



(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)



1-0031114

(51)⁷ C12N 7/04; C07K 14/085; C07K 16/10; (13) B
C12N 15/82; C12N 7/01; A01H 5/00;
C12N 15/41

(21) 1-2015-00934

(22) 29/08/2013

(86) PCT/CA2013/050666 29/08/2013

(87) WO2014/036645 13/03/2014

(30) 61/697,266 05/09/2012 US

(45) 25/02/2022 407

(43) 25/06/2015 327A

(73) MEDICAGO INC. (CA)

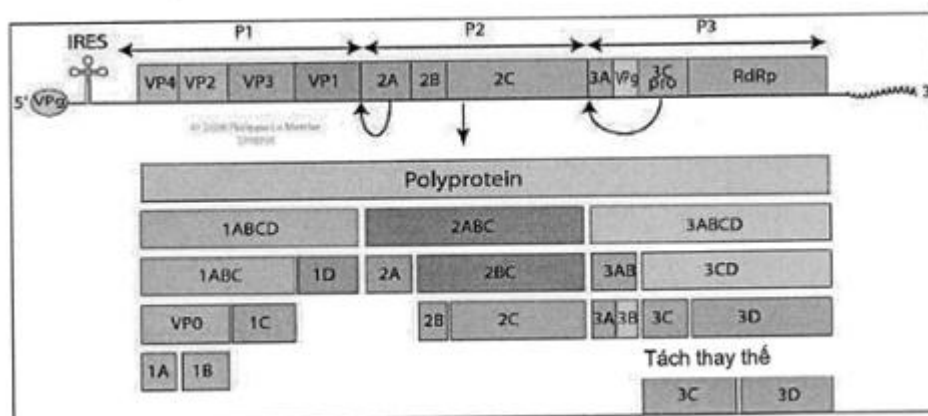
1020 route de l'Eglise, Suite 600, Quebec, Quebec G1V 3V9, Canada

(72) D'AOUST, Marc-Andre (CA); LAVOIE, Pierre-Olivier (CA); COUTURE, Manon (CA); POULIN, Lucie (CA); VEZINA, Louis-Philippe (CA).

(74) Công ty Luật TNHH T&G (TGVN)

(54) PHƯƠNG PHÁP TẠO RA HẠT TƯƠNG TỰ ENTEROVIRUT, DỊCH CHIẾT THỰC VẬT CHỨA HẠT TƯƠNG TỰ ENTEROVIRUT VÀ HẠT TƯƠNG TỰ ENTEROVIRUT THU ĐƯỢC TỪ DỊCH CHIẾT NÀY

(57) Sáng chế đề cập đến phương pháp tạo ra hạt tương tự picomavirut (PVLP) ở thực vật. Phương pháp này bao gồm việc đưa axit nucleic thứ nhất và axit nucleic thứ hai vào thực vật, bộ phận của thực vật hoặc tế bào thực vật. Axit nucleic thứ nhất bao gồm vùng điều hoà thứ nhất hoạt động ở thực vật được liên kết theo cách hoạt động được với trình tự nucleotit mã hoá polyprotein. Axit nucleic thứ hai bao gồm vùng điều hoà thứ hai hoạt động ở thực vật và liên kết theo cách hoạt động được với trình tự nucleotit mã hoá một hoặc nhiều proteaza. Thực vật, bộ phận của thực vật, hoặc tế bào thực vật được ủ trong các điều kiện mà cho phép biểu hiện axit nucleic, nhờ đó tạo ra PVLP. Sáng chế cũng đề cập đến PVLP bao gồm polyprotein. Sáng chế cũng đề cập đến chế phẩm chứa PVLP và kháng thể đa dòng được tạo ra bằng cách sử dụng PVLP này.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp tạo ra protein cấu trúc picornavirut ở thực vật. Cụ thể hơn là, sáng chế cũng đề cập đến phương pháp tạo ra hạt tương tự virut bao gồm protein cấu trúc picornavirut ở thực vật.

Tình trạng kỹ thuật của sáng chế

Picornavirut là ARN virut sợi dương không bọc nhỏ mà có thể gây ra một khoảng rộng các biểu hiện lâm sàng ở con người và động vật. Dựa trên nhiều đặc tính bao gồm tính tương đồng trình tự và tính nhạy axit. Picornavirut được tách thành nhiều chi trong số chúng là nhiều mầm bệnh quan trọng của con người và động vật.

Picornavirut có nucleocapsit trần. Capsit là sự sắp xếp của 60 đoạn khởi đầu trong cấu trúc hai mươi mặt được đóng gói mặt cách chặt chẽ. Mỗi đoạn khởi đầu gồm có 4 polypeptit đã được biết đến như VP (protein virut) 1, 2, 3 và 4. VP2 và VP4 polypeptit có nguồn gốc từ một tiền chất đã được biết đến như VP0, mà được tách sau khi tiếp thu ARN hệ gen virut vào tế bào. VP4 được định vị ở phía bên trong của capsit. Phụ thuộc vào kiểu và mức độ khử nước, hạt virut có đường kính nằm trong khoảng từ 27 đến 30nm.

Picornavirut có hệ gen ssARN(+) được polyadenyl hoá một rãnh thẳng có từ 7,1 đến 8,9kb, mà gồm có ORF đơn mã hoá polyprotein. ARN hệ gen virut có protein virut (VPg) ở đầu 5' của nó thay cho cấu trúc nắp nucleotit được metyl hoá. UTR dài ở đầu 5' chứa vị trí đi vào ribosom bên trong (IRES). Vùng P1 mã hoá polypeptit cấu trúc. Các vùng P2 và P3 mã hoá protein không cấu trúc kết hợp với sự sao chép. 3' UTR ngắn hơn là quan trọng trong việc tổng hợp sợi (-). L là protein dẫn đầu đầu cuối N bổ sung có mặt trong một số chi mà có thể là proteaza (aphthovirut, erbovirut) hoặc có chức năng khác (kobuvirut, cardiovirut).

ARN virion là lây nhiễm và có tác dụng như hệ gen và ARN thông tin virut. IRES cho phép sự dịch mã trực tiếp của polyprotein. Polyprotein được xử lý ban đầu bằng (các) proteaza virut thành các protein tiền chất và trưởng thành khác nhau để thu được protein

cấu trúc, replicaza, VPg, và nhiều protein mà cải biến tế bào vật chủ, cuối cùng dẫn đến sự tiêu tế bào.

Enterovirut 71 (EV71) là thành viên của họ *Picornaviridae* của ARN virus sợi đơn. Nó cũng là virus không bọc và capsit của nó được cấu trúc từ nhiều protein vỏ được tạo ra như các mảnh sản phẩm dịch mã virus đơn lẻ. Việc xử lý polyprotein virus thành các thành phần cấu trúc và không cấu trúc được thể hiện trên Fig.1 (kỹ thuật hiện nay). Vùng P1 của gen polyprotein mã hoá protein cấu trúc trong khi các vùng P2 và P3 mã hoá các thành phần không cấu trúc của virus. Sau khi tách tiền chất protein cấu trúc P1 (1ABCD trên Fig.1) ra khỏi polyprotein bởi proteaza virus 2A, tiền chất P1 được xử lý thành các capsit protein VP0, VP1 (mảnh 1D trên Fig.1) và VP3 (mảnh 1C trên Fig.1). Thành phần 3C và tiền chất 3CD của nó – được mã hoá bởi vùng P3 – là proteaza virus chịu trách nhiệm cho việc xử lý tiền chất P1 thành capsit protein. Các đoạn khởi đầu VP0, VP1 và VP3 tập hợp một cách tự động thành capsit rỗng và các tác giả sáng chế đã cho rằng ARN virus được đóng gói thành hạt sau khi tập hợp hạt rỗng. Sự kết hợp của capsit rỗng với ARN hệ gen dẫn đến sự dịch chuyển cấu trúc, sự tiếp thu ARN, sự tách tự xúc tác VP0 thành VP2 (mảnh 1B trên Fig.1) và VP4 (mảnh 1A trên Fig.1), và sự trưởng thành thành 150S virion ổn định. Các capsit rỗng, chứa tiền chất VP0 không tách, được tìm thấy một cách thông thường trong quá trình lây nhiễm picornavirus.

Việc sản xuất EV71 VLP ở tế bào côn trùng đã thu được từ sự biểu hiện đồng thời của protein tiền chất P1 với 3CD proteaza (Hu et al., 2003, *Biotechnology Letters* 25: 919-925). Việc sử dụng vector baculovirus đơn lẻ để tạo ra P1 và 3CD được mô tả bởi Chung et al. (2008, *Vaccine* 26: 1855-1862). Các nghiên cứu miễn dịch ở chuột đã thể hiện rằng EV71 VLP được tinh chế tạo ra sự bảo vệ đối với sự thử thách bằng các liều gây chế của virus.

VP1 protein từ EV71 đã được tạo ra ở quả cà chua chuyển gen và cho chuột ăn với quả chuyển gen chứa VP1 dẫn đến sự phát triển IgA phân đặc hiệu VP1 và IgG huyết thanh (Chen et al., 2006, *Vaccine* 24: 2944-2951).

Protein tiền chất P1 và proteaza 3C của virus gây bệnh chân và miệng (FMDV) được biểu hiện đồng thời ở cỏ linh lăng chuyển gen (Dus Santos et al. 2005, *Vaccine* 23: 1838-1843). Cỏ linh lăng được biến nạp một cách ổn định với vector đơn lẻ bao gồm vùng

hệ gen của FMDV P1 (1A, 1B, 1C, 1D), 2A, 16 gốc axit amin thứ nhất có đầu cuối N của 2B, trình tự đầy đủ của 3B1, 3B2, 3B3, 3C và 16 gốc axit amin đầu tiên có đầu cuối N của 3D. Tính miễn dịch của protein thô chiết từ thực vật chuyển gen được chứng minh bằng việc dùng trong bụng ở chuột Balb/c. Chuột được gây miễn dịch cũng được bảo vệ kháng lại thử thách FMDV gây chết. Mức biểu hiện kháng nguyên là thấp đối với mục đích thực tiễn.

Đơn yêu cầu cấp bằng Argentina số AR078257 bộc lộ thực vật chuyển gen biểu hiện capsit virus rỗng, trong đó thực vật chuyển gen bao gồm trong hệ gen của nó cấu trúc ADN mã hoá polypeptit tiền chất P1 được liên kết với 2A proteaza tự xúc tác. Cấu trúc ADN còn có thể chứa mảnh protein 2B được gắn vào trình tự mã hoá 3C proteaza liên kết với mảnh trình tự mã hoá mảnh protein 3D.

Bản chất kỹ thuật của sáng chế

Sáng chế cũng đề cập đến việc tạo ra protein cấu trúc picornavirus ở thực vật. Cụ thể hơn là, sáng chế cũng đề cập đến việc tạo ra hạt tương tự virus bao gồm protein cấu trúc picornavirus ở thực vật.

Do đó, sáng chế đề xuất phương pháp (A) tạo ra hạt tương tự picornavirus (PVLP) ở thực vật, phương pháp này bao gồm các bước:

a) đưa axit nucleic thứ nhất bao gồm vùng điều hoà thứ nhất hoạt động ở thực vật liên kết theo cách hoạt động được với trình tự nucleotit mã hoá một hoặc nhiều polyprotein picornavirus, vào thực vật, hoặc bộ phận của thực vật,

b) đưa axit nucleic thứ hai bao gồm vùng điều hoà thứ hai hoạt động ở thực vật và liên kết theo cách hoạt động được với trình tự nucleotit thứ hai mã hoá một hoặc nhiều proteaza;

c) ủ thực vật, bộ phận của thực vật trong các điều kiện mà cho phép biểu hiện axit nucleic thứ nhất và thứ hai, nhờ đó tạo ra PVLP.

Sáng chế cũng đề xuất phương pháp (B) tạo ra hạt tương tự picornavirus (PVLP), phương pháp này bao gồm các bước:

a) cung cấp thực vật, bộ phận của thực vật, hoặc tế bào thực vật bao gồm axit nucleic thứ nhất bao gồm vùng điều hoà thứ nhất hoạt động ở thực vật liên kết theo cách

hoạt động được với trình tự nucleotit thứ nhất mã hoá một hoặc nhiều polypeptid picornavirus và axit nucleic thứ hai bao gồm vùng điều hoà thứ hai hoạt động ở thực vật liên kết theo cách hoạt động được với trình tự nucleotit thứ hai mã hoá một hoặc nhiều proteaza;

b) ủ thực vật, bộ phận của thực vật, hoặc tế bào thực vật trong các điều kiện mà cho phép biểu hiện axit nucleic, nhờ đó tạo ra PVLP.

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

Hơn nữa, trong phương pháp (A) hoặc (B), tỷ lệ phần trăm của axit nucleic thứ nhất với axit nucleic thứ hai được đưa vào thực vật, bộ phận của thực vật, hoặc tế bào thực vật có thể nằm trong khoảng từ 95%:5% đến 50%:50%, hoặc từ 20:1 đến 0,5:1.

Sáng chế cũng bao gồm các phương pháp (A) hoặc (B) như được mô tả trên đây, trong đó trình tự axit nucleic thứ nhất bao gồm vùng điều hoà thứ nhất được liên kết theo cách hoạt động được với một hoặc nhiều hơn một vùng tăng cường comovirus, trình tự nucleotit mã hoá polypeptid, và một hoặc nhiều hơn một geminivirus yếu tố khuếch đại, và axit nucleic thứ ba mã hoá geminivirus replicaza được đưa vào thực vật hoặc bộ phận của thực vật. Một hoặc nhiều hơn một vùng tăng cường comovirus có thể là comovirus UTR, ví dụ, virus khảm cây đậu đũa có thể dịch mã (CPMV-HT) UTR như CPMV-HT 5', 3'UTR, hoặc tổ hợp của chúng. Một hoặc nhiều hơn một yếu tố khuếch đại geminivirus có thể được chọn từ vùng trong gen dài của virus lùn vàng đậu (BeYDV LIR), và vùng trong gen ngắn BeYDV (BeYDV SIR).

Các phương pháp như được mô tả trên đây (Phương pháp A) cũng có thể bao gồm việc đưa trình tự axit nucleic khác mã hoá gen ngăn chặn việc câm, ví dụ HcPro hoặc p19.

Các phương pháp như được mô tả trên đây (Phương pháp B) cũng có thể còn bao gồm việc cung cấp thực vật, bộ phận của thực vật hoặc tế bào thực vật bao gồm trình tự axit nucleic khác mã hoá gen ngăn chặn việc câm, 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 (bước a), axit nucleic được biểu hiện một cách tạm thời ở thực vật. Theo cách khác, ở bước đưa vào (bước a), axit nucleic được biểu hiện một cách ổn định ở thực vật.

Các phương pháp (A) và (B) như được mô tả trên đây còn có thể bao gồm bước thu hoạch thực vật và tinh chế PVLP.

Sáng chế bao gồm được phẩm bao gồm liệu hữu hiệu của PVLP như được mô tả trên đây để gây ra sự đáp ứng miễn dịch, và chất mang được dùng.

Sáng chế cũng bao gồm phương pháp gây miễn dịch đối với sự lây truyền picornavirut ở đối tượng, bao gồm việc cho đối tượng dùng PVLP như được mô tả trên đây. PVLP có thể được dùng cho đối tượng qua đường miệng, trong da, trong mũi, trong cơ, trong bụng, trong tĩnh mạch, hoặc dưới da.

Sáng chế cũng đề xuất vật liệu thực vật bao gồm PVLP được tạo ra bằng phương pháp (A) và/hoặc (B) được mô tả trên đây. Vật liệu thực vật có thể được sử dụng trong việc gây miễn dịch đối với sự lây nhiễm picornavirut ở đối tượng. Vật liệu thực vật cũng có thể được trộn lẫn như phần bổ sung thức ăn.

Bản chất kỹ thuật của sáng chế không nhất thiết 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 của sáng chế sẽ trở nên rõ ràng hơn dựa vào phần mô tả sau đây trong đó việc tham chiếu được thực hiện đến các hình vẽ kèm theo trong đó:

Fig.1 thể hiện phép biểu diễn của kỹ thuật hiện nay về hệ gen picornavirut (enterovirut 71) và polyprotein xử lý các chất trung gian. (từ trang mạng ViralZone)

Fig.2 thể hiện phép phân tích thẩm tách Western về sự biểu hiện P1 và xử lý bởi 3CD. Mười microgam protein chiết từ thực vật được biến nạp với vectơ biểu hiện được nhận dạng trên đây được nạp và được điện di trong các điều kiện không khử. Các kháng thể đơn dòng kháng VP1 của chuột được sử dụng cho việc phát hiện miễn dịch. Tỷ lệ này cho biết tỷ lệ của P1 (cấu trúc 1300 hoặc 1301 xem bảng 1) với 3CD (cấu trúc 1310, 1311 hoặc 1315 xem bảng 1) chủng *Agrobacterium* trong huyền phù vi khuẩn được sử dụng cho sự biến nạp. Vị trí được mong đợi của P1 và VP1 được chỉ định.

Fig.3 thể hiện việc sàng lọc các chiến lược biểu hiện 3CD đối với sự tích tụ tối đa của VP1. Năm microgam protein chiết từ thực vật được biến nạp với các vectơ biểu hiện

được nhận dạng trên đây được nạp và được điện di trong các điều kiện không khử. Các kháng thể đơn dòng kháng VP1 của chuột được sử dụng cho việc phát hiện miễn dịch. Các tỷ lệ này cho biết tỷ lệ của P1 (1301) với 3CD (1310, 1311, 1312 và 1313) chủng *Agrobacterium* trong huyền phù vi khuẩn được sử dụng đối với việc biến nạp.

Fig.4 thể hiện việc đánh giá bộ capsit EV71. (A) SDS-PAGE được nhuộm Coomassie và phép phân tích thẩm tách western của các phân đoạn rửa giải từ việc tách sắc ký loại cỡ (SEC) của protein chiết từ thực vật biểu hiện đồng thời P1 và 3CD (cấu trúc 1301+1310 (4:0.5)). Dải giả định tương ứng với VP1 trong gen được nhuộm Coomassie được chỉ định. (B) Thử nghiệm kính hiển vi điện tử truyền qua nhuộm màu của phân đoạn rửa giải SEC 12. Cột là 100nm.

Fig.5 thể hiện đặc điểm của EV71 PVLP được tinh chế. (A) SDS-PAGE được nhuộm Coomassie và phép phân tích thẩm tách western của EV71 PVLP đã tinh chế. Dải tương ứng với VP1 trong gel được nhuộm Coomassie được chỉ định. Các dải khác, tương ứng theo phân tử lượng với EV71 capsit protein khác cũng được nhận dạng. (B) Thử nghiệm kính hiển vi điện tử truyền nhuộm màu âm tính của EV71 PVLP được tinh chế. Mẫu được pha loãng 1/100 trước khi thử nghiệm. Cột là 100nm.

Fig.6 thể hiện đặc điểm của dây 479-23-018 bằng kính hiển vi điện tử.

Fig.7 thể hiện phép phân tích kính hiển vi điện tử cryo của EV71 PVLP được chiết bằng phương pháp và quy trình được trợ giúp enzym, được chọn bằng HIC (dãy số 479-31-020).

Fig.8 thể hiện phép phân tích kính hiển vi điện tử cryo của EV71 PVLP được chiết bằng phương pháp chiết cơ học (độ pH=8,0) kèm xử lý nhiệt (dãy số 479-32-020).

Fig.9A thể hiện 3CD từ chủng EV71 HK08 bao gồm axit amin 1549-2193 (SEQ ID NO: 1), như được nêu trong GenBank ID ADG57603. Fig.9B thể hiện 3CD từ chủng EV71 HK08 bao gồm nucleotit 5387 -7321 (SEQ ID NO: 2) được nêu trong GenBank ID GQ279369. Fig.9C thể hiện 3CD từ chủng EV71 GDFS08 bao gồm các axit amin 1549-2193 (SEQ ID NO: 3) như được nêu trong GenBank ID ACI25378. Fig.9D thể hiện 3CD từ chủng EV71 GDFS08 bao gồm nucleotit 5387 -7321 (SEQ ID NO: 4) được nêu trong GenBank ID FJ194964. Fig.9E thể hiện P1 các trình tự axit amin GenBank ID ADG57603 (các axit amin 1-862) (SEQ ID NO: 5). Fig.9F thể hiện trình tự P1 nucleotit GenBank ID

GQ279369 (nucleotit 743-3328) (SEQ ID NO: 6). Fig.9G thể hiện trình tự PVgp1 polyprotein nucleotit từ kiểu huyết thanh enterovirut C PV-1 của người (GenBank ID NC_002058 đối với hệ gen và NP_041277 đối với polyprotein: nt 5438-7369) (SEQ ID NO: 7). Fig.9H thể hiện trình tự axit amin của polyprotein từ Poliovirut (aa 1566-2209 từ GenBank ID NP_041277) (SEQ ID NO: 8). Fig.9I thể hiện trình tự nucleotit của PVgp1 polyprotein [*Human enterovirus C*] (nt 743-3385 từ GenBank ID NC_002058) (SEQ ID NO: 9). Fig.9J thể hiện trình tự axit amin của polyprotein [enterovirut C của người] GenBank ID NP_041277 (aa 1-881 từ GenBank ID NP_041277) (SEQ ID NO: 10).

Mô tả chi tiết sáng chế

Phần mô tả sau đây mô tả các phương án ưu tiên.

Sáng chế đề cập đến hạt tương tự virus (VLP) bao gồm một hoặc nhiều protein cấu trúc picornavirut (nghĩa là protein tương tự picornavirut hoặc PVLP), và phương pháp tạo ra PVLP ở thực vật hoặc ở các bộ phận của thực vật. Do đó, PVLP có thể bao gồm một hoặc nhiều hơn một protein cấu trúc picornavirut. Ví dụ, PVLP có thể bao gồm một hoặc nhiều hơn một protein cấu trúc enterovirut.

Picornavirut có thể được chọn từ nhóm gồm có Aphthovirut, Avihepatovirut, Cardiovirut, Enterovirut, Erbovirut, Hepatovirut, Kobuvirut, Parechovirut, Sapelovirut, Senecavirut, Teschovirut và Tremovirut. Ví dụ không hạn chế về picornavirut có thể là Enterovirut, ví dụ Enterovirut 71 (EV71) hoặc enterovirut C của người (cũng được biết đến như poliovirut).

Sáng chế một phần đề xuất phương pháp tạo ra VLP, ví dụ PVLP hoặc hạt tương tự enterovirut ở thực vật. Phương pháp có thể bao gồm việc đưa axit nucleic thứ nhất bao gồm vùng điều hoà thứ nhất hoạt động ở thực vật liên kết theo cách hoạt động được với trình tự nucleotit thứ nhất mã hoá một hoặc nhiều polyprotein picornavirut vào thực vật, hoặc bộ phận của thực vật và đưa axit nucleic thứ hai bao gồm vùng điều hoà thứ hai hoạt động ở thực vật liên kết theo cách hoạt động được với trình tự nucleotit thứ hai mã hoá proteaza. Tiếp theo bằng việc ủ thực vật hoặc bộ phận của thực vật trong các điều kiện mà cho phép biểu hiện axit nucleic, nhờ đó tạo ra PVLP.

Thuật ngữ "hạt tương tự virus" (VLP), hoặc "các hạt tương tự virus" hoặc "VLP" dùng để chỉ các cấu trúc mà tự tập hợp và bao gồm một hoặc nhiều hơn một protein cấu

trúc, ví dụ một hoặc nhiều hơn một protein cấu trúc picornavirut, hoặc một hoặc nhiều hơn một protein cấu trúc enterovirut hoặc tổ hợp của chúng, ví dụ nhưng không chỉ giới hạn ở protein cấu trúc VP0, VP1, VP2, VP3, VP4 hoặc tổ hợp của chúng. VLP thường là tương tự về mặt hình thái học và về mặt kháng nguyên với các virion được tạo ra trong sự lây truyền, nhưng thiếu các thông tin di truyền đủ để sao chép và do đó là không lây nhiễm. VLP có thể được tạo ra trong tế bào vật chủ thích hợp bao gồm tế bào vật chủ thực vật. Việc chiết tiếp theo từ tế bào vật chủ và phụ thuộc vào việc phân lập và tinh chế tiếp trong các điều kiện thích hợp, VLP có thể được tinh chế như cấu trúc nguyên vẹn.

Thuật ngữ "hạt tương tự Picornavirut" (PVLP), dùng để chỉ VLP hoặc các VLP bao gồm một hoặc nhiều hơn một protein cấu trúc picornavirut. Thuật ngữ hoặc "hạt tương tự enterovirut" dùng để chỉ VLP hoặc các VLP bao gồm một hoặc nhiều hơn một protein cấu trúc enterovirut. Ví dụ về protein cấu trúc picornavirut có thể bao gồm, nhưng không chỉ giới hạn ở protein cấu trúc VP0, VP1, VP2, VP3, VP4, hoặc tổ hợp của chúng. Ví dụ về protein cấu trúc enterovirut có thể bao gồm, nhưng không chỉ giới hạn ở protein cấu trúc VP0, VP1, VP2, VP3, VP4, hoặc tổ hợp của chúng.

Thuật ngữ polyprotein có nghĩa là protein mà bao gồm một hoặc nhiều hơn một protein hoặc tiền chất protein, mà khi được xử lý thủy phân protein tạo ra một hoặc nhiều protein. Ví dụ polyprotein có thể bao gồm một hoặc nhiều hơn một protein cấu trúc. Một hoặc nhiều protein ví dụ protein cấu trúc, trong polypeptit có thể ví dụ được tách bằng các vị trí tách, ví dụ vị trí tách proteaza. Ví dụ không hạn chế đối với "polyprotein" là tiền chất protein cấu trúc P1 cũng được dùng để chỉ "vùng P1". Vùng P1 được xác định như phần polyprotein picornavirut mà tạo ra "protein cấu trúc" hoặc "protein vỏ" ví dụ VP0, VP1, VP2, VP3, VP4 hoặc tổ hợp của chúng. Các ví dụ không hạn chế về picornavirut P1, hoặc mảnh P1 mà có thể được sử dụng theo sáng chế bao gồm P1 từ enterovirut ví dụ enterovirut 71.

Ví dụ về vùng P1, mà không được xét đến làm giới hạn, là trình tự axit amin được nêu trong GenBank ID ADG57603 bao gồm các axit amin 1-862 (SEQ ID NO: 5) hoặc trình tự có ít nhất khoảng từ 90 đến 100% trình tự tương tự với nó, bao gồm tỷ lệ phần trăm tương tự bất kỳ năm trong khoảng này, như 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100% trình tự tương tự với nó. Hơn nữa, ví dụ không hạn chế về trình tự nucleotit mã hoá vùng P1 được nêu trong GenBank ID GQ279369 bao gồm nucleotit 743-3328 (SEQ ID

NO: 6) hoặc trình tự có ít nhất khoảng từ 90 đến 100% trình tự tương tự với nó, bao gồm tỷ lệ phần trăm tương tự bất kỳ nằm trong khoảng này, như 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100% trình tự tương tự với nó.

Ví dụ khác về vùng P1, mà không được xét đến làm giới hạn, là trình tự axit amin được nêu trong GenBank ID NP_041277 bao gồm các axit amin 1566-2209 (SEQ ID NO: 8) hoặc trình tự có ít nhất khoảng từ 90 đến 100% trình tự tương tự với nó, bao gồm tỷ lệ phần trăm bất kỳ tương tự nằm trong khoảng này, như 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100% trình tự tương tự với nó. Hơn nữa, ví dụ không hạn chế về trình tự nucleotit mã hoá vùng P1 được nêu trong GenBank ID NC_002058 bao gồm nucleotit 5438-7369 (SEQ ID NO: 7) hoặc trình tự có ít nhất khoảng từ 90 đến 100% trình tự tương tự với nó, bao gồm tỷ lệ phần trăm bất kỳ tương tự nằm trong khoảng này, như 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100% trình tự tương tự với nó. Theo ví dụ khác mà không được xét đến làm giới hạn vùng P1 có trình tự axit amin được nêu trong GenBank ID NP_041277 bao gồm các axit amin 1-881 (SEQ ID NO: 10) hoặc trình tự có ít nhất khoảng từ 90 đến 100% trình tự tương tự với nó, bao gồm tỷ lệ phần trăm bất kỳ tương tự nằm trong khoảng này, như 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100% trình tự tương tự với nó. Hơn nữa, ví dụ không hạn chế về trình tự nucleotit mã hoá vùng P1 được nêu trong GenBank ID NC_002058 bao gồm nucleotit 743-3385 (SEQ ID NO: 9) hoặc trình tự có ít nhất khoảng từ 90 đến 100% trình tự tương tự với nó, bao gồm tỷ lệ phần trăm bất kỳ tương tự nằm trong khoảng này, như 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100% trình tự tương tự với nó.

