



(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)



1-0030754

(51)⁷ A61K 39/145; C12N 7/00; A61P 31/04 (13) B

(21) 1-2015-00535

(22) 11/07/2013

(86) PCT/US2013/050000 11/07/2013

(87) WO2014/014726 23/01/2014

(30) 61/672,398 17/07/2012 US

(45) 25/01/2022 406

(43) 27/04/2015 325A

(73) MERIAL LIMITED (US)

3239 Satellite Blvd., Duluth, GA 30096, United States of America

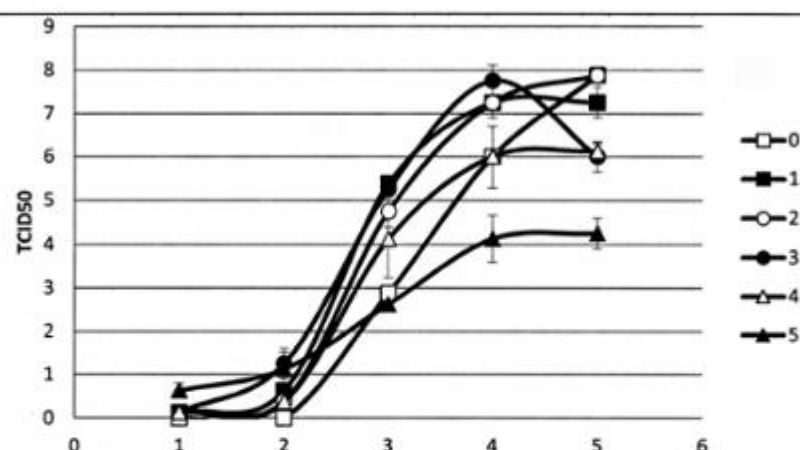
(72) HAUSE, Benjamin, M. (US).

(74) Công ty Luật TNHH Phạm và Liên danh (PHAM & ASSOCIATES)

(54) CHỨNG VIRUT CÚM LỢN NHƯỢC ĐỘC VÀ CHẾ PHẨM MIỄN DỊCH CHỨA
CHỨNG VIRUT NÀY

(57) Sáng chế đề cập đến chủng virut cúm lợn nhược độc, đặc biệt là chủng được tạo ra thông qua phương pháp di truyền ngược, chế phẩm chứa chủng này, và phương pháp tạo ra chủng này. Chủng nhược độc này được xử lý kỹ thuật để mã hóa protein HA có các vị trí glycosyl hóa bổ sung so với protein HA được mã hóa bởi virut gốc có tính độc tương ứng. Chủng virut cúm nhược độc này có thể được dùng một cách thuận lợi.

Nghiên cứu sinh trưởng của chủng tăng glycosyl hóa. Các đường thể hiện các thể đột biến có các vị trí glycosyl hóa được liên kết N bổ sung được xử lý kỹ thuật



	Virut thu được	Vị trí glycosyl hóa bổ sung (n)
0	Virut gốc RG (10-0036-2)	0
1	S71N	1
2	S71N, K90N	2
3	S71N, K90N, L173T	3
4	S71N, K90N, L173T, P287T	4
5	S71N, K90N, L173T, P287T, K294T	5

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến vaccin chứa virus nhược độc, cụ thể là vaccin có tác dụng bảo vệ rộng, an toàn, và hiệu quả cho loài lợn chống bệnh nhiễm trùng/bệnh do virus cúm lợn gây ra. Sáng chế còn đề cập đến các phương pháp tạo ra virus nhược độc, và xác định các biến đổi của axit nucleic có liên quan đến việc giảm độc lực của virus.

Sáng chế còn đề cập đến chế phẩm sinh miễn dịch hoặc vaccin chứa virus theo sáng chế; ví dụ, virus sống nhược độc. Virus này cũng có thể được làm vô hoạt trong chế phẩm; nhưng có thể có lợi nếu virus này là virus cúm lợn sống nhược độc. Tiếp theo, sáng chế đề cập đến phương pháp sản xuất và/hoặc bào chế chế phẩm này; ví dụ, nuôi cấy hoặc phát triển hoặc nhân giống virus trên hoặc trong môi trường thích hợp, thu hoạch virus này, tùy ý làm vô hoạt virus này, và tùy ý trộn với chất mang, tá dược, chất pha loãng hoặc chất dẫn và/hoặc tá chất và/hoặc chất ổn định dùng trong thú y hoặc được dùng thích hợp.

Tình trạng kỹ thuật của sáng chế

Virus cúm có một số cơ chế để tránh đáp ứng miễn dịch của vật chủ chống hemagglutinin (HA) là chất trợ giúp cho sự xâm nhập của virus. Việc sắp xếp lại các đoạn của hệ gen có thể dẫn đến thay đổi tính kháng nguyên đồng thời tích lũy dần dần các đột biến dẫn đến biến động di truyền cuốn trôi (Desselberger U, và các đồng tác giả, 1978, Schild G, và các đồng tác giả, 1974). Ngoài đột biến trong các epitop của kháng nguyên HA có thể ảnh hưởng đến khả năng nhận diện HA đột biến của các kháng thể đã tồn tại từ trước, các đột biến trong quá trình glycosyl hóa được liên kết N của HA có thể giúp virus tránh bị nhận diện bởi kháng thể bằng cách thay đổi lớp oligosacarit bao quanh HA (Schulze IT, 1997). Các gốc glycan có thể hạn chế việc gắn kết một số kháng thể với epitop của chúng, dẫn đến làm mất khả năng nhận diện và sinh miễn dịch của kháng thể trong hiện tượng được gọi là che chắn glycan (Wanzeck K, và các đồng tác giả, 2011, Wei C-J, và các đồng tác giả, 2010). Ngoài ra, sẽ tốt hơn nếu đột biến và biến động di truyền xảy ra ở các vị trí trong đầu hình cầu của HA không được glycan bảo vệ (Das SR, và các đồng tác giả, 2010). Các gốc glycan còn ảnh hưởng đến việc gắn kết thụ thể và do đó có thể ảnh hưởng đến động học sao chép của virus.

Sự tiến hóa của virus cúm thường bao gồm các thay đổi về số lượng và địa điểm của các vị trí glycosyl hóa (Long J, và các đồng tác giả, 2011). Đối với virus H3N2, số vị trí glycosyl hóa đã tăng từ hai đến mười trong 40 năm gần đây (4). Tuy nhiên, HA được glycosyl hóa tăng dần có liên quan đến sự giảm độc lực ở chuột và giảm khả năng sao chép của virus (Das SR, và các đồng tác giả, 2011, Vigerust DJ, và các đồng tác giả, 2007). Trong khi việc glycosyl hóa tăng dần có thể bảo vệ HA khỏi các kháng thể trung hòa, thì ái lực của virus đối với các thụ thể tế bào đường nhiên bị giảm (Abe Y, và các đồng tác giả, 2004, Das SR, và các đồng tác giả, 2011). Ngoài ra, các đột biến bù trong HA hoặc neuraminidaza (NA) là cần thiết để cân bằng việc gắn kết thụ thể và giải phóng hoạt tính (Wagner R, và các đồng tác giả, 2000). Các đột biến bù này trong NA liên quan đến việc thu nhận tính kháng tự nhiên với các chất ức chế NA (Hensley SE, và các đồng tác giả, 2011).

Nhiều giống di truyền và khác biệt về mặt kháng nguyên của virus cúm lợn hiện đang lưu hành ở lợn, bao gồm các virus H1N1 và H1N2 nhóm β , γ và δ , cũng như phân typ H3N2 (Lorusso A, và các đồng tác giả, 2012). Nhóm δ ban đầu được chia nhỏ thành hai phân nhóm (δ -I và δ -II), tuy nhiên, nghiên cứu gần đây đã chứng minh năm phân nhóm di truyền khác biệt (δ -A, δ -B, δ -C, δ -D, δ -E) thể hiện ít nhất ba nhóm khác biệt về mặt kháng nguyên (Hause BM, và các đồng tác giả, 2011, Vincent AL, và các đồng tác giả, 2009).

Trong khi các vaccin chứa virus chết có hiệu quả khi các chủng vaccin này tương hợp với virus trên thực địa, thì mức độ đột biến nhanh của virus cúm đòi hỏi phải thay đổi chủng thường xuyên để bảo đảm việc tương hợp di truyền và kháng nguyên giữa chủng vaccin và các virus đang lưu hành. Ngoài ra, nhiều phân typ và giống đang lưu hành đồng thời ở lợn thường cần năm chủng vaccin hoặc nhiều hơn để bao gồm các đại diện của mỗi loại. Vaccin sống đơn tạo ra phổ miễn dịch rộng đem lại tiến bộ đáng kể trong các chương trình sức khỏe đàn lợn.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất vaccin nhược độc cũng như mô tả phương pháp điều trị và dự phòng nhiễm virus cúm lợn.

Theo một phương án, vaccin này chứa virus cúm nhược độc, đã được biến đổi để biểu hiện gen HA có các vị trí glycosyl hóa bổ sung so với chủng gốc của nó.

Theo một phương án cụ thể, các gen HA và/hoặc sản phẩm chứa gen này có các biến đổi giống nhau như được nêu trong các trình tự SEQ ID NO:17 và 21;

SEQ ID NO:17 và 23; SEQ ID NO:18 và 22; hoặc SEQ ID NO:18 và 24. Theo một phương án khác nữa, các gen HA và/hoặc sản phẩm chứa gen này có ít nhất 1 trong số 5 biến đổi. Theo phương án cụ thể hơn, các gen HA và/hoặc sản phẩm chứa gen này có 4 hoặc 5 biến đổi, dẫn đến +4 và +5 vị trí glycosyl hóa, so với các sản phẩm chứa gen HA gốc.

Một mục đích khác của sáng chế là đề xuất ADN bổ trợ và/hoặc plasmit để sử dụng trong kỹ thuật di truyền ngược nhằm tạo ra virus cúm nhược độc theo sáng chế. Theo một phương án, ít nhất một trong số các ADN bổ trợ và/hoặc plasmit chứa trình tự HA có số vị trí glycosyl hóa tăng so với trình tự như được nêu trong SEQ ID NO:18.

Theo một phương án cụ thể, trình tự HA là như được nêu trong các SEQ ID NO:21 hoặc 23.

Sáng chế còn đề xuất các chủng virus cúm nhược độc mới, tạo ra tác dụng miễn dịch bảo vệ an toàn, có hiệu quả, và rộng. So với chủng virus cúm gốc, các chủng nhược độc có thể có các vị trí glycosyl hóa bổ sung và/hoặc các vị trí glycosyl hóa giả định được mã hóa bởi các gen HA của chúng, mà sự có mặt của chúng liên quan đến độc lực giảm.

Do đó, sáng chế đề xuất virus đột biến chứa (các) đột biến trong một hoặc nhiều trình tự axit nucleic, so với virus kiểu dại/gốc, chúng làm cho virus đột biến này bị nhược độc so với các virus gốc, virus gốc này chứa các axit nucleic mã hóa protein HA kiểu dại. Trong bản mô tả này, protein HA “kiểu dại” là protein có số vị trí glycosyl hóa và/hoặc số vị trí glycosyl hóa giả định giống với trình tự như được nêu trong SEQ ID NO:18.

Theo một phương án cụ thể, virus đột biến chứa trình tự axit nucleic vRNA tương ứng với (tức là bù ngược và có uraxil thay cho thymidin) trình tự ADN được nêu trong các SEQ ID NO:21 hoặc 23, mã hóa các peptit lần lượt như được nêu trong các SEQ ID NO:22 hoặc 24, và làm cho virus đột biến này bị nhược độc/không có tính độc so với virus kiểu dại/gốc có tính độc.

Trong bản mô tả này, thuật ngữ “gen” sẽ được sử dụng với nghĩa rộng, và sẽ bao gồm cả các trình tự mã hóa và không mã hóa (tức là các trình tự điều hòa trước và sau, trình tự khởi động, 5’/3’ UTR, các intron, và exon). Khi chỉ đề cập đến trình tự mã hóa của gen, thì thuật ngữ “trình tự mã hóa của gen” hoặc “CDS” (coding sequence) sẽ được sử dụng thay thế cho nhau trong suốt bản mô tả này. Khi axit nucleic cụ thể được bàn luận, ví dụ, trình tự như được nêu trong SEQ ID NO:17 (tương đương trình tự ADN của sợi “có nghĩa” của ARN bổ sung của virus gốc), người có hiểu biết trung bình trong lĩnh vực sẽ có ngay tất cả các dạng có

thể thu được từ trình tự này (ARN thông tin, vARN, ARN bổ sung, ADN, protein, v.v.). Ví dụ, virus cúm là virus có ARN đơn sợi (single strand RNA - ssRNA) âm. Để sao chép, ssRNA âm của nó (trong bản mô tả này được gọi là “vARN”) phải được phiên mã thành ARN dương hoặc có nghĩa (trong bản mô tả này được gọi là “ARN bổ sung”). Bộ máy tế bào chủ được kết nạp để sử dụng ARN bổ sung để tạo ra các protein của virus và vARN. Người có hiểu biết trung bình trong lĩnh vực kỹ thuật này bằng cách sử dụng mã di truyền đã biết có thể dễ dàng thu được vARN, ARN bổ sung, và các trình tự peptit từ trình tự ADN.

Theo một phương án cụ thể, các vaccine nhược độc còn chứa tá chất. Tá chất có thể là chất bất kỳ làm tăng và/hoặc tăng thêm đáp ứng miễn dịch so với vaccine nhược độc đơn lẻ. Các tá chất dùng qua niêm mạc, bao gồm chitosan và các dẫn xuất của nó, là đặc biệt hữu ích đối với các vaccine nhược độc dùng qua đường miệng theo sáng chế.

Sáng chế còn mô tả phương pháp gây đáp ứng miễn dịch (hoặc sinh miễn dịch) hoặc bảo vệ chống virus cúm, cũng như phương pháp phòng ngừa hoặc điều trị bệnh cúm, hoặc (các) tình trạng bệnh do virus cúm gây ra, bao gồm bước cho các động vật có nhu cầu dùng virus nhược độc, hoặc chế phẩm chứa virus nhược độc này.

Bộ kit chứa ít nhất là chủng virus cúm nhược độc và hướng dẫn sử dụng cũng được đề xuất.

Các phương án này và các phương án khác được bộc lộ hoặc là hiển nhiên từ và được bao gồm trong phần mô tả chi tiết dưới đây.

Mô tả vắn tắt các hình vẽ

Sáng chế, bao gồm các phương án tốt nhất đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật này được mô tả đầy đủ và rõ ràng trong phần còn lại của bản mô tả với chú giải bằng các hình vẽ kèm theo, trong đó:

Fig. 1 thể hiện kiểu gen của virus được tạo ra bằng cách di truyền ngược;

Fig. 2 là sự bắt cặp axit amin giữa SEQ ID NO:18 (protein HA của virus cúm gốc) và SEQ ID NO:22 (protein HA của thể đột biến glycosyl hóa n+5);

Fig. 3 là đồ thị sự sinh trưởng của các tế bào nuôi cấy của virus cúm gốc và virus cúm thể đột biến glycosyl hóa có 1 đến 5 vị trí glycosyl hóa bổ sung trong gen HA của chúng so với virus cúm gốc;

Fig. 4 là bảng liệt kê các SEQ ID trong sáng chế.

Mô tả chi tiết sáng chế

Sáng chế đề xuất các trình tự nucleotit và các gen liên quan đến việc làm giảm độc lực của vi sinh vật, như virus, ví dụ, virus cúm, các sản phẩm (ví dụ, protein, kháng nguyên, chất kháng nguyên, epitop) được mã hóa bởi các trình tự nucleotit, phương pháp tạo ra các trình tự nucleotit này, các sản phẩm, vi sinh vật này, và sử dụng chúng, như để bào chế vaccin hoặc các chế phẩm sinh miễn dịch hoặc để gây ra đáp ứng miễn dịch học hoặc đáp ứng miễn dịch hoặc làm vector, ví dụ, làm vector biểu hiện (ví dụ, vector biểu hiện in vitro hoặc in vivo).

Các đột biến được đưa vào các trình tự nucleotit và gen của vi sinh vật tạo ra các thể đột biến nhược độc mới và không hiển nhiên. Các thể đột biến này là hữu ích để sản xuất chế phẩm sinh miễn dịch sống nhược độc hoặc vaccin sống nhược độc có mức độ sinh miễn dịch cao.

Việc nhận diện các đột biến tạo ra các trình tự nucleotit và các gen mới và không hiển nhiên, cũng như các sản phẩm gen mới và không hiển nhiên được mã hóa bởi các trình tự nucleotit và các gen này.

Theo một phương án, sáng chế đề xuất chủng virus cúm lợn tăng glycosyl hóa nhược độc có khả năng tạo ra đáp ứng miễn dịch an toàn và hiệu quả ở loài lợn chống virus cúm hoặc các bệnh do virus cúm gây ra. Theo một phương án, chủng này có thể mã hóa gen HA có ít nhất 1 vị trí glycosyl hóa bổ sung so với chủng gốc có tính độc.

Theo một phương án khác, chủng nhược độc này có thể mã hóa gen HA có 4 hoặc 5 vị trí glycosyl hóa bổ sung so với chủng gốc có tính độc. Theo một phương án cụ thể, các vị trí glycosyl hóa của chủng này được chọn từ S71N, K90N, L173T, P287T, và K294T, với vị trí của các thay đổi của axit amin dựa trên gen HA có trình tự như được nêu trong SEQ ID NO:18.

Theo một phương án, protein HA được tạo ra bởi chủng virus cúm nhược độc có trình tự như được nêu trong SEQ ID NO:22 hoặc trình tự có độ đồng nhất bằng ít nhất 90% so với SEQ ID NO:22 có các vị trí sau có các axit amin như sau: 71N, 90N, 173T, 287T, và 294T.

Theo một phương án khác, virus cúm nhược độc này tạo ra protein HA có trình tự như được nêu trong SEQ ID NO:24 hoặc trình tự có độ đồng nhất bằng ít nhất 90% so với SEQ ID NO:24 có các vị trí sau có các axit amin như sau: 71N, 90N, 173T, 287T.

Theo một khía cạnh khác, sáng chế đề xuất chế phẩm miễn dịch chứa chủng virus cúm nhược độc, mã hóa protein HA tăng glycosyl hóa. Theo một phương án, chế phẩm này có thể còn chứa chất dẫn, chất pha loãng hoặc tá dược được dùng hoặc dùng trong thú y.

Theo một phương án, chế phẩm này tạo ra đáp ứng miễn dịch bảo vệ ở loài lợn chống nhiễm virus cúm lợn có tính độc. Theo một số phương án, chế phẩm này còn chứa ít nhất một kháng nguyên bổ sung liên quan đến mầm bệnh không phải là virus cúm lợn.

Theo một phương án khác, ít nhất một kháng nguyên bổ sung này được chọn từ M. hyo, PCV2, PRRSV, SIV hoặc mầm bệnh khác có khả năng nhiễm và gây ra chứng bệnh hoặc mãn cảm với chứng bệnh ở loài lợn, hoặc kết hợp của chúng.

Theo một phương án, sáng chế mô tả phương pháp chủng ngừa động vật bao gồm ít nhất một lần dùng chế phẩm chứa các trình tự mã hóa protein HA tăng glycosyl hóa của virus cúm. Theo một phương án khác, loài lợn này là lợn nái từ 3 tuần đến 6 tuần trước khi đẻ. Theo một phương án khác nữa, lợn con thu được có thể có tỷ lệ mắc bệnh và/hoặc tỷ lệ chết giảm so với lợn con thu được từ lợn nái chưa được chủng ngừa.

Theo một phương án khác, sáng chế đề xuất chế phẩm chứa nhiều vector để tạo ra virus cúm lợn nhược độc bao gồm vector chứa trình tự khởi động được liên kết có điều khiển với ADN bổ sung của HA của virus cúm, trong đó ADN bổ sung của HA mã hóa các vị trí glycosyl hóa bổ sung so với HA được mã hóa bởi chủng virus cúm lợn gốc có tính độc. Theo một phương án cụ thể, các vị trí glycosyl hóa bổ sung được chọn từ S71N, K90N, L173T, P287T, và K294T, và vị trí của các thay đổi của axit amin là dựa trên gen HA có trình tự như được nêu trong SEQ ID NO:18.

Theo một phương án, ADN bổ sung của HA để tạo ra virus cúm nhược độc mã hóa protein như được nêu trong SEQ ID NO:22. Theo một phương án khác, ADN bổ sung của HA mã hóa protein như được nêu trong SEQ ID NO:24.

Theo một phương án, sáng chế mô tả phương pháp tạo ra virus cúm, bao gồm bước: cho tế bào tiếp xúc với một trong các chế phẩm theo sáng chế với lượng hữu hiệu để thu được virus cúm có tính lây nhiễm. Theo một phương án, phương pháp này còn bao gồm bước phân lập virus.

Theo một phương án khác, sáng chế mô tả phương pháp tạo ra phương tiện phân phối gen, bao gồm bước: cho tế bào tiếp xúc với chế phẩm theo sáng chế với lượng hữu hiệu để thu được virus cúm, và phân lập virus này. Sáng chế còn đề xuất tế bào được tiếp xúc với chế phẩm theo sáng chế.

Theo một phương án, sáng chế đề xuất tế bào của động vật có xương sống chứa nhiều vector để tạo ra virus cúm lợn nhược độc bao gồm vector chứa trình tự khởi động được liên kết có điều khiển với ADN bổ sung của HA của virus cúm,

trong đó ADN bổ sung của HA mã hóa các vị trí glycosyl hóa bổ sung so với HA được mã hóa bởi chủng virus cúm lợn gốc có tính độc.

Sáng chế cũng đề xuất các sản phẩm gen, tạo ra các kháng nguyên, chất kháng nguyên và epitop, và hữu ích làm các sản phẩm gen được phân lập.

Các sản phẩm gen được phân lập này, cũng như các epitop của chúng, cũng hữu ích để tạo ra kháng thể, các kháng thể này hữu ích trong các ứng dụng chẩn đoán.

Các sản phẩm gen, có thể cung cấp hoặc tạo ra epitop, kháng nguyên hoặc chất kháng nguyên, cũng hữu ích cho các chế phẩm sinh miễn dịch hoặc miễn dịch, cũng như các vaccin.

Theo một phương án, sáng chế đề xuất virus chứa đột biến giảm độc lực trong trình tự nucleotit hoặc gen trong đó đột biến này biến đổi hoạt tính sinh học của polypeptit hoặc protein được mã hóa bởi gen, dẫn đến độc lực của virus được làm giảm.

