



(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)



1-0028548

(51)⁷ C12N 9/08

(13) B

(21) 1-2015-00873

(22) 15/08/2013

(86) PCT/US2013/055199 15/08/2013

(87) WO 2014/028773 A2 20/02/2014

(30) 61/683,913 16/08/2012 US

(45) 25/06/2021 399

(43) 25/08/2015 329A

(73) BANGLADESH JUTE RESEARCH INSTITUTE (BD)

Manik Mia Avenue, Dhaka, 1207 (BD) Bangladesh, People's Republic

(72) ISLAM, Mohammed, Shahidul (BD); HOSEN, Mohammed, Mosaddeque (BD);
HAQUE, Mohammed, Samiul (BD); ALAM, Mohammed, Monjurul (BD).

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

(54) POLYNUCLEOTIT ĐƯỢC PHÂN LẬP MÃ HÓA CHO LIGNIN PEROXIDAZA
TỪ NẤM MACROPHOMINA PHASEOLINA VÀ CẤU TRÚC BIỂU HIỆN CHỨA
POLYNUCLEOTIT NÀY

(57) Sáng chế đề cập đến lĩnh vực gây giống cây trồng và tính kháng bệnh liên quan đến sự thoái biến lignin của cây trồng bởi *Macrophomina phaseolina* ("*M. phaseolina*"). Sáng chế đề cập đến polynucleotit được phân lập mã hóa cho enzym làm thoái biến lignin, được tạo ra bởi *Macrophomina phaseolina*, gồm có và có chứa trình tự nucleotit nêu trong SEQ ID NO. 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 35, 38, 41, 44, 47, 50, 53, 56, 59, 62, 65, 68, 71, 74, 77, 80, 83, 86, 89, 92, 95, 98, 101, 104, 107, 110, 1, 4, 7, 10, 13, 16, 19, 22, 25, 28, 31, 34, 37, 40, 43, 46, 49, 52, 55, 58, 61, 64, 67, 70, 73, 76, 79, 82, 85, 88, 91, 94, 97, 100, 103, 106 và 109 hoặc trình tự bổ sung của các trình tự này. Sáng chế cũng đề cập đến polypeptit được phân lập được mã hóa bởi trình tự polynucleotit nêu trong SEQ ID NO. 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 39, 42, 45, 48, 51, 54, 57, 60, 63, 66, 69, 72, 75, 78, 81, 84, 87, 90, 93, 96, 99, 102, 105, 108 và 111; cấu trúc gen tái tổ hợp có chứa polynucleotit; nấm biến nạp và chuyển gen có chứa cấu trúc gen tái tổ hợp, có khả năng sản xuất enzym làm thoái biến lignin được tăng cường. Sáng chế đề xuất polypeptit có hoạt tính oligomeraza, ví dụ, enzym chuyển hóa oligomeraza hòa tan khó phân giải thành đường có thể lên men được trong quy trình sacarit hóa của sinh khối.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến protein/enzym làm thoái biến lignin và gen của *M. phaseolina*. Cụ thể hơn, sáng chế đề cập đến trình tự axit nucleotit của các protein này và tạo điều kiện cho việc làm giảm hàm lượng lignin của nguyên liệu lignoxenluloza. Sáng chế cũng đề cập đến phương pháp chọn lọc và phân lập enzym từ *M. phaseolina* mà có khả năng làm thoái biến lignin, quy trình tách dòng mảnh gen từ *M. phaseolina*, và phương pháp sử dụng sản phẩm enzym của mảnh gen này. Cụ thể là, nó có thể được sử dụng trong việc sản xuất hoặc xử lý bột giấy, bột gỗ, metanol và vải dệt.

Tình trạng kỹ thuật của sáng chế

Lignin là polyme hữu cơ phong phú thứ hai, chỉ sau xenluloza. Nó liên kết với sợi xenluloza tạo thành xấp xỉ một phần tư khối lượng của gỗ khô. Polyme thơm này chống lại sự thoái biến. Lignin được kết hợp cộng hóa trị với hemixenluloza trong thành tế bào thông qua một số loại liên kết. Trong số các liên kết này, cấu trúc móng lignin phong phú nhất là β -aryl ete, chiếm xấp xỉ 40% liên kết giữa các phenylpropan (Higuchi T 1990. Lignin biochemistry: biosynthesis and biodegradation. Wood Sci. Technol. 24: 23-63). Ngoài ra, lignin không mang đơn vị lặp lại bất kỳ như các polyme sinh học khác chẳng hạn như xenluloza, protein, tinh bột v.v.. Liên kết ete thường là phổ biến, tạo ra nhiều liên kết với nhiều gốc thơm. Đặc điểm cấu trúc này gây ra sự kiềm chế đối với sự thoái biến của lignin. Một phần của cấu trúc thông thường được minh họa trong Fig. 1 và cấu trúc arylglyxerol β -aryl ete có màu đỏ trong hình vẽ là liên kết quan trọng nhất về mặt định lượng, tạo thành ít nhất 40% polyme. Nhưng, sự thoái biến sinh học của lignin là điều kiện tiên quyết để xử lý nhiên liệu sinh học, và bột gỗ từ nguyên liệu thô thực vật. Việc cải thiện đối với sự thoái biến lignin sẽ dẫn đến sản lượng của việc xử lý nhiên liệu sinh học tăng lên hoặc hiệu quả hơn.

Lignin là thành phần của lignoxenluloza, mà phải được làm thoái biến để cho phép sử dụng nguyên liệu xenluloza hiệu quả cho sự sacarit hóa, sự sản xuất giấy, sự sản xuất nhiên liệu sinh học hoặc sự nâng cao chất lượng của cỏ khô. Sự thoái biến

lignoxenuloza là quy trình đa enzym bao gồm cả enzym thủy phân và enzym oxy hóa. Nhìn chung, lignin được tạo thành bởi ba khối xây dựng cơ bản cụ thể là, rượu *p*-coumaryl, rượu coniferyl và rượu sinapyl. Ngoài ra, lignin cỏ và cây hai lá mầm cũng có chứa lượng lớn axit phenolic chẳng hạn như axit *p*-coumaric và axit ferulic, mà được este hóa thành nhóm rượu.

Lignin peroxidaza (LiP), mangan peroxidaza (MnP) và lacaza là ba họ enzym tham gia vào sự thoái biến sinh học của lignin. LiP oxy hóa cấu trúc móng lignin không phenol bằng cách rút ra một electron và tạo ra gốc cation (Kirk TK, Tien M, Kersten PJ, Mozuch MD and Kalyanaraman B 1986. Ligninase of *Phanerochaete chrysosporium*. Mechanism of its degradation of the non-phenolic arylglycerol β -aryl ether substructure of lignin. Biochem J 236: 279-287; Kirk TK and Farrell RL 1987. Enzymatic “combustion”: The microbial degradation of lignin. Annu Rev Microbiol 41: 465-505) và MnP oxy hóa vòng phenol thành gốc phenoxyl dẫn đến sự phân hủy của Lignin (Gold MH, Wariishi H and Valli K 1989. Extracellular peroxidases involved in lignin degradation by the white-rot basidiomycete *Phanerochaete chrysosporium*. In: Biocatalysis in Agricultural Biotechnology, pp. 127-140. Edited by Whitaker JR and Sonnet PE. American Chemical Society, Washington, DC). Trong khi đó lacaza có khả năng oxy hóa cả gốc phenol và gốc không phenol của lignin nhưng khả năng oxy hóa gốc không phenol phụ thuộc vào sự đồng có mặt của các chất nền lacaza sơ cấp (Bourbonnais R and Paice MG 1990. Oxidation of non-phenolic substrates. An expanded role for laccase in lignin biodegradation. FEBS Lett 267(1): 99-102). Fig. 2 thể hiện hợp chất kiểu dime đại diện cho cấu trúc arylglyxerol β -aryl ete lignin chính đã trải qua sự phân cắt $C\alpha$ - $C\beta$ khi oxy hóa bằng LiP (Kirk *et al.* 1986. Ligninase of *Phanerochaete chrysosporium*. Mechanism of its degradation of the non-phenolic arylglycerol β -aryl ether substructure of lignin. Biochem J 236: 279-287).

Nghiên cứu về sự thoái biến sinh học lignin đã tăng lên mạnh mẽ trong những năm gần đây. Đặc biệt là kể từ khi phát hiện ra nấm làm thoái biến lignoxenuloza đầu tiên, hệ gen *Phanerochaete chrysosporium* được đọc trình tự do quan tâm đến sự thoái biến sinh học của lignin (Martínez *et al.* 2004. Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. Nature Biotechnol 22(6): 695-700). Có lẽ đối tượng tốt nhất được nghiên cứu là sự thoái biến của lignin trong lignoxenuloza bởi nấm thối trắng, *Phanerochaete chrysosporium* (Aust SD 1995.

Mechanisms of Degradation by White Rot Fungi. Environ Health Perspect 103:59-61; Leisola MSA, Ulmer D and Fiechter A 1984. Factors affecting lignin degradation in lignocellulose by *Phanerochaete chrysosporium*. Arch Microbiol 137: 171-175; Ulmer D, Leisola M, Puhakka J and Fiechter A 1983. *Phanerochaete chrysosporium*: Growth pattern and lignin degradation. Appl Microbiol Biotechnol 18: 153-157; Chua MGS, Chen CL, Chang HM and Kirk TK 1982. ¹³CNMR spectroscopic study of lignin degraded by *Phanerochaete chrysosporium*. I. New structures. Holzforschung 36: 165-172). Nấm thối trắng sản xuất ra nhiều loại enzym phân giải lignin ngoại bào, bao gồm lignin peroxidaza phụ thuộc hem, mangan peroxidaza, và peroxidaza đa tác dụng và lacaza phụ thuộc đồng (Sanchez C 2009. Lignocellulosic residues: biodegradation and bioconversion by fungi. Biotechnol Adv 27: 185-194; Ten Have R, Teunissen PJM 2001. Oxidative mechanisms involved in lignin degradation by white-rot fungi. Chem Rev 101: 3397-3413). Với sự hỗ trợ của các enzym peroxidaza và lacaza ngoại bào, nấm thối trắng làm thoái biến lignin (Timothy DHB, Ahmad M, Hardiman EM and Rahmanpour R 2001. Pathways for degradation of lignin in bacteria and fungi. Nat. Prod. Rep 28: 1883-1896).

Sự quan tâm ngày càng tăng đối với việc khai thác sinh khối thực vật làm nguồn tài nguyên tái tạo được đã thúc đẩy nghiên cứu về sự thoái biến bằng vi khuẩn của lignoxenluloza. Lượng lớn sinh khối (chủ yếu là nguyên liệu lignoxenluloza) được tạo ra từ lâm nghiệp, nông nghiệp và công nghiệp thực phẩm mà không được sử dụng trong các quy trình sản phẩm phụ. Để sử dụng hiệu quả sinh khối thực vật làm nguồn tài nguyên tái tạo được, phương pháp thay thế của sự thoái biến của lignoxenluloza cần phải được khai thác. Điều mong muốn là xác định nấm mới có khả năng làm thoái biến lignin để sử dụng trong việc sản xuất sản phẩm xenluloza từ nguyên liệu lignoxenluloza.

Để có hiệu quả, sự thoái biến của lignin đòi hỏi một số loại enzym hoạt động hợp tác. Cần ít nhất hai loại enzym để làm thoái biến lignin: Lignin peroxidaza oxy hóa cấu trúc móng lignin không phenol và lacaza oxy hóa cả cấu trúc móng lignin phenol và không phenol do đó chúng có thể làm thoái biến theo cách hợp tác. Để thu được và thương mại hóa các enzym kể trên, cần nguồn cung cấp thích hợp của các protein làm thoái biến lignin khác nhau. Do đó, điều mong muốn đối với ngành công nghiệp này là xác định hoàn toàn các gen làm thoái biến lignin và protein được mã hóa

của chúng, do đó sử dụng thông tin di truyền để làm thoái biến lignin từ nguyên liệu lignoxenuloza.

Bản chất kỹ thuật của sáng chế

Đã bất ngờ phát hiện ra rằng nấm *M. phaseolina*, mà không có mối liên quan gần gũi đối với *Trichoderma reesi* hoặc *Phanerochaete chrysosporium*, có khả năng sản xuất ra enzym làm thoái biến lignin.

Theo đó, ngoài những cái khác, sáng chế đề cập đến việc xác định và việc sử dụng tương ứng của enzym làm thoái biến lignin mà có nguồn gốc từ *M. phaseolina*. Sáng chế cũng đề cập đến việc sử dụng nấm *M. phaseolina* trong việc làm thoái biến hàm lượng lignin của nguyên liệu xenuloza.

Mục đích chính theo sáng chế là bộc lộ bộ trình tự nucleotit mã hóa cho lignin peroxidaza (SEQ ID NO. 1, 2, 4, 5, 7, 8, và/hoặc hỗn hợp/dạng kết hợp bất kỳ của chúng), của nấm *M. phaseolina*. Sáng chế cũng bộc lộ bộ trình tự nucleotit mã hóa cho đa đồng oxidaza (SEQ ID NO. 46, 47, 49, 50, 52, 53, 55, 56, 58, 59, 61, 62, 64, 65, 67, 68, 70, 71, 73, 74, 76, 77, 79, 80, 82, 83, 85, 86, 88, 89, 91, 92, 94, 95, 97, 98, 100, 101, 103, 104, 106, 107, 109, 110, và/hoặc hỗn hợp/dạng kết hợp bất kỳ của chúng) của nấm *M. phaseolina*.

Được ưu tiên là polynucleotit được phân lập theo sáng chế có chứa và/hoặc bao gồm trình tự axit nucleic được chọn từ nhóm có chứa và/hoặc gồm có SEQ ID NO. 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 35, 38, 41, 44, 47, 50, 53, 56, 59, 62, 65, 68, 71, 74, 77, 80, 83, 86, 89, 92, 95, 98, 101, 104, 107, 110, 1, 4, 7, 10, 13, 16, 19, 22, 25, 28, 31, 34, 37, 40, 43, 46, 49, 52, 55, 58, 61, 64, 67, 70, 73, 76, 79, 82, 85, 88, 91, 94, 97, 100, 103, 106 và 109, và/hoặc hỗn hợp/dạng kết hợp bất kỳ của chúng, mà mã hóa cho polypeptit được chọn từ nhóm có chứa và/hoặc gồm có SEQ ID NO. 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 39, 42, 45, 48, 51, 54, 57, 60, 63, 66, 69, 72, 75, 78, 81, 84, 87, 90, 93, 96, 99, 102, 105, 108 và 111, và/hoặc hỗn hợp/dạng kết hợp bất kỳ của chúng. Sáng chế cũng đề cập đến polynucleotit được phân lập gồm có và/hoặc chứa trình tự bổ sung của trình tự nucleotit được mô tả ở trên.

Mục đích khác của sáng chế là đề xuất thông tin sinh học phân tử và di truyền của gen và enzym nêu trong mục đích chính để sử dụng cho sự điều hòa và sự chuyển

hóa của sự thoái biến lignin để sản xuất sản phẩm có giá trị từ nguyên liệu lignoxenluloza.

Theo khía cạnh khác, để làm thuận tiện cho sự sản xuất *in vitro* của polypeptit làm thoái biến lignin, sáng chế đề xuất cấu trúc biểu hiện có khả năng biểu hiện polypeptit có chứa ít nhất 70% axit amin liên tiếp như nêu trong SEQ ID NO. 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 39, 42, 45, 48, 51, 54, 57, 60, 63, 66, 69, 72, 75, 78, 81, 84, 87, 90, 93, 96, 99, 102, 105, 108 và 111. Tốt hơn là, cấu trúc biểu hiện đã cài xen ADN hoặc cADN bằng các nucleotit liên tiếp như nêu trong SEQ ID NO. 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 35, 38, 41, 44, 47, 50, 53, 56, 59, 62, 65, 68, 71, 74, 77, 80, 83, 86, 89, 92, 95, 98, 101, 104, 107, 110, 1, 4, 7, 10, 13, 16, 19, 22, 25, 28, 31, 34, 37, 40, 43, 46, 49, 52, 55, 58, 61, 64, 67, 70, 73, 76, 79, 82, 85, 88, 91, 94, 97, 100, 103, 106 và 109.

Mục đích khác nữa của sáng chế bộc lộ cấu trúc gen tái tổ hợp có chứa khuôn polynucleotit có trình tự nucleotit nêu trong SEQ ID NO. 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 35, 38, 41, 44, 47, 50, 53, 56, 59, 62, 65, 68, 71, 74, 77, 80, 83, 86, 89, 92, 95, 98, 101, 104, 107, 110, 1, 4, 7, 10, 13, 16, 19, 22, 25, 28, 31, 34, 37, 40, 43, 46, 49, 52, 55, 58, 61, 64, 67, 70, 73, 76, 79, 82, 85, 88, 91, 94, 97, 100, 103, 106 và 109, trong đó khuôn polynucleotit có thể biểu hiện ở tế bào chủ để tạo ra enzym mà oxy hóa lignin từ nguyên liệu lignoxenluloza. Tốt hơn là, cấu trúc gen tái tổ hợp còn có chứa vùng trình tự khởi đầu được liên kết có điều khiển để tăng cường sự biểu hiện của khuôn polynucleotit.

Theo một phương án được ưu tiên của sáng chế, chủng ms6 của nấm *M. phaseolina* được phân lập từ cây đay bị nhiễm. Polypeptit được phân lập cũng tốt hơn là có nguồn gốc từ chủng này.

Mục đích khác nữa của sáng chế là đề xuất phương pháp tiềm năng khả thi về mặt thương mại để phân lập enzym làm thoái biến lignin từ *M. phaseolina* để bắt kịp nhu cầu ngày càng tăng về việc khai thác/sử dụng sinh khối thực vật làm nguồn tài nguyên tái tạo được để sản xuất sản phẩm xenluloza.

Mục đích khác của sáng chế là đề cập đến việc sử dụng chất làm thoái biến lignin trong các ngành công nghiệp thức ăn gia súc, nhiên liệu sinh học, bột gỗ, vải dệt và giấy.

Người có hiểu biết trung bình trong lĩnh vực sẽ dễ dàng nhận ra rằng sáng chế được làm cho thích hợp để thực hiện các mục đích và đạt được kết quả và ưu điểm đã đề cập, cũng như là kết quả và ưu điểm vốn có trong đó. Các phương án được mô tả trong bản mô tả này không nhằm làm giới hạn phạm vi của sáng chế.

Mô tả vắn tắt các hình vẽ

Các hình vẽ kèm theo, mà được kết hợp vào và tạo thành một phần của bản mô tả, minh họa các phương án của sáng chế và, cùng với phần mô tả, giải thích nguyên lý của sáng chế.

Fig. 1 thể hiện cấu trúc chung của lignin gỗ mềm. Cấu trúc arylglycerol- β -aryl ete chính được thể hiện có màu đỏ. Bản đồ lồng thể hiện rượu coniferyl, khối xây dựng phenylpropanoit của lignin gỗ mềm.

Fig. 2 thể hiện sự phân cắt ở bên trong cấu trúc arylglycerol- β -aryl ete lignin không phenol bằng cách oxy hóa bằng lignin peroxidaza.

Fig. 3 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của lignin peroxidaza của SEQ ID NO. 1 và đường M là thang khối lượng phân tử ADN.

Fig. 4 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của lignin peroxidaza của SEQ ID NO. 4 và đường M là thang khối lượng phân tử ADN.

Fig. 5 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của lignin peroxidaza của SEQ ID NO. 7 và đường M là thang khối lượng phân tử ADN.

Fig. 6 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của cloperoxidaza của SEQ ID NO. 13 và đường M là thang khối lượng phân tử ADN.

Fig. 7 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của cloperoxidaza của SEQ ID NO. 16 và đường M là thang khối lượng phân tử ADN.

Fig. 8 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của cloperoxidaza của SEQ ID NO. 19 và đường M là thang khối lượng phân tử ADN.

Fig. 9 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của cloperoxidaza của SEQ ID NO. 22 và đường M là thang khối lượng phân tử ADN.

Fig. 10 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của cloperoxidaza của SEQ ID NO. 25 và đường M là thang khối lượng phân tử ADN.

Fig. 11 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của haemperoxidaza của SEQ ID NO. 28 và đường M là thang khối lượng phân tử ADN.

Fig. 12 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của haemperoxidaza của SEQ ID NO. 31 và đường M là thang khối lượng phân tử ADN.

Fig. 13 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của haemperoxidaza của SEQ ID NO. 34 và đường M là thang khối lượng phân tử ADN.

Fig. 14 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của haemperoxidaza của SEQ ID NO. 37 và đường M là thang khối lượng phân tử ADN.

Fig. 15 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của haemperoxidaza của SEQ ID NO. 40 và đường M là thang khối lượng phân tử ADN.

Fig. 16 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của haemperoxidaza của SEQ ID NO. 43 và đường M là thang khối lượng phân tử ADN.

Fig. 17 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 46 và đường M là thang khối lượng phân tử ADN.

Fig. 18 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 49 và đường M là thang khối lượng phân tử ADN.

Fig. 19 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 52 và đường M là thang khối lượng phân tử ADN.

Fig. 20 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 55 và đường M là thang khối lượng phân tử ADN.

Fig. 21 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 58 và đường M là thang khối lượng phân tử ADN.

Fig. 22 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 64 và đường M là thang khối lượng phân tử ADN.

Fig. 23 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 67 và đường M là thang khối lượng phân tử ADN.

Fig. 24 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 70 và đường M là thang khối lượng phân tử ADN.

Fig. 25 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 73 và đường M là thang khối lượng phân tử ADN.

Fig. 26 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 76 và đường M là thang khối lượng phân tử ADN.

Fig. 27 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 79 và đường M là thang khối lượng phân tử ADN.

Fig. 28 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 82 và đường M là thang khối lượng phân tử ADN.

Fig. 29 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 85 và đường M là thang khối lượng phân tử ADN.

Fig. 30 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 88 và đường M là thang khối lượng phân tử ADN.

Fig. 31 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 91 và đường M là thang khối lượng phân tử ADN.

Fig. 32 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 97 và đường M là thang khối lượng phân tử ADN.

Fig. 33 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 100 và đường M là thang khối lượng phân tử ADN.

Fig. 34 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 103 và đường M là thang khối lượng phân tử ADN.

Fig. 35 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 106 và đường M là thang khối lượng phân tử ADN.

Fig. 36 là ảnh điện di trên gel agarosa thể hiện kết quả khuếch đại PCR trong đó đường 1 là polynucleotit của đa đồng oxidaza của SEQ ID NO. 109 và đường M là thang khối lượng phân tử ADN.

Mô tả chi tiết sáng chế

Các định nghĩa và/hoặc phương pháp được nêu trong bản mô tả này xác định sáng chế và hướng dẫn người có hiểu biết trung bình trong lĩnh vực thực hiện sáng

chế. Trừ khi có chỉ dẫn khác, các thuật ngữ cần được hiểu theo cách dùng thông thường của người có hiểu biết trung bình trong lĩnh vực thích hợp. Ở mức độ mà định nghĩa và/hoặc phương pháp bất kỳ được phát hiện là không thống nhất với định nghĩa và/hoặc phương pháp bất kỳ nêu trong tài liệu tham khảo sáng chế hoặc không phải sáng chế bất kỳ được viện dẫn trong bản mô tả này hoặc trong tài liệu tham khảo bất kỳ được tìm thấy ở nơi khác, cần hiểu rằng định nghĩa và/hoặc phương pháp đã được đề cập mà được cung cấp/được thừa nhận rõ ràng trong đơn này sẽ được sử dụng trong bản mô tả này. Các thuật ngữ số ít bao gồm cả nghĩa số nhiều trừ khi nội dung chỉ ra rõ ràng theo cách khác. Tương tự, từ "hoặc" được dự định là bao gồm cả nghĩa "và" trừ khi nội dung chỉ ra rõ ràng theo cách khác. Do đó "gồm có A hoặc B" nghĩa là bao gồm A, hoặc B, hoặc A và B. Cần hiểu thêm rằng tất cả các giá trị kích thước bazơ hoặc kích thước axit amin, và tất cả trọng lượng phân tử hoặc khối lượng phân tử, được nêu đối với axit nucleic hoặc polypeptit là xấp xỉ, và được nêu để mô tả. Mặc dù phương pháp và nguyên liệu tương tự hoặc tương đương với phương pháp và nguyên liệu được mô tả trong bản mô tả này có thể được sử dụng trong việc thực hiện hoặc thử nghiệm sáng chế, phương pháp và nguyên liệu thích hợp được mô tả dưới đây.

Sáng chế đề xuất trình tự nucleotit của *M. phaseolina* mã hóa cho protein/enzym tham gia vào sự thoái biến lignin. Gen mã hóa cho protein có hoạt tính enzym mà để sử dụng trong công nghiệp hoặc được quan tâm trong công nghiệp. Trình tự hệ gen theo sáng chế mã hóa cho enzym được xác định trước hết bằng cách so sánh trình tự nucleotit của ADN hệ gen *M. phaseolina* và trình tự nucleotit của gen enzym đã biết của vi sinh vật khác. Trước sáng chế này, trình tự nucleotit của gen *M. phaseolina*, khung đọc mở, vị trí của các exon và các intron, cấu trúc của enzym, và sự hữu ích tiềm năng của chúng trong các ngành công nghiệp khác nhau, chẳng hạn như các trình tự nucleotit tham gia vào việc sản xuất thực phẩm và thức ăn chăn nuôi, đồ uống, vải dệt và thuốc tẩy, chưa được biết đến.

Phân tích trình tự hệ gen của *M. phaseolina* phát hiện ra sự phong phú của gen mã hóa cho enzym mà làm thoái biến lignin ở thực vật. Trình tự nucleotit được chú thích ban đầu bằng chương trình phần mềm, chẳng hạn như Genecan và Glimmer M (The Institute of Genome Research, Rockville, NM), mà có thể xác định các vùng mã hóa giả định, các intron, và chỗ ghép nối. Sự lựa chọn thêm tự động và bằng tay của

trình tự nucleotit được thực hiện để tinh chế và thiết lập đặc điểm chính xác của các vùng mã hóa và các đặc điểm gen khác.

Trên 14000 cADN từ *M. phaseolina* được đọc trình tự một phần hoặc toàn bộ. Trong số chúng ba mươi bảy cADN mã hóa cho enzym mới có vai trò giả định trong sự thoái biến lignin được phát hiện.

Khung đọc mở (ORF) được phân tích theo sự đọc trình tự hoàn toàn hoặc một phần của các dòng của thư viện cADN có nguồn gốc từ mARN *M. phaseolina* và được phân tích thêm bằng cách sử dụng phần mềm phân tích trình tự, và bằng cách xác định độ tương đồng với các trình tự đã biết trong cơ sở dữ liệu (công cộng/ tư nhân).

Trong bản mô tả này, một số thuật ngữ trong toàn bộ bản mô tả có nghĩa như được chỉ ra trừ khi được chỉ ra rõ ràng là có nghĩa khác.

Như được dùng trong bản mô tả này, "polynucleotit" là trình tự nucleotit chẳng hạn như mảnh axit nucleic. Polynucleotit có thể là polyme của ARN hoặc ADN tức là mạch đơn hoặc mạch kép, tùy ý có chứa bazơ nucleotit tổng hợp, không tự nhiên hoặc được thay đổi. Polynucleotit ở dạng polyme của ADN có thể được bao gồm một hoặc nhiều đoạn của cADN, ADN hệ gen, ADN tổng hợp, hoặc hỗn hợp của chúng. Polynucleotit được phân lập theo sáng chế có thể bao gồm ít nhất một trong số 150 nucleotit liên tiếp (cả ngược chiều và xuôi chiều) có nguồn gốc từ SEQ ID NO. 1, 4, 7, 10, 13, 16, 19, 22, 25, 28, 31, 34, 37, 40, 43, 46, 49, 52, 55, 58, 61, 64, 67, 70, 73, 76, 79, 82, 85, 88, 91, 94, 97, 100, 103, 106 và 109, hoặc trình tự bổ sung của các trình tự này.

"Polypeptit" như được dùng trong bản mô tả này, là mạch đơn thẳng của axit amin được liên kết cùng nhau bằng liên kết peptit, và thường có trình tự có chiều dài lớn hơn 100 axit amin.

"Được phân lập" nghĩa được làm thay đổi "bởi bàn tay con người" từ tình trạng tự nhiên. Nếu hợp phần hoặc chất có trong tự nhiên, nó "được phân lập" nếu nó đã được thay đổi hoặc được lấy ra từ môi trường ban đầu của nó, hoặc cả hai. Ví dụ, polynucleotit hoặc polypeptit có mặt tự nhiên ở thực vật hoặc động vật sống không "được phân lập," nhưng cùng polynucleotit hoặc polypeptit đó được tách ra từ nguyên liệu đồng tồn tại của tình trạng tự nhiên của nó "được phân lập", là thuật ngữ được sử dụng trong bản mô tả này.

Thuật ngữ “gen”, như được dùng trong bản mô tả này, được xác định là trình tự hệ gen của nấm *M. phaseolina*, cụ thể là trình tự polynucleotit mã hóa cho polypeptit của một loạt các enzym tham gia vào con đường thoái biến lignin.

"Trình tự mã hóa" hoặc "vùng mã hóa" dùng để chỉ phân tử axit nucleic có thông tin trình tự cần để tạo ra sản phẩm gen, chẳng hạn như axit amin hoặc polypeptit, khi trình tự này được biểu hiện. Trình tự mã hóa có thể bao gồm các trình tự không được dịch mã (ví dụ, các intron hoặc các vùng không được dịch mã 5' hoặc 3') ở trong các vùng được dịch mã, hoặc có thể thiếu các trình tự không được dịch mã xen kẽ này (ví dụ, như trong cADN).

Thuật ngữ “đoạn mồi” như được dùng trong bản mô tả này, là oligonucleotit có khả năng liên kết với trình tự axit nucleic đích và làm mồi cho sự tổng hợp axit nucleic. Oligonucleotit khuếch đại như được xác định trong bản mô tả này tốt hơn là có chiều dài nằm trong khoảng từ 10 đến 50, tốt nhất là từ 15 đến 25 nucleotit. Mặc dù vậy, oligonucleotit khuếch đại theo sáng chế có thể được tổng hợp hóa học và oligonucleotit này không phải là axit nucleic có trong tự nhiên.

Chữ viết tắt trong toàn bộ bản mô tả dùng để chỉ axit nucleic có chứa trình tự nucleotit là chữ viết tắt một ký tự thông thường. Do đó, khi bao gồm trong axit nucleic, nucleotit mã hóa có trong tự nhiên được viết tắt như sau: adenin (A), guanin (G), xytosin (C), tymin (T) và uraxin (U). Ngoài ra, trừ khi được xác định theo cách khác, trình tự axit nucleic được thể hiện trong bản mô tả này là theo chiều 5' → 3'.

Như được dùng trong bản mô tả này, thuật ngữ “bổ sung” và từ phát sinh của nó được sử dụng để chỉ sự bắt cặp của axit nucleic theo quy tắc đã biết là A bắt cặp với T hoặc U và C bắt cặp với G. Sự bổ sung có thể là “một phần” hoặc “hoàn toàn”. Trong sự bổ sung một phần, chỉ một số bazơ axit nucleic được bắt cặp theo quy tắc bắt cặp bazơ; trong khi đó trong sự bổ sung hoàn toàn hoặc toàn bộ, tất cả các bazơ được bắt cặp theo quy tắc bắt cặp. Mức độ bổ sung giữa các sợi axit nucleic có thể có ảnh hưởng đáng kể lên hiệu quả và độ mạnh của sự lai hóa giữa các sợi axit nucleic như đã biết trong lĩnh vực. Hiệu quả và độ mạnh của sự lai hóa phụ thuộc vào phương pháp phát hiện.

Thuật ngữ "tế bào chủ", như được dùng trong bản mô tả này, bao gồm dạng tế bào bất kỳ mà dễ bị biến nạp, chuyển nạp, tải nạp, biểu hiện và dạng tương tự với cấu trúc axit nucleic hoặc vật truyền biểu hiện có chứa polynucleotit theo sáng chế.

Thuật ngữ "được liên kết có điều khiển" trong bản mô tả này có nghĩa là cấu hình trong đó trình tự điều khiển được đặt ở vị trí thích hợp so với trình tự mã hóa của trình tự polynucleotit sao cho trình tự điều khiển điều khiển sự biểu hiện của trình tự mã hóa của polypeptit.

"Vật truyền" thường dùng để chỉ đơn vị sao chép, chẳng hạn như plasmid, phago, cosmit, nấm men hoặc virus mà mảnh axit nucleic khác có thể được cài xen có điều khiển vào đó sao cho mang lại sự sao chép hoặc sự biểu hiện của mảnh này. Thuật ngữ "vật truyền" cũng nhằm dùng để chỉ phân tử axit nucleic có khả năng vận chuyển axit nucleic khác vào nơi mà nó được liên kết. Một dạng của vật truyền là "plasmid," mà dùng để chỉ thông lượng ADN mạch kép dạng vòng mà đoạn ADN bổ sung có thể được nối vào đó. Dạng khác của vật truyền là vật truyền virus, trong đó đoạn ADN bổ sung có thể được nối vào hệ gen virus. Các vật truyền nhất định có khả năng tự sao chép ở tế bào chủ mà chúng được đưa vào (ví dụ, vật truyền vi khuẩn có vị trí khởi đầu sao chép vi khuẩn và vật truyền động vật có vú thể kết hợp). Vật truyền khác có thể được hợp nhất vào hệ gen của tế bào chủ khi đưa vào trong tế bào chủ, và bằng cách đó được sao chép cùng với hệ gen vật chủ. Ngoài ra, vật truyền nhất định có khả năng định hướng sự biểu hiện của gen mà chúng được liên kết có điều khiển với. Vật truyền này trong bản mô tả này được đề cập đến là "vật truyền biểu hiện tái tổ hợp" (hoặc đơn giản là, "vật truyền biểu hiện"). Nhìn chung, vật truyền biểu hiện để sử dụng trong kỹ thuật ADN tái tổ hợp thường ở dạng plasmid. Trong bản mô tả này, "plasmid" và "vật truyền" có thể được dùng thay thế lẫn nhau vì plasmid là dạng thường dùng nhất của vật truyền. Tuy nhiên, sáng chế được dự định là bao gồm các dạng khác của vật truyền biểu hiện, chẳng hạn như vật truyền virus (ví dụ, thiếu khả năng tự sao chép retrovirus, adenovirus và virus được kết hợp adeno), mà có chức năng tương đương.

Thuật ngữ "cấu trúc axit nucleic" hoặc "cấu trúc ADN" đôi khi được sử dụng để chỉ trình tự mã hóa hoặc trình tự được liên kết có điều khiển với trình tự điều hòa thích hợp và được cài xen vào vật truyền để biến nạp tế bào. Thuật ngữ này có thể được dùng thay thế lẫn nhau với thuật ngữ "ADN biến nạp" hoặc "gen chuyển".

Thuật ngữ "trình tự khởi đầu" như được dùng trong bản mô tả này, dùng để chỉ trình tự axit nucleic mà có chức năng điều khiển sự phiên mã của gen xuôi chiều. Trình tự khởi đầu thường thích hợp đối với tế bào chủ trong đó gen đích sẽ được biểu hiện. Trình tự khởi đầu cùng với trình tự axit nucleic điều hòa phiên mã và dịch mã khác (cũng được gọi là "trình tự điều khiển") là cần thiết để biểu hiện gen nhất định. Nhìn chung, trình tự điều hòa phiên mã và dịch mã bao gồm, nhưng không giới hạn ở, trình tự khởi đầu, vị trí liên kết ribosom, trình tự mở đầu và kết thúc phiên mã, trình tự mở đầu và kết thúc dịch mã, và trình tự tăng cường hoặc trình tự kích hoạt.

Thuật ngữ "*in vitro*" như được dùng trong bản mô tả này, dùng để chỉ phản ứng sinh học xảy ra trong môi trường nhân tạo bên ngoài sinh vật sống, mà thường được thực hiện trong phòng thí nghiệm bằng cách sử dụng các thành phần của sinh vật mà đã được phân lập từ hoàn cảnh sinh học thông thường của chúng để cho phép thực hiện phân tích chi tiết hơn và thuận tiện hơn.

Thuật ngữ "độ tương đồng theo %" được sử dụng thay thế lẫn nhau trong bản mô tả này với thuật ngữ "độ đồng nhất theo %" trong bản mô tả này và dùng để chỉ mức độ của trình tự axit nucleic hoặc axit amin độ đồng nhất giữa trình tự axit nucleic mà mã hóa cho bất kỳ trong số polypeptit theo sáng chế hoặc trình tự axit amin của polypeptit theo sáng chế, khi được sắp thẳng hàng bằng cách sử dụng chương trình sắp thẳng hàng trình tự.

Ví dụ, như được dùng trong bản mô tả này, độ tương đồng 80% có nghĩa giống như là độ đồng nhất trình tự 80% được xác định bằng thuật toán xác định, và theo đó thể tương đồng của trình tự nhất định có độ đồng nhất trình tự lớn hơn 80% trên chiều dài của trình tự nhất định. Mức độ ví dụ của độ đồng nhất trình tự bao gồm, nhưng không giới hạn ở, độ đồng nhất trình tự 80, 85, 90, 95, 98% hoặc lớn hơn so với trình tự nhất định, *ví dụ*, trình tự mã hóa đối với polypeptit bất kỳ trong số các polypeptit theo sáng chế, như được mô tả trong bản mô tả này.

Chương trình máy tính ví dụ mà có thể được sử dụng để xác định độ đồng nhất giữa hai trình tự bao gồm, nhưng không giới hạn ở, bộ chương trình BLAST, *ví dụ*, BLASTN, BLASTX, và TBLASTX, BLASTP và TBLASTN, có thể truy cập công cộng tại www.ncbi.nlm.nih.gov/BLAST.

Tra cứu trình tự thường được thực hiện bằng cách sử dụng chương trình BLASTN khi đánh giá trình tự axit nucleic nhất định so với trình tự axit nucleic trong trình tự ADN ngân hàng gen và cơ sở dữ liệu công cộng khác. Chương trình BLASTX được ưu tiên để tra cứu trình tự axit nucleic mà đã được dịch mã trong tất cả các khung đọc mở đối với trình tự axit amin trong trình tự protein ngân hàng gen và cơ sở dữ liệu công cộng khác.

Sự sắp thẳng hàng được ưu tiên của trình tự được chọn để xác định "độ đồng nhất theo %" giữa hai hoặc nhiều trình tự được tạo thành bằng cách sử dụng, ví dụ, chương trình CLUSTAL-W.

Như nêu trên, một phương án của sáng chế đề cập đến polynucleotit được phân lập mã hóa cho polypeptit làm thoái biến lignin được phát hiện ở nấm *M. phaseolina* có chứa trình tự nucleotit như nêu trong SEQ ID NO. 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 35, 38, 41, 44, 47, 50, 53, 56, 59, 62, 65, 68, 71, 74, 77, 80, 83, 86, 89, 92, 95, 98, 101, 104, 107, 110, 1, 4, 7, 10, 13, 16, 19, 22, 25, 28, 31, 34, 37, 40, 43, 46, 49, 52, 55, 58, 61, 64, 67, 70, 73, 76, 79, 82, 85, 88, 91, 94, 97, 100, 103, 106 và 109. Do đó, polypeptit làm thoái biến lignin tương ứng được mã hóa bởi các trình tự nucleotit này sẽ mang trình tự axit amin như nêu trong SEQ ID NO. 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 39, 42, 45, 48, 51, 54, 57, 60, 63, 66, 69, 72, 75, 78, 81, 84, 87, 90, 93, 96, 99, 102, 105, 108 và 111.

Theo một phương án, polynucleotit dài 939bp được minh họa trong SEQ ID NO. 2 là dòng cADN nguyên chiều dài mã hóa cho protein lignin peroxidaza biểu thị khung đọc mở mã hóa cho polypeptit có 312 axit amin, như trong SEQ ID NO. 3, có khối lượng phân tử tính được là khoảng 32kD. Thông qua phân tích SMART của SEQ ID NO. 2, nó phát hiện ra sự có mặt của miền Pfam peroxidaza trong trình tự này. Họ protein có miền Pfam có chứa và/hoặc gồm có peroxidaza nấm gây tiết. Chúng là glycoprotein dạng monome tham gia vào việc làm thoái biến lignin.

Tốt hơn là, polynucleotit dài 1116bp được minh họa trong SEQ ID NO. 5 là dòng cADN nguyên chiều dài mã hóa cho protein lignin peroxidaza biểu thị khung đọc mở mã hóa cho polypeptit có 371 axit amin, như trong SEQ ID NO. 6, có khối lượng phân tử tính được là khoảng 40kD. Thông qua phân tích SMART của SEQ ID NO. 5, nó phát hiện ra sự có mặt của miền Pfam peroxidaza trong trình tự này. Họ protein có

miền Pfam có chứa và/hoặc gồm có peroxidaza nấm gây tiết. Chúng là glycoprotein dạng monome tham gia vào việc làm thoái biến lignin.

Tương tự, theo một khía cạnh khác, polynucleotit dài 1023bp được minh họa trong SEQ ID NO. 8 là dòng cADN nguyên chiều dài mã hóa cho protein lignin peroxidaza biểu thị khung đọc mở mã hóa cho polypeptit có 340 axit amin, như trong SEQ ID NO. 9, có khối lượng phân tử tính được là khoảng 35kD. Thông qua phân tích SMART của SEQ ID NO. 8, nó phát hiện ra sự có mặt của miền Pfam peroxidaza trong trình tự này. Họ protein có miền Pfam có chứa và/hoặc gồm có peroxidaza nấm gây tiết. Chúng là glycoprotein dạng monome tham gia vào việc làm thoái biến lignin.

Theo khía cạnh khác, polynucleotit mã hóa cho cloperoxidaza có 1260bp như được minh họa trong SEQ ID NO. 11 trong khi protein được mã hóa là polypeptit có 419 axit amin có khối lượng phân tử tính được là khoảng 45kD như được minh họa trong SEQ ID NO. 12. SEQ ID NO. 11 có chứa miền Pfam, cụ thể là peroxidaza_2. Miền này tham gia vào sự thoái biến lignin.

Theo khía cạnh khác, polynucleotit mã hóa cho cloperoxidaza có 690bp như được minh họa trong SEQ ID NO. 14 trong khi protein được mã hóa là polypeptit có 229 axit amin có khối lượng phân tử tính được là khoảng 25kD như được minh họa trong SEQ ID NO. 15. SEQ ID NO. 14 có chứa miền Pfam, cụ thể là peroxidaza_2. Miền này tham gia vào sự thoái biến lignin.

Theo khía cạnh khác, polynucleotit mã hóa cho cloperoxidaza có 1194bp như được minh họa trong SEQ ID NO. 17 trong khi protein được mã hóa là polypeptit có 397 axit amin có khối lượng phân tử tính được là khoảng 43kD như được minh họa trong SEQ ID NO. 18. SEQ ID NO. 17 có chứa miền Pfam, cụ thể là peroxidaza_2. Miền này tham gia vào sự thoái biến lignin.

Theo khía cạnh khác, polynucleotit mã hóa cho cloperoxidaza có 1317bp như được minh họa trong SEQ ID NO. 20 trong khi protein được mã hóa là polypeptit có 438 axit amin có khối lượng phân tử tính được là khoảng 47kD như được minh họa trong SEQ ID NO. 21. SEQ ID NO. 20 có chứa miền Pfam, cụ thể là peroxidaza_2. Miền này tham gia vào sự thoái biến lignin.

Theo khía cạnh khác, polynucleotit mã hóa cho cloperoxidaza có 780bp như được minh họa trong SEQ ID NO. 23 trong khi protein được mã hóa là polypeptit có

259 axit amin có khối lượng phân tử tính được là khoảng 29kD như được minh họa trong SEQ ID NO. 24. SEQ ID NO. 23 có chứa miền Pfam, cụ thể là peroxidaza_2. Miền này tham gia vào sự thoái biến lignin.

Theo khía cạnh khác, polynucleotit mã hóa cho cloperoxidaza có 951bp như được minh họa trong SEQ ID NO. 26 trong khi protein được mã hóa là polypeptit có 316 axit amin có khối lượng phân tử tính được là khoảng 35kD như được minh họa trong SEQ ID NO. 27. SEQ ID NO. 26 có chứa miền Pfam, cụ thể là peroxidaza_2. Miền này tham gia vào sự thoái biến lignin.

Theo khía cạnh khác, polynucleotit mã hóa cho haemperoxidaza có 1116bp như được minh họa trong SEQ ID NO. 29 trong khi protein được mã hóa là polypeptit có 371 axit amin có khối lượng phân tử tính được là khoảng 41kD như được minh họa trong SEQ ID NO. 30. SEQ ID NO. 29 có chứa miền Pfam, cụ thể là peroxidaza. Haemperoxidaza oxy hóa tiểu đơn vị lignin bằng cách sử dụng hydro peroxit ngoại bào được tạo ra bởi oxidaza không liên quan làm đồng chất nền. Miền này tham gia vào sự thoái biến lignin.

Theo khía cạnh khác, polynucleotit mã hóa cho haemperoxidaza có 1623bp như được minh họa trong SEQ ID NO. 32 trong khi protein được mã hóa là polypeptit có 540 axit amin có khối lượng phân tử tính được là khoảng 57kD như được minh họa trong SEQ ID NO. 33. SEQ ID NO. 32 có chứa miền Pfam, cụ thể là peroxidaza. Haemperoxidaza oxy hóa tiểu đơn vị lignin bằng cách sử dụng hydro peroxit ngoại bào được tạo ra bởi oxidaza không liên quan làm đồng chất nền. Miền này tham gia vào sự thoái biến lignin.

Theo khía cạnh khác, polynucleotit mã hóa cho haemperoxidaza có 1599bp như được minh họa trong SEQ ID NO. 35 trong khi protein được mã hóa là polypeptit có 532 axit amin có khối lượng phân tử tính được là khoảng 57kD như được minh họa trong SEQ ID NO. 36. SEQ ID NO. 35 có chứa miền Pfam, cụ thể là peroxidaza. Haemperoxidaza oxy hóa tiểu đơn vị lignin bằng cách sử dụng hydro peroxit ngoại bào được tạo ra bởi oxidaza không liên quan làm đồng chất nền. Miền này tham gia vào sự thoái biến lignin.

Theo khía cạnh khác, polynucleotit mã hóa cho haemperoxidaza có 2049bp như được minh họa trong SEQ ID NO. 38 trong khi protein được mã hóa là polypeptit có

682 axit amin có khối lượng phân tử tính được là khoảng 70kD như được minh họa trong SEQ ID NO. 39. SEQ ID NO. 38 có chứa miền Pfam, cụ thể là peroxidaza. Haemperoxidaza oxy hóa tiểu đơn vị lignin bằng cách sử dụng hydro peroxit ngoại bào được tạo ra bởi oxidaza không liên quan làm đồng chất nền. Miền này tham gia vào sự thoái biến lignin.

Theo khía cạnh khác, polynucleotit mã hóa cho haemperoxidaza có 1605bp như được minh họa trong SEQ ID NO. 41 trong khi protein được mã hóa là polypeptit có 534 axit amin có khối lượng phân tử tính được là khoảng 57kD như được minh họa trong SEQ ID NO. 42. SEQ ID NO. 41 có chứa miền Pfam, cụ thể là peroxidaza. Haemperoxidaza oxy hóa tiểu đơn vị lignin bằng cách sử dụng hydro peroxit ngoại bào được tạo ra bởi oxidaza không liên quan làm đồng chất nền. Miền này tham gia vào sự thoái biến lignin.

Theo khía cạnh khác, polynucleotit mã hóa cho haemperoxidaza có 960bp như được minh họa trong SEQ ID NO. 44 trong khi protein được mã hóa là polypeptit có 319 axit amin có khối lượng phân tử tính được là khoảng 35kD như được minh họa trong SEQ ID NO. 45. SEQ ID NO. 44 có chứa miền Pfam, cụ thể là peroxidaza. Haemperoxidaza oxy hóa tiểu đơn vị lignin bằng cách sử dụng hydro peroxit ngoại bào được tạo ra bởi oxidaza không liên quan làm đồng chất nền. Miền này tham gia vào sự thoái biến lignin.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 1713bp như được minh họa trong SEQ ID NO. 47 trong khi protein được mã hóa là polypeptit có 570 axit amin có khối lượng phân tử tính được là khoảng 64kD như được minh họa trong SEQ ID NO. 48. SEQ ID NO. 47 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 1860bp như được minh họa trong SEQ ID NO. 50 trong khi protein được mã hóa là polypeptit có 619 axit amin có khối lượng phân tử tính được là khoảng 69kD như được minh họa trong SEQ ID NO. 51. SEQ ID NO. 50 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 1824bp như được minh họa trong SEQ ID NO. 53 trong khi protein được mã hóa là polypeptit có 607 axit amin có khối lượng phân tử tính được là khoảng 67kD như được minh họa trong SEQ ID NO. 54. SEQ ID NO. 53 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 1737bp như được minh họa trong SEQ ID NO. 56 trong khi protein được mã hóa là polypeptit có 578 axit amin có khối lượng phân tử tính được là khoảng 63kD như được minh họa trong SEQ ID NO. 57. SEQ ID NO. 56 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 1665bp như được minh họa trong SEQ ID NO. 59 trong khi protein được mã hóa là polypeptit có 554 axit amin có khối lượng phân tử tính được là khoảng 61kD như được minh họa trong SEQ ID NO. 60. SEQ ID NO. 59 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 1803bp như được minh họa trong SEQ ID NO. 62 trong khi protein được mã hóa là polypeptit có 600 axit amin có khối lượng phân tử tính được là khoảng 65kD như được minh họa trong SEQ ID NO. 63. SEQ ID NO. 62 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 1635bp như được minh họa trong SEQ ID NO. 65 trong khi protein được mã hóa là polypeptit có 544 axit amin có khối lượng phân tử tính được là khoảng 61 kD như được minh họa trong SEQ ID NO. 66. SEQ ID NO. 65 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 1743bp như được minh họa trong SEQ ID NO. 68 trong khi protein được mã hóa là polypeptit có 580 axit amin có khối lượng phân tử tính được là khoảng 63kD như được minh họa trong SEQ ID NO. 69. SEQ ID NO. 68 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 1836bp như được minh họa trong SEQ ID NO. 71 trong khi protein được mã hóa là polypeptit có 611 axit amin có khối lượng phân tử tính được là khoảng 67kD như được minh họa trong SEQ ID NO. 72. SEQ ID NO. 71 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 702bp như được minh họa trong SEQ ID NO. 74 trong khi protein được mã hóa là polypeptit có 233 axit amin có khối lượng phân tử tính được là khoảng 26kD như được minh họa trong SEQ ID NO. 75. SEQ ID NO. 74 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 1854bp như được minh họa trong SEQ ID NO. 77 trong khi protein được mã hóa là polypeptit có 617 axit amin có khối lượng phân tử tính được là khoảng 68kD như được minh họa trong SEQ ID NO. 78. SEQ ID NO. 77 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 2073bp như được minh họa trong SEQ ID NO. 80 trong khi protein được mã hóa là polypeptit có 690 axit amin có khối lượng phân tử tính được là khoảng 77kD như được minh họa trong SEQ ID NO. 81. SEQ ID NO. 80 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 1824bp như được minh họa trong SEQ ID NO. 83 trong khi protein được mã hóa là polypeptit có 607 axit amin có khối lượng phân tử tính được là khoảng 67kD như được minh họa trong SEQ ID NO. 84. SEQ ID NO. 83 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 1932bp như được minh họa trong SEQ ID NO. 86 trong khi protein được mã hóa là polypeptit có 643 axit amin có khối lượng phân tử tính được là khoảng 72kD như được minh họa trong SEQ ID NO. 87. SEQ ID NO. 86 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 1950bp như được minh họa trong SEQ ID NO. 89 trong khi protein được mã hóa là polypeptit có 649 axit amin có khối lượng phân tử tính được là khoảng 72kD như được minh họa trong SEQ ID NO. 90. SEQ ID NO. 89 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 2016bp như được minh họa trong SEQ ID NO. 92 trong khi protein được mã hóa là polypeptit có 671 axit amin có khối lượng phân tử tính được là khoảng 74kD như được minh họa trong SEQ ID NO. 93. SEQ ID NO. 92 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 2085bp như được minh họa trong SEQ ID NO. 95 trong khi protein được mã hóa là polypeptit có 694 axit amin có khối lượng phân tử tính được là khoảng 78kD như được minh họa trong SEQ ID NO. 96. SEQ ID NO. 95 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 1821bp như được minh họa trong SEQ ID NO. 98 trong khi protein được mã hóa là polypeptit có 606 axit amin có khối lượng phân tử tính được là khoảng 68kD như được minh họa trong SEQ ID NO. 99. SEQ ID NO. 98 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 1854bp như được minh họa trong SEQ ID NO. 101 trong khi protein được mã hóa là polypeptit có 617 axit amin có khối lượng phân tử tính được là khoảng 68kD như được minh họa trong SEQ ID NO. 102. SEQ ID NO. 101 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 1815bp như được minh họa trong SEQ ID NO. 104 trong khi protein được mã hóa là polypeptit có 604 axit amin có khối lượng phân tử tính được là khoảng 67kD như được minh họa trong SEQ ID NO. 105. SEQ ID NO. 104 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 1632bp như được minh họa trong SEQ ID NO. 107 trong khi protein được mã hóa là polypeptit có 543 axit amin có khối lượng phân tử tính được là khoảng 61kD như được minh họa trong SEQ ID NO. 108. SEQ ID NO. 107 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Theo khía cạnh khác, polynucleotit mã hóa cho đa đồng oxidaza (lacaza) có 1752bp như được minh họa trong SEQ ID NO. 110 trong khi protein được mã hóa là polypeptit có 583 axit amin có khối lượng phân tử tính được là khoảng 64kD như được minh họa trong SEQ ID NO. 111. SEQ ID NO. 110 có chứa miền Pfam, cụ thể là đa đồng oxidaza. Đa đồng oxidaza (lacaza) này là protein có chứa đồng, không haem, ngoại bào mà xúc tác cho sự oxy hóa một electron của phenol thành gốc phenoxy.

Các trình tự được đề xuất bởi sáng chế cũng có thể được dùng làm nguyên liệu điều chế cho sự cải biến hợp lý hoặc thiết kế của enzym mới có các tính chất làm cho enzym có thể thực hiện tốt hơn trong các quy trình đòi hỏi khắt khe.

Bản mô tả này bao gồm nội dung trong các yêu cầu bảo hộ kèm theo, cũng như là nội dung của phần mô tả trên đây. Mặc dù sáng chế này được mô tả ở dạng được ưu tiên với mức độ cụ thể nhất định, cần hiểu rằng việc mô tả dạng được ưu tiên này chỉ làm ví dụ và nhiều thay đổi về chi tiết của cấu trúc và sự kết hợp và sự sắp xếp của các phần có thể được áp dụng mà không nằm ngoài phạm vi của sáng chế và yêu cầu bảo hộ.

Các tài liệu khác nhau được viện dẫn trong bản mô tả này chỉ nhằm mục đích tham khảo.

Ví dụ thực hiện sáng chế

Các ví dụ sau đây nhằm minh họa thêm cho sáng chế, không nhằm làm giới hạn sáng chế ở các phương án cụ thể được mô tả trong đó.

Ví dụ 1: Thiết kế và tổng hợp đoạn môi

Đoạn môi được sử dụng trong nghiên cứu này được thiết kế từ hệ phiên mã được chọn lọc bằng tay và “mô hình gen” được dự đoán từ trình tự hệ gen của *M. phaseolina* ms6, bằng cách lựa chọn các trình tự bằng tay với các ORF hoàn chỉnh hoặc bằng cách sử dụng cơ sở dữ liệu trong đó gen tương tự đã được phân lập thành công từ thực vật khác. Phân tích thông tin sinh học so sánh của các trình tự nucleotit thu được từ hệ phiên mã được thực hiện bằng cách sử dụng NCBI BLAST, BLASTP, RPS-BLAST, BLASTX và PSI-BLAST để xác định dạng tương đồng của gen liên quan và để xác định chính xác gen. Sự sắp thẳng hàng trình tự nucleotit được thực hiện thông qua clustalW phiên bản 1.82 bất cứ khi nào tìm thấy nhiều trình tự từ “tổ hợp gen”. Sau đó, sự sắp thẳng hàng này được hiệu chỉnh. Đoạn môi đặc hiệu gen (cả xuôi và ngược) được chọn lọc bằng tay hoặc thông qua Đoạn môi 3 cộng với công cụ và đoạn môi được tổng hợp tùy chỉnh.

Tất cả các oligonucleotit được sử dụng trong nghiên cứu này được tổng hợp và được tinh chế bằng HPLC bởi nhà cung cấp và mua của hãng Integrated DNA

Technologies (IDT). Dung dịch dự trữ 100pmol được điều chế trong ddH₂O đã thanh trùng và được lưu trữ ở nhiệt độ -20°C, trong các phân phân ước để sử dụng.