“Polyprotein picornavirut” dùng để chỉ tất cả hoặc một phần polyprotein picornavirut được phân lập từ picornavirut, có mặt trong chủng picornavirut xuất hiện tự nhiên hoặc biến thể bất kỳ hoặc phân lập, ví dụ polyprotein enterovirut. Tương tự, “protein cấu trúc picornavirut” có thể dùng để chỉ tất cả hoặc một phần protein cấu trúc picornavirut được phân lập từ picornavirut, có mặt trong chủng picornavirut xuất hiện tự nhiên hoặc biến thể bất kỳ hoặc phân lập, ví dụ cấu trúc protein enterovirut, ví dụ thu được từ poliovirut hoặc enterovirut 71. Do đó, thuật ngữ “polyprotein picornavirut” và “protein cấu trúc picornavirut” và tương tự bao gồm các biến thể xuất hiện tự nhiên của polyprotein picornavirut, protein cấu trúc picornavirut, hoặc tổ hợp của chúng, được tạo ra bằng việc đột biến trong vòng đời của virut hoặc được tạo ra trong sự đáp ứng với áp lực lựa chọn

(ví dụ, liệu pháp điều trị dược chất, sự kéo dài tính hướng kích thích của tế bào vật chủ hoặc tính lây nhiễm, v.v). Thuật ngữ “polyprotein picornavirut” còn bao gồm “polyprotein enterovirut” và “cấu trúc protein enterovirut” và tương tự bao gồm các biến thể xuất hiện tự nhiên của polyprotein enterovirut, cấu trúc protein enterovirut, hoặc tổ hợp của chúng, được tạo ra bằng việc đột biến trong vòng đời của virut hoặc được tạo ra trong sự đáp ứng với áp lực lựa chọn (ví dụ, liệu pháp điều trị dược chất, sự kéo dài tính hướng kích thích của tế bào vật chủ hoặc tính lây nhiễm, v.v). Thuật ngữ “polyprotein picornavirut” cũng có thể bao gồm “poliovirut polyprotein” và “poliovirut protein cấu trúc” và tương tự bao gồm các biến thể xuất hiện tự nhiên của polyprotein poliovirut, cấu trúc protein poliovirut, hoặc tổ hợp của chúng, được tạo ra bằng việc đột biến trong vòng đời của virut hoặc được tạo ra trong sự đáp ứng với áp lực lựa chọn (ví dụ, liệu pháp điều trị dược chất, sự kéo dài tính hướng kích thích của tế bào vật chủ hoặc tính lây nhiễm, v.v). Như người có hiểu biết trung bình về lĩnh vực này thấy được, nguyên thể và biến thể của polyprotein picornavirut, enterovirut hoặc poliovirut hoặc picornavirut, cấu trúc protein enterovirut hoặc poliovirut cũng có thể được tạo ra bằng cách sử dụng các kỹ thuật tái tổ hợp.

Polyprotein có thể bao gồm một hoặc nhiều protein cấu trúc ví dụ capsit protein. Các ví dụ không hạn chế về protein cấu trúc picornavirut hoặc capsit protein là protein picornavirut VP0, VP1, VP2, VP3 và VP4 và mảnh VP0, VP1, VP2, VP3 và VP4. Các ví dụ không hạn chế về VP0, VP1, VP2, VP3 và VP4, hoặc mảnh của VP0, VP1, VP2, VP3 và VP4 protein mà có thể được sử dụng theo sáng chế bao gồm VP0, VP1, VP2, VP3 và VP4 protein từ enterovirut, ví dụ poliovirut hoặc enterovirut 71. Hơn nữa, cấu trúc polyprotein protein, hoặc tổ hợp của chúng có thể là ví dụ từ chủng enterovirut 71 HK08 hoặc chủng GDFS08. Theo ví dụ không hạn chế khác, cấu trúc polyprotein protein hoặc tổ hợp của chúng có thể là từ enterovirut C của người, cũng được biết đến là poliovirut.

Trình tự axit amin tương tự hoặc giống nhau có thể được tính toán bằng cách sử dụng chương trình BLASTP và TBLASTN mà dùng thuật toán BLAST (công cụ tra cứu thuật toán quỹ tích căn bản) 2.0. Các kỹ thuật để tính toán trình tự axit amin tương tự hoặc giống nhau là đã biết đối với người có hiểu biết trung bình về lĩnh vực này và sử dụng thuật toán BLAST được mô tả trong ấn phẩm: ALTSCHUL et al. (1990, J Mol. Biol. 215: 403- 410) và ALTSCHUL et al. (1997, Nucleic Acids Res. 25: 3389-3402).

Theo sáng chế, picornavirut, enterovirut (bao gồm poliovirut) VLP được tạo ra ở thực vật, bộ phận của thực vật hoặc tế bào thực vật, bằng cách biểu hiện đồng thời axit nucleic (axit nucleic thứ nhất) mã hoá polyprotein picornavirut, enterovirut hoặc poliovirut, ví dụ nhưng không chỉ giới hạn ở P1, với axit nucleic thứ hai mã hoá proteaza, ví dụ picornavirut, enterovirut hoặc poliovirut proteaza như ví dụ nhưng không chỉ giới hạn ở 3CD, và nhờ đó tạo ra VLP.

Ví dụ về proteaza, mà không được xét đến làm giới hạn, là trình tự axit amin từ chủng EV71 HK08 bao gồm các axit amin 1549-2193, như được nêu trong GenBank ID ADG57603 (SEQ ID NO:1). Trình tự nucleotit được nêu trong GenBank ID GQ279369 (SEQ ID NO:2) từ nucleotit 5387 đến nucleotit 7321 hoặc trình tự có ít nhất khoảng từ 90 đến 100% trình tự tương tự với SEQ ID NO:2, bao gồm tỷ lệ phần trăm bất kỳ tương tự nằm trong khoảng này, như 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100% trình tự tương tự với nó. Ví dụ không hạn chế khác là trình tự axit amin từ chủng EV71 GDFS08 bao gồm các axit amin 1549-2193 như được nêu trong GenBank ID ACI25378 (SEQ ID NO:3). Trình tự nucleotit được nêu trong GenBank ID FJ194964 (SEQ ID NO:4), từ nucleotit 5387 đến nucleotit 7321 có thể được sử dụng để tạo ra trình tự axit amin. Hơn nữa, trình tự có khoảng từ 90 đến 100% trình tự giống với SEQ ID NO:4, bao gồm tỷ lệ phần trăm bất kỳ tương tự nằm trong khoảng này, như 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100% trình tự tương tự với nó.

Axit nucleic thứ nhất và axit nucleic thứ hai, có thể được đưa vào thực vật trong cùng bước hoặc có thể được đưa vào thực vật theo trình tự.

Trình tự

Ví dụ không hạn chế về trình tự mà có thể được sử dụng với sáng chế bao gồm:

Trình tự P1 từ Aphthovirut, Avihepatovirut, Cardiovirut, Enterovirut, Erbovirut, Hepatovirut, Kobuvirut, Parechovirut, Sapelovirut, Senecavirut, Teschovirut và Tremovirut có thể được sử dụng để tạo ra P1 polyprotein. Theo ví dụ không hạn chế, trình tự mã hoá P1 polyprotein có thể là từ Enterovirut, ví dụ Enterovirut 71 hoặc enterovirut C của người (cũng đã được biết đến như poliovirut).

Hơn nữa, các ví dụ không hạn chế về trình tự mà có thể được sử dụng để mã hoá proteaza để sử dụng như được mô tả trong bản mô tả này bao gồm trình tự proteaza 3CD,

ví dụ, thu được từ Aphthovirut, Avihepatovirut, Cardiovirut, Enterovirut, Erbovirut, Hepatovirut, Kobuvirut, Parechovirut, Sapelovirut, Senecavirut, Teschovirut và Tremovirut. Theo ví dụ không hạn chế, trình tự mã hoá 3CD proteaza có thể là từ Enterovirut, ví dụ Enterovirut 71 hoặc Poliovirut (cũng được biết đến như enterovirut C của người).

Các tác giả sáng chế đã phát hiện ra rằng bằng việc đưa vào và biểu hiện đồng thời polyprotein và proteaza ở thực vật hoặc bộ phận của thực vật mà sản lượng của VLP được tạo ra có thể được điều biến. Polyprotein và proteaza có thể được tạo ra trên các cấu trúc axit nucleic riêng rẽ và biểu hiện đồng thời hoặc chúng có thể được tạo ra trên cùng cấu trúc nhưng mỗi trình tự được biểu hiện theo cách khác nhau, nếu cần, để tối ưu hoá việc sản xuất VLP như được mô tả dưới đây.

Thuật ngữ “biểu hiện đồng thời” có nghĩa là hai hoặc nhiều hơn hai, trình tự nucleotit được biểu hiện ở khoảng cùng thời gian trong thực vật, trong cùng mô thực vật và trong cùng tế bào ở thực vật. Hơn nữa, hai hoặc nhiều hơn hai, trình tự nucleotit có thể cần được biểu hiện trong cùng khoang tế bào như, ví dụ, mạng lưới nội chất, bộ máy Golgi, thể không hợp bào, xytosol, ti thể, lập lục, peroxysom. Trình tự nucleotit không cần phải được biểu hiện ở cùng thời gian chính xác. Tốt hơn là, hai hoặc nhiều trình tự nucleotit được biểu hiện theo cách sao cho các sản phẩm được mã hoá có cơ hội tương tác. Ví dụ, proteaza có thể được biểu hiện trước hoặc trong khoảng thời gian khi polyprotein được biểu hiện sao cho việc tách polyprotein vào cấu trúc protein có thể đạt được. Hai hoặc nhiều hơn hai trình tự nucleotit có thể được biểu hiện đồng thời bằng cách sử dụng hệ biểu hiện tạm thời, trong đó hai hoặc nhiều trình tự được đưa vào thực vật ở khoảng cùng thời gian trong các điều kiện mà cả hai trình tự được biểu hiện. Hai hoặc nhiều hơn hai trình tự có thể có mặt trên các cấu trúc khác nhau và việc biểu hiện đồng thời đòi hỏi việc đưa vào của mỗi trong số các cấu trúc vào thực vật, bộ phận của thực vật hoặc tế bào thực vật, hoặc hai hoặc nhiều hơn hai trình tự có thể có mặt trên một cấu trúc và cấu trúc được đưa vào thực vật, bộ phận của thực vật hoặc tế bào thực vật.

Theo cách khác, thực vật bao gồm một trong số các trình tự nucleotit, ví dụ trình tự mã hoá proteaza có thể được biến nạp, một cách tạm thời hoặc theo cách ổn định, với trình tự bổ sung mã hoá polyprotein. Trong trường hợp này, trình tự mã hoá proteaza có thể được biểu hiện trong mô mong muốn, trong giai đoạn phát triển mong muốn hoặc sự biểu

hiện của nó có thể được cảm ứng bằng cách sử dụng đoạn khởi đầu có thể cảm ứng và trình tự bổ sung mã hoá polyprotein có thể được biểu hiện trong các điều kiện tương tự và trong cùng mô, để bảo đảm rằng các trình tự nucleotit được biểu hiện đồng thời. Ngoài ra, trình tự mã hoá polyprotein có thể được biến nạp, theo cách tạm thời hoặc theo cách ổn định, với trình tự bổ sung mã hoá proteaza. Trong trường hợp này, trình tự mã hoá polyprotein có thể được biểu hiện trong mô mong muốn, trong giai đoạn phát triển mong muốn hoặc sự biểu hiện của nó có thể được cảm ứng bằng cách sử dụng đoạn khởi đầu có thể cảm ứng và trình tự bổ sung mã hoá proteaza có thể được biểu hiện trong các điều kiện tương tự và trong cùng mô, để bảo đảm rằng các trình tự nucleotit được biểu hiện đồng thời.

Như có thể thấy trên các Fig.2 và 3, mức kết tụ VLP ở thực vật, bộ phận của thực vật hoặc tế bào thực vật, bị tác động bởi tỷ lệ của *Agrobacterium* chứa polyprotein với *Agrobacterium* chứa proteaza được thâm nhập vào thực vật, bộ phận của thực vật hoặc tế bào thực vật. Tỷ lệ của *Agrobacterium* chứa polyprotein với chứa proteaza có thể nằm trong khoảng ví dụ từ 20:1 đến 0,5:1 (polyprotein:proteaza), hoặc lượng bất kỳ giữa chúng, ví dụ khoảng 20:1, 18:1, 16:1, 14:1, 12:1, 10:1, 9:1, 8:1, 7:1, 6:1, 5:1, 4:1, 3:1, 2:1, 1:1, 0,5:1 (polyprotein:proteaza), hoặc lượng bất kỳ giữa chúng.

Tỷ lệ của polyprotein với proteaza có thể được thay đổi ví dụ bằng cách đưa các tỷ lệ khác nhau của *Agrobacterium* chứa axit nucleic thứ nhất với *Agrobacterium* chứa axit nucleic thứ hai vào thực vật, bộ phận của thực vật hoặc tế bào thực vật. Theo cách khác, nếu polyprotein và proteaza có mặt trên cùng cấu trúc và do đó được đưa vào cùng *Agrobacterium*, chúng có thể được biểu hiện một cách khác nhau trong thực vật, bộ phận của thực vật hoặc tế bào thực vật bằng cách sử dụng các đoạn khởi đầu thích hợp sao cho tỷ lệ mong muốn của polyprotein với proteaza thu được.

Do đó, sáng chế cũng đề xuất phương pháp làm tăng sản lượng sản xuất PVLP bằng việc điều biến tỷ lệ giữa axit nucleic thứ nhất và thứ hai.

Theo một phương án, tỷ lệ phần trăm của *Agrobacterium* chứa proteaza có thể nằm trong khoảng từ 0,5% đến 50% tổng *Agrobacterium* được thâm nhiễm hoặc lượng bất kỳ giữa chúng. Ví dụ, tỷ lệ phần trăm của *Agrobacterium* chứa proteaza có thể bằng 0, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%,

19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49% hoặc 50% hoặc lượng bất kỳ giữa chúng.

Tỷ lệ phần trăm của *Agrobacterium* chứa polyprotein với *Agrobacterium* chứa proteaza có thể bằng từ 95%:5% đến 40%:60% tổng *Agrobacterium* được thâm nhiễm hoặc lượng bất kỳ giữa chúng. Ví dụ tỷ lệ phần trăm của *Agrobacterium* chứa polyprotein nằm trong tổng lượng của *Agrobacterium* được thâm nhiễm có thể bằng 99%, 98%, 97%, 96%, 95%, 94%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, 85%, 84%, 83%, 82%, 81%, 80%, 79%, 78%, 77%, 76%, 75%, 74%, 73%, 72%, 71%, 70%, 69%, 68%, 67%, 66%, 65%, 64%, 63%, 62%, 61%, 60%, 59%, 58%, 57%, 56%, 55%, 54%, 53%, 52% hoặc 51%. Ví dụ, tỷ lệ phần trăm của *Agrobacterium* chứa polyprotein với *Agrobacterium* chứa proteaza có thể nằm trong khoảng từ 50%:50% đến 95%:5%, hoặc tỷ lệ phần trăm bất kỳ nằm trong khoảng này hoặc tỷ lệ phần trăm bất kỳ nằm trong khoảng từ *Agrobacterium* chứa polyprotein đến *Agrobacterium* chứa proteaza có thể bằng 50%:50%, 55%:45%, 60%:40%, 65%:35%, 70%:30%, 75%:25%, 80%:20%, 85%:15%, 90%:10%, 95%:5%, hoặc tỷ lệ phần trăm bất kỳ giữa chúng.

Sự biểu hiện của trình tự nucleotit thứ nhất và thứ hai trong tế bào thực vật tạo ra VLP, và VLP có thể được sử dụng ví dụ để tạo ra kháng thể mà có khả năng gắn kết protein virus ví dụ protein cấu trúc picornavirus, bao gồm nhưng không chỉ giới hạn ở VP0, VP1, VP2, VP3 và/hoặc VP4. VLP, khi được dùng cho đối tượng, gây ra đáp ứng miễn dịch.

Như được mô tả hơn nữa dưới đây, tỷ lệ của polyprotein với proteaza còn có thể được thay đổi ví dụ bằng cách biểu hiện theo cách khác nhau polyprotein và proteaza. Sự biểu hiện có thể được thay đổi bằng cách điều biến ví dụ sao chép, phiên mã, dịch mã hoặc tổ hợp của chúng, của polyprotein, proteaza, hoặc cả protein và proteaza. Ví dụ về các nhân tố điều hoà khác nhau, bao gồm các đoạn khởi đầu, yếu tố khuếch đại, đoạn tăng cường hoặc tổ hợp của chúng, có thể được sử dụng ngoài việc thay đổi tỷ lệ của *Agrobacterium* chứa polyprotein với *Agrobacterium* chứa proteaza được thâm nhiễm như được mô tả trên đây. Tập hợp thứ nhất hoặc tổ hợp của các nhân tố điều hoà có thể được sử dụng để điều hoà sự sao chép, phiên mã hoặc tổ hợp của chúng, của axit nucleic thứ nhất và tập hợp thứ hai hoặc tổ hợp của các nhân tố điều hoà có thể được sử dụng để điều

hoà sự sao chép, phiên mã hoặc tổ hợp của chúng, của axit nucleic thứ hai. Tập hợp thứ nhất hoặc tổ hợp của các nhân tố điều hoà là khác với tập hợp thứ hai hoặc tổ hợp của các nhân tố điều hoà và cho phép sự biểu hiện khác nhau của axit nucleic thứ nhất và thứ hai để cho phép điều biến tỷ lệ của polyprotein:protease *in vivo*. Ví dụ, mà không được xét đến làm giới hạn, một tập hợp hoặc tổ hợp của các nhân tố điều hoà, ví dụ tập hợp thứ nhất, có thể bao gồm yếu tố khuếch đại ví dụ yếu tố thu được từ BeYDV, trong khi yếu tố khuếch đại, ví dụ các yếu tố thu được từ BeYDV, có thể không có mặt trong tập hợp khác hoặc tổ hợp của các nhân tố điều hoà, ví dụ tập hợp thứ hai. Theo cách khác, tập hợp thứ hai có thể bao gồm yếu tố khuếch đại (ví dụ yếu tố thu được từ BeYDV), trong khi yếu tố khuếch đại (ví dụ yếu tố thu được từ BeYDV) có thể không có mặt trong tập hợp thứ nhất hoặc tổ hợp của các nhân tố điều hoà. Theo cách tương tự, cường độ của các đoạn khởi đầu có thể khác nhau giữa tập hợp thứ nhất và tập hợp thứ hai hoặc tổ hợp của các nhân tố điều hoà, hoặc một trong số các đoạn khởi đầu có thể cảm ứng và cấu tử khác, sao cho sự biểu hiện khác nhau giữa polyprotein với proteaza là đạt được *in vivo*.

Kích cỡ

Việc xuất hiện của VLP có thể được phát hiện bằng cách sử dụng phương pháp thích hợp bất kỳ, ví dụ, sucroza gradient hoặc sắc khí loại cỡ. VLP có thể được đánh giá đối với cấu trúc và kích cỡ bởi, ví dụ kính hiển vi điện tử hoặc sắc ký loại cỡ.

Đối với việc sắc ký loại cỡ, tổng protein có thể hoà tan có thể được chiết từ mô thực vật bằng việc tạo đồng nhất (Polytron) mẫu vật liệu thực vật được ép đông lạnh trong chất đệm chiết và vật liệu không thể hoà tan được loại bỏ bằng việc ly tâm. Nồng độ của chất kết tủa trợ giúp PEG cũng có thể có lợi. VLP cũng có thể được tạo ra bằng cách tạo ra thể nguyên sinh hoặc phân đoạn thể nguyên sinh bằng cách sử dụng các phương pháp được mô tả trong công bố đơn quốc tế số WO 2011/035422. Protein có thể hoà tan được định lượng và dịch chiết được chuyển qua cột Sephacryl™, ví dụ cột Sephacryl™ S500. Dextran 2000 xanh có thể được sử dụng làm tiêu chuẩn định cỡ.

Mảnh vụn tế bào có thể được loại bỏ bằng việc ly tâm. Sau đó, dịch chiết ly tâm có thể được lọc. Không muốn bị ràng buộc bởi lý thuyết, các tác giả sáng chế đã cho rằng bước lọc hoặc các bước lọc có thể loại bỏ các chất rắn trong huyền phù, làm giảm gánh nặng sinh học và làm ổn định và quy định dịch chiết trước khi tinh chế tiếp. Do kích cỡ

của chúng, PVLP có thể được tinh chế hơn nữa bằng cách sử dụng việc lọc dòng tiếp tuyến (TFF). Không muốn bị ràng buộc bởi lý thuyết, TFF loại trừ một cách hữu hiệu và theo cách chọn lọc các protein có thể hoà tan có phân tử lượng thấp hơn được tìm thấy trong dịch chiết sạch, bao gồm enzym được sử dụng cho quá trình polyme hoá khử vách tế bào. Hơn nữa, bước TFF cũng cô đặc VLP và cho phép trao đổi chất đệm trong khi chuẩn vị sắc ký. Bước TFF có thể được tiếp theo bởi vài bước sắc ký, ví dụ trao đổi anion, trao đổi cation, sắc ký tương tác kỵ khí (HIC) và/hoặc ái lực giả. Các bước TFF bổ sung có thể được bổ sung vào sau bước sắc ký. Tiếp theo sắc ký và/hoặc TFF, các phân đoạn có thể được phân tích hơn nữa bằng thẩm tách miễn dịch để xác định thành phần protein của phân đoạn.

Phân đoạn riêng biệt có thể ví dụ là dịch nổi (nếu được ly tâm, sa lắng hoặc kết tủa), hoặc phân lọc (nếu được lọc), và được làm giàu đối với protein, hoặc protein siêu cấu trúc, như ví dụ hạt phân tử lượng cao, bậc cao hơn hoặc VLP hoàn chỉnh. Phân đoạn riêng biệt còn có thể được xử lý để phân lập, tinh chế, cô đặc hoặc tổ hợp của chúng, các protein, protein siêu cấu trúc hoặc hạt bậc cao hơn bằng, ví dụ, các bước ly tâm bổ sung, kết tủa, bước sắc ký (ví dụ sắc ký loại cỡ, trao đổi ion, ái lực), lọc dòng tiếp tuyến hoặc tổ hợp của chúng. Sự có mặt của các protein được tinh chế, protein siêu cấu trúc hoặc hạt bậc cao như VLP, có thể được xác nhận bởi, ví dụ, phân tích Western nguyên thể hoặc SDS-PAGE bằng cách sử dụng kháng thể phát hiện thích hợp, điện di mao dẫn, kính hiển vi điện tử hoặc các phương pháp khác bất kỳ như sẽ là rõ ràng đối với người có hiểu biết trung bình về lĩnh vực này.

Fig.4A thể hiện ví dụ về profin rửa giải của phép phân tích sắc ký loại cỡ của chiết phẩm thực vật chứa PVLP. Trong trường hợp này, các VLP bao gồm enterovirut EV71 capsit, rửa giải trong các phân đoạn 9 đến khoảng 14, đỉnh trong phân đoạn 12.

Các VLP có thể được tinh chế hoặc được chiết bằng cách sử dụng phương pháp thích hợp bất kỳ ví dụ chiết hoá học hoặc hoá sinh. Các VLP có thể tương đối nhạy đối với việc làm khô, nhiệt, độ pH, chất hoạt động bề mặt và chất tẩy rửa. Do đó, nó có thể là hữu ích để sử dụng các phương pháp mà làm tối đa hoá sản lượng, giảm thiểu sự nhiễm bẩn của phân đoạn VLP với các protein tế bào, duy trì tính nguyên vẹn của protein hoặc VLP và các phương pháp nói lỏng vách tế bào để giải phóng protein hoặc VLP. Ví dụ, các phương pháp mà tạo ra thể nguyên sinh và/hoặc tế bào trần có thể được sử dụng (xem ví

dụ WO 2011/035422) để thu được VLP như được mô tả trong bản mô tả này. Việc làm giảm thiểu hoặc loại trừ việc sử dụng chất tẩy rửa hoặc chất hoạt động bề mặt như ví dụ SDS hoặc Triton X-100 có thể là có lợi để cải thiện sản lượng của việc chiết VLP. Sau đó, các VLP có thể được đánh giá đối với cấu trúc và kích cỡ bằng, ví dụ, kính hiển vi điện tử hoặc sắc ký loại cỡ như được kể đến trên đây và được cho vào siêu ly tâm phân tích.

Kích cỡ (nghĩa là đường kính) của các PVLP được xác định trên đây, có thể được đo ví dụ bằng cách tán sắc ánh sáng động học (DLS) hoặc kỹ thuật kính hiển vi điện tử (EM), thường nằm trong khoảng từ 20 đến 50nm, hoặc kích cỡ bất kỳ giữa chúng. Ví dụ, kích cỡ của cấu trúc PVLP nguyên vẹn có thể nằm trong khoảng từ 25nm đến 35nm, hoặc kích cỡ bất kỳ giữa chúng hoặc từ 20nm, 21nm, 22nm, 23nm, 24nm, 25nm, 26nm, 27 nm, 28nm, 29nm, 30nm, 31nm, 32nm, 33nm, 34nm, 35nm, 36nm, 37nm, 38nm, 39nm, 40nm, 41nm, 42nm, 43nm, 44nm, 45nm, 46nm, 47nm, 48nm, 49nm, 50nm, hoặc kích cỡ bất kỳ giữa chúng.

PVLP có thể được tổng hợp ở lượng đến 2g/kg khối lượng tươi của thực vật, tương ứng với khoảng 40% tổng hàm lượng protein của thực vật. Ví dụ, như được mô tả trong bản mô tả này, lượng VLP tổng hợp có thể nằm trong khoảng từ 10mg đến 2,0g/kg khối lượng tươi hoặc lượng bất kỳ giữa chúng, như 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 750, 775, 800, 825, 850, 875, 900, 925, 950, 975, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900 hoặc 2000mg/kg khối lượng tươi hoặc lượng bất kỳ giữa chúng.

Vật chủ

Một hoặc nhiều hơn một cấu trúc gen của sáng chế có thể được biểu hiện trong vật chủ thực vật thích hợp bất kỳ hoặc bộ phận của thực vật, ví dụ, một hoặc nhiều lá, thân cùng với một hoặc nhiều lá hoặc rễ, mà được biến nạp bởi trình tự nucleotit hoặc các cấu trúc hoặc vectơ của sáng chế. Các ví dụ về vật chủ thích hợp bao gồm, nhưng không chỉ giới hạn ở, cây trồng trong nông nghiệp bao gồm cỏ linh lăng, canola, *Brassica* spp., ngô, *Nicotiana* spp., khoai tây, nhân sâm, đậu Hà Lan, yến mạch, lúa gạo, đậu tương, lúa mỳ, lúa mạch, hướng dương, bông và tương tự.

Một hoặc nhiều cấu trúc gen của sáng chế còn có thể bao gồm vùng không dịch mã 3'. Vùng không dịch mã 3' dùng để chỉ phần của gen bao gồm đoạn ADN mà chứa tín hiệu polyadenyl hoá và tín hiệu điều hoà khác bất kỳ có khả năng tác động đến việc xử lý ARN thông tin hoặc biểu hiện gen. Tín hiệu polyadenyl hoá thường được đặc trưng bằng cách tác động việc bổ sung của axit polyadenylic theo vết đến đầu 3' của tiền chất ARN thông tin. Các tín hiệu polyadenyl hoá được nhận biết một cách thông thường bởi sự có mặt của tính tương đồng với dạng thích hợp 5' AATAAA-3' mặc dù các biến thể là không phổ biến. Các ví dụ không hạn chế về các vùng 3' thích hợp là các vùng không dịch mã được sao chép 3' chứa tín hiệu polyadenyl hoá của khối u *Agrobacterium* cảm ứng gen (Ti) plasmid, như gen nopal synthase (NOS), gen thực vật như gen protein lưu trữ đậu tương cấu trúc dưới đơn vị nhỏ của gen ribulose-1, 5-bisphosphat carboxylase (ssRUBISCO; bằng Mỹ số US 4,962,028), đoạn khởi đầu được sử dụng trong việc điều hoà sự biểu hiện plastocyanin, được mô tả trong bằng Mỹ số US 7,125,978.

Một hoặc nhiều trong số các cấu trúc di truyền của sáng chế cũng có thể bao gồm các đoạn tăng cường khác, các đoạn tăng cường dịch mã hoặc sao chép, như có thể được đòi hỏi. Các đoạn tăng cường có thể được định vị 5' hoặc 3' với trình tự được sao chép. Các vùng tăng cường cũng đã biết đối với người có hiểu biết trung bình về lĩnh vực này và có thể bao gồm codon khởi đầu ATG, trình tự liên kết hoặc tương tự. Codon khởi đầu, nếu có, có thể trong pha với khung đọc ("trong khung") của trình tự mã hoá để tạo ra sự dịch mã đúng của trình tự sao chép.

