



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

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

(11)



1-0034374

(51)⁷ A61K 39/12

(13) B

(21) 1-2015-03806

(22) 12/03/2014

(86) PCT/US2014/024603 12/03/2014

(87) WO2014/150939 25/09/2014

(30) 61/800,204 15/03/2013 US

(45) 26/12/2022 417

(43) 25/01/2016 334A

(73) 1. TAKEDA VACCINES, INC. (US)

75 Sidney Street, Cambridge, MA 02139, United States of America

2. THE GOVERNMENT OF THE UNITED STATES OF AMERICA as represented by THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)

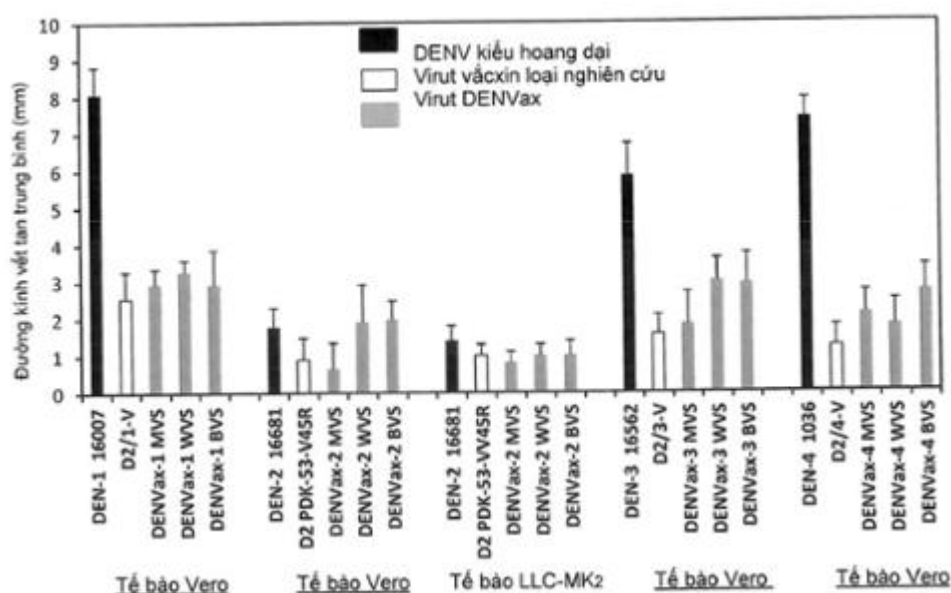
4770 Buford Highway, Mailstop K-79, Atlanta, GA 30341, United States of America

(72) STINCHCOMB, Dan, T. (US); KINNEY, Claire (US); KINNEY, Richard, M. (US); LIVENGOOD, Jill, A. (US).

(74) Công ty Luật TNHH T&G (TGVN)

(54) PHÂN TỬ POLYNUCLEOTIT MÃ HÓA PHÂN TỬ POLYPEPTIT CHỨNG VIRUT DENGUE-2 SỐNG ĐƯỢC CẢI BIẾN GIẢM ĐỘC LỰC PDK-53-V, PHÂN TỬ POLYPEPTIT ĐƯỢC MÃ HÓA BỞI POLYNUCLEOTIT NÀY, ĐƯỢC PHẨM, VẬT TRUYỀN VÀ TẾ BÀO CHỨA POLYNUCLEOTIT NÀY

(57) Sáng chế đề cập đến thể khảm axit nucleic, polypeptit được mã hóa bởi thể khảm axit nucleic này, được phẩm chứa chúng và vật truyền mã hóa chúng. Sáng chế cũng đề cập đến các axit nucleic hoặc polypeptit được mã hóa có liên quan, và chế phẩm gây miễn dịch chứa các axit nucleic hoặc polypeptit được mã hóa này. Chế phẩm gây miễn dịch này có khả năng gây ra đáp ứng miễn dịch ở đối tượng đối với tất cả các kiểu huyết thanh của virus dengue. Sáng chế cũng đề cập đến kit và virus dengue sống giảm độc lực có liên quan.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến dược phẩm, và quy trình sản xuất các cấu trúc virus dengue và dược phẩm vắc xin chứa chúng. Sáng chế đề cập đến dược phẩm mà bao gồm, nhưng không chỉ giới hạn ở, các cấu trúc virus flavivirus khả năng mà đơn lẻ hoặc trong tổ hợp với các cấu trúc khác có thể được sử dụng trong dược phẩm vắc xin. Theo một số phương án, dược phẩm có thể bao gồm các cấu trúc của nhiều hơn một kiểu huyết thanh của virus dengue, như virus dengue-1 (DEN-1), virus dengue-2 (DEN-2), virus dengue-3 (DEN-3) và/hoặc virus dengue-4 (DEN-4). Theo các phương án khác, chiến lược sản xuất mà có thể cải thiện tính an toàn và độ ổn định di truyền của virus vắc xin dengue khả năng sống giảm độc lực tái tổ hợp (DENVax). Một số phương án bao gồm ít nhất một virus dengue sống giảm độc lực trong tổ hợp với các cấu trúc khả năng virus dengue được nhận dạng để có sự an toàn và hiệu quả trong dược phẩm vắc xin mà ở đó các cấu trúc đã trải qua việc chuyển tiếp bổ sung vào trong môi trường nuôi cấy tế bào.

Tình trạng kỹ thuật của sáng chế

Sự lây nhiễm với virus dengue có thể dẫn đến cơn sốt đau đớn có mức trầm trọng khác nhau. Cho đến nay, bốn kiểu huyết thanh của virus dengue đã được nhận dạng: dengue-1 (DEN-1), dengue-2 (DEN-2), hoặc dengue-3 (DEN-3) trong tổ hợp với dengue-4 (DEN-4). Cơn sốt dengue do sự lây nhiễm virus dengue gây ra. Các kiểu huyết thanh khác có thể được khám phá trong tương lai (ví dụ DEN-5). Các kiểu huyết thanh virus dengue từ 1 đến 4 cũng có thể gây ra sốt xuất huyết do dengue (DHF), và triệu chứng sốc dengue (DSS). Các hậu quả trầm trọng nhất của việc lây nhiễm, DHF và DSS, có thể gây chết người. Virus dengue gây ra từ 50 đến 100 triệu ca sốt dengue suy yếu, 500.000 ca nhiễm DHF/DSS, và nhiều hơn 20.000 ca tử vong mỗi năm. Cho đến nay, không có vắc xin hữu hiệu nào để bảo vệ kháng lại sốt dengue và cũng không có phương pháp điều trị dược chất nào đối với bệnh này. Các nỗ lực phòng trừ muỗi là không hữu hiệu để ngăn ngừa sự bùng phát dengue ở các khu vực địa phương hoặc để ngăn ngừa sự lan rộng địa lý hơn nữa của bệnh này. Các tác giả sáng chế đã ước tính được rằng 3,5 tỷ người bị đe dọa bởi sự lây nhiễm virus dengue. Ngoài ra, virus dengue

là nguyên nhân gây sốt hàng đầu ở khách du lịch đến khu vực địa phương, như châu Á, Trung Mỹ và Nam Mỹ và khu vực Ca-ri-bê.

Tất cả bốn kiểu huyết thanh virus dengue là đặc hữu trong khắp tất cả các vùng nhiệt đới trên thế giới và cấu thành mối đe dọa virus do muỗi gây ra đáng kể nhất đối với con người ở các vùng nhiệt đới, trên toàn thế giới. Virus dengue được truyền sang người chủ yếu bởi muỗi *Aedes aegypti*. Việc lây nhiễm với một kiểu huyết thanh virus dengue dẫn đến việc bảo vệ suốt đời không bị lây nhiễm lại bởi kiểu huyết thanh đó, nhưng không ngăn ngừa sự lây nhiễm thứ cấp bởi một trong số ba kiểu huyết thanh virus dengue khác. Trên thực tế, sự lây nhiễm trước với một kiểu huyết thanh virus dengue dẫn đến nguy cơ bệnh trầm trọng gia tăng (DHF/DSS) khi lây nhiễm thứ cấp với kiểu huyết thanh khác nhau. Sự phát triển của vaccine hữu hiệu là phương pháp quan trọng để ngăn ngừa và kiểm soát bệnh xuất hiện trên toàn cầu. Nhiều lần gây miễn dịch tạo ra khó khăn bao phủ vaccine đầy đủ đối với các nỗ lực sức khỏe cộng đồng ở các quốc gia đặc hữu virus dengue cũng như những khách du lịch.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất được phẩm hữu hiệu có khả năng gây ra đáp ứng miễn dịch ở đối tượng đối với tất cả các typ huyết thanh virus dengue.

Sáng chế đề xuất được phẩm, bao gồm các cấu trúc virus dengue khả. Theo một số phương án, được phẩm có thể bao gồm các cấu trúc virus dengue khả có khung chính virus dengue giảm độc lực với gen cấu trúc từ ít nhất một kiểu huyết thanh virus dengue khác. Các phương án khác đề cập đến ít nhất một virus sống giảm độc lực trong tổ hợp với một hoặc nhiều virus dengue khả. Các phương án khác có thể bao gồm được phẩm chứa virus dengue khả có khung chính DEN-2 được cải biến (ví dụ PDK-53 là khung chính bắt đầu trong PI (chuyển tiếp-1) và tính biến thiên chuyển tiếp (sau khi chuyển tiếp và được sinh trưởng in vitro trên dòng tế bào cho phép) như được thể hiện đối với P2, P3,...P8..P10 v.v) và một hoặc nhiều thành phần cấu trúc của DEN-1, DEN-2, DEN-3 hoặc DEN-4. Theo các phương án khác, được phẩm gây miễn dịch được tạo ra mà ở đó khi được đưa vào trong đối tượng, được phẩm tạo ra đáp ứng miễn dịch đối với một hoặc nhiều virus dengue ở đối tượng. Do đó, các cấu trúc được dự định trong bản mô tả này có thể được tạo ra và chuyển tiếp in vitro, và mỗi trong số các lần chuyển tiếp tạo ra virus dengue giảm độc lực được dự định sử dụng trong được phẩm vaccine được dụng. Theo một số phương án virus sống giảm độc lực có thể là virus dengue-2 sống giảm độc lực đơn lẻ hoặc trong tổ hợp với một hoặc nhiều virus dengue khả.

Trong một số ví dụ, các cấu trúc virus dengue khả của kiểu huyết thanh virus dengue có thể bao gồm sự chuyển tiếp 7 (P7) virus sống giảm độc lực hoặc virus khả có trình tự axit nucleic được nhận dạng bởi SEQ ID NOS: 7, 15, 23 và 31 hoặc các trình tự polypeptit được chỉ định bởi SEQ ID NOS: 8, 3, 16, 10, 24, 34 và 32. Được dự định trong bản mô tả này là bất kỳ trong số các lần chuyển tiếp đối với bất kỳ trong số virus sống giảm độc lực được mô tả trong bản mô tả này có thể được sử dụng trong dược phẩm gây miễn dịch để gây ra đáp ứng miễn dịch đối với virus dengue đại diện (ví dụ các kiểu huyết thanh từ 1 đến 4). Theo các phương án này, dược phẩm gây miễn dịch mà bao gồm virus sống giảm độc lực được phân lập P-8 có thể được dùng cho đối tượng để gây ra đáp ứng miễn dịch kháng lại một hoặc nhiều kiểu huyết thanh dengue phụ thuộc vào cấu trúc được chọn. Ngoài ra, virus sống giảm độc lực có thể được tổ hợp với một hoặc nhiều trong số các virus khả này. Việc này được dự định đối với mỗi trong số các virus sống giảm độc lực được phân lập/được tạo ra trong mỗi lần chuyển tiếp tế bào tiếp theo (ví dụ tạo ra tế bào Vero khỉ xanh châu Phi, sau đây được gọi là: tế bào Vero). Được dự định trong bản mô tả này rằng dòng tế bào bất kỳ (ví dụ ngân hàng tế bào cung cấp GMP, FDA hoặc EMA được chấp thuận) có khả năng tạo ra virus dengue sử dụng để chuyển tiếp bất kỳ trong số các cấu trúc virus ở quy mô sản xuất hoặc nếu thích hợp được dự định trong bản mô tả này để sử dụng tiếp theo trong dược phẩm vắc xin hoặc gây miễn dịch kháng lại virus dengue.

Theo các phương án khác, dược phẩm được dự định trong bản mô tả này có thể được tổ hợp với các dược phẩm gây miễn dịch khác kháng lại các Flavivirus khác như virus West Nile, virus viêm não Nhật bản hoặc cấu trúc khả flavivirus khác bất kỳ và/hoặc virus sống giảm độc lực. Theo một số phương án, dược phẩm đơn có thể được sử dụng kháng lại nhiều flavivirus.

Theo một số phương án, dược phẩm gây miễn dịch của sáng chế có thể bao gồm virus dengue khả kháng lại một hoặc nhiều trong số DEN-1, DEN-2, DEN-3 và/hoặc DEN-4, đơn lẻ hoặc trong tổ hợp với dược phẩm virus dengue sống, giảm độc lực.

Theo các phương án khác, cấu trúc có thể bao gồm cấu trúc có các đột biến thích ứng trong các vùng cấu trúc hoặc không cấu trúc của virus mà làm tăng sự sinh trưởng hoặc sản xuất mà không tác động đến việc giảm độc lực hoặc tính an toàn của virus khi được đưa vào đối tượng. Theo một số phương án, bất kỳ trong số các cấu trúc virus dengue khả được dự định có thể bao gồm virus DEN-2 sống giảm độc lực có các đột biến đặc hiệu được sử dụng làm khung chính trong đó virus DEN-2 PDK sống giảm độc lực còn bao gồm các protein cấu trúc của một hoặc nhiều protein cấu trúc

prM (tiền màng) và E (bao màng) của các kiểu huyết thanh dengue khác nhau. Ngoài ra, khung chính DEN-2 có thể bao gồm các đột biến bổ sung để làm tăng việc sản xuất hoặc tăng cường đáp ứng miễn dịch để xác định sơ bộ được phẩm ở đối tượng khi dùng (ví dụ virus dengue khảm 2/1, 2/3 hoặc 2/4).

Theo một số phương án, các gen protein cấu trúc có thể bao gồm các gen prM và E của DEN-1, DEN-2, DEN-3 hoặc DEN-4 trên khung chính DEN-2 có một hoặc hai đột biến mà là một phần của virus dengue sống giảm độc lực. Ví dụ, cấu trúc dengue, theo một số phương án có thể bao gồm các cấu trúc được gọi là DENVax-1-A, DENVax-2-F, DENVax-3-F, và DENVax-4-F (xem phần ví dụ thực hiện sáng chế) trong đó khung chính DEN-2 có một hoặc nhiều đột biến (ví dụ không được tìm thấy trong PI hoặc virus chuyển tiếp trước khác hoặc PDK-53) từ virus sống giảm độc lực DEN-2 được chứng minh trước để là an toàn và hữu hiệu để gây ra đáp ứng miễn dịch. Virus sống giảm độc lực DEN-2 của sáng chế là phiên bản được cải thiện của virus sống giảm độc lực DEN-2 được sử dụng ban đầu. Cấu trúc khảm của sáng chế có thể bao gồm khung chính DEN-2 PDK-53 giảm độc lực được cải biến, có một hoặc nhiều protein cấu trúc của kiểu huyết thanh virus dengue thứ hai trong đó protein cấu trúc có thể bao gồm các đột biến bổ sung để làm tăng đáp ứng miễn dịch đối với cấu trúc khảm. Theo một số phương án, một số đột biến đạt được bởi DEN-2 PDK-53 giảm độc lực có thể tạo ra sự thay đổi axit amin bảo tồn hoặc không trong các cấu trúc khác với cấu trúc PI mà có thể dẫn đến tính trạng mong muốn để sản xuất, v.v.

Theo các phương án khác, hệ gen DEN-2 sống giảm độc lực có thể được sử dụng để tạo ra các cấu trúc của kiểu huyết thanh virus dengue 1 (DEN-1) và kiểu huyết thanh virus dengue 3 (DEN-3), kiểu huyết thanh virus dengue 4 (DEN-4) trong đó một hoặc nhiều gen protein cấu trúc của hệ gen virus DEN-2 có thể được thay thế bằng một hoặc nhiều gen protein cấu trúc lần lượt của DEN-1, DEN-3 hoặc DEN-4. Theo một số phương án, protein cấu trúc có thể là C, prM hoặc protein E của virus dengue thứ hai. Theo một số phương án, gen protein cấu trúc bao gồm các gen prM và E của DEN-1, DEN-3 hoặc DEN-4. Các virus lai này biểu hiện kháng nguyên bề mặt của DEN-1, DEN-3 hoặc DEN-4 trong khi giữ lại kiểu hình giảm độc lực của DEN-2 giảm độc lực cha mẹ.

Các cấu trúc được bộc lộ trong bản mô tả này có thể bao gồm các cấu trúc khảm của DEN-4, DEN-2, DEN-1, và DEN-3 biểu hiện kháng nguyên bề mặt của DEN-1, DEN-3 và DEN-4 bằng cách sử dụng virus DEN-2 giảm độc lực làm khung chính.

Theo một số phương án, được phẩm của sáng chế có thể bao gồm được phẩm mà bao gồm cấu trúc virus dengue khảm đơn lẻ được bộc lộ trong bản mô tả này và

chất mang được dụng hoặc tá được. Theo cách khác, được phẩm của sáng chế có thể bao gồm được phẩm mà bao gồm hai hoặc nhiều hoặc ba hoặc nhiều cấu trúc virus dengue khả dụng bộc lộ trong bản mô tả này, và chất mang được dụng hoặc tá được. Theo các phương án này, một hoặc nhiều các cấu trúc khả virus dengue được dự định trong bản mô tả này có thể được tổ hợp với một hoặc nhiều, virus dengue sống giảm độc lực. Theo một số phương án, virus sống giảm độc lực có thể là virus DEN-2 sống giảm độc lực trong đó các đột biến bổ sung trong các vùng NCR, NS1 hoặc các vùng khác làm tăng đáp ứng miễn dịch, làm tăng sự sinh trưởng virus hoặc sự cải thiện khác đối với virus dengue sống giảm độc lực được cải thiện.

Theo một số phương án, các locus giảm độc lực, nucleotit 5'NCR-57-T, NS1-53-Asp, và NS3-250-Val, của vacxin DENV-2 đã được xác định trước và tất cả trong số các thay đổi này được dùng chung bởi khung chính di truyền đặc hiệu virus PDK-53 chung của bốn virus DENVax. Trình tự di truyền của ba locus giảm độc lực cũng như được thiết lập trước in vitro và in vivo kiểu hình giảm độc lực của các ứng viên vacxin này được kiểm tra một cách cẩn thận đối với hạt DENVax được sản xuất bởi cGMP. Báo cáo này mô tả chiến lược được sử dụng để tạo ra hạt virus chủ (MVS) cũng như mô tả đặc điểm di truyền và kiểu hình của chúng sử dụng trong sản xuất được phẩm vacxin virus dengue. MVS có thể được sử dụng để sản xuất các vật liệu lâm sàng và cuối cùng là cung cấp vacxin thương mại.

Mô tả vắn tắt các hình vẽ

Các hình vẽ sau đây tạo thành một phần của sáng chế và được bao gồm để chứng minh hơn nữa một số phương án. Một số phương án có thể được hiểu tốt hơn bằng việc tham khảo đến một hoặc nhiều hình vẽ này đơn lẻ hoặc trong tổ hợp với phần mô tả chi tiết của các phương án cụ thể được thể hiện.

Fig.1 thể hiện biểu đồ lấy làm ví dụ phản ánh cấu trúc khả lấy làm ví dụ của sáng chế, DEN-2/DEN-4 so với các cấu trúc được tạo ra trước và virus dengue kiểu hoang dại.

Fig.2 thể hiện biểu đồ lấy làm ví dụ so sánh các đáp ứng khác nhau bằng cách sử dụng khung chính DEN-2 sống giảm độc lực (với các đột biến bổ sung) và kiểu huyết thanh virus dengue thứ hai là thành phần cấu trúc được thế đối với các thành phần cấu trúc dengue-2 (ví dụ DENVax-1 MVS). Biểu đồ này minh họa kích cỡ vết tan của DENVax MVS. Virus dengue kiểu hoang dại và virus ứng viên vacxin loại nghiên cứu được công bố trước được bao gồm để đối chứng và so sánh. Biểu đồ này

minh hoạ việc sản xuất được cải thiện của các cấu trúc virut dengue so với các cấu trúc khảm virut dengue đối chứng.

Fig.3 thể hiện biểu đồ lấy làm ví dụ mà thể hiện tính nhạy nhiệt độ của DENVax MVS (Master Virus Seed). Virut dengue kiểu hoang dại và virut ứng viên vacxin loại nghiên cứu được công bố trước được bao gồm để so sánh với loại MVS.

Fig.4 thể hiện biểu đồ lấy làm ví dụ mà thể hiện sự sinh trưởng virut của DENVax MVS trong tế bào C6/36 so với các đối chứng. Virut dengue kiểu hoang dại và virut ứng viên vacxin loại nghiên cứu được bao gồm để so sánh với DENVax MVS.

Các hình vẽ từ Fig.5A đến Fig.5C thể hiện biểu đồ lấy làm ví dụ của độc lực thần kinh ở chuột mới sinh. Các kết quả được nhóm lại của vài thử nghiệm tổng kết độc lực thần kinh của virut kiểu hoang dại DENV-2 16681 trong CDC-ICR ($n=72$) và chuột mới sinh Taconic-ICR ($w=32$) được thử thách ic với 10^4 pfu của virut (A). Độc lực thần kinh của DENVax MVS được thử nghiệm ở chuột Taconic-ICR với liều 10^4 pfu (B) hoặc 10^3 pfu (C). Số lượng động vật được thử nghiệm ở mỗi nhóm trong một thử nghiệm ($n=16$) hoặc hai thử nghiệm gộp lại ($n=31$ hoặc 32) được chỉ định.

Fig.6 thể hiện biểu đồ lấy làm ví dụ minh hoạ kích cỡ vết tan của DENVax MVS, WVS, và BVS. Đường kính vết tan trung bình $\pm$ SD (cột báo lỗi) của vết tan virut trong tế bào Vero hoặc LLC-MK2 trong lớp phủ agarosa được đo vào ngày 9pi. DENV kiểu hoang dại và virut ứng viên vacxin loại nghiên cứu được công bố trước được bao gồm để đối chứng và so sánh.

Fig.7 thể hiện biểu đồ lấy làm ví dụ minh hoạ sự sinh trưởng của DENVax MSV, WVS, và BVS trong tế bào C6/36 ở hai nhiệt độ ủ để xác nhận sự duy trì của chúng của gen đánh dấu giảm độc lực in vitro sau khi sản xuất quy mô lớn.

Fig.8 thể hiện biểu đồ lấy làm ví dụ giới hạn sự sinh trưởng của DENVax MVS, WVS, và BVS trong tế bào C6/36. Độ chuẩn trung bình $\pm$ SD (cột báo lỗi) của virut được sao chép trong tế bào C6/36 7 ngày pi. Virut dengue kiểu hoang dại và virut ứng viên vacxin loại nghiên cứu được công bố trước được bao gồm để so sánh.

Các hình vẽ từ Fig.9A đến Fig.9B thể hiện biểu đồ lấy làm ví dụ của dữ liệu độc lực thần kinh của DENVax MVS ở chuột ICR mới sinh. (A) tiêm truyền IC của virut ở liều 10^4 PFU. (B) tiêm truyền IC của virut ở liều 10^3 PFU.

Fig.10 thể hiện biểu đồ lấy làm ví dụ so sánh virut sống giảm độc lực mới với virut dengue sống giảm độc lực được tạo ra trước.

Mô tả chi tiết sáng chế

Các định nghĩa

Như được sử dụng trong bản mô tả này, thuật ngữ số ít “một” có thể có nghĩa là một hoặc nhiều hơn một điểm.

Như được sử dụng trong bản mô tả này, "đối tượng" hoặc "các đối tượng" có thể bao gồm, nhưng không chỉ giới hạn ở, động vật có vú như con người hoặc động vật có vú, động vật thuần hoá hoặc kiểu hoang dại, ví dụ chó, mèo, động vật kiểng trong gia đình khác (ví dụ chuột đồng, chuột lang, chuột nhắt, chuột cống), chồn, lợn, ngựa, gia súc, sóc chó, động vật gặm nhấm hoang dại hoặc động vật vườn thú.

Như được sử dụng trong bản mô tả này, thuật ngữ "thể khảm virus," "virus khảm," "thể khảm flavivirus" và "flavivirus khảm" có thể có nghĩa là cấu trúc bao gồm phần trình tự nucleotit của virus dengue-2 và trình tự nucleotit khác mà không phải từ virus dengue-2 hoặc từ flavivirus khác nhau. "Thể khảm dengue" bao gồm ít nhất hai kiểu huyết thanh dengue khác nhau nhưng không phải flavivirus khác nhau. Do đó, các ví dụ về các virus dengue khác hoặc flavivirus bao gồm, nhưng không chỉ giới hạn ở, các trình tự từ virus dengue-1, virus dengue-3, virus dengue-4, virus West Nile, virus viêm não Nhật bản, virus viêm não St. Louis, virus viêm não truyền từ ve, virus sốt vàng và tổ hợp bất kỳ của chúng.

Như được sử dụng trong bản mô tả này, "thể khảm axit nucleic" có thể có nghĩa là cấu trúc của sáng chế bao gồm axit nucleic bao gồm phần trình tự nucleotit của virus dengue-2 và trình tự nucleotit khác mà không có cùng nguồn gốc như trình tự nucleotit của virus dengue-2. Do đó, flavivirus khảm bất kỳ hoặc thể khảm flavivirus được bộc lộ trong bản mô tả này có thể được nhận biết như ví dụ về thể khảm axit nucleic.

Như được sử dụng trong bản mô tả này, thuật ngữ "virus sống giảm độc lực" có thể có nghĩa là virus kiểu hoang dại, được đột biến hoặc được chọn đối với các tính trạng sử dụng trong vaccin hoặc được phẩm gây miễn dịch khác trong đó một vài tính trạng có thể bao gồm tính độc giảm, tính an toàn, hiệu quả hoặc sự sinh trưởng được cải thiện, v.v.

Trong các phần sau đây, các được phẩm lấy làm ví dụ khác nhau và phương pháp được mô tả để mô tả chi tiết các phương án khác nhau. Sẽ hiển nhiên đối với chuyên gia trong lĩnh vực này rằng việc thực hiện các phương án khác nhau không đòi hỏi phải dùng tất cả hoặc thậm chí một vài phần mô tả chi tiết cụ thể được nêu trong bản mô tả này, nhưng tốt hơn là nồng độ, thời gian và các phần mô tả chi tiết cụ thể khác có thể được cải biến qua việc thử nghiệm thông thường. Trong một số trường

hợp, các phương pháp hoặc thành phần đã được biết đến vẫn không được bao gồm trong bản mô tả này.

Theo các phương án của sáng chế, có thể dùng các kỹ thuật sinh học phân tử thông thường, hoá học protein, vi sinh vật và các kỹ thuật ADN tái tổ hợp nằm trong kỹ năng của lĩnh vực này. Các kỹ thuật này được giải thích một cách đầy đủ trong tài liệu chuyên ngành. Xem, ví dụ, ấn phẩm: Sambrook, Fritsch & Maniatis, *Molecular Cloning: A Laboratory Manual*, Second Edition 1989, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.; *Animal Cell Culture*, R. I. Freshney, ed., 1986).

Các phương án trong bản mô tả này đề cập đến được phẩm, phương pháp và sử dụng để gây ra đáp ứng miễn dịch kháng lại một hoặc nhiều kiểu huyết thanh dengue ở đối tượng, theo cách riêng rẽ hoặc theo cách đồng thời. Theo các phương án này, virus dengue giảm độc lực và thể khảm axit nucleic được tạo ra và được sử dụng trong các được phẩm vaccin được bộc lộ trong bản mô tả này.

Một số phương án đề cập đến các cấu trúc dengue được cải biến hoặc đột biến hoặc các thể khảm. Các phương án khác đưa các đột biến vào để cải biến các trình tự axit amin của các protein cấu trúc của virus dengue trong đó đột biến làm tăng tính gây miễn dịch đối với virus.

Các virus dengue sống giảm độc lực của tất cả bốn kiểu huyết thanh đã được phát triển bằng việc chuyển tiếp virus kiểu hoang dại vào môi trường nuôi cấy tế bào. Có một vài trong số các ứng viên vaccin sống giảm độc lực có triển vọng nhất để gây miễn dịch kháng flavivirus và cụ thể là sự lây nhiễm và/hoặc bệnh virus dengue. Các ứng viên vaccin này đã được ký hiệu bởi tổ hợp của kiểu huyết thanh của chúng, dòng tế bào mà qua đó chúng được chuyển tiếp và khoảng thời gian chúng chuyển tiếp. Do đó, virus kiểu hoang dại có kiểu huyết thanh dengue 1 chuyển tiếp trong tế bào PDK 13 lần được ký hiệu là virus DEN-1 PDK-13. Các ứng viên vaccin khác là DEN-2 PDK-53, DEN-3 PGMK-30/FRhL-3 (ví dụ lần chuyển tiếp thứ ba trong tế bào thận của khỉ xanh nguyên thủy, tiếp theo ba lần chuyển tiếp trong tế bào phổi bào thai khỉ nâu và DEN-4 PDK-48). Có bốn virus vaccin ứng viên thu được bằng việc chuyển tiếp nuôi cấy mô lần lượt của các virus DEN-1 16007, DEN-2 16681, DEN-3 16562 và DEN-4 1036 cha mẹ kiểu hoang dại.

Theo một số phương án, virus vaccin dengue-2 PDK-53 sống giảm độc lực chứa hỗn hợp gồm virus, với quần thể chứa sự khác nhau nucleotit thay đổi. Sau khi mô tả đặc điểm di truyền của các đột biến giảm độc lực, một số đặc điểm giảm độc lực được nêu rõ và được thiết kế thành đơn dòng lây nhiễm ADN bổ trợ. ARN được phiên

mã từ thể đơn dòng lây nhiễm này và được đưa vào trong tế bào Vero như lần chuyển tiếp 1 của virus PDK-53-Vero-DEN-2-P 1 được mô tả đặc điểm và thu được mới (xem ví dụ, bảng 1). Virus giảm độc lực này được tạo ra đối với mỗi kiểu huyết thanh DEN, nhưng đối với DEN-1, DEN-3 và DEN-4, gen prM và E được thiết kế thành 3 dòng lây nhiễm ADN bổ trợ riêng biệt, do đó tạo ra bốn virus PDK-53-Vero riêng biệt (được gọi trong bản mô tả này là: PDK-53-Vero-DEN-2-P 1, PDK-53-Vero-DEN-1-P 1, PDK-53-Vero-DEN-3-P 1, và PDK-53-Vero-DEN-4-P 1). Các chủng virus vaccine giảm độc lực này chuyển tiếp vào tế bào Vero 10 lần (bảng 1) và mỗi dòng riêng biệt đòi hỏi đột biến phụ thuộc vào sự thích ứng của chúng sinh trưởng trong tế bào Vero (bảng 3). Một số phương án trong bản mô tả này đề cập đến sự đạt được và sử dụng đối với virus dengue sống giảm độc lực này.

Các thử nghiệm lâm sàng trước của người với virus giảm độc lực đã thể hiện rằng DEN-2 PDK-53 có liều lây nhiễm thấp nhất (50% liều lây nhiễm tối thiểu của 5 đơn vị tạo vết tan hoặc PFU) ở người, là gây miễn dịch mạnh và không gây ra các mối quan tâm an toàn rõ rệt. Các ứng viên virus vaccine DEN-1 PDK-13, DEN-3 PGMK-30/FRhL-3 và DEN-4 PDK-48 có liều lây nhiễm tối thiểu lớn hơn 50% lần lượt của 10.000, 3500 và 150PFU, ở người. Mặc dù chỉ có một lần gây miễn dịch với virus DEN-2 PDK-53 hoá trị một hoặc virus DEN-4 PDK-48 được đòi hỏi để đạt được 100% sự chuyển hoá huyết thanh ở đối tượng là con người, liều cao hơn được đòi hỏi để đạt được cùng tốc độ chuyển hoá huyết thanh đối với virus DEN-1 PDK-13 và DEN-3 PGMK-30/FRhL-3, mà có hai liều lây nhiễm cao nhất đối với con người.

Ứng viên vaccine virus DEN-2 PDK-53, cũng được viết tắt là PDK-53, có vài gen đánh dấu sinh học có thể đo kết hợp với sự giảm độc lực, bao gồm tính nhạy nhiệt độ, kích cỡ vết tan nhỏ, sự sao chép giảm ở môi trường nuôi cấy tế bào muỗi C6136, sự sao chép giảm ở muỗi nguyên vẹn, mất độc lực thần kinh ở chuột đang bú và tỷ lệ mắc phải virus huyết giảm ở khỉ. Các thử nghiệm lâm sàng của vaccine ứng viên PDK-53 đã chứng minh tính an toàn của nó và tính gây miễn dịch ở người. Hơn nữa, vaccine PDK-53 gây ra các đáp ứng nhớ tế bào T đặc hiệu virus dengue ở đối tượng nhận vaccine là con người. Một số phương án trong bản mô tả này mô tả sự cải thiện đối với DEN-2 PDK-53 được sử dụng trong các cấu trúc khả năng được mô tả trong bản mô tả này.

Thể khả năng flavivirus gây miễn dịch có khung chính virus dengue-2 và ít nhất một protein cấu trúc của kiểu huyết thanh virus dengue khác có thể được sử dụng để tạo ra thể khả năng virus dengue và các phương pháp tạo ra thể khả năng virus dengue được mô tả. Các thể khả năng virus dengue gây miễn dịch được đề xuất, đơn lẻ hoặc trong tổ

hợp, với chất mang được dụng như được phẩm gây miễn dịch để giảm thiểu, ức chế hoặc gây miễn dịch cá thể kháng lại sự lây nhiễm bởi một hoặc nhiều kiểu huyết thanh, như kiểu huyết thanh dengue DEN-1, DEN-2, DEN-3 và DEN-4, đơn lẻ hoặc trong tổ hợp. Khi tổ hợp, các thể khảm virus dengue gây miễn dịch có thể được sử dụng làm vaccin đa hoá trị (ví dụ hoá trị hai, ba và bốn) để tạo ra sự bảo vệ đồng thời kháng lại sự lây nhiễm bởi nhiều hơn một loài hoặc chủng flavivirus. Theo một số phương án, thể khảm virus dengue được tổ hợp vào được phẩm gây miễn dịch hữu ích làm vaccin hoá trị hai, hoá trị ba hoặc hoá trị bốn kháng lại kiểu huyết thanh dengue đã biết hoặc tạo ra tính miễn dịch đối với các flavivirus mầm bệnh khác bằng cách bao gồm axit nucleic mã hoá một hoặc nhiều protein từ các flavivirus khác nhau.

Theo một số phương án, các thể khảm virus dengue không độc lực gây miễn dịch được đề xuất trong bản mô tả này chứa các gen protein không cấu trúc của virus dengue-2 giảm độc lực (ví dụ PDK-53), hoặc dạng tương đương của nó và một hoặc nhiều gen protein cấu trúc hoặc các phần gây miễn dịch của nó của flavivirus kháng lại mà tính miễn dịch được gây ra ở đối tượng. Ví dụ, một số phương án đề cập đến thể khảm có hệ gen PDK-53 virus dengue-2 giảm độc lực làm khung chính virus và một hoặc nhiều gen protein cấu trúc mã hoá capsit, tiền màng/màng hoặc vỏ bọc của hệ gen PDK-53 hoặc tổ hợp của nó, được thay thế bằng một hoặc nhiều gen protein cấu trúc tương ứng từ DEN-1, DEN-3 hoặc DEN-4 hoặc flavivirus khác để được bảo vệ kháng lại, như flavivirus khác nhau hoặc kiểu huyết thanh virus dengue khác nhau. Theo các phương án này, thể khảm axit nucleic được bộc lộ trong bản mô tả này có thể có các đặc tính chức năng của virus dengue-2 giảm độc lực và là không độc lực, nhưng biểu hiện epitop kháng nguyên của các sản phẩm gen cấu trúc của DEN-1, DEN-3 hoặc DEN-4 ngoài các flavivirus khác và là gây miễn dịch (ví dụ gây ra đáp ứng miễn dịch đối với các sản phẩm gen ở đối tượng). Sau đó, các cấu trúc ADN này được sử dụng để phiên mã ARN từ dòng lây nhiễm, ARN này được đưa vào trong tế bào Vero một lần nữa tạo ra virus thế hệ mới ở PI. Các virus thế hệ mới này có thể khác biệt với PDK-53. (xem ví dụ P 1-P10).

Theo phương án khác, thể khảm axit nucleic có thể là thể khảm axit nucleic có, nhưng không chỉ giới hạn ở, trình tự nucleotit thứ nhất mã hoá protein không cấu trúc từ virus dengue-2 giảm độc lực và trình tự nucleotit thứ hai mã hoá protein cấu trúc từ virus dengue-4 đơn lẻ hoặc trong tổ hợp với flavivirus khác. Theo các phương án khác, virus dengue-2 giảm độc lực có thể là chủng vaccin PDK-53 có một hoặc nhiều axit amin đột biến (xem phần ví dụ thực hiện sáng chế). Các đột biến bổ sung này tạo ra các tính trạng mong muốn để sử dụng như dengue-2 sống giảm độc lực hoặc như các

cấu trúc khản được mô tả trong bản mô tả này. Một số phương án bao gồm protein cấu trúc của một hoặc nhiều C, prM hoặc protein E của virus dengue thứ hai.

Các khía cạnh khác bao gồm rằng virus khản có thể bao gồm nucleotit và sự thay thế, loại bớt hoặc xen vào axit amin ví dụ, trong hệ gen PDK-53 dengue-2 đối chứng để làm giảm sự cản trở các đáp ứng gây miễn dịch với kiểu huyết thanh virus dengue đích. Các cải biến này có thể được tạo thành trong các protein cấu trúc và không cấu trúc đơn lẻ hoặc trong tổ hợp với các cải biến ví dụ được bộc lộ trong bản mô tả này và có thể được tạo ra bằng cách chuyển tiếp virus giảm độc lực và thu được được phẩm được cải thiện để gây ra đáp ứng miễn dịch kháng lại một hoặc nhiều kiểu huyết thanh dengue.

Một số phương án được bộc lộ trong bản mô tả này đề xuất phương pháp tạo ra virus khản của sáng chế bằng cách sử dụng các kỹ thuật tái tổ hợp, bằng cách xen các thay thế đòi hỏi vào trong hệ gen khung chính thích hợp. Các phương án khác trong bản mô tả này đề cập đến việc chuyển tiếp virus khản sống giảm độc lực được xác nhận (ví dụ an toàn và hữu hiệu) đối với các cải thiện bổ sung. Theo một số phương án, khung chính dengue-2 được sử dụng trong bản mô tả này có thể bao gồm một hoặc nhiều đột biến được thể hiện trong bảng 3. Theo các phương án khác, thể khản dengue-dengue của sáng chế có thể bao gồm một hoặc nhiều đột biến như được thể hiện trong bảng 3. Theo các phương án khác nữa, thể khản dengue-dengue có thể bao gồm tất cả trong số các đột biến đối với mỗi thể khản như được thể hiện trong bảng 3 đối với Den-2/Den-1, Den-2/Den-3 hoặc Den-2/Den-4.

Dược phẩm mà bao gồm virus sống giảm độc lực được thể hiện bởi các cấu trúc của bảng 3 được dự định. Ví dụ, dược phẩm hoá trị một, hai, ba hoặc bốn được dự định sử dụng trong bản mô tả này bằng cách sử dụng các thể khản và virus dengue-2 sống giảm độc lực như được thể hiện trong bảng 3.

Theo một số phương án, biến thể DEN-2 sống giảm độc lực được dự định trong bản mô tả này có thể được bào chế thành dược phẩm trong đó dược phẩm có thể được dùng đơn lẻ hoặc trong tổ hợp với các thể khản dengue-dengue hoặc thể khản dengue-flavivirus. Theo một số phương án, dược phẩm hoá trị hai, ba hoặc bốn có thể được dùng trong ứng dụng đơn lẻ hoặc trong nhiều ứng dụng cho đối tượng.

Thể khản flavivirus

Các kiểu virus dengue từ 1 đến 4 (từ DEN-1 đến DEN-4) là mầm bệnh flavivirus do muỗi gây ra. Hệ gen flavivirus chứa vùng không mã hoá 5' (5'-NC), tiếp theo vùng mã hoá capsit protein (C), tiếp theo vùng mã hoá protein tiền màng/màng

(prM), tiếp theo vùng mã hoá protein bao ngoài (E), tiếp theo vùng mã hoá các protein không cấu trúc (NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5) và cuối cùng vùng không mã hoá 3' (3TSTC). Các protein cấu trúc virus là C, prM và E, và protein không cấu trúc là NS1-NS5. Các protein cấu trúc và không cấu trúc được dịch mã như polyprotein đơn lẻ và được xử lý bởi proteaza tế bào và virus.

Các thể khảm flavivirus có thể là các cấu trúc được tạo thành bằng cách dung hợp các gen protein không cấu trúc từ một kiểu hoặc kiểu huyết thanh, của virus dengue hoặc loài virus của flaviviridae, với gen protein, ví dụ, gen protein cấu trúc, từ kiểu hoặc kiểu huyết thanh khác nhau, của virus dengue hoặc các loài virus của flaviviridae. Theo cách khác, thể khảm flavivirus của sáng chế là cấu trúc được tạo thành bằng cách dung hợp gen protein không cấu trúc từ một kiểu hoặc kiểu huyết thanh, của virus dengue hoặc các loài virus của flaviviridae, với các trình tự nucleotit khác mà định hướng tổng hợp polypeptit hoặc protein được chọn từ các kiểu huyết thanh dengue khác hoặc virus khác của flaviviridae.

Theo các phương án khác, thể khảm flavivirus gây miễn dịch không độc lực được đề xuất trong bản mô tả này chứa các gen protein không cấu trúc của virus dengue-2 giảm độc lực hoặc dạng tương đương của nó và một hoặc nhiều trong số các gen protein cấu trúc hoặc các phần kháng nguyên của nó, của flavivirus kháng lại mà tính gây miễn dịch để được gây ra. Các flavivirus thích hợp bao gồm, nhưng không chỉ giới hạn ở các flavivirus được liệt kê trong bảng 1.

Các virus dengue thích hợp khác để sử dụng trong việc tạo cấu trúc các thể khảm có thể là DEN-1 16007, DEN-2 16681, DEN-3 16562 và DEN-4 1036 kiểu hoang dại độc lực và chủng vắc xin giảm độc lực DEN-1 PDK-13, DEN-2 PDK-53, DEN-3 PMK-30/FRhL-3 và DEN-4 PDK-48. Sự khác nhau di truyền giữa các cặp virus DEN-1, DEN-2, DEN-3 và DEN-4 kiểu hoang dại/giảm độc lực được dự định cùng với các thay đổi trong trình tự axit amin được mã hoá bởi hệ gen virus.

Danh mục trình tự đối với DEN-2 PDK-53 tương ứng với biến thể DEN-2 PDK-53-V, trong đó vị trí nucleotit hệ gen 5270 được đột biến từ A thành T và vị trí axit amin 1725 của polyprotein hoặc vị trí axit amin 250 của NS3 protein chứa gốc valin. Biến thể DEN-2 PDK-53 mà không có đột biến nucleotit này, DEN-2 PDK-53-E, khác với PDK-53-V chỉ ở một vị trí này. DEN-2 PDK-53-E có A ở vị trí nucleotit 5270 và glutamat ở polyprotein vị trí axit amin 1725, NS3 protein vị trí axit amin 250. Nên được hiểu rằng các phương án trong bản mô tả này bao gồm PDK 53 được cải biến mà bao gồm một hoặc nhiều lần chuyển tiếp trong tế bào vật chủ riêng biệt (ví dụ

tế bào Vero, xem bảng 1) trong đó các chủng mong muốn sử dụng trong được phẩm vaccin được dự định trong bản mô tả này được tạo ra.

Theo một số phương án, việc thiết kế thể khảm có thể được đưa vào khung chính được cải biến dòng lây nhiễm đặc hiệu virus DEN-2 và gen cấu trúc (prM-E hoặc C-prM-E) xen vào virus dengue khác hoặc các flavivirus khác. DEN-2 đối với khung chính dengue-2, tiếp theo chủng mà từ đó gen cấu trúc được xen vào. Một biến thể khung chính DEN-2 được phản chiếu trong chữ cái tiếp theo sau số ký hiệu. Một biến thể khung chính DEN-2 cụ thể mà từ đó thể khảm được tạo cấu trúc được chỉ định bởi chữ cái tiếp theo được đặt sau dấu gạch nối, 16681 (P) cha mẹ, PDK-53-E (E), hoặc PDK-53-V (V); chữ cái sau cùng cho biết gen cấu trúc C-prM-E từ chủng (P) cha mẹ hoặc dẫn xuất vắ xin của nó (V) hoặc gen cấu trúc prM-E từ (P) cha mẹ hoặc dẫn xuất vaccin của nó (VI). Ví dụ; DEN-2/1-VP có nghĩa là thể khảm bao gồm khung chính DEN-2 PDK-53V giảm độc lực bao gồm valin ở NS3-250 và gen C-prM-E từ DEN-1 16007 kiểu hoang dại; DEN-2/1-VV có nghĩa là khung chính DEN-2 PDK-53V với chủng vaccin của dengue-1, DEN-1 PDK-13; DEN-2/1-VP1 có nghĩa là khung chính DEN-2 PDK-53V và gen prM-E từ DEN-1 16007 kiểu hoang dại; DEN-2/3-VP1 có nghĩa là khung chính DEN-2 PDK-53 V và gen prM-E từ DEN-3 16562 kiểu hoang dại; DEN-2/4VP 1 có nghĩa là khung chính DEN-2 PDK-53V và gen prM-E từ DEN-4 1036 kiểu hoang dại. Các thể khảm khác được bộc lộ trong bản mô tả này được chỉ định theo cùng cách.

Theo một phương án, các thể khảm được bộc lộ trong bản mô tả này chứa hệ gen PDK-53 virus dengue-2 giảm độc lực làm khung chính virus, trong đó gen protein cấu trúc mã hoá C, prM và protein E của hệ gen PDK-53 hoặc các tổ hợp của chúng, có thể được thay thế bằng gen protein cấu trúc tương ứng từ virus dengue-1, dengue-3 hoặc dengue-4 và tùy ý, flavivirus khác để được bảo vệ kháng lại, như flavivirus khác nhau hoặc chủng virus dengue khác nhau.

Trong vùng protein không cấu trúc, đột biến Gly- thành -Asp (từ kiểu hoang dại thành PDK-53) được phát hiện ở protein không cấu trúc NS1-53 (vị trí nucleotit hệ gen 2579); đột biến Leu- thành -Phe (kiểu hoang dại thành -PDK-53) được phát hiện ở protein không cấu trúc NS2A-181 (vị trí nucleotit hệ gen 4018); đột biến Glu- thành -Val (kiểu hoang dại thành -PDK-53) được phát hiện ở protein không cấu trúc NS3-250 (vị trí nucleotit hệ gen 5270); và đột biến Gly- thành -Ala (kiểu hoang dại thành -PDK-53) được phát hiện ở protein không cấu trúc NS4A-75 (vị trí nucleotit hệ gen 6599). Virus DEN-2 sống giảm độc lực của sáng chế còn bao gồm các đột biến như được thể hiện trong các thể khảm bất kỳ hoặc virus dengue-2 sống giảm độc lực của bảng 3.

Chủng virus PDK-53 có kiểu di truyền hỗn hợp ở nucleotit hệ gen 5270. Phần đáng kể (khoảng 29%) của quần thể virus mã hoá NS3-250-Glu không đột biến mà có mặt trong virus DEN-2 16681 kiểu hoang dại tốt hơn là đột biến NS3-250-Val. Đối với cả hai biến thể di truyền là không độc lực, đột biến này có thể là không nhất thiết trong thể khả không độc lực.

Trước đây, các tác giả sáng chế đã phát hiện ra rằng tính không độc lực của chủng virus PDK-53 giảm độc lực có thể được cho là do các đột biến trong trình tự nucleotit mã hoá protein không cấu trúc và trong vùng không mã hoá 5'. Ví dụ, đột biến đơn lẻ ở NSI-53, đột biến kép ở NSI-53 và ở 5TSTC-57, đột biến kép ở NSI-53 và ở NS3-250 và đột biến bộ ba ở NSI-53, ở 5TSTC-57 và ở NS3-250, dẫn đến việc giảm độc lực của virus DEN-2. Do đó, hệ gen của virus dengue-2 bất kỳ chứa các thay thế axit amin không bảo tồn hoặc các thay thế nucleotit ở các locus này có thể được sử dụng làm trình tự gốc để thu được các virus PDK-53 được cải biến được bộc lộ trong bản mô tả này. Đột biến khác xét đến cấu trúc thân/vòng ở vùng không mã hoá 5' sẽ tạo ra độ ổn định kiểu hình không độc lực bổ sung, nếu muốn. Các đột biến đối với vùng này phá vỡ các cấu trúc thứ cấp tiềm năng quan trọng đối với sự sao chép virus. Đột biến đơn lẻ trong cấu trúc thân ngắn này (chiều dài chỉ có 6 gốc nucleotit) trong DEN và virus gây viêm não ngựa Venezuela phá vỡ sự hình thành cấu trúc kẹp tóc. Các đột biến khác trong cấu trúc thân này làm giảm khả năng nghịch đảo ở locus này, trong khi duy trì khả năng sống của virus.

Các đột biến được bộc lộ trong bản mô tả này có thể đạt được bằng phương pháp bất kỳ đã được biết đến trong lĩnh vực này bao gồm, nhưng không chỉ giới hạn ở, các dòng xuất hiện tự nhiên hoặc chọn lọc có các dấu hiệu bổ sung khi chuyển tiếp vào dòng tế bào quan tâm (ví dụ tế bào Vero). Sẽ được hiểu bởi chuyên gia trong lĩnh vực này rằng các thử nghiệm sàng lọc độc lực, như được mô tả trong bản mô tả này và như đã được biết đến trong lĩnh vực này, có thể được sử dụng để phân biệt giữa các cấu trúc khung chính độc lực và không độc lực.

Cấu trúc của thể khả flavivirus

Các thể khả flavivirus được mô tả trong bản mô tả này có thể được tạo ra bằng cách ghép một hoặc nhiều gen protein cấu trúc của flavivirus kháng lại mà tính miễn dịch được mong muốn vào trong khung chính hệ gen virus PDK-53 dengue hoặc các phương pháp khác đã được biết đến trong lĩnh vực này, bằng cách sử dụng thao tác tái tổ hợp để loại bỏ gen PDK-53 tương ứng và thay thế nó bằng gen virus dengue-1, dengue-3 hoặc dengue-4 hoặc các gen khác đã được biết đến trong lĩnh vực này.

Theo cách khác, bằng cách sử dụng các trình tự được đề xuất trong danh mục trình tự, các phân tử axit nucleic mã hoá các protein flavivirut có thể được tổng hợp bằng cách sử dụng các kỹ thuật tổng hợp axit nucleic đã biết và xen vào trong vật truyền thích hợp. Do đó, virut gây miễn dịch không độc lực được tạo ra bằng cách sử dụng các kỹ thuật thao tác tái tổ hợp đã được biết đến đối với chuyên gia trong lĩnh vực này.

Gen đích có thể được xen vào trong khung chính mà mã hoá protein cấu trúc flavivirut quan tâm đối với DEN-1, DEN-3, DEN-4 hoặc flavivirut khác. Gen flavivirut để được xen vào có thể là gen mã hoá C protein, PrM protein và/hoặc E protein. Trình tự xen vào trong khung hình dengue-2 có thể mã hoá PrM và protein E cấu trúc. Trình tự xen vào trong khung chính dengue-2 có thể mã hoá tất cả hoặc một trong số các C, prM và protein E cấu trúc.

Các virut khả thích hợp hoặc thể khả axit nucleic chứa trình tự nucleotit mã hoá protein cấu trúc của các flavivirut khác hoặc kiểu huyết thanh dengue có thể được đánh giá đối với tính hữu ích làm vaccin bằng cách sàng lọc chúng đối với các gen đánh dấu kiểu hình trên đây của việc giảm độc lực mà cho biết sự không độc lực và bằng cách sàng lọc chúng đối với tính gây miễn dịch. Tính kháng nguyên và tính gây miễn dịch có thể được đánh giá bằng cách sử dụng khả năng phản ứng in vitro hoặc in vivo với các kháng thể flavivirut hoặc huyết thanh phản ứng miễn dịch bằng cách sử dụng quy trình sàng lọc thông thường đã được biết đến đối với chuyên gia trong lĩnh vực này.

Vaccin virut dengue

Theo một số phương án, các virut khả và thể khả axit nucleic có thể tạo ra virut sống giảm độc lực hữu ích làm chất gây miễn dịch hoặc vaccin. Một số phương án bao gồm các thể khả mà thể hiện tính gây miễn dịch ở mức cao đối với virut dengue-4 trong khi tạo ra các tác dụng gây bệnh hoặc gây chết không nguy hiểm.

Để làm giảm sự xuất hiện của DHF/DSS ở đối tượng, vaccin hoá trị bốn là cần thiết để tạo ra tính gây miễn dịch đồng thời đối với tất cả bốn kiểu huyết thanh của virut. Vaccin hoá trị bốn được tạo ra bằng cách tổ hợp virut dengue-2 sống giảm độc lực của sáng chế với các thể khả dengue-2/1, dengue-2/3, và dengue-2/4 được mô tả trên đây trong chất mang dùng trong dược phẩm thích hợp để dùng làm vaccin đa hoá trị.

Các virut khả hoặc thể khả axit nucleic của sáng chế có thể bao gồm các gen cấu trúc của virut kiểu hoang dại hoặc sống giảm độc lực trong khung chính virut

DEN-2 độc lực hoặc giảm độc lực. Ví dụ, thể khảm có thể biểu hiện gen protein cấu trúc của virus DEN-4 1036 kiểu hoang dại, vaccine ứng viên của nó tạo dẫn xuất trong khung chính DEN-2.

Virus được sử dụng trong các thể khảm này được mô tả trong bản mô tả này có thể được sinh trưởng bằng cách sử dụng các kỹ thuật đã được biết đến trong lĩnh vực này. Các độ chuẩn vết tan virus sau đó được thực hiện và đếm vết tan để đánh giá khả năng sống và các đặc điểm kiểu hình của môi trường nuôi cấy sinh trưởng. Virus kiểu hoang dại có thể chuyển tiếp qua dòng tế bào nuôi cấy để tạo dẫn xuất các vật liệu khởi đầu ứng viên giảm độc lực.

Các dòng lây nhiễm khảm có thể được tạo cấu trúc từ các dòng có kiểu huyết thanh dengue khác nhau sẵn có. Việc tạo dòng các đoạn ADN bổ trợ đặc hiệu virus cũng có thể được hoàn thành, nếu muốn. Các đoạn ADN bổ trợ chứa gen protein cấu trúc hoặc protein không cấu trúc được khuếch đại bằng phản ứng chuỗi transcriptaza ngược-polymeraza (RT-PCR) từ ARN virus dengue với các đoạn mồi khác nhau. Các đoạn được khuếch đại được tạo dòng vào trong các vị trí tách của các dòng trung gian khác. Các dòng virus dengue trung gian, khảm sau đó được tạo trình tự để xác nhận độ chính xác của ADN bổ trợ đặc hiệu virus dengue được xen vào.

Plasmid khảm có chiều dài đầy đủ của hệ gen được tạo cấu trúc bằng cách xen vùng gen protein cấu trúc và/hoặc protein không cấu trúc của virus có kiểu huyết thanh dengue vào trong vật truyền có thể thu được bằng cách sử dụng các kỹ thuật tái tổ hợp đã được biết đến đối với chuyên gia trong lĩnh vực này.

Phân tích nucleotit và axit amin

Đột biến NS1-53 trong virus vaccine DEN-2 PDK-53 là đáng kể đối với kiểu hình giải độc lực của virus này, bởi vì NS1-53-Gly của virus DEN-2 16681 được bảo tồn trong gần như tất cả flavivirus, bao gồm virus do ve bét gây ra, được tạo trình tự cho đến nay. Virus vaccine DEN-4 cũng có thể chứa đột biến axit amin trong NS 1 protein ở vị trí 253. Locus này, mà là đột biến từ Gln- thành -His trong virus vaccine DEN-4 PDK-48, là Gln trong tất cả bốn kiểu huyết thanh kiểu hoang dại của virus dengue. Gốc Gln này là đơn nhất đối với virus dengue nằm trong gen flavivirus. NS1 protein là glycoprotein mà được sàng lọc từ tế bào lây nhiễm flavivirus. Nó có mặt trên bề mặt của tế bào lây nhiễm và kháng thể đặc hiệu NS1 có mặt trong huyết thanh của cá thể bị nhiễm virus. Việc bảo vệ động vật được gây miễn dịch bằng NS 1 protein hoặc theo cách thụ động với kháng thể đặc hiệu NS1 đã được báo cáo.

NS1 protein xuất hiện để tham gia vào sự sao chép ARN virus sớm.

Các đột biến mà xảy ra trong NS2A, NS2B, NS4A, và NS4B protein của các chủng giảm độc lực DEN-1, -2, -3 và -4 là bảo tồn trong tự nhiên. Các đột biến NS4A-75 và NS4A-95 lần lượt của virus vaccine DE-2 và DEN-4, xuất hiện ở các vị trí bảo tồn axit amin trong số các virus dengue, nhưng không phải trong số các flavivirus, nói chung.

NS3 protein flavivirus có ít nhất hai chức năng nhận biết: proteinase virus và ARN helicase/NTPase. NS3 protein dài 698-aa (virus DEN-2) chứa miền serine protease đầu cuối amino (bộ ba xúc tác NS3-51-His, -75-Asp, -135-Ser) mà tiếp theo các motif trình tự đối với các chức năng ARN helicase/NTPase (NS3-196-GAGKT, -284-DEAH, -459-GRIGR). Không có đột biến nào trong NS3 protein của DEN-1, DEN-2, hoặc DEN-3 virus xuất hiện trong motif được nhận biết. Đột biến NS3-510 Tyr- thành -Phe trong virus DEN-1 PDK-13 là bảo tồn. Do các virus DEN-2, -3 và -4 kiểu hoang dại chứa Phe ở vị trí này, không xảy ra rằng đột biến Tyr- thành -Phe đóng vai trò trong việc giảm độc lực virus DEN-1. Đột biến NS3-182 Glu-thành -Lys trong virus DEN-1 PDK-13 xảy ra ở vị trí mà được bảo tồn như Asp hoặc Glu ở hầu hết flavivirus do muỗi gây ra và có thể đóng một số vai trò trong việc giảm độc lực. Đột biến này được định vị 15 gốc axit amin ngược dòng của motif GAGKT helicase. Như được lưu ý trong các báo cáo trước đây, NS3-250-Glu trong virus DEN-2 16681 được bảo tồn trong tất cả các flavivirus do muỗi gây ra ngoại trừ đối với virus sốt vàng.

Các đầu dò axit nucleic lai theo cách lựa chọn với các phân tử axit nucleic mã hoá virus DEN-1, DEN-3 và DEN-4 hoặc các trình tự bổ trợ của nó. Thuật ngữ "lựa chọn" hoặc "tính lựa chọn" có nghĩa là trình tự mà không lai với các axit nucleic khác để ngăn ngừa việc phát hiện thích hợp của virus dengue. Do đó, trong việc thiết kế lai axit nucleic, tính lựa chọn sẽ phụ thuộc vào các thành phần khác có mặt trong mẫu. Việc lai axit nucleic sẽ có ít nhất 70% bổ trợ với đoạn axit nucleic mà nó lai. Như được sử dụng trong bản mô tả này để mô tả axit nucleic, thuật ngữ "lai theo cách lựa chọn" ngoại trừ việc lai theo cách ngẫu nhiên thông thường axit nucleic và do đó, có cùng nghĩa như "lai theo cách đặc hiệu." Việc lai theo cách lựa chọn axit nucleic của sáng chế có thể có ít nhất 70%, 80%, 85%, 90%, 95%, 97%, 98%, và 99% bổ trợ với đoạn trình tự mà nó lai, tốt hơn là 85% hoặc lớn hơn.

Các trình tự, đầu dò và đoạn mồi mà lai theo cách lựa chọn để mã hoá axit nucleic hoặc bổ trợ hoặc đối diện, sợi axit nucleic được dự định. Việc lai đặc hiệu với axit nucleic có thể xảy ra với các cải biến nhỏ hoặc các thay thế trong axit nucleic, miễn là khả năng lai đặc hiệu loài chức năng là được duy trì. Thuật ngữ "đầu dò" có nghĩa là trình tự axit nucleic mà có thể được sử dụng làm đầu dò hoặc đoạn mồi để lai

lựa chọn với trình tự axit nucleic bổ trợ đối với việc phát hiện hoặc khuếch đại của chúng, mà các đầu dò có thể thay đổi về chiều dài khoảng từ 5 đến 100 nucleotit hoặc tốt hơn là khoảng từ 10 đến 50 nucleotit hoặc tốt nhất là khoảng từ 18 đến 24 nucleotit.

Nếu được sử dụng làm đoạn môi, tốt hơn là được phẩm bao gồm ít nhất hai phân tử axit nucleic mà lai với các vùng khác nhau của phân tử đích sao cho để khuếch đại vùng mong muốn. Phụ thuộc vào chiều dài của đầu dò hoặc đoạn môi, vùng đích có thể nằm trong khoảng từ 70% bazơ bổ trợ đến bổ trợ đầy đủ và vẫn lai trong các điều kiện nghiêm ngặt. Ví dụ, đối với mục đích phát hiện sự có mặt của virus dengue, mức độ bổ trợ giữa việc lai axit nucleic (đầu dò hoặc đoạn môi) và trình tự mà nó lai là ít nhất đủ để phân biệt sự lai với axit nucleic từ các sinh vật khác.

Trình tự axit nucleic mã hoá virus DEN-4, DEN-3 hoặc DEN-1 (ví dụ yếu tố cấu trúc) có thể được xen vào trong vật truyền, như plasmit và biểu hiện về mặt tái tổ hợp trong sinh vật sống (ví dụ vào trong khung chính dengue-2) để tạo ra peptit virus dengue tái tổ hợp và/hoặc polypeptit và/hoặc virus.

Phương pháp phát hiện axit nucleic

Thử nghiệm di truyền nhanh mà là phép chẩn đoán đối với mỗi trong số các virus vaccin được mô tả trong bản mô tả này được đề xuất bởi sang chế. Phương án này của sang chế làm tăng phép phân tích của virus được phân lập từ huyết thanh của người được dùng vaccin mà phát triển virus huyết, cũng như làm tăng đặc điểm của virus huyết ở động vật linh trưởng không phải con người được gây miễn dịch bằng virus vaccin ứng viên.

Các trình tự này bao gồm đầu dò TaqMan chẩn đoán mà có tác dụng như báo cáo việc phát hiện amplicon ADN bổ trợ được khuếch đại từ mẫu ARN hệ gen virus bằng cách sử dụng phản ứng chuỗi transcriptaza ngược/polymeraza (RT/PCR), cũng như các amplice thuận và nghịch mà được thiết kế để khuếch đại amplicon ADN bổ trợ được mô tả dưới đây. Trong một số trường hợp, một trong số các amplice đã được thiết kế để chứa đột biến đặc hiệu virus vaccin ở đầu cuối 3' của amplice, mà thực hiện một cách hữu hiệu thử nghiệm thậm chí đặc hiệu hơn đối với chủng vaccin bởi vì sự mở rộng của đoạn môi ở vị trí đích và do đó sự khuếch tán, sẽ chỉ xảy ra nếu mẫu ARN virus chứa đột biến đặc hiệu đó.

Hệ phát hiện trình tự axit nucleic dựa trên PCR tự động có thể được sử dụng hoặc kỹ thuật đã biết khác để phát hiện axit nucleic. Thử nghiệm TaqMan là đặc hiệu ở mức cao và thử nghiệm nhạy mà cho phép hiển thị hoá thời gian thực tự động và định lượng các amplicon được tạo ra bằng PCR từ mẫu axit nucleic. TaqMan có thể xác

định sự có mặt hoặc không có mặt của trình tự đặc hiệu. Trong thử nghiệm này, đoạn mỗi thuận và nghịch được thiết kế để lần lượt ghép ngược dòng và xuôi dòng vị trí đột biến đích. Đầu dò phát hiện đặc hiệu, mà được thiết kế để có nhiệt độ nóng chảy bằng khoảng 10°C cao hơn so với của các amplime và chứa đột biến nucleotit đặc hiệu virus vaccin hoặc phần bổ trợ của nó (phụ thuộc vào sợi RT/PCR amplicon mà được phát hiện), cấu thành thành phần đoạn mỗi thứ ba của thử nghiệm này.

Đầu dò được thiết kế để phát hiện một cách đặc hiệu locus đột biến ở một trong số các hệ gen virus vaccin sẽ chứa nucleotit đặc hiệu vaccin ở giữa đầu dò. Đầu dò này sẽ dẫn đến huỳnh quang có thể phát hiện trong thử nghiệm TaqMan nếu mẫu ARN virus là đặc hiệu virus vaccin. Tuy nhiên, mẫu ARN hệ gen từ virus DEN kiểu hoang dại sẽ được mô tả một cách hữu hiệu về việc lai đầu dò do không khớp nucleotit đơn lẻ (trong trường hợp về virus DEN cha mẹ) hoặc có thể nhiều hơn một lần không khớp (như có thể xảy ra ở virus DEN kiểu dại khác) và sẽ không dẫn đến huỳnh quang đáng kể. ADN polymeraza là tương tự hơn để thể hiện đầu dò không khớp từ mẫu RT/PCR amplicon để tách đầu dò không khớp để giải phóng thuốc nhuộm thông báo (thử nghiệm TaqMan Allelic Discrimination, Applied Biosystems).

Một chiến lược để tiến hành thử nghiệm di truyền chẩn đoán sử dụng mốc hiệu phân tử.

Chiến lược mốc hiệu phân tử cũng sử dụng các đoạn mỗi để khuếch đại RT/PCR của amplicon và phát hiện trình tự đặc hiệu nằm trong amplicon bởi đầu dò chứa gen thông báo và thuốc nhuộm ngừng ở đầu cuối đầu dò. Trong thử nghiệm này, đầu dò tạo thành cấu trúc thân-vòng. Mốc hiệu phân tử dùng thuốc nhuộm ngừng và thông báo từ các thuốc nhuộm được sử dụng trong thử nghiệm TaqMan.

Dược phẩm

Các phương án trong bản mô tả này đề xuất việc dùng dược phẩm cho đối tượng ở dạng tương thích về mặt sinh học thích hợp để dùng trong dược phẩm in vivo. Thuật ngữ "dạng tương thích về mặt sinh học thích hợp để dùng in vivo" có nghĩa là dạng hoạt chất (ví dụ chất hoá học dùng trong dược phẩm, protein, gen của các phương án) để được dùng trong đó tác động độc bất kỳ có tác dụng lớn hơn tác dụng điều trị của hoạt chất. Việc dùng lượng có hoạt tính về mặt điều trị của dược phẩm điều trị được xác định là lượng hữu hiệu, ở các liều và trong khoảng thời gian cần để đạt được kết quả mong muốn. Ví dụ, lượng có hoạt tính về mặt điều trị của hợp chất có thể thay đổi theo các nhân tố như tình trạng bệnh, tuổi, giới tính và trọng

lượng của cá thể và khả năng của kháng thể gây ra đáp ứng mong muốn ở cá thể. Chế độ liều có thể được điều chỉnh để tạo ra đáp ứng điều trị tối ưu.

Theo một phương án, hợp chất (ví dụ chất hoá học dùng trong dược phẩm, protein, peptit v.v của các phương án) có thể được dùng theo cách thuận tiện ví dụ, dưới da, trong tĩnh mạch, bằng việc dùng qua đường miệng, xông, trong da, dùng qua da, dùng trong âm đạo, dùng khu trú, trong mũi hoặc dùng qua trực tràng. Phụ thuộc vào đường dùng, hoạt chất có thể được chứa trong chất đệm bảo vệ (ví dụ FTA, F127/trehaloza/albumin). Theo một phương án, dược phẩm có thể được dùng qua đường miệng. Theo phương án khác, dược phẩm có thể được dùng trong tĩnh mạch. Theo một phương án, dược phẩm có thể được dùng trong mũi, như xông. Theo phương án khác nữa, dược phẩm có thể được dùng trong da bằng cách sử dụng hệ không chứa kim (ví dụ Pharmajet®) hoặc hệ dùng trong da khác.

Dược phẩm có thể được dùng cho đối tượng trong chất mang hoặc chất pha loãng thích hợp, đồng dùng với chất ức chế enzym hoặc trong chất mang thích hợp như liposom. Thuật ngữ "chất mang dược dụng" như được sử dụng trong bản mô tả này được dự định bao gồm các chất pha loãng như nước muối hoặc dung dịch đệm trong nước. Có thể cần phải bao ngoài hợp chất bằng hoặc đồng dùng hợp chất với, vật liệu để ngăn ngừa sự bất hoạt của nó. Hoạt chất cũng có thể được dùng ngoài đường tiêu hoá hoặc trong màng bụng. Thể phân tán cũng có thể được tạo ra trong glyxerol, polyetylen glycol dạng lỏng và hỗn hợp chứa chúng và trong dầu. Trong điều kiện thông thường của việc lưu trữ và sử dụng, dược phẩm này có thể chứa chất bảo quản để ngăn ngừa sự sinh trưởng của các vi sinh vật hoặc dược phẩm làm ổn định khác (ví dụ FTA).

Dược phẩm thích hợp để sử dụng cho việc tiêm có thể được dùng bằng các phương tiện đã được biết đến trong lĩnh vực này. Ví dụ, dung dịch trong nước vô trùng (khi có thể hoà tan trong nước) hoặc thể phân tán và bột vô trùng đối với dược phẩm tùy ý chứa dung dịch vô trùng có thể tiêm hoặc thể phân tán có thể được sử dụng. Trong tất cả các trường hợp, dược phẩm có thể là vô trùng và có thể là dạng lỏng đến mức mà khả năng tiêm dễ dàng tồn tại. Nó thể là ổn định trong các điều kiện sản xuất và lưu trữ và có thể được bảo quản kháng lại tác dụng gây ô nhiễm của vi sinh vật như vi khuẩn và nấm. Chất mang dược dụng có thể là dung môi hoặc môi trường phân tán chứa, ví dụ, nước, etanol, polyol (ví dụ, glyxerol, propylen glycol, và polyetylen glycol dạng lỏng và tương tự) và các hỗn hợp thích hợp của chúng. Tính lỏng đúng đắn có thể được duy trì, ví dụ, bằng việc dùng chất bao ngoài như lecithin, bằng việc duy trì kích cỡ hạt đòi hỏi trong trường hợp thể phân tán và sử dụng chất hoạt động bề

mặt. Việc ngăn ngừa vi sinh vật có thể đạt được bằng cách gia nhiệt, phơi trần chất với chất tẩy rửa, bức xạ hoặc chất kháng vi khuẩn hoặc kháng nấm khác bổ sung.

Dung dịch vô trùng có thể tiêm có thể được tạo ra bằng cách hợp nhất hoạt chất (ví dụ hợp chất mà gây ra đáp ứng miễn dịch đối với một hoặc nhiều kiểu huyết thanh dengue) với lượng được đòi hỏi trong dung môi thích hợp với một hoặc tổ hợp của các thành phần được liệt kê trên đây, nếu cần, tiếp theo lọc vô trùng.

Khi bào chế, dung dịch sẽ được dùng theo cách tương thích với dược phẩm dạng liều và với lượng như là hữu hiệu về mặt điều trị. Dược phẩm được dùng một cách dễ dàng theo nhiều dạng liều, như loại dung dịch có thể tiêm được mô tả trên đây. Dược dự định rằng dược phẩm là thích hợp một cách đặc biệt để dùng trong cơ, dưới da, trong da, trong mũi và trong màng bụng. Tỷ lệ cụ thể có thể được phát hiện là 1:1, 1:2 hoặc tỷ lệ khác (ví dụ PFU của kiểu huyết thanh virus dengue đã nêu).

Chất có hoạt tính điều trị có thể được bào chế nằm trong tỷ lệ định trước của hỗn hợp. Liều đơn lẻ hoặc đa liều cũng có thể được dùng đối với chế độ liều thích hợp trong trường hợp đã nêu (ví dụ trước chuyến đi, khởi phát sốt dengue).

Theo phương án khác, dung dịch dùng cho mũi hoặc phun, sol khí hoặc chất xông có thể được sử dụng để phân phối hợp chất quan tâm. Dược phẩm bổ sung mà là thích hợp đối với các phương thức dùng khác bao gồm thuốc đạn và viên thuốc đặt âm đạo.

Một số dược phẩm có thể bao gồm tá dược ví dụ loại dùng trong dược phẩm của manitol, lactoza, tinh bột, magie stearat, natri sacarin, xenluloza, magie carbonat và tương tự.

Dược phẩm có thể được tạo ra với chất mang mà bảo vệ hoạt chất kháng lại sự loại bỏ nhanh ra khỏi cơ thể, như dược phẩm hoặc lớp bao ngoài giải phóng theo thời gian. Các chất mang này bao gồm dược phẩm giải phóng có kiểm soát, như, nhưng không chỉ giới hạn ở, hệ phân phối vi bao nang và polyme có thể thoái biến sinh học, tương thích sinh học, như etylen vinyl axetat, polyanhydrua, axit polyglycolic, polyorthoeste, axit polylactic và các chất khác là đã được biết đến.

Các dược phẩm được dùng với lượng và với tần số mà là hữu hiệu để ức chế hoặc làm giảm tác dụng phụ của mô ghép và/hoặc làm giảm hoặc ngăn ngừa sự thải loại. Liều chính xác và khoảng thời gian điều trị có thể được xác định theo kinh nghiệm bằng cách sử dụng phương thức thử nghiệm đã biết hoặc bằng cách thử nghiệm dược phẩm trong hệ mô hình đã biết trong lĩnh vực này và ngoại suy từ đó. Các liều cũng có thể thay đổi với mức trầm trọng của tình trạng bệnh. Dược phẩm

thường được bào chế và được dùng để gây ra tác dụng hữu hiệu về mặt điều trị trong khi giảm thiểu các tác dụng phụ không mong muốn. Nói chung, khoảng liều nằm trong khoảng từ 10^2 đến 10^6 PFU có thể được dùng ban đầu và tùy ý, tiếp theo lần dùng thứ hai trong khoảng 30 ngày hoặc đến 180 ngày sau đó, nếu cần. Theo một số phương án, đối tượng có thể nhận lần dùng kép của được hoá trị một, hai, ba hoặc bốn được bọc lộ trong bản mô tả này trong đó được phẩm là hỗn hợp được phẩm đơn lẻ hoặc có được phẩm định trước chứa các kiểu huyết thanh dengue khác nhau. Theo một số phương án, thể khảm DEN2/4 có thể có mặt với nồng độ cao hơn so với kiểu huyết thanh dengue khác như dengue-1 sống giảm độc lực.

Sẽ rõ ràng rằng, đối với đối tượng cụ thể bất kỳ, các chế độ liều đặc hiệu có thể được điều chỉnh theo thời gian theo nhu cầu của cá thể.

Theo một phương án, được phẩm được bọc lộ trong bản mô tả này có thể được dùng cho đối tượng dưới da hoặc trong da.

Được phẩm chứa virus dengue sống giảm độc lực có thể được dùng cho các thể, cụ thể là con người, ví dụ bằng việc dùng dưới da, trong cơ, trong mũi, qua đường miệng, khu trú, qua da, ngoài đường tiêu hoá, trong đường dạ dày-ruột, qua phế quản và trong phế nang. Việc dùng khu trú được hoàn thành qua kem, gel, dịch rửa được áp dụng theo cách khu trú, v.v chứa lượng hữu hiệu về mặt điều trị của chất ức chế serin proteaza. Việc dùng qua da được hoàn thành bằng việc áp dụng kem, dịch rửa, gel, v.v có khả năng cho phép các chất ức chế của serin proteaza để thẩm thấu vào da và đi vào dòng máu chung. Ngoài ra, bơm thẩm thấu có thể được sử dụng để dùng. Liều cần thiết sẽ thay đổi với tình trạng bệnh cụ thể sẽ được điều trị, phương pháp dùng và tốc độ sạch của phân tử từ cơ thể.

Theo một số phương án về phương pháp của sáng chế, đối tượng có thể là động vật có vú như con người hoặc thú y và/hoặc động vật thuần hoá hoặc thú nuôi hoặc động vật kiểu đại.

Phương pháp điều trị

Bản mô tả bọc lộ phương pháp gây ra đáp ứng miễn dịch đối với (các) kiểu huyết thanh virus dengue bằng cách sử dụng được hoá trị một, hai, ba hoặc bốn chứa các cấu trúc virus sống giảm độc lực và/hoặc khảm được dự định trong bản mô tả này.

Các phương án của sáng chế được minh hoạ hơn nữa bằng các ví dụ không hạn chế sau đây, mà không được tạo cấu trúc theo cách bất kỳ như đưa ra các giới hạn phụ thuộc vào phạm vi của nó. Trái lại, sẽ được hiểu rõ ràng rằng phương thức có thể có các phương án khác nhau khác, cải biến và dạng tương đương của nó mà, sau khi đọc

phần mô tả trong bản mô tả này, có thể gợi ý đối với chuyên gia trong lĩnh vực này mà không tách rời ý tưởng của sáng chế hoặc phạm vi của các điểm yêu cầu bảo hộ kèm theo.

Ví dụ thực hiện sáng chế

Các ví dụ sau đây được bao gồm để chứng minh một số phương án được thể hiện trong bản mô tả này. Sẽ được đánh giá cao bởi chuyên gia trong lĩnh vực này rằng các kỹ thuật được bộc lộ trong phần ví dụ thực hiện sáng chế mà tiếp theo thể hiện các kỹ thuật được phát hiện để cũng có chức năng để thực hiện sáng chế được bộc lộ và do đó có thể được xét đến để gó phần vào các phương thức được ưu tiên để thực hiện sáng chế. Tuy nhiên, chuyên gia trong lĩnh vực này sẽ, theo quan điểm của sáng chế, đánh giá cao rằng nhiều thay đổi có thể được tạo ra trong các phương án cụ thể mà được bộc lộ và vẫn thu được kết quả tương tự hoặc tương đương mà không tách rời ý tưởng và phạm vi trong bản mô tả này.

Ví dụ 1

Theo một số phương pháp lấy làm ví dụ, được phẩm được sử dụng để tạo ra như được đề cập đến trong bản mô tả này là "hạt virus chủ (MVS)" được bộc lộ. Các được phẩm này có thể thu được từ một hoặc nhiều virus dengue sống giảm độc lực, như DEN-1, DEN-2, DEN-3, và DEN-4. Theo một số phương pháp, được phẩm có thể thu được từ một hoặc nhiều virus dengue sống giảm độc lực bao gồm nhưng không chỉ giới hạn ở, các cấu trúc đặc hiệu được bộc lộ trong bản mô tả này được dùng để chỉ DENVax- 1, DENVax-2, DENVax-3, và DENVax-4. Theo các phương pháp lấy làm ví dụ khác, sáng chế đề xuất các chiến lược được sử dụng để tạo ra và mô tả đặc điểm các được phẩm này. Theo các phương án khác nữa, sáng chế đề xuất các được phẩm virus dengue hoá trị bốn và phần mô tả đặc điểm kiểu hình của các được phẩm này.

Sản xuất và phân tích virus DENVax tiền chủ

Một số quy trình được thực hiện để tạo ra hạt virus dengue tiền chủ, như sự khuếch đại theo dây và tinh chế virus dengue (ví dụ DENVax). Thứ nhất, virus DENVax được tạo dẫn xuất lại bằng việc truyền nhiễm ARN virus được sao chép từ ADN bổ trợ DENVax tái tổ hợp có chiều dài đầy đủ thành tế bào được xác nhận sản xuất (ví dụ tế bào Vero), dẫn đến hạt virus PI (lần chuyển tiếp 1). Bốn virus PI từ mỗi từ dengue-1 đến dengue-4 sau đó được khuếch đại và vết tan được tinh chế để thu được hạt vacxin P7 tiền chủ ứng viên (xem bảng 1). Một số thử nghiệm được thực hiện để phân tích sự chuyển tiếp của virus dengue. Ví dụ, việc tạo trình tự hệ gen có chiều dài đầy đủ đã chứng minh rằng tất cả bốn virus hạt P2 (lần chuyển tiếp 2) là giống về

mặt di truyền với tổ tiên tương đồng của chúng, virus vacxin ứng viên thu được từ nghiên cứu, loại dùng cho nghiên cứu. Kiểu hình vết tan ban đầu cũng giữ lại trong virus P2. Sáu virus tinh chế vết tan (P3 A-F) được phân lập đối với mỗi kiểu huyết thanh của virus dengue (ví dụ DENVax1-4) từ hạt P2 và mỗi vết tan được phân lập là vết tan được tinh chế một cách trực tiếp hai lần nữa. Lần tinh chế vết tan thứ ba (P5) của mỗi virus được khuếch đại hai lần (P6 A-F và P7 A-F) trong tế bào Vero để tạo ra hạt P7 DENVax tiền chính tiềm ẩn (bảng 1).

Bảng 1 Ví dụ về việc tạo dẫn xuất lại cGMP của virus DENVax trong tế bào WCB-Vero

Chuyển tiếp	Tạo ra/tinh chế hạt	Mô tả đặc điểm
P1	Lây truyền WCB-Vero với ARN virus được phiên mã	Chuẩn độ vết tan
P2	Khuếch đại P1 virus	Trình tự hệ gen đầy đủ
P3	Chọn 6 vết tan (A-F)/kiểu huyết thanh từ thử nghiệm vết tan P2	Tinh chế vết tan
P4	Chọn vết tan A-F từ thử nghiệm vết tan P3	Tinh chế vết tan
P5	Chọn vết tn A-F từ thử nghiệm vết tan P4	Tinh chế vết tan
P6	Khuếch đại vết tan P5 A-F	Chuẩn độ vết tan
P7	Hạt tiền chủ: khuếch đại P6 A-F	Trình tự hệ gen đầy đủ, kiểu hình vết tan TaqMAMA
P8*	MVS: khuếch đại hạt virus P7 được chọn	Mô tả đặc điểm di truyền và kiểu hình đầy đủ
P9	WVS: khuếch đại virus hạt chủ P8	Trình tự hệ gen đầy đủ, TaqMAMA
P10	BVS: khuếch đại virus hạt hoạt động P9	Trình tự hệ gen đầy đủ, TaqMAMA
* một hạt P7 tối ưu (A, B, C, D, E hoặc F) được chọn dựa vào phép phân tích di truyền và vết tan để tạo ra P8 MVS		

Một số thử nghiệm còn được thực hiện để mô tả đặc điểm hạt P7 DENVax, như phép phân tích trình tự hệ gen và kiểu hình vết tan của hạt P7 và so sánh với hạt P2 (bảng 2). Kiểu hình vết tan của virus P7 là tương tự về mặt di truyền với kiểu hình của hạt P2. Theo một số phương án lấy làm ví dụ, độ chuẩn virus được kiểm tra. Độ chuẩn virus đạt đến hơn 6,0log pfu/ml đối với hầu hết các hạt P7, ngoại trừ trong 5 phút. Việc tạo trình tự hệ gen của nhiều hơn 60 hạt virus vacxin ứng viên sau 10 lần chuyển tiếp theo dây hoặc nhiều hơn trong tế bào Vero được nhận dạng không có sự kiện nghịch

đảo ở NS1-53 và NS3-250 của ba thể quyết định giảm độc lực chính của vật truyền di truyền DENV-2 PDK-53, gợi ý rằng hai locus này là tương đối ổn định trong hạt virus vaccin ứng viên. Tất cả các sắc phổ trình tự của 24 chủng ứng viên được tạo ra từ cả hai lần tạo trình tự thuận và nghịch đối với hai vị trí này là tương đồng mà không cần có các quần thể nucleotit nhỏ bất kỳ hiển nhiên ở locus di truyền NS1 -53 và NS3-250. Ngược với các vị trí NS1 và NS3, mức nghịch chuyển khác nhau ở locus giảm độc lực 5'NCR-57 được nhận dạng từ nhiều virus vaccin loại nghiên cứu được chuyển tiếp theo đây, gợi ý rằng locus này có thể là không ổn định như NS1 và NS3 sau nhiều lần chuyển tiếp ở môi trường nuôi cấy tế bào. Do đó, thử nghiệm khuếch đại không khớp tính nhạy (TaqMAMA) được phát triển để đo một cách chính xác tốc độ nghịch chuyển ở locus 5'NCR-57 bằng RT-PCR theo thời gian thực. Theo một số nghiên cứu, tốc độ nghịch chuyển 5'NCR-57 của tất cả 24 trong số các hạt P7 được đo bằng TaqMAMA. Phụ thuộc vào nồng độ của ARN virus nguồn vào đối với mỗi virus trong thử nghiệm, giới hạn tính nhạy của TaqMAMA nằm trong khoảng từ 0,01% đến 0,07% đột biến ngược, mà là nhạy hơn nhiều so với từ 10 đến 30% giới hạn tính nhạy đột biến ngược có thể phát hiện bằng phép phân tích trình tự hệ gen liên ứng. Dữ liệu thu được minh họa rằng 15 trong số 24 virus P7 có sự đột biến ngược giảm thiểu hoặc không thể phát hiện ($< 0,07\%$), một virus (DENVax-3-D) có hầu như 100% đột biến ngược và 8 virus (1 DENVax-1, 1 DENVax-2, 2 DENVax-3, và 4 DENVax-4) có đột biến ngược riêng phần nằm trong khoảng từ 0,08% đến 12,85% (bảng 2). Việc tạo trình tự hệ gen có chiều dài đầy đủ được thực hiện đối với 16 trong số 24 virus P7 với mức đột biến ngược 5'NCR57 thấp như được đo bằng TaqMAMA. Tất cả virus được tạo trình tự duy trì hai thể quyết định giảm độc lực DENVax khác (NS1-53, NS3-250), và tất cả có các đột biến bổ sung đạt được mà không có mặt trong các dòng ADN bổ trợ ban đầu tái tổ hợp được thao tác (bảng 2). Trong một phép so sánh vaccin đích lấy làm ví dụ, DENVax-1-A, DENVax-2-F, DENVax-3-F, và DENVax-4-F là hạt tiền chủ đích đối với mỗi kiểu huyết thanh bởi vì kiểu di truyền của chúng và kiểu di truyền vết tan hầu như tương tự một cách chặt chẽ với các hạt của các thể tái tổ hợp vaccin được thiết kế ban đầu. DENVax-1-A, DENVax-2-F, và DENVax-4-F có hai đột biến không cùng nghĩa và DENVax-3-F có một. Bảng chứng này gợi ý các đột biến bổ sung này được quan sát trong bốn hạt tiền chủ này không gây ra mối quan tâm về tính an toàn hoặc các thay đổi về tính miễn dịch đối với virus. Các hạt tiền chủ này còn được khuếch đại để tạo ra MVS (hạt chủ, được ký hiệu là P7, bảng 1).

Các phương pháp lấy làm ví dụ được đề xuất trong bản mô tả này sử dụng ARN virus được phiên mã in-vitro được tinh chế từ plasmit ADN bổ trợ được tạo dòng làm

nguồn tinh khiết để truyền nhiễm tế bào Vero xác nhận vacxin để tạo ra virus vacxin. Việc tinh chế vết tan theo dây và phép phân tích trình tự hệ gen đầy đủ được hợp nhất vào trong quy trình sản xuất để bảo đảm hạt virus được sản xuất với độ tinh khiết tối ưu và độ ổn định di truyền. Sau virus tạo dòng được tạo ra là hạt tiền chủ tiềm ẩn đối với mỗi kiểu huyết thanh của DENVax. Thông qua phép phân tích hệ gen, bao gồm TaqMAMA và tạo trình tự hệ gen đầy đủ, cũng như mô tả đặc điểm của kiểu hình vết tan virus, hạt tiền chủ được chọn trước để sản xuất hạt virus chủ đối với mỗi kiểu huyết thanh (kiểu huyết thanh từ 1 đến 4). Hạt tiền chủ được chọn có các đột biến ngược không thể phát hiện (<0,01% hoặc <0,07%) ở locus 5'NCR-57, với 1 hoặc 2 thay thế axit amin trong hệ gen của chúng và giữ lại kiểu hình vết tan nhỏ được quan sát trước đây.

Bảng 2. Mô tả đặc điểm của hạt tiền chủ (P7)

Virut	Dòng ^a	TaqMAMA ^b	Log ₁₀ pfu/ml	Vết tan ^c	Đột biến được nhận dạng trong hệ gen ^d
DENVax-1	A	**	6,85	P2	NS2A-116 I-L, NS2B-92 E-D, một câm
	B	*	6,93	P2	nd ^e
	C	*	6,93	D	nd
	D	**	7,02	D	C-67 K-A; một câm
	E	0,57%	7,28	P2	nd
	F	**	7,18	P2	E473 T-M; một câm
DENVax-2	A	0,03%	6,33	P2	NS1-341 K-N
	B	*	6,33	P2	E-305 K-T, hai câm
	C	*	5,84	L	NS4A-18 T-A, bốn câm
	D	0,08%	6,20	P2	NS2B-99 I-L, một 3'NCR
	E	0,03%	6,31	P2	prM-52 K-E, NS5-412 I-V, hai câm
	F	**	6,15	P2	prM-52 K-E, NS5-412 I-V
DENVax-3	A	*	6,00	P2	NS5-200 K-N, một câm, một 3'NCR
	B	0,05%	6,27	P2	NS2A-33 I-T, NS2A-59 M-T
	C	0,30%	6,25	P2	nd
	D	100,00%	6,27	P2	nd
	E	0,31%	6,00	P2	nd
	F	**	6,30	P2	E-223 T-S, một câm
DENVax-4	A	0,47%	5,60	P2	E323 K-R/K, NS2B-21 L-F/L, NS2B-39 T-S, một câm
	B	*	5,65	D	NS2A-126 A-V; NS4A-5 N-D; NS5-383 K-R, một câm

C	4,50%	5,90	P2	nd
D	12,85%	5,97	D	nd
E	0,52%	6,85	S	prM-85 E-D, NS2B-45 T-A, NS5-320 M-T, NS5-551 E-G, hai câm
F	0,02%	6,93	S	NS2A-66 D-G, NS4A-21 A-V, bốn câm

^aVirut tạo dòng (bằng việc tinh chế vết tan theo dây) được chọn để phát triển tiếp MVS được ký hiệu nét đậm.

