(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) **1-0042443**

(51)[7] **A01K 67/027**; C12N 5/10; C12N 15/85; A01K 67/02; C07K 14/705

(13) **B**

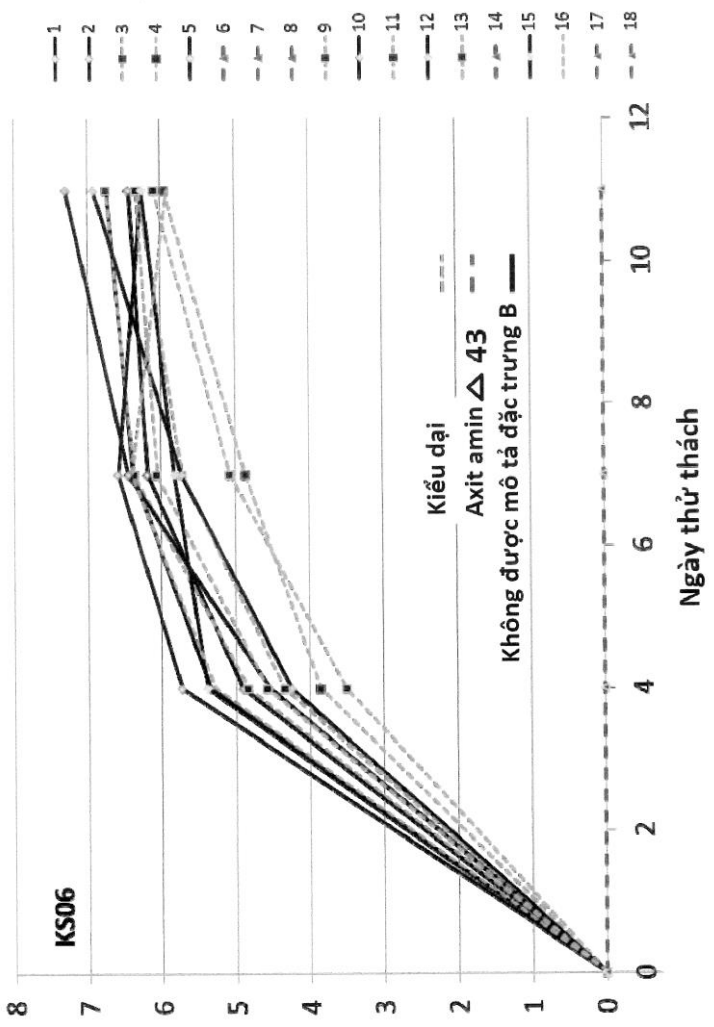
(21) 1-2018-00911
(22) 22/07/2016
(86) PCT/US2016/043467 22/07/2016
(87) WO2017/023570 09/02/2017
(30) 62/202,145 06/08/2015 US
(45) 27/01/2025 442
(43) 25/07/2018 364A
(73) THE CURATORS OF THE UNIVERSITY OF MISSOURI (US)
316 University Hall, Columbia, Missouri 65211, United States of America
(72) PRATHER, Randall, S. (US); WELLS, Kevin, D. (US); WHITWORTH, Kristin, M. (US).
(74) Công ty Luật TNHH T&G (TGVN)

(54) ĐỘNG VẬT LÀ LỢN HOẶC THẾ HỆ CON CỦA NÓ CHỨA TRÌNH TỰ NHIỄM SẮC THỂ SỬA ĐỔI Ở ÍT NHẤT MỘT ALEN CỦA GEN MÃ HÓA CHO PROTEIN CD163 VÀ TẾ BÀO LỢN CHỨA TRÌNH TỰ NHIỄM SẮC THỂ NÀY

(21) 1-2018-00911

(57) Sáng chế đề cập đến động vật không phải người và thế hệ thế hệ con bao gồm ít nhất một trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163. Sáng chế cũng đề cập đến các tế bào động vật chứa trình tự nhiễm sắc thể sửa đổi nêu trên. Động vật và tế bào có tính kháng cao đối với các tác nhân gây bệnh, bao gồm virut gây hội chứng hô hấp và sinh sản ở lợn (porcine reproductive and respiratory syndrome virus - PRRSV). Động vật và thế hệ con có các biến đổi về nhiễm sắc thể của gen CD163. Sáng chế còn đề cập đến phương pháp chọn giống để tạo ra động vật và quần thể động vật kháng tác nhân gây bệnh được tạo ra sử dụng phương pháp này.

Fig. 15



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến các động vật không phải người và thế hệ con của động vật này chứa ít nhất một trình tự nhiễm sắc thể cải biến ở trong gen mã hóa cho protein CD163. Sáng chế còn đề cập đến tế bào động vật chứa các trình tự nhiễm sắc thể sửa đổi này. Động vật và tế bào có tính kháng cao đối với các tác nhân gây bệnh, bao gồm virut gây hội chứng hô hấp và sinh sản ở lợn (porcine reproductive and respiratory syndrome virus - PRRSV). Động vật hoặc thế hệ thế hệ con có các biến đổi về nhiễm sắc thể của gen CD163, khiến cho sự xâm nhập và sao chép của PRRSV được ngăn chặn và động vật thu được thể hiện tính kháng đối với bệnh và hội chứng gây ra bởi virut. Sáng chế còn đề cập đến phương pháp chọn giống để tạo ra động vật và quần thể động vật kháng tác nhân gây bệnh được tạo ra sử dụng phương pháp này. Sáng chế cũng đề cập đến phương pháp sửa chữa gen của gen CD163 bao gồm việc bơm trực tiếp phôi và phát triển động vật, động vật và dòng biến đổi gen kháng lại các tác nhân gây bệnh như PRRSV.

Tình trạng kỹ thuật của sáng chế

Virut gây hội chứng hô hấp và sinh sản ở lợn (PRRSV) thuộc nhóm arterivirut ở động vật có vú, nhóm này cũng bao gồm virut làm tăng lactat dehydrogenaza ở chuột, virut gây sốt xuất huyết ở khỉ và virut gây viêm khớp ở ngựa. Các arterivirut có chung các đặc tính quan trọng liên quan đến sinh bệnh học do virut, bao gồm hướng kích thích đối với các đại thực bào và khả năng gây ra bệnh trầm trọng và nhiễm trùng dai dẳng. Các hội chứng bệnh lâm sàng gây ra bởi nhiễm trùng do virut gây hội chứng hô hấp và sinh sản ở lợn (PRRSV) được thông báo lần đầu tại Mỹ năm 1987 (Keffaber, 1989) và sau đó tại châu Âu năm 1990 (Wensvoort et al., 1991). Nhiễm trùng bởi PRRSV dẫn đến bệnh về hô hấp bao gồm ho và sốt, mất khả năng sinh sản trong thời gian mang thai muộn, và đặc tính sinh trưởng giảm. Virut cũng tham gia vào nhiều tương tác hội chứng bệnh do nhiều vi khuẩn trong khi duy trì nhiễm trùng cận lâm sàng lâu dài (Rowland et al., 2012).

Từ khi xuất hiện, hội chứng PRRS đã trở thành bệnh quan trọng nhất ở lợn thương mại tại Bắc Mỹ, châu Âu và châu Á, chỉ duy nhất các lục địa châu Úc và châu Nam Cực là không bị bệnh này. Tại Bắc Mỹ, chỉ duy nhất sự tổn thất liên quan đến PRRSV được lượng tính gây tốn kém cho các nhà sản xuất 664 triệu đô la Mỹ mỗi năm (Holtkamp et al., 2013). Năm 2006, một dạng bệnh trầm trọng hơn nữa, được biết là PRRS sinh bệnh cao (HP-PRRS), đã giết chết một phần mười số quần thể lợn trên toàn quốc gia Trung Quốc. Đa dạng di truyền đã hạn chế sự phát triển các vacxin cần thiết để kiểm soát và tiệt trừ bệnh một cách hiệu quả. Trong khi chọn lọc di truyền về tính kháng tự nhiên phải là sự lựa chọn, kết quả cho tới hiện này vẫn bị giới hạn (Boddicker et al., 2014).

Việc so sánh về phân tử giữa các virut ở Bắc Mỹ và châu Âu đưa tất cả các thể phân lập PRRSV vào một trong hai kiểu gen, lần lượt là Typ 2 hoặc Typ. Thậm chí ngay cả hai kiểu gen chỉ có độ tương đồng khoảng 70% ở mức nucleotit (Nelsen et al., 1999), thì cả hai có chung hướng kích thích đối với các tế bào dương tính CD163, thiết lập các nhiễm trùng kéo dài, và tạo ra các dấu hiệu lâm sàng giống nhau.

CD163 là một protein màng typ 1 130 kDa bao gồm 9 miền giàu xystein thụ thể dọn dẹp (SRCR) (Fabriek et al., 2005). *CD163* lợn chứa 17 exon mã hóa cho trình tự tín hiệu peptit, tiếp theo bằng chín miền SRCR, hai miền liên kết, cũng được gọi là các miền prolin serin threonin (PST), nằm ở sau SRCR 6 và SRCR 9, và miền tế bào chất, tiếp theo bởi đuôi tế bào chất ngắn. Sự biểu hiện trên bề mặt của CD163 được giới hạn ở các tế bào thuộc dòng tế bào đơn nhân-đại thực bào. Protein đầu tiên được nhận diện trong các mô của người nhờ khả năng liên kết các phức hemoglobin-haptoglobin (HbHp) của nó (Kristiansen et al., 2001). Dọn dẹp làm sạch HbHp là chức năng chính của CD163 và định vị tại SRCR 3 (Madsen et al., 2004). Sản phẩm chuyển hóa giải phóng từ các đại thực bào sau khi thoái giáng HbHp bao gồm bilirubin, CO, và sắt tự do. Một chức năng quan trọng của CD163, phòng ngừa độc tính oxy hóa xuất phát từ hemoglobin tự do (Kristiansen et al., 2001; Soares et al., 2009).

CD163, ở dạng thụ thể của PRRSV, lần đầu được mô tả bởi tác giả Calvert et. al. (2007). Việc gây nhiễm các dòng tế bào không cho phép bằng các cADN CD163 từ nhiều loài, bao gồm khỉ, người, chó, và chuột có thể khiến cho các tế bào cho phép chấp nhận nhiễm trùng PRRSV (Calvert et al., 2007). Ngoài CD163, protein thụ thể thứ hai,

CD169 (còn được biết là sialoadhesin hoặc SIGLEC1), được nhận diện là thụ thể PRRSV cơ bản liên quan đến việc hình thành sự tương tác đầu tiên với heterodime chất nền GP5 (M), protein chính trên bề mặt của virion (Delputte et al., 2002). Trong mẫu này, sự tương tác sau đó giữa CD163 và GP2, 3, 4 heterotrime trong khoang nội thể sẽ trung gian cho sự tháo vỏ và giải phóng hệ gen virut vào trong tế bào chất (Van Breedam et al., 2010, Allende et al., 1999). Mẫu trước đó mô tả nhiễm trùng PRRSV của các đại thực bào phế nang nhận diện SIGLEC1 (CD169) là thụ thể virut chính trên bề mặt các đại thực bào; tuy nhiên, công trình trước đó sử dụng lợn *SIGLEC1*$^{-/-}$ không chứng minh sự khác nhau về sao chép virut so với lợn kiểu hoang dại (Prather et a., 2013). Các kết quả này hỗ trợ cho các nghiên cứu *in vitro* trước đó chứng minh rằng, các dòng tế bào kháng PRRSV không có CD169 và CD163 trên bề mặt hỗ trợ cho sự sao chép virut sau khi gây nhiễm bằng plasmit CD163 (Welch et al., 2010).

Nhiều đặc điểm của cả sinh bệnh học PRRSV (nhất là ở mức phân tử) và bệnh học dịch súc vật chưa được hiểu đúng do đó khiến cho các nỗ lực kiểm soát trở nên khó khăn. Hiện tại, các nhà sản xuất thường tiêm chủng lợn chống lại PRRSV bằng các chủng sống giảm độc lực cải biến hoặc vacxin virut chết, tuy nhiên, các vacxin hiện tại thường không tạo ra sự bảo vệ thỏa mãn. Điều này là do cả sự biến đổi chủng và sự kích thích không tương xứng của hệ thống miễn dịch. Ngoài những quan tâm về hiệu quả của các vacxin PRRSV có sẵn, có bằng chứng rõ ràng rằng, vacxin sống cải biến hiện đang sử dụng có thể tiếp tục tồn tại ở các cá thể lợn riêng biệt và các đàn lợn và tích luỹ đột biến (Mengeling et al., Am. J. Vet. Res, 60(3): 334–340 (1999)), như đã được chứng minh bằng các thể phân lập diện độc sau khi gây nhiễm trùng thực nghiệm lợn (Rowland et al., Virology, 259:262–266 (1999)). Ngoài ra, đã được chứng minh rằng, virut vacxin phát tán trong tinh dịch của lợn đực đã chủng ngừa (Christopher-Hennings et al., Am. J. Vet. Res, 58(1): 40–45 (1997)). Đối với phương pháp thay thế chủng ngừa, một số chuyên gia ủng hộ chiến lược "thử nghiệm và loại bỏ" trong các đàn chọn giống (Dee and Molitor, Vet. Rec., 143:474–476 (1998)). Việc sử dụng thành công chiến lược này phụ thuộc vào việc loại bỏ toàn bộ lợn mà bị nhiễm trùng cấp tính hoặc nhiễm trùng dai dẳng bởi PRRSV, tiếp theo bằng cách kiểm soát chặt chẽ để phòng ngừa sự tái xâm nhập của virut. Khó khăn và tốn nhiều chi phí kết hợp với chiến lược này là ở chỗ, hiểu biết rất ít về sinh bệnh học của nhiễm trùng PRRSV dai dẳng và do đó, không có kỹ thuật đáng tin cậy nào để nhận diện lợn bị nhiễm trùng dai dẳng.

Như có thể quan sát thấy, trong lĩnh vực có nhu cầu cần phải phát triển các chiến lược để kích thích tính kháng PRRSV ở động vật.

Bản chất kỹ thuật của sáng chế

Sáng chế đề cập đến động vật không phải người, thế hệ thế hệ con, và các tế bào động vật mà chứa ít nhất một trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163.

Sáng chế cũng đề cập đến các phương pháp chọn giống để tạo ra động vật hoặc các dòng mà có tính nhạy giảm đối với nhiễm trùng bởi tác nhân gây bệnh. Phương pháp bao gồm bước sửa đổi noãn bào hoặc tế bào tinh trùng về mặt di truyền để đưa trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163 vào trong ít nhất một trong số noãn bào và tế bào tinh trùng, và làm thụ tinh noãn bào bằng tế bào tinh trùng để tạo ra trứng đã thụ tinh chứa trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163. Theo cách khác, phương pháp bao gồm bước sửa đổi về mặt di truyền trứng đã thụ tinh để đưa trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163 vào trong trứng đã thụ tinh. Phương pháp bao gồm thêm bước chuyển trứng đã thụ tinh vào trong động vật cái lựa chọn, trong đó thời gian mang thai và sinh đúng kỳ hạn tạo ra động vật thế hệ con, sàng lọc động vật thế hệ con về tính nhạy đối với tác nhân gây bệnh, và lựa chọn động vật thế hệ con mà có tính nhạy giảm đối với tác nhân gây bệnh khi so sánh với các động vật mà không chứa trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163.

Sáng chế cũng đề cập đến các quần thể động vật được tạo ra bằng phương pháp chọn giống.

Sáng chế cũng đề cập đến phương pháp gia tăng tính kháng của động vật nuôi đối với nhiễm trùng bởi tác nhân gây bệnh. Phương pháp bao gồm bước sửa chữa về mặt di truyền ít nhất một trình tự nhiễm sắc thể từ gen mã hóa cho protein CD163 sao cho sự tạo ra hoặc hoạt tính của protein CD163 giảm xuống khi so sánh với sự tạo ra hoặc hoạt tính của protein CD63 ở động vật nuôi mà không chứa trình tự nhiễm sắc thể đã sửa chữa ở gen mã hóa cho protein CD163.

Những sự sửa đổi đối với trình tự nhiễm sắc thể ở gen mã hóa cho protein CD163 đề cập ở đây làm giảm tính nhạy của động vật, thế hệ con, tế bào hoặc quần thể (ví dụ,

động vật là lợn, thế hệ con, tế bào hoặc quần thể) đối với tác nhân gây bệnh (ví dụ, virut như virut gây hội chứng hô hấp và sinh sản ở lợn (PRRSV)).

Ở bất kỳ trong số các động vật, thế hệ con, tế bào, quần thể, và các phương pháp được đề cập ở đây, trình tự nhiễm sắc thể sửa đổi có thể dẫn đến sự tạo ra protein CD163 cơ bản không có chức năng bởi động vật, thế hệ con, tế bào hoặc quần thể.

Ở bất kỳ trong số các động vật, thế hệ con, tế bào, quần thể, và các phương pháp được đề cập ở đây, trình tự nhiễm sắc thể sửa đổi có thể bao gồm một đoạn xóa trong khung ở gen mã hóa cho protein CD163.

Ở bất kỳ trong số các động vật là lợn, thế hệ con, tế bào, quần thể, và các phương pháp được đề cập ở đây, trình tự nhiễm sắc thể sửa đổi có thể bao gồm SEQ ID NO: 118. Theo cách khác, sự sửa đổi của trình tự nhiễm sắc thể ở gen mã hóa cho protein CD163 có thể bao gồm: đoạn xóa 11 cặp bazơ từ nucleotit 3137 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xen 2 cặp bazơ giữa các nucleotit 3149 và 3150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, với đoạn xóa 377 cặp bazơ từ nucleotit 2573 đến nucleotit 2949 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trên cùng alen; đoạn xóa 124 cặp bazơ từ nucleotit 3024 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 123 cặp bazơ từ nucleotit 3024 đến nucleotit 3146 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xen 1 cặp bazơ giữa các nucleotit 3147 và 3148 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 130 cặp bazơ từ nucleotit 3030 đến nucleotit 3159 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 132 cặp bazơ từ nucleotit 3030 đến nucleotit 3161 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1506 cặp bazơ từ nucleotit 1525 đến nucleotit 3030 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xen 7 cặp bazơ giữa nucleotit 3148 và nucleotit 3149 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1280 cặp bazơ từ nucleotit 2818 đến nucleotit 4097 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1373 cặp bazơ từ nucleotit 2724 đến nucleotit 4096 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1467 cặp bazơ từ nucleotit 2431 đến nucleotit 3897 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2417 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 12 cặp bazơ bắt đầu tại nucleotit 488, và trong đó có thêm đoạn

xóa 129 cặp bazơ trong exon 7 từ nucleotit 3044 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 28 cặp bazơ từ nucleotit 3145 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1387 cặp bazơ từ nucleotit 3145 đến nucleotit 4531 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1382 cặp bazơ từ nucleotit 3113 đến nucleotit 4494 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 11 cặp bazơ bắt đầu tại nucleotit 3113; đoạn xóa 1720 cặp bazơ từ nucleotit 2440 đến nucleotit 4160 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 452 cặp bazơ từ nucleotit 3015 đến nucleotit 3466 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; hoặc bất kỳ sự kết hợp nào của chúng.

Sáng chế cũng đề cập đến các axit nucleic. Các phân tử axit nucleic chứa trình tự nucleotit được chọn từ nhóm bao gồm: (a) trình tự nucleotit bao gồm SEQ ID NO: 47; (b) trình tự nucleotit có độ tương đồng trình tự ít nhất 80% với trình tự của SEQ ID NO: 47, trong đó trình tự nucleotit này chứa ít nhất một đoạn thế, đoạn xen hoặc đoạn xóa liên quan đến SEQ ID NO: 47; và (c) trình tự cADN của (a) hoặc (b).

Ví dụ, phân tử axit nucleic có thể bao gồm: (a) trình tự nucleotit có độ tương đồng trình tự ít nhất 87,5% với trình tự nêu trong SEQ ID NO: 47, trong đó trình tự nucleotit này chứa ít nhất một đoạn thế, đoạn xen hoặc đoạn xóa liên quan đến SEQ ID NO: 47; hoặc (b) trình tự cADN của (a).

Ngoài ra, sáng chế cũng đề cập đến các axit nucleic. Axit nucleic có thể bao gồm SEQ ID NO: 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 118, hoặc 119.

Phân tử axit nucleic bất kỳ có thể là phân tử axit nucleic phân lập.

Các đối tượng và các dấu hiệu khác sẽ phần nào được thể hiện rõ ràng và phần nào được chỉ ra dưới đây.

Mô tả vắn tắt các hình vẽ

FIG. 1. Các vectơ hướng đích và CRISPR được sử dụng để sửa đổi CD163. Panen A mô tả các exon 7, 8 và 9 kiểu hoang dại của gen CD163 được hướng đích để sửa đổi sử dụng các CRISPR. Panen B shows vectơ hướng đích được thiết kế để thay thế exon 7 của lợn (miền SRCR5 lợn của CD163) bằng ADN mã hóa cho SRCR8 người

của CD163L. Vectơ hướng đích này được sử dụng trong gây nhiễm bằng sự lựa chọn thuốc bởi G418. Các đoạn mồi PCR trong thử nghiệm nhánh trái và nhánh phải, tầm xa được đánh dấu bằng các mũi tên đối với 1230, 3752, 8791, 7765 và 7775. Panen C mô tả vectơ hướng đích giống như vectơ hướng đích được thể hiện trong panen B, nhưng trong đó, catxet *Neo* được loại bỏ. Vectơ hướng đích này được sử dụng để hướng đích CD163 vào trong các tế bào mà chúng đã kháng lại neomycin. Đoạn mồi sử dụng trong các thử nghiệm khuyết đoạn nhỏ được minh họa bằng các mũi tên và GCD163F và GCD163R đánh dấu. Panen D nhấn mạnh các exon được hướng đích bởi các CRISPR. Vị trí của các CRISPR 10, 131, 256 và 282 được biểu thị bằng các mũi tên quay xuống trên exon 7. Số lượng CRISPR biểu thị số cặp bazơ từ điểm nối intron-exon của intron 6 và exon 7.

FIG. 2. Vectơ hướng đích và các CRISPR sử dụng để sửa đổi CD1D. Panen A mô tả các exon 3, 4, 5, 6 và 7 kiểu hoang dại của gen CD1D được hướng đích để sửa đổi bằng các CRISPR. Panen B thể hiện vectơ hướng đích được thiết kế để thay thế exon 3 bằng gen chỉ thị chọn lọc *Neo*. Vectơ hướng đích này được sử dụng kết hợp với các CRISPR để sửa đổi CD1D. Các đoạn mồi PCR trong thử nghiệm nhánh trái và nhánh phải, tầm xa được đánh dấu bằng các mũi tên đối với 3991, 4363, 7373 và 12806. Panen C mô tả các exon được hướng đích bởi các CRISPR. Vị trí của các CRISPR 4800, 5350, 5620 và 5626 được biểu thị bằng các mũi tên quay xuống trên exon 3. Các đoạn mồi sử dụng trong các thử nghiệm khuyết đoạn nhỏ được minh họa bằng các mũi tên và GCD1DF và GCD1DR đánh dấu.

FIG. 3. Thế hệ lợn bất hoạt CD163 và CD1D bởi CRISPR/Cas9 và SCNT. A) đoạn xóa hướng đích của CD163 trong các tế bào sinh dưỡng sau khi gây nhiễm bằng CRISPR/Cas9 và ADN cho. Kiểu gen hoang dại (WT) sinh ra dải 6545 cặp bazơ (bp). Các lằn 1–6 biểu thị 6 đàn khác nhau từ việc gây nhiễm duy nhất bằng CRISPR 10 với Cas9 và ADN cho chứa Neo. Các lằn 1, 4, và 5 thể hiện đoạn xóa đồng hợp tử lớn 1500–2000 bp. Lằn 2 biểu thị đoạn xóa đồng hợp tử nhỏ hơn. Các lằn 3 và 6 biểu thị alen WT và đoạn xóa nhỏ hoặc sự sửa đổi hai alen của cả hai alen. Những sự sửa đổi chính xác của mỗi một đàn chỉ được xác định bằng cách giải trình tự đối với các đàn sử dụng trong SCNT. Dải WT mờ nhạt ở một số lằn có thể biểu hiện sự lây nhiễm chéo của các nguyên bào sợi của thai từ đàn WT bên cạnh. NTC = đối chứng không có khuôn. B) Đoạn xóa hướng đích của CD1D trong các tế bào sinh dưỡng sau khi gây nhiễm bằng

CRISPR/Cas9 và ADN cho. Kiểu gen WT sinh ra dải 8729 bp. Các lần 1–4 biểu thị các đàn có đoạn xóa 500–2000 bp của CD1D. Lần 4 có vẻ là đàn WT. NTC = đối chứng không có khuôn. C) Hình ảnh của lợn bất hoạt CD163 được tạo ra bởi SCNT trong khi nghiên cứu. Lợn con đực này chứa đoạn xóa đồng hợp tử 1506 bp của CD163. D) Hình ảnh của lợn CD1D được tạo ra trong khi nghiên cứu. Các lợn con này chứa đoạn xóa 1653 bp của CD1D. E) Kiểu gen của hai lứa đẻ SCNT chứa đoạn xóa 1506 bp của CD163. Các lần 1–3 (lứa đẻ 63) và các lần 1–4 (lứa đẻ 64) biểu thị kiểu gen của mỗi một lợn con từ mỗi một lứa đẻ. Lợn nái chỉ ra con cái tiếp nhận các phôi SCNT và WT biểu thị đối chứng WT. NTC = đối chứng không có khuôn. F) Kiểu gen của hai lứa đẻ SCNT chứa đoạn khuyết 1653 bp của CD1D. Các lần 1–7 (lứa đẻ 158) và các lần 1–4 (lứa đẻ 159) biểu thị kiểu gen của mỗi một lợn con.

FIG. 4. Tác dụng của hệ CRISPR/Cas9 trong các phôi lợn. A) Tần suất hình thành túi phôi sau khi tiêm hệ CRISPR/Cas9 với nồng độ khác nhau vào trong các hợp tử. Độc tính của hệ CRISPR/Cas9 thấp nhất tại nồng độ 10ng/µl. B) Hệ CRISPR/Cas9 có thể làm gián đoạn thành công sự biểu hiện của eGFP trong các túi phôi khi đưa vào trong các hợp tử. Khuếch đại ban đầu X4. C) Các kiểu đột biến trên eGFP phát sinh sử dụng hệ CRISPR/Cas9: Kiểu gen WT (SEQ ID NO:16), #1 (SEQ ID NO:17), #2 (SEQ ID NO:18), và #3 (SEQ ID NO:19).

FIG. 5. Tác dụng của hệ CRISPR/Cas9 trong việc hướng đích CD163 trong phôi lợn. A) Các ví dụ về đột biến phát sinh trên CD163 bởi hệ CRISPR/Cas9: Kiểu gen WT (SEQ ID NO:20), #1-1 (SEQ ID NO:21), #1-4 (SEQ ID NO:22), và #2-2 (SEQ ID NO:23). Tất cả các phôi kiểm tra bằng giải trình tự ADN đều thể hiện đột biến trên CD163 (18/18). CRISPR 131 được in đậm. B) Đọc giải trình tự của đoạn xóa đồng hợp tử gây ra bởi hệ CRISPR/Cas9. Hình ảnh biểu thị # 1–4 từ Panen A mang đoạn xóa 2 bp của CD163.

FIG. 6. Tác dụng của hệ CRISPR/Cas9 khi được đưa vào bằng hai kiểu CRISPR. A) Khuếch đại PCR của CD163 trong các túi phôi được tiêm bằng CRISPR/ Cas9 ở dạng các hợp tử. Các lần 1, 3, 6, và 12 thể hiện đoạn xóa thiết kế giữa hai CRISPR khác nhau. B) Khuếch đại PCR của CD1D trong các túi phôi được tiêm bằng CRISPR/Cas9 ở dạng các hợp tử. CD1D có tần suất đoạn xóa thấp hơn khi được xác định bằng điện di trên gel khi so sánh với CD163 (3/23); các lần 1, 8, và 15 thể hiện các đoạn xóa hiển

nhiên trong CD1D. C) Hệ CRISPR/Cas9 hướng đích thành công hai gen khi hệ được cung cấp bằng hai CRISPR hướng đích CD163 và eGFP. Những sửa đổi của CD163 và eGFP được chứng minh: CD163 WT (SEQ ID NO:24), CD163 #1 (SEQ ID NO:25), CD163 #2 (SEQ ID NO:26), CD163 #3 (SEQ ID NO:27), eGFP WT (SEQ ID NO:28), eGFP #1-1 (SEQ ID NO:29), eGFP #1-2 (SEQ ID NO: 30), eGFP #2 (SEQ ID NO:31), và eGFP #3 (SEQ ID NO:32).

FIG. 7. Lợn bất hoạt CD163 phát sinh bởi hệ CRISPR/Cas9 được bơm vào trong các hợp tử. A) Khuếch đại PCR của CD163 từ các lợn bất hoạt; dấu hiệu xóa đoạn rõ ràng được phát hiện trong các đàn 67-2 và 67-4. B) Hình ảnh về lợn bất hoạt CD163 bằng gen thay thế. Tất cả các động đều khỏe mạnh và không thể hiện dấu hiệu bất thường nào. C) Kiểu gen của lợn bất hoạt CD163. Trình tự kiểu hoang dại (WT) được chứng minh là SEQ ID NO: 33. Hai động vật (từ các đàn 67-1 (SEQ ID NO:34) và 67-3 (SEQ ID NO:37)) mang đoạn xóa đồng hợp tử hoặc đoạn xen trong CD163. Hai động vật khác (từ các đàn 67-2 và 67-4) mang sửa đổi hai alen của CD163: #67-2 A1 (SEQ ID NO:35), #67-2 A2 (SEQ ID NO:36), #67-4 A1 (SEQ ID NO:38), và #67-4 a2 (SEQ ID NO:39). Xóa đoạn được gây ra bằng cách đưa vào hai CRISPR khác nhau bằng hệ Cas9. Không có động vật nào từ việc tiêm hợp tử của CD163 thể hiện kiểu gen khảm.

FIG. 8. Lợn bất hoạt CD1D phát sinh bởi hệ CRISPR/Cas9 được bơm vào trong các hợp tử. A) Khuếch đại PCR của CD1D từ các lợn bất hoạt; 166-1 thể hiện kiểu gen khảm của CD1D. 166-2, 166-3, và 166-4 không thể hiện sự thay đổi về kích thước của amplicon, nhưng việc giải trình tự của amplicon lộ suất sự sửa đổi. WT FF = các nguyên bào sợi của thai kiểu hoang dại. B) Khuếch đại PCR của thử nghiệm tầm dài thể hiện đoạn xóa rõ ràng của một alen ở các lợn con 166-1 và 166-2. C) Hình ảnh của các lợn bất hoạt CD1D bằng gen thay thế. D) Dữ liệu trình tự của lợn bất hoạt CD1D; WT (SEQ ID NO:40), #166-1.1 (SEQ ID NO: 41), #166-1.2 (SEQ ID NO:42), #166-2 (SEQ ID NO:43), #166-3.1 (SEQ ID NO:44), #166-3.2 (SEQID NO:45), và #166-4 (SEQ ID NO:46). Codon khởi đầu atg trong exon 3 được thể hiện bằng in đậm và cả chữ nhỏ.

FIG. 9. Các dấu hiệu lâm sàng trong nhiễm trùng PRRSV cấp. Các kết quả của việc đánh giá hàng ngày về sự có mặt của các dấu hiệu hô hấp và sốt đối với CD163 +/+ (n=6) và CD163 –/– (n=3).

FIG. 10. Mô bệnh học về phổi trong nhiễm trùng PRRSV cấp. Ảnh chụp hiển vi đại diện của các mô nhuộm H và E từ lợn hoang dại và lợn bất hoạt. Panen bên trái thể hiện phù và thâm nhiễm tế bào đơn nhân. Panen bên phải từ lợn bất hoạt thể hiện cấu trúc phổi của phổi bình thường.

FIG. 11. Virut huyết trong các kiểu gen khác nhau. Lưu ý rằng, dữ liệu về lợn con CD163-/- nằm dọc theo trục X.

FIG. 12. Sản xuất kháng thể ở các lợn kiểu hoang dại, vô hiệu và có alen không biểu thị đặc điểm.

FIG. 13. Sự biểu hiện trên bề mặt tế bào của CD163 ở từng lợn riêng biệt. Các dòng xuất hiện ở phía bên phải trong các panen A không biểu thị đặc điểm, B không biểu thị đặc điểm, và CD163 +/+ biểu thị kháng thể CD163 trong khi các dòng xuất hiện ở phía bên trái của các panen này là các đối chứng không kháng thể (nền). Lưu ý rằng, ở các động vật CD163-/-, nhuộm CD163 chồng với đối chứng đường nền và nhuộm CD163 ở các alen không biểu thị đặc điểm nằm ở khoảng giữa mức WT và mức nền (cũng lưu ý rằng, đây là thang log, do đó dưới ~10%).

FIG. 14. Mức CD169 ở các đại thực bào phế nang từ 3 lợn đại diện và đối chứng không có kháng thể (kháng CD169 đánh dấu bằng FITC).

FIG. 15. Virut huyết trong các kiểu gen khác nhau. Lưu ý rằng, dữ liệu về lợn con axit amin Δ43 nằm dọc theo trục X.

FIG. 16. Trình tự hệ gen của các exon CD163 kiểu hoang dại 7-10 sử dụng làm trình tự tham chiếu (SEQ ID NO: 47). Trình tự bao gồm đoạn thượng dòng 3000 bp của exon 7 tới bazơ cuối cùng của exon 10. Các vùng gạch chân thể hiện vị trí lần lượt của các exon 7, 8, 9, và 10.

FIG. 17. Sơ đồ của các biến đổi gen CD163 minh họa một số biến đổi gen CD163, sản phẩm protein dự đoán của mỗi một biến đổi và sự biểu hiện đại thực bào liên quan của mỗi một biến đổi, khi được đo bằng mức CD163 bề mặt ở các đại thực bào phế nang lợn (PAMs). Các vùng màu đen chỉ ra các intron và các vùng màu trắng chỉ ra các exon. Vùng gạch bóng chỉ ra h*CD163L1* exon 11 mô phỏng, tương tự exon 7 của lợn. Vùng màu xám chỉ ra intron tổng hợp bằng cấu trúc PGK Neo.

FIG. 18. Sơ đồ của protein CD163 lợn và trình tự gen. A) protein CD163 các vùng SRCR (hình ovan) và PST (hình vuông) cùng với các exon gen tương ứng. B) So sánh CD163 SRCR 5 lợn (SEQ ID NO: 120) với CD163L1 SRCR 8 người (SEQ ID NO: 121) tương tự.

FIG. 19. Các kết quả đại diện về sự biểu hiện trên bề mặt của CD163 và CD169 trên PAMs từ các lợn kiểu hoang dại và biến đổi CD163. Các panen A–E thể hiện kết quả của biến đổi gen CD163 như được minh họa trong Fig. 17. Dữ liệu tập hợp của d7(1467) và d7(1280) được thể hiện trong panen D.

FIG. 20. Các mức haptoglobin huyết thanh ở lợn kiểu hoang dại và biến đổi CD163.

FIG. 21. Tính cho phép tương đối của các PAM kiểu hoang dại và *HL11m* với sự gây nhiễm bởi các thể phân lập PRRSV Typ 2.

FIG. 22. Gây nhiễm lợn biến đổi CD163 bằng các thể phân lập PRRSV Typ 1 và Typ 2.

FIG. 23. Tải lượng virut của lợn biến đổi WT và CD163 được gây nhiễm bằng các virut Typ 2.

Mô tả chi tiết sáng chế

Sáng chế đề cập đến động vật và phương pháp để tạo ra động vật được chỉnh sửa gen, động vật này có các biến đổi về gen CD163 và kháng lại nhiễm trùng do PRRSV và các nhiễm trùng do virut đường hô hấp liên quan khác. Động vật có các biến đổi về nhiễm sắc thể (xen đoạn hoặc xóa đoạn) mà chúng bất hoạt hoặc theo cách khác, điều biến hoạt tính gen CD163. CD163 là cần thiết đối với PRRSV để xâm nhập vào tế bào và để sao chép virut. Do đó, động vật có CD163 vô hiệu thể hiện sự kháng lại nhiễm trùng PRRSV khi được thử thách. Các động vật này có thể được tạo ra sử dụng phương thức bất kỳ trong số nhiều phương thức tận dụng sự chỉnh sửa gen.

Sáng chế cũng đề cập đến phương pháp tạo ra lợn bao gồm bước đưa vào tế bào lợn hoặc phôi lợn một tác nhân mà nó liên kết đặc hiệu với vị trí đích trên nhiễm sắc thể của tế bào và gây ra sự đứt ADN sợi kép hoặc theo cách khác, bất hoạt hoặc làm giảm hoạt tính của gen CD163 hoặc protein trong đó sử dụng các phương pháp chỉnh sửa gen như các phương pháp "trình tự sắp xếp kiểu lặp đi lặp lại xen khoảng trống"

(Clustered Regularly Interspaced Short Palindromic Repeats: CRISPR)/Cas system, men nucleaza tác động giống chất kích hoạt phiên mã (Sự phiên mã Activator-Like Effector Nucleases: TALENs), nucleaza ngón tay kẽm (Zinc Finger Nucleases: ZFN), protein dung hợp recombinaza, hoặc meganucleaza.

Cũng được mô tả ở đây là việc sử dụng một hoặc nhiều vị trí gen CD163 cụ thể kiểu tiếp đôi gen trước gen sau cùng với polypeptit có khả năng thực hiện việc phân cắt và/hoặc gắn các trình tự axit nucleic đặc hiệu bên trong các vị trí gen CD163. Ví dụ về sử dụng các vị trí gen CD163 kiểu tiếp đôi gen trước gen sau cùng với polypeptit hoặc ARN có khả năng thực hiện việc phân cắt và/hoặc gắn các vị trí gen CD163 bao gồm polypeptit được chọn từ nhóm bao gồm các protein ngón tay kẽm, meganucleaza, các vùng TAL, TALENs, các men recombinaza CRISPR/Cas dẫn hướng bằng ARN, khóa kéo leuxin, và các loại khác đã biết đối với chuyên gia trong lĩnh vực. Các ví dụ cụ thể bao gồm protein khảm ("dung hợp") bao gồm polypeptit vùng liên kết ADN đặc hiệu vị trí và polypeptit vùng phân cắt (ví dụ, nucleaza), như protein ZFN bao gồm polypeptit ngón tay kẽm và polypeptit nucleaza FokI. Được mô tả ở đây là các polypeptit bao gồm vùng liên kết ADN mà vùng này liên kết đặc hiệu với gen CD163. Polypeptit này cũng có thể bao gồm vùng hoặc nửa vùng nucleaza (phân cắt) (ví dụ, endonucleaza "tìm nhà", bao gồm endonucleaza tìm nhà có vùng liên kết ADN sửa đổi), và/hoặc vùng ligaza, sao cho polypeptit có thể gây ra đứt sợi kép hướng đích, và/hoặc tạo thuận lợi cho sự tái tổ hợp của axit nucleic mong muốn tại vị trí đứt gãy. Vùng liên kết ADN mà vùng này hướng đích vào vị trí gen CD163 có thể là vùng chức năng phân cắt ADN. Các polypeptit nêu trên có thể được sử dụng để đưa một axit nucleic ngoại sinh vào trong hệ gen của sinh vật chủ (ví dụ, loài động vật) tại một hoặc nhiều vị trí gen CD163. Vùng liên kết ADN có thể bao gồm protein ngón tay kẽm với một hoặc nhiều ngón tay kẽm (ví dụ, 2, 3, 4, 5, 6, 7, 8, 9 hoặc nhiều ngón tay kẽm), nó được thiết kế (không xuất hiện một cách tự nhiên) để liên kết với trình tự bất kỳ bên trong gen CD163. Bất kỳ trong số các protein ngón tay kẽm mô tả ở đây có thể liên kết với một vị trí đích bên trong trình tự mã hóa của gen đích hoặc bên trong các trình tự liền kề (ví dụ, đoạn khởi động hoặc các yếu tố biểu hiện khác). Protein ngón tay kẽm có thể liên kết với vị trí đích trong gen CD163.

Phần định nghĩa

Đơn vị, tiền tố và ký hiệu có thể được biểu thị ở dạng được chấp nhận theo hệ tiêu chuẩn quốc tế (SI) của chúng. Trừ khi được chỉ rõ theo cách khác, các axit nucleic được viết từ trái sang phải theo chiều 5' sang 3' và các trình tự axit amin được viết từ trái sang phải theo chiều amino sang carboxy. Các khoảng số học viện dẫn trong bản mô tả bao gồm cả các số phân định khoảng và bao gồm cả mỗi một số nguyên bên trong khoảng xác định. Các axit amin có thể được đề cập ở đây bằng các ký hiệu ba chữ cái đã biết thông thường của chúng hoặc bằng các ký hiệu một chữ cái được khuyến cáo bởi Ủy ban danh pháp IUPAC-IUB Biochemical Nomenclature Commission. Tương tự, các nucleotit có thể được đề cập bằng mã chữ cái đơn được chấp nhận phổ biến của chúng. Trừ khi được cung cấp theo cách khác, các thuật ngữ về phần mềm, điện và điện tử khi được sử dụng ở đây như được định nghĩa trong tài liệu The New IEEE Standard Dictionary of Electrical and Electronics Terms (5th edition, 1993). Các thuật ngữ định nghĩa dưới đây được định nghĩa đầy đủ hơn bằng việc tham chiếu toàn bộ bản mô tả.

Theo cách hiểu của một người có kỹ năng trong lĩnh vực, với mục đích bất kỳ và tất cả mọi mục đích, nhất là liên quan đến phần mô tả, toàn bộ các khoảng viện dẫn ở đây cũng bao hàm bất kỳ và tất cả các phân khoảng có thể và tổ hợp các phân khoảng của chúng, cũng như từng giá trị riêng biệt tạo nên khoảng, cụ thể là các giá trị nguyên. Một khoảng viện dẫn bao gồm cả mỗi một giá trị cụ thể, số nguyên, thập phân hoặc đồng nhất bên trong khoảng. Khoảng liệt kê bất kỳ có thể được nhận diện dễ dàng là mô tả đầy đủ và cho phép khoảng tương tự được phân nhỏ thành ít nhất một phần hai, phần ba, phần bốn, phần năm hoặc phần mười bằng nhau. Đối với ví dụ không hạn chế, mỗi một khoảng thảo luận ở đây có thể được phân nhỏ dễ dàng phần ba dưới, phần ba giữa và phần ba trên, v.v..

Trong sáng chế này, cụm từ chứa danh từ và các từ "nêu trên", "đã nêu" được dự tính là gồm một hoặc nhiều danh từ tương ứng. Các thuật ngữ "chứa", "bao gồm" và "có" được dự tính bao gồm cả và có nghĩa là có thể có các yếu tố bổ sung ngoài các yếu tố đã liệt kê.

Thuật ngữ "và/hoặc" có nghĩa là bất kỳ một trong số các hạng mục, tổ hợp bất kỳ của các hạng mục hoặc toàn bộ các hạng mục mà thuật ngữ này kết hợp với. Câu "một hoặc nhiều" được hiểu một cách dễ dàng bởi một người có kỹ năng trong lĩnh vực, nhất là khi được đọc trong ngữ cảnh sử dụng câu nói trên.

"Protein liên kết" là một protein có khả năng liên kết với phân tử khác. Protein liên kết có thể liên kết với, ví dụ, phân tử ADN (protein liên kết ADN), phân tử ARN (protein liên kết ARN) và/hoặc phân tử protein (protein liên kết protein). Trong trường hợp protein liên kết protein, nó có thể liên kết với chính nó (để tạo ra các homodime, homotrime, v.v..) và/hoặc nó có thể liên kết với một hoặc nhiều phân tử của protein hoặc các protein khác. Protein liên kết có thể có nhiều hơn một kiểu hoạt tính liên kết. Ví dụ, protein ngón tay kẽm có hoạt tính liên kết ADN, hoạt tính liên kết ARN và hoạt tính liên kết protein.

Thuật ngữ "các thể biến dị sửa đổi bảo tồn" áp dụng cho cả trình tự axit amin và trình tự axit nucleic. Liên quan đến các trình tự axit nucleic cụ thể, "các thể biến dị sửa đổi bảo tồn" đề cập đến các axit nucleic mà chúng mã hóa cho các thể biến dị tương đồng hoặc sửa đổi bảo tồn của các trình tự axit amin. Do sự suy biến của mã di truyền, nên nhiều axit nucleic tương đồng về chức năng mã hóa cho protein xác định bất kỳ. Ví dụ, các đơn vị mã GCA, GCC, GCG và GCU tất cả đều mã hóa cho axit amin alanine. Do đó, tại mọi vị trí mà ở đó, alanine được chỉ rõ bằng một đơn vị mã, đơn vị mã có thể được thay đổi sang đơn vị mã tương ứng bất kỳ được mô tả mà không làm thay đổi polypeptit được mã hóa. Các biến dị axit nucleic như vậy là "biến dị câm" và biểu hiện một loài có biến dị sửa đổi bảo tồn. Mọi trình tự axit nucleic ở đây mà mã hóa cho polypeptit, bằng sự tham chiếu mã di truyền, cũng mô tả mọi biến dị câm có thể của axit nucleic.

Người có hiểu biết trung bình trong lĩnh vực này sẽ hiểu rằng, mỗi một đơn vị mã trong axit nucleic (ngoại trừ AUG, đây thường là đơn vị mã duy nhất của methionin; và UGG, đây thường là đơn vị mã duy nhất của tryptophan) có thể được sửa đổi để sinh ra một phân tử tương đồng về mặt chức năng. Do đó, mỗi một biến dị câm của axit nucleic mà mã hóa cho polypeptit theo sáng chế là bao hàm trong mỗi một trình tự polypeptit được mô tả và nằm trong phạm vi bảo hộ của sáng chế.

Đối với các trình tự axit amin, một người có kỹ năng sẽ hiểu rằng, từng sự thế đoạn, xóa đoạn hoặc thêm đoạn vào trình tự axit nucleic, peptit, polypeptit, hoặc protein mà làm thay đổi, bổ sung hoặc xóa một axit amin đơn hoặc một tỷ lệ nhỏ các axit amin trong trình tự được mã hóa đều là "thể biến dị sửa đổi bảo tồn" trong đó sự thay đổi dẫn đến thay thế axit amin bằng axit amin tương tự về hóa học. Do đó, ví dụ, số lượng gốc

axit amin bất kỳ được lựa chọn từ nhóm các số nguyên bao gồm từ 1 đến 15 có thể cũng được thay đổi. Do đó, ví dụ, 1, 2, 3, 4, 5, 7, hoặc 10 sửa đổi có thể được thực hiện.

Các thể biến dị sửa đổi bảo tồn điển hình cung cấp hoạt tính sinh học tương tự như trình tự polypeptit chưa sửa đổi mà chúng xuất phát từ trình tự này. Ví dụ, tính đặc hiệu cơ chất, hoạt tính enzym, hoặc liên kết phối tử/thụ thể thông thường ít nhất là 30%, 40%, 50%, 60%, 70%, 80%, hoặc 90% của protein tự nhiên đối với cơ chất tự nhiên của nó. Các bảng về sự thế bảo tồn cung cấp các axit amin tương tự về chức năng đã được biết rõ trong lĩnh vực.

Sáu nhóm sau, mỗi nhóm chứa các axit amin là những sự thế bảo tồn của một nhóm khác: [1] Alanine (A), Serine (S), Threonine (T); [2] axit aspartic (D), axit glutamic (E); [3] Asparagine (N), Glutamine (Q); [4] Arginine (R), Lysine (K); [5] Isoleucine (I), Leucine (L), Methionin (M), Valine (V); và [6] Phenylalanine (F), Tyrosine (Y), Tryptophan (W). *See also*, Creighton (1984) Proteins W. H. Freeman and Company.

Thuật ngữ "CRISPR" là viết tắt của "clustered regularly interspaced short palindromic repeats." Thuật ngữ "Cas9" đề cập đến "protein kết hợp CRISPR 9." Các thuật ngữ "CRISPR/Cas9" hoặc "hệ CRISPR/Cas9" đề cập đến hệ nucleaza có thể lập trình được trong kỹ thuật di truyền mà bao gồm protein Cas9, hoặc dẫn xuất của nó, và một hoặc nhiều ARN không mã hóa mà có thể cung cấp chức năng của ARN CRISPR (crARN) và ARN kích hoạt trans (tracrARN) của Cas9. crARN và tracrARN có thể được sử dụng riêng biệt hoặc có thể được kết hợp để tạo ra "ARN định hướng" (gARN). crARN hoặc gARN cung cấp trình tự bổ trợ cho đích hệ gen. Các hệ CRISPR/Cas9 được mô tả thêm dưới đây.

Ở đây các tham chiếu đến sự xóa đoạn trong trình tự nucleotit từ nucleotit x tới nucleotit y có nghĩa là, toàn bộ các nucleotit nằm trong khoảng này đã bị xóa, bao gồm cả x và y. Do đó, ví dụ, câu "đoạn xóa 11 cặp bazơ từ nucleotit 3137 đến nucleotit 3147 khi so sánh với SEQ ID NO: 47" có nghĩa là, mỗi một trong số các nucleotit từ 3317 đến 3147 đã được xóa, bao gồm cả nucleotit 3317 và 3147.

"Tính kháng bệnh" là đặc tính của động vật, trong đó động vật tránh được các triệu chứng bệnh mà chúng là hậu quả của các tương tác động vật - tác nhân gây bệnh, như các tương tác giữa lợn và PRRSV. Tức là, các tác nhân gây bệnh được ngăn chặn

không gây ra bệnh ở động vật và các triệu chứng bệnh kết hợp, hoặc theo cách khác, làm giảm tỷ lệ mắc và/hoặc mức độ nặng của các dấu hiệu lâm sàng hoặc làm giảm các triệu chứng lâm sàng. Một người có kỹ năng trong lĩnh vực sẽ hiểu rằng, các chế phẩm và phương pháp bộc lộ ở đây có thể được sử dụng cùng với các chế phẩm và phương pháp khác có sẵn trong lĩnh vực để bảo vệ động vật tránh sự tấn công của tác nhân gây bệnh.