Các cấu trúc của sáng chế có thể được đưa vào tế bào thực vật bằng cách sử dụng Ti plasmid, Ri plasmid, vectơ virus thực vật, sự biến nạp ADN trực tiếp, vi tiêm, điện di, v.v. Để xem xét về các kỹ thuật này, xem ví dụ ấn phẩm: Weissbach and Weissbach, *Methods for Plant Molecular Biology*, Academy Press, New York VIII, pp. 421-463 (1988); Geierstein và Corey, *Plant Molecular Biology*, 2d Ed. (1988); và Miki and Iyer, *Fundamentals of Gene Transfer in Plant*. In *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 việc thu nhận ADN trực tiếp, việc sử dụng liposom, điện di, ví dụ sử dụng thể nguyên sinh, vi tiêm, vi bắn hoặc sợi tinh thể và thâm nhiễm chân không. Xem, ví dụ ấn phẩm: 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 and Weissbach, eds., Academic Press Inc., 1988), Methods in Plant Molecular Biology (Schuler and Zielinski, eds., Academic Press Inc., 1989), Liu and Lomonosoff (J Virol Meth, 105:343-348, 2002.), các bằng Mỹ số 4,945,050; 5,036,006; và 5,100,792, đơn yêu cầu cấp bằng Mỹ số 08/438,666, được nộp ngày 10 tháng 5 năm 1995, và 07/951,715, được nộp ngày 25 tháng 9 năm 1992.

Biểu hiện tạm thời

Không muốn bị ràng buộc bởi lý thuyết, nồng độ protein và tỷ lệ của protein cấu trúc picornavirut khác nhau, polyprotein picornavirut và/hoặc proteaza có thể là quan trọng đối với hiệu quả tập hợp của PVLP. Do đó, số bội và thời gian lây truyền, có thể là quan trọng để thao tác nồng độ protein và toàn bộ hiệu quả tập hợp của các VLP ở thực vật.

Cấu trúc của sáng chế có thể được biểu hiện một cách tạm thời ở thực vật, bộ phận của thực vật, hoặc tế bào thực vật. Hệ biểu hiện tạm thời dựa vào sự biểu hiện nhiễm sắc thể của polyprotein tái tổ hợp được đưa vào, qua việc thâm nhiễm *Agrobacterium tumefaciens*, vào thực vật, bộ phận của thực vật, hoặc tế bào thực vật có thể được sử dụng để biểu hiện protein cấu trúc picornavirut, polyprotein picornavirut và/hoặc proteaza, được tạo đích vào các khoang tế bào khác nhau hoặc các khoang phụ. Hệ biểu hiện tạm thời cho phép đối với tốc độ sản xuất ở mức cao. Hơn nữa, lượng lớn protein có thể đạt được trong vài ngày sau khi thâm nhiễm *Agrobacterium* tái tổ hợp ở thực vật (Rybicki, 2010; Fischer et al., 1999). Cũng có thể biểu hiện trình tự gen dài và có nhiều hơn một gen biểu hiện một cách đồng thời trong cùng tế bào, cho phép việc tập hợp hiệu quả của nhiều protein (Lombardi et al., 2009).

Tuy nhiên, trong quá trình làm câm gen sau dịch mã biểu hiện tạm thời có thể làm giới hạn sự biểu hiện của các protein khác loại ở thực vật. Việc biểu hiện đồng thời của các đoạn ngăn chặn câm, ví dụ, nhưng không chỉ giới hạn ở Nss từ virus héo đốm cà chua có thể được sử dụng để kháng lại sự thoái biến đặc hiệu của các ARN thông tin chuyển gen (Brigneti et al., 1998). Đoạn ngăn chặn khác nhau về việc làm câm là đã được biết đến

trong lĩnh vực này và có thể được sử dụng như được mô tả trong bản mô tả này (Chiba et al., 2006, Virology 346:7-14, ví dụ nhưng không chỉ giới hạn ở HcPro, TEV -p1/HC-Pro (virut khắc thuốc lá-p1/HC-Pro), BYV -p21, p19 của virut còi rậm cà chua (TBSV p19), capsit protein của virut quăn cà chua (TCV-CP), 2b của virut khảm dưa chuột; CMV-2b), p25 của virut khoai tây X (PVX-p25), p11 của virut khoai tây M (PVM-p11), p11 của virut khoai tây S (PVS-p11), p16 của virut gây cháy cây việt quất, (BScV-p16), p23 của virut tristexa giống cam quýt (CTV-p23), p24 của virut kết hợp cuộn lá nho-2, (GLRaV-2 p24), p10 của virut nho A, (GVA-p10), p14 của virut nho B (GVB-p14), p10 của virut tiềm năng Heracleum (HLV-p10), hoặc p16 của virut tiềm năng chung của tỏi (GCLV-p16). Do đó, đoạn ngăn chặn việc câm, ví dụ 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-p16or GVA-p10, có thể được biểu hiện đồng thời cùng với một hoặc nhiều protein cấu trúc picornavirut, polyprotein picornavirut và/hoặc proteaza để bảo đảm tiếp mức sản xuất protein ở mức cao trong thực vật, bộ phận của thực vật hoặc tế bào thực vật.

Sáng chế cũng đề xuất phương pháp như được mô tả trên đây, trong đó trình tự nucleotit (thứ ba) bổ sung được biểu hiện trong thực vật, trình tự nucleotit (thứ ba) bổ sung mã hoá đoạn ngăn chặn câm được liên kết theo cách hoạt động được với vùng điều hoà (thứ ba) bổ sung mà hoạt động ở thực vật. Trình tự nucleotit mã hoá đoạn ngăn chặn câm có thể là, ví dụ Nss, 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.

Như được mô tả dưới đây, các phương pháp biểu hiện tạm thời có thể được sử dụng để biểu hiện các cấu trúc của sáng chế (xem Liu and Lomonosoff, 2002, Journal of Virological Methods, 105:343-348). Theo cách khác, phương pháp biểu hiện tạm thời dựa trên chân không, như được mô tả bởi Kapila et al., 1997 có thể được sử dụng. Các phương pháp này có thể bao gồm, ví dụ, nhưng không chỉ giới hạn ở phương pháp ghép trong nông nghiệp hoặc thẩm nhiễm trong nông nghiệp, thẩm nhiễm ống tiêm, tuy nhiên, các phương pháp tạm thời khác cũng có thể được sử dụng như được lưu ý trên đây. Với việc ủ trong nông nghiệp, thẩm nhiễm trong nông nghiệp hoặc thẩm nhiễm ống tiêm, hỗn hợp gồm *Agrobacteria* bao gồm axit nucleic mong muốn đi vào khoảng trống nội bào của mô,

ví dụ là, phần trên mặt đất của thực vật (bao gồm thân, lá và hoa), các phần khác của thực vật (thân, rễ, hoa) hoặc toàn bộ thực vật. Sau khi lai biểu bì *Agrobacteria* lây nhiễm và truyền các bản sao t-ADN vào tế bào. t-ADN được sao chép theo cách bổ sung và ARN thông tin được dịch mã, dẫn đến việc tạo ra protein quan tâm ở tế bào được lây nhiễm, tuy nhiên, việc đi qua của t-ADN bên trong nhân là tạm thời.

Ngoài ra, phần được xét đến của sáng chế là thực vật chuyển gen, tế bào thực vật hoặc hạt chứa axit nucleic hoặc một hoặc nhiều hơn một cấu trúc gen của sáng chế. Phương pháp tạo ra toàn bộ thực vật từ tế bào thực vật cũng đã được biết đến trong lĩnh vực này. Nói chung, tế bào thực vật được biến nạp được nuôi cấy trong môi trường thích hợp, mà có thể chứa các chất chọn lọc như thuốc kháng sinh, mà ở đó các chất đánh dấu có thể chọn lọc được sử dụng để tạo thuận lợi cho việc nhận dạng tế bào thực vật được biến nạp một cách ổn định. Để hỗ trợ cho việc nhận dạng tế bào thực vật được biến nạp một cách ổn định, thì các cấu trúc của sáng chế còn có thể được thao tác để bao gồm các chất đánh dấu có thể chọn lọc của thực vật. Các chất đánh dấu có thể chọn lọc hữu ích bao gồm các enzym mà tạo ra sức kháng đối với các chất hoá học như thuốc kháng sinh ví dụ, gentamycin, hygromycin, kanamycin, hoặc thuốc diệt cỏ như phosphinothryxin, glyphosat, closulfuron, và tương tự. Tương tự, các enzym tạo ra việc sản xuất hợp chất có thể nhận dạng bằng việc thay đổi màu sắc như GUS (beta-glucuronidaza), hoặc huỳnh quang, như luxiferaza hoặc GFP, có thể được sử dụng. Khi thể sản tạo ra, sự hình thành rễ có thể được khuyến khích bằng việc dùng các hoóc môn thực vật thích hợp theo các tatg gactaactat gactccagct 16 và chồi non được chuyển vào môi trường tạo rễ để tạo ra thực vật. Sau đó, các thực vật có thể được sử dụng để thiết lập các thể hệ lặp lại, từ hạt hoặc bằng cách sử dụng các kỹ thuật gây giống sinh dưỡng. Các thực vật chuyển gen cũng có thể được tạo ra mà không cần phải sử dụng môi trường nuôi cấy mô.

Yếu tố khuếch đại

Tỷ lệ của polyprotein với proteaza có thể được thay đổi ví dụ bằng cách sử dụng các yếu tố điều hoà khác nhau hoặc tổ hợp của các nhân tố điều hoà, trong trình tự axit nucleic được sử dụng để điều khiển sự biểu hiện của polyprotein và proteaza. Ví dụ, tập hợp thứ nhất hoặc tổ hợp của các nhân tố điều hoà có thể được sử dụng để điều hoà sự sao chép, phiên mã hoặc tổ hợp của chúng, của axit nucleic thứ nhất và tập hợp thứ hai hoặc tổ hợp của các nhân tố điều hoà có thể được sử dụng để điều hoà sự sao chép, phiên mã hoặc

tổ hợp của chúng, của axit nucleic thứ hai sao cho sự khác nhau về sự biểu hiện của axit nucleic thứ nhất và thứ hai đạt được nhờ đó điều biến tỷ lệ của polyprotein:proteasea *in vivo*. Ví dụ, mà không được xét đến làm giới hạn tập hợp thứ nhất hoặc tổ hợp của các nhân tố điều hoà có thể bao gồm yếu tố khuếch đại, ví dụ, yếu tố thu được từ BeYDV, trong khi yếu tố khuếch đại có thể là không có mặt trong tập hợp thứ hai hoặc tổ hợp của các nhân tố điều hoà. Theo cách khác, tập hợp thứ hai có thể bao gồm yếu tố khuếch đại, ví dụ, yếu tố thu được từ BeYDV, trong khi yếu tố khuếch đại có thể là không có mặt trong tập hợp thứ nhất hoặc tổ hợp của các nhân tố điều hoà. Thuật ngữ “cát xét biểu hiện” dùng để chỉ trình tự nucleotit bao gồm axit nucleic quan tâm trong sự kiểm soát của và được liên kết theo cách hoạt động được (hoặc theo cách thao tác) với, đoạn khởi đầu thích hợp hoặc các yếu tố điều hoà khác đối với việc sao chép của axit nucleic quan tâm trong tế bào vật chủ.

Hệ biểu hiện như được mô tả trong bản mô tả này có thể bao gồm cát xét biểu hiện dựa trên virus hai rãnh hoặc virus với hệ gen hai rãnh. Ví dụ, các virus hai rãnh có thể thuộc họ Comoviridae. Chi của họ Comoviridae bao gồm Comovirus, Nepovirus, Fabavirus, Chervirus và Sadwavirus. Comovirus bao gồm virus khảm cây đậu đũa (CPMV), virus khảm nghiêm trọng cây đậu đũa (CPSMV), virus khảm cây bí (SqMV), virus vằn cỏ ba lá đỏ (RCMV), virus vằn vỏ đậu (BPMV), virus gây bệnh đốm củ cải (TuRSV), virus khảm cây đậu tằm (BBtMV), virus nhuộm đậu tằm (BBSV), virus khảm củ cải (RaMV). Các ví dụ về các trình tự comovirus RNA-2 bao gồm các yếu tố tăng cường mà có thể là hữu ích đối với các khía cạnh khác nhau của sáng chế bao gồm, nhưng không chỉ giới hạn ở: CPMV RNA-2 (Số lưu trữ GenBank NC_003550), RCMV RNA-2 (Số lưu trữ GenBank NC_003738), BPMV RNA-2 (Số lưu trữ GenBank NC_003495), CPSMV RNA-2 (Số lưu trữ GenBank NC_003544), SqMV RNA-2 (Số lưu trữ GenBank NC_003800), TuRSV RNA-2 (Số lưu trữ GenBank NC_013219.1), BBtMV RNA-2 (Số lưu trữ GenBank GU810904), BBSV RNA2 (Số lưu trữ GenBank FJ028650), RaMV (Số lưu trữ GenBank NC_003800).

Các đoạn của hệ gen ARN comovirus hai rãnh được dùng để chỉ RNA-1 và RNA-2. RNA-1 mã hoá protein có tham gia vào việc sao chép trong khi RNA-2 mã hoá protein cần thiết đối với sự di chuyển tế bào-tế bào và hai capsit proteins. Cát xét được tạo thành chủ yếu từ comovirus thích hợp bất kỳ có thể được sử dụng bao gồm CPMV, CPSMV,

SqMV, RCMV, hoặc BPMV, ví dụ, cát xét biểu hiện có thể được tạo thành chủ yếu từ CPMV.

Các hệ biểu hiện cũng có thể bao gồm các yếu tố khuếch đại từ geminivirut chẳng hạn, yếu tố khuếch đại từ virut còi vàng đậu (BeYDV). BeYDV thuộc chi Mastrevirut được làm thích ứng với thực vật có hai lá mầm. BeYDV là một rãnh có hệ gen ADN tròn sợi đơn và có thể sao chép thành số lượng bản sao rất nhiều bằng cơ chế tuần hoàn trực lẫn. Các hệ vectơ ADN replicon thu được từ BeYDV đã được sử dụng để tạo ra protein với sản lượng cao nhanh chóng ở thực vật.

Như được sử dụng trong bản mô tả này, cụm từ “yếu tố khuếch đại” dùng để chỉ đoạn axit nucleic bao gồm ít nhất một phần trong số một hoặc nhiều vùng trong gen dài (LIR) của hệ gen geminivirut. Như được sử dụng trong bản mô tả này, “vùng trong gen dài” dùng để chỉ vùng của vùng trong gen dài mà chứa vị trí gắn kết rep có khả năng làm trung gian cắt và sao chép bởi protein geminivirut Rep. Theo một số khía cạnh, đoạn axit nucleic bao gồm một hoặc nhiều LIR, còn có thể bao gồm vùng trong gen ngắn (SIR) của hệ gen geminivirut. Như được sử dụng trong bản mô tả này, “vùng trong gen ngắn” dùng để chỉ sợi bù (IR ngắn (SIR) của Mastrevirut). Yếu tố khuếch đại thu được từ geminivirut thích hợp bất kỳ có thể được sử dụng trong bản mô tả này. 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).

Yếu tố điều hoà

Sáng chế còn đề cập đến cấu trúc gen bao gồm axit nucleic mã hoá polyprotein, như một hoặc nhiều protein picornavirut hoặc proteaza, ví dụ nhưng không chỉ giới hạn ở picornavirut proteaza, như được mô tả trên đây, liên kết theo cách hoạt động được với yếu tố điều hoà mà hoạt động ở thực vật.

Việc sử dụng các thuật ngữ "vùng điều hoà", "yếu tố điều hoà" hoặc "đoạn khởi đầu" trong sáng chế có nghĩa đề cập đến phần axit nucleic thông thường, nhưng không luôn luôn, ngược dòng của vùng mã hoá protein của gen, mà có thể gồm có ADN hoặc ARN hoặc cả ADN và ARN. Khi vùng điều hoà hoạt động và trong sự kết hợp hoạt động hoặc được liên kết theo cách hoạt động được, với gen quan tâm, điều này có thể dẫn đến

sự biểu hiện của gen quan tâm. Yếu tố điều hoà có thể có khả năng làm trung gian tính đặc hiệu cơ quan hoặc kiểm soát sự phát triển hoặc sự hoạt hoá gen theo thời gian. "Vùng điều hoà" có thể bao gồm các yếu tố khởi đầu, yếu tố khởi đầu lõi thể hiện hoạt tính khởi đầu cơ bản, các yếu tố mà có khả năng gây ra đáp ứng với tác nhân kích thích bên ngoài, các yếu tố mà làm trung gian hoạt tính khởi đầu như các yếu tố điều hoà âm tính hoặc các yếu tố tăng cường sao chép. "Vùng điều hoà", như được sử dụng trong bản mô tả này, cũng có thể bao gồm các yếu tố mà hoạt động sau sao chép, ví dụ, các yếu tố điều hoà mà điều biến sự biểu hiện gen như các đoạn tăng cường dịch mã và phiên mã, các đoạn ngăn chặn dịch mã và phiên mã, trình tự hoạt hoá ngược dòng và yếu tố quyết định tính không ổn định ARN thông tin. Vài trong số các yếu tố sau này có thể được định vị gần với vùng mã hoá.

Các ví dụ về các yếu tố điều hoà hoạt động trong tế bào thực vật và mà có thể được sử dụng theo sáng chế bao gồm nhưng không chỉ giới hạn ở vùng điều hoà plastocyanin (US 7,125,978) hoặc vùng điều hoà của Ribuloza 1,5-bisphosphat carboxylaza/oxygenaza (RuBisCO; US 4,962,028), protein gắn kết clorofyl a/b (CAB; Leutwiler et al; 1986), ST-LS1 (kết hợp với phức hệ phát ra oxy của hệ quang hợp II và được mô tả bởi Stockhaus et al. 1987, 1989).

Trong bản mô tả này, thuật ngữ "yếu tố điều hoà" hoặc "vùng điều hoà" thường dùng để chỉ trình tự ADN, thông thường, nhưng không luôn luôn, ngược dòng (5') với trình tự mã hoá của gen cấu trúc, mà kiểm soát sự biểu hiện của vùng mã hoá bằng cách tạo ra sự nhận biết đối với ARN polymeraza và/hoặc các nhân tố khác được đòi hỏi cho việc phiên mã để bắt đầu ở vị trí cụ thể. Tuy nhiên, sẽ được hiểu rằng các trình tự nucleotit khác, được định vị trong intron hoặc 3' của trình tự cũng có thể góp phần vào việc điều hoà sự biểu hiện của vùng mã hoá quan tâm. Ví dụ về yếu tố điều hoà mà tạo ra để nhận biết đối với ARN polymeraza hoặc các nhân tố phiên mã khác để bảo đảm sự khởi đầu ở vị trí cụ thể là nhân tố khởi đầu. Hầu hết, nhưng không phải tất cả, các nhân tố khởi đầu có nhân thật chứa hộp TATA, trình tự axit nucleic được bảo tồn gồm có adenosin và cặp bazơ thymidin nucleotit thường được đặt gần 25 cặp bazơ ngược dòng của vị trí khởi đầu phiên mã. Yếu tố khởi đầu bao gồm yếu tố khởi đầu cơ bản, chịu trách nhiệm đối với việc khởi đầu phiên mã, cũng như các yếu tố điều hoà khác (như được liệt kê trên đây) mà cải biến sự biểu hiện của gen.

Có vài kiểu vùng điều hoà, bao gồm các kiểu mà được điều hoà theo cách phát triển, cảm ứng hoặc cấu thành. Vùng điều hoà mà được điều hoà theo cách phát triển hoặc kiểm soát sự biểu hiện khác nhau của gen trong sự kiểm soát của nó, được hoạt hoá trong một số cơ quan hoặc mô của cơ quan ở thời gian cụ thể trong quá trình phát triển của cơ quan hoặc mô đó. Tuy nhiên, một số vùng điều hoà mà được điều hoà theo cách phát triển có thể hoạt động theo cách ưu tiên trong một số cơ quan hoặc mô ở giai đoạn phát triển cụ thể, chúng cũng có thể hoạt động theo cách được điều hoà theo cách phát triển hoặc ở mức cơ bản trong các cơ quan hoặc mô khác trong thực vật. Các ví dụ về vùng điều hoà đặc hiệu mô, ví dụ xem vùng điều hoà đặc hiệu, bao gồm đoạn khởi đầu napin và đoạn khởi đầu cruciferin (Rask et al., 1998, *J. Plant Physiol.* 152: 595-599; Bilodeau et al., 1994, *Plant cell* 14: 125-130). Ví dụ về đoạn khởi đầu đặc hiệu lá bao gồm đoạn khởi đầu plastoxyanin (xem bằng Mỹ số US 7,125,978).

Vùng điều hoà có thể cảm biến là vùng mà có khả năng hoạt hoá sự phiên mã một cách trực tiếp hoặc gián tiếp của một hoặc nhiều trình tự ADN hoặc gen trong sự đáp ứng với nhân tố cảm ứng. Với sự không có mặt của nhân tố cảm ứng, trình tự ADN hoặc gen sẽ không được phiên mã. Thông thường, nhân tố protein mà gắn kết một cách đặc hiệu với vùng điều hoà có thể cảm ứng để hoạt hoá sự phiên mã có thể có mặt ở dạng bất hoạt, mà sau đó được chuyển hoá một cách trực tiếp hoặc gián tiếp thành dạng hoạt hoá bởi nhân tố cảm ứng. Tuy nhiên, nhân tố protein cũng có thể không có mặt. Nhân tố cảm ứng có thể là chất hoá học như protein, chất chuyển hoá, chất điều hoà sự sinh trưởng, thuốc diệt cỏ hoặc hợp chất phenola hoặc ứng suất sinh lý bị áp đặt một cách trực tiếp bởi nhiệt nóng, lạnh, muối hoặc yếu tố độc tố hoặc một cách gián tiếp thông qua tác dụng của mầm bệnh hoặc chất gây bệnh như a virus. Tế bào thực vật chứa vùng điều hoà có thể cảm ứng có thể được phơi trần với tác nhân cảm ứng bằng cách áp dụng theo cách bên ngoài tác nhân cảm ứng cho tế bào hoặc thực vật như bằng cách phun, tưới nước, gia nhiệt hoặc các phương pháp tương tự. Các yếu tố điều hoà có thể cảm ứng có thể thu được từ gen thực vật hoặc không phải thực vật (ví dụ Gatz, C. and Lenk, L.R.P., 1998, *Trends Plant Sci.* 3, 352-358). Các ví dụ về đoạn khởi đầu có thể cảm ứng hiệu lực bao gồm, nhưng không chỉ giới hạn ở, đoạn khởi đầu có thể cảm ứng tetracyclin (Gatz, C., 1997, *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 48, 89-108), đoạn khởi đầu có thể cảm ứng steroid (Aoyama, T. and Chua, N.H., 1997, *Plant* 1, 2, 397-404) và đoạn khởi đầu có thể cảm ứng ethanol (Salter, M.G., et

aI, 1998, *Plant Journal* 16, 127-132; Caddick, M.X., et al, 1998, *Nature Biotech.* 16, 177-180) các gen IB6 và CKI 1 có thể cảm ứng xytokinin (Brandstatter, I. and Kieber, I., 1998, *Plant cell* 10, 1009-1019; Kakimoto, T., 1996, *Science* 274, 982-985) và yếu tố có thể cảm ứng auxin, DR5 (Ulmasov, T., et al., 1997, *Plant cell* 9, 1963-1971).

Vùng điều hoà cơ bản chỉ dẫn sự biểu hiện của gen thông qua các phần khác nhau của thực vật và theo cách liên tục thông qua sự phát triển của thực vật. Các ví dụ về các yếu tố điều hoà cơ bản bao gồm các đoạn khởi đầu kết hợp với bản sao CaMV 35S (Odell et al., 1985, *Nature*, 313: 810-812), actin 1 của lúa gạo (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 gen 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), gen *Arabidopsis* ubiquitin 1 và 6 (Holtorf et al, 1995, *Plant Mol. Biol.* 29: 637-646), và gen nhân tố khởi đầu phiên mã của thuốc lá 4A (Mandel et al, 1995, *Plant Mol. Biol.* 29: 995-1004).

Thuật ngữ "cơ bản" như được sử dụng trong bản mô tả này không nhất thiết cho biết rằng gen dưới sự kiểm soát của vùng điều hoà cơ bản được biểu hiện ở cùng mức trong tất cả các kiểu tế bào, nhưng mà gen được biểu hiện trong khoảng rộng của các kiểu tế bào thậm chí mặc dù sự biến đổi đa dạng thường được quan sát. Các yếu tố điều hoà cơ bản có thể được ghép cặp với các trình tự khác để làm tăng cường hơn nữa sự phiên mã và/hoặc dịch mã của trình tự nucleotit mà chúng được liên kết theo cách hoạt động được. Ví dụ, hệ CPMV-HT thu được từ các vùng không dịch mã của virus khảm cây đậu đũa (CPMV) và chứng minh việc dịch mã được tăng cường của trình tự mã hoá kết hợp. Thuật ngữ "nguyên thể" có nghĩa rằng trình tự axit nucleic hoặc axit amin là xuất hiện tự nhiên hoặc "kiểu hoang dại". Thuật ngữ "được liên kết theo cách hoạt động được" có nghĩa rằng trình tự cụ thể, ví dụ yếu tố điều hoà và vùng mã hoá quan tâm, tương tác một cách trực tiếp hoặc gián tiếp để thực hiện chức năng được dự định như việc làm trung gian hoặc điều biến sự biểu hiện của gen. Sự tương tác của các trình tự được liên kết theo cách hoạt động được có thể, ví dụ, được qua trung gian bởi protein mà tương tác với các trình tự được liên kết theo cách hoạt động được.

Tỷ lệ của polyprotein với proteaza còn có thể được thay đổi ví dụ bằng cách sử dụng các yếu tố điều hoà, yếu tố khuếch đại và/hoặc yếu tố tăng cường. Ví dụ axit nucleic thứ nhất có thể bao gồm các yếu tố điều hoà, yếu tố khuếch đại và/hoặc yếu tố tăng cường.

Axit nucleic thứ hai có thể hoặc có thể không bao gồm cùng tổ hợp của các yếu tố điều hoà, yếu tố khuếch đại và/hoặc yếu tố tăng cường.

Ví dụ, các đoạn khởi đầu khác nhau có thể được sử dụng để điều khiển sự biểu hiện khác nhau giữa polyprotein so với proteaza *in vivo*. Ví dụ, tập hợp thứ nhất hoặc tổ hợp của các nhân tố điều hoà có thể bao gồm đoạn khởi đầu có thể cảm biến, trong khi đoạn khởi đầu trong tập hợp thứ hai hoặc tổ hợp của các nhân tố điều hoà có thể là cơ bản hoặc tập hợp thứ hai hoặc tổ hợp của các nhân tố điều hoà có thể bao gồm đoạn khởi đầu có thể cảm ứng, trong khi đoạn khởi đầu trong tập hợp thứ nhất hoặc tổ hợp của các nhân tố điều hoà có thể là cơ bản. Cường độ của đoạn khởi đầu cũng có thể khác nhau giữa tập hợp thứ nhất và thứ hai hoặc tổ hợp của các nhân tố điều hoà, sao cho sự biểu hiện khác nhau giữa polyprotein với proteaza đạt được *in vivo*.

Ví dụ thực hiện sáng chế

Sáng chế sẽ được minh hoạ chi tiết hơn trong các ví dụ sau đây.

Ví dụ 1 - Sự biểu hiện EV71

Tổng hợp gen

Các đoạn ADN mã hoá cấu trúc EV71 protein P1 và proteaza 3CD được sử dụng. Các trình tự ứng viên đối với P1 và 3CD là có sẵn trong GenBank. Các ví dụ không hạn chế về các trình tự này là:

- Đối với P1 trình tự aa: các trình tự axit amin GenBank ID ADG57603 (các axit amin 1-862) (SEQ ID NO:5); trình tự nucleotit: GenBank ID GQ279369 (nucleotit 743-3328) (SEQ ID NO:6);
- Đối với 3CD (chủng HK08): trình tự axit amin: GenBank ID ADG57603 (các axit amin 1549-2193) (SEQ ID NO: 1); trình tự nucleotit: GenBank ID GQ279369 (nucleotit 5386-7321) (SEQ ID NO:2);
- Đối với 3CD (chủng GDFS08): trình tự axit amin GenBank ID ACI25378 (các axit amin 1549-2193)(SEQ ID NO: 3); trình tự nucleotit: GenBank ID FJ194964 (nucleotit 5387-7321) (SEQ ID NO: 4).

