



(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)



1-0025827

(51)⁷ C07K 14/11; A23K 1/14; A23L 1/30; (13) B
A61K 39/145; A61P 31/16; C12N 7/00;
C07K 16/08; C12N 15/00; C12N 15/44;
C12N 15/82; A01H 5/00; A61P 37/04

(21) 1-2014-01334

(22) 28/09/2012

(86) PCT/CA2012/050681 28/09/2012

(87) WO2013/044390 04/04/2013

(30) 61/541,780 30/09/2011 US

(45) 26/10/2020 391

(43) 25/06/2014 315A

(73) MEDICAGO INC. (CA)

600 - 1020 Route de l'Eglise, Quebec, Québec G1V 3V9, Canada

(72) D'AOUST, Marc-Andre (CA); COUTURE, Manon (CA); VEZINA, Louis-Philippe (CA).

(74) Công ty Luật TNHH T&G (TGVN)

(54) HẠT GIỐNG NHƯ VIRUT VÀ PHƯƠNG PHÁP SẢN XUẤT HẠT GIỐNG NHƯ VIRUT NÀY

(57) Sáng chế đề cập đến hạt giống như virut và phương pháp sản xuất hạt giống như virut này ở thực vật. Phương pháp này bao gồm bước đưa axit nucleic thứ nhất và axit nucleic thứ hai vào thực vật, hoặc phần của thực vật. Axit nucleic thứ nhất bao gồm vùng điều hòa thứ nhất hoạt động ở thực vật và được liên kết theo kiểu điều khiển được với trình tự nucleotit mã hóa protein virut cấu trúc. Axit nucleic thứ hai bao gồm vùng điều hòa thứ hai hoạt động ở thực vật và được liên kết theo kiểu điều khiển được với trình tự nucleotit mã hóa protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton. Thực vật hoặc phần của thực vật được ủ dưới các điều kiện cho phép sự biểu hiện của axit nucleic, theo đó tạo ra VLP.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp sản xuất protein virus ở thực vật. Cụ thể hơn là, sáng chế đề cập đến phương pháp sản xuất và làm tăng sự tạo ra hạt giống như virus ở thực vật.

Tình trạng kỹ thuật của sáng chế

Bệnh cúm do virus ARN thuộc họ orthomyxoviridae gây ra. Có ba loại virus và chúng gây ra ba typ cúm khác nhau: typ A, B và C. Các virus cúm typ A lây nhiễm cho động vật có vú (người, lợn, chồn sương, ngựa) và chim. Typ này rất quan trọng đối với nhân loại do đây là loại virus gây ra các dịch bệnh trên toàn thế giới. Virus cúm typ B (còn được gọi đơn giản là cúm B) chỉ lây nhiễm cho người. Nó thường gây ra sự bùng phát cục bộ bệnh cúm. Các virus gây bệnh cúm C cũng chỉ lây nhiễm cho người. Chúng lây nhiễm cho hầu hết mọi người ở độ tuổi trẻ và hiếm khi gây ốm nghiêm trọng.

Việc tiêm vaccin mang lại sự bảo vệ chống lại bệnh do tác nhân tương tự gây ra bằng cách làm cho đối tượng có sự phòng thủ trước khi lây nhiễm. Thông thường, việc này đạt được thông qua việc sử dụng các dạng sống giảm độc lực hoặc bị bất hoạt hoàn toàn của các tác nhân lây nhiễm làm các chất sinh miễn dịch. Để tránh sự nguy hiểm của virus hoàn chỉnh (như các virus bị giết hoặc được giảm độc lực) làm vaccin, protein virus tái tổ hợp, ví dụ, các tiểu đơn vị, được theo đuổi làm các vaccin. Cả vaccin peptit và tiểu đơn vị đều có những hạn chế tiềm tàng. Các vaccin tiểu đơn vị có thể có tính sinh miễn dịch kém, do việc cuộn gập không chính xác hoặc trình diện kháng nguyên kém. Vấn đề chính là khó đảm bảo rằng cấu dạng của các protein được biến đổi mô phỏng giống cấu dạng của các kháng nguyên trong môi trường tự nhiên của chúng. Các chất phụ trợ thích hợp và, trong trường hợp của peptit, protein chất mang, phải được sử dụng để kích thích đáp ứng miễn dịch. Ngoài ra, các vaccin này chủ yếu tạo ra các đáp ứng dịch thể, và do đó có thể không đạt được tính miễn dịch hữu hiệu. Các vaccin tiểu đơn vị thường không hữu

hiệu với các bệnh mà các virus hoàn chỉnh bị bất hoạt có thể được chứng minh là tạo ra sự bảo vệ.

Hạt giống như virus (virus like particle - VLP) là ứng cử viên tiềm năng để được bao gồm trong các chế phẩm sinh miễn dịch. VLP gần giống các hạt virus (virion) trưởng thành, nhưng chúng không chứa vật liệu di truyền của virus. Do đó, các VLP không có khả năng sao chép trong tự nhiên, mà khiến chúng an toàn để dùng làm vaccine. Ngoài ra, VLP có thể được biến đổi để biểu hiện các glycoprotein virus trên bề mặt của VLP, là cấu hình sinh lý nguyên bản nhất của chúng. Hơn nữa, do VLP tương tự các virion nguyên vẹn và là các cấu trúc dạng hạt đa hóa trị, nên các VLP có thể hữu hiệu trong việc kích thích các kháng thể trung tính thành glycoprotein hơn so với các kháng nguyên protein hòa tan.

Các VLP đã được tạo ra ở thực vật (các công bố đơn sáng chế quốc tế số WO2009/009876; WO 2009/076778; WO 2010/003225; WO 2010/003235; WO 2011/03522; WO 2010/148511), và ở côn trùng và các hệ của động vật có vú (Noad, R. và Roy, P., 2003, *Trends Microbiol* 11: 438-44; Neumann et al., 2000, *J. Virol.*, 74, 547-551). Latham và Galarza (2001, *J. Virol.*, 75, 6154-6165) đã báo cáo rằng sự tạo ra VLP cúm ở các tế bào côn trùng bị lây nhiễm baculovirus tái tổ hợp đồng biểu hiện các gen hemagglutinin (HA), neuraminidase (NA), M1, và M2. Nghiên cứu này chứng minh rằng các protein virion cúm tự lắp ráp nhờ sự đồng biểu hiện ở các tế bào nhân chuẩn và protein nền M1 được yêu cầu để tạo ra VLP. Tuy nhiên, Gomez-Puertas et al., (1999, *J. Gen. Virol.*, 80, 1635-1645) cũng đã cho thấy rằng sự biểu hiện quá mức M2 ngăn chặn hoàn toàn việc chuyển ARN CAT vào môi trường nuôi cấy MDCK.

M2 đóng vai trò là protein kênh ion và đã được chứng minh rằng, khi protein này bị biểu hiện quá mức, sự vận chuyển nội bào HA được đồng biểu hiện bị ức chế và sự tích lũy hemagglutinin (HA) tại màng sinh chất bị giảm từ 75 đến 80% (Sakaguchi et al., 1996; Henkel & Weisz, 1998). Hơn nữa, bằng cách biểu hiện quá mức M2, sự tích lũy protein nền virus tại màng sinh chất bị giảm và do đó giảm mạnh số các VLP có chức năng được tạo ra.

Protein M2 là được biểu hiện chủ yếu tại bề mặt tế bào của các tế bào lây nhiễm cúm A (Lamb et al. (1985) *Cell*, 40, 627 to 633). Protein này cũng được phát hiện ở màng của chính hạt virus, nhưng với lượng nhỏ hơn nhiều, từ 14 đến 68 phân tử M2 trên mỗi virion (Zebedee và Lamb (1988) *J. Virol.* 62, 2762 đến 72). Protein M2 được cải biến sau dịch mã bằng cách bổ sung axit palmitic trên cysteine tại vị trí 50 (Sugrue et al. (1990) *Virology* 179, 51 đến 56).

Protein M2 là homotetramer bao gồm hai dimer được liên kết bằng disulfua, mà được giữ với nhau bằng các tương tác không cộng hóa trị (Sugrue và Hay (1991) *Virology* 180, 617 đến 624). Bằng cách tạo đột biến định hướng điểm, Holsinger và Lamb, (1991) *Virology* 183, 32 đến 43, đã chứng minh rằng các gốc cysteine tại vị trí 17 và 19 liên quan đến sự hình thành cầu disulfua. Chỉ cysteine tại vị trí 17 có mặt ở tất cả các virus được phân tích. Trong các chủng virus mà cysteine 19 cũng có mặt, không xác định được liệu cầu disulfua thứ hai có được tạo thành trong cùng dimer hay không (đã được liên kết bằng Cys 17-Cys 17) hoặc với dimer khác.

Smith et al. (công bố đơn yêu cầu cấp patent Mỹ số 2010/0143393) và Song et al. (*Plos ONE* 2011 6(1):e14538) mô tả các vaccine và VLP bao gồm protein M2 cúm. Các VLP bao gồm ít nhất protein lõi virus như M1. Protein lõi này điều khiển việc nảy chồi (budding) của virus và giải phóng các hạt từ các tế bào vật chủ côn trùng.

Szecsni et al. (*Virology Journal*, 2006, 3:70) đã ghép các VLP trên các hạt lõi bị hỏng khả năng sao chép thu được từ virus gây ung thư bạch cầu ở chuột (murine leukaemia virus - MLV). VLP cúm được biến đổi thu được bằng cách đồng biểu hiện nhất thời các thành phần virus bề mặt tế bào (HA, NA, M2) và các thành phần virus bên trong (hệ gen chỉ thị Gag, GFP) và chứa tại bề mặt của chúng HA, HA và NA hoặc M2, hoặc tất cả ba protein thu được từ virus H7N1 hoặc H5N1. Theo Szecsni et al., việc biểu hiện M2 trong quá trình tạo ra VLP cúm (Flu-VLP) không ảnh hưởng đến sự kết hợp HA hoặc NA lên trên các hạt virus (trang 2, cột bên phải, đoạn thứ hai trong tài liệu của Szecsni et al.).

Bản chất kỹ thuật của sáng chế

Sáng chế đề cập đến việc tạo ra protein virus ở thực vật. Cụ thể hơn là, sáng chế đề xuất việc tạo ra và làm tăng sự tạo ra hạt giống như virus ở thực vật.

Mục đích của sáng chế là đề xuất phương pháp được cải thiện để làm tăng sự tạo ra hạt giống như virus ở thực vật.

Sáng chế đề xuất phương pháp (A) tạo ra hạt giống như virus (VLP) ở thực vật bao gồm các bước:

a) đưa axit nucleic thứ nhất bao gồm vùng điều hòa thứ nhất hoạt động ở thực vật và được liên kết theo kiểu điều khiển được với trình tự nucleotit mã hóa protein virus cấu trúc vào thực vật, hoặc phần của thực vật,

b) đưa axit nucleic thứ hai bao gồm vùng điều hòa thứ hai hoạt động ở thực vật và được liên kết theo kiểu điều khiển được với trình tự nucleotit mã hóa protein kênh

c) ủ thực vật hoặc phần của thực vật trong các điều kiện cho phép biểu hiện axit nucleic, theo đó tạo ra VLP.

Vùng điều hòa thứ nhất hoạt động ở thực vật, và vùng điều hòa thứ hai hoạt động ở thực vật có thể giống hoặc khác nhau.

Protein kênh của phương pháp (A) được mô tả trên đây có thể là protein kênh proton. Protein kênh proton có thể được chọn từ M2 hoặc BM2. Hơn nữa, protein kênh proton có thể bao gồm trình tự tín hiệu kênh proton HXXXW. Protein M2 có thể là protein M2 thu được từ cúm A/Puerto Rico/8/1934 (SEQ ID NO:14) hoặc từ cúm A/New Caledonia/20/1999 (SEQ ID NO:11).

Sáng chế cũng đề xuất phương pháp (A) như được mô tả trên đây, trong đó protein virus cấu trúc bao gồm miền trimer hóa. Hơn nữa, trình tự nucleotit mã hóa protein virus cấu trúc bao gồm trình tự nucleotit khảm mã hóa, theo thứ tự, protein virus kháng nguyên hoặc mảnh của chúng, miền xuyên màng cúm, và đuôi tế bào chất. Protein virus cấu trúc

có thể bao gồm protein HA cúm. Hơn nữa, một hoặc nhiều vòng phân hủy protein của protein HA cúm có thể bị xóa bỏ.

Sáng chế đề xuất phương pháp (A) như được mô tả trên đây trong đó, trình tự nucleotit mã hóa protein virus cấu trúc có thể được chọn từ nhóm bao gồm B HA, C, HA, H2, H3, H4, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, và H16. Ví dụ, trình tự nucleotit mã hóa protein virus cấu trúc có thể là HA hoặc H3 typ B. Trình tự nucleotit mã hóa protein virus cấu trúc có thể là, ví dụ, HA từ cúm B/Brisbane/60/2008, B/Malaysia/2506/2004 hoặc B/Wisconsin/1/2010, hoặc H3 từ cúm A/Perth/16/2009 hoặc A/Victoria/361/2011. Hơn nữa, trình tự nucleotit mã hóa protein virus cấu trúc độ tương đồng trình tự ít nhất là 70% với SEQ ID NO: 23, 28, 43, 46, 51, 57 hoặc 61. Trình tự của protein virus cấu trúc cũng có thể bao gồm trình tự SEQ ID NO:25, 30, 41, 48, 54, 58 hoặc 64.

Sáng chế cũng bao gồm phương pháp (A) như được mô tả trên đây, trong đó trình tự axit nucleic thứ nhất bao gồm vùng điều hòa thứ nhất được liên kết theo kiểu điều khiển được với một hoặc nhiều vùng tăng cường comovirus, trình tự nucleotit mã hóa protein virus cấu trúc, và một hoặc nhiều yếu tố khuếch đại geminivirus, và axit nucleic thứ ba mã hóa replicaza geminivirus được đưa vào thực vật hoặc phần của thực vật. Một hoặc nhiều vùng tăng cường comovirus có thể là UTR của comovirus, ví dụ, UTR virus khảm của cây đậu đũa có khả năng siêu dịch mã (CPMV-HT) như CPMV-HT 5' và/hoặc 3'UTR. Một hoặc nhiều yếu tố khuếch đại geminivirus có thể được chọn từ vùng liên gen dài lùn của virus gây vàng, lùn ở cây đậu (BeYDV LIR), và vùng liên gen ngắn BeYDV (BeYDV SIR). Hơn nữa, trình tự nucleotit mã hóa protein virus cấu trúc có thể là HA Typ B hoặc H3, ví dụ, trình tự nucleotit mã hóa protein virus cấu trúc có thể độ tương đồng trình tự ít nhất là 70% với SEQ ID NO: 23, 28, 43, 46, 51, 57 hoặc 61. Trình tự của protein virus cấu trúc cũng có thể bao gồm trình tự SEQ ID NO:25, 30, 41, 48, 54, 58 hoặc 64.

Phương pháp như được mô tả trên đây (Phương pháp A) cũng có thể bao gồm việc đưa trình tự axit nucleic khác mã hóa gen ức chế bất hoạt, ví dụ, HcPro hoặc p19.

Sáng chế cũng bao gồm phương pháp (A) như được mô tả trên đây, trong đó ở bước đưa vào thực vật (bước a), axit nucleic được biểu hiện nhất thời ở thực vật. Theo cách khác, trong bước đưa vào thực vật (bước a), axit nucleic được biểu hiện ổn định ở thực vật.

Phương pháp (A) như được mô tả trên đây có thể còn bao gồm bước:

d) thu hoạch thực vật và tinh chế các VLP.

Sáng chế cũng bao gồm phương pháp (A) như được mô tả trên đây, trong đó VLP không chứa protein nền virus hoặc protein lõi.

Sáng chế đề xuất VLP được tạo ra bằng phương pháp (A) như được mô tả trên đây. VLP có thể còn bao gồm một hoặc nhiều lipid thu được từ thực vật. VLP cũng có thể đặc trưng ở chỗ không chứa protein kênh. Hơn nữa, protein virus cấu trúc của VLP có thể là protein HA0. Một hoặc nhiều protein virus của VLP có thể bao gồm các N-glycan đặc hiệu thực vật, hoặc các N-glycan được biến đổi. Sáng chế cũng đề xuất kháng thể đa dòng được tạo ra bằng cách sử dụng VLP này.

Sáng chế bao gồm chế phẩm bao gồm liệu hữu hiệu của VLP như vừa được mô tả để kích thích đáp ứng miễn dịch, và chất mang được dùng.

Sáng chế cũng mô tả phương pháp gây miễn dịch đối với sự lây nhiễm virus cúm ở đối tượng, bao gồm dùng VLP như vừa được mô tả cho đối tượng. VLP có thể được dùng cho đối tượng qua đường miệng, qua nội bì, trong mũi, trong cơ, trong màng bụng, trong tĩnh mạch hoặc dưới da.

Sáng chế cũng đề xuất vật liệu thực vật bao gồm VLP được tạo ra bằng phương pháp (A) được mô tả trên đây. Vật liệu thực vật có thể được sử dụng để gây miễn dịch đối với sự lây nhiễm virus cúm ở đối tượng. Vật liệu thực vật cũng có thể được trộn làm chất bổ sung cho thực phẩm.

Sáng chế cũng đề xuất phương pháp (B) sản xuất hạt giống như virus (VLP) bao gồm các bước:

a) tạo ra thực vật hoặc phần của thực vật bao gồm axit nucleic thứ nhất bao gồm vùng điều hòa thứ nhất hoạt động ở thực vật và được liên kết theo kiểu điều khiển được với trình tự nucleotit mã hóa protein virus cấu trúc vào thực vật, hoặc phần của thực vật, và axit nucleic thứ hai bao gồm vùng điều hòa thứ hai hoạt động ở thực vật và được liên kết theo kiểu điều khiển được với trình tự nucleotit mã hóa protein kênh

b) ủ thực vật hoặc phần của thực vật trong các điều kiện cho phép biểu hiện axit nucleic, theo đó tạo ra VLP.

Vùng điều hòa thứ nhất hoạt động ở thực vật, và vùng điều hòa thứ hai hoạt động ở thực vật có thể giống hoặc khác nhau.

Protein kênh của phương pháp (B) được mô tả trên đây có thể là protein kênh proton. Protein kênh proton có thể được chọn từ M2 hoặc BM2. Hơn nữa, protein kênh proton có thể bao gồm trình tự tín hiệu kênh proton HXXXW.

Sáng chế cũng đề xuất phương pháp (B) như được mô tả trên đây, trong đó protein virus cấu trúc bao gồm miền trimer hóa. Hơn nữa, trình tự nucleotit mã hóa protein virus cấu trúc bao gồm trình tự nucleotit thể khảm mã hóa, theo thứ tự, protein virus của kháng nguyên hoặc mảnh của nó, miền xuyên màng cúm, và đuôi tế bào chất. Protein virus cấu trúc có thể bao gồm protein HA cúm. Hơn nữa, một hoặc nhiều vòng phân hủy protein của protein HA cúm có thể bị xóa bỏ.

Sáng chế đề xuất phương pháp (B) như được mô tả trên đây trong đó, trình tự nucleotit mã hóa protein virus cấu trúc có thể được chọn từ nhóm gồm có B HA, C, HA, H2, H3, H4, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, và H16. Ví dụ, trình tự nucleotit mã hóa protein virus cấu trúc có thể là HA hoặc H3 Typ B. Trình tự nucleotit mã hóa protein virus cấu trúc có thể là, ví dụ, HA từ cúm B/Brisbane/60/2008, B/Malaysia/2506/2004 hoặc B/Wisconsin/1/2010, hoặc H3 từ cúm A/Perth/16/2009 hoặc A/Victoria/361/2011.

Sáng chế cũng bao gồm phương pháp (B) như được mô tả trên đây, trong đó trình tự axit nucleic thứ nhất bao gồm vùng điều hòa thứ nhất được liên kết theo kiểu điều khiển

được bằng một hoặc nhiều vùng tăng cường comovirut, trình tự nucleotit mã hóa protein virut cấu trúc, và một hoặc nhiều yếu tố khuếch đại geminivirut, và axit nucleic thứ ba mã hóa replicaza geminivirut được đưa vào thực vật hoặc phần của thực vật. Một hoặc nhiều vùng tăng cường comovirut có thể là UTR của comovirut, ví dụ, UTR virut khảm của cây đậu đũa có khả năng siêu dịch mã (CPMV-HT) như CPMV-HT 5' và/hoặc 3'UTR. Ngoài ra, một hoặc nhiều yếu tố khuếch đại geminivirut có thể được chọn từ vùng liên gen dài virut gây bệnh lùn, vàng của cây đậu (BeYDV LIR), và vùng liên gen ngắn BeYDV (BeYDV SIR).

Phương pháp như được mô tả trên đây (Phương pháp B) cũng có thể bao gồm việc đưa trình tự axit nucleic khác mã hóa gen ức chế bất hoạt, ví dụ, HcPro hoặc p19.

Sáng chế cũng bao gồm phương pháp (B) như được mô tả trên đây, trong đó trong bước đưa (bước a), axit nucleic được biểu hiện nhất thời ở thực vật. Theo cách khác, trong bước đưa (bước a), axit nucleic được biểu hiện ổn định ở thực vật.

Phương pháp (B) như được mô tả trên đây có thể còn bao gồm bước:

d) thu hoạch thực vật và tinh chế các VLP.

Sáng chế cũng bao gồm phương pháp (B) như được mô tả trên đây, trong đó VLP không chứa protein nền virut hoặc protein lõi.

Sáng chế đề xuất VLP được tạo ra bằng phương pháp (B) như được mô tả trên đây. VLP có thể còn bao gồm một hoặc nhiều lipid thu được từ thực vật. VLP cũng có thể đặc trưng ở chỗ không chứa protein kênh. Hơn nữa, protein virut cấu trúc của VLP có thể là protein HA0. Một hoặc nhiều protein virut bao gồm VLP có thể bao gồm các N-glycan đặc hiệu thực vật, hoặc các N-glycan được biến đổi. Sáng chế cũng đề xuất kháng thể đa dòng được tạo ra bằng cách sử dụng VLP.

Sáng chế bao gồm chế phẩm bao gồm liệu hữu hiệu của VLP được tạo ra bằng phương pháp (B) như được vừa được mô tả, để kích thích đáp ứng miễn dịch, và chất mang được dùng.

Sáng chế cũng mô tả phương pháp gây miễn dịch đối với sự lây nhiễm virus cúm ở đối tượng, bao gồm dùng VLP như được vừa được mô tả, cho đối tượng. VLP có thể được dùng cho đối tượng qua đường miệng, trong da, trong mũi, trong cơ, trong màng bụng, trong tĩnh mạch, hoặc dưới da.

Sáng chế cũng đề xuất vật liệu thực vật bao gồm VLP được tạo ra bằng phương pháp (B) được mô tả trên đây. Vật liệu thực vật có thể được sử dụng để gây miễn dịch với lây nhiễm virus cúm ở đối tượng. Vật liệu thực vật cũng có thể được trộn làm chất bổ sung cho thực phẩm.

Sáng chế đề xuất polypeptit bao gồm trình tự axit amin SEQ ID NO:41 (PDISP/HA từ cúm B/Brisbane/60/2008 có vòng phân hủy protein bị xóa bỏ), và trình tự axit nucleic mã hóa polypeptit của SEQ ID NO:41. Trình tự axit nucleic có thể bao gồm trình tự nucleotit của SEQ ID NO:43. Sáng chế đề xuất VLP bao gồm polypeptit bao gồm trình tự axit amin SEQ ID NO:41. VLP có thể còn bao gồm một hoặc nhiều lipid thu được từ thực vật. VLP cũng có thể đặc trưng ở chỗ không chứa protein kênh. VLP có thể bao gồm các N-glycan đặc hiệu thực vật, hoặc các N-glycan được biến đổi. Sáng chế đề xuất chế phẩm bao gồm liều hữu hiệu của VLP bao gồm trình tự axit amin SEQ ID NO:41, để kích thích đáp ứng miễn dịch, và chất mang được dùng. Sáng chế cũng bao gồm phương pháp gây miễn dịch đối với sự lây nhiễm virus cúm ở đối tượng, bao gồm dùng VLP bao gồm trình tự axit amin SEQ ID NO:41, cho đối tượng. VLP có thể được dùng cho đối tượng qua đường miệng, trong da, trong mũi, trong cơ, trong màng bụng, trong tĩnh mạch, hoặc dưới da. Sáng chế cũng đề xuất vật liệu thực vật bao gồm VLP bao gồm trình tự axit amin SEQ ID NO:41. Vật liệu thực vật có thể được sử dụng trong gây miễn dịch đối với sự lây nhiễm virus cúm ở đối tượng. Vật liệu thực vật cũng có thể được trộn làm chất bổ sung cho thực phẩm.

Bằng cách biểu hiện đồng thời protein virus cấu trúc cùng với protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton, hiệu suất của protein virus cấu trúc và các VLP tăng được quan sát. Protein HA được biết là trải qua sự biến đổi cấu dạng phụ thuộc vào độ pH. Không mong muốn bị ràng buộc bởi lý thuyết, độ pH nằm trong bộ máy Golgi của các tế bào tạo ra HA trong quá trình trưởng thành và di trú có thể ảnh hưởng đến việc

cuộn gập HA, ảnh hưởng đến độ ổn định và làm tăng sự phân hủy, hoặc sự kết hợp của chúng, của HA. Bằng cách biểu hiện đồng thời protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton, cùng với HA, độ pH nằm trong bộ máy Golgi có thể làm tăng, và dẫn đến tăng sự ổn định, giảm sự phân hủy hoặc sự kết hợp của chúng, và tăng hiệu suất HA.

Bản chất kỹ thuật của sáng chế này không nhất thiết phải mô tả tất cả các dấu hiệu của sáng chế.

Mô tả vắn tắt các hình vẽ

Các dấu hiệu này và các dấu hiệu khác theo sáng chế sẽ trở nên rõ ràng dựa vào phần mô tả sau đây mà trong đó tham chiếu đến các hình vẽ kèm theo, trong đó:

Fig.1A thể hiện đoạn môi IF-H5A-I-05.s1+3c (SEQ ID NO: 2). **Fig.1B** thể hiện đoạn môi IF-H5dTm.r (SEQ ID NO: 3). **Fig.1C** thể hiện dạng sơ đồ của cấu trúc 1191. **Fig.1D** thể hiện Cấu trúc 1191 (SEQ ID NO 4). **Fig.1E** thể hiện cat-xet biểu hiện số 489 (SEQ ID NO 5). **Fig.1F** thể hiện trình tự axit amin của H5 từ cúm A/Indonesia/5/2005 (H5N1) (SEQ ID NO: 6). **Fig.1G** thể hiện trình tự nucleotit mã hóa H5 từ cúm A/Indonesia/5/2005 (H5N1) (SEQ ID NO: 42).

Fig.2A thể hiện đoạn môi IF-S1-M1+M2ANC.c (SEQ ID NO:7). **Fig.2B** thể hiện đoạn môi IF-S1-4-M2ANC.r (SEQ ID NO: 8). **Fig.2C** thể hiện trình tự nucleotit của gen M2 được tổng hợp (tương ứng với nt 1-26 được nối với 715-982 từ số truy cập Genbank DQ508860) (SEQ ID NO: 9). **Fig.2D** thể hiện cat-xet biểu hiện số 1261 từ promoter (trình tự khởi đầu) 2X35S đến vùng kết thúc NOS. M2 từ cúm A/New Caledonia/20/1999 (H1N1) được gạch dưới. (SEQ ID NO: 10). **Fig.2E** thể hiện trình tự axit amin của M2 từ cúm A/New Caledonia/20/1999 (H1N1) (SEQ ID NO: 11).

Fig.3A thể hiện trình tự nucleotit của gen M2 được tổng hợp (tương ứng với nt 26-51 được liên kết với nt 740-1007 từ số truy cập Genebank EF467824) (SEQ ID NO: 12). **Fig.3B** thể hiện cat-xet biểu hiện số 859 từ promoter 2X35S đến vùng kết thúc NOS. M2

từ Cúm A/Puerto Rico/8/1934 (H1N1) được gạch dưới. (SEQ ID NO: 13). **Fig.3C** thể hiện trình tự axit amin của M2 từ cúm A/Puerto Rico/8/1934 (H1N1) (SEQ ID NO:14).

Fig.4A thể hiện đoạn môi IF-H1A-C-09.s2+4c (SEQ ID NO: 15). **Fig.4B** thể hiện đoạn môi IF-H1A-C-09.s1-4r (SEQ ID NO: 16). **Fig.4C** thể hiện trình tự nucleotit của gen H1 được tổng hợp (Số truy cập Genbank FJ966974) (SEQ ID NO: 17). **Fig.4D** thể hiện dạng sơ đồ của cấu trúc 1192. Các vị trí của enzym giới hạn SacII và StuI được sử dụng để cắt thẳng plasmit được chú thích trên hình này. **Fig.4E** thể hiện cấu trúc 1192 từ các bờ T-ADN trái sang phải (được gạch dưới). 2X35S/CPMV-HT/PDISP/NOS có cat-xet biểu hiện chất ức chế bất hoạt Plastoxyanin-P19-Plastoxyanin (SEQ ID NO: 18). **Fig.4F** thể hiện cat-xet biểu hiện số 484 từ promoter 2X35S đến vùng kết thúc NOS. PDISP/H1 từ cúm A/California/7/2009 (H1N1) được gạch dưới. (SEQ ID NO: 19). **Fig.4G** thể hiện trình tự axit amin của PDISP-H1 từ cúm A/California/7/2009 (H1N1) (SEQ ID NO: 20).

Fig.5A thể hiện đoạn môi IF-S2+S4-H3 Per.c (SEQ ID NO: 21). **Fig.5B** thể hiện đoạn môi IF-S1a4-H3 Per.r (SEQ ID NO: 22). **Fig.5C** thể hiện trình tự nucleotit của gen H3 được tổng hợp (tương ứng với nt 26-1726 từ số truy cập Genbank GQ293081) (SEQ ID NO: 23). **Fig.5D** thể hiện cat-xet biểu hiện số 1019 từ promoter 2X35S đến vùng kết thúc NOS. PDISP/H3 từ cúm A/Perth/16/2009 (H3N2) được gạch dưới. (SEQ ID NO: 24). **Fig.5E** thể hiện trình tự axit amin của PDISP/H3 từ cúm A/Perth/16/2009 (H3N2) (SEQ ID NO: 25).

Fig.6A thể hiện đoạn môi IF-S2+S4-B Bris.c (SEQ ID NO: 26). **Fig.6B** thể hiện đoạn môi IF-S1a4-B Bris.r (SEQ ID NO: 27). **Fig.6C** thể hiện trình tự nucleotit của gen HA B Brisbane được tổng hợp (tương ứng với nt 34-1791 từ số truy cập Genbank FJ766840) (SEQ ID NO: 28). **Fig.6D** thể hiện trình tự nucleotit của cat-xet biểu hiện số 1029 từ promoter 2X35S đến vùng kết thúc NOS. PDISP/HA từ cúm B/Brisbane/60/2008 được gạch dưới. (SEQ ID NO: 29). **Fig.6E** thể hiện trình tự axit amin của PDISP/HA từ cúm B/Brisbane/60/2008 (SEQ ID NO: 30). **Fig.6F** thể hiện bản đồ của cấu trúc 1194. Các vị trí của enzym giới hạn SacII và StuI được sử dụng để cắt thẳng plasmit được chú thích trên bản đồ này. **Fig.6G** thể hiện cấu trúc 1194 từ các bờ t-ADN trái sang phải (được gạch

dưới). 2X35S/CPMV-HT/PDISP/NOS trong khuếch đại BeYDV+Replicaza có cat-xet biểu hiện chất ức chế bất hoạt Plastoxyanin-P19-Plastoxyanin (SEQ ID NO: 31). **Fig.6H** thể hiện cat-xet biểu hiện số 1008 từ LIR bên trái BeYDV sang LIR bên phải BeYDV. PDISP/HA từ cúm B/Brisbane/60/2008 được gạch dưới. (SEQ ID NO: 32).

Fig.7A thể hiện đoạn môi dTmH5I-B Bris.r (SEQ ID NO: 33). **Fig.7B** thể hiện đoạn môi B Bris-dTmH5I.c (SEQ ID NO: 34). **Fig.7C** thể hiện đoạn môi IF-S1aS4-dTmH5I.r (SEQ ID NO: 35). **Fig. 7D** thể hiện cat-xet biểu hiện số 1009 từ LIR bên trái BeYDV sang LIR bên phải BeYDV. PDISP/HA B Brisbane/H5Indo TMCT được gạch dưới. (SEQ ID NO:36). **Fig.7E** thể hiện trình tự axit amin của PDISP/HA B Brisbane/H5Indo TMCT (SEQ ID NO: 37).

Fig.8A thể hiện đoạn môi 1039+1059.r (SEQ ID NO: 38). **Fig.8B** thể hiện đoạn môi 1039+1059.c (SEQ ID NO: 39). **Fig.8C** thể hiện cat-xet biểu hiện số 1059 từ LIR bên trái BeYDV sang LIR bên phải BeYDV. PDISP/HA từ cúm B/Brisbane/60/2008 có vòng protein thủy phân bị xóa bỏ được gạch dưới. (SEQ ID NO: 40). **Fig.8D** thể hiện trình tự axit amin của PDISP/HA từ cúm B/Brisbane/60/2008 có vòng phân hủy protein bị xóa bỏ (SEQ ID NO: 41). **Fig.8E** thể hiện trình tự nucleotit của PDISP/HA từ cúm B/Brisbane/60/2008 có vòng phân hủy protein bị xóa bỏ (SEQ ID NO: 43).

Fig.9 thể hiện bản đồ plasmit của cấu trúc số 1008. Cấu trúc số 1008 điều khiển sự biểu hiện HA kiểu đại từ chủng cúm B/Brisbane/60/2008. Cấu trúc này bao gồm các yếu tố thu được từ BeYDV để khuếch đại ADN.

Fig.10 thể hiện bản đồ plasmit của cấu trúc số 1009. Cấu trúc số 1009 điều khiển sự biểu hiện của HA thể khảm từ chủng cúm B/Brisbane/60/2008 mà trong đó miền xuyên màng và đuôi tế bào chất được thay thế bằng của H5 từ cúm A/Indonesia/05/2005. Cấu trúc này bao gồm các yếu tố thu được từ BeYDV để khuếch đại ADN.

Fig.11 thể hiện bản đồ plasmit của cấu trúc số 1029. Cấu trúc số 1029 điều khiển sự biểu hiện HA kiểu đại từ chủng cúm B/Brisbane/60/2008.

Fig.12 thể hiện bản đồ plasmit của cấu trúc số 1059. Cấu trúc số 1059 điều khiển sự biểu hiện của HA đột biến từ chủng cúm B/Brisbane/60/2008 có vòng phân hủy protein bị xóa bỏ. Cấu trúc này bao gồm các yếu tố thu được từ BeYDV để khuếch đại ADN.

Fig.13 thể hiện bản đồ plasmit của cấu trúc số 1019. Cấu trúc số 1019 điều khiển sự biểu hiện H3 kiểu đại từ chủng cúm A/Perth/16/2009 (H3N2).

Fig.14 thể hiện bản đồ plasmit của cấu trúc số 484. Cấu trúc số 484 điều khiển sự biểu hiện H1 kiểu đại từ chủng cúm A/California/07/2009 (H1N1).

Fig.15 thể hiện bản đồ plasmit của cấu trúc số 489. Cấu trúc số 489 điều khiển sự biểu hiện H5 kiểu đại từ chủng cúm A/Indonesia/05/2005 (H5N1).

Fig.16 thể hiện bản đồ plasmit của cấu trúc số 1261. Cấu trúc số 1261 điều khiển sự biểu hiện M2 kiểu đại từ chủng cúm A/New Caledonia/20/99 (H1N1).

Fig.17 thể hiện bản đồ plasmit của cấu trúc số 859. Cấu trúc số 859 điều khiển sự biểu hiện M2 kiểu đại từ chủng cúm A/Puerto Rico/8/34 (H1N1).

Fig.18 thể hiện phân tích thẩm tách Western của sự biểu hiện protein HA trong các lá *Nicotiana benthamiana* được thẩm Argobacteria (Agroinfiltration). HA từ B/Brisbane/60/2008 được biểu hiện đồng thời bằng M2 từ A/New Caledonia/20/99. “C+”: đối chứng dương, virus B/Brisbane/60/2008 được bán tinh chế từ Therapeutic Goods Administration, Australia; “C-”: đối chứng âm, các thực vật được giả thẩm; “1008”: biểu hiện HA kiểu đại từ B/Brisbane/60/2008 với sự có mặt của các yếu tố khuếch đại (BeYDV); “1008+1261”: biểu hiện đồng thời HA kiểu đại từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại (BeYDV) bằng M2; “1009+1261”: biểu hiện đồng thời HA thể khảm từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại (BeYDV) bằng M2; “1029”: biểu hiện HA kiểu đại từ B/Brisbane/60/2008 với sự vắng mặt của các yếu tố khuếch đại (BeYDV); “1029+1261”: biểu hiện đồng thời HA kiểu đại từ B/Brisbane/60/2008 với sự vắng mặt của các yếu tố khuếch đại(BeYDV) bằng M2 từ A/New Caledonia/20/99. Các tỷ lệ chỉ ra tỷ lệ môi trường nuôi cấy *Agrobacterium* được sử dụng trong các thử nghiệm biểu hiện đồng thời.

Fig.19 thể hiện phân tích thấm tách Western của biểu hiện protein HA ở các lá *Nicotiana benthamiana* được thấm bằng *Agrobacteria*. Giếng “C+”: Đối chứng dương, virus A/Wisconsin/15/2009 (H3N2) được bán tinh chế từ Therapeutic Goods Administration, Australia; “C-”: đối chứng âm, các thực vật được giả thấm; “1019”: biểu hiện HA kiểu đại từ A/Perth/16/2009 (H3N2); “1019+1261”: biểu hiện đồng thời HA kiểu đại từ A/Perth/16/2009 (H3N2) bằng M2 từ A/New Caledonia/20/99. Tỷ lệ này chỉ ra tỷ lệ của các môi trường nuôi cấy *Agrobacterium* được sử dụng trong các thử nghiệm biểu hiện đồng thời.

Fig.20 thể hiện phân tích thấm tách Western của Biểu hiện protein HA ở các lá *Nicotiana benthamiana* được thấm bằng *Agrobacteria*. Giếng “C+”: Đối chứng dương, A/California/7/2009 (H1N1) NYMC X-179A được bán tinh chế từ virus NIBSC (mã NIBSC 09/146); “C-”: đối chứng âm, các thực vật được giả thấm; “484”: biểu hiện HA kiểu đại from A/California/7/2009 (H1N1); “484+1261”: biểu hiện đồng thời HA kiểu đại từ A/California/7/2009 (H1N1) bằng M2 từ A/New Caledonia/20/99. Tỷ lệ chỉ ra tỷ lệ của các môi trường nuôi cấy *Agrobacterium* được sử dụng trong các thử nghiệm biểu hiện đồng thời.

Fig.21 thể hiện phân tích thấm tách Western của biểu hiện protein HA trong các lá *Nicotiana benthamiana* được thấm bằng *Agrobacteria*. Giếng “C+”: Đối chứng dương, H5 tái tổ hợp được tinh chế từ A/Indonesia/05/2005, Immune Technology Corporation (sản phẩm số IT-003-052p) ; “C-”: đối chứng âm, các thực vật được giả thấm; “489”: biểu hiện HA kiểu đại từ A/Indonesia/5/05 (H5N1); “489+1261”: biểu hiện đồng thời HA kiểu đại từ A/Indonesia/5/05 (H5N1) bằng M2 từ A/New Caledonia/20/99.

Fig.22A thể hiện phân tích thấm tách Western của biểu hiện protein HA ở lá *Nicotiana benthamiana* được thấm bằng *Agrobacteria*. “1008”: biểu hiện HA kiểu đại từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại (BeYDV); “1008+1261”: biểu hiện đồng thời HA kiểu đại từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại (BeYDV) bằng M2 từ A/New Caledonia/20/99; “1059”: biểu hiện HA đột biến từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại (BeYDV); “1059+1261”: biểu hiện đồng thời HA đột biến từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại

(BeYDV) bằng M2 từ A/New Caledonia/20/99. Các thực vật từ ba lần thăm riêng biệt đã được phân tích (A, B và C). Các tỷ lệ chỉ ra tỷ lệ của các môi trường nuôi cấy *Agrobacterium* được sử dụng trong các thử nghiệm biểu hiện đồng thời. **Fig.22B** thể hiện so sánh khả năng hemagglutinin hóa của protein thô chiết từ các thực vật tạo ra HA.

Fig.23 thể hiện phân tích thâm tách Western của biểu hiện protein HA ở lá *Nicotiana benthamiana* được thâm bằng *Agrobacteria*. **Fig.23A**: “1059”: biểu hiện HA đột biến từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại (BeYDV); “1059+1261”: biểu hiện đồng thời HA đột biến từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại (BeYDV) với M2 từ A/New Caledonia/20/99. “1059+859”: biểu hiện đồng thời của HA đột biến từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại (BeYDV) với M2 từ A/Puerto Rico/8/34. Các thực vật từ ba lần thăm riêng biệt đã được phân tích (A, B và C). Các tỷ lệ chỉ ra tỷ lệ của các môi trường nuôi cấy *Agrobacterium* được sử dụng trong các thử nghiệm biểu hiện đồng thời. **Fig.23B**: “1019”: biểu hiện HA kiểu dại từ A/Perth/16/2009 (H3N2); “1019+1261”: biểu hiện đồng thời HA kiểu dại từ A/Perth/16/2009 (H3N2) với M2 từ A/New Caledonia/20/99; “1019+859”: biểu hiện đồng thời HA kiểu dại từ A/Perth/16/2009 (H3N2) với M2 từ A/Puerto Rico/8/34. Các tỷ lệ chỉ ra tỷ lệ của các môi trường nuôi cấy *Agrobacterium* được sử dụng trong các thử nghiệm biểu hiện đồng thời.

Fig.24 thể hiện sự sắp xếp trình tự của các HA từ một vài chủng cúm. Vị trí phân cắt của tiền phân tử HA0 được chỉ ra bằng mũi tên.

Fig.25A thể hiện đoạn môi IF-H3V36111.S2+4c (SEQ ID NO: 44). **Fig.25B** thể hiện đoạn môi IF-H3V36111.s1-4r (SEQ ID NO: 45). **Fig.25C** thể hiện trình tự nucleotit của gen H3 được tổng hợp (tương ứng với nt 25 đến 1725 từ thể phân lập GISAIID số EPI_ISL_101506 trình tự HA) (SEQ ID NO: 46). **Fig.25D** thể hiện trình tự nucleotit của cat-xet biểu hiện số 1391 từ promoter 2X35S đến vùng kết thúc NOS. PDISP/H3 từ cúm A/Victoria/361/2011 (H3N2) được gạch dưới. (SEQ ID NO: 47). **Fig.25E** thể hiện trình tự axit amin của PDISP-H3 từ cúm A/Victoria/361/2011 (H3N2) (SEQ ID NO: 48). **Fig.25F** thể hiện bản đồ của cấu trúc 1391.

Fig.26A thể hiện đoạn môi IF-HAB110.S1+3c (SEQ ID NO: 49). **Fig.26B** thể hiện đoạn môi IF-HAB110.s1-4r (SEQ ID NO: 50). **Fig.26C** thể hiện trình tự nucleotit của HA B Wisconsin được tổng hợp (Số truy cập Genbank JN993010) (SEQ ID NO: 51). **Fig.26D** thể hiện bản đồ của cấu trúc 193. **Fig.26E** thể hiện cấu trúc 193 từ các bờ t-ADN từ trái sang phải (được gạch dưới). 2X35S/CPMV-HT/NOS vào hệ khuếch đại BeYDV(m)+Replicaza có cat-xet biểu hiện chất ức chế bất hoạt Plastoxyanin-P19-Plastoxyanin (SEQ ID NO: 52). **Fig.26F** thể hiện trình tự nucleotit của cat-xet biểu hiện số 1462 từ promoter 2X35S đến vùng kết thúc NOS. HA từ cúm B/Wisconsin/1/2010 được gạch dưới (SEQ ID NO: 53). **Fig.26G** thể hiện trình tự axit amin của HA từ cúm B/Wisconsin/1/2010 (SEQ ID NO: 54). **Fig.26H** thể hiện bản đồ của cấu trúc 1462.

Fig.27A thể hiện đoạn môi HAB110(PrL-).r (SEQ ID NO: 55). **Fig.27B** thể hiện đoạn môi HAB110(PrL-).c (SEQ ID NO: 56). **Fig.27C** thể hiện trình tự nucleotit của cat-xet biểu hiện số 1467 từ promoter 2X35S đến vùng kết thúc NOS. HA từ cúm B/Wisconsin/1/2010 có vòng phân hủy protein bị xóa bỏ được gạch dưới (SEQ ID NO: 57). **Fig.27D** thể hiện trình tự axit amin của cúm B/Wisconsin/1/2010 có vòng phân hủy protein bị xóa bỏ (SEQ ID NO: 58). **Fig.27E** thể hiện bản đồ của cấu trúc 1467.

Fig.28A thể hiện đoạn môi IF-HB-M-04.s2+4c (SEQ ID NO: 59). **Fig.28B** thể hiện đoạn môi IF-HB-M-04.s1-4r (SEQ ID NO: 60). **Fig.28C** thể hiện trình tự nucleotit của HA B Malaysia được tổng hợp (tương ứng với nt 31-1743 từ số truy cập Genbank EU124275) bằng T759C và các đột biến C888G được gạch dưới. (SEQ ID NO: 61). **Fig.28D** thể hiện bản đồ của cấu trúc 194, bằng các vị trí của enzym giới hạn SacII và StuI được sử dụng để cắt thẳng plasmit được chú thích trên bản đồ này. **Fig.28E** thể hiện cấu trúc 194 từ các bờ T-ADN từ trái sang phải (được gạch dưới). 2X35S/CPMV-HT/NOS vào hệ khuếch đại BeYDV(m)+Replicaza với cat-xet biểu hiện chất ức chế bất hoạt Plastoxyanin-P19-Plastoxyanin (SEQ ID NO: 62). **Fig.28F** thể hiện trình tự nucleotit của cat-xet biểu hiện số 1631 từ promoter 2X35S đến vùng kết thúc NOS. PDISP-HA từ cúm B/ Malaysia/2506/2004 được gạch dưới. (SEQ ID NO: 63). **Fig.28G** thể hiện trình tự axit amin của PDISP-HA từ cúm B/Malaysia/2506/2004 (SEQ ID NO: 64). **Fig.28H** thể hiện bản đồ của cấu trúc 1631.

Fig.29 thể hiện phân tích thẩm tách Western của biểu hiện protein HA ở các lá *Nicotiana benthamiana* được thẩm bằng *Agrobacteria*. HA từ B/Malaysia/2506/2004 được biểu hiện đồng thời với M2 từ A/New Caledonia/20/99. 20mg dịch chiết protein được nạp vào mỗi giếng. “C+”: đối chứng dương, virus B/Malaysia/2506/2004 được bán tinh chế từ National Institute for Biological Standards and Control, United Kingdom; “1631”: biểu hiện HA kiểu dại từ B/Malaysia/2506/2004 với sự có mặt của yếu tố khuếch đại (BeYDV); “1631+1261”: biểu hiện đồng thời HA kiểu dại từ B/Malaysia/2506/2004 với sự có mặt của yếu tố khuếch đại (BeYDV) bằng M2. Các tỷ lệ chỉ ra tỷ lệ của các môi trường nuôi cấy *Agrobacterium* được sử dụng trong các thử nghiệm biểu hiện đồng thời.

Fig.30A thể hiện phân tích thẩm tách Western của biểu hiện protein HA ở các lá *Nicotiana benthamiana* được thẩm bằng *Agrobacteria*. HA từ B/Wisconsin/1/2010 được biểu hiện đồng thời bằng M2 từ A/New Caledonia/20/99. 10mg dịch chiết protein được nạp vào trên mỗi giếng. “C+”: đối chứng dương, virus B/Wisconsin/1/2010 được bán tinh chế từ National Institute for Biological Standards and Control, United Kingdom; “1462”: biểu hiện HA kiểu dại từ B/Wisconsin/1/2010 với sự có mặt của yếu tố khuếch đại (BeYDV); “1467”: biểu hiện HA đột biến từ B/Wisconsin/1/2010 với sự có mặt của yếu tố khuếch đại (BeYDV); “1462+1261”: biểu hiện đồng thời HA kiểu dại từ B/Wisconsin/1/2010 với sự có mặt của yếu tố khuếch đại (BeYDV) bằng M2; “1467+1261”: biểu hiện đồng thời HA đột biến từ B/Wisconsin/1/2010 với sự có mặt của yếu tố khuếch đại (BeYDV) bằng M2. Các tỷ lệ chỉ ra mật độ quang đối với mỗi môi trường nuôi cấy *Agrobacterium* được sử dụng trong các thử nghiệm biểu hiện và biểu hiện đồng thời. **Fig.30B** thể hiện khả năng hemagglutinin hóa của dịch chiết protein thô từ thực vật được biến nạp bằng AGL1/1462, AGL1/1467, AGL1/1462+AGL1/1261 và AGL1/1467+AGL1/1261.

Fig.31 thể hiện phân tích thẩm tách Western của biểu hiện protein HA ở lá *Nicotiana benthamiana* được thẩm bằng *Agrobacteria*. HA từ H3/Victoria/361/2011 được biểu hiện đồng thời với M2 từ A/New Caledonia/20/99. 20mg dịch chiết protein được nạp vào trên mỗi giếng. “C+”: đối chứng dương, virus H3/Wisconsin/15/2009 được bán tinh chế từ Therapeutic Goods Administration, Australia; “1391”: biểu hiện HA kiểu dại từ

H3/Victoria/361/2011; “1391+1261”: biểu hiện đồng thời HA kiểu đại từ H3/Victoria/361/2011 bằng M2. Các tỷ lệ chỉ ra mật độ quang đối với mỗi môi trường nuôi cấy *Agrobacterium* được sử dụng trong các thử nghiệm biểu hiện và biểu hiện đồng thời.

Mô tả chi tiết sáng chế

Phần mô tả sau đây bao gồm phương án được ưu tiên.

Sáng chế đề cập đến các hạt giống như virus (VLP) và các phương pháp tạo ra và làm tăng sản lượng và sự tạo ra VLP ở thực vật.

Sáng chế đề xuất, cụ thể là, phương pháp sản xuất hạt giống như virus (VLP) ở thực vật, hoặc phần của thực vật. Phương pháp này bao gồm đưa axit nucleic thứ nhất và axit nucleic thứ hai vào thực vật. Axit nucleic thứ nhất bao gồm vùng điều hòa thứ nhất hoạt động ở thực vật hoặc phần của thực vật, và được liên kết theo kiểu điều khiển được với trình tự nucleotit mã hóa protein virus cấu trúc. Axit nucleic thứ hai bao gồm vùng điều hòa thứ hai hoạt động ở thực vật và được liên kết theo kiểu điều khiển được với trình tự nucleotit mã hóa protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton. Vùng điều hòa thứ nhất và vùng điều hòa thứ hai có thể giống hoặc khác nhau. Thực vật hoặc phần của thực vật được ủ dưới các điều kiện cho phép biểu hiện axit nucleic, theo đó tạo ra VLP. Nếu muốn, thực vật hoặc phần của thực vật có thể được thu hoạch và VLP được tinh chế. Tốt hơn là, VLP không chứa M1, protein nền virus hoặc protein lõi. Sáng chế cũng đề xuất VLP được tạo ra bằng phương pháp này. VLP có thể bao gồm một hoặc nhiều hơn một lipid thu được từ thực vật. VLP có thể được sử dụng để điều chế chế phẩm bao gồm liều hữu hiệu của VLP để kích thích đáp ứng miễn dịch, và chất mang được dùng.

Sáng chế cũng đề xuất vật liệu thực vật bao gồm VLP được tạo ra bằng cách biểu hiện axit nucleic thứ nhất và thứ hai được mô tả trên đây. Vật liệu thực vật có thể được sử dụng để gây miễn dịch đối với sự lây nhiễm virus cúm ở đối tượng. Vật liệu thực vật cũng có thể được trộn làm chất bổ sung cho thực phẩm.

The VLP theo sáng chế cũng có thể được tạo ra bằng cách tạo ra thực vật hoặc phần của thực vật bao gồm axit nucleic thứ nhất và axit nucleic thứ hai như được xác định trên đây, và thực vật hoặc phần của thực vật dưới các điều kiện cho phép biểu hiện axit nucleic thứ nhất và thứ hai, theo đó tạo ra VLP. VLP có thể bao gồm một hoặc nhiều lipid thu được từ thực vật. VLP có thể được sử dụng để điều chế chế phẩm bao gồm liệu hữu hiệu của VLP để kích thích đáp ứng miễn dịch, và chất mang được dụng. Sáng chế cũng đề xuất vật liệu thực vật bao gồm VLP được tạo ra bằng cách biểu hiện axit nucleic thứ nhất và thứ hai. Vật liệu thực vật có thể được sử dụng trong gây miễn dịch đối với sự lây nhiễm virus cúm ở đối tượng. Vật liệu thực vật cũng có thể được trộn làm chất bổ sung cho thực phẩm.

VLP theo sáng chế bao gồm một hoặc nhiều protein virus. Ví dụ, không bị coi là giới hạn sáng chế, một hoặc nhiều protein virus có thể là protein virus cấu trúc như hemagglutinin cúm (HA) hoặc protein kênh, ví dụ, nhưng không giới hạn ở, protein kênh proton, như, ví dụ, M2. HA có thể là HA bất kỳ, ví dụ, H2, H3, H4, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16 hoặc HA typ B như được mô tả trong các công bố đơn sáng chế quốc tế số WO 2009/009876; WO 2009/076778; WO 2010/003225; WO 2010/003235; WO 2011/03522).

Như được mô tả chi tiết dưới đây, các VLP có thể được tạo ra ở thực vật bằng cách biểu hiện đồng thời axit nucleic thứ nhất mã hóa protein virus với axit nucleic thứ hai mã hóa protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton. Axit nucleic thứ nhất và thứ hai có thể được đưa vào thực vật trong cùng bước, hoặc chúng có thể được đưa vào thực vật theo thứ tự. Axit nucleic thứ nhất và thứ hai có thể được đưa vào thực vật theo cách nhất thời, hoặc theo cách ổn định. Hơn nữa, thực vật biểu hiện axit nucleic thứ nhất mã hóa protein virus có thể được biến nạp protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton, (axit nucleic thứ hai) sao cho cả hai axit nucleic thứ nhất và thứ hai được biểu hiện đồng thời ở thực vật. Theo cách khác, thực vật biểu hiện protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton, (axit nucleic thứ hai) có thể được biến nạp axit nucleic thứ nhất mã hóa protein virus sao cho cả hai axit nucleic thứ nhất và thứ hai được biểu hiện đồng thời ở thực vật. Ngoài ra, thực vật thứ nhất biểu hiện axit nucleic

thứ nhất mã hóa protein virus, có thể lai với thực vật thứ hai biểu hiện axit nucleic thứ hai mã hóa protein kênh, nhưng không giới hạn ở protein kênh proton, để tạo ra cây con lần lượt biểu hiện đồng thời axit nucleic thứ nhất và thứ hai mã hóa protein virus và protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton.

Sáng chế cũng đề xuất phương pháp làm tăng sự biểu hiện và hiệu suất của protein virus ở thực vật bằng cách biểu hiện đồng thời axit nucleic thứ nhất mã hóa protein virus với axit nucleic thứ hai mã hóa protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton. Axit nucleic thứ nhất và thứ hai có thể được đưa vào thực vật trong cùng bước, hoặc chúng có thể được đưa vào thực vật theo thứ tự. Axit nucleic thứ nhất và thứ hai có thể được đưa vào thực vật theo cách nhất thời hoặc theo cách ổn định. Hơn nữa, thực vật biểu hiện axit nucleic thứ nhất mã hóa protein virus có thể được biến nạp protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton (axit nucleic thứ hai) sao cho cả hai axit nucleic thứ nhất và thứ hai được biểu hiện đồng thời ở thực vật. Theo cách khác, thực vật biểu hiện protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton (axit nucleic thứ hai) có thể được biến nạp axit nucleic thứ nhất mã hóa protein virus sao cho cả hai axit nucleic thứ nhất và thứ hai được biểu hiện đồng thời ở thực vật. Ngoài ra, thực vật thứ nhất biểu hiện axit nucleic thứ nhất mã hóa protein virus, có thể được lai với thực vật thứ hai biểu hiện axit nucleic thứ hai mã hóa protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton, để tạo ra cây con lần lượt biểu hiện axit nucleic thứ nhất và thứ hai mã hóa protein virus và protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton.

Protein kênh

“Protein kênh” nghĩa là protein có khả năng tạo kênh qua màng phospholipid cho phép sự đi qua của các ion và/hoặc phân tử nhỏ qua màng này. Các protein kênh có thể có tính chọn lọc theo kích cỡ và/hoặc điện tích của các ion và/hoặc phân tử nhỏ. Các ví dụ không giới hạn của các protein kênh là protein kênh không đặc hiệu biến đổi độ thấm của màng với các hợp chất và protein kênh ion có trọng lượng phân tử thấp, như, ví dụ, kênh clorua, kênh kali, kênh natri, kênh canxi và kênh proton.

“Protein kênh proton” nghĩa là protein có khả năng tạo thành kênh chọn lọc proton đi qua lớp kép phospholipit. Protein kênh proton có thể là protein xuyên màng một lần có miền xuyên màng (TM) được kẹp giữa bởi các miền kỵ nước. Miền TM của kênh proton có thể bao gồm trình tự HXXXW (SEQ ID NO. 1).

Sau khi phân cắt HA0, HA trở nên nhạy cảm với độ pH, trải qua sự thay đổi cấu dạng không thuận nghịch ở độ pH của endosom ($\text{pH}=6,0$). Cấu dạng của tiền phân tử HA0 ổn định ở độ pH thấp, nhưng dạng HA1-HA2 bị phân cắt, có thể bán bền vững (Bullough PA et. al., 1994, Nature. Vol 371 :37-43). Các nghiên cứu về ngưỡng độ pH kích thích sự thay đổi cấu dạng ở các HA khác nhau, cho thấy rằng ngưỡng này là độ pH nằm trong khoảng từ 5,8 đến 5,9 đối với các chủng B, trong khi đó nó mang tính axit hơn (độ pH từ 5,1 đến 5,3) đối với các HA typ A (Beyer WEP et al, 1986, Archives Virol, vol 90: 173). Trong quá trình chiết sinh khối thực vật (độ pH nằm trong khoảng từ 5 đến 6), sự thay đổi cấu dạng của HA1-HA2 có thể xảy ra với HA typ B.

Không mong muốn bị giới hạn bởi lý thuyết, độ pH của gian bào bao gồm HA, bao gồm bộ máy Golgi, do đó có thể quan trọng cho quá trình cuộn gập, sự ổn định và /hoặc sự phân hủy protein HA. Protein kênh proton, như, ví dụ, các protein M2 và BM2 cúm có thể điều hòa pH trong các gian bào. Ví dụ, M2 điều hòa tiềm năng dung hợp màng bằng cách tạo đệm các gian nội bào cả ở giai đoạn đầu và sau của quá trình sao chép virus cúm. Ở giai đoạn đầu của quá trình lây nhiễm các tế bào mới sau khi hấp thụ nội bào các hạt virus, sự hoạt hóa hoạt động của kênh proton M2 dẫn đến sự axit hóa bên trong virion trong quá trình mở vỏ. Ở giai đoạn sau của quá trình lây nhiễm, trong quá trình tạo ra virus, M2 đóng vai trò tăng độ pH trong quá trình chuyển tiếp qua mạng lưới chuyển Golgi và ngăn chặn sự bất hoạt cảm ứng bởi độ pH thấp của các protein được vận chuyển đồng thời, như HA trong trường hợp cúm. Bằng cách biểu hiện đồng thời protein virus cấu trúc cùng với protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton, cho thấy tăng hiệu suất của protein virus cấu trúc và VLP. HA được biết là trải qua sự thay đổi xác nhận phụ thuộc vào độ. Không mong muốn bị ràng buộc bởi lý thuyết, độ pH trong bộ máy Golgi của các tế bào tạo ra HA trong quá trình trưởng thành và di trú có thể ảnh hưởng đến việc cuộn gập HA, ảnh hưởng đến sự ổn định và làm tăng sự phân hủy, hoặc sự kết hợp của chúng,

của HA. Bằng cách biểu hiện đồng thời protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton, cùng với HA, độ pH trong bộ máy Golgi có thể tăng, và dẫn đến tăng sự ổn định, giảm phân hủy, hoặc sự kết hợp của chúng, và tăng mức biểu hiện và hiệu suất của HA và/hoặc VLP.

Bằng cách biểu hiện đồng thời protein virus cấu trúc cùng với protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton, ở thực vật, cho thấy tăng hiệu suất của protein virus cấu trúc và/hoặc VLP, khi so với thực vật đã biểu hiện protein virus cấu trúc mà không biểu hiện đồng thời protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton.

Hơn nữa, bằng cách biểu hiện đồng thời protein virus cấu trúc như HA với protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton, ở thực vật, protein HA có thể có hoạt tính tăng như được thể hiện bằng khả năng hemagglutinin hóa lớn hơn, khi được so sánh với protein HA mà không được biểu hiện đồng thời với protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton. Bằng cách tăng hoạt tính, nghĩa là tăng khả năng hemagglutinin hóa nằm trong khoảng từ 2% đến 100%, hoặc lượng bất kỳ nằm trong đó như được xác định bằng cách sử dụng các kỹ thuật chuẩn trong lĩnh vực này, ví dụ, nằm trong khoảng từ 10% đến 50% hoặc giá trị bất kỳ nằm trong đó, ví dụ, khoảng 2, 5, 8, 10, 12, 15, 18, 20, 22, 24, 25, 26, 28, 30, 32, 34, 35, 36, 38, 40, 42, 44, 45, 46, 48, 50, 52, 54, 55, 56, 58, 60, 65, 70, 75, 80, 85, 90, 95, hoặc 100%, khi được so sánh với hoạt tính của cùng protein HA được tạo ra với sự vắng mặt của protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton.

Như được sử dụng ở đây, thuật ngữ “M2,” “protein M2,” “trình tự M2” và “miền M2” đề cập đến tất cả hoặc một phần trình tự của protein M2 được phân lập từ, dựa trên hoặc có trong chủng virus cúm hoặc thể phân lập có trong tự nhiên hoặc được tạo ra nhân tạo. Do đó, thuật ngữ M2 và thuật ngữ tương tự bao gồm các thể biến đổi trình tự M2 xuất hiện trong tự nhiên bằng đột biến trong suốt vòng đời của virus hoặc được tạo ra để đáp ứng lại áp lực chọn lọc (ví dụ, liệu pháp điều trị bằng thuốc, mở rộng tính hướng kích thích tế bào chủ hoặc tính lây nhiễm, v.v.), cũng như các trình tự M2 được tạo ra tái tổ hợp hoặc tổng hợp. Các ví dụ về các protein kênh có thể được sử dụng bao gồm, nhưng không

giới hạn ở protein kênh proton, ví dụ các protein được liệt kê trong Bảng 1. Ví dụ không giới hạn của các trình tự mà có thể được sử dụng theo sáng chế bao gồm M2 từ A/Puerto Rico/8/1934 và M2 từ A/New Caledonia/20/1999. Protein M2 minh họa bao gồm trình tự axit amin như được thể hiện trong SEQ ID NO: 11 hoặc 14.

Như được sử dụng ở đây, các thuật ngữ “BM2,” “protein BM2,” “trình tự BM2” và “miền BM2” đề cập đến tất cả hoặc một phần trình tự của protein BM2 được phân lập từ, dựa trên hoặc có mặt trong chủng virus cúm hoặc thể phân lập xuất hiện trong tự nhiên hoặc được tạo ra nhân tạo. Do đó, thuật ngữ BM2 và thuật ngữ tương tự bao gồm các thể biến thể trình tự BM2 xuất hiện trong tự nhiên được tạo ra bằng đột biến trong suốt vòng đời của virus hoặc được tạo ra để đáp ứng áp lực chọn lọc (ví dụ, liệu pháp điều trị bằng thuốc, mở rộng tính hướng kích thích tế bào chủ hoặc tính lây nhiễm, v.v), cũng như các trình tự BM2 được tạo ra tái tổ hợp hoặc tổng hợp. Các ví dụ về các protein kênh có thể được sử dụng bao gồm, nhưng không giới hạn ở protein kênh proton được liệt kê trong Bảng 2.

Các trình tự protein kênh proton minh họa bổ sung bao gồm các trình tự được nộp lưu theo các số truy cập Genbank được thể hiện trong Bảng 1 và Bảng 2.

Bảng 1: Các số truy cập với các trình tự axit amin protein M2 của kênh proton

Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank
ABA42438.1	ABB54697.1	ABI36079.1	ADM95491.1	ADM29632.1
ABA42436.1	AAA43253.1	ABI36077.1	ADM95489.1	ADM29566.1
ABA42434.1	BAB19809.1	ABI36075.1	ADM95487.1	ADM29555.1
AAD51268.1	ABD59884.1	ABI36073.1	ADM95485.1	ADM29544.1
AAD51264.1	ABD59882.1	ABI36071.1	ADM95483.1	ADM29533.1
AAC60735.1	ABD59880.1	ABI36069.1	ADM95481.1	ADM29445.1
BAI77393.1	BAD89348.1	ABI36067.1	ADM95479.1	ADM29434.1
BAI77450.1	BAD89338.1	ABI36065.1	ADM95477.1	ADM29423.1
CAP58009.1	BAD89328.1	ABI36063.1	ADM95475.1	ADM29412.1

Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank
CAP58007.1	BAE47133.1	ABI36061.1	ADM95473.1	ADM29401.1
CAP58005.1	ABD59890.1	ABI36059.1	ADM95471.1	ADM29379.1
BAH84754.1	ABD59888.1	ABI36037.1	BAF37390.1	ADM29368.1
BAH86619.1	ABD59886.1	ABI36027.1	ADG59188.1	ADM29357.1
BAH86616.1	ABD59900.1	ABI36016.1	ADG59186.1	ADM29313.1
BAH84985.1	ABD59898.1	ABI36005.1	ADG59184.1	ADM29302.1
YP_308853.1	ABD59896.1	AAV87447.1	ADG59182.1	ADM29291.1
AAD49092.1	ABD59894.1	AAV87431.1	ADG59180.1	ADM29280.1
AAD49090.1	ABD59892.1	AAV32647.1	ADG59178.1	ADM29269.1
AAD49088.1	AAC79578.1	AAV32639.1	ADG59176.1	ADM29258.1
ABQ12378.1	ABB51968.1	AAU00829.1	ADG59174.1	ADM29698.1
AAO33518.1	ABY75105.1	AAU00827.1	ADG59172.1	ADM29687.1
AAO33516.1	ABY75039.1	AAA91324.1	ADG59170.1	ADM29676.1
AAO33514.1	ABY75037.1	ABG75620.1	ADG59168.1	ADM29665.1
AAO33512.1	AAL60446.1	ACR09361.1	ADG59166.1	ADM29654.1
AAO33510.1	ABB00351.1	ACR09359.1	ADG59164.1	ADM29621.1
AAO33508.1	AAA43312.1	ACR09355.1	ADG59162.1	ADM29610.1
AAO33506.1	ABB00339.1	ACR09353.1	ADG59160.1	ADM29599.1
AAO33504.1	ABW38094.1	ACQ99604.1	ADG59158.1	ADM29522.1
AAO33502.1	AAM09299.1	ACQ99602.1	ADG59156.1	ADM29511.1
ABS52607.1	AAM09297.1	ACQ99600.1	ADG59154.1	ADM29500.1
ABS52597.1	ABS00915.1	ACQ99592.1	ADG59152.1	ADM29489.1
ABS52587.1	ABS00914.1	ACQ99590.1	ADG59150.1	ADM29478.1
ABM21873.1	ABS00913.1	ACQ99588.1	ADG59148.1	ADM29467.1
ABM21871.1	ABS00912.1	ACQ99586.1	ADG59146.1	ADM29456.1
ABM21869.1	ABS00911.1	ACP41965.1	ADG59144.1	ADM29390.1
ABM21867.1	ABS00910.1	ACP41955.1	ADG59142.1	ADM29346.1

Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank
ABM21865.1	ABS00909.1	ACP41951.1	ADG59140.1	ADM29335.1
ABM21863.1	ABS00908.1	ACP41946.1	ADG59138.1	ADM29324.1
ABM21861.1	ABS00907.1	ACR49258.1	ADG59136.1	ADM29247.1
AAO33500.1	ABS00906.1	ACR49256.1	ADG59134.1	AEE73588.1
AAD49094.1	ABS00905.1	ACR49254.1	ADG59132.1	AEB89880.1
AAD49086.1	ABS00904.1	ACR49252.1	ADG59130.1	AEB89869.1
AAD49084.1	ABS00903.1	ACR49250.1	ADG59128.1	AEB89858.1
AAD49082.1	ABS00902.1	ACR49248.1	ADG59126.1	AEA74023.1
AAD49080.1	ABB51974.1	ACR49246.1	ADG59124.1	AEA74013.1
AAD49078.1	ABB51972.1	ACR49244.1	ADG59122.1	ADF42731.1
AAD49076.1	ABB51970.1	ACR38840.1	ADG59120.1	ADF42721.1
AAD49074.1	AAD00150.1	ACR38838.1	ADG59118.1	ADF28007.1
AAD49072.1	AAD00148.1	ACR38836.1	ADG59116.1	ADF27997.1
AAD49070.1	AAD00146.1	ACR38834.1	ADG59114.1	ADF27987.1
AAD49068.1	AAD00144.1	ACR38832.1	ADG59112.1	ADF27977.1
ACA25333.1	AAD00142.1	ACR18965.1	ADG59110.1	ADF27967.1
ACA25323.1	AAD00140.1	ACR18963.1	ADG59108.1	ADF27957.1
ACA25313.1	AAD00138.1	ACR18958.1	ADG59106.1	ADF27947.1
CAJ12148.1	AAD00136.1	ACR18957.1	ADG59104.1	ADF27937.1
CAJ12154.1	AAD00134.1	ACR18953.1	ADG59102.1	ADF27927.1
CAJ12152.1	AAD00132.1	ACR18949.1	ADG59100.1	ADF27917.1
CAJ12150.1	AAD00130.1	ACR18946.1	ADG59098.1	ADF27907.1
ACP41109.1	AAC80168.1	ACR18945.1	ADG59096.1	ADF27897.1
ADG59536.1	AAC80166.1	ACR18943.1	ADG59094.1	ADF27887.1
AAK14988.1	AAC80164.1	ACR08560.1	ABO21713.1	ACS87931.1
AAK14984.1	AAC80162.1	ACR08556.1	ADG59717.1	ACU44926.1
ACR67209.1	AAC80160.1	NP_040979.2	ADG59706.1	ACU44922.1

Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank
ACP41929.2	AAC80158.1	ABZ91697.1	AAF74335.1	ACU44920.1
ACR18961.1	AAC80156.1	ABZ91685.1	AAF74333.1	ACU44918.1
ACR18955.1	ABY75159.1	ACB54711.1	ADF56637.1	ACU44916.1
ACR18941.1	ABY75157.1	ABM90504.1	ADF56636.1	ACU44914.1
ACR08564.1	ABY75155.1	ABM90493.1	ADF56635.1	ACU44912.1
ACR08562.1	ABY75153.1	ABM90482.1	ADF29921.1	ACU44910.1
ACR08558.1	ABY75151.1	ABM90471.1	ADE75385.1	ACU44908.1
ACQ99594.1	ABY75149.1	ABM90460.1	ADE75374.1	ACU44906.1
ACQ83308.1	ABY75147.1	ABM90449.1	ADE75365.1	ACU44904.1
ACQ76400.1	ABY75145.1	ABM90438.1	ADE75354.1	ACU44902.1
ACQ76382.1	ABY75143.1	ABI49411.1	ADE75344.1	ACU44900.1
ACQ76375.1	ABY75141.1	ABI49400.1	ADE75327.1	ACU44898.1
ACQ76369.1	ABY75139.1	ABO31433.1	ADE75298.1	ACU44896.1
ACQ76361.1	ABY75137.1	ABM90548.1	ADE75287.1	ACU44894.1
ACQ76355.1	ABY75135.1	ABM90537.1	ADE75276.1	ACU44892.1
ACQ76346.1	ABY75133.1	ABM90526.1	ADE75265.1	ACU44890.1
ACQ76332.1	ABY75131.1	ABM90515.1	ADE75254.1	ACU44888.1
ACQ76325.1	ABY75129.1	ABL31784.1	ADE75244.1	ACU44886.1
ACQ76313.1	ABY75127.1	ABL31770.1	ADE75235.1	ACU44883.1
ACQ76303.1	ABY75125.1	ABL31759.1	ADE75228.1	ACU44881.1
ACQ76293.1	ABY75123.1	ABL31748.1	ADE75218.1	ACU44879.1
ACQ63288.1	ABY75121.1	ABI49419.1	ADE75207.1	ACU44877.1
ACQ63259.1	ABY75119.1	ABL07034.1	ADE75196.1	ACU44875.1
ACQ63250.1	ABY75117.1	ABL07023.1	ADE75187.1	ACU44873.1
ACQ63217.1	ABY75115.1	ABL07012.1	ADE75178.1	ACU44871.1
ACQ63211.1	ABY75113.1	ACC55276.2	ADE75170.1	ACU44869.1
ACQ55364.1	ABY75111.1	ABV53559.1	ADE75152.1	ACU44867.1

Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank
ACQ55353.1	ABY75109.1	AEB71385.1	ADE75143.1	ACU44865.1
ACP44171.1	ABY75107.1	AEB66897.1	ADE75134.1	ACU44863.1
ACP44160.1	ABY75103.1	AEB40208.1	ADE75124.1	ACU44861.1
ACP44153.1	ABY75101.1	ADX36111.1	ADE75115.1	ACU44859.1
ACP44149.1	ABY75099.1	ADX21100.1	ADE75095.1	ACU44857.1
ACR18951.2	ABY75097.1	ADX21090.1	ADE75085.1	ACU44855.1
AAAY87421.1	ABY75095.1	ADX21080.1	ADE75075.1	ACU44853.1
AAAY87413.1	ABY75093.1	ADW93762.1	ADE75057.1	ACU44851.1
ACQ63284.1	ABY75091.1	ADW82270.1	ADE75046.1	ACU44849.1
ACQ63275.1	ABY75089.1	ADW82260.1	ADE75030.1	ACU44847.1
ACQ63266.1	ABY75087.1	ADW82250.1	ACL11961.1	ACU44845.1
ACQ63225.1	ABY75085.1	ADW82240.1	ABY40439.1	ACU44843.1
ACP44185.1	ABY75083.1	ADW82230.1	ABY40432.1	ACU44841.1
ACP44178.1	ABY75081.1	ADW82220.1	AAD25212.1	ACU44839.1
ACA28776.1	ABY75079.1	ADW82210.1	AAD25206.1	ACU44837.1
ACA28772.1	ABY75077.1	ADW82200.1	AAD25172.1	ACU44835.1
ACA28768.1	ABY75075.1	ADW82190.1	BAF36962.1	ACU44833.1
ACR49240.1	ABY75073.1	ADW82179.1	ABI94583.1	ACU44831.1
ACQ84453.1	ABY75071.1	ADW82168.1	ACT21522.1	ACU44829.1
ACU00946.2	ABY75069.1	ADW82157.1	ABY81638.1	ACU44827.1
ACR46665.1	ABY75067.1	ADW82148.1	ACF40971.1	ACU44825.1
ACZ81655.1	ABY75065.1	ADW82137.1	ACD88518.1	ACU44823.1
ACZ81651.1	ABY75063.1	ADW82126.1	ACD88507.1	ACU44821.1
ACR46675.1	ABY75061.1	ADV19021.1	ABW97453.1	ACU44819.1
ACU00956.1	ABY75059.1	ADL41167.1	ACZ81646.1	ACU44817.1
ACU00936.1	ABY75057.1	AAF74337.1	YP_308670.1	ACU44815.1
ACT21587.1	ABY75055.1	ACS92616.1	AAA56808.1	ACU44813.1

Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank
ACT21581.1	ABY75053.1	ACC94117.1	AAA56806.1	ACU44811.1
ACT21576.1	ABY75051.1	ACC94089.1	ABS00311.1	ACU44809.1
ACR19302.2	ABY75049.1	ACC94087.1	ABS00320.1	ACU44807.1
ACR19300.2	ABY75047.1	ACC94085.1	ACR08491.1	ACU44805.1
ACR19298.2	ABY75045.1	ACC94071.1	ACR01010.1	ACU44803.1
ACR19296.2	ABY75043.1	ACC94067.1	ACR01006.1	ACU44801.1
ABX10529.1	ABY75041.1	ACC94065.1	ACG80612.1	ACU44799.1
ABJ90284.2	ABY75035.1	ACC94059.1	ABG78553.1	ACU44797.1
ABJ90273.2	ABY75033.1	ACC94057.1	ABG78550.1	ACU44795.1
ABJ90230.1	ABY75031.1	ACC94051.1	ACD37773.1	ACU44793.1
ADD21567.1	ABY75029.1	ACC94041.1	ACD37763.1	ACU44791.1
ACU44924.1	ABY75027.1	ACC94033.1	ACA64013.1	ACU44789.1
ACU44779.1	ABY75025.1	ABW97496.1	ABX10519.1	ACU44787.1
ACU44773.1	ABY75023.1	ACA28780.1	ABW95953.1	ACU44785.1
ADE48138.1	ABY75021.1	ACA28778.1	ABW95942.1	ACU44783.1
ACG80349.1	ABY75019.1	ACA28774.1	ABJ90263.1	ACU44781.1
ADN34731.1	ABY75017.1	ACA28770.1	ABJ90251.1	ACU44777.1
ADN34711.1	ABY75015.1	ACA28766.1	ABJ90241.1	ACU44775.1
ADG59534.1	ABY75013.1	ACZ81636.1	BAF38386.1	ACU44771.1
ADG59532.1	ABY75011.1	ACU27045.1	BAF37824.1	ACU44769.1
ADG59530.1	ABY75009.1	ACR54040.1	BAF33431.1	ACU44767.1
ACX43975.1	ABY75007.1	ACH68522.1	BAF33417.1	ACU44765.1
ACX43973.1	ABY75005.1	ACF04730.1	BAF33412.1	ACU44763.1
ABG91471.1	ABY75003.1	ACF04728.1	BAF33401.1	ACU44761.1
ABG91467.1	ABY75001.1	ACF04726.1	ACN22341.1	ACU44759.1
ABF21313.1	ABY74999.1	ACF04724.1	ACV49525.1	ACU44757.1
ABF21301.1	ABY74997.1	ACF04722.1	ACV49503.1	ACU44755.1

Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank
ABF21299.1	ABY74995.1	ACC69091.1	ACU79906.1	AEA92622.1
ABF21297.1	ABY74993.1	ABV53579.1	ACU79895.1	AEA35548.1
ABQ57382.1	ABY74991.1	ABV53569.1	ACU79884.1	ADM29588.1
ACR09357.1	ABY74989.1	ABV53549.1	ACU79873.1	ADM29577.1
ACQ99606.1	ABY74987.1	ABV53539.1	ACI25792.1	BAK08628.1
ACQ99598.1	ABY74985.1	ABV53529.1	ACI25781.1	BAK08626.1
ACQ99596.1	ABY74983.1	ABV53519.1	ACI25770.1	ADZ75331.1
ACU43624.2	ABV45404.1	ABV53509.1	ACI25759.1	ADZ75320.1
ACR67240.1	AAC63486.1	ABV53499.1	ACI25748.1	ADP07242.1
ACR67238.1	AAC63484.1	ABV53489.1	ACF54468.1	ACZ54004.1
ACR67235.1	AAC63482.1	ABV53479.1	ACF54457.1	ACX93288.1
ACR67234.1	AAC63480.1	ABV53470.1	ACF54446.1	ACX93222.1
ACR67232.1	ABB00355.1	ADP37370.1	ACF54435.1	ACD65198.1
ACR67230.1	ABB00353.1	ADG21464.1	ACF54424.1	ACD65196.1
ACR67228.1	ABB00349.1	ADG21457.1	ACF54413.1	ACD65194.1
ACR67226.1	ABB00347.1	ACF17953.1	ACF54402.1	ACD65191.1
ACR67224.1	ABB00345.1	ACF17943.1	ACF41825.1	ACD65189.1
ACR67222.1	ABB00343.1	ADM95569.1	ACF41814.1	ACX93277.1
ACR67220.1	ABB00341.1	ADM95567.1	ACF41803.1	ACX93269.1
ACR67218.1	ABB00337.1	ADM95565.1	ACF41792.1	ABD79034.1
ACR67216.1	ABB00335.1	ADM95563.1	ACF41781.1	ABJ16853.1
ACR67214.1	ABB00333.1	ADM95561.1	ACF41770.1	ABJ16842.1
ACR67212.1	CAA30889.1	ADM95559.1	ACF41759.1	ADI99547.1
ACR67208.1	CAA30887.1	ADM95557.1	ACF41748.1	ADI99536.1
ACR67206.1	CAA30885.1	ADM95555.1	ACF41737.1	ADI34045.1
ACR54054.1	CAA30893.1	ADM95553.1	ACF22399.1	ADD21471.1
ACR09363.1	CAA30891.1	ADM95551.1	ACF22388.1	ADD21461.1

Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank
ACP41961.1	AAA43091.1	ADM95549.1	ACF22377.1	ADD21451.1
ACP41938.1	AAA43577.1	ADM95547.1	ACF22366.1	ACZ48112.1
ABI36484.1	BAB19808.1	ADM95545.1	ACF22355.1	ACF25678.1
ABI36475.1	CAA30883.1	ADM95543.1	ACF22344.1	ACF25666.1
ABI36464.1	AAD51929.1	ADM95541.1	ACF22333.1	ACF25466.1
ABI36456.1	BAA99398.1	ADM95539.1	ACF22322.1	ACF25065.1
ABI36445.1	AAF99673.1	ADM95537.1	ACF22311.1	ACF25057.1
ABI36434.1	AAF99671.1	ADM95535.1	ACF22300.1	ACF24971.1
ABI36425.1	ABH04389.1	ADM95533.1	ACF22278.1	ACV74288.1
ABI36412.1	ABB90274.1	ADM95531.1	ACF22256.1	ACV74286.1
ABI36401.1	AAD51270.1	ADM95529.1	ACF22245.1	ABV56243.1
ABI36390.1	AAM70004.1	ADM95527.1	ACF22234.1	ACI25712.1
ABI36379.1	AAB19772.1	ADM95525.1	ACF22223.1	ACI25710.1
ABI36368.1	AAD51266.1	ADM95523.1	ACF22212.1	ADA81213.1
ABI36357.1	ABD59885.1	ADM95521.1	ACF22201.1	ACZ56084.1
ABI36346.1	AAM70001.1	ADM95519.1	ACF22190.1	ACZ45024.1
ABI36335.1	AAM69992.1	ADM95517.1	ACF22172.1	ACV91683.1
ABI36324.1	AAM69982.1	ADM95515.1	ADL41185.1	ACV91679.1
ABI36313.1	AAM69972.1	ADM95513.1	ADM07115.1	ACV91675.1
ABI36297.1	AAM69961.1	ADM95511.1	ADM07104.1	ACV72405.1
ABI36277.1	AAZ38741.1	ADM95509.1	ADM07093.1	ACV72403.1
ABI36202.1	AAZ38739.1	ADM95507.1	ADM07082.1	ACV72401.1
ABI36191.1	AAZ38737.1	ADM95505.1	ADM07071.1	ACV72399.1
ABI36181.1	AAZ38735.1	ADM95503.1	ADM07060.1	ACV72397.1
ABI36170.1	AAZ38733.1	ADM95501.1	ADE62289.1	ACV72395.1
ABI36159.1	AAZ38731.1	ADM95499.1	AEG65177.1	ACV72393.1
ABI36148.1	AAZ38729.1	ADM95497.1	AEC46386.1	ACV72391.1

Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank
ABI36083.1	ABA42442.1	ADM95495.1	ADR78653.1	ACV72389.1
ABI36081.1	ABA42440.1	ADM95493.1	ADM29643.1	ACV72349.1

Bảng 2: Các số truy cập các trình tự axit amin của protein kênh proton BM2

Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank	Số truy cập Genbank
AAU01002.1	ACR15701.1	ACA96576.1	ABN50549.1
BAC54010.1	ACR15690.1	ACA96565.1	ABN50538.1
BAC53999.1	ACR15679.1	ACA96554.1	ABN50527.1
P0C0X4.1	ACR15668.1	ACA65099.1	ABN50516.1
P03493.2	ACR15657.1	ACA65088.1	ABN50505.1
P08383.2	ACR15646.1	ACA65077.1	ABN50494.1
P13882.2	ACR15635.1	ACA65066.1	ABN50483.1
P13881.2	ACR15624.1	ACA65055.1	ABN50472.1
Q80DN6.1	ACO94663.1	ACA65044.1	ABN50450.1
ABF21319.1	ACO06025.1	ACA65033.1	ABN50439.1
ABN50461.1	ACO06014.1	ACA65022.1	ABN50428.1
YP_419283.1	ACO06003.1	ACA65011.1	ABN50417.1
ACN32784.1	ACO05992.1	ACA65000.1	ABN50406.1
ACN32773.1	ACO05981.1	ACA64989.1	ABN50395.1
ACN32719.1	ACO05970.1	ACA64978.1	ABN50384.1
ACN32613.1	ACO05959.1	ACA64967.1	ABL77389.1
ACN32602.1	ACO05937.1	ACA64956.1	ABL77378.1
ACN32591.1	ACO05926.1	ACA64945.1	ABL77367.1
ACN32580.1	ACF54369.1	ACA64934.1	ABL77356.1
ACN32569.1	ACF54358.1	ACA64923.1	ABL77345.1

ACN32558.1	ACF54347.1	ACA64912.1	ABL77334.1
ABL77103.1	ACF54336.1	ACA64901.1	ABL77323.1
ABN50725.1	ACF54325.1	ABR16019.1	ABL77312.1
ABX71689.1	ACF54314.1	ABR16008.1	ABL77301.1
ABF21321.1	ACF54303.1	ABR15997.1	ABL77290.1
AAD29209.1	ACF54292.1	ABR15986.1	ABL77279.1
AAD29207.1	ACF54281.1	ABR15975.1	ABL77268.1
AAD29205.1	ACF54270.1	ABO72379.1	ABL77257.1
AAD29203.1	ACF54259.1	ABN50637.1	ABL77246.1
AAD29201.1	ACF54248.1	ABN59447.1	ABL77235.1
AAD29199.1	ACF54226.1	ABN58663.1	ABL77224.1
AAD29197.1	ACF54215.1	ABN51197.1	ABL77213.1
AAD29195.1	ACF54204.1	ABN51186.1	ABL77202.1
AAD29193.1	ACF54182.1	ABN50747.1	ABL77191.1
AAD29191.1	ACF54160.1	ABN50736.1	ABL77180.1
AAD29189.1	ACF54149.1	ABN50714.1	ABL77169.1
AAD29185.1	ACF54138.1	ABN50703.1	ABL77158.1
AAD29183.1	ACF41660.1	ABN50692.1	ABL77147.1
AAD29181.1	ACD56579.1	ABN50681.1	ABL77136.1
AAD29179.1	ACD56568.1	ABN50670.1	ABL77125.1
AAD29177.1	ACB06477.1	ABN50659.1	ABL77114.1
AAD29175.1	ACA96664.1	ABN50648.1	ABL77092.1
AAD29173.1	ACA96653.1	ABN50626.1	ABL77081.1
AAT69452.1	ACA96642.1	ABN50615.1	ABL77070.1
AAT69441.1	ACA96631.1	ABN50604.1	ABL77059.1
AAT69430.1	ACA96620.1	ABN50593.1	ABL77048.1
ACR39338.1	ACA96609.1	ABN50582.1	ABL77037.1
ACR15734.1	ACA96598.1	ABN50571.1	ABL77026.1
ACR15723.1	ACA96587.1	ABN50560.1	ABL77015.1

ACR15712.1	ACA96576.1	ABN50549.1	ABL77004.1
------------	------------	------------	------------

Protein virut cấu trúc

Protein virut cấu trúc có thể là protein kháng nguyên virut hoặc mảnh của nó, ví dụ, nhưng không giới hạn ở glycoprotein virut hoặc protein vỏ virut. Protein virut cấu trúc có thể protein virut thể khảm. Protein virut có thể tồn tại ở dạng monome, dime, trime, hoặc tổ hợp của chúng. Trime là phức hợp đại phân tử được tạo thành bằng ba protein được gắn kết không cộng hóa trị. Không mong muốn bị ràng buộc bởi lý thuyết, miền trime hóa của protein có thể quan trọng để tạo thành các trime này. Do đó, protein virut cấu trúc hoặc mảnh của nó có thể bao gồm miền trime hóa. Ví dụ không giới hạn của protein virut cấu trúc là hemagglutinin cúm (HA), hoặc mảnh của HA. Các ví dụ không giới hạn của HA, hoặc các mảnh của HA có thể được sử dụng theo sáng chế bao gồm các mảnh được mô tả trong các công bố đơn sáng chế quốc tế số WO2009/009876, WO 2009/076778; WO 2010/003225, WO 2010/003235, WO 2011/03522, WO 2010/006452, WO 2010/148511, WO 2011/035422.