Thuật ngữ "mã hóa" hoặc "được mã hóa", liên quan đến axit nucleic đặc trưng, có nghĩa bao gồm thông tin về dịch mã thành protein đặc trưng. Một axit nucleic mã hóa cho protein có thể bao gồm các trình tự can thiệp (ví dụ, intron) bên trong các vùng được dịch mã của axit nucleic, hoặc có thể không có các trình tự can thiệp không dịch mã này (ví dụ, như trong cADN). Thông tin mà nhờ nó, protein được mã hóa được xác định nhờ sử dụng các đơn vị mã. Điển hình, trình tự axit amin được mã hóa bởi axit nucleic sử dụng mã di truyền "toàn năng". Khi axit nucleic được bào chế hoặc làm thay đổi theo phương pháp tổng hợp, lợi thế có thể được lấy từ các ưu tiên đơn vị mã đã biết của vật chủ dự định mà axit nucleic cần được biển hiện trong vật chủ này.

Như được sử dụng ở đây, "chỉnh sửa gen," "được chỉnh sửa gen" "được chỉnh sửa về di truyền" và "tác nhân chỉnh sửa gen" đề cập đến việc sử dụng công nghệ tìm nhà bằng các nucleaza xuất hiện tự nhiên hoặc được thiết kế nhân tạo, cũng được gọi là "các kéo cắt phân tử, "các endonucleaza tìm nhà," hoặc "các endonucleaza hướng đích." Các nucleaza tạo ra những sự đứt gãy nhiễm sắc thể sợi kép đặc hiệu (DSBs) tại các vị trí mong muốn trong hệ gen, trong một số trường hợp, nó kiểm soát các cơ chế nội sinh của tế bào để sửa chữa sự đứt gãy cảm ứng bởi các quá trình tái tổ hợp tương đồng tự nhiên (HR) và/hoặc quá trình nối đầu không tương đồng (NHEJ). Các tác nhân chỉnh sửa gen bao gồm nucleaza ngón tay kẽm (ZFNs), men nucleaza tác động giống chất kích hoạt phiên mã (TALENs), hệ trình tự sắp xếp kiểu lặp đi lặp lại xen khoảng trống/CAS9 (CRISPR/Cas9), và meganucleaza (ví dụ, meganucleaza tái thiết kế ở dạng endonucleaza tìm nhà). Các thuật ngữ cũng bao gồm việc sử dụng các phương thức và kỹ thuật chuyển gen, bao gồm, ví dụ, trường hợp sự thay đổi là xóa đoạn hoặc xen đoạn tương đối nhỏ (điển hình dưới 20nt) và/hoặc không đưa ADN vào từ các loài ngoại lai. Thuật ngữ cũng bao gồm động vật thế hệ con như động vật được tạo ra bởi giao phối chéo hoặc nhân giống vô tính từ động vật được chỉnh sửa gen ban đầu.

Như được sử dụng ở đây, "khác loại" tham chiếu tới axit nucleic là axit nucleic mà nó có nguồn gốc từ loại lạ hoặc nếu từ cùng loài, thì cơ bản được sửa đổi từ dạng từ nhiên của nó về thành phần và/hoặc vị trí gen bởi sự can thiệp chủ tâm của con người. Ví dụ, đoạn khởi động liên kết theo kiểu hoạt động được với gen cấu trúc khác loại là từ loài khác với loài mà gen cấu trúc được xuất phát từ loài này hoặc nếu từ cùng loài, thì một hoặc cả hai cơ bản được sửa đổi từ dạng ban đầu của chúng. Protein khác loại có thể có nguồn gốc từ loài ngoại lai hoặc nếu từ cùng loài, thì cơ bản được sửa đổi từ dạng ban đầu của nó bởi sự can thiệp chủ tâm của con người.

Như được sử dụng ở đây "công nghệ ADN tìm nhà," "công nghệ tìm nhà" và "endonucleaza tìm nhà" bao gồm cơ chế bất kỳ mà cho phép phân tử đặc hiệu được hướng đích vào trình tự ADN đặc hiệu bao gồm các protein ngón tay kẽm (ZF), các meganucleaza tác động giống chất kích hoạt phiên mã (Sự phiên mã Activator-Like Effectors - TALEs), và hệ CRISPR/Cas9.

Các thuật ngữ "tính kháng tăng" và "tính nhạy giảm" ở đây có nghĩa, nhưng không bị giới hạn bởi, sự giảm có ý nghĩa thống kê về tỷ lệ mắc và/hoặc mức độ nặng của các dấu hiệu lâm sàng hoặc các triệu chứng lâm sàng được kết hợp với nhiễm trùng bởi tác nhân gây bệnh. Ví dụ, "tính kháng tăng" hoặc "tính nhạy giảm" có thể đề cập đến sự giảm có ý nghĩa thống kê về tỷ lệ mắc và/hoặc mức độ nặng của các dấu hiệu lâm sàng hoặc các triệu chứng lâm sàng được kết hợp với nhiễm trùng bởi PRRSV ở động vật chứa ít nhất một trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163 khi so sánh với động vật đối chứng có trình tự nhiễm sắc thể chưa sửa đổi. Thuật ngữ "sự giảm có ý nghĩa thống kê của các triệu chứng lâm sàng" có nghĩa là, nhưng không bị giới hạn bởi, tần suất về tỷ lệ mắc của ít nhất một triệu chứng lâm sàng ở nhóm các đối tượng được chỉnh sửa ít nhất 10%, tốt hơn ít nhất 20%, tốt hơn nữa ít nhất 30%, thậm chí tốt hơn nữa ít nhất 50%, và thậm chí tốt hơn nữa ít nhất 70% thấp hơn so với nhóm đối chứng không chỉnh sửa sau khi thử thách bằng tác nhân gây nhiễm trùng.

Như được sử dụng ở đây, thuật ngữ "gắn gen" có nghĩa là thay thế gen nội sinh bằng gen chuyển hoặc bằng cùng gen nội sinh với một số sửa đổi về cấu trúc, nhưng giữ nguyên sự điều khiển phiên mã của gen nội sinh.

"Bất hoạt gen" có nghĩa là phá vỡ cấu trúc hoặc cơ chế điều hòa của gen. Bất hoạt gen có thể được phát sinh thông qua tái tổ hợp tương đồng của các vectơ hướng

đích, vectơ thay thế hoặc vectơ tìm tới và rời đi hoặc sự xen ngẫu nhiên của vectơ bẫy gen dẫn đến sự mất chức năng gen hoàn toàn một phần hoặc có điều kiện.

Thuật ngữ "động vật" bao gồm động vật bất kỳ không phải người, ví dụ động vật nuôi trong gia đình (ví dụ, động vật nuôi). Thuật ngữ "động vật nuôi" bao gồm động vật bất kỳ được nuôi theo cách truyền thống trong các trang trại , ví dụ lợn, bò (ví dụ, bò sữa), cừu, dê, ngựa (ví dụ, ngựa hoặc lừa), trâu, lạc đà, hoặc động vật thuộc loài chim (ví dụ, gà, gà tây, vịt, ngỗng, gà Nhật, hoặc bồ câu non). Thuật ngữ "động vật nuôi" không bao gồm chuột cống, chuột nhắt hoặc các động vật gặm nhấm khác.

Như được sử dụng ở đây, thuật ngữ "đột biến" bao gồm những thay đổi về trình tự nucleotit của polynucleotit, ví dụ như trình tự gen hoặc ADN mã hóa (CDS), so với trình tự kiểu hoang dại. Thuật ngữ bao gồm, không giới hạn, thay thế, xen đoạn, xê dịch khung, xóa đoạn, đảo đoạn, chuyển đoạn, lặp đoạn, các đột biến vị trí ghép nối - cho, đột biến điểm và tương tự.

Như được sử dụng ở đây "liên kết theo kiểu hoạt động được" bao gồm tham chiếu tới sự liên kết chức năng giữa hai trình tự axit nucleic, ví dụ trình tự khởi động và trình tự thứ hai, trong đó trình tự khởi động khởi đầu và trung gian cho sự phiên mã của trình tự ADN tương ứng với trình tự thứ hai. Nói chung, liên kết theo kiểu hoạt động được có nghĩa là, các trình tự axit nucleic liên kết là tiếp giáp nhau và ở đó, sự nối cần thiết hai vùng mã hóa cho protein, theo cách liền kề nhau và trong cùng khung đọc.

Như được sử dụng ở đây, "polynucleotit" bao gồm tham chiếu tới deoxyribopolynucleotit, ribopolynucleotit, hoặc các thể biến dị sửa đổi bảo tồn; thuật ngữ cũng có thể đề cập đến các chất tương tự của chúng mà có bản chất cốt yếu của ribonucleotit tự nhiên ở chỗ, chúng lai, trong các điều kiện lai nghiêm ngặt, với trình tự nucleotit cơ bản tương tự như các nucleotit xuất hiện tự nhiên và/hoặc cho phép dịch mã thành (các) axit amin tương tự như (các) nucleotit xuất hiện tự nhiên . Polynucleotit có thể có chiều dài đầy đủ hoặc trình tự con của gen tự nhiên hoặc gen cấu trúc khác loại hoặc gen điều hòa. Trừ khi được chỉ rõ theo cách khác, thuật ngữ bao gồm tham chiếu tới trình tự đặc hiệu cũng như trình tự bổ trợ của nó. Do đó, các ADN hoặc ARN có các khung được sửa đổi về độ ổn định hoặc về các lý do khác là "các polynucleotit" như thuật ngữ này được dự định ở đây. Ngoài ra, các ADN hoặc ARN bao gồm các bazơ không phổ biến, như inosine, hoặc các bazơ sửa đổi, như các bazơ trityl hóa, để

gọi tên chính xác hai ví dụ, là polynucleotit như thuật ngữ được sử dụng ở đây. Sẽ được hiểu rằng, rất nhiều sửa đổi đã được thực hiện đối với ADN và ARN dùng cho nhiều mục đích hữu dụng khác nhau đã biết đối với người có kỹ năng trong lĩnh vực.

Thuật ngữ polynucleotit như được sử dụng ở đây bao hàm cả các dạng sửa đổi theo phương thức hóa học, enzym hoặc theo phương thức chuyển hóa của các polynucleotit, cũng như các dạng hóa học của ADN và ARN đặc trưng của các virut và tế bào, bao gồm trong số các thứ khác, các tế bào đơn và các tế bào phức hợp.

Các thuật ngữ "polypeptit", "peptit" và "protein" được sử dụng thay đổi lẫn nhau ở đây đề cập đến polyme của các gốc axit amin. Các thuật ngữ cũng có thể áp dụng cho các thể biến dị sửa đổi bảo tồn và cho các polyme axit amin trong đó, một hoặc nhiều gốc axit amin là gốc hóa học nhân tạo của axit amin xuất hiện tự nhiên tương ứng, cũng như đối với các polyme axit amin xuất hiện tự nhiên. Bản chất cốt yếu của các chất tương tự như vậy của axit amin xuất hiện tự nhiên là ở chỗ, khi được kết hợp vào trong protein, thì protein có tính phản ứng đặc hiệu với các kháng thể gây ra bởi cùng protein nhưng hoàn toàn bao gồm các axit amin xuất hiện tự nhiên.

Các thuật ngữ "polypeptit", "peptit" và "protein" cũng bao hàm cả các sửa đổi bao gồm, nhưng không bị giới hạn bởi, glycosyl hóa, gắn lipit, sulfat hóa, gama-carboxyl hóa của các gốc axit glutamic, hydroxyl hóa và ADP-ribosyl hóa. Sẽ được hiểu rằng, như đã biết và như đã lưu ý ở trên, các polypeptit không phải luôn luôn hoàn toàn mạch thẳng. Ví dụ, polypeptit có thể được phân nhánh là kết quả của ubiquitin hóa, và chúng có thể mạch vòng, có hoặc không có phân nhánh, thường là kết quả của các sự kiện sau dịch mã, bao gồm sự kiện và các sự kiện xử lý tự nhiên gây ra bởi thao tác của con người không xuất hiện trong tự nhiên. Các polypeptit mạch vòng, nhánh và vòng nhánh có thể được tổng hợp bằng quá trình tự nhiên không dịch mã và cũng như bằng các phương pháp tổng hợp hoàn toàn. Ngoài ra, sáng chế này dự định việc sử dụng cả các thể biến thể amin tận cùng chứa methionin và ít methionin của protein theo sáng chế.

Ở đây, thuật ngữ "giảm về tỷ lệ mắc và/hoặc mức độ nặng của các dấu hiệu lâm sàng" hoặc "giảm về các triệu chứng lâm sàng" có nghĩa là, nhưng không bị giới hạn bởi, giảm về số lượng đối tượng bị nhiễm trong nhóm, giảm hoặc loại trừ số lượng đối tượng biểu hiện các dấu hiệu lâm sàng của nhiễm trùng hoặc giảm về mức độ nặng của

các dấu hiệu lâm sàng bất kỳ có mặt trong một hoặc nhiều đối tượng, so với nhiễm trùng kiểu hoang dại. Ví dụ, các thuật ngữ này bao gồm các dấu hiệu lâm sàng bất kỳ trong số nhiễm trùng, bệnh ở phổi, virut huyết, sản xuất kháng thể, giảm tải lượng tác nhân gây bệnh, sự thải loại tác nhân gây bệnh, giảm về sự lan truyền tác nhân gây bệnh, hoặc giảm về dấu hiệu triệu chứng lâm sàng bất kỳ của PRRSV. Tốt hơn các dấu hiệu lâm sàng này được giảm ở một hoặc nhiều động vật theo sáng chế tới ít nhất 10% so với các đối tượng không có sửa đổi về gen CD163 và trở nên bị nhiễm. Tốt hơn nữa các dấu hiệu lâm sàng được giảm ở các đối tượng theo sáng chế tới ít nhất 20%, tốt hơn ít nhất 30%, tốt hơn nữa ít nhất 40%, và thậm chí tốt hơn nữa ít nhất 50%.

Các thuật ngữ "gốc" hoặc "gốc axit amin" hoặc "axit amin" được sử dụng thay đổi lẫn nhau ở đây đề cập đến axit amin mà nó được kết hợp vào trong protein, polypeptit, hoặc peptit (gọi chung là "protein"). Axit amin có thể là axit amin xuất hiện tự nhiên và trừ khi được giới hạn theo cách khác, có thể bao gồm các chất tương tự không tự nhiên của các axit amin tự nhiên mà có thể hoạt động theo cách tương tự như các axit amin xuất hiện tự nhiên.

Thuật ngữ "lai chọn lọc" bao gồm tham chiếu đến phương pháp lai, trong các điều kiện lai nghiêm ngặt, của trình tự axit nucleic với trình tự axit nucleic khác hoặc các sản phẩm sinh học khác. Khi sử dụng hệ phát hiện dựa vào sự lai, đoạn dò axit nucleic được chọn là đoạn bổ trợ cho trình tự axit nucleic tham chiếu, và sau đó, bằng cách lựa chọn các điều kiện thích hợp, đoạn dò và trình tự tham chiếu lai hoặc liên kết theo cách chọn lọc với nhau để tạo ra phân tử phức kép.

Thuật ngữ "điều kiện nghiêm ngặt" hoặc "điều kiện lai nghiêm ngặt" bao gồm tham chiếu đến các điều kiện trong đó, đoạn dò sẽ lai với trình tự đích của nó ở mức độ lớn hơn theo cách có thể phát hiện so với các trình tự khác (ví dụ, ít nhất 2 lần so với mốc cơ sở). Các điều kiện nghiêm ngặt là phụ thuộc trình tự và sẽ khác nhau trong các tình huống khác nhau. Bằng cách kiểm soát độ nghiêm ngặt của phép lai và/hoặc điều kiện rửa, trình tự đích có thể được nhận diện có mức bổ trợ 100% với đoạn dò (dò tương đồng).

Theo cách khác, các điều kiện nghiêm ngặt có thể được điều chỉnh để cho phép sự không tương hợp nào đó về các trình tự sao cho mức độ tương đồng thấp hơn được

phát hiện (dò dị tương đồng). Nói chung, đoạn dò có độ dài dưới 1000 nucleotit, tuỳ ý có độ dài dưới 500 nucleotit.

Điển hình, điều kiện nghiêm ngặt sẽ là các điều kiện, trong đó nồng độ muối là dưới 1,5 M ion Na, tốt hơn nồng độ từ 0,01 đến 1,0 M ion Na (hoặc các muối khác) ở độ pH từ 7,0 đến 8,3 và nhiệt độ ít nhất khoảng 30°C đối các đoạn dò ngắn (*ví dụ*, từ 10 đến 50 nucleotit) và ít nhất khoảng 60°C đối với các đoạn dò dài (ví dụ, trên 50 nucleotit). Các điều kiện nghiêm ngặt cũng có thể đạt được bằng cách bổ sung các chất làm mất ổn định như formamit. Tính đặc hiệu điển hình là chức năng của dịch rửa sau khi lai, các yếu tố quyết định là cường độ ion và nhiệt độ của dung dịch rửa cuối cùng. Đối với các đoạn lai ADN /DNA, điểm nóng chảy nhiệt (Tm) có thể được tính xấp xỉ từ phương trình của Meinkoth and Wahl, Anal. Biochem., 138: 267-284 (1984): T_m [°C] = 81,5 + 16,6 (log M) + 0,41(% GC)-0,61 (dạng %)-500/L; trong đó M là nồng độ mol của các cation hóa trị một, % GC là tỷ lệ % của các nucleotit guanosine và cytosine trong ADN, dạng % là tỷ lệ % của formamit trong dung dịch lai và L là độ dài của đoạn lai trong các cặp bazơ. T_m là nhiệt độ (theo cường độ ion và độ pH xác định) tại đó, 50% các trình tự đích bổ trợ lai với đoạn dò tương hợp hoàn toàn. Tm được giảm khoảng 1°C trong mỗi 1% không tương hợp; do đó, các điều kiện Tm, lai và/hoặc điều kiện rửa có thể được điều chỉnh để lai với các trình tự có độ tương đồng mong muốn. Ví dụ, nếu các trình tự với độ tương đồng > 90% được tìm kiếm, thì T_m có thể giảm 10°C. Nói chung, các điều kiện nghiêm ngặt được lựa chọn khoảng 5°C thấp hơn T_m đối với trình tự đặc hiệu và bổ thể của nó tại cường độ ion và độ pH xác định. Tuy nhiên, các điều kiện đặc biệt nghiêm ngặt có thể sử dụng phép lai và/hoặc rửa ở nhiệt độ từ 1 đến 4°C thấp hơn T_m; các điều kiện nghiêm ngặt vừa phải có thể sử dụng phép lai và/hoặc rửa ở nhiệt độ từ 6 đến 10°C thấp hơn T_m; các điều kiện nghiêm ngặt thấp có thể sử dụng phép lai và/hoặc rửa ở nhiệt độ từ 11 đến 20°C thấp hơn T_m. Sử dụng phương trình, phép lai và các thành phần rửa, và T_m mong muốn, người có hiểu biết trung bình trong lĩnh vực này sẽ hiểu rằng, những thay đổi về độ nghiêm ngặt của phép lai và/hoặc dung dịch rửa đã được mô tả. Một hướng dẫn bao quát đối với phép lai của các axit nucleic được tìm thấy trong tài liệu Tijssen, Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with axit nucleic Probes, Part I, Chapter 2 "Overview of principles of hybridization and the strategy of axit nucleic probe assays",

Elsevier, New York (1993); và tài liệu Current Protocols in Molecular Biology, Chapter 2, Ausubel, *et al.*, Eds., Greene Publishing và Wiley-Interscience, New York (1995).

"Miền liên kết ADN TALE" hoặc "TALE" là một polypeptit bao gồm một hoặc nhiều miền / đơn vị lặp TALE. Các miền lặp liên quan đến liên kết của TALE với trình tự ADN đích cùng nguồn gốc của nó. "Đơn vị lặp" đơn (còn được gọi là "đoạn lặp") điển hình có độ dài 33-35 axit amin và thể hiện ít nhất sự tương đồng trình tự nào đó với các trình tự lặp TALE khác bên trong protein TALE xuất hiện tự nhiên. Các miền ngón tay kẽm và liên kết TALE có thể được "thiết kế" để liên kết với trình tự nucleotit định trước, ví dụ thông qua thiết kế (thay đổi một hoặc nhiều axit amin) của vùng xoắn nhận diện của các protein ngón tay kẽm hoặc TALE xuất hiện tự nhiên. Do đó, các protein liên kết ADN thiết kế (ngón tay kẽm hoặc TALE) là các protein mà chúng không xuất hiện một cách tự nhiên. Các ví dụ không giới hạn về phương pháp thiết kế protein liên kết ADN là thiết kế và lựa chọn. Protein liên kết ADN thiết kế là protein không xuất hiện trong tự nhiên mà thiết kế/thành phần của nó chủ yếu xuất phát từ các tiêu chuẩn hợp lý. Tiêu chuẩn hợp lý đối với thiết kế bao gồm việc ứng dụng các quy tắc thế và các thuật toán vi tính để xử lý thông tin trong cơ sở dữ liệu lưu trữ thông tin của các thiết kế và dữ liệu liên kết ZFP và/hoặc TALE hiện hữu. Ví dụ, xem trong tài liệu U.S. Pat. số 6,140,081; 6,453,242; và 6,534,261; cũng xem trong tài liệu WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 và WO 03/016496 và công bố đơn Mỹ số 20110301073.

Như được sử dụng ở đây, "vectơ" bao gồm các tham chiếu đến axit nucleic sử dụng trong cấy nhiễm tế bào chủ và polynucleotit có thể được đưa vào trong đó. Các vectơ thường là các đơn vị sao chép (replicon). Các vectơ biểu hiện cho phép sự phiên mã của axit nucleic được xen vào trong đó.

"Kiểu hoang dại" có nghĩa những động vật này và túi phôi, phôi hoặc tế bào dẫn xuất từ đó, chúng không bị chỉnh sửa về mặt di truyền hoặc theo cách khác, được sửa đổi di truyền và thường là các chủng nội phối và ngoại phối được phát triển từ các chủng xuất hiện tự nhiên.

"Protein liên kết ADN ngón tay kẽm" (hoặc miền liên kết) là protein, hoặc miền bên trong protein lớn hơn, nó liên kết ADN theo cách đặc hiệu trình tự thông qua một hoặc nhiều protein ngón tay kẽm, đây là các vùng của trình tự axit amin bên trong miền

liên kết mà cấu trúc của miền này được làm ổn định thông qua sự sắp xếp phối vị của ion kẽm. Thuật ngữ protein liên kết ADN ngón tay kẽm thường được viết tắt là protein ngón tay kẽm hoặc ZFP.

Protein ngón tay kẽm "lựa chọn" hoặc TALE là protein không được tìm thấy trong tự nhiên mà việc sản xuất ra nó chủ yếu xuất phát từ quy trình thực nghiệm như hiển thị thể thực khuẩn, bẫy tương tác hoặc chọn lọc thể lai. Ví dụ, xem trong tài liệu patent Mỹ số 5,789,538; Patent Mỹ số 5,925,523; Patent Mỹ số 6,007,988; Patent Mỹ số 6,013,453; Patent Mỹ số 6,200,759; tài liệu WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970 WO 01/88197, WO 02/099084 và công bố đơn Mỹ số 20110301073.

Các thuật ngữ sau được sử dụng để mô tả mối quan hệ trình tự giữa polynucleotit/polypeptit theo sáng chế bằng tham chiếu polynucleotit/polypeptit: (a)"trình tự tham chiếu", (b)"cửa sổ so sánh", (c) "độ tương đồng trình tự", và (d)"tỷ lệ % về độ tương đồng trình tự".

(a) khi được sử dụng ở đây, "trình tự tham chiếu" là trình tự xác định được sử dụng làm cơ sở để so sánh trình tự với polynucleotit/polypeptit theo sáng chế. Trình tự tham chiếu có thể là trình tự con hoặc toàn bộ trình tự đặc hiệu; ví dụ, như đoạn cADN hoặc trình tự gen độ dài đầy đủ, hoặc cADN hoặc trình tự gen đầy đủ.

(b) Như được sử dụng ở đây, "cửa sổ so sánh" bao gồm các tham chiếu tới đoạn tiếp giáp và đặc hiệu của trình tự polynucleotit/polypeptit, trong đó trình tự polynucleotit/polypeptit có thể được so sánh với trình tự tham chiếu và trong đó, phần trình tự polynucleotit/polypeptit trong cửa sổ so sánh có thể bao gồm các thêm đoạn hoặc xóa đoạn (tức là, các khoảng trống) so với trình tự tham chiếu (các trình tự này không bao gồm thêm đoạn hoặc xóa đoạn) để sắp thẳng hàng tối ưu hai trình tự. Nói chung, cửa sổ so sánh có độ dài ít nhất 20 các nucleotit/gốc axit amin kế tiếp về độ dài, và tùy ý có thể là 30, 40, 50, 100, hoặc dài hơn. Những người có kỹ năng trong lĩnh vực hiểu rằng, để tránh mức tương đồng cao đối với trình tự tham chiếu do bao gồm cả các khoảng trống trong trình tự polynucleotit/polypeptit, điểm phạt khoảng trống điển hình được đưa vào và được khấu trừ từ số gốc giống hệt nhau.

Các phương pháp sắp thẳng hàng trình tự để so sánh đã được biết rõ trong lĩnh vực này. Việc sắp thẳng hàng tối ưu các trình tự để so sánh có thể được thực hiện bằng

toán thuật tương đồng tại chỗ của Smith và Waterman, Adv. Appl. Math. 2: 482(1981); bằng toán thuật sắp thẳng tương đồng của Needleman và Wunsch, J. Mol. Biol. 48: 443 (1970); bằng phương pháp tra cứu phương pháp tương đồng của Pearson và Lipman, Proc. Natl. Acad. Sci. 85: 2444 (1988); và bằng cách thực hiện vi tính hóa các thuật toán này, bao gồm, nhưng không bị giới hạn bởi: CLUSTAL trong chương trình PC/Gene bởi nhà cung cấp Intelligenetics, Mountain View, California; GAP, BESTFIT, BLAST, FASTA, và TFASTA, và các chương trình liên qua trong gói GCG Wisconsin Genetics Software Package, Version 10 (có tại Accelrys Inc., 9685 Scranton Road, San Diego, California, USA). Chương trình CLUSTAL sẽ được mô tả bởi tác giả Higgins và Sharp, Gene 73: 237-244 (1988); Higgins and Sharp, CABIOS 5: 151-153 (1989); Corpet, et al., axit nucleics Research 16: 10881-90 (1988); Huang, et al., Computer Applications in the Biosciences 8 : 155-65 (1992), and Pearson, et al., Methods in Molecular Biology 24: 307-331 (1994).

Nhóm các chương trình BLAST mà có thể được sử dụng trong tra cứu tương đông cơ sở dữ liệu bao gồm: BLASTN đối với trình tự truy vấn nucleotit dựa vào trình tự cơ sở dữ liệu nucleotit; BLASTX đối với trình tự truy vấn nucleotit dựa vào trình tự cơ sở dữ liệu protein ; BLASTP đối với trình tự truy vấn protein dựa vào trình tự cơ sở dữ liệu protein; TBLASTN đối với trình tự truy vấn protein dựa vào trình tự cơ sở dữ liệu nucleotit; và TBLASTX đối với trình tự truy vấn nucleotit dựa vào trình tự cơ sở dữ liệu nucleotit. Xem trong tài liệu Current Protocols in Molecular Biology, Chapter 19, Ausubel, *et al.*, Eds., Greene Publishing and Wiley-Interscience, New York (1995); Altschul *et al.*, J. Mol. Biol., 215: 403-410 (1990); and, Altschul *et al.*, Nucleic Acids Res. 25: 3389-3402 (1997). Phần mềm để thực hiện phân tích BLAST có sẵn trên, ví dụ National Center for Biotechnology Information (ncbi.nlm.nih.gov/). Thuật toán này đã được mô tả đầy đủ trong nhiều tài liệu công bố. Ví dụ, xem trong Altschul SF *et al.*, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, 25 Nucleic Acids Res. 3389 (1997); National Center for Biotechnology Information ,The NCBI Handbook [Internet], Chapter 16: The BLAST Sequence Analysis Tool (McEntyre J, Ostell J, eds., 2002), *available at* http://www.ncbi.nlm.nih.gov/books/NBK21097/pdf/ch16.pdf. Chương trình BLASTP đối với các trình tự axit amin cũng đã được mô tả đầy đủ (ví dụ, tài liệu Henikoff & Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915).

Ngoài tính toán tỷ lệ % tương đồng trình tự, thuật toán BLAST cũng thực hiện phân tích thống kê về mức tương đồng giữa hai trình tự (*ví dụ, xem trong* Karlin & Altschul, Proc. Nat'l. Acad. Sci. USA 90: 5873-5877 (1993)). Nhiều chương trình bộ lọc phức tạp thấp có thể được sử dụng để làm giảm các sắp xếp thẳng hàng độ phức tạp thấp như vậy. Ví dụ, the SEG (Wooten and Federhen, Comput. Chem., 17: 149-163 (1993)) and XNU (Claverie and States, Comput. Chem., 17: 191-201 (1993)) các bộ lọc phức tạp thấp có thể được sử dụng riêng biệt hoặc kết hợp.

Trừ khi được nêu rõ theo cách khác, các giá trị tương đồng/đồng nhất nucleotit và protein được nêu ở đây được tính toán sử dụng chương trình GAP (GCG Version 10) theo các giá trị hằng định. Chương trình GAP (Global Alignment Program) cũng có thể được sử dụng để so sánh polynucleotit hoặc polypeptit theo sáng chế với một trình tự tham chiếu. Chương trình GAP sử dụng thuật toán của Needleman and Wunsch (J. Mol. Biol. 48: 443-453, 1970) để tìm sự căn thẳng hàng hai trình tự đầy đủ tối đa hóa số lượng các gốc tương hợp và tối thiểu hóa số lượng khoảng trống. Chương trình GAP thể hiện một thành phần trong tập hợp các căn thẳng hàng thích hợp nhất. Có nhiều thành phần trong một tập hợp, nhưng không có thành phần khác nào có chất lượng tốt hơn. Chương trình GAP thể hiện 4 con số về giá trị căn thẳng: Chất lượng, Tỷ lệ, Độ tương đồng và Mức đồng nhất. Chất lượng là giá trị tối đa hóa để căn hàng các trình tự. Tỷ lệ là số lượng bazơ ở đoạn ngắn hơn chia cho chất lượng. Tỷ lệ % tương đồng là tỷ lệ của các ký hiệu thực tiễn tương hợp. Tỷ lệ % đồng nhất là tỷ lệ của các ký hiệu tương tự. Các ký hiệu mà chéo từ các khoảng trống được bỏ qua. Mức đồng nhất được ghi khi giá trị ma trận ghi điểm của một cặp ký hiệu lớn hơn hoặc tương đương 0,50, ngưỡng đồng nhất. Ma trận ghi điểm sử dụng trong phần mềm Version 10 của gói cung cấp Wisconsin Genetics Software Package là BLOSUM62 (*xem trong* Henikoff & Henikoff (1989) Proc. Natl. Acad. Sci. USA 89: 10915).

Việc căn thẳng phức tạp của các trình tự có thể được thực hiện sử dụng phương pháp căn thẳng CLUSTAL (Higgins and Sharp (1989) CABIOS. 5: 151-153) với các thông số hằng định (GAPPENALTY=10, GAP LENGTH PENALTY=10). Các thông số hằng định để căn thẳng từng cặp sử dụng phương pháp CLUSTAL bao gồm KTUPLE 1, GAP PENALTY=3, WINDOW=5 và DIAGONALS SAVED=5.

(c) như được sử dụng ở đây, "độ tương đồng trình tự" hoặc "độ tương đồng" trong ngữ cảnh của hai trình tự axit nucleic hoặc trình tự polypeptit bao gồm tham chiếu đến các gốc ở hai trình tự mà giống nhau khi được căn thẳng hàng để tương ứng cực đại qua cửa sổ so sánh đặc trưng. Khi tỷ lệ % về độ tương đồng trình tự được sử dụng trong tham chiếu tới các protein, có thể hiểu rằng, các vị trí gốc không giống nhau thường khác nhau bởi các thay thế axit amin bảo tồn, trong đó các gốc axit amin được thay thế cho các gốc axit amin khác có các đặc tính hóa học tương tự (ví dụ, điện tích hoặc tính kị nước) và do đó, không làm thay đổi các đặc tính chức năng của phân tử. Trường hợp các trình tự khác nhau trong các thay thế bảo tồn, tỷ lệ % tương đồng trình tự có thể được điều chỉnh về phía trên để hiệu chỉnh bản chất bảo tồn của sự thay thế. Các trình tự mà khác nhau bởi những sự thế bảo tồn như vậy được cho là có "mức đồng nhất trình tự" hoặc "mức đồng nhất". Các phương tiện để thực hiện việc điều chỉnh này là đã biết đối với người có kỹ năng trong lĩnh vực. Điển hình, phương tiện này liên quan đến việc ghi điểm sự thế bảo tồn ở dạng không tương hợp một phần, hơn là hoàn toàn không tương hợp, bằng cách đó gia tăng tỷ lệ % tương đồng trình tự. Do đó, ví dụ, trường hợp axit amin tương đồng được xác định điểm ghi là 1 và không có sự thế bảo tồn được xác định điểm ghi là 0, sự thế bảo tồn được xác định điểm ghi là nằm trong khoảng giữa 0 và 1. Việc ghi điểm sự thế bảo tồn có thể được tính toán theo thuật toán của Meyers và Miller, Computer Applic. Biol. Sci., 4: 11-17 (1988), ví dụ như được thực hiện trong chương trình PC/GENE (Intelligenetics, Mountain View, California, USA).

(d) khi được sử dụng ở đây, "tỷ lệ % về độ tương đồng trình tự" có nghĩa là giá trị được xác định bằng cách so sánh hai trình tự được sắp thẳng tối ưu theo cửa sổ so sánh, trong đó phần trình tự polynucleotit trong cửa sổ so sánh có thể bao gồm các thêm đoạn hoặc xóa đoạn (tức là các khoảng trống) khi so sánh với trình tự tham chiếu (trình tự này không bao gồm thêm đoạn hoặc xóa đoạn) để sắp thẳng hàng tối ưu hai trình tự. Tỷ lệ % được tính toán bằng cách xác định số vị trí mà tại đó, gốc bazơ axit nucleic hoặc gốc axit amin tương đồng xuất hiện ở cả hai trình tự để tạo ra số lượng vị trí tương hợp, lấy tổng số vị trí trong cửa sổ so sánh chia cho số lượng vị trí tương hợp và nhân kết quả với 100 để tạo ra tỷ lệ % về độ tương đồng trình tự.

Động vật và tế bào có trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163

CD163 có 17 exon và protein bao gồm vùng ngoại bào có 9 miền giàu xystein thụ thể dọn dẹp (SRCR), đoạn xuyên màng, và đuôi tế bào chất ngắn. Nhiều biến thể khác nhau xuất phát từ sự ghép nối riêng biệt của một gen đơn (Ritter et al. 1999a; Ritter et al. 1999b). Phần lớn sự biến đổi này được giải thích bởi độ dài của đuôi tế bào chất.

CD163 có nhiều chức năng quan trọng, bao gồm hoạt động như là thụ thể dọn dẹp haptoglobin-hemoglobin. Việc loại bỏ hemoglobin tự do trong máu là một chức năng quan trọng của CD163 ở dạng nhóm heme có thể rất độc (Kristiansen et al. 2001). CD163 có đuôi tế bào chất mà nó tạo thuận lợi cho sự nhập bào. Đột biến ở vùng đuôi này dẫn đến giảm hấp thụ phức haptoglobin-hemoglobin (Nielsen et al. 2006). Các chức năng khác của C163 bao gồm kết dính nguyên hồng cầu (SRCR2), làm thụ thể TWEAK (SRCR1-4 & 6-9), thụ thể vi khuẩn (SRCR5), thụ thể virut ở lợn châu Phi (Sanchez-Tones et al. 2003), và vai trò tiềm năng làm tác nhân điều biến miễn dịch (được thảo luận trong Van Gorp et al. 2010a). Theo quan điểm về các chức năng quan trọng này, trước đây cho rằng, việc bất hoạt hoàn toàn CD163 sẽ dẫn đến các động vật sẽ không có khả năng sống sót được hoặc sẽ bị tổn thương nghiêm trọng (ví dụ, xem trong công bố đơn PCT số 2012/158828).

CD163 là thành viên của siêu họ giàu xystein thụ thể dọn dẹp (SRCR) và gồm có miền nội bào và 9 miền SRCR ngoại bào. Ở người, sự nhập bào hấp thụ hemoglobin-heme trung gian bởi CD163 thông qua SRCR3 bảo vệ tế bào tránh được sự kích ứng oxy hóa (Schaer et al., 2006a; Schaer et al., 2006b; Schaer et al., 2006c). CD163 cũng dùng làm thụ thể của tác nhân cảm ứng yếu giống yếu tố hoại tử khối u của quá trình chết tế bào theo lập trình (TWEAK: SRCR1-4 & 6-9), thụ thể tác nhân gây bệnh (Virut gây sốt ở lợn châu Phi; vi khuẩn: SRCR2), và liên kết nguyên hồng cầu (SRCR2).

CD163 đóng vai trò trong nhiễm trùng bởi nhiều tác nhân gây bệnh khác nhau và do đó, sáng chế không chỉ giới hạn ở động vật có tính nhạy giảm với nhiễm trùng PRRSV, mà còn bao gồm cả động vật có tính nhạy giảm với tác nhân gây bệnh bất kỳ mà dựa vào CD163 để gây nhiễm tế bào hoặc để sao chép sau đó và/hoặc tồn tại dai dẳng trong tế bào. Quá trình nhiễm trùng của PRRSV bắt đầu bằng sự liên kết đầu tiên với heparan sulfat trên bề mặt của đại thực bào phế nang. Trước khi 2013, được cho rằng, sau đó sự liên kết chắc chắn xuất hiện với sialoadhesin (SIGLEC1, cũng được gọi là CD169 hoặc SN). Sau đó,virut được nội nhập thông qua sự nhập bào qua trung gian

clatherin. Phân tử khác, CD163, sau đó tạo thuận lợi cho sự tháo vỏ của virut trong nội thể (Van Breedam et al. 2010a). Hệ gen virut được giải phóng và tế bào bị nhiễm

Được mô tả ở đây là động vật và thế hệ thế hệ con và các tế bào chứa ít nhất một trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163, ví dụ, xen đoạn hoặc xóa đoạn ("INDEL"), mà các đoạn này truyền tính kháng cải thiện hoặc tính kháng đầy đủ đối với nhiễm trùng bởi tác nhân gây bệnh (ví dụ, PRRSV) cho lợn. Người nộp đơn đã chứng minh rằng, CD163 là gen quyết định trong nhiễm trùng PRRSV và đã tạo ra các động vật và các dòng chuyển gen kháng bệnh.

Sáng chế này đề cập đến động vật biến đổi gen, thế hệ thế hệ con, hoặc các tế bào động vật chứa ít nhất một trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163. Sáng chế này không bao gồm việc bất hoạt hoặc chỉnh sửa gen SIGLEC1 (CD169), gen này trước đây đã được thừa nhận là quyết định đối với tính kháng PRRSV.

Trình tự nhiễm sắc thể chỉnh sửa có thể (1) được bất hoạt, (2) sửa đổi hoặc (3) chứa trình tự tích hợp dẫn đến đột biến vô hiệu. Trình tự nhiễm sắc thể bất hoạt được thay đổi sao cho chức năng protein CD163 của nó liên quan đến nhiễm trùng PRRSV bị ảnh hưởng, suy giảm hoặc bị triệt tiêu. Do đó, động vật chỉnh sửa gen chứa trình tự nhiễm sắc thể bất hoạt có thể được gọi là "bất hoạt" hoặc "bất hoạt có điều kiện." Tương tự, động vật chỉnh sửa gen bao gồm trình tự tích hợp có thể được gọi là "gắn gen" hoặc "gắn gen có điều kiện." Ngoài ra, động vật chỉnh sửa gen chứa trình tự nhiễm sắc thể sửa đổi có thể bao gồm (các) đột biến điểm đích hoặc sửa đổi khác, sao cho sản phẩm protein biến đổi được tạo ra. Một cách vắn tắt, quy trình có thể bao gồm bước đưa vào trong phôi hoặc tế bào ít nhất một phân tử ARN mã hóa cho nucleaza ngón tay kẽm đích và tùy ý, ít nhất một polynucleotit phụ. Phương pháp còn bao gồm bước ủ phôi hoặc tế bào để cho phép sự biểu hiện của nucleaza ngón tay kẽm, trong đó đoạn đứt sợi kép được đưa vào trong trình tự nhiễm sắc thể đích bằng nucleaza ngón tay kẽm được sửa chữa bằng quy trình sửa chữa ADN nối các đầu không tương đồng sai úp sấp hoặc quy trình sửa chữa ADN định hướng tương đồng. Phương pháp chỉnh sửa trình tự nhiễm sắc thể mã hóa cho protein liên quan đến sự phát triển dòng mầm sử dụng công nghệ nucleaza ngón tay kẽm đích là nhanh chóng, chính xác và hiệu quả cao.

Theo cách khác, quy trình có thể bao gồm bước sử dụng hệ CRISPR/Cas9 để sửa đổi trình tự hệ gen để sử dụng Cas9 để sửa đổi trình tự hệ gen, protein có thể được phân

phối trực tiếp tới tế bào. Theo cách khác, mARN mã hóa cho Cas9 có thể được phân phối tới tế bào hoặc gen cung cấp sự biểu hiện của mARN mã hóa cho Cas9 có thể được phân phối tới tế bào. Ngoài ra, crARN đặc hiệu đích và tracrARN có thể được phân phối trực tiếp tới tế bào hoặc (các) gARN đặc hiệu đích có thể tới tế bào (các ARN này có thể được tạo ra theo cách lựa chọn bằng gen được cấu trúc để biểu hiện các ARN này). Việc lựa chọn các vị trí đích và thiết kế crARN/gARN là đã biết trong lĩnh vực. Thảo luận về cấu trúc và tách dòng gARN có thể được tìm thấy tại http://www.genome-engineering.org/crispr/wp-content/uploads/2014/05/CRISPR-Reagent-Description-Rev20140509.pdf.

Ít nhất một vị trí gen CD163 có thể được sử dụng làm vị trí đích để chỉnh sửa đặc hiệu điểm. Chỉnh sửa đặc hiệu điểm có thể bao gồm việc đưa vào axit nucleic ngoại sinh (ví dụ, axit nucleic bao gồm trình tự nucleotit mã hóa polypeptit mong muốn) hoặc các đoạn xóa của các axit nucleic từ vị trí. Ví dụ, việc tích hợp axit nucleic ngoại sinh và/hoặc xóa đoạn một phần của axit nucleic hệ gen có thể sửa đổi vị trí để tạo ra gen CD163 đứt đoạn (tức là, hoạt tính của protein CD163 giảm).

Được đề cập ở đây là động vật không phải người, thế hệ con của động vật này, và các tế bào động vật chứa ít nhất một trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163.

Sáng chế đề cập đến động vật không phải người hoặc thế hệ thế hệ con hoặc tế bào động vật chứa ít nhất một trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163. Trình tự nhiễm sắc thể sửa đổi dẫn đến sự tạo ra protein CD163 cơ bản không có chức năng bởi động vật, thế hệ con, hoặc tế bào của nó.

Động vật không phải người hoặc thế hệ thế hệ con hoặc tế bào động vật khác chứa ít nhất một trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163 được đề cập. Trình tự nhiễm sắc thể sửa đổi bao gồm xóa trong khung ở gen mã hóa cho protein CD163.

Lợn hoặc thế hệ thế hệ con hoặc tế bào lợn chứa ít nhất một trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163 được đề cập. Trình tự nhiễm sắc thể sửa đổi bao gồm: (a) SEQ ID NO: 118; hoặc (b) sửa đổi được chọn từ nhóm bao gồm: đoạn xóa 11 cặp bazơ từ nucleotit 3137 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xen 2 cặp bazơ giữa các nucleotit 3149 và 3150 khi so sánh với

trình tự tham chiếu SEQ ID NO: 47, với đoạn xóa 377 cặp bazơ từ nucleotit 2573 đến nucleotit 2949 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trên cùng alen; đoạn xóa 124 cặp bazơ từ nucleotit 3024 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 123 cặp bazơ từ nucleotit 3024 đến nucleotit 3146 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xen 1 cặp bazơ giữa các nucleotit 3147 và 3148 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 130 cặp bazơ từ nucleotit 3030 đến nucleotit 3159 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 132 cặp bazơ từ nucleotit 3030 đến nucleotit 3161 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1506 cặp bazơ từ nucleotit 1525 đến nucleotit 3030 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xen 7 cặp bazơ giữa nucleotit 3148 và nucleotit 3149 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1280 cặp bazơ từ nucleotit 2818 đến nucleotit 4097 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1373 cặp bazơ từ nucleotit 2724 đến nucleotit 4096 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1467 cặp bazơ từ nucleotit 2431 đến nucleotit 3897 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2417 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 12 cặp bazơ bắt đầu tại nucleotit 488, và trong đó có thêm đoạn xóa 129 cặp bazơ trong exon 7 từ nucleotit 3044 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 28 cặp bazơ từ nucleotit 3145 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1387 cặp bazơ từ nucleotit 3145 đến nucleotit 4531 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1382 cặp bazơ từ nucleotit 3113 đến nucleotit 4494 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 11 cặp bazơ bắt đầu tại nucleotit 3113; đoạn xóa 1720 cặp bazơ từ nucleotit 2440 đến nucleotit 4160 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 452 cặp bazơ từ nucleotit 3015 đến nucleotit 3466 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; và tổ hợp của chúng.

Sự sửa đổi về trình tự nhiễm sắc thể ở gen mã hóa cho protein CD163 làm giảm tính nhạy của động vật, thế hệ con, hoặc tế bào đối với nhiễm trùng bởi tác nhân gây bệnh (ví dụ, virut như PRRSV), khi so sánh với tính nhạy của động vật, thế hệ con, hoặc tế bào mà không chứa trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163 đối với nhiễm trùng bởi tác nhân gây bệnh.

Ví dụ, sự sửa đổi về trình tự nhiễm sắc thể ở gen mã hóa cho protein CD163 có thể làm giảm tính nhạy của động vật, thế hệ con, hoặc tế bào với virut PRRSV Type 1, PRRSV Type 2, hoặc với cả virut PRRSV Type 1 và Type 2.

Sự sửa đổi về trình tự nhiễm sắc thể ở gen mã hóa cho protein CD163 có thể làm giảm động vật, thế hệ con hoặc tế bào với thể phân lập PRRSV được chọn từ nhóm bao gồm NVSL 97-7895, KS06-72109, P129, VR2332, CO90, AZ25, MLV-ResPRRS, KS62-06274, KS483 (SD23983), CO84, SD13-15, Lelystad, 03-1059, 03-1060, SD01-08, 4353PZ, và tổ hợp của chúng.

Động vật hoặc thế hệ con có thể là phôi, con non hoặc đối tượng trưởng thành. Tương tự, tế bào có thể bao gồm tế bào phôi, tế bào xuất phát từ động vật non, hoặc tế bào xuất phát từ động vật trưởng thành.

Động vật hoặc thế hệ con có thể bao gồm động vật nuôi trong gia đình. Tương tự, tế bào có thể bao gồm tế bào xuất phát từ động vật nuôi trong gia đình. Động vật nuôi trong gia đình có thể bao gồm động vật nuôi, ví dụ lợn, bò (ví dụ, bò thịt hoặc bò sữa), cừu, dê, ngựa (ví dụ, ngựa hoặc lừa), trâu, lạc đà, hoặc động vật thuộc loài chim (ví dụ, gà, gà tây, vịt, ngỗng, gà Nhật, hoặc bồ câu non). Động vật nuôi tốt hơn là bò hoặc lợn, và tốt nhất là lợn.

Động vật hoặc thế hệ con có thể bao gồm động vật chỉnh sửa gen. Tế bào có thể bao gồm tế bào chỉnh sửa gen.

Động vật hoặc tế bào có thể được chỉnh sửa gen sử dụng endonucleaza tìm nhà. Endonucleaza tìm nhà có thể là endonucleaza xuất hiện tự nhiên, nhưng tốt hơn là endonucleaza tìm nhà không xuất hiện một cách tự nhiên, được thiết kế hợp lý mà có trình tự nhận diện ADN đã được thiết kế sao cho endonucleaza hướng đích vào trình tự nhiễm sắc thể ở gen mã hóa cho protein CD163. Do đó, endonucleaza tìm nhà có thể là endonucleaza tìm nhà được thiết kế. Endonucleaza tìm nhà có thể bao gồm, ví dụ, hệ trình tự sắp xếp kiểu lặp đi lặp lại xen khoảng trống (CRISPR) /Cas9, men nucleaza tác động giống chất kích hoạt phiên mã (TALEN), nucleaza ngón tay kẽm (ZFN), protein dung hợp recombinaza, meganucleaza, hoặc tổ hợp của chúng. Động vật hoặc tế bào tốt hơn là động vật hoặc tế bào đã được chỉnh sửa gen sử dụng hệ CRISPR/Cas9.

Động vật chỉnh sửa gen, thế hệ thế hệ con, hoặc tế bào chỉnh sửa gen tốt hơn thể hiện tính kháng tăng đối với tác nhân gây bệnh (ví dụ, virut như PRRSV) khi so sánh với động vật không được chỉnh sửa.

Ví dụ, động vật chỉnh sửa gen có thể thể hiện tính kháng cao với virut PRRSV Type 1, PRRSV Type 2, hoặc với cả virut PRRSV Type 1 và Type 2.

Động vật chỉnh sửa gen có thể thể hiện tính kháng cao với thể phân lập PRRSV được chọn từ nhóm bao gồm NVSL 97-7895, KS06-72109, P129, VR2332, CO90, AZ25, MLV-ResPRRS, KS62-06274, KS483 (SD23983), CO84, SD13-15, Lelystad, 03-1059, 03-1060, SD01-08, 4353PZ, và tổ hợp của chúng.

Động vật, thế hệ con, hoặc tế bào có thể là dị hợp tử về trình tự nhiễm sắc thể sửa đổi. Theo cách khác, động vật, thế hệ con, hoặc tế bào có thể là đồng hợp tử về trình tự nhiễm sắc thể sửa đổi.

Ở bất kỳ trong số các động vật, thế hệ con, hoặc tế bào, trình tự nhiễm sắc thể sửa đổi có thể chứa đoạn xen ở gen mã hóa cho protein CD163, đoạn xóa ở gen mã hóa cho protein CD163, hoặc tổ hợp của chúng. Ví dụ, trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xóa ở gen mã hóa cho protein CD163 (ví dụ, xóa trong khung). Theo cách khác, trình tự nhiễm sắc thể sửa đổi có thể chứa đoạn xen ở gen mã hóa cho protein CD163.