Hai gen P1 được tổng hợp. Thứ nhất được tạo ra bằng cách sử dụng trình tự kiểu hoang dại trong khi thứ hai được tạo ra chủ yếu từ trình tự tối ưu (sử dụng codon của

người) được xác định bằng cách sử dụng các phương pháp tiêu chuẩn như đã được biết đến trong lĩnh vực này. Hai gen 3CD được tổng hợp dựa trên các trình tự kiểu hoang dại của chúng. 3 gen kiểu hoang dại được tổng hợp bằng Invitrogen™ (trước đây là GeneArt®) và gen P1 tối ưu được tối ưu hóa và được tổng hợp bằng DNA2.0.

Tạo dòng phân tử

Các gen được tổng hợp được tạo dòng vào vector biểu hiện thực vật. Các thành phần vector được chọn bao gồm các yếu tố điều hoà phiên mã và dịch mã từ virus khảm cây đậu đũa (CPMV) hoặc gen plastoxyanin cỏ linh lăng. Cả hai yếu tố điều hoà đã được sử dụng với sự thành công trong diễn thuyết của chúng tôi đối với sự biểu hiện ở mức cao của các protein tái tổ hợp. Yếu tố khuếch đại ADN từ geminivirus lùn vàng đậu (BeYDV) là dấu hiệu khác mà có thể được tích hợp vào các vector biểu hiện của thực vật. Nó đã dẫn đến làm tăng rất lớn về sự biểu hiện protein đối với một số ứng viên. Do đó, chúng tôi đã tạo dòng mỗi cấu trúc gen vào vector biểu hiện có hoặc không có yếu tố khuếch đại ADN. Bảng 1 thể hiện các cát xét biểu hiện thực vật được tập hợp đối với mục đích.

Bảng 1. Các cát xét biểu hiện thực vật được tập hợp đối với sự biểu hiện của EV71 polyprotein cấu trúc P1 và proteaza 3CD ở *N. benthamiana*.

Vùng mã hoá	Yếu tố điều hoà	Yếu tố khuếch đại ADN	Số vector
P1 (Wt HK08)	CPMV HT	-	1300
P1 (Wt HK08)	CPMV HT	BeYDV+rep	1301
P1 (Wt HK08)	Plastoxyanin	-	1302
P1 (Wt HK08)	Plastoxyanin	BeYDV+rep	1303
P1 (Opt HK08)	CPMV HT	-	1305
P1 (Opt HK08)	CPMV HT	BeYDV+rep	1306
P1 (Opt HK08)	Plastoxyanin	-	1307
P1 (Opt HK08)	Plastoxyanin	BeYDV+rep	1308
3CD (Wt HK08)	CPMV HT	-	1310
3CD (Wt HK08)	CPMV HT	BeYDV+rep	1311
3CD (Wt HK08)	Plastoxyanin	-	1312
3CD (Wt HK08)	Plastoxyanin	BeYDV+rep	1313
3CD (Wt GDFS08)	CPMV HT	-	1315
3CD (Wt GDFS08)	CPMV HT	BeYDV+rep	1316
3CD (Wt GDFS08)	Plastoxyanin	-	1317
3CD (Wt GDFS08)	Plastoxyanin	BeYDV+rep	1318

Phân tích sự biểu hiện – Chọn lọc các cấu trúc gen tái tổ hợp tốt nhất

Mỗi cát xét biểu hiện được tách dòng vào vector plasmit mà sau đó được chuyển vào *Agrobacterium tumefaciens*. Sự biểu hiện tạm thời được khởi đầu bằng việc thâm nhiễm chân không của chất tiêm chủng *Agrobacterium* chuyển gen mà dẫn đến chuyển các bản sao ADN di động của các cấu trúc ADN vào tế bào thực vật. Sự biểu hiện tạm thời của nhiều thành phần (biểu hiện đồng thời) được thực hiện bằng việc thâm nhiễm hỗn hợp gồm chất tiêm chủng *Agrobacterium* (đồng thâm nhiễm). Do một thành phần được đưa vào thực vật là cấu trúc (P1) và chất của thành phần thứ hai, 3CD proteaza, mức biểu hiện của hai thành phần là được điều biến. Việc này được thực hiện bằng việc sử dụng các đoạn nối khác nhau, các hệ khuếch đại ADN có cường độ biến đổi, bằng cách làm thay đổi sự đa dạng tương đối của mỗi chất tiêm chủng (P1 và 3CD) tại thời gian thâm nhiễm hoặc tổ hợp của chúng.

Các vector biểu hiện từ 1300 đến 1308 được sàng lọc đối với khả năng của chúng để biểu hiện P1 đơn lẻ và sau đó được kết hợp với các vector từ 1310 đến 1318, đối với khả năng của chúng để tạo ra hàm lượng mảnh phân giải protein ở mức cao VP1-4. Do chỉ có kháng thể kháng VP1 là sẵn có (Abnova, MAB1255-M05), sự tích tụ của các mảnh phân giải protein được kiểm tra thông qua việc tích tụ của VP1 và sự biến mất của P1 không được xử lý. Như được thể hiện trên Fig.2, sự biểu hiện của P1 đơn lẻ (vector số 1300) dẫn đến sự tích tụ của sản phẩm chứa VP1 có phân tử lượng rõ ràng tương ứng với phân tử lượng của các protein cấu trúc không được xử lý (98 kDa), cho biết rằng proteaza thực vật có thể không tách P1 để tạo ra các capsit protein virus. Tuy nhiên, khi P1 được biểu hiện đồng thời với 3CD (các vector số 1300 + 1310 và 1300 + 1315), tín hiệu 98kDa biến mất một cách hoàn toàn và sản phẩm mới được phát hiện mà tương ứng về phân tử lượng với VP1 (33,5kDa). Kết quả này thể hiện rằng proteaza virus được tạo ra và hoạt động ở mức cao ở thực vật và nó nhận biết và tách chất đồng tạo ra của trong tế bào thực vật để tạo ra EV71 capsit protein.

Các kết quả thu được cho biết rằng mức tích tụ VP1 trong thực vật bị tác động bởi tỷ lệ của *Agrobacterium* chứa P1 protein với *Agrobacterium* chứa 3CD proteaza, với sự tích tụ ở mức cao hơn thu được với tỷ lệ thấp hơn của *Agrobacterium* chứa 3CD proteaza (Fig.2: so sánh 1300+1315 (4:2), 1300+1315 (4:1) và 1300+1315 (4:0,5)). Nguồn gốc của 3CD, HK08 đối lại GDFS08, cũng tác động đến mức tích tụ của VP1 ở thực vật (Fig.2: 1300+1310 (4:0,5) đối lại 1300+1315 (4:0,5)). Cuối cùng, các tác giả sáng chế đã quan sát

được rằng mức tích tụ VP1 cao nhất thu được từ các vector biểu hiện bao gồm các yếu tố khuếch đại ADN (Fig.2: 1301+1311 (4:2)).

Trong thử nghiệm sau đây, P1 được duy trì dưới sự kiểm soát của CPMV-HT+BeYDV (1301) trong khi các cát xét biểu hiện 3CD khác nhau được đồng biến nạp ở các dung dịch pha loãng khác nhau tại thời gian thâm nhiễm. Phép phân tích thẩm tách Western sử dụng kháng thể đơn dòng kháng VP1 đối với dịch chiết protein thô từ các thực vật được biến nạp cho biết rằng sự tích tụ VP1 được quan sát trong khoảng tỷ lệ từ P1 đến proteaza các thành phần cấu trúc. Tổ hợp vector dẫn đến mức tích tụ VP1 lớn nhất bằng 1301 + 1310, với tỷ lệ nồng độ chủng *Agrobacterium* 4:0,5 (protein cấu trúc: proteaza) trong huyền phù vi khuẩn (Fig.3).

Phân tích sự hình thành VLP

Sự hợp nhất của VP1 vào VLP được đánh giá với việc sử dụng sắc ký loại cỡ (SEC) của các dịch chiết được cô đặc. Hạt chất keo được cô đặc từ dịch chiết lọc thô bằng việc ly tâm tốc độ cao (75 000 xg trong 20 phút). Rửa hạt và tạo huyền phù lại trong 1/6 thể tích chất đệm tạo huyền phù lại (50mM PBS độ pH=7,4, 150mM NaCl) và nạp vào cột lọc gel Sephacryl S-500. Rửa giải cột bằng chất đệm tạo huyền phù lại và các phân đoạn rửa giải được đặc trưng bởi SDS-PAGE và thẩm tách western.

Dịch chiết protein từ thực vật được biến nạp một cách tạm thời với 1301+1310 (4:0,5) được cho vào tách SEC phân tích các phân đoạn rửa giải bằng SDS-PAGE nhuộm Coomassie và thẩm tách western kháng VP1. Các kết quả được thể hiện trên Fig.4A thể hiện rằng hầu hết protein vật chủ được rửa giải từ cột trong các phân đoạn sau, đóng gói ở phân đoạn 16 trong khi tín hiệu đặc hiệu VP1 được tìm thấy trong các phân đoạn sớm hơn, đóng gói ở đỉnh 12, nhờ đó rất ít protein vật chủ được tìm thấy. VP1 là protein tương đối nhỏ, nó được mong đợi để rửa giải từ cột bằng phần lớn protein vật chủ nếu không được hợp nhất vào các cấu trúc phân tử lượng ở mức cao. Trong bản mô tả này, protein rửa giải được quan sát đối với VP1 cho biết rõ ràng rằng VP1 được tích hợp vào cấu trúc phân tử lượng cao. Tổ hợp của thẩm tách western và gen nhuộm Coomassie cũng gợi ý rằng protein dựa thừa được nhận dạng bởi mũi tên trong SDS-PAGE nhuộm Coomassie trên Fig.4A có thể là VP1.

Mẫu phân đoạn rửa giải 12 từ thử nghiệm này được gửi đến Institut Armand-Frappier (IAF, Laval, Québec) để phân tích bằng kính hiển vi điện tử truyền động (TEM). Mẫu được thử nghiệm sau khi nhuộm âm tính bằng 3% axit phosphotungstic. Fig.4B thể hiện rằng hạt hình cầu 30nm giống nhau về kích cỡ và vẻ bề ngoài với hạt EV71 rõ ràng được tìm thấy trong môi trường tế bào Vero được nhiễm EV71 (Liu et al., PLoS ONE 6, e20005) được quan sát trong phân đoạn rửa giải 12. Kết quả này cho biết rằng các cấu trúc phân tử lượng cao trong đó VP1 được hợp nhất là các EV71 VLP thật.

Tinh chế riêng phần

Phương pháp tinh chế VLP của nền sàng lọc *VLPEXpress* được phát triển để tinh chế VLP được bọc (đường kính 140nm) từ sinh khối thực vật được biến nạp. Phương pháp sử dụng sự tiêu hoá enzym của vách tế bào để giải phóng hàm lượng ngoại bào và bào thể và dịch chiết thu được được cho vào lọc sâu và vi lọc trước khi được ly tâm ở 16 000g trong 6 giờ để tạo viên VLP. Viên được tạo huyền phù lại trong 1/60 thể tích dung dịch tạo huyền phù lại (100mM Na/KPO₄ độ pH=7,4, 150mM NaCl, 0,01% Tween-80) và lọc vô trùng (0,2µm).

Phương pháp tinh chế *VLPEXpress* được thử nghiệm đối với khả năng của nó để cô đặc 30nm EV71 VLP không bọc. Phương pháp tinh chế được áp dụng cho thực vật được biến nạp với các vector biểu hiện 1301 + 1310 (4:0,5). Sản phẩm thu được được phân tích bằng SDS-PAGE nhuộm Coomassie và thẩm tách western kháng VP1 (Fig.5A). Phép phân tích SDS-PAGE nhuộm Coomassie của sản phẩm tinh chế thể hiện sự có mặt của protein tương ứng về phân tử lượng với protein vỏ EV71 (được chỉ định bằng các mũi tên, Fig.5A, panen phía bên phải). Việc nhận dạng VP1 đã được xác nhận bằng thẩm tách western (Fig.5A, panen phía bên trái). Đối với các capsit protein khác, việc nhận dạng được dựa trên phân tử lượng được ước lượng; 37,5kDa đối với VP0 và 26,5kDa đối với VP3. VP4 và VP2 được mong đợi được tìm thấy ở dạng VP0 không được tách từ khi sự hình thành hạt virus, việc tách giữa VP4 và VP2 chỉ xảy ra sau sự tiếp thu ARN virus. Phép phân tích kính hiển vi điện tử truyền động của sản phẩm được tinh chế cho biết các cấu trúc hình cầu thừa 30nm, tương ứng về kích cỡ và hình dạng với EV71 VLP (Fig.5B).

Kết luận đối với sự biểu hiện

Công việc được thực hiện để chứng minh khả năng của nền biểu hiện tạm thời dựa trên thực vật để tạo ra EV71 VLP dẫn đến các kết luận sau đây:

- EV71 P1 và 3CD protein được tạo ra một cách hiệu quả trong hệ thống
- 3CD hoạt động ở giới thực vật và xử lý một cách đúng đắn P1 thành capsit protein
- EV71 capsit protein tập hợp thành VLP
- EV71 VLP có thể chiết và có thể được tinh chế nguyên vẹn từ sinh khối thực vật.

Ví dụ 2 – Sự biểu hiện Poliovirut

Tổng hợp gen

Các đoạn ADN mã hoá poliovirut (PV) protein cấu trúc P1 và proteaza 3CD từ PV-1 kiểu huyết thanh enterovirut C của người có thể được sử dụng. Các trình tự ứng viên đối với P1 và 3CD là có mặt trong GenBank. Các ví dụ không hạn chế về các trình tự này là:

Đối với P1: các trình tự axit amin GenBank ID NP_041277 (các axit amin 1-881) (SEQ ID NO:10); trình tự nucleotit: GenBank ID NC_002058 (nucleotit 743-3385) (SEQ ID NO: 9);

Đối với 3CD: trình tự axit amin GenBank ID NP_041277 (các axit amin 1566-2209) (SEQ ID NO:8); trình tự nucleotit: GenBank ID NC_002058 (nucleotit 5438-7369) (SEQ ID NO:7).

Hai gen P1 có thể được tổng hợp. Thứ nhất có thể được tạo ra bằng cách sử dụng trình tự kiểu hoang dại trong khi thứ hai có thể được dựa trên trình tự tối ưu (sử dụng codon của người) được xác định bằng cách sử dụng các phương pháp tiêu chuẩn như đã được biết đến trong lĩnh vực này. Gen 3CD có thể được tổng hợp dựa trên trình tự kiểu hoang dại của nó. Cả hai gen kiểu hoang dại (P1 và 3CD) có thể được tổng hợp bằng Invitrogen™ (trước đây là GeneArt®) và gen P1 tối ưu được tối ưu hoá và được tổng hợp bằng DNA2.0.

Tạo dòng phân tử

Các gen được tổng hợp có thể được tạo dòng vào vector biểu hiện thực vật. Các thành phần vector được chọn bao gồm các yếu tố biểu hiện được tạo thành chủ yếu từ các yếu tố điều hoà phiên mã và dịch mã từ virut khảm cây đậu đũa (CPMV) hoặc gen

plastoxyanin cỏ linh lăng, do cả hai yếu tố điều hoà đã được sử dụng trước đây với sự thành công đối với việc biểu hiện ở mức cao của protein tái tổ hợp. Yếu tố khuếch đại ADN từ geminivirut cỏ vàng đậu (BeYDV) cũng có thể được tích hợp vào vector biểu hiện thực vật. Do đó, mỗi cấu trúc gen có thể được tạo dòng trong các vector biểu hiện có hoặc không có yếu tố khuếch đại ADN. Bảng 2 thể hiện các cát xét biểu hiện thực vật mà đã được tập hợp.

Bảng 2. Cát xét biểu hiện thực vật đối với sự biểu hiện của polypeptide P1 cấu trúc PV và proteaza 3CD ở *N. benthamiana*.

Vùng mã hoá	Yếu tố điều hoà	Yếu tố khuếch đại ADN
P1 (Wt)	CPMV HT	-
P1 (Wt)	CPMV HT	BeYDV+rep
P1 (Wt)	Plastoxyanin	-
P1 (Wt)	Plastoxyanin	BeYDV+rep
P1 (Opt)	CPMV HT	-
P1 (Opt)	CPMV HT	BeYDV+rep
P1 (Opt)	Plastoxyanin	-
P1 (Opt)	Plastoxyanin	BeYDV+rep
3CD	CPMV HT	-
3CD	CPMV HT	BeYDV+rep
3CD	Plastoxyanin	-
3CD	Plastoxyanin	BeYDV+rep

Phân tích sự biểu hiện – Chọn lọc các cấu trúc gen tái tổ hợp tốt nhất

Mỗi cát xét biểu hiện có thể được tạo dòng vào vector plasmit mà sau đó có thể được chuyển vào *Agrobacterium tumefaciens*. Sự biểu hiện tạm thời có thể được khởi đầu bằng việc thẩm lọc chân không của chất tiêm chủng *Agrobacterium* chuyển gen mà dẫn đến truyền các bản sao ADN di động của cấu trúc ADN vào tế bào thực vật. Sự biểu hiện tạm thời của nhiều thành phần (biểu hiện đồng thời) có thể được thực hiện bằng việc thẩm lọc hỗn hợp gồm các chất tiêm truyền *Agrobacterium* (đồng thẩm lọc). Do một thành phần được đưa vào thực vật là cấu trúc (P1), và chất nền của thành phần thứ hai, 3CD proteaza, mức biểu hiện của hai thành phần có thể được điều biến. Việc này có thể được thực hiện bằng cách sử dụng các đoạn khởi đầu khác nhau, các hệ khuếch đại ADN có cường độ

thay đổi, bằng cách làm thay đổi lượng dư tương đối của mỗi chất tiêu chuẩn (P1 và 3CD) tại thời gian nhiễm hoặc tổ hợp của chúng.

Các vector biểu hiện với P1 có thể thứ nhất được sàng lọc đối với khả năng của chúng để biểu hiện P1 đơn lẻ và khi kết hợp với các vector 3CD, đối với khả năng của chúng để tạo ra hàm lượng cao của các mảnh phân giải protein. Sự tích tụ của các mảnh phân giải protein có thể được kiểm tra thông qua sự biến mất của P1 không xử lý. Proteaza virus được thể hiện để được tạo ra và hoạt động ở mức cao ở thực vật, cũng như có khả năng nhận biết và tách chất nền đồng tạo ra của nó trong tế bào thực vật để tạo ra PV capsid protein.

ng tích tụ đoạn phân giải protein trong thực vật có thể bị tác động bởi tỷ lệ của *Agrobacterium* chứa P1 protein với *Agrobacterium* chứa 3CD proteaza, với sự tích tụ cao hơn thu được với tỷ lệ thấp hơn của *Agrobacterium* chứa 3CD proteaza. Sự quan sát được thực hiện đối với sự có mặt của các yếu tố khuếch đại ADN và việc sử dụng các yếu tố điều hòa khác nhau đối với việc xử lý P1 và sự tích tụ mảnh phân giải protein.

Phân tích sự hình thành VLP

Việc hợp nhất của VP1 vào VLP có thể được đánh giá bằng việc sử dụng sắc ký loại cỡ (SEC) của các dịch chiết được cô đặc. Hạt chất keo có thể được cô đặc từ các dịch chiết sạch thô bằng việc ly tâm tốc độ cao. Viên có thể được rửa và tạo huyền phù lại trong 1/6 thể tích của chất đệm tạo huyền phù lại và được nạp vào cột lọc gel. Cột có thể được rửa giải bằng chất đệm tạo huyền phù lại và các phân đoạn rửa giải được đặc trưng bởi SDS-PAGE và thẩm tách western.

Dịch chiết protein từ thực vật được biến nạp một cách tạm thời có thể được cho vào việc tách SEC và các phân đoạn rửa giải được phân tích bằng SDS-PAGE nhuộm Coomassie. Các kết quả có thể thể hiện rằng hầu hết protein vật chủ được rửa giải từ cột trong các phân đoạn sau, trong khi tín hiệu đặc hiệu VP1 có thể được tìm thấy trong các phân đoạn sớm hơn. VP1 là protein tương đối nhỏ, nó được mong đợi để rửa giải từ cột với phần lớn protein vật chủ nếu không được hợp nhất vào các cấu trúc phân tử lượng cao. Trong bản mô tả này, protein rửa giải được quan sát đối với VP1 có thể cho biết một cách rõ ràng rằng VP1 đã được tích hợp vào cấu trúc phân tử lượng cao. Việc tổ hợp của thẩm

tách western và gel nhuộm Coomassie cũng có thể gợi ý rằng protein dư được quan sát trong SDS-PAGE nhuộm Coomassie có thể là VP1.

Mẫu từ thử nghiệm này có thể được gửi đến Institut Armand-Frappier (IAF, Laval, Québec) để phân tích bằng kính hiển vi điện tử truyền động (TEM). Các kết quả cho biết rằng cấu trúc phân tử lượng cao trong đó VP1 được hợp nhất là genuine PV VLP.

Tinh chế riêng phần

Phương pháp tinh chế VLP của nền sàng lọc VLPEXpress được phát triển đối với việc tinh chế VLP bọc (đường kính 140nm) từ sinh khối thực vật được biến nạp. Phương pháp sử dụng sự tiêu hoá enzym của vách tế bào để giải phóng hàm lượng ngoại bào và bào thể và dịch chiết thu được được cho vào lọc sâu và vi lọc trước khi được cho vào ly tâm để tạo viên VLP. Viên được tạo huyền phù lại trong dung dịch tạo huyền phù lại và lọc vô trùng.

Phương pháp tinh chế VLPEXpress có thể được thử nghiệm đối với khả năng của nó để cô đặc PV VLP không bọc. Phép phân tích SDS-PAGE nhuộm Coomassie của sản phẩm tinh chế có thể thể hiện sự có mặt của protein tương ứng về phân tử lượng với protein vỏ PV. Việc nhận dạng của capsit protein có thể được dựa trên phân tử lượng được ước lượng của chúng.

Ví dụ 3-Tinh chế

Việc chiết protein được thực hiện bằng cách sử dụng kỹ thuật chiết cơ học hoặc thoái biến enzym của vách tế bào như được mô tả trong WO 2011/035422 và PCT/CA2012/050180. Việc chiết enzym là có lợi so với việc chiết cơ học trong đó nó dẫn đến việc giải phóng tăng sản phẩm với việc giải phóng tối thiểu các protein nhiễm, với các chất gây nhiễm chủ yếu trong dịch chiết thu được là các enzym được sử dụng đối với việc phá vỡ vách tế bào, mà có thể được loại bỏ bằng cách sử dụng các bước xuôi dòng tiếp theo thích hợp.

Dịch chiết cơ học hoặc enzym được cho vào ly tâm để loại bỏ mảnh vụn tế bào, *Agrobacteria*, ADN và hạt lớn hơn. Sau đó, dịch chiết được ly tâm được cho qua bước lọc được thực hiện để loại bỏ các phần rắn trong huyền phù, làm giảm gánh nặng sinh học và làm ổn định và tác động đến dịch chiết trước khi xử lý xuôi dòng. Mặc dù việc thu gom EV71 VLP trong phần lọc có thể không được đánh giá với sự không có mặt của thử

nghiệm định lượng, phép phân tích thẩm tách Western cho biết rằng VLP bị mất trong bước lọc là giảm thiểu. Dịch chiết sạch thu được còn được xử lý tiếp bằng cách sử dụng việc lọc dòng tiếp tuyến (TFF) hoặc nạp một cách trực tiếp vào môi trường sắc ký nếu thích hợp.

Kích cỡ của VLP cho phép sử dụng TFF để loại bỏ đầy đủ và chọn lọc các protein có thể hoà tan được tìm thấy trong dịch chiết sạch, bao gồm các enzym được sử dụng cho việc polyme hoá khur vách tế bào. Bước TFF cũng cô đặc VLP và cho phép trao đổi chất đệm trong khi chuẩn bị sắc ký.

Vài phương pháp sắc ký (sắc ký tương tác trao đổi anion, trao đổi cation, kỵ nước (HIC) và ái lực pseudo), phương thức (gắn kết hoặc dòng chảy) và các điều kiện của chất đệm (độ pH = từ 5 đến 8, tính dẫn từ 10 đến 80mS/cm) được đánh giá đối với khả năng của chúng để làm tăng độ tinh khiết và làm giảm ADN lây nhiễm và nội độc tố, trong khi bảo tồn các đặc tính mong muốn của VLP. Các tác giả sáng chế đã phát hiện ra rằng trong một số điều kiện, POROS® D (nhựa trao đổi anion yếu) được sử dụng theo phương thức thông dòng có thể tạo ra việc loại bỏ ADN và nội độc tố hữu hiệu nhất ra khỏi EV71 VLP được cô đặc.

Bước TFF thứ hai có thể được bổ sung vào sau khi sắc ký trong quy trình tinh chế EV71 VLP. Vai trò của bước TFF này là để cô đặc và tạo ra sản phẩm trong chất đệm mong muốn. Kích cỡ lỗ và điều kiện vận hành đối với bước TFF thứ hai được xác định dựa trên các tham số được nhận dạng đối với TFF thứ nhất. Cuối cùng, chất nền có hạt EV71 tinh khiết được cô đặc một cách rõ ràng thu được sau khi lọc 0,22- μ m. Sản phẩm được tạo ra trong PBS chứa 0,01% Polysorbat 80.

Mô tả đặc điểm VLP

Dây EV71 VLP thứ nhất được tạo ra bằng quy trình được làm thích ứng được mô tả trên đây (dãy số 479-23-018) và sản phẩm được mô tả đặc điểm một cách hoàn toàn (Bảng 3, dãy số 479-23-018). Độ tinh khiết được xác định bằng mật độ kế từ các lớp cắt gel nhuộm Coomassie mà ở đó chỉ có các dải mà thể hiện các tín hiệu dương tính đối với thẩm tách Western (kháng VP1-VP2), và còn có thể được xác nhận bằng phép đo phổ khối, được xét đến như là một phần của sản phẩm. Phép phân tích protein chất lượng sản phẩm cho biết rằng được phẩm chứa EV71 VLP tinh khiết ở mức cao.

Bảng 3. Thuộc tính chất lượng của EV71 VLP, dãy số 479-23-018.

Thuộc tính	EV71 VLP Quy trình khởi đầu
Số lô	479-23-018
Độ tinh khiết	96,4%
Nồng độ protein (BCA, µg/ml)	1192,4
SEC-HPLC (% thể tích rỗng (cấu trúc phân tử lượng cao)	100%
Tán sắc ánh sáng	
Kích cỡ hạt (nm)	48,3
Kính hiển vi điện tử	Hạt tròn Khoảng 30nm Giếng được phân tán
Tạo bản đồ trypsin /MS	2
Số lượng tạp chất được phát hiện ($p < 0,05$ và > 2 peptit)	Ubiquitin (4 pep) Peroxidaza (2 pep)
3 tạp chất đầu tiên	
Gánh nặng sinh học (CFU/ml)	< 10

* Sự ước lượng sơ bộ được tính toán dựa trên phép thử đơn lẻ.

Phép phân tích khác về sản phẩm bằng kính hiển vi điện tử đã xác nhận rằng các EV71 VLP được tinh chế là nguyên vẹn (Fig.6A) và việc tạo bản đồ trypsin bằng phép đo phổ khối xác nhận độ tinh khiết của sản phẩm.

Ví dụ 4 – Quy trình cải biến

Độ tinh khiết của VLP với VP1 nguyên vẹn bằng HIC

Trong quá trình sàng lọc ban đầu của phương pháp sắc ký để tinh chế EV71 VLP, các tác giả sáng chế đã lưu ý rằng nhựa HIC có thể chia tách VLP chứa VP1 nguyên vẹn ra khỏi hạt chứa VP1 được tạo mảnh. Trong một số điều kiện, trong khi hạt chứa mảnh LMW VP1 được gắn kết một cách mạnh mẽ vào nhựa, hạt nguyên vẹn là đang chảy qua cột. Do đó, bước HIC này được lồng vào như bước đánh bóng tiếp theo sắc ký POROS® D. Hạt EV71 được tinh chế qua quy trình này là đồng nhất về kích cỡ (tán sắc ánh sáng), gần với độ tinh khiết 100%, không có chất nhiễm protein nào có thể phát hiện bằng phép đo phổ khối. Các thuộc tính chất lượng sản phẩm của sản phẩm này (dãy số 479-31-020) được thể hiện trong bảng 3, cột trung tâm.

Quy trình chiết được cải biến

Chiết phẩm thực vật có thể được làm sạch bằng việc axit hoá ở độ pH khoảng 5,2 hoặc bằng việc xử lý nhiệt và đông tụ được loại trừ bằng việc ly tâm. Việc xử lý nhiệt được lồng vào giữa các bước chiết cơ học và ly tâm. Việc sử dụng xử lý nhiệt 10 phút ở 60°C (độ pH=8,0) loại trừ nhiều hơn 90% protein có thể hoà tan mà không làm ảnh hưởng đến độ tan của EV71 VLP đến mức có thể phát hiện. VP1 còn lại nguyên vẹn khi được chiết và làm sạch trong các điều kiện này.