Trình tự oligonucleotit được sử dụng làm đoạn mồi cho PCR

Tên gen	SEQ ID NO.	Nhận dạng trình tự	Trình tự đoạn mồi	Sản phẩm khuếch đại (bp)
lignin peroxidaza	1	112	Xuôi GAGACCGCTACACCCACCCCT	1155
		113	Ngược CACCGCACTACGACCTCGCT	
lignin peroxidaza	4	114	Xuôi GTCGTCGCGTGGCTGCTAGA	1276
		115	Ngược CCAGGCATCGGGGAACCTTCGG	
lignin peroxidaza	7	116	Xuôi GGCGGCTCTCTCGCAGACGTA	1234
		117	Ngược GCCCTGCCCCAACCGATTCA	
Cloperoxidaza	10	118	Xuôi TGCTGCCTCCGCTCTGTGCGC	1769
		119	Ngược CGCACCATGTGCGCTCTGCC	
Cloperoxidaza	13	120	Xuôi AACCGCTTTACCTGCCAGCCA	964
		121	Ngược ATTGGGGTTCGGTGCTCAGGAGT	
Cloperoxidaza	16	122	Xuôi ACGGAGCACATGAACACCGTCC	1777
		123	Ngược CTTCGCACCGCGAGCAGAGG	
Cloperoxidaza	19	124	Xuôi TCCCGCGAGCCCTTGGTCTG	1644
		125	Ngược GCCCTCGCTGGTTCCTTTGCT	
Cloperoxidaza	22	126	Xuôi ATGTTTTGTTTCGCGCCGCT	1265

		127	Ngược AATCTACCACTCCCGTCCCGC	
Cloperoxidaza	25	128	Xuôi ACCGCCGCCTTCGTTTACGTC	475
		129	Ngược GCCCCTTACTTTGCCGGTC	
Haem peroxidaza	28	130	Xuôi TCGCCTGTGCTCACACCACG	1235
		131	Ngược TTAACGTCCGCCAAGCACGC	
Tên gen	SEQ ID NO.		Trình tự đoạn mồi	Sản phẩm khuếch đại (bp)
Haem peroxidaza	31	132	Xuôi CGTTCCCAAGCCCGACGACTC	1177
		133	Ngược GGGCAAGTCCCAAGCCCATC	
Haem peroxidaza	34	134	Xuôi AATGCGGTTTCTCGGGGGCT	1762
		135	Ngược TGTTGCTGGCCCTATGAAGGCAT	
Haem peroxidaza	37	136	Xuôi CCATGGCAAGGCATCCCGGC	1172
		137	Ngược TGTGGCATCCCAACAGGGGC	
Haem peroxidaza	40	138	Xuôi TGGACCTGGCCGTCAAACCG	1488
		139	Ngược GCCTCACAGAGCCGCACACTC	
Haem peroxidaza	43	140	Xuôi CAGGCATGCGCTACGAGGCT	800
		141	Ngược ACCAGCTTACAGACGTGCCCTG A	
Đa đồng oxidaza	46	142	Xuôi TGGGCGGGCAGGTACGTGAAT	1935

		143	Ngược TCCGTGCTCTGGCCTCGCAT	
Đa đồng oxidaza	49	144	Xuôi GCTGGCGACGACAAGTGGCT	2033
		145	Ngược CCACAGAGTTCGCGAGGCCC	
Đa đồng oxidaza	52	146	Xuôi CATCCCACCGCGGGAAGCCT	2172
		147	Ngược ATCCCCGCCGTCACGGTTTT	
Đa đồng oxidaza	55	148	Xuôi TCGTTGAAGTCGCTGTCCCGT	1944
		149	Ngược GCCCCGCACACCTGCCATAG	
Đa đồng oxidaza	58	150	Xuôi TGCTTGCTCAAGGGCGCTCA	2177
		151	Ngược TCAACTCAGACTACTGTGAAG TGC	
Tên gen	SEQ ID NO.		Trình tự đoạn môi	Sản phẩm khuếch đại (bp)
Đa đồng oxidaza	61	152	Xuôi CACTGGCACGGCTTCACGCA	4930
		153	Ngược CCGCTACGCGGTCGACTCCT	
Đa đồng oxidaza	64	154	Xuôi CACCTCCGGTCGGGTAAGT	2116
		155	Ngược TACAGGTGCTATGCCAGCGTGC	
Đa đồng oxidaza	67	156	Xuôi TAGCATCGGCACAACGCCCAT	1246
		157	Ngược GAACCGGTGGCCGTGAAGGTG	
Đa đồng oxidaza	70	158	Xuôi ACCCACCGCTCGCTCTACA	2083
		159	Ngược TGGAAATGCGCAGAAGGACCGT	

Đa đồng oxidaza	73	160	Xuôi GCACGACATGTGGATTGCGGC	788
		161	Ngược TTGCCAGCCGTGCTCCGTCA	
Đa đồng oxidaza	76	162	Xuôi ACGACGTTGCAAGCTCCGCC	2022
		163	Ngược CCATCGGGCATAGAAGTCGCCG	
Đa đồng oxidaza	79	164	Xuôi TTCACCGGAGTCGCCTTCCCA	2249
		165	Ngược CGACGGGCTGCAGTACGAGA	
Đa đồng oxidaza	82	166	Xuôi AATCCCCTCTCACCTCGCCGC	2103
		167	Ngược GGCCACCCCTCAGAGACCGGA	
Đa đồng oxidaza	85	168	Xuôi CGCCGAACCAAAGCCTCCTCC	2858
		169	Ngược GCACAGGAGAAAGAGCTCACCC C	
Đa đồng oxidaza	88	170	Xuôi TCGCCCCGTCCAGGAGAGATA	2208
		171	Ngược CCCCCATCTACCGGCCATTC	
Đa đồng oxidaza	91	172	Xuôi ATGAGGGGTAAATCGCGACGG	2382
		173	Ngược CCCTCTCACAACCTGACCCTGT	
Đa đồng oxidaza	94	174	Xuôi TCACTCAGTGCCCTACCGCTCC	1299
		175	Ngược CCGCAACGACTCCGTCCGGT	
Đa đồng oxidaza	97	176	Xuôi TGGGCTGATCCCGTTGCAGGA	1405
		177	Ngược GTAGCCGTGGCTGAGCGTGTT	

Đa đồng oxidaza	100	178	Xuôi TGCCGTTGCTGTAACATGCCGT	2099
		179	Ngược GACGGCGCTTTGCTCTTGCG	
Đa đồng oxidaza	103	180	Xuôi AGCATGCAATACTCGGTCGGTC T	633
		181	Ngược CGGGCAGCAGGATGTTTGCCAT	
Đa đồng oxidaza	106	182	Xuôi GGATACTCTCCGGGCACGTTTCG	1703
		183	Ngược ATGGCAGTGGACTGCGCGAC	
Đa đồng oxidaza	109	184	Xuôi GGAGGGAGCACTGATGCGCT	1873
		185	Ngược GGCTGCTCCTCCGCTCATGG	

Ví dụ 2: Khuếch đại, tách dòng và đọc trình tự lignin peroxidaza, cloperoxidaza, haemperoxidaza và đa đồng peroxidaza từ *M. phaseolina* ms6

ARN tổng số được phân lập từ hệ sợi ba ngày tuổi phát triển trên môi trường lỏng như đã được mô tả từ trước bởi Chomczynski P and Sacchi N, Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. (Anal Biochem 1987, 162: 156-159). Chất lượng hoặc tính toàn vẹn của ARN được kiểm tra bằng cách điện di trên gel agarosa và được định lượng bằng cách sử dụng Thermo Scientific Nano Drop 2000 theo quy trình tiêu chuẩn. Sợi cADN thứ nhất được tổng hợp bằng cách sử dụng transcriptaza ngược SuperScript III (Invitrogen) theo hướng dẫn của nhà sản xuất. Gen được khuếch đại từ cADN bằng PCR bằng cách sử dụng đoạn mồi đặc hiệu gen. Phản ứng PCR (50µl) có chứa 1µl cADN, 20pmol của mỗi đoạn mồi, 5µl 10X đệm PCR, 5µl hỗn hợp dNTP 2,5mM và 1,0 đơn vị ADN polymeraza PfuTaq. PCR được thực hiện trong Thermal Cycler (Applied Biosystems) bằng cách sử dụng các điều kiện sau đây: làm biến tính ban đầu trong thời gian 5 phút ở nhiệt độ 95°C sau đó là 35 chu kỳ biến tính ở nhiệt độ 95°C trong thời gian 30 giây, ủ ở nhiệt độ 59-61°C trong thời gian 30 giây và kéo dài ở nhiệt

độ 72°C trong thời gian từ 1 đến 2,0 phút tùy thuộc vào chiều dài của gen đích, với sự kéo dài cuối cùng ở nhiệt độ 72°C trong thời gian 7 phút. Sản phẩm PCR được phân tích bằng gel agarosa 1% bằng cách sử dụng 1X đệm TAE và sản phẩm khuếch đại được rửa giải từ gel bằng cách sử dụng bộ kit chiết gel QIAGEN theo hướng dẫn của nhà sản xuất. Sản phẩm PCR đã tinh sạch được nối vào bộ kit tách dòng pCR®8/GW/TOPO® TA (Invitrogen) và được biến nạp vào tế bào *E. coli* đủ khả năng (Invitrogen). Plasmid được phân lập từ các khuẩn lạc giả định bằng cách sử dụng bộ kit QIAprep Spin Miniprep (QIAGEN) theo hướng dẫn của nhà sản xuất. Sự có mặt của vật cài xen được kiểm tra bằng cách sử dụng đoạn môi đặc hiệu gen và plasmid dương tính được đưa đi đọc trình tự.

Ví dụ 3: Phân tích trình tự

Trình tự nucleotit và trình tự axit amin lần lượt được phân tích bằng chương trình BLASTN và chương trình BLASTP. Các trình tự được báo cáo từ thực vật khác được sắp thẳng hàng với ClustalW. Phân tích phát sinh loài được thực hiện bằng cách sử dụng Neighbour Joining (NJ).

```
SEQ ID NO      : 1
CHIỀU DÀI     : 1379 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)
LOẠI           : ADN
SINH VẬT       : M. phaseolina
CTGCTCAGCAGATCTTATCCCTCCAGAGCCAGAGCATCTTTAACTGACCAGCTTGCCTTCGTCCCATCGTTTGCA
TTTTTTTACACTTGACCAGTCAAGGCCCTCTTCCTTCTTTGAGACCGCTACACCCACCCCTTCGTACTGTCAAA
ATGTTATTCTCAAAGTCTTCCATCTTTCTCCTCTCCACTGCGGCCAGTGTGCAAGCACTCAGCCTCTCCGATGTC
TCCTCTGCCGCTCGGTCTGAAGCGTGAAGCTTCCGGCTTGGGGAACAACCTTCTATCTCTCGTTCACCGTCGG
GACTCTTGCCCTGATGTTTGGCAGAAAGTCGCATCCGAGCTGAAGGGCTGGTTCTTGATGGTTCGGTGTGCAGT
GACGACGCACGCGCTGCCATCCGCTCTCTTTCCACGACTGCTTTTCCGGCGGCTGCGATGGGTCCATCATCCTT
GCCCACGAGTACACCCGCTCCGACAACGCTGGCTTAGCAGACTTTGCTATGAAGCTAGCGCCTCTCGCGGACCAG
TACGAGGTCGGAACAGCTGACCTGATCCAATTCGCTGGCGGTAAGCCTTATCTATCCACCTTAAACAAAAA
AAAaGGAaaaaaGACTATCTCTAACACCTACCTCGACCAGCCCTCGCCACGGCCACCTGTCCCCTCGGCCCCCG
CATAGCCGTCAAAGTCGGCCGCCAGGACTCGTCAACGCCCTCGGCAGAGGGACAGCTCCCCTCGTCGCGATCCTC
GGCCTCCGTCTGATCGACAGTTTCGCGGCGAAGGGTTCTCTGAGATAGACCTCGTCGCCCTCGTCGGTGCCCA
CAGCACTGCCAAGCAGTTCTTCGATCAGCCCGACAAGGCCGGCCAAAGCCTCGACTCCACCCCGGCACTTGGGA
CACCAACTTTTACCGCCAGACGACGCTCGGCACTGCGCCGTTTACCCTCGAGAGCGATAAGAACCTCGCCACGGA
TCTGAGGACCGCGGTGCAGTGGACGGCGTTAATGCCCAGGGCGTGTGGGCCGCGGCGTATGTGAGCGCGTAAGT
AGTGCATCATATGTTTCTTCTCTGCTCTCAGATTGTTGAGAACTAATGGTCAATGTAGGATGAACAAGATGACCG
TTTTGGGCAATGATGTGACGAGCTTGACGATTGCACGAGTGTTATCAGTGCGGCAACGAGCAAGCGCGACATCA
AGGCCGCACCTATTGCGGATAGGATTTGAGCTATGCAAGCAGTTCTGGGTGAGCGAGGTGCTAGTGCGGTGACTG
GAGTTCTGTTTTTTTTTTTTTTtCTTAATTGGTAATGTCTGCTCTTGTGTTGATACCTGCCAGTATTGCGGATTTAT
TCATCGGGAAGAGTTTAGAGAAGGGGGCA
```

```
SEQ ID NO      : 2
CHIỀU DÀI     : 939
LOẠI           : ADN
SINH VẬT       : M. phaseolina
TÊN/KHÓA DẤU HIỆU : CDS
VỊ TRÍ         : (1).....(939)
atgttatttctcaaagtcttccatctttctcctctccactgcgccagtggtgcaagcactc
M L F S K S S I F L L S T A A S V Q A L
agcctctccgatgtctcctctgcccgcctcggtcctgaagcgtgaagcttccggcttgggg
```

S L S D V S S A A S V L K R E A S G L G
 aacaaccttctatctctcgttcaccgtcgggactcttgccctgatgtttggcagaaagtc
 N N L L S L V H R R D S C P D V W Q K V
 gcatccgagctgaagggctggttcttgatggttccgtgtgcagtgcgacgcacgcgcgt
 A S E L K G W F L D G S V C S D D A R A
 gccatccgcctctctttccacgactgcttttccggcggctgcgatgggtccatcatcctt
 A I R L S F H D C F S G G C D G S I I L
 gccacgagtagacccgcgtccgacaacgctggcttagcagactttgctatgaagctagcg
 A H E Y T R S D N A G L A D F A M K L A
 cctctcgcggaccagtagcaggtcgggaacagctgacctgatccaattcgctggcgccctc
 P L A D Q Y E V G T A D L I Q F A G A L
 gccacggccacctgtccctcggccccgcacatagccgtcaaagtcggccgccaggactcg
 A T A T C P L G P R I A V K V G R Q D S
 tcaacgccctcggcagagggacagctccctcgtcgcgatcctcggcctccgtcctgatc
 S T P S A E G Q L P S S R S S A S V L I
 gaccagttcgcggcgaaggggttctctgagatagacctcgtcgcctcgtcgggtgccac
 D Q F A A K G F S E I D L V A L V G A H
 agcactgccaagcagttcttcgatcagcccgaagaaggccggccaaagcctcgactccacc
 S T A K Q F F D Q P D K A G Q S L D S T
 cccggcacttgggacaccaacttttaccgccagacgacgctcggcactgcgcccgttacc
 P G T W D T N F Y R Q T T L G T A P V T
 ctcgagagcgataagaacctcgccacggatctgaggaccgcggtgcagtggacggcgttt
 L E S D K N L A T D L R T A V Q W T A F
 aatgcccagggcgtgtggggcgcggcgtatgtgagcgcgatgaacaagatgaccgttttg
 N A Q G V W A A A Y V S A M N K M T V L
 ggcaatgatgtcagcagcttgacggattgcacgagtggttatcagtgcggaacgagcaag
 G N D V S S L T D C T S V I S A A T S K
 cgcgacatcaaggccgcacctattgcggaataggatttga
 R D I K A A P I A D R I -

SEQ ID NO : 3

CHIỀU DÀI : 312

LOẠI : PRT

SINH VẬT : *M. phaseolina*

MLFSKSSIFLLSTAASVQALSLSDVSSAASVLKREASGLGNLLSLVHRRDSCPDVWQKVASELKGWFLDGSVCS
 DDARAAIRLSFHDCFSGGCDGSIILAHEYTRSADNGLADLFAMKLAPLADQYEVGTADLIQFAGALATATCPLGPR
 IAVKVGRQDSSTPSAEGQLPSSRSSASVLIDQFAAKGFSEIDLVALVGAHSTAKQFFDQPKAGQSLDSTPGTWD
 TNFYRQTTLGTAPVTLES DKNLATDLRTAVQWTAFAAQGVWAAAYVSAMNKMTVLGNDVSSSLTDCTSVISAATSK
 RDIKA APIADRI*

SEQ ID NO : 4

CHIỀU DÀI : 1416 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

TCTATATCTTGCCCTTCGCCTGTTTCGCTTCAGGAGAACCCAAACGCCAACATTCTCAACCTCCTCTTTGAGT
 TACAATCGCTACTCTGGGAAATTTGTCCCAGCCTTTTCATTCAACAACCTGTTTCTGTTTCAAACCAATCCGACGCC
 ATGAAGTTCTCCACAGTCATCTCGAGCGTTGCTCTCACTTCTCTACTCCAGCCCGCCCTTGCCCTACCCTGGCATG
 GCAAATGTCGTCTCGGAGATCAAGGCCCGTCAAAACACCAATAATGATGGTGACTCAAATCCCGAGATGATTGGC
 GATCTCGCCACCACCGGCCCAACCACCCCTGTGGGCCAAAGCATATACAACATCCTGATGGGGACCGAGTCGGCC
 GAGACCAAGCAGGCTGGCTACATCCCCCTCTTATCGGCACCAACGCCTGCAAGAGGGACACCTGCTGCATCTGG
 GCCTACATCGCCGCCGAACCTGACCCTCAATTTCAAGGGCATCACGGGCCGCTGCAACAAGAACCGCGCTGCCGCC
 ATTCGGCTCGGCTTCCACGACGCGGGGACTTGGTCCAAGAGCAGCAACGGCGGCGGCGCGGACGGCTCGATCGCG
 CTGTCTGGGCACGGAGATCAACAAGGCCGAGAACAACGGGCTGCAGGACATCATCGGCAAGATGATCACGTGGCAG
 AAGCGGTACGGGGTGGGCATGGCGGATCTGATCCAGTTTCGCGGCCATCCACGCCCGTGGTGACGTGCCCGCTGGGG
 CCGCGCATCCGCTTCTTCGTCTGGGCGCAAGGACAGCAAAACGGCCAACGACGTACGCCTGCTGCCGGGCGTCAAC
 GATTCGGCCGACAAGCTGATCGCGCTCTTCCAGGACAAGACCATCACGCCGCACGAGCTCGCCGCCCTGCTCGGC
 GCCCACACCACCTCACAGCAGTTCTTCGTGACACCACCGCGCCGCGGCCCGCCCGGAGACAGCACCCCGGCGTC
 TGGGACACCCGCTTCTACAACCAAACACCTCCGACCAAGTTCCCAAGAAGGTCTTCCGCTTCGCCAGCGACGTT
 GTGCTGGCCAAGGACCCGCGTATGAGTGATGAGTGGGCGCCTTCGCCGACCCCGTCAAGGGCCAGAACCACTGG
 AATGAGGATTACGCCACCGCCTATACCCGCCTCAGCCTGCTCGGCGTCAACAACATCAATAACTGAGTGTAGTC
 AGCAAGGTGCTGCCGTGGGCGCAACCGAAGTTCCCGATGCCTGCGGCTCAACAACATCAATAACTGAGTGTAGTT
 TCTTCTTAATATCTGGCGTTCTTGGGGTTTGGTCAAAAAGGAGAAAGATATTATTGTGATATTTGGTATACCC
 TAATGTTGGGAACATACATACGTATACAGGCTTCTTGTAATTACTCGGGCATCTTTTGCACAGAG

SEQ ID NO : 5
 CHIỀU DÀI : 1116
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 TÊN/KHÓA DẤU HIỆU : CDS
 VỊ TRÍ : (1).....(1116)

atgaagttctccacagtcacgtctcgagcgttgctctcacttctctactccagccccgccctt
 M K F S T V I S S V A L T S L L Q P A L
 gcctaccctggcatggcaaatgtcgtctcggagatcaaggcccggtcaaaacaccaataat
 A Y P G M A N V V S E I K A R Q N T N N
 gatggtgactcaaatccccgagatgattggcgatctcgccaccaccggcccaaccaccct
 D G D S N P E M I G D L A T T G P T T P
 gtggggccaaagcatataacaacatcctgatggggaccgagtcggccgagaccaagcaggct
 V G Q S I Y N I L M G T E S A E T K Q A
 ggctacatccccctcttatcggcaccaacgcctgcaagagggacacctgctgcatctgg
 G Y I P P L I G T N A C K R D T C C I W
 gcctacatcgccgccgaactgacctcaatttcaagggcatcacggggccgctgcaacaag
 A Y I A A E L T L N F K G I T G R C N K
 aacgcgcgtgccgccattcggtcgggttcacgacgcggggacttggtccaagagcagc
 N A R A A I R L G F H D A G T W S K S S
 aacggcgggcggcgcggacggctcgatcgcgctgtcgggcacggagatcaacaaggccgag
 N G G G A D G S I A L S G T E I N K A E
 aacaacgggctgcaggacatcatcggaagatgatcacgtggcagaagcggtagcggggtg
 N N G L Q D I I G K M I T W Q K R Y G V
 ggcatggcggtatctgatccagttcgcggccatccacgccgtggtgacgtgcccgtggtggg
 G M A D L I Q F A A I H A V V T C P L G
 ccgcgcacccgcttcttcgtcggggcgcaaggacagcaaaacggccaacgacgtcagcctg
 P R I R F F V G R K D S K T A N D V S L
 ctgcccggcgctcaacgattcgggccgacaagctgatcgcgctcttcaggacaagaccatc
 L P G V N D S A D K L I A L F Q D K T I
 acgcccgcacgagctcgccgccctgctcggcgcccacaccacctcacagcagttcttcgtc
 T P H E L A A L L G A H T T S Q Q F F V
 gacaccaccgcgcggcgccccccaggacagcaccgccggcgtctgggacaccgcgttc
 D T T R A G A P Q D S T P G V W D T R F
 tacaaccaaaccacctccgaccaagttcccaagaaggtcttcgcttcgccagcgacgtt
 Y N Q T T S D Q V P K K V F R F A S D V
 gtgctggccaaggacccgcgtatgagtgatgagtgggccgccttcgccgaccccgctcaag
 V L A K D P R M S D E W A A F A D P V K
 ggccagaaccactggaatgaggattacgccaccgcctataccgcctcagcctgctcggc
 G Q N H W N E D Y A T A Y T R L S L L G
 gtcaacaacatcaataacttgactgagtgacgaaggtgctgccgtggggcgcaaccgaag
 V N N I N N L T E C S K V L P W A Q P K
 ttccccgatgcctggaacctgttcttgaccagtag
 F P D A W N L F L D Q -

SEQ ID NO : 6
 CHIỀU DÀI : 371
 LOẠI : PRT
 SINH VẬT : *M. phaseolina*

MKFSTVISSVALTSLQPALAYPGMANVVS EIKARQNTNNDGDSNPEMIGDLATTGPTTPVGQSIYNILMGTESA
 ETQAGYIIPPLIGTNACKRDTCCIWAYIAAELTLNFKGITGRCNKNARAAIRLGFHDAGTWSKSSNGGGADGSIA
 LSGTEINKAENNLQDIIGKMITWQKRYGVGMADLIQFAAIAHAVVTCPLGPRIRFFVGRKDSKTANDVSLPGVN
 DSADKLIALFQDKTITPHELAALLGAHTTSQQFFVDTTAGAPQDSTPGVWDTRFYNQTTSDQVPKKVFRFASDV
 VLAKDPRMSDEWAAAFADPVKGQNHWNEDYATAYTRL SLLGVNNINNLTECSKVL PWAQPKFPDAWNLF LDQ*

SEQ ID NO : 7
 CHIỀU DÀI : 1323 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*

TCGAGATGGCGGCTCTCTCGCAGACGTAAATCTATAAAGCCTTGCCGCGACTTTCTTTTGTGTATCTCATCAGCA
 TCGAGTTCCACAAGAGCTTGCTTCCTCTCCTCACTGAGCCTAGCGAGGAGCATCACTCAGAACCCTCCAATCACA
 ATGCGGACCTCATCCCTGTTCTTGTCTCCGCATGCGGAACATCTGCTTACACTCTCGTCTCGCTCGACTCTCTG
 CCAAGCACTCTTCATGACATAACCTCCCGCACCATCTCCAACCTCGACCCCGCAACCTCCTCTCCGCGCGCAAG
 ACGCCCCGACTGCCCGGCCATCTGGAGGACCATCTCAGCCGACCTGACCAAGAGCTTCTCGCCAACGGCGAGTGC

ACCGACCTCGCCCGCGCCGCCATCCGCTACGCCTTCCACGACGCGGGCACCTTCTCGCTCAAGCTGCCACCTAC
GCGCCGGCCTCCGGCGGGCGCCGACGGCTCGCTGCTGCTCGTTCGAGATCCAGCGGCCCGAGAACACGGG
CTGCAGGCGTACAACGACTTCATCAAGGCCAAGTACAGCACGTACAAGTCTCGGGCGTCCGGCGCCGCCGACCTG
ATCCAGTTTCGCCGGCAACCACGCCGTGGTGACGTGCCCCGGGCGGGCCACGGTCAAGACGCTCGTTCGGCCGCGGC
GACAGCACGACCGCGTCCCGCTGAACGTGATGCCGCCGGGGTTTCGGCGCGGGCAGCGACCACGACTCGCTGCTC
CAGCTCTTCCAGGACAAGGGGTTTCAGCGCCGTGACCTGGCCGCGCTGATCGGCGCCACACCACCTCCACGAAC
ATCGCGGAGGCGCAGATCCCCGTGGCGCGCCGAGGACAGCACGCCGGGCAGGTGGGACGTCAAGTACTACGCC
GAGACGTACGCCCCGCGCGGGCGTCTCCCGCTTCGCCTCCGACATCAACCTCTCCGACCCGACGAAAGCGGTC
GGCAAAGAGTTCCAGGGATTTCGTCAACAACCAAGGGTAAGTGGACGGGCAAGTTTGCCGACGCCATGTTCCGTCTG
AGTGTGTTGGGCATCCCGCCGGCGACGTACAAGAAATTCGCGGACTGCACCGCTGCGCTGCCCCAAGGGCACGAGC
GCCAAGCGGGACATCCCGCAGCGCCCCGATCAACGACCGCGCAAGGTAGAGGAGGGGAAAAAGAAAGGAAGAAAAGA
AAATAAAAAAGCGCAGCGAGGATGAATCGGTTGGGGCAGGGCGTTTCGGTTGAGGTTGTTTGTTCGCTCGCCTGCC
TTTTTTTTTTTTTtAATCCCTCTCATGATCCATCGAATGAGAACACTT

SEQ ID NO : 8
CHIỀU DÀI : 1023
LOẠI : ADN
SINH VẬT : *M. phaseolina*
TÊN/KHÓA DẤU HIỆU : CDS
VỊ TRÍ : (1).....(1023)

atgcggacctcatccctgttccttgcttccgcatgcggaacatctgcttacactctcgtc
M R T S S L F L A S A C G T S A Y T L V
tcgctcgactctctgccaagcactcttcatgacataacctccccgcaccatctccaacctc
S L D S L P S T L H D I T S R T I S N L
gacccccgaacctcctctccgcgcgaagacgcccgaactgcccggccatctggaggacc
D P R N L L S A R K T P D C P A I W R T
atctcagccgacctgaccaagagcttcctcgccaacggcgagtgcaccgacctcgccgc
I S A D L T K S F L A N G E C T D L A R
gccgccatccgctacgccttccacgacgcgggcaccttctcgtcgaagctgccacctac
A A I R Y A F H D A G T F S L K L P T Y
gcgcccgcctccggcgccgcgcgacggctcgtgctgctcgtcgattcggagatccagcgg
A P A S G G A D G S L L L V D S E I Q R
cccgagaacaacgggctgcaggcgtacaacgacttcatcaaggccaagtacagcacgtac
P E N N G L Q A Y N D F I K A K Y S T Y
aagtcctcgggcgtcggcgcccgccgacctgatccagttcgccggcaaccacgccgtggg
K S S G V G A A D L I Q F A G N H A V V
acgtgccccggcgggccacgggtcaagacgctcgtcgccgcggcgacagcacgaccgcg
T C P G G P T V K T L V G R G D S T T A
tcgcccgtgaacgtgatccgcccgggttcggcgcgggcagcgaccacgactcgtgctc
S P L N V M P P G F G A G S D H D S L L
cagctcttccaggacaaggggttcagcgccgtcgacctggccgcgctgatcggcgccac
Q L F Q D K G F S A V D L A A L I G A H
accacctccacgaacatcgcgaggcgagatccccgtcgccgcgcgcgaggacagcacg
T T S T N I A E A Q I P V G A P Q D S T
ccgggcaggtgggacgtcaagtactacgcccagacgtacgccccgcccggggcgtctcc
P G R W D V K Y Y A E T Y A P P A G V S
cgcttcgcctccgacatcaacctctccgacccgacgaaagcggtcggcaaagagttccag
R F A S D I N L S D P T K A V G K E F Q
ggattcgtcaacaaccagggttaagtggacgggcaagtttgccgacgccatgttccgtctg
G F V N N Q G K W T G K F A D A M F R L
agtgtgttgggcatcccgccggcgacgtacaagaatttcgcggaactgcaccgctgcgctg
S V L G I P P A T Y K N F A D C T A A L
cccaagggcacgagcgccaagcgggacatccgcagcgccccgatcaacgaccgcgcaagg
P K G T S A K R D I R S A P I N D R A R
tag
-

SEQ ID NO : 9
CHIỀU DÀI : 340
LOẠI : PRT
SINH VẬT : *M. phaseolina*

MRTSSLFLASACGTSAYTLVSLDSLPLSTLHDITSRTISNLDPRNLLSARKTPDCPAIWRITISADLTKSFLANGEC
TDLARAIRYAFHDAGTFSLKLPTYAPASGGADGSLLLVDSEIQRPENNGLQAYNDFIKAKYSTYKSSGVGAADL
IQFAGNHAVVTCPGGPTVKTLVGRGDSTTASPLNVMPPGFGAGSDHDSLLQLFQDKGFSAVDLAALIGAHTTSTN

IAEAQIPVGAPQDSTPGRWDVKYYAETYAPPAGVSRFASDINLSDPKAVGKEFQGFVNNQGWTKGFADAMFRL
SVLGIPPATYKNFADCTAALPKGTSAKRDIRSAPINDRAR*

SEQ ID NO : 10

CHIỀU DÀI : 1930 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

GAGCACGGCAGGTAGGTTGCATCCGAGAGGTCGTGGTAGGTAGGTATATATTGCTGCCTCCGCTCTGTGCTCTA
TAGAGGCAGGTCTTGGACATCTGTGAACAGGTTACCTTCTCATAATTTGATCCATTGGCGAAGATATACTCAAGA
ATGCCGTCTACCTGGATGATTGCATTGGGCGCACTACCcTTGCCGGCCAGTCCGCAGCCTTTCTGTCTGTGGCA
GAACAATACGCTGCTCAGACATCTCACAAGGAGAGAAAAGAGTAAATGCAATTAGCCCCGGGTTCAATGCGGCGGCA
CAGAGGATTGACGTCTCGGGAGCTCACACTTTTCGTGCCCCCTGGCCCCGGTGACCAAGAGAGGGCCCTGCCCGGGT
CTGAATGCCCTTAGCGAACCAGTAAGTGCTCTGCAAGTCTCACAGGTCTTCTTGTACTGACAAAAGTGTAGCAACTA
TTTGCCCCAGTGAGTTCTCCACAACCTTCATTCCCGAAAACCTTGTCACTGATTGTCTTCAGCAATGGCGTCGCAAC
AATCACGCAGTTCTGTCCAGGCCACAAAACAGGGTTAGTACAGCATCACCGCCACGGAAAGCTACTACGTCAGCTGA
ACCACATACGCAGTATACGGCATGGGTCTCGACCTTGGCACGTTCTGTCCGTCTACGGCGCGGTAATGGATGGC
GACGGCTCAGCTGGTCCATCGGCGGCGCGCCAGTACCGCAAACCTACTCAACCTCCTTACTCAGCCGCAAGGC
CTCTCAGGTTTCGCACAACAAGTACGAGACAGATGCATCGCTACGCGCGGAGACCTGTACCAATAGTGCGTCTAT
CCCCTTCTTCTGCCTCCCGCCACTCCGCTGACCTGCACCGCAGTGGCGACAACCTCCCGGGTCTCATTTTCGCAAT
GGGACGCCCTCTTCGCGAAGCAAGCGGCACTGCCCAATGACCAGTCCAATTATGGCCTAGGTGTGCTGACCGACT
TCCGAGTCGAGCGTTTCCAGCAGAGCGTCGATGAAAACCCCTACTTCTTCAACGCGCCCTTTTCCGGTGTGCTCG
TGCAGCCTGCCGCTACACATTCATCTACCGCTTCATGGGCAACAAGAGCGCCGAGAAAACCCGAGGGTGTGCTGA
CGAAAGAGGTGCTGAAGAGCTTCTTCGGCTTCACTGGGCTTACGACGACATGACATACGTGAGTCCCATCCCCG
CGCCTTTCTCATTTCTTCCCACTCGAGTCTCCACATCCCCCCCCCCCCCCCCCCCCCCCCcACcATTTGTTGTC
GCCCTTCCCTCCAGCCACGCCACTGCCACAACGCCACCGTCTGACCCACCGCAGAACCCCGGCCACGAGCGCAT
CCCCGAAAACCTGGTACAAGCGCGCCCCCGGCGACGAATACACGATCCCCCTTCTACGCACTTGACCTCAACGCCGC
GGCGCTGCAGCACCCGCAATTCTGTCCGTGCGCGGCAACACGGGCACCACTCCTTCGCGGCGTTCGACCT
GCAGGACCTGAGCGGCGGCTGTACAACGCCGCCAGCCTGTCTCAGAGGCAACAACCTCGCCTGCTTCGGTTTCCA
GGCGGCCGTCCAATTTCGCGCCGACCTGTTGAAAGGGCTGGTGAGCGATTTGACCAAGCGGTTGGGTGCTTCCG
AGATCGGTTGGCGAGTGCCTTGAATGGATTAGGGTGTCCGCACTTAGGGCGGCGGCGTGGGATGATAGCGCGCT
GGCGCAGTTTCCGGGGTATGCTAGGCTGAGGGCGGATGGGACGTACGGGAAGTAATGGACGCGGAGAAAGTAGTGT
GGCAGAGGCGACATGGTGCATGGTGGACTGCGGCTTCAACGACAATGCAGAGATGATGGTATGAGGAACTGC
ACAGCAGTCTGCATGGATTGCTAGACTAGACGAATTTTATGATTTTGAATGAAC

SEQ ID NO : 11

CHIỀU DÀI : 1260

LOẠI : ADN

SINH VẬT : *M. phaseolina*

TÊN/KHÓA DẤU HIỆU : CDS

VỊ TRÍ : (1).....(1260)

atgccgtctacctggatgattgcattgggcgcaactcacccttgccggccagtcgcgcagcc
M P S T W M I A L G A L T L A G Q S A A
tttcctgctgtggcagaacaatacgtgctcagacatctcacaaggagaaaagagtaaat
F P A V A E Q Y A A Q T S H K E K R V N
gcaattagccccgggttcaatgcggcggcacagaggattgacgtctcgggagctcacact
A I S P G F N A A Q R I D V S G A H T
ttcgtgccccctggccccggtgaccagagagggccctgccccgggtctgaatgccttagcg
F V P P G P G G D Q R G P C P G L N A L A
aaccacaatggcgtcgcaacaatcacgcagttcgtccaggccacaaaccaggtatacggc
N H N G V A T I T Q F V Q A T N Q V Y G
atgggtctcgcaccttggcacgttccgtgtccgtctacggcgcggtaatggatggcgacggc
M G L D L G T F L S V Y G A V M D G D G
ctcagctggtccatcgggcgcgccccagtagccgcaaaccctactcaacctccttactcag
L S W S I G G A P S T A N L L N L L T Q
ccgcaaggcctctcaggttcgcacaacaagtacgagacagatgcatcgccctacgcgcgga
P Q G L S G S H N K Y E T D A S P T R G
gacctgtaccaatatggcgacaactcccgggtcgtcatttcgcaatgggacgcccctcttc
D L Y Q Y G D N S R V V I S Q W D A L F
gcgaagcaagcggcactgccaatgaccagtccaattatggcctaggtgtgctgaccgac
A K Q A A L P N D Q S N Y G L G V L T D
ttccgagtcgagcgtttccagcagagcgtcgatgaaaacccctacttcttcaacgcgccc
F R V E R F Q Q S V D E N P Y F F N A P
ttttccgggtgtgctcgtgcagcctgcccgcgtacacattcatctaccgcttcatgggcaac
F S G V L V Q P A A Y T F I Y R F M G N
aagagcgccgagaaacccgaggggtgtgctgacgaaagaggtgctgaagagcttcttcggc

K S A E K P E G V L T K E V L K S F F G
 ttactggtggcgctgacgacgacatgacatacaaccccgccacgagcgcatccccgaaaac
 F T G P D D D M T Y N P G H E R I P E N
 tggtagaagcgcgcccccgcgacgaatacacgatcccccttctacgcacttgacctcaac
 W Y K R A P G D E Y T I P F Y A L D L N
 gccgcggcgctgcagcaccgcaattcctgtccgctcggcggcaacacgggcaccaccaac
 A A A L Q H P Q F L S V G G N T G T N
 tccttcgcccgcgtcgacctgcaggacctgagcggcgccgtgtacaacgccgcccagcctg
 S F A G V D L Q D L S G G L Y N A A S L
 ctgcagggcaacaacctcgctgcttcggtttccaggcgccgtccaattcgcgccggac
 L E G N N L A C F G F Q A A V Q F A P D
 ctgttgaaaagggtggtgagcgatttgaccaagccgttgggtgtcctgggagatgcgttg
 L L K G L V S D L T K P L G V L G D A L
 gcgagtgcggtgaatggattagggtgtccgcagttaggcgccaggcggtgggatgatagc
 A S A L N G L G C P Q L G G Q A W D D S
 gcgctggcgagtttccgggtatgctaggctgagggcggtgggacgtacgggaagtaa
 A L A Q F P G Y A R L R A D G T Y G K -

SEQ ID NO : 12

CHIỀU DÀI : 419

LOẠI : PRT

SINH VẬT : *M. phaseolina*

MPSTWMIALGALTLAGQSAAFPVAEQYAAQTSHKEKRVNAISPFGNAAQRIDVSGAHTFVPPGPGDQRGPCPG
 LNALANHNGVATITQFVQATNQVYGMGLDLGTFLSVYGAVMDGDLWSISIGGAPSTANLLNLLTQPPQLSGSHNK
 YETDASPTRGDLQYGDNSRVVISQWDALFAKQAALPNDQSNYGLGVLTDFRVERFQQSVDENPYFFNAPFSGVL
 VQPAAYTFIYRFMGNKSAEKPEGVLTKEVLKSFEGFTGPDDDMTYNPGHERIPENWYKRAPGDEYTI PFYALDLN
 AALQHPQFLSVGGNTGTTNSFAGVDLQDLSSGLYNAASLLEGNLACFGFQAAVQFAPDLLKGLVSDLT KPLGV
 LGDALASALNGLGCPQLGGQAWDDSLAQFPGYARLRADGTYGK*

SEQ ID NO : 13

CHIỀU DÀI : 1108 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

CTGATGGTATGTACTTTGGTAAGCACGCCATTAACACTACTCGCTCCATCGCCTCGAGTTACCCTGTTGGAGCAGCA
 ACCACAACCTTGAGCGAACAATTCGGCCATTAATACCATCGTACGAAGAAGGCTTAACCGCTTTACCTGCCAGCC
 ATGAATCCAATCTGCTTTTTATCCCTGCTCACGGCCATGCTAGGCATGGCCATGGGCGGTGGAAACCCCTCTGAAC
 CATGCTGAACCTTTTGATCCACGAACAGCTTGTGGGTGTGACTGATGCGCATGAGCCCATTCTTCCTGATATCA
 CCAACGCTCGCAGCCCCGTGTCCCGGCCTTAACACGTTGGCGAACCAGGAAGCTCATTCGCCCCCTTGCTATGAAG
 AAAACCTGGAATTTACGCTCCTCCCGGTGGCTTCGACGGTGTGCGGAACACTACGATGAACCTGTTAAGGCTCTGGT
 TGATAGTGTGTTGCTTTTGATCGGGGGTCACCACCTTTCGGTTGACGAGAACATCACGATTCGCCCTAATAATTT
 CACACGCAGCCTCAGCGGCACACAAATCTCCCTAGATTCTGAGGCCCTCGCCTACACGCCACGACGCGTACGACCC
 TCGGGCTACTCCGGCTCGAGCAGCATCGAAATGAAGTGGAACCTTCTCAAACAGCTGTACGAGAAGCAGGCCGG
 GATCCACGCGATGCGGTAACTTTGCACCTGGATGTTCTTGCTCAAACAGCTGTACGAGAAGCAGGCCGG
 CAAGAACAACCCGAACCTTCTTCCTGAGCCCAACGACATCGCATTCGACCATCGACGCCCCACATGTGCATTC
 CAACTTGTTCGCGAACCACAGCACTGAGCATTTGGCTGACCTCTCTCCTGAATAGAGAGGGAGGGTCTCGCCATTCT
 ATTCCATGCTCAAACATATGACGCTTTGTAGTCTTCAGCTCAGATTGTCACCAATCTGAAGGCGAGATCCCACTTC
 TATACTCAGATTAAAGACCTAATAATACAATCGTATACTCAGCGAGAGAAGCAGAGAAGGACCGAGCCAACATT
 CACCTCATCTTCACACAACACACTCCTGAGCACCGACCCCAATTCTAAATCAGCCCC

SEQ ID NO : 14

CHIỀU DÀI : 690

LOẠI : ADN

SINH VẬT : *M. phaseolina*

TÊN/KHÓA DẤU HIỆU : CDS

VỊ TRÍ : (1).....(690)

atgaatccaatctgctttttatccctgctcacggccatgctaggcatggccatgggcgggt
 M N P I C F L S L L T A M L G M A M G G
 ggaaacctctgaacctgctgaaccttttgatcccacgaacagcttaaaacctggaatt
 G N P L N H A E P F D P T N S L K P G I
 tacgctcctcccgggtggcttcgacgggtgcggaactacgatgaacttggttaaggctctg
 Y A P P G G F D G V A N Y D E L V K A L
 gttgatagtgtgttgcctttgatcggggggtcaccacctttcggttgacgagaacatcacg
 V D S V L L L I G G H H L S V D E N I T
 attcgcctaataatttcacacgcagcctcagcggcacacaaatctccctagattctgag
 I P P N N F T R S L S G T Q I S L D S E

gcctcgcttacacgccacgacgcgtacgaccctcgggcctactccggctcgagcagcatc
 A S P T R H D A Y D P R A Y S G S S S I
 gaaatgaagtggaaacttcttcaaacagctgtacgagaagcaggccgggatcccacgcgat
 E M K W N F F K Q L Y E K Q A G I P R D
 gcggttaaactttgcactggatgttcttgcctcaaaacatgttggagctggtggtggtcagc
 A V L N F A L D V L A Q N M L E L V V S
 atcaagaacaaccgaaacttcttccctgagcccaacgcacatcgcatcgcgaccatcgacg
 I K N N P N F F L S P T H I A F G P S T
 gccacatgtgcattcccaacttggttcggaaccacagcactgagcattggctgacctct
 A H M C I P N L F A N H S T E H W L T S
 ctctgaatagagagggagggtctcgccattcattccatgctcaaactatgacgctttgt
 L L N R E G G S R H S F H A Q T M T L C
 agtcttcagctcagattgtcaccaatctga
 S L Q L R L S P I -

SEQ ID NO : 15
 CHIỀU DÀI : 229
 LOẠI : PRT
 SINH VẬT : *M. phaseolina*
 MNPICFLSLLTAMLGMMAMGGNPLNHAEPFDPTNSLKPGIYAPPGGFDGVANYDELVKALVDSVLLLLIGGHHLSV
 DENITIPPNNFTRSLSGTQISLDSEASPTRHDAYDPRAYSGSSSIEMKWNFFKQLYEKQAGIPRAVNFALDVLA
 QNMLELVVVSIIKNNPNFFLSPTHIAFGPSTAHMCIPNLFANHSTEHWLTSLLNREGGSRHSFHAQTMTLCSLQLR
 LSPI*

SEQ ID NO : 16
 CHIỀU DÀI : 1554 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 AAGCCTTGCTGGATCGACCTTTTCAATTCCGGCAGATACACCCAAGTCATTGAACATCGAAACAGTACTTGCCCT
 TCGTCAAATATGCGGTCACTCTTCTTGCTTCTTATTACTTTCTGCGGCGTCCGCTTTCCCTTTCTGTCGAGAT
 ATGCCAGAGGTAGACTCTTCTCTTCCGTGAAGCACCCGTACGTAGGCAACAACCTGGCGGCAACCAACCTGGC
 GGAGCGGCGACTTGCCCTTTCAATGCCAACACAGTCCCCGCTGCGCCAGTGACAGCTCGATTTCCCTATAACAAC
 GCAAAGAACGGAGTTCCCGGCAACGGAAAGGGCGGTTACCAGGTTCCAGCGCCTGGTGACACGGCTCATCAGTTC
 ATTGACACCAACAGCGCACGATATCCGTGGGCCTTGCCCGGGCCTGAACGCTGCGGCCAATCACGGCGTGAGTCTT
 TCCTGTAGTCCCCATTCGAGTTTCAAGTCTCCTGACGGTGACCCATAAACAGTTCCTCGCGCGGACGGCATAGTG
 ACCTTCAACGAACTGGTCGACGCCCAGCAGAATGTCTACAATGTCGGCTACGACCTCTCTGTGCTGCTCGCCGTC
 CTCGGCCTCACGCTCACCGACGGTGACCCCATCACCCAAAACTGTCTATCGGCTGCGACGCAACGACACGCACA
 TCTGTGGCCCCCTGCTGACTGGCAGTCAGCCCGGTCTGGATGGCCACAACAAGTTTCGAAGCGGACACGTCGCTC
 ACAGCAACGACTACTTCTGGCGGGCGGCGACAACCTTCAACTTCAACGGCACGCTCTTCGGCATGATGGTGGAT
 ACGTGCCAGGGCAACTTCAACCGTGAGAACCTGGCGCTGTACCGCAAGCAGCGCTACGACCAGAGCCTACGCGAC
 AACGAGAACTTCTACTTCGGCCCCGCTAAGCCTGCTGCTGTTCGGCGCCGCGCAGCTTCCCTTTACGAGCTGATGCC
 AGCGGCACGCACAACCTACGCGCCGATCTCGACACCATCTCGTCCCTTCTTCGGCGCCGAGCAGGCGCCCGATGGC
 TCCTGGCGCTTACCGCGGAGCGTATCCCGGACAACCTGGACCAACCGTGTGCTGCGGTACACCAATGAGGACGTG
 ACGCGCGAAATCCTGGCTATGTACCTCCTCAACCCCTGTGCTATTTGGCGGCGCCACAGGCGACGGCGGCTTCAAC
 ACGCTGCCGAAGTTTGGCTCCATCCAGGACGGCAAGATCGTTGAGGCACCCAATACGCTGTGCCTGCTGTACAG
 CTGTGACGCGAGAGCTGCCGAGCTCGCTGAATGGCATCATCACGCCGACTGTGGATGCGCTGAACCTGGTCGCG
 GATAAGCTAGCGCCGAGTTCAAGAACTTGGGGTGCCCGAATCCGTTGACTTGAATGAGTGATGGCCGTGTAGAG
 CCTCTGCTCGCGGTGCGAAGGAAAAAGAAAAAGTGAGGGGATTTTGGCGAGAAATGAAATTCACGAAGATACTG
 CCCGATGGTAGAGGAAACTTGGTCATGCATGTGCATGCGAAGGCGTTGATAT

SEQ ID NO : 17
 CHIỀU DÀI : 1194
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 TÊN/KHÓA DẤU HIỆU : CDS
 VỊ TRÍ : (1).....(1194)
 atgccagaggtagactcttctcttccgtgaagcaccgtagctaggcaacaacctggc
 M P E V D S S L F R E A P V R R Q Q P G
 ggcaaccaacctggcgagcggaacttgcccttcaatgccaccacgtcccgcgtgcg
 G N Q P G G A A T C P F N A N H V P A A
 ccagtgcagctcgatttccctataacaacgcaaagaacggagttcccggaacggaaaag
 P V T A R F P Y N N A K N G V P G N G K
 ggcggttaccaggttccagcgccctggtgacacggctcatcagttcattgcaccaacagcg

G G Y Q V P A P G D T A H Q F I A P T A
 cacgatataccgtgggaccttggccgggacctgaacgctgcggccaatcacggcttcctcgcg
 H D I R G P C P G L N A A A N H G F L A
 cgcgacggcatagtgaccttcaacgaactggtcgacgcccagcagaatgtctacaatgtc
 R D G I V T F N E L V D A Q Q N V Y N V
 ggctacgacctctctgtgctgctcgccgtcctcggcctcacgctcaccgacggtgacccc
 G Y D L S V L L A V L G L T L T D G D P
 atcacccaaaaactgtctatcggtcgacgcaacgacacgcacatctgtggccccctgtg
 I T Q K L S I G C D A T T R T S V A P L
 ctgactggcagtcagcccggtctggatggccacaacaagttcgaagcggacacgctcgctc
 L T G S Q P G L D G H N K F E A D T S L
 acacgcaacgactacttccctggcgggcgcgacaacttcaacttcaacggcacgctcttc
 T R N D Y F L A G G D N F N F N G T L F
 ggcatgatggtggatacgtgccagggaacttcaaccgtgagaacctggcgctgtaccgc
 G M M V D T C Q G N F N R E N L A L Y R
 aagcagcgctacgaccagagcctacgcgacaacgagaacttctacttccggcccgctaagc
 K Q R Y D Q S L R D N E N F Y F G P L S
 ctgctgctgttcggcgccgagcttccctttacgagctgatgcccgagcggcacgcacaaac
 L L L F G A A S F L Y E L M P S G T H N
 tacgcgccccgatctcgacacccatctcgctccttcttcggcgccgagcaggcgccccgatggc
 Y A P D L D T I S S F F G A E Q A P D G
 tcctggcgcttcaccgcccagcgctatcccggaactggaccaaccgtgtgctgcccgtac
 S W R F T A E R I P D N W T N R V L P Y
 accaatgaggacgtgacgcgcgaaatcctggctatgtacctcctcaaccctgtgctatctt
 T N E D V T R E I L A M Y L L N P V L F
 ggcgccgacgacgagcggcgttcaacacgctgccgaagtgtgctccatccaggac
 G G A T G G G C F N T L P K F G S I Q D
 ggcaagatcggttgaggcaccccaatacgcgtgtgcctgctgtaccagctgtcgacgcgagc
 G K I V E A P N T L C L L Y Q L S T Q S
 gtgccgagctcgctgaatggcatcatcacgccgactgtggatgcgctgaacctggctcgcg
 V P S S L N G I I T P T V D A L N L V A
 gataagctagcgccgcgagttcaagaacttggggtgcccgaatccggttgacttga
 D K L A P Q F K N L G C P N P L T -

SEQ ID NO : 18

CHIỀU DÀI : 397

LOẠI : PRT

SINH VẬT : *M. phaseolina*

MPEVDSSSLFREAPVRRQQPGGNQPGGAATCPFNNANHVPAAPVTARFPYNNAKNGVPGNGKGGYQVPAPGDTAHQF
 IAPTAHDIRGPCPLNAAANHGFLLARDGIVTFNELVDAQQNVYNVGYDLSVLLAVLGLTLTDGDPITQKLSIGCD
 ATTRTSVAPLLTGSQPLDGHNKFEADTSLTRNDYFLAGGDNFNFNGLTFGMMVDTCQGNFNRENLLALYRKQRYD
 QSLRDNENFYFGPLSLLLFGAASFLYELMPSGTHNYAPDLDTISSFFGAEQAPDGSWRFTAERIPDNWNTNRVLPY
 TNEDVTREILAMYLNPVLFGGATGDGGFNTLPKFGSIQDGKIVEAPNTLCLLYQLSTQSVPSNLNGIITPTVDA
 LNLVADKLAPQFKNLGCNPPLT*

SEQ ID NO : 19

CHIỀU DÀI : 1792 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

ACCGACCACCATCCATTCGCTCTTTTCGCTGGCGGCTGGCGTCCAAGGCTCACTCTCACTCTAGTCACTCTTTGCA
 ATCTCTGAATTCCTCGAGACCCCTTGGTCTGCGTTGTTAAGGATTTGTTGTCCGGGCATCCCCCTCTCTTTTCCGCA
 ATGAAGTTCTCGTCCGCCCTTCTCCTGCTCTCCTCGAGCTCCCTTGTCTCGATGCCTTCCCCGCCCTCGGTGCG
 CAGAACCTTGAGGGCCTCACTCCGGAAGGTTGACCGCTGCCCTCAAGACGGTCGAGAAAGTACAGGAAGGAGAAG
 CGCCTTATCATCGACTCCAGCAAGCTATCGATACCACCGGCGACCACGCTTTCCAGCCACCCAGCGAAACCGAC
 CAGCGCGGCCCTGTCTGGCTTGAACGCCCTGGCCAACCATGGCTACATCTCTCGCGACGGCATCACCAGCTTC
 GCCGAGGTTGTACCCGCCATTAAACCAAGGTTCTCGTGGATTCTGTCGAGAGAAGGCGTGGCTTGGCACTGACAAG
 GATGTGAAGTGATGGGCATGGGCATCGAGCTCTCTCTGATTCTCGGTGTTATGGGTACCGTGTGGACGGGTAACC
 CGCTTTTCGCTGGACCTGGCTTCTCTATCGGTGGGACCGCCCCCGGTGATGGCTCCGACAACATTTCTGGGCAACC
 TTGTCCGGCTGCTCGGTACGTAACGTGCCCTCATCGAGCCACGCCGCCGAGCTAAGCTGAGACGTCTGCAGGT
 GACCTTCGTGGTCTGCAAGGCTCCCAAACTGGATTGAGTCTGACTCCTCTCTGACGCGTGATGATCTGTACCTC
 ACCGGAGATGCCTGGACGATGAACATGACGCTCTCCGCGAGATCTACGACCGCGCGGATGAGGATGGCGTCATC
 TCCATGGATCTGCTCGCCGACCAGGCCGCCGCTGCTGGGAGTACAGCATCGGCCACAACCCCACTTCTACTAC
 GGCCCTGTGACCGGCATGGTCAGCCGTAATGCCGGCTACTTTTTCTCTCGGCCGCTGCTGTCAAACACACCGAT
 GAACATCCGGACGGAATTTCTCACTCAAGAAGTTTTCAAGAAGTTCTTCGCCGTCTACGAGGACGAGCAGGGCAAC
 ATGGAATACCGCAAGGGCCACGAGACCTTCCCGGACAAC'TGGTACCGCAAGCCGGTCGAGTATGGCCTGGTCCCG

CTCAACTTGGACCTCGTTGGCTGGGTCTTGAAGCACCCCTGAGCTGGGAAGGTACGTCGTCCCTTCTACCCCAAG
 ATGGGAAGGCATGTGAACTGACTCGGCTTCCACAGCATCGGCGGTAACACTGGCACCGTCAACTCCTTCTCCGG
 CCTCGATCTGCACAGCATCACCGGCGGCGTCTCAATGCCACTTCGCTCCTCGAGAACAACAACCTGCTGTGTTT
 TGTCTTTGAAGTTCTCAAGACCTTCGCCCCAACTCCCTCTCCCCGCTCCTGTGACGCTCGAAGTGCCGCTCAA
 GCTTATCGCCGACACCCCTGGCCACCCCGCTCTTGAGCCTGGCCTGCCCTGCCCTGGAAGGATATGACCGACGGTGG
 CGAGCCGCTGTGGGATGGCATTGAGAACAGGTTCCCTGGCGCGAGCAAGGCCGGATCGAGTTTGTAGAGTTGCTC
 GAGAGGACACAGGACGTCTGGAGCATGACGTCTGGGTGAACGTACTGCGTGGAAGAAGAGCAAAGGAACCAGCGA
 GGGCGAGAAATGATGTAGTTAGCGTCTTGATTCTAATTCGATACCATTACATTCTTCGCCTTATCC

SEQ ID NO : 20
 CHIỀU DÀI : 1317
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 TÊN/KHÓA DẤU HIỆU : CDS
 VỊ TRÍ : (1).....(1317)

atgaagtctcgtccgcccttctcctgctctcctcgagctcccttgtcgtcgatgccttc
 M K F S S A L L L L S S S S L V V D A F
 cccgccctcgggtgcgacagaaccttgagggcctcactccggaaagggtgaccgctgccttc
 P A L G A Q N L E G L T P E R L T A A L
 aagacggtcgagaagtacaggaaggagaagcgccttatcatcgactccagcaagcctatc
 K T V E K Y R K E K R L I I D S S K P I
 gataaccacggcgaccacgctttccagccaccagcgaaaccgaccagcgcgggccctgt
 D T T G D H A F Q P P S E T D Q R G P C
 cctggccttgaaacgccttgccaacctggctacatctctcgcgacggcatcaccagcttc
 P G L N A L A N H G Y I S R D G I T S F
 gccgaggttggtcaccgccattaaccaagtgatgggcatgggcatcgagctctctctgatt
 A E V V T A I N Q V M G M G I E L S L I
 ctcggtgttatgggtaccgtgtggacgggtaacccgctttcgtggtgaccctggcttctct
 L G V M G T V W T G N P L S L D P G F S
 atcggtgggaccgcccccggtgatggctccgacaacattctgggcaaccttgctcggcctg
 I G G T A P G D G S D N I L G N L V G L
 ctcggtgaccctcgtggctctgcaaggctcccacaactggattgagtctgactcctctctg
 L G D P R G L Q G S H N W I E S D S S L
 acgcgtgatgatctgtacctcaccggagatgcctggacgatgaacatgacgctcttccgc
 T R D D L Y L T G D A W T M N M T L F R
 gacatctacgaccgcgcggatgaggatggcgtcatctccatggatctgctcgccgaccag
 D I Y D R A D E D G V I S M D L L A D Q
 gccgcccgtcgctgggagtacagcatcgccacaaccccaacttctactacggccctgtc
 A A R R W E Y S I G H N P N F Y Y G P V
 accggcatggtcagccgtaatgccggtactttttcctcgccgcctgctgtcaaaccac
 T G M V S R N A G Y F F L G R L L S N H
 accgatgaacatccggacggaattctcactcaagaagttttcaagaagttcttcgccgtc
 T D E H P D G I L T Q E V F K K F F A V
 tacgaggacgagcagggcaacatggaataccgcaagggccacgagaccttcccgacaac
 Y E D E Q G N M E Y R K G H E T F P D N
 tggtagccgaagccggtcgagtatggcctgggtcccgctcaacttggacctcggtggctgg
 W Y R K P V E Y G L V P L N L D L V G W
 gtcttgaagcaccctgagctgggaagcatcgcggttaacactggcaccgtcaactccttc
 V L K H P E L G S I G G N T G T V N S F
 tccggcctcgatctgcacagcatcaccggcggtcctcaatgccacttcgctcctcgag
 S G L D L H S I T G G V L N A T S L L E
 aacaacaacctgctgtgttttgaagttctcaagaccttcgcccccaactccctc
 N N N L L C F V F E V L K T F A P N S L
 tccccgtcctgtcgacgctcgaagtgcggtcaagcttatcgccgacaccttgccacc
 S P L L S T L E V P L K L I A D T L A T
 ccgctcttgagcctggcctgcctgacctggaaggatgacccgacgggtggcgagccgctg
 P L L S L A C P A W K D M T D G G E P L
 tgggatggcattcagaacaggttccctggcgcgagcaaggccggatcgagttttag
 W D G I Q N R F P G A S K A G S S L -

SEQ ID NO : 21
 CHIỀU DÀI : 438
 LOẠI : PRT

SINH VẬT : *M. phaseolina*

MKFSSALLLLSSSSSLVVDAFPALGAQNLEGLTPERLTAALKTVEKYRKEKRLIIDSSKPIDTTGDHAFQPPSETD
QRGPCPLNALANHGYSISRDGITSFAEVVTAINVQVMGMGIELSLILGVMGTVWVGTPNPLSLDPGFSIGGTAPGDGS
DNILGNLVGLLGDPRGLQGSNHWIESDSSLTRDDLYLTGDWMTNMFLFRDIYDRADEDGVISMDLLADQAARRW
EYSIGHNPNFYYPVTGMVSRNAGYFFLGRLLSNHTDEHPDGILTQEVFKFFFAVYEDEQGNMEYRKGHETFPDN
WYRKPFVEYGLVPLNLDLVGWVLKHPGLSIGGNTGTVNSFSGLDLHSITGGVLNATSLLENNNLLCFVFEVLKTF
APNSLSPLLSTLEVPLKLIADTLATPLLSLACPAWKDMTDGGEPLWDGIQNRFPGASKAGSSSL*