Đặc biệt là, sáng chế đề cập đến các chủng virus cúm lợn nhược độc và vaccin chứa các chủng này, gây ra đáp ứng sinh miễn dịch ở động vật, cụ thể là các chủng virus cúm lợn nhược độc gây ra, sinh hoặc kích thích đáp ứng ở loài lợn.

Các chủng virus cúm lợn nhược độc cụ thể cần quan tâm có các đột biến trong gen, so với chủng gốc có tính độc kiểu đại, liên quan đến độc lực. Nhận thấy rằng ngoài các chủng có các đột biến được nêu, các chủng nhược độc có số đột biến bất kỳ trong các gen độc lực được bộc lộ có thể được sử dụng trong thực hành sáng chế này.

Theo một phương án khác, các chủng virus cúm lợn nhược độc mới được bào chế thành vaccin an toàn, hiệu quả chống virus cúm lợn và bệnh nhiễm / các bệnh do virus cúm lợn gây ra.

Theo một phương án, vaccin chứa virus cúm lợn còn chứa tá chất. Theo một phương án cụ thể, tá chất này là tá chất dùng qua niêm mạc, như chitosan, chitosan được metyl hóa, chitosan được trimetyl hóa, hoặc các dẫn xuất hoặc tổ hợp của chúng.

Theo một phương án, tá chất này bao gồm vi khuẩn và/hoặc virus nguyên vẹn, bao gồm *H. parasuis*, clostridium, virus influenza ở lợn (swine influenza virus - SIV), circovirus ở lợn (porcine circovirus - PCV), virus gây hội chứng sinh sản và hô hấp ở lợn (porcine reproductive and respiratory syndrome virus - PRRSV), Mannheimia, Pasteurella, Histophilus, Salmonella, *Escherichia coli*, hoặc tổ hợp và/hoặc biến thể của chúng. Theo một số phương án, tá chất này làm

tăng mức sản sinh IgM, IgG, IgA, và/hoặc tổ hợp của chúng ở động vật.

“Kháng nguyên” hoặc “chất kháng nguyên” có nghĩa là chất gây ra đáp ứng miễn dịch đặc hiệu ở động vật chủ. Kháng nguyên này có thể bao gồm sinh vật nguyên vẹn, được làm chết, nhược độc hoặc sống; tiểu đơn vị hoặc một phần của sinh vật; vector tái tổ hợp chứa đoạn xen có đặc tính sinh miễn dịch; đoạn hoặc mảnh ADN có khả năng gây ra đáp ứng miễn dịch khi trình diện với động vật chủ; polypeptit, epitop, hapten, hoặc tổ hợp bất kỳ của chúng. Theo cách khác, chất kháng nguyên hoặc kháng nguyên có thể bao gồm độc tố hoặc kháng độc tố.

Các thuật ngữ “protein”, “peptit”, “polypeptit” và “mảnh polypeptit” được sử dụng thay thế cho nhau trong bản mô tả này để chỉ các polyme chứa các gốc axit amin có độ dài bất kỳ. Polyme này có thể có mạch thẳng hoặc mạch nhánh, nó có thể bao gồm các axit amin được cải biến hoặc các chất tương tự axit amin, và nó có thể bị gián đoạn bởi các gốc hóa học không phải axit amin. Các thuật ngữ này cũng bao gồm polyme axit amin được cải biến một cách tự nhiên hoặc bằng cách can thiệp; ví dụ, tạo ra liên kết disulfua, glycosyl hóa, lipit hóa, axetyl hóa, phosphoryl hóa, hoặc thao tác hoặc cải biến khác bất kỳ, như liên hợp với thành phần đánh dấu hoặc có hoạt tính sinh học.

Thuật ngữ “polypeptit sinh miễn dịch hoặc kháng nguyên” được sử dụng trong bản mô tả này bao gồm các polypeptit có hoạt tính miễn dịch với nghĩa là ngay khi được dùng cho vật chủ, nó có thể gây ra đáp ứng miễn dịch kiểu dịch thể và/hoặc tế bào trực tiếp chống lại protein. Tốt hơn là mảnh protein là mảnh có hoạt tính miễn dịch gần giống như protein toàn phần. Do đó, mảnh protein theo sáng chế bao gồm hoặc về cơ bản được cấu thành từ hoặc gồm có ít nhất một epitop hoặc yếu tố quyết định kháng nguyên. Protein hoặc polypeptit “sinh miễn dịch”, được sử dụng trong bản mô tả này, bao gồm trình tự có chiều dài đầy đủ của protein, chất tương tự của chúng, hoặc các mảnh sinh miễn dịch của chúng. “Mảnh sinh miễn dịch” có nghĩa là mảnh của protein bao gồm một hoặc nhiều epitop và do đó gây ra đáp ứng miễn dịch được mô tả trên đây. Các mảnh như vậy có thể được nhận diện bằng cách sử dụng một số kỹ thuật bất kỳ trong số các kỹ thuật lập bản đồ epitop đã biết trong lĩnh vực này. Xem, ví dụ, Epitope Mapping Protocols in Methods in Molecular Biology, Vol. 66 (Glenn E. Morris, Ed., 1996). Ví dụ, các epitop thẳng có thể được xác định bằng cách, ví dụ, tổng hợp đồng thời một số lượng lớn các peptit trên giá đỡ rắn, các peptit này tương ứng với các phần của phân tử protein, và cho các peptit này phản ứng với các kháng thể trong khi các peptit này vẫn được gắn với giá đỡ. Các kỹ thuật này là đã biết trong lĩnh vực và được mô tả trong, ví dụ, patent Mỹ số 4,708,871; Geysen và các đồng tác giả,

1984; Geysen và các đồng tác giả, 1986. Tương tự, các epitop cấu hình riêng được nhận diện dễ dàng bằng cách xác định cấu hình trong không gian của các axit amin như bằng, ví dụ, tinh thể học tia x và cộng hưởng từ hạt nhân 2 chiều. Xem, ví dụ, các quy trình lập bản đồ epitop, nêu trên. Các phương pháp đặc biệt có thể áp dụng được cho các protein của *T. parva* được mô tả đầy đủ trong PCT/US2004/022605 được hợp nhất vào bản mô tả này bằng cách viện dẫn toàn bộ.

Như được bàn luận trong bản mô tả này, sáng chế đề xuất các mảnh và các biến thể có hoạt tính của polypeptit kháng nguyên. Do đó, thuật ngữ "polypeptit sinh miễn dịch hoặc kháng nguyên" còn bao gồm sự mất đoạn, thêm đoạn và thay thế vào trình tự, miễn là các polypeptit này có chức năng tạo ra đáp ứng miễn dịch như được xác định trong bản mô tả này. Thuật ngữ "sự biến đổi bảo toàn" chỉ sự thay thế gốc axit amin bằng một gốc khác tương tự về mặt sinh học, hoặc sự thay thế nucleotit trong trình tự axit nucleic sao cho gốc axit amin được mã hóa không thay đổi hoặc là một gốc khác tương tự về mặt sinh học. Về mặt này, sự thay thế được đặc biệt ưu tiên thường sẽ được bảo toàn trong tự nhiên, tức là sự thay thế diễn ra trong họ các axit amin. Ví dụ, các axit amin nói chung được chia thành bốn họ: (1) axit--aspartat và glutamat; (2) bazơ--lysin, arginin, histidin; (3) không phân cực--alanin, valin, leuxin, isoleuxin, prolin, phenylalanin, metionin, tryptophan; và (4) phân cực không tích điện--glyxin, asparagin, glutamin, xystin, serin, threonin, tyrosin. Phenylalanin, tryptophan, và tyrosin đôi khi được phân loại là các axit amin thơm. Ví dụ về các biến đổi bảo toàn bao gồm sự thay thế một gốc kỵ nước như isoleuxin, valin, leuxin hoặc metionin cho một gốc kỵ nước khác, hoặc sự thay thế một gốc phân cực cho một gốc phân cực khác, như sự thay thế arginin cho lysin, axit glutamic cho axit aspartic, hoặc glutamin cho asparagin, và các thay thế tương tự; hoặc sự thay thế bảo toàn tương tự của axit amin bằng axit amin liên quan về cấu trúc mà sẽ không ảnh hưởng chủ yếu đến hoạt tính sinh học. Các protein có trình tự axit amin gần giống như phân tử đối chiếu nhưng có sự thay thế axit amin nhỏ không ảnh hưởng đáng kể đến khả năng sinh miễn dịch của protein là, do đó, thuộc định nghĩa của polypeptit đối chiếu. Tất cả các polypeptit được tạo ra bởi các biến đổi này là được bao gồm trong bản mô tả này. Thuật ngữ "sự biến đổi bảo toàn" cũng bao gồm việc sử dụng axit amin được thế thay cho axit amin gốc không được thế miễn là các kháng thể được sinh ra bởi polypeptit được thế cũng phản ứng miễn dịch với polypeptit không được thế.

Thuật ngữ "epitop" được dùng để chỉ vị trí trên kháng nguyên hoặc hapten mà tế bào B và/hoặc tế bào T đặc hiệu đáp ứng với. Thuật ngữ này cũng được sử

dụng thay thế cho "yếu tố quyết định kháng nguyên" hoặc "vị trí yếu tố quyết định kháng nguyên". Các kháng thể nhận biết cùng một epitop có thể được nhận diện trong thử nghiệm miễn dịch đơn giản thể hiện khả năng của một kháng thể phong bế gắn kết của một kháng thể khác với kháng nguyên đích.

"Đáp ứng miễn dịch" với chế phẩm hoặc vaccin là quá trình trong vật chủ có đáp ứng miễn dịch qua trung gian tế bào và/hoặc kháng thể với chế phẩm hoặc vaccin cần quan tâm. Thông thường, "đáp ứng miễn dịch" bao gồm, nhưng không chỉ giới hạn ở, một hoặc nhiều tác động sau: sản sinh kháng thể, tế bào B, tế bào T hỗ trợ, và/hoặc tế bào T độc, hướng đặc hiệu đến kháng nguyên hoặc các kháng nguyên được bao gồm trong chế phẩm hoặc vaccin cần quan tâm. Tốt hơn là, vật chủ sẽ biểu hiện đáp ứng miễn dịch trị liệu hoặc bảo vệ sao cho tính kháng với việc nhiễm mới sẽ được tăng cường và/hoặc mức độ nghiêm trọng lâm sàng của bệnh được làm giảm. Sự bảo vệ này sẽ được minh họa bởi việc giảm hoặc mất các triệu chứng và/hoặc dấu hiệu bệnh lâm sàng thường được biểu hiện bởi vật chủ bị nhiễm, thời gian hồi phục nhanh hơn và/hoặc hiệu giá virus được giảm trong vật chủ bị nhiễm.

"Động vật" được dự định bao gồm động vật có vú, chim, và các loài tương tự. Động vật hoặc vật chủ được sử dụng trong bản mô tả này bao gồm động vật có vú và người. Động vật này có thể được chọn từ nhóm bao gồm loài ngựa (ví dụ, ngựa), loài chó (ví dụ, chó, chó sói, cáo, chó sói đồng cỏ, chó rừng), loài mèo (ví dụ, sư tử, hổ, mèo nuôi trong nhà, mèo hoang, mèo lớn khác, và các loài mèo khác bao gồm báo gêpa và mèo rừng), loài cừu (ví dụ, cừu), loài bò (ví dụ, gia súc), loài lợn (ví dụ, lợn), loài chim (ví dụ, gà, vịt, ngỗng, gà tây, chim cú, gà lôi, vịt, chim họ sẻ, diều hâu, quạ, đà điểu châu Phi, đà điểu sa mạc Úc và đà điểu đầu mèo), động vật linh trưởng (ví dụ, bộ bán hầu, phủ hầu, khỉ, vượn, khỉ không đuôi), chồn sương, chó biển, và cá. Thuật ngữ "động vật" còn bao gồm động vật riêng ở tất cả các giai đoạn phát triển, bao gồm các giai đoạn mới sinh, phôi và thai.

Trừ khi được giải thích khác, tất cả các thuật ngữ kỹ thuật và khoa học được sử dụng trong bản mô tả này có nghĩa giống như thường được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật mà sáng chế này thuộc về. Các thuật ngữ dạng số ít "một" ("a", "an", và "the" trong bản mô tả tiếng Anh) cũng bao gồm cả dạng số nhiều trừ khi ngữ cảnh chỉ rõ khác. Tương tự, từ "hoặc" được dự định bao gồm "và" trừ khi ngữ cảnh chỉ rõ khác.

Lưu ý rằng trong sáng chế này và đặc biệt là trong yêu cầu bảo hộ và/hoặc các đoạn, các thuật ngữ như "gồm có", "được bao gồm", "bao gồm" và các thuật

ngữ tương tự có thể có nghĩa như được quy định trong luật patent Mỹ; ví dụ, chúng có thể có nghĩa “gồm cả”, “được gồm cả”, “kể cả”, và các nghĩa tương tự; và các thuật ngữ như “về cơ bản được cấu thành từ” và “về cơ bản gồm có” có nghĩa như được quy định trong luật patent Mỹ, ví dụ, chúng cho phép các phần tử không cần trích dẫn cụ thể, nhưng loại trừ các phần tử được tìm thấy trong giải pháp kỹ thuật đã biết hoặc ảnh hưởng đến tính cơ bản hoặc mới của sáng chế.

Chế phẩm

Sáng chế đề xuất vaccin hoặc chế phẩm phòng bệnh cúm lợn có thể chứa chủng virus cúm lợn nhược độc và chất mang, tá dược, hoặc chất dẫn được dùng hoặc dùng trong thú y, gây ra, sinh hoặc kích thích đáp ứng ở động vật.

Thuật ngữ “axit nucleic” và “polynucleotit” được dùng để chỉ ARN hoặc ADN có mạch thẳng hoặc mạch nhánh, sợi đơn hoặc kép, hoặc thể lai của chúng. Thuật ngữ này cũng bao gồm thể lai ARN/ADN. Sau đây là các ví dụ không nhằm mục đích giới hạn về các polynucleotit: gen hoặc mảnh gen, exon, intron, mRNA, tARN, rARN, ribozym, ADN bổ sung, polynucleotit tái tổ hợp, polynucleotit phân nhánh, plasmit, vector, ADN được phân lập có trình tự bất kỳ, ARN được phân lập có trình tự bất kỳ, mẫu dò và đoạn mồi của axit nucleic. Polynucleotit có thể bao gồm các nucleotit được cải biến, như nucleotit được metyl hóa và chất tương tự nucleotit, uracyl, các đường khác và các nhóm liên kết như các nhánh floribozơ và thiolat, và nucleotit. Trình tự của các nucleotit có thể được cải biến tiếp sau khi polyme hóa, như bằng cách liên hợp, với thành phần đánh dấu. Các loại cải biến khác được bao gồm trong định nghĩa này là gắn mũ, thay thế một hoặc nhiều nucleotit có trong tự nhiên bằng chất tương tự, và đưa bằng cách nào đó để gắn polynucleotit với các protein, ion kim loại, thành phần đánh dấu, polynucleotit hoặc giá thể rắn. Các polynucleotit này có thể thu được bằng cách tổng hợp hoá học hoặc thu được từ vi sinh vật.

Thuật ngữ “gen” được sử dụng rộng rãi để chỉ đoạn polynucleotit bất kỳ liên quan đến chức năng sinh học. Do đó, các gen bao gồm các intron và exon như trong trình tự hệ gen, hoặc chỉ là trình tự mã hóa như trong các ADN bổ sung và/hoặc trình tự điều hòa cần thiết cho sự biểu hiện của chúng. Ví dụ, gen cũng được dùng để chỉ mảnh axit nucleic biểu hiện ARN thông tin hoặc ARN chức năng, hoặc mã hóa protein đặc hiệu, và bao gồm các trình tự điều hòa.

Thành phần sinh học “được phân lập” (như axit nucleic hoặc protein hoặc bào quan) được dùng để chỉ thành phần đã được tách hoặc tinh chế gần hết khỏi các thành phần sinh học khác trong tế bào của sinh vật trong đó thành phần này có tự nhiên, ví dụ, ADN và ARN nhiễm sắc thể và ngoài nhiễm sắc thể khác, protein,

và bào quan. Axit nucleic và protein đã “được phân lập” bao gồm axit nucleic và protein được tinh chế bằng các phương pháp tinh chế chuẩn. Thuật ngữ này cũng bao hàm các axit nucleic và các protein được tạo ra bằng công nghệ tái tổ hợp cũng như tổng hợp hoá học.

Thuật ngữ "sự biến đổi bảo toàn" chỉ sự thay thế gốc axit amin bằng một gốc khác tương tự về mặt sinh học, hoặc sự thay thế nucleotit trong trình tự axit nucleic sao cho gốc axit amin được mã hóa không thay đổi hoặc là một gốc khác tương tự về mặt sinh học. Về mặt này, sự thay thế được đặc biệt ưu tiên thường sẽ được bảo toàn trong tự nhiên, như đã mô tả trên đây.

Thuật ngữ “tái tổ hợp” có nghĩa là polynucleotit có nguồn gốc bán tổng hợp, hoặc tổng hợp không có trong tự nhiên hoặc được gắn kết với một polynucleotit khác theo cấu trúc không tìm thấy được trong tự nhiên.

“Dị loại” có nghĩa là phần tử có nguồn gốc từ thực thể khác biệt về mặt di truyền so với phần còn lại của thực thể mà nó được so sánh. Ví dụ, polynucleotit có thể được đưa vào plasmit hoặc vectơ có nguồn gốc từ nguồn khác bằng kỹ thuật xử lý di truyền, và là polynucleotit dị loại. Trình tự khởi động được tách khỏi trình tự mã hóa nguyên thể của nó và được liên kết có điều khiển với trình tự mã hóa khác trình tự nguyên thể là trình tự khởi động dị loại.

Các polynucleotit theo sáng chế có thể bao gồm các trình tự bổ sung, như các trình tự mã hóa bổ sung trong cùng đơn vị phiên mã, các yếu tố điều khiển như trình tự khởi động, vị trí gắn kết ribosom, 5'UTR, 3'UTR, vùng kết thúc phiên mã, vị trí polyadenyl hóa, đơn vị phiên mã bổ sung dưới sự điều khiển của trình tự khởi động cùng hoặc khác, các trình tự cho phép tách dòng, biểu hiện, tái tổ hợp tương đồng, và biến nạp của tế bào chủ, và cấu trúc như vậy bất kỳ như có thể mong muốn để đề xuất các phương án của sáng chế.

Phương pháp sử dụng và vật phẩm sản xuất

Sáng chế bao gồm các phương án về phương pháp sau. Theo một phương án, sáng chế mô tả phương pháp chủng ngừa động vật bao gồm bước cho động vật này dùng chế phẩm chứa chủng virus cúm lợn nhược độc và chất mang, tá dược, hoặc chất dẫn được dùng hoặc dùng trong thú y. Theo một khía cạnh của phương án này, động vật này là lợn.

Theo một phương án của sáng chế, chế độ chủng ngừa nhắc lại có thể được dùng, chế độ này bao gồm ít nhất một lần dùng cơ bản và ít nhất một lần dùng nhắc lại bằng cách sử dụng ít nhất một polypeptit, kháng nguyên, epitop hoặc chất kháng nguyên thông thường. Thông thường, chế phẩm miễn dịch hoặc vaccin được sử dụng trong lần dùng cơ bản về bản chất là khác với chế phẩm miễn dịch

hoặc vaccin được sử dụng để nhắc lại. Tuy nhiên, lưu ý rằng chế phẩm giống nhau có thể được sử dụng trong lần dùng cơ bản và lần dùng nhắc lại. Quy trình dùng này được gọi là “chủng ngừa nhắc lại”.

Chế độ chủng ngừa nhắc lại bao gồm ít nhất một lần dùng cơ bản và ít nhất một lần dùng nhắc lại bằng cách sử dụng ít nhất một polypeptit thông thường và/hoặc các biến thể hoặc của chúng. Vaccin được sử dụng trong lần dùng cơ bản có thể khác về bản chất với vaccin được sử dụng làm vaccin nhắc lại sau đó. Lần dùng cơ bản có thể bao gồm một hoặc nhiều lần dùng. Tương tự, lần dùng nhắc lại có thể bao gồm một hoặc nhiều lần dùng.

Thể tích liều của chế phẩm đối với loài đích là động vật có vú, ví dụ, thể tích liều của chế phẩm dùng cho lợn hoặc loài lợn, tính theo kháng nguyên của virus, thường nằm trong khoảng từ 0,1 đến 2,0 ml, nằm trong khoảng từ 0,1 đến 1,0 ml, và nằm trong khoảng từ 0,5 ml đến 1,0 ml.

Hiệu lực của vaccin có thể được kiểm tra ở khoảng 2 đến 4 tuần sau lần chủng ngừa cuối cùng bằng cách thử nghiệm gây nhiễm động vật, như loài lợn, bằng chủng virus cúm lợn có tính độc. Cả hai chủng đồng loại và dị loại được sử dụng để gây nhiễm để thử nghiệm hiệu lực của vaccin. Động vật này có thể được gây nhiễm bằng cách tiêm IM (trong cơ) hoặc SC (dưới da), phun xịt trong mũi, trong mắt, trong khí quản, và/hoặc qua đường miệng. Các mẫu từ các khớp, phổi, não, và/hoặc miệng có thể được thu trước và sau khi gây nhiễm và có thể được phân tích sự có mặt của kháng thể đặc hiệu với virus cúm lợn.

Các chế phẩm chứa chủng virus nhược độc theo sáng chế được sử dụng trong quy trình chủng ngừa nhắc lại là chứa chất dẫn, chất pha loãng hoặc tá dược được dùng hoặc dùng trong thú y. Các quy trình theo sáng chế bảo vệ động vật khỏi virus cúm lợn và/hoặc phòng ngừa tiến triển của bệnh ở động vật bị nhiễm.

Các lần dùng khác nhau tốt hơn là được thực hiện cách nhau từ 1 đến 6 tuần. Theo một phương án, khoảng thời gian được ưu tiên là từ 3 đến 5 tuần, và tối ưu là 4 tuần, việc nhắc lại hàng năm cũng được dự tính. Các động vật, ví dụ, lợn, có thể ít nhất 3-4 tuần tuổi tại thời điểm dùng đầu tiên.

Người có hiểu biết trung bình trong lĩnh vực này cần hiểu rằng sáng chế trong bản mô tả này được đưa ra bằng các ví dụ và sáng chế này không bị giới hạn ở đó. Từ bộc lộ trong bản mô tả này và hiểu biết trong lĩnh vực, người có hiểu biết trung bình trong lĩnh vực này có thể xác định số lần dùng, đường dùng, và các liều được sử dụng cho mỗi quy trình tiêm chủng, mà không cần thử nghiệm quá mức bất kỳ.

