



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0029672

(51)⁷ A01H 5/00

(13) B

(21) 1-2013-03687

(22) 25/04/2012

(86) PCT/US2012/034980 25/04/2012

(87) WO 2012/149009 A2 01/11/2012

(30) 61/480,668 29/04/2011 US

(45) 25/10/2021 403

(43) 25/04/2014 313A

(73) BANGLADESH JUTE RESEARCH INSTITUTE (BD)

Manik Mia Avenue, Dhaka, 1207 Bangladesh, People's Republic.

(72) ALAM, Maqsoodul (US); KHAN, Haseena (BD); ZAMAN, Mahboob (BD); UDDIN, Mohammed, Kamal (BD); HAQUE, Mohammed, Samiul (BD); ISLAM, Mohammed, Shahidul (BD); AZAM, Muhammad, Shafiul (BD).

(74) Công ty TNHH Trần Hữu Nam và Đồng sự (TRAN H.N & ASS.)

(54) PHÂN TỬ AXIT NUCLEIC MÃ HÓA ENZYM TỪ CON ĐƯỜNG SINH TỔNG HỢP LIGNIN TRONG CÂY ĐẬY VÀ PHƯƠNG PHÁP TẠO RA THỰC VẬT CHUYỂN GEN CHỨA PHÂN TỬ NÀY

(57) Sáng chế bộc lộ các polynucleotit mã hóa các polypeptit bao gồm con đường sinh tổng hợp lignin trong cây đậu. Sáng chế này nói chung liên quan đến lĩnh vực gen sinh tổng hợp lignin thực vật, các polypeptit được mã hóa bởi các gen đó, và việc sử dụng các trình tự polynucleotit và polypeptit đó để kiểm soát việc tạo ra lignin thực vật. Sáng chế cũng bộc lộ các phương pháp sử dụng các polynucleotit và các polypeptit để tác động lên chất lượng và số lượng sợi được tạo ra từ cây đậu.

Lĩnh vực kỹ thuật được đề cập

Sáng chế liên quan đến việc nhận diện và mô tả đặc điểm các phân khác nhau của con đường sinh tổng hợp lignin trong cây đay. Cụ thể hơn, sáng chế liên quan đến các polynucleotit từ các cây đay mà mã hóa các enzym chịu trách nhiệm tổng hợp lignin, và các phương pháp sử dụng các polynucleotit này và các enzym để điều hòa gen và thao tác tạo ra lignin để mang lại cho các sợi hàm lượng lignin như mong muốn và các đặc tính khác.

Tình trạng kỹ thuật của sáng chế

Lignin là tên gọi chung cho một heteropolyme thơm phức hệ của monolignol (các rượu hydroxycinnamyl), thường được dẫn xuất từ phenylalanin trong một quy trình đa bước. (Whetten, R. và Sederoff, R., (1995) Lignin Biosynthesis, Plant Cell, 7, các trang 1001-1013). Các polyme này, được lắng đọng chủ yếu ở các thành tế bào, đảm bảo độ bền cơ học cần thiết của các thân thực vật và quan trọng nhất là tính kỵ nước của các mô mạch của thực vật. (Vanholme, R. et al. (2010) Lignin biosynthesis and structure, Plant Physiol, 153, các trang 895-905). Do đặc tính kỵ nước, lignin đóng vai trò là thành phần chính của các mô mạch và đóng vai trò thiết yếu trong việc vận chuyển nước. Ngoài vai trò về mặt cấu tạo và vận chuyển, lignin là thành phần chủ yếu của hệ bảo vệ thực vật. (Goujon, T. et al. (2003) Genes involved in the biosynthesis of lignin precursors in Arabidopsis thaliana, Plant Physiology and Biochemistry, 41, các trang 677-687). Điều không đáng bất ngờ là các điều kiện môi trường có ảnh hưởng đến lượng lignin lắng đọng. (Boerjan, W. et al. (2003) Lignin biosynthesis, Annu Rev Plant Biol, 54, các trang 519-546). Ví dụ, sự sinh tổng hợp lignin được gây ra để phản ứng lại các điều kiện căng thẳng khác nhau như thương tích, căng thẳng phi sinh học, và nhiễm mầm bệnh. Lignin hạn chế sự xâm nhập của

mầm bệnh và bảo vệ các polysaccarit ở thành tế bào chống lại sự phân hủy do vi khuẩn. (Vanholme et al., 2010).

Phần lớn quan niệm hiện tại của chúng ta về sự sinh tổng hợp lignin bắt nguồn từ cách hiểu đầy đủ về con đường này ở cây cải xoong tai chuột *A. thaliana* và cây dương đen *P. trichocarpa*. (Goujon, et al., 2003; Shi, et al. (2010) Towards a systems approach for lignin biosynthesis in *Populus trichocarpa*: transcript abundance and specificity of the monolignol biosynthetic genes, *Plant Cell Physiol*, 51, các trang 144-163). Có ba monome monolignol cơ bản: các rượu p-coumaryl, coniferyl và sinapyl. Các monolignol này được đưa vào ba đơn vị lignin, hoặc các khối hợp nhất: p-hydroxyphenyl (H), guaiacyl (G), và syringyl (S). Xem Fig. 1. Các monolignol này khác nhau về số nhóm metoxy. P-hydroxyphenyl (H) không có nhóm metoxy, guaiacyl (G) có một nhóm metoxy, và syringyl (S) có hai nhóm metoxy. (Goujon et al., 2003). Tuy nhiên, ngoài ba monolignol này, một vài phenylpropanoid khác, chẳng hạn như các hydroxycinnamyl aldehyt, hydroxycinnamyl este, và hydroxycinnamyl axetat cũng có thể được đưa vào. (Boerjan et al., 2003).

Sau khi sinh tổng hợp các khối hợp nhất lignin cơ bản này, chúng được vận chuyển đến các vùng hóa gỗ. Trong các vùng hóa gỗ, quá trình polyme hóa diễn ra do sự ghép đôi theo gốc tự do ôxy hóa của các peroxidaza hoặc các lacaza và một cấu trúc giống như mắt lưới được tạo thành bằng cách tạo ra liên kết ngang với xenluloza và hemixenluloza. (Boerjan et al., 2003; Vanholme, R. et al. (2008) Lignin engineering, *Curr Opin Plant Biol*, 11, các trang 278-285). Sự hóa gỗ diễn ra trong các pha khác nhau trong suốt quá trình làm dày thứ cấp thành tế bào khi việc tạo thành ma trận polysaccarit hoàn thành. Sự lắng đọng lignin bị ảnh hưởng bởi tính chất của ma trận polysaccarit. Trong thành tế bào nguyên sinh, nó có dạng các cấu trúc hình cầu; trong khi ở thành tế bào thứ cấp, nó có dạng các lá mỏng. (Boerjan et al., 2003).

Mặc dù lignin có vai trò rất cần thiết trong đời sống thực vật, nó là nhân tố định ra giới hạn chính cho việc sử dụng hiệu quả/sinh lợi nguyên liệu thực vật trong các

ngành bột giấy và nhiên liệu sinh học. Lignin cũng giới hạn việc sử dụng sinh khối để sản xuất sợi, hóa chất và năng lượng. Tách chiết lignin là một quy trình rất tốn kém và các ngành này có thể hưởng lợi từ việc tiếp cận sinh khối có ít lignin hơn, hoặc lignin mà dễ phân hủy. Trong một vài thập kỷ qua, con đường sinh tổng hợp lignin đã được hiểu ở góc độ nào đó, mặc dù các phần của quy trình chưa được hiểu đầy đủ.

Mặc dù sự tổng hợp lignin đóng vai trò quan trọng đối với sự phát triển tổng thể của cây đay, cũng như có ảnh hưởng đến một số khía cạnh của chất lượng sợi, hiện tại chưa có thông tin chi tiết về sự sinh tổng hợp lignin trong cây đay. Do đó, đang có nhu cầu nhận diện, phân tách và sử dụng các gen và các enzym từ cây đay mà có liên quan đến việc sinh tổng hợp lignin. Sáng chế giúp giải quyết nhu cầu này.

Bản chất kỹ thuật của sáng chế

Theo một khía cạnh, sáng chế đề cập đến một phân tử axit nucleic được phân lập có ít nhất 90% tính tương đồng về trình tự so với trình tự axit nucleic được lựa chọn từ nhóm gồm có: SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 16, 18, 20, 22, 24, 25, 26, 28, 29, 31, 33, 35, 37, 39, 40, 42, 44, 45, 47, 49, và 51.

Theo một phương án, một phân tử axit nucleic được phân lập được lựa chọn từ nhóm gồm có: SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, và 15.

Theo một phương án, một phân tử axit nucleic được phân lập được lựa chọn từ nhóm gồm có: SEQ ID NO: 16, 18, và 20.

Theo một phương án, một phân tử axit nucleic được phân lập được lựa chọn từ nhóm gồm có: SEQ ID NO: 22, 24, 25, 26, 28, và 29.

Theo một phương án, một phân tử axit nucleic được phân lập được lựa chọn từ nhóm gồm có: SEQ ID NO: 31.

Theo một phương án, một phân tử axit nucleic được phân lập được lựa chọn từ nhóm gồm có: SEQ ID NO: 33.

Theo một phương án, một phân tử axit nucleic được phân lập được lựa chọn từ nhóm gồm có: SEQ ID NO: 35, 37 và 39.

Theo một phương án, một phân tử axit nucleic được phân lập được lựa chọn từ nhóm gồm có: SEQ ID NO: 40 và 42.

Theo một phương án, một phân tử axit nucleic được phân lập được lựa chọn từ nhóm gồm có: SEQ ID NO: 44, 45, và 47.

Theo một phương án, một phân tử axit nucleic được phân lập được lựa chọn từ nhóm gồm có: SEQ ID NO: 49.

Theo một phương án, một phân tử axit nucleic được phân lập được lựa chọn từ nhóm gồm có: SEQ ID NO: 51.

Theo một khía cạnh, sáng chế đề cập đến một phân tử polypeptit được phân lập có ít nhất 90% tính tương đồng về trình tự so với một trình tự axit amin được lựa chọn từ nhóm gồm có SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 17, 19, 21, 23, 27, 30, 32, 34, 36, 38, 41, 43, 46, 48, 50, và 52.

Theo một phương án, một cặp mỗi xuôi và ngược hữu dụng cho việc khuếch đại cADN được lựa chọn từ nhóm gồm có SEQ ID NO 53 và SEQ ID NO 54; SEQ ID NO 55 và SEQ ID NO 56; SEQ ID NO 57 và SEQ ID NO 58; SEQ ID NO 59 và SEQ ID NO 60; và SEQ ID NO 61 và SEQ ID NO 62.

Theo các phương án nhất định, sáng chế này liên quan đến bất kỳ một trong số các trình tự polynucleotit đã nêu hoặc các trình tự polypeptit, trong đó trình tự đã nêu có ít nhất 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, hoặc 100% tính tương đồng về trình tự so với bất kỳ trong số các trình tự được nhận biết bởi một SEQ ID NO.

Theo một khía cạnh, sáng chế đề cập đến một vector biểu hiện bao gồm một phân tử axit nucleic được phân lập của sáng chế này.

Theo một khía cạnh, sáng chế đề cập đến một kháng thể được phân lập hoặc đoạn liên kết kháng nguyên của nó mà liên kết đặc hiệu với một phân tử polypeptit của sáng chế này.

Theo một khía cạnh, sáng chế đề cập đến một tế bào thực vật chuyển nhiễm được chuyển nhiễm bởi một vector của sáng chế này.

Theo một khía cạnh, sáng chế đề cập đến một nguyên liệu được dẫn xuất từ một thực vật chuyển gen của sáng chế này.

Theo một khía cạnh, sáng chế đề cập đến một hạt từ một thực vật được chuyển nhiễm bởi một vector của sáng chế này.

Theo một khía cạnh, sáng chế đề cập đến phương pháp tạo ra thực vật chuyển gen, bao gồm các bước chuyển nhiễm ít nhất một tế bào thực vật bằng một vector của sáng chế này, và phát triển ít nhất một tế bào thực vật thành một thực vật.

Theo một khía cạnh, sáng chế đề cập đến phương pháp cải thiện sự tăng trưởng, sản lượng sợi, độ bền của sợi, khả năng kháng bệnh, hoặc việc sử dụng nước trong cây đay, bao gồm bước đưa vào cây đay một trình tự axit nucleic phi tự nhiên của sáng chế này.

Mô tả vắn tắt các hình vẽ

Fig. 1: thể hiện con đường sinh tổng hợp monolignol được đề xuất của cây đay;

Fig. 2a và 2b: thể hiện sự sắp xếp thẳng hàng trình tự protein của ColCAD1, ColCAD2, ColCAD3, ColCAD4, ColCAD5, ColCAD6, và ColCAD7 với các trình tự protein CAD thực vật;

Fig. 3: thể hiện sự sắp xếp thẳng hàng trình tự protein của ColCCoAOMT1, ColCCoAOMT2 và ColCCoAOMT3 với các trình tự protein CCoAOMT thực vật;

Fig. 4: thể hiện sự sắp xếp thẳng hàng trình tự protein của Col4CL1, Col4CL4 và Col4CL6 với các trình tự protein 4CL thực vật;

Fig. 5: thể hiện sự sắp xếp thẳng hàng trình tự protein của Col6HCT1 với các trình tự protein 6HCT thực vật;

Fig. 6: thể hiện sự sắp xếp thẳng hàng trình tự protein của ColC3H với các trình tự protein C3H thực vật;

Fig. 7: thể hiện sự sắp xếp thẳng hàng trình tự protein của ColC4H1 và ColC4H2 với các trình tự protein C4H thực vật;

Fig. 8: thể hiện sự sắp xếp thẳng hàng trình tự protein của ColPAL1 và ColPAL2 với các trình tự protein PAL thực vật;

Fig. 9: thể hiện sự sắp xếp thẳng hàng trình tự protein của ColCCR2 với các trình tự protein CCR thực vật;

Fig. 10: thể hiện sự sắp xếp thẳng hàng trình tự protein của ColCCR3 với các trình tự protein CCR thực vật;

Fig. 11: thể hiện sự sắp xếp thẳng hàng trình tự protein của ColF5H với các trình tự protein F5H thực vật;

Fig. 12: thể hiện sự sắp xếp thẳng hàng trình tự protein của ColCOMT với các trình tự protein COMT thực vật;

Fig. 13: thể hiện gel ADN của ColCAD2;

Fig. 14: thể hiện gel ADN của ColCCoAOMT1;

Fig. 15: thể hiện gel ADN của Col4CL1;

Fig. 16: thể hiện gel ADN của ColCCR3;

Fig. 17: thể hiện gel ADN của ColF5H.

Mô tả chi tiết sáng chế

Mười họ enzym đã biết có liên quan đến sinh tổng hợp monolignol. (Goujon và đồng tác giả, 2003). Các họ là PAL (phenylalanin ammonia-lyaza), C4H (cinnamat-4-hydroxylaza), 4CL (4-coumarat:CoA ligaza), HCT (p-hydroxycinnamoyl-CoA:shikimat/quinat p-hydroxycinnamoyl transferaza), C3H (4-coumarat 3-hydroxylaza), CCoAOMT (caffeoyl-CoA O-metyltransferaza), CCR (cinnamoyl-CoA reductaza), F5H (ferulat 5-hydroxylaza), COMT (axit caffeic O-metyltransferaza), và CAD (rượu cinnamyl dehydrogenaza). Một giản đồ được đề xuất của con đường sinh tổng hợp monolignol trong cây đay được thể hiện trên Fig. 1.

Con đường sinh tổng hợp lignin trong cây đay mang tính phức tạp, phần nào đó, đối với sự có mặt của một số enzym đa chức năng, và đối với các enzym cấu thành trải dài qua một số họ gen khác nhau. Enzym thứ nhất của con đường

phenylpropanoid là PAL (phenylalanin ammonia-lyaza), làm khử amin cho phenylalanin, tạo ra axit cinamic. Enzym thứ hai của con đường, C4H (cinnamat 4-hydroxylaza), làm chuyển hóa axit cinamic thành axit 4-hydroxycinamic, được tiếp nối bởi các bước hydroxyl hóa và metyl hóa sau đó khi con đường trở nên phân nhánh. Enzym 4CL gây xúc tác cho việc thất CoA của các axit hydroxycinamic, tạo ra các tiền chất phenolic hoạt hóa cho việc sinh tổng hợp lignin. (Hu et al. (1999) *Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees*, *Nat Biotech*, 17, các trang 808-812.).

Enzym tiếp theo trong con đường monolignol (HCT) gây xúc tác việc tạo ra các p-coumaroyl-shikimat/quinat este, làm các chất nền cho C3H. HCT được thể hiện để chuyển nhóm axyl của p-coumaroyl-CoA sang shikimat hoặc quinat. (Hoffman et al. (2005) *Plant Biosystems*, phiên bản 139, Số 1, các trang 50-53). Các bước hydroxyl hóa ở C3 và C5 được thực hiện tương ứng bởi hai enzym P450 sắc tố tế bào, 4-coumarat 3-hydroxylaza (C3H) và ferulat 5-hydroxylaza (F5H). Các bước metyl hóa được thực hiện bởi CCoAOMT (caffeoyl-coenzym A (CoA) O-metyltransferaza) và COMT (caffeic-Ometyltransferaza). CCoAOMT là một enzym song chức năng chuyển hóa caffeoyl-CoA thành feruloyl-CoA và 5-hydroxyferuloyl-CoA thành sinapoyl-CoA và đóng vai trò trong việc tổng hợp các polysaccarit được feruloyl hóa. (Inoue et al., 1998). CCoAOMT được thể hiện để tham gia vào việc sinh tổng hợp lignin trong các bộ phận ông dẫn khác nhau của cây hoa cúc cánh giấy *Zinnia elegans*. (Ye, Z. H. và Varner J. E. (1995) *Differential expression of two O-methyltransferases in lignin biosynthesis in Zinnia elegans*, *Plant Physiol.* 108, các trang 459-467). CCoAOMT tham gia vào việc củng cố thành tế bào thực vật, và cũng tham gia vào các phản ứng đối với thách thức về thương tích hoặc mầm bệnh do sự tạo thành thêm các polyme axit ferulic được bao bọc bởi thành tế bào.

Các enzym bổ sung tham gia vào con đường sinh tổng hợp monolignol là cinnamoyl coenzym A reductaza (CCR) và rượu cinnamyl dehydrogenaza (CAD). CCR gây xúc tác việc giảm bớt các este hydroxycinnamoyl CoA để tạo ra

cinnamaldehyt, trong khi đó CAD gây xúc tác việc giảm bớt rượu cinnamyl. (Goujon et al., 2003).

Một trong số các enzym cuối cùng tham gia vào con đường monolignol là rượu cinnamyl dehydrogenaza (CAD) mà xúc tác việc chuyển đổi phụ thuộc vào NADPH coniferaldehyt, 5-hydroxy-coniferaldehyt, và sinapaldehyt thành các rượu tương ứng. (Kim, S. J. et al. (2004) Functional reclassification of the putative cinnamyl alcohol dehydrogenase multigene family in Arabidopsis, Proc. Natl. Acad. Sci. USA, 101, các trang 1455-60). Ở cây Arabidopsis, các đột biến đơn lẻ của các gen CAD AtCAD-C và AtCAD-D đã được thấy là có các hoạt tính CAD thấp hơn, và đột biến kép, thu được bằng cách lai hai đột biến, giảm 40% hàm lượng lignin ở thân cây, do đó cho thấy rằng đây là các gen CAD chính tham gia vào việc tổng hợp lignin ở thân cây. (Sibout, R. et al. (2005) Cinnamyl Alcohol Dehydrogenase-C and -D are the primary genes involved in lignin biosynthesis in the floral stem of Arabidopsis, Plant Cell, 17, các trang 2059-76).

Hai enzym là đặc thù cho con đường sinh tổng hợp monolignol. Chúng là axit caffeic O-metyltransferaza (COMT) và cinnamoyl coenzym A reductaza (CCR). COMT được nhận diện lần đầu tiên ở các cây hạt kín. COMT có khả năng chuyển hóa axit caffeic thành axit ferulic, cũng như chuyển hóa axit 5-hydroxyaxit ferulic thành axit sinapic. (Dixon, R. A., et al. (2001) The biosynthesis of monolignols: a "metabolic grid," or independent pathways to guaiacyl and syringyl units? Phytochemistry, 57, các trang 1069-1084). Việc điều tiết giảm gen COMT ở ngô (*Zea mays*) đã được thể hiện là gây ra việc giảm đáng kể hoạt tính COMT (giảm 70 đến 85%), dẫn đến sự thay đổi hàm lượng và thành phần lignin, điều này chỉ ra rằng enzym này là enzym chủ yếu cho việc tổng hợp lignin.

Axit ferulic do COMT tạo ra có thể được hydroxyl hóa bởi ferulat 5 hydroxylaza (F5H), là một monooxygenaza phụ thuộc vào P450 sắc tố tế bào, để tạo thành axit 5-hydroxy-ferulic. F5H cũng có khả năng hydroxyl hóa coniferaldehyt và rượu coniferyl để tương ứng tạo ra 5-hydroxy-coniferaldehyt và rượu 5-hydroxy-

coniferyl. (Meyer, K. et al. (1996) Ferulate-5-hydroxylase from *Arabidopsis thaliana* defines a new family of cytochrome P450-dependent monooxygenases, *Proc. Natl. Acad. Sci. USA*, 93, các trang 6869-74). F5H được cho là một bước giới hạn tỷ lệ trong sinh tổng hợp lignin syringyl, một đề xuất được hỗ trợ bởi phát hiện rằng một đột biến *Arabidopsis* thiếu hụt trong biểu hiện F5H cũng bị ảnh hưởng ở mức độ tích lũy các sinapat este ở các quả cải và hạt. (Ruegger, M. et al. (1999) Regulation of ferulate-5-hydroxylase expression in *Arabidopsis* in the context of sinapate ester biosynthesis, *Plant Physiol.*, 119, các trang 101-10).

Enzym thứ hai đặc thù là tham gia vào việc sinh tổng hợp lignol, CCR, gây xúc tác cho việc chuyển hóa feruloyl CoA và 5-hydroxy-feruloyl CoA tương ứng thành coniferaldehyt và 5-hydroxy-coniferaldehyt. Bước này trực tiếp dẫn đến việc sinh tổng hợp các đơn vị lignin G (coniferaldehyt) và S (5-hydroxy-coniferaldehyt). (Ma et al., 2005). Ở thuốc lá, việc điều tiết giảm gen CCR sử dụng các cấu trúc đối nghĩa đã tạo ra các thực vật phát triển bất bình thường và làm giảm sự sinh trưởng cũng như hình thái học của lá trở nên bất bình thường và các mạch bị vỡ. Cũng có sự giảm có liên quan về mức độ hợp chất lignin G. (Ralph, J. et al. (1998) NMR characterization of altered lignins extracted from tobacco plants down-regulated for lignification enzymes cinnamylalcohol dehydrogenase and cinnamoyl-CoA reductase, *Proc. Natl. Acad. Sci USA*, 95, các trang 12803-8).

Nhận biết bằng máy điện toán các gen và các phiên mã

Đáng chú ý là chúng tôi đã xác định được các trình tự của các enzym cây đay tham gia vào sinh tổng hợp lignin. Con đường sinh tổng hợp lignin đã được mô tả rõ đặc điểm và mỗi enzym được mã hóa bởi một họ gen trong phần lớn các loài thực vật. Đã lấy được tổng số 106 trình tự gen của cây *A. thaliana* và *P. trichocarpa* từ cơ sở dữ liệu bộ gen NCBI và *P. trichocarpa* (http://genome.jgi-psfz.org/Poptr1_1/). (Goujon et al., 2003; Shi et al., 2010). Các gen sinh tổng hợp monolignol cây đay được nhận diện từ các mẫu gen của tổng hợp trình tự gen cây đay quả dài *Corchorus olitorius* và dữ

liệu hệ phiên mã của *C. olitorius* và *C. capsularis* sử dụng chương trình BLASTN với ngưỡng giá trị kỳ vọng là 1e-20. (Altschul, S.F., et al. (1990) Basic local alignment search tool, *J Mol Biol*, 215, các trang 403-410). Các chuỗi tuần tự gen gADN thu được phải trải qua bước dự báo mẫu gen sử dụng phần mềm AUGUSTUS. (Stanke, M. et al. (2004) AUGUSTUS: AUGUSTUS: a web server for gene finding in eukaryotes, *Nucleic Acids Research*, 32, W309-W312). Các mẫu gen và các chuỗi tuần tự phiên mã từ các dữ liệu hệ phiên mã của *C. olitorius* và *C. capsularis* được tìm kiếm theo cơ sở dữ liệu NCBI nr (không dư thừa) để xác nhận thêm. Đối với *C. olitorius*, các chuỗi tuần tự phiên mã được vẽ bản đồ trên các mẫu gen dự báo sử dụng GMAP (với giá trị ngưỡng 95%). (Wu, T.D. và Watanabe, C.K. (2005) GMAP: a genomic mapping and alignment program for mRNA and EST sequences, *Bioinformatics*, 21, các trang 1859-1875).

Một chương trình sắp xếp thẳng hàng trình tự axit amin của các protein giả định được mã hóa bởi các gen ColCAD với các protein CAD khác sẵn có trong cơ sở dữ liệu NCBI, trong đó sử dụng chương trình CLUSTAL W, được thể hiện trên các hình vẽ Fig. 2a và 2b. Dưới đây là danh mục các protein được sắp xếp thẳng hàng với các protein ColCAD giả định, với các số truy cập GeneBank trong các dấu ngoặc đơn: PtcCADL4 (Protein giống như *Populus trichocarpa* rượu cinnamyl dehydrogenaza, CADL4, gi224138226); RcoCAD (Cây đu đủ tía *Ricinus communis* rượu dehydrogenaza, giả định, gi25558709); FraCAD (Cây dâu tây *Fragaria x ananassa*, rượu cinnamyl dehydrogenaza, gi13507210) (Chandler et al. (2002) Cloning, expression and immunolocalization pattern of a cinnamyl alcohol dehydrogenase gene from strawberry (*Fragaria x ananassa*), *J. Exp. Bot.*, 53 (375), các trang 1723-1734); GhiCAD5 (Cây bông *Gossypium hirsutum*, rượu cinnamyl dehydrogenaza 5, gi268528129); PtcCAD (Cây dương đen *Populus trichocarpa*, gi183585165) ((2010) Towards a systems approach for lignin biosynthesis in *Populus trichocarpa*: transcript abundance and specificity of the monolignol biosynthetic genes, *Plant Cell Physiol.*, 51 (1), các trang 144-163); GhiCAD3 (Cây bông *Gossypium hirsutum*, gi229368450)

(Các gen sinh sản vô tính và biểu hiện con đường phenylpropanoid trong việc phát triển sợi bông); và GhiCAD (Cây bông *Gossypium hirsutum*, gi166865124) ((2009) Molecular and biochemical evidence for phenylpropanoid synthesis and presence of wall-linked phenolics in cotton fibers, *J Integr Plant Biol*, 51 (7), các trang 626-637).

Một chương trình sắp xếp thẳng hàng trình tự axit amin của các protein giả định được mã hóa bởi các gen ColCCoAOMT với các protein CCoAOMT khác sẵn có trong cơ sở dữ liệu NCBI, trong đó sử dụng chương trình CLUSTAL W, được thể hiện trên Fig. 3. Dưới đây là danh mục các protein được sắp xếp thẳng hàng với các protein ColCCoAOMT giả định, với các số truy cập GeneBank trong các ngoặc đơn: PtrCCoAOMT (Cây dương lá rung *Populus tremuloides*, gi3023436); GhiCCoAOMT2 (Cây bông *Gossypium hirsutum*, gi229368460); và GhiCCoAOMT1 (Cây bông *Gossypium hirsutum*, gi253509567).

Một chương trình sắp xếp thẳng hàng trình tự axit amin của các protein giả định được mã hóa bởi các gen Col4CL với các protein 4CL sẵn có trong cơ sở dữ liệu NCBI, trong đó sử dụng chương trình CLUSTAL W, được thể hiện trên Fig. 4. Dưới đây là danh mục các protein được sắp xếp thẳng hàng với các protein Col4CL giả định, với các số truy cập GeneBank trong các ngoặc đơn: Ccap4CL1 (Cây đậu quả tròn *Corchorus capsularis*, gi294514718); Rco4CL (Cây đu đủ tía *Ricinus communis*, gi255565415); và Ptc4CL (Cây dương đen *Populus trichocarpa*, gi224074401).

Một chương trình sắp xếp thẳng hàng trình tự axit amin của các protein giả định được mã hóa bởi các gen Col6HCT với các protein Col6HCT khác sẵn có trong cơ sở dữ liệu NCBI, trong đó sử dụng chương trình CLUSTAL W, được thể hiện trên Fig. 5. Dưới đây là danh mục các protein được sắp xếp thẳng hàng các protein Col6HCT giả định, với các số truy cập GeneBank trong các ngoặc đơn: CycarHCT (Cây rau ca đông *Cynara cardunculus*, gi:73671233) ((2007) Isolation and functional characterization of a cADN coding a hydroxycinnamoyltransferase involved in phenylpropanoid biosynthesis in *Cynara cardunculus*, *BMC Plant Biol.* 7, 14); và PtcHCT (Cây dương đen *Populus trichocarpa*, gi183585181).

Một chương trình sắp xếp thẳng hàng trình tự axit amin của các protein giả định được mã hóa bởi các gen ColC3H với các protein C3H sẵn có trong cơ sở dữ liệu NCBI, trong đó sử dụng chương trình CLUSTAL W, được thể hiện trên Fig. 6. Dưới đây là danh mục các protein được sắp xếp thẳng hàng với các protein ColC3H giả định, với các số truy cập GeneBank trong các ngoặc đơn: EglC3H (Cây khuynh diệp *Eucalyptus globulus*, gi:295413824); PtcC3H (Cây dương đen *Populus trichocarpa*, gi:224139664); và PalxPgrC3H (Cây bạch dương *Populus alba* X Cây dương răng lớn *Populus grandidentata*, gi166209291).

Một chương trình sắp xếp thẳng hàng trình tự axit amin của các protein giả định được mã hóa bởi các gen ColC4H với các protein C4H khác sẵn có trong cơ sở dữ liệu NCBI, trong đó sử dụng chương trình CLUSTAL W, được thể hiện trên Fig. 7. Dưới đây là danh mục các protein được sắp xếp thẳng hàng với các protein ColC4H giả định, với các số truy cập GeneBank trong các ngoặc đơn: GarC4H (Cây bông cỏ *Gossypium arborium*, gi9965897) và GarC4H (Cây bông cỏ *Gossypium arborium*, gi9965899).

Một chương trình sắp xếp thẳng hàng trình tự axit amin của các protein giả định được mã hóa bởi các gen ColPAL với các protein PAL khác sẵn có trong cơ sở dữ liệu NCBI, trong đó sử dụng chương trình CLUSTAL W, được thể hiện trên Fig. 8. Dưới đây là danh mục các protein được sắp xếp thẳng hàng với các protein ColPAL giả định, với các số truy cập GeneBank trong các ngoặc đơn: JcoPAL 20 (Cây dầu mè *Jatropha curcas*, gi113203757) và PtrPAL (Cây dương đen *Populus trichocarpa*, gi:183585195).

Một chương trình sắp xếp thẳng hàng trình tự axit amin của các protein giả định được mã hóa bởi các gen ColCCR2 với các protein CCR khác sẵn có trong cơ sở dữ liệu NCBI, trong đó sử dụng chương trình CLUSTAL W, được thể hiện trên Fig. 9. Dưới đây là danh mục các protein được sắp xếp thẳng hàng với các protein ColCCR2 giả định, với các số truy cập GeneBank trong các ngoặc đơn: AthCCR (Cây cải xoong tai chuột *Arabidopsis thaliana*, gi:15237678); CofCCR (Cây trà dầu

Camellia oleifera, gi228480464); và AlyCCR (Cây *Arabidopsis lyrata*, gi:297793385).

Một chương trình sắp xếp thẳng hàng trình tự axit amin của các protein giả định được mã hóa bởi các gen ColCCR3 với các protein CCR khác sẵn có trong cơ sở dữ liệu NCBI, trong đó sử dụng chương trình CLUSTAL W, được thể hiện trên Fig. 10. Dưới đây là danh mục các protein được sắp xếp thẳng hàng với các protein ColCCR3 giả định, với các số truy cập GeneBank trong các ngoặc đơn: RcoCCR (Cây đu đủ tía *Ricinus communis*, gi:255556687) và AthCCR (Cây cải xoong tai chuột *Arabidopsis thaliana*, gi:15226955).

Một chương trình sắp xếp thẳng hàng trình tự axit amin của các protein giả định được mã hóa bởi các gen ColF5H với các protein F5H khác sẵn có trong cơ sở dữ liệu NCBI, trong đó sử dụng chương trình CLUSTAL W, được thể hiện trên Fig. 11. Dưới đây là danh mục các protein được sắp xếp thẳng hàng với các protein ColF5H giả định, với các số truy cập GeneBank trong các ngoặc đơn: EglF5H (Cây khuynh diệp cầu *Eucalyptus globules*, gi:255970299) và PtcF5H (Cây dương đen *Populus tricarpa*, gi:6688937).

Một chương trình sắp xếp thẳng hàng trình tự axit amin của các protein giả định được mã hóa bởi các gen ColCOMT với các protein COMT khác sẵn có trong cơ sở dữ liệu NCBI, trong đó sử dụng chương trình CLUSTAL W, được thể hiện trên Fig. 12. Dưới đây là danh mục các protein được sắp xếp thẳng hàng với các protein ColCOMT giả định, với các số truy cập GeneBank trong các ngoặc đơn: GhiCOMT (Cây bông *Gossypium hirsutum*, gi:253509569) và EcaCOMT (Cây bạch đàn *Eucalyptus camaldulensis*, gi:262474806).

Phân tích môtip của các vùng khởi động

Đối với mỗi trong số các mẫu gen dự báo, cả hai sợi của các vùng ngược hướng 2000 bp được tách chiết và tìm kiếm các trình tự môtip cis theo cơ sở dữ liệu PlantCARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) (Lescot, M.,

et al. (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences, Nucleic Acids Res, 30, các trang 325-327). Nếu bất kỳ phần nào của các trình tự được lựa chọn được thấy là gói lên một gen liền kề thì phần vùng ngược hướng đó không được phân tích thêm. Một danh mục các môtip quan trọng đã được biên soạn được biết đến là tham gia vào phản ứng đối với sự căng thẳng và các quy trình phát triển khác nhau (Bảng 1).

Bảng 1: Danh mục các môtip được tìm thấy trong vùng khởi động của các gen sinh tổng hợp monolignol cây đay

	Mẫu gen	Axit salicylic	Độ phản ứng	Liên kết Myb	Thương tích	Trình tự tăng cường phiên mã	Axit abscisic	Etylen	Auxin	Yếu tố phản ứng với nấm gibberella	Độ phản ứng ở nhiệt độ thấp
PAL	contig_310473_g31	*									
	contig_301819_g4			*							
	contig_302040_g24		*	*			*			*	
	contig_309920_g2				*	*	*				
C4H	contig_311419_g56										
	contig_307848_g46		*					*			
	contig_384302_g42										
4CL	contig_331398_g112										
	contig_299467_g73								*	*	
	contig_337813_g88					*					
	contig_306165_g86									*	*
	contig_309926_g132						*		*		
	contig_304308_g125	*			*	*				*	
	contig_305998_g57	*		*							
HCT	contig_308231_g128						*				
C3H	contig_597141_g180										*
CCoAOMT	contig_306979_g168			*						*	

	contig_308402_g160										
	contig_296218_g172	*	*			*				*	
	contig_311951_g178					*					
CCR	contig_310406_g194								*		
	contig_297048_g201		*					*			
	contig_294439_g187	*									
	contig_309320_g191			*							
F5H	contig_304192_g207		*								
	contig_595270_g204						*				
COMT	contig_299952_g235		*								
CAD	contig_335778_g278							*			
	contig_95812_g280										
	contig_321021_g281		*								
	contig_808760_g282	*									
	contig_808760_g284						*				
	contig_356907_g287								*		
	contig_356907_g288					*		*			
	contig_355903_g289						*				
	contig_306706_g300		*			*				*	
	contig_304576_g307	*	*			*				*	

Các polynucleotit theo sáng chế này được phân lập bằng cách xác định trình tự cao năng các thư viện cADN bao gồm mô cây đay thu được từ cây đay quả dài *Corchorus olitorius* L. Một số polynucleotit theo sáng chế này có thể là các trình tự từng phần, ở chỗ, chúng không thể hiện một gen đủ độ dài mã hóa một polypeptit đủ độ dài. Các trình tự từng phần như thế có thể được kéo dài bằng cách phân tích và xác định trình tự các thư viện ADN khác nhau sử dụng các đoạn mồi và/hoặc các mẫu dò và các kỹ thuật lai giống và/hoặc PCR đã biết rõ. Các trình tự từng phần có thể được kéo dài cho đến khi một khung đọc mở mã hóa một polypeptit, một polynucleotit đủ độ dài, một gen có khả năng biểu hiện một polypeptit, hoặc một phần hữu dụng khác của bộ gen được nhận diện.

Có thể thực hiện việc nhận diện ADN thuộc bộ gen và ADN các loài khác loại bằng các kỹ thuật ADN tiêu chuẩn/lai giống ADN trong các điều kiện nghiêm ngặt thích hợp, trong đó sử dụng tất cả hoặc một phần trình tự trình tự polynucleotit làm mẫu dò để kiểm tra một thư viện thích hợp. Tùy chọn, các kỹ thuật PCR mà sử dụng các đoạn mồi oligonucleotit được thiết kế dựa trên ADN trong bộ gen đã biết, cADN, hoặc các trình tự protein có thể được sử dụng để khuếch đại và nhận diện các trình tự trong bộ gen và cADN.

Các polypeptit theo sáng chế này có thể được tạo ra bằng cách chèn một trình tự polynucleotit theo sáng chế này mà mã hóa polypeptit mong muốn thành một vector biểu hiện và biểu hiện polypeptit trong một vật chủ thích hợp. Bất kỳ loại vector biểu hiện nào mà những người có trình độ bình thường trong lĩnh vực kỹ thuật đã biết đến cũng có thể được sử dụng. Sự biểu hiện có thể đạt được ở bất kỳ tế bào vật chủ thích hợp nào mà đã được biến nạp hoặc được chuyển nhiễm bằng một vector biểu hiện chứa một phân tử polynucleotit mã hóa một polypeptit tái tổ hợp. Các tế bào vật chủ thích hợp bao gồm các sinh vật nhân nguyên thủy (prokaryote), nấm men, và các tế bào có nhân điển hình cao hơn.

Các polypeptit mà bao gồm các con đường sinh tổng hợp lignin được tinh lọc từ cây đay, hoặc được tạo ra bằng các phương pháp tái tổ hợp, có thể được sử dụng để

tạo ra các kháng thể đơn dòng, các đoạn hoặc dẫn xuất kháng thể như được xác định trong tài liệu này, theo các phương pháp đã biết. Các kháng thể nhận biết và liên kết các đoạn polypeptit mà bao gồm các con đường sinh tổng hợp lignin theo sáng chế cũng được dự liệu, miễn là các kháng thể là đặc thù cho các polypeptit mà bao gồm con đường sinh tổng hợp lignin.

Các cấu trúc gen theo sáng chế này cũng có thể chứa một gen đánh dấu chọn lọc mà hữu hiệu trong các tế bào thực vật, để cho phép dò ra các tế bào bị biến nạp có cấu trúc sáng tạo. Các gen đánh dấu như thế, mà đã được biết rõ trong lĩnh vực kỹ thuật, thông thường mang lại sức đề kháng đối với một hoặc nhiều độc tố hoặc tạo ra các dấu hiệu trực quan về sự có mặt của nó dưới kính hiển vi huỳnh quang. Tùy chọn, sự có mặt của cấu trúc mong muốn trong các tế bào bị biến nạp có thể được xác định bằng các kỹ thuật khác mà đã biết rõ trong lĩnh vực kỹ thuật, chẳng hạn như các kỹ thuật Southern và Western blot. Cấu trúc gen theo sáng chế này có thể được liên kết với một vectơ có ít nhất một hệ thống sao chép, chẳng hạn như, *E. coli* hoặc nấm men (*Saccharomyces cerevisiae*), nhờ đó sau mỗi lần thao tác, kết cấu mang lại có thể được nhân bản và xếp trình tự.

Các cấu trúc gen theo sáng chế này có thể được sử dụng để biến nạp nhiều loại thực vật, chẳng hạn như cây một lá mầm (ví dụ, lúa) và cây hai lá mầm (ví dụ, cây đay, *Arabidopsis*). Theo một phương án ưu tiên, các cấu trúc gen có tính sáng tạo được sử dụng để biến nạp cây đay. Như đã trình bày trên đây, việc biến nạp một thực vật với một cấu trúc gen theo sáng chế này có thể được sử dụng để tạo ra hàm lượng lignin thay đổi trong thực vật đó.

Các kỹ thuật đưa các cấu trúc gen một cách ổn định vào bộ gen của các thực vật đích đã được biết đến rõ trong lĩnh vực kỹ thuật và bao gồm việc đưa vào gián tiếp qua vi khuẩn *Agrobacterium tumefaciens*, chuyển gen bằng xung điện, tiêm vào mô phân sinh hoặc các cơ quan sinh sản, tiêm vào các phôi non, và các thao tác tương tự. Việc lựa chọn kỹ thuật sẽ phụ thuộc vào thực vật/mô/vật chủ đích cần biến nạp.

Thuật ngữ "thực vật" bao gồm toàn bộ các thực vật, các cơ quan/cấu trúc sinh dưỡng chồi cây (ví dụ, lá, cuống lá và thân củ), rễ, hoa, và các cơ quan/cấu trúc của hoa (ví dụ, lá bắc, đài hoa, cánh hoa, nhị hoa, lá noãn, bao phấn và noãn), hạt (bao gồm phôi, nội nhũ và vỏ hạt) và quả (bầu nhụy trưởng thành), một thực vật (ví dụ, mô mạch, mô dự trữ, và các bộ phận tương tự) và các tế bào (ví dụ, tế bào bảo vệ, tế bào trứng, cấu trúc hình tóc lông và các bộ phận tương tự), và thế hệ con cháu của chúng. Lớp thực vật có thể sử dụng trong phương pháp theo sáng chế thông thường là rộng như lớp thực vật cao hơn và thấp hơn theo các kỹ thuật biến nạp, bao gồm các cây hạt kín (các cây một lá mầm và cây hai lá mầm), cây hạt trần, cây dương xỉ, rêu và tảo đa bào. Nó bao gồm các thực vật có nhiều mức bội thể, bao gồm thể lệch bội lẻ, thể đa bội, thể nhị bội, thể đơn bội và bán tứ hợp.