Hơn nữa, protein virut cấu trúc có thể là tiền phân tử protein chưa được xử lý của HA. Protein HA được tổng hợp làm tiền phân tử protein (HA0) có kích thước khoảng 75kDa, lắp ráp tại bề mặt thành protein trime được kéo dài. Tiền phân tử protein được phân cắt ở vị trí phân cắt hoạt hóa bảo toàn thành 2 chuỗi polypeptit, HA1 và HA2 (bao gồm vùng chuyển màng), được liên kết bằng liên kết disulfua.

Biến đổi vòng phân hủy protein (vị trí phân cắt)

Protein virut cấu trúc có thể là protein hemagglutinin cúm B hoặc hemagglutinin cúm A có sự xóa bỏ hoặc biến đổi vòng phân hủy protein (vị trí phân cắt) trong protein hemagglutinin. Việc xóa bỏ hoặc biến đổi vòng phân hủy protein đảm bảo rằng phân tử HA hầu như được duy trì ở dạng tiền phân tử HA0.

HA được tổng hợp ở dạng tiền phân tử protein HA0, trải qua quá trình phân hủy protein thành hai các tiểu đơn vị (HA1 và HA2) được liên kết với nhau bằng cầu disulfua. Các chủng virut cúm gia cầm không có khả năng gây bệnh và các chủng virut cúm động

vật có vú gây ra các lây nhiễm cục bộ về phương diện giải phẫu do khoảng giới hạn các tế bào tiết proteaza mà có thể phân cắt tiền phân tử HA0 ngoại bào (Chen J, et. al. 1998, Cell. Vol 95 :409-417). Các proteaza chịu trách nhiệm phân cắt HA0 trong các lây nhiễm cúm của người, được tiết bằng các tế bào của đường hô hấp, hoặc bằng cách lây nhiễm đồng thời vi khuẩn hoặc mycoplasma, hoặc chúng có thể được tạo ra trong các phản ứng viêm với các lây nhiễm. Ứng viên proteaza chính là tryptaza Clara, được tạo ra bởi các tế bào Clara của biểu mô phổi, và sự phân chia mô giới hạn (đường hô hấp phía trên). Proteaza đặc hiệu với trình tự đơn bazơ Q/E-X-R được tìm thấy tại vị trí phân cắt của H1, H2, H3, và H6. Các chủng HA từ H9 và B thể hiện vị trí phân cắt đơn bazơ hơi khác nhau với trình tự tương ứng là SSR và KER (xem Fig. 24). Không có proteaza nào được xác định cho phần lớn virus cúm gây ra lây nhiễm qua đường ruột và đường hô hấp được tìm thấy ở các loài chim thủy sinh. Trong phòng thí nghiệm, hầu hết các dòng tế bào không hỗ trợ sao chép đa chu kỳ trừ trường hợp bổ sung thêm proteaza ngoại sinh (thường là trypsin).

Tuy nhiên, các chủng gia cầm gây bệnh ở mức cao được phân cắt bằng họ gồm các proteaza nội bào phổ biến hơn, dẫn đến các lây nhiễm hệ thống. Sự khác biệt về tính gây bệnh này tương quan với những khác biệt về cấu trúc tại vị trí phân cắt HA0. Các chủng gây bệnh có các đoạn chen của các axit amin đa bazơ bên trong, hoặc bên cạnh, vị trí đơn bazơ. Sự phân cắt trong trường hợp này xuất hiện nội bào và các proteaza liên quan được xác định là furin, và các enzym giống như subtilisin khác, được tìm thấy trong Golgi và liên quan đến quá trình xử lý sau dịch mã của hormon và các tiền phân tử của yếu tố sinh trưởng. Trình tự nhận biết furin R-X-R/K-R là axit amin chen thường xuyên tại các vị trí phân cắt HA0 của H5 và H7 (xem Fig.24). Sự phân bố mô rộng của enzym này, và hiệu quả phân cắt nội bào, góp phần lây nhiễm hệ thống do virus và trên diện rộng do các virus này gây ra.

Horimoto T, et.al. (2006, Vaccin, Vol 24 : 3669-3676) mô tả sự xóa bỏ vị trí phân cắt đa bazơ của H5 (RERRRKKR↓G) ở H5. Các thể đột biến được chọn được đưa vào nghiên cứu tính sinh miễn dịch ở chuột, bao gồm thể đột biến bị xóa bỏ 4 axit amin mang điện tích đầu tiên (RERR) và biến đổi để bất hoạt vị trí phân cắt đa bazơ (RKRR với

TETR). Sự xóa bỏ vị trí phân cắt này không ảnh hưởng đến các tính chất sinh miễn dịch của thể đột biến H5. Sự xóa bỏ vị trí đa bazơ (GERRRKKR↓G được thay thế bằng RETR) để tạo ra virus cúm đối chứng NIBG-23 thể đột biến NIBSC 05/240 NIBSC cũng đã được báo cáo. Hoffman et. al. (2002, 2002, *Vaccin*, Vol 20 :3165-3170) đã thay thế vị trí phân cắt đa bazơ của H5 HA bằng vị trí đơn bazơ của H6 để thúc đẩy sự biểu hiện ở trứng. 4 gốc đầu tiên bị xóa bỏ và thay thế 4 axit amin cuối cùng của vị trí đa bazơ bằng IETR (thay thế RERRRKKR↓G bằng IETR↓G). Chất đột biến H5 này thể hiện mức biểu hiện cao, sự phân hủy protein tiềm năng và sự thay đổi cấu dạng ở độ pH thấp, dữ liệu sinh miễn dịch không được báo cáo. Các nghiên cứu này thể hiện rằng sự biến đổi vị trí phân cắt có thể được sử dụng để làm giảm độc lực của các hạt virus (trong các trường hợp mà các virus thật được sao chép, cho phép virus sao chép mà không giết trứng của vật chủ. Không có các đột biến này, các virus giết trứng trước khi đạt tới hàm lượng cao.

Trong quá trình cuộn gập HA và tiết hoàn toàn qua Golgi, vị trí phân cắt tiền phân tử hemagglutinin, mà nằm trên vòng tại bề mặt của HA, có thể tiếp cận dễ dàng bằng sự phân hủy protein bằng các proteaza. Không mong muốn bị ràng buộc bởi lý thuyết, nếu sự phân hủy protein của tiền chất HA0 xảy ra tại vị trí đơn bazơ hoặc đa bazơ trong quá trình cuộn gập của HA trong ER, sự thay đổi cấu dạng của protein có thể xảy ra trong bộ máy Golgi trong quá trình tiết, do môi trường pH bên trong Golgi của thực vật và trong apoplast mang tính axit nhẹ. Cấu dạng HA ở độ pH thấp có thể được tạo ra, làm giảm cả mức biểu hiện và sự ổn định bên trong của hạt. Do đó, hầu hết protein tiền phân tử HA0 chưa được phân cắt sẽ bắt đầu nảy chồi (budding) từ màng sinh chất.

“Vòng phân hủy protein” hoặc “vị trí phân cắt” nghĩa là trình tự liên ứng của vị trí phân hủy protein mà liên quan đến sự phân cắt tiền phân tử HA0. Như được sử dụng ở đây, “liên ứng” hoặc “trình tự liên ứng” nghĩa là trình tự (axit amin hoặc trình tự nucleotit) bao gồm khả năng biến đổi trình tự của các trình tự liên quan dựa trên sự phân tích sự sắp xếp của nhiều trình tự, ví dụ, các kiểu phụ của trình tự HA0 cúm cụ thể. Trình tự liên ứng của vị trí phân cắt cúm HA0 có thể bao gồm các trình tự axit amin cúm A hemagglutinin liên ứng, bao gồm, ví dụ, H1 liên ứng, H3 liên ứng, hoặc các trình tự axit amin

hemagglutinin liên ứng cúm B. Các ví dụ không giới hạn của các trình tự liên ứng được thể hiện trong Fig.24.

Trong trình tự axit amin của HA, vòng phân hủy protein được định vị, trước khi peptit dung hợp mà bao gồm 20 axit amin đầu tiên của phần HA2. Cấu trúc tinh thể của HA0 từ A/Hong Kong/68 đã được xác định (Chen, J., 1998. Cell 95:409-417). Các gốc được tiếp xúc với dung môi mà thường được cho là một phần của vị trí phân cắt mà tạo thành vòng bề mặt được tiếp xúc ở mức cao, mở rộng. Từ trình tự peptit đặc hiệu này, trình tự liên ứng có thể được xác định trong vùng được chọn này (Bianchi et al., 2005, Journal of Virology, 79:7380-7388).

Để xóa bỏ vòng phân hủy protein, cấu trúc của B HA đã được kiểm tra. Việc xóa bỏ duy nhất vị trí phân cắt phân hủy protein của HA sẽ để lại đầu C của HA1 và đầu N của HA2 còn lại một phần và phần liên kết dài sẽ cần được thiết kế. Tuy nhiên, việc xóa bỏ một phần peptit dung hợp cùng với vị trí phân cắt protein thủy phân cho phép loại bỏ vòng phân hủy protein hoàn chỉnh và kết hợp trình tự HA1 và HA2 còn lại bằng phần liên kết peptit tối thiểu gồm 2 axit amin. Tóm lại, biến thể B bao gồm việc xóa bỏ đoạn trình tự ALKLLKER tại đầu C của HA1 ngoài việc xóa bỏ các axit amin đầu N GFFGAIAGFLEG của HA2. HA1-HA2 được thu ngắn lại được liên kết với nhau bằng phần liên kết GG.

Như được thể hiện trong Fig.22B, bằng cách xóa bỏ vòng phân hủy protein của HA0, protein HA0 thu được có hoạt tính tăng như được thể hiện bằng khả năng hemagglutinin hóa lớn hơn, khi được so sánh với protein HA mà không có vòng phân hủy protein của nó bị loại bỏ. Bằng cách tăng hoạt tính, nghĩa là tăng khả năng hemagglutinin hóa nằm trong khoảng từ 2% đến 100%, hoặc lượng bất kỳ nằm trong khoảng này như được xác định bằng cách sử dụng các kỹ thuật chuẩn trong lĩnh vực này, ví dụ, nằm trong khoảng từ 10% đến 50% hoặc giá trị bất kỳ nằm trong đó, ví dụ, khoảng 2, 5, 8, 10, 12, 15, 18, 20, 22, 24, 25, 26, 28, 30, 32, 34, 35, 36, 38, 40, 42, 44, 45, 46, 48, 50, 52, 54, 55, 56, 58, 60, 65, 70, 75, 80, 85, 90, 95, hoặc 100%, khi được so sánh với hoạt tính của cùng HA nhưng không có vòng phân hủy protein của nó bị loại bỏ.

“Protein virus thể khảm” hoặc “polypeptit virus thể khảm”, còn được gọi là “protein thể khảm” hoặc “polypeptit thể khảm” nghĩa là protein hoặc polypeptit bao gồm các trình tự axit amin từ hai hoặc nhiều hơn hai nguồn, ví dụ, nhưng không giới hạn ở, hai hoặc nhiều typ hoặc typ phụ cúm, hoặc các cúm có nguồn gốc khác nhau, mà được dung hợp làm một polypeptit. Protein thể khảm hoặc polypeptit có thể bao gồm peptit tín hiệu giống, hoặc khác với, phần còn lại của polypeptit hoặc protein. Protein thể khảm hoặc polypeptit thể khảm có thể được tạo ra ở dạng sản phẩm phiên mã từ trình tự nucleotit thể khảm, và protein thể khảm hoặc polypeptit thể khảm được phân cắt sau tổng hợp, và khi được yêu cầu, được kết hợp để tạo thành protein multimer. Do đó, protein thể khảm hoặc polypeptit thể khảm còn bao gồm protein hoặc polypeptit bao gồm các tiểu đơn vị mà được kết hợp thông qua các cầu disulphua (nghĩa là protein multimer). Ví dụ, polypeptit thể khảm bao gồm các trình tự axit amin từ hai hoặc nhiều hơn hai nguồn có thể được xử lý thành các tiểu đơn vị, và các tiểu đơn vị này được kết hợp thông qua các cầu disulphua để tạo ra protein thể khảm hoặc polypeptit thể khảm. Protein virus thể khảm cũng có thể bao gồm protein kháng nguyên hoặc mảnh của nó của virus cúm thứ nhất, và phức miễn xuyên màng (TDC) từ virus cúm thứ hai HA, bao gồm miễn xuyên màng và các miền đuôi tế bào chất (TM/CT). Polypeptit có thể là hemagglutinin (HA), và một trong hai hoặc nhiều hơn hai trình tự axit amin mà tạo ra polypeptit có thể thu được từ các trình tự HA khác nhau để tạo ra HA thể khảm, hoặc HA cúm thể khảm. HA thể khảm cũng có thể bao gồm trình tự axit amin bao gồm peptit tín hiệu khác loại (protein HA thể khảm) mà được phân cắt sau hoặc trong quá trình tổng hợp protein. Tốt hơn là, polypeptit thể khảm, hoặc HA cúm thể khảm này không xuất hiện trong tự nhiên. Axit nucleic mã hóa polypeptit thể khảm có thể được mô tả là "axit nucleic thể khảm", hoặc "trình tự nucleotit thể khảm". Hạt giống như virus bao gồm HA thể khảm có thể được mô tả là “VLP thể khảm”.

Protein hoặc polypeptit thể khảm có thể bao gồm peptit tín hiệu giống, hoặc khác loại với, phần còn lại của polypeptit hoặc protein. Thuật ngữ "peptit tín hiệu" đã được biết trong lĩnh vực này và thường đề cập đến trình tự ngắn các axit amin (khoảng từ 5 đến 30 axit amin), thường được tìm thấy tại đầu N của polypeptit mà có thể hướng sự chuyển vị của polypeptit được dịch mã mới đến bào quan cụ thể, hoặc hỗ trợ định vị các miền đặc hiệu của chuỗi polypeptit so với các miền khác. Như ví dụ không giới hạn, peptit tín hiệu

có thể hướng đến sự chuyển vị protein vào lưới nội chất và/hoặc hỗ trợ định vị miền gần đầu N so với miền neo màng của polypeptit mới sinh ra để hỗ trợ phân cắt và cuộn gập của protein trưởng thành, ví dụ, mà không được xem là giới hạn, protein HA trưởng thành.

Các ví dụ không giới hạn về các protein virus thể khảm hoặc axit nucleic virus thể khảm có thể được sử dụng theo sáng chế được mô tả trong các công bố đơn sáng chế quốc tế số WO 2009/076778, WO 2010/003235, hoặc WO 2010/148511.

Peptit tín hiệu

Peptit tín hiệu (SP) có thể là nguyên bản so với protein kháng nguyên hoặc protein virus, hoặc peptit tín hiệu có thể khác loại với trình tự thứ nhất của protein kháng nguyên hoặc protein virus được biểu hiện. Protein kháng nguyên hoặc protein virus có thể bao gồm peptit tín hiệu từ typ, typ phụ, chủng cúm thứ nhất có sự cân bằng HA từ một hoặc nhiều typ, typ phụ hoặc chủng cúm khác nhau. Ví dụ, peptit tín hiệu tự nhiên của các typ phụ HA là H1, H2, H3, H5, H6, H7, H9 hoặc typ cúm B có thể được sử dụng để biểu hiện protein virus thể khảm trong hệ thực vật. Theo một số phương án của sáng chế, SP có thể thuộc typ cúm B, H1, H3 hoặc H5; hoặc typ phụ H1/Bri, H1/NC, H5/Indo, H3/Bri hoặc B/Flo.

Peptit tín hiệu cũng có thể không nguyên bản, ví dụ, từ protein kháng nguyên, protein virus hoặc hemagglutinin của virus khác loại với protein virus, hoặc từ polypeptit thực vật, động vật hoặc vi khuẩn. Ví dụ không giới hạn của peptit tín hiệu có thể được sử dụng là của protein disulfua isomerase của cỏ linh lăng ("PDISP"; các nucleotit 32-103 của số truy nhập Z11499; còn xem các công bố đơn sáng chế quốc tế số WO 2009/076778; WO 2010/148511, hoặc WO 2010/003235). Do đó, sáng chế đề xuất protein virus thể khảm bao gồm peptit tín hiệu nguyên bản hoặc không nguyên bản, và các axit nucleic mã hóa các protein virus thể khảm này.

Do đó, sáng chế đề xuất phương pháp sản xuất VLP thể khảm ở thực vật, trong đó axit nucleic thứ nhất mã hóa protein virus thể khảm được biểu hiện đồng thời với axit nucleic thứ hai mã hóa protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton.

Axit nucleic thứ nhất và thứ hai có thể được đưa vào thực vật trong cùng bước, hoặc có thể được đưa vào thực vật một cách lần lượt.

HA

Tham chiếu đến virus cúm, như được sử dụng ở đây, thuật ngữ “hemagglutinin” hoặc “HA” đề cập đến glycoprotein được tìm thấy bên ngoài các hạt virus cúm. HA là glycoprotein typ I màng homotrimer, thường bao gồm peptit tín hiệu, miền HA1, và miền HA2 bao gồm vị trí neo mở rộng màng tại đầu C và đuôi tế bào chất nhỏ. Các trình tự nucleotit mã hóa HA đã được biết và có sẵn— xem, ví dụ, BioDefence Public Health base (Influenza Virus; xem đường dẫn: biohealthbase.org) hoặc National Center for Biotechnology Information (xem đường dẫn: ncbi.nlm.nih.gov).

Thuật ngữ “homotrimer” chỉ ra rằng oligome được tạo thành bằng ba phân tử protein HA. Không mong muốn bị ràng buộc bởi lý thuyết, protein HA được tổng hợp làm protein tiền phân tử monome (HA0) có kích thước khoảng 75 kDa trong các tế bào động vật, mà lắp ráp tại bề mặt thành protein trimer được kéo dài. Trước khi trimer hóa xảy ra, protein tiền phân tử được phân cắt tại vị trí phân cắt hoạt hóa bảo toàn (còn được gọi là peptit dung hợp) thành 2 chuỗi polypeptit, HA1 và HA2 (bao gồm vùng chuyển màng), được liên kết bằng liên kết disulfua. Đoạn HA1 có thể dài 328 axit amin, và đoạn HA2 có thể dài 221 axit amin. Mặc dù sự phân cắt này có thể quan trọng với tính lây nhiễm virus, có thể không quan trọng với sự trimer hóa protein. Việc chèn HA trong màng lưới nội chất (ER) của tế bào vật chủ, sự phân cắt peptit tín hiệu và glycosyl hóa protein là các sự kiện dịch mã đồng thời. Sự cuộn gập lại chính xác của HA yêu cầu sự glycosyl hóa protein và tạo thành 6 liên kết disulfua nội chuỗi. Trimer HA lắp ráp bên trong phức cis- và trans-Golgi, miền xuyên màng đóng vai trò trong quá trình trimer hóa. Các cấu trúc tinh thể của các protein HA được xử lý bromelain, mà thiếu miền xuyên màng, thể hiện cấu trúc bảo toàn cao trong số các chủng cúm. Cũng đã biết rằng HA trải qua sự thay đổi cấu dạng chính trong suốt quá trình lây nhiễm, yêu cầu tiền phân tử HA0 bị phân cắt thành 2 chuỗi polypeptit HA1 và HA2. Protein HA có thể đã được xử lý (nghĩa là, bao gồm các miền HA1 và HA2), hoặc có thể chưa được xử lý (nghĩa là, bao gồm miền HA0). Protein HA có

thể được sử dụng trong việc tạo ra hoặc tạo thành các VLP bằng cách sử dụng thực vật, hoặc tế bào thực vật, hệ biểu hiện.

HA theo sáng chế có thể thu được từ typ phụ bất kỳ. Ví dụ, HA có thể có typ phụ H2, H3, H4, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, hoặc HA cúm typ B. HA tái tổ hợp theo sáng chế cũng có thể bao gồm trình tự axit amin dựa trên trình tự hemagglutinin bất kỳ đã biết trong lĩnh vực này— xem, ví dụ, BioDefence Public Health base (Influenza Virus; xem đường dẫn: biohealthbase.org) hoặc National Center for Biotechnology Information (xem đường dẫn: ncbi.nlm.nih.gov). Hơn nữa, HA có thể dựa trên trình tự của hemagglutinin mà được phân lập từ một hoặc nhiều virus cúm được xác định là mới hoặc đang hoành hành.

Các ví dụ không giới hạn về HA, hoặc các mảnh HA có thể được sử dụng theo sáng chế bao gồm được mô tả trong các công bố đơn sáng chế quốc tế số WO2009/009876, WO 2009/076778; WO 2010/003225, WO 2010/003235, WO 2010/006452, WO 2011/035422 hoặc WO 2010/148511.

Như được thể hiện trong Fig.18, HA từ B/Brisbane/60/2008 được biểu hiện kém ở các lá *Nicotiana benthamiana* được thâm bằng *Agrobacteria* (xem các giếng “1008” hoặc “1029”). Tuy nhiên, sự biểu hiện đồng thời HA typ B với M2 từ A/New Caledonia/20/99, dẫn đến tăng đáng kể biểu hiện HA (xem các giếng “1008+1261”; “1009+1261” và 1029+1261”). Việc tăng biểu hiện HA đã được quan sát ở cả hai HA typ B tự nhiên hoặc HA typ B thể khảm. Sự biểu hiện HA đã được quan sát với sự vắng mặt hoặc có mặt của các yếu tố khuếch đại (BeYDV), và thông qua nhiều nồng độ pha loãng *Agrobacterium* khác nhau. Việc tăng tương tự trong sự biểu hiện H3 đã được quan sát khi H3 từ A/Perth/16/2009 được biểu hiện đồng thời với M2 từ A/New Caledonia/20/99 (Fig.19; so sánh giếng “1019” một mình H3, với “1019+1261” H3 được biểu hiện đồng thời với M2).

VLP

Thuật ngữ “hạt giống như virus” (VLP), hoặc “các hạt giống như virus” hoặc “VLP” đề cập đến các cấu trúc tự lắp ráp và bao gồm các protein virus, ví dụ, protein virus cấu trúc như protein cúm HA hoặc protein kênh, ví dụ, nhưng không giới hạn ở, protein kênh

proton, như M2 hoặc tổ hợp của các protein đó. Các VLP thường tương tự về hình thái và tính kháng nguyên với các virion được tạo ra trong quá trình lây nhiễm, nhưng thiếu thông tin di truyền đủ để sao chép và do đó không lây nhiễm. Trong một số ví dụ, các VLP có thể bao gồm các loại protein đơn, hoặc nhiều hơn một loại protein. Đối với các VLP bao gồm một hoặc nhiều loại protein, các loại protein có thể từ cùng loài virus, hoặc có thể bao gồm protein từ các loài, chi, phân họ hoặc họ khác nhau của virus (như được ký hiệu bằng danh pháp ICTV). Trong các ví dụ khác, một hoặc nhiều loại protein bao gồm VLP có thể được biến đổi từ trình tự xuất hiện trong tự nhiên. Các VLP có thể được tạo ra trong các tế bào vật chủ thích hợp bao gồm các tế bào vật chủ thực vật và côn trùng. Sau khi chiết từ tế bào vật chủ và phân lập và tinh chế tiếp trong các điều kiện thích hợp, các VLP có thể được tinh chế ở dạng các cấu trúc nguyên vẹn.

Hơn nữa, các VLP có thể được tạo ra mà bao gồm tổ hợp của các typ phụ của HA. Ví dụ, các VLP có thể bao gồm một hoặc nhiều HA từ typ phụ H2, H3, H4, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, HA typ phụ B hoặc tổ hợp của chúng. Sự lựa chọn của tổ hợp các HA có thể được xác định nhờ việc sử dụng theo dự định vacxin được sản xuất từ VLP. Ví dụ, vacxin để sử dụng trong tiêm chủng cho chim có thể bao gồm tổ hợp bất kỳ của các typ phụ HA, trong khi các VLP hữu dụng để tiêm chủng cho người bao gồm các typ phụ một hoặc nhiều typ phụ H2, H3, H4, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, HA typ phụ B. Tuy nhiên, các tổ hợp typ phụ HA khác có thể được sản xuất tùy thuộc vào việc sử dụng VLP. Để sự tạo ra VLP bao gồm các tổ hợp của các typ phụ HA, typ phụ HA mong muốn có thể được biểu hiện đồng thời trong cùng tế bào, ví dụ, tế bào thực vật.

Như được mô tả chi tiết hơn dưới đây, sự biểu hiện HA typ B hoặc H3 từ B/Brisbane/60/2008 ở các lá *Nicotiana benthamiana* được thăm bằng *Agrobacterium* tăng khi biểu hiện đồng thời M2 từ A/New Caledonia/20/99 (xem Fig.18 và 19). Việc tăng tương tự không được quan sát khi H1 hoặc H5 được biểu hiện đồng thời với M2, do mức biểu hiện cao của H1 hoặc H5 được quan sát khi có mặt hoặc vắng mặt M2 (các Fig.20 và 21 tương ứng). HA được biết là được xử lý theo cách phụ thuộc pH (xem Reed M. L. et al. Journal of Virology, tháng 02 năm 2010, trang 1527-1535, Vol. 84, No. 3), và trải qua

sự thay đổi cấu dạng phụ thuộc vào độ pH (Skehel J.J. et. al. 1982, PNAS79: 968-972). H1 và H5 có sự thay đổi cấu dạng ở độ pH thấp hơn độ pH gây ra sự thay đổi cấu dạng được quan sát ở H3 hoặc HA typ B. Không mong muốn bị ràng buộc bởi lý thuyết, độ pH trong bộ máy Golgi trong quá trình trưởng thành và di trú có thể không ảnh hưởng đến sự cuộn gập của H1 hoặc H5, tuy nhiên, độ pH thấp trong bộ máy Golgi có thể ảnh hưởng đến sự cuộn gập của H3 và HA typ B. Bằng cách biểu hiện đồng thời protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton, cùng với H3 hoặc HA typ B, độ pH trong bộ máy Golgi có thể tăng và dẫn đến cuộn gập HA làm tăng hiệu suất HA. Hơn nữa, H1 và H5 có thể ổn định hơn ở độ pH thấp, so với H3 và HA typ B. Do đó, bằng cách biểu hiện đồng thời protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton, cùng với H3 hoặc HA typ B, ít HA bị phân hủy trong bộ máy Golgi.

Theo sáng chế, các VLP được sản xuất từ các protein thu được từ cúm không bao gồm protein M1. Protein M1 được biết là gắn kết ARN (Wakefield và Brownlee, 1989) là chất nhiễm tạp của sự tạo ra VLP. Sự có mặt của ARN là không mong muốn khi xin phê chuẩn lưu hành thuốc đối với sản phẩm VLP, do đó việc sản xuất VLP mà không có mặt ARN có thể là ưu điểm.

Các VLP được sản xuất như được mô tả ở đây thường không bao gồm neuraminidaza (NA). Tuy nhiên, NA có thể được biểu hiện đồng thời với HA nếu VLP bao gồm HA và NA là mong muốn.

Sáng chế cũng bao gồm, nhưng không giới hạn ở, các VLP thu được từ virut mà thu được vỏ lipit từ màng sinh chất của tế bào mà trong đó các VLP được biểu hiện. Ví dụ, nếu VLP được biểu hiện trong hệ có nguồn gốc thực vật, VLP có thể thu được vỏ lipit từ màng sinh chất của tế bào.

Nhìn chung, thuật ngữ “lipit” đề cập đến các phân tử xuất hiện trong tự nhiên, có thể hòa tan chất béo (ưa chất béo). Thuật ngữ này cũng được sử dụng cụ thể hơn đề cập đến các axit béo và các dẫn xuất của chúng (bao gồm các tri-, di-, và monoglyxerit và phospholipit), cũng như các chất trao đổi chứa sterol hoặc sterol có thể hòa tan trong chất béo. Phospholipit là thành phần chính của tất cả các màng sinh học, cùng với các

glycolipit, sterol và protein. Các ví dụ về phospholipit bao gồm phosphatidyletanolamin, phosphatidylcholin, phosphatidylinositol, phosphatidylserin, và phospholipit tương tự. Các ví dụ về sterol bao gồm zoosterol (ví dụ, cholesterol) và phytosterol. Hơn 200 phytosterol đã được nhận biết ở các loài thực vật khác nhau, phổ biến nhất là campesterol, stigmasterol, ergosterol, brassicasterol, delta-7-stigmasterol, delta-7-avenasterol, daunosterol, sitosterol, 24-metylcholesterol, cholesterol hoặc beta-sitosterol. Vì người có hiểu biết trung bình trong lĩnh vực này sẽ hiểu rằng, thành phần lipit của màng sinh chất của tế bào có thể khác với môi trường nuôi cấy hoặc các điều kiện sinh trưởng của tế bào hoặc sinh vật từ tế bào này thu được.

Màng tế bào thường bao gồm lớp kép lipit, cũng như các protein với các chức năng khác nhau. Các nồng độ cục bộ của các lipit cụ thể có thể tìm thấy trong lớp kép lipit, được gọi là 'bè lipit'. Không mong muốn bị ràng buộc bởi lý thuyết, các bè lipit có thể có vai trò đáng kể trong nhập bào và xuất bào, vào hoặc ra của virus hoặc các chất lây nhiễm khác, tải nạp tín hiệu bên trong tế bào, tương tác với các thành phần cấu trúc khác của tế bào hoặc sinh vật, như các màng nội bào và ngoại bào.

Ở thực vật, các VLP cúm này chỗi thoát ra từ màng sinh chất, do đó thành phần lipit của VLP phản ánh nguồn gốc của chúng. Các VLP được tạo ra theo sáng chế bao gồm HA của một hoặc nhiều typ hoặc typ phụ cúm, được tạo phức với lipit thu được từ thực vật. Các lipit thực vật có thể kích thích các tế bào miễn dịch và tăng cường đáp ứng miễn dịch được kích thích. Màng thực vật được tạo ra từ lipit, phosphatidylcholin (PC) và phosphatidyletanolamin (PE), và cũng chứa glycosphingolipit, saponin, và phytosterol. Ngoài ra, các bè lipit cũng được tìm thấy trong màng sinh chất của thực vật – các miền nhỏ này được làm giàu sphingolipit và sterol. Ở thực vật, nhiều phytosterol khác nhau đã được biết đến, bao gồm stigmasterol, sitosterol, 24-metylcholesterol và cholesterol (Mongrand et al., 2004).

PC và PE, cũng như glycosphingolipit có thể gắn kết với các phân tử CD1 được biểu hiện bằng các tế bào miễn dịch của động vật có vú như các tế bào trình diện kháng nguyên (APC) như các tế bào tua (dendritic cell) và đại thực bào và các tế bào khác bao gồm tế bào lympho B và T trong tuyến ức và gan (Tsuiji M., 2006). Các phân tử CD1 có

cấu trúc tương tự với các phân tử phức tương hợp mô chính (MHC) của lớp I và vai trò của chúng là trình diện các kháng nguyên glycolipit với các tế bào NKT (các tế bào T giết tự nhiên). Nhờ hoạt hóa, các tế bào NKT hoạt hóa các tế bào miễn dịch bẩm sinh như tế bào NK và các tế bào tua và cũng hoạt hóa các tế bào miễn dịch đáp ứng giống như các tế bào B và tế bào T tạo ra kháng thể.

Nhiều phytosterol có thể được tìm thấy trong màng sinh chất – phần bổ sung đặc hiệu có thể khác nhau tùy thuộc vào các loài, điều kiện sinh trưởng, nguồn dinh dưỡng hoặc trạng thái gây bệnh, vào một vài yếu tố. Nhìn chung, beta-sitosterol là phytosterol phổ biến nhất.

Phytosterol có trong VLP cúm được tạo phức với lớp kép lipid, như lớp vỏ thu được từ màng nguyên sinh có thể đề xuất cho chế phẩm vacxin có lợi. Không mong muốn bị ràng buộc bởi lý thuyết, VLP được sản xuất từ thực vật được tạo phức với lớp kép lipid, như lớp vỏ thu được từ màng nguyên sinh, có thể kích thích phản ứng miễn dịch mạnh hơn các VLP được tạo ra ở các hệ biểu hiện khác, và có thể tương tự với phản ứng miễn dịch được kích thích bằng các vacxin virus đầy đủ sống hoặc được làm giảm độc lực.

Do đó, theo một số phương án, sáng chế đề xuất VLP được tạo phức với lớp kép lipid thu được từ thực vật. Theo một số phương án, lớp kép lipid thu được từ thực vật có thể bao gồm vỏ của VLP. Lipit thu được từ thực vật có thể bao gồm các thành phần lipid của màng sinh chất của thực vật mà trong đó VLP được tạo ra, bao gồm, nhưng không giới hạn ở, phosphatidylcholin (PC), phosphatidyletanolamin (PE), glycosphingolipit, phytosterol hoặc tổ hợp của chúng. Lipit thu được từ thực vật có thể được gọi theo cách khác là ‘lipit thực vật’. Các ví dụ về các phytosterol được biết trong lĩnh vực này, và bao gồm, ví dụ, stigmasterol, sitosterol, 24-metylcholesterol và cholesterol – xem, ví dụ tài liệu của Mongrand et al., 2004.

Các VLP có thể được đánh giá cấu trúc và kích thước bằng, ví dụ, thử nghiệm hemagglutinin hóa, kính hiển vi điện tử (electron), hoặc sắc ký loại trừ kích thước.

Đối với sắc ký loại trừ kích thước, tổng protein hòa tan có thể được chiết từ mô thực vật bằng cách đồng nhất (Polytron) mẫu của vật liệu thực vật được nghiền đông lạnh

trong dung dịch đậm chiết, và vật liệu không hòa tan được loại bỏ bằng ly tâm. Kết tủa PEG có thể được sử dụng. Protein hòa tan có thể được định lượng, và dịch chiết đi qua màng loại trừ kích thước, ví dụ nhưng không giới hạn ở Sephacryl™. Sau khi sắc ký, các phân đoạn còn có thể được phân tích bằng lai thẩm tách miễn dịch để xác định thành phần protein của đoạn này.

Không mong muốn bị ràng buộc bởi lý thuyết, khả năng HA gắn kết với RBC từ các động vật khác nhau được điều khiển bởi ái lực của HA với các axit sialic $\alpha 2,3$ hoặc $\alpha 2,6$ và sự có mặt của các axit sialic này trên bề mặt của RBC. HA ngựa và gia cầm từ các virus cúm kết dính các hồng cầu từ tất cả các loài trong số, bao gồm gà tây, gà, vịt, chuột lang, người, cừu, ngựa và bò; trong khi đó các HA của người sẽ gắn kết với các hồng cầu của gà tây, gà, vịt, chuột nang, người và cừu (cũng xem tài liệu của: Ito T. et al, 1997, Virology, vol 227, p493-499; và Medeiros R et al, 2001, Virology, vol 289 p.74-85).

Cuộn gấp (Chaperon)

Sự cuộn gấp chính xác protein virus được biểu hiện có thể quan trọng với tính ổn định của protein, tạo thành các đa monome, tạo thành VLP, chức năng của protein virus và nhận biết protein virus bằng kháng thể, trong số các đặc trưng khác. Sự cuộn gấp và tích tụ protein có thể bị ảnh hưởng bằng một hoặc nhiều yếu tố, bao gồm, nhưng không giới hạn ở, trình tự của protein, độ phong phú tương đối của protein, sự dồn đông nội bào, độ pH trong gian tế bào, khả năng sẵn có của yếu tố đồng thời mà có thể gắn kết hoặc được kết hợp nhất thời với protein được cuộn gấp, cuộn gấp một phần hoặc không được cuộn gấp, sự có mặt của một hoặc nhiều protein chaperon, hoặc tương tự.

Protein sốc nhiệt (Hsp) hoặc các protein áp lực là các ví dụ về protein chaperon, mà có thể tham gia các quá trình tế bào khác nhau bao gồm tổng hợp protein, dịch chuyển nội bào, ngăn sự mất cuộn gấp, ngăn kết tụ protein, lắp ráp và tháo bỏ các phức hợp protein, cuộn gấp protein, và phân rã protein. Các ví dụ về các protein chaperon này bao gồm, nhưng không giới hạn ở, Hsp60, Hsp65, Hsp 70, Hsp90, Hsp100, Hsp20-30, Hsp10, Hsp100-200, Hsp90, Lon, TF55, FKBP, xyclophilin, ClpP, GrpE, ubiquitin, calnexin, và protein disulfua isomerase (xem, ví dụ, Macario, A.J.L., *Cold Spring Harbor*

Laboratory Res. 25:59-70. 1995; Parsell, D.A. & Lindquist, S. *Ann. Rev. Genet.* 27:437-496 (1993); Patent US số 5,232,833). Như được mô tả ở đây, các protein chaperon, ví dụ, nhưng không giới hạn ở Hsp40 và Hsp70 có thể được sử dụng để đảm bảo sự cuộn gấp của protein virut.

Các ví dụ về Hsp70 bao gồm Hsp72 và Hsc73 từ các tế bào động vật có vú, DnaK từ vi khuẩn, cụ thể là mycobacteria như *Mycobacterium leprae*, *Mycobacterium tuberculosis*, và *Mycobacterium bovis* (như Bacille-Calmette Guerin: được đề cập ở đây là Hsp71). DnaK từ *Escherichia coli*, nấm men và các sinh vật nhân sơ khác, và BiP và Grp78 từ các sinh vật nhân thực, như *A. thaliana* (Lin et al. 2001 (*Cell Stress and Chaperones* 6:201-208)). Ví dụ cụ thể về Hsp70 là Hsp70 *A. thaliana* (được mã hóa bằng số Genbank: AY120747.1). Hsp70 có khả năng gắn kết đặc hiệu ATP cũng như các polypeptit và peptit tháo cuộn gấp, theo đó tham gia vào quá trình cuộn gấp và tháo cuộn gấp protein cũng như trong lắp ráp và tháo bỏ các phức hợp protein.

Các ví dụ về Hsp40 bao gồm DnaJ từ sinh vật nhân sơ như *E. coli* và mycobacteria và HsJ1, HDJ1 và Hsp40 từ sinh vật nhân sơ, như cây đậu đũa (Frugis et al., 1999. *Plant Molecular Biology* 40:397-408). Ví dụ cụ thể về Hsp40 là *M. sativa* MsJ1 (Genbank ref: AJ000995.1). Hsp40 đóng vai trò làm chaperon phân tử trong cuộn gấp protein, chịu nhiệt và sao chép ADN, trong số các hoạt tính tế bào khác.

Trong số các Hsp, Hsp70 và đồng chaperon của nó, Hsp40, liên quan đến sự ổn định hóa dịch mã và polypeptit được tổng hợp mới trước khi việc tổng hợp hoàn thành. Không mong muốn bị ràng buộc bởi lý thuyết, Hsp40 gắn kết với các miếng kỵ nước của các polypeptit không được cuộn gấp (mới sinh hoặc được chuyển mới), do đó đẩy nhanh sự tương tác của phức Hsp70-ATP với polypeptit. Sự thủy phân ATP dẫn đến sự hình thành phức ổn định giữa polypeptit, Hsp70 và ADP, và giải phóng Hsp40. Sự kết hợp của phức Hsp70-ADP với các miếng kỵ nước của polypeptit ngăn sự tương tác của chúng với các miếng kỵ nước khác, ngăn sự cuộn gấp không chính xác và tạo thành các kết tụ với với các protein khác (được tổng kết trong Hartl, FU. 1996. *Nature* 381:571-579).

Các protein chaperon tự nhiên có thể đẩy nhanh sự cuộn gập chính xác của protein tái tổ hợp, nhưng do mức biểu hiện tăng, sự phong phú của các chaperon tự nhiên có thể trở thành yếu tố giới hạn. Mức biểu hiện cao của protein virus ở các lá được thấm bằng *Agrobacteria* có thể dẫn đến sự tích tụ protein virus trong xytosol, và sự biểu hiện đồng thời của một hoặc nhiều protein chaperon như Hsp70, Hsp40 hoặc cả Hsp70 và Hsp40 có thể giảm mức protein bị xóa bỏ cuộn gập hoặc kết tụ, và làm tăng số protein có đặc trưng cấu trúc bậc ba và bậc bốn cho phép tạo thành các hạt giống như virus.

Do đó, sáng chế cũng đề xuất phương pháp sản xuất protein virus VLP ở thực vật, trong đó axit nucleic thứ nhất mã hóa protein virus được biểu hiện đồng thời với axit nucleic thứ hai mã hóa protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton, và axit nucleic thứ ba mã hóa chaperon. Các axit nucleic thứ nhất, thứ hai và thứ ba có thể được đưa vào thực vật trong cùng bước, hoặc có thể được đưa vào thực vật theo thứ tự.

N-Glycan

VLP được tạo ra trong thực vật có thể kích thích protein virus chứa các N-glycan đặc hiệu với thực vật. Do đó, sáng chế cũng đề xuất VLP bao gồm protein virus có N-glycan đặc hiệu với thực vật.

Hơn nữa, sự biến đổi N-glycan ở thực vật đã được biết đến (xem, ví dụ, WO 2008/151440; WO 2010/006452; hoặc U.S. 60/944,344) và protein virus có các N-glycan được biến đổi có thể được tạo ra. Protein virus bao gồm mẫu glycosyl hóa được biến đổi, ví dụ, làm giảm fucosyl hóa, xylosyl hóa, hoặc cả hai, fucosyl hóa và xylosyl hóa, N-glycan có thể thu được, hoặc protein virus có mẫu glycosyl hóa được biến đổi có thể thu được, trong đó protein thiếu fucosyl hóa, xylosyl hóa, hoặc cả hai, và bao gồm tăng galatosyl hóa. Hơn nữa, việc điều biến các biến đổi sau dịch mã, ví dụ, việc bổ sung galactose có thể làm giảm fucosyl hóa và xylosyl hóa của protein virus được biểu hiện khi so sánh với protein virus biểu hiện thực vật kiểu dại.

Ví dụ, không được xem xét làm giới hạn sáng chế, sự tổng hợp protein virus có mẫu glycosyl hóa được biến đổi có thể đạt được bằng cách biểu hiện đồng thời protein quan tâm cùng với trình tự nucleotit mã hóa beta-1.4galactosyltransferaza (GalT), ví dụ, nhưng

không giới hạn ở GalT của động vật có vú, hoặc GalT của người, tuy nhiên GalT từ các nguồn khác cũng có thể được sử dụng. Miền xúc tác của GalT cũng có thể được dung hợp với miền CTS (nghĩa là đuôi tế bào chất, miền xuyên màng, vùng thân) của N-acetylglucosaminyl transferaza (GNT1), để tạo ra enzym lai GNT1-GalT, và enzym lai có thể được biểu hiện đồng thời với protein virus. Protein virus cũng có thể được biểu hiện đồng thời cùng với trình tự nucleotit mã hóa N-acetylglucosaminyltransferaza III (GnT-III), ví dụ, nhưng không giới hạn ở GnT-III của động vật có vú hoặc GnT-III của người, GnT-III từ các nguồn khác cũng có thể được sử dụng. Ngoài ra, enzym lai GNT1-GnT-III, bao gồm CTS của GNT1 được dung hợp với GnT-III cũng có thể được sử dụng.

Do đó, sáng chế cũng bao gồm VLP bao gồm một hoặc nhiều protein virus có các N-glycan được biến đổi.

Trình tự

Ví dụ không giới hạn về các trình tự có thể được sử dụng theo sáng chế bao gồm:

Protein H2 được mã hóa bằng phân tử axit nucleic có thể từ chủng A/Singapore/1/57 (H2N2);

Protein H3 được mã hóa bằng phân tử axit nucleic có thể từ A/Brisbane 10/2007 (H3N2), chủng A/Wisconsin/67/2005 (H3N2), A/Victoria/361/2011 (H3N2) hoặc A/Perth/16/2009 (H3N2);

Protein H6 được mã hóa bằng phân tử axit nucleic có thể từ chủng A/Teal/HongKong/W312/97 (H6N1);

Protein H7 được mã hóa bằng phân tử axit nucleic cũng có thể từ chủng A/Equine/Prague/56 (H7N7);

Protein H9 được mã hóa bằng phân tử axit nucleic có thể từ chủng A/HongKong/1073/99 (H9N2);

Protein HA từ typ phụ B được mã hóa bằng axit nucleic có thể từ chủng B/Florida/4/2006, B/Malaysia/2506/2004, B/Wisconsin/1/2010, hoặc B/Brisbane/60/2008.

Ví dụ không giới hạn về các trình tự có thể được sử dụng theo sáng chế cũng có thể bao gồm các trình tự được mô tả trong các công bố đơn sáng chế quốc tế số WO 2009/009876; WO 2009/076778; WO 2010/003225; WO 2010/148511; WO 2010/003235; WO 2010/006452). Các ví dụ về các trình tự của các trình tự của các phân tử axit amin mã hóa các protein HA này từ H2, H3, H4, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16 và HA typ B, không được biết trong lĩnh vực này. Ví dụ H3 hoặc các typ phụ B bao gồm các SEQ ID No: 25 hoặc 30. Trình tự mã hóa protein virus cấu trúc có thể là, ví dụ, HA từ cúm B/Brisbane/60/2008, B/Malaysia/2506/2004 hoặc B/Wisconsin/1/2010, hoặc H3 từ cúm A/Perth/16/2009 hoặc A/Victoria/361/2011. Các ví dụ khác bao gồm các trình tự của phân tử axit nucleic mã hóa các protein HA, trong đó vòng phân hủy protein của protein HA đã bị xóa bỏ như, ví dụ, nhưng không giới hạn ở trình tự được xác định bằng SEQ ID NO: 41.

Sáng chế cũng bao gồm, nhưng không giới hạn ở, các trình tự nucleotit mã hóa HA từ, ví dụ, H2, H3, H4, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16 hoặc HA typ B. Ví dụ SEQ ID NO: 28, 43, 23, mã hóa HA từ B, B bị xóa bỏ vòng phân hủy protein hoặc H3 tương ứng, trình tự nucleotit lai trong các điều kiện lai nghiêm ngặt với SEQ ID NO: 28, 43, 23, hoặc trình tự nucleotit lai trong các điều kiện lai nghiêm ngặt với bộ thể của SEQ ID NO: 28, 43, 23, trong đó trình tự nucleotit mã hóa protein hemagglutinin mà khi được biểu hiện tạo thành VLP, và VLP kích thích sự tạo ra kháng thể. Ví dụ, sự biểu hiện trình tự nucleotit trong tế bào thực vật tạo thành VLP, và VLP có thể được sử dụng để sự tạo ra kháng thể có khả năng gắn kết HA, bao gồm HA trưởng thành từ B hoặc H3. VLP, khi dùng cho đối tượng, kích thích đáp ứng miễn dịch. Trình tự nucleotit cũng có thể được biểu hiện đồng thời với trình tự nucleotit thứ hai mã hóa protein kênh, ví dụ, nhưng không giới hạn ở, trình tự nucleotit SEQ ID NO: 9, 12, trình tự nucleotit lai trong các điều kiện lai nghiêm ngặt với SEQ ID NO: 9, 12, hoặc trình tự nucleotit lai trong các điều kiện lai nghiêm ngặt với bộ thể của SEQ ID NO: 9, 12, trong đó trình tự nucleotit thứ hai mã hóa protein kênh proton tạo thành VLP. Tốt hơn là, VLP kích thích sự tạo ra kháng thể và VLP, khi dùng cho đối tượng, kích thích đáp ứng miễn dịch.

Ví dụ, sự biểu hiện trình tự nucleotit trong tế bào thực vật tạo thành VLP, và VLP có thể được sử dụng để sự tạo ra kháng thể có khả năng gắn kết protein virus như, ví dụ, HA, bao gồm, nhưng không giới hạn ở HA0, protein HA0 có vòng phân hủy protein bị xóa bỏ hoặc bị biến đổi, HA1 hoặc HA2 của một hoặc nhiều typ hoặc typ phụ cúm, như, ví dụ nhưng không giới hạn ở các typ phụ H2, H3, H4, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, HA typ phụ B. VLP, khi dùng cho đối tượng, kích thích đáp ứng miễn dịch.

Việc lai trong các điều kiện lai nghiêm ngặt đã được biết trong lĩnh vực này (xem, ví dụ ấn phẩm: *Current Protocols in Molecular Biology*, Ausubel et al., eds. 1995 và phần phụ lục; Maniatis *et al.*, in *Molecular Cloning (A Laboratory Manual)*, Cold Spring Harbor Laboratory, 1982; Sambrook và Russell, trong *Molecular Cloning: A Laboratory Manual*, tái bản lần thứ 3 năm 2001). Ví dụ về một trong các điều kiện lai nghiêm ngặt này có thể khoảng từ 16 đến 20 giờ lai trong SSC 4X ở 65°C, sau đó rửa trong SSC 0,1X ở 65°C trong một giờ, hoặc rửa 2 lần trong SSC 0,1X ở 65°C, mỗi lần trong 20 hoặc 30 phút. Theo cách khác, điều kiện lai nghiêm ngặt minh họa có thể là qua đêm (từ 16 đến 20 giờ) trong 50% formamit, SSC 4 X ở 42°C, sau đó là rửa trong SSC 0,1 X ở 65°C trong một giờ, hoặc 2 lần rửa trong SSC 0,1 X ở 65°C, mỗi lần trong 20 hoặc 30 phút, hoặc qua đêm (từ 16 đến 20 giờ), hoặc lai trong dung dịch đệm phosphat chứa nước Church (7% SDS; dung dịch đệm NaPO₄ 0,5M pH 7,2; EDTA 10mM) ở 65°C, với 2 lần rửa ở 50°C trong SSC 0,1X, 0,1% SDS trong 20 hoặc 30 phút mỗi lần, hoặc 2 lần rửa ở 65°C trong SSC 2X, 0,1% SDS trong 20 hoặc 30 phút mỗi lần.

Ngoài ra, sáng chế bao gồm các trình tự nucleotit khác biệt ở chỗ có độ độ tương đồng trình tự, hoặc độ giống trình tự khoảng 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% hoặc trị số bất kỳ nằm trong đó với trình tự nucleotit mã hóa HA từ B (SEQ ID NO: 28), B có vòng phân hủy protein bị xóa bỏ hoặc biến đổi (SEQ ID NO: 43), H3 (SEQ ID NO:23), hoặc HA được mã hóa bằng một hoặc nhiều hoặc SEQ ID NO:23, 28, 43, 46, 51, 57, hoặc 61 bất kỳ, trong đó trình tự nucleotit mã hóa protein hemagglutinin mà khi được biểu hiện tạo thành VLP, và VLP kích thích tạo ra kháng thể. Ví dụ, sự biểu hiện trình tự nucleotit trong tế bào thực vật tạo thành VLP, và VLP có thể

được sử dụng để tạo ra kháng thể có khả năng gắn kết HA, bao gồm HA trưởng thành và/hoặc chưa được xử lý từ B hoặc H3, hoặc HA chưa được xử lý và/hoặc HA trưởng thành, trong đó vòng phân hủy protein bị xóa bỏ. VLP, khi được dùng cho đối tượng, kích thích đáp ứng miễn dịch.

Sáng chế cũng bao gồm các trình tự nucleotit được đặc trưng là có độ tương đồng trình tự, độ giống trình tự khoảng 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% hoặc lượng bất kỳ nằm trong đó với trình tự nucleotit mã hóa M2 (SEQ ID NO: 9, 12), trong đó trình tự nucleotit mã hóa protein kênh, ví dụ, nhưng không giới hạn ở protein kênh proton, mà khi được biểu hiện đồng thời với protein virus cấu trúc tạo thành VLP. Tốt hơn là, VLP kích thích sự tạo ra kháng thể và VLP, khi được dùng với đối tượng, kích thích đáp ứng miễn dịch.

Độ tương đồng trình tự hoặc độ giống trình tự có thể được xác định bằng cách sử dụng chương trình so sánh trình tự nucleotit, như được đề xuất trong DNASIS (ví dụ, sử dụng, nhưng không giới hạn ở, các thông số sau: điểm trừ GAP 5, #của các đường chéo trên 5, điểm trừ GAP cố định 10, k-tuple 2, khoảng trống nổi 10, và kích cỡ cửa sổ 5). Tuy nhiên, các phương pháp sắp xếp trình tự khác để so sánh đã biết trong lĩnh vực này, ví dụ, các thuật toán của Smith & Waterman (1981, Adv. Appl. Math. 2:482), Needleman & Wunsch (J. Mol. Biol. 48:443, 1970), Pearson & Lipman (1988, Proc. Nat'l. Acad. Sci. USA 85:2444), và bằng các hệ tính toán của các giải thuật này (ví dụ, GAP, BESTFIT, FASTA, và BLAST), hoặc bằng sự sắp xếp thủ công và quan sát bằng mắt. Ví dụ về sắp xếp trình tự của các HA từ các chủng cúm khác nhau có thể được tìm thấy trong Fig.24.

“Đáp ứng miễn dịch” thường đề cập đến sự đáp ứng của hệ miễn dịch đáp ứng. Hệ miễn dịch đáp ứng bao gồm đáp ứng thể dịch, và đáp ứng qua trung gian tế bào. Đáp ứng thể dịch là khía cạnh miễn dịch qua trung gian các kháng thể được tiết, được tạo ra trong các tế bào của dòng tế bào lympho B (tế bào B). Các kháng thể được tiết gắn kết với các kháng nguyên trên các bề mặt của vi sinh vật xâm nhiễm (như virus hoặc vi khuẩn), mà đánh dấu chúng để tiêu diệt. Miễn dịch thể dịch thường được sử dụng để đề cập đến sự tạo ra kháng thể và các quá trình kèm theo nó, cũng như các chức năng cơ quan thực hiện của các kháng thể, bao gồm hoạt hóa tế bào Th2 và sự tạo ra xytokin, tạo ra tế bào ghi nhớ,

thúc đẩy opsonin của thực bào, loại trừ mầm bệnh và tương tự. Các thuật ngữ “điều biến” hoặc “sự điều biến” hoặc thuật ngữ tương tự đề cập đến việc tăng hoặc giảm đáp ứng hoặc thông số cụ thể, như được xác định bằng một vài thử nghiệm bất kỳ thường được biết hoặc sử dụng, một trong số chúng được minh họa ở đây.

Đáp ứng qua trung gian tế bào là đáp ứng miễn dịch mà không liên quan đến kháng thể nhưng khá liên quan đến sự hoạt hóa các đại thực bào, tế bào giết tự nhiên (NK), tế bào lympho T độc tế bào đặc hiệu với kháng nguyên, và giải phóng các xytokin khác nhau đáp ứng lại với kháng nguyên. Miễn dịch qua trung gian tế bào thường được sử dụng để đề cập đến một số hoạt hóa tế bào Th, hoạt hóa tế bào Tc và các đáp ứng qua trung gian tế bào T. Miễn dịch qua trung gian tế bào quan trọng đặc biệt trong việc phản ứng lại các lây nhiễm virus.

Ví dụ, sự kích thích các tế bào lympho T dương tính với CD8 đặc hiệu có thể đo được bằng cách sử dụng thử nghiệm ELISPOT; sự kích thích các tế bào lympho T dương tính với CD4 có thể đo được bằng cách sử dụng thử nghiệm tăng sinh. Các chuẩn kháng thể kháng cúm có thể được định lượng bằng cách sử dụng thử nghiệm ELISA; các isotyp của các kháng thể đặc hiệu kháng nguyên hoặc phản ứng chéo cũng có thể được đo bằng cách sử dụng các kháng thể kháng isotyp (ví dụ kháng IgG, IgA, IgE hoặc IgM). Các phương pháp và kỹ thuật thực hiện các thử nghiệm này đã được biết trong lĩnh vực này.

Các chuẩn HAI phản ứng chéo cũng có thể được sử dụng để chứng minh hiệu quả của đáp ứng miễn dịch với các chủng khác của virus liên quan đến typ phụ của vacxin. Ví dụ, huyết thanh từ đối tượng được tạo miễn dịch với chế phẩm vacxin của chủng thứ nhất (ví dụ, VLP của A/Indonesia 5/05) có thể được sử dụng trong thử nghiệm HAI có chủng thứ hai của virus đầy đủ hoặc các hạt virus (ví dụ, A/Vietnam/1194/2004), và chuẩn HAI được xác định.

Sự xuất hiện hoặc nồng độ xytokin cũng có thể được định lượng. Ví dụ, đáp ứng tế bào giúp đỡ T (Th1/Th2) sẽ được xác định bằng cách đo các tế bào tiết IFN- γ và IL-4 bằng cách sử dụng ELISA (ví dụ, các kit OptEIA của BD Biosciences). Các tế bào đơn nhân máu ngoại biên (PBMC) hoặc tế bào lách thu được từ đối tượng có thể được nuôi cấy, và

phân tích dịch nổi. Các tế bào lympho T cũng có thể được định lượng bằng cách sắp xếp tế bào hoạt hóa do huỳnh quang (FACS), sử dụng các nhãn huỳnh quang đặc hiệu với chi thị và các phương pháp đã biết trong lĩnh vực này.

Thử nghiệm vi trung hòa cũng có thể được tiến hành để mô tả đáp ứng miễn dịch ở đối tượng, xem, ví dụ, các phương pháp của Rowe et al., 1973. Các chuẩn trung hòa virus có thể thu được theo một số cách, bao gồm: 1) đếm các tấm phân hủy (thử nghiệm tấm) sau khi cố định/nhuộm màu tím tinh thể tế bào; 2) quan sát dưới kính hiển vi sự phân hủy tế bào trong môi trường nuôi cấy; 3) ELISA và phát hiện quang phổ của protein virus NP (tương quan với lây nhiễm virus của các tế bào vật chủ).

Các cấu trúc

Sáng chế còn đề cập đến cấu trúc gen bao gồm axit nucleic mã hóa protein kênh, ví dụ nhưng không giới hạn ở protein kênh proton hoặc protein virus cấu trúc, như được mô tả trên đây, được liên kết theo kiểu điều khiển được với nhân tố điều hòa hoạt động ở thực vật. Các ví dụ về các nhân tố điều hòa hoạt động ở tế bào thực vật và có thể được sử dụng theo sáng chế bao gồm nhưng không giới hạn ở vùng điều hòa plastocyanin (US 7,125,978), hoặc vùng điều hòa của Ribulosa 1,5-bisphosphat carboxylaza/oxygenaza (RuBisCO; US 4,962,028), protein gắn kết chlorophyll a/b (CAB; Leutwiler et al; 1986), ST-LS1 (được kết hợp với phức phát ra oxy của quang hệ II và được mô tả bởi Stockhaus et al. 1987, 1989).

Các nhân tố điều hòa

Sử dụng các thuật ngữ "vùng điều hòa", "nhân tố điều hòa" hoặc "trình tự khởi đầu (promoter)" trong sáng chế nghĩa là phản ánh một phần axit nucleic, thường là, nhưng không phải trong mọi trường hợp, ngược dòng (nằm phía trước) vùng mã hóa protein của gen, mà có thể bao gồm ADN hoặc ARN, hoặc cả ADN và ARN. Khi vùng điều hòa hoạt động, và theo sự kết hợp để điều khiển, hoặc được liên kết theo kiểu điều khiển được, với gen mong muốn, việc này tạo ra sự biểu hiện gen mong muốn. Nhân tố điều hòa có thể có khả năng qua trung gian tính đặc hiệu cơ quan, hoặc điều khiển sự hoạt hóa gen tiến triển hoặc tạm thời. "Vùng điều hòa" có thể bao gồm các nhân tố promoter, các nhân tố

promoter cốt lõi có hoạt tính promoter cơ bản, các nhân tố có thể kích thích đáp ứng với sự kích thích bên ngoài, các nhân tố qua trung gian hoạt tính promoter như các nhân tố điều hòa âm hoặc các vùng tăng cường phiên mã. Như được sử dụng ở đây, "vùng điều hòa" cũng có thể bao gồm các nhân tố hoạt động sau phiên mã, ví dụ, các nhân tố điều hòa mà điều hòa sự biểu hiện gen như vùng tăng cường phiên mã và dịch mã, vùng ức chế phiên mã và dịch mã, các trình tự hoạt hóa ngược dòng, và các vùng quyết định sự mất ổn định của ARN thông tin. Một vài trong số nhân tố sau này có thể nằm gần với vùng mã hóa.

Trong ngữ cảnh của sáng chế, thuật ngữ "nhân tố điều hòa" hoặc "vùng điều hòa" thường đề cập đến trình tự ADN, thường, nhưng không phải trong mọi trường hợp, ngược dòng (5') với trình tự mã hóa gen cấu trúc, mà điều khiển sự biểu hiện của vùng mã hóa bằng cách tạo sự nhận diện cho ARN polymeraza và/hoặc các yếu tố khác được yêu cầu để khởi đầu quá trình phiên mã tại vị trí cụ thể. Tuy nhiên, sẽ được hiểu rằng các trình tự nucleotit khác, nằm trong các intron, hoặc 3' của trình tự này cũng có thể góp phần điều hòa sự biểu hiện vùng mã hóa quan tâm. Ví dụ về nhân tố điều hòa mà tạo ra sự nhận diện cho ARN polymeraza hoặc các yếu tố phiên mã khác để đảm bảo sự khởi đầu tại vị trí cụ thể là nhân tố promoter. Hầu hết, nhưng không phải tất cả, các nhân tố promoter của sinh vật nhân thực chứa TATA box (hộp TATA), trình tự axit nucleic bảo toàn bao gồm cặp bazơ nucleotit adenosin và thymidin thường được đặt khoảng 25 cặp bazơ ngược dòng so với vị trí bắt đầu phiên mã. Nhân tố promoter bao gồm nhân tố promoter cơ bản, chịu trách nhiệm khởi đầu phiên mã, cũng như các nhân tố điều hòa khác (như được liệt kê trên đây) mà biến đổi sự biểu hiện gen.

Có một số loại vùng điều hòa, bao gồm các vùng được điều hòa theo cách tiến triển, có thể cảm ứng hoặc cấu trúc. Vùng điều hòa mà được điều hòa tiến triển, hoặc điều khiển sự biểu hiện khác nhau của gen dưới sự kiểm soát của nó, được hoạt hóa bên trong các cơ quan hoặc mô của cơ quan nhất định tại các thời điểm cụ thể trong quá trình phát triển của cơ quan hoặc mô đó. Tuy nhiên, một số vùng điều hòa mà được điều hòa tiến triển có thể hoạt động ưu tiên trong một số cơ quan hoặc mô tại các giai đoạn phát triển cụ thể, chúng cũng có thể hoạt động theo cách điều hòa tiến triển, hoặc cũng ở mức cơ sở ở

các cơ quan hoặc mô khác trong thực vật. Các ví dụ về các vùng điều hòa đặc hiệu với mô, ví dụ, xem vùng điều hòa cụ thể, bao gồm promoter napin, và promoter cruciferin (Rask et al., 1998, *J. Plant Physiol.* 152: 595-599; Bilodeau et al., 1994, *Plant Cell* 14: 125-130). Ví dụ về promoter đặc hiệu với lá bao gồm promoter plastoxyanin (xem US 7,125,978).

Vùng điều hòa cảm ứng là vùng có khả năng hoạt hóa phiên mã trực tiếp hoặc gián tiếp một hoặc nhiều trình tự hoặc gen ADN đáp ứng lại nhân tố cảm ứng. Với sự vắng mặt của nhân tố cảm ứng, các trình tự hoặc gen sẽ không được phiên mã. Thông thường, yếu tố protein mà gắn kết đặc hiệu với vùng điều hòa cảm ứng để hoạt hóa phiên mã có thể có mặt ở dạng bất hoạt, mà sau đó chuyển hóa trực tiếp hoặc gián tiếp thành dạng hoạt động bằng nhân tố cảm ứng. Tuy nhiên, yếu tố protein này cũng có thể không có. Nhân tố cảm ứng có thể là chất hóa học như protein, chất trao đổi, chất điều hòa sinh trưởng, chất diệt cỏ hoặc hợp chất phenol hoặc áp lực sinh lý được áp dụng trực tiếp bởi các yếu tố nhiệt, lạnh, muối, hoặc độc tố hoặc gián tiếp thông qua hoạt động của mầm bệnh hoặc tác nhân gây bệnh như là virus. Tế bào thực vật chứa vùng điều hòa cảm ứng có thể được phơi nhiễm với nhân tố cảm ứng bằng áp dụng nhân tố cảm ứng bên ngoài với tế bào hoặc thực vật như bằng cách phun, tưới, gia nhiệt hoặc các phương pháp tương tự khác. Các nhân tố điều hòa cảm ứng có thể thu được từ các gen thực vật hoặc không là thực vật (ví dụ, Gatz, C. và Lenk, L.R.P., 1998, *Trends Plant Sci.* 3, 352-358). Ví dụ, về các promoter cảm ứng tiềm năng bao gồm, nhưng không giới hạn ở, promoter cảm ứng với tetracyclin (Gatz, C., 1997, *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 48, 89-108), promoter cảm ứng với steroid (Aoyama, T. và Chua, N.H., 1997, *Plant* 1, 2, 397-404) và promoter cảm ứng với ethanol (Salter, M.G., et al., 1998, *Plant Journal* 16, 127-132; Caddick, M.X., et al., 1998, *Nature Biotech.* 16, 177-180), các gen IB6 và CKI cảm ứng với xytokinin (Brandstatter, I. và Kieber, J.J., 1998, *Plant Cell* 10, 1009-1019; Kakimoto, T., 1996, *Science* 274, 982-985;) và các nhân tố cảm ứng với auxin, DR5 (Ulmasov, T., et al., 1997, *Plant Cell* 9, 1963-1971;).

Vùng điều hòa cấu trúc điều khiển sự biểu hiện gen thông qua các phần khác nhau của thực vật và liên tục xuyên suốt quá trình phát triển của thực vật. Các ví dụ về các nhân tố điều hòa cấu trúc đã biết bao gồm các promoter kết hợp với các thể phiên mã CaMV

35S (Odell et al., 1985, *Nature*, 313: 810-812), actin 1 của lúa (Zhang et al, 1991, *Plant Cell*, 3: 1155-1165), actin 2 (An et al., 1996, *Plant J.*, 10: 107-121), hoặc tms 2 (U.S. 5,428,147), và triosephosphat isomerase 1 (Xu et al., 1994, *Plant Physiol.* 106: 459-467), gen ubiquitin 1 của ngô (Cornejo et al, 1993, *Plant Mol. Biol.* 29: 637-646), các gen ubiquitin 1 và 6 của *Arabidopsis* (Holtorf et al, 1995, *Plant Mol. Biol.* 29: 637-646), và gen 4A yếu tố khởi đầu dịch mã của cây thuốc lá (Mandel et al, 1995, *Plant Mol. Biol.* 29: 995-1004).

Như được sử dụng ở đây, thuật ngữ "cấu trúc" không nhất thiết là đề cập đến gen dưới sự điều khiển của vùng điều hòa cấu trúc được biểu hiện ở cùng mức độ ở tất cả các loại tế bào, nhưng gen đó được biểu hiện ở nhiều loại tế bào mặc dù sự khác nhau về độ phong phú thường được quan sát thấy. Các nhân tố điều hòa cấu trúc có thể được ghép với các trình tự khác để tăng cường thêm sự phiên mã và/hoặc dịch mã trình tự nucleotit mà chúng được liên kết theo kiểu điều khiển được vào đó. Ví dụ, hệ CPMV-HT thu được từ các vùng chưa được dịch mã của virus khảm cây đậu đũa (CPMV) và chúng minh được sự tăng cường dịch mã của trình tự mã hóa được kết hợp. "Tự nhiên" nghĩa là axit nucleic hoặc trình tự axit amin xuất hiện trong tự nhiên, hoặc "kiểu đại". "Được liên kết theo kiểu điều khiển được" nghĩa là các trình tự cụ thể, ví dụ, nhân tố điều hòa và vùng mã hóa mong muốn, tương tác trực tiếp hoặc gián tiếp để thực hiện chức năng mong muốn, như điều tiết hoặc điều hòa sự biểu hiện gen. Sự tương tác của các trình tự được liên kết theo kiểu điều khiển được có thể, ví dụ, qua trung gian các protein mà tương tác với các trình tự được liên kết theo kiểu điều khiển được.

Một hoặc nhiều protein virus như protein virus cấu trúc hoặc protein kênh, ví dụ nhưng không giới hạn ở protein kênh proton có thể được biểu hiện trong hệ biểu hiện bao gồm hệ biểu hiện ADN hoặc ARN dựa trên virus, ví dụ, nhưng không giới hạn ở, cat-xet biểu hiện dựa trên comovirus và nhân tố khuếch đại dựa trên geminivirus.

Như được mô tả ở đây, hệ biểu hiện có thể bao gồm cat-xet biểu hiện dựa trên virus hai phần hoặc virus có hệ gen hai phần. Ví dụ, các virus hai phần có thể thuộc họ Comoviridae. Chi thuộc họ Comoviridae bao gồm Comovirus, Nepovirus, Fabavirus, Chervirus và Sadwavirus. Comovirus bao gồm virus khảm từ cây đậu đũa (CPMV), virus

khảm nghiêm trọng của cây đậu đũa (CPSMV), virus khảm của cây bí (SqMV), virus vằn cỏ ba lá đỏ (RCMV), virus vằn vỏ đậu (BPMV), virus đốm củ cải (TuRSV), virus khảm thật đậu rộng (BBtMV), virus nhuộm đậu rộng (BBSV), virus khảm củ cải (RaMV). Các ví dụ về các trình tự comovirus ARN-2 bao gồm các nhân tố tăng cường mã hữu dụng với các khía cạnh khác nhau theo sáng chế bao gồm, nhưng không giới hạn ở: CPMV RNA-2 (số truy cập GenBank NC_003550), RCMV RNA-2 (số truy cập GenBank NC_003738), BPMV RNA-2 (số truy cập GenBank NC_003495), CPSMV RNA-2 (số truy cập GenBank NC_003544), SqMV RNA-2 (số truy cập GenBank NC_003800), TuRSV RNA-2 (số truy cập GenBank NC_013219.1), BBtMV RNA-2 (số truy cập GenBank GU810904), BBSV RNA2 (số truy cập GenBank FJ028650), RaMV (số truy cập GenBank NC_003800)

Các đoạn của hệ gen ARN comovirus hai phần được gọi là ARN-1 và ARN-2. ARN-1 mã hóa các protein liên quan đến sự sao chép trong khi ARN-2 mã hóa các protein cần thiết với sự di chuyển từ tế bào đến tế bào và hai protein vỏ capsit. Cat-xet dựa trên comovirus thích hợp bất kỳ có thể được sử dụng bao gồm CPMV, CPSMV, SqMV, RCMV, hoặc BPMV, ví dụ, cat-xet biểu hiện có thể dựa trên CPMV.

“Cat-xet biểu hiện” đề cập đến trình tự nucleotit bao gồm axit nucleic mong muốn dưới sự điều khiển của, và được liên kết theo kiểu điều khiển được với, promoter thích hợp hoặc các nhân tố điều hòa khác để phiên mã axit nucleic mong muốn trong tế bào vật chủ.

Đã biết rằng sự biến nạp vào *Nicotiana benthamiana* của các bản sao ADN có khả năng sao chép, chiều dài đầy đủ của cả hai ARN hệ gen của CPMV có thể tạo ra sự lây nhiễm năng suất (Liu et al., 2004, *Virology* 323, 37-48). Các ví dụ về các cat-xet biểu hiện CPMV được mô tả trong WO2007/135480; WO2009/087391; và Sainsbury F. et al., (2008, *Plant Physiology*; 148: 1212-1218; Sainsbury F. et al., (2008, *Plant Biotechnology Journal*; 6: 82-92; Sainsbury F. et al., 2009, *Plant Biotechnology Journal*; 7: 682-693). Ví dụ, không được xem xét làm giới hạn, các vùng không dịch mã (UTR) thu được từ ARN 2 hệ gen của virus khảm của cây đậu đũa (CPMV) mà trong đó hai codon khởi đầu dịch mã đầu tiên được tìm thấy ở trình tự dẫn đầu 5' đã bị xóa bỏ, có thể được sử dụng như được

mô tả trong WO 2009/087391. Khi kết hợp với promoter CaMV 35S và vùng kết thúc nopalín synthase (NOS), các UTR CPMV biến đổi tăng cường phiên mã của vùng mã hóa biên. Hệ biểu hiện dựa trên CPMV có tên là CPMV-HT (có khả năng siêu dịch mã). Do đó, cat-xet biểu hiện, cấu trúc biểu hiện và hệ biểu hiện của sáng chế cũng bao gồm hệ biểu hiện dựa trên CPMV như, ví dụ, hệ biểu hiện CPMV-HT.

Như được mô tả ở đây, trình tự tăng cường biểu hiện, mà trình tự này thu được từ (hoặc có sự đồng nhất với) phân mảnh hệ gen ARN-2 của virus ARN hai phần, như comovirus, trong đó vị trí khởi đầu hướng đích đã bị đột biến, có thể được sử dụng để biểu hiện trình tự axit nucleic mong muốn. Sáng chế còn đề xuất các quy trình để làm tăng sự biểu hiện hoặc hoạt tính tăng cường dịch mã, của trình tự thu được từ phân mảnh hệ gen ARN-2 của virus hai phần, mà các quy trình bao gồm gây đột biến vị trí khởi đầu đích tại đó.

Các trình tự “tăng cường” (hoặc nhân tố tăng cường), bao gồm các trình tự thu được từ (hoặc có sự đồng nhất với) phân mảnh hệ gen ARN-2 của virus ARN hai phần, như comovirus, trong đó vị trí khởi đầu hướng đích đã bị đột biến. Các trình tự này có thể tăng cường sự biểu hiện xuôi dòng của ORF không đồng nhất mà chúng gắn vào. Không có sự giới hạn, người ta tin rằng các trình tự này có mặt trong ARN được phiên mã, có thể tăng cường sự dịch mã của ORF không đồng nhất mà chúng gắn vào.

Các hệ biểu hiện cũng có thể bao gồm các nhân tố khuếch đại từ geminivirus ví dụ, nhân tố khuếch đại từ virus gây bệnh lùn, vàng ở đậu (BeYDV). BeYDV thuộc chi Mastrevirus thích hợp với thực vật hai lá mầm. BeYDV là phần đơn riêng biệt có hệ gen ADN vòng mạch đơn và có thể sao chép với số bản sao rất cao bằng cơ chế vòng tròn lăn. Các hệ vector replicon (đơn vị tái bản) ADN thu được từ BeYDV được sử dụng cho quá trình tạo ra protein với hiệu suất cao nhanh ở thực vật.

Như được sử dụng ở đây, cụm từ “nhân tố khuếch đại” đề cập đến đoạn axit nucleic bao gồm ít nhất một hoặc nhiều vùng liên gen dài (LIR) của hệ gen geminivirus. Như được sử dụng ở đây, “vùng liên gen dài” đề cập đến một vùng của vùng liên gen dài chứa vị trí gắn kết rep có thể điều hòa việc cắt và sao chép bằng protein Rep của geminivirus. Theo

một số khía cạnh, đoạn axit nucleic bao gồm một hoặc nhiều LIR, có thể còn bao gồm vùng liên gen ngắn (SIR) của hệ gen geminivirus. Như được sử dụng ở đây, “vùng liên gen ngắn” đề cập đến mạch bổ sung (IR ngắn (SIR) của Mastrevirus). Nhân tố khuếch đại thu được từ geminivirus thích hợp bất kỳ có thể được sử dụng ở đây. Xem, ví dụ, WO2000/20557; WO2010/025285; Zhang X. et al. (2005, *Biotechnology and Bioengineering*, Vol. 93, 271-279), Huang Z. et al. (2009, *Biotechnology and Bioengineering*, Vol. 103, 706-714), Huang Z. et al. (2009, *Biotechnology and Bioengineering*, Vol. 106, 9-17)). Nếu nhiều hơn một LIR được sử dụng trong cấu trúc này, ví dụ, hai LIR, thì promoter, vùng CMPV-HT và trình tự axit nucleic mong muốn và vùng kết thúc được đóng khung bởi một trong hai LIR.

Như được mô tả ở đây, việc cung cấp đồng thời vector thu được từ virus gây lùn, vàng ở đậu (BeYDV) và vector cung cấp Rep/RepA, bằng cách thâm tách *Agrobacterium* các lá *Nicotiana benthamiana* tạo ra sự khuếch đại replicon hiệu quả và sự tạo ra protein mạnh.

Cat-xet biểu hiện dựa trên comovirus và nhân tố khuếch đại thu được từ geminivirus có thể lần lượt chứa vector thứ nhất và thứ hai, hoặc các thành phần này có thể được chứa trong một vector. Nếu hai vector được sử dụng, vector thứ nhất và thứ hai có thể được đưa vào tế bào thực vật đồng thời hoặc riêng biệt.

Như được mô tả ở đây, replicaza của virus cũng có thể được bao gồm trong hệ biểu hiện như được mô tả trên đây để tăng sự biểu hiện của axit nucleic mong muốn. Ví dụ không giới hạn về replicaza là BeYDV replicaza (pREP110) mã hóa BeYDV Rep và RepA (C2/C1; Huang et al., 2009, *Biotechnol. Bioeng.* 103, 706-714).

Bất hoạt gen sau phiên mã (PTGS) có thể liên quan đến giới hạn sự biểu hiện các gen chuyển ở thực vật, và biểu hiện đồng thời nhân tố ức chế bất hoạt từ p19 của virus làm còi rậm ở cà chua (TBSV p19) hoặc virus khoai tây Y (HcPro) có thể được sử dụng để chống lại sự phân hủy đặc hiệu của các mARN gen chuyển (Brignetti et al., 1998).

Các gen ức chế bất hoạt đã biết trong lĩnh vực này và có thể được sử dụng như được mô tả ở đây (Chiba et al., 2006, *Virology* 346:7-14), ví dụ, nhưng không giới hạn ở,

TEV -p1/HC-Pro (virus tạo vết khắc ở thuốc lá -p1/HC-Pro), BYV -p21, protein capsit của virus gây nứt ở cà chua (TCV -CP), 2b của virus khảm dưa chuột; CMV-2b), p25 của virus khoai tây X (PVX-p25), p11 của virus khoai tây M (PVM-p11), p11 của virus khoai tây S (PVS-p11), p16 của virus tạo vết khắc ở cây việt quất, (BScV -p16), p23 của virus tristexa họ cam quýt (CTV-p23), p24 của virus liên quan đến cuốn lá nhỏ -2, (GLRaV-2 p24), p10 của virus nhỏ A, (GVA-p10), p14 của virus nhỏ B (GBV-p14), p10 của virus tiềm tàng Heracleum (HLV-p10), hoặc p16 của virus tiềm tàng phổ biến của tỏi (GCLV-p16). Do đó, nhân tố ức chế bất hoạt, ví dụ, nhưng không giới hạn ở, HcPro, TEV -p1/HC-Pro, BYV-p21, TBSV p19, TCV-CP, CMV-2b, PVX-p25, PVM-p11, PVS-p11, BScV-p16, CTV-p23, GLRaV-2 p24, GBV-p14, HLV-p10, GCLV-p16 hoặc GVA-p10, có thể được biểu hiện đồng thời cùng với trình tự axit nucleic mã hóa protein mong muốn để thêm vào đó, đảm bảo sự tạo ra protein ở mức cao trong thực vật.

“Được biểu hiện đồng thời” nghĩa là hai hoặc nhiều hơn hai trình tự nucleotit được biểu hiện tại gần như là cùng thời điểm trong thực vật, và trong cùng mô thực vật. Tuy nhiên, các trình tự nucleotit không cần được biểu hiện chính xác là tại cùng thời điểm. Hơn nữa, hai hoặc nhiều trình tự nucleotit được biểu hiện theo cách sao cho các sản phẩm được mã hóa có cơ hội tương tác. Ví dụ, protein biến đổi glycosyl hóa của protein mong muốn có thể được biểu hiện trước hoặc trong quá trình khi protein mong muốn được biểu hiện sao cho sự biến đổi glycosyl hóa của protein mong muốn xảy ra. Hai hoặc nhiều hơn hai trình tự nucleotit có thể được biểu hiện bằng cách sử dụng hệ biểu hiện nhất thời, trong đó hai hoặc nhiều trình tự được đưa vào trong thực vật tại gần như là cùng thời điểm trong các điều kiện mà cả hai trình tự được biểu hiện. Theo cách khác, thực vật nền tảng bao gồm một trong các trình tự nucleotit, ví dụ, trình tự mã hóa protein mà biến đổi biên dạng (profile) glycosyl hóa của protein mong muốn, có thể được biến nạp nhất thời hoặc theo cách ổn định, với trình tự khác mã hóa protein mong muốn này. Trong trường hợp này, trình tự mã hóa protein biến đổi biên dạng glycosyl hóa của protein mong muốn có thể được biểu hiện trong mô mong muốn, trong suốt giai đoạn phát triển mong muốn, hoặc sự biểu hiện của nó có thể được cảm ứng bằng cách sử dụng promoter cảm ứng, và trình tự khác mã hóa protein mong muốn có thể được biểu hiện trong các điều kiện tương tự và trong cùng mô, để đảm bảo các trình tự nucleotit này được biểu hiện đồng thời.