Việc chiết cơ học ở độ pH=8,0, kết hợp với việc làm sạch dựa trên xử lý nhiệt, được thực hiện đối với quy trình tinh chế EV71 VLP trong khi thay thế việc chiết enzym. Dây VLP đã được tạo ra bằng quy trình này (dãy số 479-32-020). Đặc điểm của sản phẩm được thể hiện trong bảng 4 (cột thứ ba). Các kết quả thu được thể hiện rằng việc chiết cơ học, được sử dụng trong sự kết hợp với việc làm sạch chủ yếu bằng nhiệt của protein, là bước thu hồi chủ yếu hữu hiệu mà có thể tương thích một cách hoàn toàn với các bước xuôi dòng được xác định trước. Quy trình thu được là có hiệu suất ở mức cao và tạo ra sản phẩm EV71 VLP có độ tinh khiết 98%. Profin tán sắc ánh sáng của EV71 VLP được tạo ra từ quy trình này thể hiện tính đồng nhất ở mức cao. Phép phân tích Cryo TEM của sản phẩm này xác nhận rằng hạt có kích cỡ và hình dạng của EV71 VLP (Fig.8).

Bảng 4. So sánh các đặc điểm EV71 VLP đối với các dây được tạo ra bằng các quy trình bao gồm việc chiết enzymat có hoặc không có bước chiết HIC hoặc cơ học.

Thuộc tính	Chiết enzym (độ pH=5,1) mà không có HIC	Chiết enzym (độ pH=5,1) với HIC	Chiết cơ học (độ pH=8,0) kèm xử lý nhiệt
Dãy số	479-35-020	479-31-020	479-32-020
Độ tinh khiết	95,5%	100%	98,2%
Nồng độ protein (BCA, µg/ml)	352,8	297,8	715,3
SEC-HPLC (% theo thể tích rỗng (cấu trúc phân tử lượng cao))	100%	98,9%	100%
Kính hiển vi điện tử	Được xác định	Hạt tròn 25-30nm Giếng được phân phối	Hạt tròn 25-30nm Giếng được phân phối

* việc ước lượng sơ bộ được tính toán dựa trên thử nghiệm đơn lẻ đối với mỗi quy trình.

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

Sáng chế đã được mô tả xét đến một hoặc nhiều phương án. Tuy nhiên, sẽ rõ ràng đối với người có hiểu biết trung bình về lĩnh vực này rằng nhiều biến thể và cải biến có thể được tạo ra mà không tách rời phạm vi của sáng chế như được xác định trong các điểm yêu cầu bảo hộ.

YÊU CẦU BẢO HỘ

1. Phương pháp tạo ra hạt tương tự enterovirut (EVLP) ở thực vật, bộ phận của thực vật hoặc tế bào thực vật, trong đó phương pháp này bao gồm các bước:

a) đưa vào thực vật, bộ phận của thực vật hoặc tế bào thực vật axit nucleic thứ nhất chứa vùng điều hòa thứ nhất hoạt động ở thực vật được liên kết theo cách hoạt động được với trình tự nucleotit mã hóa polyprotein của enterovirut, trong đó polyprotein của enterovirut bao gồm polyprotein P1 của enterovirut 71;

b) đưa axit nucleic thứ hai chứa vùng điều hòa thứ hai hoạt động ở thực vật và được liên kết theo cách hoạt động được với trình tự nucleotit thứ hai mã hóa một hoặc nhiều proteaza 3C hoặc 3CD của enterovirut 71 vào thực vật, bộ phận của thực vật hoặc tế bào thực vật; và

c) ủ thực vật, bộ phận của thực vật hoặc tế bào thực vật trong các điều kiện cho phép sự biểu hiện của axit nucleic thứ nhất và thứ hai để tạo ra polyprotein P1 của enterovirut 71 và một hoặc nhiều proteaza 3C hoặc 3CD của enterovirut 71, polyprotein P1 của enterovirut 71 được xử lý thành các protein cấu trúc VP1, VP3 và VP0 hoặc VP1, VP2, VP3 và VP4, nhờ đó tạo ra EVLP.

2. Phương pháp theo điểm 1, trong đó tỷ lệ giữa lượng axit nucleic thứ nhất được đưa vào so với lượng axit nucleic thứ hai được đưa vào nằm trong khoảng từ 20:1 đến 0,5:1.

3. Phương pháp theo điểm 1, trong đó polyprotein của enterovirut chứa các protein cấu trúc VP1, VP2, VP3, VP4, hoặc tổ hợp của chúng.

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

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

6. Phương pháp tạo ra hạt tương tự enterovirut (EVLP) ở thực vật, bộ phận của thực vật hoặc tế bào thực vật, trong đó phương pháp này bao gồm các bước:

a) tạo ra thực vật, bộ phận của thực vật hoặc tế bào thực vật chứa axit nucleic thứ nhất chứa vùng điều hòa thứ nhất hoạt động ở thực vật được liên kết theo cách hoạt động được với trình tự nucleotit thứ nhất mã hóa polyprotein của enterovirut trong đó polyprotein của

enterovirut bao gồm polyprotein P1 của enterovirut 71 và axit nucleic thứ hai chứa vùng điều hòa thứ hai hoạt động ở thực vật được liên kết theo cách hoạt động được với trình tự nucleotit thứ hai mã hóa một hoặc nhiều proteaza 3C hoặc 3CD của enterovirut 71;

b) ở thực vật, bộ phận của thực vật hoặc tế bào thực vật trong các điều kiện cho phép sự biểu hiện của các axit nucleic để tạo ra polyprotein P1 của enterovirut 71 và một hoặc nhiều proteaza 3C hoặc 3CD của enterovirut 71, polyprotein P1 của enterovirut 71 được xử lý thành các protein cấu trúc VP1, VP3, và VP0 hoặc VP1, VP2, VP3 và VP4, nhờ đó tạo ra EVLP.

7. Phương pháp theo điểm 1, trong đó trình tự axit nucleic thứ nhất chứa vùng điều hòa thứ nhất được liên kết theo cách hoạt động được với một hoặc nhiều vùng tăng cường comovirut, trình tự nucleotit mã hóa polyprotein của enterovirut, và một hoặc nhiều yếu tố khuếch đại, và phương pháp này còn bao gồm bước: đưa axit nucleic thứ ba mã hóa replicaza vào thực vật, bộ phận của thực vật hoặc tế bào thực vật.

8. Phương pháp theo điểm 1, trong đó axit nucleic thứ hai không chứa một hoặc nhiều yếu tố khuếch đại hoặc một hoặc nhiều vùng tăng cường comovirut.

9. Phương pháp theo điểm 6, trong đó tỷ lệ axit nucleic thứ nhất so với axit nucleic thứ hai nằm trong khoảng từ 20:1 đến 0,5:1.

10. Phương pháp theo điểm 6, trong đó trình tự axit nucleic thứ nhất chứa vùng điều hòa thứ nhất được liên kết theo cách hoạt động được với một hoặc nhiều vùng tăng cường comovirut, trình tự nucleotit mã hóa polyprotein của enterovirut, và một hoặc nhiều yếu tố khuếch đại, và phương pháp này còn bao gồm bước: đưa axit nucleic thứ ba mã hóa replicaza vào thực vật, bộ phận của thực vật, hoặc tế bào thực vật.

11. Phương pháp theo điểm 6, trong đó axit nucleic thứ hai không chứa một hoặc nhiều yếu tố khuếch đại hoặc một hoặc nhiều vùng tăng cường comovirut.

12. Phương pháp theo điểm 1, trong đó phương pháp này còn bao gồm (các) bước:

(d) thu nhận thực vật, bộ phận của thực vật hoặc tế bào thực vật; và/hoặc

(e) tinh chế EVLP.

13. Phương pháp theo điểm 6, trong đó phương pháp này còn bao gồm (các) bước:

(d) thu nhận thực vật, bộ phận của thực vật hoặc tế bào thực vật; và/hoặc

(e) tinh chế EVLP.

14. Phương pháp theo điểm 1, trong đó axit nucleic thứ nhất và axit nucleic thứ hai nằm trên các cấu trúc axit nucleic riêng rẽ.

15. Phương pháp theo điểm 6, trong đó axit nucleic thứ nhất và axit nucleic thứ hai nằm trên các cấu trúc axit nucleic riêng rẽ.

16. Phương pháp theo điểm 1, trong đó axit nucleic thứ nhất và axit nucleic thứ hai nằm trên cùng cấu trúc axit nucleic.

17. Phương pháp theo điểm 6, trong đó axit nucleic thứ nhất và axit nucleic thứ hai nằm trên cùng cấu trúc axit nucleic.

18. Dịch chiết thực vật chứa hạt tương tự enterovirut (EVLP) được tạo ra bằng phương pháp bao gồm các bước:

a) đưa vào thực vật, bộ phận của thực vật hoặc tế bào thực vật axit nucleic thứ nhất chứa vùng điều hòa thứ nhất hoạt động ở thực vật được liên kết theo cách hoạt động được với trình tự nucleotit mã hóa polyprotein của enterovirut, trong đó polyprotein của enterovirut bao gồm polyprotein P1 của enterovirut 71;

b) đưa axit nucleic thứ hai chứa vùng điều hòa thứ hai hoạt động ở thực vật và được liên kết theo cách hoạt động được với trình tự nucleotit thứ hai mã hóa một hoặc nhiều proteaza 3C hoặc 3CD của enterovirut 71 vào thực vật, bộ phận của thực vật hoặc tế bào thực vật;

c) ủ thực vật, bộ phận của thực vật, hoặc tế bào thực vật trong các điều kiện cho phép sự biểu hiện axit nucleic thứ nhất và thứ hai để tạo ra polyprotein P1 của enterovirut 71 và một hoặc nhiều proteaza 3C hoặc 3CD của enterovirut 71, polyprotein P1 của enterovirut 71 được xử lý thành các protein cấu trúc VP1, VP3 và VP0 hoặc VP1, VP2, VP3 và VP4, nhờ đó tạo ra EVLP; và

d) thu nhận thực vật, bộ phận của thực vật hoặc tế bào thực vật để tạo ra dịch chiết thực vật.

19. Chế phẩm dùng để gây đáp ứng miễn dịch chống lại enterovirut 71 ở đối tượng, trong đó chế phẩm này chứa dịch chiết thực vật theo điểm 18 và chất mang được dụng.

20. Dịch chiết thực vật chứa hạt tương tự enterovirut (EVLP) được tạo ra bằng phương pháp bao gồm các bước:

a) tạo ra thực vật, bộ phận của thực vật hoặc tế bào thực vật chứa axit nucleic thứ nhất chứa vùng điều hòa thứ nhất hoạt động ở thực vật được liên kết theo cách hoạt động được với trình tự nucleotit thứ nhất mã hóa polyprotein của enterovirut trong đó polyprotein của enterovirut bao gồm polyprotein P1 của enterovirut 71 và axit nucleic thứ hai chứa vùng điều hòa thứ hai hoạt động ở thực vật được liên kết theo cách hoạt động được với trình tự nucleotit thứ hai mã hóa một hoặc nhiều proteaza 3C hoặc 3CD của enterovirut 71;

b) ủ thực vật, bộ phận của thực vật hoặc tế bào thực vật trong các điều kiện cho phép sự biểu hiện các axit nucleic để tạo ra polyprotein P1 của enterovirut 71 và một hoặc nhiều proteaza 3C hoặc 3CD của enterovirut 71, polyprotein P1 của enterovirut 71 được xử lý thành các protein cấu trúc VP1, VP3, và VP0 hoặc VP1, VP2, VP3 và VP4, nhờ đó tạo ra EVLP; và

c) thu nhận thực vật, bộ phận của thực vật hoặc tế bào thực vật để tạo ra dịch chiết thực vật.

21. Chế phẩm dùng để gây đáp ứng miễn dịch chống lại enterovirut 71 ở đối tượng, trong đó chế phẩm này chứa dịch chiết thực vật theo điểm 20 và chất mang được dùng.

22. Tế bào thực vật chứa axit nucleic thứ nhất và axit nucleic thứ hai, axit nucleic thứ nhất chứa vùng điều hòa thứ nhất hoạt động ở tế bào thực vật và được liên kết theo cách hoạt động được với trình tự nucleotit thứ nhất mã hóa polyprotein của enterovirut, trong đó polyprotein của enterovirut bao gồm polyprotein P1 của enterovirut 71, và axit nucleic thứ hai chứa vùng điều hòa thứ hai hoạt động ở tế bào thực vật và được liên kết theo cách hoạt động được với trình tự nucleotit thứ hai mã hóa một hoặc nhiều hơn một proteaza 3C hoặc 3CD của enterovirut 71.

23. Tế bào thực vật chứa hạt tương tự enterovirut (EVLP), trong đó EVLP này chứa các protein cấu trúc VP1, VP2, VP3 và VP4.

24. Hạt tương tự enterovirut (EVLP) tinh chế thu được từ dịch chiết thực vật theo điểm 18.

25. Chế phẩm chứa lượng hữu hiệu điều trị EVLP theo điểm 24 để gây đáp ứng miễn dịch ở đối tượng và chất mang được dùng.

26. EVLP tinh chế thu được từ dịch chiết thực vật theo điểm 20.

27. Chế phẩm chứa lượng hữu hiệu điều trị của EVLP theo điểm 26 để gây đáp ứng miễn dịch ở đối tượng và chất mang được dùng.

DANH MỤC TRÌNH TỰ

<110> MEDICAGO INC.

<120> PHƯƠNG PHÁP TẠO RA HẠT TUƠNG TỰ ENTEROVIRUT, DỊCH CHIẾT THỰC VẬT CHỨA HẠT TUƠNG TỰ ENTEROVIRUT VÀ HẠT TUƠNG TỰ ENTEROVIRUT THU ĐƯỢC TỪ DỊCH CHIẾT NÀY

<130> V84804WO

<150> US 61/697,266

<151> 2012-09-05

<160> 10

<170> PatentIn version 3.5

<210> 1

<211> 645

<212> PRT

<213> Enterovirut A71

<400> 1

Gly	Pro	Ser	Leu	Asp	Phe	Ala	Leu	Ser	Leu	Leu	Arg	Arg	Asn	Val	Arg
1				5					10					15	

Gln	Val	Gln	Thr	Asp	Gln	Gly	His	Phe	Thr	Met	Leu	Gly	Val	Arg	Asp
			20					25					30		

Arg	Leu	Ala	Val	Leu	Pro	Arg	His	Ser	Gln	Pro	Gly	Lys	Thr	Ile	Trp
		35					40					45			

Ile	Glu	His	Lys	Leu	Val	Asn	Val	Leu	Asp	Ala	Val	Glu	Leu	Val	Asp
	50					55				60					

Glu	Gln	Gly	Val	Asn	Leu	Glu	Leu	Thr	Leu	Ile	Thr	Leu	Asp	Thr	Asn
65					70					75					80

Glu	Lys	Phe	Arg	Asp	Ile	Thr	Lys	Phe	Ile	Pro	Glu	Asn	Ile	Ser	Ala
				85					90					95	

Ala	Ser	Asp	Ala	Thr	Leu	Val	Ile	Asn	Thr	Glu	His	Met	Pro	Ser	Met
			100					105						110	

Phe	Val	Pro	Val	Gly	Asp	Val	Val	Gln	Tyr	Gly	Phe	Leu	Asn	Leu	Ser
		115					120					125			

Gly Lys Pro Thr His Arg Thr Met Met Tyr Asn Phe Pro Thr Lys Ala
 130 135 140

Gly Gln Cys Gly Gly Val Val Thr Ser Val Gly Lys Ile Ile Gly Ile
 145 150 155 160

His Ile Gly Gly Asn Gly Arg Gln Gly Phe Cys Ala Gly Leu Lys Arg
 165 170 175

Ser Tyr Phe Ala Ser Glu Gln Gly Glu Ile Gln Trp Val Lys Pro Asn
 180 185 190

Lys Glu Thr Gly Arg Leu Asn Ile Asn Gly Pro Thr Arg Thr Lys Leu
 195 200 205

Glu Pro Ser Val Phe His Asp Ile Phe Glu Gly Asn Lys Glu Pro Ala
 210 215 220

Val Leu His Ser Lys Asp Pro Arg Leu Glu Val Asp Phe Glu Gln Ala
 225 230 235 240

Leu Phe Ser Lys Tyr Val Gly Asn Thr Leu Tyr Glu Pro Asp Glu Tyr
 245 250 255

Ile Lys Glu Ala Ala Leu His Tyr Ala Asn Gln Leu Lys Gln Leu Glu
 260 265 270

Ile Asn Thr Ser Gln Met Ser Met Glu Glu Ala Cys Tyr Gly Thr Glu
 275 280 285

Asn Leu Glu Ala Ile Asp Leu His Thr Ser Ala Gly Tyr Pro Tyr Ser
 290 295 300

Ala Leu Gly Ile Lys Lys Arg Asp Ile Leu Asp Pro Thr Thr Arg Asp
 305 310 315 320

Val Ser Lys Met Lys Phe Tyr Met Asp Lys Tyr Gly Leu Asp Leu Pro
 325 330 335

Tyr Ser Thr Tyr Val Lys Asp Glu Leu Arg Ser Ile Asp Lys Ile Lys
 340 345 350

Lys Gly Lys Ser Arg Leu Ile Glu Ala Ser Ser Leu Asn Asp Ser Val
 355 360 365

Tyr Leu Arg Met Ala Phe Gly His Leu Tyr Glu Ala Phe His Ala Asn
 370 375 380

Pro Gly Thr Ile Thr Gly Ser Ala Val Gly Cys Asn Pro Asp Thr Phe
 385 390 395 400

Trp Ser Lys Leu Pro Ile Leu Leu Pro Gly Ser Leu Phe Ala Phe Asp
 405 410 415

Tyr Ser Gly Tyr Asp Ala Ser Leu Ser Pro Val Trp Phe Arg Ala Leu
 420 425 430

Glu Leu Val Leu Arg Glu Ile Gly Tyr Ser Glu Gly Ala Val Ser Leu
 435 440 445

Ile Glu Gly Ile Asn His Thr His His Val Tyr Arg Asn Lys Thr Tyr
 450 455 460

Cys Val Leu Gly Gly Met Pro Ser Gly Cys Ser Gly Thr Ser Ile Phe
 465 470 475 480

Asn Ser Met Ile Asn Asn Ile Ile Ile Arg Ala Leu Leu Ile Lys Thr
 485 490 495

Phe Lys Gly Ile Asp Leu Asp Glu Leu Asn Met Val Ala Tyr Gly Asp
 500 505 510

Asp Val Leu Ala Ser Tyr Pro Phe Pro Ile Asp Cys Leu Glu Leu Ala
 515 520 525

Lys Thr Gly Lys Glu Tyr Gly Leu Thr Met Thr Pro Ala Asp Lys Ser
 530 535 540

Pro Cys Phe Asn Glu Val Asn Trp Gly Asn Ala Thr Phe Leu Lys Arg
 545 550 555 560

Gly Phe Leu Pro Asp Glu Gln Phe Pro Phe Leu Ile His Pro Thr Met
 565 570 575

Pro Met Arg Glu Ile His Glu Ser Ile Arg Trp Thr Lys Asp Ala Arg
580 585 590

Asn Thr Gln Asp His Val Arg Ser Leu Cys Leu Leu Ala Trp His Asn
595 600 605

Gly Lys Gln Glu Tyr Glu Lys Phe Val Ser Thr Ile Arg Ser Val Pro
610 615 620

Val Gly Arg Ala Leu Ala Ile Pro Asn Tyr Glu Asn Leu Arg Arg Asn
625 630 635 640

Trp Leu Glu Leu Phe
645

<210> 2
<211> 1935
<212> ADN
<213> Enterovirut A71

<400> 2
ggcccgagcc ttgattttgc cctctcccta ctgaggagga acgtcaggca agtccaaaca 60
gaccaggggc atttcacat gttgggtgtt agggatcgct tagcagtcct cccacgccac 120
tcacaaccgc gcaaaactat ttggattgag cacaaactcg tgaacgtcct tgatgcagtt 180
gaattggtgg atgagcaagg agtcaacctg gagttaacct tcatcactct tgacactaac 240
gaaaagttaa gggatatcac caaattcatc ccagaaaata ttagtgctgc cagtgatgcc 300
accctagtga tcaacacgga gcacatgccc tcaatgtttg tcccgggtggg tgacgttgtg 360
cagtatggct tcttgaacct cagtggcaag cctaccatc gcacatgat gtacaacttt 420
cctactaaag caggacagtg tgggggagtg gtgacatctg ttgggaagat tatcggtatt 480
cacattggtg gcaatggcag acaaggTTTT tgcgcaggcc tcaaaaggag ttactttgct 540
agtgaacaag gagagatcca gtgggttaag cccaataaag aaactggaag actcaacatc 600
aatggacca cccgcacca gctagaacct agtgtattcc atgatatctt tgagggaaat 660
aaggagccag ctgtcttgca cagtaaagac ccccgacttg aggtagattt tgaacaggcc 720
ctgttctcta agtatgtggg gaatacacta tatgagcctg acgagtacat caaagaggca 780
gctcttcatt atgcaaacca attaaagcag ctagaaatca acacctctca aatgagcatg 840

```

gaggaggcct gctacggtac tgagaatcct gaggcatttg atcttcatac tagtgcaggc 900
taccctata gtgccctggg gataaagaaa agagacatct tagaccctac caccagggac 960
gtgagtaaaa tgaagttcta catggacaaa tatgggtcttg atctccctta ctccacttat 1020
gtcaaggacg agctgcgctc aattgataaa attaagaaag ggaagtcccc tctgattgag 1080
gccagtagtt taaatgattc agtgtacctt agaatggctt tcggtcattt gtatgaggct 1140
ttccacgcaa atcctgggac tataactgga tcagccgtgg ggtgtaacct tgacacattc 1200
tggagcaagc tgccaatttt gctccctggg tcaactcttg cctttgacta ctcaggttat 1260
gatgctagcc ttagccctgt ctgggtcaga gcattagaat tggtccttag ggagataggg 1320
tatagtgaag gggcagtctc actcattgag ggaatcaacc acacacacca tgtgtatcgt 1380
aataagacct attgtgtgct tgggtgggatg cctcaggct gctcgggaac atccattttc 1440
aactcaatga tcaacaacat tattatcaga gcactgctca taaaaacatt taagggcatt 1500
gatttggatg aactcaacat ggctcgcttat ggagatgatg tgctcgctag ctacccttc 1560
ccaattgatt gcttggagtt agcgaagact ggcaaggagt atggtctaac catgaccct 1620
gcggataagt ctcttgctt taatgaagtt aattggggta atgcgacctt tctcaagagg 1680
ggctttttac ccgatgaaca gtttccattt ttgatccacc ccactatgcc aatgagggag 1740
atccatgagt ccattcgatg gaccaaggat gcacgaaaca ctcaagatca tgtgcgggtcc 1800
ttgtgcctcc tagcatggca taatggtaag caagaatatg agaaatttgt gagtacaatt 1860
aggtctgtcc cagtgggaag agcggtggct atcccaaatt atgaaaacct tagacgtaat 1920
tggtctgagt tattt 1935

```

```

<210> 3
<211> 645
<212> PRT
<213> Enterovirus A71

```

```
<400> 3
```

```

Gly Pro Ser Leu Asp Phe Ala Leu Ser Leu Leu Arg Arg Asn Ile Arg
1           5           10           15

```

```

Gln Val Gln Thr Asp Gln Gly His Phe Thr Met Leu Gly Val Arg Asp
20           25           30

```

```

Arg Leu Ala Val Leu Pro Arg His Ser Gln Pro Gly Lys Thr Ile Trp
35           40           45

```

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

Ile Asn Thr Ser Gln Met Ser Met Glu Glu Ala Cys Tyr Gly Thr Glu
 275 280 285

Asn Leu Glu Ala Ile Asp Leu His Thr Ser Ala Gly Tyr Pro Tyr Ser
 290 295 300

Ala Leu Gly Ile Lys Lys Arg Asp Ile Leu Asp Pro Thr Thr Arg Asp
 305 310 315 320

Val Ser Lys Met Lys Ser Tyr Met Asp Lys Tyr Gly Leu Asp Leu Pro
 325 330 335

Tyr Ser Thr Tyr Val Lys Asp Glu Leu Arg Ser Ile Asp Lys Ile Lys
 340 345 350

Lys Gly Lys Ser Arg Leu Ile Glu Ala Ser Ser Leu Asn Asp Ser Val
 355 360 365

Tyr Leu Arg Met Thr Phe Gly His Leu Tyr Glu Ala Phe His Ala Asn
 370 375 380

Pro Gly Thr Ile Thr Gly Ser Ala Val Gly Cys Asn Pro Asp Thr Phe
 385 390 395 400

Trp Ser Lys Leu Pro Ile Leu Leu Pro Gly Ser Leu Phe Ala Phe Asp
 405 410 415

Tyr Ser Gly Tyr Asp Ala Ser Leu Ser Pro Val Trp Phe Arg Ala Leu
 420 425 430

Glu Met Val Leu Arg Glu Ile Gly Tyr Ser Glu Glu Ala Val Ser Leu
 435 440 445

Ile Glu Gly Ile Asn His Thr His His Val Tyr Arg Asn Lys Thr Tyr
 450 455 460

Cys Val Leu Gly Gly Met Pro Ser Gly Cys Ser Gly Thr Ser Ile Phe
 465 470 475 480

Asn Ser Met Ile Asn Asn Ile Ile Ile Arg Ala Leu Leu Ile Lys Thr
 485 490 495

Phe Lys Gly Ile Asp Leu Asp Glu Leu Asn Met Val Ala Tyr Gly Asp
500 505 510

Asp Val Leu Ala Ser Tyr Pro Phe Pro Ile Asp Cys Leu Glu Leu Ala
515 520 525

Lys Thr Gly Lys Glu Tyr Gly Leu Thr Met Thr Pro Ala Asp Lys Ser
530 535 540

Pro Cys Phe Asn Glu Val Asn Trp Gly Asn Ala Thr Phe Leu Lys Arg
545 550 555 560

Gly Phe Leu Pro Asp Glu Gln Phe Pro Phe Leu Ile His Pro Thr Met
565 570 575

Pro Met Arg Glu Ile His Glu Ser Ile Arg Trp Thr Lys Asp Ala Arg
580 585 590

Asn Thr Gln Asp His Val Arg Ser Leu Cys Leu Leu Ala Trp His Asn
595 600 605

Gly Lys Gln Glu Tyr Glu Lys Phe Val Ser Thr Ile Arg Ser Val Pro
610 615 620

Ile Gly Arg Ala Leu Ala Ile Pro Asn Tyr Glu Asn Leu Arg Arg Asn
625 630 635 640

Trp Leu Glu Leu Phe
645

<210> 4
<211> 1935
<212> ADN
<213> Enterovirut A71

<400> 4
ggcccgagtc ttgattttgc tctctccctg ttaaggagga acatcaggca agtccaaaca 60
gaccaggggc atttcacccat gttgggtggt agggatcggt tagcagtcct cccacgtcac 120
tcacaacccg gcaaaactat ttggatcgaa cacaaactcg tgaacattct tgatgcagtt 180
gaattggtgg atgagcaagg agtcaacctg gaattgaccc tcatcactct tgacactaac 240

gaaaagttta gggatatcac caaattcatc ccagaaagta ttagcactgc cagtgatgcc	300
accctagtga tcaacacgga gcacatgccc tcaatgtttg tcccgggtggg tgacgtcgtg	360
cagtatggct ttttgaatct tagtggaag cccacccatc gcacatgat gtacaacttt	420
cctactaaag cgggacagtg tggaggagta gtgacatctg ttgggaaagt catcgggtatt	480
cacattgggtg gcaatggtag acaagggtttt tgcgcaggcc tcaaaaggag ttactttgct	540
agtgaacaag gggagatcca gtgggttaag cccaataaag aaactggaag actcaacatc	600
aatggaccaaa cccgcaccaa gttggaaccc agtgtattcc atgatgtctt cgagggaaat	660
aaggaaccag ctgtcttgca cggcaaagat ccccgactcg aggtagattt tgagcaggcc	720
ctgttctcta agtatgtggg aaacacgcta tatgagcctg acgagtacat caaagaggca	780
gctcttcatt atgcaaatac attaaagcaa ctagaaatta atacctccca gatgagcatg	840
gaggaagcct gctatggtac tgagaatctt gaggtatcg atcttcatac tagtgcagggt	900
taccctata gtgccttggg aataaagaaa agagacatct tagaccctac caccagggac	960
gtgagtaaaa tgaaatccta tatggacaaa tatgggtctcg atctccctta ctccacttat	1020
gtcaaggatg agctgcgctc aattgataaa attaaagaaag ggaagtcccg tctgatcgag	1080
gccagcagtt taaatgattc agtgtacctc agaatgactt tcggtcattt gtatgaggct	1140
ttccacgcaa atcctgggac gataactgga tcagccgtgg ggtgtaaccc tgacacattc	1200
tggagcaagc tgccaatctt gcttcctggt tcactctttg cctttgacta ctcaggttat	1260
gatgctagcc ttagccctgt ctgggtcaga gcattagaaa tggtccttag ggagataggg	1320
tatagtgaag aggcggtctc actcattgag ggaatcaacc acacacacca cgtgtatcgt	1380
aacaagacct attgtgtgct tgggtgggatg ccctcaggct gttcgggaac atccatcttc	1440
aactcaatga tcaacaacat tattatcaga gcaactgctc taaaaacatt taagggcatt	1500
gatttggatg aactcaacat ggtcgcttat ggggatgatg tgcttgctag ctacccttc	1560
ccaatcgatt gcttggagtt agcaaagact ggcaaggagt atggtctgac catgactcct	1620
gcagataagt ccccttgctt taatgaagtt aattggggta atgcgacctt cctcaagagg	1680
ggctttttac ctgatgagca gtttccattt ttgatccacc ctactatgcc aatgcgggag	1740
atccatgaat ccattcgatg gactaaggac gcacgaaaca ctcaagatca tgtacgggtcc	1800
ttgtgcctcc tagcatggca taatggtaag caagaatatg aaaaatttgt gagcacaatt	1860
aggtctgtcc caataggaag agctttggct atcccaaatt atgaaaatct tagacgcaat	1920

tggctcgagt tatttt

1935

<210> 5

<211> 862

<212> PRT

<213> Enterovirut A71

<400> 5

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

Asn	Ser	Ala	Thr	Glu	Gly	Ser	Thr	Ile	Asn	Tyr	Thr	Thr	Ile	Asn	Tyr
			20					25					30		

