



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

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

(11)



1-0029897

(51)⁷ A61K 39/12; C12N 7/08; C07K 14/08

(13) B

(21) 1-2013-01758

(22) 09/11/2011

(86) PCT/IB2011/055003 09/11/2011

(87) WO/2012/063212 18/05/2012

(30) 61/412,006 10/11/2010 US

(45) 25/10/2021 403

(43) 25/10/2013 307A

(73) ZOETIS SERVICES LLC (US)

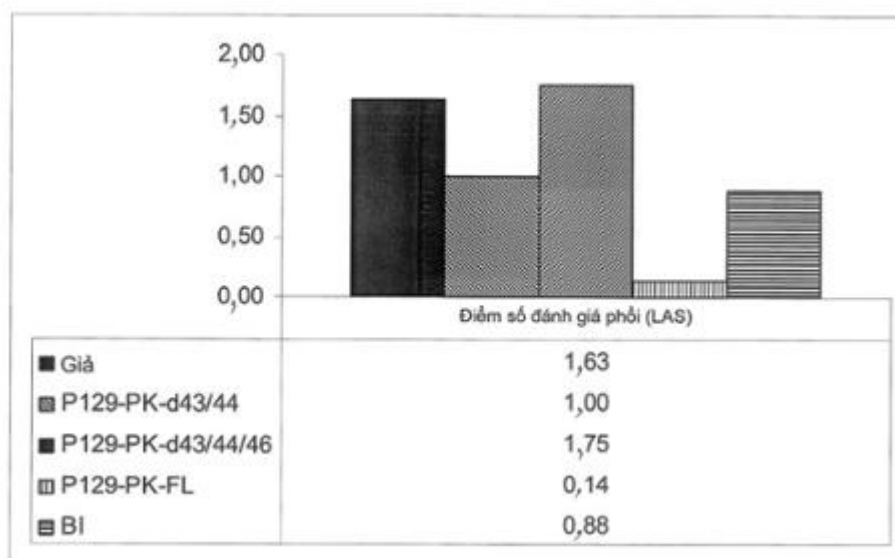
10 Sylvan Way, Parsippany, NJ 07054, United States of America

(72) CALVERT, Jay Gregory (US); SLADE, David Ewell (US); WELCH, Siao-Kun Wan (US).

(74) Công ty TNHH T&T INVENMARK Sở hữu trí tuệ Quốc tế (T&T INVENMARK CO., LTD.)

(54) PHÂN TỬ POLYNUCLEOTIT ĐÃ PHÂN LẬP BAO GỒM TRÌNH TỰ ADN MÃ HÓA PHÂN TỬ ARN GÂY NHIỄM MÃ HÓA VIRUT GÂY BỆNH RỐI LOẠN HÔ HẤP VÀ SINH SẢN TRÊN LỢN (PRRS) BẮC MỸ, VIRUT PRRS, VACCIN, VÀ PHƯƠNG PHÁP TẠO RA VIRUT NÀY

(57) Sáng chế đề xuất phân tử polynucleotit phân lập chứa trình tự ADN mã hóa trình tự ARN gây nhiễm mã hóa virut PRRS Bắc Mỹ cải biến di truyền, các phương pháp tạo ra nó và các polypeptit, polynucleotit liên quan, và các thành phần khác nhau. Sáng chế cũng đề xuất vaccin chứa virut cải biến di truyền này và polynucleotit và kit chẩn đoán để phân biệt giữa các con vật lây nhiễm tự nhiên và dùng vaccin.



Lĩnh vực kỹ thuật được đề cập

Sáng chế thuộc lĩnh vực thú y và đề cập đến dòng cADN lây nhiễm của virus ARN phân cực dương, virus ARN mới và các dạng cải biến của nó, và bào chế vaccin, cụ thể là vaccin cho lợn sử dụng dòng cADN này.

Tình trạng kỹ thuật của sáng chế

Hội chứng rối loạn hô hấp và sinh sản ở lợn (Porcine reproductive và respiratory syndrome-PRRS) được đặc trưng bởi lợn bị sảy thai, chết non, và các vấn đề sinh sản khác ở lợn nái và lợn cái con, cũng như bệnh về hô hấp ở lợn con. Tác nhân gây bệnh là virus PRRS (PRRS virus-PRRSV), thành viên của họ *Arteriviridae* và bộ *Nidovirales*. Nidovirus là virus có vỏ có hệ gen chứa sợi đơn ARN phân cực dương. ARN hệ gen của virus ARN sợi dương đáp ứng vai trò kép bao gồm trong cả bảo vệ và biểu hiện thông tin di truyền. Không có ADN tham gia vào trong sao chép hoặc dịch mã trong Nidovirus. Các protein không tham gia vào trong cấu trúc được dịch mã trực tiếp từ ARN hệ gen của nidovirus là các polyprotein lớn và sau đó được cắt bằng các proteaza virus thành các protein có chức năng. Bộ ARN dưới hệ gen (subgenomic RNA-sgRNA) ẩn nằm liền kề với đầu tận cùng 3' được tổng hợp từ hệ gen và được sử dụng làm ARN thông tin để dịch mã các protein chức năng. Do đó, sự nhân lên của ARN hệ gen nidovirus là quá trình kết hợp của sao chép hệ gen và tổng hợp sgARN.

Cuối những năm 80, hai kiểu gen riêng biệt của virus này xuất hiện gần như đồng thời, một ở Bắc Mỹ và kiểu gen kia ở châu Âu. Virus PRRS giờ đây là bệnh địa phương ở hầu hết các quốc gia sản xuất thịt lợn, và được xem là một trong những bệnh thiệt hại về mặt kinh tế gây ảnh hưởng đến ngành công nghiệp thịt lợn trên toàn cầu. Hơn nữa, kiểu gen có độ độc cao đã được phân lập ở Trung Quốc và các nước lân cận, và kiểu gen này thường được đề cập là kiểu gen Bắc Mỹ.

Mặc dù những hiểu biết về mặt sinh học của PRRSV là rất hữu ích song việc kiểm soát virus này vẫn còn khó khăn. Dùng vaccin cho động vật trong trường hợp này

đã được chứng minh là hiệu quả không cao. PRRS thường xuất hiện lại ở những đàn vật nuôi đã được gây miễn dịch, và hầu hết các chiến dịch dùng vacxin PRRSV trên trang trại cuối cùng vẫn thất bại trong việc kiểm soát bệnh này.

Không bị ràng buộc bởi lý thuyết này, lây nhiễm lợn bằng PRRSV kiểu đại hoặc dùng vacxin cho lợn ở dạng sống được làm yếu mầm bệnh này đáng tiếc là chỉ gây ra việc sản sinh nhiều kháng thể không trung hòa. Trong khoảng thời gian này, ví dụ, chỉ lượng giới hạn interferon (IFN)- γ (các tế bào tiết được tạo ra). Do đó, PRRSV vốn đã dường như kích thích đáp ứng miễn dịch không cân bằng được cho thấy là miễn dịch thể dịch (dựa trên kháng thể) luôn nhiều, và đáp ứng giống IFN- γ tế bào T trợ giúp (T helper-Th) bảo vệ tiềm năng nhưng hạn chế và thay đổi. Một đặc trưng của lây nhiễm PRRSV mà có khả năng nhất góp phần vào việc phát triển không cân bằng miễn dịch thích nghi là thiếu đáp ứng miễn dịch tự nhiên thích hợp. Thông thường, các tế bào lây nhiễm virus tiết interferon typ I “type I interferon-IFN” (bao gồm IFN- α và IFN- β), bảo vệ các tế bào bên cạnh không bị lây nhiễm. Ngoài ra, IFN typ I giải phóng ra tương tác với nhóm nhỏ các tế bào T nguyên thủy để thúc đẩy sự chuyển hóa chúng thành các tế bào tiết IFN typ II đặc hiệu virus (IFN- γ). Ngược lại, đáp ứng IFN- α của lợn đối với sự tiếp xúc PRRSV là gần như không có. Sự kích thích sản sinh IFN- α không hiệu quả bằng mầm bệnh sẽ có ảnh hưởng đáng kể lên bản chất của đáp ứng miễn dịch thích ứng của vật chủ, vì IFN- α điều hòa ngược sự biểu hiện gen IFN- γ . Do đó, xytokin trước đó điều khiển con đường chiếm ưu thế thúc đẩy sự phát triển miễn dịch thích nghi, cụ thể là, các đáp ứng IFN- γ qua trung gian tế bào T và bảo vệ miễn dịch kháng virus tối đa.

Về mặt này, đã trở nên rõ ràng rằng sự liên hệ có thể có giữa đáp ứng miễn dịch tự nhiên và miễn dịch thích ứng khi xảy ra lây nhiễm virus thông qua loại tế bào sợi nhánh đặc hiệu mà có khả năng sản xuất khối lượng lớn interferon typ I, và đóng vai trò quan trọng trong việc phân cực chức năng tế bào T. Cụ thể là, loại tế bào sợi nhánh đáng chú ý nhưng ít xuất hiện, tế bào sợi nhánh tương bào (plasmacytoid dendritic cell-PDC), cũng được biết là tế bào sản sinh IFN- α/β tự nhiên, đóng vai trò quan trọng trong miễn dịch chống virus bởi khả năng của nó gây ra sự biệt hóa tế bào T nguyên thể thành các tế bào tiết IFN- γ . Mặc dù hiếm, PDC là các tế bào sản xuất IFN- α hết sức tiềm năng, với mỗi tế bào có khả năng sản sinh từ 3-10 pg IFN- α để đáp ứng lại virus. Ngược lại, các bạch cầu đơn nhân to sản sinh ít hơn từ 5- đến 10-lần IFN- α trên mỗi tế bào. Phenotyp

này và một số đặc điểm sinh học của PDC lợn đã được mô tả (Summerfield *et al.*, 2003, Immunology 110:440). Các nghiên cứu gần đây đã xác định rằng PRRSV không kích thích PDC của lợn để tiết IFN- α (Calzada *et al.*, 2010, Veterinary Immunology và Immunopathology 135:20).

Thực tế này, kết hợp với quan sát rằng bổ sung ngoại sinh IFN- α ở thời điểm dùng vaccin đã được nhận thấy là cải thiện cường độ đáp ứng IFN- γ đặc hiệu PRRSV (W.A. Meier *et al.*, Vet. Immunol. Immunopath. 102, pp 299-314, 2004), làm nổi bật vai trò chủ chốt này của IFN- α trong quá trình lây nhiễm virus này ở lợn. Dựa trên vai trò quan trọng rõ rệt của IFN- α đối với sự phát triển của miễn dịch bảo vệ, quan trọng là xác định khả năng của các kho virus PRRS khác nhau để kích thích và/hoặc ức chế sự sản sinh IFN- α . Do đó, có nhu cầu cấp thiết về các vaccin virus sống được cải biến cải thiện và mới để bảo vệ kháng lại PRRS. Như được mô tả dưới đây, rõ ràng rằng virus có nguồn gốc từ dòng đơn tính cADN lây nhiễm mới, pCMV-S-P129-PK, và các virus khác, có phenotyp khác với virus P129 kiểu dại hoặc hai vaccin PRRS sống cải biến có trên thị trường. Không bị giới hạn như lý thuyết, sáng chế đề xuất vaccin dễ dàng gây ra đáp ứng miễn dịch dựa trên tế bào kháng lại virus này, và đưa ra thể hệ vaccin PRRS mới và hiệu quả.

Bản chất kỹ thuật của sáng chế

Theo phương án thứ nhất, sáng chế đề xuất phân tử polynucleotit phân lập chứa trình tự ADN mã hóa phân tử ARN lây nhiễm mã hóa virus PRRS thường được cải biến di truyền để làm vaccin, nó gây ra đáp ứng miễn dịch bảo vệ hiệu quả kháng lại virus PRRS ở lợn. Trong các khía cạnh nhất định, sáng chế đề xuất trình tự ADN như được nêu ra ở đây bao gồm trình tự nêu trong SEQ ID NO.:1, SEQ ID NO.:2, SEQ ID NO.:3, SEQ ID NO.:4, hoặc SEQ ID: NO:6, hoặc trình tự có ít nhất 70% tương đồng, ưu tiên là 80% giống, và ưu tiên hơn là 85%, 90%, 95%, 96%, 97%, 98% hoặc 99% tương đồng.

Trong các phương án nhất định, sáng chế đề xuất plasmit chứa phân tử polynucleotit phân lập như được nêu ra ở đây và gen khởi đầu có khả năng phiên mã phân tử polynucleotit này trong tế bào chủ thích hợp. Theo phương án khác, PRRS Bắc Mỹ hoặc Trung Quốc mã hóa trình tự plasmit ở đây mã hóa tiếp một hoặc nhiều epitop kháng nguyên khác loại có thể phát hiện. Sáng chế đề xuất tế bào chủ chuyển nhiễm để đưa vào plasmit được nêu ở đây.

Theo khía cạnh khác, sáng chế đề xuất vaccin bảo vệ lợn khỏi bị lây nhiễm virus PRRS. Vaccin này có thể chứa virus PRRS Bắc Mỹ hoặc Trung Quốc được mã hóa bởi phân tử ARN lây nhiễm, phân tử ARN lây nhiễm, hoặc plasmid, mỗi loại này được mã hóa bởi phân tử polynucleotit phân lập như được nêu ra ở đây. Theo khía cạnh khác nữa, vaccin chứa vector virus mang polynucleotit ở đây. Vaccin được nêu ra ở đây có thể tùy ý chứa chất mang dùng cho virus sử dụng trong thú y. Theo một khía cạnh quan trọng, vaccin này có hiệu quả ức chế interferon- α giảm khi so với virus PRRS P129 kiểu đại (xem ATCC 203488, 203489, bằng sáng chế Mỹ số 6,500,662).

Theo một phương án, sáng chế đề xuất kit chẩn đoán chứa phân tử polynucleotit mà phân biệt (thử nghiệm được gọi là DIVA) giữa lợn bị lây nhiễm tự nhiên chủng virus PRRS và lợn được dùng vaccin sống cải biến như được nêu ở đây.

Theo các phương án khác, sáng chế đề xuất phương pháp bảo vệ lợn khỏi bị nhiễm chủng virus PRRS bao gồm dùng cho con vật này lượng vaccin cho hiệu quả bảo vệ miễn dịch như được nêu ở đây.

Phương án nữa và ưu tiên theo sáng chế bao gồm virus gây hội chứng rối loạn sinh sản và hô hấp ở lợn (PRRS) phân lập, hoặc trình tự polynucleotit mã hóa nó, trong đó protein này được mã hóa bởi ORF1a được chọn từ nhóm gồm các trình tự chứa trình tự axit amin bất kỳ sau đây, trong đó các góc gạch chân được cho là mới: AMANVYD (SEQ ID NO: 9); IGHNAVM (SEQ ID NO: 12); TVPDGNC (SEQ ID NO: 15); CWWYLFD (SEQ ID NO: 18); HGVHGKY (SEQ ID NO: 21); AAKVDQY (SEQ ID NO: 24); PSATDTs (SEQ ID NO: 27); LNSLSK (SEQ ID NO: 30); APMCQDE (SEQ ID NO: 33); CAPTGMD (SEQ ID NO: 36); PKVAKVS (SEQ ID NO: 39); AGEIVGV (SEQ ID NO: 42); ADFNPEK (SEQ ID NO: 45); và QTPILGR (SEQ ID NO: 48). Theo phương án ưu tiên nữa của sáng chế, sáng chế đề xuất PRRS Bắc Mỹ hoặc Trung Quốc phân lập chứa trình tự nêu trên bất kỳ trong protein được mã hóa từ ORF1a, bao gồm sự kết hợp bất kỳ (2, 3, 4... lên đến 17) trình tự được xác định này.

Sáng chế đề xuất tiếp virus gây hội chứng rối loạn sinh sản và hô hấp ở lợn (PRRS) phân lập, trong đó protein của nó được mã hóa bởi ORF1a được chọn từ nhóm gồm các trình tự axit amin này chứa bất kỳ trong số: ANV (xem trình tự nêu trong SEQ ID NO: 9); HNA (xem trình tự nêu trong SEQ ID NO: 12); PDG (xem trình tự nêu trong SEQ ID NO: 15); WYL (xem trình tự nêu trong SEQ ID NO: 18); VHG (xem trình tự nêu trong SEQ ID NO: 21); KVD (xem trình tự nêu trong SEQ ID NO: 24); ATD (xem

trình tự nêu trong SEQ ID NO: 27); SLL (xem trình tự nêu trong SEQ ID NO: 30); MCQ (xem trình tự nêu trong SEQ ID NO: 33); PTG (xem trình tự nêu trong SEQ ID NO: 36); VAK (xem trình tự nêu trong SEQ ID NO: 39); EIV (xem trình tự nêu trong SEQ ID NO: 42); FNP (xem trình tự nêu trong SEQ ID NO: 45); và PIL (xem trình tự nêu trong SEQ ID NO: 48), bao gồm các sự kết hợp bất kỳ (2, 3, 4... lên đến 17) trình tự được xác định này.

Theo phương án ưu tiên nữa, sáng chế đề xuất PRRS Bắc Mỹ hoặc Trung Quốc phân lập, trong đó không kể đến tính tương đồng của các vị trí trong trình tự axit amin hoặc nucleotit đặc hiệu khác bất kỳ ở điểm bất kỳ trong polynucleotit mã hóa virus này hoặc các protein được mã hóa từ đó, protein virus ORF1a chứa:

- (a) các axit amin đặc hiệu sau bất kỳ trong các trình tự được nêu,
axit amin N trong trình tự axit amin ANV (xem trình tự nêu trong SEQ ID NO: 9);
axit amin N trong trình tự axit amin HNA (xem trình tự nêu trong SEQ ID NO: 12);
axit amin D trong trình tự axit amin PDG (xem trình tự nêu trong SEQ ID NO: 15);
axit amin Y trong trình tự axit amin WYL (xem trình tự nêu trong SEQ ID NO: 18);
axit amin H trong trình tự axit amin VHG (xem trình tự nêu trong SEQ ID NO: 21);
axit amin V trong trình tự axit amin KVD (xem trình tự nêu trong SEQ ID NO: 24);
axit amin T trong trình tự axit amin ATD (xem trình tự nêu trong SEQ ID NO: 27);
axit amin L trong trình tự axit amin SLL (xem trình tự nêu trong SEQ ID NO: 30).
axit amin C trong trình tự axit amin MCQ (xem trình tự nêu trong SEQ ID NO: 33);
axit amin T trong trình tự axit amin PTG (xem trình tự nêu trong SEQ ID NO: 36);

axit amin **A** trong trình tự axit amin VAK (xem trình tự nêu trong SEQ ID NO: 39);

axit amin **I** trong trình tự axit amin EIV (xem trình tự nêu trong SEQ ID NO: 42);

axit amin **N** trong trình tự axit amin FNP (xem trình tự nêu trong SEQ ID NO: 45); và

axit amin **I** trong trình tự axit amin PIL (xem trình tự nêu trong SEQ ID NO: 48), để đưa vào các sự kết hợp bất kỳ (2, 3, 4... lên đến 17) các trình tự được nêu này, hoặc

(b) chứa các axit amin riêng lẻ được gạch chân cụ thể này trong các trình tự peptit ORF1a ba gốc được chỉ định của các virus PRRS Bắc Mỹ hoặc Trung Quốc khác bất kỳ tương ứng với các trình tự 3 gốc như được chỉ định trên đây, không kể là các trình tự axit amin 3 gốc đặc hiệu khác này có thể có một hoặc hai sự thay đổi trình tự amino bổ sung, nhưng vẫn được nhận biết là tương ứng với các trình tự được nêu trên đây. Cho mục đích của phương án này của sáng chế, “tương ứng” nghĩa là các trình tự liên quan có thể được duỗi thẳng tối ưu sử dụng thuật toán BLOSUM như được mô tả trong Henikoff et al. Proc Natl. Acad. Sci., USA, 89, pp. 10915-10919, 1992.

Theo phương án ưu tiên nữa của sáng chế, sáng chế đề xuất virus gây hội chứng rối loạn sinh sản và hô hấp ở lợn (PRRS) phân lập trong đó protein của nó được mã hóa bởi ORF1a có trình tự axit amin chứa một hoặc nhiều biến thể (a), (b), (c) và (d), trong đó mỗi biến thể này được xác định như sau:

biến thể (a),

axit amin **N** trong trình tự axit amin ANV (xem trình tự nêu trong SEQ ID NO: 9);

axit amin **N** trong trình tự axit amin HNA (xem trình tự nêu trong SEQ ID NO: 12);

axit amin **D** trong trình tự axit amin PDG (xem trình tự nêu trong SEQ ID NO: 15),

axit amin **Y** trong trình tự axit amin WYL (xem trình tự nêu trong SEQ ID NO: 18);

axit amin **H** trong trình tự axit amin VHG (xem trình tự nêu trong SEQ ID NO: 21), hoặc bộ nhỏ bất kỳ biến thể (a);

biến thể (b),

axit amin V trong trình tự axit amin KVD (xem trình tự nêu trong SEQ ID NO: 24);

axit amin T trong trình tự axit amin ATD (xem trình tự nêu trong SEQ ID NO: 27);

axit amin L trong trình tự axit amin SLL (xem trình tự nêu trong SEQ ID NO: 30).

axit amin C trong trình tự axit amin MCQ (xem trình tự nêu trong SEQ ID NO: 33), hoặc bộ nhỏ bất kỳ biến thể (b);

biến thể (c),

axit amin T trong trình tự axit amin PTG (xem trình tự nêu trong SEQ ID NO: 36);

axit amin A trong trình tự axit amin VAK (xem trình tự nêu trong SEQ ID NO: 39), hoặc bộ nhỏ bất kỳ biến thể (c); và

biến thể (d),

axit amin I trong trình tự axit amin EIV (xem trình tự nêu trong SEQ ID NO: 42);

axit amin N trong trình tự axit amin FNP (xem trình tự nêu trong SEQ ID NO: 45);

và axit amin I trong trình tự axit amin PIL (xem trình tự nêu trong SEQ ID NO: 20), hoặc bộ nhỏ bất kỳ biến thể (d) của nó.

Virut PRRS này có thể chứa thêm hai hoặc nhiều trong số năm trình tự axit amin được xác định trong biến thể (a), và/hoặc hai hoặc nhiều trong số bốn trình tự axit amin được xác định trong biến thể (b), và/hoặc hai trình tự axit amin được xác định trong biến thể (c), và/hoặc hai hoặc nhiều trong số ba trình tự axit amin được xác định trong biến thể (d).

Sáng chế cũng đề xuất plasmit có khả năng chuyển nhiễm trực tiếp tế bào chủ thích hợp và biểu hiện virut gây hội chứng rối loạn sinh sản và hô hấp ở lợn (PRRS) từ tế bào chủ thích hợp được chuyển nhiễm như vậy, plasmit này chứa: (a) trình tự ADN mã hóa phân tử ARN lây nhiễm mã hóa virut PRRS, và (b) gen khởi đầu có khả năng phiên mã phân tử ARN lây nhiễm này, trong đó protein được mã hóa bởi ORF1a của virut này có trình tự axit amin chứa:

(1) axit amin **N** trong trình tự axit amin **ANV** (xem trình tự nêu trong SEQ ID NO: 9);

axit amin **N** trong trình tự axit amin **HNA** (xem trình tự nêu trong SEQ ID NO: 12);

axit amin **D** trong trình tự axit amin **PDG** (xem trình tự nêu trong SEQ ID NO: 15),

axit amin **Y** trong trình tự axit amin **WYL** (xem trình tự nêu trong SEQ ID NO: 18);

axit amin **H** trong trình tự axit amin **VHG** (xem trình tự nêu trong SEQ ID NO: 21), hoặc bộ nhỏ bất kỳ của nó; và/hoặc

(2) axit amin **V** trong trình tự axit amin **KVD** (xem trình tự nêu trong SEQ ID NO: 24);

axit amin **T** trong trình tự axit amin **ATD** (xem trình tự nêu trong SEQ ID NO: 27);

axit amin **L** trong trình tự axit amin **SLL** (xem trình tự nêu trong SEQ ID NO: 30).

axit amin **C** trong trình tự axit amin **MCQ** (xem trình tự nêu trong SEQ ID NO: 33), hoặc bộ nhỏ bất kỳ của nó; và/hoặc

(3) axit amin **T** trong trình tự axit amin **PTG** (xem trình tự nêu trong SEQ ID NO: 36);

axit amin **A** trong trình tự axit amin **VAK** (xem trình tự nêu trong SEQ ID NO: 39), hoặc bộ nhỏ bất kỳ của nó; và/hoặc

(4) axit amin **I** trong trình tự axit amin **EIV** (xem trình tự nêu trong SEQ ID NO: 42);

axit amin **N** trong trình tự axit amin **FNP** (xem trình tự nêu trong SEQ ID NO: 45);

và axit amin **I** trong trình tự axit amin **PIL** (xem trình tự nêu trong SEQ ID NO: 48), hoặc bộ nhỏ bất kỳ của nó.

Cần hiểu rằng ORF1a mã hóa polyprotein có chức năng proteaza, và ORF1b mã hóa polyprotein có chức năng replicaza (ARN polymeraza) và helicaza. Thông tin bổ sung liên quan đến các chức năng cho các protein được mã hóa từ các khung đọc mở ORF (open reading frame) khác nhau của PRRS có thể được tìm thấy, ví dụ, trong bảng

sáng chế Mỹ số 7,132,106. Cũng xem bằng sáng chế Mỹ số 7,544,362 về chức năng của ORF7, và các khung đọc mở khác. Như sẽ được hiểu trong ngành, các protein mã hóa bởi ORF1 được hy vọng là có các chức năng bổ sung, đã biết và chưa biết, và các sự thay đổi axit amin mới hữu ích trong thực hiện sáng chế không bị giới hạn thông qua hiệu quả của nó đối với một chức năng cụ thể bất kỳ của các protein mã hóa bởi ORF1 này.

Theo các khía cạnh được ưu tiên nữa, plasmit này chứa gen khởi đầu là gen khởi đầu của tế bào nhân chuẩn có khả năng cho phép khởi động ADN trong các tế bào nhân chuẩn đích, hoặc gen khởi đầu của tế bào nhân sơ hoặc thực khuẩn thể có khả năng điều khiển sự dịch mã *in vitro* plasmit này. Tương tự sáng chế đề xuất phương pháp tạo ra virus PRRS, phương pháp này bao gồm chuyển nhiễm vào tế bào chủ thích hợp plasmit thích hợp và thu virus PRRS được tạo ra bằng tế bào được chuyển nhiễm này.

Do đó, theo phương án cụ thể và ưu tiên, sáng chế đề xuất phân tử polynucleotit phân lập chứa trình tự ADN mã hóa phân tử ARN lây nhiễm mã hóa virus PRRS Bắc Mỹ, trong đó trình tự ADN này được chọn từ nhóm gồm:

(a) SEQ ID NO:6;

(b) trình tự có ít nhất 85% tương đồng với trình tự ADN nêu trong (a) trong đó protein được mã hóa bởi ORF1a của nó có trình tự axit amin chứa:

từ nhóm (b) (1)

axit amin **N** trong trình tự axit amin ANV (xem trình tự nêu trong SEQ ID NO: 9);

axit amin **N** trong trình tự axit amin HNA (xem trình tự nêu trong SEQ ID NO: 12);

axit amin **D** trong trình tự axit amin PDG (xem trình tự nêu trong SEQ ID NO: 15),

axit amin **Y** trong trình tự axit amin WYL (xem trình tự nêu trong SEQ ID NO: 18);

axit amin **H** trong trình tự axit amin VHG (xem trình tự nêu trong SEQ ID NO: 21), hoặc bộ nhỏ bất kỳ của nó; và/hoặc

từ nhóm (b) (2)

axit amin **V** trong trình tự axit amin KVD (xem trình tự nêu trong SEQ ID NO: 24);

axit amin T trong trình tự axit amin ATD (xem trình tự nêu trong SEQ ID NO: 27);

axit amin L trong trình tự axit amin SLL (xem trình tự nêu trong SEQ ID NO: 30).

axit amin C trong trình tự axit amin MCQ (xem trình tự nêu trong SEQ ID NO: 33), hoặc bộ nhỏ bất kỳ của nó; và/hoặc

từ nhóm (b)(3)

axit amin T trong trình tự axit amin PTG (xem trình tự nêu trong SEQ ID NO: 36);

axit amin A trong trình tự axit amin VAK (xem trình tự nêu trong SEQ ID NO: 39), hoặc bộ nhỏ bất kỳ của nó; và/hoặc

từ nhóm (b)(4)

axit amin I trong trình tự axit amin EIV (xem trình tự nêu trong SEQ ID NO: 42);

axit amin N trong trình tự axit amin FNP (xem trình tự nêu trong SEQ ID NO: 45);

và axit amin I trong trình tự axit amin PIL (xem trình tự nêu trong SEQ ID NO: 20), hoặc bộ nhỏ bất kỳ của nó; và

(c) trình tự ADN lai với phần bổ sung của trình tự ADN nêu trong (a) hoặc (b) dưới các điều kiện nghiêm ngặt cao bao gồm lai với thiết lọc liên kết ADN trong 0,5 M NaHPO₄, 7% SDS, 1mM EDTA ở 65 độ C, và rửa trong 0,1 SSC/0,1%SDS ở 68 độ C.

Sáng chế cũng đề xuất tế bào chủ được chuyển nhiễm các phân tử polynucleotit và đề xuất các vacxin để bảo vệ lợn kháng lại sự lây nhiễm virus PRRS, vacxin này chứa: (a) virus PRRS Bắc Mỹ được cải biến di truyền được mã hóa bởi các phân tử polynucleotit được đề cập trên đây, hoặc (b) phân tử lây nhiễm này, hoặc (c) phân tử polynucleotit này ở dạng plasmit, hoặc (d) vector virus chứa phân tử polynucleotit này, trong đó virus PRRS có khả năng gây ra đáp ứng miễn dịch bảo vệ hữu hiệu kháng lại sự lây nhiễm virus PRRS, với lượng hữu hiệu để tạo ra sự bảo vệ miễn dịch kháng lại sự lây nhiễm, và chất mang phù hợp để sử dụng trong thú y.

Sáng chế cũng đề xuất trình tự polynucleotit ARN tương ứng với (nghĩa là có các trình tự mã hóa bazơ bổ sung):

(a) trình tự ADN có trình tự nêu trong SEQ ID NO:6;

(b) trình tự ADN có ít nhất 85% tương đồng với trình tự ADN nêu trong (a) trong đó protein được mã hóa bởi ORF1a của nó có trình tự axit amin chứa bất kỳ trong số sau, và sự kết hợp bất kỳ trong số sau:

axit amin **N** trong trình tự axit amin ANV (xem trình tự nêu trong SEQ ID NO: 9);

axit amin **N** trong trình tự axit amin HNA (xem trình tự nêu trong SEQ ID NO: 12);

axit amin **D** trong trình tự axit amin PDG (xem trình tự nêu trong SEQ ID NO: 15),

axit amin **Y** trong trình tự axit amin WYL (xem trình tự nêu trong SEQ ID NO: 18);

axit amin **H** trong trình tự axit amin VHG (xem trình tự nêu trong SEQ ID NO: 21),

axit amin **V** trong trình tự axit amin KVD (xem trình tự nêu trong SEQ ID NO: 24);

axit amin **T** trong trình tự axit amin ATD (xem trình tự nêu trong SEQ ID NO: 27);

axit amin **L** trong trình tự axit amin SLL (xem trình tự nêu trong SEQ ID NO: 30).

axit amin **C** trong trình tự axit amin MCQ (xem trình tự nêu trong SEQ ID NO: 33),

axit amin **T** trong trình tự axit amin PTG (xem trình tự nêu trong SEQ ID NO: 36);

axit amin **A** trong trình tự axit amin VAK (xem trình tự nêu trong SEQ ID NO: 39),

(axit amin **I** trong trình tự axit amin EIV (xem trình tự nêu trong SEQ ID NO: 42);

axit amin **N** trong trình tự axit amin FNP (xem trình tự nêu trong SEQ ID NO: 45);

và axit amin **I** trong trình tự axit amin PIL (xem trình tự nêu trong SEQ ID NO: 20), hoặc

(c) trình tự ADN lai với phần bổ sung của trình tự ADN nêu trong (a) hoặc (b) dưới các điều kiện nghiêm ngặt cao bao gồm lai với thiết bị lai được liên kết ADN trong 0,5 M NaHPO₄, 7% SDS, 1mM EDTA ở 65 độ C, và rửa trong 0,1 SSC/0,1%SDS ở 68 độ C.

Do đó, sáng chế cũng đề xuất kit chẩn đoán chứa các phân tử polynucleotit phân biệt giữa các con lợn bị nhiễm tự nhiên chủng virus PRRS và các con lợn dùng vaccin bằng vaccin của sáng chế, vaccin này (virus) ưu tiên là chứng minh tác dụng ức chế interferon- α giảm khi so với chủng virus PRRS P129 kiểu dại (SEQ ID NO:5).

Mô tả vắn tắt các bảng và hình vẽ

Bảng 1 thể hiện các dòng đơn tính cADN lây nhiễm và các virus tương ứng có nguồn gốc từ sự lây nhiễm các tế bào PK-9.

Bảng 2 thể hiện hiệu quả ức chế interferon- α của virus PRRS kiểu dại và các dẫn xuất được thích ứng để sinh trưởng trong nuôi cấy tế bào.

Bảng 3 mô tả hiệu quả ức chế interferon- α của virus PRRS kiểu dại P129 và các dẫn xuất được cải biến di truyền của nó được thích ứng để sinh trưởng trong các tế bào PK-9 biểu hiện CD163.

Bảng 4 thể hiện tác dụng ức chế interferon- α giảm của các virus P129-PK-FL và P129-PK-d43/44 khi so sánh với virus P129 kiểu dại và vaccin PRRS Ingelvac.

Bảng 5 mô tả thiết kế thí nghiệm được thiết kế để đánh giá độ an toàn và hiệu quả của virus vaccin.

Bảng 6 thể hiện tất cả các sự khác biệt nucleotit và các sự khác biệt axit amin thu được giữa P129 lần cấy truyền 0 và P129-PK-FL lần cấy truyền 17, bởi vị trí trong hệ gen.

Bảng 7 thể hiện bảng tóm tắt về các sự khác biệt nucleotit và axit amin giữa P129 lần cấy truyền 0 và P129-PK-FL lần cấy truyền 17, bởi protein virus.

Bảng 8 thể hiện tất cả các sự khác biệt nucleotit và các sự khác biệt axit amin thu được giữa hệ gen PRRSV được tìm thấy trong dòng đơn tính cADN lây nhiễm pCMV-S-P129 và pCMV-S-P129-PK17-FL, bởi vị trí trong hệ gen.

Bảng 9 và 10 thể hiện các sự thay đổi axit amin trong phenotyp của virus cấy truyền lần 52 (SEQ ID NO:6).

Fig. 1 thể hiện nhiệt độ trực tràng sau khi dùng vaccin.

Fig. 2 thể hiện nhiệt độ trực tràng sau khi kích thích bằng PRRSV gây độc NADC20.

Fig. 3 thể hiện trọng lượng cơ thể sau khi dùng vaccin và sau khi kích thích.

Fig. 4 thể hiện dữ liệu sau kích thích phần trăm phổi bị các tổn thương gây ra do PRRS.

Fig. 5 thể hiện các điểm số đánh giá phổi sau khi kích thích (lung assessment score-LAS) về độ nghiêm trọng của các tổn thương quan sát được.

Fig. 6 là đồ thị mô tả các mức kháng thể kháng-PRRSV trong huyết thanh sau khi dùng vaccin và sau khi kích thích (tỷ lệ S/P ELISA).

Fig. 7 là đồ thị về lượng virus sau kích thích trong huyết thanh (log TCID50/ml trên các tế bào PAM)

Fig. 8 là sơ đồ về các phương pháp sử dụng để thu được các vaccin chứa trình tự nêu trong từ SEQ ID NO:1 đến SEQ ID NO:6, như được bộc lộ ở đây.

Mô tả vắn tắt các trình tự

SEQ ID NO:1 là hệ gen hoàn chỉnh P129-PK-FL lần cấy truyền 17.

SEQ ID NO:2 là hệ gen hoàn chỉnh P129-PK-d43/44 lần cấy truyền 17.

SEQ ID NO:3 là hệ gen hoàn chỉnh P129-PK-FL lần cấy truyền 24.

SEQ ID NO:4 là hệ gen hoàn chỉnh P129-PK-d43/44 lần cấy truyền 34.

SEQ ID NO:5 là hệ gen hoàn chỉnh P129 lần cấy truyền 0.

SEQ ID NO:6 là hệ gen hoàn chỉnh P129 lần cấy truyền 52.

Mô tả chi tiết sáng chế

Như được sử dụng trong bản mô tả và yêu cầu bảo hộ, các dạng số ít bao gồm tham chiếu đến số nhiều trừ khi ngữ cảnh được nêu rõ ràng khác. Do đó, ví dụ, tham chiếu đến "phương pháp" bao gồm một hoặc nhiều phương pháp, và/hoặc các bước là dạng như được mô tả ở đây trở nên rõ ràng đối với người có hiểu biết trung bình trong lĩnh vực khi đọc bản mô tả này.

Trừ khi được định nghĩa khác, tất cả các thuật ngữ kỹ thuật và khoa học được sử dụng ở đây có nghĩa như được hiểu thông thường bởi người có hiểu biết trung bình thuộc lĩnh vực của sáng chế. Mặc dù các phương pháp và các nguyên liệu bất kỳ tương tự hoặc tương đương với các phương pháp và nguyên liệu được mô tả ở đây có thể được

sử dụng để thực hiện hoặc thử nghiệm sáng chế, các phương pháp và nguyên liệu ưu tiên giờ đây được mô tả.

Thực hiện sáng chế này sẽ sử dụng, trừ khi được nêu cụ thể ngược lại, các phương pháp thông thường về virus học, miễn dịch, vi sinh vật, sinh học phân tử và các kỹ thuật ADN tái tổ hợp thuộc kỹ năng trong ngành, nhiều trong số các phương pháp này được mô tả dưới đây cho mục đích minh họa. Các kỹ thuật này được giải thích đầy đủ trong tài liệu này. Ví dụ xem, Sambrook, *et al.* *Molecular Cloning: A Laboratory Manual* (2nd Edition, 1989); Maniatis *et al.* *Molecular Cloning: A Laboratory Manual* (1982); *DNA Cloning: A Practical Approach*, vol. I & II (D. Glover, ed.); *Oligonucleotide Synthesis* (N. Gait, ed., 1984); *Nucleic Acid Hybridization* (B. Hames & S. Higgins, eds., 1985); *Transcription and Translation* (B. Hames & S. Higgins, eds., 1984); *Animal Cell Culture* (R. Freshney, ed., 1986); Perbal, *A Practical Guide to Molecular Cloning* (1984).

"Virus PRRS Bắc Mỹ" nghĩa là virus PRRS bất kỳ có các đặc điểm di truyền gần gũi với mẫu phân lập virus PRRS Bắc Mỹ, như, nhưng không bị giới hạn đến virus PRRS được phân lập đầu tiên ở Hoa Kỳ vào khoảng đầu những năm 90 (ví dụ xem, Collins, J. E., *et al.*, 1992, *J. Vet. Diagn. Invest.* 4:117-126); mẫu phân lập virus PRRS Bắc Mỹ MN-1b (Kwang, J. *et al.*, 1994, *J. Vet. Diagn. Invest.* 6:293-296); chủng Quebec LAF-exp91 của PRRS (Mardassi, H. *et al.*, 1995, *Arch. Virol.* 140:1405-1418); và mẫu phân lập virus PRRS Bắc Mỹ VR 2385 (Meng, X.-J *et al.*, 1994, *J. Gen. Virol.* 75:1795-1801). Các đặc điểm di truyền đề cập đến sự tương đồng trình tự nucleotide hệ gen và sự tương đồng trình tự axit amin cùng có ở các chủng virus PRRS Bắc Mỹ. Các chủng virus PRRS Trung Quốc thường có độ tương đồng khoảng từ 80-93% trình tự nucleotide với chủng Bắc Mỹ.

"Virus PRRS Châu Âu" đề cập đến chủng virus PRRS có đặc điểm di truyền gần gũi với virus PRRS mà được phân lập đầu tiên ở châu Âu vào khoảng năm 1991 (ví dụ xem, Wensvoort, G., *et al.*, 1991, *Vet. Q.* 13:121-130). "Virus PRRS châu Âu" đôi khi cũng được đề cập trong ngành là "virus Lelystad".

"Đáp ứng miễn dịch bảo vệ hữu hiệu", "bảo vệ miễn dịch", và các thuật ngữ tương tự, cho mục đích sáng chế này, nghĩa là đáp ứng miễn dịch trực tiếp kháng lại một hoặc nhiều epitop kháng nguyên của mầm bệnh để bảo vệ kháng lại sự lây nhiễm của mầm bệnh này ở động vật dùng vắc xin. Cho mục đích của sáng chế này, bảo vệ kháng

lại sự lây nhiễm mầm bệnh bao gồm không chỉ sự ngăn ngừa hoàn toàn lây nhiễm, mà còn giảm mức độ hoặc tốc độ lây nhiễm mầm bệnh này ở mức có thể nhận biết được, hoặc giảm độ nghiêm trọng của bệnh này hoặc triệu chứng bất kỳ hoặc tình trạng bệnh do sự lây nhiễm mầm bệnh này ở mức có thể nhận biết được ở động vật dùng vaccin khi so sánh với động vật bị lây nhiễm không dùng vaccin. Đáp ứng miễn dịch bảo vệ hữu hiệu có thể được cảm ứng ở các con vật mà đã không bị lây nhiễm mầm bệnh này trước đó và/hoặc không bị lây nhiễm mầm bệnh này ở thời điểm dùng vaccin. Đáp ứng miễn dịch bảo vệ hữu hiệu có thể cũng được gây ra ở con vật đã bị lây nhiễm rồi mầm bệnh này ở thời điểm dùng vaccin.

Virut PRRS cải biến di truyền đã được "làm yếu" nếu nó ít độc hơn so với chủng ban đầu chưa được cải biến. Chủng "ít độc hơn" nếu chủng này thể hiện sự giảm có ý nghĩa thống kê một hoặc nhiều thông số xác định độ nghiêm trọng của bệnh. Các thông số này có thể bao gồm mức virus huyết, sốt, độ nghiêm trọng của tình trạng nguy cấp hô hấp, độ nghiêm trọng của các triệu chứng sinh sản, hoặc số lượng hoặc độ nghiêm trọng của các tổn thương phổi, v.v..

"Tế bào chủ có khả năng hỗ trợ sao chép virus PRRS" nghĩa là tế bào có khả năng tạo ra PRRS lây nhiễm khi lây nhiễm virus theo sáng chế. Các tế bào này bao gồm các tế bào lợn của dòng bạch cầu đơn nhân to/đại thực bào như các tế bào đại thực bào phẩy nang lợn và các dẫn xuất, các tế bào thận khỉ MA-104 và các dẫn xuất như các tế bào MARC-145, và các tế bào được lây nhiễm thụ thể cho virus PRRS. Thuật ngữ "tế bào chủ có khả năng hỗ trợ sao chép virus PRRS" có thể cũng bao gồm các tế bào trong lợn sống.

"Khung đọc mở", hoặc "ORF", như được sử dụng ở đây, nghĩa là trình tự nucleotit tối thiểu cần để mã hóa protein virus PRRS cụ thể không xen vào mã dừng.

"Lợn" và "heo" được sử dụng lẫn nhau ở đây và đề cập đến động vật là thành viên của họ *Suidae* như, ví dụ, lợn nhà. Thuật ngữ "virus PRRS", như được sử dụng ở đây, trừ khi được chỉ định khác, nghĩa là chủng virus Bắc Mỹ hoặc Châu Âu.

"PRRS" bao hàm các triệu chứng bệnh ở lợn gây ra bởi nhiễm virus PRRS. Các ví dụ về các triệu chứng này bao gồm, nhưng không bị giới hạn đến, sốt, sảy thai ở những con lợn cái mang thai, tình trạng nguy cấp hô hấp, tổn thương phổi, bỏ ăn, và gây chết ở những con lợn nhỏ. Như được sử dụng ở đây, virus PRRS mà là "không có khả

năng gây bệnh PRRS" đề cập đến virus có thể lây nhiễm lợn, nhưng không tạo ra các triệu chứng bất kỳ của bệnh này mà thường liên quan đến sự lây nhiễm PRRS ở lợn.

"Protein N" PRRSV hoặc "ORF7" như được sử dụng ở đây được định nghĩa là polypeptit được mã hóa bởi ORF7 của cả hai kiểu gen của virus PRRS Châu Âu và Bắc Mỹ. Các ví dụ về isotyp đặc hiệu của protein N hiện nay được biết là polypeptit có chiều dài 123 axit amin của nguyên mẫu phân lập VR2322 PRRS Bắc Mỹ được thông báo trong Ngân hàng gen có các số lưu giữ PRU87392, và protein N 128 gốc của nguyên mẫu phân lập PRRS Châu Âu Lelystad được thông báo trong Ngân hàng gen có số lưu giữ là A26843.

"Vùng NLS-1 protein N PRRSV" hoặc "vùng ORF7 NLS-1 PRRSV" đề cập đến tín hiệu định vị nhân "pat4" hoặc "nuc1" (Nakai & Kanehisa, 1992; Rowland & Yoo, 2003) chứa bốn axit amin bazơ liên tiếp (lysine hoặc arginine), hoặc ba gốc bazơ và histidine hoặc proline, được đặt trong khoảng 15 gốc đầu tận cùng N đầu tiên của protein N thành thực chức năng. Bằng cách ví dụ trình tự vùng VR2332 NLS-1 là KRKK và được đặt ở các gốc từ 9-12, trong khi trình tự mẫu phân lập Lelystad là KKKK và được đặt ở các gốc từ 10-13 của protein N.

"Vùng NLS-2 protein N PRRSV" hoặc "vùng NLS-2 ORF7 PRRSV" đề cập đến tín hiệu định vị nhân thứ hai trong protein N mà có thể thuộc một trong số hai dạng. Trong virus PRRS Bắc Mỹ NLS-2 có mô hình mà đã được chúng tôi chỉ định là motif "pat8", mà bắt đầu bằng proline sau đó không quá ba gốc đến trình tự năm gốc chứa ít nhất ba gốc bazơ (K hoặc R) trong số năm gốc (sự cải biến nhẹ motif "pat7" hoặc "nuc2" được mô tả bởi Nakai & Kanehisa, 1992; Rowland & Yoo, 2003).-Bằng cách ví dụ trình tự này được đặt ở protein N các gốc từ 41-47 của mẫu phân lập PRRSV Bắc Mỹ VR2332, và được biểu diễn bằng trình tự P...K. Trong các virus PRRS Châu Âu NLS-2 có motif "pat4" hoặc "nuc1", mà chứa đoạn kéo dài liên tục gồm bốn axit amin bazơ hoặc ba gốc bazơ liên kết với histidine hoặc proline (Nakai & Kanehisa, 1992; Rowland & Yoo, 2003). NLS-2 của mẫu phân lập PRRSV Châu Âu Lelystad được đặt ở các gốc từ 47-50 và được biểu diễn bằng trình tự K..K.

"Vùng NoLS protein N PRRSV" hoặc "vùng NoLS ORF7 PRRSV" đề cập đến tín hiệu định vị nhân có tổng chiều dài là khoảng 32 axit amin và kết hợp với vùng NLS-2 gần đầu tận cùng amino của nó. Bằng cách ví dụ trình tự vùng NoLS VR2332 được đặt ở các gốc từ 41-72 và được biểu diễn bằng trình tự P...R (Rowland & Yoo, 2003) và

trình tự mẫu phân lập Lelystad tương ứng được đặt ở các gốc từ 42-73 và được biểu diễn bằng trình tự P..R.

"Tế bào chủ chuyển nhiễm" thực tế nghĩa là tế bào chủ bất kỳ như được mô tả trong bằng sáng chế Mỹ số 5,600,662 khi được chuyển nhiễm ARN virus PRRS có thể tạo ra ít nhất vòng thứ nhất virion PRRS.