Đoạn xen hoặc đoạn xóa có thể dẫn đến sự tạo ra hoặc hoạt tính của protein CD163 bị giảm, khi so sánh với sự tạo ra hoặc hoạt tính của protein CD163 ở động vật, thế hệ con, hoặc tế bào mà không có đoạn xen hoặc đoạn xóa.

Đoạn xen hoặc đoạn xóa có thể dẫn đến sự tạo ra protein CD163 cơ bản không có chức năng bởi động vật, thế hệ con, hoặc tế bào. Với thuật ngữ "protein CD163 cơ bản không có chức năng," nó có nghĩa rằng, mức protein CD163 ở động vật, thế hệ con, hoặc tế bào là không thể phát hiện được hoặc nếu phát hiện được, thì ít nhất thấp hơn khoảng 90% so với mức quan sát được ở động vật, thế hệ con, hoặc tế bào mà không chứa đoạn xen hoặc đoạn xóa.

Trường hợp động vật, thế hệ con, hoặc tế bào bao gồm động vật là lợn, thế hệ con, hoặc tế bào, trình tự nhiễm sắc thể sửa đổi có thể bao gồm sửa đổi trong exon 7 của gen mã hóa cho protein CD163, exon 8 của gen mã hóa cho protein CD163, một

intron mà nó liên tiếp với exon 7 hoặc exon 8 của gen mã hóa cho protein CD163, hoặc tổ hợp của chúng. Trình tự nhiễm sắc thể sửa đổi một cách thích hợp bao gồm sửa đổi trong exon 7 của gen mã hóa cho protein CD163.

Sự sửa đổi trong exon 7 của gen mã hóa cho protein CD163 có thể bao gồm đoạn xóa (ví dụ, xóa trong khung trong exon 7). Theo cách khác, sự sửa đổi trong exon 7 của gen mã hóa cho protein CD163 có thể bao gồm đoạn xen.

Ở bất kỳ trong số các động vật là lợn, thế hệ con, hoặc tế bào, trình tự nhiễm sắc thể sửa đổi có thể bao gồm SEQ ID NO: 118. Theo cách khác, trình tự nhiễm sắc thể sửa đổi có thể bao gồm sửa đổi được chọn từ nhóm bao gồm: đoạn xóa 11 cặp bazơ từ nucleotit 3137 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xen 2 cặp bazơ giữa các nucleotit 3149 và 3150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, với đoạn xóa 377 cặp bazơ từ nucleotit 2573 đến nucleotit 2949 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trên cùng alen; đoạn xóa 124 cặp bazơ từ nucleotit 3024 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 123 cặp bazơ từ nucleotit 3024 đến nucleotit 3146 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xen 1 cặp bazơ giữa các nucleotit 3147 và 3148 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 130 cặp bazơ từ nucleotit 3030 đến nucleotit 3159 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 132 cặp bazơ từ nucleotit 3030 đến nucleotit 3161 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1506 cặp bazơ từ nucleotit 1525 đến nucleotit 3030 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xen 7 cặp bazơ giữa nucleotit 3148 và nucleotit 3149 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1280 cặp bazơ từ nucleotit 2818 đến nucleotit 4097 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1373 cặp bazơ từ nucleotit 2724 đến nucleotit 4096 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1467 cặp bazơ từ nucleotit 2431 đến nucleotit 3897 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2417 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 12 cặp bazơ bắt đầu tại nucleotit 488, và trong đó có thêm đoạn xóa 129 cặp bazơ trong exon 7 từ nucleotit 3044 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 28 cặp bazơ từ nucleotit 3145 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1387 cặp bazơ từ nucleotit 3145 đến nucleotit 4531 khi so

sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1382 cặp bazơ từ nucleotit 3113 đến nucleotit 4494 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 11 cặp bazơ bắt đầu tại nucleotit 3113; đoạn xóa 1720 cặp bazơ từ nucleotit 2440 đến nucleotit 4160 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 452 cặp bazơ từ nucleotit 3015 đến nucleotit 3466 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; hoặc tổ hợp của chúng.

Ví dụ, sự sửa đổi có thể bao gồm đoạn xóa 11 cặp bazơ từ nucleotit 3137 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Sự sửa đổi có thể bao gồm đoạn xen 2 cặp bazơ giữa các nucleotit 3149 và 3150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, với đoạn xóa 377 cặp bazơ từ nucleotit 2573 đến nucleotit 2949 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trên cùng alen.

Sự sửa đổi có thể bao gồm đoạn xóa 124 cặp bazơ từ nucleotit 3024 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Sự sửa đổi có thể bao gồm đoạn xóa 123 cặp bazơ từ nucleotit 3024 đến nucleotit 3146 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Sự sửa đổi có thể bao gồm đoạn xen 1 cặp bazơ giữa các nucleotit 3147 và 3.148 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Sự sửa đổi có thể bao gồm đoạn xóa 130 cặp bazơ từ nucleotit 3030 đến nucleotit 3159 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Sự sửa đổi có thể bao gồm đoạn xóa 132 cặp bazơ từ nucleotit 3030 đến nucleotit 3161 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Sự sửa đổi có thể bao gồm đoạn xóa 1506 cặp bazơ từ nucleotit 1525 đến nucleotit 3030 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Sự sửa đổi có thể bao gồm đoạn xen 7 cặp bazơ giữa nucleotit 3148 và nucleotit 3149 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Sự sửa đổi có thể bao gồm đoạn xóa 1280 cặp bazơ từ nucleotit 2818 đến nucleotit 4097 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Sự sửa đổi có thể bao gồm đoạn xóa 1373 cặp bazơ từ nucleotit 2724 đến nucleotit 4096 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Sự sửa đổi có thể bao gồm đoạn xóa 1467 cặp bazơ từ nucleotit 2431 đến nucleotit 3897 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Sự sửa đổi có thể bao gồm đoạn xóa 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2417 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 12 cặp bazơ bắt đầu tại nucleotit 488, và trong đó có thêm đoạn xóa 129 cặp bazơ trong exon 7 từ nucleotit 3044 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Sự sửa đổi có thể bao gồm đoạn xóa 28 cặp bazơ từ nucleotit 3145 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Sự sửa đổi có thể bao gồm đoạn xóa 1387 cặp bazơ từ nucleotit 3145 đến nucleotit 4531 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Sự sửa đổi có thể bao gồm đoạn xóa 1382 cặp bazơ từ nucleotit 3113 đến nucleotit 4494 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 11 cặp bazơ bắt đầu tại nucleotit 3113.

Sự sửa đổi có thể bao gồm đoạn xóa 1720 cặp bazơ từ nucleotit 2.440 đến nucleotit 4.160 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Sự sửa đổi có thể bao gồm đoạn xóa 452 cặp bazơ từ nucleotit 3015 đến nucleotit 3466 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa tổ hợp bất kỳ của các đoạn xen và đoạn xóa nêu trên.

SEQ ID NO: 47 cung cấp trình tự nucleotit của vùng đầu dòng bắt đầu 3000 cặp bazơ (bp) của exon 7 của gen CD163 của lợn kiểu hoang dại tới bazơ cuối cùng của exon 10 của gen này. SEQ ID NO: 47 được sử dụng làm trình tự tham chiếu ở đây và được thể hiện ở hình vẽ trên Fig. 16.

Khi động vật là lợn, thế hệ con, hoặc tế bào chứa đoạn xen 2 cặp bazơ giữa các nucleotit 3149 và 3150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, đoạn xen 2 cặp bazơ có thể bao gồm đoạn xen của dinucleotit AG.

Khi động vật là lợn, thế hệ con, hoặc tế bào chứa đoạn xen 1 cặp bazơ giữa các nucleotit 3147 và 3.148 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, đoạn xen 1 bazơ có thể bao gồm đoạn xen của gốc adenine duy nhất.

Khi động vật là lợn, thế hệ con, hoặc tế bào chứa đoạn xen 7 cặp bazơ giữa nucleotit 3148 và nucleotit 3149 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, đoạn xen 7 cặp bazơ có thể bao gồm trình tự TACTACT (SEQ ID NO: 115).

Khi động vật là lợn, thế hệ con, hoặc tế bào chứa đoạn xóa 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2417 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 12 cặp bazơ bắt đầu tại nucleotit 488, và trong đó có thêm đoạn xóa 129 cặp bazơ trong exon 7 từ nucleotit 3044 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, đoạn xen 12 cặp bazơ có thể bao gồm trình tự TGTGGAGAATTC (SEQ ID NO: 116).

Khi động vật là lợn, thế hệ con, hoặc tế bào chứa đoạn xóa 1382 cặp bazơ từ nucleotit 3113 đến nucleotit 4494 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 11 cặp bazơ bắt đầu tại nucleotit 3113, đoạn xen 11 cặp bazơ có thể bao gồm trình tự AGCCAGCGTGC (SEQ ID NO: 117).

Trường hợp trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163 bao gồm đoạn xóa, đoạn xóa tốt hơn bao gồm xóa trong khung. Các đoạn xóa trong khung là các đoạn xóa không gây sự dịch chuyển về khung đọc bộ ba, và do đó, sản phẩm protein tạo ra có đoạn xóa bên trong gồm một hoặc nhiều axit amin, nhưng không bị cụt đầu. Đoạn xóa 3 cặp bazơ hoặc nhiều phần của 3 cặp bazơ bên trong exon có thể dẫn đến một đột biến trong khung, giả sử rằng, sự ghép nối xuất hiện chính xác.

Các INDEL sau mô tả ở đây về lợn và các tế bào của lợn được kỳ vọng là dẫn đến các đoạn xóa trong khung, vì đoạn xóa bên trong exon 7 của gen CD163 lợn là bội của 3: đoạn xóa 1506 cặp bazơ từ nucleotit 1525 đến nucleotit 3030 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2417 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 12 cặp bazơ bắt đầu tại nucleotit 488, và trong đó có thêm đoạn xóa 129 cặp bazơ trong exon 7 từ nucleotit 3044 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1373 cặp bazơ từ nucleotit

2724 đến nucleotit 4096 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 123 cặp bazơ từ nucleotit 3024 đến nucleotit 3146 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1467 cặp bazơ từ nucleotit 2431 đến nucleotit 3897 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1387 cặp bazơ từ nucleotit 3145 đến nucleotit 4531 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1382 cặp bazơ từ nucleotit 3113 đến nucleotit 4494 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 11 cặp bazơ bắt đầu tại nucleotit 3113; và đoạn xóa 1720 cặp bazơ từ nucleotit 2440 đến nucleotit 4160 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Do đó, ở động vật là lợn, thế hệ con, và tế bào của lợn, đoạn xen hoặc đoạn xóa ở gen mã hóa cho protein CD163 có thể bao gồm một đoạn xóa trong khung trong exon 7 được chọn từ nhóm bao gồm đoạn xóa 1506 cặp bazơ từ nucleotit 1525 đến nucleotit 3030 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2417 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 12 cặp bazơ bắt đầu tại nucleotit 488, và trong đó có thêm đoạn xóa 129 cặp bazơ trong exon 7 từ nucleotit 3044 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1373 cặp bazơ từ nucleotit 2724 đến nucleotit 4096 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 123 cặp bazơ từ nucleotit 3024 đến nucleotit 3146 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1467 cặp bazơ từ nucleotit 2431 đến nucleotit 3897 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1387 cặp bazơ từ nucleotit 3145 đến nucleotit 4531 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1382 cặp bazơ từ nucleotit 3113 đến nucleotit 4494 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 11 cặp bazơ bắt đầu tại nucleotit 3113; đoạn xóa 1720 cặp bazơ từ nucleotit 2440 đến nucleotit 4160 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; và tổ hợp của chúng.

Động vật là lợn, thế hệ con, hoặc tế bào của lợn có thể bao gồm đoạn xen hoặc đoạn xóa được chọn từ nhóm bao gồm: đoạn xen 2 cặp bazơ giữa các nucleotit 3149 và 3150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, với đoạn xóa 377 cặp bazơ từ nucleotit 2573 đến nucleotit 2949 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trên cùng alen; đoạn xóa 28 cặp bazơ từ nucleotit 3145 đến nucleotit 3172 khi so

sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 452 cặp bazơ từ nucleotit 3015 đến nucleotit 3466 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; và tổ hợp của chúng.

Ví dụ, động vật là lợn, thế hệ con, hoặc tế bào của lợn có thể bao gồm đoạn xen 2 cặp bazơ giữa các nucleotit 3149 và 3150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, với đoạn xóa 377 cặp bazơ từ nucleotit 2573 đến nucleotit 2949 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trên cùng alen.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa đoạn xóa 28 cặp bazơ từ nucleotit 3145 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể bao gồm đoạn xóa 452 cặp bazơ từ nucleotit 3015 đến nucleotit 3466 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể bao gồm tổ hợp bất kỳ của các trình tự nhiễm sắc thể sửa đổi bất kỳ mô tả ở đây.

Ví dụ, động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa đoạn xen 7 cặp bazơ giữa nucleotit 3148 và nucleotit 3149 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn xóa 11 cặp bazơ từ nucleotit 3137 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa đoạn xen 7 cặp bazơ giữa nucleotit 3148 và nucleotit 3149 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn xóa 1382 cặp bazơ từ nucleotit 3113 đến nucleotit 4494 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 11 cặp bazơ bắt đầu tại nucleotit 3113, trong alen khác của gen mã hóa cho protein CD163.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể bao gồm SEQ ID NO: 118 trong một alen của gen mã hóa cho protein CD163; và đoạn xóa 11 cặp bazơ từ nucleotit 3137 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể bao gồm SEQ ID NO: 118 trong một alen của gen mã hóa cho protein CD163; và đoạn xen 2 cặp bazơ giữa các nucleotit 3149 và 3150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, với đoạn xóa 377 cặp bazơ từ nucleotit 2573 đến nucleotit 2949 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa đoạn xóa 1280 cặp bazơ từ nucleotit 2.818 đến nucleotit 4.097 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn xóa 11 cặp bazơ từ nucleotit 3137 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa đoạn xóa 1280 cặp bazơ từ nucleotit 2.818 đến nucleotit 4.097 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn xen 2 cặp bazơ giữa các nucleotit 3149 và 3150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, với đoạn xóa 377 cặp bazơ từ nucleotit 2573 đến nucleotit 2949 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa đoạn xóa 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2417 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 12 cặp bazơ bắt đầu tại nucleotit 488, và trong đó có thêm đoạn xóa 129 cặp bazơ trong exon 7 từ nucleotit 3044 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn xen 2 cặp bazơ giữa các nucleotit 3149 và 3150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, với đoạn xóa 377 cặp bazơ từ nucleotit 2573 đến nucleotit 2949 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể bao gồm SEQ ID NO: 118 trong một alen của gen mã hóa cho protein CD163; và đoạn xóa 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2417 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 12 cặp bazơ bắt đầu tại nucleotit 488, và trong đó có thêm đoạn xóa 129 cặp bazơ trong exon 7 từ nucleotit 3044 đến

nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa đoạn xóa 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2417 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 12 cặp bazơ bắt đầu tại nucleotit 488, và trong đó có thêm đoạn xóa 129 cặp bazơ trong exon 7 từ nucleotit 3044 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn xóa 11 cặp bazơ từ nucleotit 3137 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa đoạn xóa 1467 cặp bazơ từ nucleotit 2431 đến nucleotit 3897 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn xen 2 cặp bazơ giữa các nucleotit 3149 và 3150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, với đoạn xóa 377 cặp bazơ từ nucleotit 2573 đến nucleotit 2949 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa đoạn xóa 1467 cặp bazơ từ nucleotit 2431 đến nucleotit 3897 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn xóa 11 cặp bazơ từ nucleotit 3137 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Động vật là lợn, thế hệ con, hoặc tế bào của nó chứa bất kỳ đoạn xen hoặc đoạn xóa nào mô tả ở trên có thể bao gồm trình tự nhiễm sắc thể có mức độ tương đồng trình tự cao so với SEQ ID NO: 47 ngoài đoạn xen hoặc đoạn xóa. Do đó, ví dụ, động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa trình tự nhiễm sắc thể có ít nhất 80%, ít nhất 85%, ít nhất 90%, ít nhất 95%, ít nhất 98%, ít nhất 99%, ít nhất 99,9%, hoặc 100% mức tương đồng trình tự so với SEQ ID NO: 47 ở các vùng của trình tự nhiễm sắc thể ngoài đoạn xen hoặc đoạn xóa.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa trình tự nhiễm sắc thể bao gồm SEQ ID NO: 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 118, hoặc 119. Như được mô tả thêm trong phần ví dụ dưới đây,

các SEQ ID NO. 98–114, và 119 đề cập đến các trình tự nucleotit của vùng tương ứng với vùng CD163 lợn kiểu hoang dại được nêu trong SEQ ID NO:47, và bao gồm đoạn xen hoặc đoạn xóa trong trình tự nhiễm sắc CD163 lợn được mô tả ở đây. SEQ ID NO: 118 đề cập trình tự của vùng tương ứng với vùng CD163 lợn kiểu hoang dại được nêu bởi SEQ ID NO: 47, trong đó exon 7 đã được thay thế bằng một exon tổng hợp mã hóa vùng tương đồng SRCR 8 của protein giống CD163 người 1 (hCD163L1).

Ví dụ, động vật là lợn, thế hệ con, động vật hoặc tế bào của nó có thể chứa trình tự nhiễm sắc thể bao gồm SEQ ID NO: 98, 101, 105, 109, 110, 112, 113, hoặc 114. SEQ ID NO: 98, 101, 105, 109, 110, 112, 113, và 114 cung cấp các trình tự nucleotit của các đoạn xóa trong khung trong exon 7 của trình tự nhiễm sắc CD163 lợn.

Đối với ví dụ khác, động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa trình tự nhiễm sắc thể bao gồm SEQ ID NO: 103, 111, hoặc 119.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa đoạn xóa 11 cặp bazơ trong một alen của gen mã hóa cho protein CD163 và đoạn xen 2 cặp bazơ với đoạn xóa 377 cặp bazơ trong alen khác của gen mã hóa cho protein CD163.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa đoạn xóa 124 cặp bazơ trong một alen của gen mã hóa cho protein CD163 và đoạn xóa 123 cặp bazơ trong alen khác của gen mã hóa cho protein CD163.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa đoạn xen 1 bazơ.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa đoạn xóa 130 cặp bazơ trong một alen của gen mã hóa cho protein CD163 và đoạn xóa 132 cặp bazơ trong alen khác của gen mã hóa cho protein CD163.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa đoạn xóa 1506 cặp bazơ.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa đoạn xen 7 bazơ.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa đoạn xóa 1280 cặp bazơ trong một alen của gen mã hóa cho protein CD163 và đoạn xóa 1373 cặp bazơ trong alen khác của gen mã hóa cho protein CD163.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa đoạn xóa 1467 cặp bazơ.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa đoạn xóa intron 6 gồm 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2417, với đoạn xen 12 cặp bazơ tại nucleotit 4488 và đoạn xóa 129 cặp bazơ bổ sung trong exon 7.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể chứa đoạn xóa 28 cặp bazơ trong một alen của gen mã hóa cho protein CD163 và đoạn xóa 1387 cặp bazơ trong alen khác của gen mã hóa cho protein CD163.

Động vật là lợn, thế hệ con, hoặc tế bào của nó có thể bao gồm đoạn xóa 1382 cặp bazơ với đoạn xen 11 cặp bazơ trong một alen của gen mã hóa cho protein CD163 và đoạn xóa 1720 cặp bazơ trong alen khác của gen mã hóa cho protein CD163.

Tế bào bất kỳ trong số các tế bào chứa ít nhất một trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163 có thể bao gồm tế bào tinh. Theo cách khác, tế bào bất kỳ trong số các tế bào này có thể bao gồm tế bào trứng (ví dụ, trứng đã thụ tinh).

Tế bào bất kỳ trong số các tế bào chứa ít nhất một trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163 có thể bao gồm tế bào sinh dưỡng. Ví dụ, tế bào bất kỳ trong số các tế bào có thể bao gồm nguyên bào sợi (ví dụ, nguyên bào sợi thai).

Sự gắn hướng đích của axit nucleic tại vị trí gen CD163

Sự gắn đặc hiệu vị trí của axit nucleic ngoại sinh tại vị trí gen CD163 có thể được hoàn thành bằng kỹ thuật bất kỳ đã biết đối với người có kỹ năng trong lĩnh vực. Ví dụ, sự gắn của axit nucleic ngoại sinh tại vị trí gen CD163 có thể bao gồm bước cho tế bào (ví dụ, tế bào phân lập hoặc tế bào ở mô hoặc sinh vật) tiếp xúc với phân tử axit nucleic bao gồm axit nucleic ngoại sinh. Phân tử axit nucleic như vậy có thể bao gồm trình tự nucleotit chặn axit nucleic ngoại sinh mà nó tạo thuận lợi cho sự tái tổ hợp tương đồng giữa phân tử axit nucleic và ít nhất một vị trí gen CD163. Trình tự nucleotit chặn axit nucleic ngoại sinh mà nó tạo thuận lợi cho sự tái tổ hợp tương đồng có thể bổ trợ cho các nucleotit nội sinh của vị trí gen CD163. Theo cách khác, trình tự nucleotit chặn axit nucleic ngoại sinh mà nó tạo thuận lợi cho sự tái tổ hợp tương đồng có thể bổ trợ cho các nucleotit ngoại sinh được gắn vào trước đó. Nhiều axit nucleic ngoại sinh có thể được gắn tại một vị trí gen CD163, như trong xếp đặt gen.

Sự gắn axit nucleic tại vị trí gen CD163 có thể được tạo thuận lợi (ví dụ, được xúc tác) bằng bộ máy tế bào nội sinh của tế bào chủ, ví dụ như và không bị giới hạn,

ADN nội sinh và các enzym recombinaza nội sinh. Theo cách khác, sự gắn axit nucleic tại vị trí gen CD163 có thể được tạo thuận lợi bằng một hoặc nhiều yếu tố (ví dụ, polypeptit) mà các yếu tố này được cung cấp cho tế bào chủ. Ví dụ, (các) nucleaza, (các) recombinaza, và/hoặc ligaza polypeptit có thể được cung cấp (độc lập hoặc ở dạng một phần của polypeptit khảm) bằng cách cho polypeptit tiếp xúc với tế bào chủ, hoặc bằng cách biểu hiện polypeptit bên trong tế bào chủ. Do đó, axit nucleic bao gồm trình tự nucleotit mã hóa ít nhất một nucleaza, recombinaza, và/hoặc ligaza polypeptit có thể được đưa vào trong tế bào chủ, đồng thời hoặc tuần tự cùng với axit nucleic cần được gắn theo kiểu đặc hiệu vị trí tại vị trí gen CD163, trong đó ít nhất một nucleaza, recombinaza, và/hoặc ligaza polypeptit được biểu hiện từ trình tự nucleotit trong tế bào chủ.

Các polypeptit liên kết ADN

Sự gắn đặc hiệu vị trí có thể được hoàn thành bằng cách sử dụng các yếu tố mà có khả năng nhận diện và liên kết với các trình tự nucleotit cụ thể, ví dụ, trong hệ gen của sinh vật chủ. Ví dụ, nhiều protein bao gồm các vùng polypeptit có khả năng nhận diện và liên kết với ADN theo cách đặc hiệu vị trí. Trình tự ADN mà được nhận diện bằng polypeptit liên kết ADN có thể được gọi là trình tự "đích". Các vùng polypeptit mà có khả năng nhận diện và liên kết với ADN theo cách đặc hiệu vị trí thường được gấp chính xác và hoạt động độc lập để liên kết ADN theo cách đặc hiệu vị trí, thậm chí ngay cả khi được biểu hiện trong polypeptit khác với protein mà vùng ban đầu được phân lập từ protein này. Tương tự, các trình tự đích để nhận diện và liên kết bởi polypeptit liên kết ADN thường có khả năng được nhận diện và được liên kết bằng các polypeptit như vậy, ngay cả khi có mặt trong các cấu trúc ADN lớn (ví dụ, nhiễm sắc thể), đặc biệt khi vị trí mà trình tự đích nằm ở đó là vị trí đã biết là có thể tiếp cận đối với các protein tế bào hòa tan được (ví dụ, gen).

Trong khi polypeptit liên kết ADN nhận diện từ protein tồn tại trong tự nhiên điển hình liên kết với trình tự nucleotit hoặc motif riêng biệt (ví dụ, trình tự nhận diện tương đồng), các phương pháp tồn tại và đã biết trong lĩnh vực để sửa đổi nhiều polypeptit liên kết ADN như vậy để nhận diện trình tự nucleotit hoặc motif khác. Các polypeptit liên kết ADN bao gồm, ví dụ và không giới hạn: các vùng liên kết ADN ngón

tay kẽm; khóa kéo leuxin; các vùng liên kết ADN UPA; GAL4; TAL; LexA; phần tử ức chế Tet; LacI; và thụ thể hormon steroit.

Ví dụ, polypeptit liên kết ADN có thể là ngón tay kẽm. Các motif ngón tay kẽm riêng biệt có thể được thiết kế để hướng đích và liên kết đặc hiệu với vị trí bất kỳ trong số nhiều vị trí ADN. Các polypeptit ngón tay kẽm Cys_2His_2 hợp quy (cũng như Cys_3His không hợp quy) liên kết với ADN bằng cách đưa xoắn α- vào trong nếp chính của xoắn kép ADN đích. Việc nhận diện ADN bằng ngón tay kẽm là theo khối kết cấu; mỗi một ngón tay trước hết tiếp xúc 3 cặp bazơ liên tiếp ở đích, và một số gốc chính trong polypeptit trung gian cho sự nhận diện. Bằng việc bao gồm nhiều vùng liên kết ADN ngón tay kẽm trong endonucleaza hướng đích, tính đặc hiệu liên kết ADN của endonucleaza hướng đích có thể được gia tăng thêm (và do đó, tính đặc hiệu của các tác động điều hòa gen bất kỳ có được bằng cách đó cũng có thể được gia tăng). Ví dụ, xem trong Urnov et al. (2005) Nature 435:646-51. Do đó, một hoặc nhiều polypeptit liên kết ADN ngón tay kẽm có thể được thiết kế và sử dụng sao cho endonucleaza hướng đích được đưa vào trong tế bào chủ tương tác với trình tự ADN mà nó là độc nhất bên trong hệ gen của tế bào chủ.

Tốt hơn, protein ngón tay kẽm không xuất hiện một cách tự nhiên ở chỗ, nó được thiết kế để liên kết với vị trí đích lựa chọn. Ví dụ, xem trong Beerli et al. (2002) Nature Biotechnol. 20:135-141; Pabo et al. (2001) Ann. Rev. Biochem. 70:313-340; Isalan et al. (2001) Nature Biotechnol. 19:656-660; Segal et al. (2001) Curr. Opin. Biotechnol. 12:632-637; Choo et al. (2000) Curr. Opin. Struct. Biol. 10:411-416; các patent Mỹ số 6,453,242; 6,534,261; 6,599,692; 6,503,717; 6,689,558; 7,030,215; 6,794,136; 7,067,317; 7,262,054; 7,070,934; 7,361,635; 7,253,273; và các công bố đơn yêu cầu cấp patent Mỹ số 2005/0064474; 2007/0218528; 2005/0267061.

Vùng liên kết ngón tay kẽm thiết kế có thể có tính đặc hiệu liên kết mới, so với protein ngón tay kẽm xuất hiện tự nhiên. Các phương pháp thiết kế bao gồm, nhưng không bị giới hạn bởi, thiết kế hợp lý và các kiểu lựa chọn khác. Thiết kế hợp lý bao gồm, ví dụ, sử dụng các cơ sở dữ liệu bao gồm bộ ba (hoặc bộ bốn) trình tự nucleotit và các trình tự axit amin ngón tay kẽm riêng biệt, trong đó mỗi một bộ ba hoặc bộ bốn trình tự nucleotit được kết hợp với một hoặc nhiều các trình tự axit amin của các ngón

tay kẽm mà chúng liên kết với trình tự bộ ba hoặc bộ bốn cụ thể. Ví dụ, xem trong tài liệu patent Mỹ số 6,453,242 và 6,534,261.

Các phương pháp lựa chọn minh họa bao gồm hiển thị thể thực khuẩn và hệ hai thể lai, được bộc lộ trong các patent Mỹ số 5,789,538; 5,925,523; 6,007,988; 6,013,453; 6,410,248; 6,140,466; 6,200,759; và 6,242,568; cũng như WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 và GB 2,338,237. Ngoài ra, việc tăng cường tính đặc hiệu liên kết của vùng liên kết ngón tay kẽm đã được mô tả, ví dụ, trong tài liệu WO 02/077227.

Ngoài ra, như được bộc lộ trong các tài liệu tham chiếu này và các tài liệu tham chiếu khác, các vùng ngón tay kẽm và/hoặc các protein ngón tay kẽm đa ngón tay có thể được liên kết với nhau sử dụng các trình tự liên kết thích hợp bất kỳ, bao gồm ví dụ, các trình tự liên kết có độ dài gồm 5 hoặc nhiều hơn 5 axit amin. Cũng xem trong patent Mỹ số 6,479,626; 6,903,185; và 7,153,949 về các trình tự liên kết minh họa có độ dài 6 hoặc nhiều hơn 6 axit amin. Các protein mô tả ở đây có thể bao gồm tổ hợp bất kỳ của các liên kết thích hợp giữa các ngón tay kẽm riêng biệt của protein.

Lựa chọn các vị trí đích: ZFP và các phương pháp thiết kế và cấu trúc các protein dung hợp (và các polynucleotit mã hóa chúng) là đã biết đối với người có kỹ năng trong lĩnh vực và được mô tả chi tiết trong patent Mỹ các số 6,140,0815; 789,538; 6,453,242; 6,534,261; 5,925,523; 6,007,988; 6,013,453; 6,200,759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970 WO 01/88197; WO 02/099084; WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 và WO 03/016496.

Ngoài ra, như được bộc lộ trong các tài liệu tham chiếu này và các tài liệu tham chiếu khác, các vùng ngón tay kẽm và/hoặc các protein ngón tay kẽm đa ngón tay có thể được liên kết với nhau sử dụng các trình tự liên kết thích hợp bất kỳ, bao gồm ví dụ, các trình tự liên kết có độ dài gồm 5 hoặc nhiều hơn 5 axit amin. Cũng xem trong patent Mỹ số 6,479,626; 6,903,185; và 7,153,949 về các trình tự liên kết minh họa có độ dài 6 hoặc nhiều hơn 6 axit amin. Các protein mô tả ở đây có thể bao gồm tổ hợp bất kỳ của các liên kết thích hợp giữa các ngón tay kẽm riêng biệt của protein.

Theo cách khác, polypeptit liên kết ADN là vùng liên kết ADN từ GAL4. GAL4 là tác nhân hoạt chuyển theo khối trong *Saccharomyces cerevisiae*, nhưng nó cũng hoạt

động như một tác nhân hoạt chuyển trong nhiều sinh vật khác. Ví dụ, xem trong Sadowski et al. (1988) Nature 335:563-4. Trong hệ điều hòa này, sự biểu hiện của gen mã hóa cho các enzym của con đường chuyển hóa galactoza trong *S. cerevisiae* được điều hòa một cách chặt chẽ bằng nguồn cacbon sẵn có. Johnston (1987) Microbiol. Rev. 51:458-76. Việc điều khiển phiên mã của các enzym chuyển hóa này được trung gian bởi sự tương tác giữa protein điều hòa dương, GAL4, và trình tự ADN đối xứng 17 bp mà GAL4 liên kết đặc hiệu với trình tự này (trình tự hoạt hóa thượng dòng (UAS)).

GAL4 tự nhiên gồm có 881 gốc axit amin, với phân tử lượng là 99 kDa. GAL4 bao gồm các vùng tự phát về chức năng, các hoạt tính kết hợp của các vùng này giải thích cho hoạt tính của GAL4 *in vivo*. Ma và Ptashne (1987) Cell 48:847-53); Brent and Ptashne (1985) Cell 43(3 Pt 2):729-36. 65 axit amin có N tận cùng của GAL4 bao gồm vùng liên kết ADN GAL4. Keegan et al. (1986) Science 231:699-704; Johnston (1987) Nature 328:353-5. Liên kết đặc hiệu trình tự đòi hỏi sự hiện diện của cation hóa trị hai được phối hợp bằng 6 gốc Cys có mặt trong miền liên kết ADN. Miền chứa cation phối hợp tương tác và nhận diện bộ ba CCG bảo tồn tại mỗi một đầu của UAS 17 bp thông qua các tiếp xúc trực tiếp với nếp chính của xoắn ADN. Marmorstein et al. (1992) Nature 356:408-14. Chức năng liên kết ADN của protein định vị các miền kích hoạt phiên mã tận cùng C ở vùng lân cận của vùng gen khởi động, sao cho các miền kích hoạt có thể định hướng sự phiên mã.

Các polypeptit liên kết ADN bổ sung có thể được sử dụng bao gồm, ví dụ và không giới hạn, trình tự liên kết từ gen có thể cảm ứng bởi AVRBS3; trình tự liên kết tương đồng từ gen có thể cảm ứng bởi AVRBS3 hoặc trình tự liên kết tổng hợp được thiết kế từ nó (ví dụ, vùng liên kết ADN UPA); TAL; LexA (ví dụ, xem trong Brent & Ptashne (1985), supra); LacR (ví dụ, xem trong Labow et al. (1990) Mol. Cell. Biol. 10:3343-56; Baim et al. (1991) Proc. Natl. Acad. Sci. USA 88(12):5072-6); thụ thể hormon steroit (Elliston et al. (1990) J. Biol. Chem. 265:11517-121); phần tử ức chế Tet (U.S. Pat. No. 6,271,341) và phần tử ức chế Tet đột biến mà nó liên kết với trình tự vận hành tet khi có mặt, nhưng khi không có mặt, của tetracycline (Tc); vùng liên kết ADN của NF-kappaB; và các thành phần của hệ điều hòa được mô tả trong tài liệu của tác giả Wang et al. (1994) Proc. Natl. Acad. Sci. USA 91(17):8180-4, sử dụng sự dung hợp của GAL4, thụ thể hormon, và VP16.

Vùng liên kết ADN của một hoặc nhiều nucleaza sử dụng trong các phương pháp và chế phẩm mô tả ở đây có thể bao gồm miền liên kết ADN tác động TAL xuất hiện tự nhiên hoặc được thiết kế (không xuất hiện một cách tự nhiên). Ví dụ, xem trong công bố đơn yêu cầu cấp patent Mỹ số 2011/0301073.

Theo cách khác, nucleaza có thể bao gồm hệ CRISPR/Cas. Các hệ như vậy bao gồm vị trí trình tự sắp xếp kiểu lặp đi lặp lại xen khoảng trống (CRISPR - clustered regularly interspaced short palindromic repeats), mà nó mã hóa cho các thành phần ARN của hệ, và vị trí Cas (kết hợp CRISPR), mà nó mã hóa cho các protein (Jansen et al., 2002. Mol. Microbiol. 43: 1565-1575; Makarova et al., 2002. Nucleic Acids Res. 30: 482-496; Makarova et al., 2006. Biol. Direct 1: 7; Haft et al., 2005. PLoS Comput. Biol. 1: e60). Các vị trí CRISPR trong vật chủ vi khuẩn chứa một tổ hợp các gen Cas cũng như các yếu tố ARN không mã hóa có khả năng lập trình tính đặc hiệu của sự tách axit nucleic qua trung gian CRISPR.

CRISPR Type II là một trong các hệ đặc trưng rõ nhất và thực hiện làm đứt gãy sợi kép ADN đích trong tự nhiên theo 4 bước tuần tự. Trước tiên, hai ARN không mã hóa, dàn tiền crARN và tracrARN, được phiên mã tự vị trí CRISPR. Thứ hai, tracrARN lai với các vùng lặp của tiền crARN và trung gian cho việc xử lý tiền crARN thành crARN chín chứa các trình tự đệm riêng biệt. Thứ ba, phức crARN chín:tracrARN hướng Cas9 vào ADN đích thông qua sự ghép cặp bazơ Wastson-Crick giữa đoạn đệm trên crARN và đoạn tiền đệm trên ADN đích tiếp với motif liền kề đoạn tiền đệm (PAM), yêu cầu bổ sung để nhận diện đích. Sau cùng, Cas9 trung gian cho quá trình tách ADN đích để tạo ra đứt gãy sợi kép bên trong đoạn tiền đệm.

Để sử dụng hệ CRISPR/Cas để tạo ra các đoạn xen và đoạn xóa đích, hai ARN không mã hóa (crARN và tracrARN) có thể được thay thế bằng ARN duy nhất được tham chiếu là ARN dẫn hướng (gARN). Hoạt động của hệ CRISPR/Cas bao gồm ba bước: (i) gắn các trình tự ADN ngoại sinh vào trong dàn CRISPR để ngăn chặn sự tấn công trong tương lai, trong một quá trình được gọi là "làm thích nghi," (ii) sự biểu hiện của các protein liên quan, cũng như biểu hiện và xử lý dàn, tiếp theo bằng (iii) quá trình can thiệp trung gian bởi ARN với axit nucleic ngoại lai. Trong tế bào vi khuẩn, nhiều protein Cas có liên quan đến chức năng tự nhiên của hệ CRISPR/Cas và đóng vai trò trong các chức năng như gắn ADN ngoại lai, v.v..

Protein Cas có thể là "dẫn xuất chức năng" của protein Cas xuất hiện tự nhiên. "Dẫn xuất chức năng" của polypeptit trình tự tự nhiên là một hợp chất có đặc tính sinh học định tính giống như polypeptit trình tự tự nhiên. "Các dẫn xuất chức năng" bao gồm, nhưng không bị giới hạn bởi, các đoạn có trình tự tự nhiên và các dẫn xuất của polypeptit trình tự tự nhiên và các đoạn của nó, với điều kiện chúng có hoạt tính sinh học giống như polypeptit trình tự tự nhiên tương ứng. Hoạt tính sinh học dự định ở đây là khả năng của dẫn xuất chức năng để lai cơ chất ADN vào trong các đoạn. Thuật ngữ "dẫn xuất" bao hàm cả các biến thể trình tự axit amin của polypeptit, sự sửa đổi đồng hóa trị, và sự dung hợp của chúng. Các dẫn xuất thích hợp của polypeptit Cas hoặc đoạn của nó bao gồm nhưng không bị giới hạn bởi các thể đột biến, sự dung hợp, sự sửa đổi đồng hóa trị của protein Cas hoặc đoạn của nó. Protein Cas, protei này bao gồm protein Cas hoặc đoạn của nó, cũng như các dẫn xuất của protein Cas hoặc đoạn của nó, có thể thu được từ tế bào hoặc tổng hợp theo phương thức hóa học hoặc bằng sự kết hợp của hai phương thức này. Tế bào có thể là tế bào mà tạo ra protein Cas một cách tự nhiên, hoặc tế bào mà tạo ra protein Cas một cách tự nhiên và được thiết kế theo phương thức di truyền để tạo ra protein Cas nội sinh ở mức biểu hiện cao hơn hoặc để tạo ra protein Cas tự một axit nucleic được đưa vào theo phương thức ngoại sinh, axit nucleic này mã hóa cho Cas mà nó tương tự hoặc khác với Cas nội sinh. Trong một số trường hợp, tế bào không tạo ra protein Cas một cách tự nhiên và được thiết kế theo phương thức di truyền để tạo ra protein Cas.

Polypeptit liên kết ADN có thể nhận diện và liên kết đặc hiệu với trình tự nucleotit đích được bao gồm bên trong axit nucleic hệ gen của sinh vật chủ. Số mẫu bất kỳ riêng biệt nào về trình tự nucleotit đích có thể được tìm thấy trong hệ gen vật chủ trong một số ví dụ. Trình tự nucleotit đích có thể là hiếm bên trong hệ gen của sinh vật (ví dụ, ít hơn khoảng 10, khoảng 9, khoảng 8, khoảng 7, khoảng 6, khoảng 5, khoảng 4, khoảng 3, khoảng 2, hoặc khoảng 1 bản sao của trình tự đích có thể tồn tại trong hệ gen). Ví dụ, trình tự nucleotit đích có thể nằm ở vị trí duy nhất bên trong hệ gen của sinh vật. Các trình tự nucleotit đích có thể, ví dụ và không giới hạn, được phân tán ngẫu nhiên trong toàn bộ hệ gen đối với nhau; nằm ở các nhóm liên kết khác nhau trong hệ gen; nằm ở cùng nhóm liên kết; nằm ở các nhiễm sắc thể khác nhau; nằm ở cùng nhiễm sắc thể; nằm trong hệ gen ở các vị trí mà được biểu hiện trong các điều kiện tương tự trong sinh vật (ví dụ, dưới sự điều khiển của cùng các yếu tố điều hòa, hoặc các yếu tố

điều hòa cơ bản giống nhau về mặt chức năng); và nằm sát với nhau trong hệ gen (ví dụ, các trình tự đích có thể được bao gồm bên trong các axit nucleic được hợp nhất làm các chuỗi khảm tại các vị trí hệ gen).

Các endonucleaza hướng đích

Polypeptit liên kết ADN mà nhận diện và liên kết đặc hiệu với trình tự nucleotit đích có thể được bao gồm bên trong polypeptit khảm, để truyền liên kết đặc hiệu cho trình tự đích qua polypeptit khảm. Trong các ví dụ, polypeptit khảm như vậy có thể bao gồm, ví dụ và không giới hạn, các polypeptit nucleaza, recombinaza, và/hoặc ligaza, như các polypeptit được mô tả ở trên. Các polypeptit bao gồm polypeptit liên kết ADN và polypeptit nucleaza, recombinaza, và/hoặc ligaza cũng có thể bao gồm các motif và/hoặc miền polypeptit chức năng khác, ví dụ như và không giới hạn: trình tự đệm nằm ở giữa các polypeptit chức năng trong protein khảm; peptit dẫn; peptit mà nó hướng protein dung hợp vào bào quan (ví dụ, nhân); polypeptit mà được tách bằng enzym tế bào; các thẻ peptit (ví dụ, Myc, His, v.v..); và các trình tự axit amin khác không can thiệp vào chức năng của polypeptit khảm.

Các polypeptit chức năng (ví dụ, polypeptit liên kết ADN và các polypeptit nucleaza) trong polypeptit khảm có thể được liên kết theo kiểu hoạt động được. Các polypeptit chức năng của polypeptit khảm có thể được liên kết theo kiểu hoạt động được bằng sự biểu hiện của chúng từ polynucleotit duy nhất mã hóa ít nhất các polypeptit chức năng được gắn với nhau trong khung, để tạo ra gen khảm mã hóa cho protein khảm. Theo cách khác, các polypeptit chức năng của polypeptit khảm có thể được liên kết theo kiểu hoạt động được bằng các phương tiện khác, như bằng liên kết chéo của các polypeptit biểu hiện độc lập.

Polypeptit liên kết ADN, hoặc ARN dẫn hướng và nhận diện và liên kết đặc hiệu với trình tự nucleotit đích có thể được bao gồm bên trong protein phân lập tự nhiên (hoặc thể đột biến của nó), trong đó protein phân lập tự nhiên hoặc thể đột biến của nó cũng bao gồm polypeptit nucleaza (và có thể cũng bao gồm polypeptit recombinaza và/hoặc ligaza). Các ví dụ về protein phân lập như vậy bao gồm TALEN, recombinaza (ví dụ, Cre, Hin, Tre, và FLP recombinaza), CRISPR/Cas9 dẫn hướng bằng ARN, và meganucleaza.

Như được sử dụng ở đây, thuật ngữ "endonucleaza hướng đích" đề cập đến protein phân lập tự nhiên hoặc thiết kế và các thể đột biến của nó mà bao gồm polypeptit liên kết ADN hoặc ARN dẫn hướng và polypeptit nucleaza, cũng như đến các polypeptit khảm bao gồm polypeptit liên kết ADN hoặc ARN dẫn hướng và nucleaza. Endonucleaza hướng đích bất kỳ bao gồm polypeptit liên kết ADN hoặc ARN dẫn hướng mà nhận diện và liên kết đặc hiệu với trình tự nucleotit đích được bao gồm bên trong vị trí gen CD163 (ví dụ, vì trình tự đích được bao gồm bên trong trình tự tự nhiên tại vị trí, hoặc vì trình tự đích đã được đưa vào trong vị trí, ví dụ, bằng cách tái tổ hợp) có thể được sử dụng.

Một số ví dụ về các polypeptit khảm thích hợp bao gồm, không giới hạn, các tổ hợp polypeptit sau: các polypeptit liên kết ADN ngón tay kẽm; polypeptit nucleaza FokI; các miền TALE; khóa kéo leuxin; các motif liên kết ADN yếu tố phiên mã; và các miền nhận diện và/hoặc tách ADN phân lập từ, ví dụ và không giới hạn, TALEN, recombinaza (ví dụ, Cre, Hin, RecA, Tre, và FLP recombinaza), CRISPR/Cas9 dẫn hướng bằng ARN, meganucleaza; và các yếu tố khác đã biết đối với chuyên gia trong lĩnh vực. Các ví dụ cụ thể bao gồm protein khảm bao gồm polypeptit liên kết ADN đặc hiệu vị trí và polypeptit nucleaza. Các polypeptit khảm có thể được thiết kế bằng các phương pháp đã biết đối với người có kỹ năng trong lĩnh vực để thay đổi trình tự nhận diện của polypeptit liên kết ADN được bao gồm bên trong polypeptit khảm, để hướng đích polypeptit khảm vào trình tự nucleotit mong muốn cụ thể.

Polypeptit khảm có thể bao gồm vùng liên kết ADN (ví dụ, ngón tay kẽm, miền tác động TAL, v.v..) và miền nucleaza (phân cắt). Miền phân cắt có thể là dị tương đồng với vùng liên kết ADN, ví dụ vùng liên kết ADN ngón tay kẽm và miền phân cắt từ nucleaza hoặc vùng liên kết ADN TALEN và miền phân cắt, hoặc vùng liên kết ADN meganucleaza và miền phân cắt từ nucleaza khác. Các miền phân cắt dị tương đồng có thể thu được từ endonucleaza hoặc exonucleaza bất kỳ. Các endonucleaza minh họa mà miền phân cắt có thể xuất phát từ đó bao gồm, nhưng không bị giới hạn bởi, các endonucleaza giới hạn và endonucleaza tìm nhà. Ví dụ, xem trong tài liệu 2002-2003 Catalogue, New England Biolabs, Beverly, Mass.; and Belfort et al. (1997) Nucleic Acids Res. 25:3379-3388. Các enzym bổ sung mà phân cắt ADN là đã biết (ví dụ, xem trong 51 Nuclease; mung bean nuclease; pancreatic DNAse I; micrococcal nuclease; yeast HO endonuclease; cũng xem trong tài liệu của Linn et al. (eds.) Nucleases, Cold

Spring Harbor Laboratory Press, 1993). Một hoặc nhiều trong số các enzym này (hoặc các đoạn chức năng của nó) có thể được sử dụng làm nguồn của các miền phân cắt và các miền bán phân cắt.

Tương tự, miền bán phân cắt có thể xuất phát từ nucleaza bất kỳ hoặc một phần của nó, như được nêu ở trên, nó đòi hỏi sự dime hóa cho hoạt tính phân cắt. Nói chung, hai protein dung hợp được đòi hỏi để phân cắt nếu các protein dung hợp bao gồm các miền bán phân cắt. Theo cách khác, protein đơn bao gồm hai miền bán phân cắt có thể được sử dụng. Hai miền bán phân cắt có thể xuất phát từ cùng endonucleaza (hoặc các đoạn chức năng của nó), hoặc each miền bán phân cắt có thể xuất phát from a different endonucleaza (hoặc các đoạn chức năng của nó). Ngoài ra, các vị trí đích của hai protein dung hợp tốt hơn được bố trí, liên quan với nhau, sao cho sự liên kết của hai protein dung hợp vào các vị trí đích tương ứng đặt các miền bán phân cắt theo hướng không gian với nhau, cho phép các miền bán phân cắt tạo ra miền phân cắt chức năng, ví dụ, bằng cách dime hóa. Do đó, mép gần của các vị trí đích có thể được tách riêng bằng 5-8 nucleotit hoặc bằng 15-18 nucleotit. Tuy nhiên, số nguyên bất kỳ của các nucleotit hoặc cặp nucleotit, có thể xen vào giữa hai vị trí đích (ví dụ, từ 2 đến 50 cặp nucleotit hoặc nhiều hơn). Nói chung, vị trí phân cắt nằm ở giữa các vị trí đích.

Các endonucleaza giới hạn (enzym giới hạn) có mặt trong nhiều loài và có khả năng liên kết đặc hiệu trình tự với ADN (tại vị trí nhận diện), và phân cắt ADN tại hoặc gần vị trí liên kết, ví dụ, sao cho một hoặc nhiều trình tự ngoại sinh (các gen cho/gen chuyển) được hợp nhất tại hoặc gần các vị trí liên kết (đích). Một số enzym giới hạn (ví dụ, Type IIS) phân cắt ADN ở các vị trí được di chuyển ra khỏi vị trí nhận diện và có liên kết có thể tách ra được và các miền phân cắt. Ví dụ, enzym Type IIS Fok I xúc tác cho sự phân cắt sợi kép của ADN, tại 9 nucleotit tính từ vị trí nhận diện của nó trên một sợi và 13 nucleotit tính từ vị trí nhận diện của nó trên sợi kia. Ví dụ, xem trong tài liệu patent Mỹ số 5,356,802; 5,436,150 và 5,487,994; cũng như Li et al. (1992) Proc. Natl. Acad. Sci. USA 89:4275-4279; Li et al. (1993) Proc. Natl. Acad. Sci. USA 90:2764-2768; Kim et al. (1994a) Proc. Natl. Acad. Sci. USA 91:883-887; Kim et al. (1994b) J. Biol. Chem. 269:31.978-31.982. Do đó, các protein dung hợp có thể bao gồm miền phân cắt (hoặc miền bán phân cắt) từ ít nhất một enzym giới hạn Type IIS và một hoặc nhiều vùng liên kết ngón tay kẽm, chúng có thể hoặc không thể được thiết kế.