SEQ ID NO : 22

CHIỀU DÀI : 1391 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

ATACTAACGCCGTGCAACCCCCGGGCAGCGCAGGTCTTGTTGATCGATATAAATGTTTTGTTTCGCGCCGCTATT
GAAATTGGACACTCCAACCTGTTTCCATACATTCATCTTCCCTATTGATCAATTGCTGCAAAAAACATTCCATC
ATGAAGCTCACCGTCTGTGCCACGACCCTTGTCTCGCTTTGGTTTATGGACAAGGCTCCTACGAGGGGTGGAAG
CCAGCTGGGCCAGACGACTGTATGTTTCACTGGAATTTCTACGTTTCCCTTGATTTGCGCTGATGTACATCTACAG
TTCGCGGCCCTTGCCCTATGATGAATACCTTGGCCAACCATGGCTTCTCCCCACGATGGCAGGAATATCACGA
AAGCCAACGCAATCCACGCTCTCAACACAGCTATCAACTTCAACACTTCCCTCGTGCTATCATGTGGGAGCAGG
CTATCATTGCAAACCCGGAGCCCAATGCTACGTTCTTCACTCTGTACGTGACTTCAAGCAAGCGTCTTTTGAGCT
ACCCTGAGTAGGGGGCTGCTCCGAAAGGCGACTTATCTAAACCTCCTCTTTCCACTATGTGAAGCTCTGACATTA
ACTAATATTGAACAGTGACCATCTTAATCGTCAACAACGCTCTTGGAGCACGATGCCAGCTTGAGGTGAGTGGAAGC
CGCCTTTCTCCAGCCTATTCCAACAAATCTCCTGACCAACACTCCTCTTCAGCCGATCCGACGCCCTTCTTCGGCA
ACAACCACGCTTCAACCAAACTATCTTCGACGCTCTCTCGCGCGTGGTGACGGAGGAAACCGTAGACGCCAAGA
TGCTGGCCAACAGCAAGTTGTTCCGGCAGATCGAGTCGCGAGCCGCCAACCCGAATTACACCTTCAACCAAACTA
CCGAGGCCTTTAGCTTGGGCGAGGTGGCTGCTCCCATCATCGTCTTCGGCGACCACGCGGCCGGCACCGTCAACA
GGAGCTCTGGTCGAGTACTTCTTCGGTGAGCCAAGACCTCAGAGACTTTGGTCTGCAAGGGTTCAAAATTACTAAT
TATGGGGTGGTCTTGCAGAGAACGAACGCTCCCCAGCCGAGTTGGGCTGGACTAAGCAGGCTAATGATGTGTCTC
TGGAGGTCATCCTGGAGATCCAGGACCTCGTCCGCAACGCGACCAACCTGATCACCAGTACCCCGCTGCCGGCAG
CGCCTCACAAGCGGGACCTGCACGCCCTTACAGCCTCTAGATACGAATATAAAGCTGGCATGAAGATCACCAGG
GGTTTTTGAAGAATTGTGTCTTCGGCGGGACGGGAGTGGTAGATTGAGCCTTGTGGGCTTCAGGTGTCTGGGGCA
TGGTTTGCTTTGGTCCTTTTACATGATTTCCCTAGCACGCC

SEQ ID NO : 23

CHIỀU DÀI : 780

LOẠI : ADN

SINH VẬT : *M. phaseolina*

TÊN/KHÓA DẤU HIỆU : CDS

VỊ TRÍ : (1).....(780)

atgaagctcaccgtctgtgccacgaccctttgcttcgctttggtttatggacaaggetcc
M K L T V C A T T L C F A L V Y G Q G S
tacgaggggtggaagccagctgggccagacgactttcgcggcccttgccctatgatgaat
Y E G W K P A G P D D F R G P C P M M N
accttggccaaccatggctttctccccacgatggcaggaatatcacgaaagccaacgca
T L A N H G F L P H D G R N I T K A N A
atccacgctctcaacacagctatcaacttcaacacttccctcgctgctatcatgtgggag
I H A L N T A I N F N T S L A A I M W E
caggctatcattgcaaaccggagcccaatgctacgttcttcaactcttgaccatcttaat
Q A I I A N P E P N A T F F T L D H L N
cgtcacaacgtcttgagcacgatgccagcttgagccgatccgacgccttcttcggcaac
R H N V L E H D A S L S R S D A F F G N
aaccacgtcttcaacaaaactatcttcgacgtctctcgcgctgggtggacggaggaaacc
N H V F N Q T I F D V S R A W W T E E T
gtagacgccaagatgctggccaacagcaagttgttccggcagatcgagtcgagccgccc
V D A K M L A N S K L F R Q I E S R A A
aaccggaattacaccttcacccaaactaccgagggccttttagcttgggagaggtggctgct
N P N Y T F T Q T T E A F S L G E V A A
cccatcatcgcttctcggcgaccacgcccggccacggtcaacaggagctctggctcagtagc
P I I V F G D H A A G T V N R S L V E Y
ttcttcgagaacgaacgcctcccgaccgagttgggctggactaagcaggctaatgatgtg
F F E N E R L P T E L G W T K Q A N D V
tctctggaggtcatcctggagatccaggacctcgctccgcaacgacgaacacctgatcacc
S L E V I L E I Q D L V R N A T N L I T
gatgccccgctgcccgcagcgcctcacaagcgggacctgcacgccccttacagcctctag
D A P L P A A P H K R D L H A P Y S L -

SEQ ID NO : 24
 CHIỀU DÀI : 259
 LOẠI : PRT
 SINH VẬT : *M. phaseolina*
 MKLTVCATTLICFALVYGQGSYEGWKPGAPDDFRGPCMNTLANHGFLPHDGRNITKANAIHALNTAINFNTSLA
 AIMWEQAIIANPEPNATFFTL DHLNRHNVLEHDASLSRSDAFFGNHVFNQTFIDVSRWWTEETVDAKMLANSK
 LFRQIESRAANPNYTFQTTEAFSLGEVAAPPIIVFGDHAAGTVNRSLEVEYFFENERLPTELGWTKQANDVSLEVI
 LEIQDLVRNATNLITDAPLPAAPHKRD LHAPYSL*

SEQ ID NO : 25
 CHIỀU DÀI : 1314 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 GGCCCCCGAAAGGGAGACAAAATACGAGATTCTCCTATTACGCGCCCAAGAATCGACAACACCGCCCATCAGCTA
 CCGCCGCCTTCGTTTACGTCGTCCGCAGCAGCAACGCCCCGACGCTACAGTCAGTCAACAGCCTCCAGCCGCCACG
 ATGCCCAACGCAGTGCATCTGTCAATGCTCGTCCTCACGCATGCGGTGCCGATCGCGGGTTATCCGGGCTGGGAA
 GCGCAGCAAGTTGTCGCGAGAAGCGCCCTCCAAGCACCACAACGACCTCTACGTGCCCCAACCAAGCCATCCGGTG
 CCGGGGAAAGTGCCGTACATCCCGGACGAGGAAGAGCACTACTTTGAGAAGCAGGTAAACGGCTCCGGCAATGGC
 TACTACCGGGCGGTCTCCTGCCCCGCGAGTCAACATCATGGCGAACAGGGGGCTACATCAGCCGCTCGGGCCGGGAC
 ATCAGCTACGAGGAGATAGCGATGGCATCGCGGGAGCTGTTCAACTTCGGCGACGACAACGTGAGCAGCGGCCCC
 TGCAGACCGGCAAAGTAAGCGGGCTCACCAACAATTCTTCTTGACAGATCATGATCGTGTGGGGCCAGCTTC
 GCGGCGCACCCGGGCGCGAGCGCATCGACCTCGACATGCTGGCCGACGACGCGGTGCAGCACATCACCAACTGC
 CCTGCGGCGCGGACGCGGACGGAACGCGCGCTGGGCGACAACGTGAACCTGAACACGACGCTGCTGGAGCAGCTG
 CTGGCGACGTCCAAGGACGGCGTCACGCTGACGCTCGAAGACGCGAGCCGAGCACCACCACCTGCGGCACAACCAG
 TCGCTGGCCGAGAACCCCGGCTTCCGCTTCAGCAACTCCGACGCCATCTGCTCGCTTGCAGTACGCCAACCTG
 TTCGGTATCCTGGGCCGGCAGGGCAAGCATGGGCTCAACACGCTGTATGTGGAAGACGTCAAGACCCCTGTTCTGTC
 GACGAAGACCTGCCGGACGATACGGCCGAGGGAGCTGCCGTATTTCTCGACCGAGGCGAACCACTACATCGAC
 CGTATGGCCCAACACATCGGCTTCGAGATCGAGCGGCCGTTCCCGGCCAACGACGCCGACCTGAAGGACATCGAG
 CCGGTGCAAGCCAGATTTGAAGTGGTGGACGGATGCTGAGCGGTGCAAAATTTAAAAAAAATTTTTTTTAGGCA
 GAAGTCAAGCATTTAATCGGGCTATACACACATCATTTCCGTGCAACAAAATCATTTCTACAGCTGTACCGCCTC
 GAAAAAGAATAACGGAGATTTAAAAAAAAC

SEQ ID NO : 26
 CHIỀU DÀI : 951
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 TÊN/KHÓA DẤU HIỆU : CDS
 VỊ TRÍ : (1).....(951)
 atgcccaacgcagtgcatctgtcaatgctcgtcctcacgcatgcggtgccgatcgcggt
 M P N A V H L S M L V L T H A V P I A G
 tatccgggtggaagcgcagcaagttgtcgcagaagcgccctccaagcaccacaacgac
 Y P G W E A Q Q V V A E A P S K H H N D
 ctctacgtgcccaacccaagccatccgggtgccgggaaagtgccgtacatcccgacgag
 L Y V P N P S H P V P G K V P Y I P D E
 gaagagcactactttgagaagcaggtaaacgggtccggcaatggctactaccggcggtcg
 E E H Y F E K Q V N G S G N G Y Y R R S
 tcctgcccggcagtcacatcatggcgaacaggggtacatcagccgctcgggccgggac
 S C P A V N I M A N R G Y I S R S G R D
 atcagctacgaggagatagcgatggcatcgcgaggagctgttcaacttcggcgacgacaac
 I S Y E E I A M A S R E L F N F G D D N
 atcatgatcgtgctggggcccgacttcgcggcgcccccggcgagcgcatcgacctc
 I M I V L G P S F A A H P G R E R I D L
 gacatgctggccgacgacgcggtgcagcacatcaccaactgccctgcggcgccgacgcgg
 D M L A D D A V Q H I T N C P A A P T R
 acggaccgcgcgctgggagacaacgtgaacctgaacacgacgctgctggagcagctgctg
 T D R A L G D N V N L N T T L L E Q L L
 gcgacgtccaaggacggcgtcacgctgacgctcgaagacgcagccgagcaccaccacctg
 A T S K D G V T L T L E D A A E H H H L
 cggcacaaccagtcgctggccgagaaccccggttccgcttcagcaactccgacgccatc
 R H N Q S L A E N P G F R F S N S D A I
 tgctcgcttgcgagtagccaacctgttcggtatcctggggccggcagggcaagcatggg
 C S L A Q Y A N L F G I L G R Q G K H G
 ctcaacacgctgtatgtggaagacgtcaagaccctgttcgctgcaggaagacctgccggac

L N T L Y V E D V K T L F V D E D L P D
 ggatacggccggagggagctgccgtatttctcgaccgagggcgaacaactacatcgaccgt
 G Y G R R E L P Y F S T E A N N Y I D R
 atggccccaccacatcggttcgagatcgagcgccgttcccgcccaacgacgcccgcacctg
 M A H H I G F E I E R P F P A N D A D L
 aaggacatcgagccgggtgcaagccagatttgaagtggtagcggtgctga
 K D I E P V Q A R F E V V D G C -

SEQ ID NO : 27

CHIỀU DÀI : 316

LOẠI : PRT

SINH VẬT : *M. phaseolina*

MPNAVHLSMLVLTHAVPIAGYPGWEAQVVAEAPSKHHNDLYVPNPSHPVPGKVPYIPDEEEHYFEKQVNGSGNG
 YYRRSSCPAVNIMANRGYISRSGRDISYEEIAMASRELFNFGDDNIMIVLGPFAAHPGRERIDLMLADDAVQH
 ITNCPAAPTRTDRLGDNVNLNTTLLLEQLLATSKDGVTLTLEDAAEHHHLRHNQSLAENPGFRFSNSDAICSLAQ
 YANLFGILGRQGHGLNTLYVEDVKTLFVDEDLPDGYGRRELPHYFSTEANNYIDRMAHHIGFEIERPFPANDADL
 KDIEPVQARFEVVDGC*

SEQ ID NO : 28

CHIỀU DÀI : 1480 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

ATCTGGCTTTTCTCTCTCTCTGTGCGGGCGTCCCTTGCGCAACAGATACCCCATTCATCCCTTCAGTGCTTCTT
 CTTATCTCTTTCGCGCGCTTTTCCCAACACAGCCTCTTATCGCCTGTGCTCACACCACGTTTCCCGCACACCCA
 ATGGCTTCCACCGCCAGGTCCGTCTTCGCCCCGAGCGCGCTTCTGCGCTCCGCCCCGGCCTCCATCAAGTCGAAT
 GCCGCCCGCTCATCTCGCTTCGCGTCCCTACCCAAGCATTCGCCAGCAGTCTCGCCGCGGCTACTCTTCCGAG
 GCCGGCTCCAAGTCTAACGGCCCCAACCCCGCCATCTGGATTGGTGCTCTCGCCGTCTGGGTGGCGCCGGGTAC
 TATGCCTACAGCTCGGGTGCCGGCGCCAGATCGCTCCCAAGGAACCCTTCACCCCCAAGCCCGAGGACTACCAG
 AAGGTCTACGACGCCATCGCCAAGGCCCTCGAAGAGCACGACGACTACGACGACGGCAGCTACGGCCCTGTCTTG
 CTGAGACTGGCTTGGCACGCCAGCGGAACGTGAGTGACTTCCCAACACTTCCAGCCCACCATTTGAACCACGCAC
 TGACGCCCTCCCTCACAGCTACGACAAGGAAACCGGCACCGGCGGCTCCAACGGCGCCACGATGCGCTTCGCGCC
 CGAGGCGGACCACGGCGCCAACGCCGGCTCAAGGCGGCGCGGCTTCCTCGAGCCCATCAAGCAGCAGTTCCC
 GTGGATTACGTACTCGGACCTGTGGACGCTGGCAGGCGTCGCTGCGATCCAGGAGATGCAGGGCCCCAAGGTGCC
 GTGGCGCCCCGGCGCAGCGATCGCGACGTCTCTTCTGCACGCCCCGACGGCCGCTGCCCGACGCCTCCAAGGA
 CCAGAACCACCTCCGCGCCATCTTCGGCCGCATGGGTGGAATGACCAGGAGATCGTGGCGCTGTGGGCGCGCA
 TGCGCTGGGGAGGTGCCATACGGATAGGAGTGGATTTCGATGGCCCGTGACCTTCAGCCCACGACGCTGACGAA
 CGATTATTTCAAGTTGTTGATCGACGAGAAGTGGCAGTGGCGGAAGTGGGATGGACCTAAGCAGTTGGAGGACAA
 GAAGACGAAGAGCCTGATGATGCTGCCGACGGATTACGCGTTGGTGCAGGACAAGAAGTTTAAGCCCTGGGTGCA
 GAGGTACGCGAAGGATCAGGATGCCCTTCTTCAAGGACTTCTCGAACGTGGTTATGAGGTTGTTTCGAGCTGGGCGT
 GCCGTTCCAGAGTGGTGAGGACTCGAGGATTGTGTTTTAAGAGCAGCTTCGACTAGGCGTGCTTGGCGGACGTTAA
 TTCTGATCGACGGGGTTTGATGGGAAGTTGTAAAAGGTTCTATGACGATCAGTAAAGAAGGTTGTTTTTGCTTT
 TGAGTTTCGAGGACTAAAGACTAAGACAAGAGTAGCGCAAAGGTGGGAAAGAATA

SEQ ID NO : 29

CHIỀU DÀI : 1116

LOẠI : ADN

SINH VẬT : *M. phaseolina*

TÊN/KHÓA DẤU HIỆU : CDS

VỊ TRÍ : (1).....(1116)

atggcttccaccgccaggtccgtcttcgcccgcagcgcgcttctgcgctccgccccggcc
 M A S T A R S V F A R S A L L R S A P A
 tccatcaagtcgaatgccgcccgtcatctcgcttcgcccgtccctacccaagcattccgc
 S I K S N A A R S S R F A V P T Q A F R
 cagcagctctcgccgcggtactcttcgagggccggtccaagtctaacggccccaacccc
 Q Q S R R G Y S S E A G S K S N G P N P
 gccatctggattggtgctctcgccgtcctgggtggcgccgggtactatgcctacagctcg
 A I W I G A L A V L G G A G Y Y A Y S S
 ggtgccggcgcccagatcgctccaaggaacccttcacccccaagcccaggactaccag
 G A G A Q I A S K E P F T P K P E D Y Q
 aaggctctacgacgccatcgccaaggccctcgaagagcacgacgactacgacgacggcagc
 K V Y D A I A K A L E E H D D Y D D G S
 tacggccctgtctgctgagactggcttggcacgccagcggaacctacgacaaggaaacc
 Y G P V L L R L A W H A S G T Y D K E T
 ggcaccggcggtccaacggcgccacgatgcgcttcgcccggaggcgaccacggcgcc
 G T G G S N G A T M R F A P E A D H G A

aacgcgcgcctcaaggcgccccgcgacttcctcgagcccatcaagcagcagttcccgtagg
 N A G L K A A R D F L E P I K Q Q F P W
 attacgtactcggacctgtggacgtggcaggcgctcgctgcgatccaggagatgcagggc
 I T Y S D L W T L A G V A A I Q E M Q G
 cccaaggtgccgtggcgccccggccgcagcgatcgcgacgtctctcttctgcacgccccgac
 P K V P P W R P G R S D R D V S F C T P D
 ggccgcctgccccgacgcctccaaggaccagaaccacctccgcgccatcttcggccgcatg
 G R L P D A S K D Q N H L R A I F G R M
 ggttgggaatgaccaggagatcgtagggcgctgtagggcgcgcatgtaggggaggtgccat
 G W N D Q E I V A L S G A H A L G R C H
 acggataggagtgattcgatggcccgtagaccttcagccccgacgacgtgacgaacgat
 T D R S G F D G P W T F S P T T L T N D
 tatttcaagttgttgatcgacgagaagtggcagtgccggaagtgggatggacctaaagcag
 Y F K L L I D E K W Q W R K W D G P K Q
 ttggaggacaagaagacgaagagcctgatgatgctgccgacggattacgcgttggtgcag
 L E D K K T K S L M M L P T D Y A L V Q
 gacaagaagtttaagccctgggtcgagaggtacgcgaaggatcaggatgccttcttcaag
 D K K F K P W V E R Y A K D Q D A F F K
 gacttctcgaacgtggttatgaggttggtcgagctgggcgtgccgttccagagtggtgag
 D F S N V V M R L F E L G V P F Q S G E
 gactcgaggattgtgtttaagagcagcttgcgactag
 D S R I V F K S S F D -

SEQ ID NO : 30

CHIỀU DÀI : 371

LOẠI : PRT

SINH VẬT : *M. phaseolina*

MASTARSVFARSALLRSAPASIKSNAARSSRFVPTQAFRQQSRRGYSSEAGSKSNGPNPAIWIGALAVLGGAGY
 YAYSSGAGAQIASKEPFTPKPEDYQKVYDAIAKALEEHDDYDDGSYGPVLLRLAWHASGTYDKETGTGGSNGATM
 RFAPEADHGANAGLKAARDFLEPIKQQFPWITYSDLWTLAGVAAIQEMQGPVPWRPGRSDRDVSFCTPDGRPLD
 ASKDQNLHRAIFGRMGWNDQEIVALSGAHLGRCHTD RSGFDGPWTFSP TLTNDYFKLLID EKWQWRKWDGPKQ
 LEDKKT KSLMMLPTDYALVQDKKFKPWVERYAKDQDAFFKDFSNVVMRLFELGVFPFQSGEDSRIVFKSSFD*

SEQ ID NO : 31

CHIỀU DÀI : 1981 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

TAGAAGGCGCATGGAGTCGCACGCAGGGACCATGCGCAACACGGTGCCTGCAGAGAGGCGAAAAGCCCATGCATAT
 CACATTGATTTTAAAGAAGGAAGTGCAAGAAGCATTGAGACATCGCACTCACCCCTCAACACCGCTATCGCGCG
 ATGAAAGCCATCAGCATCGCCGTGTTTCGCCGCGTCAGTGGTACCTCACACTCTGGCAGACTATGTTTGGCCCTCG
 CAACATGACTTGTGAGGATATGTTGGCCATCCAACAAGGCTATATTCGCATGGGCTTCACTGACTGTATGTCG
 ACATTCACTTGAGCCTGAAGTAATGTCTCCCTACTAAAGCATCGTCGAAGTGGTGTGTTGCCGTACGGTCACGGCAG
 CAACAGGCCCCGGAGTAAACAACGCGGCACAATGGATTTCGGACGGGTTTCCACGACTTCGCCACCCACGACTCGGC
 CGCTGGGACCGCGCGCCCTCGATGCCTCCCTCCTCTACGAGGTGCAACGGGCCGAGAACGAGGGCTCAGCCTTTAA
 CGACACCTTCGCCGACATGCACGACTTTATCAATCCCCGATCAAGCGCCTCGGATCTGATAGCGCTGGCCGTCGT
 CGCATCGGTTGCGGCTTGCGGCGGACCCAAGATCCCCCTGCGAGCGGGACGAATTGATGCTGTGGAGGCGGGCCC
 GGCCGGCGTTCCCAAGCCCGACGACTCACTGGAGAGCACGATCGATGCATTCGCGCGGACGGGGTTCAACACCAG
 CGACATGATCGCCCTCGTTGCCTGTGGCCATACCGTCGGCGGCGTGCACAGCGTAGATTTCCCCGAGATCACCGG
 CGGCGAGAAAAGACGTCTTGACGTGCCGCAATTTGACAGCAGCGGCACCATCTTCGACACCCGAGTCGTGGACGA
 ATACCTCGACAGCAACGGCGCCAACCCCTCGTCTTCGGAGCCAACGACACCACGAACTCGGACAAGCGCGTCTT
 CAGCGCCGACGGCAACAGCACCATGGCCAAGCTCAAAGACCCCGCCACATTCAAAGCCACCTGCGCCGCCCTCTT
 CGAGCGTATGATCAACACCGTCCCCCTCATCCGTACCCCTCAGCGAACCATCGAGCTGGCCGACATCAAGCCCTA
 CATCGACAAGCTCGAGCTCACACCCAACGCTCCGCCCTCGCCTTCGAAGGCCGATCCGGCTGCGCACCTCCCC
 CGTCAACCGGCCGCGACGCCGAGGGCACAGCATCGCCCTCAATGTCAACGACCGCGCTGGCGGGCGCAAGCTGGT
 GCCGGCGCCGCGCGCCGTGTTGCGCGGCGGCACCTCGTACGGCTTCTTCGACGAGCAGTTTCACTGGTTTCAATT
 TGCGACGCGAGCTCGACGTCGACGCCGGCATCCAGGCTTTGATATCCAGCTCACGACGAGGCGACCGGACACGT
 GGAGACGTTTCGACAACGCTGGCACCGGCGGCTATCCTAGTCTCGACGACTTGCTATACTTGCACTCGCAATCGTG
 CATGGACACGACTGCCACGGAGGGGAATATAACGGTGACGGTTGCGGCGGCGGTGCGCGAGGATCGCGCGAAGG
 GGGTGCGGCGCCAGTGGTGAGGATGGCGCATAAAGTTGACGATGGGTGTGATGTTGCCGAAGCTTGTGCTAGAG
 GGCAGTGCCAATGGAGAGGTCAAACGTTTCGACGGGCGGGTATGTATTGTACGAGGTGGACATCCCAATCGATGC
 CGCTGGGCTGGAGCAGCAAGTTTCGATGTTGTGTTGACGGCTGGGGGAGACGAGATTGTGTCTGGCCTGCACGGGAC
 CAGCGATCTTACCACATGTTTCGGGAAACTGACCGGTTGATGGGCTTGGGGACTTGCCCTGAGTAGCATAGCATCG
 GCGTGTTTTGGGTTTGTGTTTACAGGTGTGGAAGATCAGATGAAGGAAGAACCAACAGGTGATTGAACCTATGTCGC
 AAGTCAGCAATGCAGTGATGGTCTCGGCGTC

SEQ ID NO : 32
 CHIỀU DÀI : 1623
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 TÊN/KHÓA DẤU HIỆU : CDS
 VỊ TRÍ : (1).....(1623)

atgaaagccatcagcatcgccgtgttcgcccgcgtcagtggtacctcacactctggcagac
 M K A I S I A V F A A S V V P H T L A D
 tatgtttggccctcgcaacatgacttggttgaggatatgttggccatccaacaaggctat
 Y V W P S Q H D L L E D M L A I Q Q G Y
 attcgcatgggcttcactgacttggttggttccttgcggtcacggcagcaacaggccccgga
 I R M G F T D L V V P C G H G S N R P G
 gtaaacaacgcggcacaatggattcggaacgggtttccacgacttcgccacccacgactcg
 V N N A A Q W I R T G F H D F A T H D S
 gccgctgggacggcgccctcgatgcctccctcctctacgaggtcgaacggcccgagaac
 A A G T G G L D A S L L Y E V E R P E N
 gagggctcagcctttaacgacaccttcgccgacatgcacgactttatcaatccccgatca
 E G S A F N D T F A D M H D F I N P R S
 agcgctcggtatctgatagcgctggccgctcgtcgcatcggttgcggttgccgagccccc
 S A S D L I A L A V V A S V A A C G G P
 aagatccccctgagcgaggacgaattgatgctgtggaggcgggcccgccggcggttccc
 K I P L R A G R I D A V E A G P A G V P
 aagcccgacgactcactggagagcacgatcgatgcattcgcgcggaacgggttcaacacc
 K P D D S L E S T I D A F A R T G F N T
 agcgacatgatcgccctcggttgctgtggccataaccgtcgccggcggtgcacagcgtagat
 S D M I A L V A C G H T V G G V H S V D
 ttccccgagatcaccggcgaggagaaagacgtcctggacgtgccgcaatttgacagcagc
 F P E I T G G E K D V L D V P Q F D S S
 ggcaccatcttcgacaccgcagtcgtggacgaataacctcgacagcaacggcgccaaccccc
 G T I F D T A V V D E Y L D S N G A N P
 ctgctcttcggagccaacgacaccacgaactcggaacaagcgcgctcttcagcgccgacggc
 L V F G A N D T T N S D K R V F S A D G
 aacagcaccatggccaagctcaaagacccccgccacattcaaagccacctgcgcccgcctc
 N S T M A K L K D P A T F K A T C A A L
 ttcgagcgatgatcaaacaccgtccctcatccgtcacccctcagcgaacccatcgagctg
 F E R M I N T V P S S V T L S E P I E L
 gccgacatcaagccctacatcgacaagctcgagctcacacccaacgcctccgcccctcgcc
 A D I K P Y I D K L E L T P N A S A L A
 ttcgaaaggccgcatccggctgcgcacctcccccgctcaccggccgcgacgccgagggcacc
 F E G R I R L R T S P V T G R D A E G T
 agcatcgccctcaatgtcaccgaccgcgctggcgggcgcaagctggtgccggcgccgcgc
 S I A L N V T D R A G G R K L V P A P R
 gccgtgttgcgggcgccacctcgtacggcttcttcgacgagcagttcagctggttcgaa
 A V L R G G T S Y G F F D E Q F S W F E
 tttgcgacgcagctcgacgtcgacggcatccaggcctttgatatccagctcacgacg
 F A T Q L D V A A G I Q A F D I Q L T T
 gaggcgaccggacacgtggagacgttcgacaacgctggcaccggcggtatcctagtctc
 E A T G H V E T F D N A G T G G Y P S L
 gacgacttgctatacttgcatcgcaatcgatgcatggacacgactgccacggaggggaat
 D D L L Y L Q S Q S C M D T T A T E G N
 ataacggtgacgggttcggcgccggtgcgcgaggatgcggcgaaggcggtgcggcgcca
 I T V T V A A A V R E D A A K A G A A P
 gtggtgaggatggcgcataaggttcagcagatgggtgtgatgttgccgaagcttgctcgt
 V V R M A H K V Q Q M G V M L P K L V V
 gaggcagtgccaatggagaggtcaaacgtttcgaggcggggtatgtattgtacgaggtg
 E A V P M E R S N V S Q G G Y V L Y E V
 gacatcccaatcgatgccgctggctggagcacgaagttcgatgttggtgacggctggg
 D I P I D A A G W S T K F D V V L T A G
 ggagacgagattgtgtctggcctgcacgggaccagcgatcttaccacatgttcgggaaac
 G D E I V S G L H G T S D L T T C S G N
 tga

-

SEQ ID NO : 33
 CHIỀU DÀI : 540
 LOẠI : PRT
 SINH VẬT : *M. phaseolina*
 MKAISIAVFAASVVPHTLADYVWPSQHDLLDMLAIQQGYIRMGFTDLVVPCHGHSNRPGVNNAAQWIRTGFHDF
 ATHDSAAGTGGLDASLLYEVEPENEGSAFNDFADMHDFINPRSSASDLIALAVVASVAACGGPKIPLRAGRID
 AVEAGPAGVPKPDDSLSTIDAFARTGFNTSDMIALVACGHTVGGVHSVDFPEITGGEKDVLDVPQFDSSGTIFD
 TAVVDEYLDNSGANPLVFGANDTTNSDKRVFSADGNSTMAKLKDPATFKATCAALFERMINTVPSSVTLSEPIEL
 ADIKPYIDKLELTPNASALAFEGRIRLRTSPVTGRDAEGTSIALNVTDRAGGRKLVPAPRAVLRGGTSYGFFDEQ
 FSWFEFATQLDVAAGIQAFDIQLTTEATGHVETFADNGTGGYPSLDDLLYLQSQSCMDTTATEGNITVTVAAAVR
 EDAAKAGAAPVVRMAHKVQQMGVMLPKLVVEAVPMERSNVSQGGYVLYEVDIPIDAAGWSTKFDVVLTAGGDEIV
 SGLHGTSDLTTCSGN*

SEQ ID NO : 34
 CHIỀU DÀI : 2012 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 GGGGGAGGCCAATGCATGTCAGCAAAATACCTGTGGCTACCCCTGTTATGGCCACATCCAGCCCATTTTGAATGG
 ACGGTATAAAATCTCTTCTGAAACCTTCAGAAAATTAAGGCAGGGACTATTGAAGGTTTTGCCGTATCTGCCAAA
 ATGCGGTTTTCTCGGGGGCTTCTATCTACTCCTTGCTGCCACTCGCCACACACCTACGGCCAGAGCTGCGCTGCAC
 TACCCAAATGCCCTCATTTCTCGCATGGAACACCTGCTGGTTGATACGGACGGCTCATTCAGGTCTGGTTTCAAG
 GACGCCATCAACCCCTGCACGAACTACATTTAGGTGCCCAAACGCTGGGGCGGCAGACCTCTGCACAGTGGCTC
 CGAGTTGCCCTTCCATGATTTTCGTACGGCCCCACGTGACGAGGGCACTGGCGGGATCGACGCGTCCATCGGCTTC
 GAAACTCTCCGAGCAGAAGATTCCGGGTCCGCATTCAATGACAGTTTCGCCTTCTTTGCTCCGTTCGTGGATGCG
 CAAACTTCCAGTATCACCCATCGAGCCTTCCCCGTCGAGGGCCGTGGCATGCGCTGACGTCCTTGCACTGGCCGA
 CCTCGTAGCCCTCTCTGTGGTGACTTCTCTGGGCCACTGCGGTGGTCTGCATGTCCATATCGAGCGGGCCGTAT
 CGATGCTACGGGCGGAGGGCCGTTCGGCGTTTCCCGAACCCGAGACAAGCCTGGAAGAGACCTTGAAGAGTTTGC
 CAATGCTGGTTTCAATGCTGAAGATGCCATTGGATTAACGGCGTGCAGGGCATTCTCTCGGCCCGCTCCATCACGG
 CGGGTTCCCCAACGTGCTGCCCGAATCGGCCATAGCACCAAACAACACCGCGGGCGGCGTGAACCTGGACTCCAC
 ACGGGATAAAATTCGACATCAGCATTGTCAAAGAATACCTCGGCAACTATGGGCAGCGCGGCGGACCTCTCGTTAC
 CAGTGACAACGTGACTGTCCGCTCGGATCTTCGGCTGTACGAAAGCGATCAGAACAGGACAATGCAAGCTCTCGG
 TCAGTCAAAGAATACTTCTTTAGCACCTGTGGAAATCTATTTGAGAGGATGATCAACACCGTTCCGCGCGAGGT
 CACTTTGTCAGATGTCATCCATCCGATGACGGTGCAGCCGGTGAATTTACGTTTCGATATCATCAATGACCAGGC
 GCTGAGGTTATCAGGAGTAGTGCGGGTGAGCTTCCAAATAGAGCTGTTTCGGTCTGCACGCAACTGACCGAGCCC
 CGCAGTATTTGCCCTCAGATAATGCTGCACCGTCGACGCTCGAGGTCTCACTCGCCGACAAAGCCGGAAGACTA
 TGGCATCCATTACAGCAAGGTTATTGAAGAGAAAAGGGAATAGTTTCTGGGGGGCCACAGCTTACTACCCGGTCCG
 TTTTGTACATAAACCTTGCTGGGATTGCTTCTCGCAACAATCTTCTGGGAAGCTCCAGGTTTCGCACAGCATCTC
 CTCAGACATTTGAGCTGCAACCGGAGCTGTTTTTCATACCATCCCGTTCTAGCCCTGGCACGGGGATAAGTGTGG
 GGGCTGCCATAGGTGCTGCTGCGCGGTCCCGCAACTCGACTCTGAGTGATAGAGAGTGTTGAAGCCGTGGTCAGTG
 TGCCGGTATCCCGACGGGAACACTGGCACCGAAGGTCGAGAAGCATGAGTTGAATCTCGAAAAGAGATCAAGATA
 TTGGTTTGTACTCTATATTCAAGGGGAACCTAAGCGATGATTTGGCTTTCAACCTGCAGACGACCGTTGATATCA
 CGGCAACTTTTTTCAGACGGCACTGTCTGGCAAGATAGTTACAACAGATTTGCAGTGCCGTAGAGCATCCTGAAAA
 TGTCTTCGTCAAAATGCCTTCATAGGGCCAGCAACAAATTTTCAGGCAGATATGTAGCTAGAGAGTCATTCAAAT
 ACAAATGTCTTGTTGTGCAACAAGTTCGTGAGACGCTGCTCATGCACAATACAAGTGAGCC

SEQ ID NO : 35
 CHIỀU DÀI : 1599
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 TÊN/KHÓA DẤU HIỆU : CDS
 VỊ TRÍ : (1).....(1599)
 atgcggtttctcgggggcttctatctactccttgctgccactcgccacacacctacggcc
 M R F L G G F Y L L L A A T R H T P T A
 agagctgcgctgcactacccaaatgccctcatttctcgcatggaacacctgctggtgat
 R A A L H Y P N A L I S R M E H L L V D
 acggacggctcattcaggtctggtttcaaggacgccatcaacccctgcacgaactacatt
 T D G S F R S G F K D A I N P C T N Y I
 tcaggtgcccaaacgctggggcggcagacctctgcacagtggctccgagttgccttccat
 S G A Q T L G R Q T S A Q W L R V A F H
 gatttcgtcacggcccacgtcgacgagggcactggcgggatcgacgcgtccatcggttc
 D F V T A H V D E G T G G I D A S I G F
 gaaactctccgagcagaagattccgggtccgcattcaatgacagtttcgccttcttgc
 E T L R A E D S G S A F N D S F A F F A
 ccgttcgtggatgcgcaaaccttccatggccgacctcgtagccctctctgtggtgacttct

P F V D A Q T S M A D L V A L S V V T S
 ctggggccactgcggtggtctgcatgtcccatatcgagcgggccgtatcgatgctacgggc
 L G H C G G L H V P Y R A G R I D A T G
 ggaggggcccgttcggcggttccccgaacccgagacaagcctggaagagaccctggaagagttt
 G G P F G V P E P E T S L E E T L E E F
 gccaatgctggtttcaatgctgaagatgccattggattaacggcggtgcgggcattctctc
 A N A G F N A E D A I G L T A C G H S L
 ggccgcgtccatcacggcggttccccaacgtcgtgcccgaatcgccatagcaccaaac
 G R V H H G G F P N V V P E S A I A P N
 aacaccgcggcggtgaacctggactccacacgggataaattcgacatcagcattgtc
 N T A G G G V N L D S T R D K F D I S I V
 aaagaatacctcggcaactatgggcagcgcggcgacctctcgttaccagtgcacaacgtg
 K E Y L G N Y G Q R G G P L V T S D N V
 actgtccgctcggtatcttcggctgtacgaaagcgatcagaacaggacaatgcaagctctc
 T V R S D L R L Y E S D Q N R T M Q A L
 ggtcagtcaaaagaataacttcttttagcacctgtggaaatctattttgagaggatgatcaac
 G Q S K E Y F F S T C G N L F E R M I N
 accgttccgcgcgaggtcactttgtcagatgtcatccatccgatgacggtgcagccggtg
 T V P R E V T L S D V I H P M T V Q P V
 aatttcacgttcgatcatcaatgaccaggcgctgaggttatcaggagtagtgcggtat
 N F T F D I I N D Q A L R L S G V V R Y
 ttgccctcagataatgctgcaccgtcgacgctcgaggtctcactcgccgacaaagccgga
 L P S D N A A P S T L E V S L A D K A G
 aagactatggcatccattacagcaagggttattgaagagaaagggaatagtttctggggg
 K T M A S I T A R V I E E K G N S F W G
 gccacagcttactacccggtcgtttttgacataaaccttgctgggattgcttctcgcaac
 A T A Y Y P V V F D I N L A G I A S R N
 aatcttcctgggaagctccaggttccgacagcatctcctcagacatttgagctgcaaccg
 N L P G K L Q V R T A S P Q T F E L Q P
 gagctgtttttcataccatcccgttctagccctggcacggggataagtgtgggggctgcc
 E L F F I P S R S S P G T G I S V G A A
 ataggtgctgctgcgcggtcccgcgaactcgactctgagtgtagagagtgtggaagccgtg
 I G A A A R S R N S T L S V E S V E A V
 gtcagtggtgccggtatcccagacgggaacactggcaccgaaggtcgagaagcatgagttg
 V S V P V S Q T G T L A P K V E K H E L
 aatctcgaaagagatcaagatattggtttgtactctatattcaaggggaacctaagcgat
 N L E R D Q D I G L Y S I F K G N L S D
 gatttggtttcaacctgcagacgaccgttgatatacaggcaactttttcagacggcact
 D L A F N L Q T T V D I T A T F S D G T
 gtctggcaagatagttacaacagatttgcagtgccgtag
 V W Q D S Y N R F A V P -

SEQ ID NO : 36

CHIỀU DÀI : 532

LOẠI : PRT

SINH VẬT : *M. phaseolina*

MRFLGGFYLLLAATRHPTARAALHYPNALISRMEHLLVDTDGSFRSGFKDAINPCTNYISGAQTLGRQTSAQWL
 RVAFHDFVTAHVDEGTGGIDASIGFETLRAEDSGSAFNDSFAFFAPFVDAQTSMDLVALSVVTSLGHCGLHVP
 YRAGRIDATGGGPFVPEPETSLEETLEEFANAGFNAEDAIGLTACGHS LGRVHHGGFPNVVPESAIAPNNTAGG
 VNLDSTRDKFDISIVKEYLGNYGQRGGPLVTS DNVTVRS DLRLYESDQNR TMQALGQSKEYFFSTCGNLFERMIN
 TVPREVTLS DVIHPMTVQPVNFTFDIINDQALRLSGVVRYLPSADNAPSTLEVSLADKAGKTMASITARVIEEKG
 NSFWGATAYYPVVF DINLAGIASRNNLPGLQV RTASPQTFELQPELFFIPSRSSPGTGISVGAAIGAAARSNS
 TLSVESVEAVSVPSQ TGT LAPKVEKHELNLERDQDIGLYSIFKGNLSDDLAFNLQTTVDITATFSDGTVWQDS
 YNRFAMP*

SEQ ID NO : 37

CHIỀU DÀI : 2401 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

TTTAAATATATTCTGCCGCTAAAACGTTGCCATGTGCTCCTGATTTTTCTCGACCGCCATTTCTCGCCTCCGCA
 TCGGCTCCCTGCTTCGGCGCGCTCGTCGCTAGCATCCAGTTTCGCCAGTCCTTCCGTTTCTCCCGCTCTTGAACC
 ATGGCAAGGCATCCCGGCTGACCTGGTGTGGGTGCGAGGCGTGGTCTACCACGCCATTGGCACCGCTGCTGCC
 CCGACATGGCCATCTTCCATGGATGAAC TGGAGGACATCATGGTCTCTCAACACCGGCTACCGCGCCCGCAGCTTC
 GCAGTGCCCGTCACACCTGCGGCTTCTCGTCGCAAGGCCCGGACGCGTCCAAGCTGCCGAATGGATACGAACC
 GCCTTCCACGACATGGCTCCCGGACGCTGTATACCGGCGTCGGAGGACTGGACGCATCGATAGCTTACGAAACA

CGGAGCTTGGAAAATCTCGGCCCCGCCTTCAACACCACACTAGCTACCTACGCGCCGTATCTGACAAGCAGATCT
TCCATGGCTGACATCATTGCGCTAGGAGTCTACACGGCGGTGCGGTTCATGCGGAGGTCCAATCGTGCCTATCCGG
ACAGGCAGAGTAGACGCAAAGGCAGCAGGCCCCGAAGGTGTGCCTCTGCCGAGAATTCTATCGGAACTTTTCAA
AACCAGTTTCTCCGTACTGGCTTCAACACGACGGAGATGATCCAAGTGGTGGCGTGTGGCCACACTCTGGGCGGC
GTGCACGCATCTGCCAACCCGGAGATCGTGCCCGTGGGGTGGCGGAGGACGGCGTCTCAAGTTCGACACGACG
GACGCGTTTGACAACAAGGTTGTCAACGAGTACCTCTCGAATACAACCAAGAACTCGCTTGTGTGGGCCCTCT
ACTGCGAACGGCCGGAACCTCGGACGCTCGTGTCTTTGCGGCGGATGGAAATGCTACGGTCAGGGCTCTGGCTGAC
CCTGATACGTTCAACAGTGTGTGCGCTAGGATGCTGCAGAAGATGATTGATGTTGTTCGAACGGGCGTGGTGTCT
AAGCTTACGGGAGATATCCGTGTTAGGACTACGGAGCGCTCTGCGAGCCAGATTGAGAAGGTTCGAGCTTGTCTAC
AAGGACCGCGAGGGGGGCCGAATCCTCGACAGCTTTGAGCACCGAATCCTCAGGTTTCGGCCTCTGGGTTCGATGAC
AGCTTCGAGGTATAACTCATCCTAAGCCCCCTGTTGGGATGCCACACTAACGGCCGAGAAAGTTCTACGGGGTCTC
CGCAAACATCCCCACCGACTCCGGCATCTCCTCATTCACGTCCTCATCACCTCACCAGCGGCGAAACCGAACT
TCACGACAATAATGGCAGCGGTTTCCCCCTCCAAGACACCGTCATCTTTCAAAGCCCCGAGAGCTGCCTCAGCGG
CACCACCATGACCGTCGCCGCCGCCGTCTCAATACCGCCTCTCCGCCCCCACCCTCAGCGTCACCCTCAAGGT
CCCCAACTCCCGCTCCGTCTCTCCCGTCTCCAGGTCAGCACAGTCGCCATGACGAAGAGCTCGTCTGTGCGGCC
GTACGACCTCTACTCAACCACATACACGCTAACCTCCGCCCCAACTCGCCGACACGCGCTTTGACGTCAGGCTCGG
CGCTGCCGCCGCCGACGCTTCAAGTCTCCGCGGACCTCGGTGATGCCGTCAGGACCTGAGCCCCGAACCGCC
CACTTCTTCCAGCGCGCTAGCTCCTCGAGCACGGGTGCGCGCGCGGCCCTCGTCTCTCTGCCCCACCAAGCTC
CTCTGACCTCCCGGCCCCCTCCCCCAGCACAAACACCCACGTCTCTCCACCCAATCTCCTCATCTACAACATTAGC
ATCATCGACCACCCCAACGCCCCACAACACTCGCCTGCCCCGCCGTGACGGCGCCACCTGGACGCTATCCGGCGG
CCAGAAGTTCGCCGTCAAGTGCGGCAAGGACTACCAGGCCGGCCAGATCGGCGTCACGTGGACGGCCAGCTTCGA
GGCCTGTCTGCAGGCGTGGTGGATACGGACACGTGCCAGGCGGTGCGATTCTGTGGGGAGTGCAGAGGCGGGCGG
GCAGTGCTACTTGAAGGACCAGAGCGCGGGCAGTGTGGATGTTGAGGGGGTGTGGGGGGCGGTTCTGAAAAGCTG
AGCGGTTCGGGCGATGGAGTAGGGGATAGAGGGTTTGGGCTGGGGGGTAATGAATTACTTCTGATGAGCTTCTACT
ACTGCAGGTGGCACAAATCATGGCTCTCATGGACCTCCGACGCCGATGTGCCCGCGTTTCAAGGTATTCTTG
G

SEQ ID NO : 38
CHIỀU DÀI : 2049
LOẠI : ADN
SINH VẬT : *M. phaseolina*
TÊN/KHÓA DẤU HIỆU : CDS
VỊ TRÍ : (1).....(2049)

atggcaaggcatcccgccctgaccctgggtgtgggtcgagggcggtgtctaccacgccatt
M A R H P G L T L V W V A G V V Y H A I
ggcaccgctgctgccccgacatggccatcttccatggatgaactggaggacatcatggtc
G T A A A P T W P S S M D E L E D I M V
ctcaacaccggctaccgcgcccgacagcttcgcagtgcccgctcacaccctgcggtctctcg
L N T G Y R A R S F A V P V T P C G F S
tcgcaaggccccggacgctccaagctgccgaatggatacgaaccgccttccacgacatg
S Q G P G R V Q A A E W I R T A F H D M
gctcccggcagcggtataccggcgctcgaggactggacgcatcgatagcttacgaaaca
A P G S V Y T G V G G L D A S I A Y E T
cggagcttggaaaatctcgccccgccttcaacaccacactagctacctacgcgcgctat
R S L E N L G P A F N T T L A T Y A P Y
ctgacaagcagatcttccatggctgacatcattgcgctaggagtctacacggcggtgagg
L T S R S S M A D I I A L G V Y T A V R
tcatgaggaggtccaatcgctgcctatccggacaggcagagtagacgcaaaggcagcaggc
S C G G P I V P I R T G R V D A K A A G
ccgcaagggtgtgcctctgccgcagaattctatcggaacttttcaaaaccagtttctccgt
P Q G V P L P Q N S I G T F Q N Q F L R
actggcttcaacacgacggagatgatccaagtggtggcggtgtggccacactctgggcggc
T G F N T T E M I Q V V A C G H T L G G
gtgcacgcatctgccaaccggagatcggtgcccggtggggctcggcggaggacggcgctcgtc
V H A S A N P E I V P V G S A E D G V V
aagttcgacacgacggacgcgtttgacaacaaggttgtcaccgagtagctctcgaataca
K F D T T D A F D N K V V T E Y L S N T
accaagaactcgcttgttggggccctctactgcgaacggccggaactcggacgctcg
T K N S L V V G P S T A N G R N S D A R
gtctttgcggcgggatggaaatgctacggtcagggtctgtggtgacctgatacgttcaac
V F A A D G A T V R A L A D P D T F N
agtgtttgcgctaggatgctgcagaagatgattgatgttgttccaacgggcggtggtgctc
S V C A R M L Q K M I D V V P T G V V L

actgacccgatttcaatctatgatgtcaaaccagtgaggctgcagctgacattgcttggg
T D P I S I Y D V K P S G L Q L T L L G
ggaggagagtcggtaaagcttacgggagatatccgtgttaggactacggagcgctctgcg
G G E S V K L T G D I R V R T T E R S A
agccagattgagaaggtcgagcttgtctacaaggaccgcgagggggccgaatcctcgaca
S Q I E K V E L V Y K D R E G A E S S T
gctttgagcaccgaatcctcaggttcggcctctgggttcgatgacagcttcgagttctac
A L S T E S S G S A S G F D D S F E F Y
ggggtctccgcaaaccatccccaccgactccggcatctcctcattcaacgtcctcatcacc
G V S A N I P T D S G I S S F N V L I T
ctcaccagcgggcgaaccgaacttcacgacaataatggcagcggtttccccctccaagac
L T S G E T E L H D N N G S G F P L Q D
accgtcatctttcaaagcccgagagctgcctcagcggcaccaccatgaccgtcgccgcc
T V I F Q S P Q S C L S G T T M T V A A
gccgtcctcaataaccgcctcctccgccccaccctcagcgctcaccctcaagggtccccaac
A V L N T A S S A P T L S V T L K V P N
tcccgctccgtcctccccgtcctccaggtcagcacagtcgcatgacgaagagctcgtct
S R S V L P V L Q V S T V A M T K S S S
gtcggcccgtagcactctactcaaccacatacacgtaacctccgcccactcgccgac
V G P Y D L Y S T T Y T L T S A Q L A D
acgcgctttgacgtcaagctcggcgctgccgcccgcgacgccttcaagtcctccgcccgc
T R F D V K L G A A A A D A F K S S A D
ctcggtgatgcctgccaggacctgagccccgaaccgcccacttcttccagcgcgccctagc
L G D A C Q D L S P E P P T S S S A P S
tcctcgagcacggctgcgcgcccggcctcgtcctcctctgccgcaccaagctcctctgac
S S S T A A P P A S S S A A P S S S D
ctccccgccccctccccagcacaacaccacgctcctcccacccaatctcctcatctaca
L P A P S T P T P T S S H P I S S S T
acattagcatcatcgaccacccaacgcccacaacactcgctgccccgcccgtgacggc
T L A S S T T P T P T T L A C P A A D G
gccacctggacgctatccggcgccgagagttcgccgtcaagtgcgggcaaggactaccag
A T W T L S G G Q K F A V K C G K D Y Q
gccggccagatcggcggtcacgtggacggccagcttcgagggcctgtctgcaggcggtgcgtg
A G Q I G V T W T A S F E A C L Q A C V
gatacggacacgtgccaggcggtcgcattcgtggggagtgagagggcgggcgggcagtg
D T D T C Q A V A F V G S A E A G G Q C
tacttgaaggaccagagcgcgggcagtggtgagtggtgaggggggtggtggggggcggttctg
Y L K D Q S A G S V D V E G V W G A V L
gaaagctga
E S -

SEQ ID NO : 39
CHIỀU DÀI : 682
LOẠI : PRT
SINH VẬT : M. phaseolina

MARHPGLTLVWVAGVVYHAIGTAAAPTWPSSMDELEDIMVLNTGYRARSFAVPVTPCGFSSQGPGRVQAAEWIRT
AFHDMAPGSVYTGVGGLDASIAYETRSLNLGPAFNTTLATYAPYLTSRSSMADI IALGVYTAVRSCGGPIVPIR
TGRVDAKAAGPQGVPLPQNSIGTFQNQFLRTGFNTTEMIQVVACGHTLGGVHASANPEIVPVGSAEDGVVKFDTT
DAFDNKVVTEYLSNTTKNSLVVGPSTANGRNSDARVFAADGNATVRALADPDTFNSVCARMLQK MIDVVP TGVL
TDPISIIYDVKPSGLQLTLLGGGESVKLTGDIRVRTTERSASQIEKVELVYKDREGAESSTALSTESSGSASGFDD
SFEFYGVSANIPTDSGISSFNVLI TLTSGETELHDNNGSGFPLQD T V I F Q S P Q S C L S G T T M T V A A V L N T A S S A P
T L S V T L K V P N S R S V L P V L Q V S T V A M T K S S S A A P S S S D
Q D L S P E P P T S S S A P S S S T A A P P A S S S A A P S S S D L P A P S P T P T S S H P I S S S T L A S S T P T P T L A C P A A D G
A T W T L S G G Q K F A V K C G K D Y Q
A G Q I G V T W T A S F E A C L Q A C V
gatacggacacgtgccaggcggtcgcattcgtggggagtgagagggcgggcgggcagtg
D T D T C Q A V A F V G S A E A G G Q C
tacttgaaggaccagagcgcgggcagtggtgagtggtgaggggggtggtggggggcggttctg
Y L K D Q S A G S V D V E G V W G A V L
gaaagctga
E S -

SEQ ID NO : 40
CHIỀU DÀI : 2004 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)
LOẠI : ADN
SINH VẬT : M. phaseolina

CCTTCCTGAAATTTTCATGGCCAGCTTCTCGGTCTTTCTAATGAATTATAAGGCTATCTTCCTGCAGCTTGTCTGTC
TGCACGTTCTAGGATTCAACCACCCGTCATTCTCTCAAAATCTCAGTCCAGTCGTTACAGCTAATCACTTAAT
ATGAAGCCCCGCCGCTCTTGACGGCATTGGCCTGCTTAATGTCTACCCATTACCGCCGAATACGTCCTGGCCTTCT
AAATACGACTATTTGGAAGACCTGCTTTACCTTCAATCCGTTATCTGCGTGAGGGATTTGTCGACGGTATGCAT

ATATGAACAACCTGCTGGTGTCCCCCACTAATGGCTTTTAGGTGTGGCTCCTTGTTCGTTTTCTCCGCTGGACC
 TGGCCGTCAAACCGCAGCGGAATGGGTCCGCACTGCATACCATGATATGGCCACTCATGATGCCGATGCTGGCAC
 TGGTGGCTTGGACGCCTCTATCATGTTTGAGACCGAGCGGGACGAAAACGTGGGCGATGCGTTCAATGGTACCTT
 CGGCTTTACAAACAACCTACTACAACATCAAAGCATCCGCTGCTGATCTTCTTGCACTCTCCACTGTCATCGCCGT
 TGGAAACTGCGGTGGCCCGAAGATTCTTTCCGTGTCGGTTCGGTGGACGCCACGGAGGCTGGCCCTCTGGGTGT
 TCCCAAGCCGGATCAAGACGTTGATACGCACATTCAAATTTTTGCCAAGGCAGGCTTTAACACCAGTAGGCCTTG
 AATTACCTGAGGTCCACCTACCAGCCGCTGACTAGAATAGGCGACATGATCACCATGGTGGCCTGCGGCCACACC
 CTTGGTGGCGTCCATGGCAAGGACTTCCCCGAGATCACTTTCAACGACACGGAGACCAATTTTCGTCAAGTTCGAA
 GGCAACAACCTCCTTCTCCAACCTTTGATAACACCGTCGTGACCGAGTATCTTGGCGGAAACCCCCCAACCCCTC
 GTCACCGGCAAGAACGAGACCAACAACAGCGATAAGCGCGTCTTTGGTGCCGACAAACGCCACAATGCACTCC
 CTGTCCGACCCCTCCGTCTTCCAGTCCCTGCCAAGACATCCTCGCCCGCATGATCGACACCGTCCCCCTCCAAC
 GTCGCTCTCACCGAGCCTCTCGACCCAATCCTCATCAAGCCCTACATCCAAACCTTCTCCCTCGTCAACGCCACC
 CACCTCACCTTGACCGGCCGCATCCGCGTCCGTACCGACTGGGACAGCTACACCGACCAGTCAGTCCACCTCACC
 TACAACCCGCGCACAGCGCCCGCCAGAATGCGACCTCAACACCACCATCCCCACCACCCGCGCCACCTTCCAG
 GCGGCGACCTCCAGCGGCATCTTTGGTGAAGTCTTCGCTGGCAGAGTTCTCCGCCACCTCCCCACATCCAGC
 TCGATTACCGGCTTCACTGTCACCGTCACGCGCGGCTCCACGGGCGAATCCACCACCTACGACAACGCCGCGCAGC
 ACGAACGGCTACGCGCTCGACGACAGCTGCTCTACCAGAGCGCGCAGTCGTGCCGCGACGTCGGCACCACAACC
 ATCACCGCTGCTGTGCGCAAGGACTTCTCGCCCGAGGTGCAAAGGTCGCCGTCGAGATGGTGAACAGGGTGCCG
 CGCCAGGGCGTGACGTGCCGGCGCTAGAGGTCGAACCGTGGGGGCGGAGACGGTCAAGGAGGTCGGCGAGTGG
 GTCATCGTGCAAGGCTAAGGGTGAGTTGACGATGGAGAGCTTGAGCACAAACGTTTGACGTCGTTGCGGGCGAGAAA
 AGGGTCGAGTTCCAAAGGACGAATGTGCTGGGGGAGGAGTGTGCGGCTCTGTGAGGCGTAACTTAAAAAAAAAAAA
 AAAAGGAAGGATATATCAGTTTCTGGTACATACTTCGAATGAAAGTCTATGATTCTGTATGTCCATAACATTACTC
 TAGGCTTGAAACAACAATGCTGAATGTGCTTTATTTGTGAGATAAAGAGTGTCCCT

SEQ ID NO : 41
 CHIỀU DÀI : 1605
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 TÊN/KHÓA DẤU HIỆU : CDS
 VỊ TRÍ : (1).....(1605)

atgaagcccgccgctcttgcaggcattggcctgcttaatgtcctacccattaccgcccga
 M K P A A L A G I G L L N V L P I T A E
 tacgtctggccttctaaatacgactatgttggaagacctgctttaccttcaatccggttat
 Y V W P S K Y D Y L E D L L Y L Q S G Y
 ctgctgagggatttgcgacggtgtggctccttggttcttctccgctggacctggc
 L R E G F V D G V A P C S F S S A G P G
 cgtcaaaccgcagcggaatgggtccgcactgcataccatgatatggccactcatgatgcc
 R Q T A A E W V R T A Y H D M A T H D A
 gatgctggcactgggtggcttgagcgcctctatcatgtttgagaccgagcgggacgaaac
 D A G T G G L D A S I M F E T E R D E N
 gtgggagatgcgttcaatggtaccttcggctttacaaacaactactacaacatcaaagca
 V G D A F N G T F G F T N N Y Y N I K A
 tccgctgctgatcttcttgactctccactgtcatcgccgttgaaaactgcggtggcccg
 S A A D L L A L S T V I A V G N C G G P
 aagattccttccggtgctcggtgcggtggacgccagggctggccctctgggtgttccc
 K I P F R V G R V D A T E A G P L G V P
 aagccgatcaagacgttgatacgcacattcaaatttttgccaaggcaggctttaacacc
 K P D Q D V D T H I Q I F A K A G F N T
 agcgacatgatcaccatggtggcctgcggccacaccccttggtggcgtccatggcaaggac
 S D M I T M V A C G H T L G G V H G K D
 ttccccgagatcactttcaacgacacggagaccaatttcgtcaagttcgaaggcaacaac
 F P E I T F N D T E T N F V K F E G N N
 tccttctccaactttgataacaccgtcgtgaccgagtatcttggcgaaaccccccaac
 S F S N F D N T V V T E Y L G G N P P N
 cccctcgtcaccggcaagaacgagaccaacaacagcgataagcgcgtcttttggtgccgac
 P L V T G K N E T N N S D K R V F G A D
 aacaacgcccaatgcactccctgtcggacccctccgtcttccagtcctcctgccaagac
 N N A T M H S L S D P S V F Q S S C Q D
 atcctcgcggcatgatcgacaccgtccctccaacgtcgtctctaccgagcctctcgac
 I L A R M I D T V P S N V A L T E P L D
 ccaatcctcatcaagccctacatccaaaccttctccctcgtcaacgccaccacctcacc
 P I L I K P Y I Q T F S L V N A T H L T
 ctgaccggccgcatccggtccgtaccgactgggacagctacaccgaccagtcagtcac
 L T G R I R V R T D W D S Y T D Q S V H

ctcacctacaacccgcgcacagcgccccgccagaatgcgaccctcaacaccaccatcccc
 L T Y N P R T A P A Q N A T L N T T I P
 accacccgcgcacacctccagggcggcacctccagcgcatctttggtgaagtcttcgcc
 T T R A T F Q G G T S S G I F G E V F A
 tggcacgagttctccgccaccctccccacatccagctcgattaccggcttcactgtcacc
 W H E F S A T L P T S S I T G F T V T
 gtcacgcgcggctccacggggaatccaccacctacgacaacgccggcagcacgaacggc
 V T R G S T G E S T T Y D N A G S T N G
 tacgcgctcgacgacacgctgctctaccagagcgcgacgctcgccgcgacgtcggcacc
 Y A L D D T L L Y Q S A Q S C R D V G T
 acaaccatcaccgctgctgtgcgcaaggacttcctcgcccgaggtgcaaaggctcgccgtc
 T T I T A A V R K D F L A R G A K V A V
 gagatggtgaacagggtgcccgcgccagggcggtgtacgtgcccggcgctagaggtcgaaccg
 E M V N R V P R Q G V Y V P A L E V E P
 tggggggcgagacgggtcaaggaggtcgccgagtggtcatcgtaggctaagggtgag
 W G A E T V K E V G E W V I V Q A K G E
 ttgacgatggagagcttgagcacaacgtttgacgtcggttgcggggcgagaaaagggtcgag
 L T M E S L S T T F D V V A G E K R V E
 ttccaaaggacgaatgtgctgggggagggagtgtgcggtctctgtga
 F Q R T N V L G E E C A A L -