Theo một phương án thực hiện khác, sáng chế đề xuất bộ kit để thực hiện

phương pháp gây ra hoặc sinh đáp ứng miễn dịch hoặc bảo vệ chống virus cúm lợn ở động vật bao gồm chế phẩm miễn dịch hoặc vaccin chứa virus cúm lợn nhược độc và hướng dẫn để thực hiện phương pháp phân phối một lượng hữu hiệu để gây ra đáp ứng miễn dịch ở động vật.

Theo một phương án thực hiện khác, sáng chế đề xuất bộ kit để thực hiện phương pháp gây đáp ứng miễn dịch hoặc bảo vệ chống virus cúm lợn ở động vật bao gồm chế phẩm hoặc vaccin chứa chủng virus cúm lợn nhược độc theo sáng chế, và hướng dẫn để thực hiện phương pháp phân phối một lượng hữu hiệu để gây ra đáp ứng miễn dịch ở động vật.

Theo một khía cạnh khác nữa, sáng chế đề cập đến bộ kit để chủng ngừa nhắc lại theo sáng chế như đã mô tả trên đây. Bộ kit này có thể bao gồm ít nhất hai lọ: lọ thứ nhất chứa vaccin hoặc chế phẩm để chủng ngừa cơ bản theo sáng chế, và lọ thứ hai chứa vaccin hoặc chế phẩm để chủng ngừa nhắc lại theo sáng chế. Bộ kit này có thể chứa một cách thuận lợi lọ bổ sung thứ nhất hoặc thứ hai để chủng ngừa cơ bản bổ sung hoặc chủng ngừa nhắc lại bổ sung.

Các chất mang hoặc chất dẫn hoặc tá dược được dụng hoặc dùng trong thú y là đã biết rõ đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật này. Ví dụ, chất mang hoặc chất dẫn hoặc tá dược được dụng hoặc dùng trong thú y có thể là dung dịch NaCl 0,9% (ví dụ, nước muối sinh lý) hoặc dung dịch đệm phosphat. Chất mang hoặc chất dẫn hoặc tá dược được dụng hoặc dùng trong thú y khác mà có thể được sử dụng trong phương pháp theo sáng chế bao gồm, nhưng không giới hạn ở, poly-(L-glutamat) hoặc polyvinylpyrrolidon. Chất mang hoặc chất dẫn hoặc tá dược được dụng hoặc dùng trong thú y có thể là hợp chất bất kỳ hoặc tổ hợp các hợp chất tạo thuận lợi cho việc dùng vector (hoặc protein được biểu hiện từ vector theo sáng chế *in vitro*); một cách thuận lợi, chất mang, chất dẫn hoặc tá dược này có thể tạo thuận lợi cho việc chuyển nạp và/hoặc cải thiện việc bảo quản vector (hoặc protein). Các liều và thể tích liều được bàn luận trong bản mô tả này trong phần mô tả chung và cũng có thể được xác định bởi người có hiểu biết trung bình trong lĩnh vực từ việc đọc bản mô tả này kết hợp với hiểu biết trong lĩnh vực, mà không cần thử nghiệm quá mức bất kỳ.

Các chế phẩm miễn dịch và vaccin theo sáng chế có thể bao gồm hoặc về cơ bản bao gồm một hoặc nhiều tá chất. Các tá chất thích hợp để sử dụng trong thực hiện sáng chế là (1) polyme của axit acrylic hoặc metacrylic, anhydrit maleic và các polyme dẫn xuất alkenyl, (2) các trình tự kích thích miễn dịch (immunostimulating sequence - ISS), như trình tự oligodeoxyribonucleotit có một hoặc nhiều đơn vị CpG không được metyl hóa (Klinman et al., 1996;

WO98/16247), (3) nhũ tương dầu trong nước, như nhũ tương SPT được mô tả ở trang 147 của “Vaccine Design, The Subunit and Adjuvant Approach” được công bố bởi M. Powell, M. Newman, Plenum Press 1995, và nhũ tương MF59 được mô tả ở trang 183 của cùng tài liệu này, (4) lipid cation chứa muối amoni bậc bốn, ví dụ, DDA (5) xytokin, (6) nhôm hydroxit hoặc nhôm phosphat, (7) saponin hoặc (8) các tá chất khác được bàn luận trong tài liệu bất kỳ được trích dẫn và được hợp nhất bằng cách viện dẫn vào bản mô tả này, hoặc (9) tổ hợp hoặc hỗn hợp bất kỳ của các chất này.

Theo một phương án, các tá chất bao gồm các chất thúc đẩy sự hấp thu cải thiện qua lớp niêm mạc. Một số ví dụ bao gồm MPL, LTK63, các độc tố, các vi hạt PLG và một số chất khác (Vajdy, M. Immunology and Cell Biology (2004) 82, 617–627). Theo một phương án, tá chất này có thể là chitosan (Van der Lubben và các đồng tác giả 2001; Patel và các đồng tác giả 2005; Majithiya và các đồng tác giả 2008; patent Mỹ số 5,980,912).

Tài liệu tham khảo:

- Abe Y, Takashita E, Sugawara K, Matsuzaki Y, Muraki Y, Hongo S. 2004. Effect of the addition of oligosaccharides on the biological activities and antigenicity of influenza A/H3N2 virus hemagglutinin. *J. Virology* **78**:9605-9611.
- Altschul SF, Madden TL. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**:3389-3402.
- Bohne-Lang A, von der Lieth CW. 2005. GlyProt: in silico glycosylation of proteins. *Nucleic Acids Res.* **33**:W214-219.
- Bragstad K, Nielsen LP, Fomsgaard A. 2008. The evolution of human influenza A viruses from 1999 to 2006: A complete genome study. *Virology J.* **5**:40.
- Cai Z, Zhang T, Wan X-F. 2010. A computational framework for influenza antigenic cartography. *PLoS Computational Biology*, **6**(10):e1000949.
- Caton AJ, Brownlee GG, Yewdell JW, Gerhard W. 1982. The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). *Cell* **31**:417-427.
- Cottey R, Rowe CA, Bender BS. 2001. Influenza virus. *Curr. Protoc. Immunol.* **19**.11:1-32.
- Das SR, Hensley SE, David A, Schmidt L, Gibbs JS, Puigbò P, Ince WL, Bennick JR, Yewdell JW. 2011. Fitness costs limit influenza A virus hemagglutinin glycosylation as an immune evasion strategy. *Proc. Natl. Acad. Sci.* **108**:E1417-E1422.

- Das SR, Puigbò P, Hensley SE, Hurt DE, Bennick JR, Yewdell JW. 2010. Glycosylation focuses sequence variation in the influenza A virus H1 hemagglutinin globular domain. *PLoS Pathog* **6**:e1001211.
- Desselberger U, Nakajima K, Alfino P, Pedersen F, Haseltine W, Hannoun C, Palese P. 1978. Biochemical evidence that “new” influenza virus strains in nature may arise by recombination (reassortment). *Proc. Natl. Acad. Sci.* **75**:3341-3345.
- Hause BM, Oleson TA, Stine DL, Bey RF, Simonson RR. 2011. Genetic and antigenic characterization of recent human-like H1 (δ -cluster) swine influenza virus isolates. *J. Swine Health Prod.* **19**:268-276.
- Hause BM, Oleson TA, Bey RF, Stine DL, Simonson RR. 2010. Antigenic categorization of contemporary H3N2 swine influenza virus isolates using a high-throughput serum neutralization assay. *J. Vet. Diagn. Invest.* **22**:352-359.
- Hay AJ, Gregory V, Douglas AR, Lin YP. 2001. The evolution of human influenza viruses. *Philos. T. R. Soc. B.* **356**:1861-1870.
- Hensley SE, Das SR, Gibbs JS, Bailey AL, Schmidt LM, Bennick JR, Yewdell JW. 2011. Influenza A virus hemagglutinin antibody escape promotes neuraminidase antigenic variation and drug resistance. *PLoS ONE* **6**:e15190.
- Hoffmann E, Stech J, Guan Y, Webster RG, Perez DR. 2001. Universal primer set for the full length amplification of all influenza viruses. *Arch. Virol.* **146**:2275-2289.
- Hoffmann E, Neumann G, Kawaoka Y, Hobom G, Webster RG. 2000. A DNA transfection system for generation of influenza A virus from eight plasmids. *Proc. Natl. Acad. Sci.* **97**:6108-6113.
- Karasin AI, Carman S, Olsen CW. 2006. Identification of human H1N2 and human-swine reassortant H1N2 and H1N1 influenza A viruses among pigs in Ontario, Canada (2003-2005). *J. Clin. Microbiol.* **44**:1123-1126.
- Long J, Bushnell RV, Tobin JK, Pan K, Deem MW, Nara PL, Tobin GJ. 2011. Evolution of H3N2 influenza virus in a guinea pig model. *PLoS ONE* **6**:e20130.
- Lorusso A, Vincent AL, Gramer MR, Lager KM, Ciacchi-Zanella JR. 2012. Contemporary epidemiology of North American lineage triple reassortant influenza A viruses in pigs. *Curr. Top. Microbiol. Immunol.* Jan. 22 [epub ahead of print].
- Sali A, Potterton L. 1995. Evaluation of comparative protein modeling by MODELLER." *Proteins* **23**: 318-326.
- Schild G, Oxford J, Dowdle W, Coleman M, Pereira M, Chakraverty P. 1974. Antigenic variation in current influenza A viruses: evidence for a high frequency

- of antigenic 'drift' for the Hong Kong virus. *Bull. World Health Organ.* **51**:1-11.
- Schulze IT. 1997. Effects of glycosylation on the properties and functions of influenza virus hemagglutinin. *J. Infect. Dis.* **176**:S24-S28
- Strengell M, Ikonen N, Ziegler T, Julkunen I. 2011. Minor changes in the hemagglutinin of influenza A(H1N1)2009 virus alter its antigenic properties. *PLoS ONE* **6**:e25848.
- Sun S, Wang Q, Zhao F, Chen W, Li Z. 2011. Glycosylation site alteration in the evolution of influenza A (H1N1) viruses. *PLoS ONE* **6**:e22844.
- Sun S, Wang Q, Zhao F, Chen W, Li Z. 2012. Prediction of biological functions on glycosylation site migrations in human influenza H1N1 viruses. *PLoS ONE* **7**:e32119.
- WHO Manual on Animal Influenza Diagnosis and Surveillance, 2002 (http://whqlibdoc.who.int/hq/2002/WHO_CDS_CSR_NCS_2002.5.pdf).
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**: 2731-2739.
- Vigerust DJ, Ulett KB, Boyd KL, Madsen J, Hawgood S, McCullers JA. 2007. N-linked glycosylation attenuates H3N2 influenza viruses. *J. Virology* **81**:8593-8600.
- Vincent AL, Ma W, Lager KM, Gramer MR, Richt JA, Janke BH. 2009. Characterization of a newly emerged genetic cluster of H1N1 and H1N2 swine influenza virus in the United States. *Virus Genes* **39**:176-185.
- Wagner R, Wolff T, Herwig A, Pleschka S, Klenk H-D. 2000. Interdependence of hemagglutinin glycosylation and neuraminidase as regulators of influenza virus growth: a study by reverse genetics. *J. Virology* **74**:6316-6323.
- Wanzeck K, Boyd KL, McCullers JA. 2011. Glycan shielding of the influenza virus hemagglutinin contributes to immunopathology in mice. *Am. J. Respir. Crit. Care Med.* **183**:767-773.
- Wei C-J, Boyington JC, Dai K, Houser KV, Pearce MB, Kong W-P, Yang Z, Tumpey TM, Nabel GJ. 2010. Cross-neutralization of 1918 and 2009 influenza viruses: role of glycans in viral evolution and vaccine design. *Sci. Trans. Med.* **2**:24fa21. doi:10.1126/scitranslmed.3000799.
- Zhou B, Donnelly ME, Scholes DT, George KS, Hatta M, Kawaoka Y, Wentworth DE. 2009. Single-reaction genomic amplification accelerates sequencing and vaccine production for classical and swine origin human influenza A viruses. *J.*

Virology 83:10309-10313.

Dưới đây sáng chế sẽ được mô tả thêm bằng các ví dụ không mang tính giới hạn sau.

Ví dụ thực hiện sáng chế

Ví dụ 1 – Tạo virus cúm lợn tăng glycosyl hóa

Nguyên liệu/Phương pháp.

Các mẫu lâm sàng (gạc thối mũi hoặc mô phổi) được thu từ lợn biểu hiện bệnh giống cúm và được đem phân lập và mô tả đặc điểm của virus (thử nghiệm chẩn đoán thường quy). Việc phân lập virus được thực hiện trên các tế bào tinh hoàn lợn (swine testicle - ST) được phát triển trong DMEM chứa 5% huyết thanh thai bò ở nhiệt độ 37°C với 5% CO₂. Để nhân giống virus, huyết thanh thai bò được loại bỏ khỏi DMEM. Các tế bào 293T và MDCK được nhân giống trong DMEM chứa 10% huyết thanh thai bò. ARN được thu lại từ dịch nuôi cấy thu được của tế bào bị nhiễm bằng cách sử dụng bộ kit phân lập virus 5x MagMax-96 (Life Technologies).

Hệ gen hoàn chỉnh của virus được khuếch đại bằng cách sử dụng phương pháp PCR phiên mã ngược đa đoạn đã biết (Zhou B, và các đồng tác giả, 2009). Thư viện cADN (ADN bổ sung) của virus được tạo ra bằng cách sử dụng bộ kit NEBNext Fast DNA Fragmentation and Library Prep Set 4 theo hướng dẫn của nhà sản xuất (New England Biolabs) với ngoại lệ là đoạn thích ứng của bộ kit được thay thế bằng đoạn thích ứng được mã vạch (Ion Xpress Barcode Adaptor 1-16 Kit, Life Technologies). Các khuôn mẫu giải trình tự ADN được tạo ra bằng cách sử dụng bộ kit khuôn mẫu Ion Xpress phiên bản 2.0 (Life Technologies) và được giải trình tự bằng cách sử dụng máy giải trình tự hệ gen cá nhân Ion Torrent (Life Technologies). Các contig (đoạn liền kề) được lắp ráp bằng cách sử dụng phần mềm SeqMan NGen (DNASTar). Các contig mã hóa HA có chiều dài đầy đủ được nhận diện bằng phân tích BLAST. Trình tự ADN của HA có chiều dài đầy đủ được bắt cặp bằng cách sử dụng phương pháp ClustalW. Phân tích phát sinh loài được thực hiện bằng cách sử dụng MEGA 5.0 bằng phương pháp nổi lân cận và tổ chức cây được xác minh với 1000 bản sao khởi động (Tamura K, và các đồng tác giả, 2011). Các trình tự gen của HA được lưu giữ ở ngân hàng gen với số lưu giữ JQ638655-JQ638665.

Di truyền ngược. Trong trường hợp sinh học phân tử, “di truyền ngược” được xác định là việc tạo ra virus có hệ gen có nguồn gốc từ các ADN bổ sung tách dòng (để xem xét, xem Neumann et al., J. Gen. Viral., 83:2635; 2002). Trong

nguyên cứu này, bằng gen nội tái tổ hợp của ba dòng virus cúm lợn (triple reassortant internal gene cassette - TRIG) (A/lợn/North Carolina/3793/08(H1N1)) được sử dụng làm khuôn mẫu để tạo ra hệ di truyền ngược của virus cúm lợn TRIG. Tất cả tám đoạn được khuếch đại PCR, phân cắt bằng *BsmBI* và được cài vào pHW2000 được phân cắt tương tự như đã được mô tả trước đây (Hoffmann E, và các đồng tác giả, 2001). Các plasmit mang đoạn cài được nhận diện bằng cách phân cắt giới hạn và được giải trình tự để kiểm tra độ đồng nhất với A/lợn/North Carolina/3793/08. Các gen HA và NA (SEQ ID NO:7 & 11) từ 10-0036-2 cũng được tách dòng vào pHW2000 (Hoffmann et al 2000, PNAS 97(11):6108-6113) và chuyển nạp cùng với các plasmit mang polymeraza bazơ 2 (PB2), polymeraza bazơ 1 (PB1), polymeraza axit (PA), nucleoprotein (NP), chất nền (M) và các gen không cấu trúc (non-structural gene - NS) có nguồn gốc từ A/lợn/North Carolina/3793/08 để tạo ra 10-0036-2 có nguồn gốc di truyền ngược (RG 10-0036-2, Fig. 1). Việc gây đột biến định hướng vị trí được thực hiện trên plasmit chứa gen HA từ 10-0036-2 để tạo ra thể đột biến có 1 đến 5 vị trí glycosyl hóa được liên kết N bổ sung bằng cách sử dụng bộ kit gây đột biến định hướng vị trí Quik Change II (Agilent Technologies) (Bảng 1, Fig. 2). Các glycan được bổ sung vào protein ở các gốc asparagin (N) nằm, trong trường hợp glycosyl hóa được liên kết N, môtip của N-X-S/T, trong đó N là axit amin asparagin, X là axit amin bất kỳ và S/T là serin hoặc threonin.

Bảng 1. Các vị trí glycosyl hóa được liên kết N bổ sung được xử lý kỹ thuật đưa vào gen HA của RG10-0036-2 bằng cách gây đột biến định hướng vị trí

Số vị trí glycosyl hóa được liên kết N bổ sung	Thay đổi nucleotit	Thay đổi axit amin
1	G212A	S71N
2	G212A, G270T	S71N, K90N
3	G212A, G270T, CT517-518AC	S71N, K90N, L173T
4	G212A, G270T, CT517-518AC, C859A	S71N, K90N, L173T, P287T
5	G212A, G270T, CT517-518AC, C859A, A881C	S71N, K90N, L173T, P287T, K294T

Việc thu hồi virus tái tổ hợp được thực hiện như đã mô tả trước đây (Hoffmann E, và các đồng tác giả, 2000). Nói ngắn gọn, các tế bào 293T và MDCK được đồng nuôi cấy trong Opti-MEM I chứa 5% FBS trong đĩa 6 lỗ chứa khoảng 1×10^6 tế bào thuộc mỗi loại 293T và MDCK trong khoảng 18 giờ trước khi chuyển nạp. Một trăm nanogram của mỗi trong số tám plasmit được gộp lại

trong 100 μ L Opti-MEM I và kết hợp với 100 μ L Opti-MEM chứa 3 μ L Lipofectamine (Invitrogen) và ủ ở nhiệt độ phòng 15 phút trước khi pha loãng đến 1 ml bằng Opti-MEM I và được chuyển nạp sang một lỗ của đĩa 6 lỗ. Các đĩa được ủ ở nhiệt độ 37°C với 5% CO₂ trong 6 giờ trước khi hỗn hợp chuyển nạp được thay thế bằng Opti-MEM I. Ở 24 giờ sau khi chuyển nạp, chuyển 1,5 ml sang đĩa 6 lỗ chứa các tế bào MDCK hợp dòng và bổ sung 1,5 ml DMEM chứa 1 μ g/ml trypsin được xử lý bằng TPCK. Virut được thu lại vào ngày 5 sau khi gây nhiễm và hiệu giá của chúng được xác định bằng thử nghiệm HA. Các gen HA của virut thu được được giải trình tự để kiểm tra trình tự đúng.

Kết quả. Phân tích di truyền của các vị trí glycosyl hóa được liên kết N dự đoán (N-X-S/T) thấy rằng các vị trí ở N28, N40, N104, N142, N176, N303, N497 và N556 đối với virut 10-0036-2. Việc gây đột biến định hướng vị trí được sử dụng để thêm 1 đến 5 vị trí glycosyl hóa được liên kết N bổ sung vào phần đầu hình cầu của HA (Bảng 1). Sau khi thu hồi virut từ môi trường nuôi cấy tế bào, các virut thể đột biến được mô tả đặc điểm bằng nghiên cứu sinh trưởng trên các tế bào ST. Nghiên cứu sinh trưởng được thực hiện do các virut nhược độc thường giảm tốc độ sinh trưởng và hiệu giá *in vitro*. Virut thể đột biến có 4 hoặc 5 vị trí glycosyl hóa được liên kết N bổ sung thể hiện sự giảm sinh trưởng như vậy, gợi ý rằng việc glycosyl hóa bổ sung làm giảm độc lực virut. Các virut nhược độc được đánh giá tiếp ở lợn để đánh giá độc lực *in vivo* của chúng và mô tả đặc điểm đáp ứng miễn dịch chống các virut thể đột biến này.

Ví dụ 2 – Hiệu lực của vacxin cúm lợn nhược độc ở lợn

Nguyên liệu & Phương pháp. Sáu mươi con lợn khỏe mạnh 3 tuần tuổi (được xác nhận huyết thanh âm tính với SIV bằng IDEXX FlockChek ELISA) được tách thành 4 nhóm gồm 15 con trong các chuồng riêng. Vào ngày 0 (d0), lợn được gây nhiễm qua đường mũi bằng 2 ml virut 6,0 TCID₅₀/ml. Nhóm 1 được giả nhiễm bằng môi trường nuôi cấy tế bào (DMEM). Nhóm 2 nhận 10-0036-2 n+5. Nhóm 3 nhận 10-0036-2 n+4 (thể đột biến có 4 vị trí glycosyl hóa bổ sung; tương tự như n+5 nhưng không có đột biến A881C [K294T]). Nhóm 4 nhận chủng gốc 10-0036-2 được tạo ra bằng cách di truyền ngược (không có đột biến). Các con lợn được lau bằng gạc (mũi) ở ngày 0 và các mẫu được chạy bằng QPCR (PCR định lượng) để phát hiện SIV để xác nhận việc không nhiễm bệnh chủ động. Các kết quả được tóm tắt trong bảng 2. Trong các bảng 2 và 3, các nhóm có các chữ cái khác nhau có ý nghĩa thống kê khác nhau ($P < 0,05$). Ví dụ, “A” khác biệt về mặt thống kê với “B”, và “BC” khác biệt về mặt thống kê với “A” nhưng không

khác biệt về mặt thống kê với “B” hoặc “C”.

