......... HGTELSPFCA
T. 嗜热踩节菌                  DHLPGVDLAv SDVpyLMDLC PFETLARNh. .......... TDT.LSPFCA
M. 嗜热毁丝菌                  ANLPGANLTD ADTVaLMDLC PFETVAsSSs dpatadaggg NGrpLSPFCr

共有序列                        ADLPGVTLTD EDVV-LMDMC PFETVARTS- ---------- DATELSPFCA
共有植酸酶                      ADLPGVTLTD EDVVYLMDMC PFETVARTS. .......... DATELSPFCA
```

图 9B

```
                                  301                                                                350
A. 土曲霉 9A-1                    LFTacEWtQY NYLLSLDKYY GYGGGNPLGP VQGVGWaNEL MARLTRAPVH
A. 土曲霉 cbs                     LFTaaEWtQY NYLLSLDKYY GYGGGNPLGP VQGVGWaNEL IARLTRSPVH
A. 黑曲霉变种泡盛曲霉            LFTHdEWiHY DYLQSLkKYY GHGAGNPLGP TQGVGYaNEL IARLTHSPVH
A. 黑曲霉 T213                   LFTHdEWiHY DYLRSLkKYY GHGAGNPLGP TQGVGYaNEL IARLTHSPVH
A. 黑曲霉 NRRL3135               LFTHdEWiNY DYLQSLkKYY GHGAGNPLGP TQGVGYaNEL IARLTHSPVH
A. 烟曲霉 13073                  LFTHnEWkKY NYLQSLGKYY GYGAGNPLGP AQGIGFtNEL IARLTRSPVQ
A. 烟曲霉 32722                  LFTHnEWkKY NYLQSLGKYY GYGAGNPLGP AQGIGFtNEL IARLTRSPVQ
A. 烟曲霉 58128                  LFTHnEWkKY NYLQSLGKYY GYGAGNPLGP AQGIGFtNEL IARLTRSPVQ
A. 烟曲霉 26906                  LFTHnEWkKY NYLQSLGKYY GYGAGNPLGP AQGIGFtNEL IARLTRSPVQ
A. 烟曲霉 32239                  IFTHnEWkKY DYLQSLGKYY GYGAGNPLGP AQGIGFtNEL IARLTnSPVQ
A. 构巢曲霉                      IFTEkEWlQY DYLQSLSKYY GYGAGSPLGP AQGIGFtNEL IARLTQSPVQ
T. 嗜热踝节菌                    LaTQaEWqaY DYYQSLGKYY GnGGGNPLGP AQGVGFvNEL IARMTHSPVQ
M. 嗜热毁丝菌                    LFSEaEWraY DYLQSVGKNY GYGPGNPLGP TQGVGFvNEL LARLAqwPVR

共有序列                         LFTH-EW-QY DYLQSLGKYY GYGAGNPLGP AQGVGF-NEL IARLTRSPVQ
共有植酸酶                       LFTHDEWRQY DYLQSLGKYY GYGAGNPLGP AQGVGFANEL IARLTRSPVQ

                                  351                                                                400
A. 土曲霉 9A-1                    DHTCVNNTLD ASPATFPLNA TLYADFSHDS NLVSIFWALG LYNGTAPLSq
A. 土曲霉 cbs                     DHTCVNNTLD ANPATFPLNA TLYADFSHDS NLVSIFWALG LYNGTkPLSq
A. 黑曲霉变种泡盛曲霉            DDTSSNHTLD SNPATFPLNS TLYADFSHDN GIISILFALG LYNGTkPLST
A. 黑曲霉 T213                   DDTSSNHTLD SNPATFPLNS TLYADFSHDN GIISILFALG LYNGTkPLST
A. 黑曲霉 NRRL3135               DDTSSNHTLD SSPATFPLNS TLYADFSHDN GIISILFALG LYNGTkPLST
A. 烟曲霉 13073                  DHTSTNsTLv SNPATFPLNA TMYVDFSHDN SMVSIFFALG LYNGTEPLSr
A. 烟曲霉 32722                  DHTSTNsTLv SNPATFPLNA TMYVDFSHDN SMVSIFFALG LYNGTGPLSr
A. 烟曲霉 58128                  DHTSTNsTLv SNPATFPLNA TMYVDFSHDN SMVSIFFALG LYNGTEPLSr
A. 烟曲霉 26906                  DHTSTNsTLv SNPATFPLNA TMYVDFSHDN SMVSIFFALG LYNGTEPLSr
A. 烟曲霉 32239                  DHTSTNsTLD SDPATFPLNA TIYVDFSHDN GMIPIFFAMG LYNGTEPLSq
A. 构巢曲霉                      DNTSTNHTLD SNPATFPLDr KLYADFSHDN SMISIFFAMG LYNGTQPLSm
T. 嗜热踝节菌                    DYTTVNHTLD SNPATFPLNA TLYADFSHDN TMTSIFaALG LYNGTAkLST
M. 嗜热毁丝菌                    DqTSTNRTLD GDPrTFPLGr PLYADFSHDN DMMGVLgALG aYDGVPPLDK

共有序列                         DHTSTNHTLD SNPATFPLNA TLYADFSHDN SMISIFFALG LYNGTAPLST
共有植酸酶                       DHTSTNHTLD SNPATFPLNA TLYADFSHDN SMISIFFALG LYNGTAPLST

                                  401                                                                450
A. 土曲霉 9A-1                    TSVESVSQTD GYAAANTVPF AARAYVEMMQ C..........  ....RAEKEP
A. 土曲霉 cbs                     TTVEDITrTD GYAAANTVPF AARAYIEMMQ C..........  ....RAEKQP
A. 黑曲霉变种泡盛曲霉            TTVENITQTD GFSSANTVPF ASRIYVEMMQ C..........  ....QAEQEP
A. 黑曲霉 T213                   TTVENITQTD GFSSANTVPF ASRIYVEMMQ C..........  ....QAEQEP
A. 黑曲霉 NRRL3135               TTVENITQTD GFSSANTVPF ASRIYVEMMQ C..........  ....QAEQEP
A. 烟曲霉 13073                  TSVESaKElD GYSASWVVPF GARAYFEtMQ C..........  ....KSEKEP
A. 烟曲霉 32722                  TSVESaKElD GYSASWVVPF GARAYFEtMQ C..........  ....KSEKEP
A. 烟曲霉 58128                  TSVESaKElD GYSASWVVPF GARAYFEtMQ C..........  ....KSEKES
A. 烟曲霉 26906                  TSVESaKElD GYSASWVVPF GARAYFEtMQ C..........  ....KSEKEP
A. 烟曲霉 32239                  TSeESTKESN GYSASWAVPF GARAYFEtMQ C..........  ....KSEKEP
A. 构巢曲霉                      DSVESIQEmD GYAASWTVPF GARAYFELMQ C..........  ....DDSDEP
T. 嗜热踝节菌                    TEIKSIEETD GYSAAWTVPF GGRAYIEMMQ C..........  ....DDSDEP
M. 嗜热毁丝菌                    TArrDpEElG GYAASWAVPF AARiYVEKMR Csggggggggg  gegrQEKDEe