SEQ ID NO : 42

CHIỀU DÀI : 534

LOẠI : PRT

SINH VẬT : *M. phaseolina*

MKPAALAGIGLLNVLPIAEYVWPSKYDYLEDLLYLQSGYLREGFVDGVAPCSFSSAGPGRQTAAEWVRTAYHDM
 ATHDADAGTGGLDASIMFETERDENVGDAFNGTFGFTNNYNIKASAADLLALSTVIAVGNCGGPKIPFRVGRVD
 ATEAGPLGVKPDQDQDVTIHIQIFAKAGFNTSDMITMVACGHTLGGVHGKDFPEITFNDTETNFVKFEGNNSFSNF
 DNTVVTEYLGNNPPNPLVTGKNETNNSDKRVFGADNNATMHSLSDPVSFQSSQDILARMIDTVPSNVALTEPLD
 PILIKPYIQTFSLVNATHLTLTGRI RVRTDWDSTYDQSVHLTYNPRTPAQNATLNTTIPTRATFQGGTSSGIF
 GEVFAWHEFSATLPTSSSITGFTVTVTRGSTGESTTYADNGSTNGYALDDTLTYQSAQSCRDVGTTTITAAVRKD
 FLARGAKVAVEMVNRVPRQGVYVPALEVEPWGAETVKEVGEWVIVQAKGELTMESLSTTFDVVAGEKRVEFQRTN
 VLGECAAL*

SEQ ID NO : 43

CHIỀU DÀI : 1414 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

AGCTTGCCTCTTCTTCTGTCTTTCTGCCCCCTTCTTCTCCAGACTCCTTTTCGCCGTAGATTTCTGCCTTTTGGAC
 CTCGTTTTCTGCCCTCCACCAATAGCAAGCCAAGAGATTAGATTAGACATCGCCATCCCATCGCCGACTCCAACA
 ATGCCAGCAACCCAGGCGACTACGACGCGCTCCGCCACGACGTCAAGAACCCTCCTGCACCAGCCCGAGTACGAC
 GACGGTTCCGCCGGGCCCCGTCTCGTCCGTCTCGCATGGTATGTTACGGCTACCCGCATCCTCCAGCCCAATCCG
 CCGTTTATCTGCTGTTCTGGCTCTGCCATCTTCTCTCCAGGGGCAGCAAGTTTCCCGAGCCTTTTTTTTTTTGTC
 TTACGAACCGCATCTCGAACAGCCCGCTGACAACACACACAGGCATTCGCGAGGGACCTACGACGCCACTCTGA
 CACAGGAGGCAGCAACGGTGCAGGCATGCGCTACGAGGCTGAAGGCGGCGACCCCGCCAATGCCGGCTGCAGCA
 CGCCCGCGTCTTCTCGAGCCCATCAAGGCCGCGACCCCTGGATAACCTACTCTGATCTGTGGACGCTGGCGGG
 CGTGGTCCCATCAAGGAGATGGGCGGCCCCGACATCCCGTGGCAGCCCGGCGCACCGACTTCGTCGACGACAG
 CAAGCTGCCCGCGGGGCCCTGCCGGACGCCGCGCAGGGCGCTGACCACATCCGCTGGATTTTCTACCGCAT
 GGGCTTCAACGATCAGGAAATTGTCGCCCCAGCGGCGCCACAACTCGGCCGCTGCCACGCCGACCGCTCCGG
 CTTTCGACGGCGCCTGGGTCAACAACCCACCCGCTTCTCCAACAGTACTTTAAGCTCCTGACCTCGGTGAGTG
 GAAAGAGAAGACCTCCCCAGCGGCATCAAGCAGTTCGCCCTACTATGATGAGGACTCGGAGGAGGAGCTCATGAT
 GCTGCCCCACCGATATCGCTCTCTTGACACGACCCCTCCTTCCGGCCGTGGGTGAGAGTATGCCGAGGACAAGGA
 TGCCTTTTTTCGACAGACTTCTCAAAGGTCTTTGCCAAGCTGATTGAGCTGGGCATAGTCAGAGATGAGAGCGGTGC
 GGTAATCAACACTGATAACGTCAAGGGCGGCTACATCTCTGCGCCCAAGAAGAGTGAGCTGCCTGGTGCTCCGGG
 TAAGGCTAATGAGGAGGCTGAGCCGCTCATGAAAGAGAATGAGAGGTTGAGGGCACGCTCTGTAAGCTGGTTGGTT
 TAGATTTCTTTTTTTTTTTTTTTTCAAGACCGTTAGACTGCGCAAGCTAGAGGGGACGGATAGATGGACGATACC
 ACCTAGATTCTTATGCCCTTAAGAGGAGAAAGATAGTGCTAGAGTTTTTCATAGAGAATGCAAAAC

SEQ ID NO : 44

CHIỀU DÀI : 960

LOẠI : ADN

SINH VẬT : *M. phaseolina*

TÊN/KHÓA DẤU HIỆU : CDS

VỊ TRÍ : (1).....(960)

atgccagcaacccaggcgactacgacgccgtccgccacgacgtcaagaacctcctgcac

M P S N P G D Y D A V R H D V K N L L H
 cagcccgagtagcagcagcggttccgcccgggcccgtcctcggtccgtctcgcatggcattcc
 Q P E Y D D G S A G P V L V R L A W H S
 gcagggacctacgacgccccactctgacacaggaggcagcaacgggtgcaggcatgcgctac
 A G T Y D A H S D T G G S N G A G M R Y
 gaggtgaaggcgggcgacccccgccaatgccggcctgcagcagccccggtcttctcctcgag
 E A E G G D P A N A G L Q H A R V F L E
 cccatcaaggccgcgacccccctggataacctactctgatctgtggacgctggcgggcggtg
 P I K A A A H P W I T Y S D L W T L A G V
 gtcgccatcaaggagatgggcccggccgacatcccgtggcagccccggcgacccgacttc
 V A I K E M G G P D I P W Q P G R T D F
 gtcgacgacagcaagctgccgcccggggccgcctgccggagcggcgcgagggcgctgac
 V D D S K L P P R G R L P D A A Q G A D
 cacatccgctggattttctaccgcatgggcttcaacgatcaggaaaattgtcgccctcagc
 H I R W I F Y R M G F N D Q E I V A L S
 ggcgcccaaacctcgccgctgccacgccgaccgctccggcttcgacggcgccctgggtc
 G A H N L G R C H A D R S G F D G A W V
 aacaacccccacccgcttctccaaccagtactttaagctcctgacctcggtcgagtggaaa
 N N P T R F S N Q Y F K L L T S V E W K
 gagaagacctccccagcgcatcaagcagttcgccctactatgatgaggactcggaggag
 E K T L P S G I K Q F A Y Y D E D S E E
 gagctcatgatgctgccaccgatatcgctctcttgcacgacccccctccttcgggcccgtgg
 E L M M L P T D I A L L H D P S F R P W
 gtcgagaagtatgccgaggacaaggatgcctttttcgcagacttctcaaaggtcctttgcc
 V E K Y A E D K D A F F A D F S K V F A
 aagctgattgagctgggcatagtcagagatgagagcggtgcggtaatcaacactgataac
 K L I E L G I V R D E S G A V I N T D N
 gtcaaggggcggtacatctctgcgcccagaagagtgcgctgcctgggtgctccgggtaag
 V K G G Y I S A P K K S E L P G A P G K
 gctaattgaggaggctgagccgctcatgaaagagaatgagagggttcagggcacgtctgtaa
 A N E E A E P L M K E N E R F R A R L -

SEQ ID NO : 45

CHIỀU DÀI : 319

LOẠI : PRT

SINH VẬT : *M. phaseolina*

MPSNPGDYDAVRHDKVNLHQP EYDDGSAGPVLVRLAWHSAGTYDAHSDTGGSSNGAGMRYEAEGGDPANAGLQHA
 RVFLEPIKAAHPWITYSDLWTLAGVVAIKEMGGPDIPWQPGRTDFVDDSKLPPRGRLPDAAQAGADHRIWFYRMG
 FNDQEIVALSGAHLGRCHADRSGFDGAWVNNPTRFSNQYFKLLTSVEWKEKTLPSGIKQFAYYDEDESEELMML
 PTDIALLLHDPFSFRPWVEKYAEDKDAFFADF SKVF AKLIELGIVRDESGAVINTDNVKGYYISAPKKS ELP GAPGK
 ANEEAEPLMKENERFRARL*

SEQ ID NO : 46

CHIỀU DÀI : 2160 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

CAACTCTTTGGATCATTCGCGAGTTCTGCAGATTAGACTACTCTTTTCTCGCTCCACAATGAAGCTTGCATCATGG
 CTGCTCCTCCCTCAAGTACTGGCAGCGCTGGGCGGGCAGGTACGTGAATATAACTTGACTCTTGAGGCAAGCTGG
 ATGGCTCAAGGTAGATGTCCCCGAAGAAACAAAGTGATAGTATCTCTGCTAACATGGACAGACGGGAACCCCTCGC
 GGTGTCTTAACATATCAACGGCCAGACACCTGGTCCATTGATCTGGGGATATGAAGGAGACACACTTCGCGTCACT
 GTGACCAATAAAATGTTTATTGAGGCTACTATGCATTGGTCAGAGTGAAGTCTGTCTGGACTGGGAGAATCAACT
 AACGGAAGCAGGCACGGTGTCTATCAGGTGCACAAGTACTGGAACGACGGAGTACCTGGCGTGACTCAATGGCC
 CATTGAATCCAGGGATTCTGTATACCTACGAGTTTACTCTCACCAACCAAACCTGGAAGCTACTTCTACCATGGCCA
 CTTTGGACCCGCATTTCGCGGACGGCCAACGAGGCCCGCTGTGGATTGCACCGGCCCCCTGGAGACCCCGTCCGTA
 TGAGCTTGCGTCTGATGACCCGGCAGAGGTTGCAGCAATGCGCGCGGCCGAAGACAATCCGAGACACCTCATGGT
 TTCCGACTGGAACATATGAGGGAATGGAAGTGCTGATTGTGGGCTTCAGAGATGCAGGCATTGCTCCGGCATGTTC
 TGCGTCCCTCGTGACAAATGGAAAGGGCAGGACAACTTGCCCTCGGCCCAGATGATATCAAGAAATACGATCCCGA
 GGGTCGGAGGAATTCACTTGGGTGCCCTTCCTCCTCCAGTCGGCGCTGAGTTCACCAACAAGAGAGAATGCCGCGA
 GACTACCACCGACTTCGAGATCATTCAGGCCGAAGAAGGGGAGAAGTATATCTACATGAACTTTATCCACCCTGG
 AGCCCACCATGAACTGCCAATCGCGGTGGACGAGCAGACATGATCATCGTGGCAGCTGACGGGGATTTTGTGAT
 GCCGAAAAAAGTCCAGGTACGTTAATGCGAGAATCTGCATGAGCGGTGCTGACCTATGGCTGCAGGCAATAAAC
 CTCACATGGGCGACAGGATCAGTGTCTTGGTACCGCTAGACAAGAAGCCGGGGGAATACGCCATCCGCCCTGTGCG
 TCCATTTCCGAGGCAATTGATTACGGGCTTGAGCATCTTGCGCTACCCCGGTGTGCAGGAGCGCCGCAAGAC
 GGTATTATGCTGGCACCGGAAACAAAACCCCATATTGATCTGTTGGGGCGGATGGTCACTGAAGGAGGTGTCATG

ATGGATGAAATGACCGATTGGCCCCCTTCCGCCGCGCTACCCCCAGCGACGTCTGATCACACGTTTCGGTTT
 TTATCAAACCGCACCGGCCGAGCACATGGATGCTGTGAGCGAGCCACACCAAGGCTTCCGCCAGCAGATGCCT
 CCTATTATGTGGAACGAAGAGTCTCGTGGCCCTACAACCATTCAGGGGATGAAGAATGGATCCACCGTAGACATC
 ATCTTCGAGAGCCGCGCATACGCCATGCACCCCTTCCACAAACACAATCACAAAGGCTGGATTATTGGTAGAGGA
 AAGGGCTACTTCCGTTGGCCAGATGTTGCTACTGCCATCTCGGAAGCTCCAGAGAATTTCAACCTGATCAACCCG
 CCGTTGCGAGATGGTGCCCGGCTCGAAGCGGAAGAGGGATCCTGGACGGTGATCCGTTACACCATCACCTTTTCCT
 GCCATGAGCATGCTGCACTGCCATCGTATTGAGCACTTTGCGGTAAGTGTGTGATCTAGCCACCAAACGAATGGT
 CACTGATTGAAACAGGCTGGACAACAGATAGTCCCTTTTGGAGGGGCGAGGATGTGATGCAAAGCCCTCCTGAGTAC
 ATCAAAAAAATGACGCATGCGAGCTTTGTGCCGCCACTCCGATACGGCCCCCTTGACTGAGAGTTCCCATGCGAG
 GCCAGAGCACGGAATTGGGTAGTTGAAGTTAAAAACCTGCCACATAAAAAATAGCTCGAATAACAGGATTTTGCTA
 GCTCATGGAAGTAGCCACTTGTTCCTTTGCGTTTCACTTAATGGCTGCGGTGAAAGCGAA

SEQ ID NO : 47
 CHIỀU DÀI : 1713
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 TÊN/KHÓA DẤU HIỆU : CDS
 VỊ TRÍ : (1).....(1713)

atggctcaaggtagatgtccccgaagaaacaaagtgatagtatctctgctaacatggaca
 M A Q G R C P R R N K V I V S L L T W T
 gacgggaaccctcgcggtgtcttaactatcaacggccagacacctggccattgatctgg
 D G N P R G V L T I N G Q T P G P L I W
 ggatatgaaggagacacacttcgcgctcactgtgaccaataaaaatgtttattgaggctact
 G Y E G D T L R V T V T N K M F I E A T
 atgcattggcaggtgtctatcaggtcgacaagtaactggaacgacggagtagctggcgtg
 M H W H G V Y Q V D K Y W N D G V P G V
 actcaatggcccattgaatccagggttcgtatacttacgagtttactctcaccaaccaa
 T Q W P I E S R D S Y T Y E F T L T N Q
 actggaagctacttctaccatggccactttggaccgcattcgcgagcgccaacgaggc
 T G S Y F Y H G H F G P A F A D G Q R G
 ccgctgtggattgcaccggccccctggagaccccgctccgtatgagcttgcgctctgatgac
 P L W I A P A P W R P R P Y E L A S D D
 ccggcagaggttgagcaatgcgcgcggccgaagacaatccgagacacctcatggtttcc
 P A E V A A M R A A E D N P R H L M V S
 gactggaaactatgaggggaatggaagtgcgtgattgtgggcttcagagatgcaggcattgct
 D W N Y E G M E V L I V G F R D A G I A
 ccggcatgttctgcgtccctcgtgacaaatggaaagggcaggacaacttgccctcggccca
 P A C S A S L V T N G K G R T T C L G P
 gatgatatcaagaaatacgatccccgaggggtcggaggaattcacttggtgccttcctcct
 D D I K K Y D P E G R R N S L G C L P P
 ccagtcggcgcgtgagttaccaacaagagagaatgccgcgagactaccaccgacttcgag
 P V G A E F T N K R E C R E T T T D F E
 atcattcaggccgaagaaggggagaagtatactacatgaactttatccaccctggagcc
 I I Q A E I G E G E K Y I Y M N F I H P G A
 caccatgaactgcgaatcgcggtggacgagcacgacatgatcatcggtggcagctgacggg
 H H E L R I A V D E H D M I I V A A D G
 gattttgtcatgccgaaaaaagtccaggcaataaacctcaacatgggagacaggatcagt
 D F V M P K K V Q A I N L N M G D R I S
 gtcctggtaccgctagacaagaagccgggggaatacgccatccgcctgtcgtccatttcc
 V L V P L D K K P G E Y A I R L S S I S
 gaggagcaattgattacgggcttgagcatcttgcgctaccccggtgtgcaggagcgccgc
 E E Q L I T G L S I L R Y P G V Q E R R
 aaagacggtattatgctggcaccggaaacaaaaccccatattgatctgttggggcggtg
 K D G I M L A P E T K P H I D L L G R M
 gtcactgaaggaggtgtcatgatggatgaaatgaccgatttggccccctttccgccgcgc
 V T E G G V M M D E M T D L A P F P P R
 tcacccccagcgacgtctgatcacacgttttcggtttttatcaaaccgcaccggccccgagc
 S P P A T S D H T F R F L S N R T G P S
 acatggatgctgtcgagcgagccacaccaaggcttccgccagcagatgcctcctattatg
 T W M L S S E P H Q G F R Q Q M P P I M
 tggaacgaagagtctcgtggccctacaaccattcaggggatgaagaatggatccaccgta
 W N E E S R G P T T I Q G M K N G S T V
 gacatcatcttcgagagccgcgcatacgccatgcaccctttccacaaacacaatcacaag

D I I F E S R A Y A M H P F H K H N H K
gcctggattattggttagaggaaagggtacttccgttggccagatggtgctactgccatc
A W I I G R G K G Y F R W P D V A T A I
tcggaagctccagagaatttcaacctgatcaacccgccgttgcgagatggtgccccggctc
S E A P E N F N L I N P P L R D G A R L
gaagcggaagagggtacctggacggtgatccgttacaccatcacctttcctgccaatgagc
E A E E G S W T V I R Y T I T F P A M S
atgctgcactgccatcgatttcagcacttttgcggctggacaacagatagtccttttggag
M L H C H R I Q H F A A G Q Q I V L L E
gggcaggatgtgatgcaaagccctcctgagtacatcaaaaaaatgacgcatgagagcttt
G Q D V M Q S P P E Y I K K M T H A S F
gtgccgccactccgatacggcccccttgactga
V P P L R Y G P L D -

SEQ ID NO : 48

CHIỀU DÀI : 570

LOẠI : PRT

SINH VẬT : *M. phaseolina*

MAQGRCPRRNKVIVSLLTWTGDNPRGVLTINGQTPGPLIWGYEGDTRLVTVTNKMFIEATMHWHGVIYQVDKYWND
GVPGVTQWPIESRDSYTYEFTLTNQTSYFYHGHFGPAFADGQRGPLWIAPAPWRPRPYELASDDPAEVAAMRAA
EDNPRHLMVSDWNYEGMEVLIVGFRDAGIAPACSASLVTNGKGRITCLGPDDIKKYDPEGRNRSGLCLPPVGA
FTNKRECRETTFEIIQAEEGEKYIYMFHIGPAHHELRIAVDEHDMIIVAADGDFVMPKKVQAINLNMGDRIS
VLVPLDKKPGEYAIRLSSISSEEQLITGLSILRYPGVQERRKDGIMLAPETKPHIDLGRMVTEGGVMMDEM
DLA PFPSPSPATSDHTFRFLSNRTGPSTWMLSSEPHQGFQQMPPIMWNEESRGPTTIQGMKNGSTVDIIFES
RAYA MHPFHKNHKAWIIGRGKGYFRWPDVATAISEAPENFNLINPPLRDGARLEAEEGSWTVIRYTTTFPAM
SMLHCH RIQHFAAGQQIVLLEGQDVMQSPPEYIKKMTASFPPLRYGPLD*

SEQ ID NO : 49

CHIỀU DÀI : 2269 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

AGTATGTGCGTTGATCCCCTGTCTATTTCGAGCACCAGAGCACTGTCCAGCCACCTTGAGCTATCTCAAGGGGCT
ATCTCATCTGGTCATCTTCGCTGGGGCATCGGTTGCCTTCTCAACCCCTTCTCAATTCCTCCCTCACGCTTCTGCAAC
ATGTTCTTCGGCTCCCTCCACTTGGGCATCGGCGCCCTATTGGTTGCCGGCACTCTTGCTGGCGACGACAAGTGG
CTTAGCCCCGTCTACAAGAACTTTTACGAGTTCCTCCCTACCTAAGCCACCAATCAAGGAAGCGAAAGCGTAAGTC
AACTGATATTTCTCTCGCTCGTCATTTGTCATGCCCTAATCTTCGCGAGGAAATATACCAACCCGACTACCGGTGCTG
TGATCAACTACTACGAAATCACCATCTCACCCTGCGAGCAACAGGTTTATCCTGGCCTTGCGAAGGCAAACCTCG
TTGGCTACGATGGTATCTCTCCCGTCCCACTTTTAAGATGGAGAGGGGAGAAGAGGCTGTCTGTTCTGTTTCATCA
ACAAGGCCTCCATTTCCCAATTTCCGTCCATCTTCACGGCTCCTACTCCTTTGCCCCCTTCGATGGCTGGGCGGAGG
ATACGACCAGCCCAGGCCAATACAAGGACTACTACTACCCCAATGCCAGTCTGCCCGTACCTCTGGTACCACG
ACCATGCCGTCTTCCACACTGCCGAGAACGCCTACTACGGTCAGGCAGGTTTCTACATCCTGCACGACTCGGCTG
AGGATAGTCTGGGTCTCCCGTCTGGAGACTACGACATCCCGCTCGGTCTGAGCTCGAAGCAATACCAGTCCAACG
GTGACCTTTTTCAGCCCCGAATGGCGAGACGGATAGCCTTTTGGCGATGTTATCCATGTCAACGGCCAGCCCTGGC
CGTACCTCAAGGTCGAGCCCAGGAAGTACCGCTTCCGCTGCTTGATACAAGCATTTCCCGTGCTTCCAGCTGT
CACTCCAAGACGATAAGAGCAAGAAGATTGACTTTAACGTCATCGCCTCCGATGCCGGCTCCTGTCCAGCCCTG
TTCCGACCAACCTGCTACACATTTCCATGGCCGAACGCTGGGAAATTGTCGTCGACTTCTCCAGTACGCTGGCA
AGAACATCACCATGAAGAACGAGCGTGACGTGCAGGCCGATGAAGACTACAACAGCACTGACAAGGTCATGCGCT
TCGTAAGTAGGCAACAAGGTTACCTCGACTGCCAACAACAACCTGCCCTGGCAGCCTCCGAGCGTGCTTTCCCGC
CGAATAAGTCTGGTGTGACAGGAGCTTCAAGTTCGAGCGCAAAGGCGGTGAATGGACTATCAACGGCGTTACCT
TTGCCGACGTCGAGAACCCTATTCTAGGCAAGCCCCAACGCGGACAGGTTGAGGTCTGGGAGCTCGAGAACTCTT
CCGGCGGCTGGTCTCACCCCGTTACATCCATCTCATCGACTTCCAGGTTATTTCTCGCACTGGTGGCAAGCGTG
ATGTCCTGCCCTTACGAGAAGAACGGTCTCAAAGATGTCGTCTTGCTTGCGGTGAACGAGAAGGTCAGAGTTGTGCG
CTCGCTTCCAGCCCTGGGAAGGTGTCTACATGTTCCATTGCCACAACCTGATCCACGAGGACCACGATATGGTGC
GCTCCCAGCATTCCTCTGATTACTCGCAACAAAGACTGACAAATAATTACAGATGGCCGCTTTCAATGTATCCAC
CCTTTTCAGACTACGGCTACGACCAGAAGGAGACCTTTTTCATCGACCCGATGGAGTCCCGCTGGCGCGCAAAGGA
TGTCGAAGACCGAAGACTTCACAACCGATGCCATCCAGTCGAAGCTTGCCGCTTTTCGCTGAAATAAATCGCTACAA
GGACGTTGCGAAGATCGAAAGTGCTCTTGAAAATTTATGGGAAGACCGCTCCACCCGGGTTCAAGATTAGCACCAC
CTCCTCAACTCCGAGCTCAACCGCCGCCACCTCCACCGCTTCCAGCAGTGGCTCAAACACCTCTATTACCGCCCC
CATTCAGTCCCCGGCCACTACTTACCCCGGACCACTTCGACCAAGGCGGATGACAAGGCGAAGGCTAAGACTTC
TACGACCAAAAACGAAGTAACGAAAATGGCTTTTAGATCACGCTCATTATAGATAAAAACCTGGGAGAATTTGGGTGCG
GTGTTTTTGGGATGGTTTGGTTCGGAGGAGCGCTCTCTGTAAATATGGGCCCTCGCGAACTCTGTGGATATTTGTTT
CTCGGTGTCTGTTAATATC

SEQ ID NO : 50

CHIỀU DÀI : 1860
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 TÊN/KHÓA DẤU HIỆU : CDS
 VỊ TRÍ : (1).....(1860)

atgttcttcggtccctccacttgggcacgagccctattggttgccggcactcttgc
 M F F G S L H L G I G A L L V A G T L A
 ggcgacgacaagtggcttagccccgtctacaagaacttttacgagttccccctacctaag
 G D D K W L S P V Y K N F Y E F P L P K
 ccaccaatcaaggaagcgaaagcgaaatataccaacccgactaccgggtgctgtgatcaac
 P P I K E A K A K Y T N P T T G A V I N
 tactacgaaatcaccatctcaccctgcagcaacagggtttatcctggccttggcaaggca
 Y Y E I T I S P L Q Q Q V Y P G L G K A
 aacctcggttgctacgatggtatctctcccggtcccacttttaagatggagaggggagaa
 N L V G Y D G I S P G P T F K M E R G E
 gaggtgtcggttcgtttcatcaacaaggcctccattcccaattccgtccatcttcacggc
 E A V V R F I N K A S I P N S V H L H G
 tcctactcctttgcccccttcgatggctgggaggagatacaccagcccaggccaatac
 S Y S F A P F D G W A E D T T S P G Q Y
 aaggactactactacccaatgcccagtgctgcccgtaccctctggtaccacgaccatgcc
 K D Y Y Y P N A Q S A R T L W Y H D H A
 gtcttccacactgccgagaacgcctactacggtcaggcagggtttctacatcctgcacgac
 V F H T A E N A Y Y G Q A G F Y I L H D
 tcgggtgaggatagtgctgggtctcccggtctggagactacgacatcccgtcgggtctgagc
 S A E D S L G L P S G D Y D I P L G L S
 tcgaagcaataccagtcacacgggtgaccttttcagcccgaatggcgagacggatagcctt
 S K Q Y Q S N G D L F S P N G E T D S L
 tttggcgatggttatccatgtcaacggccagccctggccgtacctcaaggtcgagcccagg
 F G D V I H V N G Q P W P Y L K V E P R
 aagtaccgcttccgcctgcttgatacaagcatttcccggtgccttccagctgtcactccaa
 K Y R F R L L D T S I S R A F Q L S L Q
 gacgataagagcaagaagattgactttaacgtcatcgctccgatgccggcctcctgtcc
 D D K S K K I D F N V I A S D A G L L S
 agccctgttccgaccaacctgctacacatttccatggccgaacgctgggaaattgtcgtc
 S P V P T N L L H I S M A E R W E I V V
 gacttctcccagtagctggcaagaacatcaccatgaagaacgagcgtgacgtgcaggcc
 D F S Q Y A G K N I T M K N E R D V Q A
 gatgaagactacaacagcactgacaagggtcatgcgcttcgtagtaggcaacaagggttacc
 D E D Y N S T D K V M R F V V G N K V T
 tcgactgccacaacaacctgcctggcagcctccgcagcgtgcctttcccgccgaataag
 S T A N N N L P G S L R S V P F P P N K
 tctggtgttgacaggagcttcaagttcgagcgcaaaaggcggtgaatggactatcaacggc
 S G V D R S F K F E R K G G E W T I N G
 gttaccttggccgacgtcgagaaccgtattctagggcaagcccaacgcggacaggttgag
 V T F A D V E N R I L G K P Q R G Q V E
 gtctgggagctcgagaactcttccggcggtgtctcaccctgttcacatccatctcatc
 V W E L E N S S G G W S H P V H I H L I
 gacttccagggttatttctcgactggtggcaagcgtgatgtcctgccttacgagaagaac
 D F Q V I S R T G G K R D V L P Y E K N
 ggtctcaaagatgtcgctcttggcgtgaacgagaagggtcagagttgtcgctcgcttc
 G L K D V V L L G V N E K V R V V A R F
 cagccctgggaagggtgtctacatgttccattgccacaacctgatccacgaggaccacgat
 Q P W E G V Y M F H C H N L I H E D H D
 atgatggccgctttcaatgtatccaccctttcagactacggctacgaccagaaggagacc
 M M A A F N V S T L S D Y G Y D Q K E T
 cttttcatcgaccgatggagtcccgtggtggcgcaaggatgtcaagaccgaagacttc
 L F I D P M E S R W R A K D V K T E D F
 acaaccgatgccatccagtcgaagcttggcgctttcgctgaaataaatcgctacaaggac
 T T D A I Q S K L A A F A E I N R Y K D
 gttgcgaagatcgaaagtgtcttgaataattattggaagaccgctccaccgggttcaag
 V A C I E S A L E N Y W K T A P T G F K
 attagcaccacctcctcaactccgagctcaaccgcccacctccaccgcttccagcgt
 I S T T S S T P S S T A A T S T A S S S
 ggctcaaacacctctattaccgccccattcagtccccgccactacttcaccgcgacc

G S N T S I T A P I Q S P A T T S P A T
 acttcgaccaagggcgatgacaagggcaaggctaagacttctacgaccaaacgaagtaa
 T S T K A D D K G K A K T S T T K T K -

SEQ ID NO : 51
 CHIỀU DÀI : 619
 LOẠI : PRT
 SINH VẬT : *M. phaseolina*

MFFGSLHLGIGALLVAGTLAGDDKWLSPVYKNFYEFPLPKPPIKEAKAKYTNPTTGAVINYEITISPLQQQVYP
 GLGKANLVGYDGISPGPTFKMERGEEAVVRFINKASIPNSVHLHGSYSFAPFDGWAEDTSPGQYKDYYPNAQS
 ARTLWYHDHAFVHTAENAYYGQAGFYILHDSAEDSLGLPSGDYDIPGLGLSSKQYQSNGLFSPNGETDLSLFGDVI
 HVNGQPWPYLLKVEPRKYRFRLLDTSISRFAQLSLQDDKSKKIDFNVIASDAGLLSSVPVPTNLLHISMAERWEIVV
 DFSQYAGKNITMKNERDVQADEDYNSTDKVMRFVVGKVTSTANNLPGSLRSVPFPNPKSGVDRSFKFERKGG
 WTINGVTFADVENRILGKPQRGQVEVWELENSSGGWSHPVHIHLIDFQVISRTGGKRDVLPYEKNGLKDVVLLGV
 NEKVRVVARFQWEGVYMFHCHNLIHEDHDMMAAFNVSTLSYGYDQKETLFDPMESRWRAKDVKTEDFTTDAI
 QSKLAFAEINRYKDVAKIESALENYWKTAPTGFKISTTSSTPSSTAATSTASSSGSNTSITAPIQSPATTSPAT
 TSTKADDKGKAKTSTTKTK*

SEQ ID NO : 52
 CHIỀU DÀI : 2423 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*

AAAGGCCCTTTGGGCTCCCGCCAAAATTGTTTCATCCTCACGCTCACTCCTGCGCTCTTCCCTACAGACCTTTTGCCT
 GTAGCTTTAGCTTCACGGTCTTTTCATTCCCCTTCTCTGGCATTTCATCCACCGCGGGAAGCCTCAATTACCACA
 ATGGTGTCCCTTAAGCAAATCGGTGCCACTCTGTTGGCACTGACTGCTCAAACCTTCGCTGCTGCCATCCCTGAG
 GCCGAACCGGTGGACCTCGTCGCCCCGTCAAGCCACGACCCTACGTCCACCACCTCCTCAACCACCTCGAGGGTT
 CCCGATTCCCGATGTACTTGGGGCCCTTCGAGCAGGGGTTGCTGGAAGAACGGCTTCAGCATTGCCACTGACTTT
 GACACGAAGTGGCCATCTACTGGCAAGACTGTCTCATATCGCTTGGAAGTTACGAATGTCACCAACTGCCAGGAC
 TATCAAAGCAAGGGAATTGGAGATGGCTTCTGCAGGCCGATGCTGCTCATCAACAATCAGTTCCCAGGGCCTACC
 AGTCAGTTTTTCCCTCCTCGATGCCGGGTGACAAGGGCGCTAATGACTCGTACTTTCCAGTCAATGCCGAATGGG
 GAGACAACCTTGAGATTACTGTTGTCAATAGTATGCAAGACAACGGCACCTCATTTCCACTGGCACGGCATTCGTC
 AGCTGAACTCATGCCAAAACGATGGTGCCAACGGCGTCACCGAGTGCCCTATTTCCCTCCCGGCGGAAGCTTCACCT
 ATAAGTTCAAGGCTACACAGTATGGTGAGTCTTTGCCTTCTCTTCTACCTTTTCCATTTTCATGCAAGCATCTTCA
 TGTGCGTTCTGTTGAGTTTTTGAGGCGGAAATACTCTTGAGCTTGCCCCGTTCCGTTTGTGAGCCACGTGCAAGGC
 CACCTGCTCCACCCCTTGTCAATACTGACCGGCCCTTcCAGGAACAACATGGTACCATAGCCATCACTCTGCT
 CAGTACGGCGATGGGATCCAAGGTGCCATCGTGATCAATGGCCCCGGAACCGCCAATTACGACGAGGATTTAGGC
 CCCGTTGCCCTCACCGAAACCTACGATGAGACGGCATGGACGAAGAACTGGCTGGCGCTGCACGTGGCATTTCCCT
 CCTCAGCCCCCTCAACATTCTCTTCAATGGCTCCATGGTCAACAGCACCGGGCGGCGGCCGCTACAACACCATCTCA
 GTCAAGCAAGGCAAGACTTACCGGCTGCGCCTGATTAACATGAGCGTCGACACTTTCTTTCGTATTCTCCATGGAC
 GGGCACGAGTTCCAGATCATCACGGCCGACCTCGTCCCCGTGCACCTTACAATGCGACCTCGATCATGATCGGC
 ATCGGCCAGCGCTACGACATCGTCTTCAAGGCCAACCAGCCCCGTGCCAACTACTGGCTGCGTACCGAGATCGCC
 AGCTGCAGTGCCAACGCCATCACGGCCGAAGCCGACATCGTCCCCGGCGGCATCCTGAACTACGACACCATCGAC
 AAGACGGATCTGCCAGTCTCCACCAAGTCCGTTATCGAGACGACCGACTGCGCCGCCGAGCCCCACGACAAGCTG
 GTCCCCCTGGTGGGAGACGCAGGTCCCCAAGGACCAGTTCCTGACCCAGCTCGAGGGCATCGACCTGACGTTTCGCG
 GCGGGCGCCAGCGTCGGCAGCGAGACTGGTCTTGTGAGTGGTACCTGAACGACAGCGCCATGGTTCGTCGACTGG
 GCCAAACCCGACTCTGGAGTACTTCTCCGAGGGAGACACTAACTATACGTCTTCAATGAACGTTTTTCCAGATGCCC
 GCGGAGGGGAAGTGGTCGTTCTGGATCATCCACAACAACGCGGCCGCTCTGCTCGACCACCCGATCCATCTCCAC
 GGCCACGACTTTTTTCCACCTGGGCGCCGGCACCGGCACCTGGGACGGCAACGTGGACTCCTTGATCTTCGACAAT
 CCTATGCGCAGGGACGTGATGATCTTGCCACAGGATGGCTTATATCGCCTTTCCAGCGGACAAcCCCGGCGCG
 TGGTTGATGCATTGCCACATCGTAAGTACCCAACAAGCCCCATCTCTCTCTCTCTCTCATCAAAGATGACGGC
 AACTGACAAGCAACCAGGCATGGCACGTTACCGACGGGCTTTCCCTTGCAAGTTCGTTGAAAACCCAGGCTCATTTCA
 CGCAGGACCTCTCGGGCATGAAGAGCAACTGCGCCGCCCTGGAAAGAGTACGAGGAGAAGGCTTACTATGAGAAGG
 AGGTTGGCGATTCCGGCTTGTAACCGTGACGGCGGGGATGAGGAGTTTATGACGAGATGTTCACTTTATACCT
 TCTACCACCACCACCACCTTTTCCCTCTTTCTTTTACCGACGTAGGGCTGCACATCTGGCGCGTCTGTGGTCTG
 CTTTGCATATGCATAAATACCTT

SEQ ID NO : 53
 CHIỀU DÀI : 1824
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*

TÊN/KHÓA DẤU HIỆU : CDS
 VỊ TRÍ : (1).....(1824)
 atgggtgtcccttaagcaaatcggtgccactctggttgccactgactgctcaaaccttcgct
 M V S L K Q I G A T L L A L T A Q T F A

gctgccatccctgaggccgaaccggtggacctcgtcgcccggtcaagccacgaccactacg
 A A I P E A E P V D L V A R Q A T T T T
 tccaccacctcctcaaccacctcgaggggtccccgattccccgatgtacttggggcccttcg
 S T T S S T T S R V P D S R C T W G P S
 agcaggggttgctggaagaacggcttcagcattgccactgactttgacacgaagtggcca
 S R G C W K N G F S I A T D F D T K W P
 tctactggcaagactgtctcatatcgcttggaagttacgaatgtcaccaactgcgaggac
 S T G K T V S Y R L E V T N V T N C E D
 tatcaaagcaagggaattggagatggcttctgcaggccgatgctgctcatcaacaatcag
 Y Q S K G I G D G F C R P M L L I N N Q
 ttcccagggcctaccatcaatgcggaatggggagacaaccttgagattactgttgtaat
 F P G P T I N A E W G D N L E I T V V N
 agtatgcaagacaacggcacctcattccactggcacggcattcgtcagctgaactcatgc
 S M Q D N G T S F H W H G I R Q L N S C
 caaaacgatggtgccaacggcgctcaccgagtgccctattcctcccgcggaagcttact
 Q N D G A N G V T E C P I P P G G S F T
 tataagttcaaggctacacagtatggaacaacatggtaccatagccatcactctgctcag
 Y K F K A T Q Y G T T W Y H S H H S A Q
 tacggcgatgggatccaaggtgccatcgatgaatggcccggaaccgccaattacgac
 Y G D G I Q G A I V I N G P A T A N Y D
 gaggatttaggccccgttgccctcaccgaaacctacgatgagacggcatggacgaagaac
 E D L G P V A L T E T Y D E T A W T K N
 tggttggtgctgcacgtggcattccctcctcagccccctcaacattctcttcaatggctcc
 W L A L H V A F P P Q P L N I L F N G S
 atggtcaacagcaccggcgggcgccgctacaacacccatctcagtcaagcaaggcaagact
 M V N S T G G G R Y N T I S V K Q G K T
 tacggctgcgctgatttaacatgagcgctcgacacttttcttctgtattctccatggacggg
 Y R L R L I N M S V D T F F V F S M D G
 caccgagttccagatcatcacggccgacctcgtccccgtgcacccttacaatgcgacctcg
 H E F Q I I T A D L V P V H P Y N A T S
 atcatgatcggtcatcgccagcgctacgacatcgtcttcaaggccaaccagccccgtgcc
 I M I G I G Q R Y D I V F K A N Q P A A
 aactactggctgcgtaccgagatcgccagctgcagtgccaacgccatcacggccgaagcc
 N Y W L R T E I A S C S A N A I T A E A
 gacatcgccccggcggtcatcctgaactacgacacccatcgacaagacggatctgccagtc
 D I V P G G I L N Y D T I D K T D L P V
 tccaccaagtccgttatcgagacgaccgactgcgcgcgagccctacgacaagctggtc
 S T K S V I E T T D C A A E P Y D K L V
 ccctgggtgggagacgcaggtccccaaggaccagttcctgacccagctcgagggcatcgac
 P W W E T Q V P K D Q F L T Q L E G I D
 ctgacgttcgcggcgggcgccacggtcggcgagcgagactggtcttggtgcagtgggtacctg
 L T F A A G A T V G S E T G L V Q W Y L
 aacgacagcgccatggtcgtcgactgggccaaccgactctggagtacttctccgagggga
 N D S A M V V D W A K P T L E Y F S E G
 gacactaactatacgtcttcaatgaacgttttccagatgcccgcggaggggaagtggctcg
 D T N Y T S S M N V F Q M P A E G K W S
 ttctggatcatccacaacaacgcggccgctctgctcgaccacccgatccatctccacggc
 F W I I H N N A A A L L D H P I H L H G
 cacgactttttccacctgggcgcgggcaccggcacctgggacggcaacgtggactccttg
 H D F F H L G A G T G T W D G N V D S L
 atcttcgacaatcctatgcgcagggacgtgatgatcttgcccacaggatggcttattatc
 I F D N P M R R D V M I L P T G W L I I
 gcctttccagcggaacccccggcgctggttgatgcattgccacatcgcatggcacgtt
 A F P A D N P G A W L M H C H I A W H V
 accgacgggctttccttgacgttcgttgaaaaccaggtcattcacgcaggacctctcg
 T D G L S L Q F V E N P G S F T Q D L S
 ggcattgaagagcaactgcgccgctggaaagagtacgaggagaaggcttactatgagaag
 G M K S N C A A W K E Y E E K A Y Y E K
 gaggttggcgattccggcttgtaa
 E V G D S G L -

SEQ ID NO : 54
 CHIỀU DÀI : 607
 LOẠI : PRT

SINH VẬT : *M. phaseolina*

MVSLKQIGATLLALTAQTFAAAIPEAEFVDLVARQATTTTSTTSSTTSRVPDSRCTWGPSSRGCWKNGFSIATDF
DTKWPSTGKTVSYRLEVNTNVEDYQSKGIGDGFRCRPLLINNQFPGPTINAEWGDNLEITVVNSMQDNGTSFH
WHGIRQLNSCQNDGANGVTECPIPPGGSFTYKFKATQYGTWYHSHSAQYGDGIQGAIVINGPATANYDEDLGP
VALTETYDETAWTKNWLALHVAFFPQPLNILFNGSMVNSTGGGRYNTISVKQKTYRLRLINMSVDTFFVFSMDG
HEFQIITADLVVPVHPYNATSIMIGIGQRYDIVFKANQPAANYWLRTIEIASCSANAITAEADIVPGGILNYDTIDK
TDLPVSTKSVIETTDCAAEPYDKLVPWWETQVPKDQFLTQLEGIDLTFAAGATVGSETGLVQWYLND SAMVVDWA
KPTLEYFSEGDTNYTSSMNVFQMPAEGKWSFWIIHNNAALLDHPHILHGHDFHFLGAGTGTWDGNVDSLIFDNP
MRRDVMILPTGWLIIAIFPADNPGAWLMHCHIAWHVTDGLSLQFVENPGSFTQDLSGMKSNCAAWKEYEEKAYYEK
EVGDSGL*

SEQ ID NO : 55

CHIỀU DÀI : 2147 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

GCCACGGCTCCATTTGGATCGTCACTCACCACCACCACCACAACCACTCTCTACTCACCTCTTCACCAGCCCCCT
CTTTCTCCTCACTGGCATCCATTTCGTTGAAGTCGCTGTCCCCTTTTCTTGTCTTTTCGTCTTCTTCCACCCTCAGA
ATGAGGGCTTCTTACCTCTCTGCGGCTGCCTTCCTGGGCTCTCGGCTGCCGCGCCGCAAGCCGCCAGCACTTCT
TCTTCTCTGCTTCAGCCAACGCCAGTTCACCAGCTCAGCAGCTACTTCCACTTGCCTGGAACACTGCCGAC
GACCGCACTGTGTGGTGCGACTACGACATCAGCACCAGCTACTACAACGACGGACCCGACACGGGGGTACCCGC
GAGTACTACTTCGTCTGTCAGCGACGTGACCGTCTCGCCCGATGGCATCTCGCGCTCCGCTATGGCGGTGAACGGC
AGCATCCCCGGCCCCACCATTTTCGCGGCTGGGGTGACACAGTGAAAGTTACTGTCTACAACGACCTCACCAG
AGCGGCAACGGCTCTTCCATCCATTGGCAGGTATCCGGCAGAACTACACGAACCAAGATGATGGTGTGGTGTCT
ATTACGCAATGCCGATTGCGGTGCGCGAGACCTACACCTACGAGTGGAAGGCCACGCACTACGGCTCTTCTCTGG
TACCATCTCACATTGGCCTGCAGGCTGGGAGGGTGTTCGCGGTGGTATCATTATCAACGGTCCCGCTACTGCA
AATTACGACGAGGACCTCGGCATCATGTTTCTCAATGATTGGGATCACTCGACTGTTGACGAGCTCTACGATTCA
GCTCAGAGCAGCGGTCTCTACGCTTGACACCGGTCTCATTAACGGAACCAACATCTACAATGACTCCGGAACA
GTTACTGGATCTCGCTGGGAGGCCAGCCTGACCGAGGGTACCAGTACCAGGCTCCGCTCTTGTCAACGCTGCTGTA
GACTCGCACTTCAAATTTTCGATCGATAACCACACCCCTCCAGGTTATCGCCATGGACCTGGTCCCCATTGAGCCC
TACGAGACTACTGTTCTGGACATTGGCATGGGTGACGCTACGACGTCATTGTTACGGCAGACCAGGCTTCTGTT
GCTTCTGATTTCTGGCTTCGCGCAATTCGCCAGACCGCTGCTCGGACAACGATAACGCAGATGATATCAAGGGC
ATAATCCACTACGGATCATCAACTGGTACTCCGGAACCACTGCTTAGTAAGTTCGGATGTCCACATCTTCCGAG
CTCTCCGCTAATGATTCAACAGTGATTACACTGATGCCTGTGTCGATGAGGACAGCTCTGATCTCGTAAGTGCTC
GAGATTACCGTCTGCTGGAAAGAGTAGGCGCGCTGACTGTGATCTTTAGGTCCCGTATGTCTCTAAGACTGCCA
CCTCCGGTACCTCCCTGGCCGAGGCTGTTCCGTCGGCTACAACCTCGGACAACCTCTTCCGCTGGTACATGAACG
AGACCTCTATGGAGGTGCGAGTGGGAGAATCCAACCCCTTCTGCAAGTCTACAACGACAATCTGACGTTCACTGACA
CATCGGGTGTGTTCAACTTGACACCGCAGACCAATGGTACTTCTTCGTCATCGAGACCGACAACGCTGTGCCAC
ACCAATCCATCTTACGGCCACGACTTCTTCGTCCTGGCTGCGGGCACCAGGCTCTTACAGTTTACAGCTTACTC
TGACTCTGGATAACCCCTCCCCGCGGACACGGCTATGCTTGACTCCTCTGGCTACTTGGTCTCTGGCTTTTCGAGA
CCGACAACCCAGGTGCGTGGTTGATGCACTGCCACATCGGCTGGCACACCAGCGAGGGCTTTGCCCTTCAGATCT
TGGAGCGCTACACCGAGATCCAGGATAGCCTGATCGACTACGACGTCCTCAATGACACCTGCTCGACTTGGTCTA
CTTACTCCGAGGCAAACTCGATCGAGGAGGAGGACTCTGGTGTGTAAGGAAAACCTCTGAGTGACAATCCGTCTAT
GGCAGGTGTGCGGGGCGCTTTGGAGCTCTCTCTATCTCTGTGAAAGATCTTGTATATAACGTGATGCCAGCCTCT
CCCTGACATTTGGTGCTTCGGTTGATCTACATATATTTCCCTTATTTA

SEQ ID NO : 56

CHIỀU DÀI : 1737

LOẠI : ADN

SINH VẬT : *M. phaseolina*

TÊN/KHÓA DẤU HIỆU : CDS

VỊ TRÍ : (1).....(1737)

atgagggcttcttacctctctgcggtgccttcctgggctctcggtgccgcgccgcaa
M R A S Y L S A A A F L G L S A A A P Q
gccgccagcacttcttcttctctgcttcagccaacgccagttccaccagctcagcagct
A A S T S S S S A S A N A S S T S S A A
acttccacttgactggcaacactgccgacgaccgactgtgtggtgcgactacgacatc
T S T C T G N T A D D R T V W C D Y D I
agcaccgactactacaacgacgacccgacacgggggtcaccgcgagtagtactacttcgtc
S T D Y Y N D G P D T G V T R E Y Y F V
gtcagcgacgtgacgtctctgcccgatggcatctcgcgctccgctatggcggtgaacggc
V S D V T V S P D G I S R S A M A V N G
agcatccccggccccaccattttcgccgactgggggtgacacagtgaaagttactgtctac

S I P G P T I F A D W G D T V K V T V Y
 aacgacctcaccacgagcggcaacggctcttccatccattggcacggtatccggcagaac
 N D L T T S G N G S S I H W H G I R Q N
 tacacgaaccagaatgatgggtgtgggtgtctattacgcaatgcccgattgcggtcggcgag
 Y T N Q N D G V V S I T Q C P I A V G E
 acctacacctacgagtgggaagggccacgcagtagcggctcttcctgtgtaccactctcacatt
 T Y T Y E W K A T Q Y G S S W Y H S H I
 ggctgagggcctgggaggggtgttttcgggtgtatcattatcaacgggtcccgtactgca
 G L Q A W E G V F G G I I I N G P A T A
 aattacgacgaggacctcggcatcatgtttctcaatgattgggatcactcgactgttgac
 N Y D E D L G I M F L N D W D H S T V D
 gagctctacgattcagctcagagcagcgggtcctcctacgcttgacaccgggtctcattaac
 E L Y D S A Q S S G P P T L D T G L I N
 ggaaccaacatctacaatgactccggaacagttactggatctcgtctgggagggccagcctg
 G T N I Y N D S G T V T G S R W E A S L
 accgaggggtaccagctaccgggtccgtcttgtcaacgctgctgtagactcgacacttcaaa
 T E G T S Y R L R L V N A A V D S H F K
 ttttcgatcgataaaccacacctccaggttatcgccatggacctgggtccccattgagccc
 F S I D N H T L Q V I A M D L V P I E P
 tacgagactactgttctggacattggcatgggtcagcgctacgacgtcattgttacggca
 Y E T T V L D I G M G Q R Y D V I V T A
 gaccaggtcttctgttctgttctgatttctggcttcgcgcaattccccagaccgctgctcg
 D Q A S V A S D F W L R A I P Q T A C S
 gacaacgataacgcagatgatatacaagggcataatccactacggatcatcaactgggtact
 C N D N A D D I K G I I H Y G S S T G T
 ccggaaccactgcttatgattacactgatgcctgtgtcgatgaggacagctctgatctc
 P E T T A Y D Y T D A C V D E D S S D L
 gtcccgatgtctctaagactgccacctccggtacctccctggccgaggtgtttccgtc
 V P Y V S K T A T S G T S L A E A V S V
 ggctacaactcggacaacctcttccgctggtacatgaacgagacctctatggaggtcgag
 G Y N S D N L F R W Y M N E T S M E V E
 tgggagaatccaaccttctgcaggtctacaacgacaatctgacgttctactgacacatcg
 W E N P T L L Q V Y N D N L T F T D T S
 ggtgttgttcaacttgacaccgcagaccaatgggtacttcttcgtcatcgagaccgacaac
 G V V Q L D T A D Q W Y F F V I E T D N
 gctgtgccacacccaatccatcttcacggccacgacttcttcgtcctggctgcggggcacc
 A V P H P I H L H G H D F F V L A A G T
 ggctcttacagttcagacgttactctgactctggataacctccccgcgcgacacggct
 G S Y S S D V T L T L D N P P R R D T A
 atgcttgactcctctggctacttggctcctggcttttcgagaccgacaacctcaggtgcgtgg
 M L D S S G Y L V L A F E T D N P G A W
 ttgatgcactgccacatcggctggcacaccagcgagggtttgcccttcagatcttggag
 L M H C H I G W H T S E G F A L Q I L E
 cgctacccgagatccaggatagcctgactacgacgtcctcaatgacacctgctcg
 R Y T C E I Q D S L I D Y D V L N D T C S
 acttgggtctacttactccgaggcaaaactcgatcgaggaggaggactctgggtgtgtaa
 T W S T Y S E A N S I E E E D S G V -

SEQ ID NO : 57

CHIỀU DÀI : 578

LOẠI : PRT

SINH VẬT : M. phaseolina

MRASYLSAAAFGLGLSAAAPQAASTSSSSASANASSTSSAATSTCTGNTADDRTVWCDYDISTDYNDGPD TG VTR
 EYYFVVS DVT VSPDGISRSAMAVNGSIPGP TIFADWGD TVKVTVYNDLT TSGNGSSIHWHGIRQNYTNQNDGVVS
 ITQCPIAVGETYTYEWKATQYGSSWYHSHIGLQAWEGVFGGIIINGPATANYDEDLGIMFLNDWDHSTVDELYDS
 AQSSGPPTLD TGLINGTNIYNDSGTVTGSRW EASLT EGTSYRLRLVNA AVDSHF KFSIDNHTLQVIAMD LVPIEP
 YETTVLDIGMGQRYDVIVTADQASVASDFWLRAIPQTACSDNADNDDIKGIIHYGSSTGT PETTAYDYTDACVDE
 DSSDLVPYVSKTATSGTSLAEAVSVGYNSDNLFRWYMNETSMEVEWENPTLLQVYNDNLTFDTSGVVQLDTADQ
 WYFFVIETADNVPHPIHLHGHDFFVLAAGTGSYSSDVTLLDNPPRRDTAMLDSSGYLVLA FETDNPGAWLMHCH
 IGWHTSEG FALQILERYTEIQDSLIDYDVLNDTCTSTWSTYSEANSIEEEDSGV*

SEQ ID NO : 58

CHIỀU DÀI : 2302 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

TGCGCCCCCTCCCGCTGCTGGGCACTCTTTTATAAAGGCAGCTCTCCACTCCTGCCCTGTGAATGCTTGCTCAAG
GGCGCTCAGGCATAGGTAAGTAGGGTAAGTATTGAGGCGTCTGGCTTTGCGGCGAATCTTGGGACGATTTACAGCC
ATGACTCGCCTTCCGTTCTGCTCGGGCTCGTGCTACGGCTTTGGCCAAGACGGTCACTGTTGACTGGGACATT
GGCTGGGTTTCCAGGGCTCCCGATGGATTTCGAGCGGCTGTTCATCGCAATCAACGGCCAATGGCCGCTCCCTGTG
CTGGAGGCCGACGTCAACGACACCATCATCGCCACTGTCCACAATTCTCTTGGCAACGAAACGACCAGCATCCAC
TGGCAGGCGATGTGGCAGAGAGGCACGCCGAGCAAGACGGCGGGGCTGGCGTCAAGCAGTGTCCGATCCCGCCT
GGCGAGACTTTACGTACGAGTTCAAAGCATACCCGGCCGGTACTTTCTTGGTACCACTCGCATGACATGGGCGCAG
TATCCCGATGGCCTGCGCGCACCCATGATCATCCATGACCCCGACTCCGAAAACCCAGAAGAGCAGTGATGGTGAA
GTCTGTCTATGTGTCCGACTGGTACCACGACCAGATGCCGCCGCTTATCCACAGCTTCTTGACCACCCCCAAT
TTCAACGGCGCGATGCCCAACCCGAATCCAGCTTGATCAACGATCAGCAGTCCACGTCCATCAACATCCGTCCC
GGCGAGAAGAAATACGTGCGCATCATCAACAGTCCGCCCTCGCCACGTACTACCTGCAGTTTGGTGGGTTCCGC
TTCAGCCACCCCTCCCTGCGGAAGTGACTGACGATTTTACAGCAACACAACATCACCGTTGTGCAATTGACGG
TGTTGACGGTCCGTACTACCCCTGTGCTTATGCTGTGGCTTCGTGCATACCCGTGGCAGCTTCCCGCCGTTGC
ATGCCCGAGGCGATACAAGGTTTATGCTGCTGACCGTAGTCTGTCAGTGAACCGCAGAGCTGGAAGGCCCTGGAG
ATCGTCCCCGGCCAGCGGTACGACTTCATCATCGAAGGTCTCGAGAACCCCAAGGAACCTACGCATTCATCAAC
AAGATGGCTGTCTCGGTCTGCAGAACGTCAACAGCCTGGTCTACGACGAGTCCCTTCGGCGAGCCGGAGTCGTT
AGCCTGAGCTCTGGCGATCTCGGAAGTGATTTACCCCTGGTGCTCTAGATCACGAGCCTCTCTTGGAATCCGTG
GACCACACCATCACGATGGAGGTCAATAATTTGAACATTGATGGCGTTGGCTTTCCGTACGGCTGATCCCGTGGC
CGGGAGAAAATCTTGCTCACATGTGCCACAGCATCACTCAAGGCCCGGACCCGTACATTTACCCCCGCACGCCCTAC
CCTGTACACAGCCCTCAGCACCGGCTTCAACGCCACCCAGCCAGAAAATCTACGGCCAGGTGAACCCCTACGTCTG
CAACGCCCGGTGAAGTCGTCCGGCTCGTCTGCAACAGCAACGATCTCGTACCCGCCAACAACTCTGGCCGCGGGCA
CCCCATGCATCTGCACGGCCACGTTTCCAGGTGGTCCGTGAGTTCTCCGAGCACTGGGACGGCAACACCTCGTC
CTTCCCCGCCACGCCCATGAAGCGTGACACCACCGTCCATTCGTGGCGGCAGCCTGGTGCTCCAGTTCCGTGC
CGACAACCCCGGTGTTTGGCTGTGTAAGTATTTTTCGGATGCCATGGTGCCTGGACGTGAGCTGATGAGG
ATGGGCCCTTGCAAGTTCAGTCCATATTTAGTGCCACCTAGACGCCGGCATGTCCGCTACAATCATCGAGGCGCCG
CTCGACTTGCAGCGCGAGGGCATCAAGATCCCCCAGCAGCATCTTGAATCGTGCAGAGCCTTGAACCTTGACCACT
CAAGGCAATTGCGCCGGCAACACCGCCAACCTGGATGACACTGCCGCTGCAGGGTCTACGACACCGATCCATGG
GGGTAGGTATATCGGAGTTTAATTGAGCTTGAAGGACATGTGCTGACCGTTGCGTTTAGTGCGCTTATCACGGA
TGACGGTGGTGAAATAGTACCCCTGAACGGAACCACTTATAAGAGAAATTTAGTGTAATCTTAATACGTCTTCA
CATGTACACAACCTATGCATTGTATTCAAATTTACATAAGCACTTCGACAGTAGTCTGAGTTGACTGACTAGGC
TCGTGGACGGAGGACGCCGTTTCTCTGCTCTTGCCTGCTAAATTTTCGCGAT