^{b*}: Tốc độ đột biến ngược <0,07% (giới hạn phát hiện). ^{**}: Tốc độ đột biến ngược <0,01% (giới hạn phát hiện)

^cKiểu hình vết tan: P2: tương tự với virut P2; L = lớn hơn so với virut P2, D = kích cỡ tương tự, nhưng xuất hiện tương đối khác nhau về sự rõ ràng của vết tan; S = nhỏ hơn so với P2.

^dSự thay thế khác với các dòng ADN hỗ trợ DENVax được thiết kế. Các đột biến axit amin được liệt kê với vị trí gốc của protein virut và các thay đổi (đột biến kiểu hoang dại). Tổng số các đột biến câm về gen cấu trúc và không cấu trúc của mỗi hạt là được liệt kê. Các đột biến ở vùng không mã hoá (NCR) cũng được lưu ý.

^end = không được thực hiện. Các dòng này có tốc độ đột biến ngược 5'NCR-57 cao hơn (theo TaqMAMA) so với các dòng khác, do đó được loại trừ khỏi phép phân tích trình tự hơn nữa.

Ví dụ 2

Theo một số phương pháp lấy làm ví dụ, được phẩm chứa hạt virut chủ, tạo ra hạt virut và hạt virut khối cũng như phần mô tả đặc điểm di truyền và kiểu hình của chúng được mô tả. Các được phẩm này được đề xuất để sản xuất vật liệu lâm sàng và cuối cùng là cung cấp vaccin thương mại. Việc tinh chế vết tan theo dây và phân tích trình tự hệ gen đầy đủ được hợp nhất vào trong quy trình sản xuất để bảo đảm được phẩm chứa hạt vaccin với tính an toàn tối ưu và độ ổn định di truyền để sản xuất vật liệu thử nghiệm lâm sàng.

Sản xuất và kiểm soát chất lượng sản xuất đối với MVS, WVS, và BVS

Trong một số nghiên cứu, MVS của 4 DENVax được tạo ra bằng cách khuếch đại hạt P7 tiền chủ trong tế bào Vero xác nhận. Trong các nghiên cứu khác, MVS được sử dụng để tạo ra lượng lớn WVS trong các nhà máy tế bào. Hơn nữa, vật liệu gốc BVS chứa DENVax được khuếch đại từ WVS và được bào chế thành hỗn hợp sản phẩm được chất hoá trị bốn để được sử dụng cho các thử nghiệm lâm sàng của con người. Việc kiểm soát chất lượng đối với việc giải phóng sản phẩm được thực hiện trong một số phương pháp lấy làm ví dụ, bao gồm, nhưng không chỉ giới hạn ở, việc thử nghiệm tất cả MVS, WVS, và BVS để nhận dạng, độ chuẩn lây nhiễm, tính vô sinh, mycoplasma và các chất tự sinh in vitro và in vivo. Tất cả các hạt chuyển qua thử nghiệm nhận dạng virut bằng cách sử dụng thử nghiệm RT-PCR đặc hiệu kiểu huyết thanh, mà thể hiện sự khuếch đại dương tính tương ứng với kiểu huyết thanh của nó và

âm tính đối với kiểu huyết thanh khác loại (dữ liệu không được thể hiện). Không có mycoplasma có thể phát hiện hoặc chất tự sinh được phát hiện trong vật liệu gốc MVS, WVS, hoặc BVS.

Phân tích di truyền MVS, WVS, và BVS

Trong một số phương pháp lấy làm ví dụ, sau khi tạo ra MVS từ các chủng tiền MVS (P7) được chọn trên đây được tạo ra và ARN virus tương ứng được tạo trình tự lại một lần nữa. Việc tạo trình tự hệ gen có chiều dài đầy đủ cho biết rằng MVS đối với DENVax-1 là giống với hạt tiền chủ của nó, trong khi WVS và BVS tiếp theo đạt được hai lần thay thế bổ sung ở E-483 và NS4B-108 (xem các bảng 2 và 3). Sự thay thế Ala ở E-483 thể hiện một phần kiểu di truyền trong MVS, nhưng trở nên kiểu di truyền chiếm ưu thế trong BVS. DENVax-2 và DENVax-3 là giống với hạt tiền chủ tương ứng của chúng (bảng 2 và 3). DENVax-2 MVS là giống với hạt tiền chủ của nó và WVS và BVS có hai đột biến bổ sung ở NS4A-36 và NS4B-111. Cả hai đột biến là riêng phần trong WVS và là kiểu di truyền chính trong BVS. MVS của DENVax-3 một lần nữa giống với hạt tiền chủ, nhưng WVS và BVS chứa sự thay thế bổ sung aa ở NS4A-23. DENVax-4 MVS đạt được đột biến axit amin bổ sung, ở locus NS2A-99 (kiểu di truyền hỗn hợp từ Lys thành Lys/Arg) trong quá trình sản xuất MVS (bảng 3). WVS và BVS của nó giữ lại kiểu di truyền hỗn hợp NS2A-99 Lys/Arg và BVS có kiểu huyết thanh hỗn hợp NS4B-238 Ser/Phe cộng thêm. Các kết quả trình tự liên ứng cũng xác nhận rằng MVS, WVS cũng như BV giữ lại ba thể quyết định di truyền giảm độc lực ở các locus 5'NCR-57, NS1-53, và NS3-250. Việc phân tích ít nhất locus giảm độc lực ổn định bởi TaqMAMA chứng minh rằng tốc độ đột biến ngược 5'NCR-57 nằm trong khoảng từ <0,7% đến 0,13% trong số MVS, <0,07% trong số WVS, và khoảng từ <0,07 đến 0,21% trong số BVS. 3% đột biến ngược ở locus 5'NCR-57 được xét đến là tốc độ chấp nhận được tối đa của dây vacxin (bảng 3).

Bảng 3. Sự thay thế nucleotit và axit amin trong hạt DENVax

DENVax	Nucleotit	Axit amin	Tiền chủ	MVS ^a	WVS ^a	BVS ^a
DENVax-1	2384 G-C	E-483 Gly-Ala	-	-	Gly/Ala	Ala
	3823 A-C	NS2A-116 Ile-Leu	Leu	Leu	Leu	Leu
	4407 A-T	NS2B-92 Glu-Asp	Asp	Asp	Asp	Asp
	7148 C-T	NS4B-108 Thr-Ile	-	-	Ile	Ile

	7311 A-G	câm	G	G	G	G
	TaqMAMA 5’NCR-57 đột biến ngược % ^b		--	-	-	-
DENVax-2	592 A-G	prM-52 Lys-Glu	Glu	Glu	Glu	Glu
	6481 G-C	NS4A-36 Ala-Pro	-	-	Ala/Pro	Pro
	7156 C-T	NS4B-111 Leu-Phe	-	-	Leu/Phe	Phe
	8803 A-G	NS5-412 Ile-Val	Val	Val	Val	Val
	TaqMAMA 5’NCR-57 độ biến ngược % ^b		--	-	0,07%	0,21%
DENVax-3	1603 A-T	E-223 Thr-Ser	Ser	Ser	Ser	Ser
	6436 G-A	NS4A-23 Asp-Asn	-	-	Asn	Asn
	7620 A-G	Câm	G	G	G	G
	TaqMAMA 5’NCR-57 đột biến ngược % ^b		--	-	-	-
DENVax-4	225 A-T	câm	T	T	T	T
	3674 A-G	NS2A-66 Asp-Gly	Gly	Gly	Gly	Gly
	3773 A-A/G	NS2A-99 Lys-Lys/Arg	-	Lys/Arg	Lys/Arg	Lys/Arg
	5391 C-T	câm	T	T	T	T
	6437 C-T	NS4A-21 Ala-Val	Val	Val	Val	Val
	7026 T-C	câm	T/C	T/C	T/C	T/C
	7538 C-C/T	NS4B-238 Ser-Ser/Phe	-	-	Ser/Ph e	Ser/Phe
	9750 A-C	câm	C	C	C	C
	TaqMAMA 5’NCR-57 đột biến ngược % ^b		-	0,13%	-	-
^a Nét đậm: Các thay đổi được bắt đầu ở vật liệu gốc MVS.						
^b "--" cho biết độc tố đột biến ngược < 0,01% (giới hạn phát hiện), "-" cho biết tốc độ đột biến ngược <0,07% (giới hạn phát hiện)						

Phép phân tích trình tự hệ gen đầy đủ cho biết rằng đột biến axit amin bổ sung được phát triển trong DENVax-4 MVS, trong khi ba dãy DENVax MVS khác giữ lại trình tự hệ gen liên ứng của các hạt tiền chủ của chúng. Tóm lại, từ việc nhận được hạt PI đến hạt tiền chủ (P7), chỉ có 1 hoặc 2 đột biến không đồng nghĩa xảy ra ở hạt đã nêu. Từ hạt PI đến MVS (P8), từ 2 đến 7 thay thế nucleotit được nhận dạng trong hạt DENVax đã nêu bất kỳ và chỉ có từ 2 đến 3 thay thế này dẫn đến các thay đổi axit amin. Do đó, các thay đổi nhỏ xảy ra. ARN rirut có xu hướng bị lỗi trong việc sao

chép hệ gen của chúng, do đó các thay thế di truyền trong hệ gen flavivirus trong quá trình chuyển tiếp tế bào là không được mong đợi. Không có đột biến cam trong MVS nằm trong 5' hoặc 3' NCR mà có thể tác động đến sự sao chép virus. Chỉ có sự thay đổi trong prM-52 Lys-Glu của DENVax-2, và sự thay đổi trong NS2A-66 Asp-Gly của DENVax-4 là không phải các thay đổi bảo tồn. Đột biến NS2A-66 của DENVax-4 là ở phần khung chính không cấu trúc của DENV-2 PDK-53. Mặc dù locus NS2A-66 thường là Asp trong số các chủng khác nhau DENV-2, thường là Gly đối với DENV-4. Có thể có rằng sự thay đổi từ Asp thành Gly trong DENVax-4 là thích hợp để khớp DENVax-4 trong tế bào Vero. Các gốc đột biến DENVax-2 prM-52 ở phần đầu cuối C của prM mà được tách ra khỏi hạt virus chín. Theo một số phương pháp lấy loàn ví dụ, phần mô tả đặc điểm kiểu hình được thực hiện để xác nhận rằng không có đột biến trong hạt MVS làm thay đổi một cách đáng kể kiểu hình giảm độc lực của vaccine.

Virus DENVax chứng minh độ ổn định di truyền ở mức cao trong quy trình sản xuất. Ba locus giảm độc lực DENV-2 PDK-53 được xác định được định vị ở 5'NCR, NS1-53, và NS3-250 vẫn ổn định trong trình tự hệ gen liên ứng khi chuyển tiếp theo dây DENVax từ chủng tiền chủ thành được phẩm vaccine khối lớn. TaqMAMA nhạy ở mức cao của locus 5'NCR-57 chứng minh đột biến ngược tối thiểu hoặc không thể phát hiện trong MVS, WVS (P9/Working), và BVS (hạt virus khối lớn dùng cho vaccine) của kiểu huyết thanh dengue. Tốc độ đột biến ngược 5'NCR-57 của được phẩm DENVax BVS (P10-trương đương) là thấp hơn một cách đáng kể so với tốc độ đột biến ngược 5'NCR-57 mà gọi ra ứng viên vaccine loại nghiên cứu sau 10 lần chuyển tiếp theo dây trong tế bào Vero (đột biến ngược từ 4 đến 74%). Chiến lược sản xuất trên quy mô lớn của hạt DENVax được đề xuất trong bản mô tả này dẫn đến hạt vaccine ổn định về mặt di truyền mà giữ lại các gen đánh dấu giảm độc lực trong virus vaccine ứng viên.

Kiểu hình vết tan của DENVax MVS

Trong một phương pháp lấy làm ví dụ, kiểu hình vết tan của DENVax MVS được so với virus dengue kiểu hoang dại và virus khảm loại nghiên cứu tương đồng của chúng trong tế bào Vero (Fig.2). Tất cả MVS của DENVax-1, -2, và -3 tạo ra vết tan mà nhỏ hơn một cách đáng kể so với thể tương đồng kiểu hoang dại của chúng và rất tương tự (khác nhau trong 0,4-mm) với virus loại nghiên cứu tương đồng của chúng trong tế bào Vero. DENVax-4 MVS cũng nhỏ hơn một cách đáng kể so với DENV-4 kiểu hoang dại, nhưng lớn hơn một cách đáng kể (khác nhau 0,9mm) so với thể khảm D2/4-V thu được từ phòng thí nghiệm ban đầu.

Fig.2 là biểu đồ lấy làm ví dụ minh họa kích cỡ vết tan của DENVax MVS trái với virus kiểu hoang dại đối chứng và virus ứng viên vaccin loại nghiên cứu. Đường kính vết tan trung bình (mm) $\pm$ SD (cột báo lỗi) của vết tan virus trong tế bào Vero trong lớp phủ agarosa được đo vào ngày 9 pi. Virus DEN kiểu hoang dại, được đại diện bởi cột màu đen và virus ứng viên vaccin loại nghiên cứu được công bố trước, được đại diện bởi cột màu trắng, được bao gồm để đối chứng và so sánh với hạt vaccine chủ DENVax được đại diện bởi cột màu xám.

Tính nhạy nhiệt độ của DENVax MVS

Trong phương pháp lấy làm ví dụ khác, tính nhạy nhiệt độ được thử nghiệm trong tế bào Vero đối với DENVax MVS và so với kiểu hoang dại tương đồng của chúng và virus vaccin khảm loại nghiên cứu ban đầu. DENV-3 16562 kiểu hoang dại (wt) là không nhạy với nhiệt độ. Kiểu huyết thanh virus dengue 1 kiểu hoang dại và kiểu huyết thanh virus dengue-4 là nhạy với nhiệt độ ở mức vừa phải ở 39°C (độ chuẩn bằng khoảng 1,0 log₁₀pfu/ml thấp hơn ở 39°C so với ở 37°C, Fig. 3). Kiểu huyết thanh virus dengue-2 16681 kiểu hoang dại là nhạy với nhiệt độ nhất của virus dengue kiểu hoang dại được thử nghiệm và dẫn đến giảm độ chuẩn gấp 100 lần ở 39°C. DENVax-1, -2, và -3 là nhạy với nhiệt độ như virus vaccin khảm loại nghiên cứu tương đồng ban đầu của chúng (Fig.2). Các độ chuẩn ở 39°C giảm từ 2,0 đến 3,0 log₁₀ pfu/ml đối với các chủng DENVax này. DENVax-4 cũng nhạy với nhiệt độ, chứng minh giảm độ chuẩn 5 lần. Tuy nhiên, D2/4-V loại nghiên cứu ban đầu chứng minh việc giảm độ chuẩn gấp 10 lần. DENVax-4 MVS ổn định cuối cùng chứa F 127 (và các chất khác đã biết để làm ổn định các dược phẩm này (FTA)), mà được thể hiện làm tăng cường độ ổn định nhiệt của virus dengue. Sự có mặt của F127 trong DENVax-4 MVS cũng góp phần vào tính nhạy với nhiệt độ ít rõ ràng hơn của virus trong thử nghiệm nuôi cấy Vero. Trong thử nghiệm riêng biệt, tính nhạy với nhiệt độ của chủng DENVax-4 thu được từ MSV với sự không có mặt của F 127 còn được đánh giá tiếp. Để loại bỏ F127 ra khỏi chủng, ARN virus được phân lập từ dược phẩm virus khối lớn DENVax-4 và được truyền nhiễm vào trong tế bào Vero. Virus DENVax-4 này xuất hiện để là nhạy với nhiệt độ như chủng nghiên cứu D2/4 V (độ chuẩn giảm 1,5 log₁₀ pfu/ml) vào ngày 3pi với sự không có mặt của F 127 (Fig.3).

Fig.3 minh họa biểu đồ lấy làm ví dụ minh họa tính nhạy với nhiệt độ của DENVax MVS. Virus dengue kiểu hoang dại và virus ứng viên vaccin loại nghiên cứu được công bố trước đây được bao gồm để so sánh. DENVax-4 MVS chứa F-127 bổ sung mà che phủ tính nhạy với nhiệt độ là kết quả của virus trong thử nghiệm này. Thử nghiệm riêng biệt phân tích DENVax-4 thay thế với sự không có mặt của F 127 cũng

được bao gồm. Các độ chuẩn trung bình $\pm$ SD (cột báo lỗi) của các virus được sao chép trong tế bào Vero ở 37°C hoặc 39°C.

Sự sao chép DENVax MVS trong tế bào C6/36 của muỗi

Trong một số phương pháp lấy làm ví dụ, DENVax MVS được sinh trưởng trong tế bào C6/36 để kiểm tra sự duy trì của chúng của kiểu hình giảm độc lực in vitro, với sự hiểu biết rằng virus vaccin khả loại nghiên cứu giữ lại kiểu hình giảm độc lực của khung chính virus DENV-2 PDK53 trong các tế bào của muỗi này. So với virus dengue kiểu hoang dại, DENVax- 1, DENVax-2 và DENVax-4 MVS thể hiện việc giảm sinh trưởng đáng kể (giảm ít nhất 3 log₁₀ pfu/ml) trong tế bào C6/36 vào ngày 6 pi (Fig.4). DENVax-3 MSV cũng thể hiện sự sinh trưởng giảm so với DENV-3 16562 kiểu hoang dại, nhưng việc giảm là không như được đánh dấu (giảm từ 1 đến 2log₁₀ pfu/ml). Tuy nhiên, các độ chuẩn C6/36 của dây hạt DENVax-3 là tương tự (sự khác nhau trong 1log₁₀pfu/ml) với độ chuẩn C6/36 của virus vaccin D2/3-V khả loại nghiên cứu ban đầu.

Fig.4 minh họa biểu đồ lấy làm ví dụ tạo biểu đồ sự sinh trưởng giới hạn của DENVax MVS (cột màu xám) trong tế bào C6/36 khi so với virus dengue kiểu hoang dại (cột màu đen) và virus vaccin khả loại nghiên cứu (cột màu trắng). Các độ chuẩn trung bình $\pm$ SD (cột báo lỗi) của virus được sao chép trong tế bào C6/36 6 ngày pi.

Tốc độ lây nhiễm, phân toả và truyền nhiễm ở muỗi nguyên vẹn

Trong một số phương pháp lấy làm ví dụ, tốc độ lây nhiễm và phân toả của DENVax được so với virus dengue kiểu hoang dại cha mẹ của chúng. Trong một số thử nghiệm lấy làm ví dụ, các thử nghiệm lây nhiễm qua đường miệng được thực hiện ở muỗi *Ae. aegypti*. Thức ăn máu lây nhiễm được chuẩn độ lại để đo độ chuẩn virus và chỉ có các thử nghiệm với các độ chuẩn virus tương tự trong thức ăn máu (sự khác nhau nhỏ hơn 1 log₁₀ pfu/ml) giữa các virus dengue cha mẹ và DENVax đối với mỗi kiểu huyết thanh được bao gồm để so sánh trong bảng 4. DENVax-1, DENVax-2, và D2 PDK-53-VV45R khả loại nghiên cứu không lây nhiễm muỗi qua việc ăn qua đường miệng, mà khác nhau một cách đáng kể ($p < 0,0001$) với virus cha mẹ của chúng, DENV-1 16007 (lây nhiễm 44%) và DENV-2 16681 (lây nhiễm 43,3%). Bởi vì không có muỗi nào bị lây nhiễm bởi DENVax-1 và -2, có ít đến không có sự phân toả liên quan đối với hai virus vaccin này. Trong khi DENVax-4 không lây nhiễm một số muỗi qua việc ăn qua đường miệng (2 trong số 55), tốc độ lây nhiễm là giảm một cách đáng kể ($p < 0,05$) so với virus kiểu hoang dại cha mẹ của nó, DENV-4 1036 (8 trong số 50). DENVax-3 không lây nhiễm muỗi bất kỳ trong hai thử nghiệm với độ chuẩn virus thức

ăn máu $5,2 \pm 0,02 \log_{10}$ pfu/ml (bảng 4), và trong thử nghiệm riêng biệt với độ chuẩn virus thức ăn máu $6,0 \log_{10}$ pfu/ml, chỉ có 1 trong số 30 muỗi trở nên bị nhiễm (các dữ liệu không được thể hiện trên hình vẽ). Tuy nhiên, virus dengue-3 16562 kiểu hoang dại cũng có tốc độ lây nhiễm rất thấp (8%) ở $5,2 \log_{10}$ pfu/ml, và tốc độ không gia tăng trong thử nghiệm riêng biệt với độ chuẩn virus thức ăn máu cao hơn ở $6,2 \log_{10}$ pfu/ml (3%, 1 dương tính trong số 30 muỗi, các dữ liệu không được thể hiện trên hình vẽ). Mặc dù virus dengue-3 và virus dengue-4 kiểu hoang dại có tốc độ lây nhiễm thấp một cách đáng kể so với virus dengue-1 và virus dengue-2 kiểu hoang dại, các độ chuẩn virus trung bình ở muỗi bị lây nhiễm là tương tự (từ 3,1 đến 3,9 $\log_{10}$ pfu/muỗi). Ngược lại, các độ chuẩn DENVax-4 từ hai muỗi bị lây nhiễm là tối thiểu ($0,7 \log_{10}$ pfu/muỗi), mà thấp hơn gấp 1.000 lần so với độ chuẩn từ muỗi bị lây nhiễm bởi kiểu huyết thanh virus dengue-4 1036 kiểu hoang dại ($3,9 \pm 1,5$ pfu/muỗi).

Đối với các muỗi mà được lây nhiễm, phân toả ra khỏi ruột giữa có thể được đánh giá bằng cách xác định nếu virus có mặt ở chân. Bốn DENV cha mẹ dẫn đến tốc độ phân toả nằm trong khoảng từ 36,3% đến 62,5% và độ chuẩn virus trung bình của chúng (theo $\log_{10}$ pfu) từ chân nằm trong khoảng từ $0,9 \pm 0,3$ đến $2,2 \pm 0,7$ (ngoại trừ các mẫu âm tính). Cả hai muỗi đều không lây nhiễm DENVax-4 dẫn đến sự phân toả virus đến chân (bảng 4). Trong khi virus được phân toả có thể phát hiện ở chân, không có virus nào trong số bốn virus dengue kiểu hoang dại có thể phát hiện trong nước bọt của muỗi được lây nhiễm qua đường miệng, gợi ý rằng các điều kiện ăn qua đường miệng có thể không nhạy một cách đáng kể để đo tốc độ lây truyền của các DENV. Do đó, theo các phương pháp lấy làm ví dụ khác, sự lây nhiễm muỗi nhân tạo nghiêm ngặt ở mức cao bởi sự tiêm truyền IT trực tiếp được thực hiện sau đó (bảng 4). Ngoại trừ đối với DENVax-4, tất cả các virus (kiểu hoang dại và DENVax) đạt được 100% lây nhiễm *Ae. aegypti* tiêm truyền IT. Chất tiêm truyền DENVax-4 có độ chuẩn virus thấp hơn thấp hơn không đáng kể so với ba chất tiêm truyền virus khác, nhưng nó vẫn được lây nhiễm một cách thành công 70% muỗi được tiêm truyền. Mặc dù tốc độ lây nhiễm cơ thể ở mức cao đạt được bằng việc tiêm truyền IT, tất cả bốn virus DENVax thể hiện tốc độ truyền nhiễm thấp một cách đáng kể ($p < 0,005$) hoặc không thể phát hiện (từ 0 đến 10%) so với virus dengue kiểu hoang dại (từ 43 đến 87%, Bảng 4). Virus DENVax chứng minh từ ít đến không có sự lây nhiễm và phân toả sau khi ăn qua đường miệng và các kết quả IT nghiêm ngặt ở mức cao xác nhận khả năng truyền nhiễm tối thiểu của các virus DENVax này ở *Ae. aegypti*.

Bảng 4: Tốc độ lây nhiễm, phân toả và truyền nhiễm virus ở muỗi nguyên vẹn

Virus	Ăn qua đường miệng					Tiêm truyền IT				
	Thức ăn máu ^a	Lây nhiễm ^b	Độ chuẩn cơ thể ^c		Phổ biến ^c	Tiêm tuyền	Lây nhiễm ^b	Độ chuẩn cơ thể ^c	Nước bọt ^f	
	Trung bình±SD	% (P/N)	Trung bình±SD	<i>p</i> ^d	% (P/N) ^e	pfu/liều	% (P/N)	Trung bình±SD	% (P/N)	<i>p</i> ^d
DENV-1 16007	6,6	44,0% (11/25)	3,6 ± 1,5		36,3% (4/11)	53,9	100% (30/30)	4,7 ± 0,48	43% (13/30)	
DENVax-1	6,9	0% (0/30)	NA	<0,0001	NA	67,8	100% (30/30)	3,4 ± 0,39	10% (3/30)	<0,005
DENV-2 16681	6,6	43,3% (13/30)	3,1 ± 1,5		38,5% (5/13)	67,8	100% (30/30)	5,2 ± 0,34	87% (26/30)	
D2 PDK53-VV45R	6,4	0% (0/30)	NA	<0,0001	NA	56,4	100% (30/30)	4,0 ± 0,20	0% (0/30)	<0,0001
DENVax-2	6,4	0% (0/30)	NA	<0,0001	NA	52,7	100% (30/30)	3,5 ± 0,27	7% (2/30)	<0,0001
DENV-3 16562	5,2	8% (2/25)	3,8 ± 0,2		50% (1/23)	34,0	100% (30/30)	4,2 ± 0,50	67% (20/30)	
DENVax-3	5,2 ± 0,02	0% (0/50)	NA	0,108	NA	37,3	100% (30/30)	3,3 ± 0,36	3% (1/30)	<0,0001
DENV-4 1036	5,8 ± 0,5	16% (8/50)	3,9 ± 1,5		62,5% (5/8)	69,4	100% (30/30)	5,2 ± 0,45	70% (21/30)	
DENVax-4	5,4 ± 0,4	3,6% (2/55)	0,7 ± 0,0	0,033	0% (0/2)	11,8	70% (21/30)	1,1 ± 0,46	0% (0/21)	<0,0001
^a Độ chuẩn virus hoặc độ lệch tiêu chuẩn trung bình nếu từ nhiều hơn 1 thử nghiệm trong thức ăn máu (log ₁₀ pfu/ml) theo độ chuẩn ngược ^b Tốc độ của virus được phát hiện trong cơ thể muỗi. P/N= mỗi dương tính/tổng số muỗi ^c Độ chuẩn virus trung bình ± độ lệch tiêu chuẩn (log ₁₀ pfu muỗi) trong cơ thể muỗi, chỉ có mẫu dương tính là được bao gồm để tính toán ^d Phân tích thống kê của sự khác nhau giữa DENV kiểu hoang dại và DENVax theo khả năng chiết Fisher ^e Tốc độ của virus được phát hiện ở chân của muỗi bị lây nhiễm dương tính ^f Tốc độ của virus được phát hiện trong nước bọt của muỗi bị lây nhiễm dương tính. Được sử dụng để đo hiệu quả truyền nhiễm										

Khả năng của vật truyền là thành phần an toàn quan trọng đối với virus vaccin flavivirus sống giảm độc lực. Trước đây, virus DENV-2 PDK-53-VV45R loại nghiên cứu và dẫn xuất kiểu hoang dại được thử nghiệm ở *Ae. aegypti*, và đã phát hiện ra rằng đột biến giảm độc lực NS 1-53-Asp là thể quyết định chiếm ưu thế để sao chép muỗi suy yếu. Hai locus giảm độc lực chính khác của vaccin DENV-2 PDK-53, nucleotit 5'NCR-57-T và NS3-250-Val, cũng thể hiện một số ức chế tác dụng đối với sự sao chép ở muỗi, do đó tạo ra các giới hạn bổ sung, dư thừa đối với khả năng của vật truyền muỗi. Một số phương pháp lấy làm ví dụ được mô tả trong bản mô tả này được sử dụng để thử nghiệm sự lây nhiễm qua đường miệng và IT của muỗi và sự sao chép đối với tất cả bốn chủng DENVax. DENVax-1, -2, và -3 không lây nhiễm muỗi *Ae.aegypti* bất kỳ qua việc lây nhiễm qua đường miệng (bảng 4). DENVax-4 chỉ được lây nhiễm 3,6% muỗi được phơi trần qua đường miệng, mức thấp hơn một cách đáng kể so với mức của DENV-4 kiểu hoang dại với độ chuẩn trung bình sao chép trong cơ thể muỗi thấp hơn độ chuẩn trung bình của muỗi bị lây nhiễm DENV-4 kiểu hoang dại. Ngạc nhiên là, DENVax-4 được phát hiện ở chân của muỗi bị lây nhiễm, gợi ý

rằng DENVax-4 không có khả năng phân toả từ ruột giữa của muỗi sau khi lây nhiễm qua đường miệng. Tốc độ lây nhiễm đối với DENVax-1, -2, và -4 tất cả nhỏ hơn một cách đáng kể so với bản sao kiểu hoang dại của chúng, nhưng sự khác nhau là không đáng kể giữa DENVax-3 và DENV-3 16562 kiểu hoang dại do tốc độ lây nhiễm rất chậm đối với cả hai virus. So với các chủng kiểu hoang dại khác của DENV được đánh giá ở Ae. aegypti được gom từ cùng Mae Sot Province, Thailand, chủng virus dengue kiểu hoang dại cha mẹ được sử dụng để thao tác DENVax xuất hiện có tốc độ lây nhiễm và phân toả thấp hơn bằng việc lây nhiễm qua đường miệng. DENV-1 PU0359, DENV-2 PU0218, DENV-3 PaH881/88, và DENV-4 1288 kiểu hoang dại được sử dụng để thao tác vacxin ChimeriVax-DEN dựa trên vacxin cúm vàng (YF) 17D có tốc độ lây nhiễm nằm trong khoảng từ 47 đến 77%. Ngược lại, vacxin YF 17D có thể không lây nhiễm Ae. aegypti. Mặc dù các chủng ChimeriVax chứa prM-E từ DENV kiểu hoang dại lây nhiễm ở mức cao, ChimeriVax giữ lại kiểu hình giảm độc lực muỗi của khung chính sao chép YF 17D của chúng. Các kết quả được đưa ra trong bản mô tả này cũng cho biết rằng sự giảm độc lực muỗi của khung chính DENV-2 PDK-53 được duy trì trong các chủng DENVax. Ngoài ra, bằng cách sử dụng các chủng virus dengue kiểu hoang dại với khả năng lây nhiễm muỗi thấp trong các cấu trúc được bao gồm trong được phẩm được mô tả trong bản mô tả này tạo ra dấu hiệu an toàn bổ sung.

Các kết quả lây nhiễm qua đường miệng minh hoạ rằng DENVax có khả năng lây nhiễm muỗi tối thiểu và khả năng phân toả tối thiểu. Ngoài ra, các thử nghiệm lây nhiễm IT nhạy và nghiêm ngặt hơn được thực hiện để phân tích hơn nữa khả năng của DENVax để được truyền nhiễm bởi Ae. aegypti. Các kết quả IT chứng minh rằng tất cả bốn virus DENVax có khả năng truyền nhiễm muỗi không thể phát hiện hoặc tối thiểu so với các bản sao kiểu hoang dại của chúng. Sự truyền nhiễm DENVax chỉ xảy ra theo lý thuyết nếu (1) vật truyền ăn trên vacxin với độ chuẩn virus huyết đủ để lây nhiễm ruột giữa muỗi, (2) virus có khả năng sao chép ở biểu mô ruột giữa và có khả năng phân toả tiếp theo ra khỏi ruột giữa và (3) virus phân toả có thể sao chép trong tuyến nước bọt và khạc nhổ ra virus đủ trong nước bọt để truyền nhiễm. Ngưỡng virus huyết của người được đòi hỏi để lây nhiễm muỗi vẫn chưa được thiết lập một cách thích hợp, nhưng virus huyết của người có thể nằm trong khoảng từ 10^6 đến 10^8 liều lây nhiễm muỗi đến (MID50)/ml sau khi lây nhiễm DENV kiểu hoang dại tự nhiên. MID50 này được dựa vào sự tiêm truyền IT trực tiếp của muỗi với huyết tương của người được pha loãng. Phép phân tích DENVax ở động vật linh trưởng không phải người cho biết rằng độ chuẩn virus huyết sau khi gây miễn dịch DENVax là rất thấp (nhỏ hơn $2,4 \log_{10}$ pfu/ml) và kết thúc trong từ 2 đến 7 ngày. Đưa ra mức virus huyết

thấp và khả năng lây nhiễm muỗi, phân toả và truyền nhiễm thấp của DENVax, không có khả năng rằng các virus vaccin này có thể được truyền nhiễm bởi muỗi trong tự nhiên hoặc gây ra virus huyết.

Do đó, các tác giả sáng chế đề xuất rằng bất kỳ trong số các lần chuyển tiếp của bất kỳ trong số các kiểu huyết thanh (P1-P10) có thể được sử dụng trong được phẩm để tạo ra vaccin an toàn và hiệu quả kháng lại một, hai, ba hoặc tất cả bốn kiểu huyết thanh dengue.

Độc lực thần kinh ở chuột đang bú

Virus vaccin loại nghiên cứu ban đầu được giảm độc lực ở mức cao đối với độc lực thần kinh ở chuột ICR mới sinh được nhốt nội bộ ở DVBD/CDC. Tất cả trong số các chuột này sống thử thách ic (trong não) với 10^4 pfu của mỗi virus vaccin. Kiểu huyết thanh virus dengue kiểu hoang dại-2 16681, mặt khác dẫn đến tử vong từ 62,5% đến 100% ở các chuột CDC-ICR trong các thử nghiệm khác nhau. Trong một số thử nghiệm, chuột ICR thương mại thu được từ Taconic Labs (Taconic-ICR) được sử dụng để nghiên cứu độc lực thần kinh ở chuột mới sinh. Đã quan sát được rằng chuột Taconic-ICR mới sinh là có thể nhạy hơn một cách đáng kể với sự lây nhiễm kiểu huyết thanh virus dengue-2 so với chuột CDC-ICR trước. Fig.5A tổng kết độc lực thần kinh của kiểu huyết thanh virus dengue-2 16681 kiểu hoang dại trong khuôn lạc CDC-ICR và chuột mới sinh Taconic-ICR được thử thách ic với 10^4 pfu virut. Chuột Taconic-ICR (100% tử vong ở 32 chuột, thời gian sống trung bình $8,3 \pm 0,5$ ngày) là nhạy hơn với thử thách ic kiểu huyết thanh virus dengue-2 16681 so với chuột CDC-ICR trước (91% không tránh được ở 72 chuột, thời gian sống trung bình bằng $14,6 \pm 2,3$ ngày).

Theo các phương pháp lấy làm ví dụ khác, để đánh giá độc lực thần kinh của DENVax MVS, chuột Taconic-ICR ban đầu được thử thách ic (trong não) với liều khoảng 10^4 pfu huyết thanh virus dengue-2 16681 kiểu hoang dại, D2 PDK-53 VV45R, D2/3-V, hoặc virus DENVax từ 1 đến 4 trong một ($n=16$) hoặc hai ($n=31-32$) thử nghiệm (Fig.5B). Ở liều này, virus loại nghiên cứu D2/3-V, cũng như DENVax-1, và DENVax-3 MVS thể hiện kiểu hình độc lực thần kinh giảm độc lực một cách hoàn toàn (không ốm hoặc tử vong). Như được mong đợi, kiểu huyết thanh virus dengue-2 kiểu hoang dại được phát hiện là "tiền định", với thời gian sống ở chuột trung bình (AST) bằng $8,3 \pm 0,8$ ngày. Ở chuột Taconic-ICR nhạy kiểu huyết thanh virus dengue-2, virus loại nghiên cứu D2 PDK-53-VV45R dẫn đến tử vong 81,3%. DENVax-2 MVS và DENVax-4 MVS là tiền định đồng nhất ở Taconic-ICR, thể hiện giá trị AST lần lượt bằng $9,8 \pm 1,7$, $10,2 \pm 1,4$ và $11,3 \pm 0,4$ ngày.

Trong một số phương pháp lấy làm ví dụ, độc lực thần kinh của virus khiêu huyết thanh virus dengue-2 16681 kiểu hoang dại so với độc lực của D2 PDK-53 VV45R, DENVax-2 MVS và DENVax-4 MVS, cũng như virus loại nghiên cứu D2/4-V, ở liều thấp hơn 10 lần (10^3 pfu, Fig.5C). Kiểu huyết thanh virus dengue-2 kiểu hoang dại giữ lại kiểu hình độc thần kinh tiền định đồng nhất, với AST bằng $9,0 \pm 1,4$ ngày, ở liều thử thách thấp này. 4 virus khác thể hiện kiểu hình độc lực thần kinh trung gian và mức độ độc lực thần kinh là đặc hiệu kiểu huyết thanh. Virus D2 PDK-53-VV45R và nguồn gốc DENVax-2 MVS của nó thể hiện sự giảm độc lực đáng kể (lần lượt 32,3% sống với AST bằng $13,1 \pm 3,8$ ngày và 31,2% sống với AST bằng $10,5 \pm 3,4$ ngày). DENVax-4 MVS và virus D2/4-V loại nghiên cứu giảm độc lực ở mức cao đối với độc lực thần kinh (lần lượt 81,3% sống với AST bằng $18,8 \pm 5,8$ ngày và 100% sống). Các kết quả này gợi ý rằng MVS của DENVax-1 và -3 thể hiện sự giảm độc lực đầy đủ của độc lực thần kinh, trong khi các dãy DENVax-2 và -4 MVS giữ lại kiểu hình giảm độc lực mà tương tự một cách chặt chẽ với các ứng viên vắc xin virus loại nghiên cứu tương đồng của chúng.

Các hình vẽ từ Fig.5A đến 5C là biểu đồ lấy làm ví dụ minh họa độc lực thần kinh ở chuột mới sinh được thử nghiệm với các dược phẩm khác nhau bao gồm kiểu huyết thanh virus dengue-2 kiểu hoang dại và virus dengue giảm độc lực khác nhau. Các kết quả được gom của các thử nghiệm khác nhau tổng kết độc lực thần kinh của virus kiểu huyết thanh virus dengue-2 16681 kiểu hoang dại ở CDC-ICR ($n=72$) và chuột mới sinh Taconic-ICR ($n=32$) được thử thách ic với 10^4 pfu của virus (A).

Độc lực thần kinh của DENVax MVS được thử nghiệm ở chuột Taconic-ICR với liều 10^4 pfu (B) hoặc 10^3 pfu (C). Số lượng động vật được thử nghiệm đối với mỗi nhóm trong một thử nghiệm ($n=16$) hoặc hai thử nghiệm được nhóm ($n=31$ hoặc 32) được chỉ định.

Kiểu hình vết tan của WVS, và BVS

Một số nghiên cứu được thực hiện để so sánh kiểu hình vết tan của WVS và BVS với MVS, virus dengue kiểu hoang dại và các thể khả năng loại nghiên cứu thu được từ phòng thí nghiệm tương đồng của chúng ở tế bào Vero (Fig.6). Kích cỡ vết tan trung bình được tính toán dựa trên 10 vết tan đối với mỗi vắc xin virus, nhưng dựa vào số lượng giảm của DENV-1, -3, và -4 kiểu hoang dại. Tất cả các virus MVS của DENVax-1, -2, và -3 tạo ra các vết tan mà nhỏ hơn một cách đáng kể so với các thể tương đồng kiểu hoang dại của chúng và rất tương tự (khác nhau 0,4-mm) với virus loại nghiên cứu tương đồng của chúng ở virus Vero. DENVax-4 MVS cũng nhỏ hơn một cách đáng kể so với DENV-4 kiểu hoang dại, nhưng lớn hơn không đáng kể

(0,9mm) so với thể khảm D2/4-V thu được từ phòng thí nghiệm ban đầu. Với việc ngoại trừ DENVax-2, tất cả WVS và BVS của DENVax-1, -3, -4 giữ lại kích cỡ vết tan nhỏ hơn một cách đáng kể so với kích cỡ được tạo ra từ thể tương đồng kiểu hoang dại của chúng. DENVax-2 WVS và BVS tạo ra vết tan mà tương tự với vết tan của virus DENV-2 kiểu hoang dại ở tế bào Vero, nhưng khi được thử nghiệm ở tế bào LLC-MK₂, tất cả hạt sản xuất DENVax-2 tạo ra các vết tan mà đôi khi nhỏ hơn vết tan của DENV-2 kiểu hoang dại ($1,4 \pm 0,4$) và tương tự với D2 PDK-53-VV45R ($1,0 \pm 0,3$) thu được từ phòng thí nghiệm (Fig.6).

Việc đánh giá gen đánh dấu kiểu hình của việc giảm độc lực virus, bao gồm kiểu hình vết tan nhỏ, tính nhạy với nhiệt độ, sự sao chép giảm ở tế bào muỗi, sự lây nhiễm/phân toả/truyền nhiễm giảm bởi muỗi và độc lực thần kinh giảm ở chuột ICR mới sinh, được đánh đổi với các dược phẩm chứa vật liệu gốc MVS. Các kết quả cho biết rằng tất cả DENVax giữ lại kiểu hình giảm độc lực được mong đợi tương tự với virus vacxin loại nghiên cứu ban đầu. Đưa ra cá đột biến đáp ứng với việc giảm độc lực được bảo tồn ở tất cả MVS, WVS và BV, có thể được mong đợi kiểu hình giảm độc lực để được giữ lại trong vật liệu được sản xuất để thử nghiệm lâm sàng đối với con người.

Fig.6 thể hiện biểu đồ lấy làm ví dụ minh họa kích cỡ vết tan của DENVax MVS, WVS, và BVS. Đường kính vết tan trung bình $\pm$ SD (cột báo lỗi) của vết tan virus ở tế bào Vero hoặc LLC-MK₂ trong lớp phủ agarosa được đo vào ngày 9 pi. Các DENV kiểu hoang dại và virus ứng viên vacxin loại nghiên cứu được công bố trước được bao gồm để đối chứng và so sánh.

Sao chép virus ở tế bào muỗi C6/36

Các nghiên cứu trước chứng minh rằng virus vacxin khảm dựa trên PDK-53 loại nghiên cứu giữ lại kiểu hình giảm độc lực của khung chính DENV-2 PDK53 virus ở tế bào C6/36. Trong một số phương pháp lấy làm ví dụ, DENVax MSV, WVS, và BVS được sinh trưởng trong tế bào C6/36 để kiểm tra sự duy trì của chúng của gen đánh dấu giảm độc lực in vitro sau khi sản xuất trên quy mô lớn. So với virus dengue kiểu hoang dại, ngoại trừ đối với DENVax-3, hạt được sản xuất thể hiện việc giảm sinh trưởng được đánh dấu (giảm ít nhất 3 log₁₀ PFU/ml) trong tế bào C6/36 vào ngày 6 pi (Fig.7). Hạt DENVax-3 cũng thể hiện sự sinh trưởng giảm so với DENV-3 16562 kiểu hoang dại, nhưng việc giảm là không như được đánh dấu (giảm từ 1 đến 2 log₁₀ PFU/ml). Tuy nhiên, các độ chuẩn của dãy hạt DENVax-3 là tương tự (sự khác nhau trong khoảng 1 log₁₀ PFU/ml) với virus vacxin D2/3-V khảm loại nghiên cứu ban đầu.

Fig.8 thể hiện biểu đồ lấy làm ví dụ tạo biểu đồ sinh trưởng giới hạn của DENVax MVS, WVS, và BVS trong tế bào C6/36. Các độ chuẩn trung bình $\pm$ SD (cột báo lỗi) của virus được sao chép trong tế bào C6/36 7 ngày pi. Virus dengue kiểu hoang dại và virus ứng viên vaccin loại nghiên cứu được công bố trước được bao gồm để so sánh.

Độc lực thần kinh ở chuột đang bú

Các thử nghiệm bổ sung được thực hiện để phân tích độc lực thần kinh ở chuột ICR mới sinh. Ở liều trong sọ bằng 10^4 PFU, tỷ lệ sống đối với DENV-2 16681 kiểu hoang dại và D2 PDK-53-VV45R lần lượt bằng 0% và 18,8% (Fig.9A) ở chuột ICR, nhưng bằng khoảng 20% đối với DENV-2 16681 kiểu hoang dại và 100% đối với D2 PDK-53-VV45R ở chuột CDC ICR. Trong nghiên cứu này, DENVax-1 và DENVax-3 MVS được giảm độc lực (100% sống) đối với chuột ở liều 10^4 PFU, nhưng MVS của DENVax-2 và DENVax-4 gây tử vong 100% ở liều quá 10^4 PFU (Fig.5A). Tuy nhiên, khi được thử nghiệm ở liều 10^3 PFU virus, DENVax-2 (31,3% sống) và DENVax-4 (81,3% sống) thể hiện độc lực thần kinh giảm so với kiểu huyết thanh virus dengue-2 16681 kiểu hoang dại (0% sống) và tỷ lệ sống của chúng là tương tự với tỷ lệ sống của ứng viên vaccin loại nghiên cứu lần lượt D2 PKD-53-VV45R (32,3%) và D2/4-V (100%) (Fig.9B). Mặc dù, DENV-1, -3, hoặc -4 kiểu hoang dại không được bao gồm để so sánh trong nghiên cứu này, hoạt động trước chứng minh rằng DENV-1 16007 kiểu hoang dại được giảm độc lực ở chuột CDC-ICR qua đường ic, trong khi DENV-3 16562 và DENV-4 1036 kiểu hoang dại độc lực ở mức cao (0% sống) đối với chuột CDC-ICR. Cũng thích hợp rằng 3 DENV kiểu hoang dại sẽ thể hiện độc lực tương tự hoặc lớn hơn ở chuột Taconic ICR nhạy cảm hơn. Do đó, việc bao gồm virus dengue kiểu hoang dại để so sánh với DENVax MVS tương đồng của chúng được xét đến là không thông tin. Nghiên cứu này cho biết rằng tất cả 4 DENVax MVS và virus vaccin ứng viên thu được từ phòng thí nghiệm ban đầu thể hiện kiểu hình giảm độc lực chuột có thể so sánh so với DENV-2 16681 kiểu hoang dại.

Các hình vẽ từ Fig.9A đến Fig.9B thể hiện biểu đồ lấy làm ví dụ của các dữ liệu về độc lực thần kinh của DENVax MVS ở chuột ICR mới sinh. (A) tiêm truyền IC của virus ở liều 10^4 PFU. (B) tiêm truyền IC của virus ở liều 10^3 PFU.

Tất cả các dãy hạt của DENVax được thử nghiệm đối với sự giống nhau, tính vô sinh và ngẫu nhiên từ các chất không mong muốn. Phép phân tích trình tự hệ gen đầy đủ thể hiện rằng một đột biến axit amin bổ sung được gọi ra ở DENVax-4 MVS, trong khi 3 DENVax MVS khác giữ lại trình tự hệ gen liên ứng của hạt tiền chủ của chúng. Trong dãy WVS, DENVax-3 đạt được đột biến axit amin bổ sung và 3 kiểu

huyết thanh khác tích tụ 2 thay thế axit amin bổ sung, so với hạt tiền chủ của chúng. Trình tự hệ gen của tất cả 4 dây BVS là giống với các dây WVS của chúng. Tổng thể từ hạt P2 đến hạt tiền chủ (P7), chỉ có 1 hoặc 2 đột biến không phải cam xuất hiện trong hạt đã nêu. Giữa các hạt tiền chủ và BCS (P10), chỉ có từ 1 đến 2 thay thế nucleotit được quan sát, tất cả trong số chúng xảy ra ở NS2A, 4A, hoặc 4B, với việc ngoại trừ sự thay đổi nucleotit đơn lẻ dẫn đến glyxin và alanin được bảo tồn ở gốc E-483. Các hạt từ P2 đến BVS (P10), tổng cộng có từ 3 đến 8 thay thế nucleotit được nhận dạng trong hạt DENVax đã nêu bất kỳ và chỉ có từ 2 đến 4 thay thế này dẫn đến thay đổi axit amin. Không có đột biến cam nào trong BVS là nằm trong vùng 5'- hoặc 3'-NCR mà có thể tác động đến việc sao chép virus. Các kết quả này gợi ý rằng virus DENVax là ổn định ở mức cao về mặt di truyền trong khi sản xuất. Ba locus giảm độc lực DENV-2 PDK-53 được xác định được định vị ở 5'NCR, NS 1-53, và NS3-250 còn lại không đổi trong trình tự hệ gen liên ứng khi chuyển tiếp theo dây DENVax để tạo ra vật liệu gốc BVS. TaqMAMA nhạy ở mức cao của locus 5'-NCR-57 thể hiện sự đột biến ngược tối thiểu hoặc không thể phát hiện ở MVS, WVS, và BVS của DENVax. Tốc độ đột biến ngược cao nhất bằng 0,21% được nhận dạng ở DENVax-2 BVS. Tốc độ đột biến ngược của P10-BVS tương đương (từ <0,07% đến 0,21%) là thấp hơn một cách đáng kể so với tốc độ đột biến ngược mà được gợi ra trong các ứng viên vaccin khác sau khi chuyển tiếp theo dây vào tế bào Vero (đột biến ngược từ 4 đến 74% theo P10). Kết quả này gợi ý rằng chiến lược này để sản xuất trên quy mô lớn của hạt DENVax là thành công, xét đến việc duy trì độ ổn định di truyền và sự duy trì của gen đánh dấu giảm độc lực trong virus vaccin ứng viên.

Do vật liệu gốc MVS được bộc lộ trong bản mô tả này sẽ được sử dụng để sản xuất hơn nữa các dây WVS và BVS, panen đầy đủ về việc đánh giá kiểu hình giảm độc lực virus, bao gồm kiểu hình vết tan nhỏ, tính nhạy với nhiệt độ, sao chép giảm ở tế bào muỗi, lây nhiễm/phân toả/truyền nhiễm ở muỗi nguyên vẹn và độc lực thần kinh giảm ở chuột ICR mới sinh, được thực hiện đối với tất cả MVS hoặc vật liệu gốc thay thế tương đương của chúng. Đối với vật liệu gốc WVS và BVS, kích cỡ vết tan, khả năng lây nhiễm ở tế bào muỗi, cũng được thực hiện để xác nhận sự giảm độc lực của chúng. Các kết quả cho biết rằng tất cả các vật liệu gốc MVS của 4 kiểu huyết thanh của DENVax giữ lại kiểu hình giảm độc lực mong đợi, như vết tan nhỏ, sự sao chép giảm ở tế bào C6/36 và độc lực thần kinh giảm ở chuột, tương tự với virus vaccin thu được từ phòng thí nghiệm ban đầu (Fig.6, Fig.8, và Fig.9). Ngoại trừ đối với DENVax-4, tất cả 3 vật liệu gốc MVS của DENVax là TS ở 39°C như được thể hiện trên các Fig.3 và Fig.7.

Đối với các vật liệu gốc WVS và BVS, hai kiểu hình giảm độc lực, vết tan nhỏ và sao chép giới hạn ở tế bào C6/36, được phân tích và xác nhận. Do có các thay đổi di truyền rất nhỏ giữa MVS và BVS, sẽ được mong đợi rằng chúng sẽ giữ lại kiểu hình giảm độc lực như MVS. Ngoài các thử nghiệm được mô tả trong báo cáo này, tính an toàn và tính gây miễn dịch của DENVax được sản xuất ở chuột Ag129 và động vật linh trưởng không phải người đã được thử nghiệm.

Các phương pháp lấy làm ví dụ được đề xuất trong bản mô tả này để chứng minh việc sản xuất vật liệu gốc DENVax MVS, WVS, và BVS trong cGMP. Vật liệu gốc BVS được sử dụng để bào chế DENVax hoá trị bốn trong các đánh giá thử nghiệm lâm sàng của người. Chiến lược sản xuất đơn nhất để tối ưu hoá độ ổn định di truyền và tính an toàn của MVS được sản xuất được đề xuất trong một số phương pháp lấy làm ví dụ. Do các locus giảm độc lực chính của DENVax đã được mô tả đặc điểm trước và thử nghiệm SNP nhạy ở mức cao và định lượng, TaqMAMA được phát triển để tích hợp trình tự hệ gen và TaqMAMA để nhận dạng hạt tiền chủ tối ưu để sản xuất MVS. Các đặc điểm di truyền và kiểu hình của MVS được phân tích một cách đầy đủ để xác nhận rằng các virus này giữ lại việc giảm độc lực mong muốn đối với tính an toàn của vaccin. Điều này chỉ có thể là vaccin virus sống giảm độc lực mà có thể được phân tích một cách hữu hiệu đối với tất cả các locus di truyền giảm độc lực chính trong khi sản xuất từ tất cả các cách tiền chính đến vật liệu gốc BVS. Các kết quả được đề xuất trong bản mô tả này được minh hoạ là có lợi của vaccin sống giảm độc lực được thiết kế theo chiến lược về tính an toàn vaccin.

Fig.10 thể hiện bảng lấy làm ví dụ so sánh các virus sống giảm độc lực mới với virus dengue sống giảm độc lực được tạo ra trước. Các đột biến được chỉ định khi khác với virus đối chứng (ví dụ 16681), hoặc virus dengue-2 sống giảm độc lực khác.

Vật liệu và phương pháp

Virut và tế bào

DENV-1 16007, DENV-2 16681, DENV-3 16562, và DENV-4 1034 có tác dụng làm đối chứng DENV kiểu hoang dại và chúng là virus kiểu di truyền cha mẹ đối với bốn ứng viên vaccin DEVax tái tổ hợp. Các virus loại nghiên cứu tổ tiên DENVax, được ký hiệu là D2/1-V, D2 PDK-53-VV45R, D2/3-V, và D2/4-V, được tạo ra và được mô tả đặc điểm trước đây. Tế bào Vero (thận của khỉ xanh châu Phi) được sử dụng để sản xuất ngân hàng tế bào chủ và hoạt động để sản xuất vaccin có nguồn gốc từ dòng tế bào CCL81 của Bộ gom giống chuẩn Mỹ (ATCC) mà đã được mô tả đặc điểm bởi Tổ chức y tế thế giới (WHO) để sản xuất vaccin (tế bào WCB-Vero).

Tạo dẫn xuất virus DENVax sống tái tổ hợp từ các dòng ADN tái tổ hợp

Để tạo dẫn xuất lại virus vaccine ứng viên trong các điều kiện sản xuất cGMP, các dòng ADN hỗ trợ lây nhiễm DENV được thao tác trước đây, pD2-PDK-53-W45R, pD2/1-V, pD2/4-V, và pD2/3-V thất in vitro chứa ADN hỗ trợ virus hệ gen chiều dài đầy đủ được sử dụng để tạo ra các bản sao ARN virus bằng việc phiên mã in vitro như được mô tả trước đây. Tóm lại, ADN hỗ trợ hệ gen DENV tuyến tính XbaI được xử lý bằng proteinase K, được chiết bằng phenol/chloroform và tạo kết tủa trong ethanol để loại bỏ protein gốc bất kỳ và sau đó tạo huyền phù trong chất đệm Tris-EDTA không chứa Rnaza trước khi phiên mã. Sự phiên mã in vitro được thực hiện bằng cách sử dụng kit phiên mã hiệu suất cao AmpliScribe T7 (Epicentre Technologies) theo phương thức được khuyến cáo của nhà sản xuất. Thẻ tương tự nắp ARN A, m7G(5')ppp(5')A (New England BioLabs), được hợp nhất trong thời gian phản ứng phiên mã 2 giờ để bổ sung nắp A đầu cuối 5' thành bản sao ARN. Sau đó, các mẫu được xử lý bằng DNase I để tiêu hóa ADN hỗ trợ mẫu, tiếp theo chiết phenol/chloroform độ pH thấp và tạo kết tủa ethanol để loại bỏ ADN dư và protein. Các bản sao ARN tinh chế, tạo huyền phù trong nước không chứa Rnaza, được phân bố trong 20-µl dịch lỏng và lưu trữ ở -80°C cho đến khi sẵn sàng để truyền nhiễm tế bào. Tính nguyên vẹn và nồng độ của các bản sao ARN được phân tích bằng điện di gel agarose. Mỗi 20-µl dịch lỏng được đánh giá để chứa ARN virus hệ gen có chiều dài đầy đủ để cho phép truyền nhiễm từ 0,4 đến 1×10^7 tế bào Vero xác nhận sản xuất bằng việc bắn điện.

Việc truyền nhiễm của mỗi bản sao ARN vào trong tế bào WCB-Vero được thực hiện trong phòng cGMP ở Shantha Biotechnics. Các bản sao ARN DENVax được tan băng, trộn với 400µl huyền phù tế bào Vero (1×10^7 tế bào/ml), và được chuyển vào chậu thủy tinh bắn điện vô trùng được làm mát sơ bộ (khe hở 4-mm) để bắn điện bằng hệ tổng Gene Pulser Xcell (BioRad Laboratories). Mỗi mẫu được tạo xung một lần ở 250V/∞ Ohms/500µP, ủ trong từ 10 đến 15 phút ở nhiệt độ trong phòng, chuyển vào bình thót cổ 75-cm² chứa 30ml môi trường sinh trưởng tế bào (MEM với 10% FBS), và ủ ở 36°C ± 1°C, 5% CO₂ trong từ 6 đến 11 ngày. Môi trường nuôi cấy được thu hoạch, gạn bằng việc ly tâm, làm ổn định và lưu trữ trong dịch lỏng nhỏ dưới -60°C. Độ chuẩn virus của vật liệu lưu trữ vaccine ứng viên (được gọi là PI đối với mức chuyển tiếp 1) thu được từ việc lây truyền được xác định bằng thử nghiệm chuẩn độ vết tan trong tế bào Vero và được sử dụng để nhân giống hơn nữa hạt DENVax.

Sản xuất hạt virus DENVax

Hạt virus PI được sử dụng để nhân giống dây hạt virus tiền chủ, chủ, thao tác và khối lớn qua chiến lược được thiết kế để bảo đảm độ ổn định di truyền tối ưu và tính

an toàn dây sản xuất. Chiến lược này bao gồm ba lần tinh chế vết tan theo dây, cũng như phép phân tích di truyền virus ở các mức chuyển tiếp khác nhau để chọn quần thể virus đơn dòng tối ưu để sản xuất hạt tiếp tục (bảng 1). Tóm lại, hạt PI được thu hoạch từ tế bào được lây nhiễm được khuếch đại khi bằng việc lây nhiễm tế bào Vero ở MOI bằng 0,001 để tạo ra hạt P2. Dịch lỏng chứa vật liệu gốc hạt P2 được đánh giá bằng hình thái học vết tan và tạo trình tự hệ gen virus đầy đủ. Vật liệu gốc P2 xác nhận về mặt di truyền được phủ lên lớp đơn tế bào Vero bằng môi trường phủ như được mô tả trong phần chuẩn độ vết tan dưới đây để tạo ra các vết tan được phân lập. Sau khi hiển thị hoá sáu vết tan đơn lẻ màu đỏ trung tính từ mỗi trong số 4 kiểu huyết thanh của virus vacxin được phân lập (các dòng vết tan từ A đến F) và được trộn vào trong 0,5ml môi trường nuôi cấy (lần chuyển tiếp P3). Mỗi trong số sáu huyền phù vết tan được cho vào hai vòng tinh chế vết tan bổ sung, dẫn đến hạt virus được tinh chế vết tan hai lần và ba lần lần lượt ở các lần chuyển tiếp P4 và P5. Virus P5 được khuếch đại qua hai lần chuyển tiếp Vero tiếp theo để tạo ra vật liệu gốc hạt P7.

Phép phân tích di truyền của ba locus giảm độc lực DENVax chính bằng cách sử dụng đồng tạo trình tự và/hoặc thử nghiệm đột biến khuếch đại không khớp dựa trên Taqman (TaqMAMA) như được bộc lộ trước đây và phép phân tích kiểu hình vết tan được thực hiện để sàng lọc tất cả 24 hạt P7. Hạt có các đặc điểm ban đầu thích hợp sau đó được mô tả đặc điểm tiếp bằng việc tạo trình tự hệ gen đầy đủ. Đối với kết quả của các phép phân tích này, một trong số 6 (dòng từ A đến F) hạt P7 của mỗi kiểu huyết thanh DENVax được chọn để là hạt tiền chủ, dựa vào sự có mặt của các đột biến giảm độc lực DENV-2 PDK-53, các thay đổi trình tự hệ gen tối thiểu và kiểu hình vết tan mong đợi. Mỗi hạt tiền chủ được chọn được mở rộng đến hạt virus chủ (MVS hoặc P8) bằng việc chuyển tiếp một lần của virus ở MOI bằng 0,001 trong nhiều bình thót cổ 175cm² chứa tế bào Vero. Ngoại trừ đối với DENVax-4 MVS, hạt virus chủ được thu hoạch ở từ 8 đến 10 ngày sau khi lây nhiễm (pi). Vật liệu gốc MVS được thu hoạch ở từ 6 đến 10 ngày sau lây nhiễm (pi), gạn bằng việc ly tâm, làm ổn định bằng việc thêm dung dịch sucroza/phosphat/glutamat (nồng độ cuối cùng lần lượt 7,5% sucroza, 3,4mM kali dihydro phosphat, 7,2mM dikali hydro phosphat, 5,4mM mononatri glutamat) và từ 0,95 đến 1,90% FBS (nồng độ cuối cùng). DENVax-4 MVS được tạo ra theo cách khác nhau để tối ưu hoá hiệu suất của nó. Tóm lại, nhiều bình thót cổ chứa tế bào được lây nhiễm với hạt tiền chủ DENVax-4 ở MOI bằng 0,001 với sự có mặt của 0,1% F-127TM, poloxamer 407, (copolyme khối EO-PO khác đã được thử nghiệm và có thể thay thế trong bản mô tả này, xem patent đã được cấp) mà đã được chứng minh để tăng cường độ ổn định nhiệt virus DENV. Môi trường lây nhiễm được

thu hoạch các ngày từ 6 đến 10 pi, và làm ổn định bằng 17% FBS (nồng độ cuối cùng), gom và làm đông lạnh. Tất cả bốn vật liệu gốc DENVax MVS được lưu trữ như 1-ml dịch lỏng dưới -60°C .

Hạt virus thao tác DENVax (WVS) được tạo ra bằng việc chuyển tiếp một lần trong môi trường nuôi cấy tế bào Vero của MVS ở MOI bằng 0,001. Các quy trình là tương tự với việc sản xuất MVS, ngoại trừ chúng được nuôi cấy trong các nhà máy tế bào nhiều lớp (6360cm^2). Vật liệu gốc WVS được lọc qua bộ lọc $10\mu\text{M}$ và $0,45\mu\text{M}$, làm ổn định bằng cùng chất làm ổn định được sử dụng đối với MVS, tạo dịch lỏng vào trong chai 30ml PETG hoặc 2,0ml lọ nhỏ cryo và lưu trữ dưới -60°C .

Trong một số phương pháp, hạt virus khối lớn (BVS) được tạo ra bằng việc lây nhiễm nhiều nhà máy tế bào (mỗi nhà máy 6360cm^2) tế bào Vero hợp dòng với 90mL WVS được pha loãng để đạt được MOI bằng 0,001. Môi trường được sử dụng để pha loãng chất tiêm truyền WVS chứa 0,1% F-127TM mà không chứa huyết thanh. Sau 1,5 giờ hấp phụ, tế bào được rửa ba lần bằng PBS, và 800ml môi trường DMEM không chứa huyết thanh được bổ sung vào mỗi nhà máy tế bào và các nhà máy được ủ ở $36(\pm 1)^{\circ}\text{C}$ trong $5(\pm 0,5)\%$ CO_2 . Sau khi ủ trong bốn ngày, dịch lỏng nhỏ chứa môi trường được gom để thử nghiệm tính vô sinh. Virus được thu hoạch trong khoảng từ ngày 5 đến ngày 10 pi, và ngay sau đó gạn bằng việc lọc qua bộ lọc kích cỡ xốp $0,45\mu\text{m}$ và 1L mỗi nhóm virus được gạn được làm ổn định bằng việc bổ sung 500ml 3x chất đệm FTA (nồng độ cuối cùng 15% trehaloza, 1,0% Pluronic[®] F-127TM poloxamer 407, 0,1% albumin USP của người trong PBS, độ pH=7,4). Virus được làm ổn định được phân bố vào trong chai PETG 1-L và lưu trữ đông lạnh dưới -60°C để nhóm tiếp theo và thử nghiệm kiểm soát chất lượng. Tất cả các lần thu hoạch virus ổn định với độ chuẩn virus lớn hơn 10^5 PFU/ml và hàm lượng có thể chấp nhận của ADN gốc được làm tan bằng một cách nhanh chóng trong bể nước ở 32°C , sau đó nhóm theo cách vô trùng và trộn. Mỗi BVS hoá trị một được nhóm được phân bố vào trong vật chứa PETG được đánh dấu và được lưu trữ ở dưới -60°C cho đến khi sử dụng tiếp.

Kiểm soát chất lượng sản phẩm sản xuất

Hạt MVS, WVS, và BVS được thử nghiệm để nhận dạng, tính vô trùng và chất tự sinh có thể phát hiện. Việc nhận dạng của mỗi vật liệu gốc vacxin được xác nhận bằng RT-PCR với đoạn môi đặc hiệu kiểu huyết thanh DENVax. Các đoạn ADN hỗ trợ được khuếch đại chứa vị trí kết hợp khảm E/NS1 để cho phép nhận dạng mỗi trong số bốn kiểu huyết thanh DENVax. Mỗi hạt được thử nghiệm trong tất cả bốn phản ứng RT-PCR đặc hiệu kiểu huyết thanh để xác nhận việc nhận dạng virus và ngẫu nhiên từ sự ô nhiễm ngang với kiểu huyết thanh DENVax khác loại. Thử nghiệm vô trùng được

thực hiện theo USP 71 (United States Pharmacopeia, section 71). Thử nghiệm mycoplasma được thực hiện.

Các thử nghiệm in vitro và in vivo sau đây đối với sự ô nhiễm virus tất cả được thực hiện bằng cách sử dụng các lần thu hoạch DENVax không gạn không ổn định được gom trong quá trình sản xuất hạt. Môi trường lây nhiễm được thu hoạch thứ nhất được làm trung hoà bằng kháng huyết thanh đa dòng của thử DENV (Inviragen) ở $36 \pm 1^\circ\text{C}$ trong 1 giờ để làm bất hoạt DENV. Đối với thử nghiệm in vitro, hạt được làm trung hoà được tiêm truyền vào trong ba dòng tế bào chỉ định, MRC5, VERO và MA104, trong các bình thót cổ 25cm^2 . Mỗi virus (đối chứng CPE) hoặc virus quai bị (đối chứng hấp phụ hồng cầu lên tế bào) lần lượt được sử dụng làm đối chứng CPE dương tính hoặc hấp phụ lên tế bào. Tất cả các tế bào được kiểm tra hàng ngày đối với CPE trong tổng số 14 ngày. Khi kết thúc 14 ngày, dịch nuôi cấy được loại bỏ và thay thế bằng 10mL dung dịch tế bào hồng huyết cầu chuột lang (RBC) (3mL 0,5% RBC chuột lang trong nước muối đệm phosphat, được tạo thành đến 10mL với môi trường sinh trưởng tế bào). Sau đó, bình thót cổ được ủ ở $5 \pm 3^\circ\text{C}$ trong 30 phút tiếp theo ủ ở nhiệt độ trong phòng trong 30 phút. Rửa lớp đơn bằng PBS và quan sát dưới độ phóng đại 10 lần đối với sự có mặt của nhóm hình sao bất kỳ của RBC đối với sự hấp phụ lên tế bào.

Các thử nghiệm In vivo đối với các chất tự sinh được thực hiện ở chuột đang bú, chuột sau thôi bú và chuột lang. Chuột đang bú được tiêm truyền với 0,1ml hoặc 0,01ml (10 chuột trong mỗi nhóm liều) của mẫu hạt làm trung hoà kháng huyết thanh DENV qua việc thêm trong màng bụng (ip). Tương tự, 10 chuột sau thôi bú, mỗi chuột được tiêm truyền ip với 0,5ml hoặc 0,03ml mẫu. Mỗi chuột lang (5/nhóm) được tiêm truyền ip với 5,0mL. Chuột đang bú được quan sát hàng ngày đối với sự mắc bệnh và sự tử vong trong tổng cộng 14 ngày sau tiêm truyền. Chuột sau thôi bú được quan sát trong tổng cộng 28 ngày và chuột lang được quan sát trong tổng cộng 42 ngày sau tiêm truyền. Các vật dụng thử nghiệm đáp ứng các tiêu chuẩn chấp nhận nếu >80% động vật được tiêm truyền vẫn khoẻ mạnh trong suốt thời gian quan sát.

Thử nghiệm in vivo đối với chất gây ô nhiễm cũng được thử nghiệm ở trứng gà kết phôi và được thực hiện. Đối với mỗi mẫu, 10 trứng gà kết phôi (9 ngày tuổi) được tiêm truyền với 0,5mL mẫu làm trung hoà kháng huyết thanh DENV vào trong dịch túi niệu và ủ ở 35°C trong 3 ngày. Dịch túi niệu từ 10 trứng này được thu hoạch, nhóm và chuyển tiếp vào trong dịch túi niệu của 10 trứng kết phôi mới tạo ra (từ 10 đến 11 ngày tuổi; 0,5mL/trứng) và ủ ở 35°C trong 3 ngày nữa. Tương tự, đối với mỗi mẫu, 10 trứng kết phôi (từ 6 đến 7 ngày tuổi) được tiêm truyền với 0,5mL/trứng (BVS hoá trị

một DENVax-2) hoặc 0,25mL/trứng (DENVax-1, DENVax-3 và DENVax-4 BVS) bằng việc tiêm vào trong lòng đỏ trứng và ủ ở 35°C trong 9 ngày. Lòng đỏ trứng từ 10 trứng được thu hoạch và gom và 10% huyền phù được chuyển tiếp vào trong lòng đỏ trứng của 10 trứng kết phôi mới tạo ra (từ 6 đến 7 ngày tuổi; 0,5mL/trứng) và ủ ở 35°C trong 9 ngày nữa. Trứng được tiêm truyền vào trong dịch túi niệu (lần tiêm truyền ban đầu và chuyển tiếp) được quan sát đối với khả năng sống sau 3 ngày tiêm truyền. Cả hai nhóm dịch túi niệu được thử nghiệm đối với hoạt tính hemagglutinin hoá bằng cách sử dụng gà, chuột lang và hồng cầu typ O của người ở 4°C và 25°C. Trứng được tiêm truyền vào trong lòng đỏ trứng (tiêm truyền ban đầu và chuyển tiếp) được quan sát đối với khả năng sống sau 9 ngày ủ.

Thử nghiệm vết tan virus và thử nghiệm tiêu điểm miễn dịch

Các độ chuẩn virus được đo bằng thử nghiệm vết tan hoặc thử nghiệm tiêu điểm miễn dịch bằng cách sử dụng tế bào Vero. Thử nghiệm vết tan được thực hiện trong lớp phủ agarosa kép trong đĩa sáu giếng của tế bào Vero hợp dòng như được mô tả trước đây và chúng cũng được sử dụng để đánh giá kiểu hình vết tan của hạt DENVax. Để so sánh chính xác, kích cỡ vết tan của tất cả virus được đo và được so sánh trong cùng thử nghiệm. Sau khi hiển thị hoá với màu đỏ trung hoà vào ngày 9 pi, đến tối đa 10 vết tan được phân lập đối với mỗi virus được đo để tính toán kích cỡ vết tan trung bình. Vết tan ít hơn được đo đối với DENV-1, -3, và -4 kiểu hoang dại, kích cỡ vết tan lớn hơn của nó thường không cho phép đo 10 vết tan riêng biệt.

Bởi vì DENVax hoá trị bốn chứa tất cả bốn kiểu huyết thanh DENV, thử nghiệm tiêu điểm miễn dịch đặc hiệu kiểu huyết thanh DENV được thực hiện để định lượng mỗi thành phần DENVax trong được phẩm hoá trị bốn. Các thử nghiệm tiêu điểm miễn dịch của mỗi DENVax MVS riêng rẽ so với thử nghiệm vết tan để bảo đảm các kết quả chuẩn độ virus có thể so sánh giữa hai thử nghiệm. Thử nghiệm tiêu điểm miễn dịch được thực hiện trong đĩa 6 giếng chứa tế bào Vero hợp dòng được lây nhiễm với virus pha loãng theo dãy. Tế bào được phủ với môi trường muối cân bằng (môi trường BSS/YE-LAH) chứa 0,7% carboxymethyl xenluloza có độ nhớt ở mức cao (Sigma) và ủ trong 7 ngày ở 37°C với 5% CO₂. Sau khi loại bỏ lớp phủ, rửa tám tế bào 3 lần bằng PBS, cố định bằng 80% axeton mát trong 30 phút ở -20°C, rửa một lần bằng PBS, và phong bế bằng chất đệm phong bế chứa 2,5% (trọng lượng/thể tích) sữa khô không chứa chất béo, 0,5% Triton X-100, 0,05% Tween-20 trong PBS ở 37°C trong 30 phút. Tế bào được phong bế được ủ bằng MAb đặc hiệu kiểu huyết thanh DENV, 1F1 (DENV-1), 3H5 (DENV-2), 8A-1 (DENV-3), hoặc 1H10 (DENV-4) trong chất đệm phong bế ở 37°C trong 1 giờ hoặc 4°C qua đêm, rửa 3 lần bằng chất

đệm rửa (0,05% Tween-20 trong PBS), và ủ bằng IgG của dê kháng chuột tinh khiết ái lực (Jackson Immuno Research Laboratories) được tiếp hợp phosphatasa kiềm hoặc peroxidaza cá ngựa (HRP) ở IgG trong 37°C trong từ 45 đến 60 phút. Rửa đĩa 3 lần trước chất nền thích hợp, 1-Step NBT/BCIP cộng với chất ngăn chặn (Pierce) đối với phosphatasa kiềm hoặc kit Vector-VIP (Vector Labs) đối với HRP, được bổ sung vào để tạo màu. Làm ngừng tạo màu bằng cách rửa bằng nước khi tiêu điểm được phát triển một cách đầy đủ. Tiêu điểm miễn dịch được nhuộm được hiển thị hoá một cách trực tiếp và đếm trên bộp sáng.

Trình tự di truyền

Hệ gen có chiều dài đầy đủ của MVS và WVS được tạo trình tự (xem dưới đây). Tóm lại, ARN virus được chiết từ hạt DENVax bằng cách sử dụng kit ARN virus QIAamp (Qiagen), và mảnh ADN bổ trợ xếp chồng bao phủ toàn bộ hệ gen được khuếch đại bằng cách sử dụng kit Titan One Tube RT-PCR (Roche Applied Science, Inc.). Các mảnh ADN bổ trợ được khuếch đại được tinh chế gel trước khi tạo trình tự với các đoạn môi thuận và nghịch bằng cách sử dụng kit tạo trình tự tuần hoàn BigDye Terminator v3.1 (Applied Biosystems). Các phản ứng trình tự được rửa bằng cách sử dụng kit tinh chế BigDye XTerminator (Applied Biosystems), và chạy trên bộ phân tích di truyền 3130x1 (Applied Biosystems) ở DVBD/CDC. Phần mềm Lasergene SeqMan (DNASTar, Inc) được sử dụng để phân tích hệ gen và so sánh.

Thử nghiệm đột biến khuếch đại không khớp dựa trên Taqman (TaqMAMA)

TaqMAMA là thử nghiệm dạng đa hình nucleotit đơn lẻ nhạy, định lượng được phát triển để cho phép đánh giá tốt hơn mức đột biến ngược ở locus 5'NC-57 giảm độc lực và được tối ưu hoá tiếp cho nghiên cứu này. ARN virus được chiết từ MVS và WVS được phân tích bằng TaqMAMA với bộ đoạn môi/đầu dò Taqman mà đặc hiệu với vùng 5TSTC-57 kiểu hoang dại hoặc vắc xin. Các đoạn môi thuận được sử dụng để phát hiện trình tự kiểu hoang dại và vắc xin DENV-2 lần lượt là D2-41-GC và D2-40-TT. Nucleotit đầu cuối 3' của mỗi đoạn môi thuận khớp 5'NCR-57 nucleotit đặc hiệu đối với mỗi virus, trong khi nucleotit liền kề với nucleotit đầu cuối 3' trong mỗi đoạn môi khác với trình tự hệ gen virus DENV-2 để tăng cường tác dụng không khớp. Đoạn môi ngược, CD-207, và đầu dò Taqman, CD-169F, các bộ kiểu hoang dại và vắc xin là giống nhau.

Trình tự của đoạn môi và đầu dò cũng như các điều kiện tuần hoàn được mô tả trước đây. RT-PCR theo thời gian thực được thực hiện với hệ iQ5 hoặc CFX-95 (BioRad), bằng cách sử dụng kit BioRad iScript RT-PCR (đối với các đầu dò), trong

25-µl phản ứng chứa 5µl mẫu ARN virus, 0,4uM mỗi đoạn môi và 0,2uM đầu dò. Các phản ứng bộ ba đối với mỗi thử nghiệm đặc hiệu kiểu hoang dại và vacxin được thực hiện đối với mỗi mẫu. Số lượng bản sao hệ gen được xác định so với đường cong tiêu chuẩn được tạo ra đối với mỗi kiểu di truyền virus, trong đó các tiêu chuẩn ARN là các bản sao thu được các plasmit chứa nt 1-2670 của mỗi ADN hỗ trợ đặc hiệu kiểu di truyền. Ngoài ra, tính đặc hiệu của thử nghiệm được xác nhận bằng cách thử nghiệm mỗi tiêu chuẩn ARN với đoạn môi/đầu dò kiểu di truyền khác loại để bảo đảm khả năng phản ứng ngang tối thiểu trong mỗi thử nghiệm. Các kết quả được báo cáo là tỷ lệ phần trăm của hệ gen virus thể hiện sự đột biến ngược. Trước đây, do khả năng phản ứng ngang ở mức cao hơn mà giới hạn hàm lượng ARN đầu vào đối với thử nghiệm này, tính nhạy phát hiện ban đầu bằng khoảng 0,1% đột biến ngược (khả năng phân toả). Sau đó, thử nghiệm này đã được tối ưu hoá hơn nữa bằng cách sử dụng thiết bị PCR theo thời gian thực được cải thiện và kit phản ứng và khả năng phản ứng ngang được giảm một cách đáng kể ở hàm lượng ở mức rất cao (từ 7 đến 8 log₁₀ bản sao) của lượng vào mẫu ARN. Sự tối ưu hoá này dẫn đến sự cải thiện đáng kể về tính nhạy phát hiện, giảm đến từ 0,01 đến 0,07% đột biến ngược.

Sự sao chép virus ở tế bào muỗi C6/36 và tính nhạy nhiệt độ với tế bào động vật có vú Vero

Kiểu hình sao chép của bốn vật liệu gốc DENVax MVS và virus DENV-1, -2, -3, và -4 kiểu hoang dại được đánh giá ở tế bào muỗi C6/36 (*Aedes albopictus*). Tế bào C6/36 được sinh trưởng trong đĩa 6 giếng được lây nhiễm hai lần với mỗi virus ở MOI bằng 0,001 và ủ bằng 4ml/giếng môi trường DMEM chứa 2% FBS trong lò ủ 5% CO₂ ở 28°C. Dịch lỏng nhỏ chứa dịch nổi môi trường nuôi cấy được gom đối với mỗi virus vào ngày 6 pi, trộn với thể tích bằng nhau của môi trường chứa 40% FBS, và lưu trữ ở -80°C cho đến khi sẵn để chuẩn độ vết tan virus.

Tính nhạy nhiệt độ được thực hiện bằng cách so sánh sự sinh trưởng của virus ở 39°C đối lại sự sinh trưởng ở 37°C ở năm ngày pi của tế bào Vero trong đĩa 6 giếng. Tế bào được lây nhiễm bốn lần bằng mỗi virus ở MOI bằng 0,001 ở 37°C. Tiếp theo hấp phụ virus, môi trường nuôi cấy virus được ủ bằng 4ml/giếng của môi trường DMEM chứa 2% FBS trong 2 lò ủ CO₂ 5%, một lò ủ (hai đĩa) ở 37°C và lò còn lại ở 39°C.

Dịch lỏng (50-µl) dịch nổi nuôi cấy được gom vào ngày 5pi, trộn với thể tích cân bằng của DMEM chứa 40% FBS, và lưu trữ ở -80°C cho đến khi sẵn để chuẩn độ vết tan virus. Nhiệt độ lò ủ được định cỡ với nhà máy tạo vết NIST-nhiệt kế định cỡ (từ -1 đến 51°C; ERTCO).

Lây nhiễm, phân toả và truyền nhiễm muỗi

Muỗi *Aedes aegypti* được sử dụng để nghiên cứu là từ khuẩn lạc được thiết lập vào năm 2002 từ xã gần Mae Sot (16' N, 33' E), Thailand. Sau khi xuất hiện từ ấu trùng, muỗi trưởng thành được duy trì ở 28°C ở chu kỳ sáng 16:8 (sáng tối) với 10% dung dịch sucroza được cung cấp tự do. Muỗi giống cái từ năm đến bảy ngày tuổi được sử dụng cho ăn thức ăn máu lây nhiễm hoặc tiêm truyền trong ngực (IT). Dịch lỏng DENVax và DENV kiểu hoang dại mới được nuôi cấy ngay khi thu hoạch (mà không có chu kỳ đông lạnh tan đông bất kỳ) để tạo ra thức ăn máu virus như được chỉ định dưới đây để lây nhiễm qua đường miệng. Dịch nổi virus còn lại được bổ sung với FBS đến nồng độ cuối cùng 20% và dịch lỏng được lưu trữ ở -80°C để chuẩn độ vết tan virus hơn nữa và các thử nghiệm tiêm truyền IT. Hạt DENVax mới được tạo ra đối với các thử nghiệm này được khuếch đại từ hạt tiền chủ trong tế bào Vero và được xét đến là thể tương đương DENVax MVS.

Thức ăn máu lây nhiễm được tạo ra bằng cách trộn virus mới tạo ra ở tỷ lệ 1 : 1 với máu gà ròi loạn đông máu (Colorado Serum Company) vào ngày lây nhiễm qua đường miệng. Muỗi được cho đói đường qua đêm và sau đó cho hỗn hợp virus:máu trong 1 giờ bằng cách sử dụng hệ ăn màng Hemotek (Discovery Workshops). 50- μ l dịch lỏng của thức ăn máu được giữ ở -80°C trong độ chuẩn ngược của liều virus. Giống cái ứ máu đầy đủ được lưu trữ trong điều kiện gây mê lạnh và đặt vào trong hộp bìa cứng với 10% dung dịch sucroza được cung cấp tự do. Hộp bìa cứng được đặt ở 28°C với chu kỳ sáng 16:8 giờ (sáng:tối). Sau 14 ngày, từ 25 đến 30 muỗi từ mỗi nhóm virus được gây mê qua việc phơi trần với triethylamin (Flynap®, Carolina Biological Supply Company) và một chân phía sau được loại bỏ và đặt vào 0,5ml DMEM với 10% FBS và 5% penixilin/streptomycin (lần lượt 100U/ml và 100 μ g/ml). Gôm nước bọt bằng cách xen vôi của muỗi được gây mê vào trong ống mao dẫn chứa 2,5% FBS và 25% dung dịch sucroza. Muỗi được cho phép tiết nước bọt trong ít nhất 15 phút và sau đó ống mao dẫn và đặc cơ thể vào trong ống riêng biệt chứa DMEM. Cơ thể muỗi, chân và nước bọt được lưu trữ ở -80°C cho đến khi chúng được nghiền và được thử nghiệm đối với virus lây nhiễm. Để tiêm truyền IT, muỗi được gây mê lạnh và tiêm truyền với khoảng 50pfu virus trong 0,34 μ l vật tiêm truyền. Muỗi được tiêm truyền được giữ trong 7 ngày trong cùng điều kiện như được mô tả trên đây. Sau đó, muỗi được gây mê và nước bọt và thân của chúng được gom như được mô tả trên đây. Mẫu được lưu trữ ở -80°C cho đến khi xử lý tiếp.

Để xử lý mẫu để chuẩn độ virus, các mẫu cơ thể và chân được làm đồng nhất với BB được phủ đồng (Crossman Corporation, NY) ở 24 chu kỳ/giây trong 4 phút

bằng cách sử dụng máy nghiền trộn và sau đó gạn bằng cách ly tâm ở 3.000 x g trong 3 phút. Mẫu nước bọt được ly tâm ở 3.000 x g trong 3 phút để thải dịch lỏng ra khỏi ống mao dẫn. Dung dịch pha loãng mười lần của thể đồng nhất cơ thể và chân và mẫu nước bọt được thử nghiệm đối với sự có mặt của virus lây nhiễm bằng thử nghiệm vết tan. Các kết quả từ cơ thể, chân và nước bọt được sử dụng để lần lượt xác định sự lây nhiễm, phân toả và truyền nhiễm.

Độc lực thần kinh ở chuột

Chuột ICR giống cái thời kỳ mang thai thu được từ Taconic Labs, và kiểm tra vài lần mỗi ngày để xác định khoảng thời gian sinh đẻ. Trong thử nghiệm đã nêu, khoảng từ 12 đến 24 giờ sau khi sinh, hai ổ gồm tám chuột con/virus (n=16), được thử thách với từ 10^3 đến 10^4 pfu virus trong 20 μ l dịch pha loãng bằng việc tiêm truyền trong sọ (ic) bằng cách sử dụng kim 30-gauge. Động vật được kiểm tra ít nhất 3 lần hàng ngày trong ít nhất 32 ngày sau thử thách. Ở triệu chứng ốm thứ nhất (lông thô, lưng gù, giảm trọng lượng, vận động khác thường, liệt hoặc hôn mê) động vật được gây mê bằng việc gây mê gây chết bằng khí isofluran, tiếp theo làm lệch khớp cổ. Ngày sau khi lây nhiễm gây mê thể hiện "thời gian ốm/mắc bệnh" hoặc "thời gian sống" đối với động vật. Các thử nghiệm động vật được thực hiện theo quy trình động vật được chấp thuận DVBD/CDC IACUC.

Nguồn gốc virus hạt chủ

Hạt virus chủ DENvax-1 (MVS)

Trình tự nucleotit của hệ gen virus khảm và gọt ra trình tự axit amin của protein được dịch mã được đưa ra trong bản mô tả này. Hầu hết gen prM-E (nt 457 đến 2379, được gạch dưới) là (wt) DEN-1 16007 virus đặc hiệu kiểu hoang dại; hệ gen còn lại là DEN-2 PDK-53 virus đặc hiệu. Tất cả các thay thế được tạo tác khác với virus kiểu hoang dại (D1 16007 hoặc D2 16681), cũng như các đột biến bổ sung (các thay đổi từ dòng ADN bổ trợ được thao tác) được phát hiện trong MVS có bán trên thị trường.