Bảng 2. Nghiên cứu chủng ngừa

Chủng ngừa (H1N2)	Gạc thấm mũi, ngày 3 (TCID ₅₀ /ml)	Gạc thấm mũi, ngày 5 (TCID ₅₀ /ml)	Hiệu giá phổi, ngày 5 (TCID ₅₀ /ml)	Điểm phổi	Điểm IHC
Đối chứng âm tính	0,0 A	0,0 A	0,0 A	0,0 A	0,0 A
Thể đột biến n+4	3,1 B	2,8 B	0,9 A	0,0 A	0,0 A
Thể đột biến n+5	2,6 C	2,7 B	0,4 A	0,2 A	0,2 A
Gốc	4,1 D	3,2 C	5,4 B	1,5 B	2,0 B

Gạc thấm mũi được thu vào các ngày 1, 3 và 5. Năm con lợn từ mỗi nhóm được làm chết ở ngày 5 và các mẫu phổi được thu lại. Các gạc từ ngày 1 được phân tích bằng QPCR. Các gạc từ ngày 3 và 5 cũng như các mẫu phổi được hiệu giá đối với SIV. Phổi được gửi đến phòng thí nghiệm của Đại học Chẩn đoán (University Diagnostic) để phân tích mô bệnh học và IHC. Vào ngày 21, lợn được chủng ngừa lại như trên đây. Các kết quả này thể hiện các thể đột biến chứa 4 hoặc 5 vị trí glycosyl hóa được liên kết N bổ sung bị giảm độc lực và không có độc lực ở lợn, trong các nghiên cứu sinh trưởng *in vitro*. Hiệu giá gạc thấm mũi chứng tỏ rằng virus này có khả năng sao chép *in vivo*, tuy nhiên, việc không có virus trong phổi và không có tổn thương phổi chứng tỏ việc sao chép bị giới hạn ở đường hô hấp trên. Các đột biến trong các gen khác của virus cúm tạo ra tính miễn cảm nhiệt đã thể hiện là hạn chế việc nhiễm ở đường hô hấp trên và là cơ sở của vaccine cúm nhược độc sống cho người FluMist™.

Vào ngày 31, lợn được gây nhiễm bằng thể phân lập thực địa 12-1110-1 (H3N2). Các kết quả của nghiên cứu gây nhiễm được tóm tắt trong Bảng 3. Đối với việc gây nhiễm (được thực hiện so sánh với Richt et al 2006, J.Virology 80(22):11009-11018) 2mL 4.6 TCID₅₀/ml được phân phối qua đường mũi. Máu và gạc thấm mũi được thu vào ngày 31 trước khi gây nhiễm. Gạc thấm mũi được thu vào ngày 0 và 1, và được phân tích bằng QPCR. Gạc thấm mũi được thu vào ngày 3 và 5 và hiệu giá SIV được xác định bằng cách chuẩn độ. Tất cả các con lợn được làm chết vào ngày 5 và các mẫu phổi được phân tích bằng cách chuẩn độ. Các mẫu phổi được phân tích như trên đây.

Bảng 3. Nghiên cứu gây nhiễm

Gây nhiễm H3N2	Gạc thấm mũi, ngày 3	Gạc thấm mũi, ngày 5	Hiệu giá phổi, ngày 5	Điểm phổi	Điểm IHC
----------------	-------------------------	-------------------------	--------------------------	--------------	-------------

	(TCID ₅₀ /ml)	(TCID ₅₀ /ml)	(TCID ₅₀ /ml)		
Đối chứng âm tính	5,4 A	5,8 A	3,0 A	1,6 A	1,5 A
Thế đột biến n+4	2,3 B	0,7 B	0,0 B	0,0 B	0,0 B
Thế đột biến n+5	2,9 BC	1,8 C	0,1 B	0,0 B	0,0 B
Gốc	0,5 C	0,9 BC	0,0 B	0,0 B	0,0 B

Các kết quả này chứng tỏ rằng lợn được chủng ngừa bằng các thế đột biến n+4 hoặc n+5 được bảo vệ khỏi bệnh như thấy rõ bởi việc không có virus trong phổi bằng cách chuẩn độ và IHC, cũng như không có bằng chứng tổn thương phổi. Các con lợn chưa bị nhiễm (đối chứng âm tính) dễ dàng bị nhiễm virus thử nghiệm H3N2 và biểu hiện bệnh cúm điển hình. Các con lợn trước đây đã bị nhiễm virus gốc 10-0036-2 cũng được bảo vệ khỏi việc gây nhiễm H3N2.

Mặc dù các phương án được ưu tiên của sáng chế đã được mô tả chi tiết trên đây, cần hiểu rằng sáng chế được mô tả bởi các phương án này không bị giới hạn ở các chi tiết cụ thể được nêu trong phần mô tả trên do có thể có nhiều biến thể hiển nhiên của các phương án này mà không vượt quá nội dung hoặc phạm vi của sáng chế.

YÊU CẦU BẢO HỘ

1. Chủng virus cúm lợn nhược độc có khả năng tạo ra đáp ứng miễn dịch an toàn và hiệu quả ở loài lợn chống virus cúm hoặc các bệnh do virus cúm gây ra, mã hóa gen HA (hemagglutinin) có ít nhất 4 vị trí glycosyl hóa bổ sung so với chủng gốc có tính độc, trong đó các vị trí glycosyl hóa được chọn từ S71N, K90N, L173T, P287T, và K294T, và trong đó vị trí của (các) thay đổi của axit amin là dựa trên gen HA được mã hóa bởi chủng gốc có tính độc có trình tự như được nêu trong SEQ ID NO:18.
2. Chủng virus theo điểm 1, trong đó protein HA có trình tự như được nêu trong SEQ ID NO:22.
3. Chủng virus theo điểm 1, trong đó protein HA có trình tự như được nêu trong SEQ ID NO:24.
4. Chế phẩm miễn dịch chứa chủng virus nhược độc theo điểm bất kỳ trong số các điểm từ 1 đến 3.
5. Chế phẩm theo điểm 4, trong đó chế phẩm này còn chứa chất dẫn, chất pha loãng hoặc tá dược được dùng hoặc dùng trong thú y, và tạo ra đáp ứng miễn dịch bảo vệ ở loài lợn chống nhiễm virus cúm lợn có tính độc.
6. Chế phẩm theo điểm 5, trong đó chế phẩm này còn chứa ít nhất một kháng nguyên bổ sung liên quan đến hoặc có nguồn gốc từ mầm bệnh của loài lợn không phải là virus cúm lợn, và trong đó ít nhất một hoặc nhiều kháng nguyên bổ sung có khả năng gây ra đáp ứng miễn dịch ở loài lợn chống *M. hyo*, PCV2, PRRSV, SIV hoặc mầm bệnh khác có khả năng nhiễm và gây ra chứng bệnh hoặc mãn cảm với chứng bệnh ở loài lợn.
7. Chế phẩm chứa nhiều vector để tạo ra virus cúm lợn nhược độc bao gồm vector chứa trình tự khởi động được liên kết có điều khiển với ADN bổ sung của HA của virus cúm, trong đó ADN bổ sung của HA mã hóa các vị trí glycosyl hóa bổ sung so với HA được mã hóa bởi chủng virus cúm lợn gốc có tính độc, trong đó các vị trí glycosyl hóa bổ sung được chọn từ S71N, K90N, L173T, P287T, và K294T, và trong đó việc đánh số các thay đổi của axit amin là dựa trên gen HA được mã hóa bởi chủng gốc có tính độc có trình tự như được nêu trong SEQ ID NO:18; và tốt hơn là ADN bổ sung của HA để tạo ra virus cúm nhược độc mã hóa protein như được nêu trong SEQ ID NO:22 hoặc 24.

DANH MỤC TRÌNH TỰ

<110> Merial Limited
Hause, Ben

<120> Chủng virus cúm lợn nhược độc và chế phẩm miễn dịch chứa chủng virus này

<130> MER 12-202P

<160> 30

<170> PatentIn phiên bản 3.5

<210> 1
<211> 2277
<212> ADN
<213> Virus cúm A

<400> 1
atggagagaa taaaagaact aagagatctg atgtcgcagt ctgcgactcg cgagatactc 60
acaaagacca ctgtggacca tatggccata atcaaaaagt acacatcagg aaggcaagag 120
aagaaccccg cactcagaat gaagtggatg atggcaatga aatacccaat tacagcagac 180
aagagaataa tggacatgat tccagagagg aatgaacaag gacaaaccct ctggagcaaa 240
acaaacgatg ctggatcgga ccgtgtgatg gtatcacccc tggccgtaac atggtggaat 300
aggaatggcc caacaacaag cacagttcac taccctaagg tatataaaac ttatttcgaa 360
aagatcgaaa ggttaaaaca tggatatctt ggccctgtcc acttcagaaa tcaagttaaa 420
ataagaagga gggttgacac aaaccctggt catgcagatc tcagtgccaa ggaggcacag 480
gatgtgatca tggaagttgt tttcccaa atgaagtgggg caagagtact gacgtcagag 540
tcacagctga caataacaaa ggaaaagaaa gaagagctcc aggattgtaa gattgctccc 600
ctgatggtgg catacatgct agaaagagag ttggttcgca agacgaggtt tctcccgtg 660
gctggtggaa caagcagtgt ttatattgaa gtgctacact taactcaggg aacatgctgg 720
gaacaaatgt aactccagg aggagaagtg agaaatgatg atgttgacca aagtttgatt 780
atcgtgcta gaaacatagt aagaagagca gcagtgtcag cagaccatt agcatctctc 840
ttggaaatgt gccacagcac acagattgga ggaataagga tgggtggacat ccttagacag 900
aaccacacgg aggaacaagc cgtagacata tgcaaggcag caatggggct gaggattagc 960
tcttctttca gctttggtgg gttcaccttc aaaagaacaa gcggatcatc agttaagaaa 1020
gaagaagaag tgctcacggg caacctcaa aactgaaaa taagagtaca tgaaggatat 1080
gaagaattca caatggtagg gagaagagca actgctattc tcagaaaagc aaccaggaga 1140
ttgatccagt taatagtaag tgggagagac gatcaatcaa ttgctgaggg aataattgta 1200
gccatggtat ttccacaaga ggattgcatg atcaaggcag ttaggggcca tctgaacttt 1260
gtcaataggg caaaccagcg actgaatccc atgcaccaac tcttgaggca tttccaaaaa 1320
gatgcaaaaag tgcttttcca gaactgggga attgaacca tcgacagtgt gatgggaatg 1380
atcgggatat tgctgatat gacccaagc acggaaatgt cgctgagagg tataagagtc 1440
agcaaaatgg gagtagatga gtattccagc acggagagag tggtagtgag cattgaccga 1500

```

tttttgagag ttcgggatca acgaggggaac gtactattgt cccccaaga ggtcagcgag      1560
acacaaggaa ctgagaaatt gacaataact tattcgtcac caatgatgtg ggagatcaat      1620
ggctctgagt cagtgtggt caacacttat caatggatca taaggaattg ggaagcttg      1680
aaaattcaat ggtcacaaga tcccacgatg ttatacaaca aaatggaatt tgaaccattc      1740
cagtctcttg tccctaaggc aaccagaagt cgttacagtg gattcgtgag gacactgttc      1800
cagcaaatgc gggatgtgct tggaacattt gacactgtcc aaataataaa acttctcccc      1860
tttgctgctg ctccaccgga acagagtagg atacagttct cctcgctgac tgtgaatgtg      1920
agaggatcag ggctgaggat actggtgaaga ggcaattctc cagtgttcaa ttacaacaaa      1980
gcaacacaaa ggcttacaat tcttggaata gatgcagggtg cattgactga agatccagat      2040
gaaggcacag ctggagtgga gtctgctgtc ctgaggggat tcctcatttt gggtaaagaa      2100
gacaagagat atggcccagc attaagcatc aatgaactga gcaatcttgc aaaaggagag      2160
aaggctaata tgtaattgg gcaaggagac gtggtgttgg taatgaaacg gaaacggaac      2220
tctagcatac ttactgacag ccagacagcg accaaaagaa ttcggatggc catcaat      2277

```

```

<210> 2
<211> 759
<212> PRT
<213> Virut cúm A

```

```

<400> 2

```

```

Met Glu Arg Ile Lys Glu Leu Arg Asp Leu Met Ser Gln Ser Arg Thr
1           5           10          15

```

```

Arg Glu Ile Leu Thr Lys Thr Thr Val Asp His Met Ala Ile Ile Lys
          20          25          30

```

```

Lys Tyr Thr Ser Gly Arg Gln Glu Lys Asn Pro Ala Leu Arg Met Lys
          35          40          45

```

```

Trp Met Met Ala Met Lys Tyr Pro Ile Thr Ala Asp Lys Arg Ile Met
          50          55          60

```

```

Asp Met Ile Pro Glu Arg Asn Glu Gln Gly Gln Thr Leu Trp Ser Lys
65           70           75          80

```

```

Thr Asn Asp Ala Gly Ser Asp Arg Val Met Val Ser Pro Leu Ala Val
          85          90          95

```

```

Thr Trp Trp Asn Arg Asn Gly Pro Thr Thr Ser Thr Val His Tyr Pro
          100         105         110

```

```

Lys Val Tyr Lys Thr Tyr Phe Glu Lys Ile Glu Arg Leu Lys His Gly
          115         120         125

```

```

Ile Phe Gly Pro Val His Phe Arg Asn Gln Val Lys Ile Arg Arg Arg
          130         135         140

```

Val Asp Thr Asn Pro Gly His Ala Asp Leu Ser Ala Lys Glu Ala Gln
 145 150 155 160
 Asp Val Ile Met Glu Val Val Phe Pro Asn Glu Val Gly Ala Arg Val
 165 170 175
 Leu Thr Ser Glu Ser Gln Leu Thr Ile Thr Lys Glu Lys Lys Glu Glu
 180 185 190
 Leu Gln Asp Cys Lys Ile Ala Pro Leu Met Val Ala Tyr Met Leu Glu
 195 200 205
 Arg Glu Leu Val Arg Lys Thr Arg Phe Leu Pro Val Ala Gly Gly Thr
 210 215 220
 Ser Ser Val Tyr Ile Glu Val Leu His Leu Thr Gln Gly Thr Cys Trp
 225 230 235 240
 Glu Gln Met Tyr Thr Pro Gly Gly Glu Val Arg Asn Asp Asp Val Asp
 245 250 255
 Gln Ser Leu Ile Ile Ala Ala Arg Asn Ile Val Arg Arg Ala Ala Val
 260 265 270
 Ser Ala Asp Pro Leu Ala Ser Leu Leu Glu Met Cys His Ser Thr Gln
 275 280 285
 Ile Gly Gly Ile Arg Met Val Asp Ile Leu Arg Gln Asn Pro Thr Glu
 290 295 300
 Glu Gln Ala Val Asp Ile Cys Lys Ala Ala Met Gly Leu Arg Ile Ser
 305 310 315 320
 Ser Ser Phe Ser Phe Gly Gly Phe Thr Phe Lys Arg Thr Ser Gly Ser
 325 330 335
 Ser Val Lys Lys Glu Glu Glu Val Leu Thr Gly Asn Leu Gln Thr Leu
 340 345 350
 Lys Ile Arg Val His Glu Gly Tyr Glu Glu Phe Thr Met Val Gly Arg
 355 360 365
 Arg Ala Thr Ala Ile Leu Arg Lys Ala Thr Arg Arg Leu Ile Gln Leu
 370 375 380
 Ile Val Ser Gly Arg Asp Asp Gln Ser Ile Ala Glu Ala Ile Ile Val
 385 390 395 400
 Ala Met Val Phe Ser Gln Glu Asp Cys Met Ile Lys Ala Val Arg Gly
 405 410 415
 Asp Leu Asn Phe Val Asn Arg Ala Asn Gln Arg Leu Asn Pro Met His

420	425	430
Gln Leu Leu Arg His Phe Gln Lys Asp Ala Lys Val Leu Phe Gln Asn		
435	440	445
Trp Gly Ile Glu Pro Ile Asp Ser Val Met Gly Met Ile Gly Ile Leu		
450	455	460
Pro Asp Met Thr Pro Ser Thr Glu Met Ser Leu Arg Gly Ile Arg Val		
465	470	475
Ser Lys Met Gly Val Asp Glu Tyr Ser Ser Thr Glu Arg Val Val Val		
485	490	495
Ser Ile Asp Arg Phe Leu Arg Val Arg Asp Gln Arg Gly Asn Val Leu		
500	505	510
Leu Ser Pro Glu Glu Val Ser Glu Thr Gln Gly Thr Glu Lys Leu Thr		
515	520	525
Ile Thr Tyr Ser Ser Ser Met Met Trp Glu Ile Asn Gly Pro Glu Ser		
530	535	540
Val Leu Val Asn Thr Tyr Gln Trp Ile Ile Arg Asn Trp Glu Ser Leu		
545	550	555
Lys Ile Gln Trp Ser Gln Asp Pro Thr Met Leu Tyr Asn Lys Met Glu		
565	570	575
Phe Glu Pro Phe Gln Ser Leu Val Pro Lys Ala Thr Arg Ser Arg Tyr		
580	585	590
Ser Gly Phe Val Arg Thr Leu Phe Gln Gln Met Arg Asp Val Leu Gly		
595	600	605
Thr Phe Asp Thr Val Gln Ile Ile Lys Leu Leu Pro Phe Ala Ala Ala		
610	615	620
Pro Pro Glu Gln Ser Arg Ile Gln Phe Ser Ser Leu Thr Val Asn Val		
625	630	635
Arg Gly Ser Gly Leu Arg Ile Leu Val Arg Gly Asn Ser Pro Val Phe		
645	650	655
Asn Tyr Asn Lys Ala Thr Lys Arg Leu Thr Ile Leu Gly Lys Asp Ala		
660	665	670
Gly Ala Leu Thr Glu Asp Pro Asp Glu Gly Thr Ala Gly Val Glu Ser		
675	680	685
Ala Val Leu Arg Gly Phe Leu Ile Leu Gly Lys Glu Asp Lys Arg Tyr		
690	695	700

Gly Pro Ala Leu Ser Ile Asn Glu Leu Ser Asn Leu Ala Lys Gly Glu
705 710 715 720

Lys Ala Asn Val Leu Ile Gly Gln Gly Asp Val Val Leu Val Met Lys
725 730 735

Arg Lys Arg Asn Ser Ser Ile Leu Thr Asp Ser Gln Thr Ala Thr Lys
740 745 750

Arg Ile Arg Met Ala Ile Asn
755

<210> 3

<211> 2271

<212> ADN

<213> Virut cúm A

<400> 3

```

atggatgtca acccgactct acttttccta aaggttccag cgcaaaatgc cataagcacc      60
acattccctt atactggaga tcctccatac agccatggaa caggaacagg atacaccatg      120
gacacagtca acagaacaca ccagtattca gaaaaaggga aatggacgac aaacacagag      180
actggggcac ccagctcaa ccgattgat ggaccactac ccgatgataa tgaaccaagt      240
gggtatgcac aaacagactg tgtcctggag gccatggctt tccttgaaga atccccacca      300
gggatatttg agaattcatg ccttgaaaca atggaaattg tccaacaaac aagggtggat      360
aaactaactc aaggctcgca gacttatgat tggacattaa acagaaatca accggcagca      420
actgcattgg ccaacaccat agaagttttt agatcaaacg gtctaacagc taatgagtca      480
ggaaggctaa tagatttcct aaaggatgta atggaatcaa tggataagga ggaaatagag      540
ataacaacac attttcaaag aaaaaggaga gtaagagaca acatgaccaa gaagatggtc      600
acacaaagaa caatagggaa gaaaaaaca aaattgaata agagaagtta tctaataaga      660
gcactgacat tgaatacgat gaccaaagat gcagagagag gcaagttaaa aaggagggct      720
atcgcaacac ctgggatgca gattagaggg ttcgtgtact ttgttgagac tttagctaga      780
agcatctgcg aaaagcttga acagtccgga ctcccagtag ggggcaatga aaagaaagcc      840
aaattggcaa atgttgtgag aaagatgatg actaattcac aagacacaga gctttctttc      900
acaatcactg gagataacac taaatggaat gaaaaccaga atcctcgaat gttcctggcg      960
atgatcacat acattaccag aaatcaaccc gagtgggttca gaaacatact gagtatggca     1020
ccaataatgt tctcaaaca aatggcaaga ctaggaaaag ggtacatggt cgagagtaaa     1080
agaatgaagc tccgaacaca ggtaccagca gaaatgctag caagcattga tcttaagtat     1140
ttcaatgaat caacaaggaa gaaaattgag aaaataaggc ctctcctaata agatggcaca     1200
gcatcattga gccctgggat gatgatgggc atgttcaaca tgctaagtac ggttttgga      1260
gtctcaatac tgaatcttgg acaaaagaaa tacaccagga caacatactg gtgggatgga      1320
ctccaatcct cagacgattt tgccctcata gtaaatagcac caaatcatga gggaatacaa      1380
gcaggagtgg atagattcta caggacctgc aagttagtag ggatcaacat gagcaaaaag      1440

```

```

aagtcctata taaataagac tgggacattt gaattcacaa gcttttttta tcgctatggg      1500
ttttagtagcta attttagcat ggagctgccc agttttggag tgtctggaat aaacgaatca      1560
gctgatatga gcatcgaggt aacagtgata aagaacaaca tgataaataa tgatcttgga      1620
cctgcaacag cccagatggc cctccagttg ttcatcaaag actacagata cacatataga      1680
tgccatagag gggacacaca aatccagacg agaagatcat tcgagctaaa gagcctgtgg      1740
aatcaaactc aatcaaaggc aggattatta gtatctgatg gaggaccaaa tttatacaat      1800
atccggaatc ttcacattcc tgaagtctgc ttaaaatggg agctaattga tgaggattat      1860
cggggaagac tttgtaatcc cctgaatccc tttgtcagcc ataaagagat tgattctgta      1920
aacagtgtctg tgggtgatgcc agcccatggg ccagccaaaa gtatggagta tgatgccgtt      1980
gcaactacac actcctggat tccaagagg aaccgctcta ttctcaacac aagccaaagg      2040
ggaattcttg aggatgaaca gatgtaccag aagtgtgca acctgttcga gaaatttttc      2100
cctagtagtt catacagaag accagttgga atttctagca tgggtggaggc catggtgtct      2160
agggcccggg ttgatgccag gattgacttc gagtctggac ggattaagaa agaagagttc      2220
tctgagatca tgaagatctg ttccaccatt gaagaactca gacggcaaaa g      2271

```

```

<210> 4
<211> 757
<212> PRT
<213> Virut cúm A

```

```

<400> 4

```

```

Met Asp Val Asn Pro Thr Leu Leu Phe Leu Lys Val Pro Ala Gln Asn
1              5              10              15

```

```

Ala Ile Ser Thr Thr Phe Pro Tyr Thr Gly Asp Pro Pro Tyr Ser His
              20              25              30

```

```

Gly Thr Gly Thr Gly Tyr Thr Met Asp Thr Val Asn Arg Thr His Gln
35              40              45

```

```

Tyr Ser Glu Lys Gly Lys Trp Thr Thr Asn Thr Glu Thr Gly Ala Pro
50              55              60