Các gel ADN của các phản ứng PCR sử dụng các đoạn mỗi xuôi và ngược cho một số enzym cây đay được thể hiện trên các hình vẽ Fig. 13-17. Ở Fig. 13, gel ADN là của CAD2 từ cây đay quả dài *Corchorus olitorius*. Làn 1 là sản phẩm PCR của CAD2 sử dụng cADN làm bản mẫu. Đoạn mỗi xuôi và đoạn mỗi ngược tương ứng là SEQ ID NO. 53 và 54. Làn 2 là thang 1 Kb +. Ở Fig. 14, gel ADN là của CCoAOMT1 từ cây đay quả dài *Corchorus olitorius*. Làn 1 là thang 1 Kb +, và Làn 2 là sản phẩm PCR của CCoAOMT1 sử dụng cADN làm bản mẫu. Đoạn mỗi xuôi và đoạn mỗi ngược tương ứng là SEQ ID NO. 55 và 56. Ở Fig. 15, gel ADN là của 4CL1 từ cây đay quả dài *Corchorus olitorius*. Làn 1 là thang 1 Kb +, và Làn 2 là sản phẩm PCR của 4CL1 sử dụng cADN làm bản mẫu. Đoạn mỗi xuôi và đoạn mỗi ngược tương ứng là SEQ ID NO. 57 và 58. Ở Fig. 16, gel ADN là của CCR3 từ cây đay quả dài *Corchorus olitorius*. Làn 1 là thang 1 Kb +, và Làn 2 là sản phẩm PCR của CCR3 sử dụng cADN làm bản mẫu. Đoạn mỗi xuôi và đoạn mỗi ngược tương ứng là SEQ ID NO. 59 và 60. Ở Fig. 17, gel ADN là của F5H từ cây đay quả dài *Corchorus olitorius*. Làn 1 là thang 1 Kb +, và Làn 2 là sản phẩm PCR của F5H sử dụng cADN làm bản mẫu. Đoạn mỗi xuôi và đoạn mỗi ngược tương ứng là SEQ ID NO. 61 và 62.

Các định nghĩa

Một tế bào đã được "được biến nạp" hoặc "được chuyển nhiễm" bởi ADN ngoại sinh hoặc khác loại khi ADN đó đã được đưa vào trong tế bào. ADN biến nạp có thể có hoặc có thể không được tích hợp (được liên kết cộng hóa trị) vào trong bộ gen của tế bào. Ở các sinh vật nhân nguyên thủy, nấm men, và các tế bào động vật có vú chẳng hạn, ADN biến nạp có thể được duy trì trên một bộ phận thể bổ sung chẳng hạn như một plasmit. Đối với các tế bào có nhân điển hình, một tế bào được biến nạp ổn định là tế bào mà trong đó ADN biến nạp đã được tích hợp vào nhiễm sắc thể sao cho nó được kế thừa bởi các tế bào con thông qua sao chép nhiễm sắc thể. Việc thực hiện sáng chế này dự liệu rất nhiều tế bào thực vật được biến nạp ổn định.

Một "kết cấu biểu hiện" đề cập đến một cấu trúc axit nucleic, mà khi được đưa vào một vật chủ tế bào, sẽ dẫn đến việc phiên mã và/hoặc dịch mã tương ứng một ARN và/hoặc polypeptit. Kết cấu biểu hiện có thể bao gồm một axit nucleic gồm một trình tự khởi động, có hoặc không có một trình tự chứa các dấu hiệu polyadenyl hóa mARN, và một hoặc nhiều vị trí enzym giới hạn được bố trí xuôi dòng từ gen khởi động cho phép chèn các trình tự gen khác loại. Kết cấu biểu hiện có khả năng điều khiển biểu hiện của một protein khác loại khi gen mã hóa protein khác loại đó được liên kết có thể hoạt động với gen khởi động bằng cách chèn vào một trong số các vị trí giới hạn. Kết cấu biểu hiện tái tổ hợp cho phép biểu hiện protein khác loại trong một vật chủ tế bào khi kết cấu biểu hiện chứa protein khác loại được đưa vào vật chủ tế bào. Các kết cấu biểu hiện có thể được dẫn xuất từ nhiều nguồn khác nhau tùy thuộc vào vật chủ tế bào được sử dụng để biểu hiện. Ví dụ, một kết cấu biểu hiện có thể chứa các thành phần dẫn xuất từ một nguồn virus, vi khuẩn, côn trùng, thực vật hoặc động vật có vú. Trong trường hợp cả hai biểu hiện chuyển gen và ức chế gen nội sinh (ví dụ, bằng cách kim hãm đối nghĩa hoặc có nghĩa), trình tự polynucleotit được chèn vào không cần thiết phải tương đồng và có thể là "gần như tương đồng" với một trình

tự gen mà nó được dẫn xuất. Tốt hơn là, kết cấu biểu hiện tái tổ hợp cho phép biểu hiện ở một giai đoạn nhiễm trùng sớm và/hoặc nó cho phép biểu hiện trong gần như tất cả các tế bào của một sinh vật, chẳng hạn như thực vật. Có thể tìm thấy các ví dụ về các kết cấu biểu hiện phù hợp cho việc biến nạp thực vật trong Patent Mỹ số 5 880 333 và 6 002 072; các Công bố Đơn quốc tế số WO/1990/002189 và WO/2000/026388; Ainley và Key (1990) *Plant Mol. Biol.*, 14, các trang 949-967; và Birch (1997) *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 48, các trang 297-326, mà tất cả các nội dung của chúng được đưa vào đây để tham chiếu.

Thuật ngữ "vật chủ tế bào" đề cập đến một tế bào từ bất kỳ sinh vật nào. Các vật chủ tế bào ưu tiên được dẫn xuất từ thực vật, vi khuẩn, nấm men, nấm, côn trùng, hoặc các động vật khác. Thuật ngữ "vật chủ tế bào tái tổ hợp" (hay đơn giản là "vật chủ tế bào") đề cập đến một tế bào mà một vector biểu hiện tái tổ hợp đã được đưa vào đó. Nên hiểu rằng thuật ngữ "vật chủ tế bào" được nhằm để đề cập đến không chỉ tế bào cụ thể đang nói tới mà cả thế hệ con cháu của tế bào đó. Do những biến đổi nhất định có thể xảy ra trong các thế hệ tiếp theo do đột biến hoặc các ảnh hưởng của môi trường, trong thực tế các thế hệ con cháu như thế có thể không tương đồng với tế bào cha mẹ, những vẫn thuộc phạm vi của thuật ngữ "vật chủ tế bào" như được sử dụng trong tài liệu này. Các phương pháp đưa các trình tự polynucleotit vào các loại vật chủ tế bào khác nhau đã được biết rõ trong lĩnh vực kỹ thuật. Sáng chế đề xuất các tế bào vật chủ hoặc thế hệ con cháu của các tế bào vật chủ được biến nạp có các kết cấu biểu hiện tái tổ hợp của sáng chế này. Các tế bào vật chủ có thể là các tế bào thực vật. Tốt hơn là, các tế bào thực vật là các tế bào cây đay.

Thuật ngữ "được liên kết có thể hoạt động" hoặc "được chèn có thể hoạt động" có nghĩa là các trình tự điều hòa cần thiết cho việc biểu hiện trình tự mã hóa được đặt trong một phân tử axit nucleic ở các vị trí thích hợp so với trình tự mã hóa để có thể biểu hiện trình tự mã hóa. Định nghĩa đó đôi khi được áp dụng cho việc bố trí các bộ phận điều chỉnh phiên mã (ví dụ, các trình tự tăng cường) trong một kết cấu biểu hiện. Các trình tự điều chỉnh phiên mã và dịch mã là các trình tự điều hòa ADN, chẳng hạn

như các gen khởi động, trình tự tăng cường, dấu hiệu polyadenyl hóa, bộ phân giới, và các bộ phận tương tự, mà quy định việc biểu hiện một trình tự mã hóa trong một vật chủ tế bào.

Các thuật ngữ "gen khởi động", "vùng khởi động" hay "trình tự khởi động" nói chung đề cập đến các vùng điều hòa phiên mã của gen, mà có thể tìm thấy ở phía 5' hoặc 3' của vùng mã hóa, hoặc trong vùng mã hóa, hoặc trong các đoạn intron. Thông thường, một gen khởi động là một vùng điều hòa ADN có khả năng liên kết ARN polymeraza trong một tế bào và khởi động phiên mã một trình tự mã hóa xuôi dòng (hướng 3'). Trình tự khởi động 5' đặc thù bị giới hạn tại phần cuối 3' của nó bởi vị trí khởi động phiên mã và kéo dài ngược dòng (hướng 5') để có số gốc hoặc bộ phận tối thiểu cần thiết cho việc khởi động phiên mã ở các mức độ có thể dò được phía trên nền. Trong trình tự khởi động là một vị trí khởi động phiên mã (được xác định một cách thuận lợi bằng cách vẽ bản đồ với nucleaza S1), cũng như các vùng liên kết protein (các trình tự điều hòa) chịu trách nhiệm liên kết ARN polymeraza.

Thuật ngữ "cấu trúc axit nucleic" hay "cấu trúc ADN" đôi khi được sử dụng để đề cập đến một hoặc các trình tự mã hóa được liên kết có thể hoạt động với các trình tự điều hòa thích hợp và được chèn vào một kết cấu biểu hiện để biến nạp một tế bào. Thuật ngữ này có thể được sử dụng hoán đổi với thuật ngữ "biến nạp ADN" hay "chuyển gen". Một kết cấu axit nucleic như vậy có thể chứa một trình tự mã hóa cho một sản phẩm gen được quan tâm, cùng với một gen đánh dấu có thể chọn và/hoặc một gen chỉ thị. Thuật ngữ "gen đánh dấu có thể chọn" đề cập đến một gen mã hóa một sản phẩm mà, khi được biểu hiện, mang lại kiểu hình có thể chọn, chẳng hạn như khả năng đề kháng kháng sinh, trên một tế bào được biến nạp. Thuật ngữ "gen chỉ thị" đề cập đến một gen mã hóa một sản phẩm mà dễ dò ra bằng các phương pháp tiêu chuẩn, trực tiếp hoặc gián tiếp.

Một vùng "khác loại" của một cấu trúc axit nucleic là một đoạn (hoặc các đoạn) có thể nhận diện của phân tử axit nucleic trong một phân tử lớn hơn mà không được tìm thấy ở phân tử lớn hơn trong tự nhiên. Khi vùng khác loại mã hóa một gen thực

vật, gen sẽ thường bị áp sùan bởi ADN mà không nằm áp sùan ADN bộ gen thực vật trong bộ gen của sinh vật nguồn. Trong một ví dụ khác, một vùng khác loại là một cấu trúc mà ở đó bản thân trình tự mã hóa không được tìm thấy trong tự nhiên (ví dụ, một cADN mà ở đó trình tự mã hóa bộ gen chứa các đoạn intron, hoặc các trình tự tổng hợp có các thể chuông khác với gen nguyên thể). Các biến thể alen hoặc các sự kiện đột biến xảy ra một cách tự nhiên không làm phát sinh vùng khác loại của ADN như được xác định trong tài liệu này. Thuật ngữ "cấu trúc ADN" cũng được sử dụng để đề cập đến một vùng khác loại, cụ thể là vùng khác loại được tạo ra để sử dụng trong biến nạp tế bào.

Thuật ngữ "vector" nhằm đề cập đến một phân tử axit nucleic có khả năng vận chuyển một axit nucleic khác mà nó được liên kết. Một loại vector là "plasmid" mà đề cập đến một thông lọng ADN sợi kép hình tròn mà các đoạn ADN bổ sung có thể được thắt vào đó. Một loại vector là vector virus, mà ở đó các đoạn ADN bổ sung có thể được thắt vào bộ gen virus. Các vector nhất định có khả năng sao chép tự chủ trong một vật chủ tế bào mà chúng được đưa vào (ví dụ, các vector vi khuẩn có nguồn gốc sao chép vi khuẩn và các vector động vật có vú thể bổ sung). Các vector khác có thể được tích hợp vào bộ gen của một vật chủ tế bào khi đưa vào vật chủ tế bào, và nhờ đó được sao chép cùng với bộ gen vật chủ. Ngoài ra, các vector nhất định có khả năng điều khiển việc biểu hiện các gen mà chúng được liên kết có thể hoạt động. Các vector như thế được đề cập ở đây là "các vector biểu hiện tái tổ hợp" (hoặc đơn giản là "các vector biểu hiện"). Nói chung, các vector biểu hiện hữu dụng trong các kỹ thuật ADN tái tổ hợp thường ở dạng các plasmid. Trong bản mô tả này, "plasmid" và "vector" có thể được sử dụng hoán đổi cho nhau bởi vì plasmid là dạng vector thường được sử dụng nhất. Tuy nhiên, sáng chế nhằm bao gồm các dạng vector biểu hiện khác như thế, chẳng hạn như các vector virus (ví dụ, các retrovirus khiếm khuyết sao chép, virus adeno và các virus liên quan đến adeno), là các vector có các chức năng tương đương.

"Phần trăm tính tương đồng về trình tự" được xác định bằng cách so sánh hai trình tự được sắp xếp thẳng hàng một cách tối ưu trên một cửa sổ so sánh, mà ở đó

phần trình tự polynucleotit trong cửa sổ so sánh có thể bao gồm các phần bổ sung hoặc các phần đứt đoạn (tức là, các khe hở) so với trình tự tham chiếu (là trình tự không bao gồm các phần bổ sung hoặc các phần đứt đoạn) để sắp xếp thẳng hàng tối ưu hai trình tự này. Phần trăm được tính toán bằng cách xác định số vị trí mà tại đó gốc axit nucleic tương đồng hoặc gốc axit amin tìm thấy ở cả hai trình tự để tạo ra số vị trí tương ứng, chia số vị trí tương ứng cho tổng số các vị trí trong cửa sổ so sánh, và nhân kết quả với 100 để cho ra phần trăm tính tương đồng về trình tự.

Thuật ngữ "tính gần như tương đồng" của các trình tự polynucleotit nghĩa là một polynucleotit bao gồm một trình tự mà có ít nhất 25% tính tương đồng về trình tự so với một trình tự tham chiếu như được xác định bằng cách sử dụng các chương trình được mô tả trong tài liệu này; tốt hơn là BLAST sử dụng các thông số tiêu chuẩn, như được mô tả. Tùy chọn, phần trăm tính tương đồng có thể là bất kỳ số nguyên nào từ 25% đến 100%. Các phương án ưu tiên hơn bao gồm các trình tự polynucleotit mà có ít nhất: 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% hoặc 99% tính tương đồng về trình tự so với một trình tự tham chiếu. Các giá trị này có thể được điều chỉnh một cách thích hợp để xác định tính tương đồng tương ứng của các protein được mã hóa bởi hai trình tự nucleotit bằng cách xem xét sự suy biến của thể chuồng, tính tương tự của axit amin, việc định vị khung đọc, và các khía cạnh tương tự.

Thuật ngữ "gần như tương đồng" của các trình tự axit amin (và của các polypeptit có các trình tự axit amin này) thông thường có nghĩa là tính tương đồng về trình tự của ít nhất 40% so với một trình tự tham chiếu như được xác định bằng cách sử dụng các chương trình được mô tả trong tài liệu này; tốt hơn là BLAST sử dụng các thông số tiêu chuẩn, như đã được mô tả. Phần trăm tính tương đồng ưu tiên của các axit amin có thể là bất kỳ số nguyên nào từ 40% đến 100%. Các phương án ưu tiên hơn bao gồm các trình tự axit amin mà có ít nhất 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, hoặc 99% tính tương đồng về trình tự so với một trình tự tham

chiều. Các polypeptit mà "gần như tương đồng" dùng chung các trình tự axit amin như được lưu ý trên đây ngoại trừ các vị trí gốc mà không tương đồng thì có thể khác nhau do những thay đổi axit amin bảo thủ. Những thay thế axit amin bảo thủ đề cập đến khả năng hoán đổi các gốc có các mạch bên tương tự. Ví dụ, một nhóm axit amin có các mạch bên béo là glyxin, alanin, valin, leuxin, và isoleuxin; một nhóm axit amin có các mạch bên hydroxyl béo là serin và threonin; một nhóm axit amin có các mạch bên chứa amit là asparagin và glutamin; một nhóm axit amin có các mạch bên thơm là phenylalanin, tyrosin, và tryptophan; một nhóm axit amin có các mạch bên bazơ là lysin, arginin, và histidin; và một nhóm axit amin có các mạch bên chứa lưu huỳnh là cystein và methionin. Các nhóm thay thế axit amin bảo thủ ưu tiên là: valin-leuxin-isoleuxin, phenylalanin-tyrosin, lysin-arginin, alanin-valin, axit aspactic-axit glutamic, và asparagin-glutamin.

Nội dung đưa vào để tham chiếu

Tất cả các patent Mỹ, các đơn sáng chế Mỹ đã được công bố, và các đơn PCT đã được công bố có chỉ định Mỹ được viện dẫn trong tài liệu này được đưa vào đây để tham chiếu.

Các phương án tương đương

Mặc dù một số phương án của sáng chế đã được mô tả và minh họa trong tài liệu này, những người có hiểu biết trung bình trong lĩnh vực kỹ thuật sẽ dễ dàng hình dung ra nhiều phương tiện và/hoặc kết cấu khác để thực hiện các chức năng và/hoặc thu được các kết quả và/hoặc một hoặc nhiều ưu điểm được mô tả trong tài liệu này, và mỗi trong số các biến thể và/hoặc cải biến như thế được cho là nằm trong phạm vi của sáng chế này. Người có hiểu biết trung bình trong lĩnh vực kỹ thuật sẽ nhận thấy, hoặc có thể biết chắc việc không phải sử dụng thêm thực nghiệm thông thường, nhiều phương án tương đương cho các phương án cụ thể của sáng chế được mô tả trong tài liệu này. Do đó, cần hiểu rằng các phương án trên đây được trình bày chỉ để minh họa

và trong phạm vi yêu cầu bảo hộ kèm theo và các phương án tương đương của chúng; sáng chế có thể được thực hiện khác với những gì đã được mô tả cụ thể và yêu cầu bảo hộ.

Các đơn có liên quan

Đơn này xin hưởng quyền ưu tiên từ Đơn Sáng chế Tạm thời Mỹ số 61/480 668 nộp ngày 29/4/2011.

YÊU CẦU BẢO HỘ

1. Phân tử axit nucleic được phân lập có ít nhất 95% tính tương đồng về trình tự so với một trình tự axit nucleic được lựa chọn từ nhóm gồm có: SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, trong đó phân tử axit nucleic mã hóa một loại rượu cinnamyl dehydrogenaza.
2. Phân tử axit nucleic được phân lập theo điểm 1, trong đó phân tử đã nêu có ít nhất 98% tính tương đồng về trình tự so với trình tự axit nucleic đã nêu.
3. Phân tử axit nucleic được phân lập theo điểm 1, trong đó phân tử đã nêu có ít nhất 99% tính tương đồng về trình tự so với trình tự axit nucleic đã nêu.
4. Phân tử axit nucleic được phân lập theo điểm 1, trong đó phân tử đã nêu có 100% tính tương đồng về trình tự so với trình tự axit nucleic đã nêu.
5. Phân tử polypeptit được phân lập có ít nhất 90% tính tương đồng về trình tự so với một trình tự axit amin được lựa chọn từ nhóm gồm có SEQ ID NO: 2, 4, 6, 8, 10, 12 và 14.
6. Phân tử polypeptit được phân lập theo điểm 5, trong đó phân tử đã nêu có ít nhất 95% tính tương đồng về trình tự so với trình tự axit amin đã nêu.
7. Phân tử polypeptit được phân lập theo điểm 5, trong đó phân tử đã nêu có ít nhất 98% tính tương đồng về trình tự so với trình tự axit amin đã nêu.
8. Phân tử polypeptit được phân lập theo điểm 5, trong đó phân tử đã nêu có ít nhất 99% tính tương đồng về trình tự so với trình tự axit amin đã nêu.
9. Phân tử polypeptit được phân lập theo điểm 5, trong đó phân tử đã nêu có 100% tính tương đồng về trình tự so với trình tự axit amin đã nêu.
10. Cặp đoạn mồi xuôi và ngược hữu dụng cho việc khuếch đại cADN được lựa chọn từ nhóm gồm có SEQ ID NO 53 và SEQ ID NO 54.
11. Vector biểu hiện bao gồm phân tử axit nucleic được phân lập theo điểm 1.
12. Tế bào thực vật được chuyển nhiễm bao gồm vector theo điểm 11.
13. Phương pháp tạo ra thực vật chuyển gen, bao gồm các bước:

chuyển nhiễm ít nhất một tế bào thực vật với một vector biểu hiện theo điểm 11; và phát triển ít nhất một tế bào thực vật nêu trên thành cây.

14. Phương pháp cải thiện sự sinh trưởng, sản lượng sợi, độ bền của sợi, khả năng kháng bệnh, hoặc sử dụng nước trong cây đay, bao gồm việc đưa vào cây đay một phân tử axit nucleic phi tự nhiên theo điểm 1.

15. Phân tử axit nucleic được phân lập theo điểm 1, trong đó trình tự axit nucleic đã nêu là SEQ ID NO: 1.

16. Phân tử axit nucleic được phân lập theo điểm 1, trong đó trình tự axit nucleic đã nêu là SEQ ID NO: 3.

17. Phân tử axit nucleic được phân lập theo điểm 1, trong đó trình tự axit nucleic đã nêu là SEQ ID NO: 5.

18. Phân tử axit nucleic được phân lập theo điểm 1, trong đó trình tự axit nucleic đã nêu là SEQ ID NO: 7.

19. Phân tử axit nucleic được phân lập theo điểm 1, trong đó trình tự axit nucleic đã nêu là SEQ ID NO: 9.

20. Phân tử axit nucleic được phân lập theo điểm 1, trong đó trình tự axit nucleic đã nêu là SEQ ID NO: 11.

21. Phân tử axit nucleic được phân lập theo điểm 1, trong đó trình tự axit nucleic đã nêu là SEQ ID NO: 13.

22. Phân tử axit nucleic được phân lập theo điểm 1, trong đó trình tự axit nucleic đã nêu là SEQ ID NO: 15.

23. Phân tử polypeptit được phân lập theo điểm 5, trong đó trình tự axit amin đã nêu là SEQ ID NO: 2.

24. Phân tử polypeptit được phân lập theo điểm 5, trong đó trình tự axit amin đã nêu là SEQ ID NO: 4.

25. Phân tử polypeptit được phân lập theo điểm 5, trong đó trình tự axit amin đã nêu là SEQ ID NO: 6.

26. Phân tử polypeptit được phân lập theo điểm 5, trong đó trình tự axit amin đã nêu là SEQ ID NO: 8.
27. Phân tử polypeptit được phân lập theo điểm 5, trong đó trình tự axit amin đã nêu là SEQ ID NO: 10.
28. Phân tử polypeptit được phân lập theo điểm 5, trong đó trình tự axit amin đã nêu là SEQ ID NO: 12.
29. Phân tử polypeptit được phân lập theo điểm 5, trong đó trình tự axit amin đã nêu là SEQ ID NO: 14.

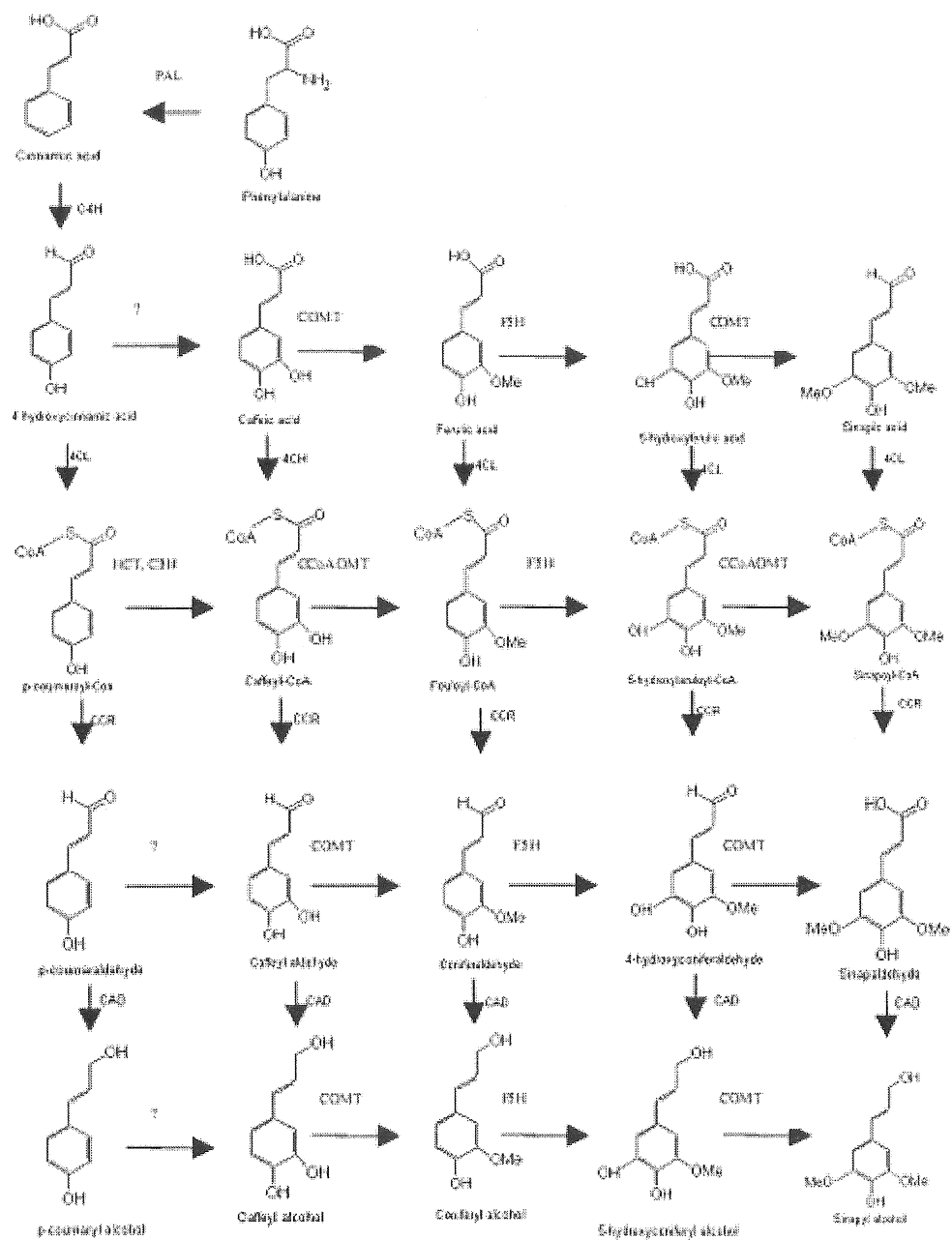


Fig. 1

```

ColCAD1      -MGSLEPEQEHFKAFGWAARDTSGHLSPTFKSRRATGEKDVAFKVLVYCGICHSDLHMVFN 59
ColCAD2      -MSRLPEEEHFNKAFGWAARDTSGVLSPTFKSRRATGEKDVAFKVLVYCGICHSDLHMVFN 59
PtcCADgi_224138226
ColCAD3      -----MIFN 4
RcoCADgi_255587709
FraCADgi_13507210
ColCAD5      MVANLPEKDHFKAFGWAARDQSGVLSPTFTSRRATGEKDVSPKVLVYCGMCHSDLHMVFN 60
PtcCADgi_183585165
ColCAD6      ---MSIEQEHFNKASGWAARDSSGVLSPTFNFSRRATGEKDVAFKVLVYCGICHSDHLMVFN 57
GhiCAD3gi_229368450
ColCAD7      MAKSSPEEEHFKAFGWAARDTSGHLSPTFNFSRRATGEKDVAFKVLVYCGICHSDLHFIIN 60
GhiCADgi_166865124
ColCAD4      MSK3-PEEEHFKAFGWAARDQSGHLSPTFNFSRRATGEEDVAFKVLVYCGICHSDLHSIKN 59
ChiCAD5gi_268528129
-----MGSLETERTTTGWAARDPSGVLSPTTYTLRNTGPEDEVFKVICCGICHTDLHQAKN 56
-----MGSLETERTTTGWAARDPSGVLSPTTYTLRNTGPEDEVFKVMCCGICHTDLHQAKN 56
-----MDSQTKSDNCLGWAARDPSGVLSPTTYFSRRPLGSDNDSIKITHCGVCYADFIWERN 56
-----MDSQTKTENCLGWAARDTSGVLSPTFKSRRPLGSDNDSIKITHCGVCYADFIWERN 56
: :

ColCAD1      ENGTAIYPMVPGHEIVGEVTEVGSKEVKEFKVGDVAGVGVNSCRSCDNAMNLENYCPQ 119
ColCAD2      ENGVSIIYPLVPGHEIVGEVTEVGSKEVQKFKVGDVGVGCMVGSCHSCDSCNMNLENYCPK 119
PtcCADgi_224138226
ColCAD3      ENGVTQYPLIPGHEIVGVVTEVGSKEVKEFKVGDVGVGCMVGSCHSCDSCDNMLENYCSK 120
RcoCADgi_255587709
FraCADgi_13507210
ColCAD5      DWGYSIYPLVPGHEIVGEVTEVGSKEVQKFKVGDVGVGCMVGSCHSCDSCDNMLENYCPK 64
PtcCADgi_183585165
ColCAD6      EWGTSTYPLVPGHEIVGVVTEVGSKEVKEFKVGDVGVGCMVGSCHSCDSCNMNLENYCPK 120
GhiCAD3gi_229368450
ColCAD7      EWGFTSYPLVPGHEIVGEVTEVGSKEVQKFKVGDVGVGCMVGSCHSCDSCNMNLENYCPK 117
GhiCADgi_166865124
ColCAD4      EWNFSIYPLVPGHEIVGEVTEVGSKEVQKFKVGDVGVGCMVGSCHSCDSCDNMLENYCPK 120
ChiCAD5gi_268528129
DWGFEMVPLVPGHEIVGEVTEVGSKEVKEFKVGDVGVGCMVGSCHSCDSCDNMLENYCPK 119
DWGITMYFVVPVPGHEITGVITKIGNNVNFKVGDVGVGCLAAASCLECEFCCKSQENYCDQ 118
DWGITMYFVVPVPGHEITGVITKIGNNVNFKVGDVGVGCLAAASCLECEFCCKSQENYCDQ 118
DLGMSNYPMVPGHEIVGEVTEVGSKEVQKFKVGDVGVGCMVGSCHSCDSCDNMLENYCPK 116
DLGMSNYPMVPGHEIVGEVTEVGSKEVQKFKVGDVGVGCLVGCCECCNCRPCDSDNEQYCLX 116
KHGDSMYFVVPVPGHEIAGVVKVWGVNHFVGVGDFVGVGVNVCSCDSCDNMLENYCPK 116
MFGDSIYPLVPGHEIAGVVKVWGVNHFVGVGDFVGVGVNVCSCDSCDNMLENYCPK 116
: : ++:++++:++ : :+..+ .....+ +..+ ..+ + + + + + + :

ColCAD1      -AVFTYAAKNYDSTITYGGYSDTMVADEHFIIRIPDTLPLDAAAPLLCAGITVYSPLRYF 178
ColCAD2      -MILTYGARYYDSTITYGGYSDTMVADEHFIIRIPENPLDAAAPLLCAGITVYSPLRYF 178
PtcCADgi_224138226
ColCAD3      -KILTYGARYYDSTITYGGYSDTMVADEHFIIRIPNNPLDAGAPLLCAGITVYSPLRYF 179
RcoCADgi_255587709
FraCADgi_13507210
ColCAD5      -MIPTCGAKYHDSITITYGGYSDTMVADEHFIIRIPNNPLDAAAPLLCAGITVYSPLRYF 123
PtcCADgi_183585165
ColCAD6      -MILTYGARYYDSTITYGGYSDTMVADEHFIIRIPNNPLDAGAPLLCAGITVYSPLRYF 179
GhiCAD3gi_229368450
ColCAD7      -QILTYGARYYDSTITYGGYSDTMVADEHFIIRIPNNPLDAGAPLLCAGITVYSPLRYF 176
GhiCADgi_166865124
ColCAD4      -AIRTNGTYDSTITYGGYSDTMVADEHFIIRIPNNPLDAGAPLLCAGITVYSPLRYF 179
ChiCAD5gi_268528129
-MILTYASITYHDSITITYGGYSDTMVADEHFIIRIPNNPLDAGAPLLCAGITVYSPLRYF 178
-IQPTYNGIFWDSITITYGGYSDTMVADEHFIIRIPNNPLDAGAPLLCAGITVYSPLRYF 177
-IQPTYNGIFWDSITITYGGYSDTMVADEHFIIRIPNNPLDAGAPLLCAGITVYSPLRYF 177
-KIMSNDVYDTSKPTQGGFAGSMVVDQKFKVVKIPEGMMAEQVAPLLCAGITVYSPLRHF 175
-KIMSNDVYDTSKPTQGGFAGSMVVDQKFKVVKIPEGMMAEQVAPLLCAGITVYSPLRHF 175
-IWLTFCNIDEDSTITYGGYSDTMVADEHFIIRIPNNPLDAGAPLLCAGITVYSPLRHF 175
GPVLTFCNIDEDSTITYGGYSDTMVADEHFIIRIPNNPLDAGAPLLCAGITVYSPLRHF 176
: : ++. + ++ : : ..: + ..+ . +++++ +..+:

ColCAD1      Q-LDKPGLHVGUVGLGSLGHMAVFAKAMGAKVTVISTSPNKKKEALENLGADSFVLSAE 237
ColCAD2      G-LDKPGLHVGUVGLGSLGHMAVFAKAMGAKVTVISTSPNKKKEALENLGADSFVLSAE 237
PtcCADgi_224138226
ColCAD3      G-LDKPGLHVGUVGLGSLGHMAVFAKAMGAKVTVISTSPNKKKEALENLGADSFVLSAE 238
RcoCADgi_255587709
FraCADgi_13507210
ColCAD5      G-LDKPGLHVGUVGLGSLGHMAVFAKAMGAKVTVISTSPNKKKEALENLGADSFVLSAE 238
PtcCADgi_183585165
ColCAD6      G-LDKPGLHVGUVGLGSLGHMAVFAKAMGAKVTVISTSPNKKKEALENLGADSFVLSAE 238
GhiCAD3gi_229368450
ColCAD7      G-LDKPGLHVGUVGLGSLGHMAVFAKAMGAKVTVISTSPNKKKEALENLGADSFVLSAE 238
GhiCADgi_166865124
ColCAD4      G-LDKPGLHVGUVGLGSLGHMAVFAKAMGAKVTVISTSPNKKKEALENLGADSFVLSAE 238
ChiCAD5gi_268528129
N-MNQFGKSLGVIGLGLGHMAVFAKAMGAKVTVISTSPNKKKEALENLGADSFVLSAE 238
: : + +:++++:++ ..: + ..+ . +++++ +..+:

```

Fig. 2a

ColCAD1		QDQLQTAMGTMDGIIIDTVSAHPHLLPLIGLKLKSHAKILLVGLPKDKPLELVFPMIIGRRT	297
ColCAD2		QQOIQAAAMDITLDGIIIDTVSAQHPILEFLIGMLKTNGKVLVVGAPKPELPFAFFLLGXRRRL	297
PtcCADgi_224138226		QDCMQAAMGTILDDGIIDTVSAHVHLIELVALLKSHGKLVLVGAPEKPPELPVFVEFLITGRKT	298
ColCAD3		EDQLKAZKGTMNGIVDTVSAGKHDLQELLGLLKONGHKLVLGVPVKPYELPAASLIILGRKL	242
RcoCADgi_255557709		QDCMGGAITGMTDGIIDTVSAGMPLSPFLIGLKSDBGKVLVVGAPKPELPFAFFLLIGGRKL	298
FraCADgi_13507210		QDHMQAAITGMTDGIIDTVSAQHPILEFLIGLKLKSHGKLMVGAPKEPPELPVFVELDMGRMT	295
ColCAD5		QEEMQAAMGTILDGIIIDTVSAVHPIMPELLGLLKSHGKLDWVGAPIEPPELPVFSLIMGART	298
PtcCADgi_183555165		QEEMQAAMGTILDGIIIDTVSAVHPILLEFLGLLKS HGKLLVVGAPKPELPFAFSLIAGRRT	297
ColCAD6		TEQMQRGKRTELVDIVDTVSARKSLGPILLELLKVNGTLVVVGAFDRPIELPSFPFLIFGKRA	297
GhiCAD3gi_229368450		AESMQIKTADSDLYIIDTVVFVPHLEPYLSVLKLDGSKILLTGVIINTPLQFVTMMMLGREY	294
ColCAD7		AEGMKRAADSLEDYIIDTVVFVPHLEPYLSVLKFDGKILLTGVIINTPLQFVSTMMMLGRES	294
GhiCADgi_166865124		QECMKGLSGKSLDFIVDTASGDHPDPFYMSLLKIAGVYVLVGFP-SEVKFPSASLNLMGRT	293
ColCAD4		QECMKGLSGKSLDFIIDDASGDHPDPFYLSLLKSGAVGYALVGFP-SEIKFPSASLNFGMRT	294
ChiCAD5gi_268528129		. . : : . : + : + : . . + : + : . . + : + : . . + : + : . . + : + : . .	
ColCAD1		VAGSGVGGIIEETQEMMFNAKKYDLKPDIEVIPVDYVNNTAMERLVRGVGVKYRFVIDIGNTL	357
ColCAD2		VAGSMIGGMKETQEMIDFAAKHNKIPADIEVIPMDYVNNTAMERLLKADVRYRFVIDIGNTL	357
PtcCADgi_224138226		VGGSBCVGGIKETQEMIDFAAKHNLTADIEVIPMDYVNNTAMERVLKADVRYRFVIDGVGTL	358
ColCAD3		VGGSBNVGGILEETQEMIDFAAKHNVTANVEVIPMDYVNNTAFERLAKADVRYRFVIDIGNTL	302
RcoCADgi_255557709		VGSSCGICGMKETQEMIDFAAKHSITADIEVIPANKYVNNTAMERMLKADVRYRFVIDIGNTL	358
FraCADgi_13507210		VGSSCIGGGMKETQEMIDFAAKHNITADIEVIPIDYLNTAMERLVKADVRYRFVIDIGNTL	355
ColCAD5		MAGSGIGGMKETQEMIDFAAKHNKITADIEVIPMDYVNNTAMERLERLGNADVRYRFVIDIGNTL	358
PtcCADgi_183555165		VAGSGSIGGMKETQEMIDFAAKHNITADIEVISTDYLNNTAMERLANNDVRYRFVIDOVGTL	357
ColCAD6		VGSMITGGMKETQEMMDVCGGHNITCDIEVFKPDQINEALDRLSKNNDVRYRFVIDIAGRS	357
GhiCAD3gi_229368450		VKGSMTGGMKETQEMMDVCGGHNITCDVLEIKPDQINQALDRLRNNDVRYRFVIDIAGTS	354
ColCAD7		ITGSPVGSMEKETEEMLDCKEKNLNSMIEVVKMDDYINTAMERLERNNDVRYRFVVDVAGSK	354
GhiCADgi_166865124		ITGSPVGSMEKETEEMLFCKEENLTSMBIEVMMDYINTAMERLERNNDVRYRFVVDVAGSK	354
ColCAD4		ISGSVTTGGSVKIQEMIDFCAAHEVYPQIEVIPQIYANEALERLERKADVRYRFVIDIENTSL	353
ChiCAD5gi_268528129		FAGSVTTGGTFMIQEMIFGCARRKIYPSQIEVIPQIYANEALERLERKADVRYRFVIDIENTSL	354
ColCAD1	KATSS	362	
ColCAD2	KFTP-	361	
PtcCADgi_224138226	KPDV-	362	
ColCAD3	KTSS-	306	
RcoCADgi_255557709	KPGH-	362	
FraCADgi_13507210	KASS-	359	
ColCAD5	ATTXP	363	
PtcCADgi_183555165	AATXP	362	
ColCAD6	KL---	359	
GhiCAD3gi_229368450	KL---	359	
ColCAD7	LE---	356	
GhiCADgi_166865124	LQ--	357	
ColCAD4	K----	354	
ChiCAD5gi_268528129	K----	355	

Fig. 2b

```

ColCCoAOMT2          -MATN-TQEQQTQAGRHOEVGHKSLLOSDALYQYILETSVYPPEPEPMKE 48
GhiCCoAOMT2gi_229368460 -MATNTTQEQQPAAGRHOEVGHKSLLOSDALYQYILETSVYPPEPEPMKE 49
PtrCCoAOMTgi_3023436   -MATN-GEEQQSQAGRHOEVGHKSLLOSDALYQYILETSVYPPEPEPMKE 48
ColCCoAOMT1          -MAPT-QAEQQTQASRHQEVGHKSLLOSDKLYQYILETSVYPPEPEPMKE 48
ColCCoAOMT3          MGSTGETQFTPTQVSDDEANLFAMQLASASVLPMLVLSAIELDLLEVMK 50
GhiCCoAOMT1gi_253509567 MGSIGETQMTPTQVSDDEANLFAMQLASASVLPMLVLSAIELDLLEIMAK 50
      :      . . . . :      * * :      : * :      * * :

ColCCoAOMT2          LRELTAHP-----WNLMTTSADEGQFLNMLLKLINAKNIMEIGVY----- 89
GhiCCoAOMT2gi_229368460 LRELTAHP-----WNLMTTSADEGQFLNMLLKLINAKNIMEIGVY----- 90
PtrCCoAOMTgi_3023436   LREVTAKHP-----WNIMTTSADEGQFLNMLLKLINAKNIMEIGVY----- 89
ColCCoAOMT1          LRELTAHP-----WNLMTTSADEGQFLNMLLKLINAKNIMEIGVY----- 89
ColCCoAOMT3          AGPGAEFLSPTEVASQLPTKNPDAPVMLDRILRLASYSILTCSLRLNLDG 100
GhiCCoAOMT1gi_253509567 AGPGAEFLSPKEVASKLPTTNPDAPVMLDRILRLASYNVLTCSLRTFPGG 100
      :      *      : : * . . *      : * : * : : : : : :

ColCCoAOMT2          -----TGYSLLATALALPDDGKILAMD--INRENYELG-LPVIQKAGVAH 131
GhiCCoAOMT2gi_229368460 -----TGYSLLATALALPDDGKILAMD--INRENYELG-LPVIRKAGVAH 132
PtrCCoAOMTgi_3023436   -----TGYSLLATALALPEDGKILAMD--INRENYELG-LPVIQKAGVAH 131
ColCCoAOMT1          -----TGYSLLATALALPEDGKILAMD--INRENYELG-LPVIQKAGVAH 131
ColCCoAOMT3          KVERLYGLGPVCKYLVKNEDGVALSALNLMNQDKVLMESWYYLKDAVLEG 150
GhiCCoAOMT1gi_253509567 KVERLYGLGPVCKFLTRNEDGVTLALSALNMNQDKVLMESWYYLKDAVLDG 150
      * . . . . * . : * * * :      : * : : : : : : : :

ColCCoAOMT2          KIEFKEGPALFVLDBKLVEDEKNHGSYDFIFVDAD-----KDNVINYHK 174
GhiCCoAOMT2gi_229368460 KIEFKEGPAMPVLDBKLVEDEKNHGSYDFIFVDAD-----KDNVINYHK 175
PtrCCoAOMTgi_3023436   KIDFKEGPALFVLDBQMIEDGKYHGSEDFIFVDAD-----KDNVINYHK 174
ColCCoAOMT1          KIDFKEGPALFVLDBQMIAGTYHGTFDFIFVDAD-----KDNVINYHK 174
ColCCoAOMT3          GIPFNKAYGMTAFEYHGTDPRFNKVFENRGMSDHSTITMKILETYDGFEG 200
GhiCCoAOMT1gi_253509567 GIPFNKAYGMTAFEYHGTDPRFNKVFENRGMSDHSTITMKILDITYDGFQ 200
      * * : : . . : : : : : : : : : * . . . . * . . .

ColCCoAOMT2          RLIDLWKVGG-----LIGYDNT 191
GhiCCoAOMT2gi_229368460 RLIELWKVGG-----LIGYDNT 192
PtrCCoAOMTgi_3023436   RLIELWKVGG-----LIGYDNT 191
ColCCoAOMT1          RLIELWKVGG-----VIGYDNT 191
ColCCoAOMT3          -LKTLDVVGCGVGATLNMIVSKHPSIKGINFDBLPHVIEDAPALP-GVEHV 248
GhiCCoAOMT1gi_253509567 -LKTLDVVGCGGTGATLSMIVSKYPTIKGINFDBLPHVIEDAPSPVGVGVHV 249
      * * . . * * *      * : : :

ColCCoAOMT2          LWNGSVVAPP-DAPLRKYVRYR-----DFVLELNKALAADP--RIEIC 232
GhiCCoAOMT2gi_229368460 LWNGSVVAPP-DAPLRKYVRYR-----DFVLELNKALAADP--RIEIC 233
PtrCCoAOMTgi_3023436   LWNGSVVAPP-DAPMRKYVRYR-----DFVLELNKALAADP--RIEIC 232
ColCCoAOMT1          LWNGSVVAPP-DAPLRKYVLYR-----DFVLELNKALAADP--RIEIC 232
ColCCoAOMT3          GGDMEFVSVPGDAIFMKWICHDSDEHCVKFLKCYEALPDNGKVIVAE 298
GhiCCoAOMT1gi_253509567 GGDMEFVSVPGDAIFMKWICHDSDEHCAKFLKNCYEALPDNGKVIVAE 299
      :      * . * * * : * : : :      : * : : * : : : *

ColCCoAOMT2          MLPVGDGITLCRRIK----- 247
GhiCCoAOMT2gi_229368460 MLPVGDGITLCRRVK----- 248
PtrCCoAOMTgi_3023436   MLPVGDGITLCRRIQ----- 247
ColCCoAOMT1          QLPVGDGITLCRRIK----- 247
ColCCoAOMT3          ILPDYPDASLATKLVVHIDCIMLAHNPGGKERTAKEFEALAKGAGFQGFQ 348
GhiCCoAOMT1gi_253509567 ILPDYPDPSLATKLVVHIDCIMLAHNPGGKERTAKEFEALAKGAGFQGFQ 349
      **      . : * . : :

ColCCoAOMT2          -----
GhiCCoAOMT2gi_229368460 -----
PtrCCoAOMTgi_3023436   -----
ColCCoAOMT1          -----
ColCCoAOMT3          VKCCAFGTYIMEFLKTV 365
GhiCCoAOMT1gi_253509567 ITCSAFGTNIMEFLKSV 366