Các thay thế được bao gồm trong hệ gen và protein:

Các vị trí ghép nối giữa khung chính D1 (prM-E) và D2:

- a. M1ul (nt 451-456): đột biến câm được thao tác, nt-453 A-thành -G
- b. NgoMIV (nt 2380-2385): đột biến được thao tác, nt-2381/2382 TG- thành -CC (dẫn đến sự thay đổi E-482 Val- thành -Ala)

Khung chính virus D2 PDK-53 (thay đổi từ D2 16681 kiểu hoang dại): tất cả có nét đậm

- a. vùng không mã hoá 5' (NCR)-57 (nt-57 C- thành -T): locus giảm độc lực chính (màu đỏ)
- b. NS1-53 Gly-thành-Asp (nt-2579 G- thành -A): locus giảm độc lực chính (màu đỏ)
- c. NS2A- 181 Leu- thành -Phe (nt-4018 C- thành -T)
- d. NS3-250 Glu-thành-Val (nt-5270 A-thành-T): locus giảm độc lực chính (màu đỏ)
- e. nt-5547 (gen NS3) đột biến câm T- thành -C
- f. NS4A-75 Gly- thành -Ala (nt-6599 G-thành-C)
- đột biến câm nt-8571 C-thành-T của PDK-53 là không được thao tác trong virus vaccine

DE-1 prM-E (thay đổi từ D1 16007 kiểu hoang dại)

- a. Đột biến câm nt-1575 T- thành-C được thao tác để loại bỏ vị trí XbaI nguyên thể

Các thay thế bổ sung được tìm thấy ở hạt vaccine (0,03% nt khác nhau từ dòng gốc)

- a. NS2A-116 Ile-thành-Leu (nt-3823 A-thành-C, nét đậm)
- b. NS2B-92 Glu-thành-Asp (nt-4407 A-thành-T, nét đậm)
- c. đột biến câm nt-7311 A-thành-G (nét đậm)

NCR-57-T, D2 PDK-53 locus giảm độc lực (wt D2 16681: C)

>5'-vùng không mã hoá|>C

10	20	30	40	50	60	70	80	90	100
AGTTGTTAGTCTACGTGGACCGACAAAGACAGATTCTTTGAGGGAGCTAAGCTCAATGTTAGTCTAACAGTTTTTTAATTAGAGAGCAGATCTCTGATGA									
M N									
110	120	130	140	150	160	170	180	190	200
ATAACCAACGGAAAAAGGCGAAAAACACGCCTTTCAATATGCTGAAACGCGAGAGAAACCGCGTGTGCGACTGTGCAACAGCTGACAAAGAGATTCTCACT									
N	Q	R	K	K	A	K	N	T	P
F	N	M	L	K	R	E	R	N	R
V	S	T	V	Q	Q	L	T	K	R
F	S	L							
210	220	230	240	250	260	270	280	290	300
TGGAATGCTGCAGGGACGAGGACCATTAATACTGTTTCATGGCCCTGGTGGCGTTCCTTCGTTTCCTAACATCCCACCAACAGCAGGGATATTGAAGAGA									
G	M	L	Q	G	R	G	P	L	K
L	F	M	A	L	V	A	F	L	R
F	L	T	I	P	P	T	A	G	I
L	K	R							
310	320	330	340	350	360	370	380	390	400
TGGGGAACAATTAAAAAATCAAAAGCTATTAATGTTTGGAGAGGGTTTCAGGAAAGAGATTGGAAGGATGCTGAACATCTTGAATAGGAGACGCAGATCTG									
W	G	T	I	K	K	S	K	A	I
N	V	L	R	G	F	R	K	E	I
G	R	M	L	N	I	L	N	R	R
R	S	A							
>prM									
bắt đầu trình tự D1 16007									
410	420	430	440	450	460	470	480	490	500
CAGGCATGATCATTATGCTGATCCAACAGTGATGGCGTTCCATTTAACACGCGTGGGGGAGAGCCGCATATGATAGTTAGCAAGCAGGAAAGAGGAAA									

G M I I M L I P T V M A F H L T T R G G E P H M I V S K Q E R G K I

Vị trí ghép nối MluI được thao tác (cảm nt-453 từ A-thành-G)

```

      510      520      530      540      550      560      570      580      590      600
GTCAC TTTTGTTC AAGACCTCTGCAGGTGTCAACATGTGCACCCCTCATTCGATGGATTTGGGAGAGTTGTGTGAGGACACGATGACCTACAAATGCCCC
S L L F K T S A G V N M C T L I A M D L G E L C E D T M T Y K C P
610      620      630      640      650      660      670      680      690      700
CGGATCACTGAGGCGGAACAGATGACGTTGACTGTTGGTGCAATGCCACGGACACATGGGTGACCTATGGAACGTGCTCTCAAACCTGGCGAACACCGAC
R I T E A E P D D V D C W C N A T D T W V T Y G T C S Q T G E H R R
>M
      710      720      730      740      750      760      770      780      790      800
GAGACAAACGTTCCGTCGCATTGGCCCCACACGTGGGGCTTGGCCTAGAAACAAGAGCCGAAACGTGGATGTCCTCTGAAGGTGCTTGGAAACAGATACA
D K R S V A L A P H V G L G L E T R A E T W M S S E G A W K Q I Q
      810      820      830      840      850      860      870      880      890      900
AAAAGTAGAGACTTGGGCTCTGAGACATCCAGGATTCACGGTGATAGCCCTTTTTCTAGCACATGCCATAGGAACATCCATCACCCAGAAAGGGATCATT
K V E T W A L R H P G F T V I A L F L A H A I G T S I T Q K G I I
>E
      910      920      930      940      950      960      970      980      990      1000
TTCATTTTGCTGATGCTGGTAACACCATCTATGGCCATGCGATGCGTGGGAATAGGCAACAGAGACTTCGTGGAAGGACTGTCAGGAGCAACATGGGTGG
F I L L M L V T P S M A M R C V G I G N R D F V E G L S G A T W V D
      1010     1020     1030     1040     1050     1060     1070     1080     1090     1100
ATGTGGTACTGGAGCATGGAAGTTGCGTCACCACCATGGCAAAAAACAAACCAACACTGGACATTGAACTCTGAAGACGGAGGTCACAAACCTGCAGT
V V L E H G S C V T T M A K N K P T L D I E L L K T E V T N P A V
      1110     1120     1130     1140     1150     1160     1170     1180     1190     1200
TCTGCGTAAATTGTGCATTGAAGCTAAAATATCAAACACCACCACCGATTTCGAGATGTCCAACACAAGGAGAAGCCACACTGGTGGAGAACAAGACGCG
L R K L C I E A K I S N T T T D S R C P T Q G E A T L V E E Q D A
      1210     1220     1230     1240     1250     1260     1270     1280     1290     1300
AACTTTGTGTGCCGACGAACGTTTCGTGGACAGAGGCTGGGGCAATGGCTGTGGGCTATTTCGAAAAGGTAGTCTAATAACGTGTGCCAAGTTTAAGTGTG
N F V C R R T F V D R G W G N G C G L F G K G S L I T C A K F K C V
      1310     1320     1330     1340     1350     1360     1370     1380     1390     1400
TGACAAAAC TAGAAGGAAAGATAGTTCAATATGAAAACCTAAAATATTTCAGTGATAGTCACCGTCCACACTGGAGATCAGCACCAGGTGGGAAATGAGAC
T K L E G K I V Q Y E N L K Y S V I V T V H T G D Q H Q V G N E T
      1410     1420     1430     1440     1450     1460     1470     1480     1490     1500
TACAGAACATGGAACAACCTGCAACCATAACACCTCAAGCTCCTACGTCGGAAATACAGCTGACCGACTACGGAACCCCTTACATTAGATTGTTACCTAGG
T E H G T T A T I T P Q A P T S E I Q L T D Y G T L T L D C S P R
      1510     1520     1530     1540     1550     1560     1570     1580     1590     1600
ACAGGGCTAGATTTTAAACGAGATGGTGTGCTGACAAATGAAAGAAAGATCATGGCTTGTCCACAAACAATGGTTCTCTAGACTTACCACTGCCTTGGACCT
T G L D F N E M V L L T M K E R S W L V H K Q W F L D L P L P W T S
I

```

Đột biến cảm được thao tác (nt-1575 T-to-C): loại bỏ vị trí xbaI đặc hiệu virus DEN-1 nguyên thể

1610 1620 1630 1640 1650 1660 1670 1680 1690 1700
CTGGGGCTTCAACATCCCAAGAGACTTGGAACAGACAAGATTTACTGGTCACATTTAAGACAGCTCATGCAAAGAAGCAGGAAGTAGTCGTACTAGGATC
 G A S T S Q E T W N R Q D L L V T F K T A H A K K Q E V V V L G S
 1710 1720 1730 1740 1750 1760 1770 1780 1790 1800
ACAAGAAGGAGCAATGCACACTGCGCTGACTGGAGCGACAGAAATCCAAACGTCAGGAACGACAACAATTTTCGCAGGACACCTAAAATGCAGACTAAAA
 Q E G A M H T A L T G A T E I Q T S G T T T I F A G H L K C R L K
 1880 1890 1900
ATGGACAAACTAACTTTAAAGGGATGTCATATGTGATGTGCACAGGCTCATTCAAGTTAGAGAAAAGTGGCTGAGACCCAGCATGGAAGTGTCTGG
 M D K L T L K G M S Y V M C T G S F K L E K E V A E T Q H G T V L V
 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000
TGCAGGTTAAATATGAAGGAACAGACGCACCATGCAAGATTCCCTTTTCGACCCAAGATGAGAAAGGAGCAACCCAGAATGGGAGATTAATAACAGCCAA
 Q V K Y E G T D A P C K I P F S T Q D E K G A T Q N G R L I T A N
 2010 2020 2030 2040 2050 2060 2070 2080 2090 2100
CCCCATAGTCACTGACAAAGAAAAACCAGTCAATATTGAGGCAGAACCCACCTTTGGTGAGAGCTACATCGTGGTAGGAGCAGGTGAAAAAGCTTTGAAA
 P I V T D K E K P V N I E A E P P F G E S Y I V V G A G E K A L K
 2110 2120 2130 2140 2150 2160 2170 2180 2190 2200
CTAAGCTGGTTCAAGAAAGGAAGCAGCATAGGGAAAATGTTTGAAGCAACTGCCCCGAGGAGCACGAAGGATGGCCATTCTGGGAGACACCGCATGGGACT
 L S W F K K G S S I G K M F E A T A R G A R R M A I L G D T A W D F
 2210 2220 2230 2240 2250 2260 2270 2280 2290 2300
TCGGTTCTATAGGAGGAGTGTTCACGTCTATGGGAAAACCTGGTACACCAGGTTTTTGGAACTGCATATGGAGTTTTGTTTAGCGAGTTTCTTGGACCAT
 G S I G G V F T S M G K L V H Q V F G T A Y G V L F S G V S W T M L
 Kết thúc trình tự **D1 16007**
 2310 2320 2330 2340 2350 2360 2370 2380 2390 2400
GAAAATAGGAATAGGGATTCTGCTGACATGGCTAGGATTAATTC AAGGAACACGTCCTTTTCGATGATGTGCATCGCAGCCGGCATTGTGACACTGTAT
 K I G I G I L L T W L G L N S R N T S L S M M C I A A G I V T L Y I

Vị trí ghép nối NgoMIV được thao tác, E-482 Val-thành-Ala (nt-2381/2382 TG-thành-CC)

>NS1

2410 2420 2430 2440 2450 2460 2470 2480 2490 2500
 TTGGGAGTCATGGTGCAGGCCGATAGTGGTTGCGTTGTGAGCTGGAAAAACAAAGAACTGAAATGTGGCAGTGGGATTTTCATCACAGACAACGTGCACA
 L G V M V Q A D S G C V V S W K N K E L K C G S G I F I T D N V H T
 2510 2520 2530 2540 2550 2560 2570 2580 2590 2600
 CATGGACAGAAACAATACAAGTTCCAACCAGAATCCCCTTCAAACTAGCTTCAGCTATCCAGAAAGCCCATGAAGAGGACATTTGTGGAATCCGCTCAGT
 W T E Q Y K F Q P E S P S K L A S A I Q K A H E E D I C G I R S V L

D2 PDK-53 NS1-53-Asp locus giảm độc lực (wt D2 16681: Gly, nt-2579-G)

2610 2620 2630 2640 2650 2660 2670 2680 2690 2700
 AACAAGACTGGAGAATCTGATGTGGAAACAAATAACACCAGAATTGAATCACATTCTATCAGAAAATGAGGTGAAGTTAACTATTATGACAGGAGACATC
 T R L E N L M W K Q I T P E L N H I L S E N E V K L T I M T G D I
 2710 2720 2730 2740 2750 2760 2770 2780 2790 2800
 AAAGGAATCATGCAGGCAGGAAACGATCTCTGCGGCCTCAGCCCACTGAGCTGAAGTATTCATGGAAAACATGGGGCAAAGCAAAAATGCTCTCTACAG

K G I M Q A G K R S L R P Q P T E L K Y S W K T W G K A K M L S T E
 2810 2820 2830 2840 2850 2860 2870 2880 2890 2900
 AGTCTCATAACCAGACCTTTCTCATTGATGGCCCCGAAACAGCAGAATGCCCCAACACAAATAGAGCTTGGGAATTCGTTGGAAGTTGAAGACTATGGCTT
 S H N Q T F L I D G P E T A E C P N T N R A W N S L E V E D Y G F
 2910 2920 2930 2940 2950 2960 2970 2980 2990 3000
 TGGAGTATTCACCACCAATATATGGCTAAAATTGAAAGAAAAACAGGATGTATTCTGCGACTCAAAACTCATGTGACGCGCCATAAAAGACAACAGAGCC
 G V F T T N I W L K L K E K Q D V F C D S K L M S A A I K D N R A
 3010 3020 3030 3040 3050 3060 3070 3080 3090 3100
 GTCCATGCCGATATGGGTTATTGGATAGAAAGTGCACCTCAATGACACATGGAAGATAGAGAAAGCCTCTTTCATTGAAGTTAAAACTGCCACTGGCCAA
 V H A D M G Y W I E S A L N D T W K I E K A S F I E V K N C H W P K
 3110 3120 3130 3140 3150 3160 3170 3180 3190 3200
 AATCACACACCCTCTGGAGCAATGGAGTGCTAGAAAGTGAGATGATAATTCCAAAGAATCTCGCTGGACCAGTGTCTCAACACAACCTATAGACCAGGCTA
 S H T L W S N G V L E S E M I I P K N L A G P V S Q H N Y R P G Y
 3210 3220 3230 3240 3250 3260 3270 3280 3290 3300
 CCATACACAAATAACAGGACCATGGCATCTAGGTAAGCTTGAGATGGACTTTGATTTCTGTGATGGAACAACAGTGGTAGTGACTGAGGACTGCGGAAAT
 H T Q I T G P W H L G K L E M D F D F C D G T T V V V T E D C G N
 3310 3320 3330 3340 3350 3360 3370 3380 3390 3400
 AGAGGACCCTCTTTGAGAACAACCACTGCCTCTGGAAAACCTATAACAGAATGGTGCTGCCGATCTTGCACATTACCACCGCTAAGATACAGAGGTGAGG
 R G P S L R T T T A S G K L I T E W C C R S C T L P P L R Y R G E D
>NS2A
 3410 3420 3430 3440 3450 3460 3470 3480 3490 3500
 ATGGGTGCTGGTACGGGATGGAATCAGACCATTGAAGGAGAAAGAAGAGAATTTGGTCAACTCCTTGGTCACAGCTGGACATGGGCAGGTCGACAACCTT
 G C W Y G M E I R P L K E K E E N L V N S L V T A G H G Q V D N F
 3510 3520 3530 3540 3550 3560 3570 3580 3590 3600
 TTCACTAGGAGTCTTGGGAATGGCATTGTTCTCTGGAGGAAATGCTTAGGACCCGAGTAGGAACGAAACATGCAATACTACTAGTTGCAGTTTCTTTTGTG
 S L G V L G M A L F L E E M L R T R V G T K H A I L L V A V S F V
 3610 3620 3630 3640 3650 3660 3670 3680 3690 3700
 ACATTGATCACAGGGAACATGTCCTTTAGAGACCTGGGAAGAGTGATGGTTATGGTAGGCGCCACTATGACGGATGACATAGGTATGGGCGTGACTTATC
 T L I T G N M S F R D L G R V M V M V G A T M T D D I G M G V T Y L
 3710 3720 3730 3740 3750 3760 3770 3780 3790 3800
 TTGCCCTACTAGCAGCCTTCAAAGTCAGACCAACTTTTGCAGCTGGACTACTCTTGAGAAAGCTGACCTCCAAGGAATTGATGATGACTACTATAGGAAT
 A L L A A F K V R P T F A A G L L L R K L T S K E L M M T T I G I
 3810 3820 3830 3840 3850 3860 3870 3880 3890 3900
 TGTACTCCTCTCCCAGAGCACCCCTACCAGAGACCATTCTTGAGTTGACTGATGCGTTAGCCTTAGGCATGATGGTCTCAAATGGTGAGAAATATGGAA
 V L L S Q S T L P E T I L E L T D A L A L G M M V L K M V R N M E I
Đột biến NS2A-116 Ile-thành-Leu (nt3823 A-thành-C) bổ sung trong hạt chủ và tiền chủ
 3910 3920 3930 3940 3950 3960 3970 3980 3990 4000
 AAGTATCAATTGGCAGTGACTATCATGGCTATCTTGTGCGTCCCAACGCAGTGATATTACAAAACGCATGGAAAGTGAGTTGCACAATATTGGCAGTGG
 K Y Q L A V T I M A I L C V P N A V I L Q N A W K V S C T I L A V V

4010 4020 4030 4040 4050 4060 4070 4080 4090 4100
 TGTCCGTTTCCCCACTGTTCTTAACATCCTCACAGCAAAAAACAGATTGGATACCATTAGCATTGACGATCAAAGGTCTCAATCCAACAGCTATTTTCT
 S V S P L F L T S S Q Q K T D W I P L A L T I K G L N P T A I F L |

NS2A-181-Phe đặc hiệu D2 PDK-53 (wt D2 16681: Leu, nt-4018-C)

>NS2B

4110 4120 4130 4140 4150 4160 4170 4180 4190 4200
 AACAAACCCTCTCAAGAACCAGCAAGAAAAGGAGCTGGCCATTAAATGAGGCTATCATGGCAGTCGGGATGGTGAGCATTTTAGCCAGTTCTCTCCTAAAA
 T T L S R T S K K R S W P L N E A I M A V G M V S I L A S S L L K
 4210 4220 4230 4240 4250 4260 4270 4280 4290 4300
 AATGATATTCCCATGACAGGACCATTAGTGGCTGGAGGGCTCCTCACTGTGTGCTACGTGCTCACTGGACGATCGGCCGATTGGAAGTGGAGAGAGCAG
 N D I P M T G P L V A G G L L T V C Y V L T G R S A D L E L E R A A
 4310 4320 4330 4340 4350 4360 4370 4380 4390 4400
 CCGATGTCAAATGGGAAGACCAGGCAGAGATATCAGGAAGCAGTCCAATCCTGTCAATAACAATATCAGAAGATGGTAGCATGTCGATAAAAAATGAAGA
 D V K W E D Q A E I S G S S P I L S I T I S E D G S M S I K N E E
 4410 4420 4430 4440 4450 4460 4470 4480 4490 4500
 GGAAGATCAAACACTGACCATACTCATTAGAACAGGATTGCTGGTGATCTCAGGACTTTTTCTGTATCAATACCAATCACGGCAGCAGCATGGTACCTG
 E D Q T L T I L I R T G L L V I S G L F P V S I P I T A A A W Y L L

Đột biến NS2B-92 Glu-thành-Asp (nt-4407 A-thành-T) bổ sung (trong hạt chủ và tiền chủ)

>NS3

4510 4520 4530 4540 4550 4560 4570 4580 4590 4600
 TGGGAAGTGAAGAAACAACGGGCCGGAGTATTGTGGGATGTTCTTCACCCCCACCCATGGGAAAGGCTGAACTGGAAGATGGAGCCTATAGAATTAAGC
 W E V K K Q R A G V L W D V P S P P P M G K A E L E D G A Y R I K Q
 4610 4620 4630 4640 4650 4660 4670 4680 4690 4700
 AAAAAGGGATTCTTGATATTCCAGATCGGAGCCGGAGTTTACAAAGAAGAACATTCCATACAATGTGGCATGTCACACGTGGCGCTGTTCTAATGCA
 K G I L G Y S Q I G A G V Y K E G T F H T M W H V T R G A V L M H
 4710 4720 4730 4740 4750 4760 4770 4780 4790 4800
 TAAAGGAAAGAGGATTGAACCATCATGGGCGGACGTCAAGAAAGACCTAATATCATATGGAGGAGGCTGGAAGTTAGAAGGAGAATGGAAGGAAGGAGAA
 K G K R I E P S W A D V K K D L I S Y G G G W K L E G E W K E G E
 4810 4820 4830 4840 4850 4860 4870 4880 4890 4900
 GAAGTCCAGGTATTGGCACTGGAGCCTGGAAAAAATCCAAGAGCCGTCCAAACGAAACCTGGTCTTTTCAAACCAACGCCGAACAATAGGTGCTGTAT
 E V Q V L A L E P G K N P R A V Q T K P G L F K T N A G T I G A V S
 4910 4920 4930 4940 4950 4960 4970 4980 4990 5000
 CTCTGGACTTTTCTCTGGAACGTCAGGATCTCCAATTATCGACAAAAAAGGAAAAGTTGTGGGTCTTTATGGTAATGGTGTGTTACAAGGAGTGGAGC
 L D F S P G T S G S P I I D K K G K V V G L Y G N G V V T R S G A
 5010 5020 5030 5040 5050 5060 5070 5080 5090 5100
 ATATGTGAGTGTATAGCCCAGACTGAAAAAGCATTGAAGACAACCCAGAGATCGAAGATGACATTTTCCGAAAGAGAAGACTGACCATCATGGACCTC
 Y V S A I A Q T E K S I E D N P E I E D D I F R K R R L T I M D L
 5110 5120 5130 5140 5150 5160 5170 5180 5190 5200
 CACCCAGGAGCGGGAAAGACGAAGAGATACCTTCCGGCCATAGTCAGAGAAGCTATAAAACGGGGTTTGAGAACATTAATCTTGGCCCCACTAGAGTTG

H P G A G K T K R Y L P A I V R E A I K R G L R T L I L A P T R V V
 5210 5220 5230 5240 5250 5260 5270 5280 5290 5300
 TGGCAGCTGAAATGGAGGAAGCCCTTAGAGGACTTCCAATAAGATACCAGACCCAGCCATCAGAGCTGTGCACACCGGGCGGGAGATTGTGGACCTAAT
 A A E M E E A L R G L P I R Y Q T P A I R A V H T G R E I V D L M L

D2 PDK-53 NS3-250-Val locus giảm độc lực (wt D2 16681: Glu, nt-5270-A)

5310 5320 5330 5340 5350 5360 5370 5380 5390 5400
 GTGTCATGCCACATTTACCATGAGGCTGCTATCACCAGTTAGAGTGCCAAACTACAACTGATTATCATGGACGAAGCCCATTTACAGACCCAGCAAGT
 C H A T F T M R L L S P V R V P N Y N L I I M D E A H F T D P A S
 5410 5420 5430 5440 5450 5460 5470 5480 5490 5500
 ATAGCAGCTAGAGGATACATCTCAACTCGAGTGGAGATGGGTGAGGCAGCTGGGATTTTATGACAGCCACTCCCCGGGAAGCAGAGACCCATTTCTCTC
 I A A R G Y I S T R V E M G E A A G I F M T A T P P G S R D P F P Q
 5510 5520 5530 5540 5550 5560 5570 5580 5590 5600
 AGAGCAATGCACCAATCATAGATGAAGAAAGAGAAATCCCTGAACGCTCGTGAATTCGGGACATGAATGGGTACGGATTTTAAAGGGAAGACTGTTTG
 S N A P I I D E E R E I P E R S W N S G H E W V T D F K G K T V W I

Đột biến câm D2 PDK-53 nt-5547-C (wt D2 16681: T)

5610 5620 5630 5640 5650 5660 5670 5680 5690 5700
 GTTCGTTCCAAGTATAAAAGCAGGAAATGATATAGCAGCTTGCCCTGAGGAAAAATGGAAAGAAAGTGATACAACCTCAGTAGGAAGACCTTTGATTCTGAG
 F V P S I K A G N D I A A C L R K N G K K V I Q L S R K T F D S E
 5710 5720 5730 5740 5750 5760 5770 5780 5790 5800
 TATGTCAAGACTAGAACCAATGATTGGGACTTCGTGGTTACAACCTGACATTTCAGAAATGGGTGCCAATTTCAAGGCTGAGAGGGTTATAGACCCAGAC
 Y V K T R T N D W D F V V T T D I S E M G A N F K A E R V I D P R R
 5810 5820 5830 5840 5850 5860 5870 5880 5890 5900
 GCTGCATGAAACCAGTCATACTAACAGATGGTGAAGAGCGGGTGATCTCTGGCAGGACCTATGCCAGTGACCCACTCTAGTGCAGCACAAAGAAGAGGGAG
 C M K P V I L T D G E E R V I L A G P M P V T H S S A A Q R R G R
 5910 5920 5930 5940 5950 5960 5970 5980 5990 6000
 AATAGGAAGAAATCCAAAAATGAGAATGACCAGTACATATACATGGGGGAACCTCTGGAAAATGATGAAGACTGTGCACACTGGAAAGAAGCTAAAATG
 I G R N P K N E N D Q Y I Y M G E P L E N D E D C A H W K E A K M
 6010 6020 6030 6040 6050 6060 6070 6080 6090 6100
 CTCCTAGATAACATCAACACGCCAGAAGGAATCATTCCTAGCATGTTTCAACCAGAGCGTGAAAAGGTGGATGCCATTGATGGCGAATACCGCTTGAGAG
 L L D N I N T P E G I I P S M F E P E R E K V D A I D G E Y R L R G
 6110 6120 6130 6140 6150 6160 6170 6180 6190 6200
 GAGAAGCAAGGAAAACCTTTGTAGACTTAATGAGAAGAGGAGACCTACCAGTCTGGTTGGCCTACAGAGTGGCAGCTGAAGGCATCAACTACGCAGACAG
 E A R K T F V D L M R R G D L P V W L A Y R V A A E G I N Y A D R
 6210 6220 6230 6240 6250 6260 6270 6280 6290 6300
 AAGGTGGTGTTTTGATGGAGTCAAGAACAACCAAATCCTAGAAGAAAACGTGGAAGTTGAAATCTGGACAAAAGAAGGGGAAAGGAAGAAATGAAACCC
 R W C F D G V K N N Q I L E E N V E V E I W T K E G E R K K L K P

>NS4A

6310 6320 6330 6340 6350 6360 6370 6380 6390 6400

AGATGGTTGGATGCTAGGATCTATTCTGACCCACTGGCGCTAAAAGAATTTAAGGAATTTGCAGCCGGAAGAAAGTCTCTGACCCTGAACCTAATCACAG
 R W L D A R I Y S D P L A L K E F K E F A A G R K S L T L N L I T E
 6410 6420 6430 6440 6450 6460 6470 6480 6490 6500
 AAATGGGTAGGCTCCCAACCTTCATGACTCAGAAGGCAAGAGACGCACTGGACAACCTTAGCAGTGTGCACACGGCTGAGGCAGGTGGAAGGGCGTACAA
 M G R L P T F M T Q K A R D A L D N L A V L H T A E A G G R A Y N
 6510 6520 6530 6540 6550 6560 6570 6580 6590 6600
 CCATGCTCTCAGTGAAGTCCCGGAGACCCTGGAGACATTGCTTTTACTGACACTTCTGGCTACAGTCACGGGAGGGATCTTTTATTCTTGATGAGCGCA
 H A L S E L P E T L E T L L L L T L L A T V T G G I F L F L M S A |
NS4A-75-Ala đặc hiệu D2 PDK-53 (wt D2 16681: Gly, nt-6599-G)
 6610 6620 6630 6640 6650 6660 6670 6680 6690 6700
 AGGGGCATAGGAAGATGACCTGGGAATGTGCTGCATAATCAGGCTAGCATCCTCCTATGGTACGCACAAATACAGCCCACTGGATAGCAGCTTCAA
 R G I G K M T L G M C C I I T A S I L L W Y A Q I Q P H W I A A S I
 6710 6720 6730 6740 6750 6760 6770 6780 6790 6800
 TAATACTGGAGTTTTTTCTCATAGTTTTTGCTTATTCCAGAACCTGAAAAACAGAGAACACCCCAAGACAACCAACTGACCTACGTTGTCATAGCCATCCT
 I L E F F L I V L L I P E P E K Q R T P Q D N Q L T Y V V I A I L
>NS4B
 6810 6820 6830 6840 6850 6860 6870 6880 6890 6900
 CACAGTGGTGGCCGCAACCATGGCAAACGAGATGGGTTTCCTAGAAAAACGAAGAAAGATCTCGGATTGGGAAGCATTGCAACCCAGCAACCCGAGAGC
 T V V A A T M A N E M G F L E K T K K D L G L G S I A T Q Q P E S
 6910 6920 6930 6940 6950 6960 6970 6980 6990 7000
 AACATCCTGGACATAGATCTACGTCCTGCATCAGCATGGACGCTGTATGCCGTGGCCACAACATTTGTTACACCAATGTTGAGACATAGCATTGAAAATT
 N I L D I D L R P A S A W T L Y A V A T T F V T P M L R H S I E N S
 7010 7020 7030 7040 7050 7060 7070 7080 7090 7100
 CCTCAGTGAATGTGTCCCTAACAGCTATAGCCAACCAAGCCACAGTGTTAATGGGTCTCGGGAAAGGATGGCCATTGTCAAAGATGGACATCGGAGTTCC
 S V N V S L T A I A N Q A T V L M G L G K G W P L S K M D I G V P
 7110 7120 7130 7140 7150 7160 7170 7180 7190 7200
 CCTTCTCGCCATTGGATGCTACTCACAAGTCAACCCATAACTCTCACAGCAGCTCTTTTCTTATTGGTAGCACATTATGCCATCATAGGGCCAGGACTC
 L L A I G C Y S Q V N P I T L T A A L F L L V A H Y A I I G P G L
 7210 7220 7230 7240 7250 7260 7270 7280 7290 7300
 CAAGCAAAAGCAACAGAGAAGCTCAGAAAAGAGCAGCGCGGGCATCATGAAAAACCAACTGTCGATGGAATAACAGTGATTGACCTAGATCCAATAC
 Q A K A T R E A Q K R A A A G I M K N P T V D G I T V I D L D P I P
 7310 7320 7330 7340 7350 7360 7370 7380 7390 7400
 CTTATGATCCGAAGTTTGAAAAGCAGTTGGGACAAGTAATGCTCCTAGTCCTCTGCGTGACTCAAGTATTGATGATGAGGACTACATGGGCTCTGTGTGA
 Y D P K F E K Q L G Q V M L L V L C V T Q V L M M R T T W A L C E L
Đột biến câm bổ sung(nt-7311 A→thành-G, trong hạt chủ và tiền chủ)
 7410 7420 7430 7440 7450 7460 7470 7480 7490 7500
 GGCTTTAACTTAGCTACCGGGCCCATCTCCACATTGTGGGAAGGAAATCCAGGGAGGTTTTGGAACACTACCATTGCGGTGTCAATGGCTAACATTTTT
 A L T L A T G P I S T L W E G N P G R F W N T T I A V S M A N I F

> NS5

```

      7510      7520      7530      7540      7550      7560      7570      7580      7590      7600
AGAGGGAGTTACTTGGCCGGAGCTGGACTTCTCTTTTCTATTATGAAGAACAACAACAAGAAGGGGAAGTGGCAACATAGGAGAGACGCTTGGAG
R G S Y L A G A G L L F S I M K N T T N T R R G T G N I G E T L G E
      7610      7620      7630      7640      7650      7660      7670      7680      7690      7700
AGAAATGGAAAAGCCGATTGAACGCATTGGGAAAAAGTGAATTCAGATCTACAAGAAAAGTGAATCCAGGAAGTGGATAGAACCTTAGCAAAAGAAGG
K W K S R L N A L G K S E F Q I Y K K S G I Q E V D R T L A K E G
      7710      7720      7730      7740      7750      7760      7770      7780      7790      7800
CATTAAAAGAGGAGAAACGGACCATCACGCTGTGTGCGGAGGCTCAGCAAAACTGAGATGGTTTCGTTGAGAGAAACATGGTCACACCAGAAGGGAAAGTA
I K R G E T D H H A V S R G S A K L R W F V E R N M V T P E G K V
      7810      7820      7830      7840      7850      7860      7870      7880      7890      7900
GTGGACCTCGGTTGTGGCAGAGGAGGCTGGTCATACTATTGTGGAGGACTAAAGAATGTAAGAGAAGTCAAAGGCCTAACAAAAGGAGGACCAGGACACG
V D L G C G R G G W S Y Y C G G L K N V R E V K G L T K G G P G H E
      7910      7920      7930      7940      7950      7960      7970      7980      7990      8000
AAGAACCCATCCCCATGTCAACATATGGGTGGAATCTAGTGCCTCTTCAAAGTGGAGTTGACGTTTTCTTCATCCCGCCAGAAAAGTGTGACACATTATT
E P I P M S T Y G W N L V R L Q S G V D V F F I P P E K C D T L L
      8010      8020      8030      8040      8050      8060      8070      8080      8090      8100
GTGTGACATAGGGGAGTCATCACCAAATCCCACAGTGAAGCAGGACGAACACTCAGAGTCCTTAACCTTAGTAGAAAATTGGTTGAACAACAACACTCAA
C D I G E S S P N P T V E A G R T L R V L N L V E N W L N N N T Q
      8110      8120      8130      8140      8150      8160      8170      8180      8190      8200
TTTTGCATAAAGGTTCTCAACCCATATATGCCCTCAGTCATAGAAAAATGGAAGCACTACAAAGGAAATATGGAGGAGCCCTTAGTGAGGAATCCACTCT
F C I K V L N P Y M P S V I E K M E A L Q R K Y G G A L V R N P L S
      8210      8220      8230      8240      8250      8260      8270      8280      8290      8300
CACGAAACTCCACACATGAGATGTACTGGGTATCCAATGCTTCCGGGAACATAGTGTGCATCAGTGAACATGATTTCAAGGATGTTGATCAACAGATTTAC
R N S T H E M Y W V S N A S G N I V S S V N M I S R M L I N R F T
      8310      8320      8330      8340      8350      8360      8370      8380      8390      8400
AATGAGATACAAGAAAGCCACTTACGAGCCGGATGTTGACCTCGGAAGCGGAACCCGTAACATCGGGATTGAAAGTGAAGATACCAAACCTAGATATAATT
M R Y K K A T Y E P D V D L G S G T R N I G I E S E I P N L D I I
      8410      8420      8430      8440      8450      8460      8470      8480      8490      8500
GGGAAAAGAATAGAAAAATAAAGCAAGAGCATGAAACATCATGGCACTATGACCAAGACCACCATACAAAACGTGGGCATACCATGGTAGCTATGAAA
G K R I E K I K Q E H E T S W H Y D Q D H P Y K T W A Y H G S Y E T
      8510      8520      8530      8540      8550      8560      8570      8580      8590      8600
CAAAACAGACTGGATCAGCATCATCCATGGTCAACGGAGTGGTCAGGCTGCTGACAAAACCTTGGGACGTCGTCCCATGGTGACACAGATGGCAATGAC
K Q T G S A S S M V N G V V R L L T K P W D V V P M V T Q M A M T
      8610      8620      8630      8640      8650      8660      8670      8680      8690      8700
AGACACGACTCCATTTGGACAACAGCGCGTTTTTAAAGAGAAAAGTGGACACGAGAACCCAAGAACCGAAAGAAGGCACGAAGAACTAATGAAAATAACA
D T T P F G Q Q R V F K E K V D T R T Q E P K E G T K K L M K I T
      8710      8720      8730      8740      8750      8760      8770      8780      8790      8800
GCAGAGTGGCTTTGGAAAAGATTAGGGAAGAAAAAGACACCCAGGATGTGCACCAGAGAAGAATTCACAAGAAAGGTGAGAAGCAATGCAGCCTTGGGGG

```

A E W L W K E L G K K K T P R M C T R E E F T R K V R S N A A L G A
 8810 8820 8830 8840 8850 8860 8870 8880 8890 8900
 CCATATTCTACTGATGAGAACAAGTGGAGTCGGCACGTGAGGCTGTTGAAGATAGTAGGTTTGGGAGCTGGTTGACAAGGAAAGGAATCTCCATCTTGA
 I F T D E N K W K S A R E A V E D S R F W E L V D K E R N L H L E
 8910 8920 8930 8940 8950 8960 8970 8980 8990 9000
 AGGAAAGTGTGAAACATGTGTGTACAACATGATGGGAAAAAGAGAGAAGAAGCTAGGGGAATTTCGGCAAGGCAAAAGGCAGCAGAGCCATATGGTACATG
 G K C E T C V Y N M M G K R E K K L G E F G K A K G S R A I W Y M
 9010 9020 9030 9040 9050 9060 9070 9080 9090 9100
 TGGCTTGGAGCACGCTTCTTAGAGTTTGAAGCCCTAGGATTCTTAAATGAAGATCACTGGTTCTCCAGAGAGAACTCCCTGAGTGGAGTGGAAAGGAGAAG
 W L G A R F L E F E A L G F L N E D H W F S R E N S L S G V E G E G
 9110 9120 9130 9140 9150 9160 9170 9180 9190 9200
 GGCTGCACAAGCTAGGTTACATTCTAAGAGACGTGAGCAAGAAAGAGGGAGGAGCAATGTATGCCGATGACACCGCAGGATGGGATACAAGAATCACACT
 L H K L G Y I L R D V S K K E G G A M Y A D D T A G W D T R I T L
 9210 9220 9230 9240 9250 9260 9270 9280 9290 9300
 AGAAGACCTAAAAAATGAAGAAATGGTAACAAACCACATGGAAGGAGAACACAAGAACTAGCCGAGGCCATTTTCAAATAACGTACCAAAACAAGGTG
 E D L K N E E M V T N H M E G E H K K L A E A I F K L T Y Q N K V
 9310 9320 9330 9340 9350 9360 9370 9380 9390 9400
 GTGCGTGTGCAAAGACCAACACCAAGAGGCACAGTAATGGACATCATATCGAGAAGAGACCAAAGAGGTAGTGGACAAGTTGGCACCTATGGACTCAATA
 V R V Q R P T P R G T V M D I I S R R D Q R G S G Q V G T Y G L N T
 9410 9420 9430 9440 9450 9460 9470 9480 9490 9500
 CTTTCACCAATATGGAAGCCCAACTAATCAGACAGATGGAGGGAGAAGGAGTCTTTAAAGCATTTCAGCACCTAACAATCACAGAAGAAATCGCTGTGCA
 F T N M E A Q L I R Q M E G E G V F K S I Q H L T I T E E I A V Q
 9510 9520 9530 9540 9550 9560 9570 9580 9590 9600
 AAAGTGGTTAGCAAGAGTGGGGCGCGAAAGGTTATCAAGAATGGCCATCAGTGGAGATGATTGTGTTGTGAAACCTTTAGATGACAGGTTTCGCAAGCGCT
 N W L A R V G R E R L S R M A I S G D D C V V K P L D D R F A S A
 9610 9620 9630 9640 9650 9660 9670 9680 9690 9700
 TTAACAGCTCTAAATGACATGGGAAAGATTAGGAAAGACATACAACAATGGGAACCTTCAAGAGGATGGAATGATTGGACACAAGTGCCCTTCTGTTTAC
 L T A L N D M G K I R K D I Q Q W E P S R G W N D W T Q V P F C S H
 9710 9720 9730 9740 9750 9760 9770 9780 9790 9800
 ACCATTTCCATGAGTTAATCATGAAAGACGGTCGCGTACTCGTTGTTCCATGTAGAAACCAAGATGAACTGATTGGCAGAGCCCGAATCTCCCAAGGAGC
 H F H E L I M K D G R V L V V P C R N Q D E L I G R A R I S Q G A
 9810 9820 9830 9840 9850 9860 9870 9880 9890 9900
 AGGGTGGTCTTTGCGGGAGACGGCCTGTTTGGGGAAGTCTTACGCCCAAATGTGGAGCTTGATGTACTTCCACAGACGCGACCTCAGGCTGGCGGCAAAAT
 G W S L R E T A C L G K S Y A Q M W S L M Y F H R R D L R L A A N
 9910 9920 9930 9940 9950 9960 9970 9980 9990 10000
 GCTATTTGCTCGGCAGTACCATCACATTGGGTTCCAACAAGTCGAACAACCTGGTCCATACATGCTAAACATGAATGGATGACAACGGAAGACATGCTGA
 A I C S A V P S H W V P T S R T T W S I H A K H E W M T T E D M L T
 10010 10020 10030 10040 10050 10060 10070 10080 10090 10100
 CAGTCTGGAACAGGGTGTGGATTCAAGAAAACCCATGGATGGAAGACAAAACCTCCAGTGGAAATCATGGGAGGAAATCCCATACTTGGGGAAAAGAGAAGA

V W N R V W I Q E N P W M E D K T P V E S W E E I P Y L G K R E D
 10110 10120 10130 10140 10150 10160 10170 10180 10190 10200
 CCAATGGTGGCTCATTGATTGGGTAAACAAGCAGGGCCACCTGGGCAAAGAACATCCAAGCAGCAATAAATCAAGTTAGATCCCTTATAGGCAATGAA
 Q W C G S L I G L T S R A T W A K N I Q A A I N Q V R S L I G N E
 >3'-vùng không mã hoá
 10210 10220 10230 10240 10250 10260 10270 10280 10290 10300
 GAATACACAGATTACATGCCATCCATGAAAAGATTGAGAAGAGAAGGAAGAAGCAGGAGTTCTGTGGTAGAAAGCAAACATAACATGAAACAAGGCTA
 E Y T D Y M P S M K R F R R E E E E A G V L W *
 10310 10320 10330 10340 10350 10360 10370 10380 10390 10400
 GAAGTCAGGTCGGATTAAGCCATAGTACGGAAAAAATATGCTACCTGTGAGCCCCGTCCAAGGACGTTAAAGAAGTCAGGCCATCATAAATGCCATAG
 10410 10420 10430 10440 10450 10460 10470 10480 10490 10500
 CTTGAGTAAACTATGCAGCCTGTAGCTCCACCTGAGAAGGTGTAAAAATCCGGGAGGCCACAAACCATGGAAGCTGTACGCATGGCGTAGTGACTAGC
 10510 10520 10530 10540 10550 10560 10570 10580 10590 10600
 GGTTAGAGGAGACCCCTCCCTTACAAATCGCAGCAACAATGGGGGCCCAAGGCGAGATGAAGCTGTAGTCTCGCTGGAAGGACTAGAGGTTAGAGGAGAC
 10610 10620 10630 10640 10650 10660 10670 10680 10690 10700
 CCCCCGAAACAAAAACAGCATATTGACGCTGGGAAAGACCAGAGATCCTGCTGTCTCCTCAGCATCATTCAGGCACAGAACGCCAGAAAATGGAATG
 10710 10720
 GTGCTGTTGAATCAACAGGTTCT

Hạt virus chủ DENvax-2 (MVS)

Trình tự nucleotit của hệ gen virus tái tổ hợp và gọi ra trình tự axit amin của protein được dịch mã được nêu trong bản mô tả này. Virus được thao tác được dựa vào virus D2 PDK-53. Tất cả các thay thế được thao tác mà khác với virus DEN-216681 kiểu hoang dại (cũng là virus cha mẹ đối với PDK-53), cũng như các đột biến bổ sung (các thay đổi từ dòng ADN bổ trợ được thao tác) được phát hiện ở MVS có bán trên thị trường.

Các thay thế được bao gồm trong hệ gen và protein:

Khung chính virus D2 PDK-53 (thay đổi từ D216681 kiểu hoang dại): tất cả có nét đậm

- vùng không mã hoá 5' (NCR)-57 (nt-57 C thành T): locus giảm độc lực chính (màu đỏ)
- prM-29 Asp-thành-Val (nt-524 A-thành-T)
- đột biến câm nt-2055 C- thành -T (gen E)
- NS1-53 Gly-thành-Asp (nt-2579 G-thành-A): locus giảm độc lực chính (màu đỏ)
- NS2A-181 Leu-thành-Phe (nt-4018 C-thành-T)

f. NS3-250 Glu-thành-Val (nt-5270 A- thành-T): locus giảm độc lực chính (màu đỏ)

g. đột biến câm nt-5547 (gen NS3) T-thành-C

h. NS4A-75 Gly-thành-Ala (nt-6599 G-thành-C)

* đột biến câm nt-8571 C-thành-T của PDK-53 là không được thao tác trong virut vắc xin

Gen đánh dấu dòng thao tác (đột biến câm):

a. đột biến câm nt-900 T-thành-C: gen đánh dấu dòng lây nhiễm

Các thay thế bổ sung được tìm thấy trong hạt vaccin (0,02% nt khác với dòng ban đầu)

a. prM-52 Lys-thành-Glu (nt-592 A-thành-G), nét đậm

b. NS5-412 Ile-thành-Val (nt-8803 A-thành-G), nét đậm

NCR-57-T, D2 PDK-53 locus giảm độc lực (wt D2 16681: C)

>5'-NC

10	20	30	40	50	60	70	80	90	100	110	120
AGTTGTTAGTCTACGTGGACCGACAAAGACAGATTCTTTGAGGGAGCTAAGCTCAATGTAGTTCTAACAGTTTTTTAATTAGAGAGCAGATCTCTGATGAATAACCAACGGAAAAAGCGC											
										M N N Q R K K A	
130	140	150	160	170	180	190	200	210	220	230	240
AAAAACACGCCCTTTCAATATGCTGAACGCGAGAGAAACCGCGTGTGCGACTGTGCAACAGCTGACAAGAGATTCTCACTTGGAAATGCTGCAGGGACGAGGACCATTAAACTGTTTCATG											
K	N	T	P	F	N	M	L	K	<u>R</u>	E	R
N	R	V	S	T	V	Q	<u>Q</u>	L	T	K	R
F	S	L	G	M	<u>L</u>	Q	G	R	G	P	L
K	L	F	M								
250	260	270	280	290	300	310	320	330	340	350	360
GCCCTGGTGGCGTTCCTTCGTTTCCTAACCAATCCCACCAACAGCAGGGATATTGAAGAGATGGGGAACAATTAAAAAATCAAAGCTATTAAATGTTTTGAGAGGGTTCAGGAAAGAGATT											
A	L	V	A	F	L	R	F	L	<u>T</u>	I	P
P	T	A	G	I	L	K	<u>R</u>	W	G	T	I
K	K	S	K	A	<u>I</u>	N	V	L	R	G	F
R	K	E	I								

>prM

370	380	390	400	410	420	430	440	450	460	470	480
GGAAGGATGCTGAACATCTTGAATAGGAGACGCAGATCTGCAGGCATGATCATTATGCTGATTCCAACAGTGATGGCGTTCACATTTAACCACACGTAACGGAGAACCACACATGATCGTC											
G	R	M	L	N	I	L	N	R	<u>R</u>	R	R
S	A	G	M	I	I	M	<u>L</u>	I	P	T	V
M	A	F	H	L	<u>T</u>	T	R	N	G	E	P
H	M	I	V								
490	500	510	520	530	540	550	560	570	580	590	600
AGCAGACAAGAGAAAGGGAAAAGTCTCTGTTTAAACAGAGGTTGGCGTGAACATGTGTACCCCTCATGGCCATGGACCTTGGTGAATTGTGTGAAGACACAATCACGTACGAGTGTCCC											
S	R	Q	E	K	G	K	S	L	<u>L</u>	F	K
T	E	V	G	V	N	M	<u>C</u>	T	L	M	A
M	D	L	G	E	<u>L</u>	C	E	D	T	I	T
Y	E	C	P								

||

prM-29 Val đặc hiệu D2 PDK-53 (wt D2 16681 Asp, nt-524-A) đột biến prM-52 Lys-to-Glu bổ sung (nt-592 A-thành-G)

>M

610	620	630	640	650	660	670	680	690	700	710	720
CTTCTCAGGCAGAATGAGCCAGAAGACATAGACTGTTGGTGCAACTCTACGTCCACGTGGGTAACCTATGGGACGTGTACCACCATGGGAGAACATAGAAGAGAAAAAGATCAGTGGCA											
L	L	R	Q	N	E	P	E	D	I	D	C
W	C	N	S	T	S	T	W	V	T	Y	G
T	C	T	T	M	G	E	H	R	R	E	K
R	S	V	A								

730 740 750 760 770 780 790 800 810 820 830 840
 CTCGTTCCACATGTGGGAATGGGACTGGAGACGACGAACCTGAAACATGGATGTCATCAGAAGGGGCCTGGAAACATGTCCAGAGAATTGAAACTTGGATCTTGAGACATCCAGGCTTCACC
 L V P H V G M G L E T R T E T W M S S E G A W K H V Q R I E T W I L R H P G F T

>E

850 860 870 880 890 900 910 920 930 940 950 960
 ATGATGGCAGCAATCCTGGCATAACCCATAGGAACGACACATTTCCAAAGAGCCCTGATCTTCATCTTACTGACAGCTGTCACTCCTTCAATGACAATGCGTTGCATAGGAATGTCAAAT
 M M A A I L A Y T I G T T H F Q R A L I F I L L T A V T P S M T M R C I G M S N
 I

Dòng đánh dấu đột biến câm: đột biến câm nt-900 T-thành-C

970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
 AGAGACTTTGTGGAAGGGGTTTCAGGAGGAAGCTGGGTTGACATAGTCTTAGAACATGGAAGCTGTGTGACGACGATGGCAAAAAACAACCAACATTGGATTTTGAACGTATAAAAAACA
 R D F V E G V S G G S W V D I V L E H G S C V T T M A K N K P T L D F E L I K T

1090 1100 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
 GAAGCCAAACAGCCTGCCACCCTAAGGAAGTACTGTATAGAGGCAAAGCTAACCAACACAAACAGAATCTCGCTGCCCAACACAAGGGGAACCCAGCCTAAATGAAGAGCAGGACAAA
 E A K Q P A T L R K Y C I E A K L T N T T T E S R C P T Q G E P S L N E E Q D K

1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 1310 1320
 AGGTTCTGCTGCAAACTCCATGGTAGACAGAGGATGGGGAATGGATGTGGACTATTTGGAAAGGGAGGCATTGTGACCTGTGCTATGTTGAGATGCAAAAAGAACATGGAAGGAAAA
 R F V C K H S M V D R G W G N G C G L F G K G G I V T C A M F R C K K N M E G K

1330 1340 1350 1360 1370 1380 1390 1400 1410 1420 1430 1440
 GTTGTGCAACCAGAAAACCTTGGAAATACACCATTGTGATAACACCTCACTCAGGGGAAGAGCATGCAGTCGGAATATGACACAGGAAAACATGGCAAGGAAATCAAATAACACCACAGAGT
 V V Q P E N L E Y T I V I T P H S G E E H A V G N D T G K H G K E I K I T P Q S

1450 1460 1470 1480 1490 1500 1510 1520 1530 1540 1550 1560
 TCATCACAGAAGCAGAATTGACAGGTTATGGCACTGTCACAATGGAGTGCTCTCCAAGAACGGGCCTCGACTTCAATGAGATGGTGTGCTGCGAGATGGAAAAATAAGCTTGGCTGGTG
 S I T E A E L T G Y G T V T M E C S P R T G L D F N E M V L L Q M E N K A W L V

1570 1580 1590 1600 1610 1620 1630 1640 1650 1660 1670 1680
 CACAGGCAATGGTTCCCTAGACCTGCCGTTACCATGGTTGCCCGGAGCGGACACACAAGGGTCAAATTTGGATACAGAAAGAGACATTGGTCACTTTCAAAAATCCCCATGCGAAGAAACAG
 H R Q W F L D L P L P W L P G A D T Q G S N W I Q K E T L V T F K N P H A K K Q

1690 1700 1710 1720 1730 1740 1750 1760 1770 1780 1790 1800
 GATGTTGTTGTTTTAGGATCCCAAGAAGGGGCCATGCACACAGCACTTACAGGGGCCACAGAAATCCAATGTCATCAGGAAACTTACTCTTCACAGACATCTCAAGTGCAGGCTGAGA
 D V V V L G S Q E G A M H T A L T G A T E I Q M S S G N L L F T G H L K C R L R

1810 1820 1830 1840 1850 1860 1870 1880 1890 1900 1910 1920
 ATGGACAAGCTACAGCTCAAAGGAATGTCATACTCTATGTGCACAGGAAGTTTAAAGTTGTGAAGGAAATAGCAGAAACACAACATGGAACAATAGTTATCAGAGTGCAATATGAAGGG
 M D K L Q L K G M S Y S M C T G K F K V V K E I A E T Q H G T I V I R V Q Y E G

1930 1940 1950 1960 1970 1980 1990 2000 2010 2020 2030 2040
 GACGGCTCTCCATGCAAGATCCCTTTTGAGATAATGGATTGGGAAAAAGACATGTCTTAGTTCGCCTGATTACAGTCAACCCAATTGTGACAGAAAAAGATAGCCCAGTCAACATAGAA
 D G S P C K I P F E I M D L E K R H V L G R L I T V N P I V T E K D S P V N I E

2050 2060 2070 2080 2090 2100 2110 2120 2130 2140 2150 2160
 GCAGAACCTCCATTGGGAGACAGCTACATCATCATTAGGAGTAGAGCCGGGACAACCTGAAGCTCAACTGGTTTAAGAAAGGAAGTTCTATCGGCCAAATGTTTGAGACAACAATGAGGGGG
 A E P P F G D S Y I I I G V E P G Q L K L N W F K K G S S I G Q M F E T T M R G I

Đột biến câm D2 PDK-53 nt-2055-T (D2 16681: C)

2170 2180 2190 2200 2210 2220 2230 2240 2250 2260 2270 2280
 GCGAAGAGAAATGGCCATTTTAGGTGACACAGCCTGGGATTTGGATCCTTGGGAGGAGTGTTTACATCTATAGGAAAGGCTCTCCACCAAGTCTTTGAGCAATCTATGGAGCTGCCTTC
 A K R M A I L G D T A W D F G S L G G V F T S I G K A L H Q V F G A I Y G A A F

2290 2300 2310 2320 2330 2340 2350 2360 2370 2380 2390 2400
 AGTGGGGTTTCATGGACTATGAAAACTCATAGGAGTCATTATCACATGGATAGGAATGAATTCACGCAGCACCTCACTGTCTGTGACACTAGTATTGGTGGGAATTGTGACACTGTAT
 S G V S W T M K I L I G V I I T W I G M N S R S T S L S V T L V L V G I V T L Y

>NS1

2410 2420 2430 2440 2450 2460 2470 2480 2490 2500 2510 2520
 TTGGGAGTCATGGTGCAGGCCGATAGTGGTTGCGTTGTGAGCTGGAAAAACAAAGAACTGAAATGTGGCAGTGGGATTTTCATCAGACACAACGTGCACACATGGACAGAACAAATACAAAG
 L G V M V Q A D S G C V V S W K N K E L K C G S G I F I T D N V H T W T E Q Y K

2530 2540 2550 2560 2570 2580 2590 2600 2610 2620 2630 2640
 TTCCAACCAGAATCCCCTTCAAAACTAGCTTCAGCTATCCAGAAAGCCCCATGAAGAGGACATTTGTGGAATCCGCTCAGTAACAAGACTGGAGAATCTGATGTGGAACAAATAACACCA
 F Q P E S P S K L A S A I Q K A H E E D I C G I R S V T R L E N L M W K Q I T P
 I

D2 PDK-53 NS1-53-Asp locus giảm độc lực (wt D2 16681: Gly, nt-2579-G)

2650 2660 2670 2680 2690 2700 2710 2720 2730 2740 2750 2760
 GAATTGAATCACATTCTATCAGAAAATGAGGTGAAGTTAACTATTATGACAGGAGACATCAAAGGAATCATGCAGGCAGGAAAACGATCTCTCGGCGCTCAGCCCACTGAGCTGAAGTAT
 E L N H I L S E N E V K L T I M T G D I K G I M Q A G K R S L R P Q P T E L K Y

2770 2780 2790 2800 2810 2820 2830 2840 2850 2860 2870 2880
 TCATGGAAAAACATGGGGCAAAGCAAAAATGCTCTCTACAGAGTCTCATAACCAGACCTTTTCTCATTGATGGCCCCGAAACAGCAGAATGCCCCAACACAAATAGAGCTTGGGAATTCGTTG
 S W K T W G K A K M L S T E S H N Q T F L I D G P E T A E C P N T N R A W N S L

2890 2900 2910 2920 2930 2940 2950 2960 2970 2980 2990 3000
 GAAGTTGAAGACTATGGCTTGGAGTATTACCACCAATATATGGCTAAAATTGAAAGAAAAACAGGATGTATTCTGCGACTCAAAACTCATGTGACGGGCCATAAAGACAACAGAGCC
 E V E D Y G F G V F T T N I W L K L K E K Q D V F C D S K L M S A A I K D N R A

3010 3020 3030 3040 3050 3060 3070 3080 3090 3100 3110 3120
 GTCCATGCCGATATGGGTTATTGGATAGAAAGTGCACTCAATGACACATGGAAGATAGAGAAAGCCTCTTTTCATTGAAGTTAAAACTGCCACTGGCCAAAATCACACACCCCTCTGGAGC
 V H A D M G Y W I E S A L N D T W K I E K A S F I E V K N C H W P K S H T L W S

3130 3140 3150 3160 3170 3180 3190 3200 3210 3220 3230 3240

AATGGAGTGCTAGAAAAGTGAGATGATAATTCCAAGAATCTCGCTGGACCACTGTCTCAACACAACCTATAGACCAGGCTACCATACACAAATAACAGGACCATGGCATCTAGGTAAGCTT
N G V L E S E M I I P K N L A G P V S Q H N Y R P G Y H T Q I T G P W H L G K L

3250 3260 3270 3280 3290 3300 3310 3320 3330 3340 3350 3360
GAGATGGACTTTGATTTCTGTGATGGAACAACAGTGGTAGTGACTGAGGACTGCGGAAATAGAGGACCCTCTTTGAGAACAACCACTGCCTCTGGAAACTCATAACAGAATGGTGCTGC
E M D F D F C D G T T V V V T E D C G N R G P S L R T T T A S G K L I T E W C C

>NS2A

3370 3380 3390 3400 3410 3420 3430 3440 3450 3460 3470 3480
CGATCTTGACATTACCACCGCTAAGATACAGAGGTGAGGATGGGTGCTGGTACGGGATGGAATCAGACCATTGAAGGAGAAAGAAGAGAATTTGGTCAACTCCTTGGTCACAGCTGGA
R S C T L P P L R Y R G E D G C W Y G M E I R P L K E K E E N L V N S L V T A G

3490 3500 3510 3520 3530 3540 3550 3560 3570 3580 3590 3600
CATGGGCAGGTGACAACTTTTCACTAGGAGTCTTGGGAATGGCATTGTTCTCTGGAGGAAATGCTTAGGACCCGAGTAGGAACGAAACATGCAATACTACTAGTTGCAGTTTCTTTTGTG
H G Q V D N F S L G V L G M A L F L E E M L R T R V G T K H A I L L V A V S F V

3610 3620 3630 3640 3650 3660 3670 3680 3690 3700 3710 3720
ACATTGATCACAGGGAACATGTCTTTAGAGACCTGGGAAGAGTGATGGTTATGGTAGGCGCCACTATGACGGATGACATAGGTATGGGCGTGACTTATCTTGCCCTACTAGCAGCCTTC
T L I T G N M S F R D L G R V M V M V G A T M T D D I G M G V T Y L A L L A A F

3730 3740 3750 3760 3770 3780 3790 3800 3810 3820 3830 3840
AAAGTCAGACCAACTTTTGCAGCTGGACTACTCTTGAGAAAGCTGACCTCCAAGGAATTGATGATGACTACTATAGGAATTGTACTCCTCTCCAGAGCACCATAACAGAGACCATTCTT
K V R P T F A A G L L L R K L T S K E L M M T T I G I V L L S Q S T I P E T I L

3850 3860 3870 3880 3890 3900 3910 3920 3930 3940 3950 3960
GAGTTGACTGATGCGTTAGCCTTAGGCATGATGGTCTCAAATGGTGAGAAATATGGAAAAGTATCAATTGGCAGTGACTATCATGGCTATCTTGTGCGTCCCAAACGCAGTGATATTA
E L T D A L A L G M M V L K M V R N M E K Y Q L A V T I M A I L C V P N A V I L

3970 3980 3990 4000 4010 4020 4030 4040 4050 4060 4070 4080
CAAAACGCATGGAAGTGAGTTGCACAATATGGCAGTGGTGTCCGTTTCCCACTGTCTTAAACATCCTCACAGCAAAAAACAGATTGGATACCATTTAGCATTGACGATCAAAGGTCTC
Q N A W K V S C T I L A V V S V S P L **F** L T S S Q Q K T D W I P L A L T I K G L
I

NS2A-181-Phe đặc hiệu D2 PDK-53 (wt D2 16681: Leu, nt-4018-C)

>NS2B

4090 4100 4110 4120 4130 4140 4150 4160 4170 4180 4190 4200
AATCCAACAGCTATTTTCTAACAACCCCTCTCAAGAACCAGCAAGAAAAGGAGCTGGCCATTAAATGAGGCTATCATGGCAGTCGGGATGGTGAGCATTTTAGCCAGTTCTCTCTCAAAA
N P T A I F L T T L S R T S K K R S W P L N E A I M A V G M V S I L A S S L L K

4210 4220 4230 4240 4250 4260 4270 4280 4290 4300 4310 4320
AATGATATCCCATGACAGGACCATTAGTGGCTGGAGGGCTCCTCACTGTGTGCTACGTGCTCACTGGACGATCGGCCGATTGGAACTGGAGAGAGCAGCCGATGTCAAATGGGAAGAC
N D I P M T G P L V A G G L L T V C Y V L T G R S A D L E L E R A A D V K W E D

4330 4340 4350 4360 4370 4380 4390 4400 4410 4420 4430 4440

CAGGCAGAGATATCAGGAAGCAGTCCAATCCTGTCAATAACAATATCAGAAGATGGTAGCATGTCGATAAAAAATGAAGAGGAAGAACAACACTGACCATACTCATTAGAACAGGATTG
Q A E I S G S S P I L S I T I S E D G S M S I K N E E E E Q T L T I L I R T G L

>NS3

4450 4460 4470 4480 4490 4500 4510 4520 4530 4540 4550 4560
CTGGTGATCTCAGGACTTTTCTCTGTATCAATACCAATCACGGCAGCAGCATGGTACCTGTGGGAAGTGAAGAAACAACGGGCCGGAGTATTGTGGGATGTTCTTCCACCCCAACCATG
L V I S G L F P V S I P I T A A A W Y L W E V K K Q R A G V L W D V P S P P P M

4570 4580 4590 4600 4610 4620 4630 4640 4650 4660 4670 4680
GGAAAGGCTGAAGTGAAGATGGAGCCTATAGAATTAAGCAAAAAGGGATTCTTGGATATTCAGATCGGAGCCGGAGTTTACAAAGAAGGAACATCCATACAATGTGGCATGTCACA
G K A E L E D G A Y R I K Q K G I L G Y S Q I G A G V Y K E G T F H T M W H V T

4690 4700 4710 4720 4730 4740 4750 4760 4770 4780 4790 4800
CGTGGCGTGTCTAATGCATAAAGGAAAGAGGATTGAACCATCATGGCGGACGTCAGAAAGACCTAATATCATATGGAGGAGGCTGGAAGTTAGAAGGAGAATGGAAGGAAGGAGAA
R G A V L M H K G K R I E P S W A D V K K D L I S Y G G G W K L E G E W K E G E

4810 4820 4830 4840 4850 4860 4870 4880 4890 4900 4910 4920
GAAGTCCAGGTATTGGCACTGGAGCCTGGAAAAATCCAAGAGCCGTCACAAACGAAACCTGGTCTTTTCAAACCAACGCCGGAACAATAGGTGCTGTATCTCTGGACTTTTCTCTCTGGA
E V Q V L A L E P G K N P R A V Q T K P G L F K T N A G T I G A V S L D F S P G

4930 4940 4950 4960 4970 4980 4990 5000 5010 5020 5030 5040
ACGTGAGATCTCCAATTATCGACAAAAAGGAAAAGTTGTGGTCTTTATGGTAATGGTGTGTTACAGGAGTGAGCATATGTGAGTGTATAGCCAGACTGAAAAAGCATTGAA
T S G S P I I D K K G K V V G L Y G N G V V T R S G A Y V S A I A Q T E K S I E

5050 5060 5070 5080 5090 5100 5110 5120 5130 5140 5150 5160
GACAACCCAGAGATCGAAGATGACATTTTCCGAAAGAGAAGACTGACCATCATGGACCTCCACCCAGGAGCGGGAAAGACGAAGAGATACCTTCCGGCCATAGTCAGAGAAGCTATAAAA
D N P E I E D D I F R K R R L T I M D L H P G A G K T K R Y L P A I V R E A I K

5170 5180 5190 5200 5210 5220 5230 5240 5250 5260 5270 5280
CGGGGTTTGAGAACATTAATCTTGGCCCCACTAGAGTTGTGGCAGCTGAAATGGAGGAAGCCCTTAGAGGACTTCCAATAAGATACCAGACCCAGCCATCAGAGCTGTGCACACCGGG
R G L R T L I L A P T R V V A A E M E E A L R G L P I R Y Q T P A I R A V H T G I

D2 PDK-53 NS3-250-Val locus giảm độc lực (D2 16681: Glu, nt-5270-A)

5290 5300 5310 5320 5330 5340 5350 5360 5370 5380 5390 5400
CGGGAGATTGTGGACCTAATGTGTCATGCCACATTTACCATGAGGCTGCTATCACCAGTTAGAGTGCCAACTACAACCTGATTATCATGGACGAAGCCCATTTACAGACCCAGCAAGT
R E I V D L M C H A T F T M R L L S P V R V P N Y N L I I M D E A H F T D P A S

5410 5420 5430 5440 5450 5460 5470 5480 5490 5500 5510 5520
ATAGCAGCTAGAGGATACATCTCAACTCGAGTGGAGATGGGTGAGGCAGCTGGGATTTTATGACAGCCACTCCCCGGGAAGCAGAGACCCATTTCTCAGAGCAATGCACCAATCATA
I A A R G Y I S T R V E M G E A A G I F M T A T P P G S R D P F P Q S N A P I I

5530 5540 5550 5560 5570 5580 5590 5600 5610 5620 5630 5640
GATGAAGAAAGAGAAATCCCTGAACGCTCGTGGAAATCCGGACATGAATGGGTGACGGATTTTAAAGGGAAGACTGTTTGGTTCGTTCCAAGTATAAAAGCAGGAAATGATATAGCAGCT
D E E R E I P E R S W N S G H E W V T D F K G K T V W F V P S I K A G N D I A A I

Đột biến cãm D2 PDK-53 nt-5547-C (D2 16681: T)

5650 5660 5670 5680 5690 5700 5710 5720 5730 5740 5750 5760
 TGCCTGAGGAAAAATGGAAGAAAGTGATACAACCTCAGTAGGAAGACCTTTGATTCTGAGTATGTCAAGACTAGAACCAATGATTGGGACTTCGTGGTTACAACCTGACATTTCAGAAATG
 C L R K N G K K V I Q L S R K T F D S E Y V K T R T N D W D F V V T T D I S E M

5770 5780 5790 5800 5810 5820 5830 5840 5850 5860 5870 5880
 GGTGCCAATTTCAAGGCTGAGAGGGTTATAGACCCAGACGCTGCATGAAACCAGTCATACTAACAGATGGTGAAGAGCGGGTGATTCTGGCAGGACCTATGCCAGTGACCCACTCTAGT
 G A N F K A E R V I D P R R C M K P V I L T D G E E R V I L A G P M P V T H S S

5890 5900 5910 5920 5930 5940 5950 5960 5970 5980 5990 6000
 GCAGCACAAAGAAGAGGGAGAATAGGAAGAAATCCAAAAATGAGAATGACCAGTACATATACATGGGGGAACCTCTGGAAAAATGATGAAGACTGTGCACACTGGAAAGAAGCTAAAATG
 A A Q R R G R I G R N P K N E N D Q Y I Y M G E P L E N D E D C A H W K E A K M

6010 6020 6030 6040 6050 6060 6070 6080 6090 6100 6110 6120
 CTCCTAGATAACATCAACACGCCAGAAGGAATCATTCTAGCATGTTTGAACCCAGAGCGTGAAAAGGTGGATGCCATTGATGGCGAATACCGCTTGAGAGGAGAAGCAAGGAAAACCTTT
 L L D N I N T P E G I I P S M F E P E R E K V D A I D G E Y R L R G E A R K T F

6130 6140 6150 6160 6170 6180 6190 6200 6210 6220 6230 6240
 GTAGACTTAATGAGAAGAGGAGACCTACCAGTCTGGTTGGCCTACAGAGTGGCAGCTGAAGGCATCAACTACGCAGACAGAAGGTGGTGTGTTTGATGGAGTCAAGAACAACCAATCCTA
 V D L M R R G D L P V W L A Y R V A A E G I N Y A D R R W C F D G V K N N Q I L

6250 6260 6270 6280 6290 6300 6310 6320 6330 6340 6350 6360
 GAAGAAAACGTGGAAGTTGAAATCTGGACAAAAGAAGGGGAAAGGAAGAAATTGAAACCCAGATGGTTGGATGCTAGGATCTATTCTGACCCACTGGCGCTAAAGAATTTAAGGAATTT
 E E N V E V E I W T K E G E R K K L K P R W L D A R I Y S D P L A L K E F K E F

>NS4A

6370 6380 6390 6400 6410 6420 6430 6440 6450 6460 6470 6480
 GCAGCCGGAAGAAAGTCTCTGACCCCTGAACCTAATCACAGAAATGGGTAGGCTCCCAACCTTCATGACTCAGAAGGCAAGAGACGCACCTGGACAACCTAGCAGTGTGCACACGGCTGAG
 A A G R K S L T L N L I T E M G R L P T F M T Q K A R D A L D N L A V L H T A E

6490 6500 6510 6520 6530 6540 6550 6560 6570 6580 6590 6600
 GCAGGTGAAGGGCGTACAACCATGCTCTCAGTGAACCTGCCGAGACCCCTGGAGACATTGCTTTTACTGACACTTCTGGCTACAGTCACGGGAGGGATCTTTTATTCTTGATGAGCGCA
 A G G R A Y N H A L S E L P E T L E T L L L L T L L A T V T G G I F L F L M S **A** |

NS4A-75-Ala đặc hiệu D2 PDK-53 (wt D2 16681: Gly, nt-6599-G)

6610 6620 6630 6640 6650 6660 6670 6680 6690 6700 6710 6720
 AGGGGCATAGGGAAGATGACCCCTGGGAATGTGCTGCATAATCACGGCTAGCATCCTCCTATGGTACGCACAAATACAGCCACACTGGATAGCAGCTTCAATAATACTGGAGTTTTTTCTC
 R G I G K M T L G M C C I I T A S I L L W Y A Q I Q P H W I A A S I I L E F F L

>NS4B

6730 6740 6750 6760 6770 6780 6790 6800 6810 6820 6830 6840
 ATAGTTTTGCTTATTCCAGAACCTGAAAAACAGAGAACACCCCAAGACAACCAACTGACCTACGTTGTGCATAGCCATCCTCACAGTGGTGCCGCAACCATGGCAACGAGATGGGTTTC
 I V L L I P E P E K Q R T P Q D N Q L T Y V V I A I L T V V A A T M A N E M G F

6850 6860 6870 6880 6890 6900 6910 6920 6930 6940 6950 6960
 CTAGAAAAACGAAGAAAGATCTCGGATTGGGAAGCATTGCAACCCAGCAACCCGAGAGCAACATCCTGGACATAGATCTACGTCCTGCATCAGCATGGACGCTGTATGCCGTGGCCACA
 L E K T K K D L G L G S I A T Q Q P E S N I L D I D L R P A S A W T L Y A V A T

6970 6980 6990 7000 7010 7020 7030 7040 7050 7060 7070 7080
 ACATTGTTACACCAATGTTGAGACATAGCATTGAAAATTCCTCAGTGAATGTGTCCTAACAGCTATAGCCAACCAAGCCACAGTGTTAATGGGTCTCGGGAAAGGATGGCCATTGTCA
 T F V T P M L R H S I E N S S V N V S L T A I A N Q A T V L M G L G K G W P L S

7090 7100 7110 7120 7130 7140 7150 7160 7170 7180 7190 7200
 AAGATGGACATCGGAGTTCCCTTCTCGCCATTGGATGCTACTCACAGTCAACCCATAACTCTCACAGCAGCTCTTTTCTTATTGGTAGCACATTATGCCATCATAGGGCCAGGACTC
 K M D I G V P L L A I G C Y S Q V N P I T L T A A L F L L V A H Y A I I G P G L

7210 7220 7230 7240 7250 7260 7270 7280 7290 7300 7310 7320
 CAAGCAAAAGCAACCAGAGAAGCTCAGAAAAAGCAGCGCGGGCATCATGAAAAACCAACTGTCGATGGAATAACAGTGATTGACCTAGATCCAATACCTTATGATCCAAGTTTGAA
 Q A K A T R E A Q K R A A A G I M K N P T V D G I T V I D L D P I P Y D P K F E

7330 7340 7350 7360 7370 7380 7390 7400 7410 7420 7430 7440
 AAGCAGTTGGGACAAGTAATGCTCCTAGTCTCTGCGTGACTCAAGTATTGATGATGAGGACTACATGGGCTCTGTGTGAGGCTTTAACCTTAGCTACCGGGCCCATCTCCACATTGTGG
 K Q L G Q V M L L V L C V T Q V L M M R T T W A L C E A L T L A T G P I S T L W

7450 7460 7470 7480 7490 7500 7510 7520 7530 7540 7550 7560
 GAAGGAAATCCAGGAGGTTTTGGAACACTACCATTGCGGTGTCAATGGCTAACATTTTAGAGGGAGTTACTTGGCCGGAGCTGGACTTCTCTTTTCTATTATGAAGAACACAACCAAC
 E G N P G R F W N T T I A V S M A N I F R G S Y L A G A G L L F S I M K N T T N

>ns5

7570 7580 7590 7600 7610 7620 7630 7640 7650 7660 7670 7680
 ACAAGAAGGGGAATCGGCAACATAGGAGAGACGCTTGGAGAGAAATGGAAAAGCCGATTGAACGCATTGGGAAAAAGTGAATTCAGATCTACAAGAAAAGTGGAAATCCAGGAAGTGGAT
 T R R G T G N I G E T L G E K W K S R L N A L G K S E F Q I Y K K S G I Q E V D

7690 7700 7710 7720 7730 7740 7750 7760 7770 7780 7790 7800
 AGAACCTTAGCAAAAAGGACATTAAGAGAGAGAAACGGACCATCACGCTGTGTGCGGAGGCTCAGCAAAACTGAGATGGTTCGTTGAGAGAAACATGGTCACACCAGAAGGGAAAGTA
 R T L A K E G I K R G E T D H H A V S R G S A K L R W F V E R N M V T P E G K V

7810 7820 7830 7840 7850 7860 7870 7880 7890 7900 7910 7920
 GTGGACCTCGGTTGTGGCAGAGGAGGCTGGTCATACTATTGTGGAGGACTAAAGAATGTAAGAGAAGTCAAAGGCCTAACAAAAGGAGGACCAGGACACGAAGAACCCATCCCCATGTCA
 V D L G C G R G G W S Y Y C G G L K N V R E V K G L T K G G P G H E E P I P M S

7930 7940 7950 7960 7970 7980 7990 8000 8010 8020 8030 8040
 ACATATGGGTGGAATCTAGTGCCTTCAAAGTGGAGTTGACGTTTTCTTCATCCCGCCAGAAAAGTGTGACACATTATTGTGTGACATAGGGGAGTCATCACCAAATCCCACAGTGGAA
 T Y G W N L V R L Q S G V D V F F I P P E K C D T L L C D I G E S S P N P T V E

8050 8060 8070 8080 8090 8100 8110 8120 8130 8140 8150 8160
 GCAGGACGAACACTCAGAGTCCCTTAAGTCTAGTAGAAAATTGGTTGAACAACAACTCAATTTTGCATAAAGGTTCTCAACCCATATATGCCCTCAGTCATAGAAAAATGGAAGCACTA
 A G R T L R V L N L V E N W L N N N T Q F C I K V L N P Y M P S V I E K M E A L

8170 8180 8190 8200 8210 8220 8230 8240 8250 8260 8270 8280
 CAAAGGAAATATGGAGGAGCCTTAGTGAGGAATCCACTCTCACGAACTCCACACATGAGATGTACTGGGTATCCAATGCTTCCGGGAACATAGTGTATCAGTGAACATGATTTCAGG

Q R K Y G G A L V R N P L S R N S T H E M Y W V S N A S G N I V S S V N M I S R

8290 8300 8310 8320 8330 8340 8350 8360 8370 8380 8390 8400

ATGTTGATCAACAGATTTACAATGAGATACAAGAAAGCCACTTACGAGCCGGATGTTGACCTCGGAAGCGGAACCCGTAACATCGGGATTGAAAGTGAGATACCAAACCTAGATATAATT

M L I N R F T M R Y K K A T Y E P D V D L G S G T R N I G I E S E I P N L D I I

8410 8420 8430 8440 8450 8460 8470 8480 8490 8500 8510 8520

GGGAAAAGAAATAGAAAAATAAAGCAAGAGCATGAAACATCATGGCACTATGACCAAGACCACCCATACAAAACGTGGGCATACCATGGTAGCTATGAAACAAAACAGACTGGATCAGCA

G K R I E K I K Q E H E T S W H Y D Q D H P Y K T W A Y H G S Y E T K Q T G S A

8530 8540 8550 8560 8570 8580 8590 8600 8610 8620 8630 8640

TCATCCATGGTCAACGGAGTGGTCAGGCTGCTGACAAAACCTTGGGACGTCGTCCCATGGTGACACAGATGGCAATGACAGACAGACTCCATTGGACAACAGCGCGTTTTTAAAGAG

S S M V N G V V R L L T K P W D V V P M V T Q M A M T D T T P F G Q Q R V F K E

8650 8660 8670 8680 8690 8700 8710 8720 8730 8740 8750 8760

AAAGTGGACACGAGAACCCAAGAACCAGAAAGAGGCACGAAGAACTAATGAAAATAACAGCAGAGTGGCTTTGGAAAGAATTAGGGAAGAAAAAGACACCCAGGATGTGCACCAGAGAA

K V D T R T Q E P K E G T K K L M K I T A E W L W K E L G K K K T P R M C T R E

8770 8780 8790 8800 8810 8820 8830 8840 8850 8860 8870 8880

GAATTCACAAGAAAGGTGAGAAGCAATGCAGCCTTGGGGGCCGTATTCACTGATGAGAAACAAGTGAAGTCGGCACGTGAGGCTGTTGAAGATAGTAGGTTTTGGGAGCTGGTTGACAAG

E F T R K V R S N A A L G A V F T D E N K W K S A R E A V E D S R F W E L V D K I

Đột biến bổ sung NS5-412 Ile-to-Val (nt-8803 A-to-G) trong hạt chủ và tiền chủ

8890 8900 8910 8920 8930 8940 8950 8960 8970 8980 8990 9000

GAAAGGAATCTCCATCTTGAAGGAAAGTGTGAAACATGTGTGTACAACATGATGGGAAAAAGAGAGAAGAAGCTAGGGGAATTCGGCAAGGCAAAAGGCAGCAGAGCCATATGGTACATG

E R N L H L E G K C E T C V Y N M M G K R E K K L G E F G K A K G S R A I W Y M

9010 9020 9030 9040 9050 9060 9070 9080 9090 9100 9110 9120

TGGCTTGGAGCAGCTTCTTAGAGTTTGAAGCCCTAGGATTCTTAAATGAAGATCACTGGTTCTCCAGAGAGAACTCCCTGAGTGGAGTGAAGGAGAAGGGCTGCACAAGCTAGGTTAC

W L G A R F L E F E A L G F L N E D H W F S R E N S L S G V E G E G L H K L G Y

9130 9140 9150 9160 9170 9180 9190 9200 9210 9220 9230 9240

ATTCTAAGAGACGTGAGCAAGAAAGAGGGAGGAGCAATGTATGCCGATGACACCGCAGGATGGGATACAGAATCACACTAGAAGACCTAAAAAATGAAGAAATGGTAACAAACCACATG

I L R D V S K K E G G A M Y A D D T A G W D T R I T L E D L K N E E M V T N H M

9250 9260 9270 9280 9290 9300 9310 9320 9330 9340 9350 9360

GAAGGAGAACACAAGAAACTAGCCGAGGCCATTTTCAAACCTAACGTACCAAAACAGGTGGTGCCTGTGCAAGACCAACACCAAGAGGCACAGTAATGGACATCATATCGAGAAGAGAC

E G E H K K L A E A I F K L T Y Q N K V V R V Q R P T P R G T V M D I I S R R D

9370 9380 9390 9400 9410 9420 9430 9440 9450 9460 9470 9480

CAAAGAGGTAGTGACAAAGTTGGACCTATGGACTCAATACTTTACCAATATGAAGCCCAACTAATCAGACAGATGGAGGGAGAAGGAGTCTTTAAAGCATTGACACCTAACCAATC

Q R G S G Q V G T Y G L N T F T N M E A Q L I R Q M E G E G V F K S I Q H L T I

9490 9500 9510 9520 9530 9540 9550 9560 9570 9580 9590 9600

ACAGAAGAAATCGCTGTGCAAAACTGGTTAGCAAGAGTGGGGCGCGAAAGGTTATCAAGAATGGCCATCAGTGGAGATGATTGTGTTGTGAAACCTTTAGATGACAGGTTTCGCAAGCGCT

T E E I A V Q N W L A R V G R E R L S R M A I S G D D C V V K P L D D R F A S A
 9610 9620 9630 9640 9650 9660 9670 9680 9690 9700 9710 9720
 TTAACAGCTCTAAATGACATGGGAAAGATTAGGAAAGACATACAACAATGGGAACCTTCAAGAGGATGGAATGATTGGACACAAGTGCCCTTCTGTTACACCATTTCATGAGTTAATC
 L T A L N D M G K I R K D I Q Q W E P S R G W N D W T Q V P F C S H H F H E L I
 9730 9740 9750 9760 9770 9780 9790 9800 9810 9820 9830 9840
 ATGAAAGACGGTCGCGTACTCGTTGTTCCATGTAGAAACCAAGATGAAGTATTGGCAGAGCCCGAATCTCCCAAGGAGCAGGGTGGTCTTTGCGGGAGACGGCCTGTTGGGGAAGTCT
 M K D G R V L V V P C R N Q D E L I G R A R I S Q G A G W S L R E T A C L G K S
 9850 9860 9870 9880 9890 9900 9910 9920 9930 9940 9950 9960
 TACGCCCAAATGTGGAGCTTGATGTACTTCCACAGACGCGACCTCAGGCTGGCGGCAAATGCTATTTGCTCGGCAGTACCATCACATTGGGTTCACAACAAGTCGAACAACCTGGTCCATA
 Y A Q M W S L M Y F H R R D L R L A A N A I C S A V P S H W V P T S R T T W S I
 9970 9980 9990 10000 10010 10020 10030 10040 10050 10060 10070 10080
 CATGCTAAACATGAATGGATGACAACGGAAGACATGCTGACAGTCTGGAACAGGGTGTGGATTCAAGAAAAACCATGGATGGAAGACAAAACCTCCAGTGAATCATGGGAGGAAATCCCA
 H A K H E W M T T E D M L T V W N R V W I Q E N P W M E D K T P V E S W E E I P
 10090 10100 10110 10120 10130 10140 10150 10160 10170 10180 10190 10200
 TACTTGGGAAAAGAGAAGACCAATGGTTCGCGCTCATTGATGGGTAAACAAGCAGGCCACCTGGGCAAAGAACATCCAAGCAGCAATAAATCAAGTTAGATCCCTTATAGGCAATGAA
 Y L G K R E D Q W C G S L I G L T S R A T W A K N I Q A A I N Q V R S L I G N E
 >3'-NC
 10210 10220 10230 10240 10250 10260 10270 10280 10290 10300 10310 10320
 GAATACACAGATTACATGCCATCCATGAAAAGATTGAGAAGAGAAGAGGAAGAAGCAGGAGTTCTGTGGTAGAAAAGCAAACCTAACATGAAACAAGGCTAGAAGTCAGGTCGGATTAAAGC
 E Y T D Y M P S M K R F R R E E E E A G V L W
 10330 10340 10350 10360 10370 10380 10390 10400 10410 10420 10430 10440
 CATAGTACGGAAAAAATATGCTACCTGTGAGCCCCGTCCAAGGACGTTAAAAGAAGTCAGGCCATCATAAATGCCATAGCTTGAGTAAACTATGCAGCCTGTAGTCCACCTGAGAAGG
 10450 10460 10470 10480 10490 10500 10510 10520 10530 10540 10550 10560
 TGTAAAAAATCCGGGAGGCCACAAACCATGGAAGCTGTACGCATGGCGTAGTGGACTAGCGGTTAGAGGAGACCCCTCCCTTACAAATCGCAGCAACAATGGGGGCCCCAAGGCGAGATGA
 10570 10580 10590 10600 10610 10620 10630 10640 10650 10660 10670 10680
 AGCTGTAGTCTCGCTGGAAGGACTAGAGGTTAGAGGAGACCCCCCGAAACAAAAACAGCATATTGACGCTGGGAAAGACCAGAGATCCTGCTGTCTCCTCAGCATCATTCAGGCACA
 10690 10700 10710 10720
 GAACGCCAGAAAATGGAATGGTGCTGTTGAATCAACAGGTTCT

Hạt virut chủ DENvax-3 (MVS)

Trình tự nucleotit của hệ gen virut khảm và gọt ra trình tự axit amin của protein được dịch mã được nêu trong bản mô tả này. Hầu hết gen prM-E (nt- 457 thành -2373, được gạch dưới) là (wt) DEN-3 16562 kiểu hoang dại đặc hiệu virut; trình tự nucleotit

còn lại là DEN-2 PDK-53 đặc hiệu virus. protein E của virus DEN-3 có ít hơn hai axit amin so với protein E của DEN-2. Do đó, vị trí nt bắt đầu từ NgoMIV là 6 nt ít hơn vị trí DEN-2 PDK-53 nt ban đầu. Tất cả thay thế được thao tác khác với virus kiểu hoang dại (DEN-3 16562 hoặc DEN-2 16681), cũng như các đột biến bổ sung (các thay đổi từ dòng ADN thao tác) có bán trên thị trường.