Các enzym giới hạn Type IIS mà miền phân cắt của nó có thể tách ra khỏi miền liên kết đó là Fok I. Enzym cụ thể này có hoạt tính như một dime. Bitinaite et al. (1998) Proc. Natl. Acad. Sci. USA 95: 10.570-10.575. Do đó, đối với mục đích của sáng chế này, phần enzym Fok I sử dụng trong các protein dung hợp đã bộc lộ được xem là miền bán phân cắt. Do đó, để phân cắt sợi kép hướng đích và/hoặc thay thế hướng đích của các trình tự tế bào sử dụng sự dung hợp ngón tay kẽm-Fok I, hai protein dung hợp, mỗi loại bao gồm miền bán phân cắt FokI, có thể được sử dụng để khôi phục lại miền phân cắt có hoạt tính xúc tác. Theo cách khác, single phân tử polypeptit chứa miền liên kết ADN và hai miền bán phân cắt Fok I cũng có thể được sử dụng.

Miền phân cắt hoặc miền bán phân cắt có thể là phần bất kỳ của protein mà giữ được hoạt tính phân cắt, hoặc giữ được khả năng multime hóa (ví dụ, dime hóa) để tạo ra miền phân cắt chức năng.

Các enzym giới hạn Type IIS được mô tả trong công bố đơn yêu cầu cấp patent Mỹ số 2007/0134796. Các enzym giới hạn bổ sung cũng chứa liên kết có thể tách ra được và các miền phân cắt, và các enzym này được dự định bởi sáng chế này. Ví dụ, xem trong Roberts et al. (2003) Nucleic Acids Res. 31:418-420.

Miền phân cắt có thể bao gồm một hoặc nhiều miền bán phân cắt thiết kế (còn được gọi là các thể đột biến miền dime hóa) mà chúng giảm thiểu đến mức thấp nhất hoặc ngăn chặn sự homodime hóa, như được mô tả, ví dụ, trong các công bố yêu cầu cấp patent Mỹ số 2005/0064474; 2006/0188987 và 2008/0131962.

Theo cách khác, các nucleaza có thể được lắp ráp *in vivo* tại vị trí đích axit nucleic sử dụng công nghệ được gọi là "tách-enzym" (ví dụ, xem trong công bố đơn yêu cầu cấp patent Mỹ số 20090068164). Các thành phần của các enzym tách này có thể được biểu hiện trên các cấu trúc biểu hiện riêng biệt hoặc có thể được liên kết trong một khung đọc mở, trong đó các thành phần riêng biệt được tách riêng, ví dụ, bằng peptit 2A bán phân cắt hoặc trình tự IRES. Các thành phần có thể là các vùng liên kết ngón tay kẽm riêng biệt hoặc các miền của miền liên kết axit nucleic meganucleaza.

Các nucleaza ngón tay kẽm

Polypeptit khảm có thể bao gồm nucleaza ngón tay kẽm thiết kế tùy chỉnh (ZFN) mà nó có thể được thiết kế để phân phối đứt gãy ADN sợi kép đặc hiệu vị trí hướng đích vào vị trí mà axit nucleic ngoại sinh, hoặc ADN cho, có thể được hợp nhất ở đó

(xem trong công bố patent Mỹ số 2010/0257638). ZFN là các polypeptit khảm chứa miền phân cắt không đặc hiệu từ endonucleaza giới hạn (ví dụ, FokI) và polypeptit vùng liên kết ADN ngón tay kẽm. Ví dụ, xem trong Huang et al. (1996) J. Protein Chem. 15:481-9; Kim et al. (1997a) Proc. Natl. Acad. Sci. USA 94:3616-20; Kim et al. (1996) Proc. Natl. Acad. Sci. USA 93:1156-60; Kim et al. (1994) Proc Natl. Acad. Sci. USA 91:883-7; Kim et al. (1997b) Proc. Natl. Acad. Sci. USA 94:12875-9; Kim et al. (1997c) Gene 203:43-9; Kim et al. (1998) Biol. Chem. 379:489-95; Nahon and Raveh (1998) Nucleic Acids Res. 26:1233-9; Smith et al. (1999) Nucleic Acids Res. 27:674-81. ZFN có thể bao gồm các miền liên kết ADN ngón tay kẽm không hợp qy (xem trong công bố patent Mỹ 2008/0182332). Endonucleaza giới hạn FokI cần phải dime hóa thông qua miền nucleaza để phân cắt ADN và đưa vào đứt gãy sợi kép. Sau đó, ZFN chứa miền nucleaza từ endonucleaza như vậy cũng đòi hỏi sự dime hóa của miền nucleaza để phân cắt ADN đích. Mani et al. (2005) Biochem. Biophys. Res. Commun. 334:1191-7; Smith et al. (2000) Nucleic Acids Res. 28:3361-9. Quá trình dime hóa của ZFN có thể được tạo thuận lợi bằng hai vị trí liên kết ADN liền kề, hướng đối diện nhau. Id.

Phương pháp để gắn đặc hiệu vị trí axit nucleic ngoại sinh vào trong ít nhất một vị trí gen CD163 của vật chủ có thể bao gồm bước đưa ZFN vào tế bào của vật chủ, trong đó ZFN nhận diện và liên kết với trình tự nucleotit đích, trong đó trình tự nucleotit đích được bao gồm bên trong ít nhất một vị trí gen CD163 của vật chủ. Trong một số ví dụ, trình tự nucleotit đích không được bao gồm bên trong hệ gen của vật chủ ở vị trí khác bất kỳ ngoài ít nhất một vị trí gen CD163. Ví dụ, polypeptit liên kết ADN của ZFN có thể được thiết kế để nhận diện và liên kết với trình tự nucleotit đích được nhận diện bên trong ít nhất một vị trí gen CD163 (ví dụ, bằng cách giải trình tự vị trí gen CD163). Phương pháp để gắn đặc hiệu vị trí của axit nucleic ngoại sinh vào trong ít nhất một vị trí đặc trưng CD163 của vật chủ, phương pháp bao gồm bước đưa vào trong tế bào của vật chủ a ZFN may also bao gồm đưa vào trong tế bào một axit nucleic ngoại sinh, trong đó sự tái tổ hợp của axit nucleic ngoại sinh vào trong axit nucleic của vật chủ bao gồm ít nhất một vị trí gen CD163 được tạo thuận lợi bằng sự nhận diện và liên kết đặc hiệu vị trí của ZFN với trình tự đích (và sự phân cắt sau đó của axit nucleic bao gồm vị trí gen CD163).

Các axit nucleic ngoại sinh tùy ý để gắn vào vị trí gen CD163

Các axit nucleic ngoại sinh để gắn vào vị trí gen CD163 bao gồm: axit nucleic ngoại sinh để gắn đặc hiệu vị trí tại ít nhất một vị trí gen CD163, ví dụ và không giới hạn, ORF; axit nucleic bao gồm trình tự nucleotit mã hóa cho endonucleaza hướng đích; và vectơ bao gồm ít nhất một hoặc cả hai loại nêu trên. Do đó, các axit nucleic cụ thể bao gồm trình tự nucleotit mã hóa cho polypeptit, các trình tự nucleotit cấu trúc, và/hoặc các vị trí nhận dạng và liên kết polypeptit liên kết ADN.

Các phân tử axit nucleic ngoại sinh tuỳ ý để gắn đặc hiệu vị trí

Như đã lưu ý ở trên, việc gắn trình tự ngoại sinh (còn được gọi là "trình tự cho" hoặc "gen cho" hoặc "gen chuyển") được đề cập, ví dụ để biểu hiện polypeptit, sửa chữa gen đột biến hoặc để biểu hiện tăng cao gen kiểu hoang dại. Rõ ràng là, trình tự cho điển hình không tương đồng với trình tự hệ gen nơi mà nó được đưa vào. Trình tự cho có thể chứa trình tự không tương đồng được chặn bởi hai vùng tương đồng để cho phép sửa chữa định hướng tương đồng hiệu quả (HDR) tại vị trí mong muốn. Ngoài ra, trình tự cho có thể bao gồm phân tử vectơ chứa các trình tự mà không tương đồng với vùng mong muốn trong chất nhiễm sắc tế bào. Phân tử cho có thể chứa nhiều vùng tương đồng không liên tục với chất nhiễm sắc tế bào. Ví dụ, để gắn định hướng đích các trình tự thông thường không có mặt ở vùng mong muốn, các trình tự này có thể có mặt ở phân tử axit nucleic cho và được chặn bởi các vùng tương đồng với trình tự ở vùng mong muốn.

Polynucleotit cho có thể là ADN hoặc ARN, sợi đơn hoặc sợi kép và có thể được đưa vào trong tế bào ở dạng thẳng hoặc vòng. Ví dụ, xem trong công bố patent Mỹ các số 2010/0047805, 2011/0281361, 2011/0207221, và 2013/0326645. Nếu được đưa vào ở dạng thẳng, các đầu của trình tự cho có thể được bảo vệ (ví dụ, tránh thoái biến phân cắt ngoài nhân) bằng các phương pháp đã biết đối với người có kỹ năng trong lĩnh vực. Ví dụ, một hoặc nhiều gốc dideoxynucleotit được bổ sung vào đầu tận cùng 3' của phân tử thẳng và/hoặc các oligonucleotit tự tương hợp được gắn với một hoặc cả hai đầu. Ví dụ, xem trong Chang et al. (1987) Proc. Natl. Acad. Sci. USA 84:4959-4963; Nehls et al. (1996) Science 272:886-889. Các phương pháp bổ sung để bảo vệ polynucleotit ngoại sinh không bị thoái biến bao gồm, nhưng không bị giới hạn bởi, bổ sung (các) nhóm amino tận cùng và sử dụng các mối liên kết giữa các nucleotit sửa đổi, ví dụ như các gốc phosphorothioat, phosphoramidat, và O-metyl riboza hoặc deoxyriboza.

Polynucleotit có thể được đưa vào trong tế bào như một phần của phân tử vectơ có các trình tự bổ sung, ví dụ như các nguồn gốc sao chép, các đoạn khởi động và các gen mã hóa tính kháng kháng sinh. Ngoài ra, các polynucleotit cho có thể được đưa vào ở dạng axit nucleic trần, ở dạng axit nucleic tạo phức với tác nhân như liposome hoặc poloxamer, hoặc có thể được phân phối bởi các virut (ví dụ, adenovirut, AAV, herpesvirut, retrovirut, lentivirut và lentivirut thiếu integraza (IDLV)).

Đối tượng cho thông thường được gắn sao cho sự biểu hiện của nó được dẫn động bằng đoạn khởi động nội sinh tại vị trí gắn, cụ thể là đoạn khởi động mà nó dẫn động sự biểu hiện của gen nội sinh mà đối tượng cho được gắn vào trong đó (ví dụ, CD163). Tuy nhiên, rõ ràng là, đối tượng cho có thể bao gồm đoạn khởi động và/hoặc đoạn gen tăng cường, ví dụ đoạn khởi động cơ định hoặc đoạn khởi động cảm ứng hoặc đặc hiệu mô.

Ngoài ra, mặc dù không cần cho sự biểu hiện, các trình tự ngoại sinh cũng có thể bao gồm các trình tự điều hòa phiên mã hoặc dịch mã, ví dụ, các đoạn khởi động, đoạn gen tăng cường, các đoạn cách ly, các vị trí nội nhập ribosom bên trong, các trình tự mã hóa peptit 2A và/hoặc các tín hiệu polyadenyl hóa

Các axit nucleic ngoại sinh mà có thể được gắn theo cách đặc hiệu vị trí vào trong ít nhất một vị trí gen CD163, để sửa đổi vị trí gen CD163 bao gồm, ví dụ và không giới hạn, các axit nucleic bao gồm trình tự nucleotit mã hóa polypeptit mong muốn; các axit nucleic bao gồm gen nông học; các axit nucleic bao gồm trình tự nucleotit mã hóa phân tử ARNi; hoặc các axit nucleic phá vỡ gen CD163.

Axit nucleic ngoại sinh có thể được gắn tại vị trí gen CD163, để sửa đổi vị trí gen CD163, trong đó axit nucleic bao gồm trình tự nucleotit mã hóa polypeptit mong muốn, sao cho trình tự nucleotit được biểu hiện trong vật chủ từ vị trí gen CD163. Trong một số ví dụ, polypeptit mong muốn (ví dụ, protein ngoại lai) được biểu hiện từ trình tự nucleotit mã hóa polypeptit mong muốn với lượng thương mại. Trong những ví dụ như vậy, polypeptit mong muốn có thể được chiết từ tế bào chủ, mô hoặc sinh khối.

Các phân tử axit nucleic bao gồm trình tự nucleotit mã hóa endonucleaza hướng đích

Trình tự nucleotit mã hóa endonucleaza hướng đích có thể được thiết kế bằng thao tác (ví dụ, nối) các nucleotit trình tự tự nhiên mã hóa các polypeptit được bao gồm bên trong endonucleaza hướng đích. Ví dụ, trình tự nucleotit của gen mã hóa cho protein

bao gồm polypeptit liên kết ADN có thể được khảo sát để nhận diện trình tự nucleotit của gen mà tương ứng với polypeptit liên kết ADN, và trình tự nucleotit có thể được sử dụng làm thành phần của trình tự nucleotit mã hóa endonucleaza hướng đích bao gồm polypeptit liên kết ADN. Theo cách khác, trình tự axit amin của endonucleaza hướng đích có thể được sử dụng để suy diễn trình tự nucleotit mã hóa endonucleaza hướng đích, ví dụ, theo sự suy thoái của mã di truyền.

Trong các phân tử axit nucleic minh họa bao gồm trình tự nucleotit mã hóa endonucleaza hướng đích, đơn vị mã cuối cùng của trình tự polynucleotit thứ nhất mã hóa polypeptit nucleaza, và đơn vị mã thứ nhất của trình tự polynucleotit thứ hai mã hóa polypeptit liên kết ADN, có thể được tách bằng số lượng bộ ba nucleotit bất kỳ, ví dụ, không mã hóa cho intron hoặc đơn vị mã kết thúc "STOP." Tương tự, đơn vị mã cuối cùng của trình tự nucleotit mã hóa trình tự polynucleotit thứ nhất mã hóa polypeptit liên kết ADN, và đơn vị mã thứ nhất của trình tự polynucleotit thứ hai mã hóa polypeptit nucleaza, có thể được tách bằng số lượng bộ ba nucleotit bất kỳ. Đơn vị mã cuối cùng (tức là, hầu hết 3' ở trình tự axit nucleic) của trình tự polynucleotit thứ nhất mã hóa polypeptit nucleaza, và trình tự polynucleotit thứ hai mã hóa polypeptit liên kết ADN, có thể được dung hợp khi đăng ký pha bằng đơn vị mã thứ nhất của polynucleotit khác mã hóa trình tự tiếp giáp trực tiếp với nó, hoặc được tách từ đó bằng không nhiều hơn một trình tự peptit ngắn, như trình tự được mã hóa bằng tác nhân liên kết nucleotit tổng hợp (ví dụ, tác nhân liên kết nucleotit mà có thể đã được sử dụng để đạt được sự dung hợp). Các ví dụ về các trình tự polynucleotit thêm như vậy bao gồm, ví dụ và không giới hạn, các thẻ, peptit hướng đích, và các vị trí phân cắt bằng enzym. Tương tự, đơn vị mã thứ nhất của hầu hết 5' (trong trình tự axit nucleic) của các trình tự polynucleotit thứ nhất và thứ hai có thể được dung hợp trong đăng ký pha bằng đơn vị mã cuối cùng của trình tự mã hóa polynucleotit thêm trực tiếp liền kề với nó, hoặc được tách từ đó bằng không nhiều hơn trình tự peptit ngắn.

Trình tự tách các trình tự polynucleotit mã hóa các polypeptit chức năng trong endonucleaza hướng đích (ví dụ, polypeptit liên kết ADN và polypeptit nucleaza) có thể, ví dụ, bao gồm trình tự bất kỳ, sao cho trình tự axit amin được mã hóa chắc chắn không thay đổi đáng kể sự dịch mã của endonucleaza hướng đích. Do bản chất tự điều chỉnh của các polypeptit nucleaza đã biết và các polypeptit liên kết ADN đã biết, nên

các trình tự can thiệp sẽ không gây trở ngại cho các chức năng tương ứng của các cấu trúc này.

Các phương pháp bất hoạt khác

Các kỹ thuật khác nhau khác đã biết trong lĩnh vực có thể được sử dụng để bất hoạt các gen để tạo ra động vật bất hoạt gen và/hoặc để đưa các cấu trúc axit nucleic vào trong động vật để tạo ra động vật chuyển gen và để tạo ra các dòng động vật, trong đó gen bất hoạt hoặc các cấu trúc axit nucleic được gắn vào trong hệ gen. Kỹ thuật như vậy bao gồm, không giới hạn, phương pháp vi tiêm giai đoạn tiền nhân (patent Mỹ số số 4,873,191), chuyển gen qua trung gian retrovirut vào trong các dòng mầm phôi (Van der Putten et al. (1985) Proc. Natl. Acad. Sci. USA 82, 6148-1652), gen hướng đích vào trong tế bào gốc phôi (Thompson et al. (1989) Cell 56, 313-321), phương pháp xung điện của phôi (Lo (1983) Mol. Cell. Biol. 3, 1803-1814), chuyển gen qua trung gian tinh trùng (Lavitrano et al. (2002) Proc. Natl. Acad. Sci. USA 99, 14230-14235; Lavitrano et al. (2006) Reprod. Fert. Develop. 18, 19-23), và sự biến nạp *in vitro* của các tế bào sinh dưỡng, như tế bào mẫu mang trứng hoặc động vật có vú, hoặc tế bào gốc trưởng thành, thai hoặc tế bào gốc phôi, tiếp theo bằng kỹ thuật chuyển nhân (Wilmut et al. (1997) Nature 385, 810-813; và Wakayama et al. (1998) Nature 394, 369-374). Phương pháp vi tiêm giai đoạn tiền nhân, phương pháp chuyển gen trung gian qua tinh trùng, và chuyển nhân từ tế bào sinh dưỡng là những kỹ thuật đặc biệt hữu dụng. Động vật mà được sửa đổi về hệ gen là động vật trong đó tất cả các tế bào của nó có sự sửa đổi về gen, bao gồm cả các tế bào dòng mầm. Khi các phương pháp được sử dụng mà tạo ra động vật mà khảm trong sửa đổi gen của nó, động vật có thể là nội phối và thế hệ con được sửa đổi trong hệ gen có thể được chọn lọc. Ví dụ, tách dòng có thể được sử dụng để tạo ra động vật khảm nếu các tế bào của nó được sửa đổi ở trạng thái túi phôi hoặc sửa đổi hệ gen có thể diễn ra khi tế bào đơn được sửa đổi. Động vật được sửa đổi khiến cho chúng không trưởng thành về giới tính có thể là đồng hợp tử hoặc dị hợp tử để sửa đổi, tùy thuộc vào phương thức cụ thể được sử dụng. Nếu như một gen cụ thể được bất hoạt bằng sửa đổi bất hoạt, tính đồng hợp tử thông thường sẽ được đòi hỏi. Nếu một gen cụ thể được bất hoạt bằng can thiệp ARN hoặc chiến lược âm tính trội, sau đó tính dị hợp tử thường là thích hợp.

Điển hình, trong kỹ thuật vi tiêm phôi/hợp tử, cấu trúc axit nucleic hoặc mARN được đưa vào trong trứng đã thụ tinh; 1 hoặc 2 tế bào trứng thụ tinh được sử dụng làm cấu trúc nhân chứa vật liệu di truyền từ đầu tinh trùng và trứng là nhìn thấy được bên trong chất nguyên sinh. Trứng thụ tinh giai đoạn tiền nhân có thể thu được *in vitro* hoặc *in vivo* (tức là, thu hồi bằng ngoại khoa từ ống dẫn trứng của động vật cho). Trứng thụ tinh *in vitro* có thể được tạo ra như sau. Ví dụ, buồng trứng lợn có thể được tập hợp tại lò mổ và được duy trì ở nhiệt độ 22-28°C trong khi vận chuyển. Các buồng trứng có thể được rửa và phân lập để hút nang, và các nang nằm trong khoảng từ 4-8mm có thể được hút vào trong các ống ly tâm hình nón thể tích 50 mL sử dụng kim cỡ 18 và trong điều kiện chân không. Dịch nang và tế bào trứng đã hút có thể được rửa qua dụng cụ lọc sơ bộ với nhãn hiệu thương mại TL-HEPES (Minitube, Verona, Wis.). Tế bào trứng được bao quanh bởi khối mang trứng chắc chắn có thể được chọn và đặt vào trong môi trường TCM-199 OOCYTE MATURATION MEDIUM (Minitube, Verona, Wis.) có bổ sung bằng 0,1 mg/mL xystein, 10 ng/mL yếu tố sinh trưởng biểu bì, 10% dịch nang trứng lợn, 50 µM 2-mercaptoethanol, 0,5 mg/ml cAMP, 10 IU/mL mỗi một gonadotropin huyết thanh ngựa chửa (PMSG) và gonadotropin nhau thai người (hCG) trong thời gian gần 22 giờ trong không khí được làm ẩm ở nhiệt độ 38,7°C và 5% CO_2. Sau đó, các noãn bào có thể được di chuyển sang môi trường thành thục tươi TCM-199, môi trường này không chứa cAMP, PMSG hoặc hCG và được ủ trong thời gian 22 giờ nữa. Tế bào trứng thành thục có thể được lấy ra khỏi các ngăn mang trứng của chúng bằng cách tạo xoáy trong 0,1% hyaluronidaza trong thời gian 1 phút.

Đối với lợn, tế bào trứng chín có thể được thụ tinh trong 500 µl Minitube PORCPRO IVF MEDIUM SYSTEM (Minitube, Verona, Wis.) trong các đĩa thụ tinh vi ống 5 giếng. Khi chuẩn bị cho thụ tinh *in vitro* (IVF), tinh dịch lợn đực vừa tập hợp hoặc đông lạnh có thể được rửa và tái tạo huyền phù trong môi trường PORCPRO IVF Medium tới 400000 tinh trùng. Nồng độ tinh trùng có thể được phân tích bằng kỹ thuật phân tích tinh dịch có hỗ trợ bằng máy vi tính (SPERMVISION, Minitube, Verona, Wis.). Sự thụ tinh *in vitro* sau cùng có thể được thực hiện với thể tích 10µl ở nồng độ cuối gần 40 tinh trùng di động/noãn bào, tùy thuộc vào lợn đực. Toàn bộ tế bào trứng thụ tinh có thể được ủ ở nhiệt độ 38.7°C trong môi trường 5,0% CO_2 trong thời gian 6 giờ. Sáu giờ sau thụ tinh, các hợp tử giả định có thể được rửa hai lần trong NCSU-23

và được di chuyển sang 0,5mL của cùng môi trường. Hệ này có thể tạo ra 20-30% túi phôi thông thường qua hầu hết lợn đực với tỷ lệ thụ tinh nhiều tinh trùng 10-30%.

Các cấu trúc axit nucleic tuyến tính hoặc mARN có thể được tiêm vào trong một trong số các tiền nhân hoặc vào trong tế bào chất. Sau đó, trứng bơm vào có thể được chuyển sang con cái tiếp nhận (ví dụ, vào trong ống dẫn trứng của con cái tiếp nhận) và được cho phép phát triển trong con cái tiếp nhận để tạo ra động vật chuyển gen hoặc động vật chỉnh sửa gen. Cụ thể, phôi đã thụ tinh *in vitro* có thể được ly tâm ở tốc độ 15000 x g trong thời gian 5 phút để làm lắng lipit cho phép quan sát tiền nhân. Phôi có thể được tiêm bằng cách sử dụng bơm tiêm Eppendorf FEMTOJET và có thể được nuôi cấy cho tới khi hình thành túi phôi. Tỷ lệ phân cắt phôi và sự hình thành túi phôi và chất lượng có thể được ghi lại.

Phôi có thể được chuyển theo phương thức ngoại khoa vào trong tử cung của những đối tượng tiếp nhận không đồng bộ. Điển hình, 100-200 (ví dụ, 150-200) phôi có thể được đặt vào trong đoạn nối bóng - eo của ống dẫn trứng sử dụng ống thông TOMCAT® 5,5 in-sơ. Sau khi phẫu thuật, kiểm tra siêu âm thời gian thực quá trình mang thai có thể được thực hiện.

Trong kỹ thuật chuyển nhân từ tế bào sinh dưỡng, tế bào chuyển gen hoặc chỉnh sửa gen như tế bào phôi, nguyên bào sợi thai, nguyên bào sợi tai trưởng thành, hoặc tế bào lớp hạt mà bao gồm cấu trúc axit nucleic mô tả ở trên, có thể được đưa vào trong noãn bào phá bỏ nhân để thiết lập tế bào kết hợp. Tế bào trứng có thể được phá bỏ nhân bằng kỹ thuật tạo lỗ thủng trên màng trong suốt gần thể cực và sau đó, ép tế bào chất tại vùng phẫu tích. Điển hình, ống pipet tiêm có đầu vát nhọn được sử dụng để bơm tế bào chuyển gen hoặc chỉnh sửa gen vào trong noãn bào phá bỏ nhân dừng tại kỳ giảm phân 2. Theo một số định chuẩn, tế bào trứng dừng tại kỳ giảm phân-2 là trứng đúng kỳ hạn. Sau khi tạo ra phôi lợn hoặc bò (ví dụ, bằng cách dung hợp và kích hoạt noãn bào), phôi được chuyển sang ống dẫn trứng của con cái tiếp nhận, khoảng 20 đến 24 giờ sau khi kích hoạt. Ví dụ, xem trong Cibelli et al. (1998) Science 280, 1256-1258 và patent Mỹ các số 6,548,741, 7,547,816, 7,989,657, hoặc 6,211,429. Đối với lợn, con cái tiếp nhận có thể được kiểm tra về sự mang thai khoảng 20-21 ngày sau khi chuyển phôi.

Kỹ thuật chọn giống chuẩn có thể được sử dụng để tạo ra động vật mà là đồng hợp tử về gen bất hoạt từ động vật chuyển gen dị hợp tử ban đầu. Tuy nhiên, tính đồng

hợp tử có thể không cần thiết. Lợn chỉnh sửa gen mô tả ở đây có thể được chọn giống với các lợn mong muốn khác.

Một khi động vật chỉnh sửa gen đã được tạo ra, việc bất hoạt axit nucleic ngoại sinh có thể được đánh giá sử dụng các kỹ thuật chuẩn. Việc sàng lọc ban đầu có thể được hoàn thành bằng phương pháp phân tích bằng thẩm tách Southern để xác định xem việc bất hoạt có diễn ra hay không. Đối với việc mô tả kỹ thuật phân tích Southern, xem trong các phần 9.37-9.52 của tài liệu Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, second edition, Cold Spring Harbor Press, Plainview; N.Y. Các kỹ thuật phản ứng chuỗi polymeraza (PCR) cũng có thể được sử dụng trong PCR sàng lọc ban đầu đề cập đến phương thức hoặc kỹ thuật trong đó các axit nucleic đích được khuếch đại. Thông thường, thông tin trình tự từ các đầu của vùng mong muốn hoặc xa hơn được sử dụng để thiết kế các đoạn mồi oligonucleotit mà chúng tương đồng hoặc tương tự về trình tự với các sợi đối của mẫu cần được khuếch đại. PCR có thể được sử dụng để khuếch đại các trình tự đặc trưng từ ADN cũng như ARN, bao gồm các trình tự từ ADN hệ gen toàn bộ hoặc ARN tế bào toàn bộ. Các đoạn mồi điển hình có độ dài 14 đến 40 nucleotit, nhưng có thể nằm trong khoảng từ 10 nucleotit đến hàng trăm nucleotits về độ dài. PCR được mô tả trong, ví dụ PCR Primer: A Laboratory Manual, ed. Dieffenbach and Dveksler, Cold Spring Harbor Laboratory Press, 1995. Các axit nucleic cũng có thể được khuếch đại bằng phản ứng chuỗi ligaza, khuếch đại thay thế sợi, sao chép trình tự tự duy trì, hoặc khuếch đại dựa vào trình tự axit nucleic. Ví dụ, xem trong Lewis (1992) Genetic Engineering News 12,1; Guatelli et al. (1990) Proc. Natl. Acad. Sci. USA 87:1874; and Weiss (1991) Science 254:1292. Ở giai đoạn túi phôi, phôi có thể được xử lý riêng biệt để phân tích bằng PCR, kỹ thuật lai Southern và splinkerette PCR (ví dụ, xem trong Dupuy et al. Proc Natl Acad Sci USA (2002) 99:4495).

Các ARN can thiệp

Nhiều hệ ARN can thiệp (RNAi) đã được biết đến. ARN sợi kép (dsRNA) cảm ứng sự thoái biến đặc hiệu trình tự của các sản phẩm phiên mã gen tương đồng. Phức làm câm cảm ứng bằng ARN (RISC) chuyển hóa dsRNA thành các ARN can thiệp nhỏ 21-23-nucleotit (siRNAs). RISC chứa RNAza sợi kép (dsRNAza, ví dụ, Dicer) và ssRNAza (ví dụ, Argonaut 2 hoặc Ago2). RISC sử dụng sợi đối nghĩa làm dẫn hướng

để tìm đích có thể phân cắt được. Cả siARN và microARN (miRNAs) đề đã được biết. Phương pháp bất hoạt gen ở động vật chỉnh sửa gen bao gồm bước cảm ứng can thiệp ARN đối với gen đích và/hoặc axit nucleic sao cho sự biểu hiện của gen đích và/hoặc axit nucleic được giảm xuống.

Ví dụ, trình tự axit nucleic ngoại sinh có thể cảm ứng can thiệp ARN đối với axit nucleic mã hóa polypeptit. Ví dụ, ARN can thiệp nhỏ sợi kép (siRNA) hoặc ARN kẹp tóc nhỏ (shRNA) tương đồng với ADN đích có thể được sử dụng để làm giảm sự biểu hiện của ADN đó. Các cấu trúc của siRNA có thể được tạo ra như được mô tả, ví dụ, trong Fire et al. (1998) Nature 391:806; Romano and Masino (1992) Mol. Microbiol. 6:3343; Cogoni et al. (1996) EMBO J. 15:3153; Cogoni and Masino (1999) Nature 399:166; Misquitta and Paterson (1999) Proc. Natl. Acad. Sci. USA 96:1451; và Kennerdell and Carthew (1998) Cell 95:1017. Các cấu trúc của shARN có thể được tạo ra như được mô tả bởi McIntyre and Fanning (2006) BMC Biotechnology 6:1. Nói chung, các shARN được phiên mã ở dạng sợi đơn phân tử ARN chứa các vùng bổ trợ, các vùng này có thể được ghép và tạo ra các kẹp tóc ngắn.

Xác suất phát hiện siARN hoặc miARN đơn, single, chức năng riêng biệt định hướng vào gen đặc hiệu là cao. Khả năng dự đoán của trình tự đặc hiệu của siARN, ví dụ vào khoảng 50% nhưng số lượng các ARN can thiệp có thể được tạo ra với độ tin cậy tốt mà ít nhất một trong số đó sẽ có hiệu quả.

Các tế bào *in vitro*, tế bào *in vivo* hoặc động vật chỉnh sửa gen như động vật nuôi mà biểu hiện ARNi định hướng vào gen mã hóa CD163 có thể được sử dụng. Ví dụ, ARNi có thể được chọn từ nhóm bao gồm siARN, shARN, dsARN, RISC và miARN.

Các hệ có thể cảm ứng

Hệ có thể cảm ứng có thể được sử dụng để bất hoạt gen CD163. Các hệ có thể cảm ứng khác nhau đã được biết là cho phép sự điều khiển không gian và thời gian về sự bất hoạt gen. Một số đã được chứng minh là có chức năng *in vivo* ở lợn.

Một ví dụ về hệ có thể cảm ứng đó là hệ tetracycline (tet)-trên đoạn khởi động, hệ này có thể được sử dụng để điều hòa sự phiên mã của axit nucleic. Trong hệ này, phần tử ức chế Tet đột biến (TetR) được dung hợp với miền kích hoạt của protein hoạt hóa phiên mã VP 16 của virut herpes để tạo ra tác nhân hoạt hóa phiên mã kiểm soát bằng tetracycline (tTA), quá trình này được điều hòa bởi tet hoặc doxycycline (dox).

Khi không có mặt kháng sinh, sự phiên mã là tối thiểu, trong khi nếu có mặt tet hoặc dox, sự phiên mãn được cảm ứng. Các hệ có thể cảm ứng thay thế bao gồm các hệ ecdysone hoặc rapamycin. Ecdysone hocmon lột xác côn trùng mà việc sản xuất suất ra hocmon này được kiểm soát bởi heterodime của thụ thể ecdysone và sản phẩm của gen thụ thể nhân của loài không xương sống (ultraspiracle gene - USP). Sự biểu hiện được cảm ứng nhờ xử lý bằng ecdysone hoặc chất tương tự ecdyson như muristeron A. Tác nhân mà được đưa vào động vật để khơi mào hệ có thể cảm ứng được gọi là tác nhân cảm ứng.

Hệ có thể cảm ứng tetracycline và hệ Cre/loxP recombinaza (cơ định hoặc có thể cảm ứng) là trong số các hệ có thể cảm ứng được sử dụng phổ biến hơn. Hệ có thể cảm ứng tetracycline bao gồm chất hoạt chuyển kiểm soát bằng tetracycline (tTA)/tTA đảo (rtTA). Phương pháp để sử dụng các hệ này *in vivo* dẫn đến sự sinh ra hai dòng động vật chỉnh sửa gen. Một dòng động vật biểu hiện tác nhân kích hoạt (tTA, rtTA, hoặc Cre recombinaza) dưới sự kiểm soát của đoạn khởi động lựa chọn. Dòng động vật kia biểu hiện tác nhân tiếp nhận, trong đó sự biểu hiện của gen mong muốn (hoặc gen cần sửa đổi) nằm dưới sự kiểm soát của trình tự đích của các chất hoạt chuyển tTA/rtTA (hoặc được chặn bởi các trình tự loxP). Ghép đôi giao phối hai động vật tạo ra sự kiểm soát biểu hiện gen.

Hệ điều hòa phụ thuộc tetracycline (các hệ tet) dựa trên hai thành phần, tức là, chất hoạt chuyển kiểm soát bằng tetracycline (tTA hoặc rtTA) và đoạn khởi động phụ thuộc tTA/rtTA mà nó kiểm soát sự biểu hiện của cADN xuôi dòng, theo cách phụ thuộc tetracycline. Khi không có mặt tetracycline hoặc các dẫn xuất của nó (như doxycycline), tTA liên kết với các trình tự tetO, cho phép kích hoạt phiên mã của đoạn khởi động phụ thuộc tTA. Tuy nhiên, khi có mặt doxycycline, tTA không thể tương tác với đích của nó và sự phiên mã không xuất hiện. Hệ tet mà hệ này sử dụng tTA được gọi là tet-OFF, vì tetracycline hoặc doxycycline cho phép sự điều hòa phiên mã theo hướng giảm. Việc sử dụng tetracycline hoặc các dẫn xuất của nó cho phép kiểm soát về thời gian của sự biểu hiện gen chuyển *in vivo*. rtTA là một biến thể của tTA mà nó không có chức năng khi không có mặt doxycycline nhưng đòi hỏi sự có mặt của phối tử để hoạt hóa vận chuyển. Do đó, hệ thống tet này được gọi là tet-ON. Các hệ thử nghiệm đã được sử dụng trong *vivo* để biểu hiện có thể cảm ứng của nhiều gen chuyển, mã hóa, ví dụ, các gen thông báo, gen sinh ung thư, hoặc protein liên quan đến dòng thác truyền tín hiệu.

Hệ Cre/lox sử dụng Cre recombinaza, hệ này xúc tác quá trình tái tổ hợp đặc hiệu vị trí bằng sự trao đổi chéo giữa hai trình tự nhận diện Cre từ xa, tức là, các vị trí loxP. Trình tự ADN được đưa vào giữa hai trình tự loxP (được gọi là ADN kẹp giữa) được cắt bởi quá trình tái tổ hợp qua trung gian Cre. Việc kiểm soát sự biểu hiện Cre trong động vật chuyển gen và/hoặc chỉnh sửa gen, sử dụng kiểm soát không gian (bằng đoạn khởi động đặc hiệu mô hoặc đặc hiệu tế bào), hoặc kiểm soát thời gian (bằng hệ có thể cảm ứng), dẫn đến kiểm soát sự cắt ADN giữa hai vị trí loxP. Một ứng dụng đó là để bất hoạt gen có điều kiện (bất hoạt có điều kiện). Phương pháp tiếp cận khác đó là để biểu hiện quá mức protein, trong đó đơn vị mã kết thúc kẹp giữa được xen vào giữa trình tự khởi động và ADN mong muốn. Động vật chỉnh sửa gen không biểu hiện gen chuyển cho tới khi Cre được biểu hiện, dẫn đến cắt đơn vị mã kết thúc kẹp giữa. Hệ này đã được áp dụng trong quá trình sinh ung thư đặc hiệu mô và sự biểu hiện thụ thể kháng nguyên được kiểm soát trong các lympho bào B. Các Cre recombinaza có thể cảm ứng cũng đã được phát triển. Cre recombinaza có thể cảm ứng được kích hoạt duy nhất bằng việc sử dụng phối tử ngoại sinh. Các Cre recombinaza có thể cảm ứng là các protein dung hợp chứa Cre recombinaza nguyên thủy và miền liên kết phối tử đặc hiệu. Hoạt tính chức năng của Cre recombinaza phụ thuộc vào phối tử bên ngoài mà nó có khả năng liên kết với miền đặc hiệu này trong protein dung hợp.

Các tế bào *in vitro*, các tế bào *in vivo* hoặc động vật chỉnh sửa gen như động vật nuôi mà bao gồm gen CD163 dưới sự kiểm soát của một hệ có thể cảm ứng có thể được sử dụng. Sử sửa đổi gen của động vật có thể là ở hệ gen hoặc thể khảm. Hệ có thể cảm ứng có thể, ví dụ, được chọn từ nhóm bao gồm Tet-On, Tet-Off, Cre-lox, và Hif1 alpha.

Các vectơ và axit nucleic

Nhiều axit nucleic có thể được đưa vào trong các tế bào nhằm mục đích bất hoạt, để bất hoạt gen, để thu được sự biểu hiện của gen, hoặc nhằm mục đích khác. Như được sử dụng ở đây, thuật ngữ axit nucleic bao gồm ADN, ARN, và các chất tương tự axit nucleic, và các axit nucleic là sợi kép hoặc sợi đơn (tức là, sợi đơn có nghĩa hoặc đối nghĩa). Các chất tương tự axit nucleic có thể được sửa đổi tại gốc bazơ, gốc đường, hoặc khung phosphat để cải thiện, ví dụ, độ ổn định, sự lai, hoặc độ hòa tan của axit nucleic. Sửa đổi tại gốc bazơ bao gồm deoxyuridin đối với deoxythymidine, và 5-metyl-2'-deoxyxytidin và 5-bromo-2'-doxyxytidin đối với deoxyxytidin. Những sửa đổi của gốc

đường bao gồm sửa đổi tại 2' hydroxyl của đường riboza để tạo ra các đường 2'-O-metyl hoặc 2'-O-alyl. Khung deoxyriboza phosphat có thể được sửa đổi để tạo ra các axit nucleic morpholino, trong đó mỗi một gốc bazơ được liên kết với vòng morpholino 6 cạnh, hoặc các axit nucleic peptit, trong đó khung deoxyphosphat được thay thế bằng khung giả peptit và 4 bazơ được giữ nguyên. Xem trong Summerton and Weller (1997) Antisense Nucleic Acid Drug Dev. 7(3):187; và Hyrup et al. (1996) Bioorgan. Med. Chem. 4:5. Ngoài ra, khung deoxyphosphat có thể được thay thế bằng, ví dụ, khung phosphorothioat hoặc khung phosphorodithioat, khung phosphoroamidit, hoặc khung alkyl phosphotrieste.

Trình tự axit nucleic đích có thể được liên kết theo kiểu hoạt động được với vùng điều hòa như đoạn khởi động. Vùng điều hòa có thể là vùng điều hòa của lợn hoặc có thể từ các loài khác. Như được sử dụng ở đây, liên kết theo kiểu hoạt động được đề cập đến việc định vị vùng điều hòa liên quan đến trình tự axit nucleic theo cách để cho phép hoặc tạo thuận lợi cho sự phiên mã của axit nucleic đích.

Kiểu đoạn khởi động bất kỳ có thể được liên kết theo kiểu hoạt động được với trình tự axit nucleic đích. Ví dụ về các đoạn khởi động bao gồm, không giới hạn, các đoạn khởi động đặc hiệu mô, các đoạn khởi động cơ định, các đoạn khởi động có thể cảm ứng, và các đoạn khởi động đáp ứng hoặc không đáp ứng với tác nhân kích thích cụ thể. Các đoạn khởi động đặc hiệu mô thích hợp có thể dẫn đến sự biểu hiện ưu tiên của sản phẩm phiên mã axit nucleic trong các tế bào beta và bao gồm, ví dụ, đoạn khởi động insulin người. Các đoạn khởi động đặc hiệu mô khác có thể dẫn đến sự biểu hiện ưu tiên, ví dụ, trong tế bào gan hoặc mô tim và có thể lần lượt bao gồm các đoạn khởi động chuỗi nặng albumin hoặc alpha-myosin. Đoạn khởi động mà tạo thuận lợi cho sự biểu hiện của phân tử axit nucleic không có tính đặc hiệu mô hoặc thời gian đáng kể có thể được sử dụng (tức là, đoạn khởi động cơ định). Ví dụ, đoạn khởi động beta-actin như đoạn khởi động gen beta-actin ở gà, đoạn khởi động ubiquitin, đoạn khởi động miniCAGs, đoạn khởi động glyxeraldehyt-3-phosphat dehydrogenaza (GAPDH) hoặc đoạn khởi động 3-phosphoglyxerat kinaza (PGK) có thể được sử dụng, cũng như các đoạn khởi động ở virut như đoạn khởi động thymidin kinaza ở virut herpes (HSV-TK), đoạn khởi động SV40 hoặc đoạn khởi động virut cự bào (CMV). Ví dụ, sự dung hợp đoạn khởi động gen beta actin ở gà và đoạn gen tăng cường CMV có thể được sử dụng

làm đoạn khởi động. Ví dụ, xem trong Xu et al. (2001) Hum. Gene Ther. 12:563; và Kiwaki et al. (1996) Hum. Gene Ther. 7:821.

Các vùng điều hòa bổ sung mà có thể hữu dụng trong các cấu trúc axit nucleic, bao gồm, nhưng không bị giới hạn bởi, các trình tự polyadenyl hóa, các trình tự kiểm soát dịch mã (ví dụ, đoạn nội nhập ribosom bên trong, IRES), đoạn gen tăng cường, các yếu tố có thể cảm ứng hoặc intron. Các vùng điều hòa như vậy có thể không cần thiết, mặc dù chúng có thể gia tăng sự biểu hiện bằng việc tác động lên sự phiên mã, độ ổn định của mARN, hiệu suất dịch mã, hoặc tương tự. Các vùng điều hòa như vậy có thể được bao gồm trong cấu trúc axit nucleic khi cần thiết để thu được sự biểu hiện tối ưu của axit nucleic trong (các) tế bào. Tuy nhiên, sự biểu hiện đầy đủ đôi khi có thể thu được không cần các yếu tố bổ sung.

Cấu trúc axit nucleic có thể được sử dụng mã hóa cho các peptit tín hiệu hoặc các dấu chuẩn có thể lựa chọn. Các peptit tín hiệu có thể được sử dụng sao cho polypeptit mã hóa được hướng vào một vị trí tế bào cụ thể (ví dụ, bề mặt tế bào). Các ví dụ không giới hạn về các dấu chuẩn có thể lựa chọn bao gồm puromycin, ganciclovir, adenosine deaminaza (ADA), aminoglycoside phosphotransferaza (neo, G418, APH), dihydrofolate reductaza (DHFR), hygromycin-B-phosphtransferaza, thymidine kinaza (TK), và xanthin-guanin phosphoribosyltransferaza (XGPRT). Các dấu chuẩn như vậy hữu dụng để lựa chọn các thể biến nạp ổn định trong nuôi cấy. Các dấu chuẩn có thể lựa chọn khác bao gồm các polypeptit huỳnh quang, như protein huỳnh quang xanh hoặc protein huỳnh quang vàng.

Trình tự mã hóa dấu chuẩn có thể lựa chọn như vậy có thể được chặn bởi trình tự nhận diện của recombinaza, ví dụ như Cre hoặc Flp. Ví dụ, gen chỉ thị chọn lọc có thể được chặn bởi các vị trí nhận diện loxP (các vị trí nhận diện 34-bp được nhận diện bởi Cre recombinaza) hoặc các vị trí nhận diện FRT sao cho, gen chỉ thị chọn lọc có thể được cắt từ cấu trúc. Xem trong tài liệu của Orban, et al., Proc. Natl. Acad. Sci. (1992) 89:6861, for a review of Cre/lox technology, và Brand and Dymecki, Dev. Cell (2004) 6:7. Gen nhảy chứa gen chuyển có thể được kích hoạt bởi Cre hoặc Flp bị làm gián đoạn bởi gen chỉ thị có thể lựa chọn cũng có thể được sử dụng để thu được động vật với sự biểu hiện có điều kiện của gen chuyển. Ví dụ, đoạn khởi động gây ra biểu hiện của gen chỉ thị/gen chuyển có thể có mặt mọi nơi hoặc đặc hiệu mô, dẫn đến sự biểu hiện mọi

chỗ hoặc đặc hiệu mô của gen chỉ thị ở động vật F0 (ví dụ, lợn). Sự kích hoạt đặc hiệu mô của gen chuyển có thể được hoàn thành, ví dụ, bằng cách lai chéo lợn mà biểu hiện gen chuyển bị gián đoạn bằng dấu chuẩn mọi nơi với lợn biểu hiện Cre hoặc Flp theo cách đặc hiệu mô hoặc bằng cách lai chéo lợn mà biểu hiện gen chuyển bị gián đoạn bằng dấu chuẩn theo cách đặc hiệu mô với lợn mà nó biểu hiện Cre hoặc Flp recombinaza mọi nơi. Sự biểu hiện có kiểm soát của gen chuyển hoặc sự cắt có kiểm soát của dấu chuẩn cho phép sự biểu hiện của gen chuyển.

Axit nucleic ngoại sinh có thể mã hóa cho polypeptit. Trình tự axit nucleic mã hóa polypeptit có thể bao gồm trình tự thẻ mà nó mã hóa cho "thẻ" được thiết kế để tạo thuận lợi cho sự thao tác sau đó của polypeptit được mã hóa (ví dụ, để tạo thuận lợi cho sự định vị hoặc phát hiện). Các trình tự thẻ có thể được xen vào trong trình tự axit nucleic mã hóa cho polypeptit sao cho, thẻ được mã hóa nằm ở đầu tận cùng carboxyl hoặc amino của polypeptit. Các ví dụ không giới hạn về thẻ được mã hóa bao gồm glutathion S-transferaza (GST) và FLAG^{TM}tag (Kodak, New Haven, Conn.).

Các cấu trúc axit nucleic có thể được metyl hóa sử dụng SssI CpG metylaza (New England Biolabs, Ipswich, Mass.). Nói chung, cấu trúc axit nucleic có thể được ủ với S-adenosylmethionin và SssI CpG-metylaza trong dung dịch đệm ở nhiệt độ 37°C. Sự metyl hóa vượt mức có thể được khẳng định bằng cách ủ cấu trúc với một đơn vị men HinP1I endonucleaza trong thời gian 1 giờ ở nhiệt độ 37°C và phân tích bằng điện di trên gel agaroza.

Các cấu trúc axit nucleic có thể được đưa vào trong tế bào phôi, thai hoặc tế bào động vật trưởng thành thuộc kiểu bất kỳ, bao gồm, ví dụ, các tế bào mầm như noãn bào hoặc trứng, tế bào tiền thân, tế bào trưởng thành hoặc tế bào mầm phôi, tế bào sinh dục mầm, tế bào thận như tế bào PK-15, tế bào đảo tụy, tế bào beta, tế bào gan, hoặc nguyên bào sợi như nguyên bào sợi da, sử dụng nhiều kỹ thuật khác nhau. Các ví dụ không giới hạn về kỹ thuật bao gồm kỹ thuật sử dụng các hệ gen nhảy, các virut tái tổ hợp có thể gây nhiễm tế bào hoặc các liposom hoặc các phương pháp không sử dụng virut khác như phương pháp xung điện, vi tiêm, hoặc kết tủa canxi phosphat, các phương pháp này có khả năng phân phối các axit nucleic vào tế bào.

Trong các hệ gen nhảy, đơn vị phiên mã của cấu trúc axit nucleic, tức là, vùng điều hòa liên kết theo kiểu hoạt động được với trình tự axit nucleic ngoại sinh, được

chặn bởi đoạn lặp đảo ngược của gen nhảy. Một số hệ gen nhảy, bao gồm, ví dụ, Sleeping Beauty (xem trong patent Mỹ số 6,613,752 và công bố đơn Mỹ số 2005/0003542); Frog Prince (Miskey et al. (2003) Nucleic Acids Res. 31:6873); Tol2 (Kawakami (2007) Genome Biology 8(Suppl.1):S7; Minos (Pavlopoulos et al. (2007) Genome Biology 8(Suppl.1):S2); Hsmar1 (Miskey et al. (2007)) Mol Cell Biol. 27:4589); và Passport đã được phát triển để đưa các axit nucleic vào trong tế bào, bao gồm các tế bào chuột, người, và lợn. Gen nhảy Sleeping Beauty là đặc biệt hữu dụng. Men transposaza có thể được phân phối ở dạng protein, được mã hóa trên cùng cấu trúc axit nucleic như axit nucleic ngoại sinh, có thể được đưa vào trên cấu trúc axit nucleic riêng biệt, hoặc được cung cấp ở dạng mARN (ví dụ, mARN có mũ hoặc phiên mã *in vitro*).

Các yếu tố cách ly cũng có thể được bao gồm trong cấu trúc axit nucleic để duy trì sự biểu hiện của axit nucleic ngoại sinh và để ngăn chặn sự phiên mã không mong muốn của các gen chủ. Ví dụ, xem trong công bố đơn Mỹ số 2004/0203158. Điển hình, yếu tố cách ly chặn một phía của đơn vị phiên mã và ở bên trong đoạn lặp đảo ngược của gen nhảy. Các ví dụ không giới hạn về yếu tố cách ly bao gồm các yếu tố cách ly kiểu vùng gắn với chất nền (MAR) và các yếu tố cách ly kiểu diềm. Ví dụ, xem trong tài liệu patent Mỹ số 6,395,549, 5,731,178, 6,100,448, và 5,610,053, và công bố đơn Mỹ số 2004/0203158.