Tyr	Lys	Asp	Ser	Tyr	Ala	Ala	Thr	Ala	Gly	Lys	Gln	Ser	Leu	Lys	Gln
		35					40					45			

Asp	Pro	Asp	Lys	Phe	Ala	Asn	Pro	Val	Lys	Asp	Ile	Phe	Thr	Glu	Met
	50					55					60				

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

Arg	Val	Ala	Gln	Leu	Thr	Ile	Gly	Asn	Ser	Thr	Ile	Thr	Thr	Gln	Glu
				85					90					95	

Ala	Ala	Asn	Ile	Ile	Val	Gly	Tyr	Gly	Glu	Trp	Pro	Ser	Tyr	Cys	Ser
		100						105					110		

Asp	Ser	Asp	Ala	Thr	Ala	Val	Asp	Lys	Pro	Thr	Arg	Pro	Asp	Val	Ser
		115					120					125			

Val	Asn	Arg	Phe	Tyr	Thr	Leu	Asp	Thr	Lys	Leu	Trp	Glu	Lys	Ser	Ser
	130					135					140				

Lys	Gly	Trp	Tyr	Trp	Lys	Phe	Pro	Asp	Val	Leu	Thr	Glu	Thr	Gly	Val
145					150					155					160

Phe	Gly	Gln	Asn	Ala	Gln	Phe	His	Tyr	Leu	Tyr	Arg	Ser	Gly	Phe	Cys
				165					170					175	

Ile	His	Val	Gln	Cys	Asn	Ala	Ser	Lys	Phe	His	Gln	Gly	Ala	Leu	Leu
			180					185					190		

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

Gly Thr Glu Asp Ser His Pro Pro Tyr Lys Gln Thr Gln Pro Gly Ala
 210 215 220

Asp Gly Phe Glu Leu Gln His Pro Tyr Val Leu Asp Ala Gly Ile Pro
 225 230 235 240

Ile Ser Gln Leu Thr Val Cys Pro His Gln Trp Ile Asn Leu Arg Thr
 245 250 255

Asn Asn Cys Ala Thr Ile Ile Val Pro Tyr Ile Asn Ala Leu Pro Phe
 260 265 270

Asp Ser Ala Leu Asn His Cys Asn Phe Gly Leu Leu Val Val Pro Ile
 275 280 285

Ser Pro Leu Asp Tyr Asp Gln Gly Ala Thr Pro Val Ile Pro Ile Thr
 290 295 300

Ile Thr Leu Ala Pro Met Cys Ser Glu Phe Ala Gly Leu Arg Gln Ala
 305 310 315 320

Val Thr Gln Gly Phe Pro Thr Glu Leu Lys Pro Gly Thr Asn Gln Phe
 325 330 335

Leu Thr Thr Asp Asp Gly Val Ser Ala Pro Ile Leu Pro Asn Phe His
 340 345 350

Pro Thr Pro Cys Ile His Ile Pro Gly Glu Val Arg Asn Leu Leu Glu
 355 360 365

Leu Cys Gln Val Glu Thr Ile Leu Glu Val Asn Asn Val Pro Thr Asn
 370 375 380

Ala Thr Ser Leu Met Glu Arg Leu Arg Phe Pro Val Ser Ala Gln Ala
 385 390 395 400

Gly Lys Gly Glu Leu Cys Ala Val Phe Arg Ala Asp Pro Gly Arg Asn
 405 410 415

Gly Pro Trp Gln Ser Thr Leu Leu Gly Gln Leu Cys Gly Tyr Tyr Thr
 420 425 430

Gln Trp Ser Gly Ser Leu Glu Val Thr Phe Met Phe Thr Gly Ser Phe
 435 440 445

Met Ala Thr Gly Lys Met Leu Ile Ala Tyr Thr Pro Pro Gly Gly Pro
 450 455 460

Leu Pro Lys Asp Arg Ala Thr Ala Met Leu Gly Thr His Val Ile Trp
 465 470 475 480

Asp Phe Gly Leu Gln Ser Ser Val Thr Leu Val Ile Pro Trp Ile Ser
 485 490 495

Asn Thr His Tyr Arg Ala His Ala Arg Asp Gly Val Phe Asp Tyr Tyr
 500 505 510

Thr Thr Gly Leu Val Ser Ile Trp Tyr Gln Thr Asn Tyr Val Val Pro
 515 520 525

Ile Gly Ala Pro Asn Thr Ala Tyr Ile Ile Ala Leu Ala Ala Ala Gln
 530 535 540

Lys Asn Phe Thr Met Lys Leu Cys Lys Asp Ala Ser Asp Ile Leu Gln
 545 550 555 560

Thr Gly Thr Ile Gln Gly Asp Arg Val Ala Asp Val Ile Glu Ser Ser
 565 570 575

Ile Gly Asp Ser Val Ser Arg Ala Leu Thr Gln Ala Leu Pro Ala Pro
 580 585 590

Thr Gly Gln Asn Thr Gln Val Ser Ser His Arg Leu Asp Thr Gly Lys
 595 600 605

Val Pro Ala Leu Gln Ala Ala Glu Ile Gly Ala Ser Ser Asn Ala Ser
 610 615 620

Asp Glu Ser Met Ile Glu Thr Arg Cys Val Leu Asn Ser His Ser Thr
 625 630 635 640

Ala Glu Thr Thr Leu Asp Ser Phe Phe Ser Arg Ala Gly Leu Val Gly
645 650 655

Glu Ile Asp Leu Pro Leu Glu Gly Thr Thr Asn Pro Asn Gly Tyr Ala
660 665 670

Asn Trp Asp Ile Asp Ile Thr Gly Tyr Ala Gln Met Arg Arg Lys Val
675 680 685

Glu Leu Phe Thr Tyr Met Arg Phe Asp Ala Glu Phe Thr Phe Val Ala
690 695 700

Cys Thr Pro Thr Gly Glu Val Val Pro Gln Leu Leu Gln Tyr Met Phe
705 710 715 720

Val Pro Pro Gly Ala Pro Lys Pro Asp Ser Arg Glu Ser Leu Ala Trp
725 730 735

Gln Thr Ala Thr Asn Pro Ser Val Phe Val Lys Leu Ser Asp Pro Pro
740 745 750

Ala Gln Val Ser Val Pro Phe Met Ser Pro Ala Ser Ala Tyr Gln Trp
755 760 765

Phe Tyr Asp Gly Tyr Pro Thr Phe Gly Glu His Lys Gln Glu Lys Asp
770 775 780

Leu Glu Tyr Gly Ala Cys Pro Asn Asn Met Met Gly Thr Phe Ser Val
785 790 795 800

Arg Thr Val Gly Thr Ser Lys Ser Lys Tyr Pro Leu Val Val Arg Ile
805 810 815

Tyr Met Arg Met Lys His Val Arg Ala Trp Ile Pro Arg Pro Met Arg
820 825 830

Asn Gln Asn Tyr Leu Phe Lys Ala Asn Pro Asn Tyr Ala Gly Asn Ser
835 840 845

Ile Lys Pro Thr Gly Thr Ser Arg Thr Ala Ile Thr Thr Leu
850 855 860

<210> 6
 <211> 2586
 <212> ADN
 <213> Enterovirut A71

<400> 6
 atgggttcgc aggtgtccac gcagcgctcc ggttctcatg aaaattcaaa ctcagccacc 60
 gagggttcca ccataaacta caccaccatt aattattaca aagactccta tgctgccaca 120
 gcaggcaaac agagtctcaa gcaggatcca gacaagtttg caaatcctgt taaagacatc 180
 ttcactgaaa tggcagcgcc actgaagtcc ccatccgctg aggcattgtg atacagtgat 240
 cgagtagcgc aattaactat tggtaactcc accatcacca cgcaagaagc ggctaacatc 300
 atagttggtt atggtgagtg gccttcctac tgctcggatt ctgacgctac agcagtggat 360
 aagccaacgc gcccggtatg ttcagtgaac aggttttata cattggacac taaattgtgg 420
 gagaaatcgt ccaagggatg gtactggaaa ttcccggatg tgttaactga aactgggggtt 480
 tttgggcaaa atgcacaatt ccactacctc taccgatcag ggttctgtat ccacgtgcag 540
 tgcaatgcta gtaaattcca ccaaggagca ctctagtcg ctgttctacc agagtacgtc 600
 attgggacag tggcaggcgg cacagggacg gaagatagtc acccccctta caagcagact 660
 caaccggcg ccgatggctt cgaattgcaa caccgctacg tgcttgatgc tggcatccca 720
 atatcacagt taacagtgtg cccacatcag tggattaatt tgagaaccaa caattgtgct 780
 acaataatag tgccatacat taacgcactg ccttttgatt ccgccttgaa ccactgcaat 840
 tttggcctat tagttgtgcc tattagcca ctagattacg accaaggagc gacgccagta 900
 atccctataa ctatcacatt agccccaatg tgttctgaat tcgcaggtct taggcaggca 960
 gtcacgcaag gatttccac cgagttgaaa cctggcacia atcaattttt aaccactgat 1020
 gatggcggtt cagcacctat tctaccaaac ttccacccca ccccggtgat ccatatacct 1080
 ggtgaagtta ggaacttgct agagttatgc caggtggaaa ccattctaga ggttaacaat 1140
 gtgcccacga atgccactag tttaatggag agactgcgct ttccagtctc agcacaagca 1200
 gggaaagggtg agctgtgtgc ggtgttcaga gctgatcctg ggcgaaatgg gccgtggcag 1260
 tccaccttgc tgggtcagtt gtgtgggtat tacacccaat ggtcaggatc attggaagtc 1320
 accttcatgt ttactggatc ctttatggct accggcaaga tgctcatagc ctatacaccg 1380
 ccaggaggcc ctttgcccaa ggaccgggcg accgccatgt tgggcacgca cgtcatctgg 1440

gatTTTgggc tgcaatcgtc cgTTaccctt gtaataccat ggatcagcaa cactcactac 1500
 agagcgcacg cccgagatgg agtgTTtgac tactacacca cagggTtagt cagtatatgg 1560
 tatcagacaa attacgtggT tccaattggg gcgcctaata cagcctatat aatagcacta 1620
 gcggcagccc aaaagaattt cactatgaag ttgtgcaagg atgctagtga tatcctacaa 1680
 acgggcacca tccaggggaga tagggtagca gatgtaattg aaagttccat aggggatagc 1740
 gtgagcagag ccctcactca agctctacca gcacccacag gccagaacac acaggtgagc 1800
 agtcatcgac tggatacagg caaggTtcca gcactccaag ctgctgaaat tggagcatca 1860
 tcaaatgcta gtgacgagag catgatcgag acacgctgtg ttcttaactc gcacagcaca 1920
 gctgagacca ctcttgatag tttcttcagc agagcgggat tagttggaga gatagatctt 1980
 cctcttgaag gcacaactaa cccaaatggT tatgccaaact gggacataga tataacaggt 2040
 tacgcacaaa tgcgcagaaa ggtggagtta ttcacctaca tgcgctttga tgcagagttc 2100
 actttcgTtg cgtgcacacc taccggggaa gttgtccac aattgctcca atatatgttt 2160
 gtaccacctg gagcccctaa gccagactcc agggagtccc tcgcatggca aaccgccacc 2220
 aaccctcag tttttgtcaa gttgtcagac cctccagcac aggtttcagt accattcatg 2280
 tcaccgcga gtgcttacca atggttctat gacggatata ccacattcgg ggaacacaaa 2340
 caggagaaaag atcttgagta tggggcgtgc cctaataaca tgatgggtac gttctcagtg 2400
 cggactgtag ggacttccaa atccaagtat cttttagtgg ttaggattta catgaggatg 2460
 aagcacgtca gggcgtggat acctcgcccg atgcgtaacc aaaactacct attcaaggcc 2520
 aacccaaatt atgctggcaa ctccattaag ccaactggta ctagtcgcac agcgatcact 2580
 actctt 2586

<210> 7
 <211> 1932
 <212> ADN
 <213> Enterovirus C

<400> 7
 ggaccagggT tcgattacgc agtggctatg gctaaaagaa acattgttac agcaactact 60
 agcaagggag agttcactat gttaggagtc cacgacaacg tggctatTTT accaaccac 120
 gcttcacctg gtgaaagcat tgtgatcgat ggcaaagaag tggagatctt ggatgccaaa 180
 gcgctcgaag atcaagcagg aaccaatctt gaaatcacta taatcactct aaagagaaat 240
 gaaaagttca gagacattag accacatata cctactcaaa tctactgagac aaatgatgga 300

gtcttgatcg tgaacactag caagtacccc aatatgtatg ttctgtgcgg tgctgtgact	360
gaacagggat atctaaatct cgggtgggcgc caaactgctc gtactctaata gtacaacttt	420
ccaaccagag caggacagtg tgggtggagtc atcacatgta ctgggaaagt catcgggatg	480
catgttggtg ggaacggttc acacggggtt gcagcggccc tgaagcgatc atacttcact	540
cagagtcaag gtgaaatcca gtggatgaga ctttcgaagg aagtgggata tccaatcata	600
aatgccccgt ccaaaaccaa gcttgaaccc agtgctttcc actatgtggt tgaaggggtg	660
aaggaaccag cagtcctcac taaaaacgat ccagggctta agacagactt tgaggaggca	720
atcttctcca agtacgtggg taacaaaatt actgaagtgg atgagtacat gaaagaggca	780
gtagaccact atgctggcca gctcatgtca ctagacatca acacagaaca aatgtgcttg	840
gaggatgcca tgtatggcac tgatggtcta gaagcacttg atttgtccac cagtgtcggc	900
tacccttatg tagcaatggg aaagaagaag agagacatct tgaacaaaca aaccagagac	960
actaaggaaa tgcaaaaact gctcgacaca tatggaatca acctcccact ggtgacttat	1020
gtaaaggatg aacttagatc caaaacaaag gttgagcagg ggaaatccag attaattgaa	1080
gcttctagtt tgaatgactc agtggcaatg agaatggctt ttgggaacct atatgctgct	1140
tttcacaaaa acccaggagt gataacaggc tcagcagtgg ggtgcgatcc agatttgttt	1200
tggagcaaaa ttccggtatt gatggaagag aagctgtttg cttttgacta cacaggggat	1260
gatgcatctc tcagccctgc ttggttcgag gcactaaaga tggtgcttga gaaaatcgga	1320
ttcggagaca gagttgacta catcgactac ctaaaccact cacaccacct gtacaagaat	1380
aaaacatact gtgtcaaggg cggtatgcca tctggctgct caggcacttc aatttttaac	1440
tcaatgatta acaacttgat tatcaggaca ctcttactga aaacctaca gggcatagat	1500
ttagaccacc taaaaatgat tgcctatggt gatgatgtaa ttgcttccta ccccatgaa	1560
gttgacgcta gtctcctagc ccaatcagga aaagactatg gactaactat gactccagct	1620
gacaaatcag ctacatttga aacagtcaca tgggagaatg taacattctt gaagagattc	1680
ttcagggcag acgagaaata cccatttctt attcatccag taatgccaat gaaggaaatt	1740
catgaatcaa ttagatggac taaagatcct aggaacactc aggatcacgt tcgctctctg	1800
tgccttttag cttggcacia tggcgaagaa gaatataaca aattcctagc taaaatcagg	1860
agtggtgcaa ttggaagagc tttattgctc ccagagtact caacattgta ccgccgttgg	1920
cttgactcat tt	1932

<210> 8
 <211> 644
 <212> PRT
 <213> Enterovirut C

<400> 8

Gly Pro Gly Phe Asp Tyr Ala Val Ala Met Ala Lys Arg Asn Ile Val
 1 5 10 15

Thr Ala Thr Thr Ser Lys Gly Glu Phe Thr Met Leu Gly Val His Asp
 20 25 30

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

Ile Asp Gly Lys Glu Val Glu Ile Leu Asp Ala Lys Ala Leu Glu Asp
 50 55 60

Gln Ala Gly Thr Asn Leu Glu Ile Thr Ile Ile Thr Leu Lys Arg Asn
 65 70 75 80

Glu Lys Phe Arg Asp Ile Arg Pro His Ile Pro Thr Gln Ile Thr Glu
 85 90 95

Thr Asn Asp Gly Val Leu Ile Val Asn Thr Ser Lys Tyr Pro Asn Met
 100 105 110

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

Gly Arg Gln Thr Ala Arg Thr Leu Met Tyr Asn Phe Pro Thr Arg Ala
 130 135 140

Gly Gln Cys Gly Gly Val Ile Thr Cys Thr Gly Lys Val Ile Gly Met
 145 150 155 160

His Val Gly Gly Asn Gly Ser His Gly Phe Ala Ala Ala Leu Lys Arg
 165 170 175

Ser Tyr Phe Thr Gln Ser Gln Gly Glu Ile Gln Trp Met Arg Pro Ser
 180 185 190

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

Glu Pro Ser Ala Phe His Tyr Val Phe Glu Gly Val Lys Glu Pro Ala
 210 215 220

Val Leu Thr Lys Asn Asp Pro Arg Leu Lys Thr Asp Phe Glu Glu Ala
 225 230 235 240

Ile Phe Ser Lys Tyr Val Gly Asn Lys Ile Thr Glu Val Asp Glu Tyr
 245 250 255

Met Lys Glu Ala Val Asp His Tyr Ala Gly Gln Leu Met Ser Leu Asp
 260 265 270

Ile Asn Thr Glu Gln Met Cys Leu Glu Asp Ala Met Tyr Gly Thr Asp
 275 280 285

Gly Leu Glu Ala Leu Asp Leu Ser Thr Ser Ala Gly Tyr Pro Tyr Val
 290 295 300

Ala Met Gly Lys Lys Lys Arg Asp Ile Leu Asn Lys Gln Thr Arg Asp
 305 310 315 320

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

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

Gln Gly Lys Ser Arg Leu Ile Glu Ala Ser Ser Leu Asn Asp Ser Val
 355 360 365

Ala Met Arg Met Ala Phe Gly Asn Leu Tyr Ala Ala Phe His Lys Asn
 370 375 380

Pro Gly Val Ile Thr Gly Ser Ala Val Gly Cys Asp Pro Asp Leu Phe
 385 390 395 400

Trp Ser Lys Ile Pro Val Leu Met Glu Glu Lys Leu Phe Ala Phe Asp
 405 410 415

Tyr Thr Gly Tyr Asp Ala Ser Leu Ser Pro Ala Trp Phe Glu Ala Leu
 420 425 430

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

Asp Tyr Leu Asn His Ser His His Leu Tyr Lys Asn Lys Thr Tyr Cys
 450 455 460

Val Lys Gly Gly Met Pro Ser Gly Cys Ser Gly Thr Ser Ile Phe Asn
 465 470 475 480

Ser Met Ile Asn Asn Leu Ile Ile Arg Thr Leu Leu Leu Lys Thr Tyr
 485 490 495

Lys Gly Ile Asp Leu Asp His Leu Lys Met Ile Ala Tyr Gly Asp Asp
 500 505 510

Val Ile Ala Ser Tyr Pro His Glu Val Asp Ala Ser Leu Leu Ala Gln
 515 520 525

Ser Gly Lys Asp Tyr Gly Leu Thr Met Thr Pro Ala Asp Lys Ser Ala
 530 535 540

Thr Phe Glu Thr Val Thr Trp Glu Asn Val Thr Phe Leu Lys Arg Phe
 545 550 555 560

Phe Arg Ala Asp Glu Lys Tyr Pro Phe Leu Ile His Pro Val Met Pro
 565 570 575

Met Lys Glu Ile His Glu Ser Ile Arg Trp Thr Lys Asp Pro Arg Asn
 580 585 590

Thr Gln Asp His Val Arg Ser Leu Cys Leu Leu Ala Trp His Asn Gly
 595 600 605

Glu Glu Glu Tyr Asn Lys Phe Leu Ala Lys Ile Arg Ser Val Pro Ile
 610 615 620

Gly Arg Ala Leu Leu Leu Pro Glu Tyr Ser Thr Leu Tyr Arg Arg Trp
 625 630 635 640

Leu Asp Ser Phe

<210> 9
 <211> 2643
 <212> ADN
 <213> Enterovirut C

<400> 9
 atgggtgctc aggtttcatc acagaaagtg ggcgcacatg aaaactcaaa tagagcgtat 60
 ggtggtttcta ccattaatta caccaccatt aattattata gagattcagc tagtaacgcg 120
 gcttcgaaac aggacttctc tcaagaccct tccaagttca ccgagcccat caaggatgtc 180
 ctgataaaaa cagccccaat gctaaactcg ccaaacatag aggcttgccg gtatagcgat 240
 agagtactgc aattaacact gggaaactcc actataacca cacaggaggc ggctaattca 300
 gtagtcgctt atgggcgttg gcctgaatat ctgagggaca gcgaagccaa tccagtggac 360
 cagccgacag aaccagacgt cgctgcacatc aggttttata cgctagacac cgtgtccttg 420
 acgaaagagt cgcgaggggtg gtggtggaag ttgcctgatg cactgaggga catgggactc 480
 tttgggcaaa atatgtacta ccactaccta ggtaggtccg ggtacaccgt gcatgtacag 540
 tgtaacgcct ccaaattcca ccagggggca ctaggggtat tcgccgtacc agagatgtgt 600
 ctggccgggg atagcaacac cactaccatg cacaccagct atcaaaatgc caatcctggc 660
 gagaaaggag gcactttcac gggtagcttc actcctgaca acaaccagac atcacctgcc 720
 cgcaggttct gcccggtgga ttacctcctt ggaaatggca cgttggtggg gaatgccttt 780
 gtgttcccgc accagataat aaacctacgg accaacaact gtgctacact ggtactccct 840
 tacgtgaact ccctctcgat agatagtatg gtaaagcaca ataattgggg aattgcaata 900
 ttaccattgg ccccatataa ttttgctagt gagtcctccc cagagattcc aatcaccttg 960
 accatagccc ctatgtgctg tgagttcaat ggattaagaa acatcacctt gccacgctta 1020
 cagggcctgc cggtcacgaa cacccttggt agcaatcaat atcttactgc agacaacttc 1080
 cagtcaccgt gtgcgctgcc tgaatttgat gtgacccac ctattgacat acccggtgaa 1140
 gtaaagaaca tgatggaatt ggcagaaatc gacaccatga ttccctttga ctttaagtgc 1200
 aaaaaaaga acaccatgga aatgtatagg gttcggttaa gtgacaaacc acatacagac 1260
 gatcccatat tctgcctgtc actctctcca gcttcagatc ctaggttgct acatactatg 1320
 cttggagaaa tcctaaatta ctacacacac tgggcaggat ccctgaagtt cacgtttctg 1380

```

ttctgtggat tcatgatggc aactggcaaa ctgttggtgt catacgcgcc tcctggagcc 1440
gaccaccaa agaagcgtaa ggaggcgatg ttgggaacac atgtgatctg ggacatagga 1500
ctgcagtcct catgtactat ggtagtgcc tggattagca acaccacgta tcggcaaacc 1560
atagatgata gtttcaccga aggcgggatac atcagcgtct tctaccaaac tagaatagtc 1620
gtccctcttt cgacaccag agagatggac atccttggtt ttgtgtcagc gtgtaatgac 1680
ttcagcgtgc gcttggtgag agataccaca catatagagc aaaaagcgct agcacagggg 1740
ttaggtcaga tgcttgaaag catgattgac aacacagtcc gtgaaacggt gggggcggca 1800
acatctagag acgctctccc aaacactgaa gccagtggac caacacactc caaggaaatt 1860
ccggcactca ccgcagtgga aactggggcc acaaatccac tagtcccttc tgatacagtg 1920
caaaccagac atgttgtaca acataggtca aggtcagagt ctagcataga gtctttcttc 1980
gcgcgggggtg catgcgtgac cattatgacc gtggataacc cagcttccac cacgaataag 2040
gataagctat ttgcagtgtg gaagatcact tataaagata ctgtccagtt acggaggaaa 2100
ttggagttct tcacctattc tagatttgat atggaactta cctttgtggt tactgcaaatt 2160
ttcactgaga ctaacaatgg gcatgcctta aatcaagtgt accaaattat gtacgtacca 2220
ccaggcgctc cagtgccga gaaatgggac gactacacat ggcaaactc atcaaatcca 2280
tcaatctttt acacctacgg aacagctcca gcccggatct cggtaccgta tgttggtatt 2340
tcgaacgcct attcacactt ttacgacggt ttttccaaag taccactgaa ggaccagtgc 2400
gcagcactag gtgactccct ttatggtgca gcatctctaa atgacttcgg tattttggct 2460
gttagagtag tcaatgatca caaccgacc aaggtcacct ccaaaatcag agtgtatcta 2520
aaacccaaac acatcagagt ctggtgcccg cgtccaccga gggcagtggc gtactacggc 2580
cctggagtgg attacaagga tggtagctt acacccctct ccaccaagga tctgaccaca 2640
tat 2643

```

```

<210> 10
<211> 881
<212> PRT
<213> Enterovirus C

```

```
<400> 10
```

```

Met Gly Ala Gln Val Ser Ser Gln Lys Val Gly Ala His Glu Asn Ser
1           5           10          15