"Phân tử ADN lây nhiễm" cho mục đích sáng chế này, là phân tử ADN mã hóa các yếu tố cần thiết để hỗ trợ sao chép, phiên mã và dịch mã thành virion chức năng từ tế bào chủ thích hợp.

Tương tự, "phân tử polynucleotit phân lập" đề cập đến thành phần vật chất chứa phân tử phân tử polynucleotit theo sáng chế được tinh chế đến mức độ có thể nhận thấy bất kỳ từ trạng thái có trong tự nhiên của nó, nếu có.

Cho mục đích sáng chế này, trình tự nucleotit của phân tử polynucleotit thứ hai (ARN hoặc ADN) là "tương đồng" với trình tự nucleotit của phân tử polynucleotit thứ nhất, hoặc có "giống" với phân tử polynucleotit thứ nhất này, trong đó trình tự nucleotit của phân tử polynucleotit thứ hai này mã hóa polyaxitamin tương tự như trình tự nucleotit của phân tử polynucleotit thứ nhất dựa trên sự thoái hóa mã di truyền, hoặc khi nó mã hóa polyaxit amin mà tương đồng thích hợp với polyaxit amin được mã hóa bởi trình tự nucleotit của phân tử polynucleotit thứ nhất để hữu ích trong việc thực hiện sáng chế. Các trình tự polynucleotit tương đồng đề cập đến các sợi có nghĩa và đối nghĩa, và trong tất cả các trường hợp đề cập đến sợi bổ sung của các sợi bất kỳ này. Cho mục đích sáng chế này, phân tử polynucleotit là hữu ích để thực hiện sáng chế, và do đó là tương đồng hoặc có giống, trong đó nó có thể được sử dụng làm môi chẩn đoán để phát hiện sự có mặt của virus PRRS hoặc polynucleotit virus trong mẫu dịch hoặc mô của lợn bị lây nhiễm, ví dụ bằng kỹ thuật lai hoặc khuếch đại chuẩn. Nhìn chung, trình tự nucleotit của phân tử polynucleotit thứ hai là tương đồng với trình tự nucleotit của phân tử polynucleotit thứ nhất nếu có ít nhất khoảng 70% trình tự nucleotit tương đồng với trình tự nucleotit của phân tử polynucleotit thứ nhất dựa trên thuật toán BLASTN (National Center for Biotechnology Information, nói cách khác được biết là NCBI, (Bethesda, Maryland, USA) của United States National Institute of Health). Trong ví dụ cụ thể để tính toán theo thực hiện sáng chế, tham khảo BLASTP 2.2.6 [Tatusova TA và TL Madden, "BLAST 2 sequences- a new tool for comparing protein và nucleotit sequences." (1999) FEMS Microbiol Lett. 174:247-250.]. Vấn đề như sau: hai trình tự

axit amin được duỗi thẳng so sánh để tối ưu các điểm số duỗi thẳng sử dụng điểm bù mở khoảng cách là 10, điểm bù kéo dài khoảng cách là 0,1, và ma trận ghi điểm “blosum62” của Henikoff và Henikoff (Proc. Nat. Acad. Sci. USA 89:10915-10919. 1992). Phần trăm giống sau đó được tính toán là: Tổng số các cặp giống nhau x 100/ được chia cho chiều dài của trình tự dài hơn + số các khoảng cách được đưa vào trong trình tự dài hơn này để sắp thẳng hàng hai trình tự này.

Ưu tiên là, trình tự nucleotit tương đồng có ít nhất khoảng 75% trình tự nucleotit giống nhau, thậm chí ưu tiên hơn ít nhất khoảng 80%, 85%, 90%, 95%, 96%, 97%, 98% và 99% trình tự nucleotit giống nhau. Vì mã di truyền là thoái biến, trình tự nucleotit tương đồng có thể bao gồm một số thay đổi bazơ “câm” bất kỳ, ví dụ các sự thay thế nucleotit tuy nhiên là mã hóa cùng một axit amin.

Trình tự nucleotit tương đồng có thể chứa thêm các đột biến không câm, nghĩa là các sự thay thế, xóa bỏ hoặc bổ sung bazơ dẫn đến các sự khác biệt axit amin trong polyaxit amin được mã hóa miễn là trình tự này vẫn có ít nhất khoảng 70% giống với polyaxit amin được mã hóa bởi trình tự nucleotit thứ nhất hoặc nếu không thì là hữu ích để thực hiện sáng chế này. Về mặt này, các sự thay thế axit amin bảo toàn nhất định có thể được tạo ra mà thường được nhận biết không làm bất hoạt toàn bộ chức năng protein: như đối với axit amin tích điện dương (và ngược lại), lysin, arginin và histidin; đối với axit amin tích điện âm (và ngược lại), axit aspartic và axit glutamic; và đối với các nhóm nhất định axit amin tích điện trung hòa (và trong tất cả các trường hợp, cũng ngược lại), (1) alanin và serin, (2) asparagin, glutamin, và histidin, (3) xystein và serin, (4) glyxin và prolin, (5) isoloxin, loxin và valin, (6) methionin, loxin và isoloxin, (7) phenylalanin, methionin, loxin, và tyrosin, (8) serin và threonin, (9) tryptophan và tyrosin, (10) và ví dụ tyrosin, tyrptophan và phenylalanin. Axit amin có thể được phân loại theo các đặc điểm vật lý và sự đóng góp vào cấu trúc protein bậc hai và bậc ba. Sự thay thế bảo toàn được nhận biết trong ngành là sự thay thế một axit amin bằng một axit amin khác có các đặc điểm tương tự. Các sự thay thế bảo toàn điển hình có thể được tìm thấy trong WO 97/09433, trang 10, công bố 13.03. 1997 (PCT/GB96/02197, nộp ngày 6.09. 1996. Theo cách khác, các axit amin bảo toàn có thể được nhóm lại như được mô tả trong Lehninger, (Biochemistry, Second Edition; Worth Publishers, Inc. NY:NY (1975), pp. 71-77). Các sự thay đổi bảo toàn thích hợp bổ sung và áp dụng của nó được mô tả dưới đây.

Trình tự nucleotit tương đồng có thể được xác định bằng cách so sánh các trình tự nucleotit, ví dụ bằng cách sử dụng BLASTN, trên đây. Theo phương án khác,, các trình tự nucleotit tương đồng có thể được xác định bằng cách lai dưới các điều kiện chọn lọc. Ví dụ, trình tự nucleotit của phân tử polynucleotit thứ hai là tương đồng với SEQ ID NO:1 (hoặc trình tự poly nucleotit cụ thể khác bất kỳ) nếu nó lai với phần bổ sung của SEQ ID NO:1 dưới các điều kiện nghiêm ngặt vừa phải, ví dụ, lai với ADN được liên kết với thiết bị lọc trong NaHPO_4 0,5 M, 7% natri dodecyl sulfat (SDS), 1 mM EDTA ở 65°C , và rửa trong 0,2xSSC/0,1% SDS ở 42°C (xem Ausubel et al editors, *Protocols in Molecular Biology*, Wiley and Sons, 1994, pp. 6.0.3 đến 6.4.10), hoặc các điều kiện mà nói cách khác sẽ dẫn đến lai các trình tự mã hóa virus PRRS như được xác định dưới đây. Các cải biến trong các điều kiện lai có thể được xác định bằng kinh nghiệm hoặc được tính toán chính xác dựa trên chiều dài và phần trăm cặp bazơ guanosin/xytosin (GC) của mỗi. Các điều kiện lai có thể được tính toán như được mô tả trong Sambrook, *et al.*, (Eds.), Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York (1989), pp. 9.47 đến 9.51.

Theo phương án khác, trình tự nucleotit thứ hai là tương đồng với trình tự nêu trong SEQ ID NO:1 (hoặc trình tự khác bất kỳ của sáng chế) nếu nó lai với phần bổ sung của trình tự nêu trong SEQ ID NO:1 dưới các điều kiện nghiêm ngặt cao, ví dụ lai với ADN liên kết với thiết bị lọc trong 0,5 M NaHPO_4 , 7% SDS, 1 mM EDTA ở 65°C , và rửa trong 0,1xSSC/0,1% SDS ở 68°C , như được biết trong ngành (Ausubel et al. *Current Protocols in Molecular Biology*, Greene Publishing and Wiley Interscience, New York, 1989).

Hơn nữa, được hiểu rằng các phân tử polynucleotit phân lập và các phân tử ARN phân lập theo sáng chế bao gồm cả phân tử tổng hợp và phân tử thu được thông qua các kỹ thuật tái tổ hợp, như bằng cách tách dòng đơn tính *in vitro* và dịch mã.

Các phân tử polynucleotit có thể được gây đột biến di truyền sử dụng các kỹ thuật tái tổ hợp được biết bởi các người có hiểu biết trung bình trong lĩnh vực, bao gồm gây đột biến trực tiếp vị trí, hoặc bằng cách gây đột biến ngẫu nhiên như bằng cách tiếp xúc với các tác nhân gây đột biến hóa học hoặc với phóng xạ, như được biết trong ngành." Các đột biến này có thể được thực hiện bằng các phương pháp chuẩn được biết nhiều trong ngành, ví dụ gây đột biến trực tiếp vị trí (xem ví dụ Sambrook et al.(1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory

Press, Cold Spring Harbor, N.Y.) bản sao lây nhiễm như được mô tả (ví dụ Meulenberg et al., Adv. Exp. Med. Biol., 1998, 440:199-206).

Do đó, đối tượng của sáng chế đề xuất tiếp phương pháp tạo ra virus PRRS Bắc Mỹ được cải biến di truyền, phương pháp này bao gồm gây đột biến trình tự ADN mã hóa phân tử ARN lây nhiễm mã hóa virus PRRS như được mô tả trên đây, và biểu hiện virus PRRS cải biến di truyền này sử dụng hệ thống biểu hiện thích hợp. Virus PRRS cải biến di truyền có thể được biểu hiện từ phân tử polynucleotit phân lập sử dụng các hệ thống biểu hiện thích hợp thường được biết trong ngành, các ví dụ về các hệ thống biểu hiện này được mô tả trong đơn này. Ví dụ, phân tử polynucleotit phân lập có thể ở dạng plasmit có khả năng biểu hiện virus mã hóa trong tế bào chủ thích hợp *in vitro*, như được mô tả chi tiết hơn dưới đây.

Các trình tự protein N virus PRRSV Bắc Mỹ được bảo toàn cao và các trình tự thông báo có khoảng từ 93-100% giống nhau. Các protein N PRRSV Bắc Mỹ và Châu Âu có khoảng từ 57-59% giống nhau và cùng có chung các motif cấu trúc chung. Thông thường, khi so sánh các trình tự mã hóa PRRS và các mẫu phân lập, mà có thể được đánh số khác nhau như với các nucleotit hoặc axit amin mã hóa cụ thể, việc xác định các vùng thích hợp đạt được dễ dàng bằng cách xác định các axit amin đặc trưng bảo toàn trong chủng PRRS quan tâm và đuổi thẳng so sánh nó với chủng tham chiếu.

Công nghệ ADN tái tổ hợp bao gồm các kỹ thuật sinh học phân tử mạnh và cực kỳ biến đổi nhằm cải biến các axit nucleic ở mức ADN và làm cho nó có thể phân tích và cải biến hệ gen ở mức phân tử. Theo khía cạnh này, các virus như virus PRRS do kích thước vừa phải của hệ gen của nó là đặc biệt có thể sửa đổi cho các thao tác này. Tuy nhiên, công nghệ ADN tái tổ hợp không thể áp dụng ngay đối với các virus ARN không phải retrovirus bởi vì các virus này không chứa bước trung gian ADN trong khi chúng sao chép. Đối với các virus này, các dòng đơn tính cADN lây nhiễm phải được phát triển trước khi công nghệ ADN tái tổ hợp có thể được áp dụng cho hệ gen của chúng để tạo ra virus cải biến. Các dòng đơn tính lây nhiễm có thể nhận được thông qua xây dựng cADN có chiều dài hoàn chỉnh (chiều dài hệ gen) (được sử dụng ở đây với nghĩa rộng là bản sao ADN của ARN và không chỉ ở nghĩa hẹp là bản sao ADN của mARN) của virus này khi nghiên cứu, sau đó bản sao lây nhiễm được tổng hợp *in vivo* trong các tế bào lây nhiễm cADN có chiều dài hoàn chỉnh, nhưng các bản sao lây nhiễm có thể cũng được thu được bằng cách phiên mã *in vitro* từ cADN có chiều dài hoàn chỉnh trong plasmit có

gen khởi đầu của tế bào chưa có nhân điển hình khi có mặt hỗn hợp phiên mã, hoặc hơn nữa *in vitro* sử dụng các mảnh cADN có chiều dài từng phần được nối chứa hệ gen virus hoàn chỉnh. Trong tất cả các trường hợp, ARN phiên mã mang tất cả các cải biến mà đã được đưa vào cADN và có thể được sử dụng để cấy truyền tiếp virus cải biến như vậy.

Điều chế dòng đơn tính lây nhiễm mẫu phân lập virus PRRS Châu Âu hoặc virus Lelystad được mô tả trong bằng sáng chế Mỹ số 6,268,199 đưa toàn bộ nội dung của nó vào đây để viện dẫn. Điều chế dòng đơn tính cADN lây nhiễm của mẫu phân lập virus PRRS Bắc Mỹ được chỉ định là P129 (Lee *et al.*, 2005; Yoo *et al.*, 2004) được mô tả trong bằng sáng chế Mỹ số 6,500,662 đưa toàn bộ nội dung của nó vào đây để viện dẫn. Trình tự cADN của P129 được bộc lộ trong Sổ lưu giữ của Genbank AF494042 và trong bằng sáng chế Mỹ số 6,500,662. Sáng chế dưới đây tạo ra việc sử dụng dòng đơn tính lây nhiễm này mà được biểu hiện trong phạm vi plasmit nhờ gen khởi đầu sớm tức thì CMV và đã được chỉ định là pCMV-S-P129 và cũng được bộc lộ trong bằng sáng chế Mỹ số 6,500,662. Như được bộc lộ trong bằng sáng chế Mỹ số 6,500,662 có các plasmit và gen khởi đầu khác thích hợp để sử dụng ở đây.

Dựa trên trình tự hoàn chỉnh của khung đọc mở quan tâm bất kỳ và vị trí của gốc axit amin quan tâm, người có hiểu biết trung bình trong lĩnh vực chỉ cần tham khảo bảng mã di truyền để thiết kế các sự thay đổi ở vị trí cụ thể mong muốn.

Các mã di truyền cấu thành trình tự nucleotit bộ ba của trong mRNA và các phân tử cADN tương ứng của chúng. Các mã di truyền được đặc trưng bởi bazơ uraxil (U) khi có mặt trong phân tử mRNA nhưng được đặc trưng bởi bazơ thymidin (T) khi có mặt trong ADN. Sự thay đổi đơn giản trong mã di truyền bằng gốc axit amin tương tự trong polynucleotit sẽ không thay đổi trình tự hoặc cấu trúc của polypeptit mã hóa. Sẽ hiển nhiên rằng khi cụm từ bắt đầu bằng trình tự 3 nucleotit cụ thể "mã hóa" axit amin cụ thể bất kỳ, người có hiểu biết trung bình trong lĩnh vực sẽ nhận thấy rằng bảng trên đây cung cấp nghĩa nhận biết các nucleotit cụ thể đang tranh cãi. Bằng cách ví dụ, nếu ba trình tự ba nucleotit cụ thể mã hóa lysin, bảng trên bộc lộ rằng hai trình tự ba có thể là AAA và AAG. Glyxin được mã hóa bởi GGA, GGC, GGT (GGU nếu trong ARN) và GGG. Để thay đổi gốc lysin thành glyxin trong protein mã hóa, có thể thay thế bộ ba AAA hoặc AAG bằng bộ ba bất kỳ trong số GGA và GGC, GGT hoặc GGG trong axit nucleic mã hóa.

Như được đề cập trước đây, sáng chế đề cập đến việc đề xuất các chủng vacxin PRRS trong đó các đáp ứng của vật chủ đối với virus này qua trung gian con đường interferon, các đáp ứng khác không được điều hòa giảm. Như được mô tả chi tiết dưới đây, có nhiều sự cải biến khác nhau đối với hệ gen virus mà là hữu hiệu về mặt này, cụ thể là các cải biến ở trong ORF1a như được bộc lộ ở đây và các sự kết hợp các cải biến này. Nên được lưu ý rằng các điểm cải biến tương tự có thể được tìm thấy trong các khung đọc mở bổ sung của hệ gen PRRS, cũng như được bộc lộ ở đây (xem bảng 9).

Đáng chú ý là các phương pháp trước đó khác nhất định để cải biến polynucleotit PRRS đã thành công để làm yếu virus PRRS, có thể phù hợp cung cấp để sử dụng trong vacxin, mặc dù lý do chính xác để làm yếu virus này thường không được biết. Ví dụ, đã được bộc lộ việc làm yếu virus PRRS độc bằng cách gây đột biến hoặc xóa bỏ vùng NLS-2, vùng NoLS, hoặc vùng NES trong nucleocapsit hoặc protein N (được mã hóa bởi ORF7) của virus này, để đưa vào sự xóa bỏ trong khung đọc mở 7 (ORF7). Theo khía cạnh khác, sự xóa bỏ ORF7 là trong trình tự mã hóa tín hiệu định vị nhân (nuclear localization signal-NLS) của protein capsit. Sự xóa bỏ ORF7 trong trình tự mã hóa NLS có thể bao gồm sự xóa bỏ một hoặc nhiều axit amin ở vị trí từ 43-48 hoặc sự xóa bỏ axit amin ở một trong hai hoặc cả hai vị trí 43 và 44. Xem, ví dụ, toàn bộ nội dung của bằng sáng chế Mỹ số 7,544,362 mà được đưa vào đây bằng cách viện dẫn. Protein nucleocapsit (N) của PRRSV, được mã hóa bởi ORF7, là protein cơ bản nhỏ được phosphoryl hóa (Wootton, Rowland, và Yoo, 2002) và tạo thành các dị dimer (Wootton và Yoo, 2003). Cấu trúc tinh thể gần đây đã được xác định (Doan và Dokland, 2003). Protein N rõ ràng có nhiều chức năng trong tế bào lây nhiễm. Ngoài việc tạo thành cấu trúc capsit hình cầu mà đóng gói trong đó ARN hệ gen, quy trình này xảy ra trong tế bào chất, một phần protein N được vận chuyển vào trong nhân và đặc biệt là vào trong nhân của tế bào bị lây nhiễm. Trình tự axit amin của protein N chứa hai tín hiệu định vị nhân (NLS), tín hiệu định vị hạch nhân (nucleolar localization signal-NoLS), và tín hiệu đưa ra ngoài nhân (nuclear export signal-NES) mà tương ứng dễ dàng vận chuyển vào trong nhân và hạch nhân, và đưa ra khỏi nhân (Rowland et al., 1999; Rowland et al., 2003; Rowland and Yoo, 2003). Trong khi trong hạch nhân, protein N tương tác với protein liên kết với ARN hạch nhân nhỏ fibrillarin và có thể điều hòa việc xử lý rARN và sự phát sinh sinh vật ribosom trong tế bào lây nhiễm này để thích hợp cho sự sao chép virus (Yoo et al., 2003). Các đột biến virus kiểu này là có giá trị, riêng rẽ hoặc kết hợp với các

đột biến làm yếu khác, để tạo ra các vaccin PRRS mới. Trong ví dụ khác về virus PRRS mà đã được làm yếu, sử dụng việc cải biến ORF1a. Sử dụng việc xóa bỏ trình tự ADN mã hóa epitop kháng nguyên ở giữa axit amin 616 đến 752 trong vùng siêu biến trong vùng mã hóa 2 protein không phải cấu trúc của ORF1a, xem bằng sáng chế Mỹ số 7,618,797, mà được đưa vào đây bằng cách viện dẫn toàn bộ nội dung của nó.

Các nghiên cứu về sinh miễn dịch học của virus PRRS gợi ý rằng sự tương tác của virus PRRS với các PDC nên được xem xét. Loại tế bào này chiếm khoảng từ 0,2%-0,8% tế bào đơn nhân máu ngoại vi ở người, chuột nhắt, chuột, lợn và khỉ. Mặc dù khan hiếm, loại tế bào này là thành phần quan trọng của hệ miễn dịch bẩm sinh và có khả năng tiết khối lượng dư dả IFN- α sau khi kích thích virus. Mặc dù việc tiết IFN- α mà PDC đóng vai trò lớn trong việc điều hòa tính miễn dịch thích ứng và bẩm sinh đối với virus vì chúng thúc đẩy chức năng của tế bào giết tự nhiên, tế bào B, và tế bào T. Hơn nữa, sự thành thực chức năng các tế bào sợi nhánh có nguồn gốc từ bạch cầu đơn nhân lớn to của lợn (monocyte derived dendritic cell-MoDC) được trợ giúp bởi IFN- α được tiết bởi PDC dẫn đến khả năng của MoDC gia tăng để trình diện kháng nguyên và hoạt hóa các tế bào T. Ở trạng thái lây nhiễm virus sau đó, sự biệt hóa PDC thành kiểu tế bào sợi nhánh thành thực chức năng duy nhất, mà trực tiếp điều hòa chức năng của tế bào T và điều khiển sự biệt hóa tế bào T thành các tế bào có khả năng tiết IFN- γ , mà là chất trung gian chính của miễn dịch chống virus kháng lại virus bao gồm virus PRRS. Không ngạc nhiên, có các virus người, như virus hợp bào hô hấp và virus sởi, mà được biết là kiểm soát khả năng của PDC để tiết IFN- α . Hiệu quả ức chế này được cho là đóng vai trò trong ưu thế của đáp ứng miễn dịch thể dịch và bệnh học miễn dịch liên quan được quan sát thấy là do sự nhiễm các virus này, cũng như tính miễn cảm gia tăng của vật chủ đối với việc nhiễm virus và vi khuẩn lần thứ hai.

Ngược lại, các mẫu phân lập PRRSV kiểu đại cũng như cả hai chủng vaccin PRRS Ingelvac (xem ví dụ 5 và sau đây, dưới đây) ức chế khả năng của chủng loại PDC lợn thuần khiết để sản xuất IFN- α , trong khi kho virus P129-PK-FL và P129-PK-d43/44 mới (xem dưới đây) thể hiện giảm thiểu tác dụng không ức chế đối với chức năng PDC này. Ý nghĩa của các quan sát này tập trung một phần vào tính quan trọng của IFN- α trong việc điều hòa sự phát triển của đáp ứng miễn dịch thích ứng đối với virus. Do đó, rất có khả năng là vaccin virus làm yếu có nguồn gốc từ virus kiểm soát IFN- α tối thiểu sẽ gây ra đáp ứng miễn dịch bảo vệ chống virus mạnh. Trước đây, đã chứng minh được

rằng hiệu quả hỗ trợ của IFN- α đối với vacxin PRRS MLV Ingelvac cảm ứng đáp ứng IFN- γ qua trung gian tế bào T đặc hiệu virus và cường độ của đáp ứng IFN- γ qua trung gian tế bào T đặc hiệu virus gây ra bởi vacxin này có tương quan rõ ràng với miễn dịch bảo vệ kháng lại virus này dưới các điều kiện thực địa và phòng thí nghiệm. Do đó, mặc dù đã không bị giới hạn đến lý thuyết, sẽ có lý để kỳ vọng rằng đáp ứng miễn dịch qua trung gian tế bào và mức bảo vệ miễn dịch gây ra bởi PRRSV ức chế không phải IFN- α sẽ lớn hơn đáng kể so với đáp ứng của mẫu phân lập PRRSV có phenotyp (ức chế IFN- α) kiểu đại.

Đề cập đến sáng chế, đáng chú ý là virus P129-PK-FL cũng như tất cả năm thể đột biến xóa bỏ có nguồn gốc từ dòng đơn tính cADN lây nhiễm pCMV-S-P129-PK bị mất khả năng ức chế sự sản sinh IFN- α . Do đó, phenotyp bất thường này không thể chỉ là do sự mất đoạn, mà còn phải là do ít nhất một phần là các sự thay đổi di truyền mà trở thành cố định trong khi xây dựng dòng đơn tính lây nhiễm này. Thú vị là, virus P129 không được tách dòng mà được cấy truyền liên tiếp 63 lần trên các tế bào PK-9 vẫn giữ khả năng ức chế cảm ứng IFN (bảng 1). Giải thích có khả năng nhất về phenotyp IFN thông thường được cho thấy ở tất cả các virus có nguồn gốc từ dòng đơn tính lây nhiễm là sự đưa vào một hoặc nhiều đột biến trong khi tạo ra dòng đơn tính lây nhiễm này. Các đột biến này có thể đã tồn tại, có thể ở các mức thấp, trong ARN virus được sử dụng để xây dựng dòng đơn tính lây nhiễm này. Cuối cùng, các đột biến này có thể đã tồn tại ở virus nguyên thể (lần cấy truyền 0) trong lợn hoặc chúng có thể đã được tạo ra và làm giàu trong quá trình thích ứng virus này để sinh trưởng trên tế bào PK-9 cho 16 lần cấy truyền. Khả năng là các đột biến này là các lỗi gây ra do PCR thu được hoặc tách dòng đơn tính các dòng nhân tạo không thể được loại bỏ. Ở mức độ bất kỳ, (các) đột biến có tác dụng làm mất chức năng ức chế IFN- α trở nên “cố định” trong khi xây dựng dòng đơn tính lây nhiễm, và sẽ được kỳ vọng là có mặt trong tất cả các virus có nguồn gốc từ dòng đơn tính lây nhiễm cụ thể này.

Khả năng là các đột biến tạo ra phenotyp ức chế IFN- α thay đổi có trước đó trong ARN virus được sử dụng để xây dựng dòng đơn tính cADN là có khả năng, cho thấy rằng PRRSV được biết để dễ dàng tạo ra sự đa dạng di truyền ngẫu nhiên do các lỗi bằng ARN polymeraza phụ thuộc ARN virus. Loài dạng virus chứa hỗn hợp khác loại gồm các biến thể có di truyền gần gũi mà xuất hiện tự nhiên trong khi sao chép virus *in vivo*. Thích hợp hơn là quan sát loài dạng virus sau nhiều lần cấy truyền *in vitro* PRRSV

có nguồn gốc từ dòng đơn tính cADN lây nhiễm. Điều này là đáng chú ý vì chủng loại hệ gen virus ban đầu trong các nghiên cứu được tiến hành trước đó bao gồm chứa chủng loại đồng nhất di truyền, và sự đa dạng trình tự được tạo ra nhanh chóng trong khi cấy truyền khi nuôi cấy tế bào. Trong nghiên cứu gần đây, mức độ không đồng nhất hệ gen sẽ cao hơn vì virus P129 ban đầu đã không được tách dòng (sinh học hoặc phân tử) trước khi lần cấy truyền 16 PK-9 hướng đến xây dựng dòng đơn tính lây nhiễm. Do đó, sự chọn lọc có thể biến thể ARN PRRSV tạo ra mất chức năng ức chế IFN- α từ trong số loài dạng này, và đưa vào trong dòng đơn tính cADN lây nhiễm pCMV-S-P129-PK17-FL dường như hợp lý. Việc đưa vào các đột biến này vào trong dòng đơn tính lây nhiễm có thể được xem là ngẫu nhiên, dưới một số hoàn cảnh, cho thấy rằng tất cả các virus dẫn xuất nên có chung phenotyp sinh học khác biệt này mà có thể chứng tỏ là quan trọng đối với sự phát triển của vacxin PRRS thể hệ tiếp theo hữu hiệu.

Virus PRRS kiểu đại chủng P129, giống như chủng khác của virus này, thể hiện hiệu quả ức chế mạnh đối với khả năng của các tế bào đơn nhân máu ngoại vi (peripheral blood mononuclear cell-PBMC) và các tế bào sợi tương bào (plasmacytoid dendritic cell-PDC) để tạo ra interferon (IFN)- α . Mặt khác, virus có nguồn gốc từ dòng đơn tính cADN lây nhiễm của P129 (pCMV-S-P129-PK17-FL) thể hiện sự giảm đáng kể phenotyp ức chế IFN- α . Dòng đơn tính lây nhiễm này được xây dựng từ virus mà được điều chỉnh trước đó để sinh trưởng trên dòng tế bào thận lợn biểu hiện CD163 PK-9 qua tiến trình 16 lần cấy truyền liên tiếp (xem bằng sáng chế Mỹ số 7,754,464 mà được đưa vào đây bằng cách viện dẫn toàn bộ nội dung của nó). Phenotyp ức chế IFN- α là P129-PK-FL và P129-PK-d43/44 được sắp xếp từ thấp đến không đáng kể và là đối lập đáng kể với phenotyp thể hiện một trong hai chủng vacxin virus sống cải biến PRRS Ingelvac, cả hai phenotyp này là ức chế cao. Các kết quả này cho thấy rằng virus P129-PK-FL và P129-PK-d43/44 là khác biệt về mặt sinh học với mẫu phân lập cấy truyền ít P129 ban đầu, các virus PRRS kiểu đại khác, và cả hai vacxin PRRS Ingelvac. Sự liên quan tiềm năng của phenotyp ức chế IFN- α thấp, cũng như các lý do thay đổi phenotyp có thể, được thảo luận.

Sự cải biến axit amin theo sáng chế

Theo sự thực hiện sáng chế, các mẫu phân lập PRRS mới, dù là kiểu gen Bắc Mỹ hoặc Châu Âu, có thể được xác định phạm vi chứa đựng cụ thể các axit amin ở các

vị trí cụ thể trong các protein được mã hóa từ ORF1, và mà tạo ra phenotyp mong muốn trên các virus này. Theo phương án khác, như được đề cập trên đây, quy trình di truyền chuẩn có thể được sử dụng để cải biến trình tự di truyền (và do đó là trình tự axit amin) của protein ORF1 được mã hóa, lại để tạo ra virus PRRS Bắc Mỹ và Trung Quốc cải biến, và các dòng lây nhiễm, và các vaccine từ đó, tất cả chúng cung cấp các phenotyp này. Trong các ví dụ ưu tiên, các phenotyp bao gồm, nhưng không bị giới hạn đến, hiệu quả ức chế interferon- α giảm khi so sánh với virus PRRS kiểu dại, và, tùy ý, khả năng sản sinh hoặc tiếp tục tồn tại trong động vật chủ (lợn) trong khi kích thích đáp ứng miễn dịch mạnh, nhưng có bệnh học có thể phát hiện rất ít.

Do đó, để thực hiện sáng chế, chủng PRRS Bắc Mỹ hoặc các mẫu phân tách mà có thể thực hiện làm các điểm khởi đầu hữu ích bao gồm các chủng được bộc lộ, ví dụ trong bằng sáng chế Mỹ số 6,500,662; 7,618,797; 7,691,389, 7,132,106; 6,773, 908; 7,264,957; 5,695,766; 5,476,778; 5,846,805; 6,042,830; 6,982,160; 6,241,990; và 6,110,468. Về chủng PRRS Trung Quốc và các mẫu phân lập mà có thể sử dụng làm các điểm khởi đầu hữu ích, xem ví dụ, công bố đơn sáng chế Trung Quốc số CN200910091233.6 từ đơn Trung Quốc số CN201633909 gắn liền với virus TJM-92.

Liên quan đến thảo luận sau, các chỉ định ba chữ cái và một chữ cái được công nhận quốc tế được sử dụng cho các axit amin thông thường nhất được mã hóa bởi ADN: alanin (Ala, A); arginin (Arg, R); asparagin (Asn, N); axit aspartic (Asp, D); cystein (Cys, C); axit glutamic (Glu, E); glutamin (Gln, Q); glyxin (Gly, G); histidin (His, H); isoleuxin (Ile, I); loxin (Leu, L); lysin (Lys, K); methionin (Met, M); phenylalanin (Phe, F); prolin (Pro, P); serin (ser, S); threonin (Thr, T); tryptophan (Trp, W); tyrosin (Tyr, Y) và Valin (Val, V).

Bảng 9 và 10 xác định các sự thay đổi axit amin quan sát tạo ra việc làm yếu độ độc của PRRS Bắc Mỹ (và Trung Quốc) mà tương quan với sự ức chế giảm hoạt tính của interferon alpha, do đó cho phép đáp ứng miễn dịch mạnh và an toàn đối với vaccine. Bảng 10 xác định các sự cải biến axit amin ưu tiên cao về mặt này, trong ORF1a, và thể hiện cách phát sinh các đột biến này đối với các lần nuôi cấy sau khi cấy truyền P129 khác (nên được lưu ý rằng kiểm tra tiến trình này cũng dễ dàng thiết kế các chiến lược phát sinh đột biến để xây dựng (lại) ADN mã hóa có các sự thay đổi axit amin bất kỳ theo sáng chế, khi cần). Về mặt này, sự cải tiến axit amin ưu tiên nhất đối với ORF1a (như được chứng tỏ bằng P129 cấy truyền lần 52 bao gồm: asparagin ở vị trí 182,

asparagin ở vị trí 189, tyrosin ở vị trí 273, histidin ở vị trí 302, threonin ở vị trí 665, xystein ở vị trí 943, threonin ở vị trí 1429, alanin ở vị trí 1505, asparagin ở vị trí 2410, mà cũng bổ sung tiềm tàng một số cơ hội glycosyl hóa và có thể thay đổi thêm nữa chức năng protein.

Do đó, sáng chế đề xuất an virus gây hội chứng rối loạn sinh sản và hô hấp ở lợn (PRRS) phân tách trong đó protein của nó được mã hóa bởi ORF1a được chọn từ nhóm gồm các trình tự axit amin này chứa bất kỳ trong số:

axit amin N trong trình tự axit amin AMANNVYD (SEQ ID NO: 9);
 axit amin N trong trình tự axit amin IGHNANVM (SEQ ID NO: 12);
 axit amin D trong trình tự axit amin TVPDGNC (SEQ ID NO: 15);
 axit amin Y trong trình tự axit amin CWWYLFD (SEQ ID NO: 18);
 axit amin H trong trình tự axit amin HGVHGKY (SEQ ID NO: 21);
 axit amin V trong trình tự axit amin AAKVDQY (SEQ ID NO: 24);
 axit amin T trong trình tự axit amin PSATTDTS (SEQ ID NO: 27);
 axit amin L trong trình tự axit amin LNSLLSK (SEQ ID NO: 30).
 axit amin C trong trình tự axit amin APMCCQDE (SEQ ID NO: 33);
 axit amin T trong trình tự axit amin CAPTTGMD (SEQ ID NO: 36);
 axit amin A trong trình tự axit amin PKVAKVS (SEQ ID NO: 39);
 axit amin I trong trình tự axit amin AGEIIVG (SEQ ID NO: 42);
 axit amin N trong trình tự axit amin ADFNNPEK (SEQ ID NO: 45); và
 axit amin I trong trình tự axit amin QTPILIGR (SEQ ID NO: 48).

Cụ thể hơn, sáng chế đề xuất virus gây hội chứng rối loạn sinh sản và hô hấp ở lợn (PRRS) được phân lập trong đó protein của nó được mã hóa bởi ORF1a được chọn từ nhóm gồm các trình tự axit amin này chứa bất kỳ trong số:

axit amin N trong trình tự axit amin ANNV (xem trình tự nêu trong SEQ ID NO: 9);
 axit amin N trong trình tự axit amin HNA (xem trình tự nêu trong SEQ ID NO: 12);
 axit amin D trong trình tự axit amin PDG (xem trình tự nêu trong SEQ ID NO: 15);
 axit amin Y trong trình tự axit amin WYL (xem trình tự nêu trong SEQ ID NO: 18);

- axit amin H trong trình tự axit amin VHG (xem trình tự nêu trong SEQ ID NO: 21);
- axit amin V trong trình tự axit amin KVD (xem trình tự nêu trong SEQ ID NO: 24);
- axit amin T trong trình tự axit amin ATD (xem trình tự nêu trong SEQ ID NO: 27);
- axit amin L trong trình tự axit amin SLL (xem trình tự nêu trong SEQ ID NO: 30).
- axit amin C trong trình tự axit amin MCQ (xem trình tự nêu trong SEQ ID NO: 33);
- axit amin T trong trình tự axit amin PTG (xem trình tự nêu trong SEQ ID NO: 36);
- axit amin A trong trình tự axit amin VAK (xem trình tự nêu trong SEQ ID NO: 39);
- axit amin I trong trình tự axit amin EIV (xem trình tự nêu trong SEQ ID NO: 42);
- axit amin N trong trình tự axit amin FNP (xem trình tự nêu trong SEQ ID NO: 45); và
- amino acid I trong trình tự axit amin PIL (xem trình tự nêu trong SEQ ID NO: 48).

Như được đề cập ở trên, có nhiều chủng đã biết và các mẫu phân lập của PRRS Bắc Mỹ và Trung Quốc, và các chủng mới này liên tục tiến hóa hoặc được phân lập. Mặc dù mức độ tương đồng trình tự axit amin cao giữa tất cả các chủng này, người có hiểu biết trung bình trong lĩnh vực sẽ nhận biết ngay rằng một số biến dị có tồn tại, và quả thực hữu ích có thể được lấy từ các sự khác biệt và các sự tương đồng này để cải thiện hơn nữa các đặc điểm phenotyp của tất cả các chủng vacxin.

Trước hết, về tất cả các motif axit amin được xác định bởi trình tự nêu trong SEQ ID NO như được chỉ định trực tiếp (trên các trang 27-28) trên đây, các axit amin được gạch chân và ưu tiên (như được cung cấp từ P129 lần cấy truyền 52) thường vẫn hoàn toàn có lợi ngay cả khi axit amin liền đã được thay đổi khác so với các trình tự SEQ ID NO được chỉ định. Do đó, đối với AMANVYD (SEQ ID NO: 9), là ví dụ cụ thể và đại diện, thường có thể kiểm tra trình tự protein biểu hiện ORF1 tương ứng từ PRRS Bắc Mỹ hoặc Trung Quốc bất kỳ, để tìm ra motif axit amin tương ứng, ngay cả khi các sự

thay đổi bổ sung đã xảy ra ở các chủng khác này, do sự tiến hóa, do các sự thay thế và/hoặc xóa bỏ hoặc bổ sung, như được hiểu bởi người có kỹ năng trong lĩnh vực này, các sự thay đổi axit amin ưu tiên được chứng tỏ bởi P129 cây truyền 52 như vậy cũng nên vẫn hoạt động mặc dù các sự thay đổi khác trong toàn bộ trình tự axit amin mà là theo hướng 5' hoặc 3' đối với axit amin cụ thể của cây truyền 52. Điều này sẽ đặc biệt đến mức nếu các sự thay đổi axit amin so sánh này được xem là bảo toàn. Do đó đối với AMANVYD (SEQ ID NO: 9), và trình tự ANV của nó, nên có thể dễ dàng xác định motif so sánh trong chủng PRRS khác nếu, ví dụ, valin trên đó được thay bằng isoloxin hoặc loxin, hoặc gốc khác bất, hoặc nếu một gốc đơn giản bị mất hoặc bổ sung thêm một gốc. Một số chương trình máy tính hiện có để xác định các sự duỗi thẳng và do đó xác định nếu các motif trình tự polypeptit tương ứng, ví dụ các bảng được gọi là Blosum (dựa trên mức phần trăm tương đồng xác định), xem S. Henikoff et al. "Amino Acid Substitution matrices from protein blocks", Proc Natl Acad Sci, USA, 89(22), pp. 10915-10919, Nov 15, 1992., và see also A. L. Lehninger et al. Principles of Biochemistry, 2005, MacMillan và Company, 4th edition. Các sự thay đổi axit amin bảo toàn cũng được nhận biết dựa trên sự phân loại thành tổng số 5 nhóm: sulfydryl (Cys); thơm (Phe, Tyr, và Trp); bazơ (Lys, Arg, His); béo (Val, Ileu, Leu, Met), và ưa nước (Ala, Pro, Gly, Glu, Asp, Gln, Asn, Ser và Thr). Do đó trong phạm vi thực hiện sáng chế để cải biến trình tự nucleotit mã hóa PRRS Bắc Mỹ hoặc Trung Quốc bất kỳ để đưa vào ở vị trí thích hợp và tương ứng, các sự thay đổi axit amin bất kỳ được chỉ định cho P129 lần cây truyền 52, ngay cả khi một hoặc nhiều axit amin khác liền kề với vị trí chỉ định đã được bổ sung, xóa bỏ hoặc thay thế.

Ngoài ra, dựa trên nguyên tắc đơn giản, người có hiểu biết trung bình trong lĩnh vực sẽ nhận biết rằng ngay khi axit amin ưu tiên được xác định từ lần cây truyền 52 cụ thể các sự thay đổi được xác định cho ORF1a theo sự thực hiện sáng chế, rằng các sự thay thế bảo toàn cho các axit amin lần cây truyền 52 này bất kỳ có thể sau đó cũng được sử dụng, trong các biến thể P129, hoặc đối với các chủng Bắc Mỹ hoặc Trung Quốc khác bất kỳ, bằng sự bảo toàn quan trọng của phenotyp cây truyền 52 chủ định. Do đó, là các ví dụ điển hình: đối với SLL (trong trình tự nêu trong SEQ ID NO: 30), gốc loxin chỉ định có thể được thay thế tiếp bằng isoloxin, valin hoặc methionin; đối với FNP (trong trình tự nêu trong SEQ ID NO: 45), asparagin chỉ định có thể được thay bằng gốc bất kỳ trong số Ala, Pro, Gly, Glu, Asp, Gln, Ser và Thr; và đối với VAK (xem

trình tự nêu trong SEQ ID NO: 39), alanin chỉ định có thể được thay bằng bất kỳ trong số Asn, Pro, Gly, Glu, Asp, Gln, Ser và Thr; tất cả và tương tự, mặc dù sẽ được nhận biết dễ dàng rằng không yêu cầu trong sáng chế này là axit amin thay thế bất kỳ này hoạt động tốt như các sự thay đổi axit amin lần cây truyền 52 duy nhất được xác định ban đầu, ở các vị trí cụ thể. Tất nhiên, sử dụng các sự thay đổi axit amin bảo toàn chuẩn theo mô hình nhận biết khác bất kỳ cũng được thực hiện trong sáng chế này. Ví dụ, và bao gồm ngược lại trong tất cả các trường hợp, Asp cho Glu và ngược lại, Asn cho Gln, Arg cho Lys, Ser cho Cys hoặc Thr, Phe cho Tyr, Val cho Leu hoặc Ileu, Ala cho Gly, và sự thay thế tương tự.

Hơn nữa, trong khi thực hiện sáng chế này, mặc dù các sự thay đổi axit amin ở lần cây truyền 52 riêng biệt (như được xác định cho ORF1a trên đây) có thể được thay hữu ích trong PRRS Bắc Mỹ và Trung Quốc bất kỳ bằng hiệu ứng phenotyp mong muốn, được ưu tiên nữa là đưa nhiều nhất có thể các sự chọn lọc axit amin trong bảng 9 hoặc bảng 10 vào trong cấu trúc cuối cùng, như được đề xuất điển hình bằng sự cải biến thích hợp trình tự poly nucleotit mã hóa. Do đó, thực hiện sáng chế bao gồm cung cấp virus PRRS Bắc Mỹ hoặc Trung Quốc (và các polynucleotit mã hóa tương ứng) cung cấp 2, 3, 4, 5, 6, 7, 8, 9, và lên đến bất kỳ trong số khoảng 17 sự thay đổi ORF1a trong lần cây truyền 52 xác định, (Bảng 9) tất cả trong trình tự virus cuối cùng, để đưa vào các cặp, các bộ ba hoặc các sự kết hợp cao hơn khác cụ thể bất kỳ tất cả tổng các sự thay đổi axit amin ở lần cây truyền 52 xác định. Tất nhiên, các sự thay đổi axit amin này có thể được đưa vào các trình tự nucleotit mã hóa tương ứng của virus này bằng sự phát sinh đột biến trực tiếp vị trí, PCR, và các kỹ thuật khác được biết nhiều trong ngành.

Để chứng minh rằng chúng cải biến di truyền cụ thể được làm yếu, có thể sử dụng thí nghiệm được mô tả như sau.

Ít nhất 10 lợn cái con cho mỗi nhóm được đưa vào trongmooix thử nghiệm, có nguồn gốc từ trang trại không bị nhiễm PRRSV. Các con vật được xét nghiệm không có kháng thể trong huyết thanh đặc hiệu virus PRRS và âm tính với PRRSV. Tất cả các con vật được đưa vào trong thử nghiệm này là cùng nguồn và được nuôi. Sự chỉ định các con vật vào các nhóm là ngẫu nhiên.

Kích thích được thực hiện vào ngày 90 của thai kỳ bằng cách dùng trong mũi 1 ml PRRSV chứa 10^5 TCID₅₀ cho mỗi lỗ mũi. Có ít nhất ba nhóm cho mỗi sơ đồ thử

nghiệm: Một nhóm kích thích bằng P129; một nhóm thử nghiệm kích thích bằng virus làm yếu; và một nhóm đối chứng nghiêm ngặt.

Nghiên cứu này dường như có giá trị khi các đối chứng nghiêm ngặt này vẫn âm tính PRRSV trong khoảng thời gian thực hiện thí nghiệm này và ít nhất ít hơn 25% số lợn con khỏe mạnh sống được sinh ra ở nhóm kích thích bằng live P129 so với các con đối chứng nghiêm ngặt.

Làm yếu, nói cách khác là làm giảm độ độc, được định nghĩa là sự thay đổi có ý nghĩa thống kê một hoặc nhiều thông số xác định hoạt động sinh sản hoặc dấu hiệu khác:

Giảm có ý nghĩa ít nhất một trong số các thông số sau đối với nhóm thử nghiệm (virus làm yếu) so với nhóm lây nhiễm chủng ban đầu chưa cải biến sẽ là sự chỉ định làm yếu:

- a) tần suất tử sản
- b) sảy thai ở hoặc trước ngày 112 của thai kỳ
- c) số lượng lợn con khô đẻ
- d) số lợn con yếu và sống sót thấp
- e) tỷ lệ chết trước khi cai sữa

Hơn nữa, sự gia tăng đáng kể một trong số các thông số sau đối với nhóm thử nghiệm so với nhóm lây nhiễm chủng ban đầu chưa cải biến được ưu tiên:

- f) số lợn con đã cai sữa cho mỗi con lợn cái
- g) số lợn con sống khỏe mạnh được sinh ra cho mỗi con lợn cái

Theo phương án khác, các triệu chứng hô hấp và các triệu chứng khác của việc lây nhiễm PRRSV có thể được xem xét để thiết lập sự làm yếu.

Chủng làm yếu là có giá trị cho các chế phẩm vaccin. Vaccin theo sáng chế là hữu hiệu nếu nó ngăn ngừa lợn khỏi bị nhiễm virus PRRS. Vaccin bảo vệ lợn khỏi bị nhiễm virus PRRS nếu sau khi dùng vaccin này cho một hoặc nhiều lợn chưa bị bệnh, sự kích thích sau đó bằng mẫu phân lập virus tinh khiết về mặt sinh học (ví dụ, VR 2385, VR 2386, P129 v.v.) dẫn đến độ trầm trọng của các thay đổi về mô bệnh học hoặc thay đổi tổng thể bất kỳ giảm bớt (ví dụ, tổn thương trong phổi) và/hoặc các triệu chứng của bệnh này, khi so sánh với các sự thay đổi hoặc các triệu chứng này thường gây ra bởi mẫu phân lập ở trong lợn tương tự mà không được bảo vệ (nghĩa là so với đối chứng thích hợp). Cụ thể hơn, vaccin theo sáng chế có thể cho thấy là hữu hiệu bằng cách dùng

vacxin nayf cho một hoặc nhiều con lợn thích hợp cần dùng, rồi thì sau khoảng thời gian thích hợp (ví dụ, 4 tuần), kích thích bằng mẫu phân lập PRRSV tinh khiết về mặt sinh học lớn ($10^{(3-7)}$ TCID₍₅₀₎). Sau đó lấy mẫu máu từ lợn đã được kích thích sau khoảng một tuần, và nỗ lực để phân lập virus từ mẫu máu này sau đó được thực hiện. Phân lập thấy khối lượng lớn virus này là chỉ định rằng vacxin này có thể không hữu hiệu, trong khi phân lập thấy khối lượng virus giảm (hoặc không có virus) là chỉ định rằng vacxin này là hữu hiệu.