Các thay thế được bao gồm trong hệ gen và protein:

Vị trí ghép nối:

- a. M1ul (nt 451-456): đột biến câm được thao tác, nt-453 A-thành-G
- b. NgoM IV (nt 2374-2379): đột biến được thao tác, nt-2375/2376 TG-thành-CC (dẫn đến thay đổi E-480 Val-thành-Ala)

Khung chính virus D2 PDK-53 (thay đổi từ D2 16681 kiểu hoang dại): nét đậm

- a. vùng không mã hoá 5' (NCR)-57 (nt-57 C-thành-T): locus giảm độc lực chính (màu đỏ)
- b. NS1-53 Gly-thành-Asp (nt-2573 G-thành-A): locus giảm độc lực chính (màu đỏ)
- c. NS2A-181 Leu-thành-Phe (nt-4012 C-thành-T)
- d. NS3-250 Glu-thành-Val (nt-5264 A-thành-T): locus giảm độc lực chính (màu đỏ)
- e. đột biến câm nt-5541 (gen NS3) T-thành-C
- f. NS4A-75 Gly-thành-Ala (nt-6593 G-thành-C)

* đột biến câm nt-8565 C-thành-T của PDK-53 là không được thao tác trong virus vaccine

Đột biến được thao tác trong DEN-3 prM-E (thay đổi từ D3 16562 kiểu hoang dại)

- a. đột biến câm nt-552 C-thành-T được thao tác: gen đánh dấu dòng
- b. E-345 His-thành-Leu (nt-1970 A-thành-T) để sao chép hiệu quả trong môi trường nuôi cấy

Các thay thế bổ sung được tìm thấy trong hạt vaccine (0,02% nt khác với dòng ban đầu)

- a. đột biến E-223 Thr-thành-Ser (nt-1603 A-thành-T, nét đậm)
- b. đột biến câm nt-7620 A-thành-G (nét đậm)

NCR-57-T, D2 PDK-53 locus giảm độc lực (wt D2 16681:C)

> 5'-vùng không mã hoá|> C

```

      10      20      30      40      50      60      70      80      90     100
AGTTGTTAGTCTACGTGGACCGACAAAGACAGATTCTTTGAGGGAGCTAAGCTCAATGTTAGTTCTAACAGTTTTTTAATTAGAGAGCAGATCTCTGATGA
                                                                 M  N
      110     120     130     140     150     160     170     180     190     200
ATAACCAACGGAAAAAGGCGAAAAACACGCCTTTCAATATGCTGAAACGCGAGAGAAACCGCGTGTGCGACTGTGCAACAGCTGACAAAGAGATTCTCACT
  N  Q  R  K  K  A  K  N  T  P  F  N  M  L  K  R  E  R  N  R  V  S  T  V  Q  Q  L  T  K  R  F  S  L
      210     220     230     240     250     260     270     280     290     300
TGGAAATGCTGCAGGGACGAGGACCATTAAACTGTTTCATGGCCCTGGTGGCGTTCCCTTCGTTTCCTAACAATCCCACCAACAGCAGGGATATTGAAGAGA
  G  M  L  Q  G  R  G  P  L  K  L  F  M  A  L  V  A  F  L  R  F  L  T  I  P  P  T  A  G  I  L  K  R
      310     320     330     340     350     360     370     380     390     400
TGGGGAACAATTAAAAAATCAAAAGCTATTAAATGTTTTGAGAGGGTTCAGGAAAGAGATTGGAAGGATGCTGAACATCTTGAATAGGAGACGCAGATCTG
  W  G  T  I  K  K  S  K  A  I  N  V  L  R  G  F  R  K  E  I  G  R  M  L  N  I  L  N  R  R  R  R  S  A

```

> prM bắt đầu trình tự D3 16562

```

      410     420     430     440     450     460     470     480     490     500
CAGGCATGATCATTATGCTGATTCCAACAGTGATGGCGTTCCATTTAACCACGCGTGATGGAGAGCCGCGCATGATTGTGGGGAAGAATGAAAGAGGAAA
  G  M  I  I  M  L  I  P  T  V  M  A  F  H  L  T  T  R  D  G  E  P  R  M  I  V  G  K  N  E  R  G  K  |

```

Vị ghép nối MluI được thao tác (đột biến câm nt-453 A-thành-G)

```

      510     520     530     540     550     560     570     580     590     600
ATCCCTACTTTTCAAGACAGCCTCTGGAATCAACATGTGCACACTCATAGCTATGGATCTGGGAGAGATGTGTGATGACACGGTCACTTACAAATGCCCC
  S  L  L  F  K  T  A  S  G  I  N  M  C  T  L  I  A  M  D  L  G  E  M  C  D  D  T  V  T  Y  K  C  P  |

```

Đột biến câm C-thành-T nt như dòng đánh dấu

```

      610     620     630     640     650     660     670     680     690     700
CACATTACCGAAGTGGAGCCTGAAGACATTGACTGCTGGTGCAACCTTACATCGACATGGGTGACTTATGGAACATGCAATCAAGCTGGAGAGCATAGAC
  H  I  T  E  V  E  P  E  D  I  D  C  W  C  N  L  T  S  T  W  V  T  Y  G  T  C  N  Q  A  G  E  H  R  R

```

> M

```

      710     720     730     740     750     760     770     780     790     800
GCGATAAGAGATCAGTGGCGTTAGCTCCCCATGTTGGCATGGGACTGGACACACGCACTCAAACCTGGATGTCGGCTGAAGGAGCTTGGAGACAAGTCGA

```

```

  D  K  R  S  V  A  L  A  P  H  V  G  M  G  L  D  T  R  T  Q  T  W  M  S  A  E  G  A  W  R  Q  V  E
      810     820     830     840     850     860     870     880     890     900
GAAGGTAGAGACATGGGCCCTTAGGCACCCAGGGTTTACCATACTAGCCCTATTTCTTGCCCATTACATAGGCACTTCCTTGACCCAGAAAGTGGTTATT
  K  V  E  T  W  A  L  R  H  P  G  F  T  I  L  A  L  F  L  A  H  Y  I  G  T  S  L  T  Q  K  V  V  I

```

> E

```

      910     920     930     940     950     960     970     980     990    1000
TTTATACTATTAATGCTGGTTACCCCATCCATGACAATGAGATGTGTAGGAGTAGGAAACAGAGATTTTGTGGAAGGCCTATCGGGAGCTACGTGGGTTG

```

```

  F  I  L  L  M  L  V  T  P  S  M  T  M  R  C  V  G  V  G  N  R  D  F  V  E  G  L  S  G  A  T  W  V  D
      1010     1020     1030     1040     1050     1060     1070     1080     1090     1100

```

```

ACGTGGTGCTCGAGCACGGTGGGTGTGTGACTACCATGGCTAAGAACAAGCCCACGCTGGACATAGAGCTTCAGAAGACCGAGGCCACCCAACTGGCGAC
  V  V  L  E  H  G  G  C  V  T  T  M  A  K  N  K  P  T  L  D  I  E  L  Q  K  T  E  A  T  Q  L  A  T

```

1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
CCTAAGGAAGCTATGCATTGAGGGAAAAATTACCAACATAACAACCGACTCAAGATGTCCCACCCAAGGGGAAGCGATTTTACCTGAGGAGCAGGACCAG
 L R K L C I E G K I T N I T T D S R C P T Q G E A I L P E E Q D Q
 1210 1220 1230 1240 1250 1260 1270 1280 1290 1300
AACTACGTGTGTAAGCATACATACGTGGACAGAGGCTGGGGAAACGGTTGTGGTTTGTGGCAAGGGAAGCTTGGTGACATGCGCGAAATTTCAATGTT
 N Y V C K H T Y V D R G W G N G C G L F G K G S L V T C A K F Q C L
 1310 1320 1330 1340 1350 1360 1370 1380 1390 1400
TAGAATCAATAGAGGGAAAAGTGGTGCAACATGAGAACCTCAAATACACCGTCATCATCACAGTGCACACAGGAGACCAACACCAGGTGGGAAATGAAAC
 E S I E G K V V Q H E N L K Y T V I I T V H T G D Q H Q V G N E T
 1410 1420 1430 1440 1450 1460 1470 1480 1490 1500
GCAGGGAGTCACGGCTGAGATAACACCCCAGGCATCAACCGCTGAAGCCATTTTACCTGAATATGGAACCCCTCGGGCTAGAATGCTCACCACGGACAGGT
 Q G V T A E I T P Q A S T A E A I L P E Y G T L G L E C S P R T G
 1510 1520 1530 1540 1550 1560 1570 1580 1590 1600
TTGGATTTCAATGAAATGATCTCATTGACAATGAAGAACAAGCATGGATGGTACATAGACAATGGTTCTTTGACTTACCCCTACCATGGACATCAGGAG
 L D F N E M I S L T M K N K A W M V H R Q W F F D L P L P W T S G A
 1610 1620 1630 1640 1650 1660 1670 1680 1690 1700
CTTCAGCAGAAACACCAACTTGGAACAGGAAAGAGCTTCTTGTGACATTTAAAAATGCACATGCAAAAAAGCAAGAAGTAGTTGTTCTTGGATCACAAGA
 S A E T P T W N R K E L L V T F K N A H A K K Q E V V V L G S Q E |
Đột biến E-233 Thr-thành-Ser bổ sung (wt D3 16562: nt-1603 A) trong hạt chủ và tiền chủ
 1710 1720 1730 1740 1750 1760 1770 1780 1790 1800
GGGAGCAATGCATACGCACTGACAGGAGCTACAGAGATCCAAACCTCAGGAGGCACAAGTATCTTTGCGGGGCACTTAAATGTAGACTCAAGATGGAC
 G A M H T A L T G A T E I Q T S G G T S I F A G H L K C R L K M D
 1810 1820 1830 1840 1850 1860 1870 1880 1890 1900
AAATTGGAACCTCAAGGGGATGAGCTATGCAATGTGCTTGAGTAGCTTTGTGTTGAAGAAAGAAGTCTCCGAAACGCAGCATGGGACAATACTCATTAAGG
 K L E L K G M S Y A M C L S S F V L K K E V S E T Q H G T I L I K V
 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000
TTGAGTACAAAGGGGAAGATGCACCCTGCAAGATTCCTTTCTCCACGGAGGATGGACAAGGAAAAGCTCTCAATGGCAGACTGATCACAGCCAATCCAGT
 E Y K G E D A P C K I P F S T E D G Q G K A L N G R L I T A N P V L

E-345 His-to-Leu (wt D3 16562: nt-1970-A) được thao tác để sinh trưởng hữu hiệu

2010 2020 2030 2040 2050 2060 2070 2080 2090 2100
GGTGACCAAGAAGGAGGAGCCTGTCAACATTGAGGCTGAACCTCCTTTTGGAGAAAAGTAACATAGTAATTGGAATTGGAGACAAAGCCCTGAAAATCAAC
 V T K K E E P V N I E A E P P F G E S N I V I G I G D K A L K I N
 2110 2120 2130 2140 2150 2160 2170 2180 2190 2200
TGGTACAAGAAGGGAAGCTCGATTGGGAAGATGTTTCGAGGCCACTGCCAGAGGTGCAAGGCGCATGGCCATCTTGGGAGACACAGCCTGGGACTTTGGAT
 W Y K K G S S I G K M F E A T A R G A R R M A I L G D T A W D F G S
 2210 2220 2230 2240 2250 2260 2270 2280 2290 2300
CAGTGGGTGGTGTGTTTGAATTCATTAGGGGAAAATGGTCCACCAAATATTTGGGAGTGCTTACACAGCCCTATTTGGTGGAGTCTCCTGGATGATGAAAAT
 V G G V L N S L G K M V H Q I F G S A Y T A L F G G V S W M M K I

Kết thúc trình tự **D3 16562**

2310 2320 2330 2340 2350 2360 2370 | 2380 2390 2400
 TGGAAATAGGTGTCCTCTTAACCTGGATAGGGTTGAACTCAAAAAATACTTCTATGTCATTTTCATGCATCGCGGCCGGCATTGTGACACTGTATTGGGA
 G I G V L L T W I G L N S K N T S M S F S C I A A G I V T L Y L G

Vị trí ghép nối **NgoMIV được thao tác, E-480 Val-thành-Ala (nt-2375/2376 TG-thành-CC)**

> **NS1**

2410 2420 2430 2440 2450 2460 2470 2480 2490 2500
 GTCATGGTGCAGGCCGATAGTGGTTGCGTTGTGAGCTGGAAAAACAAAGAACTGAAATGTGGCAGTGGGATTTTCATCAGACAACTGCACACATGGA
 V M V Q A D S G C V V S W K N K E L K C G S G I F I T D N V H T W T
 2510 2520 2530 2540 2550 2560 2570 2580 2590 2600
 CAGAACAAATACAAGTTCCAACCAGAATCCCCCTCAAACTAGCTTCAGCTATCCAGAAAGCCCATGAAGAGGACATTTGTGGAATCCGCTCAGTAACAAG
 E Q Y K F Q P E S P S K L A S A I Q K A H E E D I C G I R S V T R L

D2 PDK-53 NS1-53-Asp locus giảm độc lực (wt D2 16681: Gly, nt-2573-G)

2610 2620 2630 2640 2650 2660 2670 2680 2690 2700
 ACTGGAGAATCTGATGTGGAAACAAATAACACCAGAATTGAATCACATTCTATCAGAAAATGAGGTGAAGTTAACTATTATGACAGGAGACATCAAAGGA
 L E N L M W K Q I T P E L N H I L S E N E V K L T I M T G D I K G
 2710 2720 2730 2740 2750 2760 2770 2780 2790 2800
 ATCATGCAGGCAGGAAAACGATCTCTGCGGCCTCAGCCCACTGAGCTGAAGTATTCATGGAAAACATGGGGCAAAGCAAAAATGCTCTCTACAGAGTCTC
 I M Q A G K R S L R P Q P T E L K Y S W K T W G K A K M L S T E S H
 2810 2820 2830 2840 2850 2860 2870 2880 2890 2900
 ATAACCAGACCTTTCTCATTGATGGCCCCGAAACAGCAGAATGCCCCAACACAAATAGAGCTTGAATTGTTGGAAGTTGAAGACTATGGCTTTGGAGT
 N Q T F L I D G P E T A E C P N T N R A W N S L E V E D Y G F G V
 2910 2920 2930 2940 2950 2960 2970 2980 2990 3000
 ATTCACCACCAATATATGGCTAAAATTGAAAGAAAAACAGGATGTATTCTGCGACTCAAACTCATGTCTAGCGGCCATAAAAGACAACAGAGCCGTCCAT
 F T T N I W L K L K E K Q D V F C D S K L M S A A I K D N R A V H
 3010 3020 3030 3040 3050 3060 3070 3080 3090 3100
 GCCGATATGGGTTATTGGATAGAAAGTGCCTCAATGACACATGGAAGATAGAGAAAGCCTCTTTTCATTGAAGTTAAAACTGCCACTGGCCAAAATCAC
 A D M G Y W I E S A L N D T W K I E K A S F I E V K N C H W P K S H
 3110 3120 3130 3140 3150 3160 3170 3180 3190 3200
 ACACCTCTGGAGCAATGGAGTGCTAGAAAGTGAGATGATAATTCCAAAGAATCTCGCTGGACCAGTGTCTCAACACAACATATAGACCAGGCTACCATAC
 T L W S N G V L E S E M I I P K N L A G P V S Q H N Y R P G Y H T
 3210 3220 3230 3240 3250 3260 3270 3280 3290 3300
 ACAAATAACAGGACCATTGGCATCTAGGTAAGCTTGAGATGGACTTTGATTTCTGTGATGGAACAACAGTGGTAGTGACTGAGGACTGCGGAAATAGAGGA
 Q I T G P W H L G K L E M D F D F C D G T T V V V T E D C G N R G

3310 3320 3330 3340 3350 3360 3370 3380 3390 3400
 CCCTCTTTGAGAACAACCACTGCCTCTGGAACACTCATAACAGAATGGTGTGCCGATCTTGCACATTACCACCGCTAAGATACAGAGGTGAGGATGGGT
 P S L R T T T A S G K L I T E W C C R S C T L P P L R Y R G E D G C

> **NS2A**

3410 3420 3430 3440 3450 3460 3470 3480 3490 3500

GCTGGTACGGGATGGAAATCAGACCATTGAAGGAGAAAGAAGAGAATTTGGTCAACTCCTTGGTCACAGCTGGACATGGGCAGGTCGACAACTTTTCACT
 W Y G M E I R P L K E K E E N L V N S L V T A G H G Q V D N F S L
 3510 3520 3530 3540 3550 3560 3570 3580 3590 3600
 AGGAGTCTTGGGAATGGCATTGTTCCCTGGAGGAAATGCTTAGGACCCGAGTAGGAACGAAACATGCAATACTACTAGTTGCAGTTTCTTTTGTGACATTG
 G V L G M A L F L E E M L R T R V G T K H A I L L V A V S F V T L
 3610 3620 3630 3640 3650 3660 3670 3680 3690 3700
 ATCACAGGGAACATGTCCTTTAGAGACCTGGGAAGAGTGATGGTTATGGTAGGCCCACTATGACGGATGACATAGGTATGGGCGTGACTTATCTTGCCC
 I T G N M S F R D L G R V M V M V G A T M T D D I G M G V T Y L A L
 3710 3720 3730 3740 3750 3760 3770 3780 3790 3800
 TACTAGCAGCCTTCAAAGTCAGACCAACTTTTGCAGCTGGACTACTCTTGAGAAAGCTGACCTCCAAGGAATTGATGATGACTACTATAGGAATTGTACT
 L A A F K V R P T F A A G L L L R K L T S K E L M M T T I G I V L
 3810 3820 3830 3840 3850 3860 3870 3880 3890 3900
 CCTCTCCCAGAGCACCATAACCAGAGACCATTTCTTGAGTTGACTGATGCGTTAGCCTTAGGCATGATGGTCCTCAAAATGGTGAGAAATATGGAAAAGTAT
 L S Q S T I P E T I L E L T D A L A L G M M V L K M V R N M E K Y
 3910 3920 3930 3940 3950 3960 3970 3980 3990 4000
 CAATTGGCAGTGACTATCATGGCTATCTTGTGCGTCCCAAACGCAGTGATATTACAAAACGCATGGAAAGTGAGTTGCACAATATTGGCAGTGGTGTCCG
 Q L A V T I M A I L C V P N A V I L Q N A W K V S C T I L A V V S V
 4010 4020 4030 4040 4050 4060 4070 4080 4090 4100
 TTTCCCCACTGTTCTTAACATCCTCACAGCAAAAACAGATTGGATACCATTAGCATTGACGATCAAAGGTCTCAATCCAACAGCTATTTTCTAACAAC
 S P L F L T S S Q Q K T D W I P L A L T I K G L N P T A I F L T T L

NS2A-181-Phe đặc hiệu D2 PDK-53 (wt D2 16681: Leu, nt-4012-C)

> NS2B

4110 4120 4130 4140 4150 4160 4170 4180 4190 4200
 CCTCTCAAGAACCAGCAAGAAAAGGAGCTGGCCATTAAATGAGGCTATCATGGCAGTCGGGATGGTGAGCATTTTAGCCAGTTCTCTCCTAAAAAATGAT
 L S R T S K K R S W P L N E A I M A V G M V S I L A S S L L K N D
 4210 4220 4230 4240 4250 4260 4270 4280 4290 4300
 ATTCCCATGACAGGACCATTAGTGGCTGGAGGGCTCCTCACTGTGTGCTACGTGCTCACTGGACGATCGGCCGATTGGAAGTGGAGAGAGCAGCCGATG
 I P M T G P L V A G G L L T V C Y V L T G R S A D L E L E R A A D V
 4310 4320 4330 4340 4350 4360 4370 4380 4390 4400
 TCAAATGGGAAGACCAGGCAGAGATATCAGGAAGCAGTCCAATCCTGTCAATAACAATATCAGAAGATGGTAGCATGTGCGATAAAAAATGAAGAGGAAGA
 K W E D Q A E I S G S S P I L S I T I S E D G S M S I K N E E E E
 4410 4420 4430 4440 4450 4460 4470 4480 4490 4500
 ACAAACACTGACCATACTCATTAGAACAGGATTGCTGGTGATCTCAGGACTTTTCTGTATCAATACCAATCACGGCAGCAGCATGGTACCTGTGGGAA
 Q T L T I L I R T G L L V I S G L F P V S I P I T A A A W Y L W E

> NS3

4510 4520 4530 4540 4550 4560 4570 4580 4590 4600
 GTGAAGAAACAACGGGCCGGAGTATTGTGGGATGTTCTTCACCCCCACCATGGGAAAGGCTGAACTGGAAGATGGAGCCTATAGAATTAAGCAAAAAAG
 V K K Q R A G V L W D V P S P P P M G K A E L E D G A Y R I K Q K G

4610 4620 4630 4640 4650 4660 4670 4680 4690 4700
 GGATTCTTGGATATTCCCAGATCGGAGCCGGAGTTTACAAAGAAGGAACATTCCATACAATGTGGCATGTCACACGTGGCGCTGTTCTAATGCATAAAGG
 I L G Y S Q I G A G V Y K E G T F H T M W H V T R G A V L M H K G
 4710 4720 4730 4740 4750 4760 4770 4780 4790 4800
 AAAGAGGATTGAACCATCATGGGCGGACGTCAAGAAAGACCTAATATCATATGGAGGAGGCTGGAAGTTAGAAGGAGAATGGAAGGAAGGAGAAGAAGTC
 K R I E P S W A D V K K D L I S Y G G G W K L E G E W K E G E E V
 4810 4820 4830 4840 4850 4860 4870 4880 4890 4900
 CAGGTATTGGCACTGGAGCCTGGAAAAAATCCAAGAGCCGTCCAAACGAAACCTGGTCTTTTCAAACCAACGCCGGAACAATAGGTGCTGTATCTCTGG
 Q V L A L E P G K N P R A V Q T K P G L F K T N A G T I G A V S L D
 4910 4920 4930 4940 4950 4960 4970 4980 4990 5000
 ACTTTTCTCCTGGAACGTGAGGATCTCCAATTATCGACAAAAAAGGAAAAGTTGTGGGTCTTTATGGTAATGGTGTGTGTACAAGGAGTGGAGCATATGT
 F S P G T S G S P I I D K K G K V V G L Y G N G V V T R S G A Y V
 5010 5020 5030 5040 5050 5060 5070 5080 5090 5100
 GAGTGCTATAGCCCAGACTGAAAAAGCATTGAAGACAACCCAGAGATCGAAGATGACATTTTCCGAAAGAGAAGACTGACCATCATGGACCTCCACCCA
 S A I A Q T E K S I E D N P E I E D D I F R K R R L T I M D L H P
 5110 5120 5130 5140 5150 5160 5170 5180 5190 5200
 GGAGCGGGAAGACGAAGAGATACCTTCCGGCCATAGTCAGAGAAGCTATAAAACGGGGTTTGAGAACATTAATCTTGGCCCCACTAGAGTTGTGGCAG
 G A G K T K R Y L P A I V R E A I K R G L R T L I L A P T R V V A A
 5210 5220 5230 5240 5250 5260 5270 5280 5290 5300
 CTGAAATGGAGGAAGCCCTTAGAGGACTTCCAATAAGATACCAGACCCCAGCCATCAGAGCTGTGCACACCGGGCGGGAGATTGTGGACCTAATGTGTCA
 E M E E A L R G L P I R Y Q T P A I R A V H T G R E I V D L M C H I

D2 PDK-53 NS3-250-Val locus giảm độc lực (D2 16681: Glu, nt-5270-A)

5310 5320 5330 5340 5350 5360 5370 5380 5390 5400
 TGCCACATTTACCATGAGGCTGCTATCACCAGTTAGAGTGCCAAACTACAACTGATTATCATGGACGAAGCCCATTTACAGACCCAGCAAGTATAGCA
 A T F T M R L L S P V R V P N Y N L I I M D E A H F T D P A S I A
 5410 5420 5430 5440 5450 5460 5470 5480 5490 5500
 GCTAGAGGATACATCTCAACTCGAGTGGAGATGGGTGAGGCAGCTGGGATTTTTATGACAGCCACTCCCCGGGAAGCAGAGACCCATTTCTCTAGAGCA
 A R G Y I S T R V E M G E A A G I F M T A T P P G S R D P F P Q S N
 5510 5520 5530 5540 5550 5560 5570 5580 5590 5600
 ATGCACCAATCATAGATGAAGAAAGAGAAATCCCTGAACGCTCGTGGGAATTCGGGACATGAATGGGTCACGGATTTTAAAGGGAAGACTGTTTGGTTTCGT
 A P I I D E E R E I P E R S W N S G H E W V T D F K G K T V W F V I

Đột biến câm D2 PDK-53 nt-5541-C (D2 16681: T)

5610 5620 5630 5640 5650 5660 5670 5680 5690 5700
 TCCAAGTATAAAAGCAGGAAATGATATAGCAGCTTGCTGAGGAAAAATGGAAAGAAAGTGATACAACCTCAGTAGGAAGACCTTTGATTCTGAGTATGTC
 P S I K A G N D I A A C L R K N G K K V I Q L S R K T F D S E Y V
 5710 5720 5730 5740 5750 5760 5770 5780 5790 5800
 AAGACTAGAACCAATGATTGGGACTTCGTGGTTACAACCTGACATTTTCAGAAATGGGTGCCAATTTCAAGGCTGAGAGGGTTATAGACCCCAGACGCTGCA
 K T R T N D W D F V V T T D I S E M G A N F K A E R V I D P R R C M
 5810 5820 5830 5840 5850 5860 5870 5880 5890 5900

TGAAACCAGTCATACTAACAGATGGTGAAGAGCGGGTGATTCTGGCAGGACCTATGCCAGTGACCCACTCTAGTGCAGCACAAAGAAGAGGGAGAATAGG
 K P V I L T D G E E R V I L A G P M P V T H S S A A Q R R G R I G
 5910 5920 5930 5940 5950 5960 5970 5980 5990 6000
 AAGAAATCCAAAAATGAGAATGACCAGTACATATACATGGGGGAACCTCTGGAAAATGATGAAGACTGTGCACACTGGAAAGAAGCTAAAATGCTCCTA
 R N P K N E N D Q Y I Y M G E P L E N D E D C A H W K E A K M L L
 6010 6020 6030 6040 6050 6060 6070 6080 6090 6100
 GATAACATCAACACGCCAGAAGGAATCATTCCTAGCATGTTCTGAACAGAGCGTGAAAAGGTGGATGCCATTGATGGCGAATACCGCTTGAGAGGAGAAG
 D N I N T P E G I I P S M F E P E R E K V D A I D G E Y R L R G E A
 6110 6120 6130 6140 6150 6160 6170 6180 6190 6200
 CAAGGAAAACCTTTGTAGACTTAATGAGAAGAGGAGACCTACCAGTCTGGTTGGCCTACAGAGTGGCAGCTGAAGGCATCAACTACGCAGACAGAAGGTG
 R K T F V D L M R R G D L P V W L A Y R V A A E G I N Y A D R R W
 6210 6220 6230 6240 6250 6260 6270 6280 6290 6300
 GTGTTTTGATGGAGTCAAGAACAACCAATCCTAGAAGAAAACCTGGAAGTTGAAATCTGGACAAAAGAAGGGGAAAGGAAATGAAACCCAGATGG
 C F D G V K N N Q I L E E N V E V E I W T K E G E R K K L K P R W

> NS4A

6310 6320 6330 6340 6350 6360 6370 6380 6390 6400
 TTGGATGCTAGGATCTATTCTGACCCACTGGCGCTAAAAGAATTTAAGGAATTTGCAGCCGGAAGAAAGTCTCTGACCTGAACCTAATCACAGAAATGG
 L D A R I Y S D P L A L K E F K E F A A G R K S L T L N L I T E M G
 6410 6420 6430 6440 6450 6460 6470 6480 6490 6500
 GTAGGCTCCCAACCTTCATGACTCAGAAGCAAGAGACGCACTGGACAACCTTAGCAGTGTGCACACGGCTGAGGCAGGTGGAAGGGCGTACAACCATGC
 R L P T F M T Q K A R D A L D N L A V L H T A E A G G R A Y N H A
 6510 6520 6530 6540 6550 6560 6570 6580 6590 6600
 TCTCAGTGAAGTGCCTGGAGACCTGGAGACATTGCTTTTACTGACACTTCTGGCTACAGTCACGGGAGGATCTTTTTATTCTTGATGAGCGCAAGGGGC
 L S E L P E T L E T L L L L T L L A T V T G G I F L F L M S A R G L

NS4A-75-Ala đặc hiệu D2 PDK-53 (wt D2 16681: Gly, nt-6599-G)

6610 6620 6630 6640 6650 6660 6670 6680 6690 6700
 ATAGGGAAGATGACCCCTGGGAATGTGCTGCATAATCACGGCTAGCATCCTCCTATGGTACGCACAAATACAGCCACACTGGATAGCAGCTTCAATAATAC
 I G K M T L G M C C I I T A S I L L W Y A Q I Q P H W I A A S I I L
 6710 6720 6730 6740 6750 6760 6770 6780 6790 6800
 TGGAGTTTTTTCTCATAGTTTTTGCTTATTCAGAACCTGAAAAACAGAGAACACCCCAAGACAACCAACTGACCTACGTTGTCATAGCCATCCTCACAGT
 E F F L I V L L I P E P E K Q R T P Q D N Q L T Y V V I A I L T V

> NS4B

6810 6820 6830 6840 6850 6860 6870 6880 6890 6900
 GGTGGCCGCAACCATGGCAAACGAGATGGGTTTCCTAGAAAAAACGAAGAAAGATCTCGGATTGGGAAGCATTGCAACCCAGCAACCCGAGAGCAACATC
 V A A T M A N E M G F L E K T K K D L G L G S I A T Q Q P E S N I
 6910 6920 6930 6940 6950 6960 6970 6980 6990 7000
 CTGGACATAGATCTACGTCCTGCATCAGCATGGACGCTGTATGCCGTGGCCACAACATTTGTTACCAATGTTGAGACATAGCATTGAAAATTCCTCAG
 L D I D L R P A S A W T L Y A V A T T F V T P M L R H S I E N S S V

```

7010      7020      7030      7040      7050      7060      7070      7080      7090      7100
TGAATGTGTCCCTAACAGCTATAGCCAACCAAGCCACAGTGTTAATGGGTCTCGGGAAAGGATGGCCATTGTCAAAGATGGACATCGGAGTTCCCTTCT
N V S L T A I A N Q A T V L M G L G K G W P L S K M D I G V P L L
7110      7120      7130      7140      7150      7160      7170      7180      7190      7200
CGCCATTGGATGCTACTCACAAAGTCAACCCATAACTCTCACAGCAGCTCTTTTCTTATTGGTAGCACATTATGCCATCATAGGGCCAGGACTCCAAGCA
A I G C Y S Q V N P I T L T A A L F L L V A H Y A I I G P G L Q A
7210      7220      7230      7240      7250      7260      7270      7280      7290      7300
AAAGCAACCAGAGAAGCTCAGAAAAGAGCAGCGCGGGCATCATGAAAAACCAACTGTCGATGGAATAACAGTGATTGACCTAGATCCAATACCTTATG
K A T R E A Q K R A A A G I M K N P T V D G I T V I D L D P I P Y D
7310      7320      7330      7340      7350      7360      7370      7380      7390      7400
ATCCAAAGTTTGAAAAGCAGTTGGGACAAGTAATGCTCCTAGTCCTCTGCGTGA CTCAAGTATTGATGATGAGGACTACATGGGCTCTGTGTGAGGCTTT
P K F E K Q L G Q V M L L V L C V T Q V L M M R T T W A L C E A L
7410      7420      7430      7440      7450      7460      7470      7480      7490      7500
AACCTTAGCTACCGGGCCCATCTCCACATTGTGGGAAGGAAATCCAGGGAGGTTTGGAACTACTACCATTCGGGTGTCAATGGCTAACATTTTGTAGAGGG
T L A T G P I S T L W E G N P G R F W N T T I A V S M A N I F R G
> NS5

```

```

7510      7520      7530      7540      7550      7560      7570      7580      7590      7600
AGTTACTTGGCCGAGCTGGACTTCTCTTTTCTATTATGAAGAACACAACCAACACAAGAAGGGGAAGTGGCAACATAGGAGAGACGCTTGGAGAGAAAT
S Y L A G A G L L F S I M K N T T N T R R G T G N I G E T L G E K W
7610      7620      7630      7640      7650      7660      7670      7680      7690      7700
GGAAAAGCCGATTGAACGCGTTGGGAAAAAGTGAATTCCAGATCTACAAGAAAAGTGAATCCAGGAAGTGGATAGAACCTTAGCAAAAGAAGGCATTAA
K S R L N A L G K S E F Q I Y K K S G I Q E V D R T L A K E G I K L

```

Đột biến cảm nt-7260 A-thành-G bổ sung trong hạt chủ và tiên chủ

```

7710      7720      7730      7740      7750      7760      7770      7780      7790      7800
AAGAGGAGAAACGGACCATCACGCTGTGTCGCGAGGCTCAGCAAACTGAGATGGTTCGTTGAGAGAAACATGGTCACACCAGAAGGGAAAGTAGTGGAC
R G E T D H H A V S R G S A K L R W F V E R N M V T P E G K V V D
7810      7820      7830      7840      7850      7860      7870      7880      7890      7900
CTCGGTTGTGGCAGAGGAGGCTGGTCATACTATTGTGGAGGACTAAAGAATGTAAGAGAAGTCAAAGGCCTAACAAAAGGAGGACCAGGACACGAAGAAC
L G C G R G G W S Y Y C G G L K N V R E V K G L T K G G P G H E E P
7910      7920      7930      7940      7950      7960      7970      7980      7990      8000
CCATCCCCATGTCAACATATGGGTGGAATCTAGTGCGTCTTCAAAGTGGAGTTGACGTTTTCTTCATCCCGCCAGAAAAGTGTGACACATTATTGTGTGA
I P M S T Y G W N L V R L Q S G V D V F F I P P E K C D T L L C D
8010      8020      8030      8040      8050      8060      8070      8080      8090      8100
CATAGGGGAGTCATACCAAATCCCACAGTGGAAGCAGGACGAACACTCAGAGTCCTTAACTTAGTAGAAAATTGGTTGAACAACAACACTCAATTTTGC
I G E S S P N P T V E A G R T L R V L N L V E N W L N N N T Q F C
8110      8120      8130      8140      8150      8160      8170      8180      8190      8200
ATAAAGGTTCTCAACCCATATATGCCCTCAGTCATAGAAAAAATGGAAGCACTACAAAGGAAATATGGAGGAGCCTTAGTGAGGAATCCACTCTCACGAA
I K V L N P Y M P S V I E K M E A L Q R K Y G G A L V R N P L S R N
8210      8220      8230      8240      8250      8260      8270      8280      8290      8300

```

ACTCCACACATGAGATGTACTGGGTATCCAATGCTTCCGGGAACATAGTGTTCATCAGTGAACATGATTTCAAGGATGTTGATCAACAGATTTACAATGAG
 S T H E M Y W V S N A S G N I V S S V N M I S R M L I N R F T M R
 8310 8320 8330 8340 8350 8360 8370 8380 8390 8400
 ATACAAGAAAGCCACTTACGAGCCGGATGTTGACCTCGGAAGCGGAACCCGTAACATCGGGATTGAAAGTGAGATACCAAACCTAGATATAATTGGGAAA
 Y K K A T Y E P D V D L G S G T R N I G I E S E I P N L D I I G K
 8410 8420 8430 8440 8450 8460 8470 8480 8490 8500
 AGAATAGAAAAATAAAGCAAGAGCATGAAACATCATGGCACTATGACCAAGACCACCCATACAAAACGTGGGCATACCATGGTAGCTATGAAACAAAAC
 R I E K I K Q E H E T S W H Y D Q D H P Y K T W A Y H G S Y E T K Q
 8510 8520 8530 8540 8550 8560 8570 8580 8590 8600
 AGACTGGATCAGCATCATCCATGGTCAACGGAGTGGTCAGGCTGCTGACAAAACCTTGGGACGTCGTCCCATGGTGACACAGATGGCAATGACAGACAC
 T G S A S S M V N G V V R L L T K P W D V V P M V T Q M A M T D T
 8610 8620 8630 8640 8650 8660 8670 8680 8690 8700
 GACTCCATTTGGACAACAGCGCGTTTTTAAAGAGAAAGTGGACACGAGAACCCAAAGAACCGAAAGAAGGCACGAAGAACTAATGAAAATAACAGCAGAG
 T P F G Q Q R V F K E K V D T R T Q E P K E G T K K L M K I T A E
 8710 8720 8730 8740 8750 8760 8770 8780 8790 8800
 TGGCTTTGGAAAGAATTAGGGAAGAAAAAGACACCCAGGATGTGCACCAGAGAAGAATTCACAAGAAAGGTGAGAAGCAATGCAGCCTTGGGGGCCATAT
 W L W K E L G K K K T P R M C T R E E F T R K V R S N A A L G A I F
 8810 8820 8830 8840 8850 8860 8870 8880 8890 8900
 TCACTGATGAGAACAAGTGAAGTCGGCACGTGAGGCTGTTGAAGATAGTAGGTTTTGGGAGCTGGTTGACAAGGAAAGGAATCTCCATCTTGAAGGAAA
 T D E N K W K S A R E A V E D S R F W E L V D K E R N L H L E G K
 8910 8920 8930 8940 8950 8960 8970 8980
 8990 9000
 GTGTGAAACATGTGTGTACAACATGATGGGAAAAAGAGAGAAGAAGCTAGGGGAATTCGGCAAGGCAAAAAGGCAGCAGAGCCATATGGTACATGTGGCTT
 C E T C V Y N M M G K R E K K L G E F G K A K G S R A I W Y M W L
 9010 9020 9030 9040 9050 9060 9070 9080 9090 9100
 GGAGCACGCTTCTTAGAGTTTGAAGCCCTAGGATTCTTAAATGAAGATCACTGGTTCTCCAGAGAGAACTCCCTGAGTGGAGTGAAGGAGAAGGGCTGC
 G A R F L E F E A L G F L N E D H W F S R E N S L S G V E G E G L H
 9110 9120 9130 9140 9150 9160 9170 9180 9190 9200
 ACAAGCTAGGTTACATTCTAAGAGACGTGAGCAAGAAAGAGGGAGGAGCAATGTATGCCGATGACACCGCAGGATGGGATACAAGAATCACACTAGAAGA
 K L G Y I L R D V S K K E G G A M Y A D D T A G W D T R I T L E D
 9210 9220 9230 9240 9250 9260 9270 9280 9290 9300
 CCTAAAAATGAAGAAATGGTAACAAACCACATGGAAGGAGAACAAGAACTAGCCGAGGCCATTTTCAAATAACGTACCAAAAACAAGGTGGTGCCT
 L K N E E M V T N H M E G E H K K L A E A I F K L T Y Q N K V V R
 9310 9320 9330 9340 9350 9360 9370 9380 9390 9400
 GTGCAAAGACCAACACCAAGAGGCACAGTAATGGACATCATATCGAGAAGAGACCAAAGAGGTAGTGGACAAGTTGGCACCTATGGACTCAATACTTTCA
 V Q R P T P R G T V M D I I S R R D Q R G S G Q V G T Y G L N T F T
 9410 9420 9430 9440 9450 9460 9470 9480 9490 9500
 CCAATATGGAAGCCCAACTAATCAGACAGATGGAGGGAGAAGGAGTCTTTAAAGCATTTCAGCACCTAACAATCACAGAAGAAATCGCTGTGCAAACTG

N M E A Q L I R Q M E G E G V F K S I Q H L T I T E E I A V Q N W
 9510 9520 9530 9540 9550 9560 9570 9580 9590 9600
 GTTAGCAAGAGTGGGGCGCGAAAGGTTATCAAGAATGGCCATCAGTGGAGATGATTGTGTTGTGAAACCTTTAGATGACAGGTTGCGAAGCGCTTTAACA
 L A R V G R E R L S R M A I S G D D C V V K P L D D R F A S A L T
 9610 9620 9630 9640 9650 9660 9670 9680 9690 9700
 GCTCTAAATGACATGGGAAAGATTAGGAAAGACATACAACAATGGGAACCTTCAAGAGGATGGAATGATTGGACACAAGTGCCCTTCTGTTTACACCATT
 A L N D M G K I R K D I Q Q W E P S R G W N D W T Q V P F C S H H F
 9710 9720 9730 9740 9750 9760 9770 9780 9790 9800
 TCCATGAGTTAATCATGAAAGACGGTCGCGTACTCGTTGTTCCATGTAGAAACCAAGATGAACTGATTGGCAGAGCCCGAATCTCCCAAGGAGCAGGGTG
 H E L I M K D G R V L V V P C R N Q D E L I G R A R I S Q G A G W
 9810 9820 9830 9840 9850 9860 9870 9880 9890 9900
 GTCTTTGCGGGAGACGGCTGTTTGGGGAAGTCTTACGCCCCAAATGTGGAGCTTGATGTACTTCCACAGACGCGACCTCAGGCTGGCGGCAAATGCTATT
 S L R E T A C L G K S Y A Q M W S L M Y F H R R D L R L A A N A I
 9910 9920 9930 9940 9950 9960 9970 9980 9990 10000
 TGCTCGGCAGTACCATCACATTGGGTTCCAACAAGTGAACAACCTGGTCCATACATGCTAAACATGAATGGATGACAACGGAAGACATGCTGACAGTCT
 C S A V P S H W V P T S R T T W S I H A K H E W M T T E D M L T V W
 10010 10020 10030 10040 10050 10060 10070 10080 10090 10100
 GGAACAGGGTGTGGATTCAAGAAAACCCATGGATGGAAGACAAAACCTCCAGTGGAAATCATGGGAGGAAATCCCATACTTGGGGAAAAGAGAAGACCAATG
 N R V W I Q E N P W M E D K T P V E S W E E I P Y L G K R E D Q W
 10110 10120 10130 10140 10150 10160 10170 10180 10190 10200
 GTGCGGCTCATTGATTGGGTTAACAAGCAGGGCCACCTGGGCAAAGAACATCCAAGCAGCAATAAATCAAGTTAGATCCCTTATAGGCAATGAAGAATAC
 C G S L I G L T S R A T W A K N I Q A A I N Q V R S L I G N E E Y
> 3'-vùng không mã hoá
 10210 10220 10230 10240 10250 10260 10270 10280 10290 10300
 ACAGATTACATGCCATCCATGAAAAGATTGAGAAGAGAAGAGGAAGAAGCAGGAGTTCTGTGGTAGAAAGCAAACTAACATGAAACAAGGCTAGAAGTC
 T D Y M P S M K R F R R E E E E A G V L W *
 10310 10320 10330 10340 10350 10360 10370 10380 10390 10400
 AGGTCGGATTAAGCCATAGTACGGAATAAACTATGCTACCTGTGAGCCCCGTCCAAGGACGTTAAAAGAAGTCAGGCCATCATAAATGCCATAGCTTGAG
 10410 10420 10430 10440 10450 10460 10470 10480 10490 10500
 TAAACTATGCAGCCTGTAGCTCCACCTGAGAAGGTGTAAAAATCCGGGAGGCCACAAACCATGGAAGCTGTACGCATGGCGTAGTGACTAGCGGTTAG
 10510 10520 10530 10540 10550 10560 10570 10580 10590 10600
 AGGAGACCCCTCCCTTACAAATCGCAGCAACAATGGGGGCCCAAGGCGAGATGAAGCTGTAGTCTCGCTGGAAGGACTAGAGGTTAGAGGAGACCCCCC
 10610 10620 10630 10640 10650 10660 10670 10680 10690 10700
 GAAACAAAAACAGCATATTGACGCTGGGAAAGACCAGAGATCCTGCTGTCTCCTCAGCATCATTCAGGCACAGAACGCCAGAAAATGGAATGGTGCTG
 10710
 TTGAATCAACAGGTTCT

Hạt virut chủ DENvax-4 (MVS)

Trình tự nucleotit của hệ gen virus khảm và gọi ra trình tự axit amin của protein được dịch mã. Hầu hết gen prM-E (nt-457 đến -2379, được gạch dưới) là DEN-4 1036 kiểu hoang dại đặc hiệu virus; trình tự nucleotit còn lại là DEN-2 PDK-53 đặc hiệu virus. Tất cả các thay thế được thao tác khác với (DEN-3 16562 hoặc DEN-2 16681) virus kiểu hoang dại, cũng như các đột biến bổ sung (các thay đổi từ dòng ADN bổ trợ được thao tác) có bán trên thị trường.

Các thay thế được bao gồm trong hệ gen và protein:

Các vị trí ghép nối:

- a. MluI (nt 451-456): đột biến câm được thao tác, nt-453 A-thành-G
- b. NgoMIV (nt 2380-2385): đột biến được thao tác, nt-2381/2382 TG-thành-CC (dẫn đến thay đổi E-482 Val-thành-Ala)

Khung chính virus D2 PDK-53 (thay đổi từ D2 16681 kiểu hoang dại)

- a. vùng không mã hoá 5' (NCR)-57 (nt-57 C-thành-T): locus giảm độc lực chính (màu đỏ)
 - b. NS1-53 Gly-thành-Asp (nt-2579 G-thành-A): locus giảm độc lực chính (màu đỏ)
 - c. NS2A-181 Leu-thành-Phe (nt-4018 C-thành-T, nét đậm)
 - d. NS3-250 Glu-thành-Val (nt-5270 A-thành-T): locus giảm độc lực chính (màu đỏ)
 - e. đột biến câm nt-5547 (gen NS3) T-thành-C (nét đậm)
 - f. NS4A-75 Gly-thành-Ala (nt-6599 G-thành-C, nét đậm)
- * đột biến câm nt-8571 C-thành-T của PDK-53 được thao tác trong virus vaccin

Các thay thế được thao tác trong dòng ADN bổ trợ

- a. C-100 Arg-thành-Ser (nt-396 A-thành-C) được thao tác: có thể cải thiện sự sao chép virus trong môi trường nuôi cấy
- b. đột biến câm nt-1401 A-thành -G được thao tác
- c. E-364 Ala-thành-Val (nt-2027 C-thành-T) được thao tác: có thể cải thiện sự sao chép virus trong môi trường nuôi cấy
- d. E-447 Met-thành-Leu (nt-2275 A-thành-C) được thao tác: có thể cải thiện sự sao chép virus trong môi trường nuôi cấy

Các thay thế bổ sung được tìm thấy trong hạt vaccin (0,06% nt khác với dòng ban đầu)

- đột biến câm nt-225 (gen C) A-thành-T (nét đậm)
- đột biến NS2A-66 Asp-thành-Gly (nt-3674 A-thành-G) (nét đậm)
- hỗn hợp NS2A-99 Lys-thành-Lys/Arg (hỗn hợp nt-3773 A-thành-A/G, nét đậm)
- đột biến câm nt-5391 C-thành-T (gen NS3) (nét đậm)
- NS4A-21 Ala-thành-Val (nt-6437 C-thành-T, nét đậm)
- đột biến câm hỗn hợp nt-7026 T-thành-C/T (nét đậm)
- đột biến câm nt-9750 A-thành-C (nét đậm)

NCR-57-T, D2 PDK-53 locus giảm độc lực (wt D2 16681: C)

> 5'-vùng không mã hoá|>C

```

      10      20      30      40      50      60      70      80      90     100
AGTTGTTAGTCTACGTGGACCGACAAAGACAGATTCTTTGAGGGAGCTAAGCTCAATGTAGTTCTAACAGTTTTTTAATTAGAGAGCAGATCTCTGATGA
                                                                M  N
      110     120     130     140     150     160     170     180     190     200
ATAACCAACGGAAAAAGGCGAAAAACACGCCTTTCAATATGCTGAACGCGAGAGAAACCGCGTGTGCGACTGTGCAACAGCTGACAAAGAGATTCTCACT
N  Q  R  K  K  A  K  N  T  P  F  N  M  L  K  R  E  R  N  R  V  S  T  V  Q  Q  L  T  K  R  F  S  L
      210     220     230     240     250     260     270     280     290     300
TGAATGCTGCAGGGACGAGGACCTTTAAACTGTTTCATGGCCCTGGTGGCGTTCCTTCGTTTCCTAACAAATCCCAACAGCAGGGATATTGAAGAGA
G  M  L  Q  G  R  G  P  L  K  L  F  M  A  L  V  A  F  L  R  F  L  T  I  P  P  T  A  G  I  L  K  R  L

```

Đột biến câm nt-225 A-thành-T bổ sung trong hạt chủ và tiền chủ

```

      310     320     330     340     350     360     370     380     390     400
TGGGGAACAATTAAAAAATCAAAGCTATTAATGTTTGGAGGGGTTTCAGGAAAGAGATTGGAAGGATGCTGAACATCTTGAATAGGAGACGCAGCTCTG
W  G  T  I  K  K  S  K  A  I  N  V  L  R  G  F  R  K  E  I  G  R  M  L  N  I  L  N  R  R  R  S  S  A  L

```

C-100 Arg-thành-Ser (nt 396 A-thành-C) được thao thác

> prM Bắt đầu trình tự D4 1036

```

      410     420     430     440     450     460     470     480     490     500
CAGGCATGATCATTATGCTGATTCCAACAGTGATGGCGTTCCATTTAACCACGCGTGATGGCGAACCCTCATGATAGTGGCAAAACATGAAAGGGGGAG
G  M  I  I  M  L  I  P  T  V  M  A  F  H  L  T  T  R  D  G  E  P  L  M  I  V  A  K  H  E  R  G  R  L

```

Vị trí ghép nối MluI được thao tác (câm nt-453 A-thành-G)

```

      510     520     530     540     550     560     570     580     590     600
ACCTCTCTTGTTTAAGACAACAGAGGGGATCAACAAATGCACTCTCATTGCCATGGACTTGGGTGAAATGTGTGAGGACACTGTCTACGTATAAATGCCCC
P  L  L  F  K  T  T  E  G  I  N  K  C  T  L  I  A  M  D  L  G  E  M  C  E  D  T  V  T  Y  K  C  P
      610     620     630     640     650     660     670     680     690     700
TTACTGGTCAATACCGAACCTGAAGACATTGATTGCTGGTGCATCTCACGTCTACCTGGGTGATGTATGGGACATGCCCCAGAGCGGAGAACGGAGAC
L  L  V  N  T  E  P  E  D  I  D  C  W  C  N  L  T  S  T  W  V  M  Y  G  T  C  T  Q  S  G  E  R  R  R
> M

```

```

      710     720     730     740     750     760     770     780     790     800

```

GAGAGAAGCGCTCAGTAGCTTTAACACCACATTCAGGAATGGGATTGGAACAAGAGCTGAGACATGGATGTCATCGGAAGGGGCTTGAAGCATGCTCA
 E K R S V A L T P H S G M G L E T R A E T W M S S E G A W K H A Q
 810 820 830 840 850 860 870 880 890 900
GAGAGTAGAGAGCTGGATACTCAGAAACCCAGGATTCGCGCTCTTGGCAGGATTTATGGCTTATATGATTGGGCAAAACAGGAATCCAGCGAACTGTCTTC
 R V E S W I L R N P G F A L L A G F M A Y M I G Q T G I Q R T V F
 > E
 910 920 930 940 950 960 970 980 990 1000
TTTGTCTAATGATGCTGGTCGCCCCATCCTACGGAATGCGATGCGTAGGAGTAGGAAACAGAGACTTTGTGGAAGGAGTCTCAGGTGGAGCATGGGTCTG
 F V L M M L V A P S Y G M R C V G V G N R D F V E G V S G G A W V D
 1010 1020 1030 1040 1050 1060 1070 1080 1090 1100
ATCTGGTGCTAGAACATGGAGGATCGGTCACAACCATGGCCAGGGAAAACCAACCTTGGATTTTGAAGTACTAAGACAACAGCCAAGGAAGTGGCTCT
 L V L E H G G C V T T M A Q G K P T L D F E L T K T T A K E V A L
 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
GTTAAGAACCTATTGCATTGAAGCCTCAATATCAACATAACCACGGCAACAAGATGTCCAACGCAAGGAGAGCCTTATCTAAAAGAGGAACAAGACCAA
 L R T Y C I E A S I S N I T T A T R C P T Q G E P Y L K E E Q D Q
 1210 1220 1230 1240 1250 1260 1270 1280 1290 1300
CAGTACATTTGCCGGAGAGATGTGGTAGACAGAGGGTGGGGCAATGGCTGTGGCTTGTTTGAAAAGGAGGAGTTGTGACATGTGCCAAGTTTTCATGTT
 Q Y I C R R D V V D R G W G N G C G L F G K G G V V T C A K F S C S
 1310 1320 1330 1340 1350 1360 1370 1380 1390 1400
CGGGGAAGATAACAGGCAATTTGGTCCAAATTGAGAACCTTGAATACACAGTGGTTGTAACAGTCCACAATGGAGACACCCATGCAGTAGGAAATGACAC
 G K I T G N L V Q I E N L E Y T V V V T V H N G D T H A V G N D T
 1410 1420 1430 1440 1450 1460 1470 1480 1490 1500
GTCCAATCATGGAGTTACAGCCACGATAAATCCCAGGTACCATCGGTGGAAGTCAAATTGCCGGACTATGGAGAACTAACACTCGATTGTGAACCCAGG
 S N H G V T A T I T P R S P S V E V K L P D Y G E L T L D C E P R L
Đột biến nt-1401 A-thành-G nằm trong dòng được thao tác
 1510 1520 1530 1540 1550 1560 1570 1580 1590 1600
TCTGGAATTGACTTTAATGAGATGATTCTGATGAAAATGAAAAGAAAACATGGCTTGTGCATAAGCAATGGTTTTTGGATCTACCTCTACCATGGACAG
 S G I D F N E M I L M K M K K K T W L V H K Q W F L D L P L P W T A
 1610 1620 1630 1640 1650 1660 1670 1680 1690 1700
CAGGAGCAGACACATCAGAGGTTCACTGGAATTACAAGAGAGAATGGTGACATTTAAGGTTCTCATGCCAAGAGACAGGATGTGACAGTCTGGGATC
 G A D T S E V H W N Y K E R M V T F K V P H A K R Q D V T V L G S
 1710 1720 1730 1740 1750 1760 1770 1780 1790 1800
TCAGGAAGGAGCCATGCATTCTGCCCTCGCTGGAGCCACAGAAGTGGACTCCGGTGATGGAAATCACATGTTTGCAGGACATCTCAAGTGCAAAGTCCGT
 Q E G A M H S A L A G A T E V D S G D G N H M F A G H L K C K V R
 1810 1820 1830 1840 1850 1860 1870 1880 1890 1900
ATGGAGAAATTGAGAATCAAGGGAATGTCATACACGATGTGTTTCAGGAAAGTTCTCAATTGACAAAGAGATGGCAGAAACACAGCATGGGACAACAGTGG
 M E K L R I K G M S Y T M C S G K F S I D K E M A E T Q H G T T V V
 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000
TGAAAGTCAAGTATGAAGGTGCTGGAGCTCCGTGTAAAGTCCCCATAGAGATAAGAGATGTGAACAAGGAAAAAGTGGTTGGGCGTATCATCTCATCCAC

K V K Y E G A G A P C K V P I E I R D V N K E K V V G R I I S S T
2010 2020 2030 2040 2050 2060 2070 2080 2090 2100

CCCTTTGGCTGAGAATACCAACAGTGTAAACCAACATAGAGTTAGAACCCCTTTGGGGACAGCTACATAGTGATAGGTGTTGGAACAGTGCATTAACA

P L A E N T N S V T N I E L E P P F G D S Y I V I G V G N S A L T I

E-364 Ala-thành-Val (nt-2027 C-thành-T) được thao tác để cải thiện sự sinh trưởng virus trong nuôi trường nuôi cấy

2110 2120 2130 2140 2150 2160 2170 2180 2190 2200

CTCCATTGGTTCAGGAAAGGGAGTTCCATTGGCAAGATGTTTGGAGTCCACATACAGAGGTGCAAAACGAATGGCCATTCTAGGTGAAACAGCTTGGGATT

L H W F R K G S S I G K M F E S T Y R G A K R M A I L G E T A W D F

2210 2220 2230 2240 2250 2260 2270 2280 2290 2300

TTGGTTCCGTTGGTGGACTGTTACATCATTGGGAAAGGCTGTGCACCAGGTTTTTGGAAAGTGTGTATACAACCCCTGTTTGGAGGAGTCTCATGGATGAT

G S V G G L F T S L G K A V H Q V F G S V Y T T L F G G V S W M I

Đột biến E-447 Met-thành-Leu (nt-2275 A-thành-C) được thao tác

Kết thúc trình tự D4 1036

2310 2320 2330 2340 2350 2360 2370 | 2390 2400

TAGAACTCCTAATTGGGTTCTAGTGTGTGGATTGGCAGCAACTCAAGGAACACTTCAATGGCTATGACGTGCATAGCTGCCGGCATTGTGACACTGTAT

R I L I G F L V L W I G T N S R N T S M A M T C I A A G I V T L Y

Vị trí ghép nối NgoMIV được thao tác, E-482 Val-thành-Ala (nt-2381/2382 TG-thành-CC)

> NS1

2410 2420 2430 2440 2450 2460 2470 2480 2490 2500

TTGGGAGTCATGGTGCAGGCCGATAGTGGTTGCGTTGTGAGCTGGAAAAACAAAGAACTGAAATGTGGCAGTGGGATTTTCATCACAGACAACGTGCACA

L G V M V Q A D S G C V V S W K N K E L K C G S G I F I T D N V H T

2510 2520 2530 2540 2550 2560 2570 2580 2590 2600

CATGGACAGAACAATACAAGTTCCAACCAGAATCCCTTCAAACTAGCTTCAGCTATCCAGAAAGCCCATGAAGAGGAACATTTGTGGAATCCGCTCAGT

W T E Q Y K F Q P E S P S K L A S A I Q K A H E E D I C G I R S V

D2 PDK-53 NS1-53-Asp locus giảm độc lực (wt D2 16681: Gly, nt-2579-G)

2610 2620 2630 2640 2650 2660 2670 2680 2690 2700

AACAAGACTGGAGAATCTGATGTGGAACAAATAACACCAGAATTGAATCACATTCTATCAGAAAATGAGGTGAAGTTAACTATTATGACAGGAGACATC

T R L E N L M W K Q I T P E L N H I L S E N E V K L T I M T G D I

2710 2720 2730 2740 2750 2760 2770 2780 2790 2800

AAAGGAATCATGCAGGCAGGAAAACGATCTCTGCGGCCTCAGCCCACTGAGCTGAAGTATTCATGGAAAACATGGGGCAAAGCAAAAATGCTCTCTACAG

K G I M Q A G K R S L R P Q P T E L K Y S W K T W G K A K M L S T E

2810 2820 2830 2840 2850 2860 2870 2880 2890 2900

AGTCTCATAACCAGACCTTTCTCATTGATGGCCCCGAAACAGCAGAATGCCCCAACACAAATAGAGCTTGGAAATTCGTTGGAAGTTGAAGACTATGGCTT

S H N Q T F L I D G P E T A E C P N T N R A W N S L E V E D Y G F

2910 2920 2930 2940 2950 2960 2970 2980 2990 3000

TGGAGTATTACCACCAATATATGGCTAAAATTGAAAGAAAAACAGGATGTATTCTGCGACTCAAACTCATGTCAGCGGCCATAAAGACAACAGAGCC

G V F T T N I W L K L K E K Q D V F C D S K L M S A A I K D N R A

3010 3020 3030 3040 3050 3060 3070 3080 3090 3100

GTCCATGCCGATATGGGTTATTGGATAGAAAGTGCACATCAATGACACATGGAAGATAGAGAAAGCCTCTTTCATTGAAGTTAAAACTGCCACTGGCCAA

V H A D M G Y W I E S A L N D T W K I E K A S F I E V K N C H W P K
 3110 3120 3130 3140 3150 3160 3170 3180 3190 3200
 AATCACACACCTCTGGAGCAATGGAGTGCTAGAAAGTGAGATGATAATTCCAAAGAATCTCGCTGGACCAGTGTCTCAACACAACATATAGACCAGGCTA
 S H T L W S N G V L E S E M I I P K N L A G P V S Q H N Y R P G Y
 3210 3220 3230 3240 3250 3260 3270 3280 3290 3300
 CCATACACAAATAACAGGACCATGGCATCTAGGTAAGCTTGAGATGGACTTTGATTCTGTGATGGAACAACAGTGGTAGTGACTGAGGACTGCGGAAAT
 H T Q I T G P W H L G K L E M D F D F C D G T T V V V T E D C G N
 3310 3320 3330 3340 3350 3360 3370 3380 3390 3400
 AGAGGACCCTCTTTGAGAACAACCACTGCCTCTGGAAAACTCATAACAGAATGGTGCTGCCGATCTTGCACATTACCACCGCTAAGATACAGAGGTGAGG
 R G P S L R T T T A S G K L I T E W C C R S C T L P P L R Y R G E D
> NS2A
 3410 3420 3430 3440 3450 3460 3470 3480 3490 3500
 ATGGGTGCTGGTACGGGATGGAAATCAGACCATTGAAGGAGAAAGAAGAGAATTTGGTCAACTCCTTGGTCACAGCTGGACATGGGCAGGTCGACAACCTT
 G C W Y G M E I R P L K E K E E N L V N S L V T A G H G Q V D N F
 3510 3520 3530 3540 3550 3560 3570 3580 3590 3600
 TTCCTAGGAGTCTTTGGGAATGGCATTGTTTCCTGGAGGAAATGCTTAGGACCCGAGTAGGAACGAAACATGCAATACTACTAGTTGCAGTTTCTTTTGTG
 S L G V L G M A L F L E E M L R T R V G T K H A I L L V A V S F V
 3610 3620 3630 3640 3650 3660 3670 3680 3690 3700
 ACATTGATCACAGGGAACATGTCCTTTAGAGACCTGGGAAGAGTGATGGTTATGGTAGGCGCCACTATGACGGGTGACATAGGTATGGGCGTGACTTATC
 T L I T G N M S F R D L G R V M V M V G A T M T **G** D I G M G V T Y L I
NS2A-66 Asp-thành-Gly (đột biến nt-3674 A-thành-G) bổ sung trong hạt chủ và tiền chủ
 3710 3720 3730 3740 3750 3760 3770 3780 3790 3800
 TTGCCCTACTAGCAGCCTTCAAAGTCAGACCAACTTTTGCAGCTGGACTACTCTTGAGAAAGCTGACCTCCAGGGAATTGATGATGACTACTATAGGAAT
 A L L A A F K V R P T F A A G L L L R K L T S **K** E L M M T T I G I I
Đột biến NS2A-99 K thành R/K(hỗn hợp) (nt-3773 A-thành-G/A) bổ sung trong hạt chủ
 3810 3820 3830 3840 3850 3860 3870 3880 3890 3900
 TGTACTCCTCTCCCAGAGCACCATACCAGAGACCATTCTTGAGTTGACTGATGCGTTAGCCTTAGGCATGATGGTCCTCAAAATGGTGAGAAATATGGAA
 V L L S Q S T I P E T I L E L T D A L A L G M M V L K M V R N M E
 3910 3920 3930 3940 3950 3960 3970 3980 3990 4000
 AAGTATCAATTGGCAGTGACTATCATGGCTATCTTGTGCGTCCCAAACGCAGTGATATTACAAAACGCATGGAAAGTGAGTTGCACAATATTGGCAGTGG
 K Y Q L A V T I M A I L C V P N A V I L Q N A W K V S C T I L A V V
 4010 4020 4030 4040 4050 4060 4070 4080 4090 4100
 TGTCCGTTTCCCACTGTTCTTAACATCCTCACAGCAAAAAACAGATTGGATACCATTAGCATTGACGATCAAAGGTCTCAATCCAACAGCTATTTTCTCT
 S V S P L **F** L T S S Q Q K T D W I P L A L T I K G L N P T A I F L I
NS2A-181-Phe đặc hiệu D2 PDK-53 (wt D2 16681: Leu, nt-4018-C)
> NS2B
 4110 4120 4130 4140 4150 4160 4170 4180 4190 4200
 AACAAACCTCTCAAGAACCAGCAAGAAAAGGAGCTGGCCATTAAATGAGGCTATCATGGCAGTCGGGATGGTGAGCATTTTAGCCAGTTCTCTCTCTAAAA
 T T L S R T S K K R S W P L N E A I M A V G M V S I L A S S L L K

4210 4220 4230 4240 4250 4260 4270 4280 4290 4300
 AATGATATTTCCCATGACAGGACCATTAGTGGCTGGAGGGCTCCTCACTGTGTGCTACGTGCTCACTGGACGATCGGCCGATTTGGAACCTGGAGAGAGCAG
 N D I P M T G P L V A G G L L T V C Y V L T G R S A D L E L E R A A
 4310 4320 4330 4340 4350 4360 4370 4380 4390 4400
 CCGATGTCAAATGGGAAGACCAGGCAGAGATATCAGGAAGCAGTCCAATCCTGTCAATAACAATATCAGAAGATGGTAGCATGTCGATAAAAAATGAAGA
 D V K W E D Q A E I S G S S P I L S I T I S E D G S M S I K N E E
 4410 4420 4430 4440 4450 4460 4470 4480 4490 4500
 GGAAGAACAACACTGACCATACTCATTAGAACAGGATTGCTGGTGATCTCAGGACTTTTTCTGTATCAATACCAATCACGGCAGCAGCATGGTACCTG
 E E Q T L T I L I R T G L L V I S G L F P V S I P I T A A A W Y L

> NS3

4510 4520 4530 4540 4550 4560 4570 4580 4590 4600
 TGGGAAGTGAAGAAACAACGGGCCGGAGTATTGTGGGATGTTCTTCACCCCCACCCATGGGAAAGGCTGAACCTGGAAGATGGAGCCTATAGAATTAAGC
 W E V K K Q R A G V L W D V P S P P P M G K A E L E D G A Y R I K Q
 4610 4620 4630 4640 4650 4660 4670 4680 4690 4700
 AAAAAGGGATTCTTGGATATTTCCAGATCGGAGCCGGAGTTTACAAAGAAGGAACATTCCATACAATGTGGCATGTCACACGTGGCGCTGTTCTAATGCA
 K G I L G Y S Q I G A G V Y K E G T F H T M W H V T R G A V L M H
 4710 4720 4730 4740 4750 4760 4770 4780 4790 4800
 TAAAGGAAAGAGGATTGAACCATCATGGGCGGACGTCAAGAAAGACCTAATATCATATGGAGGAGGCTGGAAGTTAGAAGGAGAATGGAAGGAAGGAGAA
 K G K R I E P S W A D V K K D L I S Y G G G W K L E G E W K E G E
 4810 4820 4830 4840 4850 4860 4870 4880 4890 4900
 GAAGTCCAGGTATTGGCACTGGAGCCTGGAAAAAATCCAAGAGCCGTCCAACGAAACCTGGTCTTTTCAAAACCAACGCCGGAACAATAGGTGCTGTAT
 E V Q V L A L E P G K N P R A V Q T K P G L F K T N A G T I G A V S
 4910 4920 4930 4940 4950 4960 4970 4980 4990 5000
 CTCTGGACTTTTCTCTGGAACGTCAAGGATCTCCAATTATCGACAAAAAAGGAAAAGTTGTGGGTCTTTATGGTAATGGTGTGTGTACAAGGAGTGGAGC
 L D F S P G T S G S P I I D K K G K V V G L Y G N G V V T R S G A
 5010 5020 5030 5040 5050 5060 5070 5080 5090 5100
 ATATGTGAGTGCTATAGCCCAGACTGAAAAAGCATTGAAGACAACCCAGAGATCGAAGATGACATTTTCCGAAAGAGAAGACTGACCATCATGGACCTC
 Y V S A I A Q T E K S I E D N P E I E D D I F R K R R L T I M D L
 5110 5120 5130 5140 5150 5160 5170 5180 5190 5200
 CACCCAGGAGCGGGAAAGACGAAGAGATACCTTCCGGCCATAGTCAGAGAAGCTATAAAACGGGGTTTGAGAACATTAATCTTGGCCCCACTAGAGTTG
 H P G A G K T K R Y L P A I V R E A I K R G L R T L I L A P T R V V
 5210 5220 5230 5240 5250 5260 5270 5280 5290 5300
 TGGCAGCTGAAATGGAGGAAGCCCTTAGAGGACTTCCAATAAGATACCAGACCCAGCCATCAGAGCTGTGCACACCGGGCGGGAGATTGTGGACCTAAT
 A A E M E E A L R G L P I R Y Q T P A I R A V H T G R E I V D L M |
 D2 PDK-53 NS3-250-Val locus giảm độc lực (D2 16681: Glu, nt-5270-A)
 5310 5320 5330 5340 5350 5360 5370 5380 5390 5400
 GTGTCATGCCACATTTACCATGAGGCTGCTATCACCAGTTAGAGTGCCAAACTACAACCTGATTATCATGGACGAAGCCCATTTACAGATCCAGCAAGT
 C H A T F T M R L L S P V R V P N Y N L I I M D E A H F T D P A S |

Đột biến câm nt-5391 C-thành-T bổ sung ở hạt chủ và tiền chủ

5410 5420 5430 5440 5450 5460 5470 5480 5490 5500
 ATAGCAGCTAGAGGATACATCTCAACTCGAGTGGAGATGGGTGAGGCAGCTGGGATTTTTATGACAGCCACTCCCCGGGAAGCAGAGACCCATTTCCTC
 I A A R G Y I S T R V E M G E A A G I F M T A T P P G S R D P F P Q
 5510 5520 5530 5540 5550 5560 5570 5580 5590 5600
 AGAGCAATGCACCAATCATAGATGAAGAAAGAGAAATCCCTGAACGCTCGTGGAATCCGGACATGAATGGGTACCGGATTTTAAAGGGAAGACTGTTTG
 S N A P I I D E E R E I P E R S W N S G H E W V T D F K G K T V W I

Đột biến câm đặc hiệu D2 PDK-53 nt-5547-C (D2 16681: T)

5610 5620 5630 5640 5650 5660 5670 5680 5690 5700
 GTTCGTTCCAAGTATAAAAGCAGGAAATGATATAGCAGCTTGCCCTGAGGAAAAATGGAAAGAAAGTGATACAACCTCAGTAGGAAGACCTTTGATTCTGAG
 F V P S I K A G N D I A A C L R K N G K K V I Q L S R K T F D S E
 5710 5720 5730 5740 5750 5760 5770 5780 5790 5800
 TATGTCAAGACTAGAACCAATGATTGGGACTTCGTGGTTACAACCTGACATTTTCAGAAATGGGTGCCAATTTCAAGGCTGAGAGGGTTATAGACCCGAGAC
 Y V K T R T N D W D F V V T T D I S E M G A N F K A E R V I D P R R
 5810 5820 5830 5840 5850 5860 5870 5880 5890 5900
 GCTGCATGAAACCAGTCATACTAACAGATGGTGAAGAGCGGGTGATTCTGGCAGGACCTATGCCAGTGACCCACTCTAGTGCAGCACAAAGAAGAGGGAG
 C M K P V I L T D G E E R V I L A G P M P V T H S S A A Q R R G R
 5910 5920 5930 5940 5950 5960 5970 5980 5990 6000
 AATAGGAAGAAATCCAAAAATGAGAATGACCAGTACATATACATGGGGGAACCTCTGGAAAATGATGAAGACTGTGCACACTGGAAAGAAGCTAAAATG
 I G R N P K N E N D Q Y I Y M G E P L E N D E D C A H W K E A K M
 6010 6020 6030 6040 6050 6060 6070 6080 6090 6100
 CTCCTAGATAACATCAACACGCCAGAAGGAATCATTCCTAGCATGTTTGAACACAGAGCGTGAAAAGGTGGATGCCATTGATGGCGAATACCGCTTGAGAG
 L L D N I N T P E G I I P S M F E P E R E K V D A I D G E Y R L R G
 6110 6120 6130 6140 6150 6160 6170 6180 6190 6200
 GAGAAGCAAGGAAAACCTTTGTAGACTTAATGAGAAGAGGAGACCTACCAGTCTGGTTGGCCTACAGAGTGGCAGCTGAAGGCATCAACTACGCAGACAG
 E A R K T F V D L M R R G D L P V W L A Y R V A A E G I N Y A D R
 6210 6220 6230 6240 6250 6260 6270 6280 6290 6300
 AAGGTGGTGTGTTTGTATGGAGTCAAGAACAACCAAATCCTAGAAGAAAACGTGGAAGTTGAAATCTGGACAAAAGAAGGGGAAAGGAAGAAATTGAAACCC
 R W C F D G V K N N Q I L E E N V E V E I W T K E G E R K K L K P

> NS4A

6310 6320 6330 6340 6350 6360 6370 6380 6390 6400
 AGATGGTTGGATGCTAGGATCTATTCTGACCCACTGGCGCTAAAAGAATTTAAGGAATTTGCAGCCGGAAGAAAGTCTCTGACCCTGAACCTAATCACAG
 R W L D A R I Y S D P L A L K E F K E F A A G R K S L T L N L I T E
 6410 6420 6430 6440 6450 6460 6470 6480 6490 6500
 AAATGGGTAGGCTCCCAACCTTCATGACTCAGAAGGTAAAGAGACGCACTGGACAACCTTAGCAGTGTGCACACGGCTGAGGCAGGTGGAAGGGCGTACAA
 M G R L P T F M T Q K V R D A L D N L A V L H T A E A G G R A Y N I

Đột biến NS4A-21 Ala-thành-Val (nt-6437 C-thành-T) bổ sung ở hạt chủ và tiền chủ

6510 6520 6530 6540 6550 6560 6570 6580 6590 6600
 CCATGCTCTCAGTGAAGTCCCGGAGACCTGGAGACATTGCTTTTACTGACACTTCTGGCTACAGTCACGGGAGGGATCTTTTTATTCTTGATGAGCGCA

H A L S E L P E T L E T L L L L T L L A T V T G G I F L F L M S A |

NS4A-75-Ala đặc hiệu D2 PDK-53 (wt D2 16681: Gly, nt-6599-G)

6610	6620	6630	6640	6650	6660	6670	6680	6690	6700																								
AGGGGCATAGGGAAGATGACCCTGGGAATGTGCTGCATAATCACGGCTAGCATCCTCCTATGGTACGCACAAATACAGCCCACTGGATAGCAGCTTCAA																																	
R	G	I	G	K	M	T	L	G	M	C	C	I	I	T	A	S	I	L	L	W	Y	A	Q	I	Q	P	H	W	I	A	A	S	I
6710	6720	6730	6740	6750	6760	6770	6780	6790	6800																								
TAATACTGGAGTTTTTTCTCATAGTTTTGCTTATTCCAGAACCTGAAAAACAGAGAACACCCCAAGACAACCAACTGACCTACGTTGTCATAGCCATCCT																																	
I	L	E	F	F	L	I	V	L	L	I	P	E	P	E	K	Q	R	T	P	Q	D	N	Q	L	T	Y	V	V	I	A	I	L	

> NS4B

6810	6820	6830	6840	6850	6860	6870	6880	6890	6900																								
CACAGTGGTGGCCGCAACCATGGCAAACGAGATGGGTTTCTAGAAAAACGAAGAAAGATCTCGGATTGGGAAGCATTGCAACCCAGCAACCCGAGAGC																																	
T	V	V	A	A	T	M	A	N	E	M	G	F	L	E	K	T	K	K	D	L	G	L	G	S	I	A	T	Q	Q	P	E	S	
6910	6920	6930	6940	6950	6960	6970	6980	6990	7000																								
AACATCCTGGACATAGATCTACGTCCTGCATCAGCATGGACGCTGTATGCCGTGGCCACAACATTTGTTACACCAATGTTGAGACATAGCATTGAAAATT																																	
N	I	L	D	I	D	L	R	P	A	S	A	W	T	L	Y	A	V	A	T	T	F	V	T	P	M	L	R	H	S	I	E	N	S
7010	7020	7030	7040	7050	7060	7070	7080	7090	7100																								
CCTCAGTGAATGTGTCCCTAACAGCCATAGCCAACCAAGCCACAGTGTTAATGGGTCTCGGGAAGGATGGCCATTGTCAAAGATGGACATCGGAGTTCC																																	
S	V	N	V	S	L	T	A	I	A	N	Q	A	T	V	L	M	G	L	G	K	G	W	P	L	S	K	M	D	I	G	V	P	I

Đột biến câm hỗn hợp nt-7026 T→thành-C/T bổ sung ở hạt chủ và tiền chủ

7110	7120	7130	7140	7150	7160	7170	7180	7190	7200																								
CCTTCTCGCCATTGGATGCTACTCACAAGTCAACCCATAACTCTCAGCAGCTCTTTTCTTATTGGTAGCACATTATGCCATCATAGGGCCAGGACTC																																	
L	L	A	I	G	C	Y	S	Q	V	N	P	I	T	L	T	A	A	L	F	L	L	V	A	H	Y	A	I	I	G	P	G	L	
7210	7220	7230	7240	7250	7260	7270	7280	7290	7300																								
CAAGCAAAAGCAACCAGAGAAGCTCAGAAAAGAGCAGCGCGGGCATCATGAAAACCCAACTGTCGATGGAATAACAGTGATTGACCTAGATCCAATAC																																	
Q	A	K	A	T	R	E	A	Q	K	R	A	A	A	G	I	M	K	N	P	T	V	D	G	I	T	V	I	D	L	D	P	I	P
7310	7320	7330	7340	7350	7360	7370	7380	7390	7400																								
CTTATGATCCAAAGTTTGAAAAGCAGTTGGGACAAGTAATGCTCCTAGTCTCTGCGTGAAGTCAAGTATTGATGATGAGGACTACATGGGCTCTGTGTGA																																	
Y	D	P	K	F	E	K	Q	L	G	Q	V	M	L	L	V	L	C	V	T	Q	V	L	M	M	R	T	T	W	A	L	C	E	
7410	7420	7430	7440	7450	7460	7470	7480	7490	7500																								
GGCTTTAACCTTAGCTACCGGGCCCATCTCCACATTGTGGGAAGGAAATCCAGGAGGTTTTTGGAACTACTACCATTCGCGGTGTCAATGGCTAACATTTTT																																	
A	L	T	L	A	T	G	P	I	S	T	L	W	E	G	N	P	G	R	F	W	N	T	T	I	A	V	S	M	A	N	I	F	

> NS5

7510	7520	7530	7540	7550	7560	7570	7580	7590	7600																								
AGAGGGAGTTACTTGGCCGGAGCTGGACTTCTCTTTTCTATTATGAAGAACACAACCAACACAAGAAGGGGAAGTGGCAACATAGGAGAGACGCTTGGAG																																	
R	G	S	Y	L	A	G	A	G	L	L	F	S	I	M	K	N	T	T	N	T	R	R	G	T	G	N	I	G	E	T	L	G	E
7610	7620	7630	7640	7650	7660	7670	7680	7690	7700																								
AGAAATGGAAAAGCCGATTGAACGCATTGGGAAAAAGTGAATTCCAGATCTACAAGAAAAGTGGAAATCCAGGAAGTGGATAGAACCTTAGCAAAAAGAAGG																																	
K	W	K	S	R	L	N	A	L	G	K	S	E	F	Q	I	Y	K	K	S	G	I	Q	E	V	D	R	T	L	A	K	E	G	
7710	7720	7730	7740	7750	7760	7770	7780	7790	7800																								

CATTAAAGAGGAGAAACGGACCATCACGCTGTGTGCGAGGCTCAGCAAACTGAGATGGTTCGTTGAGAGAAACATGGTCACACCAGAAGGGAAAGTA
 I K R G E T D H H A V S R G S A K L R W F V E R N M V T P E G K V
 7810 7820 7830 7840 7850 7860 7870 7880 7890 7900
 GTGGACCTCGGTTGTGGCAGAGGAGGCTGGTCATACTATTGTGGAGGACTAAAGAATGTAAGAGAAGTCAAAGGCCTAACAAAAGGAGGACCAGGACACG
 V D L G C G R G G W S Y Y C G G L K N V R E V K G L T K G G P G H E
 7910 7920 7930 7940 7950 7960 7970 7980 7990 8000
 AAGAACCCATCCCCATGTCAACATATGGGTGGAATCTAGTGCCTCTTCAAAGTGGAGTTGACGTTTTCTTCATCCCGCCAGAAAAGTGTGACACATTATT
 E P I P M S T Y G W N L V R L Q S G V D V F F I P P E K C D T L L
 8010 8020 8030 8040 8050 8060 8070 8080 8090 8100
 GTGTGACATAGGGGAGTCATCACCAATCCCACAGTGAAGCAGGACGAACACTCAGAGTCCTTAACCTAGTAGAAAATTGGTTGAACAACAACACTCAA
 C D I G E S S P N P T V E A G R T L R V L N L V E N W L N N N T Q
 8110 8120 8130 8140 8150 8160 8170 8180 8190 8200
 TTTTGATAAAGGTTCTCAACCCATATATGCCCTCAGTCATAGAAAAATGGAAGCACTACAAAGGAAATATGGAGGAGCCTTAGTGAGGAATCCACTCT
 F C I K V L N P Y M P S V I E K M E A L Q R K Y G G A L V R N P L S
 8210 8220 8230 8240 8250 8260 8270 8280 8290 8300
 CACGAAACTCCACACATGAGATGTACTGGGTATCCAATGCTTCCGGGAACATAGTGTCACTGACATGATTTCAAGGATGTTGATCAACAGATTAC
 R N S T H E M Y W V S N A S G N I V S S V N M I S R M L I N R F T
 8310 8320 8330 8340 8350 8360 8370 8380 8390 8400
 AATGAGATACAAGAAAGCCACTTACGAGCCGGATGTTGACCTCGGAAGCGGAACCCGTAACATCGGGATTGAAAGTGAGATACCAAACCTAGATATAATT
 M R Y K K A T Y E P D V D L G S G T R N I G I E S E I P N L D I I
 8410 8420 8430 8440 8450 8460 8470 8480 8490 8500
 GGGAAAAGAATAGAAAAAATAAAGCAAGAGCATGAAACATCATGGCACTATGACCAAGACCACCCATACAAAACGTGGGCATACCATGGTAGCTATGAAA
 G K R I E K I K Q E H E T S W H Y D Q D H P Y K T W A Y H G S Y E T
 8510 8520 8530 8540 8550 8560 8570 8580 8590 8600
 CAAAACAGACTGGATCAGCATCATCCATGGTCAACGGAGTGGTCAGGCTGCTGACAAAACCTTGGGACGTCGTCCTCCCATGGTGACACAGATGGCAATGAC
 K Q T G S A S S M V N G V V R L L T K P W D V V P M V T Q M A M T
 8610 8620 8630 8640 8650 8660 8670 8680 8690 8700
 AGACACGACTCCATTTGGACAACAGCGCGTTTAAAGAGAAAAGTGGACACGAGAACCCAAGAACCAGAAAGGCACGAAGAACTAATGAAAATAACA
 D T T P F G Q Q R V F K E K V D T R T Q E P K E G T K K L M K I T
 8710 8720 8730 8740 8750 8760 8770 8780 8790 8800
 GCAGAGTGGCTTTGGAAAGAATTAGGGAAGAAAAAGACACCCAGGATGTGCACCAGAGAAGAATTCACAAGAAAGGTGAGAAGCAATGCAGCCTTGGGGG
 A E W L W K E L G K K K T P R M C T R E E F T R K V R S N A A L G A
 8810 8820 8830 8840 8850 8860 8870 8880 8890 8900
 CCATATTCACTGATGAGAACAAGTGAAGTCGGCACGTGAGGCTGTTGAAGATAGTAGGTTTTGGGAGCTGGTTGACAAGGAAAGGAATCTCCATCTTGA
 I F T D E N K W K S A R E A V E D S R F W E L V D K E R N L H L E
 8910 8920 8930 8940 8950 8960 8970 8980 8990 9000
 AGGAAAGTGTGAAACATGTGTGTACAACATGATGGGAAAAAGAGAGAAGAAGCTAGGGGAATTCGGCAAGGCAAAAGGCAGCAGAGCCATATGGTACATG
 G K C E T C V Y N M M G K R E K K L G E F G K A K G S R A I W Y M
 9010 9020 9030 9040 9050 9060 9070 9080 9090 9100

TGGCTTGAGCACGCTTCTTAGAGTTTGAAGCCCTAGGATTCTTAAATGAAGATCACTGGTTCTCCAGAGAGAACTCCCTGAGTGGAGTGGGAAGGAGAAG
 W L G A R F L E F E A L G F L N E D H W F S R E N S L S G V E G E G
 9110 9120 9130 9140 9150 9160 9170 9180 9190 9200
 GGCTGCACAAGCTAGGTTACATTCTAAGAGACGTGAGCAAGAAAGAGGGAGGAGCAATGTATGCCGATGACACCGCAGGATGGGATACAAGAATCACACT
 L H K L G Y I L R D V S K K E G G A M Y A D D T A G W D T R I T L
 9210 9220 9230 9240 9250 9260 9270 9280 9290 9300
 AGAAGACCTAAAAAATGAAGAAATGGTAACAAACCACATGGAAGGAGAACACAAGAACTAGCCGAGGCCATTTTCAAATAACGTACCAAAACAAGGTG
 E D L K N E E M V T N H M E G E H K K L A E A I F K L T Y Q N K V
 9310 9320 9330 9340 9350 9360 9370 9380 9390 9400
 GTGCGTGTGCAAAGACCAACACCAAGAGGCACAGTAATGGACATCATATCGAGAAGAGACCAAAGAGGTAGTGGACAAGTTGGCACCTATGGACTCAATA
 V R V Q R P T P R G T V M D I I S R R D Q R G S G Q V G T Y G L N T
 9410 9420 9430 9440 9450 9460 9470 9480 9490 9500
 CTTTCACCAATATGGAAGCCCAACTAATCAGACAGATGGAGGGAGAAGGAGTCTTTAAAAGCATTTCAGCACCTAACAAATCAGAGAAGAAATCGCTGTGCA
 F T N M E A Q L I R Q M E G E G V F K S I Q H L T I T E E I A V Q
 9510 9520 9530 9540 9550 9560 9570 9580 9590 9600
 AAAGTGGTTAGCAAGAGTGGGGCGCGAAAGGTTATCAAGAATGGCCATCAGTGGAGATGATTGTGTTGTGAAACCTTTAGATGACAGGTTTCCGAAGCGCT
 N W L A R V G R E R L S R M A I S G D D C V V K P L D D R F A S A
 9610 9620 9630 9640 9650 9660 9670 9680 9690 9700
 TTAACAGCTCTAAATGACATGGGAAAGATTAGGAAAGACATACAACAATGGGAACCTTCAAGAGGATGGAATGATTGGACACAAGTGCCCTTCTGTTTAC
 L T A L N D M G K I R K D I Q Q W E P S R G W N D W T Q V P F C S H
 9710 9720 9730 9740 9750 9760 9770 9780 9790 9800
 ACCATTTCCATGAGTTAATCATGAAAGACGGTCGCGTACTCGTTGTTCCCTGTAGAAACCAAGATGAACTGATTGGCAGAGCCCGAATCTCCCAAGGAGC
 H F H E L I M K D G R V L V V P C R N Q D E L I G R A R I S Q G A I
 Đột biến câm **nt-9750 A-thành-C bổ sung ở hạt chủ và tiền chủ**
 9810 9820 9830 9840 9850 9860 9870 9880 9890 9900
 AGGGTGGTCTTTTGGGGAGACGGCCTGTTTGGGGAAGTCTTACGCCCAAATGTGGAGCTTGATGTACTTCCACAGACGCGACCTCAGGCTGGCGGCAAAAT
 G W S L R E T A C L G K S Y A Q M W S L M Y F H R R D L R L A A N
 9910 9920 9930 9940 9950 9960 9970 9980 9990 10000
 GCTATTTGCTCGGCAGTACCATCACATTGGGTTCACAACAAGTCGAACAACCTGGTCCATACATGCTAAACATGAATGGATGACAACGGAAGACATGCTGA
 A I C S A V P S H W V P T S R T T W S I H A K H E W M T T E D M L T
 10010 10020 10030 10040 10050 10060 10070 10080 10090 10100
 CAGTCTGGAACAGGGTGTGGATTCAAGAAAACCCATGGATGGAAGACAAAACCTCCAGTGGAAATCATGGGAGGAAATCCCATACTTGGGGAAAAGAGAAGA
 V W N R V W I Q E N P W M E D K T P V E S W E E I P Y L G K R E D
 10110 10120 10130 10140 10150 10160 10170 10180 10190 10200
 CCAATGGTGC GGCTCATTGATTGGGTTAACAAGCAGGGCCACCTGGGCAAAGAACATCCAAGCAGCAATAAATCAAGTTAGATCCCTTATAGGCAATGAA
 Q W C G S L I G L T S R A T W A K N I Q A A I N Q V R S L I G N E
 > 3'-vùng không mã hoá
 10210 10220 10230 10240 10250 10260 10270 10280 10290 10300
 GAATACACAGATTACATGCCATCCATGAAAAGATTGAGAAGAGAAGAGGAAGAAGCAGGAGTTCTGTGGTAGAAAGCAAACTAACATGAAACAAGGCTA

E Y T D Y M P S M K R F R R E E E E A G V L W *

10310	10320	10330	10340	10350	10360	10370	10380	10390	10400
GAAGTCAGGTCGGATTAAGCCATAGTACGGAAAAAAGTATGCTACCTGTGAGCCCCGTCCAAGGACGTTAAAAGAAGTCAGGCCATCATAAATGCCATAG									
10410	10420	10430	10440	10450	10460	10470	10480	10490	10500
CTTGAGTAACTATGCAGCCTGTAGCTCCACCTGAGAAGGTGTAAAAAATCCGGGAGGCCACAAACCATGGAAGCTGTACGCATGGCGTAGTGGAAGTAGC									
10510	10520	10530	10540	10550	10560	10570	10580	10590	10600
GGTTAGAGGAGACCCCTCCCTTACAAATCGCAGCAACAATGGGGGCCCAAGGCGAGATGAAGCTGTAGTCTCGCTGGAAGGACTAGAGGTAGAGGAGAC									
10610	10620	10630	10640	10650	10660	10670	10680	10690	10700
CCCCCGAAACAAAAACAGCATATTGACGCTGGGAAAGACCAGAGATCCTGCTGTCTCCTCAGCATCATTCAGGCACAGAACGCCAGAAAATGGAATG									
10710	10720								
GTGCTGTTGAATCAACAGGTTCT									

YÊU CẦU BẢO HỘ

1. Phân tử polynucleotit mã hóa phân tử polypeptit chủng virus dengue-2 sống được cải biến giảm độc lực PDK-53-V, trong đó phân tử polynucleotit này chứa:
đột biến adenin thành guanin ở vị trí 592 mã hóa axit glutamic thay cho lysin trong phân tử polypeptit này ở vị trí axit amin 166 tương ứng với prM-52; và
đột biến adenin thành guanin ở vị trí 8803 mã hóa valin thay cho isoleuxin trong phân tử polypeptit này ở vị trí axit amin 2903 tương ứng với NS5-412.
2. Phân tử polynucleotit theo điểm 1 chứa polynucleotit được biểu thị bởi SEQ ID NO: 14.
3. Phân tử polynucleotit theo điểm 1 hoặc 2 mã hóa phân tử polypeptit chủng virus dengue-2 sống được cải biến giảm độc lực PDK-53-V, trong đó phân tử polypeptit này được biểu thị bởi SEQ ID NO: 6.
4. Phân tử polypeptit được mã hóa bởi phân tử polynucleotit theo điểm bất kỳ trong số các điểm từ 1 đến 3.
5. Virus dengue-2 sống được cải biến giảm độc lực chủng PDK-53-V, chứa phân tử polypeptit theo điểm 4, hoặc
được biểu thị bởi phân tử polynucleotit theo điểm bất kỳ trong số các điểm từ 1 đến 3.
6. Dược phẩm chứa phân tử polynucleotit theo điểm bất kỳ trong số các điểm từ 1 đến 3, phân tử polypeptit theo điểm 4, hoặc virus dengue-2 theo điểm 5, và tá dược được dùng.
7. Vật truyền mã hóa phân tử polynucleotit theo điểm bất kỳ trong số các điểm từ 1 đến 3.
8. Tế bào chứa phân tử polynucleotit theo điểm bất kỳ trong số các điểm từ 1 đến 3 hoặc vật truyền theo điểm 7.
9. Chế phẩm sinh miễn dịch chứa chủng virus dengue-2 sống được cải biến giảm độc lực PDK-53-V được biểu thị bởi phân tử polynucleotit theo điểm bất kỳ trong số

các điểm từ 1 đến 3 và chất mang được dụng.

10. Chế phẩm sinh miễn dịch theo điểm 9, còn chứa thể khảm dengue-1/dengue-2 được biểu thị bởi phân tử polynucleotit mã hóa thể khảm polypeptit dengue-1/dengue-2, phân tử polynucleotit chứa trình tự nucleotit thứ nhất mã hóa protein không cấu trúc từ chủng virus dengue-2 sống được cải biến giảm độc lực PDK-53-V, và trình tự nucleotit thứ hai mã hóa ít nhất một protein cấu trúc từ dengue-1, trong đó phân tử polynucleotit này chứa:

đột biến adenin thành xytosin ở vị trí 3823 mã hóa leuxin thay vì isoleuxin trong thể khảm polypeptit dengue-1/dengue-2 ở vị trí axit amin 1243 tương ứng với NS2A-116;

đột biến adenin thành thymin ở vị trí 4407 mã hóa axit aspartic thay vì axit glutamic trong thể khảm polypeptit dengue-1/dengue-2 ở vị trí axit amin 1437 tương ứng với NS2B-92; và

đột biến adenin thành guanin ở vị trí 7311.

11. Chế phẩm sinh miễn dịch theo điểm 10, trong đó phân tử polynucleotit mã hóa thể khảm polypeptit dengue-1/dengue-2 còn chứa:

đột biến xytosin thành thymin ở vị trí 7148 mã hóa isoleuxin thay vì threonin trong thể khảm polypeptit dengue-1/dengue-2 ở vị trí axit amin 2351 tương ứng với NS4B-108; và

đột biến guanine thành xytosin ở vị trí 2384 mã hóa alanin thay vì glyxin trong thể khảm polypeptit dengue-1/dengue-2 ở vị trí axit amin 763 tương ứng với E-483.

12. Chế phẩm sinh miễn dịch theo điểm 10, trong đó phân tử polynucleotit mã hóa thể khảm polypeptit dengue-1/dengue-2 bao gồm polynucleotit được biểu thị bởi SEQ ID NO: 13.

13. Chế phẩm sinh miễn dịch theo điểm 10 hoặc 12, trong đó phân tử polynucleotit mã hóa thể khảm polypeptit dengue-1/dengue-2 được biểu thị bởi SEQ ID NO: 3.

14. Chế phẩm sinh miễn dịch theo điểm 9, còn chứa thể khảm dengue-3/dengue-2 được biểu thị bởi phân tử polynucleotit mã hóa thể khảm polypeptit dengue-3/dengue-

2, phân tử polynucleotit này chứa trình tự nucleotit thứ nhất mã hóa protein không cấu trúc từ chủng virus dengue-2 sống được cải biến giảm độc lực PDK-53-V, và trình tự nucleotit thứ hai mã hóa ít nhất một protein cấu trúc từ dengue-3, trong đó phân tử polynucleotit này chứa:

đột biến adenin thành thymine ở vị trí 1603 mã hóa serine thay vì threonine trong thể khảm polypeptit dengue-3/dengue-2 ở vị trí axit amin 503 tương ứng với E-223; và

đột biến adenin thành guanine ở vị trí 7620.

15. Chế phẩm sinh miễn dịch theo điểm 14, trong đó phân tử polynucleotit mã hóa thể khảm polypeptit dengue-3/dengue-2 còn bao gồm đột biến guanine thành adenine ở vị trí 6436 mã hóa asparagine thay vì axit aspartic trong thể khảm polypeptit dengue-3/dengue-2 ở vị trí axit amin 2114 tương ứng với NS4A-23.

16. Chế phẩm sinh miễn dịch theo điểm 14, trong đó phân tử polynucleotit mã hóa thể khảm polypeptit dengue-3/dengue-2 bao gồm polynucleotit được biểu thị bởi SEQ ID NO: 15.

17. Chế phẩm sinh miễn dịch theo điểm 14 hoặc 16, trong đó phân tử polynucleotit mã hóa thể khảm polypeptit dengue-3/dengue-2 được biểu thị bởi SEQ ID NO: 9.

18. Chế phẩm sinh miễn dịch theo điểm 9, còn chứa thể khảm dengue-4/dengue-2 được biểu thị bởi phân tử polynucleotit mã hóa thể khảm polypeptit dengue-4/dengue-2, phân tử polynucleotit này chứa trình tự nucleotit thứ nhất mã hóa protein không cấu trúc từ chủng virus dengue-2 sống được cải biến giảm độc lực PDK-53-V, và trình tự nucleotit thứ hai mã hóa ít nhất một protein cấu trúc từ dengue-4, trong đó phân tử polynucleotit này chứa:

đột biến adenine thành thymine ở vị trí 225;

đột biến adenine thành guanine ở vị trí 3674 mã hóa glycine thay vì axit aspartic trong thể khảm polypeptit dengue-4/dengue-2 ở vị trí axit amin 1193 tương ứng với NS2A-66;

đột biến xytosine thành thymine ở vị trí 5391;

đột biến xytosine thành thymine ở vị trí 6437 mã hóa valine thay vì alanine trong thể

khảm polypeptit dengue-4/dengue-2 ở vị trí axit amin 2114 tương ứng với NS4A-21, đột biến adenin thành xytosin ở vị trí 9750, và

tùy ý đột biến thymin thành xytosin ở vị trí 7026, và tùy ý đột biến adenin thành guanin ở vị trí 3773 mã hóa arginin thay vì lysin trong thể khảm polypeptit dengue-4/dengue-2 ở vị trí axit amin 1226 tương ứng với NS2A-99.

19. Chế phẩm sinh miễn dịch theo điểm 18, trong đó phân tử polynucleotit mã hóa thể khảm polypeptit dengue-4/dengue-2 còn chứa đột biến xytosin thành thymin ở vị trí 7538 mã hóa phenylalanin thay vì serin trong thể khảm polypeptit dengue-4/dengue-2 ở vị trí axit amin 2481 tương ứng với NS4B-238.

20. Chế phẩm sinh miễn dịch theo điểm 18, trong đó phân tử polynucleotit mã hóa thể khảm polypeptit dengue-4/dengue-2 chứa polynucleotit được biểu thị bởi SEQ ID NO: 16.

21. Chế phẩm sinh miễn dịch theo điểm bất kỳ trong số các điểm 18 hoặc 20, trong đó phân tử polynucleotit đã nêu mã hóa thể khảm polypeptit dengue-4/dengue-2 được biểu thị bởi SEQ ID NO: 12.