```

Asn Arg Ala Tyr Gly Gly Ser Thr Ile Asn Tyr Thr Thr Ile Asn Tyr
 20 25 30

Tyr Arg Asp Ser Ala Ser Asn Ala Ala Ser Lys Gln Asp Phe Ser Gln
 35 40 45

Asp Pro Ser Lys Phe Thr Glu Pro Ile Lys Asp Val Leu Ile Lys Thr
 50 55 60

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

Arg Val Leu Gln Leu Thr Leu Gly Asn Ser Thr Ile Thr Thr Gln Glu
 85 90 95

Ala Ala Asn Ser Val Val Ala Tyr Gly Arg Trp Pro Glu Tyr Leu Arg
 100 105 110

Asp Ser Glu Ala Asn Pro Val Asp Gln Pro Thr Glu Pro Asp Val Ala
 115 120 125

Ala Cys Arg Phe Tyr Thr Leu Asp Thr Val Ser Trp Thr Lys Glu Ser
 130 135 140

Arg Gly Trp Trp Trp Lys Leu Pro Asp Ala Leu Arg Asp Met Gly Leu
 145 150 155 160

Phe Gly Gln Asn Met Tyr Tyr His Tyr Leu Gly Arg Ser Gly Tyr Thr
 165 170 175

Val His Val Gln Cys Asn Ala Ser Lys Phe His Gln Gly Ala Leu Gly
 180 185 190

Val Phe Ala Val Pro Glu Met Cys Leu Ala Gly Asp Ser Asn Thr Thr
 195 200 205

Thr Met His Thr Ser Tyr Gln Asn Ala Asn Pro Gly Glu Lys Gly Gly
 210 215 220

Thr Phe Thr Gly Thr Phe Thr Pro Asp Asn Asn Gln Thr Ser Pro Ala
 225 230 235 240

Arg Arg Phe Cys Pro Val Asp Tyr Leu Leu Gly Asn Gly Thr Leu Leu
 245 250 255

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

Asn Cys Ala Thr Leu Val Leu Pro Tyr Val Asn Ser Leu Ser Ile Asp
 275 280 285

Ser Met Val Lys His Asn Asn Trp Gly Ile Ala Ile Leu Pro Leu Ala
 290 295 300

Pro Leu Asn Phe Ala Ser Glu Ser Ser Pro Glu Ile Pro Ile Thr Leu
 305 310 315 320

Thr Ile Ala Pro Met Cys Cys Glu Phe Asn Gly Leu Arg Asn Ile Thr
 325 330 335

Leu Pro Arg Leu Gln Gly Leu Pro Val Met Asn Thr Pro Gly Ser Asn
 340 345 350

Gln Tyr Leu Thr Ala Asp Asn Phe Gln Ser Pro Cys Ala Leu Pro Glu
 355 360 365

Phe Asp Val Thr Pro Pro Ile Asp Ile Pro Gly Glu Val Lys Asn Met
 370 375 380

Met Glu Leu Ala Glu Ile Asp Thr Met Ile Pro Phe Asp Leu Ser Ala
 385 390 395 400

Thr Lys Lys Asn Thr Met Glu Met Tyr Arg Val Arg Leu Ser Asp Lys
 405 410 415

Pro His Thr Asp Asp Pro Ile Leu Cys Leu Ser Leu Ser Pro Ala Ser
 420 425 430

Asp Pro Arg Leu Ser His Thr Met Leu Gly Glu Ile Leu Asn Tyr Tyr
 435 440 445

Thr His Trp Ala Gly Ser Leu Lys Phe Thr Phe Leu Phe Cys Gly Phe
 450 455 460

Met	Met	Ala	Thr	Gly	Lys	Leu	Leu	Val	Ser	Tyr	Ala	Pro	Pro	Gly	Ala
465					470					475					480
Asp	Pro	Pro	Lys	Lys	Arg	Lys	Glu	Ala	Met	Leu	Gly	Thr	His	Val	Ile
				485					490					495	
Trp	Asp	Ile	Gly	Leu	Gln	Ser	Ser	Cys	Thr	Met	Val	Val	Pro	Trp	Ile
			500					505					510		
Ser	Asn	Thr	Thr	Tyr	Arg	Gln	Thr	Ile	Asp	Asp	Ser	Phe	Thr	Glu	Gly
		515					520					525			
Gly	Tyr	Ile	Ser	Val	Phe	Tyr	Gln	Thr	Arg	Ile	Val	Val	Pro	Leu	Ser
	530						535				540				
Thr	Pro	Arg	Glu	Met	Asp	Ile	Leu	Gly	Phe	Val	Ser	Ala	Cys	Asn	Asp
545					550					555					560
Phe	Ser	Val	Arg	Leu	Leu	Arg	Asp	Thr	Thr	His	Ile	Glu	Gln	Lys	Ala
				565					570					575	
Leu	Ala	Gln	Gly	Leu	Gly	Gln	Met	Leu	Glu	Ser	Met	Ile	Asp	Asn	Thr
			580					585					590		
Val	Arg	Glu	Thr	Val	Gly	Ala	Ala	Thr	Ser	Arg	Asp	Ala	Leu	Pro	Asn
		595					600					605			
Thr	Glu	Ala	Ser	Gly	Pro	Thr	His	Ser	Lys	Glu	Ile	Pro	Ala	Leu	Thr
	610					615					620				
Ala	Val	Glu	Thr	Gly	Ala	Thr	Asn	Pro	Leu	Val	Pro	Ser	Asp	Thr	Val
625					630					635					640
Gln	Thr	Arg	His	Val	Val	Gln	His	Arg	Ser	Arg	Ser	Glu	Ser	Ser	Ile
				645					650					655	
Glu	Ser	Phe	Phe	Ala	Arg	Gly	Ala	Cys	Val	Thr	Ile	Met	Thr	Val	Asp
			660					665					670		
Asn	Pro	Ala	Ser	Thr	Thr	Asn	Lys	Asp	Lys	Leu	Phe	Ala	Val	Trp	Lys
		675					680					685			

Ile Thr Tyr Lys Asp Thr Val Gln Leu Arg Arg Lys Leu Glu Phe Phe
690 695 700

Thr Tyr Ser Arg Phe Asp Met Glu Leu Thr Phe Val Val Thr Ala Asn
705 710 715 720

Phe Thr Glu Thr Asn Asn Gly His Ala Leu Asn Gln Val Tyr Gln Ile
725 730 735

Met Tyr Val Pro Pro Gly Ala Pro Val Pro Glu Lys Trp Asp Asp Tyr
740 745 750

Thr Trp Gln Thr Ser Ser Asn Pro Ser Ile Phe Tyr Thr Tyr Gly Thr
755 760 765

Ala Pro Ala Arg Ile Ser Val Pro Tyr Val Gly Ile Ser Asn Ala Tyr
770 775 780

Ser His Phe Tyr Asp Gly Phe Ser Lys Val Pro Leu Lys Asp Gln Ser
785 790 795 800

Ala Ala Leu Gly Asp Ser Leu Tyr Gly Ala Ala Ser Leu Asn Asp Phe
805 810 815

Gly Ile Leu Ala Val Arg Val Val Asn Asp His Asn Pro Thr Lys Val
820 825 830

Thr Ser Lys Ile Arg Val Tyr Leu Lys Pro Lys His Ile Arg Val Trp
835 840 845

Cys Pro Arg Pro Pro Arg Ala Val Ala Tyr Tyr Gly Pro Gly Val Asp
850 855 860

Tyr Lys Asp Gly Thr Leu Thr Pro Leu Ser Thr Lys Asp Leu Thr Thr
865 870 875 880

Tyr

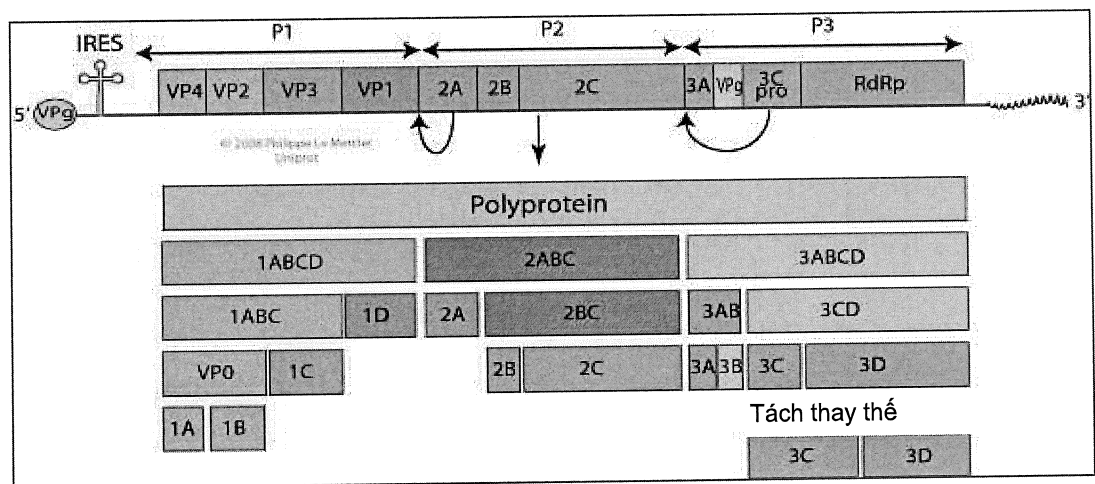


Fig.1

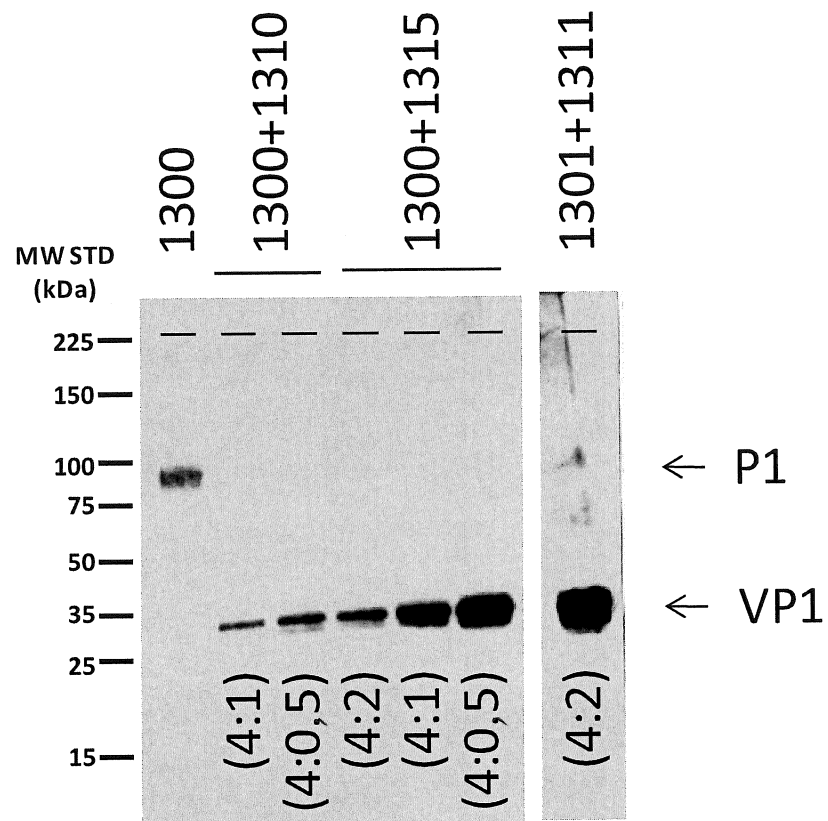


Fig.2

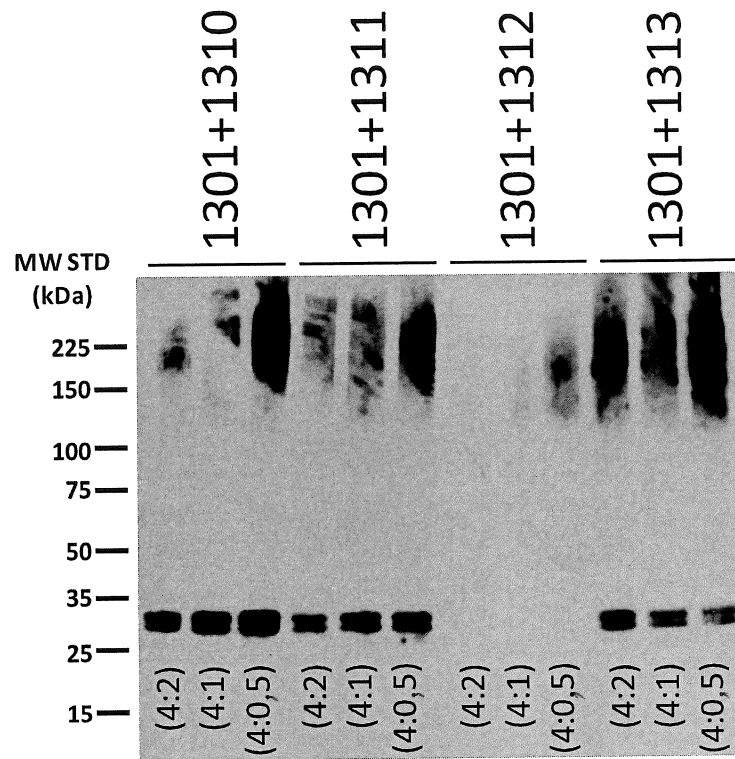


Fig.3

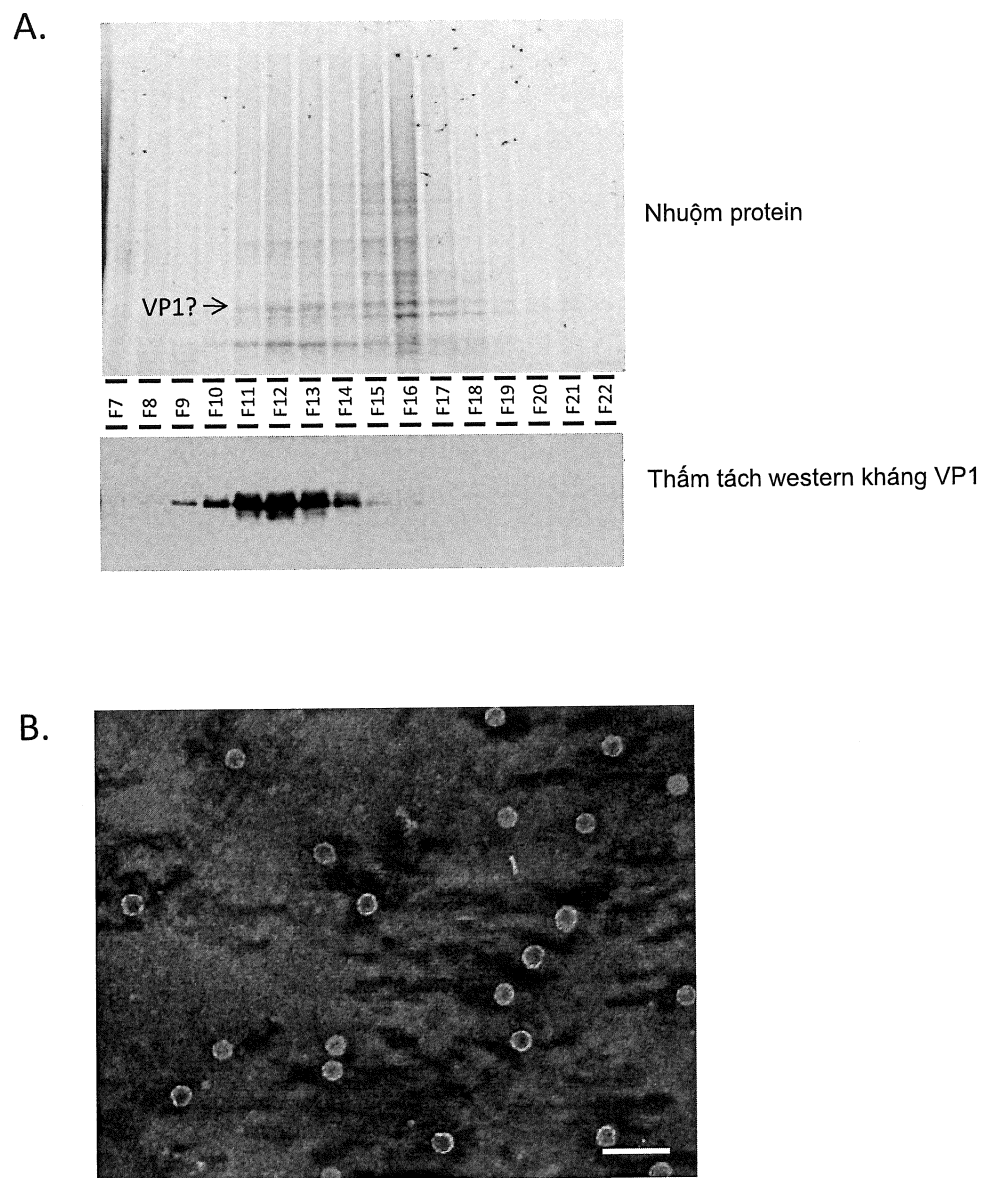


Fig.4

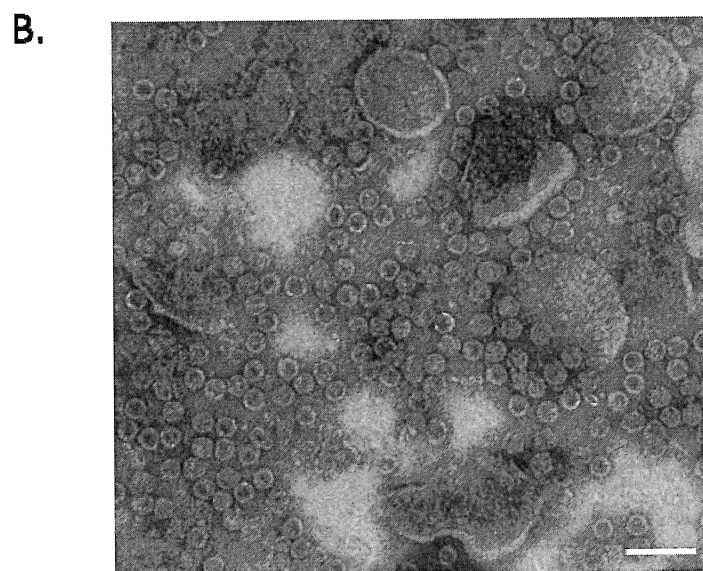
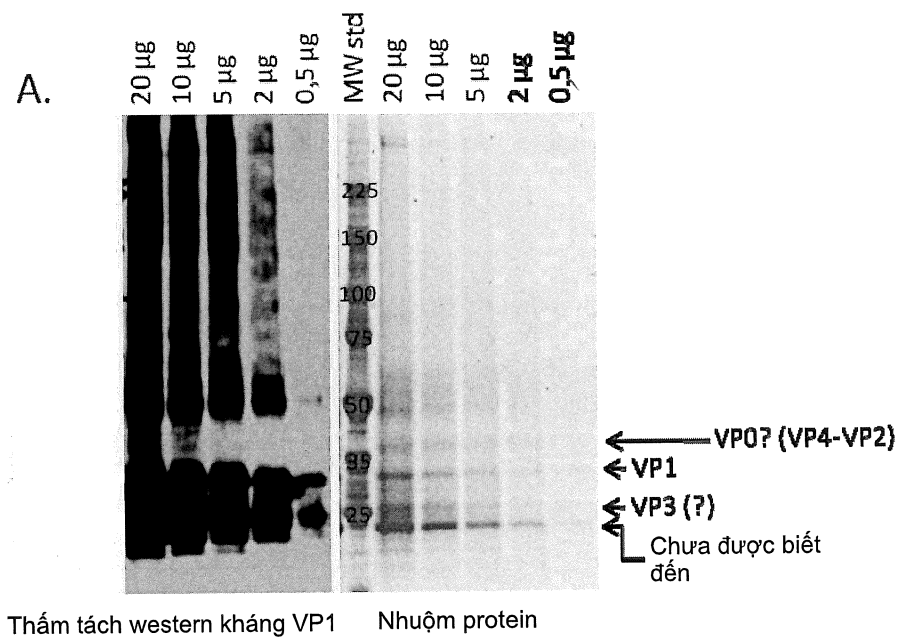


Fig.5

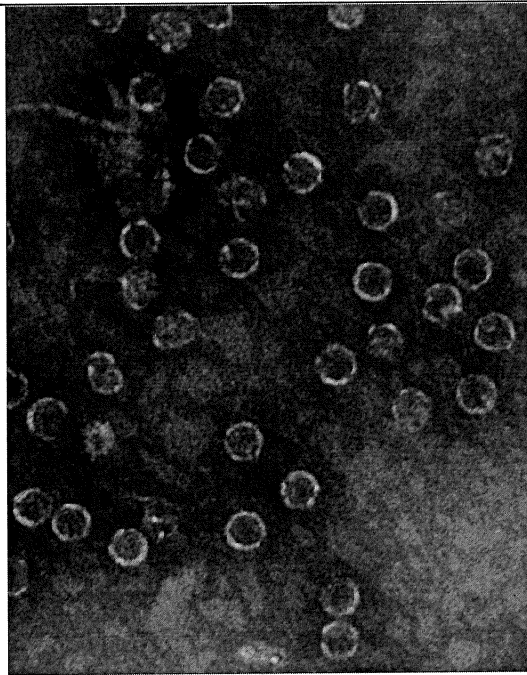


Fig.6

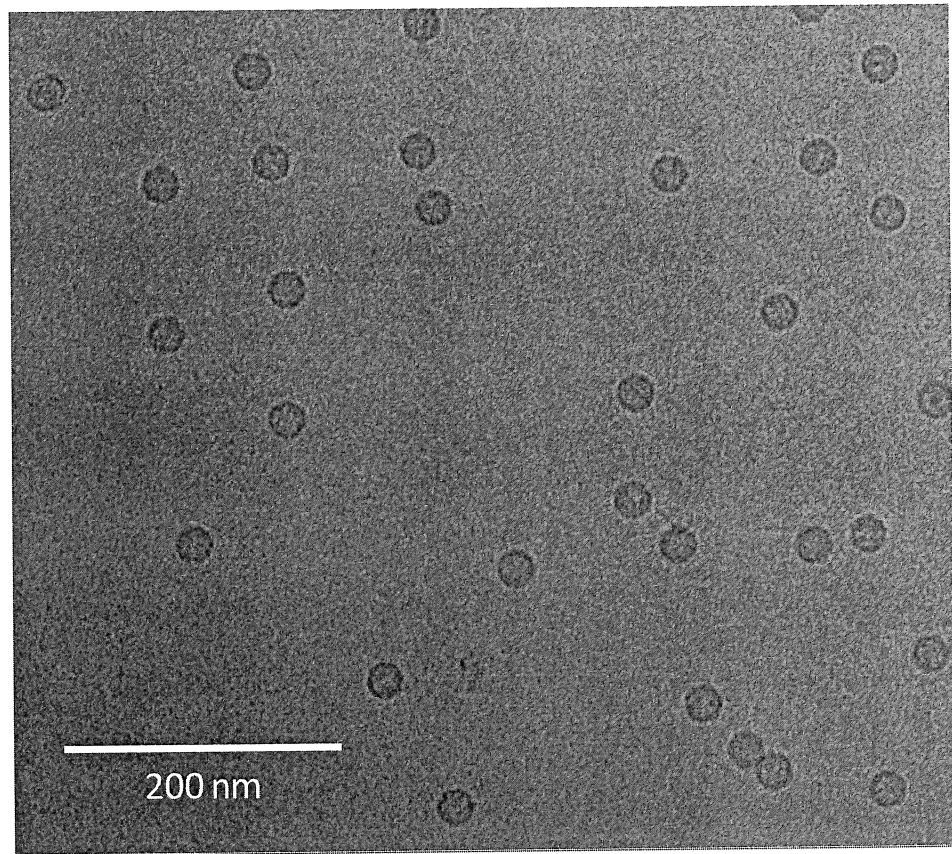


Fig.7

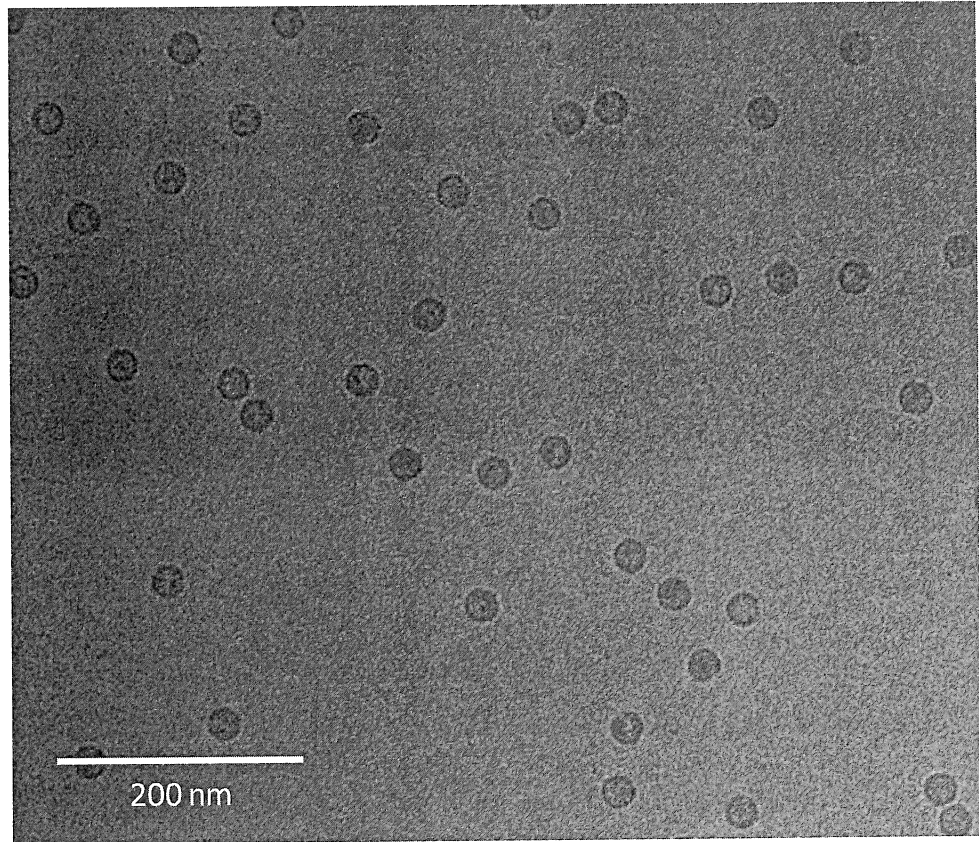


Fig.8

Fig.9A (SEQ ID NO:1)

GPSLDFALSLLRRNVRQVQTDQGHFTMLGVRDRDLAVLPRHSQPGKTIWIEHKLNVNVLDA
 VELVDEQGVNLELTITLDTNEKFRDITKFI PENISAASDATALVINTEHMPSMFVPVGD
 VVQYGFNLNLSGKPTHRTMMYNFPTKAGQCGGVVTSVGKIIGIHIHGGNGRQGFCAGLKRS
 YFASEQGEIQWVKPNKETGRLNINGPTRTKLEPSVFHDIFEGNKEPAVLHSKDPRLEVD
 FEQALFSKYVGNTLYEPDEYIKEAALHYANQLKQLEINTSQMSMEEACYGTENLEAIDL
 HTSAGYPYSALGIKKRDILDPTTRDVS KMKFYMDKYGLDLPYSTYVKDELRSIDKIKKG
 KSRLIEASSLND SVYL RMAFGHLYEAFHANPGTITGSAVGCNPDTFWSKLPILLPGSLF
 AFDYSGYDASLSPVWFRALELVLEIGYSEGAVSLIEGINHTHHVYRNKTYCVLGGMPS
 GCSGTSIFNSMINNIIIRALLIKTFKIDLDLDELNMVAYGDDVLASYPPIDCLELAKTG
 KEYGLTMT PADKSPCFNEVNWGNATFLKRGFLPDEQFPFLIHPTMPMREIHESI RWTKD
 ARNTQDHVRS LCLLAWHNGKQEYEFVSTIRSVPVGRALAI PNYENLRNWLLELF

Fig.9B (SEQ ID NO:2)

GGCCCGAGCCTTGATTTTGCCCTCTCCCTACTGAGGAGGAACGTCAGGCAAGTCCAAAC
 AGACCAGGGGCATTTACCATGTTGGGTGTTAGGGATCGCTTAGCAGTCCTCCCACGCC
 ACTCACAACCCGGCAAACTATTTGGATTGAGCACAACTCGTGAACGTCCTTGATGCA
 GTTGAATTGGTGGATGAGCAAGGAGTCAACCTGGAGTTAACCCTCATCACTCTTGACAC
 TAACGAAAAGTTTAGGGATATCACCAAATTCATCCCAGAAAATATTAGTGCTGCCAGTG
 ATGCCACCCTAGTGATCAACACGGAGCACATGCCCTCAATGTTTGTCCCGGTGGGTGAC
 GTTGTGCAGTATGGCTTCTTGAACCTCAGTGGCAAGCCTACCCATCGCACCATGATGTA
 CAACTTTCTACTAAAGCAGGACAGTGTGGGGGAGTGGTGACATCTGTTGGGAAGATTA
 TCGGTATTACATTGGTGGCAATGGCAGACAAGGTTTTTGC GCAGGCCTCAAAGGAGT
 TACTTTGCTAGTGAACAAGGAGAGATCCAGTGGGTAAAGCCCAATAAAGAACTGGAAG
 ACTCAACATCAATGGACCAACCCGCACCAAGCTAGAACCCAGTGTATTCCATGATATCT
 TTGAGGGAAATAAGGAGCCAGCTGTCTTGACAGTAAAGACCCCGACTTGAGGTAGAT
 TTTGAACAGGCCCTGTTCTCTAAGTATGTGGGAATACACTATATGAGCCTGACGAGTA
 CATCAAAGAGGCAGCTCTTCATTATGCAAACCAATTAAAGCAGCTAGAAATCAACACCT
 CTCAAATGAGCATGGAGGAGGCCTGCTACGGTACTGAGAATCTTGAGGCTATTGATCTT
 CATACTAGTGCAGGTTACCCCTATAGTGCCCTGGGGATAAAGAAAAGAGACATCTTAGA
 CCCTACCACCAGGGACGTGAGTAAATGAAGTTCTACATGGACAAATATGGTCTTGATC
 TCCCTTACTCCACTTATGTCAAGGACGAGCTGCGCTCAATTGATAAAATTAAGAAAGGG
 AAGTCCCGTCTGATTGAGGCCAGTAGTTTAAATGATTCAGTGTACCTTAGAATGGCTTT
 CGGTCAATTTGTATGAGGCTTTCCACGCAAATCCTGGGACTATAACTGGATCAGCCGTGG
 GGTGTAACCTTGACACATTCTGGAGCAAGCTGCCAATTTTGCTCCCTGGTTCACCTCTT
 GCCTTTGACTACTCAGGTTATGATGCTAGCCTTAGCCCTGTCTGGTTCAGAGCATTAGA
 ATTGGTCCTTAGGGAGATAGGGTATAGTGAAGGGGCAGTCTCACTCATTGAGGGAATCA
 ACCACACACACCATGTGTATCGTAATAAGACCTATTGTGTGCTTGGTGGGATGCCCTCA
 GGCTGCTCGGGAACATCCATTTCAACTCAATGATCAACAACATTATTATCAGAGCACT
 GCTCATAAAAACATTTAAGGGCATTGATTTGGATGAACTCAACATGGTCGCTTATGGAG
 ATGATGTGCTCGCTAGCTACCCCTTCCCAATTGATTGCTTGGAGTTAGCGAAGACTGGC
 AAGGAGTATGGTCTAACCATGACCCCTGCGGATAAGTCTCCTTGCTTTAATGAAGTTAA
 TTGGGGTAATGCGACCTTTCTCAAGAGGGGCTTTTTACCCGATGAACAGTTTCCATTTT
 TGATCCACCCCACTATGCCAATGAGGGAGATCCATGAGTCCATTGATGGACCAAGGAT
 GCACGAAACACTCAAGATCATGTGCGGTCTTGTGCCTCCTAGCATGGCATAATGGTAA