Một hoặc nhiều protein virus có thể được tạo ra ở dạng sản phẩm phiên mã từ trình tự nucleotit, và protein được phân cắt sau tổng hợp, và theo yêu cầu, được kết hợp để tạo thành protein multimer. Do đó, một hoặc nhiều protein virus cũng bao gồm protein hoặc polypeptit bao gồm các tiểu đơn vị mà được kết hợp thông qua các cầu disulphua (nghĩa là protein multimer). Ví dụ, protein bao gồm các trình tự axit amin từ hai hoặc nhiều nguồn có thể được xử lý thành các tiểu đơn vị, và các tiểu đơn vị này được kết hợp qua các cầu disulphua để tạo ra protein.

Một hoặc nhiều trình tự axit nucleic hoặc cấu trúc gen theo sáng chế có thể được biểu hiện trong vật chủ thực vật thích hợp bất kỳ mà được biến nạp trình tự nucleotit, hoặc các cấu trúc, hoặc vector theo sáng chế. Các ví dụ về các vật chủ thích hợp bao gồm, nhưng không giới hạn ở, các cây nông nghiệp bao gồm cỏ linh lăng, hạt cải dầu, *Brassica* spp., ngô, *Nicotiana* spp., khoai tây, nhân sâm, đậu, yến mạch, lúa, đậu tương, lúa mì, lúa mạch, hướng dương, bông và tương tự.

Một hoặc nhiều cấu trúc gen theo sáng chế có thể còn bao gồm vùng không dịch mã 3'. Vùng không dịch mã 3' đề cập đến phần gen bao gồm đoạn ADN mà chứa tín hiệu polyadenyl hóa và các tín hiệu điều hòa bất kỳ nào khác có khả năng ảnh hưởng đến xử lý ARN thông tin hoặc biểu hiện gen. Tín hiệu polyadenyl hóa thường được đặc trưng bằng cách ảnh hưởng đến sự bổ sung các dây axit polyadenyl vào đầu 3' của tiền phân tử mRNA. Tín hiệu polyadenyl hóa thường được nhận biết bằng sự có mặt của thể tương đồng với dạng chính tắc 5' AATAAA-3' mặc dù các biến thể không phải là không phổ biến. Các ví dụ không giới hạn của các vùng 3' thích hợp là các vùng không được dịch mã được phiên mã 3' chứa tín hiệu polyadenyl hóa của khối u *Agrobacterium* cảm ứng các gen (Ti) plasmid, như gen nopal synthase (NOS), gen thực vật như các gen protein bảo quản đậu tương, các tiểu đơn vị nhỏ của ribulose-1, gen 5-bisphosphat carboxylase (ssRUBISCO; US 4,962,028), promoter được sử dụng trong điều hòa biểu hiện plastocyanin, được mô tả trong US 7,125,978.

Một hoặc nhiều cấu trúc di truyền theo sáng chế cũng có thể bao gồm nhân tố tăng cường khác, cho dịch mã hoặc phiên mã, nêu được yêu cầu. Các nhân tố tăng cường có thể nằm ở 5' hoặc 3' so với trình tự được phiên mã. Các nhân tố tăng cường đã biết với người

có hiểu biết trung bình trong lĩnh vực này, và có thể bao gồm codon khởi đầu ATG, các trình tự liền kề hoặc tương tự. Codon khởi đầu, nếu có mặt, có thể cùng pha với khung đọc ("trong khung") của trình tự mã hóa để tạo ra sự dịch mã chính xác của trình tự được phiên mã.

“Biến nạp” nghĩa là chuyển thông tin di truyền (trình tự nucleotit) giữa các loài với nhau mà được biểu hiện kiểu gen, kiểu hình hoặc cả hai. Việc chuyển thông tin di truyền giữa các loài với nhau từ cấu trúc sang vật chủ có thể nhất thời và chuyển thông tin di truyền không thể di truyền và chuyển thông tin di truyền được coi là ổn định.

Các cấu trúc theo sáng chế có thể được đưa vào các tế bào thực vật bằng cách sử dụng Ti plasmit, Ri plasmit, vector virus thực vật, biến nạp ADN trực tiếp, vi tiêm, xung điện, v.v.. Với các tổng quan về các kỹ thuật này, xem ví dụ, Weissbach và Weissbach, *Methods for Plant Molecular Biology*, Academy Press, New York VIII, pp. 421-463 (1988); Geierman và Corey, *Plant Molecular Biology*, 2d Ed. (1988); và Miki và Iyer, Fundamentals of Gene Transfer in Plants. Trong *Plant Metabolism*, 2d Ed. DT. Dennis, DH Turpin, DD Lefebvre, DB Layzell (eds), Addison Wesley, Langmans Ltd. London, pp. 561-579 (1997). Các phương pháp khác bao gồm hấp thu ADN trực tiếp, sử dụng liposom, xung điện, ví dụ, sử dụng protoplast, vi tiêm, vi đạn hoặc sợi đơn thể (whisker), và thâm chân không. Xem, ví dụ, Bilang, et al. (Gene 100: 247-250 (1991), Scheid et al. (Mol. Gen. Genet. 228: 104-112, 1991), Guerche et al. (Plant Science 52: 111-116, 1987), Neuhauser et al. (Theor. Appl Genet. 75: 30-36, 1987), Klein et al., Nature 327: 70-73 (1987); Howell et al. (Science 208: 1265, 1980), Horsch et al. (Science 227: 1229-1231, 1985), DeBlock et al., Plant Physiology 91: 694-701, 1989), Methods for Plant Molecular Biology (Weissbach và Weissbach, eds., Academic Press Inc., 1988), Methods in Plant Molecular Biology (Schuler và Zielinski, eds., Academic Press Inc., 1989), Liu và Lomonossoff (J Virol Meth, 105:343-348, 2002), U.S. Pat. Nos. 4,945,050; 5,036,006; và 5,100,792, các chuỗi đơn yêu cầu cấp patent US số 08/438,666, nộp ngày 10/05/1995, và 07/951,715, nộp ngày 25/09/1992.

Như được mô tả dưới đây, các phương pháp biểu hiện nhất thời có thể được sử dụng để biểu hiện các cấu trúc theo sáng chế (xem, Liu và Lomonossoff, 2002, Journal of

Virological Methods, 105:343-348). Theo cách khác, phương pháp biểu hiện nhất thời dựa trên chân không, như được mô tả bởi Kapila et al., 1997) có thể được sử dụng. Các phương pháp này bao gồm, ví dụ, nhưng không giới hạn ở, phương pháp ủ bằng *Agrobacteria* hoặc thấm bằng *Agrobacteria*, thấm bằng ống tiêm, tuy nhiên, các phương pháp nhất thời khác cũng có thể được sử dụng như được lưu ý nêu trên. Bằng việc ủ bằng *Agrobacteria*, thấm bằng *Agrobacteria*, hoặc thấm bằng ống tiêm, hỗn hợp *Agrobacteria* bao gồm axit nucleic mong muốn đi vào các khoảng gian bào của mô, ví dụ, lá, một phần trên không của thực vật (bao gồm thân, lá và hoa), một phần khác của thực vật (thân, rễ, hoa), hoặc toàn bộ thực vật này. Sau khi đi qua biểu bì, *Agrobacteria* lây nhiễm và chuyển các bản sao của T-ADN vào các tế bào. T-ADN được phiên mã episom và ARN thông tin được dịch mã, dẫn đến sự tạo ra protein mong muốn trong các tế bào được lây nhiễm, tuy nhiên, việc chuyển T-ADN bên trong nhân là nhất thời.

Để hỗ trợ nhận biết các tế bào được biến nạp, các cấu trúc của sáng chế này có thể còn được thao tác để chứa các chỉ thị chọn lọc thực vật. Các chỉ thị chọn lọc hữu dụng bao gồm các enzym tạo ra đề kháng với các chất hóa học như kháng sinh, ví dụ, gentamycin, hygromycin, kanamycin, hoặc các chất diệt cỏ như phosphinothryxin, glyphosat, clorosulfuron, và chất tương tự. Tương tự, có thể sử dụng các enzym mang lại sự tạo ra hợp chất có thể nhận biết được bằng sự thay đổi màu như GUS (beta-glucuronidaza), hoặc chất phát quang, như luciferaza hoặc GFP.

Thuật ngữ "vật liệu thực vật" nghĩa là vật liệu bất kỳ thu được từ thực vật. Vật liệu thực vật có thể bao gồm toàn bộ thực vật, mô, tế bào hoặc phần bất kỳ của nó. Hơn nữa, vật liệu thực vật có thể bao gồm các thành phần thực vật nội bào, thành phần thực vật ngoại bào, sản phẩm chiết lỏng hoặc rắn của thực vật hoặc tổ hợp của chúng. Hơn nữa, vật liệu thực vật có thể bao gồm thực vật, tế bào thực vật, mô, sản phẩm chiết lỏng, hoặc tổ hợp của chúng, từ lá, thân, quả, rễ của thực vật hoặc tổ hợp của chúng. Vật liệu thực vật có thể bao gồm thực vật hoặc một phần của nó mà không trải qua bước xử lý bất kỳ. Tuy nhiên, cũng dự định rằng vật liệu thực vật có thể được đưa vào các bước xử lý tối thiểu như được xác định dưới đây, hoặc xử lý mạnh, bao gồm tinh chế protein một phần hoặc về

cơ bản là toàn bộ nhờ các kỹ thuật đã biết rộng rãi trong lĩnh vực này bao gồm, nhưng không giới hạn ở sắc ký, điện di và tương tự.

Thuật ngữ "xử lý tối thiểu" nghĩa là vật liệu thực vật, ví dụ, thực vật hoặc phần của nó bao gồm protein mong muốn mà được tinh chế một phần để thu được sản phẩm chiết, sản phẩm đồng nhất của thực vật, phân đoạn của sản phẩm đồng nhất của thực vật hoặc tương tự (nghĩa là được xử lý tối thiểu). Tinh chế một phần có thể bao gồm, nhưng không giới hạn ở phá vỡ các cấu trúc tế bào thực vật, theo đó tạo ra chế phẩm bao gồm các thành phần thực vật hòa tan, và các thành phần thực vật không hòa tan có thể được tách, ví dụ, nhưng không giới hạn ở, bằng cách ly tâm, lọc hoặc sự kết hợp của chúng. Về vấn đề này, các protein được tiết bên trong gian ngoại bào của lá hoặc các mô khác có thể dễ dàng thu được bằng cách sử dụng chiết chân không hoặc chiết ly tâm, hoặc các mô có thể được chiết dưới áp suất bằng cách cho đi qua trục lăn hoặc nghiền hoặc tương tự để ép hoặc giải phóng protein từ trong gian ngoại bào. Xử lý tối thiểu cũng có thể liên quan đến việc chế phẩm dịch chiết thô của các protein hòa tan, do các chế phẩm này bị nhiễm không đáng kể từ các sản phẩm thực vật thức cấp. Hơn nữa, xử lý tối thiểu có thể liên quan đến chiết trong nước các protein hòa tan từ lá, sau đó kết tủa bằng muối thích hợp. Các phương pháp khác có thể bao gồm ngâm trên quy mô lớn hoặc chiết dịch để cho phép sử dụng sản phẩm chiết trực tiếp.

Vật liệu thực vật, ở dạng vật liệu hoặc mô thực vật có thể là được đưa vào đối tượng qua đường miệng. Vật liệu thực vật có thể được dùng làm một phần của chất bổ sung cho chế độ ăn uống, cùng với các thực phẩm khác, hoặc được bao nang. Vật liệu thực vật hoặc mô thực vật cũng có thể được cô đặc để cải thiện hoặc tăng vị ngon, hoặc được cung cấp cùng với các vật liệu, thành phần hoặc tá dược khác, nếu yêu cầu.

Dự định rằng thực vật chứa protein mong muốn, hoặc biểu hiện VLP chứa protein mong muốn có thể được dùng với đối tượng hoặc sinh vật đích, theo nhiều cách tùy thuộc vào nhu cầu và hoàn cảnh. Ví dụ, protein mong muốn thu được từ thực vật có thể được chiết trước khi sử dụng ở dạng thô, tinh chế một phần hoặc tinh chế. Nếu protein được tinh chế, thì nó có thể được tạo ra ở thực vật ăn được hoặc không ăn được. Hơn nữa, nếu protein được dùng qua đường miệng, mô thực vật có thể được thu hoạch hoặc cho đối

trạng ăn trực tiếp, hoặc mô được thu hoạch có thể được làm khô trước khi cho ăn, hoặc động vật được phép ăn thực vật mà không cần thu hoạch trước. Cũng xem xét nằm trong phạm vi của sáng chế đối với các mô thực vật được thu hoạch được cung cấp làm chất bổ sung cho thực phẩm trong thức ăn động vật. Nếu mô thực vật này được trở thành thức ăn cho đối tượng hoặc động vật mà xử lý thêm ít hoặc không xử lý thêm thì tốt hơn nếu mô thực vật được dùng có thể ăn được.

VLP được sản xuất theo sáng chế có thể được tinh chế, được tinh chế một phần từ thực vật, phần của thực vật hoặc vật liệu thực vật, có thể được dùng làm vacxin qua đường miệng, bằng cách sử dụng các phương pháp như đã biết với người có hiểu biết trung bình trong lĩnh vực này. Việc tinh chế có thể bao gồm việc tạo ra phân đoạn apoplast như được mô tả trong WO 2011/035422. Để sắc ký loại trừ kích thước điều chế, chế phẩm bao gồm VLP có thể thu được và vật liệu không hòa tan được loại bỏ bằng ly tâm. Kết tủa với PEG cũng có thể được sử dụng. Protein được thu hồi có thể được định lượng bằng các phương pháp thông thường (ví dụ, thử nghiệm Bradford, BCA), và dịch chiết đi qua cột loại trừ kích thước, bằng cách sử dụng ví dụ SEPHACRYL™, SEPHADEX™, hoặc môi trường tương tự, và các phân đoạn được thu lại. Blue Dextran 2000 hoặc protein thích hợp, có thể được sử dụng làm chuẩn hiệu chỉnh. Dịch chiết cũng có thể đi qua cột trao đổi ion và các phân hoạt động gom được. Sau khi sắc ký, các phân đoạn có thể còn được phân tích bằng điện di protein, thẩm tách miễn dịch, hoặc cả hai, để xác nhận sự có mặt của VLP và phần bổ sung protein của phân đoạn này.

Cũng được coi là một phần của sáng chế này là thực vật chuyển gen, tế bào thực vật, hạt giống hoặc phân đoạn bất kỳ của chúng chứa các trình tự nucleotit theo sáng chế. Các phương pháp tái tạo toàn bộ thực vật từ tế bào thực vật cũng đã biết trong lĩnh vực này. Nhìn chung, các tế bào thực vật được biến nạp được nuôi cấy trong môi trường thích hợp, mà có thể chứa các chất chọn lọc như kháng sinh, trong đó các chỉ thị chọn lọc được sử dụng để tạo điều kiện cho việc nhận biết các tế bào thực vật được biến nạp. Khi các nốt sần tạo ra, việc tạo thành ngọn có thể được hỗ trợ bằng cách sử dụng các hormon thực vật thích hợp theo các phương pháp đã biết và ngọn được chuyển vào môi trường phát sinh rễ để tái tạo thực vật. Tiếp đó, thực vật được sử dụng để tạo thành các thể hệ bản

sao, hoặc từ hạt giống hoặc sử dụng các kỹ thuật nhân giống thực vật. Các thực vật chuyển gen cũng có thể được tạo ra mà không cần sử dụng các môi trường nuôi cấy mô.

Như được thể hiện trong Fig.18, HA từ B/Brisbane/60/2008 được biểu hiện kém trong các lá *Nicotiana benthamiana* nhiễm *Agrobacteria* (xem các giếng “1008” hoặc “1029”). Tuy nhiên, sự biểu hiện đồng thời HA-typ B với M2 từ A/New Caledonia/20/99, dẫn đến tăng đáng kể biểu hiện HA (xem các giếng “1008+1261”; “1009+1261” và 1029+1261”). Việc tăng biểu hiện HA đã được quan sát ở cả HA typ B tự nhiên hoặc HA typ B khảm. Sự biểu hiện HA đã được quan sát với sự có mặt hoặc vắng mặt của các nhân tố khuếch đại (BeYDV), và qua các lần pha loãng khác nhau của *Agrobacterium*. Việc tăng tương tự trong biểu hiện H3 đã được quan sát khi H3 từ A/Perth/16/2009 đã được biểu hiện đồng thời với M2 từ A/New Caledonia/20/99 (Fig.19; so sánh giếng “1019” riêng H3, với “1019+1261” H3 được biểu hiện đồng thời với M2).

Sáng chế bao gồm các trình tự nucleotit như nêu trong Bảng 3.

Bảng 3. Danh mục các số nhận biết trình tự.

SEQ ID NO:	Mô tả	Hình vẽ
1	trình tự tín hiệu kênh proton HXXXW	
2	đoạn môi IF-H5A-I-05.s1+3c	Fig. 1A
3	đoạn môi IF-H5dTm.r	Fig. 1B
4	Cấu trúc 1191	Fig. 1D
5	Cat-xet biểu hiện số 489	Fig. 1E
6	Trình tự axit amin của H5 từ cúm A/Indonesia/5/2005 (H5N1)	Fig. 1F
7	đoạn môi IF-S1-M1+M2ANC.c	Fig. 2A
8	đoạn môi IF-S1-4-M2ANC.r	Fig. 2B
9	trình tự nucleotit của gen M2 được tổng hợp (DQ508860)	Fig. 2C
10	Cat-xet biểu hiện số 1261	Fig. 2D
11	Trình tự axit amin của M2 từ cúm A/New	Fig. 2E

SEQ ID NO:	Mô tả	Hình vẽ
	Caledonia/20/1999 (H1N1)	
12	trình tự nucleotit của gen M2 được tổng hợp	Fig. 3A
13	Cat-xet biểu hiện số 859	Fig. 3B
14	Trình tự axit amin của M2 từ cúm A/Puerto Rico/8/1934 (H1N1)	Fig. 3C
15	đoạn môi IF-H1A-C-09.s2+4c	Fig. 4A
16	đoạn môi IF-H1A-C-09.s1-4r	Fig. 4B
17	trình tự nucleotit của gen H1 được tổng hợp	Fig. 4C
18	Cấu trúc 1192	Fig. 4E
19	Cat-xet biểu hiện số 484	Fig. 4F
20	Trình tự axit amin của PDISP-H1 từ cúm A/California/7/2009 (H1N1)	Fig. 4G
21	đoạn môi IF-S2+S4-H3 Per.c	Fig. 5A
22	đoạn môi IF-S1a4-H3 Per.r	Fig. 5B
23	trình tự nucleotit của gen H3 được tổng hợp	Fig. 5C
24	Cat-xet biểu hiện số 1019	Fig. 5D
25	Trình tự axit amin của PDISP/H3 từ cúm A/Perth/16/2009 (H3N2)	Fig. 5E
26	đoạn môi IF-S2+S4-B Bris.c	Fig. 6A
27	đoạn môi IF-S1a4-B Bris.r	Fig. 6B
28	trình tự nucleotit của gen HA B Brisbane được tổng hợp	Fig. 6C
29	Cat-xet biểu hiện số 1029	Fig. 6D
30	Trình tự axit amin của PDISP/HA từ cúm B/Brisbane/60/2008	Fig. 6E
31	Cấu trúc 1194	Fig. 6G
32	Cat-xet biểu hiện số 1008	Fig. 6H
33	đoạn môi dTmH5I-B Bris.r	Fig. 7A

SEQ ID NO:	Mô tả	Hình vẽ
34	đoạn mồi B Bris-dTmH5I.c	Fig. 7B
35	đoạn mồi IF-S1aS4-dTmH5I.r	Fig. 7C
36	Cat-xet biểu hiện số 1009	Fig. 7D
37	Trình tự axit amin của PDISP/HA B Brisbane/H5Indo TMCT	Fig. 7E
38	đoạn mồi 1039+1059.r	Fig. 8A
39	đoạn mồi 1039+1059.c	Fig. 8B
40	Cat-xet biểu hiện số 1059 từ LIR bên trái BeYDV sang LIR bên phải BeYDV. PDISP/HA từ cúm B/Brisbane/60/2008 có vòng phân hủy protein bị xóa bỏ	Fig. 8C
41	Trình tự axit amin của PDISP/HA từ cúm B/Brisbane/60/2008 có vòng phân hủy protein bị xóa bỏ	Fig. 8D
42	trình tự nucleotit mã hóa H5 từ cúm A/Indonesia/5/2005 (H5N1)	Fig. 1G
43	trình tự nucleotit của PDISP/HA từ cúm B/Brisbane/60/2008 có vòng phân hủy protein bị xóa bỏ	Fig 8E
44	đoạn mồi IF-H3V36111.S2+4c	Fig. 25A
45	đoạn mồi IF-H3V36111.s1-4r	Fig. 25B
46	trình tự nucleotit của gen H3 được tổng hợp	Fig. 25C
47	cat-xet biểu hiện số 1391	Fig. 25D
48	Trình tự axit amin của PDISP-H3 từ cúm A/ Victoria/361/2011	Fig. 25E
49	đoạn mồi IF-HAB110.S1+3c	Fig. 26A
50	đoạn mồi IF-HAB110.s1-4r	Fig. 26B
51	trình tự nucleotit của HA B/Wisconsin được tổng hợp (JN993010)	Fig. 26C
52	Cấu trúc 193	Fig. 26E

SEQ ID NO:	Mô tả	Hình vẽ
53	Cat-xet biểu hiện số 1462	Fig. 26F
54	Trình tự axit amin của HA từ cúm B/Wisconsin/1/2010	Fig. 26G
55	đoạn môi HAB110(PrL-).r	Fig. 27A
56	đoạn môi HAB110(PrL-).c	Fig. 27B
57	Cat-xet biểu hiện số 1467	Fig. 27C
58	Trình tự axit amin của HA từ cúm B/Wisconsin/1/2010 có vòng phân hủy protein bị xóa bỏ	Fig. 27D
59	đoạn môi IF-HB-M-04.s2+4c	Fig. 28A
60	đoạn môi IF-HB-M-04.s1-4r	Fig. 28B
61	trình tự nucleotit của HA B Malaysia được tổng hợp	Fig. 28C
62	Cấu trúc 194	Fig. 28E
63	Cat-xet biểu hiện số 1631	Fig. 28F
64	Trình tự axit amin của PDISP-HA từ cúm B/Malaysia/2506/2004	Fig. 28G

Ví dụ thực hiện sáng chế

Sáng chế sẽ còn được minh họa trong các ví dụ sau.

Vật liệu và phương pháp: Lắp ráp các cat-xet biểu hiện với protein cúm

A-2X35S/CPMV-HT/H5 Indonesia/NOS (Cấu trúc số 489)

Trình tự mã hóa H5 từ cúm A/Indonesia/5/2005 (H5N1) được tách dòng thành hệ biểu hiện 2X35S/CPMV-HT/NOS trong plasmit chứa cat-xet biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp dựa trên PCR sau. Phân đoạn chứa trình tự mã hóa H5 đầy đủ được khuếch đại bằng cách sử dụng các đoạn môi IF-H5A-I-05.s1+3c (Fig.1A, SEQ ID NO: 2) và IF-H5dTm.r (Fig.1B, SEQ ID NO: 3) bằng cách sử dụng cấu trúc số 972 (xem Fig.94, SEQ ID NO: 134 của WO 2010/003225, đối

với trình tự của cấu trúc số 972) làm khuôn. Sản phẩm PCR được tách dòng trong hệ biểu hiện 2X35S/CPMV-HT/NOS bằng cách sử dụng hệ tách dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 1191 (Fig. 1D, SEQ ID NO: 4) được cắt bằng enzym giới hạn SacII và StuI và plasmit được cắt thẳng được sử dụng trong phản ứng lắp ráp In-Fusion. Cấu trúc số 1191 là plasmit nhận dùng cho tách dòng “In Fusion” gen mong muốn trong cat-xet biểu hiện dựa trên CPMV-HT. Nó cũng kết hợp cấu trúc gen để biểu hiện đồng thời gen ức chế bất hoạt TBSV P19 dưới promoter và vùng kết thúc của gen Plastoxyanin của cỏ linh lăng. Mạch chính là plasmit mạch đôi pCAMBIA và trình tự từ bờ T-ADN từ trái sang phải được thể hiện trong Fig.1D (SEQ ID NO: 4). Cấu trúc thu được có số 489 (Fig.1E, SEQ ID NO: 5). Trình tự axit amin của H5 từ cúm A/Indonesia/5/2005 (H5N1) được thể hiện trong Fig.1F (SEQ ID NO: 6). Bản đồ của plasmit 489 được thể hiện trong Fig.15.

B-2X35S/CPMV-HT/M2 New Caledonia/NOS (Cấu trúc số 1261)

Trình tự mã hóa M2 từ cúm A/New Caledonia/20/1999 (H1N1) được tách dòng trong hệ biểu hiện 2X35S/CPMV-HT/NOS trong plasmit chứa cat-xet biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp dựa trên PCR sau. Mạch chứa trình tự mã hóa M2 đầy đủ được khuếch đại bằng cách sử dụng các đoạn mồi IF-S1-M1+M2ANC.c (Fig.2A, SEQ ID NO: 7) và IF-S1-4-M2ANC.r (Fig.2B, SEQ ID NO: 8) bằng cách sử dụng gen M2 được tổng hợp (tương ứng với nt 1-26 được kết hợp với nt 715-982 từ Số truy cập Genbank DQ508860) (Fig.2C, SEQ ID NO: 9) làm khuôn. Sản phẩm PCR được tách dòng trong hệ biểu hiện 2X35S/CPMV-HT/NOS bằng cách sử dụng hệ tách dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 1191 (Fig.1C) được cắt bằng enzym giới hạn SacII và StuI và plasmit được cắt thẳng được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 1191 plasmit nhận dùng cho tách dòng “In Fusion” gen mong muốn trong cat-xet biểu hiện dựa trên CPMV HT. Nó cũng kết hợp cấu trúc gen để biểu hiện đồng thời gen ức chế bất hoạt TBSV P19 dưới promoter và vùng kết thúc của gen Plastoxyanin của cỏ linh lăng. Mạch chính là plasmit mạch đôi pCAMBIA và trình tự từ các bờ T-ADN từ trái sang phải trong Fig.1D (SEQ ID NO: 4). Cấu trúc thu được có số 1261 (Fig.2D, SEQ ID NO: 10). Trình tự axit amin của M2 từ

cúm A/New Caledonia/20/1999 (H1N1) được thể hiện trong Fig.2E (SEQ ID NO: 11).
Bản đồ của plasmid 1261 được thể hiện trong Fig.16.

C-2X35S/CPMV-HT/M2 Puerto Rico/NOS (Cấu trúc số 859)

Trình tự mã hóa M2 từ cúm A/Puerto Rico/8/1934 (H1N1) được tách dòng trong hệ biểu hiện 2X35S/CPMV-HT/NOS trong plasmid chứa cat-xet biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp dựa trên PCR. Mảnh chứa trình tự mã hóa M2 đầy đủ được khuếch đại bằng cách sử dụng các đoạn mồi IF-S1-M1+M2ANC.c (Fig.2A, SEQ ID NO: 7) và IF-S1-4-M2ANC.r (Fig. 2B, SEQ ID NO: 8), bằng cách sử dụng gen M2 được tổng hợp (tương ứng với nt 26-51 được kết hợp với nt 740-1007 từ Số truy cập Genbank EF467824) (Fig.3A, SEQ ID NO: 12) làm khuôn. Sản phẩm PCR được tách dòng trong hệ biểu hiện 2X35S/CPMV-HT/NOS bằng cách sử dụng hệ tách dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 1191 (Fig.1C) được cắt bằng enzym giới hạn SacII và StuI và plasmid được cắt thẳng được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 1191 là plasmid nhận dùng cho tách dòng “In Fusion” gen mong muốn trong cat-xet biểu hiện dựa trên CPMV-HT. Cũng kết hợp cấu trúc gen để biểu hiện đồng thời gen ức chế bất hoạt TBSV P19 dưới promoter và vùng kết thúc của gen Plastoxanin của cỏ linh lăng. Vector là plasmid mạch đôi pCAMBIA và trình tự từ các bờ T-ADN từ trái sang phải trong Fig.1D (SEQ ID NO: 4). Cấu trúc thu được có số 859 (Fig. 3B, SEQ ID NO: 13). Trình tự axit amin của M2 từ cúm A/Puerto Rico/8/1934 (H1N1) được thể hiện trong Fig.3C (SEQ ID NO: 14). Bản đồ của plasmid 859 được thể hiện trong Fig.17.

D-2X35S/CPMV-HT/PDISP/H1 California/NOS (Cấu trúc số 484)

Trình tự mã hóa H1 từ cúm A/California/7/2009 (H1N1) được tách dòng trong hệ biểu hiện 2X35S-CPMV-HT-PDISP-NOS trong plasmid chứa cat-xet biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp dựa trên PCR. Mảnh chứa trình tự mã hóa H1 không có peptit tín hiệu kiểu đại được khuếch đại bằng cách sử dụng các đoạn mồi IF-H1A-C-09.s2+4c (Fig.4A, SEQ ID NO: 15) và IF-H1A-C-09.s1-4r (Fig.4B, SEQ ID NO: 16), bằng cách sử dụng H1 được tổng hợp (Số truy cập Genbank

FJ966974) (Fig.4C, SEQ ID NO: 17) làm khuôn. Sản phẩm PCR được tách dòng trong khung với peptit tín hiệu PDI của cỏ linh lăng trong hệ biểu hiện 2X35S/CPMV-HT/NOS bằng cách sử dụng hệ tách dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 1192 (Fig.4D) được cắt bằng enzym giới hạn SacII và StuI và plasmit được cắt thẳng được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 1192 là plasmit nhận dùng cho tách dòng “In Fusion” gen mong muốn trong khung có peptit tín hiệu PDI của cỏ linh lăng trong cat-xet biểu hiện dựa trên CPMV-HT. Cũng kết hợp cấu trúc gen để biểu hiện đồng thời TBSV P19 gen ức chế bất hoạt dưới promoter và vùng kết thúc của gen Plastoxyanin của cỏ linh lăng. Mạch chính là plasmit mạch đôi pCAMBIA và trình tự từ các bờ T-ADN từ trái sang phải Fig.4E (SEQ ID NO: 18). Cấu trúc thu được có số 484 (Fig.4F, SEQ ID NO: 19). Trình tự axit amin của PDISP/H1 từ cúm A/California/7/2009 (H1N1) được thể hiện trong Fig.4G (SEQ ID NO: 20). Bản đồ của plasmit 484 được thể hiện trong Fig.14.

E-2X35S/CPMV-HT/PDISP/H3 Perth/NOS (Cấu trúc số 1019)

Trình tự mã hóa H3 từ cúm A/Perth/16/2009 (H3N2) được tách dòng trong hệ biểu hiện 2X35S/CPMV-HT/PDISP/NOS trong plasmit chứa cat-xet biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp dựa trên PCR. Mạch chứa trình tự mã hóa H3 mà không có peptit tín hiệu kiểu đại bằng cách sử dụng các đoạn môi IF-S2+S4-H3 Per.c (Fig.5A, SEQ ID NO: 21) và IF-S1a4-H3 Per.r (Fig.5B, SEQ ID NO: 22), bằng cách sử dụng gen H3 được tổng hợp (tương ứng với nt 26-1726 từ Số truy cập Genbank GQ293081) (Fig.5C, SEQ ID NO: 23) làm khuôn. Sản phẩm PCR được tách dòng trong khung có peptit tín hiệu PDI của cỏ linh lăng trong hệ biểu hiện 2X35S/CPMV-HT/NOS bằng cách sử dụng hệ tách dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 1192 (Fig.4D) được cắt bằng enzym giới hạn SacII và StuI và plasmit được cắt thẳng được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 1192 là plasmit nhận dùng cho tách dòng “In Fusion” gen mong muốn trong khung có peptit tín hiệu PDI của cỏ linh lăng trong cat-xet biểu hiện dựa trên CPMV-HT. Cũng kết hợp cấu trúc gen để biểu hiện đồng thời gen ức chế bất hoạt TBSV P19 dưới promoter và vùng kết thúc của gen Plastoxyanin của cỏ linh lăng. Mạch chính là plasmit mạch đôi

pCAMBIA và trình tự từ các bờ T-ADN từ trái sang phải Fig.4E (SEQ ID NO: 18). Cấu trúc thu được có số 1019 (Fig. 5D, SEQ ID NO: 24). Trình tự axit amin của PDISP/H3 từ cúm A/Perth/16/2009 (H3N2) được thể hiện trong Fig.5E (SEQ ID NO: 25). Bản đồ của plasmit 1019 được thể hiện trong Fig.13.

F-2X35S/CPMV-HT/PDISP/HA B Brisbane/NOS (Cấu trúc số 1029)

Trình tự mã hóa HA từ cúm B/Brisbane/60/2008 được tách dòng trong hệ biểu hiện 2X35S/CPMV-HT/PDISP/NOS trong plasmit chứa cat-xet biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp dựa trên PCR. Mảnh chứa trình tự mã hóa HA B Brisbane không có peptit tín hiệu kiểu đại bằng cách sử dụng các đoạn mồi IF-S2+S4-B Bris.c (Fig.6A, SEQ ID NO: 26) và IF-S1a4-B Bris.r (Fig.6B, SEQ ID NO: 27), bằng cách sử dụng gen HA B Brisbane được tổng hợp (tương ứng với nt 34-1791 từ số truy cập Genbank FJ766840) (Fig.6C, SEQ ID NO: 28) làm khuôn. Sản phẩm PCR được tách dòng trong-khung có peptit tín hiệu PDI của cỏ linh lăng trong hệ biểu hiện 2X35S/CPMV-HT/NOS bằng cách sử dụng hệ tách dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 1192 (Fig. 4D) được cắt bằng enzym giới hạn SacII và StuI và plasmit được cắt thẳng được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 1192 là plasmit nhận dùng cho tách dòng “In Fusion” gen mong muốn trong khung có peptit tín hiệu PDI của cỏ linh lăng trong cat-xet biểu hiện dựa trên CPMV-HT. Cũng kết hợp cấu trúc gen để biểu hiện đồng thời gen ức chế bất hoạt TBSV P19 dưới promoter và vùng kết thúc của gen Plastoxyanin của cỏ linh lăng. Mạch chính là plasmit mạch đôi pCAMBIA và trình tự từ các bờ T-ADN từ trái sang phải trong Fig.4E (SEQ ID NO: 18). Cấu trúc thu được có số 1029 (Fig. 6D, SEQ ID NO: 29). Trình tự axit amin của PDISP/HA từ cúm B/Brisbane/60/2008 được thể hiện trong Fig.6E (SEQ ID NO: 30). Hình biểu diễn plasmit 1029 được thể hiện trong Fig.11.

G-2X35S/CPMV-HT/PDISP/HA B Brisbane/NOS trong hệ khuếch đại BeYDV+Replicaza (Cấu trúc số 1008)

Trình tự mã hóa HA từ cúm B/Brisbane/60/2008 được tách dòng trong 2X35S/CPMV-HT/PDISP/NOS bao gồm hệ khuếch đại BeYDV+replicaza trong

plasmid chứa cat-xet biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp dựa trên PCR. Mảnh chứa trình tự mã hóa HA B Brisbane mà không có peptit tín hiệu kiểu đại được khuếch đại bằng cách sử dụng các đoạn mồi IF-S2+S4-B Bris.c (Fig.6A, SEQ ID NO: 26) và IF-S1a4-B Bris.r (Fig.6B, SEQ ID NO: 27), bằng cách sử dụng gen HA B Brisbane được tổng hợp (tương ứng với nt 34-1791 từ số truy cập Genbank FJ766840) (Fig.6C, SEQ ID NO: 28) làm khuôn. Sản phẩm PCR được tách dòng trong-khung có peptit tín hiệu PDI của cỏ linh lăng trong cat-xet biểu hiện 2X35S/CPMV-HT/NOS trong hệ khuếch đại BeYDV bằng cách sử dụng hệ tách dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 1194 (xem các Fig. 6F và 6G) được cắt bằng enzym giới hạn SacII và StuI và plasmid được cắt thẳng được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 1194 là plasmid nhận dùng cho tách dòng “In Fusion” gen mong muốn trong khung có peptit tín hiệu PDI của cỏ linh lăng trong cat-xet biểu hiện dựa trên CPMV-HT trong hệ khuếch đại BeYDV. Cũng kết hợp cấu trúc gen để biểu hiện đồng thời gen ức chế bất hoạt TBSV P19 dưới promoter và vùng kết thúc của gen Plastoxyanin của cỏ linh lăng. Mạch chính là plasmid mạch đôi pCAMBIA và trình tự từ các bờ T-ADN từ trái sang phải Fig. 6G (SEQ ID NO: 31). Cấu trúc thu được có số 1008 (Fig.6H SEQ ID NO: 32). Trình tự axit amin của cúm PDISP/HA từ B/Brisbane/60/08 được thể hiện trong Fig.6E (SEQ ID NO: 30). Bản đồ của plasmid 1008 được thể hiện trong Fig.9.

Miền xuyên màng H-2X35S/CPMV-HT/PDISP/HA B Brisbane/H5 Indonesia và đuôi tế bào chất (H5Indo TMCT)/NOS vào hệ khuếch đại BeYDV+Replicaza (Cấu trúc số 1009)

Trình tự mã hóa HA từ cúm B/Brisbane/60/2008 miền ngoài được dung hợp với màng chuyển và các miền tế bào chất của H5 từ A/Indonesia/5/2005 (H5N1) được tách dòng trong 2X35S/CPMV-HT/PDISP/NOS bao gồm hệ khuếch đại BeYDV+replicaza trong plasmid chứa cat-xet biểu hiện Plasto_pro/P19/Plasto_ter do như dưới đây bằng cách sử dụng phương pháp nối dựa trên PCR được thể hiện bởi Darveau et al. (Methods in Neuroscience 26: 77-85 (1995)). Trong chu kỳ thứ nhất của PCR, mảnh chứa trình tự mã hóa ngoại miền HA B Brisbane mà không có peptit tín hiệu tự nhiên, màng chuyển và các miền tế bào chất được khuếch đại bằng cách sử dụng các đoạn mồi IF-S2+S4-B

Bris.c (Fig.6A, SEQ ID NO: 26) và dTmH5I-B Bris.r (Fig.7A, SEQ ID NO: 33), nhờ sử dụng gen HA B Brisbane được tổng hợp (tương ứng với nt 34-1791 từ Số truy cập Genbank FJ766840) (Fig.6C, SEQ ID NO: 28) làm khuôn. Mảnh thứ hai chứa màng chuyển và các miền tế bào chất của H5 Indonesia được khuếch đại bằng cách sử dụng các đoạn mồi B Bris-dTmH5I.c (Fig.7B, SEQ ID NO: 34) và IF-S1aS4-dTmH5I.r (Fig.7C, SEQ ID NO: 35), bằng cách sử dụng cấu trúc số 489 (xem Fig.1E, SEQ ID NO: 5) làm khuôn. Sản phẩm PCR từ cả hai lần khuếch đại tiếp đó được trộn và được sử dụng làm khuôn trong chu kỳ thứ hai khuếch đại bằng cách sử dụng IF-S2+S4-B Bris.c (Fig.6A, SEQ ID NO: 26) và IF-H5dTm.r (Fig.7C, SEQ ID NO: 34) làm các đoạn mồi. Mảnh thu được này được tách dòng trong khung có peptit tín hiệu PDI của cỏ linh lăng trong cat-xet biểu hiện 2X35S/CPMV-HT/NOS vào hệ khuếch đại BeYDV bằng cách sử dụng hệ tách dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 1194 (Fig.6F và 6G) được cắt bằng enzym giới hạn SacII và StuI và plasmit được cắt thẳng được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 1194 là plasmit nhận dùng cho tách dòng “In Fusion” gen mong muốn trong khung có peptit tín hiệu PDI của cỏ linh lăng trong cat-xet biểu hiện dựa trên CPMV HT vào hệ khuếch đại BeYDV. Cũng kết hợp cấu trúc gen để biểu hiện đồng thời gen ức chế bất hoạt TBSV P19 dưới promoter và vùng kết thúc của gen Plastoxyanin của cỏ linh lăng. Mạch chính là plasmit mạch đôi pCAMBIA và trình tự từ các bờ T-ADN từ trái sang phải trong Fig.6G (SEQ ID NO: 31). Cấu trúc thu được có số 1009 (Fig.7D, SEQ ID NO: 36). Trình tự axit amin của PDISP/HA B Brisbane/H5indo TMCT được thể hiện trong Fig.7E (SEQ ID NO: 37). Bản đồ của plasmit 1009 được thể hiện trong Fig.10.

I-2X35S/CPMV-HT/PDISP-HA B Brisbane có vòng phân hủy protein bị xóa bỏ trong hệ khuếch đại BeYDV+Replicaza (Cấu trúc số 1059)

Trình tự mã hóa HA từ cúm B/Brisbane/60/2008 có vòng phân hủy protein bị xóa bỏ được tách dòng trong 2X35S/CPMV-HT/PDISP/NOS bao gồm hệ khuếch đại BeYDV+replicaza trong plasmit chứa cat-xet biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp nối dựa trên PCR sau được thể hiện bởi Darveau et al. (Methods in Neuroscience 26: 77-85 (1995)). Trong chu kỳ thứ nhất của PCR, mảnh chứa

trình tự mã hóa HA B Brisbane từ nt 46 đến nt 1065 được khuếch đại bằng cách sử dụng các đoạn mồi IF-S2+S4-B Bris.c (Fig.6A, SEQ ID NO: 26) và 1039+1059.r (Fig.8A, SEQ ID NO: 38), nhờ sử dụng gen HA B Brisbane được tổng hợp (tương ứng với nt 34-1791 từ Sổ truy cập Genbank FJ766840) (Fig.6C, SEQ ID NO: 28) làm khuôn. Mẫu thứ hai, chứa trình tự mã hóa HA B Brisbane từ nt 1123 đến nt 1758, được khuếch đại bằng cách sử dụng các đoạn mồi 1039+1059.c (Fig.8B, SEQ ID NO: 39) và IF-S1a4-B Bris.r (Fig.6B, SEQ ID NO: 27), nhờ sử dụng gen HA B Brisbane được tổng hợp (tương ứng với nt 34-1791 từ Sổ truy cập Genbank FJ766840) (Fig.6C, SEQ ID NO: 28) làm khuôn. Các sản phẩm PCR từ cả hai lần khuếch đại tiếp đó được trộn và sử dụng làm khuôn trong chu kỳ khuếch đại thứ hai bằng cách sử dụng IF-S2+S4-B Bris.c (Fig.6A, SEQ ID NO: 26) và IF-H5dTm.r IF-S1a4-B Bris.r (Fig.6B, SEQ ID NO: 27) làm các đoạn mồi. Mẫu thu được (mã hóa HA B/Brisbane/60/2008 Δa.a. 356-374 có liên kết GG giữa các mảnh) được tách dòng trong khung có peptit tín hiệu PDI của cỏ linh lăng trong cat-xet biểu hiện 2X35S/CPMV-HT/NOS bao gồm hệ khuếch đại BeYDV bằng cách sử dụng hệ tách dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 1194 (Fig.6F và 6G) được cắt bằng enzym giới hạn SacII và StuI và plasmit được cắt thẳng được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 1194 là plasmit nhận dùng cho tách dòng “In Fusion” gen mong muốn trong khung có peptit tín hiệu PDI của cỏ linh lăng trong cat-xet biểu hiện dựa trên CPMV-HT vào hệ khuếch đại BeYDV. Cũng kết hợp cấu trúc gen để biểu hiện đồng thời gen ức chế bất hoạt TBSV P19 dưới promoter và vùng kết thúc của gen Plastoxyanin của cỏ linh lăng. Mạch chính là plasmit mạch đôi pCAMBIA và trình tự từ các bờ T-ADN từ trái sang phải Fig.6G (SEQ ID NO: 31). Cấu trúc thu được có số 1059 (Fig.8C, SEQ ID NO: 40). Trình tự axit amin của PDISP-HA B/Brisbane/60/2008 có vòng phân hủy protein bị xóa bỏ được thể hiện trong Fig.8D (SEQ ID NO: 41). Bản đồ của plasmit 1059 được thể hiện trong Fig.12.

A-2X35S/CPMV-HT/PDISP/H3 Victoria/NOS (Cấu trúc số 1391)

Trình tự mã hóa H3 từ cúm A/Victoria/361/2011 (H3N2) được tách dòng trong hệ biểu hiện 2X35S-CPMV-HT-PDISP-NOS trong plasmit chứa cat-xet biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp dựa trên PCR. Mẫu chứa

trình tự mã hóa H3 mà không có peptit tín hiệu kiểu dài được khuếch đại bằng cách sử dụng các đoạn mồi IF-H3V36111.S2+4c (Fig.25A, SEQ ID NO: 44) và IF-H3V36111.s1-4r (Fig.25B, SEQ ID NO: 45), nhờ sử dụng gen H3 được tổng hợp (tương ứng với nt 25 đến 1725 từ trình tự HA phân lập GISAID EPI_ISL_101506) (Fig.25C, SEQ ID NO: 46) làm khuôn. Sản phẩm PCR được tách dòng trong khung có peptit tín hiệu PDI của cỏ linh lăng trong hệ biểu hiện 2X35S/CPMV-HT/NOS bằng cách sử dụng hệ tách dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 1192 (Fig.4D) được cắt bằng enzym giới hạn SacII và StuI và plasmit được cắt thẳng được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 1192 là plasmit nhận dùng cho tách dòng “In Fusion” gen mong muốn trong khung có peptit tín hiệu PDI của cỏ linh lăng trong cat-xet biểu hiện dựa trên CPMV-HT. Cũng kết hợp cấu trúc gen để biểu hiện đồng thời gen ức chế bất hoạt TBSV P19 dưới promoter và vùng kết thúc của gen Plastoxyanin của cỏ linh lăng. Mạch chính là plasmit mạch đôi pCAMBIA và trình tự từ các bờ T-ADN từ trái sang phải Fig. 4E (SEQ ID NO: 18). Cấu trúc thu được có số 1391 (Fig. 25D, SEQ ID NO: 47). Trình tự axit amin của PDISP/H3 từ Cúm A/Victoria/361/2011 (H3N2) được thể hiện trong Fig.25E (SEQ ID NO: 48). Bản đồ của plasmit 1391 được thể hiện trong Fig.25F.

B-2X35S/CPMV-HT/HA B Wisconsin/NOS trong hệ khuếch đại BeYDV(m)+Replicaza (Cấu trúc số 1462)

Trình tự mã hóa HA từ cúm B/Wisconsin/1/2010 được tách dòng vào 2X35S/CPMV-HT/NOS bao gồm hệ khuếch đại BeYDV(m)+Replicaza trong plasmit chứa cat-xet biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp dựa trên PCR. Mảnh chứa trình tự mã hóa HA B Wisconsin đầy đủ được khuếch đại bằng cách sử dụng các đoạn mồi IF-HAB110.S1+3c (Fig.26A, SEQ ID NO: 49) và IF-HAB110.s1-4r (Fig.26B, SEQ ID NO: 50), nhờ sử dụng gen HA B Wisconsin được tổng hợp (Số truy cập Genbank JN993010) (Fig.26C, SEQ ID NO: 51) làm khuôn. Sản phẩm PCR được tách dòng trong 2X35S/CPMV-HT/NOS cat-xet biểu hiện vào hệ khuếch đại BeYDV(m) bằng cách sử dụng hệ tách dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 193 (Fig.26D) được cắt bằng enzym giới hạn SacII và StuI và plasmit được cắt

thẳng được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 193 là plasmit nhận dùng cho tách dòng “In Fusion” gen mong muốn trong cat-xet biểu hiện CPMV-HT vào hệ khuếch đại BeYDV(m). Cũng kết hợp cấu trúc gen để biểu hiện đồng thời gen ức chế bất hoạt TBSV P19 dưới promoter và vùng kết thúc của gen Plastoxyanin của cỏ linh lăng. Mạch chính là plasmit mạch đôi pCAMBIA và trình tự từ các bờ T-ADN từ trái sang phải Fig.26E (SEQ ID NO: 52). Cấu trúc thu được có số 1462 (Fig. 26F, SEQ ID NO: 53). Trình tự axit amin của PDISP/HA từ Cúm B/Wisconsin/1/2010 được thể hiện trong Fig.26G (SEQ ID NO: 54). Bản đồ của plasmit 1462 được thể hiện trong Fig.26H.

C-2X35S/CPMV-HT/HA B Wisconsin có vòng phân hủy protein bị xóa bỏ trong hệ khuếch đại BeYDV(m)+Replicaza (Cấu trúc số 1467)

Trình tự mã hóa HA từ cúm B/Wisconsin/1/2010 có vòng phân hủy protein bị xóa bỏ được tách dòng trong 2X35S/CPMV-HT/NOS bao gồm hệ khuếch đại BeYDV(m)+Replicaza trong plasmit chứa cat-xet biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp nối dựa trên PCR sau được thể hiện bởi Darveau et al. (Methods in Neuroscience 26: 77-85 (1995)). Trong chu kỳ thứ nhất của PCR, mảnh chứa trình tự mã hóa HA B Wisconsin từ nt 1 đến nt 1062 được khuếch đại bằng cách sử dụng các đoạn mồi IF-HAB110.S1+3c (Fig.26A, SEQ ID NO: 49) và HAB110(PrL-).r (Fig.27A, SEQ ID NO: 55), nhờ sử dụng gen HA B Wisconsin được tổng hợp (Số truy cập Genbank JN993010) (Fig.26C, SEQ ID NO: 51) làm khuôn. Mảnh thứ hai, chứa trình tự mã hóa HA B Wisconsin từ nt 1120 đến nt 1755, được khuếch đại bằng cách sử dụng các đoạn mồi HAB110(PrL-).c (Fig.27B, SEQ ID NO: 56) và IF-HAB110.s1-4r (Fig.26B, SEQ ID NO: 50), nhờ sử dụng gen HA B Wisconsin được tổng hợp (Số truy cập Genbank JN993010) (Fig.26C, SEQ ID NO: 51) làm khuôn. Các sản phẩm PCR từ cả hai lần khuếch đại tiếp được trộn và được sử dụng làm khuôn trong chu kỳ thứ hai bằng cách sử dụng IF-HAB110.S1+3c (Fig.26A, SEQ ID NO: 49) và IF-HAB110.s1-4r (Fig.26B, SEQ ID NO: 50) làm các đoạn mồi. Mảnh thu được (mã hóa HA B/Wisconsin/1/2010 Δa.a. 340-358 có liên kết GG giữa các mảnh) được tách dòng trong cat-xet biểu hiện 2X35S/CPMV-HT/NOS bao gồm hệ khuếch đại BeYDV(m) bằng cách sử dụng hệ tách dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 193 (Fig.

26D) được cắt bằng enzym giới hạn SacII và StuI và plasmit được cắt thẳng được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 193 là plasmit nhận dùng cho tách dòng “In Fusion” gen mong muốn trong cat-xet biểu hiện dựa trên CPMV-HT vào hệ khuếch đại BeYDV(m). Cũng kết hợp cấu trúc gen để biểu hiện đồng thời nhân tố ức chế bất hoạt TBSV P19 dưới promoter và vùng kết thúc của gen Plastoxyanin của cỏ linh lăng. Mạch chính là plasmit mạch đôi pCAMBIA và trình tự từ các bờ T-ADN từ trái sang phải Fig.26E (SEQ ID NO: 52). Cấu trúc thu được có số 1467 (Fig.27C, SEQ ID NO: 57). Trình tự axit amin của HA từ Cúm B/Wisconsin/1/2010 có vòng phân hủy protein bị xóa bỏ được thể hiện trong Fig.27D (SEQ ID NO: 58). Bản đồ của plasmit 1467 được thể hiện trong Fig.27E.

D-2X35S/CPMV-HT/PDISP/HA B Malaysia/NOS vào hệ khuếch đại BeYDV(m)+Replicaza (Cấu trúc số 1631)

Trình tự mã hóa HA từ cúm B/Malaysia/2506/2004 được tách dòng trong 2X35S/CPMV-HT/PDISP/NOS bao gồm hệ khuếch đại BeYDV(m)+Replicaza trong plasmit chứa cat-xet biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp dựa trên PCR. Mảnh chứa trình tự mã hóa HA B Malaysia mà không có peptit tín hiệu kiểu đại được khuếch đại bằng cách sử dụng các đoạn mồi IF-HB-M-04.s2+4c (Fig.28A, SEQ ID NO: 59) và IF-HB-M-04.s1-4r (Fig.28B, SEQ ID NO: 60), nhờ sử dụng gen HA B Malaysia được tổng hợp (tương ứng với nt 31-1743 từ số truy cập Genbank EU124275). Các đột biến câm T759C và C888G được chèn vào trình tự tổng hợp để biến đổi các vị trí nhận biết enzym giới hạn DraIII và BamHI (Fig.28C, SEQ ID NO: 61) làm khuôn. Sản phẩm PCR được tách dòng trong khung có peptit tín hiệu PDI của cỏ linh lăng trong cat-xet biểu hiện 2X35S/CPMV-HT/NOS vào hệ khuếch đại BeYDV(m) bằng cách sử dụng hệ tách dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 194 (Fig.28D) được cắt bằng enzym giới hạn SacII và StuI và plasmit được cắt thẳng được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 194 là plasmit nhận dùng cho tách dòng “In Fusion” gen mong muốn trong khung với peptit tín hiệu PDI của cỏ linh lăng trong CPMV-HT dựa trên cat-xet biểu hiện vào hệ khuếch đại BeYDV(m). Cũng kết hợp cấu trúc gen để biểu hiện đồng thời gen TBSV P19 ức chế bất hoạt dưới promoter và

vùng kết thúc của gen Plastoxyanin của cỏ linh lăng. Mạch chính là plasmit mạch đôi pCAMBIA và trình tự từ các bờ T-ADN từ trái sang phải Fig.28E (SEQ ID NO: 62). Cấu trúc thu được có số 1631 (Fig.28F, SEQ ID NO: 63). Trình tự axit amin của PDISP/HA từ Cúm B/Malaysia/2506/2004 được thể hiện trong Fig.28G (SEQ ID NO: 64). Bản đồ của plasmit 1631 được thể hiện trong Fig.28H.

Chuyển nhiễm *Agrobacterium*

Chủng *Agrobacterium* AGL1 được chuyển nhiễm các cấu trúc ADN bằng xung điện bằng cách sử dụng các phương pháp được mô tả bởi D'Aoust et al 2008 (Plant Biotechnology Journal 6:930-940). *Agrobacterium* chuyển nhiễm được sinh trưởng trong môi trường YEB được bổ sung axit 2-(N-morpholino)etansulfonic (MES) 10mM, axetosyringon 20 μ M, kanamycin 50 μ g/ml và carbenixilin 25 μ g/ml pH= 5,6 với OD₆₀₀ nằm trong khoảng từ 0,6 đến 1.6. Ly tâm huyền phù *Agrobacterium* trước khi sử dụng và tạo huyền phù lại trong môi trường thẩm (MgCl₂ 10mM và MES 10mM pH 5,6).

Chuẩn bị sinh khối thực vật, cấy và thẩm bằng *Agrobacterium*

Như được sử dụng ở đây, thuật ngữ "sinh khối" và "vật liệu thực vật" để chỉ vật liệu bất kỳ thu được từ thực vật. Sinh khối hoặc vật liệu thực vật có thể bao gồm toàn bộ thực vật, mô, tế bào hoặc phần bất kỳ của nó. Hơn nữa, sinh khối hoặc vật liệu thực vật có thể bao gồm các thành phần thực vật nội bào, thành phần thực vật ngoại bào, phần chiết dạng lỏng hoặc rắn của thực vật, hoặc tổ hợp của chúng. Hơn nữa, sinh khối hoặc vật liệu thực vật có thể bao gồm thực vật, tế bào thực vật, mô, phần chiết dạng lỏng, hoặc tổ hợp của chúng, từ lá cây, thân, quả, rễ hoặc tổ hợp của chúng. Một phần thực vật có thể bao gồm vật liệu thực vật hoặc sinh khối.

Cây *Nicotiana benthamiana* được sinh trưởng từ hạt trong các tấm phẳng được làm đầy bằng giá thể than bùn thương mại. Các cây được sinh trưởng trong nhà kính dưới chu kỳ chiếu sáng 16/8 và chế độ nhiệt độ là 25°C ngày/20°C đêm. Ba tuần sau khi gieo hạt, nhặt riêng các cây non, cấy vào các lọ và để sinh trưởng trong nhà kính trong thêm ba tuần dưới cùng các điều kiện môi trường.

Agrobacteria được chuyển nhiễm mỗi cấu trúc được sinh trưởng trong môi trường YEB được bổ sung axit 2-(N-morpholino)etansulfonic (MES) 10mM, axetosyringon 20 μ M, kanamycin 50 μ g/ml và carbenixilin 25 μ g/ml pH=5,6 cho đến khi chúng đạt OD₆₀₀ nằm trong khoảng từ 0,6 đến 1,6. Các huyền phù *Agrobacterium* được ly tâm trước khi dùng và được tái huyền phù trong môi trường thẩm (MgCl₂ 10mM và MES 10mM có độ pH=5,6) và được bảo quản qua đêm ở 4°C. Vào ngày thẩm, các mẻ nuôi cấy được pha loãng 2,5 thể tích nuôi cấy và cho phép làm ấm trước khi dùng. Toàn bộ các cây *N. benthamiana* được đặt lộn ngược trong huyền phù vi khuẩn trong thùng thép không gỉ kín khí trong điều kiện chân không áp suất từ 20 đến 40 Torr (~0,029-0,053atm) trong 2 phút. Đưa các cây trở lại nhà kính trong thời gian ủ từ 2 đến 6 ngày cho đến khi thu hoạch.

Thu hoạch lá và chiết protein tổng số

Sau khi ủ, phần trên tiếp xúc với không khí của cây được thu hoạch, đông lạnh ở -80°C và nghiền thành các miếng. Chiết protein hòa tan tổng số bằng cách đồng nhất hóa (Polytron), mỗi mẫu vật liệu thực vật được nghiền đông lạnh trong 3 lần thể tích của Tris lạnh 50mM độ pH=8,0, NaCl 0,15M, Triton X-100 0,1% và phenylmetansulfonyl florua 1mM. Sau khi đồng nhất hóa, ly tâm huyền phù đặc ở 10.000g (~ 9449 vòng/phút với bán kính rô-to là 100mm) trong 10 phút ở 4°C và các phần chiết thô được làm trong (dịch nổi) này được giữ cho các bước phân tích.

Phân tích protein và thẩm tách miễn dịch

Hàm lượng protein tổng của dịch chiết thô được phân loại được xác định bằng thử nghiệm Bradford (Bio-Rad, Hercules, CA) bằng cách sử dụng albumin huyết thanh của bò làm tiêu chuẩn tham chiếu. Protein được tách bằng SDS-PAGE và điện chuyển lên trên màng polyvinyliden diflorua (PVDF) (Roche Diagnostics Corporation, Indianapolis, IN) để phát hiện miễn dịch. Trước khi thẩm tách miễn dịch, khóa cố định màng bằng sữa gầy 5% và Tween-20 0,1% trong muối đệm Tris (TBS-T) trong từ 16 đến 18h ở 4°C.

Thẩm tách miễn dịch được thực hiện với lần ủ thứ nhất bằng kháng thể thứ nhất (Bảng 4 trình bày các kháng thể và các điều kiện được sử dụng để phát hiện mỗi HA), trong 2 μ g/ml trong 2% sữa gầy trong TBS-Tween 20 0,1%. Các kháng thể thứ cấp được

sử dụng để phát hiện phát quang hóa học được chỉ ra trong Bảng 4, được pha loãng như được chỉ ra trong 2% sữa gầy trong TBS-Tween 20 0,1%. Các phức phản ứng miễn dịch được phát hiện bằng phát quang hóa học bằng cách sử dụng luminol làm cơ chất (Roche Diagnostics Corporation). Thử nghiệm tiếp hợp enzym peroxidaza Cải ngựa của kháng thể IgG của người được thực hiện bằng cách sử dụng kit tiếp hợp EZ-Link Plus® Activated Peroxidaza (Pierce, Rockford, IL).

Bảng 4: Các điều kiện điện di, kháng thể, và pha loãng để thẩm tách miễn dịch các protein được biểu hiện.

Typ phụ HA	Chủng cúm	Điều kiện điện di	Kháng thể sơ cấp	Pha loãng	Kháng thể thứ cấp	Pha loãng
B	B/Brisbane/60/2008	Không khử	TGA, AS397	1:20000	Kháng cừu của thỏ (JIR 313-035-045)	1:10 000
B	B/Wisconsin/1/2010	Không khử	NIBSC 07/356	1 :2000	Kháng cừu của thỏ (JIR 313-035-045)	1:10 000
B	B/Malaysia/2506/2004	Không khử	NIBSC 07/184	1 :2000	Kháng cừu của thỏ (JIR 313-035-045)	1:10 000
H3	A/Perth/16/2009 (H3N2)	Không khử	TGA, AS400	1:20000	Kháng cừu của thỏ (JIR 313-035-045)	1:10 000
H3	A/Victoria/361/2011	Không khử	TGA, AS400	1 :2000 0	Kháng cừu của thỏ (JIR 313-035-045)	1:10 000

H1	A/California/07/2009 (H1N1)	Không khử	Sino, 11055- MMO1	1 µg/ml	Kháng chuột của đê (JIR 115-035- 146)	1:7 500
H5	A/Indonesia/05/2005 (H5N1)	Khử	CBER, S-7858	1:4000	Kháng cừu của thỏ (JIR 313- 035-045)	1:10 000

JIR: Jackson ImmunoResearch, West Grove, PA, USA;

CBER: Center for Biologics Evaluation and Research, Rockville, MD, USA.

Sino: Sino Biological inc., Beijing, China.

TGA: Therapeutic Goods Administration, Australia.

NIBSC: National Institute for Biological Standards and Control, United Kingdom

Thử nghiệm hemagglutinin hóa

Thử nghiệm hemagglutinin hóa dựa trên phương pháp được mô tả bởi Nayak và Reichl (2004). Tóm lại, pha loãng hai lần liên tiếp các mẫu thử nghiệm (100 µL) được thực hiện trong các đĩa microlit 96 giếng đáy chữ V chứa 100µL PBS, còn lại 100µL mẫu được pha loãng trên mỗi giếng. Một trăm microlit huyền phù tế bào hồng cầu đỏ của gà tây 0,25% (Bio Link Inc., Syracuse, NY) được bổ sung vào mỗi giếng, và ủ các đĩa trong 2 giờ ở nhiệt độ trong phòng. Nghịch đảo của độ pha loãng cao nhất thể hiện sự hemagglutinin hóa hoàn toàn được ghi nhận là hoạt tính HA. Đồng thời, mẫu chuẩn HA tái tổ hợp (A/Vietnam/1203/2004 H5N1) (Protein Science Corporation, Meriden, CT) được pha loãng trong PBS và làm đối chứng trên mỗi đĩa.

Ví dụ 1: Ảnh hưởng của sự biểu hiện đồng thời M2 cúm lên mức tích lũy HA và H3 typ B

Ảnh hưởng của sự biểu hiện đồng thời M2 cúm lên mức tích lũy HA từ các chủng cúm khác nhau được phân tích bằng cách biến nạp đồng thời các cấu trúc điều khiển sự biểu hiện của HA với cấu trúc biểu hiện M2 từ cúm A/New Caledonia/20/1999 (H1N1) trong hệ biến nạp nhất thời dựa trên thảm bằng Argobacteria.

Phân tích thẩm tách Western của các dịch chiết protein từ thực vật được biến nạp các cấu trúc gen điều khiển sự biểu hiện HA cúm B (từ B/Brisbane/60/2008) (các cấu trúc số 1008, 1009 và 1029) với sự có mặt hoặc vắng mặt của cấu trúc M2 (cấu trúc số 1261) đã thể hiện rằng sự biểu hiện đồng thời M2 làm tăng sự tích tụ HA cúm B (Fig.18). Tương tự, sự biểu hiện đồng thời M2 với H3 from cúm A/Perth/16/2009 (cấu trúc số 1019+1261) làm tăng sự tích tụ H3 trong các cây được biến nạp khi được so sánh với các cây chỉ được biến nạp cấu trúc biểu hiện H3 (cấu trúc số 1019) như được thể hiện trong Fig.19.

Phân tích thẩm tách Western của các dịch chiết protein từ các cây biểu hiện đồng thời M2 với H1 từ cúm A/California/07/2009 đã thể hiện rằng sự biểu hiện đồng thời M2 với H1 làm giảm nhẹ mức tích lũy H1 (Fig.20, 484 với 484+1261). Sự biểu hiện đồng thời của M2 với H5 từ cúm A/Indonesia/05/2005 cũng làm giảm sự tích lũy H5 khi được so sánh với sự biểu hiện riêng H5 (Fig.21, 489 với 489+1261).

Sự biểu hiện đồng thời M2 còn được đánh giá sự tác động của nó đến sự tích lũy HA cúm B được biến đổi. Cấu trúc số 1059 mã hóa HA cúm B mà trong đó vòng phân hủy protein được thay thế bằng 2 liên kết axit amin (GG thay cho aa 341-359). Các kết quả từ phân tích thẩm tách Western được thể hiện trong Fig.22A thể hiện rằng việc loại bỏ vòng phân hủy protein làm tăng mức tích lũy HA cúm B (so sánh 1008 với 1059) và sự biểu hiện đồng thời M2 với HA cúm B được biến đổi còn làm tăng mức tích lũy HA (Fig.22A, 1059 với 1059+1261). Phân tích hoạt tính hemagglutinin hóa trên dịch chiết protein thô từ các cây được biến nạp với HA cúm B có hoặc không có biến đổi và có hoặc không có sự biểu hiện đồng thời M2 đã xác nhận ảnh hưởng tích cực của sự biểu hiện đồng thời M2 lên mức tích lũy của HA cúm B tự nhiên (Fig.22B, 1008 với 1008+1261) và HA cúm B được biến đổi (Fig.22B, 1059 với 1059+1261).

Hiệu quả của M2 từ cúm A/Puerto Rico/8/1934 để tăng sự tích lũy HA cúm B và H3 được biến đổi đã so sánh với hiệu quả của M2 từ cúm A/New Caledonia/20/1999. Đối với HA cúm B được biến đổi, so sánh này được đảm nhiệm bằng phân tích thẩm tách Western của các dịch chiết protein từ các cây được biến nạp có các cấu trúc 1059, 1059+1261 và 1059+859. Đối với H3, so sánh tương tự được thực hiện trên các dịch

chiết protein từ các cây có 1019, 1019+1261 và 1019+859. Các kết quả thu được đã chứng minh rằng sự biểu hiện đồng thời M2 từ cúm A/Puerto Rico/8/1934 (được mã hóa bằng cấu trúc số 859) có hiệu quả giống như sự biểu hiện của M2 từ cúm A/New Caledonia/20/1999 (được mã hóa bởi cấu trúc số 1261) để làm tăng sự tích lũy của cả hai HA cúm B được biến đổi (Fig.23A) và H3 (Fig.23B).

Ví dụ 2: Ảnh hưởng của sự biểu hiện đồng thời cúm M2 lên mức tích lũy của các chủng khác nhau của HA và H3 B

Phân tích thẩm tách Western các dịch chiết protein từ các cây được biến nạp các cấu trúc gen mà điều khiển sự biểu hiện của HA cúm B (từ B/Malaysia/2506/2004) (cấu trúc số 1631) với sự có mặt hoặc vắng mặt của cấu trúc biểu hiện M2 (cấu trúc số 1261) đã thể hiện rằng sự biểu hiện đồng thời M2 làm tăng sự tích lũy HA cúm B (Fig.29).

Phân tích thẩm tách Western của các dịch chiết protein từ các cây được biến nạp các cấu trúc gen mà chạy sự biểu hiện của HA cúm B (từ B/Wisconsin/1/2010) (cấu trúc số 1462) với sự có mặt hoặc vắng mặt của cấu trúc biểu hiện M2 (cấu trúc số 1261) đã thể hiện rằng sự biểu hiện đồng thời M2 làm tăng sự tích lũy của HA cúm B (Fig.30).

Sự biểu hiện đồng thời M2 còn được đánh giá sự tác động của nó lên mức tích lũy của HA cúm B được biến đổi. Cấu trúc số 1467 mã hóa HA cúm B mà trong đó vòng phân hủy protein được thay thế bằng 2 liên kết axit amin (GG thay cho aa 341-359). Các kết quả từ phân tích thẩm tách Western được thể hiện trong Fig.30A chỉ ra rằng sự loại bỏ vòng phân hủy protein làm tăng mức tích lũy HA cúm B (so sánh 1462 với 1467) và sự biểu hiện đồng thời M2 với HA cúm B được biến đổi còn làm tăng mức tích lũy HA (Fig.30A, 1467 với 1467+1261). Phân tích hoạt tính hemagglutinin hóa trên các dịch chiết protein thô từ các cây được biến nạp HA cúm B có hoặc không biến đổi và có hoặc không biểu hiện đồng thời M2 đã xác nhận ảnh hưởng tích cực của sự biểu hiện đồng thời M2 lên mức tích lũy của HA cúm B tự nhiên (Fig.30B, 1462 với 1462+1261) và HA cúm B được biến đổi (Fig.26B, 1467 với 1467+1261).

Phân tích thẩm tách Western các dịch chiết protein từ các cây được biến nạp các cấu trúc gen điều khiển sự biểu hiện H3 cúm (từ H3/Victoria/361/2011) (cấu trúc số

1391) với sự có mặt hoặc vắng mặt cấu trúc biểu hiện M2 (cấu trúc số 1261) đã chỉ ra rằng sự biểu hiện đồng thời M2 làm tăng sự tích lũy H3 cúm (Fig.31).

Tất cả các tài liệu trích dẫn được đưa ra trong bản mô tả sáng chế này để tham khảo.

Sáng chế được mô tả liên quan đến một hoặc nhiều phương án. Tuy nhiên, sẽ rõ ràng với người có hiểu biết trung bình trong lĩnh vực này rằng một số biến đổi và biến thể có thể được thực hiện mà không tách rời phạm vi của sáng chế được xác định trong yêu cầu bảo hộ.

YÊU CẦU BẢO HỘ

1. Phương pháp sản xuất hạt giống như virus (virus like particle - VLP) ở thực vật bao gồm các bước:

a) đưa axit nucleic thứ nhất bao gồm vùng điều hòa thứ nhất hoạt động ở thực vật và được liên kết theo kiểu điều khiển được với trình tự nucleotit mã hóa protein hemagglutinin cúm (HA) được chọn từ nhóm gồm có HA cúm B và HA H3 cúm vào thực vật, hoặc phần của thực vật,

b) đưa axit nucleic thứ hai bao gồm vùng điều hòa thứ hai hoạt động ở thực vật và được liên kết theo kiểu điều khiển được với trình tự nucleotit mã hóa protein kênh proton và

c) ủ thực vật hoặc phần của thực vật trong các điều kiện cho phép sự biểu hiện của axit nucleic, theo đó tạo ra VLP.

2. Phương pháp theo điểm 1, trong đó protein kênh proton được chọn từ M2 hoặc BM2.

3. Phương pháp theo điểm 1, trong đó protein kênh proton bao gồm trình tự tín hiệu kênh proton HXXXW.

4. Phương pháp theo điểm 1, trong đó một hoặc nhiều vòng phân hủy của protein cúm HA bị xóa bỏ.

5. Phương pháp theo điểm 1, trong đó HA cúm B là từ cúm B/Brisbane/60/2008, B/Malaysia/2506/2004 hoặc B/Wisconsin/1/2010.

6. Phương pháp theo điểm 5, trong đó một hoặc nhiều vòng phân hủy của protein HA cúm B bị xóa bỏ.

7. Phương pháp theo điểm 1, trong đó HA H3 cúm là từ cúm A/Perth/16/2009 hoặc từ cúm A/Victoria/361/2011.

8. Phương pháp theo điểm 1, trong đó trình tự nucleotit mã hóa protein hemagglutinin cúm (HA) có độ tương đồng trình tự ít nhất là 90% với trình tự được chọn từ nhóm gồm có SEQ ID NO: 28, SEQ ID NO:51 và SEQ ID NO:61.

9. Phương pháp theo điểm 1, trong đó trình tự của protein hemagglutinin cúm (HA) là như được thể hiện trong ít nhất một trong số SEQ ID NO:30, SEQ ID NO:54 và SEQ ID NO:64.

10. Phương pháp theo điểm 1, trong đó trình tự nucleotit mã hóa protein hemagglutinin cúm (HA) có độ tương đồng trình tự ít nhất là 90% với trình tự được chọn từ nhóm gồm có SEQ ID NO: 43 và phần được gạch dưới của SEQ ID NO:57.

11. Phương pháp theo điểm 1, trong đó trình tự của protein hemagglutinin cúm (HA) là như được thể hiện trong ít nhất một trong SEQ ID NO:41 và SEQ ID NO:58.

12. Phương pháp theo điểm 1, trong đó trình tự nucleotit mã hóa protein hemagglutinin cúm (HA) có độ tương đồng trình tự ít nhất là 90% với trình tự được chọn từ nhóm gồm có SEQ ID NO: 23 và SEQ ID NO:46.

13. Phương pháp theo điểm 1, trong đó trình tự của protein hemagglutinin cúm (HA) là như được thể hiện trong ít nhất một trong số SEQ ID NO:25 và SEQ ID NO:48.

14. Phương pháp theo điểm 1, trong đó axit nucleic thứ nhất bao gồm vùng điều hòa thứ nhất được liên kết theo kiểu điều khiển được với một hoặc nhiều vùng tăng cường comovirut, trình tự nucleotit mã hóa protein hemagglutinin cúm (HA), và một hoặc nhiều nhân tố khuếch đại geminivirut, và axit nucleic thứ ba bao gồm trình tự nucleotit mã hóa replicaza geminivirut được đưa vào thực vật hoặc phần của thực vật.

15. Phương pháp theo điểm 14, trong đó một hoặc nhiều vùng tăng cường comovirut là UTR của comovirut.

16. Phương pháp theo điểm 15, trong đó UTR của comovirut là UTR của virus thể khảm gây bệnh ở cây đậu đũa (Cowpea Mosaic Virus - CPMV).

17. Phương pháp theo điểm 14, trong đó một hoặc nhiều nhân tố khuếch đại geminivirus được chọn từ vùng liên gen dài của virus gây lùn, vàng ở cây đậu (Bean Yellow Dwarf Virus long intergenic region - BeYDV LIR), và vùng liên gen ngắn của virus gây lùn, vàng ở cây đậu (Bean Yellow Dwarf Virus short intergenic region - BeYDV SIR).

18. Phương pháp theo điểm 14, trong đó trình tự nucleotit mã hóa protein hemagglutinin cúm (HA) có độ tương đồng trình tự ít nhất là 90% với trình tự được chọn từ nhóm gồm có SEQ ID NO: 28, SEQ ID NO: 51 và SEQ ID NO: 61.

19. Phương pháp theo điểm 14, trong đó trình tự của protein hemagglutinin cúm (HA) là như được thể hiện trong ít nhất một trong số SEQ ID NO: 30, SEQ ID NO: 54 và SEQ ID NO: 64.

20. Phương pháp theo điểm 14, trong đó trình tự nucleotit mã hóa protein hemagglutinin cúm (HA) có độ tương đồng trình tự ít nhất là 90% với trình tự được chọn từ nhóm bao gồm SEQ ID NO: 43 và phần được gạch dưới của SEQ ID NO: 57.

21. Phương pháp theo điểm 14, trong đó trình tự của protein hemagglutinin cúm (HA) là như được thể hiện trong ít nhất một trong số SEQ ID NO: 41 và SEQ ID NO: 58.

22. Phương pháp theo điểm 14, trong đó trình tự nucleotit mã hóa protein hemagglutinin cúm (HA) có độ tương đồng trình tự ít nhất là 90% với trình tự được chọn từ nhóm gồm có SEQ ID NO: 23 và SEQ ID NO: 46.

23. Phương pháp theo điểm 14, trong đó trình tự của protein hemagglutinin cúm (HA) là như được thể hiện trong ít nhất một trong số SEQ ID NO: 25 và SEQ ID NO: 48.

24. Phương pháp theo điểm 1, trong đó ở bước đưa vào (bước a), axit nucleic được biểu hiện nhất thời trong thực vật.

25. Phương pháp theo điểm 1, trong đó, ở bước đưa vào (bước a), axit nucleic được biểu hiện ổn định ở thực vật.

26. Phương pháp theo điểm 1, phương pháp này còn bao gồm bước:

d) thu hoạch thực vật và tinh chế các VLP.

27. Phương pháp theo điểm 1, trong đó VLP không chứa protein nền virus, protein lõi hoặc protein kênh.

28. Phương pháp sản xuất hạt giống như virus (VLP) bao gồm các bước:

a) tạo ra thực vật hoặc phần của thực vật bao gồm axit nucleic thứ nhất bao gồm vùng điều hòa thứ nhất hoạt động ở thực vật và được liên kết theo kiểu điều khiển được với trình tự nucleotit mã hóa protein hemagglutinin cúm (HA) được chọn từ nhóm gồm có HA cúm B và HA H3 cúm, và axit nucleic thứ hai bao gồm vùng điều hòa thứ hai hoạt động ở thực vật và được liên kết theo kiểu điều khiển được với trình tự nucleotit mã hóa protein kênh proton và

b) ủ thực vật hoặc phần của thực vật dưới các điều kiện cho phép sự biểu hiện của axit nucleic, theo đó tạo ra VLP.

29. Phương pháp theo điểm 28, trong đó protein kênh proton được chọn từ M2, BM2 hoặc protein kênh proton này bao gồm trình tự tín hiệu kênh proton HXXXW.

30. Hạt giống như virus được tạo ra bằng phương pháp theo điểm 4.

31. Hạt giống như virus theo điểm 30, trong đó VLP này còn bao gồm một hoặc nhiều lipid thu được từ thực vật.

32. Hạt giống như virus theo điểm 30, trong đó một hoặc nhiều protein virus của VLP này bao gồm các N-glycan đặc hiệu với thực vật, hoặc các N-glycan được biến đổi.

33. Chế phẩm bao gồm liều hữu hiệu của VLP theo điểm 30 để gây ra đáp ứng miễn dịch, và chất mang dược dụng.

34. Chế phẩm chứa VLP theo điểm 30 để gây miễn dịch đối với sự lây nhiễm virus cúm ở đối tượng.

35. Vật liệu thực vật bao gồm VLP được tạo ra bằng phương pháp theo điểm 1 hoặc 28.

36. Chế phẩm chứa vật liệu thực vật theo điểm 35 để sử dụng trong việc gây miễn dịch đối với sự lây nhiễm virus cúm ở đối tượng.
37. Chất bổ sung cho thực phẩm bao gồm vật liệu thực vật theo điểm 35.
38. Phương pháp theo điểm 1, phương pháp này còn bao gồm bước đưa vào axit nucleic thứ ba, axit nucleic thứ ba này chứa trình tự mã hóa gen ức chế bất hoạt.
39. Phương pháp theo điểm 14, phương pháp này còn bao gồm bước đưa vào axit nucleic thứ tư, axit nucleic thứ tư này chứa trình tự mã hóa gen ức chế bất hoạt.
40. Phương pháp theo điểm 38 hoặc 39, trong đó gen ức chế bất hoạt được chọn từ nhóm gồm có HcPro và p19.
41. Phương pháp theo điểm 1 hoặc 28, trong đó VLP không chứa protein kênh proton.
42. Phương pháp theo điểm 1 hoặc 28, trong đó protein hemagglutinin cúm (HA) là protein HA0.
43. Phương pháp theo điểm 2 hoặc 29, trong đó protein M2 được chọn từ nhóm bao gồm cúm A/Puerto Rico/8/1934 và cúm A/New Caledonia/20/1999.
44. Polypeptit bao gồm trình tự axit amin là trình tự được chọn từ nhóm bao gồm SEQ ID NO: 41 và SEQ ID NO: 58.
45. Axit nucleic chứa trình tự mã hóa polypeptit theo điểm 44.
46. Axit nucleic theo điểm 45, trong đó axit nucleic này bao gồm trình tự nucleotit của trình tự được chọn từ nhóm bao gồm SEQ ID NO: 43 và phần được gạch dưới của SEQ ID NO: 57.