22. Chế phẩm sinh miễn dịch theo điểm bất kỳ trong số các điểm từ 9 đến 21, trong đó chế phẩm này chứa cả bốn typ huyết thanh virus dengue, là chế phẩm tứ giá.

23. Chế phẩm sinh miễn dịch theo điểm bất kỳ trong số các điểm từ 9 đến 22, còn bao gồm chế phẩm sinh miễn dịch chống lại các Flavivirus khác như virus West Nile, virus gây viêm não Nhật Bản.

24. Chế phẩm chứa một hoặc nhiều virus sống giảm độc lực theo điểm 5 và chất mang được dụng.

25. Chế phẩm theo điểm 24, trong đó chế phẩm này là chế phẩm virus dengue tứ giá có khả năng gây ra đáp ứng miễn dịch ở đối tượng chống lại cả bốn typ huyết thanh virus dengue.

26. Vật truyền theo điểm 7, trong đó vật truyền này là vật truyền plasmit.

27. Vật truyền theo điểm 7, trong đó vật truyền này là dòng lây nhiễm ADN hỗ trợ.

Danh mục trình tự

- <110> Takeda Vaccines, Inc.
The Government of The United States of America as
represented by The Secretary of The Department of Health and
Human Services
- <120> PHÂN TỬ POLYNUCLEOTIT MÃ HÓA PHÂN TỬ POLYPEPTIT CHỦNG VIRUT DENGUE-2 SỐNG
ĐƯỢC CẢI BIẾN GIẢM ĐỘC LỰC PDK-53-V, PHÂN TỬ POLYPEPTIT ĐƯỢC MÃ HÓA BỞI
POLYNUCLEOTIT NÀY, THẺ KHẪM CHỨA MỘT HOẶC NHIỀU POLYPEPTIT NÀY, DƯỢC PHẨM,
VẬT TRUYỀN VÀ TẾ BÀO CHỨA POLYNUCLEOTIT NÀY
- <130> T08239WOVN/ES
- <150> US 61/800,204
<151> 2013-03-15
- <150> PCT/US2014/024603
<151> 2014-03-12
- <150> 1-2015-03806
<151> 2014-03-12
- <160> 16
- <170> PatentIn version 3.5
- <210> 1
<211> 10723
<212> ADN
<213> Trình tự nhân tạo
- <220>
<223> Virut dengue typ huyết thanh 1, BVS
- <400> 1
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60
gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120
aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180
ctgacaaaga gattctcact tggaatgctg cagggacgag gaccattaaa actgttcattg 240
gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300
tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gaggggttcag gaaagagatt 360
ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat cattatgctg 420
attccaacag tgatggcggt ccatTTaacc acgcgtgggg gagagccgca tatgatagtt 480
agcaagcagg aaagaggaaa gtcacttttg ttcaagacct ctgcaggtgt caacatgtgc 540
accctcattg cgatggattt gggagagttg tgtgaggaca cgatgaccta caaatgcccc 600
cggatcactg aggcggaacc agatgacggt gactgttggg gcaatgccac ggacacatgg 660

gtgacctatg gaacgtgctc tcaaactggc gaacaccgac gagacaaacg ttccgtcgca	720
ttggccccac acgtggggct tggcctagaa acaagagccg aaacgtggat gtcctctgaa	780
ggtgcttgga aacagataca aaaagtagag acttgggctc tgagacatcc aggattcacg	840
gtgatagccc tttttctagc acatgccata ggaacatcca tcaccagaa agggatcatt	900
ttcattttgc tgatgctggc aacaccatct atggccatgc gatgcgtggg aataggcaac	960
agagacttcg tggaaggact gtcaggagca acatgggtgg atgtgggtact ggagcatgga	1020
agttgcgtca ccaccatggc aaaaaacaaa ccaacactgg acattgaact cttgaagacg	1080
gaggtcacia accctgcagt tctgcgtaaa ttgtgcattg aagctaaaat atcaaacacc	1140
accaccgatt cgagatgtcc aacacaagga gaagccacac tggtggaaga acaagacgcg	1200
aactttgtgt gccgacgaac gttcgtggac agaggctggg gcaatggctg tgggctattc	1260
ggaaaaggta gtctaataac gtgtgccaaag tttaagtgtg tgacaaaact agaaggaaag	1320
atagttcaat atgaaaacct aaaatattca gtgatagtca ccgtccacac tggagatcag	1380
caccaggtgg gaaatgagac tacagaacat ggaacaactg caaccataac acctcaagct	1440
cctacgtcgg aaatacagct gaccgactac ggaaccctta cattagattg ttcacctagg	1500
acagggctag attttaacga gatgggtgtg ctgacaatga aagaaagatc atggcttgtc	1560
cacaaacaat ggttcctaga cttaccactg ctttgacct ctggggcttc aacatcccaa	1620
gagacttgga acagacaaga ttactggctc acatttaaga cagctcatgc aaagaagcag	1680
gaagtagtcg tactaggatc acaagaagga gcaatgcaca ctgcgctgac tggagcgaca	1740
gaaatccaaa cgtcaggaac gacaacaatt ttgcgaggac acctaaaatg cagactaaaa	1800
atggacaaac taactttaaa agggatgtca tatgtgatgt gcacaggctc attcaagtta	1860
gagaaagaag tggctgagac ccagcatgga actgttctgg tgcagggtta atatgaagga	1920
acagacgcac catgcaagat tcccttttctg acccaagatg agaaaggagc aaccagaaat	1980
gggagattaa taacagccaa ccccatagtc actgacaaag aaaaaccagt caatattgag	2040
gcagaaccac cctttggtga gagctacatc gtggtaggag cagggtgaaa agctttgaaa	2100
ctaagctggt tcaagaaagg aagcagcata gggaaaatgt ttgaagcaac tgcccagga	2160
gcacgaagga tggccattct gggagacacc gcatgggact tcggttctat aggaggagt	2220
ttcacgtcta tgggaaaact ggtacaccag gtttttgga ctgcatatgg agttttgttt	2280
agcggagttt cttggaccat gaaaatagga atagggatc tgctgacatg gctaggatta	2340

aattcaagga acacgtccct ttcgatgatg tgcacgcag cgcattgt gacactgtat	2400
ttgggagtca tgggtgcaggc cgatagtggg tgcgttgtga gctggaaaaa caaagaactg	2460
aaatgtggca gtgggatttt catcacagac aacgtgcaca catggacaga acaatacaag	2520
ttccaaccag aatccccctt aaaactagct tcagctatcc agaaagccca tgaagaggac	2580
atgtgtggaa tccgctcagt aacaagactg gagaatctga tgtggaaaca aataacacca	2640
gaattgaatc acattctatc agaaaatgag gtgaagttaa ctattatgac aggagacatc	2700
aaaggaatca tgcaggcagg aaaacgatct ctgcggcctc agcccactga gctgaagtat	2760
tcatggaaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt	2820
ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg	2880
gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa	2940
aaacaggatg tattctgcga ctcaaaactc atgtcagcgg ccataaaaga caacagagcc	3000
gtccatgccg atatgggtta ttggatagaa agtgcactca atgacacatg gaagatagag	3060
aaagcctctt tcattgaagt taaaaactgc cactggccaa aatcacacac cctctggagc	3120
aatggagtgc tagaaagtga gatgataatt ccaaagaatc tcgctggacc agtgtctcaa	3180
cacaactata gaccaggcta ccatacacia ataacaggac catggcatct aggtaagctt	3240
gagatggact ttgatttctg tgatggaaca acagtggtag tgactgagga ctgcggaaat	3300
agaggaccct ctttgagaac aaccactgcc tctggaaaac tcataacaga atggtgctgc	3360
cgatcttgca cattaccacc gctaagatac agaggtgagg atgggtgctg gtacgggatg	3420
gaaatcagac cattgaagga gaaagaagag aatttgggtca actccttggc cacagctgga	3480
catgggcagg tcgacaactt ttcactagga gtcttgggaa tggcattgtt cctggaggaa	3540
atgcttagga cccgagtagg aacgaaacat gcaatactac tagttgcagt ttcttttgtg	3600
acattgatca cagggaacat gtcctttaga gacctgggaa gaggatggg tatggtaggc	3660
gccactatga cggatgacat aggtatgggc gtgacttata ttgccctact agcagccttc	3720
aaagtcagac caacttttgc agctggacta ctcttgagaa agctgacctc caaggaattg	3780
atgatgacta ctataggaat tgtactcttc tcccagagca ccctaccaga gaccattctt	3840
gagttgactg atgcgttagc cttaggcatg atggctctca aaatgggtgag aaatatggaa	3900
aagtatcaat tggcagtgc tatcatggct atcttgtgcg tcccaaacgc agtgatatta	3960
caaaacgcat ggaaagtgc ttgcacaata ttggcagtgg tgtccgtttc cccactgttc	4020
ttaacatcct cacagcaaaa aacagattgg ataccattag cattgacgat caaaggcttc	4080

-99-

ggtgccatt tcaaggctga gaggggtata gacccagac gctgcatgaa accagtcata	5820
ctaacagatg gtgaagagcg ggtgattctg gcaggaccta tgccagtgc ccactctagt	5880
gcagcacaaa gaagagggag aataggaaga aatccaaaa atgagaatga ccagtacata	5940
tacatggggg aacctctgga aaatgatgaa gactgtgcac actggaaaga agctaaaatg	6000
ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt	6060
gaaaaggtgg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt	6120
gtagacttaa tgagaagagg agacctacca gtctgggttg cctacagagt ggcagctgaa	6180
ggcatcaact acgcagacag aaggtggtgt tttgatggag tcaagaacaa ccaaactcta	6240
gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc	6300
agatggttgg atgctaggat ctattctgac ccactggcgc taaaagaatt taaggaatth	6360
gcagccggaa gaaagtctct gacctgaac ctaatcacag aaatgggtag gctcccaacc	6420
ttcatgactc agaaggcaag agacgcactg gacaacttag cagtgtctga cacggctgag	6480
gcaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg	6540
cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca	6600
aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta	6660
tggtacgcac aaatacagcc aactggata gcagcttcaa taatactgga gtttttctc	6720
atagttttgc ttattccaga acctgaaaa cagagaacac ccaagacaa ccaactgacc	6780
tacgttgtca tagccatcct cacagtgggt gccgcaacca tggcaaacga gatgggtttc	6840
ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caaccagca acccgagagc	6900
aacatcctgg acatagatct acgtcctgca tcagcatgga cgctgtatgc cgtggccaca	6960
acatttggtta caccaatgtt gagacatagc attgaaaatt cctcagtga tgtgtcccta	7020
acagctatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca	7080
aagatggaca tcggagttcc cttctcgc attggatgct actcacaagt caaccccata	7140
actctcatag cagctctttt cttattggta gcacattatg ccatcatagg gccaggactc	7200
caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaacca	7260
actgtcgatg gaataacagt gattgacct gatccaatac cttatgatcc gaagtttgaa	7320
aagcagttgg gacaagtaat gtccttagtc ctctgcgtga ctcaagtatt gatgatgagg	7380
actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg	7440
gaaggaaatc caggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt	7500

agagggagtt acttggccgg agctggactt ctcttttcta ttatgaagaa cacaaccaac	7560
acaagaaggg gaactggcaa cataggagag acgcttggag agaaatggaa aagccgattg	7620
aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat	7680
agaaccttag caaaagaagg cattaataaaga ggagaaacgg accatcacgc tgtgtcgcga	7740
ggctcagcaa aactgagatg gttcgttgag agaaacatgg tcacaccaga agggaaagta	7800
gtggacctcg gttgtggcag aggaggctgg tcatactatt gtggaggact aaagaatgta	7860
agagaagtca aaggcctaac aaaaggagga ccaggacacg aagaacccat ccccatgtca	7920
acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcca	7980
gaaaagtgtg acacattatt gtgtgacata ggggagtcac caccaaacc caccagtggaa	8040
gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa	8100
ttttgcataa aggttctcaa cccatatatg ccctcagtca tagaaaaaat ggaagcacta	8160
caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag	8220
atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg	8280
atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac	8340
ctcggaaagc gaacccgtaa catcgggatt gaaagtgaga taccaaacct agatataatt	8400
gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac	8460
caccataca aaacgtgggc ataccatggg agctatgaaa caaaacagac tggatcagca	8520
tcatccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtcccatg	8580
gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgctg ttttaaagag	8640
aaagtggaca cgagaacca agaaccgaaa gaaggcacga agaaactaat gaaaataaca	8700
gcagagtggc ttttgaaaaga attagggag aaagagacac ccaggatgtg caccagagaa	8760
gaattcacaa gaaaggtgag aagcaatgca gccttggggg ccatattcac tgatgagaac	8820
aagtggaaagt cggcacgtga ggctgttgaa gatagtaggt tttgggagct ggttgacaag	8880
gaaaggaatc tccatcttga aggaaagtgt gaaacatgtg tgtacaacat gatgggaaaa	8940
agagagaaga agctagggga attcggcaag gcaaaaggca gcagagccat atggtacatg	9000
tggcttggag cacgcttctt agagtttgaa gccctaggat tcttaaatga agatcactgg	9060
ttctccagag agaactccct gagtggagtg gaaggagaag ggctgcacaa gctaggttac	9120
attctaagag acgtgagcaa gaaagaggga ggagcaatgt atgccgatga caccgcagga	9180

tgggatacaa gaatcacact agaagaccta aaaaatgaag aaatggtaac aaaccacatg	9240
gaaggagaac acaagaaact agccgaggcc attttcaaac taacgtacca aaacaagggtg	9300
gtgctgtgtc aaagaccaac accaagaggc acagtaatgg acatcatatc gagaagagac	9360
caaagaggta gtggacaagt tggcacctat ggactcaata ctttcaccaa tatggaagcc	9420
caactaatca gacagatgga gggagaagga gtcttttaaaa gcattcagca cctaacaatc	9480
acagaagaaa tcgctgtgca aaactgggta gcaagagtgg ggcgcgaaag gttatcaaga	9540
atggccatca gtggagatga ttgtgtttgtg aaaccttttag atgacagggtt cgcaagcgct	9600
ttaacagctc taaatgacat gggaaagatt aggaaagaca tacaacaatg ggaaccttca	9660
agaggatgga atgattggac acaagtgcc ttctgttcac accatttcca tgagttaatc	9720
atgaaagacg gtcgcgtact cgttggttcca tgtagaaacc aagatgaact gattggcaga	9780
gcccgaaatct cccaaggagc aggggtggtct ttgcgggaga cggcctgttt ggggaagtct	9840
tacgccc aaa tgtggagctt gatgtacttc cacagacgcg acctcaggct ggcggcaa	9900
gctattttgct cggcagtacc atcacattgg gttccaacaa gtcgaacaac ctgggtccata	9960
catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa caggggtgtg	10020
attcaagaaa acccatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca	10080
tacttgggga aaagagaaga ccaatgggtc ggctcattga ttgggttaac aagcagggcc	10140
acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa	10200
gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga	10260
gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc	10320
catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca	10380
ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg	10440
tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc	10500
ggttagagga gaccctccc ttacaaatcg cagcaacaat gggggcccaa ggcgagatga	10560
agctgtagtc tcgctggaag gactagaggt tagaggagac ccccccga aaacacag	10620
catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca	10680
gaacgccaga aaatggaatg gtgctgttga atcaacaggt tct	10723

<210> 2

<211> 3391

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Virut dengue typ huyết thanh 1, BVS

<400> 2

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

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
85 90 95

Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
100 105 110

Met Ala Phe His Leu Thr Thr Arg Gly Gly Glu Pro His Met Ile Val
115 120 125

Ser Lys Gln Glu Arg Gly Lys Ser Leu Leu Phe Lys Thr Ser Ala Gly
130 135 140

Val Asn Met Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Leu Cys Glu
145 150 155 160

Asp Thr Met Thr Tyr Lys Cys Pro Arg Ile Thr Glu Ala Glu Pro Asp
165 170 175

Asp Val Asp Cys Trp Cys Asn Ala Thr Asp Thr Trp Val Thr Tyr Gly
180 185 190

Thr Cys Ser Gln Thr Gly Glu His Arg Arg Asp Lys Arg Ser Val Ala
195 200 205

Leu Ala Pro His Val Gly Leu Gly Leu Glu Thr Arg Ala Glu Thr Trp
 210 215 220

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

Ala Leu Arg His Pro Gly Phe Thr Val Ile Ala Leu Phe Leu Ala His
 245 250 255

Ala Ile Gly Thr Ser Ile Thr Gln Lys Gly Ile Ile Phe Ile Leu Leu
 260 265 270

Met Leu Val Thr Pro Ser Met Ala Met Arg Cys Val Gly Ile Gly Asn
 275 280 285

Arg Asp Phe Val Glu Gly Leu Ser Gly Ala Thr Trp Val Asp Val Val
 290 295 300

Leu Glu His Gly Ser Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr
 305 310 315 320

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

Arg Lys Leu Cys Ile Glu Ala Lys Ile Ser Asn Thr Thr Thr Asp Ser
 340 345 350

Arg Cys Pro Thr Gln Gly Glu Ala Thr Leu Val Glu Glu Gln Asp Ala
 355 360 365

Asn Phe Val Cys Arg Arg Thr Phe Val Asp Arg Gly Trp Gly Asn Gly
 370 375 380

Cys Gly Leu Phe Gly Lys Gly Ser Leu Ile Thr Cys Ala Lys Phe Lys
 385 390 395 400

Cys Val Thr Lys Leu Glu Gly Lys Ile Val Gln Tyr Glu Asn Leu Lys
 405 410 415

Tyr Ser Val Ile Val Thr Val His Thr Gly Asp Gln His Gln Val Gly
 420 425 430

Asn Glu Thr Thr Glu His Gly Thr Thr Ala Thr Ile Thr Pro Gln Ala
435 440 445

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

Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Val Leu Leu Thr
465 470 475 480

Met Lys Glu Arg Ser Trp Leu Val His Lys Gln Trp Phe Leu Asp Leu
485 490 495

Pro Leu Pro Trp Thr Ser Gly Ala Ser Thr Ser Gln Glu Thr Trp Asn
500 505 510

Arg Gln Asp Leu Leu Val Thr Phe Lys Thr Ala His Ala Lys Lys Gln
515 520 525

Glu Val Val Val Leu Gly Ser Gln Glu Gly Ala Met His Thr Ala Leu
530 535 540

Thr Gly Ala Thr Glu Ile Gln Thr Ser Gly Thr Thr Thr Ile Phe Ala
545 550 555 560

Gly His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Thr Leu Lys Gly
565 570 575

Met Ser Tyr Val Met Cys Thr Gly Ser Phe Lys Leu Glu Lys Glu Val
580 585 590

Ala Glu Thr Gln His Gly Thr Val Leu Val Gln Val Lys Tyr Glu Gly
595 600 605

Thr Asp Ala Pro Cys Lys Ile Pro Phe Ser Thr Gln Asp Glu Lys Gly
610 615 620

Ala Thr Gln Asn Gly Arg Leu Ile Thr Ala Asn Pro Ile Val Thr Asp
625 630 635 640

Lys Glu Lys Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Glu Ser
645 650 655

Tyr Ile Val Val Gly Ala Gly Glu Lys Ala Leu Lys Leu Ser Trp Phe

660										665					670				
Lys	Lys	Gly	Ser	Ser	Ile	Gly	Lys	Met	Phe	Glu	Ala	Thr	Ala	Arg	Gly				
		675					680					685							
Ala	Arg	Arg	Met	Ala	Ile	Leu	Gly	Asp	Thr	Ala	Trp	Asp	Phe	Gly	Ser				
	690					695					700								
Ile	Gly	Gly	Val	Phe	Thr	Ser	Met	Gly	Lys	Leu	Val	His	Gln	Val	Phe				
705					710					715					720				
Gly	Thr	Ala	Tyr	Gly	Val	Leu	Phe	Ser	Gly	Val	Ser	Trp	Thr	Met	Lys				
				725					730					735					
Ile	Gly	Ile	Gly	Ile	Leu	Leu	Thr	Trp	Leu	Gly	Leu	Asn	Ser	Arg	Asn				
			740					745					750						
Thr	Ser	Leu	Ser	Met	Met	Cys	Ile	Ala	Ala	Ala	Ile	Val	Thr	Leu	Tyr				
		755					760					765							
Leu	Gly	Val	Met	Val	Gln	Ala	Asp	Ser	Gly	Cys	Val	Val	Ser	Trp	Lys				
	770					775					780								
Asn	Lys	Glu	Leu	Lys	Cys	Gly	Ser	Gly	Ile	Phe	Ile	Thr	Asp	Asn	Val				
785					790					795					800				
His	Thr	Trp	Thr	Glu	Gln	Tyr	Lys	Phe	Gln	Pro	Glu	Ser	Pro	Ser	Lys				
				805					810					815					
Leu	Ala	Ser	Ala	Ile	Gln	Lys	Ala	His	Glu	Glu	Asp	Ile	Cys	Gly	Ile				
			820					825					830						
Arg	Ser	Val	Thr	Arg	Leu	Glu	Asn	Leu	Met	Trp	Lys	Gln	Ile	Thr	Pro				
		835					840					845							
Glu	Leu	Asn	His	Ile	Leu	Ser	Glu	Asn	Glu	Val	Lys	Leu	Thr	Ile	Met				
	850					855					860								
Thr	Gly	Asp	Ile	Lys	Gly	Ile	Met	Gln	Ala	Gly	Lys	Arg	Ser	Leu	Arg				
865					870					875					880				
Pro	Gln	Pro	Thr	Glu	Leu	Lys	Tyr	Ser	Trp	Lys	Thr	Trp	Gly	Lys	Ala				
				885					890						895				

Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
 900 905 910

Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
 915 920 925

Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
 930 935 940

Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
 945 950 955 960

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
 965 970 975

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
 980 985 990

Ile Glu Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser
 995 1000 1005

Asn Gly Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala
 1010 1015 1020

Gly Pro Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln
 1025 1030 1035

Ile Thr Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp
 1040 1045 1050

Phe Cys Asp Gly Thr Thr Val Val Val Thr Glu Asp Cys Gly Asn
 1055 1060 1065

Arg Gly Pro Ser Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile
 1070 1075 1080

Thr Glu Trp Cys Cys Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr
 1085 1090 1095

Arg Gly Glu Asp Gly Cys Trp Tyr Gly Met Glu Ile Arg Pro Leu
 1100 1105 1110

Lys Glu	Lys Glu Glu Asn	Leu Val	Asn Ser	Leu Val	Thr Ala Gly
1115		1120		1125	
His Gly	Gln Val Asp Asn	Phe Ser	Leu Gly	Val Leu	Gly Met Ala
1130		1135		1140	
Leu Phe	Leu Glu Glu Met	Leu Arg	Thr Arg	Val Gly	Thr Lys His
1145		1150		1155	
Ala Ile	Leu Leu Val Ala	Val Ser	Phe Val	Thr Leu	Ile Thr Gly
1160		1165		1170	
Asn Met	Ser Phe Arg Asp	Leu Gly	Arg Val	Met Val	Met Val Gly
1175		1180		1185	
Ala Thr	Met Thr Asp Asp	Ile Gly	Met Gly	Val Thr	Tyr Leu Ala
1190		1195		1200	
Leu Leu	Ala Ala Phe Lys	Val Arg	Pro Thr	Phe Ala	Ala Gly Leu
1205		1210		1215	
Leu Leu	Arg Lys Leu Thr	Ser Lys	Glu Leu	Met Met	Thr Thr Ile
1220		1225		1230	
Gly Ile	Val Leu Leu Ser	Gln Ser	Thr Leu	Pro Glu	Thr Ile Leu
1235		1240		1245	
Glu Leu	Thr Asp Ala Leu	Ala Leu	Gly Met	Met Val	Leu Lys Met
1250		1255		1260	
Val Arg	Asn Met Glu Lys	Tyr Gln	Leu Ala	Val Thr	Ile Met Ala
1265		1270		1275	
Ile Leu	Cys Val Pro Asn	Ala Val	Ile Leu	Gln Asn	Ala Trp Lys
1280		1285		1290	
Val Ser	Cys Thr Ile Leu	Ala Val	Val Ser	Val Ser	Pro Leu Phe
1295		1300		1305	
Leu Thr	Ser Ser Gln Gln	Lys Thr	Asp Trp	Ile Pro	Leu Ala Leu
1310		1315		1320	

Thr Ile	Lys Gly Leu Asn	Pro	Thr Ala Ile Phe	Leu	Thr Thr Leu
1325		1330		1335	
Ser Arg	Thr Ser Lys Lys	Arg	Ser Trp Pro Leu	Asn	Glu Ala Ile
1340		1345		1350	
Met Ala	Val Gly Met Val	Ser	Ile Leu Ala Ser	Ser	Leu Leu Lys
1355		1360		1365	
Asn Asp	Ile Pro Met Thr	Gly	Pro Leu Val Ala	Gly	Gly Leu Leu
1370		1375		1380	
Thr Val	Cys Tyr Val Leu	Thr	Gly Arg Ser Ala	Asp	Leu Glu Leu
1385		1390		1395	
Glu Arg	Ala Ala Asp Val	Lys	Trp Glu Asp Gln	Ala	Glu Ile Ser
1400		1405		1410	
Gly Ser	Ser Pro Ile Leu	Ser	Ile Thr Ile Ser	Glu	Asp Gly Ser
1415		1420		1425	
Met Ser	Ile Lys Asn Glu	Glu	Glu Asp Gln Thr	Leu	Thr Ile Leu
1430		1435		1440	
Ile Arg	Thr Gly Leu Leu	Val	Ile Ser Gly Leu	Phe	Pro Val Ser
1445		1450		1455	
Ile Pro	Ile Thr Ala Ala	Ala	Trp Tyr Leu Trp	Glu	Val Lys Lys
1460		1465		1470	
Gln Arg	Ala Gly Val Leu	Trp	Asp Val Pro Ser	Pro	Pro Pro Met
1475		1480		1485	
Gly Lys	Ala Glu Leu Glu	Asp	Gly Ala Tyr Arg	Ile	Lys Gln Lys
1490		1495		1500	
Gly Ile	Leu Gly Tyr Ser	Gln	Ile Gly Ala Gly	Val	Tyr Lys Glu
1505		1510		1515	
Gly Thr	Phe His Thr Met	Trp	His Val Thr Arg	Gly	Ala Val Leu
1520		1525		1530	
Met His	Lys Gly Lys Arg	Ile	Glu Pro Ser Trp	Ala	Asp Val Lys

1535		1540		1545
Lys Asp 1550	Leu Ile Ser Tyr	Gly 1555	Gly Gly Trp Lys	Leu 1560
				Glu Gly Glu
Trp Lys 1565	Glu Gly Glu Glu	Val 1570	Gln Val Leu Ala	Leu 1575
				Glu Pro Gly
Lys Asn 1580	Pro Arg Ala Val	Gln 1585	Thr Lys Pro Gly	Leu 1590
				Phe Lys Thr
Asn Ala 1595	Gly Thr Ile Gly	Ala 1600	Val Ser Leu Asp	Phe 1605
				Ser Pro Gly
Thr Ser 1610	Gly Ser Pro Ile	Ile 1615	Asp Lys Lys Gly	Lys 1620
				Val Val Gly
Leu Tyr 1625	Gly Asn Gly Val	Val 1630	Thr Arg Ser Gly	Ala 1635
				Tyr Val Ser
Ala Ile 1640	Ala Gln Thr Glu	Lys 1645	Ser Ile Glu Asp	Asn 1650
				Pro Glu Ile
Glu Asp 1655	Asp Ile Phe Arg	Lys 1660	Arg Arg Leu Thr	Ile 1665
				Met Asp Leu
His Pro 1670	Gly Ala Gly Lys	Thr 1675	Lys Arg Tyr Leu	Pro 1680
				Ala Ile Val
Arg Glu 1685	Ala Ile Lys Arg	Gly 1690	Leu Arg Thr Leu	Ile 1695
				Leu Ala Pro
Thr Arg 1700	Val Val Ala Ala	Glu 1705	Met Glu Glu Ala	Leu 1710
				Arg Gly Leu
Pro Ile 1715	Arg Tyr Gln Thr	Pro 1720	Ala Ile Arg Ala	Val 1725
				His Thr Gly
Arg Glu 1730	Ile Val Asp Leu	Met 1735	Cys His Ala Thr	Phe 1740
				Thr Met Arg
Leu Leu 1745	Ser Pro Val Arg	Val 1750	Pro Asn Tyr Asn	Leu 1755
				Ile Ile Met

Asp	Glu	Ala	His	Phe	Thr	Asp	Pro	Ala	Ser	Ile	Ala	Ala	Arg	Gly
1760						1765					1770			
Tyr	Ile	Ser	Thr	Arg	Val	Glu	Met	Gly	Glu	Ala	Ala	Gly	Ile	Phe
1775						1780					1785			
Met	Thr	Ala	Thr	Pro	Pro	Gly	Ser	Arg	Asp	Pro	Phe	Pro	Gln	Ser
1790						1795					1800			
Asn	Ala	Pro	Ile	Ile	Asp	Glu	Glu	Arg	Glu	Ile	Pro	Glu	Arg	Ser
1805						1810					1815			
Trp	Asn	Ser	Gly	His	Glu	Trp	Val	Thr	Asp	Phe	Lys	Gly	Lys	Thr
1820						1825					1830			
Val	Trp	Phe	Val	Pro	Ser	Ile	Lys	Ala	Gly	Asn	Asp	Ile	Ala	Ala
1835						1840					1845			
Cys	Leu	Arg	Lys	Asn	Gly	Lys	Lys	Val	Ile	Gln	Leu	Ser	Arg	Lys
1850						1855					1860			
Thr	Phe	Asp	Ser	Glu	Tyr	Val	Lys	Thr	Arg	Thr	Asn	Asp	Trp	Asp
1865						1870					1875			
Phe	Val	Val	Thr	Thr	Asp	Ile	Ser	Glu	Met	Gly	Ala	Asn	Phe	Lys
1880						1885					1890			
Ala	Glu	Arg	Val	Ile	Asp	Pro	Arg	Arg	Cys	Met	Lys	Pro	Val	Ile
1895						1900					1905			
Leu	Thr	Asp	Gly	Glu	Glu	Arg	Val	Ile	Leu	Ala	Gly	Pro	Met	Pro
1910						1915					1920			
Val	Thr	His	Ser	Ser	Ala	Ala	Gln	Arg	Arg	Gly	Arg	Ile	Gly	Arg
1925						1930					1935			
Asn	Pro	Lys	Asn	Glu	Asn	Asp	Gln	Tyr	Ile	Tyr	Met	Gly	Glu	Pro
1940						1945					1950			
Leu	Glu	Asn	Asp	Glu	Asp	Cys	Ala	His	Trp	Lys	Glu	Ala	Lys	Met
1955						1960					1965			

Leu	Leu	Asp	Asn	Ile	Asn	Thr	Pro	Glu	Gly	Ile	Ile	Pro	Ser	Met
1970						1975					1980			
Phe	Glu	Pro	Glu	Arg	Glu	Lys	Val	Asp	Ala	Ile	Asp	Gly	Glu	Tyr
1985						1990					1995			
Arg	Leu	Arg	Gly	Glu	Ala	Arg	Lys	Thr	Phe	Val	Asp	Leu	Met	Arg
2000						2005					2010			
Arg	Gly	Asp	Leu	Pro	Val	Trp	Leu	Ala	Tyr	Arg	Val	Ala	Ala	Glu
2015						2020					2025			
Gly	Ile	Asn	Tyr	Ala	Asp	Arg	Arg	Trp	Cys	Phe	Asp	Gly	Val	Lys
2030						2035					2040			
Asn	Asn	Gln	Ile	Leu	Glu	Glu	Asn	Val	Glu	Val	Glu	Ile	Trp	Thr
2045						2050					2055			
Lys	Glu	Gly	Glu	Arg	Lys	Lys	Leu	Lys	Pro	Arg	Trp	Leu	Asp	Ala
2060						2065					2070			
Arg	Ile	Tyr	Ser	Asp	Pro	Leu	Ala	Leu	Lys	Glu	Phe	Lys	Glu	Phe
2075						2080					2085			
Ala	Ala	Gly	Arg	Lys	Ser	Leu	Thr	Leu	Asn	Leu	Ile	Thr	Glu	Met
2090						2095					2100			
Gly	Arg	Leu	Pro	Thr	Phe	Met	Thr	Gln	Lys	Ala	Arg	Asp	Ala	Leu
2105						2110					2115			
Asp	Asn	Leu	Ala	Val	Leu	His	Thr	Ala	Glu	Ala	Gly	Gly	Arg	Ala
2120						2125					2130			
Tyr	Asn	His	Ala	Leu	Ser	Glu	Leu	Pro	Glu	Thr	Leu	Glu	Thr	Leu
2135						2140					2145			
Leu	Leu	Leu	Thr	Leu	Leu	Ala	Thr	Val	Thr	Gly	Gly	Ile	Phe	Leu
2150						2155					2160			
Phe	Leu	Met	Ser	Ala	Arg	Gly	Ile	Gly	Lys	Met	Thr	Leu	Gly	Met
2165						2170					2175			

Cys	Cys	Ile	Ile	Thr	Ala	Ser	Ile	Leu	Leu	Trp	Tyr	Ala	Gln	Ile
2180						2185					2190			
Gln	Pro	His	Trp	Ile	Ala	Ala	Ser	Ile	Ile	Leu	Glu	Phe	Phe	Leu
2195						2200					2205			
Ile	Val	Leu	Leu	Ile	Pro	Glu	Pro	Glu	Lys	Gln	Arg	Thr	Pro	Gln
2210						2215					2220			
Asp	Asn	Gln	Leu	Thr	Tyr	Val	Val	Ile	Ala	Ile	Leu	Thr	Val	Val
2225						2230					2235			
Ala	Ala	Thr	Met	Ala	Asn	Glu	Met	Gly	Phe	Leu	Glu	Lys	Thr	Lys
2240						2245					2250			
Lys	Asp	Leu	Gly	Leu	Gly	Ser	Ile	Ala	Thr	Gln	Gln	Pro	Glu	Ser
2255						2260					2265			
Asn	Ile	Leu	Asp	Ile	Asp	Leu	Arg	Pro	Ala	Ser	Ala	Trp	Thr	Leu
2270						2275					2280			
Tyr	Ala	Val	Ala	Thr	Thr	Phe	Val	Thr	Pro	Met	Leu	Arg	His	Ser
2285						2290					2295			
Ile	Glu	Asn	Ser	Ser	Val	Asn	Val	Ser	Leu	Thr	Ala	Ile	Ala	Asn
2300						2305					2310			
Gln	Ala	Thr	Val	Leu	Met	Gly	Leu	Gly	Lys	Gly	Trp	Pro	Leu	Ser
2315						2320					2325			
Lys	Met	Asp	Ile	Gly	Val	Pro	Leu	Leu	Ala	Ile	Gly	Cys	Tyr	Ser
2330						2335					2340			
Gln	Val	Asn	Pro	Ile	Thr	Leu	Ile	Ala	Ala	Leu	Phe	Leu	Leu	Val
2345						2350					2355			
Ala	His	Tyr	Ala	Ile	Ile	Gly	Pro	Gly	Leu	Gln	Ala	Lys	Ala	Thr
2360						2365					2370			
Arg	Glu	Ala	Gln	Lys	Arg	Ala	Ala	Ala	Gly	Ile	Met	Lys	Asn	Pro
2375						2380					2385			
Thr	Val	Asp	Gly	Ile	Thr	Val	Ile	Asp	Leu	Asp	Pro	Ile	Pro	Tyr

2390		2395		2400
Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val				
2405		2410		2415
Leu Cys Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu				
2420		2425		2430
Cys Glu Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp				
2435		2440		2445
Glu Gly Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser				
2450		2455		2460
Met Ala Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu				
2465		2470		2475
Leu Phe Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr				
2480		2485		2490
Gly Asn Ile Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu				
2495		2500		2505
Asn Ala Leu Gly Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly				
2510		2515		2520
Ile Gln Glu Val Asp Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg				
2525		2530		2535
Gly Glu Thr Asp His His Ala Val Ser Arg Gly Ser Ala Lys Leu				
2540		2545		2550
Arg Trp Phe Val Glu Arg Asn Met Val Thr Pro Glu Gly Lys Val				
2555		2560		2565
Val Asp Leu Gly Cys Gly Arg Gly Gly Trp Ser Tyr Tyr Cys Gly				
2570		2575		2580
Gly Leu Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly Gly				
2585		2590		2595
Pro Gly His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn				
2600		2605		2610

Leu Val	Arg Leu Gln Ser	Gly Val Asp Val Phe Phe	Ile Pro Pro
2615		2620	2625
Glu Lys	Cys Asp Thr Leu Leu	Cys Asp Ile Gly Glu	Ser Ser Pro
2630		2635	2640
Asn Pro	Thr Val Glu Ala Gly	Arg Thr Leu Arg Val	Leu Asn Leu
2645		2650	2655
Val Glu	Asn Trp Leu Asn Asn	Asn Thr Gln Phe Cys	Ile Lys Val
2660		2665	2670
Leu Asn	Pro Tyr Met Pro Ser	Val Ile Glu Lys Met	Glu Ala Leu
2675		2680	2685
Gln Arg	Lys Tyr Gly Gly Ala	Leu Val Arg Asn Pro	Leu Ser Arg
2690		2695	2700
Asn Ser	Thr His Glu Met Tyr	Trp Val Ser Asn Ala	Ser Gly Asn
2705		2710	2715
Ile Val	Ser Ser Val Asn Met	Ile Ser Arg Met Leu	Ile Asn Arg
2720		2725	2730
Phe Thr	Met Arg Tyr Lys Lys	Ala Thr Tyr Glu Pro	Asp Val Asp
2735		2740	2745
Leu Gly	Ser Gly Thr Arg Asn	Ile Gly Ile Glu Ser	Glu Ile Pro
2750		2755	2760
Asn Leu	Asp Ile Ile Gly Lys	Arg Ile Glu Lys Ile	Lys Gln Glu
2765		2770	2775
His Glu	Thr Ser Trp His Tyr	Asp Gln Asp His Pro	Tyr Lys Thr
2780		2785	2790
Trp Ala	Tyr His Gly Ser Tyr	Glu Thr Lys Gln Thr	Gly Ser Ala
2795		2800	2805
Ser Ser	Met Val Asn Gly Val	Val Arg Leu Leu Thr	Lys Pro Trp
2810		2815	2820

Asp Val	Val Pro Met Val	Thr	Gln Met Ala Met	Thr	Asp Thr Thr
2825		2830		2835	
Pro Phe	Gly Gln Gln Arg	Val	Phe Lys Glu Lys	Val	Asp Thr Arg
2840		2845		2850	
Thr Gln	Glu Pro Lys Glu	Gly	Thr Lys Lys Leu	Met	Lys Ile Thr
2855		2860		2865	
Ala Glu	Trp Leu Trp Lys	Glu	Leu Gly Lys Lys	Lys	Thr Pro Arg
2870		2875		2880	
Met Cys	Thr Arg Glu Glu	Phe	Thr Arg Lys Val	Arg	Ser Asn Ala
2885		2890		2895	
Ala Leu	Gly Ala Ile Phe	Thr	Asp Glu Asn Lys	Trp	Lys Ser Ala
2900		2905		2910	
Arg Glu	Ala Val Glu Asp	Ser	Arg Phe Trp Glu	Leu	Val Asp Lys
2915		2920		2925	
Glu Arg	Asn Leu His Leu	Glu	Gly Lys Cys Glu	Thr	Cys Val Tyr
2930		2935		2940	
Asn Met	Met Gly Lys Arg	Glu	Lys Lys Leu Gly	Glu	Phe Gly Lys
2945		2950		2955	
Ala Lys	Gly Ser Arg Ala	Ile	Trp Tyr Met Trp	Leu	Gly Ala Arg
2960		2965		2970	
Phe Leu	Glu Phe Glu Ala	Leu	Gly Phe Leu Asn	Glu	Asp His Trp
2975		2980		2985	
Phe Ser	Arg Glu Asn Ser	Leu	Ser Gly Val Glu	Gly	Glu Gly Leu
2990		2995		3000	
His Lys	Leu Gly Tyr Ile	Leu	Arg Asp Val Ser	Lys	Lys Glu Gly
3005		3010		3015	
Gly Ala	Met Tyr Ala Asp	Asp	Thr Ala Gly Trp	Asp	Thr Arg Ile
3020		3025		3030	

Thr	Leu	Glu	Asp	Leu	Lys	Asn	Glu	Glu	Met	Val	Thr	Asn	His	Met
3035						3040					3045			
Glu	Gly	Glu	His	Lys	Lys	Leu	Ala	Glu	Ala	Ile	Phe	Lys	Leu	Thr
3050						3055					3060			
Tyr	Gln	Asn	Lys	Val	Val	Arg	Val	Gln	Arg	Pro	Thr	Pro	Arg	Gly
3065						3070					3075			
Thr	Val	Met	Asp	Ile	Ile	Ser	Arg	Arg	Asp	Gln	Arg	Gly	Ser	Gly
3080						3085					3090			
Gln	Val	Gly	Thr	Tyr	Gly	Leu	Asn	Thr	Phe	Thr	Asn	Met	Glu	Ala
3095						3100					3105			
Gln	Leu	Ile	Arg	Gln	Met	Glu	Gly	Glu	Gly	Val	Phe	Lys	Ser	Ile
3110						3115					3120			
Gln	His	Leu	Thr	Ile	Thr	Glu	Glu	Ile	Ala	Val	Gln	Asn	Trp	Leu
3125						3130					3135			
Ala	Arg	Val	Gly	Arg	Glu	Arg	Leu	Ser	Arg	Met	Ala	Ile	Ser	Gly
3140						3145					3150			
Asp	Asp	Cys	Val	Val	Lys	Pro	Leu	Asp	Asp	Arg	Phe	Ala	Ser	Ala
3155						3160					3165			
Leu	Thr	Ala	Leu	Asn	Asp	Met	Gly	Lys	Ile	Arg	Lys	Asp	Ile	Gln
3170						3175					3180			
Gln	Trp	Glu	Pro	Ser	Arg	Gly	Trp	Asn	Asp	Trp	Thr	Gln	Val	Pro
3185						3190					3195			
Phe	Cys	Ser	His	His	Phe	His	Glu	Leu	Ile	Met	Lys	Asp	Gly	Arg
3200						3205					3210			
Val	Leu	Val	Val	Pro	Cys	Arg	Asn	Gln	Asp	Glu	Leu	Ile	Gly	Arg
3215						3220					3225			
Ala	Arg	Ile	Ser	Gln	Gly	Ala	Gly	Trp	Ser	Leu	Arg	Glu	Thr	Ala
3230						3235					3240			
Cys	Leu	Gly	Lys	Ser	Tyr	Ala	Gln	Met	Trp	Ser	Leu	Met	Tyr	Phe

3245 3250 3255
 His Arg Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile Cys Ser Ala
 3260 3265 3270
 Val Pro Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp Ser Ile
 3275 3280 3285
 His Ala Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr Val
 3290 3295 3300
 Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys
 3305 3310 3315
 Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg
 3320 3325 3330
 Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala
 3335 3340 3345
 Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser
 3350 3355 3360
 Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys
 3365 3370 3375
 Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
 3380 3385 3390

 <210> 3
 <211> 3391
 <212> PRT
 <213> Trình tự nhân tạo

 <220>
 <223> Virut dengue typ huyết thanh 1, MVS

 <400> 3

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

 Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
 20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
 35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
 50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
 65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
 85 90 95

Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
 100 105 110

Met Ala Phe His Leu Thr Thr Arg Gly Gly Glu Pro His Met Ile Val
 115 120 125

Ser Lys Gln Glu Arg Gly Lys Ser Leu Leu Phe Lys Thr Ser Ala Gly
 130 135 140

Val Asn Met Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Leu Cys Glu
 145 150 155 160

Asp Thr Met Thr Tyr Lys Cys Pro Arg Ile Thr Glu Ala Glu Pro Asp
 165 170 175

Asp Val Asp Cys Trp Cys Asn Ala Thr Asp Thr Trp Val Thr Tyr Gly
 180 185 190

Thr Cys Ser Gln Thr Gly Glu His Arg Arg Asp Lys Arg Ser Val Ala
 195 200 205

Leu Ala Pro His Val Gly Leu Gly Leu Glu Thr Arg Ala Glu Thr Trp
 210 215 220

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

Ala Leu Arg His Pro Gly Phe Thr Val Ile Ala Leu Phe Leu Ala His
 245 250 255

Ala Ile Gly Thr Ser Ile Thr Gln Lys Gly Ile Ile Phe Ile Leu Leu

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

Pro Leu Pro Trp Thr Ser Gly Ala Ser Thr Ser Gln Glu Thr Trp Asn
 500 505 510

Arg Gln Asp Leu Leu Val Thr Phe Lys Thr Ala His Ala Lys Lys Gln
 515 520 525

Glu Val Val Val Leu Gly Ser Gln Glu Gly Ala Met His Thr Ala Leu
 530 535 540

Thr Gly Ala Thr Glu Ile Gln Thr Ser Gly Thr Thr Thr Ile Phe Ala
 545 550 555 560

Gly His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Thr Leu Lys Gly
 565 570 575

Met Ser Tyr Val Met Cys Thr Gly Ser Phe Lys Leu Glu Lys Glu Val
 580 585 590

Ala Glu Thr Gln His Gly Thr Val Leu Val Gln Val Lys Tyr Glu Gly
 595 600 605

Thr Asp Ala Pro Cys Lys Ile Pro Phe Ser Thr Gln Asp Glu Lys Gly
 610 615 620

Ala Thr Gln Asn Gly Arg Leu Ile Thr Ala Asn Pro Ile Val Thr Asp
 625 630 635 640

Lys Glu Lys Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Glu Ser
 645 650 655

Tyr Ile Val Val Gly Ala Gly Glu Lys Ala Leu Lys Leu Ser Trp Phe
 660 665 670

Lys Lys Gly Ser Ser Ile Gly Lys Met Phe Glu Ala Thr Ala Arg Gly
 675 680 685

Ala Arg Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser
 690 695 700

Ile Gly Gly Val Phe Thr Ser Met Gly Lys Leu Val His Gln Val Phe
 705 710 715 720

Gly Thr Ala Tyr Gly Val Leu Phe Ser Gly Val Ser Trp Thr Met Lys
725 730 735

Ile Gly Ile Gly Ile Leu Leu Thr Trp Leu Gly Leu Asn Ser Arg Asn
740 745 750

Thr Ser Leu Ser Met Met Cys Ile Ala Ala Gly Ile Val Thr Leu Tyr
755 760 765

Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
770 775 780

Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
785 790 795 800

His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
805 810 815

Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
820 825 830

Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
835 840 845

Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
850 855 860

Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
865 870 875 880

Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
885 890 895

Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
900 905 910

Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
915 920 925

Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
930 935 940

Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
 945 950 955 960

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
 965 970 975

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
 980 985 990

Ile Glu Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser
 995 1000 1005

Asn Gly Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala
 1010 1015 1020

Gly Pro Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln
 1025 1030 1035

Ile Thr Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp
 1040 1045 1050

Phe Cys Asp Gly Thr Thr Val Val Val Thr Glu Asp Cys Gly Asn
 1055 1060 1065

Arg Gly Pro Ser Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile
 1070 1075 1080

Thr Glu Trp Cys Cys Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr
 1085 1090 1095

Arg Gly Glu Asp Gly Cys Trp Tyr Gly Met Glu Ile Arg Pro Leu
 1100 1105 1110

Lys Glu Lys Glu Glu Asn Leu Val Asn Ser Leu Val Thr Ala Gly
 1115 1120 1125

His Gly Gln Val Asp Asn Phe Ser Leu Gly Val Leu Gly Met Ala
 1130 1135 1140

Leu Phe Leu Glu Glu Met Leu Arg Thr Arg Val Gly Thr Lys His
 1145 1150 1155

Ala Ile Leu Leu Val Ala Val Ser Phe Val Thr Leu Ile Thr Gly

1160		1165		1170
Asn Met Ser Phe Arg Asp Leu Gly Arg Val Met Val Met Val Gly				
1175		1180		1185
Ala Thr Met Thr Asp Asp Ile Gly Met Gly Val Thr Tyr Leu Ala				
1190		1195		1200
Leu Leu Ala Ala Phe Lys Val Arg Pro Thr Phe Ala Ala Gly Leu				
1205		1210		1215
Leu Leu Arg Lys Leu Thr Ser Lys Glu Leu Met Met Thr Thr Ile				
1220		1225		1230
Gly Ile Val Leu Leu Ser Gln Ser Thr Leu Pro Glu Thr Ile Leu				
1235		1240		1245
Glu Leu Thr Asp Ala Leu Ala Leu Gly Met Met Val Leu Lys Met				
1250		1255		1260
Val Arg Asn Met Glu Lys Tyr Gln Leu Ala Val Thr Ile Met Ala				
1265		1270		1275
Ile Leu Cys Val Pro Asn Ala Val Ile Leu Gln Asn Ala Trp Lys				
1280		1285		1290
Val Ser Cys Thr Ile Leu Ala Val Val Ser Val Ser Pro Leu Phe				
1295		1300		1305
Leu Thr Ser Ser Gln Gln Lys Thr Asp Trp Ile Pro Leu Ala Leu				
1310		1315		1320
Thr Ile Lys Gly Leu Asn Pro Thr Ala Ile Phe Leu Thr Thr Leu				
1325		1330		1335
Ser Arg Thr Ser Lys Lys Arg Ser Trp Pro Leu Asn Glu Ala Ile				
1340		1345		1350
Met Ala Val Gly Met Val Ser Ile Leu Ala Ser Ser Leu Leu Lys				
1355		1360		1365
Asn Asp Ile Pro Met Thr Gly Pro Leu Val Ala Gly Gly Leu Leu				
1370		1375		1380

Thr Val	Cys Tyr Val Leu	Thr Gly Arg Ser Ala Asp	Leu Glu Leu
1385		1390	1395
Glu Arg	Ala Ala Asp Val	Lys Trp Glu Asp Gln Ala	Glu Ile Ser
1400		1405	1410
Gly Ser	Ser Pro Ile Leu	Ser Ile Thr Ile Ser Glu	Asp Gly Ser
1415		1420	1425
Met Ser	Ile Lys Asn Glu	Glu Glu Asp Gln Thr	Leu Thr Ile Leu
1430		1435	1440
Ile Arg	Thr Gly Leu Leu Val	Ile Ser Gly Leu Phe	Pro Val Ser
1445		1450	1455
Ile Pro	Ile Thr Ala Ala Ala	Trp Tyr Leu Trp Glu	Val Lys Lys
1460		1465	1470
Gln Arg	Ala Gly Val Leu	Trp Asp Val Pro Ser	Pro Pro Pro Met
1475		1480	1485
Gly Lys	Ala Glu Leu Glu	Asp Gly Ala Tyr Arg	Ile Lys Gln Lys
1490		1495	1500
Gly Ile	Leu Gly Tyr Ser	Gln Ile Gly Ala Gly	Val Tyr Lys Glu
1505		1510	1515
Gly Thr	Phe His Thr Met	Trp His Val Thr Arg	Gly Ala Val Leu
1520		1525	1530
Met His	Lys Gly Lys Arg	Ile Glu Pro Ser Trp	Ala Asp Val Lys
1535		1540	1545
Lys Asp	Leu Ile Ser Tyr	Gly Gly Gly Trp Lys	Leu Glu Gly Glu
1550		1555	1560
Trp Lys	Glu Gly Glu Glu	Val Gln Val Leu Ala	Leu Glu Pro Gly
1565		1570	1575
Lys Asn	Pro Arg Ala Val	Gln Thr Lys Pro Gly	Leu Phe Lys Thr
1580		1585	1590

Asn	Ala	Gly	Thr	Ile	Gly	Ala	Val	Ser	Leu	Asp	Phe	Ser	Pro	Gly
1595						1600					1605			
Thr	Ser	Gly	Ser	Pro	Ile	Ile	Asp	Lys	Lys	Gly	Lys	Val	Val	Gly
1610						1615					1620			
Leu	Tyr	Gly	Asn	Gly	Val	Val	Thr	Arg	Ser	Gly	Ala	Tyr	Val	Ser
1625						1630					1635			
Ala	Ile	Ala	Gln	Thr	Glu	Lys	Ser	Ile	Glu	Asp	Asn	Pro	Glu	Ile
1640						1645					1650			
Glu	Asp	Asp	Ile	Phe	Arg	Lys	Arg	Arg	Leu	Thr	Ile	Met	Asp	Leu
1655						1660					1665			
His	Pro	Gly	Ala	Gly	Lys	Thr	Lys	Arg	Tyr	Leu	Pro	Ala	Ile	Val
1670						1675					1680			
Arg	Glu	Ala	Ile	Lys	Arg	Gly	Leu	Arg	Thr	Leu	Ile	Leu	Ala	Pro
1685						1690					1695			
Thr	Arg	Val	Val	Ala	Ala	Glu	Met	Glu	Glu	Ala	Leu	Arg	Gly	Leu
1700						1705					1710			
Pro	Ile	Arg	Tyr	Gln	Thr	Pro	Ala	Ile	Arg	Ala	Val	His	Thr	Gly
1715						1720					1725			
Arg	Glu	Ile	Val	Asp	Leu	Met	Cys	His	Ala	Thr	Phe	Thr	Met	Arg
1730						1735					1740			
Leu	Leu	Ser	Pro	Val	Arg	Val	Pro	Asn	Tyr	Asn	Leu	Ile	Ile	Met
1745						1750					1755			
Asp	Glu	Ala	His	Phe	Thr	Asp	Pro	Ala	Ser	Ile	Ala	Ala	Arg	Gly
1760						1765					1770			
Tyr	Ile	Ser	Thr	Arg	Val	Glu	Met	Gly	Glu	Ala	Ala	Gly	Ile	Phe
1775						1780					1785			
Met	Thr	Ala	Thr	Pro	Pro	Gly	Ser	Arg	Asp	Pro	Phe	Pro	Gln	Ser
1790						1795					1800			

Asn	Ala	Pro	Ile	Ile	Asp	Glu	Glu	Arg	Glu	Ile	Pro	Glu	Arg	Ser
1805						1810					1815			
Trp	Asn	Ser	Gly	His	Glu	Trp	Val	Thr	Asp	Phe	Lys	Gly	Lys	Thr
1820						1825					1830			
Val	Trp	Phe	Val	Pro	Ser	Ile	Lys	Ala	Gly	Asn	Asp	Ile	Ala	Ala
1835						1840					1845			
Cys	Leu	Arg	Lys	Asn	Gly	Lys	Lys	Val	Ile	Gln	Leu	Ser	Arg	Lys
1850						1855					1860			
Thr	Phe	Asp	Ser	Glu	Tyr	Val	Lys	Thr	Arg	Thr	Asn	Asp	Trp	Asp
1865						1870					1875			
Phe	Val	Val	Thr	Thr	Asp	Ile	Ser	Glu	Met	Gly	Ala	Asn	Phe	Lys
1880						1885					1890			
Ala	Glu	Arg	Val	Ile	Asp	Pro	Arg	Arg	Cys	Met	Lys	Pro	Val	Ile
1895						1900					1905			
Leu	Thr	Asp	Gly	Glu	Glu	Arg	Val	Ile	Leu	Ala	Gly	Pro	Met	Pro
1910						1915					1920			
Val	Thr	His	Ser	Ser	Ala	Ala	Gln	Arg	Arg	Gly	Arg	Ile	Gly	Arg
1925						1930					1935			
Asn	Pro	Lys	Asn	Glu	Asn	Asp	Gln	Tyr	Ile	Tyr	Met	Gly	Glu	Pro
1940						1945					1950			
Leu	Glu	Asn	Asp	Glu	Asp	Cys	Ala	His	Trp	Lys	Glu	Ala	Lys	Met
1955						1960					1965			
Leu	Leu	Asp	Asn	Ile	Asn	Thr	Pro	Glu	Gly	Ile	Ile	Pro	Ser	Met
1970						1975					1980			
Phe	Glu	Pro	Glu	Arg	Glu	Lys	Val	Asp	Ala	Ile	Asp	Gly	Glu	Tyr
1985						1990					1995			
Arg	Leu	Arg	Gly	Glu	Ala	Arg	Lys	Thr	Phe	Val	Asp	Leu	Met	Arg
2000						2005					2010			
Arg	Gly	Asp	Leu	Pro	Val	Trp	Leu	Ala	Tyr	Arg	Val	Ala	Ala	Glu

2015	2020	2025
Gly Ile Asn Tyr Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys 2030 2035 2040		
Asn Asn Gln Ile Leu Glu Glu Asn Val Glu Val Glu Ile Trp Thr 2045 2050 2055		
Lys Glu Gly Glu Arg Lys Lys Leu Lys Pro Arg Trp Leu Asp Ala 2060 2065 2070		
Arg Ile Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe Lys Glu Phe 2075 2080 2085		
Ala Ala Gly Arg Lys Ser Leu Thr Leu Asn Leu Ile Thr Glu Met 2090 2095 2100		
Gly Arg Leu Pro Thr Phe Met Thr Gln Lys Ala Arg Asp Ala Leu 2105 2110 2115		
Asp Asn Leu Ala Val Leu His Thr Ala Glu Ala Gly Gly Arg Ala 2120 2125 2130		
Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu 2135 2140 2145		
Leu Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu 2150 2155 2160		
Phe Leu Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met 2165 2170 2175		
Cys Cys Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile 2180 2185 2190		
Gln Pro His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu 2195 2200 2205		
Ile Val Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln 2210 2215 2220		
Asp Asn Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val 2225 2230 2235		

Ala	Ala	Thr	Met	Ala	Asn	Glu	Met	Gly	Phe	Leu	Glu	Lys	Thr	Lys
2240						2245					2250			
Lys	Asp	Leu	Gly	Leu	Gly	Ser	Ile	Ala	Thr	Gln	Gln	Pro	Glu	Ser
2255						2260					2265			
Asn	Ile	Leu	Asp	Ile	Asp	Leu	Arg	Pro	Ala	Ser	Ala	Trp	Thr	Leu
2270						2275					2280			
Tyr	Ala	Val	Ala	Thr	Thr	Phe	Val	Thr	Pro	Met	Leu	Arg	His	Ser
2285						2290					2295			
Ile	Glu	Asn	Ser	Ser	Val	Asn	Val	Ser	Leu	Thr	Ala	Ile	Ala	Asn
2300						2305					2310			
Gln	Ala	Thr	Val	Leu	Met	Gly	Leu	Gly	Lys	Gly	Trp	Pro	Leu	Ser
2315						2320					2325			
Lys	Met	Asp	Ile	Gly	Val	Pro	Leu	Leu	Ala	Ile	Gly	Cys	Tyr	Ser
2330						2335					2340			
Gln	Val	Asn	Pro	Ile	Thr	Leu	Thr	Ala	Ala	Leu	Phe	Leu	Leu	Val
2345						2350					2355			
Ala	His	Tyr	Ala	Ile	Ile	Gly	Pro	Gly	Leu	Gln	Ala	Lys	Ala	Thr
2360						2365					2370			
Arg	Glu	Ala	Gln	Lys	Arg	Ala	Ala	Ala	Gly	Ile	Met	Lys	Asn	Pro
2375						2380					2385			
Thr	Val	Asp	Gly	Ile	Thr	Val	Ile	Asp	Leu	Asp	Pro	Ile	Pro	Tyr
2390						2395					2400			
Asp	Pro	Lys	Phe	Glu	Lys	Gln	Leu	Gly	Gln	Val	Met	Leu	Leu	Val
2405						2410					2415			
Leu	Cys	Val	Thr	Gln	Val	Leu	Met	Met	Arg	Thr	Thr	Trp	Ala	Leu
2420						2425					2430			
Cys	Glu	Ala	Leu	Thr	Leu	Ala	Thr	Gly	Pro	Ile	Ser	Thr	Leu	Trp
2435						2440					2445			

Glu Gly	Asn Pro Gly Arg	Phe	Trp Asn Thr Thr	Ile	Ala Val Ser
2450		2455		2460	
Met Ala	Asn Ile Phe Arg	Gly	Ser Tyr Leu Ala	Gly	Ala Gly Leu
2465		2470		2475	
Leu Phe	Ser Ile Met Lys	Asn	Thr Thr Asn Thr	Arg	Arg Gly Thr
2480		2485		2490	
Gly Asn	Ile Gly Glu Thr	Leu	Gly Glu Lys Trp	Lys	Ser Arg Leu
2495		2500		2505	
Asn Ala	Leu Gly Lys Ser	Glu	Phe Gln Ile Tyr	Lys	Lys Ser Gly
2510		2515		2520	
Ile Gln	Glu Val Asp Arg	Thr	Leu Ala Lys Glu	Gly	Ile Lys Arg
2525		2530		2535	
Gly Glu	Thr Asp His His	Ala	Val Ser Arg Gly	Ser	Ala Lys Leu
2540		2545		2550	
Arg Trp	Phe Val Glu Arg	Asn	Met Val Thr Pro	Glu	Gly Lys Val
2555		2560		2565	
Val Asp	Leu Gly Cys Gly	Arg	Gly Gly Trp Ser	Tyr	Tyr Cys Gly
2570		2575		2580	
Gly Leu	Lys Asn Val Arg	Glu	Val Lys Gly Leu	Thr	Lys Gly Gly
2585		2590		2595	
Pro Gly	His Glu Glu Pro	Ile	Pro Met Ser Thr	Tyr	Gly Trp Asn
2600		2605		2610	
Leu Val	Arg Leu Gln Ser	Gly	Val Asp Val Phe	Phe	Ile Pro Pro
2615		2620		2625	
Glu Lys	Cys Asp Thr Leu	Leu	Cys Asp Ile Gly	Glu	Ser Ser Pro
2630		2635		2640	
Asn Pro	Thr Val Glu Ala	Gly	Arg Thr Leu Arg	Val	Leu Asn Leu
2645		2650		2655	

Val	Glu	Asn	Trp	Leu	Asn	Asn	Asn	Thr	Gln	Phe	Cys	Ile	Lys	Val
2660						2665					2670			
Leu	Asn	Pro	Tyr	Met	Pro	Ser	Val	Ile	Glu	Lys	Met	Glu	Ala	Leu
2675						2680					2685			
Gln	Arg	Lys	Tyr	Gly	Gly	Ala	Leu	Val	Arg	Asn	Pro	Leu	Ser	Arg
2690						2695					2700			
Asn	Ser	Thr	His	Glu	Met	Tyr	Trp	Val	Ser	Asn	Ala	Ser	Gly	Asn
2705						2710					2715			
Ile	Val	Ser	Ser	Val	Asn	Met	Ile	Ser	Arg	Met	Leu	Ile	Asn	Arg
2720						2725					2730			
Phe	Thr	Met	Arg	Tyr	Lys	Lys	Ala	Thr	Tyr	Glu	Pro	Asp	Val	Asp
2735						2740					2745			
Leu	Gly	Ser	Gly	Thr	Arg	Asn	Ile	Gly	Ile	Glu	Ser	Glu	Ile	Pro
2750						2755					2760			
Asn	Leu	Asp	Ile	Ile	Gly	Lys	Arg	Ile	Glu	Lys	Ile	Lys	Gln	Glu
2765						2770					2775			
His	Glu	Thr	Ser	Trp	His	Tyr	Asp	Gln	Asp	His	Pro	Tyr	Lys	Thr
2780						2785					2790			
Trp	Ala	Tyr	His	Gly	Ser	Tyr	Glu	Thr	Lys	Gln	Thr	Gly	Ser	Ala
2795						2800					2805			
Ser	Ser	Met	Val	Asn	Gly	Val	Val	Arg	Leu	Leu	Thr	Lys	Pro	Trp
2810						2815					2820			
Asp	Val	Val	Pro	Met	Val	Thr	Gln	Met	Ala	Met	Thr	Asp	Thr	Thr
2825						2830					2835			
Pro	Phe	Gly	Gln	Gln	Arg	Val	Phe	Lys	Glu	Lys	Val	Asp	Thr	Arg
2840						2845					2850			
Thr	Gln	Glu	Pro	Lys	Glu	Gly	Thr	Lys	Lys	Leu	Met	Lys	Ile	Thr
2855						2860					2865			
Ala	Glu	Trp	Leu	Trp	Lys	Glu	Leu	Gly	Lys	Lys	Lys	Thr	Pro	Arg

2870		2875		2880
Met Cys Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser Asn Ala				
2885		2890		2895
Ala Leu Gly Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser Ala				
2900		2905		2910
Arg Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys				
2915		2920		2925
Glu Arg Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr				
2930		2935		2940
Asn Met Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys				
2945		2950		2955
Ala Lys Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg				
2960		2965		2970
Phe Leu Glu Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp				
2975		2980		2985
Phe Ser Arg Glu Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu				
2990		2995		3000
His Lys Leu Gly Tyr Ile Leu Arg Asp Val Ser Lys Lys Glu Gly				
3005		3010		3015
Gly Ala Met Tyr Ala Asp Asp Thr Ala Gly Trp Asp Thr Arg Ile				
3020		3025		3030
Thr Leu Glu Asp Leu Lys Asn Glu Glu Met Val Thr Asn His Met				
3035		3040		3045
Glu Gly Glu His Lys Lys Leu Ala Glu Ala Ile Phe Lys Leu Thr				
3050		3055		3060
Tyr Gln Asn Lys Val Val Arg Val Gln Arg Pro Thr Pro Arg Gly				
3065		3070		3075
Thr Val Met Asp Ile Ile Ser Arg Arg Asp Gln Arg Gly Ser Gly				
3080		3085		3090

Gln Val	Gly Thr Tyr Gly	Leu	Asn Thr Phe Thr	Asn	Met Glu Ala
3095		3100		3105	
Gln Leu	Ile Arg Gln Met	Glu	Gly Glu Gly Val	Phe	Lys Ser Ile
3110		3115		3120	
Gln His	Leu Thr Ile Thr	Glu	Glu Ile Ala Val	Gln	Asn Trp Leu
3125		3130		3135	
Ala Arg	Val Gly Arg Glu	Arg	Leu Ser Arg Met	Ala	Ile Ser Gly
3140		3145		3150	
Asp Asp	Cys Val Val Lys	Pro	Leu Asp Asp Arg	Phe	Ala Ser Ala
3155		3160		3165	
Leu Thr	Ala Leu Asn Asp	Met	Gly Lys Ile Arg	Lys	Asp Ile Gln
3170		3175		3180	
Gln Trp	Glu Pro Ser Arg	Gly	Trp Asn Asp Trp	Thr	Gln Val Pro
3185		3190		3195	
Phe Cys	Ser His His Phe	His	Glu Leu Ile Met	Lys	Asp Gly Arg
3200		3205		3210	
Val Leu	Val Val Pro Cys	Arg	Asn Gln Asp Glu	Leu	Ile Gly Arg
3215		3220		3225	
Ala Arg	Ile Ser Gln Gly	Ala	Gly Trp Ser Leu	Arg	Glu Thr Ala
3230		3235		3240	
Cys Leu	Gly Lys Ser Tyr	Ala	Gln Met Trp Ser	Leu	Met Tyr Phe
3245		3250		3255	
His Arg	Arg Asp Leu Arg	Leu	Ala Ala Asn Ala	Ile	Cys Ser Ala
3260		3265		3270	
Val Pro	Ser His Trp Val	Pro	Thr Ser Arg Thr	Thr	Trp Ser Ile
3275		3280		3285	
His Ala	Lys His Glu Trp	Met	Thr Thr Glu Asp	Met	Leu Thr Val
3290		3295		3300	

Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys
3305 3310 3315

Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg
3320 3325 3330

Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala
3335 3340 3345

Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser
3350 3355 3360

Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys
3365 3370 3375

Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
3380 3385 3390

<210> 4

<211> 10723

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Virut dengue typ huyết thanh 2, BVS

<400> 4

```

agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta      60
gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg      120
aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag      180
ctgacaaaga gattctcact tggaatgctg cagggacgag gaccattaaa actgttcattg      240
gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga      300
tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt      360
ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat cattatgctg      420
attccaacag tgatggcggt ccatттаacc acacgtaacg gagaaccaca catgatcgtc      480
agcagacaag agaaagggaa aagtcttctg tttaaaacag aggttggcgt gaacatgtgt      540
acctcatgg ccatggacct tggatgaattg tgtgaagaca caatcacgta cgagtgtccc      600
cttctcaggc agaatgagcc agaagacata gactgttggt gcaactctac gtccacgtgg      660
gtaacttatg ggacgtgtac caccatggga gaacatagaa gagaaaaaag atcagtggca      720

```

ctcgttccac atgtgggaat gggactggag acacgaactg aaacatggat gtcacagaa	780
ggggcctgga aacatgtcca gagaattgaa acttgatct tgagacatcc aggcttcacc	840
atgatggcag caatcctggc atacaccata ggaacgacac atttccaaag agccctgac	900
ttcatcttac tgacagctgt cactccttca atgacaatgc gttgcatagg aatgtcaa	960
agagactttg tggaaggggt ttcaggagga agctgggttg acatagtctt agaacatgga	1020
agctgtgtga cgacgatggc aaaaaacaaa ccaacattgg attttgaact gataaaaaca	1080
gaagccaaac agcctgccac cctaaggaag tactgtatag aggcaaagct aaccaacaca	1140
acaacagaat ctcgtgccc aacacaaggg gaaccagcc taaatgaaga gcaggacaaa	1200
aggttcgtct gcaaactc catggtagac agaggatggg gaaatggatg tggactat	1260
ggaaagggag gcattgtgac ctgtgctatg ttcagatgca aaaagaacat ggaaggaaaa	1320
gttgtgcaac cagaaaactt ggaatacacc attgtgataa cacctcactc aggggaagag	1380
catgcagtcg gaaatgacac aggaaaacat ggcaaggaaa tcaaaataac accacagagt	1440
tccatcacag aagcagaatt gacaggttat ggcactgtca caatggagtg ctctccaaga	1500
acgggcctcg acttcaatga gatggtgttg ctgcagatgg aaaataaagc ttggctggtg	1560
cacaggcaat ggttcctaga cctgccgtta ccatgggtgc ccggagcggg cacacaaggg	1620
tcaaattgga tacagaaaga gacattggtc actttcaaaa atccccatgc gaagaaacag	1680
gatgttgttg ttttaggatc ccaagaaggg gccatgcaca cagcacttac aggggccaca	1740
gaaatccaaa tgtcatcagg aaacttactc ttcacaggac atctcaagtg caggctgaga	1800
atggacaagc tacagctcaa aggaatgtca tactctatgt gcacaggaaa gtttaaagtt	1860
gtgaaggaaa tagcagaaac acaacatgga acaatagtta tcagagtgca atatgaaggg	1920
gacggctctc catgcaagat cccttttgag ataatggatt tggaaaaaag acatgtctta	1980
ggtcgcctga ttacagtcaa cccaattgtg acagaaaaag atagcccagt caacatagaa	2040
gcagaacctc catttggaag cagctacatc atcataggag tagagccggg acaactgaag	2100
ctcaactggg ttaagaaagg aagttctatc ggccaaatgt ttgagacaac aatgaggggg	2160
gcgaagagaa tggccatfff aggtgacaca gcctgggatt ttggatcctt gggaggagt	2220
tttacatcta taggaaaggc tctccaccaa gtctttggag caatctatgg agctgccttc	2280
agtggggttt catggactat gaaaatcctc ataggagtca ttatcacatg gataggaatg	2340
aattcacgca gcacctcact gtctgtgaca ctagtattgg tgggaattgt gacactgtat	2400

ttgggagtca tgggtgcaggc cgatagtggg tgcgttgtga gctggaaaaa caaagaactg	2460
aaatgtggca gtgggatttt catcacagac aacgtgcaca catggacaga acaatacaag	2520
ttccaaccag aatcccccttc aaaactagct tcagctatcc agaaagccca tgaagaggac	2580
atgtgtggaa tccgctcagt aacaagactg gagaatctga tgtggaaaca aataacacca	2640
gaattgaatc acattctatc agaaaatgag gtgaagttaa ctattatgac aggagacatc	2700
aaaggaatca tgcaggcagg aaaacgatct ctgcggcctc agcccactga gctgaagtat	2760
tcatggaaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt	2820
ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg	2880
gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa	2940
aaacaggatg tattctgcga ctcaaaactc atgtcagcgg ccataaaaaga caacagagcc	3000
gtccatgccg atatgggtta ttggatagaa agtgcactca atgacacatg gaagatagag	3060
aaagcctctt tcattgaagt taaaaactgc cactggccaa aatcacacac cctctggagc	3120
aatggagtgc tagaaagtga gatgataatt ccaaagaatc tcgctggacc agtgtctcaa	3180
cacaactata gaccaggcta ccatacacaa ataacaggac catggcatct aggtaagctt	3240
gagatggact ttgattttctg tgatggaaca acagtggtag tgactgagga ctgcggaaat	3300
agaggaccct ctttgagaac aaccactgcc tctggaaaac tcataacaga atgggtgctgc	3360
cgatcttgca cattaccacc gctaagatac agagggtgagg atgggtgctg gtacgggatg	3420
gaaatcagac cattgaagga gaaagaagag aatttggtca actccttggc cacagctgga	3480
catgggcagg tcgacaactt ttcactagga gtcttgggaa tggcattgtt cctggaggaa	3540
atgcttagga cccgagtagg aacgaaacat gcaatactac tagttgcagt ttctttttgtg	3600
acattgatca cagggaacat gtcctttaga gacctgggaa gagtgatggc tatggtaggc	3660
gccactatga cggatgacat aggtatgggc gtgacttata ttgccctact agcagccttc	3720
aaagtcagac caacttttgc agctggacta ctcttgagaa agctgacctc caaggaattg	3780
atgatgacta ctataggaat tgtactcctc tcccagagca ccataccaga gaccattctt	3840
gagttgactg atgcgttagc cttaggcatg atggctctca aaatggtgag aaatatggaa	3900
aagtatcaat tggcagtgc tatcatggct atcttgtgcg tcccaaacgc agtgatatta	3960
caaaacgcat ggaaagtgag ttgcacaata ttggcagtgg tgtccgtttc cccactgttc	4020
ttaacatcct cacagcaaaa aacagattgg ataccattag cattgacgat caaaggcttc	4080
aatccaacag ctatTTTTTct aacaaccctc tcaagaacca gcaagaaaag gagctggcca	4140

ttaaatgagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa	4200
aatgatattc ccatgacagg accattagtg gctggagggc tcctcactgt gtgctacgtg	4260
ctcactggac gatcggccga tttggaactg gagagagcag ccgatgtcaa atgggaagac	4320
caggcagaga tatcaggaag cagtccaatc ctgtcaataa caatatcaga agatggtagc	4380
atgtcgataa aaaatgaaga ggaagaacaa aactgacca tactcattag aacaggattg	4440
ctggtgatct caggactttt tcctgtatca ataccaatca cggcagcagc atggtacctg	4500
tgggaagtga agaaacaacg ggccggagta ttgtgggatg ttccttcacc cccacccatg	4560
ggaaaggctg aactggaaga tggagcctat agaattaagc aaaaagggat tcttgatat	4620
tcccagatcg gagccggagt ttacaaagaa ggaacattcc atacaatgtg gcatgtcaca	4680
cgtggcgtg ttctaattgca taaaggaaag aggattgaac catcatgggc ggacgtcaag	4740
aaagacctaa tatcatatgg aggaggctgg aagttagaag gagaatggaa ggaaggagaa	4800
gaagtccagg tattggcact ggagcctgga aaaaatccaa gagccgtcca aacgaaacct	4860
ggtcttttca aaaccaacgc cggaacaata ggtgctgtat ctctggactt ttctcctgga	4920
acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt	4980
gttgttacaa ggagtggagc atatgtgagt gctatagccc agactgaaaa aagcattgaa	5040
gacaaccag agatcgaaga tgacattttc cgaaagagaa gactgaccat catggacctc	5100
caccaggag cgggaaagac gaagagatac cttccggcca tagtcagaga agctataaaa	5160
cggggtttga gaacattaat cttggcccc actagagttg tggcagctga aatggaggaa	5220
gcccttagag gacttccaat aagataccag accccagcca tcagagctgt gcacaccggg	5280
cgggagattg tggacctaat gtgtcatgcc acatttacca tgaggctgct atcaccagtt	5340
agagtgccaa actacaacct gattatcatg gacgaagccc atttcacaga cccagcaagt	5400
atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt	5460
atgacagcca ctccccggg aagcagagac ccatctctc agagcaatgc accaatcata	5520
gatgaagaaa gagaaatccc tgaacgctcg tggaaattccg gacatgaatg ggtcacggat	5580
tttaaaggga agactgtttg gtctgtcca agtataaaag caggaaatga tatagcagct	5640
tgctgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag	5700
tatgtcaaga ctagaaccaa tgattgggac ttogtggtta caactgacat ttcagaaatg	5760
ggtgccaat tcaaggctga gaggggtata gaccccagac gctgcatgaa accagtcata	5820

ctaacagatg gtgaagagcg ggtgattctg gcaggaccta tgccagtgac ccactctagt	5880
gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata	5940
tacatggggg aacctctgga aatgatgaa gactgtgcac actggaaaga agctaaaatg	6000
ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt	6060
gaaaagggtg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt	6120
gtagacttaa tgagaagagg agacctacca gtctggttgg cctacagagt ggcagctgaa	6180
ggcatcaact acgcagacag aagggtggtg tttgatggag tcaagaacaa ccaaactcta	6240
gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc	6300
agatggttgg atgctaggat ctattctgac ccactggcgc taaaagaatt taaggaattt	6360
gcagccggaa gaaagtctct gacctgaac ctaatcacag aatgggtag gctcccaacc	6420
ttcatgactc agaaggcaag agacgcactg gacaacttag cagtgctgca cacggctgag	6480
ccaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg	6540
cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca	6600
aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta	6660
tggtacgcac aaatacagcc aactggata gcagcttcaa taatactgga gttttttctc	6720
atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc	6780
tacgttgtca tagccatcct cacagtggtg gccgcaacca tggcaaacga gatgggtttc	6840
ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caaccagca acccgagagc	6900
aacatcctgg acatagatct acgtcctgca tcagcatgga cgctgtatgc cgtggccaca	6960
acatttgtaa caccaatgtt gagacatagc attgaaaatt cctcagtgaa tgtgtcccta	7020
acagctatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca	7080
aagatggaca tcggagttcc cttctcgcgc attggatgct actcacaagt caaccccata	7140
actctcacag cagctttttt cttattggta gcacattatg ccatcatagg gccaggactc	7200
caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaacca	7260
actgtcgatg gaataacagt gattgacctc gatccaatac cttatgatcc aaagtttgaa	7320
aagcagttgg gacaagtaat gtccttagtc ctctgcgtga ctcaagtatt gatgatgagg	7380
actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg	7440
gaaggaaatc caggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt	7500
agagggagtt acttggccgg agctggactt ctcttttcta ttatgaagaa cacaaccaac	7560

acaagaaggg gaactggcaa cataggagag acgcttggag agaaatggaa aagccgattg	7620
aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat	7680
agaaccttag caaaagaagg cattaaaaga ggagaaacgg accatcacgc tgtgtcgcga	7740
ggctcagcaa aactgagatg gttcgttgag agaaacatgg tcacaccaga agggaaagta	7800
gtggacctcg gttgtggcag aggaggctgg tcatactatt gtggaggact aaagaatgta	7860
agagaagtca aaggcctaac aaaaggagga ccaggacacg aagaacccat ccccatgtca	7920
acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcca	7980
gaaaagtgtg acacattatt gtgtgacata ggggagtcac caccaaatcc cacagtggaa	8040
gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa	8100
ttttgcataa aggttctcaa cccatatatg ccctcagtca tagaaaaaat ggaagcacta	8160
caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag	8220
atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg	8280
atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac	8340
ctcgaagcg gaacccgtaa catcgggatt gaaagtgaga taccaaacct agatataatt	8400
gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac	8460
caccataca aaacgtgggc ataccatggg agctatgaaa caaacagac tggatcagca	8520
tcatccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtcccatg	8580
gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgcgt ttttaaagag	8640
aaagtggaca cgagaaccca agaaccgaaa gaaggcacga agaaactaat gaaaataaca	8700
gcagagtggc tttggaaaga attagggag aaaaagacac ccaggatgtg caccagagaa	8760
gaattcacia gaaaggtgag aagcaatgca gccttggggg ccgtattcac tgatgagaac	8820
aagtggaagt cggcacgtga ggctgttgaa gatagtaggt tttgggagct ggttgacaag	8880
gaaaggaatc tccatcttga aggaaagtgt gaaacatgtg tgtacaacat gatgggaaaa	8940
agagagaaga agctagggga attcggcaag gcaaaaggca gcagagccat atggtacatg	9000
tggcttggag cacgcttctt agagtttgaa gccctaggat tcttaaataga agatcactgg	9060
ttctccagag agaactccct gagtggagtg gaaggagaag ggctgcacia gctaggttac	9120
attctaagag acgtgagcaa gaaagaggga ggagcaatgt atgccgatga caccgcagga	9180
tgggatacaa gaatcacact agaagaccta aaaaatgaag aaatggtaac aaaccacatg	9240

```

gaaggagaac acaagaaact agccgaggcc attttcaaac taacgtacca aaacaagggtg 9300
gtgctgtgtc aaagaccaac accaagaggc acagtaatgg acatcataatc gagaagagac 9360
caaagaggta gtggacaagt tggcacctat ggactcaata ctttcaccaa tatggaagcc 9420
caactaatca gacagatgga gggagaagga gtctttaaaa gcattcagca cctaacaatc 9480
acagaagaaa tcgctgtgca aaactgggta gcaagagtgg ggcgcgaaag gttatcaaga 9540
atggccatca gtggagatga ttgtgttgtg aaaccttttag atgacagggtt cgcaagcgct 9600
ttaacagctc taaatgacat gggaaagatt aggaaagaca tacaacaatg ggaaccttca 9660
agaggatgga atgattggac acaagtgcc ttctgttcac accatttcca tgagttaatc 9720
atgaaagacg gtgcgtact cgttggtcca tgtagaaacc aagatgaact gattggcaga 9780
gcccgaaatc cccaaggagc aggggtggtc ttgcgggaga cggcctgttt ggggaagtct 9840
tacgccccaa tgtggagctt gatgtacttc cacagacgcg acctcaggct ggcggcgaat 9900
gctatttgct cggcagtacc atcacattgg gttccaacaa gtcgaacaac ctggtccata 9960
catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa caggggtgtg 10020
attcaagaaa acccatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca 10080
tacttgggga aaagagaaga ccaatggtgc ggctcattga ttgggttaac aagcagggcc 10140
acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa 10200
gaatacacag attaccatgcc atccatgaaa agattcagaa gagaagagga agaagcagga 10260
gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320
catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380
ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440
tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500
ggttagagga gaccctccc ttacaaatcg cagcaacaat gggggcccaa ggcgagatga 10560
agctgtagtc tcgctggaag gactagaggt tagaggagac cccccgaaa caaaaaacag 10620
catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca 10680
gaacgccaga aaatggaatg gtgctgttga atcaacagggt tct 10723

```

```

<210> 5
<211> 3391
<212> PRT
<213> Trình tự nhân tạo

<220>

```

<223> Virut dengue typ huyết thanh 2, BVS

<400> 5

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

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
85 90 95

Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
100 105 110

Met Ala Phe His Leu Thr Thr Arg Asn Gly Glu Pro His Met Ile Val
115 120 125

Ser Arg Gln Glu Lys Gly Lys Ser Leu Leu Phe Lys Thr Glu Val Gly
130 135 140

Val Asn Met Cys Thr Leu Met Ala Met Asp Leu Gly Glu Leu Cys Glu
145 150 155 160

Asp Thr Ile Thr Tyr Glu Cys Pro Leu Leu Arg Gln Asn Glu Pro Glu
165 170 175

Asp Ile Asp Cys Trp Cys Asn Ser Thr Ser Thr Trp Val Thr Tyr Gly
180 185 190

Thr Cys Thr Thr Met Gly Glu His Arg Arg Glu Lys Arg Ser Val Ala
195 200 205

Leu Val Pro His Val Gly Met Gly Leu Glu Thr Arg Thr Glu Thr Trp

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

Ser Ile Thr Glu Ala Glu Leu Thr Gly Tyr Gly Thr Val Thr Met Glu
450 455 460

Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Val Leu Leu Gln
465 470 475 480

Met Glu Asn Lys Ala Trp Leu Val His Arg Gln Trp Phe Leu Asp Leu
485 490 495

Pro Leu Pro Trp Leu Pro Gly Ala Asp Thr Gln Gly Ser Asn Trp Ile
500 505 510

Gln Lys Glu Thr Leu Val Thr Phe Lys Asn Pro His Ala Lys Lys Gln
515 520 525

Asp Val Val Val Leu Gly Ser Gln Glu Gly Ala Met His Thr Ala Leu
530 535 540

Thr Gly Ala Thr Glu Ile Gln Met Ser Ser Gly Asn Leu Leu Phe Thr
545 550 555 560

Gly His Leu Lys Cys Arg Leu Arg Met Asp Lys Leu Gln Leu Lys Gly
565 570 575

Met Ser Tyr Ser Met Cys Thr Gly Lys Phe Lys Val Val Lys Glu Ile
580 585 590

Ala Glu Thr Gln His Gly Thr Ile Val Ile Arg Val Gln Tyr Glu Gly
595 600 605

Asp Gly Ser Pro Cys Lys Ile Pro Phe Glu Ile Met Asp Leu Glu Lys
610 615 620

Arg His Val Leu Gly Arg Leu Ile Thr Val Asn Pro Ile Val Thr Glu
625 630 635 640

Lys Asp Ser Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Asp Ser
645 650 655

Tyr Ile Ile Ile Gly Val Glu Pro Gly Gln Leu Lys Leu Asn Trp Phe
660 665 670

Lys Lys Gly Ser Ser Ile Gly Gln Met Phe Glu Thr Thr Met Arg Gly
 675 680 685

Ala Lys Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser
 690 695 700

Leu Gly Gly Val Phe Thr Ser Ile Gly Lys Ala Leu His Gln Val Phe
 705 710 715 720

Gly Ala Ile Tyr Gly Ala Ala Phe Ser Gly Val Ser Trp Thr Met Lys
 725 730 735

Ile Leu Ile Gly Val Ile Ile Thr Trp Ile Gly Met Asn Ser Arg Ser
 740 745 750

Thr Ser Leu Ser Val Thr Leu Val Leu Val Gly Ile Val Thr Leu Tyr
 755 760 765

Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
 770 775 780

Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
 785 790 795 800

His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
 805 810 815

Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
 820 825 830

Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
 835 840 845

Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
 850 855 860

Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
 865 870 875 880

Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
 885 890 895

Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
 900 905 910

Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
 915 920 925

Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
 930 935 940

Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
 945 950 955 960

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
 965 970 975

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
 980 985 990

Ile Glu Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser
 995 1000 1005

Asn Gly Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala
 1010 1015 1020

Gly Pro Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln
 1025 1030 1035

Ile Thr Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp
 1040 1045 1050

Phe Cys Asp Gly Thr Thr Val Val Val Thr Glu Asp Cys Gly Asn
 1055 1060 1065

Arg Gly Pro Ser Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile
 1070 1075 1080

Thr Glu Trp Cys Cys Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr
 1085 1090 1095

Arg Gly Glu Asp Gly Cys Trp Tyr Gly Met Glu Ile Arg Pro Leu
 1100 1105 1110

Lys Glu Lys Glu Glu Asn Leu Val Asn Ser Leu Val Thr Ala Gly

1115		1120		1125
His Gly Gln Val Asp Asn Phe Ser Leu Gly Val Leu Gly Met Ala				
1130		1135		1140
Leu Phe Leu Glu Glu Met Leu Arg Thr Arg Val Gly Thr Lys His				
1145		1150		1155
Ala Ile Leu Leu Val Ala Val Ser Phe Val Thr Leu Ile Thr Gly				
1160		1165		1170
Asn Met Ser Phe Arg Asp Leu Gly Arg Val Met Val Met Val Gly				
1175		1180		1185
Ala Thr Met Thr Asp Asp Ile Gly Met Gly Val Thr Tyr Leu Ala				
1190		1195		1200
Leu Leu Ala Ala Phe Lys Val Arg Pro Thr Phe Ala Ala Gly Leu				
1205		1210		1215
Leu Leu Arg Lys Leu Thr Ser Lys Glu Leu Met Met Thr Thr Ile				
1220		1225		1230
Gly Ile Val Leu Leu Ser Gln Ser Thr Ile Pro Glu Thr Ile Leu				
1235		1240		1245
Glu Leu Thr Asp Ala Leu Ala Leu Gly Met Met Val Leu Lys Met				
1250		1255		1260
Val Arg Asn Met Glu Lys Tyr Gln Leu Ala Val Thr Ile Met Ala				
1265		1270		1275
Ile Leu Cys Val Pro Asn Ala Val Ile Leu Gln Asn Ala Trp Lys				
1280		1285		1290
Val Ser Cys Thr Ile Leu Ala Val Val Ser Val Ser Pro Leu Phe				
1295		1300		1305
Leu Thr Ser Ser Gln Gln Lys Thr Asp Trp Ile Pro Leu Ala Leu				
1310		1315		1320
Thr Ile Lys Gly Leu Asn Pro Thr Ala Ile Phe Leu Thr Thr Leu				
1325		1330		1335

Ser	Arg	Thr	Ser	Lys	Lys	Arg	Ser	Trp	Pro	Leu	Asn	Glu	Ala	Ile
1340						1345					1350			
Met	Ala	Val	Gly	Met	Val	Ser	Ile	Leu	Ala	Ser	Ser	Leu	Leu	Lys
1355						1360					1365			
Asn	Asp	Ile	Pro	Met	Thr	Gly	Pro	Leu	Val	Ala	Gly	Gly	Leu	Leu
1370						1375					1380			
Thr	Val	Cys	Tyr	Val	Leu	Thr	Gly	Arg	Ser	Ala	Asp	Leu	Glu	Leu
1385						1390					1395			
Glu	Arg	Ala	Ala	Asp	Val	Lys	Trp	Glu	Asp	Gln	Ala	Glu	Ile	Ser
1400						1405					1410			
Gly	Ser	Ser	Pro	Ile	Leu	Ser	Ile	Thr	Ile	Ser	Glu	Asp	Gly	Ser
1415						1420					1425			
Met	Ser	Ile	Lys	Asn	Glu	Glu	Glu	Glu	Gln	Thr	Leu	Thr	Ile	Leu
1430						1435					1440			
Ile	Arg	Thr	Gly	Leu	Leu	Val	Ile	Ser	Gly	Leu	Phe	Pro	Val	Ser
1445						1450					1455			
Ile	Pro	Ile	Thr	Ala	Ala	Ala	Trp	Tyr	Leu	Trp	Glu	Val	Lys	Lys
1460						1465					1470			
Gln	Arg	Ala	Gly	Val	Leu	Trp	Asp	Val	Pro	Ser	Pro	Pro	Pro	Met
1475						1480					1485			
Gly	Lys	Ala	Glu	Leu	Glu	Asp	Gly	Ala	Tyr	Arg	Ile	Lys	Gln	Lys
1490						1495					1500			
Gly	Ile	Leu	Gly	Tyr	Ser	Gln	Ile	Gly	Ala	Gly	Val	Tyr	Lys	Glu
1505						1510					1515			
Gly	Thr	Phe	His	Thr	Met	Trp	His	Val	Thr	Arg	Gly	Ala	Val	Leu
1520						1525					1530			
Met	His	Lys	Gly	Lys	Arg	Ile	Glu	Pro	Ser	Trp	Ala	Asp	Val	Lys
1535						1540					1545			