Do đó, hiệu quả của vacxin theo sáng chế có thể được đánh giá định lượng (nghĩa là, tăng phần trăm mô phổi không bị bệnh khi so sánh với nhóm đối chứng thích hợp) hoặc định tính (ví dụ, phân lập PRRSV từ máu, phát hiện kháng nguyên PRRSV trong phổi, amidan hoặc mẫu mô hạch lymphô bằng thử nghiệm miễn dịch). Các triệu chứng của bệnh sinh sản và hô hấp ở lợn có thể được đánh giá định lượng (ví dụ, nhiệt độ/sốt) hoặc bán định lượng (ví dụ, có hoặc không có một hoặc nhiều triệu chứng hoặc giảm mức trầm trọng của một hoặc nhiều triệu chứng, như chứng xanh tím, viêm phổi, các tổn thương phổi v.v.).

Lợn không bị ảnh hưởng là lợn không tiếp xúc với tác nhân lây nhiễm bệnh sinh sản và hô hấp lợn, hoặc đã tiếp xúc với tác nhân lây nhiễm bệnh sinh sản và hô hấp lợn nhưng không thể hiện các triệu chứng của bệnh này. Lợn bị ảnh hưởng là lợn thể hiện các triệu chứng của PRRS hoặc có thể phân lập được từ lợn này PRRSV.

Vacxin theo sáng chế có thể được bào chế theo quy trình chấp nhận được sau để đưa vào các chất mang dùng trong thú y cho động vật, bao gồm người (nếu có thể áp dụng), như các đệm chuẩn, các chất ổn định, chất pha loãng, chất bảo quản, và/hoặc chất hòa tan, và có thể cũng được bào chế để dễ dàng giải phóng duy trì. Các chất pha loãng bao gồm nước, nước muối, dextroza, etanol, glyxerol, và chất tương tự. Các chất phụ trợ để tạo tính đẳng trương bao gồm natri clorua, dextroza, mannitol, sorbitol, và lactoza, các chất khác. Các chất ổn định bao gồm albumin, các chất khác. Các tá dược lỏng và các chất phụ trợ vacxin thích hợp khác bao gồm các chất mà là đặc biệt hữu ích để bào chế vacxin sống cải biến, được biết hoặc rõ ràng đối với người có hiểu biết trung bình trong lĩnh vực. Ví dụ xem, Remington's Pharmaceutical Science, 18th ed., 1990, Mack Publishing, được đưa vào đây bằng cách viện dẫn.

Vacxin theo sáng chế có thể chứa thêm một hoặc nhiều thành phần điều biến miễn dịch bổ sung như, ví dụ, tá dược hoặc xytokin, các thành phần khác. Các ví dụ

không giới hạn về các tá dược có thể được sử dụng trong vacxin trong sáng chế này bao gồm hệ thống tá dược RIBI (Ribi Inc., Hamilton, Mont.), phèn, các gel vô cơ như gel nhôm hydroxit, các nhũ tương dầu trong nước, các nhũ tương nước trong dầu như, ví dụ, tá dược không hoàn toàn và hoàn toàn Freund, copolyme khối (CytRx, Atlanta, Ga.), QS-21 (Cambridge Biotech Inc., Cambridge Mass.), SAF-M (Chiron, Emeryville Calif.), tá dược AMPHIGEN.RTM., saponin, Quil A hoặc phân saponin khác, monophosphoryl lipit A, và tá dược lipit-amin Avridin. Các ví dụ không giới hạn về nhũ tương dầu trong nước hữu ích trong vacxin theo sáng chế bao gồm chế phẩm SEAM62 và SEAM 1/2 cải biến. SEAM62 cải biến là nhũ tương dầu trong nước chứa 5% (thể tích) squalen (Sigma), 1% (thể tích) chất tẩy rửa SPAN.RTM. 85 (ICI Surfactants), 0,7% (thể tích) chất tẩy rửa TWEEN.RTM. 80 (ICI Surfactants), 2,5% (thể tích) etanol, 200 pg/ml Quil A, 100 [mgr]g/ml cholesterol, và 0,5% (thể tích) lexitin. SEAM ½ cải biến là nhũ tương dầu trong nước chứa 5% (thể tích) squalen, 1% (thể tích) chất tẩy rửa SPAN.RTM. 85, 0,7% (thể tích) chất tẩy rửa Tween 80, 2,5% (thể tích) etanol, 100 .mu.g/ml Quil A, và 50 .mu.g/ml cholesterol. Các tác nhân điều biến miễn dịch khác có thể được đưa vào trong vacxin bao gồm, ví dụ, một hoặc nhiều intolokin, interferon, hoặc xytokin đã biết khác.

Vacxin theo sáng chế có thể tùy ý được bào chế để giải phóng duy trì virus này, phân tử ARN lây nhiễm, plasmit, hoặc vector virus theo sáng chế. Các ví dụ về các chế phẩm giải phóng duy trì này bao gồm virus, phân tử ARN lây nhiễm, plasmit, hoặc vector virus kết hợp với các hợp phần polyme tương thích sinh học, như, ví dụ, poly(axit lactic), poly(axit lactic-co-glycolic), metylxenluloza, axit hyaluronic, collagen và chất tương tự. Cấu trúc, sự chọn lọc và việc sử dụng polyme phân hủy trong tá dược lỏng phân phối được chất đã được xem xét trong một vài công bố, bao gồm A. Domb et al., 1992, *Polymers for Advanced Technologies* 3: 279-292, được đưa vào đây bằng cách viện dẫn. Hướng dẫn bổ sung để chọn lọc và sử dụng polyme trong các dược phẩm có thể được tìm thấy trong các tài liệu đã biết trong ngành, ví dụ M. Chasin và R. Langer (eds), 1990, "Biodegradable Polymers as Drug Delivery Systems" trong: *Drugs and the Pharmaceutical Sciences*, Vol. 45, M. Dekker, N.Y., cũng được đưa vào đây bằng cách viện dẫn. Theo cách khác, hoặc ngoài ra, virus, plasmit, hoặc vector virus có thể được tạo vỏ vi nang để cải thiện việc dùng và hiệu quả. Các phương pháp tạo vỏ vi nang các kháng nguyên được biết nhiều trong ngành, và bao gồm các kỹ thuật được mô tả, ví dụ,

trong bằng sáng chế Mỹ số 3,137,631; bằng sáng chế Mỹ số 3,959,457; bằng sáng chế Mỹ số 4,205,060; bằng sáng chế Mỹ số 4,606,940; bằng sáng chế Mỹ số 4,744,933; bằng sáng chế Mỹ số 5,132,117; và công bố đơn quốc tế số WO 95/28227, tất cả các tài liệu này được đưa vào đây bằng cách viện dẫn.

Liposom có thể cũng được sử dụng để tạo ra sự giải phóng duy trì virus, plasmit, hoặc vector virus. Chi tiết liên quan đến cách tạo ra và sử dụng chế phẩm liposom có thể được tìm thấy trong, bằng sáng chế Mỹ số 4,016,100; bằng sáng chế Mỹ số 4,452,747; bằng sáng chế Mỹ số 4,921,706; bằng sáng chế Mỹ số 4,927,637; bằng sáng chế Mỹ số 4,944,948; bằng sáng chế Mỹ số 5,008,050; và bằng sáng chế Mỹ số 5,009,956, hoặc các tài liệu khác, tất cả tài liệu này được đưa vào đây bằng cách viện dẫn.

Lượng hữu hiệu vaccin bất kỳ được mô tả trên đây có thể được xác định bằng phương pháp thông thường, bắt đầu bằng liều dùng virus thấp, plasmit protein virus hoặc vector virus, và sau đó gia tăng phân liều trong khi theo dõi hiệu quả. Lượng hữu hiệu có thể thu được sau khi dùng liều đơn vaccin hoặc sau nhiều lần dùng vaccin. Các yếu tố đã biết có thể được xem xét khi xác định liều dùng tối ưu cho mỗi con vật. Các yếu tố này bao gồm loài, kích thước, tuổi và tình trạng chung của con vật này, việc sử dụng các dược chất khác trong con vật này, và yếu tố tương tự. Liều dùng thực tế ưu tiên được chọn sau khi xem xét kết quả từ các nghiên cứu động vật khác (xem, ví dụ, ví dụ 2 và 3 dưới đây).

Một phương pháp phát hiện liệu đáp ứng miễn dịch thích hợp đã thu được là để xác định sự chuyển đổi huyết thanh và chuẩn độ kháng thể trong con vật này sau khi dùng vaccin. Thời điểm dùng vaccin và số lần dùng tăng cường, nếu có, sẽ ưu tiên được xác định bởi bác sĩ hoặc bác sĩ thú y dựa trên phân tích tất cả yếu tố thích hợp, một số yếu tố này được mô tả trên đây.

Lượng liều hữu hiệu của virus, protein, phân tử ADN lây nhiễm, plasmit, hoặc vector virus, theo sáng chế có thể được xác định sử dụng các kỹ thuật đã biết, không kể đến các yếu tố có thể xác định bởi người có hiểu biết trung bình trong lĩnh vực như trọng lượng động vật được dùng vaccin. Khối lượng liều dùng virus theo sáng chế trong vaccin theo sáng chế ưu tiên là nằm trong khoảng từ 10^1 đến 10^9 pfu (plaque forming unit-đơn vị tạo thành vết tan), ưu tiên hơn là nằm trong khoảng từ 10^2 đến 10^8 pfu, và ưu tiên nhất là nằm trong khoảng từ 10^3 đến 10^7 pfu. Khối lượng liều dùng plasmit theo

sáng chế trong vacxin theo sáng chế ưu tiên là nằm trong khoảng từ 0.1mg đến 100mg, ưu tiên hơn là nằm trong khoảng từ 1mg đến 10mg, cũng ưu tiên hơn là nằm trong khoảng từ 10mg đến 1mg. Khối lượng liều dùng phân tử ADN lây nhiễm theo sáng chế trong vacxin theo sáng chế ưu tiên là nằm trong khoảng từ 0,1mg đến 100mg, ưu tiên hơn là nằm trong khoảng từ 1mg đến 10mg, cũng ưu tiên hơn là nằm trong khoảng từ 10mg đến 1mg. Khối lượng liều dùng vector virus theo sáng chế trong vacxin theo sáng chế ưu tiên là nằm trong khoảng từ 10^1 pfu đến 10^9 pfu, ưu tiên hơn là nằm trong khoảng từ 10^2 pfu đến 10^8 pfu, và cũng ưu tiên hơn là nằm trong khoảng từ 10^3 đến 10^7 pfu. Khối lượng liều dùng thích hợp nằm trong khoảng từ 0,5 ml đến 10 ml, và ưu tiên hơn là nằm trong khoảng từ 1 ml đến 5 ml.

Các liều dùng thích hợp cho vacxin protein hoặc peptit theo sáng chế thường nằm trong khoảng từ 1 đến 50 microgam cho mỗi liều dùng, hoặc khối lượng lớn hơn như có thể được xác định bởi các phương pháp chuẩn, với khối lượng tá được được xác định bởi các phương pháp nhận biết đối với mỗi chất này. Trong ví dụ ưu tiên theo sáng chế liên quan đến dùng vacxin lợn, đích tuổi tối ưu cho các con vật này là nằm trong khoảng từ 1 đến 21 ngày, trước khi cai sữa, có thể cũng tương ứng với các thời gian dùng vacxin khác như kháng lại *Mycoplasma hyopneumoniae* hoặc PCV. Ngoài ra, thời gian dùng vacxin ưu tiên đối với lợn cái nuôi sinh sản sẽ bao gồm liều dùng tương tự, với thời gian dùng lại vacxin hằng năm.

Lượng hữu hiệu vacxin được mô tả trên đây bất kỳ có thể được xác định bằng phương pháp thông thường, bắt đầu bằng liều dùng virus thấp, plasmid hoặc vector virus, và sau đó gia tăng phân liều trong khi theo dõi hiệu quả. Lượng hữu hiệu có thể thu được sau khi dùng liều đơn vacin hoặc sau nhiều lần dùng vacxin. Các yếu tố đã biết có thể được xem xét khi xác định liều dùng tối ưu cho mỗi con vật. Các yếu tố này bao gồm loài, kích thước, tuổi và tình trạng chung của con vật này, việc sử dụng các dược chất khác trong con vật này, và yếu tố tương tự. Liều dùng thực tế ưu tiên được chọn sau khi xem xét kết quả từ các nghiên cứu động vật khác.

Một phương pháp phát hiện liệu đáp ứng miễn dịch thích hợp đã thu được là để xác định sự chuyển đổi huyết thanh và chuẩn độ kháng thể trong con vật này sau khi dùng vacxin. Thời điểm dùng vacxin và số lần dùng tăng cường, nếu có, sẽ ưu tiên được xác định bởi bác sỹ hoặc bác sỹ thú y dựa trên việc phân tích tất cả yếu tố thích hợp, một số yếu tố này được mô tả trên đây.

Lượng liều hữu hiệu virus, phân tử ARN lây nhiễm, plasmit, hoặc vector virus, theo sáng chế có thể được xác định sử dụng các kỹ thuật đã biết, không kể đến các yếu tố có thể được xác định bởi người có hiểu biết trung bình trong lĩnh vực như trọng lượng của con vật được dùng vaccin. Bằng cách ví dụ, vaccin có thể được dùng qua đường miệng, ngoài đường tiêu hóa, trong chân bì, dưới da, trong cơ, trong mũi hoặc trong tĩnh mạch. Dùng qua đường miệng có thể bao hàm, ví dụ, cho chế phẩm này vào trong thức ăn hoặc nước uống của động vật. Các yếu tố ảnh hưởng đến liều dùng vaccin bao gồm, ví dụ, trọng lượng và tuổi của lợn.

Sáng chế còn đề xuất phương pháp bào chế vaccin chứa virus PRRS, phân tử ARN lây nhiễm, plasmit, hoặc vector virus được mô tả ở đây, phương pháp này bao gồm kết hợp lượng hữu hiệu một trong số virus PRRS, phân tử ARN lây nhiễm, plasmit, hoặc vector virus theo sáng chế, với chất mang được dùng hoặc sử dụng trong thú y.

Hơn nữa vaccin sống được làm yếu theo sáng chế có thể được cải biến như được mô tả trong bằng sáng chế Mỹ số 6,500,662 để mã hóa epitop kháng nguyên khác loại mà được xen vào trong hệ gen virus PRRS sử dụng các kỹ thuật tái tổ hợp đã biết. Cũng xem. bằng sáng chế Mỹ số 7,132,106 mà được đưa vào đây bằng cách viện dẫn toàn bộ nội dung của nó. Các epitop kháng nguyên hữu ích làm epitop kháng nguyên khác loại cho sáng chế bao gồm các epitop kháng nguyên từ mầm bệnh của lợn không phải là virus PRRS bao gồm, nhưng không bị giới hạn đến, epitop kháng nguyên từ mầm bệnh lợn được chọn từ nhóm bao gồm parvovirus lợn, xircovirus lợn, rotavirus lợn, virus cúm lợn, virus đại giả, virus viêm dạ dày-ruột lan truyền, coronaviruses hô hấp lợn, virus dịch tả lợn cổ điển, virus dịch tả lợn Châu Phi, virus viêm não-cơ tim, paramyxovirus lợn, virus torque teno, *Actinobacillus pleuropneumoniae*, *Actinobacillus suis*, *Bacillus anthracis*, *Bordetella bronchiseptica*, *Clostridium haemolyticum*, *Clostridium perfringens*, *Clostridium tetani*, *Escherichia coli*, *Erysipelothrix rhusiopathiae*, *Haemophilus parasuis*, *Leptospira* spp., *Mycoplasma hyopneumoniae*, *Mycoplasma hyorhinis*, *Mycoplasma hyosynoviae*, *Pasteurella multocida*, *Salmonella choleraesuis*, *Salmonella typhimurium*, *Streptococcus equisimilis*, và *Streptococcus suis*. Trình tự nucleotit mã hóa epitop kháng nguyên từ các mầm bệnh lợn được đề cập trên đây được biết trong ngành và có thể thu được từ cơ sở dữ liệu gen công bố trên trang web, như ở Genbank từ Trung tâm thông tin quốc gia về công nghệ sinh học (National Center for Biotechnology Information) (Mỹ).

Đặc điểm bổ sung và các biến đổi theo sáng chế sẽ rõ ràng đối với người có hiểu biết trung bình trong lĩnh vực từ toàn bộ bản mô tả này, bao gồm mô tả chi tiết, và tất cả các đặc điểm được nêu như các khía cạnh theo sáng chế. Tương tự, các đặc điểm theo sáng chế được mô tả ở đây có thể được kết hợp lại vào trong các phương án bổ sung mà cũng nhằm mô tả các khía cạnh theo sáng chế, không kể dù là sự kết hợp các đặc điểm được đề cập đặc biệt trên đây là một khía cạnh hoặc phương án theo sáng chế. Do đó, chỉ các giới hạn này được mô tả ở đây là quan trọng đối với sáng chế nên được xem xét là; các biến thể theo sáng chế thiếu các giới hạn mà đã không được mô tả ở đây là quan trọng nhằm để làm các khía cạnh theo sáng chế. Sẽ rõ ràng rằng sáng chế này có thể được thực hiện khác với như được mô tả cụ thể trong bản mô tả và các ví dụ ở đây.

Một số sự cải biến và các biến thể theo sáng chế có thể rõ ràng dưới sáng chế này và, do đó là thuộc phạm vi của sáng chế.

Các ví dụ sau đây nhằm để minh họa sáng chế không làm giới hạn sáng chế.

Ví dụ thực hiện sáng chế

Ví dụ 1

Làm thích ứng và làm yếu mẫu phân lập PRRSV P129 vào các tế bào PK-9

Mẫu phân lập virus PRRS P129 được phân lập từ lợn ốm ở Ấn Độ năm 1995 ở phòng thí nghiệm chẩn đoán bệnh động vật của trường đại học Purdue (Animal Disease Diagnostic Laboratory of Purdue University). Mẫu huyết thanh từ lợn này được cấy truyền một lần vào lợn có tình trạng sức khỏe tốt để mở rộng kho dịch treo mô đồng thể huyết thanh và phổi. ARN virus được chiết từ dịch treo mô đồng thể huyết thanh và phổi và được sử dụng để xác định trình tự liên ứng hệ gen hoàn chỉnh của virus P129 lần cấy truyền 0. ARN trước hết được môi bằng hexame ngẫu nhiên và được sử dụng để tổng hợp cADN. Hệ gen này được khuếch đại trong ba mẫu chồng lên nhau sử dụng PCR có độ chính xác cao (đọc và sửa). Các sản phẩm PCR từ ba phản ứng PCR riêng biệt (cho mỗi mảnh hệ gen) được tách dòng T/A và được sử dụng để phân tích trình tự ADN để tạo ra trình tự liên ứng hệ gen chiều dài hoàn chỉnh (xem trình tự nêu trong SEQ ID NO:1).

Lượng bằng nhau huyết thanh của cùng một con lợn được sử dụng để phân tích trình tự ADN, chứa P129 lần cấy truyền 0, được sử dụng để lây nhiễm ban đầu các tế bào đại thực bào phế nang lợn (porcine alveolar macrophage-PAM). Virus thế hệ sau từ

sự lây nhiễm PAM (cây truyền 1) được lọc qua thiết bị lọc xylan 0,1 micromet và được sử dụng để lây nhiễm các tế bào PK-9.

Các tế bào PK-9 là dòng tế bào chuyển gen được bắt nguồn bằng cách chuyển nhiễm ổn định dòng tế bào thận lợn PK0809 plasmit mã hóa bản sao bị xóa bớt của gen CD163 lợn và gen kháng neomycin. Chi tiết về cấu trúc và đặc trưng của dòng tế bào PK-9 đã được mô tả trước đây.

Việc làm thích ứng virus cây truyền lần 1 từ các tế bào PAM để sinh trưởng trên các tế bào PK-9 là khó khăn, và cần nỗ lực với nhiều dòng song song. Lây nhiễm được theo dõi bằng miễn dịch huỳnh quang các lỗ kép sử dụng kháng thể đơn dòng được tiếp hợp FITC SDOW17 đặc hiệu đối với protein nucleocapsit virus (Rural Technologies Inc, Brookings South Dakota). Các lần cấy truyền ban đầu thu được ổ bệnh rất nhỏ, nhưng không tạo ra đủ các hạt virus vô bào để bắt đầu lây nhiễm đơn lớp sạch. Các lần cấy truyền này được thực hiện bằng cách xử lý đơn lớp lây nhiễm bằng Accutaza (chất thay thế trypsin) và cấy lại các tế bào này vào trong nhiều lỗ bằng môi trường sạch, có hoặc không bổ sung các tế bào PK-9 không bị lây nhiễm. Sau một vài lần cấy truyền này, một số dòng thể hiện gia tăng rõ rệt tần suất và kích thước của ổ bệnh theo dõi bằng huỳnh quang. Một số dòng này đã thu được khả năng để được cấy truyền sử dụng dịch virus vô bào. Đến lần cấy truyền 17 (1 lần trên các tế bào PAM, 16 lần trên PK-9), một dòng có thể dễ dàng được duy trì sử dụng pha loãng các dịch vô bào từ các lần cấy truyền trước đó, và dẫn đến lây nhiễm toàn bộ đơn lớp trong một vài ngày. Virus này không gây ra tác dụng gây bệnh tế bào trên các tế bào PK-9 ở mức cấy truyền bất kỳ. ARN được chiết từ tế bào PK-9 lây nhiễm ở lần cấy truyền virus 17 và được sử dụng để xây dựng dòng cADN lây nhiễm.

Ví dụ 2

Xây dựng dòng cADN lây nhiễm của OF P129-PK lần cấy truyền 17

Dòng đơn tính cADN của virus P129-PK lần cấy truyền 17 được xây dựng, sử dụng khung plasmit như được mô tả trước đây. Hệ gen của virus này được khuếch đại bằng cách phiên mã ngược và PCR trong ba mảnh chồng lên nhau, có các vị trí endonucleaza giới hạn duy nhất có trong tự nhiên trong các vùng chồng lên nhau. Các sản phẩm từ ba phản ứng PCR riêng biệt được tách dòng và phân tích trình tự, và được nối thẳng để tạo ra trình tự liên ứng cho mỗi mảnh hệ gen. Nếu không một sản phẩm nào trong ba sản phẩm được tách dòng của mảnh xác định bắt cặp với trình tự axit amin

liên ứng được dự đoán cho mảnh này, một trong các dòng này được cải biến bằng cách tách dòng phụ và/hoặc gây phát sinh đột biến trực tiếp vị trí cho tới khi nó bắt cặp với trình tự axit amin liên ứng được dự đoán này. Ba mảnh của hệ gen và khung plasmid được nối sử dụng các kỹ thuật tách dòng chuẩn và các vị trí endonucleaza giới hạn. Dòng đơn tính có chiều dài hoàn chỉnh thu được, được chỉ định là pCMV-S-P129-PK17-FL, là gây nhiễm khi được chuyển nhiễm vào trong các tế bào PK-9. Trình tự của dòng cADN lây nhiễm này được xác định trong trình tự nêu trong SEQ ID NO:2. Hệ gen này là hầu như giống với virus cây truyền 17 mà từ đó nó được xây dựng, với các đầu mút đáng tin cậy, và thiếu các đoạn xen vào hoặc xóa bỏ bất kỳ liên quan đến các trình tự liên ứng của virus cây truyền 17. Không có các vị trí giới hạn được cải biến di truyền hoặc các sự thay đổi mục tiêu khác trong hệ gen virus của dòng cADN lây nhiễm.

Các sự khác biệt nucleotit và axit amin giữa các trình tự liên ứng của hệ gen hoàn chỉnh của P129 lần cây truyền 0 và trình tự hệ gen của dòng lây nhiễm pCMV-S-P129-PK17-FL được liệt kê riêng rẽ trong bảng 6. Bảng 6 nêu tất cả các sự khác biệt nucleotit và các sự khác biệt axit amin thu được trong vị trí trong hệ gen. Bộ nhỏ các đột biến này là tạo nên sự thay đổi phenotyp từ ức chế IFN γ (cây truyền 0) thành không ức chế IFN (tất cả các virus có nguồn gốc từ dòng đơn tính cADN lây nhiễm lần cây truyền 17). Bảng 7 tóm tắt nucleotit, axit amin, và các sự khác biệt axit amin không bảo toàn so với khung đọc mở (ORF) PRRSV hoặc protein không cấu trúc (non-structural protein-nsp). Cho các mục đích của bảng 7, các nhóm axit amin sau là nằm trong số các axit amin được xem là bảo toàn: [K, R], [D, E], [L, I, V, A], và [S, T].

Ví dụ 3

Các thể đột biến xóa bỏ trong P129-PK lần cây truyền 17

Xóa bỏ ở hai vùng của hệ gen được cải biến di truyền thành dòng cADN lây nhiễm pCMV-S-P129-PK17-FL, để tạo ra năm dòng lây nhiễm được cải biến di truyền.

Một vùng của hệ gen này được cải biến là trình tự định vị nhân (nuclear localization sequence-NLS) được đặt ở vị trí axit amin từ 41-47 của protein nucleocapsit (được mã hóa bởi PRRSV ORF 7). Hai kiểu xóa bỏ được tạo ra. Các sự xóa bỏ này đã được mô tả trước đó trong phạm vi dòng lây nhiễm PRRSV khác. Trình tự kiểu đại gồm các gốc axit amin từ 41-49 là PG...KN. Trong thể đột biến “d43/44”, cũng được biết là “PG--KS”, các gốc lysin 43 và 44 được xóa bỏ và gốc asparagin 49 được thay đổi thành serin. Trong thể đột biến “d43/44/46”, cũng được biết là “PG--S-KS”, các gốc lysin 43,

44, và 46 được xóa bỏ và gốc asparagin 49 được thay đổi thành serin. Các dòng lây nhiễm có nguồn gốc từ pCMV-S-P129-PK17-FL mà đưa vào các sự xóa bỏ này tương ứng là pCMV-S-P129-PK17-d43/44 và pCMV-S-P129-PK17-d43/44/46. Xem bằng sáng chế Mỹ số 7,544,362.

Vùng thứ hai của hệ gen được cải biến là vùng siêu biến nsp2, trong ORF1a. Sự xóa bỏ 131 axit amin (393 nucleotit) đã được mô tả trước đây trong phạm vi dòng lây nhiễm PRRSV khác. Dòng lây nhiễm có nguồn gốc từ pCMV-S-P129-PK17-FL mà đưa vào sự xóa bỏ này là pCMV-S-P129-PK17-nsp2.

Các dòng lây nhiễm mà kết hợp các sự xóa bỏ NLS và nsp2 cũng được tạo ra trong khung pCMV-S-P129-PK17-FL, và các dòng này được chỉ định là pCMV-S-P129-PK17-nsp2-d43/44 và pCMV-S-P129-PK17-nsp2-d43/44/46.

Ví dụ 4

Tạo ra và sinh trưởng của virus trên các tế bào PK-9

Sáu dòng lây nhiễm được mô tả trong ví dụ 3 được chuyển nhiễm vào trong các tế bào PK-9 để tạo ra sáu virus như được thể hiện trong bảng 1. Virus được tạo ra từ các dòng lây nhiễm này bằng cách lây nhiễm trực tiếp plasmit hình tròn vào trong các tế bào PK-9 sử dụng Lipofectamine 2000. Sau khi lây nhiễm, các virus thu hồi lại được cấy truyền lại liên tiếp trên các tế bào PK-9 để gia tăng nữa chuẩn độ và làm yếu virus. Các dung dịch gốc được tạo ra để thử nghiệm *in vitro* và đánh giá *in vivo* làm các dung dịch để chọn làm vaccin. Trong trường hợp virus P129-PK17-FL được bắt nguồn từ dòng lây nhiễm pCMV-S-P129-PK17-FL không cải biến, virus này được nuôi cấy cho tới khi đạt được tổng số 52 lần cấy truyền tự lộn. Hệ gen hoàn chỉnh của virus này được phân tích trình tự ở lần cấy truyền 24 (SEQ ID NO: 3) và 52 (SEQ ID NO:6).

Ví dụ 5

Các virus có nguồn gốc từ dòng cADN lây nhiễm lần cấy truyền 17 P129-PK đã được làm giảm khả năng ức chế sự cảm ứng IFN-alpha

Virus và tế bào. Tế bào MARC-145 và ST được nuôi trong môi trường Eagle cải biến (modified Eagle's medium-MEM) được bổ sung 5% huyết thanh thai bò (fetal bovine serum-FBS) và các kháng thể (50 µg/ml gentamixin, 100 UI penixilin và 100 µg/ml streptomycin). Các tế bào đại thực bào phế nang lợn ZMAC-1 được sinh trưởng trong RPMI-1640 có bổ sung 10% FBS. Chủng virus TGE của Purdue được tạo ra bằng

cách gây nhiễm các đơn lớp tế bào ST nhập dòng với số nhân là 0,01 trong môi trường Eagle cải biến. Virut cấy truyền được loại bỏ sau 1h và ủ các tế bào trong MEM có bổ sung 2,5% FBS ở 37°C trong khí quyển 5% CO₂. Giải phóng virut bằng cách làm lạnh và rã đông đơn lớp tế bào sau khi quan sát hiệu quả gây bệnh tế bào 80%. Dịch gốc virut TGE được ly tâm với 3.500 vòng/phút trong 15 phút ở 4°C và bảo quản ở -80°C cho tới khi sử dụng. Các dung dịch gốc virut (từ các tế bào PK-9) là như sau: P129-PK-FL và P129-PK-dnsp2-d43/44/46 là ở lần cấy truyền 8/25 (8 từ dòng lây nhiễm, 25 từ lợn). Bốn virut khác (P129-PK-d43/44, P129-PK-d43/44/46, P129-PK-dnsp2, và P129-PK-dnsp2-d43/44) là ở lần cấy truyền 21/38. Các dung dịch gốc hoạt động của các virut PRRS khác nhau được chuẩn bị bằng cách tạo ra một lần cấy truyền trên các tế bào ZMAC-1, ngoại trừ các vacxin thương mại Ingelvac PRRS MLV và Ingelvac PRRS ATP được hoàn nguyên theo hướng dẫn của nhà sản xuất và được sử dụng trực tiếp để tiêm.

Phân lập PBMC lợn. Máu lấy từ tĩnh mạch có heparin sạch được pha loãng bằng Hank's và PBMC được phân lập bằng ách ly tâm tỷ trọng thông qua gradien Ficoll-Hypaque 1077 (Sigma). Sau khi rửa hai lần trong Hank's, tạo huyền phù các tế bào này trong môi trường RPMI bằng L-glutamin (Mediatech) có bổ sung 5% huyết thanh thai bò (Gibco), 100 U/ml penixilin, 0,1 mg/ml streptomycin, 1 mM natri pyruvat, 1x axit amin không thiết yếu (Mediatech) 100 U/ml gentamixin và 250 mM 2-mercaptoetanol (Sigma).

Tinh chế các tế bào sợi nhánh từ tương bào lợn. Tinh chế các tế bào sợi nhánh từ tương bào lợn được thực hiện như được mô tả trước đây (Calzada-Nova, được đệ trình), và được dựa trên sự biểu hiện CD4 và CD172 đặc trưng của các tế bào này (Summerfield et al., 2003) Vắn tắt như sau: PBMC lợn sạch được tạo huyền phù trong PBS bằng 0,5% BSA và được đánh dấu bằng khối lượng tối ưu mAb nhận biết CD172 lợn (74-22-15, VMRD). Sau một lần rửa, sau đó ủ các tế bào này bằng kháng thể cừu kháng chuột thứ cấp được tiếp hợp với PE (Southern Biotech) và sau khi rửa, bằng FITC được đánh dấu kháng-CD4 (74-12-4, VMRD). PDC được phân loại trên máy phân loại tế bào phản chiếu (Reflection Cell Sorter-iCyt), các cổng phân loại được thiết lập dựa trên chủng loại CD4⁺/CD172^{thấp}. Sau khi phân loại, độ khiết của các tế bào được khẳng định bằng cách phân tích lại. Trong tất cả các trường hợp, độ tinh khiết là >95%.

Thử nghiệm đánh giá tiết xytokin. PBMC hoặc PDC được kích thích trong 16 h (37°C , 5 % CO_2) bằng các chất kích thích khác nhau hoặc được kích thích giả. Sau khi ủ, môi trường phủ các tế bào được kích thích này được thử nghiệm sự có mặt của IFN- α sử dụng ELISA kẹp được chuẩn bị có các kháng thể đơn dòng sẵn có trên thị trường (mAb kháng IFN- α lợn K9 và F17). Vắc xin như sau: các đĩa Immulon II (Dynatech Inc.) được phủ mAb kháng IFN- α lợn F17 (PBL Laboratories) qua đêm ủ ở 4°C sau đó chặn bằng môi trường RPMI có bổ sung 5% huyết thanh thai bò. Sau 1 giờ loại bỏ môi trường và năm microlit huyền phù được thử nghiệm bổ sung vào mỗi lỗ thử nghiệm theo hai mẫu. Sau 1h ủ, các lỗ thử nghiệm được rửa 4 lần và sau đó ủ lần lượt với mAb kháng IFN- α lợn, được đánh dấu biotin K9 (PBL Laboratories), streptavidin được tiếp hợp HRP (Zymed Laboratories), và cơ chất TMB (KPL). Mật độ quang học được xác định bằng máy đọc đĩa ELISA.

Virut được bắt nguồn từ dòng đơn tính cADN lây nhiễm lần cấy truyền 17 P129-PK không có khả năng ức chế sản sinh IFN-alpha. Các dịch gốc virut hoạt động được chuẩn bị từ nhóm gồm bốn mẫu phân lập virut kiểu đại PRRS khác nhau (P3412, P129, IND5, NADC20) sử dụng dòng tế bào đại thực bào phế nang lợn ZMAC-1. Các dịch gốc bổ sung cũng được điều chế từ các dẫn xuất của hai virut kiểu đại thứ nhất mà đã được thích ứng để sinh trưởng trong nuôi cấy tế bào bằng cách cấy truyền lặp lại trong các tế bào PK-9, FK.D4 hoặc MARC-145. Ba trong số bốn mẫu phân lập kiểu đại (P129, IND5, NADC20) sinh trưởng dễ dàng và hiệu quả trong các tế bào ZMAC-1 đến chuẩn độ khoảng 10^7 TCID₅₀/ml, trong khi mẫu phân lập kiểu đại P3412 đạt đến chuẩn độ chỉ là 10^5 TCID₅₀/ml. Đáng chú ý là, các dịch gốc virut P129 được điều chế trong dòng tế bào ZMAC-1 đạt được chuẩn độ cao hơn 10-lần so với chuẩn độ thu được trong các tế bào PK-9 hoặc MARC-145 mà đã thích ứng virut này. Kiểm tra khả năng của các virut này kích thích tiết IFN- α bởi PBMC thấy rằng với một ngoại trừ (mẫu phân lập P3412 dòng C), khối lượng rất nhỏ (<50 pg) IFN- α được tiết bởi các tế bào này để đáp ứng với việc tiếp xúc chúng với dịch gốc virut PRRS thử nghiệm bất kỳ, mà là không đáng kể khi so sánh với việc tiết dư thừa IFN- α (17,540 pg) được sản sinh bởi cùng các tế bào này do tiếp xúc chúng với coronavirut, virut viêm dạ dày-ruột lan truyền lợn (TGEv).

PRRSV không chỉ không có khả năng kích thích sản sinh IFN- α bởi PBMC lợn nhưng ức chế hiệu quả việc sản sinh chúng. Hiệu quả ức chế của dịch gốc PRRSV được xác định bằng cách đo khối lượng IFN- α được tiết bởi PBMC đáp ứng với việc tiếp xúc với TGEv khi có hoặc không có PRRSV. Như được thể hiện trong bảng 2, tất cả 4 mẫu phân lập virus PRRS kiểu đại được thử nghiệm, cũng như tất cả các dẫn xuất được thích nghi nuôi cấy tế bào, thể hiện hiệu quả ức chế mạnh (>80%) đối với đáp ứng IFN- α của PBMC với TGEv. Phân tích nhóm dịch gốc virus được bắt nguồn từ dòng cADN lây nhiễm (pCMV-S-P129-PK17-FL), bao gồm virus P129-PK-FL có chiều dài hoàn chỉnh và một vài thể đột biến được cải biến di truyền bằng cách xóa bỏ được tiến hành. Như được thể hiện trong bảng 3, khi so sánh với hiệu quả ức chế mạnh (95%) của mẫu phân lập kiểu đại ban đầu P129 (lần cấy truyền 1), virus P129-PK-FL và tất cả các thể đột biến bằng cách xóa bỏ thể hiện khả năng giảm mạnh ức chế sự cảm ứng IFN- α bởi TGEv trong PBMC. Để đánh giá tiếp phenotyp IFN- α của các virus này, các thí nghiệm sau đó được tập trung vào thực hiện các so sánh trực tiếp giữa các virus P129-PK-FL và P129-PK-d43/44, chủng kiểu đại P129 ban đầu, và/hoặc hai vacxin PRRSV sống cải biến sẵn có trên thị trường được sản xuất bởi Boehringer Ingelheim (Ingelvac PRRS MLV và Ingelvac PRRS ATP). Mẫu phân lập tham chiếu cấy truyền ít bổ sung, NVSL-14, cũng được thử nghiệm. Như được thể hiện trong bảng 4, trong bốn thí nghiệm độc lập, P129-PK-FL và P129-PK-d43/44 thể hiện hiệu quả ức chế IFN- α kém hơn đáng kể so với virus P129 ban đầu, hai chủng làm yếu Ingelvac, hoặc chủng tham chiếu. Trong một trường hợp, đồng lây nhiễm bằng virus P129-PK-FL hoặc P129-PK-d43/44 dẫn đến tăng cường rõ rệt đáp ứng IFN- α đối với TGEv.

Các được thể hiện trong bảng 2 là chỉ định về hiệu quả ức chế interferon- α của virus PRRS kiểu đại và các dẫn xuất được điều chỉnh để sinh trưởng trong nuôi cấy tế bào. Các dịch gốc virus PRRS chỉ định, được sinh trưởng trong các tế bào ZMAC-1 và chuẩn độ các dịch gốc mới được tạo ra này được xác định sử dụng các tế bào ZMAC-1. Khối lượng IFN- α có trong phần nổi dịch nuôi cấy các tế bào đơn nhân máu ngoại vi lợn được tiếp xúc trong 18 h với dịch gốc virus PRRS chỉ định có hoặc không có virus TGE được xác định bằng ELISA. * Đáp ứng với chỉ mình TGEv.

Các kết quả được thể hiện trong bảng 3 chứng minh hiệu quả ức chế interferon- α của virus PRRS kiểu đại P129 và các dẫn xuất cải biến di truyền của nó được thích ứng

để sinh trưởng trong các tế bào PK-9 biểu hiện CD163. Khối lượng IFN- α có trong phần nổi nuôi cấy của các tế bào đơn nhân máu ngoại vi lợn (PBMC) được tiếp xúc trong 18 h với dịch gốc virus PRRS chỉ định có hoặc không có virus TGE được xác định bằng ELISA. na = không thể áp dụng; *Đáp ứng với chỉ mình TGEV.

Bảng 4 thể hiện hiệu quả ức chế interferon- α giảm của virus P129-PK-FL và P129-PK-d43/44 khi so sánh với virus P129 kiểu dại và vaccin PRRS Ingelvac. Khối lượng IFN- α có trong phần nổi nuôi cấy của tế bào đơn nhân máu ngoại vi lợn (PBMC) được tiếp xúc trong 18 h với dịch gốc virus PRRS chỉ định có hoặc không có virus TGE được xác định bằng ELISA. na = không thể áp dụng; *đáp ứng với chỉ mình TGEV.

Khối lượng lớn IFN- α được tiết bởi PBMC để đáp ứng với sự tiếp xúc chúng với TGEV được bắt nguồn ban đầu từ nhóm nhỏ các tế bào chứa ít hơn 0,3% chủng loại PBMC. Nhóm nhỏ tế bào này quan trọng nhưng ít có bao gồm các tế bào sợi nhánh tương bào (plasmacytoid dendritic cell-PDC), mà nhận cùng tên do hình thái tương bào đặc trưng. Để xem xét tiếp phenotyp IFN- α của các virus P129-PK-FL và P129-PK-d43/44, một loạt các thí nghiệm được thực hiện tương tự như các thí nghiệm được mô tả trên đây, ngoại trừ PDC được phân lập mới từ PBMC đến độ tinh khiết >95% được sử dụng. Như được thể hiện trong Fig. 1, loạt thí nghiệm này khẳng định rằng các virus P129-PK-FL và P129-PK-d43/44 gây ra ức chế không đáng kể sự cảm ứng IFN- α do TGEV. Hơn nữa, trong một thí nghiệm hiệu quả tăng cường rõ ràng được quan sát đối với sự cảm ứng IFN- α qua trung gian TGEV bởi PDC để đáp ứng với các virus P129-PK-FL và P129-PK-d43/44 PRRS. Ngược lại, virus Ingelvac PRRS MLV thể hiện hiệu quả ức chế mạnh đối với đáp ứng IFN- α , như được thể hiện trong Fig. 1.

Các kết quả được mô tả trong phần thí nghiệm cho thấy rằng các virus P129-PK-FL và P129-PK-d43/44 PRRS, cũng như các dẫn xuất khác của dòng cADN lây nhiễm pCMV-S-P129-PK17-FL, có khả năng ức chế IFN- α bởi TGEV giảm đáng kể trong các tế bào PBMC hoặc PDC lây nhiễm. Kết quả này là đối lập rõ rệt với hiệu quả kiểm chế IFN quan sát được với virus PRRS kiểu dại (cây truyền thấp) và với hai vaccin virus sống cải biến sẵn có trên thị trường (Ingelvac PRRS MLV và Ingelvac PRRS ATP). Quan sát rằng các virus P129-PK-FL và P129-PK-d43/44 kiểm chế tối thiểu chức năng PDC quan trọng này cho thấy rõ rệt vai trò chính của các tế bào này làm trung gian cho miễn dịch bẩm sinh kháng lại việc nhiễm virus.

Cũng nên được lưu ý rằng sáng chế đề xuất các virus sử dụng cho vaccin thương mại có hiệu quả lâm sàng được điều chỉnh để sinh trưởng trên các tế bào cho phép biểu hiện tái tổ hợp thụ thể CD163, và các virus và vaccin này không phụ thuộc vào, cũng không được phát triển ở điểm bất kỳ bằng công nghệ nuôi cấy “tế bào khi” trước đó. Cụ thể xem bằng sáng chế Mỹ số 7,754,464.

Ví dụ 6

Tính an toàn và hiệu quả của các vaccin để chọn

Để đánh giá tính an toàn và hiệu quả làm vaccin kháng lại PRRS, có ba virus bắt nguồn từ dòng cADN lây nhiễm pCMV-S-P129-PK17-FL, P129-PK-FL (lần cấy truyền 7/24), P129-PK-d43/44 (lần cấy truyền 17/34) và P129-PK-d43/44/46 (lần cấy truyền 17/34) được đánh giá trong mô hình bệnh hô hấp ở lợn nhỏ. Nguồn gốc của các virus này được thể hiện trong Fig. 8, và thiết kế thí nghiệm (các nhóm điều trị) được liệt kê trong bảng 5. Mẫu phân lập virus PRRSV cấy truyền thấp NADC20 được sử dụng để kích thích khác loại khi 7 tuần tuổi (bốn tuần sau khi dùng vaccin). Các nhóm điều trị đối chứng bao gồm vaccin giả và vaccin PRRS thương mại Ingelvac MLV.

Các nhóm không điều trị (Non-treated-NT) là như sau: Lợn NT1 là lợn báo hiệu để theo dõi tình trạng sức khỏe của các con lợn nguồn. Chúng được nhốt riêng biệt và được mổ tử thi khám nghiệm trước khi kích thích PRRSV. Lợn NT2 là các đối chứng tiếp xúc được nhốt riêng biệt trong chuồng nằm giữa hai chuồng nhốt lợn được dùng vaccin, tổng số hai cho mỗi nhóm điều trị cho mỗi từ T02 đến T05. Lợn NT3 là đối chứng tiếp xúc được nhốt một con cho mỗi chuồng với lợn dùng vaccin, tổng số hai cho mỗi nhóm điều trị (T01 đến T05). Chỉ lợn NT3 được chỉ định cho nhóm T01.

Nhiệt độ trực tràng của mỗi con vật dùng vaccin được đo sau khi dùng vaccin và so sánh với nhóm điều trị T01 (vaccin giả). Các kết quả này được thể hiện trong Fig. 1. Không một vaccin nào gây sốt. Tất cả các nhóm trung bình thấp hơn 104°C trong suốt khoảng thời gian quan sát sau khi dùng vaccin.

Nhiệt độ trực tràng của lợn được đo sau khi kích thích. Các kết quả được thể hiện trong Fig. 2. Lợn T01 không được dùng vaccin cho thấy sốt kéo dài lớn hơn 104°C nhiều. Ngược lại, ba trong số các vaccin giảm đáng kể sốt sau khi kích thích. P129-PK-FL là hữu hiệu nhất giảm sốt.

Trọng lượng cơ thể động vật được ghi lại trước và sau khi dùng vaccin. Các keets quả này được thể hiện trong Fig. 3. Trọng lượng cơ thể được ghi lại vào các ngày -1 (trước khi dùng vaccin), 10, 24, 27 (sau khi kích thích), và 37 của nghiên cứu. Trong lợn không dùng vaccin, kích thích bằng virus NADC20 có virus hầu như hoàn toàn không tăng cân trong khoảng thời gian quan sát 10 ngày. Vaccin này không gây tác dụng này ở các mức khác nhau.

Phổi của các con vật bị kích thích được xem xét sau khi kích thích rồi mổ khám tử thi. Phần trăm mỗi phổi có xuất hiện các tổn thương được thể hiện trong Fig. 4. Nhóm dùng vaccin giả T01 có trung bình 25,1% phổi tổn thương. Vaccin làm giảm tổn thương phổi đến các mức độ khác nhau. P129-PK-FL là hiệu quả nhất, giảm tổn thương phổi xuống đến 1,1%.

Độ trầm trọng của các tổn thương phổi được đánh giá sử dụng điểm số đánh giá phổi (lung assessment score-LAS) như được thể hiện trong Fig. 5. Ba trong số các vaccin giảm LAS. P129-PK-FL giảm LAS trung bình từ 1,63 trong nhóm dùng vaccin giả xuống 0,14.

Các kháng thể huyết thanh đối với PRRSV được cảm ứng bởi cả hai dùng vaccin và kích thích. Tỷ lệ IDEXX ELISA S/P được đo vào các ngày 27 và 37 của nghiên cứu này. Các kết quả được thể hiện trong Fig. 6. Dùng vaccin P129-PK-FL cảm ứng mức kháng thể kháng PRRS cao nhất.

Virut huyết trong huyết thanh của lợn bị kích thích được đo chuẩn độ trên các tế bào PAM. Các kết quả ($TCID_{50}/mL$) được đưa ra trong Fig. 7. P129-PK-FL là hiệu quả nhất làm giảm virus huyết sau khi kích thích.

Mặc dù sáng chế đã được mô tả bằng cách tham khảo đến các ví dụ trên đây và các tài liệu, toàn bộ nội dung của các tài liệu này được đưa vào đây bằng cách viện dẫn toàn bộ nội dung của nó, sẽ được hiểu rằng các cải biến và biến đổi được bao hàm trong tinh thần và phạm vi của sáng chế. Mặc dù, sáng chế chỉ giới hạn đến các yêu cầu bảo hộ sau.