Các axit nucleic có thể được sáp hợp vào trong các vectơ. Vectơ là một thuật ngữ rộng bao gồm đoạn ADN bất kỳ mà được thiết kế để di chuyển từ thể mang vào trong ADN đích. Vectơ có thể được tham chiếu là vectơ biểu hiện, hoặc hệ vectơ, hệ này là một tập hợp các thành phần cần thiết để đưa đoạn xen ADN vào trong hệ gen hoặc trình tự ADN đích khác như episome, plasmit, hoặc thậm chí đoạn virut/ ADN thể thực khuẩn. Các hệ vectơ như vectơ virut (ví dụ, retrovirut, virut kết hợp adeno và các virut thể thực khuẩn hợp nhất), và các vectơ không virut (ví dụ, gen nhảy) sử dụng để phân phối gen ở động vật có hai thành phần cơ bản: 1) vectơ bao gồm ADN (hoặc ARN mà nó được phiên mã ngược thành cADN) và 2) transposaza, recombinaza, hoặc enzym integraza khác mà nó nhận diện cả vectơ và trình tự đích ADN và gắn vectơ vào trong trình tự ADN đích. Vectơ thường chứa một hoặc nhiều catxet biểu hiện bao gồm một hoặc nhiều trình tự kiểm soát biểu hiện, trong đó trình tự kiểm soát biểu hiện là trình tự

ADN mà nó kiểm soát và điều hòa sự phiên mã và/hoặc dịch mã của trình tự ADN hoặc mARN tương ứng khác.

Nhiều kiểu vectơ khác nhau là đã biết. Ví dụ, plasmit và vectơ virut, ví dụ, vectơ retrovirut là đã biết. Các plasmit biểu hiện ở động vật có vú điển hình có nguồn gốc sao chép, đoạn khởi động thích hợp và đoạn gen tăng cường tùy ý, các vị trí liên kết ribosom cần thiết, vị trí polyadenyl hóa, các vị trí cho và nhận ghép nối, các trình tự kết thúc phiên mã, và các trình tự không phiến mã chặn đầu 5'. Ví dụ về vectơ bao gồm: plasmit (chúng cũng có thể là thể mang của kiểu vectơ khác), adenovirut, virut kết hợp adeno (AAV), lentivirut (ví dụ, HIV-1, SIV hoặc FIV sửa đổi), retrovirut (ví dụ, ASV, ALV hoặc MoMLV), và gen nhảy (ví dụ, Sleeping Beauty, các yếu tố P, Tol-2, Frog Prince, piggyBac).

Như được sử dụng ở đây, thuật ngữ axit nucleic đề cập đến cả ARN và ADN, bao gồm, ví dụ, cADN, ADN hệ gen, ADN tổng hợp (ví dụ, tổng hợp theo phương thức hóa học), cũng như các axit nucleic xuất hiện tự nhiên và sửa đổi theo phương thức hóa học, ví dụ, các bazơ tổng hợp hoặc khung thay thế. Phân tử axit nucleic có thể là sợi kép hoặc sợi đơn (tức là, sợi đơn có nghĩa hoặc đối nghĩa).

Động vật chuyển gen, các dòng động vật, tính trạng và sinh sản

Động vật chuyển gen có thể được tạo ra bằng phương pháp tách dòng và các phương pháp khác mô tả ở đây. Các động vật chuyển gen có thể là đồng hợp tử về sửa đổi gen, như trong trường hợp mà hợp tử hoặc tế bào nguyên thủy trải qua sửa đổi đồng hợp tử. Tương tự, động vật chuyển gen cũng có thể được tạo ra là dị hợp tử. Trong trường hợp động vật chứa ít nhất một trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163, động vật chuyển gen tốt hơn là dị hợp tử. Động vật chuyển gen có thể được sửa đổi về hệ gen, có nghĩa là tất cả các tế bào trong hệ gen của chúng đã trải qua sửa đổi. Động vật chuyển gen có thể là khảm để sửa đổi, như có thể xảy ra khi các vectơ được đưa vào trong một trong số nhiều tế bào trong phôi, điển hình là ở giai đoạn túi phôi. Thế hệ con của động vật khảm có thể được thử nghiệm để nhận diện thế hệ con mà được sửa đổi về hệ gen. Một dòng động vật được thiết lập khi nhóm động vật đã được tạo ra mà chúng có thể được sinh sản tự nhiên hoặc bằng các kỹ thuật sinh sản trợ giúp, với thê hệ con bất đồng nhất hoặc đồng hợp tử biểu hiện sự sửa đổi thích hợp.

Đối với vật nuôi, nhiều alen được biết là có liên quan với các tính trạng khác nhau như tính trạng sinh sản, tính trạng kiểu, tính trạng về khả năng chế biến, và các tính trạng chức năng khác. Người chăn nuôi được tập quen để giám sát và định lượng các tính trạng này, ví dụ, Visscher et al., Livestock Production Science, 40 (1994) 123-137, patent Mỹ số 7,709,206, US 2001/0016315, US 2011/0023140, và US 2005/0153317. Dòng động vật có thể bao gồm tính trạng được chọn từ tính trạng trong nhóm bao gồm tính trạng sinh sản, tính trạng kiểu, tính trạng về khả năng chế biến, tính trạng thụ tinh, tình trạng sinh con và tính trạng kháng bệnh. Các tính trạng khác bao gồm sự biểu hiện của sản phẩm gen tái tổ hợp.

Động vật với tính trạng hoặc các tính trạng mong muốn có thể được sửa đổi để ngăn chặn sự thành thục về sinh dục của chúng. Vì động vật không có khả năng sinh sản cho tới khi thành thục, nên có thể điều chỉnh khả năng thành thục sinh dục như là phương tiện kiểm soát sự phân tán động vật. Động vật mà đã được chọn giống hoặc chỉnh sửa để có một hoặc nhiều tính trạng như vậy có thể được cung cấp cho đối tượng tiếp nhận với nguy cơ giảm thiểu mà đối tượng tiếp nhận sẽ nhân giống động vật và có được giá trị của các tính trạng của chúng. Ví dụ, hệ gen của động vật có thể sửa đổi về gen, trong đó sự sửa đổi bao gồm bất hoạt gen thành thục về sinh dục, trong đó gen thành thục về sinh dục ở động vật kiểu hoang dại biểu hiện một yếu tố lựa chọn về sự thành thục về sinh dục. Động vật có thể được xử lý bằng cách đưa hợp chất để khắc phục sự thiếu hụt gây ra bởi sự mất biểu hiện của gen để cảm ứng sự thành thục về sinh dục ở động vật.

Việc nhân giống động vật mà đòi hỏi sử dụng hợp chất để gây ra khả năng thành thục sinh dục có thể được hoàn thành một các thuận lợi tại cơ sở xử lý. Cơ sở xử lý có thể thực hiện các phương thức chuẩn hóa đối với nguyên liệu gốc được kiểm soát tốt để tạo ra một cách hiệu quả các động vật thích hợp. Thế hệ con động vật có thể được phân bố tới nhiều nơi để nuôi dưỡng. Các trang trại và chủ trang trại (thuật ngữ bao gồm trại chăn nuôi gia súc và chủ trại gia súc) do đó có thể yêu cầu số lượng thế hệ con mong muốn với khoảng tuổi và/hoặc trọng lượng và/hoặc tính trạng cụ thể và yêu cầu chuyển đến tại thời điểm và/hoặc địa điểm mong muốn. Đối tượng tiếp nhận, ví dụ, các chủ trang trại sau đó có thể nuôi dưỡng động vật và phân phối chúng ra thị trường khi họ muốn.

Động vật nuôi biến đổi gen có gen thành thục về sinh dục bất hoạt có thể được phân phối (ví dụ, tới một hoặc nhiều địa điểm, tới nhiều trang trại). Động vật có thể có độ tuổi nằm trong khoảng từ 1 ngày tới khoảng 180 ngày. Động vật có thể có một hoặc nhiều tính trạng (ví dụ, một động vật biểu hiện tính trạng mong muốn hoặc tính trạng giá trị cao hoặc tính trạng mới hoặc tính trạng tái tổ hợp).

Các phương pháp nhân giống và phương pháp tăng cường tính kháng ở động vật với nhiễm trùng và các quần thể động vật

Được đề cập ở đây là phương pháp nhân giống để tạo ra động vật hoặc các dòng mà có tính nhạy giảm đối với nhiễm trùng bởi tác nhân gây bệnh. Phương pháp bao gồm bước sửa đổi noãn bào hoặc tế bào tinh trùng về mặt di truyền để đưa trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163 vào trong ít nhất một trong số noãn bào và tế bào tinh trùng, và làm thụ tinh noãn bào bằng tế bào tinh trùng để tạo ra trứng đã thụ tinh chứa trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163. Theo cách khác, phương pháp bao gồm bước sửa đổi về mặt di truyền trứng đã thụ tinh để đưa trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163 vào trong trứng đã thụ tinh. Phương pháp bao gồm thêm bước chuyển trứng đã thụ tinh vào trong động vật cái lựa chọn, trong đó thời gian mang thai và sinh đúng kỳ hạn tạo ra động vật thế hệ con; sàng lọc động vật thế hệ con này về tính nhạy đối với tác nhân gây bệnh; và lựa chọn động vật thế hệ con mà có tính nhạy giảm đối với tác nhân gây bệnh khi so sánh với các động vật mà không chứa trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163.

Phương pháp nhân giống khác để tạo ra động vật hoặc các dòng mà có tính nhạy giảm đối với nhiễm trùng bởi tác nhân gây bệnh được đề cập. Phương pháp bao gồm bước sửa đổi noãn bào hoặc tế bào tinh trùng về mặt di truyền để đưa trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163 vào trong ít nhất một trong số noãn bào và tế bào tinh trùng, và làm thụ tinh noãn bào bằng tế bào tinh trùng để tạo ra trứng đã thụ tinh chứa trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163. Theo cách khác, phương pháp bao gồm bước sửa đổi về mặt di truyền trứng đã thụ tinh để đưa trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163 vào trong trứng đã thụ tinh. Phương pháp bao gồm thêm bước chuyển trứng đã thụ tinh vào trong động vật cái lựa chọn, trong đó thời gian mang thai và sinh đúng kỳ hạn tạo ra động vật thế hệ

con; sàng lọc động vật thế hệ con này về tính nhạy đối với tác nhân gây bệnh; và lựa chọn động vật thế hệ con mà có tính nhạy giảm đối với tác nhân gây bệnh khi so sánh với các động vật mà không chứa trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163. Trình tự nhiễm sắc thể sửa đổi dẫn đến việc tạo ra protein CD163 cơ bản không có chức năng bởi động vật thế hệ con.

Phương pháp nhân giống khác nữa để tạo ra động vật hoặc các dòng mà có tính nhạy giảm đối với nhiễm trùng bởi tác nhân gây bệnh được đề cập. Phương pháp bao gồm bước sửa đổi noãn bào hoặc tế bào tinh trùng về mặt di truyền để đưa trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163 vào trong ít nhất một trong số noãn bào và tế bào tinh trùng, và làm thụ tinh noãn bào bằng tế bào tinh trùng để tạo ra trứng đã thụ tinh chứa trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163. Theo cách khác, phương pháp bao gồm bước sửa đổi về mặt di truyền trứng đã thụ tinh để đưa trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163 vào trong trứng đã thụ tinh. Phương pháp bao gồm thêm bước chuyển trứng đã thụ tinh vào trong động vật cái lựa chọn, trong đó thời gian mang thai và sinh đúng kỳ hạn tạo ra động vật thế hệ con; sàng lọc động vật thế hệ con này về tính nhạy đối với tác nhân gây bệnh; và lựa chọn động vật thế hệ con mà có tính nhạy giảm đối với tác nhân gây bệnh khi so sánh với các động vật mà không chứa trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163. Trình tự nhiễm sắc thể sửa đổi bao gồm xóa trong khung ở gen mã hóa cho protein CD163.

Tác nhân gây bệnh tốt hơn bao gồm virut, ví dụ, PRRSV.

Ví dụ, sự sửa đổi có thể làm giảm tính nhạy với virut PRRSV Type 1, PRRSV Type 2, hoặc với cả virut PRRSV Type 1 và Type 2.

Sự sửa đổi có thể làm giảm tính nhạy với thể phân lập PRRSV được chọn từ nhóm bao gồm NVSL 97-7895, KS06-72109, P129, VR2332, CO90, AZ25, MLV-ResPRRS, KS62-06274, KS483 (SD23983), CO84, SD13-15, Lelystad, 03-1059, 03-1060, SD01-08, 4353PZ, và tổ hợp của chúng.

Động vật có thể là phôi, con non hoặc đối tượng trưởng thành.

Động vật có thể bao gồm động vật nuôi trong gia đình. Động vật nuôi trong gia đình có thể bao gồm động vật nuôi, ví dụ lợn, bò (ví dụ, bò thịt hoặc bò sữa), cừu, dê, ngựa (ví dụ, ngựa hoặc lừa), trâu, lạc đà, hoặc động vật thuộc loài chim (ví dụ, gà, gà

tây, vịt, ngỗng, gà Nhật, hoặc bồ câu non). Động vật nuôi tốt hơn là bò hoặc lợn, và tốt nhất là lợn.

Bước sửa đổi về gen đối với noãn bào, tế bào tinh, hoặc trứng đã thụ tinh có thể bao gồm việc chỉnh sửa gen của noãn bào, tế bào tinh, hoặc trứng đã thụ tinh. Chỉnh sửa gen có thể bao gồm bước sử dụng endonucleaza tìm nhà. Endonucleaza tìm nhà có thể là endonucleaza xuất hiện tự nhiên, nhưng tốt hơn là endonucleaza tìm nhà không xuất hiện một cách tự nhiên, được thiết kế hợp lý mà có trình tự nhận diện ADN đã được thiết kế sao cho endonucleaza hướng đích vào trình tự nhiễm sắc thể ở gen mã hóa cho protein CD163. Do đó, endonucleaza tìm nhà có thể là endonucleaza tìm nhà được thiết kế. Endonucleaza tìm nhà có thể bao gồm, ví dụ, hệ trình tự sắp xếp kiểu lặp đi lặp lại xen khoảng trống (CRISPR) /Cas9, men nucleaza tác động giống chất kích hoạt phiên mã (TALEN), nucleaza ngón tay kẽm (ZFN), protein dung hợp recombinaza, meganucleaza, hoặc tổ hợp của chúng). Chỉnh sửa gen tốt hơn bao gồm việc sử dụng hệ CRISPR/Cas9.

Noãn bào, tế bào tinh, hoặc trứng đã thụ tinh có thể là dị hợp tử về trình tự nhiễm sắc thể sửa đổi. Theo cách khác, noãn bào, tế bào tinh, hoặc trứng đã thụ tinh có thể là đồng hợp tử về trình tự nhiễm sắc thể sửa đổi.

Trình tự nhiễm sắc thể sửa đổi có thể chứa đoạn xen ở gen mã hóa cho protein CD163, đoạn xóa ở gen mã hóa cho protein CD163, hoặc tổ hợp của chúng. Ví dụ, trình tự nhiễm sắc thể sửa đổi bao gồm đoạn xóa ở gen mã hóa cho protein CD163 (ví dụ, xóa trong khung). Theo cách khác, trình tự nhiễm sắc thể sửa đổi có thể chứa đoạn xen ở gen mã hóa cho protein CD163.

Đoạn xen hoặc đoạn xóa có thể dẫn đến sự tạo ra hoặc hoạt tính của protein CD163 được giảm xuống, khi so sánh với sự tạo ra hoặc hoạt tính của protein CD163 ở động vật mà không có đoạn xen hoặc đoạn xóa.

Đoạn xen hoặc đoạn xóa có thể dẫn đến sự tạo ra protein CD163 cơ bản không có chức năng bởi động vật. Với thuật ngữ "protein CD163 cơ bản không có chức năng," có nghĩa rằng, mức protein CD163 ở động vật, thế hệ con, hoặc tế bào là không thể phát hiện được hoặc nếu phát hiện được, thì ít nhất thấp hơn khoảng 90% so với mức quan sát được ở động vật, thế hệ con, hoặc tế bào mà không chứa đoạn xen hoặc đoạn xóa.

Trường hợp động vật là lợn, trình tự nhiễm sắc thể sửa đổi có thể bao gồm sửa đổi trong exon 7 của gen mã hóa cho protein CD163, exon 8 của gen mã hóa cho protein CD163, một intron mà nó liên tiếp với exon 7 hoặc exon 8 của gen mã hóa cho protein CD163, hoặc tổ hợp của chúng. Trình tự nhiễm sắc thể sửa đổi một cách thích hợp bao gồm sửa đổi trong exon 7 của gen mã hóa cho protein CD163.

Sự sửa đổi trong exon 7 của gen mã hóa cho protein CD163 có thể bao gồm đoạn xóa (ví dụ, xóa trong khung trong exon 7). Theo cách khác, sự sửa đổi trong exon 7 của gen mã hóa cho protein CD163 có thể bao gồm đoạn xen.

Trương hợp động vật là lợn, trình tự nhiễm sắc thể sửa đổi có thể bao gồm: (a) SEQ ID NO: 118; hoặc (b) sửa đổi được chọn từ nhóm bao gồm: đoạn xóa 11 cặp bazơ từ nucleotit 3137 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xen 2 cặp bazơ giữa các nucleotit 3149 và 3150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, với đoạn xóa 377 cặp bazơ từ nucleotit 2573 đến nucleotit 2949 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trên cùng alen; đoạn xóa 124 cặp bazơ từ nucleotit 3024 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 123 cặp bazơ từ nucleotit 3024 đến nucleotit 3146 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xen 1 cặp bazơ giữa các nucleotit 3147 và 3148 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 130 cặp bazơ từ nucleotit 3030 đến nucleotit 3159 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 132 cặp bazơ từ nucleotit 3030 đến nucleotit 3161 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1506 cặp bazơ từ nucleotit 1525 đến nucleotit 3030 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xen 7 cặp bazơ giữa nucleotit 3148 và nucleotit 3149 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1280 cặp bazơ từ nucleotit 2818 đến nucleotit 4097 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1373 cặp bazơ từ nucleotit 2724 đến nucleotit 4096 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1467 cặp bazơ từ nucleotit 2431 đến nucleotit 3897 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2417 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 12 cặp bazơ bắt đầu tại nucleotit 488, và trong đó có thêm đoạn xóa 129 cặp bazơ trong exon 7 từ nucleotit 3044 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 28 cặp bazơ từ nucleotit 3145 đến nucleotit 3172 khi so sánh với trình

tự tham chiếu SEQ ID NO: 47; đoạn xóa 1387 cặp bazơ từ nucleotit 3145 đến nucleotit 4531 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1382 cặp bazơ từ nucleotit 3113 đến nucleotit 4494 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 11 cặp bazơ bắt đầu tại nucleotit 3113; đoạn xóa 1720 cặp bazơ từ nucleotit 2440 đến nucleotit 4160 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 452 cặp bazơ từ nucleotit 3015 đến nucleotit 3466 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; hoặc tổ hợp của chúng.

Khi lợn bao gồm đoạn xen 2 cặp bazơ giữa các nucleotit 3149 và 3150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, đoạn xen 2 cặp bazơ có thể bao gồm đoạn xen của dinucleotit AG.

Khi lợn bao gồm đoạn xen 1 bazơ giữa các nucleotit 3147 và 3148 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, đoạn xen 1 bazơ có thể bao gồm đoạn xen của gốc adenine duy nhất.

Khi lợn bao gồm đoạn xen 7 cặp bazơ giữa nucleotit 3148 và nucleotit 3149 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, đoạn xen 7 cặp bazơ có thể bao gồm trình tự TACTACT (SEQ ID NO: 115).

Khi lợn bao gồm đoạn xóa 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2417 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 12 cặp bazơ bắt đầu tại nucleotit 488, và trong đó có thêm đoạn xóa 129 cặp bazơ trong exon 7 từ nucleotit 3044 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, đoạn xen 12 cặp bazơ có thể bao gồm trình tự TGTGGAGAATTC (SEQ ID NO: 116).

Khi lợn bao gồm đoạn xóa 1382 cặp bazơ từ nucleotit 3113 đến nucleotit 4494 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 11 cặp bazơ bắt đầu tại nucleotit 3113, đoạn xen 11 cặp bazơ có thể bao gồm trình tự AGCCAGCGTGC (SEQ ID NO: 117).

Trường hợp trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163 bao gồm đoạn xóa, đoạn xóa tốt hơn bao gồm xóa trong khung. Do đó, trường hợp động vật là lợn, đoạn xen hoặc đoạn xóa ở gen mã hóa cho protein CD163 có thể bao gồm một đoạn xóa trong khung trong exon 7 được chọn từ nhóm bao gồm đoạn xóa 1506 cặp bazơ từ nucleotit 1525 đến nucleotit 3030 khi so sánh với trình tự tham chiếu SEQ ID

NO: 47; đoạn xóa 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2417 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 12 cặp bazơ bắt đầu tại nucleotit 488, và trong đó có thêm đoạn xóa 129 cặp bazơ trong exon 7 từ nucleotit 3044 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1373 cặp bazơ từ nucleotit 2724 đến nucleotit 4096 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 123 cặp bazơ từ nucleotit 3024 đến nucleotit 3146 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1467 cặp bazơ từ nucleotit 2431 đến nucleotit 3897 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1387 cặp bazơ từ nucleotit 3145 đến nucleotit 4531 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 1382 cặp bazơ từ nucleotit 3113 đến nucleotit 4494 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 11 cặp bazơ bắt đầu tại nucleotit 3113; đoạn xóa 1720 cặp bazơ từ nucleotit 2440 đến nucleotit 4160 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; và tổ hợp của chúng.

Khi động vật là lợn, đoạn xen hoặc đoạn xóa có thể được chọn từ nhóm bao gồm: đoạn xen 2 cặp bazơ giữa các nucleotit 3149 và 3150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, với đoạn xóa 377 cặp bazơ từ nucleotit 2573 đến nucleotit 2949 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trên cùng alen; đoạn xóa 28 cặp bazơ từ nucleotit 3145 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; đoạn xóa 452 cặp bazơ từ nucleotit 3015 đến nucleotit 3466 khi so sánh với trình tự tham chiếu SEQ ID NO: 47; và tổ hợp của chúng.

Ví dụ, trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xen 2 cặp bazơ giữa các nucleotit 3149 và 3150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, với đoạn xóa 377 cặp bazơ từ nucleotit 2573 đến nucleotit 2949 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trên cùng alen.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xóa 28 cặp bazơ từ nucleotit 3145 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xóa 452 cặp bazơ từ nucleotit 3015 đến nucleotit 3466 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm tổ hợp bất kỳ của các trình tự nhiễm sắc thể sửa đổi bất kỳ mô tả ở đây.

Ví dụ, trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xen 7 cặp bazơ giữa nucleotit 3148 và nucleotit 3149 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn xóa 11 cặp bazơ từ nucleotit 3137 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xen 7 cặp bazơ giữa nucleotit 3148 và nucleotit 3149 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn xóa 1382 cặp bazơ từ nucleotit 3113 đến nucleotit 4494 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 11 cặp bazơ bắt đầu tại nucleotit 3113 trong alen khác của gen mã hóa cho protein CD163.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm SEQ ID NO: 118 trong một alen của gen mã hóa cho protein CD163; và đoạn xóa 11 cặp bazơ từ nucleotit 3137 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm SEQ ID NO: 118 trong một alen của gen mã hóa cho protein CD163; và đoạn xen 2 cặp bazơ giữa các nucleotit 3149 và 3150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, với đoạn xóa 377 cặp bazơ từ nucleotit 2573 đến nucleotit 2949 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xóa 1280 cặp bazơ từ nucleotit 2818 đến nucleotit 4097 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn xóa 11 cặp bazơ từ nucleotit 3137 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xóa 1280 cặp bazơ từ nucleotit 2818 đến nucleotit 4097 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn xen 2 cặp bazơ giữa các nucleotit 3149 và 3150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, với đoạn xóa 377 cặp bazơ từ nucleotit 2573 đến nucleotit 2949 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xóa 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2417 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 12 cặp bazơ bắt đầu tại nucleotit 488, và trong đó có thêm đoạn xóa 129 cặp bazơ trong exon 7 từ nucleotit 3044 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn xen 2 cặp bazơ giữa các nucleotit 3149 và 3150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, với đoạn xóa 377 cặp bazơ từ nucleotit 2573 đến nucleotit 2949 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm SEQ ID NO: 118 trong một alen của gen mã hóa cho protein CD163; và đoạn xóa 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2417 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 12 cặp bazơ bắt đầu tại nucleotit 488, và trong đó có thêm đoạn xóa 129 cặp bazơ trong exon 7 từ nucleotit 3044 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xóa 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2417 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 12 cặp bazơ bắt đầu tại nucleotit 488, và trong đó có thêm đoạn xóa 129 cặp bazơ trong exon 7 từ nucleotit 3044 đến nucleotit 3172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn xóa 11 cặp bazơ từ nucleotit 3137 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xóa 1467 cặp bazơ từ nucleotit 2431 đến nucleotit 3897 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn xen 2 cặp bazơ giữa các nucleotit 3149 và 3150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, với đoạn xóa 377 cặp bazơ từ nucleotit 2573 đến nucleotit 2949 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xóa 1467 cặp bazơ từ nucleotit 2431 đến nucleotit 3897 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn xóa 11 cặp bazơ từ nucleotit 3137 đến nucleotit 3147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

Trình tự nhiễm sắc thể sửa đổi bao gồm chứa bất kỳ đoạn xen hoặc đoạn xóa nào mô tả ở trên có thể bao gồm trình tự nhiễm sắc thể có mức độ tương đồng trình tự cao so với SEQ ID NO: 47 ngoài đoạn xen hoặc đoạn xóa. Do đó, ví dụ, noãn bào, tế bào tinh, hoặc trứng đã thụ tinh có thể bao gồm trình tự nhiễm sắc thể có ít nhất 80%, ít nhất 85%, ít nhất 90%, ít nhất 95%, ít nhất 98%, ít nhất 99%, ít nhất 99,9%, hoặc 100% mức tương đồng trình tự so với SEQ ID NO: 47 ở các vùng của trình tự nhiễm sắc thể ngoài đoạn xen hoặc đoạn xóa.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm trình tự nhiễm sắc thể bao gồm SEQ ID NO: 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 118, hoặc 119. Như được mô tả thêm trong phần ví dụ dưới đây, các SEQ ID NO. 98–114, và 119 đề cập đến các trình tự nucleotit của vùng tương ứng với vùng CD163 lợn kiểu hoang dại được nêu trong SEQ ID NO:47, và bao gồm đoạn xen hoặc đoạn xóa trong trình tự nhiễm sắc CD163 lợn được mô tả ở đây. SEQ ID NO: 118 đề cập trình tự của vùng tương ứng với vùng CD163 lợn kiểu hoang dại được nêu bởi SEQ ID NO: 47, trong đó exon 7 đã được thay thế bằng một exon tổng hợp mã hóa vùng tương đồng SRCR 8 của protein giống CD163 người 1 (hCD163L1).

Ví dụ, trình tự nhiễm sắc thể sửa đổi có thể bao gồm trình tự nhiễm sắc thể bao gồm SEQ ID NO: 98, 101, 105, 109, 110, 112, 113, hoặc 114. SEQ ID NO: 98, 101, 105, 109, 110, 112, 113, hoặc 114 đề cập đến các trình tự nucleotit của các đoạn xóa trong khung trong exon 7 của trình tự nhiễm sắc CD163 lợn.

Đối với ví dụ khác, trình tự nhiễm sắc thể sửa đổi có thể bao gồm trình tự nhiễm sắc thể bao gồm SEQ ID NO: 103, 111, hoặc 119.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xóa 11 cặp bazơ trong một alen của gen mã hóa cho protein CD163 và đoạn xen 2 cặp bazơ với đoạn xóa 377 cặp bazơ trong alen khác của gen mã hóa cho protein CD163.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xóa 124 cặp bazơ trong một alen của gen mã hóa cho protein CD163 và đoạn xóa 123 cặp bazơ trong alen khác của gen mã hóa cho protein CD163.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xen 1 bazơ.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xóa 130 cặp bazơ trong một alen của gen mã hóa cho protein CD163 và đoạn xóa 132 cặp bazơ trong alen khác của gen mã hóa cho protein CD163.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xóa 1506 cặp bazơ.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xen 7 cặp bazơ.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xóa 1280 cặp bazơ trong một alen của gen mã hóa cho protein CD163 và đoạn xóa 1373 cặp bazơ trong alen khác của gen mã hóa cho protein CD163.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xóa 1467 cặp bazơ.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xóa intron 6 gồm 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2417, với đoạn xen 12 cặp bazơ tại nucleotit 4488 và đoạn xóa 129 cặp bazơ bổ sung trong exon 7.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xóa 28 cặp bazơ trong một alen của gen mã hóa cho protein CD163 và đoạn xóa 1387 cặp bazơ trong alen khác của gen mã hóa cho protein CD163.

Trình tự nhiễm sắc thể sửa đổi có thể bao gồm đoạn xóa 1382 cặp bazơ với đoạn xen 11 cặp bazơ trong một alen của gen mã hóa cho protein CD163 và đoạn xóa 1720 cặp bazơ trong alen khác của gen mã hóa cho protein CD163.

Theo phương pháp bất kỳ trong số các phương pháp chọn giống, động vật lựa chọn có thể được sử dụng làm động vật chuyển gen.

Theo phương pháp bất kỳ trong số các phương pháp chọn giống, sự thụ tinh có thể bao gồm thụ tinh nhân tạo.

Quần thể động vật được tạo ra bằng phương pháp bất kỳ trong số các phương pháp chọn giống cũng được đề cập. Quần thể động vật tốt hơn kháng lại nhiễm trùng bởi tác nhân gây bệnh, ví dụ virut như PRRSV. Ví dụ, quần thể có thể kháng lại nhiễm

trùng bởi virut PRRSV Type 1, PRRSV Type 2, hoặc với cả virut PRRSV Type 1 và Type 2. Quần thể có thể kháng lại nhiễm trùng bởi thể phân lập PRRSV được chọn từ nhóm bao gồm NVSL 97-7895, KS06-72109, P129, VR2332, CO90, AZ25, MLV-ResPRRS, KS62-06274, KS483 (SD23983), CO84, SD13-15, Lelystad, 03-1059, 03-1060, SD01-08, 4353PZ, và tổ hợp của chúng.

Phương pháp làm tăng tính kháng của động vật nuôi đối với nhiễm trùng bởi tác nhân gây bệnh cũng được đề cập. Phương pháp bao gồm bước sửa chữa về mặt di truyền ít nhất một trình tự nhiễm sắc thể từ gen mã hóa cho protein CD163 sao cho sự tạo ra hoặc hoạt tính của protein CD163 giảm xuống khi so sánh với sự tạo ra hoặc hoạt tính của protein CD63 ở động vật nuôi mà không chứa trình tự nhiễm sắc thể đã sửa chữa ở gen mã hóa cho protein CD163. Tác nhân gây bệnh tốt hơn bao gồm virut (ví dụ, PRRSV).

Phương pháp khác để gia tăng tính kháng của động vật nuôi đối với nhiễm trùng bởi tác nhân gây bệnh được đề cập. Phương pháp bao gồm bước sửa chữa về mặt di truyền ít nhất một trình tự nhiễm sắc thể từ gen mã hóa cho protein CD163 sao cho động vật nuôi tạo ra protein CD163 cơ bản không có chức năng.

Phương pháp khác nữa để gia tăng tính kháng của động vật nuôi đối với nhiễm trùng bởi tác nhân gây bệnh được đề cập. Phương pháp bao gồm bước sửa chữa về mặt di truyền ít nhất một trình tự nhiễm sắc thể từ gen mã hóa cho protein CD163 để đưa vào đoạn xóa trong khung, trong đó sự tạo ra hoặc hoạt tính của protein CD163 được giảm xuống ở động vật nuôi, khi so sánh với sự tạo ra hoặc hoạt tính của protein CD63 ở động vật nuôi mà không chứa trình tự nhiễm sắc thể đã sửa chữa ở gen mã hóa cho protein CD163. Đoạn xóa trong khung có thể là, ví dụ, đoạn xóa trong khung bất kỳ mô tả ở đây.

Các axit nucleic

Các axit nucleic được đề cập. Phân tử axit nucleic có thể bao gồm trình tự nucleotit được chọn từ nhóm bao gồm: (a) trình tự nucleotit bao gồm SEQ ID NO: 47; (b) trình tự nucleotit có độ tương đồng trình tự ít nhất 80% với trình tự của SEQ ID NO: 47, trong đó trình tự nucleotit này chứa ít nhất một đoạn thế, đoạn xen hoặc đoạn xóa liên quan đến SEQ ID NO: 47; và (c) trình tự cADN của (a) hoặc (b).

Theo cách khác, axit nucleic có thể bao gồm (a) trình tự nucleotit có độ tương đồng trình tự ít nhất 87,5% với trình tự nêu trong SEQ ID NO: 47, trong đó trình tự nucleotit này chứa ít nhất một đoạn thế, đoạn xen hoặc đoạn xóa liên quan đến SEQ ID NO: 47; và (b) trình tự cADN của (a).

Phân tử axit nucleic bất kỳ mô tả ở đây có thể là phân tử axit nucleic phân lập.

Ví dụ, axit nucleic phân lập có thể bao gồm trình tự nucleotit bao gồm SEQ ID NO: 47.

Theo cách khác, axit nucleic có thể bao gồm trình tự nucleotit có độ tương đồng trình tự ít nhất 80% với trình tự của SEQ ID NO: 47, trong đó trình tự nucleotit này chứa ít nhất một đoạn thế, đoạn xen hoặc đoạn xóa liên quan đến SEQ ID NO: 47. Axit nucleic có thể bao gồm trình tự nucleotit có ít nhất 85%, ít nhất 90%, ít nhất 95%, ít nhất 98%, ít nhất 99%, hoặc ít nhất 99,9%, độ tương đồng trình tự với trình tự SEQ ID NO: 47, trong đó trình tự nucleotit này chứa ít nhất một đoạn thế, đoạn xen hoặc đoạn xóa liên quan đến SEQ ID NO: 47.

Phân tử axit nucleic tốt hơn có ít nhất 87,5% độ tương đồng trình tự với trình tự SEQ ID NO: 47, trong đó trình tự nucleotit chứa ít nhất một đoạn thế, đoạn xen hoặc đoạn xóa liên quan đến SEQ ID NO: 47.

Đoạn thế, đoạn xen hoặc đoạn xóa tốt hơn làm giảm hoặc loại trừ sự tạo ra hoặc hoạt tính của protein CD163, khi so sánh với axit nucleic mà không chứa đoạn thế, đoạn xen hoặc đoạn xóa.

Axit nucleic có thể bao gồm SEQ ID NO: 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 118, hoặc 119.

Ví dụ, axit nucleic có thể bao gồm SEQ ID NO: 98, 101, 105, 109, 110, 112, 113, hoặc 114.

Ví dụ, axit nucleic có thể bao gồm SEQ ID NO: 103,111, hoặc 119.

Axit nucleic có thể bao gồm cADN.

Các axit nucleic khác được đề cập. Axit nucleic có thể bao gồm SEQ ID NO: 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 118, hoặc 119. ví dụ, axit nucleic có thể bao gồm SEQ ID NO: 98, 101, 105, 109, 110, 112, 113,

hoặc 114. Đối với ví dụ khác, axit nucleic có thể bao gồm SEQ ID NO: 103, 111, hoặc 119.

Sau khi đã mô tả chi tiết sáng chế, rõ ràng là các biến đổi và sửa đổi là có thể có mà không đi trệch khỏi phạm vi bảo hộ của sáng chế được xác định trong các điểm yêu cầu bảo hộ kèm theo.

Ví dụ thực hiện sáng chế

Các ví dụ không giới hạn dưới đây được đề xuất để minh họa thêm sáng chế.

Ví dụ 1: Sử dụng hệ CRISPR/Cas9 để tạo ra lợn được xử lý di truyền từ tế bào trứng và phôi thu được in vitro

Các báo báo gần đây mô tả endonucleaza tìm nhà, như các nucleaza cánh tay kẽm (Zinc- Finger Nucleazas -ZFNs), men nucleaza tác động giống chất kích hoạt phiên mã (TALENs), và các thành phần trong nhóm các đoạn ngắn đối xứng lặp lại thường xuyên (Clustered Regularly Interspaced Short Palindromic Repeat -CRISPR)/hệ (Cas9) được liên kết với CRISPR gợi ý rằng kỹ thuật di truyền (GE) ở lợn hiện tại có thể có hiệu quả hơn. Endonucleaza tìm nhà được hướng đích có thể gây ra sự đứt gãy sợi kép (DSBs) ở các vùng cụ thể trong hệ gen và gây ra các đột biến ngẫu nhiên thông qua việc kết nối đầu cuối không có tính tương đồng (Nonhomologous End Joining - NHEJ) hoặc sự kích thích tái tổ hợp tương đồng (HR) nếu ADN cho được tạo ra. Sự biến đổi hướng đích của hệ gen thông qua HR có thể thu được với endonucleaza tìm nhà nếu ADN cho được tạo cùng với nucleaza hướng đích. Sau khi đưa vào các biến đổi cụ thể trong các tế bào sinh dưỡng, các tế bào này được sử dụng để tạo ra lợn GE cho các mục đích khác nhau thông qua SCNT. Do đó, endonucleaza tìm nhà là công cụ hữu dụng để tạo ra lợn GE. Trong số các endonucleaza tìm nhà khác nhau, hệ CRISPR/Cas9, được thích nghi cho các sinh vật tiền nhân nơi nó được sử dụng làm cơ chế phản vệ, được thấy là biện pháp hữu hiệu. Trong tự nhiên, hệ Cas9 yêu cầu ba thành phần, ARN (~20 bazơ) chứa vùng bổ trợ cho trình tự đích (ARN kìm hãm cis [crARN]), ARN chứa vùng bổ trợ cho crARN (crARN hoạt hóa trans[tracrARN]), và Cas9, thành phần protein enzyme trong phức này. ARN đơn dẫn hướng (gARN) có thể được cấu trúc để đóng vai trò crARN và tracrARN cặp bazơ. Phức gARN/protein có thể quét hệ gen và xúc tác DSB ở các vùng mà bổ trợ cho crARN/gARN. Không giống các nucleaza được thiết kế khác, không chỉ oligome ngắn cần được thiết kế để cấu trúc các chất phản ứng được yêu cầu để hướng

đích gen quan tâm nhờ đó một loạt các bước tách dòng được yêu cầu để gắn ZFNs và TALENs.

Không giống các phương pháp tiêu chuẩn hiện nay cho việc phá vỡ gen, việc sử dụng các nucleaza được thiết kế đưa ra cơ hội sử dụng các hợp tử làm vật liệu khởi đầu cho GE. Các phương pháp tiêu chuẩn để phá vỡ gen ở vật nuôi liên quan đến HR trong các tế bào nuôi cấy và sự tái cấu trúc sau đó của phôi bằng cách chuyển nhân từ tế bào sinh dưỡng (SCNT). Vì động vật tách dòng được tạo ra qua SCNT đôi khi thể hiện các biểu hiện về sự khuyết thiếu về phát triển, con cháu của các thế hệ đầu SCNT/GE thường được sử dụng để nghiên cứu để tránh các dị thường SCNT và kiểu hình mà có thể xảy ra nếu động vật chuyển gen được sử dụng cho các thử nghiệm. Xem xét đến giai đoạn mang thai dài hơn và chi phí nhà xưởng cao hơn của lợn so với, có lợi ích về thời gian và chi phí đối với nhu cầu nhân giống giảm đi. Báo cáo gần đây chứng minh rằng việc tiêm trực tiếp ZFNs và TALENs vào các hợp tử lợn có thể phá gen nội sinh và tạo ra lợn với các đột biến mong muốn. Tuy nhiên, chỉ khoảng 10% lợn con thể hiện sự sửa đổi hai alen của gen đích, và một số thể hiện kiểu gen khảm. Bài báo gần đây chứng minh rằng hệ CRISPR/ Cas9 có thể gây ra các đột biến trong việc phát triển phôi và tạo ra lợn GE với hiệu quả cao hơn so với ZFNs hoặc TALENs. Tuy nhiên, lợn GE tạo ra từ hệ CRISPR/ Cas9 cũng có kiểu gen khảm. Ngoài ra, tất cả các nghiên cứu nêu trên đây sử dụng các hợp tử được tạo *in vivo* cho các thử nghiệm, điều này đòi hỏi sự lao động chuyên sâu và nhiều lợn nái để thu được số lượng hợp tử đủ.

Ví dụ của sáng chế mô tả phương pháp hữu hiệu để sử dụng hệ CRISPR/ Cas9 trong việc tạo ra lợn GE thông qua việc tiêm các hợp tử thu được *in vitro* và sự biến đổi của các tế bào sinh dưỡng sau đó là SCNT. Hai gen nội sinh (CD163 và CD1D) và một gen chuyển (eGFP) được hướng đích, và chỉ có các tế bào trước hoặc hợp tử thu được *in vitro* được sử dụng để tiêm SCNT hoặc ARN một cách tương ứng. CD163 được thấy là cần cho nhiễm trùng sinh sản bởi virut gây hội chứng hô hấp và sinh sản ở lợn, virut được biết là gây ra tổn thất kinh tế đáng kể đối với ngành công nghiệp lợn. CD1D được xem là protein phức tương hợp mô chính không kinh điển và tham gia vào việc đưa ra các kháng nguyên lipit cho các tế bào T giết tự nhiên bất biến. Lợn thiếu hụt các gen này được thiết kế là các mẫu cho ngành nông nghiệp và dược sinh. Gen chuyển eGFP được sử dụng làm đích cho các thử nghiệm kiểm chứng lí thuyết và sự tối ưu hóa của các phương pháp.

Vật liệu và phương pháp

Hóa chất và chất phản ứng. Trừ khi được nêu rõ theo cách khác, tất cả các hóa chất sử dụng trong thử nghiệm này được mua của Sigma.

Thiết kế gARNs để xây dựng các CRISPR cụ thể

Các ARNs dẫn hướng được thiết kế đối với các vùng trong exon 7 của *CD163* mà chỉ có ở *CD163* kiểu dại và không có trong vector hướng đích hoán đổi vùng (được mô tả dưới đây), sao cho CRISPR có thể dẫn đến DSB trong *CD163* kiểu dại nhưng không ở trong vectơ hướng đích hoán đổi vùng. Chỉ có bốn vùng trong đó vectơ hướng đích có thể đưa vào đa hình nucleotit đơn (Single Nucleotit polymorphism -SNP) mà có thể thay đổi (*Spy*) trình tự PAM (Protospacer Adjacent Motif -PAM) của *S. pyogenes* . Tất cả bốn đích được lựa chọn bao gồm:

(SEQ ID NO:1) GGAAACCCAGGCTGGTTGGA**gGG** (CRISPR 10),

(SEQ ID NO:2) GGAACTACAGTGCGGCACTG**tGG** (CRISPR 131),

(SEQ ID NO:3) CAGTAGCACCCCGCCCTGAC**gGG** (CRISPR 256) và

(SEQ ID NO:4) TGTAGCCACAGCAGGGACGT**cGG** (CRISPR 282). PAM có thể được nhận diện bằng phông chữ đậm trong mỗi gARN.

Đối với các đột biến *CD1D*, nghiên cứu đối với các đích CRISPR được tùy ý giới hạn đối với dải mã hóa trong 1000 bp đầu tiên của bản phiên mã đầu tiên. Tuy nhiên, RepeatMasker [26] (thư viện lặp của "lợn") xác định yếu tố lặp bắt đầu ở bazơ 943 của bản phiên mã đầu tiên. Nghiên cứu cho các đích CRISPR sau đó được giới hạn đối với 942 bp thứ nhất của bản phiên mã đầu tiên. Nghiên cứu cũng được giới hạn đối với 873 bp đầu tiên của bản phiên mã đầu tiên vì *Spy* PAM cuối cùng nằm ở bazơ 873. Đích thứ nhất (CRISPR 4800) được chọn vì nó chồng lấn với đơn vị mã khởi đầu nằm ở bazơ 42 trong bản phiên mã đầu tiên (CCAGCCTCGCCCAGCGACAT**gGG** (SEQ ID NO:5)). Hai đích bổ sung (CRISPRs 5620 và 5626) được chọn vì chúng ở xa nhất so với lựa chọn đầu tiên trong vùng lựa chọn tùy ý (CTTTCATTTATCTGAACTCA**gGG** (SEQ ID NO:6) và TTATCTGAACTCAGGGTCCC**cGG** (SEQ ID NO:7)). Các đích này chồng lấn nhau. Liên quan đến đơn vị mã khởi đầu, *Spy* PAMs ở đầu gần nhất nằm trong trình tự đơn chứa trình tự polyme đồng nhất hoàn toàn như được xác định bởi sự đánh giá bằng mắt.Đích thứ tư (CRISPR 5350) được lựa chọn vì, so với lựa chọn thứ nhất, nó ở đích gần nhất mà không chứa các vùng polyme đồng nhất hoàn toàn

(CAGCTGCAGCATATATTTAA**gGG** (SEQ ID NO:8)). Sự đặc trưng của các crARNs được thiết kế được xác nhận bằng cách tìm kiếm các trình tự lợn tương tự trong GenBank. Các oligonucleotit (bảng 1) gắn và được tách dòng thành vectơ p330X chứa hai catxet biểu hiện, *S. pyogenes* (h*Spy*) Cas9 được tối ưu hóa đơn vị mã của người và RNA. P330X dẫn hướng thể khảm được phân loại bằng *BbsI* (New England Biolabs) theo phương pháp phòng thí nghiệm Zhang (http://www.addgene.org/crispr/zhang/).

Để hướng đích *eGFP*, hai gARNs đặc hiệu hướng đích trình tự mã hóa *eGFP* được thiết kế trong 60 bp đầu tiên của đơn vị mã khởi đầu *eGFP*. Cả gARN *eGFP*1 và *eGFP*2 là trên dải đối nghĩa và eGFP1 trực tiếp hướng đích đơn vị mã khởi đầu. Trình tự gARN *eGFP*1 là CTCCTCGCCCTTGCTCACCA**tGG** (SEQ ID NO:9) và trình tự *eGFP*2 là GACCAGGATGGGCACCACCC**cGG** (SEQ ID NO:10).

Bảng 1. crARNs được thiết kế. Mồi 1 và mồi 2 được gắn theo phương pháp Zhang.

Đoạn mồi	**Trình tự (5' – 3')**	**SEQ ID NO.**
CD163 10 1	CACCGGAAACCCAGGCTGGTTGGA	48
CD163 10 2	AAACTCCAACCAGCCTGGGTTTCC	49
CD163 131 1	CACCGGAACTACAGTGCGGCACTG	50
CD163 131 2	AAACCAGTGCCGCACTGTAGTTCC	51
CD163 256 1	CACCGCAGTAGCACCCCGCCCTGAC	52
CD163 256 2	AAACGTCAGGGCGGGGTGCTACTGC	53
CD163 282 1	CACCGTGTAGCCACAGCAGGGACGT	54
CD163 282 2	AAACACGTCCCTGCTGTGGCTACAC	55
CD1D 4800 1	CACCGCCAGCCTCGCCCAGCGACAT	56
CD1D 4800 2	AAACATGTCGCTGGGCGAGGCTGGC	57
CD1D 5350 1	CACCGCAGCTGCAGCATATATTTAA	58
CD1D 5350 2	AAACTTAAATATATGCTGCAGCTGC	59
CD1D 5620 1	CACCGCTTTCATTTATCTGAACTCA	60
CD1D 5620 2	AAACTGAGTTCAGATAAATGAAAGC	61
CD1D 5626 1	CACCGTTATCTGAACTCAGGGTCCC	62
CD1D 5626 2	AAACGGGACCCTGAGTTCAGATAAC	63

eGFP 1 1	CACCGCTCCTCGCCCTTGCTCACCA	64
eGFP 1 2	AAACTGGTGAGCAAGGGCGAGGAGC	65
eGFP 2 1	CACCGGACCAGGATGGGCACCACCC	66
eGFP 2 2	AAACGGGTGGTGCCCATCCTGGTCC	67

Tổng hợp ADN cho cho gen CD163 và CD1D

Cả CD163 và CD1D của lợn đều được khuếch đại bằng PCR từ ADN phân lập từ các nguyên bào sợi của thai mà có thể được sử dụng cho chuyển nhiễm sau này để đảm bảo sự khớp về đồng hợp tử giữa vectơ hướng đích và dòng tế bào chuyển nhiễm. Tóm lại, taq LA (Clontech) sử dụng mồi xuôi CTCTCCCTCACTCTAACCTACTT (SEQ ID NO:11), và mồi ngược TATTTCTCTCACATGGCCAGTC (SEQ ID NO:12) được sử dụng để khuếch đại đoạn 9538 bp của CD163. Đoạn là trình tự ADN được phê chuẩn và được sử dụng để xây dựng vector hướng đích hoán đổi vùng (Fig. 1). Vector này bao gồm cả các đột biến 33 điểm trong exon 7 sao cho nó sẽ có thể mã hóa cho trình tự axit amin giống CD163L của người từ exon 11. Exon thay thế là 315 bp. Ngoài ra, intron sau đó được thay thế với intron myostatin B được biến đổi mà chứa gen đánh dấu có thể lựa chọn mà có thể được loại bỏ bằng Cre-recombinaza (Cre) và đã được chứng minh việc cắt bình trường khi việc neo vị trí loxP được giữ lại (Wells, các kết quả không công bố). Nhánh dài của cấu trúc là 3469 bp và bao gồm cả exon DS hoán đổi vùng. Nhánh ngắn là 1578 bp và bao gồm cả các exon 7 và 8 (Fig. 1, Panen B). Plasmit này được sử dụng trong nỗ lực để thay thế vùng mã hóa của exon 7 trong các thử nghiệm chuyển nhiễm đầu tiên và cho phép lựa chọn các sự kiện hướng đích thông qua gen chỉ thị chọn lọc (G418). Nếu việc hướng đích xảy ra, gen đánh dấu có thể bị xóa bởi Cre-recombinaza. Vector hướng đích CD163 DS sau đó được biến đổi để sử dụng với các dòng tế bào mà chứa gen SIGLEC1 được phá vỡ bằng Neo mà không thể bị xóa Cre. Trong vectơ hướng đích này, catxet Neo, loxP và intron myostatin B, không bị loại bỏ, và chỉ có exon DS exon được giữ lại với nhánh dài và nhánh ngắn WT (Fig. 1, Panen C).