```

```

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

```

```

Gly Tyr Ala Gln Thr Asp Cys Val Leu Glu Ala Met Ala Phe Leu Glu
85              90              95

```

```

Glu Ser His Pro Gly Ile Phe Glu Asn Ser Cys Leu Glu Thr Met Glu
100             105             110

```

```

Ile Val Gln Gln Thr Arg Val Asp Lys Leu Thr Gln Gly Arg Gln Thr
115             120             125

```

```

Tyr Asp Trp Thr Leu Asn Arg Asn Gln Pro Ala Ala Thr Ala Leu Ala

```

130	135	140
Asn Thr Ile Glu Val Phe Arg Ser Asn Gly Leu Thr Ala Asn Glu Ser		
145	150	155 160
Gly Arg Leu Ile Asp Phe Leu Lys Asp Val Met Glu Ser Met Asp Lys		
	165 170	175
Glu Glu Ile Glu Ile Thr Thr His Phe Gln Arg Lys Arg Arg Val Arg		
	180 185	190
Asp Asn Met Thr Lys Lys Met Val Thr Gln Arg Thr Ile Gly Lys Lys		
	195 200	205
Lys Gln Lys Leu Asn Lys Arg Ser Tyr Leu Ile Arg Ala Leu Thr Leu		
	210 215	220
Asn Thr Met Thr Lys Asp Ala Glu Arg Gly Lys Leu Lys Arg Arg Ala		
225	230 235	240
Ile Ala Thr Pro Gly Met Gln Ile Arg Gly Phe Val Tyr Phe Val Glu		
	245 250	255
Thr Leu Ala Arg Ser Ile Cys Glu Lys Leu Glu Gln Ser Gly Leu Pro		
	260 265	270
Val Gly Gly Asn Glu Lys Lys Ala Lys Leu Ala Asn Val Val Arg Lys		
	275 280	285
Met Met Thr Asn Ser Gln Asp Thr Glu Leu Ser Phe Thr Ile Thr Gly		
	290 295	300
Asp Asn Thr Lys Trp Asn Glu Asn Gln Asn Pro Arg Met Phe Leu Ala		
305	310 315	320
Met Ile Thr Tyr Ile Thr Arg Asn Gln Pro Glu Trp Phe Arg Asn Ile		
	325 330	335
Leu Ser Met Ala Pro Ile Met Phe Ser Asn Lys Met Ala Arg Leu Gly		
	340 345	350
Lys Gly Tyr Met Phe Glu Ser Lys Arg Met Lys Leu Arg Thr Gln Val		
	355 360	365
Pro Ala Glu Met Leu Ala Ser Ile Asp Leu Lys Tyr Phe Asn Glu Ser		
	370 375	380
Thr Arg Lys Lys Ile Glu Lys Ile Arg Pro Leu Leu Ile Asp Gly Thr		
385	390 395	400
Ala Ser Leu Ser Pro Gly Met Met Met Gly Met Phe Asn Met Leu Ser		
	405 410	415

Thr Val Leu Gly Val Ser Ile Leu Asn Leu Gly Gln Lys Lys Tyr Thr
 420 425 430

Arg Thr Thr Tyr Trp Trp Asp Gly Leu Gln Ser Ser Asp Asp Phe Ala
 435 440 445

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

Arg Phe Tyr Arg Thr Cys Lys Leu Val Gly Ile Asn Met Ser Lys Lys
 465 470 475 480

Lys Ser Tyr Ile Asn Lys Thr Gly Thr Phe Glu Phe Thr Ser Phe Phe
 485 490 495

Tyr Arg Tyr Gly Phe Val Ala Asn Phe Ser Met Glu Leu Pro Ser Phe
 500 505 510

Gly Val Ser Gly Ile Asn Glu Ser Ala Asp Met Ser Ile Gly Val Thr
 515 520 525

Val Ile Lys Asn Asn Met Ile Asn Asn Asp Leu Gly Pro Ala Thr Ala
 530 535 540

Gln Met Ala Leu Gln Leu Phe Ile Lys Asp Tyr Arg Tyr Thr Tyr Arg
 545 550 555 560

Cys His Arg Gly Asp Thr Gln Ile Gln Thr Arg Arg Ser Phe Glu Leu
 565 570 575

Lys Ser Leu Trp Asn Gln Thr Gln Ser Lys Ala Gly Leu Leu Val Ser
 580 585 590

Asp Gly Gly Pro Asn Leu Tyr Asn Ile Arg Asn Leu His Ile Pro Glu
 595 600 605

Val Cys Leu Lys Trp Glu Leu Met Asp Glu Asp Tyr Arg Gly Arg Leu
 610 615 620

Cys Asn Pro Leu Asn Pro Phe Val Ser His Lys Glu Ile Asp Ser Val
 625 630 635 640

Asn Ser Ala Val Val Met Pro Ala His Gly Pro Ala Lys Ser Met Glu
 645 650 655

Tyr Asp Ala Val Ala Thr Thr His Ser Trp Ile Pro Lys Arg Asn Arg
 660 665 670

Ser Ile Leu Asn Thr Ser Gln Arg Gly Ile Leu Glu Asp Glu Gln Met
 675 680 685

Tyr Gln Lys Cys Cys Asn Leu Phe Glu Lys Phe Phe Pro Ser Ser Ser

690

695

700

Tyr Arg Arg Pro Val Gly Ile Ser Ser Met Val Glu Ala Met Val Ser
705 710 715 720

Arg Ala Arg Ile Asp Ala Arg Ile Asp Phe Glu Ser Gly Arg Ile Lys
725 730 735

Lys Glu Glu Phe Ser Glu Ile Met Lys Ile Cys Ser Thr Ile Glu Glu
740 745 750

Leu Arg Arg Gln Lys
755

<210> 5

<211> 2148

<212> ADN

<213> Virut cúm A

<400> 5

atggaagact ttgtgcgaca atgcttcaat ccaatgatcg tcgagcttgc ggaaaaggca 60
atgaaagaat atggagaaga tccgaaaatt gaaactaaca aattcgctgc aatatgcaca 120
cacttggaag tatgtttcat gtattcagat ttccatttca ttgacgagcg ggtggaatca 180
atcattgtag aatctggtga tccaaatgca ttactgaagc accgatttga gataattgaa 240
ggacgagacc ggaccatggc ctggacagta gtgaacagta tctgcaacac cacaggggta 300
gagaagccta aatttcttcc ggatttatac gactacaaag aaaatcggtt cattgaaatt 360
ggagtgcac gaagggaggt ccacatatac tacctagaga aagccaacaa aataaaatcc 420
gagaagacac acattcacat tttttcatte actggagagg agatggccac caaagcagac 480
tacacccttg atgaagaaag cagggcaaga atcaaaacca ggcttttcac cataagacaa 540
gaaatggcaa gtaggggtct atgggattcc tttcgtcagt ccgaaagagg cgaggagaca 600
attgaagaaa gatttgaaat tacaggaacc atgcgcagac ttgccgacca aagtctccca 660
ccgaacttct ccagtcttga aaactttaga gcttatgtag atgggttcga accaaacggc 720
tgcatgtagg gcaagcttcc tcaaattgca aaagaagtga gcgccccaat tgaacccttc 780
ttgaagacaa caccacgccc tctaaaattg cctgatgggc ctcttgctc tcagcgggtca 840
aagttcttgc tgatggatgc tctgaaacta agtattgaag acccgagtca tgaaggagaa 900
ggaataccac tatatgatgc aatcaagtgc atgaagacat tttttggctg gaaagaaccc 960
aacataatca aaccacatga gaaaggcata aaccccaatt acctactggc ttggaagcag 1020
gtgctagcag agctccaaga cattgaaaat gaagagaaga tcccaaagac aaagaacatg 1080
aggagaacaa gccaatgaa gtgggcactc ggtgagaata tggcaccaga gaaagtagat 1140
tttgatgact gcaaagatgt tggatgattt aaacagtatg acagcgacga gccagagccc 1200
agatctctag caagttgggt ccaaaatgaa ttcaacaagg catgtgaatt gaccgattca 1260
agctggatag aacttgatga gataggagaa gatattgcac cgattgaaca catcgcaagt 1320
atgaggagga actattttac agcagaagtg tccattgtga gggctacgga atacataatg 1380

```

aagggagtgt acataaacac ggctttgctt aatgcatctt gtgcagccat ggatgacttt 1440
cagctgatcc caatgataag caaatgcagg accaaagaag gaagacgaaa aacaaatctg 1500
tatgggttca ttataaaagg aagggtcccat ctgaggaatg atactgacgt ggtgaacttt 1560
gtaagcatgg agttctccct caccgacccg agactggagc cacacaaatg ggaaaaatac 1620
tgtgttcttg aaataggaga catgctcctg aggactgcga taggccaagt gtcgaggccc 1680
atgttcttat atgtgagaac caatggaacc tccaagatca agatgaaatg gggcatggaa 1740
atgaggcgct gccttcttca atctcttcag cagattgaga gcatgattga ggctgagtct 1800
tctgtaaaag agaaagacat gaccaaggaa ttttttgaag acaaactgga aacatggcca 1860
attggagaat cacccaaagg agtggaggaa ggctctattg ggaaagtgtg caggacctta 1920
ctggcaaaat ctgtattcaa cagtctatac gcgtctccac aacttgaggg attttcggct 1980
gaatcgagaa agttgcttct cattgttcag gcacttaggg acaacctgga acctggaacc 2040
ttcgcatttg gggggctata tgaagcaatc gaggagtgcc tgattaatga tccctggggt 2100
ttgcttaatg catcttggtt caactccttc ctcacacatg cactgaaa 2148

```

```

<210> 6
<211> 716
<212> PRT
<213> Virut cúm A

```

```

<400> 6

```

```

Met Glu Asp Phe Val Arg Gln Cys Phe Asn Pro Met Ile Val Glu Leu
1           5           10          15

```

```

Ala Glu Lys Ala Met Lys Glu Tyr Gly Glu Asp Pro Lys Ile Glu Thr
          20           25           30

```

```

Asn Lys Phe Ala Ala Ile Cys Thr His Leu Glu Val Cys Phe Met Tyr
          35           40           45

```

```

Ser Asp Phe His Phe Ile Asp Glu Arg Gly Glu Ser Ile Ile Val Glu
          50           55           60

```

```

Ser Gly Asp Pro Asn Ala Leu Leu Lys His Arg Phe Glu Ile Ile Glu
65           70           75           80

```

```

Gly Arg Asp Arg Thr Met Ala Trp Thr Val Val Asn Ser Ile Cys Asn
          85           90           95

```

```

Thr Thr Gly Val Glu Lys Pro Lys Phe Leu Pro Asp Leu Tyr Asp Tyr
          100          105          110

```

```

Lys Glu Asn Arg Phe Ile Glu Ile Gly Val Thr Arg Arg Glu Val His
          115          120          125

```

```

Ile Tyr Tyr Leu Glu Lys Ala Asn Lys Ile Lys Ser Glu Lys Thr His
          130          135          140

```

Ile His Ile Phe Ser Phe Thr Gly Glu Glu Met Ala Thr Lys Ala Asp
 145 150 155 160
 Tyr Thr Leu Asp Glu Glu Ser Arg Ala Arg Ile Lys Thr Arg Leu Phe
 165 170 175
 Thr Ile Arg Gln Glu Met Ala Ser Arg Gly Leu Trp Asp Ser Phe Arg
 180 185 190
 Gln Ser Glu Arg Gly Glu Glu Thr Ile Glu Glu Arg Phe Glu Ile Thr
 195 200 205
 Gly Thr Met Arg Arg Leu Ala Asp Gln Ser Leu Pro Pro Asn Phe Ser
 210 215 220
 Ser Leu Glu Asn Phe Arg Ala Tyr Val Asp Gly Phe Glu Pro Asn Gly
 225 230 235 240
 Cys Ile Glu Gly Lys Leu Ser Gln Met Ser Lys Glu Val Ser Ala Gln
 245 250 255
 Ile Glu Pro Phe Leu Lys Thr Thr Pro Arg Pro Leu Lys Leu Pro Asp
 260 265 270
 Gly Pro Pro Cys Ser Gln Arg Ser Lys Phe Leu Leu Met Asp Ala Leu
 275 280 285
 Lys Leu Ser Ile Glu Asp Pro Ser His Glu Gly Glu Gly Ile Pro Leu
 290 295 300
 Tyr Asp Ala Ile Lys Cys Met Lys Thr Phe Phe Gly Trp Lys Glu Pro
 305 310 315 320
 Asn Ile Ile Lys Pro His Glu Lys Gly Ile Asn Pro Asn Tyr Leu Leu
 325 330 335
 Ala Trp Lys Gln Val Leu Ala Glu Leu Gln Asp Ile Glu Asn Glu Glu
 340 345 350
 Lys Ile Pro Lys Thr Lys Asn Met Arg Arg Thr Ser Gln Leu Lys Trp
 355 360 365
 Ala Leu Gly Glu Asn Met Ala Pro Glu Lys Val Asp Phe Asp Asp Cys
 370 375 380
 Lys Asp Val Gly Asp Leu Lys Gln Tyr Asp Ser Asp Glu Pro Glu Pro
 385 390 395 400
 Arg Ser Leu Ala Ser Trp Val Gln Asn Glu Phe Asn Lys Ala Cys Glu
 405 410 415
 Leu Thr Asp Ser Ser Trp Ile Glu Leu Asp Glu Ile Gly Glu Asp Ile

420	425	430
Ala Pro Ile Glu His Ile Ala Ser Met Arg Arg Asn Tyr Phe Thr Ala		
435	440	445
Glu Val Ser His Cys Arg Ala Thr Glu Tyr Ile Met Lys Gly Val Tyr		
450	455	460
Ile Asn Thr Ala Leu Leu Asn Ala Ser Cys Ala Ala Met Asp Asp Phe		
465	470	475 480
Gln Leu Ile Pro Met Ile Ser Lys Cys Arg Thr Lys Glu Gly Arg Arg		
	485	490 495
Lys Thr Asn Leu Tyr Gly Phe Ile Ile Lys Gly Arg Ser His Leu Arg		
	500	505 510
Asn Asp Thr Asp Val Val Asn Phe Val Ser Met Glu Phe Ser Leu Thr		
	515	520 525
Asp Pro Arg Leu Glu Pro His Lys Trp Glu Lys Tyr Cys Val Leu Glu		
	530	535 540
Ile Gly Asp Met Leu Leu Arg Thr Ala Ile Gly Gln Val Ser Arg Pro		
545	550	555 560
Met Phe Leu Tyr Val Arg Thr Asn Gly Thr Ser Lys Ile Lys Met Lys		
	565	570 575
Trp Gly Met Glu Met Arg Arg Cys Leu Leu Gln Ser Leu Gln Gln Ile		
	580	585 590
Glu Ser Met Ile Glu Ala Glu Ser Ser Val Lys Glu Lys Asp Met Thr		
	595	600 605
Lys Glu Phe Phe Glu Asn Lys Ser Glu Thr Trp Pro Ile Gly Glu Ser		
	610	615 620
Pro Lys Gly Val Glu Glu Gly Ser Ile Gly Lys Val Cys Arg Thr Leu		
625	630	635 640
Leu Ala Lys Ser Val Phe Asn Ser Leu Tyr Ala Ser Pro Gln Leu Glu		
	645	650 655
Gly Phe Ser Ala Glu Ser Arg Lys Leu Leu Leu Ile Val Gln Ala Leu		
	660	665 670
Arg Asp Asn Leu Glu Pro Gly Thr Phe Asp Leu Gly Gly Leu Tyr Glu		
	675	680 685
Ala Ile Glu Glu Cys Leu Ile Asn Asp Pro Trp Val Leu Leu Asn Ala		
690	695	700

Ser Trp Phe Asn Ser Phe Leu Thr His Ala Leu Lys
 705 710 715

<210> 7
 <211> 1692
 <212> ADN
 <213> Virut cúm A

<400> 7
 atgaaggcaa tactagtagt cctgctatat acattttacaa ccgcaaattgc cgacacatta 60
 tgcataggtt atcatgcgaa caattcaact gacaccgtag acacagtgtc agaaaaaaat 120
 gtaacagtaa cacactctgt caaccttcta gaaaacaggc ataatgggaa actatgtaaa 180
 ctaagagggg tagctccatt gcatttgggt aaatgtaaca ttgctggctg gctcctggga 240
 aatccagagt gtgagtcaat ctccaaagca agctcatggt cctacattgt ggaaacatct 300
 aattcagaca atgggacgtg ttaccagga gatttcatca attatgagga gctaagagag 360
 cagttgagct cagtgtcatc atttgaaaga ttgagatat tcccatgac aagttcatgg 420
 cccaatcatg acacgaacag aggtgtgacg gcagcatgtc ctacagctgg gacaaatagc 480
 ttctacaaaa atttaatatg gctggtcaaa aaaggaaatt catacccaa gatcaacaaa 540
 tctacatta acaacaaaga gaaagaagtt ctctgtctat gggccataca tcatccacct 600
 accaatgccg accaacaag cctctaccaa atgcagatg cctatgtttt tgtggggtca 660
 tcaagataca gcaggaagtt cgagccagaa atagcaacaa gaccaaggt gagagaccaa 720
 gcaggagaa tgaactatta ctggacattg gtagagcctg gagacaagat aacattcgaa 780
 gcaactggaa atctagtgtt accgagatat gccttcgat tgaagaaa ttctggatct 840
 ggtattatca ttccagatac atcagtcac gattgtgata cgacttgta gacaccaat 900
 ggtgctataa acaccagcct cccatttcaa aatatacatc cagtcacaat tggagaatgt 960
 ccaaaatatg taaaagtac taaactgaga atggccacag gattaaggaa tatcccgctc 1020
 attcaatcta gaggcctgtt tggggccatt gctggcttta ttgaagggg ctggacagga 1080
 atgatagacg gatggtacg ttaccacat caaatgagc agggatcagg atatgcagcc 1140
 gacctgaaaa gcacacagaa tgccattgac gggatcacta acaaggtaaa ttctgttatt 1200
 gaaaagatga acacacaatt cacagcagta ggtaaagagt tcagccactt ggaaagaaga 1260
 atagagaatt taaataaaaa ggttgatgat ggttttctag atatttggac ttacaatgcc 1320
 gaactgttgg ttctgttga gaatgaaaga actttggatt accacgattc aaatgtgaaa 1380
 aacttatatg aaaaagtaag aagccaacta aaaaacaatg ccaaagaaat tggaaatggc 1440
 tgctttgaat ttaccacaa atgtgatgac acgtgtatgg aaagcgtcaa aaatgggact 1500
 tatgattacc caaataactc agaggaagca aaactaaaca gagaggaaat agatggggta 1560
 aagttggaat caacaaggt ttaccaaatt ttggcgatct attcaacggt cgccagttca 1620
 ttggtactgg tagtctccct gggggcaatc agtttctgga tgtgctctaa tgggtcgcta 1680
 cagtgcagaa ta 1692

<210> 8
 <211> 564
 <212> PRT
 <213> Virut cúm A

<400> 8

Met Lys Ala Ile Leu Val Val Leu Leu Tyr Thr Phe Thr Thr Ala Asn
 1 5 10 15

Ala Asp Thr Leu Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr
 20 25 30

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

Leu Leu Glu Asn Arg His Asn Gly Lys Leu Cys Lys Leu Arg Gly Val
 50 55 60

Ala Pro Leu His Leu Gly Lys Cys Asn Ile Ala Gly Trp Leu Leu Gly
 65 70 75 80

Asn Pro Glu Cys Glu Ser Ile Ser Lys Ala Ser Ser Trp Ser Tyr Ile
 85 90 95

Val Glu Thr Ser Asn Ser Asp Asn Gly Thr Cys Tyr Pro Gly Asp Phe
 100 105 110

Ile Asn Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe
 115 120 125

Glu Arg Phe Glu Ile Phe Pro Met Thr Ser Ser Trp Pro Asn His Asp
 130 135 140

Thr Asn Arg Gly Val Thr Ala Ala Cys Pro His Ala Gly Thr Asn Ser
 145 150 155 160

Phe Tyr Lys Asn Leu Ile Trp Leu Val Lys Lys Gly Asn Ser Tyr Pro
 165 170 175

Lys Ile Asn Lys Ser Tyr Ile Asn Asn Lys Glu Lys Glu Val Leu Val
 180 185 190

Leu Trp Ala Ile His His Pro Pro Thr Asn Ala Asp Gln Gln Ser Leu
 195 200 205

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

Arg Lys Phe Glu Pro Glu Ile Ala Thr Arg Pro Lys Val Arg Asp Gln
 225 230 235 240

Ala Gly Arg Met Asn Tyr Tyr Trp Thr Leu Val Glu Pro Gly Asp Lys
 245 250 255

Ile Thr Phe Glu Ala Thr Gly Asn Leu Val Val Pro Arg Tyr Ala Phe
 260 265 270

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

Val His Asp Cys Asp Thr Thr Cys Gln Thr Pro Asn Gly Ala Ile Asn
 290 295 300

Thr Ser Leu Pro Phe Gln Asn Ile His Pro Val Thr Ile Gly Glu Cys
 305 310 315 320

Pro Lys Tyr Val Lys Ser Thr Lys Leu Arg Met Ala Thr Gly Leu Arg
 325 330 335

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

Phe Ile Glu Gly Gly Trp Thr Gly Met Ile Asp Gly Trp Tyr Gly Tyr
 355 360 365

His His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Leu Lys Ser
 370 375 380