共有序列                         TSVESIEETD GYSASWTVPF GARAYVEMMQ C----------  ----QAEKEP
共有植酸酶                       TSVESIEETD GYSASWTVPF GARAYVEMMQ C..........  ....QAEKEP
```

图 9C

00.11.20

	451	500
A. 土曲霉 9A-1	LVRVLVNDRV MPLHGCPTDK LGRCK=DAFV AGLSFAQAGG NWADCF----	
A. 土曲霉 cbs	LVRVLVNDRV MPLHGCAVDN LGRCK=DDFV EGLSFARAGG NWAECF----	
A. 黑曲霉变种泡盛曲霉	LVRVLVNDRV VPLHGCPIDa LGRCT=DSFV rGLSFARSGG DWAECsA----	
A. 黑曲霉 T213	LVRVLVNDRV VPLHGCPIDa LGRCT=DSFV rGLSFARSGG DWAECFA----	
A. 黑曲霉 NRRL3135	LVRVLVNDRV VPLHGCPVDa LGRCT=DSFV rGLSFARSGG DWAECFA----	
A. 烟曲霉 13073	LVRALINDRV VPLHGCDVDK LGRCKLNDFV KGLSWARSGG NWGECFS----	
A. 烟曲霉 32722	LVRALINDRV VPLHGCDVDK LGRCKLNDFV KGLSWARSGG NWGECFS----	
A. 烟曲霉 58128	LVRALINDRV VPLHGCDVDK LGRCKLNDFV KGLSWARSGG NWGECFS----	
A. 烟曲霉 26906	LVRALINDRV VPLHGCDVDK LGRCKLNDFV KGLSWARSGG NWGECFS----	
A. 烟曲霉 32239	LVRALINDRV VPLHGCAVDK LGRCKLKDFV KGLSWARSGG NSEQSFS----	
A. 构巢曲霉	LVRVLVNDRV VPLHGCAVDK rGRCTLDDWV EGLNFARSGG NWkTCFTl--	
T. 嗜热踝节菌	VVRVLVNDRV VPLHGCEVDS LGRCK=DDFV rGLSFARqGG NWEGCYAase	
M. 嗜热毁丝菌	MVRVLVNDRV MTLkGCGADE rGMCTLErFI ESMAFARGNG KWDlCFA----	
共有序列	LVRVLVNDRV VPLHGCAVDK LGRCK-DDFV EGLSFARSGG NWAECFA----	
共有植酸酶	LVRVLVNDRV VPLHGCAVDK LGRCKDDDFV EGLSFARSGG NWAECFA	

图 9D

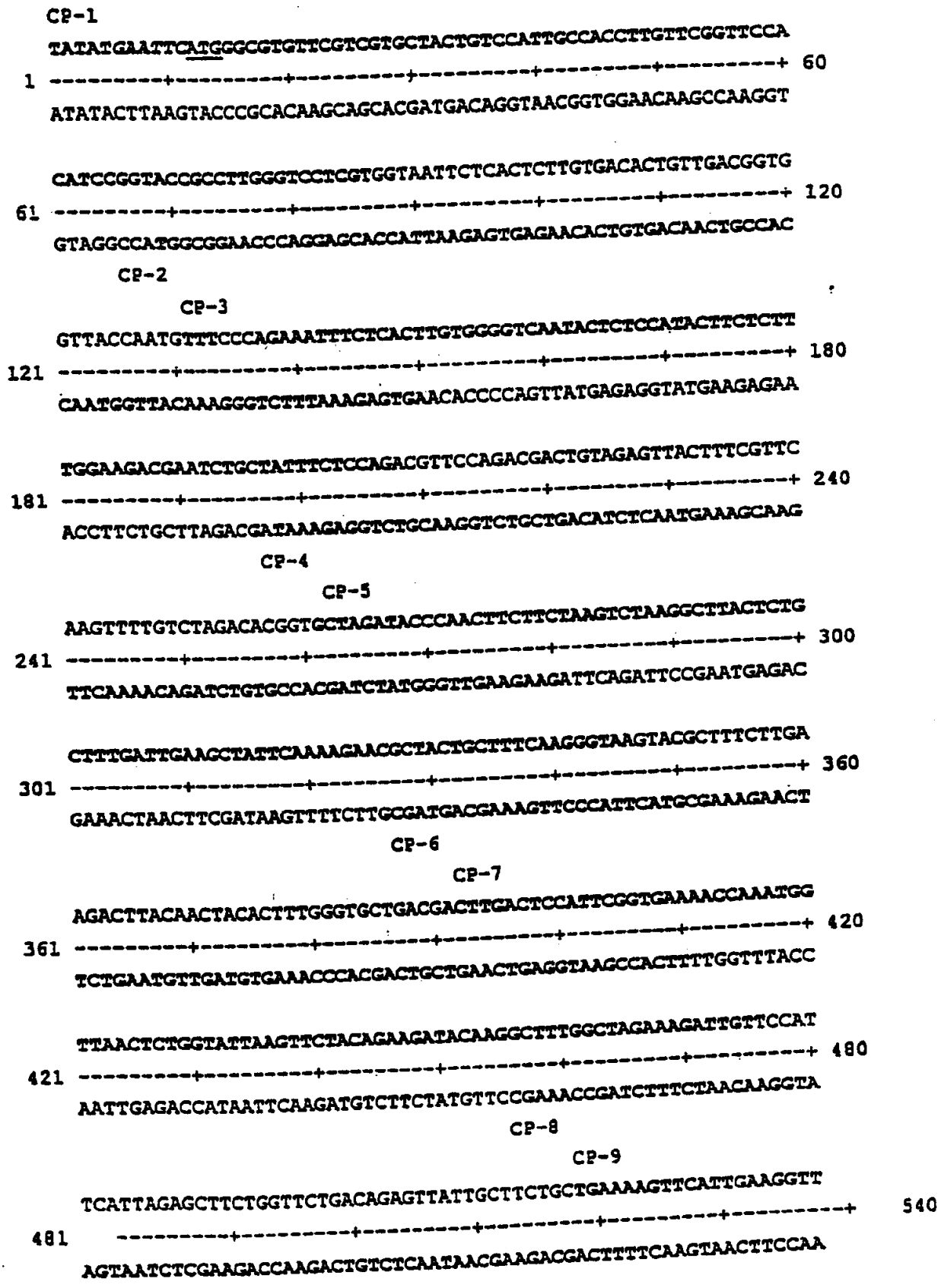

CP-1
TATATGAATTCATGGGCGTGTTCGTCGTGCTACTGTCCATTGCCACCTTGTTCGGTTCCA
1 ---------+---------+---------+---------+---------+---------+ 60
ATATACTTAAGTACCCGCACAAGCAGCACGATGACAGGTAACGGTGGAACAAGCCAAGGT
CATCCGGTACCGCCTTGGGTCCTCGTGGTAATTCTCACTCTTGTGACACTGTTGACGGTG
61 ---------+---------+---------+---------+---------+---------+ 120
GTAGGCCATGGCGGAACCCAGGAGCACCATTAAGAGTGAGAACACTGTGACAACTGCCAC
CP-2
CP-3
GTTACCAATGTTTCCCAGAAATTTCTCACTTGTGGGGTCAATACTCTCCATACTTCTCTT
121 ---------+---------+---------+---------+---------+---------+ 180
CAATGGTTACAAAGGGTCTTTAAAGAGTGAACACCCCAGTTATGAGAGGTATGAAGAGAA
TGGAAGACGAATCTGCTATTTCTCCAGACGTTCCAGACGACTGTAGAGTTACTTTCGTTC
181 ---------+---------+---------+---------+---------+---------+ 240
ACCTTCTGCTTAGACGATAAAGAGGTCTGCAAGGTCTGCTGACATCTCAATGAAAGCAAG
CP-4
CP-5
AAGTTTTGTCTAGACACGGTGCTAGATACCCAACTTCTTCTAAGTCTAAGGCTTACTCTG
241 ---------+---------+---------+---------+---------+---------+ 300
TTCAAAACAGATCTGTGCCACGATCTATGGGTTGAAGAAGATTCAGATTCCGAATGAGAC
CTTTGATTGAAGCTATTCAAAAGAACGCTACTGCTTTCAAGGGTAAGTACGCTTTCTTGA
301 ---------+---------+---------+---------+---------+---------+ 360
GAAACTAACTTCGATAAGTTTTCTTGCGATGACGAAAGTTCCCATTCATGCGAAAGAACT
CP-6
CP-7
AGACTTACAACTACACTTTGGGTGCTGACGACTTGACTCCATTCGGTGAAAACCAAATGG
361 ---------+---------+---------+---------+---------+---------+ 420
TCTGAATGTTGATGTGAAACCCACGACTGCTGAACTGAGGTAAGCCACTTTTGGTTTACC
TTAACTCTGGTATTAAGTTCTACAGAAGATACAAGGCTTTGGCTAGAAAGATTGTTCCAT
421 ---------+---------+---------+---------+---------+---------+ 480
AATTGAGACCATAATTCAAGATGTCTTCTATGTTCCGAAACCGATCTTTCTAACAAGGTA
CP-8
CP-9
TCATTAGAGCTTCTGGTTCTGACAGAGTTATTGCTTCTGCTGAAAAGTTCATTGAAGGTT
481 ---------+---------+---------+---------+---------+---------+ 540
AGTAATCTCGAAGACCAAGACTGTCTCAATAACGAAGACGACTTTTCAAGTAACTTCCAA

图 9E

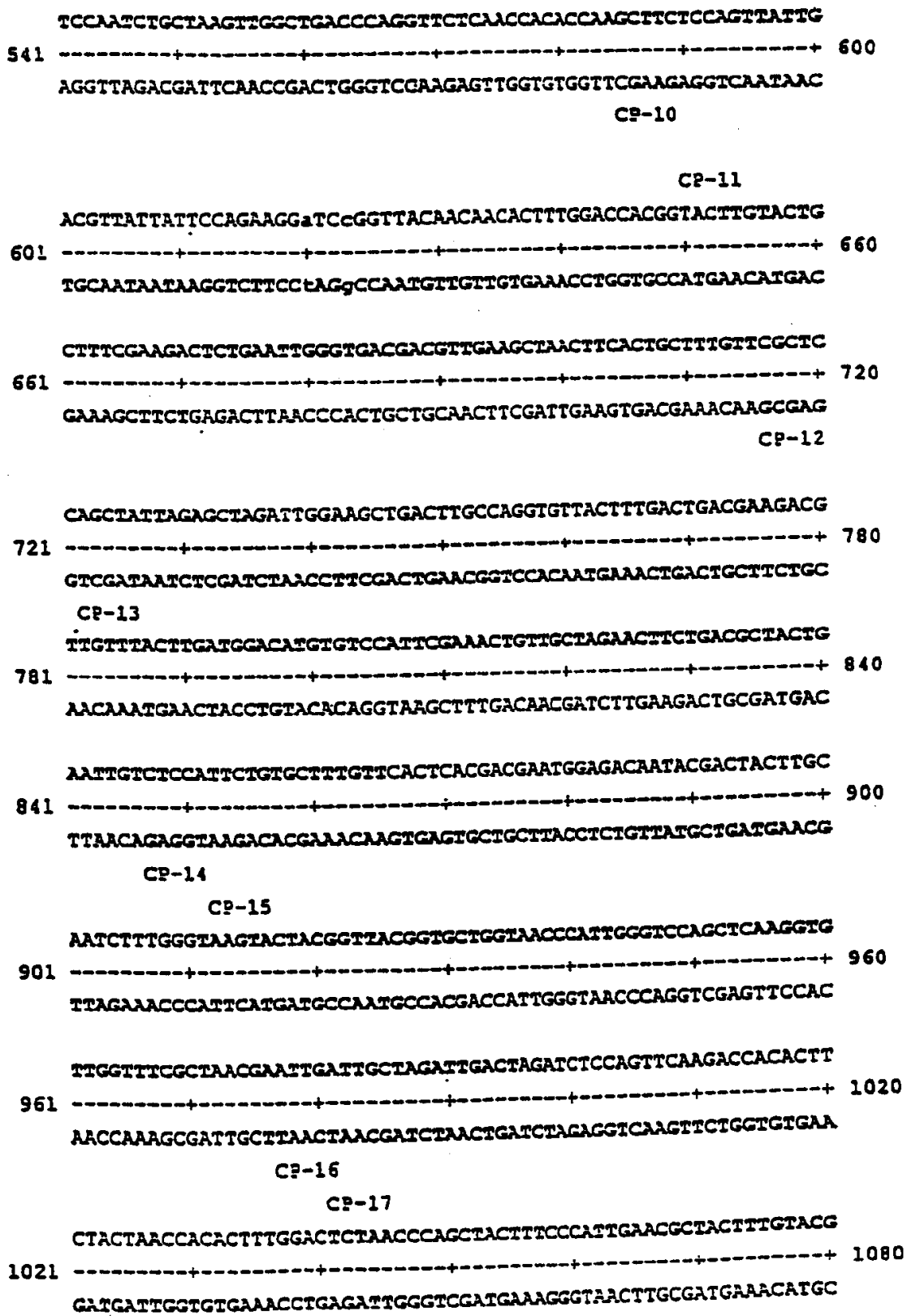

TCCAATCTGCTAAGTTGGCTGACCCAGGTTCTCAACCACACCAAGCTTCTCCAGTTATTG
541 ---------+---------+---------+---------+---------+---------+ 600
AGGTTAGACGATTCAACCGACTGGGTCCAAGAGTTGGTGTGGTTCGAAGAGGTCAATAAC
CP-10
CP-11
ACGTTATTATTCCAGAAGGaTCcGGTTACAACAACACTTTGGACCACGGTACTTGTACTG
601 ---------+---------+---------+---------+---------+---------+ 660
TGCAATAATAAGGTCTTCCtAGgCCAATGTTGTTGTGAAACCTGGTGCCATGAACATGAC
CTTTCGAAGACTCTGAATTGGGTGACGACGTTGAAGCTAACTTCACTGCTTTGTTCGCTC
661 ---------+---------+---------+---------+---------+---------+ 720
GAAAGCTTCTGAGACTTAACCCACTGCTGCAACTTCGATTGAAGTGACGAAACAAGCGAG
CP-12
CAGCTATTAGAGCTAGATTGGAAGCTGACTTGCCAGGTGTTACTTTGACTGACGAAGACG
721 ---------+---------+---------+---------+---------+---------+ 780
GTCGATAATCTCGATCTAACCTTCGACTGAACGGTCCACAATGAAACTGACTGCTTCTGC
CP-13
TTGTTTACTTGATGGACATGTGTCCATTCGAAACTGTTGCTAGAACTTCTGACGCTACTG
781 ---------+---------+---------+---------+---------+---------+ 840
AACAAATGAACTACCTGTACACAGGTAAGCTTTGACAACGATCTTGAAGACTGCGATGAC
AATTGTCTCCATTCTGTGCTTTGTTCACTCACGACGAATGGAGACAATACGACTACTTGC
841 ---------+---------+---------+---------+---------+---------+ 900
TTAACAGAGGTAAGACACGAAACAAGTGAGTGCTGCTTACCTCTGTTATGCTGATGAACG
CP-14
CP-15
AATCTTTGGGTAAGTACTACGGTTACGGTGCTGGTAACCCATTGGGTCCAGCTCAAGGTG
901 ---------+---------+---------+---------+---------+---------+ 960
TTAGAAACCCATTCATGATGCCAATGCCACGACCATTGGGTAACCCAGGTCGAGTTCCAC
TTGGTTTCGCTAACGAATTGATTGCTAGATTGACTAGATCTCCAGTTCAAGACCACACTT
961 ---------+---------+---------+---------+---------+---------+ 1020
AACCAAAGCGATTGCTTAACTAACGATCTAACTGATCTAGAGGTCAAGTTCTGGTGTGAA
CP-16
CP-17
CTACTAACCACACTTTGGACTCTAACCCAGCTACTTTCCCATTGAACGCTACTTTGTACG
1021 ---------+---------+---------+---------+---------+---------+ 1080