DANH MỤC TRÌNH TỰ

<110> Medicago Inc.
 <120> Hạt giống như virus và phương pháp sản xuất hạt giống như virus này
 <130> V83991WO
 <150> US 61/541,780
 <151> 2011-09-30
 <160> 64
 <170> PatentIn version 3.5
 <210> 1
 <211> 5
 <212> PRT
 <213> Trình tự nhân tạo
 <220>
 <223> trình tự tín hiệu kênh proton
 <220>
 <221> misc_feature
 <222> (2)..(4)
 <223> Xaa có thể là axit amin xuất hiện tự nhiên bất kỳ
 <400> 1
 His Xaa Xaa Xaa Trp
 1 5
 <210> 2
 <211> 41
 <212> ADN
 <213> Trình tự nhân tạo
 <220>
 <223> đoạn mỗi IF-H5A-I-05.s1+3c
 <400> 2
 aaatttgtcg ggcccatgga gaaaatagtg cttcttcttg c 41
 <210> 3
 <211> 50
 <212> ADN
 <213> Trình tự nhân tạo
 <220>
 <223> đoạn mỗi IF-H5dTm.r
 <400> 3
 actaaagaaa ataggccttt aaatgcaaat tctgcattgt aacgatccat 50
 <210> 4
 <211> 4903

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Cấu trúc 1191

<400> 4

tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg	60
gacgttttta atgtactgaa ttaacgccga atccccggct ggtatatatta tatgttgtca	120
aataactcaa aaaccataaa agtttaagtt agcaagtgtg tacattttta cttgaacaaa	180
aatattcacc tactactggt ataatcatt attaaacatt agagtaaaga aatatggatg	240
ataagaacaa gagtagtgat attttgacaa caattttgtt gcaacatttg agaaaatttt	300
gttgttctct cttttcattg gtcaaaaaca atagagagag aaaaaggaag agggagaata	360
aaaacataat gtgagtatga gagagaaagt tgtacaaaag ttgtaccaa atagttgtac	420
aaatatcatt gaggaatttg acaaaagcta cacaaataag ggtaattgc tgtaataaaa	480
taaggatgac gcattagaga gatgtaccat tagagaattt ttggcaagtc attaaaaaga	540
aagaataaat tttttttaa attaaaagtt gagtcatttg attaaacatg tgattattta	600
atgaattgat gaaagagttg gattaaagtt gtattagtaa ttagaatttg gtgtcaaatt	660
taatttgaca tttgatcttt tcctatatat tgccccatag agtcagttaa ctcatTTTTA	720
tatttcatag atcaaataag agaaataacg gtatatatta ccctccaaaa aaaaaaacg	780
gtatatTTac taaaaaatct aagccacgta ggaggataac aggatccccg taggaggata	840
acatccaatc caaccaatca caacaatcct gatgagataa cccactTTaa gccacgcac	900
ctgtggcaca tctacattat ctaaatacaca cattcttcca cacatctgag ccacacaaaa	960
accaatccac atctttatca ccatttctat aaaaaatcac actttgtgag tctacacttt	1020
gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag	1080
aaaatggaac gagctataca aggaaacgac gctagggaac aagctaacag tgaacgttgg	1140
gatggaggat caggaggtac cacttctccc ttcaaacttc ctgacgaaag tccgagttgg	1200
actgagtggc ggctacataa cgatgagacg aattcgaatc aagataatcc ctttggtttc	1260
aaggaaagct ggggtttcgg gaaagttgta ttaagagat atctcagata cgacaggacg	1320
gaagcttcac tgcacagagt ccttgatct tggacgggag attcggttaa ctatgcagca	1380
tctcgatttt tcggtttcga ccagatcgga tgtacctata gtattcggtt tcgaggagtt	1440
agtatcaccg tttctggagg gtgcggaact cttcagcatc tctgtgagat ggcaattcgg	1500
tctaagcaag aactgctaca gcttgcccca atcgaagtgg aaagtaatgt atcaagagga	1560
tgccctgaag gtactcaaac cttcgaaaaa gaaagcgagt aagttaaaat gcttcttcgt	1620

ctcctattta taatatgggtt tgttattggtt aattttgttc ttgtagaaga gcttaattaa	1680
tcgttggtgt tatgaaatac tatttgtatg agatgaactg gtgtaatgta attcatttac	1740
ataagtggag tcagaatcag aatgtttcct ccataactaa ctagacatga agacctgccg	1800
cgtacaattg tcttatattt gaacaactaa aattgaacat cttttgccac aactttataa	1860
gtggttaata tagctcaaat atatggtcaa gttcaataga ttaataatgg aaatatcagt	1920
tatcgaaatt cattaacaat caacttaacg ttattaacta ctaattttat atcatcccct	1980
ttgataaatg atagtacacc aattaggaag gagcatgctc gcctaggaga ttgtcgtttc	2040
ccgccttcag tttgcaagct gctctagccg tgtagccaat acgcaaaccg cctctccccg	2100
cgcgttggga attactagcg cgtgtcgaca agcttgcatg ccggtcaaca tgggtggagca	2160
cgacacactt gtctactcca aaaatatcaa agatacagtc tcagaagacc aaagggcaat	2220
tgagactttt caacaaaggg taatatccgg aaacctcctc ggattccatt gccagctat	2280
ctgtcacttt attgtgaaga tagtggaaaa ggaagggtgc tctacaaat gccatcattg	2340
cgataaagga aaggccatcg ttgaagatgc ctctgccgac agtgggtcca aagatggacc	2400
cccacccacg aggagcatcg tggaaaaaga agacgttcca accacgtctt caaagcaagt	2460
ggattgatgt gataacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga	2520
tacagtctca gaagaccaa gggcaattga gacttttcaa caaagggtaa tatccggaaa	2580
cctctcgga ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga	2640
aggtggctcc tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc	2700
tgccgacagt ggtcccaaag atggaccccc acccacgagg agcatcgttg aaaaagaaga	2760
cgttccaacc acgtcttcaa agcaagtgga ttgatgtgat atctccactg acgtaaggga	2820
tgacgcacaa tcccactatc cttcgcaaga cccttcctct atataaggaa gttcatttca	2880
tttgagaggg tattaatac ttaataggtt ttgataaaag cgaacgtggg gaaacccgaa	2940
ccaaaccttc ttctaaactc tctctcatct ctcttaaagc aaacttctct cttgtctttc	3000
ttgcgtgagc gatcttcaac gttgtcagat cgtgcttcgg caccagtaca acgttttctt	3060
tcactgaagc gaaatcaaag atctctttgt ggacacgtag tgcggcgcca ttaaataacg	3120
tgtacttgtc ctattcttgt cgggtgtggtc ttgggaaaag aaagcttgct ggaggctgct	3180
gttcagcccc atacattact tgttacgatt ctgctgactt tcggcggtg caatatctct	3240
acttctgctt gacgaggtat tgttgctgt acttctttct tcttcttctt gctgattggt	3300
tctataagaa atctagtatt ttctttgaaa cagagttttc ccgtggtttt cgaacttgga	3360
gaaagattgt taagcttctg tatattctgc ccaaatttgt cgggcccgcg gatggcgaaa	3420
aacgttgcca ttttcggctt attgttttct cttcttgtgt tggttccttc tcagatcttc	3480

gcctgcaggc tcctcagcca aaacgacacc cccatctgtc tatccactgg ccctggatc 3540
 tgctgccccaa actaactoca tggtagacct gggatgcctg gtcaagggct atttcctga 3600
 gccagtgaca gtgacctgga actctggatc cctgtccagc ggtgtgcaca ccttcccagc 3660
 tgtcctgcag tctgacctct acactctgag cagctcagtg actgtcccct ccagcacctg 3720
 gcccagcgag accgtcacct gcaacgttgc ccacccggcc agcagcacca aggtggacaa 3780
 gaaaattgtg cccagggatt gtggttgtaa gccttgcata tgtacagtcc cagaagtatc 3840
 atctgtcttc atcttcccc caaagcccaa ggatgtgtc accattactc tgactcctaa 3900
 ggtcacgtgt gttgtggtag acatcagcaa ggatgatccc gaggtccagt tcagctggtt 3960
 tgtagatgat gtggaggtgc acacagctca gacgcaaccc cgggaggagc agttcaacag 4020
 cactttccgc tcagtcaagt aacttcccat catgcaccag gactggctca atggcaagga 4080
 gogatcgctc accatcacca tcaccatcac catcaccatt aaaggcctat tttctttagt 4140
 ttgaatttac tgttattcgg tgtgcatttc tatgtttggt gagcggttt ctgtgctcag 4200
 agtgtgttta ttttatgtaa ttttaattct ttgtgagctc ctgttttagca ggtcgtccct 4260
 tcagcaagga cacaaaaaga ttttaatttt attaaaaaaa aaaaaaaaaa agaccgggaa 4320
 ttogatatca agcttatcga cctgcagatc gttcaaacat ttggcaataa agtttcttaa 4380
 gattgaatcc tgttgccggc cttgcgatga ttatcatata atttctgttg aattacgtta 4440
 agcatgtaat aattaacatg taatgcatga cgttatttat gagatgggtt tttatgatta 4500
 gagtcccgc attatacatt taatacgca tagaaaacaa aatatagcg gcaaactagg 4560
 ataaattatc gcgcgcggtg tcatctatgt tactagatct ctagagtctc aagcttggcg 4620
 cgcccacgtg actagtggca ctggccgtcg ttttacaacg tcgtgactgg gaaaaccctg 4680
 gcgttaccca acttaatcgc cttgcagcac atccccctt cgccagctgg cgtaatagcg 4740
 aagaggcccg caccgatcgc ctttcccaac agttgcgcag cctgaatggc gaatgctaga 4800
 gcagcttgag cttggatcag attgtcgttt cccgccttca gtttaaacta tcagtgtttg 4860
 acaggatata ttggcgggta aacctaagag aaaagagcgt tta 4903

<210> 5
 <211> 3447
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Cát xét biểu hiện số 489

<400> 5
 gtcaacatgg tggagcacga cacacttgct tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga	120
ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc	180
tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt	240
ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc	300
acgtcttcaa agcaagtgga ttgatgtgat aacatgggtg agcacgacac acttgtctac	360
tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa	420
agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg	480
aagatagtgg aaaaggaagg tggtcctac aaatgccatc attgcgataa aggaaaggcc	540
atcgttgaag atgcctctgc cgacagtggc cccaaagatg gacccccacc cacgaggagc	600
atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc	660
tccaactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata	720
taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga	780
acgtggggaa acccgaacca aacottcttc taaactctct ctcactctctc ttaaagcaaa	840
cttctctctt gtctttcttg cgtgagcgat cttcaacggt gtcagatcgt gcttcggcac	900
cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc	960
ggcgccatta aataacgtgt acttgtccta ttcttgtcgg tgtggctctg ggaaaagaaa	1020
gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg	1080
gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgctgtact tctttcttct	1140
tcttcttget gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg	1200
tggttttcga acttgagaaa agattgttaa gcttctgtat attctgcca aatttgcgg	1260
gcccatggag aaaatagtgc ttcttcttgc aatagtcagt cttgttaaaa gtgatcagat	1320
ttgcattggt taccatgcaa acaattcaac agagcagggt gacacaatca tggaaaagaa	1380
cgttactgtt acacatgccc aagacatact ggaaaagaca cacaacggga agctctgcga	1440
tctagatgga gtgaagcctc taattttaag agattgtagt gtagctggat ggctcctcgg	1500
gaacccaatg tgtgacgaat tcatcaatgt accggaatgg tottacatag tggagaaggc	1560
caatccaacc aatgacctct gttaccacg gagtttcaac gactatgaag aactgaaaca	1620
cctattgagc agaataaacc attttgagaa aattcaaatc atccccaaaa gttcttggtc	1680
cgatcatgaa gcctcatcag gagttagctc agcatgtcca tacctgggaa gtccctcctt	1740
ttttagaaat gtggtatggc ttatcaaaaa gaacagtaca taccacaaca taaagaaaag	1800
ctacaataat accaaccaag aggatctttt ggtactgtgg ggaattcacc atcctaata	1860
tgcggcagag cagacaaggc tatatcaaaa cccaaccacc tatatttcca ttgggacatc	1920

```

aacactaaac cagagattgg taccaaaaat agctactaga tccaaagtaa acgggcaaaag 1980
tggaaggatg gagttcttct ggacaatttt aaacctaata gatgcaatca acttcgagag 2040
taatggaaat ttcattgctc cagaatatgc atacaaaatt gtcaagaaag gggactcagc 2100
aattatgaaa agtgaattgg aatatggtaa ctgcaacacc aagtgtcaaa ctccaatggg 2160
ggcgataaac tctagtatgc cattccacaa catacaccct ctcaccatcg gggaatgcc 2220
caaatatgtg aaatcaaaca gattagtcct tgcaacaggg ctccagaaata gccctcaaag 2280
agagagcaga agaaaaaaga gaggactatt tggagctata gcaggtttta tagaggagg 2340
atggcagga atggtagatg gttggtatgg gtaccacat agcaatgagc aggggagtg 2400
gtacgctgca gacaaagaat ccactcaaaa ggcaatagat ggagtcacca ataaggtcaa 2460
ctcaatcatt gacaaaatga aactcagtt tgaggccgtt ggaagggaat ttaataactt 2520
agaaaggaga atagagaatt taaacaagaa gatggaagac gggtttctag atgtctggac 2580
ttataatgcc gaacttctgg ttctcatgga aatgagaga actctagact ttcatgactc 2640
aaatgttaag aacctctacg acaagggtccg actacagctt agggataatg caaaggagct 2700
gggtaacggt tgtttcgagt tctatcacia atgtgataat gaatgtatgg aaagtataag 2760
aaacggaacg tacaactatc cgcagtattc agaagaagca agattaaaaa gagaggaaat 2820
aagtggggta aaattggaat caataggaac ttaccaaata ctgtcaattt attcaacagt 2880
ggcgagttcc ctgacactgg caatcatgat ggctgggtcta tctttatgga tgtgctccaa 2940
tggatcggtta caatgcagaa ttgcattda aaggcctatt ttctttagtt tgaatttact 3000
gttattcggt gtgcatttct atgtttggtg agcgggtttc tgtgctcaga gtgtgtttat 3060
tttatgtaat ttaatttctt tgtgagctcc tgtttagcag gtcgtccctt cagcaaggac 3120
acaaaaagat ttaattttta ttaaaaaaaaa aaaaaaaaaa gaccgggaat tcgatatcaa 3180
gcttatcgac ctgcagatcg ttcaaacatt tggcaataaa gtttcttaag attgaatcct 3240
gttgccggtc ttgcgatgat tatcatataa tttctgttga attacgttaa gcatgtaata 3300
attaacatgt aatgcgatgac gttatttatg agatgggttt ttatgattag agtcccgcaa 3360
ttatacattt aatacgcgat agaaaacaaa atatagcgcg caaactagga taaattatcg 3420
cgcgcggtgt catctatggt actagat 3447

```

<210> 6
 <211> 568
 <212> PRT
 <213> Virut cúm

<400> 6

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

245	250	255
Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile 260 265 270		
Val Lys Lys Gly Asp Ser Ala Ile Met Lys Ser Glu Leu Glu Tyr Gly 275 280 285		
Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser 290 295 300		
Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys 305 310 315 320		
Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser 325 330 335		
Pro Gln Arg Glu Ser Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile 340 345 350		
Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr 355 360 365		
Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys 370 375 380		
Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser 385 390 395 400		
Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe 405 410 415		
Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp 420 425 430		
Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met 435 440 445		
Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu 450 455 460		
Tyr Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly 465 470 475 480		
Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu 485 490 495		

Ser Ile Arg Asn Gly Thr Tyr Asn Tyr Pro Gln Tyr Ser Glu Glu Ala
500 505 510

Arg Leu Lys Arg Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly
515 520 525

Thr Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala
530 535 540

Leu Ala Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly
545 550 555 560

Ser Leu Gln Cys Arg Ile Cys Ile
565

<210> 7
<211> 42
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> đoạn mồi IF-S1-M1+M2ANC.c

<400> 7
aaatttgtcg ggcccatgag tcttctaacc gaggtcgaaa cg

42

<210> 8
<211> 43
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> đoạn mồi IF-S1-4-M2ANC.r

<400> 8
actaaagaaa ataggccttt actccagctc tatgctgaca aaa

43

<210> 9
<211> 294
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> trình tự nucleotit của gen M2 tổng hợp (DQ508860)

<400> 9
atgagtcttc taaccgaggt cgaaacgcct atcagaaacg aatgggggtg cagatgcaac

60

gattcaagtg atcctcttgt tggtgccgca agtataattg ggattgtgca cctgatattg

120

tggattattg atcgccctttt ttccaaaagc atttatcgta tctttaaaca cggtttaaaa

180

agagggcctt ctacggaagg agtaccagag tctatgaggg aagaatatcg agaggaacag 240
 cagaatgctg tggatgctga cgatggcat tttgtcagca tagagctgga gtaa 294

 <210> 10
 <211> 2034
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Cát xét biểu hiện số 1261

 <400> 10
 gtcaacatgg tggagcacga cacacttgct tactccaaaa atatcaaaga tacagtctca 60
 gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120
 ttccattgcc cagctatctg tcaactttatt gtgaagatag tggaaaagga aggtggctcc 180
 taaaaatgcc atcattgcga taaaggaaag gccatcggtg aagatgcctc tgccgacagt 240
 ggtcccaaag atggaccccc acccacgagg agcatcggtg aaaaagaaga cgttccaacc 300
 acgtcttcaa agcaagtgga ttgatgtgat aacatgggtg agcacgacac acttgtctac 360
 tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420
 agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480
 aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540
 atcgttgaag atgcctctgc cgacagtggc cccaaagatg gacccccacc cacgaggagc 600
 atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660
 tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720
 taaggaagtt catttcattt ggagaggtat taaaatotta ataggttttg ataaaagcga 780
 acgtggggaa acccgaacca aaccttcttc taaactctct ctcactctctc ttaaagcaaa 840
 cttctctctt gtctttcttg cgtgagcgat cttcaacggt gtcagatcgt gcttcggcac 900
 cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc 960
 ggcgccatta aataacgtgt acttgtccta ttcttgctgg tgtggctctg ggaaaagaaa 1020
 gcttgctgga ggctgctgtt cagccccata cattaactgt tacgattctg ctgactttcg 1080
 gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgctgtact tctttctctt 1140
 tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg 1200
 tggttttcga acttgagaaa agattgttaa gcttctgtat attctgcca aatttgctcg 1260
 gcccatgagt cttctaaccg aggtcgaaac gcctatcaga aacgaatggg ggtgcagatg 1320
 caacgattca agtgatcctc ttgttggtgc cgcaagtata attgggattg tgcacctgat 1380
 attgtggatt attgatcgcc ttttttccaa aagcatttat cgtatcttta aacacggttt 1440

aaaaagaggg ccttctacgg aaggagtacc agagtctatg agggaagaat atcgagagga 1500
 acagcagaat gctgtggatg ctgacgatgg tcattttgtc agcatagagc tggagtaaag 1560
 gcctattttc tttagtttga atttactgtt attcgggtgtg catttctatg tttgggtgagc 1620
 ggttttctgt gctcagagtg tgtttatttt atgtaattta atttctttgt gagctcctgt 1680
 ttagcaggtc gtcccttcag caaggacaca aaaagatttt aattttatta aaaaaaaaaa 1740
 aaaaaaagac cgggaattcg atatcaagct tatcgacctg cagatcgttc aaacatttgg 1800
 caataaagtt tcttaagatt gaatcctgtt gccggtcttg cgatgattat catataattt 1860
 ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 1920
 tgggttttta tgattagagt cccgcaatta tacatttaat acgcgataga aaacaaaata 1980
 tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact agat 2034

<210> 11
 <211> 97
 <212> PRT
 <213> Virut cúm

<400> 11

Met Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly
 1 5 10 15

Cys Arg Cys Asn Asp Ser Ser Asp Pro Leu Val Val Ala Ala Ser Ile
 20 25 30

Ile Gly Ile Val His Leu Ile Leu Trp Ile Ile Asp Arg Leu Phe Ser
 35 40 45

Lys Ser Ile Tyr Arg Ile Phe Lys His Gly Leu Lys Arg Gly Pro Ser
 50 55 60

Thr Glu Gly Val Pro Glu Ser Met Arg Glu Glu Tyr Arg Glu Glu Gln
 65 70 75 80

Gln Asn Ala Val Asp Ala Asp Asp Gly His Phe Val Ser Ile Glu Leu
 85 90 95

Glu

<210> 12
 <211> 294
 <212> ADN
 <213> Trình tự nhân tạo

<220>

<223> trình tự nucleotit của gen M2 tổng hợp

<400> 12

atgagtcttc taaccgaggt cgaaacgcct atcagaaacg aatgggggtg cagatgcaac	60
ggttcaagt atcctctcac tattgccgca aatatcattg ggatcttgca cttgacattg	120
tggattcttg atcgtctttt ttcaaagtc atttaccgtc gctttaaata cggactgaaa	180
ggagggcctt ctacggaagg agtgccaaag tctatgaggg aagaatatcg aaaggaacag	240
cagagtgcctg tggatgctga cgatgggtcat tttgtcagca tagagctgga gtaa	294

<210> 13

<211> 2034

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Cát xét biểu hiện số 859

<400> 13

gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca	60
gaagacccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga	120
ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc	180
tacaaatgcc atcattgcca taaaggaaa ggcacgttg aagatgcctc tgccgacagt	240
ggccccaaa atggaccccc acccacgagg agcatcgtgg aaaaagaaga cggtccaacc	300
acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac	360
tccaaaaata tcaaagatac agtctcagaa gaccaaagg caattgagac ttttcaacaa	420
agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg	480
aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc	540
atcgttgaag atgcctctgc cgacagtggc cccaaagatg gacccccacc cacgaggagc	600
atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc	660
tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagacct ttcctctata	720
taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga	780
acgtggggaa acccgaacca aaccttcttc taaactctct ctcactctctc ttaaagcaaa	840
cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac	900
cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc	960
ggcgccatta aataacgtgt acttgtccta ttcttgctcg tgtgggtcttg ggaaaagaaa	1020
gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg	1080

```

gcggggtgcaa tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct 1140
tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg 1200
tggttttcga acttgagaa agattgttaa gcttctgtat attctgcca aatttgctgg 1260
gcccattgagt cttctaaccg aggtcgaaac gcctatcaga aacgaatggg ggtgcagatg 1320
caacggttca agtgatcctc tcactattgc cgcaaatac attgggatct tgcacttgac 1380
attgtggatt ctgatcgtc tttttttcaa atgcatttac cgtcgcttta aatacggact 1440
gaaaggaggg ccttctacgg aaggagtgcc aaagtctatg aggaagaat atcgaaagga 1500
acagcagagt gctgtggatg ctgacgatgg tcattttgtc agcatagagc tggagtaaag 1560
gcctattttc tttagtttga atttactggt attcgggtgtg catttctatg tttgggtgagc 1620
ggttttctgt gctcagagtg tgtttatttt atgtaattta atttctttgt gagctcctgt 1680
ttagcaggtc gtcccttcag caaggacaca aaaagatttt aattttatta aaaaaaaaaa 1740
aaaaaaaaagac cggaattcg atatcaagct tatcgacctg cagatcgttc aaacatttgg 1800
caataaagtt tcttaagatt gaatcctggt gccggtcttg cgatgattat catataattt 1860
ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 1920
tggggttttta tgattagagt cccgcaatta tacatttaac acgcgataga aaacaaaata 1980
tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact agat 2034

```

<210> 14
 <211> 97
 <212> PRT
 <213> Virut cúm

<400> 14

```

Met Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly
1          5          10          15

```

```

Cys Arg Cys Asn Gly Ser Ser Asp Pro Leu Thr Ile Ala Ala Asn Ile
          20          25          30

```

```

Ile Gly Ile Leu His Leu Thr Leu Trp Ile Leu Asp Arg Leu Phe Phe
35          40          45

```

```

Lys Cys Ile Tyr Arg Arg Phe Lys Tyr Gly Leu Lys Gly Gly Pro Ser
50          55          60

```

```

Thr Glu Gly Val Pro Lys Ser Met Arg Glu Glu Tyr Arg Lys Glu Gln
65          70          75          80

```

```

Gln Ser Ala Val Asp Ala Asp Asp Gly His Phe Val Ser Ile Glu Leu

```

85

90

95

Glu

<210> 15
 <211> 46
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> đoạn mỗi IF-H1A-C-09.s2+4c

<400> 15
 tctcagatct tcgccgacac attatgtata gggtatcatg cgaaca 46

<210> 16
 <211> 49
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> đoạn mỗi IF-H1A-C-09.s1-4r

<400> 16
 actaaagaaa ataggccttt aaatacatat tctacactgt agagacca 49

<210> 17
 <211> 1701
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> trình tự nucleotit của gen H1 tổng hợp

<400> 17
 atgaaggcaa tactagtagt tctgctatat acatttgcaa ccgcaaattgc agacacatta 60
 tgtataggtt atcatgcgaa caattcaaca gacactgtag acacagtact agaaaagaat 120
 gtaacagtaa cacactctgt taaccttcta gaagacaagc ataacgggaa actatgcaaa 180
 ctaagagggg tagccccatt gcatttgggt aaatgtaaca ttgctggctg gatcctggga 240
 aatccagagt gtgaatcact ctccacagca agctcatggt cctacattgt ggaaacacct 300
 agttcagaca atggaacgtg ttaccagga gatttcatcg attatgagga gctaagagag 360
 caattgagct cagtgtcatc atttgaaagg tttgagatat tccccaagac aagttcatgg 420
 cccaatcatg actcgaacaa aggtgtaacg gcagcatgtc ctcatgctgg agcaaaaagc 480
 ttctacaaaa atttaatatg gctagttaaa aaaggaaatt catacccaaa gctcagcaaa 540
 tcctacatta atgataaagg gaaagaagtc ctctgtgctat ggggcattca ccatccatct 600
 actagtgtctg accaacaag tctctatcag aatgcagatg catatgtttt tgtgggggtca 660

tcaagataca gcaagaagtt caagccggaa atagcaataa gacccaaagt gagggatcaa 720
 gaagggagaa tgaactatta ctggacacta gtagagccgg gagacaaaat aacattcgaa 780
 gcaactggaa atctagtggg accgagatat gcattcgcaa tggaaagaaa tgctggatct 840
 ggtattatca tttcagatac accagtccac gattgcaata caacttgtca aacacccaag 900
 ggtgctataa acaccagcct cccatttcag aatatacatc cgatcacaaat tggaaaatgt 960
 ccaaaatatg taaaaagcac aaaattgaga ctggccacag gattgaggaa tatcccgctct 1020
 attcaatcta gaggcctatt tggggccatt gccggtttca ttgaaggggg gtggacaggg 1080
 atggtagatg gatggtacgg ttatcaccat caaaatgagc aggggtcagg atatgcagcc 1140
 gacctgaaga gcacacagaa tgccattgac gagattacta acaaagtaaa ttctgttatt 1200
 gaaaagatga atacacagtt cacagcagta ggtaaagagt tcaaccacct ggaaaaaaga 1260
 atagagaatt taaataaaaa agttgatgat ggtttctctg acatttgagc ttacaatgcc 1320
 gaactgttgg ttctattgga aaatgaaaga actttggact accacgattc aaatgtgaag 1380
 aacttatatg aaaaggtaag aagccagcta aaaaacaatg ccaaggaaat tggaaacggc 1440
 tgctttgaat ttaccacaa atgcgataac acgtgcatgg aaagtgtcaa aaatgggact 1500
 tatgactacc caaaatactc agaggaagca aaattaaaca gagaagaaat agatggggta 1560
 aagctggaat caacaaggat ttaccagatt ttggcgatct attcaactgt cgccagttca 1620
 ttggtactgg tagtctccct gggggcaatc agtttctgga tgtgctctaa tgggtctcta 1680
 cagtgtagaa tatgtattta a 1701

<210> 18
 <211> 4897
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc 1192

<400> 18
 tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 60
 gacgttttta atgtactgaa ttaacgccga atccccgggt ggtatattta tatgttgtca 120
 aataactcaa aaaccataaa agtttaagtt agcaagtgtg tacattttta cttgaacaaa 180
 aatattcacc tactactggt ataaatcatt attaaacatt agagtaaaga aatatggatg 240
 ataagaacaa gagtagtgat attttgacaa caattttgtt gcaacatttg agaaaatttt 300
 gttgttctct cttttcattg gtcaaaaaca atagagagag aaaaaggaag agggagaata 360
 aaaacataat gtgagtatga gagagaaagt tgtacaaaag ttgtaccaaa atagttgtac 420

aaatatcatt gaggaatttg acaaaagcta cacaaataag ggttaattgc tgtaaataaa	480
taaggatgac gcattagaga gatgtacat tagagaatth ttggcaagtc attaaaaaga	540
aagaataaat tattttttaa attaaaagtt gagtcatttg attaaacatg tgattattta	600
atgaattgat gaaagagttg gattaaagtt gtattagtaa ttagaatttg gtgtcaaatt	660
taatttgaca ttgatcttt tctatatat tgcccatag agtcagttaa ctcatthttt	720
tatttcatag atcaaataag agaaataacg gtatattaat ccctccaaaa aaaaaaacg	780
gtatatttac taaaaaatct aagccacgta ggaggataac aggatccccg taggaggata	840
acatccaatc caaccaatca caacaatcct gatgagataa cccactthta gccacgcat	900
ctgtggcaca tctacattat ctaaatacaca cattcttcca cacatctgag ccacacaaaa	960
accaatccac atctthtatca cccattctat aaaaaatcac actthgtgag tctacactth	1020
gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag	1080
aaaatggaac gagctataca aggaaacgac gctagggaac aagctaacag tgaacgttg	1140
gatggaggat caggaggtac cacttctccc ttcaaacttc ctgacgaaag tccgagttg	1200
actgagtggc ggctacataa cgatgagacg aattcgaatc aagataatcc ccttggtthc	1260
aaggaaagct ggggtthcgg gaaagttgta tttaagagat atctcagata cgacaggacg	1320
gaagcttcac tgcacagagt ccttgatct tggacgggag attcggthta ctatgcagca	1380
tctcgattth tcggtthcga ccagatcgga tgtacctata gtattcgtt tcgaggagt	1440
agtatcaccg thtctggagg gtcgcgaact cttcagcatc tctgtgagat ggcaattcgg	1500
tctaagcaag aactgctaca gcttgcccca atcgaagtgg aaagtaatgt atcaagagga	1560
tgccctgaag gtactcaaac cttcgaaaaa gaaagcgagt aagthaaaa gctthctcgt	1620
ctctattta taatatggtt tgttattgtt aattthgttc ttgtagaaga gcttaattaa	1680
tcgttgthgt tatgaaatac tatttgatg agatgaactg gtgtaatgta attcatttac	1740
ataagtggag tcagaatcag aatgthtctt ccataactaa ctagacatga agacctgccg	1800
cgtacaattg tcttatattt gaacaactaa aattgaacat cthttgccac aactthataa	1860
gtggttaata tagctcaaata atatggtcaa gttcaataga ttaataatgg aaatatcagt	1920
tatcgaaatt cattaacaat caacttaacg ttattaacta ctaattthtat atcatccct	1980
ttgataaatg atagtacacc aattaggaag gagcatgctc gcctaggaga ttgtcgtthc	2040
ccgccttcag thtgcaagct gctctagccg tgtagccaat acgcaaaccg cctctccccg	2100
cgcgthtggga attactagcg cgtgtcgaca agcttgcatg ccggtcaaca tggthggagca	2160
cgacacactt gtctactcca aaaatatcaa agatacagtc tcagaagacc aaagggcaat	2220
tgagacttht caacaaaggg taatatccgg aaacctctc ggattccatt gccagctat	2280

ctgtcacttt atttgtgaaga tagtggaataa ggaaggtggc tcctacaaat gccatcattg	2340
cgataaagga aaggccatcg ttgaagatgc ctctgccgac agtgggtcca aagatggacc	2400
cccacccacg aggagcatcg tggaaaaaga agacgttcca accacgtctt caaagcaagt	2460
ggattgatgt gataacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga	2520
tacagtctca gaagaccaa gggcaattga gacttttcaa caaagggtaa tatccggaaa	2580
cctcctcgga ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga	2640
aggtggctcc tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc	2700
tgccgacagt ggtcccaaag atggaccccc acccacgagg agcatcgttg aaaaagaaga	2760
cgttccaacc acgtcttcaa agcaagtgga ttgatgtgat atctccactg acgtaaggga	2820
tgacgcacaa tcccactatc cttcgcaaga cccttctct atataaggaa gttcatttca	2880
tttgagaggg tattaatac ttaataggtt ttgataaaag cgaacgtggg gaaacccgaa	2940
ccaaaccttc ttctaaactc tctctcatct ctcttaaagc aaacttctct cttgtcttct	3000
ttgcgtgagc gatcttcaac gttgtcagat cgtgcttcgg caccagtaca acgttttctt	3060
tcactgaagc gaaatcaaag atctctttgt ggacacgtag tgcggcgcca ttaaataacg	3120
tgtacttgtc ctattcttgt cgggtgtggtc ttgggaaaag aaagcttgct ggaggtgct	3180
gttcagcccc atacattact tgttacgatt ctgctgactt tcggcggtg caatatctct	3240
acttctgctt gacgaggtat tggtgcctgt acttcttctt tcttcttctt gctgattggt	3300
tctataagaa atctagtatt ttctttgaaa cagagttttc ccgtggtttt cgaacttgga	3360
gaaagattgt taagcttctg tatattctgc ccaaatttgt cgggcccatt gcgaaaaacg	3420
ttgcgatttt cggcttattg ttttctcttc ttgtgttggt tccttctcag atcttcgccg	3480
cggctcctca gccaaaacga ccccccatc tgtctatcca ctggcccctg gatctgctgc	3540
ccaaactaac tccatggtga ccctgggatg cctggtaag ggctatttcc ctgagccagt	3600
gacagtgacc tggaaactct gatccctgtc cagcgggtgt cacaccttcc cagctgtcct	3660
gcagtctgac ctctacactc tgagcagctc agtgactgtc ccctcagca cctggcccag	3720
cgagaccgtc acctgcaacg ttgcccaccc ggccagcagc accaaggtgg acaagaaaat	3780
tgtgcccagg gattgtggtt gtaagccttg catatgtaca gtcccagaag tatcatctgt	3840
cttcattctt ccccaaagc ccaaggatgt gctcaccatt actctgactc ctaaggtcac	3900
gtgtgttggt gtagacatca gcaaggatga tcccagagtc cagttcagct ggtttgtaga	3960
tgatgtggag gtgcacacag ctgagacgca accccgggag gagcagttca acagcacttt	4020
ccgctcagtc agtgaacttc ccatcatgca ccaggactgg ctcaatggca aggagcgatc	4080

gctcaccatc accatcacca tcaccatcac cattaaaggc ctattttctt tagtttgaat 4140
 ttactgttat tcggtgtgca tttctatggt tggtagagcg ttttctgtgc tcagagtgtg 4200
 tttattttat gtaatttaat ttctttgtga gctcctgttt agcaggtcgt cccttcagca 4260
 aggacacaaa aagattttta ttttattaaa aaaaaaaaaa aaaaagaccg ggaattcgat 4320
 atcaagctta tcgacctgca gatcgttcaa acatttggca ataaagtttc ttaagattga 4380
 atcctgttgc cggctcttgc atgattatca tataatttct gttgaattac gttaagcatg 4440
 taataattaa catgtaatgc atgacgttat ttatgagatg ggtttttatg attagagtcc 4500
 cgcaattata catttaatac gcgatagaaa aaaaaatata gcgcgcaaac taggataaat 4560
 tatcgcgcg cgtgtcatct atgttactag atctctagag tctcaagctt ggcgcgccca 4620
 cgtgactagt ggcaactggc gtcgttttac aacgtcgtga ctgggaaaac cctggcggtta 4680
 cccaacttaa tcgccttgca gcacatcccc ctttcgccag ctggcgtaat agcgaagagg 4740
 cccgcaccga tcgcccttcc caacagttgc gcagcctgaa tggcgaatgc tagagcagct 4800
 tgagcttgga tcagattgtc gtttcccgcc ttcagtttaa actatcagtg tttgacagga 4860
 tatattggcg ggtaaaccta agagaaaaga gcgttta 4897

<210> 19
 <211> 3462
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Cát xét biểu hiện số 484

<400> 19
 gtcaacatgg tggagcacga cacacttgct tactccaaaa atatcaaaga tacagtctca 60
 gaagacaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120
 ttccattgcc cagctatctg tcaactttatt gtgaagatag tggaaaagga aggtggctcc 180
 tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240
 ggtcccaaag atggaccccc acccacgagg agcatcgttg aaaaagaaga cgttccaacc 300
 acgtcttcaa agcaagtgga ttgatgtgat aacatgggtg agcacgacac acttgtctac 360
 tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaaaa 420
 agggtaatat ccggaacct cctcggattc cattgccag ctatctgtca ctttattgtg 480
 aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540
 atcgttgaag atgcctctgc cgacagtggc cccaagatg gacccccacc cacgaggagc 600
 atcgtgaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660
 tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttctctata 720

taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga	780
acgtggggaa acccgaacca aaccttcttc taaactctct ctcactcttc ttaaagcaaa	840
cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac	900
cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc	960
ggcgccatta aataacgtgt acttgctcta ttcttgtcgg tgtggtcttg ggaaaagaaa	1020
gcttgctgga ggctgctgtt cagccccata cattaactgt tacgattctg ctgactttcg	1080
gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgctgtact tctttcttct	1140
tcttcttgct gattggttct ataagaaatc tagtattttc ttgaaacag agttttcccg	1200
tggttttcga acttgagaa agattgttaa gcttctgtat attctgcca aatttgtcgg	1260
gcccattggcg aaaaacgttg cgattttcgg cttattgttt tctcttcttg tgttggttc	1320
ttctcagatc ttgcgcgaca cattatgtat aggttatcat gcgaacaatt caacagacac	1380
tgtagacaca gtactagaaa agaataaac agtaacacac tctgttaacc ttctagaaga	1440
caagcataac gggaaactat gcaaactaag aggggtagcc ccattgcatt tgggtaaattg	1500
taacattgct ggctggatcc tgggaaatcc agagtgtgaa tcactctcca cagcaagctc	1560
atggtcctac attgtggaaa cacctagttc agacaatgga acgtgttacc caggagattt	1620
catcgattat gaggagctaa gagagcaatt gagctcagtg tcatcatttg aaaggtttga	1680
gatattcccc aagacaagtt catggcccaa tcatgactcg aacaaagggtg taacggcagc	1740
atgtcctcat gctggagcaa aaagcttcta caaaaattta atatggctag ttaaaaaagg	1800
aaattcatac ccaaagctca gcaaatccta cattaatgat aaagggaaaag aagtcctcgt	1860
gctatggggc attcaccatc catctactag tgctgaccaa caaagtctct atcagaatgc	1920
agatgcatat gtttttgtgg ggtcatcaag atacagcaag aagttcaagc cggaaatagc	1980
aataagaccc aaagtgaggg atcaagaagg gagaatgaac tattactgga cactagtaga	2040
gccgggagac aaaataacat tcgaagcaac tggaaatcta gtggtaccga gatatgcatt	2100
cgcaatggaa agaaatgctg gatctggtat tatcatttca gatacaccag tccacgattg	2160
caatacaact tgtcaaacac ccaaggggtgc tataaacacc agcctcccat ttcagaatat	2220
acatccgatc acaattggaa aatgtccaaa atatgtaaaa agcacaaaat tgagactggc	2280
cacaggattg aggaatatcc cgtctattca atctagaggc ctatttgggg ccattgccgg	2340
tttcattgaa ggggggtgga cagggatggt agatggatgg tacggttatc accatcaaaa	2400
tgagcagggg tcaggatatg cagccgacct gaagagcaca cagaatgcca ttgacgagat	2460
tactaacaaa gtaaattctg ttattgaaaa gatgaataca cagttcacag cagtaggtaa	2520

```

agagttcaac cacctggaaa aaagaataga gaattttaaataaaaaagttg atgatggttt 2580
cctggacatt tggacttaca atgccgaact gttggtttcta ttggaaaatg aaagaacttt 2640
ggactaccac gattcaaatg tgaagaactt atatgaaaag gtaagaagcc agctaaaaaa 2700
caatgccaaag gaaattggaa acggctgctt tgaattttac cacaaatgcg ataacacgtg 2760
catggaaagt gtcaaaaatg ggacttatga ctacccaaaa tactcagagg aagcaaaatt 2820
aaacagagaa gaaatagatg gggtaaagct ggaatcaaca aggatttacc agattttggc 2880
gatctattca actgtcgcca gttcattggg actggtagtc tccctggggg caatcagttt 2940
ctggatgtgc tctaattgggt ctctacagtg tagaatatgt atttaaaggc ctattttctt 3000
tagtttgaat ttactgttat tcggtgtgca tttctatggt tggtgagcgg ttttctgtgc 3060
tcagagtgtg tttattttat gtaatttaat ttctttgtga gtcctgttt agcagggtcgt 3120
cccttcagca aggacacaaa aagattttta ttttattaaa aaaaaaaaaa aaaaagaccg 3180
ggaattcgat atcaagctta tcgacctgca gatcgttcaa acatttggca ataaagtttc 3240
ttaagattga atcctgttgc cggctcttgcg atgattatca tataatttct gttgaattac 3300
gttaagcatg taataattaa catgtaatgc atgacgttat ttatgagatg ggtttttatg 3360
attagagtcc cgcaattata catttaatac gcgatagaaa acaaaatata gcgcgcaaac 3420
taggataaat tatcgcgcgcg ggtgtcatct atgttactag at 3462

```

<210> 20
 <211> 573
 <212> PRT
 <213> Virut cúm

<400> 20

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
 1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Thr Leu Cys Ile Gly Tyr His
 20 25 30

Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val
 35 40 45

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

Leu Cys Lys Leu Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn
 65 70 75 80

Ile Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Ser Leu Ser Thr

85	90	95
Ala Ser Ser Trp Ser Tyr Ile Val Glu Thr Pro Ser Ser Asp Asn Gly		
100	105	110
Thr Cys Tyr Pro Gly Asp Phe Ile Asp Tyr Glu Glu Leu Arg Glu Gln		
115	120	125
Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Thr		
130	135	140
Ser Ser Trp Pro Asn His Asp Ser Asn Lys Gly Val Thr Ala Ala Cys		
145	150	155
Pro His Ala Gly Ala Lys Ser Phe Tyr Lys Asn Leu Ile Trp Leu Val		
165	170	175
Lys Lys Gly Asn Ser Tyr Pro Lys Leu Ser Lys Ser Tyr Ile Asn Asp		
180	185	190
Lys Gly Lys Glu Val Leu Val Leu Trp Gly Ile His His Pro Ser Thr		
195	200	205
Ser Ala Asp Gln Gln Ser Leu Tyr Gln Asn Ala Asp Ala Tyr Val Phe		
210	215	220
Val Gly Ser Ser Arg Tyr Ser Lys Lys Phe Lys Pro Glu Ile Ala Ile		
225	230	235
Arg Pro Lys Val Arg Asp Gln Glu Gly Arg Met Asn Tyr Tyr Trp Thr		
245	250	255
Leu Val Glu Pro Gly Asp Lys Ile Thr Phe Glu Ala Thr Gly Asn Leu		
260	265	270
Val Val Pro Arg Tyr Ala Phe Ala Met Glu Arg Asn Ala Gly Ser Gly		
275	280	285
Ile Ile Ile Ser Asp Thr Pro Val His Asp Cys Asn Thr Thr Cys Gln		
290	295	300
Thr Pro Lys Gly Ala Ile Asn Thr Ser Leu Pro Phe Gln Asn Ile His		
305	310	315
Pro Ile Thr Ile Gly Lys Cys Pro Lys Tyr Val Lys Ser Thr Lys Leu		
325	330	335

Arg Leu Ala Thr Gly Leu Arg Asn Ile Pro Ser Ile Gln Ser Arg Gly
 340 345 350

Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met
 355 360 365

Val Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly
 370 375 380

Tyr Ala Ala Asp Leu Lys Ser Thr Gln Asn Ala Ile Asp Glu Ile Thr
 385 390 395 400

Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala
 405 410 415

Val Gly Lys Glu Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn
 420 425 430

Lys Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu
 435 440 445

Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser
 450 455 460

Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn
 465 470 475 480

Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp
 485 490 495

Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys
 500 505 510

Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys
 515 520 525

Leu Glu Ser Thr Arg Ile Tyr Gln Ile Leu Ala Ile Tyr Ser Thr Val
 530 535 540

Ala Ser Ser Leu Val Leu Val Val Ser Leu Gly Ala Ile Ser Phe Trp
 545 550 555 560

Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile
 565 570

<210> 21
 <211> 40
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> đoạn mỗi IF-S2+S4-H3 Per.c

<400> 21
 tctcagatct tcgccccaaa acttcctgga aatgacaaca 40

<210> 22
 <211> 45
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> đoạn mỗi IF-S1a4-H3 Per.r

<400> 22
 actaaagaaa ataggccttc aaatgcaa atgtgcaccta atgtt 45

<210> 23
 <211> 1701
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> trình tự nucleotit của gen H3 tổng hợp

<400> 23
 atgaagacta tcattgcttt gagctacatt ctatgtctgg ttttcgctca aaaacttcct 60
 ggaaatgaca acagcacggc aacgctgtgc cttgggcacc atgcagtacc aaacggaacg 120
 atagtga aaa caatcacgaa tgaccaaatt gaagttacta atgctactga gctggttcag 180
 agttcctcaa caggtgaaat atgcgacagt cctcatcaga tccttgatgg aaaaaactgc 240
 aactaataag atgctctatt gggagaccct cagtgtgatg gcttccaaaa taagaaatgg 300
 gacctttttg ttgaacgcag caaagcctac agcaactgtt acccttatga tgtgccggat 360
 tatgcctccc ttaggtcact agttgcctca tccggcacac tggagtttaa caatgaaagc 420
 ttcaattgga ctggagtcac tcaaaacgga acaagctctg cttgcataag gagatctaaa 480
 aacagtttct ttagtagatt gaattgggtg acccacttaa acttcaaata cccagcattg 540
 aacgtgacta tgccaaacaa tgaacaattt gacaaattgt acatttgggg ggttcaccac 600
 ccgggtacgg acaaagacca aatcttcctg tatgtcgaag catcaggaag aatcacagtc 660
 tctacaaaaa gaagccaaca aaccgtaagc ccgaatatcg gatctagacc cagagtaagg 720
 aatatcccta gcagaataag catctattgg acaatagtaa aaccgggaga catacttttg 780
 attaacagca cagggaatct aattgctcct aggggttact tcaaaatagc aagtgggaaa 840

agctcaataa tgagatcaga tgcacccatt ggcaaagtca attctgaatg catcactcca 900
 aatggaagca ttcccaatga caaaccattc caaaatgtaa acaggatcac atacggggcc 960
 tgtcccagat atgttaagca aaacactctg aaattggcaa cagggatgcg aaatgtacca 1020
 gagaaacaaa ctagaggcat atttggcgca atcgcggggtt tcataaaaaa tggttgggag 1080
 ggaatggtgg atggttggtg cggtttcagg catcaaaatt ctgaggggaag aggacaagca 1140
 gcagatctca aaagcactca agcagcaatc gatcaaatca atgggaagct gaatagattg 1200
 atcgggaaaa ccaacgagaa attccatcag attgaaaaag aattctcaga agtcgaaggg 1260
 agaattcagg accttgagaa atatgttgag gacactaaaa tagatctctg gtcatacaac 1320
 gcggagcttc ttgttgccct ggagaaccaa catacaattg atctaactga ctcaaaaatg 1380
 aacaaaactgt ttgaaaaaac aaagaagcaa ctgagggaaa atgctgagga tatgggcaat 1440
 ggttgtttca aaatatacca caaatgtgac aatgcctgca taggatcaat cagaaatgga 1500
 acttatgacc acgatgtata cagagatgaa gcattaaaca accggtttca gatcaaggga 1560
 gttgagctga agtcagggtg caaagattgg atcctatgga tttcctttgc catatcatgt 1620
 tttttgcttt gtgttgcttt gttggggttc atcatgtggg cctgccaaaa aggcaacatt 1680
 aggtgcaaca tttgcatttg a 1701

<210> 24
 <211> 3465
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Cát xét biểu hiện số 1019

<400> 24
 gtcaacatgg tggagcacga cacacttgct tactccaaaa atatcaaaga tacagtctca 60
 gaagacaaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120
 ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180
 tacaaatgcc atcattgcca taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240
 ggtcccaaag atggaccccc acccacgagg agcatcgttg aaaaagaaga cgttccaacc 300
 acgtcttcaa agcaagtgga ttgatgtgat aacatgggtg agcacgacac acttgtctac 360
 tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420
 agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480
 aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540
 atcgttgaag atgcctctgc cgacagtggg cccaagatg gacccccacc cagaggagc 600

atcgtgga	aaaagaacgt	tccaaccacg	tcttcaaagc	aagtggattg	atgtgatata	660
tccactgacg	taagggatga	cgcacaatcc	cactatcctt	cgcaagaccc	tctctctata	720
taaggaagtt	catttcattt	ggagaggtat	taaaatctta	ataggttttg	ataaaagcga	780
acgtggggaa	acccgaacca	aaccttcttc	taaactctct	ctcatctctc	ttaaagcaaa	840
cttctctctt	gtctttcttg	cgtgagcgat	cttcaacgtt	gtcagatcgt	gcttcggcac	900
cagtacaacg	ttttctttca	ctgaagcgaa	atcaaagatc	tctttgtgga	cacgtagtgc	960
ggcgccatta	aataacgtgt	acttgtccta	ttcttgtcgg	tgtggctctg	ggaaaagaaa	1020
gcttgctgga	ggctgctgtt	cagccccata	cattacttgt	tacgattctg	ctgactttcg	1080
gcgggtgcaa	tatctctact	tctgcttgac	gaggtattgt	tgctgtact	tctttcttct	1140
tcttcttgct	gattggttct	ataagaaatc	tagtatcttc	tttgaaacag	agttttcccg	1200
tggttttcga	acttggagaa	agattgttaa	gcttctgtat	attctgcca	aatttgtcgg	1260
gcccattggc	aaaaacgttg	cgattttcgg	cttattgttt	tctcttcttg	tgttggttcc	1320
ttctcagatc	ttcgccaaa	aacttctctg	aaatgacaac	agcacggcaa	cgctgtgcct	1380
tgggcaccat	gcagtaccaa	acggaacgat	agtgaaaaca	atcacgaatg	accaaattga	1440
agttactaat	gctactgagc	tggttcagag	ttcctcaaca	ggtgaaatat	gcgacagtcc	1500
tcactcagatc	cttgatggaa	aaaactgcac	actaatagat	gctctattgg	gagaccctca	1560
gtgtgatggc	ttccaaaata	agaaatggga	cctttttgtt	gaacgcagca	aagcctacag	1620
caactgttac	ccttatgatg	tgccggatta	tgctccctt	aggtcactag	ttgcctcatc	1680
cggcacactg	gagtttaaca	atgaaagctt	caattggact	ggagtactc	aaaacggaac	1740
aagctctgct	tgcataagga	gatctaataa	cagtttcttt	agtagattga	attggttgac	1800
ccacttaaac	ttcaaatacc	cagcattgaa	cgtgactatg	ccaacaatg	aacaatttga	1860
caaattgtac	atttgggggg	ttcaccaccc	gggtacggac	aaagaccaa	tcttctgtga	1920
tgctcaagca	tcaggaagaa	tcacagtctc	taccaaaga	agccaacaaa	ccgtaagccc	1980
gaatatcgga	tctagacca	gagtaaggaa	tatccctagc	agaataagca	tctattggac	2040
aatagtaaaa	ccgggagaca	tacttttgat	taacagcaca	gggaatctaa	ttgctcctag	2100
gggttacttc	aaaatacgaa	gtgggaaaag	ctcaataatg	agatcagatg	caccatttgg	2160
caaattgcaat	tctgaatgca	tactccaaa	tggaagcatt	ccaatgaca	aaccattcca	2220
aaatgtaaac	aggatcacat	acggggcctg	tccagatat	gttaagcaaa	acactctgaa	2280
attggcaaca	gggatgcgaa	atgtaccaga	gaaacaaact	agaggcatat	ttggcgcaat	2340
cgcgggtttc	atagaaaatg	gttgggaggg	aatggtggat	ggttggtagc	gtttcaggga	2400
tcaaaattct	gagggaagag	gacaagcagc	agatctcaaa	agcactcaag	cagcaatcga	2460

tcaaatcaat gggaagctga atagattgat cgggaaaacc aacgagaaat tccatcagat 2520
 tgaaaaagaa ttctcagaag tcgaaggag aattcaggac cttgagaaat atgttgagga 2580
 cactaaaata gatctctggt catacaacgc ggagcttctt gttgccctgg agaaccaaca 2640
 tacaattgat ctaactgact cagaaatgaa caaactgttt gaaaaaacaagaagcaact 2700
 gagggaaaat gctgaggata tgggcaatgg ttgtttcaaa atataccaca aatgtgacaa 2760
 tgccctgcata ggatcaatca gaaatggaac ttatgaccac gatgtatata gagatgaagc 2820
 attaaacaac cggtttcaga tcaaggagat tgagctgaag tcagggtaca aagattggat 2880
 cctatggatt tcctttgccca tatcatgttt tttgctttgt gttgctttgt tgggggtcat 2940
 catgtgggcc tgccaaaaag gcaacattag gtgcaacatt tgcatttgaa ggcctatattt 3000
 ctttagtttg aatttactgt tattcggtgt gcatttctat gtttggtgag cggttttctg 3060
 tgctcagagt gtgtttatatt tatgtaattt aatttctttg tgagctcctg ttttagcaggt 3120
 cgtcccttca gcaaggacac aaaaagattt taattttatt aaaaaaaaaa aaaaaaaga 3180
 ccgggaattc gatatcaagc ttatcgacct gcagatcggt caaacatttg gcaataaagt 3240
 ttcttaagat tgaatcctgt tgccggtctt gcgatgatta tcatataatt tctgttgaat 3300
 tacgttaagc atgtaataat taacatgtaa tgcagacgt tatttatgag atgggttttt 3360
 atgattagag tcccgcgaatt atacatttaa tacgcgatag aaaacaaaat atagcgcgca 3420
 aactaggata aattatogcg cgcggtgtca tctatgttac tagat 3465

<210> 25
 <211> 574
 <212> PRT
 <213> Virut cúm

<400> 25

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Gln Lys Leu Pro Gly Asn Asp Asn
20 25 30

Ser Thr Ala Thr Leu Cys Leu Gly His His Ala Val Pro Asn Gly Thr
35 40 45

Ile Val Lys Thr Ile Thr Asn Asp Gln Ile Glu Val Thr Asn Ala Thr
50 55 60

Glu Leu Val Gln Ser Ser Ser Thr Gly Glu Ile Cys Asp Ser Pro His
65 70 75 80

Gln Ile Leu Asp Gly Lys Asn Cys Thr Leu Ile Asp Ala Leu Leu Gly
 85 90 95

Asp Pro Gln Cys Asp Gly Phe Gln Asn Lys Lys Trp Asp Leu Phe Val
 100 105 110

Glu Arg Ser Lys Ala Tyr Ser Asn Cys Tyr Pro Tyr Asp Val Pro Asp
 115 120 125

Tyr Ala Ser Leu Arg Ser Leu Val Ala Ser Ser Gly Thr Leu Glu Phe
 130 135 140

Asn Asn Glu Ser Phe Asn Trp Thr Gly Val Thr Gln Asn Gly Thr Ser
 145 150 155 160

Ser Ala Cys Ile Arg Arg Ser Lys Asn Ser Phe Phe Ser Arg Leu Asn
 165 170 175

Trp Leu Thr His Leu Asn Phe Lys Tyr Pro Ala Leu Asn Val Thr Met
 180 185 190

Pro Asn Asn Glu Gln Phe Asp Lys Leu Tyr Ile Trp Gly Val His His
 195 200 205

Pro Gly Thr Asp Lys Asp Gln Ile Phe Leu Tyr Ala Gln Ala Ser Gly
 210 215 220

Arg Ile Thr Val Ser Thr Lys Arg Ser Gln Gln Thr Val Ser Pro Asn
 225 230 235 240

Ile Gly Ser Arg Pro Arg Val Arg Asn Ile Pro Ser Arg Ile Ser Ile
 245 250 255

Tyr Trp Thr Ile Val Lys Pro Gly Asp Ile Leu Leu Ile Asn Ser Thr
 260 265 270

Gly Asn Leu Ile Ala Pro Arg Gly Tyr Phe Lys Ile Arg Ser Gly Lys
 275 280 285

Ser Ser Ile Met Arg Ser Asp Ala Pro Ile Gly Lys Cys Asn Ser Glu
 290 295 300

Cys Ile Thr Pro Asn Gly Ser Ile Pro Asn Asp Lys Pro Phe Gln Asn
 305 310 315 320

Val Asn Arg Ile Thr Tyr Gly Ala Cys Pro Arg Tyr Val Lys Gln Asn
 325 330 335

Thr Leu Lys Leu Ala Thr Gly Met Arg Asn Val Pro Glu Lys Gln Thr
 340 345 350

Arg Gly Ile Phe Gly Ala Ile Ala Gly Phe Ile Glu Asn Gly Trp Glu
 355 360 365

Gly Met Val Asp Gly Trp Tyr Gly Phe Arg His Gln Asn Ser Glu Gly
 370 375 380

Arg Gly Gln Ala Ala Asp Leu Lys Ser Thr Gln Ala Ala Ile Asp Gln
 385 390 395 400

Ile Asn Gly Lys Leu Asn Arg Leu Ile Gly Lys Thr Asn Glu Lys Phe
 405 410 415

His Gln Ile Glu Lys Glu Phe Ser Glu Val Glu Gly Arg Ile Gln Asp
 420 425 430

Leu Glu Lys Tyr Val Glu Asp Thr Lys Ile Asp Leu Trp Ser Tyr Asn
 435 440 445

Ala Glu Leu Leu Val Ala Leu Glu Asn Gln His Thr Ile Asp Leu Thr
 450 455 460

Asp Ser Glu Met Asn Lys Leu Phe Glu Lys Thr Lys Lys Gln Leu Arg
 465 470 475 480

Glu Asn Ala Glu Asp Met Gly Asn Gly Cys Phe Lys Ile Tyr His Lys
 485 490 495

Cys Asp Asn Ala Cys Ile Gly Ser Ile Arg Asn Gly Thr Tyr Asp His
 500 505 510

Asp Val Tyr Arg Asp Glu Ala Leu Asn Asn Arg Phe Gln Ile Lys Gly
 515 520 525

Val Glu Leu Lys Ser Gly Tyr Lys Asp Trp Ile Leu Trp Ile Ser Phe
 530 535 540

Ala Ile Ser Cys Phe Leu Leu Cys Val Ala Leu Leu Gly Phe Ile Met
 545 550 555 560

Trp Ala Cys Gln Lys Gly Asn Ile Arg Cys Asn Ile Cys Ile
 565 570

<210> 26
 <211> 40
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> đoạn mỗi IF-S2+S4-B Bris.c

<400> 26
 tctcagatct tcgccgatcg aatctgcact ggaataacat

40

<210> 27
 <211> 43
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> đoạn mỗi IF-S1a4-B Bris.r

<400> 27
 actaaagaaa ataggccttt atagacagat ggagcaagaa aca

43

<210> 28
 <211> 1758
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> trình tự nucleotit của gen HA B Brisbane tổng hợp

<400> 28
 atgaaggcaa taattgtact actcatggta gtaacatcca atgcagatcg aatctgcact 60
 ggaataacat cgtcaaactc accacatgtc gtcaaaactg ctactcaagg ggaggtcaat 120
 gtgactggtg taataccact gacaacaaca cccaccaaact ctcattttgc aaatctcaaa 180
 ggaacagaaa ccagggggaa actatgcccc aaatgcctca actgcacaga tctggacgta 240
 gccttgggca gaccaaactg cacggggaaa ataccctcgg caagagtttc aatactccat 300
 gaagtcagac ctgttacatc tgggtgcttt cctataatgc acgacagaac aaaaattaga 360
 cagctgccta accttctccg aggatacgaa catatcaggt tatcaacca taacgttatc 420
 aatgcagaaa atgcaccagg aggaccctac aaaattggaa cctcaggggtc ttgccctaac 480
 attaccaatg gaaacggatt ttctgcaaca atggcttggg ccgtcccaaa aaacgacaaa 540
 aacaaaacag caacaaatcc attaacaata gaagtacat acatttgtac agaaggagaa 600
 gaccaaatta ccgtttgggg gttccactct gacaacgaga cccaaatggc aaagctctat 660
 ggggactcaa agccccagaa gttcacctca tctgccaacg gagtgaccac acattacgtt 720

tcacagattg gtggcttccc aaatcaaaca gaagacggag gactaccaca aagtggtaga 780
 attgttggtt attacatggt gcaaaaatct gggaaaacag gaacaattac ctatcaaagg 840
 ggtatTTTTat tgcctcaaaa ggtgtggtgc gcaagtggca ggagcaagggt aataaaagga 900
 tccttgcctt taattggaga agcagattgc ctccacgaaa aatacgggtg attaaacaaa 960
 agcaagcctt actacacagg ggaacatgca aaggccatag gaaattgccc aatatgggtg 1020
 aaaacaccct tgaagctggc caatggaacc aaatatagac ctcttgcaaa actattaaag 1080
 gaaaggggtt tcttcggagc tattgctggt ttcttagaag gaggatggga aggaatgatt 1140
 gcaggttggc acggatacac atcccatggg gcacatggag tagcgggtggc agcagacctt 1200
 aagagcactc aagaggccat aaacaagata acaaaaaatc tcaactcttt gagtgagctg 1260
 gaagtaaaga atcttcaaag actaagcggg gccatggatg aactccacaa cgaaatacta 1320
 gaactagatg agaaagtgga tgatctcaga gctgatacaa taagctcaca aatagaactc 1380
 gcagtctctg tttccaatga aggaataata aacagtgaag atgaacatct cttggcgctt 1440
 gaaagaaaagc tgaagaaaat gctgggcccc tctgctgtag agataggga tggatgcttt 1500
 gaaaccaaac acaagtgcaa ccagacctgt ctgcacagaa tagctgctgg tacctttgat 1560
 gcaggagaat tttctctccc cacctttgat tcaactgaata ttactgctgc atctttaaat 1620
 gacgatggat tggataatca tactatactg ctttactact caactgctgc ctccagtttg 1680
 gctgtaacac tgatgatagc tatctttggt gtttatatgg tctccagaga caatgtttct 1740
 tgctccatct gtctataa 1758

<210> 29

<211> 3525

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Cát xét biểu hiện số 1029

<400> 29

gtcaacatgg tggagcacga cacacttgct tactccaaaa atatcaaaga tacagtctca 60
 gaagaccaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120
 ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180
 tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240
 ggtcccaaag atggaccccc acccagagg agcatcgtgg aaaaagaaga cgttccaacc 300
 acgtcttcaa agcaagtgga ttgatgtgat aacatgggtg agcacgacac acttgtctac 360
 tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420
 agggtaatat ccggaaacct cctcggaatt cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc	540
atcgttgaag atgcctctgc cgacagtggc cccaaagatg gacccccacc cagcaggagc	600
atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc	660
tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata	720
taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga	780
acgtggggaa acccgaacca aaccttcttc taaactctct ctcactctct ttaaagcaaa	840
cttctctctt gtctttcttg cgtgagcgat cttcaacggt gtcagatcgt gcttcggcac	900
cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc	960
ggcgccatta aataacgtgt acttgtccta ttcttgcgg tgtggtcttg ggaaaagaaa	1020
gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg	1080
gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct	1140
tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg	1200
tggttttoga acttgagaa agattgttaa gcttctgtat attctgcca aatttgcgg	1260
gcccattggc aaaaacgttg cgattttcgg cttattgttt tctcttcttg tgttggttcc	1320
ttctcagatc ttgcgcgac gaatctgcac tggaaataaca tcgtcaaact caccacatgt	1380
cgtcaaaact gctactcaag gggaggtcaa tgtgactggc gtaataccac tgacaacaac	1440
accacacaaa tctcattttg caaatctcaa aggaacagaa accaggggga aactatgcc	1500
aaaatgcctc aactgcacag atctggacgt agccttgggc agacaaaat gcacggggaa	1560
aataccctcg gcaagagttt caatactcca tgaagtcaga cctgttacat ctgggtgctt	1620
tcctataatg cagcagagaa caaaaattag acagctgcct aaccttctcc gaggatacga	1680
acatatcagg ttatcaaccc ataacgttat caatgcagaa aatgcaccag gaggacccta	1740
caaaattgga acctcagggt cttgccctaa cattaccaat ggaaacggat ttttcgcaac	1800
aatggcttgg gccgtcccaa aaaacgacaa aaacaaaaca gcaacaaatc cattaacaat	1860
agaagtacca tacatttgta cagaaggaga agaccaaatt accgtttggg ggttccactc	1920
tgacaacgag acccaaattg caaagctcta tggggactca aagccccaga agttcacctc	1980
atctgccaac ggagtgacca cacattacgt ttcacagatt ggtggcttcc caaatcaaac	2040
agaagacgga ggactaccac aaagtggtag aattgttggt gattacatgg tgcaaaaatc	2100
tgggaaaaca ggaacaatta cctatcaaag gggatattta ttgcctcaa aggtgtggtg	2160
cgcaagtggc aggagcaagg taataaaagg atccttgcct ttaattggag aagcagattg	2220
cctccacgaa aaatacgggt gattaaacaa aagcaagcct tactacacag gggaacatgc	2280

```

aaaggccata ggaaattgcc caatatgggt gaaaacaccc ttgaagctgg ccaatggaac 2340
caaatataga cctcctgcaa aactattaaa ggaaaggggt ttcttcggag ctatttgctgg 2400
tttcttagaa ggaggatggg aaggaatgat tgcaggttgg cacggatata catcccatgg 2460
ggcacatgga gtagcgggtg cagcagacct taagagcact caagaggcca taaacaagat 2520
aacaaaaaat ctcaactctt tgagtgaagt ggaagtaaag aatcttcaaa gactaagcgg 2580
tgccatggat gaactccaca acgaaatact agaactagat gagaaagtgg atgatctcag 2640
agctgataca ataagctcac aaatagaact cgcagtcctg ctttccaatg aaggaataat 2700
aaacagtga gatgaacatc tcttggcgct tgaaagaaag ctgaagaaaa tgctgggccc 2760
ctctgctgta gagataggga atggatgctt tgaaacaaaa cacaagtga accagacctg 2820
tctcgacaga atagctgctg gtaccttga tgcaggagaa ttttctctcc ccaccttga 2880
ttcactgaat attactgctg catctttaa tgacgatgga ttggataatc atactatact 2940
gctttactac tcaactgctg cctccagttt ggctgtaaca ctgatgatag ctatctttgt 3000
tgtttatatg gtctccagag acaatgtttc ttgctccatc tgtctataaa ggcctatttt 3060
ctttagtttg aatttactgt tattcgggtg gcatttctat gtttggtgag cggttttctg 3120
tgctcagagt gtgtttattt tatgtaattt aatttctttg tgagctcctg tttagcaggt 3180
cgtcccttca gcaaggacac aaaaagattt taattttatt aaaaaaaaaa aaaaaaaga 3240
ccgggaattc gatatcaagc ttatcgacct gcagatcggt caaacatttg gcaataaagt 3300
ttcttaagat tgaatcctgt tgccgggtctt gcgatgatta tcatataatt tctgttgaat 3360
tacgttaagc atgtaataat taacatgtaa tgcagacgt tatttatgag atgggttttt 3420
atgattagag tcccgcatt atacatttaa tacgcgatag aaaacaaaat atagcgcgca 3480
aactaggata aattatcgcg cgcggtgtca tctatgttac tagat 3525

```

<210> 30
 <211> 594
 <212> PRT
 <213> Virut cúm

<400> 30

```

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1           5           10           15

```

```

Leu Val Pro Ser Gln Ile Phe Ala Asp Arg Ile Cys Thr Gly Ile Thr
                20                25                30

```

```

Ser Ser Asn Ser Pro His Val Val Lys Thr Ala Thr Gln Gly Glu Val
          35          40          45

```

Asn Val Thr Gly Val Ile Pro Leu Thr Thr Thr Pro Thr Lys Ser His
 50 55 60

Phe Ala Asn Leu Lys Gly Thr Glu Thr Arg Gly Lys Leu Cys Pro Lys
 65 70 75 80

Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Lys Cys
 85 90 95

Thr Gly Lys Ile Pro Ser Ala Arg Val Ser Ile Leu His Glu Val Arg
 100 105 110

Pro Val Thr Ser Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile
 115 120 125

Arg Gln Leu Pro Asn Leu Leu Arg Gly Tyr Glu His Ile Arg Leu Ser
 130 135 140

Thr His Asn Val Ile Asn Ala Glu Asn Ala Pro Gly Gly Pro Tyr Lys
 145 150 155 160

Ile Gly Thr Ser Gly Ser Cys Pro Asn Ile Thr Asn Gly Asn Gly Phe
 165 170 175

Phe Ala Thr Met Ala Trp Ala Val Pro Lys Asn Asp Lys Asn Lys Thr
 180 185 190

Ala Thr Asn Pro Leu Thr Ile Glu Val Pro Tyr Ile Cys Thr Glu Gly
 195 200 205

Glu Asp Gln Ile Thr Val Trp Gly Phe His Ser Asp Asn Glu Thr Gln
 210 215 220

Met Ala Lys Leu Tyr Gly Asp Ser Lys Pro Gln Lys Phe Thr Ser Ser
 225 230 235 240

Ala Asn Gly Val Thr Thr His Tyr Val Ser Gln Ile Gly Gly Phe Pro
 245 250 255

Asn Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val
 260 265 270

Asp Tyr Met Val Gln Lys Ser Gly Lys Thr Gly Thr Ile Thr Tyr Gln
 275 280 285

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

Lys Val Ile Lys Gly Ser Leu Pro Leu Ile Gly Glu Ala Asp Cys Leu
 305 310 315 320

His Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys Pro Tyr Tyr Thr Gly
 325 330 335

Glu His Ala Lys Ala Ile Gly Asn Cys Pro Ile Trp Val Lys Thr Pro
 340 345 350

Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Ala Lys Leu Leu
 355 360 365

Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala Gly Phe Leu Glu Gly Gly
 370 375 380

Trp Glu Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His Gly Ala
 385 390 395 400

His Gly Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala Ile
 405 410 415

Asn Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val Lys
 420 425 430

Asn Leu Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu Ile
 435 440 445

Leu Glu Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile Ser
 450 455 460

Ser Gln Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile Asn
 465 470 475 480

Ser Glu Asp Glu His Leu Leu Ala Leu Glu Arg Lys Leu Lys Lys Met
 485 490 495

Leu Gly Pro Ser Ala Val Glu Ile Gly Asn Gly Cys Phe Glu Thr Lys
 500 505 510

His Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala Ala Gly Thr Phe
 515 520 525

Asp Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser Leu Asn Ile Thr

530 535 540
 Ala Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn His Thr Ile Leu Leu
 545 550 555 560
 Tyr Tyr Ser Thr Ala Ala Ser Ser Leu Ala Val Thr Leu Met Ile Ala
 565 570 575
 Ile Phe Val Val Tyr Met Val Ser Arg Asp Asn Val Ser Cys Ser Ile
 580 585 590

Cys Leu

<210> 31
 <211> 6861
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc 1194

<400> 31
 tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 60
 gacgttttta atgtactgaa ttaacgccga atccccgggct ggtatattta tatgtttgtca 120
 aataactcaa aaaccataaa agtttaagtt agcaagtgtg tacattttta cttgaacaaa 180
 aatattcacc tactactggt ataaatcatt attaaacatt agagtaaaga aatatggatg 240
 ataagaacaa gagtagtgat attttgacaa caattttgtt gcaacatttg agaaaatttt 300
 gttgttctct cttttcattg gtcaaaaaca atagagagag aaaaaggaag agggagaata 360
 aaaacataat gtgagtatga gagagaaagt tgtacaaaag ttgtaccaa atagttgtac 420
 aaatatcatt gaggaatttg acaaaagcta cacaaataag ggtaattgc tgtaaataaa 480
 taaggatgac gcattagaga gatgtacat tagagaattt ttggcaagtc attaaaaaga 540
 aagaataaat tttttttaa attaaaagtt gagtcatttg attaaacatg tgattattta 600
 atgaattgat gaaagagttg gattaaagtt gtattagtaa ttagaatttg gtgtcaaatt 660
 taatttgaca tttgatcttt tcttatatat tgcccatag agtcagttaa ctcatTTTTA 720
 tttttcatag atcaaataag agaaataacg gtatattaat ccctccaaaa aaaaaaacg 780
 gtatattttac taaaaaatct aagccacgta ggaggataac aggatccccg taggaggata 840
 acatccaatc caaccaatca caacaatcct gatgagataa cccacttta gccacgcat 900
 ctgtggcaca tctacattat ctaaatacaca cattcttcca cacatctgag ccacacaaaa 960
 accaatccac atctttatca ccattctat aaaaaatcac actttgtgag tctacacttt 1020

gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag	1080
aaaatggaac gagctataca aggaaacgac gctaggaac aagctaacag tgaacgttgg	1140
gatggaggat caggaggtac cacttctccc ttcaaacttc ctgacgaaag tccgagttgg	1200
actgagtggc ggctacataa cgatgagacg aattcgaatc aagataatcc ccttggtttc	1260
aaggaaagct ggggtttcgg gaaagttgta tttaagagat atctcagata cgacaggacg	1320
gaagcttcac tgcacagagt ccttggatct tggacgggag attcggttaa ctatgcagca	1380
tctcgatttt tcggtttcga ccagatcgga tgtacctata gtattcggtt tcgaggagtt	1440
agtatcacccg tttctggagg gtcgcgaact cttcagcatc tctgtgagat ggcaattcgg	1500
tctaagcaag aactgctaca gcttgcccca atcgaagtgg aaagtaatgt atcaagagga	1560
tgccctgaag gtactcaaac cttcgaaaaa gaaagcgagt aagttaaaat gcttcttcgt	1620
ctcctatttta taatatggtt tgttattggt aattttgttc ttgtagaaga gcttaattaa	1680
tcgttggtgt tatgaaatac tatttgatg agatgaactg gtgtaatgta attcatttac	1740
ataagtggag tcagaatcag aatgtttcct ccataactaa ctagacatga agacctgccg	1800
cgtacaattg tcttatattt gaacaactaa aattgaacat cttttgccac aactttataa	1860
gtggttaata tagctcaa atatgggtcaa gttcaataga ttaataatgg aaatatcagt	1920
tatcgaaatt cattaacaat caacttaacg ttattaacta ctaattttat atcatcccct	1980
ttgataaatg atagtacacc aattaggaag gagcatgctc gcctaggaga ttgtcgtttc	2040
ccgccttcag tttgcaagct gctctagccg tgtagccaat acgcaaaccg cctctccccg	2100
cgcggtggga attactagcg cgtgtcgaca cgcggtggcg gccctagcag aaggcatggt	2160
gttgtgactc cgaggggttg cctcaaactc tatcttataa cggcggtgga ggcatggagg	2220
caagggcatt ttggtaat ttagtagttg tggaaaatga cgtcatttac ttaaagacga	2280
agtcttgcca caaggggggc ccacgccga ttttaatat accggcgtgg cccacctta	2340
tcgagagtgc tttagcacga gcggtccaga tttaaagtag aaaagttccc gccactagg	2400
gttaaaggtg ttcacactat aaaagcatat acgatgtgat ggtatttgat aaagcgtata	2460
ttgtatcagg tatttcgctc ggatacgaat tattcgtaca agcttcttaa gccggtcaac	2520
atggtggagc acgacacact tgtctactcc aaaaatatca aagatacagt ctgagaagac	2580
caaagggcaa ttgagacttt tcaacaaagg gtaatatccg gaaacctcct cggattccat	2640
tgcccagcta tctgtcactt tattgtgaag atagtggaaa aggaaggtgg ctctacaaa	2700
tgccatcatt gcgataaagg aaaggccatc gttgaagatg cctctgccga cagtgggtccc	2760
aaagatggac cccacccac gaggagcatc gtggaaaaag aagacgttcc aaccacgtct	2820

tcaaagcaag tggattgatg tgataacatg gtggagcacg acacacttgt ctactccaaa	2880
aatatcaaag atacagtctc agaagaccaa agggcaattg agacttttca acaaagggtg	2940
atatccggaa acctcctcgg attccattgc ccagctatct gtcactttat tgtgaagata	3000
gtggaaaagg aaggtggctc ctacaaatgc catcattgcg ataaaggaaa ggccatcggt	3060
gaagatgcct ctgccgacag tggccccaaa gatggacccc caccacagag gagcatcggtg	3120
gaaaaagaag acgttccaac cacgtcttca aagcaagtgg attgatgtga tatctccact	3180
gacgtaagggt atgacgcaca atcccactat ccttcgcaag acccttcctc tatataagga	3240
agttcatttc atttgagag gtattaaaat cttaataggt tttgataaaa gcgaacgtgg	3300
ggaaacccga accaaacctt cttctaaact ctctctcatc tctcttaaag caaacttctc	3360
tcttgctttt cttgcgtgag cgatcttcaa cgttgtcaga tcgtgcttcg gcaccagtac	3420
aacgttttct ttcactgaag cgaaatcaaa gatctctttg tggacacgta gtgcggcgcc	3480
attaaataac gtgtacttgt cctattcttg tcggtgtggt cttgggaaaa gaaagcttgc	3540
tggaggctgc tgttcagccc catacattac ttgttacgat tctgctgact ttcggcgggt	3600
gcaatatctc tacttctgct tgacgaggtg ttgttgctg tacttctttc ttcttcttct	3660
tgctgattgg ttctataaga aatctagtat tttctttgaa acagagtttt cccgtgggtt	3720
tcgaacttgg agaaagattg ttaagcttct gtatatctg cccaaatttg tcgggcccac	3780
ggcgaaaaac gttgcgattt tcggcttatt gttttctctt cttgtgttggt ttccttctca	3840
gatcttcgcc gcggctcctc agccaaaacg acacccccat ctgtctatcc actggcccct	3900
ggatctgctg cccaaactaa ctccatggtg accctgggat gcctgggtcaa gggctatttc	3960
cctgagccag tgacagtgac ctggaactct ggatccctgt ccagcgggtg gcacaccttc	4020
ccagctgtcc tgcagtctga cctctacact ctgagcagct cagtgactgt cccctccagc	4080
acctggccca gcgagaccgt cacctgcaac gttgccacc cggccagcag caccaagggtg	4140
gacaagaaaa ttgtgcccag ggattgtggt tgtaagcctt gcatatgtac agtcccagaa	4200
gtatcatctg tcttcatctt cccccaaaag cccaaggatg tgctcaccat tactctgact	4260
cctaagggtca cgtgtgttgt ggtagacatc agcaaggatg atcccagggt ccagttcagc	4320
tggttttag atgatgtgga ggtgcacaca gctcagacgc aaccccgga ggagcagttc	4380
aacagcactt tccgctcagt cagtgaactt cccatcatgc accaggactg gctcaatggc	4440
aaggaaggcc tattttcttt agtttgaatt tactgttatt cgggtgtgcat ttctatgttt	4500
ggtgagcgggt tttctgtgct cagagtgtgt ttattttatg taatttaatt tctttgtgag	4560
ctcctgttta gcaggtcgtc ccttcagcaa ggacacaaaa agattttaat ttatttaaaa	4620
aaaaaaaaa aaaagaccgg gaattcgata tcaagcttat cgacctgcag atcgttcaaa	4680

catttggcaa taaagtttct taagattgaa tcctgttgcc ggtcttgcca tgattatcat	4740
ataatttctg ttgaattacg ttaagcatgt aataattaac atgtaatgca tgacgttatt	4800
tatgagatgg gtttttatga ttagagtccc gcaattatac atttaatacg cgatagaaaa	4860
caaaatatag cgcgcaaact aggataaatt atcgcgcgcg gtgtcatcta tgttactaga	4920
tctctagagt ctcaagcttg gcgcggggta ccgagctcga attccgagtg tacttcaagt	4980
cagttggaaa tcaataaaat gattatttta tgaatatatt tcattgtgca agtagataga	5040
aattacatat gttacataac acacgaaata aacaaaaaaaa cacaatccaa aacaaacacc	5100
ccaaacaaaa taacactata tatatcctcg tatgaggaga ggcacgttca gtgactcgac	5160
gattcccgag caaaaaaagt ctccccgtca cacatatagt gggtgacgca attatcttca	5220
aagtaatcct tctgttgact tgtcattgat aacatccagt cttcgtcagg attgcaaaga	5280
attatagaag ggatcccacc ttttattttc ttcttttttc catatttagg gttgacagtg	5340
aaatcagact ggcaacctat taattgcttc cacaatggga cgaacttgaa ggggatgtcg	5400
tcgatgatat tataggtggc gtgttcatcg tagttggtga agtcgatggc cccgttccag	5460
tagttgtgtc gcccgagact tctagcccag gtggtctttc cggtagcagt tgggtccgag	5520
atgtagaggc tggggtgtct gaccccagtc cttccctcat cctgggttaga tcggccatcc	5580
actcaaggtc agattgtgct tgatcgtagg agacaggatg tatgaaagtg taggcatcga	5640
tgcttacatg atataggtgc gtctctctcc agttgtgcag atcttcgtgg cagcggagat	5700
ctgattctgt gaagggcgac acgtactgct caggttgtgg aggaaataat ttgttggtcg	5760
aatattccag ccattgaagc tttgttgccc attcatgagg gaattcttct ttgatcatgt	5820
caagatactc ctcttagac gttgcagtct ggataatagt tcgccatcgt gcgtcagatt	5880
tgcgaggaga gaccttatga tctcgaaat ctctctggt tttaatatct ccgtcctttg	5940
atatgtaatc aaggacttgt ttagagtttc tagctggctg gatattaggg tgatttcctt	6000
caaaatcgaa aaaagaagga tccctaatac aagggttttt atcaagctgg ataagagcat	6060
gatagtgggt agtgccatct tgatgaagct cagaagcaac accaaggaag aaaataagaa	6120
aagggtgtgag tttctcccag agaaactgga ataaatcatc tctttgagat gagcacttgg	6180
ggtaggtaag gaaaacatat ttagattgga gtctgaagtt cttgctagca gaaggcatgt	6240
tgttgtgact ccgagggggt gcctcaaact ctatcttata accggcgtgg aggcatggag	6300
gcaagggcat tttggtaatt taagtagtta gtggaaaatg acgtcattta cttaaagacg	6360
aagtcttgcg acaagggggg cccacgccga attttaatat tacggcgtg gccccacctt	6420
atcgcgagtg ctttagcacg agcgggtccag atttaaagta gaaaagttcc cgcccactag	6480

gggttaaaggt gttcacacta taaaagcata tacgatgtga tggatatttga tggagcgtat 6540
 attgtatcag gtatttccgt cggatacgaa ttattcgtac ggccggccac tagtggcact 6600
 ggccgtcgtt ttacaacgtc gtgactggga aaaccctggc gttacccaac ttaatcgcct 6660
 tgcagcacat ccccctttcg ccagctggcg taatagcgaa gaggcccgca ccgatcgccc 6720
 ttcccaacag ttgcgagcc tgaatggcga atgctagagc agcttgagct tggatcagat 6780
 tgtcgtttcc cgccttcagt ttaactatc agtgtttgac aggatatatt ggcgggtaaa 6840
 cctaagagaa aagagcgttt a 6861

<210> 32

<211> 5555

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Cát xét biểu hiện số 1008

<400> 32

ctagcagaag gcatgttggt gtgactccga ggggttgctt caaactctat cttataaccg 60
 gcgtggaggc atggaggcaa gggcattttg gtaatttaag tagttagtgg aaaatgacgt 120
 catttactta aagacgaagt cttgcgacaa ggggggcccc cgccgaattt taatattacc 180
 ggcggtggccc caccttatcg cgagtgtttt agcacgagcg gtccagattt aaagtagaaa 240
 agttcccgcc cactaggggt aaaggtgttc aactataaaa agcatatacg atgtgatggt 300
 atttgataaaa gcgtatatgt tatcaggtat ttccgtcgga tacgaattat tcgtacaagc 360
 ttcttaagcc ggtaacatg gtggagcacg acacacttgt ctactccaaa aatatcaaag 420
 atacagtctc agaagaccaa agggcaattg agacttttca acaaagggta atatccggaa 480
 acctcctcgg attccattgc ccagctatct gtcactttat tgtgaagata gtggaaaagg 540
 aaggtggctc ctacaaatgc catcattgag ataaaggaaa ggccatcggt gaagatgcct 600
 ctgccgacag tgggtccaaa gatggacccc caccacgag gagcatcgtg gaaaaagaag 660
 acgttccaac cacgtcttca aagcaagtgg attgatgtga taacatggtg gagcacgaca 720
 cacttgctca ctccaaaaat atcaaagata cagtctcaga agaccaaagg gcaattgaga 780
 cttttcaaca aagggttaata tccggaaacc tcctcggatt ccattgcca gctatctgtc 840
 actttattgt gaagatagtg gaaaaggag gtggctccta caaatgccat cattgagata 900
 aaggaaaggc catcgttgaa gatgcctctg ccgacagtgg tcccaaagat ggacccccac 960
 ccacgaggag catcgttgaa aaagaagacg ttccaaccac gtcttcaaag caagtggatt 1020
 gatgtgatat ctccactgac gtaagggatg acgcacaatc ccactatcct tcgcaagacc 1080
 cttcctctat ataaggaagt tcatttcatt tggagaggta ttaaaatcct aatagggttt 1140

gataaaagcg aacgtgggga aaccggaacc aaaccttctt ctaaactctc tctcatctct	1200
cttaaagcaa acttctctct tgtctttctt gcgtgagcga tcttcaacgt tgtcagatcg	1260
tgcttcggca ccagtacaac gttttctttc actgaagcga aatcaaagat ctctttgtgg	1320
acacgtagtg cggcgccatt aaataacgtg tacttgtcct attcttgtcg gtgtggtctt	1380
gggaaaagaa agcttgctgg aggctgctgt tcagcccat acattacttg ttacgattct	1440
gctgactttc ggcggtgca atatctctac ttctgcttga cgaggatttg ttgcctgtac	1500
ttctttcttc ttcttcttgc tgattggttc tataagaaat ctagtatttt ctttgaaaca	1560
gagttttccc gtggttttcg aacttggaga aagattgtta agcttctgta tattctgccc	1620
aaatttgtcg ggcccatggc gaaaaacgtt gcgattttcg gcttattgtt ttctcttctt	1680
gtgttggttc cttctcagat cttcgccgat cgaatctgca ctggaataac atcgtaaac	1740
tcaccacatg tcgtcaaaac tgctactcaa ggggaggtca atgtgactgg tgtaatacca	1800
ctgacaacaa caccaccaa atctcatttt gcaaactctc aaggaacaga aaccaggggg	1860
aaactatgcc caaaatgcct caactgcaca gatctggacg tagccttggg cagacaaaaa	1920
tgcacgggga aaataccctc ggcaagagtt tcaatactcc atgaagtcag acctgttaca	1980
tctgggtgct ttctataat gcacgacaga acaaaaatta gacagctgcc taaccttctc	2040
cgaggatacg aacatatcag gttatcaacc cataacgtta tcaatgcaga aaatgcacca	2100
ggaggaccct acaaaatttg aacctcaggg tcttgcccta acattaccaa tggaaacgga	2160
tttttcgcaa caatggcttg ggccgtccca aaaaacgaca aaaacaaaac agcaacaaat	2220
ccattaacaa tagaagtacc atacatttgt acagaaggag aagaccaaat taccgttttg	2280
gggttccact ctgacaacga gacccaaatg gcaaagctct atggggactc aaagccccag	2340
aagttcacct catctgcaa cggagtgacc acacattacg ttccacagat tgggtggcttc	2400
ccaaatcaaa cagaagacgg aggactacca caaagtggta gaattgttgt tgattacatg	2460
gtgcaaaaat ctgggaaaac aggaacaatt acctatcaaa ggggtatttt attgcctcaa	2520
aaggtgtggt gcgcaagtgg caggagcaag gtaataaaaag gatccttgcc ttttaattgga	2580
gaagcagatt gcctccacga aaaatacggg ggattaaaca aaagcaagcc ttactacaca	2640
ggggaacatg caaaggccat aggaaattgc ccaatatggg tgaaaacacc cttgaagctg	2700
gccaatggaa ccaaatatag acctcctgca aaactattaa aggaaagggg tttcttcgga	2760
gctattgctg gtttcttaga aggaggatgg gaaggaatga ttgcaggttg gcacggatac	2820
acatcccatg gggcacatgg agtagcggtg gcagcagacc ttaagagcac tcaagaggcc	2880
ataaacaaga taacaaaaaa tctcaactct ttgagtgagc tggaagtaaa gaatcttcaa	2940

agactaagcg gtgccatgga tgaactccac aacgaaatac tagaactaga tgagaaagtg	3000
gatgatctca gagctgatac aataagctca caaatagaac tcgcagtcct gctttccaat	3060
gaaggaataa taaacagtga agatgaacat ctcttggcgc ttgaaagaaa gctgaagaaa	3120
atgctggggc cctctgctgt agagataggg aatggatgct ttgaaaccaa acacaagtgc	3180
aaccagacct gtctcgacag aatagctgct ggtacctttg atgcaggaga attttctctc	3240
cccacctttg attcactgaa tattactgct gcatctttaa atgacgatgg attggataat	3300
catactatac tgctttacta ctcaactgct gcctccagtt tggctgtaac actgatgata	3360
gctatctttg ttgtttatat ggtctccaga gacaatgttt cttgctccat ctgtctataa	3420
aggcctatit tctttagttt gaatttactg ttattcgggtg tgcatttcta tgtttggtga	3480
gcggttttct gtgctcagag tgtgtttatt ttatgtaatt taatttcttt gtgagctcct	3540
gtttagcagg tcgtcccttc agcaaggaca caaaaagatt ttaattttat taaaaaaaa	3600
aaaaaaaaag accgggaatt cgatatcaag cttatcgacc tgcagatcgt tcāaacatit	3660
ggcaataaag tttcttaaga ttgaatcctg ttgccgggtc tgcgatgatt atcatataat	3720
ttctgttgaa ttacgttaag catgtaataa ttaacatgta atgcatgacg ttatttatga	3780
gatgggtttt tatgattaga gtcccgaat tatacattta atacgcgata gaaaacaaaa	3840
tatagcgcgc aaactaggat aaattatcgc gcgcgggtgc atctatgtta ctagatctct	3900
agagtctcaa gcttggcgcg gggtagcgag ctccaattcc gagtgactt caagtcagtt	3960
ggaaatcaat aaaatgatta ttttatgaat atatttcatt gtgcaagtag atagaaatta	4020
catatgttac ataacacacg aaataaacia aaaaacacia tccaaaacia acaccccaaa	4080
caaaataaca ctatatatat cctcgtatga ggagaggcac gttcagtgac tcgacgattc	4140
ccgagcaaaa aaagtctccc cgtcacacat atagtgggtg acgcaattat cttcaaagta	4200
atccttctgt tgacttgtca ttgataacat ccagtcttcg tcaggattgc aaagaattat	4260
agaagggatc ccacctttta ttttcttctt ttttccatat ttagggttga cagtgaatc	4320
agactggcaa cctattaatt gcttcacaa tgggacgaac ttgaagggga tgtcgtcgat	4380
gatattatag gtggcgtgtt catcgtagtt ggtgaagtcg atgggtccgt tccagtagtt	4440
gtgtcgcccg agacttctag ccaggtgggt ctttccggtg cgagttggtc cgcagatgta	4500
gaggctgggg tgtctgacct cagtccttcc ctcatcctgg ttagatcggc catccactca	4560
aggtcagatt gtgcttgatc gtaggagaca ggatgtatga aagtgtaggc atcgatgctt	4620
acatgatata ggtgcgtctc tctccagttg tgcagatctt cgtggcagcg gagatctgat	4680
tctgtgaagg gcgacacgta ctgctcaggt tgtggaggaa ataatttggt ggctgaatat	4740
tccagccatt gaagctttgt tgcccattca tgagggaatt cttctttgat catgtcaaga	4800

tactcctcct tagacgttgc agtctggata atagttcgcc atcgtgcgtc agatttgcca 4860
ggagagacct tatgatctcg gaaatctcct ctggttttta tatctccgtc ctttgatatg 4920
taatcaagga cttgttttaga gtttctagct ggctggatat tagggtgatt tccttcaaaa 4980
tcgaaaaaag aaggatccct aatacaaggt tttttatcaa gctggataag agcatgatag 5040
tgggtagtgcatccttgatg aagctcagaa gcaacaccaa ggaagaaaat aagaaaaggt 5100
gtgagtttct cccagagaaa ctggaataaa tcatctcttt gagatgagca cttggggtag 5160
gtaaggaaaa catattttaga ttggagtctg aagttcttgc tagcagaagg catgttggtg 5220
tgactccgag ggggttgctc aaactctatc ttataaccgg cgtggaggca tggaggcaag 5280
ggcatttttg taatttaagt agttagtga aaatgacgtc atttacttaa agacgaagtc 5340
ttgcgacaag gggggccac gccgaatttt aatattaccg gcgtggcccc accttatcgc 5400
gagtgtttta gcacgagcgg tccagattta aagtagaaaa gttcccgcgc actaggggta 5460
aaggtgttca cactataaaa gcatatacga tgtgatggta tttgatggag cgtatattgt 5520
atcaggtatt tccgtcggat acgaattatt cgtac 5555