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

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

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

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

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

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

Cys Phe Glu Phe Tyr His Lys Cys Asp Asp Thr Cys Met Glu Ser Val
 485 490 495

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

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

Gln Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Val

530

535

540

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

Gln Cys Arg Ile

<210> 9

<211> 1512

<212> ADN

<213> Virut cúm A

<400> 9

atgagtgaca tcgaagccat ggcgtctcaa ggcaccaaac gatcatatga acaaattggag . 60
actggtggtg aacgccagga tgccacagaa atcagagcat ctgtcggaaag aatgattggt 120
ggaatcggga aattctacat ccaaattgtgc actgaactca aactcagtga ctatgaggga 180
cgactaatcc aaaatagcat aacaatagag agaattggtgc tctctgcttt tgatgagaga 240
agaaataaat acctagaaga gcatcccatg gctgggaaag accctaagaa aactggagga 300
cccatatata gaagagtaga cggaaagtgg atgagagaac tcattcttta tgacaaagaa 360
gaaataagga gagtttggcg ccaggcaaac aatggtgatg atgcaacagc tgggtcttact 420
catatcatga tttggcattc caatctgaat gatgccacgt accagagaac aagagcactt 480
gttcgcaccg gaatggatcc cagaattgtgc tctctaattgc aaggttcaac acttcccaga 540
agggtctggag cagcaggtgc tgcagtgaag ggagttggaa caataacaat ggaattaatc 600
agaatgatca aacgggggat caatgaccga aatttctgga gaggtgaaaa tggagaagg 660
acaaggattg catatgaaag aatgtgcaat attctcaaag gaaaatttca gacagctgcc 720
caaagggcaa tgatggatca agtgagagaa agtcggaacc cagggaacgc tgagattgaa 780
gatctcattt tcttggcagc gtcagcactt atcctaaggg gatcagttgc acataagtct 840
tgcttgcctg cttgcgtgta tgggcttgca gtggcaagtg ggcattgactt tgaaagggaa 900
gggtattcgc tggtcgggat agaccattt aaattactcc agaacagtca agtggttcagc 960
ctggtaagac caaatgaaaa ccagctcac aagagtcaat tagtgtggat ggcattgccac 1020
tctgctgcat ttgaggatct aagggtctca agtttcataa gaggaagaa agtgattcca 1080
aggggaaagc tttccacaag aggggttcag attgcttcaa atgagaatgt ggaagccatg 1140
gattccaata ccttagagct gagaagcaga tactgggcta taaggaccag aagtggagga 1200
aataactaatc aacagaaagc atccgcaggc cagatcagtg tgcaacctac attctcagtg 1260
caacgggaatc tcccttttga aagagcaacc gttatggcag ctttcagcgg aaacaatgaa 1320
ggacggacat ccgatattgc gacagaaatt ataaggatga tggaaaatgc aaaaccagaa 1380
gatttgcctt tccagggcg gggagtcttc gagctctcgg acgaaaaggc aacgagcccg 1440
atcgtgcctt cctttgacat gagtaatgaa ggggtcttatt tcttcggaga caatgcagag 1500
gagtatgaca gt 1512

<210> 10
 <211> 504
 <212> PRT
 <213> Virut cúm A

<400> 10

```

Met Ser Asp Ile Glu Ala Met Ala Ser Gln Gly Thr Lys Arg Ser Tyr
1           5           10           15

Glu Gln Met Glu Thr Gly Gly Glu Arg Gln Asp Ala Thr Glu Ile Arg
          20           25           30

Ala Ser Val Gly Arg Met Ile Gly Gly Ile Gly Lys Phe Tyr Ile Gln
          35           40           45

Met Cys Thr Glu Leu Lys Leu Ser Asp Tyr Glu Gly Arg Leu Ile Gln
          50           55           60

Asn Ser Ile Thr Ile Glu Arg Met Val Leu Ser Ala Phe Asp Glu Arg
65           70           75           80

Arg Asn Lys Tyr Leu Glu Glu His Pro Ser Ala Gly Lys Asp Pro Lys
          85           90           95

Lys Thr Gly Gly Pro Ile Tyr Arg Arg Val Asp Gly Lys Trp Met Arg
          100          105          110

Glu Leu Ile Leu Tyr Asp Lys Glu Glu Ile Arg Arg Val Trp Arg Gln
          115          120          125

Ala Asn Asn Gly Asp Asp Ala Thr Ala Gly Leu Thr His Ile Met Ile
          130          135          140

Trp His Ser Asn Leu Asn Asp Ala Thr Tyr Gln Arg Thr Arg Ala Leu
145          150          155          160

Val Arg Thr Gly Met Asp Pro Arg Met Cys Ser Leu Met Gln Gly Ser
          165          170          175

Thr Leu Pro Arg Arg Ser Gly Ala Ala Gly Ala Ala Val Lys Gly Val
          180          185          190

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

Asp Arg Asn Phe Trp Arg Gly Glu Asn Gly Arg Arg Thr Arg Ile Ala
          210          215          220

Tyr Glu Arg Met Cys Asn Ile Leu Lys Gly Lys Phe Gln Thr Ala Ala
225          230          235          240

Gln Arg Ala Met Met Asp Gln Val Arg Glu Ser Arg Asn Pro Gly Asn
          245          250          255

```

Ala Glu Ile Glu Asp Leu Ile Phe Leu Ala Arg Ser Ala Leu Ile Leu
 260 265 270
 Arg Gly Ser Val Ala His Lys Ser Cys Leu Pro Ala Cys Val Tyr Gly
 275 280 285
 Leu Ala Val Ala Ser Gly His Asp Phe Glu Arg Glu Gly Tyr Ser Leu
 290 295 300
 Val Gly Ile Asp Pro Phe Lys Leu Leu Gln Asn Ser Gln Val Phe Ser
 305 310 315 320
 Leu Val Arg Pro Asn Glu Asn Pro Ala His Lys Ser Gln Leu Val Trp
 325 330 335
 Met Ala Cys His Ser Ala Ala Phe Glu Asp Leu Arg Val Ser Ser Phe
 340 345 350
 Ile Arg Gly Lys Lys Val Ile Pro Arg Gly Lys Leu Ser Thr Arg Gly
 355 360 365
 Val Gln Ile Ala Ser Asn Glu Asn Val Glu Ala Met Asp Ser Asn Thr
 370 375 380
 Leu Glu Leu Arg Ser Arg Tyr Trp Ala Ile Arg Thr Arg Ser Gly Gly
 385 390 395 400
 Asn Thr Asn Gln Gln Lys Ala Ser Ala Gly Gln Ile Ser Val Gln Pro
 405 410 415
 Thr Phe Ser Val Gln Arg Asn Leu Pro Phe Glu Arg Ala Thr Val Met
 420 425 430
 Ala Ala Phe Ser Gly Asn Asn Glu Gly Arg Thr Ser Asp Met Arg Thr
 435 440 445
 Glu Ile Ile Arg Met Met Glu Asn Ala Lys Pro Glu Asp Leu Ser Phe
 450 455 460
 Gln Gly Arg Gly Val Phe Glu Leu Ser Asp Glu Lys Ala Thr Ser Pro
 465 470 475 480
 Ile Val Pro Ser Phe Asp Met Ser Asn Glu Gly Ser Tyr Phe Phe Gly
 485 490 495
 Asp Asn Ala Glu Glu Tyr Asp Ser
 500

<210> 11
 <211> 1407
 <212> ADN
 <213> Virut cúm A

<400> 11
 atgaatacaa atcaaagaat aataaccatt gggacagttt gtctgatagt tggaataatt 60
 agtctattgt tacagatagg aaacatgggt tcgttatgga tcagccattc aattcagact 120
 gaagggaaaa atcatactga gatgtgcaat caaaatgtca ttacatatgt aaataacaca 180
 tgggtgaacc gaacttatgt aaacattagc aataccaaaa ttgttaatgt acaggacgtg 240
 gtttcagtaa tattaaccgg caattcctct ctctgcccaa taagtgggtg ggctatatac 300
 agcaaagaca atagcataag gattgggttct aaaggggaca tttttgtcat aagagaacca 360
 ttcatttcat gctctcactt ggaatgcaga actttttttc tgaccaagg cgctttgctg 420
 aatgacaagc attctaattg aaccgtcaag gacaggagtc cctatagaac tttaatgagc 480
 tgtcccatcg gtgaagctcc atctocatat aactcaagggt tcgaatcagt tgcttgggtca 540
 gcaagtgcac gccatgatgg gatgggatgg ctaacaatcg gaatctccgg tccagataat 600
 ggagcagtag ctgtttttaa atacaacgggt ataataacag atacaataaa aagttggaga 660
 aacaaaatat taagaacaca agagtcagaa tgtgtttgta tgaacgggtc ttgttttact 720
 gtattaactg atggcccaag caatgggcaa gcatcgtaca aaatattcaa gatggaaaaa 780
 ggaaaaataa ttaaatcaat tgagctggat gcaccaatt accactatga ggaatgctcc 840
 tgttatcctg atgcaggcaa agtaatgtgt gtttgcagag acaactggca tgctcgaac 900
 cggccatggg tctctttcga tcagaatctt aattatcaaa tagggtacat atgcagtggg 960
 gttttcggtg ataaccgcg ttctaattgat ggaaagggca attgtggccc agtacattct 1020
 aatggagcaa atggagtga aggattctca tataaatatg gtaatggtgt ttggatagga 1080
 aggactaaaa gtatcaactc cagaagtgga tttgaaatga tttgggatcc aaatgggtg 1140
 actggaactg atagtagttt ctctatgaag caggatatta tagcattaac tgattggtca 1200
 ggatacagtg gaagttttgt ccaacatcct gaattaacag gaatgaattg cataaggccc 1260
 tgtttctggg tagaattaat cagagggcaa cccaaggaaa acaccatctg ggctagcgga 1320
 agcagcatct ctttctgtgg tgtaaatggt gaaaccgcaa gctgggtcat gccagacgga 1380
 gctgatctgc cattcaccat tgacaag 1407

<210> 12
 <211> 469
 <212> PRT
 <213> Virut cúm A

<400> 12

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

Val Gly Ile Ile Ser Leu Leu Leu Gln Ile Gly Asn Met Val Ser Leu
 20 25 30

Trp Ile Ser His Ser Ile Gln Thr Glu Gly Lys Asn His Thr Glu Met
 35 40 45

Cys Asn Gln Asn Val Ile Thr Tyr Val Asn Asn Thr Trp Val Asn Arg

50	55	60
Thr Tyr Val Asn Ile Ser Asn Thr Lys Ile Val Asn Val Gln Asp Val		
65	70	75 80
Val Ser Val Ile Leu Thr Gly Asn Ser Ser Leu Cys Pro Ile Ser Gly		
	85	90 95
Trp Ala Ile Tyr Ser Lys Asp Asn Ser Ile Arg Ile Gly Ser Lys Gly		
	100	105 110
Asp Ile Phe Val Ile Arg Glu Pro Phe Ile Ser Cys Ser His Leu Glu		
	115	120 125
Cys Arg Thr Phe Phe Leu Thr Gln Gly Ala Leu Leu Asn Asp Lys His		
	130	135 140
Ser Asn Gly Thr Val Lys Asp Arg Ser Pro Tyr Arg Thr Leu Met Ser		
	145	150 155 160
Cys Pro Ile Gly Glu Ala Pro Ser Pro Tyr Asn Ser Arg Phe Glu Ser		
	165	170 175
Val Ala Trp Ser Ala Ser Ala Cys His Asp Gly Met Gly Trp Leu Thr		
	180	185 190
Ile Gly Ile Ser Gly Pro Asp Asn Gly Ala Val Ala Val Leu Lys Tyr		
	195	200 205
Asn Gly Ile Ile Thr Asp Thr Ile Lys Ser Trp Arg Asn Lys Ile Leu		
	210	215 220
Arg Thr Gln Glu Ser Glu Cys Val Cys Met Asn Gly Ser Cys Phe Thr		
	225	230 235 240
Val Leu Thr Asp Gly Pro Ser Asn Gly Gln Ala Ser Tyr Lys Ile Phe		
	245	250 255
Lys Met Glu Lys Gly Lys Ile Ile Lys Ser Ile Glu Leu Asp Ala Pro		
	260	265 270
Asn Tyr His Tyr Glu Glu Cys Ser Cys Tyr Pro Asp Ala Gly Lys Val		
	275	280 285
Met Cys Val Cys Arg Asp Asn Trp His Ala Ser Asn Arg Pro Trp Val		
	290	295 300
Ser Phe Asp Gln Asn Leu Asn Tyr Gln Ile Gly Tyr Ile Cys Ser Gly		
	305	310 315 320
Val Phe Gly Asp Asn Pro Arg Ser Asn Asp Gly Lys Gly Asn Cys Gly		
	325	330 335

Pro Val His Ser Asn Gly Ala Asn Gly Val Lys Gly Phe Ser Tyr Lys
 340 345 350

Tyr Gly Asn Gly Val Trp Ile Gly Arg Thr Lys Ser Ile Asn Ser Arg
 355 360 365

Ser Gly Phe Glu Met Ile Trp Asp Pro Asn Gly Trp Thr Gly Thr Asp
 370 375 380

Ser Ser Phe Ser Met Lys Gln Asp Ile Ile Ala Leu Thr Asp Trp Ser
 385 390 395 400

Gly Tyr Ser Gly Ser Phe Val Gln His Pro Glu Leu Thr Gly Met Asn
 405 410 415

Cys Ile Arg Pro Cys Phe Trp Val Glu Leu Ile Arg Gly Gln Pro Lys
 420 425 430

Glu Asn Thr Ile Trp Ala Ser Gly Ser Ser Ile Ser Phe Cys Gly Val
 435 440 445

Asn Gly Glu Thr Ala Ser Trp Ser Trp Pro Asp Gly Ala Asp Leu Pro
 450 455 460

Phe Thr Ile Asp Lys
 465

<210> 13
 <211> 756
 <212> ADN
 <213> Virut cúm A

<400> 13
 atgagtcttc taaccgaggt cgaaacgtat gttctctcta tcgtcccgtc agggcccccctc 60
 aaagccgaga tagcacagag gctcgaagac gtttttgcag ggaaaaaacac cgatcttgag 120
 gctctcatgg aatggctaaa gacaagacca atcctgtcac ctctgactaa agggatttta 180
 gggtttgtgt tcacgctcac cgtgcccagt gagcgaggac tgcagcgtag acgttttgtc 240
 cagaatgccc tcaatgggaa tggtagacca aacaacatgg acaaggcggg aaaactgtac 300
 aggaaactaa aaagggaaat aacattccat ggggccaaagg aagtagcgct cagttactct 360
 gctggtgcac ttgccagttg catgggcctc atatacaaca gaatggggac tgtcgccact 420
 gaggtggcat ttggtctggt atgcgcaacc tgtgaacaaa ttgctgattc tcagcatcga 480
 tctcatagac aaatgggtgac aacaaccaat ccaactaatca ggcacgagaa cagaatggta 540
 atagccagca caacagctaa ggccatggaa caaatggctg gatcaagtga acaagcagca 600
 gaggctatgg aggttgccag tcaggctaga caaatggtac aggcaatgag aacaattggg 660
 actcacccta gttccagcac tgggtctaaaa gatgatcttc ttgaaaattt acaggcctat 720
 cagaaacgga tgggagtgca aatgcaacga ttcaag 756

<210> 14
 <211> 252
 <212> PRT
 <213> Virut cúm A

<400> 14

Met Ser Leu Leu Thr Glu Val Glu Thr Tyr Val Leu Ser Ile Val Pro
 1 5 10 15

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

Ala Gly Lys Asn Thr Asp Leu Glu Ala Leu Met Glu Trp Leu Lys Thr
 35 40 45

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

Thr Leu Thr Val Pro Ser Glu Arg Gly Leu Gln Arg Arg Arg Phe Val
 65 70 75 80

Gln Asn Ala Leu Asn Gly Asn Gly Asp Pro Asn Asn Met Asp Lys Ala
 85 90 95

Val Lys Leu Tyr Arg Lys Leu Lys Arg Glu Ile Thr Phe His Gly Ala
 100 105 110

Lys Glu Val Ala Leu Ser Tyr Ser Ala Gly Ala Leu Ala Ser Cys Met
 115 120 125

Gly Leu Ile Tyr Asn Arg Met Gly Thr Val Ala Thr Glu Val Ala Phe
 130 135 140

Gly Leu Val Cys Ala Thr Cys Glu Gln Ile Ala Asp Ser Gln His Arg
 145 150 155 160

Ser His Arg Gln Met Val Thr Thr Thr Asn Pro Leu Ile Arg His Glu
 165 170 175

Asn Arg Met Val Ile Ala Ser Thr Thr Ala Lys Ala Met Glu Gln Met
 180 185 190

Ala Gly Ser Ser Glu Gln Ala Ala Glu Ala Met Glu Val Ala Ser Gln
 195 200 205

Ala Arg Gln Met Val Gln Ala Met Arg Thr Ile Gly Thr His Pro Ser
 210 215 220

Ser Ser Thr Gly Leu Lys Asp Asp Leu Leu Glu Asn Leu Gln Ala Tyr
 225 230 235 240

Gln Lys Arg Met Gly Val Gln Met Gln Arg Phe Lys
 245 250

<210> 15
 <211> 657
 <212> ADN
 <213> Virut cúm A

<400> 15
 atggactcca atactgtgtc aagctttcag gtagactggt tcctttggca catccgcaaa 60
 cgatttgcag acaatggatt gggatgatgcc ccattccttg atcggtccg ccgagatcaa 120
 aagtcacctaa aaggaagggg caacaccctt agcctagaca tcgaaacagc cactcttggt 180
 gggaaacaaa ttgttgagtg gatattgaaa gaggaatcca gcgatatact taagatgacc 240
 attgcatctg tgcctacttc gcgtacctt gctgacatga ccctcgagga aatgtcacga 300
 gactgggttca tgctaatagcc taggcaaaag ataataaggcc ctctttgtgt gcgagtggac 360
 caggcgatca tggaaaagaa catcactactg aaagcgaact tcagtgtgat ctttaaccga 420
 ttagagactt tgatactact aagggtcttc actgaggagg gagcaatcgt tggagaaatt 480
 tcaccattac cttatcttcc aggacatact aatgaggatg tcaaaaatgc agttggggtc 540
 ctcatcgagg ggcttgaatg gaatggtaac acggttcgag gctctgaaaa tctacagaga 600
 ttcgcttgga gaaaccataa tgaggatggg agatcttcac tacctccaga acagaaa 657

<210> 16
 <211> 219
 <212> PRT
 <213> Virut cúm A

<400> 16

Met Asp Ser Asn Thr Val Ser Ser Phe Gln Val Asp Cys Phe Leu Trp
 1 5 10 15

His Ile Arg Lys Arg Phe Ala Asp Asn Gly Leu Gly Asp Ala Pro Phe
 20 25 30

Leu Asp Arg Leu Arg Arg Asp Gln Lys Ser Leu Lys Gly Arg Gly Asn
 35 40 45

Thr Leu Ser Leu Asp Ile Glu Thr Ala Thr Leu Val Gly Lys Gln Ile
 50 55 60

Val Glu Trp Ile Leu Lys Glu Glu Ser Ser Asp Ile Leu Lys Met Thr
 65 70 75 80

Ile Ala Ser Val Pro Thr Ser Arg Tyr Leu Ala Asp Met Thr Leu Glu
 85 90 95

Glu Met Ser Arg Asp Trp Phe Met Leu Met Pro Arg Gln Lys Ile Ile
 100 105 110

Gly Pro Leu Cys Val Arg Val Asp Gln Ala Ile Met Glu Lys Asn Ile
 115 120 125

Ile Leu Lys Ala Asn Phe Ser Val Ile Phe Asn Arg Leu Glu Thr Leu

130 135 140
 Ile Leu Leu Arg Ala Phe Thr Glu Glu Gly Ala Ile Val Gly Glu Ile
 145 150 155 160
 Ser Pro Leu Pro Tyr Leu Pro Gly His Thr Asn Glu Asp Val Lys Asn
 165 170 175
 Ala Val Gly Val Leu Ile Gly Gly Leu Glu Trp Asn Gly Asn Thr Val
 180 185 190
 Arg Gly Ser Glu Asn Leu Gln Arg Phe Ala Trp Arg Asn His Asn Glu
 195 200 205
 Asp Gly Arg Ser Ser Leu Pro Pro Glu Gln Lys
 210 215
 <210> 17
 <211> 1695
 <212> ADN
 <213> Virut cúm A
 <400> 17
 atgaaagtaa aactaatggt tctgttatgt acattttacag ctacatatgc agacacaata 60
 tgtgtaggct accatgccaa caactcaact gacactgttg acacagtact tgagaagaat 120
 gtgacagtga cacactctgt caacctactt gaggacagcc acaatggaaa actatgtcta 180
 ctaaaaggaa tagctccact acaattgggt agttgcagcg ttgccggatg gatcttagga 240
 aaccagagt gcgaattgct gatttccaag gaatcttgggt cctacattgt agaaacacca 300
 aatcctgaga atggaacatg ttaccagggt tatttcacag actatgaaga actgagggag 360
 caattgagtt cagtatcttc atttaagagg ttcgaaatat tccccaaaga gagctcatgg 420
 cccaaccaca ccgtaaccgg agtgtcatca tcatgtctcc ataacgggaa aagcagcttc 480
 tacagaaatt tgctatggct gacggtgaag aacggtctgt acccaaacct gagcaagtcc 540
 tatacaaaca aaaaggagaa agaagtcctt gtactatggg gtgttcatca cccatctaac 600
 ataggggacc aaagggccct ctatcataca gaaaatgctt atgtctctgt agtgtcttca 660
 cattatagca gaagattcac cccagaaata gccaaaagac ccaaggtgag aaatcaggaa 720
 ggaagaatca actactactg gaccctgcta gaaccggggg atacaataat atttgaggca 780
 aatggaaatc taatagcacc aaggtatgcc ttcgaaactga gtaagggttt tggatcagga 840
 atcatcacat caaatgcacc aatgggtgaa tgtaatgcaa agtgtcaaac acctcaggga 900
 gctataaaca gcagtcttcc tttccagaat gtacaccag taacaatagg agagtgccca 960
 aagtatgtca aaagtgcaaa attaaggatg gttacaggac taaggaacac cccatccatt 1020
 caatccagag gtttggttgg agccattgcc ggtttcattg aaggaggggtg gactggaatg 1080
 gtagatgggt ggtatggta tcaccatcag aatgagcaag gatctgggta tgctgcagac 1140
 caacaaagca caaaaatgc cattaatggg attacaaaca aggtgaattc tgtgattgaa 1200
 aaaatgaaca ctcaattcac agctgtgggc aaagaattca acaaactgga aagaagaatg 1260

gaaaaacttaa ataaaaaggt tgatgatggg tttctagaca tttggacata taatgcagaa 1320
 ttgttagttc tactggaaaa tgaaaggact ttggatttcc atgactccaa cgtgaagaat 1380
 ctgtatgaga aagtaaaaag ccaattaaaa aataatgcca aagaaatagg aaacgggtgt 1440
 tttgaattct atcataagtg taacgatgaa tgcattggaga gtgtgaaaaa tggaacttat 1500
 gactatccaa aatattccga agaatacaag ttaaacaggg agaaaattga tggagtgaag 1560
 ttggaatcaa tgggagtcta taatatcctg gcgactctact caacagtcgc cagttcccta 1620
 gttcttttag tctccctggg ggcaatcagc ttctggatgt gttccaatgg gtctttacag 1680
 tgtagaatat gcatac 1695

<210> 18
 <211> 565
 <212> PRT
 <213> Virut cúm A

<400> 18

Met Lys Val Lys Leu Met Val Leu Leu Cys Thr Phe Thr Ala Thr Tyr
1 5 10 15

Ala Asp Thr Ile Cys Val Gly Tyr His Ala Asn Asn Ser Thr Asp Thr
20 25 30

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

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

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

Asn Pro Glu Cys Glu Leu Leu Ile Ser Lys Glu Ser Trp Ser Tyr Ile
85 90 95

Val Glu Thr Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly Tyr Phe
100 105 110