GATGATTGGTGTGAAACCTGAGATTGGGTCGATGAAAGGGTAACTTGCGATGAAACATGC

图 9F

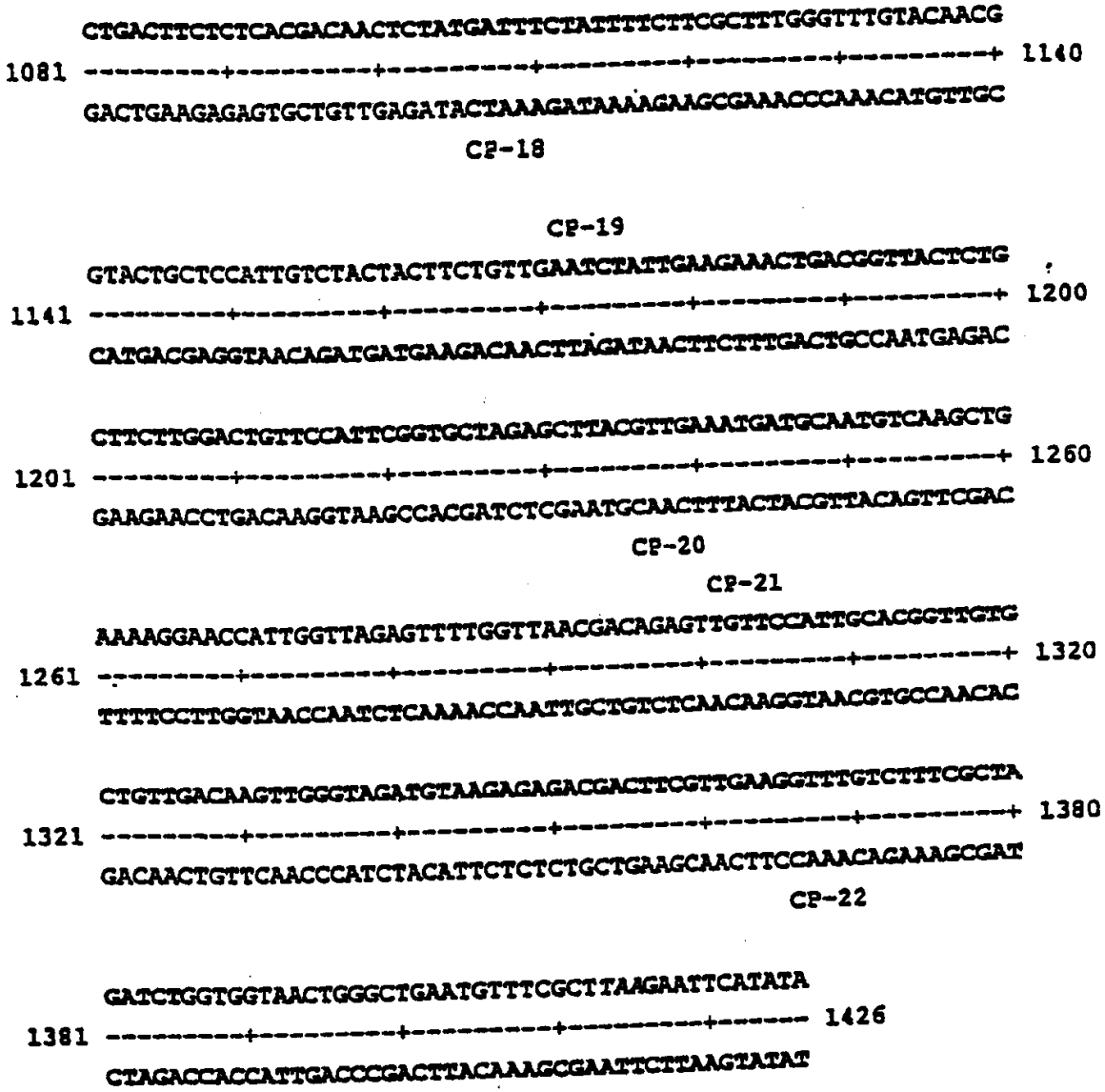

CTGACTTCTCTCACGACAACTCTATGATTTCTATTTTCTTCGCTTTGGGTTTGTACAACG
1081 ---------+---------+---------+---------+---------+---------+ 1140
GACTGAAGAGAGTGCTGTTGAGATACTAAAGATAAAAGAAGCGAAACCCAAACATGTTGC
CP-18
CP-19
GTACTGCTCCATTGTCTACTACTTCTGTTGAATCTATTGAAGAAACTGACGGTTACTCTG
1141 ---------+---------+---------+---------+---------+---------+ 1200
CATGACGAGGTAACAGATGATGAAGACAACTTAGATAACTTCTTTGACTGCCAATGAGAC
CTTCTTGGACTGTTCCATTCGGTGCTAGAGCTTACGTTGAAATGATGCAATGTCAAGCTG
1201 ---------+---------+---------+---------+---------+---------+ 1260
GAAGAACCTGACAAGGTAAGCCACGATCTCGAATGCAACTTTACTACGTTACAGTTCGAC
CP-20
CP-21
AAAAGGAACCATTGGTTAGAGTTTTGGTTAACGACAGAGTTGTTCCATTGCACGGTTGTG
1261 ---------+---------+---------+---------+---------+---------+ 1320
TTTTCCTTGGTAACCAATCTCAAAACCAATTGCTGTCTCAACAAGGTAACGTGCCAACAC
CTGTTGACAAGTTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTTCGCTA
1321 ---------+---------+---------+---------+---------+---------+ 1380
GACAACTGTTCAACCCATCTACATTCTCTCTGCTGAAGCAACTTCCAAACAGAAAGCGAT
CP-22
GATCTGGTGGTAACTGGGCTGAATGTTTCGCTTAAGAATTCATATA
1381 ---------+---------+---------+---------+------ 1426
CTAGACCACCATTGACCCGACTTACAAAGCGAATTCTTAAGTATAT

图 9G

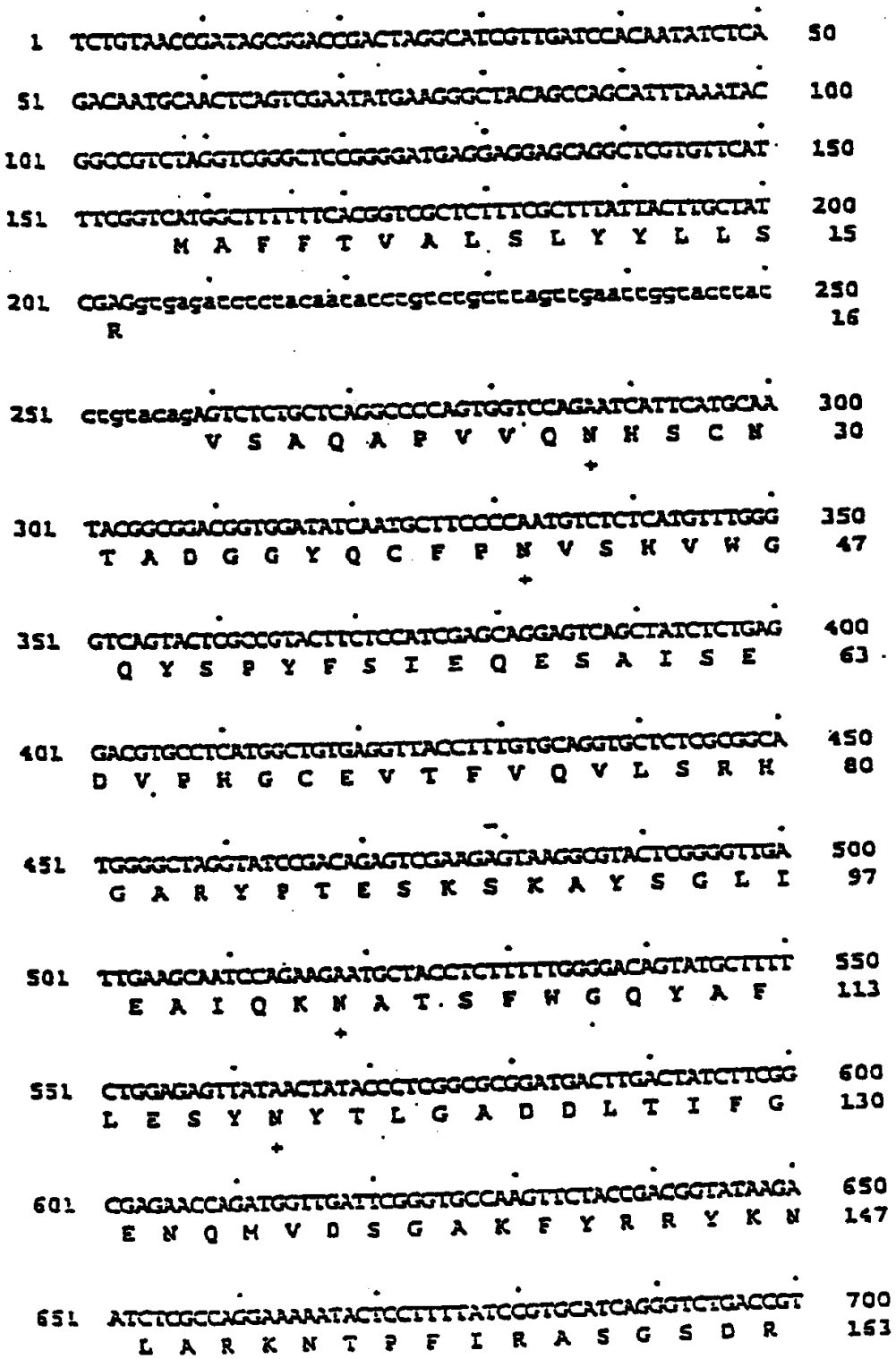

1 TCTGTAACCGATAGCGGACCGACTAGGCATCGTTGATCCACAATATCTCA 50
51 GACAATGCAACTCAGTCGAATATGAAGGGCTACAGCCAGCATTTAAATAC 100
101 GGCCGTCTAGGTCGGGCTCCGGGGATGAGGAGGAGCAGGCTCGTGTTCAT 150
151 TTCGGTCATGGCTTTTTTCACGGTCGCTCTTTCGCTTTATTACTTGCTAT 200
M A F F T V A L S L Y Y L L S 15
201 CGAGgtgagatccctacaatacctgtctgcttagctgaatcggtaccttac 250
R 16
251 ctgtacagAGTCTCTGCTCAGGCCCCAGTGGTCCAGAATCATTCATGCAA 300
V S A Q A P V V Q N H S C N 30
301 TACGGCGGACGGTGGATATCAATGCTTCCCCAATGTCTCTCATGTTTGGG 350
T A D G G Y Q C F P N V S H V W G 47
351 GTCAGTACTCGCCGTACTTCTCCATCGAGCAGGAGTCAGCTATCTCTGAG 400
Q Y S P Y F S I E Q E S A I S E 63
401 GACGTGCCTCATGGCTGTGAGGTTACCTTTGTGCAGGTGCTCTCGCGGCA 450
D V P H G C E V T F V Q V L S R H 80
451 TGGGGCTAGGTATCCGACAGAGTCGAAGAGTAAGGCGTACTCGGGGTTGA 500
G A R Y P T E S K S K A Y S G L I 97
501 TTGAAGCAATCCAGAAGAATGCTACCTCTTTTTGGGGACAGTATGCTTTT 550
E A I Q K N A T S F W G Q Y A F 113
551 CTGGAGAGTTATAACTATACCCTCGGCGCGGATGACTTGACTATCTTCGG 600
L E S Y N Y T L G A D D L T I F G 130
601 CGAGAACCAGATGGTTGATTCGGGTGCCAAGTTCTACCGACGGTATAAGA 650
E N Q M V D S G A K F Y R R Y K N 147
651 ATCTCGCCAGGAAAAATACTCCTTTTATCCGTGCATCAGGGTCTGACCGT 700
L A R K N T P F I R A S G S D R 163

图 10A

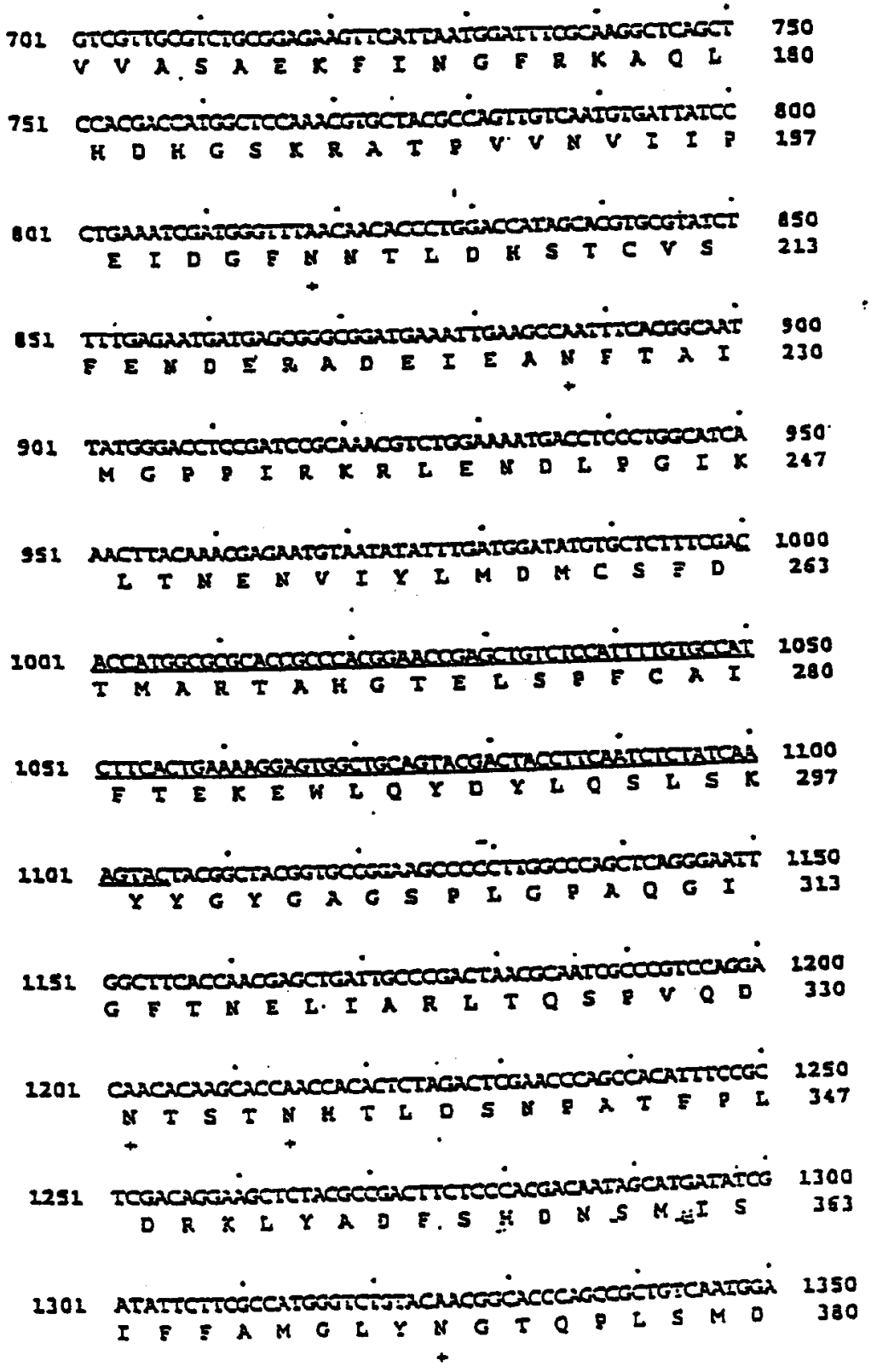

701 GTCGTTGCGTCTGCCGAGAAGTTCATTAATGGATTTCGCAAGGCTCAGCT 750
V V A S A E K F I N G F R K A Q L 180
751 CCACGACCATGGCTCCAAACGTGCTACGCCAGTTGTCAATGTGATTATCC 800
H D H G S K R A T P V V N V I I P 197
801 CTGAAATCGATGGGTTTAACAACACCCTGGACCATAGCACGTGCGTATCT 850
E I D G F N N T L D H S T C V S 213
851 TTTGAGAATGATGAGCGGGCGGATGAAATTGAAGCCAATTTCACGGCAAT 900
F E N D E R A D E I E A N F T A I 230
901 TATGGGACCTCCGATCCGCAAACGTCTGGAAAATGACCTCCCTGGCATCA 950
M G P P I R K R L E N D L P G I K 247
951 AACTTACAAACGAGAATGTAATATATTTGATGGATATGTGCTCTTTCGAC 1000
L T N E N V I Y L M D M C S F D 263
1001 ACCATGGCGCGCACTGCCCACGGAACCGAGCTGTCTCCATTTTGTGCCAT 1050
T M A R T A H G T E L S P F C A I 280
1051 CTTCACTGAAAAGGAGTGGCTGCAGTACGACTACCTTCAATCTCTATCAA 1100
F T E K E W L Q Y D Y L Q S L S K 297
1101 AGTACTACGGCTACGGTGCCGGAAGCCCCCTTGGCCCAGCTCAGGGAATT 1150
Y Y G Y G A G S P L G P A Q G I 313
1151 GGCTTCACCAACGAGCTGATTGCCCGACTAACGCAATCGCCCGTCCAGGA 1200
G F T N E L I A R L T Q S P V Q D 330
1201 CAACACAAGCACCAACCACACTCTAGACTCGAACCCAGCCACATTTCCGC 1250
N T S T N H T L D S N P A T F P L 347
1251 TCGACAGGAAGCTCTACGCCGACTTCTCCCACGACAATAGCATGATATCG 1300
D R K L Y A D F S H D N S M I S 363
1301 ATATTCTTCGCCATGGGTCTGTACAACGGCACCCAGCCGCTGTCAATGGA 1350
I F F A M G L Y N G T Q P L S M D 380

图 10B

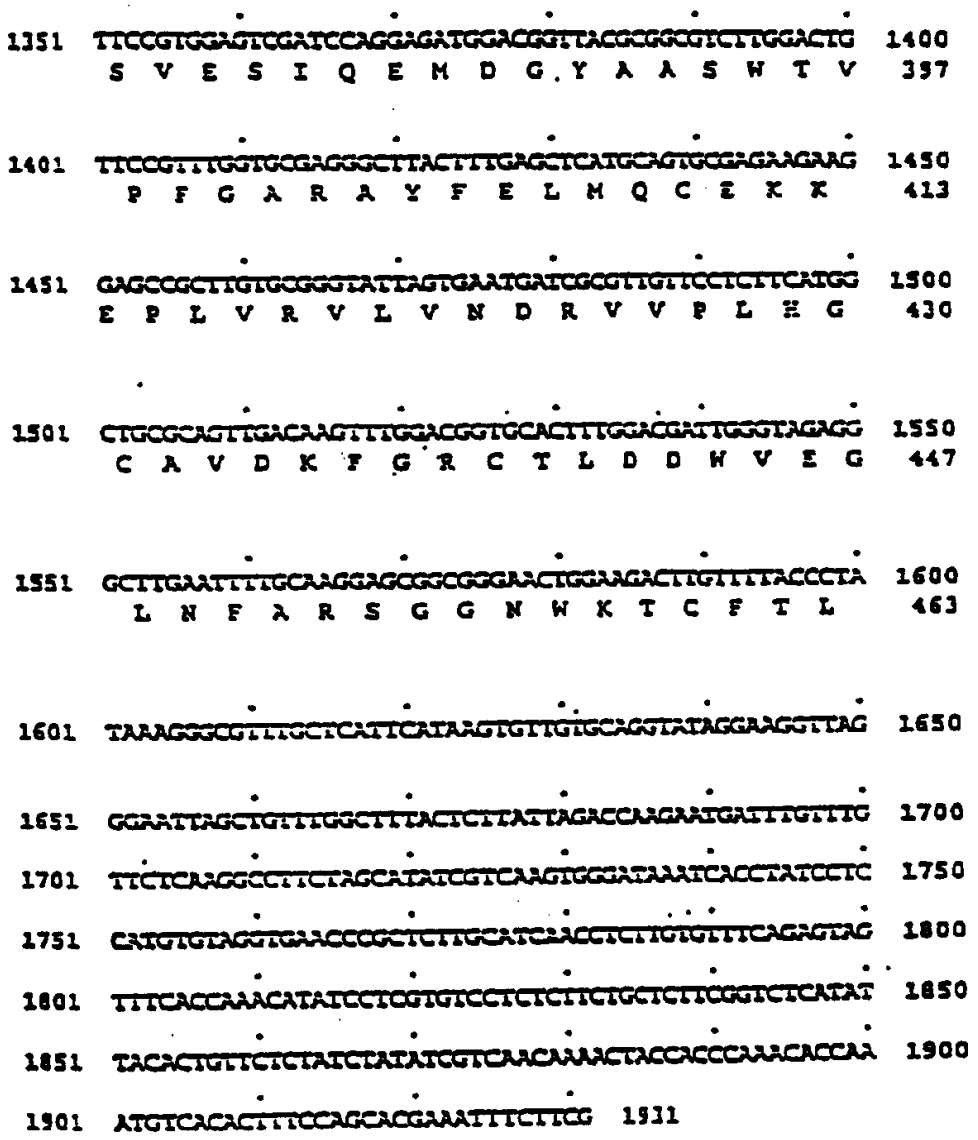

1351 TTCCGTGGAGTCGATCCAGGAGATGGACGGTTACGCGGCGTCTTGGACTG 1400
S V E S I Q E M D G Y A A S W T V 397
1401 TTCCGTTTGGTGCGAGGGCTTACTTTGAGCTCATGCAGTGCGAGAAGAAG 1450
P F G A R A Y F E L M Q C E K K 413
1451 GAGCCGCTTGTGCGGGTATTAGTGAATGATCGCGTTGTTCCTCTTCATGG 1500
E P L V R V L V N D R V V P L H G 430
1501 CTGCGCAGTTGACAAGTTTGGACGGTGCACTTTGGACGATTGGGTAGAGG 1550
C A V D K F G R C T L D D W V E G 447
1551 GCTTGAATTTTGCAAGGAGCGGCGGGAACTGGAAGACTTGTTTTACCCTA 1600
L N F A R S G G N W K T C F T L 463
1601 TAAAGGGCGTTTGCTCATTCATAAGTGTTGTGCAGGTATAGGAAGGTTAG 1650
1651 GGAATTAGCTGTTTGGCTTTACTCTTATTAGACCAAGAATGATTTGTTTG 1700
1701 TTCTCAAGGCCTTCTAGCATATCGTCAAGTGGGATAAATCACCTATCCTC 1750
1751 CATGTGTAGGTGAACCCGCTCTTGCATCAACCTCTTGTGTTTCAGAGTAG 1800
1801 TTTCACCAAACATATCCTCGTGTCCTCTCTTCTGCTCTTCGGTCTCATAT 1850
1851 TACACTGTTCTCTATCTATATCGTCAACAAACTACCACCCAAACACCAA 1900
1901 ATGTCACACTTTCCAGCACGAAATTTCTTCG 1931

图 10C

00.11.20