Lys	Asp	Leu	Ile	Ser	Tyr	Gly	Gly	Gly	Trp	Lys	Leu	Glu	Gly	Glu
1550						1555					1560			
Trp	Lys	Glu	Gly	Glu	Glu	Val	Gln	Val	Leu	Ala	Leu	Glu	Pro	Gly
1565						1570					1575			
Lys	Asn	Pro	Arg	Ala	Val	Gln	Thr	Lys	Pro	Gly	Leu	Phe	Lys	Thr
1580						1585					1590			
Asn	Ala	Gly	Thr	Ile	Gly	Ala	Val	Ser	Leu	Asp	Phe	Ser	Pro	Gly
1595						1600					1605			
Thr	Ser	Gly	Ser	Pro	Ile	Ile	Asp	Lys	Lys	Gly	Lys	Val	Val	Gly
1610						1615					1620			
Leu	Tyr	Gly	Asn	Gly	Val	Val	Thr	Arg	Ser	Gly	Ala	Tyr	Val	Ser
1625						1630					1635			
Ala	Ile	Ala	Gln	Thr	Glu	Lys	Ser	Ile	Glu	Asp	Asn	Pro	Glu	Ile
1640						1645					1650			
Glu	Asp	Asp	Ile	Phe	Arg	Lys	Arg	Arg	Leu	Thr	Ile	Met	Asp	Leu
1655						1660					1665			
His	Pro	Gly	Ala	Gly	Lys	Thr	Lys	Arg	Tyr	Leu	Pro	Ala	Ile	Val
1670						1675					1680			
Arg	Glu	Ala	Ile	Lys	Arg	Gly	Leu	Arg	Thr	Leu	Ile	Leu	Ala	Pro
1685						1690					1695			
Thr	Arg	Val	Val	Ala	Ala	Glu	Met	Glu	Glu	Ala	Leu	Arg	Gly	Leu
1700						1705					1710			
Pro	Ile	Arg	Tyr	Gln	Thr	Pro	Ala	Ile	Arg	Ala	Val	His	Thr	Gly
1715						1720					1725			
Arg	Glu	Ile	Val	Asp	Leu	Met	Cys	His	Ala	Thr	Phe	Thr	Met	Arg
1730						1735					1740			
Leu	Leu	Ser	Pro	Val	Arg	Val	Pro	Asn	Tyr	Asn	Leu	Ile	Ile	Met
1745						1750					1755			

Asp	Glu	Ala	His	Phe	Thr	Asp	Pro	Ala	Ser	Ile	Ala	Ala	Arg	Gly
1760						1765					1770			
Tyr	Ile	Ser	Thr	Arg	Val	Glu	Met	Gly	Glu	Ala	Ala	Gly	Ile	Phe
1775						1780					1785			
Met	Thr	Ala	Thr	Pro	Pro	Gly	Ser	Arg	Asp	Pro	Phe	Pro	Gln	Ser
1790						1795					1800			
Asn	Ala	Pro	Ile	Ile	Asp	Glu	Glu	Arg	Glu	Ile	Pro	Glu	Arg	Ser
1805						1810					1815			
Trp	Asn	Ser	Gly	His	Glu	Trp	Val	Thr	Asp	Phe	Lys	Gly	Lys	Thr
1820						1825					1830			
Val	Trp	Phe	Val	Pro	Ser	Ile	Lys	Ala	Gly	Asn	Asp	Ile	Ala	Ala
1835						1840					1845			
Cys	Leu	Arg	Lys	Asn	Gly	Lys	Lys	Val	Ile	Gln	Leu	Ser	Arg	Lys
1850						1855					1860			
Thr	Phe	Asp	Ser	Glu	Tyr	Val	Lys	Thr	Arg	Thr	Asn	Asp	Trp	Asp
1865						1870					1875			
Phe	Val	Val	Thr	Thr	Asp	Ile	Ser	Glu	Met	Gly	Ala	Asn	Phe	Lys
1880						1885					1890			
Ala	Glu	Arg	Val	Ile	Asp	Pro	Arg	Arg	Cys	Met	Lys	Pro	Val	Ile
1895						1900					1905			
Leu	Thr	Asp	Gly	Glu	Glu	Arg	Val	Ile	Leu	Ala	Gly	Pro	Met	Pro
1910						1915					1920			
Val	Thr	His	Ser	Ser	Ala	Ala	Gln	Arg	Arg	Gly	Arg	Ile	Gly	Arg
1925						1930					1935			
Asn	Pro	Lys	Asn	Glu	Asn	Asp	Gln	Tyr	Ile	Tyr	Met	Gly	Glu	Pro
1940						1945					1950			
Leu	Glu	Asn	Asp	Glu	Asp	Cys	Ala	His	Trp	Lys	Glu	Ala	Lys	Met
1955						1960					1965			
Leu	Leu	Asp	Asn	Ile	Asn	Thr	Pro	Glu	Gly	Ile	Ile	Pro	Ser	Met

1970		1975		1980
Phe Glu 1985	Pro Glu Arg Glu Lys 1990	Val Asp Ala Ile Asp 1995	Gly Glu Tyr	
Arg Leu 2000	Arg Gly Glu Ala Arg 2005	Lys Thr Phe Val Asp 2010	Leu Met Arg	
Arg Gly 2015	Asp Leu Pro Val Trp 2020	Leu Ala Tyr Arg Val 2025	Ala Ala Glu	
Gly Ile 2030	Asn Tyr Ala Asp Arg 2035	Arg Trp Cys Phe Asp 2040	Gly Val Lys	
Asn Asn 2045	Gln Ile Leu Glu Glu 2050	Asn Val Glu Val Glu 2055	Ile Trp Thr	
Lys Glu 2060	Gly Glu Arg Lys Lys 2065	Leu Lys Pro Arg Trp 2070	Leu Asp Ala	
Arg Ile 2075	Tyr Ser Asp Pro Leu 2080	Ala Leu Lys Glu Phe 2085	Lys Glu Phe	
Ala Ala 2090	Gly Arg Lys Ser Leu 2095	Thr Leu Asn Leu Ile 2100	Thr Glu Met	
Gly Arg 2105	Leu Pro Thr Phe Met 2110	Thr Gln Lys Ala Arg 2115	Asp Ala Leu	
Asp Asn 2120	Leu Ala Val Leu His 2125	Thr Ala Glu Pro Gly 2130	Gly Arg Ala	
Tyr Asn 2135	His Ala Leu Ser Glu 2140	Leu Pro Glu Thr Leu 2145	Glu Thr Leu	
Leu Leu 2150	Leu Thr Leu Leu Ala 2155	Thr Val Thr Gly Gly 2160	Ile Phe Leu	
Phe Leu 2165	Met Ser Ala Arg Gly 2170	Ile Gly Lys Met Thr 2175	Leu Gly Met	
Cys Cys 2180	Ile Ile Thr Ala Ser 2185	Ile Leu Leu Trp Tyr 2190	Ala Gln Ile	

Gln	Pro	His	Trp	Ile	Ala	Ala	Ser	Ile	Ile	Leu	Glu	Phe	Phe	Leu
2195						2200					2205			
Ile	Val	Leu	Leu	Ile	Pro	Glu	Pro	Glu	Lys	Gln	Arg	Thr	Pro	Gln
2210						2215					2220			
Asp	Asn	Gln	Leu	Thr	Tyr	Val	Val	Ile	Ala	Ile	Leu	Thr	Val	Val
2225						2230					2235			
Ala	Ala	Thr	Met	Ala	Asn	Glu	Met	Gly	Phe	Leu	Glu	Lys	Thr	Lys
2240						2245					2250			
Lys	Asp	Leu	Gly	Leu	Gly	Ser	Ile	Ala	Thr	Gln	Gln	Pro	Glu	Ser
2255						2260					2265			
Asn	Ile	Leu	Asp	Ile	Asp	Leu	Arg	Pro	Ala	Ser	Ala	Trp	Thr	Leu
2270						2275					2280			
Tyr	Ala	Val	Ala	Thr	Thr	Phe	Val	Thr	Pro	Met	Leu	Arg	His	Ser
2285						2290					2295			
Ile	Glu	Asn	Ser	Ser	Val	Asn	Val	Ser	Leu	Thr	Ala	Ile	Ala	Asn
2300						2305					2310			
Gln	Ala	Thr	Val	Leu	Met	Gly	Leu	Gly	Lys	Gly	Trp	Pro	Leu	Ser
2315						2320					2325			
Lys	Met	Asp	Ile	Gly	Val	Pro	Leu	Leu	Ala	Ile	Gly	Cys	Tyr	Ser
2330						2335					2340			
Gln	Val	Asn	Pro	Ile	Thr	Leu	Thr	Ala	Ala	Phe	Phe	Leu	Leu	Val
2345						2350					2355			
Ala	His	Tyr	Ala	Ile	Ile	Gly	Pro	Gly	Leu	Gln	Ala	Lys	Ala	Thr
2360						2365					2370			
Arg	Glu	Ala	Gln	Lys	Arg	Ala	Ala	Ala	Gly	Ile	Met	Lys	Asn	Pro
2375						2380					2385			
Thr	Val	Asp	Gly	Ile	Thr	Val	Ile	Asp	Leu	Asp	Pro	Ile	Pro	Tyr
2390						2395					2400			

Asp	Pro	Lys	Phe	Glu	Lys	Gln	Leu	Gly	Gln	Val	Met	Leu	Leu	Val
2405						2410					2415			
Leu	Cys	Val	Thr	Gln	Val	Leu	Met	Met	Arg	Thr	Thr	Trp	Ala	Leu
2420						2425					2430			
Cys	Glu	Ala	Leu	Thr	Leu	Ala	Thr	Gly	Pro	Ile	Ser	Thr	Leu	Trp
2435						2440					2445			
Glu	Gly	Asn	Pro	Gly	Arg	Phe	Trp	Asn	Thr	Thr	Ile	Ala	Val	Ser
2450						2455					2460			
Met	Ala	Asn	Ile	Phe	Arg	Gly	Ser	Tyr	Leu	Ala	Gly	Ala	Gly	Leu
2465						2470					2475			
Leu	Phe	Ser	Ile	Met	Lys	Asn	Thr	Thr	Asn	Thr	Arg	Arg	Gly	Thr
2480						2485					2490			
Gly	Asn	Ile	Gly	Glu	Thr	Leu	Gly	Glu	Lys	Trp	Lys	Ser	Arg	Leu
2495						2500					2505			
Asn	Ala	Leu	Gly	Lys	Ser	Glu	Phe	Gln	Ile	Tyr	Lys	Lys	Ser	Gly
2510						2515					2520			
Ile	Gln	Glu	Val	Asp	Arg	Thr	Leu	Ala	Lys	Glu	Gly	Ile	Lys	Arg
2525						2530					2535			
Gly	Glu	Thr	Asp	His	His	Ala	Val	Ser	Arg	Gly	Ser	Ala	Lys	Leu
2540						2545					2550			
Arg	Trp	Phe	Val	Glu	Arg	Asn	Met	Val	Thr	Pro	Glu	Gly	Lys	Val
2555						2560					2565			
Val	Asp	Leu	Gly	Cys	Gly	Arg	Gly	Gly	Trp	Ser	Tyr	Tyr	Cys	Gly
2570						2575					2580			
Gly	Leu	Lys	Asn	Val	Arg	Glu	Val	Lys	Gly	Leu	Thr	Lys	Gly	Gly
2585						2590					2595			
Pro	Gly	His	Glu	Glu	Pro	Ile	Pro	Met	Ser	Thr	Tyr	Gly	Trp	Asn
2600						2605					2610			

Leu Val	Arg Leu Gln Ser Gly	Val Asp Val Phe Phe	Ile Pro Pro
2615	2620	2625	
Glu Lys	Cys Asp Thr Leu Leu	Cys Asp Ile Gly Glu	Ser Ser Pro
2630	2635	2640	
Asn Pro	Thr Val Glu Ala Gly	Arg Thr Leu Arg Val	Leu Asn Leu
2645	2650	2655	
Val Glu	Asn Trp Leu Asn Asn	Asn Thr Gln Phe Cys	Ile Lys Val
2660	2665	2670	
Leu Asn	Pro Tyr Met Pro Ser	Val Ile Glu Lys Met	Glu Ala Leu
2675	2680	2685	
Gln Arg	Lys Tyr Gly Gly Ala	Leu Val Arg Asn Pro	Leu Ser Arg
2690	2695	2700	
Asn Ser	Thr His Glu Met Tyr	Trp Val Ser Asn Ala	Ser Gly Asn
2705	2710	2715	
Ile Val	Ser Ser Val Asn Met	Ile Ser Arg Met Leu	Ile Asn Arg
2720	2725	2730	
Phe Thr	Met Arg Tyr Lys Lys	Ala Thr Tyr Glu Pro	Asp Val Asp
2735	2740	2745	
Leu Gly	Ser Gly Thr Arg Asn	Ile Gly Ile Glu Ser	Glu Ile Pro
2750	2755	2760	
Asn Leu	Asp Ile Ile Gly Lys	Arg Ile Glu Lys Ile	Lys Gln Glu
2765	2770	2775	
His Glu	Thr Ser Trp His Tyr	Asp Gln Asp His Pro	Tyr Lys Thr
2780	2785	2790	
Trp Ala	Tyr His Gly Ser Tyr	Glu Thr Lys Gln Thr	Gly Ser Ala
2795	2800	2805	
Ser Ser	Met Val Asn Gly Val	Val Arg Leu Leu Thr	Lys Pro Trp
2810	2815	2820	
Asp Val	Val Pro Met Val Thr	Gln Met Ala Met Thr	Asp Thr Thr

2825		2830		2835
Pro Phe Gly Gln Gln Arg Val Phe Lys Glu Lys Val Asp Thr Arg				
2840		2845		2850
Thr Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu Met Lys Ile Thr				
2855		2860		2865
Ala Glu Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys Thr Pro Arg				
2870		2875		2880
Met Cys Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser Asn Ala				
2885		2890		2895
Ala Leu Gly Ala Val Phe Thr Asp Glu Asn Lys Trp Lys Ser Ala				
2900		2905		2910
Arg Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys				
2915		2920		2925
Glu Arg Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr				
2930		2935		2940
Asn Met Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys				
2945		2950		2955
Ala Lys Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg				
2960		2965		2970
Phe Leu Glu Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp				
2975		2980		2985
Phe Ser Arg Glu Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu				
2990		2995		3000
His Lys Leu Gly Tyr Ile Leu Arg Asp Val Ser Lys Lys Glu Gly				
3005		3010		3015
Gly Ala Met Tyr Ala Asp Asp Thr Ala Gly Trp Asp Thr Arg Ile				
3020		3025		3030
Thr Leu Glu Asp Leu Lys Asn Glu Glu Met Val Thr Asn His Met				
3035		3040		3045

Glu Gly 3050	Glu His Lys Lys 3055	Leu Ala Glu Ala Ile Phe 3060	Lys Leu Thr
Tyr Gln 3065	Asn Lys Val Val Arg 3070	Val Gln Arg Pro Thr 3075	Pro Arg Gly
Thr Val 3080	Met Asp Ile Ile Ser 3085	Arg Arg Asp Gln Arg 3090	Gly Ser Gly
Gln Val 3095	Gly Thr Tyr Gly Leu 3100	Asn Thr Phe Thr Asn 3105	Met Glu Ala
Gln Leu 3110	Ile Arg Gln Met Glu 3115	Gly Glu Gly Val Phe 3120	Lys Ser Ile
Gln His 3125	Leu Thr Ile Thr Glu 3130	Glu Ile Ala Val Gln 3135	Asn Trp Leu
Ala Arg 3140	Val Gly Arg Glu Arg 3145	Leu Ser Arg Met Ala 3150	Ile Ser Gly
Asp Asp 3155	Cys Val Val Lys Pro 3160	Leu Asp Asp Arg Phe 3165	Ala Ser Ala
Leu Thr 3170	Ala Leu Asn Asp Met 3175	Gly Lys Ile Arg Lys 3180	Asp Ile Gln
Gln Trp 3185	Glu Pro Ser Arg Gly 3190	Trp Asn Asp Trp Thr 3195	Gln Val Pro
Phe Cys 3200	Ser His His Phe His 3205	Glu Leu Ile Met Lys 3210	Asp Gly Arg
Val Leu 3215	Val Val Pro Cys Arg 3220	Asn Gln Asp Glu Leu 3225	Ile Gly Arg
Ala Arg 3230	Ile Ser Gln Gly Ala 3235	Gly Trp Ser Leu Arg 3240	Glu Thr Ala
Cys Leu 3245	Gly Lys Ser Tyr Ala 3250	Gln Met Trp Ser Leu 3255	Met Tyr Phe

His Arg Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile Cys Ser Ala
3260 3265 3270

Val Pro Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp Ser Ile
3275 3280 3285

His Ala Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr Val
3290 3295 3300

Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys
3305 3310 3315

Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg
3320 3325 3330

Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala
3335 3340 3345

Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser
3350 3355 3360

Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys
3365 3370 3375

Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
3380 3385 3390

<210> 6

<211> 3391

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Virut dengue typ huyết thanh 2, MVS

<400> 6

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

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
 50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
 65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
 85 90 95

Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
 100 105 110

Met Ala Phe His Leu Thr Thr Arg Asn Gly Glu Pro His Met Ile Val
 115 120 125

Ser Arg Gln Glu Lys Gly Lys Ser Leu Leu Phe Lys Thr Glu Val Gly
 130 135 140

Val Asn Met Cys Thr Leu Met Ala Met Asp Leu Gly Glu Leu Cys Glu
 145 150 155 160

Asp Thr Ile Thr Tyr Glu Cys Pro Leu Leu Arg Gln Asn Glu Pro Glu
 165 170 175

Asp Ile Asp Cys Trp Cys Asn Ser Thr Ser Thr Trp Val Thr Tyr Gly
 180 185 190

Thr Cys Thr Thr Met Gly Glu His Arg Arg Glu Lys Arg Ser Val Ala
 195 200 205

Leu Val Pro His Val Gly Met Gly Leu Glu Thr Arg Thr Glu Thr Trp
 210 215 220

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

Ile Leu Arg His Pro Gly Phe Thr Met Met Ala Ala Ile Leu Ala Tyr
 245 250 255

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

Thr Ala Val Thr Pro Ser Met Thr Met Arg Cys Ile Gly Met Ser Asn
 275 280 285

Arg Asp Phe Val Glu Gly Val Ser Gly Gly Ser Trp Val Asp Ile Val
 290 295 300

Leu Glu His Gly Ser Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr
 305 310 315 320

Leu Asp Phe Glu Leu Ile Lys Thr Glu Ala Lys Gln Pro Ala Thr Leu
 325 330 335

Arg Lys Tyr Cys Ile Glu Ala Lys Leu Thr Asn Thr Thr Thr Glu Ser
 340 345 350

Arg Cys Pro Thr Gln Gly Glu Pro Ser Leu Asn Glu Glu Gln Asp Lys
 355 360 365

Arg Phe Val Cys Lys His Ser Met Val Asp Arg Gly Trp Gly Asn Gly
 370 375 380

Cys Gly Leu Phe Gly Lys Gly Gly Ile Val Thr Cys Ala Met Phe Arg
 385 390 395 400

Cys Lys Lys Asn Met Glu Gly Lys Val Val Gln Pro Glu Asn Leu Glu
 405 410 415

Tyr Thr Ile Val Ile Thr Pro His Ser Gly Glu Glu His Ala Val Gly
 420 425 430

Asn Asp Thr Gly Lys His Gly Lys Glu Ile Lys Ile Thr Pro Gln Ser
 435 440 445

Ser Ile Thr Glu Ala Glu Leu Thr Gly Tyr Gly Thr Val Thr Met Glu
 450 455 460

Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Val Leu Leu Gln
 465 470 475 480

Met Glu Asn Lys Ala Trp Leu Val His Arg Gln Trp Phe Leu Asp Leu
 485 490 495

Pro Leu Pro Trp Leu Pro Gly Ala Asp Thr Gln Gly Ser Asn Trp Ile
 500 505 510

Gln Lys Glu Thr Leu Val Thr Phe Lys Asn Pro His Ala Lys Lys Gln
 515 520 525

Asp Val Val Val Leu Gly Ser Gln Glu Gly Ala Met His Thr Ala Leu
 530 535 540

Thr Gly Ala Thr Glu Ile Gln Met Ser Ser Gly Asn Leu Leu Phe Thr
 545 550 555 560

Gly His Leu Lys Cys Arg Leu Arg Met Asp Lys Leu Gln Leu Lys Gly
 565 570 575

Met Ser Tyr Ser Met Cys Thr Gly Lys Phe Lys Val Val Lys Glu Ile
 580 585 590

Ala Glu Thr Gln His Gly Thr Ile Val Ile Arg Val Gln Tyr Glu Gly
 595 600 605

Asp Gly Ser Pro Cys Lys Ile Pro Phe Glu Ile Met Asp Leu Glu Lys
 610 615 620

Arg His Val Leu Gly Arg Leu Ile Thr Val Asn Pro Ile Val Thr Glu
 625 630 635 640

Lys Asp Ser Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Asp Ser
 645 650 655

Tyr Ile Ile Ile Gly Val Glu Pro Gly Gln Leu Lys Leu Asn Trp Phe
 660 665 670

Lys Lys Gly Ser Ser Ile Gly Gln Met Phe Glu Thr Thr Met Arg Gly
 675 680 685

Ala Lys Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser
 690 695 700

Leu Gly Gly Val Phe Thr Ser Ile Gly Lys Ala Leu His Gln Val Phe
 705 710 715 720

Gly Ala Ile Tyr Gly Ala Ala Phe Ser Gly Val Ser Trp Thr Met Lys

				725						730						735			
Ile	Leu	Ile	Gly	Val	Ile	Ile	Thr	Trp	Ile	Gly	Met	Asn	Ser	Arg	Ser				
			740					745					750						
Thr	Ser	Leu	Ser	Val	Thr	Leu	Val	Leu	Val	Gly	Ile	Val	Thr	Leu	Tyr				
		755					760					765							
Leu	Gly	Val	Met	Val	Gln	Ala	Asp	Ser	Gly	Cys	Val	Val	Ser	Trp	Lys				
	770					775					780								
Asn	Lys	Glu	Leu	Lys	Cys	Gly	Ser	Gly	Ile	Phe	Ile	Thr	Asp	Asn	Val				
785					790					795					800				
His	Thr	Trp	Thr	Glu	Gln	Tyr	Lys	Phe	Gln	Pro	Glu	Ser	Pro	Ser	Lys				
				805					810					815					
Leu	Ala	Ser	Ala	Ile	Gln	Lys	Ala	His	Glu	Glu	Asp	Ile	Cys	Gly	Ile				
			820					825					830						
Arg	Ser	Val	Thr	Arg	Leu	Glu	Asn	Leu	Met	Trp	Lys	Gln	Ile	Thr	Pro				
		835					840					845							
Glu	Leu	Asn	His	Ile	Leu	Ser	Glu	Asn	Glu	Val	Lys	Leu	Thr	Ile	Met				
	850					855					860								
Thr	Gly	Asp	Ile	Lys	Gly	Ile	Met	Gln	Ala	Gly	Lys	Arg	Ser	Leu	Arg				
865					870					875				880					
Pro	Gln	Pro	Thr	Glu	Leu	Lys	Tyr	Ser	Trp	Lys	Thr	Trp	Gly	Lys	Ala				
				885					890					895					
Lys	Met	Leu	Ser	Thr	Glu	Ser	His	Asn	Gln	Thr	Phe	Leu	Ile	Asp	Gly				
			900					905					910						
Pro	Glu	Thr	Ala	Glu	Cys	Pro	Asn	Thr	Asn	Arg	Ala	Trp	Asn	Ser	Leu				
		915					920					925							
Glu	Val	Glu	Asp	Tyr	Gly	Phe	Gly	Val	Phe	Thr	Thr	Asn	Ile	Trp	Leu				
	930					935					940								
Lys	Leu	Lys	Glu	Lys	Gln	Asp	Val	Phe	Cys	Asp	Ser	Lys	Leu	Met	Ser				
945					950					955				960					

--

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
 965 970 975

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
 980 985 990

Ile Glu Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser
 995 1000 1005

Asn Gly Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala
 1010 1015 1020

Gly Pro Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln
 1025 1030 1035

Ile Thr Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp
 1040 1045 1050

Phe Cys Asp Gly Thr Thr Val Val Val Thr Glu Asp Cys Gly Asn
 1055 1060 1065

Arg Gly Pro Ser Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile
 1070 1075 1080

Thr Glu Trp Cys Cys Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr
 1085 1090 1095

Arg Gly Glu Asp Gly Cys Trp Tyr Gly Met Glu Ile Arg Pro Leu
 1100 1105 1110

Lys Glu Lys Glu Glu Asn Leu Val Asn Ser Leu Val Thr Ala Gly
 1115 1120 1125

His Gly Gln Val Asp Asn Phe Ser Leu Gly Val Leu Gly Met Ala
 1130 1135 1140

Leu Phe Leu Glu Glu Met Leu Arg Thr Arg Val Gly Thr Lys His
 1145 1150 1155

Ala Ile Leu Leu Val Ala Val Ser Phe Val Thr Leu Ile Thr Gly
 1160 1165 1170

Asn Met	Ser Phe Arg Asp	Leu	Gly Arg Val Met	Val	Met Val Gly
1175		1180		1185	
Ala Thr	Met Thr Asp Asp	Ile	Gly Met Gly Val	Thr	Tyr Leu Ala
1190		1195		1200	
Leu Leu	Ala Ala Phe Lys	Val	Arg Pro Thr Phe	Ala	Ala Gly Leu
1205		1210		1215	
Leu Leu	Arg Lys Leu Thr	Ser	Lys Glu Leu Met	Met	Thr Thr Ile
1220		1225		1230	
Gly Ile	Val Leu Leu Ser	Gln	Ser Thr Ile Pro	Glu	Thr Ile Leu
1235		1240		1245	
Glu Leu	Thr Asp Ala Leu	Ala	Leu Gly Met Met	Val	Leu Lys Met
1250		1255		1260	
Val Arg	Asn Met Glu Lys	Tyr	Gln Leu Ala Val	Thr	Ile Met Ala
1265		1270		1275	
Ile Leu	Cys Val Pro Asn	Ala	Val Ile Leu Gln	Asn	Ala Trp Lys
1280		1285		1290	
Val Ser	Cys Thr Ile Leu	Ala	Val Val Ser Val	Ser	Pro Leu Phe
1295		1300		1305	
Leu Thr	Ser Ser Gln Gln	Lys	Thr Asp Trp Ile	Pro	Leu Ala Leu
1310		1315		1320	
Thr Ile	Lys Gly Leu Asn	Pro	Thr Ala Ile Phe	Leu	Thr Thr Leu
1325		1330		1335	
Ser Arg	Thr Ser Lys Lys	Arg	Ser Trp Pro Leu	Asn	Glu Ala Ile
1340		1345		1350	
Met Ala	Val Gly Met Val	Ser	Ile Leu Ala Ser	Ser	Leu Leu Lys
1355		1360		1365	
Asn Asp	Ile Pro Met Thr	Gly	Pro Leu Val Ala	Gly	Gly Leu Leu
1370		1375		1380	

Thr Val	Cys Tyr Val Leu	Thr Gly Arg Ser Ala Asp	Leu Glu Leu
1385		1390	1395
Glu Arg	Ala Ala Asp Val	Lys Trp Glu Asp Gln Ala	Glu Ile Ser
1400		1405	1410
Gly Ser	Ser Pro Ile Leu	Ser Ile Thr Ile Ser Glu	Asp Gly Ser
1415		1420	1425
Met Ser	Ile Lys Asn Glu	Glu Glu Glu Gln Thr Leu	Thr Ile Leu
1430		1435	1440
Ile Arg	Thr Gly Leu Leu Val	Ile Ser Gly Leu Phe	Pro Val Ser
1445		1450	1455
Ile Pro	Ile Thr Ala Ala Ala	Trp Tyr Leu Trp Glu	Val Lys Lys
1460		1465	1470
Gln Arg	Ala Gly Val Leu Trp	Asp Val Pro Ser Pro	Pro Pro Met
1475		1480	1485
Gly Lys	Ala Glu Leu Glu Asp	Gly Ala Tyr Arg Ile	Lys Gln Lys
1490		1495	1500
Gly Ile	Leu Gly Tyr Ser Gln	Ile Gly Ala Gly Val	Tyr Lys Glu
1505		1510	1515
Gly Thr	Phe His Thr Met Trp	His Val Thr Arg Gly	Ala Val Leu
1520		1525	1530
Met His	Lys Gly Lys Arg Ile	Glu Pro Ser Trp Ala	Asp Val Lys
1535		1540	1545
Lys Asp	Leu Ile Ser Tyr Gly	Gly Gly Trp Lys Leu	Glu Gly Glu
1550		1555	1560
Trp Lys	Glu Gly Glu Glu Val	Gln Val Leu Ala Leu	Glu Pro Gly
1565		1570	1575
Lys Asn	Pro Arg Ala Val Gln	Thr Lys Pro Gly Leu	Phe Lys Thr
1580		1585	1590
Asn Ala	Gly Thr Ile Gly Ala	Val Ser Leu Asp Phe	Ser Pro Gly

1595		1600		1605
Thr Ser Gly Ser Pro Ile Ile Asp Lys Lys Gly Lys Val Val Gly				
1610		1615		1620
Leu Tyr Gly Asn Gly Val Val Thr Arg Ser Gly Ala Tyr Val Ser				
1625		1630		1635
Ala Ile Ala Gln Thr Glu Lys Ser Ile Glu Asp Asn Pro Glu Ile				
1640		1645		1650
Glu Asp Asp Ile Phe Arg Lys Arg Arg Leu Thr Ile Met Asp Leu				
1655		1660		1665
His Pro Gly Ala Gly Lys Thr Lys Arg Tyr Leu Pro Ala Ile Val				
1670		1675		1680
Arg Glu Ala Ile Lys Arg Gly Leu Arg Thr Leu Ile Leu Ala Pro				
1685		1690		1695
Thr Arg Val Val Ala Ala Glu Met Glu Glu Ala Leu Arg Gly Leu				
1700		1705		1710
Pro Ile Arg Tyr Gln Thr Pro Ala Ile Arg Ala Val His Thr Gly				
1715		1720		1725
Arg Glu Ile Val Asp Leu Met Cys His Ala Thr Phe Thr Met Arg				
1730		1735		1740
Leu Leu Ser Pro Val Arg Val Pro Asn Tyr Asn Leu Ile Ile Met				
1745		1750		1755
Asp Glu Ala His Phe Thr Asp Pro Ala Ser Ile Ala Ala Arg Gly				
1760		1765		1770
Tyr Ile Ser Thr Arg Val Glu Met Gly Glu Ala Ala Gly Ile Phe				
1775		1780		1785
Met Thr Ala Thr Pro Pro Gly Ser Arg Asp Pro Phe Pro Gln Ser				
1790		1795		1800
Asn Ala Pro Ile Ile Asp Glu Glu Arg Glu Ile Pro Glu Arg Ser				
1805		1810		1815

Trp	Asn	Ser	Gly	His	Glu	Trp	Val	Thr	Asp	Phe	Lys	Gly	Lys	Thr
1820						1825					1830			
Val	Trp	Phe	Val	Pro	Ser	Ile	Lys	Ala	Gly	Asn	Asp	Ile	Ala	Ala
1835						1840					1845			
Cys	Leu	Arg	Lys	Asn	Gly	Lys	Lys	Val	Ile	Gln	Leu	Ser	Arg	Lys
1850						1855					1860			
Thr	Phe	Asp	Ser	Glu	Tyr	Val	Lys	Thr	Arg	Thr	Asn	Asp	Trp	Asp
1865						1870					1875			
Phe	Val	Val	Thr	Thr	Asp	Ile	Ser	Glu	Met	Gly	Ala	Asn	Phe	Lys
1880						1885					1890			
Ala	Glu	Arg	Val	Ile	Asp	Pro	Arg	Arg	Cys	Met	Lys	Pro	Val	Ile
1895						1900					1905			
Leu	Thr	Asp	Gly	Glu	Glu	Arg	Val	Ile	Leu	Ala	Gly	Pro	Met	Pro
1910						1915					1920			
Val	Thr	His	Ser	Ser	Ala	Ala	Gln	Arg	Arg	Gly	Arg	Ile	Gly	Arg
1925						1930					1935			
Asn	Pro	Lys	Asn	Glu	Asn	Asp	Gln	Tyr	Ile	Tyr	Met	Gly	Glu	Pro
1940						1945					1950			
Leu	Glu	Asn	Asp	Glu	Asp	Cys	Ala	His	Trp	Lys	Glu	Ala	Lys	Met
1955						1960					1965			
Leu	Leu	Asp	Asn	Ile	Asn	Thr	Pro	Glu	Gly	Ile	Ile	Pro	Ser	Met
1970						1975					1980			
Phe	Glu	Pro	Glu	Arg	Glu	Lys	Val	Asp	Ala	Ile	Asp	Gly	Glu	Tyr
1985						1990					1995			
Arg	Leu	Arg	Gly	Glu	Ala	Arg	Lys	Thr	Phe	Val	Asp	Leu	Met	Arg
2000						2005					2010			
Arg	Gly	Asp	Leu	Pro	Val	Trp	Leu	Ala	Tyr	Arg	Val	Ala	Ala	Glu
2015						2020					2025			

Gly Ile	Asn Tyr	Ala Asp	Arg	Arg Trp	Cys Phe	Asp	Gly Val	Lys
2030			2035			2040		
Asn Asn	Gln Ile	Leu Glu	Glu	Asn Val	Glu Val	Glu	Ile Trp	Thr
2045			2050			2055		
Lys Glu	Gly Glu	Arg Lys	Lys	Leu Lys	Pro Arg	Trp	Leu Asp	Ala
2060			2065			2070		
Arg Ile	Tyr Ser	Asp Pro	Leu	Ala Leu	Lys Glu	Phe	Lys Glu	Phe
2075			2080			2085		
Ala Ala	Gly Arg	Lys Ser	Leu	Thr Leu	Asn Leu	Ile	Thr Glu	Met
2090			2095			2100		
Gly Arg	Leu Pro	Thr Phe	Met	Thr Gln	Lys Ala	Arg	Asp Ala	Leu
2105			2110			2115		
Asp Asn	Leu Ala	Val Leu	His	Thr Ala	Glu Ala	Gly	Gly Arg	Ala
2120			2125			2130		
Tyr Asn	His Ala	Leu Ser	Glu	Leu Pro	Glu Thr	Leu	Glu Thr	Leu
2135			2140			2145		
Leu Leu	Leu Thr	Leu Leu	Ala	Thr Val	Thr Gly	Gly	Ile Phe	Leu
2150			2155			2160		
Phe Leu	Met Ser	Ala Arg	Gly	Ile Gly	Lys Met	Thr	Leu Gly	Met
2165			2170			2175		
Cys Cys	Ile Ile	Thr Ala	Ser	Ile Leu	Leu Trp	Tyr	Ala Gln	Ile
2180			2185			2190		
Gln Pro	His Trp	Ile Ala	Ala	Ser Ile	Ile Leu	Glu	Phe Phe	Leu
2195			2200			2205		
Ile Val	Leu Leu	Ile Pro	Glu	Pro Glu	Lys Gln	Arg	Thr Pro	Gln
2210			2215			2220		
Asp Asn	Gln Leu	Thr Tyr	Val	Val Ile	Ala Ile	Leu	Thr Val	Val
2225			2230			2235		

Ala	Ala	Thr	Met	Ala	Asn	Glu	Met	Gly	Phe	Leu	Glu	Lys	Thr	Lys
2240						2245					2250			
Lys	Asp	Leu	Gly	Leu	Gly	Ser	Ile	Ala	Thr	Gln	Gln	Pro	Glu	Ser
2255						2260					2265			
Asn	Ile	Leu	Asp	Ile	Asp	Leu	Arg	Pro	Ala	Ser	Ala	Trp	Thr	Leu
2270						2275					2280			
Tyr	Ala	Val	Ala	Thr	Thr	Phe	Val	Thr	Pro	Met	Leu	Arg	His	Ser
2285						2290					2295			
Ile	Glu	Asn	Ser	Ser	Val	Asn	Val	Ser	Leu	Thr	Ala	Ile	Ala	Asn
2300						2305					2310			
Gln	Ala	Thr	Val	Leu	Met	Gly	Leu	Gly	Lys	Gly	Trp	Pro	Leu	Ser
2315						2320					2325			
Lys	Met	Asp	Ile	Gly	Val	Pro	Leu	Leu	Ala	Ile	Gly	Cys	Tyr	Ser
2330						2335					2340			
Gln	Val	Asn	Pro	Ile	Thr	Leu	Thr	Ala	Ala	Leu	Phe	Leu	Leu	Val
2345						2350					2355			
Ala	His	Tyr	Ala	Ile	Ile	Gly	Pro	Gly	Leu	Gln	Ala	Lys	Ala	Thr
2360						2365					2370			
Arg	Glu	Ala	Gln	Lys	Arg	Ala	Ala	Ala	Gly	Ile	Met	Lys	Asn	Pro
2375						2380					2385			
Thr	Val	Asp	Gly	Ile	Thr	Val	Ile	Asp	Leu	Asp	Pro	Ile	Pro	Tyr
2390						2395					2400			
Asp	Pro	Lys	Phe	Glu	Lys	Gln	Leu	Gly	Gln	Val	Met	Leu	Leu	Val
2405						2410					2415			
Leu	Cys	Val	Thr	Gln	Val	Leu	Met	Met	Arg	Thr	Thr	Trp	Ala	Leu
2420						2425					2430			
Cys	Glu	Ala	Leu	Thr	Leu	Ala	Thr	Gly	Pro	Ile	Ser	Thr	Leu	Trp
2435						2440					2445			
Glu	Gly	Asn	Pro	Gly	Arg	Phe	Trp	Asn	Thr	Thr	Ile	Ala	Val	Ser

2450	2455	2460
Met Ala Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu 2465 2470 2475		
Leu Phe Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr 2480 2485 2490		
Gly Asn Ile Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu 2495 2500 2505		
Asn Ala Leu Gly Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly 2510 2515 2520		
Ile Gln Glu Val Asp Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg 2525 2530 2535		
Gly Glu Thr Asp His His Ala Val Ser Arg Gly Ser Ala Lys Leu 2540 2545 2550		
Arg Trp Phe Val Glu Arg Asn Met Val Thr Pro Glu Gly Lys Val 2555 2560 2565		
Val Asp Leu Gly Cys Gly Arg Gly Gly Trp Ser Tyr Tyr Cys Gly 2570 2575 2580		
Gly Leu Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly Gly 2585 2590 2595		
Pro Gly His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn 2600 2605 2610		
Leu Val Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro 2615 2620 2625		
Glu Lys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser Pro 2630 2635 2640		
Asn Pro Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu 2645 2650 2655		
Val Glu Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys Val 2660 2665 2670		

Leu	Asn	Pro	Tyr	Met	Pro	Ser	Val	Ile	Glu	Lys	Met	Glu	Ala	Leu
2675						2680					2685			
Gln	Arg	Lys	Tyr	Gly	Gly	Ala	Leu	Val	Arg	Asn	Pro	Leu	Ser	Arg
2690						2695					2700			
Asn	Ser	Thr	His	Glu	Met	Tyr	Trp	Val	Ser	Asn	Ala	Ser	Gly	Asn
2705						2710					2715			
Ile	Val	Ser	Ser	Val	Asn	Met	Ile	Ser	Arg	Met	Leu	Ile	Asn	Arg
2720						2725					2730			
Phe	Thr	Met	Arg	Tyr	Lys	Lys	Ala	Thr	Tyr	Glu	Pro	Asp	Val	Asp
2735						2740					2745			
Leu	Gly	Ser	Gly	Thr	Arg	Asn	Ile	Gly	Ile	Glu	Ser	Glu	Ile	Pro
2750						2755					2760			
Asn	Leu	Asp	Ile	Ile	Gly	Lys	Arg	Ile	Glu	Lys	Ile	Lys	Gln	Glu
2765						2770					2775			
His	Glu	Thr	Ser	Trp	His	Tyr	Asp	Gln	Asp	His	Pro	Tyr	Lys	Thr
2780						2785					2790			
Trp	Ala	Tyr	His	Gly	Ser	Tyr	Glu	Thr	Lys	Gln	Thr	Gly	Ser	Ala
2795						2800					2805			
Ser	Ser	Met	Val	Asn	Gly	Val	Val	Arg	Leu	Leu	Thr	Lys	Pro	Trp
2810						2815					2820			
Asp	Val	Val	Pro	Met	Val	Thr	Gln	Met	Ala	Met	Thr	Asp	Thr	Thr
2825						2830					2835			
Pro	Phe	Gly	Gln	Gln	Arg	Val	Phe	Lys	Glu	Lys	Val	Asp	Thr	Arg
2840						2845					2850			
Thr	Gln	Glu	Pro	Lys	Glu	Gly	Thr	Lys	Lys	Leu	Met	Lys	Ile	Thr
2855						2860					2865			
Ala	Glu	Trp	Leu	Trp	Lys	Glu	Leu	Gly	Lys	Lys	Lys	Thr	Pro	Arg
2870						2875					2880			

Met Cys Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser Asn Ala
2885 2890 2895

Ala Leu Gly Ala Val Phe Thr Asp Glu Asn Lys Trp Lys Ser Ala
2900 2905 2910

Arg Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys
2915 2920 2925

Glu Arg Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr
2930 2935 2940

Asn Met Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys
2945 2950 2955

Ala Lys Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg
2960 2965 2970

Phe Leu Glu Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp
2975 2980 2985

Phe Ser Arg Glu Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu
2990 2995 3000

His Lys Leu Gly Tyr Ile Leu Arg Asp Val Ser Lys Lys Glu Gly
3005 3010 3015

Gly Ala Met Tyr Ala Asp Asp Thr Ala Gly Trp Asp Thr Arg Ile
3020 3025 3030

Thr Leu Glu Asp Leu Lys Asn Glu Glu Met Val Thr Asn His Met
3035 3040 3045

Glu Gly Glu His Lys Lys Leu Ala Glu Ala Ile Phe Lys Leu Thr
3050 3055 3060

Tyr Gln Asn Lys Val Val Arg Val Gln Arg Pro Thr Pro Arg Gly
3065 3070 3075

Thr Val Met Asp Ile Ile Ser Arg Arg Asp Gln Arg Gly Ser Gly
3080 3085 3090

Gln Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr Asn Met Glu Ala
 3095 3100 3105
 Gln Leu Ile Arg Gln Met Glu Gly Glu Gly Val Phe Lys Ser Ile
 3110 3115 3120
 Gln His Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn Trp Leu
 3125 3130 3135
 Ala Arg Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser Gly
 3140 3145 3150
 Asp Asp Cys Val Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala
 3155 3160 3165
 Leu Thr Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln
 3170 3175 3180
 Gln Trp Glu Pro Ser Arg Gly Trp Asn Asp Trp Thr Gln Val Pro
 3185 3190 3195
 Phe Cys Ser His His Phe His Glu Leu Ile Met Lys Asp Gly Arg
 3200 3205 3210
 Val Leu Val Val Pro Cys Arg Asn Gln Asp Glu Leu Ile Gly Arg
 3215 3220 3225
 Ala Arg Ile Ser Gln Gly Ala Gly Trp Ser Leu Arg Glu Thr Ala
 3230 3235 3240
 Cys Leu Gly Lys Ser Tyr Ala Gln Met Trp Ser Leu Met Tyr Phe
 3245 3250 3255
 His Arg Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile Cys Ser Ala
 3260 3265 3270
 Val Pro Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp Ser Ile
 3275 3280 3285
 His Ala Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr Val
 3290 3295 3300
 Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys

--

3305 3310 3315
 Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg
 3320 3325 3330

 Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala
 3335 3340 3345

 Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser
 3350 3355 3360

 Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys
 3365 3370 3375

 Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
 3380 3385 3390

<210> 7
 <211> 10717
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Virut dengue typ huyết thanh 3, BVS

<400> 7
 agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60
 gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120
 aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180
 ctgacaaaga gattctcact tggaatgctg cagggacgag gaccattaaa actgttcattg 240
 gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300
 tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360
 ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat cattatgctg 420
 attccaacag tgatggcggt ccatttaacc acgcgtgatg gagagccgcg catgattgtg 480
 gggaagaatg aaagaggaaa atccctactt ttcaagacag cctctggaat caacatgtgc 540
 acactcatag ctatggatct gggagagatg tgtgatgaca cggtcactta caaatgcccc 600
 cacattaccg aagtggagcc tgaagacatt gactgctggg gcaaccttac atcgacatgg 660
 gtgacttatg gaacatgcaa tcaagctgga gagcatagac gcgataagag atcagtggcg 720
 ttagctcccc atgttggcat gggactggac acacgcactc aaacctggat gtcggctgaa 780

ggagcttgga gacaagtcga gaaggtagag acatgggccc ttaggcaccc agggtttacc	840
atactagccc tattttcttgc ccattacata ggcacttcct tgaccagaa agtggttatt	900
tttatactat taatgctggt taccatcc atgacaatga gatgtgtagg agtaggaaac	960
agagattttg tggaaggcct atcgggagct acgtgggttg acgtggtgct cgagcacggt	1020
gggtgtgtga ctaccatggc taagaacaag cccacgctgg acatagagct tcagaagacc	1080
gaggccaccc aactggcgac cctaaggaag ctatgcattg agggaaaaat taccaacata	1140
acaaccgact caagatgtcc caccgaagg gaagcgattt tacctgagga gcaggaccag	1200
aactacgtgt gtaagcatac atacgtggac agaggctggg gaaacggttg tggtttgttt	1260
ggcaagggaa gcttggtgac atgcgcgaaa tttcaatgtt tagaatcaat agagggaaaa	1320
gtggtgcaac atgagaacct caaatacacc gtcacatca cagtgcacac aggagaccaa	1380
caccaggtgg gaaatgaaac gcaggagtc acggctgaga taacacccca ggcacatcaacc	1440
gctgaagcca ttttacctga atatggaacc ctgggctag aatgctcacc acggacaggt	1500
ttggatttca atgaaatgat ctcatgaca atgaagaaca aagcatggat ggtacataga	1560
caatggttct ttgacttacc cctaccatgg acatcaggag cttcagcaga aacaccaact	1620
tggaacagga aagagcttct tgtgacattt aaaaatgcac atgcaaaaaa gcaagaagta	1680
gttgttcttg gatcacaaga gggagcaatg catacagcac tgacaggagc tacagagatc	1740
caaacctcag gaggcacaag tatctttgcg gggcacttaa aatgtagact caagatggac	1800
aaattggaac tcaaggggat gagctatgca atgtgcttga gtagctttgt gttgaagaaa	1860
gaagtctccg aaacgcagca tgggacaata ctcatlaagg ttgagtacaa aggggaagat	1920
gcaccctgca agattccttt ctccacggag gatggacaag gaaaagctct caatggcaga	1980
ctgatcacag ccaatccagt ggtgaccaag aaggaggagc ctgtcaacat tgaggctgaa	2040
cctccttttg gagaaagtaa catagtaatt ggaattggag acaaagccct gaaaatcaac	2100
tggtacaaga agggaagctc gattgggaag atgttcgagg cactgccag aggtgcaagg	2160
cgcattggcca tcttgggaga cacagcctgg gactttggat cagtgggtgg tgttttgaat	2220
tcattagga aaatggtcca ccaaataatt gggagtgtt acacagccct atttggtgga	2280
gtctcctgga tgatgaaaat tggaataggt gtctcttaa cctggatagg gttgaactca	2340
aaaaatactt ctatgtcatt ttcatgcac gcggccggca ttgtgacact gtatttggga	2400
gtcatggtgc aggccgatag tggttgcgtt gtgagctgga aaaacaaaga actgaaatgt	2460

```

ggcagtggga ttttcatcac agacaacgtg cacacatgga cagaacaata caagttccaa 2520
ccagaatccc cttcaaaact agcttcagct atccagaaag cccatgaaga ggacatttgt 2580
ggaatccgct cagtaacaag actggagaat ctgatgtgga aacaaataac accagaattg 2640
aatcacattc tatcagaaaa tgaggtgaag ttaactatta tgacaggaga catcaaagga 2700
atcatgcagg caggaaaacg atctctgcgg cctcagccca ctgagctgaa gtattcatgg 2760
aaaacatggg gcaaagcaaa aatgctctct acagagtctc ataaccagac ctttctcatt 2820
gatggccccg aaacagcaga atgccccaac acaaatagag cttggaattc gttggaagtt 2880
gaagactatg gctttggagt attcaccacc aatatatggc taaaattgaa agaaaaacag 2940
gatgtattct gcgactcaaa actcatgtca gcggccataa aagacaacag agccgtccat 3000
gccgatatgg gttattggat agaaagtgca ctcaatgaca catggaagat agagaaagcc 3060
tctttcattg aagttaaaaa ctgccactgg ccaaaatcac acaccctctg gagcaatgga 3120
gtgctagaaa gtgagatgat aattccaaag aatctcgctg gaccagtgtc tcaacacaac 3180
tatagaccag gctaccatac acaaataaca ggaccatggc atctaggtaa gcttgagatg 3240
gactttgatt tctgtgatgg aacaacagtg gtagtgactg aggactgcgg aaatagagga 3300
ccctctttga gaacaaccac tgccctctgga aaactcataa cagaatgggtg ctgccgatct 3360
tgcacattac caccgctaag atacagaggt gaggatgggt gctggtacgg gatggaaatc 3420
agaccattga aggagaaaga agagaatttg gtcaactcct tggtcacagc tggacatggg 3480
caggtcgaca acttttctact aggagtcttg ggaatggcat tgttcctgga ggaaatgctt 3540
aggacccgag taggaacgaa acatgcaata ctactagttag cagtttcttt tgtgacattg 3600
atcacaggga acatgtcctt tagagacctg ggaagagtga tggttatggg aggcgccact 3660
atgacggatg acataggtat gggcgtgact tatcttgccc tactagcagc cttcaaagtc 3720
agaccaactt ttgcagctgg actactcttg agaaagtga cctccaagga attgatgatg 3780
actactatag gaattgtact cctctcccag agcaccatac cagagaccat tcttgagttg 3840
actgatgcgt tagccttagg catgatgggt ctcaaatgg tgagaaatat ggaaaagtat 3900
caattggcag tgactatcat ggctatcttg tgcggtcccaa acgcagtgat attacaaaac 3960
gcatggaaag tgagttgcac aatattggca gtggtgtccg tttccccact gttcttaaca 4020
tcttcacagc aaaaaacaga ttggatacca ttagcattga cgatcaaagg tctcaatcca 4080
acagctattt ttctaacaac cctctcaaga accagcaaga aaaggagctg gccattaaat 4140
gaggctatca tggcagtcgg gatggtgagc attttagcca gttctctcct aaaaaatgat 4200

```

--

attcccatga caggaccatt agtggctgga gggctcctca ctgtgtgcta cgtgctcact	4260
ggacgatcgg ccgatttgga actggagaga gcagccgatg tcaaatggga agaccaggca	4320
gagatatcag gaagcagtcc aatcctgtca ataacaatat cagaagatgg tagcatgtcg	4380
ataaaaaatg aagaggaaga acaaacactg accatactca ttagaacagg attgctgggtg	4440
atctcaggac tttttcctgt atcaatacca atcacggcag cagcatggta cctgtgggaa	4500
gtgaagaaac aacggggccgg agtattgtgg gatgttcctt cacccccacc catgggaaag	4560
gctgaactgg aagatggagc ctatagaatt aagcaaaaag ggattcttgg atattcccag	4620
atcggagccg gagtttacia agaaggaaca ttccatacaa tgtggcatgt cacacgtggc	4680
gctgttctaa tgcataaagg aaagaggatt gaaccatcat gggcggacgt caagaaagac	4740
ctaatatcat atggaggagg ctggaagtta gaaggagaat ggaaggaagg agaagaagtc	4800
caggtattgg cactggagcc tggaaaaaat ccaagagccg tccaaacgaa acctggtctt	4860
ttcaaaacca acgccggaac aataggtgct gtatctctgg acttttctcc tggaacgtca	4920
ggatctccaa ttatcgacaa aaaaggaaaa gttgtgggtc tttatggtaa tgggtgtgtt	4980
acaaggagtg gagcatatgt gagtgctata gcccagactg aaaaaagcat tgaagacaac	5040
ccagagatcg aagatgacat tttccgaaag agaagactga ccatcatgga cctccacca	5100
ggagcgggaa agacgaagag ataccttccg gccatagtca gagaagctat aaaacggggt	5160
ttgagaacat taatcttggc cccactaga gttgtggcag ctgaaatgga ggaagccctt	5220
agaggacttc caataagata ccagaccca gccatcagag ctgtgcacac cgggcgggag	5280
attgtggacc taatgtgtca tgccacattt accatgaggc tgctatcacc agttagagtg	5340
ccaaactaca acctgattat catggacgaa gccatttca cagaccagc aagtatagca	5400
gctagaggat acatctcaac tcgagtggag atgggtgagg cagctgggat ttttatgaca	5460
gccactcccc cggaagcag agaccattt cctcagagca atgcaccaat catagatgaa	5520
gaaagagaaa tccctgaacg ctcggtggaat tccggacatg aatgggtcac ggattttaaa	5580
gggaagactg tttggttcgt tccaagtata aaagcaggaa atgatatagc agcttgctg	5640
aggaaaaatg gaaagaaagt gatacaactc agtaggaaga cctttgattc tgagtatgtc	5700
aagactagaa ccaatgattg ggacttcgtg gttacaactg acatttcaga aatgggtgcc	5760
aatttcaagg ctgagagggt tatagacccc agacgctgca tgaaaccagt catactaaca	5820
gatggtgaag agcgggtgat tctggcagga cctatgccag tgaccactc tagtgcagca	5880

caaagaagag ggagaatagg aagaaatcca aaaaatgaga atgaccagta catatacatg	5940
ggggaacctc tggaaaatga tgaagactgt gcacactgga aagaagctaa aatgctccta	6000
gataacatca acacgccaga aggaatcatt cctagcatgt tcgaaccaga gcgtgaaaag	6060
gtggatgcca ttgatggcga ataccgcttg agaggagaag caaggaaaac ctttgtagac	6120
ttaatgagaa gaggagacct accagtctgg ttggcctaca gagtggcagc tgaaggcatc	6180
aactacgcag acagaagggtg gtgttttgat ggagtcaaga acaaccaaact cctagaagaa	6240
aacgtggaag ttgaaatctg gacaaaagaa ggggaaagga agaaattgaa acccagatgg	6300
ttggatgcta ggatctattc tgaccactg gcgctaaaag aatttaagga atttgcagcc	6360
ggaagaaagt ctctgacctt gaacctaatc acagaaatgg gtaggctccc aaccttcatg	6420
actcagaagg caagaaacgc actggacaac ttagcagtgc tgcacacggc tgaggcaggt	6480
ggaagggcgt acaaccatgc tctcagtga ctgccggaga cctggagac attgctttta	6540
ctgacacttc tggctacagt cacgggaggg atctttttat tcttgatgag cgcaaggggc	6600
ataggggaaga tgaccctggg aatgtgctgc ataatcacgg ctagcatcct cctatggtac	6660
gcacaaatac agccacactg gatagcagct tcaataatac tggagttttt tctcatagtt	6720
ttgcttattc cagaacctga aaaacagaga acacccaag acaaccaact gacctacgtt	6780
gtcatagcca tcttcacagt ggtggccgca accatggcaa acgagatggg tttcctagaa	6840
aaaacgaaga aagatctcgg attgggaagc attgcaaccc agcaaccga gagcaacatc	6900
ctggacatag atctacgtcc tgcacagca tggacgctgt atgccgtggc cacaacattt	6960
gttacaccaa tggtgagaca tagcattgaa aattcctcag tgaatgtgtc cctaacagct	7020
atagccaacc aagccacagt gttaatgggt ctcgggaaag gatggccatt gtcaaagatg	7080
gacatcggag tcccccttct cgccattgga tgctactcac aagtcaaccc cataactctc	7140
acagcagctc ttttcttatt ggtagcacat tatgccatca tagggccagg actccaagca	7200
aaagcaacca gagaagctca gaaaagagca gcggcgggca tcatgaaaaa cccaactgtc	7260
gatggaataa cagtgattga cctagatcca ataccttatg atccaaagtt tgaaaagcag	7320
ttgggacaag taatgctcct agtcctctgc gtgactcaag tattgatgat gaggactaca	7380
tgggctctgt gtgaggcttt aaccttagct accgggcca tctccacatt gtgggaagga	7440
aatccagga ggttttgaa cactaccatt gcggtgtcaa tggctaacat ttttagaggg	7500
agttacttgg cggagctgg acttctcttt tctattatga agaacacaac caacacaaga	7560
aggggaactg gcaacatagg agagacgctt ggagagaaat ggaaaagccg attgaacgcg	7620

ttgggaaaaa gtgaattcca gatctacaag aaaagtggaa tccaggaagt ggatagaacc	7680
ttagcaaaag aaggcattaa aagaggagaa acggaccatc acgctgtgtc gcgaggctca	7740
gcaaaactga gatggttcgt tgagagaaac atggtcacac cagaagggaa agtagtggac	7800
ctcggttgtg gcagaggagg ctggtcatac tattgtggag gactaaagaa tgtaagagaa	7860
gtcaaaaggcc taacaaaagg aggaccagga cacgaagaac ccatcccat gtcaacatat	7920
gggtggaatc tagtgcgtct tcaaagtgga gttgacgttt tcttcatccc gccagaaaag	7980
tgtgacacat tattgtgtga cataggggag tcatcaccaa atcccacagt ggaagcagga	8040
cgaacactca gagtccctaa cttagtagaa aattgggtga acaacaacac tcaattttgc	8100
ataaagggttc tcaaccata tatgccctca gtcatagaaa aaatggaagc actacaaagg	8160
aaatatggag gagccttagt gaggaatcca ctctcacgaa actccacaca tgagatgtac	8220
tgggtatcca atgcttccgg gaacatagtg tcatcagtga acatgatttc aaggatgttg	8280
atcaacagat ttacaatgag atacaagaaa gccacttacg agccggatgt tgacctcgga	8340
agcggaaccc gtaacatcgg gattgaaagt gagataccaa acctagatat aattgggaaa	8400
agaatagaaa aaataaagca agagcatgaa acatcatggc actatgacca agaccacca	8460
tacaaaacgt gggcatacca tggtagctat gaaacaaaac agactggatc agcatcatcc	8520
atggtcaacg gagtggtcag gctgctgaca aaaccttggg acgtcgtccc catggtgaca	8580
cagatggcaa tgacagacac gactccatth ggacaacagc gcgtttttta agagaaagtg	8640
gacacgagaa cccaagaacc gaaagaaggc acgaagaaac taatgaaaat aacagcagag	8700
tggcttttga aagaattagg gaagaaaaag acaccagga tgtgcaccag agaagaattc	8760
acaagaaagg tgagaagcaa tgcagccttg ggggccatat tcatgatga gaacaagtgg	8820
aagtcggcac gtgaggctgt tgaagatagt aggttttggg agctggttga caaggaaagg	8880
aatctccatc ttgaaggaaa gtgtgaaaca tgttgttaca acatgatggg aaaaagagag	8940
aagaagctag ggaattcgg caaggcaaaa ggcagcagag ccatatggta catgtggctt	9000
ggagcacgct tcttagagtt tgaagcccta ggattcttaa atgaagatca ctggttctcc	9060
agagagaact ccctgagtgg agtggaagga gaagggtgc acaagctagg ttacattcta	9120
agagacgtga gcaagaaaga gggaggagca atgtatgccg atgacaccgc aggatgggat	9180
acaagaatca cactagaaga ctaaaaaat gaagaaatgg taacaaacca catggaagga	9240
gaacacaaga aactagccga ggccattttc aaactaacgt accaaaacaa ggtggtgcgt	9300

```

gtgcaaagac caacaccaag aggcacagta atggacatca tatcgagaag agaccaaaga 9360
ggtagtggac aagttggcac ctatggactc aatactttca ccaatatgga agcccaacta 9420
atcagacaga tggagggaga aggagtcttt aaaagcattc agcacctaac aatcacagaa 9480
gaaatcgctg tgcaaaactg gttagcaaga gtggggcgcg aaagggtatc aagaatggcc 9540
atcagtggag atgatttgtgt tgtgaaacct ttagatgaca ggttcgcaag cgctttaaca 9600
gctctaaatg acatgggaaa gattaggaaa gacatacaac aatgggaacc ttcaagagga 9660
tggaatgatt ggacacaagt gcccttctgt tcacaccatt tccatgagtt aatcatgaaa 9720
gacggtcgcg tactcgttgt tccatgtaga aaccaagatg aactgattgg cagagcccga 9780
atctcccaag gagcaggggtg gtctttgcgg gagacggcct gtttggggaa gtcttacgcc 9840
caaatgtgga gcttgatgta cttccacaga cgcgacctca ggctggcggc aaatgctatt 9900
tgctcggcag taccatcaca ttgggttcca acaagtcgaa caacctggtc catacatgct 9960
aaacatgaat ggatgacaac ggaagacatg ctgacagtct ggaacagggt gtggattcaa 10020
gaaaacccat ggatggaaga caaaactcca gtggaatcat gggaggaaat cccatacttg 10080
gggaaaagag aagaccaatg gtgcggctca ttgattgggt taacaagcag ggccacctgg 10140
gcaaagaaca tccaagcagc aataaatcaa gttagatccc ttataggcaa tgaagaatac 10200
acagattaca tgccatccat gaaaagattc agaagagaag aggaagaagc aggagtcttg 10260
tggtagaaag caaaactaac atgaaacaag gctagaagtc aggtcggatt aagccatagt 10320
acggaaaaaa ctatgctacc tgtgagcccc gtccaaggac gttaaaagaa gtcaggccat 10380
cataaatgcc atagcttgag taaactatgc agcctgtagc tccacctgag aaggtgtaaa 10440
aaatccggga ggccacaaac catggaagct gtacgcatgg cgtagtggac tagcggttag 10500
aggagacccc tcccttacia atcgagcaa caatgggggc ccaaggcgag atgaagctgt 10560
agtctcgctg gaaggactag aggttagagg agaccccccc gaaacaaaaa acagcatatt 10620
gacgctggga aagaccagag atcctgctgt ctctcagca tcattccagg cacagaacgc 10680
cagaaaatgg aatggtgctg ttgaatcaac aggttct 10717

```

<210> 8

<211> 3389

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Virut dengue typ huyết thanh 3, BVS

<400> 8

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

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
 20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
 35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
 50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
 65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
 85 90 95

Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
 100 105 110

Met Ala Phe His Leu Thr Thr Arg Asp Gly Glu Pro Arg Met Ile Val
 115 120 125

Gly Lys Asn Glu Arg Gly Lys Ser Leu Leu Phe Lys Thr Ala Ser Gly
 130 135 140

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

Asp Thr Val Thr Tyr Lys Cys Pro His Ile Thr Glu Val Glu Pro Glu
 165 170 175

Asp Ile Asp Cys Trp Cys Asn Leu Thr Ser Thr Trp Val Thr Tyr Gly
 180 185 190

Thr Cys Asn Gln Ala Gly Glu His Arg Arg Asp Lys Arg Ser Val Ala
 195 200 205

Leu Ala Pro His Val Gly Met Gly Leu Asp Thr Arg Thr Gln Thr Trp
 210 215 220

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

Ala Leu Arg His Pro Gly Phe Thr Ile Leu Ala Leu Phe Leu Ala His
 245 250 255

Tyr Ile Gly Thr Ser Leu Thr Gln Lys Val Val Ile Phe Ile Leu Leu
 260 265 270

Met Leu Val Thr Pro Ser Met Thr Met Arg Cys Val Gly Val Gly Asn
 275 280 285

Arg Asp Phe Val Glu Gly Leu Ser Gly Ala Thr Trp Val Asp Val Val
 290 295 300

Leu Glu His Gly Gly Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr
 305 310 315 320

Leu Asp Ile Glu Leu Gln Lys Thr Glu Ala Thr Gln Leu Ala Thr Leu
 325 330 335

Arg Lys Leu Cys Ile Glu Gly Lys Ile Thr Asn Ile Thr Thr Asp Ser
 340 345 350

Arg Cys Pro Thr Gln Gly Glu Ala Ile Leu Pro Glu Glu Gln Asp Gln
 355 360 365

Asn Tyr Val Cys Lys His Thr Tyr Val Asp Arg Gly Trp Gly Asn Gly
 370 375 380

Cys Gly Leu Phe Gly Lys Gly Ser Leu Val Thr Cys Ala Lys Phe Gln
 385 390 395 400

Cys Leu Glu Ser Ile Glu Gly Lys Val Val Gln His Glu Asn Leu Lys
 405 410 415

Tyr Thr Val Ile Ile Thr Val His Thr Gly Asp Gln His Gln Val Gly
 420 425 430

Asn Glu Thr Gln Gly Val Thr Ala Glu Ile Thr Pro Gln Ala Ser Thr
 435 440 445

Ala Glu Ala Ile Leu Pro Glu Tyr Gly Thr Leu Gly Leu Glu Cys Ser
 450 455 460

Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Ile Ser Leu Thr Met Lys
 465 470 475 480

Asn Lys Ala Trp Met Val His Arg Gln Trp Phe Phe Asp Leu Pro Leu
 485 490 495

Pro Trp Thr Ser Gly Ala Ser Ala Glu Thr Pro Thr Trp Asn Arg Lys
 500 505 510

Glu Leu Leu Val Thr Phe Lys Asn Ala His Ala Lys Lys Gln Glu Val
 515 520 525

Val Val Leu Gly Ser Gln Glu Gly Ala Met His Thr Ala Leu Thr Gly
 530 535 540

Ala Thr Glu Ile Gln Thr Ser Gly Gly Thr Ser Ile Phe Ala Gly His
 545 550 555 560

Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Glu Leu Lys Gly Met Ser
 565 570 575

Tyr Ala Met Cys Leu Ser Ser Phe Val Leu Lys Lys Glu Val Ser Glu
 580 585 590

Thr Gln His Gly Thr Ile Leu Ile Lys Val Glu Tyr Lys Gly Glu Asp
 595 600 605

Ala Pro Cys Lys Ile Pro Phe Ser Thr Glu Asp Gly Gln Gly Lys Ala
 610 615 620

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

Glu Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Glu Ser Asn Ile
 645 650 655

Val Ile Gly Ile Gly Asp Lys Ala Leu Lys Ile Asn Trp Tyr Lys Lys
 660 665 670

Gly Ser Ser Ile Gly Lys Met Phe Glu Ala Thr Ala Arg Gly Ala Arg

675	680	685
Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser Val Gly		
690	695	700
Gly Val Leu Asn Ser Leu Gly Lys Met Val His Gln Ile Phe Gly Ser		
705	710	715
Ala Tyr Thr Ala Leu Phe Gly Gly Val Ser Trp Met Met Lys Ile Gly		
	725	730
		735
Ile Gly Val Leu Leu Thr Trp Ile Gly Leu Asn Ser Lys Asn Thr Ser		
	740	745
		750
Met Ser Phe Ser Cys Ile Ala Ala Gly Ile Val Thr Leu Tyr Leu Gly		
	755	760
		765
Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys Asn Lys		
	770	775
		780
Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val His Thr		
785	790	795
		800
Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys Leu Ala		
	805	810
		815
Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile Arg Ser		
	820	825
		830
Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro Glu Leu		
	835	840
		845
Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met Thr Gly		
	850	855
		860
Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg Pro Gln		
865	870	875
		880
Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala Lys Met		
	885	890
		895
Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly Pro Glu		
	900	905
		910

Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu Glu Val
 915 920 925

Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu Lys Leu
 930 935 940

Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser Ala Ala
 945 950 955 960

Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp Ile Glu
 965 970 975

Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe Ile Glu
 980 985 990

Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser Asn Gly
 995 1000 1005

Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala Gly Pro
 1010 1015 1020

Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln Ile Thr
 1025 1030 1035

Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp Phe Cys
 1040 1045 1050

Asp Gly Thr Thr Val Val Val Thr Glu Asp Cys Gly Asn Arg Gly
 1055 1060 1065

Pro Ser Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile Thr Glu
 1070 1075 1080

Trp Cys Cys Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr Arg Gly
 1085 1090 1095

Glu Asp Gly Cys Trp Tyr Gly Met Glu Ile Arg Pro Leu Lys Glu
 1100 1105 1110

Lys Glu Glu Asn Leu Val Asn Ser Leu Val Thr Ala Gly His Gly
 1115 1120 1125

..

Gln Val	Asp Asn Phe Ser	Leu	Gly Val Leu Gly	Met	Ala Leu Phe
1130		1135		1140	
Leu Glu	Glu Met Leu Arg	Thr	Arg Val Gly Thr	Lys	His Ala Ile
1145		1150		1155	
Leu Leu	Val Ala Val Ser	Phe	Val Thr Leu Ile	Thr	Gly Asn Met
1160		1165		1170	
Ser Phe	Arg Asp Leu Gly	Arg	Val Met Val Met	Val	Gly Ala Thr
1175		1180		1185	
Met Thr	Asp Asp Ile Gly	Met	Gly Val Thr Tyr	Leu	Ala Leu Leu
1190		1195		1200	
Ala Ala	Phe Lys Val Arg	Pro	Thr Phe Ala Ala	Gly	Leu Leu Leu
1205		1210		1215	
Arg Lys	Leu Thr Ser Lys	Glu	Leu Met Met Thr	Thr	Ile Gly Ile
1220		1225		1230	
Val Leu	Leu Ser Gln Ser	Thr	Ile Pro Glu Thr	Ile	Leu Glu Leu
1235		1240		1245	
Thr Asp	Ala Leu Ala Leu	Gly	Met Met Val Leu	Lys	Met Val Arg
1250		1255		1260	
Asn Met	Glu Lys Tyr Gln	Leu	Ala Val Thr Ile	Met	Ala Ile Leu
1265		1270		1275	
Cys Val	Pro Asn Ala Val	Ile	Leu Gln Asn Ala	Trp	Lys Val Ser
1280		1285		1290	
Cys Thr	Ile Leu Ala Val	Val	Ser Val Ser Pro	Leu	Phe Leu Thr
1295		1300		1305	
Ser Ser	Gln Gln Lys Thr	Asp	Trp Ile Pro Leu	Ala	Leu Thr Ile
1310		1315		1320	
Lys Gly	Leu Asn Pro Thr	Ala	Ile Phe Leu Thr	Thr	Leu Ser Arg
1325		1330		1335	

Thr	Ser	Lys	Lys	Arg	Ser	Trp	Pro	Leu	Asn	Glu	Ala	Ile	Met	Ala
1340						1345					1350			
Val	Gly	Met	Val	Ser	Ile	Leu	Ala	Ser	Ser	Leu	Leu	Lys	Asn	Asp
1355						1360					1365			
Ile	Pro	Met	Thr	Gly	Pro	Leu	Val	Ala	Gly	Gly	Leu	Leu	Thr	Val
1370						1375					1380			
Cys	Tyr	Val	Leu	Thr	Gly	Arg	Ser	Ala	Asp	Leu	Glu	Leu	Glu	Arg
1385						1390					1395			
Ala	Ala	Asp	Val	Lys	Trp	Glu	Asp	Gln	Ala	Glu	Ile	Ser	Gly	Ser
1400						1405					1410			
Ser	Pro	Ile	Leu	Ser	Ile	Thr	Ile	Ser	Glu	Asp	Gly	Ser	Met	Ser
1415						1420					1425			
Ile	Lys	Asn	Glu	Glu	Glu	Glu	Gln	Thr	Leu	Thr	Ile	Leu	Ile	Arg
1430						1435					1440			
Thr	Gly	Leu	Leu	Val	Ile	Ser	Gly	Leu	Phe	Pro	Val	Ser	Ile	Pro
1445						1450					1455			
Ile	Thr	Ala	Ala	Ala	Trp	Tyr	Leu	Trp	Glu	Val	Lys	Lys	Gln	Arg
1460						1465					1470			
Ala	Gly	Val	Leu	Trp	Asp	Val	Pro	Ser	Pro	Pro	Pro	Met	Gly	Lys
1475						1480					1485			
Ala	Glu	Leu	Glu	Asp	Gly	Ala	Tyr	Arg	Ile	Lys	Gln	Lys	Gly	Ile
1490						1495					1500			
Leu	Gly	Tyr	Ser	Gln	Ile	Gly	Ala	Gly	Val	Tyr	Lys	Glu	Gly	Thr
1505						1510					1515			
Phe	His	Thr	Met	Trp	His	Val	Thr	Arg	Gly	Ala	Val	Leu	Met	His
1520						1525					1530			
Lys	Gly	Lys	Arg	Ile	Glu	Pro	Ser	Trp	Ala	Asp	Val	Lys	Lys	Asp
1535						1540					1545			
Leu	Ile	Ser	Tyr	Gly	Gly	Gly	Trp	Lys	Leu	Glu	Gly	Glu	Trp	Lys

..

1550		1555		1560
Glu Gly	Glu Glu Val Gln Val	Leu Ala Leu Glu Pro	Gly Lys Asn	
1565	1570	1575		
Pro Arg	Ala Val Gln Thr Lys	Pro Gly Leu Phe Lys	Thr Asn Ala	
1580	1585	1590		
Gly Thr	Ile Gly Ala Val Ser	Leu Asp Phe Ser Pro	Gly Thr Ser	
1595	1600	1605		
Gly Ser	Pro Ile Ile Asp Lys	Lys Gly Lys Val Val	Gly Leu Tyr	
1610	1615	1620		
Gly Asn	Gly Val Val Thr Arg	Ser Gly Ala Tyr Val	Ser Ala Ile	
1625	1630	1635		
Ala Gln	Thr Glu Lys Ser Ile	Glu Asp Asn Pro Glu	Ile Glu Asp	
1640	1645	1650		
Asp Ile	Phe Arg Lys Arg Arg	Leu Thr Ile Met Asp	Leu His Pro	
1655	1660	1665		
Gly Ala	Gly Lys Thr Lys Arg	Tyr Leu Pro Ala Ile	Val Arg Glu	
1670	1675	1680		
Ala Ile	Lys Arg Gly Leu Arg	Thr Leu Ile Leu Ala	Pro Thr Arg	
1685	1690	1695		
Val Val	Ala Ala Glu Met Glu	Glu Ala Leu Arg Gly	Leu Pro Ile	
1700	1705	1710		
Arg Tyr	Gln Thr Pro Ala Ile	Arg Ala Val His Thr	Gly Arg Glu	
1715	1720	1725		
Ile Val	Asp Leu Met Cys His	Ala Thr Phe Thr Met	Arg Leu Leu	
1730	1735	1740		
Ser Pro	Val Arg Val Pro Asn	Tyr Asn Leu Ile Ile	Met Asp Glu	
1745	1750	1755		
Ala His	Phe Thr Asp Pro Ala	Ser Ile Ala Ala Arg	Gly Tyr Ile	
1760	1765	1770		

~

Ser	Thr	Arg	Val	Glu	Met	Gly	Glu	Ala	Ala	Gly	Ile	Phe	Met	Thr
1775						1780					1785			
Ala	Thr	Pro	Pro	Gly	Ser	Arg	Asp	Pro	Phe	Pro	Gln	Ser	Asn	Ala
1790						1795					1800			
Pro	Ile	Ile	Asp	Glu	Glu	Arg	Glu	Ile	Pro	Glu	Arg	Ser	Trp	Asn
1805						1810					1815			
Ser	Gly	His	Glu	Trp	Val	Thr	Asp	Phe	Lys	Gly	Lys	Thr	Val	Trp
1820						1825					1830			
Phe	Val	Pro	Ser	Ile	Lys	Ala	Gly	Asn	Asp	Ile	Ala	Ala	Cys	Leu
1835						1840					1845			
Arg	Lys	Asn	Gly	Lys	Lys	Val	Ile	Gln	Leu	Ser	Arg	Lys	Thr	Phe
1850						1855					1860			
Asp	Ser	Glu	Tyr	Val	Lys	Thr	Arg	Thr	Asn	Asp	Trp	Asp	Phe	Val
1865						1870					1875			
Val	Thr	Thr	Asp	Ile	Ser	Glu	Met	Gly	Ala	Asn	Phe	Lys	Ala	Glu
1880						1885					1890			
Arg	Val	Ile	Asp	Pro	Arg	Arg	Cys	Met	Lys	Pro	Val	Ile	Leu	Thr
1895						1900					1905			
Asp	Gly	Glu	Glu	Arg	Val	Ile	Leu	Ala	Gly	Pro	Met	Pro	Val	Thr
1910						1915					1920			
His	Ser	Ser	Ala	Ala	Gln	Arg	Arg	Gly	Arg	Ile	Gly	Arg	Asn	Pro
1925						1930					1935			
Lys	Asn	Glu	Asn	Asp	Gln	Tyr	Ile	Tyr	Met	Gly	Glu	Pro	Leu	Glu
1940						1945					1950			
Asn	Asp	Glu	Asp	Cys	Ala	His	Trp	Lys	Glu	Ala	Lys	Met	Leu	Leu
1955						1960					1965			
Asp	Asn	Ile	Asn	Thr	Pro	Glu	Gly	Ile	Ile	Pro	Ser	Met	Phe	Glu
1970						1975					1980			

Pro Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr Arg Leu
1985 1990 1995

Arg Gly Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg Arg Gly
2000 2005 2010

Asp Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu Gly Ile
2015 2020 2025

Asn Tyr Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys Asn Asn
2030 2035 2040

Gln Ile Leu Glu Glu Asn Val Glu Val Glu Ile Trp Thr Lys Glu
2045 2050 2055

Gly Glu Arg Lys Lys Leu Lys Pro Arg Trp Leu Asp Ala Arg Ile
2060 2065 2070

Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe Lys Glu Phe Ala Ala
2075 2080 2085

Gly Arg Lys Ser Leu Thr Leu Asn Leu Ile Thr Glu Met Gly Arg
2090 2095 2100

Leu Pro Thr Phe Met Thr Gln Lys Ala Arg Asn Ala Leu Asp Asn
2105 2110 2115

Leu Ala Val Leu His Thr Ala Glu Ala Gly Gly Arg Ala Tyr Asn
2120 2125 2130

His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu Leu Leu
2135 2140 2145

Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu Phe Leu
2150 2155 2160

Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met Cys Cys
2165 2170 2175

Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile Gln Pro
2180 2185 2190

His	Trp	Ile	Ala	Ala	Ser	Ile	Ile	Leu	Glu	Phe	Phe	Leu	Ile	Val
2195						2200					2205			
Leu	Leu	Ile	Pro	Glu	Pro	Glu	Lys	Gln	Arg	Thr	Pro	Gln	Asp	Asn
2210						2215					2220			
Gln	Leu	Thr	Tyr	Val	Val	Ile	Ala	Ile	Leu	Thr	Val	Val	Ala	Ala
2225						2230					2235			
Thr	Met	Ala	Asn	Glu	Met	Gly	Phe	Leu	Glu	Lys	Thr	Lys	Lys	Asp
2240						2245					2250			
Leu	Gly	Leu	Gly	Ser	Ile	Ala	Thr	Gln	Gln	Pro	Glu	Ser	Asn	Ile
2255						2260					2265			
Leu	Asp	Ile	Asp	Leu	Arg	Pro	Ala	Ser	Ala	Trp	Thr	Leu	Tyr	Ala
2270						2275					2280			
Val	Ala	Thr	Thr	Phe	Val	Thr	Pro	Met	Leu	Arg	His	Ser	Ile	Glu
2285						2290					2295			
Asn	Ser	Ser	Val	Asn	Val	Ser	Leu	Thr	Ala	Ile	Ala	Asn	Gln	Ala
2300						2305					2310			
Thr	Val	Leu	Met	Gly	Leu	Gly	Lys	Gly	Trp	Pro	Leu	Ser	Lys	Met
2315						2320					2325			
Asp	Ile	Gly	Val	Pro	Leu	Leu	Ala	Ile	Gly	Cys	Tyr	Ser	Gln	Val
2330						2335					2340			
Asn	Pro	Ile	Thr	Leu	Thr	Ala	Ala	Leu	Phe	Leu	Leu	Val	Ala	His
2345						2350					2355			
Tyr	Ala	Ile	Ile	Gly	Pro	Gly	Leu	Gln	Ala	Lys	Ala	Thr	Arg	Glu
2360						2365					2370			
Ala	Gln	Lys	Arg	Ala	Ala	Ala	Gly	Ile	Met	Lys	Asn	Pro	Thr	Val
2375						2380					2385			
Asp	Gly	Ile	Thr	Val	Ile	Asp	Leu	Asp	Pro	Ile	Pro	Tyr	Asp	Pro
2390						2395					2400			
Lys	Phe	Glu	Lys	Gln	Leu	Gly	Gln	Val	Met	Leu	Leu	Val	Leu	Cys

2405		2410		2415
Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu Cys Glu				
2420		2425		2430
Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp Glu Gly				
2435		2440		2445
Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser Met Ala				
2450		2455		2460
Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu Leu Phe				
2465		2470		2475
Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr Gly Asn				
2480		2485		2490
Ile Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu Asn Ala				
2495		2500		2505
Leu Gly Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly Ile Gln				
2510		2515		2520
Glu Val Asp Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg Gly Glu				
2525		2530		2535
Thr Asp His His Ala Val Ser Arg Gly Ser Ala Lys Leu Arg Trp				
2540		2545		2550
Phe Val Glu Arg Asn Met Val Thr Pro Glu Gly Lys Val Val Asp				
2555		2560		2565
Leu Gly Cys Gly Arg Gly Gly Trp Ser Tyr Tyr Cys Gly Gly Leu				
2570		2575		2580
Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly Gly Pro Gly				
2585		2590		2595
His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn Leu Val				
2600		2605		2610
Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro Glu Lys				
2615		2620		2625

Cys	Asp	Thr	Leu	Leu	Cys	Asp	Ile	Gly	Glu	Ser	Ser	Pro	Asn	Pro
2630						2635					2640			
Thr	Val	Glu	Ala	Gly	Arg	Thr	Leu	Arg	Val	Leu	Asn	Leu	Val	Glu
2645						2650					2655			
Asn	Trp	Leu	Asn	Asn	Asn	Thr	Gln	Phe	Cys	Ile	Lys	Val	Leu	Asn
2660						2665					2670			
Pro	Tyr	Met	Pro	Ser	Val	Ile	Glu	Lys	Met	Glu	Ala	Leu	Gln	Arg
2675						2680					2685			
Lys	Tyr	Gly	Gly	Ala	Leu	Val	Arg	Asn	Pro	Leu	Ser	Arg	Asn	Ser
2690						2695					2700			
Thr	His	Glu	Met	Tyr	Trp	Val	Ser	Asn	Ala	Ser	Gly	Asn	Ile	Val
2705						2710					2715			
Ser	Ser	Val	Asn	Met	Ile	Ser	Arg	Met	Leu	Ile	Asn	Arg	Phe	Thr
2720						2725					2730			
Met	Arg	Tyr	Lys	Lys	Ala	Thr	Tyr	Glu	Pro	Asp	Val	Asp	Leu	Gly
2735						2740					2745			
Ser	Gly	Thr	Arg	Asn	Ile	Gly	Ile	Glu	Ser	Glu	Ile	Pro	Asn	Leu
2750						2755					2760			
Asp	Ile	Ile	Gly	Lys	Arg	Ile	Glu	Lys	Ile	Lys	Gln	Glu	His	Glu
2765						2770					2775			
Thr	Ser	Trp	His	Tyr	Asp	Gln	Asp	His	Pro	Tyr	Lys	Thr	Trp	Ala
2780						2785					2790			
Tyr	His	Gly	Ser	Tyr	Glu	Thr	Lys	Gln	Thr	Gly	Ser	Ala	Ser	Ser
2795						2800					2805			
Met	Val	Asn	Gly	Val	Val	Arg	Leu	Leu	Thr	Lys	Pro	Trp	Asp	Val
2810						2815					2820			
Val	Pro	Met	Val	Thr	Gln	Met	Ala	Met	Thr	Asp	Thr	Thr	Pro	Phe
2825						2830					2835			

Gly	Gln	Gln	Arg	Val	Phe	Lys	Glu	Lys	Val	Asp	Thr	Arg	Thr	Gln
2840						2845					2850			
Glu	Pro	Lys	Glu	Gly	Thr	Lys	Lys	Leu	Met	Lys	Ile	Thr	Ala	Glu
2855						2860					2865			
Trp	Leu	Trp	Lys	Glu	Leu	Gly	Lys	Lys	Lys	Thr	Pro	Arg	Met	Cys
2870						2875					2880			
Thr	Arg	Glu	Glu	Phe	Thr	Arg	Lys	Val	Arg	Ser	Asn	Ala	Ala	Leu
2885						2890					2895			
Gly	Ala	Ile	Phe	Thr	Asp	Glu	Asn	Lys	Trp	Lys	Ser	Ala	Arg	Glu
2900						2905					2910			
Ala	Val	Glu	Asp	Ser	Arg	Phe	Trp	Glu	Leu	Val	Asp	Lys	Glu	Arg
2915						2920					2925			
Asn	Leu	His	Leu	Glu	Gly	Lys	Cys	Glu	Thr	Cys	Val	Tyr	Asn	Met
2930						2935					2940			
Met	Gly	Lys	Arg	Glu	Lys	Lys	Leu	Gly	Glu	Phe	Gly	Lys	Ala	Lys
2945						2950					2955			
Gly	Ser	Arg	Ala	Ile	Trp	Tyr	Met	Trp	Leu	Gly	Ala	Arg	Phe	Leu
2960						2965					2970			
Glu	Phe	Glu	Ala	Leu	Gly	Phe	Leu	Asn	Glu	Asp	His	Trp	Phe	Ser
2975						2980					2985			
Arg	Glu	Asn	Ser	Leu	Ser	Gly	Val	Glu	Gly	Glu	Gly	Leu	His	Lys
2990						2995					3000			
Leu	Gly	Tyr	Ile	Leu	Arg	Asp	Val	Ser	Lys	Lys	Glu	Gly	Gly	Ala
3005						3010					3015			
Met	Tyr	Ala	Asp	Asp	Thr	Ala	Gly	Trp	Asp	Thr	Arg	Ile	Thr	Leu
3020						3025					3030			
Glu	Asp	Leu	Lys	Asn	Glu	Glu	Met	Val	Thr	Asn	His	Met	Glu	Gly
3035						3040					3045			

Glu His Lys Lys Leu Ala Glu Ala Ile Phe Lys Leu Thr Tyr Gln
 3050 3055 3060
 Asn Lys Val Val Arg Val Gln Arg Pro Thr Pro Arg Gly Thr Val
 3065 3070 3075
 Met Asp Ile Ile Ser Arg Arg Asp Gln Arg Gly Ser Gly Gln Val
 3080 3085 3090
 Gly Thr Tyr Gly Leu Asn Thr Phe Thr Asn Met Glu Ala Gln Leu
 3095 3100 3105
 Ile Arg Gln Met Glu Gly Glu Gly Val Phe Lys Ser Ile Gln His
 3110 3115 3120
 Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn Trp Leu Ala Arg
 3125 3130 3135
 Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser Gly Asp Asp
 3140 3145 3150
 Cys Val Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala Leu Thr
 3155 3160 3165
 Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln Gln Trp
 3170 3175 3180
 Glu Pro Ser Arg Gly Trp Asn Asp Trp Thr Gln Val Pro Phe Cys
 3185 3190 3195
 Ser His His Phe His Glu Leu Ile Met Lys Asp Gly Arg Val Leu
 3200 3205 3210
 Val Val Pro Cys Arg Asn Gln Asp Glu Leu Ile Gly Arg Ala Arg
 3215 3220 3225
 Ile Ser Gln Gly Ala Gly Trp Ser Leu Arg Glu Thr Ala Cys Leu
 3230 3235 3240
 Gly Lys Ser Tyr Ala Gln Met Trp Ser Leu Met Tyr Phe His Arg
 3245 3250 3255
 Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile Cys Ser Ala Val Pro

~

3260 3265 3270
 Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp Ser Ile His Ala
 3275 3280 3285
 Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr Val Trp Asn
 3290 3295 3300
 Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys Thr Pro
 3305 3310 3315
 Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg Glu Asp
 3320 3325 3330
 Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala Thr Trp
 3335 3340 3345
 Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser Leu Ile
 3350 3355 3360
 Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys Arg Phe
 3365 3370 3375
 Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
 3380 3385

 <210> 9
 <211> 3389
 <212> PRT
 <213> Trình tự nhân tạo

 <220>
 <223> Virut dengue typ huyết thanh 3, MVS

 <400> 9

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

 Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
 20 25 30

 Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
 35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
 50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
 65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
 85 90 95

Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
 100 105 110

Met Ala Phe His Leu Thr Thr Arg Asp Gly Glu Pro Arg Met Ile Val
 115 120 125

Gly Lys Asn Glu Arg Gly Lys Ser Leu Leu Phe Lys Thr Ala Ser Gly
 130 135 140

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

Asp Thr Val Thr Tyr Lys Cys Pro His Ile Thr Glu Val Glu Pro Glu
 165 170 175

Asp Ile Asp Cys Trp Cys Asn Leu Thr Ser Thr Trp Val Thr Tyr Gly
 180 185 190

Thr Cys Asn Gln Ala Gly Glu His Arg Arg Asp Lys Arg Ser Val Ala
 195 200 205

Leu Ala Pro His Val Gly Met Gly Leu Asp Thr Arg Thr Gln Thr Trp
 210 215 220

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

Ala Leu Arg His Pro Gly Phe Thr Ile Leu Ala Leu Phe Leu Ala His
 245 250 255

Tyr Ile Gly Thr Ser Leu Thr Gln Lys Val Val Ile Phe Ile Leu Leu
 260 265 270

Met Leu Val Thr Pro Ser Met Thr Met Arg Cys Val Gly Val Gly Asn

275	280	285
Arg Asp Phe Val Glu Gly Leu Ser Gly Ala Thr Trp Val Asp Val Val		
290	295	300
Leu Glu His Gly Gly Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr		
305	310	315 320
Leu Asp Ile Glu Leu Gln Lys Thr Glu Ala Thr Gln Leu Ala Thr Leu		
	325	330 335
Arg Lys Leu Cys Ile Glu Gly Lys Ile Thr Asn Ile Thr Thr Asp Ser		
	340	345 350
Arg Cys Pro Thr Gln Gly Glu Ala Ile Leu Pro Glu Glu Gln Asp Gln		
	355	360 365
Asn Tyr Val Cys Lys His Thr Tyr Val Asp Arg Gly Trp Gly Asn Gly		
	370	375 380
Cys Gly Leu Phe Gly Lys Gly Ser Leu Val Thr Cys Ala Lys Phe Gln		
385	390	395 400
Cys Leu Glu Ser Ile Glu Gly Lys Val Val Gln His Glu Asn Leu Lys		
	405	410 415
Tyr Thr Val Ile Ile Thr Val His Thr Gly Asp Gln His Gln Val Gly		
	420	425 430
Asn Glu Thr Gln Gly Val Thr Ala Glu Ile Thr Pro Gln Ala Ser Thr		
	435	440 445
Ala Glu Ala Ile Leu Pro Glu Tyr Gly Thr Leu Gly Leu Glu Cys Ser		
	450	455 460
Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Ile Ser Leu Thr Met Lys		
465	470	475 480
Asn Lys Ala Trp Met Val His Arg Gln Trp Phe Phe Asp Leu Pro Leu		
	485	490 495
Pro Trp Thr Ser Gly Ala Ser Ala Glu Thr Pro Thr Trp Asn Arg Lys		
	500	505 510