Thr Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe
115 120 125

Lys Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr
130 135 140

Val Thr Gly Val Ser Ser Ser Cys Ser His Asn Gly Lys Ser Ser Phe
145 150 155 160

Tyr Arg Asn Leu Leu Trp Leu Thr Val Lys Asn Gly Leu Tyr Pro Asn
165 170 175

Leu Ser Lys Ser Tyr Thr Asn Lys Lys Glu Lys Glu Val Leu Val Leu

180	185	190
Trp Gly Val His His Pro Ser Asn Ile Gly Asp Gln Arg Ala Leu Tyr		
195	200	205
His Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg		
210	215	220
Arg Phe Thr Pro Glu Ile Ala Lys Arg Pro Lys Val Arg Asn Gln Glu		
225	230	235
Gly Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile		
	245	250
		255
Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Arg Tyr Ala Phe Glu		
	260	265
		270
Leu Ser Lys Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Pro Met		
	275	280
		285
Gly Glu Cys Asn Ala Lys Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser		
290	295	300
Ser Leu Pro Phe Gln Asn Val His Pro Val Thr Ile Gly Glu Cys Pro		
305	310	315
		320
Lys Tyr Val Lys Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn		
	325	330
		335
Thr Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe		
	340	345
		350
Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His		
355	360	365
His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Gln Ser Thr		
370	375	380
Gln Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu		
385	390	395
		400
Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu		
	405	410
		415
Glu Arg Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu		
420	425	430
Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu		
435	440	445
Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys		
450	455	460

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

Phe Glu Phe Tyr His Lys Cys Asn Asp Glu Cys Met Glu Ser Val Lys
485 490 495

Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn
500 505 510

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

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

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

Cys Arg Ile Cys Ile
565

<210> 19
<211> 1407
<212> ADN
<213> Virut cúm A

<400> 19
atgaatccaa atcaaaagat aataacgatt ggctctgttt ctcttactat tgccacaatg 60
tgcttcctta tgcaaattgc catcctggta actaatgtaa cattgcactt caatcaatat 120
gaatgcaact acccccaaaa caaccaagta atactgtgtg aaccaacaat aatagaaaga 180
aacataacag agatagtgtg tctgaccaac accaccatag agaaggaaat atgccccaaa 240
ctagcagaat acagaaattg gtcaaagccg caatgtaaaa ttacaggggtt tgcacctttt 300
tccaaggaca attcgattag gctttctgct ggtggggaca tttgggtgac gagagaacct 360
tatgtgtcat gcgatcctga taagtgttat cagtttgccc ttggacaagg aacaacatta 420
aacaacaggc attcaaatga cacagtacat gataggaccc cttatcgaac cctattgatg 480
aatgagttgg gtattccatt ccatttgggg accaaacaag tgtgcatagc atggtccagc 540
tcaagttgtc atgatggaaa agcatggctt cacgtttgta ttactgggca tgaatrraat 600
gcaactgcta gcttcattta caatgggagg cttgtagata gtattgggtc atggtccaaa 660
aaaatactca ggacacagga gtcggaatgt gtttgcata atggaacttg tacagtagta 720
atgactgatg ggagtgttc aggaatagct gacactaaaa tattattcat tgaagagggg 780
aaaatcggtc atattagccy attgttagga agtgctcagc atgtagagga gtgctcctgt 840
tatccccgat atcctggtgt cagatgcata ttagagagaca actggaaagg ttccaataga 900
cccgctgtag atataaatgt aaaggattat agcattgttt ccagttatgt gtgctcagga 960
cttggtggag atacaccag aaaagacgac agatccagca gtagcgattg tctgaatcct 1020
aacaatgagg aaggggggca tggagtgaat ggctgggcct ttgatgatgg aaatgatgtg 1080

```

tggatgggaa gaacaatcaa cgagacgtta cgctcaggtt atgaaacctt caaagtcatt      1140
gaaggctggt ccaaacctaa ttccaaattg cagataaata ggcaagtcac agttgaaaga      1200
ggtgataggt ccggttatcc tggcattttc tctgttgaag gcaaaagctg tatcaatcgg      1260
tgctttttatg tggagttgat aagaggaagg aaacaggaaa ctgcagtatg gtggacgtca      1320
aacagtattg ttgtgttttg tggcacctca ggtacatatg gaacaggctc atggcctgat      1380
ggggcgaaca tcaatctcat gcctgta                                          1407

```

```

<210> 20
<211> 469
<212> PRT
<213> Virut cúm A

```

```

<220>
<221> misc_feature
<222> (199)..(199)
<223> Xaa có thể là axit amin bất kỳ có trong tự nhiên

```

```

<220>
<221> misc_feature
<222> (267)..(267)
<223> Xaa có thể là axit amin bất kỳ có trong tự nhiên

```

```

<400> 20

```

```

Met Asn Pro Asn Gln Lys Ile Ile Thr Ile Gly Ser Val Ser Leu Thr
1              5              10              15

```

```

Ile Ala Thr Met Cys Phe Leu Met Gln Ile Ala Ile Leu Val Thr Asn
              20              25              30

```

```

Val Thr Leu His Phe Asn Gln Tyr Glu Cys Asn Tyr Pro Pro Asn Asn
              35              40              45

```

```

Gln Val Ile Leu Cys Glu Pro Thr Ile Ile Glu Arg Asn Ile Thr Glu
50              55              60

```

```

Ile Val Tyr Leu Thr Asn Thr Thr Ile Glu Lys Glu Ile Cys Pro Lys
65              70              75              80

```

```

Leu Ala Glu Tyr Arg Asn Trp Ser Lys Pro Gln Cys Lys Ile Thr Gly
              85              90              95

```

```

Phe Ala Pro Phe Ser Lys Asp Asn Ser Ile Arg Leu Ser Ala Gly Gly
              100              105              110

```

```

Asp Ile Trp Val Thr Arg Glu Pro Tyr Val Ser Cys Asp Pro Asp Lys
              115              120              125

```

```

Cys Tyr Gln Phe Ala Leu Gly Gln Gly Thr Thr Leu Asn Asn Arg His
              130              135              140

```

```

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

```

Asn Glu Leu Gly Ile Pro Phe His Leu Gly Thr Lys Gln Val Cys Ile
 165 170 175

Ala Trp Ser Ser Ser Ser Cys His Asp Gly Lys Ala Trp Leu His Val
 180 185 190

Cys Ile Thr Gly His Asp Xaa Asn Ala Thr Ala Ser Phe Ile Tyr Asn
 195 200 205

Gly Arg Leu Val Asp Ser Ile Gly Ser Trp Ser Lys Lys Ile Leu Arg
 210 215 220

Thr Gln Glu Ser Glu Cys Val Cys Ile Asn Gly Thr Cys Thr Val Val
 225 230 235 240

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

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

Gln His Val Glu Glu Cys Ser Cys Tyr Pro Arg Tyr Pro Gly Val Arg
 275 280 285

Cys Ile Cys Arg Asp Asn Trp Lys Gly Ser Asn Arg Pro Val Val Asp
 290 295 300

Ile Asn Val Lys Asp Tyr Ser Ile Val Ser Ser Tyr Val Cys Ser Gly
 305 310 315 320

Leu Val Gly Asp Thr Pro Arg Lys Asp Asp Arg Ser Ser Ser Ser Asp
 325 330 335

Cys Leu Asn Pro Asn Asn Glu Glu Gly Gly His Gly Val Lys Gly Trp
 340 345 350

Ala Phe Asp Asp Gly Asn Asp Val Trp Met Gly Arg Thr Ile Asn Glu
 355 360 365

Thr Leu Arg Ser Gly Tyr Glu Thr Phe Lys Val Ile Glu Gly Trp Ser
 370 375 380

Lys Pro Asn Ser Lys Leu Gln Ile Asn Arg Gln Val Ile Val Glu Arg
 385 390 395 400

Gly Asp Arg Ser Gly Tyr Ser Gly Ile Phe Ser Val Glu Gly Lys Ser
 405 410 415

Cys Ile Asn Arg Cys Phe Tyr Val Glu Leu Ile Arg Gly Arg Lys Gln
 420 425 430

Glu Thr Ala Val Trp Trp Thr Ser Asn Ser Ile Val Val Phe Cys Gly
 435 440 445

Thr Ser Gly Thr Tyr Gly Thr Gly Ser Trp Pro Asp Gly Ala Asn Ile
 450 455 460

Asn Leu Met Pro Val
 465

<210> 21
 <211> 1695
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Gen HA của virus cúm 10-0036-2 (n+5, tương ứng với virus gốc)

<400> 21
 atgaaagtaa aactaatggt tctgttatgt acatttacag ctacatatgc agacacaata 60
 tgtgtaggct accatgcaa caactcaact gacactgttg acacagtact tgagaagaat 120
 gtgacagtga cacactctgt caacctactt gaggacagcc acaatggaaa actatgtcta 180
 ctaaaaggaa tagctccact acaattgggt aattgcagcg ttgccggatg gatcttagga 240
 aaccagagt gcgaattgct gatttccaat gaatcttggc cctacattgt agaaacacca 300
 aatcctgaga atggaacatg ttaccagagg tatttcacag actatgaaga actgagggag 360
 caattgagtt cagtatcttc atttaagagg ttcgaaatat tcccaaaga gagctcatgg 420
 cccaaccaca ccgtaaccgg agtgtcatca tcatgtctcc ataacgggaa aagcagcttc 480
 tacagaaatt tgctatggct gacgggtgaag aacgggtacgt acccaaacct gagcaagtcc 540
 tatacaaaca aaaaggagaa agaagtcctt gtactatggg gtgttcatca cccatctaac 600
 ataggggacc aaagggccct ctatcataca gaaaatgctt atgtctctgt agtgtcttca 660
 cattatagca gaagattcac ccagaaaata gccaaaagac ccaaggtgag aaatcaggaa 720
 ggaagaatca actactactg gaccctgcta gaaccggggg atacaataat atttgaggca 780
 aatggaaatc taatagcacc aagggtatgcc ttcgaaactga gtaagggttt tggatcagga 840
 atcatcacat caaatgcaac aatgggtgaa tgtaatgcaa cgtgtcaaac acctcaggga 900
 gctataaaca gcagtcttcc tttccagaat gtacacccag taacaatagg agagtgccca 960
 aagtatgtca aaagtgcaaa attaaggatg gttacaggac taaggaaacac cccatccatt 1020
 caatccagag gtttgtttgg agccattgcc ggtttcattg aaggaggggtg gactggaatg 1080
 gtagatgggt ggtatgggtt tcaccatcag aatgagcaag gatctgggta tgctgcagac 1140
 caacaaagca caaaaaatgc cattaatggg attacaaaca aggtgaattc tgtgattgaa 1200
 aaaatgaaca ctcaattcac agctgtgggc aaagaattca acaaactgga aagaagaatg 1260
 gaaaacttaa ataaaaaggt tgatgatggg tttctagaca tttggacata taatgcagaa 1320
 ttgttagttc tactggaaaa tgaaaggact ttggatttcc atgactccaa cgtgaagaat 1380
 ctgtatgaga aagtaaaaag ccaattaaaa aataatgcca aagaaatagg aaacgggtgt 1440
 tttgaattct atcataagtg taacgatgaa tgcattggaga gtgtgaaaaa tggaacttat 1500

gactatccaa aatattccga agaatacaaag ttaaacaggg agaaaattga tggagtgaaa 1560
 ttggaatcaa tgggagtcta taatatcctg gcgatctact caacagtcgc cagttcccta 1620
 gttcttttag tctccctggg ggcaatcagc ttctggatgt gttccaatgg gtctttacag 1680
 tgtagaatat gcatc 1695

<210> 22

<211> 565

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> HA dự đoán AA của virus cúm 10-0036-2 (n+5, tương ứng với virus gốc)

<400> 22

Met Lys Val Lys Leu Met Val Leu Leu Cys Thr Phe Thr Ala Thr Tyr
 1 5 10 15

Ala Asp Thr Ile Cys Val Gly Tyr His Ala Asn Asn Ser Thr Asp Thr
 20 25 30

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

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

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

Asn Pro Glu Cys Glu Leu Leu Ile Ser Asn Glu Ser Trp Ser Tyr Ile
 85 90 95

Val Glu Thr Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly Tyr Phe
 100 105 110

Thr Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe
 115 120 125

Lys Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr
 130 135 140

Val Thr Gly Val Ser Ser Ser Cys Ser His Asn Gly Lys Ser Ser Phe
 145 150 155 160

Tyr Arg Asn Leu Leu Trp Leu Thr Val Lys Asn Gly Thr Tyr Pro Asn
 165 170 175

Leu Ser Lys Ser Tyr Thr Asn Lys Lys Glu Lys Glu Val Leu Val Leu
 180 185 190

Trp Gly Val His His Pro Ser Asn Ile Gly Asp Gln Arg Ala Leu Tyr
 195 200 205

His Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg
 210 215 220
 Arg Phe Thr Pro Glu Ile Ala Lys Arg Pro Lys Val Arg Asn Gln Glu
 225 230 235 240
 Gly Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile
 245 250 255
 Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Arg Tyr Ala Phe Glu
 260 265 270
 Leu Ser Lys Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Thr Met
 275 280 285
 Gly Glu Cys Asn Ala Thr Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser
 290 295 300
 Ser Leu Pro Phe Gln Asn Val His Pro Val Thr Ile Gly Glu Cys Pro
 305 310 315 320
 Lys Tyr Val Lys Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn
 325 330 335
 Thr Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
 340 345 350
 Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His
 355 360 365
 His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Gln Ser Thr
 370 375 380
 Gln Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu
 385 390 395 400
 Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu
 405 410 415
 Glu Arg Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu
 420 425 430
 Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu
 435 440 445
 Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys
 450 455 460
 Val Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys
 465 470 475 480

Phe Glu Phe Tyr His Lys Cys Asn Asp Glu Cys Met Glu Ser Val Lys
485 490 495

Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn
500 505 510

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

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

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

Cys Arg Ile Cys Ile
565

<210> 23
<211> 1695
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Gen HA của virus cúm 10-0036-2 (n+4, tương ứng với virus gốc)

<400> 23
atgaaagtaa aactaatggt tctgttatgt acatttacag ctacatatgc agacacaata 60
tgtgtaggct accatgccaa caactcaact gacactgttg acacagtact tgagaagaat 120
gtgacagtga cacactctgt caacctactt gaggacagcc acaatggaaa actatgtcta 180
ctaaaaggaa tagctccact acaattgggt aattgcagcg ttgccggatg gatcttagga 240
aaccagagt gcgaattgct gatttccaat gaatcttggc cctacattgt agaaacacca 300
aatcctgaga atggaacatg ttaccagggt tatttcacag actatgaaga actgaggagg 360
caattgagtt cagtatcttc atttaagagg ttcgaaatat tccccaaaga gagctcatgg 420
cccaaccaca ccgtaaccgg agtgtcatca tcatgtctcc ataacgggaa aagcagcttc 480
tacagaaatt tgctatggct gacggtgaag aacggtacgt acccaaacct gagcaagtcc 540
tatacaaaca aaaaggagaa agaagtcctt gtactatggg gtgttcatca cccatctaac 600
ataggggacc aaagggccct ctatcataca gaaaatgctt atgtctctgt agtgtcttca 660
cattatagca gaagattcac ccagaaata gccaaaagac ccaagggtgag aaatcaggaa 720
ggaagaatca actactactg gaccctgcta gaaccggggg atacaataat atttgaggca 780
aatggaaatc taatagcacc aagggtatgcc ttcgaaactga gtaagggttt tggatcagga 840
atcatcacat caaatgcaac aatgggtgaa tgtaatgcaa agtgtcaaac acctcaggga 900
gctataaaca gcagtcttcc tttccagaat gtacaccag taacaatagg agagtgccca 960
aagtatgtca aaagtgcaaa attaaggatg gttacaggac taaggaacac cccatccatt 1020
caatccagag gtttggttg agccattgcc ggtttcattg aaggagggtg gactggaatg 1080
gtagatggtt ggtatggtta tcaccatcag aatgagcaag gatctgggta tgctgcagac 1140

```

caacaaagca cacaaaatgc cattaatggg attacaaaca aggtgaattc tgtgattgaa      1200
aaaatgaaca ctcaattcac agctgtgggc aaagaattca acaaactgga aagaagaatg      1260
gaaaacttaa ataaaaaggt tgatgatggg tttctagaca tttggacata taatgcagaa      1320
ttgttagttc tactggaaaa tgaaaggact ttggatttcc atgactccaa cgtgaagaat      1380
ctgtatgaga aagtaaaaaag ccaattaaaa aataatgcc aagaaatagg aaacgggtgt      1440
tttgaattct atcataagtg taacgatgaa tgcattggaga gtgtgaaaaa tggaacttat      1500
gactatccaa aatattccga agaatcaaag ttaaacaggg agaaaattga tggagtgaaa      1560
ttggaatcaa tgggagtcta taatatcctg gcgatctact caacagtcgc cagttcccta      1620
gttcttttag tctccctggg ggcaatcagc ttctggatgt gttccaatgg gtctttacag      1680
tgtagaatat gcattc                                     1695

```

<210> 24

<211> 565

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> HA dự đoán AA của virus cúm 10-0036-2 (n+4, tương ứng với virus gốc)

<400> 24

```

Met Lys Val Lys Leu Met Val Leu Leu Cys Thr Phe Thr Ala Thr Tyr
1           5           10           15

```

```

Ala Asp Thr Ile Cys Val Gly Tyr His Ala Asn Asn Ser Thr Asp Thr
          20           25           30

```

```

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

```

```

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

```

```

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

```

```

Asn Pro Glu Cys Glu Leu Leu Ile Ser Asn Glu Ser Trp Ser Tyr Ile
85           90           95

```

```

Val Glu Thr Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly Tyr Phe
100          105          110

```

```

Thr Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe
115          120          125

```

```

Lys Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr
130          135          140

```

```

Val Thr Gly Val Ser Ser Ser Cys Ser His Asn Gly Lys Ser Ser Phe
145          150          155          160

```

Tyr Arg Asn Leu Leu Trp Leu Thr Val Lys Asn Gly Thr Tyr Pro Asn
 165 170 175

Leu Ser Lys Ser Tyr Thr Asn Lys Lys Glu Lys Glu Val Leu Val Leu
 180 185 190

Trp Gly Val His His Pro Ser Asn Ile Gly Asp Gln Arg Ala Leu Tyr
 195 200 205

His Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg
 210 215 220

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

Gly Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile
 245 250 255

Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Arg Tyr Ala Phe Glu
 260 265 270

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

Gly Glu Cys Asn Ala Lys Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser
 290 295 300

Ser Leu Pro Phe Gln Asn Val His Pro Val Thr Ile Gly Glu Cys Pro
 305 310 315 320

Lys Tyr Val Lys Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn
 325 330 335

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

Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His
 355 360 365

His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Gln Ser Thr
 370 375 380

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

Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu
 405 410 415

Glu Arg Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu
 420 425 430

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

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

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

Phe Glu Phe Tyr His Lys Cys Asn Asp Glu Cys Met Glu Ser Val Lys
485 490 495

Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn
500 505 510

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

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

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

Cys Arg Ile Cys Ile
565

<210> 25

<211> 1695

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Gen HA của virus cúm 10-0036-2 (n+3, tương ứng với virus gốc)

<400> 25

atgaaagtaa aactaatggt tctgttatgt acatttacag ctacatatgc agacacaata	60
tgtgtaggct accatgcaa caactcaact gacactgttg acacagtact tgagaagaat	120
gtgacagtga cacactctgt caacctactt gaggacagcc acaatggaaa actatgtcta	180
ctaaaaggaa tagctccact acaattgggt aattgcagcg ttgccgatg gatcttagga	240
aaccagagt gcgaattgct gatttccaat gaatcttggt cctacattgt agaaacacca	300
aatcctgaga atggaacatg ttaccagggt tatttcacag actatgaaga actgagggag	360
caattgagtt cagtatcttc atttaagagg ttcgaaatat tcccaaaga gagctcatgg	420
cccaaccaca ccgtaaccgg agtgtcatca tcatgtcccc ataacgggaa aagcagcttc	480
tacagaaatt tgctatggct gacggtgaag aacggtacgt acccaaacct gagcaagtcc	540
tatacaaaca aaaaggagaa agaagtcctt gtactatggg gtgttcatca cccatctaac	600
ataggggacc aaagggccct ctatcataca gaaaatgctt atgtctctgt agtgtcttca	660
cattatagca gaagattcac ccagaaata gccaaaagac ccaaggtgag aaatcaggaa	720
ggaagaatca actactactg gaccctgcta gaaccgggg atacaataat atttgaggca	780

```

aatggaaatc taatagcacc aaggtatgcc ttogaactga gtaagggttt tggatcagga      840
atcatcacat caaatgcacc aatgggtgaa tgtaatgcaa agtgtcaaac acctcaggg      900
gctataaaca gcagtcttcc tttccagaat gtacaccag taacaatagg agagtgccca      960
aagtatgtca aaagtgcaaa attaaggatg gttacaggac taaggaacac cccatccatt     1020
caatccagag gtttgtttgg agccattgcc ggtttcattg aaggagggtg gactggaatg     1080
gtagatgggtt ggtatgggta tcaccatcag aatgagcaag gatctgggta tgctgcagac     1140
caacaaagca cacaaaatgc cattaatggg attacaaaca aggtgaattc tgtgattgaa     1200
aaaatgaaca ctcaattcac agctgtgggc aaagaattca acaaactgga aagaagaatg     1260
gaaaacttaa ataaaaaggt tgatgatggg tttctagaca tttggacata taatgcagaa     1320
ttgttagttc tactggaaaa tgaaaggact ttggatttcc atgactccaa cgtgaagaat     1380
ctgtatgaga aagtaaaaag ccaattaaaa aataatgcca aagaaatagg aaacgggtgt     1440
ttgaattct atcataagtg taacgatgaa tgcattggaga gtgtgaaaaa tggaacttat     1500
gactatccaa aatattccga agaatcaaag ttaaacaggg agaaaattga tggagtga    1560
ttggaatcaa tgggagtcta taatatcctg gcgatctact caacagtcgc cagttcccta     1620
gttcttttag tctccctggg ggcaatcagc ttctggatgt gttccaatgg gtctttacag     1680
tgtagaatat gcac                                     1695

```

<210> 26
 <211> 565
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> HA dự đoán AA của virus cúm 10-0036-2 (n+3, tương ứng với virus gốc)
 <400> 26

```

Met Lys Val Lys Leu Met Val Leu Leu Cys Thr Phe Thr Ala Thr Tyr
1           5           10           15