```
  1 ATGGGCGTCTCTGCTGTTCTACTTCCTTTGTATCTCCTGTCTGGAGTCACCTCCGGACTG
-23  M  G  V  S  A  V  L  L  F  L  Y  L  L  S  G  V  T  S  G  L

 61 GCAGTCCCCGCCTCGAGAAATCAATCCAGTTGCGATACGGTCGATCAGGGGTATCAATGC
     A  V  P  A  S  R  N  Q  S  S  C  D  T  V  D  Q  G  Y  Q  C
        -1 +1

121 TTCTCCGAGACTTCGCATCTTTGGGGTCAATACGCACCGTTCTTCTCTCTGGCAAACGAA
 18  F  S  E  T  S  H  L  W  G  Q  Y  A  P  F  F  S  L  A  N  E

181 TCGGTCATCTCCCCTGAGGTGCCCGCCGGATGCAGAGTCACTTTCGCTCAGGTCCTCTCC
 38  S  V  I  S  P  E  V  P  A  G  C  R  V  T  F  A  Q  V  L  S

241 CGTCATGGAGCGCGGTATCCGACCGACTCCAAGGGCAAGAAATACTCCGCTCTCATTGAG
 58  R  H  Q  A  R  Y  P  T  D  S  K  G  K  K  Y  S  A  L  I  E

301 GAGATCCAGCAGAACGCGACCACCTTTGACGGAAAATATGCCTTCCTGAAGACATACAAC
 78  E  I  Q  Q  N  A  T  T  F  D  G  K  Y  A  F  L  K  T  Y  N

361 TACAGCTTGGGTGCAGATGACCTGACTCCCTTCGGAGAACAGGAGCTAGTCAACTCCGGC
 98  Y  S  L  G  A  D  D  L  T  P  F  G  E  Q  E  L  V  N  S  G

421 ATCAAGTTCTACCAGCGGTACGAATCGCTCACAAGGAACATCGTTCCATTCATCCGATCC
118  I  K  F  Y  Q  R  Y  E  S  L  T  R  N  I  V  P  F  I  R  S

481 TCTGGCTCCAGCCGCGTGATCGCCTCCGGCAAGAAATTCATCGAGGGCTTCCAGAGCACC
138  S  G  S  S  R  V  I  A  S  G  K  K  F  I  E  G  F  Q  S  T

541 AAGCTGAAGGATCCTCGTGCCCAGCCCGGCCAATCGTCGCCCAAGATCGACGTGGTCATT
158  K  L  K  D  P  R  A  Q  P  G  Q  S  S  P  K  I  D  V  V  I

601 TCCGAGGCCAGCTCATCCAACAACACTCTCGACCCAGGCACCTGCACTGTGTTCGAAGAC
178  S  E  A  S  S  S  N  N  T  L  D  P  G  T  C  T  V  F  E  D

661 AGCGAATTGGCCGATACCGTCGAAGCCAATTTCACCGCCACGTTCGTCCCCTCCATTCGT
198  S  E  L  A  D  T  V  E  A  N  F  T  A  T  F  V  P  S  I  R

721 CAACGTCTGGAGAACGACCTGTCCGGTGTGACTCTCACAGACACAGAAGTGACCTACCTC
218  Q  R  L  E  N  D  L  S  G  V  T  L  T  D  T  E  V  T  Y  L

781 ATGGACATGTGCTCCTTCGACACCATCTCCACCAGCACCGTCGACACCAAGCTGTCCCCC
238  M  D  M  C  S  F  D  T  I  S  T  S  T  V  D  T  K  L  S  P
```

图 11A

00.11.20

841 TTCTGTGACCTGTTCACCCATGACGAATGGATCAACTACGACTACCTCCAGTCCTTGAAA
258 F C D L F T H D E W I N Y D Y L Q S L K

901 AAGTATTACGGCCATGGTGCAGGTAACCCGCTCGGCCCGACCCAGGGCGTCGGCTACGCT
278 K Y Y G H G A G N P L G P T Q G V G Y A

961 AACGAGCTCATCGCCCGTCTGACCCACTCGCCTGTCCACGATGACACCAGTTCCAACCAC
298 N E L I A R L T H S P V H D D T S S N H

1021 ACTTTGGACTCGAGCCCGGCTACCTTTCCGCTCAAGTCTACTCTCTACGCGGACTTTTCG
318 T L D S S P A T F P L N S T L Y A D F S

1081 CATGACAACGGCATCATCTCGATTCTCTTTGCTTTAGGTCTGTACAACGGCACTAAGCCG
338 H D N G I I S I L F A L G L Y N G T K P

1141 CTATCTACCACGACCGTGGAGAATATCACCCAGACAGATGGATTCTCGTCTGCTTGGACG
358 L S T T T V E N I T Q T D G F S S A W T

1201 GTTCCGTTTGCTTCGCGTTTGTACGTCGAGATGATGCAGTGTCAGGCGGAGCAGGAGCCG
378 V P F A S R L Y V E M M Q C Q A E Q E P

1261 CTGGTCCGTGTCTTGGTTAATGATCGCGTTGTCCCGCTGCATGGGTGTCCGGTTGATGCT
398 L V R V L V N D R V V P L H G C P V D A

1321 TTGGGGAGATGTACCCGGGATAGCTTTGTGAGGGGGTTGAGCTTTGCTAGATCTGGGGGT
418 L G R C T R D S F V R G L S F A R S G G

1381 GATTGGGCGGAGTGTTTTGCTTAG
438 D W A E C F A *

图 11B