SEQ ID NO : 59
CHIỀU DÀI : 1665
LOẠI : ADN
SINH VẬT : *M. phaseolina*
TÊN/KHÓA DẤU HIỆU : CDS
VỊ TRÍ : (1).....(1665)

atgactcgccttccgttcgtcctcgggctcgtggctacggctttggccaagacgggtcact
M T R L P F V L G L V A T A L A K T V T
gttgactgggacattggctgggttccagggtcccgatggattcgagcggcctgtcatc
V D W D I G W V S R A P D G F E R P V I
gcaatcaacggccaatggcgctccctgtgctggaggccgacgtcaacgacaccatcatc
A I N G Q W P L P V L E A D V N D T I I
gccactgtccacaattctcttggcaacgaaacgaccagcatccactggcacggcatgtgg
A T V H N S L G N E T T S I H W H G M W
cagagaggcacgcccagcaagacggcggggtggcgctcacgcagtgatcccgatcccgct
Q R G T P E Q D G G A G V T Q C P I P P
ggcgagactttcacgtacgagttcaaagcatacccgccggtactttctggtaccactcg
G E T F T Y E F K A Y P A G T F W Y H S
catgacatggggcagtatcccgatggcctgcgcgacccatgatcatccatgaccccgac
H D M G Q Y P D G L R A P M I I H D P D
tccgaaaccagaagagcagtgatggtgaagtcgtgctctatgtgtccgactggtaccac
S E T Q K S S D G E V V L Y V S D W Y H
gaccagatggcgccgttatccacagcttccctgaccaccccccaatttcaacggcgcgatg
D Q M P P L I H S F L T T P N F N G A M
cccaacccgaactccagcttgatcaacgatcagcagtcacgtccatcaacatccgtccc
P N P N S S L I N D Q Q S T S I N I R P
ggcgagaagaaatcagtgcgcatcaacacgtccgcccctcgccacgtactacctgcag
G E K K Y V R I I N T S A L A T Y Y L Q
tttgaccaacacaacatcaccgttgctcgcaattgacgggtgttgacgtcgaaccgcagagc
F D Q H N I T V V A I D G V D V E P Q S
tggaaggccctggagatcgtccccggccagcgggtacgacttcatcatcgaaggtctcgag

W K A L E I V P G Q R Y D F I I E G L E
aaccccacaaggaactacgcattcatcaacaagatggctgttctcggtctgcagaacgtc
N P T R N Y A F I N K M A V L G L Q N V
aacagcctgggtctacgacgagtccttcggcgagccggagtcgttcagcctgagctctggc
N S L V Y D E S F G E P E S F S L S S G
gatctcgggaagtgatttcaccctgggtgcctctagatcacgagcctctcttggaaatccgtg
D L G S D F T L V P L D H E P L L E S V
gaccacaccatcacgatggaggtcaataatttgaacattgatggcgttggctttcgcatc
D H T I T M E V N N L N I D G V G F R I
actcaaggcccggaaccgtacatttcaccccgacgcctaccctgtacacagccctcaggc
T Q G P D P Y I S P R T P T L Y T A L S
accggcttcaacgccaccgacccagaaatctacggccaggtgaacccctacgtcgtcaac
T G F N A T D P E I Y G Q V N P Y V V N
gccggtgaagtgcgtccggctcgtcgtcaacagcaacgatctcgtcaccgccaacaactct
A G E V V R L V V N S N D L V T A N N S
ggccgcggggcaccatgcatctgcacggccacgttttccaggtgggtcggtcagttctcc
G R G H P M H L H G H V F Q V V G Q F S
gagcactgggacggcaacacctcgtccttccccgccacgcccattgaagcgtgacaccacc
E H W D G N T S S F P A T P M K R D T T
gtcctattcgtcgtggcggcagcctgggtgctccagttccgtgccgacaaccccggtgtttgg
V L F A G G S L V L Q F R A D N P G V W
ctgtttcactgccatattgagtggtgacacctagacgccggcatgtccgctacaatcatcgag
L F H C H I E W H L D A G M S A T I I E
gcgccgctcgacttgcagcgcgagggcatcaagatccccagcagcatcttgaatcgtgc
A P L D L Q R E G I K I P Q Q H L E S C
agagccttgaacttgaccactcaaggcaattgcgcccggcaacaccgccaacctggatgac
R A L N L T T Q G N C A G N T A N L D D
actgccgcctgcagggctctacgacacccgatccatgggggtgcgcttatcacggatgacggg
T A A C R V Y D T D P W G A L I T D D G
ggtggaaatagtagccctgaacggaaccacttataagagaatttag
G G N S T L N G T T Y K R I -

SEQ ID NO : 60

CHIỀU DÀI : 554

LOẠI : PRT

SINH VẬT : M. phaseolina

MTRLPLFVLGLVATALAKTVTVVDWDIGWVSRAPDGFERPVIAINGQWPLPVLEADVNDTIIATVHNSLGNETTSIH
WHGMWQRGTPEQDGGAGVTQCPIPPGETFTTYEFKAYPAGTFWYHSHDMGQYPDGLRAPMIIHDPDSETQKSSDGE
VVLVSDWYHDQMPPLIHSFLTTPNFNGAMPNPNSSSLINDQQSTSINIRPGEKKYVRIINTSALATYYLQFDQHN
ITVVAIDGVDVEPQSWKALEIVPGQRYDFIIEGLENPTRNYAFINKMAVLGLQNVNSLVYDESFGEPEFSLSG
DLGSDFTLVLPLDHEPLLESVDHTITMEVNNLNIDGVGFRITQGPDPYISPRTPPLYTALSTGFNATDPEIYGQVN
PYVVNAGEVVRLVNSNDLVTANNSGRGHMHLHGHVQVVGQFSEHWDGNTSSFAPATPMKRDTTVLFAGGSLVL
QFRADNPGVWLFHCHIEWHL DAGMSATII EAPLDLQREGIKIPQQHLESCRALNLTQGN CAGNTANLDDTAACR
VYDTPW GALITDDGGNSTLNGTTYKRI*

SEQ ID NO : 61

CHIỀU DÀI : 2460 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : M. phaseolina

CTCCGTATCTACTGCTGACTTCACATCCACAGCATGCAAGGACTCGAGGGCTTCTGATTTAACGAGCACTCACTT
GGCTTTTGTCTGCGTTTTCTTTTCGTTTTCTACTATTTTCATCTTGATCCATTCTGTTCTCGCCGTCGCTTTCCAG
ATGTCCTTTCTGTTTCGAAAATTTTCACAGGCCTCATCGCTTTTACCGCTGGCTTAGGCCAAGATGAGACTAATGGG
TGAGTATGCTGAACCTTGCGCAGGATATCTGAATAATAGTTGTAGGGTTTCGAAATGGGGCACGTTTCGATGCACC
GAGGCTCCAGAAGTTTCTGGATCGGGTGAGGTTATTTCTGAAACATTACTGAGCTCGGCACTGATTGGTCGGCAGG
GTCATGGCATTCCCTGGGGCGGTATGACTTGTACCAACGCCAACCCGTATACTGAGTAGGACTCGGCCCCGTGGTC
CACAATCTTCTCTGCGCTGACGTTTCGGAAAGAGCTCCAGACACAGGACAGACTGTGCGGTATGATTTACAGTCC
AGCGGCATCCGGTGTCTCCGGACGGCTACAAGAAGAAGCTTTTGTCTGTGAACGGGCAATTTCCAGGCCCGCTCA
TGGAAGCTAACTGGGGAGACACAATTGAAGGTAGGATGAGAGTAAATACGGGCAGCAACAACCTGCTGACAGATAC
AGTGACTGTGCACAACAACATAGCCGGACCTGAGGAAGGCACACAAATCCACTGGCAGGGCTTCACGCAGAGAGG
GACGCCGTTTCATGGATGGTATCCCTTCCGTATCAAGCTGCCCCATTGCGCCCAACAATACCTTTGTGTATACCTT
CAAGGCAGACCTTTACGGCACTGGCTGGTACCCTCTCATTACTCTGGGCAATCCACCGCGGCCCTCCTCGGCC
AATCGTCGTCATGGTCCCAGTGCCTTGACATGATATTTGACCTTGGCCCTGTGTTTTGAATGATGGTACCA
CAAGGACTACTTGCAGCTTATTGACGGTGGTGAGACTTTCCCGTAGGTCAAAGGGCTCGATAGCTGAAACCTCCA
ACAGTCTCGGAACGGACCCAGCCTATGGCATCCCAAGGCGGACAACAACATGATAAACGGGAAGATGGACTAC

GATTGCTCCCTTGTCACGTGACGGCAGCCCTTGCGTCTCTAATGCCGGCTTGGCCACGTTTCAGCTTCACCAAGGGC
GCCACGCACCGTCTCAGGCTCATAAACGGAGGATCCGCTCCCTGCAGCACTTCAGTATCGATGGACACGAGATG
ACAGTCATTTCCAATGACTTCGTAGCCGTAGAGCCATACCAGACGAAGACAGTGACTCTTGCGTAAGCCGCCCG
ATTTTTCCAGATATCCCACATTACTAACACCCACCAGGTTCGGCCAGCGGACAGACGTCTCTCGTCACGGCCAACGG
CGACGCCACCGGCGCCTACTGGATGCGCAGCACCGTTCGCGGACGATGAATCCTGCAACTGGTCCAACCAGCCCCG
CGCGCTCGCCGCGAGTCTACTACGACGCGGCAAAACCCACCGTCAAGCCCCAACAGCACCGGCTGGCCCCCGTTCG
CAACCAGCAAGGCAGCTGCGACAACGACCCGCTTACCCAAACCATCCCCCTCTTCCCCATCCCCGCGGACCCAG
CCCCCAACGACGCTTGAGCTCGACTTCGGCTGGACGCAAAACGCAACCGGACACCAAGTCTGGACGGTCAACGA
CCGCGGCTTCCGCGGCAACTACAACCGCCCCGTGCTGCAGCTCGCCGCGGGCTCGGACACCGCGTCTCGCGCTA
CGCGTGGGAGCCGAATGGAACGTCTACGACACGGGCGGCAACCGCACCGTCCGCATCGTCATGCACAACAACCTC
CTCCATGTACCACGTACGTGAGCCCTTGCTCTCCCCCCCCCCCCCCCCCTCCCCCTCTCGCTAATGGAAAACACATG
CAGCCCATGCACCTCCACGGGCACAACGCGCAGATCCTCGCCGCGGGCGCCAACGGCCCCCTGGGACGGCCGCACC
GTCGCGCGCCCCGCCAACCCGGCCCCGCGCGACGTCTACCAGTCTCCGCGGAACGGGCACCTGGTGATCCAGTAC
GCGCAGGACAACCCGGGCGTGTGGCCGCTGCACTGCCACATCGCGTGGCACGCGAGCGCGGGCATGTTTGCGAGC
GTGCTGGAGCGGGCGGGGACATTGTGGGGAGCAGCGGGGCCGGCGGGTGGAGGGAGGAGATGGCGGGCGTGTGT
GCCGGGTGGGAGGCGTACACGCGGATGAACGTGGTGGAGCAGGTGGACAGTGGGGTGTGAGTTGGGAGCAAGGAG
GGGGGTTTGAGATGCGCGTATATGGTAAGCGGTGGGGTTCTTGTGTTGACTCGATTGCTCTGCAGGCGGTGTCT
CTTTTCGACTGTGGCGCGTGCAATCTCGTTTCGTGCATCCTGGGTGGATAGGGAGCGGCG

SEQ ID NO : 62
CHIỀU DÀI : 1803
LOẠI : ADN
SINH VẬT : M. phaseolina
TÊN/KHÓA DẤU HIỆU : CDS
VỊ TRÍ : (1).....(1803)

atgtctttcgtttcgaaaattttcacaggcctcatcgcttttaccgctggcttaggcaa
M S F V S K I F T G L I A F T A G L G Q
gatgagactaatgggttttcgaaatggggcacgttcgatgcaccgagggtccagaagttt
D E T N G V S K W G T F D A P R L Q K F
ctggatcggggtcatggcattccctggggcggtatgacttgtaccaacgccaacccgtat
L D R G H G I P W G G M T C T N A N P Y
actgaagctccagacacaggacagactgtgcggtatgatttcacagtccagcggcatccg
T E A P D T G Q T V R Y D F T V Q R H P
gtgtctccggacggctacaagaagaacgttttgcttgtgaacgggcaatttcaggccccg
V S P D G Y K K N V L L V N G Q F P G P
ctcatggaagctaactggggagacacaattgaagtgactgtgcacaacaacatagccgga
L M E A N W G D T I E V T V H N N I A G
cctgaggaaggcacacaaatccaactggcacggcttcacgcagagaggagcgcgttcgat
P E E G T Q I H W H G F T Q R G T P F M
gatggtatcccttccgtatcaagctgccccattgcgccaacaatacctttgtgtatacc
D G I P S V S S C P I A P N N T F V Y T
ttcaaggcagacctttacggcactggctggtaccactctcattactctgggcaatccacc
F K A D L Y G T G W Y H S H Y S G Q S T
ggcggcctcctcggcccaatcgctcgccatgggtccagtgcgcttgactatgatattgac
G G L L G P I V V H G P S A L D Y D I D
cttggccctgtgtttttgaatgactggtaccacaaggactacttgacgcttattgacggt
L G P V F L N D W Y H K D Y L Q L I D G
gtcgtcggaacggacccccagcctatggcatcccaaggcgggacaacaacatgataaacggg
V V G T D P S L W H P K A D N N M I N G
aagatggactacgattgtcccttgtcactgacggcacgccttgctctctaattgccggc
K M D Y D C S L V T D G T P C V S N A G
ttggccacgttcagcttcaccaaggcgccacgcaccgtctcaggctcataaacggagga
L A T F S F T K G A T H R L R L I N G G
tccgcctccctgcagcacttcagtatcgatggacacgagatgacagtcatttccaatgac
S A S L Q H F S I D G H E M T V I S N D
ttcgtagccgtagagccataaccagacgaagacagtgactcttgacgtcggccagcggaca
F V A V E P Y Q T K T V T L A V G Q R T
gacgtcctcgtcacggccaacggcgacgccaccggcgccctactggatgcgcagcaccgtc
D V L V T A N G D A T G A Y W M R S T V
gcggacgatgaatcctgcaactggtccaaccagccccgcgcgtcgccgcagtcactac
A D D E S C N W S N Q P A A L A A V Y Y
gacgcggcaaacccaccgtcaagcccaacagcaccggctggcccccgctcgccaaccag
D A A G N P T V K P N S T G W P P V A N Q
caaggcagctgcgacaacgacccgcttacccaaaccatccccctcttccccatccccgcg

Q G S C D N D P L T Q T I P L F P I P A
gacccccagccccctcaacgacgcttgagctcgacttcggtggacgcaaaacgcaaccgga
D P S P S T T L E L D F G W T Q N A T G
caccaagtctggacggtcaacgaccgcggttccgcggaactacaaccgccccgtgctg
H Q V W T V N D R G F R G N Y N R P V L
cagctcgccgcggtcgacaccgcgctcgctccgctacgctgggagcccgaatggaac
Q L A A G S D T A S S A Y A W E P E W N
gtctacgacagggcgcaaccgcaccgtccgcatcgatgcacaacaactcctccatg
V Y D T G R N R T V R I V M H N N S S M
taccacccccatgcacctccacgggcacaacgcgcagatcctcgccgcgggcgccaacggc
Y H P M H L H G H N A Q I L A A G A N G
ccctgggacggccgcaccgtcgcgcgccccgccaacccggcccgccgacgtctaccag
P W D G R T V A R P A N P A R R D V Y Q
ctcccgccgaacgggcacctggtgatccagtacgcgaggaacaacccggcggtgtggccg
L P P N G H L V I Q Y A Q D N P G V W P
ctgcaactgccacatcgctggcacgagcgcgggcatgtttgagcgctgctggagcgg
L H C H I A W H A S A G M F A S V L E R
gcgggggacattgtggggagcagcggggccggcggtggagggaggagatggcgggcggtg
A G D I V G S S G A G G W R E E M A G V
tgtgccgggtgggagggctacacgcggatgaacgtggtggagcaggtggacagtggggtg
C A G W E A Y T R M N V V E Q V D S G V
tga
-

SEQ ID NO : 63

CHIỀU DÀI : 600

LOẠI : PRT

SINH VẬT : *M. phaseolina*

MSFVSKIIFTGLIAFTAGLGQDETNGVSKWGTFDAPRLQKFLDRGHGIPWGGMTCTNANPYTEAPDTGQTVRYDFT
VQRHPVSPDGYKKNVLLVNGQFPGPLMEANWGDTEIVTVHNNIAGPEEGTQIHHWGFTQRGTFPMDGIPSVSSCP
IAPNNTFVYTFKADLYGTGWYHSHYSGQSTGGLLGPIVHGPSALDYDIDLGPVFLNDWYHKDYLQLIDGVVGT
PSLWHPKADNNMINGKMDYDCSLVTDGTPCVSNAGLATFSFTKGATHRLRLINGGSASLQHFSIDGHEMTVISND
FVAVEPYQTKTVTLAVGQRTDVLVTANGDATGAYWMRSTVADDESCNWSNQPAAALAAVYDAANPTVKPNSTGWP
PVANQQGSCDNDPLTQTIPFPADPSPSTTLELDFGWTQONATGHQVWTVNDRGFRGNYNRPVLQLAAGSDTAS
SAYAWEPENWYDTRNRTRIVMHNNSMYHPMHLHGHNQILAAAGANGPWDGRTPANPANPARRDVYQLPPNG
HLVIQYAQDNPGVWPLHCHIAWHASAGMFASVLERAGDIVGSSGAGGWREEMAGVCAGWEAYTRMNVVEQVDSGV
*

SEQ ID NO : 64

CHIỀU DÀI : 2285 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

TGAAGCCCTCTCTGACGCAGGTATCTATTTTATTAAACCAGTTTCACCCCTCCGGTCCGGTAAGTATTTGCTCAAA
CAAACCTTGTCCTCAAGTGTCTAGTTAGGTAAGCTAGGAGGCTCACGAATTCAAGAGACTCTCCAGAGATATCTTCGCC
ATGATTTCGCCCTTTCTCTCGTGCTCGGCTTCATGGCCACGACACTGGCCAAGACCGTCACCTCTTAACCTGGGATATT
GGGTGGGTCTTCTGCGGCTCCAGATGGATTTACACGGCCCGTCAATTGGCATCAACGGGGAGTGGCCGCCCCcGTT
TTGGAAGCCGACGTCAACGACACTATTATAGTAATTACGCGGAACCTCCTGCGTAACGAGACGACGAGCTTGAC
TGGCATGGTATGTGGCACTATAACTCGACCCACATGGACGGCGGAGCCAGGATTTACAGTGTGAAATCCCTCCG
GGGGGGACATTACGTACAAGTTCAAGGCGTACCCGGCCGGCACCTTTTGGTATCATTTCTCACGATATGGGCCAA
TATCCCGACGGCTGCGGCCCAATGATCATTCACGACCTAAGGCTGCCGCGGAGCGGGACACTGATAAGGAG
TACGTACTTACGGTCTCCGACTGGTACCGTGACCAGATGCCGTCACTTATCCACCGCTACTTGACAACCTCCACT
TATAATAGTACTATGCCAAATCCAAACTCGAGCTTGATCAACGATCAGCAGTCTACAACGCTAAACATCCGCCCC
GGGCAGAAGATATATGTTCCGATTATTAACATGTCAGCCCTCGCAACGTACTATCTACAGTTCCGGTAGGTGTATC
CATATGGCCCTGTTACACGGACGAAGCTAACGATTAAGATCAACACCACCTTGACTGTTATTGCCATTGACGGTGT
CGACGGTTGGTAACCTCTATTACCTATTCTGCGCTAAGATGCATGTCGGCATATGTGCCGTTAGCTGCATGCAAC
GCGTTACAAGGTTAAGTGCTAACAACGGGCAGTCGATCCCCAGACTTGGGAGGCTCTGGAGATTATCCCTGGACA
GCGGTACGACGTCAATTATCACCGGCCCTAGAGAACCCGAAAGGAACATGCATTTATCAACAAGATGGCTACTCT
TGGCTTCCAGAACAATAACGTCCTTAGTTACGATTCGTCTTGGCTGTCCCGGAGCCGTTGAACGTGGGCAGCTCT
CAATCTTAGAAGCGATTTCACCTGACCCCGCTTAGTGAGGAGCTACTTGGAGCCCGTGGACACACCTTCAC
CATGGAGGTCAATAACGTGAACGTCGACGGCGTAGGCTCCCGTGAGATATTTCCCTACCCCGGCCCCCGGGG
CGGCTCCCGAAGTATGAGGCTTACTCACTCCAATCCGCAGCATCACGCAGGGACCGGACCTTACATCGCC
CCGCGTACGCCCACTTATACACCACCTGAGCACCGGCTCTAATGCTATTAACCCCGCATATACGGCCAGGCG
AACGCTTACGTAGTGAAGCCGGCGATATTGTCCAGCTTGTGCTCAATAGTAACGAACCCGCTCACTACCAACACT
TCCGGTCGTGGGCACCTTATGCACCTTGCACGGCCACACTTCCAAGTGGTTGGCCAATATGGCAGCCCTTGGGAC
GGCGATGCCTCTAAATTCCTGTGTTTCTATGAAGCGGGATACCACTGTTCTGTTTACTGGCGGGAGCTTGGTG

ATCCGGTTCCGCGCGGACAATCCTGGAGTTTGGATGTGTACGGCTCCGCTTTGCTCGAAGACCCGCGTTTGAAAG
 CGTTGACTGACCTGTAACCTCGCAGTCCATTGCCACAACGAATGGCACCTTGACGCCGGCATGGCTGGAACGATTA
 TCGAAGCGCCACTCGAGCTTCAACAAAGCGGTCTGACGATTCCGCCGCGAGCACCTCGCGTCGTGCAGGGCGTTAA
 ACTTAACGACCCGGGGCAATTGTGCCGGTAATACTGCGAACCTAGAGGATACGGCTGCATGCAGAGTCTACGACA
 CTGAGCCTTGGGGGTGAGTTACTCCATTATTGTTTTGTTGTTATAACCCCTTGCTGATGACAATACTTCTCTAGTG
 CACTTATCAAGAGAGATGAGGAAACAGCATATTAAATAGCACGCTGGCATAGCACCTGTAGATGTAGACTGATTT
 CTAGTATTTATTACCTGTGTACACTGAAGAAAATTGTTGTAAAATGATATGTCCCCATGCAACTGAACATGGCC
 TACGTGGCGAGAGTTTATCAGGCGCCTTCTGCTCC

SEQ ID NO : 65
 CHIỀU DÀI : 1635
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 TÊN/KHÓA DẤU HIỆU : CDS
 VỊ TRÍ : (1).....(1635)

atgattcgccttttctctcgtgctcggcttcatggccacgacactggccaagaccgtcact
 M I R L S L V L G F M A T T L A K T V T
 cttactgggatattgggtgggtttctcgcggctccagatggatttacacggcccgctcatt
 L N W D I G W V S A A P D G F T R P V I
 ggcattcaacggggagtgccgcccccggttttggaaagccgacgtcaacgacactattata
 G I N G E W P P P V L E A D V N D T I I
 gtaattacgcggaacctcctgcgtaacgagacgacgagcttgcaactggcatgggtatgtgg
 V I T R N L L R N E T T S L H W H G M W
 cactataactcgaccacatggacggcggagccaggatttcacagtgtgaaatccctccg
 H Y N S T H M D G G A R I S Q C E I P P
 ggggggacattcacgtacaagttcaaggcgtacccggccggcaccttttgggtatcattct
 G G T F T Y K F K A Y P A G T F W Y H S
 cactagatggggccaatatcccgacggcctgcgcgccccaatgatcattcacgaccctaag
 H D M G Q Y P D G L R A P M I I H D P K
 gctgccgcgagcgggacactgataaggagtacgtacttacgggtctccgactgggtaccgt
 A A A E R D T D K E Y V L T V S D W Y R
 gaccagatgccgtcacttatccaccgctacttgacaacttccacttataatagtactatg
 D Q M P S L I H R Y L T T S T Y N S T M
 ccaaatccaaactcgagcttgatcaacgatcagcagctctacaacgctaaacatccgcccc
 P N P N S S L I N D Q Q S T T L N I R P
 gggcagaagatatatgttcggtattattaacatgtcagccctcgcaacgtactatctacag
 G Q K I Y V R I I N M S A L A T Y Y L Q
 ttcgatcaacaccacttgactgttattgccattgacggtgtcgacgtcgatccccagact
 F D Q H H L T V I A I D G V D V D P Q T
 tgggaggtcctggagattatccctggacagcgggtacgacgtcattatcaccggcctagag
 W E A L E I I P G Q R Y D V I I T G L E
 aaccccgaaaggaactatgcatttatcaacaagatggctactcttggttccagaacaat
 N P E R N Y A F I N K M A T L G F Q N N
 aacgtcccttagttacgattcgtccttggtcctgcccggagccggtgaacgtgggcagcttc
 N V L S Y D S S W P V P E P L N V G S F
 aatcttagaagcgaatttcaacctgaccccgcttgatgaggagctactactggagcccggtg
 N L R S D F N L T P L D E E L L L E P V
 gaccacaccttcacatggaggtcaataacgtgaacgtcgacggcgtaggctcccgcatc
 D H T F T M E V N N V N V D G V G S R I
 acgcaggggaccggacccttacatcgcccccggtacgcccaccctatacaccacctgagc
 T Q G P D P Y I A P R T P T L Y T T L S
 accggctctaattgctattaaccccgccatatacggccaggcgaacgcttacgtagtggaa
 T G S N A I N P A I Y G Q A N A Y V V E
 gccggcgatattgtccagcttgctcgtcaatagtaacgaacccgtcactaccaacacttcc
 A G D I V Q L V V N S N E P V T T N T S
 ggtcgtgggcaccctatgcacttgacggccacaccttccaagtggttggccaatatggc
 G R G H P M H L H G H T F Q V V G Q Y G
 agcccttgggacggcgatgcctctaaattccctgctgttccctatgaagcgggataccact
 S P W D G D A S K F P A V P M K R D T T
 gttctgtttactggcgggagcttggtgatccggttccgcgcggacaatcctggagtttgg
 V L F T G G S L V I R F R A D N P G V W
 atgttccattgccacaacgaatggcaccttgacgccggcatggctggaacgattatcgaa
 M F H C H N E W H L D A G M A G T I I E
 gcgcactcgagcttcaacaaagcggctgacgattccgccgcagcacctcgcgctcgtgc

A P L E L Q Q S G L T I P P Q H L A S C
 agggcggttaaacttaacgacccggggcaattgtgccggttaatactgcgaacctagaggat
 R A L N L T T R G N C A G N T A N L E D
 acggctgcatgcagagtctacgacactgagccttgggggtgcacttatcaagagagatgag
 T A A C R V Y D T E P W G A L I K R D E
 gaaacagcatattaa
 E T A Y -

SEQ ID NO : 66

CHIỀU DÀI : 544

LOẠI : PRT

SINH VẬT : *M. phaseolina*

MIRLSVLVLFMATTLAKTVTLNWDIGWVSAAPDGFTRPVGINGEWPPPVLEADVNDTIIVITRNLRLNETTSLH
 WHGMWHYNSTHMDGGARISQCEIPPGGTFTYKFKAYPAGTFWYHSHDMGQYPDGLRAPMI IHDPKAAAERD TDKE
 YVLTVSDWYRDQMPSLIHRYLTTSTYNSTMPNPNSSLINDQQSTTLNIRPGQKIYVRIINMSALATYYLQFDQHH
 LTVIAIDGVDVDPQTWEALEIIPGQRYDVIIITGLENPERNYAFINKMATLGFQNNNVLSYDSSWPVPEPLNVGSF
 NLRSDFNLTPLDEELLLEPVDHTFTMEVNNVNVVDGVGSRITQGPDPYIAPRPTLYTTTLSTGNSAINPAIYGQAN
 AYVVEAGDIVQLVVSNEPVTNTNTSGRGHPMHLHGHTFQVVGQYGSPPWDGDASKFPAPVPMKRDTTVLF TTGGS LVI
 RFRADNPGVWMFHCHNEWHL DAGMAGTII EAPLELQQSGLTIPPQHLASCRALNLTTRGNCAGNTANLEDTAACR
 VYDTEPWGALIKRDEETAY*

SEQ ID NO : 67

CHIỀU DÀI : 2412 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

GGGAGACGATCTGGAAGTGTTTGTTCAGAATGATTTGGCCGTTGAAACCACCATTTCATTGGCACGGTCGGTATCC
 ATGGTATACTTTTTACGAAGGCTAGTCAGCTGAAGGAGAAGTTCGGCAGGCATCCTTCAGCAAGGAACACCACAC
 ATGGACGGAGTTCCGGGAGTAACACAGGTTGGTTGATGACCTTCTATCGGGATGACTAGGGCCTAACTTTTGGCG
 CAGGAGCCGATCCCACCAGGAGGAACTTCACGTACCGCTTCTCGCTCAAAAACGAATACGGCTTCTACTGGTAT
 CACTCGCATTTCCGAGCATACTCGGACGACGCCATCCGCGGACCACTAGTCATCCATCCATCCTCACAACGCCCC
 CGGCCATACGAGACTCTCGCGAGGAACCAGACTGAGCTTACTGCTTTGCAAGAAGCTGAACCGGAAGCGGTGCC
 ATCCTCCTATCTGACTGGTACCATCGCGTTTCGGACGACATCTTCAACGAATACCTAACAACAGGCGCGTTCCCC
 AGTTGCGTGGACAGTCTCCTCGCCAATGGATACGGAAGGTTGCGGTGTCTACCGGAATACATTCTGGCAGCCGGA
 GCTGGGCTAGGCATGGAGCCTGCACCTGTCAATGCTACCGCGACTAGCATCGGCACAACGCCCATGTCTTCCATG
 GCGATGGGTACAAAGTACATGGAGACCAATACCAAGGAGTCGATGCGTATGGAGTCTATGGCAC TAGACACTACG
 TCCATGGAGCATCACATGCGACGCATGGATTCCATGT CAGCGGAGGATATGTCCATGAATACGATGACTGCTCAT
 TCTACGCCTGCGGCTGCGCCCGACTCAGGCATGCCGATGGGCTCCATGTCAATGGCTTCTCTACCCGTATCAGGC
 ATATCGGGCACGTCGGGCCCTGTCGGGCATGTCAAATATGGCGAGCGGTCCCCCTTGGTCCCTCGCGGCTGCAGTGCT
 CCCATGATGTT CAGACCGGGCTACAACATCAGCTCTCTCCCGCCAGAGACTTGCACAGACACGT CAGCGCCGCTG
 CTGACTGTGCGATGCAAACTATACACGAGGCTGGCTTGCAATTAAACCTAGTCAACTCCGGTT CAGTAACCAAGCTA
 AGCGTCTCACTAGACGCCCATTCGATGTTTGTCTATGCGGCCGACGGATTCTTTGTGAAACCCCAAGAAGTAGAA
 GTAAGTTGCACCTCCCCCTAACTTGTGTCGCGGAAGACTTACCGCGTCTAGGTATTACAAATCTCAATCGGGCAAC
 GTACTCAGTAATGATCAAACCTTAATCAACGGCCCGGAACTACACTCTACGTTTGTGCTCGTATCCCTACGGCG
 ATATGCAACAGGTTATCGAAGGCCAAGCGACCGTCTCATATAAGGTAAGCAGCAAGCTTTTCCCCCACTCGGTAT
 AGAACTAACTTCGTTAGGTTGATGCCGCGAGGACATTATGCCGGTGGACTTGACAAACGACCCCTACTGCAACGT
 GGATGCTTGTCAACGGTTTCGGCAAAGTCGAACGCTTCTGAATTGAAGACGGACATGCTCGCCCCGTTCGAGGCAA
 TTGCGCCGCCATCCCAAGCGGATATAACTTACGACTTCACGATCAGTCAGACAGAGATCGTAACTTGGGTGT TAA
 ACGGATATCCCTATTCGGAACCTTCGACGCCTATTATCTACGGCAATGCGTCGGAAGCGTGGAACGCAAATACTA
 CAATCCGCATTCCTTCCAACCTCGACCGTAGACATTATAATGCGGATTGCCAACGACTCAATGGATACGGCAAGTT
 GCAGTCCAGGCCCGCGTCAAAAACGCTGCATAAGCTAACC CGATAATGCCGTAGATGGGT CATCCTATGCACCTT
 CACGGCCACCGGTTCTTCGCCCTCGGCTCTGGATCGGGCTCCTTCCCATATCAGAACGCCGTCGACGCGCCTCCA
 TCCCTCATTAACCTCGAAAACCTCCGTACCGGGACACAACCGATTTACCACCTTCAGGCTGGGCAGTCATTTCGC
 TATGTAGCCAAACAATCCAGGCGCATGGATGTTTCACTGCCACATCCAGTGGCACCTCGTGAGCGGCATGGCATTG
 GTGTTTGTGCGAAGGAGAAGAGCAGCTGCCTGGTTTGGTGGGTGCGGCTGCGAACGGAACAAGCAATGCGAATAGC
 GCGTCACCGGCGCGTAGCACTCGAGAGCATGCGGCCCTTTGCCGTTGTGCTACCCCTAAGCACTGTTTCTTCGCG
 TATGGCTACTAATGCGCGAAGCGCCATGTAAGAGCGCGCTGAGGTAATAAAGGTGGCGACTGCAAAAAGCTCAT
 GTCCCGCAATGATGGGAGTTCTTTGGAGATTCTGAATCAAAGACGTCGATATTCTGGGTCCAGTCGGTGCAGAA
 CAGACCCCATGG

SEQ ID NO : 68

CHIỀU DÀI : 1743

LOẠI : ADN

SINH VẬT : *M. phaseolina*

TÊN/KHÓA DẤU HIỆU : CDS

VỊ TRÍ : (1).....(1743)

atggacggaggttccgggagtaaacacaggagccgatcccaccaggaggaaacttcacgtac
M D G V P G V T Q E P I P P G G N F T Y
cgcttctcgcgtcaaaaacgaatacggcttctactggtatcactcgcatttccgagcatac
R F S L K N E Y G F Y W Y H S H F R A Y
tcggacgacgccatccgcggaccactagtcacatccatccctcacaacgccccggcca
S D D A I R G P L V I H P S S Q R P R P
tacgagactctcgcgaggaaccagactgagcttactgctttgcaagaagctgaacgcgaa
Y E T L A R N Q T E L T A L Q E A E R E
gcggtgcctatcctcctatctgactggtaccatcgcgtttcggacgacatcttcaacgaa
A V P I L L S D W Y H R V S D D I F N E
tacctaacaacaggcgcggttccccagttgcggtggacagtctcctcgccaatggatacggg
Y L T T G A F P S C V D S L L A N G Y G
agggtgcggtgtctaccggaatacattctggcagccggagctgggctaggcatggagcct
R V R C L P E Y I L A A G A G L G M E P
gcacctgtcaatgctaccgcgactagcatcggcacaacgcccattgtcttccatggcgatg
A P V N A T A T S I G T T P M S S M A M
ggtacaaagtacatggagaccaataccaaggagtcgatgcgtatggagtctatggcacta
G T K Y M E T N T K E S M R M E S M A L
gacactacgtccatggagcatcacatgcgacgcatggattccatgtcagcggaggatag
D T T S M E H H M R R M D S M S A E D M
tccatgaatacgtgactgctcattctacgcctgcggctgcgcccgaactcaggcatgccg
S M N T M T A H S T P A A A P D S G M P
atgggctccatgtcaatggcttctctacccgtatcaggcatatcgggcacgtcgggcctg
M G S M S M A S L P V S G I S G T S G L
tcgggcatgtcaaatatggcgagcggtccccttggtcctcgcggctgcagtgtcccatg
S G M S N M A S G P L G P R G C S A P M
atgttcagaccgggctacaacatcagctctctcccgccagagacttgacagacacgtca
M F R P G Y N I S S L P P E T C T D T S
gcgcccgtgctgactgtcgatgcaaactatacacgaggctggcttgcatataaacctagtc
A P L L T V D A N Y T R G W L A L N L V
aactccggttcagtaaccaagctcagcgtctcactagacgccattcgtatgtttgtctat
N S G S V T K L S V S L D A H S M F V Y
gcggccgacggattctttgtgaaaccccaagaagtagaagattacaaatctcaatcggg
A A D G F F V K P Q E V E V L Q I S I G
caacggtactcagtaatgatcaaaccttaacacggcccgaaactacactctacggttt
Q R Y S V M I K L N Q R P G N Y T L R F
gcgtcgtatccctacggcgatatgcaacagggttatcgaaggccaagcgaccgtctcatat
A S Y P Y G D M Q Q V I E G Q A T V S Y
aagggttgatgccgcagaggacattatgccggtggacttgacaaacgacctactgcaacg
K V D A A E D I M P V D L T N D P T A T
tggtatgcttgcaacggttcggcaagtcgaacgcttctgaattgaagacggacatgctc
W M L V N G S A K S N A S E L K T D M L
gccccgttcgaggcaattgcgcccgcattcccaagcgatataacttacgacttcacgatc
A P F E A I A P P S Q A D I T Y D F T I
agtcagacagagatcatgggtcatcctatgcaccttcacggccaccggttcttcgccctc
S Q T E I M G H P M H L H G H R F F A L
ggctctggatcgggctccttcccatatcagaacgccgtcgacgcgcctccatccctcatt
G S G S G S F P Y Q N A V D A P P S L I
aacctcgaaaaccctccgtaccgggacacaaccgatttaccaccttcaggctgggcagtc
N L E N P P Y R D T T D L P P S G W A V
attcgtatgttagccaacaatccaggcgcatggatgtttcactgccacatccagtggcac
I R Y V A N N P G A W M F H C H I Q W H
ctcgtgagcggcatggcattgggtgtttgtcgaaggagaagagcagctgcctggtttggtg
L V S G M A L V F V E G E E Q L P G L V
ggtgcgggtgcgaacggaacaagcaatgcgaatagcgcggtcaccggcgcgtagcactcga
G A A A N G T S N A N S A S P A R S T R
gagcatgcggcctttgcggttggtgctaccctaagcactgttttcttcgcgtatggctac
E H A A F A V V A T L S T V F F A Y G Y
taa
-

SEQ ID NO : 69
CHIỀU DÀI : 580

LOẠI : PRT
 SINH VẬT : M. phaseolina
 MDGVPGVGTQEPPIPPGGNFETYRFSLKNEYGFYWHSHFRAYSDDAIRGPLVIHPSSQRPRPYETLARNQTELTAQ
 EAEREAVPILLSDWYHRVSDDI FNEYLTGTGAFPPSCVDSLLANGYGRVRCLPEYILAAGAGLGMEPAPVNATATSI
 GTTPMSSMAMGTKYMETNTKESMRMESMALDTTSMEEHMRMRDMSAEDMSMNTMTAHSTPAAAPDSGMPMGSMS
 MASLPVSGISGTSGLSGMSNMASGPLGPRGCSAPMMFRPGYNISSLPPETCTDTSAPLLTVDANYTRGWLALNLV
 NSGSVTKLSVSLDAHSMFVYAADGFFVKPQEVEVLQISIGQRYSVMIKLNQRPGNYTLRFASYPYGDMQQVIEGQ
 ATVSYKVDAAEDIMPVDLTNDPTATWMLVNGSAKSNASELKTDM LAPFEAIAPPSQADITYDFTISQTEIMGHPM
 HLHGHRFFALGSGSGSFYPYQNAVDAPPSLINLENPPYRDTTDLPPSGWAVIRYVANNPGAWMFHCHI QWHLVSGM
 ALVFVEGEEQLPGLVGAAANGTSNANSASPARSTREHAFAVVATLSTVFFAYGY*

SEQ ID NO : 70
 CHIỀU DÀI : 2242 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*

AGTCCACCCCTTTCCCTATTTCCTTTCTCCCTCACGTCACATTAGATACCCGGGCTCCTTTCAACTGCCAAATTTCCCTT
TTCCCCGGGACTCATTTCCCTCCCCGACTCTCTTACTCTTCTCTCAATTTCTACCACCCACCGCTCGCTCTCACA
ATGTCTTTTCGTTGCCAGGTCCTAGTAGAAGCGGCTCGCCAGCAGAGCCGCTTTACTAATGGATCATGGACCGGT
GCCTGGGAGACTCATGGCTCCACCAAGGTTCCATCTCACGCAACTACGCAACGGTACGTACCAGGTTTCAGTTG
TTTAAATCCCAAATTTGACTGATCTTTGTTGCAGGCGGCGGCTCCTTGGTACCTTGGACGCTCCAAACTTGCCT
GGCTGTATTGGTTCTCCACCGTGGGGCAACCATGACTCTCATGACCCTTTTCGGCATGCCTGATACTGGCATTACT
CGGGAGTATGACTTTACTCTCACTTATCAAGACATCGCTCCCAGTGGCGTCACTAAGAGGGGGCGTGGTTGTCAAT
GGTCAATATCCCGGGCCACGATCGAGGCCAACTGGTGCGTACATCACTACATCACGCACATGACATCAGGCTAA
CACTTCGTTTCAGGGGCGACTGGATCCAGGTCACGGTCCACAACGGTCTAGGCGAGGACGAAGGTGAGGGTACCGC
GATGCACTGGCACGGCTTTTTGTCAGAAGGAGAGCCAGTGGATGGATGGTGTCCCCGGTGTTCACAATGTCTAT
TCCTCCCGGAGAGAGCTTACCCTACCGCTTCCGCGCTGAGCAGTATGGCACTTCTTGGTACCACAGCCACTACAG
CGCGCAGTACTCGGGCGGTGCTGCTGGTCTCTCATCGTATACGGTCCGGACAGCCAGAGTTACGACGCTTGACCT
CGGTCCCGTTCATGGGTATCTGACTGGTACCACCTCGCAATACTACGATATTGTCAAGCAGACCATGCAGGCCGACC
CACGGGACCAACCCCTCCGCCGCCGCTCAGTCCGACAACAACCTGATCCAGGGCTTTGGCGAATTCGACTGCTC
GCTGACGACGAAGCCCTGCATTCTGACGCCGGTGTAGCCAAGTTCAAGTTCACTTCCGGCAAGAAGCATCGCT
CAGGCTCATCAACTCCGGCTCCGAGGCGATGCAGAGGTTCTCTATCGACGGCCACACCATGAAGATCATCGCACA
CGACTTCGTCCCGATCGAGCCCTACGAGGTCACGGCCCTGACCTTGGCGTCGGACAGCGGGCCGACGTGGTTGT
CGAAGCCACGGGCAAGCCCTCGGACGCGTACTGGATGCGCTCAGAGATCGGGCTCAACAAGTGCAACGTCTTCAA
TGCGAACGCCCTCCGAGGCCCTCGCCGTATCTCTACGAGGATGCAGACCTGCTGGCCGTGCCAGCTCCTCCGC
CCAGCCCAGTGCAGGAGCTGACATCGTGCACCAACGACGACATCTCCGTTCGGCTGCCGCTCCAGCACATCCTGCC
CGACCCCAACCCGTCCGTACGACCGAGCTGCACATCGAGAACAAAGTACAACGGCACGCACTGGCTCTGGCACTT
CGGCGGCCCCAGCTACCGCGCCGACTTCAACGACGCCCTCCTCTACCAGCTCCAACAGGGGCACGCCCGACTTCGG
CCCGCAGTCCAACGTGCACGACTTCGGCGCCAACAAATCCGTCCGCTTCGTTCGTCTACAACCACGTCCCGGCTCA
GCATCCCATGCACTTGCATGGCCACAACCTTCTGGGTCTTCGCCGACGGCGTTCGGCACCTGGGACGGCGCCATCGC
CAACCCGCAGAACCCCCAGCGCCGCGACGTGCACGTCTTCAGCCCCGCCAGGGCAGCACCCCGTTCGTACATGGT
CGTACAGATTGAGCTGGACAACCCGGGCCCTGTGGCCGTTCCACTGCCACATCGCGTGGCACGTCTCGGCGGGGT
GTATTTGAACGTGCTGGAGCGCCCGGACGACATCAAGGCGCTGCAGATTCGGGCCGCGGTGTGGTGATCAGTGCAG
GGCTGGGCGGACTACACGCGCAAGAAGTACGGTGATCAGATCGACTCGGGGTGAAGAAGAGTAGAAGGCGTG
ATAGTTTTGCTATTGTTCTTTCTCTCTCTCTCTTTTTTTTTTTTTTTTTTACACCTTCCCTTCTTGATACAGCATG
TTTTTGGGCGTTATTTACGGTCTTCTGCGCATTTCCAAGGCGTTATGTCTTTCCCTTGGCGATAAG

SEQ ID NO : 71
CHIỀU DÀI : 1836
LOẠI : ADN
SINH VẬT : *M. phaseolina*
TÊN/KHÓA DẤU HIỆU : CDS
VỊ TRÍ : (1).....(1836)

atgtcttttcggttgccagggtccgtagtaagcggcctcgccagcagagccgcctttactaat
M S F V A R S V V S G L A S R A A F T N
ggatcatggaccggtgcctgggagactcatggctccaccaaggttccatctcaccgcaa
G S W T G A W E T H G S T Q G S I S P Q
ctacgcaacggcgggcgggcctcttggtaccttgagcgtccaaacttgcttggtgtatt
L R N G G G P L G T L D A P N L P G C I
ggttctccaccgtggggcaaccatgactctcatgaccctttcggcatgcttgatactggc
G S P P W G N H D S H D P F G M P D T G
attactcgggagtagtactttactctcacttatcaagacatcgctcccgatggcgctcact
I T R E Y D F T L T Y Q D I A P D G V T
aagaggggcgctcgttgtcaatggtcaatatccggggccacgatcgaggccaactggggc
K R G V V V N G O Y P G P T I E A N W G

gactggatccaggtcacggtccacaacggtctagcgaggacgaaggtgaggggtaccgcg
 D W I Q V T V H N G L G E D E G E G T A
 atgcaactggcagggctttttgcagaaggagagccagtggatggatgggtgtccccgggtgtt
 M H W H G F L Q K E S Q W M D G V P G V
 caacaatgtcctattcctccccggagagagcttcacctaccgcttccgcgctgagcagtat
 Q Q C P I P P G E S F T Y R F R A E Q Y
 ggcacttcttgggtaccacagccactacagcgcgcagtagctcgggcggtgctgctggctcct
 G T S W Y H S H Y S A Q Y S G G A A G P
 ctcatcgtatacgggtccggacagccagagttacgacggttgacctcgggtcccgctcatggta
 L I V Y G P D S Q S Y D V D L G P V M V
 tctgactgggtaccactcgcaatactacgatattgtcaagcagaccatgcaggccgacccc
 S D W Y H S Q Y Y D I V K Q T M Q A D P
 acggggcaccacccctccgcccgcgcctcagtcggacaacaacctgatccaggggtttggc
 T G T T P P P P P Q S D N N L I Q G F G
 gaattcgactgctcgtgacgacgaagccctgcattcctgacgccggtgtagccaagttc
 E F D C S L T T K P C I P D A G V A K F
 aagttcacttccggcaagaagcatcgccctcaggtcatcaactccgggtccgaggcgatg
 K F T S G K K H R L R L I N S G S E A M
 cagaggttctctatcgacggccacaccatgaagatcatcgacacgacttcgtcccgcac
 Q R F S I D G H T M K I I A H D F V P I
 gagccctacgaggtcacggccctgacccttggcgtcggacagcgggcccgcagtggtgtc
 E P Y E V T A L T L G V G Q R A D V V V
 gaagccacgggcaagccctcggacgcgtactggatgcgctcagagatcggggtcaacaag
 E A T G K P S D A Y W M R S E I G L N K
 tgcaacgtcttcaatgcgaacgcctccgagggccctcggcgtcatcctctacgaggatga
 C N V F N A N A S E A L A V I L Y E D A
 gacctgctggcgtgccagctcctccgcccagcccgatgcggagctgacatcgtgcacc
 D L L A V P S S S A Q P D A E L T S C T
 aacgacgacatctccgtcggcctgccgctccagcacatcctgcccgaccccaacccgtcc
 N D D I S V G L P L Q H I L P D P N P S
 gtcacgaccgagctgcacatcgagaacaagtacaacggcacgcactggctctggcacttc
 V T T E L H I E N K Y N G T H W L W H F
 ggcggccccagctaccgcgcgcgacttcaacgacgccctcctctaccagctccaacagggc
 G G P S Y R A D F N D A L L Y Q L Q Q G
 acgcccgacttcggcccgcagtcacaacgtgcacgacttcggcgccaacaaatccgtccgc
 T P D F G P Q S N V H D F G A N K S V R
 ttcgtcgtctacaaccacgtcccgggtcagcatcccatgcacttgcatggccacaacttc
 F V V Y N H V P A Q H P M H L H G H N F
 tgggtcctcgcgcgacggcgtcggcacctgggacggcgccatcgccaacccgcagaacccc
 W V L A D G V G T W D G A I A N P Q N P
 cagcgccgcgacgtgcacgtccttcagcccgcagggcagcaccctcgtacatggctc
 Q R A R D V H V L Q P A Q G S T P S Y M V
 gtacagattgagctggacaacccgggcctgtggccgttcactgccacatcgcggtggac
 V Q I E L D N P G L W P F H C H I A W H
 gtctcggcggggttgatttgaacgtgctggagcgcccggacgacatcaaggcgctgcag
 V S A G L Y L N V L E R P D D I K A L Q
 attccggccgcggttggtgatcagtcagggcggtggcgggactacacggcaaagaatacg
 I P A A V G D Q C R A W A D Y T A K N T
 gtggatcagatcgactcgggggtgaagaaggagtag
 V D Q I D S G L K K E -

SEQ ID NO : 72

CHIỀU DÀI : 611

LOẠI : PRT

SINH VẬT : M. phaseolina

MSFVARSVVSGLASRAAFTNGSWTGAWETHGSTQGSISPQLRNGGGPLGTLTLDAPNLPGCIGSPPWGNHDSHPFG
 MPDTGITREYDFTLTYQDIAPDGVTKRGVVVNGQYPGPTIEANWGDWIQVTVHNGLGEDEGEGETAMHWHGFLQKE
 SQWMDGVPGVQQCPPIPPGESFTYRFRAEQYGTSWYHSHYSAQYSGGAAGPLIVYGPDSQSYDVLDGPMVSDWYH
 SQYYDIVKQTMQADPTGTTTPPPPPQSDNNLIQGFGEFDCSLTTKPCIPDAGVAKFKFTSGKKHRLRLINSGSEAM
 QRFSIDGHTMKIIAHDFVPIEPYEVLTALTLGVGQRADVVEATGKPSDAYWMRSEI GLNKNVFNANASEALAVI
 LYEDADLLAVPSSAQPDALTSCTNDDISVGLPLQHILPDPNPSVTTELHIENKYNHGLHWHFVGGPSYRADFN
 DALLYQLQQGTPDFGPQSNVHDFGANKSVRFVVYNHVP AQHPMLHGHNFVWLADGVGTWDGAIANPQNPRRDV
 HVLQPAQGSTPSYMVVQIELDNPGLPWFHCHIAWHVSAGLYLNVLERPD DIKALQIPAAVGDQCRAWADYTAKNT
 VDQIDSGLKKE*

SEQ ID NO : 73
 CHIỀU DÀI : 1048 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 GGACCGTTGCGGCGTTACTGATCTTCGGGTGAACGTCTCTCTTCGTACCAGGAAGCCGCCAACGCCGAAGAAAGG
 GAGAAATACGTGTGGCTCAACTTTATCCACCCTGGCGCCCATCATGAGCTGCGGATCTCCGTTGACGAGCACGAC
 ATGTGGATTGCGGCCGAGATGGAGACTTCGTAAAACCaAAGAAAGTCCAAGTAAGTTTCCACGTCTCGGTTTCG
 AGCACGCGCTGACAATCTCAAGGCCATCAATGTCAACATGGGCGAGAGGATAAGCGTCTTATTCCTCTTACCCA
 AAGTCCTGGAGAGTACGCCATCCGCATGGTCTCTCTCGCGGAAGAGCAGCTCATCTGGGGCTTGGGAATTCTCCG
 GTACCCCGGTGTTCAAGAGAGACGAGATGAGAATGGCATAATGATCCTGCCAGAAAGCCAGCCACACATTGATGT
 TCAGGATAAACCTGTTGACTGACGGAATTGTGATGGACGAAATGACCGACCTGATCCCGTTTCCCGCGCGCCGTCC
 TCCAGCCAAGGCCGACCACACCTTTTCGCTTCGCCATCAAGCGGCCGAATCCAAGCACGTGGATTTTGGCATCGGA
 GCCGCATCAAGGATTTCAGACAACAGCTTCCGCCGGTGCTTTGGAACAAGGATTCCCGTGGCCCTACGACGTTTCGG
 CGGAATGAAGAACGGCTCGGTTGTGGACATCATTTATGAAAACGGAGCATTTGGGATGCATCCGTTTACCAGTG
 GATGAACGAATCGCATCACAGTATGGCAACTGCTGACATTTTCATGCACAACCACAAAGCGTTCATCATCGGCATG
 GGGGATGGGTTCTTCCGGTGGCCAGACGTTGCCTCGGCCCTCAAGGAGGCCCTGAAAACTTTAACATGGTGAAC
 CCTCCTCTCCGTGACGGAGCACGGCTGGCAAAGGGAGAAGGATCGTGGACTGTAATCCGCTACCAGATCACCTCT
 CCCGCAATGTCCATGTTACACTGTAAGCAGCTCTCTGCAAAGTGCCCTGACGGTGTAAGCACTGACGACCTCTC

SEQ ID NO : 74
 CHIỀU DÀI : 702
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 TÊN/KHÓA DẤU HIỆU : CDS
 VỊ TRÍ : (1).....(702)

atgtggattgcgccgcagatggagacttcgtaaaaccaaagaaagtccaagccatcaat
 M W I A A A D G D F V K P K K V Q A I N
 gtcaacatggcgagagagataagcgcttattccttacccttaccctgagagtag
 V N M G E R I S V L I P L T Q S P G E Y
 gccatccgcatgggtctctctcgcggaagagcagctcatctggggcttgggaattctccgg
 A I R M V S L A E E Q L I W G L G I L R
 taccctgggtgttcaagagagacgagatgagaatggcataatgatcctgccagaaagccag
 Y P G V Q E R R D E N G I M I L P E S Q
 ccacacattgatgttcaggataacctgttgactgacggaattgtgatggacgaaatgacc
 P H I D V Q D N L L T D G I V M D E M T
 gacctgatcccggtttcccgcgcgccgtcctccagccaaggccgaccacacctttcgcttc
 D L I P F P A R R P P A K A D H T F R F
 gccatcaagcgccgaatccaagcacgtggattttggcatcggagccgcatcaaggattc
 A I K R P N P S T W I L A S E P H Q G F
 agacaacagcttccgcccgtgctttggaacaaggattcccgtggccctacgacgttcggc
 R Q Q L P P V L W N K D S R G P T T F G
 ggaatgaagaacggctcggttggtgacatcatttatgaaaacggagcatttgggatgcat
 G M K N G S V V D I I Y E N G A F G M H
 ccgtttaccagtggtgaacgaatcgcatcacagtatggcaactgctgacattttcatgc
 P F F H Q W M N E S H H S M A T A D I S C
 acaaccacaaagcgttcatcatcgcatgggggatgggttcttccggtggccagacgttg
 T T T K R S S S A W G M G S S G G Q T L
 cctcgccctcaaggaggccctgaaaactttaacatggtga
 P R P S R R P L K T L T W -

SEQ ID NO : 75
 CHIỀU DÀI : 233
 LOẠI : PRT
 SINH VẬT : *M. phaseolina*

MWIAAADGDFVKPKKVQAINVNMGERISVLIPLTQSPGEYAIRMVSLAEEQLIWGLGILRYPGVQERRDENGIMI
 LPESQPHIDVQDNLITDGIVMDEMTDLIPFPARRPPAKADHTFRFAIKRPNPSTWILASEPHQGFRRQQLPVLWN
 KDSRGPTTFGGMKNGSVVDIIYENGAFGMHPFHQWMNESHSMTADISCTTTKRSSSAWGMGSSGGQTLPRPSR
 RPLKTLTW*

SEQ ID NO : 76
 CHIỀU DÀI : 2269 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)
 LOẠI : ADN

SINH VẬT : *M. phaseolina*

GCTGCGAGGCATTCGAGAGGATTACCGAGTATTGGAGAGGCCGCGTGAAGGCAGCAGAGCATTAATAACGTGCTT
TGGAGCACGCGAATTACTTTCAAGTTGAACAAGATATCACCTGTCTTGCACGACGTTGCAAGCTCCGCCGAGCC
ATGAAGTGTGCAACTCTTTGGAGCTATCTTGCCCTCTGTCTGACCGTTCGGCGCTTCTGCGAGGTTATCCATTCA
CATAGACTTCAAACATTACCTAAATGCTGACAACTCATTTACAGAAGCTTGACGCGGTCTTGCCCTCAGCACAAG
ACCAACGGAGGGTCCAATTGGGGAACTTTGGACTGTCCAAAGTTACCCGATTTTTTTGACCTCGAACCCGCTGCCC
GGCGGCTTCCCGTGGGGTGACAGAAGCGGCCATAAGCAACGATCCCTACACTGATGTGCCGAACACCGGGGTGACG
AGGTACTACGATTTTCAAGCTCGCACGTGGCTATCTCGCTCCAGATGGCTACAACAAAAGTGGCATCTTCATCAAC
GGCGAGTTCCCTGGGCGTCCCATTTGAGGCCAATTGGGGCGACATGATTGAAGTACGAGTGACACAACAACATCGTC
GGCCCTGAAGAAGGCACCTGCGTTCCACTGGCATGGCATTAATCTCAGAAGGGCACGCAATGGTTTGACGGCGTTCCC
GGCGTGTCCCAATGCCCCATTGCCCCCTGGATCCTCTTTTACCTACCGCTTCCGTGCCGACGTCTACGGCACTTCT
TGGTGGCACTCGCACTTCTCTGCGCAATATACCGTGGTGCTTTTGGGCCCCCTCATTATCTACGGCCCCAAGCAT
GTTCTTACGATGTTGATGTGCGCCCTGTAATTTCTCGGTGATTACTACCACCGTGACTACTTTGATGTTCTGGAG
GATGCTGCCAGCAACACCACTGACTTCAACATCTACGTCCCTTGGTCCGACAACAATCTGATCAATGGCAAGAAC
AACTATAATTGCTCCATGGTAGCTGGAACCTCTACGAGCTTCGCCAACGCTACAAGCTCCTCGAACGCCACGTGC
TTCTCCAACGCTGGCCTTGCCCAGTTCCGTTTCGAGCCAGGCAAAGTGACCCGCTGCGTCTGATGAACGTGGGC
GCAGCAGCACTGCTGCACTTCTCAATCGACGGGCACAAAATGCAAGTCATCGCCACGACTTCGAACCTGTTGTC
CCGTACGAGGCAGACGTATCACGCTGGGCGCCGCCAACGCACCGACATCCTCGTCACTGCAGATGCCAACCCC
AACGAGACATACTGGATCCGCTCCACCATCTCGCTCAACTGCTCTGTCTCGCACAAACACCAACGCGTGGCCGTT
CTCTCTTACGAAGGCAATGACCACATAGAAGAGCCACGCAGCCGCATTAGCGCCGCCGCGGCCGCTGCTGACGAG
AAGAGCTTCTCTGCAAGAACGACGACCTGTCCCAGACGGTGCCCTTCTTCCCCAAGCCCGTCGCCGAGCCCGAT
GTGACCGAGACGATCGAAGTCGACCTCTTACCAATGCGACCGGCCACCATGTGTGGATCATGAACAACCGCACG
CAGCGCACGAACCTACACGAGCCCGTCTTGCTGCTCGCCAACCAGGGCAACAGCACCTTCCCGGACGAGTGGAAC
GTTTACGACTTTGGGCGCAACAAGACCATCCGCATCGTCTCAACACCGTCTACCAGTCCGCCACCCGATGCAC
TTGCACGGCCACTCTTTTCGTAAGACCTTTCATCCCTTCGCCACGCATGTGTGCCCACTCTGACCTCTTCTTCGCA
GCAAGTCTTCGCCGAAGGCCCCGGCGCCTGGGACGGCACGACCATCACCAACCCATCCAATCCCGCCCCGCCGCA
CACGCACATGCAGCCCGGTACGGGCACCTGGTAATCCAGTTTCGAGGGCCGACAACCCGGGCGGTGGAGCTACCA
CTGCCACATCGCCTGGCACGCCAGCATGGGCTACAACATCGAGATCCTCGAGCGCGGCGACGAGCTGGCGGCCGC
CGGCGCTATTCTTATGGTCATGCAGCAGACGTGTGATGATTGGAAGGAGTGGAGCGGCAGGAATGTGGTCAATCA
GATCGACGCGGGCATTTAGATTATCGGCGACTTCTATGCCCGATGGTATAATGTTTTTACCGAGCGTGGGATGAT
GTGGGTTTGGGGTCGGAATTTGTAGATATGGGCGGGGATTGCCTGCACAAAATAGACCGATGCACATGTTTATA
AAGCAAAAACCTCTTCCCAT