Ví dụ 7

Thu dòng cADN lây nhiễm PCMV-S-P129-PK17-FL từ dòng cADN lây nhiễm PCMV-S-P129

Dòng cADN lây nhiễm PRRSV theo sáng chế pCMV-S-P129-PK17-FL có thể thu được dễ dàng từ cADN lây nhiễm PRRSV pCMV-S-P129 được mô tả trước đây bởi người có hiểu biết trung bình trong lĩnh vực, sử dụng kỹ thuật phát sinh đột biến trực tiếp vị trí. Dòng cADN lây nhiễm PRRSV pCMV-S-P129 được mô tả trong bằng sáng chế Mỹ số 6,500,662 và được lưu giữ bởi ATCC dưới số lưu giữ 203489. Trình tự ADN của hệ gen PRRSV trong dòng này cũng có sẵn trong cơ sở dữ liệu Genbank (NCBI) là số hiệu AF494042. Kit gây phát sinh đột biến trực tiếp vị trí là sẵn có trên thị trường từ một số nhà cung cấp, và có khả năng tạo ra một số sự thay đổi nucleotit đồng thời ở nhiều vị trí trong plasmit lớn. Các kit này bao gồm, nhưng không bị giới hạn đến, Change-IT™ Multiple Mutation Site Directed Mutagenesis Kit (Affymetrix / USB), QuikChange Lightning Multi Site-Directed Mutagenesis Kit (Agilent Technologies - Stratagene Products), và AMAP Multi Site-directed Mutagenesis Kit (MBL International).

Danh sách các thay đổi nucleotit giữa dòng cADN lây nhiễm PRRSV pCMV-S-P129 (sẵn có từ ATCC) và dòng cADN lây nhiễm PRRSV theo sáng chế pCMV-S-P129-PK17-FL được trình bày trong bảng 8. Tất cả các sự thay đổi là ở trong vùng mã hóa protein của hệ gen. Có tổng số 74 thay đổi nucleotit, mà có thể được đưa vào trong dòng lây nhiễm pCMV-S-P129 sử dụng 74 môi đột biến và nhiều phản ứng liên tiếp với kit phát sinh đột biến trực tiếp vị trí thương mại, thu được phân tử plasmit mà là giống trong trình tự với pCMV-S-P129-PK17-FL được mô tả ở đây. Trong thực tế, có thể nhận được cùng kết quả với ít hơn 74 môi đột biến, vì các đám đột biến trong khoảng từ 50-60 nucleotit của nhau có thể được thay đổi sử dụng một môi phát sinh đột biến. Ví dụ, nucleotit 735, 750, và 756 có thể được thay đổi sử dụng một môi phát sinh đột biến, như có thể làm với các nucleotit 965, 992, và 1009. Do đó số các môi giảm xuống khoảng là 60.

Trong số 74 thay đổi nucleotit, phần lớn (42) là đồng nghĩa hoặc “câm”, nghĩa là chúng mã hóa axit amin giống nhau. Các sự thay đổi nucleotit này là không chắc chắn có hiệu quả có thể thấy được đối với phenotype cảm ứng hoặc ức chế interferon của virus này. Sự thay đổi 32 nucleotit còn lại là không đồng nghĩa hoặc “không câm”, và dẫn đến các sự thay đổi axit amin các protein virus. Các sự thay đổi 32 nucleotit này được dự đoán là trực tiếp tạo ra phenotyp cảm ứng/ức chế interferon của virus này, và nên được thay đổi để chuyển virus được mã hóa bởi dòng lây nhiễm pCMV-S-P129

thành phenotyp interferon tương tự như được thể hiện bởi virus được mã hóa bởi dòng lây nhiễm pCMV-S-P129-PK17-FL. Sự thay đổi này sẽ cần ít nhất 32 mỗi đột biến, ít hơn nếu tính đến tập hợp gồm một số nucleotit thích hợp.

Ví dụ 8

Tổng hợp mới dòng cADN PCMV-S-P129-PK17-FL

Theo cách khác để phát sinh đột biến trực tiếp vị trí, hệ gen virus PRRS theo sáng chế có thể được tổng hợp mới bằng phương pháp hóa học, bằng các trình tự gá lắp ở đầu 5' và 3' thích hợp, và được tách dòng vào trong khung plasmit được sử dụng cho dòng cADN lây nhiễm PRRS pCMV-S-P129 (sẵn có từ ATCC là số hiệu 203489) hoặc khung plasmit tương tự. Tổng hợp theo đặt hàng các gen dài hơn 50 kb (hệ gen PRRSV là khoảng 15,5 kb) sẵn có là dịch vụ thương mại từ một số nhà bán lẻ, bao gồm (nhưng không bị giới hạn đến): GenScript, DNA 2.0, và Bio Basic Inc. Hệ gen virus tổng hợp được tách dòng trực tiếp vào trong vector pCMV-S bằng cách thay thế hệ gen virus trong dòng lây nhiễm pCMV-S-P129 sử dụng các vị trí enzyme giới hạn 5' *PacI* và 3' *SpeI* đặt ở bên hệ gen này. Để cắt hệ gen tổng hợp, đoạn kéo dài 24-nucleotit (5'-GCAGAGCTCGTTAATAAACCGTC-hệ gen-3, đưa vào vị trí *PacI* được gạch chân) được xây dựng thành đầu 5' của hệ gen tổng hợp, và đoạn kéo dài 83-nucleotit (5'-hệ gen-

AAAAAAAAAAAAAAAAAAAAAAAAATGCATATTTAAATCCCAAGCCGAATTCCAG CAACTGGCGGCCGTTACTAGTGAGCGGCCGC-3', mà đưa vào vị trí *SpeI* được gạch chân) được xây dựng thành đầu 3' của hệ gen tổng hợp. Sau khi cắt plasmit và hệ gen tổng hợp bằng *PacI* và *SpeI*, tinh chế mảnh thu được, nối sử dụng ADN ligaza, và được biến nạp vào trong *Escherichia coli* để tiết và sàng lọc và nhân lên sử dụng các kỹ thuật tách dòng chuẩn được biết nhiều bởi người có hiểu biết trung bình trong lĩnh vực.

Lưu giữ nguyên liệu sinh học

Nguyên liệu sinh học sau (xem bằng sáng chế Mỹ số 6,500,662) được lưu giữ ở bảo tàng giống chuẩn Mỹ (American Type Culture Collection-ATCC) ở 10801 University Blvd., Manassas, Virginia, 20110-2209, USA vào ngày 19.11. 1998, và được chỉ định các số hiệu sau.

Plasmit pT7P129A, số chỉ định 203488

Plasmit pCMV-S-P129, số chỉ định 203489

Toàn bộ nội dung của các bằng sáng chế Mỹ số sau được đưa vào đây bằng cách viện dẫn: US 6,500,662, và US 7,618,797.

Bảng 1.

Dòng lây nhiễm	Virus
pCMV-S-P129-PK17-FL	P129-PK-FL
pCMV-S-P129-PK17-d43/44	P129-PK-d43/44
pCMV-S-P129-PK17-d43/44/46	P129-PK-d43/44/46
pCMV-S-P129-PK17-nsp2	P129-PK-nsp2
pCMV-S-P129-PK17-nsp2-d43/44/46	P129-PK-nsp2-d43/44
pCMV-S-P129-PK17-nsp2-d43/44/46	P129-PK-nsp2-d43/44/46

Bảng 2.

M Ã U	Dịch gốc PRRSv	DÒNG NG	Lần cấy truyền/ Loại tế bào	Chuẩn độ virus PRRS ước tính (TCID ₅₀) trong dung dịch gốc ban đầu	Chuẩn độ virus PRRS (TCID ₅₀) trong dịch gốc virus PRRS được tạo ra trong các tế bào ZMAC-1	IFN- α (ng/ml) được tạo ra bởi PBMC đáp ứng với chỉ mình PRRSv	IFN- α (ng/ml) được tạo ra bởi PBMC đáp ứng với PRSSv + TGEv	Ước chế đáp ứng IFN- α với TGEv (%)
1	P3412	wt	0 (huyết thanh)	nd	10 ⁵	< 0,04	< 0,04	99
2	P3412	A	41/PK	10 ³	10 ³	< 0,04	1,166	93
3	P3412	C	41/PK	10 ³	10 ³	0,464	3,376	81
4	P3412	A	43/FK	10 ⁵	10 ⁶	< 0,04	< 0,04	99
5	P3412	B	43/FK	10 ⁵	10 ⁵	< 0,04	1,86	89
6	P129	wt	0 (huyết thanh)	nd	10 ⁷	< 0,04	0,327	98
7	P129	A	63/PK	10 ⁴	10 ⁸	< 0,04	< 0,04	99
8	P129	B	63/PK	10 ⁴	10 ⁸	< 0,04	2,3	87
9	P129	A	60/FK	10 ⁵	10 ⁶	0,05	2,151	88

10	P129	B	60/FK	10 ⁵	10 ⁶	< 0,04	< 0,04	99
11	P129	A-1	51/MARC	10 ⁵	10 ⁸	< 0,04	0,686	96
12	P129	A-1	151/MARC	10 ⁵	10 ⁸	< 0,04	0,04	99
13	NADC 20	wt	0 (huyết thanh)	nd	10 ⁷	< 0,04	< 0,04	99
14	IND5	wt	0 (huyết thanh)	nd	10 ⁷	< 0,04	< 0,04	99
-	Giả	-	-	na	na	0	17,54*	na

Bảng 3.

Thí nghiệm	Dung dịch gốc virus PRRS	IFN- α (ng/ml) được tạo ra bởi PBMC đáp ứng với chỉ mình PRRSv	IFN- α (ng/ml) được tạo ra bởi PBMC đáp ứng với PRRSv + TGEv	Sự ức chế TGEv cảm ứng đáp ứng IFN- α (%)
1	P129	0,09	0,64	95
	P129-PK-FL	0,53	7,37	38
	P129-PK-d43/44	0,85	8,59	28
	P129-PK-d43/44/46	1,01	7,65	36
	P129-PK- dnsp2d43/44	1,34	10,28	24
	P129-PK- dnsp2d43/44/46	0,38	8,28	31
	P129-PK-dnsp2	0,04	4,86	60
	Giả	na	12,16*	na

Bảng 4

Thí nghiệm	Dung dịch gốc virus PRRS	IFN- α (ng/ml) được tạo ra bởi PBMC đáp ứng với chỉ mình PRRSv	IFN- α (ng/ml) được tạo ra bởi PBMC đáp ứng với PRRSv + TGEv	Ức chế TGEv cảm ứng đáp ứng IFN- α (%)
1	P129	0,04	9,75	81

Thí nghiệm	Dung dịch gốc virus PRRS	IFN- α (ng/ml) được tạo ra bởi PBMC đáp ứng với chỉ mình PRRSv	IFN- α (ng/ml) được tạo ra bởi PBMC đáp ứng với PRRSv + TGEv	Ức chế TGEv cảm ứng đáp ứng IFN- α (%)
	P129-PK-FL	0,07	60,22	0
	P129-PK-d43/44	0,13	70,10	0
	Ingelvac PRRS	0,04	13,05	74
	Ingelvac PRRS ATP	0,04	9,10	82
	NVSL-14	0,01	13,35	74
	Giả	na	50,06*	na
2	P129-PK-FL	0,126	13,52	26
	P129-PK-d43/44	0,127	17,18	6
	Ingelvac PRRS	0,04	3,09	83
	Ingelvac PRRS ATP	0,04	3,11	83
	Giả	na	18,21*	na
3	P129-PK-FL	0,04	8,28	6
	P129-PK-d43/44	0,06	8,26	6
	Ingelvac PRRS	0,04	3,56	60
	Ingelvac PRRS ATP	0,04	4,15	53
	Giả	na	8,84*	na
4	P129-PK-FL	0,05	12,97	7
	P129-PK-d43/44	0,05	13,57	3
	Ingelvac PRRS	0,04	6,07	57
	Ingelvac PRRS ATP	0,04	5,30	62
	Giả	na	13,95*	na

Bảng 5

TX	IVP	Lần cấy truyền hoặc Loạt số	Dòng tế bào	Chế độ	Thử tích cho mỗi liều IM	Số lợn
NT1*	NA	NA	NA	NA	NA	3
NT2*	NA	NA	NA	NA	NA	8
NT3*	NA	NA	NA	NA	NA	10

T01	Giả	NA	NA	Ngày 0	2 mL	12
T02	P129-PK-d43/44	17/34	PK9	Ngày 0	2 mL	12
T03	P129-PK-d43/44/46	17/34	PK9	Ngày 0	2 mL	12
T04	P129-PK-FL	7/24	PK9	Ngày 0	2 mL	12
T05	BI (Ingelvac MLV)	NA	Thận khi	Ngày 0	2 mL	12

Bảng 6

Vị trí trong hệ gen	Lần cấy truyền 0 nucleotit	Lần cấy truyền 17 nucleotit	Protein virus bị ảnh hưởng	Vị trí axit amin	Lần cấy truyền 0 axit amin	Lần cấy truyền 17 axit amin
407	C	T	Nsp1a	72	P	P
612	C	T	Nsp1a	141	H	Y
735	G	A	Nsp1b	2	D	N
750	A	G	Nsp1b	7	S	G
756	G	A	Nsp1b	9	D	N
992	A	T	Nsp1b	87	E	D
1009	G	A	Nsp1b	93	C	Y
1096	C	A	Nsp1b	122	P	H
1215	C	T	Nsp1b	162	L	L
1620	G	A	Nsp2	94	A	T
1786	C	A	Nsp2	149	T	K
1793	C	T	Nsp2	151	G	G
1808	T	C	Nsp2	156	D	D
1841	C	T	Nsp2	167	C	C
2106	A	G	Nsp2	256	I	V
2164	T	G	Nsp2	275	M	R
2185	T	C	Nsp2	282	M	T
2318	A	G	Nsp2	326	S	S
2403	G	T	Nsp2	355	V	L
2591	G	A	Nsp2	417	L	L
2804	C	T	Nsp2	488	D	D

Vị trí trong hệ gen	Lần cây truyền 0 nucleotit	Lần cây truyền 17 nucleotit	Protein virut bị ảnh hưởng	Vị trí axit amin	Lần cây truyền 0 axit amin	Lần cây truyền 17 axit amin
3019	A	G	Nsp2	560	Y	C
3074	T	C	Nsp2	578	S	S
3167	C	T	Nsp2	609	D	D
3214	G	A	Nsp2	625	R	K
3563	C	T	Nsp2	741	I	I
3740	A	G	Nsp2	800	A	A
4154	T	C	Nsp2	938	C	C
4477	A	C	Nsp3	72	K	T
4643	A	G	Nsp3	127	V	V
4705	T	C	Nsp3	148	V	A
4736	C	T	Nsp3	158	P	P
5231	C	T	Nsp3	323	I	I
5324	C	T	Nsp3	354	L	L
5393	G	A	Nsp3	377	L	L
5498	A	G	Nsp3	412	L	L
5851	C	A	Nsp4	84	A	E
5855	T	C	Nsp4	85	D	D
5909	C	T	Nsp4	103	V	V
5917	G	A	Nsp4	106	S	N
5985	C	A	Nsp4	129	L	I
6505	C	T	Nsp5	98	A	V
6644	T	C	Nsp5	144	F	F
6653	T	C	Nsp5	147	R	R
7419	G	A	Nsp7	217	D	N
8032	A	G	Nsp9	162	A	A
8074	C	T	Nsp9	176	G	G
8200	A	G	Nsp9	218	G	G
8593	T	C	Nsp9	349	P	P
8831	G	A	Nsp9	429	V	I
8911	T	C	Nsp9	455	N	N
9160	A	G	Nsp9	538	E	E

Vị trí trong hệ gen	Lần cây truyền 0 nucleotit	Lần cây truyền 17 nucleotit	Protein virus bị ảnh hưởng	Vị trí axit amin	Lần cây truyền 0 axit amin	Lần cây truyền 17 axit amin
9568	A	G	Nsp9	674	L	L
9714	T	C	Nsp10	38	I	T
10,271	C	T	Nsp10	224	L	L
10,627	T	C	Nsp10	342	V	V
11,265	G	A	Nsp11	114	G	E
11,512	T	C	Nsp11	196	S	S
11,913	T	C	Nsp12	107	I	T
12,487	A	T	ORF2a	144	E	V
12,876	C	T	ORF3	66	A	V
12,949	A	G	ORF3	90	L	L
13,575	G	A	ORF4	117	V	V
13,857	T	G	ORF5	29	V	G
13,869	A	G	ORF5	33	N	S
13,872	C	G	ORF5	34	T	S
14,102	G	A	ORF5	111	V	I
14,143	C	T	ORF5	124	V	V
14,257	G	A	ORF5	162	E	E
14,287	C	T	ORF5	172	N	N
14,379	T	C	ORF6	7	D	D
14,546	T	C	ORF6	63	V	A
14,578	A	G	ORF6	74	T	A
14,780	C	T	ORF6	141	T	I
14,932	T	C	ORF7	20	N	N
14,973	G	A	ORF7	34	S	N
15,124	T	C	ORF7	84	N	N
15,288	G	A	3'UTR	-	-	-

Bảng 7

ORF hoặc nsp	Các sự khác biệt nucleotit	Các sự khác biệt axit amin	Các sự khác biệt axit amin không bảo toàn
Nsp1a	2	1	1

Nsp1b	7	6	5
Nsp2	19	8	5
Nsp3	8	2	1
Nsp4	5	3	2
Nsp5	3	1	0
Nsp6	0	0	0
Nsp7	1	1	1
Nsp8	0	0	0
Nsp9	8	1	0
Nsp10	3	1	1
Nsp11	2	1	1
Nsp12	1	1	1
ORF2a	1	1	1
ORF2b	0	0	0
ORF3	2	1	0
ORF4	1	0	0
ORF5	7	4	2
ORF6	4	3	2
ORF7	3	1	1

Bảng 8.

Vị trí trong hệ gen	Lần cây truyền 10 trên các tế bào MARC- 145 nucleotit	Cây truyền 17 trên cá tế bào PK-9 nucleotit	Protein virut bị ảnh hưởng	Vị trí axit amin	Cây truyền 10 (trên các tế bào MARC- 145) axit amin	Cây truyền 17 (trên các tế bào PK- 9) axit amin
407	C	T	Nsp1a	72	P	P
612	C	T	Nsp1a	141	H	Y
735	G	A	Nsp1b	2	D	N
750	A	G	Nsp1b	7	S	G
756	G	A	Nsp1b	9	D	N
965	T	C	Nsp1b	78	G	G

Vị trí trong hệ gen	Lần cây truyền 10 trên các tế bào MARC-145 nucleotit	Cây truyền 17 trên cá tế bào PK-9 nucleotit	Protein virus bị ảnh hưởng	Vị trí axit amin	Cây truyền 10 (trên các tế bào MARC-145) axit amin	Cây truyền 17 (trên các tế bào PK-9) axit amin
992	A	T	Nsp1b	87	E	D
1009	G	A	Nsp1b	93	C	Y
1096	C	A	Nsp1b	122	P	H
1215	C	T	Nsp1b	162	L	L
1376	A	G	Nsp2	12	A	A
1395	T	C	Nsp2	19	C	R
1871	A	G	Nsp2	177	L	L
2185	T	C	Nsp2	282	M	T
2235	G	A	Nsp2	299	D	N
2403	G	T	Nsp2	355	V	L
2731	A	G	Nsp2	464	Y	C
2804	C	T	Nsp2	488	D	D
2918	T	C	Nsp2	526	S	S
3019	A	G	Nsp2	560	Y	C
3067	G	A	Nsp2	576	G	E
3074	T	C	Nsp2	578	S	S
3214	G	A	Nsp2	625	R	K
3256	T	C	Nsp2	639	L	S
3563	C	T	Nsp2	741	I	I
3740	A	G	Nsp2	800	A	A
4154	T	C	Nsp2	938	C	C
4477	A	C	Nsp3	72	K	T
4643	A	G	Nsp3	127	V	V
4705	T	C	Nsp3	148	V	A
4736	C	T	Nsp3	158	P	P
4784	C	T	Nsp3	174	V	V
5231	C	T	Nsp3	323	I	I

Vị trí trong hệ gen	Lần cấy truyền 10 trên các tế bào MARC-145 nucleotit	Cấy truyền 17 trên cá tế bào PK-9 nucleotit	Protein virus bị ảnh hưởng	Vị trí axit amin	Cấy truyền 10 (trên các tế bào MARC-145) axit amin	Cấy truyền 17 (trên các tế bào PK-9) axit amin
5324	C	T	Nsp3	354	L	L
5498	A	G	Nsp3	412	L	L
5855	T	C	Nsp4	85	D	D
5862	G	A	Nsp4	88	A	T
5985	C	A	Nsp4	129	L	I
6155	T	C	Nsp4	185	D	D
6505	C	T	Nsp5	98	A	V
6776	A	G	Nsp7	2	L	L
7419	G	A	Nsp7	217	D	N
7521	T	C	Nsp7	251	S	P
8032	A	G	Nsp9	162	A	A
8074	C	T	Nsp9	176	G	G
8200	A	G	Nsp9	218	G	G
8263	T	C	Nsp9	239	S	S
8485	T	C	Nsp9	313	H	H
8593	T	C	Nsp9	349	P	P
8831	G	A	Nsp9	429	V	I
8911	T	C	Nsp9	455	N	N
9022	A	G	Nsp9	492	L	L
9714	T	C	Nsp10	38	I	T
9934	C	T	Nsp10	111	N	N
10,237	G	A	Nsp10	212	L	L
10,271	C	T	Nsp10	224	L	L
10,333	T	C	Nsp10	244	L	L
10,520	A	G	Nsp10	307	M	V
10,627	T	C	Nsp10	342	V	V
10,847	A	C	Nsp10	416	I	L

Vị trí trong hệ gen	Lần cây truyền 10 trên các tế bào MARC-145 nucleotit	Cây truyền 17 trên cá tế bào PK-9 nucleotit	Protein virus bị ảnh hưởng	Vị trí axit amin	Cây truyền 10 (trên các tế bào MARC-145) axit amin	Cây truyền 17 (trên các tế bào PK-9) axit amin
10,867	C	T	Nsp10	422	F	F
10,936	C	T	Nsp11	4	S	S
11,512	T	C	Nsp11	196	S	S
12,949	A	G	ORF3	90	L	L
13,452	C	T	ORF4	76	P	P
13,575	G	A	ORF4	117	V	V
13,843	T	G	ORF5	24	C	W
13,860	G	A	ORF5	30	S	N
14,287	C	T	ORF5	172	N	N
14,481	A	G	ORF6	41	L	L
14,546	T	C	ORF6	63	V	A
14,578	A	G	ORF6	74	T	A
14,780	C	T	ORF6	141	T	I
14,932	T	C	ORF7	20	N	N

Bảng 9

Sự thay đổi axit amin tạo nên sự ức chế interferon giảm và làm yếu virus

Vị trí trong hệ gen (nt)	ORF và vị trí (aa)	Nsp và vị trí (aa)	Lần cây truyền 0 axit amin	Lần cây truyền 17 axit amin	Lần cây truyền 52 axit amin
735	ORF1a: 182	Nsp1b: 2	AMAD <u>V</u> YD	AMAN <u>V</u> YD	AMAN <u>V</u> YD
756	ORF1a: 189	Nsp1b: 9	ISHD <u>A</u> VM	IGHN <u>A</u> VM	IGHN <u>A</u> VM
992	ORF1a: 267	Nsp1b: 87	TVPE <u>G</u> NC	TVPD <u>G</u> NC	TVPD <u>G</u> NC
1009	ORF1a: 273	Nsp1b: 93	CWWC <u>L</u> FD	CWWY <u>L</u> FD	CWWY <u>L</u> FD
1096	ORF1a: 302	Nsp1b: 122	HGV <u>P</u> GKY	HGVH <u>G</u> KY	HGVH <u>G</u> KY
2106	ORF1a: 639	Nsp2: 256	AAKID <u>Q</u> Y	AAKV <u>D</u> QY	AAKV <u>D</u> QY
2185	ORF1a: 665	Nsp2: 282	PSAM <u>D</u> T	PSAT <u>D</u> T	PSAT <u>D</u> T

2403	ORF1a: 738	Nsp2: 355	LVS <u>V</u> LSK	LNSL <u>L</u> SK	LNSL <u>L</u> SK
3019	ORF1a: 943	Nsp2: 560	APMY <u>Q</u> DE	APMC <u>Q</u> DE	APMC <u>Q</u> DE
4477	ORF1a: 1429	Nsp3: 72	CAP <u>K</u> GMD	CAPT <u>G</u> MD	CAPT <u>G</u> MD
4705	ORF1a: 1505	Nsp3: 148	PKV <u>V</u> KVS	PKV <u>A</u> KVS	PKV <u>A</u> KVS
5985	ORF1a: 1932	Nsp4: 129	AGEL <u>V</u> GV	AGE <u>I</u> GV	AGE <u>I</u> GV
7419	ORF1a: 2410	Nsp7: 217	ADFD <u>P</u> EK	ADFN <u>P</u> EK	ADFN <u>P</u> EK
8831	ORF1a/1b: 2881	Nsp9: 429	QTP <u>V</u> LGR	QTP <u>I</u> LGR	QTP <u>I</u> LGR
13,857	ORF5: 29	-	AVL <u>V</u> NAN	AVL <u>G</u> NAN	AVL <u>G</u> NAN
14,578	ORF6: 74	-	VAL <u>T</u> MGA	VAL <u>A</u> MGA	VAL <u>A</u> MGA
14,780	ORF6: 141	-	PGST <u>T</u> VN	PGS <u>I</u> VN	PGS <u>I</u> VN

Bảng 10

Sự thay đổi axit amin tạo ra sự ức chế interferon giảm và làm yếu virut

Vị trí trong hệ gen (nt)	ORF và vị trí (aa)	Nsp và vị trí (aa)	Vị trí 0 axit amin	Lần cây truyền 17 axit amin	Lần cây truyền 52 axit amin
735	ORF1a: 182	Nsp1b: 2	AMAD <u>V</u> YD	AMAN <u>V</u> YD	AMAN <u>V</u> YD
756	ORF1a: 189	Nsp1b: 9	ISHD <u>A</u> VM	IGHN <u>A</u> VM	IGHN <u>A</u> VM
1009	ORF1a: 273	Nsp1b: 93	CWW <u>C</u> LFD	CWW <u>Y</u> LFD	CWW <u>Y</u> LFD
1096	ORF1a: 302	Nsp1b: 122	HGVP <u>G</u> KY	HGVH <u>G</u> KY	HGVH <u>G</u> KY
2185	ORF1a: 665	Nsp2: 282	PSAM <u>D</u> TS	PSAT <u>D</u> TS	PSAT <u>D</u> TS
3019	ORF1a: 943	Nsp2: 560	APMY <u>Q</u> DE	APMC <u>Q</u> DE	APMC <u>Q</u> DE
4477	ORF1a: 1429	Nsp3: 72	CAP <u>K</u> GMD	CAPT <u>G</u> MD	CAPT <u>G</u> MD
4705	ORF1a: 1505	Nsp3: 148	PKV <u>V</u> KVS	PKV <u>A</u> KVS	PKV <u>A</u> KVS
7419	ORF1a: 2410	Nsp7: 217	ADFD <u>P</u> EK	ADFN <u>P</u> EK	ADFN <u>P</u> EK

YÊU CẦU BẢO HỘ

1. Phân tử polynucleotit đã phân lập bao gồm trình tự ADN mã hóa phân tử ARN gây nhiễm mã hóa virus gây bệnh rối loạn hô hấp và sinh sản trên lợn (PRRS - Porcine Reproductive and Respiratory Syndrome) Bắc Mỹ, trong đó trình tự ADN nêu trên được chọn từ nhóm bao gồm SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4 và SEQ ID NO:6.
2. Vật chủ được gây nhiễm bằng phân tử polynucleotit như được xác định trong điểm 1.
3. Vacxin bảo vệ lợn chống lại việc nhiễm virus PRRS, trong đó vacxin này bao gồm (a) virus PRRS Bắc Mỹ được mã hóa bởi phân tử polynucleotit như được xác định trong điểm 1, (b) phân tử ARN gây nhiễm như được xác định trong điểm 1, (c) phân tử polynucleotit như được xác định trong điểm 1 ở dạng plasmit, hoặc (d) vector virus bao gồm phân tử polynucleotit như được xác định trong điểm 1, và chất mang thích hợp để sử dụng trong thú y.
4. Phân tử polynucleotit đã phân lập bao gồm trình tự nucleotit mã hóa virus PRRS, trong đó trình tự nucleotit nêu trên là trình tự ARN có trình tự ADN là SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4 hoặc SEQ ID NO:6.
5. Virus PRRS Bắc Mỹ đã phân lập được mã hóa bởi phân tử ARN gây nhiễm được xác định trong điểm 1.
6. Vacxin theo điểm 3, trong đó vacxin này có hiệu quả ức chế interferon- α thấp hơn so với virus PRRS P129 kiểu dại.
7. Phân tử polynucleotit đã phân lập theo điểm 1 hoặc điểm 4 mà bao gồm trình tự polynucleotit SEQ ID NO: 6.
8. Phân tử polynucleotit ARN bổ trợ trên toàn bộ chiều dài của trình tự polynucleotit ADN bất kỳ theo điểm 1 hoặc 7.
9. Phương pháp *in-vitro* tạo ra virus PRRS, trong đó phương pháp này bao gồm việc chuyển nhiễm tế bào vật chủ thích hợp bằng plasmit bao gồm trình tự polynucleotit theo điểm 1, và thu được virus PRRS được tạo ra bởi tế bào đã gây nhiễm.

10. Phương pháp *in-vitro* tạo ra virus PRRS, trong đó phương pháp này bao gồm việc phiên mã từ plasmit của phân tử polynucleotit có trình tự theo điểm 1.
11. Plasmit bao gồm phân tử polynucleotit đã phân lập theo điểm 1 hoặc 7 và đoạn khởi đầu có khả năng phiên mã phân tử polynucleotit trong tế bào chủ thích hợp.