```
tctagaacaataacaggtactccctaggtacccgaaggaccttgtggaaaatgtatggag     60
gtggacacggcaccaaccaccacccgcgatggcgcacgtggtgccctaaccccttgctcc    120
ctcaggatggaatccatgtcgactctttaccctcaccatcgcctggatgaaacctccccg    180
ctaagctcacgacgatcgctatttccgaccgatttgaccgtcatggtggagggctgattc    240
ggtcgatgctcctgccttcatttcggagttcggagacatgaaaggcttatatgaggacgt    300
cccaggtcggggacgaaatccgccctgggctgtgctccttcgtcggaaacatctgctgtc    360
cgtgatggctaccatgggctttcttgccattgtgctctccgtcgccttgctctttagaag    420
                M  G  F  L  A  I  V  L  S  V  A  L  L  F  R  S     16

gtatgcacccctctacgtccaattctctgggcactgacaacggcgcagcacatcgggcac    480
                                                T  S  G  T     20

cccgttgggccccgggggcaaacatagcgactgcaactcagtcgatcacggctatcaatg    540
 P  L  G  P  R  G  K  H  S  D  C  N  S  V  D  H  G  Y  Q  C     40

ctttcctgaactctctcataaatggggactctacgcgccctacttctccctccaggacga    600
 F  P  E  L  S  H  K  W  G  L  Y  A  P  Y  F  S  L  Q  D  E     60

gtctccgtttcctctggacgtcccagaggactgtcacatcaccttcgtgcaggtgctggc    660
 S  P  F  P  L  D  V  P  E  D  C  H  I  T  F  V  Q  V  L  A     80

ccgccacggcgcgcggagcccaacccatagcaagaccaaggcgtacgcggcgaccattgc    720
 R  H  G  A  R  S  P  T  H  S  K  T  K  A  Y  A  A  T  I  A    100

ggccatccagaagagtgccactgcgtttccgggcaaatacgcgttcctgcagtcatataa    780
 A  I  Q  K  S  A  T  A  F  P  G  K  Y  A  F  L  Q  S  Y  N    120

ctactccttggactctgaggagctgactcccttcgggcggaaccagctgcgagatctggg    840
 Y  S  L  D  S  E  E  L  T  P  F  G  R  N  Q  L  R  D  L  G    140

cgcccagttctacgagcgctacaacgccctcaccgacacatcaacccttcgtccgcgc    900
 A  Q  F  Y  E  R  Y  N  A  L  T  R  H  I  N  P  F  V  R  A    160

caccgatgcatcccgcgtccacgaatccgccgagaagttcgtcgagggcttccaaaccgc    960
 T  D  A  S  R  V  H  E  S  A  E  K  F  V  E  G  F  Q  T  A    180

tcgacaggacgatcatcacgccaatccccaccagccttcgcctcgcgtggacgtggccat   1020
 R  Q  D  D  H  H  A  N  P  H  Q  P  S  P  R  V  D  V  A  I    200

ccccgaaggcagcgcctacaacaacacgctggagcacagcctctgcaccgccttcgaatc   1080
 P  E  G  S  A  Y  N  N  T  L  E  H  S  L  C  T  A  F  E  S    220

cagcaccgtcggcgacgacgcggtcgccaacttcaccgccgtgttcgcgccggcgatcgc   1140
 S  T  V  G  D  D  A  V  A  N  F  T  A  V  F  A  P  A  I  A    240

ccagcgcctggaggccgatcttcccggcgtgcagctgtccaccgacgacgtggtcaacct   1200
 Q  R  L  E  A  D  L  P  G  V  Q  L  S  T  D  D  V  V  N  L    260

gatggccatgtgtccgttcgagacggtcagcctgaccgacgacgcgcacacgctgtcgcc   1260
 M  A  M  C  P  F  E  T  V  S  L  T  D  D  A  H  T  L  S  P    280

gttctgcgacctcttcacggccactgagtggacgcagtacaactacctgctctcgctgga   1320
 F  C  D  L  F  T  A  T  E  W  T  Q  Y  N  Y  L  L  S  L  D    300

caagtactacggctacggcgggggcaatccgctgggtccggtgcaggggtcggctgggc   1380
 K  Y  Y  G  Y  G  G  G  N  P  L  G  P  V  Q  G  V  G  W  A    320

gaacgagctgatggcgcggctaacgcgcgccccccgtgcacgaccacacctgcgtcaacaa   1440
 N  E  L  M  A  R  L  T  R  A  P  V  H  D  H  T  C  V  N  N    340

caccctcgacgcgagtccggccaccttcccgctgaacgccacccctctacgccgacttctc   1500
 T  L  D  A  S  P  A  T  F  P  L  N  A  T  L  Y  A  D  F  S    360
```

图 12A

00.11.20

```
ccacgacagcaacctggtgtcgatcttctgggcgctgggcctgtacaacggcaccgcgcc 1560
 H  D  S  N  L  V  S  I  F  W  A  L  G  L  Y  N  G  T  A  P   380

gctgtcgcagacctccgtcgagagcgtctcccagacggacgggtacgccgccgcctggac 1620
 L  S  Q  T  S  V  E  S  V  S  Q  T  D  G  Y  A  A  A  W  T   400

ggtgccgttcgccgctcgcgcgtacgtcgagatgatgcagtgtcgcgccgagaaggagcc 1680
 V  P  F  A  A  R  A  Y  V  E  M  M  Q  C  R  A  E  K  E  P   420

gctggtgcgcgtgctggtcaacgaccgggtcatgccgctgcatggctgccctacggacaa 1740
 L  V  R  V  L  V  N  D  R  V  M  P  L  H  G  C  P  T  D  K   440

gctggggcggtgcaagcgggacgctttcgtcgcggggctgagctttgcgcaggcgggcgg 1800
 L  G  R  C  K  R  D  A  F  V  A  G  L  S  F  A  Q  A  G  G   460

gaactgggcggattgtttctgatgttgagaagaaaggtagatagataggtagtacatatg 1860
 N  W  A  D  C  F                                              466

gattgctcggctctgggtcgttgcccacaatgcatattacgcccgtcaactgccttgcgc 1920
catccacctctcaccctggacgcaaccgagcggtctaccctgcacacggcttccaccgcg 1980
acgcgcacggataaggcgcttttgttacggggttggggctgggggcagccggagccggag 2040
agagagaccagcgtgaaaaacgacagaacatagatatcaattcgacgccaattcatgcag 2100
agtagtatacagacgaactgaaacaaacacatcacttccctcgctcctctcctgtagaag 2160
acgctcccaccagccgcttctggcccttattcccgtacgctaggtagaccagtcagccag 2220
acgcatgcctcacaagaacgggggcgggggacacactccgctcgtacagcacccacgacg 2280
tgtacaggaaaaccggcagcgccacaatcgtcgagagccatctgcag               2327
```

图 12B

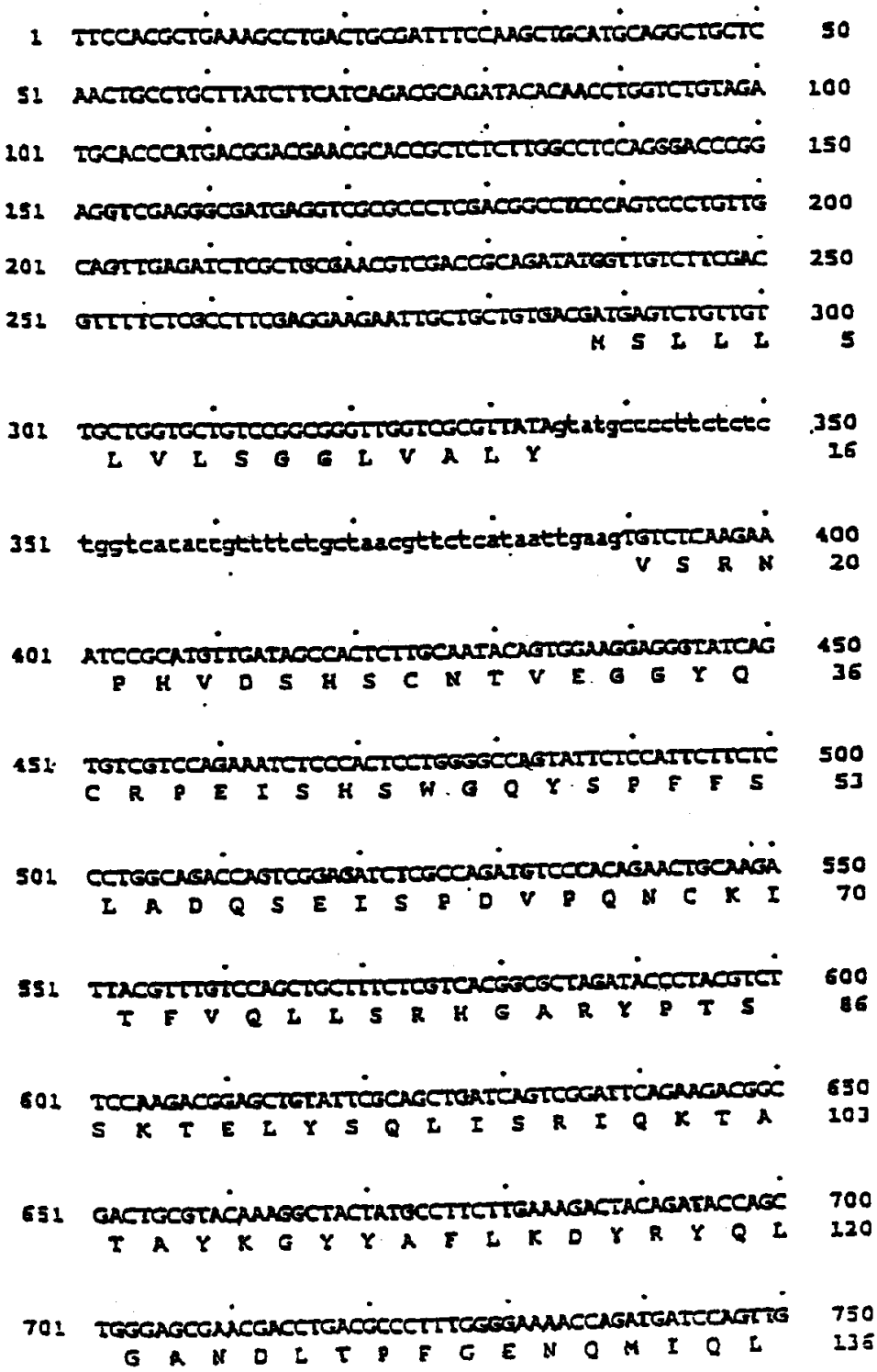

1 TTCCACGCTGAAAGCCTGACTGCGATTTCCAAGCTGCATGCAGGCTGCTC 50
51 AACTGCCTGCTTATCTTCATCAGACGCAGATACACAACCTGGTCTGTAGA 100
101 TGCACCCATGACGGACGAACGCACCGCTCTCTTGGCCTCCAGGGACCCGG 150
151 AGGTCGAGGGCGATGAGGTCGCGCCCTCGACGGCCTCCCAGTCCCTGTTG 200
201 CAGTTGAGATCTCGCTGCGAACGTCGACCGCAGATATGGTTGTCTTCGAC 250
251 GTTTTCTCGCCTTCGAGGAAGAATTGCTGCTGTGACGATGAGTCTGTTGT 300
M S L L L 5
301 TGCTGGTGCTGTCCGGCGGGTTGGTCGCGTTATAGtatgcccctttctctc 350
L V L S G G L V A L Y 16
351 tggtcatactgttttctgctaacgttctcataattgaagTGTCTCAAGAA 400
V S R N 20
401 ATCCGCATGTTGATAGCCACTCTTGCAATACAGTGGAAGGAGGGTATCAG 450
P H V D S H S C N T V E G G Y Q 36
451 TGTCGTCCAGAAATCTCCCACTCCTGGGGCCAGTATTCTCCATTCTTCTC 500
C R P E I S H S W G Q Y S P F F S 53
501 CCTGGCAGACCAGTCGGAGATCTCGCCAGATGTCCCACAGAACTGCAAGA 550
L A D Q S E I S P D V P Q N C K I 70
551 TTACGTTTGTCCAGCTGCTTTCTCGTCACGGCGCTAGATACCCTACGTCT 600
T F V Q L L S R H G A R Y P T S 86
601 TCCAAGACGGAGCTGTATTCGCAGCTGATCAGTCGGATTCAGAAGACGGC 650
S K T E L Y S Q L I S R I Q K T A 103
651 GACTGCGTACAAAGGCTACTATGCCTTCTTGAAAGACTACAGATACCAGC 700
T A Y K G Y Y A F L K D Y R Y Q L 120
701 TGGGAGCGAACGACCTGACGCCCTTTGGGGAAAACCAGATGATCCAGTTG 750
G A N D L T P F G E N Q M I Q L 136

图 13A

```
751  GGCATCAAGTTTTATAACCATTACAAGAGTCTCGCCAGGAATGCCGTCCC   800
     G  I  K  F  Y  N  H  Y  K  S  L  A  R  N  A  V  P    153

801  ATTCGTTCGTTGCTCCGGCTCTGATCGGGTCATTGCCTCGGGGAGACTTT   850
     F  V  R  C  S  G  S  D  R  V  I  A  S  G  R  L  F    170

851  TCATCGAAGGTTTCCAGAGCGCCAAAGTGCTGGATCCTCATTCAGACAAG   900
     I  E  G  F  Q  S  A  K  V  L  D  P  H  S  D  K      186

901  CATGACGCTCCTCCCACGATCAACGTGATCATCGAGGAGGGTCCGTCCTA   950
     H  D  A  P  P  T  I  N  V  I  I  E  E  G  P  S  Y    203

951  CAATAACACGCTCGACACCGGCAGCTGTCCAGTCTTTGAGGACAGCAGCG  1000
     N  N  T  L  D  T  G  S  C  P  V  F  E  D  S  S  G    220
     +

1001 GGGGACATGACGCACAGGAAAAGTTCGCAAAGCAATTCGCACCAGCTATC  1050
     G  H  D  A  Q  E  K  F  A  K  Q  F  A  P  A  I      236

1051 CTGGAAAAGATCAAGGACCATCTTCCCGGCGTGGACCTGGCCGTGTCGGA  1100
     L  E  K  I  K  D  H  L  P  G  V  D  L  A  V  S  D    253

1101 TGTACCGTACTTGATGGACTTGTGTCCGTTTGAGACCTTGGCTCGCAACC  1150
     V  P  Y  L  M  D  L  C  P  F  E  T  L  A  R  N  H    270
                                                  +

1151 ACACAGACACGCTGTCTCCGTTCTGCGCTCTTTCCACGCAAGAGGAGTGG  1200
     T  D  T  L  S  P  F  C  A  L  S  T  Q  E  E  W      286

1201 CAAGCATATGACTACTACCAAAGTCTGGGGAAATACTATGGCAATGGCGG  1250
     Q  A  Y  D  Y  Y  Q  S  L  G  K  Y  Y  G  N  G  G    303

1251 GGGTAACCCGTTGGGGCCAGCCCAAGGCGTGGGGTTTGTCAACGAGTTGA  1300
     G  N  P  L  G  P  A  Q  G  V  G  F  V  N  E  L  I    320

1301 TTGCTCGCATGACCCATAGCCCTGTCCAGGACTACACCACGGTCAACCAC  1350
     A  R  M  T  H  S  P  V  Q  D  Y  T  T  V  N  H      336