Glu Leu Leu Val Thr Phe Lys Asn Ala His Ala Lys Lys Gln Glu Val
 515 520 525

Val Val Leu Gly Ser Gln Glu Gly Ala Met His Thr Ala Leu Thr Gly
 530 535 540

Ala Thr Glu Ile Gln Thr Ser Gly Gly Thr Ser Ile Phe Ala Gly His
 545 550 555 560

Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Glu Leu Lys Gly Met Ser
 565 570 575

Tyr Ala Met Cys Leu Ser Ser Phe Val Leu Lys Lys Glu Val Ser Glu
 580 585 590

Thr Gln His Gly Thr Ile Leu Ile Lys Val Glu Tyr Lys Gly Glu Asp
 595 600 605

Ala Pro Cys Lys Ile Pro Phe Ser Thr Glu Asp Gly Gln Gly Lys Ala
 610 615 620

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

Glu Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Glu Ser Asn Ile
 645 650 655

Val Ile Gly Ile Gly Asp Lys Ala Leu Lys Ile Asn Trp Tyr Lys Lys
 660 665 670

Gly Ser Ser Ile Gly Lys Met Phe Glu Ala Thr Ala Arg Gly Ala Arg
 675 680 685

Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser Val Gly
 690 695 700

Gly Val Leu Asn Ser Leu Gly Lys Met Val His Gln Ile Phe Gly Ser
 705 710 715 720

Ala Tyr Thr Ala Leu Phe Gly Gly Val Ser Trp Met Met Lys Ile Gly
 725 730 735

Ile Gly Val Leu Leu Thr Trp Ile Gly Leu Asn Ser Lys Asn Thr Ser
740 745 750

Met Ser Phe Ser Cys Ile Ala Ala Gly Ile Val Thr Leu Tyr Leu Gly
755 760 765

Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys Asn Lys
770 775 780

Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val His Thr
785 790 795 800

Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys Leu Ala
805 810 815

Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile Arg Ser
820 825 830

Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro Glu Leu
835 840 845

Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met Thr Gly
850 855 860

Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg Pro Gln
865 870 875 880

Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala Lys Met
885 890 895

Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly Pro Glu
900 905 910

Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu Glu Val
915 920 925

Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu Lys Leu
930 935 940

Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser Ala Ala
945 950 955 960

Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp Ile Glu
 965 970 975

Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe Ile Glu
 980 985 990

Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser Asn Gly
 995 1000 1005

Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala Gly Pro
 1010 1015 1020

Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln Ile Thr
 1025 1030 1035

Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp Phe Cys
 1040 1045 1050

Asp Gly Thr Thr Val Val Val Thr Glu Asp Cys Gly Asn Arg Gly
 1055 1060 1065

Pro Ser Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile Thr Glu
 1070 1075 1080

Trp Cys Cys Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr Arg Gly
 1085 1090 1095

Glu Asp Gly Cys Trp Tyr Gly Met Glu Ile Arg Pro Leu Lys Glu
 1100 1105 1110

Lys Glu Glu Asn Leu Val Asn Ser Leu Val Thr Ala Gly His Gly
 1115 1120 1125

Gln Val Asp Asn Phe Ser Leu Gly Val Leu Gly Met Ala Leu Phe
 1130 1135 1140

Leu Glu Glu Met Leu Arg Thr Arg Val Gly Thr Lys His Ala Ile
 1145 1150 1155

Leu Leu Val Ala Val Ser Phe Val Thr Leu Ile Thr Gly Asn Met
 1160 1165 1170

Ser Phe Arg Asp Leu Gly Arg Val Met Val Met Val Gly Ala Thr

1175		1180		1185
Met Thr Asp Asp Ile Gly Met Gly Val Thr Tyr Leu Ala Leu Leu				
1190		1195		1200
Ala Ala Phe Lys Val Arg Pro Thr Phe Ala Ala Gly Leu Leu Leu				
1205		1210		1215
Arg Lys Leu Thr Ser Lys Glu Leu Met Met Thr Thr Ile Gly Ile				
1220		1225		1230
Val Leu Leu Ser Gln Ser Thr Ile Pro Glu Thr Ile Leu Glu Leu				
1235		1240		1245
Thr Asp Ala Leu Ala Leu Gly Met Met Val Leu Lys Met Val Arg				
1250		1255		1260
Asn Met Glu Lys Tyr Gln Leu Ala Val Thr Ile Met Ala Ile Leu				
1265		1270		1275
Cys Val Pro Asn Ala Val Ile Leu Gln Asn Ala Trp Lys Val Ser				
1280		1285		1290
Cys Thr Ile Leu Ala Val Val Ser Val Ser Pro Leu Phe Leu Thr				
1295		1300		1305
Ser Ser Gln Gln Lys Thr Asp Trp Ile Pro Leu Ala Leu Thr Ile				
1310		1315		1320
Lys Gly Leu Asn Pro Thr Ala Ile Phe Leu Thr Thr Leu Ser Arg				
1325		1330		1335
Thr Ser Lys Lys Arg Ser Trp Pro Leu Asn Glu Ala Ile Met Ala				
1340		1345		1350
Val Gly Met Val Ser Ile Leu Ala Ser Ser Leu Leu Lys Asn Asp				
1355		1360		1365
Ile Pro Met Thr Gly Pro Leu Val Ala Gly Gly Leu Leu Thr Val				
1370		1375		1380
Cys Tyr Val Leu Thr Gly Arg Ser Ala Asp Leu Glu Leu Glu Arg				
1385		1390		1395

Ala	Ala	Asp	Val	Lys	Trp	Glu	Asp	Gln	Ala	Glu	Ile	Ser	Gly	Ser
1400						1405					1410			
Ser	Pro	Ile	Leu	Ser	Ile	Thr	Ile	Ser	Glu	Asp	Gly	Ser	Met	Ser
1415						1420					1425			
Ile	Lys	Asn	Glu	Glu	Glu	Glu	Gln	Thr	Leu	Thr	Ile	Leu	Ile	Arg
1430						1435					1440			
Thr	Gly	Leu	Leu	Val	Ile	Ser	Gly	Leu	Phe	Pro	Val	Ser	Ile	Pro
1445						1450					1455			
Ile	Thr	Ala	Ala	Ala	Trp	Tyr	Leu	Trp	Glu	Val	Lys	Lys	Gln	Arg
1460						1465					1470			
Ala	Gly	Val	Leu	Trp	Asp	Val	Pro	Ser	Pro	Pro	Pro	Met	Gly	Lys
1475						1480						1485		
Ala	Glu	Leu	Glu	Asp	Gly	Ala	Tyr	Arg	Ile	Lys	Gln	Lys	Gly	Ile
1490						1495					1500			
Leu	Gly	Tyr	Ser	Gln	Ile	Gly	Ala	Gly	Val	Tyr	Lys	Glu	Gly	Thr
1505						1510					1515			
Phe	His	Thr	Met	Trp	His	Val	Thr	Arg	Gly	Ala	Val	Leu	Met	His
1520						1525					1530			
Lys	Gly	Lys	Arg	Ile	Glu	Pro	Ser	Trp	Ala	Asp	Val	Lys	Lys	Asp
1535						1540					1545			
Leu	Ile	Ser	Tyr	Gly	Gly	Gly	Trp	Lys	Leu	Glu	Gly	Glu	Trp	Lys
1550						1555					1560			
Glu	Gly	Glu	Glu	Val	Gln	Val	Leu	Ala	Leu	Glu	Pro	Gly	Lys	Asn
1565						1570					1575			
Pro	Arg	Ala	Val	Gln	Thr	Lys	Pro	Gly	Leu	Phe	Lys	Thr	Asn	Ala
1580						1585					1590			
Gly	Thr	Ile	Gly	Ala	Val	Ser	Leu	Asp	Phe	Ser	Pro	Gly	Thr	Ser
1595						1600					1605			

Gly Ser	Pro Ile Ile Asp	Lys	Lys Gly Lys Val Val	Gly Leu Tyr
1610		1615		1620
Gly Asn	Gly Val Val Thr	Arg	Ser Gly Ala Tyr Val	Ser Ala Ile
1625		1630		1635
Ala Gln	Thr Glu Lys Ser	Ile	Glu Asp Asn Pro Glu	Ile Glu Asp
1640		1645		1650
Asp Ile	Phe Arg Lys Arg	Arg	Leu Thr Ile Met Asp	Leu His Pro
1655		1660		1665
Gly Ala	Gly Lys Thr Lys	Arg	Tyr Leu Pro Ala Ile	Val Arg Glu
1670		1675		1680
Ala Ile	Lys Arg Gly Leu Arg	Thr	Leu Ile Leu Ala	Pro Thr Arg
1685		1690		1695
Val Val	Ala Ala Glu Met	Glu	Glu Ala Leu Arg Gly	Leu Pro Ile
1700		1705		1710
Arg Tyr	Gln Thr Pro Ala	Ile	Arg Ala Val His Thr	Gly Arg Glu
1715		1720		1725
Ile Val	Asp Leu Met Cys	His	Ala Thr Phe Thr Met	Arg Leu Leu
1730		1735		1740
Ser Pro	Val Arg Val Pro	Asn	Tyr Asn Leu Ile Ile	Met Asp Glu
1745		1750		1755
Ala His	Phe Thr Asp Pro	Ala	Ser Ile Ala Ala Arg	Gly Tyr Ile
1760		1765		1770
Ser Thr	Arg Val Glu Met	Gly	Glu Ala Ala Gly Ile	Phe Met Thr
1775		1780		1785
Ala Thr	Pro Pro Gly Ser	Arg	Asp Pro Phe Pro Gln	Ser Asn Ala
1790		1795		1800
Pro Ile	Ile Asp Glu Glu	Arg	Glu Ile Pro Glu Arg	Ser Trp Asn
1805		1810		1815

Ser	Gly	His	Glu	Trp	Val	Thr	Asp	Phe	Lys	Gly	Lys	Thr	Val	Trp
1820						1825					1830			
Phe	Val	Pro	Ser	Ile	Lys	Ala	Gly	Asn	Asp	Ile	Ala	Ala	Cys	Leu
1835						1840					1845			
Arg	Lys	Asn	Gly	Lys	Lys	Val	Ile	Gln	Leu	Ser	Arg	Lys	Thr	Phe
1850						1855					1860			
Asp	Ser	Glu	Tyr	Val	Lys	Thr	Arg	Thr	Asn	Asp	Trp	Asp	Phe	Val
1865						1870					1875			
Val	Thr	Thr	Asp	Ile	Ser	Glu	Met	Gly	Ala	Asn	Phe	Lys	Ala	Glu
1880						1885					1890			
Arg	Val	Ile	Asp	Pro	Arg	Arg	Cys	Met	Lys	Pro	Val	Ile	Leu	Thr
1895						1900					1905			
Asp	Gly	Glu	Glu	Arg	Val	Ile	Leu	Ala	Gly	Pro	Met	Pro	Val	Thr
1910						1915					1920			
His	Ser	Ser	Ala	Ala	Gln	Arg	Arg	Gly	Arg	Ile	Gly	Arg	Asn	Pro
1925						1930					1935			
Lys	Asn	Glu	Asn	Asp	Gln	Tyr	Ile	Tyr	Met	Gly	Glu	Pro	Leu	Glu
1940						1945					1950			
Asn	Asp	Glu	Asp	Cys	Ala	His	Trp	Lys	Glu	Ala	Lys	Met	Leu	Leu
1955						1960					1965			
Asp	Asn	Ile	Asn	Thr	Pro	Glu	Gly	Ile	Ile	Pro	Ser	Met	Phe	Glu
1970						1975					1980			
Pro	Glu	Arg	Glu	Lys	Val	Asp	Ala	Ile	Asp	Gly	Glu	Tyr	Arg	Leu
1985						1990					1995			
Arg	Gly	Glu	Ala	Arg	Lys	Thr	Phe	Val	Asp	Leu	Met	Arg	Arg	Gly
2000						2005					2010			
Asp	Leu	Pro	Val	Trp	Leu	Ala	Tyr	Arg	Val	Ala	Ala	Glu	Gly	Ile
2015						2020					2025			
Asn	Tyr	Ala	Asp	Arg	Arg	Trp	Cys	Phe	Asp	Gly	Val	Lys	Asn	Asn

2030		2035		2040
Gln Ile Leu Glu Glu Asn Val Glu Val Glu Ile Trp Thr Lys Glu				
2045		2050		2055
Gly Glu Arg Lys Lys Leu Lys Pro Arg Trp Leu Asp Ala Arg Ile				
2060		2065		2070
Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe Lys Glu Phe Ala Ala				
2075		2080		2085
Gly Arg Lys Ser Leu Thr Leu Asn Leu Ile Thr Glu Met Gly Arg				
2090		2095		2100
Leu Pro Thr Phe Met Thr Gln Lys Ala Arg Asp Ala Leu Asp Asn				
2105		2110		2115
Leu Ala Val Leu His Thr Ala Glu Ala Gly Gly Arg Ala Tyr Asn				
2120		2125		2130
His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu Leu Leu				
2135		2140		2145
Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu Phe Leu				
2150		2155		2160
Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met Cys Cys				
2165		2170		2175
Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile Gln Pro				
2180		2185		2190
His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu Ile Val				
2195		2200		2205
Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln Asp Asn				
2210		2215		2220
Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val Ala Ala				
2225		2230		2235
Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys Lys Asp				
2240		2245		2250

Leu Gly 2255	Leu Gly Ser Ile 2260	Ala Thr Gln Gln Pro Glu 2265	Ser Asn Ile
Leu Asp 2270	Ile Asp Leu Arg Pro 2275	Ala Ser Ala Trp Thr 2280	Leu Tyr Ala
Val Ala 2285	Thr Thr Phe Val Thr 2290	Pro Met Leu Arg His 2295	Ser Ile Glu
Asn Ser 2300	Ser Val Asn Val Ser 2305	Leu Thr Ala Ile Ala 2310	Asn Gln Ala
Thr Val 2315	Leu Met Gly Leu Gly 2320	Lys Gly Trp Pro Leu 2325	Ser Lys Met
Asp Ile 2330	Gly Val Pro Leu Leu 2335	Ala Ile Gly Cys Tyr 2340	Ser Gln Val
Asn Pro 2345	Ile Thr Leu Thr Ala 2350	Ala Leu Phe Leu Leu 2355	Val Ala His
Tyr Ala 2360	Ile Ile Gly Pro Gly 2365	Leu Gln Ala Lys Ala 2370	Thr Arg Glu
Ala Gln 2375	Lys Arg Ala Ala Ala 2380	Gly Ile Met Lys Asn 2385	Pro Thr Val
Asp Gly 2390	Ile Thr Val Ile Asp 2395	Leu Asp Pro Ile Pro 2400	Tyr Asp Pro
Lys Phe 2405	Glu Lys Gln Leu Gly 2410	Gln Val Met Leu Leu 2415	Val Leu Cys
Val Thr 2420	Gln Val Leu Met Met 2425	Arg Thr Thr Trp Ala 2430	Leu Cys Glu
Ala Leu 2435	Thr Leu Ala Thr Gly 2440	Pro Ile Ser Thr Leu 2445	Trp Glu Gly
Asn Pro 2450	Gly Arg Phe Trp Asn 2455	Thr Thr Ile Ala Val 2460	Ser Met Ala

Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu Leu Phe	2465	2470	2475
Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr Gly Asn	2480	2485	2490
Ile Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu Asn Ala	2495	2500	2505
Leu Gly Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly Ile Gln	2510	2515	2520
Glu Val Asp Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg Gly Glu	2525	2530	2535
Thr Asp His His Ala Val Ser Arg Gly Ser Ala Lys Leu Arg Trp	2540	2545	2550
Phe Val Glu Arg Asn Met Val Thr Pro Glu Gly Lys Val Val Asp	2555	2560	2565
Leu Gly Cys Gly Arg Gly Gly Trp Ser Tyr Tyr Cys Gly Gly Leu	2570	2575	2580
Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly Gly Pro Gly	2585	2590	2595
His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn Leu Val	2600	2605	2610
Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro Glu Lys	2615	2620	2625
Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser Pro Asn Pro	2630	2635	2640
Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu Val Glu	2645	2650	2655
Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys Val Leu Asn	2660	2665	2670

Pro	Tyr	Met	Pro	Ser	Val	Ile	Glu	Lys	Met	Glu	Ala	Leu	Gln	Arg
2675						2680					2685			
Lys	Tyr	Gly	Gly	Ala	Leu	Val	Arg	Asn	Pro	Leu	Ser	Arg	Asn	Ser
2690						2695					2700			
Thr	His	Glu	Met	Tyr	Trp	Val	Ser	Asn	Ala	Ser	Gly	Asn	Ile	Val
2705						2710					2715			
Ser	Ser	Val	Asn	Met	Ile	Ser	Arg	Met	Leu	Ile	Asn	Arg	Phe	Thr
2720						2725					2730			
Met	Arg	Tyr	Lys	Lys	Ala	Thr	Tyr	Glu	Pro	Asp	Val	Asp	Leu	Gly
2735						2740					2745			
Ser	Gly	Thr	Arg	Asn	Ile	Gly	Ile	Glu	Ser	Glu	Ile	Pro	Asn	Leu
2750						2755					2760			
Asp	Ile	Ile	Gly	Lys	Arg	Ile	Glu	Lys	Ile	Lys	Gln	Glu	His	Glu
2765						2770					2775			
Thr	Ser	Trp	His	Tyr	Asp	Gln	Asp	His	Pro	Tyr	Lys	Thr	Trp	Ala
2780						2785					2790			
Tyr	His	Gly	Ser	Tyr	Glu	Thr	Lys	Gln	Thr	Gly	Ser	Ala	Ser	Ser
2795						2800					2805			
Met	Val	Asn	Gly	Val	Val	Arg	Leu	Leu	Thr	Lys	Pro	Trp	Asp	Val
2810						2815					2820			
Val	Pro	Met	Val	Thr	Gln	Met	Ala	Met	Thr	Asp	Thr	Thr	Pro	Phe
2825						2830					2835			
Gly	Gln	Gln	Arg	Val	Phe	Lys	Glu	Lys	Val	Asp	Thr	Arg	Thr	Gln
2840						2845					2850			
Glu	Pro	Lys	Glu	Gly	Thr	Lys	Lys	Leu	Met	Lys	Ile	Thr	Ala	Glu
2855						2860					2865			
Trp	Leu	Trp	Lys	Glu	Leu	Gly	Lys	Lys	Lys	Thr	Pro	Arg	Met	Cys
2870						2875					2880			
Thr	Arg	Glu	Glu	Phe	Thr	Arg	Lys	Val	Arg	Ser	Asn	Ala	Ala	Leu

2885		2890		2895
Gly Ala Ile Phe Thr Asp	Glu Asn Lys Trp Lys Ser	Ala Arg Glu		
2900	2905	2910		
Ala Val Glu Asp Ser Arg	Phe Trp Glu Leu Val Asp	Lys Glu Arg		
2915	2920	2925		
Asn Leu His Leu Glu Gly	Lys Cys Glu Thr Cys Val	Tyr Asn Met		
2930	2935	2940		
Met Gly Lys Arg Glu Lys	Lys Leu Gly Glu Phe Gly	Lys Ala Lys		
2945	2950	2955		
Gly Ser Arg Ala Ile Trp	Tyr Met Trp Leu Gly Ala	Arg Phe Leu		
2960	2965	2970		
Glu Phe Glu Ala Leu Gly	Phe Leu Asn Glu Asp His	Trp Phe Ser		
2975	2980	2985		
Arg Glu Asn Ser Leu Ser	Gly Val Glu Gly Glu Gly	Leu His Lys		
2990	2995	3000		
Leu Gly Tyr Ile Leu Arg	Asp Val Ser Lys Lys Glu	Gly Gly Ala		
3005	3010	3015		
Met Tyr Ala Asp Asp Thr	Ala Gly Trp Asp Thr Arg	Ile Thr Leu		
3020	3025	3030		
Glu Asp Leu Lys Asn Glu	Glu Met Val Thr Asn His	Met Glu Gly		
3035	3040	3045		
Glu His Lys Lys Leu Ala	Glu Ala Ile Phe Lys Leu	Thr Tyr Gln		
3050	3055	3060		
Asn Lys Val Val Arg Val	Gln Arg Pro Thr Pro Arg	Gly Thr Val		
3065	3070	3075		
Met Asp Ile Ile Ser Arg	Arg Asp Gln Arg Gly Ser	Gly Gln Val		
3080	3085	3090		
Gly Thr Tyr Gly Leu Asn	Thr Phe Thr Asn Met Glu	Ala Gln Leu		
3095	3100	3105		

Ile	Arg	Gln	Met	Glu	Gly	Glu	Gly	Val	Phe	Lys	Ser	Ile	Gln	His
3110						3115					3120			
Leu	Thr	Ile	Thr	Glu	Glu	Ile	Ala	Val	Gln	Asn	Trp	Leu	Ala	Arg
3125						3130					3135			
Val	Gly	Arg	Glu	Arg	Leu	Ser	Arg	Met	Ala	Ile	Ser	Gly	Asp	Asp
3140						3145					3150			
Cys	Val	Val	Lys	Pro	Leu	Asp	Asp	Arg	Phe	Ala	Ser	Ala	Leu	Thr
3155						3160					3165			
Ala	Leu	Asn	Asp	Met	Gly	Lys	Ile	Arg	Lys	Asp	Ile	Gln	Gln	Trp
3170						3175					3180			
Glu	Pro	Ser	Arg	Gly	Trp	Asn	Asp	Trp	Thr	Gln	Val	Pro	Phe	Cys
3185						3190					3195			
Ser	His	His	Phe	His	Glu	Leu	Ile	Met	Lys	Asp	Gly	Arg	Val	Leu
3200						3205					3210			
Val	Val	Pro	Cys	Arg	Asn	Gln	Asp	Glu	Leu	Ile	Gly	Arg	Ala	Arg
3215						3220					3225			
Ile	Ser	Gln	Gly	Ala	Gly	Trp	Ser	Leu	Arg	Glu	Thr	Ala	Cys	Leu
3230						3235					3240			
Gly	Lys	Ser	Tyr	Ala	Gln	Met	Trp	Ser	Leu	Met	Tyr	Phe	His	Arg
3245						3250					3255			
Arg	Asp	Leu	Arg	Leu	Ala	Ala	Asn	Ala	Ile	Cys	Ser	Ala	Val	Pro
3260						3265					3270			
Ser	His	Trp	Val	Pro	Thr	Ser	Arg	Thr	Thr	Trp	Ser	Ile	His	Ala
3275						3280					3285			
Lys	His	Glu	Trp	Met	Thr	Thr	Glu	Asp	Met	Leu	Thr	Val	Trp	Asn
3290						3295					3300			
Arg	Val	Trp	Ile	Gln	Glu	Asn	Pro	Trp	Met	Glu	Asp	Lys	Thr	Pro
3305						3310					3315			

Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg Glu Asp
3320 3325 3330

Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala Thr Trp
3335 3340 3345

Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser Leu Ile
3350 3355 3360

Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys Arg Phe
3365 3370 3375

Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
3380 3385

<210> 10
<211> 10723
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Virut dengue typ huyết thanh 4, BVS

<220>
<221> n
<222> (3773)..(3773)
<223> a hoặc g

<220>
<221> n
<222> (7026)..(7026)
<223> t hoặc c

<220>
<221> n
<222> (7538)..(7538)
<223> t hoặc c

<400> 10
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60
gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120
aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180
ctgacaaaga gattctcact tggaatgctg cagggacgag gacctttaaa actgttcatg 240
gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300
tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360

ggaaggatgc tgaacatctt gaataggaga cgcagctctg caggcatgat cattatgctg 420
 attccaacag tgatggcggt ccatttaacc acgcgtgatg gcgaaccctt catgatagtg 480
 gcaaaacatg aaagggggag acctctcttg tttaagacaa cagaggggat caacaaatgc 540
 actctcattg ccatggactt ggggtgaaatg tgtgaggaca ctgtcacgta taaatgcccc 600
 ttactgggtca ataccgaacc tgaagacatt gattgctggg gcaatctcac gtctacctgg 660
 gtcattgtatg ggacatgcac ccagagcgga gaacggagac gagagaagcg ctcagtagct 720
 ttaacaccac attcaggaat gggattggaa acaagagctg agacatggat gtcattcgga 780
 ggggcttgga agcatgctca gagagtagag agctggatac tcagaaacc aggattcgcg 840
 ctcttgccag gatattatggc ttatatgatt gggcaaacag gaatccagcg aactgtcttc 900
 tttgtcctaa tgatgctggg cggcccatcc tacggaatgc gatgcgtagg agtaggaaac 960
 agagactttg tggaaggagt ctcagggtgga gcatgggtcg atctgggtgct agaacatgga 1020
 ggatgcgtca caaccatggc ccagggaaaa ccaaccttgg attttgaact gactaagaca 1080
 acagccaagg aagtggctct gttaagaacc tattgcattg aagcctcaat atcaaacata 1140
 accacggcaa caagatgtcc aacgcaagga gagccttacc taaaagagga acaagaccaa 1200
 cagtacattt gccggagaga tgtggtagac agaggggtgg gcaatggctg tggcttgttt 1260
 ggaaaaggag gagttgtgac atgtgcgaag ttttcatgtt cggggaagat aacaggcaat 1320
 ttgggtccaaa ttgagaacct tgaatacaca gtgggtgtaa cagtccacaa tggagacacc 1380
 catgcagtag gaaatgacac gtccaatcat ggagttacag ccacgataac tcccagggtca 1440
 ccatcgggtg aagtcaaatt gccggactat ggagaactaa cactcgattg tgaaccagg 1500
 tctggaattg actttaatga gatgattctg atgaaaatga aaaagaaaac atggcttgtg 1560
 cataagcaat ggttttttga tctacctcta ccatggacag caggagcaga cacatcagag 1620
 gttcactgga attacaaaga gagaatggtg acatttaagg ttcctcatgc caagagacag 1680
 gatgtgacag tgctgggatc tcaggaagga gccatgcatt ctgccctcgc tggagccaca 1740
 gaagtggact ccggtgatgg aaatcacatg tttgcaggac atctcaagtg caaagtccgt 1800
 atggagaaat tgagaatcaa gggaatgtca tacacgatgt gttcaggaaa gttctcaatt 1860
 gacaaagaga tggcagaaac acagcatggg acaacagtgg tgaaagtcaa gtatgaagg 1920
 gctggagctc cgtgtaaaagt ccccatagag ataagagatg tgaacaagga aaaagtgggt 1980
 gggcgatatc tctcatccac ccctttggct gagaatacca acagtgtaac caacatagag 2040
 ttagaacccc cctttgggga cagctacata gtgatagggt ttggaaacag tgcattaaca 2100

ctccattggt	tcaggaaagg	gagttccatt	ggcaagatgt	ttgagtccac	atacagaggt	2160
gcaaaacgaa	tggccattct	aggtgaaaca	gcttgggatt	ttggttccgt	tgggtggactg	2220
ttcacatcat	tgggaaaggc	tgtgcaccag	gtttttggaa	gtgtgtatac	aaccctgttt	2280
ggaggagtct	catggatgat	tagaatccta	attgggttcc	tagtgttgtg	gattggcacg	2340
aactcaagga	acacttcaat	ggctatgacg	tgcatactg	ccggcattgt	gacactgtat	2400
ttgggagtca	tgggtgcaggc	cgatagtgg	tgcgttgtga	gctggaaaaa	caaagaactg	2460
aatgtggca	gtgggatttt	catcacagac	aacgtgcaca	catggacaga	acaatacaag	2520
ttccaaccag	aatccccctt	aaaactagct	tcagctatcc	agaaagccca	tgaagaggac	2580
atttgtggaa	tccgctcagt	aacaagactg	gagaatctga	tgtggaaaca	aataacacca	2640
gaattgaatc	acattctatc	agaaaatgag	gtgaagttaa	ctattatgac	aggagacatc	2700
aaaggaatca	tgcaggcagg	aaaacgatct	ctgcggcctc	agcccactga	gctgaagtat	2760
tcattgaaaa	catggggcaa	agcaaaaatg	ctctctacag	agtctcataa	ccagaccttt	2820
ctcattgatg	gccccgaaac	agcagaatgc	cccaacacaa	atagagcttg	gaattcgctg	2880
gaagttgaag	actatggctt	tggagtattc	accaccaata	tatggctaaa	attgaaagaa	2940
aaacaggatg	tattctgcga	ctcaaaactc	atgtcagcgg	ccataaaaga	caacagagcc	3000
gtccatgccg	atatgggtta	ttggatagaa	agtgactca	atgacacatg	gaagatagag	3060
aaagcctctt	tcattgaagt	taaaaactgc	cactggccaa	aatcacacac	cctctggagc	3120
aatggagtgc	tagaaagtga	gatgataatt	caaagaatc	tcgctggacc	agtgtctcaa	3180
cacaactata	gaccaggcta	ccatacacia	ataacaggac	catggcatct	aggtaagctt	3240
gagatggact	ttgatttctg	tgatggaaca	acagtggtag	tgactgagga	ctgcggaaat	3300
agaggaccct	ctttgagaac	aaccactgcc	tctggaaaac	tcataacaga	atggtgctgc	3360
cgatcttgca	cattaccacc	gctaagatac	agaggtgagg	atgggtgctg	gtacgggatg	3420
gaaatcagac	cattgaagga	gaaagaagag	aatttgggtca	actccttgg	cacagctgga	3480
catgggcagg	tcgacaactt	ttcactagga	gtcttgggaa	tggcattgtt	cctggaggaa	3540
atgcttagga	cccagtagg	aacgaaacat	gcaatactac	tagttgcagt	ttcttttgtg	3600
acattgatca	cagggaacat	gtcctttaga	gacctgggaa	gagtgatgg	tatggtaggc	3660
gccactatga	cgggtgacat	aggtatgggc	gtgacttata	ttgccctact	agcagccttc	3720
aaagtcagac	caacttttgc	agctggacta	ctcttgagaa	agctgacctc	canggaattg	3780

atgatgacta ctataggaat tgtactcttc tcccagagca ccataccaga gaccattctt	3840
gagttgactg atgcgttagc cttaggcatg atggtcctca aaatgggtgag aaatatggaa	3900
aagtatcaat tggcagtgac tatcatggct atcttgtgcg tcccaaacgc agtgatatta	3960
caaaacgcat ggaaagtgag ttgcacaata ttggcagtggtgtccgtttc cccactgttc	4020
ttaacatcct cacagcaaaa aacagattgg ataccattag cattgacgat caaaggcttc	4080
aatccaacag ctatTTTTtct aacaaccctc tcaagaacca gcaagaaaag gagctggcca	4140
ttaaagtagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa	4200
aatgatattc ccatgacagg accattagtgt gctggagggtc tctcactgt gtgtacgtg	4260
ctcactggac gatcggccga ttggaactg gagagagcag ccgatgtcaa atgggaagac	4320
caggcagaga tatcaggaag cagtccaatc ctgtcaataa caatatcaga agatggtagc	4380
atgtcgataa aaaatgaaga ggaagaacaa aactgacca tactcattag aacaggattg	4440
ctggtgatct caggactttt tctgtatca ataccaatca cggcagcagc atggtacctg	4500
tgggaagtga agaaacaacg ggccggagta ttgtgggatg ttccttcacc cccacccatg	4560
ggaaaggctg aactggaaga tggagcctat agaattaagc aaaaagggat tcttgatat	4620
tcccagatcg gagccggagt ttacaaagaa ggaacattcc atacaatgtg gcatgtcaca	4680
cgtggcgctg ttctaattgca taaaggaaag aggattgaac catcatgggc ggacgtcaag	4740
aaagacctaa tatcatatgg aggaggctgg aagttagaag gagaatggaa ggaaggagaa	4800
gaagtccagg tattggcact ggagcctgga aaaaatccaa gagccgtcca aacgaaacct	4860
ggtcttttca aaaccaacgc cggaacaata ggtgctgtat ctctggactt ttctcctgga	4920
acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt	4980
gttgttacia ggagtgagc atatgtgagt gctatagccc agactgaaaa aagcattgaa	5040
gacaacccag agatcgaaga tgacattttc cgaaagagaa gactgaccat catggacctc	5100
caccaggag cgggaaagac gaagagatac cttccggcca tagtcagaga agctataaaa	5160
cggggtttga gaacattaat cttggcccc actagagttg tggcagctga aatggaggaa	5220
gcccttagag gacttccaat aagataccag accccagcca tcagagctgt gcacaccggg	5280
cgggagattg tggacctaat gtgtcatgcc acatttacca tgaggctgct atcaccagtt	5340
agagtgccaa actacaacct gattatcatg gacgaagccc atttcacaga tccagcaagt	5400
atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt	5460
atgacagcca ctccccggg aagcagagac ccatttcctc agagcaatgc accaatcata	5520

gatgaagaaa gagaaatccc tgaacgctcg tgggaattccg gacatgaatg ggtcacggat	5580
tttaaaggga agactgtttg gttcgtttcca agtataaaag caggaaatga tatagcagct	5640
tgcctgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag	5700
tatgtcaaga ctagaaccaa tgattgggac ttcgtggtta caactgacat ttcagaaatg	5760
ggtgccaat tcaaggctga gaggggttata gacccagac gctgcatgaa accagtcata	5820
ctaacagatg gtgaagagcg ggtgattctg gcaggaccta tgccagtgc cactctagt	5880
gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata	5940
tacatggggg aacctctgga aaatgatgaa gactgtgcac actggaaaga agctaaaatg	6000
ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt	6060
gaaaagggtg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt	6120
gtagacttaa tgagaagagg agacctacca gtctggttgg cctacagagt ggcagctgaa	6180
ggcatcaact acgcagacag aaggtggtgt tttgatggag tcaagaacaa ccaaactcta	6240
gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc	6300
agatggttgg atgctaggat ctattctgac cactggcgc taaaagaatt taaggaatth	6360
gcagccggaa gaaagtctct gacctgaac ctaatcacag aaatgggtag gctcccaacc	6420
ttcatgactc agaaggtaag agacgcactg gacaacttag cagtgtgca cacggctgag	6480
gcaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg	6540
cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca	6600
aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta	6660
tggtacgcac aaatacagcc aactggata gcagcttcaa taatactgga gttttttctc	6720
atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc	6780
tacgttgtca tagccatcct cacagtgggtg gccgcaacca tggcaaacga gatgggtttc	6840
ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caaccagca acccgagagc	6900
aacatcctgg acatagatct acgtcctgca tcagcatgga cgctgtatgc cgtggccaca	6960
acatttgta caccaatgtt gagacatagc attgaaaatt cctcagtga tgtgtcccta	7020
acagcnatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca	7080
aagatggaca tcggagtcc cttctcggc attggatgct actcacaagt caaccccata	7140
actctcacag cagctctttt cttattggta gcacattatg ccatcatagg gccaggactc	7200

caagcaaaag	caaccagaga	agctcagaaa	agagcagcgg	cgggcatcat	gaaaaaccca	7260
actgtcgatg	gaataacagt	gattgacctt	gatccaatac	cttatgatcc	aaagtttgaa	7320
aagcagttgg	gacaagtaat	gctcctagtc	ctctgcgtga	ctcaagtatt	gatgatgagg	7380
actacatggg	ctctgtgtga	ggctttaacc	ttagctaccg	ggcccatctc	cacattgtgg	7440
gaaggaaatc	cagggaggtt	ttggaacact	accattgcgg	tgtcaatggc	taacattttt	7500
agagggagtt	acttggccgg	agctggactt	ctcttttnta	ttatgaagaa	cacaaccaac	7560
acaagaaggg	gaactggcaa	cataggagag	acgcttgagg	agaaatggaa	aagccgattg	7620
aacgcattgg	gaaaaagtga	attccagatc	tacaagaaaa	gtggaatcca	ggaagtggat	7680
agaaccttag	caaaaagaag	cattaaaaga	ggagaaacgg	accatcacgc	tgtgtcgcga	7740
ggctcagcaa	aactgagatg	gttcgttgag	agaaacatgg	tcacaccaga	agggaaagta	7800
gtggacctcg	gttgtggcag	aggaggctgg	tcatactatt	gtggaggact	aaagaatgta	7860
agagaagtca	aaggcctaac	aaaaggagga	ccaggacacg	aagaacccat	ccccatgtca	7920
acatatgggt	ggaatctagt	gcgtcttcaa	agtggagttg	acgttttctt	catcccgccca	7980
gaaaagtgtg	acacattatt	gtgtgacata	ggggagtcac	caccaaattc	cacagtggaa	8040
gcaggacgaa	cactcagagt	ccttaactta	gtagaaaatt	ggttgaacaa	caacactcaa	8100
ttttgcataa	aggttctcaa	cccatatatg	ccctcagtca	tagaaaaaat	ggaagcacta	8160
caaaggaaat	atggaggagc	cttagtgagg	aatccactct	cacgaaactc	cacacatgag	8220
atgtactggg	tatccaatgc	ttccgggaac	atagtgtcat	cagtgaacat	gatttcaagg	8280
atgttgatca	acagatttac	aatgagatac	aagaaagcca	cttacgagcc	ggatgttgac	8340
ctcggaaagc	gaacccgtaa	catcgggatt	gaaagtgaga	taccaaacct	agatataatt	8400
gggaaaagaa	tagaaaaaat	aaagcaagag	catgaaacat	catggcacta	tgaccaagac	8460
cacccataca	aaacgtgggc	ataccatggg	agctatgaaa	caaaacagac	tggatcagca	8520
tcattccatg	tcaacggagt	ggtcaggctg	ctgacaaaac	cttgggacgt	cgtcccatg	8580
gtgacacaga	tggcaatgac	agacacgact	ccatttggac	aacagcgcgt	ttttaaagag	8640
aaagtggaca	cgagaaccca	agaaccgaaa	gaaggcacga	agaaactaat	gaaaataaca	8700
gcagagtggc	tttggaagaa	attagggaag	aaaaagacac	ccaggatgtg	caccagagaa	8760
gaattcacia	gaaaggtgag	aagcaatgca	gccttggggg	ccatattcac	tgatgagaac	8820
aagtggaagt	cggcacgtga	ggctgttgaa	gatagtaggt	tttgggagct	ggttgacaag	8880
gaaaggaatc	tccatcttga	aggaaagtgt	gaaacatgtg	tgtacaacat	gatgggaaaa	8940

agagagaaga agctagggga attcggcaag gcaaaaggca gcagagccat atggtacatg	9000
tggcttggag cacgcttctt agagtttgaa gccctaggat tcttaaatga agatcactgg	9060
ttctccagag agaactccct gagtggagtg gaaggagaag ggctgcacaa gctaggttac	9120
attctaagag acgtgagcaa gaaagaggga ggagcaatgt atgccgatga caccgcagga	9180
tgggatacaa gaatcacact agaagaccta aaaaatgaag aaatggtaac aaaccacatg	9240
gaaggagaac acaagaaact agccgaggcc attttcaaac taacgtacca aaacaagggtg	9300
gtgcgtgtgc aaagaccaac accaagaggc acagtaatgg acatcatatc gagaagagac	9360
caaagaggta gtggacaagt tggcacctat ggactcaata ctttcaccaa tatggaagcc	9420
caactaatca gacagatgga gggagaagga gtctttaaaa gcattcagca cctaacaatc	9480
acagaagaaa tcgctgtgca aaactgggta gcaagagtgg ggcgcgaaaag gttatcaaga	9540
atggccatca gtggagatga ttgtgttgtg aaacctttag atgacagggt cgcaagcgct	9600
ttaacagctc taaatgacat gggaaagatt aggaaagaca tacaacaatg ggaaccttca	9660
agaggatgga atgattggac acaagtgcc ttctgttcac accatttcca tgagttaatc	9720
atgaaagacg gtcgctact cgttggttccc tgtagaaacc aagatgaact gattggcaga	9780
gcccgaatct cccaaggagc aggggtggtct ttgcgggaga cggcctgttt ggggaagtct	9840
tacgccaaa tgtggagctt gatgtacttc cacagacgcg acctcagggt ggcggcaaat	9900
gctatttgct cggcagtacc atcacattgg gttccaacaa gtcgaacaac ctggtccata	9960
catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa cagggtgtgg	10020
attcaagaaa acccatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca	10080
tacttgggga aaagagaaga ccaatggtgc ggctcattga ttgggttaac aagcagggcc	10140
acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa	10200
gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga	10260
gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcagggt cggattaagc	10320
catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca	10380
ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg	10440
tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc	10500
ggttagagga gacccctccc ttacaaatcg cagcaacaat gggggcccaa ggcgagatga	10560
agctgtagtc tcgctggaag gactagaggt tagaggagac ccccccgaac caaaaaacag	10620

catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca 10680
gaacgccaga aaatggaatg gtgctgttga atcaacaggt tct 10723

<210> 11
<211> 3391
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Virut dengue typ huyết thanh 4, BVS

<220>
<221> Xaa
<222> (1226)..(1226)
<223> Lys hoặc Arg

<220>
<221> Xaa
<222> (2481)..(2481)
<223> Ser hoặc Phe

<400> 11

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

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
85 90 95

Arg Arg Arg Ser Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
100 105 110

Met Ala Phe His Leu Thr Thr Arg Asp Gly Glu Pro Leu Met Ile Val
115 120 125

Ala Lys His Glu Arg Gly Arg Pro Leu Leu Phe Lys Thr Thr Glu Gly
130 135 140

Ile Asn Lys Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Met Cys Glu
145 150 155 160

Asp Thr Val Thr Tyr Lys Cys Pro Leu Leu Val Asn Thr Glu Pro Glu
165 170 175

Asp Ile Asp Cys Trp Cys Asn Leu Thr Ser Thr Trp Val Met Tyr Gly
180 185 190

Thr Cys Thr Gln Ser Gly Glu Arg Arg Arg Glu Lys Arg Ser Val Ala
195 200 205

Leu Thr Pro His Ser Gly Met Gly Leu Glu Thr Arg Ala Glu Thr Trp
210 215 220

Met Ser Ser Glu Gly Ala Trp Lys His Ala Gln Arg Val Glu Ser Trp
225 230 235 240

Ile Leu Arg Asn Pro Gly Phe Ala Leu Leu Ala Gly Phe Met Ala Tyr
245 250 255

Met Ile Gly Gln Thr Gly Ile Gln Arg Thr Val Phe Phe Val Leu Met
260 265 270

Met Leu Val Ala Pro Ser Tyr Gly Met Arg Cys Val Gly Val Gly Asn
275 280 285

Arg Asp Phe Val Glu Gly Val Ser Gly Gly Ala Trp Val Asp Leu Val
290 295 300

Leu Glu His Gly Gly Cys Val Thr Thr Met Ala Gln Gly Lys Pro Thr
305 310 315 320

Leu Asp Phe Glu Leu Thr Lys Thr Thr Ala Lys Glu Val Ala Leu Leu
325 330 335

Arg Thr Tyr Cys Ile Glu Ala Ser Ile Ser Asn Ile Thr Thr Ala Thr
340 345 350

Arg Cys Pro Thr Gln Gly Glu Pro Tyr Leu Lys Glu Glu Gln Asp Gln

355 360 365
 Gln Tyr Ile Cys Arg Arg Asp Val Val Asp Arg Gly Trp Gly Asn Gly
 370 375 380
 Cys Gly Leu Phe Gly Lys Gly Gly Val Val Thr Cys Ala Lys Phe Ser
 385 390 395 400
 Cys Ser Gly Lys Ile Thr Gly Asn Leu Val Gln Ile Glu Asn Leu Glu
 405 410 415
 Tyr Thr Val Val Val Thr Val His Asn Gly Asp Thr His Ala Val Gly
 420 425 430
 Asn Asp Thr Ser Asn His Gly Val Thr Ala Thr Ile Thr Pro Arg Ser
 435 440 445
 Pro Ser Val Glu Val Lys Leu Pro Asp Tyr Gly Glu Leu Thr Leu Asp
 450 455 460
 Cys Glu Pro Arg Ser Gly Ile Asp Phe Asn Glu Met Ile Leu Met Lys
 465 470 475 480
 Met Lys Lys Lys Thr Trp Leu Val His Lys Gln Trp Phe Leu Asp Leu
 485 490 495
 Pro Leu Pro Trp Thr Ala Gly Ala Asp Thr Ser Glu Val His Trp Asn
 500 505 510
 Tyr Lys Glu Arg Met Val Thr Phe Lys Val Pro His Ala Lys Arg Gln
 515 520 525
 Asp Val Thr Val Leu Gly Ser Gln Glu Gly Ala Met His Ser Ala Leu
 530 535 540
 Ala Gly Ala Thr Glu Val Asp Ser Gly Asp Gly Asn His Met Phe Ala
 545 550 555 560
 Gly His Leu Lys Cys Lys Val Arg Met Glu Lys Leu Arg Ile Lys Gly
 565 570 575
 Met Ser Tyr Thr Met Cys Ser Gly Lys Phe Ser Ile Asp Lys Glu Met
 580 585 590

Ala Glu Thr Gln His Gly Thr Thr Val Val Lys Val Lys Tyr Glu Gly
595 600 605

Ala Gly Ala Pro Cys Lys Val Pro Ile Glu Ile Arg Asp Val Asn Lys
610 615 620

Glu Lys Val Val Gly Arg Ile Ile Ser Ser Thr Pro Leu Ala Glu Asn
625 630 635 640

Thr Asn Ser Val Thr Asn Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser
645 650 655

Tyr Ile Val Ile Gly Val Gly Asn Ser Ala Leu Thr Leu His Trp Phe
660 665 670

Arg Lys Gly Ser Ser Ile Gly Lys Met Phe Glu Ser Thr Tyr Arg Gly
675 680 685

Ala Lys Arg Met Ala Ile Leu Gly Glu Thr Ala Trp Asp Phe Gly Ser
690 695 700

Val Gly Gly Leu Phe Thr Ser Leu Gly Lys Ala Val His Gln Val Phe
705 710 715 720

Gly Ser Val Tyr Thr Thr Leu Phe Gly Gly Val Ser Trp Met Ile Arg
725 730 735

Ile Leu Ile Gly Phe Leu Val Leu Trp Ile Gly Thr Asn Ser Arg Asn
740 745 750

Thr Ser Met Ala Met Thr Cys Ile Ala Ala Gly Ile Val Thr Leu Tyr
755 760 765

Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
770 775 780

Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
785 790 795 800

His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
805 810 815

Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
820 825 830

Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
835 840 845

Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
850 855 860

Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
865 870 875 880

Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
885 890 895

Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
900 905 910

Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
915 920 925

Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
930 935 940

Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
945 950 955 960

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
965 970 975

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
980 985 990

Ile Glu Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser
995 1000 1005

Asn Gly Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala
1010 1015 1020

Gly Pro Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln
1025 1030 1035

Ile Thr	Gly Pro Trp His	Leu	Gly Lys Leu Glu	Met	Asp Phe Asp
1040		1045		1050	
Phe Cys	Asp Gly Thr Thr	Val	Val Val Thr Glu	Asp	Cys Gly Asn
1055		1060		1065	
Arg Gly	Pro Ser Leu Arg	Thr	Thr Thr Ala Ser	Gly	Lys Leu Ile
1070		1075		1080	
Thr Glu	Trp Cys Cys Arg	Ser	Cys Thr Leu Pro	Pro	Leu Arg Tyr
1085		1090		1095	
Arg Gly	Glu Asp Gly Cys	Trp	Tyr Gly Met Glu	Ile	Arg Pro Leu
1100		1105		1110	
Lys Glu	Lys Glu Glu Asn	Leu	Val Asn Ser Leu	Val	Thr Ala Gly
1115		1120		1125	
His Gly	Gln Val Asp Asn	Phe	Ser Leu Gly Val	Leu	Gly Met Ala
1130		1135		1140	
Leu Phe	Leu Glu Glu Met	Leu	Arg Thr Arg Val	Gly	Thr Lys His
1145		1150		1155	
Ala Ile	Leu Leu Val Ala	Val	Ser Phe Val Thr	Leu	Ile Thr Gly
1160		1165		1170	
Asn Met	Ser Phe Arg Asp	Leu	Gly Arg Val Met	Val	Met Val Gly
1175		1180		1185	
Ala Thr	Met Thr Gly Asp	Ile	Gly Met Gly Val	Thr	Tyr Leu Ala
1190		1195		1200	
Leu Leu	Ala Ala Phe Lys	Val	Arg Pro Thr Phe	Ala	Ala Gly Leu
1205		1210		1215	
Leu Leu	Arg Lys Leu Thr	Ser	Xaa Glu Leu Met	Met	Thr Thr Ile
1220		1225		1230	
Gly Ile	Val Leu Leu Ser	Gln	Ser Thr Ile Pro	Glu	Thr Ile Leu
1235		1240		1245	
Glu Leu	Thr Asp Ala Leu	Ala	Leu Gly Met Met	Val	Leu Lys Met

1250		1255		1260
Val Arg Asn Met Glu Lys Tyr Gln Leu Ala Val Thr Ile Met Ala				
1265		1270		1275
Ile Leu Cys Val Pro Asn Ala Val Ile Leu Gln Asn Ala Trp Lys				
1280		1285		1290
Val Ser Cys Thr Ile Leu Ala Val Val Ser Val Ser Pro Leu Phe				
1295		1300		1305
Leu Thr Ser Ser Gln Gln Lys Thr Asp Trp Ile Pro Leu Ala Leu				
1310		1315		1320
Thr Ile Lys Gly Leu Asn Pro Thr Ala Ile Phe Leu Thr Thr Leu				
1325		1330		1335
Ser Arg Thr Ser Lys Lys Arg Ser Trp Pro Leu Asn Glu Ala Ile				
1340		1345		1350
Met Ala Val Gly Met Val Ser Ile Leu Ala Ser Ser Leu Leu Lys				
1355		1360		1365
Asn Asp Ile Pro Met Thr Gly Pro Leu Val Ala Gly Gly Leu Leu				
1370		1375		1380
Thr Val Cys Tyr Val Leu Thr Gly Arg Ser Ala Asp Leu Glu Leu				
1385		1390		1395
Glu Arg Ala Ala Asp Val Lys Trp Glu Asp Gln Ala Glu Ile Ser				
1400		1405		1410
Gly Ser Ser Pro Ile Leu Ser Ile Thr Ile Ser Glu Asp Gly Ser				
1415		1420		1425
Met Ser Ile Lys Asn Glu Glu Glu Glu Gln Thr Leu Thr Ile Leu				
1430		1435		1440
Ile Arg Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro Val Ser				
1445		1450		1455
Ile Pro Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys Lys				
1460		1465		1470

Gln Arg	Ala Gly Val Leu	Trp	Asp Val Pro Ser	Pro	Pro Pro Met
1475		1480		1485	
Gly Lys	Ala Glu Leu Glu	Asp	Gly Ala Tyr Arg	Ile	Lys Gln Lys
1490		1495		1500	
Gly Ile	Leu Gly Tyr Ser	Gln	Ile Gly Ala Gly	Val	Tyr Lys Glu
1505		1510		1515	
Gly Thr	Phe His Thr Met	Trp	His Val Thr Arg	Gly	Ala Val Leu
1520		1525		1530	
Met His	Lys Gly Lys Arg	Ile	Glu Pro Ser Trp	Ala	Asp Val Lys
1535		1540		1545	
Lys Asp	Leu Ile Ser Tyr	Gly	Gly Gly Trp Lys	Leu	Glu Gly Glu
1550		1555		1560	
Trp Lys	Glu Gly Glu Glu	Val	Gln Val Leu Ala	Leu	Glu Pro Gly
1565		1570		1575	
Lys Asn	Pro Arg Ala Val	Gln	Thr Lys Pro Gly	Leu	Phe Lys Thr
1580		1585		1590	
Asn Ala	Gly Thr Ile Gly	Ala	Val Ser Leu Asp	Phe	Ser Pro Gly
1595		1600		1605	
Thr Ser	Gly Ser Pro Ile	Ile	Asp Lys Lys Gly	Lys	Val Val Gly
1610		1615		1620	
Leu Tyr	Gly Asn Gly Val	Val	Thr Arg Ser Gly	Ala	Tyr Val Ser
1625		1630		1635	
Ala Ile	Ala Gln Thr Glu	Lys	Ser Ile Glu Asp	Asn	Pro Glu Ile
1640		1645		1650	
Glu Asp	Asp Ile Phe Arg	Lys	Arg Arg Leu Thr	Ile	Met Asp Leu
1655		1660		1665	
His Pro	Gly Ala Gly Lys	Thr	Lys Arg Tyr Leu	Pro	Ala Ile Val
1670		1675		1680	

Arg	Glu	Ala	Ile	Lys	Arg	Gly	Leu	Arg	Thr	Leu	Ile	Leu	Ala	Pro
1685						1690					1695			
Thr	Arg	Val	Val	Ala	Ala	Glu	Met	Glu	Glu	Ala	Leu	Arg	Gly	Leu
1700						1705					1710			
Pro	Ile	Arg	Tyr	Gln	Thr	Pro	Ala	Ile	Arg	Ala	Val	His	Thr	Gly
1715						1720					1725			
Arg	Glu	Ile	Val	Asp	Leu	Met	Cys	His	Ala	Thr	Phe	Thr	Met	Arg
1730						1735					1740			
Leu	Leu	Ser	Pro	Val	Arg	Val	Pro	Asn	Tyr	Asn	Leu	Ile	Ile	Met
1745						1750					1755			
Asp	Glu	Ala	His	Phe	Thr	Asp	Pro	Ala	Ser	Ile	Ala	Ala	Arg	Gly
1760						1765					1770			
Tyr	Ile	Ser	Thr	Arg	Val	Glu	Met	Gly	Glu	Ala	Ala	Gly	Ile	Phe
1775						1780					1785			
Met	Thr	Ala	Thr	Pro	Pro	Gly	Ser	Arg	Asp	Pro	Phe	Pro	Gln	Ser
1790						1795					1800			
Asn	Ala	Pro	Ile	Ile	Asp	Glu	Glu	Arg	Glu	Ile	Pro	Glu	Arg	Ser
1805						1810					1815			
Trp	Asn	Ser	Gly	His	Glu	Trp	Val	Thr	Asp	Phe	Lys	Gly	Lys	Thr
1820						1825					1830			
Val	Trp	Phe	Val	Pro	Ser	Ile	Lys	Ala	Gly	Asn	Asp	Ile	Ala	Ala
1835						1840					1845			
Cys	Leu	Arg	Lys	Asn	Gly	Lys	Lys	Val	Ile	Gln	Leu	Ser	Arg	Lys
1850						1855					1860			
Thr	Phe	Asp	Ser	Glu	Tyr	Val	Lys	Thr	Arg	Thr	Asn	Asp	Trp	Asp
1865						1870					1875			
Phe	Val	Val	Thr	Thr	Asp	Ile	Ser	Glu	Met	Gly	Ala	Asn	Phe	Lys
1880						1885					1890			

Ala	Glu	Arg	Val	Ile	Asp	Pro	Arg	Arg	Cys	Met	Lys	Pro	Val	Ile
1895						1900					1905			
Leu	Thr	Asp	Gly	Glu	Glu	Arg	Val	Ile	Leu	Ala	Gly	Pro	Met	Pro
1910						1915					1920			
Val	Thr	His	Ser	Ser	Ala	Ala	Gln	Arg	Arg	Gly	Arg	Ile	Gly	Arg
1925						1930					1935			
Asn	Pro	Lys	Asn	Glu	Asn	Asp	Gln	Tyr	Ile	Tyr	Met	Gly	Glu	Pro
1940						1945					1950			
Leu	Glu	Asn	Asp	Glu	Asp	Cys	Ala	His	Trp	Lys	Glu	Ala	Lys	Met
1955						1960					1965			
Leu	Leu	Asp	Asn	Ile	Asn	Thr	Pro	Glu	Gly	Ile	Ile	Pro	Ser	Met
1970						1975					1980			
Phe	Glu	Pro	Glu	Arg	Glu	Lys	Val	Asp	Ala	Ile	Asp	Gly	Glu	Tyr
1985						1990					1995			
Arg	Leu	Arg	Gly	Glu	Ala	Arg	Lys	Thr	Phe	Val	Asp	Leu	Met	Arg
2000						2005					2010			
Arg	Gly	Asp	Leu	Pro	Val	Trp	Leu	Ala	Tyr	Arg	Val	Ala	Ala	Glu
2015						2020					2025			
Gly	Ile	Asn	Tyr	Ala	Asp	Arg	Arg	Trp	Cys	Phe	Asp	Gly	Val	Lys
2030						2035					2040			
Asn	Asn	Gln	Ile	Leu	Glu	Glu	Asn	Val	Glu	Val	Glu	Ile	Trp	Thr
2045						2050					2055			
Lys	Glu	Gly	Glu	Arg	Lys	Lys	Leu	Lys	Pro	Arg	Trp	Leu	Asp	Ala
2060						2065					2070			
Arg	Ile	Tyr	Ser	Asp	Pro	Leu	Ala	Leu	Lys	Glu	Phe	Lys	Glu	Phe
2075						2080					2085			
Ala	Ala	Gly	Arg	Lys	Ser	Leu	Thr	Leu	Asn	Leu	Ile	Thr	Glu	Met
2090						2095					2100			
Gly	Arg	Leu	Pro	Thr	Phe	Met	Thr	Gln	Lys	Val	Arg	Asp	Ala	Leu

2105	2110	2115
Asp Asn Leu Ala Val Leu His Thr Ala Glu Ala Gly Gly Arg Ala 2120 2125 2130		
Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu 2135 2140 2145		
Leu Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu 2150 2155 2160		
Phe Leu Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met 2165 2170 2175		
Cys Cys Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile 2180 2185 2190		
Gln Pro His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu 2195 2200 2205		
Ile Val Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln 2210 2215 2220		
Asp Asn Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val 2225 2230 2235		
Ala Ala Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys 2240 2245 2250		
Lys Asp Leu Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser 2255 2260 2265		
Asn Ile Leu Asp Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu 2270 2275 2280		
Tyr Ala Val Ala Thr Thr Phe Val Thr Pro Met Leu Arg His Ser 2285 2290 2295		
Ile Glu Asn Ser Ser Val Asn Val Ser Leu Thr Ala Ile Ala Asn 2300 2305 2310		
Gln Ala Thr Val Leu Met Gly Leu Gly Lys Gly Trp Pro Leu Ser 2315 2320 2325		

Lys Met	Asp Ile Gly Val	Pro	Leu Leu Ala Ile	Gly	Cys Tyr Ser
2330		2335		2340	
Gln Val	Asn Pro Ile Thr	Leu	Thr Ala Ala Leu	Phe	Leu Leu Val
2345		2350		2355	
Ala His	Tyr Ala Ile Ile	Gly	Pro Gly Leu Gln	Ala	Lys Ala Thr
2360		2365		2370	
Arg Glu	Ala Gln Lys Arg	Ala	Ala Ala Gly Ile	Met	Lys Asn Pro
2375		2380		2385	
Thr Val	Asp Gly Ile Thr	Val	Ile Asp Leu Asp	Pro	Ile Pro Tyr
2390		2395		2400	
Asp Pro	Lys Phe Glu Lys	Gln	Leu Gly Gln Val	Met	Leu Leu Val
2405		2410		2415	
Leu Cys	Val Thr Gln Val	Leu	Met Met Arg Thr	Thr	Trp Ala Leu
2420		2425		2430	
Cys Glu	Ala Leu Thr Leu	Ala	Thr Gly Pro Ile	Ser	Thr Leu Trp
2435		2440		2445	
Glu Gly	Asn Pro Gly Arg	Phe	Trp Asn Thr Thr	Ile	Ala Val Ser
2450		2455		2460	
Met Ala	Asn Ile Phe Arg	Gly	Ser Tyr Leu Ala	Gly	Ala Gly Leu
2465		2470		2475	
Leu Phe	Xaa Ile Met Lys	Asn	Thr Thr Asn Thr	Arg	Arg Gly Thr
2480		2485		2490	
Gly Asn	Ile Gly Glu Thr	Leu	Gly Glu Lys Trp	Lys	Ser Arg Leu
2495		2500		2505	
Asn Ala	Leu Gly Lys Ser	Glu	Phe Gln Ile Tyr	Lys	Lys Ser Gly
2510		2515		2520	
Ile Gln	Glu Val Asp Arg	Thr	Leu Ala Lys Glu	Gly	Ile Lys Arg
2525		2530		2535	

Gly	Glu	Thr	Asp	His	His	Ala	Val	Ser	Arg	Gly	Ser	Ala	Lys	Leu
2540						2545					2550			
Arg	Trp	Phe	Val	Glu	Arg	Asn	Met	Val	Thr	Pro	Glu	Gly	Lys	Val
2555						2560					2565			
Val	Asp	Leu	Gly	Cys	Gly	Arg	Gly	Gly	Trp	Ser	Tyr	Tyr	Cys	Gly
2570						2575					2580			
Gly	Leu	Lys	Asn	Val	Arg	Glu	Val	Lys	Gly	Leu	Thr	Lys	Gly	Gly
2585						2590					2595			
Pro	Gly	His	Glu	Glu	Pro	Ile	Pro	Met	Ser	Thr	Tyr	Gly	Trp	Asn
2600						2605					2610			
Leu	Val	Arg	Leu	Gln	Ser	Gly	Val	Asp	Val	Phe	Phe	Ile	Pro	Pro
2615						2620					2625			
Glu	Lys	Cys	Asp	Thr	Leu	Leu	Cys	Asp	Ile	Gly	Glu	Ser	Ser	Pro
2630						2635					2640			
Asn	Pro	Thr	Val	Glu	Ala	Gly	Arg	Thr	Leu	Arg	Val	Leu	Asn	Leu
2645						2650					2655			
Val	Glu	Asn	Trp	Leu	Asn	Asn	Asn	Thr	Gln	Phe	Cys	Ile	Lys	Val
2660						2665					2670			
Leu	Asn	Pro	Tyr	Met	Pro	Ser	Val	Ile	Glu	Lys	Met	Glu	Ala	Leu
2675						2680					2685			
Gln	Arg	Lys	Tyr	Gly	Gly	Ala	Leu	Val	Arg	Asn	Pro	Leu	Ser	Arg
2690						2695					2700			
Asn	Ser	Thr	His	Glu	Met	Tyr	Trp	Val	Ser	Asn	Ala	Ser	Gly	Asn
2705						2710					2715			
Ile	Val	Ser	Ser	Val	Asn	Met	Ile	Ser	Arg	Met	Leu	Ile	Asn	Arg
2720						2725					2730			
Phe	Thr	Met	Arg	Tyr	Lys	Lys	Ala	Thr	Tyr	Glu	Pro	Asp	Val	Asp
2735						2740					2745			

Leu Gly 2750	Ser Gly Thr Arg 2755	Asn Ile Gly Ile Glu Ser 2760	Glu Ile Pro
Asn Leu 2765	Asp Ile Ile Gly Lys 2770	Arg Ile Glu Lys Ile 2775	Lys Gln Glu
His Glu 2780	Thr Ser Trp His Tyr 2785	Asp Gln Asp His Pro 2790	Tyr Lys Thr
Trp Ala 2795	Tyr His Gly Ser Tyr 2800	Glu Thr Lys Gln Thr 2805	Gly Ser Ala
Ser Ser 2810	Met Val Asn Gly Val 2815	Val Arg Leu Leu Thr 2820	Lys Pro Trp
Asp Val 2825	Val Pro Met Val Thr 2830	Gln Met Ala Met Thr 2835	Asp Thr Thr
Pro Phe 2840	Gly Gln Gln Arg Val 2845	Phe Lys Glu Lys Val 2850	Asp Thr Arg
Thr Gln 2855	Glu Pro Lys Glu Gly 2860	Thr Lys Lys Leu Met 2865	Lys Ile Thr
Ala Glu 2870	Trp Leu Trp Lys Glu 2875	Leu Gly Lys Lys Lys 2880	Thr Pro Arg
Met Cys 2885	Thr Arg Glu Glu Phe 2890	Thr Arg Lys Val Arg 2895	Ser Asn Ala
Ala Leu 2900	Gly Ala Ile Phe Thr 2905	Asp Glu Asn Lys Trp 2910	Lys Ser Ala
Arg Glu 2915	Ala Val Glu Asp Ser 2920	Arg Phe Trp Glu Leu 2925	Val Asp Lys
Glu Arg 2930	Asn Leu His Leu Glu 2935	Gly Lys Cys Glu Thr 2940	Cys Val Tyr
Asn Met 2945	Met Gly Lys Arg Glu 2950	Lys Lys Leu Gly Glu 2955	Phe Gly Lys
Ala Lys	Gly Ser Arg Ala Ile	Trp Tyr Met Trp Leu	Gly Ala Arg

2960	2965	2970
Phe Leu Glu Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp 2975 2980 2985		
Phe Ser Arg Glu Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu 2990 2995 3000		
His Lys Leu Gly Tyr Ile Leu Arg Asp Val Ser Lys Lys Glu Gly 3005 3010 3015		
Gly Ala Met Tyr Ala Asp Asp Thr Ala Gly Trp Asp Thr Arg Ile 3020 3025 3030		
Thr Leu Glu Asp Leu Lys Asn Glu Glu Met Val Thr Asn His Met 3035 3040 3045		
Glu Gly Glu His Lys Lys Leu Ala Glu Ala Ile Phe Lys Leu Thr 3050 3055 3060		
Tyr Gln Asn Lys Val Val Arg Val Gln Arg Pro Thr Pro Arg Gly 3065 3070 3075		
Thr Val Met Asp Ile Ile Ser Arg Arg Asp Gln Arg Gly Ser Gly 3080 3085 3090		
Gln Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr Asn Met Glu Ala 3095 3100 3105		
Gln Leu Ile Arg Gln Met Glu Gly Glu Gly Val Phe Lys Ser Ile 3110 3115 3120		
Gln His Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn Trp Leu 3125 3130 3135		
Ala Arg Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser Gly 3140 3145 3150		
Asp Asp Cys Val Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala 3155 3160 3165		
Leu Thr Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln 3170 3175 3180		

Gln Trp	Glu Pro Ser Arg	Gly	Trp Asn Asp Trp Thr	Gln Val Pro
3185		3190		3195
Phe Cys	Ser His His Phe His	Glu Leu Ile Met Lys	Asp Gly Arg	
3200		3205	3210	
Val Leu	Val Val Pro Cys Arg	Asn Gln Asp Glu Leu	Ile Gly Arg	
3215		3220	3225	
Ala Arg	Ile Ser Gln Gly Ala	Gly Trp Ser Leu Arg	Glu Thr Ala	
3230		3235	3240	
Cys Leu	Gly Lys Ser Tyr Ala	Gln Met Trp Ser Leu	Met Tyr Phe	
3245		3250	3255	
His Arg	Arg Asp Leu Arg Leu	Ala Ala Asn Ala Ile	Cys Ser Ala	
3260		3265	3270	
Val Pro	Ser His Trp Val Pro	Thr Ser Arg Thr Thr	Trp Ser Ile	
3275		3280	3285	
His Ala	Lys His Glu Trp Met	Thr Thr Glu Asp Met	Leu Thr Val	
3290		3295	3300	
Trp Asn	Arg Val Trp Ile Gln	Glu Asn Pro Trp Met	Glu Asp Lys	
3305		3310	3315	
Thr Pro	Val Glu Ser Trp Glu	Glu Ile Pro Tyr Leu	Gly Lys Arg	
3320		3325	3330	
Glu Asp	Gln Trp Cys Gly Ser	Leu Ile Gly Leu Thr	Ser Arg Ala	
3335		3340	3345	
Thr Trp	Ala Lys Asn Ile Gln	Ala Ala Ile Asn Gln	Val Arg Ser	
3350		3355	3360	
Leu Ile	Gly Asn Glu Glu Tyr	Thr Asp Tyr Met Pro	Ser Met Lys	
3365		3370	3375	
Arg Phe	Arg Arg Glu Glu Glu	Glu Ala Gly Val Leu	Trp	
3380		3385	3390	

<210> 12
 <211> 3391
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Virut dengue typ huyết thanh 4, MVS

<220>
 <221> Xaa
 <222> (1226)..(1226)
 <223> Lys hoặc Arg

<400> 12

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

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
 20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
 35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
 50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
 65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
 85 90 95

Arg Arg Arg Ser Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
 100 105 110

Met Ala Phe His Leu Thr Thr Arg Asp Gly Glu Pro Leu Met Ile Val
 115 120 125

Ala Lys His Glu Arg Gly Arg Pro Leu Leu Phe Lys Thr Thr Glu Gly
 130 135 140

Ile Asn Lys Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Met Cys Glu
 145 150 155 160

Asp Thr Val Thr Tyr Lys Cys Pro Leu Leu Val Asn Thr Glu Pro Glu

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

Cys Ser Gly Lys Ile Thr Gly Asn Leu Val Gln Ile Glu Asn Leu Glu
 405 410 415

Tyr Thr Val Val Val Thr Val His Asn Gly Asp Thr His Ala Val Gly
 420 425 430

Asn Asp Thr Ser Asn His Gly Val Thr Ala Thr Ile Thr Pro Arg Ser
 435 440 445

Pro Ser Val Glu Val Lys Leu Pro Asp Tyr Gly Glu Leu Thr Leu Asp
 450 455 460

Cys Glu Pro Arg Ser Gly Ile Asp Phe Asn Glu Met Ile Leu Met Lys
 465 470 475 480

Met Lys Lys Lys Thr Trp Leu Val His Lys Gln Trp Phe Leu Asp Leu
 485 490 495

Pro Leu Pro Trp Thr Ala Gly Ala Asp Thr Ser Glu Val His Trp Asn
 500 505 510

Tyr Lys Glu Arg Met Val Thr Phe Lys Val Pro His Ala Lys Arg Gln
 515 520 525

Asp Val Thr Val Leu Gly Ser Gln Glu Gly Ala Met His Ser Ala Leu
 530 535 540

Ala Gly Ala Thr Glu Val Asp Ser Gly Asp Gly Asn His Met Phe Ala
 545 550 555 560

Gly His Leu Lys Cys Lys Val Arg Met Glu Lys Leu Arg Ile Lys Gly
 565 570 575

Met Ser Tyr Thr Met Cys Ser Gly Lys Phe Ser Ile Asp Lys Glu Met
 580 585 590

Ala Glu Thr Gln His Gly Thr Thr Val Val Lys Val Lys Tyr Glu Gly
 595 600 605

Ala Gly Ala Pro Cys Lys Val Pro Ile Glu Ile Arg Asp Val Asn Lys
 610 615 620

Glu Lys Val Val Gly Arg Ile Ile Ser Ser Thr Pro Leu Ala Glu Asn
625 630 635 640

Thr Asn Ser Val Thr Asn Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser
645 650 655

Tyr Ile Val Ile Gly Val Gly Asn Ser Ala Leu Thr Leu His Trp Phe
660 665 670

Arg Lys Gly Ser Ser Ile Gly Lys Met Phe Glu Ser Thr Tyr Arg Gly
675 680 685

Ala Lys Arg Met Ala Ile Leu Gly Glu Thr Ala Trp Asp Phe Gly Ser
690 695 700

Val Gly Gly Leu Phe Thr Ser Leu Gly Lys Ala Val His Gln Val Phe
705 710 715 720

Gly Ser Val Tyr Thr Thr Leu Phe Gly Gly Val Ser Trp Met Ile Arg
725 730 735

Ile Leu Ile Gly Phe Leu Val Leu Trp Ile Gly Thr Asn Ser Arg Asn
740 745 750

Thr Ser Met Ala Met Thr Cys Ile Ala Ala Gly Ile Val Thr Leu Tyr
755 760 765

Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
770 775 780

Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
785 790 795 800

His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
805 810 815

Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
820 825 830

Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
835 840 845

Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
 850 855 860

Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
 865 870 875 880

Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
 885 890 895

Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
 900 905 910

Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
 915 920 925

Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
 930 935 940

Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
 945 950 955 960

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
 965 970 975

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
 980 985 990

Ile Glu Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser
 995 1000 1005

Asn Gly Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala
 1010 1015 1020

Gly Pro Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln
 1025 1030 1035

Ile Thr Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp
 1040 1045 1050

Phe Cys Asp Gly Thr Thr Val Val Val Thr Glu Asp Cys Gly Asn
 1055 1060 1065

Arg Gly Pro Ser Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile

1070		1075		1080
Thr Glu Trp Cys Cys Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr				
1085		1090		1095
Arg Gly Glu Asp Gly Cys Trp Tyr Gly Met Glu Ile Arg Pro Leu				
1100		1105		1110
Lys Glu Lys Glu Glu Asn Leu Val Asn Ser Leu Val Thr Ala Gly				
1115		1120		1125
His Gly Gln Val Asp Asn Phe Ser Leu Gly Val Leu Gly Met Ala				
1130		1135		1140
Leu Phe Leu Glu Glu Met Leu Arg Thr Arg Val Gly Thr Lys His				
1145		1150		1155
Ala Ile Leu Leu Val Ala Val Ser Phe Val Thr Leu Ile Thr Gly				
1160		1165		1170
Asn Met Ser Phe Arg Asp Leu Gly Arg Val Met Val Met Val Gly				
1175		1180		1185
Ala Thr Met Thr Gly Asp Ile Gly Met Gly Val Thr Tyr Leu Ala				
1190		1195		1200
Leu Leu Ala Ala Phe Lys Val Arg Pro Thr Phe Ala Ala Gly Leu				
1205		1210		1215
Leu Leu Arg Lys Leu Thr Ser Xaa Glu Leu Met Met Thr Thr Ile				
1220		1225		1230
Gly Ile Val Leu Leu Ser Gln Ser Thr Ile Pro Glu Thr Ile Leu				
1235		1240		1245
Glu Leu Thr Asp Ala Leu Ala Leu Gly Met Met Val Leu Lys Met				
1250		1255		1260
Val Arg Asn Met Glu Lys Tyr Gln Leu Ala Val Thr Ile Met Ala				
1265		1270		1275
Ile Leu Cys Val Pro Asn Ala Val Ile Leu Gln Asn Ala Trp Lys				
1280		1285		1290

Val Ser Cys Thr Ile Leu Ala Val Val Ser Val Ser Pro Leu Phe
1295 1300 1305

Leu Thr Ser Ser Gln Gln Lys Thr Asp Trp Ile Pro Leu Ala Leu
1310 1315 1320

Thr Ile Lys Gly Leu Asn Pro Thr Ala Ile Phe Leu Thr Thr Leu
1325 1330 1335

Ser Arg Thr Ser Lys Lys Arg Ser Trp Pro Leu Asn Glu Ala Ile
1340 1345 1350

Met Ala Val Gly Met Val Ser Ile Leu Ala Ser Ser Leu Leu Lys
1355 1360 1365

Asn Asp Ile Pro Met Thr Gly Pro Leu Val Ala Gly Gly Leu Leu
1370 1375 1380

Thr Val Cys Tyr Val Leu Thr Gly Arg Ser Ala Asp Leu Glu Leu
1385 1390 1395

Glu Arg Ala Ala Asp Val Lys Trp Glu Asp Gln Ala Glu Ile Ser
1400 1405 1410

Gly Ser Ser Pro Ile Leu Ser Ile Thr Ile Ser Glu Asp Gly Ser
1415 1420 1425

Met Ser Ile Lys Asn Glu Glu Glu Glu Gln Thr Leu Thr Ile Leu
1430 1435 1440

Ile Arg Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro Val Ser
1445 1450 1455

Ile Pro Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys Lys
1460 1465 1470

Gln Arg Ala Gly Val Leu Trp Asp Val Pro Ser Pro Pro Pro Met
1475 1480 1485

Gly Lys Ala Glu Leu Glu Asp Gly Ala Tyr Arg Ile Lys Gln Lys
1490 1495 1500

...

Gly	Ile	Leu	Gly	Tyr	Ser	Gln	Ile	Gly	Ala	Gly	Val	Tyr	Lys	Glu
1505						1510					1515			
Gly	Thr	Phe	His	Thr	Met	Trp	His	Val	Thr	Arg	Gly	Ala	Val	Leu
1520						1525					1530			
Met	His	Lys	Gly	Lys	Arg	Ile	Glu	Pro	Ser	Trp	Ala	Asp	Val	Lys
1535						1540					1545			
Lys	Asp	Leu	Ile	Ser	Tyr	Gly	Gly	Gly	Trp	Lys	Leu	Glu	Gly	Glu
1550						1555					1560			
Trp	Lys	Glu	Gly	Glu	Glu	Val	Gln	Val	Leu	Ala	Leu	Glu	Pro	Gly
1565						1570					1575			
Lys	Asn	Pro	Arg	Ala	Val	Gln	Thr	Lys	Pro	Gly	Leu	Phe	Lys	Thr
1580						1585					1590			
Asn	Ala	Gly	Thr	Ile	Gly	Ala	Val	Ser	Leu	Asp	Phe	Ser	Pro	Gly
1595						1600					1605			
Thr	Ser	Gly	Ser	Pro	Ile	Ile	Asp	Lys	Lys	Gly	Lys	Val	Val	Gly
1610						1615					1620			
Leu	Tyr	Gly	Asn	Gly	Val	Val	Thr	Arg	Ser	Gly	Ala	Tyr	Val	Ser
1625						1630					1635			
Ala	Ile	Ala	Gln	Thr	Glu	Lys	Ser	Ile	Glu	Asp	Asn	Pro	Glu	Ile
1640						1645					1650			
Glu	Asp	Asp	Ile	Phe	Arg	Lys	Arg	Arg	Leu	Thr	Ile	Met	Asp	Leu
1655						1660					1665			
His	Pro	Gly	Ala	Gly	Lys	Thr	Lys	Arg	Tyr	Leu	Pro	Ala	Ile	Val
1670						1675					1680			
Arg	Glu	Ala	Ile	Lys	Arg	Gly	Leu	Arg	Thr	Leu	Ile	Leu	Ala	Pro
1685						1690					1695			
Thr	Arg	Val	Val	Ala	Ala	Glu	Met	Glu	Glu	Ala	Leu	Arg	Gly	Leu
1700						1705					1710			

Pro	Ile	Arg	Tyr	Gln	Thr	Pro	Ala	Ile	Arg	Ala	Val	His	Thr	Gly
1715						1720					1725			
Arg	Glu	Ile	Val	Asp	Leu	Met	Cys	His	Ala	Thr	Phe	Thr	Met	Arg
1730						1735					1740			
Leu	Leu	Ser	Pro	Val	Arg	Val	Pro	Asn	Tyr	Asn	Leu	Ile	Ile	Met
1745						1750					1755			
Asp	Glu	Ala	His	Phe	Thr	Asp	Pro	Ala	Ser	Ile	Ala	Ala	Arg	Gly
1760						1765					1770			
Tyr	Ile	Ser	Thr	Arg	Val	Glu	Met	Gly	Glu	Ala	Ala	Gly	Ile	Phe
1775						1780					1785			
Met	Thr	Ala	Thr	Pro	Pro	Gly	Ser	Arg	Asp	Pro	Phe	Pro	Gln	Ser
1790						1795					1800			
Asn	Ala	Pro	Ile	Ile	Asp	Glu	Glu	Arg	Glu	Ile	Pro	Glu	Arg	Ser
1805						1810					1815			
Trp	Asn	Ser	Gly	His	Glu	Trp	Val	Thr	Asp	Phe	Lys	Gly	Lys	Thr
1820						1825					1830			
Val	Trp	Phe	Val	Pro	Ser	Ile	Lys	Ala	Gly	Asn	Asp	Ile	Ala	Ala
1835						1840					1845			
Cys	Leu	Arg	Lys	Asn	Gly	Lys	Lys	Val	Ile	Gln	Leu	Ser	Arg	Lys
1850						1855					1860			
Thr	Phe	Asp	Ser	Glu	Tyr	Val	Lys	Thr	Arg	Thr	Asn	Asp	Trp	Asp
1865						1870					1875			
Phe	Val	Val	Thr	Thr	Asp	Ile	Ser	Glu	Met	Gly	Ala	Asn	Phe	Lys
1880						1885					1890			
Ala	Glu	Arg	Val	Ile	Asp	Pro	Arg	Arg	Cys	Met	Lys	Pro	Val	Ile
1895						1900					1905			
Leu	Thr	Asp	Gly	Glu	Glu	Arg	Val	Ile	Leu	Ala	Gly	Pro	Met	Pro
1910						1915					1920			
Val	Thr	His	Ser	Ser	Ala	Ala	Gln	Arg	Arg	Gly	Arg	Ile	Gly	Arg

1925		1930		1935
Asn Pro Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu Pro 1940 1945 1950				
Leu Glu Asn Asp Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met 1955 1960 1965				
Leu Leu Asp Asn Ile Asn Thr Pro Glu Gly Ile Ile Pro Ser Met 1970 1975 1980				
Phe Glu Pro Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr 1985 1990 1995				
Arg Leu Arg Gly Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg 2000 2005 2010				
Arg Gly Asp Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu 2015 2020 2025				
Gly Ile Asn Tyr Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys 2030 2035 2040				
Asn Asn Gln Ile Leu Glu Glu Asn Val Glu Val Glu Ile Trp Thr 2045 2050 2055				
Lys Glu Gly Glu Arg Lys Lys Leu Lys Pro Arg Trp Leu Asp Ala 2060 2065 2070				
Arg Ile Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe Lys Glu Phe 2075 2080 2085				
Ala Ala Gly Arg Lys Ser Leu Thr Leu Asn Leu Ile Thr Glu Met 2090 2095 2100				
Gly Arg Leu Pro Thr Phe Met Thr Gln Lys Val Arg Asp Ala Leu 2105 2110 2115				
Asp Asn Leu Ala Val Leu His Thr Ala Glu Ala Gly Gly Arg Ala 2120 2125 2130				
Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu 2135 2140 2145				

Leu	Leu	Leu	Thr	Leu	Leu	Ala	Thr	Val	Thr	Gly	Gly	Ile	Phe	Leu
2150						2155					2160			
Phe	Leu	Met	Ser	Ala	Arg	Gly	Ile	Gly	Lys	Met	Thr	Leu	Gly	Met
2165						2170					2175			
Cys	Cys	Ile	Ile	Thr	Ala	Ser	Ile	Leu	Leu	Trp	Tyr	Ala	Gln	Ile
2180						2185					2190			
Gln	Pro	His	Trp	Ile	Ala	Ala	Ser	Ile	Ile	Leu	Glu	Phe	Phe	Leu
2195						2200					2205			
Ile	Val	Leu	Leu	Ile	Pro	Glu	Pro	Glu	Lys	Gln	Arg	Thr	Pro	Gln
2210						2215					2220			
Asp	Asn	Gln	Leu	Thr	Tyr	Val	Val	Ile	Ala	Ile	Leu	Thr	Val	Val
2225						2230					2235			
Ala	Ala	Thr	Met	Ala	Asn	Glu	Met	Gly	Phe	Leu	Glu	Lys	Thr	Lys
2240						2245					2250			
Lys	Asp	Leu	Gly	Leu	Gly	Ser	Ile	Ala	Thr	Gln	Gln	Pro	Glu	Ser
2255						2260					2265			
Asn	Ile	Leu	Asp	Ile	Asp	Leu	Arg	Pro	Ala	Ser	Ala	Trp	Thr	Leu
2270						2275					2280			
Tyr	Ala	Val	Ala	Thr	Thr	Phe	Val	Thr	Pro	Met	Leu	Arg	His	Ser
2285						2290					2295			
Ile	Glu	Asn	Ser	Ser	Val	Asn	Val	Ser	Leu	Thr	Ala	Ile	Ala	Asn
2300						2305					2310			
Gln	Ala	Thr	Val	Leu	Met	Gly	Leu	Gly	Lys	Gly	Trp	Pro	Leu	Ser
2315						2320					2325			
Lys	Met	Asp	Ile	Gly	Val	Pro	Leu	Leu	Ala	Ile	Gly	Cys	Tyr	Ser
2330						2335					2340			
Gln	Val	Asn	Pro	Ile	Thr	Leu	Thr	Ala	Ala	Leu	Phe	Leu	Leu	Val
2345						2350					2355			

Ala	His	Tyr	Ala	Ile	Ile	Gly	Pro	Gly	Leu	Gln	Ala	Lys	Ala	Thr
2360						2365					2370			
Arg	Glu	Ala	Gln	Lys	Arg	Ala	Ala	Ala	Gly	Ile	Met	Lys	Asn	Pro
2375						2380					2385			
Thr	Val	Asp	Gly	Ile	Thr	Val	Ile	Asp	Leu	Asp	Pro	Ile	Pro	Tyr
2390						2395					2400			
Asp	Pro	Lys	Phe	Glu	Lys	Gln	Leu	Gly	Gln	Val	Met	Leu	Leu	Val
2405						2410					2415			
Leu	Cys	Val	Thr	Gln	Val	Leu	Met	Met	Arg	Thr	Thr	Trp	Ala	Leu
2420						2425					2430			
Cys	Glu	Ala	Leu	Thr	Leu	Ala	Thr	Gly	Pro	Ile	Ser	Thr	Leu	Trp
2435						2440					2445			
Glu	Gly	Asn	Pro	Gly	Arg	Phe	Trp	Asn	Thr	Thr	Ile	Ala	Val	Ser
2450						2455					2460			
Met	Ala	Asn	Ile	Phe	Arg	Gly	Ser	Tyr	Leu	Ala	Gly	Ala	Gly	Leu
2465						2470					2475			
Leu	Phe	Ser	Ile	Met	Lys	Asn	Thr	Thr	Asn	Thr	Arg	Arg	Gly	Thr
2480						2485					2490			
Gly	Asn	Ile	Gly	Glu	Thr	Leu	Gly	Glu	Lys	Trp	Lys	Ser	Arg	Leu
2495						2500					2505			
Asn	Ala	Leu	Gly	Lys	Ser	Glu	Phe	Gln	Ile	Tyr	Lys	Lys	Ser	Gly
2510						2515					2520			
Ile	Gln	Glu	Val	Asp	Arg	Thr	Leu	Ala	Lys	Glu	Gly	Ile	Lys	Arg
2525						2530					2535			
Gly	Glu	Thr	Asp	His	His	Ala	Val	Ser	Arg	Gly	Ser	Ala	Lys	Leu
2540						2545					2550			
Arg	Trp	Phe	Val	Glu	Arg	Asn	Met	Val	Thr	Pro	Glu	Gly	Lys	Val
2555						2560					2565			

Val Asp 2570	Leu Gly Cys Gly Arg 2575	Gly Gly Trp Ser Tyr 2580	Tyr Cys Gly
Gly Leu 2585	Lys Asn Val Arg Glu 2590	Val Lys Gly Leu Thr 2595	Lys Gly Gly
Pro Gly 2600	His Glu Glu Pro Ile 2605	Pro Met Ser Thr Tyr 2610	Gly Trp Asn
Leu Val 2615	Arg Leu Gln Ser Gly 2620	Val Asp Val Phe Phe 2625	Ile Pro Pro
Glu Lys 2630	Cys Asp Thr Leu Leu 2635	Cys Asp Ile Gly Glu 2640	Ser Ser Pro
Asn Pro 2645	Thr Val Glu Ala Gly 2650	Arg Thr Leu Arg Val 2655	Leu Asn Leu
Val Glu 2660	Asn Trp Leu Asn Asn 2665	Asn Thr Gln Phe Cys 2670	Ile Lys Val
Leu Asn 2675	Pro Tyr Met Pro Ser 2680	Val Ile Glu Lys Met 2685	Glu Ala Leu
Gln Arg 2690	Lys Tyr Gly Gly Ala 2695	Leu Val Arg Asn Pro 2700	Leu Ser Arg
Asn Ser 2705	Thr His Glu Met Tyr 2710	Trp Val Ser Asn Ala 2715	Ser Gly Asn
Ile Val 2720	Ser Ser Val Asn Met 2725	Ile Ser Arg Met Leu 2730	Ile Asn Arg
Phe Thr 2735	Met Arg Tyr Lys Lys 2740	Ala Thr Tyr Glu Pro 2745	Asp Val Asp
Leu Gly 2750	Ser Gly Thr Arg Asn 2755	Ile Gly Ile Glu Ser 2760	Glu Ile Pro
Asn Leu 2765	Asp Ile Ile Gly Lys 2770	Arg Ile Glu Lys Ile 2775	Lys Gln Glu
His Glu	Thr Ser Trp His Tyr	Asp Gln Asp His Pro	Tyr Lys Thr

2780		2785		2790
Trp Ala Tyr His Gly Ser Tyr Glu Thr Lys Gln Thr Gly Ser Ala				
2795		2800		2805
Ser Ser Met Val Asn Gly Val Val Arg Leu Leu Thr Lys Pro Trp				
2810		2815		2820
Asp Val Val Pro Met Val Thr Gln Met Ala Met Thr Asp Thr Thr				
2825		2830		2835
Pro Phe Gly Gln Gln Arg Val Phe Lys Glu Lys Val Asp Thr Arg				
2840		2845		2850
Thr Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu Met Lys Ile Thr				
2855		2860		2865
Ala Glu Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys Thr Pro Arg				
2870		2875		2880
Met Cys Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser Asn Ala				
2885		2890		2895
Ala Leu Gly Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser Ala				
2900		2905		2910
Arg Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys				
2915		2920		2925
Glu Arg Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr				
2930		2935		2940
Asn Met Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys				
2945		2950		2955
Ala Lys Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg				
2960		2965		2970
Phe Leu Glu Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp				
2975		2980		2985
Phe Ser Arg Glu Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu				
2990		2995		3000