```

Fig. 3

[illegible]

Fig. 4

[illegible]

Fig. 5

[illegible]

Fig. 6

```

GarC4Hgi_9965897      MDLLFLEKALLGLFVAVVLAITISKLRGKRFKLPPGSLPVFVFGNWLQVGDOLNHRNLTD 60
GhiC4H2gi_268528127  MDLLFLEKALLGLFVAVVLAITISKLRGKRFKLPPGSLPVFVFGNWLQVGDOLNHRNLTD 60
ColC4H1               MDLLFLEKALLGLFVAVVLAITISKLRGKRFKLPPGSLPVFVFGNWLQVGDOLNHRNLTD 60
ColC4H2               MDLLFLEKALLGLFVAVVLAITISKLRGKRFKLPPGSLPVFVFGNWLQVGDOLNHRNLTD 60
*****:***:***:*****:*****:*****:*****

GarC4Hgi_9965897      LARKFGDIFLLMGQSNLWVVSPELAKVLTQGVFSGRTANVVFDFITGKGQDMVTT 120
GhiC4H2gi_268528127  LARKFGDIFLLMGQSNLWVVSPELAKVLTQGVFSGRTANVVFDFITGKGQDMVTT 120
ColC4H1               LARKFGDIFLLMGQSNLWVVSPELAKVLTQGVFSGRTANVVFDFITGKGQDMVTT 120
ColC4H2               LARKFGDIFLLMGQSNLWVVSPELAKVLTQGVFSGRTANVVFDFITGKGQDMVTT 120
*****:*****:*****:*****:*****:*****

GarC4Hgi_9965897      VYGEHWKRMRRIMTVPFFTNKVQQYRFGWEDEAARVVEDVRKNPEAATNGIVLRRLQL 180
GhiC4H2gi_268528127  VYGEHWKRMRRIMTVPFFTNKVQQYRFGWEDEAARVVEDVRKNPEAATNGIVLRRLQL 180
ColC4H1               VY-----VQQYRFGWEDEAARVVEDVRKNPEAATNGIVLRRLQL 161
ColC4H2               VYGEHWKRMRRIMTVPFFTNKVQQYRFGWEDEAARVVEDVRKNPEAATNGIVLRRLQL 180
**          *****:***:***:*****:*****:*****

GarC4Hgi_9965897      MMYNMYRIMFDRFESEDDPLFVRLKALNGERSRLAQSFYNYGDFIFILRPFIRGYLK 240
GhiC4H2gi_268528127  MMYNMYRIMFDRFESEDDPLFVRLKALNGERSRLAQSFYNYGDFIFILRPFIRGYLK 240
ColC4H1               MMYNMYRIMFDRFESEDDPLFVRLKALNGERSRLAQSFYNYGDFIFILRPFIRGYLK 221
ColC4H2               MMYNMYRIMFDRFESEDDPLFVRLKALNGERSRLAQSFYNYGDFIFILRPFIRGYLK 240
*****:*****:*****:*****:*****:*****

GarC4Hgi_9965897      ICKEVKDRRLQLFRDHFVEERKLGSTKSMNNDGLKCAIDHILDAQQGEINEDNVLYIV 300
GhiC4H2gi_268528127  ICKEVKDRRLQLFRDHFVEERKLGSTKSMNNDGLKCAIDHILDAQQGEINEDNVLYIV 300
ColC4H1               ICKEVKDRRLQLFRDHFVEERKLGSTKSMNNDGLKCAIDHILDAQQGEINEDNVLYIV 281
ColC4H2               LCKEVKDRRLQLFRDHFVEERKLGSTKSMNNDGLKCAIDHILDAQQGEINEDNVLYIV 300
:*****:*****:***:***:*****:***:***:*****:*****

GarC4Hgi_9965897      ENINVAAIETTLWSIEWGIAELVNHFETQKRLRHELDTVLGPQNQITEPDTHKLPYLQAV 360
GhiC4H2gi_268528127  ENINVAAIETTLWSIEWGIAELVNHFETQKRLRHELDTVLGPQNQITEPDTHKLPYLQAV 360
ColC4H1               ENINVAAIETTLWSIEWGIAELVNHFETQKRLRHELDTVLGPQNQITEPDTHKLPYLQAV 341
ColC4H2               ENINVAAIETTLWSIEWGIAELVNHFETQKRLRHELDTVLGPQNQITEPDTHKLPYLQAV 360
*****:*****:***:***:*****:***:***:*****:*****

GarC4Hgi_9965897      IKETLRLRMAIPLLVPBMNLHDAKLGGYDIPAESKILVNANWLANNPANWKNPEEFRPER 420
GhiC4H2gi_268528127  IKETLRLRMAIPLLVPBMNLHDAKLGGYDIPAESKILVNANWLANNPANWKNPEEFRPER 420
ColC4H1               IKETLRLRMAIPLLVPBMNLHDAKLGGYDIPAESKILVNANWLANNPANWKNPEEFRPER 401
ColC4H2               IKETLRLRMAIPLLVPBMNLHDAKLGGYDIPAESKILVNANWLANNPANWKNPEEFRPER 420
*****:*****:*****:*****:*****:*****

GarC4Hgi_9965897      FFEESAKVEANGNDFRYLPFGVGRRCPCGIILALPILGITLGRVQNFELLFPFGQSQID 480
GhiC4H2gi_268528127  FFEESAKVEANGNDFRYLPFGVGRRCPCGIILALPILGITLGRVQNFELLFPFGQSQID 480
ColC4H1               FFEESAKVEANGNDFRYLPFGVGRRCPCGIILALPILGITLGRVQNFELLFPFGQSQID 461
ColC4H2               FFEESAKVEANGNDFRYLPFGVGRRCPCGIILALPILGITLGRVQNFELLFPFGQSQID 480
*****:*****:*****:*****:*****:*****

GarC4Hgi_9965897      TTEKGGQFSLHILEHSTIVAKPRQF 505
GhiC4H2gi_268528127  TTEKGGQFSLHILEHSTIVAKPRQF 505
ColC4H1               TSEKGGQFSLHILEHSTIVL----- 481
ColC4H2               TSEKGGQFSLHILEHSTIVAKPRVF 505
+:*****:*****:*****:*****:*****:*****

```

Fig. 7

```

ColPAL1
JcoPALgi_113203757
PtcPALgi_183585195
ColPAL2
-----
MATIIGNGHNGSGLEGLCITRDPLSWGVAAESMRGSHLDEVKMWSEYRKPLVVLGGETL 60
MEFCQDSBRNGNGSLG--FNTNDPLNWGMRAESLKGSHLDEVKRMIEEYRKPFVVLGGETL 58
-----

ColPAL1
JcoPALgi_113203757
PtcPALgi_183585195
ColPAL2
-----
TVAQVAAIASHDAGVTVVLAESARAGVKAASDQWMDSMNKGTDSYGVITGPGATSHRATK 120
TIGQVTAIASRDWGVVMVLESEEARAGVKAASDQWMDSMNKGTDSYGVITGPGATSHRATK 118
-----MLEFSAKALS----- 10

ColPAL1
JcoPALgi_113203757
PtcPALgi_183585195
ColPAL2
-----
-----MLVRLNTLLQGYSGIRFEILEAIT 24
QGRLQRELIREFLNAGIFGNGTETCHTLPHSATRAAMLVRLNTLLQGYSGIRFEILEAIT 180
QGSELQRELIREFLNAGIFGNGTESHTLPRSATRAAMLVRLNTLLQGYSGIRFEILEAIT 178
-----RAHTLPHTATRAAMLVRLNTLLQGYSGIRFEILEAIT 47
*****:*****

ColPAL1
JcoPALgi_113203757
PtcPALgi_183585195
ColPAL2
-----
KFLNQWITPCLELRGTITASGDVPLSYIAGLLTGRPNKAVGPNGESINAEERFNRAGI 84
KLLNHNITPCLELRGTITASGDVPLSYIAGLLTGRPNKAVGPNGESINAEERFNRAGI 240
KLLNHNITPCLELRGTITASGDVPLSYIAGLLTGRPNKAVGPNGESINAEERFNRAGI 238
KFLNQWITPCVFLRGSITASGDVPLSYIAGLLTGRPNKAVGPNGESINAEERFNRAGI 107
*:*:*****:*****:*****:*****:*****:*****:*****:*****

ColPAL1
JcoPALgi_113203757
PtcPALgi_183585195
ColPAL2
-----
ESGFFTLQPKRGLALVNGTAVSGMARMVLFANILAVLSEVLSAIPAEVMQKKEPTDH 144
DSGFFELQPKRGLALVNGTAVSGGLASMLVFANVLAIVLSEILSAIPAEVMQKKEPTDH 300
DGGFFELQPKRGLALVNGTAVSGGLASMLVFANVLAIVLSEVLSAIPAEVMQKKEPTDH 298
HGGFFELQPKRGLALVNGTAVSGGLASMLVFANVLAIVLSEVLSAIPAEVMQKKEPTDH 167
*:*:*****:*****:*****:*****:*****:*****:*****:*****

ColPAL1
JcoPALgi_113203757
PtcPALgi_183585195
ColPAL2
-----
LTHRLKHHHPGQIERAAIMEHILDGSGYVKAARKLHEMDPLQPKQDRYALRTSPQWLGPQ 204
LTHRLKHHHPGQIERAAIMEHILDGSGYVKAARKLHEMDPLQPKQDRYALRTSPQWLGPQ 360
LTHRLKHHHPGQIERAAIMEHILDGSGYVKAARKLHEMDPLQPKQDRYALRTSPQWLGPQ 358
LTHRLKHHHPGQIERAAIMEHILDGSGYVKAARKLHEMDPLQPKQDRYALRTSPQWLGPQ 227
*****:*****:*****:*****:*****:*****:*****:*****

ColPAL1
JcoPALgi_113203757
PtcPALgi_183585195
ColPAL2
-----
IEVIRPATKSIEREINSVMDNPLIDVSRNKALHGGNFQGTPIGVMDNARLAIASIGKLM 264
IEVIRPATKSIEREINSVMDNPLIDVSRNKALHGGNFQGTPIGVMDNARLAIASIGKLM 420
IEVIRPATKSIEREINSVMDNPLIDVSRNKALHGGNFQGTPIGVMDNARLAIASIGKLM 418
IEVIRPATKSIEREINSVMDNPLIDVSRNKALHGGNFQGTPIGVMDNARLAIASIGKLM 287
*****:*****:*****:*****:*****:*****:*****:*****

ColPAL1
JcoPALgi_113203757
PtcPALgi_183585195
ColPAL2
-----
FAQFSELVNDYNNGLPSNLSGGRNPSLOYGFKGAEIAMSAYCSELQFLANPVTNHVQSA 324
FAQFSELVNDYNNGLPSNLSGGRNPSLOYGFKGAEIAMSAYCSELQFLANPVTNHVQSA 480
FAQFSELVNDYNNGLPSNLSGGRNPSLOYGFKGAEIAMSAYCSELQFLANPVTNHVQSA 478
FAQFSELVNDYNNGLPSNLSGGRNPSLOYGFKGAEIAMSAYCSELQFLANPVTNHVQSA 347
*****:*****:*****:*****:*****:*****:*****:*****

ColPAL1
JcoPALgi_113203757
PtcPALgi_183585195
ColPAL2
-----
EQHNQDVNSLGLISARKTSEAVDILKIMSSTYLVALCQAIDLRHLEENLNNTVNTVSIQI 384
EQHNQDVNSLGLISARKTSEAVDILKIMSSTYLVALCQAIDLRHLEENLNNTVNTVSIQI 540
EQHNQDVNSLGLISARKTSEAVDILKIMSSTYLVALCQAIDLRHLEENLNNTVNTVSIQI 538
EQHNQDVNSLGLISARKTSEAVDILKIMSSTYLVALCQAIDLRHLEENLNNTVNTVSIQI 407
*****:*****:*****:*****:*****:*****:*****:*****

ColPAL1
JcoPALgi_113203757
PtcPALgi_183585195
ColPAL2
-----
AKRVLTGANGELHPSRFCEKDLLKAVDREYVFAYIDDPSCSATYFLMQRLRQVLVEHALT 444
AKRVLTGANGELHPSRFCEKDLLKAVDREYVFAYIDDPSCSATYFLMQRLRQVLVEHALT 600
AKRVLTGANGELHPSRFCEKDLLKAVDREYVFAYIDDPSCSATYFLMQRLRQVLVEHALV 598
AKRVLTGANGELHPSRFCEKDLLKAVDREYVFAYIDDPSCSATYFLMQRLRQVLVEHALM 467
*:*:*****:*****:*****:*****:*****:*****:*****

ColPAL1
JcoPALgi_113203757
PtcPALgi_183585195
ColPAL2
-----
NGESEKNASTSIFQRIAAFEELKTLTPKEVESARVALENGSNVAVPNRIKECRSYPLYK 504
NGESEKNASTSIFQRIAAFEELKTLTPKEVESARVALENGSNVAVPNRIKECRSYPLYK 659
NGESEKNASTSIFQRIAAFEELKTLTPKEVESARVALENGSNVAVPNRIKECRSYPLYK 657
NGESEKNASTSIFQRIAAFEELKTLTPKEVESARVALENGSNVAVPNRIKECRSYPLYK 526
*:*:*****:*****:*****:*****:*****:*****:*****

ColPAL1
JcoPALgi_113203757
PtcPALgi_183585195
ColPAL2
-----
FVREELGTGLLTGEKVSFSGEEFDRVFTAMCQSKLIDEMLECKLEWNGAPLPIC 558
FVREELGTGLLTGEKVSFSGEEFDRVFTAMCQSKLIDEMLECKLEWNGAPLPIC 713
FVREELGTGLLTGEKVSFSGEEFDRVFTAMCQSKLIDEMLECKLEWNGAPLPIC 711
FVREELGTGLLTGEKVSFSGEEFDRVFTAMCQSKLIDEMLECKLEWNGAPLPIC 580
*****:*****:*****:*****:*****:*****:*****:*****

```

Fig. 8

```

ColCCR2      -MSK-----QGEAVCVTGASGAIGSNLVRLLLRGYTVHGTVRNIDDERETHHLEALEG 53
CofCCRgi_228480464 -MSSNTRAGGQGQVVCVTGGSGFIGSNLVRLLLRGYTVHATVKDLDEXETHHLEALEG 59
AthCCRgi_15237678 MLTD-----EREVVCVTGASGCIGSNLVHQLLLRGYSVHATVKNLQDEXETHHLEALEG 54
AlyCCRgi_297793385 -MST-----EREVVCVTGASGCIGSNLVHLLHRRGYSVHATVKNLQDEXETHHLEALEG 53
      ::      : :.+++++.++ ++++++: ++ ++++++.+++:.+++++.+++++

ColCCR2      AESRLRLRFQIDLLDYDAISAAIEGCAGVFHLASFCIVDQVHDPQKELLDBPAINGTLNVLT 113
CofCCRgi_228480464 AESRLRLRFQIDLLDYDSIVAAVTGSSGVFHLASFCIVDQVKDPERELLEBPARGTLNVLT 119
AthCCRgi_15237678 AATRLHLFEMDLLQYDTVSAaingcsgvfhlaspfcivdevqdpqkqlldbpavgtinvtl 114
AlyCCRgi_297793385 AATRLHLFEMDLLQYDTVSAAVNGCsgvfhlaspfcivdevqdpqkqlldbpavgtinvtl 113
      + :+::+::+::+::+::+:: ++: +: ++++++ +::+::+::+::+::+::+::

ColCCR2      AAKELGVRVWVVTSSVSSTIPSENNWPAKIKTEDCWTDIDYCRQNELWYBISKTIAEKAA 173
CofCCRgi_228480464 AAKELGVRVWVVTSSNTAITPSENNWPAKVNNEDCWTDVEYCKQNGLWYPLSKTLAEKAA 179
AthCCRgi_15237678 AAKEASVRVWVVTSSISAITPSENNWPAKIKNEECWAAEDYCRQNGLWYPLSKTLAEKAA 174
AlyCCRgi_297793385 AAKEAGVRVWVVTSSISAITPSENNWPAKIKNEECWADQDYCKQNGLWYPLSKTLAEKAA 173
      +::+::+::+::+::+::+:: +::+::+::+::+::+::+::+:: +::+::+::+::+::+::+::

ColCCR2      WEFSEKERGLDQVVVWNPQTVMGSPNIPFTLTASMMMLLRLLQGCTETETQDFFMGSVHFEDVA 233
CofCCRgi_228480464 WEFAKEKGLDQVVVWNPQTVMSPPIPPALNASMIMLLRFLQSCTEIYENFFMGFVHFEDVA 239
AthCCRgi_15237678 WEPAEERGLDQVVVWNPQTVMGSPVIPPALNASMMMLLRLLQGCTETETENFFMGSVHFEDVA 234
AlyCCRgi_297793385 WEPAEQKGLDQVVVWNPQTVMGSPVIPPALNASMIMLLLRLLQGCTETETENFFMGSVHFEDVA 233
      +::+::+::+::+::+::+:: +::+::+::+::+::+::+::+:: +::+::+::+::+::+::+::

ColCCR2      LAHILVYENPSASGRHMCLEAISHYGDFAKVAELYPYNNVPSLPRDTQPGLLRAKNGGQ 293
CofCCRgi_228480464 LAHILVYENTSATGRHLCVEAISHYGDFTAMVAELYPYNNVPSLPRDTQPGLLRATKDGSK 299
AthCCRgi_15237678 LAHILVYEDPYSGRHLCEAISHYGDFAKVAELYPYNNVPSLPRDTQPGLLRDRNAK 294
AlyCCRgi_297793385 LAHILVYENPSAKGRHLCVEAISHYGDFAKVAELYPYNNVPSLPRDTQGLLRAKNAK 292
      +::+::+::+::+::+::+:: +::+::+::+::+::+::+::+:: +::+::+::+::+::+::+::

ColCCR2      KIMDLGLEFIFMEQIIKDAVESLKSRGLI- 322
CofCCRgi_228480464 KIMDLGFQFIFMEQIIKETVESLKSRGYIS 329
AthCCRgi_15237678 KIIDLGLKFISMEEIIKEGVESLKSRGFIS 324
AlyCCRgi_297793385 KIMELGLEFSSMEDIIKEGVESLKSRGFIS 323
      ++::+::+::+::+::+::+:: +::+::+::+::+::+::+::+::

```

Fig. 9

```

ColCCR3
RooCCRgi_255556687
AthCCRgi_15226955
-----
MATQNKKEAVCVTGANGFIGSWLIQTLLQHGYYTTHASITYPASDFSHLFHLISSSSHGDI 60
----MAKETVCVTGANGFIGSWIIRTLIERGYTKIHASITYPGSDPTLLQLPGSDSR--- 53

ColCCR3
RooCCRgi_255556687
AthCCRgi_15226955
-----
INLELYEADLLDYDAICKAVEGCGGVFHVASPTCLEEEKDFEKEIWLPAVQSTINVLAA 120
--IKIFEADLLDSDAISRAIDGCCAGVFHVASPTCLDPFVDFPEKELVEPAVKGTINVLAA 111

ColCCR3
RooCCRgi_255556687
AthCCRgi_15226955
-----
AIVFNPNWNPQTNGAFDETSWTDLEYCKSRQKWYFVSTMAEKTA 45
RKFEKRRVWLTSSISALVFNPNWP--AGKVFDESSWTDLDYCKSRQKWYFVSKSLAEKAA 178
KRFNVRVWLTSSISALVFNPNWP--EKVVFDESSWSDLDYCKSRQKWYFISKTLAEKAA 169
      +:+++++      +:+++:+++:+++++++:+++:+++++

ColCCR3
RooCCRgi_255556687
AthCCRgi_15226955
-----
WEFAEKHGMDVVAINPATCLGFLLPQNLNASCAYLLQLLEGSKDTQEHYHNLGAVHVRDVA 105
WEFAEKHGMDVVAIHPTCIGFLLPQNLNASCAYLLQLLEGSKDTQEHYHNLGAVHVRDVA 238
WEPSEKHGTNIVTIHPSTCLGFLLPQNLNASCAYLLQLLEGSKDTQEHYHNLGAVHVRDVA 229
+++:+:+++:+++:+++:+++:+++:+++:+++:+++:+++:+++:+++:+++:+++:+++:

ColCCR3
RooCCRgi_255556687
AthCCRgi_15226955
-----
KQILLFESPSASGRYLCTNGIYQFGTFAETVSHLFPQYFVHRPTGDTQEGLVSKKDAK 165
KQVLLFEPASASGRYLCTNGIYQFGDFADKVSKLFPFEPVHSPFIGETQPGLTTCCKDAK 298
KGVMLFETPDASGRFLCTNGIYQFSEFALVSKLPPEFAVHKFDEKETQPGLTSCNDAK 289
+:::+++:+:::+++:+++:+++:+++:+++:+++:+++:+++:+++:+++:+++:+++:

ColCCR3
RooCCRgi_255556687
AthCCRgi_15226955
-----
RLIELGLIFTFVEEAVRETVESLQARGFLKQQQPSSES 202
RLIELGLVFTFVEDAVGESVESLQARGFLKHKT-SES 334
RLIELGLVFTAVEDAVKETVQSLRDKGFL----- 318
+++++++:+++:+++:+++:+++:+++:+++:+++:+++:+++:+++:+++:+++:+++:

```

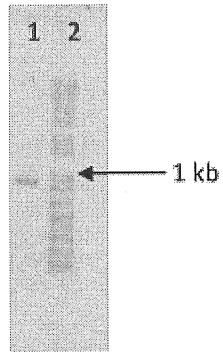
Fig. 10

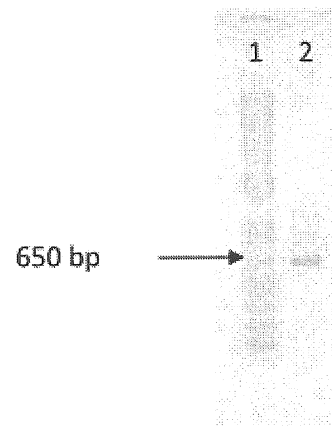
[illegible]

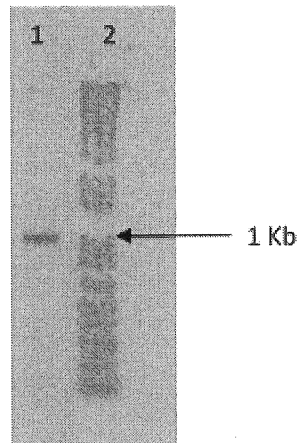
Fig. 11

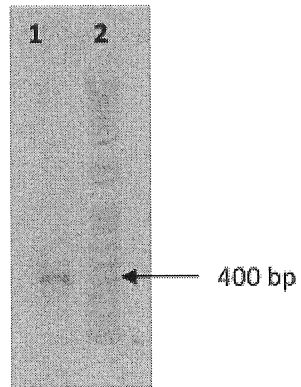
[illegible]

Fig. 12

**Fig.13**

**Fig.14**

**Fig.15**

**Fig.16**

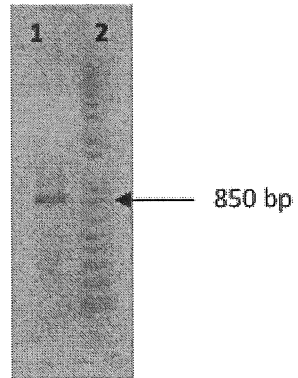


Fig.17

DANH MỤC TRÌNH TỰ

<110> BANGLADESH JUTE RESEARCH INSTITUTE
ALAM, MAQSUDUL

<120> PHÂN TỬ AXIT NUCLEIC MÃ HÓA ENZYM TỪ CON ĐƯỜNG SINH TỔNG HỢP LIGNIN TRONG
CÂY ĐAY VÀ PHƯƠNG PHÁP TẠO RA THỰC VẬT CHUYỂN GEN CHỨA PHÂN TỬ NÀY

<130> JGX-003.25

<140> PCT/US2012/034980

<141> 2012-04-25

<150> 61/480,668

<151> 2011-04-29

<160> 93

<170> PatentIn phiên bản 3.5

<210> 1

<211> 1089

<212> ADN

<213> Corchorus olitorius

<400> 1

```

atgggcagtt taccagagca agagcacccc aaagaggctt ttggatgggc agctagagac      60
acttctggcc acctttctcc cttcaaattc tccagaaggg caacaggaga aaaagatgtg      120
gcattcaaag tgctttattg tgggatatgc cactcagatc ttcatatgat caagaatgaa      180
tggggcactg ccatctaccc tatggtccct gggcatgaga ttgttgagga agtaacagag      240
gtaggaagca aggtagaaaa gttcaaagtt ggagacaaag ttggagctgg agtcttggtt      300
aatcgtgcc gctcttgcca taactgtgct aataatcttg aaaactactg cccacaagcc      360
gttttcactt atgctgcaaa aaactacgat ggaaccatta cctatggagg ctactccgac      420
accatgggtg ccgatgagca cttcataatc cgaattccag acactctgcc tctcgacgcc      480
gccgctcccc tgctctgcgc cggaatcaca gtttatagtc ctttgagata tttccaactc      540
gacaaaccgg gtttccatat tgggtgtggtt ggccttggtg gtttaggcca tatggctgtc      600
aaatttgcca aggctatggg ggccaaggct acagtgatta gcacctctcc caacaagaag      660
aaggaagctt tggaaaatct tgggtgctgat tcgtttttga ttagcgcaga gcaggatcag      720
ctccagactg ccatgggaac aatggatggt atcattgata cagtgtctgc tccacaccct      780
ttactgccat tgatcggttt gttaaagtct catgcaaagc ttattttggt tgggtcttcca      840
gacaaaccac ttgagctaca tgtcttccct atgatcatag ggaggaagac ggtggcagga      900
agtggcgtgg gagggataga ggagactcag gagatgatga attttgcagc caaatatgac      960
ttgaaaccag acattgaagt tatacccggt gactatgtca acactgccat ggaacgcctt     1020

```

gtcaaaggcg atgttaaata cagatttgtc atcgacattg ggaacacact gaaggctaca 1080
tcattcttaa 1089

<210> 2
<211> 362
<212> PRT
<213> Corchorus olitorius

<400> 2

Met Gly Ser Leu Pro Glu Gln Glu His Pro Lys Glu Ala Phe Gly Trp
1 5 10 15

Ala Ala Arg Asp Thr Ser Gly His Leu Ser Pro Phe Lys Phe Ser Arg
20 25 30

Arg Ala Thr Gly Glu Lys Asp Val Ala Phe Lys Val Leu Tyr Cys Gly
35 40 45

Ile Cys His Ser Asp Leu His Met Ile Lys Asn Glu Trp Gly Thr Ala
50 55 60

Ile Tyr Pro Met Val Pro Gly His Glu Ile Val Gly Glu Val Thr Glu
65 70 75 80

Val Gly Ser Lys Val Glu Lys Phe Lys Val Gly Asp Lys Val Gly Ala
85 90 95

Gly Val Leu Val Asn Ser Cys Arg Ser Cys Asp Asn Cys Ala Asn Asn
100 105 110

Leu Glu Asn Tyr Cys Pro Gln Ala Val Phe Thr Tyr Ala Ala Lys Asn
115 120 125

Tyr Asp Gly Thr Ile Thr Tyr Gly Gly Tyr Ser Asp Thr Met Val Ala
130 135 140

Asp Glu His Phe Ile Ile Arg Ile Pro Asp Thr Leu Pro Leu Asp Ala
145 150 155 160

Ala Ala Pro Leu Leu Cys Ala Gly Ile Thr Val Tyr Ser Pro Leu Arg
165 170 175

Tyr Phe Gln Leu Asp Lys Pro Gly Phe His Ile Gly Val Val Gly Leu
180 185 190

Gly Gly Leu Gly His Met Ala Val Lys Phe Ala Lys Ala Met Gly Ala

195	200	205
Lys Val Thr Val Ile Ser Thr Ser Pro Asn Lys Lys Lys Glu Ala Leu		
210	215	220
Glu Asn Leu Gly Ala Asp Ser Phe Leu Ile Ser Ala Glu Gln Asp Gln		
225	230	235 240
Leu Gln Thr Ala Met Gly Thr Met Asp Gly Ile Ile Asp Thr Val Ser		
	245	250 255
Ala Pro His Pro Leu Leu Pro Leu Ile Gly Leu Leu Lys Ser His Ala		
	260	265 270
Lys Leu Ile Leu Val Gly Leu Pro Asp Lys Pro Leu Glu Leu His Val		
	275	280 285
Phe Pro Met Ile Ile Gly Arg Lys Thr Val Ala Gly Ser Gly Val Gly		
	290	295 300
Gly Ile Glu Glu Thr Gln Glu Met Met Asn Phe Ala Ala Lys Tyr Asp		
305	310	315 320
Leu Lys Pro Asp Ile Glu Val Ile Pro Val Asp Tyr Val Asn Thr Ala		
	325	330 335
Met Glu Arg Leu Val Lys Gly Asp Val Lys Tyr Arg Phe Val Ile Asp		
	340	345 350
Ile Gly Asn Thr Leu Lys Ala Thr Ser Ser		
	355	360

<210> 3
 <211> 1086
 <212> ADN
 <213> Corchorus olitorius

<400> 3
 atgagcagat tgccagagga agagcaccct aacaaggctt ttggatgggc agccagagac 60
 acttctggag ttctctctcc cttcaaattt tccagaaggg caacagggga gaaggatgtg 120
 gcattcaagg tgctttattg tgggatatgc cattctgac ttcatatggt caagaatgaa 180
 tggggcgttt ccatctaccc tcttgtccct gggcacgaga ttgttgagaa agtgacagaa 240
 gtgggaagca aggtacaaaa gttcaaagtt ggagacagag ttggtgtcgg ctgcatgggt 300
 gggatcatgtc attcctgtga tagctgcacc aacaatctcg agaactattg cccaaaaaatg 360

atacttactt atggggcaaa gtactatgat ggaactatta catatggagg ttactctgac 420
 actatgggtt cagatgaaca cttcattgtc cgaattcctg aaaatttgcc acttgatgct 480
 gctgctcctc ttctttgtgc tggaatcaca gtttatagcc cattgaaata ttatggactt 540
 gacaagcctg gtttgcattg ggggtgtggtt ggccttggtg ggctaggcca tatggctgtc 600
 aaatttgcca aggctatggg ggccaagggtc acagtaatta gcacttctcc taataagaag 660
 aaggaagctt tggaaaatct tgggtgctgat tcatttttgg tcagcaaaga ccaagatcag 720
 attcaggccg ccatggacac attggatgga atcattgata cagtgtcagc tcaacatcct 780
 atcctgccat tgcttgggat gttaaagact aatggaaagc ttgttctggt tgggtgcaccg 840
 gagaaaccac ttgagttgcc ggcgttcctt ttactcggaa agaggaggct agtagcagga 900
 agcatgattg ggggaatgaa ggagacacaa gagatgattg attttgcagc taaacacaac 960
 attaaaccag acattgaagt tatagctatg gattatgtca aactgccat ggaccgcctt 1020
 ctcaaagctg atgtcaaata cagatttgtc attgacattg gcaacacatt gaaaccaacc 1080
 ccttaa 1086

<210> 4

<211> 361

<212> PRT

<213> *Corchorus olitorius*

<400> 4

Met Ser Arg Leu Pro Glu Glu Glu His Pro Asn Lys Ala Phe Gly Trp
 1 5 10 15

Ala Ala Arg Asp Thr Ser Gly Val Leu Ser Pro Phe Lys Phe Ser Arg
 20 25 30

Arg Ala Thr Gly Glu Lys Asp Val Ala Phe Lys Val Leu Tyr Cys Gly
 35 40 45

Ile Cys His Ser Asp Leu His Met Val Lys Asn Glu Trp Gly Val Ser
 50 55 60

Ile Tyr Pro Leu Val Pro Gly His Glu Ile Val Gly Glu Val Thr Glu
 65 70 75 80

Val Gly Ser Lys Val Gln Lys Phe Lys Val Gly Asp Arg Val Gly Val
 85 90 95

Gly Cys Met Val Gly Ser Cys His Ser Cys Asp Ser Cys Thr Asn Asn
 100 105 110

Leu Glu Asn Tyr Cys Pro Lys Met Ile Leu Thr Tyr Gly Ala Lys Tyr
 115 120 125

Tyr Asp Gly Thr Ile Thr Tyr Gly Gly Tyr Ser Asp Thr Met Val Ala
 130 135 140

Asp Glu His Phe Ile Val Arg Ile Pro Glu Asn Leu Pro Leu Asp Ala
 145 150 155 160

Ala Ala Pro Leu Leu Cys Ala Gly Ile Thr Val Tyr Ser Pro Leu Lys
 165 170 175

Tyr Tyr Gly Leu Asp Lys Pro Gly Leu His Val Gly Val Val Gly Leu
 180 185 190

Gly Gly Leu Gly His Met Ala Val Lys Phe Ala Lys Ala Met Gly Ala
 195 200 205

Lys Val Thr Val Ile Ser Thr Ser Pro Asn Lys Lys Lys Glu Ala Leu
 210 215 220

Glu Asn Leu Gly Ala Asp Ser Phe Leu Val Ser Lys Asp Gln Asp Gln
 225 230 235 240

Ile Gln Ala Ala Met Asp Thr Leu Asp Gly Ile Ile Asp Thr Val Ser
 245 250 255

Ala Gln His Pro Ile Leu Pro Leu Leu Gly Met Leu Lys Thr Asn Gly
 260 265 270

Lys Leu Val Leu Val Gly Ala Pro Glu Lys Pro Leu Glu Leu Pro Ala
 275 280 285

Phe Pro Leu Leu Gly Lys Arg Arg Leu Val Ala Gly Ser Met Ile Gly
 290 295 300

Gly Met Lys Glu Thr Gln Glu Met Ile Asp Phe Ala Ala Lys His Asn
 305 310 315 320

Ile Lys Pro Asp Ile Glu Val Ile Ala Met Asp Tyr Val Asn Thr Ala
 325 330 335

Met Asp Arg Leu Leu Lys Ala Asp Val Lys Tyr Arg Phe Val Ile Asp
 340 345 350

Ile Gly Asn Thr Leu Lys Pro Thr Pro
355 360

<210> 5
<211> 921
<212> ADN
<213> *Corchorus olitorius*

<400> 5
atgatcaaga acgactgggg ctattccatc tatectctcg tccctggaca tgaaatagtt 60
ggggaagtaa cagaggtggg ttctaaggta agcaagttca aagttggaga caaggttgga 120
gttgataca tggttggatc atgccgttct tgcgatgatt gctccgataa tcttgaaaac 180
tactgtccga aaatgatacc tacttgtggg gcaaagtacc atgatggaac tattacatat 240
ggaggtttct ctgacactat ggttgccgat gaacacttcg ttgtccgaat tccggacaac 300
atgcccttg atgcggcggc tctcttctt tgtgccgga ccaagttta cagtccaatg 360
aaatattatg gactcgacaa gcctggtttg catttggttg ttgttggaatt gggagggctt 420
ggccatgttg ctgtgaaatt tgcaaaagct atgggggcca aggtgacagt gatcagcacc 480
tctcctagta agaagcagga agctttggaa attcttggtg ctgattcgtt tttggtttagc 540
cgggacgaag atcagcttaa ggctgccaa ggacacaatga atggtatagt ggatacggta 600
tctgccaaac atgatctgca gccattactt ggactgttga agaatcatgg aaagcttggt 660
ctaattggtg ttccggtaaa gccatatgag ctaccagcag cttctttgat cttggggagg 720
aagctagtag gaggaagtaa cgtcggagga ttggaagaga ctcaagagat gattgatttt 780
gcagcaaaac acaatgtgac agcaaacggt gaagtgattc cgatggatta tgtgaacact 840
gcctttgagc gacttgcaaa agccgatgtt agatatcgat ttgtcgttga cattggcaac 900
accttgaaga cttcgtctta a 921

<210> 6
<211> 306
<212> PRT
<213> *Corchorus olitorius*

<400> 6
Met Ile Lys Asn Asp Trp Gly Tyr Ser Ile Tyr Pro Leu Val Pro Gly
1 5 10 15

His Glu Ile Val Gly Glu Val Thr Glu Val Gly Ser Lys Val Ser Lys
20 25 30

Phe Lys Val Gly Asp Lys Val Gly Val Gly Tyr Met Val Gly Ser Cys

35	40	45
Arg Ser Cys Asp Asp Cys Ser Asp Asn Leu Glu Asn Tyr Cys Pro Lys		
50	55	60
Met Ile Pro Thr Cys Gly Ala Lys Tyr His Asp Gly Thr Ile Thr Tyr		
65	70	75
Gly Gly Phe Ser Asp Thr Met Val Ala Asp Glu His Phe Val Val Arg		
	85	90
Ile Pro Asp Asn Met Pro Leu Asp Ala Ala Ala Pro Leu Leu Cys Ala		
	100	105
Gly Thr Thr Val Tyr Ser Pro Met Lys Tyr Tyr Gly Leu Asp Lys Pro		
	115	120
Gly Leu His Leu Gly Val Val Gly Leu Gly Gly Leu Gly His Val Ala		
	130	135
Val Lys Phe Ala Lys Ala Met Gly Ala Lys Val Thr Val Ile Ser Thr		
	145	150
Ser Pro Ser Lys Lys Gln Glu Ala Leu Glu Ile Leu Gly Ala Asp Ser		
	165	170
Phe Leu Val Ser Arg Asp Glu Asp Gln Leu Lys Ala Ala Lys Gly Thr		
	180	185
Met Asn Gly Ile Val Asp Thr Val Ser Ala Lys His Asp Leu Gln Pro		
	195	200
Leu Leu Gly Leu Leu Lys Asn His Gly Lys Leu Val Leu Ile Gly Val		
	210	215
Pro Val Lys Pro Tyr Glu Leu Pro Ala Ala Ser Leu Ile Leu Gly Arg		
	225	230
Lys Leu Val Gly Gly Ser Asn Val Gly Gly Leu Glu Glu Thr Gln Glu		
	245	250
Met Ile Asp Phe Ala Ala Lys His Asn Val Thr Ala Asn Val Glu Val		
	260	265
Ile Pro Met Asp Tyr Val Asn Thr Ala Phe Glu Arg Leu Ala Lys Ala		

275

280

285

Asp Val Arg Tyr Arg Phe Val Val Asp Ile Gly Asn Thr Leu Lys Thr
 290 295 300

Ser Ser
 305

<210> 7
 <211> 1065
 <212> ADN
 <213> *Corchorus olitorius*

<400> 7
 atggattctc aaaccaaadc tgataactgc cttgggtggg ccgccagaga tccatctgga 60
 gttctgtcac cctacacatt tagtcgccgg cctcttgga gtaatgatgt ttccataaag 120
 attaccact gtggagtttg ttatgctgat ttcatttgg ttaggaataa gcatggagat 180
 agtatgtacc ctgtgggtccc cgggcatgaa attgctgggt ttgtcaagga ggttggctcg 240
 aatgttcacc gcttcaaggt tggatgacct gttggagtgg gaacttatgt taactcatgc 300
 agagattgtg agtactgcaa tgatggcctt gaagttcatt gtgaaaaaat agttcttacc 360
 ttttaattgta ttgatgagga tggaacagtc actaaaggcg gttattctaa ccatatcatt 420
 gtccacgaaa ggtactgctt aagaatacct agcaattatc ctttggcttc agcagcacct 480
 ttgctttgtg ctggcatcac tgtttatgca ccaatgatgc gtcataacat gaatcagcct 540
 ggcaaatcat taggagtgat tgggctcggt ggccttgggt acatggcagt gaagtttggg 600
 aaggcttttg ggttgcatgt aacagtttta agcacaagca tatctaagaa agatgaggcg 660
 ttgagtcttc ttggtgcaga caactttggt gtctcctctg accaagagca gatgaagggc 720
 ctatcgaagt cattggactt tatagttgac actgcatctg gtgatcatcc ctttgatccc 780
 tacatgtcac tcctgaagat tgctggtggt tatgtccttg ttgggttccc aagtgaagtc 840
 aaattcagcc ctgcaagtct caatctgggt atgagaacaa tctcaggaag cgtaacgggg 900
 ggagtaaaag tgatccagga gatgatagac ttttgtgctg ctcataaagt ttacccgcag 960
 atagaagtaa tccaattca atatgcaaat gaagctcttg agaggctaga aaagagggat 1020
 gtgaagtaca ggtttgtgat tgacatcgag aacagcctga aatga 1065