1351 ACTCTTGACTCGAATCCGGCGACATTCCCTTTGAACGCGACGCTGTACGC  1400
     T  L  D  S  N  P  A  T  F  P  L  N  A  T  L  Y  A    353
```

图 13B

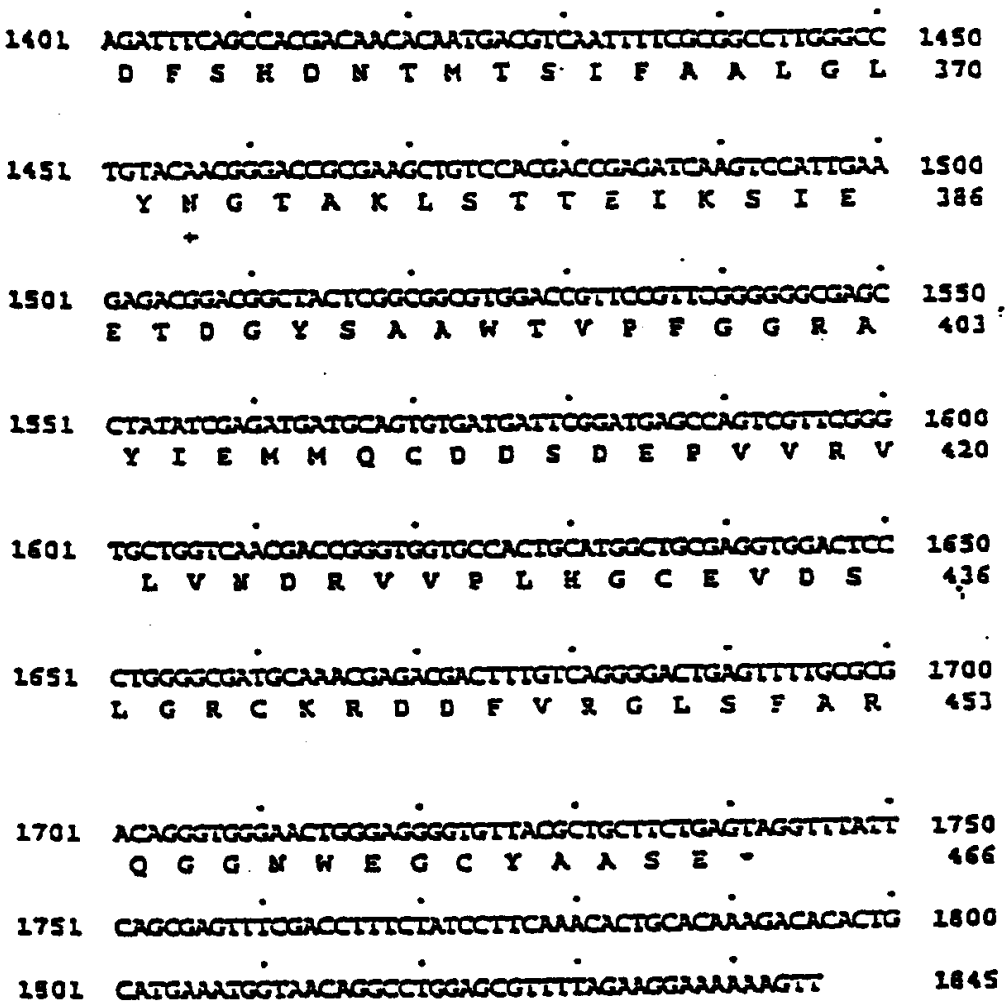

1401 AGATTTCAGCCACGACAACACAATGACGTCAATTTTCGCGGCCTTGGGCC 1450
D F S H D N T M T S I F A A L G L 370
1451 TGTACAACGGGACCGCGAAGCTGTCCACGACCGAGATCAAGTCCATTGAA 1500
Y N G T A K L S T T E I K S I E 386
1501 GAGACGGACGGCTACTCGGCGGCGTGGACCGTTCCGTTCGGGGGGCGAGC 1550
E T D G Y S A A W T V P F G G R A 403
1551 CTATATCGAGATGATGCAGTGTGATGATTCGGATGAGCCAGTCGTTCGGG 1600
Y I E M M Q C D D S D E P V V R V 420
1601 TGCTGGTCAACGACCGGGGTGGTGCCACTGCATGGCTGCGAGGTGGACTCC 1650
L V N D R V V P L H G C E V D S 436
1651 CTGGGGCGATGCAAACGAGACGACTTTGTCAGGGGACTGAGTTTTGCGCG 1700
L G R C K R D D F V R G L S F A R 453
1701 ACAGGGTGGGAACTGGGAGGGGTGTTACGCTGCTTCTGAGTAGGTTTATT 1750
Q G G N W E G C Y A A S E * 466
1751 CAGCGAGTTTCGACCTTTCTATCCTTCAAACACTGCACAAGACACACTG 1800
1801 CATGAAATGGTAACAGGCCTGGAGCGTTTTAGAAGGAAAAAAGTT 1845

图 13C

```
AATTACGGAGTAGTTGCCATTCGATGTTCATTGATCAACAGTCAACCGCAAGTTTCGTAGTATTTTCCAAACTCCTCCACTGGCCGTGCG      90

TTGCGACACGACCTGCATGAGAATCGATCGATCGATCGATGATCGCTCAGGATGATCTGATCATCTCGGGTTGGAAGAGTCCACTTTATG     180

ACCAGGGGATTGATTTTTCAATGCGTTGGTTGTTGTTCATCCGATTCATGAACAAGTGGACATTATTATTATGATTGCACGTGTCCTAAG     270

CTGCAAGTACTATTGAATAGTGCTTCAATGCTACCATGATCGGACACCAACACTCATGGAAGCCCGCCCCTAGCCGGCAGATCTGGCACA     360

CGCATCGTGCTGATATAAAAAGACTGCCAAATGCCGAAGACGAAATGCAGCAACGTTCAGCCCGCAGAGTGATTGCCGTCATGGCGGGGA     450
                                                                                  M  A  G

TAGGTTTGGGGTCCTTTCTGGTCCTGCTGCTGCAATTgtacgcattcttctagaccctaattatagaggtctgttgctgatatctctgact     540
I  G  L  G  S  F  L  V  L  L  L  Q  F  ————————————————————————————————————————————————

agTTCGGCATTATTGACGGCCTCGCCGGCCATTCCTCCTTTCTGGAGGAAGAAGCATCCCAACGTGGACATTGCCCGCCACTGGGGCCAG     630
—  S  A  L  L  T  A  S  P  A  I  P  P  F  W  R  K  K  H  P  N  V  D  I  A  R  H  W  G  Q

TACTCGCCCTTCTTCTCGCTGGCCGAGGTCTCTGAAATCTCGCCTGCGGTGCCCAAGGGCTGTCGTGTCGAGTTTGTGCAGGTGCTGTCC     720
Y  S  P  F  F  S  L  A  E  V  S  E  I  S  P  A  V  P  K  G  C  R  V  E  F  V  Q  V  L  S

CGGCACGGAGCTCGGTATCCTACTGCTCACAAGAGTGAAGTCTACGCCGAGTTGCTTCAAAGGATCCAGGACACTGCGACCGAGTTCAAG     810
R  H  G  A  R  Y  P  T  A  H  K  S  E  V  Y  A  E  L  L  Q  R  I  Q  D  T  A  T  E  F  K

GGCGATTTTGCCTTTCTCCGAGACTATGCCTATCATCTCGGTGCCGATAATTTGACGCGCTTTGGCGAGGAGCAGATGATGGAATCGGGC     900
G  D  F  A  F  L  R  D  Y  A  Y  H  L  G  A  D  N  L  T  R  F  G  E  E  Q  M  M  E  S  G
```

图 14A

```
CGCCAGTTCTACCACCGGTATCGTGAGCAGGCCCGAGAGATTGTGCCATTTGTGCGTGCGGCAGGCTCCGCGCGAGTCATTGCGTCGGCA   990
 R  Q  F  Y  H  R  Y  R  E  Q  A  R  E  I  V  P  F  V  R  A  A  G  S  A  R  V  I  A  S  A

GAGTTCTTCAACCGCGGATTCCAGGATGCCAAAGACCGGGATCCCAGGTCGAACAAGGACCAGGCAGAGCCTGTGATCAACGTGATCATT  1080
 E  F  F  N  R  G  F  Q  D  A  K  D  R  D  P  R  S  N  K  D  Q  A  E  P  V  I  N  V  I  I

TCCGAAGAAACTGGCAGTAACAATACTCTGGATGGGCTGACGTGCCCGGCGGCCGAGGAGGCACCGGACCCAACCCAGCCCGCAGAGTTC  1170
 S  E  E  T  G  S  N  N  T  L  D  G  L  T  C  P  A  A  E  E  A  P  D  P  T  Q  P  A  E  F

CTGCAAGTTTTCGGCCCGCGTGTCTTGAAAAAGATCACTAAACACATGCCGGGTGTGAACCTCACCTTGGAGGATGTCCCGTTGTTCATG  1260
 L  Q  V  F  G  P  R  V  L  K  K  I  T  K  H  M  P  G  V  N  L  T  L  E  D  V  P  L  F  M

GATCTTTGTCCGTTTGACACGGTGGGCTCCGACCCAGTTCTTTTCCCACGGCAGCTCTCTCCGTTTTGTCACTTGTTCACGGCCGACGAT  1350
 D  L  C  P  F  D  T  V  G  S  D  P  V  L  F  P  R  Q  L  S  P  F  C  H  L  F  T  A  D  D

TGGATGGCCTACGATTACTACTACACCCTCGACAAATACTACAGCCACGGCGGCGGCAGCGCATTTGGCCCGTCCCGCGGCGTCGGGTTC  1440
 W  M  A  Y  D  Y  Y  Y  T  L  D  K  Y  Y  S  H  G  G  G  S  A  F  G  P  S  R  G  V  G  F

GTCAACGAGCTGATTGCGCGTATGACGGGAAATCTTCCCGTCAAGGACCACACAACAGTCAACCACACTCTCGATGACAACCCGGAAACT  1530
 V  N  E  L  I  A  R  M  T  G  N  L  P  V  K  D  H  T  T  V  N  H  T  L  D  D  N  P  E  T

TTCCCGTTGGACGCTGTCCTCTACGCAGACTTTTCGCACGACAACACCATGACGGGCATCTTTTCCGCAATGGGCCTGTACAACGGCACA  1620
 F  P  L  D  A  V  L  Y  A  D  F  S  H  D  N  T  M  T  G  I  F  S  A  M  G  L  Y  N  G  T

AAGCCGCTGTCGACGTCCAAGATTCAGCCTCCGACGGGTGCAGCAGCGGATGGATATGCGGCATCGTGGACGGTGCCGTTCGCAGCGAGG  1710
 K  P  L  S  T  S  K  I  Q  P  P  T  G  A  A  A  D  G  Y  A  A  S  W  T  V  P  F  A  A  R
```

图 14B

GCGTATGTGGAGTTGCTGCGATGTGAGACGGAAACGAGCTCTGAGGAGGAGGAGGAGGGGGAGGACGAGCCGTTCGTGCGGGTTCTGGTG 1800
A Y V E L L R C E T E T S S E E E E E E G E D E P F V R V L V

AATGATCGGGTTGTGCCGCTGCATGGTTGTCGGGTTGATCGATGGGGGAGGTGTCGGAGGGATGAGTGGATTAAGGGACTCACGTTTGCT 1890
N D R V V P L H G C R V D R W G R C R R D E W I K G L T F A

CGACAGGGTGGGCATTGGGATCGCTGCTTTTGATTAGATGCTCATAGACATAACCCCATGATTCCGAATTGATGTTTTTAGATACAATCA 1980
R Q G G H W D R C F

CTGCGGAAAGGGAAATGATCCAAAAAGCGCCAGTCTAGTATAACTTTGCGAATCCGTTGACTTGTTCAGTCCTTGGTGTCGCCATCAACC 2070

AGGCCTGCCACAAGGTCCAATGTTCCCGCTCTACATGGAGTCCGTCGTCGCCGAGATCATCCACGCCCAGCGCACGGAGCTGTTCCGTTG 2160

AGGGTATCTGCCGTGGTTGACCCCCGTGCTCACAGTCACA 2200

图 14C

```
AAGCTTGGGCAAACTCATCATGCTCATCTTGATGATTCCACTGTTCAGCTACCTGGCTGCTGCTTCTCTGTGGGTTCATC
 80
HindIII          M  L  I  L  M  I  P  L  F  S  Y  L  A  A  A  S  L

CTTTGCCCCTGTCTCGATGTTAAAATACTAAACATATTTCACCAGACGTGTACTCTCCCCTCAGCCAGTGTCCTGTGACA
 160
                                                R  V  L  S  P  Q  P  V  S  C  D

GCCCGGAGCTTGGTTACCAATGCGACCAGCAGACAACGCACACCTGGGGTCAATACTCACCCTTCTTCTCTGTCCCGTCA
 240
S  P  E  L  G  Y  Q  C  D  Q  Q  T  T  H  T  W  G  Q  Y  S  P  F  F  S  V  P  S

GAGATCTCCCCTTCCGTTCCTGATGGCTGCCGCCTCACCTTCGCCCAAGTTCTCTCCCGCCACGGCGCCCGCTTCCCAAC
 320
E  I  S  P  S  V  P  D  G  C  R  L  T  F  A  Q  V  L  S  R  H  G  A  R  F  P  T