SEQ ID NO : 77
CHIỀU DÀI : 1854
LOẠI : ADN
SINH VẬT : *M. phaseolina*
TÊN/KHÓA DẤU HIỆU : CDS
VỊ TRÍ : (1).....(1854)

atgaagtgtgcaactcttttgagctatcttgccctctgtcctgaccgtcggcgcttctgcg
M K C A T L W S Y L A S V L T V G A S A
agaagcttgacgcggtccttgccctcagcacaagaccaacggagggtccaattggggaact
R S L T R S L P Q H K T N G G S N W G T
ttggactgtccaaagttacccgattttttgacctcgaaccgctgcccggcggttcccg
L D C P K L P D F L T S N P L P G G F P
tggggtgacagacggcctaagcaacgatccctacactgatgtgccgaacaccggggtg
W G D R S G L S N D P Y T D V P N T G V
acgaggtactacgatttcagcgtcgcacgtggctatctcgtccagatggctacaacaaa
T R Y Y D F S V A R G Y L A P D G Y N K
agtggcatcttcatcaacggcgagttccctgggcctgccattgaggccaattggggcgac
S G I F I N G E F P G P A I E A N W G D
atgattgaagtacgagtgcacaacaacatcgtcgccctgaagaaggcactgcgttccac
M I E V R V H N N I V G P E E G T A F H
tggcatggcattactcagaagggcacgcaatggtttgacggcggttcccgcggtgtcccaa
W H G I T Q K G T Q W F D G V P G V S Q
tgccccattgcccctggatcctctttttacctaccgcttccgtgccgacgtctacggcact
C P I A P G S S F T Y R F R A D V Y G T
tcttggtggcactcgcacttctctgcgcaatataaccgctggtgcttttgggccccctcatt
S W W H S H F S A Q Y T A G A F G P L I
atctacggccccaagcatgttcccttacgatgttgatgtcgccctgtaattctcgggtgat
I Y G P K H V P Y D V D V G P V I L G D
tactaccacggtgactactttgatgttctgaggatgctgccagcaacaccactgacttc
Y Y H R D Y F D V L E D A A S N T T D F
aacatctacgtccccttggtccgacaacaatctgatcaatggcaagaacaactataattgc

N I Y V P W S D N N L I N G K N N Y N C
 tccatggtagctggaaactctacgagcttcgccaacgctacaagctcctcgaacgccacg
 S M V A G N S T S F A N A T S S S N A T
 tgcttctccaacgctggccttgcccagttccggttcgagccaggcaaagtgcaccgcctg
 C F S N A G L A Q F R F E P G K V H R L
 cgtctgatgaacgtgggcgcagcagcactgctgcacttctcaatcgacgggcacaaaatg
 R L M N V G A A A L L H F S I D G H K M
 caagtcatcgccacgacttcgaacctgtgtcccgtagcaggcagacgtcatcacgctg
 Q V I A H D F E P V V P Y E A D V I T L
 ggccgcccccaacgcaccgacatcctcgtcactgcagatgccaaacccaacgagacatac
 G A A Q R T D I L V T A D A N P N E T Y
 tggatccgctccaccatctcgtcgaactgctcgtcgtcgcacaacaccaacgcgctggcc
 W I R S T I S L N C S V S H N T N A L A
 gttctctcctacgaaggcaatgaccacatagaagagccacgcagccgcattagcgccgcc
 V L S Y E G N D H I E E P R S R I S A A
 gcggccgctgctgacgagaagagcttcctctgcaagaacgacgacctgtcccagacggtg
 A A A A D E K S F L C K N D D L S Q T V
 cccttcttccccaagcccgtcgccgagcccgatgtgaccgagacgatcgaagtgcacctc
 P F F P K P V A E P D V T E T I E V D L
 ttcaccaatgcgaccggccaccatgtgtggatcatgaacaaccgcacgcagcgcacgaac
 F T N A T G H H V W I M N N R T Q R T N
 tacaacgagcccgtcttctgctcgtcgcgaaccagggcaacagcaccttcccggacgagtg
 Y N E P V L L L A N Q G N S T F P D E W
 aacgtttacgactttggcgcaacaagaccatccgcatcgtcctcaacaccgtctaccag
 N V Y D F G R N K T I R I V L N T V Y Q
 tccgcccacccgacttgacgtgacggccactcttccaagtcctcgccgaaggccccggc
 S A H P M H L H G H S F Q V L A E G P G
 gcctgggacggcagcaccatcaccaacccatccaatcccgcgccgacgacgcacatg
 A W D G T T I T N P S N P A R R D T H M
 cagcgccggtacgggcacctggtaatccagttcgaggccgacaacccgggcgcgctggagc
 Q R R Y G H L V I Q F E A D N P G A W S
 taccactgccacatcgctggcagccagcatgggctacaacatcgagatcctcgagcgc
 Y H C H I A W H A S M G Y N I E I L E R
 ggcgacgagctggcgccgcccggcgctattcctatggatgcagcagacgtgtgatgat
 G D E L A A A G A I P M V M Q Q T C D D
 tggaaggagtggagcggcaggaatgtggtcaatcagatcgacgcgggcatttag
 W K E W S G R N V V N Q I D A G I -

SEQ ID NO : 78

CHIỀU DÀI : 617

LOẠI : PRT

SINH VẬT : *M. phaseolina*

MKCATLWSYLASVLTVGASARSLTRSLPQHKTNNGGSNWGTLDCPKLPDFLTSNPLPGGFPPWGDRSGLSNDPYTDV
 PNTGVTRYDFSVARGYLAPDGYNKSGIFINGEFPGPAIEANWGDMEIVRVHNNIVGPEEGTAFHWHGITQKGTQ
 WFDGVPVGSQCPIAPGSSFTYRFRADVYGTSSWWHSHFSAQYTAGAFGLIYGPKHVPYDVDVGPVILGDYYHRD
 YFDVLEDAASNTTDFNIYVPWSDNNLINGKNNYNCSMVAGNSTSFANATSSSNATCFSNAGLAQFRFEPGKVHRL
 RLMNVGAAALLHFSIDGHKMQVIAHDFEPVVPYEADVITLGAQRTDILVTADANPNETYWIRSTISLNCVSHN
 TNALAVLSYEGNDHIEEPRSRISAAAAAADEKSFLCKNDDLSQTVPPFPKPAEPDVTETIEVDLFTNATGHHVW
 IMNNRTQRTNYNEPVLNLANQGNSTFPDEWNVYDFGRNKTIRIVLNTVYQSAHPMHLHGHSFQVLAEGPGAWDGT
 TITNPSNPARRDTHMQRRYGHLLVIQFEADNPGAWSYHCHIAWHASMGYNIEILERGDELAAGAI PMVMQQTCD
 WKEWSGRNVVNQIDAGI*

SEQ ID NO : 79

CHIỀU DÀI : 2373 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

GACAATCCCTCCGCTAAATCCGACATCCGGTACTCTCGCCTCCCGGATCCCGAGCGGACACACAGCTTCGCCTGT
 ATACCACTTTACCGGAGTCGCCTTCCCACCCCTATAATTCTTTTCTGTGAGTCTACTCCGCGCGCGGTATC
 ATGATAGATACCAAGAGCACAGGGTCCGGCGCAGACGGCGGTGTTACGCCGCCTTAAGGCAAGACGAGTCAGAG
 CTATACGAGCAGAAAGAACATACGCAGTGTACACAGCCCTCCGCGCCGCGTACCTCGGTGGAAATGGCAGAGAA
 GAAGCCCCTGGGATTGATCGACGAAAATGTCGTGCTCGAACAAGACGCTGTGGACCCACCCGGCGGTACTCT
 GTCTTCTTGAATTTGCCATCCTCGGTCTGGTACTCATAATCGCTCTACTCGGCGCTCTCGCTTGGTCCAGAGGC
 TCTCATCATCATACCGACCCGGTATCTGGCAGTTCAACCATCGTCCAAGGGTCGCCGCGGAAAGTATGTCCTG
 GACCTGCCTGGGATTTCTGCTGCACCGCCGCGAGGTCCGCAATACCACTGGACGATCAGGGACATCGAGCTTCGC

CCAGACGGCGTGGAAGCGGCCGCTGATCACCATCAATAACGAGTTCCCGGGGCGGACCATCGAATGCAACCAGGGG
 GATACCGTGCGGGTTGAGGTTTCATAACGAAGCCGTCAATTCAACTTCCTTTTACTGGCACGGCATTACCAGAAC
 GGAACCACGTACATGGATGGCACGGTCGGCATCAGCCAGTGCTTATCACGTCTGGGTCTAGCATGACATATGAG
 TTCAAAGTCGACAGAGAATCCGGCACCTATTGGTATCATGCGCACATGGCTATGCAGGGCTCAGATGGTCTTTT
 GGTCTCTGATCGTCCATTGAAAAATGAGCGGAAGCTGCAGCAACTCGAATATGCCCTCCGATCAAGTCATCATG
 GTCCACGATTACTACCACGATCTGACCAGCGCGCTGATACCACACTACTTAGCGCCGGATAACGAGAACACAGAG
 CCTGTCCCTGACGGAGGTCTCATCAACGGAATGAATAAGAGAAAATTGCGAGCTCCTCCGTGGTTCGAGACTGTGAT
 GCCACTGATGCACAGCTTGCCACATTGCGCCTCGAACCGAACAAGAACCCGCTCTTCGAATTATCAATACTGGA
 GCATTTGCTGAATTCCAGGTGAAGATTGACGAGCACACGTTTCGCTGTGACGGAGGTGGATGGGACCGAAGTTGCT
 CCCGCTACTACCACAGGCTCAACATCAATCCCGGACAGCGCTACAGCATCGTAATAAACACCAATGTCACGGAT
 CGTGACTCCTTCTGGCTGAGAGCTAAGATGATTGAGGCCTGTTTCGCTGAGGAGAACCCAAATCTGGATCCCGAA
 GTGCGCGCCATTATCCAATATACTCGCAAGGACGAGGATACCCAGCCCAAGGAACCTTCGAGCAGAGACTGGGAC
 GACATCGTGGACATGCAGTGTCTCGACATGAACGTGACAGAGCTCCAGCCCGTAGAGAAGGCAACACCTCCACCT
 GCAGACACCACACTATACCTCCGCTCCAACCTTCGAGATCGGCAACTGGCGTCTGAGCCGCGGCTTCTTCAACAGC
 TCGTCTGGCGTCCAACACTCTCATCCCCAAGCCTGCACCGCATGATCGACGGCCTCCACAGCCAAAACGCCAGC
 TTCTTCCCCGACCGAGCGTACCCCTTCCAGATCAACTCGGCCGGCTTCGACACTGGGCCCGAGCTGGTCTACCAG
 ACCAGCGGCATCCGCACCATCGACATCCTCGTTTCCAACCTTCGACGACGGCAACCACCCGCTCCACCTGCACGGC
 TACAAGTACTTCGTCTTTCGCTCGGGCCACGGCTACCCGCCCGCCGACCTCTACGCGCATCTCGACATCTCGAAC
 CCGCTGCGCCGCGACACCGCCTCGATCGAGGCGTTTCGGCTGGATCCTACTGCGTCTCGTCGCCGACAACCCGGGC
 GTCTGGGCTTCCACTGCCACATCGGCTGGCACACCGAGGCCGGCATGCTGATGCAGTTCGCCACGCGCGTTCGAC
 GTGCTCGCATCCAGCCAAATCCCGGATACGCACCTCGCGCTCTGCGCGGCCGACGGGCTCGACCGCGCGCGTTCG
 CCGCCAGACTCGACGTGGTTTGGGGATTTTGGGGATCTAGATCCTTGAAAAGTTCTGGAGAGGAGTGGGAGTTTC
 ACCTGGCATGCAAGCGTCTCCGTTTCTTCAGGCCAGGCTGACCCACTTCCCGCGCACGCCGATCTCGTACTGCAG
 CCCGTCGCTTCCATGTTTAGGTCCACCTGCGGTACTCATCTTCGAC

SEQ ID NO : 80
 CHIỀU DÀI : 2073
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 TÊN/KHÓA DẤU HIỆU : CDS
 VỊ TRÍ : (1).....(2073)

atgatagataccaagagcacagggtccggcgcagacggcggtcgttacgccgccttaagg
 M I D T K S T G S G A D G G R Y A A L R
 caagacgagtcagagctatacgagcagaaagaacatacgagtcgtacacagccctccggc
 Q D E S E L Y E Q K E H T Q C T Q P S G
 gccgcgtacctcgggtgaaatggcagagaagaagccctgggattgatcgacgaaaatgtc
 A A Y L G G N G R E E A L G L I D E N V
 gtcgctcgaacaaaaagacgctgtggacccaccggcggtactctgtcttccttgaattt
 V A R T K R R C G P T R R Y S V F L E F
 gccatectcgggtctgtgactcataatcgctctactcggcgctctcgttgggtccagaggc
 A I L G L V L I I A L L G A L A W S R G
 tctcatcatcaccgaccgggtatctggcagttcacaaccatcgccaagggtcgccgc
 S H H T D P V S G S Q P S S K G R G
 ggaaagtatgtcctggaccctgcctgggatttcgctgcaccgccgcaggtccgcaaatac
 G K Y V L D P A W D F A A P P Q V R K Y
 cactggacgatcagggacatcgagcttcgcccagacggcgtgaagcggccgctgatcacc
 H W T I R D I E L R P D G V K R P L I T
 atcaataacgagttcccggggcccagccatcgaatgcaaccagggggataccgtgcgggtt
 I N N E F P G P T I E C N Q G D T V R V
 gaggttcataacgaagccgtcaattcaacttcctttcactggcacggcatttaccagaac
 E V H N E A V N S T S F H W H G I Y Q N
 ggaaccacgtacatggatggcacgggtcggcatcagccagtgtcctatcacgtctgggtct
 G T T Y M D G T V G I S Q C P I T S G S
 agcatgacatatgagttcaaagtcgacagagaatccggcacctattggtatcatgcgcac
 S M T Y E F K V D R E S G T Y W Y H A H
 atggctatgcagggtcagatggtcttttggctcctctgatcgtccattcgaaaaatgag
 M A M Q G S D G L F G P L I V H S K N E
 cggaagctgcagcaactcgaatatgcctccgatcaagtcacatcatggtccacgattactac
 R K L Q Q L E Y A S D Q V I M V H D Y Y
 cacgatctgaccagcgcgtgataccacactacttagcgccggataacgagaacacagag
 H D L T S A L I P H Y L A P D N E N T E
 cctgtccctgacggaggtctcatcaacggaatgaataagagaaattgagagctcctccgt
 P V P D G G L I N G M N K R N C E L L R
 ggtcgagactgtgatgccactgatgcacagcttgccacattcggcctcgaaccgaacaag

G R D C D A T D A Q L A T F G L E P N K
aaccaccgtcttcgaattatcaatactggagcatttgctgaattccagggtgaagattgac
N H R L R I I N T G A F A E F Q V K I D
gagcacacgttcgctgtgacggaggtggatgggaccgaagttgctcccgccactaccac
E H T F A V T E V D G T E V A P A Y Y H
aggctcaacatcaatcccggacagcgctacagcatcgtaataaacaccaatgtcacggat
R L N I N P G Q R Y S I V I N T N V T D
cgtgactccttctggctgagagctaagatgattgaggcctgtttcgtgaggagaaccca
R D S F W L R A K M I E A C F A E E N P
aatctggatccgaagtgcgcgccattatccaatatactcgcaaggacgaggatacccg
N L D P E V R A I I Q Y T R K D E D T Q
cccaaggaaccttcgagcagagactgggacgacatcgtagacatgcagtgctcgcacatg
P K E P S S R D W D D I V D M Q C L D M
aacgtgacagagctccagcccgtagagaaggcaacacctccacctgcagacaccacacta
N V T E L Q P V E K A T P P P A D T T L
tacctccgctccaacttcgagatcggaactggcgtctgagccgcggttcttcaacagc
Y L R S N F E I G N W R L S R G F F N S
tcgtcctggcgtccaactctcatccccaaagcctgcaccgcatgatcgacggcctccac
S S W R P T L S S P S L H R M I D G L H
agccaaaacgccagcttcctccccgaccgagcgtaccccttcagatcaactcggccggc
S Q N A S F L P D R A Y P F Q I N S A G
ttcgacactgggcccagctgggtctaccagaccagcggcatccgcaccatcgacatcctc
F D T G P E L V Y Q T S G I R T I D I L
gtttccaacttcgacgacggcaaccacccgctccacctgcacggctacaagtacttcgtc
V S N F D D G N H P L H L H G Y K Y F V
cttgcgctgggcccacggctacccgcccgcgacctctacgcgcatctcgacatctcgaac
L A S G H G Y P P A D L Y A H L D I S N
ccgctgcgccgcgacaccgcctcgatcgaggcgttcggctggatcctactgcgtctcgtc
P L R R D T A S I E A F G W I L L R L V
gccgacaacccgggctgtgggccttcactgccacatcggtggcacaccgaggccggc
A D N P G V W A F H C H I G W H T E A G
atgctgatgcagttcgccacgcgcgtcgacgtgctcgcatccagccaaatcccggtacg
M L M Q F A T R V D V L A S S Q I P D T
cacctcgcgctctgcgcgccgacgggctcgaccgcgcgctcgccgcccagactcgacg
H L A L C A A D G L D R G A S P P D S T
tggtttggggattttggggatctagatccttga
W F G D F G D L D P -

SEQ ID NO : 81
CHIỀU DÀI : 690
LOẠI : PRT
SINH VẬT : M. phaseolina

MIDTKSTGSGADGGRYAALRQDESELYEQKEHTQCTQPSGAAYLGGNGREEALGLIDENNVARTKRRCGPTRRYS
VFLEFAILGLVLIIALLGALAWSRGSHHHTDPVSGSSQPSSKGRRGKYVLDPAWDFAAPPQVRKYHWTIRDIELR
PDGVKRPLITINNEFPGPITIECNQGDTVRVEVHNEAVNSTSFHWHGIYQNGTTYMDGTGVISQCPITSGSSMTYE
FKVDRESGTYWYHAHMAMQGS DGLFGLPLIVHSKNERKLQLEYASDQVIMVHDYYHDLTSALIPHYLAPDNENTE
PVPDGGILINGMKNRNCCELLRGRDCDATDAQLATFGLPKNHRLRIINTGAFAEFQVKIDEHTFAVTEVDGTEVA
PAYYHRLNINPGQRYISIVINTNVTD RDSFWLRAKMIEACFAEENPNLDPEVRAIIQYTRKDEDTQPKEPSSRDWD
DIVDMQCLDMNVTTELQPVEKATPPPADTTLYLRNFEIGNWRLSRGFFNSSWRPTLSSPSLHRMIDGLHSQNAS
FLPDRAYPFQINSAGFDTPGPELVYQTS GIRTIDILVSNFDDGNHPLHLHGKYFVLASGHGYPPADLYAHLDISN
PLRRDTASIEAFGWILLRLVADNPGVWAFHCHIGWHTEAGMLMQFATRVDVLASSQIPDTHLALCAADGLDRGAS
PPDSTWFGDFGDLDP*

SEQ ID NO : 82
CHIỀU DÀI : 2372 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)
LOẠI : ADN
SINH VẬT : M. phaseolina

TATATGTTCCCTCCGTCCTCAACGGAAATCCGGCAGTGC CGCAAGACGAGCCCCCTTTTCCTTCTGCTTCGCCTGCT
TTTCGTAGTCTCGTCCTCCGCCCTAGACGTCGTCTTCTCAGACAATTTGCTCAAATCCCCCTCTCACCTCGCCGCC
ATGCGCTTCGCGAGCCTCGCCGTGGCATGGCTCACAACGTGCGTTGTCCAATCACTTGCCCAACGAGCCCATCCCCG
TTCGACAGGATCCTCTGGGGTGAGAAATGGCCCCGCCGGTAACCATATGGTGAAGCGTCAGGCCAGCTCCAGCTCC
CTTGCCCTCATCCACCACAAGGGCCCCGATTCCGCCGTGCACAAACGGCCCTTTGACGAGGAGCTGCTGGTCCAAT
GGCTTTTCCATAGCCACTGACTTTTGACGCCAAGTGGCCCAATACCGGAAAGACCGTCCATGTAAGCGCGCCTGCA
GATATGCATCCGTCCTGCTGCTGACGTGCAGCCCATAGTATGACCTGACCATCAACAACGCCACCTGCAGCCCCG

ACGGCGGGCCCCAGCCGCCCCGTGCCTGATGTTCAACAACAAGATCCCTGGGCGCTACGCTTTACGCCAATTGGGGCG
ACATGATCTCCGTTACCATCACCAACAAGATGCCCAACAACGGCACCAGCGTGCACTGGCATGGTCTGCGTCAGT
ACAATACAAACACCCAGGATGGCGTCAACGGAATCACGGAATGCCCTCTGGCTCCCGGCGATTCCAAGACCTACC
TATTCCAGGCTACACAGTTCGGCACGACCTGGTTCCACAGCCATTTCTCTGCGCAGTATGGCGACGGTGCCGTCG
GCCAGCTCATCATCAATGGCCCCGCTCGGCCGAATTACGATTTTCGATCTCGGCACCTACACTATGACCGACTGGT
ACTACAGCACTGCGTTCCAGGTGCAAGATCAATTTGATGCCGCCCTTCAGAGGAAGGCCCCGGCCCCGCCAGGTG
ACACCATCTCTGGTGAACGGTACGATGAAGTCTCCGGATGGCTCTGCTGGTAGCTACAGCCAAGTCAAGGGCCTTG
TCAAGGGAAAGAAGTACCGCCTGCGTCTCATCAACACCTCGGTGGACAACAACATCCGCGTCTCGCTGGACAACC
ACCCATTACCGTTCGTCACTTCCGACTTCGTCCCCGAGCAAGCCTTGGACTACTGACTGGCTTCTCTTAGCCATCG
GCCAGCGCTACGATGTCTCTTACAGGCCAATCAACCCGCGGCCAATTATTGGTTCCGTGCAGAAGTTGCCACCG
CATGTGCTAGCGCCAACAAGTACCGCGGGCCGCGGTATATTACGCTACGTTGGTGCCGATGGCAGCGCTCCCCCAG
AGACCGCCGCTGACTGTTCCAGGTGGCTGTACCGAGCCTCTGCCTGCGCCTTTCGTCGCAAACCAGGTTCCAAGCC
AAGTCTTCTCGACCAAGTGAAGACCCCTTAGCGTCGATGTTTATGCGGCAAACGTCTCGACCAACCAGAAGAACA
TTGTCTTCTGGGGCATCAACATGACTGCCATTGACATTGATTGGGAGAAGCCGACGCTGGAATACGTCAGGACAA
AAAATACCAGTACCCCCACGTTTACAACCTTGATCGAGTTGCCACAGAGAACATTGTAAGCGGAGTTAAGCTTC
AACCCCGAGAAATGTTATTAAAGATCGCAACAGTGGACCTACTGGATCATCCAAGAACTCCCGGCACTCCCCCA
ATTCCGCATCCAATTCACCTGCACGGTAAGCTGCAGGTCTCAACCTTGCAACCCCGTAGAGATACTGATGGTCGC
CTCTGTCCAGGCCACGACTTCTACATCCTCGGAACCGGCTCTGGTGCCCTTCGACCGCAGCACCTCGCCGCTCCTCC
CTCAATTTCAATAACCCCAACCCGCGCGACGTGCGACTGGTTCCCGGCGGCGGTTGGTTGGCCATTGCCTTCCCG
ACCGACAACCCAGGCGCTTGGCTCATGCACTGCCATATCGTAAGTTTACCTATCTCTCCGACTACGATCAATTGT
GTATATTTCATGCGCACGGCGCTGACACGTGGATTTCCCTCGCCACAGGCTTGGCACATTAGCGAAGGTCCTCGGAG
TTCAGTTCTCTCGAGGGCAAGGATAAGATCAATCTTCCCGACGCTGCGTGGGAGACGACCTGCTCGAATTGGGACA
AGTACTGGGACACCACTATCTACCCCAAGCAGGATTCGGTCTCTGAGGGGTGGCCGGTTTTTTTTTCTTCTTTAC
GGGCTGAGTTTAAGGGGGGTTTTGATGTTTTGCTGAAAAAAGACCCGTAGTGATTCTTGTATTATACTTTTCGTT
TTATTCATATTGATAGGAAGCAACCTGATCGCACTTCGTAAACCGCA

SEQ ID NO : 83
CHIỀU DÀI : 1824
LOẠI : ADN
SINH VẬT : *M. phaseolina*
TÊN/KHÓA DẤU HIỆU : CDS
VỊ TRÍ : (1).....(1824)

atgcgcttcgcgagcctcgccgtggcatggctcacaacgtgcgttgccaatcacttgcc
M R F A S L A V A W L T T C V V Q S L A
aacgagcccatcccgttcgacaggatcctctgggggtgagaatggccccgcccgttaacct
N E P I P F D R I L W G E N G P A G N H
atggtgaagcgtcaggccagctccagctcccctgcctcatccaccacaagggccccggat
M V K R Q A S S S S P A S S T T R A P D
tccgcctgcacaaacggccctttgacgaggagctgctggtccaatggcttttccatagcc
S A C T N G P L T R S C W S N G F S I A
actgactttgacgccaagtggcccaataaccggaagaccgtccattatgacctgaccatc
T D F D A K W P N T G K T V H Y D L T I
aacaacgcccactgcagccccgacggcgggcccgcccgctgcctgatgttcaacaac
N N A T G C S P D G G P S R P C L M F N
aagatcccctgggcttacgcttttacgcaattggggcgacatgatctccgttaccatcacc
K I P G P T L Y A N W G D M I S V T I T
aacaagatgcccacaacggcaccagcgtgcaactggcatggtctgcgtcagtacaataca
N K M P N N G T S V H W H G L R Q Y N T
aacacccaggatggcgtcaacggaatcacggaatgccctctggctcccggcgattccaag
N T Q D G V N G I T E C P L A P G D S K
acctacctattccaggctacacagttcggcacgacctggttcacagccatttctctgcg
T Y L F Q A T Q F G T T W F H S H F S A
cagtatggcgacggtgccgtcggccagctcatcatcaatggccccgcctcggcgaattac
Q Y G D G A V G Q L I I N G P A S A N Y
gatttcgatctcggcacctacactatgaccgactggactacagcactgcgttccagggtc
D F D L G T Y T M T D W Y Y S T A F Q V
gaagatcaatttgatgccgcccttcagaggaaggcccccgccccgccagggtgacaccatc
E D Q F D A A L Q R K A P G P P G D T I
ctggtgaacggtacgatgaagtctccggatggctctgctggttagctacagccaagtcaag
L V N G T M K S P D G S A G S Y S Q V K
ggccttgtcaagggaagaagtaccgcctgcgtctcatcaacacctcgggtggacaacaac
G L V K K L R L I N T S V D N
atccgcgtctcgtggaacacccattccacgctcgtcacttccgacttctgctcccgagc
I R V S L D N H P F T V V T S D F V P S

aagccttggactactgactggcttctcttagccatcggccagcgctacgatgtcatcttc
 K P W T T D W L L L A I G Q R Y D V I F
 acggccaatcaacccgcggccaattattggttccgtgcagaagttgccaccgcatgtgct
 T A N Q P A A N Y W F R A E V A T A C A
 agcgccaacaagtaccgcggcgccggtatattcagctacgttgggtgccgatggcagcgct
 S A N K Y R G R G I F S Y V G A D G S A
 cccccagagaccgccgtgactgttccaggtggctgtaccgagcctctgcctgcgcctttc
 P P E T A V T V P G G C T E P L P A P F
 gtcgcaaaccaggttccaagccaagcttctcctcgaccaagtgaagacccttagcgctgat
 V A N Q V P S Q V F L D Q V K T L S V D
 gtttatgctggcaaacgtctcgcaccaaccagaagaacattgtcttctggggcatcaacatg
 V Y A A N V S T N Q K N I V F W G I N M
 actgccattgacattgattgggagaagccgacgctggaatacgtcaggacaaaaaataacc
 T A I D I D W E K P T L E Y V R T K N T
 agctacccccacgtttacaacttgatcgagttgccacagagaacatttggacctactgg
 S Y P H V Y N L I E L P T E N I W T Y W
 atcatccaagaaactcccggcactcccccaattccgcatccaattcacttgcacggccac
 I I Q E T P G T P P I P H P I H L H G H
 gacttctacatcctcggaaccggctctggtgccttcgaccgcagcacctcgccgtcctcc
 D F Y I L G T G S G A F D R S T S P S S
 ctcaatttcaataacccccaccggcgcgacgtcgcactggttcccggcgcggttggtg
 L N F N N P T R R D V A L V P G G G W L
 gccattgccttcccggaccgacaacccaggcgcttggtcatgcactgccatatcgcttg
 A I A F P T D N P G A W L M H C H I A W
 cacattagcgaaggtctcggagttcagttcctcgagggcaaggataagatcaatcttccc
 H I S E G L G V Q F L E G K D K I N L P
 gacgtgcgtgggagacgacgtgctgaattgggacaagtactgggacaccactatctac
 D A A W E T T C S N W D K Y W D T T I Y
 cccaagcaggattccggtctctga
 P K Q D S G L -

SEQ ID NO : 84

CHIỀU DÀI : 607

LOẠI : PRT

SINH VẬT : *M. phaseolina*

MRFASLAVAWLTTTCVVQSLANEPIPFDRILWGENGPAAGNHMVKRQASSSSPASSTTRAPDSACTNGPLTRSCWSN
 GFSIATDFDAKWPNITGKTVHYDLTINNATCSPDGGPSRCLMFNNKIPGPTLYANWGDMSVITITNKMPPNNGTSV
 HWHGLRQYNTNTQDGVNGITECLAPGDSKTYLFQATQFGTTWFHSHFSAQYGDGAVGQLIINGPASANYDFDLG
 TYTMTDWYYSTAFQVEDQFDAALQRKAPGPPGDTILVNGTMKSPDGSAGSYSQVKGKLYRLRLINTSVDNN
 IRVSLDNHPFTVVTSDFFVPSKPWTTDWLLAIGQRYDVIETANQPAANYWFRAEVATACASANKYRGRGIFSYVG
 ADGSAPPETAFTVPGGCTEPLPAPFVANQVPSQVFLDQVKTLSDVYAANVSTNQKNIVFWGINMTAIDIDWEKP
 TLEYVRTKNTSYPHVYNLIELPTENIWTYWI IQETPGTPPIPHPIHLHGHDIFYILGTGSGAFDRSTSPSSLNFNN
 PTRRDVALVPGGWLAIAPFTDNPGAWLMHCHIAWHISEGLGVQFLEKDKINLPDAAWETTC SNWDKYWDTTIY
 PKQDSGL*

SEQ ID NO : 85

CHIỀU DÀI : 3008 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

GAGAAGAGCTTTTCGAAGCCATGTCAGCAGGATTAGCTCCTGAGACAGCCCTGCCCAAAGTCCCGTATTTCTAT
 GGCCGCAGTTCTCTCTCTCCTCGCCGAACCAAAGCCTCCTCCAGTACTTTCTTCTCAAACCTGCTCATACAAGTGCA
 ATGGGGTTCTTCAATGCCTGCTGGGACTTGCTGCTCCACCTCCCTTTCTTTGCTCCGCGGGGCGGGCGGCTTCGAG
 CGGGATCGCTCGCAGCTTCCGATCTCCTCCGGTCCGAGTGGCGGCGTAGTCATCCACCCGGAGAATGCTTCCCTT
 GGCTTACCTGCTCTTACCCGAGCATGGAAGGGTGGGAAAGCTGCAACTACCCGACGACAGGAGCTGCTGGCTT
 AAAGATGGGAGGGCGAGTCAGCCTTACTTCTCTCAATACGATATCCACACAGACTGTAAGTCTCGATATGGTTG
 CCCGCATTGCATGCTTTACGCTTTACCCAGGCCTTCGTGCTAAATTTATGAAGACGAGACCGTCTGGCCTCAGGGT
 GTAACCAGAGAAGTAAGTGTGGAAAATCAAAGCCCTGGGGCGCGGGGAACCCGGCCTCTGAGCCTTCATCTATGG
 CAGCCCGTACACATCGAGAGACGAAGTATTTATCGCGAAGAGAATGCCCTGATCTCTTTAAGCACATACATCTG
 CATGAATATTTCAATCCCCCTTGTCTATCAGGATGATCAAGGGGCCATGCCGTTACCAAAATGTCTGGGCGAGAT
 CCCTCTTTTTACGCTGCAGCAGATTTGCTTTACTCGGTGTTTACTGAACATTTGCAGTACTGGATCAACCTCAAG
 GACCAAGTGGTAAGTGCTGCCATTTCCCTTGAGTCCGCTGTAGCTGCTCATACGCCAGCTTTTCCCCGACGGTTA
 TTCCAAGCCTTATGGCAAAGTAATCAACGACACCTACCCAGGTCCTTTAATCGAGGCTTGCTGGGGCGATGAAGT
 AGTTGTTTATGTGACAAACTACCTGCAGACGAATGGCACTACGTATTCACACGTCTATATCAGTTGTGCCGCGGT

ATGAGCTGATGATGGGGCAGTATTTCACTGGGCACGGTGTGAGACAGCAGTTTCAAGCAACGAAATGGACGGAGTCAAC
 GGTTTGTCTTGCAAGATCCCACTATTTTACCTCCTCGGCTCACGTCATGTAGGTATTACACAGTGTCTATCGCTT
 ATGGTGACACCTTACCTACCGCTTTTCGTGTTACTCAATATGGAACCTACGTGTAAGTGTTCGACTCCTGATCT
 GGTCTCAGCCATTGACACGGAAGTAGGGTACCATTACACTACTCGCTCCAATACCCAGATGGCGTTGCTGGCCC
 TCTCGTATTTACAGGCCCTACAGCAGCCGATTGGGATGAAGAGTGGGAGACCCCGCTGATGATCACTGATTGGGT
 TCACGATTTCGGCTTTTGGGGTCTTCTCCCAGGAACCTCTCGCTCGGATCCTGCCAACCCGGAACGTCACACCACC
 CGTGGGGGACAGCATTCTGCTTAACGGACACGGACATTACAACCTGCAGCCTTTCTCAGGACCAAAACCGCTGCGC
 GCCGGGATATGGATCTTACTACACGCAAGATTCCAAAAGGGTAAAAGGTACTTGATCAGGCTAATTAACTCCTC
 AGCTGGAGCAGCATTTATCTTCTCCATAGACGGACACAAAATGAAGGTCATCTCCACGGATCTCGTTCCCATTTGA
 GCCGTACGAGACAAATGCAGTGCTCCTCAACATAGGTATATAGCGTTTGGTGCTCGTCTAGCATGTCTCCAGA
 CACTAACTTCCCGCAGGCCAACGCTACAACATCATCGTCGAAGCCAAACGCCGAACCCGGCGACTACTGGATCCGT
 ACCGAGATACCCGGCGGGCCAGGCGGCTGTGGCAGCGTGCACGACCGGGCCGGTAACGTGACGGGCATCCTGCGC
 TACGACGGACGACGAGTACCGCGCTACCGACCTCATCGAAGAATGACTACCCGTCGGACTGCCACGATGAGCCAGCG
 GAGCTGCTGCACCAATCCTGCCGTGGACGGTGGATCCGCACCCGCAGAACGACGTACACAACAACACGTACGAG
 GTCGGCATTTCAGACGCCCAGTTCCACAAGGCCTTCCGCTGGGACCTGACCGACACGCCCATGTGGCTCGACTTC
 TCAAACCCGACCATCCTCAACCTGTACAACACCACCTGGAACCCGGAGTACGCCGTATCGACTGTGAGTTGTCC
 TTCTTACCTCCCGTTCTCCCCAAAAAACAACCTAGCAACGAGTGGATGAAGACTGACAGGTACGAAAACAGACAAC
 ATGACCGCGGGCTTCGTCTACCTCGTCATCACGGCCAACCTGACACGGCTGGGCGACAACAAGCGCGAGATCCCCG
 CCGGCCACCCCATCCACCTGCACGGCCACGACTTCGCCGTCTCTCGCCAGTCCAACCTCGACCTATGACGAGCGCT
 CCGACCCGCTCAATTTACCCCTCGCCAACCCGCCGCGCCGCGATGTCGTCTTCTCGCCGAGCAACGGCTACGTCG
 CGCTGGCGTTCAAGCCGGACAATCCCGGCATATGGTTGGTGCAATTGCCATATCGCTTGGCATGCCAGTTCTGGTG
 AGTGTTTTTCTTTTTTCTTTTTGGTAACTTTGTTTTCTGTGGCTTTTTTAAAGTGTGTGAATGCATGCTGATGA
 TTTTATACGATGCACAGGCTGGCACTGCAGATTCTGGAGAGGCAGCCAGATATCCTGGATTTCGATTGGCACGC
 TCGAGGCGACGAATAAGACGTGTGCTGGATGGGATACGTACGAGAGGGCGCATCCGATCGAGCAAGACGACAGCG
 GTATCTGATTAGACGTACCTCTGCGTCAGATCCGACGAGTTGGGAGTGAATGATGCCCTCTGGAAGGATTGTGAA
 GGATGGGGTGAGCTCTTCTCCTGTGCATTTTCTCTCATGTTGAGGAGCTCTGTTCTGAGCTGGAGTACGCTCAT
 TTATTTCC

SEQ ID NO : 86
 CHIỀU DÀI : 1932
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 TÊN/KHÓA DẤU HIỆU : CDS
 VỊ TRÍ : (1).....(1932)

atgggggttcttcaatgcctgctgggacttgctgctccacctccctttctttgctccgcg
 M G F F N A C W D L L L H L P F F A P R
 ggcgggcggttcgagcgggatcgctcgcagcttccgatctcctccgggtccgagtgggc
 G G G F E R D R S Q L P I S S G P S G G
 gtagtcatccaccggagaatgcttcccctggcttcaectgctcttaccgagcatggaa
 V V I H P E N A S P G F T C S Y P S M E
 ggggtgggaaagctgcaactcaccgacgacaggagctgctggcttaaagatgggagggcg
 G W E S C N S P D D R S C W L K D G R A
 agtcagccttactctctcaatacgcatacgcacagactacgagaccgtctggcctcag
 S Q P Y F S Q Y D I H T D Y E T V W P Q
 ggtgtaaccagagaataactggatcaacctcaaggaccaagtgcttttcccgacggttat
 G V T R E Y W I N L K D Q V L F P D G Y
 tccaagccttatggcaaagtaataacgacacctaccaggtcctttaatcgaggcttgc
 S K P Y G K V I N D T Y P G P L I E A C
 tggggcgatgaagtagttgttcatgtgacaaactacctgcagacgaatggcactactatt
 W G D E V V V H V T N Y L Q T N G T T I
 cactggcacggtgtgagacagcagttcagcaacgaaatggacggagtcaacggtattaca
 H W H G V R Q Q F S N E M D G V N G I T
 cagtgtcctatcgcttatggtgacaccttcacctaccgctttcgtgttactcaatatgga
 Q C P I A Y G D T F T Y R F R V T Q Y G
 actacgtggtaccattcacactactcgctccaataccagatggcggttgctggccctctc
 T T W Y H S H Y S L Q Y P D G V A G P L
 gtatttcacggccctacagcagccgattgggatgaagagtgggagaccccgctgatgatc
 V F H G P T A A D W D E E W E T P L M I
 actgattgggttcacgattcggttttgggtcttctcccaggaactcctcgcgctcgat
 T D W V H D S A F G V F S Q E L L A S D
 cctgccaaaccggaacgtcacaccacccggtgggggacagcattctgcttaacggacagga
 P A N R N V T P P V G D S I L L N G H G
 cattacaactgcagcctttctcaggacaaaaaccgctgcgcgcccgggatatggatcttac

H Y N C S L S Q D Q N R C A P G Y G S Y
 tacacgcaaagattccaaaagggtaaaagggtacttgatcaggctaattaactcctcagct
 Y T Q R F Q K G K R Y L I R L I N S S A
 ggagcagcatttatcttctccatagacggacacaaaatgaagggtcatctccacggatctc
 G A A F I F S I D G H K M K V I S T D L
 gttcccatgtgagccgtacgagacaaatgcagtgctcctcaacataggccaacgctacaac
 V P I E P Y E T N A V L L N I G Q R Y N
 atcatcgctgaagccaacgccaacccggcgactactggatccgtaccgagatacccggc
 I I V E A N A E P G D Y W I R T E I P G
 gggccaggcggtgtggcagcgtgcacgaccggcggtaacgtgacgggcatcctgcgc
 G P G G C G S V H D R A G N V T G I L R
 tacgacggacgcagtagcgtaccgacctcatcgaagaatgactaccgctcggactgc
 Y D G R S T A L P T S S K N D Y P S D C
 cacgatgagccagcggagctgctgcacccaatcctgccgtggacgggtggatccgcacccg
 H D E P A E L L H P I L P W T V D P H P
 cagaacgacgtacacaacaacacgtacgaggtcggcatttcagacgcccagttccacaag
 Q N D V H N N T Y E V G I S D A Q F H K
 gccttccgctgggacctgaccgacacgcccattgtggctcgacttctcaaaccggaccatc
 A F R W D L T D T P M W L D F S N P T I
 ctcaacctgtacaacaccacctggaacccggagtagccgtcatcgactacaactatgac
 L N L Y N T T W N P E Y A V I D Y N Y D
 cgcggtctcgtctacctcgtcatcacggccaacctgacacggctgggcgacaacaagcgc
 R G F V Y L V I T A N L T R L G D N K R
 gagatccccgcggccaccccatccacctgcacggccacgacttcgccgtcctcgcccag
 E I P A G H P I H L H G H D F A V L A Q
 tccaaactgcacctgacgagcgtccgacccgctcaatttcaccctcgccaacccggcg
 S N S T Y D E R S D P L N F T L A N P P
 cgccgcgatgtcgtcttctcctgccgagcaacggctacgtcgcgctggcggttcaagccggac
 R R D V V F L P S N G Y V A L A F K P D
 aatccccgcataatggttggtgcattgccatcgtctggcatgccagttctggcctggca
 N P G I W L V H C H I A W H A S S G L A
 ctgcagattctggagaggcagccagatatcctggattcgattggcacgctcgaggcgacg
 L Q I L E R Q P D I L D S I G T L E A T
 aataagacgtgtgctggatgggatacgtacgagaggcgcatccgatcgagcaagacgac
 N K T C A G W D T Y E R A H P I E Q D D
 agcggtatctga
 S G I -

SEQ ID NO : 87

CHIỀU DÀI : 643

LOẠI : PRT

SINH VẬT : *M. phaseolina*

MGFFNACWDLHLPLFFAPRGGGFERDRSQLPISSGSPGGVVIHPENASPGFTCSYPSMEGWESCNSPDDRSCWL
 KDGRASQPYFSQYDIHTDYETVWPQGVTR EYWINLKDQVLPDGYSKPYGKVINDTYPGLIEACWGDEVVVHVT
 NYLQTNGTTHWHGVRQQFSNEMDGVNGITQCPIAYGDTFTYRFRVTQYGT TWYHSHYSLQYPDGVAGPLVFHGP
 TAADWDEEWETPLMITDWVHDSAFGVFSQELLASDPANRNVTPPVGDSILLNGHGHYNC SLSQDQNR CAPGYGSY
 YTQRFRQKGKRYLIRLINSSAGAAFI F SIDGHMKMVISTDLVPIEPYETNAVLLNIGQRYNIIVEANAEPGDYWIR
 TEIPGGPGGCGSVHDRAGNVTGILRYDGRSTALPTSSKNDYPSDCHDEPAELLHPILPWTVDPHQNDVHNNTYE
 VGISDAQFHKA FRWDLTDT PMWLD FSNPTILNLYNTTWNPEYAVIDYNYDRGFVYL VITANLTRLGDNKREIPAG
 HPIHLHGHDFAVLAQSNSTYDERSDPLNFTLANPPRRDVVFLPSNGYVALAFKPDNPGIWL VHCHIAWHASSGLA
 LQILERQPDILDSIGTLEATNKTCAGWD TYERAHP IEQDDSGI*

SEQ ID NO : 88

CHIỀU DÀI : 2404 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

AAATGCCATGGACAGCTCGCCCGGGATCGATGTCCTTCTCCCCGAGCCCCCGTCGCCCCGTCCAGGAGAGATAAA
 CCTTCTCATAGGGCCTCATTCCTACTCTCTCCTTCCCTTTTtCGTTACCCTCCCCGACCGCGGTCCACCTTCGAC
 ATGTGGTGGTTGGCATTGTTCTCATTGCTTGTGGCCGCCACGGCCTGGGCCAAGGAGCCCTACCTCAAGGTCCAC
 GATGATACCTTTATTCCAGATGCCGTGTTGCGGGTAACAGAGGAGAGTGCCTCCATTGGTTGCATCGAGGAGACC
 TCCGCCGTGGTCAACGGCACAGTTCCGGGGCCGATTTTGGAAATTCCAATCCGGCAGCGTGGTATGGGTTTCGTGTT
 TACAATGATATGGCAGACAAAAACCTCACCATGGTAAGCCCCGACGATATGCCTGCGGTGAATCAGACTAGTCAA
 TTAACCTCTCGGTCCACAGCATTGGCATGGCCTCACCATGGCAGCCGCTCCATTGCGCCGACGGTTCCGTTGCAGC
 CAGCCAGTGGGCGATCGAGCCGTTCAAATCTTCGACTACGAGCTCAACCTATTCGACATCAAGCCTGGGACGTA

TTTCTACCACTCCCATGTCTGGATTCCAGGCAATTACTGCCACGGGTCCCCTGCTCATTACCAAAAAAGCCGGGAGA
GGAGCCGCCGTATGAGTACGAGGAGGAACGCATCGTCCTCTTTTCCGACCTTTTACAACACGACGGACCACGACAT
TGAAACCCGGACTGGTAGCCAGCCCTTTCAAATGGAGCGGTGAGGTCTGGAGACGTCTCTGTCACCGGATATGGCAT
ATCGCAGTACCCGGCCACGAACGACGCCGAGAGTTGCAACCTTGCACAGATTCCCGTGgAAGCTGGGAAGACGTA
CCGTTTTCGTTTTCATTGGTGCCACTGCCCTTTCTTTCTCTCCGTTGGCTTTGAGAAGCACAAATCTTACCATCAT
CGAGGCGGATGGCCACTACACGGAACCTGCAGAAATCAGCTTCCTTCAAATCGGCGGCGGCCAACGGTATTCTGC
TCTCTAAAGACATGGACTTGCAGGAGCTCGCCGCCAAGACTGCTGGTTCGAAACCAGTTCTACATCCAAATCGA
GACTCGTGACAGGCCCAAGAACCTCACCACCTACGCCATTCTGGACTACAGCGACAGCTGCGCCACCAATAGCAT
CAGCGTCGAGGCCACCATCGGCAAGGACTCAACATCGCCCCCTTCTACGCCCCCTCTTCGCTCAGCAAACCCGG
CAACAACAACAAAAACAACAACAACACCACGCTCTCCCGCAAGACGCCGCCCAAGACCCCCACCGCTGCACCTCCC
GCCACCGTGCAAGGCTGGCTCGACCATGACCTCCACCCGCTCGAAACCTACACGGACTTCCCAGCGGCGGACGA
GGTCACGCGCACCGTCTACATGGACATCTTCCAGCTAGGGCAGGACGGCTACGTCAAGTGGGCGCAGAACAACT
GTCTGGTACGAGCACACGCCCAAGGTTCCCTACCTCGTCGCGCTCTACACCAACAGCACGCGAGTACCTGCCCCG
CTACGACTATGCCGTCGCATCGGGGACGGGCCACGACGACCGCGTCTGGTGCCTGGCCCGCCAAGATGGGCGAGGT
GCTCGAGATCATCGTCGTCAACACGGGCAGCTACAGCGCGGCATGGACGTGCATCCCATGCACCTGCACGGCGC
GCACCCGTTCTATTTGGGGAGTGGGAAGGGCACGTACAACAGAGAGGAGAACGAGAAGAAGTTGGCGGGGAGGGT
GCCGGTCACGAGGGATTTCGATGATGCTGTATCGGTATGGCGAGAAGGAGGAGCCGCACAAGGATAATAGCTGGAT
TGCGCTGAGGATAAGGGTGACGCAGCCGGGGGTGTGGATGTTCCATTGCCATACGCTGGCGCATATGATTATGGG
TGAGTTTGACCCCTTTTtCCCTTTTTTTTTTTTTTTTTtCTtATTGCGCATGTGGTTTGATGAAAACCATGGTGCT
GATGTGATTTGGAACAGGTATGCAAACGGTCTGGGTGTTTGGCGATTTCGAAGGATATCCTCACGCTGCCGCTGCC
GATGGTGGAAGGTTACTTGGTGCCTGGCGGAGATGTGTTCCGTGATGATGATCATGATCCGGTGGTGGTGCATT
CTTTGACCTGGACGACGACGACGATGATGATGATGCCAATAAGACCGACGGGAACGGCGGGAAGAATGGCCG
GTAGATGGGGGGAGGAGAGTTACTTGATTTCGAATGGAAGAGGTTACTGCGGAGGGGATTTGTAATGCATGCTAC
AGTTTTTATGTTATATGCCAGGTATAACCAGAAGGAAAGCCGGATGATTAACTGTGGAACAAGAGAAAAGAAGCA
TTTA

SEQ ID NO : 89
CHIỀU DÀI : 1950
LOẠI : ADN
SINH VẬT : *M. phaseolina*
TÊN/KHÓA DẤU HIỆU : CDS
VỊ TRÍ : (1).....(1950)

atgtggtggttggcattgttctcattgcttgtggccgccacggcctgggccaaggagccc
M W W L A L F S L L V A A T A W A K E P
tacctcaaggtccacgatgatacctttattccagatgccgtgttgcggttaacagaggag
Y L K V H D D T F I P D A V L R V T E E
agtgcctccattggttgcacgagaggacctccgccgtggtcaacggcacagtccgggg
S A S I G C I E R T S A V V N G T V P G
ccgatttttggaattccaatccggcagcgtggtatgggttcgtgtttacaatgatatggca
P I L E F Q S G S V V W V R V Y N D M A
gacaaaaaacctcaccatgcattggcatggcctcaccatggcagccgctccattcgccgac
D K N L T M H W H G L T M A A A P F A D
ggttccgttgacagccagccagtgggcgatcgagccgttcaaattcttcgactacgagctc
G S V A A S Q W A I E P F K F F D Y E L
aacctattcgacatcaagcctgggacgtatttctaccactcccatgtcggattccaggca
N L F D I K P G T Y F Y H S H V G F Q A
attactgccacgggtcccctgctcattacaaaaagccgggagaggagccgccgtatgag
I T A T G P L L I T K K P G E E P P Y E
tacgaggaggaacgcacatcgtcctcttttccgacctttacaacacgacggaccacgacatt
Y E E E R I V L F S D L Y N T T D H D I
gaaaccgggactggtagccagccctttcaaattggagcggtgaggtcgggagacgtcctcgtc
E T G L V A S P F K W S G E V G D V L V
aacggatatggcatatcgagtagccggccacgaacgacgccgagagttgcaaccttgca
N G Y G I S Q Y P A T N D A E S C N L A
cagattcccgtggaagctgggaagacgtaccgtttgcgtttcattggtgccactgccctt
Q I P V E A G K T Y R L R F I G A T A L
tctttcctctccgttggctttgagaagcacaatcttaccatcatcgaggcggtatggccac
S F L S V G F E K H N L T I I E A D G H
tacacggaacctgcagaaatcagcttccttcaaattcggcgggcgccaacggtattctgct
Y T E P A E I S F L Q I G G G Q R Y S A
ctcctaaagacatggacttgcgaggagctcgccgccaagactgctggtcgaaaccagttc
L L K T W T C E E L A A K T A G R N Q F
tacatccaaatcgagactcgtgacaggcccaagaacctcaccacctacgccattctggac
Y I Q I E T R D R P K N L T T Y A I L D

tacagcgacagctgcgccaccaatagcatcagcgtcgaggccaccatcggcaaggactca
 Y S D S C A T N S I S V E A T I G K D S
 acatcgcccccttctctacgccccctcttcgctcagcaaacccggcaacaacaacaaaaac
 T S P P S Y A P P S L S K P G N N N K N
 aacaacaacaccacgctctcccgcaagacgcccgaagacccccaccgctgcacctcccg
 N N N T T L S R K T P P K T P P L H L P
 cccaccgtgcaaggctggctcgaccatgacctccaccgctcgaaacctacacggacttc
 P T V Q G W L D H D L H P L E T Y T D F
 ccgacggcgacgaggtcacgcgcaccgtctacatggacatcttccagctagggcaggac
 P T A D E V T R T V Y M D I F Q L G Q D
 ggctacgtcaagtggcgcgagaacaacctgtcgtgggtacgagcacacgcccgaaggttccc
 G Y V K W A Q N N L S W Y E H T P K V P
 tacctcgtcgcgctctacaccaacagcacgcagtagctgcccgaactacgactatgcccgtc
 Y L V A L Y T N S T Q Y L P D Y D Y A V
 gcatcgggggacggggccacgacgaccgcgtcggtgcctggcccggcaagatgggaggggtg
 A S G T G H D D R V G A W P A K M G E V
 ctcgagatcatcgtcgtcaacacgggcagctacagcggcgcatggacgtgcatcccatg
 L E I I V V N T G S Y S G G M D V H P M
 cacctgcacggcgcgaccccgcttctatttggggagtggaagggcacgtacaacagagag
 H L H G A H P F Y L G S G K G T Y N R E
 gagaacgagaagaagttggcggggaggggtgccggtcacgagggattcgatgatgctgtat
 E N E K K L A G R V P V T R D S M M L Y
 cggtatggcgagaaggaggagccgcacaaggataatagctggattgcgctgaggataagg
 R Y G E K E E P H K D N S W I A L R I R
 gtgacgcagccgggggtgtggatgttccattgccatacgtggcgcatatgattatgggt
 V T Q P G V W M F H C H T L A H M I M G
 atgcaaacggctgtgggtgtttggcgattcgaaggatcctcacgctgccgctgccgatg
 M Q T V W V F G D S K D I L T L P L P M
 gtggaaggttacttggtgcctggcgagatgtgttcggtgatgatgatcatgatccgggtg
 V E G Y L V P G G D V F G D D D H D P V
 gtgggtgcacttctttgacctggacgacgacgacgatgatgatgatgccaataagacc
 V V H F F D L D D D D D D D D D A N K T
 gacgggaacggcggaagaatggccggtag
 D G N G G K N G R -

SEQ ID NO : 90

CHIỀU DÀI : 649

LOẠI : PRT

SINH VẬT : *M. phaseolina*

MWWLALFSLVAATAWAKEPYLKVHDDTFIPDAVLRVTEESASIGCIERTSAVVNGTVPGPILEFQSGSVVWVRV
 YNDMADKNLTMHWHGLTMAAAPFADGSVAASQWAIIEPFKFFDYELNLFDIKPGTYFYHSHVGFQAITATGPLLIT
 KKPGEPPYEEERIVLFSDLYNTHDHIETGLVASPFKWSGEVGDVLVNGYGISQYPATNDAESCNAQIPVE
 AGKTYRLRFIGATALSFLSVGFKEHNLTII EADGHYTEPAEISFLQIGGGQRY SALLKTWTCEELAAKTAGRNQF
 YIQIETRDRPKNLTTYAILDYS DSCATNSISVEATIGKDSTSPPSYAPPSLSKPGNNNNNNNTTLSRKT PPKTP
 PLHLPPTVQWLDHDLHPLETYTDFPTADEVTRTVYMDIFQLGQDGYVKWAQNNLSWYEHTPKVPYLVALYTNST
 QYLPDYDYAVASGTGHHDRVGAWPAKMGEVLEIIVVNTGSYSGGMDVHPMHLHGAHPFYLGS GKGTYNRENEKK
 LAGRPVPTRDSMMLYRYGEKEEPPHKD NSWIALRIRVTQPGVWMFHCHTLAHMIMGMQTVWVFGDSKDILTLPLPM
 VEGYLVPGGDVFGDDDHDPVVVHFFDLDDDDDDDDANKTDGNGGKNR*

SEQ ID NO : 91

CHIỀU DÀI : 2569 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

CGCCCGATTACACCATTCCTCTGACACTCTCCTCCGGAGCTCGCATCAGGGCGCTCTCTCGCCCTCCTGACCTT
 CCGTCCGGCGCCGATATCGTCTGTTCCTGATATGAGGGGTAATCGCGACGGTGCTCCCGCCCACTCTCAGCC
 ATGGCCAGTCTCTGTTCTCTGCTTGCTGCTGCCCTGTGCAGCCGCGCTGCCACAGTCACGTATGACTTTAAAC
 GTGACCTGGGTACCCGCCAACCTGATGCTGCTTTTCGTCGCACTACCATTTGGCATCAATGGCCAGTGGCCGCTC
 CCTGCAATCGACGTCACCAAGGGCGACCGCGTTGTATCAACGTCAACAACAGCTGGAACGGAGAGTACCAGT
 TTGCATTTCCATGGAATCTACATGAATGGCACCAACCACATGGATGGCCCGACTGGCGTGACCCAGTGCAGATA
 CCGCCCGGGAGCTCATTACATAACAATTTTACGGTATAGCACGATCCCACCCCTTGTCTTACTTATAGCTAACCGC
 ACGCAGGTCGACCAACCGGGAACCTTACTGGTGAGTAGCCAGCTGTTGGCGTCAAGACTGGGAAAGCTGACGTAGC
 TTCAGGTATCATTCTACAACCGCGCCAGTATCCCGATGGATGCGAGGGGCCTTTTATTGTCTCAGAGACCCCGAC
 AATCCGTTCAAGGATGACTATGATGAGGAAGTCGTCCTGACATTCTCCGACTGGTACCACGACCGGATCCCCACC
 CTCATGAAGAGTTTCATTAGTGTCACAAACCCACCGGCGCGGAGCCTGTCCCGAACGCGGCACTGATGAATGAC