<210> 8
 <211> 354
 <212> PRT
 <213> *Corchorus olitorius*

<400> 8

Met Asp Ser Gln Thr Lys Ser Asp Asn Cys Leu Gly Trp Ala Ala Arg
1 5 10 15

Asp Pro Ser Gly Val Leu Ser Pro Tyr Thr Phe Ser Arg Arg Pro Leu
20 25 30

Gly Ser Asn Asp Val Ser Ile Lys Ile Thr His Cys Gly Val Cys Tyr
35 40 45

Ala Asp Phe Ile Trp Ser Arg Asn Lys His Gly Asp Ser Met Tyr Pro
50 55 60

Val Val Pro Gly His Glu Ile Ala Gly Val Val Lys Glu Val Gly Ser
65 70 75 80

Asn Val His Arg Phe Lys Val Gly Asp Pro Val Gly Val Gly Thr Tyr
85 90 95

Val Asn Ser Cys Arg Asp Cys Glu Tyr Cys Asn Asp Gly Leu Glu Val
100 105 110

His Cys Glu Lys Ile Val Leu Thr Phe Asn Cys Ile Asp Glu Asp Gly
115 120 125

Thr Val Thr Lys Gly Gly Tyr Ser Asn His Ile Ile Val His Glu Arg
130 135 140

Tyr Cys Leu Arg Ile Pro Ser Asn Tyr Pro Leu Ala Ser Ala Ala Pro
145 150 155 160

Leu Leu Cys Ala Gly Ile Thr Val Tyr Ala Pro Met Met Arg His Asn
165 170 175

Met Asn Gln Pro Gly Lys Ser Leu Gly Val Ile Gly Leu Gly Gly Leu
180 185 190

Gly His Met Ala Val Lys Phe Gly Lys Ala Phe Gly Leu His Val Thr
195 200 205

Val Leu Ser Thr Ser Ile Ser Lys Lys Asp Glu Ala Leu Ser Leu Leu
210 215 220

Gly Ala Asp Asn Phe Val Val Ser Ser Asp Gln Glu Gln Met Lys Gly
225 230 235 240

Leu Ser Lys Ser Leu Asp Phe Ile Val Asp Thr Ala Ser Gly Asp His
 245 250 255

Pro Phe Asp Pro Tyr Met Ser Leu Leu Lys Ile Ala Gly Val Tyr Val
 260 265 270

Leu Val Gly Phe Pro Ser Glu Val Lys Phe Ser Pro Ala Ser Leu Asn
 275 280 285

Leu Gly Met Arg Thr Ile Ser Gly Ser Val Thr Gly Gly Val Lys Val
 290 295 300

Ile Gln Glu Met Ile Asp Phe Cys Ala Ala His Lys Val Tyr Pro Gln
 305 310 315 320

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

Glu Lys Arg Asp Val Lys Tyr Arg Phe Val Ile Asp Ile Glu Asn Ser
 340 345 350

Leu Lys

<210> 9

<211> 1092

<212> ADN

<213> Corchorus olitorius

<400> 9

atggctaagt catcacctga agaagagcac cctgttaagg cctttggctg ggctgccagg	60
gacacttctg gtcactcttc tcctttcaac ttctccagaa gggcaactgg tgaaggagat	120
gtgaggttta aggtgttata ttgtggaata tgccattctg accttcattt tatcaagaat	180
gaatggaact tctccatcta tcctttgggt cccgggcatg aaattgtggg ggaagtgaca	240
gaagttggca aaaaagttca aaaggttaag attggagaca aagtgggtgt gggatgcatt	300
attggtgctt gtcacacttg tgagagttgt gccaatgacc ttgaaaatta ctgtcctaaa	360
gcaattgcaa cctacaatgg gacttactat gatggaacca tgacatacgg aggctactcc	420
gattcaatgg ttgccgatga acgatacgtc gttcagattc ctgatggcat ggcccttgac	480
tctgctgccc cattgctctg tgctggaatt actgtttaca gtccattgaa gtattttggg	540
ttaggtgaag ctggtaatca cattggcatt gttggccttg gtggccttgg ccatgtagct	600
gttaagtttg ccaaggcttt ggggtccaaa gttacagtaa ttagcacttc ccctggtaag	660

aagaaggaag ccttgaact tcttggggct gattcggttct tggttagccg tgaccaagac 720
gagatgcagg ctgccatggg cacattggat gggatcatag acacagtctc tgccgttcat 780
ccaattatgc cattgcttgg tcttctgaaa tcccatggaa aactcattat ggtgggcgca 840
ccaatcgaac cacttgagtt acctgtcttt tctttgatca tgggaaggaa gacgatggct 900
gggagtggaa ttggaggaat gaaagagaca caagagatga ttgattttgc agcaaaacac 960
aacataaaag cagacatcga agtgattccg atggattatg tgaacaaagc tatggaaaga 1020
cttgaaaagg gtgatgttag atacagattc gtaattgaca ttggaaacac cttggctact 1080
accaagcctt ag 1092

<210> 10

<211> 363

<212> PRT

<213> *Corchorus olitorius*

<400> 10

Met Ala Lys Ser Ser Pro Glu Glu Glu His Pro Val Lys Ala Phe Gly
1 5 10 15

Trp Ala Ala Arg Asp Thr Ser Gly His Leu Ser Pro Phe Asn Phe Ser
20 25 30

Arg Arg Ala Thr Gly Glu Gly Asp Val Arg Phe Lys Val Leu Tyr Cys
35 40 45

Gly Ile Cys His Ser Asp Leu His Phe Ile Lys Asn Glu Trp Asn Phe
50 55 60

Ser Ile Tyr Pro Leu Val Pro Gly His Glu Ile Val Gly Glu Val Thr
65 70 75 80

Glu Val Gly Lys Lys Val Gln Lys Val Lys Ile Gly Asp Lys Val Gly
85 90 95

Val Gly Cys Ile Ile Gly Ala Cys His Thr Cys Glu Ser Cys Ala Asn
100 105 110

Asp Leu Glu Asn Tyr Cys Pro Lys Ala Ile Ala Thr Tyr Asn Gly Thr
115 120 125

Tyr Tyr Asp Gly Thr Met Thr Tyr Gly Gly Tyr Ser Asp Ser Met Val
130 135 140

Ala Asp Glu Arg Tyr Val Val Gln Ile Pro Asp Gly Met Ala Leu Asp

145						150						155						160
Ser	Ala	Ala	Pro	Leu	Leu	Cys	Ala	Gly	Ile	Thr	Val	Tyr	Ser	Pro	Leu			
				165					170					175				
Lys	Tyr	Phe	Gly	Leu	Gly	Glu	Ala	Gly	Asn	His	Ile	Gly	Ile	Val	Gly			
				180					185					190				
Leu	Gly	Gly	Leu	Gly	His	Val	Ala	Val	Lys	Phe	Ala	Lys	Ala	Leu	Gly			
				195					200					205				
Ser	Lys	Val	Thr	Val	Ile	Ser	Thr	Ser	Pro	Gly	Lys	Lys	Lys	Glu	Ala			
				210					215					220				
Leu	Glu	Leu	Leu	Gly	Ala	Asp	Ser	Phe	Leu	Val	Ser	Arg	Asp	Gln	Asp			
				225					230					235				
Glu	Met	Gln	Ala	Ala	Met	Gly	Thr	Leu	Asp	Gly	Ile	Ile	Asp	Thr	Val			
				245					250					255				
Ser	Ala	Val	His	Pro	Ile	Met	Pro	Leu	Leu	Gly	Leu	Leu	Lys	Ser	His			
				260					265					270				
Gly	Lys	Leu	Ile	Met	Val	Gly	Ala	Pro	Ile	Glu	Pro	Leu	Glu	Leu	Pro			
				275					280					285				
Val	Phe	Ser	Leu	Ile	Met	Gly	Arg	Lys	Thr	Met	Ala	Gly	Ser	Gly	Ile			
				290					295					300				
Gly	Gly	Met	Lys	Glu	Thr	Gln	Glu	Met	Ile	Asp	Phe	Ala	Ala	Lys	His			
				305					310					315				
Asn	Ile	Lys	Ala	Asp	Ile	Glu	Val	Ile	Pro	Met	Asp	Tyr	Val	Asn	Lys			
				325					330					335				
Ala	Met	Glu	Arg	Leu	Glu	Lys	Gly	Asp	Val	Arg	Tyr	Arg	Phe	Val	Ile			
				340					345					350				
Asp	Ile	Gly	Asn	Thr	Leu	Ala	Thr	Thr	Lys	Pro								
				355					360									

```
<210> 11
<211> 1080
<212> ADN
<213> Corchorus olitorius
```

<400> 11

```

atggctatgg aaacacctaa ccacactcag acggtagcag ggtgggctgc tcataattcc      60
tcaggcaaga tcgtccctta caccttcaaa agaagggaaa atggcgtgaa cgatgtgacc      120
attaaagtga tgtattgtgg gatctgccat actgatctcc accatgttaa gaacgattgg      180
ggtatcacca tgtatcctgt agttcccggg catgaaatta ctgggggtgat caccaagatt      240
ggaaacaatg tgaagaattht caaagtggga gacagggtag gtgtgggttg cttggcagca      300
tcctgttttg aatgcgagtt ctgtaaaagc tcgcaagaga actactgtga ccaaatccag      360
ttcacttaca atggcatctt ttgggatggt agcggtactt atggcggcta ttcccaaagt      420
ttagtcgccg atcaccggtg cgttggtcgt gtgccggata acctgccgat ggacgccgca      480
gcgccactgt tgtgtgccgg gatcaccggt ttcagcccca tgaaagatag ccaactgctc      540
gagtcaccgg gcaaaaaagt gggcatagtt ggtttaggcg gtctcggtea tgcgctgtc      600
aaaatggcaa aggcatttgg tcatcatgtg accgtgataa gcacttctcc atcaaaagaa      660
aaggaagcta aacagcgtht gggcgcagat gatttcatag ttagcaccaa caccgaacaa      720
atgcagagag gaaagcgaac gctggatggt attttggaca cagtttcagc taaacactcg      780
ctcggaacaa tcttggaaact gcttaaagtg aatgggtactt tagtggttgt gggagcacca      840
gacaagccaa tcgaccttcc ttcatttcca ttaatatttg ggaaaagagc agtgaagggg      900
agcatgacag gggggatgaa agagacacaa gaaatgatgg atgtgtgtgg caaacacaa      960
attacatgtg atatagaagt aattaaacct gatcagataa atgaagccct tgatcgactt     1020
tccaaaaatg atgttcgata cagatttgta atcgacattg ctggaagggt caagctttaa     1080

```

<210> 12

<211> 359

<212> PRT

<213> *Corchorus olitorius*

<400> 12

```

Met Ala Met Glu Thr Pro Asn His Thr Gln Thr Val Ala Gly Trp Ala
1           5           10          15

```

```

Ala His Asn Ser Ser Gly Lys Ile Val Pro Tyr Thr Phe Lys Arg Arg
          20           25           30

```

```

Glu Asn Gly Val Asn Asp Val Thr Ile Lys Val Met Tyr Cys Gly Ile
          35           40           45

```

```

Cys His Thr Asp Leu His His Val Lys Asn Asp Trp Gly Ile Thr Met
          50           55           60

```

Tyr Pro Val Val Pro Gly His Glu Ile Thr Gly Val Ile Thr Lys Ile
65 70 75 80

Gly Asn Asn Val Lys Asn Phe Lys Val Gly Asp Arg Val Gly Val Gly
85 90 95

Cys Leu Ala Ala Ser Cys Leu Glu Cys Glu Phe Cys Lys Ser Ser Gln
100 105 110

Glu Asn Tyr Cys Asp Gln Ile Gln Phe Thr Tyr Asn Gly Ile Phe Trp
115 120 125

Asp Gly Ser Val Thr Tyr Gly Gly Tyr Ser Gln Met Leu Val Ala Asp
130 135 140

His Arg Tyr Val Val Arg Val Pro Asp Asn Leu Pro Met Asp Ala Ala
145 150 155 160

Ala Pro Leu Leu Cys Ala Gly Ile Thr Val Phe Ser Pro Met Lys Asp
165 170 175

Ser Gln Leu Leu Glu Ser Pro Gly Lys Lys Val Gly Ile Val Gly Leu
180 185 190

Gly Gly Leu Gly His Val Ala Val Lys Met Ala Lys Ala Phe Gly His
195 200 205

His Val Thr Val Ile Ser Thr Ser Pro Ser Lys Glu Lys Glu Ala Lys
210 215 220

Gln Arg Leu Gly Ala Asp Asp Phe Ile Val Ser Thr Asn Thr Glu Gln
225 230 235 240

Met Gln Arg Gly Lys Arg Thr Leu Asp Val Ile Leu Asp Thr Val Ser
245 250 255

Ala Lys His Ser Leu Gly Pro Ile Leu Glu Leu Leu Lys Val Asn Gly
260 265 270

Thr Leu Val Val Val Gly Ala Pro Asp Lys Pro Ile Asp Leu Pro Ser
275 280 285

Phe Pro Leu Ile Phe Gly Lys Arg Ala Val Lys Gly Ser Met Thr Gly
290 295 300

Gly Met Lys Glu Thr Gln Glu Met Met Asp Val Cys Gly Lys His Asn
 305 310 315 320

Ile Thr Cys Asp Ile Glu Val Ile Lys Pro Asp Gln Ile Asn Glu Ala
 325 330 335

Leu Asp Arg Leu Ser Lys Asn Asp Val Arg Tyr Arg Phe Val Ile Asp
 340 345 350

Ile Ala Gly Arg Ser Lys Leu
 355

<210> 13

<211> 1071

<212> ADN

<213> *Corchorus olitorius*

<400> 13

atgggcagtc ttgaaactga gagaacaacc acaggctggg ctgccagaga cccttctgga	60
gtcttgtctc cttacactta cactctcaga aacacaggtc cagaagatgt ttttatcaag	120
gttatttgct gtggaatctg ccacactgat cttcatcaag ccaaaaatga tcttggcatg	180
tcaaactacc caatggttcc tgggcatgaa gtggttgggg aggtggtgga agtaggatca	240
caggtgacca aattcagagt aggagagata gttggcgttg gttgtattgt tgggtgttgc	300
agaaactgcc gcccatgcaa cactgacaat gaacaatact gcaacaagaa gatttggctt	360
tacaatgatg tctacactga tggcaaacc actcaagggtg gctttgctgc ctccatggtc	420
gctgatcaaa agtttgtggt gaaaatccct gatggaatgg cagcagaaca ggtggctcca	480
ctcttgtgtg ctggtgtgac agtttacagc ccactgaaac actttggctt aatggagagt	540
gggttaagag gagggatttt ggggcttggg ggagtaggtc acatgggagt gaagatagcc	600
aaagcaatgg gacaccatgt gactgttatc agctcttcag acaagaaaa agttgaggcc	660
ttggagcatc ttggtgctga tgattatgta gtcagctctg atgctgaaag catgcaaaag	720
attgctgatt cactcgacta tatcatcgat accgtgcctg tttttcatcc ccttgagcct	780
tacctttcag tgttgaaact tgatggaaag ttgatcttga ctggtgttat caatactcct	840
cttcagtttg ttacccccat ggtcatgctt gggagaaagg taattacagg gagtttcgtt	900
gggagcatga aggaaacaga ggagatgctt gatttctgta aagagaagga tttaaactca	960
atgattgaag ttgtgaagat ggattatatc aacacagcca tggagaggct cgagaagaat	1020
gatgttcgct acaggttcgt cgtggatggt gccggaagca aacttgagta a	1071

<210> 14

<211> 356

<212> PRT

<213> *Corchorus olitorius*

<400> 14

Met Gly Ser Leu Glu Thr Glu Arg Thr Thr Thr Gly Trp Ala Ala Arg
 1 5 10 15

Asp Pro Ser Gly Val Leu Ser Pro Tyr Thr Tyr Thr Leu Arg Asn Thr
 20 25 30

Gly Pro Glu Asp Val Phe Ile Lys Val Ile Cys Cys Gly Ile Cys His
 35 40 45

Thr Asp Leu His Gln Ala Lys Asn Asp Leu Gly Met Ser Asn Tyr Pro
 50 55 60

Met Val Pro Gly His Glu Val Val Gly Glu Val Leu Glu Val Gly Ser
 65 70 75 80

Gln Val Thr Lys Phe Arg Val Gly Glu Ile Val Gly Val Gly Cys Ile
 85 90 95

Val Gly Cys Cys Arg Asn Cys Arg Pro Cys Asn Thr Asp Asn Glu Gln
 100 105 110

Tyr Cys Asn Lys Lys Ile Trp Ser Tyr Asn Asp Val Tyr Thr Asp Gly
 115 120 125

Lys Pro Thr Gln Gly Gly Phe Ala Ala Ser Met Val Ala Asp Gln Lys
 130 135 140

Phe Val Val Lys Ile Pro Asp Gly Met Ala Ala Glu Gln Val Ala Pro
 145 150 155 160

Leu Leu Cys Ala Gly Val Thr Val Tyr Ser Pro Leu Lys His Phe Gly
 165 170 175

Leu Met Glu Ser Gly Leu Arg Gly Gly Ile Leu Gly Leu Gly Gly Val
 180 185 190

Gly His Met Gly Val Lys Ile Ala Lys Ala Met Gly His His Val Thr
 195 200 205

Val Ile Ser Ser Ser Asp Lys Lys Lys Val Glu Ala Leu Glu His Leu
 210 215 220

Gly Ala Asp Asp Tyr Val Val Ser Ser Asp Ala Glu Ser Met Gln Lys
 225 230 235 240

Ile Ala Asp Ser Leu Asp Tyr Ile Ile Asp Thr Val Pro Val Phe His
 245 250 255

Pro Leu Glu Pro Tyr Leu Ser Val Leu Lys Leu Asp Gly Lys Leu Ile
 260 265 270

Leu Thr Gly Val Ile Asn Thr Pro Leu Gln Phe Val Thr Pro Met Val
 275 280 285

Met Leu Gly Arg Lys Val Ile Thr Gly Ser Phe Val Gly Ser Met Lys
 290 295 300

Glu Thr Glu Glu Met Leu Asp Phe Cys Lys Glu Lys Asp Leu Asn Ser
 305 310 315 320

Met Ile Glu Val Val Lys Met Asp Tyr Ile Asn Thr Ala Met Glu Arg
 325 330 335

Leu Glu Lys Asn Asp Val Arg Tyr Arg Phe Val Val Asp Val Ala Gly
 340 345 350

Ser Lys Leu Glu
 355

<210> 15

<211> 2492

<212> ADN

<213> *Corchorus olitorius*

<400> 15

attatcttgg gcatataccc catatttgaa ggcattgggat gtccctccat gaccatctgc	60
ctctaatttc agattcccat caatgtaaac cttcaacttt gctgcatcaa catcatggat	120
aacattcagc ttgaaccact tgtcatagat gtcccgaacc aggaccggac tttttagta	180
cgtgagcgaa ccattgtaga cccggagcat taggggtggtg gctcgaggag ggcttgctcc	240
aaaaacttgc atgatgcaaa atcctgatgt tccctggggc acataccagt atccttcaaa	300
ctgccacacc cctgacgagt aattgtatcc atgtatggca atctcggttc ggggctttgt	360
tttgctagta ggagtatgcg gcttgtctgt ggagtaaacc caacatctat gaacgcctc	420
aatgaagcta taacgctgcg cctcaggcac gttgtagggc ctctggatat ggtagtatgt	480
tcggttcagg ggaagagaga tgaacccttt cgtaagatca acctgttcgt ccgaagcaaa	540

ggcttgatac	aaaagtgtac	tgactagaat	aattaatatc	aagtagaagc	aagtcatttt	600
ggttaaggct	ggtgacaatt	gcatgtatga	aaaaagtga	tgtaaatgat	tgatggggtt	660
aattagtatg	ggacattgat	ctttgcatat	gagtaattaa	tatatagaaa	tgagaaggaa	720
tttgtaaatg	attcaatgct	tattgtatct	gtagagtggc	agatgatcat	aatagcttaa	780
ttttgatggg	ttgattccta	gtggaaccat	atatatagct	ttattaaata	gctaggccag	840
aataattttg	caaaatatca	gcaataaaga	gacacaataa	gataaagatt	ctatcgatta	900
acagctcaag	aatggacatc	agcaagcaaa	cagcagcaaa	acatttcata	tgaagtcatg	960
agtcacgggc	ttaaacaatg	catagcaagg	gggaaacaat	gtaaaagaca	aaagcagtag	1020
ttagagggaac	taaagaagca	aaatccaata	ccatgaatca	gattcccagc	agataaaaaca	1080
actaaatggg	catagaagcg	agggatcgac	gatttatcag	ttctttaatt	atgctagcgc	1140
gtgatcgatg	aactttattc	taacatctag	actaccagc	tggtattcta	atatctggac	1200
ttcctagtgt	aaagtgaacc	tagctactct	gacaaggaat	ttgccacatc	gatcacaaaa	1260
cggatatctga	catcagatgt	tgcaagcctg	tccattgctg	tgtaaatctc	atccatccga	1320
atcagctcaa	tgtctgcact	aatattgtgc	ttggcacaaa	agtctagcat	ctcttgtgtt	1380
tctttcatcc	ctccaatatc	acttcctoca	acaagcttcc	ggcctaaaac	taaaggaaaag	1440
atgggcaatt	caaggggctt	attgggcaat	cccaaagtga	ccagttttcc	attcaccttc	1500
agcagactaa	gcagtggaag	caggggatga	actgcagaca	ccgtgtcaat	gatatagtcc	1560
atggtgccaa	cggccgactt	catttttgca	ggatcattgg	aaagaagaaa	tgagtgcagca	1620
ccaagtctat	taattgcttc	ttcttccttc	tttggggagc	tactaatgac	agtaactttc	1680
aacccaaagg	ccttaccgat	tttcacagca	acatgaccaa	gccaccaag	tcctgcaact	1740
cccaaagtgc	tgctgcctc	agtcattcca	tagtatttca	tagggctgta	cactgtgatc	1800
ccggcacata	atagtgtg	acccgcatcc	agtggcatgt	tatcaggaaa	ccgaaccaca	1860
aagcgctggg	caacaacaat	catatctgaa	taaccacat	aatttcttgt	cccatccaca	1920
taattagagt	tataggtaaa	tatcatctga	gggcaatagt	tctccaggtc	ctgctggcag	1980
cactcacatt	tcttgagga	gccaaccatg	acccaaccc	caaccgggtc	tcctctttg	2040
aattttgtca	ccttattccc	cacttttgtt	gcaacaccaa	caatttcatg	cccaggaaca	2100
acagggtaac	gggtgaaacc	ccattcattt	ctaagggtat	gcaggtcgga	atggcacact	2160
ccacagtaga	gtattttcaa	ggtgacatct	tcctcaccat	tttcccttct	ggagaactgg	2220
aaaggagaga	gagttccgga	ggaatcccga	gcagcataac	caaaagcctt	ttgtgggtgc	2280
tcttcttctg	gtgattttga	catttctttt	togactctgt	tttttgtttt	tactttactc	2340

tttttttgtt ctcttcactg ttgaagaagt ggttcttatt gttttaaggt tgcaaaatct 2400
 aacagatggg tataggacaa tttataggag gagatggtgg ttccttagac actgtttacg 2460
 tacacccctt gtataccttt attagttata ga 2492

<210> 16
 <211> 744
 <212> ADN
 <213> *Corchorus olitorius*

<400> 16
 atggctcaa ctcaggcaga acagcaaact caagctagta ggcaccagga agttggccac 60
 aagagtctct tgcagagtga taaactctac caatatatac ttgagaccag tgtttatcct 120
 agggagcctg aagccatgaa agagctcagg gagcttactg caaaacatcc atggaatctc 180
 atgactactt cagctgatga agggcagttt ttgaacatgt tgcttaagct tattaatgcc 240
 aagaacacca tggaaattgg tgtttacact ggctattcac tccttgccac tgccttgct 300
 ctgcctgaag atggcaagat tttggccatg gacatcaacc gtgaaaacta tgaattgggt 360
 ttaccagtaa tccaaaaagc cggtgttgct cacaagattg acttcaaaga aggccctgct 420
 cttcctgttc ttgaccaaatt gattgaagct gggacatacc atggaacatt cgatttcac 480
 tttgttgatg ctgacaagga caactacatt aactaccaca agaggctgat tgagctagtt 540
 aaggttgggg gagtcatcgg ctacgacaac accctatgga acggttccgt ggtgggcct 600
 cccgatgctc cattaaggaa gtatgtcttg tattacagag actttgtctt ggagcttaac 660
 aaagccctag ctgctgatcc aaggattgaa atctgccaac ttctgtttgg cgatggaatt 720
 accctctgcc gtcggattaa gtga 744

<210> 17
 <211> 247
 <212> PRT
 <213> *Corchorus olitorius*

<400> 17
 Met Ala Pro Thr Gln Ala Glu Gln Gln Thr Gln Ala Ser Arg His Gln
 1 5 10 15
 Glu Val Gly His Lys Ser Leu Leu Gln Ser Asp Lys Leu Tyr Gln Tyr
 20 25 30
 Ile Leu Glu Thr Ser Val Tyr Pro Arg Glu Pro Glu Ala Met Lys Glu
 35 40 45
 Leu Arg Glu Leu Thr Ala Lys His Pro Trp Asn Leu Met Thr Thr Ser

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

<210> 18
 <211> 744
 <212> ADN
 <213> Corchorus olitorius

<400> 18
 atggcaacca atacccaaga gcagcaaact caggctggca gacaccagga ggttgccac 60
 aagagtcttt tgcaaagtga tgctctttac cagtatatcc ttgagacaag tgtgtatcca 120

```

agagagcctg aacccatgaa ggagctcagg gaattgactg ccaagcatcc atggaatctg      180
atgacaacat cagcagatga agggcaattc ctgaacatgc ttctgaagct gatcaatgcc      240
aagaacacca tggaaattgg tgtttacact ggctactctc tcttagccac agcccttgct      300
cttcctgacg atggcaagat cttggccatg gacattaaca gggaaaacta cgagttgggt      360
ctgcctgtaa tccaaaaagc aggcgttgca cacaaaattg aattcaaaga gggccctgct      420
ttgcctgttc ttgacaaact agtcgaagat gaaaagaatc atggatcata tgacttcacg      480
ttcgtggatg ctgacaaaga caactacata aactaccaca agaggttaat agaccttgct      540
aaagttggag gcttaatcgg ctacgacaac actctatgga atggctctgt tgcgcacca      600
cctgatgctc ctctcaggaa atatgtcagg tattacagag actttgtttt ggaactcaac      660
aaggctctcg ccgccgaccc caggattgag atctgtatgc ttcccgctcg cgatggaatc      720
accctttgcc gtcgaatcaa atga                                             744

```

<210> 19

<211> 247

<212> PRT

<213> *Corchorus olitorius*

<400> 19

```

Met Ala Thr Asn Thr Gln Glu Gln Gln Thr Gln Ala Gly Arg His Gln
1              5              10              15

```

```

Glu Val Gly His Lys Ser Leu Leu Gln Ser Asp Ala Leu Tyr Gln Tyr
              20              25              30

```

```

Ile Leu Glu Thr Ser Val Tyr Pro Arg Glu Pro Glu Pro Met Lys Glu
              35              40              45

```

```

Leu Arg Glu Leu Thr Ala Lys His Pro Trp Asn Leu Met Thr Thr Ser
              50              55              60

```

```

Ala Asp Glu Gly Gln Phe Leu Asn Met Leu Leu Lys Leu Ile Asn Ala
65              70              75              80

```

```

Lys Asn Thr Met Glu Ile Gly Val Tyr Thr Gly Tyr Ser Leu Leu Ala
              85              90              95

```

```

Thr Ala Leu Ala Leu Pro Asp Asp Gly Lys Ile Leu Ala Met Asp Ile
              100              105              110

```

```

Asn Arg Glu Asn Tyr Glu Leu Gly Leu Pro Val Ile Gln Lys Ala Gly
115              120              125

```

Val Ala His Lys Ile Glu Phe Lys Glu Gly Pro Ala Leu Pro Val Leu
130 135 140

Asp Lys Leu Val Glu Asp Glu Lys Asn His Gly Ser Tyr Asp Phe Ile
145 150 155 160

Phe Val Asp Ala Asp Lys Asp Asn Tyr Ile Asn Tyr His Lys Arg Leu
165 170 175

Ile Asp Leu Val Lys Val Gly Gly Leu Ile Gly Tyr Asp Asn Thr Leu
180 185 190

Trp Asn Gly Ser Val Val Ala Pro Pro Asp Ala Pro Leu Arg Lys Tyr
195 200 205

Val Arg Tyr Tyr Arg Asp Phe Val Leu Glu Leu Asn Lys Ala Leu Ala
210 215 220

Ala Asp Pro Arg Ile Glu Ile Cys Met Leu Pro Val Gly Asp Gly Ile
225 230 235 240

Thr Leu Cys Arg Arg Ile Lys
245

<210> 20

<211> 1098

<212> ADN

<213> Corchorus olitorius

<400> 20

atgggttcaa ctggtgaaac ccaattcact ccaactcaag tctccgatga ggaagcaaac	60
ttgttcgcca tgcaattggc tagtgccctca gttcttccca tggctcctcaa atctgccata	120
gaacttgacc tacttgaagt catggccaag gctggacctg gtgctttctt gtccccaaca	180
gaagtagctt cccaattgcc caccaagaac cctgatgcac ccgtcatgct cgaccgtatc	240
ttgcggctcc ttgctagtta ctccatttta acttgctcct taaggaatct tcctgatggc	300
aaagttgaga ggctctatgg ccttgccct gtctgtaaat acctggtcaa gaatgaagat	360
ggtgtcgctc tttccgccct taatctcatg aatcaagaca aggtcctaata ggagagctgg	420
tactacttga aagatgcagt gttggaagg ggaattccat tcaacaaggc ctatggcatg	480
accgcgttcg agtaccatgg cactgaccct agattcaaca aggttttcaa caggggaatg	540
tctgatcact caactatcac catgaagaag attctcgaga cctacgatgg attcgagggg	600

```

ctcaaaacat tggttgacgt tgggtggtggt gttggtgccg cgcttaacat gatcgtctcc      660
aagcaccctt ccattaaggg cattaacttt gatttgcctc atgtcattga ggatgctcca      720
gctcttcctg gtgttgagca tgttggtgga gatatgtttg taagtgttcc aaaaggagat      780
gccattttca tgaagtggat atgtcatgat tggagcgaat aacactgcgt aaaattcttg      840
aagaagtgct atgaagcttt gccagacaat gggaaagtca tcgttgccga atgcattctt      900
cctgattacc cagatgctag ccttgccaca aagctagttg ttcatatcga ttgtatcatg      960
ttggctcaca accctggttg gaaagaaagg acagagaagg aatttgaagc cttggcaaag     1020
ggggcaggtt ttcaagggtt ccaagtaaag tgttgtgctt ttggcactta catcatggag     1080
ttcctcaaaa ctgtttaa                                           1098

```

<210> 21
 <211> 365
 <212> PRT
 <213> Corchorus olitorius

<400> 21

```

Met Gly Ser Thr Gly Glu Thr Gln Phe Thr Pro Thr Gln Val Ser Asp
1              5              10              15

```

```

Glu Glu Ala Asn Leu Phe Ala Met Gln Leu Ala Ser Ala Ser Val Leu
              20              25              30

```

```

Pro Met Val Leu Lys Ser Ala Ile Glu Leu Asp Leu Leu Glu Val Met
              35              40              45

```

```

Ala Lys Ala Gly Pro Gly Ala Phe Leu Ser Pro Thr Glu Val Ala Ser
50              55              60

```

```

Gln Leu Pro Thr Lys Asn Pro Asp Ala Pro Val Met Leu Asp Arg Ile
65              70              75              80

```

```

Leu Arg Leu Leu Ala Ser Tyr Ser Ile Leu Thr Cys Ser Leu Arg Asn
              85              90              95

```

```

Leu Pro Asp Gly Lys Val Glu Arg Leu Tyr Gly Leu Gly Pro Val Cys
              100              105              110

```

```

Lys Tyr Leu Val Lys Asn Glu Asp Gly Val Ala Leu Ser Ala Leu Asn
115              120              125

```

```

Leu Met Asn Gln Asp Lys Val Leu Met Glu Ser Trp Tyr Tyr Leu Lys
130              135              140

```

Asp Ala Val Leu Glu Gly Gly Ile Pro Phe Asn Lys Ala Tyr Gly Met
 145 150 155 160

Thr Ala Phe Glu Tyr His Gly Thr Asp Pro Arg Phe Asn Lys Val Phe
 165 170 175

Asn Arg Gly Met Ser Asp His Ser Thr Ile Thr Met Lys Lys Ile Leu
 180 185 190

Glu Thr Tyr Asp Gly Phe Glu Gly Leu Lys Thr Leu Val Asp Val Gly
 195 200 205

Gly Gly Val Gly Ala Thr Leu Asn Met Ile Val Ser Lys His Pro Ser
 210 215 220

Ile Lys Gly Ile Asn Phe Asp Leu Pro His Val Ile Glu Asp Ala Pro
 225 230 235 240

Ala Leu Pro Gly Val Glu His Val Gly Gly Asp Met Phe Val Ser Val
 245 250 255

Pro Lys Gly Asp Ala Ile Phe Met Lys Trp Ile Cys His Asp Trp Ser
 260 265 270

Asp Glu His Cys Val Lys Phe Leu Lys Lys Cys Tyr Glu Ala Leu Pro
 275 280 285

Asp Asn Gly Lys Val Ile Val Ala Glu Cys Ile Leu Pro Asp Tyr Pro
 290 295 300

Asp Ala Ser Leu Ala Thr Lys Leu Val Val His Ile Asp Cys Ile Met
 305 310 315 320

Leu Ala His Asn Pro Gly Gly Lys Glu Arg Thr Glu Lys Glu Phe Glu
 325 330 335

Ala Leu Ala Lys Gly Ala Gly Phe Gln Gly Phe Gln Val Lys Cys Cys
 340 345 350

Ala Phe Gly Thr Tyr Ile Met Glu Phe Leu Lys Thr Val
 355 360 365

<210> 22

<211> 1635

<212> ADN

<213> *Corchorus olitorius*

<400> 22

atggcacc	aaagcaccaga	attgccacaa	gaagatatca	tttttcgatc	aaaacttcct	60
gatatactata	ttcctaagca	tcttccttta	cattcatatt	gcttcgagaa	catttcgaag	120
gttgcttcta	agccttggtt	gatcaatggg	acgacgggtc	agatttatac	ttacgaagaa	180
gttgaactca	cagctcgccg	agtcgccgct	gggcttcaca	aactcggcgt	tcaacaacgt	240
caagttataa	tgctttttatt	acccaatact	cccgagtttg	tcctgtcttt	cctcggcgct	300
tcattcctcg	gcgccgtttg	cacggcggtg	aaccggtttt	tcaccgcccc	ggagggttgc	360
aagcaagcca	aagcttccaa	cgccaggatc	ataatcactc	aagcttcgta	cgtcgacaaa	420
gtgaaggaat	tcgcccagga	aaatgatgtc	aaggatcatgt	gcattgactc	agctccggaa	480
ggatgtttac	atttctccga	gttaactcaa	gctgacgaga	acgatctccc	ggaagtcgaa	540
atcaatcccc	acgatgtcgt	ggcacttcct	tattcgtcgg	gaaccaccgg	gctcccaaaa	600
ggtgtgatgt	taactcacia	aggtttggtt	accagcgtgg	ctcaacaggt	cgacggtgaa	660
aacccaaatt	tgtacttcca	cagcgatgac	gtcatcttat	gtactctgcc	catgttccat	720
atctatgccc	tgaactcgat	catgctttgc	gggcttcggg	ccggagctgc	gattttgatc	780
atgcagaagt	ttgacatcgg	attattgttg	gatttgattc	agaaatacaa	aattacgatt	840
gctccgatgg	tgccacccat	agttttggct	attgccaaagt	catcggaac	tgaaaaatac	900
gacttgctct	cgataaggat	ggtgaaatcc	ggcgccgcgc	cgttgggtaa	agagctggaa	960
gatgctgtga	gagccaagtt	tcctggtgcc	aaactcggtc	agggatatgg	gatgacagaa	1020
gcaggaccag	ttctagcaat	gtgcttggtg	tttgccaagg	aaccatttga	aatcaaatcc	1080
ggtgcctgtg	ggacggtggt	tagaaatgca	gagatgaaaa	tcgttgacct	agacaccggt	1140
gcctcacttc	caagaaacca	ggctggagag	atttgcatca	gaggggatca	gatcatgaaa	1200
ggttacctaa	atgaccacga	ggccacagct	aggaccattg	acaaagatgg	ctggttacat	1260
accggtgata	tcggttacat	tgatgacgac	gatgaactct	tcacgttgga	tcgtttgaag	1320
gaattgatca	aatataaggg	tttccagggt	gctcctgctg	agcttgaagc	tatgcttatt	1380
gccaccctg	agattattga	tgctgctgtc	gtcgcaatga	aggatgaggt	agctggagaa	1440
gttcctgttg	catttggtgt	gaaatcagag	aaatcaggaa	tactgagga	tgaaatcaag	1500
caatatat	caaagcaggt	tgtgttctac	aagagaataa	gccgtgtgtt	cttcatggaa	1560
tcaattccaa	aggcaccatc	aggcaagatt	ttgagaaagg	aattgagagc	taaattggct	1620
tctggaaact	actga					1635

<210> 23

<211> 544

<212> PRT

<213> Corchorus olitorius

<400> 23

Met Ala Pro Gln Ala Pro Glu Leu Pro Gln Glu Asp Ile Ile Phe Arg
 1 5 10 15

Ser Lys Leu Pro Asp Ile Tyr Ile Pro Lys His Leu Pro Leu His Ser
 20 25 30

Tyr Cys Phe Glu Asn Ile Ser Lys Val Ala Ser Lys Pro Cys Leu Ile
 35 40 45

Asn Gly Thr Thr Gly Gln Ile Tyr Thr Tyr Glu Glu Val Glu Leu Thr
 50 55 60

Ala Arg Arg Val Ala Ala Gly Leu His Lys Leu Gly Val Gln Gln Arg
 65 70 75 80

Gln Val Ile Met Leu Leu Leu Pro Asn Thr Pro Glu Phe Val Leu Ser
 85 90 95

Phe Leu Gly Ala Ser Phe Leu Gly Ala Val Cys Thr Ala Ala Asn Pro
 100 105 110

Phe Phe Thr Ala Pro Glu Val Ala Lys Gln Ala Lys Ala Ser Asn Ala
 115 120 125

Arg Ile Ile Ile Thr Gln Ala Ser Tyr Val Asp Lys Val Lys Glu Phe
 130 135 140

Ala Gln Glu Asn Asp Val Lys Val Met Cys Ile Asp Ser Ala Pro Glu
 145 150 155 160

Gly Cys Leu His Phe Ser Glu Leu Thr Gln Ala Asp Glu Asn Asp Leu
 165 170 175

Pro Glu Val Glu Ile Asn Pro Asp Asp Val Val Ala Leu Pro Tyr Ser
 180 185 190

Ser Gly Thr Thr Gly Leu Pro Lys Gly Val Met Leu Thr His Lys Gly
 195 200 205

Leu Val Thr Ser Val Ala Gln Gln Val Asp Gly Glu Asn Pro Asn Leu
 210 215 220

Tyr Phe His Ser Asp Asp Val Ile Leu Cys Thr Leu Pro Met Phe His
 225 230 235 240

Ile Tyr Ala Leu Asn Ser Ile Met Leu Cys Gly Leu Arg Ala Gly Ala
 245 250 255

Ala Ile Leu Ile Met Gln Lys Phe Asp Ile Gly Leu Leu Leu Asp Leu
 260 265 270

Ile Gln Lys Tyr Lys Ile Thr Ile Ala Pro Met Val Pro Pro Ile Val
 275 280 285

Leu Ala Ile Ala Lys Ser Ser Glu Thr Glu Lys Tyr Asp Leu Ser Ser
 290 295 300

Ile Arg Met Val Lys Ser Gly Ala Ala Pro Leu Gly Lys Glu Leu Glu
 305 310 315 320

Asp Ala Val Arg Ala Lys Phe Pro Gly Ala Lys Leu Gly Gln Gly Tyr
 325 330 335

Gly Met Thr Glu Ala Gly Pro Val Leu Ala Met Cys Leu Gly Phe Ala
 340 345 350

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

Asn Ala Glu Met Lys Ile Val Asp Pro Asp Thr Gly Ala Ser Leu Pro
 370 375 380

Arg Asn Gln Ala Gly Glu Ile Cys Ile Arg Gly Asp Gln Ile Met Lys
 385 390 395 400

Gly Tyr Leu Asn Asp Pro Glu Ala Thr Ala Arg Thr Ile Asp Lys Asp
 405 410 415

Gly Trp Leu His Thr Gly Asp Ile Gly Tyr Ile Asp Asp Asp Asp Glu
 420 425 430

Leu Phe Ile Val Asp Arg Leu Lys Glu Leu Ile Lys Tyr Lys Gly Phe
 435 440 445

Gln Val Ala Pro Ala Glu Leu Glu Ala Met Leu Ile Ala His Pro Glu
 450 455 460

Ile Ile Asp Ala Ala Val Val Ala Met Lys Asp Glu Val Ala Gly Glu
465 470 475 480

Val Pro Val Ala Phe Val Val Lys Ser Glu Lys Ser Gly Ile Thr Glu
485 490 495

Asp Glu Ile Lys Gln Tyr Ile Ser Lys Gln Val Val Phe Tyr Lys Arg
500 505 510

Ile Ser Arg Val Phe Phe Met Glu Ser Ile Pro Lys Ala Pro Ser Gly
515 520 525

Lys Ile Leu Arg Lys Glu Leu Arg Ala Lys Leu Ala Ser Gly Asn Tyr
530 535 540

<210> 24

<211> 964

<212> ADN

<213> *Corchorus olitorius*

<400> 24

tgtgttaggg atatgggctg actgagactg gaggaggagc gactaggggtg atagggcctg	60
aggagtcagc acggtatggg acggtcgggc gccttgacaga aaatatggaa gccaagatag	120
ttgaccctga aactggagag gccctgcctc ctgggcagag aggggaatta tggttgcgag	180
ggccaacagt aatgaagggt tatataggag atgagaaggc aactgctgaa accttggtgatt	240
cagaaggctg gttaaaaaact ggtgatatat gttattttga ctctgagggg tttctctata	300
ttgtagatag attgaaggaa ttgatcaaat acaaggcata tcagggttcct cccgctgaat	360
tggaacagtt acttcattcc catcctaaaa ttgccgatgc agctgtgatt ccgtaccctg	420
atgaagaagc agggcagatt cccatggcct atgttgtaag aaatgccgga agtagcatca	480
ccgaggcaga agtcatggat ttcgttgcaa aacaggttgc accatacaag aagatccgac	540
gtgttgcttt tatcgattct attccaaaat ctccggcagg aaagatctta aggagggagc	600
tgattaaaca ttctctttcc agtggtttat caaagttata acaatctgga acatacaaca	660
gtcaacagag gagtttgccg actgctgaga aagacctctt agtgaacatg aaaacctact	720
tggccatctg ctgaaggctt tgattccact gttggttatt gttacagttc aaaaacttca	780
gattgaaagt aaattctttt tagatttttag acagcttggg attctaaatg agttccctta	840
ctggcgcat tgatcaccat agcaacagaa caataactaa attgtatact atttgtgttt	900
caagttttca ggggaaaaaa gaaattagtt gagtggtgtt gaacatgaca tgagatgagg	960

gaca

964

<210> 25

<211> 1154

<212> ADN

<213> *Corchorus olitorius*

<400> 25

tctcggactt aaaatgttct tccaccctta caggtttgag tttaaaccctt aagtcaagaa	60
tggttttcagc cttcaatatt cctctttatt tcaagaatac aacagaacaa ttttcatgac	120
aagcaaagga ccaaggggag gaatagtaag aaaggacata attttaagca taaggactac	180
caataagtaa taacatggtg attttttagta gtaacatttt tgactcacta cattaattaa	240
ttagcaaact acaatatcat catgattata aatccaaata aaaatctaca aatactagaa	300
ggaagaaatg aaaatttcaa ggcagggtag cagattgagc agccagctta gctctgagat	360
cttttctcaa tatctttcct gaaggagact tgggaattgc atgaacaaag taaatcttat	420
gcaatctctt gtagaaaacc acctgttttg caatgaattc tttgacagct tcttcagtaa	480
gttcaaaacc atttgatcta accacaaatg caacaggaac ttccccagca acttcatctt	540
tttgcggaac tacagctgca tctgcaattg atggatggct tacaaggaga gattcaagct	600
cagctggtgg cacttgaag cctttgaatt tgatgatttc ctttacccta tcaacaatga	660
aaatctcatc atcttcatca acataaccaa tgtcacctgt atgaagccaa ccctccacat	720
ctatggtggc tgctgtggcc gcgacatcat tcaaataacc taaaatttgt tcattagaaa	780
tgttgaacaa atcaggattg acagaatttc ttttatcaga tcaaacagtt ttaaagatgt	840
ataattagtt aacctttcat gatttgggat ccgcgaatgc aaatttcgcc aggttgattg	900
tagccgaggg agcagccggt ttcagggcca atgaccttaa gctcagcatt tctaaccaca	960
gtgccacatg aaccagactt ggttggaag ggttgcttag caaatcctag gcacattgat	1020
agaactggtc cagcttctgt catcccatat ccctatattt gtgtggaaat gacatagtta	1080
atgcattgta aaaggtagag tgtgtaacac tacatggtat gattctatat aaatgtaata	1140
tggaagagaca taac	1154