CCCGGGTAAAGCCGCCGCCATCTCCGCTGTCCTCACCAAAATCAAAACCTCTGCCACCTGGTACGGTTCCGACTTTCAGT
 400
  P  G  K  A  A  A  I  S  A  V  L  T  K  I  K  T  S  A  T  W  Y  G  S  D  F  Q

TCATCAAGAACTACGACTATGTACTTGGCGTAGACCACCTGACCGCGTTCGGCGAGCAAGAAATGGTCAACTCCGGCATC
 480
F  I  K  N  Y  D  Y  V  L  G  V  D  H  L  T  A  F  G  E  Q  E  M  V  N  S  G  I

AAGTTCTACCAGCGCTACTCCTCCCTCATCCAGACAGAAGACTCGGATACGCTCCCCTTCGTCCGCGCCTCTGGCCAGGA
 560
K  F  Y  Q  R  Y  S  S  L  I  Q  T  E  D  S  D  T  L  P  F  V  R  A  S  G  Q  E

ACGCGTCATCGCCTCCGCCGAGAACTTCACCACCGGCTTCTACTCGGCCCTCTCAGCCGACAAGAACCCTCCTTCCTCCT
 640
  R  V  I  A  S  A  E  N  F  T  T  G  F  Y  S  A  L  S  A  D  K  N  P  P  S  S

TACCAAGACCAGAAATGGTCATCATTTCTGAGGAGCCAACAGCCAACAACACCATGCACCACGGCCTCTGCCGCTCCTTT
 720
L  P  R  P  E  M  V  I  I  S  E  E  P  T  A  N  N  T  M  H  H  G  L  C  R  S  F

GAAGATTCCACCACCGGCGACCAAGCCCAAGCGGAATTCATCGCCGCCACCTTCCCACCCATCACCGCCCGTCTCAACGC
 800
E  D  S  T  T  G  D  Q  A  Q  A  E  F  I  A  A  T  F  P  P  I  T  A  R  L  N  A

CCAAGGTTTCAAAGGCGTCACCCTCTCCAACACCGACGTCCTATCACTAATGGACCTCTGCCCCTTTGACACCGTCGCCT
 880
  Q  G  F  K  G  V  T  L  S  N  T  D  V  L  S  L  M  D  L  C  P  F  D  T  V  A

ACCCCCTTTCCTCCCTCACCACCACCTCTTCCGTTTCTGGAGGCGGCAAGTTATCCCCCTTCTGCTCTCTTTTCACTGCC
 960
Y  P  L  S  S  L  T  T  T  S  S  V  S  G  G  G  K  L  S  P  F  C  S  L  F  T  A

AGCGACTGGACAATCTACGATTACCTCCAGTCCCTAGGGAAATACTACGGTTTCGGCCCCGGTAATTCCCTAGCTGCCAC
 1040
S  D  W  T  I  Y  D  Y  L  Q  S  L  G  K  Y  Y  G  F  G  P  G  N  S  L  A  A  T

CCAGGGGGTAGGGTACGTCAACGAGCTTATCGCCCGCTTGATCCGTGCTCCCGTCGTAGATCACACGACGACCAACTCTA
 1120
  Q  G  V  G  Y  V  N  E  L  I  A  R  L  I  R  A  P  V  V  D  H  T  T  T  N  S

CTCTTGATGGCGACGAAAAAACGTTTCCGTTGAACAGAACGGTGTATGCGGATTTTTCCCATGATAATGATATGATGAAT
 1200
T  L  D  G  D  E  K  T  F  P  L  N  R  T  V  Y  A  D  F  S  H  D  N  D  M  M  N
```

图 15A

00.1.20

```
ATCCTGACTGCTTTGCGGATATTCGAGCATATCAGTCCGATGGATAACACCACTATCCCGACCAACTATGGCCAGACAGG
 1280
 I  L  T  A  L  R  I  F  E  H  I  S  P  M  D  N  T  T  I  P  T  N  Y  G  Q  T  G

AGATGACGGGGTGAAGGAAAGGGATTTGTTCAAGGTTAGTTGGGCGGTGCCCTTTGCTGGGAGGGTGTACTTTGAGAAAA
 1360
  D  D  G  V  K  E  R  D  L  F  K  V  S  W  A  V  P  F  A  G  R  V  Y  F  E  K

TGGTTTGTGATGCGGATGGGGATGGCAAGATTGATAGTGATGAGGCTCAGAAAGAGTTGGTGAGGATTTTGGTTAATGAT
 1400
M  V  C  D  A  D  G  D  G  K  I  D  S  D  E  A  Q  K  E  L  V  R  I  L  V  N  D

CGGGTGATGAGATTGAATGGGTGTGATGCTGATGAACAGGGTAGGTGTGGATTGGAGAAGTTTGTGGAGAGTATGGAGTT
 1520
 R  V  M  R  L  N  G  C  D  A  D  E  Q  G  R  C  G  L  E  K  F  V  E  S  M  E  F

TGCGAGGAGAGGGGGGGAGTGGGAGGAGAGGTGTTTTGTTTAGCTCTAGA
  A  R  R  G  G  E  W  E  E  R  C  F  V     XbaI
```

图 15B

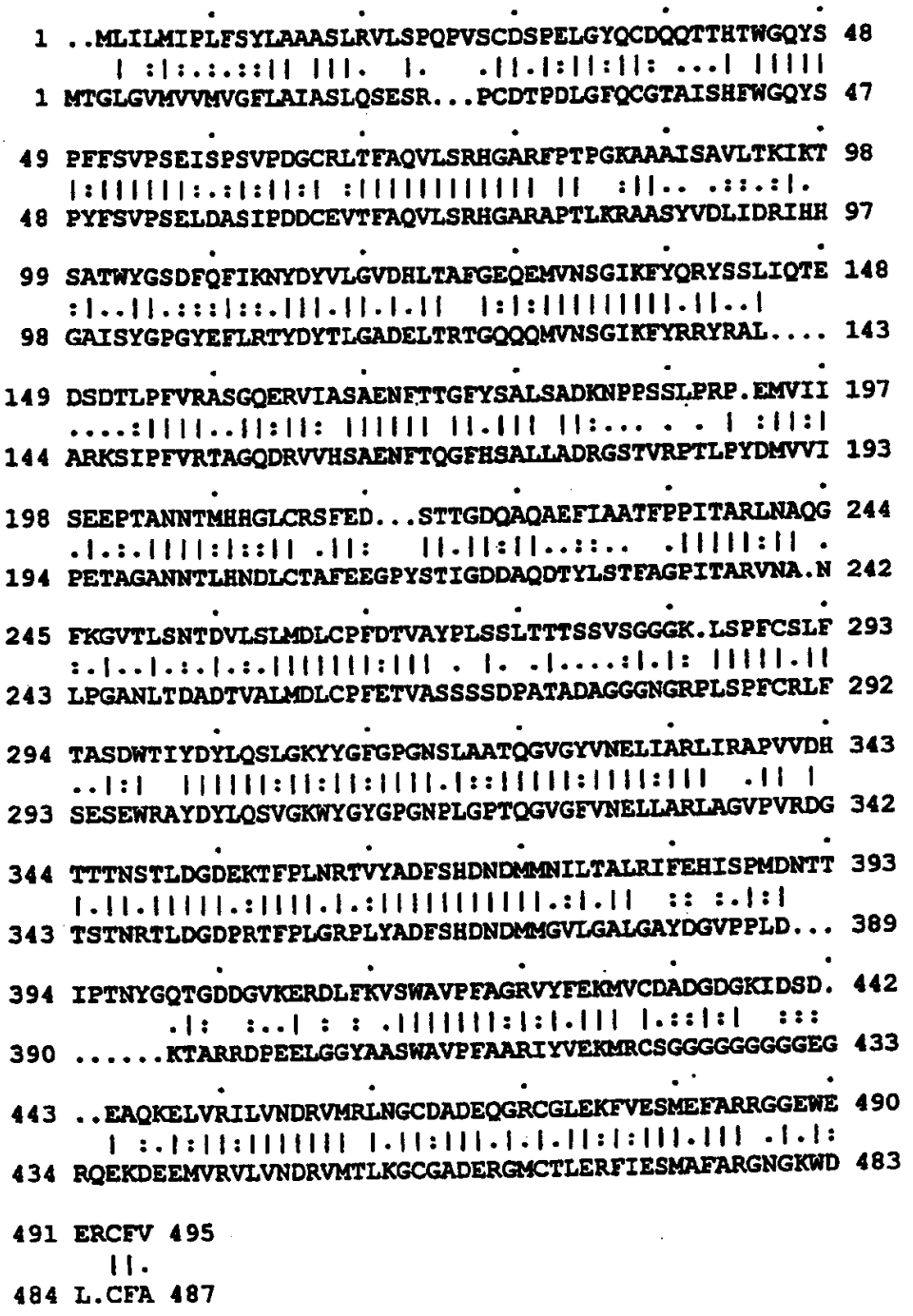


图 16

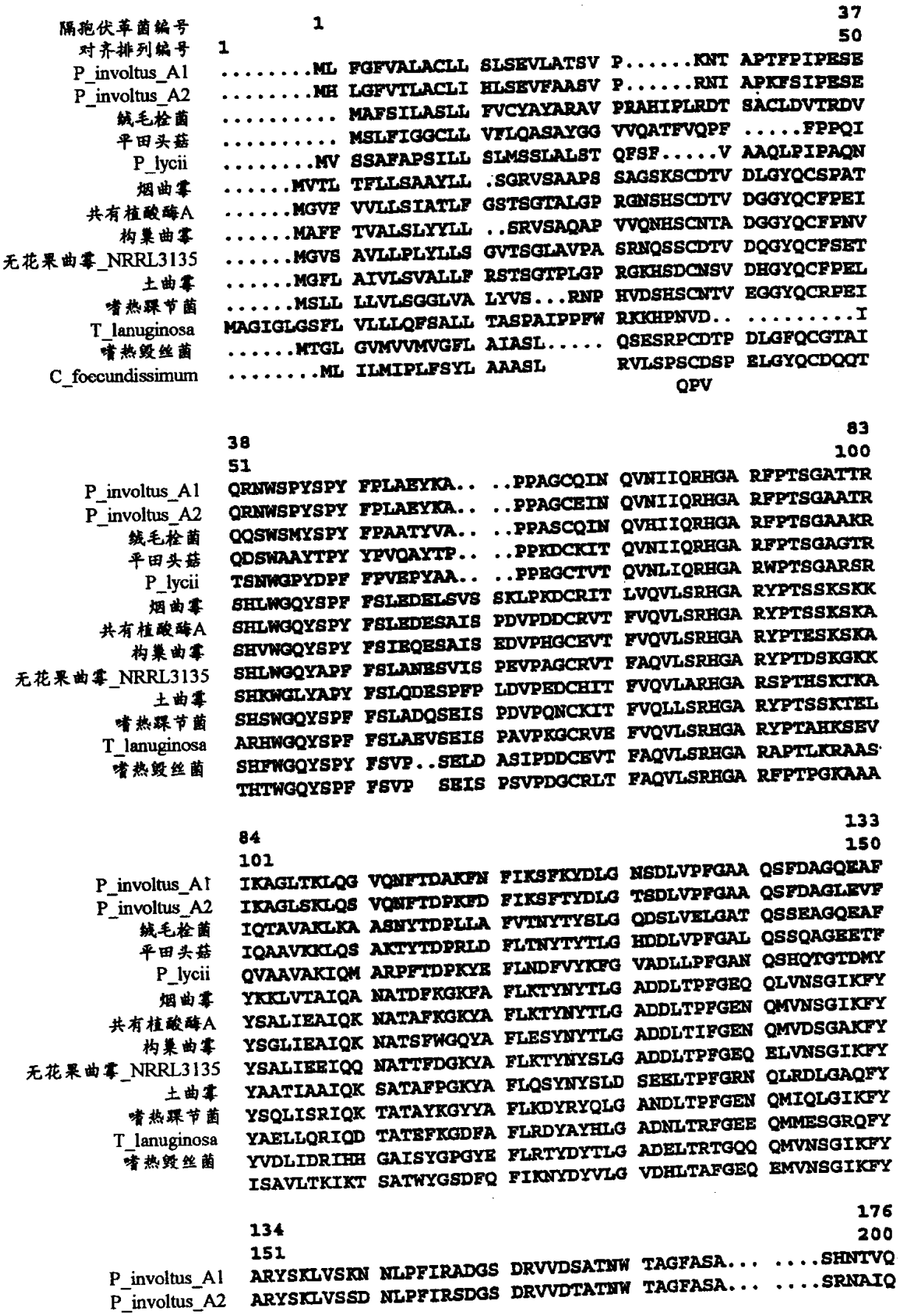

隔孢伏革菌编号
对齐排列编号
P_involtus_A1
P_involtus_A2
绒毛栓菌
平田头菇
P_lycii
烟曲霉
共有植酸酶A
构巢曲霉
无花果曲霉_NRRL3135
土曲霉
嗜热踝节菌
T_lanuginosa
嗜热毁丝菌
C_foecundissimum

图 17A