Trình tự hệ gen cho CD1D của lợn được khuếch đại với taq LA taq sử dụng mồi xuôi CTCTCCCTCACTCTAACCTACTT(SEQ ID NO:13) và mồi ngược GACTGGCCATGTGAGAGAAATA (SEQ ID NO:14), thu được đoạn 8729 bp. Đoạn là trình tự ADNd và được sử dụng để xây dựng vectơ hướng đích được thể hiện trên

Fig. 2. Catxet Neo nằm dưới sự kiểm soát của đoạn gen khởi động kinaza phosphoglyxerol (PGK) và được gắn phía cánh với các trình tự loxP, mà được đưa vào để lựa chọn. Nhánh dài của cấu trúc là 4832 bp và nhánh ngắn là 3563 bp, và bao gồm cả các exon 6 và 7. Nếu HR xảy ra thành công, các exon 3, 4, và 5 có thể được loại bỏ và được thay thế bằng Catxet Neo. Nếu sự sửa chữa NHEJ xảy ra không chính xác, thì exon 3 có thể bị phá vỡ.

Thu nguyên bào sợi của bào thai

Mô bào thai lợn được thu vào ngày 35 của kỳ thai để tạo ra các dòng tế bào. Hai dòng tế bào nguyên bào sợi của bào thai đực và cái kiểu dại (WT) được thiết lập từ đoạn giao lớn màu trắng bên trong. Các nguyên bào sợi của thai đực và cái mà đã được biến đổi trước đó để chứa catxet Neo (SIGLEC1-/- di truyền học) cũng được sử dụng trong các nghiên cứu này. Các nguyên bào sợi của bào thai được thu như được mô tả với các biến đổi nhỏ; mô được băm nhỏ từ mỗi bào thai được phân loại trong 20 ml môi trường phân loại (môi trường Eagle được biến đổi Dulbecco [DMEM] chứa L-glutamin và 1 g/L D-glucoza [Cellgro] được bổ sung bằng 200 đơn vị/ml colagenaza và 25 đơn vị Kunitz /ml ADNseI) trong 5 giờ ở 38,5°C. Sau khi phân loại, các nguyên bào sợi bào thai được rửa và được nuôi cấy với DMEM, 15% huyết thanh bào thai bò (FBS), và 40 µg/ml gentamicin. Sau khi nuôi cấy qua đêm, các tế bào được typsin hóa và làm đông ở –80°C trong các phần phân ước trong FBS với 10% dimetyl sulfoxit và được lưu trong nitơ lỏng.

Gây nhiễm tế bào và kiểu gen

Các điều kiện chuyển nhiễm về cơ bản như đã được báo cáo trước đây. ADN cho luôn luôn được sử dụng với lượng không đổi là 1 µg với lượng thay đổi của plasmit CRISPR/Cas9 (được liệt kê dưới đây). ADN cho được xếp thẳng hàng bằng MLUI (CD163) (NEB) hoặc AFLII (CD1D) (NEB) trước khi chuyển nhiễm. Giới tính của các dòng tế bào được thiết lập được xác định bằng PCR như được mô tả trước đó trước khi chuyển nhiễm. Cả dòng tế bào đực và cái đều được chuyển nhiễm, dữ liệu biến đổi hệ gen được phân tích cùng giữa chuyển nhiễm. Các dòng tế bào nguyên bào sợi bào thai có số passage tương tự (2–4) được nuôi cấy trong 2 ngày và được phát triển đến 75%–85% hợp lưu trong DMEM chứa L-glutamin và 1 g/L D-glucoza (Cellgro) được bổ sung bằng 15% FBS, 2,5 ng/ml yếu tố sinh trưởng nguyên bào sợi cơ sở, và 10 mg/ml gentamicin. Các nguyên bào sợi được rửa bằng nước muối đệm phosphat (PBS) (Life

Technologies) và được trypsin hóa. Ngay khi các tế bào rời ra, các tế bào được rửa bằng môi trường trong phương pháp xung điện (75% muối dùng cho tế bào [120 mM KCl, 0,15 mM $CaCl_2$, 10 mM K_2HPO_4, pH 7,6, 5 Mm $MgCl_2$]) và 25% Opti-MEM (LifeTechnologies). Nồng độ tế bào được định lượng bằng cách sử dụng huyết cầu kế. Các tế bào được tạo pellet ở 600 X g trong 5 phút và được tái tạo huyền phù ở nồng độ là 1 X 10^6 trong môi trường của phương pháp xung điện. Mỗi phương pháp xung điện sử dụng 200 µl tế bào trong 2 mm cuvet hở với ba (1 m giây) xung sóng vuông được sử dụng thông qua BTX ECM 2001 ở 250 V. Sau phương pháp xung điện, các tế bào được tái tạo huyền phù trong DMEM mô tả ở trên. Để lựa chọn, 600 µg/ml G418 (Life Technologies) được thêm vào 24 giờ sau khi chuyển nhiễm, và môi trường được thay đổi vào ngày 7. Khuẩn lạc được lấy vào ngày 14 sau khi chuyển nhiễm. Các nguyên bào sợi bào thai được cấy ở mật độ 10,000 tế bào/đĩa nếu sự lựa chọn G418 được sử dụng và ở 50 tế bào/đĩa nếu không có sự lựa chọn G418 được sử dụng. Các khuẩn lạc nguyên bào sợi bào thai được thu bằng cách áp dụng xylanh tách dòng 10 mm được hấp được gắn kín xung quanh mỗi khuẩn lạc bằng mỡ chân không được hấp. Các khuẩn lạc được rửa bằng PBS và thu qua trypsin; sau khi tái tạo huyền phù trong môi trường nuôi cấy DMEM. Một phần (1/3) khuẩn lạc được tái tạo huyền phù được chuyển vào đĩa PCR 96 giếng, và các tế bào còn lại (2/3) được nuôi cấy trong giếng của đĩa 24 giếng. Các pallet tế bào được tái tạo huyền phù trong 6 µl đệm dung giải (40 mM Tris, pH 8,9, 0,9% Triton X-100, 0,4 mg/ml proteinaza K [NEB]), được ủ ở 65°C trong 30 phút để dung giải tế bào, sau đó là 85°C trong 10 phút để bất hoạt proteinaza K.

Sàng lọc PCR cho DS và đoạn xóa nhỏ và lớn

Phát hiện việc sửa chữa được định hướng bởi HR. PCRs dải dài được sử dụng để nhận diện các đột biến trên CD163 hoặc CD1D. Ba thử nghiệm PCR khác nhau được sử dụng để nhận diện các sự kiện HR: khuếch đại PCR của các vùng mở rộng từ các trình tự CD163 hoặc CD1D trong ADN cho đến các trình tự CD163 hoặc CD1D nội sinh trên PCR phía trái hoặc phía phải và PCR dải dài mà khuếch đại các vùng rộng lớn của CD163 hoặc CD1D bao gồm các ADN cho được thiết kế. Việc tăng kích thước của sản phẩm PCR, 1,8 kb (CD1D) hoặc 3,5 kb (CD163), nảy sinh từ việc thêm các trình tự Neo ngoại sinh được xem là bằng chứng của việc sửa chữa được định hướng bởi HR của các gen. Tất cả các điều kiện PCR bao gồm cả sự biến tính ban đầu ở 95°C trong 2 phút sau đó là 33 chu trình 30 giây ở 94°C, 30 giây ở 50°C, và 7–10 phút ở 68°C. Taq

LA được sử dụng cho tất cả các thử nghiệm theo khuyến cáo của nhà sản xuất. Các đoạn mồi được thể hiện trong bảng 2.

Bảng 2. Các đoạn mồi được sử dụng để nhận diện sự sửa chữa được định hướng bởi HR của CD163 và CD1D

Đoạn mồi	**Trình tự (5' – 3')**	**SEQ ID NO.**
Mồi thử nghiệm dải dài của CD163 1230F	TTGTTGGAAGGCTCACTGTCCTTG	68
Mồi thử nghiệm dải dài của CD163 7775 R	ACAACTAAGGTGGGGCAAAG	69
Mồi thử nghiệm nhánh trái của CD163 1230 F	TTGTTGGAAGGCTCACTGTCCTTG	70
Mồi thử nghiệm nhánh trái của CD163 8491 R	GGAGCTCAACATTCTTGGGTCCT	71
Mồi thử nghiệm nhánh phải của CD163 3752 F	GGCAAAATTTTCATGCTGAGGTG	72
Mồi thử nghiệm nhánh phải của CD163 7765 R	GCACATCACTTCGGGTTACAGTG	73
Mồi thử nghiệm dải dài của CD1D F 3991 F	CCCAAGTATCTTCAGTTCTGCAG	74
Mồi thử nghiệm dải dài của CD1D R 12806 R	TACAGGTAGGAGAGCCTGTTTTG	75
Mồi thử nghiệm nhánh trái của *CD1D* F 3991 F	CCCAAGTATCTTCAGTTCTGCAG	76
Mồi thử nghiệm nhánh trái của *CD1D* 7373 R	CTCAAAAGGATGTAAACCCTGGA	77
Mồi thử nghiệm nhánh phải của *CD1D* 4363 F	TGTTGATGTGGTTTGTTTGCCC	78
Mồi thử nghiệm nhánh phải của *CD1D* 12806 R	TACAGGTAGGAGAGCCTGTTTTG	79

Thử nghiệm xóa đoạn nhỏ (NHEJ). Sự xóa đoạn nhỏ được xác định bằng khuếch đại PCR của CD163 hoặc CD1D. Điểm cắt phía cánh được xác lập bởi hệ CRISPR/Cas9. Kích thước của các sản phẩm khuếch đại tương ứng là 435 bp và 1244 bp cho CD163 và CD1D. Dịch dung giải từ cả phôi và các nguyên bào sợi của thai được khuếch đại bằng PCR với taq LA. Các điều kiện PCR của thử nghiệm là trước tiên biến tính ở 95°C trong 2 phút sau đó là 33 chu kỳ 30 giây ở 94°C, 30 giây ở 56°C, và 1 phút ở 72°C. Để đọc kiểu gen các tế bào chuyển nhiễm, sự xen vào và xóa đoạn (INDELs) được nhận diện bằng cách phân tách các sản phẩm khuếch đại PCR bằng điện di trên gel agaroza. Để độc kiểu gen của phôi, các sản phẩm PCR thu được sau đó được đọc trình tự ADN để nhận diện các đoạn xóa nhỏ sử dụng các đoạn mồi xuôi được sử dụng trong PCR. Thông tin mồi được thể hiện trong bảng 3.

Bảng 3. Các đoạn mồi được sử dụng để nhận diện các đột biến thông qua NHEJ trên CD163 và CD1D

Đoạn mồi	**Trình tự (5' – 3')**	**SEQ ID NO.**
G*CD163*F	GGAGGTCTAGAATCGGCTAAGCC	80
G*CD163*R	GGCTACATGTCCCGTCAGGG	81
G*CD1D*F	GCAGGCCACTAGGCAGATGAA	82
G*CD1D*R	GAGCTGACACCCAAGAAGTTCCT	83
*eGFP*1	GGCTCTAGAGCCTCTGCTAACC	84
*eGFP*2	GGACTTGAAGAAGTCGTGCTGC	85

Sự chuyển nhân tế bào soma (SCNT)

Để tạo ra SCNT phôi, tế bào trứng thu được từ lợn nái (ART, Inc.) hoặc tế bào trứng thu được từ lợn cái con từ lò mổ địa phương được sử dụng. Tế bào trứng thu được từ lợn nái được vận chuyển qua đêm trong môi trường thành thục (TCM-199 với 2,9 mM Hepes, 5 µg/ml insulin, 10 ng/ml yếu tố sinh trưởng biểu bì [EGF], 0,5 µg/ml hocmon kích thích nang lợn [p-FSH], 0,91 mM pyruvat, 0,5 mM xystein, 10% dịch nang trứng lợn, và 25 ng/ml gentamicin) và được chuyển vào môi trường tươi sau 24 h. Sau 40–42 giờ làm chín, các tế bào kết tụ được loại bỏ từ noãn bào bằng vortex với sự có mặt của 0,1% hyaluronidaza. Các tế bào trứng thu được từ lợn cái non được làm trưởng thành như mô tả dưới đây để thụ tinh *in vitro* (IVF). Trong quá trình thao tác,

các tế bào trứng được đặt trong môi trường thao tác (TCM-199 [Life Technologies] với 0,6 mM $NaHCO_3$, 2,9 mM Hepes, 30 mM NaCl, 10 ng/ml gentamicin, và 3 mg/ml BSA, với nồng độ mol os là 305 mOsm) bổ sung bằng 7,0 µg/ml cytochalasin B. Thể cực cùng với phần tế bào chất gần kề, được cho là chứa đĩa metaphase II, được loại bỏ, và tế bào cho được đặt trong khoang quanh phôi bằng cách sử dụng mao dẫn thủy tinh mỏng. Phôi được tái cấu trúc sau đó được dung hợp trong môi trường dung hợp (0,3 M mannitol, 0,1 mM $CaCl_2$, 0,1 mM $MgCl_2$, và 0,5 mM Hepes) với hai xung DC (cách nhau 1 giây) ở 1,2 kV/cm trong 30 giây sử dụng dụng cụ thao tác tế bào BTX Electro (Harvard Apparatus). Sau khi dung hợp, phôi dung hợp được hoạt hóa hoàn toàn bằng 200 µM thimerosal trong 10 phút trong bóng tối và 8 mM dithiothreitol trong 30 phút. Phôi sau đó được ủ trong môi trường hợp tử của lợn được cải biến PZM3-MU1 với 0,5 µM Scriptaid (S7817; Sigma-Aldrich), chất ức chế deaxetylaza histon, trong 14–16 giờ, như được mô tả trước đây.

Thụ tinh in vitro (IVF)

Để IVF, các tế bào trứng từ lợn cái non trước tuổi dậy thì thu được từ lò mổ (Farmland Foods Inc.). Các tế bào trứng non này được hút ra từ các nang kích thước trung bình (3–6 mm) sử dụng kim tiêm dưới da 18-gauge được gắn với xy lanh 10 ml. Tế bào trứng với tế bào chất sậm màu đều nhau và các tế bào trần kết tụ xung quanh sau đó được lựa chọn để trưởng thành hóa. Khoảng 50 phức hệ noãn bào kết tụ được đặt trong giếng chứa 500 µl môi trường thành thục, TCM-199 (Invitrogen) với 3,05 mM glucoza, 0,91 mM natri pyruvat, 0,57 mM xystein, 10 ng/ml EGF, 0,5 µg/ml hocmon lutein hóa (LH), 0,5 µg/ml FSH, 10 ng/ml gentamicin (APP Pharm), và 0,1% rượu polyvinylic trong 42–44 giờ ở 38,5°C, 5% CO_2, trong không khí được làm ẩm. Khi kết thúc việc trưởng thành hóa, các tế bào kết tụ xung quanh được loại bỏ khỏi noãn bào bằng cách vortex trong 3 phút với sự có mặt của 0,1% hyaluronidaza. Sau đó, các tế bào trứng được trưởng thành *in vitro* được đặt trong 50 µl giọt nhỏ môi trường IVF (môi trường đệm Tris được cải biến chứa 113,1 mM NaCl, 3 mM KCl, 7,5 mM CaCl2, 11 mM glucoza, 20 mM Tris, 2 mM cafein, 5 mM natri pyruvat, và 2 mg/ml albumin huyết thanh bào thai bò [BSA]) trong các nhóm gồm 25–30 tế bào trứng. Một pellet tinh dịch lạnh đông 100 µl được làm tan trong 3 ml PBS Dulbecco bổ sung bằng 0,1% BSA. Tinh dịch WT lạnh đông hoặc tinh dịch eGFP tươi được rửa trong 60% Percoll trong 20 phút ở 650 3 g và trong môi trường đệm Tris được cải biến trong 10 phút bằng cách ly tâm.

Trong một số trường hợp, tinh dịch được thu mới mà dị hợp đối với gen chuyển eGFP được mô tả trên đây được rửa ba lần trong PBS. Pellet tinh dịch sau đó được tái tạo huyền phù với môi trường IVF thành 0,5 X 10^6 tế bào/ml. 50 µl huyền phù tinh dịch được đưa vào trong các giọt tế bào trứng. Các giao tử được đồng thời ủ trong 5 giờ ở 38,5°C trong môi trường 5% CO_2 trong không khí. Sau khi thụ tinh, phôi được ủ trong PZM3-MU1 ở 38,5°C và 5% CO_2 trong không khí.

Chuyển phôi

Phôi được tạo để tạo ra GE CD163 hoặc lợn CD1D được chuyển vào các vật thay thế vào ngày 1 (SCNT) hoặc ngày 6 (hợp tử được nạp vào) sau thời kì động dục thứ nhất. Đối với việc chuyển vào ngày 6, các hợp tử được nuôi cấy thêm 5 ngày trong PZM3-MU1 với sự có mặt của 10 ng/ml ps48 (Stemgent, Inc.). Phôi được chuyển theo cách phẫu thuật vào ampullary-isthmic junction của ống dẫn trứng của vật thay thế.

Tổng hợp In Vitro của ARN cho hệ CRISPR/Cas9

Mẫu ADN để phiên mã *in vitro* được khuếch đại sử dụng PCR (bảng 4). Plasmit CRISPR/Cas9 được sử dụng cho các thử nghiệm chuyển nhiễm tế bào dùng làm mẫu cho PCR. Để biểu hiện Cas9 trong các hợp tử, kit mMESSAGE mMACHINE Ultra (Ambion) được sử dụng để tạo ra mARN của Cas9. Sau đó tín hiệu poly A được thêm vào mARN Cas9 sử dụng kit gắn đuôi Poly (A) (Ambion). ARNs dẫn hướng CRISPR được tạo ra bởi MEGAshortscript (Ambion). Chất lượng của ARNs được tổng hợp được quan sát bằng mắt trên gel agaroza 1,5% và sau đó pha loãng đến nồng độ cuối là 10 ng/µl (cả gARN và Cas9) và được chia thành ba phần 3 µl.

Bảng 4. Các đoạn mồi được sử dụng để khuếch đại các mẫu để phiên mã *in vitro*.

Các đoạn mồi	**Trình tự (5' – 3')**	**SEQ ID NO.**
Cas9	F: TAATACGACTCACTATAGGGAGAATGGACTATAAGGACCACG AC R: GCGAGCTCTAGGAATTCTTAC	86 87
eGFP 1	F: TTAATACGACTCACTATAGGCTCCTCGCCCTTGCTCACCA R: AAAAGCACCGACTCGGTGCC	88 89
CD163 10	F: TTAATACGACTCACTATAGGAAACCCAGGCTGGTTGGA R: AAAAGCACCGACTCGGTGCC	90 91
CD163 131	F: TTAATACGACTCACTATAGGAACTACAGTGCGGCACTG R: AAAAGCACCGACTCGGTGCC	92 93
CD1D 4800	F: TTAATACGACTCACTATAGGCCAGCCTCGCCCAGCGACAT R: AAAAGCACCGACTCGGTGCC	94 95
CD1D 5350	F: TTAATACGACTCACTATAGGCAGCTGCAGCATATATTTAA R: AAAAGCACCGACTCGGTGCC	96 97

Vi tiêm hệ CRISPR/Cas9 được thiết kế trong các hợp tử

ARN thông tin mã hóa Cas9 và gARN được tiêm vào tế bào chất của tế bào trứng được thụ tinh 14 giờ sau khi thụ tinh (hợp tử ước đoán) sử dụng dụng cụ vi tiêm FemtoJet (Eppendorf). Việc vi tiêm được tiến hành trong môi trường thao tác ở giai đoạn đốt nóng của kính hiển vi đảo ngược của Nikon (Nikon Corporation; Tokyo, Japan). Các hợp tử được tiêm sau đó được chuyển vào PZM3-MU1 với 10 ng/ml ps48 cho đến sử dụng tiếp theo.

Phân tích thống kê

Số lượng khuẩn lạc với hệ gen sửa đổi được phân loại là 1, và các khuẩn lạc mà không có sửa đổi của hệ gen được phân loại là 0. Các khác biệt được xác định bằng cách sử dụng PROC GLM (SAS) với giá trị P là 0,05 được xem là có nghĩa. Các giá trị trung bình được tính toán là giá trị bình phương tối thiểu. Dữ liệu được thể hiện là trung bình số học ± SEM.

KẾT QUẢ

Hạ gục CD163 và CD1D được trung gian bởi CRISPR/Cas9 trong các tế bào sinh dưỡng

Hiệu quả của bốn plasmit CRISPR khác nhau (dẫn hướng 10, 131, 256, và 282) hướng đích CD163 được thử nghiệm ở lượng là 2 μg/μl of ADN cho (bảng 5). CRISPR 282 dẫn đến sự tạo ra khuẩn lạc nhiều hơn mức trung bình đáng kể hơn là xử lý bằng CRISPR 10 và 256 ($P < 0,05$). Từ thử nghiệm PCR dải dài được mô tả ở trên, việc xóa đoạn lớn được thấy trong khoảng từ 503 bp đến nhiều khoảng 1506 bp thay vì DS thông qua HR như ban đầu dự định ban đầu (Fig. 3, Panen A). Điều này không được mong đợi vì các báo cáo trước đây với các hệ chỉnh sửa ADN khác thể hiện việc xóa đoạn nhỏ hơn nhiều là 6–333 bp sử dụng ZFN ở lợn. CRISPR 10 và tổng hợp của tất cả bốn CRISPR dẫn đến số lượng khuẩn lạc cao hơn với hệ gen sửa đổi so với CRISPR 256 và 282 (bảng 5, $P < 0,002$). Việc chuyển nhiễm của CRISPR 10 và plasmit chứa Neo nhưng không có sự tương đồng đối với CD163 dẫn đến không có khuẩn lạc nào thể hiện sự xóa đoạn lớn. Thú vị là một xóa đoạn một alen cũng được phát hiện khi ADN cho được đưa vào mà không có bất kỳ CRISPR. Thử nghiệm này có lẽ thể hiện sự đánh giá thấp tỷ lệ đột biến vì việc đoạn xóa nhỏ tiềm năng bằng cách giải trình tự mà có thể không được phát hiện trên gel agaroza trong các tế bào sinh dưỡng được chuyển nhiễm là không được sàng lọc.

Bảng 5. Hiệu quả của bốn plasmit CRISPR khác nhau (dẫn hướng 10, 131, 256, và 282) hướng đích CD163. Bốn CRISPR khác nhau được thử nghiệm với lượng 2 µg đến 1 µg ADN cho (thể hiện trên Fig. 1).

Xử lý*	Tổng số khuẩn lạc	Tổng số đĩa	Số lượng khuẩn lạc/đĩa trung bình †	Số lạc NHEJ	Khuẩn lạc với HR	Phần trăm khuẩn lạc có hệ gen sửa đổi†	Reps
10+ADN cho	76	102	$0,75^{bc}$	11	1‡	$15,79^{a}$	4
131+ ADN cho	102	51	$2,00^{ab}$	11	0	$10,78^{ab}$	3
256+ ADN cho	43	49	$0,88^{c}$	2	0	$4,65^{bc}$	3
282+ ADN cho	109	46	$2,37^{a}$	3	0	$2,75^{bc}$	3
Hỗn hợp 4+ ADN cho	111	55	$2,02^{ab}$	20	0	$18,02^{a}$	3
ADN cho	48	52	$0,92^{bc}$	1	0	$2,08^{bc}$	3
10 + *Neo* (no *CD163*)	26	20	$1,3^{n/a}$	0	0	$0,00^{c}$	1

* Hỗn hợp của 4 + ADN cho thể hiện việc trộn lẫn giống nhau của 0,5µg của mỗi CRISPR với 1 µg ADN cho. Xử lý ADN cho dùng làm đối chứng không có CRISPR và xử lý 10 + *Neo* chứng minh rằng, các xóa đoạn lớn quan sát thấy trong các xử lý CRISPR chỉ có mặt khi *CD163* ADN cho cũng có mặt.

† ANOVA được tiến hành so sánh số lượng trung bình của cụm /đĩa để đánh giá độc tính CRISPR và về tỷ lệ cụm có hệ gen sửa đổi. Giá trị P- lần lượt là 0,025 và 0,0002. n/a = Không có sự sao chép cho xử lý này dodos không có dữ liệu thống kê được tiến hành.

‡ Một khuẩn lạc với HR thể hiện sự kiện HR một phần.

$^{a–c}$ Các chữ viết nhỏ phía trên chỉ ra sự khác biệt đáng kể giữa các xử lý cho số lượng khuẩn lạc trung bình/đĩa và phần trăm khuẩn lạc với hệ gen sửa đổi ($P < 0,05$).

42443

95/283

Mục đích ban đầu là thu được sự kiện hướng định hoán đổi vùng (DS) bởi HR cho CD163, nhưng các CRISPR không làm tang hiệu quả của việc hướng đích CD163. Nên chú ý rằng các tổ hợp khác nhau của vectơ hướng đích này đã được sử dụng để sửa đổi CD163 bởi HR bằng các chuyển nhiễm thông thường và thu được 0 sự kiện hướng đích sau khi sàng lọc 3399 khuẩn lạc (Whitworth và Prather, các kết quả không công bố). Hai lợn được thu được với DS đầy đủ thu được từ HR chứa tất cả 33 đột biến mà đã được nỗ lực để đưa vào bằng cách chuyển nhiễm với CRISPR 10 và vectơ hướng đích DS làm ADN cho.

Tiếp theo, hiệu quả của các đột biến gây ra bằng CRISPR/Cas9 không cần chọn lọc thuốc được thử nghiệm; dòng tế bào nguyên bào sợi bào thai sử dụng trong nghiên cứu này đã có sự tích hợp của catxet hỗ trợ Neo và làm gục SIGLEC1. Việc liệu tỷ lệ CRISPR/Cas9 và ADN cho có thể làm tăng sự biến đổi của hệ gen hoặc dẫn đến hiệu quả độc hại ở nồng độ cao là cũng được thử nghiệm. CRISPR 131 được lựa chọn cho thử nghiệm này vì trong thử nghiệm trước, nó dẫn đến số lượng khuẩn lạc tổng cao và tỷ lệ % gia tăng của về các khuẩn lạc có hệ gen sửa đổi. Việc tăng lượng CRISPR 131 ADN từ 3:1 đến 20:1 không có hiệu quả đáng kể đối với khả năng sống sót của nguyên bào sợi bào thai. Phần trăm khuẩn lạc với hệ gen được sửa đổi bởi NHEJ không khác biệt đáng kể giữa các nồng độ CRISPR khác nhau nhưng có số lượng cao nhất của NHEJ ở tỷ lệ 10:1 (bảng 6, P = 0,33). Thậm chí ở tỷ lệ cao nhất của ADN CRISPR đối với ADN cho (20:1), HR là không quan sát thấy.

Bảng 6. Hiệu quả của các đột biến gây ra bằng CRISPR/Cas9 không cần chọn lọc thuốc. Bốn tỷ lệ khác nhau của ADN cho trên CRISPR 131 ADN được so sánh ở dòng tế bào sửa đổi trước đó không cần sử dụng chọn lọc G418.

Tỷ lệ ADN cho: CRISPR	Số đĩa	Số khuẩn lạc	Số khuẩn lạc trung bình/đĩa	Số khuẩn lạc NHEJ	Phần trăm khuẩn lạc với NHEJ	Khuẩn lạc với HR	Phần trăm khuẩn lạc với HR	Reps
1:0	30	79	2.6	1	1,3[a]	0	0.0	2
1:3	30	84	2.8	1	1,2[a]	0	0.0	2
1:5	27	76	2.8	2	2,6[a]	0	0.0	2
1:10 SA	32	63	2.0	5	7,9[a]	0	0.0	2
1:20	35	77	2.2	3	3,9[a]	0	0.0	2

[a] Sự khác biệt đáng kể giữa các điều trị về phần trăm khuẩn lạc với sự chỉnh sửa NHEJ (P>0.05).

[b] Không có sự khác biệt đáng kể về số lượng khuẩn lạc được biến đổi hệ gen với sự tăng nồng độ của CRISPR (P>0,33).

42443

97/283

Dựa trên thí nghiệm này, sự đứt gãy hướng đích của CD1D trong các tế bào sinh dưỡng là bất thành. Bốn CRISPR khác nhau được thiết kế và thử nghiệm trên cả tế bào đực và cái. Các biến đổi của CD1D có thể được phát hiện từ ba CRISPR được áp dụng, tuy nhiên việc sử dụng CRISPR 5350 không dẫn đến sự biến đổi CD1D với đoạn khuyết đủ lớn để phát hiện bằng cách điện di trên gel agaroza (Bảng 7). Thú vị là, không thu được sự biến đổi gen thông qua HR mặc dù ADN cho được đưa vào. Tuy nhiên, quan sát thấy các đoạn khuyết lớn tương tự với các thí nghiệm knockout CD163 (Fig. 3, ô B). Không phát hiện thấy sự biến đổi hướng đích của CD1D với đoạn khuyết lớn khi CRISPR/Cas9 không được dùng với ADN cho. Tỷ lệ biến đổi CD1D từ quá trình hướng đích được điều khiển bởi CRISPR/Cas9 là 4/121 và 3/28 ở các khuẩn lạc tế bào từ lợn cái và lợn đực, tương ứng. Chỉ các INDEL có thể phát hiện được bằng cách điện di trên gel agaroza được bao hàm trong dữ liệu chuyển nhiễm.

Bảng 7. Bốn CRISPR khác nhau được thử nghiệm với lượng 2 µg cho 1 µg ADN cho (được chỉ ra trên Fig. 2). Việc xử lý ADN cho được dùng làm mẫu đối chứng không có CRISPR.

Giới tính	**Xử lý**	**Tổng số khuẩn lạc**	**INDEL**	**Hiệu lực (%)**
đực	4800 +ADN cho	29	2	6,9
đực	5350+ADN cho	20	0	0
đực	5620+ADN cho	43	1	2,33
đực	5626+ADN cho	29	2	6,9
đực	ADN cho	28	0	0
cái	4800 +ADN cho	2	0	0
cái	5350+ADN cho	8	0	0
cái	5620+ADN cho	10	0	0
cái	5626+ADN cho	8	3	37,5
cái	ADN cho	7	0	0

Sản xuất lợn CD163 và CD1D nhờ SCNT có sử dụng tế bào GE

Các tế bào thể hiện sự biến đổi CD163 hoặc CD1D được dùng cho SCNT để tạo ra lợn knockout CD163 và CD1D (Fig. 3). Thực hiện bảy lần chuyển phôi (CD163 Bảng 8), sáu lần chuyển phôi (CD163-No Neo), và năm lần chuyển phôi (CD1D) vào lợn nái con nhận với các phôi SCNT từ các nguyên bào sợi của lợn đực và cái sơ sinh đã được chuyển nhiễm với hệ CRISPR/ Cas9. Sáu (CD163), hai (CD163-No Neo), và bốn (CD1D) (Bảng 9) của lợn nái con nhận vẫn mang thai đến hạn do tỷ lệ thụ thai là 85,7%. 33,3%, và 80%, tương ứng. Năm con lợn con khỏe mạnh đã được sinh ra từ các cá thể nhận CD163 bằng cách mổ đẻ. Một con (O044) được đẻ tự nhiên. Cỡ lứa nằm trong khoảng từ một đến tám. Bốn con lợn được làm chết không đau đớn do chúng không lớn mau sau khi đẻ. Một con lợn con được làm chết không đau đớn do hở hàm ếch nghiêm trọng. Tất cả các con lợn con còn lại tỏ ra khỏe mạnh (Fig. 3, ô C). Hai đàn lợn con đực tạo ra từ các nguyên bào sợi của thai được chuyển nhiễm với CRISPR 10 và ADN cho được mô tả trên Fig. 3, ô B có đoạn khuyết 30 bp trong exon 7 liền kề với CRISPR 10 và một đoạn khuyết 1476 bp nữa của intron đứng trước, do đó tháo rời mối nối intron 6/exon 7 của CD163 (Fig. 3, ô E). Kiểu gen và sự dịch mã được dự báo được tóm tắt trong bảng 10. Một lợn con đực và một lứa đẻ gồm các con lợn cái (4 lợn con) thu được từ sự chuyển nhiễm CD163-No Neo các tế bào SIGLEC1 đã được biến đổi trước đây. Cả năm lợn con được bất hoạt kép SIGLEC1 và CD163. Lợn con đực có sự biến đổi hai alen của CD163 bằng đoạn khuyết 28 bp trong exon 7 trên một alen và đoạn khuyết 1387 bp trên alen còn lại mà bao gồm cả sự khuyết đoạn một phần của exon 7 và sự khuyết hoàn toàn của exon 8 và intron hoạt động, vì vậy tháo rời mối nối intron exon. Lợn con cái mang đột biến hai alen của CD163, bao gồm đoạn khuyết 1382 bp với đoạn cài xen 11 bp trên một alen và đoạn khuyết 1720 bp của CD163 trên alen còn lại. Tóm tắt về sự biến đổi CD163 và sự dịch mã dự báo có thể được tìm thấy trong Bảng 10. Tóm tắt về sự biến đổi CD1D và sự dịch mã dự báo do sự biến đổi CRISPR có thể được tìm thấy trong Bảng 11. Tóm lại, các lứa gồm một lợn cái và hai lợn đực được sinh ra, tạo ra 13 lợn con. Một lợn con chết ngay sau khi sinh. Mười hai trong số 13 lợn con chứa đoạn khuyết hai alen hoặc đồng hợp tử của CD1D (Fig. 3, ô F). Một con lợn con là WT.

Bảng 8. Dữ liệu chuyển phôi đối với *CD163*.

ID lợn	Dòng*	Giới tính	# phôi được chuyển	Nguồn noãn bào †	Ngày động dục	Kết quả đối với lợn con
O047	*CD163* CRISPR NT	Đực	240	ART	2	4 lợn con sống (2 được làm cho chết không đau đớn sau khi sinh)
O015	*CD163* CRISPR NT	Đực	267	ART	1	3 lợn con sống (tất cả đều khỏe mạnh)
O044	*CD163* CRISPR NT	Đực	206	ART	1	7 lợn con sống (1 sinh ra đã chết, 1 được làm cho chết không đau đớn sau khi sinh)
O053	*CD163* CRISPR NT	Đực	224	ART	2	1 lợn con đực (được làm cho chết không đau đớn ở ngày 13)
O08	*CD163* CRISPR NT	Đực	226	ART	1	0 lợn con
O094	*CD163* CRISPR NT	Đực	193	MU	2	8 lợn con sống (1 được làm cho chết không đau đớn do FTT)
O086	*CD163* CRISPR NT	Đực	213	MU	1	9 lợn con sống (2 được làm cho chết không đau đớn ở ngày 0, 2 do FTT)
O082	*CD163* Được bơm CRISPR 10/131	Đực /Cái	50 Blast	MU	5	0 lợn con
O083	*CD163* Được bơm CRISPR 10/131	Đực	46 Blast	MU	5	4 lợn con sống

42443

100/283

O99	*CD163* CRISPR NT-no *Neo*	Đực	156	ART	1	1 lợn con sống, 1 lợn con chết
O128	*CD163* CRISPR NT-no *Neo*	Đực	196	ART	2	0 lợn con
O100	*CD163* CRISPR NT-no *Neo*	Đực	261	MU	3	0 lợn con
O134	*CD163* CRISPR NT-no *Neo*	Đực/Cái	181	MU	1	0 lợn con
200889	*CD163* CRISPR NT-no *Neo*	Cái	202	ART	1	4 lợn con sống
O135	*CD163* CRISPR NT-no *Neo*	Cái	169	ART	2	0 lợn con

* Dòng *CD163* CRISPR NT là phôi được tạo ra bằng cách NT với dòng nguyên bào sợi thai đã được biến đổi bằng chuyển nhiễm. Phôi được bơm CRISPR là phôi IVF được bơm ở giai đoạn 1 của tế bào với ARN dẫn hướng *CD163* với ARN CAS9. Dòng *CD163* CRISPR NT-no *Neo* bào thai là phôi được tạo ra bằng NT với nguyên bào sợi thai đã được biến đổi trước đây mà là dòng kháng *Neo* được biến đổi bằng cách chuyển nhiễm không cần sử dụng gen đánh dấu có thể chọn lọc.

† MU dùng để chỉ tế bào trứng lợn nái con được hút và hoàn thiện ở University of Missouri như được mô tả trong phần IVF thuộc phần Vật liệu và phương pháp. ART dùng để chỉ tế bào trứng lợn nái được mua và hoàn thiện như được mô tả trong phần SCNT thuộc phần Vật liệu và phương pháp.

42443

Bảng 9. Dữ liệu chuyển phôi đối với *CD1D*.

ID Lợn	Dòng*	Giới tính	# phôi được chuyển nhiễm	Nguồn noãn bào †	Ngày động dục	Kết quả
200888	*CD1D* CRISPR NT	Đực	201	ART	2	7 lợn con sống
O61	*CD1D* CRISPR NT	Đực	239	ART	0	4 lợn con sống
O164	*CD1D* CRISPR NT	Cái	199	MU	2	0 lợn con
O156	*CD1D* CRISPR NT	Cái	204	MU	2	0 lợn con
O165	Được bơm *CD1D* 4800/5350	Đực/Cái	55 phôi	MU	6	4 lợn con (1 cái, 3 đực)
O127	Được bơm *CD1D* 4800/5350	Đực/Cái	55 phôi	MU	6	0 lợn con
O121	*CD1D* CRISPR NT	Cái	212	ART	1	2 lợn con sống

* Dòng *CD1D* CRISPR NT là phôi được tạo ra bằng cách NT với dòng nguyên bào sợi thai đã được biến đổi bằng chuyển nhiễm. Phôi được bơm CRISPR là phôi IVF được bơm ở giai đoạn 1 của tế bào với ARN dẫn hướng *CD1D* với ARN CAS9.

† MU dùng để chỉ tế bào trứng lợn nái con được hút và hoàn thiện ở University of Missouri như được mô tả trong phần IVF thuộc phần Vật liệu và phương pháp. ART dùng để chỉ tế bào trứng lợn nái được mua và hoàn thiện như được mô tả trong phần SCNT thuộc phần Vật liệu và phương pháp.

42443

102/283

Bảng 10. Kiểu gen và dự báo dịch mã đối với lợn được biến đổi CD163. Một số lợn chứa kiểu biến đổi hai alen, nhưng chỉ có một alen được mô tả và một alen được biến đổi khác không được khuếch đại bởi PCR.

Lứa	Số lợn con	Cơ chế sửa chữa	Loại	Kích thước các INDEL	Mô tả	Sự dịch mã protein *	Condon kết thúc sớm	Tham chiếu đến SEQ ID NO: 47	SEQ ID số†
63 & 64	7	NHEJ	Hai alen	Đoạn khuyết 1506 bp Alen khác	đoạn khuyết 30 bp trong exon 7 không được mô tả, không thể khuếch đại	KO hoặc $CD163^{\Delta 422-527}$	Không	Khuyết đoạn từ nt 1525 đến nt 3030	98
65	3	NHEJ	Hai alen	Đoạn cài xen 7 bp	Cài xen vào exon 7	KO	Có (491)	Cài xen giữa nt 3148 & 3149 [a]	99
65	2	NHEJ	Hai alen	Đoạn khuyết 503 bp Alen khác	Khuyết một phần exon 7 và 8 không được mô tả	KO	Có (491)	**	**
65	2	NHEJ	Hai alen	Đoạn khuyết 1280 bp Đoạn khuyết 1373 bp	Khuyết hoàn toàn exon 7 và 8 Khuyết hoàn toàn exon 7 và 8	$CD163^{\Delta 422-631}$ $CD163^{\Delta 422-631}$	Không Không	Khuyết đoạn từ nt 2818 đến nt 4097 Khuyết đoạn từ nt 2724 đến nt 4096	100 101
66	1	NHEJ	Đồng hợp tử	Đoạn khuyết 2015 bp	Sự cài xen của bộ khung vectơ hướng đích vào exon 7			**	**

42443

67-1	1	NHEJ	Hai alen	Đoạn khuyết 11 bp Đoạn khuyết 2 bp, đoạn khuyết 377 bp trong intron 6	Khuyết đoạn trong exon 7 Cài xen trong exon 7	KO	Có (485)	Khuyết đoạn từ nt 3137 đến nt 3147 Cải xen 2 bp giữa nt 3149 & nt 3150 [b] với khuyết đoạn 377 bp từ nt 2573 đến nt 2949	102 103
67-2	1	NHEJ	Hai alen	Đoạn khuyết 124 bp Đoạn khuyết 123 bp	Khuyết đoạn trong exon 7 Khuyết đoạn trong exon 7	KO CD163$^{\Delta 429-470}$	Có (464) Không	Khuyết đoạn từ nt 3024 đến nt 3147 Khuyết đoạn từ nt 3024 đến nt 3146	104 105
67-3	1	NHEJ	Hai alen	Đoạn cài xen 1 bp Alen khác	Cài xen vào exon 7 Không được mô tả, không thể khuếch đại	KO	Có (489)	Cài xen giữa nt 3147 & 3148[c]	106
67-4	1	NHEJ	Hai alen	Đoạn khuyết 130 bp	Khuyết đoạn trong exon 7	KO	Có (462)	Khuyết đoạn từ nt 3030 đến nt 3159	107 108

42443

				Đoạn khuyết 132 bp	Khuyết đoạn trong exon 7	CD163$^{\Delta 430-474}$	Không	Khuyết đoạn từ nt 3030 đến nt 3161	
68 & 69	6	NHEJ	Hai alen	Đoạn khuyết 1467 bp Alen khác	Khuyết hoàn toàn exon 7 và 8 Không được mô tả, không thể khuếch đại	CD163$^{\Delta 422-631}$	Không	Khuyết đoạn từ nt 2431 đến nt 3897	109
68 & 69	2	NHEJ	Hai alen	Đoạn khuyết 129 bp, Đoạn khuyết 1930 bp trong intron 6 Alen khác	Khuyết đoạn trong exon 7 Không được mô tả, không thể khuếch đại	CD163$^{\Delta 435-478}$	Không	Khuyết đoạn từ nt 488 đến nt 2417 trong exon 6, trình tự đã khuyết đoạn được thay thế bằng đoạn cài xen 12 bp [d] bắt đầu ở nt 488 & khuyết đoạn 129 bp nữa từ nt 3044 đến nt 3172	110

42443

65 & 69	3	WT			Lợn kiểu hoang được tạo ra từ khuẩn lạc kết hợp			SEQ ID NO: 47	47
70	2	NHEJ	On *SIGLEC1* -/- Hai alen	Đoạn khuyết 28 bp Đoạn khuyết 1387 bp	Khuyết đoạn trong exon 7 Khuyết đoạn một phần trong exon 7 và tất cả exon 8	KO KO	Có (528) Không	Khuyết đoạn từ nt 3145 đến nt 3172 Khuyết đoạn từ nt 3145 đến nt 4531	111 112
73	4	NHEJ	On *SIGLEC1* -/- Hai alen	Đoạn khuyết 1382 bp + đoạn cài xen 11 bp Đoạn khuyết 1720 bp	Khuyết đoạn một phần trong exon 7 và tất cả exon 8 Khuyết hoàn toàn exon 7 và 8	$CD163^{\Delta 422-631/KO}$	Không	Khuyết đoạn từ nt 3113 đến nt 4494, trình tự đã khuyết đoạn được thay thế bằng đoạn cài xen 11 bp[e] bắt đầu ở nt 3113 Khuyết đoạn từ nt 2440 đến nt 4160	113 114

*KO, knock-out – bất hoạt gen

** Không được bao hàm do lợn con được làm cho chết không đau đớn.

† SEQ ID số. trong cột này dùng để chỉ các SEQ ID NO. đối với các trình tự mà thể hiện INDEL so với SEQ ID NO: 47.

42443

106/283

[a] Trình tự được cài xen là TACTACT (SEQ ID NO: 115)

[b] Trình tự được cài xen là AG.

[c] Trình tự được cài xen là một gốc adenin (A) duy nhất.

[d] Trình tự được cài xen là TGTGGAGAATTC (SEQ ID NO:116).

[e] Trình tự được cài xen là AGCCAGCGTGC (SEQ ID NO: 117).

42443

Bảng 11. Kiểu gen và dự báo dịch mã đối với lợn được biến đổi *CD1D*

Lứa	**Số Lợn con**	**Cơ chế sửa chữa**	**Loại**	**Kích thước INDEL**	**Mô tả**	**Sự dịch mã protein**
158, 159	11	NHEJ	đồng hợp tử	Đoạn khuyết 1653 bp	Sự khuyết đoạn của exon 3, 4 và 5	KO*
167	2	NHEJ	đồng hợp tử	Đoạn khuyết 1265 bp	Sự khuyết đoạn của exon 5 và 72 bp của exon 6	KO
166-1	1	NHEJ	Hai alen	Đoạn khuyết 24 bp \	Sự loại bỏ codon mở đầu trong exon 3	KO
				Đoạn khuyết 27 bp \	Sự phá vỡ codon mở đầu trong exon 3	
				Đoạn khuyết 362 bp + 5 bp	Sự khuyết đoạn của exon 3	
166-2	1	NHEJ	Hai alen	Đoạn cài xen 6 bp +đoạn bắt cặp sai 2 bp	Sự bổ sung 6 bp trước codon mở đầu trong exon 3	$CD1D^{ko/+}$
				Đoạn khuyết 1598 bp	Sự loại bỏ codon mở đầu trong exon 3 và sự khuyết đoạn của các exon 4,5	
166-3	1	NHEJ	Hai alen	Đoạn cài xen 1 bp	Sự bổ sung G/T trong exon 3 trước codon mở đầu trong exon 3	$CD1D^{+/+}$
166-4	1	NHEJ	đồng hợp tử	Đoạn cài xen 1 bp	Sự bổ sung A trong exon 3 trước codon mở đầu trong exon 3	$CD1D^{+/+}$

*KO, bất hoạt gen

42443

Hiệu lực của hệ CRISPR/Cas9 trong các hợp tử của lợn

Dựa trên sự đứt gãy hướng đích của CD163 và CD1D trong các tế bào sinh dưỡng sử dụng hệ CRISPR/Cas9, phương pháp này được áp dụng cho việc tạo ra phôi lợn. Trước tiên, hiệu lực của hệ CRISPR/Cas9 trong việc phát triển phôi được đánh giá. Hệ CRISPR/Cas9 hướng đích eGFP được đưa vào trong các hợp tử đã được thụ tinh với tinh dịch từ lợn lòi đực dị hợp tử đối với gen chuyển eGFP. Sau khi bơm, phôi sau đó biểu hiện eGFP được theo dõi. Các nồng độ khác nhau của hệ CRISPR/Cas9 được thử nghiệm và tính độc tế bào của hệ phân phối CRISPR/Cas9 được quan sát (Fig. 4, ô A); sự phát triển phôi sau khi bơm CRISPR/Cas9 là kém hơn so với mẫu đối chứng. Tuy nhiên, tất cả các nồng độ của CRISPR/Cas9 mà được thử nghiệm đều hữu hiệu trong việc tạo ra sự biến đổi eGFP do không tìm thấy phôi với sự biểu hiện eGFP trong nhóm được bơm CRISPR/Cas9 (Fig. 4, ô B); trong số phôi đối chứng không được bơm, 67,7% là còn sống, điều đó cho thấy sự biểu hiện của eGFP. Khi các túi phôi riêng rẽ được xác định kiểu gen, có thể nhận biết đột biến nhỏ gần các vị trí liên kết CRISPR (Fig. 4, ô C). Dựa trên tính độc và hiệu quả, 10 ng/µl gARN và Cas9 mARN được sử dụng cho các thí nghiệm tiếp theo.

Khi các hợp phần CRISPR/Cas9 được thiết kế để hướng đích CD163 được đưa vào trong các hợp tử dự kiến, sự biến đổi hướng đích của các gen trong các túi phôi sau đó được quan sát. Khi các túi phôi riêng rẽ được xác định kiểu gen đối với đột biến CD163, các đột biến đặc hiệu đã được tìm thấy trong tất cả các phôi (Hiệu suất GE 100%). Quan trọng hơn là, trong khi phôi có thể được phát hiện với các biến đổi đồng hợp tử hoặc hai alen (8/18 và 3/18, tương ứng) (Fig. 5), kiểu gen khảm (biến đổi một alen) cũng được phát hiện (4/18 phôi). Một số phôi (8/10) từ phần gom được bơm bằng 2 ng/µl Cas9 và 10 ng/µl CRISPR và không nhận thấy sự khác biệt về hiệu quả sinh đột biến. Tiếp theo, căn cứ vào các kết quả *in vitro*, hai CRISPR biểu hiện gARN khác nhau được đưa vào để phá vỡ CD163 hoặc CD1D trong quá trình tạo phôi để tạo ra đoạn khuyết đặc hiệu của các gen đích. Kết quả là, có thể tạo ra một cách thành công đoạn khuyết đã được thiết kế của CD163 và CD1D bằng cách đưa vào hai đoạn dẫn. Đoạn khuyết đã được thiết kế được định nghĩa là đoạn khuyết mà loại bỏ trình tự hệ gen giữa hai đoạn dẫn được đưa vào. Trong số các phôi nhận hai CRISPR hướng đích CD163, toàn bộ nhưng ngoại trừ một phôi tạo ra sự sửa đổi hướng đích của CD163. Ngoài ra, 5/13 phôi được nhận thấy là có đoạn khuyết đã được thiết kế trên CD163 (Fig. 6, ô A)

và 10/13 phôi dường như có sự biến đổi CD163 theo kiểu đồng hợp tử hoặc hai alen. Việc hướng đích CD1D với hai CRISPR cũng hữu hiệu do tất cả các phôi (23/23) chỉ có thể được phát hiện trên hai phôi (2/23) (Fig. 6, ô B). 5 trong số 23 phôi có kiểu gen khảm cũng được phát hiện, nhưng các phôi còn lại có biến đổi đồng hợp tử hoặc biến đổi hai alen của CD1D. Cuối cùng, nhiều gen có thể được hướng đích bởi hệ CRISPR/Cas9, trong cùng một phôi được thử nghiệm. Nhằm mục đích này, tiến hành hướng đích cả CD163 và eGFP trong các hợp tử mà đã được thụ tinh với tinh dịch eGFP dị hợp tử. Khi các túi phôi từ phôi đã được bơm được xác định kiểu gen đối với CD163 và eGFP, nhận thấy rằng CD163 và eGFP đã được hướng đích thành công trong quá trình tạo phôi. Kết quả trình tự đã chứng tỏ rằng nhiều gen có thể được hướng đích bằng cách đưa nhiều CRISPR vào Cas9 (Fig. 6, ô C).