<210> 26

<211> 1644

<212> ADN

<213> *Corchorus olitorius*

<400> 26

atggagaaat ctgggtatgg aagagatggg atttacaggt ctctaaggcc acctgtaatt	60
cttcccaagg atccaaatct ttccatgatt tcattttctgt ttagaaacat ttcttcttac	120

```

cctcaccatc cagccctcat tgatgctgat tccaatgaaa ttctaacttt ttctcagttc      180
aaatccactg tcatcaagct ttcccatgct tttctcaatc tgggtatcaa caaaaaagac      240
cgtgtcttga tttttgcacc aaattcgatc caattccctc tctgtttctt cgctgtaacc      300
gccattggcg ctattgctac aaccgccaac cccatttaca ctgtcaatga actctcgaaa      360
caaatcaaag attcgactcc caagcttctt gttactgttc ctgaattggt cgacaaagtt      420
aaggatttca agcttcctgt tatattgctt ggtcctaaac agaacaaacc cccatctcca      480
tctgatgtaa aaaatgtccc aaaaatctta tcttttcatg accttctcga tttagcgggg      540
agcgtgacag agcttcccg cgtttctgtt aagcaaactg atacggcgct acttttatac      600
tcctccggca cgacaggggt aagtaaagggt gttgttttga cgcataaggaa tttcattgca      660
gcggtcttga tgataaccaa ggaccaagaa cttgcggcg ataagcaccg ggttttcttg      720
tgtgttttgc ccttgttcca tgtctttgga ttggcgggta ttgcgttttc acagctgcag      780
ataggaaaca ctttggtttc tatggcaaag ttcgattttg ggttggtttt gaagaatgca      840
gagaagtata aagccacca tttgtgggtt gtgcgcgcaa ttgtgcttgc catggctaag      900
cagagtgtgg ttaagaagtt tgatctttcc tcagtgaggc aaattggctc tggcgctgct      960
cctcttgga aggatttgat ggaggaatgt gcaaaaaatt ttcctcaggc tgtggttatg     1020
caggggtttg gaatgactga aacttggtgc attgtctcag tggagaatcc tacagttggt     1080
gtccgacata ctggttcagc tggaatgctt gtttcaagca ttgaagctca aataatcagt     1140
actgagagtc taaagcctct tcctcccaat caattagggg aaatatgggt tcgagggcct     1200
aatatgatgc aaggttacta caacaatcca gaggcaacaa aactaacaat agataaaaaag     1260
ggttggttac atacaggaga tcttggtac tttgatgaag atgggaatct ttatgttggt     1320
gaccgaatta aagagttgat caaatataaa ggattccaga ttgcaccagc cgaacttgaa     1380
ggactacttg tatctcatcc tgaaatattg gatgctgttg tcatcccgta tcctgacgct     1440
gaagctggtg aggttccggt tgcatatggt gttcgctccc ctaacagctc actgactgag     1500
gaggatgtcc aaaattttat agctaaacag gtggcaccgt tcaaaagact aaggagagtt     1560
acattcataa cgagtgtccc aaagtcgggt tcaggaaaaa tcctgaggag agagcttata     1620
gcggaagtaa gatccaagat gtga                                             1644

```

<210> 27

<211> 547

<212> PRT

<213> *Corchorus olitorius*

<400> 27

Met Glu Lys Ser Gly Tyr Gly Arg Asp Gly Ile Tyr Arg Ser Leu Arg

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

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

485

490

495

Ser Leu Thr Glu Glu Asp Val Gln Asn Phe Ile Ala Lys Gln Val Ala
 500 505 510

Pro Phe Lys Arg Leu Arg Arg Val Thr Phe Ile Thr Ser Val Pro Lys
 515 520 525

Ser Ala Ser Gly Lys Ile Leu Arg Arg Glu Leu Ile Ala Glu Val Arg
 530 535 540

Ser Lys Met
 545

<210> 28

<211> 1089

<212> ADN

<213> *Corchorus olitorius*

<400> 28

tccaacacca acggcggcac caccgccgcc acggacacct tatgcctctg tatcaattcc	60
aacaacgccc ctatctcaaa tttctgcata agcaaaaccg cagcccctgc tctcaaggaa	120
cacaaaagca cgctgttgag tgagtaaata tgaaacaatg gcaaaacgca taacacaacg	180
tcgtcgtttt tcaggtaaag attgggggtc tctccatcta cttgctgcgc tacgctcgtg	240
atcaagcttt tgtgggtcaa aacgactcct ttgggtaacc ctgttggtcc tgaagagaat	300
ggtaaagcaa cagggtcatc agggtaaatg gaaacttcgg ggatatcgtt ttcatttccc	360
tctgataaca ccgtgaaatg caagcagttt tctgggggat catcgatggg tacaactttg	420
aaatcttggc caattttagg gaagttttgg ttttgggtcg tctcatcctt aagcttgctg	480
acatattggg attgcgtgat aataatctta gcccagacag ctttgaattg tttgaagatt	540
tcgtttgatg tgtaaaaagg gttggcagtt gtggaaactg ctccaatcat ggaagctccc	600
atgaaagaga agacgaattc agcacagttt gggaggagaa tcatgatgac atcgcccttc	660
tcgatgccta aatttgacaa accagcggct gtctttcgag aaatcaagtg agtttcagag	720
aaagtgtagg ttttggcaga ggatattcca gagatcaaac atggcttgctc agggaaagag	780
gagagttttt caaagcagta agtatggaga gggaggtgat tggagatggg aatgtcaggc	840
aatttggacc tgaaaatgtg atgattagtt tgaggggaag gttttggtgt tgtcgaaagc	900
tctgctgcag gcttttgtgg ttcaacagga tgatcagcta tggagatcat ggtggtaatt	960
tggtttgatg agtacttggt gggatttttg ggttagaggg aaagtgaatg aaaggaaaga	1020
aaagagaaat aaatataagg taataagaag tggcaagcaa gggccaagga gaggtagggt	1080

gaagaggaa

1089

<210> 29

<211> 1689

<212> ADN

<213> *Corchorus olitorius*

<400> 29

atggcaacga gcttgaactc ccatttcagc ttcccaacct ccgaatcaa aacaactcgt	60
tttcctgact ggtattcacc ggaaacagga atctatacca gcaagcatgc ctccgtatcc	120
atccccacta atccatatct tgatgttggt tcttccattt tctctcacca acatcaaggg	180
gtcactgctt tcattgattc ctcatctggg ttttcaatat cttactcaaa gctcttacct	240
ttggccaat ctatggcttc tggctccac cacctgggtg tttccaaagg tgacgtggtc	300
ttgcttttgt tgccaaattc tcttcaactat cctattattt tcttcagtgt tttatattta	360
ggcgcaatcg ttaccctat gaatccactg agtagcattt ccgagatcaa gaaacagatt	420
gctgattcta atgtgcgttt cgctttcact cttcttgaaa cgggtgacaa actggagaag	480
ttgggtgttc atgcaattgg ggtaccggaa aacatgaact tggattcaga aaagggtgat	540
tttttacctt tttataagct tatggcgggg caatctggta ataaggcccc aaggccagt	600
attaagcagc aagacactgc ggcgataatg tattcatcgg gaactacagg aacgagtaag	660
ggagttgtat taacacatgg gaatttcata gcaatgattg agctttttgt aaaatttgaa	720
gcttcacagt atgaatatcc aggttcagag attgtatatt tagctgctct accaatgttc	780
catatatatg ggctatcact gtttgtgggt ggattgttat cgttggggtc tacagttgtt	840
gtcatgagga aatttaatgc tggatgaatt gtaaaagtaa ttgataagtt tgggatcacc	900
cactttccag ttgttccacc tatactcaca gcattgacaa taagcgccaa ggggtgttgt	960
gaaaataact tcaagagctt gaaacagggt tcttgcggtg ctgctcctat aagcaggaaa	1020
tccatagagg attttgttca ggctttccct catgttgatt tcattcaggg ctatgggatg	1080
acagaatcaa ctgcagtagg aactcggggc ttcaacaccg gaaaacatca taaatattct	1140
tcaataggac ttctagcacc aaacatgcaa gctaaagtgg tggatctgaa ttctggttct	1200
tctatgcctc ctgggtgatta cggcgagctt tgggttaagag gacctgcaat tatgcaacga	1260
tacttgaata atgttgaagc caccctgatg tcaatccaca gagatgggtg gctacgtact	1320
ggtgacattg cttgttttga tgaagatggc tatatgtatt tatctgaccg cttaaaagag	1380
attataaagt acaaaggcta tcagatagct cctgccgatt tagaggccat attaattacc	1440
catcctgaga tacttgacgc tgctgtaact ggagccagtg atgaagcatg tggtgagatt	1500

cctgtggcat ttgtggtgag gaggcattggt tgcacactga cccatggcgc tgtcatggac 1560
 ttcgtggcta agcaggttgc accttataag aaagtaagaa aggtggtggt ttcaaattca 1620
 ataccgaggt ctgctgcagg aaagatcctc cgaagagaac tcaagaagtt cttatgttca 1680
 aggctttaa 1689

<210> 30

<211> 562

<212> PRT

<213> Corchorus olitorius

<400> 30

Met Ala Thr Ser Leu Asn Ser His Phe Ser Phe Pro Thr Ser Glu Ser
 1 5 10 15

Lys Thr Thr Arg Phe Pro Asp Trp Tyr Ser Pro Glu Thr Gly Ile Tyr
 20 25 30

Thr Ser Lys His Ala Ser Val Ser Ile Pro Thr Asn Pro Tyr Leu Asp
 35 40 45

Val Val Ser Phe Ile Phe Ser His Gln His Gln Gly Val Thr Ala Phe
 50 55 60

Ile Asp Ser Ser Ser Gly Phe Ser Ile Ser Tyr Ser Lys Leu Leu Pro
 65 70 75 80

Leu Val Gln Ser Met Ala Ser Gly Leu His His Leu Gly Val Ser Lys
 85 90 95

Gly Asp Val Val Leu Leu Leu Leu Pro Asn Ser Leu His Tyr Pro Ile
 100 105 110

Ile Phe Phe Ser Val Leu Tyr Leu Gly Ala Ile Val Thr Pro Met Asn
 115 120 125

Pro Leu Ser Ser Ile Ser Glu Ile Lys Lys Gln Ile Ala Asp Ser Asn
 130 135 140

Val Arg Phe Ala Phe Thr Leu Leu Glu Thr Val Asp Lys Leu Glu Lys
 145 150 155 160

Leu Gly Val His Ala Ile Gly Val Pro Glu Asn Met Asn Leu Asp Ser
 165 170 175

Glu Lys Val Asp Phe Leu Pro Phe Tyr Lys Leu Met Ala Gly Gln Ser

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

420 425 430
 His Arg Asp Gly Trp Leu Arg Thr Gly Asp Ile Ala Cys Phe Asp Glu
 435 440 445
 Asp Gly Tyr Met Tyr Leu Ser Asp Arg Leu Lys Glu Ile Ile Lys Tyr
 450 455 460
 Lys Gly Tyr Gln Ile Ala Pro Ala Asp Leu Glu Ala Ile Leu Ile Thr
 465 470 475 480
 His Pro Glu Ile Leu Asp Ala Ala Val Thr Gly Ala Ser Asp Glu Ala
 485 490 495
 Cys Gly Glu Ile Pro Val Ala Phe Val Val Arg Arg His Gly Cys Thr
 500 505 510
 Leu Thr His Gly Ala Val Met Asp Phe Val Ala Lys Gln Val Ala Pro
 515 520 525
 Tyr Lys Lys Val Arg Lys Val Val Phe Ser Asn Ser Ile Pro Arg Ser
 530 535 540
 Ala Ala Gly Lys Ile Leu Arg Arg Glu Leu Lys Lys Phe Leu Cys Ser
 545 550 555 560
 Arg Leu

<210> 31
 <211> 1299
 <212> ADN
 <213> Corchorus olitorius

<400> 31
 atgattgtga acgtgaaaga atcaacaatg gtacctccgg cggaggagac tccacgggtt 60
 tgcttatgga actccaatgt ggacttggtg gtgccaggt tccatacccc gagcgtctac 120
 ttctataggc catcgggtgc atccaacttc tttgacccaa aggtcatgaa ggaggctctg 180
 ggcaaggcct tggtgccgtt ttaccccatg gcggggcggt tgaagaggga tgaagatgga 240
 aggattgaga ttgattgcaa tggatgcagggt gtgctctttg ttgaggctga gactaatgct 300
 gtcattgatg attttggtga ttttgctccc actttggagc tcaggcagct cattccaact 360
 gttgattatt ctggtggcat cgagacttac ccgctcttgg ttttgcaggt cacttatttc 420
 aaatgtggtg gagcatcact tgggtgttggc atgcaacatc atgcggcaga tggcttttct 480

```

ggtctccatt ttatcaatac atgggtccgat atgggtcgtg gtcttgacct cacaattcca      540
ccattcatcg atcgtaacct gctccgtgcc cgggatccac cgcaacctgc attcgagcac      600
attgaatacc aaccacctcc tgcattgaaa tctgcacctg aatccacagg ttctgaaggt      660
gcagcagtct ccattttcaa attgacctga gaacagctaa atgcacttaa agctaagtcc      720
aaggaagatg ggaacactat tgcttatagc tcatatgaga tggtgtcagg tcatgtatgg      780
agatcagtct gcaaagcacg tggacttcct gatgatcaag agtcaaaatt gtacattgcc      840
actgatggaa gggctaggtt gcgcccccca cttccacctg gttactttgg aaatgttatt      900
ttcacgccta cccaattgc agtggccggt gagctaattgt caaagccaac atgggtatgct      960
gctgggaaaa ttcatgatgc cttgggttcgc atggacaatg attatctaaa gtcagccctc     1020
gattacctag aacttcagcc tgatttatct gcccttggtc gtggagcaca tacatttaag     1080
tgtccgaatc ttgggattac tagttggtca aggctgccaa tccacgatgc agattttgga     1140
tgggggccgac ccatatttat gggtcctggt ggaatccctt atgagggggtt atcttttgtg     1200
ttaccaagtc caaccaatga tgggagctta tcagttgccca tcgctctgca aaccgaacac     1260
atgaaactgt ttgagaagat cttttatgat gacatataa                               1299

```

<210> 32

<211> 432

<212> PRT

<213> *Corchorus olitorius*

<400> 32

```

Met Ile Val Asn Val Lys Glu Ser Thr Met Val Pro Pro Ala Glu Glu
1           5           10           15

```

```

Thr Pro Arg Val Cys Leu Trp Asn Ser Asn Val Asp Leu Val Val Pro
          20           25           30

```

```

Arg Phe His Thr Pro Ser Val Tyr Phe Tyr Arg Pro Ser Gly Ala Ser
          35           40           45

```

```

Asn Phe Phe Asp Pro Lys Val Met Lys Glu Ala Leu Gly Lys Ala Leu
50           55           60

```

```

Val Pro Phe Tyr Pro Met Ala Gly Arg Leu Lys Arg Asp Glu Asp Gly
65           70           75           80

```

```

Arg Ile Glu Ile Asp Cys Asn Gly Ala Gly Val Leu Phe Val Glu Ala
          85           90           95

```

Glu Thr Asn Ala Val Ile Asp Asp Phe Gly Asp Phe Ala Pro Thr Leu
 100 105 110

Glu Leu Arg Gln Leu Ile Pro Thr Val Asp Tyr Ser Gly Gly Ile Glu
 115 120 125

Thr Tyr Pro Leu Leu Val Leu Gln Val Thr Tyr Phe Lys Cys Gly Gly
 130 135 140

Ala Ser Leu Gly Val Gly Met Gln His His Ala Ala Asp Gly Phe Ser
 145 150 155 160

Gly Leu His Phe Ile Asn Thr Trp Ser Asp Met Ala Arg Gly Leu Asp
 165 170 175

Leu Thr Ile Pro Pro Phe Ile Asp Arg Thr Leu Leu Arg Ala Arg Asp
 180 185 190

Pro Pro Gln Pro Ala Phe Glu His Ile Glu Tyr Gln Pro Pro Pro Ala
 195 200 205

Leu Lys Ser Ala Pro Glu Ser Thr Gly Ser Glu Gly Ala Ala Val Ser
 210 215 220

Ile Phe Lys Leu Thr Arg Glu Gln Leu Asn Ala Leu Lys Ala Lys Ser
 225 230 235 240

Lys Glu Asp Gly Asn Thr Ile Ala Tyr Ser Ser Tyr Glu Met Leu Ser
 245 250 255

Gly His Val Trp Arg Ser Val Cys Lys Ala Arg Gly Leu Pro Asp Asp
 260 265 270

Gln Glu Ser Lys Leu Tyr Ile Ala Thr Asp Gly Arg Ala Arg Leu Arg
 275 280 285

Pro Pro Leu Pro Pro Gly Tyr Phe Gly Asn Val Ile Phe Thr Ala Thr
 290 295 300

Pro Ile Ala Val Ala Gly Glu Leu Met Ser Lys Pro Thr Trp Tyr Ala
 305 310 315 320

Ala Gly Lys Ile His Asp Ala Leu Val Arg Met Asp Asn Asp Tyr Leu
 325 330 335

Lys Ser Ala Leu Asp Tyr Leu Glu Leu Gln Pro Asp Leu Ser Ala Leu
 340 345 350

Val Arg Gly Ala His Thr Phe Lys Cys Pro Asn Leu Gly Ile Thr Ser
 355 360 365

Trp Ser Arg Leu Pro Ile His Asp Ala Asp Phe Gly Trp Gly Arg Pro
 370 375 380

Ile Phe Met Gly Pro Gly Gly Ile Pro Tyr Glu Gly Leu Ser Phe Val
 385 390 395 400

Leu Pro Ser Pro Thr Asn Asp Gly Ser Leu Ser Val Ala Ile Ala Leu
 405 410 415

Gln Thr Glu His Met Lys Leu Phe Glu Lys Ile Phe Tyr Asp Asp Ile
 420 425 430

<210> 33

<211> 1062

<212> ADN

<213> *Corchorus olitorius*

<400> 33

atgccaat	cgatgcaga	aagcaagaat	aaaactttgc	aagtaaggga	ttacttggga	60
gcagtggcat	tcaacaacat	aacaaggcta	gtatttggga	agcgttttat	gaactctgag	120
ggcataattg	acgagcaagg	caaagaattc	aagggcattg	tgtcaaattg	aactaagatt	180
ggtgcatccc	ttgccatggc	agagcacatt	ccatggcttc	gttgatgtt	tcccttggaa	240
gaggaagcat	ttgccaagca	tgcagcaagg	agggacaatc	ttaccagaac	catcatggag	300
gagcacaccg	ctgctcgcaa	gaagagcggg	ggtgctaagc	agcattttgt	tgatgccttg	360
ctcacattgc	aagaaaagta	tgacctcagc	gacgacacag	ttattggact	actttgggac	420
atgattacag	caggcatgga	tacaacagca	attgcagcag	agtgggcaat	ggcagagtta	480
atcaagaacc	caagagtgca	gcaaaaggca	caagaggagc	tagatcgtgt	ggtaggattc	540
gaaaggggtga	tgtccgaaac	tgattttctca	agcctgcctt	accttcaaag	tgtaaccaag	600
gaggcattca	gaatgcaccc	cccaactcct	ctaattgctac	cccacaaagc	caacgccaat	660
gtcaaaatcg	gaggttatga	catccccaag	ggatcaaattg	tgcattgtcaa	cgtctgggca	720
gtggccaatg	atccggctgt	atggaaggac	cctgaagtgt	tccggccaga	gcgattcctg	780
gaggaggatg	tggacatgaa	gggtcatgat	tatcgcttgc	ttccttttgg	tgcggggagg	840
agggatatgcc	ctggagcaca	acttgggatc	aacctgggtca	catccatgtt	gggtcactta	900

ctgcaccatt ttgtttggac accaccagag ggagtaaagg ccgaggaaat cgacatggct 960
 gaaaatcccg gacttgttgc ctacatgaag actcctgtgc aggtgtgtggc cactcctagg 1020
 ctgccttccg atctctacaa acgtgtagct gttgacatat aa 1062

<210> 34

<211> 353

<212> PRT

<213> Corchorus olitorius

<400> 34

Met Pro Ile Ser Asp Ala Glu Ser Lys Asn Lys Thr Leu Gln Val Arg
 1 5 10 15

Asp Tyr Leu Gly Ala Val Ala Phe Asn Asn Ile Thr Arg Leu Val Phe
 20 25 30

Gly Lys Arg Phe Met Asn Ser Glu Gly Ile Ile Asp Glu Gln Gly Lys
 35 40 45

Glu Phe Lys Gly Ile Val Ser Asn Gly Thr Lys Ile Gly Ala Ser Leu
 50 55 60

Ala Met Ala Glu His Ile Pro Trp Leu Arg Trp Met Phe Pro Leu Glu
 65 70 75 80

Glu Glu Ala Phe Ala Lys His Ala Ala Arg Arg Asp Asn Leu Thr Arg
 85 90 95

Thr Ile Met Glu Glu His Thr Ala Ala Arg Lys Lys Ser Gly Gly Ala
 100 105 110

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

Leu Ser Asp Asp Thr Val Ile Gly Leu Leu Trp Asp Met Ile Thr Ala
 130 135 140

Gly Met Asp Thr Thr Ala Ile Ala Ala Glu Trp Ala Met Ala Glu Leu
 145 150 155 160

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

Val Val Gly Phe Glu Arg Val Met Ser Glu Thr Asp Phe Ser Ser Leu
 180 185 190

Pro Tyr Leu Gln Ser Val Thr Lys Glu Ala Phe Arg Met His Pro Pro
 195 200 205

Thr Pro Leu Met Leu Pro His Lys Ala Asn Ala Asn Val Lys Ile Gly
 210 215 220

Gly Tyr Asp Ile Pro Lys Gly Ser Asn Val His Val Asn Val Trp Ala
 225 230 235 240

Val Ala Asn Asp Pro Ala Val Trp Lys Asp Pro Glu Val Phe Arg Pro
 245 250 255

Glu Arg Phe Leu Glu Glu Asp Val Asp Met Lys Gly His Asp Tyr Arg
 260 265 270

Leu Leu Pro Phe Gly Ala Gly Arg Arg Val Cys Pro Gly Ala Gln Leu
 275 280 285

Gly Ile Asn Leu Val Thr Ser Met Leu Gly His Leu Leu His His Phe
 290 295 300

Val Trp Thr Pro Pro Glu Gly Val Lys Ala Glu Glu Ile Asp Met Ala
 305 310 315 320

Glu Asn Pro Gly Leu Val Ala Tyr Met Lys Thr Pro Val Gln Ala Val
 325 330 335

Ala Thr Pro Arg Leu Pro Ser Asp Leu Tyr Lys Arg Val Ala Val Asp
 340 345 350

Ile

<210> 35

<211> 1446

<212> ADN

<213> *Corchorus olitorius*

<400> 35

atggatcttc tcttcttgga gaaggccctt attggtcttt tcgtggctgt catcttagcc 60

atcgccatct ctaaactccg gggaaagcgt tacaagctcc ctctgggtcc ttaccgggtt 120

cccgctcttcg gcaactggct ccaagtgggc gacgacttga accaccgtaa cctaactgac 180

ttggccaaga aatatggcga catattcctc ctccgaatgg ggcagcgcaa cctgggtggtg 240

gtgtcgtctc cggaactagc caaagaggtg ctccacaccc agggagtggga attcgggtca 300

```

agaacccgga atgtggtggt tgacatatct accggcaagg gtcaagacat ggtgttcacg      360
gtgtacgttg tccagcagta ccggtttgga tgggaggagg aagccgctcg cggtgtggag      420
gatgtcaaga aaaatcccga ggctgccacc aatggcatcg ttctgaggag gaggttgacg      480
ctcatgatgt acaacaatat gtacagaatc atgttcgacc ggaggttcga gagcgaggag      540
gatcctctgt ttgttaaact caaggctttg aacggggaga ggagtcgatt ggctcagagt      600
tttgagtaca attatggaga ttttattcct attttgaggc ctttcttgag aggttacttg      660
aagatctgca aggagggtta agagaggagg ttgcaactct tcaaggacta ctttgtcgaa      720
gagagaaaga aacttgcaag cacgaagagc atgagcaacg aaggattgaa atgtgccata      780
gatcatatct tggatgctca gcagaaaggg gagatcaacg aggacaatgt tctgtatatc      840
gtcgagaaca tcaatgttgc cgcaattgag acaacattat ggtcgatcga gtggggcatt      900
gcagagctgg tgaaccaccc tgaaatccag aagaagctgc gagatgaact tgacactctt      960
cttggaaccc gccaccagat caccgaaccc gacacctaca aactccctta ccttcaggct     1020
gtcatcaagg agactttgag gcttcgtatg gccattcctc tgctcgctcc ccacatgaac     1080
ctccacgatg ctaagcttgc cggctatgat atccccgctg agagcaaaat cttggtcaat     1140
gcatggtggc tcgccaacaa ccctgcccag tggaaaaaac cccaggagtt taggcccag      1200
aggttctttg aagaggaatc taaggttgag gccaatggca atgacttcag gtaccttcca     1260
tttggggttg gaagaagaag ttgccctgga attattctcg ctttgcccat ccttggtatc     1320
actttgggcc gcttagtaca gaatttcgag ctcttgccct caccaggaca gtccaagatt     1380
gatacctccg agaaaggtgg acagttcagt ttgcacatct tgaagcattc caccattggt     1440
ttgtaa                                           1446

```

<210> 36

<211> 481

<212> PRT

<213> *Corchorus olitorius*

<400> 36

```

Met Asp Leu Leu Phe Leu Glu Lys Ala Leu Ile Gly Leu Phe Val Ala
1           5           10           15

```

```

Val Ile Leu Ala Ile Ala Ile Ser Lys Leu Arg Gly Lys Arg Tyr Lys
20           25           30

```

```

Leu Pro Pro Gly Pro Leu Pro Val Pro Val Phe Gly Asn Trp Leu Gln
35           40           45

```

Val Gly Asp Asp Leu Asn His Arg Asn Leu Thr Asp Leu Ala Lys Lys
 50 55 60

Tyr Gly Asp Ile Phe Leu Leu Arg Met Gly Gln Arg Asn Leu Val Val
 65 70 75 80

Val Ser Ser Pro Glu Leu Ala Lys Glu Val Leu His Thr Gln Gly Val
 85 90 95

Glu Phe Gly Ser Arg Thr Arg Asn Val Val Phe Asp Ile Phe Thr Gly
 100 105 110

Lys Gly Gln Asp Met Val Phe Thr Val Tyr Val Val Gln Gln Tyr Arg
 115 120 125

Phe Gly Trp Glu Glu Glu Ala Ala Arg Val Val Glu Asp Val Lys Lys
 130 135 140

Asn Pro Glu Ala Ala Thr Asn Gly Ile Val Leu Arg Arg Arg Leu Gln
 145 150 155 160

Leu Met Met Tyr Asn Asn Met Tyr Arg Ile Met Phe Asp Arg Arg Phe
 165 170 175

Glu Ser Glu Glu Asp Pro Leu Phe Val Lys Leu Lys Ala Leu Asn Gly
 180 185 190

Glu Arg Ser Arg Leu Ala Gln Ser Phe Glu Tyr Asn Tyr Gly Asp Phe
 195 200 205

Ile Pro Ile Leu Arg Pro Phe Leu Arg Gly Tyr Leu Lys Ile Cys Lys
 210 215 220

Glu Val Lys Glu Arg Arg Leu Gln Leu Phe Lys Asp Tyr Phe Val Glu
 225 230 235 240

Glu Arg Lys Lys Leu Ala Ser Thr Lys Ser Met Ser Asn Glu Gly Leu
 245 250 255

Lys Cys Ala Ile Asp His Ile Leu Asp Ala Gln Gln Lys Gly Glu Ile
 260 265 270

Asn Glu Asp Asn Val Leu Tyr Ile Val Glu Asn Ile Asn Val Ala Ala
 275 280 285

Ile Glu Thr Thr Leu Trp Ser Ile Glu Trp Gly Ile Ala Glu Leu Val
 290 295 300

Asn His Pro Glu Ile Gln Lys Lys Leu Arg Asp Glu Leu Asp Thr Leu
 305 310 315 320

Leu Gly Pro Gly His Gln Ile Thr Glu Pro Asp Thr Tyr Lys Leu Pro
 325 330 335

Tyr Leu Gln Ala Val Ile Lys Glu Thr Leu Arg Leu Arg Met Ala Ile
 340 345 350

Pro Leu Leu Val Pro His Met Asn Leu His Asp Ala Lys Leu Ala Gly
 355 360 365

Tyr Asp Ile Pro Ala Glu Ser Lys Ile Leu Val Asn Ala Trp Trp Leu
 370 375 380

Ala Asn Asn Pro Ala Gln Trp Lys Asn Pro Gln Glu Phe Arg Pro Glu
 385 390 395 400

Arg Phe Phe Glu Glu Glu Ser Lys Val Glu Ala Asn Gly Asn Asp Phe
 405 410 415

Arg Tyr Leu Pro Phe Gly Val Gly Arg Arg Ser Cys Pro Gly Ile Ile
 420 425 430

Leu Ala Leu Pro Ile Leu Gly Ile Thr Leu Gly Arg Leu Val Gln Asn
 435 440 445

Phe Glu Leu Leu Pro Pro Pro Gly Gln Ser Lys Ile Asp Thr Ser Glu
 450 455 460

Lys Gly Gly Gln Phe Ser Leu His Ile Leu Lys His Ser Thr Ile Val
 465 470 475 480

Leu

<210> 37

<211> 1518

<212> ADN

<213> Corchorus olitorius

<400> 37

atggacctcc tcttcctcga aaaagctctc atttcccttt tcgtcaccat aattgtagcc

60

```

attgtagtct ccaagctccg cggcaagcgt tacaaacttc cccttgcccc aatcccagtc      120
cctgttttcg gcaactggct ccaagtgcgc gatgacttga accaccgcaa cctcaccgac      180
ttggccaaga agttcggaga cattttcttg ctccgtatgg gacagcgcaa cctcgtcgtt      240
gtgtcttccc cggagctggc caaagaagtc ctccataccc agggagttga gtttggttca      300
agaactagaa atgtggtggt tgatattttc acagggaaag gtcaggacat ggttttctact      360
gtctatggtg aacattggcg caaatgagg agaatcatga ccgtcccttt cttaccaaac      420
aaggttgttc aacagtatag acatgggtgg gaagcggagg ttgccgccgt cgttgaggat      480
gtgaagaaga acccagagtc ggccaccact gggattgttc tgaggaagag attgcagctt      540
atgatgtaca acaatatgta ccggatcatg tttgatagaa ggtttgagag tgaagatgat      600
cctttgtttg ttaagctcaa ggctttgaat ggtgagagga gtagattggc acagagcttt      660
gattacaact atggtgattt catcccaatt ttgaggcctt tcttgagagg gtatttgaag      720
ttatgcaagg aagtcaaaga aatgagattg caactcttta gggaccattt ccttgaggag      780
aggaagaagc tttcaagcac aaaaaggcct gacaacaatg ctctgaagtg tgccattgat      840
cacattcttg atgctcagca gaaaggagaa atcaatgaag ataatgttct ctacattggt      900
gagaatatca atgttgctgc cattgaaaca actttgtggt caattgaatg gggaattgct      960
gagcttgtga accatcctga gatccagcag aagctccgca atgaaatcga cactgtactc     1020
ggaccaggag tgcaagttac cgaacccgac acccacaagc ttccatatct ccaggcagtg     1080
atcaaggaga ccctccggct ccggatggcc atccctctat tagtgccaca catgaacctc     1140
catgatgcta agcttggtgg ctatgacatc ccagctgaaa gcaagatcct tgttaacgca     1200
tgggtggttag ccaacaaccc ggctcaatgg aagaaccggg aagagttcag gcccgaaagg     1260
tttttcgagg aggaagctaa ggttgaagcc aatggaaatg acttcaggta ccttccattt     1320
ggtggttgaa ggaggagttg cccaggaatt attcttgccct tgccaatctt aggaatcaca     1380
ttgggacggt tggtgcaaaa ctttgagcta ttgcctcctc ctggacagtc aaagcttgat     1440
acctcagaga aaggaggaca attcagcttg cacattctca agcattcaac tattgttgca     1500
aagccacgag tcttttaa                                     1518

```

<210> 38

<211> 505

<212> PRT

<213> Corchorus olitorius

<400> 38

```

Met Asp Leu Leu Phe Leu Glu Lys Ala Leu Ile Ser Leu Phe Val Thr
1           5           10          15

```

Ile Ile Val Ala Ile Val Val Ser Lys Leu Arg Gly Lys Arg Tyr Lys
20 25 30

Leu Pro Pro Gly Pro Ile Pro Val Pro Val Phe Gly Asn Trp Leu Gln
35 40 45

Val Gly Asp Asp Leu Asn His Arg Asn Leu Thr Asp Leu Ala Lys Lys
50 55 60

Phe Gly Asp Ile Phe Leu Leu Arg Met Gly Gln Arg Asn Leu Val Val
65 70 75 80

Val Ser Ser Pro Glu Leu Ala Lys Glu Val Leu His Thr Gln Gly Val
85 90 95

Glu Phe Gly Ser Arg Thr Arg Asn Val Val Phe Asp Ile Phe Thr Gly
100 105 110

Lys Gly Gln Asp Met Val Phe Thr Val Tyr Gly Glu His Trp Arg Lys
115 120 125

Met Arg Arg Ile Met Thr Val Pro Phe Phe Thr Asn Lys Val Val Gln
130 135 140

Gln Tyr Arg His Gly Trp Glu Ala Glu Val Ala Ala Val Val Glu Asp
145 150 155 160

Val Lys Lys Asn Pro Glu Ser Ala Thr Thr Gly Ile Val Leu Arg Lys
165 170 175

Arg Leu Gln Leu Met Met Tyr Asn Asn Met Tyr Arg Ile Met Phe Asp
180 185 190

Arg Arg Phe Glu Ser Glu Asp Asp Pro Leu Phe Val Lys Leu Lys Ala
195 200 205

Leu Asn Gly Glu Arg Ser Arg Leu Ala Gln Ser Phe Asp Tyr Asn Tyr
210 215 220

Gly Asp Phe Ile Pro Ile Leu Arg Pro Phe Leu Arg Gly Tyr Leu Lys
225 230 235 240

Leu Cys Lys Glu Val Lys Glu Met Arg Leu Gln Leu Phe Arg Asp His
245 250 255

Phe Leu Glu Glu Arg Lys Lys Leu Ser Ser Thr Lys Arg Pro Asp Asn
 260 265 270

Asn Ala Leu Lys Cys Ala Ile Asp His Ile Leu Asp Ala Gln Gln Lys
 275 280 285

Gly Glu Ile Asn Glu Asp Asn Val Leu Tyr Ile Val Glu Asn Ile Asn
 290 295 300

Val Ala Ala Ile Glu Thr Thr Leu Trp Ser Ile Glu Trp Gly Ile Ala
 305 310 315 320

Glu Leu Val Asn His Pro Glu Ile Gln Gln Lys Leu Arg Asn Glu Ile
 325 330 335

Asp Thr Val Leu Gly Pro Gly Val Gln Val Thr Glu Pro Asp Thr His
 340 345 350

Lys Leu Pro Tyr Leu Gln Ala Val Ile Lys Glu Thr Leu Arg Leu Arg
 355 360 365

Met Ala Ile Pro Leu Leu Val Pro His Met Asn Leu His Asp Ala Lys
 370 375 380

Leu Gly Gly Tyr Asp Ile Pro Ala Glu Ser Lys Ile Leu Val Asn Ala
 385 390 395 400

Trp Trp Leu Ala Asn Asn Pro Ala Gln Trp Lys Asn Pro Glu Glu Phe
 405 410 415

Arg Pro Glu Arg Phe Phe Glu Glu Glu Ala Lys Val Glu Ala Asn Gly
 420 425 430

Asn Asp Phe Arg Tyr Leu Pro Phe Gly Val Gly Arg Arg Ser Cys Pro
 435 440 445

Gly Ile Ile Leu Ala Leu Pro Ile Leu Gly Ile Thr Leu Gly Arg Leu
 450 455 460

Val Gln Asn Phe Glu Leu Leu Pro Pro Pro Gly Gln Ser Lys Leu Asp
 465 470 475 480

Thr Ser Glu Lys Gly Gly Gln Phe Ser Leu His Ile Leu Lys His Ser
 485 490 495

Thr Ile Val Ala Lys Pro Arg Val Phe
500 505

<210> 39
<211> 1580
<212> ADN
<213> *Corchorus olitorius*

<400> 39
ggattgttat ctctctcagg tcataatddd aaaacatggc tcacccctct aagaaagcaa 60
ttcttgccat tataacacta tccttgatct ttgttcttgc taccaccaa aacttgcccc 120
tcaccctddd gccactcttt ctgtgcacct actttttatt gcagaaacgt tctaattctt 180
ccaatgccct cctccaggt cctctctctg ttcccatatt tggaaactgg ctccaggttg 240
gaaatgacct aaaccaccgg ctccatagctt ccttggtga aacctatggc ccagtcttcc 300
tcctaaaact tggctccaag aatctagcag tggtttcaga ccagagcta gcctcccagg 360
ttcttcacac ccaaggggtc gaattcgggt ccaggccacg caacgtgggtg ttgacattt 420
tcacgggcaa tgggcaggac atgggtgttca cagtttacgg agaccattgg cgaaaaatgc 480
gcagaattat gactttgcca tttttcacta acaaagttgt gcacaattac agtaacatgt 540
gggaggaaga gatggagctt gtggttagtg acttgaagag agatgaggaa ttggtgaaga 600
gcaaagggat tgttatcaga aaacgtctgc agcttatgct ttataacatc atgtatagga 660
tgatgtttga tgccaagttc gagtccatgg aagacccttt attcgttgag gcaaccaggt 720
tcaattccga gagaagccgc ttggctcaga gttttgagta caattatggt gatttcattc 780
cattgctcag accctttttg agaggggtact tgaacaagtg cagggatttg cagagcaggc 840
ggcttgccct cttcaacaac tattatgtta agaaaagaag ggaaattatg ggtgctaattg 900
gagagaagca caaaatcagc tgtgcaattg attacataat agatgctgaa atgaaggag 960
agattagtga agaaaatgtg ctctacattg tggaaaacat caatgttgca gccattgaaa 1020
ccactctgtg gtcaatggag tgggcaatag ctgaggtggt gaaccacca aatgtgcagc 1080
aaaagatccg ccaagaaatc tcacaagtcc tcaaaggaga ggctgtcaca gaatcaaacc 1140
tccttgaatt gccttacttg caagccactg tcaaggagac actacgggta cacaccccaa 1200
ttcctctgtt gggttcctac atgaaccttg aagaggcaaa attgggagggg ttcacaattc 1260
caaaggagtc caaggttgta gtcaatgcct ggtggctggc caacaatcca aatggtggg 1320
aaaatccaga ggaattcagg ccagagcggg ttttgaaga agaatcggcc acacaagccg 1380
tcgccggagg gaaagttgat ttcaggtatt tgccatttg aatgggaagg cgtagctgcc 1440
ctggtatcat actggcactg ccaatcctgg ggcttatcat tgccaaattg gttacaaatt 1500

ttgaaatcaa agctccccaa ggaacacaca agattgatgt gagtgagaaa ggagggcaat 1560
 tcagtttaca catagcaaac 1580

<210> 40
 <211> 1677
 <212> ADN
 <213> *Corchorus olitorius*

<400> 40
 atgcttgtca gaatcaacac tttgttgcag ggttattctg gaatcagatt cgaaattttg 60
 gaagcaatta ccaaatttct gaaccagaac ataactccat gtttgccact tcgtggaaca 120
 attacagctt caggggatct ggttccactt tcttacattg caggacttct taccggaaga 180
 ccaaattcca aggcggttgg acctaacggt gaatcgtaa acgccgagga agctttcaat 240
 cgtgccggga tcgaatctgg gttctttacc ttgcagccca aggaaggtct tgctcttgtt 300
 aatgggacag cagttgggtc tggaatggct tctatgggtc tttttgaagc caacatttta 360
 gctgtattgt ccgaagtttt gtctgcaatt ttcgctgaag tcatgaatgg taaacctgaa 420
 tttaccgacc atttgacaca taaattgaag catcatcctg gacaaatcga agctgctgca 480
 attatggaac atattttgga tggaagtggg tatgttaaag cggcgaagaa attacatgaa 540
 atggatccat tgcagaaacc aaagcaagat cgttatgctt taagaacttc ccacaaatgg 600
 cttgggtccac agattgaagt gatcagattt gcaacaaagt caattgaaag agaaatcaat 660
 tctgttaatg acaacccttt gatcgatgtt tctagaaaca aggctcttca tgggtgaaac 720
 ttccaaggaa ctctatttgg tgtttctatg gataatgctc gtttggtctat tgcttcaatt 780
 gggaaactca tgtttgctca attttctgaa cttgttaatg atttttacaa caatgggctc 840
 ccatcaaatc tgtcaggtgg aaggaatccc agtttgatt atggattcaa aggtgctgaa 900
 attgccatgg cttcttattg ttccgaactt caattccttg ctaatcctgt taccaatcat 960
 gtccaaagtg ctgagcaaca taatcaggat gttaattcct tgggattgat ttccgcaagg 1020
 aaaacttcag aagctgttga tattttgaag cttatgtctt ctacttactt agttgctctg 1080
 tgccaagcta ttgatttgag gcatttggaa gaaaacttga ggaacacggt aaagaacact 1140
 gtgagccaga ttgctaagaa ggttttgacc actggtgcc aatggtgaact tcacccttca 1200
 agattctgtg agaaggactt gctcaaagcc gttgaccgag aatacgtttt cgcttacatt 1260
 gatgatcctt gcagtgttac ttaccattg atgcagaaat tgagacaagt tcttgttgag 1320
 catgcattga caaatggtga gagtgagaag aatgcaagca cttcaatfff ccagaaaatt 1380
 gcagcatttg aagaggaatt gaagactttg ttgcctaagg aagttgagag tgcaagagtg 1440

gcacttgaga atgggtcaaa tgtggcagtg ccaaacagaa tcaaggaatg cagaagttac 1500
ccattgtata aatttgtgag ggaagagctt ggaactgggc ttttgactgg tgaaaaagtt 1560
aggtcacctg gtgaggaatt tgacaagggtt ttcacagcta tgtgccaggg gaaactcatt 1620
gatccaatgc ttgagtgtct caaggaatgg gacgggtgcc ctcttcccat ctgctag 1677