ACTCAGAACCTCACCTTCCAGATGACTCCCGGGCGGAGATACATGTTCCGACTAATAAACATTGGTGCATTTGCT
GCTCAATACGTCTGGTTCGAAGGCCATACCATGCGTATCGTAGAAGTCGACGGCGTGACACCGAAGCCGCAGAT
GCAGAGCGAATTTACATGACTGCCGCTCAACGCTACAGCGTAATCATCACCGCAAAGAACGAATCTACCTCAAAC
TTCGCCTTTGTTGGAAGCATGGATCAGGTCAGCTATCCTTTTACTGGGCCGACATTTTCCTTCTTCTAACATAAT
GCAGGATCTCTTTGATACTATTCCAGCAGGCCCTTAATAATAACGTGACCGGGTGGCTTGCTACAATCAACAGAA
CGGCTTGTTGCCACCTTTAGCTATCGGGGACTACGATCCGTTTCGATGACTTCACACTTGTCCTCAAGATGGTAT
GGAGCTCTACGATCACGTCGATCATTCCATCACCCCTCGATATGAAGATGGACAATCTCGGGGACGGAGCAAATTA
GTATGTTGACTGTAGTTGACAACGGCCAAATGCCATGACTGACGAGCGCAGTGCCTTTTTCAACGACGTGACCTA
CGTCGAGCCCAAGGTGCCGACTCTGTACACGGTGTTGTCTACTGGCAATAATGCCACAGACTCGAGAATCTACGG
TAGCAACACCAACAGCTTCATATTGGCTAAGGACGAAGTCGTCGAGATCATCTCAACAACAATGATCCGGGAAA
GCACCCTTTCCATCTGCATGGGCACGCATTCCAAGCAATCGTTCGCTCCGAAGAAGCCGGCGCGTACGTGGC
AAACGAAACCTTCCCGCAGACGCCAATGCGCCGCGATACAATTCTTGTCCGGCCCAATGGCAATATAGTACTGAG
ATTCAAGGCTGACAACTCTGGTGTCTGGCTGTTCCATTGTCTATTTGAATGGTATGTCCTTGGTATCTTTTATC
CTAGCTCTACTGACCTCTTACTGTTAGGCATGTTGCGTCGGGCCCTTATTGCCACCATGATAGAGGCACCGCTTGA
CCTGCAATCTTCTCTCGGCAACAGCATTCCCTGCTGATCACTGGCGAGCTTGTGCCGCCGCCGGCACACCTACTGC
AGGCAACGCAGCAGGCAACACGATTGATTATCTTGACCTGACTGGTGAAAACAAGAGCCCCGGCCCGCTCCCTC
TGGCTTTGAAGCGAAAGGCATTGTGCGCTTGTCTTCAGTTGTATTGCTGCTGCTTGGGCATGGCAGCCATCGT
ATGGTACGGGATGGCGCTCTGACAGATGGAAGTGGTCAGCAACAACACGTTGTGGATCAACAGCCGCAAGGACG
CTCCGCACAAGTTGGCGTCATGCCCCTAGCGGTGCGGGCAAAACGTGAGCAAAGGATGGCCAGCCGACTGATTT
GGAGGCAGAGATGCGGCAAGAAAAGAGCGGAGAAATTACCATGGTGGAGCAGAGGAGGAGCAGCCGTGGGTATC
GAACATGCTGGATGTCAGGAGGAGTAGCCGCGGTTCTGCTAACATGCTAGAGACCAGAAGAAGCAGTCGTGGCAG
TGAGCAGATGCTGTCATGAGGACCAACATCATGGTAGCAAGGTAAGTGGAAATCGGCTGTGAAGAAATACAGGGT
CAGTTGTGAGAGGGAAGCTATCACGAATGGCTGAAGAGCTTTGGGCCCTTGTACGATGTAATGTTGTCCGAAAA
TGTTGTGGTAATTATTGTA

SEQ ID NO : 92
CHIẾU DÀI : 2016
LOẠI : ADN
SINH VẬT : *M. phaseolina*
TÊN/KHÓA DẤU HIỆU : CDS
VỊ TRÍ : (1).....(2016)

atggccagtcctctgttccctcctgcttgctgctgacctgtgcagccgctgccacagtc
M A Q S L F L L L A A A L C S R A A T V
acgtatgactttaacgtgacctgggtcacccgcaaccctgatgctgcttttcgtcgact
T Y D F N V T W V T A N P D A A F R R T
accattggcatcaatggccagtgccgctccctgcaatcgacgtcaccaagggcgaccgc
T I G I N G Q W P L P A I D V T K G D R
gttgatcatcaacgtcaacaaccagctggaacggagagtaccagtttgatttccatgga
V V I N V N N Q L E T E S T S L H F H G
atctacatgaatggcaccaaccacatggatggcccgactggcgtgacctgagagata
I Y M N G T N H M D G P T G V T Q C E I
ccgccccgggagctcattcacatacaattttacggctcgaccaaccgggaacttactggtat
P P G S S F T Y N F T V D Q P G T Y W Y
cattctcacaaccgcgccagtatcccgatggattgcgagggccttttattgtcagagac
H S H N R G Q Y P D G L R G P F I V R D
cccgacacccggttcaaggatgactatgatgaggaagtcgctcctgacattctccgactgg
P D N P F K D D Y D E E V V L T F S D W
taccagaccggtccccaccctcatgaagagtttcattagtgtcaciaaacccccaccggc
Y H D R I P T L M K S F I S V T N P T G
gcgagcctgtcccgaacgcggcactgatgaatgacactcagaacctcaccttccagatg
A E P V P N A A L M N D T Q N L T F Q M
actccccgggagagatacatgttccgactaataaacattgggtgatttgctgctcaatac
T P G R R Y M F R L I N I G A F A A Q Y
gtctggttcgaaggccataccatgcgtatcgtagaagtcgacggcgtgtacaccgaagcc
V W F E G H T M R I V E V D G V Y T E A
gcagatgcagagcgaatttacatgactgccgctcaacgctacagcgtaatcatcaccgca
A D A E R I Y M T A A Q R Y S V I I T A
aagaacgaatctacctcaaacttcgctttgttggaagcatggatcaggatctctttgat
K N E S T S N F A F V G S M D Q D L F D
actattccagcaggccttaataataacgtgaccgggtggcttgcttacaatcaacagaac
T I P A G L N N N V T G W L V Y N Q Q N
ggcttggtgccaccttagctatcggggactacgatccggttcgatgacttcacacttggtg
G L L P P L A I G D Y D P F D D F T L V
cctcaagatggatggagctctacgatcacgtcgatcattccatcaccctcgatatgaag

P Q D G M E L Y D H V D H S I T L D M K
atggacaatctcggggacggagcaaattatgcctttttcaacgacgtgacctacgtcgag
M D N L G D G A N Y A F F N D V T Y V E
cccaagggtgccgactctgtacacggtgttgtctactggcaataatgccacagactcgaga
P K V P T L Y T V L S T G N N A T D S R
atctacggtagcaacaccaacagcttcatattggctaaggacgaagtcgtcgagatcatc
I Y G S N T N S F I L A K D E V V E I I
ctcaacaacaatgatccgggaaagcaccctttccatctgcatgggcacgcattccaagca
L N N N D P G K H P F H L H G H A F Q A
atcggttcgctccgaagaagaagccggcgctacgtggcaaacgaaaccttccccagacg
I V R S E E E A G A Y V A N E T F P Q T
ccaatgcgccgcgatacaattcttgtccggcccaatggcaatatagtactgagattcaag
P M R R D T I L V R P N G N I V L R F K
gctgacaatcctgggtgtctggctgttccattgtcatattgaatggcatgttgcgtcgggc
A D N P G V W L F H C H I E W H V A S G
cttattgccaccatgatagaggcaccgcttgacctgcaatcttctctcggaacagcatt
L I A T M I E A P L D L Q S S L G N S I
cctgctgatcactggcgagcttgtgccgcccgcggcacacctactgcaggcaacgcagca
P A D H W R A C A A A G T P T A G N A A
ggcaacacgattgattatcttgacctgactggtagaaaacaagagccccggcccgcttccc
G N T I D Y L D L T G E N K S P G P L P
tctggctttgaagcgaaaggcattgtcgcgcttgtcttcagttgtattgctgctgtcttg
S G F E A K G I V A L V F S C I A A V L
ggcatggcgagccatcgatggtacgggatggcgcccttgacagatggaactggtcagcaa
G M A A I V W Y G M A P L T D G T G Q Q
caacacgcttgtggatcaacagccgcaaggcgtccgcacaaagttggcgatgccccta
Q H V V D Q Q P Q G R S A Q V G V M P L
gcggtcgcggaacagtgagcaaggatggcagccggactgatttggaggcagagatg
A V A A K R E Q R M A S R T D L E A E M
cggcaagaaaagagcggagaaattaccatggtggagcagaggaggagcagccgtgggtca
R Q E K S G E I T M V E Q R R S S R G S
tcgaacatgctggatgtcaggaggagtagccgcggttctgctaacatgctagagaccaga
S N M L D V R R S S R G S S N M L E T R
agaagcagtcgtggcagtgagcagatgctgtcatga
R S S R G S E Q M L S -

SEQ ID NO : 93

CHIỀU DÀI : 671

LOẠI : PRT

SINH VẬT : *M. phaseolina*

MAQSLFLLAAALCSRAATVTYDFNVTWVTANPDAAFRRTTIGINGQWPLPAIDVTKGDRVINVNQLETESTS
LHFHGIYMGNTNMDGPTGVTQCEIPGSSFTYNFTVDQPGTYWYHSHNRGQYPDGLRGPFIVRDPDNPFKDDYD
EEVVLTFSDWYHDRIPITLMSFISVTNPTGAEPVPNAALMNDTQNLTFQMTPGRRYMFRLINIGAFAAQYVWFEG
HTMRIVEVDGVYTEADAERIYMTAAQRYSVIITAKNESTSNFAFVGSMQDLFDITIPAGLNNNVGTGLVYNQQN
GLLPPLAIGDYDPFDFTLVPQDGMELYDHVDHSITLDMKMDNLGDGANYAFFNDVTYVEPKVPTLYTVLSTGNN
ATDSRIYGSNTNSFILAKDEVVEIILNNNDPGKHPFHLHGHAFAQIVRSEEEAGAYVANETFPQTPMRRDTILVR
PNGNIVLRFKADNPVWLFHCHIEWHVASGLIATMIEAPLDLQSSLGNSIPADHWRACAAAGTPTAGNAAGNTID
YLDLTGENKSPGPLSPSGFEAKGIVALVFSCIAAVLGMAAIVWYGMAPLTDGTGQQQHVVDQPPQGRSAQVGVMPPL
AVAAKREQRMASRTDLEAEMRQEKSGEITMVEQRRSSRGSSNMLDVRRSSRGSSNMLETTRSSRGSEQMLS*

SEQ ID NO : 94

CHIỀU DÀI : 3063 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

ACGCCTCGCCAACCTCATCTTCGTGCAACAGGCTTAGGTATCAAGCCTCCTTTGCAATCACTTACGCGGCCCTCT
CTACCTCCTGCTCATGTTTGAACAAATCACACCTCTTTGAAACGCTAGAAACACACTGTTTCTATACCGTCGAGA
ATGGCCTGGACCTCAGCTCTAGATCGATCCTCGTGGCTTGCCCCAGCTTCCCCACCAAGGGGGTTTGGTAAGTCG
TAAAAGTTTCTCACAGTTCGCAAATTCAGAGTACCGGTTTACGCTCTGATCTTTGCCACTGTACGGCGGCAGC
TGAAGCCGCGTGTACGTGCAATCTCCCCGAATACGTATGTGCCGGACTTGTATTTCTTCATACCCAGCCGTCT
AGGATCGTTCCGTATCGGTCTCTGCAACGACTACCACTACGACCACTGGGAAATCCTCCACGTTCTCTACGCTGA
CCAGCACTTCATCTCTCGTCTACAGCGAAGACTAGCTCGGCCAGCTCGGCATCGCTTTACGCACTACTCTTTCAC
CTGCCAGCTCTTCTCCGTGAGCTCTTCGTGATCAGATCCTCTTCTGGCGGATCTTCATCTGCTAGTGGTACTTCTT
CTGCCGCTCATATGTCAGCTCTTCTTTGAGCTCGCGTTCTGCTGTTTCTTCCCTAGTGACCAGCATATCCT
CTTCTCTCACCTCAACTGGCACGGTATCGTCTCGCCCGCTTTCGAAGAGCTCCAGCTCTTCATCTTCCACTCTTT

CAGCCTCTTCAACCACAAAGGTTGCTACTGGTACCCCGTGTGCTGGAAACACTGCGGCTTCCAGGACAGCCTGGT
 GCGATCACACCATTGACGATGACTACTATCCCATTTATCCCAGATACCCGGAGTCACACGCGAATACTGGTTCCGACC
 TTGTGCGAGGTGACTGTTGCCCCGTGATGGCATTGAGAGAGCTGCCATGGCGGTCAACGGCAGCATCCCTGGACCCA
 CCATTGAGGCGGATTGGGGCGATATCGCCGTCATGTCATGTGACCAACAGCTTGATCTCGAGCAAGAACGGCACTA
 GCATTCATTTCCACGGTATCCAGCAGAACTTCACAAATCAAACGGACGGTGTTATTTCCATCACTCAGTGCCCTA
 CCGCTCCTGGCGAGAACTACACGTATACCTGGAGGGCTGAGCAGTACGGAACATACATGGTGAGCTTGATCCGAGG
 ATACAGATGTAAAGGAGATGCTAAGATATCACAGGTACCATTCTCACTTTGCGCTTCAAGCTTGGGAAGGTGTCT
 TTGGAGGTACATACATCCCCGCTCCAGGACCTCCAACATTTATTAACGAAGCCCTGCAGGTATCAAGATCAATGG
 TCCGGCGAGTGCCAACTACGATTACGATCTTGGCCATGTCTTCCTGTGAGTTTCGACGCCCCCTTTTCCACTCAT
 GTATAAAGTAAGCTTATGCGACAAAAGCCAACGATTGGAGCCACGAGACTTCGAGCTCTCTCGAGATCGTTTCTG
 CAATTCGAGGTCCGCCAACACTTGAGAATGCCCCATTATCAACGGTACCAACGTGTACAACAACTCGGGAACCATAA
 CGGGCTCTCGTTTTGAGACAACATTTGAGGAAGGCAAACTCTACAGGTTGAGGTTGGTCAGCGGCCCATCGACA
 CGCATTTCAAGGTGTCATTGGGTAAGTTGAGCCAAAACATTGAGATTTCCCTGAGATCTAAAGCTGATTGTGTTGC
 AGATAACCACAGCATGCTTGTATAGCCAACGACCTTGTACCCATCGTGCCGTACAACACGACCGTCTGAACAT
 CGGAATGGGTATCCACTTTCTATCAGGAACTGGCCATCAAACAGACTGAACCCAGGCCAACGCTACGATGT
 AATCATCACCGCCAACCAAGCCGTAGTCGCCACCGACTTCTGGCTGCGTGCCGTCCCGCAGACGGCCTGCTCCAA
 TAACGCCAACCAGACAATATCAAAGGAATAATCCGGTACAGCACCTCCACCTCCGCTACGACTGGACGAACGA
 ATGCGTCGACGAGGCCCTCACCAACCTCGTGCCATGGGTGACTAAGAACGCGGCCCTCGGGCAGAGTCTATCTGA
 AGTGGTGACGCTGGGTGCGAACGTTGACAACCTCAACCGCTGGATGATGAACAGCACATCGATGGTTGTGAGTG
 GAATGATCCCTCGTTGCTGTCAGGTTTGGGAATAATGATACCAATTTACGGATACGAGCGGCGTGGTCAGGCTTGG
 TACGGCGGACGAGTGGGTGATGTTTGTATCGAGATGACGCTGCCGATTCCGCGTCCCATCCATCTTACGGGCA
 TGATGTAGATTTCTTCCAGCCGCCCCCCCCCCCCCTTTTTTTTGGGCTCTACTTCCTAGAATACCGGACGGAG
 TCGTTGCGGAGCGCGCCTTCGCCGACTTTTTCGCTTCCGATGTCCTCAAGAAAGGTGTCATGTTGTAATGTGTATC
 CCAGAGGCTTCCCGCTGACTTGTGTTTGTCAACACAGTTCAACATCCTCGCCCAAGGGACCGGCACGTACGATTC
 ATCCGTGTCGCTCACCTGTCAAACCCGCCGCGGAGGACGTTGCGCTGCTGCCGCGAGCCGGGTACTTGGTGAT
 TGCATTGCGCACCGATAACCTGGAGCCTGTGAGGACTTCTCTTTTCTTGTGCATCCCTGGATGTCATCTTGC
 TGTGAAATTTTTTGGCTGACGCAATTCGTCTCAATGCAGGGCTCATGCACATCGGCCGCGCACCTACGGAG
 GGCTTCGCAATCCAGATCCTCGAGCGCTGGGACGAGATCTCGCCGCTTATCGACTACGAGACGCTTGAAGGCAAC
 TGCAACAGATGGGACGCGTACGTGTCGGTCAGCGATGTCGTACAGGACGATTCCTGGCGTATGAGCAATGAAGGTT
 AGAGTATTTGCGACGTGTTTTATCATATATACAACCTTTTTTAATCGTAGTCGCTCCTTAACCATAATGTTAATCC
 ATCACTTGGGCCATCCAAGCACCAACAAAACATTTGAGCCTTCCCAATCAGGACTGGCACAAG

SEQ ID NO : 95
 CHIỀU DÀI : 2085
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 TÊN/KHÓA DẤU HIỆU : CDS
 VỊ TRÍ : (1).....(2085)

atggcctggacctcagctctagatcgatcctcgtggcttgccccagcttccccaccaagg
 M A W T S A L D R S S W L A P A S P P R
 gggtttgctgaagccgctgtacgtcgaattctccctgaatacgtatgtgccggacttg
 G F A E A A C T S N S P L N T Y V P D L
 tatttcttcataccagccgtctaggatcggttcggtatcggtctctgcaacgactaccac
 Y F F I P S R L G S F G I G L C N D Y H
 tacgaccactgggaaatcctccacgttctctacgctgaccagcattcatcttctcttc
 Y D H W E I L H V L Y A D Q H F I F L F
 ttccgtgagctcttcgctcgatcagatcctcttctggcgatcttcatctgctggtacttc
 F R E L F V D Q I L F W R I F I C W Y F
 ttctgcccgtcatcatatgtcagctcttctttgagctcgcttctgctgtttcttccct
 F C P L I I C Q L F F E L A F C C F F P
 agtgaccagcatatcctcttctctcacctcaactggcacggtatcgctcctcgcccgcttc
 S D Q H I L F S H L N W H G I V L A R F
 gaagagctccagctcttcatcttccactcttccagcctcttcaaccacaaagcctgggtgc
 E E L Q L F I F H S F S L F N H K A W C
 gatcacaccattgacgatgactactatcccattatcccagataccggagtcacacgcgaa
 D H T I D D D Y Y P I I P D T G V T R E
 tactggttcgaccttgctcgaggtgactgttgcccctgatggcattgagagagctgccatg
 Y W F D L V E V T V A P D G I E R A A M
 gcggtcaacggcagcatccctggacccaccattgaggcggattggggcgatatcgccgtc
 A V N G S I P G P T I E A D W G D I A V
 atgcatgtgaccaacagcttgatctcgagcaagaacggcactagcattcatttccacggt
 M H V T N S L I S S K N G T S I H F H G
 atccagcagaacttcacaaatcaaacggacggtgttatttccatcactcagtgccctacc
 I Q Q N F T N Q T D G V I S I T Q C P T

gctcctggcgagaactacacgtatacctggagggctgagcagtacggaactacatggtac
 A P G E N Y T Y T W R A E Q Y G T T W Y
 cattctcacttttgcgcttcaagcttggaaggtgtctttggaggtacatacatccccgct
 H S H F A L Q A W E G V F G G T Y I P A
 ccaggacctccaacattttattaacgaagccctgcaggtatcaagatcaatggtccggcga
 P G P P T F I N E A L Q V S R S M V R R
 gtgccaaactacgattacgatcttggccatgtcttctccaacgattggagccacgagact
 V P T T I T I L A M S S S N D W S H E T
 tcgagctctctcgagatcgtttctgcaattcgaggtccgccaacacttgagaatgccctt
 S S S L E I V S A I R G P P T L E N A L
 atcaacggtaccaacgtgtacaacaactcggaaccataacgggctctcgttttgagaca
 I N G T N V Y N N S G T I T G S R F E T
 acatttgaggaaggcaaatacctacaggttgaggttggtcagcggcgccatcgacacgcat
 T F E E G K S Y R L R L V S G A I D T H
 ttcaaggtgtcattggataaccacagcatgcttgtcatagccaacgaccttgtagccatc
 F K V S L D N H S M L V I A N D L V P I
 gtgccgtacaacacgacgcctcgaacatcggaatgggccaacgctacgatgtaatcatc
 V P Y N T T V L N I G M G Q R Y D V I I
 accgccaaccaagccgtagtcgccaccgacttctggctgctgctgccgtcccgcagacggcc
 T A N Q A V V A T D F W L R A V P Q T A
 tgctccaataacgccaaccagacaatatcaaaggaataatccggtacagcacctccacc
 C S N N A N P D N I K G I I R Y S T S T
 tccgcctacgactggacgaacgaatgcgtcgacgaggccctcaccaacctcggtgccatgg
 S A Y D W T N E C V D E A L T N L V P W
 gtgactaagaacgcggcctcgggcacgagtcctatctgaagtggtagcgtgggtcgcaac
 V T K N A A S G T S L S E V V T L G R N
 gttgacaacctcaaccgctggatgatgaacagcacatcgatggttgcgagtggatgat
 V D N L N R W M M N S T S M V V E W N D
 ccctcggttgcgtgcaggtttggaataatgataccaatttcacggatacgagcggcgtggtc
 P S L L Q V W N N D T N F T D T S G V V
 aggcttggttacggcggacgagtggtcatgtttgttatcgagatgacgctgccgattccg
 R L G T A D E W V M F V I E M T L P I P
 cgtcccatccatcttcacgggcatgatttcaacatcctcgcccaagggaacgggcacgtac
 R P I H L H G H D F N I L A Q G T G T Y
 gattcatccgtgtcgtccacctgtcaaaccgcgcggagggacgttgcgctgctgcc
 D S S V S L T L S N P P R R D V A L L P
 gcagccgggtacttggtgattgcattcgccaccgataaccctggagcctggctcatgcac
 A A G Y L V I A F A T D N P G A W L M H
 tgccacatcgccggcgccactacggagggcttcgcaatccagatcctcgagcgtgggac
 C H I G R R T T E G F A I Q I L E R W D
 gagatctcgccgcttatcgactacgagacgcttgaaggcaactgcaacagatgggacgcg
 E I S P L I D Y E T L E G N C N R W D A
 tacgtgtcggtcagcgatgtcgtacaggacgattctggcgtatga
 Y V S V S D V V Q D D S G V -

SEQ ID NO : 96

CHIỀU DÀI : 694

LOẠI : PRT

SINH VẬT : M. phaseolina

MAWTSALDRSSWLAPASPPRGFAEAACTSNSPLNTYVPDLYFFIPSRGLSGFGLCNDYHYDHWELHVLVYADQH
 FIFLFFRELFDQILFWRIFICWYFFCPLIICQLFFELAFCCFFPSDQHILFSLNWNHGVILARFEELQLFIFHS
 FSLFNHKAUCDHTIDDDYPIIPDTGVTREYWFDLVEVTVPAPDGIERAAMAVNGSIPGPTIEADWGDIAVMHVTN
 SLISSKNGTSIHFGIQNFNTQTDGVISITQCPTAPGENYTYTWRAEQYGTWYHSHFALQAWEGVFGGTYIPA
 PGPPTFINEALQVSRSMVRRVPTTITILAMSSSNDWSHETSSSLEIVSAIRGPPTLENALINGTNVYNNSGTITG
 SRFETTFEEGKSRYRLRVSGAIDTHFKVSLDNHSMVLVIANLVPVYPYNTTVLNIGMGQRYDVIIITANQAVVATD
 FWLRAVPQTACSNNANPDNIKGIIRYSTSTSDYDWTNECVDEALTNLVPWVTKNAASGTSLSVVTLGRNVNDNLN
 RWMNSTSMVVEWNDPSLLQVWNNNTNFTDTSGLVRLGTADEWVMFVIEMTLPIPRPIHLHGHDFNILAQGTGT
 DSSVSLTSLNPPRRDVALLPAAGYLVIAFATDNPGAWLMHCHIGRRTTEGFQILERWDEISPLIDYETLEGNC
 NRWDAYVSVSDVVQDDSGV*

SEQ ID NO : 97

CHIỀU DÀI : 2449 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

GCGCACCCCCGAATTTTCATGTGGCACTTTGCTGGTGGCGATGCATAAAGTCTCCCAAACCTTGATCTTCAGCGTC
GAGAGCAGCTCAACGTCCGATCCGATTTTTTCGTTACCCGCGCTTCTTGTCCCTCAAGCCCCAATACAATTC AAC
ATGTTTTCTTTTGTCTCCAGAACCTCAACTGCCGCTATGGCTGGACTTTGGCCTTTCAATATGGTCAATGTTTC
GGCGGAAACAGTCAAGACTGCGTACACTCGCCCCAGAGCAGACACTGCTGGCATGATGGGTTCGATATCAACACC
GATTATGAAGCTAAGATACCGCCAGGGAAGCTCGTGGAGGTAAGATAAACCAGCCACAGACGGCTCAGGCCATCG
CAACTCACCTCCCTTAGTACGATTTCACTATCTCGGAAGCAGTCCCTTGTCCGGACGGATACTTGACGAATGTCA
CTCTGGTAAATGGAGTATTCCTTGGTCCAACCTCGAAGCTGAATGGGGAGATACGATCAGTATGCCTTCTTCAT
ATGAAATCTTTCTGTGTGCCCTTTGACGTGATTGCTGACATTTCCCTGCCAGGGATAACGATCCACAACAACCTCA
CAAACCACAATGGCACATCCATACATTGGCATGGAATTCGCCAATTCGAGACCAATTGGCTTGATGGTGTTCCTCG
GCGTCACTCAATGCCCGTCAAAGGTACTACGGCTGGTTTGTCTTGGGAGAGAAACAGGGGAGCTAAGTGTCAACCA
GCCTGGAGACTCGCAAGTTGTGCAATTCCGAGCGATGCAATATGGGACCGCATGGTATCATTACACTACAGTCT
TCAGTGTCAGTATCAATCACAGAGCTCTTGAGCGGCACTTACGCTAACCATGATTAGATACAAACGGAGTTCTCG
GTATTTGCTATCTCCCTTCCATACAGTTCACAACCTGTGGGCTGATCCCGTTGCAGGACCCATTACATCAAGGG
ACCCTCGAGCATGAACACGACGTGGATCTCGGGCCACTGTTGATCAGTGATTGGTATCACCACGATGCCTTCGG
CCTATTCATTATGAGATCGCTTACCAGCAGCGCCGCTTCCGGTCACAACATCTTGAATGGCAAGGGAGTCTT
TGATTGCGATCCAGCCAGCGATGCTCGTTGTACGGGAGAGCACAACGGCACGAGATAGTGTTCGAAGAGGGTAA
AAGATACAAAATTGGACTAATAAAATACCGGCAGCCTTCTGACATACAAGTTCTGGATCGATGGCCATAATTTTAC
AGTTGTGCAGACGGATTTCTGTTCCCATCAAACCATACGTCACCGACGTTCTGATCGTGGGATAGGTATGTTTCAG
TGCCTACAAAGAATTGGCGGTAAACTGATCATTTCCAGCTCAACGATACGAGATCATTATCGAAGCAAATGTGAC
ATTCACACGTGGCTCCAACCTCTGGATTACGCAACGTACTGTGACGATGATGACATGTTGGACTCGAGAGTTGG
CATAGTCCGCTACGACGGCAGCGACGGTCTGATCCGCACACGCCGCCCAAGAGTGAGCAACACCCCGGATACGG
GTGTCGTGATCCAGCCACGGAGAATCTTGTTCCTATCGTGAAGAGGGAAGTAGGCAAGAGAGTGAACGGGCTCAG
CCCTGCTGATTACCTCAGGATCGGCCTGCAGGGCTGGCCCAACATCTCGGACACAGATTCGCTCGTACACAAATG
GACACTTACCAACAGAACCAGTACATTGATTGGAGGGAGCCAACAATCAAGGCGCTCACTTCGGATTCGGGGC
TGATTTTGCAGGTGAGACATGCCCCATATACCTGGACTACGAGACTGGCGAGTGGGTGTACTTCGTATCGAGAA
CAACTACACGCTGAGCGACGCCAACACGCCCGCACCATCCCCCGCTCGGTCCATCCCATCCACTTACACGGGCA
TGACTTTCGTGATCCTCGCCCAGGGTGACGGCATGTTTCGACCCCGTCGACGTGGTGCCGAACCTCCACAACCCAAAC
CAGGAGGGACGTGGTCAATTGCCCGATTGGCGGGTACGTATGGATCGCATTCAGGTCAACAATCCAGGAGCGTG
GCTGATGCATTGCCATATCGCCTGGCATGCCAGCGCCGGACTCTCACTGCAGTTCATTGAGCAACCTGGCCTGAT
CAAGGGGTTAATGGAGCAGGCAGGAGCCTTGCCGGAGCTTGCTGATCGATGCGAGGACTGGACTGAGTACTACAA
TACTGTAAACATACCAAAGGCGCACTGCAAGATGATTTCGGGCATATGAGTTGGTATTGAACCTCCGGCTGAACA
CGCTCAGCCACGGCTACTTGCTTGCCCAAGATTGAGGGTTGGCCGAGAAATGAGTAGTGGGGATTATCACTTTCT
TAGAATGTGCACGAAAGGGTAGTGTTCGGTCATGTCTCTTAGCTTGAAA

SEQ ID NO : 98

CHIỀU DÀI : 1821

LOẠI : ADN

SINH VẬT : *M. phaseolina*

TÊN/KHÓA DẤU HIỆU : CDS

VỊ TRÍ : (1).....(1821)

atgttttcttttgtcgtccagaagcctcaactgccgctatggctggactttggcctttca
M F S F V V Q K P Q L P L W L D F G L S
atatgggtcaatgttcggcggaacagtcgaagactgcgtacactcgccccagagcagacac
I W S M F G G N S Q D C V H S P Q S R H
tgctggcatgatgggttcgatatcaacaccgattatgaagctaagataccgccagggaag
C W H D G F D I N T D Y E A K I P P G K
ctcgtggagtacgatttcactatctcggaagcagtccttgctccggacggatacttgacg
L V E Y D F T I S E A V L A P D G Y L T
aatgtcactctggtaaatggagtattccctgggtccaaccctcgaagctgaatggggagat
N V T L V N G V F P G P T L E A E W G D
acgatcaggataacgatccacaacaacctcacaaaccacaatggcacatccatacattgg
T I R I T I H N N L T N H N G T S I H W
catggaattcgccaattcgagaccaattggcttgatgggtgttcccggtcactcaatgc
H G I R Q F E T N W L D G V P G V T Q C
ccgtcaaagcctggagactcgcaagttgtcgaattccgagcgatgcaatatgggaccgca
P S K P G D S Q V V E F R A M Q Y G T A
tggtatcattcacactacagtcttcagtatacaaacggagttctcgacccattcacatc
W Y H S H Y S L Q Y T N G V L G P I H I
aagggaccctcgagcatgaactacgacgtggatctcgggccactgttgatcagtgattgg
K G P S S M N Y D V D L G P L L I S D W
tatcaccacgatgccttcggcctattccattatgagatcgcttcaccgcacgcgcccgtt
Y H H D A F G L F H Y E I A S P H A P L
ccggtcacaaactatcttgaatggcaagggagtctttgattgcatccagccagcgatgct

P V T T I L N G K G V F D C D P A S D A
 cggtgtacgggagagcaccaacgggcacgagatagtggttcgaagagggtaaaagatacaaa
 R C T G E H Q R H E I V F E E G K R Y K
 attggactaataaaataccggcagccttctgacatacaagttctggatcgatggccataat
 I G L I N T G S L L T Y K F W I D G H N
 tttaacagttgtgcagacggatttctgttcccatcaaaccatacgtcaccgacggttctgatc
 F T V V Q T D F V P I K P Y V T D V L I
 gtcgggatagctcaacgatacagagatcattatcgaagcaaagtgtgacattcacacgtggc
 V G I A Q R Y E I I I E A N V T F T R G
 tccaacttctggattcacgcaacgtactgtgacgatgatgacatggttgactcgagagtt
 S N F G W I H A T Y C D D D D M L D S R V
 ggcatagtcgctacgagcgacgggtcgtgatccgcacacgcccgaagagtga
 G I V R Y D G S D G R D P H T P P K S E
 caacaccccgatacgggtgtcgtgatccagccacggagaatcttgttcccatcgtgaag
 Q H P G Y G C R D P A T E N L V P I V K
 aggggaagtaggcaagagagtgaacggggtcagccctgctgattacctcaggatcggcctg
 R E V G K R V N G L S P A D Y L R I G L
 cagggtggtgccccacatctcggacacagattcgtcgtacacaaatggacacttaccaac
 Q G W P N I S D T D S L V H K W T L T N
 agaaccagtagcattgattggaggaggccaacaatcaaggcgctcacttcggatgtcggg
 R T Q Y I D W R E P T I K A L T S D V G
 gctgattttgcgatgagacatgccccatatacctggactacgagactggcgagtgggtg
 A D F A D E T C P I Y L D Y E T G E W V
 tacttctgcatcgagaacaactacacgctgagcgacgccaacacgccccgcaccatcccc
 Y F V I E N N Y T L S D A N T P R T I P
 cgctcgggtccatcccatccacttacacgggcatgacttctgctgatcctcgcccagggtgac
 R S V H P I H L H G H D F V I L A Q G D
 ggcattgttcgaccccgctcgacgtggtgccgaacctccacaaccaaccaggagggtgacgtg
 G M F D P V D V P N L H N P T R R D V
 gtcaattgccccgattggcggtacgtatggatcgcatccaggtcaacaatccaggagcg
 V N C P I G G Y V W I A F Q V N N P G A
 tggctgatgcattgccatatacgctggcatgccagcgccggactctcactgcagttcatt
 W L M H C H I A W H A S A G L S L Q F I
 gagcaacctggcctgatcaaggggttaatggagcaggcaggagccttgccggagcttgct
 E Q P G L I K G L M E Q A G A L P E L A
 gatcgatgcgaggactggactgagtactacaatactgtaaacataccaaaaggcgactg
 D R C E D W T E Y Y N T V N I P K G A L
 caagatgattcgggcatatga
 Q D D S G I -

SEQ ID NO : 99

CHIỀU DÀI : 606

LOẠI : PRT

SINH VẬT : M. phaseolina

MFSFVVQKPLPLWLDGFLSIWSMFGGNSQDCVHSPQSRHCWHDGFDINTDYEAKIPPGKLVEYDFTISEAVLAP
 DGYLTNVTLVNGVFPGPITLEAEWGDITIRITIHNNLTNHNHGTISIWHHGIRQFETNWLGDVPGVGTQCPSKPGDSQVV
 EFRAMQYGTAWYHSHYSLQYTNGVLGPIHIKGPSSMNYDVLGPLLISDWYHHDAFGLFHYEIASPHAPLPVTTI
 LNKGVFDCDPASDARCTGEHQRHEIVFEEGKRYKIGLINTGSLTYKFWIDGHNFTVVTQDFVPIKPYVTDVLI
 VGIAQRYEIIIEANVTFTTRGSNFWIHATYCDDDMLDSRVGIVRYDGS DGRDPHTPPKSEQHPGYGCRDPATENL
 VPIVKREVVKRVNGLSPADYLRIGLQGWPNISDTSLVHKWTLTNRTQYIDWREPTIKALTS DVGADFADETCPI
 YLDYETGEWVYFVIENNYT LSDANTPRTIPRSVHPIHLHGHD FVILAQGDGMFDPVDVVPNLHNPTRRDVVNCPI
 GGYVWIAFQVNNPGAWLMHCHIAWHASAGLSLQFIEQPLIKGLMEQAGALPELADRCEDWTEYYNTVNI PKGAL
 QDDSGI*

SEQ ID NO : 100

CHIỀU DÀI : 2347 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : M. phaseolina

GATACAACCTCTTCTCCCCGCCTCTCTTTCACAAACTGTCTATCGCTCCTTTGTTTTCTGTCCCTCCTTAAGGTC
 ACATACTACTGAGACTATTCGTGTCTCGTGCCTCTCATTTAGTTTTTAGTTGATTATCTTGCCGTTGCTGTAAC
 ATGCCGTTTTCTCCACGCCTAAGTATATGCCTTTTAGCATTTTGCTCTCACATTGTGTTTCGCACTTTTCGATACCA
 AATGTAAAGGATAGGACCCCCCACTTAAGCCCCGCGCAATTACGGATTTCTGTTTCGGCAACCTGTCTTGGGGCGGG
 GCGGAGCCTTCCGACCCACTCGCACAAAGGTGCAGAACGTTTGGATGACATCAGTCTCCCGGCCCTGAAGATGGC
 TGTCTTTTCTCCAAAGATGCTCGGAACGCTGGCGCGATAACTTCAACATCGACACCGATTTCGACGAGCGCTTC

CCCACGACCGGGAAGACGGTCACTGTAAGGAATCATCGCACCTCTTGGAACCCCTTACGCTAATACCGAGGGCAGT
 ATAATCTGGAGATCACAAACACTACCATGGCCCCGTGACGGAATCGAACGCGTCGTCATGGCCGTGAATGGCCAAT
 ATCCCGGCCCCGACCCCTTATTGCTGATTGGGGGGACACGATGGTTATCAACGTCAAGAATAGCCTGGACCACAACG
 GTACGGGTCTCCATTTCCATGGACTGCGCCAATACAAAAGCAACGGCGCCGATGGGGCGAACGGGTATCACAGAAT
 GCCCAATTGCCCCCGGCGAGACCAAGACCTACACCTTTCAATGTACCCAGCACGGTAGCTCGTGGTACCACTCGC
 ATTACTCTGTCCAGTACTCCGACGGCGTTCTCGGCGGCATTATCATCAACGGCCCCCGCCGACGCGCACTACGACC
 ACGATCTTGGTGTGTACATGCTTTCTGACTGGTACCACACTCCAATGTTTGAACTAGCCGAAGCTGCCAGGCATT
 CGACAAGGGGGCCACCGAAGGCGGATAACGGACTCATCAACGGGACGATGAAGAGCCCTGACGGTTCTCTTGGAG
 CCTATGGCCAGATCCATGTGAAGAAGGGTCTGCGGTACAGGATTGCGGTAATGAATGTTGGCACTAACGACCACT
 ACCTCTTCTCTGTTGATGGGCACAACCTCACCGTGATCGCAAGTGATTTTCGTGCCGGTCTGCCCCTTTTTCGGCGT
 CCAGCATTTCCCTCGGTGTTGGTGAGATATCCGATCTTTATGTTGTTATTTTCGTGCTGACCGTCACCTTAGGAC
 AGAGGTACGACGTGATCCTTATCGCCGACCAAGACATTGACAACTACTGGATCCGCTCCGACCCGGACTCTGCCCT
 GTAGCGTTAAACGGCAACGCCGGCAACATAAAAGCCATCCTCTCGTATGACACGGCTCCCGCGGACGCTCAGCCGA
 ATAGCACGCGCCACAGCATCTCCTCGGGCTGCAAGGACATGGCAGTCGTGCCAAGAGTTGCTAATACCGTGCCCT
 CTGATCGCTTCGCGGACGCCGTCCAGAGCCTGGCGATGAGCGTCAATATCACGCAGCAGAACGGCCCCGCTCGTCC
 AGTGGTATATCAACGGCTCGGCTATGGAGGTGACTGGAGCTACCCGACGGTCCAATATGTCTTAGATGGGAACA
 CGTCGTACCCGCGTGAGCTGAACCTAGTCCAGCTGGATGAGGCGGACCAAGTGGTACTACTTTCGTATCCAGACCG
 TGCAAGGCTTGCGTGTCAACCTGCCGCACCCAAATTCATCTTCATGGCCACGTACGTTCCCTCCCTTTCCCTTCCTC
 CCTAAACGAAACAGTAAAAGAGGGACAAGTGAATGCGACGTACTGATTTGTTGTCCAATTGGTCCAGGACTTCTA
 CATCCTCGGCGCCGGTCCCGGCGAGTGGGACGGCAACATCGATGGCCTGCAGTTCGACAACCCCTCCGCGGCGCGA
 TACGGCAATGCTACCTGCGGGTGGCTACCTCATCCTCGCCTTTCCGGCCGACAACCCCTGGCGCCTGGCTCATGCA
 TTGCCACATCCCGTTCCACGTCCAGCAGGGCTTCGGGCTGCAGTTCCTTGGAGCGGCGGATGAGATTGAGGGCGT
 CATGGGCGATACGAGCCCGTTTTACAACGAGTGTGCGGCTTGGAGGACTACTATGGTGGGGGCCAGGCTTTTCA
 GCAGTCCGATTTCGGGCTTGTAAATTGGTCGTGGTGGCTTGCAGCAAGAGCAAAGCGCCGTCGGTGCTCTGGCGGTGA
 TTTTGTAGTGGCACGGTTTTTGTACTCATCAACTCCTAATCACTAAGTCAGCCACCCTCGTTTTCCGCCACCCTGT
 ACAGTCAGCATCATAACGTAAC

SEQ ID NO : 101
 CHIỀU DÀI : 1854
 LOẠI : ADN
 SINH VẬT : M. phaseolina
 TÊN/KHÓA DẤU HIỆU : CDS
 VỊ TRÍ : (1).....(1854)

atgccgtttttctccacgcctaagtatatgccttttagcattttgctctcacattgtgttc
 M P F S P R L S I C L L A F C S H I V F
 gcactttcgataccaaatgtaaaggataggacccccacttaagcccgcgcaattacgga
 A L S I P N V K D R T P H L S P R N Y G
 tttcgtttcggcaacctgtcttggggcgggcgaggaccttcgacccactcgacacaagg
 F R F G N L S W G G A E P S D P L A Q G
 gcagaacgtttggatgacatcagtcctcccgccccctgaagatggctgtcttttctccaaa
 A E R L D D I S L P A P E D G C L F S K
 gatgctcggaactgctggcgcgataacttcaacatcgacaccgatttcgacgagcgcttc
 D A R N C W R N N F N I D T D F D E R F
 cccacgaccgggaagacgggtcacttataatctggagatcacaaacactaccatggccccct
 P T T G K T V T Y N L E I T N T T M A P
 gacggaatcgaacgcgtcgtcatggccgtgaatggccaatatcccgccccgacccttatt
 D G I E R V V M A V N G Q Y P G P T L I
 gctgattggggggacacgatggttatcaacgtcaagaatagcctggaccacaacggtacg
 A D W G D T M V I N V K N S L D H N G T
 ggtctccattttccatggactgcgccaatacaaaagcaacggcgccgatggggcgaaacggt
 G L H F H G L R Q Y K S N G A D G A N G
 atcacagaatgcccaattgccccggcgagaccaagacctaacacctttcaatgtaccag
 I T E C P I A P G E T K T Y T F Q C T Q
 cacggtagctcgtggtaccactcgacttactctgtccagtactccgacggcggttctcggc
 H G S S W Y H S H Y S V Q Y S D G V L G
 ggcattatcatcaacggccccgcccgcgacgcgcactacgaccacgatcttgggtgtgtacatg
 G I I I N G P A D A H Y D H D L G V Y M
 ctttctgactggtaccacactccaatgtttgaactagccgaagctgccaggcatttcgaca
 L S D W Y H T P M F E L A E A A R H S T
 agggggccaccgaaggcgataacggactcatcaacgggacgatgaagagccctgacggt
 R G P P K A D N G L I N G T M K S P D G
 tctcttgagcctatggccagatccatgtgaagaagggtctgcggtacaggatttcgcta
 S L G A Y G Q I H V K K G L R Y R I R V
 atgaatgttggcactaacgaccactacctcttctctgttgatgggcacaacctcaccgtg

M N V G T N D H Y L F S V D G H N L T V
 atcgcaagtgatttcgtgccggtcgtgcccttttcggcgccagcatttcctcgggtgtt
 I A S D F V P V V P F S A S S I S L G V
 ggacagaggtacgacgtgatccttatcgccgaccaagacattgacaactactggatccgc
 G Q R Y D V I L I A D Q D I D N Y W I R
 tccgacccggactctgcctgtagcgttaacggcaacgccggcaacataaaagccatcctc
 S D P D S A C S V N G N A G N I K A I L
 tcgtatgacacgggtcccgcgacgctcagccgaatagcacgcgccacagcatctcctcg
 S Y D T A P A D A Q P N S T R H S I S S
 ggctgcaaggacatggcagtcgtgccaagagttgctaataccgtgccctctgacgttc
 G C K D M A V V P R V A N T V P S D R F
 gcggacgcgtccagagcctggcgatgagcgtcaatatcacgcagcagaacggcccgctc
 A D A V Q S L A M S V N I T Q Q N G P L
 gtccagtggtatatcaacgggtcggtatggaggtcgactggagctacccgacgggtccaa
 V Q W Y I N G S A M E V D W S Y P T V Q
 tatgtcctagatgggaacacgtcgtacccgcgtgagctgaacctagtccagctggatgag
 Y V L D G N T S Y P R E L N L V Q L D E
 gcggaccagtggtactacttcgtcatccagaccgtgcaaggcttgctgtcaacctgccg
 A D Q W Y Y F V I Q T V Q G L R V N L P
 cacccaattcatcttcattggccacgacttctacatcctcggcgcgggtcccgggcgagtg
 H P I H L H G H D F Y I L G A G P G E W
 gacggcaacatcgatggcctgcagttcgacaaccctccgcggcgcgatacggcaatgcta
 D G N I D G L Q F D N P P R R D T A M L
 cctgcgggtgggtacctcatcctcgctttccggccgacaaccctggcgctgggtcatg
 P A G G Y L I L A F P A D N P G A W L M
 cattgccacatcccgttcacgtccagcagggcttcgggctgcagttcttgagcggccg
 H C H I P F H V Q Q G F G L Q F L E R P
 gatgagattgagggcgtcatgggcatagcagccggttttacaacgagtggtgcggcttg
 D E I E G V M G D T S P F Y N E C A A W
 aaggactactagtggtggggcaggcttttcagcagtcctcgattcgggcttgtaa
 K D Y Y G G G Q A F Q Q S D S G L -

SEQ ID NO : 102

CHIỀU DÀI : 617

LOẠI : PRT

SINH VẬT : *M. phaseolina*

MPFSPRLSICLLAFCSHIVFALSIPNVKDRTPHLSRPNYGFRFGNLSWGGAEPSDPLAQAERLDDISLPAPEDG
 CLFSKDARNCWRDNFNIDTDFDERFPTTGKTVTYNLEITNTTMAPDGIERVVMAVNGQYPGP TLIADWGD TMVIN
 VKNSLDHNGTGLHFHGLRQYKSNAGDANGITECPIAPGETKTYTFQCTQHGSWYHSHYSVQYSDGVLGGIIIN
 GPADAHYDHD LGVYMLSDWYHTPMFELAEAAHSTRGPPKADNGLINGTMKSPDGS LGAYGQIHVKKGLRYRIRV
 MNVGTNDHYLFSVDGHNLTVIASDFVPVVPFSASSISLGVGQRYDVILIADQDIDNYWIRSDPSACSVNGNAGN
 IKAILS YDTAPADAQPNSTRHSISSGCKDMAVPRVANTVPSDRFADAVQSLAMSVNITQQNGPLVQWYINGSAM
 EVDWSYPTVQYVLDGNTSYPRELNLVQLDEADQWYFVIQTVQGLRVNLPHPIHLHGHD FYILGAGPGEW DGNID
 GLQFDNPPRRDTAML PAGGYLILAFPADNPGAWLMHCHIPFHVQQGFLQLFLERPDEIEGVMGDTSPFYNECAAW
 KDYYGGGQAFQQSDSGL*

SEQ ID NO : 103

CHIỀU DÀI : 2415 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

TCGAGCAAGATGAATGTATCCTCCGTGGAAAGTCTTTTTTCTCAGGGCGGCCGTCCGCTTTGACCTCTGCCCTCT
 TGCTGGATTATAGAATACCCAGACGCCCATGTAGTGCATGGCCTTGTGAACCTACCCACCACCTCTCTCGCGAAG
 ATGATTGGCCATTCTCTCGTAGCTGTAGCACTCGGCAGCGCAGCCATGGCACTGGCTCTTCCCTCAGAGCCACGTC
 CCCGCTGCTTGCATGAACGGTCCGCACCTCTCGGAAGTGTGGGGCAACTACTCCATCGACACGGACTGGTACACC
 GAAACTCCATACACGGGCGTGGTGAGAGAATACTGGTTCCTGGTTCGAGAATACCACCGTAGCACCAGATGTACAC
 GCCCCCTCTGGGTTGCGCCATTGGCAACGCATGGCAGGCTGACTTGGGGTGGCAGGGATATGAGACATGGGCTCTC
 ACGGTAAACCGCTCGATTCTTGGACCAACAATCGAGGCCAACTGGGGCGACGAAGGTAAAGTTGTTTCGACTGGG
 ATTCGCCCGGCCGATAATGCTTACTGCTACGCAGTAATCGTTTCATGTACCAATGGCATGGAGCGGAATGGTACCG
 CAATCCATTTCACGGCCTAAGGCAGCTGGGAGCCCACGAAATGGATGGTGTCCCTGGGGTTACGCAGTGTCTTA
 TAGTACGCCCGTACCCTCTCTCTCTCGCACTGCTGCGCGCAAGCTATTGATACTTCTCAGGCTCCCGGACACTCC
 TACACATACAAGTGGCGAGCTACTCAATACGGAACGGTGAATCCCCCTCGAGGCAGAAATCGGCCATACTTTAAC
 TGACGGGTCTGAAAGAGCTGGTATCATTTCCCACTTCAGCATGCAATACTCGGTTCGGTCTACAGGGCCCCGATCGTC
 ATTCACGGACCTGCTACGGCAGACTATGACGAGGACCTGGGAACGGTTCGTCTTGACGAGCTGGAGTACACCTCT
 CCGTTTCGCCATGTGGTGGTACGCACGTGTGCCTTCCGGACCGCCCTCGCTCTCAAACCTCGCTCATCAACGGCAA

AACGTCTTCTATTGCGACAGTACGACCGATTTCGAGATGCTACGGTAACGGCACACGGTCGGAGTGGCGCTTTGAG
 CAGGGGAAAAAGTATCGAATGAGGCTCATCAATACAGGCCCTTACTCAAACCTCCGTTTGGCATTGACAACCAC
 AACCTCACCGTTATTGCGACCGACTTTGTTCCTATCAAACCCCTACACCACAGATAACGTGAGTCACCAACCCCTT
 GCTTTCGTGCACCTCTAACGAGACATTTCTCCCAGGTGGCAATTTCAATGGGGCAACGCTACGACATCGTCGTTGA
 GGCAAATCAACCGGAAGGTGATTACTGGCTGCGGTAAGGCTCTAATGAATTGATCAGACATTGAAAAAAAGGGG
 GCGGGGGGAACCCCATCAACTGACATTC AACCCCTACCTAGGGCTATATGGCAAACATCCTGCTGCCCCAACGAC
 TACTCAAACAATACCCCTTGGCATAATCCGCTACACAGCAAACCTCCACTGCCGAACCAACACAACAAGTCCTGCA
 CTATCCTACCCGGACACATGCGGCGACGAGCCGGCAGCGAGCCTCGTGCCGCACCTGGCCCTCAACGCGAGCACG
 CCGGCCGTGTCGCGACCTACGACCTGTCCAAAGTACACACTCGAGCTGCCAAAGGGCTTCCCTCTGGACACTGAAC
 GACACCTACCTCTGGATCAACTGGTTCGTGCGCAACGAACCTTGAGGCTGGCCGAGGGCGCGCCGCCGCTGCGGCG
 AGCCTGCCCGCCGAATATCTCGCCGTGACAGCCGGCCGGCAACGAGGGACGCTGGGCGTATCTCGTGTTC AAC
 GACGTCTCCGCCCGCAATCGCTCGCACCCGATGCACCTGCACGGCCACGACTTCTTCTCGCTCGGCACCGGCCCT
 GGCTATTTTGTAGTACGGCAGCAACAGCAGCAGCCTGGCGATGCTCAACCTGCACAACCCGCCCTCGTCGCGACACG
 GCGACCTGGCCCGAGTCTGGCTGGATGGTTCGTGCGCTTCTCATGGACAACCCGGGGAGCTGGCTGATCCACTGC
 CACATCGCCTGGCACTCGAGCGAGTTCGTGCGCCTGCAGTTCTTGGAGAGCCCTGAGACGTATGTCCCTCGATTG
 GAAGGCCAGAGATTGCGGGAGACGTGCGAGGCGTGGGATGCATTCTGGAATCGCCACGACTCATACGAGCAAGAG
 GATGCGGGGATATAAGGGCCGCGAGGACTGCATGAACGCATGCTTTGGTCTGCATGGAGGCTTCGTGAATTGGC
 GGGAAAGATGTGACTAACTTCTTCTCGCATGTTCTTGGCTCTCTCCTTCTCTACACTCTCTAAATAATCCGCAA
 ATTTCCATGTGCGTA

SEQ ID NO : 104
 CHIỀU DÀI : 1815
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 TÊN/KHÓA DẤU HIỆU : CDS
 VỊ TRÍ : (1).....(1815)

atgattggccattctctcgtagctgtagcactcggcagcgcagccatggcactggctctt
 M I G H S L V A V A L G S A A M A L A L
 cctcagagccacgtccccgctgcttgcatgaacgggtccgcactctcggaagtgcggggc
 P Q S H V P A A C M N G P H S R K C W G
 aactactccatcgacacggactggtacaccgaaactccatacacgggctggtgagagaa
 N Y S I D T D W Y T E T P Y T G V V R E
 tactggttctcgtcgagaataaccacgtagcaccagatgtacacgcccctctgggttgc
 Y W F L V E N T T V A P D V H A P L G C
 gccattggcaacgcacatggcagggctgacttgggggtggcagggatatgagacatgggctctc
 A I G N A W Q A D L G W Q G Y E T W A L
 acggtaaaccgctcgattcctggaccaacaatcgaggccaactggggcgacgaagtaatc
 T V N R S I P G P T I E A N W G D E V I
 gttcatgtcaccaatggcatggagcggaatggtaccgcaatccatttccacggcctaagg
 V H V T N G M E R N G T A I H F H G L R
 cagctgggagcccacgaaatggatggtgtccctggggttacgcagtgctcctatagctccc
 Q L G A H E M D G V P G V T Q C P I A P
 ggacactcctacacatacaagtggcgagctactcaatacggaaacgagctggtatcattcc
 G H S Y T Y K W R A T Q Y G T S W Y H S
 cacttcagcatgcaatactcggtcggtctacagggcccgatcgctcattcacggacgtgct
 H F S M Q Y S V G L Q G P I V I H G P A
 acggcgactatgacgaggacctgggaacggctcgtcttgaggactggagtcacacctct
 T A D Y D E D L G T V V L Q D W S H T S
 ccgttcgccatgtggtggtacgcacgtgtgccttccggaccgccctcgctctcaaactcg
 P F A M W W Y A R V P S G P P S L S N S
 ctcacacggaacacgtcttctatttgcgacagtagcaccgattcgagatgctacggg
 L I N G K N V F Y C D S T T D S R C Y G
 aacggcacacggctcggagtggtgctttgagcaggggaaaaagtatcgaatgaggctcatc
 N G T R S E W R F E Q G K K Y R M R L I
 aatacaggcctctactcaaacttccgttttgcattgacaaccacaacctcaccgttatt
 N T G L Y S N F R F A I D N H N L T V I
 gcgaccgactttgttcctatcaaaccctacaccacagataacgtggcaatttcaatgggg
 A T D F V P I K P Y T T D N V A I S M G
 caacgctacgacatcgctggttgggcaaatcaaccggaaggtgattactggctgcgggct
 Q R Y D I V V E A N Q P E G D Y W L R A
 atatggcaacatcctgctgcccgaacgactactcaacaatacccttggcataatccgc
 I W Q T S C C P N D Y S N N T L G I I R
 tacacagcaaacctcactgccgaaccaaacacaacaagtctgcactatcctacccggac
 Y T A N S T A E P N T T S P A L S Y P D

acatgcgggcgacgagccggcagcgagcctcgtgcccgcacctggccctcaacgcgagcacg
 T C G D E P A A S L V P H L A L N A S T
 ccggccgctcgtgcgacacctacgacctgtccaaagtcacactcgagctgccaaagggcttc
 P A V V R T Y D L S K V T L E L P K G F
 ctctggacactgaacgacacctacctctggatcaactggcgtcgccaacgaacttgagg
 L W T L N D T Y L W I N W S S P T N L R
 ctggccgagggcgccgcccgtgcggcgagcctgcccgcgaatatctcgccgtcgac
 L A E G G A A A A A S L P A E Y L A V D
 agccggggccggcaacgagggacgctggcgctatctcgtgttcaacgacgtctccgcccgc
 S R A G N E G R W A Y L V F N D V S A R
 aatcgtcgcacccgatgcacctgcacggccacgacttcttctcgtcggcaccggccct
 N R S H P M H L H G H D F F L L G T G P
 ggctatcttgagtacggcagcaacagcagcagcctggcgatgctcaacctgcacaaccgg
 G Y F E Y G S N S S S L A M L N L H N P
 cctcgtcgcgacacggcgacctggcccgagtctggctggatggcgtcgcgttcctcatg
 P R R D T A T W P E S G W M V V A F L M
 gacaaccgggggagctggctgatccactgccacatcgccctggcactcgagcgagtcgctc
 D N P G S W L I H C H I A W H S S E S L
 ggccctgcagttcctggagagccctgagacgtatgtccctcgattggaaggccagagattg
 G L Q F L E S P E T Y V P R L E G Q R L
 cgggagacgtgcgagggcgtgggatgcattctggaatcgccacgactcatacgagcaagag
 R E T C E A W D A F W N R H D S Y E Q E
 gatgcggggatataa
 D A G I -