Fig. 1

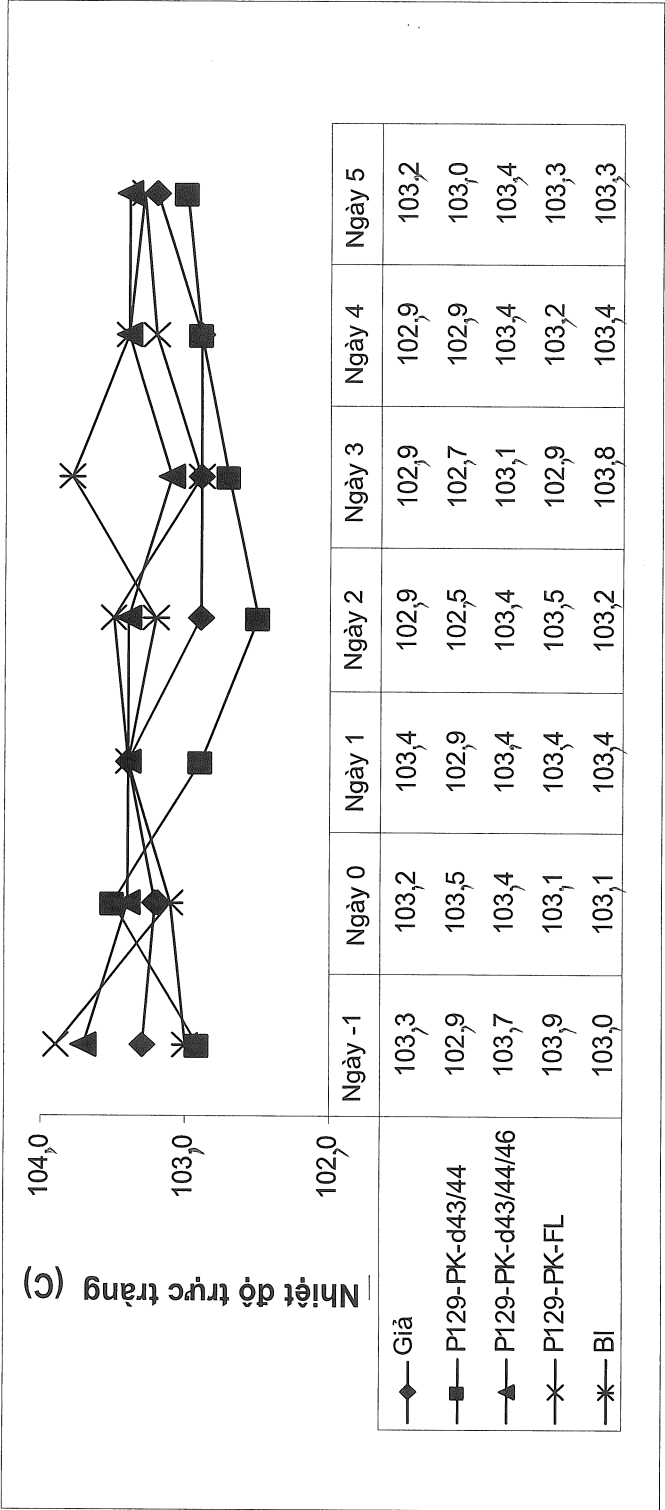


Fig. 2

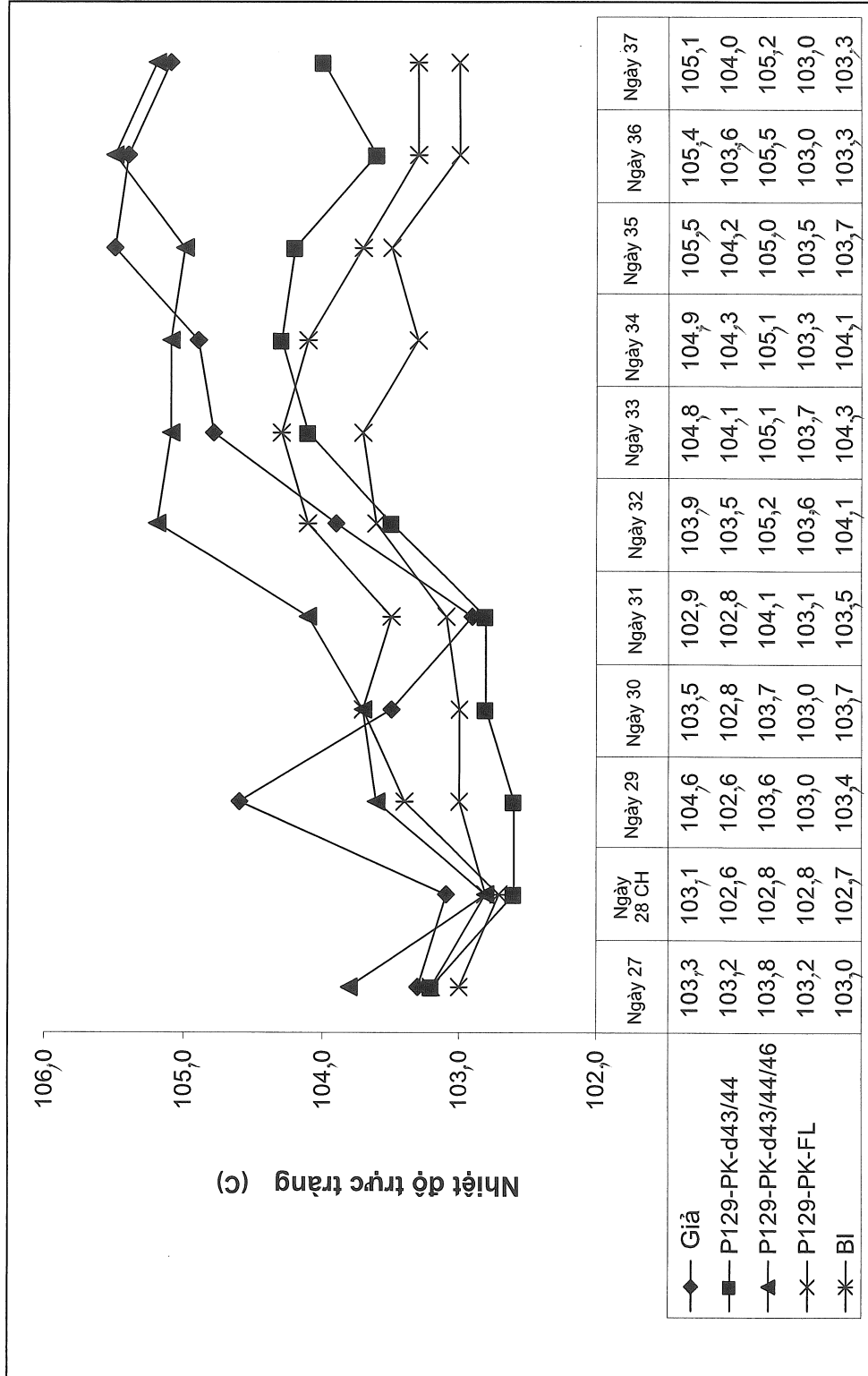


Fig. 3

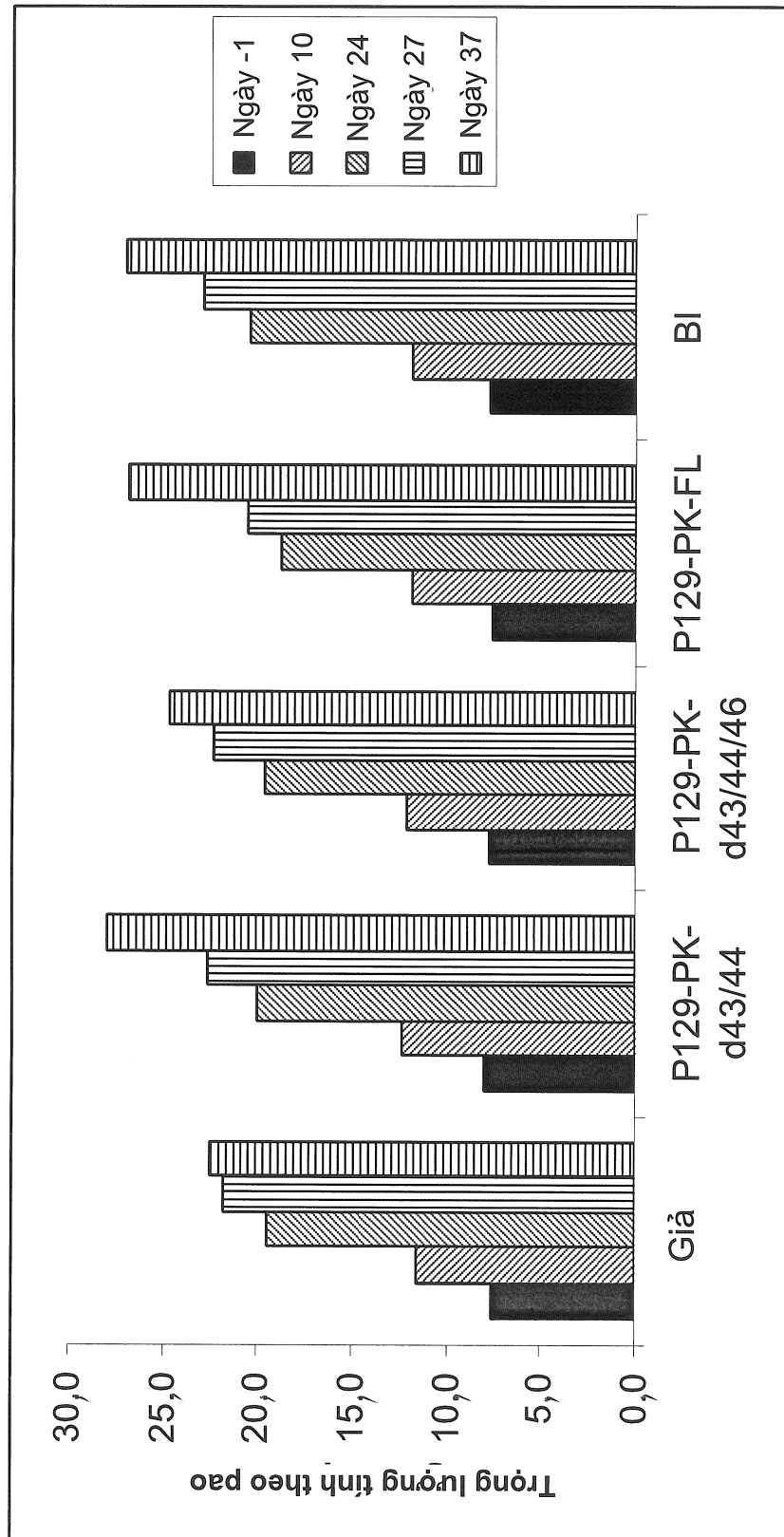


Fig. 4

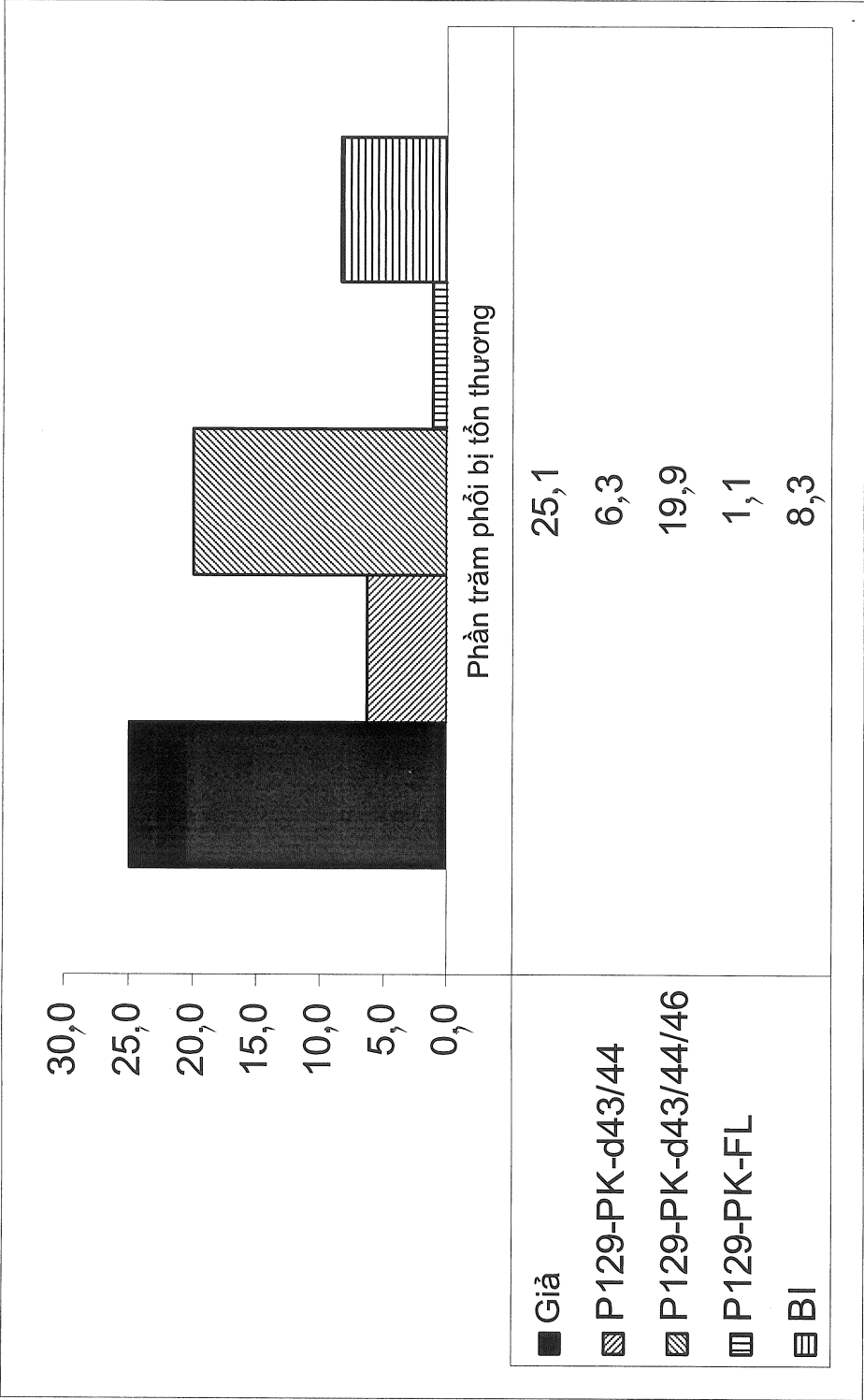


Fig. 5

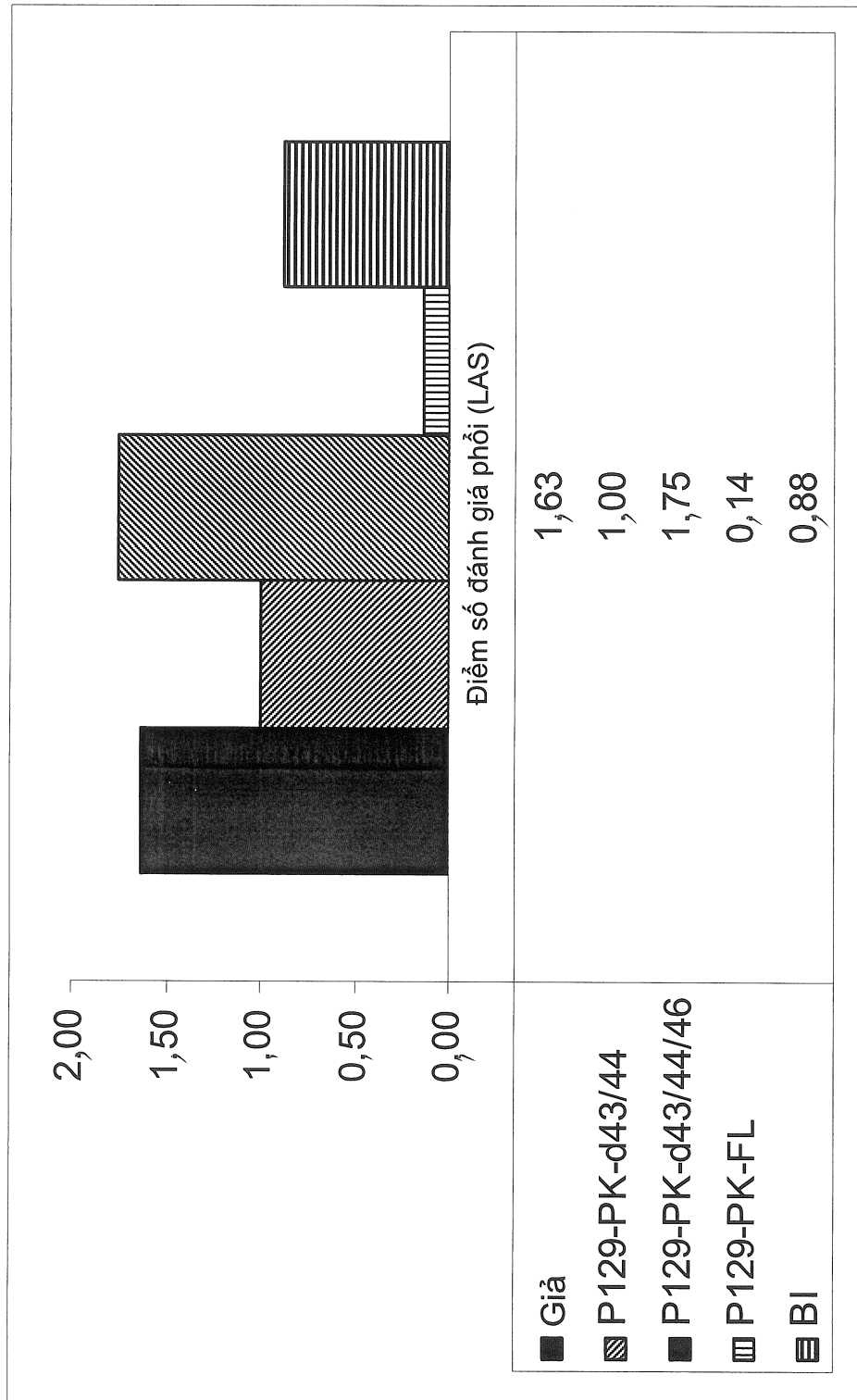


Fig. 6

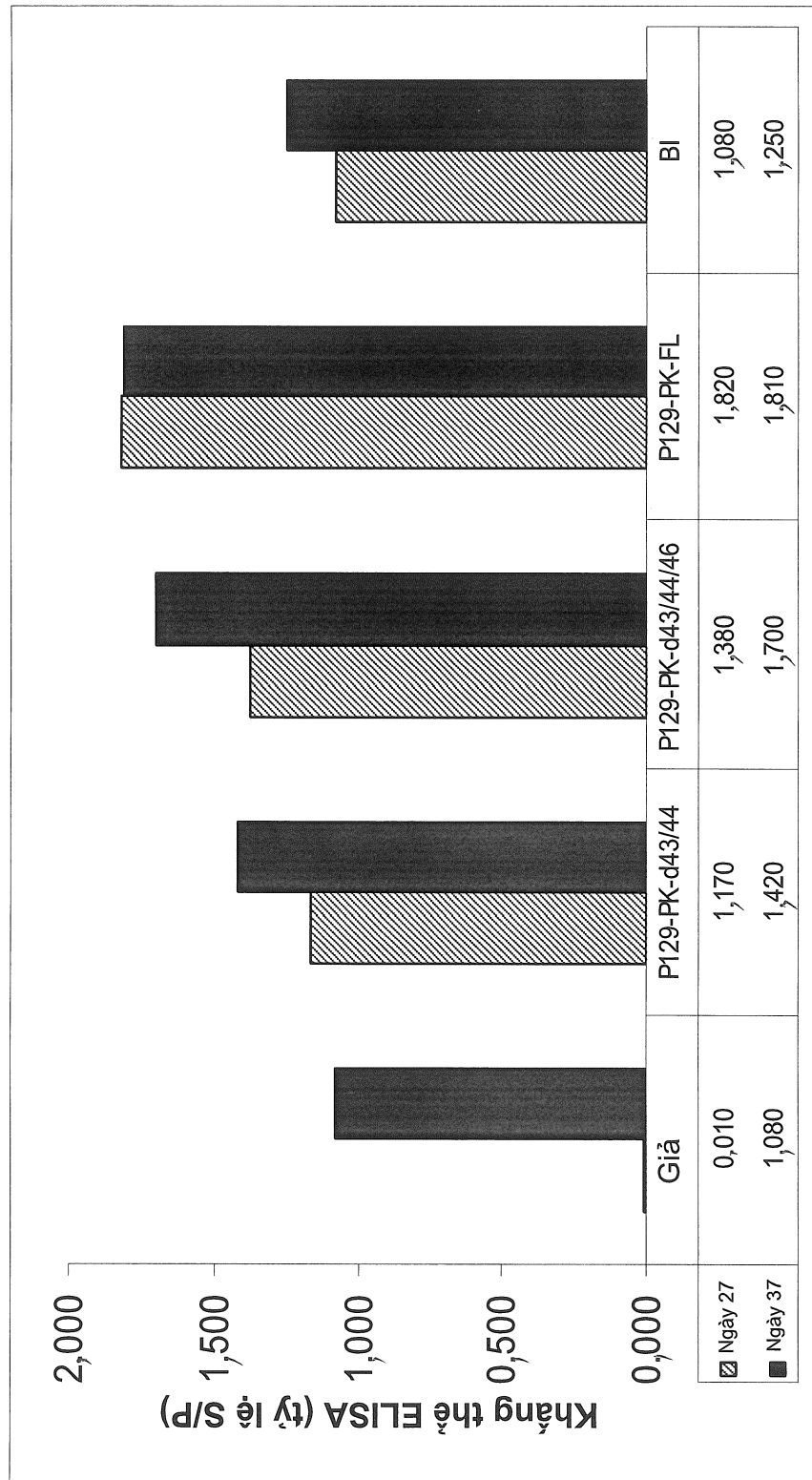


Fig. 7

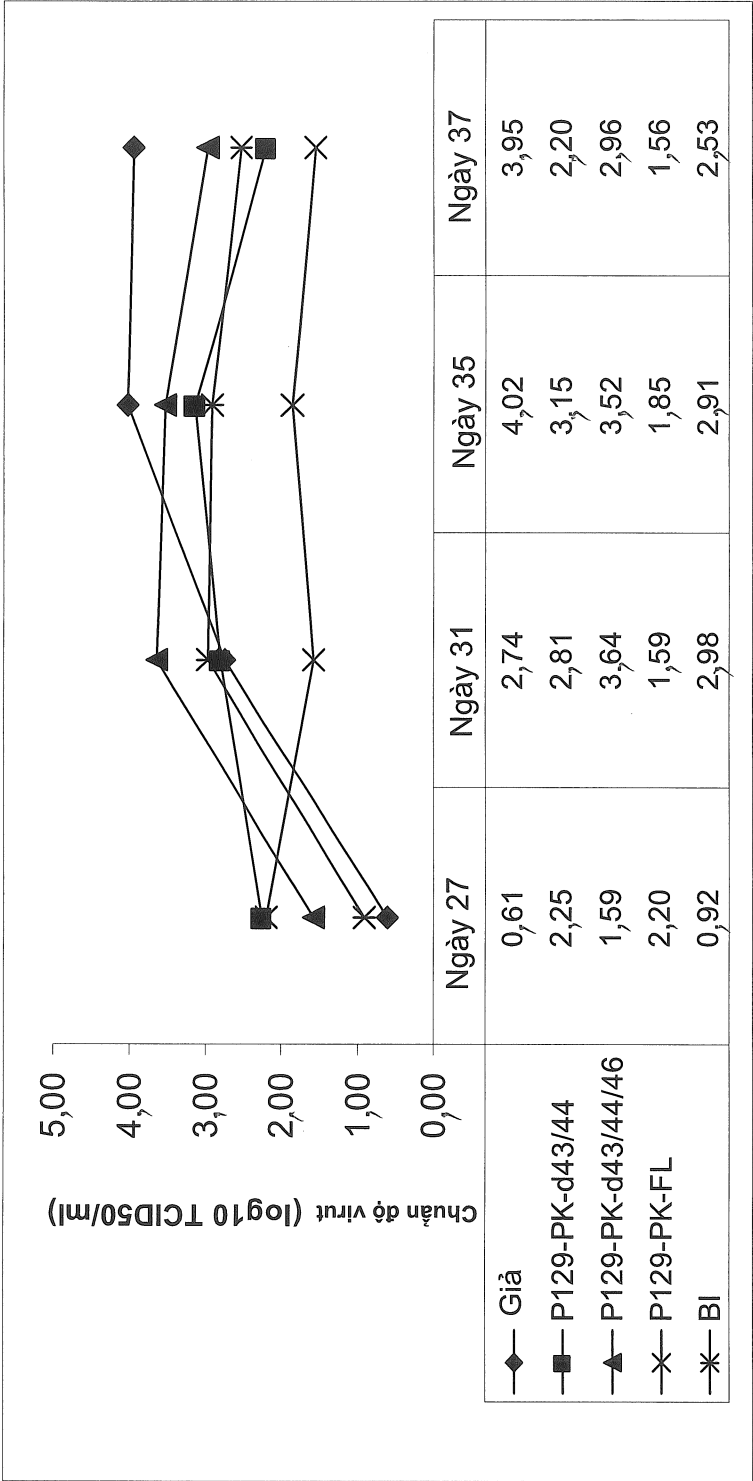
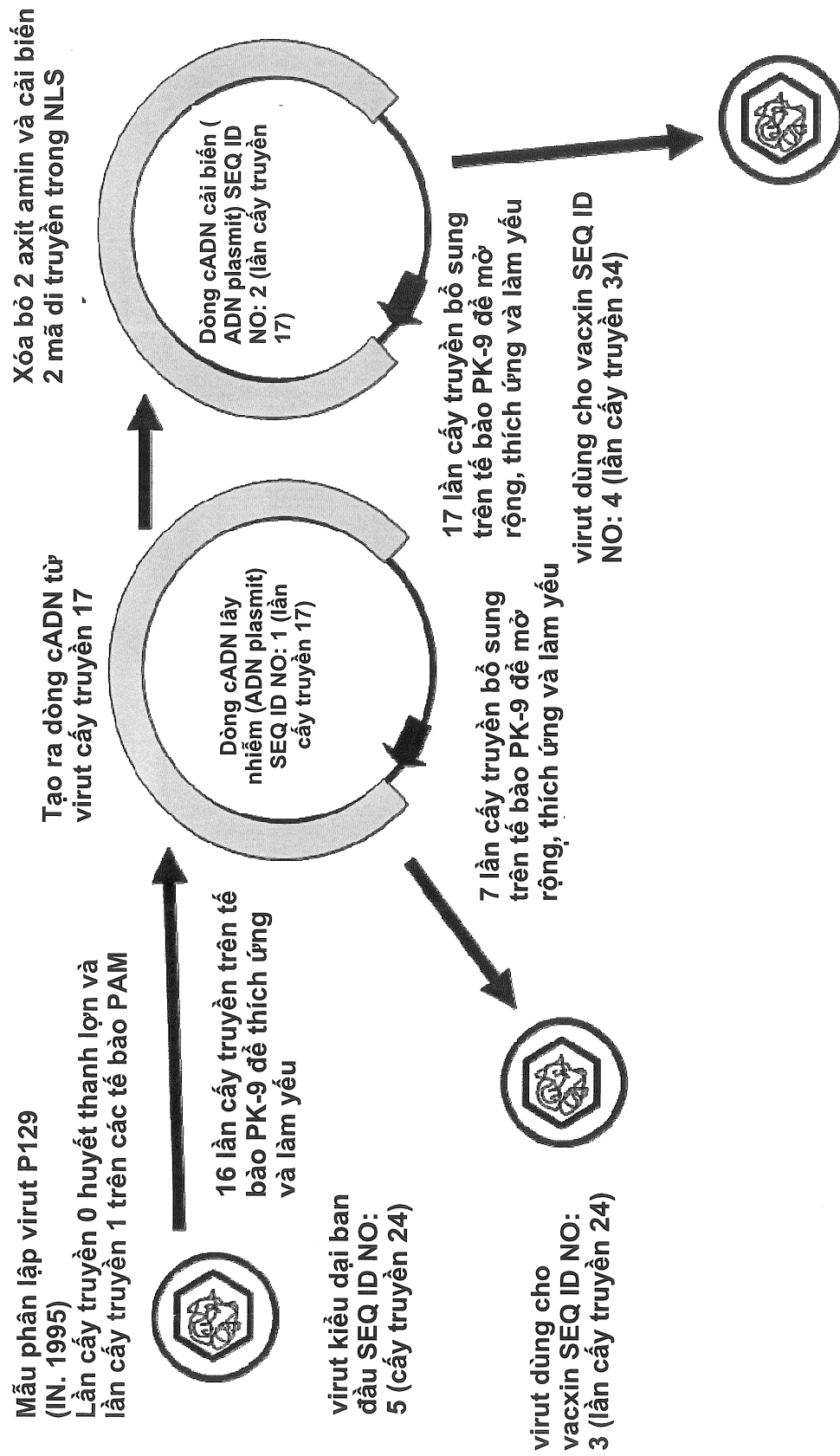


Fig. 8



DANH MỤC TRÌNH TỰ

- <110> Calvert, Jay G.
Slade, David E.
- <120> Phân tử polynucleotit đã phân lập bao gồm trình tự adn mã hóa phân tử arn gây nhiễm mã hóa virus gây bệnh rối loạn hô hấp và sinh sản trên lợn (PRRS) Bắc Mỹ, virus PRRS, vaccin, và phương pháp tạo ra virus này
- <130> PC33979A
- <150> 61/412,006
- <151> 2010-11-10
- <160> 57
- <170> Phiên bản PatentIn 3.5
- <210> 1
- <211> 15450
- <212> ADN
- <213> Trình tự nhân tạo
- <220>
- <223> Dòng lây nhiễm với hệ gen hoàn chỉnh P129-PK-FL lần cấy truyền 17 (15,450 nt) có chiều dài đầy đủ
- <400> 1
- | | |
|--|------|
| atgacgtata ggtgttggt ctatgccacg gcatttgtat tgtcaggagc tgtgaccatt | 60 |
| ggcacagccc aaaacttgct gcacggaaaa cgcccttctg tgacagcctt cttcagggga | 120 |
| gcttaggggt ctgtccctag caccttgctt ctggagttgc actgctttac ggtctctcca | 180 |
| cccctttaac catgtctggg atacttgatc ggtgcacgtg ccccccaat gccagggtgt | 240 |
| ttatggcgga gggccaagtc tactgcacac gatgtctcag tgcacggtct ctccttcctc | 300 |
| tgaatctcca agttcctgag cttgggggtgc tgggcctatt ttataggccc gaagagccac | 360 |
| tccggtggac gttgccacgt gcattcccca ctgtcgagtg ctcccctgcc ggggcctgct | 420 |
| ggctttctgc gatctttcca attgcacgaa tgaccagtgg aaacctgaac tttcaacaaa | 480 |
| gaatggtgcg ggttgacgct gagatttaca gagccggcca actcaccctt gcagttttga | 540 |
| aggctctaca agtttatgaa cgggggtgtc gctgggtaccc cattgtcgga cctgtccctg | 600 |
| gagtggccgt ttacgccaac tccctacatg tgagtgacaa acctttcccg ggagcaactc | 660 |
| atgtgttaac caacctaccg ctcccgaga gggccaagcc tgaagacttt tgcccttttg | 720 |
| agtgtgctat ggctaacgtc tatgacattg gccataacgc cgtcatgtat gtggccagag | 780 |
| ggaaagtctc ctgggcccct cgtggcgggg atgaagtga atttgaaacc gtccccgaag | 840 |
| agttgaagtt gattgcgaac cgactccaca tctccttccc gccccaccac gcagtgga | 900 |
| tgtctgagtt tgccttcata gcccctggga gtgggtgtct cttgcgggtc gagcaccaac | 960 |
| acggctgcct tcccgctgat actgtccctg atgggaactg ctggtggtac ttgtttgact | 1020 |
| tgctcccacc ggaagttcag aataaagaaa ttcgccgtgc taaccaattt ggctatcaaa | 1080 |

ccaagcatgg tgtccatggc aagtacctac agcggaggct gcaagttaat ggtctccgag	1140
cagtgactga tacagatgga cctattgtcg tacagtactt ctctgttagg gagagttgga	1200
tccgccactt cagattggcg gaagaaccta gcctccctgg gtttgaagac ctctcagaa	1260
taagggtaga gcctaatacg tcgccattgg gtggcaaggg tgaaaaaato ttccggtttg	1320
gcagtcacaa gtggtacggg gctggaaaga gagcaaggag agcacgctct ggtgcgactg	1380
ccacggtcgc tcaccgcgct ttgcccgtc gcgaagccca gcaggccaag aagctcgagg	1440
ttgccagcgc caacagggct gagcatctca agtactattc cccgcctgcc gacgggaact	1500
gtggttgga ctgcatttcc gccattacca accggatggg gaattccaaa ttgaaacca	1560
ctcttcccga gagagtgaga cttcagatg actgggctac tgacgaggat cttgtgaata	1620
ccatccaaat cctcaggctc cccgcggcct tggacaggaa cggtgcttgt gctggcgcca	1680
agtacgtgct caagctggaa ggtgagcact ggaccgtctc tgtgaccctt gggatgacct	1740
cttctttgct ccccttgaa tgtgttcagg gttgttgtga gcataagagc ggtcttggtt	1800
tcccagacgt ggtcgaagtt tccgatttg accctgcctg tcttgaccga cttgctgaga	1860
taatgcactt gcctagcagt gtcacccag ctgctctggc cgagatgtcc gacgacttca	1920
atcgtctggc ttccccggcc gccactgtgt ggactgtttc gcaattcttt gcccgccaca	1980
gaggaggaga gcacctgac caggtgtgct tagggaaaat tatcaacctt tgtcaggtga	2040
ttgaggaatg ctgctgttcc cggaacaaag ccaaccgggc taccgccgaa gaggttgcg	2100
caaaagttga ccagtacctc cgtggtgcag caagccttgg agaatgcttg gccaaagctt	2160
agagggctcg cccgccgagc gcgacggaca cctcctttga ttggaatgtt gtgcttctg	2220
gggttgagac ggcgaatcag acaaccaaac agtccatgt caaccagtgc cgcgctctgg	2280
ttctgtcgt gactcaagag ctttggaca gagactcggc ccctctgacc gccttctcgc	2340
tgtccaattg ctactacct gcacaagggt acgaggtccg tcaccgtgag aggctaaact	2400
ccttgcctc taagttggag ggggtgttc gtgaggaata tgggctcac ccaactggac	2460
ctggcccgcg acccgactg ccgaacgggc tcgacgagct taaagaccag atggaggagg	2520
atctgtgaa attagtcaac gcccaggcaa cttcagaaat gatggcctgg gcagccgagc	2580
aggttgatct aaaagcttgg gtcaaaaatt acccaggtg gacaccgcca cccctccac	2640
caagagttca gcctcgaaaa acgaagtctg tcaagagctt gctagagAAC aagcctgtcc	2700
ctgctccgcg caggaaggtc agatctgatt gtggcagccc gattttgatg ggcgacaatg	2760
ttcctaacgg ttgggaagat tcgactgttg gtgggtccct tgatctttcg gcaccatccg	2820
agccgatgac acctctgagt gagcctgtac ttatttccag gccagtgaca tctttgagt	2880
tgcggcccc agttcctgca ccgcgtagag ctgtgtcccg accgatgacg cctcgagt	2940
agccaatttt tgtgtctgca ctgcgacaca aatttcagca ggtggaaaaa gcaaatctgg	3000

cggcagcagc gccgatgtgc caggacgaac ccttagattt gtctgcatcc tcacagactg 3060
 aatatgagge ttcccccta acaccaccgc agaacgtggg cattctggag gtaagggggc 3120
 aagaagctga ggaagtcttg agtgaaatct cggatattct gaatgatacc aaccctgcac 3180
 ctgtgtcatc aagcagctcc ctgtcaagtg ttaagatcac acgccccaaa tactcagctc 3240
 aagccattat cgactcgggc gggccctgca gtgggcacct ccaaaggga aaagaagcat 3300
 gcctccgcat catgcgtgag gcttgtgatg cggccaagct tagtgaccct gccacgcagg 3360
 aatggctttc tcgcatgtgg gatagggtgg acatgctgac ttggcgcaac acgtctgctt 3420
 accaggcggt tcgcacctta gatggcaggt ttgggtttct ccaaagatg atactcgaga 3480
 cgccgccgcc ctaccctgtg gggtttgtga tgttgccctca caccctgca ccttccgtga 3540
 gtgcagagag cgaccttacc attggttcag tcgccactga agatattcca cgcatacctcg 3600
 ggaaaataga aaataccggt gagatgatca accagggacc cttggcatcc totgaggaag 3660
 aaccggtata caaccaacct gccaaagact cccggatata gtcgcggggg tctgacgaga 3720
 gcacagcagc tccgtccgcg ggtacaggtg gcgccggctt atttactgat ttgccacctt 3780
 cagacggcgt agatgcggac ggtggggggc cgttgacagc ggtaagaaag aaagctgaaa 3840
 ggctcttcga ccaattgagc cgtcagggtt ttaacctcgt ctccatctc cctgttttct 3900
 tctcacacct cttcaaactc gacagtgggt attctccggg tgattggggg tttgcagctt 3960
 ttactctatt ttgcctcttt ttgtgttaca gctaccatt cttcggtttc gttccccctc 4020
 tgggtgtatt ttctgggtct tctcggcgtg tgcgcatggg ggtttttggc tgctggctgg 4080
 cttttgctgt tggcctgttc aagcctgtgt ccgaccagc cggcactgct tgtgagtttg 4140
 actcgccaga gtgcaggaac gtccttcatt cttttgagct tctcaaacct tgggaccctg 4200
 ttcgcagcct tgttgtgggc cccgtcggtc tcgggtcttg cattcttggc aggttactgg 4260
 gcggggcacg ctacatctgg cattttttgc ttaggcttgg cattgttgca gattgtatct 4320
 tggctggagc ttatgtgctt tctcaaggta ggtgtaaaaa gtgctgggga tcttgataa 4380
 gaactgctcc taatgaaatc gccttcaacg tgttcccttt tacacgtgcg accaggctcg 4440
 cactcatcga cctgtgcgat cggttttgtg cgccaacagg catggacccc attttcctcg 4500
 ccaactgggtg gcgtgggtgc tggaccggcc gaagtcccat tgagcaaccc tctgaaaaac 4560
 ccatcgctt cgcccagttg gatgaaaaga ggattacggc tagaactgtg gtcgctcagc 4620
 cttatgatcc taatcaagcc gtgaagtgtg tgcgggtgtt acaggcgggt ggggcgatgg 4680
 tggccgagge agtccccaaa gtggccaaag tttctgctat tccattccga gcccttttt 4740
 ttcccaccgg agtgaaagt gatcccgagt gcaggatcgt gggtgacccc gatactttta 4800
 ctacagccct ccggtctggt tactctacca caaacctcgt ccttggtgtg ggggactttg 4860
 cccagctgaa tggactaaag atcaggcaaa tttccaagcc ttcgggagga ggccacacc 4920

tcattgctgc cctgcatgtt gcoctgctoga tggcgttgca catgcttgcg ggggtttatg	4980
taacttcagt ggggtcttgc ggtgccggca ccaacgatcc atggtgcact aatccgtttg	5040
cogttcctgg ctacggacca ggctctctct gcacgtccag attgtgcac tcccaacatg	5100
gccttaccct gcccttgaca gcacttgtgg cgggattcgg tottcaggaa atcgcttgg	5160
tcgttttgat tttcgtttcc atcggaggca tggctcatag gttgagttgt aaggctgata	5220
tgctgtgcat tttacttgca atcgccagct atgtttgggt accccttacc tggttgcttt	5280
gtgtgtttcc ttgttggttg cgctggttct ctttgcaccc ccttaccatc ctatggttgg	5340
tgtttttctt gatttctgta aatatgcctt cggaatctt ggccgtggtg ttattggttt	5400
ctctttggct tttgggacgt tatactaaca ttgctggtct tgtcaccccc tatgatattc	5460
atcattacac cagtggcccc cgcggtgttg ccgccttggc taccgcacca gatggaacct	5520
acttggtgc cgtccgccgc gctgcgttga ctggtgcac catgctgttc acccgtctc	5580
agcttgggtc cttcttgag ggcgctttca gaactcgaag gccctcactg aacaccgtca	5640
atgtggttg gtcctccatg ggctctggtg gagtgttcac catcgacggg aaaattaggt	5700
gcgtgactgc cgcacatgtc cttacgggta attcggctag ggtttccgga gtcggcttca	5760
atcaaagtct tgactttgat gtgaaagggg acttcgccat agctgattgc ccgaattggc	5820
aaggagctgc tcccaagacc caattctgcg aggacggatg gactggccgt gcctattggc	5880
tgacatctc tggcgctgaa cccggtgtta ttgggaatgg attcgccttc tgcttcaccg	5940
cgtgcggcga ttccgggtcc ccagtgatca ccgaagctgg tgagattgtc ggcgttcaca	6000
caggatcaaa taaacaagga ggtggcatcg tcacgcgccc ttcaggccag ttttgtaacg	6060
tggcacccat caagctgagc gaattaagtg aattctttgc tggacccaag gtcccgtctg	6120
gtgatgtgaa ggttggcagc cacataatta aagacacgtg cgaagtacct tcagatcttt	6180
gcgccttgct tgctgcaaaa cctgaactgg agggaggcct ctccaccgtc caacttctgt	6240
gtgtgttttt cctactgtgg agaattgatg gacatgcctg gacgcccttg gttgctgtgg	6300
ggtttttcat tctgaatgag gttctcccag ctgtcctggt tcggagtgtt ttctcctttg	6360
ggatgtttgt gctatcttgg ctacacccat ggtctgcgca agttctgatg atcaggcttc	6420
taacagcagc tcttaacagg aacagatggt cacttgccct ttacagcctt ggtgcggtga	6480
ccggttttgt cgcagatctt gcggttaactc aagggcaccg gttgcaggca gtaatgaatt	6540
tgagcaccta tgccttctct cctcggatga tggttgtgac ctaccagtc ccagtgattg	6600
cgtgtggtgt tgtgcaccta cttgccatca ttttgactt gttcaagtac cgcggcctgc	6660
acaatgttct tgttggtgat ggagcgtttt ctgcagcttt cttcttgcca tactttgccg	6720
agggaaagt ggaggaaggg gtgtcgcaat cctgcggaat gaatcatgag tcattgactg	6780
gtgccctcgc tatgagactc aatgacgagg acttggaact ccttacgaaa tggactgatt	6840

ttaagtgctt tgtttctgcg tccaacatga ggaatgcagc aggccaattc atcgaggctg	6900
cctatgcaaa agcacttaga attgaacttg ccagttggg gcaggttgat aaggttcgag	6960
gtactttggc caagcttgag gcttttgctg ataccgtggc accccaactc tcgcccgggtg	7020
acattgttgt tgctcttggc catacgctg ttggcagcat cttcgacctt aaggttggtg	7080
gtaccaagca tactctccaa gtcattgaga ccagagtcct tgccgggtcc aaaatgaccg	7140
tggcgcgcggt cgttgacca acccccacgc cccacccgc acccgtgcc atccccctcc	7200
caccgaaagt tctagagaat ggtcccaacg cctgggggga tggggaccgt ttgaataaga	7260
agaagaggcg taggatggaa accgtcggca tctttgtcat ggggtgggaag aagtaccaga	7320
aattttggga caagaattcc ggtgatgtgt ttacgagga ggtccatgac aacacagatg	7380
cgtgggagtg cctcagagtt ggtgaccctg ccgactttaa ccctgagaag ggaactctgt	7440
gtgggcatac tactattgaa gataaggatt acaaagtcta cgcctcccca tctggcaaga	7500
agttcctggg ccccgctaac ccagagagcg gaagagccca atgggaagct gcaaagcttt	7560
ccgtggagca ggcccttggc atgatgaatg tcgacggtga actgacggcc aaagaagtgg	7620
agaaactgaa aagaataatt gacaaacttc agggcctgac taaggagcag tgtttaaact	7680
gctagccgcc agcggcttga cccgctgtgg tcgcgggcggc ttggttggtt ctgagacagc	7740
ggtaaaaaata gtcaaatttc acaaccggac ttccacctt gggcctgtga atttaaaagt	7800
ggccagttag gttgagctga aagacgcggt cgagcacaac caacacccgg ttgcaagacc	7860
ggttgacggt ggtgttgtgc tctgcgttc cgagttcct tcgcttatag atgtcctgat	7920
ctccggtgct gacgcatctc ctaagttact cgctcgtcac gggccgggga aactgggat	7980
cgatggcacg ctttgggact ttgaggccga ggccaccaa gaggaaattg cgctcagtgc	8040
gcaaataata caggcttgtg acattaggcg cggtagcgca cctgaaattg gtctccctta	8100
caagctgtac cctgttaggg gcaaccctga gcgggtaaaa ggagttttac agaatacaag	8160
gtttggagac ataccttaca aaacccccag tgacactggg agcccagtgc acgcggtgc	8220
ctgcctcacg cccaatgcca ctccggtgac tgatgggcgc tccgtcttgg ctactaccat	8280
gccctccggt tttgaattgt atgtaccgac cattccagcg tctgtccttg attatcttga	8340
ctctaggcct gactgcccc aacagttgac agagcacggc tgtgaggatg ccgcattgag	8400
agacctctcc aagtatgact tgtccacca aggctttgtt ttacctgggg ttcttcgcct	8460
tgtgcgtaag tacctgtttg ccacgtggg taagtgccg cccgttcac ggccctccac	8520
ttacctgcc aagaattcta tggctggaat aaatgggaac aggtttccaa ccaaggacat	8580
tcagagcgtc cccgaaatcg acgttctgtg cgcacaggcc gtgcgagaaa actggcaaac	8640
tgttaccctt tgtaccctca agaaacagta ttgtgggaag aagaagacta ggacaatact	8700
cggcaccaat aatttcattg cgttggccca ccgggcagcg ttgagtgggt tcacccaggg	8760

cttcatgaaa aaggcgttta actcgcccat cgccctcggg aaaaacaaat ttaaggagct 8820
 acagactccg atcttaggca ggtgccttga agctgatctt gcatcctgtg atcgatccac 8880
 acctgcaatt gtccgctggt ttgcgcgcaa ccttctttat gaacttgcct gtgctgaaga 8940
 gcacctaccg tcgtacgtgc tgaactgctg ccatgacctt ttggtcacgc agtcgggcgc 9000
 agtgactaag aggggtggcc tgtcgtctgg cgacccgatc acttctgtgt ctaacaccat 9060
 ttacagcttg gtgatatatg cacagcacat ggtgcttagt tactttaaaa gtggtcaccc 9120
 tcatggcctt ctgttcctac aagaccagct gaagtctgag gacatgctca aagtccaacc 9180
 cctgatcgtc tattcggacg acctcgtgct gtatgccgaa tctccacca tgccgaacta 9240
 ccactgggtg gtcgaacatc tgaatttgat gctgggtttt cagacggacc caaagaagac 9300
 agccataacg gactcgccat catttctagg ctgtaggata ataaatggac gccagctagt 9360
 cccaaccgt gacaggatcc tcgcggccct cgcttaccat atgaaggcaa gcaatgtttc 9420
 tgaatactac gccgcggcgg ctgcaatact catggacagc tgtgcttggt tagagtatga 9480
 tcctgaatgg tttgaagagc ttgtggttgg gatagcgagc tgcgccgca aggacggcta 9540
 cagctttccc ggcccgcgt tcttcttgtc catgtgggaa aaactcagat ccaatcatga 9600
 ggggaagaag tccagaatgt gcgggtattg cggggccccg gtcctgtacg ccactgcctg 9660
 tggcctcgac gtctgtattt accacacca ctccaccag cattgtccag tcacaatctg 9720
 gtgtggccac ccggtggtt ctggttcttg tagtgagtgc aaaccccccc tagggaaagg 9780
 cacaagccct ctagatgagg tgtagaaca agtcccgtat aagcctccac ggactgtaat 9840
 catgcatgtg gagcagggtc tcacccctct tgaccaggc agataccaga ctgcgcgcg 9900
 attagtctcc gttaggcgtg gcatcagagg aaatgaagtt gacctaccag acggtgatta 9960
 tgctagcacc gccctactcc ccacttgtaa agagatcaac atggtcgctg tcgcctctaa 10020
 tgtgttgccg agcaggttca tcatcgggtc gcccggtgct gggaaaacat actggctcct 10080
 tcagcaggtc caggatggtg atgtcattta cacaccgact caccagacca tgctcgacat 10140
 gattagggct ttggggacgt gccgggtcaa cgtcccagca ggtacaacgc tgcaattccc 10200
 tgccccctcc cgtaccggcc cgtgggttcg catcctagcc ggcgggttgg gtccctgtaa 10260
 gaattccttc ttggatgaag cagcgtattg taatcacctt gatgtcttga ggctccttag 10320
 caaaaccacc ctcacctgtc tgggagactt caaacaactc caccagtggt gttttgatcc 10380
 tcattgctat gtttttgaca tcatgcctca gaccagttg aagaccatct ggagattcgg 10440
 acagaacatc tgtgatgcca tccaaccaga ttacagggac aaacttgtgt ccatgggtaa 10500
 cacaaccgt gtaacctacg tggaaaaacc tgtcaagtat gggcaagtcc tcacccctta 10560
 ccacagggac cgagaggacg gcgccatcac aattgactcc agtcaaggcg ccacatttga 10620
 tgtggtcaca ctgcatttgc ccactaaaga ttactcaac aggcaaagag cccttggtgc 10680

tatcaccagg gcaagacatg ctatctttgt gtatgacca cacaggcaat tgcagagcat	10740
gtttgatctt cctgcgaagg gcacaccogt caacctcgca gtgcaccgtg atgagcagct	10800
gatcgtagtg gatagaaata ataaagaatg cacagttgct caggctctag gcaacggaga	10860
taaatttagg gccaccgaca agcgcgttgt agattctctc cgcgccatth gtgctgatct	10920
ggaagggctg agctctccgc tccccagggt cgcacacaac ttgggattth atttctcacc	10980
tgatttgaca cagtttgcta aactcccgtt agaccttgca cccactggc ccgtgggtgac	11040
aaccagaac aatgaaaagt ggccggatcg gctgggtgac agccttcgcc ctgtccataa	11100
gtatagccgt gcgtgcattg gtgccggcta tatgggtggc ccctcgggtg ttctaggcac	11160
ccctgggggtc gtgtcatact acctcacaaa atttgtcaag ggcgaggctc aagtgttcc	11220
ggagacagtc ttcagcaccg gccgaattga ggtggattgc cgggagtatc ttgatgacag	11280
ggagcgagaa gttgctgagt ccctcccaca tgccttcatt ggcgacgtca aaggcaccac	11340
cggtggggga tgtcatcatg tcacctcaa ataccttcgc cgcttccttc ccaaggaatc	11400
agtcgcggta gtcggggtht cgagccccgc gaaagccgca aaagcagtgt gcacattgac	11460
ggatgtgtac ctcccagacc ttgaggccta cctccacca gagactcagt ccaagtgtg	11520
gaaagttatg ttggacttca aggaagttcg actgatggc tggaaagaca agacggccta	11580
tttccaactt gaaggccgtt atttcacctg gtatcagctt gcaagctacg cctcgtacat	11640
ccgtgttctt gtcaactcca cgggtgtatc ggacccctgc atgggccctg ccctttgcaa	11700
cagaagagtt gtcgggtcca ccattgggg agctgacctc gcagtcaccc cttatgatta	11760
cgggtgctaaa atcatcttgt ctacgcctta ccatggtgaa atgcctcctg gatacaagat	11820
tctggcgtgc gcggagttct cgctcgacga ccagtcgaag taaaaacaca cctggggtht	11880
tgaatcggat acagcgtatc tgtatgagtt caccggaaac ggtgaggact gggaggatta	11940
caatgatgcg tttcgtgcgc gccagaaagg gaaaatttat aaggccactg ctaccagcat	12000
gaagtthtat tttccccgcg gccccgtcat tgaaccaact ttaggcctga attgaaatga	12060
aatgggggtt atacaaagcc tcttcgacaa aattggccag cthttttgtgg atgctttcac	12120
ggaattthtt gtgtccattg ttgatatcat catattthtt gccattthtt ttggcttcac	12180
catcgccggt tggctgggtg tctthttgcat cagattgggt tgctccgcgg tattccgtgc	12240
gcgccctgcc attcaccctg agcaattaca gaagatccta tgaggcctth cthttctcagt	12300
gccgggtgga cattcccacc tggggggtaa aacaccctth ggggatgtht tggcaccata	12360
aggtgtcaac cctgattgat gaaatgggtg cgcgtcgaat gtaccgcata atggaaaaag	12420
cagggaagc tgcttgaaa cagggtgtga gcgaggctac gctgtctcgc attagtagtt	12480
tggatgtggt ggctcattht caacatcttg ccgccattga agccgagacc tgtaaatatt	12540
tggcttctcg actgcccatg ctacacaacc tgcgcatgac aggttcaaat gtaaccatag	12600

tgtataatag cacttttaaat caggtgtttg ctatTTTTcc aacccttggg tcccggccaa	12660
agcttcatga ttttcagcaa tggctaatag ctgtacattc ctccatattt tcctctgttg	12720
cagcttcttg tactcttttt gttgtgctgt gggtgcgggt tccaatgcta cgtactgttt	12780
ttggtttccg ctggttaggg gcaatttttc tttcgaactc atggtgaatt acacggtgtg	12840
tccaccttgc ctccccgac aagcagccgc tgaggctcct gaacccggtg ggtctctttg	12900
gtgcaggata gggcatgacc gatgtgggga ggacgatcac gacgaactgg ggttcatggt	12960
tccgcctggc ctctccagcg aaagccactt gaccagtgtt taagcctggg tggcgttcct	13020
gtccttcagc tacacggccc agttccatcc cgagatattt gggatagggg acgtgagtga	13080
agtttatgtt gacatcaagc accaattcat ctgcgccgtt catgacgggc agaacaccac	13140
cttgccctgc catgacaata tttcagccgt atttcagacc tactatcaac atcaggctga	13200
cggcggcaat tggtttcacc tagaatggct gcgtcccttc ttttctctt ggttggtttt	13260
aaatgtttcg tggtttctca ggcgttcgcc tgcaagccat gtttcagttc gagtctttca	13320
gacatcaaaa ccaacactac cgcagcatca ggctttgttg tcctccagga catcagctgc	13380
cttaggcatg gcgactcgtc ctttccgacg attcgcaaaa gctctcaatg ccgcacggcg	13440
atagggacac ctgtgtatat caccatcaca gccaatgtga cagatgagaa ttacttacat	13500
tcttctgac tcctcatgct ttcttcttgc ctttctatg cttctgagat gagtgaag	13560
ggattcaagg tggattttg caatgtgtca ggcatcgtg ctgtgtgtgt caactttacc	13620
agctacgtcc aacatgtcaa agagtttacc caacgctcct tggtggtcga tcatgtgcgg	13680
ctgcttcatt tcatgacacc tgagaccatg aggtgggcaa ccgttttagc ctgtctttt	13740
gccatcctac tggcaatttg aatgttcaag tatgttggg aaatgcttga ccgcgggctg	13800
ttgctcgcga ttgctttctt tgtggtgtat cgtgccgttc tggtttgctg tgctcggcaa	13860
cgccaacagc agcagcagct ctcatTTTca gttgatttat aacttgacgc tatgtgagct	13920
gaatggcaca gattggctgg cagaaaaatt tgattgggca gtggagactt ttgtcatctt	13980
tcccggtgtg actcacattg tttcctatgg tgcactcacc accagccatt tccttgacac	14040
agttggtctg gttactgtgt ccaccgccgg gttttatcac gggcggtatg tcttgagtag	14100
catctacgcg gtctgtgctc tggctgcgtt gatttgcttc gttattagc ttgcgaagaa	14160
ctgcatgtcc tggcgctact cttgtaccag atataccaac ttccttctgg acactaagg	14220
cagactctat cgttggcggt cgcccgttat catagaaaaa gggggtaagg ttgaggtcga	14280
aggtcatctg atcgacctca aaagagttgt gcttgatggt tccgtggcaa cccctttaac	14340
cagagtttca gcggaacaat ggggtcgtct ctagacgact tttgccatga tagcacggt	14400
ccacaaaagg tgcttttggc gttttccatt acctacacgc cagtaatgat atatgtctc	14460
aaggtaagtc gcggccgact gctagggctt ctgcacctt tgatctttct gaattgtgct	14520

```

tttaccttcg ggtacatgac attcgcgcac tttcagagca caaatagggg cgcgctcgt 14580
atgggagcag tagttgcact tctttggggg gtgtactcag ccatagaaac ctggaaattc 14640
atcacctcca gatgccgttt gtgcttgcta ggccgcaagt acattctggc cctgcccac 14700
cacgtcgaaa gtgccgctgg ctttcacccg attgcggaac atgataacca cgcatttgct 14760
gtccggcgct cccgctccat tacggttaac ggcacattgg tgcccggtt gaaaagcctc 14820
gtgttgggtg gcagaaaagc tgttaaacag ggagtggtaa accttgtaa atatgcaaaa 14880
taacaacggc aagcagcaaa agaaaaagaa ggggaatggc cagccagtca accagctgtg 14940
ccagatgctg ggtaaaatca tcgccagca aaaccagtcc agaggcaagg gaccgggcaa 15000
gaaaagtaag aagaaaaacc cggagaagcc ccattttcct ctagcgaccg aagatgacgt 15060
caggcatcac ttcacccctg gtgagcggca attgtgtctg tcgtcgatcc agactgcctt 15120
taaccagggc gctggaactt gtaccctgtc agattcaggg aggataagtt aactgtgga 15180
gtttagtttg ccgacgcac atactgtgcg cctgatccgc gtcacagcat caccctcagc 15240
atgatgggct ggcattcttt aggcacctca gtgtcagaat tggaagaatg tgtggtggat 15300
ggcactgatt gacattgtgc ctctaagtca cctattcaat tagggcgacc gtgtgggggt 15360
aaaatttaat tggcgagaac catgcgggcg caattaaaaa aaaaaaaaaa aaaaaaaaaa 15420
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa 15450

```

<210> 2

<211> 15444

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Dòng lây nhiễm xóa bỏ 2aa P129-PK-d43/44 lần cấy truyền 17 (15,444 nt) với hệ gen hoàn chỉnh

<400> 2

```

atgacgtata ggtgttggct ctatgccacg gcatttgtat tgtcaggagc tgtgaccatt 60
ggcacagccc aaaacttgct gcacggaaaa cgccttctg tgacagcctt cttcagggga 120
gcttaggggt ctgtccctag caccttgctt ctggagttgc actgctttac ggtctctcca 180
ccctttaac catgtctggg atacttgatc ggtgcacgtg ccccccaat gccaggggtg 240
ttatggcgga gggccaagtc tactgcacac gatgtctcag tgcaaggctt ctccctcctc 300
tgaatctcca agttcctgag cttgggggtg tgggcctatt ttataggccc gaagagccac 360
tccggtggac gttgccacgt gcattcccca ctgtcgagtg ctcccctgcc ggggcctgct 420
ggctttctgc gatctttcca attgcacgaa tgaccagtgg aaacctgaac tttcaacaaa 480
gaatggtgcg ggttgcagct gagatttaca gagccggcca actcaccctt gcagttttga 540
aggctctaca agtttatgaa cgggggtgtc gctggtaccc cattgtcgga cctgtccctg 600