<210> 41

<211> 558

<212> PRT

<213> *Corchorus olitorius*

<400> 41

Met Leu Val Arg Ile Asn Thr Leu Leu Gln Gly Tyr Ser Gly Ile Arg
1 5 10 15

Phe Glu Ile Leu Glu Ala Ile Thr Lys Phe Leu Asn Gln Asn Ile Thr
20 25 30

Pro Cys Leu Pro Leu Arg Gly Thr Ile Thr Ala Ser Gly Asp Leu Val
35 40 45

Pro Leu Ser Tyr Ile Ala Gly Leu Leu Thr Gly Arg Pro Asn Ser Lys
50 55 60

Ala Val Gly Pro Asn Gly Glu Ser Leu Asn Ala Glu Glu Ala Phe Asn
65 70 75 80

Arg Ala Gly Ile Glu Ser Gly Phe Phe Thr Leu Gln Pro Lys Glu Gly
85 90 95

Leu Ala Leu Val Asn Gly Thr Ala Val Gly Ser Gly Met Ala Ser Met
100 105 110

Val Leu Phe Glu Ala Asn Ile Leu Ala Val Leu Ser Glu Val Leu Ser
115 120 125

Ala Ile Phe Ala Glu Val Met Asn Gly Lys Pro Glu Phe Thr Asp His
130 135 140

Leu Thr His Lys Leu Lys His His Pro Gly Gln Ile Glu Ala Ala Ala
145 150 155 160

Ile Met Glu His Ile Leu Asp Gly Ser Gly Tyr Val Lys Ala Ala Lys
165 170 175

Lys Leu His Glu Met Asp Pro Leu Gln Lys Pro Lys Gln Asp Arg Tyr

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

420	425	430
Lys Leu Arg Gln Val Leu Val Glu His Ala Leu Thr Asn Gly Glu Ser		
435	440	445
Glu Lys Asn Ala Ser Thr Ser Ile Phe Gln Lys Ile Ala Ala Phe Glu		
450	455	460
Glu Glu Leu Lys Thr Leu Leu Pro Lys Glu Val Glu Ser Ala Arg Val		
465	470	475 480
Ala Leu Glu Asn Gly Ser Asn Val Ala Val Pro Asn Arg Ile Lys Glu		
	485	490 495
Cys Arg Ser Tyr Pro Leu Tyr Lys Phe Val Arg Glu Glu Leu Gly Thr		
	500	505 510
Gly Leu Leu Thr Gly Glu Lys Val Arg Ser Pro Gly Glu Glu Phe Asp		
	515	520 525
Lys Val Phe Thr Ala Met Cys Gln Gly Lys Leu Ile Asp Pro Met Leu		
	530	535 540
Glu Cys Leu Lys Glu Trp Asp Gly Ala Pro Leu Pro Ile Cys		
545	550	555

<210> 42
 <211> 1743
 <212> ADN
 <213> Corchorus olitorius

<400> 42
 atgctggagt tttcggccaa ggcactgagt cgtgcccata cgttgccgca cacggcgacg 60
 agggcggtcta tgcttgtag gatcaacact cttttgcaag gatattctgg cattagattt 120
 gagatcctag aagctataac caagttcctc aatgtgaaca ttaccccatg tgtgccactt 180
 aggggatcaa ttactgcatc tggtgacttg gtgccattgt cctacatagc cgggcttttg 240
 acaggcaggc ctaattccaa ggctcttggg cctaacggtg aggccatgaa cccacagaa 300
 gcttttagcc gggccggaat ccatggtggg tttttcgagt tgcagccaa agagggtctt 360
 gcttttagtca atggtacagc cgttgggtca ggcttagcct ctttggctct ttatgaggcc 420
 aatgttttag cagtcctctc agaagtttta tcagcaattt ttgctgaagt tatgcaaggc 480
 aaacctgaat tcatgacca tttaactcac aagttgaagc atcatccagg acaaattgaa 540
 gctgctgccca ttatggaaca tattttgat ggtagctctt acattaaagc agcacaaaaa 600

```

ttgcatgaaa tggatcctct tcagaaacca aaacaggaca gatatgctct tagaacttcc      660
cctcaatggc ttggccctca aattgaagta atcaggtcag caacccaaat gattgaaagg      720
gaaatcaatt cagtgaatga taatccttta attgatgttt caagagacaa ggctttacat      780
ggaggggaatt tccaaggtac cccaattggt gtttcaatgg acaacactcg tttagccatt      840
gctgcaattg gtaaaactcat gtttgctcaa ttttctgagc ttgttaatga ttattacaac      900
aatggtttgc cttcaaactct gtctgctage cgcaaccgca gtttggatta tggtttcaaa      960
ggagctgaaa ttgcaatggc ttcttattgc tcagagcttc aattccttgg taatcctgtc    1020
actaaccatg tccaaagtgc tgagcaacat aaccaagatg tcaactcatt agggttgatc    1080
tcagcaagaa aaacagctga agctattgat atattgaagc tgatgtcttc aactttcttg    1140
attgctttat gccaaagcaat tgacttgagg catttggaag agaatttgaa gaacactgtc    1200
aagaacacag ttagtcaaat tgccaagagg gtcttaacca tgggatcaaa cggtgaactt    1260
catccatcaa gattctgtga aaaagatctt ctcagagttg tcgatcgcgga acatctctat    1320
gcctatattg atgatccttg cagtgcaagt taccatttaa tgcagaagtt gagacaagta    1380
ctggtagacc atgccttgat gaatggtgac aatgagaaga actcaaccac ctccattttc    1440
cagaagattg gtgcctttga agaagaattg aaaacccttt tgcctaaaga agttgagagt    1500
gctagaatcg aattcgagaa tggaaatgca gcaattccta acagaatcaa agaatgcagg    1560
tcctatccat tgtacaagtt tgtgagggaa gttcttgggg ctagcttggt aactggtgaa    1620
aaggatgatct ctccagggga agaattgtgac aaggttttct cagcaatttg tgcaggaaaag    1680
ttgattgatc cattgtttca atgcttgaag gagtggaatg gtgctccttt gcctatatgc    1740
taa                                                                    1743

```

<210> 43

<211> 580

<212> PRT

<213> *Corchorus olitorius*

<400> 43

```

Met Leu Glu Phe Ser Ala Lys Ala Leu Ser Arg Ala His Thr Leu Pro
1           5           10           15

```

```

His Thr Ala Thr Arg Ala Ala Met Leu Val Arg Ile Asn Thr Leu Leu
          20           25           30

```

```

Gln Gly Tyr Ser Gly Ile Arg Phe Glu Ile Leu Glu Ala Ile Thr Lys
          35           40           45

```

Phe Leu Asn Val Asn Ile Thr Pro Cys Val Pro Leu Arg Gly Ser Ile
50 55 60

Thr Ala Ser Gly Asp Leu Val Pro Leu Ser Tyr Ile Ala Gly Leu Leu
65 70 75 80

Thr Gly Arg Pro Asn Ser Lys Ala Leu Gly Pro Asn Gly Glu Ala Met
85 90 95

Asn Pro Thr Glu Ala Phe Ser Arg Ala Gly Ile His Gly Gly Phe Phe
100 105 110

Glu Leu Gln Pro Lys Glu Gly Leu Ala Leu Val Asn Gly Thr Ala Val
115 120 125

Gly Ser Gly Leu Ala Ser Leu Val Leu Tyr Glu Ala Asn Val Leu Ala
130 135 140

Val Leu Ser Glu Val Leu Ser Ala Ile Phe Ala Glu Val Met Gln Gly
145 150 155 160

Lys Pro Glu Phe Thr Asp His Leu Thr His Lys Leu Lys His His Pro
165 170 175

Gly Gln Ile Glu Ala Ala Ala Ile Met Glu His Ile Leu Asp Gly Ser
180 185 190

Ser Tyr Ile Lys Ala Ala Gln Lys Leu His Glu Met Asp Pro Leu Gln
195 200 205

Lys Pro Lys Gln Asp Arg Tyr Ala Leu Arg Thr Ser Pro Gln Trp Leu
210 215 220

Gly Pro Gln Ile Glu Val Ile Arg Ser Ala Thr Lys Met Ile Glu Arg
225 230 235 240

Glu Ile Asn Ser Val Asn Asp Asn Pro Leu Ile Asp Val Ser Arg Asp
245 250 255

Lys Ala Leu His Gly Gly Asn Phe Gln Gly Thr Pro Ile Gly Val Ser
260 265 270

Met Asp Asn Thr Arg Leu Ala Ile Ala Ala Ile Gly Lys Leu Met Phe
275 280 285

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

Arg Glu Val Leu Gly Thr Ser Leu Leu Thr Gly Glu Lys Val Ile Ser
530 535 540

Pro Gly Glu Glu Cys Asp Lys Val Phe Ser Ala Ile Cys Ala Gly Lys
545 550 555 560

Leu Ile Asp Pro Leu Phe Gln Cys Leu Lys Glu Trp Asn Gly Ala Pro
565 570 575

Leu Pro Ile Cys
580

<210> 44

<211> 3455

<212> ADN

<213> Corchorus olitorius

<400> 44

tgaaatttca aaacatcata ttttactagc gaattaaaat ttagatatag aggaacatta	60
aatggtaggc taaatatacc catttaatat tgcttgatca attaaaagat caaaatatta	120
atgtcgacga gattatTTTT cagtccaata ttctatctat aaaaaagaaa aggaaaataa	180
aaaaacaact ttgacatta aatctttctt gatattcata tcttcttcag cagcttaaca	240
caaaatatca tgatgagagg caacctaaaga ctgaatccga agtgattgtt cctcttgttg	300
tgctggcaga ggaagatgac ctttctcttg caagctcttc actgtttcat ataggcattg	360
cttcactgtg gtgaattcca accccaagtc cctcagcttt tggtttgtga acttgtaggg	420
ctttgctctt gggttcttct catccgagca cctgcatgga tttttaattg taattatata	480
tcataattag ttatgaaaaa aggattgaga attcttattt aatactaata ttttctatta	540
gcgttcgagt ctattagaaa attcattctt gataattaa actacgagtc cagtactcta	600
caatcagacc tgaccggtat gtatgagttt taatcatagc ttgttcacct aacttcatt	660
tctatcattc aaagacaaaa agcctaaagc tatcttagag agtgtggaat tttcaaagac	720
aattctcaaa ggagaaatgg gaggacccat tttggccatt gttgcagccc cccatagttc	780
atagaaaatt ccaagaaaag aggccttttc ttcaccata tttcccttaa ttgggaaatc	840
tccatgattg aactactttg agtaggtgtg gatagaatag aatagaatat ctttttcatt	900
tttatcctaa gtctcttgaa atgtttgttt tgaaagtgac attagggttag gtgaagttct	960
ctctgcacca ctactattta tactaactca gtatatagta taccatttac atttaatttt	1020
caacttgaaa aagaatagaa agaaaattca gtccaaaaaa cctacacttc ttcttgtagt	1080
aggtaagatt atgtcatttg tcaaaactgg tgtactcact gcccttcccc ttcccactc	1140

tgaattaaga aaaacaaatg aaaatttgot gtccatgctt atccacttaa ttatcttcca	1200
ataatctttt caatggcttc tacatcttat tcttatcttg tttttatttt ccttccaaat	1260
tttcttcttc tccttttata taggaagtgg tgatgtcatg agtgatgaga aaattcttta	1320
ataaaattca atgtcataat ctttaatat attttctttt tcataatggg tttcaccttt	1380
ggatttcatt aagtttggtt tggacccatg attgccctcc actttttctt caataattat	1440
ttattatatac attatacttt ttattttcct tttccatttt ctcaagaaat tctccttggtg	1500
aacttttaac aaactttaga agaaaccata tataatctacc aagatgatga aagccaatat	1560
ttctaattcc aaatctacct ttggaagtaa aatcattttc atatatatat ttcttctaca	1620
tttttatgga ccatatatat ttaatcttga agatgagaac atcaaagttc ttctgtcac	1680
tatcttccaa caaacaagg ttgacattat atatttaacc aagttgtgga taatatctaa	1740
agccttggtt ttttattttc gaatattgtc gcacatttaa gacttagagc aagcaaatat	1800
atattgaata cgtgttgatg ttcaaattat tgacaattca aatgacccaa ttaaaacaat	1860
gaataagatt atcatctaata ctaagttttc aatttctttt ttgtttttta gatttaggtg	1920
tgccactctt actataagaa acgtgttcat ataaataatt catttcaaag agaagtaaag	1980
ttaggttaaa gaaagtgaag tgatctttcc ctctagatat ttctagatgt taaaataatg	2040
gaattcggct ccaaacttac cagaaattta gtgcaagccc ctctaaattt atttttttgt	2100
tctatttttt gtctataata atattaataa taactagaca ttgggcttaa tctttggtac	2160
atgaaccctt aaatattccc ttatattcct ctaatgatat tccatgccat gtaaacttaa	2220
aagtggaacc taataaatcc tcagatatgt ctatttatta tgttttttgt gactataatg	2280
acaacatatac gaacctgtta atgtttgtta gtatgatatt tttctcccct gtttgcaa	2340
tggaatagaaa aaattgcaag aagcaagttt acattattcc tctgcaaaca aggggtgtcc	2400
cacttaccta ctacagcact agaaaatcac taataaaatt tgattactat actaaaatta	2460
attcttaatt agcaattttt tgtaatttaa gtatacttaa aaaggaaatt agagaaatta	2520
cttacttggt agggatggga tactcaggga agaacttggc aagaatgtcc accacctctc	2580
cacggtggag gacgctctcg gcgcagaggt aacggccgga ggaggaggga ttctcgaaga	2640
caagaatgtg tgctaataca acatctctaa catgaacata ggcttgaact gaattggcat	2700
aggtctttgc agagccagtt aagtacttaa gaatgtgaat aatgctagca ttacagttg	2760
attgtagcaa tggaccaagc accaaaactg gagttattgt gactaagtca accccttttt	2820
ccttggtgtt ttcccaggct gcctgctctg ccacagtctt ccataacaa taccaattct	2880
gcaatttttg ttttaaacat gaaaatttgt taattaagga aaggaaatgt ttaattttgt	2940

gctaattata gaatgaaaat ggtggttaac ttaaateccct ttgttgtttc tattttgggt 3000
 aggggagatt attaaattgg taataactaac aatagattgt agccataagt ttacacattc 3060
 attgcattaa ttagccaaaa catgcatggt gttgggattt ttattactaa aacaaaaaag 3120
 ggaaagacca atttgacat ttaaaacaag catttgatca ggtgatagac caagccagca 3180
 aaaccagact ttttgtctgg ttccctgtcc ttccttaatt tcccagaaac agaacaggaa 3240
 atttcctggt caaacacaag gtcagttaat tttgactttt caatactctc actaatcatt 3300
 tatgttatga cccaccacta tttttaacat tggcctgttt attactatgt ttttaattaa 3360
 atacgcgtga cttggattct taatttcaat taattatttg tgattagtgg attcaattcc 3420
 tgataatctc aagtatttac ttaattaaga aggtg 3455

<210> 45

<211> 969

<212> ADN

<213> *Corchorus olitorius*

<400> 45

atgtcgaagc aaggagaggc cgtgtgcgtt accggagcca gcggcgccat tggctcttgg 60
 ctctcgaagc tctcctggc tcgcggttac actgtccacg gaaccgtcag aaatatcaag 120
 gatgagaaag aaacgaagca tctagaatct ttggaaggag cagaatctcg tctccgtctc 180
 ttccagatcg atctcctcga ttatgatgca atctccgcgc ctatcgaagg ctgcgctggc 240
 gtcttccacc tcgcctcccc ctgcaccgtc gatcaagttc acgaccctca gaaggagctt 300
 ttggatcctg caattaaagg aacactcaat gtactgacag ctgccaagga gctcgggtgtt 360
 aagcgcgtgg ttgtcacgtc ttctgtctcc tccataaac ctagtccgaa ctggcctgca 420
 gataagatta aaacagagga ttgctggact gatattgact actgcaagca aaatgagtta 480
 tggatatcaa tttccaaaac actggctgag aaggcagcat gggaattttc caaggagaaa 540
 ggtttggatg tgggtggtgt gaaccctggc actgtgatgg gtccaaatat tctccaacc 600
 cttactgcta gcatgtggat gttgttgccg ctattgcaag gctgcacaga gacatatcaa 660
 gactttttca tgggatctgt ccatttcaaa gatgttgcat tagcacatat tttggtgtat 720
 gagaacccat cagcaagtgg aaggcacatg tgtcttgagg ccataatctca ttatggtgat 780
 tttgtcgcca aagttgctga actttaccct gaatataatg tgcccagttt gccgagggac 840
 actcaacctg gactattaag ggctaagaat ggaggtcaaa agctgatgga tttgggggtg 900
 gaattcattc ccatggaaca gataatcaaa gatgctgtcg agagtttaaa aagcaagggc 960
 cttatttaa 969

<210> 46

<211> 322

<212> PRT

<213> Corchorus olitorius

<400> 46

Met Ser Lys Gln Gly Glu Ala Val Cys Val Thr Gly Ala Ser Gly Ala
 1 5 10 15

Ile Gly Ser Trp Leu Val Lys Leu Leu Leu Ala Arg Gly Tyr Thr Val
 20 25 30

His Gly Thr Val Arg Asn Ile Lys Asp Glu Lys Glu Thr Lys His Leu
 35 40 45

Glu Ser Leu Glu Gly Ala Glu Ser Arg Leu Arg Leu Phe Gln Ile Asp
 50 55 60

Leu Leu Asp Tyr Asp Ala Ile Ser Ala Ala Ile Glu Gly Cys Ala Gly
 65 70 75 80

Val Phe His Leu Ala Ser Pro Cys Thr Val Asp Gln Val His Asp Pro
 85 90 95

Gln Lys Glu Leu Leu Asp Pro Ala Ile Lys Gly Thr Leu Asn Val Leu
 100 105 110

Thr Ala Ala Lys Glu Leu Gly Val Lys Arg Val Val Val Thr Ser Ser
 115 120 125

Val Ser Ser Ile Thr Pro Ser Pro Asn Trp Pro Ala Asp Lys Ile Lys
 130 135 140

Thr Glu Asp Cys Trp Thr Asp Ile Asp Tyr Cys Lys Gln Asn Glu Leu
 145 150 155 160

Trp Tyr Pro Ile Ser Lys Thr Leu Ala Glu Lys Ala Ala Trp Glu Phe
 165 170 175

Ser Lys Glu Lys Gly Leu Asp Val Val Val Val Asn Pro Gly Thr Val
 180 185 190

Met Gly Pro Asn Ile Pro Pro Thr Leu Thr Ala Ser Met Trp Met Leu
 195 200 205

Leu Arg Leu Leu Gln Gly Cys Thr Glu Thr Tyr Gln Asp Phe Phe Met
 210 215 220

Gly Ser Val His Phe Lys Asp Val Ala Leu Ala His Ile Leu Val Tyr
225 230 235 240

Glu Asn Pro Ser Ala Ser Gly Arg His Met Cys Leu Glu Ala Ile Ser
245 250 255

His Tyr Gly Asp Phe Val Ala Lys Val Ala Glu Leu Tyr Pro Glu Tyr
260 265 270

Asn Val Pro Ser Leu Pro Arg Asp Thr Gln Pro Gly Leu Leu Arg Ala
275 280 285

Lys Asn Gly Gly Gln Lys Leu Met Asp Leu Gly Leu Glu Phe Ile Pro
290 295 300

Met Glu Gln Ile Ile Lys Asp Ala Val Glu Ser Leu Lys Ser Lys Gly
305 310 315 320

Leu Ile

<210> 47

<211> 787

<212> ADN

<213> *Corchorus olitorius*

<400> 47

tgcgatcgtg ccaaacccaa actggaatcc tcaaaccaac ggggccttcg atgagacatc	60
atggactgac ttggagtact gtaagagtcg acagaaatgg tatccagttt cgaagacgat	120
ggcagagaaa acagcatggg aatttgcaga gaagcatggg atggatgttg tggctataaa	180
cccagccacg tgtttaggcc ctctcttgca acccaacttg aatgccagct gcgctgtgtt	240
gttgcaattg ttggaaggat ccaaagatac ccaggaatac cattggttgg gagctgtgca	300
tgttaaagat gtcgccaagg ctccagatctt gttgtttgag tctccttctg cttctggtag	360
atatctttgc accaatggca tttatcagtt tggaactttc gctgaaactg tctccacct	420
cttccctcag tatectgtcc acaggtttac tggagatata caacctggct tagtttcttg	480
caaagatgca gcaaagcgat taattgagct aggtctaata ttcacccag ttgaagaagc	540
cgtccgagag accgtggaga gtctgcaagc caaaggcttc ttgaagcagc agcaaccatc	600
agagtcttag aaattaagtc tcttcttttt ttatcctttc tgtctctcaa aacagttaga	660
agctttatag ttatactag ttttttaggt ttatttagct gtaggtaaag catggattca	720

atcagaagtt ttgtgcctaa agcttctatt attgtttcat tgtgataagt tataacaagt 780
ataaata 787

<210> 48
<211> 202
<212> PRT
<213> Corchorus olitorius

<400> 48

Ala Ile Val Pro Asn Pro Asn Trp Asn Pro Gln Thr Asn Gly Ala Phe
1 5 10 15

Asp Glu Thr Ser Trp Thr Asp Leu Glu Tyr Cys Lys Ser Arg Gln Lys
20 25 30

Trp Tyr Pro Val Ser Lys Thr Met Ala Glu Lys Thr Ala Trp Glu Phe
35 40 45

Ala Glu Lys His Gly Met Asp Val Val Ala Ile Asn Pro Ala Thr Cys
50 55 60

Leu Gly Pro Leu Leu Gln Pro Asn Leu Asn Ala Ser Cys Ala Val Leu
65 70 75 80

Leu Gln Leu Leu Glu Gly Ser Lys Asp Thr Gln Glu Tyr His Trp Leu
85 90 95

Gly Ala Val His Val Lys Asp Val Ala Lys Ala Gln Ile Leu Leu Phe
100 105 110

Glu Ser Pro Ser Ala Ser Gly Arg Tyr Leu Cys Thr Asn Gly Ile Tyr
115 120 125

Gln Phe Gly Thr Phe Ala Glu Thr Val Ser His Leu Phe Pro Gln Tyr
130 135 140

Pro Val His Arg Phe Thr Gly Asp Thr Gln Pro Gly Leu Val Ser Cys
145 150 155 160

Lys Asp Ala Ala Lys Arg Leu Ile Glu Leu Gly Leu Ile Phe Thr Pro
165 170 175

Val Glu Glu Ala Val Arg Glu Thr Val Glu Ser Leu Gln Ala Lys Gly
180 185 190

Phe Leu Lys Gln Gln Gln Pro Ser Glu Ser

195

200

<210> 49
 <211> 1161
 <212> ADN
 <213> *Corchorus olitorius*

<400> 49
 atggctcttc ttttcatcgt ccctttctta ctcttccttg gcctgggtctc tcgacttcgg 60
 cgaaaaccat tccctccggg gcccaaaggg ttaccaatca tcggaaacat gatgatgatg 120
 gaccaattaa cccaccgtgg tctggccaag ctggccaaca aatacggcgg catattccat 180
 atgaagatgg ggtattttaca tatgggtcgcc gtttcaaadc ccgacatggc tcgccaagtt 240
 ttgcaagttc aagacaacat tttctcaaac agaccagcca ccatagccat aagttattta 300
 acttatgaca gagctgatat ggcttttgcc cattacggac cgttttggag acagatgcgt 360
 aagctttgtg taatgaagct tttcagtcgg aaaagagccg aatcatggga gtctgtaaga 420
 gatgaagttg attccatggg tcgaaccgtg tcggccaaca ctgggaaatc cattaatggt 480
 ggtgagttga tttttaacct tacgaaaaac attattttaca gggcggtttt tggttccagt 540
 tcacaagaag gacaagacga gttcatcgga atcttacaag agttttcgaa actgttcggg 600
 gcttttaata ttgccgattt tatccctctg ctgacgtggg ttgatectca aggacttaac 660
 aacaggctta agaacgctcg tcaagcttta gacaagttca tcgataccat tattgatgaa 720
 catattcaga agaggaacaa caagaacaat gtttctgatg atgttgatac cgatatggtc 780
 gatgatttgc ttgcttttta cagtgaagaa gctaaagtaa atgaatcgga ggatcttcaa 840
 aacgccatca gactcactag agaaaatata aaagccatta tcatggacgt tatgtttgga 900
 gggacagaga cgggtggcatc agcaatagaa tgggccttat ctgagctgat gagaagccca 960
 gaagatatga agagagtcca acaagagttg gctgacgtgg tgggtcttga ccggaaagtc 1020
 gaagaatccg atttcgacaa actaaccttc cttaaagtca ctttgaaaga aaccctccgc 1080
 ctccaccac caatcccact tctcctccat gagaccgccg aggacgccga ggtcggcgga 1140
 tacagaatcc ccgccaagtc g 1161

<210> 50
 <211> 387
 <212> PRT
 <213> *Corchorus olitorius*

<400> 50
 Met Ala Leu Leu Phe Ile Val Pro Phe Leu Leu Leu Gly Leu Val
 1 5 10 15

Ser Arg Leu Arg Arg Lys Pro Phe Pro Pro Gly Pro Lys Gly Leu Pro
 20 25 30

Ile Ile Gly Asn Met Met Met Met Asp Gln Leu Thr His Arg Gly Leu
 35 40 45

Ala Lys Leu Ala Asn Lys Tyr Gly Gly Ile Phe His Met Lys Met Gly
 50 55 60

Tyr Leu His Met Val Ala Val Ser Asn Pro Asp Met Ala Arg Gln Val
 65 70 75 80

Leu Gln Val Gln Asp Asn Ile Phe Ser Asn Arg Pro Ala Thr Ile Ala
 85 90 95

Ile Ser Tyr Leu Thr Tyr Asp Arg Ala Asp Met Ala Phe Ala His Tyr
 100 105 110

Gly Pro Phe Trp Arg Gln Met Arg Lys Leu Cys Val Met Lys Leu Phe
 115 120 125

Ser Arg Lys Arg Ala Glu Ser Trp Glu Ser Val Arg Asp Glu Val Asp
 130 135 140

Ser Met Val Arg Thr Val Ser Ala Asn Thr Gly Lys Ser Ile Asn Val
 145 150 155 160

Gly Glu Leu Ile Phe Asn Leu Thr Lys Asn Ile Ile Tyr Arg Ala Ala
 165 170 175

Phe Gly Ser Ser Ser Gln Glu Gly Gln Asp Glu Phe Ile Gly Ile Leu
 180 185 190

Gln Glu Phe Ser Lys Leu Phe Gly Ala Phe Asn Ile Ala Asp Phe Ile
 195 200 205

Pro Trp Leu Thr Trp Val Asp Pro Gln Gly Leu Asn Asn Arg Leu Lys
 210 215 220

Asn Ala Arg Gln Ala Leu Asp Lys Phe Ile Asp Thr Ile Ile Asp Glu
 225 230 235 240

His Ile Gln Lys Arg Asn Asn Lys Asn Asn Val Ser Asp Asp Val Asp
 245 250 255

Thr Asp Met Val Asp Asp Leu Leu Ala Phe Tyr Ser Glu Glu Ala Lys
 260 265 270

Val Asn Glu Ser Glu Asp Leu Gln Asn Ala Ile Arg Leu Thr Arg Glu
 275 280 285

Asn Ile Lys Ala Ile Ile Met Asp Val Met Phe Gly Gly Thr Glu Thr
 290 295 300

Val Ala Ser Ala Ile Glu Trp Ala Leu Ser Glu Leu Met Arg Ser Pro
 305 310 315 320

Glu Asp Met Lys Arg Val Gln Gln Glu Leu Ala Asp Val Val Gly Leu
 325 330 335

Asp Arg Lys Val Glu Glu Ser Asp Phe Asp Lys Leu Thr Phe Leu Lys
 340 345 350

Cys Thr Leu Lys Glu Thr Leu Arg Leu His Pro Pro Ile Pro Leu Leu
 355 360 365

Leu His Glu Thr Ala Glu Asp Ala Glu Val Gly Gly Tyr Arg Ile Pro
 370 375 380

Ala Lys Ser
 385

<210> 51

<211> 1098

<212> ADN

<213> Corchorus olitorius

<400> 51

atgggttcaa ctggtgaaac ccaattcact ccaactcaag tctccgatga ggaagcaaac	60
ttgttgcgca tgcaattggc tagtgcctca gttcttccca tggctctcaa atctgccata	120
gaacttgacc tacttgaagt catggccaag gctggacctg gtgctttctt gtccccaaca	180
gaagtagctt cccaattgcc caccaagaac cctgatgcac ccgtcatgct cgaccgtatc	240
ttggggtccc ttgctagtta ctccatttta acttgctcct taaggaatct tctgatggc	300
aaagttgaga ggctctatgg ccttggccct gtctgtaaat acctgggtcaa gaatgaagat	360
gggtgctgctc tttccgccct taatctcatg aatcaagaca aggtcctaata ggagagctgg	420
tactacttga aagatgcagt gttggaaggt ggaattccat tcaacaaggc ctatggcatg	480
accgcgttcg agtaccatgg cactgaccct agattcaaca aggttttcaa caggggaatg	540

tctgatcact caactatcac catgaagaag attctcgaga cctacgatgg attcgagggg 600
 ctcaaaacat tggttgacgt tgggtggtggt gttggtgccg cgcttaacat gatcgtctcc 660
 aagcaccctt ccattaaggg cattaacttt gatttgcctc atgtcattga ggatgctcca 720
 gctcttcctg gtgttgagca tgttggtgga gatatgtttg taagtgttcc aaaaggagat 780
 gccattttca tgaagtggat atgtcatgat tggagcgatg aacactgcgt aaaattcttg 840
 aagaagtgct atgaagcttt gccagacaat gggaaagtca tcgttgccga atgcattctt 900
 cctgattacc cagatgctag ccttgccaca aagctagtgtg ttcatatcga ttgtatcatg 960
 ttggctcaca accctggtgg gaaagaaagg acagagaagg aatttgaagc cttggcaaag 1020
 ggggcaggtt ttcaaggttt ccaagtaaag tgttggtgctt ttggcactta catcatggag 1080
 ttctcaaaa ctgtttaa 1098

<210> 52

<211> 365

<212> PRT

<213> *Corchorus olitorius*

<400> 52

Met Gly Ser Thr Gly Glu Thr Gln Phe Thr Pro Thr Gln Val Ser Asp
 1 5 10 15

Glu Glu Ala Asn Leu Phe Ala Met Gln Leu Ala Ser Ala Ser Val Leu
 20 25 30

Pro Met Val Leu Lys Ser Ala Ile Glu Leu Asp Leu Leu Glu Val Met
 35 40 45

Ala Lys Ala Gly Pro Gly Ala Phe Leu Ser Pro Thr Glu Val Ala Ser
 50 55 60

Gln Leu Pro Thr Lys Asn Pro Asp Ala Pro Val Met Leu Asp Arg Ile
 65 70 75 80

Leu Arg Leu Leu Ala Ser Tyr Ser Ile Leu Thr Cys Ser Leu Arg Asn
 85 90 95

Leu Pro Asp Gly Lys Val Glu Arg Leu Tyr Gly Leu Gly Pro Val Cys
 100 105 110

Lys Tyr Leu Val Lys Asn Glu Asp Gly Val Ala Leu Ser Ala Leu Asn
 115 120 125

Leu Met Asn Gln Asp Lys Val Leu Met Glu Ser Trp Tyr Tyr Leu Lys

130		135		140											
Asp	Ala	Val	Leu	Glu	Gly	Gly	Ile	Pro	Phe	Asn	Lys	Ala	Tyr	Gly	Met
145					150					155					160
Thr	Ala	Phe	Glu	Tyr	His	Gly	Thr	Asp	Pro	Arg	Phe	Asn	Lys	Val	Phe
				165					170					175	
Asn	Arg	Gly	Met	Ser	Asp	His	Ser	Thr	Ile	Thr	Met	Lys	Lys	Ile	Leu
			180					185					190		
Glu	Thr	Tyr	Asp	Gly	Phe	Glu	Gly	Leu	Lys	Thr	Leu	Val	Asp	Val	Gly
		195					200					205			
Gly	Gly	Val	Gly	Ala	Thr	Leu	Asn	Met	Ile	Val	Ser	Lys	His	Pro	Ser
	210					215					220				
Ile	Lys	Gly	Ile	Asn	Phe	Asp	Leu	Pro	His	Val	Ile	Glu	Asp	Ala	Pro
225					230					235					240
Ala	Leu	Pro	Gly	Val	Glu	His	Val	Gly	Gly	Asp	Met	Phe	Val	Ser	Val
				245					250					255	
Pro	Lys	Gly	Asp	Ala	Ile	Phe	Met	Lys	Trp	Ile	Cys	His	Asp	Trp	Ser
			260					265					270		
Asp	Glu	His	Cys	Val	Lys	Phe	Leu	Lys	Lys	Cys	Tyr	Glu	Ala	Leu	Pro
		275					280					285			
Asp	Asn	Gly	Lys	Val	Ile	Val	Ala	Glu	Cys	Ile	Leu	Pro	Asp	Tyr	Pro
	290					295					300				
Asp	Ala	Ser	Leu	Ala	Thr	Lys	Leu	Val	Val	His	Ile	Asp	Cys	Ile	Met
305					310					315					320
Leu	Ala	His	Asn	Pro	Gly	Gly	Lys	Glu	Arg	Thr	Glu	Lys	Glu	Phe	Glu
				325					330					335	
Ala	Leu	Ala	Lys	Gly	Ala	Gly	Phe	Gln	Gly	Phe	Gln	Val	Lys	Cys	Cys
			340					345					350		
Ala	Phe	Gly	Thr	Tyr	Ile	Met	Glu	Phe	Leu	Lys	Thr	Val			
	355						360					365			

<210> 53

<211> 20
 <212> ADN
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Synthetic
 primer

 <400> 53
 gaatggggcg tttccatcta 20

<210> 54
 <211> 20
 <212> ADN
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Synthetic
 primer

 <400> 54
 ttggcaacac attgaaacca 20

<210> 55
 <211> 20
 <212> ADN
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Synthetic
 primer

 <400> 55
 aagagctcag ggagcttact 20

<210> 56
 <211> 20
 <212> ADN
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Synthetic
 primer

 <400> 56
 tctgccgtcg gattaagtga 20

<210> 57
 <211> 20
 <212> ADN
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: Synthetic
 primer

<400> 57
 cttccaacgc caggatcata 20

<210> 58
 <211> 20
 <212> ADN
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic primer

<400> 58
 ctatgcttat tgcccaccct 20

<210> 59
 <211> 21
 <212> ADN
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic primer

<400> 59
 aataccattg gttgggagct g 21

<210> 60
 <211> 25
 <212> ADN
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic primer

<400> 60
 ggattcaatc agaagttttg tgcct 25

<210> 61
 <211> 20
 <212> ADN
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic primer

<400> 61
 ccagccacca tagccataag 20

<210> 62
 <211> 18
 <212> ADN
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
primer

<400> 62

acagaatccc cgccaagt

18

<210> 63

<211> 362

<212> PRT

<213> Populus trichocarpa

<400> 63

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

Arg Tyr Phe Gly Leu Asp Lys Pro Gly Met His Val Gly Ile Val Gly

180 185 190
 Leu Gly Gly Leu Gly His Val Ala Val Lys Phe Ala Arg Ala Met Gly
 195 200 205
 Val Lys Val Thr Val Ile Ser Thr Ser Pro Asn Lys Lys Gln Glu Ala
 210 215 220
 Leu Glu Asn Leu Gly Ala Asp Ser Phe Leu Val Ser Arg Asp Gln Asp
 225 230 235 240
 Gln Met Gln Ala Ala Met Gly Thr Leu Asp Gly Ile Ile Asp Thr Val
 245 250 255
 Ser Ala Val His Pro Leu Leu Pro Leu Val Ala Leu Leu Lys Ser His
 260 265 270
 Gly Lys Leu Val Leu Val Gly Ala Pro Glu Lys Pro Leu Glu Leu Pro
 275 280 285
 Val Phe Pro Leu Ile Thr Gly Arg Lys Thr Val Gly Gly Ser Cys Val
 290 295 300
 Gly Gly Ile Lys Glu Thr Gln Glu Met Ile Asp Phe Ala Ala Lys His
 305 310 315 320
 Asn Ile Thr Ala Asp Ile Glu Val Ile Pro Met Asp Tyr Val Asn Thr
 325 330 335
 Ala Met Glu Arg Val Leu Lys Ala Asp Val Arg Tyr Arg Phe Val Ile
 340 345 350
 Asp Val Gly Lys Thr Leu Lys Pro Asp Val
 355 360
 <210> 64
 <211> 362
 <212> PRT
 <213> Ricinus communis
 <400> 64
 Met Val Ala Asn Leu Pro Glu Lys Asp His Pro Arg Lys Ala Phe Gly
 1 5 10 15
 Trp Ala Ala Arg Asp Gln Ser Gly Val Leu Ser Pro Phe Thr Phe Ser
 20 25 30

Arg Arg Glu Thr Gly Glu Lys Asp Val Ser Phe Lys Val Leu Tyr Cys
 35 40 45

Gly Met Cys His Ser Asp Leu His Met Val Lys Asn Glu Trp Gly Thr
 50 55 60

Ser Thr Tyr Pro Leu Val Pro Gly His Glu Ile Val Gly Val Val Thr
 65 70 75 80

Glu Val Gly Ser Lys Val Glu Lys Ile Lys Val Gly Asp Lys Val Gly
 85 90 95

Val Gly Cys Met Val Gly Ser Cys Arg Ser Cys Asn Asn Cys Asn Lys
 100 105 110

Asp Leu Glu Asn Tyr Cys Pro Lys Met Ile Leu Thr Tyr Gly Ala Lys
 115 120 125

Tyr Tyr Asp Gly Thr Thr Thr Tyr Gly Gly Tyr Ser Asp Ile Met Val
 130 135 140

Ser Asp Glu His Phe Val Val Arg Ile Pro Asp Asn Leu Pro Leu Asp
 145 150 155 160

Ala Thr Ala Pro Leu Leu Cys Ala Gly Ile Thr Val Tyr Ser Pro Leu
 165 170 175

Lys Tyr Tyr Gly Leu Asp Lys Pro Gly Met Gln Leu Gly Val Val Gly
 180 185 190

Leu Gly Gly Leu Gly His Met Ala Val Lys Phe Ala Lys Ala Met Gly
 195 200 205

Ala Lys Val Thr Val Ile Ser Thr Ser Pro Asn Lys Lys Gln Glu Ala
 210 215 220

Ile Glu Arg Leu Gly Ala Asp Ser Phe Leu Val Ser Arg Asp Gln Asp
 225 230 235 240

Gln Met Lys Gly Ala Ile Gly Thr Met Asp Gly Ile Ile Asp Thr Val
 245 250 255

Ser Ala Met His Pro Leu Ser Pro Leu Ile Gly Leu Leu Lys Ser Asp
 260 265 270

Gly Lys Leu Val Leu Val Gly Ala Pro Glu Lys Pro Leu Glu Leu Pro
 275 280 285

Ala Phe Pro Leu Ile Gly Gly Arg Lys Leu Val Gly Gly Ser Cys Ile
 290 295 300

Gly Gly Met Lys Glu Thr Gln Glu Met Ile Asp Phe Ala Ala Lys His
 305 310 315 320

Ser Ile Thr Ala Asp Ile Glu Val Ile Pro Ala Asn Tyr Val Asn Thr
 325 330 335

Ala Met Glu Arg Met Leu Lys Ala Asp Val Arg Tyr Arg Phe Val Ile
 340 345 350

Asp Ile Gly Asn Thr Leu Lys Pro Gly His
 355 360

<210> 65

<211> 359

<212> PRT

<213> Fragaria x ananassa

<400> 65

Met Ser Ile Glu Gln Glu His Pro Asn Lys Ala Ser Gly Trp Ala Ala
 1 5 10 15

Arg Asp Ser Ser Gly Val Leu Ser Pro Phe Asn Phe Ser Arg Arg Glu
 20 25 30

Thr Gly Glu Lys Asp Val Met Phe Lys Val Leu Tyr Cys Gly Ile Cys
 35 40 45

His Ser Asp His His Met Val Lys Asn Glu Trp Gly Phe Ser Thr Tyr
 50 55 60

Pro Leu Val Pro Gly His Glu Ile Val Gly Glu Val Thr Glu Val Gly
 65 70 75 80

Ser Lys Val Gln Lys Phe Lys Val Gly Asp Arg Val Gly Val Gly Cys
 85 90 95

Ile Val Gly Ser Cys Arg Ser Cys Glu Asn Cys Thr Asp His Leu Glu
 100 105 110

Asn Tyr Cys Pro Lys Gln Ile Leu Thr Tyr Gly Ala Asn Tyr Tyr Asp

115						120						125					
Gly	Thr	Thr	Thr	Tyr	Gly	Gly	Cys	Ser	Asp	Ile	Met	Val	Ala	His	Glu		
130						135					140						
His	Phe	Val	Val	Arg	Ile	Pro	Asp	Asn	Leu	Pro	Leu	Asp	Gly	Ala	Ala		
145					150					155					160		
Pro	Leu	Leu	Cys	Ala	Gly	Ile	Thr	Thr	Tyr	Ser	Pro	Leu	Arg	Tyr	Phe		
				165					170					175			
Gly	Leu	Asp	Lys	Pro	Gly	Met	His	Val	Gly	Val	Val	Gly	Leu	Gly	Gly		
			180					185					190				
Leu	Gly	His	Val	Ala	Val	Lys	Phe	Ala	Lys	Ala	Met	Gly	Val	Lys	Val		
		195					200					205					
Thr	Val	Ile	Ser	Thr	Ser	Pro	Lys	Lys	Glu	Glu	Glu	Ala	Leu	Lys	His		
	210					215						220					
Leu	Gly	Ala	Asp	Ser	Phe	Leu	Val	Ser	Arg	Asp	Gln	Asp	His	Met	Gln		
225					230					235					240		
Ala	Ala	Ile	Gly	Thr	Met	Asp	Gly	Ile	Ile	Asp	Thr	Val	Ser	Ala	Gln		
				245					250					255			
His	Pro	Leu	Leu	Pro	Leu	Ile	Gly	Leu	Leu	Lys	Ser	His	Gly	Lys	Leu		
			260					265					270				
Val	Met	Val	Gly	Ala	Pro	Glu	Lys	Pro	Leu	Glu	Leu	Pro	Val	Phe	Pro		
		275					280					285					
Leu	Leu	Met	Gly	Arg	Lys	Met	Val	Ala	Gly	Ser	Gly	Ile	Gly	Gly	Met		
	290					295					300						
Met	Glu	Thr	Gln	Glu	Met	Ile	Asp	Phe	Ala	Ala	Lys	His	Asn	Ile	Thr		
305					310					315					320		
Ala	Asp	Ile	Glu	Val	Ile	Pro	Ile	Asp	Tyr	Leu	Asn	Thr	Ala	Met	Glu		
				325					330					335			
Arg	Leu	Val	Lys	Ala	Asp	Val	Arg	Tyr	Arg	Phe	Val	Ile	Asp	Ile	Gly		
			340					345					350				
Asn	Thr	Leu	Lys	Ala	Ser	Ser											