<210> 33
<211> 50
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> đoạn mỗi dTmH5I-B Bris.r

<400> 33
ttgacagtat ttggtaatta tccaatccat cgtcatttaa agatgcagca 50

<210> 34
<211> 50
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> đoạn mỗi B Bris-dTmH5I.c

<400> 34
catctttaaa tgacgatgga ttggataatt accaaatact gtcaatttat 50

<210> 35
<211> 50
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> đoạn mỗi IF-SlaS4-dTmH5I.r

<400> 35

actaaagaaa ataggccttt aaatgcaaat tctgcattgt aacgatccat

50

<210> 36

<211> 5555

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Cát xét biểu hiện số 1009

<400> 36

ctagcagaag gcatgttgtt gtgactccga ggggttgctt caaactctat cttataaccg	60
gcgtggaggc atggaggcaa gggcattttg gtaatttaag tagttagtgg aaaatgacgt	120
catttactta aagacgaagt cttgcgacaa ggggggcccc cgccgaattt taatattacc	180
ggcgtggccc caccttatcg cgagtgtttt agcacgagcg gtccagattt aaagtagaaa	240
agttcccgcc cactaggggt aaaggtgttc acactataaa agcatatacg atgtgatggg	300
atttgataaa gcgtatatgt tatcaggtat ttccgtcgga tacgaattat tcgtacaagc	360
ttcttaagcc ggtcaacatg gtggagcacg acacacttgt ctactccaaa aatatcaaag	420
atacagtctc agaagaccaa agggcaattg agacttttca acaaagggtt atatccggaa	480
acctcctcgg attccattgc ccagctatct gtcactttat tgtgaagata gtggaaaagg	540
aaggtggctc ctacaaatgc catcattgcg ataaaggaaa ggccatcgtt gaagatgcct	600
ctgccgacag tgggtccaaa gatggacccc caccacgag gagcatcgtg gaaaaagaag	660
acgttccaac cacgtcttca aagcaagtgg attgatgtga taacatgggt gagcacgaca	720
cacttgtcta ctccaaaaat atcaaagata cagtctcaga agaccaaagg gcaattgaga	780
cttttcaaca aagggttaata tccggaaacc tcctcggatt ccattgcca gctatctgtc	840
actttattgt gaagatagtg gaaaaggaag gtggctccta caaatgccat cattgcgata	900
aaggaaaggc catcgttgaa gatgcctctg ccgacagtgg tcccaaagat ggacccccac	960
ccacgaggag catcgtggaa aaagaagacg ttccaaccac gtcttcaaag caagtggatt	1020
gatgtgatat ctccactgac gtaagggatg acgcacaatc ccactatcct tcgcaagacc	1080
cttcctctat ataaggaagt tcatttcatt tggagaggta ttaaaatctt aataggtttt	1140
gataaaagcg aacgtgggga aaccgaacc aaaccttctt ctaaactctc tctcatctct	1200
cttaaagcaa acttctctct tgtctttctt gcgtgagcga ttttcaacgt tgtcagatcg	1260
tgcttcggca ccagtacaac gttttctttc actgaagcga aatcaaagat ctctttgtgg	1320
acacgtagtg cggcgccatt aaataacgtg tacttgtcct attcttgtcg gtgtgggtctt	1380
gggaaaagaa agcttgctgg aggtgctgtt tcagcccat acattacttg ttacgattct	1440
gctgactttc ggcggtgca atatctctac ttctgcttga cgaggtattg ttgcctgtac	1500

ttctttcttc ttcttcttgc tgattgggtc tataagaaat ctagtatttt ctttgaaca	1560
gagttttccc gtggttttcg aacttgaga aagattgtta agcttctgta tattctgccc	1620
aaatttgctg ggcccatggc gaaaaacgtt gcgattttcg gcttattgtt ttctcttctt	1680
gtgttggttc cttctcagat cttcgccgat cgaatctgca ctggaataac atcgtcaaac	1740
tcaccacatg tcgtcaaaac tgctactcaa ggggaggtca atgtgactgg tgtaatacca	1800
ctgacaacaa caccaccaa atctcathtt gcaaatctca aaggaacaga aaccaggggg	1860
aaactatgcc caaaatgcct caactgcaca gatctggacg tagccttggg cagacaaaaa	1920
tgcacgggga aaataccctc ggcaagagtt tcaatactcc atgaagtcag acctgttaca	1980
tctgggtgct ttcctataat gcacgacaga acaaaaatta gacagctgcc taaccttctc	2040
cgaggatacg aacatatcag gttatcaacc cataacgtta tcaatgcaga aaatgcacca	2100
ggaggaccct acaaaattgg aacctcaggg tcttgcccta acattaccaa tggaaacgga	2160
tttttcgcaa caatggcttg ggccgtccca aaaaacgaca aaaacaaaac agcaacaaat	2220
ccattaacaa tagaagtacc atacatttgt acagaaggag aagaccaaat taccgtttgg	2280
gggttccact ctgacaacga gacccaaatg gcaaagctct atggggactc aaagccccag	2340
aagttcacct catctgccaa cggagtgacc acacattacg tttcacagat tgggtggcttc	2400
ccaaatcaaa cagaagacgg aggactacca caaagtggta gaattgttgt tgattacatg	2460
gtgcaaaaat ctgggaaaac aggaacaatt acctatcaaa ggggtatttt attgcctcaa	2520
aaggtgtggt gcgcaagtgg caggagcaag gtaataaaag gatccttgcc ttttaattgga	2580
gaagcagatt gcctccacga aaaatacggg ggattaaaca aaagcaagcc ttactacaca	2640
ggggaacatg caaaggccat aggaaattgc ccaatatggg tgaacacacc cttgaagctg	2700
gccaatggaa ccaaatatag acctcctgca aaactattaa aggaaagggg tttcttcgga	2760
gctattgctg gtttcttaga aggaggatgg gaaggaatga ttgcaggttg gcacggatac	2820
acatcccatg gggcacatgg agtagcggg gcagcagacc ttaagagcac tcaagaggcc	2880
ataaacaaga taacaaaaaa tctcaactct ttgagtgagc tggaagtaaa gaatcttcaa	2940
agactaagcg gtgccatgga tgaactccac aacgaaatac tagaactaga tgagaaagtg	3000
gatgatctca gagctgatac aataagctca caaatagaac tcgcagtcct gctttccaat	3060
gaaggaataa taaacagtga agatgaacat ctcttgccgc ttgaaagaaa gctgaagaaa	3120
atgctggggc cctctgctgt agagataggg aatggatgct ttgaaaccaa acacaagtgc	3180
aaccagacct gtctcgacag aatagctgct ggtacctttg atgcaggaga attttctctc	3240
cccacctttg attcaactgaa tattactgct gcactcttaa atgacgatgg attggataat	3300

taccaaatac tgtcaattta ttcaacagtg gcgagttccc tagcactggc aatcatgatg	3360
gctggtctat ctttatggat gtgctccaat ggatcgttac aatgcagaat ttgcatttaa	3420
aggcctatTT tctttagttt gaatttactg ttattcgggtg tgcatttcta tgtttggtga	3480
gcggTTTTct gtgctcagag tgtgtttatt ttatgtaatt taatttcttt gtgagctcct	3540
gtttagcagg tCGTcccttc agcaaggaca caaaaagatt ttaattttat taaaaaaaa	3600
aaaaaaaaag accgggaatt cgatatcaag cttatcgacc tgcagatcgt tcaaacattt	3660
ggcaataaag tttcttaaga ttgaatcctg ttgccgggtct tgcgatgatt atcatataat	3720
ttctgttgaa ttacgttaag catgtaataa ttaacatgta atgcatgacg ttattttatga	3780
gatgggtttt tatgattaga gtcccgcaat tatacattta atacgcgata gaaaacaaaa	3840
tatagcgcgc aaactaggat aaattatcgc gcgcgggtgtc atctatgtta ctagatctct	3900
agagtctcaa gcttggcgcg gggtagcgag ctCGaattcc gagtgtactt caagtcagtt	3960
ggaaatcaat aaaatgatta ttttatgaat atatttcatt gtgcaagtag atagaaatta	4020
catatgttac ataacacacg aaataaaca aaaaacacaa tccaaaaca acaccccaaa	4080
caaaataaca ctatatatat cctcgtatga ggagaggcac gttcagtgac tgcagattc	4140
ccgagcaaaa aaagtctccc cgtcacacat atagtgggtg acgcaattat cttcaaagta	4200
atccttctgt tgacttgtca ttgataacat ccagtcttcg tcaggattgc aaagaattat	4260
agaagggatc ccacctttta ttttcttctt ttttccatat ttagggttga cagtgaatc	4320
agactggcaa cctattaatt gcttcacaa tgggacgaac ttgaagggga tgcgtcgat	4380
gatattatag gtggcgtgtt catcgtagtt ggtgaagtcg atggtcccgt tccagtagtt	4440
gtgtcgcccg agacttctag cccagggtgt ctttccggtg cgagttggtc cgcagatgta	4500
gaggctgggg tgtctgacct cagtccttcc ctcatcctgg ttagatcggc catccactca	4560
aggtcagatt gtgcttgatc gtaggagaca ggatgtatga aagtgtaggc atcgatgctt	4620
acatgatata ggtgcgtctc tctccagttg tgcagatctt cgtggcagcg gagatctgat	4680
tctgtgaagg gcgacacgta ctgctcaggt tgtggaggaa ataatttggt ggctgaatat	4740
tccagccatt gaagctttgt tgcccattca tgagggaatt cttctttgat catgtcaaga	4800
tactctcct tagacgttgc agtctggata atagttcgcc atcgtgcgtc agatttgca	4860
ggagagacct tatgatctcg gaaatctcct ctggttttaa tatctccgtc ctttgatatg	4920
taatcaagga cttgtttaga gtttctagct ggctggatat tagggtgatt tccttcaaaa	4980
tcgaaaaaag aaggatccct aatacaaggt tttttatcaa gctggataag agcatgatag	5040
tgggtagtgc catcttgatg aagctcagaa gcaacaccaa ggaagaaaat aagaaaaggt	5100
gtgagtttct cccagagaaa ctggaataaa tcatctcttt gagatgagca cttggggtag	5160

gtaaggaaaa catatattaga ttggagctctg aagttcttgc tagcagaagg catgttggtg 5220
 tgactccgag gggttgcctc aaactctatc ttataaccgg cgtggaggca tggaggcaag 5280
 ggcatttttg taatttaagt agttagtga aaatgacgtc atttacttaa agacgaagtc 5340
 ttgcgacaag gggggcccac gccgaatttt aatattaccg gcgtggcccc accttatcgc 5400
 gagtgcttta gcacgagcgg tccagattta aagtagaaaa gttcccggcc actagggtta 5460
 aaggtgttca cactataaaa gcatatacga tgtgatggta tttgatggag cgtatatattgt 5520
 atcagggtatt tccgtcggat acgaattatt cgtac 5555

<210> 37
 <211> 594
 <212> PRT
 <213> Virut cúm

<400> 37

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
 1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Arg Ile Cys Thr Gly Ile Thr
 20 25 30

Ser Ser Asn Ser Pro His Val Val Lys Thr Ala Thr Gln Gly Glu Val
 35 40 45

Asn Val Thr Gly Val Ile Pro Leu Thr Thr Thr Pro Thr Lys Ser His
 50 55 60

Phe Ala Asn Leu Lys Gly Thr Glu Thr Arg Gly Lys Leu Cys Pro Lys
 65 70 75 80

Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Lys Cys
 85 90 95

Thr Gly Lys Ile Pro Ser Ala Arg Val Ser Ile Leu His Glu Val Arg
 100 105 110

Pro Val Thr Ser Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile
 115 120 125

Arg Gln Leu Pro Asn Leu Leu Arg Gly Tyr Glu His Ile Arg Leu Ser
 130 135 140

Thr His Asn Val Ile Asn Ala Glu Asn Ala Pro Gly Gly Pro Tyr Lys
 145 150 155 160

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

His Gly Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala Ile
405 410 415

Asn Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val Lys
420 425 430

Asn Leu Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu Ile
435 440 445

Leu Glu Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile Ser
450 455 460

Ser Gln Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile Asn
465 470 475 480

Ser Glu Asp Glu His Leu Leu Ala Leu Glu Arg Lys Leu Lys Lys Met
485 490 495

Leu Gly Pro Ser Ala Val Glu Ile Gly Asn Gly Cys Phe Glu Thr Lys
500 505 510

His Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala Ala Gly Thr Phe
515 520 525

Asp Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser Leu Asn Ile Thr
530 535 540

Ala Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn Tyr Gln Ile Leu Ser
545 550 555 560

Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu Ala Ile Met Met Ala
565 570 575

Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile
580 585 590

Cys Ile

<210> 38

<211> 52

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> đoạn mỗi 1039+1059.r

<400> 38
 cttcccatcc tccaccagga ggtctatatatt tgggttcatt ggccagcttc aa 52

<210> 39
 <211> 52
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> đoạn mỗi 1039+1059.c

<400> 39
 caaatataga cctcctggtg gaggatggga aggaatgatt gcaggttggc ac 52

<210> 40
 <211> 5504
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Cát xét biểu hiện số 1059 từ BeYDV trái LIR sang BeYDV
 phải LIR. PDISP/HA từ cúm B/Brisbane/60/2008 với vòng phân hủy
 protein bị mất

<400> 40
 ctagcagaag gcatgttggt gtgactccga ggggttgctt caaactctat cttataaccg 60
 gcgtggagggc atggaggcaa gggcattttg gtaatttaag tagttagtgg aaaatgacgt 120
 catttactta aagacgaagt cttgcgacaa gggggggccca cgccgaattt taatattacc 180
 ggcgtagggc caccttatcg cgagtgttt agcacgagcg gtccagattt aaagtagaaa 240
 agttcccgcc cactaggggt aaaggtgttc aactataaa agcatatacg atgtgatggt 300
 atttgataaa gcgtatatgt tatcaggtat ttccgtcgga tacgaattat tcgtacaagc 360
 ttcttaagcc ggtcaacatg gtggagcacg acacacttgt ctactccaaa aatatcaaag 420
 atacagtctc agaagaccaa agggcaattg agacttttca acaaagggtg atatccggaa 480
 acctcctcgg attccattgc ccagctatct gtcaactttat tgtgaagata gtggaaaagg 540
 aaggtggctc ctacaaatgc catcattgcg ataaaggaaa ggccatcgtt gaagatgcct 600
 ctgccgacag tgggtcccaa gatggacccc caccacgag gagcatcgtg gaaaaagaag 660
 acgttccaac cacgtcttca aagcaagtgg attgatgtga taacatggtg gagcacgaca 720
 cacttgtcta ctcaaaaat atcaaagata cagtctcaga agaccaaagg gcaattgaga 780
 cttttcaaca aagggttaata tccggaaacc tctcggatt ccattgcca gctatctgtc 840
 actttattgt gaagatagtg gaaaaggaag gtggctccta caaatgccat cattgcgata 900
 aaggaaaggc catcgttgaa gatgcctctg ccgacagtgg tcccaaagat ggacccccac 960
 ccacgaggag catcgtggaa aaagaagacg ttccaaccac gtcttcaaag caagtggatt 1020

gatgtgatat ctccactgac gtaagggatg acgcacaatc ccactatcct tcgcaagacc	1080
cttctcttat ataaggaagt tcatttcatt tggagaggta ttaaaatctt aataggtttt	1140
gataaaagcg aacgtgggga aaccgaacc aaaccttctt ctaaactctc tctcatctct	1200
cttaaagcaa acttctctct tgtctttctt gcgtgagcga tcttcaacgt tgtcagatcg	1260
tgcttcggca ccagtacaac gttttctttc actgaagcga aatcaaagat ctctttgtgg	1320
acacgtagtg cggcgccatt aaataacgtg tacttgctct attcttgctg gtgtggtctt	1380
gggaaaagaa agcttgctgg aggctgctgt tcagcccat acattacttg ttacgattct	1440
gctgactttc ggcggtgca atatctctac ttctgcttga cgaggatttg ttgcctgtac	1500
ttctttcttc ttcttcttgc tgattggttc tataagaaat ctagtatttt ctttgaaaca	1560
gagttttccc gtggttttcg aacttggaaga aagattgtta agcttctgta tattctgccc	1620
aaatttgctg ggcccatggc gaaaaacgtt gcgattttcg gcttattgtt ttctcttctt	1680
gtgttggttc cttctcagat cttcgccgat cgaatctgca ctggaataac atcgtaaac	1740
tcaccacatg tcgtcaaac tgctactcaa ggggaggtca atgtgactgg tgtaatacca	1800
ctgacaacaa caccaccaa atctcatttt gcaaatctca aaggaacaga aaccagggg	1860
aaactatgcc caaaatgcct caactgcaca gatctggacg tagccttggg cagacaaaa	1920
tgcacgggga aaataccctc ggcaagagtt tcaatactcc atgaagtcag acctgttaca	1980
tctgggtgct ttctataat gcacgacaga aaaaaatta gacagctgcc taaccttctc	2040
cgaggatacg aacatatcag gttatcaacc cataacgtta tcaatgcaga aaatgcacca	2100
ggaggaccct acaaaattgg aacctcaggg tcttgcccta acattaccaa tggaaacgga	2160
tttttcgcaa caatggcttg ggccgtccca aaaaacgaca aaaacaaaac agcaacaaat	2220
ccattaacaa tagaagtacc atacatttgt acagaaggag aagaccaaat taccgtttgg	2280
gggttcact ctgacaacga gacccaaag gcaaagctct atggggactc aaagccccag	2340
aagttcacct catctgcaa cggagtgacc acacattacg ttacacagat tgggtggcttc	2400
ccaaatcaaa cagaagacgg aggactacca caaagtggta gaattgttgt tgattacatg	2460
gtgcaaaaat ctgggaaaac aggaacaatt acctatcaaa ggggtatttt attgcctcaa	2520
aaggtgtggt gcgcaagtgg caggagcaag gtaataaaag gatccttgcc tttaattgga	2580
gaagcagatt gcctccacga aaaatacggg ggattaaaca aaagcaagcc ttactacaca	2640
ggggaacatg caaaggccat aggaaattgc ccaatatggg tgaacacacc cttgaagctg	2700
gccaatggaa ccaaatatag acctcctggt ggaggatggg aaggaatgat tgcaggttgg	2760
cacggataca catcccatgg ggcacatgga gtagcgggtg cagcagacct taagagcact	2820
caagaggcca taaacaagat acaaaaaaat ctcaactctt tgagtgaact ggaagtaaag	2880

aatcttcaaa gactaagcgg tgccatggat gaactccaca acgaaatact agaactagat	2940
gagaaagtgg atgatctcag agctgataca ataagctcac aaatagaact cgcagtcctg	3000
ctttccaatg aaggaataat aaacagtga gatgaacatc tcttggcgct tgaaagaaag	3060
ctgaagaaaa tgctgggccc ctctgctgta gagataggga atggatgctt tgaaaccaa	3120
cacaagtgca accagacctg tctcgacaga atagctgctg gtacctttga tgcaggagaa	3180
ttttctctcc ccacctttga ttcactgaat attactgctg catctttaa tgacgatgga	3240
ttggataatc atactatact gctttactac tcaactgctg cctccagttt ggctgtaaca	3300
ctgatgatag ctatctttgt tgtttatatg gtctccagag acaatgtttc ttgctccatc	3360
tgtctataaa ggcctatttt ctttagtttg aatttactgt tattcgggtg gcatttctat	3420
gtttggtgag cggttttctg tgctcagagt gtgtttattt tatgtaattt aatttctttg	3480
tgagctcctg tttagcaggt cgtcccttca gcaaggacac aaaaagattt taattttatt	3540
aaaaaaaaa aaaaaaaga ccgggaattc gatatcaagc ttatcgacct gcagatcggt	3600
caaacatttg gcaataaagt ttcttaagat tgaatcctgt tgccgggtctt gcgatgatta	3660
tcatataatt tctgttgaat tacgttaagc atgtaataat taacatgtaa tgcatacgt	3720
tatttatgag atgggttttt atgattagag tcccgcaatt atacatttaa tacgcgatag	3780
aaaacaaaat atagcgcgca aactaggata aattatcgcg cgcggtgtca tctatgttac	3840
tagatctcta gagtctcaag cttggcgcg ggtaccgagc tcgaattccg agtgtacttc	3900
aagtcagttg gaaatcaata aaatgattat tttatgaata tatttcattg tgcaagtaga	3960
tagaaattac atatgttaca taacacacga aataaacaaa aaaacacaat ccaaaacaaa	4020
caccccaaac aaaataacac tatatatatc ctcgtagag gagaggcacg ttcagtgact	4080
cgacgattcc cgagcaaaaa aagtctcccc gtcacacata tagtgggtga cgcaattatc	4140
ttcaaagtaa tccttctggt gacttgtcat tgataacatc cagtcttcgt caggattgca	4200
aagaattata gaagggatcc caccttttat tttcttcttt tttccatatt taggggtgac	4260
agtgaaatca gactggcaac ctattaattg cttccacaat gggacgaact tgaaggggat	4320
gtcgtcgatg atattatagg tggcgtgttc atcgtagttg gtgaagtcga tgggtccggt	4380
ccagtagttg tgtcgcccga gacttctagc ccaggtggtc tttccggtac gagttggtcc	4440
gcagatgtag aggctggggt gtctgacccc agtccttccc tcatcctggt tagatcggcc	4500
atccactcaa ggtcagattg tgcttgatcg taggagacag gatgtatgaa agtgtaggca	4560
tcgatgctta catgatatag gtgcgtctct ctccagttgt gcagatcttc gtggcagcgg	4620
agatctgatt ctgtgaaggg cgacacgtac tgctcaggtt gtggaggaaa taatttggtg	4680

gctgaatatt ccagccattg aagctttgtt gccattcat gagggaattc ttctttgatc 4740
 atgtcaagat actcctcctt agacgttgca gtctggataa tagttcgcca tcgtgcgta 4800
 gatttgcgag gagagacctt atgatctcgg aaatctcctc tggttttaat atctccgtcc 4860
 tttgatatgt aatcaaggac ttgttttagag tttctagctg gctggatatt agggtgattt 4920
 ccttcaaaat cgaaaaaaga aggatcccta atacaagggt ttttatcaag ctggataaga 4980
 gcatgatagt gggtagtgcc atcttgatga agctcagaag caacaccaag gaagaaaata 5040
 agaaaagggtg tgagtttctc ccagagaaaac tggaataaat catctctttg agatgagcac 5100
 ttggggtagg taaggaaaac atatttagat tggagtctga agttcttgct agcagaaggc 5160
 atgttgttgt gactccgagg ggttgccctca aactctatct tataaccggc gtggaggcat 5220
 ggaggcaagg gcatttttgt aatttaagta gttagtggaa aatgacgtca tttacttaaa 5280
 gacgaagtct tgcgacaagg ggggccacg ccgaatttta atattaccgg cgtggcccca 5340
 ccttatcgcg agtgcttttag cagcagcggc ccagatttaa agtagaaaag ttcccgccca 5400
 ctaggggttaa aggtgttcac actataaaag catatacgat gtgatggtat ttgatggagc 5460
 gtatattgta tcaggatatt ccgtcggata cgaattattc gtac 5504

<210> 41
 <211> 577
 <212> PRT
 <213> Virut cúm

<400> 41

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
 1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Arg Ile Cys Thr Gly Ile Thr
 20 25 30

Ser Ser Asn Ser Pro His Val Val Lys Thr Ala Thr Gln Gly Glu Val
 35 40 45

Asn Val Thr Gly Val Ile Pro Leu Thr Thr Thr Pro Thr Lys Ser His
 50 55 60

Phe Ala Asn Leu Lys Gly Thr Glu Thr Arg Gly Lys Leu Cys Pro Lys
 65 70 75 80

Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Lys Cys
 85 90 95

Thr Gly Lys Ile Pro Ser Ala Arg Val Ser Ile Leu His Glu Val Arg

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

Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Gly Gly Gly Trp
 355 360 365

Glu Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His Gly Ala His
 370 375 380

Gly Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala Ile Asn
 385 390 395 400

Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val Lys Asn
 405 410 415

Leu Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu Ile Leu
 420 425 430

Glu Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile Ser Ser
 435 440 445

Gln Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile Asn Ser
 450 455 460

Glu Asp Glu His Leu Leu Ala Leu Glu Arg Lys Leu Lys Lys Met Leu
 465 470 475 480

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

Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala Ala Gly Thr Phe Asp
 500 505 510

Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser Leu Asn Ile Thr Ala
 515 520 525

Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn His Thr Ile Leu Leu Tyr
 530 535 540

Tyr Ser Thr Ala Ala Ser Ser Leu Ala Val Thr Leu Met Ile Ala Ile
 545 550 555 560

Phe Val Val Tyr Met Val Ser Arg Asp Asn Val Ser Cys Ser Ile Cys
 565 570 575

Leu

<210> 42
 <211> 1707
 <212> ADN
 <213> Virut cúm

<400> 42
 atggagaaaa tagtgcttct tcttgcaata gtcagtcttg ttaaaagtga tcagatttgc 60
 attggttacc atgcaaaca ttcaacagag cagggttgaca caatcatgga aaagaacggt 120
 actgttacac atgcccaga catactggaa aagacacaca acgggaagct ctgcatctca 180
 gatggagtga agcctctaatt ttaagagat tgtagtgtag ctggatggct cctcgggaac 240
 ccaatgtgtg acgaattcat caatgtaccg gaatgggtctt acatagtgga gaaggccaat 300
 ccaaccaatg acctctgtta ccaggagggt ttcaacgact atgaagaact gaaacaccta 360
 ttgagcagaa taaaccattt tgagaaaatt caaatcatcc caaaagttc ttggtccgat 420
 catgaagcct catcaggagt tagctcagca tgtccatacc tgggaagtcc ctcttttttt 480
 agaaatgtgg tatggcttat caaaaagaac agtacatacc caacaataaa gaaaagctac 540
 aataatacca accaagagga tcttttggtg ctgtggggaa ttcacatcc taatgatgag 600
 gcagagcaga caaggctata tcaaaaccca accacctata tttccattgg gacatcaaca 660
 ctaaaccaga gattggtacc aaaaatagct actagatcca aagtaaacgg gcaaagtgga 720
 aggatggagt tcttctggac aattttaaaa cctaattgatg caatcaactt cgagagtaat 780
 ggaaatttca ttgctccaga atatgcatac aaaattgtca agaaagggga ctgagcaatt 840
 atgaaaagtg aattggaata tggtaactgc aacaccaagt gtcaaactcc aatgggggag 900
 ataaactcta gtatgccatt ccacaacata caccctctca ccatcgggga atgccccaaa 960
 tatgtgaaat caaacagatt agtccttgca acagggtca gaaatagccc tcaaagagag 1020
 agcagaagaa aaaagagagg actatttgga gctatagcag gttttataga gggaggatgg 1080
 cagggaatgg tagatgggtg gtatgggtac caccatagca atgagcaggg gagtgggtac 1140
 gctgcagaca aagaatccac tcaaaaggca atagatggag tcaccaataa ggtcaactca 1200
 atcattgaca aaatgaacac tcagtttgag gccgttgga gggaatttaa taacttagaa 1260
 aggagaatag agaattttaa caagaagatg gaagacgggt ttctagatgt ctggacttat 1320
 aatgccgaac ttctggttct catggaaaat gagagaactc tagactttca tgactcaaat 1380
 gttagaacc tctacgaca ggtccgacta cagcttaggg ataatgcaa ggagctgggt 1440
 aacggttggt tcgagttcta tcacaaatgt gataatgaat gtatggaaag tataagaaac 1500
 ggaacgtaca actatccgca gtattcagaa gaagcaagat taaaagaga ggaaataagt 1560
 ggggtaaaat tggaatcaat aggaacttac caaatactgt caatttatcc aacagtggcg 1620

agttccctag cactggcaat catgatggct ggtctatctt tatggatgtg ctccaatgga 1680
 tcgttacaat gcagaatttg catttaa 1707

 <210> 43
 <211> 1734
 <212> ADN
 <213> Virut cúm

 <400> 43
 atggcgaaaa acgttgcgat tttcggtta ttgttttctc ttcttgtgtt ggttccttct 60
 cagatcttcg ccgatcgaat ctgcactgga ataacatcgt caaactcacc acatgtcgtc 120
 aaaactgcta ctcaagggga ggtcaatgtg actggtgtaa taccactgac aacaacaccc 180
 accaaatctc attttgcaaa tctcaaagga acagaaacca gggggaaact atgccccaaa 240
 tgctcaact gcacagatct ggacgtagcc ttgggcagac caaatgcac ggggaaaata 300
 ccctcggcaa gagtttcaat actccatgaa gtcagacctg ttacatctgg gtgctttcct 360
 ataatgcacg acagaacaaa aattagacag ctgcctaacc ttctccgagg atacgaacat 420
 atcaggttat caaccataa cgttatcaat gcagaaaatg caccaggagg accctacaaa 480
 attggaacct caggggtcttg ccctaacatt accaatggaa acggattttt cgcaacaatg 540
 gcttggggccg tccccaaaaa cgacaaaaac aaaacagcaa caaatccatt aacaatagaa 600
 gtaccataca tttgtacaga aggagaagac caaattaccg tttggggggt cactctgac 660
 aacgagaccc aaatggcaaa gctctatggg gactcaaagc cccagaagtt cacctcatct 720
 gccaacggag tgaccacaca ttacgtttca cagattggtg gcttcccaa tcaaacagaa 780
 gacggaggac taccacaaag tggtagaatt gttgttgatt acatggtgca aaaatctggg 840
 aaaacaggaa caattacct tcaaaggggt attttattgc ctcaaagggt gtggtgcgca 900
 agtggcagga gcaaggtaat aaaaggatcc ttgcctttaa ttggagaagc agattgcctc 960
 cacgaaaaat acggtggatt aaacaaaagc aagccttact acacagggga acatgcaaag 1020
 gccataggaa attgccaat atgggtgaaa acacccttga agctggccaa tggaaccaa 1080
 tatagacctc ctggtggagg atgggaagga atgattgcag gttggcacgg atacacatcc 1140
 catggggcac atggagtagc ggtggcagca gaccttaaga gactcaaga ggccataaac 1200
 aagataacaa aaaatctcaa ctctttgagt gagctggaag taaagaatct tcaaagacta 1260
 agcggtgcca tggatgaact ccacaacgaa atactagaac tagatgagaa agtggatgat 1320
 ctgagagctg atacaataag ctacaaata gaactcgcag tcttgctttc caatgaagga 1380
 ataataaaca gtgaagatga acatctcttg gcgcttgaaa gaaagctgaa gaaatgctg 1440
 ggcccctctg ctgtagagat agggaatgga tgctttgaaa ccaaacacaa gtgcaaccag 1500

acctgtctcg acagaatagc tgctggtacc ttgatgcag gagaattttc tctccccacc 1560
 ttgattcac tgaatattac tgctgcatct ttaaatgacg atggattgga taatcatact 1620
 atactgcttt actactcaac tgctgcctcc agtttggtg taacactgat gatagctatc 1680
 tttgttggtt atatggtctc cagagacaat gtttcttgct ccactgtct ataa 1734

<210> 44
 <211> 48
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> đoạn mỗi IF-H3V36111.S2+4c

<400> 44
 tctcagatct tcgccccaaa acttcctgga aatgacaaca gcacggca 48

<210> 45
 <211> 51
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> đoạn mỗi IF-H3V36111.s1-4r

<400> 45
 actaaagaaa ataggccttc aaatgcaa atgtgcaccta atgttgccct t 51

<210> 46
 <211> 1701
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> trình tự nucleotit của gen H3 tổng hợp

<400> 46
 atgaagacta tcattgcttt gagccacatt ctatgtctgg ttttcgctca aaaacttcct 60
 ggaaatgaca acagcacggc aacgctgtgc cttgggcacc atgcagtacc aaacggaacg 120
 atagtgaata caatcacgaa tgaccaaatt gaagttacta atgctactga gctgggttcag 180
 aattcctcaa taggtgaaat atgcgacagt cctcatcaga tccttgatgg agaaaactgc 240
 acactaatag atgctctatt gggagaccct cagtgtgatg gcttccaaaa taagaaatgg 300
 gacctttttg ttgaacgaag caaagcctac agcaactgtt acccttatga tgtgccggat 360
 tatgcctccc ttaggtcact agttgcctca tccggcacac tggagtttaa caatgaaagc 420
 ttcaattgga ctggagtcac tcaaaacgga acaagttctg cttgcataag gagatctaata 480
 aatagtttct ttagtagatt aaattgggtg acccacttaa acttcaaata cccagcattg 540
 aacgtgacta tgccaaacaa tgaacaattt gacaaattgt acatttgggg gggttcaccac 600

ccgggtacgg acaaggacca aatcttctctg tatgctcaat catcaggaag aatcacagta 660
 tctacaaaaa gaagccaaca agctgtaatc ccgaatatcg gatctagacc cagaataagg 720
 aatatcccta gcagaataag catctattgg acaatagtaa aaccgggaga catacttttg 780
 attaacagca cagggaatct aattgctcct aggggttact tcaaaatacg aagtgggaaa 840
 agctcaataa tgagatcaga tgcaccatt ggcaaatgca attctgaatg catcactcca 900
 aatggaagca ttcccaatga caaacattc caaaatgtaa acaggatcac atacggggcc 960
 tgtcccagat atgttaagca aagcactctg aaattggcaa caggaatgcg aaatgtacca 1020
 gagaaacaaa ctagaggcat atttggcgca atagcgggtt tcatagaaaa tggttgggag 1080
 ggaatggtgg atggttggtg cggtttcagg catcaaaatt ctgaggggaag aggacaagca 1140
 gcagatctca aaagcactca agcagcaatc gatcaaatca atgggaagct gaatcgattg 1200
 atcgggaaaaa ccaacgagaa attccatcag attgaaaaag aattctcaga agtcgaaggg 1260
 agaattcagg accttgagaa atatgttgag gacactaaaa tagatctctg gtcatacaac 1320
 gcggagcttc ttgttgccct ggagaaccaa catacaattg atctaactga ctcagaaatg 1380
 aacaaactgt ttgaaaaaac aaagaagcaa ctaagggaaa atgctgagga tatgggcaat 1440
 ggttgtttca aaatatacca caaatgtgac aatgcctgca taggatcaat cagaaatgga 1500
 acttatgacc acgatgtata cagagatgaa gcattaaaca accggttcca gatcaaggga 1560
 gttgagctga agtcagggtg caaagattgg atcctatgga tttcctttgc catatcatgt 1620
 tttttgcttt gtgttgcttt gttgggggtc atcatgtggg cctgccaaaaa gggcaacatt 1680
 aggtgcaaca tttgcatttg a 1701

<210> 47

<211> 3465

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> cát xét biểu hiện số 1391

<400> 47

gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60
 gaagacaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120
 ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180
 tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240
 ggtcccaaag atggaccccc acccagcagg agcatcgtgg aaaaagaaga cgttccaacc 300
 acgtcttcaa agcaagtgga ttgatgtgat aacatgggtg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa	420
agggtaatat ccggaacact cctcggattc cattgcccag ctatctgtca ctttattgtg	480
aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc	540
atcgttgaag atgcctctgc cgacagtggc cccaaagatg gacccccacc cacgaggagc	600
atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc	660
tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagacct ttcctctata	720
taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga	780
acgtggggaa acccgaacca aaccttcttc taaactctct ctcactcttc ttaaagcaaa	840
cttctctctt gtctttcttg cgtgagcgat cttcaacggt gtcagatcgt gcttcggcac	900
cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc	960
ggcgccatta aataacgtgt acttgctcta ttcttgctgg tgtggctctg ggaaaagaaa	1020
gcttgctgga ggctgctggt cagccccata cattacttgt tacgattctg ctgactttcg	1080
gcggtgcaa tatctctact tctgcttgac gaggtattgt tgctgtact tctttcttct	1140
tcttcttgct gattggttct ataagaaatc tagtatcttc tttgaaacag agttttcccg	1200
tggttttcga acttgagaaa agattgttaa gcttctgtat attctgcca aatttgctcg	1260
gcccatggcg aaaaacgttg cgattttcgg cttattgttt tctcttcttg tgttggttcc	1320
ttctcagatc ttcgccccaa aacttcctgg aaatgacaac agcacggcaa cgctgtgcct	1380
tgggcacat gtagtaccac acggaacgat agtgaaaaca atcacgaatg accaaattga	1440
agttactaat gctactgagc tgggttcagaa ttcctcaata ggtgaaatat gcgacagtcc	1500
tcatcagatc cttgatggag aaaactgcac actaatagat gctctattgg gagaccctca	1560
gtgtgatggc ttccaaaata agaaatggga cttttttgtt gaacgaagca aagcctacag	1620
caactgttac cttatgatg tgccggatta tgctccctt aggtcactag ttgcctcatc	1680
cggcacactg gaggtttaaca atgaaagctt caattggact ggagtcactc aaaacggaac	1740
aagttctgct tgcataagga gatctaataa tagtttcttt agtagattaa attggttgac	1800
ccacttaaac ttcaaatacc cagcattgaa cgtgactatg ccaaacaatg aacaatttga	1860
caaattgtac atttgggggg ttcaccaccc gggtagcgac aaggaccaa tcttcctgta	1920
tgctcaatca tcaggaagaa tcacagtatc taccaaaaga agccaacaag ctgtaatccc	1980
gaatatcgga tctagacca gaataaggaa tatccctagc agaataagca tctattggac	2040
aatagtaaaa ccgggagaca tacttttgat taacagcaca gggaatctaa ttgctcctag	2100
gggttacttc aaaatacgaa gtgggaaaag ctcaataatg agatcagatg caccatttgg	2160
caaatgcaat tctgaatgca tcaactccaa tggaagcatt cccaatgaca aaccattcca	2220

```

aatgtaaac aggatcacat acggggcctg tcccagatat gttaagcaaa gcactctgaa 2280
attggcaaca ggaatgcgaa atgtaccaga gaaacaaact agaggcatat ttggcgcaat 2340
agcgggtttc atagaaaatg gttgggaggg aatggtggat ggttggtacg gtttcaggca 2400
tcaaaattct gagggagag gacaagcagc agatctcaaa agcactcaag cagcaatcga 2460
tcaaatcaat gggaagctga atcgattgat cgggaaaacc aacgagaaat tocatcagat 2520
tgaaaaagaa ttctcagaag tcgaaggag aattcaggac cttgagaaat atgttgagga 2580
cactaaaata gatctctggt catacaacgc ggagcttctt gttgccctgg agaaccaaca 2640
tacaattgat ctaactgact cagaaatgaa caaactgttt gaaaaaaca agaagcaact 2700
aagggaaaat gctgaggata tgggcaatgg ttgtttcaaa atataccaca aatgtgacaa 2760
tgctgcata ggatcaatca gaaatggaac ttatgaccac gatgtataca gagatgaagc 2820
attaaacaac cggttccaga tcaaggagat tgagctgaag tcagggtaca aagattggat 2880
cctatggatt tcctttgcc aatcatgttt tttgctttgt gttgctttgt tggggttcat 2940
catgtgggcc tgccaaaagg gcaacattag gtgcaacatt tgcatttgaa ggcctatatt 3000
ctttagtttg aatttactgt tattcgggtg gcatttctat gtttggtgag cggttttctg 3060
tgctcagagt gtgtttatatt tatgtaattt aatttctttg tgagctcctg tttagcaggt 3120
cgtcccttca gcaaggacac aaaaagattt taattttatt aaaaaaaaaa aaaaaaaga 3180
ccgggaattc gatatcaagc ttatcgacct gcagatcggt caaacatttg gcaataaagt 3240
ttcttaagat tgaatcctgt tgccggtctt gcgatgatta tcatataatt tctgttgaat 3300
tacgttaagc atgtaataat taacatgtaa tgcagacgt tatttatgag atgggttttt 3360
atgattagag tcccgaatt atacatttaa tacgcgatag aaaacaaaat atagcgcgca 3420
aactaggata aattatcgcg cgcggtgtca tctatgttac tagat 3465

```

<210> 48
 <211> 574
 <212> PRT
 <213> Virut cúm

<400> 48

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
 1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Gln Lys Leu Pro Gly Asn Asp Asn
 20 25 30

Ser Thr Ala Thr Leu Cys Leu Gly His His Ala Val Pro Asn Gly Thr
 35 40 45

Ile Val Lys Thr Ile Thr Asn Asp Gln Ile Glu Val Thr Asn Ala Thr
 50 55 60

Glu Leu Val Gln Asn Ser Ser Ile Gly Glu Ile Cys Asp Ser Pro His
 65 70 75 80

Gln Ile Leu Asp Gly Glu Asn Cys Thr Leu Ile Asp Ala Leu Leu Gly
 85 90 95

Asp Pro Gln Cys Asp Gly Phe Gln Asn Lys Lys Trp Asp Leu Phe Val
 100 105 110

Glu Arg Ser Lys Ala Tyr Ser Asn Cys Tyr Pro Tyr Asp Val Pro Asp
 115 120 125

Tyr Ala Ser Leu Arg Ser Leu Val Ala Ser Ser Gly Thr Leu Glu Phe
 130 135 140

Asn Asn Glu Ser Phe Asn Trp Thr Gly Val Thr Gln Asn Gly Thr Ser
 145 150 155 160

Ser Ala Cys Ile Arg Arg Ser Asn Asn Ser Phe Phe Ser Arg Leu Asn
 165 170 175

Trp Leu Thr His Leu Asn Phe Lys Tyr Pro Ala Leu Asn Val Thr Met
 180 185 190

Pro Asn Asn Glu Gln Phe Asp Lys Leu Tyr Ile Trp Gly Val His His
 195 200 205

Pro Gly Thr Asp Lys Asp Gln Ile Phe Leu Tyr Ala Gln Ser Ser Gly
 210 215 220

Arg Ile Thr Val Ser Thr Lys Arg Ser Gln Gln Ala Val Ile Pro Asn
 225 230 235 240

Ile Gly Ser Arg Pro Arg Ile Arg Asn Ile Pro Ser Arg Ile Ser Ile
 245 250 255

Tyr Trp Thr Ile Val Lys Pro Gly Asp Ile Leu Leu Ile Asn Ser Thr
 260 265 270

Gly Asn Leu Ile Ala Pro Arg Gly Tyr Phe Lys Ile Arg Ser Gly Lys
 275 280 285

Ser Ser Ile Met Arg Ser Asp Ala Pro Ile Gly Lys Cys Asn Ser Glu
 290 295 300

Cys Ile Thr Pro Asn Gly Ser Ile Pro Asn Asp Lys Pro Phe Gln Asn
 305 310 315 320

Val Asn Arg Ile Thr Tyr Gly Ala Cys Pro Arg Tyr Val Lys Gln Ser
 325 330 335

Thr Leu Lys Leu Ala Thr Gly Met Arg Asn Val Pro Glu Lys Gln Thr
 340 345 350

Arg Gly Ile Phe Gly Ala Ile Ala Gly Phe Ile Glu Asn Gly Trp Glu
 355 360 365

Gly Met Val Asp Gly Trp Tyr Gly Phe Arg His Gln Asn Ser Glu Gly
 370 375 380

Arg Gly Gln Ala Ala Asp Leu Lys Ser Thr Gln Ala Ala Ile Asp Gln
 385 390 395 400

Ile Asn Gly Lys Leu Asn Arg Leu Ile Gly Lys Thr Asn Glu Lys Phe
 405 410 415

His Gln Ile Glu Lys Glu Phe Ser Glu Val Glu Gly Arg Ile Gln Asp
 420 425 430

Leu Glu Lys Tyr Val Glu Asp Thr Lys Ile Asp Leu Trp Ser Tyr Asn
 435 440 445

Ala Glu Leu Leu Val Ala Leu Glu Asn Gln His Thr Ile Asp Leu Thr
 450 455 460

Asp Ser Glu Met Asn Lys Leu Phe Glu Lys Thr Lys Lys Gln Leu Arg
 465 470 475 480

Glu Asn Ala Glu Asp Met Gly Asn Gly Cys Phe Lys Ile Tyr His Lys
 485 490 495

Cys Asp Asn Ala Cys Ile Gly Ser Ile Arg Asn Gly Thr Tyr Asp His
 500 505 510

Asp Val Tyr Arg Asp Glu Ala Leu Asn Asn Arg Phe Gln Ile Lys Gly
 515 520 525

Val Glu Leu Lys Ser Gly Tyr Lys Asp Trp Ile Leu Trp Ile Ser Phe
530 535 540

Ala Ile Ser Cys Phe Leu Leu Cys Val Ala Leu Leu Gly Phe Ile Met
545 550 555 560

Trp Ala Cys Gln Lys Gly Asn Ile Arg Cys Asn Ile Cys Ile
565 570

<210> 49
<211> 46
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> đoạn mỗi IF-HAB110.S1+3c

<400> 49
aaatttgtagc ggcccatgaa ggcaataatt gtactactca tggtag 46

<210> 50
<211> 52
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> đoạn mỗi IF-HAB110.s1-4r

<400> 50
actaaagaaa ataggccttt atagacagat ggagcatgaa acgttgcttc tg 52

<210> 51
<211> 1755
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> trình tự nucleotit của HA B/Wisconin (JN993010) tổng hợp

<400> 51
atgaaggcaa taattgtact actcatggta gtaacatcca atgcagatcg aatctgcact 60
gggataacat cttcaaactc acctcatgtg gtcaaaacag ctactcaagg ggaggtcaat 120
gtgactggcg tgataccact gacaacaaca ccaacaaaat cttattttgc aaatctcaaa 180
ggaacaagga ccagagggaa actatgcccg gactgtctca actgtacaga tctggatgtg 240
gccttgggca ggccaatgtg tgtggggacc acaccttctg cttaaagcttc aatactccac 300
gaggtcagac ctgttacatc cgggtgcttt cctataatgc acgacagaac aaaaatcagg 360
caactaccca atcttctcag aggatatgaa aatatcaggt tatcaacca aaacgttatc 420
gatgcagaaa aagcaccagg aggaccctac agacttgga cctcaggatc ttgccctaac 480

gctaccagta aaatcggtatt ttttgcaaca atggcttggg ctgtcccaaa ggacaactac 540
 aaaaatgcaa cgaaccact aacagtagaa gtaccataca tttgtacaga aggggaagac 600
 caaattactg tttgggggtt ccattcagat aacaaaaccc aaatgaagag cctctatgga 660
 gactcaaata ctcaaaagtt cacctcatct gctaattggag taaccacaca ttatgtttct 720
 cagattggcg acttcccaga tcaaacagaa gacggaggac taccacaaag cggcagaatt 780
 gttgttgatt acatgatgca aaaacctggg aaaacaggaa caattgtcta tcaaagaggt 840
 gttttgttgc ctcaaaaggt gtggtgcgcg agtggcagga gcaaagtaat aaaaggggtca 900
 ttgcctttaa ttggtgaagc agattgcctt catgaaaaat acggtggatt aaacaaaagc 960
 aagccttact acacaggaga acatgcaaaa gccataggaa attgcccaat atgggtaaaa 1020
 acacctttga agcttgccaa tggaaccaa tatagacctc ctgcaaaact attgaaggaa 1080
 aggggtttct tcggagctat tgctggtttc ctagaaggag gatgggaagg aatgattgca 1140
 ggttggcacg gatacacatc tcacggagca catggagtgg cagtggcggc agaccttaag 1200
 agtacacaag aagctataaa taagataaca aaaaatctca attctttgag tgagctagaa 1260
 gtaaagaacc ttcaagact aagtgggtgc atggatgaac tccacaacga aatactcgag 1320
 ctggatgaga aagtggatga tctcagagct gacactataa gctcaciaat agaacttgca 1380
 gtcttgcttt ccaacgaagg aataataaac agtgaagacg agcatctatt ggcacttgag 1440
 agaaaactaa agaaaatgct gggccctct gctgtagaca taggaaacgg atgcttcgaa 1500
 accaaacaca aatgcaacca gacctgtta gacaggatag ctgctggcac ctttaatgca 1560
 ggagaatttt ctctcccac ttttgattca ttgaacatta ctgctgcac tttaatgat 1620
 gatggattgg ataaccatac tatactgctc tattactcaa ctgctgcttc tagtttggt 1680
 gtaacattaa tgctagctat ttttattgtt tatatggtct ccagagacaa cgtttcatgc 1740
 tccatctgtc tataa 1755

<210> 52
 <211> 6745
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc 193

<400> 52
 tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 60
 gacgttttta atgtactgaa ttaacgccga atcccggtt ggtatatatta tatgttgtca 120
 aataactcaa aaaccataaa agtttaagtt agcaagtgtg tacattttta cttgaacaaa 180
 aatattcacc tactactgtt ataaatcatt attaaacatt agagtaaaga aatatggatg 240

ataagaacaa gagtagtgat attttgacaa caatthttgtt gcaacatttg agaaaatttt	300
gttgthttctct cthtttcattg gtcaaaaaca atagagagag aaaaagggaag agggagaata	360
aaaacataat gtgagtatga gagagaaaagt tgtacaaaag ttgtaccaaa atagttgtac	420
aaatatcatt gaggaatttg acaaaagcta cacaaataag ggttaattgc tgtaaaaaa	480
taaggatgac gcattagaga gatgtacat tagagaattt ttggcaagtc attaaaaaga	540
aagaataaat tathththaaa attaaaagtt gagtcatttg attaaacatg tgattattta	600
atgaattgat gaaagagttg gattaaagtt gtattagtaa ttagaatttg gtgtcaaatt	660
taatttgaca thttgatcttt tcttatatat tgcccatag agtcagttaa ctcatththta	720
taththcatag atcaaataag agaaataacg gtatattaat ccttccaaaa aaaaaaacg	780
gtataththac taaaaaatct aagccacgta ggaggataac aggatccccg taggaggata	840
acatccaatc caaccaatca caacaatcct gatgagataa cccactthta gccacgcat	900
ctgtggcaca tctacattat ctaaatcaca cattctthca cacatctgag ccacacaaaa	960
accaatccac atctthtatca cccattctat aaaaaatcac actthtgtag tctacactth	1020
gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag	1080
aaaatggaac gagctataca aggaaacgac gctagggaac aagctaacag tgaacgttg	1140
gatggaggat caggaggtac cactthctcc thcaaacttc ctgacgaaag tccgagttg	1200
actgagtggc ggctacataa cgatgagacg aattcgaatc aagataatcc ccttggtthc	1260
aaggaaagct ggggtthcgg gaaagttgta thtaagagat atctcagata cgacaggacg	1320
gaagcttcac tgcacagagt ccttgatct tggacgggag attcggthta ctatgcagca	1380
tctcgathth thcgththcga ccagatcgga tgtacctata gtattcggth thcaggagth	1440
agtatcaccg ththctggagg gtgcggaact cthcagcatc ththtgagat ggcaattcgg	1500
tctaagcaag aactgctaca gcttgccca atcgaagtgg aaagtaatgt atcaagagga	1560
tgccctgaag gtactcaaac cthcgaaaaa gaaagcgagt aagthtaaat gctththcgt	1620
ctctathth taatatggt thgtththgt aathththth ththtagaaga gcttaattaa	1680
thgtththth tatgaaatac tathththth agatgaactg gtgtaatgta atthththac	1740
ataagtggag tcagaatcag aatgththct ccataactaa ctagacatga agacctgccg	1800
cgtacaattg ththththth gaacaactaa aathgaacat cththththth aactththth	1860
gtggttaata tagctcaaat atatggtcaa gthththth thtaaatgg aaatatcagt	1920
tatcgaaatt cattaacaat caactthth ththththth ththththth atthththth	1980
ththththth atagththth aathththth gagththth gcctaggaga ththththth	2040

ccgccttcag tttgcaagct gctctagccg tgtagccaat acgcaaaccg cctctccccg	2100
cgcggttgga attactagcg cgtgtcgaga cgcggttggtg ttgtgactcc gaggggttgc	2160
ctcaaactct atcttataac cggcgtggag gcatggaggc aggggtattht ttgtcatttht	2220
aatagatagt ggaaaatgac gtggaattta cttaaagacg aagtctthtgc gacaaggggg	2280
ggcccacgcc gaatttaata ttaccggcgt ggccccccct tatcgcgagt gctthtagcac	2340
gagcgggtcca gattthaaagt agaaaatttc ccgcccacta gggttaaagg tgttcacact	2400
ataaaagcat atacgatgtg atggtattht gtcgacaagc ttgcatgccg gtcaacatgg	2460
tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca gaagacaaaa	2520
gggcaattga gactthttcaa caaagggtaa tatccggaaa cctcctcgga ttccattgcc	2580
cagctatctg tcactthtatt gtgaagatag tggaaaagga aggtggctcc taaaaatgcc	2640
atcattgcga taaaggaaaag gccatcgttg aagatgcctc tgccgacagt ggtcccaaag	2700
atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgthccaacc acgtcttcaa	2760
agcaagtgga ttgatgtgat aacatgggtg agcacgacac acttgctctac tccaaaaata	2820
tcaaagatac agtctcagaa gaccaaaagg caattgagac thttcaacaa agggtaatat	2880
ccggaaacct cctcggattc cattgcccag ctatctgtca thttattgtg aagatagtgg	2940
aaaaggaaag tggctcctac aaatgccatc attgcgataa aggaaaggcc atcgttgaag	3000
atgcctctgc cgacagtggc cccaaagatg gacccccacc cacgaggagc atcgtggaaa	3060
aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc tccactgacg	3120
taagggatga cgcacaatcc cactatcctt cgcaagacct thctctata taaggaaagt	3180
catttcattt ggagaggtat taaaatctta ataggthttg ataaaagcga acgtggggaa	3240
acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa cttctctctt	3300
gtctthtctg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac cagtacaacg	3360
thttctthtca ctgaagcgaa atcaaagatc thttgtgga cacgtagtgc ggccgcat	3420
aataacgtgt acttgctcta thttgtcgg tgtggtcttg ggaaaagaaa gcttgctgga	3480
ggctgctgtt cagccccata cattacttgt tacgattctg ctgactthtcg gcgggtgcaa	3540
tatctctact tctgcttgac gaggtattgt tgctgtact thttcttht tcttcttgc	3600
gattggttct ataagaaatc tagtatthtc thtgaaacag agthttcccg tggthttcga	3660
acttgagaaa agattgttaa gcttctgtat attctgcca aatttgctcg gccgcggat	3720
ggcgaaaaac gttgcgattt tcggcttatt gthttctctt cttgtgttg thcttctca	3780
gatcttcgcc tgcaggctcc tcagccaaaa cgacaccccc atctgtctat cactggccc	3840
ctggatctgc tgccaaaact aactccatgg tgacctggg atgcctggtc aagggtatt	3900

tccctgagcc agtgacagtg acctggaact ctggatccct gtccagcggg gtgcacacct	3960
tcccagctgt cctgcagtct gacctctaca ctctgagcag ctcaagtact gtcccccca	4020
gcacctggcc cagcgagacc gtcacctgca acgttgccca cccggccagc agcaccaagg	4080
tggacaagaa aattgtgccc agggattgtg gttgtaagcc ttgcatatgt acagtcccag	4140
aagtatcatc tgtcttcac tccccccaa agcccaagga tgtgctcacc attactctga	4200
ctcctaagggt cacgtgtgtt gtggtagaca tcagcaagga tgatcccag gtccagttca	4260
gctggtttgt agatgatgtg gaggtgcaca cagctcagac gcaaccccg gaggagcagt	4320
tcaacagcac tttccgctca gtcagtgaac ttcccatcat gcaccaggac tggctcaatg	4380
gcaaggagcg atcgctcacc atcaccatca ccatcaccat caccattaaa ggcctatttt	4440
ctttagtttg aatttactgt tattcgggtg gcatttctat gtttggtgag cggttttctg	4500
tgctcagagt gtgtttat tttatgtat aatttctttg tgagctcctg tttagcaggt	4560
cgtcccttca gcaaggacac aaaaagattt taattttatt aaaaaaaaaa aaaaaaaga	4620
ccgggaattc gatatcaagc ttatcgacct gcagatcggt caaacatttg gcaataaagt	4680
ttcttaagat tgaatcctgt tgccgggtctt gcgatgatta tcatataatt tctgttgaat	4740
tacgttaagc atgtaataat taacatgtaa tgcatgacgt tatttatgag atgggttttt	4800
atgattagag tcccgcatt atacatttaa tacgcgatag aaaacaaaat atagcgcgca	4860
aactaggata aattatcgcg cgcggtgtca tctatgttac tagatctcta gagtctcaag	4920
cttggcgcgc cataaaatga ttattttatg aatataattt attgtgcaag tagatagaaa	4980
ttacatatgt tacataacac acgaaataaa caaaaaaaga caatccaaaa acaaacaccc	5040
caaaaaaat aatcacttta gataaactcg tatgaggaga ggcacgttca gtgactcgac	5100
gattcccag caaaaaaagt ctccccgtca cacatatagt gggtgacgca attatcttta	5160
aagtaatcct tctgttgact tgtcattgat aacatccagt cttcgtcagg attgcaaaga	5220
attatagaag ggatcccacc ttttattttt ttcttttttc catatttagg gttgacagtg	5280
aaatcagact ggcaacctat taattgcttc cacaatggga cgaacttgaa ggggatgtcg	5340
tcgatgatat tatagggtggc gtgttcacgt tagttggtga aatcgatgg accgttccaa	5400
tagttgtgtc gtccgagact tctagcccag gtggtctttc cggtagcagt tgggtccgag	5460
atgtagaggc tgggggtgtc gattccattc cttccattgt cctgggttaa tcggccatcc	5520
attcaaggtc agattgagct tgttggtatg agacaggatg tatgtaagta taagcgtcta	5580
tgcttacatg gtatagatgg gtttccctcc aggagtgtag atcttcgtgg cagcgaagat	5640
ctgattctgt gaagggcgac acatacgggt cagggtgtgg agggaataat ttgttggtcg	5700

aatattccag ccattgaagt ttgttggccc attcatgagg gaattcttcc ttgatcatgt 5760
caagatattc ctcttagac gttgcagtct ggataatagt tctccatcgt gcgtcagatt 5820
tgcgaggaga gaccttatga tctcggaat ctctctgtgt ttaatatct cgtctcttg 5880
atatgtaatc aaggacttgt ttagagtttc tagctggctg gatattaggg tgatttcctt 5940
caaaatcgaa aaaagaagga tccctaatac aagggttttt atcaagctgg agaagagcat 6000
gatagtgggt agtgccatct tgatgaagct cagaagcaac accaaggaag aaaataagaa 6060
aagggtgtgag tttctcccag agaaactgga ataatcatc tctttgagat gagcacttgg 6120
gataggtaag gaaaacatat ttagattgga gtctgaagtt cttactagca gaaggcatgt 6180
tgttgtgact ccgaggggtt gcctcaaact ctatcttata accggcgtgg aggcattggg 6240
gcaggggtat ttgtgtcatt ttaatagata gtggaaaatg acgtggaatt tacttaaaga 6300
cgaagtcttt gcgacaaggg ggggccacg ccgaatttaa tattaccggc gtggccccc 6360
cttatcgca gtgcttttagc acgagcgtc cagatttaaa gtagaaaatt tcccgcac 6420
tagggtaaa ggtgttcaca ctataaaagc atatacgatg tgatggtatt tgactagtgg 6480
cactggccgt cgttttataa cgtcgtgact gggaaaacc tggcgttacc caacttaatc 6540
gccttgagc acatccccct ttcgccagct ggcgtaatag cgaagaggcc cgcaccgatc 6600
gcccttcca acagttgcgc agcctgaatg gcgaatgcta gagcagcttg agcttgatc 6660
agattgtcgt tccccgctt cagtttaaac tatcagtgtt tgacaggata tattggcggg 6720
taaacctaag agaaaagagc gttta 6745

<210> 53

<211> 3495

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Cát xét biểu hiện số 1462

<400> 53

gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60
gaagacaaa gggcaattga gacttttcaa caaagggtaa tatccgaaa cctcctcgga 120
ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180
tacaaatgcc atcattgca taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240
ggtcccaaag atggaccccc acccagagg agcatcgttg aaaaagaaga cgttccaacc 300
acgtcttcaa agcaagtgga ttgatgtgat aacatggttg agcacgacac acttgtctac 360
tccaaaaata tcaaagatac agtctcagaa gaccaaagg caattgagac ttttcaacaa 420
agggtaatat ccggaacct cctcggttc cattgccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc	540
atcgttgaag atgcctctgc cgacagtggc cccaaagatg gacccccacc cacgaggagc	600
atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc	660
tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagacc ttctctata	720
taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga	780
acgtggggaa acccgaacca aaccttcttc taaactctct ctcactctctc ttaaagcaaa	840
cttctctctt gtctttcttg cgtgagcgat cttcaacggt gtcagatcgt gcttcggcac	900
cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc	960
ggcgccatta aataacgtgt acttgtccta ttcttgtcgg tgtggctctg ggaaaagaaa	1020
gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg	1080
gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct	1140
tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg	1200
tggttttcga acttggagaa agattgttaa gcttctgtat attctgcca aatttgtcgg	1260
gcccataag gcaataattg tactactcat ggtagtaaca tccaatgcag atcgaatctg	1320
cactgggata acatcttcaa actcacctca tgttggtcaaa acagctactc aaggggaggt	1380
caatgtgact ggcgtagtac cactgacaac aacaccaaca aaatcttatt ttgcaaactc	1440
caaaggaaca aggaccagag ggaaactatg cccggactgt ctcaactgta cagatctgga	1500
tgtggccttg ggcaggccaa tgtgtgtggg gaccacacct tctgctaaag cttcaatact	1560
ccacgaggtc agacctgtta catccgggtg ctttctata atgcacgaca gaacaaaaat	1620
caggcaacta cccaatcttc tcagaggata tgaaaatata aggttatcaa cccaaaacgt	1680
tatcgatgca gaaaaagcac caggaggacc ctacagactt ggaacctcag gatcttgccc	1740
taacgctacc agtaaaatcg gattttttgc aacaatggct tgggctgtcc caaaggacaa	1800
ctacaaaaat gcaacgaacc cactaacagt agaagtacca tacatttgta cagaagggga	1860
agaccaaatt actgtttggg ggttccattc agataacaaa acccaaata agagcctcta	1920
tggagactca aatcctcaaa agttcacctc atctgctaata ggagtaacca cacattatgt	1980
ttctcagatt ggcgacttcc cagatcaaac agaagacgga ggactaccac aaagcggcag	2040
aattgttggt gattacatga tgcaaaaacc tgggaaaaca ggaacaattg tctatcaaag	2100
aggtgttttg ttgcctcaaa aggtgtgggt gcgagtggtg aggagcaaag taataaaagg	2160
gtcattgcct ttaattggtg aagcagattg cttcatgaa aaatacgggt gattaaacaa	2220
aagcaagcct tactacacag gagaacatgc aaaagccata ggaaattgcc caatatgggt	2280

```

aaaaacacct ttgaagcttg ccaatggaac caaatataga cctcctgcaa aactattgaa 2340
ggaaaggggt ttcttcggag ctattgctgg tttcctagaa ggaggatggg aaggaatgat 2400
tgcaggttgg cacggataca catctcacgg agcacatgga gtggcagtgg cggcagacct 2460
taagagtaca caagaagcta taaataagat aacaaaaaat ctcaattctt tgagtgaagct 2520
agaagtaaag aaccttcaaa gactaagtgg tgccatggat gaactccaca acgaaatact 2580
cgagctggat gagaaagtgg atgatctcag agctgacact ataagctcac aaatagaact 2640
tgcagtcttg ctttccaacg aaggaataat aaacagtgaa gacgagcatc tattggcact 2700
tgagagaaaa ctaaagaaaa tgctgggtcc ctctgctgta gacataggaa acggatgctt 2760
cgaaacccaaa cacaaatgca accagacctg cttagacagg atagctgctg gcacctttaa 2820
tgcaggagaa ttttctctcc ccacttttga ttcattgaac attactgctg catctttaa 2880
tgatgatgga ttggataacc atactatact gctctattac tcaactgctg cttctagttt 2940
ggctgtaaca ttaatgctag ctatttttat tgtttatatg gtctccagag acaacgtttc 3000
atgctccatc tgtctataaa ggcctatttt ctttagtttg aatttactgt tattcggtgt 3060
gcattttctat gtttggtgag cggttttctg tgctcagagt gtgtttattt tatgtaattt 3120
aattttcttg tgagctcctg tttagcaggt cgtcccttca gcaaggacac aaaaagattt 3180
taattttatt aaaaaaaaaa aaaaaaaga ccgggaattc gatatcaagc ttatcgacct 3240
gcagatcggt caaacatttg gcaataaagt ttcttaagat tgaatcctgt tgccggtctt 3300
gcgatgatta tcatataatt tctgttgaat tacgttaagc atgtaataat taacatgtaa 3360
tgcatgacgt tatttatgag atgggttttt atgattagag tcccgcaatt atacatttaa 3420
tacgcgatag aaaacaaaat atagcgcgca aactaggata aattatcgcg cgcggtgtca 3480
tctatgttac tagat 3495

```

<210> 54

<211> 584

<212> PRT

<213> Virut cúm

<400> 54

```

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

```

```

Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val Lys
20           25           30

```

```

Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu Thr
35           40          45

```

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

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

Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys Ser
 305 310 315 320

Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn Cys Pro
 325 330 335

Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg
 340 345 350

Pro Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala
 355 360 365

Gly Phe Leu Glu Gly Gly Trp Glu Gly Met Ile Ala Gly Trp His Gly
 370 375 380

Tyr Thr Ser His Gly Ala His Gly Val Ala Val Ala Ala Asp Leu Lys
 385 390 395 400

Ser Thr Gln Glu Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn Ser Leu
 405 410 415

Ser Glu Leu Glu Val Lys Asn Leu Gln Arg Leu Ser Gly Ala Met Asp
 420 425 430

Glu Leu His Asn Glu Ile Leu Glu Leu Asp Glu Lys Val Asp Asp Leu
 435 440 445

Arg Ala Asp Thr Ile Ser Ser Gln Ile Glu Leu Ala Val Leu Leu Ser
 450 455 460

Asn Glu Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala Leu Glu
 465 470 475 480

Arg Lys Leu Lys Lys Met Leu Gly Pro Ser Ala Val Asp Ile Gly Asn
 485 490 495

Gly Cys Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu Asp Arg
 500 505 510

Ile Ala Ala Gly Thr Phe Asn Ala Gly Glu Phe Ser Leu Pro Thr Phe
 515 520 525

Asp Ser Leu Asn Ile Thr Ala Ala Ser Leu Asn Asp Asp Gly Leu Asp

530 535 540

Asn His Thr Ile Leu Leu Tyr Tyr Ser Thr Ala Ala Ser Ser Leu Ala
545 550 555 560

Val Thr Leu Met Leu Ala Ile Phe Ile Val Tyr Met Val Ser Arg Asp
565 570 575

Asn Val Ser Cys Ser Ile Cys Leu
580

<210> 55
<211> 56
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> đoạn mỗi HAB110(PrL-).r

<400> 55
tccttcccat cctccaccag gaggtctata tttggttcca ttggcaagct tcaaag 56

<210> 56
<211> 52
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> đoạn mỗi HAB110(PrL-).c

<400> 56
atatagacct cctggtggag gatgggaagg aatgattgca ggttggcacg ga 52

<210> 57
<211> 3444
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cát xét biểu hiện số 1467

<400> 57
gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60
gaagacaaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120
ttccattgcc cagctatctg tcaactttatt gtgaagatag tggaaaagga aggtggctcc 180
tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240
ggtcccaaag atggaccccc acccagcagg agcatcgttg aaaaagaaga cgttccaacc 300
acgtcttcaa agcaagtgga ttgatgtgat aacatggttg agcacgacac acttgtctac 360
tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaacct cctcggttc cattgcccag ctatctgtca ctttattgtg	480
aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc	540
atcgttgaag atgcctctgc cgacagtggc cccaaagatg gacccccacc cacgaggagc	600
atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc	660
tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttctctata	720
taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga	780
acgtggggaa acccgaacca aaccttcttc taaactctct ctcactcttc ttaaagcaaa	840
cttctctctt gtctttcttg cgtgagcgt cttcaacgtt gtcagatcgt gcttcggcac	900
cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc	960
ggcgccatta aataacgtgt acttgctcta ttcttgctgg tgtggctctg ggaaaagaaa	1020
gcttgctgga ggctgctgtt cagccccata cttacttgtt tacgattctg ctgactttcg	1080
gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct	1140
tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttccc	1200
tggttttcga acttgagaaa agattgttaa gcttctgtat attctgcca aatttgctcg	1260
gcccataag gcaataattg tactactcat ggtagtaaca tccaatgcag atcgaatctg	1320
cactgggata acatcttcaa actcacctca tgtggtcaaa acagctactc aaggggaggt	1380
caatgtgact ggcgtgatac cactgacaac aacaccaaca aaatcttatt ttgcaaatct	1440
caaaggaaca aggaccagag ggaaactatg cccggactgt ctcaactgta cagatctgga	1500
tgtggccttg ggcaggccaa tgtgtgtggg gaccacacct tctgctaaag cttcaatact	1560
ccacgaggtc agacctgtta catccgggtg ctttcctata atgcacgaca gaacaaaaat	1620
caggcaacta cccaatcttc tcagaggata tgaaaatatc aggttatcaa cccaaaacgt	1680
tatcgatgca gaaaaagcac caggaggacc ctacagactt ggaacctcag gatcttgccc	1740
taacgctacc agtaaaatcg gattttttgc aacaatggct tgggctgtcc caaaggacaa	1800
ctacaaaaat gcaacgaacc cactaacagt agaagtacca tacatttgta cagaagggga	1860
agaccaaatt actgtttggg ggttcatttc agataacaaa acccaaatga agagcctcta	1920
tggagactca aatcctcaaa agttcacctc atctgctaag ggagtaacca cacattatgt	1980
ttctcagatt ggcgacttcc cagatcaaac agaagacgga ggactaccac aaagcggcag	2040
aattgttgtt gattacatga tgcaaaaacc tgggaaaaca ggaacaattg tctatcaaag	2100
aggtgttttg ttgcctcaaa aggtgtgggt gcgagtggtc aggagcaaag taataaaagg	2160
gtcattgcct ttaattggtg aagcagattg cttcatgaa aaatacgggt gattaaacaa	2220

```

aagcaagcct tactacacag gagaacatgc aaaagccata ggaaattgcc caatatgggt      2280
aaaaacacct ttgaagcttg ccaatggaac caaatataga cctcctgggtg gaggatggga      2340
aggaatgatt gcaggttggc acggatacac atctcacgga gcacatggag tggcagtggc      2400
ggcagacctt aagagtacac aagaagctat aaataagata acaaaaaatc tcaattcttt      2460
gagtgaagcta gaagtaaaga accttcaaag actaagtggg gccatggatg aactccacaa      2520
cgaaatactc gagctggatg agaaagtgga tgatctcaga gctgacacta taagctcaca      2580
aatagaactt gcagtcttgc tttccaacga aggaataata aacagtgaag acgagcatct      2640
attggcactt gagagaaaac taaagaaaat gctgggtccc tctgctgtag acataggaaa      2700
cggatgcttc gaaaccaaac acaaatgcaa ccagacctgc ttagacagga tagctgctgg      2760
cacctttaat gcaggagaat tttctctccc cacttttgat tcattgaaca ttactgctgc      2820
atctttaaat gatgatggat tggataacca tactatactg ctctattact caactgctgc      2880
ttctagtttg gctgtaacat taatgctagc tatttttatt gtttatatgg tctccagaga      2940
caacgtttca tgctccatct gtctataaag gcctattttc tttagtttga atttactgtt      3000
attcgggtgtg catttctatg tttggtgagc ggttttctgt gctcagagtg tgtttatfff      3060
atgtaattta atttctttgt gagctcctgt ttagcagggtc gtcccttcag caaggacaca      3120
aaaagatttt aattttatta aaaaaaaaaa aaaaaaagac cggaattcg atatcaagct      3180
tatcgacctg cagatcggtc aaacatttgg caataaagtt tcttaagatt gaatcctggt      3240
gccggctctt cgatgattat catataattt ctgttgaatt acgttaagca tgtaataatt      3300
aacatgtaat gcatgacgtt atttatgaga tgggttttta tgattagagt cccgcaatta      3360
tacatttaat acgcgataga aaacaaaata tagcgcgcaa actaggataa attatcgcg      3420
gcggtgtcat ctatgttact agat      3444

```

<210> 58
 <211> 567
 <212> PRT
 <213> Virut cúm

<400> 58

```

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

```

```

Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val Lys
          20           25           30

```

```

Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu Thr
35           40           45

```

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

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

Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys Ser
 305 310 315 320

Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn Cys Pro
 325 330 335

Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg
 340 345 350

Pro Pro Gly Gly Gly Trp Glu Gly Met Ile Ala Gly Trp His Gly Tyr
 355 360 365

Thr Ser His Gly Ala His Gly Val Ala Val Ala Ala Asp Leu Lys Ser
 370 375 380

Thr Gln Glu Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser
 385 390 395 400

Glu Leu Glu Val Lys Asn Leu Gln Arg Leu Ser Gly Ala Met Asp Glu
 405 410 415

Leu His Asn Glu Ile Leu Glu Leu Asp Glu Lys Val Asp Asp Leu Arg
 420 425 430

Ala Asp Thr Ile Ser Ser Gln Ile Glu Leu Ala Val Leu Leu Ser Asn
 435 440 445

Glu Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala Leu Glu Arg
 450 455 460

Lys Leu Lys Lys Met Leu Gly Pro Ser Ala Val Asp Ile Gly Asn Gly
 465 470 475 480

Cys Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile
 485 490 495

Ala Ala Gly Thr Phe Asn Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp
 500 505 510

Ser Leu Asn Ile Thr Ala Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn
 515 520 525

His Thr Ile Leu Leu Tyr Tyr Ser Thr Ala Ala Ser Ser Leu Ala Val

530

535

540

Thr Leu Met Leu Ala Ile Phe Ile Val Tyr Met Val Ser Arg Asp Asn
 545 550 555 560

Val Ser Cys Ser Ile Cys Leu
 565

<210> 59
 <211> 44
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> đoạn mỗi IF-HB-M-04.s2+4c

<400> 59
 tctcagatct tcgccgatcg aatctgcact gggataacat cgtc 44

<210> 60
 <211> 46
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> đoạn mỗi IF-HB-M-04.s1-4r

<400> 60
 actaaagaaa ataggccttt atagacagat ggagcaagaa acattg 46

<210> 61
 <211> 1713
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> trình tự nucleotit của HA B Malaysia tổng hợp

<400> 61
 gatcgaatct gcaactgggat aacatcgta aactcaccac atgttgtcaa aactgctact 60
 caaggggagg tcaatgtgac tgggtgtaata ccactgacaa caacacccac caaatctcat 120
 tttgcaaadc tcaaaggaac agaaaccaga gggaaactat gcccaaatg cctcaactgc 180
 acagatctgg acgtggcctt gggcagacca aatgcacgg ggaacatacc ctcggcaaga 240
 gtttcaatac tccatgaagt cagacctgtt acatctgggt gctttcctat aatgcacgac 300
 agaacaaaaa ttagacagct gcctaaactt ctcagaggat acgaacatat caggttatca 360
 actcataacg ttatcaatgc agaaaatgca ccaggaggac cctacaaaat tggaacctca 420
 ggggtcttgcc ctaacgttac caatggaaac ggatttttcg caacaatggc ttggggccgtc 480
 ccaaaaaacg acaacaacaa aacagcaaca aattcattaa caatagaagt accatacatt 540

tgtacagaag gagaagacca aattaccggt tgggggttcc actctgataa cgaaacccaa 600
 atggcaaagc tctatgggga ctcaaagccc cagaagttca cctcatctgc caacggagtg 660
 accacacatt acgtttcaca gattggtggc ttcccaaatac aaacagaaga cggaggacta 720
 ccacaaagcg gtagaattgt tgttgattac atgggtgcaaa aatctgggaa aacaggaaca 780
 attacctatc aaagaggtat tttattgcct caaaaagtgt ggtgcgcaag tggcaggagc 840
 aaggtaataa aaggatcggt gcctttaatt ggagaagcag attgcctcca cgaaaaatac 900
 ggtggattaa acaaaagcaa gccttactac acaggggaac atgcaaaggc cataggaaat 960
 tgcccaatat ggggtgaaaac acccttgaag ctggccaatg gaaccaaata tagacctcct 1020
 gcaaaactat taaaggaaag gggtttcttc ggagctattg ctggtttctt agaaggagga 1080
 tgggaaggaa tgattgcagg ttggcacgga tacacatccc atggggcaca tggagtagcg 1140
 gtggcagcag accttaagag cactcaagag gccataaaca agataacaaa aaatctcaac 1200
 tctttgagtg agctggaagt aaagaatctt caaagactaa gcggtgccat ggatgaactc 1260
 cacaacgaaa tactagaact agacgagaaa gtggatgatc tcagagctga tacaataagc 1320
 tcacaaatag aactgcaggt cctgctttcc aatgaaggaa taataaacag tgaagatgag 1380
 catctcttgg cgcttgaaag aaagctgaag aaaatgctgg gccctctgc ttagagata 1440
 gggaatggat gctttgaaac caaacacaag tgcaaccaga cctgtctcga cagaatagct 1500
 gctggtacct ttgatgcagg agaattttct ctccccactt ttgattcact gaatattact 1560
 gctgcatctt taaatgacga tggattggat aatcatacta tactgcttta ctactcaact 1620
 gctgcctcca gtttggtgtg aacattgatg atagctatct ttgttgttta tatggtctcc 1680
 agagacaatg tttcttgctc catctgtcta taa 1713