```

gagtggccgt ttacgccaac tccctacatg tgagtgacaa acctttcccg ggagcaactc	660
atgtgttaac caacctaccg ctcccgcaga ggcccaagcc tgaagacttt tgcccttttg	720
agtgtgctat ggctaacgtc tatgacattg gccataacgc cgtcatgtat gtggccagag	780
ggaaagtctc ctgggcccct cgtggcgggg atgaagtga atttgaaacc gtccccgaag	840
agttgaagtt gattgcgaac cgactccaca tctccttccc gccccaccac gcagtggaca	900
tgtctgagtt tgccttcata gccctggga gtggtgtctc cttgcgggtc gagcaccaac	960
acggctgcct tcccgtgat actgtccctg atgggaactg ctggtggtac ttgtttgact	1020
tgtcccacc ggaagttcag aataaagaaa ttcgccgtgc taaccaattt ggctatcaaa	1080
ccaagcatgg tgtccatggc aagtacctac agcggaggct gcaagttaat ggtctccgag	1140
cagtgactga tacagatgga cctattgtcg tacagtactt ctctgttagg gagagttgga	1200
tccgccactt cagattggcg gaagaaccta gcctccctgg gtttgaagac ctctcagaa	1260
taagggtaga gcctaatacg tcgccattgg gtggcaaggg tgaaaaaatc ttccggtttg	1320
gcagtcacaa gtggtacggg gctggaaaga gagcaaggag agcacgctct ggtgcgactg	1380
ccacggtcgc tcaccgcgct ttgcccgctc gcgaagcca gcaggccaag aagctcgagg	1440
ttgccagcgc caacagggct gagcatctca agtactattc cccgcctgcc gacgggaact	1500
gtggttgcca ctgcatttcc gccattacca accggatggt gaattccaaa ttgaaacca	1560
ctcttcccga gagagtgaga ccttcagatg actgggctac tgacgaggat cttgtgaata	1620
ccatccaaat cctcaggctc cccgcggcct tggacaggaa cggtgcttgt gctggcgcca	1680
agtacgtgct caagctggaa ggtgagcact ggaccgtctc tgtgaccctt gggatgacct	1740
cttctttgct ccccttgaa tgtgttcagg gttgttgtga gcataagagc ggtcttggtt	1800
tcccagacgt ggtcgaagtt tccggatttg accctgcctg tcttgaccga cttgctgaga	1860
taatgcactt gcctagcagt gtcattccag ctgctctggc cgagatgtcc gacgacttca	1920
atcgtctggc ttccccggcc gccactgtgt ggactgtttc gcaattcttt gcccgccaca	1980
gaggaggaga gcatcctgac cagggtgtgt tagggaaaat tatcaacctt tgtcaggtga	2040
ttgaggaatg ctgctgttcc cggaacaaag ccaaccgggc taccocggaa gaggttgccg	2100
caaaagttga ccagtacctc cgtggtgcag caagccttg agaatgcttg gccaaagctg	2160
agagggctcg cccgccgagc gcgacggaca cctcctttga ttggaatgtt gtgcttctg	2220
gggttgagac ggcgaatcag acaaccaaac agctccatgt caaccagtgc cgcgctctgg	2280
ttcctgtcgt gactcaagag cctttggaca gagactcggc ccctctgacc gccttctcgc	2340
tgtccaattg ctactaccct gcacaagggt acgaggtccg tcaccgtgag aggctaaact	2400
ccttgctctc taagttggag ggggttggtc gtgaggaata tgggctcacg ccaactggac	2460
ctggcccgcg acccgactg ccgaacgggc tcgacgagct taaagaccag atggaggagg	2520

atctgctgaa attagtcaac gccagggcaa cttcagaaat gatggcctgg gcagccgagc	2580
aggttgatct aaaagcttgg gtcaaaaatt acccacgggtg gacaccgccca cccctccac	2640
caagagttca gcctcgaaaa acgaagtctg tcaagagctt gctagagAAC aagcctgtcc	2700
ctgctccgcg caggaaggtc agatctgatt gtggcagccc gattttgatg ggcgacaatg	2760
ttcctaacgg ttgggaagat tcgactgttg gtggccccct tgatctttcg gcaccatccg	2820
agccgatgac acctctgagt gagcctgtac ttatttccag gccagtgaca tctttgagtg	2880
tgccggcccc agttcctgca ccgcgtagag ctgtgtcccg accgatgacg ccctcgagtg	2940
agccaatttt tgtgtctgca ctgcgacaca aatttcagca ggtggaaaaa gcaaactctgg	3000
cggcagcagc gccgatgtgc caggacgaac ccttagattt gtctgcatcc tcacagactg	3060
aatatgaggc ttcccccta acaccaccgc agaacgtggg cattctggag gtaagggggc	3120
aagaagctga ggaagtctg agtgaaatct cggatattct gaatgatacc aaccctgcac	3180
ctgtgtcatc aagcagctcc ctgtcaagtg ttaagatcac acgccccaaa tactcagctc	3240
aagccattat cgactcgggc gggccctgca gtgggcacct ccaaaggga aaagaagcat	3300
gcctccgcat catgcgtgag gcttgtgatg cggccaagct tagtgaccct gccacgcagg	3360
aatggctttc tcgcatgtgg gataggggtg acatgctgac ttggcgcaac acgtctgctt	3420
accaggcgtt tcgcacctta gatggcaggt ttgggtttct ccaaagatg atactcgaga	3480
cgccgccgcc ctaccctgtg gggtttgtga tgttgccctca caccctgca cttccgtga	3540
gtgcagagag cgaccttacc attggttcag tcgccactga agatattcca cgcctcctcg	3600
ggaaaataga aaataccggt gagatgatca accagggacc cttggcatcc tctgaggaag	3660
aaccggtata caaccaacct gccaaagact cccggatata gtcgcggggg tctgacgaga	3720
gcacagcagc tccgtccgcg ggtacagggtg gcgcgggctt atttactgat ttgccacctt	3780
cagacggcgt agatgcggac ggtggggggc cgttgcagac ggtaagaaag aaagctgaaa	3840
ggctcttcga ccaattgagc cgtcagggtt ttaacctcgt ctcccatctc cctgttttct	3900
tctcacacct cttcaaatct gacagtgggt attctccggg tgattggggg tttgcagctt	3960
ttactctatt ttgcctcttt ttgtgttaca gctaccatt cttcggtttc gttccccctc	4020
tgggtgtatt ttctgggtct tctcggcgtg tgcgcatggg ggtttttggc tgctggctgg	4080
cttttgctgt tggcctgttc aagcctgtgt ccgaccagc cggcactgct tgtgagtttg	4140
actcgccaga gtgcaggAAC gtccttcatt cttttgagct tctcaaacct tgggaccctg	4200
ttcgcagcct tgttgtgggc ccgtcggtc tcggctcttg cattcttggc aggttactgg	4260
gcggggcacg ctacatctgg ctttttttgc ttaggcttgg cattgttgca gattgtatct	4320
tggctggagc ttatgtgctt tctcaaggta ggtgtaaaaa gtgctgggga tcttgtataa	4380
gaactgctcc taatgaaatc gccttcaacg tgttcccttt tacacgtgcg accaggctcg	4440

cactcatcga cctgtgcgat cggttttgtg cgccaacagg catggacccc attttcctcg	4500
ccactgggtg gcgtgggtgc tggaccggcc gaagtcccat tgagcaaccc tctgaaaaac	4560
ccatcgcggtt cgcccagttg gatgaaaaga ggattacggc tagaactgtg gtcgctcagc	4620
cttatgatcc taatcaagcc gtgaagtgtc tgcgggtgtt acaggcgggt ggggcgatgg	4680
tggccgaggg agtcccaaaa gtggccaaag tttctgctat tccattccga gccctttttt	4740
ttcccaccgg agtgaaagtt gatcccgagt gcaggatcgt ggttgacccc gatactttta	4800
ctacagccct cgggtctggt tactctacca caaacctcgt ccttggtgtg ggggactttg	4860
cccagctgaa tggactaaag atcaggcaaa ttccaagcc ttcgggagga ggcccacacc	4920
tcattgctgc cctgcatgtt gcctgctcga tggcgttgca catgcttgcg ggggtttatg	4980
taacttcagt ggggtcttgc ggtgccggca ccaacgatcc atggtgcact aatccgtttg	5040
ccgttcctgg ctacggacca ggctctctct gcacgtccag attgtgcac tcccaacatg	5100
gccttaccct gcccttgaca gcacttgtgg cgggattcgg tcttcaggaa atcgccctgg	5160
tcgttttgat tttcgtttcc atcggaggca tggctcatag gttgagttgt aaggctgata	5220
tgctgtgcat tttacttgca atcgccagct atgtttgggt accccttacc tggttgcttt	5280
gtgtgtttcc ttgttggttg cgctggttct ctttgacccc cttaccatc ctatggttgg	5340
tgtttttctt gatttctgta aatatgcctt cgggaatctt ggccgtgggtg ttattggttt	5400
ctctttggct tttgggacgt tatactaaca ttgctggtct tgtcaccccc tatgatattc	5460
atcattacac cagtggcccc cgcggtgttg ccgccttggc taccgcacca gatggaacct	5520
acttggtgc cgccgcgcg gctgcgttga ctggtgcac catgctgttc acccgtctc	5580
agcttgggtc cttctttgag ggcgctttca gaactcgaaa gccctcactg aacaccgtca	5640
atgtggttgg gtcctccatg ggctctggtg gagtggtcac catcgacggg aaaattaggt	5700
gcgtgactgc cgcacatgtc cttacgggta attcggttag ggtttccgga gtcggcttca	5760
atcaaatgct tgactttgat gtgaaagggg acttcgccat agctgattgc ccgaattggc	5820
aaggagctgc toccaagacc caattctgcg aggacggatg gactggccgt gcctattggc	5880
tgacatcctc tggcgtcgaa ccggtgtta ttgggaatgg attcgcttc tgcttcaccg	5940
cgtgcggcga ttccgggtcc ccagtgatca ccgaagctgg tgagattgtc ggcgttcaca	6000
caggatcaaa taaacaagga ggtggcatcg tcacgcgccc ttcaggccag ttttgtaacg	6060
tggcacccat caagctgagc gaattaagtg aattctttgc tggacccaag gtcccgcctc	6120
gtgatgtgaa ggttggcagc cacataatta aagacacgtg cgaagtacct tcagatcttt	6180
gcgccttgct tgctgcaaaa cctgaactgg agggaggcct ctccaccgtc caacttctgt	6240
gtgtgttttt cctactgtgg agaattgatg gacatgcctg gacgcccttg gttgctgtgg	6300
ggtttttcat tctgaatgag gttctcccag ctgtcctggt tcggagtgtt ttctcctttg	6360

ggatgtttgt gctatcttgg ctacacccat ggtctgcgca agttctgatg atcaggcttc	6420
taacagcagc tcttaacagg aacagatggt cacttgccct ttacagcctt ggtgcggtga	6480
ccggttttgt cgcagatctt gcggttaactc aagggcaccc gttgcaggca gtaatgaatt	6540
tgagcaccta tgccttcttg cctcggtatga tggttgtgac ctaccagtc ccagtgattg	6600
cgtgtggtgt tgtgcaccta cttgccatca ttttgtactt gttcaagtac cgcggcctgc	6660
acaatgttct tgttggtgat ggagcgtttt ctgcagcttt cttcttgca tactttgccg	6720
agggaaagt gaggggaagg gtgtcgcaat cctgcggaat gaatcatgag tcattgactg	6780
gtgccctcgc tatgagactc aatgacgagg acttggaact ccttacgaaa tggactgatt	6840
ttaagtgcct tgtttctgcg tccaacatga ggaatgcagc aggccaattc atcgaggctg	6900
cctatgcaaa agcacttaga attgaacttg ccagttggt gcaggttgat aaggttcgag	6960
gtactttggc caagcttgag gcttttgctg ataccgtggc accccaactc tcgcccggtg	7020
acattgttgt tgctcttggc catacgcctg ttggcagcat cttcgaccta aaggttggtg	7080
gtaccaagca tactctcaa gtcattgaga ccagagtcct tgccgggtcc aaaatgaccg	7140
tggcgcgcgt cgttgacca acccccacgc cccacccgc acccgtgcc atccccctcc	7200
caccgaaagt tctagagaat ggtcccaacg cctgggggga tggggaccgt ttgaataaga	7260
agaagaggcg taggatggaa accgtcggca tctttgtcat gggtggaag aagtaccaga	7320
aattttggga caagaattcc ggtgatgtgt ttacgagga ggtccatgac aacacagatg	7380
cgtgggagtg cctcagagtt ggtgaccctg ccgactttaa ccctgagaag ggaactctgt	7440
gtgggcatac tactattgaa gataaggatt acaaagtcta cgcctcccca tctggcaaga	7500
agttcctggt ccccgtaac ccagagagcg gaagagccca atgggaagct gcaaagcttt	7560
ccgtggagca ggcccttggc atgatgaatg tcgacggtga actgacggcc aaagaagtgg	7620
agaaactgaa agaataatt gacaaacttc agggcctgac taaggagcag tgtttaaact	7680
gctagccgcc agcggcttga cccgctgtgg tcgcggcggc ttggttgta ctgagacagc	7740
ggtaaaaata gtcaaatttc acaaccggac ttccacccta gggcctgtga atttaaaagt	7800
ggccagtgag gttgagctga aagacgcggt cgagcacaa caacaccccg ttgcaagacc	7860
ggttgacggt ggtgttgtgc tcctgcgttc cgcagttcct tcgcttatag atgtcctgat	7920
ctccggtgct gacgatctc ctaagttact cgctcgtcac gggccgggga aactgggat	7980
cgatggcacg ctttgggact ttgaggccga ggccacaaa gaggaattg cgctcagtgc	8040
gcaaataata caggcttgtg acattaggcg cggtgacgca cctgaaattg gtctccctta	8100
caagctgtac cctgttaggg gcaaccctga gcgggtaaaa ggagttttac agaatacaag	8160
gtttggagac ataccttaca aaacccccag tgacactggg agcccagtgc acgcggtgc	8220
ctgcctcacg cccaatgcca ctccggtgac tgatgggcgc tccgtcttgg ctactaccat	8280

gccctccggt tttgaattgt atgtaccgac cattccagcg tctgtccttg attatcttga 8340
 ctctaggcct gactgcccc aacagttgac agagcacggc tgtgaggatg ccgcattgag 8400
 agacctctcc aagtatgact tgtccaccca aggcctttgtt ttacctgggg ttcttcgcct 8460
 tgtgcgtaag tacctgtttg cccacgtggg taagtgcccg cccgttcacg gcccttccac 8520
 ttacctgcc aagaattcta tggctggaat aaatgggaac aggtttccaa ccaaggacat 8580
 tcagagcgtc cccgaaatcg acgttctgtg cgcacaggcc gtgcgagaaa actggcaaac 8640
 tgttaccctt tgtaccctca agaaacagta ttgtgggaag aagaagacta ggacaatact 8700
 cggcaccaat aatttcattg cgttggccca cggggcagcg ttgagtgggtg tcaccagggt 8760
 cttcatgaaa aaggcgttta actcgcccat cgccctcggg aaaaacaaat ttaaggagct 8820
 acagactccg atcttaggca ggtgccttga agctgatctt gcacctctgtg atcgatccac 8880
 acctgcaatt gtccgctggg ttgccgcaa ccttctttat gaacttgcct gtgctgaaga 8940
 gcacctaccg tcgtacgtgc tgaactgctg ccatgacctt ttggtcacgc agtccggcgc 9000
 agtgactaag aggggtggcc tgcgtctgtg cgacccgatc acttctgtgt ctaacaccat 9060
 ttacagcttg gtgatatatg cacagcacat ggtgcttagt tactttaaaa gtggtcaccc 9120
 tcatggcctt ctgttctac aagaccagct gaagttcgag gacatgctca aagtccaacc 9180
 cctgatcgtc tattcgagc acctcgtgct gtatgccga tctccacca tgccgaacta 9240
 ccaactgggtg gtcgaacatc tgaatttgat gctgggtttt cagacggacc caaagaagac 9300
 agccataacg gactcgccat catttctagg ctgtaggata ataaatggac gccagctagt 9360
 cccaaccgt gacaggatcc tcgcggccct cgcttaccat atgaaggcaa gcaatgtttc 9420
 tgaatactac gccgcggcgg ctgcaatact catggacagc tgtgcttgtt tagagtatga 9480
 tcctgaatgg tttgaagagc ttgtggttgg gatagcgagc tgcgcccga aggacggcta 9540
 cagctttccc ggcccgccgt tcttcttgc catgtgggaa aaactcagat ccaatcatga 9600
 ggggaagaag tcagaaatgt gcgggtattg cggggccccc gctccgtacg ccaactgcctg 9660
 tggcctcgac gtctgtatgt accacaccca ctccaccag cattgtccag tcacaatctg 9720
 gtgtggccac ccggtgggt ctggttcttg tagtgagtgc aaaccccc tagggaaag 9780
 cacaagccct ctagatgagg tgtagaaca agtcccgat aagcctccac ggactgtaat 9840
 catgcatgtg gagcagggtc tcacccctct tgaccaggc agataccaga ctgcgcgcg 9900
 attagtctcc gttaggcgtg gcacagagg aaatgaagtt gacctaccag acggtgatta 9960
 tgctagcacc gccctactcc ccaattgtaa agagatcaac atggtcgctg tcgcctctaa 10020
 tgtgttgcc agcaggttca tcacgggtcc gcccggtgct gggaaaacat actggctcct 10080
 tcagcaggtc caggatgggt atgtcattta cacaccgact caccagacca tgctcgacat 10140
 gattagggct ttggggacgt gccggttcaa cgtccagca ggtacaacgc tgcaattccc 10200

tgccccctcc cgtaccggcc cgtgggttcg catcctagcc ggcggttggt gtccctggtaa	10260
gaattcccttc ttggatgaag cagcgtattg taatcacctt gatgtcttga ggctcccttag	10320
caaaaccacc ctcacctgtc tgggagactt caaacaactc caccagtggt gttttgattc	10380
tcattgctat gtttttgaca tcatgcctca gaccagttg aagaccatct ggagattcgg	10440
acagaacatc tgtgatgcca tccaaccaga ttacagggaac aaacttggtt ccatgggtcaa	10500
cacaaccctg gtaacctacg tggaaaaacc tgtcaagtat gggcaagtcc tcaccctta	10560
ccacagggaac cgagaggacg gcgccatcac aattgactcc agtcaaggcg ccacatttga	10620
tgtggtcaca ctgcatttgc ccaactaaaga ttactcaac aggcaaagag cccttggtgc	10680
tatcaccagg gcaagacatg ctatctttgt gtatgacca cacaggcaat tgcagagcat	10740
gtttgatctt cctgcgaagg gcacaccctg caacctcgca gtgcaccgtg atgagcagct	10800
gatcgtactg gatagaaata ataaagaatg cacagttgct caggctctag gcaacggaga	10860
taaatttagg gccaccgaca agcgcgttgt agattctctc cgcgccattt gtgctgatct	10920
ggaagggtcg agctctccgc tccccaaagt cgcacacaac ttgggatttt atttctcacc	10980
tgatttgaca cagtttgcta aactcccgtt agaccttgca cccactggc ccgtggtgac	11040
aacccagaac aatgaaaagt ggccggatcg gctggttgcc agccttcgcc ctgtccataa	11100
gtatagccgt gcgtgcattg gtgccggcta tatggtgggc ccctcgggtt ttctaggcac	11160
ccctggggtc gtgtcatact acctcacaaa atttgtcaag ggcgaggctc aagtgttcc	11220
ggagacagtc ttcagcaccg gccgaattga ggtggattgc cgggagtatc ttgatgacag	11280
ggagcgagaa gttgctgagt ccctcccaca tgccttcatt ggcgacgtca aaggcaccac	11340
cgttggggga tgtcatcatg tcacctcaa ataccttccg cgcttccttc ccaaggaatc	11400
agtcgcggta gtcgggggtt cgagccccgg gaaagccgca aaagcagtggt gcacattgac	11460
ggatgtgtac ctcccagacc ttgaggccta cctccacca gagactcagt ccaagtgtg	11520
gaaagttatg ttggacttca aggaagttcg actgatggtc tggaaagaca agacggccta	11580
tttccaactt gaaggccgct atttcacctg gtatcagctt gcaagctacg cctcgtacat	11640
ccgtgttctt gtcaactcca cgggtgtatct ggacccctgc atgggccctg ccctttgcaa	11700
cagaagagtt gtcgggtcca ccattgggg agctgacctc gcagtcaccc cttatgatta	11760
cgggtgctaaa atcatcttgt ctagcgctta ccatggtgaa atgcctcctg gatacaagat	11820
tctggcgtgc gcggagttct cgctcgacga ccagtcgaag tacaaacaca cctgggggtt	11880
tgaatcggat acagcgtatc tgtatgagtt caccggaaac ggtgaggact gggaggatta	11940
caatgatgcg tttcgtgcgc gccagaaagg gaaaatttat aaggccactg ctaccagcat	12000
gaagttttat tttccccggg gcccgtcat tgaaccaact ttaggcctga attgaaatga	12060
aatggggtct atacaaagcc tcttcgacaa aattggccag ctttttgtgg atgctttcac	12120

ggaatTTTTg gtgtccattg ttgatatcat catatTTTTg gccatTTTgt ttggcttcac	12180
catcgccgggt tggtcggtgg tcttttgcac cagattgggt tgctccgagg tattccgtgc	12240
gogccctgcc attcaccctg agcaattaca gaagatccta tgaggccttt ctttctcagt	12300
gccgggtgga cattcccacc tggggggtaa aacacccttt ggggatgttt tggcaccata	12360
aggtgtcaac cctgattgat gaaatgggtg cgcgtcgaat gtaccgcac atggaaaaag	12420
cagggcaagc tgcctggaaa caggtgggtga gcgaggctac gctgtctcgc attagtagtt	12480
tggatgtgggt ggctcatttt caacatcttg ccgccattga agccgagacc tgtaaatatt	12540
tggcttctcg actgcccatg ctacacaacc tgcgcacgac aggggtcaaat gtaaccatag	12600
tgtataatag cactttaaat caggtgtttg ctatTTTTcc aaccctgggt tcccggccaa	12660
agcttcatga ttttcagcaa tggctaatag ctgtacattc ctccatattt tcctctgttg	12720
cagcttcttg tactcttttt gttgtgctgt ggttgccgggt tccaatgcta cgtactgttt	12780
ttggtttccg ctggttaggg gcaatTTTTc tttcgaactc atgggtgaatt acacgggtgtg	12840
tccaccttgc ctaccccgac aagcagccgc tgaggtcctt gaaccggga ggtctctttg	12900
gtgcaggata gggcatgacc gatgtgggga ggacgatcac gacgaactgg ggttcatgg	12960
tccgcctggc ctctccagcg aaagccactt gaccagtgtt tacgcctgggt tggcgcttct	13020
gtccttcagc tacacggccc agttccatcc cgagatattt gggataggga acgtgagtga	13080
agtttatgtt gacatcaagc accaattcat ctgcgccgtt catgacgggc agaaccacc	13140
cttgccctgc catgacaata tttcagccgt atttcagacc tactatcaac atcaggtcga	13200
cggcggcaat tggtttcacc tagaatggct gcgtcccttc ttttctctt ggttggtttt	13260
aaatgtttcg tggttttctca ggcgttcgcc tgcaagccat gtttcagttc gagtctttca	13320
gacatcaaaa ccaacactac cgcagcatca ggctttgttg tcctccagga catcagctgc	13380
cttaggcatg gcgactcgtc ctttccgacg attcgcaaaa gctctcaatg ccgcacggcg	13440
atagggacac ctgtgtatat caccatcaca gccaatgtga cagatgagaa ttacttacat	13500
tcttctgac tcctcatgct ttcttcttgc cttttctatg cttctgagat gagtgaag	13560
ggattcaagg tggtatTTgg caatgtgtca ggcatcgtgg ctgtgtgtgt caactttacc	13620
agctacgtcc aacatgtcaa agagtttacc caacgctcct tgggtggcga tcatgtgcgg	13680
ctgcttcatt tcatgacacc tgagaccatg aggtgggcaa ccgTTTTagc ctgtcttttt	13740
gccatcctac tggcaatttg aatgttcaag tatgttgggg aaatgcttga ccgcgggctg	13800
ttgctcgcga ttgctttctt tgtgggtgat cgtgccgttc tggtttgctg tgctcggcaa	13860
cgccaacagc agcagcagct ctcatTTtca gttgatttat aacttgacgc tatgtgagct	13920
gaatggcaca gattggctgg cagaaaaatt tgattgggca gtggagactt ttgtcatctt	13980
tcccggtgtg actcacattg tttcctatgg tgcactcacc accagccatt tccttgacac	14040

```

agttggtctg gttactgtgt ccaccgccgg gttttatcac gggcggtatg tcttgagtag      14100
catctacgcg gtctgtgctc tggctgcgtt gatttgcttc gttattaggc ttgcgaagaa      14160
ctgcatgtcc tggcgctact cttgtaccag atataccaac ttccttcttg acactaaggg      14220
cagactctat cgttggcggt cgcccggttat catagaaaaa gggggtaagg ttgaggtcga      14280
aggtcactctg atcgacctca aaagagttgt gcttgatggt tccgtggcaa cccctttaac      14340
cagagtttca gcggaacaat ggggtcgtct ctagacgact tttgccatga tagcacggct      14400
ccacaaaagg tgcttttggc gttttccatt acctacacgc cagtaatgat atatgctcta      14460
aaggtaagtc ggggcgcact gctagggctt ctgcaccttt tgatctttct gaattgtgct      14520
tttaccttcg ggtacatgac attcgcgcac tttcagagca caaatagggg cgcgctcgct      14580
atgggagcag tagttgcact tctttggggg gtgtactcag ccatagaaac ctggaaattc      14640
atcacctcca gatgccgttt gtgcttgcta ggccgcaagt acattctggc ccctgccac      14700
cacgtcgaaa gtgccgcggg ctttcatccg attgcggaac atgataacca cgcatttgct      14760
gtccggcgtc ccggctccat tacggttaac ggcacattgg tgcccggggt gaaaagcctc      14820
gtgttgggtg gcagaaaagc tgtaaacag ggagtggtaa accttgtaa atatgccaaa      14880
taacaacggc aagcagcaaa agaaaaagaa ggggaatggc cagccagtca accagctgtg      14940
ccagatgctg ggtaaaatca tcgcccagca aaaccagtcc agaggcaagg gaccgggctc      15000
taagaagaaa tccccggaga agccccatth tctctagcg accgaagatg acgtcaggca      15060
tcacttcacc cctggtgagc ggcaattgtg tctgtcgtcg atccagactg cctttaacca      15120
gggcgctgga acttgtaacc tgtcagattc agggaggata agttacactg tggagttag      15180
tttgccgacg catcactactg tgcgcctgat ccgcgtcaca gcacaccct cagcatgatg      15240
ggctggcatt ctttaggcac ctcaagtgtc gaattggaag aatgtgtggt ggatggcact      15300
gattgacatt gtgcctctaa gtcacctatt caattagggc gaccgtgtgg gggtaaaatt      15360
taattggcga gaacctgcg gccgcaatta aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa      15420
aaaaaaaaaa aaaaaaaaaa aaaa                                     15444

```

<210> 3

<211> 15450

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Vacxin chứa hệ gen hoàn chỉnh của P129-PK-FL lần cấy truyền 24 (15,450 nt)

<400> 3

```

atgacgtata ggtgttggct ctatgccacg gcatttgtat tgtcaggagc tgtgaccatt      60
ggcacagccc aaaacttgct gcacggaaaa cgcccttctg tgacagcctt cttcagggga      120
gcttaggggt ctgtccctag caccttgctt ctggagttgc actgctttac ggtctctcca      180

```

cccctttaac catgtctggg atacttgatc ggtgcacgtg ccccccaat gccagggtgt 240
 ttatggcgga gggccaagtc tactgcacac gatgtctcag tgcacgggtct ctcttctctc 300
 tgaatctcca agttcctgag cttgggggtgc tgggcctatt ttataggccc gaagagccac 360
 tccggtggac gttgccacgt gcattcccca ctgtcgagtgc ctccctgcc ggggcctgct 420
 ggctttctgc gatctttcca attgcacgaa tgaccagtgg aaacctgaac tttcaacaaa 480
 gaatggtgcg ggttgacagt gagatttaca gagccggcca actcaccct gcagttttga 540
 aggctctaca agtttatgaa cgggggtgtc gctgggtacc cattgtcgga cctgtccctg 600
 gagtggccgt ttacgccaac tccctacatg tgagtgacaa acctttcccg ggagcaactc 660
 atgtgttaac caacctaccg ctcccgaga ggccaagcc tgaagacttt tgcccttttg 720
 agtgtgctat ggctaacgtc tatgacattg gccataacgc cgtcatgtat gtggccagag 780
 ggaaagtctc ctgggccct cgtggcgggg atgaagtga atttgaaacc gtccccgaag 840
 agttgaagtt gattgcgaac cgactccaca tctccttccc gcccaccac gcagtggaca 900
 tgtctgagtt tgccttcata gccctggga gtggtgtctc cttgcgggtc gagcaccaac 960
 acggctgcct tcccgctgat actgtccctg atgggaactg ctggtggtac ttgtttgact 1020
 tgctcccacc ggaagttcag aataaagaaa ttcgccgtgc taaccaattt ggctatcaaa 1080
 ccaagcatgg tgtccatggc aagtacctac agcggaggct gcaagttaat ggtctccgag 1140
 cagtgactga tacagatgga cctattgtcg tacagtactt ctctgttagg gagagttgga 1200
 tccgccactt cagactggcg gaagaaccta gcctccctgg gtttgaagac ctctcagaa 1260
 taagggtaga gcctaatacg tcgccaatgg gtggcaaggg tgaaaaaatc ttccggtttg 1320
 gcagtcacaa gtggtacggt gctggaaaga gagcaaggag agcacgctct ggtgcgactg 1380
 ccacggtcgc tcaccgcgt ttgccgctc gcgaagcca gcaggccaag aagctcgagg 1440
 ttgccagcgc caacagggt gagcatctca agtactattc cccgcctgcc gacgggaact 1500
 gtggttgga ctgcatttcc gccattacca accggatggt gaattccaaa tttgaaacca 1560
 ctcttccga gagagtgaga ccttcagatg actgggctac tgacgaggat cttgtgaata 1620
 ccatccaaat cctcaggctc cccgcggcct tggacaggaa cgggtgcttg gctggcgcca 1680
 agtacgtgct caagctggaa ggtgagcact ggaccgtctc tgtgaccct gggatgacct 1740
 cttctttgct ccccttgaa tgtgttcagg gttgttgtga gcataagagc ggtcttggtt 1800
 tcccagacgt ggtcgaagtt tccgatttg acctgcctg tcttgaccga cttgctgaga 1860
 taatgcactt gcctagcagt gtcattccag ctgctctggc cgagatgtcc gacgacttca 1920
 atcgtctggc tccccggcc gccactgtgt ggactgttcc gcaattcttt gcccgccaca 1980
 gaggaggaga gcctcctgac caggtgtgct tagggaaaat tatcaacctt tgtcaggtga 2040
 ttgaggaatg ctgctgttcc cggaacaaag ccaaccgggc taccggga gaggttgcgg 2100

caaaagttga ccagtacctc cgtggtgcag caagccttgg agaagcttg gccaaagcttg	2160
agagggctcg cccgccgagc gcgacggaca cctcctttga ttggaatgtt gtgcttcctg	2220
gggttgagac ggcgaaatcag acaaccaaac agctccatgt caaccagtgc cgcgctctgg	2280
ttcctgtcgt gactcaagag cctttggaca gagactcggc cctctgacc gccttctcgc	2340
tgtccaattg ctactaccct gcacaaggtg acgaggtccg tcaccgtgag aggctaaact	2400
ccttgctctc taagttggag ggggttggtc gtgaggaata tgggctcacg ccaactggac	2460
ctggcccgcg acccgcaactg ccgaacgggc tcgacgagct taaagaccag atggaggagg	2520
atctgctgaa attagtcaac gcccaggcaa cttcagaaat gatggcctgg gcagccgagc	2580
aggttgatct aaaagcttgg gtcaaaaatt acccacggtg gacaccgcca cccctccac	2640
caagagttca gcctcgaaaa acgaagtctg tcaagagctt gctagagaac aagcctgtcc	2700
ctgctccgcg caggaaggtc agatctgatt gtggcagccc gattttgatg ggcgacaatg	2760
ttcctaacgg ttgggaagat tcgactgttg gtggtcccct tgatctttcg gcaccatccg	2820
agccgatgac acctctgagt gagcctgtac ttatttccag gccagtgaca tctttgagtg	2880
tgccggcccc agttcctgca ccgcgtagag ctgtgtcccg accgatgacg ccctcgagtg	2940
agccaatttt tgtgtctgca ctgcgacaca aatttcagca ggtggaaaaa gcaaactctg	3000
cggcagcagc gccgatgtgc caggacgaac ccttagattt gtctgcatcc tcacagactg	3060
aatatgaggc ttcccccta acaccaccgc agaacgtggg cattctggag gtaagggggc	3120
aagaagctga ggaagtctg agtgaaatct cggatattct gaatgatacc aaccctgcac	3180
ctgtgtcatc aagcagctcc ctgtcaagtg ttaagatcac acgccccaaa tactcagctc	3240
aagccattat cgactcgggc gggccctgca gtgggcacct ccaaagggaa aaagaagcat	3300
gcctccgcat catgcgtgag gcttgtgatg cggccaagct tagtgaccct gccacgcagg	3360
aatggctttc tcgcatgtgg gataggggtg acatgctgac ttggcgcaac acgtctgctt	3420
accaggcgtt tcgcacctta gatggcaggt ttgggtttct ccaaagatg atactcgaga	3480
cgcgcgcgcc ctaccctgtg gggtttgtga tggtgcctca caccctgca ccttccgtga	3540
gtgcagagag cgaccttacc attggttcag tcgccactga agatattcca cgcacctcg	3600
ggaaaataga aaataccggt gagatgatca accagggacc cttggcatcc tctgaggaag	3660
aaccggtata caaccaacct gccaaagact cccgatatc gtcgcggggg tctgacgaga	3720
gcacagcagc tccgtccgcg ggtacaggtg gcgccggctt atttactgat ttgccacctt	3780
cagacggcgt agatgcggac ggtggggggc cgttgcagac ggtaagaaaag aaagctgaaa	3840
ggctcttcga ccaattgagc cgtcagggtt ttaacctcgt ctcccatctc cctgttttct	3900
tctcacacct cttcaaatct gacagtgggt attctccggg tgattggggg tttgcagctt	3960
ttactctatt ttgcctcttt ttgtgttaca gctaccatt cttcggtttc gttccctct	4020

tgggtgtatt ttctgggtct tctcggcgtg tgcgcacatggg ggTTTTTggc tgctggctgg	4080
cttttgctgt tggcctgttc aagcctgtgt ccgacccagt cggcactgct tgtgagtttg	4140
actcgccaga gtgcaggaac gtccttcatt cttttgagct tctcaaacct tgggaccctg	4200
ttcgcagcct tgttggtggc cccgtcggtc toggctttgc cattcttggc aggttactgg	4260
gcggggcacg ctacatctgg cattttttgc ttaggcttgg cattgttgca gattgtatct	4320
tggctggagc ttatgtgctt tctcaaggta ggtgtaaaaa gtgctgggga tcttgataa	4380
gaactgctcc taatgaaatc gccttcaacg tgttcccttt tacacgtgcg accaggctgt	4440
cactcatcga cctgtgcgat cggTTTTgtg cgccaacagg catggacccc attttcctcg	4500
ccactgggtg gcgtgggtgc tggaccggcc gaagtcccat tgagcaaccc tctgaaaaac	4560
ccatcgctt cgcccagttg gatgaaaaga ggattacggc tagaactgtg gtcgctcagc	4620
cttatgatcc taatcaagcc gtgaagtgt tgcgggtgtt acaggcgggt ggggcgatgg	4680
tggccgaggc agtcccaaaa gtggccaaag tttctgctat tccattccga gccccttttt	4740
ttcccaccgg agtgaaagt gatcccgagt gcaggatcgt ggttgacccc gatactttta	4800
ctacagccct cgggtctggt tactctacca caaacctcgt ccttggtgtg ggggactttg	4860
cccagctgaa tggactaaag atcaggcaaa tttccaagcc ttcgggagga ggcccacacc	4920
tcattgctgc cctgcatgtt gcctgctcga tggcgttgca catgcttgcg ggggtttatg	4980
taacttcagt ggggtcttgc ggtgccggca ccaacgatcc atggtgcact aatccgtttg	5040
ccgttcctgg ctacggacca ggctctctct gcacgtccag attgtgcac tcccaacatg	5100
gccttaccct gcccttgaca gcaattgtgg cgggattcgg tcttcaggaa atgccttgg	5160
tcgttttgat tttcgtttcc atcggaggca tggctcatag gttgagttgt aaggctgata	5220
tgctgtgcat tttacttgca atcgccagct atgtttgggt accccttacc tggttgcttt	5280
gtgtgtttcc ttgttggttg cgctggttct ctttgacccc ccttaccatc ctatggttg	5340
tgTTTTtctt gatctctgta aatatgcctt cgggaatctt ggccgtggtg ttattggttt	5400
ctctttggct tttgggacgt tatactaaca ttgctggtct tgtcaccccc tatgatattc	5460
atcattacac cagtggcccc cggggtgttg ccgccttggc tacgcacca gatggaacct	5520
acttggtgc cgtccgccgc gctgcgttga ctggctgcac catgctgttc acccgtctc	5580
agcttgggtc ccttcttgag ggcgctttca gaactcgaaa gccctcactg aacaccgtca	5640
atgtggttgg gtcctccatg ggctctggtg gagtggtcac catcgacggg aaaattaggt	5700
gcgtgactgc cgcacatgtc cttacgagta attcggctag ggtttccgga gtcggcttca	5760
atcaaatgct tgactttgat gtgaaagggg acttcgccat agctgattgc ccgaattggc	5820
aaggagctgc tcccaagacc caattctgcg aggacggatg gactggccgt gcctattggc	5880
tgacatctc tggcgtcgaa cccggtgtta ttgggaatgg attcgccttc tgcttcaccg	5940

cgtgcggcga ttccgggtcc ccagtgatca ccgaagctgg tgagattgtc ggcgttcaca	6000
caggatcaaa taaacaagga ggtggcatcg tcacgcgccc ttcaggccag ttttgtaacg	6060
tggcaccat caagctgagc gaattaagtg aattctttgc tggacccaag gtcccgcctcg	6120
gtgatgtgaa ggttggcagc cacataatta aagacacgtg cgaagtacct tcagatcttt	6180
gcgccttgct tgctgcaaaa cctgaactgg agggaggcct ctccaccgtc caacttctgt	6240
gtgtgttttt cctactgtgg agaattgatgg gacatgcctg gacgcccttg gttgctgtgg	6300
ggtttttcat tctgaatgag gttctccag ctgtcctggg tcggagtgtt ttctccttg	6360
ggatgtttgt gctatcttgg ctacacccat ggtctgcgca agttctgatg atcaggcttc	6420
taacagcagc tcttaacagg aacagatggg cacttgcctt ttacagcctt ggtgcggtga	6480
ccggttttgt cgcagatctt gcggttaactc aagggcaccg gttgcaggca gtaatgaatt	6540
tgagcaccta tgccttctcg cctcgatga tggttgtgac ctaccagtc ccagtgattg	6600
cgtgtggtgt tgtgcaccta cttgccatca ttttgtactt gttcaagtac cgcggcctgc	6660
acaatgttct tgttggtgat ggagcgtttt ctgcagcttt cttcttgca tactttgccg	6720
agggaaagt gaggggaagg gtgtcgcaat cctgcggaat gaatcatgag tcattgactg	6780
gtgccctcgc tatgagactc aatgacgagg acttgactt ccttacgaaa tggactgatt	6840
ttaagtgtt tgtttctgcg tccaacatga ggaatgcagc aggccaattc atcgaggctg	6900
cctatgcaaa agcacttaga attgaacttg ccagttggg gcaggttgat aaggttcgag	6960
gtactttggc caagcttgag gcttttgctg ataccgtggc accccaactc tcgcccgtg	7020
acattgttgt tgctcttggc catacgctg ttggcagcat cttcgaccta aaggttggtg	7080
gtaccaagca tactctccaa gtcattgaga ccagagtcct tgccgggtcc aaaatgaccg	7140
tggcgcgcgt cgttgacca acccccacgc cccacccgc acccgtgcc atccccctcc	7200
caccgaaagt tctagagaat ggtcccaacg cctgggggga tggggaccgt ttgaataaga	7260
agaagaggcg taggatggaa accgtcggca tctttgtcat ggggtgggaag aagtaccaga	7320
aattttggga caagaattcc ggtgatgtgt ttacagagga ggtccatgac aacacagatg	7380
cgtgggagtg cctcagagtt ggtgacctg ccgactttaa ccctgagaag ggaactctgt	7440
gtgggcatac tactattgaa gataaggatt acaaagtcta cgctcccca tctggcaaga	7500
agttcctggg ccccgtaac ccagagagcg gaagagccca atgggaagct gcaaagcttt	7560
ccgtggagca ggcccttggc atgatgaatg tcgacggtga actgacggcc aaagaagtgg	7620
agaaactgaa aagaataatt gacaaacttc agggcctgac taaggagcag tgtttaaact	7680
gctagccgcc agcggcttga cccgctgtgg tcgcgggcgc ttggttgta ctgagacagc	7740
ggtaaaaata gtcaaatttc acaaccggac ttccacccta gggcctgtga atttaaaagt	7800
ggccagtgag gttgagctga aagacgcggt cgagcacaac caacaccggt ttgcaagacc	7860

ggttgacggt ggtgttggtgc tctgcggttc cgcagttcct tcgcttatag atgtcctgat	7920
ctccggtgct gacgcatctc ctaagttact cgctcgtcac gggccgggga aactgggat	7980
cgatggcacg ctttgggact ttgaggccga ggccaccaa gaggaattg cgctcagtgc	8040
gcaaataata caggcttggt acattaggcg cggtgacgca cctgaaattg gtctccctta	8100
caagctgtac cctgttaggg gcaaccctga gcgggtaaaa ggagttttac agaatacaag	8160
gtttggagac atacottaca aaacccccag tgacactggg agcccagtgc acgcggtgc	8220
ctgctcagc ccaatgcc ctccggtgac tgatgggcgc tccgtcttgg ctactaccat	8280
gccctccggt tttgaattgt atgtaccgac cattccagcg tctgtccttg attatcttga	8340
ctctaggcct gactgcccc aacagttgac agagcacggc tgtgaggatg ccgcattgag	8400
agacctctcc aagtatgact tgtccacca aggctttgtt ttacctgggg ttcttcgcct	8460
tgtgcgtaag tacctgtttg ccacgtggg taagtgccg cccgttcac ggcttccac	8520
ttacctgcc aagaattcta tggctggaat aaatgggaac aggtttccaa ccaaggacat	8580
tcagagcgtc cccgaaatcg acgttctgtg cgcacaggcc gtgcgagaaa actggcaaac	8640
tgttaccct tgtaccctca agaaacagta ttgtgggaag aagaagacta ggacaatact	8700
cggcaccaat aatttcattg cgttgcccc cggggcagcg ttgagtgggtg tcaccaggg	8760
cttcatgaaa aaggcgttta actcgcccat cgccctcggg aaaaacaaat ttaaggagct	8820
acagactccg atcttaggca ggtgccttga agctgatctt gcacctgtg atcgatccac	8880
acctgcaatt gtccgctggt ttgccgcaa ccttctttat gaacttgcct gtgctgaaga	8940
gcacctaccg tcgtacgtgc tgaactgctg ccatgaccta ttggtcacgc agtccggcgc	9000
agtgactaag aggggtggcc tgtcgtctgg cgaccgac acttctgtgt ctaacaccat	9060
ttacagcttg gtgatatatg cacagcacat ggtgcttagt tactttaaaa gtggtcacc	9120
tcatggcctt ctgttcctac aagaccagct gaagttcgag gacatgctca aagtccaacc	9180
cctgatcgtc tatteggacg acctcggtg gtatgccgaa tctcccacca tgccgaacta	9240
ccactgggtg gtgcgaacat tgaatttgat gctgggttt cagacggacc caaagaagac	9300
agccataacg gactcgccat ctttctagg ctgtaggata ataaatggac gccagctagt	9360
ccccaccgt gacaggatcc tcgcgccct cgcttaccat atgaaggcaa gcaatgttct	9420
tgaatactac gccgcgcggt ctgcaatact catggacagc tgtgcttgt tagagtatga	9480
tcctgaatgg tttgaagagc ttgtggttgg gatagcgag tcgccccgca aggacggcta	9540
cagctttccc ggccgcggt tcttcttgtc catgtgggaa aaactcagat ccaatcatga	9600
ggggaagaag tccagaatgt gcgggtattg cggggcccc gctccgtac cactgcctg	9660
tggcctcgac gtctgtattt accacacca cttccaccag cattgtccag tcacaatctg	9720
gtgtggccac ccggtggtt ctggttcttg tagtgagtgc aaaccccc tagggaaagg	9780

cacaagccct ctagatgagg tgtagaaca agtcccgtat aagcctccac ggactgtaat	9840
catgcatgtg gagcagggtc tcacccctct tgaccaggc agataccaga ctgcccgcg	9900
attagtctcc gttaggcgtg gcatcagagg aatgaagtt gacctaccag acggtgatta	9960
tgctagcacc gccctactcc ccacttgtaa agagatcaac atggtcgctg tcgcctctaa	10020
tgtgttgccg agcaggttca tcatcggtec gcccggtgct gggaaaacat actggctcct	10080
tcagcaggtc caggatgggt atgtcattta cacaccgact caccagacca tgctcgacat	10140
gattagggct ttggggacgt gccggttcaa cgtcccagca ggtacaacgc tgcaattccc	10200
tgccccctcc cgtaccggcc cgtgggttcg catcctagcc ggcgggttgggt gtccctggtaa	10260
gaattccttc ttggatgaag cagcgtattg taatcacctt gatgtcttga ggctccttag	10320
caaaaccacc ctacctgtc tgggagactt caaacaactc caccagtggt gttttgattc	10380
tcattgctat gtttttgaca tcatgcctca gaccagttg aagaccatct ggagattcgg	10440
acagaacatc tgtgatgcca tccaaccaga ttacaggac aaacttgtgt ccatggtcaa	10500
cacaaccgt gtaacctacg tggaaaaacc tgtcaagtat gggcaagtcc tcaccctta	10560
ccacaggac cgagaggacg gcgccatcac aattgactcc agtcaaggcg ccacatttga	10620
tgtggtcaca ctgcatttgc ccactaaaga ttactcaac aggcaaagag cccttggtgc	10680
taccagcagg gcaagacatg ctatctttgt gtatgacca cacaggcaat tgcagagcat	10740
gtttgatctt cctgcgaagg gcacaccgt caacctcgca gtgcaccgtg atgagcagct	10800
gatcgactg gatagaaata ataaagaatg cacagttgct caggctctag gcaacggaga	10860
taaatttagg gccaccgaca agcgcgttgt agattctctc cgcgccattt gtgctgatct	10920
ggaagggtcg agctctccgc tccccagggt cgcacacaac ttgggatttt atttctcacc	10980
tgatttgaca cagtttgcta aactcccgt agaccttgca cccactggc ccgtggtgac	11040
aaccagaac aatgaaaagt ggccggatcg gctggttgcc agccttcgcc ctgtccataa	11100
gtatagccgt gcgtgcattg gtgccggcta tatggtgggc ccctcgggtg ttctaggcac	11160
ccctggggtc gtgtcatact acctcacaaa atttgtcaag ggcgaggctc aagtgtctcc	11220
ggagacagtc ttcagcaccg gccgaattga ggtggattgc cgggagtatc ttgatgacag	11280
ggagcgagaa gttgctgagt cctcccaca tgccttcatt ggcgacgtca aaggcaccac	11340
cgttggggga tgtcatcatg tcacctcaa ataccttcog cgcttccttc ccaaggaatc	11400
agtcgcggta gtcggggttt cgagccccgg gaaagccgca aaagcagtgt gcacattgac	11460
ggatgtgtac ctcccagacc ttgaggccta cctccacca gagactcagt ccaagtgtg	11520
gaaagttatg ttggacttca aggaagttcg actgatggtc tggaaagaca agacggccta	11580
tttccaactt gaaggccgt atttcacctg gtatcagctt gcaagctacg cctcgtacat	11640
ccgtgttct gtcaactcca cgggtgatct ggaccctgc atgggcctg ccctttgcaa	11700

cagaagagtt gtcgggtcca cccattgggg agctgacctc gcagtcaccc cttatgatta	11760
cggtgctaaa atcatcttgt ctagcgctta ccatggtgaa atgcctcctg gatacaagat	11820
tctggcgtgc gcgagttct cgctcgacga cccagtcaag taaaaacaca cctgggggttt	11880
tgaatcggat acagcgtatc tgtatgagtt caccggaaac ggtgaggact gggaggatta	11940
caatgatgcg tttcgtgcgc gccagaaagg gaaaatttat aaggccactg ctaccagcat	12000
gaagttttat tttcccccg gcccgctcat tgaaccaact ttaggcctga attgaaatga	12060
aatggggctc atacaaagcc tcttcgacaa aattggccag ctttttgtgg atgctttcac	12120
ggaatTTTTg gtgtccattg ttgatatcat catatTTTTg gccatTTTgt ttggcttcac	12180
catcgccggt tggctgggtg tcttttgcac cagattgggt tgctccgcgg tattccgtgc	12240
gcgccctgcc attcacctg agcaattaca gaagatccta tgaggccttt ctttctcagt	12300
gccgggtgga cattcccacc tggggggtaa aacacccttt ggggatgttt tggcaccata	12360
aggtgtcaac cctgattgat gaaatggtgt cgcgtcgaat gtaccgcac atggaaaaag	12420
cagggcaagc tgcctgaaa caggtggtga gcgaggctac gctgtctcgc attagtagtt	12480
tggatgtggt ggctcatTTT caacatcttg ccgccattga agccgagacc tgtaaatatt	12540
tggcttctcg actgcccatt ctacacaacc tgcgcatgac agggTcaaat gtaaccatag	12600
tgtataatag cactTTaaat caggtgtttg ctatTTTTcc aaccctggt tcccgccaa	12660
agcttcatga ttttcagcaa tggctaatac ctgtacattc ctccatattt tcctctgttg	12720
cagcttcttg tactctTTTT gttgtgctgt ggttgccgggt tccaatgcta cgtactgttt	12780
ttggtttccg ctgggttaggg gcaatTTTTc tttcgaactc atggtgaatt acacggtgtg	12840
tccaccttgc ctcacccgac aagcagccgc tgaggTcctt gaaccggga ggtctctttg	12900
gtgcaggata gggcatgacc gatgtgggga ggacgatcac gacgaactgg ggttcatggt	12960
tccgcctggc ctctccagcg aaagccactt gaccagtgtt tacgcctggt tggcgTcct	13020
gtccttcagc tacacggccc agttccatcc cgagatattt gggataggga acgtgagtga	13080
agtttatgtt gacatcaagc accaattcat ctgcgccgtt catgacgggc agaaccac	13140
cttgctcgc catgacaata tttcagccgt atttcagacc tactatcaac atcaggTga	13200
cggcggcaat tggtttcacc tagaatggct gcgtcccttc ttttctctt ggttggtttt	13260
aaatgtttcg tggtttctca ggcgttcgcc tgcaagccat gtttcagttc gagtctttca	13320
gacatcaaaa ccaacactac cgcagcatca ggctttgttg tcctccagga catcagctgc	13380
cttaggcatg gcgactcgtc ctttccgacg attcgcaaaa gctctcaatg ccgcacggcg	13440
atagggacac ctgtgtatat caccatcaca gccaatgtga cagatgagaa ttacttacat	13500
tcttctgac tcctcatgct ttcttcttgc cttttctatg cttctgagat gagtgaag	13560
ggattcaagg tggatTTtg caatgtgtca ggcacgtgg ctgtgtgtgt caactttacc	13620

agctacgtcc aacatgtcaa agagtttacc caacgctcct tggtaggtcga tcatgtgcgg	13680
ctgcttcatt tcatgacacc tgagaccatg aggtgggcaa ccgttttagc ctgtcttttt	13740
gccatcctac tggcaatttg aatgttcaag tatgttgggg aaatgcttga ccgcgggctg	13800
ttgctcgcga ttgctttctt tgtggtgtat cgtgccgttc tggtttgctg tgctcggcaa	13860
cgccaacagc agcagcagct ctcatTTTTca gttgatttat aacttgacgc tatgtgagct	13920
gaatggcaca gattggctgg cagaaaaatt tgattgggca gtggagactt ttgtcatctt	13980
tcccggttg actcacattg tttcctatgg tgcactcacc accagccatt tccttgacac	14040
agttgggtctg gttactgtgt ccaccgccgg gttttatcac gggcggtatg tcttgagtag	14100
catctacgcg gtctgtgctc tggctgcgtt gatttgcttc gttattaggc ttgcgaagaa	14160
ctgcatgtcc tggcgctact cttgtaccag atataccaac ttccttctgg aactaaggg	14220
cagactctat cgttggcggg cgcccgttat catagaaaaa gggggtaagg ttgaggtcga	14280
aggtcatctg atcgacctca aaagagttgt gcttgatggg tccgtggcaa cccctttaac	14340
cagagtttca gcggaacaat ggggtcgtct ctagacgact ttgccatga tagcacggct	14400
ccacaaaagg tgcttttggc gttttccatt acctacacgc cagtaatgat atatgctcta	14460
aaggtaagtc gcggccgact gctagggctt ctgcaccttt tgatctttct gaattgtgct	14520
tttaccttcg ggtacatgac attcgcgcac tttcagagca caaatagggg cgcgctcgct	14580
atgggagcag tagttgcact tctttggggg gtgtactcag ccatagaaac ctggaaattc	14640
atcacctcca gatgccgttt gtgcttgcta ggccgcaagt acattctggc cctgcccac	14700
cacgtcga aa gtgccgcggg ctttcatccg attgcggcaa atgataacca cgcatttgc	14760
gtccggcgtc ccggtccat tacggttaac ggcacattgg tgcccggtt gaaaagcctc	14820
gtgttgggtg gcagaaaagc tgttaaacag ggagtggtaa accttgtaa atatgcaaaa	14880
taacaacggc aagcagcaaa agaaaaagaa ggggaatggc cagccagtca accagctgtg	14940
ccagatgctg ggtaaaatca tcgccagca aaaccagtcc agaggcaagg gaccgggcaa	15000
gaaaagtaag aagaaaaacc cggagaagcc ccattttcct ctagegaccg aagatgacgt	15060
caggcatcac ttcacccctg gtgagcggca attgtgtctg tcgtcgatcc agactgcctt	15120
taaccagggc gctggaactt gtaccctgtc agattcaggg aggataagtt aactgtgga	15180
gtttagtttg ccgacgcatc atactgtgcg cctgatccgc gtcacagcat caccctcagc	15240
atgatgggct ggcattcttt aggcacctca gtgtcagaat tggaagaatg tgtggtggat	15300
ggcactgatt gacattgtgc ctctaagtca cctattcaat tagggcgacc gtgtgggggt	15360
aaaatttaat tggcgagaac catgcggccg caattaaaaa aaaaaaaaaa aaaaaaaaaa	15420
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	15450