His Lys	Leu Gly Tyr Ile	Leu	Arg Asp Val Ser	Lys	Lys Glu Gly
3005		3010		3015	
Gly Ala	Met Tyr Ala Asp	Asp	Thr Ala Gly Trp	Asp	Thr Arg Ile
3020		3025		3030	
Thr Leu	Glu Asp Leu Lys	Asn	Glu Glu Met Val	Thr	Asn His Met
3035		3040		3045	
Glu Gly	Glu His Lys Lys	Leu	Ala Glu Ala Ile	Phe	Lys Leu Thr
3050		3055		3060	
Tyr Gln	Asn Lys Val Val	Arg	Val Gln Arg Pro	Thr	Pro Arg Gly
3065		3070		3075	
Thr Val	Met Asp Ile Ile	Ser	Arg Arg Asp Gln	Arg	Gly Ser Gly
3080		3085		3090	
Gln Val	Gly Thr Tyr Gly	Leu	Asn Thr Phe Thr	Asn	Met Glu Ala
3095		3100		3105	
Gln Leu	Ile Arg Gln Met	Glu	Gly Glu Gly Val	Phe	Lys Ser Ile
3110		3115		3120	
Gln His	Leu Thr Ile Thr	Glu	Glu Ile Ala Val	Gln	Asn Trp Leu
3125		3130		3135	
Ala Arg	Val Gly Arg Glu	Arg	Leu Ser Arg Met	Ala	Ile Ser Gly
3140		3145		3150	
Asp Asp	Cys Val Val Lys	Pro	Leu Asp Asp Arg	Phe	Ala Ser Ala
3155		3160		3165	
Leu Thr	Ala Leu Asn Asp	Met	Gly Lys Ile Arg	Lys	Asp Ile Gln
3170		3175		3180	
Gln Trp	Glu Pro Ser Arg	Gly	Trp Asn Asp Trp	Thr	Gln Val Pro
3185		3190		3195	
Phe Cys	Ser His His Phe	His	Glu Leu Ile Met	Lys	Asp Gly Arg
3200		3205		3210	

Val Leu Val Val Pro Cys Arg Asn Gln Asp Glu Leu Ile Gly Arg
3215 3220 3225

Ala Arg Ile Ser Gln Gly Ala Gly Trp Ser Leu Arg Glu Thr Ala
3230 3235 3240

Cys Leu Gly Lys Ser Tyr Ala Gln Met Trp Ser Leu Met Tyr Phe
3245 3250 3255

His Arg Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile Cys Ser Ala
3260 3265 3270

Val Pro Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp Ser Ile
3275 3280 3285

His Ala Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr Val
3290 3295 3300

Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys
3305 3310 3315

Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg
3320 3325 3330

Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala
3335 3340 3345

Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser
3350 3355 3360

Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys
3365 3370 3375

Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
3380 3385 3390

<210> 13
<211> 10723
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Virut dengue typ huyết thanh 1, MVS

<400> 13
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60
gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120
aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180
ctgacaaaga gattctcact tggaatgctg cagggacgag gaccattaaa actgttcatg 240
gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300
tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360
ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat cattatgctg 420
attccaacag tgatggcggt ccatttaacc acgcgtgggg gagagccgca tatgatagtt 480
agcaagcagg aaagaggaaa gtcacttttg ttcaagacct ctgcagggtg caacatgtgc 540
accctcattg cgatggattt gggagagttg tgtgaggaca cgatgacct acaatgcccc 600
cggatcactg aggcggaacc agatgacgtt gactgttggg gcaatgccac ggacacatgg 660
gtgacctatg gaacgtgctc tcaaactggc gaacaccgac gagacaaacg ttccgtcgca 720
ttggccccac acgtggggct tggcctagaa acaagagccg aaacgtggat gtcctctgaa 780
ggtgcttgga aacagataca aaaagtagag acttgggctc tgagacatcc aggattcacg 840
gtgatagccc tttttctagc acatgccata ggaacatcca tcaccagaa agggatcatt 900
ttcattttgc tgatgctggt aacaccatct atggccatgc gatgcgtggg aataggcaac 960
agagacttcg tggaaggact gtcaggagca acatgggtgg atgtggtact ggagcatgga 1020
agttgcgtca ccaccatggc aaaaaacaaa ccaacactgg acattgaact cttgaagacg 1080
gaggtcacia accctgcagt tctgcgtaaa ttgtgcattg aagctaaaat atcaaacc 1140
accaccgatt cgagatgtcc aacacaagga gaagccacac tgggtggaaga acaagacgcg 1200
aactttgtgt gccgacgaac gttcgtggac agaggctggg gcaatggctg tgggctattc 1260
ggaaaaggta gtctaataac gtgtgccaag ttttaagtgtg tgacaaaact agaaggaaag 1320
atagttcaat atgaaaacct aaaatattca gtgatagtca ccgtccacac tggagatcag 1380
caccagggtg gaaatgagac tacagaacat ggaacaactg caaccataac acctcaagct 1440
cctacgtcgg aaatacagct gaccgactac ggaaccctta cattagattg ttcacctagg 1500
acagggctag attttaacga gatggtgttg ctgacaatga aagaaagatc atggcttgctc 1560
cacaacaat ggttcctaga cttaccactg ccttggacct ctggggcttc aacatcccaa 1620
gagacttgga acagacaaga ttactggctc acatttaaga cagctcatgc aaagaagcag 1680

gaagtagtcg tactaggatc acaagaagga gcaatgcaca ctgcgctgac tggagcgaca	1740
gaaatccaaa cgtcaggaac gacaacaatt ttgcgaggac acctaaaatg cagactaaaa	1800
atggacaaac taacttttaa agggatgtca tatgtgatgt gcacaggctc attcaagtta	1860
gagaaagaag tggctgagac ccagcatgga actgttcttg tgcaggttaa atatgaagga	1920
acagacgcac catgcaagat tcccttttcg acccaagatg agaaaggagc aaccagaat	1980
gggagattaa taacagccaa ccccatagtc actgacaaag aaaaaccagt caatattgag	2040
gcagaaccac cctttggtga gagctacatc gtggtaggag caggtgaaaa agctttgaaa	2100
ctaagctggt tcaagaaagg aagcagcata gggaaaatgt ttgaagcaac tgcccagga	2160
gcacgaagga tggccattct gggagacacc gcatgggact tcggttctat aggaggagt	2220
ttcacgtcta tgggaaaact ggtacaccag gtttttgga ctgcatatgg agttttgttt	2280
agcggagttt cttggaccat gaaaatagga atagggattc tgctgacatg gctaggatta	2340
aattcaagga acacgtccct ttogatgatg tgcacgcag ccggcattgt gacactgtat	2400
ttgggagtca tgggtgcaggc cgatagtggg tgcgttgga gctggaaaaa caaagaactg	2460
aaatgtggca gtgggatttt catcacagac aacgtgcaca catggacaga acaatacaag	2520
ttccaaccag aatccccttc aaaactagct tcagctatcc agaaagccca tgaagaggac	2580
atltgtggaa tccgctcagt aacaagactg gagaatctga tgtggaaaca aataacacca	2640
gaattgaatc acattctatc agaaaatgag gtgaagttaa ctattatgac aggagacatc	2700
aaaggaatca tgcaggcagg aaaacgatct ctgcggcctc agccactga gctgaagtat	2760
tcattggaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt	2820
ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg	2880
gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa	2940
aaacaggatg tattctgcga ctcaaaactc atgtcagcgg ccataaaaga caacagagcc	3000
gtccatgccg atatgggtta ttggatagaa agtgactca atgacacatg gaagatagag	3060
aaagcctctt tcattgaagt taaaaactgc cactggccaa aatcacacac cctctggagc	3120
aatggagtgc tagaaagtga gatgataatt ccaaagaatc tcgctggacc agtgtctcaa	3180
cacaactata gaccaggcta ccatacacia ataacaggac catggcatct aggttaagctt	3240
gagatggact ttgatttctg tgatggaaca acagtggtag tgactgagga ctgcggaaat	3300
agaggaccct ctttgagaac aaccactgcc tctggaaaac tcataacaga atgggtgctgc	3360
cgatcttgca cattaccacc gctaagatac agaggtgagg atgggtgctg gtacgggatg	3420

gaaatcagac cattgaagga gaaagaagag aatttgggtca actccttgggt cacagctgga	3480
catgggcagg tcgacaactt ttcactagga gtcttgggaa tggcattgtt cctggaggaa	3540
atgcttagga cccgagtagg aacgaaacat gcaatactac tagttgcagt ttcttttgtg	3600
acattgatca cagggaacat gtcctttaga gacctgggaa gagtgatgggt tatggtaggc	3660
gccactatga cggatgacat aggtatgggc gtgacttattc ttgccctact agcagccttc	3720
aaagtcagac caacttttgc agctggacta ctcttgagaa agctgacctc caaggaattg	3780
atgatgacta ctataggaat tgtactcttc tcccagagca ccctaccaga gaccattctt	3840
gagttgactg atgcgtagc cttaggcatg atggctctca aaatggtgag aaatatggaa	3900
aagtatcaat tggcagtgac tatcatggct atcttgtgcg tcccaaacgc agtgatatta	3960
caaaacgcat ggaaagtgag ttgcacaata ttggcagtggt tgtccgtttc cccactgttc	4020
ttaacatcct cacagcaaaa aacagattgg ataccattag cattgacgat caaaggcttc	4080
aatccaacag ctatttttct aacaaccctc tcaagaacca gcaagaaaag gagctggcca	4140
ttaaatgagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa	4200
aatgatattc ccatgacagg accattagtg gctggagggc tcctcactgt gtgctacgtg	4260
ctcactggac gatcggccga tttggaactg gagagagcag ccgatgtcaa atgggaagac	4320
caggcagaga tatcaggaag cagtccaatc ctgtcaataa caatatcaga agatggtagc	4380
atgtcgataa aaaatgaaga ggaagatcaa aactgacca tactcattag aacaggattg	4440
ctgggtgatct caggactttt tcctgtatca ataccaatca cggcagcagc atggtacctg	4500
tgggaagtga agaaacaacg ggccggagta ttgtgggatg ttcccttcacc cccacccatg	4560
ggaaaggctg aactggaaga tggagcctat agaattaagc aaaaagggat tcttgatat	4620
tcccagatcg gagccggagt ttacaaagaa ggaacattcc atacaatgtg gcatgtcaca	4680
cgtggcgctg ttctaattgca taaaggaaaag aggattgaac catcatgggc ggacgtcaag	4740
aaagacctaa tatcatatgg aggaggctgg aagttagaag gagaatggaa ggaaggagaa	4800
gaagtccagg tattggcact ggagcctgga aaaaatccaa gagccgtcca aacgaaacct	4860
ggtcttttca aaaccaacgc cggaacaata ggtgctgtat ctctggactt ttctcctgga	4920
acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt	4980
gttgttacaa ggagtggagc atatgtgagt gctatagccc agactgaaaa aagcattgaa	5040
gacaaccag agatcgaaga tgacattttc cgaaagagaa gactgaccat catggacctc	5100

cacccaggag cgaggaaagac gaagagatac cttccggcca tagtcagaga agctataaaa	5160
cgggggtttga gaacattaat cttggccccc actagagttg tggcagctga aatggaggaa	5220
gcccttagag gacttccaat aagataccag accccagcca tcagagctgt gcacaccggg	5280
cgggagattg tggacctaata gtgtcatgcc acatttacca tgaggctgct atcaccagtt	5340
agagtgccaa actacaacct gattatcatg gacgaagccc atttcacaga cccagcaagt	5400
atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt	5460
atgacagcca ctcccccggg aagcagagac ccatttcctc agagcaatgc accaatcata	5520
gatgaagaaa gagaaatccc tgaacgctcg tggaattccg gacatgaatg ggtcacggat	5580
tttaaaggga agactgtttg gttcgttcca agtataaaag caggaaatga tatagcagct	5640
tgcctgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag	5700
tatgtcaaga ctagaaccaa tgattgggac ttcgtggtta caactgacat ttcagaaatg	5760
ggtgccaat tcaaggctga gaggggtata gacccagac gctgcatgaa accagtcata	5820
ctaacagatg gtgaagagcg ggtgattctg gcaggaccta tgccagtga ccaactctagt	5880
gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata	5940
tacatggggg aacctctgga aaatgatgaa gactgtgcac actggaaaga agctaaaatg	6000
ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt	6060
gaaaagggtg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt	6120
gtagacttaa tgagaagagg agacctacca gtctgggttg cctacagagt ggcagctgaa	6180
ggcatcaact acgcagacag aagggtggtg tttgatggag tcaagaacaa ccaaactcta	6240
gaagaaaacg tggaagtga aatctggaca aaagaagggg aaaggaagaa attgaaaccc	6300
agatgggttg atgctaggat ctattctgac cactggcgc taaaagaatt taaggaattt	6360
gcagccggaa gaaagtctct gacctgaac ctaatcacag aaatgggtag gctccaacc	6420
ttcatgactc agaaggcaag agacgactg gacaacttag cagtgtgca cagggtgag	6480
gcaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg	6540
cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca	6600
aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta	6660
tggtacgcac aaatacagcc aactggata gcagcttcaa taatactgga gttttttctc	6720
atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc	6780
tacgttgtca tagccatcct cacagtgggt gccgcaacca tggcaaacga gatgggtttc	6840

ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caaccagca acccgagagc	6900
aacatcctgg acatagatct acgtcctgca tcagcatgga cgctgtatgc cgtggccaca	6960
acatttggtta caccaatggt gagacatagc attgaaaatt cctcagtga tgtgtcccta	7020
acagctatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca	7080
aagatggaca tcggagttcc ctttctcgcc attggatgct actcacaagt caaccata	7140
actctcacag cagctctttt cttattggtta gcacattatg ccatcatagg gccaggactc	7200
caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaacca	7260
actgtcgatg gaataacagt gattgacctt gatccaatac cttatgatcc gaagtttgaa	7320
aagcagttgg gacaagtaat gctcctagtc ctctgcgtga ctcaagtatt gatgatgagg	7380
actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg	7440
gaaggaaatc cagggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt	7500
agagggagtt acttggccgg agctggactt ctcttttcta ttatgaagaa cacaaccaac	7560
acaagaaggg gaactggcaa cataggagag acgcttgagg agaaatggaa aagccgattg	7620
aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat	7680
agaaccttag caaaagaagg cattaanaaga ggagaaacgg accatcacgc tgtgtcgcga	7740
ggctcagcaa aactgagatg gttcgttgag agaaacatgg tcacaccaga agggaaagta	7800
gtggacctcg gttgtggcag aggaggctgg tcatactatt gtggaggact aaagaatgta	7860
agagaagtca aaggcctaac aaaaggagga ccaggacacg aagaacccat ccccatgtca	7920
acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcca	7980
gaaaagtgtg acacattatt gtgtgacata ggggagtcac caccaaatcc cacagtggaa	8040
gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa	8100
ttttgcataa aggttctcaa cccatatatg ccctcagtca tagaaaaaat ggaagcacta	8160
caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag	8220
atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg	8280
atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac	8340
ctcggaagcg gaaccgtaa catcgggatt gaaagtgaga taccaaact agatataatt	8400
gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac	8460
caccatata aaacgtgggc ataccatggt agctatgaaa caaaacagac tggatcagca	8520

tcatccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtcccatg	8580
gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgcgt ttttaaagag	8640
aaagtggaca cgagaaccca agaaccgaaa gaaggcacga agaaactaat gaaaataaca	8700
gcagagtggc tttggaaaga attaggggaag aaaaagacac ccaggatgtg caccagagaa	8760
gaattcacaa gaaaggtgag aagcaatgca gccttggggg ccatattcac tgatgagaac	8820
aagtggaagt cggcacgtga ggctgttgaa gatagtaggt tttgggagct ggttgacaag	8880
gaaaggaatc tccatcttga aggaaagtgt gaaacatgtg tgtacaacat gatgggaaaa	8940
agagagaaga agctagggga attcggcaag gcaaaaggca gcagagccat atggtacatg	9000
tggcttgag cagccttctt agagtgtgaa gccctaggat tcttaaatga agatcactgg	9060
ttctccagag agaactccct gagtggagtg gaaggagaag ggctgcacaa gctaggttac	9120
attctaagag acgtgagcaa gaaagaggga ggagcaatgt atgccgatga caccgcagga	9180
tgggatacaa gaatcacact agaagacctt aaaaatgaag aaatggtaac aaaccacatg	9240
gaaggagaac acaagaaact agccgaggcc attttcaaac taacgtacca aaacaagggtg	9300
gtgcgtgtgc aaagaccaac accaagaggc acagtaatgg acatcatatc gagaagagac	9360
caaagaggta gtggacaagt tggcacctat ggactcaata ctttcaccaa tatggaagcc	9420
caactaatca gacagatgga gggagaagga gtcttttaaaa gcattcagca cctaacaatc	9480
acagaagaaa tcgctgtgca aaactgggta gcaagagtgg ggcgcgaaag gttatcaaga	9540
atggccatca gtggagatga ttgtgttggtg aaaccttttag atgacagggt cgcaagcgct	9600
ttaacagctc taaatgacat gggaaagatt aggaaagaca tacaacaatg ggaaccttca	9660
agaggatgga atgattggac acaagtggcc ttctgttcac accatttcca tgagttaatc	9720
atgaaagacg gtcgcgtact cgttgttcca tgtagaaacc aagatgaact gattggcaga	9780
gcccgaatct cccaaggagc aggggtgtct ttgcgggaga cggcctgttt ggggaagtct	9840
tacgccccaa tgtggagctt gatgtacttc cacagacggg acctcaggct ggcggcaaat	9900
gctatttgct cggcagtacc atcacattgg gttccaacaa gtcgaacaac ctggtccata	9960
catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa cagggtgtgg	10020
attcaagaaa acccatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca	10080
tacttgggga aaagagaaga ccaatggtgc ggctcattga ttgggttaac aagcagggcc	10140
acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa	10200
gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga	10260

gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320
catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380
ggccatcata aatgccatag cttgagtaaa ctatgcagcc ttagctcca cctgagaagg 10440
tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500
ggttagagga gaccctccc ttacaaatcg cagcaacaat gggggcccaa ggcgagatga 10560
agctgtagtc tcgctggaag gactagaggt tagaggagac cccccgaaa caaaaaacag 10620
catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca 10680
gaacgccaga aaatggaatg gtgctgttga atcaacaggt tct 10723

<210> 14

<211> 10723

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Virut dengue typ huyết thanh 2, MVS

<400> 14

agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60
gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120
aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180
ctgacaaaga gattctcact tggaatgctg caggagcag gaccattaaa actgttcattg 240
gccctggtgg cgctccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300
tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360
ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat cattatgctg 420
attccaacag tgatggcggt ccatttaacc acacgtaacg gagaaccaca catgatcgtc 480
agcagacaag agaaagggaa aagtcttctg tttaaaacag aggttggcgt gaacatgtgt 540
accctcatgg ccatggacct tggatgaattg tgtgaagaca caatcacgta cgagtgtccc 600
cttctcaggc agaatgagcc agaagacata gactgttggg gcaactctac gtccacgtgg 660
gtaacttatg ggacgtgtac caccatggga gaacatagaa gagaaaaaag atcagtggca 720
ctcgttccac atgtgggaat gggactggag acacgaactg aaacatggat gtcacagaa 780
ggggcctgga aacatgtcca gagaattgaa acttggatct tgagacatcc aggcttcacc 840
atgatggcag caatcctggc atacaccata ggaacgacac atttccaaag agccctgatc 900

ttcatcttac tgacagctgt cactccttca atgacaatgc gttgcatagg aatgtcaa	960
agagactttg tggaaggggt ttcaggagga agctgggttg acatagtctt aga	1020
agctgtgtga cgacgatggc aaaaaacaaa ccaacattgg attttgaact gata	1080
gaagccaaac agcctgccac cctaaggaag tactgtatag aggcaaagct aac	1140
acaacagaat ctgctgccc aacacaaggg gaaccagcc taaatgaaga gcagg	1200
aggttcgtct gcaaacactc catggtagac agaggatggg gaaatggatg tgg	1260
ggaaaggag gcattgtgac ctgtgctatg ttcagatgca aaaagaacat gga	1320
gttgtgcaac cagaaaactt ggaatacacc attgtgataa cacctcactc ag	1380
catgcagtcg gaaatgacac aggaaaacat ggcaaggaaa tcaaaataac acc	1440
tccatcacag aagcagaatt gacaggttat ggcaactgtca caatggagtg ct	1500
acgggcctcg acttcaatga gatgggtgtg ctgcagatgg aaaataaagc ttg	1560
cacaggcaat ggttcctaga cctgccgtta ccatggttgc cggagcgga cac	1620
tcaaattgga tacagaaaga gacattggtc actttcaaaa atcccatgc ga	1680
gatgttgttg ttttaggatc ccaagaaggg gccatgcaca cagcacttac ag	1740
gaaatccaaa tgtcatcagg aaacttactc ttcacaggac atctcaagtg cag	1800
atggacaagc tacagctcaa aggaatgtca tactctatgt gcacaggaaa g	1860
gtgaaggaaa tagcagaaac acaacatgga acaatagtta tcagagtgca at	1920
gacggctctc catgcaagat cccttttgag ataatggatt tggaaaaaag ac	1980
ggtcgctga ttacagtcaa cccaattgtg acagaaaaag atagcccagt ca	2040
gcagaacctc catttgaga cagctacatc atcataggag tagagccggg aca	2100
ctcaactggg ttaagaaagg aagttctatc ggccaaatgt ttgagacaac a	2160
gcgaagagaa tggccatttt aggtgacaca gcctgggatt ttggatcctt g	2220
tttacatcta taggaaaggc tctccaccaa gtctttggag caatctatgg ag	2280
agtggggttt catggactat gaaaatcctc ataggagtca ttatcacatg ga	2340
aattcacgca gcacctcact gtctgtgaca ctagtattgg tgggaattgt ga	2400
ttgggagtca tgggtgcaggc cgatagtggg tgcgttgtga gctggaaaaa ca	2460
aatgtggca gtgggatttt catcacagac aacgtgcaca catggacaga aca	2520
ttccaaccag aatccccttc aaaactagct tcagctatcc agaaagcca tga	2580
atttgtggaa tccgctcagt aacaagactg gagaatctga tgtggaaaca aa	2640

gaattgaatc acattctatc agaaaatgag gtgaagttaa ctattatgac aggagacatc	2700
aaaggaatca tgcaggcagg aaaacgatct ctgcggcctc agcccactga gctgaagtat	2760
tcatggaaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt	2820
ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg	2880
gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa	2940
aaacaggatg tattctgcga ctcaaaactc atgtcagcgg ccataaaaga caacagagcc	3000
gtccatgccg atatgggtta ttggatagaa agtgcactca atgacacatg gaagatagag	3060
aaagcctctt tcattgaagt taaaaactgc cactggccaa aatcacacac cctctggagc	3120
aatggagtgc tagaaagtga gatgataatt ccaaagaatc tcgctggacc agtgtctcaa	3180
cacaactata gaccaggcta ccatacacia ataacaggac catggcatct aggtaagctt	3240
gagatggact ttgatttctg tgatggaaca acagtggtag tgactgagga ctgcggaaat	3300
agaggacctt ctttgagaac aaccactgcc tctggaaaac tcataacaga atggtgctgc	3360
cgatcttgca cattaccacc gctaagatac agaggtgagg atgggtgctg gtacgggatg	3420
gaaatcagac cattgaagga gaaagaagag aatttgggtca actccttggt cacagctgga	3480
catgggcagg tcgacaactt ttcactagga gtcttgggaa tggcattgtt cctggaggaa	3540
atgcttagga cccgagtagg aacgaaacat gcaatactac tagttgcagt ttcttttgtg	3600
acattgatca cagggaaacat gtcctttaga gacctgggaa gagtgatggt tatggtaggc	3660
gccactatga cggatgacat aggtatgggc gtgacttatac ttgccctact agcagccttc	3720
aaagtcagac caacttttgc agctggacta ctcttgagaa agctgacctc caaggaattg	3780
atgatgacta ctataggaat tgtactcctc tcccagagca ccataccaga gaccattctt	3840
gagttgactg atgcgttagc cttaggcatg atggtcctca aaatggtgag aaatatggaa	3900
aagtatcaat tggcagtgc tatcatggct atcttgtgcg tcccaaagc agtgatatta	3960
caaaacgcat ggaaagtgc ttgcacaata ttggcagtgg tgtccgtttc cccactgttc	4020
ttaacatcct cacagcaaaa aacagattgg ataccattag cattgacgat caaaggtctc	4080
aatccaacag ctatTTTTTct aacaaccctc tcaagaacca gcaagaaaag gagctggcca	4140
ttaaagtagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa	4200
aatgatattc ccatgacagg accattagtg gctggagggc tctcactgt gtgctacgtg	4260
ctcactggac gatcggccga tttggaactg gagagagcag ccgatgtcaa atgggaagac	4320

caggcagaga tatcaggaag cagtccaatc ctgtcaataa caatatcaga agatggtagc	4380
atgtcgataa aaaatgaaga ggaagaacaa aactgacca tactcattag aacaggattg	4440
ctggtgatct caggactttt tcctgtatca ataccaatca cggcagcagc atggtacctg	4500
tgggaagtga agaaacaacg ggccggagta ttgtgggatg ttccttcacc cccacccatg	4560
ggaaaggctg aactggaaga tggagcctat agaattaagc aaaaagggat tcttgatat	4620
tcccagatcg gagccggagt ttacaaagaa ggaacattcc atacaatgtg gcatgtcaca	4680
cgtggcgctg ttctaattgca taaaggaaag aggattgaac catcatgggc ggacgtcaag	4740
aaagacctaa tatcatatgg aggaggctgg aagttagaag gagaatggaa ggaaggagaa	4800
gaagtccagg tattggcact ggagcctgga aaaaatccaa gagccgtcca aacgaaacct	4860
ggtcttttca aaaccaacgc cggaacaata ggtgctgtat ctctggactt ttctcctgga	4920
acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt	4980
gttgttacaa ggagtggagc atatgtgagt gctatagccc agactgaaaa aagcattgaa	5040
gacaaccag agatcgaaga tgacattttc cgaaagagaa gactgaccat catggacctc	5100
caccagagag cgggaaagac gaagagatac cttccggcca tagtcagaga agctataaaa	5160
cggggtttga gaacattaat cttggcccc actagagttg tggcagctga aatggaggaa	5220
gcccttagag gacttccaat aagataccag accccagcca tcagagctgt gcacaccggg	5280
cgggagattg tggacctaat gtgtcatgcc acatttacca tgaggctgct atcaccagtt	5340
agagtgccaa actacaacct gattatcatg gacgaagccc atttcacaga ccagcaagt	5400
atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt	5460
atgacagcca ctccccggg aagcagagac ccatttcctc agagcaatgc accaatcata	5520
gatgaagaaa gagaaatccc tgaacgctcg tgaattccg gacatgaatg ggtcacggat	5580
tttaaaggga agactgtttg gttcgttcca agtataaaag caggaaatga tatagcagct	5640
tgctgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag	5700
tatgtcaaga ctagaaccaa tgattgggac ttcgtggtta caactgacat ttcagaaatg	5760
ggtgcccaatt tcaaggctga gaggggtata gacccagac gctgcatgaa accagtcata	5820
ctaacagatg gtgaagagcg ggtgattctg gcaggaccta tgccagtgac cactctagt	5880
gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata	5940
tacatggggg aacctctgga aatgatgaa gactgtgcac actggaaaga agctaaaatg	6000
ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt	6060

gaaaagggtgg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt	6120
gtagacttaa tgagaagagg agacctacca gtctgggttg cctacagagt ggcagctgaa	6180
ggcatcaact acgcagacag aagggtggtgt tttgatggag tcaagaacaa ccaaatocta	6240
gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc	6300
agatgggttg atgctaggat ctattctgac ccactggcgc taaaagaatt taaggaattt	6360
gcagccggaa gaaagtctct gaccctgaac ctaatcacag aaatgggtag gctcccaacc	6420
ttcatgactc agaaggcaag agacgcactg gacaacttag cagtgtctgca cacggctgag	6480
gcaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg	6540
cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca	6600
aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta	6660
tggtacgcac aaatacagcc aactggata gcagcttcaa taatactgga gttttttctc	6720
atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc	6780
tacgttgtca tagccatcct cacagtgggtg gccgcaacca tggcaaacga gatgggtttc	6840
ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caaccagca acccgagagc	6900
aacatcctgg acatagatct acgtcctgca tcagcatgga cgctgtatgc cgtggccaca	6960
acatttggtta caccaatgtt gagacatagc attgaaaatt cctcagtga tgtgtcccta	7020
acagctatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca	7080
aagatggaca tcggagttcc cttctcggc attggatgct actcacaagt caaccccata	7140
actctcacag cagctctttt cttattggtg gcacattatg ccatcatagg gccaggactc	7200
caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaacca	7260
actgtcgatg gaataacagt gattgaccta gatccaatac cttatgatcc aaagtttgaa	7320
aagcagttgg gacaagtaat gtccttagtc ctctgcgtga ctcaagtatt gatgatgagg	7380
actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg	7440
gaaggaaatc cagggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt	7500
agagggagtt acttggccgg agctggactt ctcttttcta ttatgaagaa cacaaccaac	7560
acaagaaggg gaactggcaa cataggagag acgcttgag agaaatggaa aagccgattg	7620
aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat	7680
agaaccttag caaaagaagg cattaaaaga ggagaaacgg accatcacgc tgtgtcgcga	7740

ggctcagcaa aactgagatg gttcgttgag agaaacatgg tcacaccaga agggaaagta	7800
gtggacctcg gttgtggcag aggaggctgg tcatactatt gtggaggact aaagaatgta	7860
agagaagtca aaggcctaac aaaaggagga ccaggacacg aagaacccat ccccatgtca	7920
acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcca	7980
gaaaagtgtg acacattatt gtgtgacata ggggagtcac caccaaatcc cacagtggaa	8040
gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa	8100
ttttgcataa aggttctcaa cccatatatg ccctcagtca tagaaaaaat ggaagcacta	8160
caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag	8220
atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg	8280
atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac	8340
ctcggaaagcg gaacccgtaa catcgggatt gaaagtgaga taccaaacct agatataatt	8400
gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac	8460
caccataca aaacgtgggc ataccatggg agctatgaaa caaaacagac tggatcagca	8520
tcatccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtccccatg	8580
gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgcgt ttttaaagag	8640
aaagtggaca cgagaacca agaaccgaaa gaaggcacga agaaactaat gaaaataaca	8700
gcagagtggc tttggaaaga attagggaag aaaaagacac ccaggatgtg caccagagaa	8760
gaattcaciaa gaaaggtgag aagcaatgca gccttggggg ccgtattcac tgatgagaac	8820
aagtggaagt cggcacgtga ggctgttgaa gatagtaggt tttgggagct ggttgacaag	8880
gaaaggaatc tccatcttga aggaaagtgt gaaacatgtg tgtacaacat gatgggaaaa	8940
agagagaaga agctagggga attcggcaag gcaaaaggca gcagagccat atggtacatg	9000
tggcttggag cacgcttctt agagtgtgaa gccctaggat tcttaaataga agatcactgg	9060
ttctccagag agaactccct gagtggagtg gaaggagaag ggctgcacaa gctaggttac	9120
attctaagag acgtgagcaa gaaagagga ggagcaatgt atgccgatga caccgcagga	9180
tgggatacaa gaatcacact agaagaccta aaaaatgaag aatggtaac aaaccacatg	9240
gaaggagAAC acaagaaact agccgaggcc attttcaaac taacgtacca aaacaagggtg	9300
gtgcgtgtgc aaagaccaac accaagaggc acagtaatgg acatcataatc gagaagagac	9360
caaagaggta gtggacaagt tggcacctat ggactcaata ctttcaccaa tatggaagcc	9420
caactaatca gacagatgga gggagaagga gtctttaaaa gcattcagca cctaacaatc	9480

acagaagaaa tcgctgtgca aaactgggtta gcaagagtgg ggcgcgaaaag gttatcaaga 9540
 atggccatca gtggagatga ttgtgttgtg aaaccttttag atgacagggtt cgcaagcgct 9600
 ttaacagctc taaatgacat gggaaagatt aggaaagaca tacaacaatg ggaaccttca 9660
 agaggatgga atgattggac acaagtggcc ttctgttcac accatttcca tgagttaatc 9720
 atgaaagacg gtcgcgtact cgttggttcca tgtagaaacc aagatgaact gattggcaga 9780
 gcccgaaatct cccaaggagc aggggtggtct ttgcgggaga cggcctgttt ggggaagtct 9840
 tacgccccaa tgtggagctt gatgtacttc cacagacgcg acctcaggct ggcggcaaata 9900
 gctatttgcg cggcagtacc atcacattgg gttccaacaa gtcgaacaac ctggtccata 9960
 catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa cagggtgtgg 10020
 attcaagaaa acccatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca 10080
 tacttgggga aaagagaaga ccaatggtgc ggctcattga ttgggttaac aagcagggcc 10140
 acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa 10200
 gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga 10260
 gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320
 catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380
 ggccatcata aatgccatag cttgagttaa ctatgcagcc tgtagctcca cctgagaagg 10440
 tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500
 ggttagagga gaccctccc ttacaaatcg cagcaacaat gggggcccaa ggcgagatga 10560
 agctgtagtc tcgctggaag gactagaggt tagaggagac cccccgaaa caaaaaacag 10620
 catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca 10680
 gaacgccaga aatggaatg gtgctgttga atcaacaggt tct 10723

<210> 15
 <211> 10717
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Virut dengue typ huyết thanh 3, MVS

<400> 15
 agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60
 gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120

aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag	180
ctgacaaaaga gattctcact tggaatgctg cagggacgag gaccattaaa actgttcatg	240
gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga	300
tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt	360
ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat cattatgctg	420
attccaacag tgatggcggt ccatttaacc acgcgtgatg gagagccgcg catgattgtg	480
gggaagaatg aaagaggaaa atccctactt ttcaagacag cctctggaat caacatgtgc	540
acactcatag ctatggatct gggagagatg tgtgatgaca cggtcactta caaatgcccc	600
cacattaccg aagtggagcc tgaagacatt gactgctggt gcaaccttac atcgacatgg	660
gtgacttatg gaacatgcaa tcaagctgga gagcatagac gcgataagag atcagtggcg	720
ttagctcccc atgttggcat gggactggac acacgcactc aaacctggat gtcggctgaa	780
ggagcttgga gacaagtcga gaaggtagag acatggggccc ttaggcaccc agggtttacc	840
atactagccc tatttcttgc ccattacata ggcacttcct tgaccagaa agtggttatt	900
tttatactat taatgctggt taccocatcc atgacaatga gatgtgtagg agtaggaaac	960
agagattttg tggaaggcct atcgggagct acgtgggttg acgtggtgct cgagcacggt	1020
gggtgtgtga ctacatggc taagaacaag cccacgctgg acatagagct tcagaagacc	1080
gaggccaccc aactggcgac cctaaggaag ctatgcattg agggaaaaat taccaacata	1140
acaaccgact caagatgtcc caccgaagg gaagcgattt tacctgagga gcaggaccag	1200
aactacgtgt gtaagcatac atacgtggac agaggctggg gaaacggttg tggtttgttt	1260
ggcaagggaa gcttggtgac atgcgcgaaa tttcaatgtt tagaatcaat agagggaaaa	1320
gtggtgcaac atgagaacct caaatacacc gtcacatca cagtgcacac aggagaccaa	1380
caccaggtgg gaaatgaaac gcagggagtc acggctgaga taacaccca ggcacatca	1440
gctgaagcca ttttacctga atatggaacc ctcggtgtag aatgctcacc acggacaggt	1500
ttggatttca atgaaatgat ctattgaca atgaagaaca aagcatggat ggtacataga	1560
caatggttct ttgacttacc cctaccatgg acatcaggag cttcagcaga aacaccaact	1620
tggaacagga aagagcttct tgtgacattt aaaaatgcac atgcaaaaaa gcaagaagta	1680
gttggttcttg gatcacaaga gggagcaatg catacagcac tgacaggagc tacagagatc	1740
caaacctcag gaggcacaag tatctttgcg gggcacttaa aatgtagact caagatggac	1800
aaattggaac tcaaggggat gagctatgca atgtgcttga gtagctttgt gttgaagaaa	1860

gaagtctccg aaacgcagca tgggacaata ctcattaagg ttgagtacaa aggggaagat	1920
gcaccctgca agattccttt ctccacggag gatggacaag gaaaagctct caatggcaga	1980
ctgatcacag ccaatccagt ggtgaccaag aaggaggagc ctgtcaacat tgaggctgaa	2040
cctccttttg gagaaagtaa catagtaatt ggaattggag acaaagccct gaaaatcaac	2100
tgggtacaaga agggaagctc gattgggaag atgttcgagg cactgccag aggtgcaagg	2160
cgcattggcca tcttgggaga cacagcctgg gactttggat cagtgggtgg tgttttgaat	2220
tcattagga aaatggtcca ccaaataatt gggagtgtt acacagccct atttgggtga	2280
gtctcctgga tgatgaaaat tggaataggt gtcctcttaa cctggatagg gttgaactca	2340
aaaaatactt ctatgtcatt ttcattgcac gcggccggca ttgtgacact gtatttgga	2400
gtcatgggtgc agccgatag tggttgcgtt gtgagctgga aaaacaaaga actgaaatgt	2460
ggcagtggga ttttcattcac agacaacgtg cacacatgga cagaacaata caagttccaa	2520
ccagaatccc cttcaaaact agcttcagct atccagaaag cccatgaaga ggacatttgt	2580
ggaatccgct cagtaacaag actggagaat ctgatgtgga acaaataac accagaattg	2640
aatcacattc tatcagaaaa tgaggtgaag ttaactatta tgacaggaga catcaaagga	2700
atcatgcagg caggaaaacg atctctgcgg cctcagccca ctgagctgaa gtattcatgg	2760
aaaacatggg gcaaagcaaa aatgctctct acagagtctc ataaccagac ctttctcatt	2820
gatggccccg aaacagcaga atgccccaac acaaatagag cttggaattc gttggaagtt	2880
gaagactatg gctttggagt attcaccacc aatatatggc taaaattgaa agaaaaacag	2940
gatgtattct gcgactcaa actcatgtca gcggccataa aagacaacag agccgtccat	3000
gcgatatgg gttattggat agaaagtga ctcaatgaca catggaagat agagaaagcc	3060
tctttcattg aagttaaaaa ctgccactgg ccaaaatcac acaccctctg gagcaatgga	3120
gtgctagaaa gtgagatgat aattccaaag aatctcgctg gaccagtgtc tcaacacaac	3180
tatagaccag gctaccatac acaataaca ggaccatggc atctaggtaa gcttgagatg	3240
gactttgatt tctgtgatgg aacaacagtg gtagtgactg aggactgagg aaatagagga	3300
ccctctttga gaacaaccac tgcctctgga aaactcataa cagaatggtg ctgccgatct	3360
tgcacattac caccgctaag atacagaggt gaggatgggt gctggtagcg gatggaaatc	3420
agaccattga aggagaaaga agagaatttg gtcaactcct tggtcacagc tggacatggg	3480
caggctcgaca acttttcact aggagtcttg ggaatggcat tgttcctgga ggaaatgctt	3540

aggacccgag taggaacgaa acatgcaata ctactagttg cagtttcttt tgtgacattg	3600
atcacaggga acatgtcctt tagagacctg ggaagagtga tggttatggt aggcgccact	3660
atgacggatg acataggtat gggcgtgact tatcttgccc tactagcagc cttcaaagtc	3720
agaccaactt ttgcagctgg actactcttg agaaagctga cctccaagga attgatgatg	3780
actactatag gaattgtact cctctcccag agcaccatac cagagaccat tcttgagttg	3840
actgatgcgt tagccttagg catgatggtc ctcaaatgg tgagaaatat ggaaaagtat	3900
caattggcag tgactatcat ggctatcttg tgcgtcccaa acgcagtgat attacaaaac	3960
gcatggaaag tgagttgcac aatattggca gtggtgtccg tttccccact gttcttaaca	4020
tcctcacagc aaaaaacaga ttggatacca ttagcattga cgatcaaagg tctcaatcca	4080
acagctattt ttctaacaac cctctcaaga accagcaaga aaaggagctg gccattaaat	4140
gaggctatca tggcagtcgg gatggtgagc attttagcca gttctctcct aaaaaatgat	4200
attcccatga caggaccatt agtggctgga gggctcctca ctgtgtgcta cgtgctcact	4260
ggacgatcgg ccgatttgga actggagaga gcagccgatg tcaaattggga agaccaggca	4320
gagatatcag gaagcagtcc aatcctgtca ataacaatat cagaagatgg tagcatgtcg	4380
ataaaaaatg aagaggaaga acaaacactg accatactca ttagaacagg attgctggtg	4440
atctcaggac tttttcctgt atcaatacca atcacggcag cagcatggta cctgtgggaa	4500
gtgaagaaac aacgggccgg agtattgtgg gatgttcctt cccccacc catgggaaag	4560
gctgaactgg aagatggagc ctatagaatt aagcaaaaag ggattcttgg atattcccag	4620
atcggagccg gagtttacia agaaggaaca ttccatacaa tgtggcatgt cacacgtggc	4680
gctgttctaa tgcataaagg aaagaggatt gaaccatcat gggcggacgt caagaaagac	4740
ctaatatcat atggaggagg ctggaagtta gaaggagaat ggaaggaagg agaagaagtc	4800
caggtattgg cactggagcc tggaaaaaat ccaagagccg tccaaacgaa acctgggtctt	4860
ttcaaaacca acgccggaac aataggtgct gtatctctgg acttttctcc tggaacgtca	4920
ggatctccaa ttatcgacaa aaaaggaaaa gttgtgggtc tttatggtaa tgggtgttgtt	4980
acaaggagtg gagcatatgt gagtgtata gccagactg aaaaagcat tgaagacaac	5040
ccagagatcg aagatgacat tttccgaaag agaagactga ccatcatgga cctccacca	5100
ggagcgggaa agacgaagag ataccttccg gccatagtca gagaagctat aaaacggggt	5160
ttgagaacat taatcttggc cccactaga gttgtggcag ctgaaatgga ggaagccctt	5220
agaggacttc caataagata ccagacccca gccatcagag ctgtgcacac cgggcgggag	5280

attgtggacc taatgtgtca tgccacattt accatgaggc tgctatcacc agttagagtg	5340
ccaaactaca acctgattat catggacgaa gcccatittca cagacccagc aagtatagca	5400
gctagaggat acatctcaac tcgagtggag atgggtgagg cagctgggat ttttatgaca	5460
gccactcccc cggaagcag agaccattt cctcagagca atgcaccaat catagatgaa	5520
gaaagagaaa tccctgaacg ctctggaat tccggacatg aatgggtcac ggattttaaa	5580
gggaagactg tttggttcgt tccaagtata aaagcaggaa atgatatagc agcttgccctg	5640
aggaaaaatg gaaagaaagt gatacaactc agtaggaaga cctttgattc tgagtatgtc	5700
aagactagaa ccaatgattg ggacttcgtg gttacaactg acatttcaga aatgggtgcc	5760
aatttcaagg ctgagagggg tatagacccc agacgctgca tgaaccagt cataactaaca	5820
gatggtgaag agcgggtgat tctggcagga cctatgccag tgaccactc tagtgcagca	5880
caaagaagag ggagaatagg aagaaatcca aaaaatgaga atgaccagta catatacatg	5940
ggggaacctc tggaaaatga tgaagactgt gcacactgga aagaagctaa aatgctccta	6000
gataacatca acacgccaga aggaatcatt cctagcatgt tcgaaccaga gcgtgaaaag	6060
gtggatgcca ttgatggcga ataccgcttg agaggagaag caaggaaaac cttttagac	6120
ttaatgagaa gaggagacct accagtctgg ttggcctaca gagtggcagc tgaaggcatc	6180
aactacgcag acagaagggtg gtgttttgat ggagtcaaga acaaccaaact cctagaagaa	6240
aacgtggaag ttgaaatctg gacaaaagaa ggggaaagga agaaattgaa acccagatgg	6300
ttggatgcta ggatctattc tgaccactg gcgctaaaag aatttaagga atttgcagcc	6360
ggaagaaagt ctctgaccct gaacctaatc acagaaatgg gtaggctccc aaccttcag	6420
actcagaagg caagagacgc actggacaac ttagcagtgc tgcacacggc tgaggcaggt	6480
ggaagggcgt acaaccatgc tctcagtga ctgccggaga ccctggagac attgctttta	6540
ctgacacttc tggctacagt cacgggaggg atctttttat tcttgatgag cgcaaggggc	6600
atagggaaga tgaccctggg aatgtgctgc ataatacagg ctagcatcct cctatggtac	6660
gcacaaatac agccacactg gatagcagct tcaataatac tggagttttt tctcatagtt	6720
ttgcttattc cagaacctga aaaacagaga acacccaag acaaccaact gacctacgtt	6780
gtcatagcca tcctcacagt ggtggccgca accatggcaa acgagatggg tttcctagaa	6840
aaaacgaaga aagatctcgg attgggaagc attgcaacc agcaaccga gagcaacatc	6900
ctggacatag atctacgtcc tgcacagca tggacgctgt atgccgtggc cacaacattt	6960

gttacaccaa tgttgagaca tagcattgaa aattcctcag tgaatgtgtc cctaacagct	7020
atagccaacc aagccacagt gttaatgggt ctcgggaaag gatggccatt gtcaaagatg	7080
gacatcggag ttcccccttct cgccattgga tgctactcac aagtcaaccc cataactctc	7140
acagcagctc ttttcttatt ggtagcacat tatgccatca tagggccagg actccaagca	7200
aaagcaacca gagaagctca gaaaagagca gcggcgggca tcatgaaaaa cccaactgtc	7260
gatggaataa cagtgattga cctagatcca ataccttatg atccaaagtt tgaaaagcag	7320
ttgggacaag taatgctcct agtcctctgc gtgactcaag tattgatgat gaggactaca	7380
tgggctctgt gtgaggcttt aaccttagct accgggcccc tctccacatt gtgggaagga	7440
aatccagga ggTTTTggaa cactaccatt gcggtgtcaa tggctaacat ttttagaggg	7500
agttacttgg ccggagctgg acttctcttt tctattatga agaacacaac caacacaaga	7560
aggggaactg gcaacatagg agagacgctt ggagagaaat ggaaaagccg attgaacgcg	7620
ttgggaaaaa gtgaattcca gatctacaag aaaagtggaa tccaggaagt ggatagaacc	7680
ttagcaaaaag aaggcattaa aagaggagaa acggaccatc acgctgtgtc gcgaggctca	7740
gcaaaaactga gatggttcgt tgagagaaac atggtcacac cagaaggga agtagtgga	7800
ctcggttgtg gcagaggagg ctggtcatac tattgtggag gactaaagaa tgtaagagaa	7860
gtcaaaggcc taacaaaagg aggaccagga cacgaagaac ccatcccat gtcaacatat	7920
gggtggaatc tagtgctct tcaaagtga gttgacgttt tcttcatccc gccagaaaag	7980
tgtgacacat tattgtgtga cataggggag tcatcaccaa atcccacagt ggaagcagga	8040
cgaacactca gagtccttaa cttagtagaa aattggttga acaacaacac tcaattttgc	8100
ataaaggttc tcaaccata tatgccctca gtcatagaaa aaatggaagc actacaaagg	8160
aaatatggag gagccttagt gaggaatcca ctctcacgaa actccacaca tgagatgtac	8220
tgggtatcca atgcttccgg gaacatagtg tcatcagtga acatgatttc aaggatgttg	8280
atcaacagat ttacaatgag atacaagaaa gccacttacg agccggatgt tgacctcgga	8340
agcggaaacc gtaacatcgg gattgaaagt gagataccaa acctagatat aattgggaaa	8400
agaatagaaa aaataaagca agagcatgaa acatcatggc actatgacca agaccacca	8460
tacaaaacgt gggcatacca tggtagctat gaaacaaaac agactggatc agcatcatcc	8520
atgggtcaacg gagtgggtcag gctgctgaca aaaccttggg acgtcgtccc catggtgaca	8580
cagatggcaa tgacagacac gactccattt ggacaacagc gcgtttttta agagaaagtg	8640
gacacgagaa cccaagaacc gaaagaaggc acgaagaaac taatgaaaat aacagcagag	8700

tggcttttga aagaattagg gaagaaaaag acaccagga tgtgcaccag agaagaattc	8760
acaagaaagg tgagaagcaa tgcagccttg ggggccatat tctactgatga gaacaagtgg	8820
aagtcggcac gtgaggctgt tgaagatagt aggttttggg agctgggtga caaggaaagg	8880
aatctccatc ttgaaggaaa gtgtgaaaca tgttgttaca acatgatggg aaaaagagag	8940
aagaagctag gggaattcgg caaggcaaaa ggcagcagag ccatatggta catgtggctt	9000
ggagcacgct tcttagagtt tgaagcccta ggattcttaa atgaagatca ctggttctcc	9060
agagagaact ccctgagtgg agtggaagga gaagggtgc acaagctagg ttacattcta	9120
agagacgtga gcaagaaaga gggaggagca atgtatgccg atgacaccgc aggatgggat	9180
acaagaatca cactagaaga cctaaaaaat gaagaaatgg taacaaacca catggaagga	9240
gaacacaaga aactagccga ggccattttc aaactaacgt accaaaacaa ggtggtgcgt	9300
gtgcaaagac caacaccaag aggcacagta atggacatca tatcgagaag agaccaaaga	9360
ggtagtggac aagttggcac ctatggactc aatactttca ccaatatgga agcccaacta	9420
atcagacaga tggagggaga aggagtcttt aaaagcattc agcacctaac aatcacagaa	9480
gaaatcgctg tgcaaaaactg gttagcaaga gtggggcgcg aaaggttatc aagaatggcc	9540
atcagtggag atgatttgtt tgtgaaacct ttagatgaca ggttcgcaag cgctttaaca	9600
gctctaaatg acatgggaaa gattaggaaa gacatacaac aatgggaacc ttcaagagga	9660
tggaatgatt ggacacaagt gcccttctgt tcacaccatt tccatgagtt aatcatgaaa	9720
gacggtcgcg tactcgttgt tccatgtaga aaccaagatg aactgattgg cagagccga	9780
atctcccaag gagcaggtgt gtctttgcgg gagacggcct gtttggggaa gtcttacgcc	9840
caaatgtgga gcttgatgta cttccacaga cgcgacctca ggctggcggc aaatgctatt	9900
tgctcggcag taccatcaca ttgggttcca acaagtcgaa caacctggtc catacatgct	9960
aaacatgaat ggatgacaac ggaagacatg ctgacagtct ggaacagggg gtggattcaa	10020
gaaaacccat ggatggaaga caaaactcca gtggaatcat gggaggaaat ccatacttg	10080
gggaaaagag aagaccaatg gtgcggctca ttgattgggt taacaagcag ggccacctgg	10140
gcaaagaaca tccaagcagc aataaatcaa gttagatccc ttataggcaa tgaagaatac	10200
acagattaca tgccatccat gaaaagattc agaagagaag aggaagaagc aggagttctg	10260
tggtagaaag caaaactaac atgaaacaag gctagaagtc aggtcggatt aagccatagt	10320
acggaaaaaa ctatgctacc tgtgagcccc gtccaaggac gttaaaagaa gtcaggccat	10380

cataaatgcc atagcttgag taaactatgc agcctgtagc tccacctgag aaggtgtaaa 10440
 aaatccggga ggccacaaac catggaagct gtacgcatgg cgtagtggac tagcggttag 10500
 aggagacccc tcccttacia atcgagcaa caatgggggc ccaaggcgag atgaagctgt 10560
 agtctcgctg gaaggactag aggttagagg agaccccc gaaacaaaaa acagcatatt 10620
 gacgctggga aagaccagag atcctgctgt ctctcagca tcattccagg cacagaacgc 10680
 cagaaaatgg aatggtgctg ttgaatcaac aggttct 10717

<210> 16
 <211> 10723
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Virut dengue typ huyết thanh 4, MVS

<220>
 <221> n
 <222> (3773)..(3773)
 <223> a hoặc g

<220>
 <221> n
 <222> (7026)..(7026)
 <223> t hoặc c

<400> 16
 agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60
 gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120
 aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180
 ctgacaaaga gattctcact tggaatgctg caggagcgag gacctttaa actgttcatg 240
 gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300
 tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360
 ggaaggatgc tgaacatctt gaataggaga cgcagctctg caggcatgat cattatgctg 420
 attccaacag tgatggcggt ccatttaacc acgcgtgatg gcgaaccct catgatagt 480
 gcaaaacatg aaagggggag acctctcttg tttaagacaa cagaggggat caacaaatgc 540
 actctcattg ccatggactt gggtgaaatg tgtgaggaca ctgtcacgta taaatgcccc 600
 ttactggtca ataccgaacc tgaagacatt gattgctggt gcaatctcac gtctacctgg 660
 gtcattgtatg ggacatgcac ccagagcgga gaacggagac gagagaagcg ctcatgtagct 720
 ttaacaccac attcaggaat gggattggaa acaagagctg agacatggat gtcacggaa 780

ggggcttgga agcatgctca gagagtagag agctggatac tcagaaaccc aggattcgcg	840
ctcttggcag gatttatggc ttatatgatt gggcaaacag gaatccagcg aactgtcttc	900
tttgcctaa tgatgctggt cgccccatcc tacggaatgc gatgcgtagg agtaggaaac	960
agagactttg tggaaggagt ctcaggtgga gcatgggtcg atctggtgct agaacatgga	1020
ggatgcgtca caaccatggc ccagggaaaa ccaaccttgg attttgaact gactaagaca	1080
acagccaagg aagtggctct gttaagaacc tattgcattg aagcctcaat atcaaacata	1140
accacggcaa caagatgtcc aacgcaagga gagccttata taaaagagga acaagaccaa	1200
cagtacattt gccggagaga tgtggtagac agaggggtgg gcaatggctg tggcttgttt	1260
ggaaaaggag gagttgtgac atgtgcgaag ttttcatgtt cggggaagat aacaggcaat	1320
ttgggtccaaa ttgagaacct tgaatacaca gtgggttgtaa cagtccacaa tggagacacc	1380
catgcagtag gaaatgacac gtccaatcat ggagttacag ccacgataac tcccaggta	1440
ccatcggtgg aagtcaaatt gccggactat ggagaactaa cactcgattg tgaacccagg	1500
tctggaattg actttaatga gatgattctg atgaaaatga aaaagaaaac atggcttgtg	1560
cataagcaat ggtttttgga tctacctcta ccatggacag caggagcaga cacatcagag	1620
gttacttgga attacaaaga gagaatggtg acatttaagg ttcctcatgc caagagacag	1680
gatgtgacag tgctgggatc tcaggaagga gccatgcatt ctgccctcgc tggagccaca	1740
gaagtggact ccggtgatgg aaatcacatg tttgcaggac atctcaagtg caaagtccgt	1800
atggagaaat tgagaatcaa gggaatgtca tacacgatgt gttcaggaaa gttctcaatt	1860
gacaaagaga tggcagaaac acagcatggg acaacagtgg tgaaagtcaa gtatgaaggt	1920
gctggagctc cgtgtaaagt ccccatagag ataagagatg tgaacaagga aaaagtgggt	1980
gggcgtatca tctcatccac ccctttggct gagaatacca acagtgtaac caacatagag	2040
ttagaacccc cctttgggga cagctacata gtgatagggt ttggaaacag tgcattaaca	2100
ctccattgggt tcaggaaagg gagttccatt ggcaagatgt ttgagtccac atacagaggt	2160
gcaaaacgaa tggccattct aggtgaaaca gcttgggatt ttggttccgt tgggtggactg	2220
ttcacatcat tgggaaaggc tgtgcaccag gtttttgtaa gtgtgtatac aacctgttt	2280
ggaggagtct catggatgat tagaatccta attgggttcc tagtgttgtg gattggcacg	2340
aactcaagga acacttcaat ggctatgacg tgcatagctg ccggcattgt gacactgtat	2400
ttgggagtca tgggtgcaggc cgatagtgggt tgcgttgtga gctggaaaaa caaagaactg	2460

aaatgtggca gtgggatttt catcacagac aacgtgcaca catggacaga acaatacaag	2520
ttccaaccag aatcccccttc aaaactagct tcagctatcc agaaagccca tgaagaggac	2580
atttgtggaa tccgctcagt aacaagactg gagaatctga tgtggaaaca aataacacca	2640
gaattgaatc acattctatc agaaaatgag gtgaagttaa ctattatgac aggagacatc	2700
aaaggaatca tgcaggcagg aaaacgatct ctgcggcctc agcccactga gctgaagtat	2760
tcatggaaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt	2820
ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg	2880
gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa	2940
aaacaggatg tattctgcga ctcaaaactc atgtcagcgg ccataaaaga caacagagcc	3000
gtccatgccg atatgggtta ttggatagaa agtgcactca atgacacatg gaagatagag	3060
aaagcctctt tcattgaagt taaaaactgc cactggccaa aatcacacac cctctggagc	3120
aatggagtgc tagaaagtga gatgataatt ccaagaatc tcgctggacc agtgtctcaa	3180
cacaactata gaccaggcta ccatacacia ataacaggac catggcatct aggtaagctt	3240
gagatggact ttgatttctg tgatggaaca acagtggtag tgactgagga ctgcggaaat	3300
agaggaccct ctttgagaac aaccactgcc tctggaaaac tcataacaga atgggtgctgc	3360
cgatcttgca cattaccacc gctaagatac agaggtgagg atgggtgctg gtacgggatg	3420
gaaatcagac cattgaagga gaaagaagag aatttgggtca actccttggc cacagctgga	3480
catgggcagg tcgacaactt ttcactagga gtcttgggaa tggcattgtt cctggaggaa	3540
atgcttagga cccgagtagg aacgaaacat gcaatactac tagttgcagt ttcttttgtg	3600
acattgatca cagggaacat gtcctttaga gacctgggaa gagtgatggt tatggtaggc	3660
gccactatga cgggtgacat aggtatgggc gtgacttata ttgccctact agcagccttc	3720
aaagtcagac caacttttgc agctggacta ctcttgagaa agctgacctc canggaattg	3780
atgatgacta ctataggaat tgtactcctc tcccagagca ccataccaga gaccattctt	3840
gagttgactg atgcgttagc cttaggcattg atggctctca aaatgggtgag aaatatggaa	3900
aagtatcaat tggcagtgac tatcatggct atcttgtgcg tcccaaacgc agtgatatta	3960
caaaacgcat ggaaagtgag ttgcacaata ttggcagtggt tgtccgtttc cccactgttc	4020
ttaacatcct cacagcaaaa aacagattgg ataccattag cattgacgat caaaggcttc	4080
aatccaacag ctatttttct aacaaccctc tcaagaacca gcaagaaaag gagctggcca	4140
ttaaatgagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa	4200

aatgatattc ccatgacagg accattagtg gctggagggc tcctcactgt gtgctacgtg	4260
ctcactggac gatcggccga tttggaactg gagagagcag ccgatgtcaa atgggaagac	4320
caggcagaga tatcaggaag cagtccaatc ctgtcaataa caatatcaga agatggtagc	4380
atgtcgataa aaaatgaaga ggaagaacaa aactgacca tactcattag aacaggattg	4440
ctggtgatct caggactttt tcctgtatca ataccaatca cggcagcagc atggtacctg	4500
tgggaagtga agaaacaacg ggccggagta ttgtgggatg ttccttcacc cccacccatg	4560
ggaaaggctg aactggaaga tggagcctat agaattaagc aaaaagggat tcttggatat	4620
tcccagatcg gagccggagt ttacaaagaa ggaacattcc atacaatgtg gcatgtcaca	4680
cgtggcgctg ttctaattgca taaaggaaag aggattgaac catcatgggc ggacgtcaag	4740
aaagacctaa tatcatatgg aggaggctgg aagttagaag gagaatggaa ggaaggagaa	4800
gaagtccagg tattggcact ggagcctgga aaaaatccaa gagccgtcca aacgaaacct	4860
ggtcttttca aaaccaacgc cggaacaata ggtgctgtat ctctggactt ttctcctgga	4920
acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt	4980
gttgttacaa ggagtggagc atatgtgagt gctatagccc agactgaaaa aagcattgaa	5040
gacaaccag agatcgaaga tgacattttc cgaaagagaa gactgaccat catggacctc	5100
caccaggag cgggaaaagac gaagagatac cttccggcca tagtcagaga agctataaaa	5160
cggggtttga gaacattaat cttggccccc actagagttg tggcagctga aatggaggaa	5220
gcccttagag gacttccaat aagataccag accccagcca tcagagctgt gcacaccggg	5280
cgggagattg tggacctaat gtgtcatgcc acatttacca tgaggctgct atcaccagtt	5340
agagtgccaa actacaacct gattatcatg gacgaagccc atttcacaga tccagcaagt	5400
atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt	5460
atgacagcca ctccccggg aagcagagac ccatttcctc agagcaatgc accaatcata	5520
gatgaagaaa gagaaatccc tgaacgctcg tggaattccg gacatgaatg ggtcacggat	5580
tttaaaggga agactgtttg gttcgttcca agtataaaag caggaaatga tatagcagct	5640
tgcttgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag	5700
tatgtcaaga ctagaaccaa tgattgggac ttcgtggtta caactgacat ttcagaaatg	5760
ggtgccaat tcaaggctga gaggggtata gacccagac gctgcatgaa accagtcata	5820
ctaacagatg gtgaagagcg ggtgattctg gcaggacctg tgccagtgc ccactctagt	5880

gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata	5940
tacatggggg aacctctgga aaatgatgaa gactgtgcac actggaaaga agctaaaatg	6000
ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt	6060
gaaaagggtg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt	6120
gtagacttaa tgagaagagg agacctacca gtctggttgg cctacagagt ggcagctgaa	6180
ggcatcaact acgcagacag aagggtggtg tttgatggag tcaagaacaa ccaaactcta	6240
gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc	6300
agatggttgg atgctaggat ctattctgac cactggcgc taaaagaatt taaggaattt	6360
gcagccggaa gaaagtctct gacctgaac ctaatcacag aaatgggtag gctcccaacc	6420
ttcatgactc agaaggtaag agacgcactg gacaacttag cagtgtctgca cacggctgag	6480
gcaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg	6540
cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca	6600
aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta	6660
tggtacgcac aaatacagcc aactggata gcagcttcaa taatactgga gttttttctc	6720
atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc	6780
tacgttgtca tagccatcct cacagtggtg gccgcaacca tggcaaacga gatgggtttc	6840
ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caaccagca acccgagagc	6900
aacatcctgg acatagatct acgtcctgca tcagcatgga cgctgtatgc cgtggccaca	6960
acatttgtaa caccaatgtt gagacatagc attgaaaatt cctcagtga tgtgtcccta	7020
acagcnatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca	7080
aagatggaca tcggagttcc cttctcgcc attggatgct actcacaagt caaccccata	7140
actctcacag cagctctttt cttattggta gcacattatg ccatcatagg gccaggactc	7200
caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaacca	7260
actgtcgatg gaataacagt gattgaccta gatccaatac cttatgatcc aaagtttgaa	7320
aagcagttgg gacaagtaat gtcctagtc ctctgcgtga ctcaagtatt gatgatgagg	7380
actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg	7440
gaaggaaatc caggagaggt ttggaacact accattgcgg tgtcaatggc taacattttt	7500
agagggagtt acttggccgg agctggactt ctcttttcta ttatgaagaa cacaaccaac	7560
acaagaaggg gaactggcaa cataggagag acgcttgag agaaatggaa aagccgattg	7620

aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat	7680
agaaccttag caaaagaagg cattaaaaga ggagaaacgg accatcacgc tgtgtcgcga	7740
ggctcagcaa aactgagatg gttcgttgag agaaacatgg tcacaccaga agggaaagta	7800
gtggacctcg gttgtggcag aggaggctgg tcatactatt gtggaggact aaagaatgta	7860
agagaagtca aaggcctaac aaaaggagga ccaggacacg aagaacccat ccccatgtca	7920
acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcc	7980
gaaaagtgtg acacattatt gtgtgacata ggggagtcac caccaaacc cacagtggaa	8040
gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa	8100
ttttgcataa aggttctcaa cccatatatg ccctcagtca tagaaaaaat ggaagcacta	8160
caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag	8220
atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg	8280
atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac	8340
ctcgggaagcg gaaccgtaa catcgggatt gaaagtgaga taccaaacc agatataatt	8400
gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac	8460
caccataca aaacgtgggc ataccatggt agctatgaaa caaacagac tggatcagca	8520
tcattccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtcccatg	8580
gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgcgt ttttaaagag	8640
aaagtggaca cgagaacca agaaccgaaa gaaggcacga agaaactaat gaaaataaca	8700
gcagagtggc tttggaaaga attagggaag aaaaagacac ccaggatgtg caccagagaa	8760
gaattcaciaa gaaaggtgag aagcaatgca gccttggggg ccatattcac tgatgagaac	8820
aagtggaagt cggcacgtga ggctgttgaa gatagtaggt tttgggagct ggttgacaag	8880
gaaaggaatc tccatcttga aggaaagtgt gaaacatgtg tgtacaacat gatgggaaaa	8940
agagagaaga agctagggga attcggcaag gcaaaaggca gcagagccat atggtacatg	9000
tggcttggag cacgcttctt agagtttgaa gccctaggat tcttaaatga agatcactgg	9060
ttctccagag agaactccct gagtggagtg gaaggagaag ggctgcacaa gctaggttac	9120
attctaagag acgtgagcaa gaaagaggga ggagcaatgt atgccgatga caccgcagga	9180
tgggatacaa gaatcacact agaagaccta aaaaatgaag aaatggtaac aaaccacatg	9240
gaaggagaac acaagaaact agccgaggcc attttcaaac taacgtacca aaacaagggtg	9300

gtgctgtgtgc aaagaccaac accaagaggc acagtaatgg acatcatatc gagaagagac	9360
caaagaggta gtggacaagt tggcacctat ggactcaata ctttcaccaa tatggaagcc	9420
caactaatca gacagatgga gggagaagga gtctttaaaa gcattcagca cctaacaatc	9480
acagaagaaa tcgctgtgca aaactgggta gcaagagtgg ggcgcgaaaag gttatcaaga	9540
atggccatca gtggagatga ttgtgttggtg aaaccttttag atgacagggt cgcaagcgct	9600
ttaacagctc taaatgacat gggaaagatt aggaaagaca tacaacaatg ggaaccttca	9660
agaggatgga atgattggac acaagtggcc ttctgttcac accatttcca tgagttaatc	9720
atgaaagacg gtcgcgtact cggtgttccc tgtagaaacc aagatgaact gattggcaga	9780
gcccgaatct cccaaggagc aggggtggtct ttgcgggaga cggcctgttt ggggaagtct	9840
tacgccccaa tgtggagctt gatgtacttc cacagacgcg acctcaggct ggcggcaa	9900
gctatttgct cggcagtacc atcacattgg gttccaacaa gtcgaacaac ctggtccata	9960
catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa cagggtgtgg	10020
attcaagaaa acctatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca	10080
tacttgggga aaagagaaga ccaatgggtgc ggctcattga ttgggttaac aagcagggcc	10140
acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa	10200
gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga	10260
gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcagggt cggattaagc	10320
catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca	10380
ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg	10440
tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc	10500
ggttagagga gaccctccc ttacaaatcg cagcaacaat gggggcccaa ggcgagatga	10560
agctgtagtc tcgctggaag gactagaggt tagaggagac cccccgaaa caaaaaacag	10620
catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca	10680
gaacgccaga aaatggaatg gtgctgttga atcaacaggt tct	10723

Fig. 1

Biến thể di truyền trong số các thể khảm D2/4 (so với D2 16681 và D4-1036 kiểu hoang dại

Hệ gen	D2		D4		kết hợp	kết hợp	D2						NS5
	NCR	C	Eng	E			NS1	NS2A	NS3	NS4A	NS4B	PDK53	
Gen	PDK-53	Eng	Eng	Eng	Eng	Eng	PDK-53	PDK-53	PDK-53	PDK-53	NS4B	NS5	
Kiểu đột biến *	C57T	A225T	A396C	C2027T	A2275	TG2380/1CC	G2579A	A3674G	A3773A/G	C4018T	C6437T	C8571T**	
Hệ gen NT	C57T	A225T	A396C	C2027T	A2275	TG2380/1CC	G2579A	A3674G	A3773A/G	C4018T	C6437T	C8571T**	
Vị trí	C57T	A225T	A396C	C2027T	A2275	TG2380/1CC	G2579A	A3674G	A3773A/G	C4018T	C6437T	C8571T**	
Vị trí protein AA	C57T	A225T	A396C	C2027T	A2275	TG2380/1CC	G2579A	A3674G	A3773A/G	C4018T	C6437T	C8571T**	
D2-16681	C	A	A (R)	A	E	TG (V)	G (G0)	A (D)	A (K)	C (L)	C (A)	C	
D2-PDK-53	T	-	-	-	E	A (D)	A (D)	-	-	T (F)	-	t	
D4-1036	T	-	-	-	E	CC (A)	A (D)	-	-	T (F)	-	-	
Dòng D2/4-V1	T	-	C (S)	T (V)	C (L)	CC (A)	A (D)	G (G)	A (G) (K/R)	T (F)	T (V)	-	
(pD2/4-VF1)	T	-	C (S)	T (V)	C (L)	CC (A)	A (D)	G (G)	A (G) (K/R)	T (F)	T (V)	-	
DENVa-v-4 (NVS)	T	t	C (S)	T (V)	C (L)	CC (A)	A (D)	G (G)	A (G) (K/R)	T (F)	T (V)	c/t	

"-" tương tự như D2 16681 hoặc D4-1036 kiểu hoang dại; đột biến câm trong vùng đọc mở
*: PDK-53: Kiểu di truyền đặc hiệu D2-PDK-53 (đối lại 16681), Mẫu đồ: locus PDK-53 giảm độc lực chính; Hat: đột biến chỉ được tìm thấy trong hạt virus xác định và không tìm thấy trong dòng gốc; Eng: đột biến được thao tác đối với các dòng D2/4; Mlu và NgoMV: vị trí RE thao tác kết hợp D2/4;
**: C8571T (đột biến câm PDK-53) không được bao gồm trong hầu hết dòng khảm D2/4

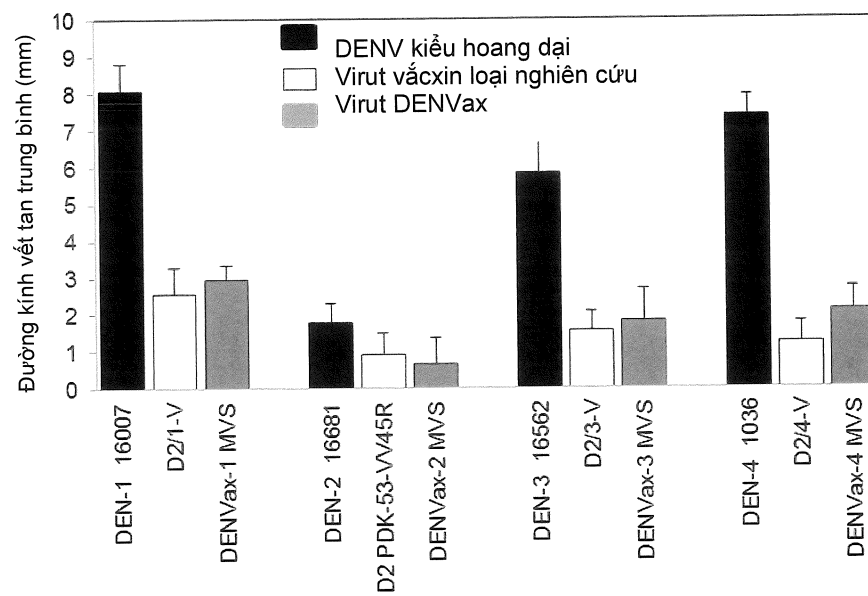
Fig. 2

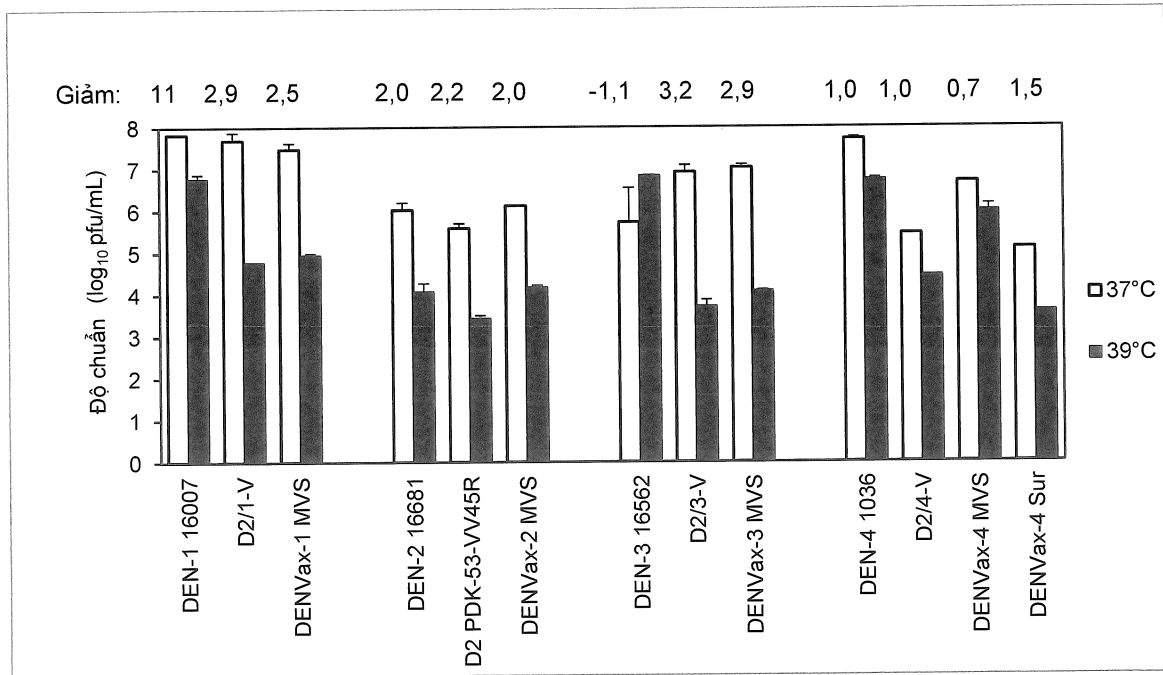
Fig. 3

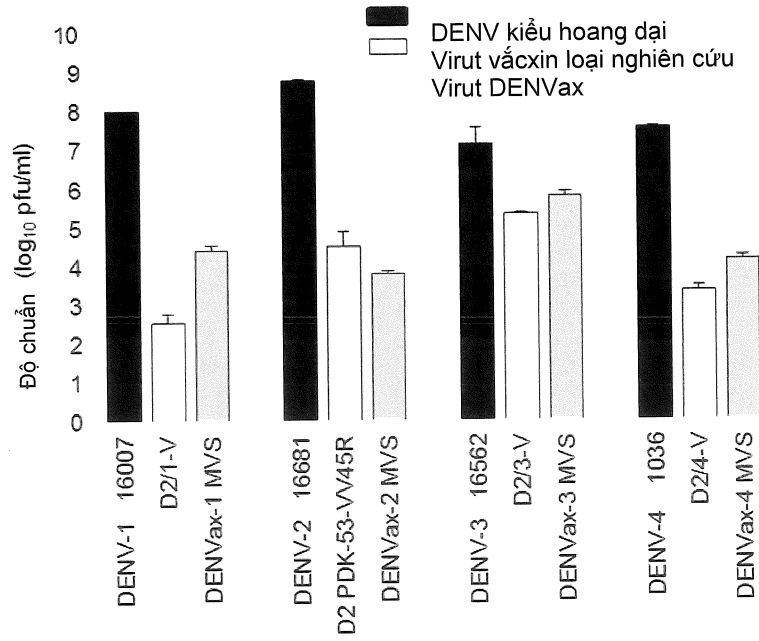
Fig. 4

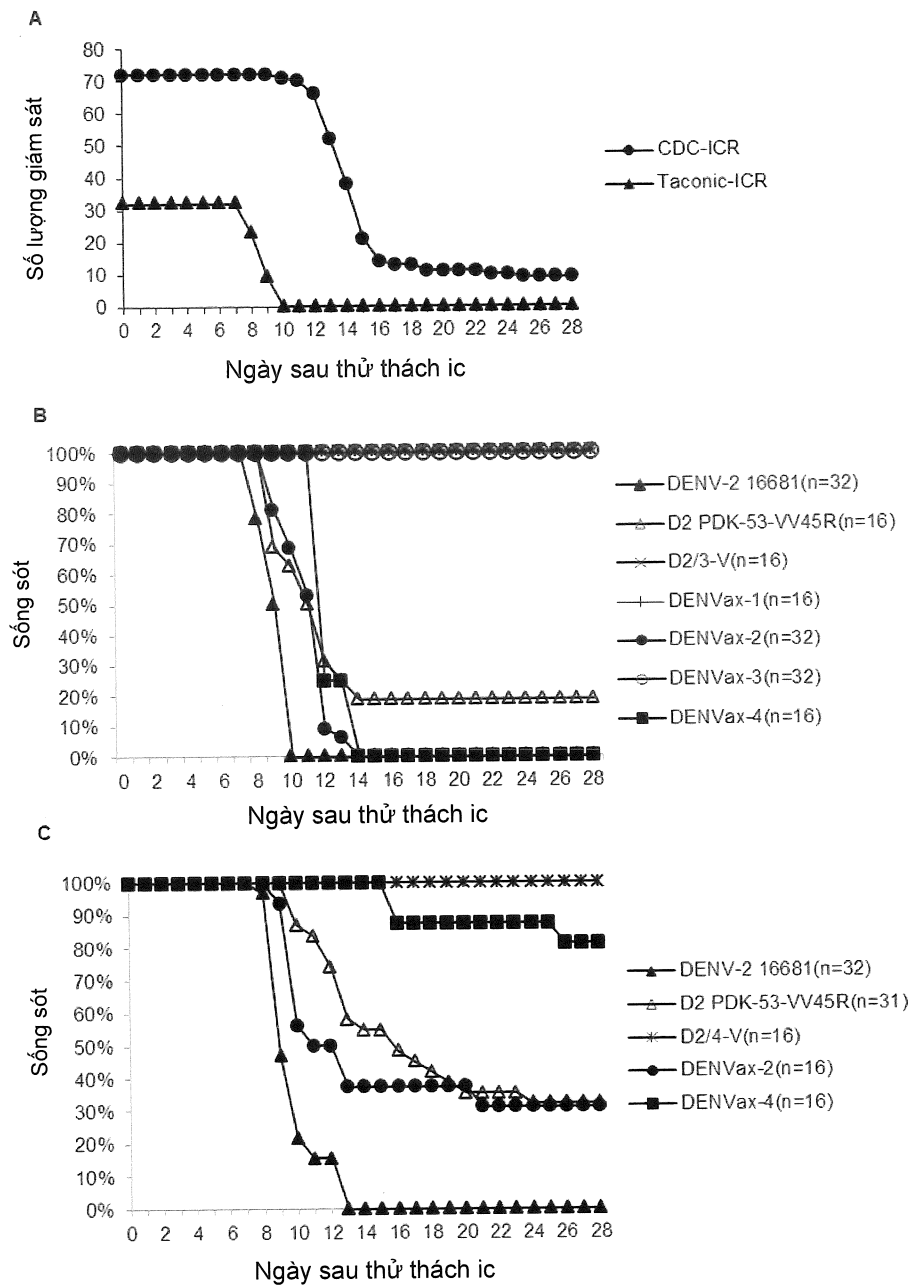
Fig. 5

Fig. 6

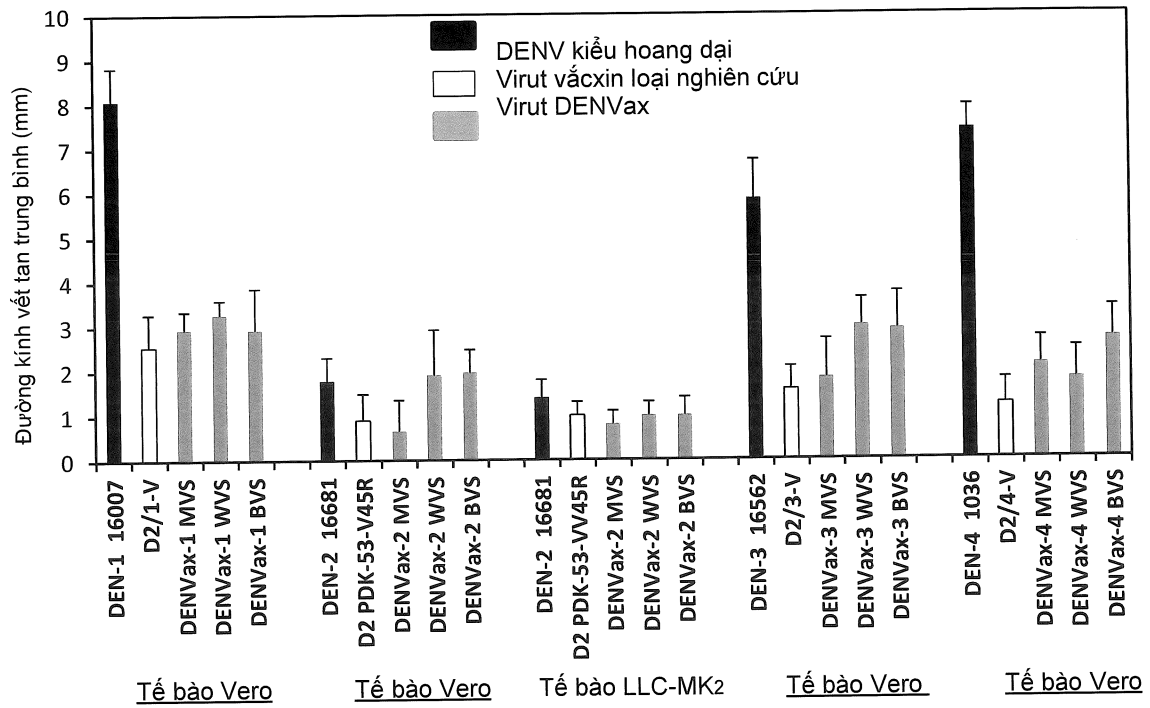


Fig. 7

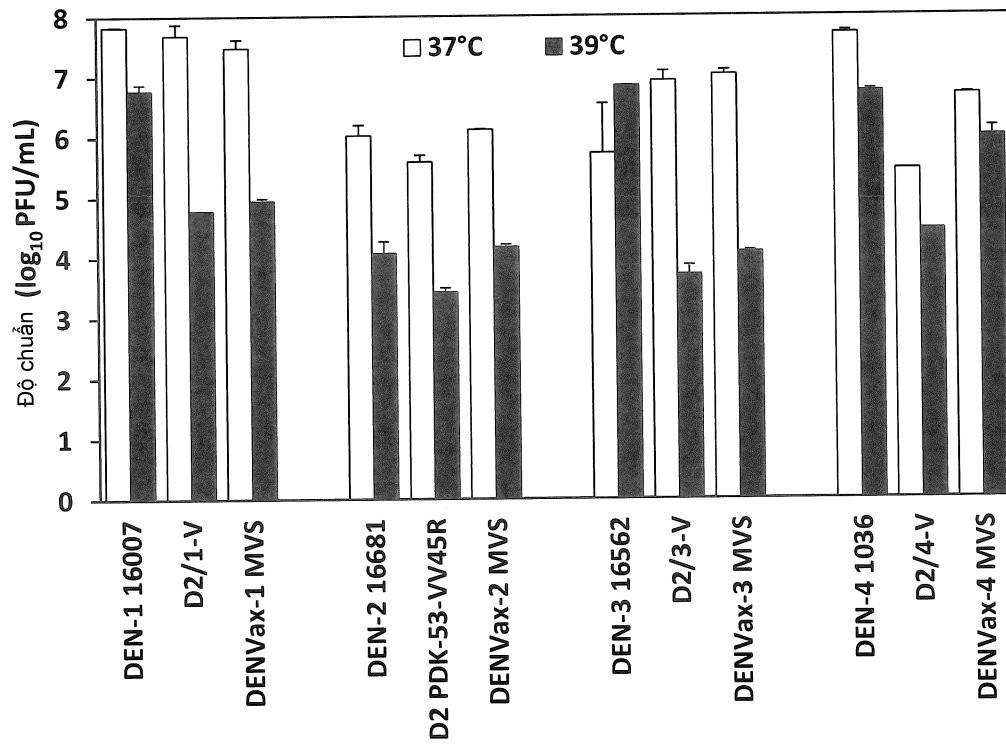


Fig. 8

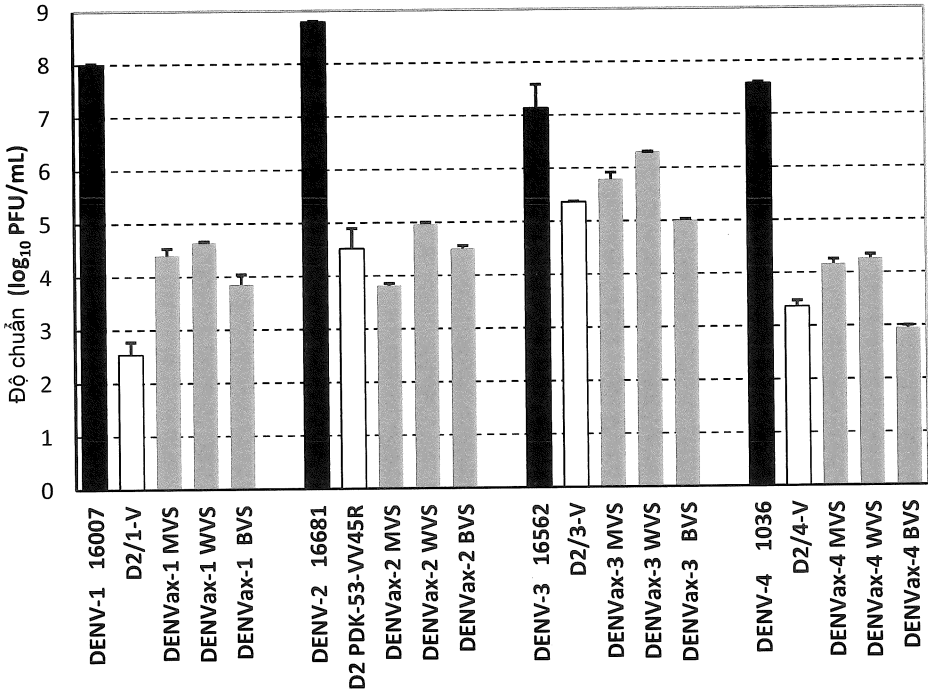


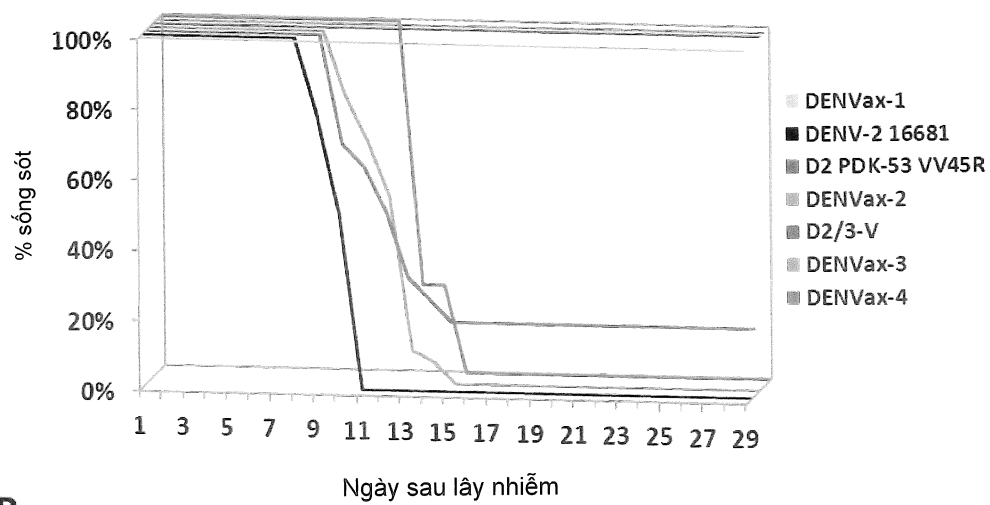
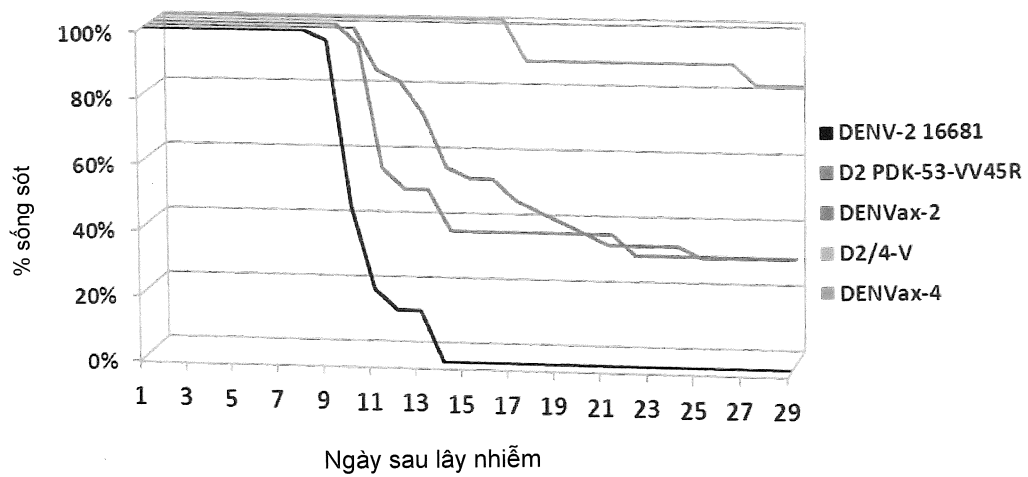
Fig. 9**A****B**

Fig. 10

Kiểu huyết Chủng thanh	Nguồn gốc virus	CSTI	A524T	T900C ^a	C205ST	G1579A	C401ST	A52703	T5547C	G6599C	C8571T
		SYNCR	pm4-D29V	M (cảm)	E (cảm)	NS3-G53D	NS2A-L181	NS3-E380V	NS3 (cảm)	NS4A-G75A	NS5 (cảm)
16681	Nguồn gốc virus	C	A	T	C	G	C	A	T	G	C
PDK-53	phân lập từ người	T	T	.	t	A	T	hỗn hợp T/A	ε	C	t
DENV-2	tế bào PDK của 16681	T	T	ε	t	A	T	T	ε	C	C
PDK53-V(VV49R)	PDK-53-V tái tổ hợp	T	T	ε	t	A	T	A	ε	C	C
PDK53-E(VF48R)	PDK-53-E tái tổ hợp	T	T	ε	t	A	T	A	ε	C	C

Đột biến không gạch dưới: 3 locus giảm độc lực hậu hết quan trọng của PDK-53

Màu đỏ: trình tự đặc hiệu PDK-53 (thay đổi từ 16681)

Nét đậm: Trình tự khác nhau giữa PDK-53 và virus V hoặc E thu được từ dòng

^a Gen đánh dấu dòng cảm được thao tác để phân biệt PDK-53 ban đầu và virus tái tổ hợp (thu được từ dòng)