<210> 62
 <211> 6739
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc 194

<400> 62
 tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 60
 gacgttttta atgtactgaa ttaacgccga atccccggct ggtatattta tatgttgtca 120
 aataactcaa aaaccataaa agtttaagtt agcaagtgtg tacattttta cttgaacaaa 180
 aatattcacc tactactggt ataaatcatt attaaacatt agagtaaaga aatatggatg 240
 ataagaacaa gagtagtgat attttgacaa caattttgtt gcaacatttg agaaaatttt 300

gttggtctct cttttcattg gtcaaaaaca atagagagag aaaaaggaag agggagaata	360
aaaacataat gtgagtatga gagagaaagt tgtacaaaag ttgtaccaa atagttgtac	420
aaatatcatt gaggaatttg acaaaagcta cacaaataag ggttaattgc tgtaataaa	480
taaggatgac gcattagaga gatgtacat tagagaattt ttggcaagtc attaaaaaga	540
aagaataaat tttttttaa attaaaagtt gagtcatttg attaaacatg tgattattta	600
atgaattgat gaaagagttg gattaaagtt gtattagtaa ttagaatttg gtgtcaaatt	660
taatttgaca tttgatcttt tcctatatat tgccccatag agtcagttaa ctcattttta	720
tatttcatag atcaaataag agaaataacg gtatattaat ccctccaaa aaaaaaacg	780
gtatatttac taaaaatct aagccacgta ggaggataac aggatccccg taggaggata	840
acatccaatc caaccaatca caacaatcct gatgagataa cccacttta gccacgcat	900
ctgtggcaca tctacattat ctaaatcaca cattcttcca cacatctgag ccacacaaaa	960
accaatccac atctttatca cccattctat aaaaaatcac actttgtgag tctacacttt	1020
gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag	1080
aaaatggaac gagctataca aggaaacgac gctagggaac aagctaacag tgaacgttgg	1140
gatggaggat caggaggta cacttctccc ttcaaacttc ctgacgaaag tccgagttgg	1200
actgagtggc ggctacataa cgatgagacg aattcgaatc aagataatcc ctttggttcc	1260
aaggaaagct ggggtttcgg gaaagttgta tttaagagat atctcagata cgacaggacg	1320
gaagcttcac tgcacagagt ccttgatct tggacgggag attcggttaa ctatgcagca	1380
tctcgatttt tcggtttcga ccagatcgga tgtacctata gtattcggtt tcgaggagtt	1440
agtatcaccg tttctggagg gtgcggaact cttcagcatc tctgtgagat ggcaattcgg	1500
tctaagcaag aactgctaca gcttgcccca atcgaagtgg aaagtaatgt atcaagagga	1560
tgccctgaag gtactcaaac cttcgaaaaa gaaagcgagt aagttaaat gcttcttcgt	1620
ctcctattta taatatggtt tgttattggt aattttgttc ttgtagaaga gcttaattaa	1680
tcgttggtgt tatgaaatac tatttgatg agatgaactg gtgtaatgta attcatttac	1740
ataagtggag tcagaatcag aatgtttcct ccataactaa ctagacatga agacctgccg	1800
cgtacaattg tcttatattt gaacaactaa aattgaacat cttttgccac aactttataa	1860
gtggttaata tagctcaa ataatggtaa gttcaataga ttaataatgg aaatatcagt	1920
tatcgaaatt cattaacaat caacttaacg ttattaacta ctaattttat atcatcccct	1980
ttgataaatg atagtacacc aattaggaag gagcatgctc gcctaggaga ttgtcgtttc	2040
ccgccttcag tttgcaagct gctctagccg tgtagccaat acgcaaaccg cctctccccg	2100
cgcgttggga attactagcg cgtgtcgaga cgcgttggtt ttgtgactcc gaggggttgc	2160

ctcaaactct atcttataac cggcgtggag gcatggaggc aggggtattht tggatcattht	2220
aatagatagt ggaaaatgac gtggaattta cttaaagacg aagtctthtg gacaaggggg	2280
ggccccacgcc gaatttaata ttaccggcgt ggccccccct tatcgcgagt gctthtagcac	2340
gagcgggtcca gattttaaagt agaaaatttc ccgcccacta gggttaaagg tgttcacact	2400
ataaaagcat atacgatgtg atggatthtg gtcgacaagc ttgcatgccg gtcaacatgg	2460
tggagcacga cacacttgte tactccaaaa atatcaaaga tacagtctca gaagacaaaa	2520
gggcaattga gactthtcaa caaagggtaa tatccggaaa cctcctcgga ttccattgcc	2580
cagctatctg tcactthatt gtgaagatag tggaaaagga aggtggctcc tacaaatgcc	2640
atcattgcga taaaggaaag gccatcgthg aagatgcctc tgccgacagt ggtcccaaag	2700
atggaccccc acccacgagg agcatcgthg aaaaagaaga cgttccaacc acgtcttcaa	2760
agcaagtga ttgatgtgat aacatggthg agcacgacac acttgcttac tccaaaaata	2820
tcaaagatac agtctcagaa gaccaaagg caattgagac thttcaacaa agggtaatat	2880
ccggaaacct cctcggahtc cattgcccag ctatctgtca thttatttg agatagthg	2940
aaaaggaaag tggctcctac aaatgccatc attgcgataa aggaaaggcc atcgthgaag	3000
atgcctctgc cgacagthg cccaaagatg gacccccacc cacgaggagc atcgthgaaa	3060
aagaagacgt tccaaccacg tcttcaaagc aagthgattg atgtgatatc tccactgacg	3120
taagggatga cgcacaatcc cactatcctt cgcaagaccc thctctata taaggaaht	3180
catttcattt ggagaggtat taaaatctta atagththtg ataaaagcga acgtggggaa	3240
acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa thtctctctt	3300
gtctthtctg cgtgagcgat thtcaacgtt gtcagatcgt gcttcggcac cagtacaacg	3360
thttctthtca ctgaagcgaa atcaaagatc thttgtgga cacgtagthg ggcgccatta	3420
aataacgtgt acttgthcta thttgtcgg thtggtcttg ggaaaagaaa gcttgctgga	3480
ggctgctgth cagccccata cattactthg tacgattctg ctgactthcg gcgggtgcaa	3540
tatctctact thtgcttgac gaggtattgt tgctgtact thttcttht thttctthgt	3600
gattgthtct ataagaaatc tagtaththt thtgaaacag agththcccg thgtthtctga	3660
actthgagaa agattgthaa gctthtctat thtctgcca aattthtctg gccatggcg	3720
aaaaacgttg cgaththtctg thtattgtht thtctthtthg thtggtthcc thtctcagatc	3780
thtgcgcgth thtctcagcc aaaacgacac cccatctgt ctatccactg gccctggat	3840
ctgctgcca aactaactcc atggthgacc thggatgcct ggtcaagggc taththctctg	3900
agccagthgac agthgacctg aactctggat cctgtctcag cgtgtgthc acctthccag	3960

ctgtcctgca gtctgacctc tacactctga gcagctcagt gactgtcccc tccagcacct	4020
ggcccagcga gaccgtcacc tgcaacgttg cccacccggc cagcagcacc aaggtggaca	4080
agaaaattgt gcccagggat tgtggttgta agccttgcat atgtacagtc ccagaagtat	4140
catctgtctt catcttcccc ccaaagccca aggatgtgct caccattact ctgactccta	4200
aggtcacgtg tgttgtggta gacatcagca aggatgatcc cgaggtccag ttcagctggg	4260
ttgtagatga tgtggagggtg cacacagctc agacgcaacc ccgggaggag cagttcaaca	4320
gcactttccg ctcagtcagt gaacttccca tcatgcacca ggactggctc aatggcaagg	4380
agcgatcgct caccatcacc atcaccatca ccatcaccat taaaggccta ttttcttttag	4440
tttgaattta ctgttattcg gtgtgcattt ctatgtttgg tgagcggttt tctgtgctca	4500
gagtgtgttt attttatgta atttaatttc tttgtgagct cctgttttagc aggtcgtccc	4560
ttcagcaagg acacaaaaag attttaattt tattaaaaaa aaaaaaaaaa aagaccggga	4620
attcgatatc aagcttatcg acctgcagat cgttcaaaca tttggcaata aagtttctta	4680
agattgaatc ctggtgocgg tcttgcatg attatcatat aatttctgtt gaattacgtt	4740
aagcatgtaa taattaacat gtaatgcatg acgttattta tgagatgggt ttttatgatt	4800
agagtcccg c aattatacat ttaatacgcg atagaaaaca aaatatagcg cgcaaaactag	4860
gataaattat cgcgcgcgggt gtcactctatg ttactagatc tctagagtct caagcttggc	4920
gcgccataaa atgattattht tatgaatata tttcattgtg caagtagata gaaattacat	4980
atgttacata acacacgaaa taaacaaaaa aagacaatcc aaaaacaaac accccaaaaa	5040
aaataatcac tttagataaa ctcgatatgag gagaggcacg ttcagtgact cgacgattcc	5100
cgagcaaaaa aagtctcccc gtcacacata tagtgggtga cgcaattatc tttaaagtaa	5160
tccttctgtt gacttgtcat tgataacatc cagtcttctg caggattgca aagaattata	5220
gaagggatcc caccttttat tttcttcttt tttccatatt taggggtgac agtgaaatca	5280
gactggcaac ctattaattg cttccacaat gggacgaact tgaaggggat gtcgtcgatg	5340
atattatagg tggcgtgttc atcgtagttg gtgaaatcga tggtagcggt ccaatagttg	5400
tgtcgtccga gacttctagc ccagggtggc tttccggtac gagttgggtc gcagatgtag	5460
aggctgggggt gtcggattcc attccttcca ttgtcctgggt taaatcggcc atccattcaa	5520
ggtcagattg agcttggttg tatgagacag gatgtatgta agtataagcg tctatgctta	5580
catggtatag atgggtttcc ctccaggagt gtagatcttc gtggcagcga agatctgatt	5640
ctgtgaaggg cgacacatac ggttcagggt gtggagggaa taatttggtg gctgaatatt	5700
ccagccattg aagttttgtt gccattcat gaggggaattc ttccttgatc atgtcaagat	5760
attcctcctt agacgttgca gtctggataa tagttctcca tcgtgcgtca gatttgcgag	5820

gagagacctt atgatctcgg aaatctcctc tggttttaat atctccgtcc tttgatatgt 5880
 aatcaaggac ttgttttagag tttctagctg gctggatatt agggtgattt ccttcaaaat 5940
 cgaaaaaaga aggatcccta atacaagggtt ttttatcaag ctggagaaga gcatgatagt 6000
 gggtagtgcc atcttgatga agctcagaag caacaccaag gaagaaaata agaaaagggtg 6060
 tgagtttctc ccagagaaac tggaataaat catctctttg agatgagcac ttgggatagg 6120
 taaggaaaac atatttagat tggagtctga agttcttact agcagaaggc atgttggtgt 6180
 gactccgagg ggttgccctca aactctatct tataaccggc gtggaggcat ggaggcaggg 6240
 gtatttttgt cattttaata gatagtggaa aatgacgtgg aatttactta aagacgaagt 6300
 ctttgcgaca agggggggcc cagccgaat ttaatattac cggcgtggcc cccccttacc 6360
 gcgagtgtt tagcacgagc ggtccagatt taaagtagaa aatttccgc ccactagggt 6420
 taaagggtgtt cacactataa aagcatatac gatgtgatgg tatttgacta gtggcactgg 6480
 ccgtcgtttt acaacgtcgt gactgggaaa accctggcgt tacccaactt aatcgccttg 6540
 cagcacatcc ccctttcgcc agctggcgta atagcgaaga ggcccgacc gatcgcctt 6600
 cccaacagtt gcgcagcctg aatggcgaat gctagagcag cttgagcttg gatcagattg 6660
 tcgtttcccg ccttcagttt aaactatcag tgtttgacag gatatatgtg cgggtaaacc 6720
 taagagaaaa gagcgttta 6739

<210> 63
 <211> 3525
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Cát xét biểu hiện số 1631

<400> 63
 gtcaacatgg tggagcacga cacacttgct tactccaaaa atatcaaaga tacagtctca 60
 gaagaccaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120
 ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaggga aggtggctcc 180
 tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240
 ggtcccaaag atggaccccc acccagcagg agcatcgttg aaaaagaaga cgttccaacc 300
 acgtcttcaa agcaagtgga ttgatgtgat aacatgggtg agcacgacac acttgtctac 360
 tccaaaaata tcaaagatac agtctcagaa gaccaaagg caattgagac ttttcaacaa 420
 agggtaatat ccggaaacct cctcggattc cattgccag ctatctgtca ctttattgtg 480
 aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggc cccaaagatg gacccccacc cacgaggagc	600
atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc	660
tccactgaag taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata	720
taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga	780
acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa	840
cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac	900
cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc	960
ggcgccatta aataacgtgt acttgtccta ttcttgtcgg tgtggctctg ggaaaagaaa	1020
gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg	1080
gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgctgtact tctttcttct	1140
tcttcttgtt gattggttct ataagaaatc tagtatcttc tttgaaacag agttttcccg	1200
tggttttoga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcgg	1260
gccccatggc aaaaacgttg cgattttcgg cttattgttt tctcttcttg tgttggttcc	1320
ttctcagatc ttgcgcgac gaatctgcac tgggataaca tcgtcaaact caccacatgt	1380
tgtcaaaact gctactcaag gggaggtcaa tgtgactggg gtaataccac tgacaacaac	1440
accacacaaa tctcatcttg caaatctcaa aggaacagaa accagaggga aactatgccc	1500
aaaatgcctc aactgcacag atctggacgt ggccttgggc agacaaaaat gcacggggaa	1560
cataccctcg gcaagagttt caatactcca tgaagtcaga cctgttacat ctgggtgctt	1620
tcctataatg cacgacagaa caaaaattag acagctgcct aaacttctca gaggatacga	1680
acatatcagg ttatcaactc ataacgttat caatgcagaa aatgcaccag gaggacccta	1740
caaaaattgga acctcagggt cttgccctaa cgttaccaat ggaaacggat ttttcgcaac	1800
aatggcttgg gccgtcccaa aaaacgacaa caacaaaaca gcaacaaatt cattaacaat	1860
agaagtacca tacatttgta cagaaggaga agaccaaatt accgtttggg ggttccactc	1920
tgataacgaa acccaaattg caaagctcta tggggactca aagccccaga agttcacctc	1980
atctgccaac ggagtgaaca cacattacgt ttcacagatt ggtggcttcc caaatcaaac	2040
agaagacgga ggactaccac aaagcggtag aattgttggt gattacatgg tgcaaaaatc	2100
tgggaaaaca ggaacaatta cctatcaaag aggtatttta ttgcctcaaa aagtgtggtg	2160
cgcaagtggc aggagcaagg taataaaagg atcgttgccct ttaattggag aagcagattg	2220
cctccacgaa aaatacgttg gattaaacaa aagcaagcct tactacacag gggaacatgc	2280
aaaggccata ggaaattgcc caatatgggt gaaaacaccc ttgaagctgg ccaatggaac	2340
caaatataga cctcctgcaa aactattaaa ggaaaggggt ttcttcggag ctattgctgg	2400

```

tttcttagaa ggaggatggg aaggaatgat tgcaggttgg cacggataca catcccatgg      2460
ggcacatgga gtagcgggtg cagcagacct taagagcact caagaggcca taaacaagat      2520
aacaaaaaat ctcaactctt tgagtgaagt ggaagtaaag aatcttcaaa gactaagcgg      2580
tgccatggat gaactccaca acgaaatact agaactagac gagaaagtgg atgatctcag      2640
agctgataca ataagctcac aaatagaact cgcagtcctg ctttccaatg aaggaataat      2700
aaacagtga  gatgagcatc tcttggcgct tgaaagaaag ctgaagaaaa tgctggggccc      2760
ctctgctgta gagataggga atggatgctt tgaaacaaaa cacaagtgca accagacctg      2820
tctcgacaga atagctgctg gtacctttga tgcaggagaa ttttctctcc ccacttttga      2880
ttcactgaat attactgctg catcttttaa tgacgatgga ttggataatc atactatact      2940
gctttactac tcaactgctg cctccagttt ggctgtaaca ttgatgatag ctatctttgt      3000
tgtttatatg gtctccagag acaatgtttc ttgctccatc tgtctataaa ggcctatttt      3060
ctttagtttg aatttactgt tattcgggtg gcatttctat gtttggtgag cggttttctg      3120
tgctcagagt gtgtttattt tatgtaattt aatttctttg tgagctcctg tttagcaggt      3180
cgtcccttca gcaaggacac aaaaagattt taattttatt aaaaaaaaaa aaaaaaaaga      3240
ccgggaattc gatatcaagc ttatcgacct gcagatcggt caaacatttg gcaataaagt      3300
ttcttaagat tgaatcctgt tgccggtcct gcgatgatta tcatataatt tctgttgaat      3360
tacgttaagc atgtaataat taacatgtaa tgcagacgtt tatttatgag atgggttttt      3420
atgattagag tcccgcattt atacatttaa tacgcgatag aaaacaaaat atagcgcgca      3480
aactaggata aattatcgcg cgcggtgtca tctatgttac tagat                        3525

```

<210> 64
 <211> 594
 <212> PRT
 <213> Virut cúm

<400> 64

```

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1           5           10           15

```

```

Leu Val Pro Ser Gln Ile Phe Ala Asp Arg Ile Cys Thr Gly Ile Thr
          20           25           30

```

```

Ser Ser Asn Ser Pro His Val Val Lys Thr Ala Thr Gln Gly Glu Val
          35           40           45

```

```

Asn Val Thr Gly Val Ile Pro Leu Thr Thr Thr Pro Thr Lys Ser His
50           55           60

```

Phe Ala Asn Leu Lys Gly Thr Glu Thr Arg Gly Lys Leu Cys Pro Lys
 65 70 75 80
 Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Lys Cys
 85 90 95
 Thr Gly Asn Ile Pro Ser Ala Arg Val Ser Ile Leu His Glu Val Arg
 100 105 110
 Pro Val Thr Ser Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile
 115 120 125
 Arg Gln Leu Pro Lys Leu Leu Arg Gly Tyr Glu His Ile Arg Leu Ser
 130 135 140
 Thr His Asn Val Ile Asn Ala Glu Asn Ala Pro Gly Gly Pro Tyr Lys
 145 150 155 160
 Ile Gly Thr Ser Gly Ser Cys Pro Asn Val Thr Asn Gly Asn Gly Phe
 165 170 175
 Phe Ala Thr Met Ala Trp Ala Val Pro Lys Asn Asp Asn Asn Lys Thr
 180 185 190
 Ala Thr Asn Ser Leu Thr Ile Glu Val Pro Tyr Ile Cys Thr Glu Gly
 195 200 205
 Glu Asp Gln Ile Thr Val Trp Gly Phe His Ser Asp Asn Glu Thr Gln
 210 215 220
 Met Ala Lys Leu Tyr Gly Asp Ser Lys Pro Gln Lys Phe Thr Ser Ser
 225 230 235 240
 Ala Asn Gly Val Thr Thr His Tyr Val Ser Gln Ile Gly Gly Phe Pro
 245 250 255
 Asn Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val
 260 265 270
 Asp Tyr Met Val Gln Lys Ser Gly Lys Thr Gly Thr Ile Thr Tyr Gln
 275 280 285
 Arg Gly Ile Leu Leu Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser
 290 295 300

Lys Val Ile Lys Gly Ser Leu Pro Leu Ile Gly Glu Ala Asp Cys Leu
305 310 315 320

His Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys Pro Tyr Tyr Thr Gly
325 330 335

Glu His Ala Lys Ala Ile Gly Asn Cys Pro Ile Trp Val Lys Thr Pro
340 345 350

Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Ala Lys Leu Leu
355 360 365

Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala Gly Phe Leu Glu Gly Gly
370 375 380

Trp Glu Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His Gly Ala
385 390 395 400

His Gly Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala Ile
405 410 415

Asn Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val Lys
420 425 430

Asn Leu Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu Ile
435 440 445

Leu Glu Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile Ser
450 455 460

Ser Gln Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile Asn
465 470 475 480

Ser Glu Asp Glu His Leu Leu Ala Leu Glu Arg Lys Leu Lys Lys Met
485 490 495

Leu Gly Pro Ser Ala Val Glu Ile Gly Asn Gly Cys Phe Glu Thr Lys
500 505 510

His Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala Ala Gly Thr Phe
515 520 525

Asp Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser Leu Asn Ile Thr
530 535 540

Ala Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn His Thr Ile Leu Leu
545 550 555 560

Tyr Tyr Ser Thr Ala Ala Ser Ser Leu Ala Val Thr Leu Met Ile Ala
565 570 575

Ile Phe Val Val Tyr Met Val Ser Arg Asp Asn Val Ser Cys Ser Ile
580 585 590

Cys Leu

A-2X35S/CPMV-HT-/H5 Indonesia/NOS (Cấu trúc số 489)

Fig.1A

IF-H5A-I-05.s1+3c SEQ ID NO: 2

AAATTTGTCGGGCCCATGGAGAAAATAGTGCTTCTTCTTGC

Fig.1B

IF-H5dTm.r SEQ ID NO: 3

ACTAAAGAAAATAGGCCTTTAAATGCAAATTCTGCATTGTAACGATCCAT

Fig.1C

Bản đồ của cấu trúc 1191. Các vị trí enzym giới hạn SacII và StuI được sử dụng để kéo dài plasmit được giảm độc lực trên bản đồ này

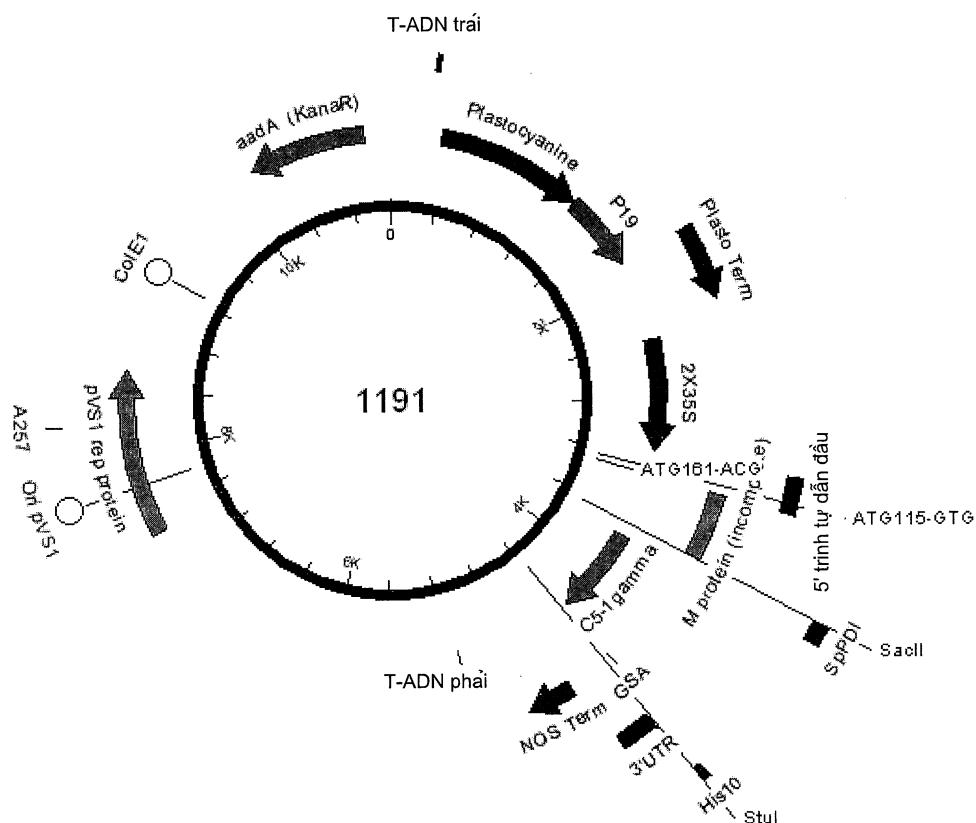


Fig.1D

Cấu trúc 1191 từ các bờ ADN vận chuyển từ trái sang phải (được gạch chân). 2X35S/CPMV-HT/NOS có băng biểu hiện chất ức chế bất hoạt Plastoxyanin-P19-Plastoxyanin

TGGCAGGATATATTGTGGTGTAAACAAATTGACGCTTAGACAACCTTAATAACACATTGCGGACGTTTTTA
ATGTACTGAATTAACGCCGAATCCCGGGCTGGTATATTTATATGTTGTCAAATAACTCAAAAACCATAAA
AGTTTAAAGTTAGCAAGTGTGTACATTTTACTTGAACAAAAATATTCACCTACTACTGTTATAAATCATT
TTAAACATTAGAGTAAAGAAAATATGGATGATAAGAACAAGAGTAGTGATATTTTGACAACAATTTTGTG
CAACATTTGAGAAAAATTTGTTGTTCTCTCTTTTCATTGGTCAAAAACAATAGAGAGAGAAAAAGGAAGA
GGGAGAATAAAAAACATAATGTGAGTATGAGAGAGAAAAGTTGTACAAAAGTTGTACCAAAATAGTTGTAC
AAATATCATTGAGGAATTTGACAAAAGCTACACAAATAAGGGTTAATTGCTGTAAATAAATAAGGATGA
CGCATTAGAGAGATGTACCATTAGAGAATTTTGGCAAGTCATTAAGAAAGAAAGAAATAAATTATTTTAA
AATTAAGTTGAGTCATTTGATTAAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGT
TGTATTAGTAATTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCTATATATTGCCCATAGA
GTCAGTTAACTCATTTTTATATTTTCATAGATCAAAAGAGAGAAATAACGGTATATTAATCCCTCCAAAAA
AAAAAACCGGTATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTAGGAGGATA
ACATCCAATCCAACCAATCACAACAATCCTGATGAGATAACCCACTTTAAGCCACGCATCTGTGGCACA
TCTACATTATCTAAATCACACATTCTTCCACACATCTGAGCCACACAAAAACCAATCCACATCTTTATCAC
CCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAG
ACTAATTAATTAATTAATCATCTTGAGAGAAAAATGGAACGAGCTATACAAGGAAACGACGCTAGGGAAC
AAGCTAACAGTGAACGTTGGGATGGAGGATCAGGAGGTACCACTTCTCCCTTCAAACCTCCTGACGAAAG
TCCGAGTTGGACTGAGTGGCGGCTACATAACGATGAGACGAATTCGAATCAAGATAATCCCTTGGTTTC
AAGGAAAGCTGGGGTTTCGGGAAAGTTGTATTTAAGAGATATCTCAGATACGACAGGACGGAAGCTTCA
CTGCACAGAGTCCTTGATCTTGACGGGAGATTCCGTTAACTATGCAGCATCTCGATTTTTCGGTTTCGA
CCAGATCGGATGTACCTATAGTATTCGGTTTCGAGGAGTTAGTATCACCGTTTCTGGAGGGTCGCGAACT
CTTCAGCATCTCTGTGAGATGGCAATTCGGTCTAAGCAAGAACTGCTACAGCTTGCCCAATCGAAGTGG
AAAGTAATGTATCAAGAGGATGCCCTGAAGGTACTCAAACCTTCGAAAAAGAAAGCGAGTAAGTTAAAA
TGCTTCTTCGTCTCCTATTTATAATATGGTTTGTATTGTTAATTTTGTCTGTAGAAGAGCTTAATTAAT
CGTTGTTGTTATGAAATACTATTTGTATGAGATGAAGTGGTGAATGTAAATTCATTTACATAAGTGGAGTC
AGAATCAGAATGTTTCCCTCCATAAATACTAGACATGAAGACCTGCCGCGTACAATTGTCTTATATTTGA
ACAATAAAATTGAACATCTTTTGCCACAACCTTTATAAGTGGTTAATATAGCTCAAAATATATGGTCAAGTT
CAATAGATTAATAATGGAAATATCAGTTATCGAAATTCATTAACAATCAACTTAACGTTATTAACCTACTA
ATTTTATATCATCCCTTTGATAAATGATAGTACCAATTAGGAAGGAGCATGCTCGCCTAGGAGATTG
TCGTTTCCCGCCTTCAGTTTGCAAGCTGCTCTAGCCGTGTAGCCAATACGCAAACCGCCTCTCCCCGCGCG
TTGGGAATTACTAGCGCGTGTGACAAGCTTGCATGCCGTCAACATGGTGGAGCACGACACACTTGTCT
ACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAGGGCAATTGAGACTTTTCAACAAAGGGTAA
TATCCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGA
AGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGT
GGTCCCAAAGATGGACCCCAACCCACGAGGAGCATCGTGGAAGAAAGAACGTTCCAACCACGCTCTCA
AAGCAAGTGGATTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGAT
ACAGTCTCAGAAGACCAAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGA
TTCCATTGCCCAGCTATCTGTCACTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCC
ATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCC
CACCCACGAGGAGCATCGTGGAAGAAAGAACGTTCCAACCACGCTCTTCAAAGCAAGTGGATTGATGTG
ATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCTCTATATAAGG
AAGTTCAATTTCAATTTGGAGAGGTATTAATCTTAATAGGTTTTGATAAAAGCGAACGTGGGGAAACCCG
AACCAAACTTCTTCTAACTCTCTCTCATCTCTCTTAAAGCAAACCTCTCTCTTGTCTTTCTTGCGTGAGC
GATCTTCAACGTTGTGATCGTGTCTCGGCACCACTACAACGTTTTCTTTCACTGAAGCGAAATCAAAG
ATCTCTTTGTGGACAGTAGTGCGGCCCATTAATAACGTGTACTTGTCTATTCTTGTGGTGTGGTCT
TGGGAAAAGAAAGCTTGTGAGGGCTGCTGTTTCAAGCCCATACATTACTTGTACGATTCTGCTGACTTTG
GGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTCTTCTTCTTGTGCTG

ATTGGTTCTATAAGAAATCTAGTATTTCTTTGAAACAGAGTTTCCCGTGGTTTTCGAACTGGAGAAAAG
 ATTGTTAAGCTTCTGTATATCTGCCAAATTTGTGCGGGCCGCGGATGGCGAAAAACGTTGCGATTTTTCG
 GCTTATTGTTTTCTCTTCTTGTGTTGGTTCTTCTCAGATCTTCGCCTGCAGGCTCCTCAGCCAAAACGACA
 CCCCCATCTGTCTATCCACTGGCCCTGGATCTGTGCCCCAACTAACTCCATGGTGACCCTGGGATGCCT
 GGTCAAGGGCTATTTCCCTGAGCCAGTGACAGTGACCTGGAACCTCTGGATCCCTGTCCAGCGGTGTGCAC
 ACCTTCCCAGCTGTCTGCAGTCTGACCTCTACACTCTGAGCAGCTCAGTGACTGTCCCCTCCAGCACCTG
 GCCCAGCGAGACCGTCACCTGCAACGTTGCCACCCGCCAGCAGCACCAAGGTGGACAAGAAAATTGT
 GCCCAGGGATTGTGGTTGTAAGCCTTGCATATGTACAGTCCCAGAAGTATCATCTGTCTTCATCTTCCCCC
 CAAAGCCCCAAGGATGTGCTCACCATTACTCTGACTCCTAAGGTCACGTGTGTTGTGGTAGACATCAGCAA
 GGATGATCCCCGAGGTCCAGTTCAGCTGGTTGTAGATGATGTGGAGGTGCACACAGCTCAGACGCAACCC
 CGGGAGGAGCAGTTCAACAGCACTTTCCGCTCAGTCAGTGAACCTTCCATCATGCACCAGGACTGGCTCA
 ATGGCAAGGAGCGATCGCTCACCATCACCATCACCATCACCATCACCATTAAAGGCCTATTTTCTTTAGTT
 TGAATTTACTGTTATTCGGTGTGCATTTCTATGTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTT
 TATGTAATTTAATTTCTTGTGAGCTCCTGTTTAGCAGGTCTCCCTTCAGCAAGGACACAAAAAGATTTT
 AATTTTATTAAAAAAGACCGGGAATTTCGATATCAAGCTTATCGACCTGCAGATCGTT
 CAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTT
 CTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTAT
 GATTAGAGTCCCGAATTATACATTTAATACGCGATAGAAAAAAAATATAGCGCGCAAACTAGGATAA
 ATTATCGCGCGCGGTGTCTATGTTACTAGATCTCTAGAGTCTCAAGCTTGGCGCGCCACGTGACTAG
 TGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCAACCTAATCGCCTTGC
 AGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTG
 CGCAGCCTGAATGGCGAATGCTAGAGCAGCTTGAGCTTGATCAGATTGTCGTTTCCCGCCTTCAGTTTA
 AACTATCAGTGTTTGACAGGATATATTGGCGGTAACCTAAGAGAAAAAGAGCGTTTA

Fig.1E

Bảng biểu hiện số 489 từ vùng khởi đầu 2X35S đến vùng kết thúc. H5 từ cúm A/Indonesia/5/2005 (H5N1) được gạch chân. SEQ ID NO: 5

GTCAACATGGTGGAGCAGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAA
 AGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAACCTCCTCGGATTCCATTGCCAGCTATCT
 GTCACTTTATTGTGAAGATAGTGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAA
 AGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCG
 TGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACG
 ACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAAGGGCAATTGAGACTTTTCA
 ACAAAGGGTAATATCCGGAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATA
 GTGAAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCC
 TCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAACGTTCCA
 ACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACT
 ATCCTTCGCAAGACCCTTCTCTATATAAGGAAGTTCAATTTCAATTTGGAGAGGTATTAATAATCTTAATAGG
 TTTTGATAAAAGCGAACGTGGGGAAACCCGAACCAACCTTCTTCTAAACTCTCTCATCTCTCTTAAAG
 CAAACTTCTCTCTTGTCTTCTTGTGCGTGAGCGATCTTCAACGTTGTGAGATCGTGCTTCGGCACCAGTACA
 ACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAGTGCGGCGCCATTAAATAACG
 TGTACTTGTCTATTCTTGTGCGGTGTGGTCTTGGGAAAAGAAAGCTTGCTGGAGGCTGCTGTTAGCCCCA
 TACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTG
 TTGCCTGTACTTCTTCTTCTTCTTGTGCTGATTGGTTCTATAAGAAATCTAGTATTTCTTTGAAACAGA
 GTTTTCCCGTGGTTTTGCAACTGGAGAAAAGATTGTTAAGCTTCTGTATATTCTGCCAAATTTGTGCGGC
 CCATGGAGAAAATAGTGCTTCTTCTTGCAATAGTCAGTCTTGTTAAAAGTGATCAGATTGCGATTGGTTAC
 CATGCAAAACAATTCAACAGAGCAGGTTGACACAATCATGGAAAAGAACGTTACTGTTACACATGCCCAA
 GACATACTGGAAAAGACACACAACGGGAAGCTCTGCGATCTAGATGGAGTGAAGCCTCTAATTTAAGA
 GATTGTAGTGTAGCTGGATGGCTCCTCGGGAACCAATGTGTGACGAATTCATCAATGTACCGGAATGGT
 CTTACATAGTGGAGAAGGCCAATCCAACCAATGACCTCTGTTACCCAGGGAGTTTCAACGACTATGAAGA

ACTGAAACACCTATTGAGCAGAATAAACCATTTTGAGAAAAATTCAAATCATCCCCAAAAGTTCTTGGTCC
 GATCATGAAGCCTCATCAGGAGTTAGCTCAGCATGTCCATACCTGGGAAGTCCCTCCTTTTGTAGAAATGT
 GGTATGGCTTATCAAAAAGAACAGTACATACCCAACAATAAAAGAAAAGCTACAATAATACCAACCAAGA
 GGATCTTTTGGTACTGTGGGGAATTCACCATCCTAATGATGCGGCAGAGCAGACAAGGCTATATCAAAAC
 CCAACCACCTATATTTCCATTGGGACATCAACACTAAACCAGAGATTGGTACCAAAAAATAGCTACTAGAT
 CCAAAGTAAACGGGCAAAGTGGAAGGATGGAGTTCTTCTGGACAATTTTAAACCTAATGATGCAATCA
 ACTTCGAGAGTAATGGAAATTTTCATTGCTCCAGAATATGCATACAAAATTGTCAAGAAAGGGGACTCAGC
 AATTATGAAAAGTGAATTGGAATATGGTAAGTCAACACCAAGTGTCAAACTCCAATGGGGGCGATAAA
 CTCTAGTATGCCATTCCACAACATACACCCTCTCACCATCGGGGAATGCCCAAATATGTGAAATCAAAC
 AGATTAGTCCTTGCAACAGGGCTCAGAAATAGCCCTCAAAGAGAGAGCAGAAAGAAAAAGAGAGGACT
 ATTTGGAGCTATAGCAGGTTTTATAGAGGGAGGATGGCAGGGAATGGTAGATGGTTGGTATGGGTACCA
 CCATAGCAATGAGCAGGGGAGTGGGTACGCTGCAGACAAAGAATCCACTCAAAAGGCAATAGATGGAGT
 CACCAATAAGGTCAACTCAATCATTGACAAAATGAACACTCAGTTTGAGGCCGTTGGAAGGGAATTAAT
 AACTTAGAAAGGAGAATAGAGAATTTAAACAAGAAGATGGAAGACGGGTTTCTAGATGTCTGGACTTAT
 AATGCCGAACCTCTGGTTCTCATGGAAAATGAGAGAATCTAGACTTTTCATGACTCAAATGTTAAGAACC
 TCTACGACAAGGTCGACTACAGCTTAGGGATAATGCAAGGAGCTGGGTAACGGTTGTTTCGAGTTCTA
 TCACAAATGTGATAATGAATGTATGGAAAGTATAAGAAACGGAACGTACAACATATCCGCAGTATTGAGA
 AGAAGCAAGATTAAAAAGAGAGGAAATAAGTGGGGTAAAAATTGGAATCAATAGGAACCTACCAAATAC
 TGTCAAATTTATTCAACAGTGGCGAGTCCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATG
 TGCTCCAATGGATCGTTACAATGCAGAATTTGCATTTAAAGGCCTATTTCTTTAGTTTGAATTTACTGTT
 ATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTATTTTATGTAATTTAAT
 TTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCTTCAGCAAGGACACAAAAAGATTTTAAATTTATTAATA
 AAAAAAAAAAAAAAAAAAGACCGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCA
 ATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACG
 TTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCG
 CAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCG
 GTGTCATCTATGTTACTAGAT

Fig.1F

Trình tự axit amin của H5 từ cúm A/Indonesia/5/2005 (H5N1) SEQ ID NO: 6

MEKIVLLLAIIVSLVKSDQICIGYHANNSTEQVDTIMEKNVTVTHAQDILEKTHNGKLCDL
 DGVKPLILRDCSVAGWLLGNPMCDEFINVPESYIVEKANPTNDLCYPGSFNDYEELKH
 LLSRINHFEKIQIIPKSSWSDHEASSGVSSACPYLGSPSFFRNVVWLIKKNSTYPTIKKSYN
 NTNQEDLLVLWGIHPNDAAEQTRL YQNPTTYISIGTSTLNQRLVPKIATRISKVNGQSGR
 MEFFWTILKPNDAINFESNGNFIAP EYAYKIVKKGDSAIMKSELEYGNCNTKCQTPMGAI
 NSSMPFHNIHPLTIGECPKYVKS NRLVLATGLRNSPQRESRRKKRGLFGAIAFGIEGGWQ
 GMVDGWYGYHHSNEQSGSYAADKESTQKAIDGVTNKVNSIIDKMNTQFEAVGREFNN
 LERRIENLNKKMEDGFLDVWTYNAELLVLMENERTLDFHDSNVKNLYDKVRLQLRDN
 AKELGNGCFEFYHKCDNECMESIRNGTYNYPQYSEEARLKREEISGVKLESIGTYQILSIY
 STVASSLALAIMMAGLSLWMCSNGSLQCRICI

Fig.1G trình tự nucleotit mã hóa H5 từ cúm A/Indonesia/5/2005/ (H5N1) SEQ ID NO:42)

ATGGAGAAAATAGTGCTTCTTCTTGCAATAGTCAGTCTTGTTAAAAAGTGATCAGATTGTCATTGGTTACCA
 TGCAACAATTCAACAGAGCAGGTTGACACAATCATGGAAAAGAACGTTACTGTTACACATGCCCCAAGA
 CATACTGGAAAAGACACACAACGGGAAGCTCTGCGATCTAGATGGAGTGAAGCCTCTAATTTTAAGAGA
 TTGTAGTGTAGCTGGATGGCTCCTCGGGAACCCAATGTGTGACGAATTCATCAATGTACCGGAATGGTCT
 TACATAGTGGAGAAGGCCAATCCAACCAATGACCTCTGTTACCCAGGGAGTTTCAACGACTATGAAGAA
 CTGAAACACCTATTGAGCAGAATAAACCATTTTGAGAAAAATTCAAATCATCCCCAAAAGTTCTTGGTCCG
 ATCATGAAGCCTCATCAGGAGTTAGCTCAGCATGTCCATACCTGGGAAGTCCCTCCTTTTTTAGAAATGTG
 GTATGGCTTATCAAAAAGAACAGTACATACCCAACAATAAAGAAAAGTACAATAATACCAACCAAGAG
 GATCTTTTGGTACTGTGGGGAATTCACCATCCTAATGATGCGGCAGAGCAGACAAGGTATATCAAAACC
 CAACCACCTATATTTCCATTGGGACATCAACACTAAACCAGAGATTGGTACCAAAAAATAGCTACTAGATC
 CAAAGTAAACGGGCAAAGTGGAAGGATGGAGTTCTTCTGGACAATTTTAAACCTAATGATGCAATCAA
 CTTTCGAGAGTAATGGAAATTTCAATTGCTCCAGAATATGCATACAAAATTGTCAAGAAAGGGGACTCAGCA
 ATTATGAAAAGTGAATTGGAATATGGTAACTGCAACACCAAGTGTCAAACCTCCAATGGGGGCGATAAAC
 TCTAGTATGCCATTCCACAACATACACCCTCTCACCATCGGGGAATGCCCAAATATGTGAAATCAAACA
 GATTAGTCCTTGCAACAGGGCTCAGAAATAGCCCTCAAAGAGAGAGCAGAAGAAAAAAGAGAGGACTA
 TTTGGAGCTATAGCAGGTTTTATAGAGGGAGGATGGCAGGGAATGGTAGATGGTTGGTATGGGTACCAC
 CATAGCAATGAGCAGGGGAGTGGGTACGCTGCAGACAAAGAATCCACTCAAAGGCAATAGATGGAGTC
 ACCAATAAGGTCAACTCAATCATTGACAAAATGAACACTCAGTTTGAGGCCGTTGGAAGGGAATTTAATA
 ACTTAGAAAGGAGAATAGAGAATTTAAACAAGAAGATGGAAGACGGGTTTCTAGATGTCTGGACTTATA
 ATGCCGAACTTCTGGTTCTCATGGAAAATGAGAGAACTCTAGACTTTCATGACTCAAATGTTAAGAACCT
 CTACGACAAGGTCCGACTACAGCTTAGGGATAATGCAAAGGAGCTGGGTAAACGGTTGTTTCGAGTTCTAT
 CACAAATGTGATAATGAATGTATGGAAAGTATAAGAAACGGAACGTACAACCTATCCGCAGTATTCAGAA
 GAAGCAAGATTAAAAAGAGAGGAAATAAGTGGGGTAAAAATTGGAATCAATAGGAACTTACCAAATACT
 GTCAATTTATTCAACAGTGGCGAGTTCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGT
 GCTCCAATGGATCGTTACAATGCAGAATTTGCATTTAA

B-2X35S/CPMV New Caledonia/NOS (Cấu trúc số 1261)

Fig.2A

IF-S1-M1+M2ANC.c (SEQ ID NO:7)

AAATTTGTCTGGGCCCATGAGTCTTCTAACCGAGGTCGAAACG

Fig.2B

IF-S1-4-M2ANC.r (SEQ ID NO: 8)

ACTAAAGAAAATAGGCCTTTACTCCAGCTCTATGCTGACAAAA

Fig.2C

Gen M2 được tổng hợp (tương ứng với nt 1-26 được kết hợp với 715-892 từ số truy cập Genbank DQ508860) (SEQ ID NO:9)



ATGAGTCTTCTAACCGAGGTCGAAACGCCTATCAGAAACGAATGGGGGTGCAGATGCAACGATTCAAGT
GATCCTCTTGTGTGTTGCCGCAAGTATAATTGGGATTGTGCACCTGATATTGTGGATTATTGATCGCCTTTT
TCCAAAAGCATTATCGTATCTTTAAACACGGTTTAAAAAGAGGGCCTTCTACGGAAGGAGTACCAGAGT
CTATGAGGGAAGAATATCGAGAGGAACAGCAGAATGCTGTGGATGCTGACGATGGTCATTTTGTGAGCA
TAGAGCTGGAGTAA

Fig.2D

Bảng biểu hiện số 1261 từ vùng khởi đầu 2X35S đến vùng kết thúc. M2 từ cúm A/New Caledonia/20/1999 (H1N1) được gạch chân (SEQ ID NO:10)

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAA
AGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCT
GTCACCTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAAGGAA
AGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCG
TGGAAAAAGAAGACGTTCCAACACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACG
ACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCA
ACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATA
GTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAAGGAAAGGCCATCGTTGAAGATGCC
TCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCA
ACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACT
ATCCTTCGCAAGACCTTCCTCTATATAAGGAAGTTCATTTTCATTTGGAGAGGTATTAATAATCTTAATAGG
TTTTGATAAAAGCGAACGTGGGGAAACCCGAACCAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAG
CAAACCTTCTCTTGTCTTTCTTGGGTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAAGTACA
ACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAGTGCGGCGCCATTAAATAACG
TGTACTTGTCTTATTCTTGTGCGGTGTGGTCTTGGGAAAAGAAAGCTTGCTGGAGGCTGCTGTTCAGCCCCA

TACATTACTTGTTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTG
 TTGCCTGTACTTCTTCTTCTTCTTCTTGCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTTGAAACAGA
 GTTTTCCCGTGGTTTTTCGAACTTGGAGAAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGC
CCATGAGTCTTCTAACCGAGGTCGAAACGCCTATCAGAAACGAATGGGGGTGCAGATGCAACGATTCAA
GTGATCCTCTTGTGTGTCGCAAGTATAATTGGGATTGTGCACCTGATATTGTGGATTATTGATCGCCTT
TTTTCCAAAAGCATTTATCGTATCTTTAAACACGGTTAAAAAGAGGGCCTTCTACGGAAGGAGTACCAG
AGTCTATGAGGGAAGAATATCGAGAGGAACAGCAGAATGCTGTGGATGCTGACGATGGTCATTTTGTCA
GCATAGAGCTGGAGTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTTCGGTGTGCATTTCTATGTTT
 GGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGC
 AGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTTATTAAGGACCGGG
 AATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATC
 CTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATG
 TAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGAT
 AGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGGGTGCATCTATGTTACTAGAT

Fig.2E

Trình tự axit amin của M2 từ cúm A/New Caledonia/20/1999 (H1N1) (SEQ ID NO: 11))

MSLLTEVETPIRNEWGCRCNDSSDPLVVAASIIGIVHLILWIIDRLFSKSIYRIFKHGLKRGF
 STEGVPESMREEYREEQNVAVDADDGHFVSIELE

C-2X35S/CPMV-HT/M2 Puerto Rico/NOS (Cấu trúc số 859)

Fig.3A

Gen M2 được tổng hợp (tương ứng với nt 26-51 được kết hợp với nt 740-1007 từ số truy cập Genebank¹ số EF467824) (SEQ ID NO: 12)

ATGAGTCTTCTAACCGAGGTGCGAAACGCCTATCAGAAACGAATGGGGGTGCAGATGCAACGGTTCAAGT
GATCCTCTCACTATTGCCGCAAATATCATTGGGATCTTGCACTTGACATTGTGGATTCTTGATCGTCTTTT
TTCAAATGCATTTACCGTCGCTTTAAATACGGACTGAAAGGAGGGCCTTCTACGGAAGGAGTGCCAAAGT
CTATGAGGGAAGAATATCGAAAGGAACAGCAGAGTGCTGTGGATGCTGACGATGGTCATTTTGTGAGCA
TAGAGCTGGAGTAA

Fig.3B

Bảng biểu hiện số 859 từ vùng khởi đầu 2X35S đến vùng kết thúc NOS. M2 từ Cúm A/Puerto Rico/8/1934 (H1N1) được gạch chân (SEQ ID NO: 13)

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAA
AGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAACCTCCTCGGATTCCATTGCCAGCTATCT
GTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAAGGAA
AGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCG
TGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACG
ACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCA
ACAAAGGGTAATATCCGGAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATA
GTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAAGGCCATCGTTGAAGATGCC
TCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCA
ACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCCT
ATCCTTCGCAAGACCCCTTCCTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTTAAATCTTAATAGG
TTTTGATAAAAGCGAACGTGGGGAACCCGAACCAACCTTCTTCTAACTCTCTCTCATCTCTCTTAAAG
CAAACCTCTCTCTTGTCTTTCTTGGCTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACA
ACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAAGTGCAGCGCCATTAAATAACG
TGTACTTGTCTTATTCTTGTCTGGTGTGGTCTTGGGAAAAGAAAGCTTGTCTGGAGGCTGCTGTTTCAAGCCCCA
TACATTACTTGTACGATTCTGCTGACTTTTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTG
TTGCCTGTACTTCTTTCTTCTTCTTGTCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTTGAAACAGA
GTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCTGGGC
CCATGAGTCTTCTAACCGAGGTGCGAAACGCCTATCAGAAACGAATGGGGGTGCAGATGCAACGGTTCAA
GTGATCCTCTCACTATTGCCGCAAATATCATTGGGATCTTGCACTTGACATTGTGGATTCTTGATCGTCTTT
TTTTCAAATGCATTTACCGTCGCTTTAAATACGGACTGAAAGGAGGGCCTTCTACGGAAGGAGTGCCAAA
GTCTATGAGGGAAGAATATCGAAAGGAACAGCAGAGTGCTGTGGATGCTGACGATGGTCATTTTGTGAG
CATAGAGCTGGAGTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTG
GTGAGCGGTTTTCTGTGCTCAGAGTGTGTTATTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCA
GGTCGTCCTTCAGCAAGGACACAAAAAGATTTTAAATTTATTAATAAAAAAAAAAAAAAAAAAGACCGGGA
ATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCC
TGTGCGCGTCTTGCGATGATTATCATATAATTTCTGTTGTAATTACGTTAAGCATGTAATAATTAACATGT
AATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATA
GAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGAT

Fig. 3C

Trình tự axit amin của M2 từ cúm A/Puerto Rico/8/1934 (H1N1) (SEQ ID NO: 14)

MSLLTEVETPIRNEWGCRCNGSSDPLTIAANIIGILHLTLWILDRLFFKCIYRRFKYGLKGG
PSTEGVPKSMREEYRKEQQSAVDADDGHFVSIELE

D-2X35S/CPMV-HT/PDISP/H1 California/NOS (Cấu trúc số 484)

Fig.4A

IF-H1A-C-09.s2+4c (SEQ ID NO: 15)

TCTCAGATCTTCGCCGACACATTATGTATAGGTTATCATGCGAACA

Fig.4B

IF-H1A-C-09.s1-4r (SEQ ID NO: 16)

ACTAAAGAAAATAGGCCTTTAAATACATATTCTACACTGTAGAGACCCA

Fig.4C

Gen H1 được tổng hợp (số truy cập Genbank FJ966974)(SEQ ID NO: 17)

ATGAAGGCAATACTAGTAGTTCTGCTATATACATTTGCAACCGCAAATGCAGACACATTATGTATAGGTT
ATCATGCGAACAATTCAACAGACACTGTAGACACAGTACTAGAAAAGAATGTAACAGTAACACACTCTG
TTAACCTTCTAGAAGACAAGCATAACGGGAACTATGCAAATAAGAGGGGTAGCCCCATTGCATTTGG
GTAAATGTAACATTGCTGGCTGGATCCTGGGAAATCCAGAGTGTGAATCACTCTCCACAGCAAGCTCATG
GTCCTACATTGTGGAACACCTAGTTCAGACAATGGAACGTGTTACCCAGGAGATTTTCATCGATTATGAG
GAGCTAAGAGAGCAATTGAGCTCAGTGTATCATTTGAAAGGTTTGAGATATTCCCCAAGACAAGTTCAT
GGCCCAATCATGACTCGAACAAGGTGTAACGGCAGCATGTCCTCATGCTGGAGCAAAAAGCTTCTACA
AAAATTTAATATGGCTAGTTAAAAAAGGAAATTCATACCCAAAGCTCAGCAAAATCCTACATTAATGATAA
AGGGAAAGAAATCCTCGTGCTATGGGGCATTACCATCCATCTACTAGTGCTGACCAACAAAGTCTCTAT
CAGAATGCAGATGCATATGTTTTGTGGGGTCATCAAGATACAGCAAGAAGTTCAAGCCGGAAATAGCA
ATAAGACCCAAAGTGAGGGATCAAGAAGGGAGAATGAACTATTACTGGACACTAGTAGAGCCGGGAGA
CAAAATAACATTGCAAGCAACTGGAAATCTAGTGGTACCGAGATATGCATTTCGCAATGGAAAGAAATGC
TGGATCTGGTATTATCATTTTCAGATACACCAAGTCCAGATTGCAATACAACCTTGTCAAACACCCAAGGGT
GCTATAAACACCAAGCCTCCCATTTTCAGAATATACATCCGATCACAATTGGAAAATGTCCAAAATATGTAA
AAAGCACAAAATTGAGACTGGCCACAGGATTGAGGAATATCCCGTCTATTCAATCTAGAGGCCTATTGG
GGCCATTGCCGTTTCATTGAAGGGGGGTGGACAGGGATGGTAGATGGATGGTACGGTTATCACCATCA
AAATGAGCAGGGGTCAGGATATGCAGCCGACCTGAAGAGCACACAGAATGCCATTGACGAGATTACTAA
CAAAGTAAATTCTGTTATTGAAAAGATGAATACACAGTTCACAGCAGTAGGTAAAGAGTTCAACCACCTG
GAAAAAGAATAGAGAATTTAAATAAAAAAGTTGATGATGGTTTCCTGGACATTTGGACTTACAATGCC
GAACTGTTGGTTCTATTGGAAAATGAAAGAACTTTGGACTACCACGATTCAAATGTGAAGAACTTATATG
AAAAGGTAAGAAGCCAGCTAAAAACAATGCCAAGGAAATTGGAAACGGCTGCTTTGAATTTTACCACA
AATGCGATAACACGTGCATGGAAAGTGTCAAAAATGGGACTTATGACTACCCAAAATACTCAGAGGAAG
CAAAATTAAACAGAGAAGAAATAGATGGGGTAAAGCTGGAATCAACAAGGATTTACCAGATTTTGGCGA
TCTATTCAACTGTCGCCAGTTCATTGGTACTGGTAGTCTCCCTGGGGGCAATCAGTTTCTGGATGTGCTCT
AATGGGTCTCTACAGTGTAGAATATGTATTTAA

Fig.4D

Bản đồ của cấu trúc 1192. Các vị trí enzyme giới hạn SacII và StuI được sử dụng để kéo dài plasmid được giảm độc lực trên bản đồ này

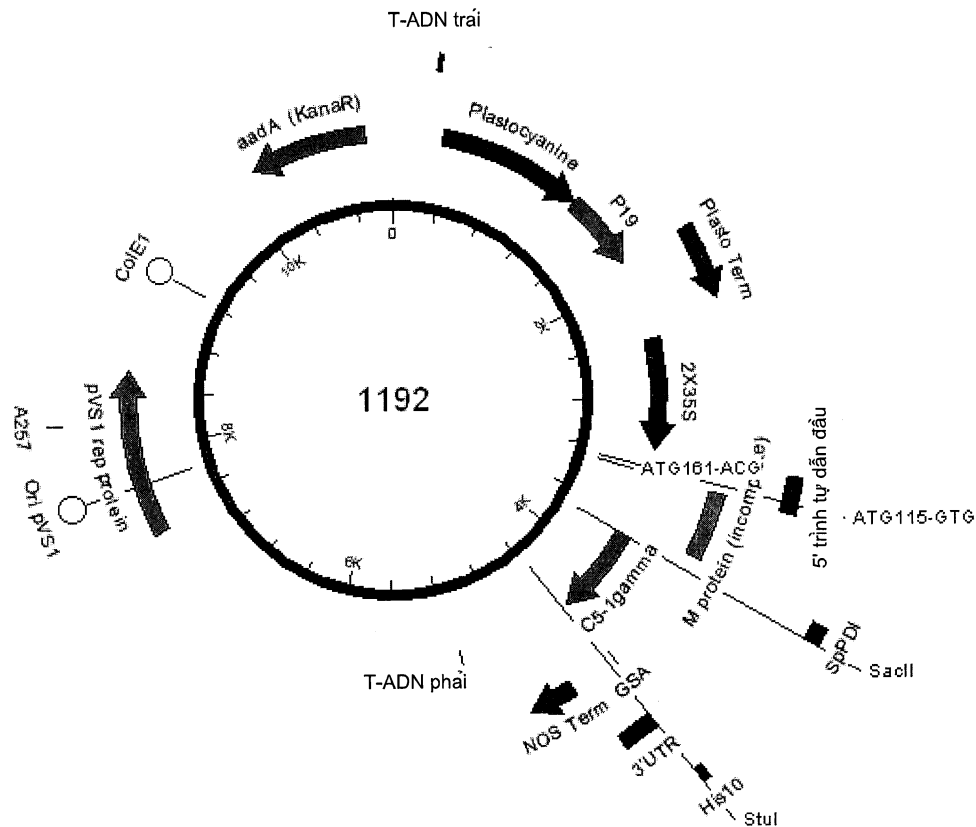


Fig.4E

Cấu trúc 1192 từ các bờ ADN vận chuyển từ trái sang phải (được gạch chân). 2X35S/CPMV-HT/PDISP/NOS có bằng biểu hiện chất ức chế bất hoạt

TGGCAGGATATATTGTGGTGTAACAAATTGACGCTTAGACAACCTTAATAACACATTGCGGACGTTTAA
ATGTACTGAATTAACGCCGAATCCCGGGCTGGTATATTATATGTTGTCAAATAACTCAAAAAACATAAA
AGTTTAAGTTAGCAAGTGTGTACATTTTACTTGAACAAAAATATTCACCTACTACTGTTATAAATCATT
TTAAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTGATATTTTGACAACAATTTTGTG
CAACATTTGAGAAAATTTGTGTTCTCTCTTTTCATTGGTCAAAAAACAATAGAGAGAGAAAAAGGAAGA
GGGAGAATAAAAAACATAATGTGAGTATGAGAGAGAAAGTTGTACAAAAGTTGTACCAAAATAGTTGTAC
AAATATCATTGAGGAATTTGACAAAAGCTACACAAATAAGGGTTAATTGCTGTGAAAATAAATAAGGATGA
CGCATTAGAGAGATGTACCATTAGAGAATTTTGGCAAGTCATTAAGAAAGAAAGAAATAAATTATTTTAA
AATTAAGAGTTGAGTCATTTGATTAAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGT
TGTATTAGTAATTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCCTATATATTGCCCATAGA
GTCAGTTAACTCATTTTATATTTTCATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAA
AAAAAACGGTATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTAGGAGGATA

ACATCCAATCCAACCAATCACAACAATCCTGATGAGATAACCCACTTTAAGCCCACGCATCTGTGGCACA
 TCTACATTATCTAAATCACACATTCTTCCACACATCTGAGCCACACAAAAACCAATCCACATCTTTATCAC
 CCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAG
 ACTAATTAATTAATTAATCATCTTGAGAGAAAATGGAACGAGCTATACAAGGAAACGACGCTAGGGAAC
 AAGCTAACAGTGAACGTTGGGATGGAGGATCAGGAGGTACCACTTCTCCCTTCAAACCTCTGACGAAAG
 TCCGAGTTGGACTGAGTGGCGGTACATAACGATGAGACGAATTCGAATCAAGATAATCCCTTGGTTTC
 AAGGAAAGCTGGGGTTTCGGGAAAGTTGTATTTAAGAGATATCTCAGATACGACAGGACGGAAGCTTCA
 CTGCACAGAGTCTTGGATCTTGGACGGGAGATTGCGTTAACTATGCAGCATCTCGATTTTTCGGTTTCGA
 CCAGATCGGATGTACCTATAGTATTCGGTTTCGAGGAGTTAGTATCACCGTTTCTGGAGGGTCGCGAACT
 CTTCAGCATCTCTGTGAGATGGCAATTCGGTCTAAGCAAGAACTGCTACAGCTTGCCCAATCGAAGTGG
 AAAGTAATGTATCAAGAGGATGCCCTGAAGGTACTCAAACCTTCGAAAAAGAAAGCGAGTAAGTTAAAA
 TGCTTCTTCGTCTCCTATTTATAATATGGTTTGTATTGTTAATTTTGTCTTGTAGAAGAGCTTAATTAAT
 CGTTGTTGTTATGAAATACTATTTGTATGAGATGAACTGGTGTAAATGTAATTCATTACATAAGTGGAGTC
 AGAATCAGAATGTTTCTCCATAACTAAGTACATGAAGACCTGCCGCGTACAATTGTCTTATATTGTA
 ACAACTAAAATTGAACATCTTTTCCACAACCTTTATAAGTGGTTAATATAGTCAAAATATATGGTCAAGT
 CAATAGATTAAATAATGGAATATCAGTTATCGAAATTCATTAACAATCAACTTAACGTTATTAAGTACTA
 ATTTATATATCCTCCCTTTGATAAATGATAGTACACCAATTAGGAAGGAGCATGCTCGCCTAGGAGATTG
 TCGTTTCCGCGCTTCAGTTGCAAGCTGCTCTAGCCGTGTAGCCAATACGCAAACCGCTCTCCCGCGCG
 TTGGGAATTACTAGCGCGTGTGACAAGCTTGCATGCCGTCAACATGGTGGAGCAGACACACTTGTCT
 ACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAA
 TATCCGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGA
 AGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGT
 GGTCCCAAAGATGGACCCCAACACGAGGAGCATCGTGGAAAAAGAACGTTCCAACACAGTCTTCA
 AAGCAAGTGGATTGATGTGATAACATGGTGGAGCAGACACACTTGTCTACTCCAAAAATATCAAAGAT
 ACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGAAACCTCCTCGGA
 TTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCC
 ATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCCTCTGCCGACAGTGGTCCCAAAGATGGACCCC
 CACCCACGAGGAGCATCGTGGAAAAAGAACGTTCCAACACAGTCTTCAAAGCAAGTGGATTGATGTG
 ATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTCCTCTATATAAGG
 AAGTTCATTTCATTTGGAGAGGTATTAAATCTTAATAGGTTTGTATAAAGCGAACGTTGGGGAACCCG
 AACCAAACCTTCTTCTAACTCTCTCTCATCTCTCTTAAAGCAAACTTCTCTTGTCTTTCTTGCCTGAGC
 GATCTTCAACGTTGTGAGATCGTGCTTCGGCACCAGTACAACGTTTCTTTCACTGAAGCGAAATCAAAG
 ATCTCTTTGTGGACACGTAGTGGCGGCCATTAAATAACGTGTACTTGTCTATTCTTGTGCGGTGTGGTCT
 TGGGAAAAGAAAGCTTGTGAGGCTGCTGTTACGCCCCATACATTACTTGTACGATTCTGCTGACTTTC
 GCGGGGTGCAATATCTCTACTTCTGCTTACGAGGTATTGTTGCCTGTACTTCTTCTTCTTCTTGTCTG
 ATTGTTCTATAAGAAATCTAGTATTTCTTTGAAACAGAGTTTCCCGTGGTTTTCGAACTTGGAGAAAG
 ATTGTTAAGCTTCTGTATATTCTGCCAAATTTGTGCGGCCCATGGCGAAAAACGTTGCGATTTTCGGCTT
 ATTGTTTCTCTTCTTGTGTTGGTTCCCTTCTCAGATCTTCGCCGCGGCTCCTCAGCCAAAACGACACCCCA
 TCTGTCTATCCACTGGCCCCGGATCTGCTGCCCAAACCTCACTCCATGGTGACCCTGGGATGCCCTGGTCAA
 GGGCTATTTCCCTGAGCCAGTGACAGTGACCTGGAACCTCTGGATCCCTGTCCAGCGGTGTGCACACCTTC
 CCAGCTGTCTGACGTCTGACCTCTACACTCTGAGCAGCTCAGTGACTGTCCCTCCAGCACCTGGCCCA
 GCGAGACCGTCACCTGCAACGTTGCCACCCGCGCCAGCAGCACCAGGTGGACAAAGAAATTTGTGCCCA
 GGGATTGTGGTTGTAAGCCTTGCATATGTACAGTCCCAGAAGTATCATCTGTCTTCTTCTTCTTCTTCTG
 CCAAGGATGTGCTCACCATTACTCTGACTCCTAAGGTACAGTGTGTTGTGGTAGACATCAGCAAGGATG
 ATCCCGAGGTCCAGTTCAGCTGGTTTGTAGATGATGTGGAGGTGCACACAGCTCAGACGCAACCCCGGGA
 GGAGCAGTTCAACAGCACTTCCGCTCAGTCACTGAACTTCCCATCATGCACCAGGACTGGCTCAATGGC
 AAGGAGCGATCGCTCACCATCACCATCACCATCACCATCACCATTAAAGGCCTATTTCTTTAGTTTGAAT
 TTAAGTTATTTCGGTGTGCATTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTATTTTATGT
 AATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCTCCCTTCAGCAAGGACACAAAAAGATTTTAAATTT
 TATTAATAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAAC
 ATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGCTTTCGCGATGATTATCATATAATTTCTGTT
 GAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTATGATTA
 GAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTAT
 CGCGCGCGGTGTCATCTATGTTACTAGATCTCTAGAGTCTCAAGCTTGGCGCGCCACGCTGACTAGTGGC

ACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCAACCTAATCGCCTTGACGCA
 CATCCCCCTTTGCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCA
 GCCTGAATGGCGAATGCTAGAGCAGCTTGAGCTTGATCAGATTGTCGTTTCCCGCCTTCAGTTAAACT
 ATCAGTGTTTGACAGGATATATTGGCGGGTAAACCTAAGAGAAAAAGAGCGTTTA

Fig.4F

Bảng biểu hiện số 484 từ vùng khởi đầu 2X35S đến vùng kết thúc NOS. PDISP/H1 từ cúm A/
 California/7/2009 (H1N1) được gạch chân. (SEQ ID NO: 19)

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAA
 AGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAACCTCCTCGGATTCCATTGCCAGCTATCT
 GTCACCTTTATTGTGAAGATAGTGGAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAA
 AGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCG
 TGGAAAAAGAAGACGTTCCAACACGCTCTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACG
 ACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCA
 ACAAAGGGTAATATCCGGAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATA
 GTGGAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCC
 TCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCA
 ACCACGCTCTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACT
 ATCCTTCGCAAGACCCCTTCCTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTAATACTTAATAGG
 TTTTGATAAAAGCGAACGTTGGGGAACCCGAACCAACCTTCTTCTAACTCTCTCTCATCTCTCTTAAAG
 CAAACTTCTCTCTTGTCTTTCTTGCCTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACA
 ACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAAGTGGCGGCCATTAAATAACG
 TGTACTTGTCTTCTTGTCTGGTGTGGTCTTGGGAAAAAGAAAGCTTGTCTGGAGGCTGCTGTTACGCCCCA
 TACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTG
 TTGCCGTGACTTCTTTCTTCTTCTTCTGCTGATTGGTCTCTATAAGAAATCTAGTATTTTCTTTGAAACAGA
 GTTTTCCCGTGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCTGGGC
 CCATGGCGAAAAACGTTGCGATTTCGGCTTATTGTTTTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCG
 CCGACACATTATGTATAGGTTATCATGCGAACAATTCAACAGACACTGTAGACACAGTACTAGAAAAAG
 ATGTAACAGTAACACACTCTGTAAACCTTCTAGAAGACAAGCATAACGGGAAACTATGCAACTAAGAG
 GGGTAGCCCCATTGCATTTGGGTAAATGTAACATTGCTGGCTGGATCCTGGGAAATCCAGAGTGTGAATC
 ACTCTCCACAGCAAGCTCATGGTCTACATTGTGGAACACCTAGTTTACAGACAATGGAACGTGTTACCCA
 GGAGATTTTCATCGATTATGAGGAGCTAAGAGAGCAATTGAGCTCAGTGTATCATTTGAAAGGTTTGAGA
 TATCCCAAGACAAGTTCATGGCCCAATCATGACTCGAACAAGGTGTAACGGCAGCATGTCCTCATGC
 TGGAGCAAAAAGCTTCTACAAAAATTAATATGGCTAGTTAAAAAGGAAATTCATACCCAAAGCTCAG
 CAAATCCTACATTAATGATAAAGGGAAGAGTCTCGTGTATGGGGCATTACCATCCATCTACTAGT
 GCTGACCAACAAAGTCTCTATCAGAAATGCAGATGCATATGTTTTGTGGGGTCATCAAGATACAGCAAGA
 AGTTCAAGCCGGAATAGCAATAAGACCCAAAGTGAGGGATCAAGAAGGGAGAATGAACTATTACTGG
 ACAGTAGTAGAGCCGGGAGACAAAAAATCAATTCGAAGCAACTGGAATCTAGTGGTACCGAGATATGCA
 TTCGCAATGGAAGAAATGCTGGATCTGGTATTATCATTTTTCAGATACACCAAGTCCACGATTGCAATACAA
 CTTGTCAAACACCCAAAGGGTGCTATAAACACACAGCCTCCCATTTTTCAGAAATATACATCCGATCACAATTGG
 AAAATGTCCAAAATATGTAAAAAGCACAAAAATTGAGACTGGCCACAGGATTGAGGAATATCCCGTCTAT
 TCAATCTAGAGGCCATTTTGGGGCCATTGCCGGTTTCATTGAAGGGGGGTGGACAGGGATGGTAGATGGA
 TGGTACGGTTATCACCATCAAAATGAGCAGGGGTGAGGATATGCAGCCGACCTGAAGAGCACACAGAAT
 GCCATTGACGAGATTACTAACAAAGTAAATTCTGTTATTGAAAAGATGAATACACAGTTCACAGCAGTAG
 GTAAAGAGTTCAACCACCTGGAAAAAGAAATAGAGAATTTAAATAAAAAAGTTGATGATGGTTTCTCGG
 ACATTTGGAATTACAATGCCGAACCTGTTGGTCTATTGGAAAAATGAAAGAACTTTGGACTACCACGATT
 AAATGTGAAGAACTTATATGAAAAGGTAAGAAAGCCAGCTAAAAACAATGCCAAGGAAATTTGGAACG
 GCTGCTTTGAATTTTACCACAAATGCGATAACACGTGCATGGAAAGTGTCAAAAATGGGACTTATGACTA
 CCAAAATACTCAGAGGAAGCAAAATTAACAGAGAGAAATAGATGGGGTAAAGCTGGAATCAACAA
 GGATTTACCAGATTTTGGCGATCTATTCAACTGTCGCCAGTTCATTGGTACTGGTAGTCTCCCTGGGGCA

ATCAGTTTCTGGATGTGCTCTAATGGGTCTCTACAGTGTAGAATATGTATTTAAAGGCCTATTTTCTTTAG
 TTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTAT
 TTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATT
 TTAATTTTATTAATAAAAAAAAAAAAAAAAAAGACCGGAATTCGATATCAAGCTTATCGACCTGCAGATCGT
 TCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATT
 TCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTA
 TGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATA
 AATTATCGCGCGCGGTGTCATCTATGTTACTAGAT

Fig.4G

Trình tự axit amin của PDISP-H1 từ cúm A/California/7/2009 (H1N1) (SEQ ID NO:20)

MAKNVAIFGLLFSLLVLVPSQIFADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDK
 HNGKLCKLRGVAPLHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFI
 DYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAACPHAGAKSFYKNLIWLVKKGN
 SYPKLSKSYINDKGKEVLVLWGIHPSTSADQQSLYQNADAYVFGSSRYSKKFKPEIAI
 RPKVRDQEGRMNYYWTLVEPGDKITFEATGNLVPVRYAFAMERNAGSGIIISDTPVHDC
 NTTCTPKGAINSTLFPQNIHPITIGKCPKYVKSTKLRLATGLRNIPSIQSRGLFGAIAGFIE
 GGWTGMVDGWYGYHHQNEQSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGK
 EFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERLTDYHDSNVKNLYEKVRSQK
 NNAKEIGNGCFEFYHKCDNTCMESVKNNGTYDYPKYSEEAKLNREEIDGVKLESTRIYQIL
 AIYSTVASSLVLVVSLSLGAISFWMCSNGLQCRICI

E-2X35S/CPMV-HT/PDISP/H3 Perth/NOS (Cấu trúc số 1019)

Fig.5A

IF-S2+S4-H3 Per.c (SEQ ID NO: 21)

TCTCAGATCTTCGCCCAAAAACCTTCCTGGAAATGACAACA

Fig.5B

IF-S1a4-H3 Per.r (SEQ ID NO: 22)

ACTAAAGAAAATAGGCCTTCAAATGCAAATGTTGCACCTAATGTT

Fig.5C

Gen H3 được tổng hợp (tương ứng với nt 26-1726 từ số truy cập Genbank GQ293081) (SEQ ID NO: 23)

ATGAAGACTATCATTGCTTTGAGCTACATTCTATGTCTGGTTTTTCGCTCAAAAACCTTCCTGGAAATGACAA
CAGCACGGCAACGCTGTGCCTTGGGCACCATGCAGTACCAAACGGAACGATAGTGAAAACAATCACGAA
TGACCAAATTGAAGTTACTAATGCTACTGAGCTGGTTCAGAGTTCCTCAACAGGTGAAATATGCGACAGT
CCTCATCAGATCCTTGATGGAAAAAAGTGCACACTAATAGATGCTCTATTGGGAGACCCTCAGTGTGATG
GCTTCCAAAATAAGAAATGGGACCTTTTTGTTGAACGCAGCAAAGCCTACAGCAACTGTTACCCTTATGA
TGTGCCGGATTATGCCTCCCTTAGGTCACTAGTTGCCTCATCCGGCACACTGGAGTTTAAACAATGAAAGCT
TCAATTGGACTGGAGTCACTCAAAACGGAACAAGCTCTGCTTGCATAAGGAGATCTAAAAACAGTTTCTT
TAGTAGATTGAATTGGTTGACCCACTTAACTTCAAATACCCAGCATTGAACGTGACTATGCCAAACAAT
GAACAATTTGACAAATTGTACATTTGGGGGGTTCACCAACCGGGTACGGACAAAGACCAAATCTTCTGT
ATGCTCAAGCATCAGGAAGAATCACAGTCTCTACCAAAAGAAGCCAACAAACCGTAAGCCCGAATATCG
GATCTAGACCCAGAGTAAGGAATATCCCTAGCAGAATAAGCATCTATTGGACAATAGTAAAACCGGGAG
ACATACTTTTGATTAACAGCACAGGGAATCTAATTGCTCCTAGGGGTTACTTCAAAATACGAAGTGGGAA
AAGCTCAATAATGAGATCAGATGCACCCATTGGCAAATGCAATTCTGAATGCATCACTCCAAATGGAAGC
ATTCCCAATGACAAACCATTCAAAATGTAAACAGGATCACATACGGGGCCTGTCCAGATATGTAAAGC
AAAACACTCTGAAATTGGCAACAGGGATGCGAAATGTACCAGAGAAACAACTAGAGGCATATTTGGCG
CAATCGCGGGTTTCATAGAAAATGGTTGGGAGGAATGGTGGATGGTTGGTACGGTTTCAGGCATCAAA
ATTCTGAGGGAAGAGGACAAGCAGCAGATCTCAAAAGCACTCAAGCAGCAATCGATCAAATCAATGGGA
AGCTGAATAGATTGATCGGGAAAACCAACGAGAAATTCATCAGATTGAAAAAGAATTCTCAGAAGTCG
AAGGGAGAATTCAGGACCTTGAGAAATATGTTGAGGACACTAAAATAGATCTCTGGTCATACAACGCGG
AGCTTCTTGTGCCCCTGGAGAACCAACATAACAATTGATCTAACTGACTCAGAAATGAACAACTGTTTGA
AAAAACAAAGAAGCAACTGAGGGAAAATGCTGAGGATATGGGCAATGGTTGTTTCAAAATATACCACAA
ATGTGACAATGCCTGCATAGGATCAATCAGAAATGGAACCTATGACCACGATGTATACAGAGATGAAGC
ATTAAACAACCGGTTTCAGATCAAGGGAGTTGAGCTGAAGTCAGGGTACAAAGATTGGATCCTATGGATT
TCCTTTGCCATATCATGTTTTTGTGTTGCTTTGTTGGGGTTCATCATGTGGGCCTGCCAAAAAGGC
AACATTAGGTGCAACATTTGCATTTGA

Fig.5D

Bảng biểu hiện số 1019 từ vùng khởi đầu 2X35S đến vùng kết thúc.PDISP/H3 từ cúm A/Perth/16/2009 (H3N2) được gạch chân. (SEQ ID NO: 24)

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAA
 AGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCAGCTATCT
 GTCACTTTATTGTGAAGATAGTGAAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAA
 AGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCG
 TGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACG
 ACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCA
 ACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATA
 GTGAAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCC
 TCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCA
 ACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCCT
 ATCCTTCGCAAGACCCCTTCTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTAATCTTAATAGG
 TTTTGATAAAAGCGAACGTGGGGAACCCGAACCAACCTTCTTCTAACTCTCTCATCTCTCTTAAAG
 CAACTTCTCTCTTGTCTTCTTGTGCGTGAGCGATCTTCAACGTTGTGAGATCGTGCTTCGGCACCAGTACA
 ACGTTTCTTCTTCACTGAAGCGAAATCAAAGATCTCTTGTGGACACGTAAGTGGCGGCCATTAAATAACG
 TGTACTTGTCTTCTTGTGCGGTGTGGTCTTGGGAAAAAGAAAGCTTGTGGAGGCTGCTGTTGAGCCCCA
 TACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTG
 TTGCTGTACTTCTTCTTCTTCTTCTTGTGCTGATTGGTCTATAAGAAATCTAGTATTTCTTTGAAACAGA
 GTTTTCCCGTGGTTTTGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTGCGGC
 CCATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCG
CCAAAAACTCTCGAAATGACAACAGCACGGCAACGCTGTGCCTTGGGCACCATGCAGTACCAAAACG
GAACGATAGTGAAAACAATCACGAATGACCAAAATTGAAGTTACTAATGCTACTGAGCTGGTTCAGAGTTC
CTCAACAGGTGAAAATATGCGACAGTCCTCATCAGATCCTTGATGGAAAAAAGTGCACACTAATAGATGCT
CTATTGGGAGACCTCAGTGTGATGGCTTCCAAAAATAAGAAATGGGACCTTTTGTGTAACGCAGCAAAG
CCTACAGCAACTGTTACCCTTATGATGTGCCGATTATGCCTCCCTTAGGTCACTAGTTGCCCTCATCCGGC
ACACTGGAGTTAACAATGAAAGCTTCAATTGGACTGGAGTCACTCAAAACGGAACGAGCTCTGCTTGCA
TAAGGAGATCTAAAAAACAGTTTCTTTAGTAGATTGAATTGGTTGACCCACTTAAACTCAAAATACCCAGC
ATTGAACGTGACTATGCCAAACAATGAACAATTGACAAATTGTACATTGGGGGGTTACCAACCCGGGT
ACGGCAAAAGACCAAAATCTCCTGTATGCTCAAGCATCAGGAAGAATCACAGTCTCTACCAAAAGAAGC
CAACAAACCGTAAGCCCGAATATCGGATCTAGACCCAGAGTAAGGAATATCCCTAGCAGAATAAGCATC
TATTGGACAATAGTAACCCGGGAGACATACTTTTGATTAAACAGCACAGGGAATCTAATTGCTCCTAGGG
GTTACTTCAAAAATACGAAGTGGGAAAAGCTCAATAATGAGATCAGATGCACCCATTGGCAAATGCAATT
CTGAATGCATCACTCCAAATGGAAGCATTCCCAATGACAAACCATTCAAAAATGTAACAGGATCACATA
CGGGGCCTGTCCCAGATATGTTAAGCAAAACACTCTGAAATTGGCAACAGGGATGCGAAATGTACCAGA
GAAACAAACTAGAGGCATATTTGGCGCAATCGCGGGTTTCATAGAAAAATGGTTGGGAGGGAATGGTGGA
TGGTTGGTACGGTTTCAGGCATCAAAATCTGAGGGAAGAGGACAAGCAGCAGATCTCAAAAGCACTCA
AGCAGCAATCGATCAAAATCAATGGGAAGCTGAATAGATTGATCGGGAAAACCAACGAGAAATTCATCA
GATTGAAAAAGAATTCTCAGAAGTCGAAGGGAGAATTGAGGACCTTGAGAAATATGTTGAGGACACTAA
AATAGATCTCTGGCATACAACGCGGAGCTTCTTGTGCCCCTGGAGAACCAACATACAATTGATCTAACT
GACTCAGAAATGAACAACTGTTTGAACAAACAAAGAAGCAACTGAGGGAAAAATGCTGAGGATATGGG
CAATGGTTGTTTCAAAATATACCACAAATGTGACAATGCCTGCATAGGATCAATCAGAAATGGAACCTTAT
GACCACGATGTACAGAGATGAAGCATTAACAACACCGGTTTCAGATCAAGGGAGTTGAGCTGAAGTCA
GGGTACAAAGATTGGATCCTATGGATTTCTTTGCCATATCATGTTTTTGTCTTGTGTTGCTTTGTGGGG
TTCATCATGTGGGCCTGCCAAAAAGGCAACATTAGGTGCAACATTGCAATTTGAAGGCCTATTTCTTTAG
 TTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTAT
 TTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTCAGCAAGGACACAAAAAGATT
 TTAATTTTATTAACAAAAAAGACCGGGAATTTCGATATCAAGCTTATCGACCTGCAGATCGT
 TCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGCTTTCGCGATGATTATCATATAATT
 TCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTA

TGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATA
AATTATCGCGCGCGGTGTCATCTATGTTACTAGAT

Fig.5E

Trình tự axit amin của PDISP/H3 từ cúm A/Perth/16/2009 (H3N2) (SEQ ID NO:25)

MAKNVAIFGLLFSLVLVPSQIFAQKLPGNDNSTATLCLGHHAVPNGTIVKTITNDQIEVT
NATELVQSSSTGEICDSPHQILDGKNCTLIDALLGDPQCDGFQNKKWDLFVERSKAYSNC
YPYDVPDYASLRSLVASSGTLEFNNEFNFNWTGVTQNGTSSACIRRSKNSFFSRLNWLTHL
NFKYPALNVTMPNNEQFDKLYIWGVHHPGTDKDQIFLYAQASGRITVSTKRSQQTVSPNI
GSRPRVRNIPSRISYWTIVKPGDILLINSTGNLIAPRGYFKIRSGKSSIMRSDAPIGKCNSEC
ITPNGSIPNDKPFQNVNRITYGACPRYVKQNTLKLATGMRNVPEKQTRGIFGAIAGFIENG
WEGMVDGWYGFRHQNSEGRGQAADLKSTQAAIDQINGKLNRLIGKTNEKFHQIEKEFS
EVEGRIQDLEKYVEDTKIDLWSYNAELLVALENQHTIDLTDSMNKLFKTKKQLRENA
EDMGNGCFKIYHKCDNACIGSIRNGTYDHDVYRDEALNNRFQIKGVELKSGYKDWILWI
SFAISCFLLCVALLGFIMWACQKGNIRCNICI

F-2X35S/CPMV-HT/PDSP/HA B Brisbane/NOS (Cấu trúc số 1029)

Fig.6A

IF-S2+S4-B Bris.c (SEQ ID NO: 26)

TCTCAGATCTTCGCCGATCGAATCTGCACTGGAATAACAT

Fig.6B

IF-S1a4-B Bris.r (SEQ ID NO: 27)

ACTAAAGAAAATAGGCCTTTATAGACAGATGGAGCAAGAAACA

Fig.6C

Gen HA B Brisbane được tổng hợp (tương ứng với nt 34-1791 từ số truy cập Genbank FJ766840) (SEQ ID NO: 28)

ATGAAGGCAATAATTGTACTACTCATGGTAGTAACATCCAATGCAGATCGAATCTGCACTGGAATAACAT
CGTCAAACCTACCCACATGTCGTCAAACTGCTACTCAAGGGGAGGTCAATGTGACTGGTGTAAATACCACT
GACAACAACCCACCAAATCTCATTTTGCAAATCTCAAAGGAACAGAAACCAGGGGGAAACTATGCC
AAAATGCCTCAACTGCACAGATCTGGACGTAGCCTTGGGCAGACCAAAATGCACGGGGAAAATACCCTC
GGCAAGAGTTTCAATACTCCATGAAGTCAGACCTGTTACATCTGGGTGCTTTCCTATAATGCACGACAGA
ACAAAAATTAGACAGCTGCCTAACCTTCTCCGAGGATACGAACATATCAGGTTATCAACCCATAACGTTA
TCAATGCAGAAAATGCACCAGGAGGACCCTACAAAATTGGAACCTCAGGGTCTTGCCCTAACATTACCA
ATGGAAACGGATTTTTCGCAACAATGGCTTGGGCCGTCCTCAAAAAACGACAAAAACAAAACAGCAACAA
ATCCATTAACAATAGAAGTACCATACATTTGTACAGAAGGAGAAAGACCAAATTACCGTTTGGGGGTTC
CTCTGACAACGAGACCCAAATGGCAAAAGCTCTATGGGGACTCAAAGCCCCAGAAAGTTCACCTCATCTGCC
AACGGAGTGACCACACATTACGTTTCACAGATTGGTGGCTTCCCAAATCAAACAGAAGACGGAGGACTA
CCACAAAGTGGTAGAATTGTTGTTGATTACATGGTGCAAAAATCTGGGAAAACAGGAACAATTACCTATC
AAAGGGGTATTTTATTGCCTCAAAAGGTGTGGTGCGCAAGTGGCAGGAGCAAGGTAATAAAAGGATCCT
TGCCTTTAATTGGAGAAGCAGATTGCCTCCACGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTA
CACAGGGGAACATGCAAAGGCCATAGGAAATTGCCCAATATGGGTGAAAACACCCTTGAAAGCTGGCCAA
TGGAACCAATATAGACCTCCTGCAAAACTATTAAAGGAAAGGGGTTTCTTCGGAGCTATTGCTGGTTTC
TTAGAAGGAGGATGGGAAGGAATGATTGCAGTTGGCACGGATACACATCCCATGGGGCACATGGAGTA
GCGGTGGCAGCAGACCTTAAGAGCACTCAAGAGGCCATAAACAAGATAACAAAAAATCTCAACTCTTTG
AGTGAGCTGGAAGTAAAGAATCTTCAAAGACTAAGCGGTGCCATGGATGAACTCCACAACGAAATACTA
GAACTAGATGAGAAAGTGGATGATCTCAGAGCTGATACAATAAGCTCACAAATAGAACTCGCAGTCCTG
CTTTCCAATGAAGGAATAATAAACAGTGAAGATGAACATCTCTTGGCGCTTGAAAGAAAGCTGAAGAAA
ATGCTGGGCCCCCTGCTGTAGAGATAGGGAATGGATGCTTTGAAACCAAAACACAAGTGAACCAAGACC
TGTCTCGACAGAATAGCTGCTGGTACCTTTGATGCAGGAGAATTTTCTCTCCCCACCTTTGATTCACTGAA
TATTACTGCTGCATCTTTAAATGACGATGGATTGGATAATCATACTATACTGCTTTACTACTCAACTGCTG
CCTCCAGTTTGGCTGTAACACTGATGATAGCTATCTTTGTTGTTTATATGGTCTCCAGAGACAAATGTTTCT
GCTCCATCTGTCTATAA