355

<210> 66

<211> 362

<212> PRT

<213> *Populus trichocarpa*

<400> 66

Met Ser Lys Ser Pro Glu Glu Glu His Pro Val Lys Ala Phe Gly Trp
 1 5 10 15

Ala Ala Arg Asp Gln Ser Gly His Leu Ser Pro Phe Asn Phe Ser Arg
 20 25 30

Arg Ala Thr Gly Glu Glu Asp Val Arg Phe Lys Val Leu Tyr Cys Gly
 35 40 45

Ile Cys His Ser Asp Leu His Ser Ile Lys Asn Asp Trp Gly Phe Ser
 50 55 60

Met Tyr Pro Leu Val Pro Gly His Glu Ile Val Gly Glu Val Thr Glu
 65 70 75 80

Val Gly Ser Lys Val Lys Lys Val Asn Val Gly Asp Lys Val Gly Val
 85 90 95

Gly Cys Leu Val Gly Ala Cys His Ser Cys Glu Ser Cys Ala Ser Asp
 100 105 110

Leu Glu Asn Tyr Cys Pro Lys Met Ile Leu Thr Tyr Ala Ser Ile Tyr
 115 120 125

His Asp Gly Thr Ile Thr Tyr Gly Gly Tyr Ser Asp His Met Val Ala
 130 135 140

Asn Glu Arg Tyr Ile Ile Arg Phe Pro Asp Asn Met Pro Leu Asp Gly
 145 150 155 160

Gly Ala Pro Leu Leu Cys Ala Gly Ile Thr Val Tyr Ser Pro Leu Lys
 165 170 175

Tyr Phe Gly Leu Asp Glu Pro Gly Lys His Ile Gly Ile Val Gly Leu
 180 185 190

Gly Gly Leu Gly His Val Ala Val Lys Phe Ala Lys Ala Phe Gly Ser
 195 200 205

Lys Val Thr Val Ile Ser Thr Ser Pro Ser Lys Lys Glu Glu Ala Leu
210 215 220

Lys Asn Leu Gly Ala Asp Ser Phe Leu Val Ser Arg Asp Gln Glu Gln
225 230 235 240

Met Gln Ala Ala Ala Gly Thr Leu Asp Gly Ile Ile Asp Thr Val Ser
245 250 255

Ala Val His Pro Leu Leu Pro Leu Phe Gly Leu Leu Lys Ser His Gly
260 265 270

Lys Leu Ile Leu Val Gly Ala Pro Glu Lys Pro Leu Glu Leu Pro Ala
275 280 285

Phe Ser Leu Ile Ala Gly Arg Lys Thr Val Ala Gly Ser Gly Ile Gly
290 295 300

Gly Met Lys Glu Thr Gln Glu Met Ile Asp Phe Ala Ala Lys His Asn
305 310 315 320

Ile Thr Ala Asp Ile Glu Val Ile Ser Thr Asp Tyr Leu Asn Thr Ala
325 330 335

Met Glu Arg Leu Ala Lys Asn Asp Val Arg Tyr Arg Phe Val Ile Asp
340 345 350

Val Gly Asn Thr Leu Ala Ala Thr Lys Pro
355 360

<210> 67

<211> 359

<212> PRT

<213> Gossypium hirsutum

<400> 67

Met Ala Leu Glu Thr Pro Asn His Thr Gln Thr Val Ala Gly Trp Ala
1 5 10 15

Ala Tyr Asp Ser Ser Gly Lys Ile Ala Pro Tyr Ile Phe Lys Arg Arg
20 25 30

Glu Asn Gly Val Asn Asp Val Thr Ile Gln Val Met Tyr Cys Gly Ile
35 40 45

Cys His Thr Asp Leu His His Val Lys Asp Asp Trp Gly Ile Thr Met

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

290 295 300
 Gly Met Lys Glu Thr Gln Glu Met Met Asp Val Cys Gly Lys His Asn
 305 310 315 320

 Ile Thr Cys Asp Val Glu Leu Ile Lys Pro Asp Lys Ile Asn Gln Ala
 325 330 335

 Leu Asp Arg Leu Ala Arg Asn Asp Val Arg Tyr Arg Phe Val Ile Asp
 340 345 350

 Ile Ala Gly Thr Ser Lys Leu
 355

 <210> 68
 <211> 357
 <212> PRT
 <213> Gossypium hirsutum

 <400> 68
 Met Gly Ser Leu Glu Thr Glu Arg Thr Thr Thr Gly Trp Ala Ala Arg
 1 5 10 15

 Asp Pro Ser Gly Val Leu Ser Pro Tyr Thr Tyr Thr Leu Arg Asn Thr
 20 25 30

 Gly Pro Glu Asp Val Phe Val Lys Val Met Cys Cys Gly Ile Cys His
 35 40 45

 Thr Asp Leu His Gln Ala Lys Asn Asp Leu Gly Met Ser Asn Tyr Pro
 50 55 60

 Met Val Pro Gly His Glu Val Val Gly Glu Val Leu Glu Val Gly Ser
 65 70 75 80

 Asp Val Ser Lys Phe Arg Val Gly Asp Ile Val Gly Val Gly Cys Leu
 85 90 95

 Val Gly Cys Cys Arg Asn Cys Arg Pro Cys Asp Ser Asp Asn Glu Gln
 100 105 110

 Tyr Cys Leu Lys Lys Ile Trp Ser Tyr Asn Asp Val Tyr Thr Asp Gly
 115 120 125

 Lys Pro Thr Gln Gly Gly Phe Ala Gly Ser Met Val Val Asp Gln Lys
 130 135 140

Phe Val Val Lys Ile Pro Glu Gly Met Ala Pro Glu Gln Val Ala Pro
145 150 155 160

Leu Leu Cys Ala Gly Val Thr Val Tyr Ser Pro Leu Asn His Phe Gly
165 170 175

Leu Met Gly Ser Gly Leu Arg Gly Gly Ile Leu Gly Leu Gly Gly Val
180 185 190

Gly His Met Gly Val Lys Ile Ala Lys Ala Met Gly His His Val Thr
195 200 205

Val Ile Ser Ser Ser Asp Lys Lys Lys Val Glu Ala Leu Glu His Leu
210 215 220

Gly Ala Asp Asp Tyr Leu Val Ser Ser Asp Ala Glu Gly Met Gln Lys
225 230 235 240

Ala Ala Asp Ser Leu Asp Tyr Ile Ile Asp Thr Val Pro Val Phe His
245 250 255

Pro Leu Glu Pro Tyr Leu Ser Leu Leu Lys Phe Asp Gly Lys Leu Ile
260 265 270

Leu Thr Gly Val Ile Asn Thr Pro Leu Gln Phe Val Ser Pro Met Val
275 280 285

Met Leu Gly Arg Lys Ser Ile Thr Gly Ser Phe Ile Gly Ser Met Lys
290 295 300

Glu Thr Glu Glu Met Leu Asn Phe Cys Lys Glu Glu Asn Leu Thr Ser
305 310 315 320

Met Ile Glu Val Val Lys Met Asp Tyr Ile Asn Thr Ala Met Glu Arg
325 330 335

Leu Glu Lys Asn Asp Val Arg Tyr Arg Phe Val Val Asp Val Ala Gly
340 345 350

Ser Lys Leu Asp Gln
355

<210> 69
<211> 355
<212> PRT

<213> Gossypium hirsutum

<400> 69

Met Asp Ser Gln Thr Lys Thr Glu Asn Cys Leu Gly Trp Ala Ala Thr
 1 5 10 15

Asp Pro Thr Gly Val Leu Ser Pro Tyr Lys Phe Ser Arg Arg Pro Leu
 20 25 30

Gly Ser Asp Asp Val Ser Ile Lys Ile Thr His Cys Gly Val Cys Tyr
 35 40 45

Ala Asp Val Ile Trp Ser Arg Asn Met Phe Gly Asp Ser Ile Tyr Pro
 50 55 60

Leu Val Pro Gly His Glu Ile Ala Gly Ile Val Lys Glu Val Gly Ser
 65 70 75 80

Asn Val Gln Arg Ile Lys Val Gly Asp Leu Val Gly Val Gly Thr Tyr
 85 90 95

Val Asn Ser Cys Arg Asn Cys Glu Tyr Cys Asn Asp Gly Val Glu Val
 100 105 110

Gln Cys Val Lys Gly Pro Val Leu Thr Phe Asn His Ile Asp Ile Asp
 115 120 125

Gly Thr Val Thr Lys Gly Gly Tyr Ser Ser His Ile Val Val His Glu
 130 135 140

Arg Tyr Cys Phe Lys Ile Pro Asn Asn Tyr Pro Leu Ala Ser Ala Ala
 145 150 155 160

Pro Leu Leu Cys Ala Gly Ile Thr Val Tyr Thr Pro Met Met Arg Tyr
 165 170 175

Asn Met Asn Gln Pro Gly Lys Ser Leu Gly Val Ile Gly Leu Gly Gly
 180 185 190

Leu Gly His Met Ala Val Lys Phe Gly Lys Ala Phe Gly Leu Ser Val
 195 200 205

Thr Val Leu Ser Thr Ser Ile Ser Lys Lys Glu Glu Ala Leu Ser Leu
 210 215 220

Leu Gly Ala Asp Asn Phe Val Val Thr Ser Asp Gln Glu Gln Met Lys

225						230						235						240
Gly	Leu	Ser	Lys	Ser	Leu	Asp	Phe	Ile	Ile	Asp	Thr	Ala	Ser	Gly	Asp			
			245						250					255				
His	Pro	Phe	Asp	Pro	Tyr	Leu	Ser	Leu	Leu	Lys	Ser	Ala	Gly	Val	Tyr			
			260					265					270					
Ala	Leu	Val	Gly	Phe	Pro	Ser	Glu	Ile	Lys	Phe	Ser	Pro	Ala	Ser	Leu			
		275					280					285						
Asn	Pro	Gly	Met	Lys	Thr	Phe	Ala	Gly	Ser	Val	Thr	Gly	Gly	Thr	Lys			
	290					295					300							
Met	Ile	Gln	Glu	Met	Ile	Gly	Phe	Cys	Ala	Ala	Arg	Lys	Ile	Tyr	Pro			
305					310					315					320			
Gln	Ile	Glu	Val	Ile	Pro	Ile	Gln	Tyr	Ala	Asn	Glu	Ala	Leu	Glu	Arg			
				325					330					335				
Leu	Val	Lys	Lys	Asp	Val	Lys	Tyr	Arg	Phe	Val	Ile	Asp	Ile	Glu	Asn			
			340					345					350					
Thr	Leu	Lys																
		355																
<210>	70																	
<211>	248																	
<212>	PRT																	
<213>	Gossypium	hirsutum																
<400>	70																	
Met	Ala	Thr	Asn	Thr	Thr	Gln	Glu	Gln	Gln	Pro	Ala	Ala	Gly	Arg	His			
1			5						10					15				
Gln	Glu	Val	Gly	His	Lys	Ser	Leu	Leu	Gln	Ser	Asp	Ala	Leu	Tyr	Gln			
			20					25					30					
Tyr	Ile	Leu	Glu	Thr	Ser	Val	Tyr	Pro	Met	Glu	Pro	Glu	Pro	Met	Lys			
		35					40					45						
Glu	Leu	Arg	Glu	Leu	Thr	Ala	Lys	His	Pro	Trp	Asn	Leu	Met	Thr	Thr			
	50					55					60							
Ser	Ala	Asp	Glu	Gly	Gln	Phe	Leu	Asn	Met	Leu	Leu	Lys	Leu	Ile	Asn			
65					70					75					80			

Ala Lys Asn Thr Met Glu Ile Gly Val Tyr Thr Gly Tyr Ser Leu Leu
85 90 95

Ala Thr Ala Leu Ala Leu Pro Asp Asp Gly Lys Ile Leu Ala Met Asp
100 105 110

Ile Asn Arg Glu Asn Tyr Glu Leu Gly Leu Pro Val Ile Arg Lys Ala
115 120 125

Gly Val Ala His Lys Ile Glu Phe Lys Glu Gly Pro Ala Met Pro Val
130 135 140

Leu Asp Lys Leu Val Glu Asp Glu Lys Asn His Gly Ser Tyr Asp Phe
145 150 155 160

Ile Phe Val Asp Ala Asp Lys Asp Asn Tyr Leu Asn Tyr His Lys Arg
165 170 175

Leu Ile Glu Leu Val Lys Val Gly Gly Leu Ile Gly Tyr Asp Asn Thr
180 185 190

Leu Trp Asn Gly Ser Val Val Ala Pro Pro Asp Ala Pro Leu Arg Lys
195 200 205

Tyr Val Arg Tyr Tyr Arg Asp Phe Val Leu Glu Leu Asn Lys Ala Leu
210 215 220

Ala Val Asp Pro Arg Ile Glu Ile Cys Met Leu Pro Val Gly Asp Gly
225 230 235 240

Ile Thr Leu Cys Arg Arg Val Lys
245

<210> 71

<211> 247

<212> PRT

<213> Populus tremuloides

<400> 71

Met Ala Thr Asn Gly Glu Glu Gln Gln Ser Gln Ala Gly Arg His Gln
1 5 10 15

Glu Val Gly His Lys Ser Leu Leu Gln Ser Asp Ala Leu Tyr Gln Tyr
20 25 30

Ile Leu Glu Thr Ser Val Tyr Pro Arg Glu Pro Glu Cys Met Lys Glu

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

<210> 72

<211> 366

<212> PRT

<213> Gossypium hirsutum

<400> 72

Met Gly Ser Ile Gly Glu Thr Gln Met Thr Pro Thr Gln Val Ser Asp
 1 5 10 15

Asp Glu Ala Asn Leu Phe Ala Met Gln Leu Ala Ser Ala Ser Val Leu
 20 25 30

Pro Met Val Leu Lys Ser Ala Ile Glu Leu Asp Leu Leu Glu Ile Met
 35 40 45

Ala Lys Ala Gly Pro Gly Ala Phe Leu Ser Pro Lys Glu Val Ala Ser
 50 55 60

Lys Leu Pro Thr Thr Asn Pro Asp Ala Pro Val Met Leu Asp Arg Ile
 65 70 75 80

Leu Arg Leu Leu Ala Ser Tyr Asn Val Leu Thr Cys Ser Leu Arg Thr
 85 90 95

Phe Pro Gly Gly Lys Val Glu Arg Leu Tyr Gly Leu Gly Pro Val Cys
 100 105 110

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

Leu Met Asn Gln Asp Lys Val Leu Met Glu Ser Trp Tyr Tyr Leu Lys
 130 135 140

Asp Ala Val Leu Asp Gly Gly Ile Pro Phe Asn Lys Ala Tyr Gly Met
 145 150 155 160

Thr Ala Phe Glu Tyr His Gly Thr Asp Pro Arg Phe Asn Lys Val Phe
 165 170 175

Asn Arg Gly Met Ser Asp His Ser Thr Ile Thr Met Lys Lys Ile Leu
 180 185 190

Asp Thr Tyr Asp Gly Phe Gln Gly Leu Lys Thr Leu Val Asp Val Gly
 195 200 205

Gly Gly Thr Gly Ala Thr Leu Ser Met Ile Val Ser Lys Tyr Pro Thr
 210 215 220

Ile Lys Gly Ile Asn Phe Asp Leu Pro His Val Ile Glu Asp Ala Pro
 225 230 235 240

Ser Cys Pro Val Gly Val Glu His Val Gly Gly Asp Met Phe Val Ser
245 250 255

Val Pro Lys Gly Asp Ala Ile Phe Met Lys Trp Ile Cys His Asp Trp
260 265 270

Ser Asp Glu His Cys Ala Lys Phe Leu Lys Asn Cys Tyr Glu Ala Leu
275 280 285

Pro Asp Asn Gly Lys Val Ile Val Ala Glu Cys Ile Leu Pro Asp Tyr
290 295 300

Pro Asp Pro Ser Leu Ala Thr Lys Leu Val Val His Ile Asp Cys Ile
305 310 315 320

Met Leu Ala His Asn Pro Gly Gly Lys Glu Arg Thr Ala Lys Glu Phe
325 330 335

Glu Ala Leu Ala Lys Gly Ala Gly Phe Gln Gly Phe Gln Ile Thr Cys
340 345 350

Ser Ala Phe Gly Thr Asn Ile Met Glu Phe Leu Lys Ser Val
355 360 365

<210> 73

<211> 545

<212> PRT

<213> Corchorus capsularis

<400> 73

Met Ala Pro Gln Ala Pro Glu Leu Pro Gln Glu Asp Phe Ile Phe Arg
1 5 10 15

Ser Lys Leu Pro Asp Ile Tyr Ile Pro Lys His Leu Pro Leu His Ser
20 25 30

Tyr Cys Phe Glu Asn Ile Ser Lys Val Ala Ser Lys Pro Cys Leu Ile
35 40 45

Asn Gly Thr Thr Gly Gln Ile Tyr Thr Tyr Glu Glu Val Glu Leu Thr
50 55 60

Ala Arg Arg Val Ala Ala Gly Leu His Lys Leu Gly Val Gln Gln Arg
65 70 75 80

Gln Val Ile Met Leu Leu Leu Pro Asn Thr Pro Glu Phe Val Leu Ser

85					90					95					
Phe	Leu	Gly	Ala	Ser	Phe	Leu	Gly	Ala	Val	Cys	Thr	Ala	Ala	Asn	Pro
			100					105					110		
Phe	Phe	Thr	Ala	Pro	Glu	Val	Ala	Lys	Gln	Ala	Lys	Ala	Ser	Asn	Ala
		115					120					125			
Arg	Ile	Ile	Ile	Thr	Gln	Ala	Ser	Tyr	Val	Asp	Lys	Val	Lys	Glu	Phe
	130					135					140				
Ala	Gln	Glu	Asn	Val	Asp	Val	Lys	Val	Met	Cys	Ile	Asp	Ser	Ala	Pro
145					150					155					160
Glu	Gly	Cys	Leu	His	Phe	Ser	Glu	Leu	Thr	Gln	Ala	Asp	Glu	Asn	Asp
				165					170					175	
Leu	Pro	Glu	Val	Glu	Ile	Asn	Pro	Asp	Asp	Val	Val	Ala	Leu	Pro	Tyr
			180					185					190		
Ser	Ser	Gly	Thr	Thr	Gly	Leu	Pro	Lys	Gly	Val	Met	Leu	Thr	His	Lys
		195					200					205			
Gly	Leu	Val	Thr	Ser	Val	Ala	Gln	Gln	Val	Asp	Gly	Glu	Asn	Pro	Asn
	210					215					220				
Leu	Tyr	Phe	His	Ser	Asp	Asp	Val	Ile	Leu	Cys	Thr	Leu	Pro	Leu	Phe
225					230					235					240
His	Ile	Tyr	Ala	Leu	Asn	Ser	Ile	Met	Leu	Cys	Gly	Leu	Arg	Ala	Gly
				245					250					255	
Ala	Ala	Ile	Leu	Ile	Met	Gln	Lys	Phe	Glu	Ile	Gly	Leu	Leu	Leu	Asp
			260					265					270		
Leu	Ile	Gln	Lys	Tyr	Lys	Ile	Thr	Ile	Ala	Pro	Met	Val	Pro	Pro	Ile
		275					280					285			
Val	Leu	Ala	Ile	Ala	Lys	Ser	Ser	Glu	Thr	Glu	Lys	Tyr	Asp	Leu	Ser
	290					295					300				
Ser	Ile	Arg	Met	Val	Lys	Ser	Gly	Ala	Ala	Pro	Leu	Gly	Lys	Glu	Leu
305					310					315					320
Glu	Asp	Ala	Val	Arg	Ala	Lys	Phe	Pro	Gly	Ala	Lys	Leu	Gly	Gln	Gly

```

          325                      330                      335

Tyr Gly Met Thr Glu Ala Gly Pro Val Leu Ala Met Cys Leu Gly Phe
      340                      345                      350

Ala Lys Glu Pro Phe Glu Ile Lys Ser Gly Ala Cys Gly Thr Val Val
      355                      360                      365

Arg Asn Ala Glu Met Lys Ile Val Asp Pro Asp Thr Gly Ala Ser Leu
      370                      375                      380

Pro Arg Asn Gln Ala Gly Glu Ile Cys Ile Arg Gly Asp Gln Ile Met
      385                      390                      395                      400

Lys Gly Tyr Leu Asn Asp Pro Glu Ala Thr Ala Arg Thr Ile Asp Lys
      405                      410                      415

Asp Gly Trp Leu His Thr Gly Asp Ile Gly Tyr Ile Asp Asp Asp Asp
      420                      425                      430

Glu Leu Phe Ile Val Asp Arg Leu Lys Glu Leu Ile Lys Tyr Lys Gly
      435                      440                      445

Phe Gln Val Ala Pro Ala Glu Leu Glu Ala Met Leu Ile Ala His Pro
      450                      455                      460

Glu Ile Ile Asp Ala Ala Val Val Ala Met Lys Asp Glu Val Ala Gly
      465                      470                      475                      480

Glu Val Pro Val Ala Phe Val Val Lys Ser Glu Lys Ser Glu Ile Thr
      485                      490                      495

Glu Asp Glu Ile Lys Gln Tyr Ile Ser Lys Gln Val Val Phe Tyr Lys
      500                      505                      510

Arg Ile Ser Arg Val Phe Phe Met Glu Ala Ile Pro Lys Ala Pro Ser
      515                      520                      525

Gly Lys Ile Leu Arg Lys Glu Leu Arg Ala Lys Leu Ala Ser Gly Asn
      530                      535                      540

Tyr
545

```

<210> 74

<211> 542

<212> PRT

<213> Ricinus communis

<400> 74

Met Glu Lys Ser Gly Tyr Gly Arg Asp Gly Ile Tyr Arg Ser Leu Arg
 1 5 10 15

Pro Leu Leu Val Leu Pro Lys Asp Pro Asn Leu Ser Met Val Ser Phe
 20 25 30

Leu Phe Arg Asn Cys Asn Ser Tyr Pro His Lys Pro Ala Leu Ile Asp
 35 40 45

Ala Asp Leu Ser Lys Thr Leu Ser Phe Ser Gln Leu Lys Ser Ile Val
 50 55 60

Ile Lys Val Ser His Gly Leu Leu Lys Leu Gly Ile Ser Lys Asn Asp
 65 70 75 80

Val Val Leu Ile Phe Ala Pro Asn Ser Tyr Gln Phe Pro Ile Cys Phe
 85 90 95

Leu Ala Ile Thr Ser Ile Gly Ala Ile Ala Thr Thr Ala Asn Pro Leu
 100 105 110

Tyr Thr Thr Thr Glu Ile Ser Lys Gln Ile Lys Asp Ser Asn Pro Lys
 115 120 125

Leu Val Ile Thr Val Pro Glu Leu Trp Asn Lys Val Lys Asp Phe Asn
 130 135 140

Leu Pro Ala Val Phe Leu Gly Ala Lys Glu Ser Leu Leu Ile Glu Pro
 145 150 155 160

Asn Ser Arg Ile Lys Ser Phe Asp His Leu Val Glu Leu Gly Gly Ser
 165 170 175

Asn Ser Glu Phe Pro Thr Ile Asn Val Lys Gln Thr Asp Ile Ala Thr
 180 185 190

Leu Leu Tyr Ser Ser Gly Thr Thr Gly Ile Ser Lys Gly Val Ile Leu
 195 200 205

Thr His Gly Asn Phe Ile Ala Ala Ser Gln Met Ile Thr Met Asp Gln
 210 215 220

Glu Ile Ala Gly Glu Leu His Asn Val Phe Leu Cys Phe Leu Pro Met
 225 230 235 240

Phe His Val Phe Gly Leu Ala Val Ile Ala Tyr Ser Gln Leu Gln Thr
 245 250 255

Gly Asn Ala Val Val Ser Met Gly Lys Phe Asp Phe Glu Leu Val Leu
 260 265 270

Lys Ala Val Glu Lys Tyr Arg Ile Thr His Leu Trp Val Val Pro Pro
 275 280 285

Val Ile Leu Ala Leu Ala Lys Gln Ser Leu Val Lys Lys Tyr Asp Leu
 290 295 300

Ser Ser Leu Gln His Val Gly Ser Gly Ala Ala Pro Leu Ser Lys Glu
 305 310 315 320

Leu Met Glu Glu Cys Ala Lys Thr Ile Pro His Ala Ala Ile Ala Gln
 325 330 335

Gly Tyr Gly Met Thr Glu Thr Thr Gly Ile Val Ser Val Glu Asn Pro
 340 345 350

Arg Ile Gly Val Arg His Ser Gly Ser Ala Gly Thr Leu Ala Ala Gly
 355 360 365

Ile Glu Ala Gln Ile Ile Ser Val Asp Thr Leu Lys Pro Leu Pro Pro
 370 375 380

Asn Gln Leu Gly Glu Ile Trp Val Arg Gly Pro Asn Met Met Arg Gly
 385 390 395 400

Tyr Phe Asn Asn Pro Gln Ala Thr Lys Gln Thr Ile Asp Lys Lys Gly
 405 410 415

Trp Leu His Thr Gly Asp Leu Gly Tyr Phe Asp Glu Asp Gly Lys Leu
 420 425 430

Tyr Val Val Asp Arg Ile Lys Glu Leu Ile Lys Tyr Lys Gly Phe Gln
 435 440 445

Val Ala Pro Ala Glu Leu Glu Gly Leu Leu Val Ser His Pro Glu Leu
 450 455 460

Leu Asp Ala Val Val Ile Pro Phe Pro Asp Ala Glu Ala Gly Glu Val
465 470 475 480

Pro Val Ala Tyr Val Val Arg Ser Pro Asn Ser Ser Leu Thr Glu Glu
485 490 495

Glu Val Gln Lys Tyr Ile Ala Asp Gln Val Ala Pro Phe Lys Arg Leu
500 505 510

Arg Arg Val Thr Phe Ile Asn Thr Val Pro Lys Ser Ala Ser Gly Lys
515 520 525

Ile Leu Arg Arg Glu Leu Ile Glu Lys Val Lys Ser Lys Leu
530 535 540

<210> 75

<211> 544

<212> PRT

<213> Populus trichocarpa

<400> 75

Met Glu Lys Leu Thr Ala Thr Ala Ser Lys Thr Thr Leu Pro Glu Trp
1 5 10 15

Tyr Ser Pro Glu Thr Gly Ile Tyr Ser Ser Lys His Pro Pro Ile Pro
20 25 30

Leu Pro Ala Asp Pro Phe Leu Asp Val Val Ser Phe Ile Phe Ser His
35 40 45

His Asn His Asn Gly Leu Thr Ala Leu Ile Asp Ser Ser Ser Gly Phe
50 55 60

Ser Ile Ser Tyr Ser Lys Ile Leu Pro Leu Val Lys Ser Val Ala Ser
65 70 75 80

Gly Leu Ser Asn Met Gly Ile Lys Gln Gly Asp Val Val Leu Leu Leu
85 90 95

Leu Pro Asn Ser Ile His Phe Pro Ile Val Phe Leu Gly Val Leu Tyr
100 105 110

Leu Gly Gly Ile Val Ser Thr Met Asn Pro Leu Ser Ser Glu Leu Glu
115 120 125

Val Lys Gln Arg Ile Val Asp Cys Asn Ala Cys Ile Ala Phe Val Glu

130				135				140							
Leu 145	Glu	Lys	Val	Cys	Lys 150	Phe	Gln	Pro	Leu	Gly 155	Ile	Pro	Val	Ile	Gly 160
Val	Pro	Glu	Asn	Val 165	Asn	Phe	Asp	Glu	Lys 170	Ile	Tyr	Ser	Lys	Gly 175	Asp
Val	Gly	Val	Lys 180	Pro	Val	Ile	Arg	Gln 185	Gln	Asp	Thr	Ala	Ala 190	Ile	Met
Tyr	Ser	Ser 195	Gly	Thr	Thr	Ala	Ala 200	Ser	Lys	Gly	Val	Val 205	Leu	Thr	His
Arg	Asn 210	Phe	Ile	Ser	Met	Val 215	Glu	Leu	Phe	Val	Lys 220	Phe	Glu	Ala	Ser
Gln 225	Tyr	Glu	Tyr	Ser	Ser 230	Thr	Asp	Asn	Val	Tyr 235	Leu	Ala	Val	Leu	Pro 240
Met	Phe	His	Ile	Tyr 245	Gly	Leu	Ser	Leu	Phe 250	Val	Val	Gly	Leu	Leu 255	Ser
Leu	Gly	Ser	Ser 260	Ile	Val	Val	Met	Arg 265	Lys	Phe	Asp	Val	Ser 270	Glu	Met
Val	Lys 275	Val	Ile	Asp	Arg	Tyr	Gly 280	Val	Thr	His	Phe	Pro 285	Val	Val	Pro
Pro	Ile 290	Leu	Thr	Ala	Leu	Thr 295	Arg	Thr	Ala	Lys	Gly 300	Val	Cys	Gly	Asn
Ser 305	Leu	Lys	Cys	Leu	Lys 310	Leu	Val	Ser	Cys	Gly 315	Ala	Ala	Pro	Leu	Phe 320
Gly	Lys	Thr	Ile	Gln 325	Asp	Phe	Val	Glu	Val 330	Leu	Pro	His	Val	Asp 335	Phe
Ile	Gln	Gly	Tyr 340	Gly	Leu	Thr	Glu	Ser 345	Thr	Ala	Val	Gly	Thr 350	Arg	Gly
Leu	Asn 355	Thr	Glu	Lys	Phe	Gln	Lys 360	Tyr	Ser	Ser	Ile	Gly 365	Leu	Leu	Ala
Pro	Asn	Ile	Glu	Ala	Lys	Val	Val	Asp	Trp	Ile	Thr	Gly	Ala	Leu	Leu

370 375 380
 Pro Pro Gly Gly Ser Gly Glu Leu Trp Ile Arg Gly Pro Gly Val Met
 385 390 395 400

 Lys Glu Tyr Leu Asn Gly Gly Glu Ala Thr Ala Leu Thr Ile Asp Lys
 405 410 415

 Asp Gly Trp Leu His Thr Gly Asp Val Val Tyr Ala Asp His Asp Gly
 420 425 430

 Tyr Leu Tyr Val Val Asp Arg Leu Lys Glu Ile Ile Lys Tyr Lys Gly
 435 440 445

 Phe Gln Ile Ala Pro Ala Asp Leu Glu Ala Val Leu Ile Ser His Cys
 450 455 460

 Glu Ile Leu Asp Ala Ala Val Ile Pro Val Val Asp Lys Glu Cys Gly
 465 470 475 480

 Glu Ile Pro Val Ala Phe Val Val Lys Arg Gln Gly Ser Met Leu Thr
 485 490 495

 Gln Glu Ala Ile Ile Asn Tyr Val Ala Glu Gln Val Ala Pro Tyr Lys
 500 505 510

 Lys Val Arg Lys Val Ile Phe Thr Gln Ser Ile Pro Lys Ser Ala Ala
 515 520 525

 Gly Lys Ile Leu Arg Arg Glu Leu Lys Cys Ser Leu Thr Ser Lys Leu
 530 535 540

 <210> 76
 <211> 433
 <212> PRT
 <213> Populus trichocarpa

 <400> 76
 Met Ile Ile Asn Val Lys Glu Ser Thr Met Val Gln Pro Ala Glu Glu
 1 5 10 15

 Thr Pro Arg Arg Gly Leu Trp Asn Ser Asn Val Asp Leu Val Val Pro
 20 25 30

 Arg Phe His Thr Pro Ser Val Tyr Phe Tyr Arg Pro Thr Gly Ala Ser
 35 40 45

Asn Phe Phe Asp Ala Lys Val Leu Lys Glu Ala Leu Ser Lys Ala Leu
 50 55 60

Val Pro Phe Tyr Pro Met Ala Gly Arg Leu Arg Arg Asp Asp Asp Gly
 65 70 75 80

Arg Ile Glu Ile Asp Cys Asn Ala Glu Gly Val Leu Phe Val Glu Ala
 85 90 95

Gly Thr Ala Ser Val Val Ala Asp Phe Gly Asp Phe Ala Pro Thr Leu
 100 105 110

Glu Leu Lys Gln Leu Ile Pro Thr Val Asp Tyr Ser Gly Gly Ile Ser
 115 120 125

Thr Tyr Pro Leu Leu Val Leu Gln Val Thr Tyr Phe Lys Cys Gly Gly
 130 135 140

Val Ser Leu Gly Val Gly Met Gln His His Ala Ala Asp Gly Phe Ser
 145 150 155 160

Gly Leu His Phe Val Asn Thr Trp Ser Asp Met Ala Arg Gly Leu Asp
 165 170 175

Leu Thr Ile Pro Pro Phe Ile Asp Arg Thr Leu Leu Arg Ala Arg Asp
 180 185 190

Pro Pro Gln Pro Ala Phe His His Val Glu Tyr Gln Pro Pro Pro Ala
 195 200 205

Met Lys Thr Val Leu Glu Thr Ser Lys Pro Glu Ser Thr Ala Val Ser
 210 215 220

Ile Phe Lys Leu Thr Arg Asp Gln Leu Asn Thr Leu Lys Ala Lys Ala
 225 230 235 240

Lys Glu Gly Gly Asn Asn Ile Gly Tyr Ser Ser Tyr Glu Met Leu Ala
 245 250 255

Gly His Val Trp Arg Ser Ala Cys Lys Ala Arg Gly Leu Pro Asp Asp
 260 265 270

Gln Glu Thr Lys Leu Tyr Ile Ala Thr Asp Gly Arg Ser Arg Leu Arg
 275 280 285

Pro Thr Leu Pro Pro Gly Tyr Phe Gly Asn Val Ile Phe Thr Ala Thr
 290 295 300

Pro Ile Ala Val Ala Gly Glu Ile Gln Ser Lys Pro Thr Trp Tyr Ala
 305 310 315 320

Ala Gly Lys Ile His Asp Ser Leu Val Arg Met Asp Asn Asp Tyr Leu
 325 330 335

Arg Ser Ala Leu Asp Phe Leu Glu Leu Gln Pro Asp Leu Ser Ala Leu
 340 345 350

Val Arg Gly Ala His Thr Phe Arg Cys Pro Asn Leu Gly Ile Thr Ser
 355 360 365

Trp Val Arg Leu Pro Ile His Asp Ala Asp Phe Gly Trp Gly Arg Pro
 370 375 380

Ile Phe Met Gly Pro Gly Gly Ile Ala Tyr Glu Gly Leu Ser Phe Ile
 385 390 395 400

Ile Pro Ser Ser Thr Asn Asp Gly Ser Leu Ser Val Ala Ile Ser Leu
 405 410 415

Gln Ala Glu His Met Lys Leu Phe Glu Lys Phe Ile Tyr Asp Ile Lys
 420 425 430

Glu

<210> 77

<211> 436

<212> PRT

<213> Cynara cardunculus

<400> 77

Met Lys Ile Glu Val Arg Glu Ser Thr Met Val Arg Pro Ala Glu Glu
 1 5 10 15

Thr Pro Arg Ile Asn Leu Trp Asn Ser Asn Val Asp Leu Val Val Pro
 20 25 30

Asn Phe His Thr Pro Ser Val Tyr Phe Tyr Arg Pro Asn Gly Ala Ala
 35 40 45

Asn Phe Phe Asp Pro Lys Val Met Lys Asp Ala Leu Ser Arg Ala Leu

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

290 295 300
 Ile Phe Thr Thr Thr Pro Ile Ala Val Ala Gly Asp Leu Gln Ser Lys
 305 310 315 320
 Pro Thr Trp Tyr Ala Ala Ser Lys Ile His Asp Ala Leu Ala Arg Met
 325 330 335
 Asp Asp Asp Tyr Leu Lys Ser Ala Leu Asp Tyr Leu Glu Leu Gln Pro
 340 345 350
 Asp Leu Lys Ala Leu Val Arg Gly Ala His Thr Phe Lys Cys Pro Asn
 355 360 365
 Leu Gly Ile Thr Ser Trp Ala Arg Leu Pro Ile His Asp Ala Asp Phe
 370 375 380
 Gly Trp Gly Arg Pro Ile Phe Met Gly Pro Gly Gly Ile Ala Tyr Glu
 385 390 395 400
 Gly Leu Ser Phe Val Leu Pro Ser Pro Ile Asn Asp Gly Ser Leu Ser
 405 410 415
 Ile Val Ile Ser Leu Gln Ala Glu His Met Lys Leu Phe Ser Lys Phe
 420 425 430
 Leu Tyr Asp Ile
 435
 <210> 78
 <211> 509
 <212> PRT
 <213> Eucalyptus globulus
 <400> 78
 Met Ala Leu Pro Leu Ile Leu Leu Ser Ile Pro Leu Leu Phe Leu Leu
 1 5 10 15
 Leu Ala His Gln Leu Tyr Gln Arg Leu Arg Phe Lys Leu Pro Pro Gly
 20 25 30
 Pro Arg Ala Trp Pro Val Val Gly Asn Leu Tyr Asp Ile Lys Pro Val
 35 40 45
 Arg Phe Arg Cys Phe Ala Glu Trp Ser Gln Ala Tyr Gly Pro Ile Ile
 50 55 60

Ser Val Trp Phe Gly Ser Thr Leu Asn Val Val Val Ser Ser Ser Glu
65 70 75 80

Leu Ala Lys Glu Val Leu Lys Glu Asn Asp Gln Gln Leu Ala Asp Arg
85 90 95

His Arg Ser Arg Ser Ala Ala Lys Phe Ser Arg Asp Gly Gln Asp Leu
100 105 110

Ile Trp Ala Asp Tyr Gly Pro His Tyr Val Lys Val Arg Lys Val Cys
115 120 125

Thr Leu Glu Leu Phe Thr Pro Lys Arg Leu Glu Ala Leu Arg Pro Ile
130 135 140

Arg Glu Asp Glu Val Thr Ala Met Val Glu Ser Ile Phe Lys Asp Cys
145 150 155 160

Thr Asn Pro Asp Asn Ser Gly Lys Thr Leu Leu Val Lys Lys Tyr Leu
165 170 175

Gly Ala Val Ala Phe Asn Asn Ile Thr Arg Leu Ala Phe Gly Lys Arg
180 185 190

Phe Met Asn Ala Glu Gly Val Ile Asp Glu Gln Gly Leu Glu Phe Lys
195 200 205

Ala Ile Val Ser Asn Gly Leu Lys Leu Gly Ala Ser Leu Ala Met Ala
210 215 220

Glu His Ile Pro Trp Leu Arg Trp Met Phe Pro Leu Glu Glu Glu Ala
225 230 235 240

Phe Ala Lys His Ser Ala Arg Arg Asp Arg Leu Thr Arg Ala Ile Met
245 250 255

Glu Glu His Thr Val Ala Arg Gln Lys Ser Gly Ala Lys Gln His Phe
260 265 270

Val Asp Ala Leu Leu Thr Leu Lys Asp Lys Tyr Asp Leu Ser Glu Asp
275 280 285

Thr Ile Ile Gly Leu Leu Trp Asp Met Ile Thr Ala Gly Met Asp Thr
290 295 300

Thr Ala Ile Ser Val Glu Trp Ala Met Ala Glu Leu Ile Lys Asn Pro
 305 310 315 320

Arg Val Gln Gln Lys Ala Gln Glu Glu Leu Asp Arg Val Val Gly Phe
 325 330 335

Glu Arg Val Val Thr Glu Pro Asp Phe Ser Asn Leu Pro Tyr Leu Gln
 340 345 350

Cys Ile Ala Lys Glu Ala Leu Arg Leu His Pro Pro Thr Pro Leu Met
 355 360 365

Leu Pro His Arg Ser Asn Ser His Val Lys Ile Gly Gly Tyr Asp Ile
 370 375 380

Pro Lys Gly Ser Asn Val His Val Asn Val Trp Ala Ile Ala Arg Asp
 385 390 395 400

Pro Ala Val Trp Asn Ser Pro Leu Glu Phe Arg Pro Glu Arg Phe Leu
 405 410 415

Glu Glu Asp Val Asp Met Lys Gly His Asp Phe Arg Leu Leu Pro Phe
 420 425 430

Gly Ala Gly Arg Arg Val Cys Pro Gly Ala Gln Leu Gly Ile Asn Leu
 435 440 445

Val Thr Ser Met Leu Gly His Leu Leu His His Phe Val Trp Thr Pro
 450 455 460

Pro Gln Gly Thr Lys Pro Glu Glu Ile Asp Met Ser Glu Asn Pro Gly
 465 470 475 480

Leu Val Thr Tyr Met Ser Thr Pro Val Gln Ala Val Ala Thr Pro Arg
 485 490 495

Leu Pro Ser Glu Leu Tyr Lys Arg Val Pro Tyr Glu Met
 500 505

<210> 79

<211> 508

<212> PRT

<213> Populus alba x Populus grandidentata

<400> 79

Met Asn Leu Leu Leu Ile Pro Ile Ser Phe Ile Thr Ile Leu Leu Thr

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

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

485

490

495

Pro Ser His Leu Tyr Lys Arg Val Ala Val Asp Ile
500 505

<210> 80

<211> 509

<212> PRT

<213> Populus trichocarpa

<400> 80

Met Ala Leu Pro Leu Leu Val Leu Val Ser Ile Phe Val Leu Leu Leu
1 5 10 15

Ala Tyr Ile Leu Tyr Gln Arg Leu Arg Phe Lys Leu Pro Pro Gly Pro
20 25 30

Arg Pro Trp Pro Ile Val Gly Asn Leu Tyr Ala Ile Lys Pro Ile Arg
35 40 45

Phe Arg Cys Phe Ala Glu Trp Ala Gln Ala Tyr Gly Pro Val Val Ser
50 55 60

Val Trp Phe Gly Ser Thr Leu Asn Val Val Val Cys Asn Ala Glu Leu
65 70 75 80

Ala Lys Gln Val Leu Lys Glu Asn Asp Gln Gln Leu Ala Asp Arg His
85 90 95

Arg Ser Arg Leu Ala Ala Arg Phe Ser Arg Asp Gly Lys Asp Leu Ile
100 105 110

Trp Ala Asp Tyr Gly Pro His Tyr Val Lys Val Arg Arg Val Ser Thr
115 120 125

Leu Glu Leu Phe Ser Ala Lys Arg Leu Glu Glu Leu Arg Pro Ile Arg
130 135 140

Glu Asp Glu Val Thr Phe Met Ala Glu Ser Ile Phe Lys Asp Cys Thr
145 150 155 160

Asn Pro Glu Asn His Gly Lys Ser Leu Leu Val Lys Lys Tyr Leu Gly
165 170 175

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

Met	Asn	Ser	Glu	Gly	Ile	Ile	Asp	Glu	Gln	Gly	Gln	Glu	Phe	Lys	Ala	195	200	205
Ile	Val	Ser	Asn	Gly	Val	Arg	Leu	Gly	Gly	Ser	Leu	Thr	Met	Ala	Glu	210	215	220
His	Ile	Pro	Trp	Leu	Gln	Trp	Met	Phe	Pro	Leu	Glu	Glu	Glu	Ala	Val	225	230	235
Glu	Lys	His	Asn	Ala	Arg	Arg	Asp	Gly	Leu	Thr	Arg	Val	Ile	Met	Glu	245	250	255
Glu	His	Thr	Asn	Ala	Arg	Lys	Lys	Ser	Gly	Gly	Ala	Lys	Lys	His	Phe	260	265	270
Val	Asp	Ala	Leu	Leu	Thr	Leu	Gln	Glu	Lys	Tyr	Asp	Leu	Ser	Glu	Val	275	280	285
Thr	Ile	Thr	Gly	Leu	Leu	Trp	Asp	Met	Ile	Thr	Ala	Gly	Met	Asp	Thr	290	295	300
Thr	Ala	Ile	Thr	Val	Glu	Trp	Ala	Met	Ala	Glu	Leu	Ile	Lys	Asn	Pro	305	310	315
Arg	Val	Gln	Gln	Lys	Ala	Gln	Asp	Glu	Leu	Asp	Arg	Val	Val	Gly	Phe	325	330	335
Glu	Arg	Val	Met	Thr	Glu	Ala	Asp	Phe	Pro	Asn	Leu	Pro	Tyr	Leu	Gln	340	345	350
Ala	Val	Val	Lys	Glu	Ser	Leu	Arg	Leu	His	Pro	Pro	Thr	Pro	Leu	Met	355	360	365
Leu	Pro	His	Arg	Ala	Asn	Thr	Thr	Val	Lys	Ile	Gly	Gly	Tyr	Asp	Ile	370	375	380
Pro	Lys	Gly	Ser	Val	Val	His	Val	Asn	Val	Trp	Ala	Val	Ala	Arg	Asp	385	390	395
Pro	Ala	Leu	Trp	Lys	Asn	Pro	Leu	Glu	Phe	Arg	Pro	Glu	Arg	Phe	Phe	405	410	415
Glu	Glu	Asp	Val	Asp	Met	Arg	Gly	His	Asp	Phe	Arg	Leu	Leu	Pro	Phe	420	425	430