Tạo ra các thể đột biến CD163 và CD1D từ các hợp tử được bơm CRISPR/ Cas9

Dựa trên thành công từ nghiên cứu *in vitro* trước đây, một số hợp tử được bơm CRISPR/Cas9 đã được tạo ra và 46–55 túi phôi được chuyển trên một cá thể nhận (do số này được nhận thấy là hữu hiệu trong việc tạo ra lợn từ phôi có nguồn gốc *in vitro*). Bốn lần chuyển phôi được thực hiện, hai lần đối với mỗi CD163 và CD1D, và thu được một lần thụ thai cho mỗi một biến đổi. Bốn con lợn con khỏe mạnh được tạo ra mang các biến đổi trên CD163 (Bảng 8). Tất cả các con lợn con, lứa 67 từ lợn nái nhận có ID O083 thể hiện biến đổi đồng hợp tử hoặc biến đổi đổi hai alen của CD163 (Fig. 7). Hai con lợn con thể hiện đoạn khuyết đã được thiết kế của CD163 bởi hai CRISPR đã được phân phối. Tất cả lợn con đều khỏe mạnh. Đối với CD1D, một lần thụ thai cũng tạo ra bốn lợn con (lứa 166 từ lợn nái nhận có số nhận dạng O165): một lợn cái và ba lợn đực (Bảng 9). Một con lợn con (166-1) mang một đột biến thể khảm của CD1D, bao gồm đoạn khuyết 362 bp mà loại bỏ hoàn toàn exon 3 chứa codon mở đầu (Fig. 8). Một con lợn con chứa đoạn cài xen 6 bp với đoạn bắt cặp sai 2 bp trên một alen với một đoạn khuyết lớn trên alen còn lại. Hai con lợn con khác có đoạn cài xen 1 bp duy nhất trên hai alen. Không có đột biến dạng khảm nào được phát hiện đối với CD163.

Thảo luận

Việc gia tăng hiệu quả tạo ra lợn GE có thể tác động lớn đến việc tạo ra nhiều lợn GE hơn cho nông nghiệp và y sinh. Dữ liệu mô tả ở trên chỉ ra rằng bằng cách sử dụng hệ CRISPR/Cas9, lợn GE với đột biến đặc hiệu có thể được tạo ra ở hiệu suất cao.

Hệ CRISPR/Cas9 đã được ứng dụng thành công để sửa đổi gen trong cả tế bào sinh dưỡng và trong phôi trước khi cấy ghép.

Khi hệ CRISPR/Cas9 được đưa vào trong các tế bào sinh dưỡng, nó gây ra một cách thành công sự phá vỡ hướng đích gen đích bởi NHEJ nhưng không làm tăng khả năng hướng đích bởi HR. Hiệu quả hướng đích của CRISPR/Cas9 riêng rẽ trong các tế bào sinh dưỡng là thay đổi, điều này chỉ ra rằng thiết kế của đoạn dẫn có thể ảnh hưởng đến hiệu quả hướng đích. Cụ thể là, không thể nhận thấy sự biến đổi hướng đích của CD1D khi CRISPR 5350 và Cas9 được đưa vào trong các tế bào sinh dưỡng. Điều này gợi ý rằng có thể có lợi nếu thiết kế nhiều gARN và hợp lý hóa hiệu quả của chúng trước khi đưa vào lợn. Vẫn chưa rõ tại sao không có sự sửa chữa hướng đến HR với sự có mặt của ADN cho. Sau khi sàng lọc 886 khuẩn lạc (cả CD163 và CD1D) được chuyển nhiễm với CRISPR và ADN cho, chỉ có một khuẩn lạc có bằng chứng về sự kiện HR một phần. Kết quả này đã chứng tỏ rằng hệ CRISPR/Cas9 hoạt động với ADN cho được đưa vào để tạo ra các đoạn khuyết lớn không mong muốn trên gen đích nhưng đã không làm tăng hiệu quả HR đối với hai vectơ hướng đích cụ thể này. Tuy nhiên, vẫn chưa biết cơ chế đặc hiệu để quan sát đoạn khuyết lớn này. Các báo cáo trước đây từ nhóm tác giả sáng chế đã gợi ý rằng ADN cho có thể được sử dụng một cách hữu hiệu với ZFN để tạo ra sự sửa chữa hướng đến HR. Tương tự, đã nhận thấy sự gia tăng hiệu quả hướng đích khi ADN cho được sử dụng với hệ CRISPR/ Cas9, nhưng không quan sát thấy sự sửa chữa hướng đến HR hoàn toàn. Trong một nghiên cứu trước đây sử dụng ZFN, đã thấy rằng sự biến đổi hướng đích có thể xảy ra nhờ sự kết hợp giữa HR và NHEJ do sự tái tổ hợp một phần của ADN cho được đưa vào được phát hiện sau khi DSB được tạo ra bởi ZFN. Có thể lý giải rằng các con đường HR và NHEJ không độc lập mà có thể hoạt động cùng nhau để hoàn thiện quá trình sửa chữa sau khi DSB được tạo ra bởi homing endonucleaza. Nồng độ cao hơn của các CRISPR có thể cải thiện hiệu quả hướng đích trong các tế bào sinh dưỡng mặc dù không nhận thấy sự khác biệt có nghĩa thống kê ở các kết quả thí nghiệm này. Điều này có thể gợi ý rằng CRISPR là yếu tố hạn chế trong hệ CRISPR/Cas9, nhưng vẫn cần phải xác nhận thêm nữa. Các tế bào hướng đích được sử dụng một cách thành công để tạo ra lợn GE nhờ SCNT, cho thấy việc sử dụng hệ CRISPR/Cas9 không ảnh hưởng đến khả năng tạo dòng các tế bào. Một vài con lợn con đã được làm chết không đau đớn do các vấn đề về sức khỏe; tuy nhiên, đây không phải là hiện tượng bất thường ở các con lợn con bắt nguồn từ SCNT.

Khi hệ CRISPR/Cas9 được đưa vào trong phôi đang phát triển bằng cách bơm hợp tử, gần 100% phôi và lợn chứa INDEL ở gen đích, chứng tỏ rằng kỹ thuật này là rất hữu hiệu trong quá trình tạo phôi. Hiệu quả quan sát được trong nghiên cứu này lớn hơn hiệu quả thông thường được công bố trong các nghiên cứu khác có sử dụng homing endonucleaza trong quá trình tạo phôi. Sự giảm số lượng phôi đạt đến giai đoạn túi phôi đã gợi ý rằng nồng độ của CRISPR/Cas9 được đưa vào nghiên cứu này có thể gây độc cho phôi. Việc tối ưu hóa hơn nữa hệ phân phối này có thể làm gia tăng khả năng sống sót của phôi và vì vậy cải thiện hiệu quả nói chung của quá trình. Tỷ lệ đột biến đạt gần 100% quan sát được ở đây khác với báo cáo trước đây trong quá trình bất hoạt qua trung gian CRISPR/Cas9 ở lợn; tuy nhiên, sự khác biệt về mặt hiệu quả giữa các nghiên cứu có thể là ở sự phối hợp giữa đoạn dẫn và đích đích được lựa chọn. Trong nghiên cứu này, nồng độ CRISPR/Cas9 thấp hơn (10 ng/µl mỗi loại) là hữu hiệu trong việc tạo ra các đột biến ở phôi đang phát triển và tạo ra lợn GE. Nồng độ này thấp hơn nồng độ được báo cáo trước đây trong các hợp tử của lợn (Cas9 125 ng/µl và CRISPR 12,5 ng/µl). Nồng độ của các hợp phần CRISPR/Cas9 thấp hơn có thể có lợi đối với phôi đang phát triển do việc đưa những lượng dư axit nucleic vào phôi đang phát triển có thể gây độc. Một số kiểu gen khảm đã được nhận thấy ở phôi được bơm CRISPR/Cas9 từ thử nghiệm *in vitro*; tuy nhiên, chỉ một con lợn con được tạo ra nhờ phương pháp này có kiểu gen khảm. Có khả năng là, việc bơm các hợp phần CRISPR/Cas9 vào có thể hữu hiệu hơn việc đưa các homing endonucleaza khác vào do kiểu gen khảm được xem là rào cản chính của việc sử dụng hệ CRISPR/Cas9 trong các hợp tử. Một lợi ích khác nữa của việc sử dụng hệ CRISPR/Cas9 đã được chứng minh bởi các kết quả này là ở chỗ không có con lợn bất hoạt CD163s nào được tạo từ các hợp tử có nguồn gốc IVF được bơm hệ CRISPR/Cas9 bị chết, trong khi một vài lợn con thu được từ SCNT được làm chết không đau đớn sau vài ngày. Điều này gợi ý rằng kỹ thuật này có thể không chỉ không cần đến SCNT trong việc tạo ra lợn bất hoạt mà cũng có thể khắc phục được các vấn đề về sức khỏe nói chung có liên quan đến SCNT. Khi việc bơm CRISPR/Cas9 mARN vào các hợp tử đã được tối ưu hóa, các thí nghiệm tiếp theo cũng sẽ bao gồm bước bơm đồng thời ADN cho.

Nghiên cứu hiện thời chứng tỏ rằng việc đưa hai CRISPR với Cas9 vào các hợp tử có thể gây ra sự xóa đoạn nhiễm sắc thể ở phôi đang phát triển và tạo ra lợn với đoạn khuyết dự định, nghĩa là, đoạn khuyết cụ thể giữa hai đoạn dẫn CRISPR. Các đoạn

khuyết đã được thiết kế này có thể có lợi do có thể định rõ kích thước của đoạn khuyết thay vì dựa vào các sự kiện ngẫu nhiên do NHEJ gây ra. Cụ thể là, nếu có sự cài xen /khuyết đoạn của các nucleotit trong bộ ba gây ra bởi homing endonucleaza, đột biến này có thể dẫn đến đột biến giảm liều do sự dịch chuyển khung sẽ không xảy ra. Tuy nhiên, bằng cách đưa vào hai CRISPR, có thể gây ra đoạn xóa lớn hơn mà nó sẽ có cơ hội lớn hơn trong việc tạo ra protein không chức năng. Thú vị là, các CD1D CRISPR đã được thiết kế trong hệ gen trên một diện tích lớn hơn so với CD163; có một khoảng cách 124 bp giữa CD163 CRISPR 10 và 131 trong khi có một khoảng cách 550 bp giữa CRISPR 4800 và 5350 đối với CD1D. Khoảng cách lớn hơn giữa các CRISPR không thực sự hiệu quả trong việc tạo ra đoạn khuyết như được chỉ ra trong nghiên cứu. Tuy nhiên, do nghiên cứu này chỉ bao gồm số lượng hữu hạn các nhận xét và cần phải xem xét tính hiệu quả của các CRISPR riêng rẽ, mà không được đề cập đến ở đây, cần có một nghiên khác nữa để xác nhận mối quan hệ giữa khoảng cách giữa các CRISPR và khả năng gây ra sự xóa đoạn mong đợi.

Hệ CRISPR/Cas9 cũng hữu hiệu trong việc hướng đích hai gen một cách đồng thời trong cùng một phôi chỉ với một bước thêm vào là đưa vào một CRISPR nữa với crARN. Điều này giải thích sự dễ dàng phá vỡ nhiều gen so với các homing endonucleaza khác. Các kết quả này gợi ý rằng kỹ thuật này có thể được sử dụng để hướng đích các cụm gen hoặc họ gen mà có thể có tác dụng bù trừ, vì vậy khó có thể xác định vai trò của các gen riêng rẽ trừ khi tất cả các gen đều bị phá vỡ. Kết quả này chứng tỏ rằng kỹ thuật CRISPR/Cas9 có thể được ứng dụng trong việc tạo ra lợn GE bằng cách gia tăng hiệu suất hướng đích gen trong các tế bào sinh dưỡng và bằng cách bơm hợp tử trực tiếp.

Ví dụ 2: Tính kháng PRRSV tăng ở lợn có trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163

Virut gây hội chứng hô hấp và sinh sản ở lợn (PRRSV) gây thiệt hại cho ngành công nghiệp lợn trong suốt một phần tư thế kỷ qua. Suy đoán về phương thức xâm nhập của virut bao gồm cả SIGLEC1 và CD163. Trong khi sự bất hoạt *SIGLEC1* không ảnh hưởng đến đáp ứng đối với thử thách virut, trong ví dụ này đã chỉ ra rằng động vật không mang *CD163* không thể hiện các dấu hiệu lâm sàng của sự nhiễm trùng, bệnh ở phổi, virut huyết hoặc sản xuất kháng thể mà đều là dấu hiệu của sự nhiễm trùng PRRSV. Không chỉ có chất trung gian xâm nhập PRRSV được xác nhận; mà nếu động

vật được tạo ra theo cách tương tự được phép đi vào chuỗi cung cấp thực phẩm, chiến lược phòng ngừa thiệt hại đáng kể về mặt kinh tế và sự chịu đựng của động vật được mô tả.

Vật liệu và phương pháp

Xác định kiểu gen

Việc xác định kiểu gen được dựa trên cả phương pháp giải trình tự ADN và phương pháp giải trình tự mARN. Kiểu gen của con đực giống có đoạn khuyết 11 bp trong một alen mà khi dịch mã 45 axit amin được dự báo trước thành miền 5, tạo ra codon kết thúc sớm ở axit amin 64. Trong alen còn lại, có sự thêm đoạn 2 bp vào exon 7 và khuyết đoạn 377 bp trong intron trước exon 7, mà khi dịch mã 49 axit amin đầu tiên của miền 5, tạo ra codon kết thúc sớm ở axit amin 85. Một con lợn nái có sự thêm đoạn 7 bp trong một alen mà khi dịch mã 48 axit amin đầu tiên của miền 5, tạo ra codon kết thúc sớm ở axit amin 70. Alen khác không được mô tả (A), do không có băng nào từ exon 7 theo PCR hoặc PCR khoảng dài 6.3 kb. 3 con lợn nái khác được tạo dòng và có đoạn khuyết 129 bp trong exon 7 mà được dự đoán là tạo ra đoạn khuyết gồm 43 axit amin từ miền 5. Alen khác không được mô tả (B).

Sự sinh trưởng của PRRSV trong môi trường nuôi cấy và sản xuất chất chủng ngừa virut để lây nhiễm cho lợn được bao hàm trong ứng dụng 973 đã được phê chuẩn bởi IBC

Chủng PRRSV, thể phân lập NVSL 97-7895 (GenBank # AF325691 2001-02-11), được sinh trưởng như đã được mô tả trong quy trình 973 đã được IBC phê chuẩn. Thể phân lập thí nghiệm này đã được sử dụng trong các nghiên cứu thí nghiệm trong khoảng 20 năm (Ladinig et al., 2015). Thể phân lập thứ hai được sử dụng cho thử nghiệm lần thứ 2, KS06-72109 như đã mô tả trước đây (Prather et al., 2013).

Lây nhiễm lợn với PRRSV

Quy trình lây nhiễm tiêu chuẩn đối với PRRSV được sử dụng để lây nhiễm cho lợn. Lợn con 3 tuần tuổi được chủng ngừa với khoảng 10^4 TCID50 virut PRRS được dùng qua đường trong cơ (IM) và trong mũi (IN). Lợn được theo dõi hằng ngày và những con lợn biểu hiện triệu chứng ốm được điều trị theo khuyến cáo của bác sĩ thú y CMG. Lợn mà thể hiện sự kiệt sức nghiêm trọng và có nguy cơ ngừng chống chọi với sự lây nhiễm được làm chết không đau đớn một cách nhân đạo và các mẫu được thu

nhận. Các nhân viên và bác sĩ thú y đều không thấy được tình trạng di truyền của lợn để hạn chế sự sai lệch trong đánh giá hoặc điều trị. PRRSV có mặt trong dịch thể trong quá trình lây nhiễm; Do đó, các mẫu máu được thu nhận và được bảo quản ở –80°C cho đến khi được phân tích nhằm xác định lượng hoặc mức độ nhiễm virut huyết trong mỗi con lợn. Ở cuối thí nghiệm, lợn được cân và được làm cho chết không đau đớn một cách nhân đạo, và các mô được thu nhận và cố định trong formalin đệm 10%, được gắn trên parafin, và được xử lý mô bệnh học bởi các nhà nghiên cứu mô bệnh học được công nhận rộng rãi.

Cho điểm kiểu hình của lợn đã được thử thách

Kiểu hình của lợn được cho điểm mù hằng ngày như sau: Thái độ của lợn như thế nào? Điểm thái độ: 0: BAR, 1: QAR, 2: Suy nhược không đáng kể, 3: Suy nhược, 4: Hấp hối. Tình trạng cơ thể của lợn như thế nào? Điểm tình trạng: 1: Hốc hác, 2: Gầy yếu, 3: Lý tưởng, 4: Béo, 5: Quá béo/Béo phì. Nhiệt độ trực tràng của lợn như thế nào? Thân nhiệt bình thường 101,6–103,6°F (≥ 104°F được xem là sốt). Có đi khập khiễng không (mức độ)? Chân như thế nào? Đánh giá các chân về việc sung khớp và thương tổn ở móng (kiểm tra dưới móng và các mặt bên của móng). Điểm khập khiễng: 1: Không khấp khiễng, 2: Không đều một chút khi đi bộ, xuất hiện sự cứng ở một số khớp nhưng không đi khập khiễng, 3: Đi khập khiễng nhẹ, lê nhẹ trong khi đi bộ, 4: khập khiễng vừa phải, lê rõ rệt bao gồm khập khiễng có liên quan đến ngón chân, 5: khập khiễng rõ rệt, không gánh trọng lượng cơ thể trên chân, cần độn viên để đứng lên/đi lại. Có sự khó khăn nào về hô hấp không (mức độ)? Có sự thở há mồm không? Có dịch tiết ra ở mũi không (màu của dịch tiết ra, lượng dịch tiết ra: ít/vừa phải/nghiêm trọng)? Có nghe thấy động vật ho không? Có dịch tiết ra ở mắt không? Điểm hô hấp: 0: Bình thường, 1: khó thở nhẹ và/hoặc thở gấp khi căng thẳng (khi bị điều khiển), 2: khó thở nhẹ và/hoặc thở gấp khi nghỉ ngơi, 3: khó thở vừa phải và/hoặc thở gấp khi căng thẳng (khi bị điều khiển), 4: khó thở vừa phải và/hoặc thở gấp khi nghỉ ngơi, 5: khó thở nghiêm trọng và/hoặc thở gấp khi căng thẳng (khi bị điều khiển), 6: khó thở nghiêm trọng và/hoặc thở gấp khi nghỉ ngơi. Có dấu hiệu của bệnh tiêu chảy (mức độ) hoặc nôn? Có lẫn mãu hoặc chất nhầy không? Điểm tiêu chảy: 0: Không thấy có phân, 1: phân rắn bình thường, 2: phân mềm nhưng thành khuôn (độ đặc sệt và mượt, tạo ra phân giống như của bò), 3: tiêu chảy dạng lỏng có màu nâu/nâu vàng với các hạt phân, 4: tiêu chảy

dạng lỏng có màu nâu/nâu vàng không chứa các hạt phân, 5: tiêu chảy dạng lỏng xuất hiện giống như nước.

Hệ thống cho điểm này được phát triển bởi Dr. Megan Niederwerder ở KSU và được dựa trên các công bố sau (Halbur et al., 1995; Merck; Miao et al., 2009; Patience and Thacker, 1989; Winckler and Willen, 2001). Điểm và nhiệt độ được phân tích bằng cách sử dụng ANOVA được tách riêng dựa trên kiểu gen khi điều trị.

Xác định lượng virut PRRSV nhiễm trong máu

Lượng virut huyết được xác định bằng hai phương pháp. Độ chuẩn virut được tiến hành bằng cách bổ sung dịch pha loãng theo dãy nồng độ 1:10 của huyết thanh để nhập dòng tế bào MARC-145 vào đĩa 96 lỗ. Huyết thanh được pha loãng trong môi trường thiết yếu tối thiểu của Eagle có bổ sung 8% huyết thanh bào thai bò, penicillin, streptomycin, và amphotericin B như đã mô tả trước đây (Prather et al., 2013). Tế bào được kiểm tra sự có mặt của tác dụng gây bệnh tế bào sau 4 ngày ủ bằng cách sử dụng kính hiển vi. Dịch pha loãng ở mức cao nhất thể hiện tác dụng gây bệnh tế bào được cho điểm là điểm kết thúc chuẩn độ. ARN tổng được phân lập từ huyết thanh bằng cách sử dụng kit phân lập ARN virus Life Technologies MagMAX-96 để xác định axit nucleic của virut. Phản ứng chuỗi polymeraza phiên mã ngược được tiến hành bằng cách sử dụng kit EZ-PRRSV MPX 4.0 từ Tetracore trên hệ PCR thời gian thực CFX-96 (Bio-Rad) theo hướng dẫn của nhà sản xuất. Mỗi phản ứng (25 µl) chứa ARN từ 5,8 µl huyết thanh. Đường cong tiêu chuẩn được xây dựng bằng cách chuẩn bị các dịch pha loãng theo dãy nồng độ của mẫu đối chứng ARN được cung cấp trong kit (Tetracore). Số khuôn trên PCR được báo cáo.

Nhuộm SIGLEC1 và CD163 của tế bào PAM

Các đại thực bào phế nang lợn (PAM) được thu nhận bằng cách cắt phổi và và lấp đầy phổi bằng ~100 ml nước muối đệm phosphat lạnh. Sau khi thu hồi dung dịch rửa muối đệm phosphat, tế bào được kết viên và tái tạo huyền phù trong 5 ml dung dịch muối đệm phosphat lạnh và được bảo quản trên đá. Khoảng 10^7 PAM được ủ trong 5 ml kháng thể khác nhau (CD169 kháng lợn (dòng 3B11/11; AbD Serotec); CD163 kháng lợn (dòng 2A10/11; AbD Serotec)) được pha loãng trong dung dịch muối đệm phosphat với 5% huyết thanh bào thai bò và 0,1% natri azit trong 30 phút trên đá. Tế

bào được rửa và tái tạo huyền phù trong dịch pha loãng 1/100 của isothioxyanat huỳnh quang (FITC)-được tiếp hợp với IgG dê kháng chuột (life Technologies) được pha loãng trong dung dịch đệm nhuộm màu và được ủ trong 30 phút trên đá. Ít nhất 10^4 tế bào được phân tích bằng cách sử dụng thiết bị đếm tế bào trong dòng chảy FACSCalibur và phần mềm Cell Quest (Becton Dickinson).

Xác định Ig đặc hiệu PRRSV

Để xác định Ig đặc hiệu PRRSV, PRRSV N protein tái tổ hợp được biểu hiện ở vi khuẩn (Trible et al., 2012) và được tiếp hợp với các hạt Luminex từ tính bằng cách sử dụng kit (Luminex Corporation). Các hạt kết hợp N protein được pha loãng trong dung dịch muối đệm phosphat chứa 10% huyết thanh dê thành 2.500 hạt/50 µl và được đặt vào các lỗ thuộc đĩa polystyren đáy tròn 96 lỗ. Huyết thanh được pha loãng 1:400 trong dung dịch muối đệm phosphat chứa 10% huyết thanh dê và 50 µl được bổ sung vào các lỗ giống hệt và được ủ trong 30 phút đồng thời lắc nhẹ ở nhiệt độ trong phòng. Tiếp theo, đĩa được rửa (3X) bằng dung dịch muối đệm phosphat chứa 10% huyết thanh dê và 50 µl kháng thể thứ cấp dê kháng lợn được tinh chế về mặt ái lực được tiếp hợp với biotin-SP (IgG, Jackson ImmunoResearch) hoặc IgM dê kháng lợn được tinh chế về mặt ái lực có gắn nhãn biotin (KPL) được pha loãng thành 2 µg/ml trong dung dịch muối đệm phosphat chứa 10% huyết thanh dê được bổ sung vào. Đĩa được rửa (3X) sau 30 phút ủ và sau đó 50 µl phycoerythrin được tiếp hợp với streptavidin (2 µg/ml (Moss, Inc.) trong dung dịch muối đệm phosphat chứa 10% huyết thanh dê) được bổ sung. Đĩa được rửa 30 phút sau đó và các hạt nhỏ hình cầu được tái tạo huyền phù trong 100 µl dung dịch muối đệm phosphat chứa 10% huyết thanh dê và được phân tích bằng cách sử dụng MAGPIX và phần mềm Luminex xPONENT 4.2. Cường độ huỳnh quang trung bình (MFI) được báo cáo.

KẾT QUẢ

Đột biến ở *CD163* được tạo ra bằng cách sử dụng kỹ thuật CRISPR/Cas9 như được mô tả ở trên trong ví dụ 1. Một số động vật chuyển gen được tạo ra từ quá trình bơm hợp tử và từ quá trình chuyển nhân từ tế bào sinh dưỡng. Một vài trong số những con tạo thành này được giao phối tạo thành con cháu để nghiên cứu. Một con đực tạo thành được giao phối với các con cái với hai kiểu gen. Con đực tạo thành (67-1) có sự

khuyết đoạn 11 bp trong exon 7 trên một alen và sự bổ sung 2 bp trong exon 7 (và sự khuyết đoạn 377 bp trong intron đứng trước) của một alen khác và được dự báo là động vật khuyết thiểu gen (*CD163*$^{-/-}$). Một con cái tạo thành (65-1) có sự bổ sung 7 bp trong exon 7 trong một alen và một alen tương ứng không được mô tả và vì vậy được dự báo là dị hợp tử đối với knockout (*CD163*$^{-/?}$). Một kiểu gen thứ hai của con cái được tạo ra (3 động vật mà được tạo dòng) chứa một alen cũng không được mô tả và một alen với sự khuyết đoạn 129 bp trong exon 7. Đoạn khuyết này được dự báo tạo ra đoạn khuyết gồm 43 axit amin trong miền 5. Sự giao phối giữa các động vật này tạo ra tất cả lợn con mà nhận alen không từ lợn lòi đực và đoạn khuyết gồm 43 axit amin hoặc một trong các alen không được mô tả từ lợn nái. Ngoài lợn con kiểu hoang mà dùng làm mẫu đối chứng dương tính để thử thách virut, quy trình này tạo ra 4 kiểu gen khác nữa (bảng 8).

Bảng 12. Các kiểu gen được thử nghiệm tính kháng đối với thử thách PRRSV (các chủng NVSL và KS06)

alen		**Tính kháng đối với thử thách PRRSV như được xác định bởi sự nhiễm virut huyết**	
Bố	Mẹ	NVSL	KS06
Không	Không	Kháng	Không xác định
Không	Δ43 axit amin	Không xác định	Kháng
Không	Không được mô tả A	Dễ bị mắc	Không xác định
Không	Không được mô tả B	Dễ bị mắc	Dễ bị mắc
Kiểu hoang	Kiểu hoang	Dễ bị mắc	Dễ bị mắc

Ở lợn con được chỉnh sửa gen mới cai sữa và lợn con phù hợp với tuổi kiểu hoang dại được đưa đến Kansas State University để thử thách PRRSV. Thử thách PRRSV được tiến hành như đã được mô tả trước đây (Prather et al., 2013). Các con lợn con, ở ba tuần tuổi, được đem đến cơ sở thử thách và được duy trì dưới dạng nhóm đơn. Tất cả các thử nghiệm được bắt đầu sau khi được chấp thuận bởi các ủy ban an toàn và sử dụng động vật có tổ chức. Sau khi thích nghi, các con lợn được thử thách bằng thể phân lập PRRSV, NVSL 97-7895 (Ladinig et al., 2015), được truyền trên các tế bào MARC-145 (Kim et al., 1993). Các con lợn được thử thách bằng khoảng 10^5 $TCID_{50}$

virut. Một nửa thuốc chủng được phân phát trong cơ và phần còn lại được phân phát trong mũi. Tất cả các con lợn được gây nhiễm được duy trì trong một nhóm duy nhất, điều này cho phép sự tiếp xúc liên tục với virut từ các con cùng chuồng bị nhiễm. Các mẫu máu được thu gom ở các ngày khác nhau cho tới 35 ngày sau khi nhiễm và đến khi kết thúc, ngày 35. Các con lợn được mổ và mô được cố định trong dung dịch formalin được đệm 10%, nhúng trong parafin và được xử lý về mô bệnh học. Các dấu hiệu lâm sàng kết hợp với PRRSV được ghi lại trong quá trình gây nhiễm bao gồm cả suy hô hấp, ăn không ngon, hôn mê và sốt. Kết quả về các dấu hiệu lâm sàng trong thời gian thử nghiệm được tổng kết trên Fig.9. Như mong đợi, các con lợn Wild Type kiểu hoang dại (*CD163*+/+) thể hiện các dấu hiệu sớm về sự nhiễm trùng PRRSV, các dấu hiệu này đạt đỉnh điểm ở các ngày từ ngày 5 đến ngày 14 và dai dẳng trong nhóm trong suốt phần còn lại của quá trình thử nghiệm. Tỷ lệ % các con lợn bị sốt đạt đỉnh vào khoảng ngày 10. Ngược lại, các con lợn Vô hiệu (*CD163*-/-) không thể hiện bằng chứng về các dấu hiệu lâm sàng trong toàn bộ thời gian thử nghiệm. Các dấu hiệu hô hấp trong quá trình nhiễm trùng PRRSV cấp được phản ánh ở sự thay đổi mô bệnh học đáng kể ở phổi (Bảng 9). Sự nhiễm trùng của lợn kiểu hoang dại thể hiện mô bệnh học phù hợp với PRRS bao gồm phù kẽ với sự thâm nhiễm tế bào đơn nhân (Fig.10). Ngược lại, không có dấu hiệu về sự thay đổi ở phổi ở các con lợn Vô hiệu (*CD163*-/-). Cỡ mẫu đối với các kiểu gen khác nhau là nhỏ; tuy nhiên điểm số trung bình là 3,85 (n=7) đối với kiểu hoang dại, 1,75 (n=4) đối với A không được đặc trưng hóa, 1,33 (n=3) đối với B không được đặc trưng hóa, và 0 (n=3) và đối với Vô hiệu (CD163-/-).

Bảng 13. Đánh giá phổi qua kính hiển vi

Lợn	**Kiểu gen**	**Mô tả**	**Điểm**
41	Kiểu hoang dại	100% sung huyết. Các vùng đa ổ phù. Thâm nhiễm số lượng vừa phải các lympho bào và các đại thực bào	3
42	Kiểu hoang dại	100% sung huyết. Các vùng đa ổ phù. Thâm nhiễm số lượng vừa phải các lympho bào và các đại thực bào	3
47	Kiểu hoang dại	75% thâm nhiễm đa ổ các tế bào đơn nhân và phù nhẹ	2

50	Kiểu hoang dại	75% thâm nhiễm đa ổ các tế bào đơn nhân trong không gian túi phổi và xung quanh các mạch máu nhỏ phù quanh mạch	3
51	Kiểu hoang dại	25% chứng xẹp phổi với sự thâm nhiễm vừa phải các tế bào đơn nhân	1
52	Kiểu hoang dại	10% không gian túi phổi bị xẹp với sự thâm nhiễm của một số lượng nhỏ các tế bào đơn nhân	1
56	Kiểu hoang dại	100% thâm nhiễm khuếch tán vừa phải qua kẽ của các tế bào đơn nhân. Các vách gian phế nang bị làm đầy vừa phải bởi sự xuất huyết và phù.	4
45	A không được đặc trưng hóa	75% thâm nhiễm đa ổ của các tế bào đơn nhân, đặc biệt là xung quanh phế quản, các mạch máu, các khoang dưới màng phổi, và các vách gian phế nang.	3
49	A không được đặc trưng hóa	75% thâm nhiễm đa ổ từ vừa phải đến mức độ lớn các tế bào đơn nhân. Một số mạch máu bị phù nhẹ.	2
53	A không được đặc trưng hóa	10% thâm nhiễm đa ổ mức độ nhỏ các tế bào đơn nhân	1
57	A không được đặc trưng hóa	15% thâm nhiễm tế bào đơn nhân	1
46	B không được đặc trưng hóa	Viêm phổi kẽ mức độ vừa phải	2
48	B không được đặc trưng hóa	Phù quanh mạch và thâm nhiễm tế bào đơn nhân xung quanh các mạch cỡ nhỏ và trung bình và xung quanh các vách gian phế nang	2
54	B không được đặc trưng hóa	Không thay đổi	0
40	Vô hiệu	Không thay đổi	0
43	Vô hiệu	Không thay đổi	0
55	Vô hiệu	Không thay đổi	0

Các dấu hiệu lâm sàng đỉnh điểm tương ứng với các mức độ PRRSV trong máu. Việc đo axit nucleic virut được tiến hành bằng cách phân lập ARN tổng số từ huyết thanh sau đó khuếch đại ARN PRRSV bằng cách sử dụng thử nghiệm PCR PRRSV thời gian thực transcriptaza nghịch đảo có bán trên thị trường (Tetracore, Rockville, MD). Đường cong chuẩn được tạo ra bằng cách điều chế các dung dịch pha loãng theo dãy của đối chứng ARN PRRSV, được cung cấp trong kit RT-PCR và kết quả được chuẩn hóa dưới dạng các khuôn số trên 50 µl dung dịch phản ứng PCR. Thể phân lập PRRSV sau quá trình PRRSV virut huyết ở các con lợn kiểu hoang dại *CD163*+/+ (Fig.11). Chỉ số virut huyết là rõ ràng ở ngày bốn, đạt đến đỉnh điểm ở ngày 11 và suy giảm cho đến khi kết thúc thử nghiệm. Ngược lại, ARN virut không được phát hiện ở các con lợn *CD163*$^{-/-}$ ở thời điểm bất kỳ trong quá trình thử nghiệm. Phù hợp với chỉ số virut huyết, sự sản xuất kháng thể bởi các con lợn vô hiệu và các con lợn có alen không biểu thị đặc điểm được phát hiện ở khoảng ngày 14 và tăng đến ngày 28. Không có sự sản xuất kháng thể ở các con vật vô hiệu (Fig.12). Cùng nhau, các dữ liệu này thể hiện rằng lợn kiểu hoang dại hỗ trợ cho sự sao chép PRRSV với sự tạo ra các dấu hiệu lâm sàng phù hợp với PRRS. Ngược lại, các con lợn bất hoạt không sản sinh virut huyết và không có các dấu hiệu lâm sàng, ngay cả khi các con lợn này được chủng và tiếp xúc liên tục với các con cùng chuồng bị nhiễm.

Cuối quá trình thử nghiệm, các đại thực bào phế nang của lợn được lấy ra bằng cách rửa phổi và được nhuộm để biểu hiện trên bề mặt SIGLEC1 (CD169, dòng 3B11/11) và CD163 (dòng 2A10/11), như đã được mô tả (Prather et al., 2013). Mức độ biểu hiện tương đối cao của CD163 được phát hiện trên các con vật kiểu hoang dại *CD163*+/+ (Fig.13). Ngược lại, các con lợn *CD163*-/- chỉ thể hiện mức độ cơ bản vết kháng CD163, do đó xác nhận kiểu hình bất hoạt. Mức độ biểu hiện đối với dấu chuẩn đại thực bào khác CD169 là tương tự đối với cả lợn kiểu hoang dại và lợn bất hoạt (Fig.14). Các dấu chuẩn đại thực bào bề mặt khác, bao gồm MHC II và CD172 là giống nhau về cả hai kiểu gen (dữ liệu không được thể hiện).

Mặc dù cỡ mẫu là nhỏ nhưng các con lợn kiểu hoang dại có xu hướng tăng trọng ít hơn sau quá trình thử nghiệm (mức tăng trọng trung bình hàng ngày là 0,81 kg ± 0,33, n=7) so với các con lợn thuộc ba kiểu gen khác (A không được đặc trưng hóa 1,32 kg ± 0,17, n=4; B không được đặc trưng hóa 1,20 kg ± 0,16, n =3; vô hiệu 1,21 kg ± 0,16, n=3).

Trong thử nghiệm thứ hai, 6 kiểu hoang dại, 6 Δ43 axit amin, và 6 con lợn với alen không được đặc trưng hóa (B) được thử thách như được mô tả ở trên, ngoại trừ việc KS06-72109 được sử dụng để chủng cho các con lợn con. Tương tự như dữ liệu NVSL, các con lợn con kiểu hoang dại và B không được đặc trưng hóa đã phát triển virut huyết. Tuy nhiên, ở các con lợn Δ43 axit amin, KS06 không dẫn đến virut huyết (Fig.15; Bảng 7).

Gợi ý và kết luận

Bệnh liên quan nhất trên lâm sàng trong lĩnh vực công nghiệp chăn nuôi lợn là PRRS. Trong khi các chương trình tiêm chủng đã thành công trong việc ngăn chặn hoặc làm giảm hầu hết các tác nhân gây bệnh cho lợn, PRRSV đã chứng tỏ là một thử thách khó khăn. Ở đây CD163 được xác định là yếu tố trung gian xâm nhập đối với chủng virut này. Lợn đực què được tạo ra bằng cách tiêm CRISPR/Cas9 vào các hợp tử (Whitworth et al., 2014) và do đó không có gen chuyển. Ngoài ra, một trong số các alen từ lợn nái (cũng được tạo ra bằng cách sử dụng CRISPR/Cas9) không chứa chứa gen chuyển. Do đó, lợn con #40 mang đoạn bổ sung 7 bp trong một alen và đoạn xóa 11 bp trong alen khác, nhưng không mang gen chuyển. Các alen CD163 kháng virut này thể hiện sự chỉnh sửa hệ gen nhỏ với việc xem xét rằng hệ gen của lợn là khoảng 2,8 tỷ bp (Groenen et al., 2012). Nếu con vật được tạo ra một cách tương tự được đưa vào làm nguồn cung cấp thực phẩm, sự tổn thất kinh tế đáng kể sẽ được ngăn chặn.

Ví dụ 3: Tăng tính kháng đối với các virut PRRS và sinh sản ở lợn có kiểu gen 1 ở lợn có miền SRCR 5 CD163 được thay thế bằng miền SRCR 8 tương đồng giống CD163 của người

CD163 được xem như là thụ thể chính đối với virut gây hội chứng hô hấp và sinh sản ở lợn (PRRSV). Trong thử nghiệm này, các con lợn được biến đổi gen (GM) để có một trong số các kiểu gen sau: bất hoạt (knock out - KO) hoàn toàn CD163, xóa đoạn trong miền 5 giàu xystein của thụ thể dọn dẹp (scavenger receptor cysteine-rich - SRCR) của CD163, hoặc thay thế đoạn (tráo đổi miền) của miền SRCR 5 bằng một exon tổng hợp mã hóa vùng tương đồng của miền SRCR 8 (hCD163L1) giống như CD163 của người. Việc xác định kiểu hình miễn dịch của các đại thực bào phế nang của lợn (PAMs) chỉ ra rằng các con lợn KO hoặc xóa đoạn miền SRCR 5 không biểu hiện CD163 và PAMs không hỗ trợ sự nhiễm trùng PRRSV. PAMs từ các con lợn mà đã có vùng tương

đồng miền hCD163L1 8 biểu hiện CD163 và hỗ trợ sự sao chép các virut kiểu gen Typ 2, mà không phải là các virut kiểu gen Typ 1. Sự nhiễm các virut Typ 1 và Typ 2 đại diện của các con lợn được biến đổi *CD163* tạo ra các kết quả tương tự. Mặc dù các virut Typ 1 và Typ 2 được xem xét về mặt di truyền và kiểu hình là tương tự nhau ở một số mức độ, bao gồm yêu cầu đối với CD163 làm thụ thể, nhưng kết quả đã chứng minh sự khác biệt rõ rệt giữa các kiểu gen PRRSV trong việc nhận biết phân tử CD163.

Vật liệu và phương pháp

Các biến đổi hệ gen của gen CD163 của lợn

Các thử nghiệm liên quan đến động vật và virut được thực hiện theo Liên hiệp hội khoa học động vật: Hướng dẫn chăm sóc và sử dụng động vật nông nghiệp trong nghiên cứu và giảng dạy, USDA Animal Welfare Act and Animal Welfare Regulations, và được chấp thuận bởi Kansas State University và University of Missouri Institutional Animal Care and Use Committees và Institutional Biosafety Committees. Các đột biến trong *CD163* sử dụng trong thử nghiệm này được tạo ra bằng cách sử dụng công nghệ CRISPR/Cas9 như được mô tả trên đây trong các ví dụ trên. Các đột biến này được lập thành biểu đồ trên Fig.17. Vùng thuộc hệ gen được lập biểu đồ được thể hiện trên Fig.17 bao gồm trình tự từ intron 6 đến intron 8 của gen CD163 của lợn. Các intron và các exon được lập biểu đồ trên Fig.17 không được vẽ theo tỷ lệ. Sản phẩm protein được dự đoán được minh họa ở bên phải của mỗi cấu trúc hệ gen. Sự biểu hiện đại thực bào, như được đo bằng mức độ CD163 bề mặt trên PAMs, được thể hiện ở phần phía xa bên phải của Fig.17. Các vùng màu đen thể hiện các intron; các vùng màu trắng thể hiện các exon; vùng gạch bóng thể hiện dạng bắt chước h*CD163L1* exon 11, vùng tương đồng exon 7 của lợn; và vùng màu xám thể hiện intron tổng hợp với cấu trúc PGK Neo như được thể hiện trên Fig.17.

Cấu trúc gen CD163 *KO-d7(11)* được thể hiện trên Fig.17 có đoạn xóa 11 cặp bazơ trong exon 7 từ nucleotit 3137 đến nucleotit 3147. Cấu trúc gen CD163 *KO-i7(2)*, có đoạn xen 2 cặp bazơ trong exon 7 giữa các nucleotit 3149 và 3150 cũng như đoạn xóa 377 cặp bazơ ở ngược dòng intron của exon 7, từ nucleotit 2573 đến nucleotit 2949. Các chỉnh sửa này được dự đoán là gây ra các đột biến xê dịch khung và các codon kết thúc sớm, dẫn đến chỉ dịch mã một phần SRCR 5 và kiểu hình KO. Ba đột biến khác tạo ra các đoạn xóa trong exon 7. Đột biến đầu tiên, *d7(129)*, có đoạn xóa 129 cặp bazơ

trong exon 7 từ nucleotit 3044 đến nucleotit 3172. Cấu trúc *d7(129)* còn có đoạn xóa từ nucleotit 488 đến nucleotit 2417 trong exon 6, trong đó trình tự khuyết được thay thế bằng đoạn xen 12 bp. Hai cấu trúc đoạn xóa khác, *d7(1467)* và *d7(1280)*, có đoạn xóa hoàn toàn của các exon 7 và 8 như được minh họa trong Fig.17. *d7(1467)* có đoạn xóa 1467 cặp bazơ từ nucleotit 2431 đến nucleotit 3897, và *d7(1280)* có đoạn xóa 1280 cặp bazơ từ nucleotit 2818 đến nucleotit 4097. Đối với các cấu trúc đoạn xóa này, các exon *CD163* khác vẫn còn nguyên vẹn.

Cấu trúc cuối cùng được thể hiện trên Fig.17, *HL11m*, được sản xuất bằng cách sử dụng sự kiện đích mà xóa exon 7 và thay thế nó bằng một exon tổng hợp mà mã hóa cho vùng tương đồng SRCR 8 của protein giống CD163 người 1 (miền hCD163L1 8 được mã hóa bởi h*CD163L1* exon 11). Trình tự peptit SRCR 8 được tạo ra bằng cách tạo sự thay đổi 33 nucleotit trong trình tự exon 7 của lợn. Cassette neomycin được chứa trong exon tổng hợp để cho phép sàng lọc cho sự biến đổi. SEQ ID NO: 118 cung cấp trình tự nucleotit cho cấu trúc *HL11m* trong vùng tương ứng với vùng như vậy trong trình tự tham chiếu SEQ ID NO: 47.

Biểu đồ protein CD163 và gen của lợn được cung cấp trên Fig.18. Protein CD163 SCRC (hình trái xoan) và các miền PST (hình vuông) cùng với các exon gen tương ứng được thể hiện trong Panen A của Fig.18. Kết quả so sánh trình tự peptit đối với CD163 SRCR 5 của lợn (SEQ ID NO: 120) và vùng tương đồng CD163 SRCR 8 của người (SEQ ID NO: 121) được thể hiện trên Panen B của Fig.18. Fig này được dựa trên các số truy cập GenBank AJ311716 (CD163 lợn) và GQ397482 (hCD163-L1).

Các virut

Panen của các virut được sử dụng trong ví dụ này được liệt kê trong bảng 14. Các phần phân lập được truyền và chuẩn độ trên các tế bào MARC-145 (Kim et al., 1993). Để chuẩn độ, mỗi virut được pha loãng theo dãy 1:10 trong MEM bổ sung bằng 7% FBS, Pen-Strep (lần lượt 80 đơn vị/ml và 80 µg/ml), 3 µg/ml FUNGIZONE (amphotericin B), và 25 mM HEPES. Các mẫu được pha loãng được bổ sung thành bốn phần vào các tế bào MARC-145 hợp dòng trong đĩa 96 lỗ đến thể tích cuối là 200 µl trên lỗ và được ủ trong bốn ngày ở 37°C trong 5% CO_2. Các điểm đầu mút chuẩn độ được xác định như là lỗ cuối cùng với hiệu quả gây bệnh tế bào (cytopathic effect - CPE). Liều gây nhiễm môi trường mô 50% (50% tissue culture infectious dose -

$TCID_{50}$/ml) được tính sử dụng phương pháp như đã được mô tả (Reed and Muench 1938).

Bảng 14. Các phần phân lập PRRSV.

Virut	**Kiểu gen**	**Năm phân lập**	**GenBank số**
NVSL 97-7895	2	1997	AY545985
KS06-72109	2	2006	KM252867
P129	2	1995	AF494042
VR2332	2	1992	AY150564
CO90	2	2010	KM035799
AZ25	2	2010	KM035800
MLV-ResPRRS	2	NA*	AF066183
KS62-06274	2	2006	KM035798
KS483 (SD23983)	2	1992	JX258843
CO84	2	2010	KM035802
SD13-15	1	2013	Không có
Lelystad	1	1991	M96262
03-1059	1	2003	Không có
03-1060	1	2003	Không có
SD01-08	1	2001	DQ489311
4353PZ	1	2003	Không có

*NA, Không có (Not available)

Sự gây nhiễm các đại thực bào phế nang

Việc điều chế và gây nhiễm các đại thực bào được thực hiện như đã được mô tả (Gaudreault, et al., 2009 và Patton, et al., 2008). Phổi được lấy ra từ các con lợn làm chết êm ái và được rửa bằng cách rót 100 ml nước muối đệm phosphat lạnh (phosphate buffered saline - PBS) vào khí quản. Các khí quản được kẹp lại và phổi được xoa bóp

nhẹ. Các thành phần trong phế nang được rót vào ống ly tâm dung tích 50 ml và bảo quản trên đá. Các đại thực bào phế nang của lợn (PAMs) được lắng cặn bằng cách ly tâm ở tốc độ 1200 x g trong 10 phút ở 4°C. Viên cặn ly tâm được tái tạo huyền phù và được rửa một lần trong PBS vô trùng lạnh. Các viên cặn tế bào được tái tạo huyền phù trong môi trường đông lạnh chứa 45% RPMI 1640, 45% huyết thanh thai bò (FBS), và 10% dimetylsulfoxit (DMSO) và được bảo quản trong nitơ lỏng cho đến khi sử dụng. Các tế bào đóng băng được làm tan băng trên đá, được đếm và điều chỉnh đến mức $5x10^5$ tế bào /ml trong môi trường (RPMI 1640 bổ sung bằng 10% FBS, PenStrep, và FUNGIZONE; RPMI-FBS). Khoảng 10^3 PAMs trên lỗ được bổ sung vào các đĩa 96 lỗ và ủ qua đêm ở 37°C trong 5% CO_2. Các tế bào được rửa nhẹ để loại bỏ các tế bào không bám dính. Các phần pha loãng theo chuỗi 1:10 của virut được bổ sung vào các lỗ sao ba bản. Sau khi ủ qua đêm, các tế bào được rửa bằng PBS và được cố định trong 10 phút bằng 80% axeton. Sau khi làm khô, các lỗ được được nhuộm bằng SDOW-17 mAb đặc hiệu với protein PRRSV N (Rural Technologies Inc.) đã được pha loãng tỷ lệ 1:1000 trong PBS với 1% gelatin của cá (PBS-FG; Sigma Aldrich). Sau quá trình ủ 30 phút ở 37°C, các tế bào được rửa bằng PBS và được nhuộm bằng ALEXAFLUOR 488-chất kháng IgG chuột được đánh dấu (Thermofisher Scientific) đã được pha loãng 1:200 trong PBS-FG. Các đĩa được ủ trong 30 phút trong bóng tối ở 37°C, rửa bằng PBS, và được xem xét dưới kính hiển vi huỳnh quang. Liều gây nhiễm môi trường mô 50% ($TCID_{50}$)/ml được tính theo phương pháp đã được mô tả (Reed and Muench 1938).

Đo sự biểu hiện CD169 và CD163 trên bề mặt trên PAMs

Sự nhuộm để biểu hiện trên bề mặt của CD169 và CD163 được tiến hành như đã được mô tả trước đây (Prather et al., 2013). Khoảng $1X10^6$ PAMs được đặt trong các ống đếm tế bào theo dòng chảy kích thước 12 mm x 75 mm bằng polystyren (FACS) và ủ trong 15 phút ở nhiệt độ trong 1 ml PBS với 10% huyết thanh chuột bình thường để phong bế các thụ thể Fc. Các tế bào được kết viên bằng cách ly tâm và tái tạo huyền phù trong 5 µl mAb của chuột nhắt kháng CD169 của lợn, được liên hợp với FITC (dòng 3B11/11; AbD Serotec) và 5 µl mAb của chuột nhắt kháng CD163 của lợn, được liên hợp với PE (Dòng: 2A10/11, AbD Serotec). Sau 30 phút ủ, các tế bào được rửa hai lần bằng PBS chứa 1% albumin huyết thanh bò (BSA Fraction V; Hyclone) và lập tức được phân tích trên thiết bị đếm tế bào theo dòng BD LSR Fortessa (BD Biosciences)

với phần mềm FCS Express 5 (De Novo Software). Tối thiểu 10.000 tế bào được phân tích đối với mỗi mẫu.