Fig.6D

Bảng biểu hiện số 1029 từ vùng khởi đầu 2X35S đến vùng kết thúc NOS.PDISP/HA từ cúm B/ Brisbane/60/2008 được gạch chân. (SEQ ID NO: 29)

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAA
 AGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCAGCTATCT
 GTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAA
 AGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCAACCCACGAGGAGCATCG
 TGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACG
 ACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCA
 ACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATA
 GTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCC
 TCTGCCGACAGTGGTCCCAAAGATGGACCCCAACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCA
 ACCACGTCTTCAAAGCAAGTGGATTGATGTGATACTCCACTGACGTAAGGGATGACGCACAAATCCCCT
 ATCCTTCGCAAGACCCCTTCTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTAATCTTAATAGG
 TTTTGATAAAAGCGAACGTGGGGAACCCGAACCAACCTTCTTCTAACTCTCTCTCATCTCTCTTAAAG
 CAAACTTCTCTCTTGTCTTCTTGGCTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACA
 ACGTTTTCTTCTACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAAGTGGCGGCCATTAAATAACG
 TGTACTTGTCTATTCTGTGCGGTGTGGTCTTGGGAAAAAGAAAGCTTGCTGGAGGCTGCTGTTCAAGCCCA
 TACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTG
 TTGCCTGTACTTCTTCTTCTTCTTCTTCTGCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTTGAAACAGA
 GTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGC
 CCATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCG
CCGATCGAATCTGCACTGGAATAACATCGTCAAACTCACCACATGTCGTCAAACTGCTACTCAAGGGGA
GGTCAATGTGACTGGTGAATACCACTGACAACAACCCACCAAATCTCATTGTCAAATCTCAAAGGA
ACAGAAACCAGGGGGAAACTATGCCCAAAATGCCTCAACTGCACAGATCTGGACGTAGCCTTGGGCAGA
CCAAATGCACGGGGAAATACCCCTCGGCAAGAGTTTCAATACTCCATGAAGTCAGACCTGTTACATCTG
GGTGCTTTCCTATAATGCACGACAGAACAAAAATTAGACAGCTGCCTAACCTTCTCCGAGGATACGAACA
TATCAGGTTATCAACCCATAACGTTATCAATGCAGAAAAATGCACCAGGAGGACCCTACAAAATTGGAAC
CTCAGGGTCTTGCCCTAACATTACCAATGGAAACGGATTTTCGCAACAATGGCTTGGGCCGTCCCAAAA
AACGACAAAAACAAAAACAGCAACAAATCCATTAAACATAGAAAGTACCATAACATTTGTACAGAAGGAGAA
GACCAAAATTACCGTTTGGGGGTTCCACTCTGACAACGAGACCCAAATGGCAAAGCTCTATGGGGACTCAA
AGCCCCAGAAAGTTCACCTCATCTGCCAACGGAGTGACCACACATTACGTTTACAGATTGGTGGCTTCCC
AAATCAAACAGAAGACGGAGGACTACCACAAAGTGGTAGAATTGTTGTTGATTACATGGTGCAAAAATC
TGGGAAAAACAGGAACAATTACCTATCAAAGGGGTATTTTATTGCCTCAAAGGTGTGGTGCAGCAAGTGG
CAGGAGCAAGGTAATAAAAGGATCCTTGCCCTTAATTGGAGAAGCAGATTGCCTCCACGAAAAATACGG
TGGATTAACAAAAAGCAAGCCTTACTACACAGGGGAACATGCAAAGGCCATAGGAAATTGCCCAATATG
GGTGAAAAACACCCTTGAAGCTGGCCAATGGAACCAAATATAGACCTCCTGCAAACTATTAAAGGAAAG
GGGTTTCTTCGGAGCTATTGCTGGTTCTTAGAAGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGA
TACACATCCCATGGGGCACATGGAGTAGCGGTGGCAGCAGACCTTAAGAGCACTCAAGAGGCCATAAAC
AAGATAACAAAAATCTCAACTCTTTGAGTGAGCTGGAAGTAAAGAATCTTCAAAGACTAAGCGGTGCC
ATGGATGAAGTCCACAACGAAATACTAGAACTAGATGAGAAAGTGGATGATCTCAGAGCTGATACAATA
AGCTCACAATAAGAACTCGCAGTCTGCTTTCCAATGAAGGAATAATAACAGTGAAGATGAACATCTCT
TGGCGCTTGAAAGAAAGCTGAAGAAAAATGCTGGGCCCTCTGCTGTAGAGATAGGGAATGGATGCTTTG
AAACCAAAACACAAGTGCAACCAGACCTGTCTCGACAGAATAGCTGCTGGTACCTTTGATGCAGGAGAAT
TTTCTCTCCCCACCTTTGATTCAGTGAATATTACTGCTGCATCTTTAAATGACGATGGATTGGATAATCAT
ACTATACTGCTTTACTACTCAACTGCTGCCTCCAGTTTGGCTGTAACACTGATGATAGCTATCTTTGTTGTT
TATATGGTCTCCAGAGACAATGTTTCTTGCTCCATCTGTCTATAAAGGCCTATTTTCTTTAGTTGAATTA
CTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTATTTTATGTAAT
TTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTTAT
TAAAAAAGAAAAAAGACCGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATT
TGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGCTTTCGATGATTATCATATAATTTCTGTTGAA
TTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAG

TCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAAGTAGGATAAATTATCGCG
CGCGGTGTCATCTATGTTACTAGAT

Fig.6E

Trình tự axit amin của PDISP/HA từ cúm B/Brisbane/60/2008 SEQ ID NO:30

MAKNVAIFGLLFSLLVLVPSQIFADRICTGITSSNSPHVVKTATQGEVNVTVGVPLTTTPTK
SHFANLKGTTETRGKLCPLNCTDLVALGRPKCTGKIPSARVSILHEVRPVTSGCFPIM
HDRTKIRQLPNLLRGYEHIRLSTHNVINAENAPGGPYKIGTSGSCPNTNGNGFFATMAW
AVPKNDKNKTATNPLTIEVPYICTEGEDQITVWGFHSDNETQMAKLYGDSKPQKFTSSA
NGVTTHYVSQIGGFPNQTEDGGLPQSGRIVVDYMVQKSGKTGTITYQRGILLPQKVWCA
SGRSKVIKGSPLIGEADCLHEKYGGLNKSHPYYTGEHAKAIGNCPIWVKTPCLKLANGTK
YRPPAKLLKERGFFGAIAGFLEGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINK
ITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDLADTISSQIELAVLLSNEGIIN
SEDEHLLALERKLLKMLGPSAVEIGNGCFETKHKCNQTCLDRIAAGTFDAGEFSLPTFDS
LNITAASLNDDGLDNHTILLYYSTAASSLAVTLMIAIFVVMVSRDNVSCSICL

G-2X35S/CPMV-HT/PDISP/HA B Brisbane/NOS vào hệ khuếch đại BeYDV+Replicaza (Cấu trúc số 1008)

Fig.6F

Bản đồ của cấu trúc 1194. Các vị trí enzym giới hạn SacII và StuI được sử dụng để kéo dài plasmid được giảm độc lực trên bản đồ này

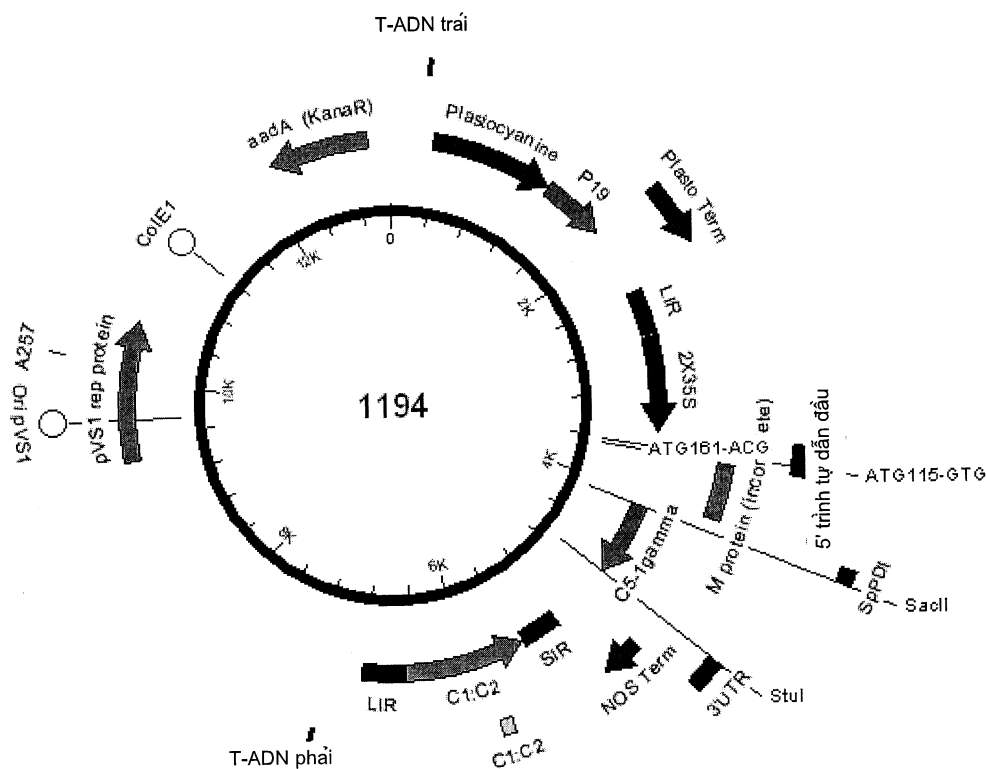


Fig.6G

Cấu trúc 1194 từ các bờ ADN vận chuyển từ trái sang phải (được gạch chân). 2X35S/CPMV-HT/PDISP/NOS vào hệ khuếch đại BeYDV+Replicaza có băng biểu hiện chất ức chế bất hoạt Plastocyanin-P19-Plastocyanin

TGGCAGGATATATTGTGGTGTAACAAATTGACGCTTAGACAACCTTAATAACACATTGCGGACGTTTTTA
 ATGTACTGAATTAACGCCGAATCCCGGGCTGGTATATTTATATGTTGTCAAATAACTCAAAAACCATAAA
 AGTTTAAGTTAGCAAGTGTGTACATTTTACTTGAACAAAAATTCACCTACTACTGTTATAAATCATT
 TTAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTGATATTTTGACAACAATTTTGTG
 CAACATTTGAGAAAATTTTGTGTTCTCTCTTTTCATTGGTCAAAAACAATAGAGAGAGAAAAAGGAAGA
 GGGAGAATAAAAAACATAATGTGAGTATGAGAGAGAAAAGTTGTACAAAAGTTGTACCAAAATAGTTGTAC
 AAATATCATTGAGGAATTTGACAAAAGCTACACAAATAAGGGTTAATTGCTGTAAATAAATAAGGATGA

CGCATTAGAGAGATGTACCATTAGAGAATTTTTGGCAAGTCATTAAAAAGAAAGAATAAATTATTTTTTAA
 AATTAAGTTGAGTCATTTGATTAAACATGTGATTATTAATGAATTGATGAAAGAGTTGGATTAAAGT
 TGTATTAGTAATTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCTATATATTGCCCATAGA
 GTCAGTTAACTCATTTTTATATTTTCATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAA
 AAAAAACGGTATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTAGGAGGATA
 ACATCCAATCCAACCAATCACACAATCCTGATGAGATAACCCACTTTAAGCCACGCATCTGTGGCACA
 TCTACATTATCTAAATCACACATTCTCCACACATCTGAGCCACACAAAAACCAATCCACATCTTTATCAC
 CCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAG
 ACTAATTAATTAATTAATCATCTTGAGAGAAAATGGAACGAGCTATACAAGGAAACGACGCTAGGGAAC
 AAGCTAACAGTGAAACGTTGGGATGGAGGATCAGGAGGTACCATTCTCCCTTCAAACCTTCTGACGAAAG
 TCCGAGTTGGACTGAGTGGCGGTACATAACGATGAGACGAATTCGAATCAAGATAATCCCCTTGGTTTC
 AAGGAAAGCTGGGGTTTCGGGAAAGTTGTATTTAAGAGATATCTCAGATACGACAGGACGGAAGCTTCA
 CTGCACAGAGTCCCTTGGATCTTGGACGGGAGATTCCGTTAACTATGCAGCATCTCGATTTTTCGGTTTCGA
 CCAGATCGGATGTACCTATAGTATTCCGTTTCGAGGAGTTAGTATCACCGTTTCTGGAGGGTCGCGAACT
 CTTACAGCATCTCTGTGAGATGGCAATTCGGTCTAAGCAAGAACTGCTACAGCTTGCCCCAATCGAAGTGG
 AAAGTAATGTATCAAGAGGATGCCCTGAAGGTACTCAAACCTTCGAAAAAGAAAGCGAGTAAGTTAAAA
 TGCTTCTTCGTCTCTTATTTATAATATGGTTTGTATTGTAAATTTTGTCTTGTAGAAGAGCTTAATTAAT
 CGTTGTGTATGAAATACTATTTGTATGAGATGAACTGGTGAATGTAATTCATTACATAAGTGGAGTCT
 AGAATCAGAATGTTTCTCCATACTAACTAGACATGAAGACCTGCCGCTACAATTGTCTTATATTTGA
 ACACTAAAATGAACATCTTTTCCCACTTTATAAGTGGTTAATATAGCTCAAATATATGGTCAAGTT
 CAATAGATTAATAATGGAATATCAGTTATCGAAATTCATTAACAATCACTTAACGTTATTAACACTACTA
 ATTTTATATCATCCCCTTTGATAAATGATAGTACACCAATTAGGAAGGAGCATGCTCGCCTAGGAGATTG
 TCGTTTCCCGCCTTCAGTTTGAAGCTGCTCTAGCCGTGTAGCCAATACGCAAACCGCCTCTCCCCGCGCG
 TTGGGAATTACTAGCGCGTGTGACACGCGTGGCGCGCCCTAGCAGAAGGCATGTTGTTGTGACTCCGAG
 GGGTTGCCTCAAACCTCTATCTTATAACCGCGGTGGAGGCATGGAGGCAAGGGCATTTTGGTAATTTAAGT
 AGTTAGTGGAAAATGACGTCATTTACTTAAAGACGAAGTCTTGCGACAAGGGGGGCCACGCCGAATTT
 AATATTACCGCGTGGCCCCACCTTATCGCGAGTGCTTTAGCACGAGCGGTCCAGATTTAAAGTAGAAAA
 GTTCCCGCCCACTAGGGTTAAAGGTGTTCACTATAAAAGCATATACGATGTGATGTTTATGATAAAG
 CGTATATTGTATCAGGTATTTCCGTCCGATACGAATTATTCGTACAAGCTTCTTAAGCCGGTCAACATGGT
 GGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGA
 GACTTTTCAACAAAGGGTAATATCCGGAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTG
 TGAAGATAGTGGAAAAGGAAGGTGGCTCTACAAATGCCATCATTGCGATAAAGGAAAAGGCCATCGTTG
 AAGATGCCTCTGCCGACAGTGGTCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAGAAAGAA
 ACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACGACACACTTGTCTA
 CTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAAT
 ATCCGGAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAA
 GGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTG
 GTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAGAAAGAACGTTCCAACCACGTCTTCAA
 AGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGA
 CCCTTCCTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTAATCTTAATAGGTTTGGATAAAAGC
 GAACGTGGGGAACCCGAACCAACCTTCTTAAACTCTCTCATCTCTCTTAAAGCAAACTTCTCTCT
 TGTCTTTCTTGGTGTGAGCATCTTCAACGTTGTGAGATCGTGCTTCGGCACCAAGTACAACGTTTCTTTCA
 CTGAAGCGAAATCAAAGATCTCTTGTGGACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCTTA
 TTCTTGTGCGTGTGGTCTTGGGAAAAGAAAGCTTGTGGAGGCTGCTGTTTCAGCCCCATACATTACTTGT
 ACGATTCTGCTGACTTTCGGCGGGTGAATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTC
 TTTCTTCTTCTTCTTGTGATTGGTCTATAAGAAATCTAGTATTTTCTTGAACAGAGTTTCCCGTGGT
 TTTGCAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTGCGGGCCATGGCGAAAA
 ACGTTGCGATTTTCGGCTTATTGTTTCTTCTTGTGTTGGTTCCCTTCTCAGATCTTCGCGCGGGCTCCTC
 AGCCAAAACGACACCCCATCTGTCTATCCACTGGCCCTGGATCTGCTGCCCAAACCTAATCCATGGTG
 ACCCTGGGATGCCTGGTCAAGGGCTATTTCCCTGAGCCAGTGACAGTGACCTGGAACCTCTGGATCCCTGT
 CCAGCGGTGTGCACACCTTCCAGCTGTCTGACGTCTGACCTCTACACTCTGAGCAGCTCAGTGACTGTG
 CCTCCAGCACCTGGCCAGCGAGACCGTCACTGCAACGTTGCCACCCGCGCAGCAGCACCAAGGTG
 GACAAGAAAATTGTGCCAGGGATTGTGGTTGTAAGCCTTGCAATGTACAGTCCCAAGATATCATCTG
 TCTTCATCTTCCCCCAAAGCCCAAGGATGTGCTCACCATTACTCTGACTCCTAAGGTCACGTGTGTTGTG

GTAGACATCAGCAAGGATGATCCCGAGGTCCAGTTCAGCTGGTTTGTAGATGATGTGGAGGTGCACACA
 GCTCAGACGCAACCCCGGGAGGAGCAGTTC AACAGCACCTTCCGCTCAGTCAGTGAACCTCCCATCATGC
 ACCAGGACTGGCTCAATGGCAAGGAAGGCCTATTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATT
 TCTATGTTTGGTGAGCGGTTTCTGTGCTCAGAGTGTGTTTATTTATGTAATTTAATTTCTTTGTGAGCTC
 CTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAAATTTATTA AAAAAAAAAAAAAAAAAA
 AGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAA
 GATTGAATCCTGTTGCCGGTCTTGCATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATA
 ATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTA
 ATACGCGATAGAAAAACAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCTATGTT
 ACTAGATCTCTAGAGTCTCAAGCTTGGCGCGGGGTACCGAGCTCGAATTCAGTGTACTTCAAGTCAGT
 TGGAAATCAATAAAATGATTATTTTATGAATATATTTTATTGTGCAAGTAGATAGAAATTACATATGTTAC
 ATAACACACGAAATAAAACAAAAACACAATCCAAACAAACACCCCAAAACAAATAACACTATATAT
 ATCCTCGTATGAGGAGAGGCACGTTTCAAGTACTCGACGATTCCCGAGCAAAAAAGTCTCCCGTCACAC
 ATATAGTGGGTGACGCAATTATCTTCAAAGTAATCCTTCTGTTGACTTGTGATTGATAACATCCAGTCTTC
 GTCAGGATTGCAAGAATTATAGAAGGGATCCCACTTTTATTTTCTTTTCCATATTTAGGGTTGA
 CAGTGAAATCAGACTGGCAACCTATTAATTGCTTCCACAATGGGACGAACCTGAAGGGGATGTCGTCGAT
 GATATTATAGGTGGCGTGTTCATCGTAGTTGGTGAAGTCGATGGTCCCGTTCAGTAGTTGTGTGCGCCGA
 GACTTCTAGCCAGGTGGTCTTTCCGGTACGAGTTGGTCCGCAGATGTAGAGGCTGGGGTGTCTGACCCC
 AGTCCTTCCCTCATCCTGGTTAGATCGGCCATCCACTCAAGGTCAGATTGTGCTTGATCGTAGGAGACAG
 GATGTATGAAAGTGTAGGCATCGATGCTTACATGATATAGGTGCGTCTCTCTCCAGTTGTGCAGATCTTCG
 TGGCAGCGGAGATCTGATTCTGTGAAGGGCGACACGTACTGCTCAGGTTGTGGAGGAAATAATTTGTTGG
 CTGAATATTCAGCCATTGAAGCTTTGTGCCCCATTATGAGGGAATCTTCTTTGATCATGTCAAGATAC
 TCCTCCTTAGACGTTGCAGTCTGGATAATAGTTTCGCCATCGTGCCTCAGATTTGCGAGGAGAGACCTTAT
 GATCTCGGAAATCTCCTCTGGTTTAAATATCTCCGTCCTTTGATATGTAATCAAGGACTTGTTTAGAGTTTC
 TAGCTGGCTGGATAATTAGGGTGATTTCCCTTCAAAATCGAAAAAGAAGGATCCCTAATAACAAGGTTTTT
 ATCAAGCTGGATAAGAGCATGATAGTGGGTAGTGCCATCTTGATGAAGCTCAGAAGCAACACCAAGGAA
 GAAAAATAAGAAAAGGTGTGAGTTTCTCCAGAGAAACTGGAATAAATCATCTCTTTGAGATGAGCACTTG
 GGGTAGGTAAGGAAAAACATATTTAGATTGGAGTCTGAAGTTCTTGCTAGCAGAAGGCATGTTGTTGTGAC
 TCCGAGGGGTGCCTCAAACCTATCTTATAACCGGCGTGAGGCATGGAGGCAAGGGCATTTTGGTAAT
 TTAAGTAGTTAGTGGAATAAGACGTCATTTACTTAAAGACGAAGTCTTGCGACAAGGGGGGCCACGCC
 GAATTTTAATATTACCGGCGTGCCCCACCTATCGCGAGTGCTTTAGCACGAGCGGTCCAGATTTAAAG
 TAGAAAAGTTCCCGCCCACTAGGGTTAAAGGTGTTACACTATAAAAGCATATACGATGTGATGGTATTT
 GATGGAGCGTATATTGTATCAGGTATTTCCGTCGGATACGAATTATTCGTACGCGCGGCCACTAGTGGCA
 CTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGCGGTTACCCAACCTAATCGCCTTGCAGCAC
 ATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAG
 CCTGAATGGCGAATGCTAGAGCAGCTTGAGCTTGGATCAGATTGTCGTTTCCCGCCTTCAGTTTAAACTAT
 CAGTGTTTGACAGGATATATTGGCGGTTAAACCTAAGAGAAAAGAGCGTTTA

Fig.6H

Bảng biểu hiện số 1008 từ LIR bên trái BeYDV sang LIR bên phải BeYDV. PDISP/HA từ cúm B/
 Brisbane/60/2008 được gạch chân (SEQ ID NO:32)

CTAGCAGAAGGCATGTTGTTGTGACTCCGAGGGGTTGCCTCAAACCTCTATCTTATAACCGGCGTGAGGC
 ATGGAGGCAAGGGCATTTTGGTAATTTAAGTAGTTAGTGGAATAAGACGTCATTTACTTAAAGACGAAGT
 CTTGCGACAAGGGGGGCCACGCCGAATTTAATATTACCGGCGTGCCCCACCTTATCGCGAGTGCTTT
 AGCACGAGCGGTCCAGATTTAAAGTAGAAAAGTCCCGCCCACTAGGGTTAAAGGTGTTACACTATAA
 AAGCATATACGATGTGATGGTATTTGATAAAGCGTATATTGTATCAGGTATTTCCGTCGGATACGAATTA
 TTCGTACAAGCTTCTTAAGCCGGTCAACATGGTGGAGCACGACACACTTGTCTACTCAAAAAATATCAAA
 GATACAGTCTCAGAAGACCAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTC
 GGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAGAAGGTGGCTCCTACAAAT
 GCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAGATGGAC

CCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGAT
 GTGATAACATGGTGGAGCAGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACC
 AAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCAGCTAT
 CTGTCACTTTATTGTGAAGATAGTGGAAAAAGGAAGTGGCTCCTACAAATGCCATCATTGCGATAAAGGA
 AAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATC
 GTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAA
 GGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCTATATAAGGAAGTTCATTTTCATTGGAG
 AGGTATTAAAACTTAATAGGTTTTGATAAAAGCGAACGTGGGGAAACCCGAACCAACCTTCTTCTAAA
 CTCTCTCTCATCTCTCTTAAAGCAAACCTTCTCTTGTCTTCTTGCCTGAGCGATCTTCAACGTTGTCAGA
 TCGTGCTTCGGCACCAGTACAACGTTTCTTTCCTGAAAGCGAAATCAAAGATCTCTTGTGGACACGTAG
 TGGCGGCCATTAAATAACGTGTACTTGTCTTCTTGTCTGCTGACTTTCGGCGGGTGCAATATCTCTAC
 TCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTCTTCTTCTTGTGCTGATTGGTCTATAAGAAATCT
 AGTATTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTGCAACTTGGAGAAAGATTGTTAAGCTTCTGTATAT
 TCTGCCCAAATTTGTGCGGCCCATGGCGAAAAACGTTGCGATTTCGGCTTATTGTTTCTTCTTGTGTT
GGTTCCTTCTCAGATCTTCGCCGATCGAATCTGCACTGGAATAACATCGTCAAACTCACCACATGTCGTCA
AAATGCTACTCAAGGGGAGGTCAATGTGACTGGTGTGAATACCACTGACAAACAACCCACCAAAATCTC
ATTTTGCAAAATCTCAAAGGAACAGAAACCAGGGGAAACTATGCCCAAAATGCCTCACTGCACAGATC
TGGACGTAGCCTTGGGCAGACCAAAATGCACGGGAAAAATACCCTCGGCAAGAGTTCAATACTCCATG
AAGTCAGACCTGTTACATCTGGGTGCTTTCCTATAATGCACGACAGAACAAAAATTAGACAGTGCCTAA
CCTTCTCCGAGGATACGAACATATCAGGTTATCAACCCATAACGTTATCAATGCAGAAAAATGCACCAGGA
GGACCCTACAAAATTGGAACCTCAGGGTCTTGCCCTAACATTACCAATGGAAACGGATTTTCGCAACAA
TGGCTTGGGCGCTCCAAAAAACGACAAAAACAAAAACAGCAACAAATCCATTAACAATAGAAGTACCAT
ACATTTGTACAGAAGGAGAAGACCAAAATACCGTTTGGGGGTTCCACTCTGACAACGAGACCCAAATGG
CAAAGCTCTATGGGGACTCAAAGCCCCAGAAGTTCACCTCATCTGCCAACGGAGTGACCACACATTACGT
TTCACAGATTGGTGGCTTCCCAAATCAAACAGAAGACGGAGGACTACCACAAAGTGGTAGAATTGTTGTT
GATTACATGGTGCAAAAATCTGGGAAAAACAGGAACAATTACCTATCAAAGGGGTATTTATTGCCTCAAA
AGGTGTGGTGCAGCAAGTGGCAGGAGCAAGGTAATAAAAGGATCCTTGCCTTAATTGGAGAAGCAGATT
GCCTCCACGAAAAATACGGTGGATTAACAAAAAGCAAGCCTTACTACACAGGGGAACATGCAAAAGGCCA
TAGAAATTTGCCAATATGGGTGAAAAACACCCTTGAAGCTGGCCAATGGAACCAATATAGACCTCCTG
CAAAACTATTAAAGGAAAGGGGTTTCTTCGGAGCTATTGCTGGTTTCTTAGAAGGAGGATGGGAAGGAA
TGATTGCAGGTTGGCACGGATACACATCCCATGGGGCACATGGAGTAGCGGTGGCAGCAGACCTTAAGA
GCACTCAAGAGGCCATAAACAAGATAACAAAAATCTCAACTCTTTGAGTGAGCTGGAAGTAAAGAATC
TTCAAAGACTAAGCGGTGCCATGGATGAACTCCACAACGAAATACTAGAACTAGATGAGAAAAGTGGATG
ATCTCAGAGCTGATACAATAAGCTCACAAATAGAACTCGCAGTCTGCTTCCAATGAAGGAATAATAAA
CAGTGAAGATGAACATCTCTTGGCGCTTGAAGAAAGCTGAAGAAAATGCTGGGCCCCCTGCTGTAGA
GATAGGGAATGGATGCTTTGAAACCAACACAAGTGCAACCAGACCTGTCTCGACAGAATAGCTGCTGG
TACCTTTGATGCAGGAGAATTTCTCTCCCCACCTTTGATTCACTGAATATTACTGCTGCATCTTTAAATG
ACGATGGATTGGATAATCATACTATACTGCTTTACTACTCAACTGCTGCCTCCAGTTGGCTGTAACACTG
ATGATAGCTATCTTTGTTGTTTATATGGTCTCCAGAGACAATGTTTCTTGTCTCCATCTGTCTATAAAGGCCT
 ATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAG
 AGTGTGTTTATTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCTTCAGCAAGGACA
 CAAAAAGATTTTAAATTTATTAATAAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGAC
 CTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCATGATT
 ATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAG
 ATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCA
 AACTAGGATAAAATTATCGCGCGCGGTGCATCTATGTTACTAGATCTCTAGAGTCTCAAGCTTGGCGCGG
 GGTACCGAGCTCGAATTCAGTGTACTTCAAGTCAGTTGGAATCAATAAAATGATTATTTTATGAATA
 TATTTTCATTGTGCAAGTAGATAGAAATTACATATGTTACATAACACACGAAATAAACAACAAAAACACAAT
 CCAAAACAAACACCCCAAAACAAAATAAACAATATATCTCTCGTATGAGGAGAGGCACGTTCAAGTAC
 TCGACGATTCCCAGCAAAAAAAGTCTCCCCGTACACATATAGTGGGTGACGCAATTATCTTCAAAGTA
 ATCCTTCTGTTGACTTGTCAATTGATAACATCCAGTCTTCGTCAGGATTGCAAGAATTATAGAAGGGATCC
 CACCTTTTATTTTCTTCTTTTCCATATTTAGGGTGTGACAGTGAAATCAGACTGGCAACCTATTAATTGCT
 TCCACAATGGGACGAACCTGAAGGGGATGTCGTCGATGATATTATAGGTGGCGTGTTCATCGTAGTTGGT

GAAGTCGATGGTCCCGTTCCAGTAGTTGTGTGCGCCCGAGACTTCTAGCCCAGGTGGTCTTTCCGGTACGA
GTTGGTCCGCAGATGTAGAGGCTGGGGTGTCTGACCCAGTCCTTCCCTCATCCTGGTTAGATCGGCCATC
CACTCAAGGTCAGATTGTGCTTGATCGTAGGAGACAGGATGTATGAAAGTGTAGGCATCGATGCTTACAT
GATATAGGTGCGTCTCTCTCCAGTTGTGCAGATCTTCGTGGCAGCGGAGATCTGATTCTGTGAAGGGCGA
CACGTA CTGCTCAGGTTGTGGAGGAAATAATTTGTTGGCTGAATATTCCAGCCATTGAAGCTTTGTTGCCC
ATTCATGAGGGAATTCTTCTTTGATCATGTCAAGATACTCCTCCTTAGACGTTGCAGTCTGGATAATAGTT
CGCCATCGTGCGTCAGATTTGCGAGGAGAGACCTTATGATCTCGGAAATCTCCTCTGGTTTAAATATCTCC
GTCCTTTGATATGTAATCAAGGACTTGTTTTAGAGTTTCTAGCTGGCTGGATATTAGGGTGATTTCTTCAA
AATCGAAAAAAGAAGGATCCCTAATACAAGGTTTTTTATCAAGCTGGATAAGAGCATGATAGTGGGTAG
TGCCATCTTGATGAAGCTCAGAAGCAACACCAAGGAAGAAAAATAAGAAAAGGTGTGAGTTTCTCCCAGA
GAAACTGGAATAAATCATCTCTTTGAGATGAGCACTTGGGGTAGGTAAGGAAAACATATTTAGATTGGA
GTCTGAAGTTCTTGCTAGCAGAAGGCATGTTGTTGTGACTCCGAGGGGTTGCCTCAAACCTCTATCTTATAA
CCGGCGTGGAGGCATGGAGGCAAGGGCATTGTTGGTAATTTAAGTAGTTAGTGGAATGACGTCATTTAC
TTAAAGACGAAGTCTTGCGACAAGGGGGGGCCACGCCGAATTTAATATTACCGGCGTGGCCCCACCTTA
TCGCGAGTGCTTTAGCACGAGCGGTCCAGATTTAAAGTAGAAAAAGTTCCCGCCCACTAGGGTTAAAGGTG
TTCACACTATAAAAGCATATACGATGTGATGGTATTTGATGGAGCGTATATTGTATCAGGTATTTCCGTCG
GATACGAATTATTCGTAC

Miền chuyển mào H-2X35S/CPMV-HT/PDISP/HA B Brisbane/H5 Indonesia và đoạn tế bào chất (H5Indo TMCT)/NOS vào hệ khuếch đại BeYDV+Replicaza (Cấu trúc số 1009)

Fig.7A

dTmH5I-B Bris.r (SEQ ID NO: 33)

TTGACAGTATTTGGTAATTATCCAATCCATCGTCATTTAAAGATGCAGCA

Fig.7B

B Bris-dTmH5I.c (SEQ ID NO: 34)

CATCTTTAAATGACGATGGATTGGATAATTACCAAATACTGTCAATTTAT

Fig.7C

IF-S1aS4-dTmH5I.r (SEQ ID NO: 35)

ACTAAAGAAAATAGGCCTTTAAATGCAAATTCTGCATTGTAACGATCCAT

Fig.7D

Bảng biểu hiện số 1009 từ LIR bên trái BeYDV sang LIR bên phải BeYDV. PDISP/HA B Brisbane/H5Indo TMCT được gạch chân. SEQ ID NO: 36

CTAGCAGAAGGCATGTTGTTGTGACTCCGAGGGGTTGCCTCAAACCTCTATCTTATAACCGGCGTGGAGGC
ATGGAGGCAAGGGCATTGTTGGTAATTTAAGTAGTTAGTGGAATAAGACGTCATTTACTTAAAGACGAAGT
CTTGCGACAAGGGGGGCCACGCCGAATTTAATATTACCGGCGTGGCCCCACCTTATCGCGAGTGCTT
AGCACGAGCGGTCCAGATTTAAAGTAGAAAAGTCCCGCCCACTAGGGTTAAAGGTGTCACACTATAA
AAGCATATACGATGTGATGGTATTTGATAAAGCGTATATTGTATCAGGTATTTCCGTCGGATACGAATTA
TTCGTACAAGCTTCTTAAGCCGGTCAACATGGTGGAGCAGCACACACTTGTCTACTCCAAAAATATCAAA
GATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGTAATATCCGGAAACCTCCTC
GGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAGGAAGGTGGCTCCTACAAAT
GCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAGATGGAC
CCCCACCCACGAGGAGCATCGTGGAAGAAAGACGTTCCAACACGTCCTCAAAGCAAGTGGATTGAT
GTGATAACATGGTGGAGCAGCACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACC
AAAGGGCAATTGAGACTTTTCAACAAAGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCAGCTAT
CTGTCACTTTATTGTGAAGATAGTGGAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGA
AAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAGATGGACCCCCACCCACGAGGAGCATC
GTGGAAGAAAGAACGTTCCAACACGTCCTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAA
GGGATGACGCACAATCCCACTATCCTTCGCAAGACCTTCCTCTATATAAGGAAGTTCATTTATTGGAG
AGGTATTAATCTTAATAGGTTTTGATAAAGCGAACGTGGGGAAACCCGAACCAACCTTCTCTAA

CTCTCTCATCTCTCTTAAAGCAAACCTCTCTCTTGTCTTTCTTGGCTGAGCGATCTTCAACGTTGTCAGA
 TCGTGCTTCGGCACCAGTACAACGTTTTCTTTCTACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAG
 TGCGGCGCCATTAAATAACGTGTACTTGTCTATTCTTGTCTGGTGTGGTCTTGGGAAAAAGAAAGCTTGCTG
 GAGGCTGCTGTTTACGCCCCATACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTAC
 TTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTTCTGCTGATTGGTTCTATAAGAAATCT
 AGTATTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATAT
 TCTGCCCAAATTTGTCTGGGCCCCATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTT
GGTTCCTTCTCAGATCTTCGCCGATCGAATCTGCACTGGAATAACATCGTCAAACCTACCCACATGTCGTC
AAACTGCTACTCAAGGGGAGGTCAATGTGACTGGTGAATACCACTGACAACAACACCCACCAAAATCTC
ATTTTGCAAATCTCAAAGGAACAGAAACCAGGGGAAACTATGCCCAAAATGCCTCAACTGCACAGATC
TGGACGTAGCCTTGGGCAGACCAAAATGCACGGGAAAAATACCCCTCGGCAAGAGTTTCAATACTCCATG
AAGTCAGACCTGTTACATCTGGGTGCTTTCTTATAATGCACGACAGAAACAAAAATTAGACAGCTGCCTAA
CCTTCTCCGAGGATACGAACATATCAGGTTATCAACCCATAACGTTATCAATGCAGAAAAATGCACAGGA
GGACCCTACAAAATTGGAACCTCAGGGTCTTGCCCTAACATTACCAATGGAAACGGATTTTTCGCAACAA
TGGCTTGGGCGCTCCCAAAAAACGACAAAAACAGCAACAAATCCATTAAACAATAGAAGTACCAT
ACATTTGTACAGAAGGAGAAGACCAAAATTACCGTTTGGGGGTTCCTACTCTGACAACGAGACCCAAATGG
CAAAGTCTATGGGGACTCAAAGCCCCAGAAAGTTCACCTCATCTGCCAACGGAGTGACCACACATTACGT
TTACACAGATTGGTGGCTTCCCAAAATCAAACAGAAAGACGGAGGACTACCACAAAGTGGTAGAATTGTTGTT
GATTACATGGTGCAAAAAATCTGGGAAAAACAGGAACAATTACCTATCAAAGGGGTATTTTATTGCCTCAAA
AGGTGTGGTGCAGCAAGTGGCAGGAGCAAGGTAATAAAAGGATCCTTGCCTTAATTGGAGAAGCAGATT
GCCTCCACGAAAAATACGGTGGATTAACAAAAAGCAAGCCTTACTACACAGGGGAACATGCAAAAGGCCA
TAGGAAATTGCCCAATATGGGTGAAAAACACCCTTGAAGCTGGCCAATGGAAACCAAAATATAGACCTCCTG
CAAAACTATTAAAGGAAAGGGGTCTTCTCGGAGCTATTGCTGGTTTCTTAGAAGGAGGATGGGAAGGAA
TGATTGCAGTTTGGCACGGATACACATCCCATGGGGCACATGGAGTAGCGGTGGCAGCAGACCTTAAGA
GCACTCAAGAGGCCATAAAACAAGATAACAAAAAATCTCAACTCTTGAGTGAGCTGGAAGTAAAGAATC
TTCAAAGACTAAGCGGTGCCATGGATGAACTCCACAACGAAATACTAGAACTAGATGAGAAAGTGGATG
ATCTCAGAGCTGATACAATAAGCTCACAAATAGAACTCGCAGTCCTGCTTTCCAATGAAGGAATAATAAA
CAGTGAAGATGAACATCTCTTGGCGCTTGAAACCAACACAAGTGCAACCAGACCTGTCTCGACAGAATAGCTGCTGG
GATAGGGAATGGATGCTTTGAAACCAACACAAGTGCAACCAGACCTGTCTCGACAGAATAGCTGCTGG
TACCTTTGATGCAGGAGAATTTTCTCTCCCACTTTGATTCACTGAATATTACTGCTGCATCTTTAAATG
ACGATGGATTGGATAATTACCAAAATACTGTCAATTTATTCAACAGTGGCGAGTTCCTAGCACTGGCAAT
CATGATGGCTGGTCTATCTTTATGGATGTGCTCCAATGGATCGTTACAATGCAGAATTTGCATTTAAAGG
 CTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTCTGTGCTC
 AGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTGCTCCCTTCAGCAAGG
 ACACAAAAAGATTTTAAATTTATTAATAAAAAAAAAAAAAAAAAAGACCGGAATTCGATATCAAGCTTATC
 GACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATG
 ATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTAT
 GAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAAACAAAATATAGCGC
 GCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCTCTAGAGTCTCAAGCTTGGCG
 CGGGGTACCGAGCTCGAATTCAGTGTACTTCAAGTCAGTTGGAATCAATAAAATGATTATTTATGA
 ATATATTTTATTGTGCAAGTAGATAGAAATTACATATGTTACATAACACACGAAATAAACAAAAAACAC
 AATCCAAACAAACACCCCAACAAAATAACACTATATATCCTCGTATGAGGAGAGGCACGTTCACT
 GACTCGACGATTCCCGAGCAAAAAAAGTCTCCCGTCACACATATAGTGGGTGACGCAATTATCTTCAAA
 GTAATCCTTCTGTTGACTTGTCTTGTGATAACATCCAGTCTTCGTCAGGATTGCAAAGAATTATAGAAGGG
 ATCCCACTTTTATTTTCTTTTCTTTTCCATATTTAGGGTGTGACAGTGAAATCAGACTGGCAACCTATTAAT
 TGCTTCCACAATGGGACGAACTTGAAGGGGATGTCGTCGATGATATTATAGGTGGCGTGTTCATCGTAGT
 TGGTGAAGTCGATGGTCCCGTCCAGTAGTTGTGTCGCGGAGACTTCTAGCCCAGGTGGTCTTTCCGGTA
 CGAGTTGGTCCGAGATGTAGAGGCTGGGGTGTCTGACCCAGTCCTTCCCTCATCCTGGTTAGATCGGC
 CATCCACTCAAGGTCAGATTGTGCTTATCGTAGGAGACAGGATGTATGAAAGTGTAGGCATCGATGCTT
 ACATGATATAGGTGCGTCTCTCTCCAGTTGTGTCAGATCTTCGTGGCAGCGGAGATCTGATTCTGTGAAGG
 GCGACACGTAAGTCTCAGGTTGTGGAGGAAATAATTTGTTGGCTGAATATCCAGCCATTGAAGCTTTGT
 TGCCCATTCATGAGGGAATCTTCTTTGATCATGTCAAGATACTCCTCCTTAGACGTTGCAGTCTGGATAA
 TAGTTCCGCATCGTGCCTCAGATTGCGAGGAGAGACCTTATGATCTCGGAAATCTCCTCTGGTTTTAATA
 TCTCCGTCCTTGATATGTAATCAAGGACTTGTTTAGAGTTTCTAGCTGGCTGGATATTAGGGTGATTTC

TTCAAAATCGAAAAAAGAAGGATCCCTAATACAAGGTTTTTTATCAAGCTGGATAAGAGCATGATAGTG
 GGTAGTGCCATCTTGATGAAGCTCAGAAGCAACACCAAGGAAGAAAATAAGAAAAGGTGTGAGTTTCTC
 CCAGAGAAACTGGAATAAATCATCTCTTTGAGATGAGCACTTGGGGTAGGTAAGGAAAACATATTTAGA
 TTGGAGTCTGAAGTTCTTGCTAGCAGAAGGCATGTTGTTGTGACTCCGAGGGGTTGCCTCAAACCTCTATCT
 TATAACCGGCGTGGAGGCATGGAGGCAAGGGCATTGTTGTAATTTAAGTAGTTAGTGAAAAATGACGTC
 ATTTACTTAAAGACGAAGTCTTGCGACAAGGGGGGCCCCACGCCGAATTTTAATATTACCGGCGTGGCCCC
 ACCTTATCGCGAGTGCTTTAGCACGAGCGGTCCAGATTTAAAGTAGAAAAGTTCCCGCCCACTAGGGTTA
 AAGGTGTTACACTATAAAAGCATATACGATGTGATGGTATTTGATGGAGCGTATATTGTATCAGGTATT
 TCCGTCGGATACGAATTATTCGTAC

Fig.7E

Trình tự axit amin của PDISP/HA B Brisbane/H5Indo TMCT (SEQ ID NO: 37)

MAKNV AIFGLLFSLLVLVPSQIFADRICTGITSSNSPHVVK TATQGEVNVTVGVIPLTTTPTK
 SHFANLKGTTETRGKLC PKCLNCTDL DVALGRPKCTGKIPSARVSILHEVRPVTSGCFPIM
 HDRTKIRQLPNLLRGYEHIRLSTHNVINAENAPGGPYKIGTSGSCP NITNGNGFFATMAW
 AVPKNDKNKTATNPLTIEVPYICTEGEDQITVWGFHSDNETQMAKLYGDSKPQKFTSSA
 NGVTTHYVSQIGGFPNQTEDGGLPQSGRIVVDY MVQKSGKTGTITYQRGILLPQK VWCA
 SGRSKVIKGSPLIGEADCLHEKYGGLNKS KPYTGEHAKAIGNCPIWVK TPLKLANGTK
 YRPPAKLLKERGFFGAIAGFLEGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQEANK
 ITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDL RADTISSQIELAVLLSNEG IIN
 SEDEHLLALERKLK KMLGPSAVEIGNGCFETKHKCNQTCLDRIAAGTFDAGEFSLPTFDS
 LNITAASLND DGLDNYQILSIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

I-2X35S/CPMV-HT/PDISP/ HA B Brisbane có vòng thủy phân protein bị mất/NOS vào hệ
khuếch đại BeYDV+Replicaza (Cấu trúc số 1059)

Fig.8A

1039+1059.r (SEQ ID NO: 38)

CTTCCCATCCTCCACCAGGAGGTCTATATTTGGTTCCATTGGCCAGCTTCAA

Fig.8B

1039+1059.c (SEQ ID NO: 39)

CAAATATAGACCTCCTGGTGGAGGATGGGAAGGAATGATTGCAGGTTGGCAC

Fig.8C

Bảng biểu hiện số 1059 từ LIR bên trái BeYDV sang LIR bên phải BeYDV. PDISP/HA từ cúm
B/Brisbane/60/2008 có vòng thủy phân protein bị mất được gạch chân. (SEQ ID NO: 40)

CTAGCAGAAGGCATGTTGTTGTGACTCCGAGGGGTTGCCTCAAACCTCTATCTTATAACCGGCGTGAGGC
ATGGAGGCAAGGGCATTGTTGTAATTTAAGTAGTTAGTGAAAAATGACGTCATTTACTTAAAGACGAAGT
CTTGCGACAAGGGGGGGCCACGCCGAATTTTAATATTACCGGCGTGCGCCACCTTATCGCGAGTGCTTT
AGCACGAGCGGTCCAGATTTAAAGTAGAAAAAGTCCCGCCCACTAGGGTTAAAGGTGTTCACTATAA
AAGCATATACGATGTGATGGTATTTGATAAAGCGTATATTGTATCAGGTATTTCCGTCGGATACGAATTA
TTCGTACAAGCTTCTTAAGCCGGTCAACATGGTGGAGCAGCAGACACTTGTCTACTCCAAAAATATCAAA
GATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTC
GGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGAAAAAGGAAGGTGGCTCCTACAAAT
GCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGAC
CCCCACCCACGAGGAGCATCGTGAAAAAAGAAGACGTCCAACACAGTCTTCAAAGCAAGTGATTGAT
GTGATAACATGGTGGAGCAGCAGACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACC
AAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCAGCTAT
CTGTCACTTTATTGTGAAGATAGTGAAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGA
AAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATC
GTGAAAAAAGAAGACGTTCCAACACAGTCTTCAAAGCAAGTGATTGATGTGATATCTCCACTGACGTAA
GGGATGACGCACAATCCCACTATCCTTCGCAAGACCTTCCTCTATATAAGGAAGTTCAATTCATTGGAG
AGGTATTAATCTTAATAGGTTTGTATAAAGCGAACGTGGGGAAACCCGAACCAACCTTCTCTAAA
CTCTCTCATCTCTCTTAAAGCAAACCTCTCTCTTGTCTTTCTTGCGTGAGCGATCTTCAACGTTGTCAGA
TCGTGCTTCGGCACCAAGTACAACGTTTTCTTCACTGAAGCGAAATCAAAGATCTCTTGTGGACAGTAG
TGCGGCGCCATTAAATAACGTGTAATTGTCTATTCTTGTGCGGTGGTCTTGGGAAAAGAAAGCTTGCTG
GAGGCTGCTGTTAGCCCCATACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTAC
TTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTCTCTTCTTCTGCTGATTGGTTCTATAAGAAATCT
AGTATTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAACCTGGAGAAAGATTGTTAAGCTTCTGTATAT
TCTGCCCAATTTGTGGGGCCATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTGTGTT
GGTTCCTTCTCAGATCTTCGCCGATCGAATCTGCACTGGAATAACATCGTCAAACCTACCCACATGTCTCA
AAACTGCTACTCAAGGGGAGGTCAATGTGACTGGTGTAAATACCACTGACAACAACCCACCAATCTC
ATTTTGCAATCTCAAAGGAACAGAAACCAGGGGAAACTATGCCCAAAATGCCTCAACTGCACAGATC

TGGACGTAGCCTTGGGCAGACCAAAATGCACGGGGAAAAATACCCCTCGGCAAGAGTTTCAATACTCCATG
 AAGTCAGACCTGTTACATCTGGGTGCTTTCTATAATGCACGACAGAAACAAAAATTAGACAGCTGCCTAA
 CCTTCTCCGAGGATACGAACATATCAGGTTATCAACCCATAACGTTATCAATGCAGAAAATGCACCAGGA
 GGACCCCTACAAAATTGGAACCTCAGGGTCTTGCCCTAACATTACCAATGGAAACGGATTTTTTCGCAACAA
 TGGCTTGGGCCGTCCCAAAAAACGACAAAAACAAAACAGCAACAAATCCATTAACAATAGAAGTACCAT
 ACATTGTACAGAAGGAGAAGACCAAAATTACCGTTTGGGGGTTCCACTCTGACAACGAGACCCAAATGG
 CAAAGCTCTATGGGGACTCAAAGCCCCAGAAGTTCACCTCATCTGCCAACGGAGTGACCACACATTACGT
 TTCACAGATTGGTGGCTTCCCAATCAAACAGAAGACGGAGGACTACCACAAAGTGGTAGAATTGTTGTT
 GATTACATGGTGCAAAAATCTGGGAAAACAGGAACAATTACCTATCAAAGGGGTATTTTATTGCCTCAAA
 AGGTGTGGTGCGCAAGTGGCAGGAGCAAGGTAATAAAAGGATCCTTGCTTTAATTGGGAGAAGCAGATT
 GCCTCCACGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGGGAACATGCAAAAGGCCA
 TAGGAAATTGCCCAATATGGGTGAAAACACCCCTGAAGCTGGCCAATGGAACCAATATAGACCTCTTG
 GTGGAGGATGGGAAGGAATGATTGCAGGTTGGCAGGATACACATCCCATGGGGCAGCATAGGAGTAGCGG
 TGGCAGCAGACCTTAAGAGCACTCAAGAGGCCATAAACAAAGATAACAAAAAATCTCAACTCTTTGAGTG
 AGTGGAAAGTAAAGAAATCTTCAAAGACTAAGCGGTGCCATGGATGAACTCCACAACGAAATACTAGAAC
 TAGATGAGAAAAGTGGATGATCTCAGAGCTGATACAATAAGCTCACAATAAGAACTCGCAGTCTGCTTTC
 CAATGAAGGAATAATAAACAGTGAAGATGAACATCTCTTGGCGCTTGAAAGAAAGCTGAAGAAAATGCT
 GGGCCCCCTCTGCTGTAGAGATAGGGAATGGATGCTTTGAAACCAAAACACAAGTGCAACCAGACCTGTCTC
 GACAGAATAGCTGCTGGTACCTTTGATGCAGGAGAATTTCTCTCCCCACCTTTGATTCACTGAATATTAC
 TGCTGCATCTTTAAATGACGATGGATTGGATAATCATACTATACTGCTTTACTACTCAACTGCTGCCTCCA
 GTTTGGCTGTAACACTGATGATAGCTATCTTTGTTGTTATATGGTCTCCAGAGACAAATGTTTCTTGCTCC
 ATCTGTCTATAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGA
 GCGGTTTTCTGTGCTCAGAGTGTGTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTAGCAGGTC
 GTCCTTCAGCAAGGACACAAAAAGATTTTAAATTTATTAATAAAAAAAAAAAAAAAAAAGACCGGGAATTC
 GATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTTAAGATTGAATCCTGTT
 GCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTTACGTTAAGCATGTAATAATTAACATGTAATG
 CATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAA
 ACAAATATATGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCTCTAGAG
 TCTCAAGCTTGGCGCGGGGTACCGAGCTCGAATTCCGAGTGTACTTCAAGTCAGTTGGAATCAATAAAA
 TGATTATTTTATGAATATATTTCAATTGTGCAAGTAGATAGAAATTACATATGTTACATAACACACGAAATA
 AACAAAAAACACAATCCAAACAAACACCCCAACAAAATAACACTATATATATCCTCGTATGAGGAG
 AGGCACGTTCACTGACTCGACGATTCGGAGCAAAAAAGTCTCCCGTCACACATATAGTGGGTGACG
 CAATTATCTCAAAGTAATCCTTCTGTTGACTTGTGATTGATAACATCCAGTCTTCGTCAGGATTGCAAG
 AATTATAGAAGGGATCCACCTTTTATTTCTTCTTTTTCATATTTAGGGTTGACAGTGAAATCAGACT
 GGCAACCTATTAATTGCTTCCACAATGGGACGAACCTGAAGGGGATGTCGTCGATGATATTATAGGTGGC
 GTGTTTCATCGTAGTTGGTGAAGTCGATGGTCCCGTTCCAGTAGTTGTGTCGCCCCGAGACTTCTAGCCCAGG
 TGGTCTTTCCGGTACGAGTTGGTCCGCAGATGTAGAGGCTGGGGTGTCTGACCCAGTCTTCCCTCATCC
 TGGTTAGATCGGCCATCCACTCAAGGTCAGATTGTGCTTGATCGTAGGAGACAGGATGTATGAAAGTGTA
 GGCATCGATGCTTACATGATATAGGTGCGTCTCTCCAGTTGTGCAGATCTTCGTGGCAGCGGAGATCT
 GATTCTGTGAAGGGCGACACGTAAGTCTGCTCAGGTTGTGGAGGAAATAATTTGTTGGCTGAATATCCAGCC
 ATTGAAGCTTTGTTGCCCATTCATGAGGGAATCTTCTTTGATCATGTCAAGATACTCTCCTTAGACGTT
 GCAGTCTGGATAATAGTTCGCCATCGTGCTCAGATTGCGAGGAGACCTTATGATCTCGGAAATCTC
 CTCTGGTTTTAATATCTCCGCTCTTTGATATGTAATCAAGGACTTGTTTAGAGTTTCTAGCTGGCTGGATAT
 TAGGGTGATTTCTTCAAATCGAAAAAGAAGGATCCCTAATAACAAGGTTTTTTATCAAGCTGGATAAG
 AGCATGATAGTGGGTAGTGCCATCTTGATGAAGCTCAGAAGCAACACCAAGGAAGAAAATAAGAAAAG
 GTGTGAGTTTCTCCAGAGAACTGGAATAAATCATCTCTTTGAGATGAGCACTTGGGGTAGGTAAGGAA
 AACATATTTAGATTGGAGTCTGAAGTCTTGCTAGCAGAAGGCATGTTGTTGTGACTCCGAGGGGTTGCC
 TCAAACCTATCTTATAACCGGCGTGGAGGCATGGAGGCAAGGGCATTTTGGTAATTTAAGTAGTTAGTG
 GAAAAATGACGTCATTTACTTAAAGACGAAGTCTTGCGACAAGGGGGGGCCACGCCGAATTTAATATTAC
 CGGCGTGGCCCCACCTTATCGCGAGTGCTTTAGCACGAGCGGTCCAGATTTAAAGTAGAAAAGTTCCCGC
 CCACTAGGGTTAAAGGTGTTCACTATAAAAGCATATACGATGTGATGGTATTTGATGGAGCGTATATT
 GTATCAGGTATTTCCGTCGGATACGAATTATTCGTAC

Fig.8D

Trình tự axit amin của PDISP/HA từ cúm B/Brisbane/60/2008 có vòng thủy phân protein bị mất
(SEQ ID NO: 41)

MAKNVAIFGLLFSLLVLVPSQIFADRICTGITSSNSPHVVKATATQGEVNVTVGIPLTTTPTK
SHFANLKGTETRGKLCPKCLNCTDLVALGRPCKTGKIPSARVSILHEVRPVTSGCFPIM
HDRTKIRQLPNLLRGYEHIRLSTHNVINAENAPGGPYKIGTSGSCPNTNGNGFFATMAW
AVPKNDKNKTATNPLTIEVPYICTEGEDQITVWGFHSDNETQMAKLYGDSKPQKFTSSA
NGVTTHYVSQIGGFNPQTEDGGLPQSGRIVVDYMVQKSGKTGTITYQRGILLPQKVWCA
SGRSKVIKGSPLIGEADCLHEKYGGLNLSKPYTGEHAKAIGNCPIWVKTPLKLANGTK
YRPPGGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLSLSELEVKNLQ
RLSGAMDELHNEILEDEKVDDLADTISSQIELAVLLSNEGIINSEDEHLLALERKLLKM
LGPSAVEIGNGCFETKHKCNQTCLDRIAAGTFDAGEFSLPTFDSL NITAASLNDDGLDNH
TILLYYSTAASSLAVTLMIAIFVVYVMVSRDNVSCSICL

Fig.8E trình tự nucleotit của PDISP/HA từ cúm B/Brisbane/60/2008/ có vòng thủy phân protein (SEQ ID NO: 43)

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTGTGTTGGTTCCTTCTCAGATCTTCGCC
GATCGAATCTGCACTGGAATAACATCGTCAAACCTACCACATGTCGTCAAACCTGCTACTCAAGGGGAGG
TCAATGTGACTGGTGAATACCACTGACAACAACCCACCAAACTCATTTCGAAATCTCAAAGGAAC
AGAAACCAGGGGGAAAATATGCCAAAATGCCTCAACTGCACAGATCTGGACGTAGCCTTGGGCAGACC
AAAATGCACGGGGAAAATACCTCGGCAAGAGTTTCAATACTCCATGAAGTCAGACCTGTTACATCTGGG
TGCTTTCCTATAATGCACGACAGAACAATAATTAGACAGCTGCCTAACCTTCTCCGAGGATACGAACATA
TCAGGTTATCAACCCATAACGTTATCAATGCAGAAAATGCACAGGAGGACCCTACAAAATTGGAACCTC
AGGGTCTTGCCCTAACATTACCAATGGAAACGGATTTTCGCAACAATGGCTTGGGCCGTCCCAAAAAAC
GACAAAAACAAAACAGCAACAAATCCATTAACAATAGAAGTACCATAACATTTGTACAGAAGGAGAAGAC
CAAATTACCGTTTGGGGGTTCCTACTCTGACAACGAGACCCAAATGGCAAAGCTCTATGGGGACTCAAAGC
CCCAGAAGTTCACCTCATCTGCCAACGGAGTGACCACACATTACGTTTCACAGATTGGTGGCTTCCCAA
TCAAACAGAAGACGGGAGTACCACAAAGTGGTAGAATTGTTGTTGATTACATGGTGCAAAAATCTGG
GAAAACAGGAACAATTACCTATCAAAGGGGTATTTATTGCCTCAAAGGTGTGGTGCGCAAGTGGCAG
GAGCAAGGTAATAAAAGGATCCTTGCTTTAATTGGAGAAGCAGATTGCCTCCACGAAAAATACGGTGG
ATTAACAAAAAGCAAGCCTTACTACACAGGGGAACATGCAAAGGCCATAGGAAATTGCCAATATGGGT
GAAAACACCCCTTGAAGCTGGCCAATGGAACCAATATAGACCTCCTGGTGGAGGATGGGAAGGAATGAT
TGCAGGTTGGCACGGATACACATCCCATGGGGCACATGGAGTAGCGGTGGCAGCAGACCTTAAGAGCAC
TCAAGAGGCCATAAACAAGATAACAAAAATCTCAACTCTTTGAGTGAGCTGGAAGTAAAGAATCTTCA
AAGACTAAGCGGTGCCATGGATGAATCCACAACGAAATACTAGAACTAGATGAGAAAGTGGATGATCT
CAGAGCTGATACAATAAGCTCACAAATAGAACTCGCAGTCCTGCTTCCAATGAAGGAATAATAACAG
TGAAGATGAACATCTCTTGCGCTTGAAAGAAAGCTGAAGAAAATGCTGGGCCCCCTGCTGTAGAGAT
AGGGAATGGATGCTTTGAAACCAACACAAGTGCAACCAGACCTGTCTCGACAGAATAGCTGCTGGTAC
CTTTGATGCAGGAGAATTTCTCTCCCCACCTTGATTCACTGAATATTACTGCTGCATCTTAAATGACG
ATGGATTGGATAATCATACTATACTGCTTTACTACTCAACTGCTGCCTCCAGTTTGGCTGTAACACTGATG
ATAGCTATCTTTGTTGTTTATATGGTCTCCAGAGACAATGTTTCTTGCTCCATCTGTCTATAA

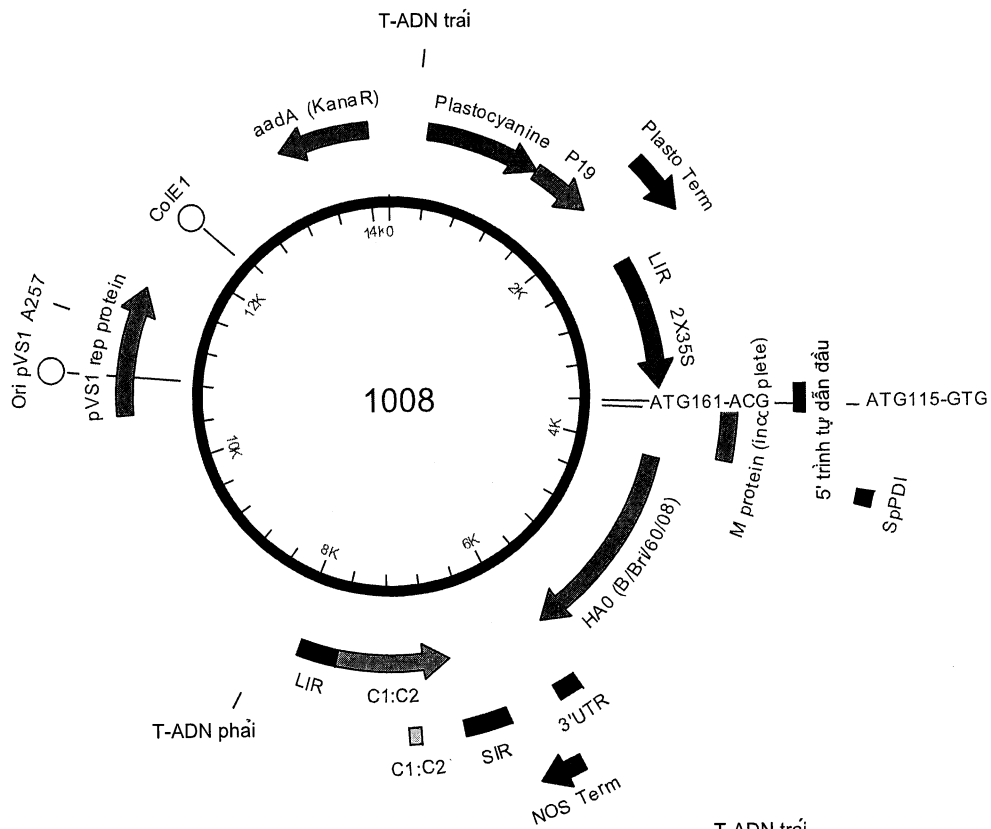


Fig.9

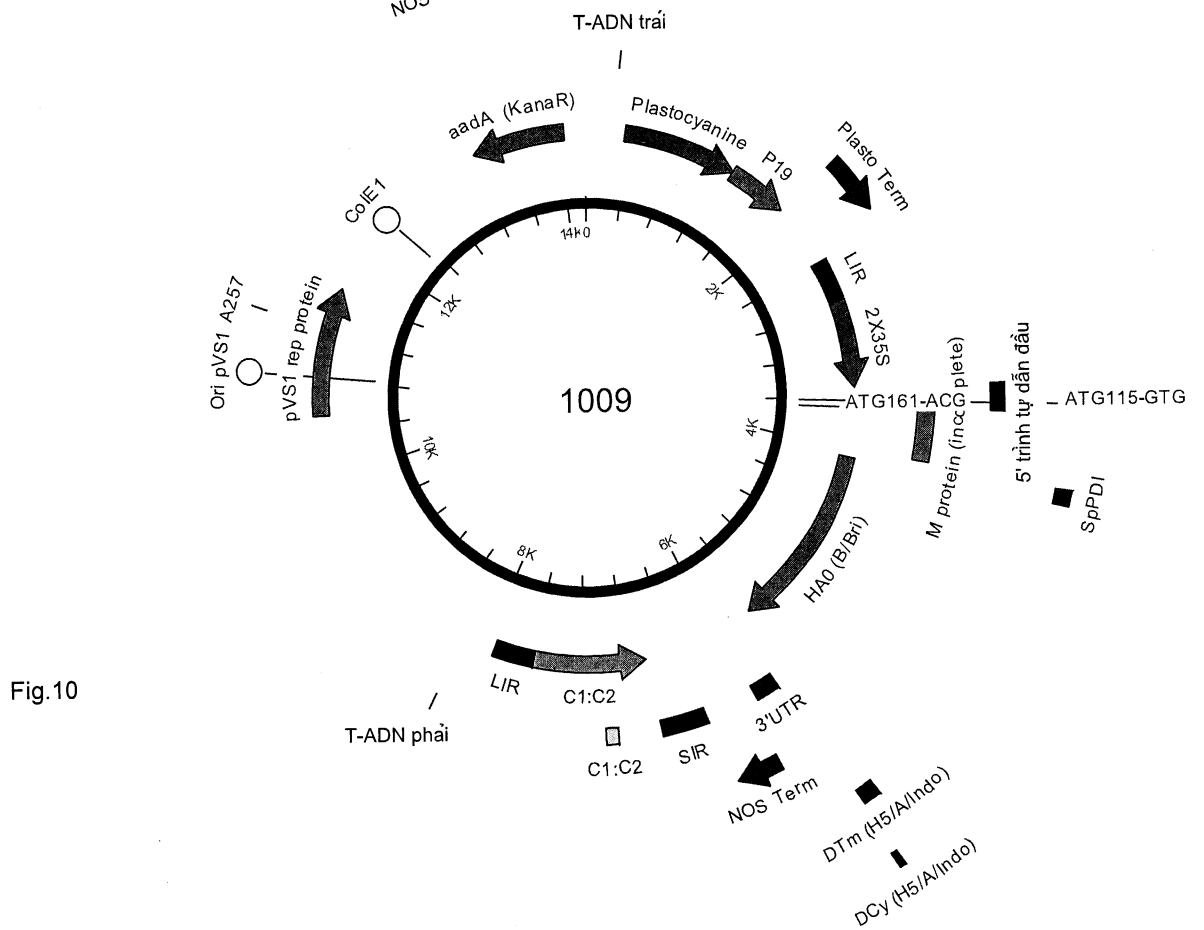


Fig.10

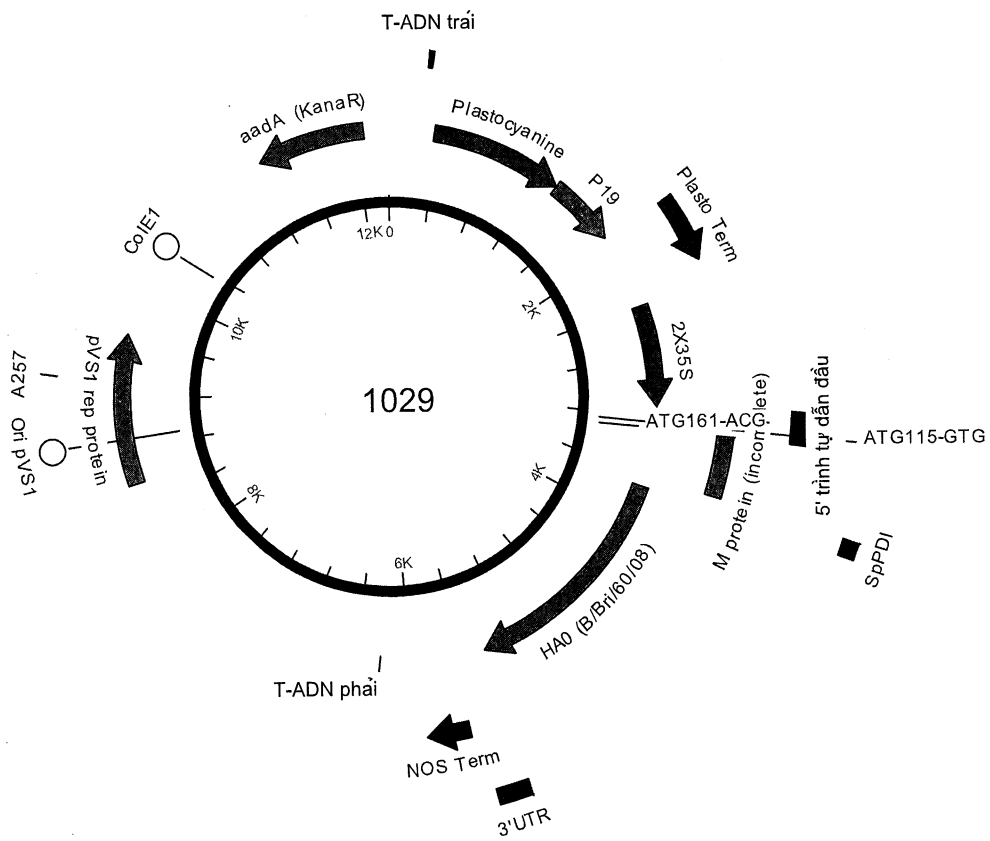


Fig.11

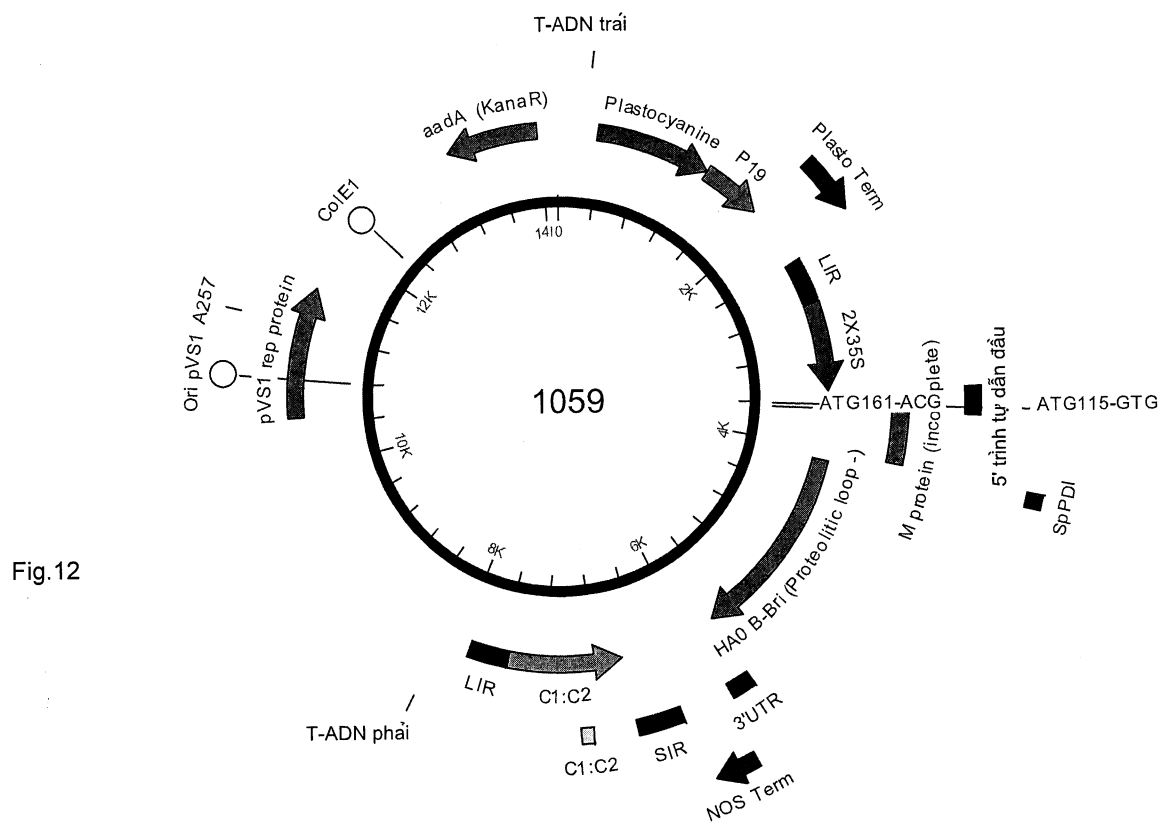


Fig.12

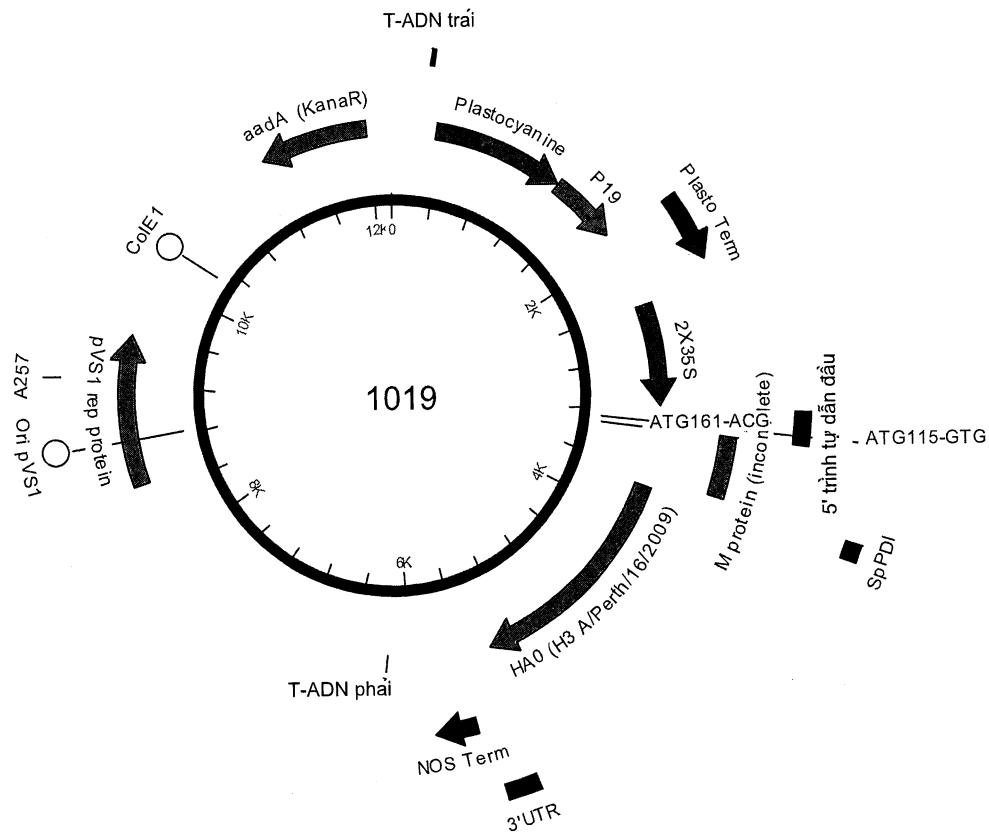


Fig.13

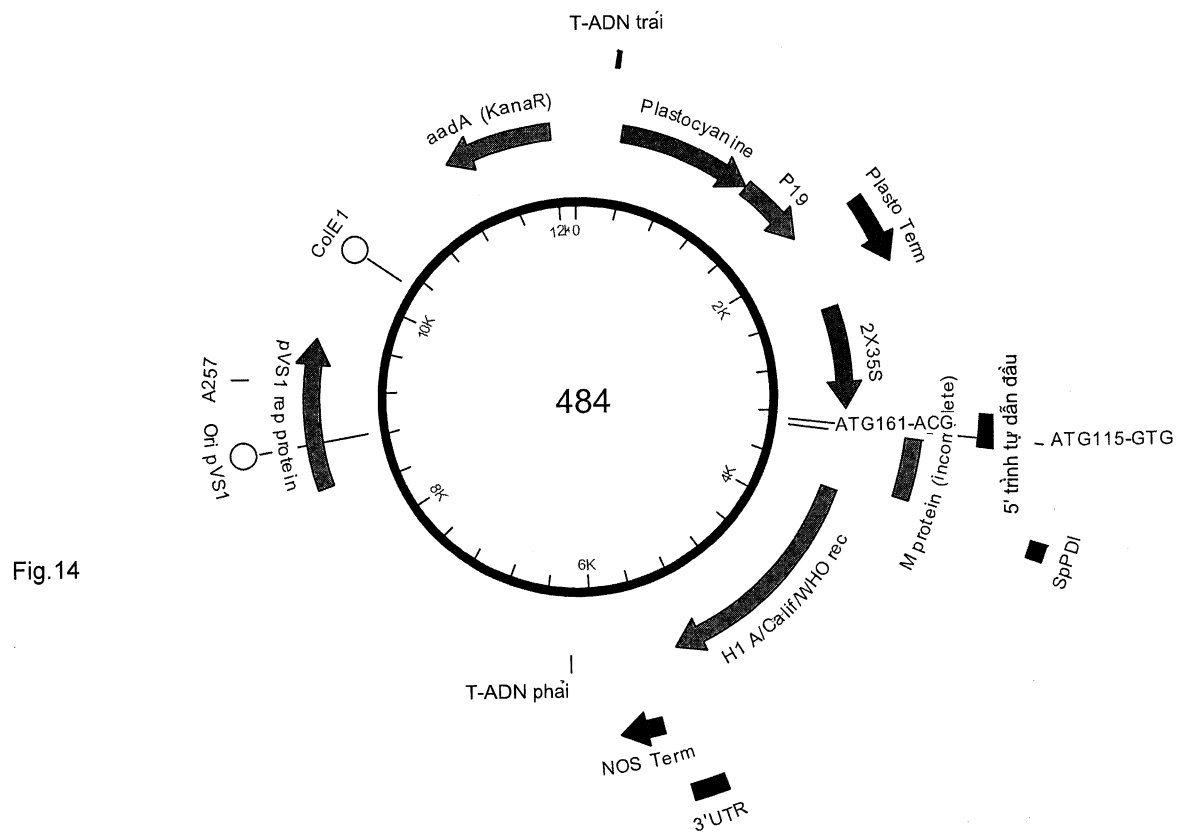


Fig.14

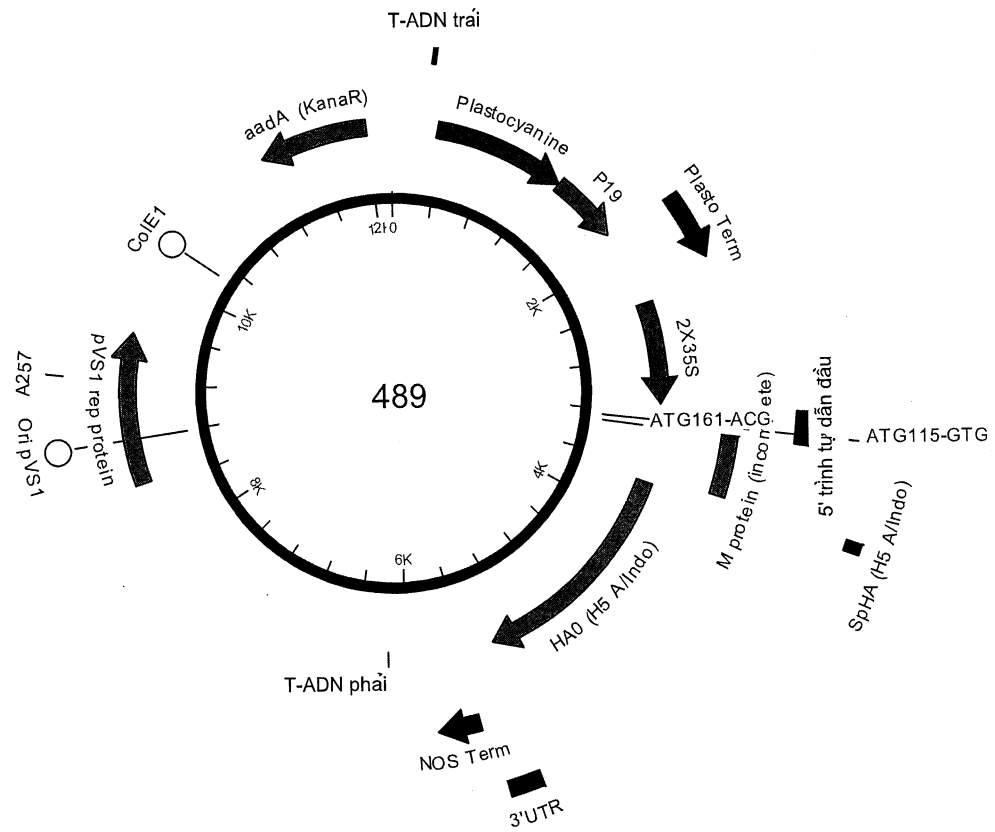


Fig.15

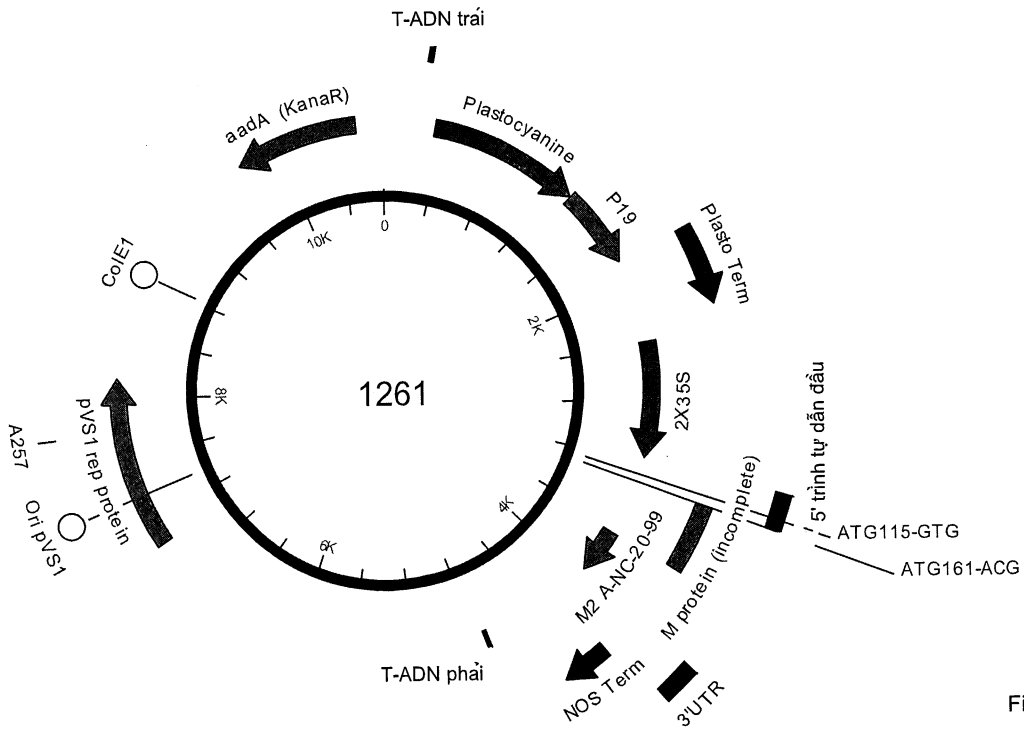


Fig.16

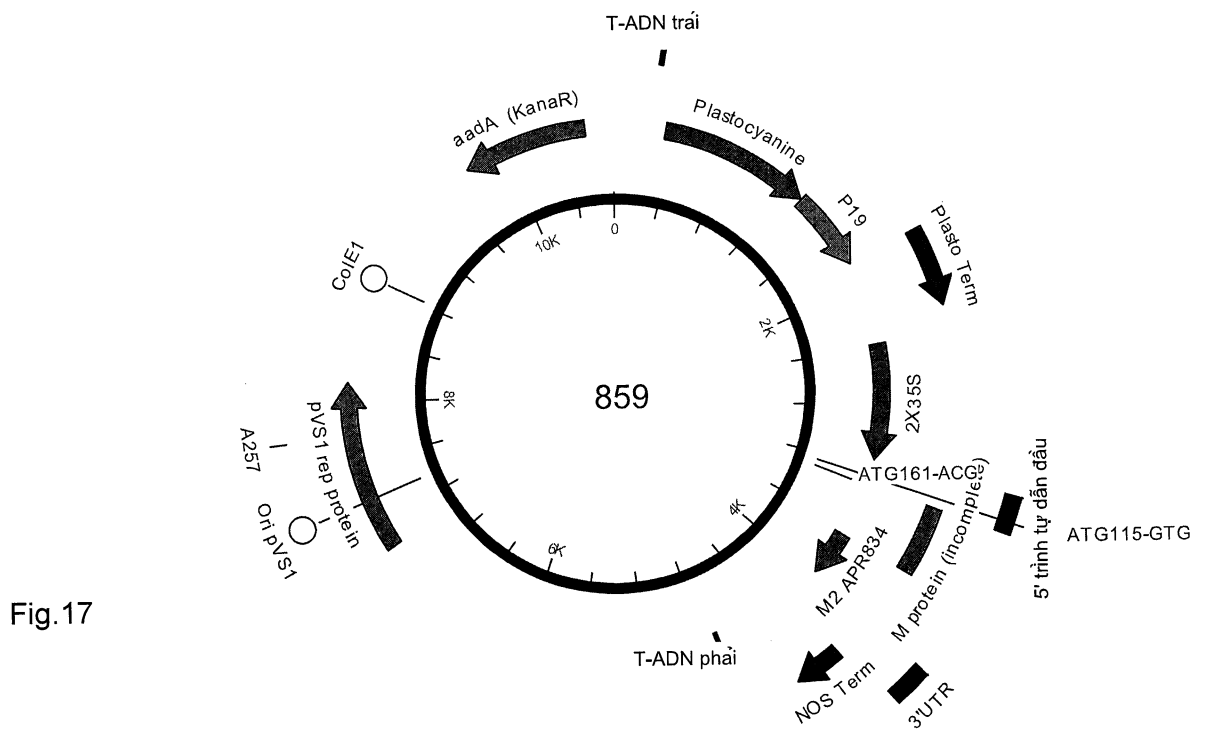


Fig.17

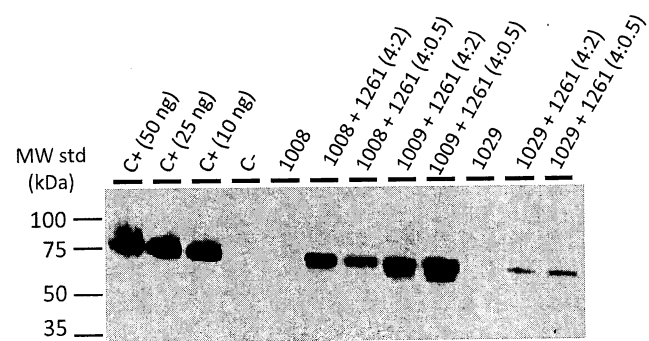


Fig.18

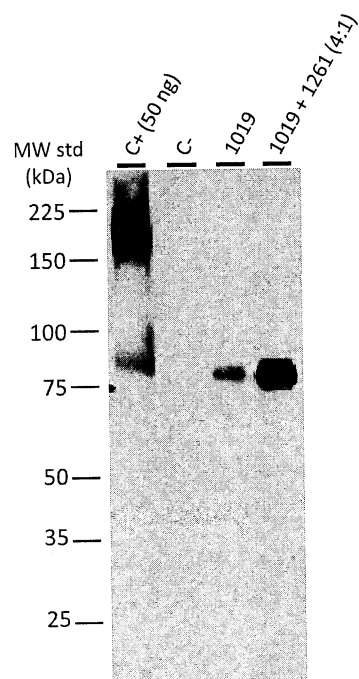


Fig.19

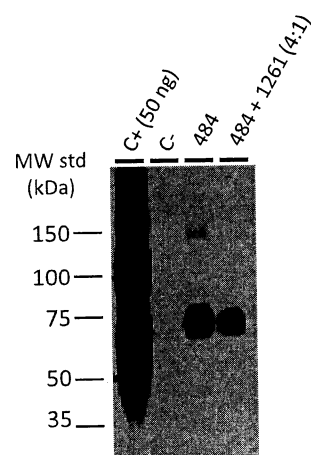


Fig.20

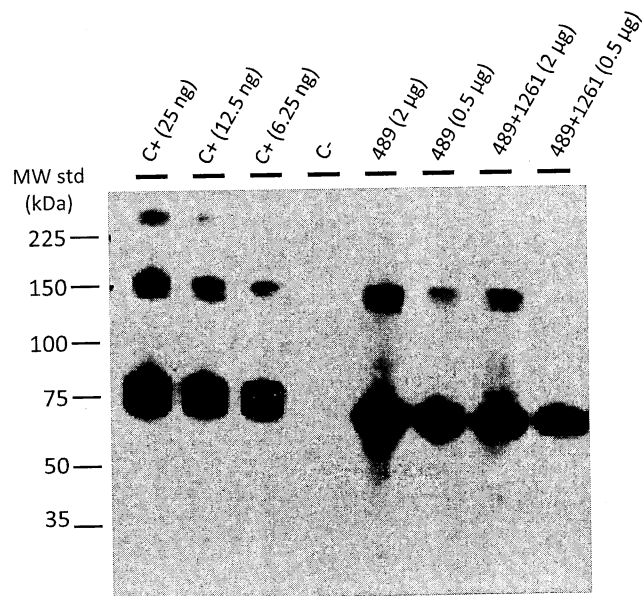


Fig.21

Fig.22A

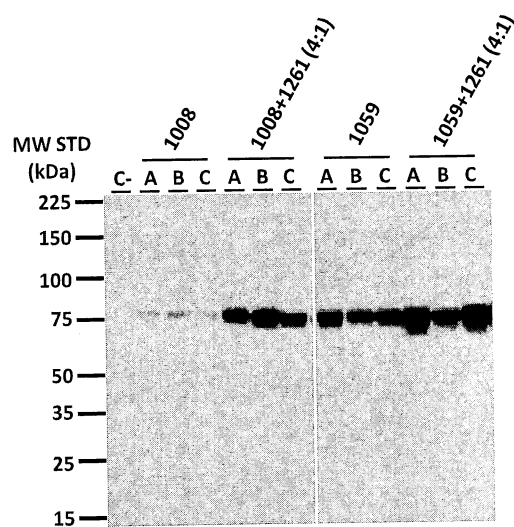


Fig.22B

Cấu trúc số	Khả năng hemagglutinin hóa (đơn vị HA/mg protein) ^a
1008	533
1008 + 1261 (4:1)	3200
1059	4267
1059 + 1261 (4:1)	34133