<210> 4
 <211> 15444
 <212> ADN
 <213> Trình tự nhân tạo

<220>

<223> Vacxin chứa hệ gen hoàn chỉnh P129-PK-d43/44 lần cấy truyền 34
 (15,444 nt xóa 2aa

<400> 4
 atgacgtata ggtgttggt ctatgccacg gcatttgtat tgtcaggagc tgtgaccatt 60
 ggcacagccc aaaacttgct gcacggaaaa cgcccttctg tgacagcctt cttcagggga 120
 gcttaggggt ctgtccctag caccttgctt ctggagttgc actgctttac ggtctctcca 180
 cccctttaac catgtctggg atacttgatc ggtgcacgtg caccaccaat gccagggtgt 240
 ttatggcgga gggccaagtc tactgcacac gatgtctcag tgcacggctc ctccttctc 300
 tgaatctcca agttctgag cttgggggtgc tgggcctatt ttataggccc gaagagccac 360
 tccggtggac gttgccacgt gcattcccca ctgtcgagtg ctcctctgcc ggggcctgct 420
 ggctttctgc gatctttcca attgcacgaa tgaccagtgg aaacctgaac tttcaacaaa 480
 gaatggtgcg ggttgacgct gagatttaca gagccggcca actcaccctt gcagttttga 540
 aggctctaca agtttatgaa cggggttgtc gctggtaccc cattgtcgga cctgtccctg 600
 gagtggccgt ttacgccaac tccctacatg tgagtgacaa acctttcccg ggagcaactc 660
 atgtgttaac caacctaccg ctcccgaga ggccaagcc tgaagacttt tgcccttttg 720
 agtgtgctat ggctaacgtc tatgacattg gccataacgc cgtcatgtat gtggccagaa 780
 ggaaagtctc ctgggcccct cgtggcgggg atgaagtga atttgaaacc gtccccgaag 840
 agttgaagtt gattgcgaac cgactccaca tctccttccc gcccaccac gcagtggaca 900
 tgtctgagtt tgccttcata gccctggga gtggtgtctc cttgcgggtc gagcaccaac 960
 acggctgcct tcccgctgat actgtccctg atgggaactg ctggtgttac ttgtttgact 1020
 tgctcccacc ggaagttcag aataaagaaa ttgcgcgtgc taaccaatth ggctatcaaa 1080
 ccaagcatgg tgtccatggc aagtacctac agcggaggct gcaagttaat ggtctccgag 1140
 cagtgaactga tacagatgga cctattgtcg tacagtactt ctctgttagg gagagttgga 1200
 tccgccactt cagactggcg gaagaacctg gcctccctgg gtttgaagac ctcctcagaa 1260
 taagggtaga gcctaatacg tcgccaatgg gtggcaaggg tgaaaaaatc ttccggtttg 1320
 gcagtcacaa gtggtacggt gctggaaaga gagcaaggag agcacgctct ggtgcgactg 1380
 ccacggtcgc tcaccgcgct ttgcccgtc gcgaagccca gcaggccaag aagctcgagg 1440
 ttgccagcgc caacagggct gagcatctca agtactattc cccgcctgcc gacgggaact 1500
 gtggttgga ctgcatttcc gccattacca accgatgggt gaattccaaa tttgaaacca 1560
 ctcttcccga gagagtgaga ccttcagatg actgggctac tgacaggat cttgtgaata 1620

ccatccaaat cctcaggctc cccgcggcct tggacaggaa cggtgcttgt gctggcgcca	1680
agtacgtgct caagctggaa ggtgagcact ggaccgtctc tgtgaccctt gggatgaccc	1740
cttcttttgc ccccttgaa tgtgttcagg gttgttgatga gcataagagc ggtcttggtt	1800
tcccagacgt ggtcgaagtt tccggatttg accctgcctg tcttgaccga cttgctgaga	1860
taatgcactt gcctagcagt gtcaccccag ctgctctggc cgagatgtcc gacgacttca	1920
atcgtctggc ttccccggcc gccactgtgt ggactgtttc gcaattcttt gcccgccaca	1980
gaggaggaga gcaccttgac cagggtgtgt tagggaaaat tatcaacctt tgtcagggtga	2040
ttgaggaatg ctgctgttcc cggaacaaag ccaaccgggc taccocggaa gaggttgccg	2100
caaaagttga ccagtacctc cgtggtgcag caagccttgg agaattgctt gccaaagctt	2160
agagggctcg cccgcggagc gcgacggaca cctcctttga ttggaatgtt gtgcttctcg	2220
gggttgagac ggcgaatcag acaaccaaac agctccatgt caaccagtgc cgcgctctgg	2280
ttcctgtcgt gactcaagag cctttggaca gagactcggg cctctgacc gccttctcgc	2340
tgtccaattg ctactacct gcacaagggt acgaggtccg tcaccgtgag aggctaaact	2400
ccttgctctc taagttggag ggggttggtc gtgaggaata tgggctcacg ccaactggac	2460
ctggcccgcg acccgactg ccgaacgggc tcgacgagct taaagaccag atggaggagg	2520
atctgctgaa attagtcaac gcccaggcaa cttcagaaat gatggcctgg gcagccgagc	2580
aggttgatct aaaagcttgg gtcaaaaatt acccacggtg gacaccgcca cccctccac	2640
caagagttca gcctcgaaaa acgaagtctg tcaagagctt gctagagaac aagcctgtcc	2700
ctgctccgcg caggaaggtc agatctgatt gtggcagccc gattttgatg ggcgacaatg	2760
ttcctaacgg ttgggaagat tcgactgttg gtggccccct tgatctttcg gcaccatccg	2820
agccgatgac acctctgagt gagcctgtac ttatttccag gccagtgaca tctttgagt	2880
tgccggcccc agttcctgca ccgcgtagag ctgtgtcccc accgatgacg ccctcgagt	2940
agccaatttt tgtgtctgca ctgcgacaca aatttcagca ggtggaaaaa gcaaactctg	3000
cggcagcagc gccgatgtgc caggacgaac ccttagattt gtctgcatcc tcacagactg	3060
aatatgaggc ttcccccta acaccaccgc agaacgtggg cattctggag gtaagggggc	3120
aagaagctga ggaagtctg agtgaaatct cggatattct gaatgatacc aaccctgcac	3180
ctgtgtcatc aagcagctcc ctgtcaagtg ttaagatcac acgccccaaa tactcagctc	3240
aagccattat cgactcgggc gggccctgca gtgggcacct ccaaaggga aaagaagcat	3300
gcctccgat catgcgtgag gcttgatgag cggccaagct tagtgaccct gccacgcagg	3360
aatggctttc tcgcatgtgg gataggggtg acatgctgac ttggcgcaac acgtctgctt	3420
accaggcgtt tcgcacctta gatggcaggt ttgggtttct ccaaagatg atactcgaga	3480
cgccgcggcc ctaccogtgt gggtttgatga tgttgcctca caccocgca ccttccgtga	3540

gtgcagagag cgaccttacc attggttcag tcgccactga agatattcca cgcgtcctcg	3600
ggaaaataga aaataccggt gagatgatca accaggggacc cttggcatcc tctgaggaag	3660
aaccggtata caaccaacct gccaaagact cccggatatac gtgcgggggg tctgacgaga	3720
gcacagcagc cccgtccgcg ggtacagggtg gcgcccggctt atttactgat ttgccacctt	3780
cagacggcgt agatgcggac ggtggggggc cgttgcagac ggtaagaaag aaagctgaaa	3840
ggctcttcga ccaattgagc cgtcaggttt ttaacctcgt ctcccatctc cctgttttct	3900
tctcacacct cttcaaactc gacagtgggtt attctccggg tgattggggg tttgcagctt	3960
ttactctatt ttgcctcttt ttgtgttaca gctaccatt cttcggtttc gttcccctct	4020
tgggtgtatt ttctgggtct tctcggcgtg tgogcatggg ggtttttggc tgctggctgg	4080
cttttgctgt tggcctgttc aagcctgtgt ccgaccagc cggcactgct tgtgagtttg	4140
actcgccaga gtgcaggaaac gtcccttcat cttttgagct tctcaaacct tgggacctg	4200
ttcgcagcct tgttgtgggc cccgtcggtc tcggctcttg cttctttggc aggttactgg	4260
gcggggcacg ctacatctgg ctttttttgc ttaggcttgg cattgttgca gattgtatct	4320
tggctggagc ttatgtgctt tctcaaggta ggtgtaaaaa gtgctgggga tcttgataa	4380
gaactgctcc taatgaaatc gccttcaacg tgttcccttt tacacgtgcg accaggctgt	4440
cactcatcga cctgtgcgat cggttttgtg cgccaacagg catggacccc attttctctg	4500
ccactgggtg gcgtgggtgc tggaccggcc gaagtcccat tgagcaacct tctgaaaaac	4560
ccatcgcggt cgcccagttg gatgaaaaga ggattacggc tagaactgtg gtcgctcagc	4620
cttatgatcc taatcaagcc gtgaagtgtc tgcgggtgtt acaggcgggt ggggcgatgg	4680
tggccgaggc agtcccaaaa gtggccaaag tttctgctat tccattccga gcccctttt	4740
ttcccaccgg agtgaaagt gatcccgagt gcaggatcgt ggttgacccc gatactttta	4800
ctacagccct ccggtctggt tactctacca caaacctcgt ccttggtgtg ggggactttg	4860
cccagctgaa tggactaaag atcaggcaaa tttccaagcc ttcgggagga ggcccacacc	4920
tcattgctgc cctgcatgtt gcctgctcga tggcgttgca catgcttgtc ggggtttatg	4980
taacttcagt ggggtcttgc ggtgccggca ccaacgatcc atggtgcact aatccgtttg	5040
ccgttcctgg ctacggacca ggctctctct gcacgtccag attgtgcac tcccaacatg	5100
gccttacct gcccttgaca gcaattgtgg cgggattcgg tcttcaggaa atcgccttgg	5160
tcgttttgat tttcgtttcc atcggaggca tggctcatag gttgagttgt aaggctgata	5220
tgctgtgcat tttacttgca atcgccagct atgtttgggt accccttacc tggttgcttt	5280
gtgtgtttcc ttgttggttg cgtggttct ctttgacccc ccttaccatc ctatggttgg	5340
tgtttttctt gatttctgta aatatgcctt cgggaatctt ggccgtgggtg ttattggttt	5400
ctctttggct tttgggacgt tatactaaca ttgctgggtc tgtcaccccc tatgatattc	5460

atcattacac cagtggcccc cgcggtgttg ccgccttggc taccgcacca gatggaacct	5520
acttggtgctgc cgtccgcccgc gctgcgttga ctggtcgcac catgctgttc acccgtctc	5580
agcttggggtc ccttcttgag ggcgctttca gaactcgaaa gccctcactg aacaccgtca	5640
atgtgggttg gtcctccatg ggctctggtg gagtgttcac catcgacggg aaaattaggt	5700
gcgtgactgc cgcacatgtc cttacgagta attcggctag ggtttccgga gtcggcttca	5760
atcaaagtgt tgactttgat gtgaaagggg acttcgcoat agctgattgc ccgaattggc	5820
aaggagctgc tcccaagacc caattctgcg aggacggatg gactggccgt gcctattggc	5880
tgacatcctc tggcgctgaa ccggtgtta ttgggaatgg attcgccttc tgcttcaccg	5940
cgtgcggcga ttccgggtcc ccagtgatca ccgaagctgg tgagattgtc ggcgttcaca	6000
caggatcaaa taaacaagga ggtggcatcg tcacgcgcc ttcaggccag ttttgtaacg	6060
tggcaccat caagctgagc gaattaagtg aattctttgc tggacccaag gtcccgtcgc	6120
gtgatgtgaa ggttggcagc cacataatta aagacacgtg cgaagtacct tcagatcttt	6180
gcgccttgct tgctgcaaaa cctgaactgg agggaggcct ctccaccgtc caacttctgt	6240
gtgtgttttt cctactgtgg agaattgatg gacatgcctg gacgcccttg gttgctgtgg	6300
ggtttttcat tctgaatgag gttctcccag ctgtcctggg tcggagtgtt ttctcctttg	6360
ggatgtttgt gctatcttgg ctacacccat ggtctgcgca agttctgatg atcaggcttc	6420
taacagcagc tcttaacagg aacagatggt cacttgcctt ttacagcctt ggtgcggtga	6480
ccggttttgt cgcagatctt gcgtaactc aagggcaccc gttgcaggca gtgatgaatt	6540
tgagcaccta tgccttctcg cctcggatga tggttgtgac ctaccagtc ccagtgattg	6600
cgtgtggtgt tgtgcaccta cttgccatca ttttgtactt gttcaagtac cgcggcctgc	6660
acaatgttct tgttggtgat ggagcgtttt ctgcagcttt cttcttgca tactttgccg	6720
agggaaagtt gaggaaggg gtgtcgcaat cctgcggaat gaatcatgag tcattaactg	6780
gtgccctcgc tatgggactc aatgacgagg acttggactt ccttacgaaa tggactgatt	6840
ttaagtgtt tgtttctgcg tccaacatga ggaatgcagc aggccaattc atcgaggctg	6900
cctatgcaaa agcacttaga attgaacttg ccagtttgt gcaggttgat aaggttcgag	6960
gtactttggc caagcttgag gcttttctg ataccgtggc accccaactc tcgcccgtg	7020
acattgttgt tgctcttggc cacacgcctg ttggcagcat cttcgaccta aaggttggtg	7080
gtaccaagca tactctccaa gtcattgaga ccagagtcct tgccgggtcc aaaatgaccg	7140
tggcgcgcgt cgttgacca acccccacgc cccacccgc acccgtgcc atccccctcc	7200
caccgaaagt tctagagaat ggtcccaacg cctgggggga tggggaccgt ttgaataaga	7260
agaagaggcg taggatggaa accgtcggca tctttgtcat ggggtgggaag aagtaccaga	7320
aattttggga caagaattcc ggtgatgtgt ttacgagga ggtccatgac aacacagatg	7380

cgtgggagtg cctcagagtt ggtgaccctg cgcactttaa ccctgagaag ggaactctgt	7440
gtgggcatac tactattgaa gataaggatt acaaagtcta cgcctcccca totggcaaga	7500
agttcctggt ccccgctcaac ccagagagcg gaagagccca atgggaagct gcaaagcttt	7560
ccgtggagca ggcccttggc atgatgaatg tcgacggtga actgacggcc aaagaagtgg	7620
agaaactgaa aagaataatt gacaaacttc agggcctgac taaggagcag tgtttaaact	7680
gctagccgcc agcggttga cccgctgtgg tcgcggcggc ttggttgta ctgagacagc	7740
ggtaaaaata gtcaaatttc acaaccggac ttccacctta gggcctgtga atttaaaagt	7800
ggccagtgag gttgagctga aagacgcggt cgagcacaac caacacccgg ttgcaagacc	7860
ggttgacggt ggtgttggtc tcctgcgttc cgcagttcct tcgcttatag atgtcctgat	7920
ctccggtgct gacgcatttc ctaagttact cgctcgtcac gggccgggga aactgggat	7980
cgatggcacg ctttgggact ttgaggccga ggccaccaa gaggaattg cgctcagtgc	8040
gcaaataata caggcttgtg acattaggcg cggtgacgca cctgaaattg gtctccctta	8100
caagctgtac cctgttaggg gcaaccctga gcgggtaaaa ggagttttac agaatacaag	8160
gtttggagac ataccttaca aaacccccag tgacactggg agcccagtgc acgcggctgc	8220
ctgcctcacg cccaatgcca ctccggtgac tgatgggcgc tccgtcttgg ctactaccat	8280
gccctccggt tttgaattgt atgtaccgac cattccagcg tctgtccttg attatcttga	8340
ctctaggcct gactgcccc aacagttgac agagcacggc tgtgaggatg ccgcattgag	8400
agacctctcc aagtatgact tgtccacca aggctttgtt ttacctgggg ttcttcgcct	8460
tgtgcgtaag tacctgtttg cccacgtggg taagtgcccg cccgttcac ggccctccac	8520
ttaccctgcc aagaattcta tggctggaat aaatgggaac aggtttccaa ccaaggacat	8580
tcagagcgtc ccgaaatcg acgttctgtg cgcacaggcc gtgcgagaaa actggcaaac	8640
tgttaccocct tgtaccctca agaaacagta ttgtgggaag aagaagacta ggacaatact	8700
cggcaccaat aatttcattg cgttggccca ccgggcagcg ttgagtgggtg tcaccagggt	8760
cttcatgaaa aaggcgttta actcgcccat cgccctcggg aaaaacaaat ttaaggagct	8820
acagactccg atcttaggca ggtgccttga agctgatctt gcacccgtgtg atcgatccac	8880
acctgcaatt gtccgctggt ttgccgcaa ccttctttat gaacttgcct gtgctgaaga	8940
gcacctaccg tcgtacgtgc tgaactgtg ccatgacctt ttggtcacgc agtccggcgc	9000
agtgactaag aggggtggcc tgtcgtctgg cgaccgcac acttctgtgt ctaacaccat	9060
ttacagcttg gtgatatatg cacagcacat ggtgcttagt tactttaaaa gtggtcaccc	9120
tcatggcctt ctgttctac aagaccagct gaagtccgag gacatgctca aagtccaacc	9180
cctgatcgtc tattcggacg acctcgtgct gtatgccgaa tctcccacca tgccgaacta	9240
ccactggtgg gtcgaacatc tgaatttgat gctgggtttt cagacggacc caaagaagac	9300

agccataacg gactcgccat cattttctagg ctgtaggata ataaatggac gccagctagt	9360
ccccaaaccgt gacaggatcc tcgcggccct cgcttaccat atgaaggcaa gcaatgtttc	9420
tgaatactac gccgcggcgg ctgcaatact catggacagc tgtgcttggt tagagtatga	9480
tcctgaatgg tttgaagagc ttgtggttgg gatagcgcag tgcgcccgcga aggacggcta	9540
cagctttccc ggcccgcctg tcttcttgtc catgtgggaa aaactcagat ccaatcatga	9600
ggggaagaag tccagaatgt gcgggtattg cggggccccc gctccgtacg ccaactgcctg	9660
tggcctcgac gtctgtatct accacaccca cttccaccag cattgtccag tcacaatctg	9720
gtgtggccac ccggctgggt ctggttcttg tagtgagtgc aaaccccccc tagggaaagg	9780
cacaagccct ctagatgagg tgttagaaca agtcccgtat aagcctccac ggactgtaat	9840
catgcatgtg gagcagggtc tcacccctct tgaccaggc agataccaga ctgcgcgcgg	9900
attagtctcc gttaggcgtg gcatcagagg aatgaagtt gacctaccag acggtgatta	9960
tgctagcacc gccctactcc ccacttgtaa agagatcaac atggtcgctg tcgcctctaa	10020
tgtgttgccg agcaggttca tcatcgggtc gcccggtgct gggaaaacat actggctcct	10080
tcagcaggtc caggatggtg atgtcattta cacaccgact caccagacca tgctcgacat	10140
gattagggct ttggggacgt gccggttcaa cgtcccagca ggtacaacgc tgcaattccc	10200
tgccccctcc cgtaccggcc cgtgggttcg catcctagcc ggcggttggt gtccctggtaa	10260
gaattccttc ctggatgaag cagcgtattg taatcacctt gatgtcttga ggctccttag	10320
caaaaccacc ctcacctgtc tgggagactt caaacaactc caccagtggt gttttgatcc	10380
tcattgctat gtttttgaca tcatgcctca gaccagttg aagaccatct ggagattcgg	10440
acagaacatc tgtgatgcca tccaaccaga ttacaggagc aaacttgtgt ccatggtcaa	10500
cacaaccctg gtaacctacg tggaaaaacc tgtcaagtat gggcaagtcc tcaccctta	10560
ccacaggagc cgagaggagc gcgccatcac aattgactcc agtcaaggcg ccacatttga	10620
tgtggtcaca ctgcatttgc ccactaaaga ttactcaac aggcaaagag cccttggtgc	10680
tatcaccagg gcaagacatg ctatctttgt gtatgacca cacaggcaat tgcagagcat	10740
gtttgatctt cctgcgaagg gcacaccctg caacctcgca gtgcaccgtg atgagcagct	10800
gatcgtactg gatagaaata ataaagaatg cacagttgct caggctctag gcaacggaga	10860
taaatttagg gccaccgaca agcgcgttgt agattctctc cgcgccattt gtgctgatct	10920
ggaagggtcg agctctccgc tccccaaagg cgcacacaac ttgggatttt atttctcacc	10980
tgatttgaca cagtttgcta aactcccggg agaccttgca cccactggc ccgtggtgac	11040
aaccagaac aatgaaaagt ggccggatcg gctgggtgcc agccttcgcc ctgtccataa	11100
gtatagccgt gcgtgcattg gtgccggcta tatggtgggc cctcgggtgt ttctaggcac	11160
ccctggggtc gtgtcatact acctcacaaa atttgtcaag ggcgaggctc aagtgcctcc	11220

ggagacagtc ttcagcaccg gccgaattga ggtggattgc cgggagtatc ttgatgacag	11280
ggagcgagaa gttgctgagt cctccccaca tgccttcatt ggcgacgtca aaggcaccac	11340
cgttggggga tgtcatcatg tcacctccaa ataccttccg cgcttccttc ccaaggaatc	11400
agtcgcggtg gtcgggggtt cgagccccgg gaaagccgca aaagcagtgt gcacattgac	11460
ggatgtgtac ctcccagacc ttgaggccta cctccacca gagactcagt ccaagtgtctg	11520
gaaagttatg ttggacttca aggaagtctg actgatggtc tggaaagaca agacggccta	11580
tttccaactt gaaggccgct atttcacctg gtatcagctt gcaagctacg cctcgtacat	11640
ccgtgttcct gtcaactcca cggtgtatct ggaccctgc atgggacctg ccctttgcaa	11700
cagaagagtt gtcgggtcca ccattgggg agctgacctc gcagtcaccc cttatgatta	11760
cggtgctaaa atcatcttgt ctagcgctta ccatggtgaa atgcctcctg gatacaagat	11820
tctggcgtgc gcggagtctt cgctcgacga ccagtcgaag taaaaacaca cctgggggtt	11880
tgaatcggat acagcgtatc tgtatgagtt caccggaaac ggtgaggact gggaggatta	11940
caatgatgcg tttcgtgcgc gccagaaagg gaaaatttat aaggccactg ctaccagcat	12000
gaagttttat tttcccccg gcccgctcat tgaaccaact ttaggcctga attgaaatga	12060
aatggggctt atacaaagcc tcttcgacaa aattggccag ctttttgtgg atgctttcac	12120
ggaatTTTTg gtgtccattg ttgatatcat catatTTTTg gccatTTTgt ttggcttcac	12180
catgcgcggt tggctgggtg tcttttgcac cagattgggt tgctccgcgg tattccgtgc	12240
gcgccctgcc attcaccctg agcaattaca gaagatccta tgaggccttt ctttctcagt	12300
gccgggtgga cattcccacc tggggggtaa aacacccttt ggggatgttt tggcaccata	12360
aggtgtcaac cctgattgat gaaatggtgt cgcgtcgaat gtaccgcac atggaaaaag	12420
cagggcaagc tgcctgaaa caggtggtga gcgaggctac gctgtctgc attagtagtt	12480
tggatgtggt ggctcatTTT caacatcttg ccgccattga agccgagacc tgtaaatatt	12540
tggcttctcg actgcccatt ctacacaacc tgcgcatgac aggttcaa atgtaaccatag	12600
tgtataatag cactttaaat caggtgtttg ctatTTTTcc aaccctggt tccggccaa	12660
agcttcatga ttttcagcaa tggctaata ctgtacattc ctccatattt tcctctgttg	12720
cagcttcttg tactctTTTT gttgtgctgt ggttgccgggt tccaatgcta cgtactgttt	12780
ttggtttccg ctgggttaggg gcaatTTTTt tttcgaactc atggtgaatt acacggtgtg	12840
tccaccttgc ctccccgac aagcagccgc tgaggctcct gaaccggga ggtctctttg	12900
gtgcaggata gggcatgacc gatgtgggga ggacgatcac gacgaactgg ggttcatggt	12960
tccgctggc ctctccagcg aaagccactt gaccagtgtt tacgcctggt tggcgcttct	13020
gtccttcagc tacacggccc agttccatcc cgagatattt gggataggga acgtgagtga	13080
agtttatggt gacatcaagc accaattcat ctgcgccgtt catgacgggc agaaccac	13140

cttgccctcgc catgacaata tttcagccgt atttcagacc tactatcaac atcaggtcga	13200
cggcggcaat tggtttcacc tagaatggct gcgtcccttc ttttcctctt ggttggtttt	13260
aaatgtttcg tggttttctca ggcgttcgcc tgcaagccat gtttcagttc gagtctttca	13320
gacatcaaaa ccaacactac cgcagcatca ggctttgttg tcctccagga catcagctgc	13380
cttaggcatg gcgactcgtc ctttccgacg attcgcaaaa gctctcaatg ccgcacggcg	13440
atagggacac ctgtgtatat caccatcaca gccaatgtga cagatgagaa ttacttacat	13500
tcttctgac tcctcatgct ttcttcttgc cttttctatg cttctgagat gagtgaaaag	13560
ggattcaagg tggatatttg caatgtgtca ggcacgtgg ctgtgtgtgt caactttacc	13620
agctacgtcc aacatgtcaa agagtttacc caacgctcct tggtggtcga tcatgtgcgg	13680
ctgcttcatt tcatgacacc tgagaccatg aggtgggcaa ccgttttagc ctgtcttttt	13740
gccatcctac tggcaatttg aatgttcaag tatgttgggg aaatgcttga ccgcgggctg	13800
ttgctcgcga ttgctttctt tgtggtgtat cgtgccgttc tggtttgctg tgctcggcaa	13860
cgccaacagc agcagcagct ctcattttca gttgatttat aacttgacgc tatgtgagct	13920
gaatggcaca gattggctgg caggaaaatt tgattgggca gtggagactt ttgtcatctt	13980
tcccggttg actcacattg tttcctatgg tgcactcacc accagccatt tccttgacac	14040
agttggctctg gttactgtgt ccaccgccgg gttttatcac gggcggtatg tcttgagtag	14100
catctacgcg gtctgtgctc tggctgcgtt gatttgcttc gttattaggc ttgcgaagaa	14160
ctgcatgtcc tggcgctact cttgtaccag atataccaac ttccttctgg aactaaggg	14220
cagactctat cgttggcggg cgcccgttat catagaaaaa gggggtaagg ttgaggtcga	14280
aggtcatctg atcgacctca aaagagttgt gcttgatggg tccgtggcaa cccctttaac	14340
cagagtttca gcggaacaat ggggtcgtct ctagacgact tttgcatga tagcacggct	14400
ccacaaaagg tgcttttggc gttttccatt acctacacgc cagtaatgat atatgctcta	14460
aaggtaagtc gcggccgact gctagggctt ctgcaccttt tgatctttct gaattgtgct	14520
tttaccttcg ggtacatgac attcgcgcac tttcagagca caaatagggg cgcgctcgct	14580
atgggagcag tagttgcact tctttggggg gtgtactcag ccatagaaac ctggaaattc	14640
atcacctcca gatgccgttt gtgcttgcta ggccgcaagt acattctggc cctgcccac	14700
cacgtcgaaa gtgccgcggg ctttcatccg attgcggcaa atgataacca cgcatttgtc	14760
gtccggcgtc ccggctccat tacggttaac ggcacattgg tgcccgggtt gaaaagcctc	14820
gtgttgggtg gcagaaaagc tgttaaacgg ggagtggtaa accttgtaa atatgccaaa	14880
taacaacggc aagcagcaaa agaaaaagaa ggggaatggc cagccagtca accagctgtg	14940
ccagatgctg ggtaaaatca tcgccagca aaaccagtcc agaggcaagg gaccgggctc	15000
taagaagaaa tccccggaga agccccattt tcctctagcg accgaagatg acgtcaggca	15060

tcacttcacc cctggtgagc ggcaattgtg tctgtcgtcg atccagactg cctttaacca 15120
 gggcgctgga acttgtaccc tgtcagattc agggaggata agttacactg tggagttag 15180
 tttgccgacg catcatactg tgcgcctgat ccgcgtcaca gcatcaccct cagcatgatg 15240
 ggctggcatt ctttaggcac ctacagtgtca gaattggaag aatgtgtggt ggatggcact 15300
 gattgacatt gtgcctctaa gtcacctatt caattagggc gaccgtgtgg gggtaaaatt 15360
 taattggcga gaacatgcg gccgcaatta aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 15420
 aaaaaaaaaa aaaaaaaaaa aaaa 15444

<210> 5
 <211> 15450
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Virut ban đầu có hệ gen với chiều dài hoàn chỉnh P129 lần cấy truyền 0 (15,450 nt)

<400> 5
 atgacgtata ggtgttggct ctatgccacg gcatttgtat tgtcaggagc tgtgaccatt 60
 ggcacagccc aaaacttgct gcacggaaaa cgcccttctg tgacagcctt cttcagggga 120
 gcttaggggt ctgtccctag caccttgctt ctggagttgc actgctttac ggtctctcca 180
 cccctttaac catgtctggg atacttgatc ggtgcacgtg ccccccaat gccaggggtg 240
 ttatggcgga gggccaagtc tactgcacac gatgtctcag tgcacggtct ctcttctctc 300
 tgaatctcca agttcctgag cttgggggtgc tgggcctatt ttataggccc gaagagccac 360
 tccggtggac gttgccacgt gcattcccca ctgtcgagtg ctccccgcc ggggcctgct 420
 ggctttctgc gatctttcca attgcacgaa tgaccagtgg aaacctgaac tttcaacaaa 480
 gaatggtgcg ggttgcagct gagatttaca gagccggcca actcaccctt gcagttttga 540
 aggtcttaca agtttatgaa cgggggtgtc gctggtaccc cattgtcgga cctgtccctg 600
 gagtggccgt tcacgccaac tccctacatg tgagtgacaa acctttcccg ggagcaactc 660
 atgtgttaac caacctaccg ctcccgaga ggccaagcc tgaagacttt tgcccttttg 720
 agtgtgctat ggctgacgtc tatgacatta gccatgacgc cgtcatgtat gtggccagag 780
 ggaaagtctc ctgggcccct cgtggcgggg atgaagtga atttgaaacc gtccccgaag 840
 agttgaagtt gattgcgaac cgactccaca tctccttccc gccccaccac gcagtggaca 900
 tgtctgagtt tgccctcata gccctggga gtggtgtctc cttgcgggtc gagcaccaac 960
 acggctgcct tccgctgat actgtccctg aagggaactg ctggtggtgc ttgtttgact 1020
 tgctccacc ggaagttcag aataaagaaa ttgcgcgtgc taaccaattt ggctatcaaa 1080
 ccaagcatgg tgtccctggc aagtacctac agcggaggct gcaagttaat ggtctccgag 1140
 cagtgactga tacagatgga cctattgtcg tacagtactt ctctgttagg gagagttgga 1200

tccgccactt cagactggcg gaagaaccta gcctccctgg gtttgaagac ctctcagaa	1260
taagggtaga gcctaatacg tcgccattgg gtggcaaggg tgaaaaaatc ttccggtttg	1320
gcagtcacaa gtggtacggt gctggaaaga gagcaaggag agcacgctct ggtgcgactg	1380
ccacggtcgc tcaccgcgct ttgcccgctc gcgaagccca gcaggccaag aagctcgagg	1440
ttgccagcgc caacagggct gagcatctca agtactattc cccgcctgcc gacgggaact	1500
gtggttggca ctgcatttcc gccattacca accggatggt gaattccaaa tttgaaacca	1560
ctcttccoga gagagtgaga ccttcagatg actgggctac tgacgaggat cttgtgaatg	1620
ccatccaaat cctcaggctc cccgcggcct tggacaggaa cggtgcttgt gctggcgcca	1680
agtacgtgct caagctggaa ggtgagcact ggaccgtctc tgtgacctct gggatgacct	1740
cttctttgct ccccttgaa tgtgttcagg gttgttgtga gcatacgagc ggccttggtt	1800
tcccagatgt ggtcgaagtt tccgatttg accctgcctg ccttgaccga cttgctgaga	1860
taatgcactt gcctagcagt gtcattccag ctgctctggc cgagatgtcc gacgacttca	1920
atcgtctggc ttccccggcc gccactgtgt ggactgttct gcaattcttt gcccgccaca	1980
gaggaggaga gcattctgac caggtgtgct tagggaaaat tatcaacctt tgtcaggtga	2040
ttgaggaatg ctgctgttcc cggaacaaag ccaaccgggc taccctggaa gaggttgcg	2100
caaaaattga ccagtacctc cgtggtgcag caagccttgg agaatgcttg gccaaagcttg	2160
agatggctcg cccgccgagc gcgatggaca cctcctttga ttggaatgtt gtgcttctg	2220
gggttgagac ggcgaatcag acaaccaaac agctccatgt caaccagtgc cgcgctctgg	2280
ttctgtcgt gactcaagag ctttggaca gagactcagt ccctctgacc gccttctcgc	2340
tgtccaattg ctactaccct gcacaagggt acgaggtccg tcaccgtgag aggctaaact	2400
ccgtgctctc taagttggag ggggttgctc gtgaggaata tgggctcacg ccaactggac	2460
ctggcccgcg acccgactg ccgaacgggc tcgacgagct taaagaccag atggaggagg	2520
atctgctgaa attagtcaac gccaggcaa cttcagaaat gatggcctgg gcagccgagc	2580
aggttgatct gaaagcttgg gtcaaaaatt acccacggtg gacaccgcca cccctccac	2640
caagagttca gcctcgaaaa acgaagtctg tcaagagctt gctagagaac aagcctgtcc	2700
ctgctccgcg caggaaggct agatctgatt gtggcagccc gattttgatg ggcgacaatg	2760
ttcctaacgg ttgggaagat tcgactgttg gtggtcccct tgacctttcg gcaccatccg	2820
agccgatgac acctctgagt gagcctgtac ttatttccag gccagtgaca tctttgagtg	2880
tgcgggcccc agttcctgca ccgcgtagag ctgtgtcccg accgatgacg ccctcgagtg	2940
agccaatttt tgtgtctgca ctgcgacaca aatttcagca ggtggaaaaa gcaaatctgg	3000
cggcagcagc gccgatgtac caggacgaac ccttagattt gtctgcatcc tcacagactg	3060
aatatgaggc ttctccccta acaccaccgc agaacgtggg cattctggag gtaagggggc	3120

aagaagctga ggaagttctg agtgaaatct cggatattct gaatgacacc aaccctgcac	3180
ctgtgtcatc aagcagctcc ctgtcaagtg ttaggatcac acgccccaaa tactcagctc	3240
aagccattat cgactcgggc gggccctgca gtgggcacct ccaaagggaa aaagaagcat	3300
gcctccgcat catgcgtgag gcttgtgatg cggccaagct tagtgaccct gccacgcagg	3360
aatggctttc tcgcatgtgg gataggggtg acatgctgac ttggcgcaac acgtctgctt	3420
accaggcggt tcgcacctta gatggcagggt ttgggtttct cccaaagatg atactcgaga	3480
cgccgcccgc ctaccctgtg gggtttgtga tgttgccca caccctgca cttccgtga	3540
gtgcagagag cgaccttacc atcggttcag tcgccactga agatattcca cgcacccctg	3600
ggaaaataga aaataccggt gagatgatca accagggacc cttggcatcc tctgaggaag	3660
aaccggtata caaccaacct gccaaagact ccgcatatc gtgcgggggg tctgacgaga	3720
gcacagcagc tccgtccgca ggtacagggt gcgcgggctt atttactgat ttgccacctt	3780
cagacggcgt agatgcggac ggtggggggc cgttgacagc ggtaagaaag aaagctgaaa	3840
ggctcttcga ccaattgagc cgtcagggtt ttaacctcgt ctccatctc cctgttttct	3900
tctcacacct cttcaaactc gacagtgggt attctccggg tgattgggggt tttgcagctt	3960
ttactctatt ttgcctcttt ttgtgttaca gctaccatt cttcggtttc gttccctct	4020
tgggtgtatt ttctgggtct tctcggcgtg tgcgcatggg ggtttttggc tgctggctgg	4080
cttttgctgt tggcctgttc aagcctgtgt ccgaccagc cggcactgct tgtgagtttg	4140
actcgccaga gtgtaggaac gtccttcatt cttttgagct tctcaaacct tgggaccctg	4200
ttcgcagcct tgttgtgggc ccgctcggtc tcggtcttgc cattcttggc aggttactgg	4260
gcggggcacg ctacatctgg cattttttgc ttaggcttgg cattgttgca gattgtatct	4320
tggctggagc ttatgtgctt tctcaaggta ggtgtaaaaa gtgctgggga tcttgataa	4380
gaactgctcc taatgaaatc gccttcaacg tgttcccttt tacacgtgcg accaggctgt	4440
cactcatcga cctgtgcgat cggttttgtg cgccaaaagg catggacccc atttccctcg	4500
ccactgggtg gcgtgggtgc tggaccggcc gaagtcccat tgagcaacct tctgaaaaac	4560
ccatcgcggt cgcccagttg gatgaaaaga ggattacggc tagaactgtg gtcgctcagc	4620
cttatgatcc taatcaagcc gtaaagtgtg tgcgggtgtt acaggcgggt ggggcgatgg	4680
tggccgaggc agtcccaaaa gtggtcaaag tttctgctat tccattccga gccccctttt	4740
ttcccaccgg agtgaaagtgt gatcccgagt gcaggatcgt gggtgacccc gatactttta	4800
ctacagccct ccggtctggt tactctacca caaacctcgt ccttggtgtg ggggactttg	4860
cccagctgaa tggactaaag atcaggcaaa tttccaagcc ttccgggagga ggcccacacc	4920
tcattgctgc cctgcatgtt gcctgctcga tggcgttgca catgcttgct ggggtttatg	4980
taacttcagt ggggtcttgc ggtgcgggca ccaacgatcc atggtgcact aatccgtttg	5040

ccgttcctgg ctacggacca ggctctctct gcacgtccag attgtgcac tcccaacatg 5100
 gccttacccct gcccttgaca gcacttgtag cgggattcgg tcttcaggaa atcgcccttg 5160
 tcgttttgat tttcgtttcc atcggaggca tggctcatag gttgagttgt aaggctgata 5220
 tgctgtgcat cttacttgca atcgccagct atgtttgggt accccttacc tggttgcttt 5280
 gtgtgtttcc ttgttggttg cgctggttct ctttgacccc cctcaccatc ctatggttgg 5340
 tgtttttctt gatttctgta aatatgcctt cggaatctt ggccgtggtg ttgttggttt 5400
 ctctttggct tttgggacgt tatactaaca ttgctgggtct tgtcaccccc tatgatattc 5460
 atcattacac cagtggcccc cgcggtggtg ccgccttagc taccgcacca gatggaacct 5520
 acttggtgc cgtccgcgc gctgcgttga ctggtcgcac catgctgttc acccgtctc 5580
 agcttgggtc cttcttgag ggcgtttca gaactcgaaa gccctcactg aacaccgtca 5640
 atgtggttg gtcctccatg ggctctggtg gagtgttcac catcgacggg aaaattaggt 5700
 gcgtgactgc cgcacatgtc cttacgggta attcggttag ggtttccgga gtcggcttca 5760
 atcaaagct tgactttgat gtgaaagggg acttcgccat agctgattgc ccgaattggc 5820
 aaggagctgc tcccaagacc caattctgcg cggatggatg gactggcgt gcctattggc 5880
 tgacatctc tggcgctgaa cccggtgtca ttgggagtgg attcgcttc tgcctcaccg 5940
 cgtgcggcga ttccgggtcc ccagtgatca ccgaagctgg tgagcttgtc ggcgttcaca 6000
 caggatcaaa taaacaagga ggtggcatcg tcacgcgcc ttcaggccag ttttgtaacg 6060
 tggcacccat caagctgagc gaattaagtg aattctttgc tggacccaag gtcccgctcg 6120
 gtgatgtgaa ggttggcagc cacataatta aagacacgtg cgaagtacct tcagatcttt 6180
 gcgccttgct tgctgcaaaa cctgaactgg agggaggcct ctccaccgtc caacttctgt 6240
 gtgtgttttt cctactgtgg agaattgatg gacatgcctg gacgcccttg gttgctgtgg 6300
 ggtttttcat tctgaatgag gttctccag ctgtcctggt tcggagtgtt ttctcctttg 6360
 ggatgtttgt gctatcttg ctcacaccat ggtctgcgca agttctgatg atcaggettc 6420
 taacagcagc tcttaacagg aacagatggt cacttgctt ttacagcctt ggtgcggtga 6480
 ccggttttgt cgcagatctt gcggcaactc aagggcaccc gttgcaggca gtaatgaatt 6540
 tgagcaccta tgccttctg cctcgatga tggttgtgac ctcaccagtc ccagtgattg 6600
 cgtgtggtgt tgtgcaccta cttgccatca ttttgactt gtttaagtac cgtggcctgc 6660
 acaatgttct tgttggtgat ggagcgttt ctgcagctt cttcttgca tactttgccg 6720
 agggaaagt gaggaagg gtgtcgcaat cctgcggaat gaatcatgag tcattgactg 6780
 gtgccctgc tatgagactc aatgacgagg acttgactt cttacgaaa tggactgatt 6840
 ttaagtgctt tgtttctgcg tccaacatga ggaatgcagc aggccaattc atcgaggctg 6900
 cctatgcaaa agcacttaga attgaacttg ccagttggt gcaggttgat aaggttcgag 6960

gtacttttggc caagcttgag gcttttggctg ataccgtggc accccaactc tcgcccgggtg	7020
acattgttgt tgctcttggc catacgctg ttggcagcat cttcgaccta aaggttggtg	7080
gtaccaagca tactctccaa gtcattgaga ccagagtcct tgccgggtcc aaaatgaccg	7140
tggcgcgcggt cgttgaccca acccccacgc cccacccgc acccgtgccc atccccctcc	7200
caccgaaagt tctagagaat ggtcccaacg cctgggggga tggggaccgt ttgaataaga	7260
agaagaggcg taggatggaa accgtcggca tctttgtcat gggtggaag aagtaccaga	7320
aattttggga caagaattcc ggtgatgtgt ttacgagga ggtccatgac aacacagatg	7380
cgtgggagtg cctcagagtt ggtgaccctg ccgactttga ccctgagaag ggaactctgt	7440
gtgggcatac tactattgaa gataaggatt acaaagtcta cgcctcccca tctggcaaga	7500
agttcctggt ccccgtaac ccagagagcg gaagagccca atgggaagct gcaaagcttt	7560
ccgtggagca ggcccttggc atgatgaatg tcgacggtga actgacggcc aaagaagtgg	7620
agaaactgaa aagaataatt gacaaacttc agggcctgac taaggagcag tgtttaaact	7680
gctagccgcc agcggttga cccgctgtgg tcgcgcgggc ttggttgta ctgagacagc	7740
ggtaaaaata gtcaaatttc acaaccggac ttccacctta gggcctgtga atttaaaagt	7800
ggccagtgag gttgagctga aagacgcggt cgagcacaac caacaccggg ttgcaagacc	7860
ggttgacggt ggtgttgctc tcctgcgttc cgcagttcct tcgcttatag atgtcctgat	7920
ctccggtgct gacgcatctc ctaagttact cgctcgtcac gggccgggga aactgggat	7980
cgatggcacg ctttgggact ttgaggccga ggccaccaa gaggaaattg cactcagtgc	8040
gcaaataata caggcttgtg acattaggcg cggcgacgca cctgaaattg gtctccctta	8100
caagctgtac cctgttaggg gcaaccctga gcgggtaaaa ggagttttac agaatacaag	8160
gtttggagac ataccttaca aaacccccag tgacactgga agcccagtgc acgcggtgc	8220
ctgcctcacg cccaatgcca ctccggtgac tgatgggagc tccgtcttgg ctactaccat	8280
gccctcgggt tttgaattgt atgtaccgac cattccagcg tctgtccttg attatcttga	8340
ctctaggcct gactgcccc aacagttgac agagcacggc tgtgaggatg ccgcattgag	8400
agacctctcc aagtatgact tgtccacca aggctttgtt ttacctggg ttcttcgcct	8460
tgtgcgtaag tacctgtttg cccacgtggg taagtgccg cccgttcac ggccttcac	8520
ttacctgcc agaattcta tggctggaat aaatgggaac aggtttccaa ccaaggacat	8580
tcagagcgtc cctgaaatcg acgttctgtg cgcacaggcc gtgcgagaaa actggcaaac	8640
tgttaccct tgtaccctca agaaacagta ttgtgggaag aagaagacta ggacaatact	8700
cggcaccaat aatttcattg cgttggccca ccgggcagcg ttgagtggg tcacccaggg	8760
cttcataaaa aaggcgttta actcgcccat cgccctcggg aaaaacaaat ttaaggagct	8820
acagactccg gtcttaggca ggtgccttga agctgatctt gcatcctgtg atcgatccac	8880

acctgcaatt gtccgctggg ttgccgcaa tcttctttat gaacttgcct gtgctgaaga	8940
gcacctaccg tcgtacgtgc tgaactgctg ccatgaccta ttggtcacgc agtccggcgc	9000
agtgactaag aggggtggcc tgtcgtctgg cgacccgatc acttctgtgt ctaacaccat	9060
ttacagcttg gtgatatatg cacagcacat ggtgcttagt tactttaaaa gtggtcaccc	9120
tcattggcctt ctgttccctac aagaccagct gaagttcgaa gacatgctca aagtccaacc	9180
cctgatcgtc tattcggacg acctcgtgct gtatgccgaa tctcccacca tgccgaacta	9240
ccactgggtg gtcgaacatc tgaatttgat gctgggtttt cagacggacc caaagaagac	9300
agccataacg gactcgccat catttctagg ctgtaggata ataaatggac gccagctagt	9360
ccccaacgt gacaggatcc tcgcggccct cgcttaccat atgaaggcaa gcaatgtttc	9420
tgaatactac gccgcggcgg ctgcaatact catggacagc tgtgcttggt tagagtatga	9480
tctgaatgg tttgaagagc ttgtggttgg gatagcgcag tgcgcccga aggacggcta	9540
cagctttccc ggcccgccgt tcttcttctc catgtgggaa aaactcagat ccaatcatga	9600
ggggaagaag tccagaatgt gcgggtattg cggggccccc gctccgtacg ccactgcctg	9660
tggcctcgac gtctgtatct accacacca ctccaccag cattgtccag tcataatctg	9720
gtgtggccac ccggctggtt ctggttcttg tagtgagtgc aaaccccc tagggaaagg	9780
cacaagccct ctagatgagg tgttagaaca agtcccgtat aagcctccac ggactgtaat	9840
catgcatgtg gagcagggtc tcacccctct tgaccaggc agataccaga ctgcgcggc	9900
attagtctcc gttaggcgtg gcatcagagg aatgaagtt gacctaccag acggtgatta	9960
tgctagcacc gccctactcc ccacttgtaa agagatcaac atggtcgctg tcgcctctaa	10020
tgtgttgccg agcaggttca tcatcggctc gcccggtgct gggaaaacat actggctcct	10080
tcagcaggtc caggatggtg atgtcattta cacaccgact caccagacca tgctcgacat	10140
gattagggct ttggggacgt gccggttcaa cgtcccagca ggtacaacgc tgcaattccc	10200
tgccccctcc cgtaccggcc cgtgggttcg catcctagcc ggcggttggt gtcctggtaa	10260
gaattccttc ctggatgaag cagcgtattg taatcacctt gatgtcttga ggctccttag	10320
caaaaccacc ctcacctgtc tgggagactt caaacaactc caccagtggt gttttgatc	10380
tcattgctat gtttttgaca tcatgcctca gaccagttg aagaccatct ggagattcgg	10440
acagaacatc tgtgatgcca tccaaccaga ttacagggac aaacttgtgt ccatggtcaa	10500
cacaaccgt gtaacctacg tggaataacc tgtcaagtat gggcaagtcc tcaccctta	10560
ccacagggac cgagaggacg gcgccatcac aattgactcc agtcaaggcg ccacatttga	10620
tgtggttaca ctgcatttgc ccactaaaga ttcactcaac aggcaaagag cccttggtgc	10680
tatcaccagg gcaagacatg ctatctttgt gtatgacca cacaggcaat tgcagagcat	10740
gtttgatctt cctgcgaagg gcacaccgt caacctcgca gtgcacogtg atgagcagct	10800