GCAAGAATATGAGAAATTTGTGAGTACAATTAGGTCTGTCCCAGTGGGAAGAGCGTTGG
CTATCCCAAATTATGAAAACCTTAGACGTAATTGGCTCGAGTTATTT

Fig.9C (SEQ ID NO:3)

GPSLDFALSLLRRNIRQVQTDQGHFTMLGVRDRDLAVLPRHSQPCKTIWIEHKLVNILD
VELVDEQGVNLELTLITLDTNEKFRDITKFIPESTASDATLVINTEHMPSMFVPVGD
VVQYGFNLNLGKPTHRMTMYNFPTKAGQCGGVVTSVGKVIGIHIGGNRQGFCAGLKRS
YFASEQGEIQWVKPNKETGRNLNINGPTRKLEPSVFHDVFEKNKEPAVLHGKDPRLVD
FEQALFSKYVGNTLYEPDEYIKEAALHYANQLKQLEINTSQMSMEEACYGTENLEAIDL
HTSAGYPYSALGIKKRDILDPTTRDVSKMKSMDKYGLDLPYSTYVKDELRSIDKIKG
KSRLIEASSLNDVSVYLRMTFGHLYEAFHANPGTITGSAVGCNPDTFWSKLPILLPGSLF
AFDYSGYDASLSPVWFRALEMVLRIGYSEEAVSLIEGINHTHHVYRNKTYCVLGGMPS
GCSGTSIFNSMINNIIIRALLIKTFKIDLDLNMVAYGDDVLASYFPFIDCLELAKTG
KEYGLTMTTPADKSPCFNEVNWGNATFLKRGFLPDEQFPFLIHPTMPMREIHESIRWTKD
ARNTQDHVRSLLAWHNGKQEYEFVSTIRSVPIGRALAI PNENLRNWLLELF

Fig.9D (SEQ ID NO:4)

GGCCCGAGTCTTGATTTTGCTCTCTCCCTGTTAAGGAGGAACATCAGGCAAGTCCAAAC
AGACCAGGGGCATTTACCATGTTGGGTGTTAGGGATCGTTAGCAGTCCTCCCACGTC
ACTCACAACCCGGCAAACTATTTGGATCGAACACAACTCGTGAACATTCTTGATGCA
GTTGAATTGGTGGATGAGCAAGGAGTCAACCTGGAATTGACCCTCATCACTCTTGACAC
TAACGAAAAGTTTAGGGATATCACCAATTCATCCCAGAAAGTATTAGCACTGCCAGTG
ATGCCACCCTAGTGATCAACACGGAGCACATGCCCTCAATGTTTGTCCCGGTGGGTGAC
GTCGTGCAGTATGGCTTTTTGAATCTTAGTGGCAAGCCCACCCATCGCACCATGATGTA
CAACTTTCTACTAAAGCGGGACAGTGTGGAGGAGTAGTGACATCTGTTGGGAAAGTCA
TCGGTATTACATTGGTGGCAATGGTAGACAAGGTTTTTGCAGGCGCTCAAAAGGAGT
TACTTTGCTAGTGAACAAGGGGAGATCCAGTGGGTAAAGCCCAATAAAGAACTGGAAG
ACTCAACATCAATGGACCAACCCGCACCAAGTTGGAACCCAGTGTATTCCATGATGTCT
TCGAGGGAAATAAGGAACCAGCTGTCTTGACGGCAAGATCCCCGACTCGAGGTAGAT
TTTGAGCAGGCCCTGTTCTCTAAGTATGTGGGAAACACGCTATATGAGCCTGACGAGTA
CATCAAAGAGGCAGCTCTTCATTATGCAAATCAATTAAAGCAACTAGAAATTAATACCT
CCCAGATGAGCATGGAGGAAGCCTGCTATGGTACTGAGAATCTTGAGGCTATCGATCTT
CATACTAGTGCAGGTTACCCCTATAGTGCCCTGGGAATAAAGAAAAGAGACATCTTAGA
CCCTACCACCAGGGACGTGAGTAAATGAAATCCTATATGGACAAATATGGTCTCGATC
TCCCTTACTCCACTTATGTCAAGGATGAGCTGCGCTCAATTGATAAAATTAAGAAAGGG
AAGTCCCGTCTGATCGAGGCCAGCAGTTTAAATGATTCAAGTGTACCTCAGAATGACTTT
CGGTCAATTTGTATGAGGCTTTCCACGCAAATCCTGGGACGATAACTGGATCAGCCGTGG
GGTGTAACCCTGACACATTCTGGAGCAAGCTGCCAATCTTGCTTCCTGGTTCACCTTT
GCCTTTGACTACTCAGGTTATGATGCTAGCCTTAGCCCTGTCTGGTTCAGAGCATTAGA
AATGGTCCTTAGGGAGATAGGGTATAGTGAAGAGGCGGTCTCACTCATTGAGGGAATCA
ACCACACACACCACGTGTATCGTAACAAGACCTATTGTGTGCTTGGTGGGATGCCCTCA
GGCTGTTGCGGAACATCCATCTTCAACTCAATGATCAACAACATTATTATCAGAGCACT
GCTCATAAAAACATTTAAGGGCATTGATTTGGATGAACCTCAACATGGTTCGCTTATGGGG
ATGATGTGCTTGCTAGCTACCCCTTCCCAATCGATTGCTTGGAGTTAGCAAAGACTGGC
AAGGAGTATGGTCTGACCATGACTCCTGCAGATAAGTCCCCTTGCTTTAATGAAGTTAA

TTGGGGTAATGCGACCTTCCTCAAGAGGGGCTTTTTACCTGATGAGCAGTTTCCATTTT
 TGATCCACCCTACTATGCCAATGCGGGAGATCCATGAATCCATTTCGATGGACTAAGGAC
 GCACGAAACACTCAAGATCATGTACGGTCCTTGTGCCTCCTAGCATGGCATAATGGTAA
 GCAAGAATATGAAAAATTTGTGAGCACAATTAGGTCTGTCCCAATAGGAAGAGCTTTGG
 CTATCCCAAATTATGAAAACTTAGACGCAATTGGCTCGAGTTATTT

Fig.9E (SEQ ID NO:5)

MGSQVSTQSRSGSHENSNSATEGSTINYTTINYKDSYAATAGKQSLKQDPDKFANPVKD
 IFTEMAAPLKSPSAEACGYSDRVAQLTIGNSTITTQEAANIIVGYGEWPSYCSDSDATA
 VDKPTRPDVSVNRFYTLDTKLWEKSSKGWYWKFPDVLTTETGVFGQNAQFHYLYRSGFCI
 HVQCNAASKFHQGALLVAVLPEYVIGTVAGGTGTEDSHPPYKQTQPGADGFELQHPYVLD
 AGIPISQLTVCPHQWINLRTNNCATIIVPYINALPFDSALNHCNFGLLVVPISPLDYDQ
 GATPVIPIITITLAPMCSEFAGLRQAVTQGFPTTELKPGTNQFLTTDDGVSAFILPNFHPT
 PCIHIPGEVRNLELCQVETILEVNNVPTNATSLMERLRFPVSAQAGKGELCAVFRADP
 GRNGPWQSTLLGQLCGYYTQWSGSLEVTFMFTGSFMATGKMLIAYTPPGGPLPKDRATA
 MLGTHVIWDFGLQSSVTLVIPWISNTHYRAHARDGVFDYYTTGLVSIWYQTNVVPIGA
 PNTAYIIIALAAQKNFTMKLCKDASDILQGTIQGDRVADVIESSIGDSVSRALTQALP
 APTGQNTQVSSHRLDTGKVPALQAAEIGASSNASDESMIETRCVLNSHSTAETTLDSFF
 SRAGLVGEIDLPLEGTTNPNGYANWDIDITGYAQMRKVELFTYMRFDAEFTFVACTPT
 GEVVPQLLQYMFVPPGAPKPSRESLAWQTATNPVSFVKLSDPQAQVSVPFMSPASAYQ
 WFYDGYPTFGEHKQEKDLEYGACPNMMGTFSVRTVGTSSKSKYPLVVRIYMRMKHVRW
 IPRPMRNQNYLFGANPNYAGNSIKPTGTSRTAITTL

Fig.9F (SEQ ID NO:6)

ATGGGTTTCGCAGGTGTCCACGCAGCGCTCCGGTTCTCATGAAAATTCAAACTCAGCCAC
 CGAGGGTTCCACCATAAACTACACCACCATTAATTATTACAAAGACTCCTATGCTGCCA
 CAGCAGGCAAACAGAGTCTCAAGCAGGATCCAGACAAGTTTGCAAATCCTGTTAAAGAC
 ATCTTCACTGAAATGGCAGCGCCACTGAAGTCCCCATCCGCTGAGGCATGTGGATACAG
 TGATCGAGTAGCGCAATTAACATATTGGTAACTCCACCATCACCACGCAAGAAGCGGCTA
 ACATCATAGTTGGTTATGGTGAGTGGCCTTCCTACTGCTCGGATTCTGACGCTACAGCA
 GTGGATAAGCCAACGCGCCCGGATGTTTCAGTGAACAGGTTTTATACATTGGACACTAA
 ATTGTTGGGAGAAATCGTCCAAGGGATGGTACTGGAAATTCCTCGGATGTGTTAACTGAAA
 CTGGGGTTTTTGGGCAAAATGCACAATTCCTACTACCTCTACCGATCAGGGTTCTGTATC
 CACGTGCAGTGCAATGCTAGTAAATTCACCAAGGAGCACTCCTAGTCGCTGTTCTACC
 AGAGTACGTCAATTGGGACAGTGGCAGGCGGCACAGGGACGGAAGATAGTCACCCCCCTT
 ACAAGCAGACTCAACCCGGCGCCGATGGCTTCGAATTGCAACACCCGTACGTGCTTGAT
 GCTGGCATCCCAATATCACAGTTAACAGTGTGCCACATCAGTGGATTAATTTGAGAAC
 CAACAATTGTGCTACAATAATAGTGCCATACATTAACGCACTGCCTTTTGATTCCGCCT
 TGAACCACTGCAATTTTGGCCTATTAGTTGTGCCTATTAGCCCACTAGATTACGACCAA
 GGAGCGACGCCAGTAATCCCTATACTATCACATTAAGCCCCAATGTGTTCTGAATTCGC
 AGGTCTTAGGCAGGCAGTCACGCAAGGATTTCCACCGAGTTGAAACCTGGCACAAATC
 AATTTTTAACCACTGATGATGGCGTTTCAGCACCTATTCTACCAAACCTCCACCCACC
 CCGTGTATCCATATACCTGGTGAAGTTAGGAACTTGCTAGAGTTATGCCAGGTGGAAC
 CATTCTAGAGGTAAACAATGTGCCCACGAATGCCACTAGTTTAATGGAGAGACTGCGCT
 TTCCAGTCTCAGCACAAGCAGGGAAAGGTGAGCTGTGTGCGGTGTTTACAGAGCTGATCCT

GGGCGAAATGGGCCGTGGCAGTCCACCTTGCTGGGTGAGTTGTGTGGGTATTACACCCA
 ATGGTCAGGATCATTTGGAAGTCACCTTCATGTTTACTGGATCCTTTATGGCTACCGGCA
 AGATGCTCATAGCCTATACACCGCCAGGAGGCCCTTTGCCCCAAGGACCGGGCGACCGCC
 ATGTTGGGCACGCACGTCATCTGGGATTTTGGGCTGCAATCGTCCGTTACCTTGTAAT
 ACCATGGATCAGCAACACTCACTACAGAGCGCATGCCGAGATGGAGTGTGTTGACTACT
 ACACCACAGGGTTAGTCAGTATATGGTATCAGACAAATTACGTGGTTCCAATTGGGGCG
 CCTAATACAGCCTATATAATAGCACTAGCGGCAGCCCCAAAAGAATTTCACTATGAAGTT
 GTGCAAGGATGCTAGTGATATCCTACAAACGGGCACCATCCAGGGAGATAGGGTAGCAG
 ATGTAATTGAAAGTTCCATAGGGGATAGCGTGAGCAGAGCCCTCACTCAAGCTCTACCA
 GCACCCACAGGCCAGAACACACAGGTGAGCAGTCATCGACTGGATACAGGCAAGGTTCC
 AGCACTCCAAGCTGCTGAAATTGGAGCATCATCAAATGCTAGTGACGAGAGCATGATCG
 AGACACGCTGTGTTCTTAACTCGCACAGCACAGCTGAGACCACTCTTGATAGTTTCTTC
 AGCAGAGCGGGATTAGTTGGAGAGATAGATCTTCTCTTGAAGGCACAATAACCCAAA
 TGGTTATGCCAACTGGGACATAGATATAACAGGTTACGCACAAATGCGCAGAAAGGTGG
 AGTTATTACCTACATGCGCTTTGATGCAGAGTTCACTTTTCGTTGCGTGACACCTACC
 GGGGAAGTTGTCCACAATTGCTCCAATATATGTTTGTACCACCTGGAGCCCCCTAAGCC
 AGACTCCAGGGAGTCCCTCGCATGGCAAACCGCCACCAACCCCTCAGTTTTTGTCAAGT
 TGTGACACCTCCAGCACAGGTTTCAGTACCATTTCATGTCACCCGCGAGTGCTTACCAA
 TGGTTCTATGACGGATATCCACATTGGGGAACACAAACAGGAGAAAGATCTTGAGTA
 TGGGGCGTGCCCTAATAACATGATGGGTACGTTCTCAGTGCGGACTGTAGGGACTTCCA
 AATCCAAGTATCCTTTAGTGTTAGGATTTACATGAGGATGAAGCACGTCAGGGCGTGG
 ATACCTCGCCCGATGCGTAACCAAACTACCTATTCAAGGCCAACCCAAATTATGCTGG
 CAACTCCATTAAGCCAACTGGTACTAGTCGCACAGCGATCACTACTCTT

Fig.9G (SEQ ID NO:7)

GGACCAGGGTTCGATTACGCAGTGGCTATGGCTAAAAGAAACATTGTTACAGCAACTAC
 TAGCAAGGGAGAGTTCACTATGTTAGGAGTCCACGACAACGTGGCTATTTTACCAACCC
 ACGCTTACCTGGTGAAAGCATTGTGATCGATGGCAAAGAAGTGGAGATCTTGATGCC
 AAAGCGCTCGAAGATCAAGCAGGAACCAATCTTGAAATCACTATAATCACTCTAAAGAG
 AAATGAAAAGTTTCAGAGACATTAGACCACATATACCTACTCAAATCACTGAGACAAATG
 ATGGAGTCTTGATCGTGAACACTAGCAAGTACCCCAATATGTATGTTTCTGTCGGTGCT
 GTGACTGAACAGGGATATCTAAATCTCGGTGGGCGCCAACTGCTCGTACTCTAATGTA
 CAACCTTCCAACCAGAGCAGGACAGTGTGGTGGAGTCATCACATGTACTGGGAAAGTCA
 TCGGGATGCATGTTGGTGGGAACGGTTCACACGGGTTTGCAGCGGCCCTGAAGCGATCA
 TACTTCACTCAGAGTCAAGGTGAAATCCAGTGGATGAGACCTTCGAAGGAAGTGGGATA
 TCCAATCATAAATGCCCCGTCCAAAACCAAGCTTGAACCCAGTGCTTTCCACTATGTGT
 TTGAAGGGGTGAAGGAACAGCAGTCCTCACTAAAAACGATCCCAGGCTTAAGACAGAC
 TTTGAGGAGGCAATTTTCTCCAAGTACGTGGGTAACAAAATTACTGAAGTGGATGAGTA
 CATGAAAGAGGCAGTAGACCACTATGCTGGCCAGCTCATGTCACTAGACATCAACACAG
 AACAAATGTGCTTGGAGGATGCCATGTATGGCACTGATGGTCTAGAAGCACTTGATTTG
 TCCACCAGTGCTGGCTACCTTATGTAGCAATGGGAAAGAAGAAGAGAGACATCTTGAA
 CAAACAAACCAGAGACACTAAGGAAATGCAAAAACCTGCTCGACACATATGGAATCAACC
 TCCCCTGGTGACTTATGTAAAGGATGAACCTTAGATCCAAAACAAAGGTTGAGCAGGGG
 AAATCCAGATTAATTGAAGCTTCTAGTTTGAATGACTCAGTGGCAATGAGAATGGCTTT
 TGGGAACCTATATGCTGCTTTTCAAAAACCCAGGAGTGATAACAGGTTTCAGCAGTGG
 GGTGCGATCCAGATTTGTTTTGGAGCAAAATTCCGGTATTGATGGAAGAGAAGCTGTTT

GCTTTTGACTACACAGGGTATGATGCATCTCTCAGCCCTGCTTGGTTCGAGGCACATAAA
 GATGGTGCTTGAGAAAAATCGGATTCGGAGACAGAGTTGACTACATCGACTACCTAAACC
 ACTCACACCACCTGTACAAGAATAAAACATACTGTGTCAAGGGCGGTATGCCATCTGGC
 TGCTCAGGCACTTCAATTTTTTAACCAATGATTAACAACCTTGATTATCAGGACACTCTT
 ACTGAAAACCTACAAGGGCATAGATTTAGACCACCTAAAAATGATTGCCTATGGTGATG
 ATGTAATTGCTTCCTACCCCATGAAGTTGACGCTAGTCTCCTAGCCCAATCAGGAAAA
 GACTATGGACTAACTATGACTCCAGCTGACAAATCAGCTACATTTGAAACAGTCACATG
 GGAGAATGTAACATTCTTGAAGAGATTCTTCAGGGCAGACGAGAAATACCCATTTCTTA
 TTCATCCAGTAATGCCAATGAAGGAAATTCATGAATCAATTAGATGGACTAAAGATCCT
 AGGAACACTCAGGATCACGTTTCGCTCTCTGTGCCTTTTAGCTTGGCACAATGGCGAAGA
 AGAATATAACAAATTCCTAGCTAAATCAGGAGTGTGCCAATTGGAAGAGCTTTATTGC
 TCCCAGAGTACTCAACATTGTACCGCCGTTGGCTTGACTCATTT

Fig.9H (SEQ ID NO:8)

PGPGFDYAVAMAKRNIVTATTSKGEFTMLGVHDNVAILPHTASPGESIVIDGKEVEILDA
 KALEDQAGTNLEITIIITLKRNEKFRDIRPHIPTQITETNDGVLIVNTSKYPNMYVPVGA
 VTEQGYLNLGGRQTARTLMYNFPTRAGQCQGVITCTGKVIGMHVGGNGSHGFAAALKRS
 YFTQSQGEIQWMRPSKEVGYPIINAPSKTKLEPSAFHYVFEVKEPAVLTKNDPRLKTD
 FEEAIFSKYVGNKITEVDEYMKAEVDHYAGQLMSLDINTEQMCLEDAMYGTGDLALDL
 STSAGYPYVAMGKKRDI LNKQTRDTKEMQKLLDTYGINLPLVITYVKDELRSKTKVEQG
 KSRLIEASSLNDVSAMRMAFGNLYAAFHKNPGVITGSAVGCDPDLFWSKIPVLMEEKLF
 AFDYTGYSASLSPAWFEALKMVLEKIGFGDRVDYIDYLNHSHHLYKNKTYCVKGGMPSG
 CSGTSIFNSMINNLIIRTL LKTYKGIDLDHLKMIAYGDDVIASYPHEVDASLLAQSGK
 DYGLTMTPADKSATFETVTWENVTF LKRFFRADEKYPFLIHPVMPMKEIHESIRWTKDP
 RNTQDHRVSLCLLAWHNGEEYNKFLAKIRSVPIGRALLLPEYSTLYRRWLDSF

Fig.9I (SEQ ID NO:9)

ATGGGTGCTCAGTTTTCATCACAGAAAGTGGGCGCACATGAAAACCTCAAATAGAGCGTA
 TGGTGCTTCTACCATTAAATTACACCACCTTAATTATTATAGAGATTCAGCTAGTAACG
 CGGCTTCGAAACAGGACTTCTCTCAAGACCCTTCCAAGTTCACCGAGCCCATCAAGGAT
 GTCCTGATAAAAACAGCCCCAATGCTAAACTCGCCAAACATAGAGGCTTGCGGGTATAG
 CGATAGAGTACTGCAATTAACACTGGGAAACTCCACTATAACCACACAGGAGGCGGCTA
 ATTCAGTAGTCGCTTATGGGCGTTGGCCTGAATATCTGAGGGACAGCGAAGCCAATCCA
 GTGGACCAGCCGACAGAACCAGACGTCGCTGCATGCAGGTTTTATACGCTAGACACCGT
 GTCTTGGACGAAAGAGTCGCGAGGGTGGTGGTGGAAAGTTGCCTGATGCACTGAGGGACA
 TGGGACTCTTTGGGCAAAATATGTACTACCTACCTAGGTAGGTCCGGGTACACCGTG
 CATGTACAGTGTAACGCCTCCAAATTCACCAGGGGGCACTAGGGGTATTCGCCGTACC
 AGAGATGTGTCTGGCCGGGGATAGCAACACCACTACCATGCACACCAGCTATCAAATG
 CCAATCCTGGCGAGAAAGGAGGCACTTTCACGGGTACGTTCACTCCTGACAACAACCAG
 ACATCACCTGCCCGCAGGTTCTGCCCGGTGGATTACCTCCTTGGAATGGCACGTTGTT
 GGGGAATGCCTTTGTGTTCCCGCACCAGATAATAAACCTACGGACCAACAACCTGTGCTA
 CACTGGTACTCCCTTACGTGAACTCCCTCTCGATAGATAGTATGGTAAAGCACAATAAT
 TGGGGAATTGCAATATTACCATTGGCCCCATTAAATTTTGCTAGTGAGTCCTCCCCAGA
 GATTCCAATCACCTTGACCATAGCCCCTATGTGCTGTGAGTTCAATGGATTAAGAAACA

TCACCCTGCCACGCTTACAGGGCCTGCCGGTCATGAACACCCCTGGTAGCAATCAATAT
 CTTACTGCAGACAACTTCCAGTCACCGTGTGCGCTGCCTGAATTTGATGTGACCCCACC
 TATTGACATAACCCGGTGAAGTAAAGAACATGATGGAATTGGCAGAAATCGACACCATGA
 TTCCCTTTGACTTAAGTGCCACAAAAAGAACACCATGGAAATGTATAGGGTTTCGGTTA
 AGTGACAAACCACATACAGACGATCCCATACTCTGCCTGTCACTCTCTCCAGCTTCAGA
 TCCTAGGTTGTACATACTATGCTTGGAGAAATCCTAAATTACTACACACACTGGGCAG
 GATCCCTGAAGTTCACGTTTCTGTTCTGTGGATTTCATGATGGCAACTGGCAAACCTGTTG
 GTGTCATACGCGCCTCCTGGAGCCGACCCACCAAAGAAGCGTAAGGAGGCGATGTTGGG
 AACACATGTGATCTGGGACATAGGACTGCAGTCCTCATGTACTATGGTAGTGCCATGGA
 TTAGCAACACCACGTATCGGCAAACCATAGATGATAGTTTACCCGAAGGCGGATACATC
 AGCGTCTTCTACCAAACCTAGAATAGTCGTCCCTCTTTTCGACACCCAGAGAGATGGACAT
 CCTTGGTTTTGTGTGTCAGCGTGTAATGACTTCAGCGTGCCTTGTGCGAGATACACAC
 ATATAGAGCAAAAAGCGCTAGCACAGGGGTTAGGTCAGATGCTTGAAAGCATGATTGAC
 AACACAGTCCGTGAAACGGTGGGGGCGGCAACATCTAGAGACGCTCTCCCAAACACTGA
 AGCCAGTGGACCAACACACTCCAAGGAAATTCGGGCACTCACCGCAGTGGAAACTGGGG
 CCACAAATCCACTAGTCCCTTCTGATACAGTGCAAACCAGACATGTTGTACAACATAGG
 TCAAGGTCAGAGTCTAGCATAGAGTCTTCTTCGCGCGGGGTGCATGCGTGACCATTAT
 GACCGTGGATAACCCAGCTTCCACCACGAATAAGGATAAGCTATTTGCAGTGTGGAAGA
 TCACTTATAAAGATACTGTCCAGTTACGGAGGAAATTGGAGTCTTTCACCTATTCTAGA
 TTTGATATGGAACCTTACCTTTGTGGTTACTGCAAATTTCACTGAGACTAACAATGGGCA
 TGCCTTAAATCAAGTGTACCAAATTATGTACGTACCACCAGGCGCTCCAGTGCCCGAGA
 AATGGGACGACTACACATGGCAAACCTCATCAAATCCATCAATCTTTTACACCTACGGA
 ACAGCTCCAGCCCGGATCTCGGTACCGTATGTTGGTATTTTCGAACGCCTATTCACACTT
 TTACGACGGTTTTTCCAAAGTACCACTGAAGGACCAGTCCGGCAGCACTAGGTGACTCCC
 TTTATGGTGCAGCATCTCTAAATGACTTCGGTATTTTGGCTGTTAGAGTAGTCAATGAT
 CACAACCCGACCAAGGTCACCTCCAAATCAGAGTGTATCTAAAACCCAAACACATCAG
 AGTCTGGTGCCCGCTCCACCGAGGGCAGTGGCGTACTACGGCCCTGGAGTGGATTACA
 AGGATGGTACGCTTACACCCCTCTCCACCAAGGATCTGACCACATAT

Fig.9J (SEQ ID NO:10)

MGAQVSSQKVGAHENSNRAYGGSTINYTTINYRDSASNAASKQDFSQDPSKFTEPIKD
 VLIKTAPMLNSPNIACGYSDRVLQLTLGNSTITTQEAANSVVAYGRWPEYLRDSEANP
 VDQPTPEPDVAACRFYTLDTVSWTKESRGWWWKLPDALRDMGLFGQNMYYHYLGRSGYTV
 HVQCNASKFHHQALGVFAVPEMCLAGDSNTTMMHTSYQANANPGEKGGTFTGTFTPDNNQ
 TSPARRFCPVDYLLGNLTLLGNAFVFPHQIINLRTNNCATLVLPYVNSLSIDSMVKHNN
 WGIAILPLAPLNFASESSPEIPITLTIAPMCCEFNGLRNITLPRLOGLPVMNTPGSNQY
 LTADNFQSPCALPEFDVTPPIDIPGEVKNMMELEIDTMI PFDLSATKNTMEMYRVRL
 SDKPHTDDPILCLSLSPASDPRLSHTMLGEILNYYTHWAGSLKFTFLFCGFMMATGKLL
 VSYAPPGADPPKKRKEAMLGTHVIWDIGLQSSCTMVVPWISNTTYRQTIDDSFTEGGYI
 SVFYQTRIVVPLSTPREMDILGFVSACNDFSVRLLRDTTHIEQKALAQGLGQMLESMID
 NTVRETVGAATSRDALPNTEASGPTHSKEIPALTAVETGATNPLVPSDTVQTRHVVQHR
 SRSESSIESFFARGACVTIMTVDNPASTTNKDKLFAVWKITYKDTVQLRRKLEFFTYSR
 FDMELTFVVTANFTETNNGHALNQVYQIMYVPPGAPVPEKWDDYTWQTSSNPSIFYTYG
 TAPARISVPYVGISNAYSHFYDGF SKVPLKQSAALGDSLYGAASLNDFGILAVRVVND
 HNPTKVT SKIRVYLKPKHIRVWCPRPPRAVAYYGPGVDYKDGTLTPLSTKDLTTY