^a Nghịch đảo của lượng nhỏ nhất của protein tổng yêu cầu để phản ứng hemagglutinin hóa tích cực trong thể tích cuối là 200microlit

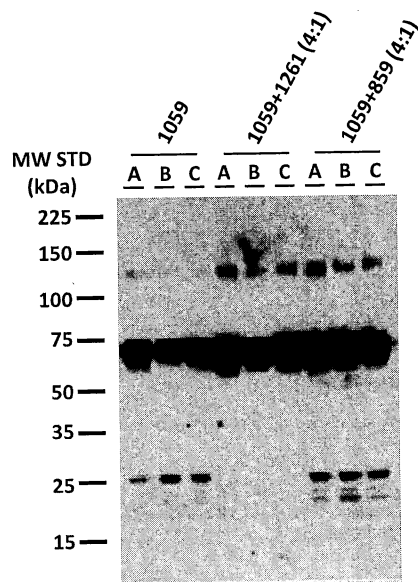


Fig.23A

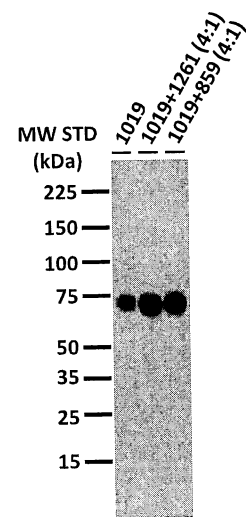


Fig.23B

	250	260	270	280	290
H1 New Cal	V R S A K L R M V T G L R N	- - - - -	- - - - -	I P S I Q S R G L F G A I A G F I E G G	
H1 Brisbane	V R S A K L R M V T G L R N	- - - - -	- - - - -	I P S I Q S R G L F G A I A G F I E G G	
H1 Sol Islands	V R S A K L R M V T G L R N	- - - - -	- - - - -	I P S I Q S R G L F G A I A G F I E G G	
H2 A. Singapore	V K S E K L V L A T G L R N	- - - - -	- - - - -	V P Q I E S R G L F G A I A G F I E G G	
H3 A. Brisbane	V K Q N T L K L A T G M R N	- - - - -	- - - - -	V P E K Q T R G I F G A I A G F I E G G	
H3 A. WCN	V K Q N T L K L A T G M R N	- - - - -	- - - - -	V P E K Q T R G I F G A I A G F I E G G	
H5 Anhui	V K S N K L V L A T G L R N	- - - - -	- - - - -	3 P L R E R R R K R G L F G A I A G F I E G G	
H5 Indo	V K S N R L V L A T G L R N	- - - - -	- - - - -	S P Q R E R R R K R G L F G A I A G F I E G G	
H5 Viet Nam	V K S N R L V L A T G L R N	- - - - -	- - - - -	S P Q R E R R R K R G L F G A I A G F I E G G	
H6 Teal HK	V K S E S L R L A T G L R N	- - - - -	- - - - -	V P Q I E T R G L F G A I A G F I E G G	
H7 Eq Prague	V K Q K S L M L A T G M K N	- - - - -	- - - - -	V P E A P A H K Q L T H H M R K K R G L F G A I A G F I E G G	
H9 A. HK	V R V N S L K L A V G L R N	- - - - -	- - - - -	V P A R S S R G L F G A I A G F I E G G	
B. Florida	A K - A I G N C P I W V K T P L K L A N G T K Y R P P A K L L K E R G F F G A I A G F L E G G				
B. Malaysia	A K - A I G N C P I W V K T P L K L A N G T K Y R P P A K L L K E R G F F G A I A G F L E G G				

Fig.24

A-2X35S/CPMV-HT/PDISP/H3 Victoria/NOS (Cấu trúc số 1391)

Fig.25A, SEQ ID NO: 44

IF-H3V36111.S2+4c

TCTCAGATCTTCGCCCAAAAACCTTCCTGGAAATGACAACAGCACGGCA

Fig.25B, SEQ ID NO:45

IF-H3V36111.s1-4r

ACTAAAGAAAATAGGCCTTCAAATGCAAATGTTGCACCTAATGTTGCCCTT

Fig.25C, SEQ ID NO: 46

Gen H3 được tổng hợp (tương ứng với nt 25 đến 1725 từ số phân lập trình tự GIS AID EPI_ISL101506 HA

ATGAAGACTATCATTGCTTTGAGCCACATTCTATGTCTGGTTTTCGCTCAAAAACCTTCCTGGAAATGACAA
CAGCACGGCAACGCTGTGCCTTGGGCACCATGCAGTACCAAACGGAACGATAGTGAAAACAATCACGAA
TGACCAAATTGAAGTTACTAATGCTACTGAGCTGGTTCAGAATTCCTCAATAGGTGAAATATGCGACAGT
CCTCATCAGATCCTTGATGGAGAAAACCTGCACACTAATAGATGCTCTATTGGGAGACCCCTCAGTGTGATG
GCTTCCAAAATAAGAAATGGGACCTTTTTGTTGAACGAAGCAAAAGCCTACAGCAACTGTTACCCTTATGA
TGTGCCGGATTATGCCTCCCTTAGGTCACTAGTTGCCTCATCCGGCACACTGGAGTTTAAACAATGAAAGCT
TCAATTGGACTGGAGTCACTCAAAACGGAACAAGTTCTGCTTGACATAAGGAGATCTAATAATAGTTTCTT
TAGTAGATTAAATTGGTTGACCCACTTAACTTCAAATACCCAGCATTGAACGTGACTATGCCAAACAAT
GAACAATTTGACAAATTGTACATTTGGGGGGTTCACCACCCGGGTACGGACAAGGACCAAAATCTTCCTGT
ATGCTCAATCATCAGGAAGAATCACAGTATCTACCAAAAAGAAGCCAACAAGCTGTAATCCCGAATATCG
GATCTAGACCCAGAATAAGGAATATCCCTAGCAGAATAAGCATCTATTGGACAATAGTAAAACCGGGAG
ACATACTTTTGATTAAACAGCACAGGGAATCTAATTGCTCCTAGGGGTTACTTCAAAAATACGAAGTGGGAA
AAGCTCAATAATGAGATCAGATGCACCCATTGGCAAATGCAATTCTGAATGCATCACTCCAAATGGAAGC
ATTCCCAATGACAAACCATTCCAAAATGTAAACAGGATCACATACGGGGCCTGTCCCAGATATGTTAAGC
AAAGCACTCTGAAATTGGCAACAGGAATGCGAAATGTACCAGAGAAACAACTAGAGGCATATTTGGCG
CAATAGCGGGTTTCATAGAAAATGGTTGGGAGGGAATGGTGGATGGTTGGTACGGTTTCAGGCATCAAA
ATTCTGAGGGAAGAGGACAAGCAGCAGATCTCAAAAGCACTCAAGCAGCAATCGATCAAATCAATGGGA
AGCTGAATCGATTGATCGGGAAAACCAACGAGAAATTCATCAGATTGAAAAAGAATTCTCAGAAGTCG
AAGGGAGAATTACAGGACCTTGAGAAATATGTTGAGGACACTAAAATAGATCTCTGGTCATACAACGCGG
AGCTTCTTGTGGCCTGGAGAACCAACATACAATTGATCTAACTGACTCAGAAATGAACAACTGTTTGA
AAAAACAAAGAAGCAACTAAGGGAAAAATGCTGAGGATATGGGCAATGGTTGTTTCAAAATATACCACAA
ATGTGACAATGCCTGCATAGGATCAATCAGAAATGGAAGTTATGACCACGATGTATACAGAGATGAAGC
ATTAAACAACCGGTTCCAGATCAAGGGAGTTGAGCTGAAGTCAGGGTACAAAGATTGGATCCTATGGAT
TTCCTTTGCCATATCATGTTTTTGTGTTGCTTTGTTGGGGTTCATCATGTGGGCCTGCCAAAAGGG
CAACATTAGGTGCAACATTGCATTGA

Fig.25D,SEQ ID NO: 47

Bảng biểu hiện số 1391 từ vùng khởi đầu 2X35S đến vùng kết thúc NOS. PDISP/H3 từ cúm A/ Victoria/361/2011 (H3N2) được gạch chân

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAA
 AGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCAGCTATCT
 GTCACCTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAA
 AGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCG
 TGGAAAAAAGGAAGACGTTCCAACACGTCCTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACG
 ACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCA
 ACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATA
 GTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCC
 TCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCA
 ACCACGTCCTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCCT
 ATCCTTCGCAAGACCCTTCCTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTTAAATCTTAATAGG
 TTTTGATAAAAGCGAACGTGGGGAACCCGAACCAACCTTCTTCTAACTCTCTCTCATCTCTCTTAAAG
 CAAACTTCTCTCTTGTCTTTCTTGCCTGAGCGATCTTCAACGTTGTGAGATCGTGCTTCGGCACCAGTACA
 ACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAAGTGGCGGCCATTAAATAACG
 TGTACTTGTCTATTTCTGTCTGGTGTGGTCTTGGGAAAAGAAAGCTTGTGGAGGCTGCTGTTTCAAGCCCCA
 TACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTG
 TTGCCTGTACTTCTTTCTTCTTCTTGTGCTGATTGGTCTTATAAGAAATCTAGTATTTTCTTTGAAACAGA
 GTTTTCCCGTGGTTTTCGAACTTGGAGAAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTGGGC
 CCTATGGCGAAAAACGTTGCGATTTCGGCTTATTGTTTTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCG
CCAAAAAACTTCTCGGAAATGACAACAGCACGGCAACGCTGTGCCTTGGGCACCATGCAGTACCAAACG
GAACGATAGTGAAAACAATCACGAATGACCAATTGAAGTTACTAATGCTACTGAGCTGGTTCAGAAATC
CTCAATAGGTGAAAATATGCGACAGTCTCATCAGATCCTTGATGGAGAAAAGTGCACACTAATAGATGCT
CTATTGGGAGACCCTCAGTGTGATGGCTTCCAAAATAAGAAATGGGACCTTTTGTGTAACGAAGCAAAG
CCTACAGCAACTGTTACCTTATGATGTGCCGGATTATGCCTCCCTTAGGTCACTAGTTGCCATCCTGGC
ACACTGGAGTTAACAAATGAAAGCTTCAATTGGACTGGAGTCACTCAAAACGGAACAAGTTCTGCTTGCA
TAAGGAGATCTAATAATAGTTTCTTTAGTAGATTAATTTGGTTGACCCACTTAAACTTCAAATACCCAGC
ATTGAACGTGACTATGCCAAACAATGAACAATTTGACAAATTGTACATTTGGGGGGTTCCACCCCGGGT
ACGGACAAGGACCAATCTTCCTGTATGCTCAATCAGGAAGAATCACAGTATCTACCAAAAGAAGC
CAACAAGCTGTAATCCCGAATATCGGATCTAGACCCAGAATAAGGAATATCCCTAGCAGAATAAGCATC
TATTGGACAATAGTAAAACCGGGAGACATACTTTTGATTAAACAGCACAGGGAATCTAATTGCTCTAGGG
GTTACTTCAAAAACGAAAGTGGGAAAAGCTCAATAATGAGATCAGATGCACCCATTGGCAAATGCAATT
CTGAATGCATCACTCCAAATGGAAAGCATTCCCAATGACAAACCATTCCAAATGTAAACAGGATCACATA
CGGGGCCTGTCCCAGATATGTTAAGCAAAGCACTCTGAAATTGGCAACAGGAATGCGAAATGTACCAGA
GAAAAAACTAGAGGCATATTTGGCGCAATAGCGGGTTTCATAGAAAATGGTTGGGAGGGAAATGGTGGA
TGGTTGGTACGGTTTCAGGCATCAAAATTCTGAGGGAAAGAGGACAAGCAGCAGATCTCAAAAGCACTCA
AGCAGCAATCGATCAAAATCAATGGGAAGCTGAATCGATTGATCGGGAAAACCAACGAGAAATTCATCA
GATTGAAAAAGAATTCTCAGAAGTCGAAGGGAGAAATTCAGGACCTTGAGAAATATGTTGAGGACACTAA
AATAGATCTCTGGTCATACAACGCGGAGCTTCTTGTGCTTGGAGAACCAACATACAAATTGATCTAACT
GACTCAGAAATGAAACAACTGTTTGAAAAAACAAGCAACTAAGGGAAAAATGCTGAGGATATGGG
CAATGGTTGTTCAAAATATACCACAAATGTGACAAATGCCTGCATAGGATCAATCAGAAATGGAACCTAT
GACCACGATGTATACAGAGATGAAGCATTAAACAACCGGTTCCAGATCAAGGGAGTTGAGCTGAAGTCA
GGGTACAAAGATTGGATCCTATGGATTTCCTTTGCCATATCATGTTTTTGTCTTGTGCTTTGTTGGGG
TTCATCATGTGGGCCTGCCAAAAGGGCAACATTAGGTGCAACATTTGCATTTGAAGGCCTATTTCTTTAG
 TTTGAATTTACTGTTATTTCGGTGTGATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTAT
 TTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCTGCCCTCAGCAAGGACACAAAAAGATT
 TTAATTTTATTAATAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGT
 TCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCAGTATTATCATATAATT
 TCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTA

TGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATA
AATTATCGCGCGCGGTGTCATCTATGTTACTAGAT

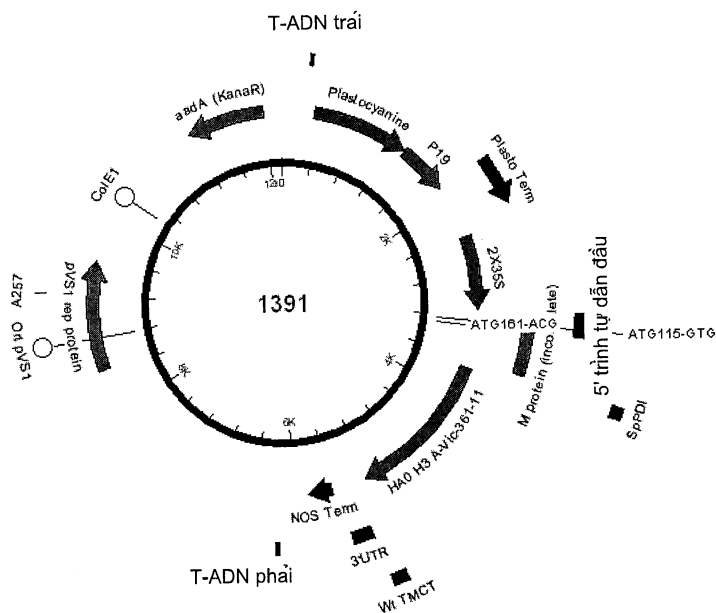
Fig.25E, SEQ ID NO:48

Trình tự axit amin củ PDISP-H3 từ cúm A/Victoria/361/2011 (H3N2)

MAKNVAIFGLLFSLVLVPSQIFAQKLPGNDNSTATLCLGHHAVPNGTIVKTITNDQIEVT
NATELVQNSSIGEICDSPHQILDGENCTLIDALLGDPQCDGFQNKKWDLFVERSKAYSNC
YPYDVPDYASLRSLVASSGTLEFNNEFNWTGVTQNGTSSACIRRSNNSFFSRLNWLTHL
NFKYPALNVTMPNNEQFDKLYIWGVHHPGTDKQIFLYAQSSGRITVSTKRSQQAVIPNI
GSRPRIRNIPSRISYWTIVKPGDILLINSTGNLIAPRGYFKIRSGKSSIMRSDAPIGKCNSECI
TPNGSIPNDKPFQNVNRITYGACPRYVKQSTLKLATGMRNVPEKQTRGIFGAIAGFIENG
WEGMVDGWYGFRHQNSEGRGQAADLKSTQAAIDQINGKLNRLIGKTNEKFHQIEKEFS
EVEGRIQDLEKYVEDTKIDLWSYNAELLVALENQHTIDLTSEMKNLFEKTKKQLRENA
EDMGNGCFKIYHKCDNACIGSIRNGTYDHDVYRDEALNNRFQIKGVELKSGYKDWILWI
SFAISCFLLCVALLGFIMWACQKGNIRCNICI*

Fig.25F

Bản đồ của cấu trúc số 1391



B-2X35S/CPMV-HT/HA B Wisconsin/NOS vào hệ khuếch đại BeYDV(m)+Replicaza (Cấu trúc số 1462

Fig.26A, SEQ ID NO: 49

IF-HAB110.S1+3c

AAATTTGTCGGGCCCATGAAGGCAATAATTGTACTACTCATGGTAG

Fig.26B, SEQ ID NO:50

IF-HAB110.s1-4r

ACTAAAGAAAATAGGCCTTTATAGACAGATGGAGCATGAAACGTTGTCTCTG

Fig.26C,SEQ ID NO: 51

HA B Wisconsin được tổng hợp (số truy cập Genbank JN993010)

ATGAAGGCAATAATTGTACTACTCATGGTAGTAACATCCAATGCAGATCGAATCTGCACTGGGATAACAT
CTTCAAACCTCACCTCATGTGGTCAAAACAGCTACTCAAGGGGAGGTCAATGTGACTGGCGTGATACCACT
GACAACAACACCAACAAAATCTTATTTTGCAAATCTCAAAGGAACAAGGACCAGAGGGAAACTATGCC
GGACTGTCTCAACTGTACAGATCTGGATGTGGCCTTGGGCAGGCCAATGTGTGTGGGGACCACACCTTCT
GCTAAAGCTTCAATACTCCACGAGGTGACAGCTGTTACATCCGGGTGCTTTCCTATAATGCACGACAGAA
CAAAAATCAGGCAACTACCCAATCTTCTCAGAGGATATGAAAATATCAGGTTATCAACCCAAAACGTTAT
CGATGCAGAAAAAGCACCAGGAGGACCTACAGACTTGGAACTCAGGATCTTGCCTAACGCTACCAAG
TAAAATCGGATTTTGTGCAACAATGGCTTGGGCTGTCCCAAAGGACAACCTACAAAATGCAACGAACCCA
CTAACAGTAGAAGTACCATACATTTGTACAGAAGGGGAAGACCAAATTACTGTTTGGGGGTTCCATTAG
ATAACAAAACCCAAATGAAGAGCCTCTATGGAGACTCAAATCCTCAAAGTTACCTCATCTGCTAATGG
AGTAACCACACATTATGTTTCTCAGATTGGCGACTTCCAGATCAAACAGAAGACGGAGGACTACCACAA
AGCGGCAGAATTGTTGTTGATTACATGATGCAAAAACCTGGGAAAACAGGAACAATTGTCTATCAAAGA
GGTGTGTTTGTGCTTCAAAGGTGTGGTGCAGGAGTGGCAGGAGCAAAGTAATAAAAGGGTCATTGCCTT
TAATTGGTGAAGCAGATTGCCTTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAG
GAGAACATGCAAAAAGCCATAGGAAATTGCCAATATGGGTAAAAACACCTTTGAAGCTTGCCAATGGAA
CCAAATATAGACCTCCTGCAAAACTATTGAAGGAAAGGGGTTTCTTCGGAGCTATTGCTGGTTTCCTAGA
AGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGATACACATCTCACGGAGCACATGGAGTGGCAGT
GGCGGCAGACCTTAAGAGTACACAAGAAGCTATAAAATAAGATAACAAAAAATCTCAATTCTTTGAGTGA
GCTAGAAGTAAAGAACCTTCAAAGACTAAGTGGTGCCATGGATGAACTCCACAACGAAATACTCGAGCT
GGATGAGAAAAGTGGATGATCTCAGAGCTGACACTATAAGCTCACAATAGAACTTGCAGTCTTGCTTTCC
AACGAAGGAATAATAAACAGTGAAGACGAGCATCTATTGGCACTTGAGAGAAAACCTAAAGAAAATGCTG
GGTCCCTCTGCTGTAGACATAGGAAACGGATGCTTCGAAACCAAACACAAATGCAACCAGACCTGCTTA
GACAGGATAGCTGCTGGCACCTTTAATGCAGGAGAATTTCTCTCCCACTTTTGATTCAATTGAACATTAC
TGCTGCATCTTTAAATGATGATGGATTGGATAACCATACTATACTGCTCTATTACTCAACTGCTGCTCTA
GTTTGGCTGTAAACATTAATGCTAGCTATTTTATTGTTTATATGGTCTCCAGAGACAACGTTTCATGCTCC
ATCTGTCTATAA

Fig.26D

Bản đồ của cấu trúc 193. Các vị trí enzym giới hạn SacII và StuI được sử dụng để kéo dài plasmid được giảm độc lực trên bản đồ này

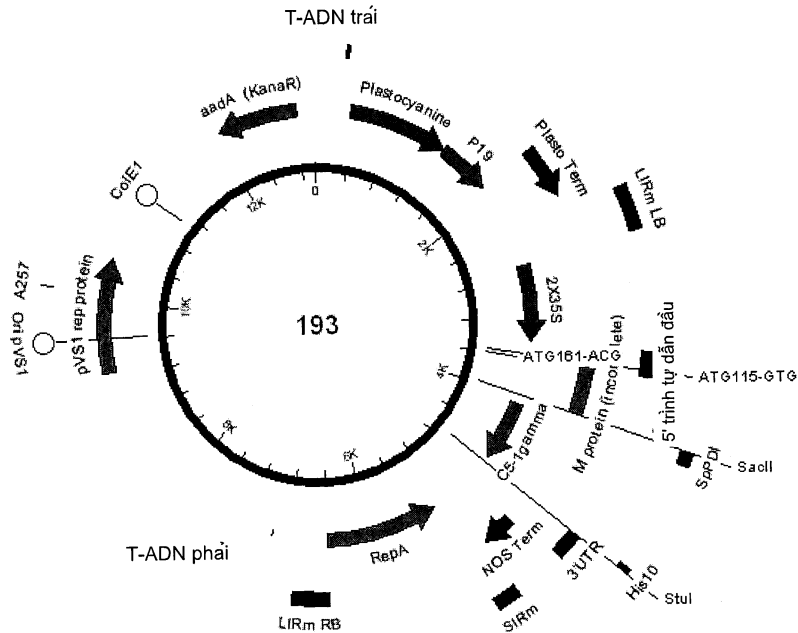


Fig.26E, SEQ ID NO: 52

Cấu trúc 193 từ các bộ ADN vận chuyển từ trái sang phải (được gạch chân). 2X35S/CPMV-HT/NOS vào hệ khuếch đại BeYDV(m)+Replicaza có băng biểu hiện chất ức chế bất hoạt Plastoxyanin-P19-Plastoxyanin

TGGCAGGATATATTGTGGTGTAACAAATTGACGCTTAGACAACCTTAATAACACATTGCGGACGTTTTTA
ATGTACTGAATTAACGCCGAATCCCGGGCTGGTATATTTATATGTTGTCAAATAACTCAAAAACCATAAA
AGTTTAAGTTAGCAAGTGTGTACATTTTTACTTGAACAAAAATATTCACCTACTACTGTTATAAATCATT
TTAAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTGATATTTTGACAACAATTTTGTG
CAACATTTGAGAAAAATTTTGTGTTCTCTCTTTTCATTGGTCAAAAACAATAGAGAGAGAAAAAGGAAGA
GGGAGAATAAAAAACATAATGTGAGTATGAGAGAGAAAGTTGTACAAAAGTTGTACAAAATAGTTGTAC
AAATATCATTGAGGAATTTGACAAAAGCTACACAAATAAGGGTTAATTGCTGTAATAAATAAGGATGA
CGCATTAGAGAGATGTACCATTAGAGAATTTTGGCAAGTCATTAAAAAGAAAGAATAAATTATTTTTAA
AATTAAGAGTTGAGTCATTTGATTAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGT
TGTATTAGTAATTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCCTATATATTGCCCATAGA
GTCAGTTAACTCATTTTATATTTTCATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAA
AAAAAACGGTATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTAGGAGGATA
ACATCCAATCCAACCAATCACAACAATCCTGATGAGATAACCCACTTTAAGCCACGCATCTGTGGCACA
TCTACATTATCTAAATCACACATTCTTCCACACATCTGAGCCACACAAAAACCAATCCACATCTTTATCAC
CCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAG
ACTAATTAATTAATATCATCTTGAGAGAAAATGGAACGAGCTATACAAGGAAACGACGCTAGGGAAC

AAGCTAACAGTGAACAGTGGGATGGAGGATCAGGAGTACCATTCTCCCTTCAAACCTCTCTGACGAAAG
TCCGAGTTGGACTGAGTGGCGGCTACATAACGATGAGACGAATTCTGAATCAAGATAATCCCCTTGGTTTC
AAGGAAAGCTGGGGTTTCGGGAAAGTTGTATTTAAGAGATATCTCAGATACGACAGGACGGAAGCTTCA
CTGCACAGAGTCTTGGATCTTGGACGGGAGATTTCGGTAACTATGCAGCATCTCGATTTTTCGGTTTCGA
CCAGATCGGATGTACCTATAGTATTCGGTTTCGAGGAGTTAGTATCACCGTTTCTGGAGGGTCGCGAACT
CTTCAGCATCTCTGTGAGATGGCAATTCGGTCTAAGCAAGAACTGCTACAGCTTGCCCCAATCGAAGTGG
AAAGTAATGTATCAAGAGGATGCCCTGAAGGTACTCAAACCTTCGAAAAAGAAAGCGAGTAAGTTAAAA
TGCTTCTTCGTCTCCTATTTATAATATGGTTTGTATTGTTAATTTTGTCTTGTAGAAGAGCTTAATTAAT
CGTTGTTGTTATGAAATACTATTTGTATGAGATGAACTGGTGTAAATGTAATTCATTACATAAGATGGAGTC
AGAATCAGAATGTTTCTCCATAACTAACTAGACATGAAGACCTGCCGCGTACAATTGTCTTATATTGTA
ACAACTAAAAATTGAACATCTTTTGCCACAACCTTATAAGTGGTTAATATAGCTCAAATATGATTGTCAGGTT
CAATAGATTAATAATGGAATATCAGTTATCGAAATTCATTAACAATCAACCTTAACGTTATTAECTACTA
ATTTTATATCATCCCCCTTTGATAAATGATAGTACACCAATTAGGAAGGAGCATGCTCGCCTAGGAGATTG
TCGTTTCCCGCCTTCAGTTTGGCAAGCTGCTCTAGCCGTGTAGCCAATACGCAAACCGCCTCTCCCCGCGCG
TGTTGAATTACTAGCGCGTGTGAGACGCGTTGTTGTTGTGACTCCGAGGGGTGCGCTCAAACCTCTATCTT
ATAACCGCGCTGGAGGCATGGAGGCAGGGGTATTTTGGTCATTTTAATAGATAGTGGAAAATGACGTGG
AATTTACTTAAAGACGAAGTCTTTGCGACAAGGGGGGGGCCACGCCGAATTAATATTACCGCGGTGGCC
CCCCCTTATCGCGAGTGCTTTAGCACGAGCGGTCCAGATTTAAAGTAGAAAATTTCCCGCCCACTAGGGT
TAAAGGTGTTACACTATAAAAGCATATACGATGTGATGGTATTTGGTCGACAAGCTTGCATGCCGGTCA
ACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGG
CAATTGAGACTTTTCAACAAAGGGTAATATCCGGAACCTCCTCGGATTCCATTGCCAGCTATCTGTCA
CTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCA
CATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCAACCCAGGAGGAGCATCGTGGGA
AAAAGAAGACGTTCCAACCACGTTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACGACAC
ACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAA
AGGGTAATATCCGGAACCTCCGGAATTTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGG
AAAAGGAAGGTGGTCTCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTG
CCGACAGTGGTCCCAAAGATGGACCCCAACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCA
CGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCC
TTCGCAAGACCCCTTCTCTATATAAGGAAGTTTCAATTCATTTGGAGAGGTATTAATACTTAATAGGTTTT
GATAAAAGCGAACGTGGGGAAACCCGAACCAAACCTTCTTCAAACCTCTCTCTCATCTCTCTTAAAGCAA
ACTTCTCTCTTGTCTTTCTTGGGTGAGCGATCTTCAACGTTGTGAGATCGTGCTTCGGCACCAAGTACAACG
TTTTCTTCACTGAAGCGAAATCAAAGATCTCTTTGTGGACAGTAGTGCGGCGCCATTAAATAACGTGT
ACTTGTCTTATTCTTGTGCGGTGTGGTCTTGGGAAAAGAAAGCTTGTCTGGAGGCTGCTGTTCAGCCCCATC
ATTACTTGTTACGATTCTGCTGACTTTTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGATTGTTG
CCTGTACTTCTTTCTTCTTCTTCTTGTGCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTGAACAGAGTT
TTCCCGTGGTTTTCGAAGCTTGGAGAAAGATTTGTAAGCTCTGTATATTCTGCCCCAAATTTGTGCGGCCCC
CGGATGGCGAAAAACGTGGCAGTTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTC
GCCGTGAGGCTCCTCAGCCAAAACGACACCCCATCTGTCTATCCACTGGCCCCCTGGATCTGCTGCCCAA
ACTAAGTCCATGGTGACCCTGGGATGCCTGGTCAAGGGCTATTTCCCTGAGCCAGTGACAGTGACCTGGA
ACTCTGGATCCCTGTCCAGCGGTGTGCACACCTTCCCAGCTGTCTGACGTCTGACCTCTACACTCTGAGC
AGCTCAGTGACTGTCCCTCCAGCACCTGGCCCAGCGAGACCGTCACCTGCAACGTTGCCACCCGGCCA
GCAGCACCAAGGTGGACAAGAAAATTGTGCCAGGGATTGTGGTTGTAAGCCTTGCAATGTACAGTCCC
AGAAGTATCATCTGTCTTCATCTTCCCCCAAAGGCCAAAGGATGTGCTCACCATTACTCTGACTCCTAAGG
TCAGTGTGTTGTGGTAGACATCAGCAAGGATGATCCCGAGGTCCAGTTCAGCTGGTTGTAGATGATGT
GGAGGTGCACACAGCTCAGACGCAACCCCGGGAGGAGCAGTTCAACAGCACTTTCGGCTCAGTCAGTGA
ACTTCCCATCATGCACCAGGACTGGCTCAATGGCAAGGAGCGATCGCTCACCATCACCATCACCATCACC
ATCACCATTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAG
CGGTTTTCTGTGCTCAGAGTGTGTTTATTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTG
TCCCTTCAGCAAGGACACAAAAAGATTTTATTTATTAATAAAAAAAAAAAAAAAAAAGACCGGAATTG
ATATCAAGCTTATCGACATCTGAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTG
CCGGTCTTGCATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGC
ATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAA
ACAAAATATAGCGCGCAAACCTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCTCTAGAG

TCTCAAGCTTGGCGCGCCATAAAATGATTATTTTATGAATATATTTTCATTGTGCAAGTAGATAGAAATTAC
 ATATGTTACATAACACACGAAATAAACAAAAAAGACAATCCAAAAACAAACACCCCAAAAAAATAA
 TCACTTTAGATAAACTCGTATGAGGAGAGGCACGTTTCAGTGAAGTACGATTCCTCGTGTGACTTGTTCATTGATAAC
 CCCCCTCACACATATAGTGGGTGACGCAATTATCTTTAAAGTAATCCTTCTGTGACTTGTTCATTGATAAC
 ATCCAGTCTTCGTCAGGATTGCAAAGAATTATAGAAGGGATCCACCTTTTATTTTCTTTTTCATAT
 TTAGGGTTGACAGTGAAATCAGACTGGCAACCTATTAATTGCTTCCACAATGGGACGAACCTGAAGGGGA
 TGTCGTCGATGATATTATAGGTGGCGTGTTCATCGTAGTTGGTGAAATCGATGGTACCGTTCCAATAGTTG
 TGTCGTCGAGACTTCTAGCCAGGTGGTCTTTCCGGTACGAGTTGGTCCGAGATGTAGAGGCTGGGGT
 GTCGGATTCCATTCTTCCATTGTCTGGTTAAATCGGCCATCCATTCAAGGTCAGATTGAGCTTGTGGT
 ATGAGACAGGATGTATGTAAGTATAAGCGTCTATGCTTACATGGTATAGATGGGTTTCCCTCCAGGAGTG
 TAGATCTTCGTGGCAGCGAAGATCTGATTCTGTGAAGGGCGACACATACGGTTCAGGTTGTGGAGGGAAT
 AATTTGTTGGCTGAATATTCAGCCATTGAAGTTTGTGTTGCCATTTCATGAGGGAATTCTTCTTGATCAT
 GTCAAGATATTCCTCCTTAGACGTTGCAGTCTGGATAATAGTTCTCCATCGTGCCTCAGATTGTGCGAGGAG
 AGACCTTATGATCTCGGAAATCTCCTCTGGTTTAAATATCTCCGTCCTTGATATGTAATCAAGGACTTGT
 TTAGAGTTTCTAGCTGGCTGGATATTAGGGTGATTTCCTTCAAAATCGAAAAAAGAAGGATCCCTAATAC
 AAGGTTTTTATCAAGCTGGAGAAGAGCATGATAGTGGGTAGTGCCATCTTGATGAAGCTCAGAAGCAAC
 ACCAAGGAAGAAAATAAGAAAAGGTGTGAGTTTCTCCAGAGAACTGGAATAAAATCATCTCTTTGAGA
 TGAGCACTTGGGATAGGTAAGGAAAAACATATTAGATTGGAGTCTGAAGTTCTTACTAGCAGAAGGCATG
 TTGTTGTGACTCCGAGGGGTGCTTCAAACTCTATCTTATAACCGGCGTGGAGGCATGGAGGCAGGGGTA
 TTTTGGTCATTTTAAATAGATAGTGGAATAAGAGTGAATTTACTTAAAGACGAAGTCTTTGCGACAAGG
 GGGGGCCACGCCGAATTTAATATTACCGCGTGGCCCCCTTATCGCGAGTGCTTTAGCAGGAGCGGT
 CCAGATTAAAGTAGAAAAATTTCCCGCCCACTAGGGTTAAAGGTGTTCACTATAAAAGCATATACGAT
 GTGATGGTATTGACTAGTGGCACTGGCCGTCGTTTACAACGTCGTGACTGGGAAAAACCTGGCGTTAC
 CCAACTTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCCGACCGAT
 CGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGCTAGAGCAGCTTGAGCTTGGATCAGATTGTGCG
 TTTCCCGCCTTCAGTTTAACTATCAGTGTGTTGACAGGATATATTGGCGGGTAAACCTAAGAGAAAAGAG
CGTTTA

Fig.26F, SEQ ID NO: 53

Bảng biểu hiện số 1462 từ vùng 2X35S đến vùng kết thúc NOS. HA từ cúm B/Wisconsin/1/2010 được
 gạch chân

GTCAACATGGTGGAGCAGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAA
 AGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAACCTCCTCGGATTCCATTGCCAGCTATCT
 GTCACCTTATTGTGAAGATAGTGGAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAA
 AGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCAACCCACGAGGAGCATCG
 TGGAAAAAGAAGACGTTCCAACACGCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACG
 ACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCA
 ACAAAGGGTAATATCCGGAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATA
 GTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCC
 TCTGCCGACAGTGGTCCCAAAGATGGACCCCAACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCA
 ACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACT
 ATCCTTCGCAAGACCCTTCCTCTATATAAGGAAGTTCAATTTGAGAGGTTATTAATCTTAAATAGG
 TTTTGATAAAAGCGAACGTGGGGAAACCCGAACCAACCTTCTTCTAAACTCTCTCTCATCTCTTAAAG
 CAAACTCTCTCTTGTCTTTCTTGGCTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACA
 ACGTTTTCTTCACTGAAGCGAAATCAAAGATCTCTTGTGGACACGATAGTGGCGGCCATTAAATAACG
 TGTACTTGTCTATTCTTGTGGTGTGGTCTTGGGAAAAGAAAGCTTGCTGGAGGCTGCTGTCAGCCCCA
 TACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTTACTTCTGCTTGACGAGGTATTG
 TTGCCTGTACTTCTTCTTCTTCTTGTGATTGGTTCTATAAGAAATCTAGTATTTCTTTGAAACAGA
 GTTTTCCCGTGGTTTTCGAACCTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCAAATTTGTCGGGC
 CCATGAAGGCAATAATTGTACTACTCATGGTAGTAACATCCAATGCAGATCGAATCTGCACTGGGATAAC

ATCTTCAAACCTCACCTCATGTGGTCAAAACAGCTACTCAAGGGGAGGTCAATGTGACTGGCGTGATACCA
 CTGACAACAACACCAACAAAATCTTATTTGCAAATCTCAAAGGAACAAGGACCAGAGGGAAACTATGC
 CCGGACTGTCTCAACTGTACAGATCTGGATGTGGCCTTGGGCAGGCCAATGTGTGTGGGGACCACACCTT
 CTGCTAAAGCTTCAATACTCCACGAGGTGACACCTGTTACATCCGGGTGCTTTCTATAATGCACGACAG
 AACAAAAATCAGGCAACTACCCAATCTTCTCAGAGGATATGAAAATATCAGGTTATCAACCCAAAAACGTT
 ATCGATGCAGAAAAAGCACCAGGAGGACCCTACAGACTTGGAACTCAGGATCTTGGCCTAACGCTACC
 AGTAAATCGGATTTTTTGAACAATGGCTTGGGCTGTCCCAAAGGACAACCTACAAAAATGCAACGAAC
 CCACTAACAGTAGAAGTACCATACATTTGTACAGAAGGGGAAGACCAAATTACTGTTGGGGGTTCCATT
 CAGATAACAAAACCCAAATGAAGAGCCTCTATGGAGACTCAAATCCTCAAAAAGTTACCTCATCTGCTAA
 TGGAGTAACCACACATTATGTTTCTCAGATTGGCGACTTCCAGATCAAAACAGAAGACGGAGGACTACCA
 CAAAGCGGCAGAAATTGTTGTTGATTACATGATGCAAAAACTGGGAAAAACAGGAACAATTGTCTATCAA
 AGAGGTGTTTGTGGCTCAAAGGTGTGGTGCAGGAGTGGCAGGAGCAAAGTAATAAAAGGGTCATTG
 CCTTAATTGGTGAAGCAGATTGCCTTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACA
 CAGGAGAACATGCAAAAGCCATAGGAAATTGCCCAATATGGGTAAAAACACCTTTGAAGCTTGGCAATG
 GAACCAATATAGACCTCCTGCAAACTATTGAAGGAAAGGGGTTCTTCGGAGCTATTGCTGGTTTCTT
 AGAAGGAGGATGGGAAGGAATGATTGCAGGTTGGCAGGATACACATCTCACGGAGCACATGGAGTGGC
 AGTGGCGGCAGACCTTAAGAGTACACAAGAAGCTATAAATAAGATAACAAAAAATCTCAATTCTTTGAG
 TGAGCTAGAAGTAAAGAACCTTCAAAGACTAAGTGGTGCCATGGATGAACTCCACAACGAAATACTCGA
 GCTGGATGAGAAAAGTGGATGATCTCAGAGCTGACACTATAAGCTCACAATAGAACTTGCAGTCTTGCTT
 TCCAACGAAGGAATAATAAACAGTGAAGACGAGCATCTATTGGCACTTGAGAGAAAACCTAAAGAAAATG
 CTGGGTCCCTCTGCTGTAGACATAGGAAACGGATGCTTCGAAACCAACACAAATGCAACCAGACCTGC
 TTAGACAGGATAGCTGCTGGCACCTTTAATGCAGGAGAATTTTCTCTCCCACTTTTGATTCACTGAACAT
 TACTGCTGCATCTTTAAATGATGATGGATTGGATAACCATACTATACTGCTCTATTACTCAACTGCTGCTT
 CTAGTTTGGCTGTAACATTAATGCTAGCTATTTTTATTGTTTATATGGTCTCCAGAGACAACGTTTCATGCT
 CCATCTGTCTATAAAGGCCTATTTCTTTAGTTTGAATTTACTGTTATTCCGGTGTGCATTTCTATGTTTGGT
 GAGCGGTTTTCTGTGCTCAGAGTGTGTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGG
 TCGTCCCTTCAGCAAGGACACAAAAAGATTTTAAATTTATTAATAAAAAAAAAAAAAAAAAAGACCGGAAT
 TCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTCTTAAGATTGAATCCTG
 TTGCCGCTTTCGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAA
 TGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGA
 AAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGGGTGCATCTATGTTACTAGAT

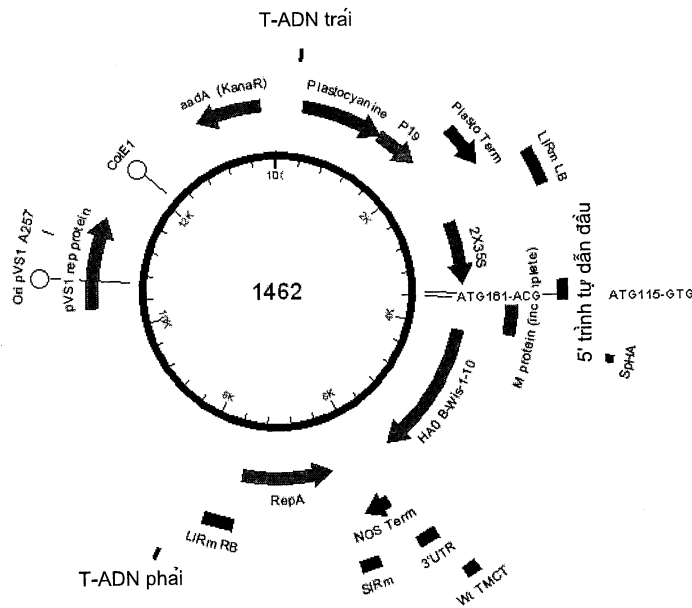
Figure 26G, SEQ ID NO: 54

Trình tự axit amin của HA từ cúm B/Wisconsin/1/2010

MKAIIVLLMVVTSNADRICTGITSSNSPHVVKATQGEVNVTVIPLTTPTKSYFANLKGTRTRGKLCPLDCLN
 CTDLDVALGRPMC VGTTPSAKASILHEVRPVTSGCFPIIMHDKIRQLPNLLRGYENIRLSTQNVIDAEKAPGG
 PYRLGTSGSCP NATSKIGFFATMAWAVPKDNYKNATNPLTVEVPYICTEGEDQITVWGFHSDNKTQMKSLYG
 DSNPQKFTSSANGVTTHYVSQIGDFPDQTEDGGLPQSGRIVVDYMMQKPGKTGTIVYQRGVLLPQK VWCASG
 RSKVIKGSPLIGEADCLHEKYGGLNKS KPYTGEHAKAIGNCPIWVKTPKLKLANGTKYRPPAKLLKERGFFG
 AIAGFLEGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLSLSELEVKNLQRLSGAMDELHN
 EILELDEKVDDL RADTISSQIELAVLLSNEGINSEDEHLLALERKLLKMLGPSAVDIGNGCFETKHKCNQTCLD
 RIAAGTFNAGEFSLPTFDSL NITAASLNDDGLDNHTILLYSTAASSLAVTLM LAIFIVYMVSRDNVSCSICL*

Fig.26H

Bản đồ của cấu trúc số 1462



C-2X35S/CPMV-HT/HA B Wisconsin có vòng thủy phân protein bị mất/NOS vào hệ khuếch
daaij BeYDV(m)+ Replicaza (Cấu rúc số 1467)

Fig.27A, SEQ ID NO: 55

HAB110(PrL-).r

TCCTTCCCATCCTCCACCAGGAGGTCTATATTTGGTTCCATTGGCAAGCTTCAAAG

Fig.27B, SEQ ID NO:56

HAB110(PrL-).c

ATATAGACCTCCTGGTGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGA

Fig.27C, SEQ ID NO: 57

Bảng biểu hiện số 1467 từ vùng khởi đầu đến vùng kết thúc. HA từ cúm B/Wisconsin/1/2010 có vòng
thủy phân protein bị mất được gạch chân

GTCAACATGGTGGAGCAGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAA
AGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGAAACCTCCTCGGATTCCATTGCCAGCTATCT
GTCACTTTATTGTGAAGATAGTGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAA

AGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCG
 TGGAAAAAGAAGACGTTCCAACCACGCTCTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACG
 ACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCA
 ACAAAGGGTAATATCCGGAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATA
 GTGGAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCC
 TCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCA
 ACCACGCTCTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACT
 ATCCTTCGCAAGACCCTTCCTCTATATAAGGAAGTTCATTTCATTGGAGAGGTATTAAAACTCTTAATAGG
 TTTTGATAAAAGCGAACGTGGGGAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAG
 CAAACTCTCTCTTGTCTTTCTTGCCTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACA
 ACGTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAAGTGGCGGCCATTAAATAACG
 TGTACTTGTCTATTCTTGTGCGGTGTGGTCTTGGGAAAAGAAAGCTTGTGGAGGCTGCTGTTACGCCCCA
 TACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTGACGAGGTATTG
 TTGCTGTACTTCTTCTTCTTCTTGTGATTGGTCTATAAGAAATCTAGTATTTTCTTGTAAACAGA
 GTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTGGGC
 CCATGAAGGCAATAATTGTAATACTACTCATGGTATTGAACATCCAATGCAGATCGAATCTGCACTGGGATAAC
 ATCTTCAAACCTCACTCATGTGGTCAAAAACAGCTACTCAAGGGGAGGTCAATGTGACTGGCGTGATACCA
 CTGACAACAACACCAACAAAAATCTTATTTTGCAAATCTCAAAGGAACAAGGACCAGAGGGGAACTATGC
 CCGGACTGTCTCAACTGTACAGATCTGGATGTGGCTTGGGCAGGCCAATGTGTGTGGGGACCACACCTT
 CTGCTAAAGCTTCAATACTCCACGAGGTGACAGCTGTTACATCCGGGTGCTTTCCTATAATGCACGACAG
 AACAAAAATCAGGCAACTACCAATCTTCTCAGAGGATATGAAAATATCAGGTTATCAACCCAAAACGTT
 ATCGATGCAGAAAAAGCACCAGGAGGACCTACAGACTTGGAACTCAGGATCTTGGCCTAACGCTACC
 AGTAAAAATCGGATTTTTTGCAACAATGGCTTGGGCTGTCCCAAAGGACAACACTACAAAAATGCAACGAAC
 CCACTAACAGTAGAAGTACCATACATTTGTACAGAAGGGGAAGACCAAATTAAGTTGGGGGTTCATT
 CAGATAACAAACCCAAATGAAGAGCCTCTATGGAGACTCAAATCCTCAAAAGTTACCTCATCTGTAA
 TGGAGTAACCACACATTATGTTTCTCAGATTGGCGACTTCCCAGATCAAAACAGAAGACGGAGGACTACCA
 CAAAGCGGCAGAAATTGTTGTTGATTACATGATGCAAAAACCTGGGAAAACAGGAACAATTGTCTATCAA
 AGAGGTGTTTTGTTGCCTCAAAAGGTGGTGGCGAGTGGCAGGAGCAAAAGTAATAAAAGGGTCATTG
 CCTTTAATTGGTGAAGCAGATTGCCTTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACA
 CAGGAGAACATGCAAAAGCCATAGGAAATGGCCAATATGGGTAAAAACACCTTTGAAGCTTGCCAATG
 GAACCAAAATATAGACCTCCTGGTGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGATACACATCTC
 ACGGAGCATATGGAGTGGCAGTGGCGGCAGACCTTAAGAGTACACAAGAAGCTATAAATAAGATAACA
 AAAAAATCTCAATTCTTTGAGTGAGCTAGAAGTAAAGAACCTTCAAAGACTAAGTGGTGCCATGGATGAA
 CTCCACAACGAAATACTCGAGCTGGATGAGAAAGTGGATGATCTCAGAGCTGACACTATAAGCTCACA
 ATAGAACTTGAGTCTTGCTTTCCAACGAAGGAATAATAAACAGTGAAGACGAGCATCTATTGGCACTTG
 AGAGAAAACTAAAGAAAATGCTGGGTCCCTCTGCTGTAGACATAGGAAACGGATGCTTCGAAACCAAC
 ACAATGCAACCAGACCTGCTTAGACAGGATAGCTGCTGGCACCTTTAATGCAGGAGAATTTCTCTCCC
 CACTTTTGATTCAATTGAACATTACTGCTGCATCTTAAATGATGATGGATTGGATAACCATACTATACTGC
 TCTATTACTCAACTGCTGCTTCTAGTTTGGCTGTAACATTAATGCTAGCTATTTTATTGTTTATATGGTCT
 CCAGAGACAACGTTTCATGCTCCATCTGTCTATAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTG
 GTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTATGTAATTTAATTTCTT
 TGTGAGCTCCTGTTTAGCAGGTCGTCCCTCAGCAAGGACACAAAAAGATTTTAATTTTATTAACAAAAA
 AAAAAAAGACCGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAA
 AGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGGCATGATTATCATATAATTTCTGTTGAATTACGTTAA
 GCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAAT
 TATACATTTAATACGCGATAGAAAAACAAATATAGCGCGCAAACCTAGGATAAATTATCGCGCGCGGTGT
 CATCTATGTTACTAGAT

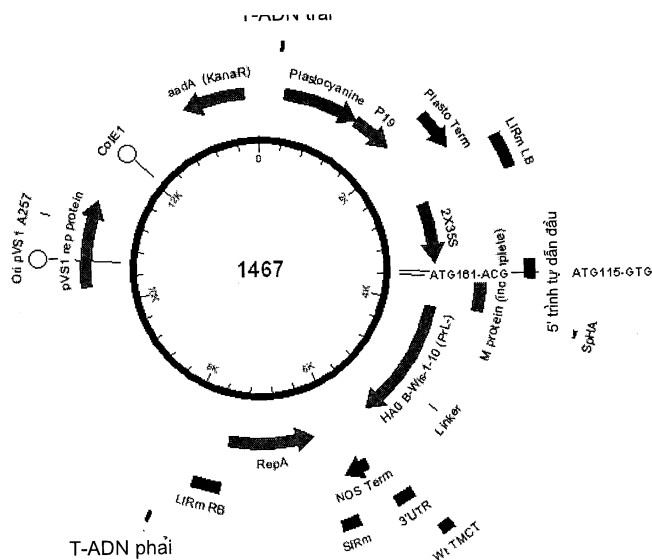
Fig.27D, SEQ ID NO: 58

Trình tự axit amin của HA từ cúm B/Wisconsin/1/2010 có vòng thủy phân protein bị mất

MKAIIVLLMVVTSNADRICTGITSSNSPHVVKATATQGEVNVTVGVPLTTTPTKSYFANLKG
 TRTRGKLCPCDCLNCTDLDDVALGRPMC VGTTPSAKASILHEVRPVTSGCFPIMHDRTKIRQ
 LPNLLRGYENIRLSTQNVIDAEKAPGGPYRLGTSGSPNATSKIGFFATMAWAVPKDNY
 KNATNPLTVEVPYICTEGEDQITVWGFHSDNKTQMKSLYGDSNPQKFTSSANGVTTHYV
 SQIGDFPDQTEDGGLPQSGRIVVDYMMQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIK
 GSLPLIGEADCLHEKYGGLNKSHPYYTGEHAKAIGNCPIWVKTPKLKLANGTKYRPPGGG
 WEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLNSLSELEVKNLQRLSGAMD
 ELHNEILELDEKVDLDRADTISSQIELAVLLSNEGIINSEDEHLLALERKLKMLGPSAVDI
 GNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDSLNTAASLNDGDLNHTILLYYST
 AASSLAVTLMIAIFIVYMVSRDNVSCSICL*

Fig.27E

Bản đồ của cấu trúc số 1467



D-2X35S/CPMV-HT/PDISP/HA B Malaysia/NOS vào hệ khuếch đại (Cấu trúc số 1631)

Fig.28A, SEQ ID NO:59

IF-HB-M-04.s2+4c

TCTCAGATCTTCGCCGATCGAATCTGCACTGGGATAACATCGTC

Figure 28B, SEQ ID NO: 60

IF-HB-M-04.s1-4r

ACTAAAGAAAATAGGCCTTTATAGACAGATGGAGCAAGAAACATTG

Fig.28C, SEQ ID NO:61

HA B Malaysia được tổng hợp (tương ứng với nt 31-1743 từ số truy cập Genbank EU124275). Các đột biến T759C và C888G được gạch chân.

GATCGAATCTGCACTGGGATAACATCGTCAAACCTCACCACATGTTGTCAAACTGCTACTCAAGGGGAGG
TCAATGTGACTGGTGTAATACCACTGACAACAACACCCACCAAATCTCATTTTGCAAATCTCAAAGGAAC
AGAAACCAGAGGGGAAACTATGCCCAAAATGCCTCAACTGCACAGATCTGGACGTGGCCTTGGGCAGACC
AAAATGCACGGGGAAACATACCCTCGGCAAGAGTTTCAATACTCCATGAAGTCAGACCTGTTACATCTGGG
TGCTTTCCTATAATGCACGACAGAACAAAAATTAGACAGCTGCCTAACTTCTCAGAGGATACGAACATA
TCAGGTTATCAACTCATAACGTTATCAATGCAGAAAATGCACCAGGAGGACCCTACAAAATTGGAAACCTC
AGGGTCTTGGCCTAACGTTACCAATGGAAACGGATTTTTCGCAACAATGGCTTGGGCCGTCCCCAAAAAC
GACAACAACAAAACAGCAACAAATTCATTAACAATAGAAAGTACCATACATTTGTACAGAAGGAGAAGAC
CAAATTACCGTTTGGGGGTTCCACTCTGATAACGAAACCCAAATGGCAAAGCTCTATGGGGACTCAAAGC
CCCAGAAGTTCACCTCATCTGCCAACGGAGTGACCACACATTACGTTTCACAGATTGGTGGCTTCCCAAA
TCAAACAGAAGACGGAGGACTACCACAAAGCGGTAGAATTGTTGTTGATTACATGGTGCAAAAATCTGG
GAAAACAGGAACAATTACCTATCAAAGAGGTATTTTATTGCCTCAAAAAGTGTGGTGCGCAAGTGCGCAG
GAGCAAGGTAATAAAAGGATCGTTGCCTTTAATTGGAGAAGCAGATTGCCTCCACGAAAAATACGGTGG
ATTAACAAAAGCAAGCCTTACTACACAGGGGAACATGCAAAGGCCATAGGAAATTGCCCAATATGGGT
GAAAACACCCTTGAAGCTGGCCAATGGAACCAAAATATAGACCTCCTGCAAAACTATTAAAGGAAAGGGG
TTTCTTCGGAGCTATTGCTGGTTTCTTAGAAGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGATAC
ACATCCCATGGGGCACATGGAGTAGCGGTGGCAGCAGACCTTAAGAGCACTCAAGAGGCCATAAACAAG
ATAACAAAAAATCTCAACTCTTTGAGTGAGCTGGAAGTAAAGAATCTTCAAAGACTAAGCGGTGCCATG
GATGAACTCCACAACGAAATACTAGAACTAGACGAGAAAGTGGATGATCTCAGAGCTGATACAATAAGC
TCACAAATAGAACTCGCAGTCCTGCTTTCCAATGAAGGAATAATAAACAGTGAAGATGAGCATCTCTTG
CGCTTGAAAGAAAGCTGAAGAAAATGCTGGGGCCCTCTGCTGTAGAGATAGGGAATGGATGCTTTGAAA
CCAAACACAAGTGCAACCAGACCTGTCTCGACAGAATAGCTGCTGGTACCTTTGATGCAGGAGAATTTTC
TCTCCCCACTTTTGATTCACTGAATATTACTGCTGCATCTTTAAATGACGATGGATTGGATAATCATACTA
TACTGCTTTACTACTCAACTGCTGCCTCCAGTTTGGCTGTAACATTGATGATAGCTATCTTTGTTGTTTATA
TGGTCTCCAGAGACAATGTTTCTTGCTCCATCTGTCTATAA

Fig.28D

Bản đồ của cấu trúc 194. Các vị trí enzyme giới hạn SacII và StuI được sử dụng để kéo dài plasmit được giảm độc lực trên bản đồ này

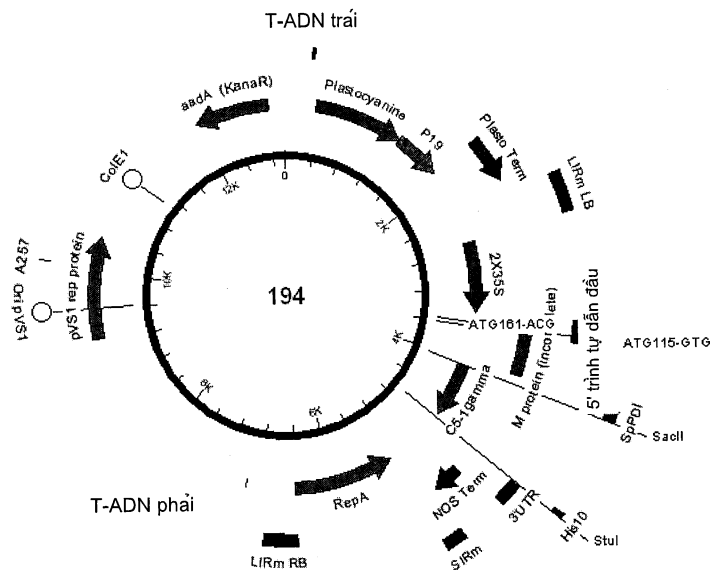


Fig.28E, SEQ ID NO: 62

Cấu trúc 194 từ các bờ ADN vận chuyển từ trái sang phải (được gạch chân). 2X35S/CPMV-HT/NOS vào hệ khuếch đại BeYDV(m)+Replicaza có bằng biểu hiện chất ức chế bất hoạt Plastoxyanin-P19-Plastoxyanin

TGGCAGGATATATTGTGGTGTAAACAATTGACGCTTAGACAACCTTAATAACACATTGCGGACGTTTTTA
 ATGTACTGAATTAACGCCGAATCCGGGCTGGTATATTTATATGTTGTCAAATAACTCAAAAACCATAAA
 AGTTTAAAGTTAGCAAGTGTGTACATTTTTACTTGAACAAAAATATTCACCTACTACTGTTATAAATCATT
 TTAAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTGATATTTTGACAACAATTTTGTG
 CAACATTGAGAAAAATTTGTTGTTCTCTTTTCATTGGTCAAAAACAATAGAGAGAGAAAAAGGAAGA
 GGGAGAATAAAAAACATAATGTGAGTATGAGAGAGAAAAGTTGTACAAAAGTTGTACCAAAATAGTTGTAC
 AAATATCATTGAGGAATTTGACAAAAGCTACACAAATAAGGGTTAATTGCTGTAAATAAATAAGGATGA
 CGCATTAGAGAGATGTACCATTAGAGAATTTTTGGCAAGTCATTAAGAAAGAAAGAAATAAATTATTTTAA
 AATTAAGTTGAGTCATTGATTAAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGT
 TGTATTAGTAATTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCTATATATTGCCCATAGA
 GTCAGTTAACTCATTTTTATATTTTCATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAA
 AAAAAACGGTATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTAGGAGGATA
 ACATCCAATCCAACCAATCACACAATCCTGATGAGATAACCCACTTTAAGCCCACGCATCTGTGGCACA
 TCTACATTATCTAAATCACACATTCTCCACACATCTGAGCCACACAAAAACCAATCCACATCTTTATCAC
 CCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAG
 ACTAATTAATTAATTAATCATCTTGAGAGAAAATGGAACGAGCTATACAAGGAAACGACGCTAGGGAAC
 AAGCTAACAGTGAACGTTGGGATGGAGGATCAGGAGGTACCACTTCTCCCTTCAAACCTTCTGACGAAAG
 TCCGAGTTGGACTGAGTGGCGGTACATAACGATGAGACGAATTCGAATCAAGATAATCCCTTGGTTTC
 AAGGAAAGCTGGGTTTCGGGAAAAGTTGTATTTAAGAGATATCTCAGATACGACAGGACGGAAGCTTCA
 CTGCACAGAGTCCTTGATCTTGACGGGAGATTTCGGTTAACTATGCAGCATCTCGATTTTTCGGTTTCGA
 CCAGATCGGATGTACCTATAGTATTCGGTTTCGAGGAGTTAGTATCACCGTTTCTGGAGGGTCGCGAACT
 CTTACAGCATCTCTGTGAGATGGCAATTCGGTCTAAGCAAGAACTGCTACAGCTTGCCCCAATCGAAGTGG
 AAAGTAATGTATCAAGAGGATGCCCTGAAGGTACTCAAACCTTCGAAAAAGAAAGCGAGTAAGTTAAAA
 TGCTTCTTCGTCCTCTATTATAATATGGTTTGTATTGTAAATTTGTTCTTGTAGAAGAGCTTAATTAAT
 CGTTGTTGTTATGAAATACTATTTGTATGAGATGAACTGGTGTAATGTAATTCATTTACATAAGTGGAGTC

AGAATCAGAATGTTTCCCTCCATACTAATACTAGACATGAAGACCTGCCGCGTACAATTGTCTTATATTGGA
ACAACCTAAAATTGAACATCTTTTGCCACAACCTTTATAAGTGGTTAATATAGCTCAAATATATGGTCAAGTT
CAATAGATTAATAATGGAAATATCAGTTATCGAAATTCATTAACAATCAACTTAACGTTATTAACCTACTA
ATTTTATATCATCCCCCTTTGATAAATGATAGTACACCAATTAGGAAGGAGCATGCTCGCCTAGGAGATTG
TCGTTTCCCGCCTTCAGTTTGCAAGCTGCTCTAGCCGTGTAGCCAATACGCAAACCGCTCTCCCCGCGCG
TTGGGAATTACTAGCGCGTGTGAGACGCGTGTGTGTGTGACTCCGAGGGGTGCTCAAACCTCTATCTT
ATAACCGCGCTGGAGGCATGGAGGCAGGGGTATTTTGGTCATTTTAATAGATAGTGGAAAAATGACGTGG
AATTTACTTAAAGACGAAGTCTTTGCGACAAGGGGGGGGCCACGCCGAATTTAATATTACCGGCGTGGCC
CCCCCTTATCGCGAGTGCTTTAGCACGAGCGGTCCAGATTTAAAGTAGAAAATTTCCCGCCCACTAGGGT
TAAAGGTGTTCACTATAAAAAGCATATACGATGTGATGGTATTTGGTCGACAAGCTTGCATGCCGGTCA
ACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAAGGG
CAATTGAGACTTTTCAACAAAAGGGTAATATCCGGAAACCTCTCGGATTCCATTGCCCAGCTATCTGTCA
CTTTATTGTGAAGATGTGAAAAAGGAAGTGGTCTCTACAAATGCCATCATTGCGATAAAAGGAAAGGC
CATCGTTGAAGATGCCTTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGA
AAAAGAAGACGTTCCAACCAAGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACGACAC
ACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAAGGGCAATTGAGACTTTTCAACAA
AGGGTAATATCCGGAAACCTCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGG
AAAAGGAAGGTGGCTCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTG
CCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGA AAAAAGAACGTTCCAACCA
CGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCC
TTCGCAAGACCTTCTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTA AAAATCTTAATAGGTTTT
GATAAAAGCGAACGTGGGGGAAAACCCGAACCAAACTTCTTCTAAACTCTCTCTCATCTCTCTTAAGCAA
ACTTCTCTCTTGTCTTTCTTTCGCGTGAGCGATCTTCAACGTTGTGAGATCGTGTCTCGGCACGACCAACG
TTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTGGACAGTAGTGC GCGCCATTAAATAACGTGT
ACTTGTCTATTCTTTCGCTGGTCTTGGGAAAAAGAAAGCTTGTGGAGGCTGCTGTTTACGCCCCATAC
ATTACTTGTATACGATTCTGCTGACTTTTCGGCGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTTG
CTCTGACTTCTTTCTTCTTCTTCTTCTGCTGATTGGTTCTATAAGAAATCTAGATTTTTCTTTGAAACAGAGTT
TTCCCGTGGTTTTTCGAACCTGGAGAAAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCCCA
TGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTTCTTCTGTGTTGGTTCTTCTCAGATCTTCGCCG
CGGCTCCTCAGCCAAAACGACACCCCCATCTGTCTATCCACTGGCCCCCTGGATCTGCTGCCCAAACCTAAC
TCCATGGTGACCCTGGGATGCCTGGTCAAGGGCTATTTCCCTGAGCCAGTGACAGTGACCTGGAACCTCTG
GATCCCTGTCCAGCGGTGTGCACACCTTCCCAGCTGTCTGCACTGACCTCTACACTCTGAGCAGCTCA
GTGACTGTCCCTCCAGCACCTGGCCCAGCGAGACCGTCACCTGCAACGTTGCCACCCCGGCCAGCAGCA
CCAAGGTGGACAAGAAAAATTGTGCCCAGGGATTGTGGTTGTAAGCCTTGCAATGTACAGTCCCAGAAGT
ATCATCTGTCTTCATCTTCCCCCAAAGGCCAAGGATGTGCTCACCATTACTCTGACTCCTAAGGTACAGT
GTGTTGTGGTAGACATCAGCAAGGATGATCCCGAGGTCCAGTTTCACTGAGTGTGATAGTGTGGAGGT
GCACACAGCTCAGACGCAACCCCGGAGGAGCAGTTCAACAGCACTTTCCGCTCAGTCAAGTGAACCTCCC
ATCATGCACCAGGACTGGCTCAATGGCAAGGAGCGATCGCTCACCATCACCATCACCATCACCATCACC
TTAAAGGCTCATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTT
CTGTGCTCAGAGTGTGTTTATTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCTGCCCTTC
AGCAAGGACACAAAAAGATTTTAATTTTATTA AAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAA
GCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCT
TGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGT
TATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACCGGATAGAAAAACAAAT
ATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCTCTAGAGTCTCAAG
CTTGGCGCGCCATAAAATGATTATTTTATGAATATATTTCAATTGTGCAAGTAGATAGAAATTACATATGTT
ACATAACACACGAAATAAACAAAAAAGACAATCCAAAAACAAACACCCCAAAAAAATCACTTT
AGATAAACTCGTATGAGGAGAGGCACGTTTCACTGACTCGACGATCCCCGAGCAAAAAAAGTCTCCCCGT
CACACATATAGTGGGTGACGCAATTTATCTTTAAAGTAATCCTTCTGTGACTGTGATGATAACATCCAG
TCTTCGTGAGGATTGCAAGAATTATAGAAGGATCCCACTTTTATTTTCTCTTTTTTCCATATTTAGGG
TTGACAGTGAATCAGACTGGCAACCTATTAATGTCTTCCACAATGGGACGAACTTGAAGGGGATGTGCT
CGATGATATTATAGTGGCGGTGTTTCATCGTAGTTGGTGAAATCGATGGTACCGTTCCAATAGTTGTGCTGT
CCGAGACTTCTAGCCAGGTGGTCTTTCCGGTACGAGTTGGTCCGAGATGTAGAGGCTGGGGGTGTCGGA
TTCCATTCTTCCATTGTCTGGTTAAATCGGCCATCCATTCAAGGTGAGATTGAGCTTGTGTTGATGAGA

CAGGATGTATGTAAGTATAAGCGTCTATGCTTACATGGTATAGATGGGTTTCCCTCCAGGAGTGTAGATC
 TTCGTGGCAGCGAAGATCTGATTCTGTGAAGGGCGACACATACGGTTCAGGTTGTGGAGGGAATAATTTG
 TTGGCTGAATATTTCCAGCCATTGAAGTTTTGTTGCCCATTCATGAGGGAATTCTTCCTTGATCATGTCAAG
 ATATTCTCCTTAGACGTTGCAGTCTGGATAATAGTTCTCCATCGTGCCTCAGATTTGCGAGGAGAGACCT
 TATGATCTCGGAAATCTCCTCTGGTTTTAATATCTCCGCTCTTTGATATGTAATCAAGGACTTGTTTAGAG
 TTTCTAGCTGGCTGGATATTAGGGTGATTTCTTCAAAATCGAAAAAGAAGGATCCCTAATACAAGGTT
 TTTTATCAAGCTGGAGAAGAGCATGATAGTGGGTAGTGCCATCTTGATGAAGCTCAGAAGCAACACCAA
 GGAAGAAAATAAGAAAAGGTGTGAGTTTCTCCAGAGAACTGGAATAAATCATCTCTTTGAGATGAGC
 ACTTGGGATAGGTAAGGAAAACATATTTAGATTGGAGTCTGAAGTTCTTACTAGCAGAAGGCATGTTGTT
 GTGACTCCGAGGGGTTGCCTCAAACCTCTATCTTATAACCGCGTGGAGGCATGGAGGCAGGGGTATTTTG
 GTCATTTTAATAGATAGTGGAATAAGCTGGAATTTACTTAAAGACGAAGTCTTTGCGACAAGGGGGG
 GCCACGCGCAATTTAATATTACCGCGTGGCCCCCTTATCGCGAGTGCTTTAGCACGAGCGGTCCAG
 ATTTAAAGTAGAAAAATTTCCCGCCCACTAGGGTTAAAGGTGTTTCACTATAAAAGCATATACGATGTGA
 TGGTATTTGACTAGTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCAA
 CTTAATCGCCTTGCAGCACATCCCCCTTTCCGCGAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCC
 CTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGCTAGAGCAGCTTGAGCTTGGATCAGATTGTGCTTTC
 CCGCCTTCAGTTTAACTATCAGTGTTTGACAGGATATATTGGCGGGTAAACCTAAGAGAAAAGAGCGTT
TA

Fig. 28F, SEQ ID NO:63

Bảng biểu hiện số 1631 từ vùng khởi đầu 2X35S đến vùng kết thúc NOS.PDISP-HA từ cúm
 B/Malaysia/2506/2004 được gạch chân

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAA
 AGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCAGCTATCT
 GTCACTTTATTGTGAAGATAGTGGAAGGAAGGTGGCTCCTACAAATGCCATCATTGGGATAAAGGAA
 AGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAGATGGACCCCAACCGAGGAGCATCG
 TGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACG
 ACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCA
 ACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATA
 GTGGAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCC
 TCTGCCGACAGTGGTCCCAAGATGGACCCCAACCGAGGAGCATCGTGAAAAAGAAGACGTTCCA
 ACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCACTGACGTAAGGGATGACGCACAATCCCACT
 ATCCTTCGCAAGACCCTTCTCTATATAAGGAAGTTCATTTTATTGGAGAGGTATTAATAATCTTAATAGG
 TTTTGATAAAAGCGAACGTGGGGAACCCGAACCAACCTTCTTCTAACTCTCTCTCATCTCTCTTAAAG
 CAAACTTCTCTTGTCTTTCTTGCGTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAAGTACA
 ACGTTTTCTTCTACTGAAGCGAAATCAAAGATCTCTTGTGGACACGTAAGTGGCGGCCATTAAATAACG
 TGTACTTGTCTATTCTTGTGCGGTGTGGTCTTGGGAAAAGAAAGCTTGCTGGAGGCTGCTGTTACGCCCA
 TACATTACTTGTACGATTCTGTGACTTTCCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTG
 TTGCCTGTACTTCTTCTTCTTCTTCTTGCTGATTGGTTCTATAAGAAATCTAGTATTTCTTTGAAACAGA
 GTTTTCCCGTGGTTTTGCAACTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTGCGGC
 CCATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCG
CCGATCGAATCTGCACTGGGATAACATCGTCAAACTCACCACATGTTGTCAAACTGCTACTCAAGGGGA
GGTCAATGTGACTGGTGTAAATACCACTGACAACAACCCACCAAAATCTCATTTTGCAAATCTCAAAGGA
ACAGAAACCAGAGGGGAACTATGCCAAAAATGCCTCAACTGCACAGATCTGGACGTGGCCTTGGGCAGA
CCAAAAATGCACGGGGAAACATACCCTCGGCAAGAGTTCAATACTCCATGAAGTCAGACCTGTTACATCTG
GGTGCTTTCCTATAATGCACGACAGAACAAAAATTAGACAGCTGCCTAAACTTCTCAGAGGATACGAACA
TATCAGGTTATCAACTCATAACGTTATCAATGCAGAAAAATGCACCAGGAGGACCCTACAAAAATTGGAACC
TCAGGGTCTTGCCCTAACGTTACCAATGGAAACGGATTTTCGCAACAATGGCTTGGGCCGTCCCAAAAA
ACGACAACAACAAAACAGCAACAAATTCATTAACAATAGAAGTACCATACATTTGTACAGAAGGAGAAG
ACCAAAATTACCGTTTGGGGGTTCCTACTGTATAACGAAACCCAAATGGCAAAGCTCTATGGGGACTCAAA

GCCCCAGAAGTTCACCTCATCTGCCAACGGAGTGACCACACATTACGTTTCACAGATTGGTGGCTTCCCA
AATCAAAACAGAAAGACGGAGGACTACCACAAAGCGGTAGAATTGTTGTTGATTACATGGTGCAAAAATCT
GGGAAAACAGGAACAATTACCTATCAAAGAGGTATTTATTGCCTCAAAAAGTGTGGTGCGCAAGTGGC
AGGAGCAAGGTAATAAAAGGATCGTTGCCTTTAATTGGAGAAGCAGATTGCCTCCACGAAAAATACGGT
GGATTAAACAAAAGCAAGCCTTACTACACAGGGGAACATGCAAAGGCCATAGGAAATTGCCCAATATGG
GTGAAAACACCCTTGAAGCTGGCCAATGGAACCAAATATAGACCTCCTGCAAAACTATTAAAGGAAAGG
GGTTTCTTCGGAGCTATTGCTGGTTTCTTAGAAGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGAT
ACACATCCCATGGGGCACATGGAGTAGCGGTGGCAGCAGACCTTAAGAGCACTCAAGAGGCCATAAACA
AGATAACAAAAAATCTCAACTCTTTGAGTGAGCTGGAAGTAAAGAATCTTCAAAGACTAAGCGGTGCCA
TGGATGAACTCCACAACGAAATACTAGAACTAGACGAGAAAGTGGATGATCTCAGAGCTGATACAATAA
GCTCACAAATAGAACTCGCAGTCCTGCTTTCCAATGAAGGAATAATAAACAGTGAAGATGAGCATCTCTT
GGCGCTTGAAAGAAAGCTGAAGAAAATGCTGGGCCCCCTGCTGTAGAGATAGGGAATGGATGCTTTGA
AACCAAAACACAAGTGCAACCAGACCTGTCTCGACAGAATAGCTGCTGGTACCTTTGATGCAGGAGAATT
TCTCTCCCCACTTTTGATTCACTGAATATTACTGCTGCATCTTTAAATGACGATGGATTGGATAATCATA
TATACTGCTTTACTACTCAACTGCTGCCTCCAGTTTGGCTGTAAACATTGATGATAGCTATCTTTGTTGTTA
TATGGTCTCCAGAGACAATGTTTCTTGCTCCACTGTCTATAAAGGCCTATTTTCTTTAGTTTGAATTTACT
GTTATTCCGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTT
AATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCTTCAGCAAGGACACAAAAAGATTTTAAATTTATTA
AAAAAAAAAAAAAAAAAAGACCGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTCAAACATTTG
GCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATT
ACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTC
CCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGC
GCGGTGTCATCTATGTTACTAGAT

Fig.28G, SEQ ID NO: 64

Trình tự axit amin của PDISP-HA từ cúm B/Malaysia/2506/2004

MAKNVAFGLFLSLLVLVPSQIFADRICTGITSSNSPHVVKATQGEVNVTVIPLTTTPTKSHFANLKGTETRG
 KLCPKCLNCTDLVALGRPKCTGNIPSARVSILHEVRPVTSGCFPIIMHDRTKIRQLPKLLRGYEHIRLSTHNVIN
 AENAPGGPYKIGTSGSCPNVTNGNGFFATMAWAVPKNDNNKTATNSLTIEVPYICTEGEDQITVWGFHSDNET
 QMAKLYGDSKPQKFTSSANGVTTHYVSQIGGFPNQTEDGGLPQSGRIVVDYMVQKSGKTGTITYQRGILLPQ
 KVWCASGRSKVIKGSPLIGEADCLHEKYGGLNKSPPYYTGEHAKAIGNCPIWVKPLKLANGTKYRPPAKL
 LKERGFAGIAAGFLEGGWEGMIAGWHGYTSHGAHGVAADLKSTQEAINKITKNLSLSELEVKNLQRLSG
 AMDELHNEILELDEKVDDLADTISSQIELAVLLSNEGINSEDEHLLALERKLLKMLGPSAVEIGNGCFETKH
 KCNQTCLDRIAAGTFDAGEFSLPTFDSL NITAASLNDDGLDNHTILLYSTAASSLAVTLMAIFVVYMVSRDN
 VSCSICL*

Fig.28H

Bản đồ của cấu trúc số 1631

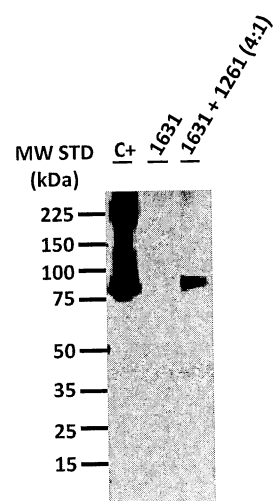
**Fig.29**

Fig.30A

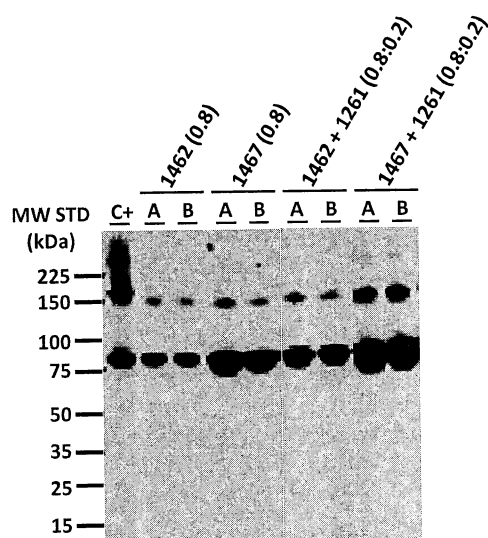


Fig.30B

Cấu trúc số	Khả năng hemagglutin hóa (độ chuẩn HA) ^a
1462 (0.8)	136 ± 48
1467 (0.8)	2660 ± 384
1462 + 1261 (0.8:0.2)	941 ± 192
1467 + 1261 (0.8:0.2)	4344 ± 1536

^a Nghịch đảo của lần pha loãng cao nhất của ngưỡng kết tủa các tế bào hồng cầu

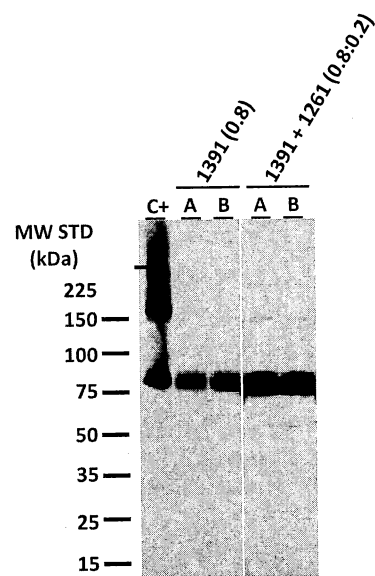


Fig.31