Gly Ala Gly Arg Arg Val Cys Pro Gly Ala Gln Leu Gly Ile Asn Leu
 435 440 445

Val Thr Ser Ile Ile Gly His Leu Leu His His Phe His Trp Thr Thr
 450 455 460

Pro Asp Gly Val Lys Pro Glu Glu Ile Asp Met Ser Glu Arg Pro Gly
 465 470 475 480

Leu Val Thr Tyr Met Met Thr Pro Leu Gln Ala Val Ala Thr Pro Arg
 485 490 495

Leu Pro Ser His Leu Tyr Lys Arg Met Ala Ser Asp Met
 500 505

<210> 81

<211> 505

<212> PRT

<213> Gossypium arboreum

<400> 81

Met Asp Leu Leu Phe Leu Glu Lys Ala Leu Leu Gly Leu Phe Val Ala
 1 5 10 15

Val Val Leu Ala Ile Thr Ile Ser Lys Leu Arg Gly Lys Arg Phe Lys
 20 25 30

Leu Pro Pro Gly Pro Leu Pro Val Pro Val Phe Gly Asn Trp Leu Gln
 35 40 45

Val Gly Asp Asp Leu Asn His Arg Asn Leu Thr Asp Leu Ala Lys Lys
 50 55 60

Phe Gly Asp Ile Phe Leu Leu Arg Met Gly Gln Arg Asn Leu Val Val
 65 70 75 80

Val Ser Ser Pro Glu Leu Ala Lys Glu Val Leu His Thr Gln Gly Val
 85 90 95

Glu Phe Gly Ser Arg Thr Arg Asn Val Val Phe Asp Ile Phe Thr Gly
 100 105 110

Lys Gly Gln Asp Met Val Phe Thr Val Tyr Gly Glu His Trp Arg Lys
 115 120 125

Met Arg Arg Ile Met Thr Val Pro Phe Phe Thr Asn Lys Val Val Gln

130																			
Gln	Tyr	Arg	Phe	Gly	Trp	Glu	Asp	Glu	Ala	Ala	Arg	Val	Val	Glu	Asp				
145					150					155					160				
Val	Arg	Lys	Asn	Pro	Glu	Ala	Ala	Thr	Asn	Gly	Ile	Val	Leu	Arg	Arg				
				165					170					175					
Arg	Leu	Gln	Leu	Met	Met	Tyr	Asn	Asn	Met	Tyr	Arg	Ile	Met	Phe	Asp				
			180					185					190						
Thr	Arg	Phe	Glu	Ser	Glu	Asp	Asp	Pro	Leu	Phe	Val	Arg	Leu	Lys	Ala				
		195					200					205							
Leu	Asn	Gly	Glu	Arg	Ser	Arg	Leu	Ala	Gln	Ser	Phe	Glu	Tyr	Asn	Tyr				
	210					215					220								
Gly	Asp	Phe	Ile	Pro	Ile	Leu	Arg	Pro	Phe	Leu	Arg	Gly	Tyr	Leu	Lys				
225					230					235					240				
Ile	Cys	Lys	Glu	Val	Lys	Asp	Arg	Arg	Leu	Gln	Leu	Phe	Lys	Asp	His				
				245					250					255					
Phe	Val	Glu	Glu	Arg	Lys	Lys	Leu	Gly	Ser	Thr	Lys	Ser	Met	Asn	Asn				
			260					265					270						
Asp	Gly	Leu	Lys	Cys	Ala	Ile	Asp	His	Ile	Leu	Asp	Ala	Gln	Gln	Lys				
		275					280					285							
Gly	Glu	Ile	Asn	Glu	Asp	Asn	Val	Leu	Tyr	Ile	Val	Glu	Asn	Ile	Asn				
	290					295					300								
Val	Ala	Ala	Ile	Glu	Thr	Thr	Leu	Trp	Ser	Ile	Glu	Trp	Gly	Ile	Ala				
305					310					315					320				
Glu	Leu	Val	Asn	His	Pro	Glu	Ile	Gln	Lys	Lys	Leu	Arg	His	Glu	Leu				
				325					330					335					
Asp	Thr	Val	Leu	Gly	Pro	Gly	Asn	Gln	Ile	Thr	Glu	Pro	Asp	Thr	His				
			340				345						350						
Lys	Leu	Pro	Tyr	Leu	Gln	Ala	Val	Ile	Lys	Glu	Thr	Leu	Arg	Leu	Arg				
		355					360					365							
Met	Ala	Ile	Pro	Leu	Leu	Val	Pro	His	Met	Asn	Leu	His	Asp	Ala	Lys				

370 375 380
 Leu Gly Gly Tyr Asp Ile Pro Ala Glu Ser Lys Ile Leu Val Asn Ala
 385 390 395 400
 Trp Trp Leu Ala Asn Asn Pro Ala Asn Trp Lys Asn Pro Glu Glu Phe
 405 410 415
 Arg Pro Glu Arg Phe Phe Glu Glu Glu Ala Lys Val Glu Ala Asn Gly
 420 425 430
 Asn Asp Phe Arg Tyr Leu Pro Phe Gly Val Gly Arg Arg Ser Cys Pro
 435 440 445
 Gly Ile Ile Leu Ala Leu Pro Ile Leu Gly Ile Thr Leu Gly Arg Leu
 450 455 460
 Val Gln Asn Phe Glu Leu Leu Pro Pro Pro Gly Gln Ser Gln Ile Asp
 465 470 475 480
 Thr Thr Glu Lys Gly Gly Gln Phe Ser Leu His Ile Leu Lys His Ser
 485 490 495
 Thr Ile Val Ala Lys Pro Arg Gln Phe
 500 505

 <210> 82
 <211> 505
 <212> PRT
 <213> Gossypium hirsutum

 <400> 82
 Met Asp Leu Leu Phe Leu Glu Lys Ala Leu Leu Gly Leu Phe Val Ala
 1 5 10 15
 Val Val Leu Ala Ile Thr Ile Ser Lys Leu Arg Gly Lys Arg Phe Lys
 20 25 30
 Leu Pro Pro Gly Pro Leu Pro Val Pro Val Phe Gly Asn Trp Leu Gln
 35 40 45
 Val Gly Asp Asp Leu Asn His Arg Asn Leu Thr Asp Leu Ala Lys Lys
 50 55 60
 Phe Gly Asp Ile Phe Leu Leu Arg Met Gly Gln Arg Asn Leu Val Val
 65 70 75 80

Val Ser Ser Pro Glu Leu Ala Lys Glu Val Leu His Thr Gln Gly Val
85 90 95

Glu Phe Gly Ser Arg Thr Arg Asn Val Val Phe Gly Ile Phe Thr Gly
100 105 110

Lys Gly Gln Asp Met Val Phe Thr Val Tyr Gly Glu His Trp Arg Lys
115 120 125

Met Arg Arg Ile Met Thr Val Pro Phe Phe Thr Asn Lys Val Val Gln
130 135 140

Gln Tyr Arg Phe Gly Trp Glu Asp Glu Ala Ala Arg Val Val Glu Asp
145 150 155 160

Val Arg Lys Asn Pro Glu Ala Ala Thr Asn Gly Ile Val Leu Arg Arg
165 170 175

Arg Leu Gln Leu Met Met Tyr Asn Asn Met Tyr Arg Ile Met Phe Asp
180 185 190

Thr Arg Phe Glu Ser Glu Asp Asp Pro Leu Phe Val Arg Leu Lys Ala
195 200 205

Leu Asn Gly Glu Arg Ser Arg Leu Ala Gln Ser Phe Glu Tyr Asn Tyr
210 215 220

Gly Asp Phe Ile Pro Ile Leu Arg Pro Phe Leu Arg Gly Tyr Leu Lys
225 230 235 240

Ile Cys Lys Glu Val Lys Asp Arg Arg Leu Gln Leu Phe Lys Asp His
245 250 255

Phe Val Glu Glu Arg Lys Lys Leu Gly Ser Thr Lys Ser Met Asn Asn
260 265 270

Asp Gly Leu Lys Cys Ala Ile Asp His Ile Leu Asp Ala Gln Gln Lys
275 280 285

Gly Glu Ile Asn Glu Asp Asn Val Leu Tyr Ile Val Glu Asn Ile Asn
290 295 300

Val Ala Ala Ile Glu Thr Thr Leu Trp Ser Ile Glu Trp Gly Ile Ala
305 310 315 320

Glu Leu Val Asn His Pro Glu Thr Gln Lys Lys Leu Arg His Glu Leu
325 330 335

Asp Thr Val Leu Gly Pro Gly Asn Gln Ile Thr Glu Pro Asp Thr His
340 345 350

Lys Leu Pro Tyr Leu Gln Ala Val Ile Lys Glu Thr Leu Arg Leu Arg
355 360 365

Met Ala Ile Pro Leu Leu Val Pro His Met Asn Leu His Asp Ala Lys
370 375 380

Leu Gly Gly Tyr Asp Ile Pro Ala Glu Ser Lys Ile Leu Val Asn Ala
385 390 395 400

Trp Trp Leu Ala Asn Asn Pro Ala Asn Trp Lys Asn Pro Glu Glu Phe
405 410 415

Arg Pro Glu Arg Phe Phe Glu Glu Glu Ala Lys Val Glu Ala Asn Gly
420 425 430

Asn Asp Phe Arg Tyr Leu Pro Phe Gly Val Gly Arg Arg Ser Cys Pro
435 440 445

Gly Ile Ile Leu Ala Leu Pro Ile Leu Gly Ile Thr Leu Gly Arg Leu
450 455 460

Val Gln Asn Phe Glu Leu Leu Pro Pro Pro Gly Gln Ser Gln Ile Asp
465 470 475 480

Thr Thr Glu Lys Gly Gly Gln Phe Ser Leu His Ile Leu Lys His Ser
485 490 495

Thr Ile Val Ala Lys Pro Arg Gln Phe
500 505

<210> 83

<211> 711

<212> PRT

<213> Populus trichocarpa

<400> 83

Met Glu Phe Cys Gln Asp Ser Arg Asn Gly Asn Gly Ser Leu Gly Phe
1 5 10 15

Asn Thr Asn Asp Pro Leu Asn Trp Gly Met Ala Ala Glu Ser Leu Lys

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

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

	500		505		510														
Gly	Leu	Cys	Gln	Ala	Val	Asp	Leu	Arg	His	Ile	Glu	Glu	Asn	Leu	Lys				
	515						520					525							
Ser	Thr	Val	Lys	Asn	Thr	Val	Ser	Gln	Val	Ala	Lys	Arg	Val	Leu	Thr				
	530					535					540								
Met	Gly	Phe	Asn	Gly	Glu	Leu	His	Pro	Ser	Arg	Phe	Cys	Glu	Lys	Asp				
545					550					555					560				
Leu	Leu	Lys	Val	Val	Asp	Arg	Glu	His	Val	Phe	Ser	Tyr	Ile	Asp	Asp				
				565					570					575					
Pro	Cys	Ser	Ala	Thr	Tyr	Pro	Leu	Met	Gln	Lys	Leu	Arg	Gln	Val	Leu				
			580					585					590						
Val	Glu	His	Ala	Leu	Val	Asn	Gly	Glu	Arg	Glu	Arg	Asn	Ser	Thr	Thr				
		595					600					605							
Ser	Ile	Phe	Gln	Lys	Ile	Gly	Ser	Phe	Glu	Glu	Glu	Leu	Lys	Thr	Leu				
	610					615					620								
Leu	Pro	Lys	Glu	Val	Glu	Ser	Ala	Arg	Leu	Glu	Val	Glu	Asn	Gly	Asn				
625					630					635					640				
Pro	Ala	Ile	Pro	Asn	Arg	Ile	Lys	Glu	Cys	Arg	Ser	Tyr	Pro	Leu	Tyr				
				645					650					655					
Lys	Phe	Val	Arg	Glu	Glu	Leu	Gly	Thr	Ser	Leu	Leu	Thr	Gly	Glu	Lys				
			660					665					670						
Val	Lys	Ser	Pro	Gly	Glu	Glu	Phe	Asp	Lys	Val	Phe	Thr	Ala	Ile	Cys				
		675					680					685							
Ala	Gly	Lys	Leu	Ile	Asp	Pro	Leu	Leu	Glu	Cys	Leu	Lys	Glu	Trp	Asp				
	690					695					700								
Gly	Ala	Pro	Leu	Pro	Ile	Cys													
705					710														

<210> 84

<211> 713

<212> PRT

<213> Jatropha curcas

<400> 84

Met Ala Thr Ile Ile Gly Asn Gly His Gln Asn Gly Ser Leu Glu Gly
 1 5 10 15

Leu Cys Ile Thr Arg Asp Pro Leu Ser Trp Gly Val Ala Ala Glu Ser
 20 25 30

Met Lys Gly Ser His Leu Asp Glu Val Lys Lys Met Val Ser Glu Tyr
 35 40 45

Arg Lys Pro Leu Val Lys Leu Gly Gly Glu Thr Leu Thr Val Ala Gln
 50 55 60

Val Ala Ala Ile Ala Ser His Asp Ala Gly Val Lys Val Glu Leu Ala
 65 70 75 80

Glu Ser Ala Arg Ala Gly Val Lys Ala Ser Ser Asp Trp Val Met Asp
 85 90 95

Ser Met Asn Lys Gly Thr Asp Ser Tyr Gly Val Thr Thr Gly Phe Gly
 100 105 110

Ala Thr Ser His Arg Arg Thr Lys Gln Gly Ala Ala Leu Gln Arg Glu
 115 120 125

Leu Ile Arg Phe Leu Asn Ala Gly Ile Phe Gly Asn Gly Thr Glu Thr
 130 135 140

Cys His Thr Leu Pro His Ser Ala Thr Arg Ala Ala Met Leu Val Arg
 145 150 155 160

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

Glu Ala Ile Thr Lys Leu Leu Asn His Asn Ile Thr Pro Cys Leu Pro
 180 185 190

Leu Arg Gly Thr Ile Thr Ala Ser Gly Asp Leu Val Pro Leu Ser Tyr
 195 200 205

Ile Ala Gly Leu Leu Thr Gly Arg Pro Asn Ser Lys Ala Ile Gly Pro
 210 215 220

Ser Gly Glu Ser Leu Asp Ala Val Glu Ala Phe Arg Leu Ala Asp Ile
 225 230 235 240

Asp Ser Gly Phe Phe Glu Leu Gln Pro Lys Glu Gly Leu Ala Leu Val
 245 250 255

Asn Gly Thr Ala Val Gly Ser Gly Leu Ala Ser Met Val Leu Phe Glu
 260 265 270

Ala Asn Val Leu Ala Val Leu Ser Glu Ile Leu Ser Ala Ile Phe Ala
 275 280 285

Glu Val Met Asn Gly Lys Pro Glu Phe Thr Asp His Leu Thr His Lys
 290 295 300

Leu Lys His His Pro Gly Gln Ile Glu Ala Ala Ala Ile Met Glu His
 305 310 315 320

Ile Leu Asp Gly Ser Ser Tyr Ile Lys Ala Ala Lys Gln Leu His Glu
 325 330 335

Ile Asp Pro Leu Gln Lys Pro Lys Gln Asp Arg Tyr Ala Leu Arg Thr
 340 345 350

Ser Pro Gln Trp Leu Gly Pro Gln Ile Glu Val Ile Arg Phe Ser Thr
 355 360 365

Lys Ser Ile Glu Arg Glu Ile Asn Ser Val Asn Asp Asn Pro Leu Ile
 370 375 380

Asp Val Ser Arg Asn Lys Ala Leu His Gly Gly Asn Phe Gln Gly Thr
 385 390 395 400

Pro Ile Gly Val Ser Met Asp Asn Ala Arg Leu Ala Ile Ala Ser Ile
 405 410 415

Gly Lys Leu Met Phe Ala Gln Phe Ser Glu Leu Val Asn Asp Phe Tyr
 420 425 430

Asn Asn Gly Leu Pro Ser Asn Leu Ser Gly Gly Arg Asn Pro Ser Leu
 435 440 445

Asp Tyr Gly Phe Lys Gly Ala Glu Ile Ala Met Ala Ser Tyr Cys Ser
 450 455 460

Glu Leu Gln Phe Leu Ala Asn Pro Val Thr Ser His Val Gln Ser Ala
 465 470 475 480

Glu Gln His Asn Gln Asp Val Asn Ser Leu Gly Leu Ile Ser Ser Arg
 485 490 495

Lys Thr Gln Glu Ala Ile Asp Ile Leu Lys Leu Met Ser Ser Thr Phe
 500 505 510

Leu Val Ala Leu Cys Gln Ala Ile Asp Leu Arg His Leu Glu Glu Asn
 515 520 525

Leu Lys His Ala Val Lys Asn Thr Val Thr Gln Val Ala Lys Arg Val
 530 535 540

Leu Thr Thr Gly Ala Asn Gly Glu Leu His Pro Ser Arg Phe Cys Gly
 545 550 555 560

Lys Asp Leu Leu Lys Val Val Asp Arg Glu Gln Val Phe Ala Tyr Ile
 565 570 575

Asp Asp Pro Cys Ser Ala Thr Tyr Pro Leu Met Gln Lys Leu Arg Gln
 580 585 590

Val Leu Val Glu His Ala Leu Ala Asn Gly Glu Asn Glu Lys Asn Ala
 595 600 605

Ser Thr Ser Val Phe Gln Lys Ile Gly Ala Phe Glu Glu Glu Leu Lys
 610 615 620

Thr Leu Leu Pro Lys Glu Val Glu Ser Ala Arg Glu Ala Tyr Glu Ser
 625 630 635 640

Gly Ser Ala Ala Ile Gly Asn Lys Ile Lys Glu Cys Arg Ser Tyr Pro
 645 650 655

Leu Tyr Lys Phe Val Arg Glu Glu Leu Gly Ser Gly Leu Leu Thr Gly
 660 665 670

Glu Lys Val Arg Ser Pro Gly Glu Glu Phe Asp Lys Val Phe Thr Ala
 675 680 685

Met Cys Glu Gly Lys Ile Ile Asp Pro Met Met Glu Cys Leu Lys Glu
 690 695 700

Trp Asn Gly Ala Pro Leu Pro Ile Cys
 705 710

<210> 85

<211> 324

<212> PRT

<213> Arabidopsis thaliana

<400> 85

Met Leu Thr Asp Glu Arg Glu Val Val Cys Val Thr Gly Ala Ser Gly
 1 5 10 15

Cys Ile Gly Ser Trp Leu Val His Gln Leu Leu Leu Arg Gly Tyr Ser
 20 25 30

Val His Ala Thr Val Lys Asn Leu Gln Asp Glu Lys Glu Thr Lys His
 35 40 45

Leu Glu Gly Leu Glu Gly Ala Ala Thr Arg Leu His Leu Phe Glu Met
 50 55 60

Asp Leu Leu Gln Tyr Asp Thr Val Ser Ala Ala Ile Asn Gly Cys Ser
 65 70 75 80

Gly Val Phe His Leu Ala Ser Pro Cys Ile Val Asp Glu Val Gln Asp
 85 90 95

Pro Gln Lys Gln Leu Leu Asp Pro Ala Val Lys Gly Thr Ile Asn Val
 100 105 110

Leu Thr Ala Ala Lys Glu Ala Ser Val Lys Arg Val Val Val Thr Ser
 115 120 125

Ser Ile Ser Ala Ile Thr Pro Ser Pro Asn Trp Pro Ala Asp Lys Ile
 130 135 140

Lys Asn Glu Glu Cys Trp Ala Ala Glu Asp Tyr Cys Arg Gln Asn Gly
 145 150 155 160

Leu Trp Tyr Pro Leu Ser Lys Thr Leu Ala Glu Lys Ala Ala Trp Glu
 165 170 175

Phe Ala Glu Glu Lys Gly Leu Asp Val Val Val Val Asn Pro Gly Thr
 180 185 190

Val Met Gly Pro Val Ile Pro Pro Ser Leu Asn Ala Ser Met His Met
 195 200 205

Leu Leu Arg Leu Leu Gln Gly Cys Thr Glu Thr Tyr Glu Asn Phe Phe

```

210                215                220

Met Gly Ser Val His Phe Lys Asp Val Ala Leu Ala His Ile Leu Val
225                230                235                240

Tyr Glu Asp Pro Tyr Ser Lys Gly Arg His Leu Cys Val Glu Ala Ile
                245                250                255

Ser His Tyr Gly Asp Phe Val Ala Lys Val Ala Glu Leu Tyr Pro Asn
                260                265                270

Tyr Asn Val Pro Lys Leu Pro Arg Glu Thr Gln Pro Gly Leu Leu Arg
                275                280                285

Asp Lys Asn Ala Ser Lys Lys Leu Ile Asp Leu Gly Leu Lys Phe Ile
                290                295                300

Ser Met Glu Glu Ile Ile Lys Glu Gly Val Glu Ser Leu Lys Ser Lys
305                310                315                320

Gly Phe Ile Ser

```

```

<210> 86
<211> 323
<212> PRT
<213> Arabidopsis lyrata

```

```

<400> 86
Met Ser Thr Glu Arg Glu Val Val Cys Val Thr Gly Ala Ser Gly Cys
1                5                10                15

Ile Gly Ser Trp Leu Val His Leu Leu Leu His Arg Gly Tyr Ser Val
                20                25                30

His Ala Thr Val Lys Asn Leu Gln Asp Glu Lys Glu Thr Lys His Leu
                35                40                45

Glu Ala Leu Glu Gly Ala Ala Thr Arg Leu His Leu Phe Glu Met Asp
50                55                60

Leu Leu Gln Tyr Asp Thr Val Ser Ala Ala Val Asn Gly Cys Ser Gly
65                70                75                80

Val Phe His Leu Ala Ser Pro Cys Ile Val Asp Glu Val Gln Asp Pro
                85                90                95

```

Gln Lys Gln Leu Leu Asp Pro Ala Val Lys Gly Thr Ile Asn Val Leu
 100 105 110

Thr Ala Ala Lys Glu Ala Gly Val Lys Arg Val Val Val Thr Ser Ser
 115 120 125

Ile Ser Ala Ile Thr Pro Ser Pro Asn Trp Pro Ala Asp Lys Ile Lys
 130 135 140

Asn Glu Glu Cys Trp Ala Asp Gln Asp Tyr Cys Lys Gln Asn Gly Leu
 145 150 155 160

Trp Tyr Pro Leu Ser Lys Thr Leu Ala Glu Lys Ala Ala Trp Glu Phe
 165 170 175

Ala Glu Gln Lys Gly Leu Asp Val Val Val Val Asn Pro Gly Thr Val
 180 185 190

Met Gly Pro Val Ile Pro Pro Ser Ile Asn Ala Ser Met Leu Met Leu
 195 200 205

Leu Arg Leu Leu Gln Gly Cys Thr Glu Thr Tyr Glu Asn Phe Phe Met
 210 215 220

Gly Ser Val His Phe Lys Asp Val Ala Leu Ala His Ile Leu Val Tyr
 225 230 235 240

Glu Asn Pro Ser Ala Lys Gly Arg His Leu Cys Val Glu Ala Ile Ser
 245 250 255

His Tyr Gly Asp Phe Val Ala Lys Val Ala Glu Leu Tyr Pro Asn Tyr
 260 265 270

Ser Val Pro Lys Leu Pro Arg Glu Thr Gln Leu Gly Leu Leu Arg Ala
 275 280 285

Lys Asn Ala Ala Lys Lys Leu Met Glu Leu Gly Leu Glu Phe Ser Ser
 290 295 300

Met Glu Asp Ile Ile Lys Glu Gly Val Glu Ser Leu Lys Ser Lys Gly
 305 310 315 320

Phe Ile Ser

<210> 87

<211> 329

<212> PRT

<213> Camellia oleifera

<400> 87

Met Ser Ser Asn Thr Lys Ala Gly Gly Asp Gly Gln Val Val Cys Val
 1 5 10 15

Thr Gly Gly Ser Gly Phe Ile Gly Ser Trp Leu Val Arg Leu Leu Leu
 20 25 30

Asp Arg Gly Tyr Thr Val His Ala Thr Val Lys Asp Leu Lys Asp Glu
 35 40 45

Lys Glu Thr Lys His Leu Glu Ala Leu Glu Gly Ala Glu Ser Arg Leu
 50 55 60

Arg Leu Phe Gln Ile Asp Leu Leu Asp Tyr Asp Ser Ile Val Ala Ala
 65 70 75 80

Val Thr Gly Ser Ser Gly Val Phe His Leu Ala Ser Pro Cys Ile Val
 85 90 95

Asp Gln Val Lys Asp Pro Glu Arg Glu Leu Leu Glu Pro Ala Ile Lys
 100 105 110

Gly Thr Leu Asn Val Leu Thr Ala Ala Lys Glu Leu Gly Val Arg Arg
 115 120 125

Val Val Val Thr Ser Ser Asn Thr Ala Ile Thr Pro Ser Pro Asn Trp
 130 135 140

Pro Ala Asp Lys Val Lys Asn Glu Asp Cys Trp Thr Asp Val Glu Tyr
 145 150 155 160

Cys Lys Gln Asn Gly Leu Trp Tyr Pro Leu Ser Lys Thr Leu Ala Glu
 165 170 175

Lys Ala Ala Trp Glu Phe Ala Lys Glu Lys Gly Leu Asp Val Val Val
 180 185 190

Val Asn Pro Gly Thr Val Met Gly Pro Ile Ile Pro Pro Ala Leu Asn
 195 200 205

Ala Ser Met Leu Met Leu Leu Arg Phe Leu Gln Gly Cys Thr Glu Ile

210 215 220
 Tyr Glu Asn Phe Phe Met Gly Pro Val His Val Lys Asp Val Ala Leu
 225 230 235 240
 Ala His Ile Leu Val Tyr Glu Asn Thr Ser Ala Thr Gly Arg His Leu
 245 250 255
 Cys Val Glu Ala Ile Ser His Tyr Gly Asp Phe Thr Ala Met Val Ala
 260 265 270
 Glu Leu Tyr Pro Glu Tyr Asn Val Pro Arg Leu Pro Lys Asp Thr Gln
 275 280 285
 Pro Gly Leu Leu Arg Thr Lys Asp Gly Ser Lys Lys Leu Met Asp Leu
 290 295 300
 Gly Phe Gln Phe Ile Pro Met Glu Gln Ile Ile Lys Glu Thr Val Glu
 305 310 315 320
 Ser Leu Lys Ser Lys Gly Tyr Ile Ser
 325
 <210> 88
 <211> 334
 <212> PRT
 <213> Ricinus communis
 <400> 88
 Met Ala Thr Gln Asn Lys Lys Glu Ala Val Cys Val Thr Gly Ala Asn
 1 5 10 15
 Gly Phe Ile Gly Ser Trp Leu Ile Gln Thr Leu Leu Gln His Gly Tyr
 20 25 30
 Thr Thr Ile His Ala Ser Ile Tyr Pro Ala Ser Asp Pro Ser His Leu
 35 40 45
 Phe His Leu Ile Ser Ser Ser Ser His Gly Asp Ile Ile Asn Leu Lys
 50 55 60
 Leu Tyr Glu Ala Asp Leu Leu Asp Tyr Asp Ala Ile Cys Lys Ala Val
 65 70 75 80
 Glu Gly Cys Gln Gly Val Phe His Val Ala Ser Pro Cys Thr Leu Glu
 85 90 95

Glu Pro Lys Asp Pro Glu Lys Glu Leu Val Leu Pro Ala Val Gln Gly
 100 105 110

Thr Ile Asn Val Leu Glu Ala Ala Arg Lys Phe Lys Val Arg Arg Val
 115 120 125

Val Leu Thr Ser Ser Ile Ser Ala Leu Val Pro Asn Pro Asn Trp Pro
 130 135 140

Ala Gly Lys Val Phe Asp Glu Ser Ser Trp Thr Asp Leu Asp Tyr Cys
 145 150 155 160

Lys Ser Arg Gln Lys Trp Tyr Pro Val Ser Lys Ser Leu Ala Glu Lys
 165 170 175

Ala Ala Trp Glu Phe Ala Glu Lys His Gly Met Asp Val Val Ala Ile
 180 185 190

His Pro Ser Thr Cys Ile Gly Pro Leu Leu Gln Pro Ser Leu Asn Ala
 195 200 205

Ser Ser Ala Val Leu Gln Gln Leu Leu Glu Gly Ser Lys Asp Thr Gln
 210 215 220

Glu Tyr His Trp Leu Gly Ala Val His Val Lys Asp Val Ala Lys Ala
 225 230 235 240

Gln Val Leu Leu Phe Glu Ala Pro Ser Ala Ser Gly Arg Tyr Leu Cys
 245 250 255

Thr Asn Gly Ile Tyr Gln Phe Gly Asp Phe Ala Asp Arg Val Ser Lys
 260 265 270

Leu Phe Pro Glu Phe Pro Val His Ser Phe Ile Gly Glu Thr Gln Pro
 275 280 285

Gly Leu Thr Thr Cys Lys Asp Ala Ala Lys Arg Leu Ile Glu Leu Gly
 290 295 300

Leu Val Phe Thr Pro Val Glu Asp Ala Val Gly Glu Ser Val Glu Ser
 305 310 315 320

Leu Gln Ala Lys Gly Phe Leu Lys His Lys Thr Ser Glu Ser
 325 330

<210> 89

<211> 318

<212> PRT

<213> Arabidopsis thaliana

<400> 89

Met Ala Lys Glu Thr Val Cys Val Thr Gly Ala Asn Gly Phe Ile Gly
 1 5 10 15

Ser Trp Ile Ile Arg Thr Leu Ile Glu Lys Gly Tyr Thr Lys Ile His
 20 25 30

Ala Ser Ile Tyr Pro Gly Ser Asp Pro Thr His Leu Leu Gln Leu Pro
 35 40 45

Gly Ser Asp Ser Lys Ile Lys Ile Phe Glu Ala Asp Leu Leu Asp Ser
 50 55 60

Asp Ala Ile Ser Arg Ala Ile Asp Gly Cys Ala Gly Val Phe His Val
 65 70 75 80

Ala Ser Pro Cys Thr Leu Asp Pro Pro Val Asp Pro Glu Lys Glu Leu
 85 90 95

Val Glu Pro Ala Val Lys Gly Thr Ile Asn Val Leu Glu Ala Ala Lys
 100 105 110

Arg Phe Asn Val Arg Arg Val Val Ile Thr Ser Ser Ile Ser Ala Leu
 115 120 125

Val Pro Asn Pro Asn Trp Pro Glu Lys Val Pro Val Asp Glu Ser Ser
 130 135 140

Trp Ser Asp Leu Asp Phe Cys Lys Ser Arg Gln Lys Trp Tyr Pro Ile
 145 150 155 160

Ser Lys Thr Leu Ala Glu Lys Ala Ala Trp Glu Phe Ser Glu Lys His
 165 170 175

Gly Thr Asn Ile Val Thr Ile His Pro Ser Thr Cys Leu Gly Pro Leu
 180 185 190

Leu Gln Pro Asn Leu Asn Ala Ser Cys Ala Val Leu Leu Gln Leu Leu
 195 200 205

Gln Gly Ser Thr Glu Thr Gln Glu His His Trp Leu Gly Val Val His

210 215 220
 Val Lys Asp Val Ala Lys Gly His Val Met Leu Phe Glu Thr Pro Asp
 225 230 235 240
 Ala Ser Gly Arg Phe Leu Cys Thr Asn Gly Ile Tyr Gln Phe Ser Glu
 245 250 255
 Phe Ala Ala Leu Val Ser Lys Leu Phe Pro Glu Phe Ala Val His Lys
 260 265 270
 Phe Asp Lys Glu Thr Gln Pro Gly Leu Thr Ser Cys Asn Asp Ala Ala
 275 280 285
 Lys Arg Leu Ile Glu Leu Gly Leu Val Phe Thr Ala Val Glu Asp Ala
 290 295 300
 Val Lys Glu Thr Val Gln Ser Leu Arg Asp Lys Gly Phe Leu
 305 310 315

 <210> 90
 <211> 529
 <212> PRT
 <213> Eucalyptus globules

 <400> 90
 Met Asp Ile Phe Tyr Phe Tyr Ser Gln Leu Gln Ser Leu Val Gln Thr
 1 5 10 15

 Gln Leu Gln Gln Ser Pro Met Thr Leu Leu Leu Ser Val Val Pro Leu
 20 25 30

 Leu Leu Phe Leu Gly Leu Val Ala Arg Leu Arg Arg Lys Pro Pro Phe
 35 40 45

 Pro Pro Gly Pro Arg Gly Leu Pro Val Ile Gly Asn Met Leu Met Met
 50 55 60

 Gly Glu Leu Thr His Arg Gly Leu Ala Ser Leu Ala Lys Lys Tyr Gly
 65 70 75 80

 Gly Ile Phe His Leu Arg Met Gly Phe Leu His Met Val Ala Val Ser
 85 90 95

 Ser Pro Asp Val Ala Arg Gln Val Leu Gln Val His Asp Gly Ile Phe
 100 105 110

Ser Asn Arg Pro Ala Thr Ile Ala Ile Ser Tyr Leu Thr Tyr Asp Arg
 115 120 125

Ala Asp Met Ala Phe Ala His Tyr Gly Pro Phe Trp Arg Gln Met Arg
 130 135 140

Lys Leu Cys Val Met Lys Leu Phe Ser Arg Lys Arg Ala Glu Ser Trp
 145 150 155 160

Glu Ser Val Arg Asp Glu Val Asp Thr Met Val Arg Thr Val Ala Gly
 165 170 175

Ser Glu Gly Thr Ala Val Asn Ile Gly Glu Leu Val Phe Glu Leu Thr
 180 185 190

Arg Asp Ile Ile Tyr Arg Ala Ala Phe Gly Thr Ser Ser Thr Glu Gly
 195 200 205

Gln Asp Glu Phe Ile Ser Ile Leu Gln Glu Phe Ser Lys Leu Phe Gly
 210 215 220

Ala Phe Asn Ile Ala Asp Phe Ile Pro Tyr Leu Ser Trp Ile Asp Pro
 225 230 235 240

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

Phe Ile Asp His Ile Ile Asp Asp His Met Asp Lys Lys Arg Asn Lys
 260 265 270

Thr Ser Ser Gly Gly Gly Asp Gln Glu Val Asp Thr Asp Met Val Asp
 275 280 285

Asp Leu Leu Ala Phe Tyr Ser Asp Glu Ala Lys Val Asn Glu Ser Asp
 290 295 300

Asp Leu Gln Asn Ser Ile Arg Leu Thr Arg Asp Asn Ile Lys Ala Ile
 305 310 315 320

Ile Met Asp Val Met Phe Gly Gly Thr Glu Thr Val Ala Ser Ala Ile
 325 330 335

Glu Trp Ala Met Ala Glu Leu Met Arg Ser Pro Glu Asp Leu Lys Lys
 340 345 350

Val Gln Gln Glu Leu Ala Asp Val Val Gly Leu Asp Arg Arg Val Glu
 355 360 365

Glu Ser Asp Phe Glu Lys Leu Thr Tyr Leu Lys Cys Cys Leu Lys Glu
 370 375 380

Thr Leu Arg Leu His Pro Pro Ile Pro Leu Leu Leu His Glu Thr Ala
 385 390 395 400

Glu Asp Ala Val Ile Ser Gly Tyr Arg Ile Pro Ala Arg Ser Arg Val
 405 410 415

Met Ile Asn Ala Trp Ala Ile Gly Arg Asp Pro Gly Ser Trp Thr Glu
 420 425 430

Pro Asp Lys Phe Lys Pro Ser Arg Phe Leu Glu Ser Gly Met Pro Asp
 435 440 445

Tyr Lys Gly Ser Asn Phe Glu Phe Ile Pro Phe Gly Ser Gly Arg Arg
 450 455 460

Ser Cys Pro Gly Met Gln Leu Gly Leu Tyr Ala Leu Asp Met Ala Val
 465 470 475 480

Ala His Leu Leu His Cys Phe Thr Trp Glu Leu Pro Asp Gly Met Lys
 485 490 495

Pro Ser Glu Met Asp Met Gly Asp Val Phe Gly Leu Thr Ala Pro Arg
 500 505 510

Ser Thr Arg Leu Val Ala Val Pro Thr Pro Arg Leu Val Gly Ala Leu
 515 520 525

Tyr

<210> 91

<211> 443

<212> PRT

<213> Populus trichocarpa

<400> 91

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

Ser Pro Glu Val Ala Arg Gln Val Leu Gln Val Gln Asp Asn Met Phe

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

260	265	270
Ala Asp Val Val Gly Leu Glu Arg Arg Val Glu Glu Ser Asp Phe Asp		
275	280	285
Lys Leu Thr Phe Phe Lys Cys Thr Leu Lys Glu Thr Leu Arg Leu His		
290	295	300
Pro Pro Ile Pro Leu Leu Leu His Glu Thr Ser Glu Asp Ala Glu Val		
305	310	315
Ala Gly Tyr Tyr Val Pro Lys Lys Thr Arg Val Met Ile Asn Ala Tyr		
325	330	335
Ala Ile Gly Arg Asp Lys Asn Ser Trp Glu Asp Pro Asp Ser Phe Lys		
340	345	350
Pro Ser Arg Phe Leu Glu Pro Gly Val Pro Asp Phe Lys Gly Asn His		
355	360	365
Phe Glu Phe Ile Pro Phe Gly Ser Gly Arg Arg Ser Cys Pro Gly Met		
370	375	380
Gln Leu Gly Leu Tyr Ala Leu Asp Leu Ala Val Ala His Leu Leu His		
385	390	395
Cys Phe Thr Trp Glu Leu Pro Asp Gly Met Lys Pro Ser Glu Leu Asp		
405	410	415
Met Thr Asp Met Phe Gly Leu Thr Ala Pro Arg Ala Thr Arg Leu Val		
420	425	430
Ala Val Pro Arg Lys Arg Val Val Cys Pro Leu		
435	440	
<210> 92		
<211> 365		
<212> PRT		
<213> Gossypium hirsutum		
<400> 92		
Met Gly Ser Thr Gly Glu Thr Gln Met Thr Pro Thr Gln Val Ser Asp		
1	5	10
Glu Glu Ala Asn Leu Phe Ala Met Gln Leu Thr Ser Ala Ser Val Leu		
20	25	30

Pro Met Val Leu Lys Ser Ala Ile Glu Leu Asp Leu Leu Glu Ile Met
 35 40 45

Ala Lys Ala Gly Pro Gly Ala Phe Leu Ser Pro Lys Glu Leu Ala Ser
 50 55 60

Gln Leu Pro Thr Ser Asn Pro Asp Ala Pro Val Met Leu Asp Arg Ile
 65 70 75 80

Leu Arg Leu Leu Ala Thr Tyr Ser Ile Leu Thr Cys Ser Leu Arg Thr
 85 90 95

Leu Pro Asp Gly Lys Val Glu Arg Leu Tyr Gly Leu Gly Pro Val Cys
 100 105 110

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

Leu Met Asn Gln Asp Lys Val Leu Met Glu Ser Trp Tyr Tyr Leu Lys
 130 135 140

Asp Ala Val Leu Glu Gly Gly Ile Pro Phe Asn Lys Val Tyr Gly Met
 145 150 155 160

Thr Ala Phe Glu Tyr His Gly Thr Asp Pro Arg Phe Asn Lys Val Phe
 165 170 175

Asn Arg Gly Met Ser Asp His Ser Thr Ile Thr Met Lys Lys Ile Leu
 180 185 190

Glu Thr Tyr Asp Gly Phe Glu Gly Leu Lys Thr Leu Val Asp Val Gly
 195 200 205

Gly Gly Thr Gly Ala Thr Leu Asn Met Ile Val Thr Lys His Pro Ser
 210 215 220

Ile Lys Gly Ile Asn Phe Asp Leu Pro His Val Ile Glu Asp Ala Pro
 225 230 235 240

Ala Tyr Pro Gly Val Glu His Val Gly Gly Asp Met Phe Glu Ser Val
 245 250 255

Pro Lys Gly Asp Ala Ile Phe Met Lys Trp Ile Cys His Asp Trp Ser
 260 265 270

Asp Glu His Cys Ser Lys Phe Leu Lys Lys Cys Tyr Glu Ala Leu Pro
 275 280 285

Asp Ser Gly Lys Val Ile Val Ala Glu Cys Ile Leu Pro Asp Tyr Pro
 290 295 300

Asp Pro Ser Leu Ala Thr Lys Leu Val Val His Ile Asp Cys Ile Met
 305 310 315 320

Leu Ala His Asn Pro Gly Gly Lys Glu Arg Thr Glu Lys Glu Phe Glu
 325 330 335

Ala Leu Ala Arg Ser Ala Gly Phe Gln Gly Phe Gln Val Lys Cys Cys
 340 345 350

Ala Phe Gly Thr Tyr Ile Met Glu Phe Val Lys Arg Val
 355 360 365

<210> 93

<211> 366

<212> PRT

<213> Eucalyptus camaldulensis

<400> 93

Met Gly Ser Thr Gly Ser Glu Thr Gln Met Thr Pro Thr Gln Val Ser
 1 5 10 15

Asp Glu Glu Ala Asn Leu Phe Ala Met Gln Leu Ala Ser Ala Ser Val
 20 25 30

Leu Pro Met Val Leu Lys Ala Ala Ile Glu Leu Asp Leu Leu Glu Ile
 35 40 45

Met Ala Lys Ala Gly Pro Gly Ala Phe Leu Ser Pro Gly Glu Val Ala
 50 55 60

Ala Gln Leu Pro Thr Gln Asn Pro Glu Ala Pro Val Met Leu Asp Arg
 65 70 75 80

Ile Phe Arg Leu Leu Ala Ser Tyr Ser Val Leu Thr Cys Thr Leu Arg
 85 90 95

Asp Leu Pro Asp Gly Lys Val Glu Arg Leu Tyr Gly Leu Ala Pro Val
 100 105 110

Cys Lys Phe Leu Val Lys Asn Glu Asp Gly Val Ser Ile Ala Ala Leu

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