SEQ ID NO : 105

CHIỀU DÀI : 604

LOẠI : PRT

SINH VẬT : *M. phaseolina*

MIGHSLVAVALGSAAMALALPQSHVPAACMNGPHSRKCGWNYSIDTDWYTETPYTGTVVREYWFLENTTVAPDVH
 APLGCAIGNAWQADLGWQGYETWALTVNRSIPGPTIEANWGVDEVIVHTNGMERNGTAIHFHGLRQLGAHEMDGV
 PGVTQCPIAPGHSYTYKWRATQYGTSWYHSHFSMQYSVGLQGPIVIHGPATADYDEDLGTVVLDWSHTSPFAMW
 WYARVPSGPPSLNSLINGKNVFCSTTDSRCYGNTRSEWRFEQGGKRYMRLINTGLYSNFRFAIDNHNLTVI
 ATDFVPIKPYTTDNVAISMGQRYDIVVEANQPEGDYWLRAIWQTSCCPNDSNNTLGIIRYTANSTAEPNNTTSPA
 LSYPTCGDEPAASLVPHLALNASTPAVVRTYDLSKVTLELPKGFLLWTLNDTYLWINWSSPTNLRLAEGGAAAAA
 SLPAEYLAVDSRAGNEGRWAYLVFNDVSARNRSHPMHLHGHDFLLGTGPGYFEYGSNSSSLAMLNLHNPPRRDT
 ATWPESGWMVVAFLMDNPGSWLIHCHIAWHSSSLGLQFLESPEYVPRLEGQRLRETCEAWDAFWNRHDSYEQE
 DAGI*

SEQ ID NO : 106

CHIỀU DÀI : 2275 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

TCGATGCATCCCTCTCTGACGCAGATCTCTTTTATAAAACCGACTGTCATCCTCGATTGAGTCAGTCTTTGCTCA
 AACCGACTTATTCAAAGGTGAGTTGGGTGAGTTGGGAGGCTGATGGATTACAGATACTCTCCGGGCACGTTCCGC
 ATGATTCGCCTTTTTCTCGTGCTCGGCTTTGTGGCCACGACATTGGCCAAGACTGTTACTTTTAACTGGAACATC
 GGCTGGGTTTTCTGCGGCTCCAGATGGATTTACACGGCCCGTCATCGGCATCAACGGGCTGTGGCCGCTCCCGTC
 TTAGAAGCCGACGTTAACGACACCATTATCGTAACCGCGCATAACTCGCTAGGCAATGAGACGACGAGCTTGCAC
 TGGCATGGCATGTGGCAGAACAACCTGACTCATATAGACGGCGGAAGTAGAGTCTCGCAGTGCGAAATCCCTCTA
 GGGGGCACCTTTACGTACAGGTTTAAGGCATACCCGGCAGGCACTTTTTGGTACCATTCTCACGCTATGGGCCAA
 TATCCCGACGGCCTACGGGCCCCGATGATTATTACGACCCCTGATTCTGCTGCGGAGCAGGACGCCGATGAGCAG
 CATATACTTACGGTCTCGGACTGGTACCCTAACGATGCCGCCGCTTATCCACCGCTATCTAACTACTCCTAAC
 TATAACGGCGCCATGCCGAAACCAACTCGAGCTTAGTTAACGACCAGCAGTCCAAGAGGCTAAACATCCGCCCG
 GGGCAGAAGAGCTATATCCGCATTATTAATATATCAGCGCTAGCAACGTTCTATCTACAGTTCGGTAGGTACATC
 CCCCTTGTTGTACAGACAAAGCTAACGATTTAGATCAACACAACATGACTGTGCTTTCTATCGACGATGTTAACG
 GTTGGAATCCTGTTACTTATGCTACGCTAAGATGCATGTGCGCATATAAGCCGTAGTCGTATGCAAACACGTAA
 GATTAAAGTGCTAACGACAGGTAGTCGAGCCCCAGAGTTGGGAGGCCCTAGAGATTACCCCCGGACAACGGTACG
 ACGTTATCATCACCGGTCTAGAGAACCCCCAAAGAACTATGCATTTATCAATAAGATGGCCGTTCTCGGCTTGC
 AGAACAACAACATCCTCAGTTATAATTCTGCTCTGGCCTGACCCAGAGCCATTGGCCGTGAGTAGGTTCAACCTCG
 GAAGCGATATCAATCTAACTCCGCTTGACCACGAGCCACTGCTGGAGGCCGTGGACAAAACCTTCACCATGGAG
 TCAACAACCTCAACATCGACGGCGTAGGCTACCGGTGAGATTGCCCCGCGGCCCTTAATCCGATCCCGCTTGCT
 AACTTCGATACACGATCAGCAAGGCCGTACCCCTACATTACCCCGCGCACACCCACTCTGTACACTGCCCT
 AACCACCGCTTTAATGCTACCAACCCCGCCATCTACGGCCAGACCAACTCTTACATCGTAGAAGCCGGCGATAT
 CGTCCAGCTCGTTGTCAACAGCAACGACCTTGTACAACTAACACCTCCGGCCGCGGGCACCCCATGCACTTGCA

CGGCCACACCTTCCAAGTCGTGGGCCAATACGGCACCCACTGGGACGGCGACACCGCAAAATTCCCTACC GTTCC
 AATGAAGCGGGACACGACCGTTCTCTTCGCCGGCGGGAGCTTGGTCATTTCGGTTCAGGCGAACAATCCTGGTGT
 CTGGATGTGTACGGCTCTCCTCCACTTTGCACGAAAACCCAGTCCCCCGGCTGCTTGAAAAGCTTTTTGAACTG
 ACCTTGTGAGTCGCGCAGTCCACTGCCATATCGAATGGCATCTCGACGCCGGCATGGCCGCCACAATCATCGAAG
 CGCCGCTCGAGTTCAGCGAAGCGGTCTGCGGATCCCGCCACAGCACCTCGCGGCGTGGCGGGCATTAAACTTAA
 CGACCCGGGGCAATTGTGCCGGCAACACCGTTAACCTGGAGGATACGGCTGCGTGCGAGAATCTACGACACTGATC
 CTTGGGGGTGAGTTGCTCTGATTTTCGCGTTAACGCAGCCTTTGCTGACGACTACCGCTTTCTGCACTGCGCTTAT
 CGGGGAACGTGAGACAGCACGTTGAATAACACGCTGGCGTAGCACCGTATTTTGTATAGACTACTTTTCCATGTT
 AAATTTTCTGTATACAGTTCGAAATAGATTCAATTTAGGGACAAATACCAGAAACAGGCTCATCCGCAACTCATG
 TGCACCTGCGTAGATCGTTATGCT

SEQ ID NO : 107
 CHIỀU DÀI : 1632
 LOẠI : ADN
 SINH VẬT : *M. phaseolina*
 TÊN/KHÓA DẤU HIỆU : CDS
 VỊ TRÍ : (1).....(1632)

atgattcgcctttttctcgtgctcggctttgtggccacgacattggccaagactgttact
 M I R L F L V L G F V A T T L A K T V T
 ttttaactggaacatcggctgggtttctcggctccagatggatttacacggcccgtcatc
 F N W N I G W V S A A P D G F T R P V I
 ggcatcaacgggctgtggccgcctcccgctcttagaagccgacgttaacgacaccattatc
 G I N G L W P P P V L E A D V N D T I I
 gtaaccgcgcataactcgctaggcaatgagacgacgagcttgactggcatggcatgtgg
 V T A H N S L G N E T T S L H W H G M W
 cagaacaactcgactcatatagacggcggaagtagagtctcgcagtgcgaaatccctcta
 Q N N S T H I D G G S R V S Q C E I P L
 gggggcacctttacgtacaggtttaaggcatacccggcaggcactttttggtaccattct
 G G T F T Y R F K A Y P A G T F W Y H S
 cacgctatgggccaatatcccgacggcctacggggcccgatgattattcacgaccctgat
 H A M G Q Y P D G L R A P M I I H D P D
 tctgctgcgagcaggacgccgatgagcagcatatacttacgggtctcggactggtagcgt
 S A A E Q D A D E Q H I L T V S D W Y R
 aaccagatgccgcccgttatccaccgctatctaactactcctaactataacggcgccatg
 N Q M P P L I H R Y L T T P N Y N G A M
 ccgaaacccaactcgagcttagttaacgaccagcagtcgaagaggctaaacatccgcccg
 P K P N S S L V N D Q Q S K R L N I R P
 gggcagaagagctatatccgcattattaatatatcagcgctagcaacgttctatctacag
 G Q K S Y I R I I N I S A L A T F Y L Q
 ttcgatcaacacaacatgactgtcgtttctatcgacgatgttaacgtcgagccccagagt
 F D Q H N M T V V S I D D V N V E P Q S
 tgggaggccctagagattacccccggacaacgggtacgacgttatcatcaccgggtctagag
 W E A L E I T P G Q R Y D V I I T G L E
 aacccccaaagaaactatgcattttatcaataagatggccgttctcggcttgacagaaca
 N P Q R N Y A F I N K M A V L G L Q N
 aacatcctcagttataattcgtcctggcctgacccagagccattggccgtgagtaggttc
 N I L S Y N S S W P D P E P L A V S R F
 aacctcggaagcgatatcaatctaactccgcttgaccacgagccactgctggagcccgtg
 N L G S D I N L T P L D H E P L L E P V
 gacaaaaccttcacatggaggtcaacaacctaactcgacggcgtaggctaccgcac
 D K T F T M E V N N L N I D G V G Y R I
 acgcaaggcccgtagccctacattaccccgcgacacccactctgtacactgccttaacc
 T Q G P Y P Y I T P R T P T L Y T A L T
 accggctttaatgctaccaacccccgcatctacggccagaccaactcttacatcgtagaa
 T G F N A T N P A I Y G Q T N S Y I V E
 gccggcgatatcgtccagctcgttgtcaacagcaacgaccttgtcacaactaacacctcc
 A G D I V Q L V V N S N D L V T T N T S
 ggccgcgggcaccccatgcacttgacggccacaccttccaagtcgtgggccaatacggc
 G R G H P M H L H G H T F Q V V G Q Y G
 acccactgggacggcgacaccgcaaaattccctaccgttccaatgaagcgggacacgacc
 T H W D G D T A K F P T V P M K R D T T
 gttctcttcgcccggggagcttggtcattcgggttcaggcgcaacaatcctgggtgtctgg
 V L F A G G S L V I R F Q A N N P G V W
 atgttcactgccatatcgaatggcatctcgacgcggcatggccgccacaatcatcgaa

M F H C H I E W H L D A G M A A T I I E
 gcgccgctcgagttccagcgaagcggtctgcggatcccgccacagcacctcgcggtgc
 A P L E F Q R S G L R I P P Q H L A A C
 cgggcattaaacttaacgacccggggcaattgtgccggcaacaccgttaacctggagat
 R A L N L T T R G N C A G N T V N L E D
 acggctgcgtgcagaatctacgacactgatccttggggtgcgcttatcggggaacgtgag
 T A A C R I Y D T D P W G A L I G E R E
 acagcacgttga
 T A R -

SEQ ID NO : 108

CHIỀU DÀI : 543

LOẠI : PRT

SINH VẬT : *M. phaseolina*

MIRLFLVLGFVATTLAKTVTFNWNIGWVSAAPDGFTRPVGINGLWPPPVLADVNDTIIVTAHNSLGNETTSLH
 WHGMWQNNSTHIDGGSRVSQCEIPLGGTFTYRFKAYPAGTFWYHSHAMGQYPDGLRAPMI IHDPDSAAEQDADEQ
 HILTVSDWYRNQMPPLIHRYLTTPNYNGAMPKPNSSLVNDQQSKRLNIRPGQKSYIRIINISALATFYLQFDQHN
 MTVVSIDDVNEPQSWEALEITPGQRYDVIITGLENPQRNYAFINKMAVLGLQNNNLSYNSSWPDPEPLAVSRF
 NLGSDINLTPLDHEPLLEPVDKTFTEVNNLNIDGVGYRITQGPYPYITPRTPTLYTALTTFGNATNPAYGQTN
 SYIVEAGDIVQLVNSNDLVTNTSGRGHPMHLHGHTFQVVGQYGTHTWDGDTAKFPPTVPMKRDTTVLFAAGSLVI
 RFQANNPGVWMFHCHIEWHLDAAGMAATII EAPLEFQRSGLRIPPQHLAACRALNLTTRGNCAGNTVNLEDTAACR
 IYDTPW GALIGERETAR*

SEQ ID NO : 109

CHIỀU DÀI : 2052 (bao gồm 150 bp 5' UTR và 150 bp 3' UTR)

LOẠI : ADN

SINH VẬT : *M. phaseolina*

GCAGCAACTACATCCGGGTAGGCATCAAACGCAAAGATGGTACGGAGGGAGCACTGATGCGCTTTTATAGCTCAT
 ATCAGCACAGCTGCGTGTAGTTGAATGGGACCGACCGAGTTTTTTTGTAGCCTCAATCTTTCACGCTCCGAACC
 ATGTCTCAGATTCTCAGAAGCATGGCGCTGTGCTCCGTAGGCGCAGCATACCACAGCAACCAGAAGAAGAAACC
 CCGCCTTCAACAACCCGGAAGTCATCGGAAAATCGACATAGCGTTCTATCATGGTTTCTCTTTGTCTTTTCATGC
 GCCATCTTCGCGACGTTTATTTCTTATCTCAATAGCGCAACGGCATAACCAAACCTGCGGGGTCTTATTACACAATA
 ACTGGCCTCAAGGCTTTTCTGTCCCATGGGAACCTCTGACAACAATTTTCGACAGTCATCCTGGCAAGGGCCCATAC
 GCGGATCTCTGGGCCAGAACTTGCATCCTCGAGAGCAGTGGTCCGTGCTCCTAGTGTGAGGCACTACAGTTGG
 AAGGTTACCAAGGCTTTTCGATACCCAGATGGGGTGAAAAGGCTGTTTATCTGATTAACGACGGCTTTCTGGGA
 CCCACAGTCGAAGCTCGTTCTGGCGATAGACTGGTGATTGAGGTCCAAAATGCGTTGGAAGACGAAGGTCTCTCC
 TTCCACTGGCACGGTCTTTTAATGAGAGGTGCCAACTACATGGACGGTGCCGTGCGATTTACGCAAGACGCCATT
 CACCCGGGCGCCAACCTCACGTATGAGTTTCGATATCGCGGATGACCAAGCCGGCACATTTTGGTATCACGCTCAT
 GACCAAGTGCAGCGGGCGGATGGCCTGTTCCGAGGACTGATCATCCATCGCCCGGAAACCGCAACTGGAGTCCGC
 GATTTGGATAGATATGGGTACGATGAAGAGAGATTGCTGCTTATCGGACACTGGTACCATCGTTCCGCACAAGAT
 GTCCTAGCGTGGTACATCAGTGTGGTCCCTTTGGAAATGAGCCTGTGCCGATTCACTCCTCATCAATGGAATG
 GGAGCATTCATTTGCTCAAAAGCTATTCCTGCGAGACCTGTGGAATGCATTAACCTTTGAAGGAACGCCACACCT
 AATCTACAGTTCAACTTCACCAGACGTCACCGGCTGAGACTCGTCAACACTGGTACATTGGCTGGATTACCTTG
 AGCATTCGGGTGCGGCCATGCAGGTCATCGAAGTTGACGGCGGCAATGCCGTTACCAGCGACTCTGAGAACGAC
 ACTTCAGTAGGGAGCTTATATCCAGGCCAACGTGCCGACCTAATCCTCTCTTGGCCAGAGGATACCTCTAGAAGCC
 TCAAAAATTTCAATCACCCCTTGACGGGGAGGACTTCAAGTACCCCAATCCAGCCCTGACTCGCACTCAGCACTTC
 TCCATATTTGCTCTGCCCCGTGCCGAAGGAAAAGAGCGAAGCCTCTGCCTCTTCGGAACAAGGCAGGCCAGAA
 AGCCACGCACCACAAACCCCTCATCGACCTCAATGCCCTTGTGAGCGCAGATATCATCACCCCATCACTGCCCCCT
 GCCACTGAACACACTCTCGTTCTATACGCCAACACCCCTCAAACCTCTCCACCGCGGCAACAACCCACGGCTAC
 ATGAACCAAACAGCTGGTCGCCGCAATCCTCTCCGCCCCGTCCGCTCATCGCGCTGCCGCGCTCCTCTTGGGAC
 GCCAACCAGTTCTGTCGCCGCGCATCCCTCTTCCCAACGGCACCTCAGACGCGCCCTGGGTACCATCGTCTCAAC
 AACCTGGACGACGGCTCGCACCCCTTCCACCTACATGGCCACGGCTTCTGGGTGCTCCAAACCCACGCCGAGGG
 TGGGGCTGGGGGTCTGGAATCCATGAGCGGAGGAGCAGCCGCCGGGCGGCCCACTGGAGCTGCAGCGCGCGGTG
 ACGAGGGATACGGTCATGGTGCCGCGGAGGGGGTATGCGGTGCTGCGATTTAGAGCGGACAATGAGGGGCTCTGG
 ATGCTGCACTGCCACAATTTGTGGCAT

SEQ ID NO : 110

CHIỀU DÀI : 1752

LOẠI : ADN

SINH VẬT : *M. phaseolina*

TÊN/KHÓA DẤU HIỆU : CDS

Vị TRÍ : (1).....(1752)

atgtcctcagattctcagaagcatggcgctgtgctccgtaggcgagcataccacagcaa

M S S D S Q K H G A V L R R R S I P Q Q
 ccagaagaagaaaacccgccttcaacaacccggaagtcacgcgaaaatcgacatagcggtt
 P E E E T P P S T T R K S S E N R H S V
 ctatcatgggtttctctttgtcctttcatgcgccatcttcgcgcagctttatttcctatctc
 L S W F L F V L S C A I F A T F I S Y L
 aatagcgcaacggcataccaaaactgcgggggtcttattacacaataactggcctcaaggct
 N S A T A Y Q T A G S Y Y T I T G L K A
 tttctgtcccactgtgaaactctgacaacaatttcgacagtcacccctggcaagggcccatat
 F L S H G N S D N N F D S H P G K G P Y
 ggcggtatctctggggccagaacttgcatcctcgagagcacgtgggtccgtgctcctagtgtc
 G G S L G Q N L H P R E H V V R A P S V
 aggactacagttggaaggttaccaaggcttttcgatacccagatgggggtgaaaaaggct
 R H Y S W K V T K A F R Y P D G V K K A
 gtttatctgattaacgacggctttctgggacccacagtcgaagctcgttctggcgataga
 V Y L I N D G F L G P T V E A R S G D R
 ctgggtgattgaggtccaaaatgcgttggaagacgaaggtctctccttccactggcacgggt
 L V I E V Q N A L E D E G L S F H W H G
 cttttaatgagaggtgccaactacatggacgggtgccgtcggatttacgcaagacgccatt
 L L M R G A N Y M D G A V G F T Q D A I
 caccggggcgccaacttcacgtatgagttcgatatcgcggtatgaccaagccgggcacattt
 H P G A N F T Y E F D I A D D Q A G T F
 tgggtatcacgctcatgaccaagtgcagcggggcggtatggcctgttcggaggactgatcatc
 W Y H A H D Q V Q R A D G L F G G L I I
 catcgcccggaacccgcaactggagtcgcccgttttgatagatatgggtacgatgaagag
 H R P E T A T G V A D L D R Y G Y D E E
 agattgctgcttactcgacactgggtaccatcggtccgcacaaagatgtcctagcgtggtag
 R L L L I G H W Y H R S A Q D V L A W Y
 atcagtgcgtgggtcctttggaaatgagcctgtgccggattcactcctcatcaatggaatg
 I S A G S F G N E P V P D S L L I N G M
 ggagcattcaattgctcaaaagctattcctgcgagacctgtggaatgcattaactttgaa
 G A F N C S K A I P A R P V E C I N F E
 ggaactgccacacctaattctacagttcaacttcaccagacgtcaccgggtgagactcgtc
 G T A T P N L Q F N F T R R H R L R L V
 aacactgggtacattgggtggttcccttgagcattccgggtgcggccatgcaggtcatc
 N T G T L A G F T L S I P G A A M Q V I
 gaagttgacggcggaatgccgttaccagcgcactctgagaacgacacttcagttagggagc
 E V D G G N A V T S D S E N D T S V G S
 ttatatccaggccaacgtgccgacctaatectctcttggccagaggataactctagaagcc
 L Y P G Q R A D L I L S W P E D T L E A
 tcaaaaatttcaatcaccttgacggggaggacttcaagtacccaatccagccctgact
 S K I S I T L D G E D F K Y P N P A L T
 cgcactcagcacttctccatatttcgctctgcccctgtcccgaaggaaaagagcgaagcc
 R T Q T S I F R S A P V P K E K S E A
 tctgcctcttcggaacaaggcaggccagaaagccacgcaccacaaaacctcatcgacctc
 S A S S E Q G R P E S H A P Q T L I D L
 aatgcccttgtcagcgcagatatcatcaccccatcactgccccctgccactgaacacact
 N A L V S A D I I T P S L P P A T E H T
 ctcgttctatacgccaacacctcaaactctcccaccgcggaacaaacccccacggctac
 L V L Y A N T L K L S H R G N K P H G Y
 atgaaccaaaccagctgggtcgccgcaatcctctccgccccgtccgctcatcgcgctgccg
 M N Q T S W S P Q S S P P R P L I A L P
 cgctcctcttgggacgccaaccagttcgtcccgcgcacccctcttcccaacggcacctca
 R S S W D A N Q F V P R I P L P N G T S
 gacgcgccctgggtcaccatcgctcctcaacaacctggacgacggctcgaccccccttccac
 D A P W V T I V L N N L D D G S H P F H
 ctacatggccacgcgttctgggtgctccaaacccacgcccaggggtggggctgggggtcg
 L H G H A F W V L Q T H A A G W G W G S
 tggaatccatga
 W N P -

SEQ ID NO : 111
 CHIỀU DÀI : 583
 LOẠI : PRT
 SINH VẬT : M. phaseolina

MSSDSQKHGAVLRRRSIPQQPEEETPPSTTRKSSSENHRHSVLSWFLFVLSCAIFATFISYLSATAYQTAGSYTYI
 TGLKAFLSHGNSDNNFDSHPGKGPYGGSLGQNLHPREHVVRAPSVRHYSWKVTKAFRYPDGVKKAVYLINDGFLG
 PTVEARSGDRLVIEVQNALEDEGLSFHWHGLLMRGANYMDGAVGFTQDAIHPGANFTYEFDIADDQAGTFWYHAH
 DQVQRADGLFGGLIIHRPETATGVADLDYGYDEERLLLIGHWYHRSAQDVLAWYISAGSFGNEPVPDSLLINGM
 GAFNCSKAIPARPEVINFEGTATPNLQFNFTRRHRLRLVNTGTLAGFTLSIPGAAMQVIEVDGGNAVTSSEND
 TSVGSLYPGQRADLILSWPEDTLEASKISITLDGEDFKYPNPALTRTQHFSIFRSAPVPKEKSEASASSEQGRPE
 SHAPQTLIDLNALVSADIITPSLPPATEHTLVLYANTLKLSHRGKPHGYMNQTSWSPQSSPPRPLIALPRSSWD
 ANQFVPRIPLPNGTSDAPWVTIVLNNLDDGSHPFHLHGHAFWVLQTHAAGWGWGSWNP*

Tài liệu tham khảo

Tất cả các bằng sáng chế Hoa Kỳ, đơn sáng chế đã công bố của Hoa Kỳ và đơn PCT đã công bố được viện dẫn trong bản mô tả này chỉ nhằm mục đích tham khảo.

Dạng tương đương

Mặc dù một số phương án của sáng chế đã được mô tả và được minh họa trong bản mô tả này, người có hiểu biết trung bình trong lĩnh vực sẽ dễ dàng nhận thấy nhiều phương tiện và/hoặc cấu trúc khác để thực hiện các chức năng và/hoặc đạt được kết quả và/hoặc một hoặc nhiều ưu điểm được mô tả ở đây, và mỗi biến đổi và/hoặc cải biến này đều thuộc phạm vi của sáng chế. Người có hiểu biết trung bình trong lĩnh vực sẽ nhận ra, hoặc có khả năng xác định bằng cách sử dụng không gì khác hơn các thí nghiệm thông thường, nhiều dạng tương đương đối với các phương án cụ thể của sáng chế được mô tả trong bản mô tả này. Do đó, cần hiểu rằng các phương án nêu trên được thể hiện chỉ để làm ví dụ và, trong phạm vi của các yêu cầu bảo hộ kèm theo và các dạng tương tự của chúng; sáng chế có thể được thực hiện theo cách khác với những gì đã được mô tả và được yêu cầu bảo hộ cụ thể.

Đơn liên quan

Đơn này hưởng quyền ưu tiên của Đơn đăng ký sáng chế tạm thời Hoa Kỳ số 61/683913, nộp ngày 16 tháng 8 năm 2012, nội dung của đơn được đưa vào bằng cách tham chiếu.

YÊU CẦU BẢO HỘ

1. Polynucleotit được phân lập mã hóa cho lignin peroxidaza có chứa trình tự nucleotit có độ tương đồng ít nhất là 80% so với trình tự nucleotit như nêu trong SEQ ID NO. 1, 2, 4, 5, 7 hoặc 8, hoặc hỗn hợp bất kỳ của chúng.
2. Polypeptit lignin peroxidaza được phân lập có chứa trình tự axit amin có độ tương đồng ít nhất là 70% so với trình tự axit amin như nêu trong SEQ ID NO. 3, 6 hoặc 9, hoặc hỗn hợp bất kỳ của chúng.
3. Cấu trúc biểu hiện chứa polynucleotit được phân lập theo điểm 1, trong đó polynucleotit nói trên được liên kết có điều khiển với ít nhất một trình tự điều hòa dịch mã, trình tự điều hòa dịch mã này có vị trí liên kết ribosom, trình tự khởi đầu dịch mã, hoặc trình tự kết thúc dịch mã, hoặc sự kết hợp của chúng, làm tăng cường sự biểu hiện của polypeptit mà có độ tương đồng ít nhất là 70% so với trình tự axit amin như nêu trong SEQ ID NO. 3, 6 hoặc 9, hoặc hỗn hợp bất kỳ của chúng.
4. Cấu trúc gen tái tổ hợp chứa polynucleotit theo điểm 1, trong đó polynucleotit có thể biểu hiện trong tế bào chủ để tạo ra một enzym làm thoái biến lignin.
5. Cấu trúc gen tái tổ hợp theo điểm 4, trong đó cấu trúc này còn bao gồm thêm vùng trình tự khởi đầu được liên kết có điều khiển để tăng cường sự biểu hiện của khuôn polynucleotit.
6. Thể biến nạp chứa các cấu trúc biểu hiện theo điểm 3 hoặc cấu trúc gen tái tổ hợp theo điểm 4, trong đó thể biến nạp nói trên tạo ra enzym làm tăng tốc sự thoái biến lignin.
7. Nấm chuyển gen của *M. phaseolina* có sự thoái biến lignin được tăng cường có chứa cấu trúc biểu hiện theo điểm 3 hoặc cấu trúc gen tái tổ hợp theo điểm 4.

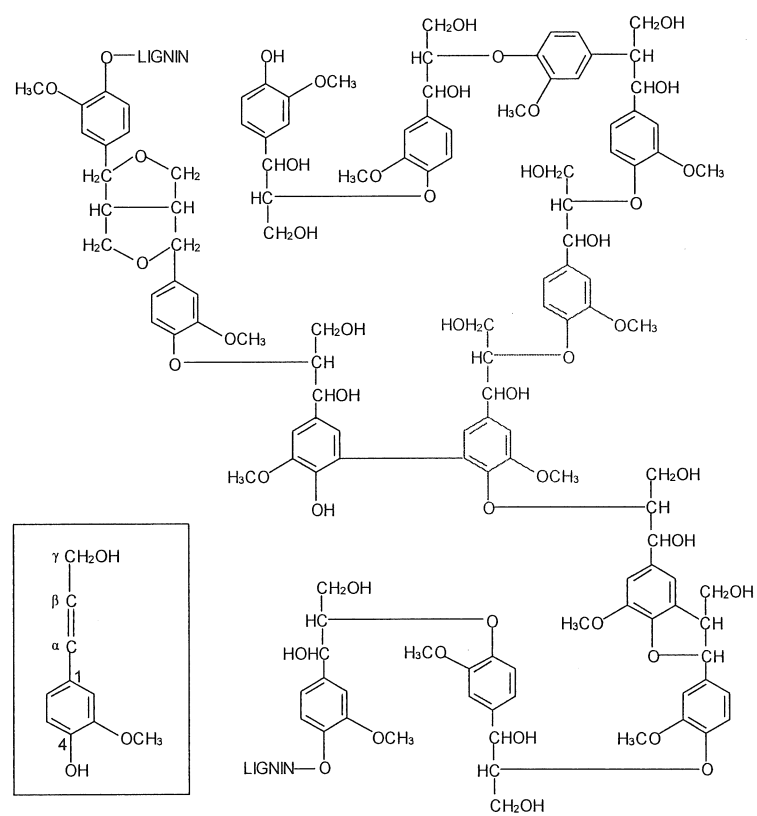


Fig. 1

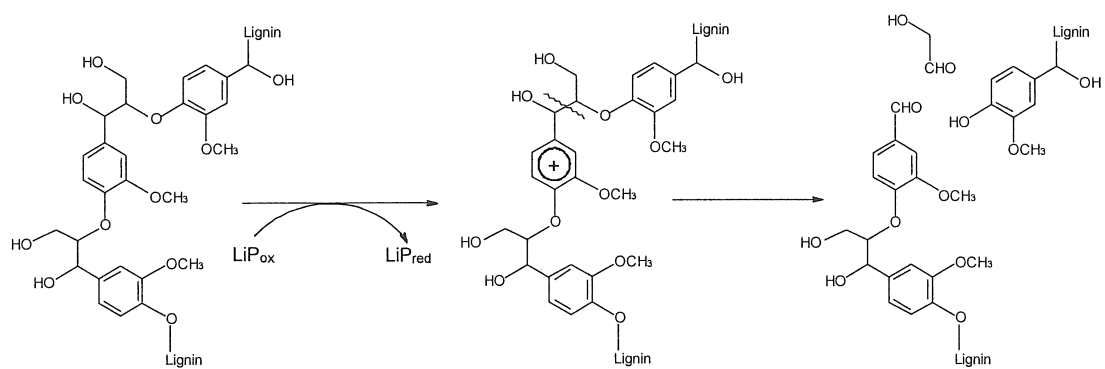


Fig. 2

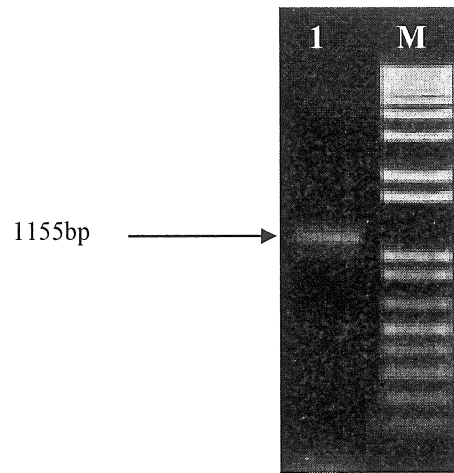


Fig. 3

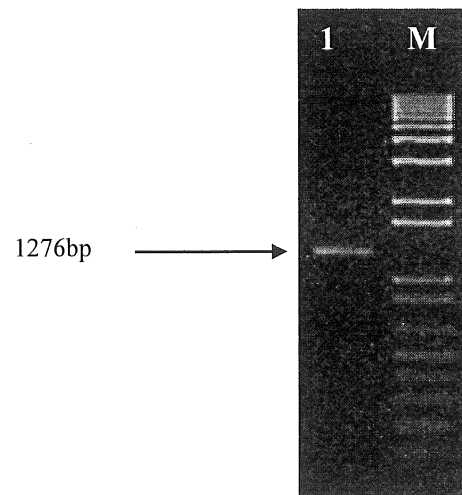


Fig. 4

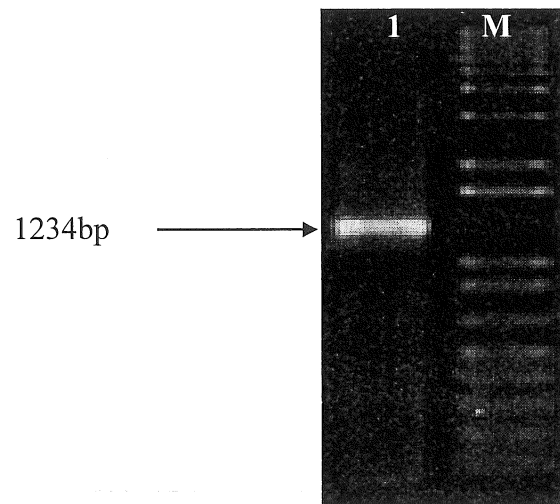


Fig. 5



Fig. 6

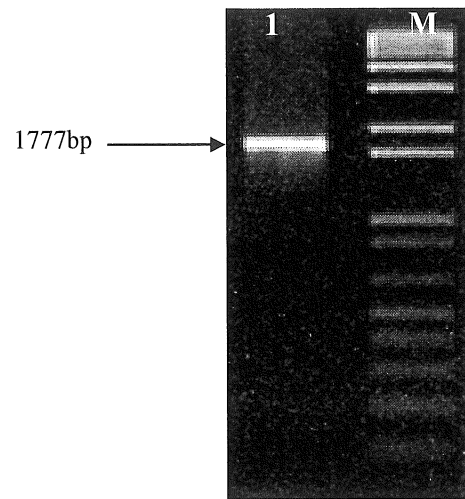


Fig. 7

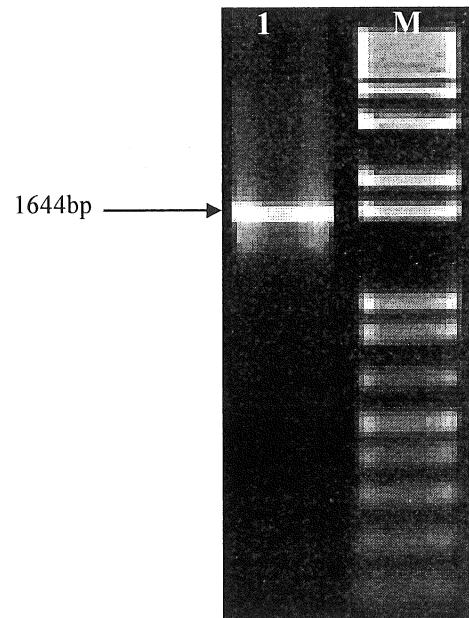


Fig. 8

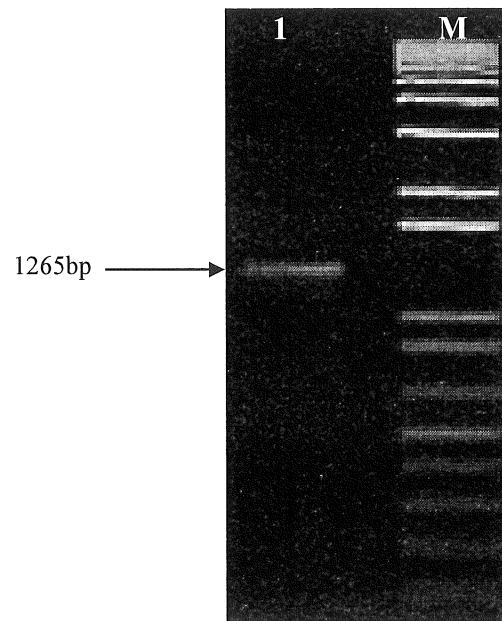


Fig. 9

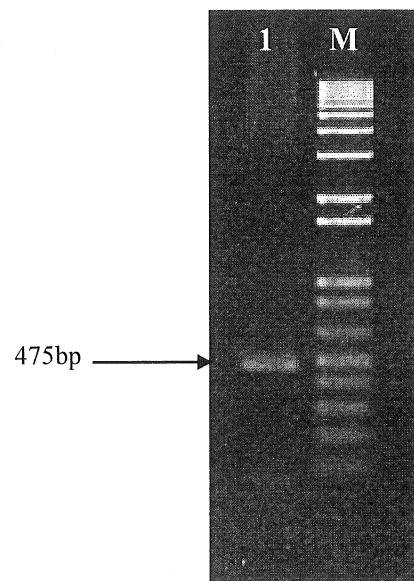


Fig. 10

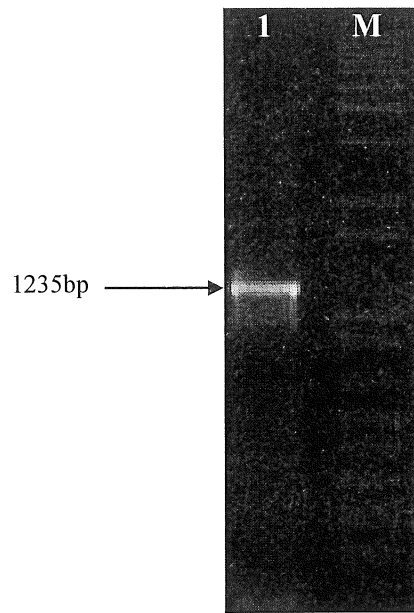


Fig. 11

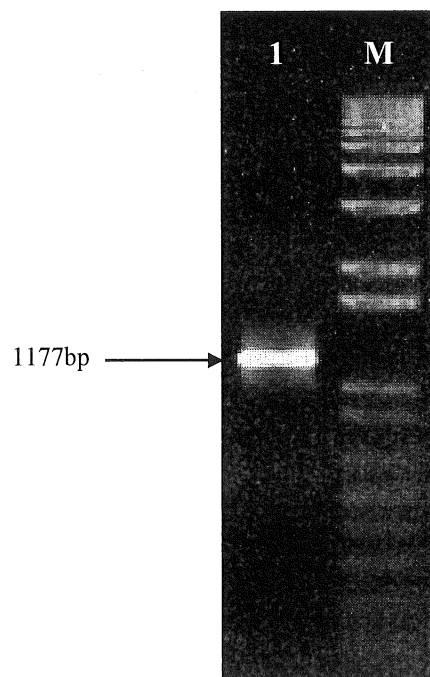


Fig. 12

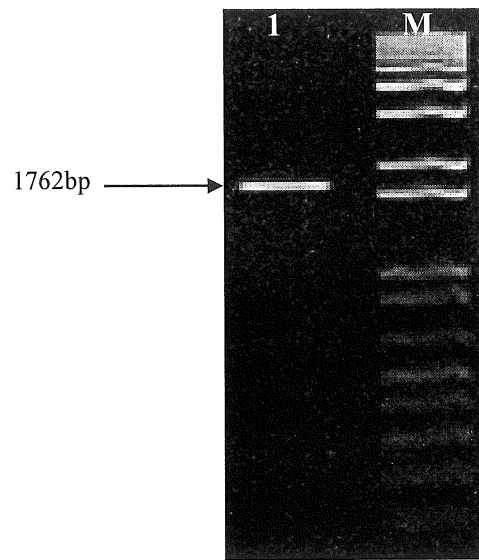


Fig. 13

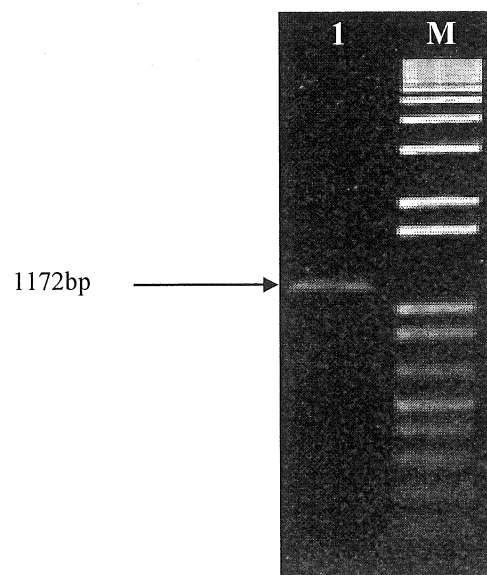


Fig. 14

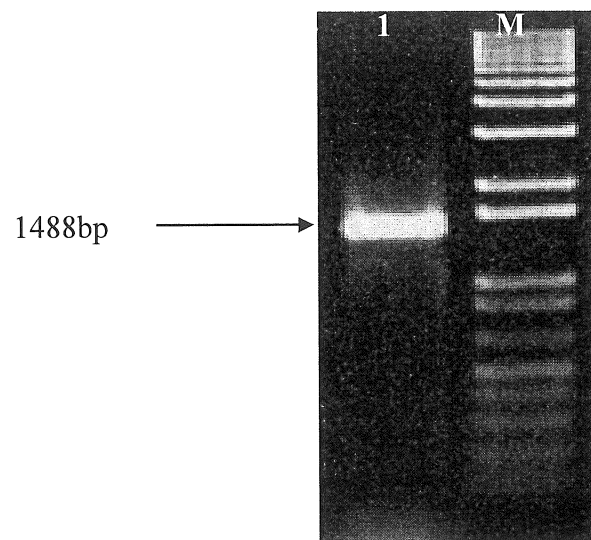


Fig. 15

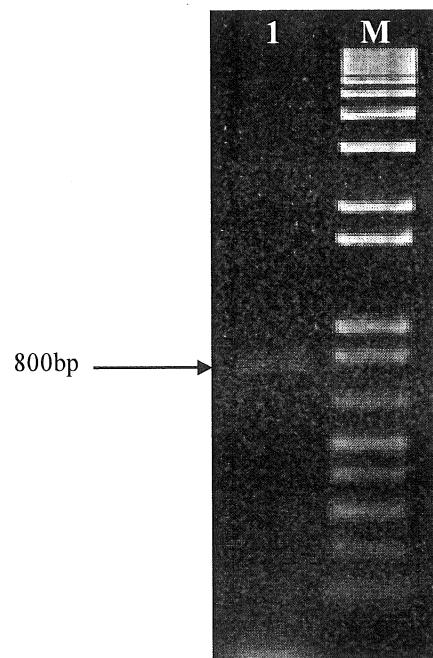


Fig. 16

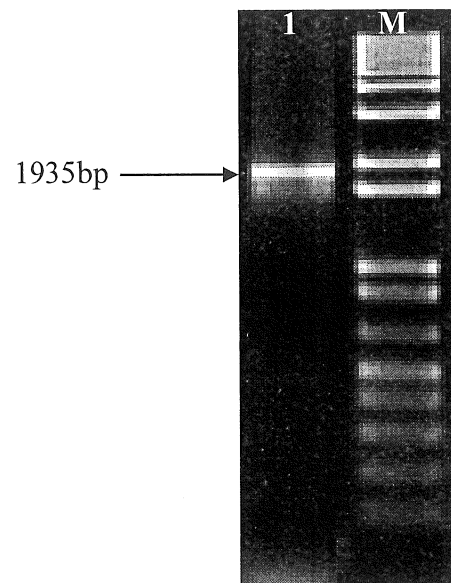


Fig. 17

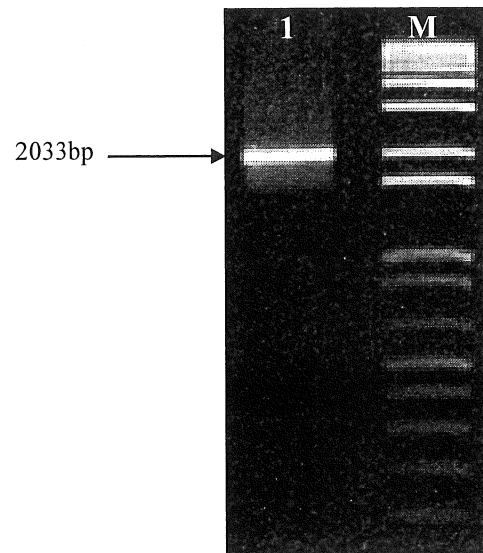


Fig. 18

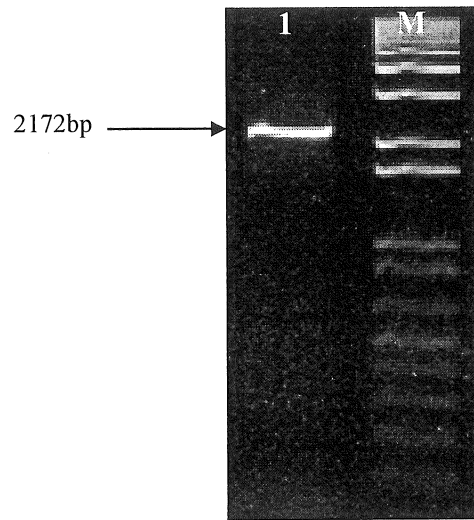


Fig. 19

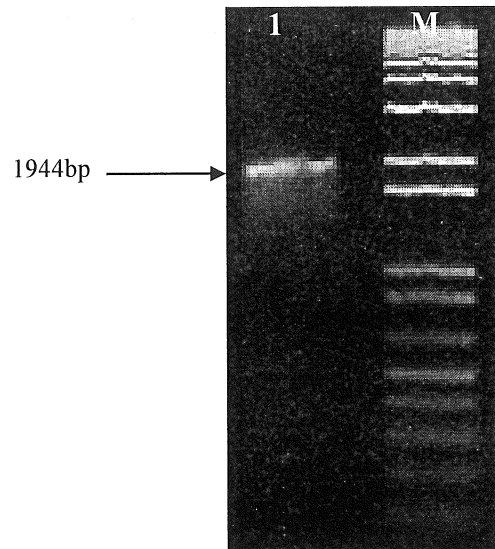


Fig. 20

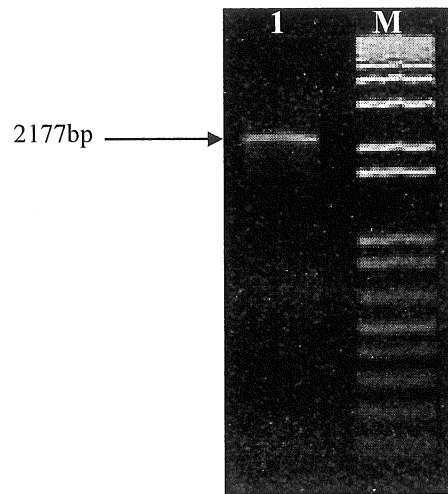


Fig. 21

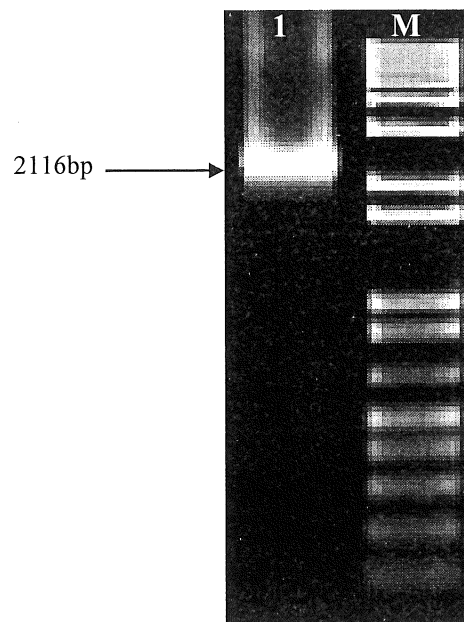


Fig. 22

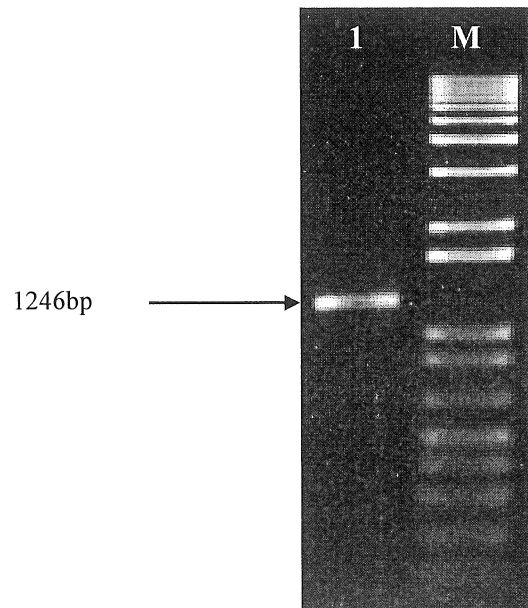


Fig. 23

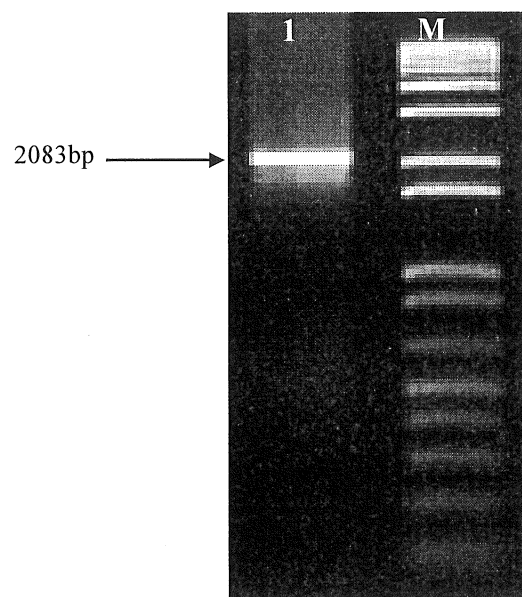


Fig. 24

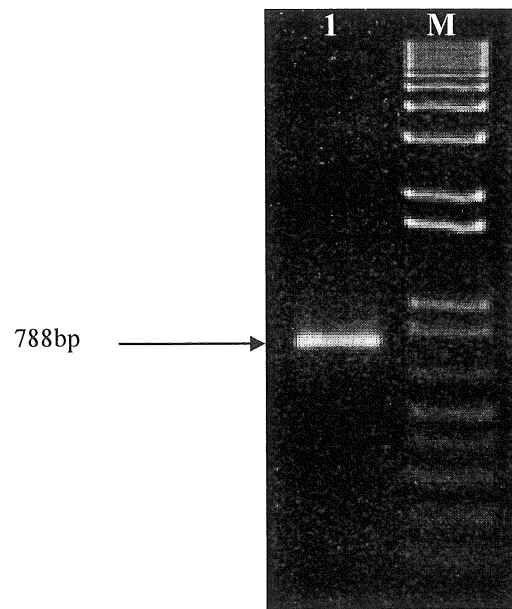


Fig. 25

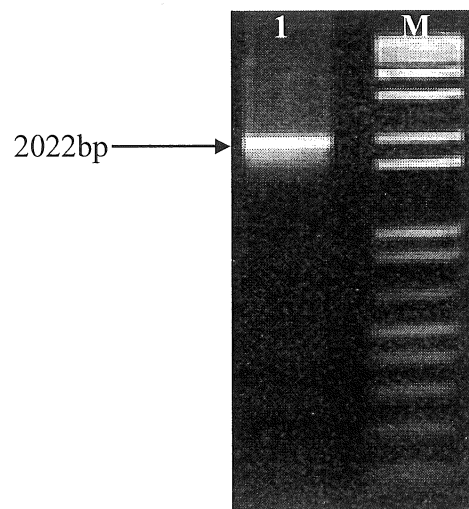


Fig. 26

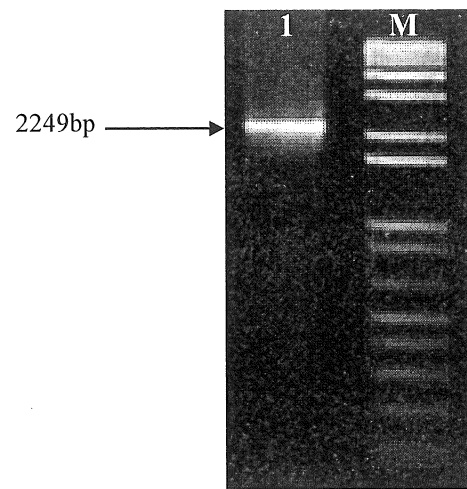


Fig. 27

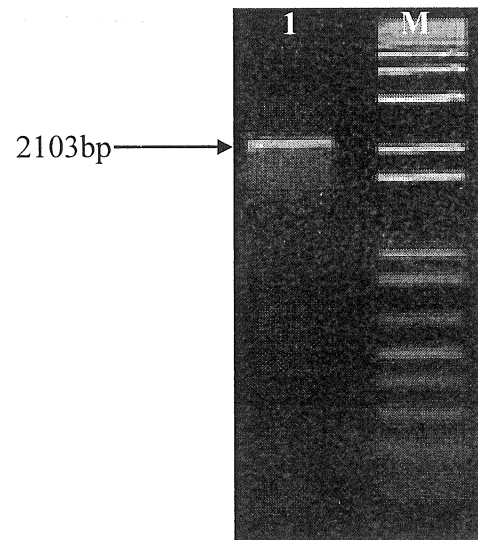


Fig. 28

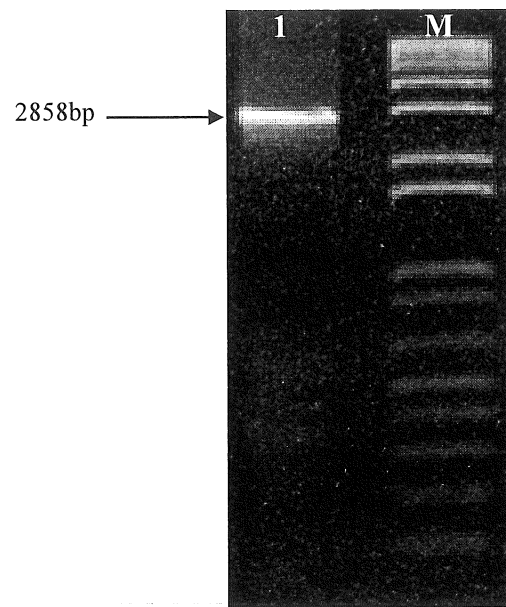


Fig. 29

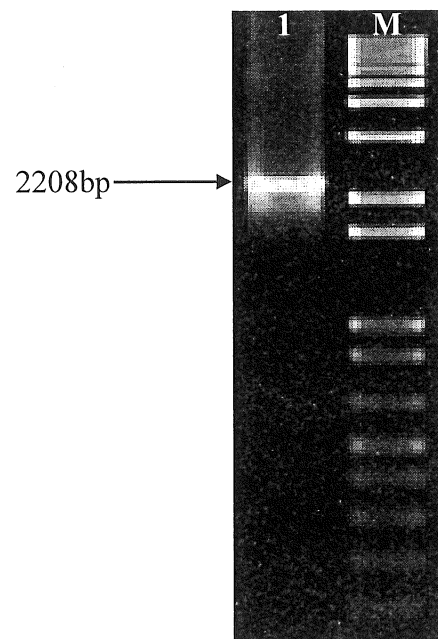


Fig. 30

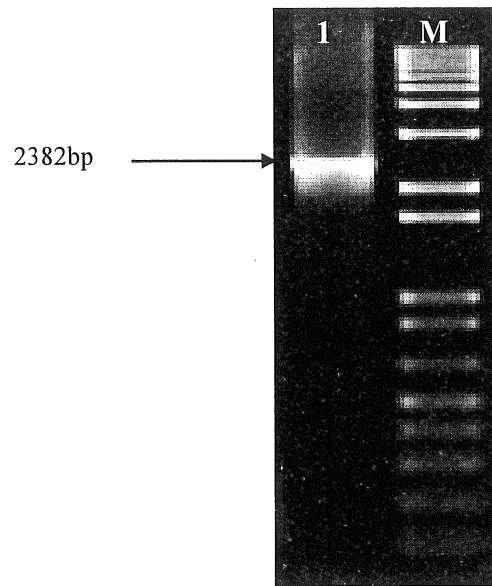


Fig. 31

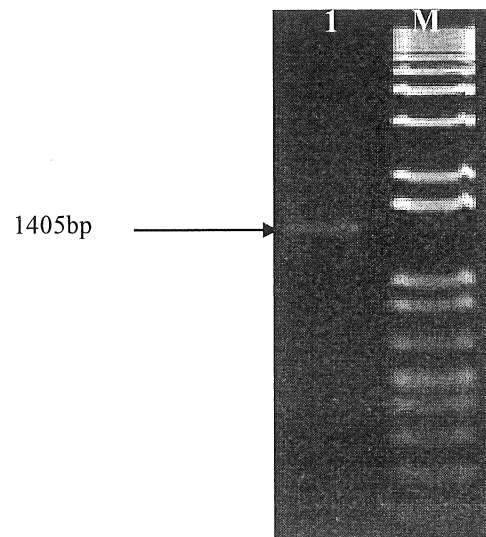


Fig. 32

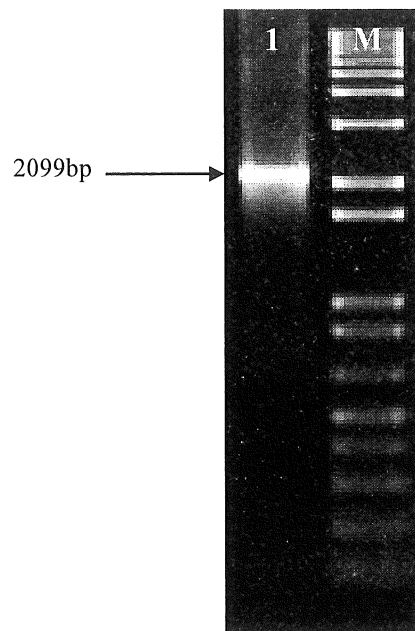


Fig. 33

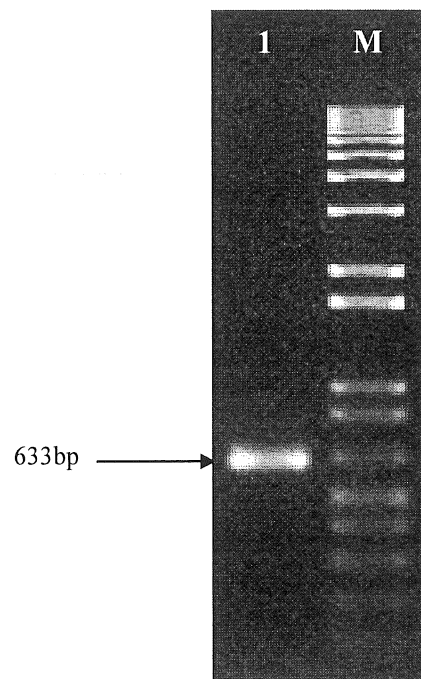


Fig. 34

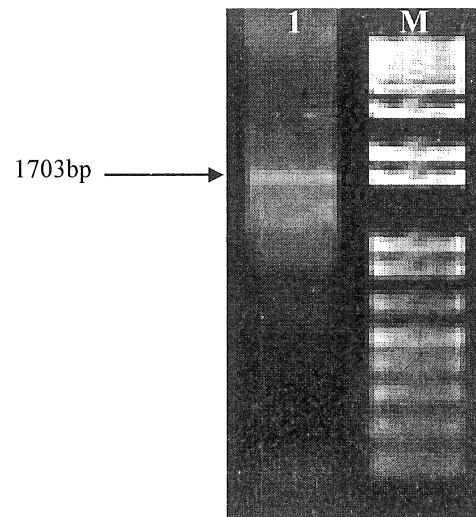


Fig. 35

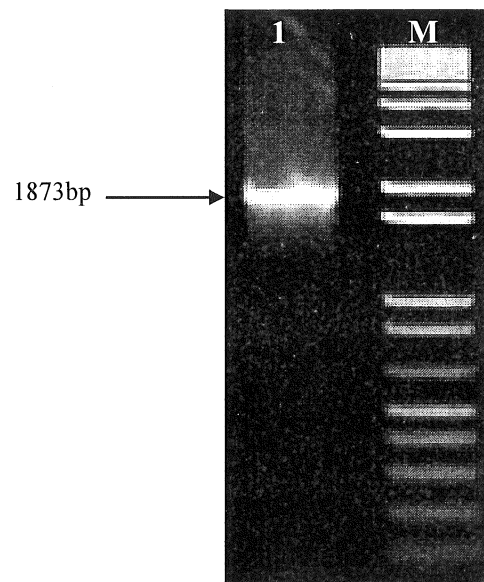


Fig. 36