```

```

Ala Asp Thr Ile Cys Val Gly Tyr His Ala Asn Asn Ser Thr Asp Thr
           20           25           30

```

```

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

```

```

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

```

```

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

```

```

Asn Pro Glu Cys Glu Leu Leu Ile Ser Asn Glu Ser Trp Ser Tyr Ile
           85           90           95

```

```

Val Glu Thr Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly Tyr Phe
           100          105          110

```

Thr Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe
 115 120 125

Lys Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr
 130 135 140

Val Thr Gly Val Ser Ser Ser Cys Ser His Asn Gly Lys Ser Ser Phe
 145 150 155 160

Tyr Arg Asn Leu Leu Trp Leu Thr Val Lys Asn Gly Thr Tyr Pro Asn
 165 170 175

Leu Ser Lys Ser Tyr Thr Asn Lys Lys Glu Lys Glu Val Leu Val Leu
 180 185 190

Trp Gly Val His His Pro Ser Asn Ile Gly Asp Gln Arg Ala Leu Tyr
 195 200 205

His Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg
 210 215 220

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

Gly Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile
 245 250 255

Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Arg Tyr Ala Phe Glu
 260 265 270

Leu Ser Lys Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Pro Met
 275 280 285

Gly Glu Cys Asn Ala Lys Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser
 290 295 300

Ser Leu Pro Phe Gln Asn Val His Pro Val Thr Ile Gly Glu Cys Pro
 305 310 315 320

Lys Tyr Val Lys Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn
 325 330 335

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

Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His
 355 360 365

His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Gln Ser Thr
 370 375 380

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

Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu
405 410 415

Glu Arg Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu
420 425 430

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

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

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

Phe Glu Phe Tyr His Lys Cys Asn Asp Glu Cys Met Glu Ser Val Lys
485 490 495

Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn
500 505 510

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

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

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

Cys Arg Ile Cys Ile
565

<210> 27
<211> 1695
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Gen HA của virus cúm 10-0036-2 (n+2, tương ứng với virus gốc)

<400> 27
atgaaagtaa aactaatggt tctgttatgt acatttacag ctacatatgc agacacaata 60
tgtgtaggct accatgcca caactcaact gacactgttg acacagtact tgagaagaat 120
gtgacagtga cacactctgt caacctactt gaggacagcc acaatggaaa actatgtcta 180
ctaaaaggaa tagctccact acaattgggt aattgcagcg ttgccggatg gatcttagga 240
aaccagagt gcgaattgct gatttccaat gaatcttggt cctacattgt agaaacacca 300
aatcctgaga atggaacatg ttaccagggt tatttcacag actatgaaga actgaggagg 360
caattgagtt cagtatcttc atttaagagg ttcgaaatat tcccaaaga gagctcatgg 420


```

cccaaccaca ccgtaaccgg agtgtcatca tcatgtctcc ataacgggaa aagcagcttc      480
tacagaaatt tgctatggct gacggtgaag aacggtctgt acccaaacct gagcaagtcc      540
tatacaaaca aaaaggagaa agaagtcctt gtactatggg gtgttcatca cccatctaac      600
ataggggacc aaagggccct ctatcataca gaaaatgctt atgtctctgt agtgtcttca      660
cattatagca gaagattcac ccagaaata gccaaaagac ccaaggtgag aaatcaggaa      720
ggaagaatca actactactg gaccctgcta gaaccggggg atacaataat atttgaggca      780
aatggaaatc taatagcacc aaggtatgcc ttcgaactga gtaaggggtt tggatcagga      840
atcatcacat caaatgcacc aatgggtgaa tgtaatgcaa agtgtcaaac acctcagga      900
gctataaaca gcagtcttcc tttccagaat gtacaccag taacaatagg agagtgccca      960
aagtatgtca aaagtgcaaa attaaggatg gttacaggac taaggaacac cccatccatt     1020
caatccagag gtttgtttgg agccattgcc ggtttcattg aaggaggggtg gactggaatg     1080
gtagatgggt ggtatgggta tcaccatcag aatgagcaag gatctgggta tgctgcagac     1140
caacaaagca cacaaaatgc cattaatggg attacaaaca aggtgaattc tgtgattgaa     1200
aaaatgaaca ctcaattcac agctgtgggc aaagaattca acaaactgga aagaagaatg     1260
gaaaacttaa ataaaaaggt tgatgatggg tttctagaca tttggacata taatgcagaa     1320
ttgttagttc tactggaaaa tgaaaggact ttggatttcc atgactccaa cgtgaagaat     1380
ctgtatgaga aagtaaaaag ccaattaaaa aataatgcc aagaaatagg aaacgggtgt     1440
tttgaattct atcataagtg taacgatgaa tgcatggaga gtgtgaaaaa tggaaacttat     1500
gactatccaa aatattccga agaatcaaag ttaaacaggg agaaaattga tggagtga aa     1560
ttggaatcaa tgggagtcta taatatcctg gcatctact caacagtcgc cagttcccta     1620
gttcttttag tctccctggg ggcaatcagc ttctggatgt gttccaatgg gtctttacag     1680
tgtagaatat gcac                                     1695

```

<210> 28

<211> 565

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> HA dự đoán AA của virus cúm 10-0036-2 (n+2, tương ứng với virus gốc)

<400> 28

```

Met Lys Val Lys Leu Met Val Leu Leu Cys Thr Phe Thr Ala Thr Tyr
1           5           10          15

```

```

Ala Asp Thr Ile Cys Val Gly Tyr His Ala Asn Asn Ser Thr Asp Thr
          20          25          30

```

```

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

```

```

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

```

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

Asn Pro Glu Cys Glu Leu Leu Ile Ser Asn Glu Ser Trp Ser Tyr Ile
85 90 95

Val Glu Thr Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly Tyr Phe
100 105 110

Thr Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe
115 120 125

Lys Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr
130 135 140

Val Thr Gly Val Ser Ser Ser Cys Ser His Asn Gly Lys Ser Ser Phe
145 150 155 160

Tyr Arg Asn Leu Leu Trp Leu Thr Val Lys Asn Gly Leu Tyr Pro Asn
165 170 175

Leu Ser Lys Ser Tyr Thr Asn Lys Lys Glu Lys Glu Val Leu Val Leu
180 185 190

Trp Gly Val His His Pro Ser Asn Ile Gly Asp Gln Arg Ala Leu Tyr
195 200 205

His Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg
210 215 220

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

Gly Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile
245 250 255

Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Arg Tyr Ala Phe Glu
260 265 270

Leu Ser Lys Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Pro Met
275 280 285

Gly Glu Cys Asn Ala Lys Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser
290 295 300

Ser Leu Pro Phe Gln Asn Val His Pro Val Thr Ile Gly Glu Cys Pro
305 310 315 320

Lys Tyr Val Lys Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn
325 330 335

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

Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His
355 360 365

His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Gln Ser Thr
370 375 380

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

Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu
405 410 415

Glu Arg Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu
420 425 430

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

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

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

Phe Glu Phe Tyr His Lys Cys Asn Asp Glu Cys Met Glu Ser Val Lys
485 490 495

Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn
500 505 510

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

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

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

Cys Arg Ile Cys Ile
565

<210> 29

<211> 1695

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Gen HA của virus cúm 10-0036-2 (n+1, tương ứng với virus gốc)

<400> 29

atgaaagtaa aactaatggt tctgttatgt acatttacag ctacatatgc agacacaata

60

```

tgtgtaggct accatgccaa caactcaact gacactggtg acacagtact tgagaagaat      120
gtgacagtga cacactctgt caacctactt gaggacagcc acaatggaaa actatgtcta      180
ctaaaaggaa tagctccact acaattgggt agttgcagcg ttgccggatg gatcttagga      240
aaccagagt gcgaattgct gatttccaat gaatcttggc cctacattgt agaaacacca      300
aatcctgaga atggaacatg ttaccagggt tatttcacag actatgaaga actgagggag      360
caattgagtt cagtatcttc atttaagagg ttcgaaatat tccccaaaga gagctcatgg      420
cccaaccaca ccgtaaccgg agtgtcatca tcatgtctcc ataacgggaa aagcagcttc      480
tacagaaatt tgctatggct gacggtgaag aacggtctgt acccaaacct gagcaagtcc      540
tatacaaaca aaaaggagaa agaagtcctt gtactatggg gtgttcatca cccatctaac      600
atagggggacc aaagggccct ctatcataca gaaaatgctt atgtctctgt agtgtcttca      660
cattatagca gaagattcac cccagaaata gccaaaagac ccaagggtgag aaatcaggaa      720
ggaagaatca actactactg gaccctgcta gaaccggggg atacaataat atttgaggca      780
aatggaaaac taatagcacc aaggtatgcc ttcgaaactg gtaagggttt tggatcagga      840
atcatcacat caaatgcacc aatgggtgaa tgtaatgcaa agtgtcaaac acctcaggga      900
gctataaaca gcagtcttcc tttccagaat gtacaccagc taacaatagg agagtgccca      960
aagtatgtca aaagtgcaaa attaaggatg gttacaggac taaggaacac cccatccatt     1020
caatccagag gtttggttgg agccattgcc ggtttcattg aaggaggggtg gactggaatg     1080
gtagatgggt ggtatggtta tcaccatcag aatgagcaag gatctgggta tgctgcagac     1140
caacaaagca cacaaaatgc cattaatggg attacaaaca aggtgaattc tgtgattgaa     1200
aaaatgaaca ctcaattcac agctgtgggc aaagaattca acaaactgga aagaagaatg     1260
gaaaacttaa ataaaaaggt tgatgatggg tttctagaca tttggacata taatgcagaa     1320
ttgttagttc tactggaaaa tgaaaggact ttggatttcc atgactccaa cgtgaagaat     1380
ctgtatgaga aagtaaaaag ccaattaaaa aataatgcc aagaaatagg aaacgggtgt     1440
tttgaattct atcataagtg taacgatgaa tgcatggaga gtgtgaaaaa tggaaacttat     1500
gactatccaa aatattccga agaatcaaag ttaaacaggg agaaaattga tggagtgaaa     1560
ttggaatcaa tgggagtcta taatatcctg gcgatctact caacagtcgc cagttcccta     1620
gttcttttag tctccctggg ggcaatcagc ttctggatgt gttccaatgg gtctttacag     1680
tgtagaatat gcac                                     1695

```

<210> 30

<211> 565

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> HA dự đoán AA của virus cúm 10-0036-2 (n+1, tương ứng với virus gốc)

<400> 30

```

Met Lys Val Lys Leu Met Val Leu Leu Cys Thr Phe Thr Ala Thr Tyr
1          5          10          15

```

Ala Asp Thr Ile Cys Val Gly Tyr His Ala Asn Asn Ser Thr Asp Thr
20 25 30

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

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

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

Asn Pro Glu Cys Glu Leu Leu Ile Ser Lys Glu Ser Trp Ser Tyr Ile
85 90 95

Val Glu Thr Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly Tyr Phe
100 105 110

Thr Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe
115 120 125

Lys Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr
130 135 140

Val Thr Gly Val Ser Ser Ser Cys Ser His Asn Gly Lys Ser Ser Phe
145 150 155 160

Tyr Arg Asn Leu Leu Trp Leu Thr Val Lys Asn Gly Leu Tyr Pro Asn
165 170 175

Leu Ser Lys Ser Tyr Thr Asn Lys Lys Glu Lys Glu Val Leu Val Leu
180 185 190

Trp Gly Val His His Pro Ser Asn Ile Gly Asp Gln Arg Ala Leu Tyr
195 200 205

His Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg
210 215 220

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

Gly Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile
245 250 255

Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Arg Tyr Ala Phe Glu
260 265 270

Leu Ser Lys Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Pro Met
275 280 285

Gly Glu Cys Asn Ala Lys Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser
 290 295 300

Ser Leu Pro Phe Gln Asn Val His Pro Val Thr Ile Gly Glu Cys Pro
 305 310 315 320

Lys Tyr Val Lys Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn
 325 330 335

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

Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His
 355 360 365

His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Gln Ser Thr
 370 375 380

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

Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu
 405 410 415

Glu Arg Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu
 420 425 430

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

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

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

Phe Glu Phe Tyr His Lys Cys Asn Asp Glu Cys Met Glu Ser Val Lys
 485 490 495

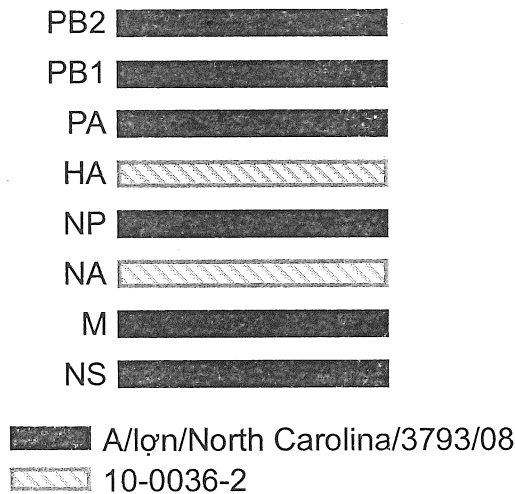
Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn
 500 505 510

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

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

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

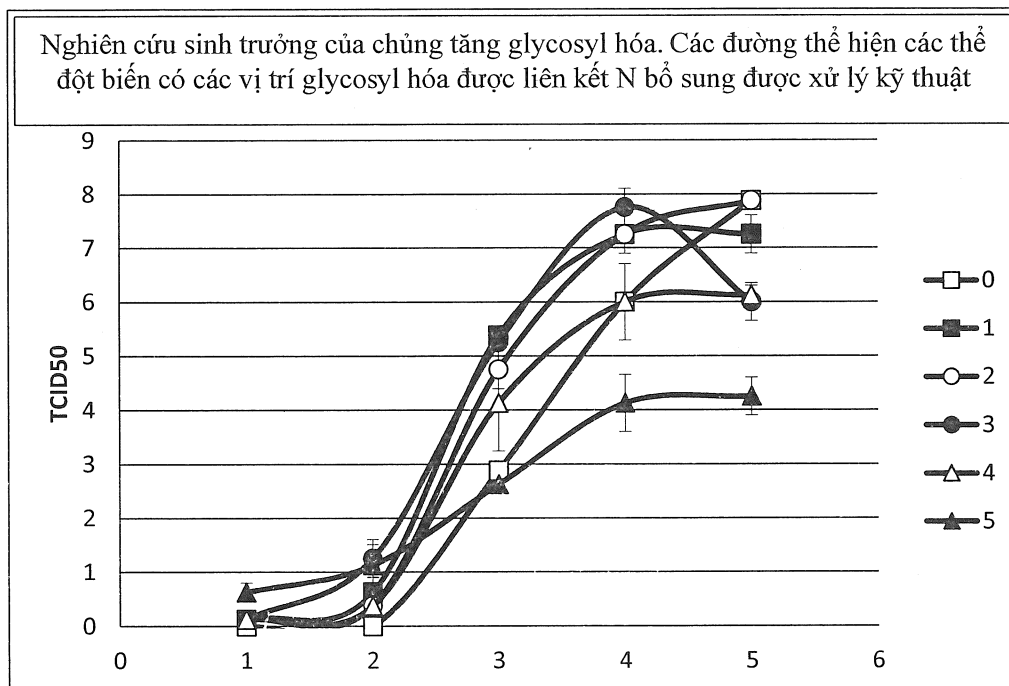
Cys Arg Ile Cys Ile
 565

**FIG. 1**

Seq 18 = HA (gốc); Seq 22 = HA (n+5 thế đột biến)

SEQ 18	MKVKLMVLLCTFTATYADTICVGYHANNSTDTVDTVLEKNVTVTHSVNLLLED SHNGKLCL	60
SEQ 22	MKVKLMVLLCTFTATYADTICVGYHANNSTDTVDTVLEKNVTVTHSVNLLLED SHNGKLCL	60
SEQ 18	LKGIAPLQLGSCSVAGWILGNPECELLISKESWSYIVETPNPENGTCPGYFTDYEELRE	120
SEQ 22	LKGIAPLQLGNCSCSVAGWILGNPECELLISNESWSYIVETPNPENGTCPGYFTDYEELRE	120
SEQ 18	QLSSVSSFKRFEIFPKESSWPNHVTGTGVSSSCSHNGKSSFYRNLLWLTVKNGLYPNLSKS	180
SEQ 22	QLSSVSSFKRFEIFPKESSWPNHVTGTGVSSSCSHNGKSSFYRNLLWLTVKNGLYPNLSKS	180
SEQ 18	YTNKKEKEVLVLWGVHHPNSNIGDQRALYHTENAYVSVVSSHYSRRFTPEIAKRPKVRNQE	240
SEQ 22	YTNKKEKEVLVLWGVHHPNSNIGDQRALYHTENAYVSVVSSHYSRRFTPEIAKRPKVRNQE	240
SEQ 18	GRINYYWTLLEPGDTIIFEANGNLIAPRYAFELSKGFGSGIITSNAPMGECAKQCQTPQG	300
SEQ 22	GRINYYWTLLEPGDTIIFEANGNLIAPRYAFELSKGFGSGIITSNAPMGECAKQCQTPQG	300
SEQ 18	AINSSLPFQNVHPVTIGECPKYVKS AKLRMVTGLRNTPSIQSRGLFGAIA GFIEGGWTGM	360
SEQ 22	AINSSLPFQNVHPVTIGECPKYVKS AKLRMVTGLRNTPSIQSRGLFGAIA GFIEGGWTGM	360
SEQ 18	VDGWYGYHHQNEQSGYAADQQSTQNAINGITNKVNSVIEKMNTQFTAVGKEFNKLERRM	420
SEQ 22	VDGWYGYHHQNEQSGYAADQQSTQNAINGITNKVNSVIEKMNTQFTAVGKEFNKLERRM	420
SEQ 18	ENLNKKVDDGFLDIWTYNAELLVLL ENERTLDFHDSNVKNLYEKVKS QLKNNAKEIGNGC	480
SEQ 22	ENLNKKVDDGFLDIWTYNAELLVLL ENERTLDFHDSNVKNLYEKVKS QLKNNAKEIGNGC	480
SEQ 18	FEFYHKCNDECMESVKNGTYDYPKYSEESKLNREKIDGVKLESMGVYNILAIYSTVASSL	540
SEQ 22	FEFYHKCNDECMESVKNGTYDYPKYSEESKLNREKIDGVKLESMGVYNILAIYSTVASSL	540
SEQ 18	VLLVSLGAISFWMC SNGSLQCRICI	565
SEQ 22	VLLVSLGAISFWMC SNGSLQCRICI	565

FIG. 2



	Virut thu được	Vị trí glycosyl hóa bổ sung (n)
0	Virut gốc RG (10-0036-2)	0
1	S71N	1
2	S71N, K90N	2
3	S71N, K90N, L173T	3
4	S71N, K90N, L173T, P287T	4
5	S71N, K90N, L173T, P287T, K294T	5

FIG. 3

SEQ ID NO	Loại	Mô tả
1	ADN	Gen PB2 của virus cúm A/lợn/NorthCarolina/3793/08(H1N1) - 10-0036-2
2	PRT	PB2 dự đoán AA của virus cúm A/lợn/NorthCarolina/3793/08(H1N1) - 10-0036-2
3	ADN	Gen PB1 của virus cúm A/lợn/NorthCarolina/3793/08(H1N1) - 10-0036-2
4	PRT	PB1 dự đoán AA của virus cúm A/lợn/NorthCarolina/3793/08(H1N1) - 10-0036-2
5	ADN	Gen PA của virus cúm A/lợn/NorthCarolina/3793/08(H1N1) - 10-0036-2
6	PRT	PA dự đoán AA của virus cúm A/lợn/NorthCarolina/3793/08(H1N1) - 10-0036-2
7	ADN	Gen HA của virus cúm A/lợn/NorthCarolina/3793/08(H1N1) - 10-0036-2
8	PRT	HA dự đoán AA của virus cúm A/lợn/NorthCarolina/3793/08(H1N1) - 10-0036-2
9	ADN	Gen NP của virus cúm A/lợn/NorthCarolina/3793/08(H1N1) - 10-0036-2
10	PRT	NP dự đoán AA của virus cúm A/lợn/NorthCarolina/3793/08(H1N1) - 10-0036-2
11	ADN	Gen NA của virus cúm A/lợn/NorthCarolina/3793/08(H1N1) - 10-0036-2
12	PRT	NA dự đoán AA của virus cúm A/lợn/NorthCarolina/3793/08(H1N1) - 10-0036-2
13	ADN	Gen M của virus cúm A/lợn/NorthCarolina/3793/08(H1N1) - 10-0036-2
14	PRT	M dự đoán AA của virus cúm A/lợn/NorthCarolina/3793/08(H1N1) - 10-0036-2
15	ADN	Gen NS của virus cúm A/lợn/NorthCarolina/3793/08(H1N1) - 10-0036-2
16	PRT	NS dự đoán AA của virus cúm A/lợn/NorthCarolina/3793/08(H1N1) - 10-0036-2
17	ADN	Gen HA của virus cúm 10-0036-2 (n+0, gốc) - RG 10-0036-2
18	PRT	HA dự đoán AA của virus cúm 10-0036-2 (n+0, = gốc) - RG 10-0036-2
19	ADN	Gen NA của virus cúm 10-0036-2 (n+0, gốc) - RG 10-0036-2
20	PRT	NA dự đoán AA của virus cúm 10-0036-2 (n+0, = gốc) - RG 10-0036-2
21	ADN	Gen HA của virus cúm 10-0036-2 (n+5, tương ứng với RG 10-0036-2)
22	PRT	HA dự đoán AA của virus cúm 10-0036-2 (n+5, tương ứng với RG 10-0036-2)
23	ADN	Gen HA của virus cúm 10-0036-2 (n+4, tương ứng với RG 10-0036-2)
24	PRT	HA dự đoán AA của virus cúm 10-0036-2 (n+4, tương ứng với RG 10-0036-2)
25	ADN	Gen HA của virus cúm 10-0036-2 (n+3, tương ứng với RG 10-0036-2)
26	PRT	HA dự đoán AA của virus cúm 10-0036-2 (n+3, tương ứng với RG 10-0036-2)
27	ADN	Gen HA của virus cúm 10-0036-2 (n+2, tương ứng với RG 10-0036-2)
28	PRT	HA dự đoán AA của virus cúm 10-0036-2 (n+2, tương ứng với RG 10-0036-2)
29	ADN	Gen HA của virus cúm 10-0036-2 (n+1, tương ứng với RG 10-0036-2)
30	PRT	HA dự đoán AA của virus cúm 10-0036-2 (n+1, tương ứng với RG 10-0036-2)

FIG. 4