```
                  TRYSSLVSAD ELPFVRASGS DRVVATANNW TAGFALA... ....SSNSIT
                  QRYSFLVSKE NLPFVRASSS NRVVDSATNW TEGFSAA... ....SHHVLN
                  TRYSTLFEGG DVPFVRAAGD QRVVDSSTNW TAGFGDA... ....SGETVL
                  QRYKAL.ARS VVPFIRASGS DRVIASGEKF IEGFQQAKLA DPGA.TNRAA
                  RRYKAL.ARK IVPFIRASGS DRVIASAEKF IEGFQSAKLA DPGSQPHQAS
                  RRYKNL.ARK NTPFIRASGS DRVVASAEKF INGFRKAQLH DHGS..KRAT
                  QRYESL.TRN IVPFIRSSGS SRVIASGKKF IEGFQSTKLK DPRAQPGQSS
                  ERYNAL.TRH INPFVRATDA SRVHESAEKF VEGFQTARQD DHHANPHQPS
                  NHYKSL.ARN AVPFVRCSGS DRVIASGRLF IEGFQSAKVL DPHSDKHDAP
                  HRYREQ.ARE IVPFVRAAGS ARVIASAEFF NRGFQDAKDR DPRSNKDQAE
                  RRYRAL.ARK SIPFVRTAGQ DRVVHSAENF TQGFHSALLA DRGSTVRPTL
                  QRYSSLIDSD TLPFVRASGQ ERVIASAENF TTGFYSALSA DKNPPSSLPR
                        QTE

                  177                                                 217
                  201                                                 250
P_involtus_A1     PKLNLILPQT G..NDTLEDN MCPAAGD... ...SDPQVNA WLAVAFPSIT
P_involtus_A2     PKLDLILPQT G..NDTLEDN MCPAAGE... ...SDPQVDA WLASAFPSVT
绒毛栓菌          PVLSVIISEA G..NDTLDDN MCPAAGD... ...SDPQVNQ WLAQFAPPMT
平田头菇          PILFVILSES L..NDTLDDA MCPNAGS... ...SDPQTGI WTSIYGTPIA
P_lycii           PTLQVVLQEE G..NCTLCNN MCPNEVD... ...GD.ESTT WLGVFAPNIT
烟曲霉            PAISVIIPES ETFNNTLDHG VCTKFEA... SQLGDEVAAN FTALFAPDIR
共有植酸酶A       PVIDVIIPEG SGYNNTLDHG TCTAFED... SELGDDVEAN FTALFAPAIR
构巢曲霉          PVVNVIIPEI DGFNNTLDHS TCVSFEN... DERADEIEAN FTAIMGPPIR
无花果曲霉_NRRL3135 PKIDVVISEA SSSNNTLDPG TCTVFED... SELADTVEAN FTATFVPSIR
土曲霉            PRVDVAIPEG SAYNNTLEHS LCTAFES... STVGDDAVAN FTAVFAPAIA
嗜热踝节菌        PTINVIIEEG PSYNNTLDTG SCPVFED... SSGGHDAQEK FAKQFAPAIL
T_lanuginosa      PVINVIISEE TGSNNTLDGL TCPAAEE... AP.DPTQPAE FLQVFGPRVL
嗜热毁丝菌        PYDMVVIPET AGANNTLHND LCTAFEEGPY STIGDDAQDT YLSTFAGPIT
                  P.EMVIISEE PTANNTMHHG LCRSFED    STTGDQAQAE FIAATFPPIT

                  218                                                 252
                  251                                                 300
P_involtus_A1     ARLNAAAPSV NLTDTDAFNL VSLCAFLTVS KEKK...... .........S
P_involtus_A2     AQLNAAAPGA NLTDADAFNL VSLCPFMTVS KEQK...... .........S
绒毛栓菌          ARLNAGAPGA NLTDTDTYNL LTLCPFETVA TERR...... .........S
平田头菇          NRLNQQAPGA NITAADVSNL IPLCAFETIV KETP...... .........S
P_lycii           ARLNAAAPSA NLSDSDALTL MDMCPFDTLS SGNA...... .........S
烟曲霉            ARAEKHLPGV TLTDEDVVSL MDMCSFDTVA RTSD..ASQ. ........LS
共有植酸酶A       ARLEADLPGV TLTDEDVVYL MDMCPFETVA RTSD..ATE. ........LS
构巢曲霉          KRLENDLPGI KLTNENVIYL MDMCSFDTMA RTAH..GTE. ........LS
无花果曲霉_NRRL3135 QRLENDLSGV TLTDTEVTYL MDMCSFDTIS TSTV..DTK. ........LS
土曲霉            QRLEADLPGV QLSTDDVVNL MAMCPFETVS LTDD..AHT. ........LS
嗜热踝节菌        EKIKDHLPGV DLAVSDVPYL MDLCPFETLA RNHT..DT.. ........LS
T_lanuginosa      KKITKHMPGV NLTLEDVPLF MDLCPFDTVG SDPVLFPRQ. ........LS
嗜热毁丝菌        ARVNANLPGA NLTDADTVAL MDLCPFETVA SSSSDPATAD AGGGNGRPLS
                  ARLNAGFKGV TLSNTDVLSL MDLCPFDTVA YPLSSLTTTS SVSGGGK LS
                     Q

                  253                                                 300
                  301                                                 350
P_involtus_A1     DFCTLFEGIP GSFEAFAYGG DLDKFYGTGY GQELGPVQGV GYVNELIARL
P_involtus_A2     DFCTLFEGIP GSFEAFAYAG DLDKFYGTGY GQALGPVQGV GYINELLARL
绒毛栓菌          EFCDIYEELQ AE.DAFAYNA DLDKFYGTGY GQPLGPVQGV GYINELIARL
平田头菇          PFCNLFT..P EEFAQFEYFG DLDKFYGTGY GQPLGPVQGV GYINELLARL
P_lycii           PFCDLFT..A EEYVSYEYYY DLDKYYGTGP GNALGPVQGV GYVNELLARL
烟曲霉            PFCQLFT..H NEWKKYNYLQ SLGKYYGYGA GNPLGPAQGI GFTNELIARL
```

图 17B



```
共有植酸酶A                PFCALFT..H DEWRQYDYLQ SLGKYYGYGA GNPLGPAQGV GFANELIARL
构巢曲霉                  PFCAIFT..E KEWLQYDYLQ SLSKYYGYGA GSPLGPAQGI GFTNELIARL
无花果曲霉_NRRL3135        PFCDLFT..H DEWINYDYLQ SLKKYYGHGA GNPLGPTQGV GYANELIARL
土曲霉                    PFCDLFT..A TEWTQYNYLL SLDKYYGYGG GNPLGPVQGV GWANELMARL
嗜热踝节菌                PFCALST..Q EEWQAYDYYQ SLGKYYGNGG GNPLGPAQGV GFVNELIARM
T_lanuginosa              PFCHLFT..A DDWMAYDYYY TLDKYYSHGG GSAFGPSRGV GFVNELIARM
嗜热毁丝菌                PFCRLFS..E SEWRAYDYLQ SVGKWYGYGP GNPLGPTQGV GFVNELLARL
                          PFCSLFT  A SDWTIYDYLQ SLGKYYGFGP GNSLAATQGV GYVNELIARL

                          301                                                 349
                          351                                                 400
P_involtus_A1             TNS.AVRDNT QTNRTLDASP VTFPLNKTFY ADFSHDNLMV AVFSAMGLFR
P_involtus_A2             TNS.AVNDNT QTNRTLDAAP DTFPLNKTMY ADFSHDNLMV AVFSAMGLFR
绒毛栓菌                  TAQ.NVSDHT QTNSTLDSSP ETFPLNRTLY ADFSHDNQMV AIFSAMGLFN
平田头菇                  TEM.PVRDNT QTNRTLDSSP LTFPLDRSIY ADLSHDNQMI AIFSAMGLFN
P_lycii                   TGQ.AVRDET QTNRTLDSDP ATFPLNRTFY ADFSHDNTMV PIFAALGLFN
烟曲霉                    TRS.PVQDHT STNSTLVSNP ATFPLNATMY VDFSHDNSMV SIFFALGLYN
共有植酸酶A                TRS.PVQDHT STNHTLDSNP ATFPLNATLY ADFSHDNSMI SIFFALGLYN
构巢曲霉                  TQS.PVQDNT STNHTLDSNP ATFPLDRKLY ADFSHDNSMI SIFFAMGLYN
无花果曲霉_NRRL3135        THS.PVHDDT SSNHTLDSSP ATFPLKSTLY ADFSHDNGII SILFALGLYN
土曲霉                    TRA.PVHDHT CVNNTLDASP ATFPLNATLY ADFSHDSNLV SIFWALGLYN
嗜热踝节菌                THS.PVQDYT TVNHTLDSNP ATFPLNATLY ADFSHDNTMT SIFAALGLYN
T_lanuginosa              TGNLPVKDHT TVNHTLDDNP ETFPLDAVLY ADFSHDNTMT GIFSAMGLYN
嗜热毁丝菌                A.GVPVRDGT STNRTLDGDP RTFPLGRPLY ADFSHDNDMM GVLGALGAYD
                          I RAPVVDHT TTNSTLDGDE KTFPLNRTVY ADFSHDNDMM NILTALRIFE

                          350                                           383
                          401                                                 450
P_involtus_A1             QPAPLSTSVP NPWR.....T WRTSSLVPFS GRMVVERLSC ..........
P_involtus_A2             QSAPLSTSTP DPNR.....T WLTSSVVPFS ARMAVERLSC ..........
绒毛栓菌                  QSAPLDPTTP DPAR.....T FLVKKIVPFS ARMVVERLDC ..........
平田头菇                  QSSPLDPSFP NPKR.....T WVTSRLTPFS ARMVTERLLC QRDGTGSGGP
P_lycii                   ATA.LDPLKP DENR.....L WVDSKLVPFS GHMTVEKLAC ..........
烟曲霉                    GTEPLSRTSV ESAKE..LDG YSASWVVPFG ARAYFETMQC ..........
共有植酸酶A                GTAPLSTTSV ESIEE..TDG YSASWTVPFG ARAYVEMMQC ..........
构巢曲霉                  GTQPLSMDSV ESIQE..MDG YAASWTVPFG ARAYFELMQC ..........
无花果曲霉_NRRL3135        GTKPLSTTTV ENITQ..TDG FSSAWTVPFA SRLYVEMMQC ..........
土曲霉                    GTAPLSQTSV ESVSQ..TDG YAAAWTVPFA ARAYVEMMQC ..........
嗜热踝节菌                GTAKLSTTEI KSIEE..TDG YSAAWTVPFG GRAYIEMMQC ..........
T_lanuginosa              GTKPLSTSKI QPPTGAAADG YAASWTVPFA ARAYVELLRC ETETSSEEEE
嗜热毁丝菌                GVPPLDKTAR RDPEE..LGG YAASWAVPFA ARIYVEKMRC SGGGGGGGGG
                          HISPMDQTGD DGVKE  RDL FKVSWAVPFA GRVYFEKMVC DADGDGKIDS
                               NTTIPTNYG

                                384                                           425
                          451                                                 500
P_involtus_A1             ........FGT TKVRVLVQDQ VQPLEFCGGD RNGLCTLAKF VESQTFARSD
P_involtus_A2             ........AGT TKVRVLVQDQ VQPLEFCGGD QDGLCALDKF VESQAYARSG
绒毛栓菌                  ........GGA QSVRLLVNDA VQPLAFCGAD TSGVCTLDAF VESQAYARND
平田头菇                  SRIMRNGNVQ TFVRILVNDA LQPLKFCGGD MDSLCTLEAF VESQKYARED
P_lycii                   ........SGK EAVRVLVNDA VQPLEFCGG. VDGVCELSAF VESQTYAREN
烟曲霉                    K..S...EKE PLVRALINDR VVPLHGCDVD KLGRCKLNDF VKGLSWARSG
共有植酸酶A                Q..A...EKE PLVRVLVNDR VVPLHGCAVD KLGRCKRDDF VEGLSFARSG
构巢曲霉                  E......KKE PLVRVLVNDR VVPLHGCAVD KFGRCTLDDW VEGLNFARSG
无花果曲霉_NRRL3135        Q..A...EQA PLVRVLVNDR VVPLHGCPVD ALGRCTRDSF VRGLSFARSG
土曲霉                    R..A...EKE PLVRVLVNDR VMPLHGCPTD KLGRCKRDAF VAGLSFAQAG
嗜热踝节菌                D..D...SDE PVVRVLVNDR VVPLHGCEVD SLGRCKRDDF VRGLSFARQG
```

图 17C

09.11.20

```
T_lanuginosa   E..G...EDE PFVRVLVNDR VVPLHGCRVD RWGRCRRDEW IKGLTFARQG
嗜热毁丝菌     E..GRQEKDE EMVRVLVNDR VMTLKGCGAD ERGMCTLERF IESMAFARGN
               D     EAQK ELVRILVNDR VMRLNGCDAD EQGRCGLEKF VESMEFARRG

                426          439
                501          514
P_involtus_A1  GAGDFEKCFA TSA.
P_involtus_A2  GAGDFEKCLA TTV.
绒毛栓菌       GEGDFEKCFA T...
平田头菇       GQGDFEKCFD ....
P_lycii        GQGDFAKCGF VPSE
烟曲霉         ..GNWGECFS ....
共有植酸酶A    ..GNWAECFA *...
构巢曲霉       ..GNWKTCFT L...
无花果曲霉_NRRL3135 ..GDWAECFA ....
土曲霉         ..GNWADCF. ....
嗜热栊节菌     ..GNWEGCYA ASE.
T_lanuginosa   ..GHWDRCF. ....
嗜热毁丝菌     ..GKWDLCFA ....
                 GEWEECFV
                    R
```

图 17D