gatcgtactg gatagaaata ataaagaatg cacagttgct caggctctag gcaacggaga	10860
taaatttagg gccaccgaca agcgcgttgt agattctctc cgcgccattt gtgctgatct	10920
ggaagggctg agctctccgc tccccaaagt cgcacacaac ttgggatttt atttctcacc	10980
tgatttgaca cagtttgcta aactcccggg agaccttgca cccactggc ccgtggtgac	11040
aacccagaac aatgaaaagt ggccggatcg gctggttgcc agccttcgcc ctgtccataa	11100
gtatagccgt gcgtgcattg gtgccggcta tatggtgggc ccctcgggtg ttctaggcac	11160
ccctgggggtc gtgtcatact acctcacaaa atttgtcaag ggcgaggctc aagtgttcc	11220
ggagacagtc ttcagcaccg gccgaattga ggtggattgc cgggggtatc ttgatgacag	11280
ggagcgagaa gttgctgagt ccctcccaca tgccttcatt ggcgacgtca aaggcaccac	11340
cgttggggga tgtcatcatg tcacctcaa ataccttcg cgttccttc ccaaggaatc	11400
agtcgcggta gtcggggttt cgagccccgg gaaagccgca aaagcagtgt gcacattgac	11460
ggatgtgtac ctcccagacc ttgaggccta cctccacca gagactcagt ctaagtgtg	11520
gaaagttatg ttggacttca aggaagttcg actgatggtc tggaaagaca agacggccta	11580
tttccaactt gaaggccgct atttcacctg gtatcagctt gcaagctacg cctcgtacat	11640
ccgtgttcct gtcaactcca cgggtgatct ggaccctgc atgggccctg ccctttgcaa	11700
cagaagagtt gtcgggtcca ccattgggg agctgacctc gcagtcaccc cttatgatta	11760
cggtgctaaa atcatcttgt ctagcgctta ccatggtgaa atgcctcctg gatacaagat	11820
tctggcgtgc gcggagttct cgctcgacga ccagtcgaag taaaacaca cctggggttt	11880
tgaatcggat acagcgtatc tgtatgagtt catcggaac ggtgaggact gggaggatta	11940
caatgatgcg tttcgtgcgc gccagaaagg gaaaatttat aaggccactg ctaccagcat	12000
gaagttttat tttcccccg gccccgtcat tgaaccaact ttaggcctga attgaaatga	12060
aatgggggtct atacaaagcc tcttcgacaa aattggccag ctttttgtgg atgctttcac	12120
ggaatttttg gtgtccattg ttgatatcat catatttttg gccattttgt ttggcttcac	12180
catcgccggt tggctggtgg tcttttgcac cagattgggt tgctccggcg tattccgtgc	12240
gcgccctgcc attcaccctg agcaattaca gaagatccta tgaggccttt ctttctcagt	12300
gccgggtgga cattcccacc tggggggtaa aacacccttt ggggatgttt tggcaccata	12360
aggtgtcaac cctgattgat gaaatggtgt cgcgtcgaat gtaccgcac atggaaaaag	12420
cagggcaagc tgcttgaaa caggtggtga gcgaggctac gctgtctcgc attagtagtt	12480
tggatgaggt ggctcatttt caacatcttg ccgccattga agccgagacc tgtaaataat	12540
tggcttctcg actgcccacg ctacacaacc tgcgcatgac aggttcaa atgtaaccatag	12600
tgtataatag cactttaaat caggtgtttg ctatttttcc aaccctggt tcccgccaa	12660
agcttcacga ttttcagcaa tggctaata ctgtacattc ctccatattt tcctctgttg	12720

cagcttcttg tactcttttt gttgtgctgt ggttgcggtg tccaatgcta cgtactgttt	12780
ttggtttccg ctggttagg gcaatttttc tttcgaactc atggtgaatt acacgggtgtg	12840
tccaccttgc ctcacccgac aagcagccgc tgaggccctt gaacccggta ggtctctttg	12900
gtgcaggata gggcatgacc gatgtgggga ggacgatcac gacgaactag ggttcatggt	12960
tccgcctggc ctctccagcg aaagccactt gaccagtgtt tacgcctggt tggcgttcct	13020
gtccttcagc tacacggccc agttccatcc cgagatattt gggatagggga acgtgagtga	13080
agtttatgtt gacatcaagc accaattcat ctgcgccgtt catgacgggc agaacaccac	13140
cttgcctcgc catgacaata tttcagccgt atttcagacc tactatcaac atcaggtcga	13200
cggcggaat tggtttcacc tagaatggct gcgtcccttc ttttcctctt ggttggtttt	13260
aaatgtttcg tggtttctca ggcgttcgcc tgcaagccat gtttcagttc gagtctttca	13320
gacatcaaaa ccaacactac cgcagcatca ggctttgttg tctccagga catcagctgc	13380
cttaggcatg gcgactcgtc ctttccgacg attcgcaaaa gctctcaatg ccgcacggcg	13440
atagggacac ctgtgtatat caccatcaca gccaatgtga cagatgagaa ttacttacat	13500
tcttctgac tctcatgct ttcttcttgc cttttctatg cttctgagat gagtgaag	13560
ggattcaagg tgggtgttg caatgtgtca ggcacgtgg ctgtgtgtgt caactttacc	13620
agctacgtcc aacatgtcaa agagtttacc caacgctcct tgggtggtcga tcatgtgcgg	13680
ctgcttcatt tcatgacacc tgagaccatg aggtgggcaa ccgttttagc ctgtcttttt	13740
gccatcctac tggcaatttg aatgttcaag tatgttgggg aaatgcttga ccgcgggctg	13800
ttgctcgcga ttgctttctt tgtggtgtat cgtgccgttc tggtttgctg tgctcgtcaa	13860
cgccaacaac accagcagct ctattttca gttgatttat aacttgacgc tatgtgagct	13920
gaatggcaca gattggctgg cagaaaaatt tgattgggca gtggagactt ttgtcatctt	13980
tcccggttg actcacattg tttcctatgg tgcactcacc accagccatt tccttgacac	14040
agttggtctg gttactgtgt ccaccgccg gttttatcac gggcggtatg tcttgagtag	14100
cgtctacgcg gtctgtgctc tggctgcgtt gatttgcttc gtcattagc ttgcaagaa	14160
ctgcatgtcc tggcgctact cttgtaccag atataccaac ttccttctgg aactaagg	14220
cagactctat cgttggcggt cgcccgttat catagagaaa gggggttaagg ttgaggtcga	14280
aggtcacctg atcgacctca aaagagttgt gcttgatggt tccgtggcaa cccctttaac	14340
cagagtttca gcggaacaat ggggtcgtct ctagacgatt tttgccatga tagcacggct	14400
ccacaaaagg tgcttttggc gttttccatt acctacacgc cagtaatgat atatgctcta	14460
aaggtaagtc gcggccgact gctagggctt ctgcacctt tgatctttct gaattgtgct	14520
tttaccttcg ggtacatgac attcgtgcac tttcagagca caaatagggt cgcgctcact	14580
atgggagcag tagttgcaact tctttggggg gtgtactcag ccatagaaac ctggaaattc	14640

```

atcacctcca gatgccgttt gtgcttgcta ggccgcaagt acattctggc cctgcccac 14700
cacgtcgaaa gtgccgcggg ctttcatccg attgcggcaa atgataacca cgcatttgtc 14760
gtccggcgctc ccggctccac tacggttaac ggcacattgg tgcccggggtt gaaaagcctc 14820
gtgttgggtg gcagaaaagc tgttaaacag ggagtggtaa accttgtcaa atatgccaaa 14880
taacaacggc aagcagcaaa agaaaaagaa ggggaatggc cagccagtca atcagctgtg 14940
ccagatgctg ggtaaaatca tcgcccagca aagccagtcc agaggcaagg gaccgggcaa 15000
gaaaagtaag aagaaaaacc cggagaagcc ccattttcct ctagcgaccg aagatgacgt 15060
caggcatcac ttaccccctg gtgagcggca atttgtgtctg tcgtcgatcc agactgcctt 15120
taatcagggc gctggaactt gtaccctgtc agattcaggg aggataagtt aactgtgga 15180
gttttagtttg ccgacgcata atactgtgcg cctgatccgc gtcacagcat caccctcagc 15240
atgatgggct ggcattcttt aggcacctca gtgtcagaat tggaagagtg tgtgggtgat 15300
ggcactgatt gacattgtgc ctctaagtca cctattcaat tagggcgacc gtgtgggggt 15360
aaaatttaat tggcgagaac catgcggccg caattaaaaa aaaaaaaaaa aaaaaaaaaa 15420
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 15450

```

<210> 6

<211> 15450

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Virut dùng cho vacxin có hệ gen hoàn chỉnh chiều dài hoàn chỉnh
P129-PKC12-FL lần cấy truyền 52(15,450 nt)

<400> 6

```

atgacgtata ggtgttggct ctatgccacg gcatttgtat tgtcaggagc tgtggccatt 60
ggcacagccc aaaacttgct gcacggaaaa cgcccttctg tgacagcctt cttcagggga 120
gcttaggggt ctgtccctag caccttgctt ctggagttgc actgctttac ggtctctcca 180
cccctttaac catgtctggg atacttgatc ggtgcacgtg ccccccaat gccaggggtg 240
ttatggcgga gggccaagtc tactgcacac gatgtctcag tgcacgggtc ctcccttcctc 300
tgaatctcca agttcctgag cttgggggtg tgggcctatt ttataggccc gaagagccac 360
tccggtggac gttgccacgt gcattcccca ctgtcgagtg ctccctgcc ggggcctgct 420
ggctttctgc gatctttcca attgcacgaa tgaccagtgg aaacctgaac tttcaacaaa 480
gaatgggtgc ggttgcagct gagatttaca gagccggcca actcaccct gcagttttga 540
aggctctaca agtttatgaa cgggggttgc gctggtaccc cattgtcgga cctgtccctg 600
gagtggccgt ttacgccaac tccctacatg tgagtgacaa acctttcccg ggagcaacte 660
atgtgttaac caacctaccg ctcccgaga ggccaagcc tgaagacttt tgcccttttg 720

```

agtgtgctat ggctaacgtc tatgacattg gccataacgc cgtcatgtat gtggccagag	780
ggaaagtctc ctgggccctt cgtggcgagg atgaagtga atttgaaacc gtccccgaag	840
agttgaagtt gattgcgaac cgactccaca tctccttccc gcccaccac gcagtggaca	900
tgtctgagtt tgccttcata gccctggga gtggtgtctc cttgcgggtc gagcaccaac	960
acggctgcct tcccgctgat actgtccctg atgggaactg ctggtggtac ttgtttgact	1020
tgctcccacc ggaagttcag aataaagaaa ttcgccgtgc taaccaattt ggctatcaaa	1080
ccaagcatgg tgtccatggc aagtacctac agcggaggct gcaagttaat ggtctccgag	1140
cagtgactga tacagatgga cctattgtcg tacagtactt ctctgttagg gagagttgga	1200
tccgccactt cagactggcg gaagaacctt gcctccctgg gtttgaagac ctctcagaa	1260
taagggtaga gcctaatacg tcgccaatgg gtggcaaggg tgaaaaaatc ttccggtttg	1320
gcagtcacaa gtggtacggt gctggaaaga gagcaaggag agcacgctct ggtgcgactg	1380
ccacggtcgc tcaccgcgct ttgcccgtc gcgaagccca gcaggccaag aagctcgagg	1440
ttgccagcgc caacagggt gagcatctca agtactattc cccgcctgcc gacgggaact	1500
gtggttggca ctgcatttcc gccattacca accggatggt gaattccaaa ttgaaacca	1560
ctcttcccga gagagtgaga ccttcagatg actgggctac tgacgaggat cttgtgaata	1620
ccatccaaat cctcaggctc cccgcggcct tggacaggaa cggtgcttgt gctggcgcca	1680
agtacgtgct caagctggaa ggtgagcact ggaccgtctc tgtgaccctt gggatgacct	1740
cttctttgct ccccttgaa tgtgttcagg gttgttgtga gcataagagc ggtcttggtt	1800
tcccagacgt ggtcgaagtt tccggatttg accctgcctg tcttgaccga cttgctgaga	1860
taatgcactt gcctagcagt gtcattccag ctgctctggc cgagatgtcc gacgacttca	1920
atcgtctggc ttccccggcc gccactgtgt ggactgtttc gcaattcttt gcccgccaca	1980
gaggaggaga gcattcctgac cagggtgtgt tagggaaaat tatcaacctt tgtcaggtga	2040
ttgaggaatg ctgctgttcc cggaacaaag ccaaccgggc taccgccgaa gaggttgagg	2100
caaaagttga ccagtacctc cgtggtgcag caagccttgg agaattgctt gccaaagctt	2160
agagggctcg cccgccgagc gcgacggaca cctcctttga ttggaatgtt gtgcttctg	2220
gggttgagac ggccaatcag acaaccaaac agctccatgt caaccagtgc cgcgctctgg	2280
ttctgtcgt gactcaagag cctttggaca gagactcggc cctctgacc gccttctcgc	2340
tgtccaattg ctactacct gcacaagggt acgagggtccg tcaccgtgag aggctaaact	2400
ccttgctctc taagttggag ggggttgttc gtgaggaata tgggctcacg ccaactggac	2460
ctggcccgcg acccgactg ccgaacgggc tcgacgagct taaagaccag atggaggagg	2520
atctgctgaa attagtcaac gccaggcaa cttcagaaat gatggcctgg gcagccgagc	2580
aggttgatct aaaagcttgg gtcaaaaatt acccagggtg gacaccgcca cccctccac	2640

caagagttca gcctcgaaaa acgaagtctg tcaagagctt gctagagAAC aagcctgtcc	2700
ctgctccgCG caggaaggTC agatctgatt gtggcagccc gattttgatg ggcgacaatg	2760
ttcctaacgg ttgggaagat tcgactgttg gtgggtccct tgatctttcg gcaccatccg	2820
agccgatgac acctctgagt gagcctgtac ttattttccag gccagtgaca tctttgagtG	2880
tgccggcccc agttcctgca ccgcgtagag ctgtgtcccg accgatgacg cctcagagtG	2940
agccaatttt tgtgtctgca ctgcgacaca aatttcagca ggtggaaaaa gcaaatctgg	3000
cggcagcagc gccgatgtgc caggacgaac ccttagattt gtctgcatcc tcacagactg	3060
aatatgaggc ttcccccta acaccaccgc agaacgtggg cattctggag gtaagggggc	3120
aagaagctga ggaagtctg agtgaaatct cggatattct gaatgatacc aaccctgcac	3180
ctgtgtcatc aagcagctcc ctgtcaagtG ttaagatcac acgccccaaa tactcagctc	3240
aagccattat cgactcgggc gggccctgca gtgggcacct ccaaaggga aaagaagcat	3300
gcctccgcat catgcgtgaa gcttgtgatg cggccaagct tagtgacct gccacgcagg	3360
aatggctttc tcgcatgtgg gatagggtgg acatgctgac ttggcgcaac acgtctgctt	3420
accaggcgtt tcgcacctta gatggcaggt ttgggtttct ccaaagatg atactcgaga	3480
cgcgcgcgcc ctaccctgtg gggtttgtga tgttgccctca caccctgca ccttccgtga	3540
gtgcagagag cgaccttacc attggttcag tcgccactga agatattcca cgcctcctcg	3600
ggaaaataga aaataccggt gagatgatca accagggacc cttggcatcc tctgaggaag	3660
aaccggtata caaccaacct gccaaagact cccggatata gtcgcggggg tctgacgaga	3720
gcacagcagc tccgtccgCG ggtacaggTg gcgcgggctt atttactgat ttgccacctt	3780
cagacggcgt agatgcggac ggtggggggc cgttgCagac ggtaagaaag aaagctgaaa	3840
ggctcttcga ccaattgagc cgtcaggttt ttaacctcgt ctcccatctc cctgttttct	3900
tctcacacct cttcaaactc gacagtggtt attctccggg tgattggggg tttgcagctt	3960
ttactctatt ttgcctcttt ttgtgttaca gctaccatt ctccggtttc gttcccctct	4020
tgggtgtatt ttctgggtct tctcggcgtg tgcgcagtgg ggtttttggc tgctggctgg	4080
cttttgctgt tggcctgttc aagcctgtgt ccgaccagT cggcactgct tgtgagtttg	4140
actcgccaga gtgcaggaac gtccttcatt cttttgagct totcaaacct tgggaccctg	4200
ttcgcagcct tgttgtgggc cccgtcggtc tcggctcttg cattcttggc aagttactgg	4260
gcggggcacg ctacatctgg cattttttgc ttaggcttgg cattgttgca gattgtatct	4320
tggctggagc ttatgtgctt tctcaaggta ggtgtaaaaa gtgctgggga tcttgtataa	4380
gaactgctcc taatgaaatc gccttcaacg tgttcccttt tacacgtgcg accaggctcg	4440
cactcatcga cctgtgcgat cggttttgtg cgccaacagg catggacccc attttcctcg	4500
ccactgggtg gcgtgggtgc tggaccggcc gaagtcccat tgagcaacct tctgaaaaac	4560

ccatcgcggtt cgcccagttg gatgaaaaga ggattacggc tagaactgtg gtcgctcagc	4620
cttatgatcc taatcaagcc gtgaagtgtc tgcgggtgtt acaggcggtt ggggcgatgg	4680
tggccgaggg agtcccaaaa gtggccaaag tttctgctat tccattccga gccctttttt	4740
ttcccaccgg agtgaaagtt gatcccagat gcaggatcgt ggttgacccc gatactttta	4800
ctacagccct ccggtctggt tactctacca caaacctcgt ccttggtgtg ggggactttg	4860
cccagctgaa tggactaaag atcaggcaaa tttccaagcc ttcgggagga ggcccacacc	4920
tcattgctgc cctgcatgtt gcctgctcga tggcggttga catgcttgct ggggtttatg	4980
taacttcagt ggggtcttgc ggtgccggca ccaacgatcc atggtgcaat aatccgtttg	5040
ccgttcctgg ctacggacca ggctctctct gcacgtccag attgtgcac tcccaacatg	5100
gccttaccct gcccttgaca gcaattgtgg cgggattcgg tcttcaggaa atcgccctgg	5160
togttttgat tttcgtttcc atcggaggca tggctcatag gttgagttgt aaggctgata	5220
tgctgtgcat tttacttgca atcgccagct atgtttgggt accccttacc tggttgcttt	5280
gtgtgtttcc ttgttggttg cgctgggtct ctttgacccc ccttaccatc ctatgggttg	5340
tgtttttctt gatttctgta aatatgcctt cgggaatctt ggccgtggtg ttattggttt	5400
ctctttggct tttgggacgt tatactaaca ttgctggtct tgtcaccccc tatgatattc	5460
atcattacac cagtggcccc cgcggtgttg ccgccttggc taccgcacca gatggaacct	5520
acttggtgc cgtccgccgc gctgcgttga ctggtcgac catgctgttc acccgtctc	5580
agcttgggtc ccttcttgag ggcgctttca gaactcgaag gccctcactg aacaccgtca	5640
atgtgggttg gtcctccatg ggctctggtg gagtgttcac catcgacggg aaaattaggt	5700
gcgtgactgc cgcacatgtc cttacgggta attcggctag ggtttccgga gtcggcttca	5760
atcaaagtct tgactttgat gtgaaagggg acttcgccat agctgattgc ccgaattggc	5820
aaggagctgc tcccaagacc caattctgcg aggacggatg gactggccgt gcctattggc	5880
tgacatcctc tggcgctcga cccggtgtta ttgggaatgg attcgccttc tgcttcaccg	5940
cgtgcggcga ttccgggtcc ccagtgatca ccgaagctgg tgagattgtc ggcgttcaca	6000
caggatcaaa taaacaagga ggtggcatcg tcacgcgccc ttcaggccag ttttgtaacg	6060
tggcaccat caagctgagc gaattaagtg aattctttgc tggaccaag gtcccgtcgc	6120
gtgatgtgaa ggttggcagc cacataatta aagacacgtg cgaagtacct tcagatcttt	6180
gcgccttgct tgctgccaaa cctgaactgg agggaggcct ctccaccgtc caacttctgt	6240
gtgtgttttt cctactgtgg agaattgatg gacatgcctg gacgcccttg gttgctgtgg	6300
ggtttttcat tctgaatgag gttctcccag ctgtcctggt tcggagtgtt ttctcctttg	6360
ggatgtttgt gctatcttgg ctacacccat ggtctgcga agttctgatg atcaggcttc	6420
taacagcagc tcttaacagg aacagatggt cacttgcctt ttacagcctt ggtgcggtga	6480

ccggttttgt cgcagatctt gcggttaactc aagggcaccc gttgcaggca gtaatgaatt	6540
tgagcaccta tgccttcctg cctcggatga tggttgtgac ctaccagtc ccagtgattg	6600
cgtgtggtgt tgtgcaccta cttgccatca ttttgtactt gttcaagtac cgcggcctgc	6660
acaatgttct tgttggtgat ggagcgtttt ctgcagcttt cttcttgcca tactttgccg	6720
agggaaagt gaggaagg gtgtcgcaat cctgcggaat gaatcatgag tcattgactg	6780
gtgccctgc tatgagactc aatgacgagg acttggaactt ccttacgaaa tggactgatt	6840
ttaagtgcct tgtttctgcg tccaacatga ggaatgcagc aggccaattc atcgaggctg	6900
cctatgcaaa agcacttaga attgaacttg ccagttggt gcaggttgat aaggttcgag	6960
gtactttggc caagcttgag gcttttgctg ataccgtggc accccaactc tcgcccgggtg	7020
acattgttgt tgccttggc catacgcctg ttggcagcat ctccgaccta aaggttggtg	7080
gtaccaagca tactctcaa gtcattgaga ccagagtcct tgccgggtcc aaaatgaccg	7140
tggcgcgcgt cgttgacca acccccacgc cccacccgc acccgtgcc atccccctcc	7200
caccgaaagt tctagagaat ggtcccaacg cctgggggga tggggaccgt ttgaataaga	7260
agaagaggcg taggatggaa accgtcgga tctttgtcat ggggtgggaag aagtaccaga	7320
aattttggga caagaattcc ggtgatgtgt ttacgagga ggtccatgac aacacagatg	7380
cgtgggagtg cctcagagtt ggtgaccctg ccgactttaa ccctgagaag ggaactctgt	7440
gtgggcatac tactattgaa gataaggatt acaaagtcta cgcctccca tctggcaaga	7500
agttcctggt ccccgtaac ccagagagcg gaagagcca atgggaagct gcaaagcttt	7560
ccgtggagca ggcccttggc atgatgaatg tcgacggtga actgacggcc aaagaagtgg	7620
agaaactgaa aagaataatt gacaaacttc agggccttac taaggagcag tgtttaaact	7680
gctagccgcc agcggcttga cccgctgtgg tcgcggcggc ttggttgta ctgagacagc	7740
ggtaaaaata gtcaaatttc acaaccggac tttcaccta gggcctgtga atttaaaagt	7800
ggccagtgag gttgagctga aagacgcggt cgagcacaac caacaccgg ttgcaagacc	7860
ggttgacggt ggtgttgtgc tcctgcgttc cgcagttcct tcgcttatag atgtcctgat	7920
ctccggtgct gacgcatctc ctaagttact cgctcgtcac gggccgggga aactgggat	7980
cgatggcacg ctttgggact ttgaggcca ggccaccaa gaggaattg cgctcagtgc	8040
gcaaataata caggcttgtg acattaggcg cggtagcga cctgaaattg gtctccctta	8100
caagctgtac cctgttaggg gcaacctga gcgggtaaaa ggagttttac agaatacaag	8160
gtttggagac ataccttaca aaaccccag tgacactggg agcccagtgc acgcggctgc	8220
ctgcctcacg cccaatgcca ctccggtgac tgatgggcgc tccgtcttgg ctactaccat	8280
gccctccggt tttgaattgt atgtaccgac cattccagcg tctgtccttg attatcttga	8340
ctctaggcct gactgcccc aacagttgac agagcacggc tgtgaggatg ccgcattgag	8400

agacctctcc aagtatgact tgtccacca aggctttgtt ttacctgggg ttcttcgcct	8460
tgtgcgtaag tacctgtttg cccacgtggg taagtgcccg cccgttcacg ggccttcac	8520
ttacctgcc aagaattcta tggctggaat aaatgggaac aggtttccaa ccaaggacat	8580
tcagagcgtc cccgaaatcg acgtttctgtg cgcacaggcc gtgcgagaaa actggcaaac	8640
tgttaccct tgtaccctca agaaacagta ttgtgggaag aagaagacta ggacaatact	8700
cggcaccaat aatttcattg cgttggtcca cggggcagcg ttgagtgggtg tcaccaggg	8760
cttcatgaaa aaggcgttta actcgcccat cgccctcggg aaaaacaaat ttaaggagct	8820
acagactccg atcttaggca ggtgccttga agctgatctt gcacctgtg atcgatccac	8880
acctgcaatt gtccgtggt ttgccgcaa ccttctttat gaacttgct gtgctgaaga	8940
gcacctaccg tcgtacgtgc tgaactgctg ccatgacctt ttggtcacgc agtccggcgc	9000
agtgactaag aggggtggcc tgcgtctgtg cgaccgacg acttctgtgt ctaacaccat	9060
ttacagcttg gtgatatatg cacagcacat ggtgcttagt tactttaaaa gtggtcaccc	9120
tcattggcctt ctgttctac aagaccagct gaagtctgag gacatgctca aagtccaacc	9180
cctgatcgtc tattcggaag acctcgtgct gtatgccgaa tctccacca tgccgaacta	9240
ccactggtgg gtcgaacatc tgaatttgat gctgggtttt cagacggacc caaagaagac	9300
agccataacg gactcgccat catttctagg ctgtaggata ataatggac gccagctagt	9360
ccccaacgt gacaggatcc tcgcggccct cgcttaccat atgaaggcaa acaatgtttc	9420
tgaatactac gccgcggcgg ctgcaatact catggacagc tgtgcttgtt tagagtatga	9480
tcctgaatgg tttgaagagc ttgtggttgg gatagcgcat tgcgccgca aggacggcta	9540
cagctttccc ggcccgcgt tcttcttctc catgtgggaa aaactcagat ccaatcatga	9600
ggggaagaag tccagaatgt gcgggtattg cggggccctg gctccgtacg ccactgcctg	9660
tggcctcgac gtctgtattt accacacca cttccaccag cattgtccag tcacaatctg	9720
gtgtggccac ccggtggtt ctggttcttg tagtgagtgc aaaccccc tagggaaagg	9780
cacaagccct ctagatgagg tgttagaaca agtcccgat aagcctccac ggactgtaat	9840
catgcatgtg gagcagggtc tcaccctct tgaccagggc agataccaga ctgcgccg	9900
attagtctcc gttaggcgtg gcacagagg aaatgaagtt gacctaccag acggtgatta	9960
tgctagcacc gccctactcc ccacttgtaa agagatcaac atggtcgctg tcgcctctaa	10020
tgtgttgccg agcaggttca tcacgggtcc gcccggtgct gggaaaacat actggctcct	10080
tcagcaggtc caggatggtg atgtcattta cacaccgact caccagacca tgctcgacat	10140
gattagggct ttggggacgt gccggttcaa cgtcccagca ggtgcaacgc tgcaattccc	10200
tgccccctcc cgtaccggcc cgtgggttcg catcctagcc ggcggttggt gtcttgtaa	10260
gaattccttc ttggatgaag cagcgtattg taatcacctt gatgtcttga ggctccttag	10320

caaaaccacc ctcacctgtc tgggagactt caaacaactc caccagtggt gttttgatcc	10380
tcatttgctat gtttttgaca tcatgcctca gaccagttg aagaccatct ggagattcgg	10440
acagaacatc tgtgatgcca tccaaccaga ttacagggtac aaacttgtgt ccatgggtcaa	10500
cacaaccctg gtaaccacg tggaaaaacc tgtcaagtat gggcaagtcc tccccctta	10560
ccacagggtac cgagaggacg gcgccatcac aattgactcc agtcaaggcg ccacatttga	10620
tgtggtcaca ctgcatttgc ccactaaaga ttcactcaac aggcaaagag cccttggtgc	10680
tatcaccagg gcaagacatg ctatctttgt gtatgaccca cacaggcaat tgcagagcat	10740
gtttgatctt cctgcgaagg gcacaccctg caacctcgca gtgcaccgtg atgagcagct	10800
gatcgactg gatagaaata ataaagaatg cacagttgct caggctctag gcaacggaga	10860
taaatttagg gccaccgaca agcgcgttgt agattctctc cgcgccattt gtgctgatct	10920
ggaagggctg agctctccgc tccccagggt cgcacacaac ttgggatttt atttctcacc	10980
tgatttgaca cagtttgcta aactcccgtt agaccttgca cccactggc ccgtgggtgac	11040
aaccagaac aatgaaaagt ggccggatcg gctgggtgcc agccttcgcc ctgtccataa	11100
gtatagccgt gcgtgcattg gtgccggcta tatgggtggc ccctcggtgt ttctaggcac	11160
ccctggggtc gtgtcatact acctcacaaa atttgtcaag ggcgaggctc aagtgttcc	11220
ggagacagtc ttcagcaccg gccgaattga ggtggattgc cgggagtatc ttgatgacag	11280
ggagcgagaa gttgctgagt ccctcccaca tgccttcatt ggcgacgtca aaggcaccac	11340
cggtggggga tgtcatcatg tcacctcaa ataccttcg cgcttccttc ccaaggaatc	11400
agtcgcggta gtcggggttt cgagccccgg gaaagccgca aaagcagtgt gcacattgac	11460
ggatgtgtac ctcccagacc ttgaggccta cctccacca gagactcagt ccaagtgtg	11520
gaaagttatg ttggacttca aggaagttcg actgatgtc tggaaagaca agacggccta	11580
tttccaactt gaaggccgt atttcacctg gtatcagctt gcaagctacg cctcgtacat	11640
ccgtgttctt gtcaactcca cggtgtatct ggaccctgc atgggccctg ccctttgcaa	11700
cagaagagtt gtcgggtcca ccattgggg agctgacctc gcagtcaccc cttatgatta	11760
cggtgctaaa atcatcttgt ctacgcctta ccatggtgaa atgcctcctg gatacaagat	11820
tctggcgtgc gcggagttct cgctcgacga ccagtcagg taaaaacaca cctggggttt	11880
tgaatcggat acagcgtatc tgtatgagtt caccggaaac ggtgaggact gggaggatta	11940
caatgatgcg tttcgtgcgc gccagaaagg gaaaatttat aaggccactg ctaccagcat	12000
gaagttttat tttccccgg gccccgtcat tgaaccaact ttaggcctga attgaaatga	12060
aatggggctt atacaaagcc tottcgacaa aattggccag ctttttgtgg atgctttcac	12120
ggaatttttg gtgtccattg ttgatatacat catatttttg gccattttgt ttggcttcac	12180
catcgccggg tggctggtgg tcttttgcac cagattgggt tgctccggg tattccgtgc	12240

gcgccttgcc attcaccctg agcaattaca gaagatccta tgaggccttt ctttctcagt	12300
gccgggtgga cattcccacc tggggggtaa aacacccttt ggggatgttt tggcaccata	12360
aggtgtcaac cctgattgat gaaatgggtg cgcgtcgaat gtaccgcgtc atggataaag	12420
cagggcaagc tgcttggaag caggtgggtg gcgaggctac gctgtctcgc attagtagtc	12480
tggtatgtgt ggctcatttt caacatcttg ccgccattga agccgagacc tgtaaataatt	12540
tggcttctcg actgcccatt ctacacaacc tgcgcatgac aggggtcaaatt gtaaccatag	12600
tgtataatag cactttaaat caggtgtttg ctatttttcc aacccttggt tcccggccaa	12660
agcttcatga ttttcagcaa tggctaatag ctgtacattc ctccatattt tcctctgttg	12720
cagcttcttg tactcttttt gttgtgctgt ggttgcggtt tccaatgcta cgtactgttt	12780
ttggtttccg ctgggttaggg gcaatttttc ttccgaactc atggtgaatt acacgggtgtg	12840
tccaccttgc ctcccccagc aagcagccgc tgaggctcctt gaaccgcgta ggtctctttg	12900
gtgcaggata gggcatgacc gatgtgggga ggacgatcac gacgaactgg ggttcatggt	12960
tccgcctggc ctctccagcg aaagccactt gaccagtgtt tacgcctggt tggcgttcct	13020
gtccttcagc tacacggccc agttccatcc cgagatattt gggataggga acgtgagtga	13080
agtttatgtt gacatcaagc accaattcat ctgcgcggtt catgacgggc agaaccacc	13140
cttgccctgc catgacaata tttcagccgt atttcagacc tactatcaac atcaggctga	13200
cggcggcaat tggtttcacc tagaatggct gcgtcccttc ttttccctctt ggttggtttt	13260
aaatgtttcg tggtttctca ggcgttcgcc tgcaagccat gtttcagttc gagtctttca	13320
gacatcaaaa ccaacactac cgcagcatca ggctttgttg tcctccagga catcagctgc	13380
cttaggcatg gcgactcgtc ctttccgacg attcgcaaaa gctctcaatg ccgcacggcg	13440
ataggagcac ctgtgtatat caccatcaca gccaatgtga cagatgagaa ttacttacat	13500
tcttctgac tcctcatgct ttcttcttgc ctttctatg cttctgagat gagtgaagag	13560
ggattcaagg tggatatttg caatgtgtca ggcacgttgg ctgtgtgtgt caactttacc	13620
agctacgtcc aacatgtcaa agagtttact caacgctcct tggtggtcga tcatgtgcgg	13680
ctgcttcatt tcatgacacc tgagaccatg aggtgggcaa ccgttttagc ctgtcttttt	13740
gccatcctac tggcaatttg aatgttcaag tatgttgggg aaatgcttga ccgcgggctg	13800
ttgctcgcga ttgctttctt tgtggtgtat cgtgccgttc tggtttgctg tgctcggcaa	13860
cgccaacagc agcagcagct ctcatctcca gttgatttat aacttgacgc tatgtgagct	13920
gaatggcaca gattggctgg cagaaaaatt tgattgggag gtggagactt ttgtcatctt	13980
tcccgtgttg actcacattg tttcctattg tgcaactcacc accagccatt tccttgacac	14040
agttgggtcg gttactgtgt ccaccgccgg gttttatcac gggcggtatg tcttgagtag	14100
catctacgcg gtctgtgctc tggctgcgtt gatttgcttc gttattagga ttgcgaagaa	14160

ctgcatgtcc tggcgctact cttgtaccag atataccaac ttccttctgg acactaaggg 14220
 cagactctat cgttggcggt cgcccgttat catagaaaaa aggggtaagg ttgaggtcga 14280
 aggtcatctg atcgacctca aaagagttgt gcttgatggg tccgtggcaa cccctttaac 14340
 cagagtttca gcggaacaat ggggtcgtct ctagacgact tttgccatga tagcactgct 14400
 ccacaaaagg tgcttttggc gttttccatt acctacacgc cagtaatgat atatgctcta 14460
 aaggtaagtc gcggccgact gctagggctt ctgcaccttt tgatctttct gaattgtgct 14520
 tttaccttcg ggtacatgac attcgcgcac tttcagagca caaatagggt cgcgctcgct 14580
 atgggagcag tagttgcact tctttggggg gtgtactcag ccatagaaac ctggaaattc 14640
 atcacctcca gatgccgttt gtgcttgcta ggccgcaagt acattctggc ccctgccac 14700
 cacgtcgaaa gtgccgcggg ctttcatccg attgcggcaa atgataacca cgcatttgct 14760
 gtccggcgctc ccggctccat tacggttaac ggcacattgg tgcccgggtt gaaaagcctc 14820
 gtgttggttg gcagaaaagc tgtaaagcag ggagtggtaa accttgtcaa atatgccaaa 14880
 taacaacggc aagcagcaaa agaaaaagaa ggggaatggc cagccagtca accagctgtg 14940
 ccagatgctg ggtaaaatca tcgccagca aaaccagtcc agaggcaagg gaccgggcaa 15000
 gaaaagtaag aagaaaaacc cggagaagcc ccattttcct ctagcgaccg aagatgacgt 15060
 caggcatcac ttcacctctg gtgagcggca attgtgtctg tcgtcgatcc agactgcctt 15120
 taaccagggc gctggaactt gtaccctgtc agattcaggg aggataagtt aactgtgga 15180
 gtttagtttg ccgacgcac atactgtgcg cctgatccgc gtcacagcat caccctcagc 15240
 atgatgggct ggcattcttt aggcacctca gtgtcagaat tggaagaatg tgtggtggat 15300
 ggcactgatt gacattgtgc ctctaagtca cctattcaat tagggcgacc gtgtgggggt 15360
 aaaatttaat tggcgagaac catgcggccg caattaaaaa aaaaaaaaaa aaaaaaaaaa 15420
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 15450

<210> 7
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> axit amin vị trí trong hệ gen 735, ORF1a: 182, Nsp1b: 2, lần cấy
 truyền 0
 <400> 7

Ala Met Ala Asp Val Tyr Asp
 1 5

<210> 8
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 735, ORF1a: 182, Nsp1b: 2, lần cấy truyền 17

<400> 8

Ala Met Ala Asn Val Tyr Asp
1 5

<210> 9

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 735, ORF1a: 182, Nsp1b: 2, lần cấy truyền 52

<400> 9

Ala Met Ala Asn Val Tyr Asp
1 5

<210> 10

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 756, ORF1a: 189, Nsp1b: 9, lần cấy truyền 0

<400> 10

Ile Ser His Asp Ala Val Met
1 5

<210> 11

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 756, ORF1a: 189, Nsp1b: 9, lần cấy truyền 17

<400> 11

Ile Gly His Asn Ala Val Met
1 5

<210> 12

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 756, ORF1a: 189, Nsp1b: 9, lần cấy truyền 52

<400> 12

Ile Gly His Asn Ala Val Met
1 5

<210> 13

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 992, ORF1a: 267, Nsp1b: 87, lần cấy truyền 0

<400> 13

Thr Val Pro Glu Gly Asn Cys
1 5

<210> 14

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 992, ORF1a: 267, Nsp1b: 87, lần cấy truyền 17

<400> 14

Thr Val Pro Asp Gly Asn Cys
1 5

<210> 15

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 992, ORF1a: 267, Nsp1b: 87, lần cấy truyền 52

<400> 15

Thr Val Pro Asp Gly Asn Cys
1 5

<210> 16

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 1009, ORF1a: 273, Nsp1b: 93, lần cấy truyền 0

<400> 16

Cys Trp Trp Cys Leu Phe Asp
1 5

<210> 17
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> axit amin vị trí trong hệ gen 1009, ORF1a: 273, Nsp1b: 93, lần cấy truyền 17

<400> 17

Cys Trp Trp Tyr Leu Phe Asp
 1 5

<210> 18
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> axit amin vị trí trong hệ gen 1009, ORF1a: 273, Nsp1b: 93, lần cấy truyền 52

<400> 18

Cys Trp Trp Tyr Leu Phe Asp
 1 5

<210> 19
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> axit amin vị trí trong hệ gen 1096, ORF1a: 302, Nsp1b: 122, lần cấy truyền 0

<400> 19

His Gly Val Pro Gly Lys Tyr
 1 5

<210> 20
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> axit amin vị trí trong hệ gen 1096, ORF1a: 302, Nsp1b: 122, lần cấy truyền 17

<400> 20

His Gly Val His Gly Lys Tyr
 1 5

<210> 21
 <211> 7
 <212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 1096, ORF1a: 302, Nsp1b: 122, lần cấy truyền 52

<400> 21

His Gly Val His Gly Lys Tyr
1 5

<210> 22

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 2106, ORF1a: 639, Nsp2: 256, lần cấy truyền 0

<400> 22

Ala Ala Lys Ile Asp Gln Tyr
1 5

<210> 23

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 2106, ORF1a: 639, Nsp2: 256, lần cấy truyền 17

<400> 23

Ala Ala Lys Val Asp Gln Tyr
1 5

<210> 24

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 2106, ORF1a: 639, Nsp2: 256, lần cấy truyền 52

<400> 24

Ala Ala Lys Val Asp Gln Tyr
1 5

<210> 25

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 2185, ORF1a: 665, Nsp2: 282, lần cấy truyền 0

<400> 25

Pro Ser Ala Met Asp Thr Ser
1 5

<210> 26

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 2185, ORF1a: 665, Nsp2: 282, lần cấy truyền 17

<400> 26

Pro Ser Ala Thr Asp Thr Ser
1 5

<210> 27

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 2185, ORF1a: 665, Nsp2: 282, lần cấy truyền 52

<400> 27

Pro Ser Ala Thr Asp Thr Ser
1 5

<210> 28

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 2403, ORF1a: 738, Nsp2: 355, lần cấy truyền 0

<400> 28

Leu Val Ser Val Leu Ser Lys
1 5

<210> 29

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 2403, ORF1a: 738, Nsp2: 355, lần cấy truyền 17

<400> 29

Leu Asn Ser Leu Leu Ser Lys
1 5

<210> 30
<211> 7
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> axit amin vị trí trong hệ gen 2403, ORF1a: 738, Nsp2: 355, lần cấy truyền 52

<400> 30

Leu Asn Ser Leu Leu Ser Lys
1 5

<210> 31
<211> 7
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> axit amin vị trí trong hệ gen 3019, ORF1a: 943, Nsp2: 560, lần cấy truyền 0

<400> 31

Ala Pro Met Tyr Gln Asp Glu
1 5

<210> 32
<211> 7
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> axit amin vị trí trong hệ gen 3019, ORF1a: 943, Nsp2: 560, lần cấy truyền 17

<400> 32

Ala Pro Met Cys Gln Asp Glu
1 5

<210> 33
<211> 7
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> axit amin vị trí trong hệ gen 3019, ORF1a: 943, Nsp2: 560, lần cấy truyền 52

<400> 33

Ala Pro Met Cys Gln Asp Glu
1 5

<210> 34
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> axit amin vị trí trong hệ gen 4477, ORF1a: 1429, Nsp3: 72, lần cấy truyền 0

<400> 34

Cys Ala Pro Lys Gly Met Asp
 1 5

<210> 35
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> axit amin vị trí trong hệ gen 4477, ORF1a: 1429, Nsp3: 72, lần cấy truyền 17

<400> 35

Cys Ala Pro Thr Gly Met Asp
 1 5

<210> 36
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> axit amin vị trí trong hệ gen 4477, ORF1a: 1429, Nsp3: 72, lần cấy truyền 52

<400> 36

Cys Ala Pro Thr Gly Met Asp
 1 5

<210> 37
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> axit amin vị trí trong hệ gen 4705, ORF1a: 1505, Nsp3: 148, lần cấy truyền 0

<400> 37

Pro Lys Val Val Lys Val Ser
 1 5

<210> 38
 <211> 7
 <212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 4705, ORF1a: 1505, Nsp3: 148, lần cấy truyền 17

<400> 38

Pro Lys Val Ala Lys Val Ser
1 5

<210> 39

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 4705, ORF1a: 1505, Nsp3: 148, lần cấy truyền 52

<400> 39

Pro Lys Val Ala Lys Val Ser
1 5

<210> 40

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 5985, ORF1a: 1932, Nsp4: 129, lần cấy truyền 0

<400> 40

Ala Gly Glu Leu Val Gly Val
1 5

<210> 41

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 5985, ORF1a: 1932, Nsp4: 129, lần cấy truyền 17

<400> 41

Ala Gly Glu Ile Val Gly Val
1 5

<210> 42

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 5985, ORF1a: 1932, Nsp4: 129, lần cấy truyền 52

<400> 42

Ala Gly Glu Ile Val Gly Val
1 5

<210> 43

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 7419, ORF1a: 2410, Nsp7: 217, lần cấy truyền 0

<400> 43

Ala Asp Phe Asp Pro Glu Lys
1 5

<210> 44

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 7419, ORF1a: 2410, Nsp7: 217, lần cấy truyền 17

<400> 44

Ala Asp Phe Asn Pro Glu Lys
1 5

<210> 45

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 7419, ORF1a: 2410, Nsp7: 217, lần cấy truyền 52

<400> 45

Ala Asp Phe Asn Pro Glu Lys
1 5

<210> 46

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 8831, ORF1a/1b: 2881, Nsp9: 429, lần cấy truyền 0

<400> 46

Gln Thr Pro Val Leu Gly Arg
1 5

<210> 47
<211> 7
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> axit amin vị trí trong hệ gen 8831, ORF1a/1b: 2881, Nsp9: 429, lần cấy truyền 17

<400> 47

Gln Thr Pro Ile Leu Gly Arg
1 5

<210> 48
<211> 7
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> axit ami vị trí trong hệ gen 8831, ORF1a/1b: 2881, Nsp9: 429, lần cấy truyền 52

<400> 48

Gln Thr Pro Ile Leu Gly Arg
1 5

<210> 49
<211> 7
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> axit amin vị trí trong hệ gen 13,857, ORF5: 29, lần cấy truyền 0

<400> 49

Ala Val Leu Val Asn Ala Asn
1 5

<210> 50
<211> 7
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> axit amin vị trí trong hệ gen 13,857, ORF5: 29, lần cấy truyền 17

<400> 50

Ala Val Leu Gly Asn Ala Asn
1 5

<210> 51
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> axit amin vị trí trong hệ gen 13,857, ORF5: 29, lần cấy truyền 52

<400> 51

Ala Val Leu Gly Asn Ala Asn
 1 5

<210> 52
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> axit amin vị trí trong hệ gen 14,578, ORF6: 74, lần cấy truyền 0

<400> 52

Val Ala Leu Thr Met Gly Ala
 1 5

<210> 53
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> axit amin vị trí trong hệ gen 14,578, ORF6: 74, lần cấy truyền 17

<400> 53

Val Ala Leu Ala Met Gly Ala
 1 5

<210> 54
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> axit amin vị trí trong hệ gen 14,578, ORF6: 74, lần cấy truyền 52

<400> 54

Val Ala Leu Ala Met Gly Ala
 1 5

<210> 55
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> axit amin vị trí trong hệ gen 14,780, ORF6: 141, lần cấy truyền 0

<400> 55

Pro Gly Ser Thr Thr Val Asn
1 5

<210> 56

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 14,780, ORF6: 141, lần cấy truyền 17

<400> 56

Pro Gly Ser Ile Thr Val Asn
1 5

<210> 57

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> axit amin vị trí trong hệ gen 14,780, ORF6: 141, lần cấy truyền 52

<400> 57

Pro Gly Ser Ile Thr Val Asn
1 5