Đo PRRS virut huyết

ARN được phân lập từ 50 µl huyết thanh sử dụng bộ kit phân lập virut của MagMAX 96 Ambion (Applied Biosystems) theo các hướng dẫn của nhà sản xuất. PRRSV ARN được định lượng sử dụng các tác nhân phản ứng đặc hiệu đối với đích RT-PCR thời gian thực EZ-PRRSV MPX 4.0 (Tetracore) được thực hiện theo các hướng dẫn của nhà sản xuất. Mỗi đĩa chứa các Chuẩn định lượng Tetracore và Bộ đối chứng được thiết kế để sử dụng với các tác nhân phản ứng RT-PCR. PCR được thực hiện trên Hệ thống phát hiện PCR Thời gian thực Tiếp xúc CFX96 (Bio-Rad) dưới dạng 96 lỗ sử dụng các thông số chu kỳ được khuyến nghị. Kết quả phân tích PCR được báo cáo dưới dạng $\log_{10}$ số bản sao PRRSV ARN trên 50 µl thể tích phản ứng, mà xấp xỷ với số lượng bản sao trên một ml huyết thanh. Diện tích dưới đường cong (AUC) đối với mức virut huyết theo thời gian được tính sử dụng phần mềm GraphPad Prism phiên bản 6.00 cho Windows.

Đo kháng thể PRRSV

Thử nghiệm miễn dịch huỳnh quang trung thể (microsphere fluorescent immunoassay - FMIA) để phát hiện các kháng thể kháng protein PRRSV nucleocapsid (N) được tiến hành như đã được mô tả trước đây (Stephenson et al., 2015). Protein PRRSV N tái tổ hợp được kết hợp với các giọt trung thể polystyren Luminex MAGPLEX được carboxyl hóa theo hướng dẫn của nhà sản xuất. Đối với FMIA, khoảng 2500 giọt được phủ kháng nguyên, được tạo huyền phù trong 50 µL PBS với 10% huyết thanh dê (PBS-GS), được đặt vào mỗi lỗ của đĩa 96 lỗ đáy tròn polystyren. Các huyết tương được pha loãng 1:400 trong PBS-GS và 50 µl được bổ sung vào mỗi lỗ. Đĩa được bọc trong lá kim loại và được ủ trong 30 phút ở nhiệt độ trong phòng với sự lắc nhẹ. Đĩa được đặt trên máy khuấy từ và các giọt được rửa ba lần bằng 190 µl PBS-GS. Để phát hiện IgG, 50 µl kháng thể dê kháng kháng thể thứ cấp của lợn được tinh chế ái lực được liên hợp với biotin-SP (IgG, Jackson ImmunoResearch) được pha loãng đến 2 µg/ml trong PBS-GS và 100 µl được bổ sung vào mỗi lỗ. Đĩa được ủ ở nhiệt độ trong phòng trong 30 phút và được rửa ba lần sau đó bổ sung 50 µl

phycoerythrin được liên hợp với streptavidin (2 ug/ml trong PBS-GS; SAPE). Sau 30 phút, các trung thể được rửa, tái tạo huyền phù trong 100 µl PBS-GS, và được phân tích sử dụng thiết bị MAGPIX (LUMINEX) và phần mềm LUMINEX xPONENT 4.2. Cường độ huỳnh quang trung bình (MFI) được tính trên tối thiểu 100 giọt trung thể.

Đo haptoglobin (HP)

Lượng Hp trong huyết thanh được đo sử dụng kit Hp ELISA đặc hiệu đối với lợn (Genway Biotech Inc.) và các bước được thực hiện các hướng dẫn của nhà sản xuất. Các mẫu huyết thanh được pha loãng 1:10.000 trong dung dịch 1X và được hút bằng pipet thành hai bản sao trên đĩa ELISA 96 lỗ được phủ trước kháng Hp lợn, được ủ ở nhiệt độ trong phòng trong 15 phút, sau đó được rửa ba lần. Thể liên hợp HRP kháng Hp được bổ sung vào mỗi lỗ và được ủ trong trong bóng tối ở nhiệt độ trong phòng trong 15 phút. Đĩa được rửa và 100 µl dung dịch chất nền chromogen được bổ sung vào mỗi lỗ. Sau khi ủ trong bóng tối trong 10 phút, 100 µl dung dịch chấm dứt được bổ sung vào mỗi lỗ. Đĩa được đọc ở 450 nm trên máy đọc vi đĩa trên cơ sở máy lọc Fluostar Omega (BMG Labtech).

KẾT QUẢ

Các tính chất kiểu hình của PAMs từ các con lợn được biến đổi gen CD163

Các tính chất phân tán tiếp và phân tán biên của các tế bào trong nguyên liệu rửa phổi được sử dụng để đóng mở trên quần thể phụ các tế bào đơn nhân. Kết quả nhuộm CD169 và CD163 đại diện cho các mẫu biến đổi gen khác nhau mà được thể hiện trên Fig.17 được biểu thị trên Fig.19. Trong ví dụ đại diện được biểu thị trên Panen A của Fig.19, trên 91% của PAMs từ các con lợn WT là dương tính đối với cả CD169 và CD163. Các kết quả đối với 12 con lợn sử dụng trong thử nghiệm này thể hiện giá trị trung bình là 85 +/– 8% số tế bào dương tính kép. Như được thể hiện trên Panen B của Fig.19, PAMs từ các con lợn KO *CD163* không thể hiện dấu hiệu của CD163, nhưng duy trì các mức CD169 bề mặt bình thường. Mặc dù đã dự đoán rằng các polypeptit CD163 thu được từ kiểu gen xóa đoạn *d7(1467) và d7(1280)* nên sản sinh các polypeptit CD163 được biến đổi được neo với bề mặt PAM, các kết quả nhuộm miễn dịch không thể hiện sự biểu hiện trên bề mặt của CD163 (xem Fig.19, Panen D). Vì MAb 2A10 nhận diện epitop nằm trong ba miền SRCR thứ nhất, nên không phát hiện không phải

là kết quả của sự xóa epitop phản ứng miễn dịch. Kiểu gen *d7(129)* được dự đoán là có đoạn xóa 43 axit amin trong SRCR 5 (xem Fig.17). Trong ví dụ được biểu thị trên Panen C của Fig.19, chỉ 2,4% tế bào nằm trong góc phần tư dương tính kép. Sự phân tích PAMs từ chín con lợn *d7(129)* sử dụng trong thử nghiệm này thể hiện phần trăm các tế bào dương tính kép nằm trong khoảng từ 0% đến 3,6% (trung bình = 0,9%). Sự biểu hiện trên bề mặt của CD169 vẫn tương tự như PAMs WT. Nhằm mục đích của thử nghiệm này, tất cả các con lợn có kiểu gen KO, *d7(1467)*, *d7(1280)*, *và d7(129)* được phân loại là có kiểu hình CD163-vô hiệu.

Sự biến đổi CD163 chứa trình tự peptit miền hCD163L1 8 *HL11m*, thể hiện sự biểu hiện kép của $CD163^+$ và $CD169^+$ trên PAMs (Panen E của Fig.19). Tuy nhiên, ở tất cả các con lợn *HL11m* được phân tích trong thử nghiệm này, sự biểu hiện trên bề mặt của CD163 được giảm đi một cách rõ rệt so với PAMs WT. Mức CD163 trong vùng biểu hiện liên tục, nằm trong khoảng từ CD163 không phát hiện được đến mức CD163 của các con lợn có mức CD163 vừa phải. Trong ví dụ được thể hiện trong Panen E của Fig.19, khoảng 60% tế bào nằm trong góc phần tư dương tính kép trong khi 40% số tế bào được nhuộm chỉ dương tính với CD169. Phân tích PAMs từ tổng số 24 con lợn *HL11m* thể hiện 38 +/– 12% tế bào PAM chỉ dương tính đối với CD169 và 54+/– 14% là dương tính kép ($CD169^+CD163^+$).

Các mức haptoglobin tuần hoàn ở lợn WT và lợn được biến đổi CD163

Dưới dạng phân tử làm sạch, CD163 chịu trách nhiệm cho việc loại bỏ các phức chất HbHp ra khỏi máu (Fabriek, et al., 2005; Kristiansen et al., 2001; và Madsen et al., 2004). Mức Hp trong huyết thanh cung cấp phương pháp thuận tiện để xác định các tính chất chức năng nói chung của các đại thực bào biểu hiện CD163. Mức Hp trong các huyết thanh từ các con lợn WT, *HL11m* và CD163-vô hiệu được đo ở ba đến bốn tuần tuổi, ngay trước khi nhiễm trùng bởi PRRSV. Kết quả, được biểu thị trên Fig.20, chỉ ra rằng các huyết thanh từ các con lợn WT có lượng Hb thấp nhất (A450 trung bình =23+/– 0,18, n=10). Trung bình và độ lệch chuẩn đối với mỗi nhóm là WT, 0,23+/– 0,18, n=10; *HL11m*, 1,63+/– 0,8, n=11; và 2,06 +/– 0,57, n=9, đối với nhóm vô hiệu. Nhóm vô hiệu gồm kiểu gen mà không biểu hiện CD163 (các con lợn có kiểu hình CD163 vô hiệu). Việc đo Hp được thực hiện trên đĩa ELISA đơn. Các nhóm với chữ cái giống nhau không khác biệt đáng kể ($p>0,05$, ANOVA một chiều Kruskal-Wallis với sự hậu kiểm

Dunnett). Giá trị A450 trung bình đối với các con lợn WT là khác biệt đáng kể với giá trị trung bình của các con lợn *HL11m* và các con lợn CD163-vô hiệu ($p < 0,05$). Mặc dù giá trị A450 trung bình đối với nhóm *HL11m* là thấp hơn so với nhóm CD163-vô hiệu (A450 = 1,6+/-0,8 so với 2,1+/–0,6), sự khác biệt là không có ý nghĩa thống kê. Vì tương tác giữa HbHp và CD163 xuất hiện qua SRCR 3 (Madsen et al., 2004), Hp tuần hoàn gia tăng ở các con lợn *HL11m* so với các con lợn WT dường như không phải là kết quả của ái lực giảm của CD163 đối với Hb/Hp, mà là kết quả của số lượng các đại thực bào $CD163^+$ bị giảm cùng với sự biểu hiện CD163 giảm trên các đại thực bào còn lại (xem Panen E của Fig.19).

Sự nhiễm của PAMs với các virut Typ 1 và Typ 2

Tính dễ mắc PRRSV của các con lợn được biến đổi CD163 ban đầu được đánh giá bằng cách gây nhiễm các tế bào PAM *in vitro* panen gồm sáu thể phân lập PRRSV typ 1 và chín thể phân lập PRRSV typ 2 (xem bảng 14 về danh sách các virut). Các virut trong panen này thể hiện kiểu gen khác nhau, cũng như sự khác nhau về các trình tự nucleotit và peptit, sinh bệnh học, và năm phân lập. Dữ liệu được biểu trên bảng 15 thể hiện kết quả từ các thử nghiệm sử dụng PAMs từ ba con lợn đối với nhóm kiểu gen *CD163*. Các virut được liệt kê tương ứng với các thể phân lập PRRSV được liệt kê trong bảng 14. Kết quả được thể hiện là giá trị trung bình +/– độ lệch chuẩn của phần trăm PAMs nhiễm. Các PAMs CD163-vô hiệu là từ các con lợn biểu hiện alen *d7(129)* (xem các Fig.17 và Fig.19 lần lượt đối với các cấu trúc gen CD163 và sự biểu hiện CD163 trên PAMs).

Bảng 15. Sự nhiễm các thể phân lập PRRSV khác nhau vào PAMs từ các con lợn kiểu hoang dại và các con lợn GM

	Kiểu gen/Kiểu hình (% nhiễm)		
Typ 1	**WT (%)**	***HL11m***	**Vô hiệu**
13-15	56 +/–9	0	0
Lelystad	62 +/–15	0	0
03-1059	50 +/–18	0	0
03-1060	61 +/–12	0	0
01-08	64 +/–20	0	0
4353-PZ	62 +/–15	0	0
Typ 2	**WT (%)**	***HL11m***	**Vô hiệu**
NVSL 97	59 +/–15	8 +/–08	0
KS-06	56 +/–20	12 +/–09	0
P129	64 +/–11	8 +/–06	0
VR2332	54 +/–05	6 +/–03	0
CO 10-90	43 +/–18	8 +/–08	0
CO 10-84	51 +/–22	7 +/–04	0
MLV-ResP	55 +/–12	3 +/–01	0
KS62	49 +/–03	10 +/–11	0
KS483	55 +/–23	6 +/– 03	0

Như được mong đợi, PAMs WT bị nhiễm bởi tất cả các virut. Ngược lại, các con lợn kiểu hình CD163-vô hiệu là âm tính đối với sự nhiễm tất cả các virut. Sự khác nhau rõ rệt được quan sát thấy ở đáp ứng của PAMs từ các con lợn *HL11m*. Không có con virut Typ 1 nào có thể nhiễm vào PAMs *HL11m*; trong khi đó, tất cả các con virut trong panen Typ 2 nhiễm vào PAMs *HL11m*, mặc dù ở mức phần trăm thấp hơn nhiều so với PAMs WT.

Sự dễ nhiễm cũng được đánh giá bằng cách so sánh các điểm đầu mút chuẩn độ virut giữa PAMs WT và PAMs *HL11m* đối với cùng virut Typ 2. Các kết quả được thể

hiện đối với hai con lợn WT và hai con lợn *HL11m* (Fig.21). Các giá trị $log_{10}TCID_{50}$ được tính dựa trên sự nhiễm các mẫu virut giống nhau của các môi trường đại thực bào. Kết quả nhiễm biểu thị hai con lợn khác nhau đối với mỗi kiểu gen. Các virut được sử dụng để gây nhiễm được liệt kê trong bảng 14. Các giá trị $log_{10}TCID_{50}$ đối với PAMs từ các con lợn *HL11m* là thấp hơn 1–3 giá trị log so với PAMs WT bị nhiễm bởi cùng virut này. Ngoại lệ duy nhất là sự nhiễm chủng vắc xin virut được biến đổi sự sống. Khi sử dụng cùng nhau, các kết quả gợi ý rằng PAMs từ các con lợn *HL11m* có tính nhạy giảm hoặc dễ nhiễm các virut Typ 2.

Sự nhiễm virut Typ 1 và Typ 2 của các con lợn được biến đổi CD163

Các con lợn WT (hình tròn), *HL11m* (hình vuông), và CD163-vô hiệu (hình tam giác) được gây nhiễm các virut đại diện Typ 1 (SD13-15) (Fig.22, Panen A, biểu đồ bên trái) và Typ 2 (NVSL 97-7895) (Fig.22, Panen A, biểu đồ bên phải). Các con lợn kiểu hình vô hiệu thu được từ các alen *d(1567)* và KO (xem Fig.17). Các con lợn từ ba kiểu gen được chủng bằng các virut giống nhau được trộn lẫn với nhau trong một chuồng, điều này cho phép các con lợn được biến đổi CD163 tiếp xúc liên tục với virut tỏa ra từ các con lợn WT cùng chuồng. Số lượng các con lợn bị nhiễm virut Typ 1 đại diện là: WT (n=4), *HL11m* (n=5), và Vô hiệu (n=3); và virut Typ 2: WT (n=4), *HL11m* (n=4), và Vô hiệu (n=3). Như được thể hiện trên Fig.22, các con lợn CD163-vô hiệu bị nhiễm virut Typ 1 hoặc virut Typ 2 là có virut huyết âm tính tại tất cả các thời điểm và không chuyển dạng huyết thanh. Như trông đợi, các con lợn WT bị nhiễm phát sinh có mức virut huyết trung bình gần 10^6 mẫu trên 50 µl phản ứng PCR ở thời điểm 7 ngày sau khi nhiễm đối với cả hai virut. Khoảng 14 ngày, tất cả các con lợn WT đã chuyển dạng huyết thanh (xem Fig.22, panen B). Phù hợp với kết quả nhiễm PAM (bảng 15), năm con lợn *HL11m* bị nhiễm virut Typ 1 không thể hiện dấu hiệu virut huyết hoặc kháng thể PRRSV. Tất cả các con lợn *HL11m* bị nhiễm thể phân lập Typ 2, NVSL, đều hỗ trợ sự nhiễm và chuyển dạng huyết thanh (Fig. 22, panen B). Sự giảm tính dễ nhiễm của các con lợn *HL11m* là không rõ ràng. Mức virut huyết trung bình đối với ba trong số bốn con lợn *HL11m* là tương tự như các con lợn WT. Tuy nhiên, đối với một con lợn *HL11m*, #101 (các hình vuông trống trong Fig. 22, panen A biểu đồ bên phải), mức virut huyết được giảm đi nhiều so với các con lợn khác trong nhóm kiểu gen *HL11m*. Sự giải thích đối với việc giảm mức virut huyết từ 3 đến 4 giá trị log đối với con lợn

#101 là không rõ ràng, nhưng gợi ý rằng một số con lợn *HL11m* có thể ít dễ mắc PRRSV hơn, quan sát hỗ trợ kết quả nhiễm PAM *in vitro* (bảng 15). Vì tất cả các con lợn được chủng bằng lượng virut như nhau và được giữ chung chuồng với các con lợn WT, mức virut huyết thấp hơn ở con lợn #101 không là kết quả của việc nhận được lượng virut thấp hơn hoặc ít tiếp xúc hơn với virut. Sự đếm tế bào theo dòng đối với các đại thực bào thể hiện rằng sự biểu hiện CD163 đối với con lợn #101 là có thể so sánh với các con lợn *HL11m* khác (số liệu không được thể hiện). Không có khác biệt về trình tự trong trình tự bắt chước exon 11.

Các thử nghiệm nhiễm virut bổ sung được tiến hành sử dụng hai virut, NVSL 97-7895 và KS06-72109. Kết quả được thể hiện trong Fig.23. Các con lợn được theo dõi trong 35 ngày sau khi nhiễm và số liệu được báo cáo dưới dạng diện tích dưới đường cong (AUC) đối với các số đo mức virut huyết được lấy ở thời điểm 3, 7, 11, 14, 21, 28 và 35 ngày sau khi nhiễm. Như được thể hiện trên Fig.23, đối với NVSL, giá trị AUC trung bình đối với bảy con lợn WT bị nhiễm NVSL là 168 +/– 8 so với 165 +/– 15 đối với bảy con lợn *HL11m*. Đối với KS06, các giá trị AUC trung bình đối với sáu con lợn WT và sáu con lợn *HL11m* lần lượt là 156 +/– 9 và 163 +/–13. Đối với cả hai virut, không có sự khác biệt có ý nghĩa thống kê giữa các con lợn WT và các con lợn *HL11m* ($p > 0,05$). Khi xử lý cùng nhau, kết quả thể hiện rằng các con lợn *HL11m* không hỗ trợ sự nhiễm PRRSV Typ 1, nhưng vẫn dễ nhiễm virut Typ 2. Mặc dù có sự giảm tính dễ nhiễm PRRSV của PAMs từ các con lợn *HL11m* đã nhiễm *in vitro* thể phân lập Typ 2, khác biệt này không chuyển sang cho lợn. Đối với các kết quả được thể hiện trên Fig.23, tải lượng virut được xác định bằng cách tính diện tích dưới đường cong (AUC) đối với mỗi con lợn trong quá trình nhiễm 35 ngày. Tính toán AUC được tiến hành sử dụng các số đo PCR virut huyết $\log_{10}$ ở thời điểm 0, 4, 7, 10, 14, 21, 28 và 35 ngày sau khi nhiễm. Các đường nằm ngang thể hiện giá trị trung bình và độ lệch chuẩn. Các chữ viết tắt: WT = lợn kiểu hoang dại, HL11 = lợn kiểu gen *HL11m*; Vô hiệu = kiểu gen CD163-vô hiệu.

Bàn luận

CD163 là protein bề mặt đại thực bào quan trọng cho việc làm sạch lượng dư Hb từ máu và điều biến sự viêm đáp ứng lại sự tổn thương mô. Nó cũng có tác dụng như là thụ thể virut. CD163 tham gia vào cả đáp ứng tiền viêm và đáp ứng chống viêm (Van

Gorp et al., 2010). Các đại thực bào dương tính với CD163 được đặt trong nhóm các đại thực bào M2 được hoạt hóa theo cách khác, mà thường được mô tả là có tính thực bào cao và chống viêm. Các đại thực bào M2 tham gia vào quá trình làm sạch và sửa chữa sau khi tổn thương mô cơ học hoặc sau khi nhiễm trùng (Stein et al., 1992). Về khả năng chống viêm, sự biểu hiện CD163 được điều hòa ngược bằng các protein chống viêm, như IL-10 (Sulahian, et al., 2002). Trong quá trình viêm, CD163 làm giảm sự viêm bằng cách giảm chất oxy hóa thông qua việc loại bỏ hem tuần hoàn ra khỏi máu. Các sản phẩm thoái biến hem, như bilverdin, bilirubin, và cacbon monoxit là các phân tử chống viêm có hiệu lực (Soares và Bach, 2009 và Jeney et al., 2002). Về khả năng tiền viêm, sự liên kết ngang của CD163 trên bề mặt đại thực bào bằng kháng thể kháng CD163 hoặc vi khuẩn dẫn đến sự giải phóng cục bộ các xytokin tiền viêm, bao gồm IL-6, GM-CSF, TNFα và IL-1β (Van den Heuvel et al., 1999 và Fabriek et al., 2009).

Các con lợn GM mà thiếu CD163 không hỗ trợ sự sao chép của thể phân lập PRRSV Typ 2 (Whitworth et al., 2016). Trong nghiên cứu này, các thử nghiệm nhiễm *in vitro* chứng minh sự kháng của các đại thực bào kiểu hình CD163 vô hiệu đối với panen rộng của các thể phân lập PRRSV Typ 1 và Typ 2, mở rộng hơn nữa sự kháng có hiệu lực đối với tất cả các thể phân lập PRRSV (bảng 15). Sự kháng của các đại thực bào kiểu hình CD163-vô hiệu đối với các virut Typ 1 và Typ 2 được xác nhận *in vivo* (Fig.22 và Fig.23). Dựa trên các kết quả này, sự đóng góp của các thụ thể PRRSV khác đã được mô tả trong tài liệu tham khảo (Zhang và Yoo, 2015) có thể được loại trừ. Ví dụ, Shanmukhappa et al. (2007) thể hiện rằng các tế bào BHK không dễ nhiễm được chuyển nhiễm bằng plasmit CD151 đạt được khả năng hỗ trợ sự sao chép PRRSV, và sự ủ với kháng thể đa dòng kháng CD151 đã thể hiện sự giảm đáng kể sự nhiễm của các tế bào MARC-145. Ngoài ra, dòng tế bào khỉ, SJPL, được phát triển nguồn gốc để sử dụng trong việc truyền virut cúm lợn, đã được thể hiện là hỗ trợ sự sao chép PRRSV (Provost, et al., 2012). Các tính chất quan trọng của dòng tế bào SJPL bao gồm sự có mặt của CD151 và sự không có mặt của sialoadhesin và CD163. Khi xử lý cùng nhau, các số liệu này cung cấp bằng chứng chứng minh rằng sự có mặt của một mình CD151 là đủ để hỗ trợ sự sao chép PRRSV. Các kết quả từ nghiên cứu này thể hiện sự không nhiễm trùng PRRSV ở các đại thực bào và các con lợn có kiểu hình CD163 vô hiệu chỉ ra rằng CD151 dưới dạng là thụ thể khác đối với PRRSV là không liên quan về mặt sinh học.

Các protein virut GP2a và GP4, mà tạo thành một phần của phức chất heterotrime GP2a, GP3, GP4 trên bề mặt PRRSV, có thể được đồng kết tủa với CD163 trong các thử nghiệm đẩy xuống từ các tế bào được chuyển nhiễm bằng các plasmit GP2 và GP4 (Das, et al., 2009). Giả định rằng, GP2 và GP4 tạo thành tương tác với một hoặc nhiều miền CD163 SRCR. Thử nghiệm mức độ nhiễm *in vitro* kết hợp khung cADN CD163 lợn chứa miền hoán đổi giữa SRCR 5 lợn và homolog từ hCD163-L1 SRCR 8 còn định vị vùng được sử dụng bởi các virut Typ 1 cho SRCR 5 (Van Gorp, et al., 2010). Thú vị là suy xét thấy rằng tương tác ổn định giữa GP2/GP4 và CD163 xuất hiện thông qua SRCR 5. Các glycoprotein virut bổ sung, như GP3 và GP5, có thể còn ổn định hóa phức hợp virut-thụ thể hoặc có thể có tác dụng như là các phân tử đồng thụ thể. Yêu cầu đối với SRCR 5 được khảo sát trong nghiên cứu này bằng cách gây nhiễm các đại thực bào và các con lợn có alen *HL11m*, mà tái tạo hoán đổi miền CD163L1 SRCR 8 bằng cách tạo ra sự thay thế 33 bp trong exon 7 của lợn. Alen *HL11m* còn bao gồm cassette neomycin để chọn lọc các tế bào dương tính đối với các biến đổi gen (Fig.17). Các con lợn *HL11m* biểu hiện CD163 trên PAMs, mặc dù ở mức giảm so với PAMs WT (Fig.19, so sánh các panen A và E). Sự biểu hiện giảm dường như là do sự có mặt của cassette neomycin, mà nằm giữa phần bắt chước exon 11 và intron sau đó. Các con lợn *HL11m* không dễ nhiễm virut Typ 1, xác nhận sự quan trọng của SRCR 5. Tuy nhiên, các đại thực bào *HL11m* và các con lợn *HL11m* hỗ trợ sự nhiễm virut Typ 2. Dựa vào kết quả chuẩn độ virut và phần trăm nhiễm, PAMs từ các con lợn *HL11m* thể hiện sự giảm nói chung về tính dễ nhiễm virut so với các đại thực bào WT (bảng 15 và Fig.17). Tính dễ nhiễm giảm có thể do mức CD163 giảm trên các đại thực bào *HL11m*, kết hợp với ái lực giảm của virut đối với protein CD163 được biến đổi. Giả định rằng các virut Typ 2 có yêu cầu của SRCR 5 và rằng L1 SRCR 8 có thể có tác dụng như là chất thay thế thích hợp, ái lực thấp hơn có thể được giải thích bằng sự khác biệt về các trình tự peptit giữa SRCR 8 người và SRCR 5 lợn (xem Fig.18, Panen B). Tuy nhiên, tính dễ mắc giảm của PAMs không được chuyển sang cho lợn. Mức virut huyết trung bình đối với các con lợn *HL11m* khong khác biệt đáng kể khi so với các con lợn WT (Fig.23). Ngoài PAMs, sự nhiễm trùng PRRSV của các đại thực bào mô nội mạch, vách ngăn và lymphoid góp phần vào mức virut huyết (Lawson et al., 1997 và Morgan et al., 2014). Sự đóng góp tiềm năng của các đại thực bào này và các quần thể tế bào dương tính với

CD163 khác trong việc duy trì tải lượng virut nói chung ở lợn *HL11m* xứng đáng được nghiên cứu thêm.

Mặc dù các plasmit CD163 có sự xóa các miền SRCR được biểu hiện ổn định ở các tế bào HEK (Van Gorp et al., 2010), nhưng sự xóa các exon 7 và 8 trong *d7(1467)* và *d7(1280)* dẫn đến sự thiếu sự biểu hiện trên bề mặt có thể phát hiện được của CD163 (Fig.19, Panen D). Vì mAB 2A10 được sử dụng cho việc đếm tế bào theo dòng nhận diện ba miền SRCR đầu tận cùng N (Van Gorp et al., 2010), và có thể nhận diện các miền thứ 7 và thứ 8 (Sanchez, et al., 1999), nên sự không phát hiện được không phải do sự loại bỏ epitop 2A10 trong các protein đột biến. Mặc dù lượng biểu hiện CD163 nhỏ có thể được phát hiện trên PAMs từ một số con lợn *d7(129)* (xem Fig. 19, Panen C), nhưng lượng protein được biểu hiện không đủ để hỗ trợ sự nhiễm trùng PRRSV ở PAMs hoặc lợn. Sự không biểu hiện CD163 trên các thể đột biến xóa exon 7 và 8 không được hiểu biết đầy đủ, nhưng dường như là kết quả của sự thoái biến mARN và/hoặc protein.

Năm 2003, CD163 được nhận diện là thụ thể của virut sốt lợn châu Phi (ASFV; Sánchez-Torres et al., 2003). Kết luận này là dựa trên quan sát rằng các đại thực bào bị nhiễm có kiểu hình dương tính với CD163 trưởng thành, và các kháng thể kháng CD163, như 2A10, phong bế sự nhiễm ASFV của các đại thực bào *in vitro*. Nhưng vẫn cần xác định liệu các con lợn CD163-vô hiệu có kháng lại sự nhiễm ASFV hay không.

Các mô hình nuôi cấy tế bào kết hợp với các biến đổi đối với thụ thể PRRSV đã cung cấp nhận thức có giá trị đối với các cơ chế xâm nhập, sao chép và sinh bệnh học PRRSV. Một khía cạnh duy nhất của thử nghiệm này là sự thực hiện thử nghiệm song song *in vivo* sử dụng các con lợn được biến đổi thụ thể. Nghiên cứu này có tác động quan trọng đối với tính khả thi của việc phát triển các phương thức chữa bệnh phòng ngừa đối với một trong số các bệnh nguy hiểm nhất từng phải đối mặt trong ngành công nghiệp chăn nuôi lợn trên toàn cầu.

Các ví dụ được bộc lộ ở đây được cung cấp bằng cách ví dụ và không được dự tính là để giới hạn phạm vi của sáng chế.

Dựa vào phần mô tả trên đây, người ta sẽ thấy rằng một số đối tượng theo sáng chế đạt được và các kết quả có lợi khác cũng đạt được.

Các thay đổi khác có thể được tạo ra đối với các sản phẩm trên và các phương pháp nêu trên mà không trệch khỏi phạm vi của sáng chế, dự tính rằng tất cả các vấn đề

được nêu trong phần mô tả này và được thể hiện trong phần hình vẽ kèm theo sẽ được hiểu là nhằm mục đích minh họa và không có nghĩa giới hạn.

Tài liệu tham khảo

Allende R, et al., 1999. North American and European porcine reproductive and respiratory syndrome viruses differ in non-structural protein coding regions. J. Gen. Virol. 80: 307–315.

Bauer BK, et al., Arginine supplementation in vitro increases porcine embryo development and affects mRNA transcript expression. Reprod Fertil Dev 2011; 23:107.

Beaton BP, et al., Compound transgenics: recombinase- mediated gene stacking. In: Pinkert CA (ed.). Transgenic Animal Technology: A Laboratory Handbook, 3rd ed. Waltham, MA: Elsevier; 2014:565–578.

Benfield DA, et al., 1992. Characterization of swine infertility and respiratory syndrome (SIRS) virus (Isolate ATCC VR-2332). J Vet Diag Invest 4:127-133.

Boddicker, et al., 2014. Genome-wide association and genomic prediction for host response to porcine reproductive and respiratory syndrome virus infection. Genetics, selection, evolution : GSE 46, 18.

Borg NA, et al., CD1d-lipid- antigen recognition by the semi-invariant NKT T-cell receptor. Nature 2007; 448:44–49.

Brinster RL, et al., Factors affecting the efficiency of introducing foreign DNA into mice by microinjecting eggs. Proc Natl Acad Sci U S A 1985; 82:4438–4442.

Calvert JG, et al., 2007. CD163 expression confers susceptibility to porcine reproductive and respiratory syndrome viruses. J Virol. 81:7371-7379.

Carter DB, et al., Phenotyping of transgenic cloned piglets. Cloning Stem Cells 2002; 4:131–145.

Cong L, et al., Multiplex genome engineering using CRISPR/Cas systems. Science 2013; 339:819–823.

Dai Y, et al., Targeted disruption of the alpha1,3-galactosyltransferase gene in cloned pigs. Nat Biotechnol 2002; 20:251–255.

Das PB, et al., 2009. The minor envelope glycoproteins GP2a and GP4 of porcine reproductive and respiratory syndrome virus interact with the receptor CD163. J Virol. 84:1731-40.

Delputte PL, et al., 2002. Involvement of the matrix protein in attachment of porcine reproductive and respiratory syndrome virus to a heparin-like receptor on porcine alveolar macrophages. J Virol. 76:4312-4320.

Etzerodt, A., et al., 2013. Plasma clearance of hemoglobin and haptoglobin in mice and effect of CD163 gene targeting disruption. Antioxidants & redox signaling 18, 2254-2263.

Etzerodt, A., et al., 2013. CD163 and inflammation: biological, diagnostic, and therapeutic aspects. Antioxidants & redox signaling 18, 2352-2363.

Fabriek BO, et al., 2005. The macrophage scavenger receptor CD163. Immunobiology. 210:153-160.

Fabriek BO, et al., 2009. The macrophage scavenger receptor CD163 functions as an innate immune sensor for bacteria. Blood. 113:887-892.

Gaj T, et al., ZFN, TALEN, and CRISPR/Cas- based methods for genome engineering. Trends Biotechnol 2013; 31: 397–405.

Gaudreault N, et al., 2009. Factors affecting the permissiveness of porcine alveolar macrophages for porcine reproductive and respiratory syndrome virus. Archiv Virol. 154:133-136.

Graversen, J.H., et al., 2012. Targeting the hemoglobin scavenger receptor CD163 in macrophages highly increases the anti-inflammatory potency of dexamethasone. Molecular therapy : the journal of the American Society of Gene Therapy 20, 1550-1558.

Groenen, M.A., et al., 2012. Analyses of pig genomes provide insight into porcine demography and evolution. Nature 491, 393-398.

Hai T, et al., One-step generation of knockout pigs by zygote injection of CRISPR/Cas system. Cell Res 2014; 24: 372–375.

Halbur, P.G., et al., 1995. Comparison of the pathogenicity of two US porcine reproductive and respiratory syndrome virus isolates with that of the Lelystad virus. Veterinary pathology 32, 648-660.

Hammer RE, et al., Production of transgenic rabbits, sheep and pigs by microinjection. Nature 1985; 315:680–683.

Hauschild J, et al., Efficient generation of a biallelic knockout in pigs using zinc-finger nucleases. Proc Natl Acad Sci U S A 2011; 108:12013–12017.

Holtkamp, D.J., et al., 2013. Assessment of the economic impact of porcine reprodutive and respiratory syndrome virus on United States pork producers. Journal of Swine Health and Production 21, 72-84.

Hwang WY, et al., Efficient genome editing in zebrafish using a knockout pigs via zinc-finger nucleases and nuclear transfer cloning. Cell Res 2011; 21:979–982.

Hwang WY, et al., Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol 2013; 31:227–229.

Jeney V, et al., 2002. Pro-oxidant and cytotoxic effects of circulating heme. Blood. 100:879-887.

Keffaber, K.K., 1989. Reproductive failure of unknown etiology. American Association of Swine Practitioners Newsletter 1, 1-10.

Kim, H.S., et al.,1993. Enhanced replication of porcine reproductive and respiratory syndrome (PRRS) virus in a homogeneous subpopulation of MA-104 cell line. Arch Virol 133, 477-483.

Kristiansen M, et al., 2001. Identification of the haemoglobin scavenger receptor. Nature. 409:198-201.

Kwon DN, et al., Production of biallelic CMP-Neu5Ac hydroxylase knock-out pigs. Sci Rep 2013; 3:1981.

Ladinig, A., et al., 2015. Pathogenicity of three type 2 porcine reproductive and respiratory syndrome virus strains in experimentally inoculated pregnant gilts. Virus Res 203, 24-35.

Lai L, et al., Production of alpha-1,3-galactosyltransferase knockout pigs by nuclear transfer cloning. Science 2002; 295:1089–1092.

Lai L, et al., Creating genetically modified pigs by using nuclear transfer. Reprod Biol Endocrinol 2003; 1:82.

Lai L, et al., Production of cloned pigs by using somatic cells as donors. Cloning Stem Cells 2003; 5:233–241.

Lawson SR, et al.,1997. Porcine reproductive and respiratory syndrome virus infection of gnotobiotic pigs: sites of virus replication and co-localization with MAC-387 staining at 21 days post-infection. Virus Res. 51:105-113.

Lee K, et al., Engraftment of human iPS cells and allogeneic porcine cells into pigs with inactivated RAG2 and accompanying severe combined immunodeficiency. Proc Natl Acad Sci U S A 2014; 111:7260–7265.

Lee K, et al., Piglets produced from cloned blastocysts cultured in vitro with GM-CSF. Mol Reprod Dev 2013; 80: 145–154.

Li D, et al., Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. Nat Biotechnol 2013; 31:681–683.

Lillico SG, et al., Live pigs produced from genome edited zygotes. Sci Rep 2013; 3:2847.

Machaty Z, et al., Complete activation of porcine oocytes induced by the sulfhydryl reagent, thimerosal. Biol Reprod 1997; 57:1123–1127.

Madsen M, et al., 2004. Molecular characterization of the haptoglobin-hemoglobin receptor CD163. Ligand binding properties of the scavenger receptor cysteine-rich domain region. J Biol Chem. 279:51561-51567.

Merck, The Merck Veterinary Manual. http://www.merckmanuals.com/vet/appendixes/reference_guides/normal_rectal_temperature_ranges.html.

Miao, Y.L., et al., 2009. Centrosome abnormalities during porcine oocyte aging. Environmental and molecular mutagenesis 50, 666-671.

Morgan SB, et al., 2014. Pathology and Virus Distribution in the Lung and Lymphoid Tissues of Pigs Experimentally Inoculated with Three Distinct Type 1 PRRS Virus Isolates of Varying Pathogenicity. Transbound Emerg Dis. Nov 10. pp 1-11.

Nelsen CJ, et al., 1999. Porcine reproductive and respiratory syndrome virus comparison: divergent evolution on two continents. J. Virol. 73, 270–280.

Neumann EJ, et al., Assessment of the economic impact of porcine reproductive and respiratory syndrome on swine production in the United States. J Am Vet Med Assoc 2005; 227:385–392.

Niu Y, et al., Generation of gene-modified Cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. Cell 2014; 156:836–843.

Patience, J.F., et al.,1989. Swine Nutrition Guide, in: Center, P.S. (Ed.), University of Saskatchewan, Saskatoon, CA, pp. 149-171.

Patton JP, et al., 2008. Modulation of CD163 receptor expression and replication of porcine reproductive and respiratory syndrome virus in porcine macrophages. Virus Res. 140: 161-171.

Prather RS, et al., An intact sialoadhesin (Sn/SIGLEC1/CD169) is not required for attachment/ internalization of the porcine reproductive and respiratory syndrome virus. J Virol 2013; 87:9538–9546.

Prather, R.S., et al., 2013. An intact sialoadhesin (Sn/SIGLEC1/CD169) is not required for attachment/internalization of the porcine reproductive and respiratory syndrome virus. J Virol 87, 9538-9546.

Provost C, et al., 2012. Identification of a new cell line permissive to porcine reproductive and respiratory syndrome virus infection and replication which is phenotypically distinct from MARC-145 cell line. Virol J. 9:267.

Reed JL, et al., 1938. A simple method of estimating fifty percent endpoints. The American Journal of Hygiene 27:493–497.

Ross JW, et al., Optimization of square-wave electroporation for transfection of porcine fetal fibroblasts. Transgenic Res 2010; 19:611–620.

Rowland, R.R., et al., 2012. Control of porcine reproductive and respiratory syndrome (PRRS) through genetic improvements in disease resistance and tolerance. Frontiers in genetics 3, 260.

Sanchez C, et al., 1999. The porcine 2A10 antigen is homologous to human CD163 and related to macrophage differentiation. Journal of Immunology 162:5230-5237.

Sánchez-Torres C, et al., 2003. Expression of porcine CD163 on monocytes/macrophages correlates with permissiveness to African swine fever infection. Arch Virol. 148:2307-2323.

Schaer, C.A., et al., 2006a. Constitutive endocytosis of CD163 mediates hemoglobin-heme uptake and determines the noninflammatory and protective transcriptional response of macrophages to hemoglobin. Circulation research 99, 943-950.

Schaer, D.J., et al., 2006b. CD163 is the macrophage scavenger receptor for native and chemically modified hemoglobins in the absence of haptoglobin. Blood 107, 373-380.

Schaer, D.J., et al., 2006c. Hemophagocytic macrophages constitute a major compartment of heme oxygenase expression in sepsis. European journal of haematology 77, 432-436.

Shanmukhappa, K, et al., 2007. Role of CD151, A tetraspanin, in porcine reproductive and respiratory syndrome virus infection. Virol J. 4:62.

Shimozawa N, et al., Abnormalities in cloned mice are not transmitted to the progeny. Genesis 2002; 34:203–207.

Smit, AFA, et al., RepeatMasker Open-3.0. 1996–2010. Current Version: open-4.0.5 (RMLib: 20140131 and Dfam: 1.2). http:// www.repeatmasker.org&gt. CD163: accessed July 25, 2014; CD1D: accessed August 27, 2013.

Soares MP, et al., 2009. Heme oxygenase-1: from biology to therapeutic potential. Trends Mol Med. 15:50-58.

Stein M, et al.,1992. Interleukin 4 potently enhances murine macrophage mannose receptor activity: a marker of alternative immunologic macrophage activation. J Exp Med. 176:287-292.

Stephenson R, et al., 2015. Multiplex serology for common viral infections in feral pigs in Hawaii between 2007 and 2010. Accepted for publication. J Wildlife Dis 51:239-2343.

Sulahian TH1, et al., 2002. 2002. Human monocytes express CD163, which is upregulated by IL-10 and identical to p155. Cytokine. 2000 12:1312-1321.

Terns MP, et al., CRISPR-based adaptive immune systems. Curr Opin Microbiol 2011; 14:321–327.

Trible, B.R., et al., 2012. Recognition of the different structural forms of the capsid protein determines the outcome following infection with porcine circovirus type 2. J Virol 86, 13508-13514.

U.S. Department of Agriculture and U.S. Department of Health and Human Services. Dietary Guidelines for Americans, 2010. 7th Edition,

Van Breedam W, et al., Porcine reproductive and respiratory syndrome virus entry into the porcine macrophage. J Gen Virol 2010; 91:1659–1667.

Van Breedam, W., et al., 2010. Porcine reproductive and respiratory syndrome virus entry into the porcine macrophage. J Gen Virol 91, 1659-1667.

Van den Heuvel MM, et al., 1999. Regulation of CD 163 on human macrophages: cross-linking of CD163 induces signaling and activation. J Leukoc Biol. 66:858-866.

van den Hoff MJ, et al., Electroporation in 'intracellular' buffer increases cell survival. Nucleic Acids Res 1992; 20:2902.

Van Gorp, H., et al., 2010. Scavenger receptor CD163, a Jack-of-all-trades and potential target for cell-directed therapy. Mol Immunol 47, 1650-1660.

Van Gorp H, et al., 2010. Identification of the CD163 protein domains involved in infection of the porcine reproductive and respiratory syndrome virus. J Virol. 84:3101-3105.

Walters EM, et al., Advancing swine models for human health and diseases. Mo Med 2013; 110:212–215.

Wang H, et al., One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell 2013; 153: 910–918.

Welch SK, et al., 2010. A brief review of CD163 and its role in PRRSV infection. Virus Res. 154:98-103.

Wells, K.D., et al., 2014. Use of the CRISPR/Cas9 system to produce genetically engineered pigs from in vitro-derived oocytes and embryos. Biol Reprod 91, 78.

Wensvoort, G., et al., 1991. Mystery swine disease in The Netherlands: the isolation of Lelystad virus. Veterinary Quarterly 13, 121-130.

Whitworth KM, et al., 2016. CD163 facilitates both entry and replication of porcine reproductive and respiratory syndrome virus. Nature Biotech. 34:20-22.

Whitworth KM, et al., Method of oocyte activation affects cloning efficiency in pigs. Mol Reprod Dev 2009;76:490–500.

Whitworth KM, et al., Activation method does not alter abnormal placental gene expression and development in cloned pigs. Mol Reprod Dev 2010; 77:1016–1030.

Whitworth KM, et al., Scriptaid corrects gene expression of a few aberrantly reprogrammed transcripts in nuclear transfer pig blastocyst stage embryos. Cell Reprogram 2011; 13:191–204.

Whitworth, K.M., et al., 2014. Use of the CRISPR/Cas9 system to produce genetically engineered pigs from in vitro-derived oocytes and embryos. Biol Reprod 91, 78.

Whyte JJ, et al., Genetic modifications of pigs for medicine and agriculture. Mol Reprod Dev 2011; 78:879–891.

Whyte JJ, et al., Gene targeting with zinc finger nucleases to produce cloned eGFP knockout pigs. Mol Reprod Dev 2011; 78:2.

Wiedenheft B, et al., RNA-guided genetic silencing systems in bacteria and archaea. Nature 2012; 482:331–338.

Winckler, C., et al., 2001. The reliability and repeatability of a lameness scoring system for use as an indicator of welfare in dairy cattle. Acta Agricultura Scandinavica Section A. Animal Science, 103-107.

Yang D, et al., Generation of PPARgamma mono-allelic knockout pigs via zinc-finger nucleases and nuclear transfer cloning. Cell Res 2011; 21:979–982.

Yoshioka K, et al., Birth of piglets derived from porcine zygotes cultured in a chemically defined medium. Biol Reprod 2002; 66:112–119.

Zhang Q, et al., 2015. PRRS virus receptors and their role for pathogenesis. Vet Microbiol. 177:229-241.

Zhao J, et al., Histone deacetylase inhibitors improve in vitro and in vivo developmental competence of somatic cell nuclear transfer porcine embryos. Cell Reprogram 2010; 12:75–83.

Zhao J, et al., Significant improvement in cloning efficiency of an inbred miniature pig by histone deacetylase inhibitor treatment after somatic cell nuclear transfer. Biol Reprod 2009; 81:525–530.

Bảng các trình tự

SEQ	KIỂU	MÔ TẢ
SEQ ID NO:1	nucleotit	CRISPR 10
SEQ ID NO:2	nucleotit	CRISPR 131
SEQ ID NO:3	nucleotit	CRISPR 256
SEQ ID NO:4	nucleotit	CRISPR 282
SEQ ID NO:5	nucleotit	CRISPR 4800
SEQ ID NO:6	nucleotit	CRISPR 5620
SEQ ID NO:7	nucleotit	CRISPR 5626
SEQ ID NO:8	nucleotit	CRISPR 5350
SEQ ID NO:9	nucleotit	eGFP1
SEQ ID NO:10	nucleotit	eGFP2
SEQ ID NO:11	nucleotit	đoạn mồi xuôi 9538
SEQ ID NO:12	nucleotit	đoạn mồi ngược 9538
SEQ ID NO:13	nucleotit	đoạn mồi xuôi 8729
SEQ ID NO:14	nucleotit	đoạn mồi xuôi 8729

SEQ ID NO:15	nucleotit	CD163 kiểu hoang dại
SEQ ID NO:16	nucleotit	Fig. 4, Panen C WT
SEQ ID NO:17	nucleotit	Fig. 4, Panen C #1
SEQ ID NO:18	nucleotit	Fig. 4, Panen C #2
SEQ ID NO:19	nucleotit	Fig. 4, Panen C #3
SEQ ID NO:20	nucleotit	Fig. 5, Panen A WT
SEQ ID NO:21	nucleotit	Fig. 5, Panen A #1-1
SEQ ID NO:22	nucleotit	Fig. 5, Panen A #1-4
SEQ ID NO:23	nucleotit	Fig. 5, Panen A #2-2
SEQ ID NO:24	nucleotit	Fig. 6, Panen C CD163 WT
SEQ ID NO:25	nucleotit	Fig. 6, Panen C CD163 #1
SEQ ID NO:26	nucleotit	Fig. 6, Panen C CD163 #2
SEQ ID NO:27	nucleotit	Fig. 6, Panen C CD163 #3
SEQ ID NO:28	nucleotit	Fig. 6, Panen C eGFP WT
SEQ ID NO:29	nucleotit	Fig. 6, Panen C eGFP #1-1
SEQ ID NO: 30	nucleotit	Fig. 6, Panen C eGFP #1-2
SEQ ID NO:31	nucleotit	Fig. 6, Panen C eGFP #2
SEQ ID NO:32	nucleotit	Fig.6, Panen C eGFP #3
SEQ ID NO:33	nucleotit	Fig. 7, Panen C WT
SEQ ID NO:34	nucleotit	Fig. 7, Panen C #67-1
SEQ ID NO:35	nucleotit	Fig. 7, Panen C #67-2 a1
SEQ ID NO:36	nucleotit	Fig. 7, Panen C #67-2 a2
SEQ ID NO:37	nucleotit	Fig. 7, Panen C #67-3
SEQ ID NO:38	nucleotit	Fig. 7, Panen C #67-4 a1
SEQ ID NO:39	nucleotit	Fig. 7, Panen C #67-4 a2

SEQ ID NO:40	nucleotit	Fig. 8, Panen D WT
SEQ ID NO:41	nucleotit	Fig. 8, Panen D #166-1.1
SEQ ID NO:42	nucleotit	Fig. 8, Panen D #166-1.2
SEQ ID NO:43	nucleotit	Fig. 8, Panen D #166-2
SEQ ID NO:44	nucleotit	Fig. 8, Panen D #166-3.1
SEQ ID NO:45	nucleotit	Fig. 8, Panen D #166-3.2
SEQ ID NO:46	nucleotit	Fig. 8, Panen D #166-4
SEQ ID NO:47	nucleotit	Fig. 16 WT CD163 partial
SEQ ID NOs. 48–67	nucleotit	Các trình tự đoạn mồi (Bảng 1)
SEQ ID NOs. 68–79	nucleotit	Các trình tự đoạn mồi (Bảng 2)
SEQ ID NOs. 80–85	nucleotit	Các trình tự đoạn mồi (Bảng 3)
SEQ ID NOs. 86–97	nucleotit	Các trình tự đoạn mồi (Bảng 4)
SEQ ID NO: 98	nucleotit	alen với đoạn xóa 1506 bp
SEQ ID NO: 99	nucleotit	alen với đoạn xen 7 bp
SEQ ID NO: 100	nucleotit	alen với đoạn xóa 1280 bp
SEQ ID NO: 101	nucleotit	alen với đoạn xóa 1373 bp
SEQ ID NO: 102	nucleotit	alen với đoạn xóa 11 bp
SEQ ID NO: 103	nucleotit	alen với đoạn xen 2 bp và với đoạn xóa 377 bp
SEQ ID NO: 104	nucleotit	alen với đoạn xóa 124 bp
SEQ ID NO: 105	nucleotit	alen với đoạn xóa 123 bp
SEQ ID NO: 106	nucleotit	alen với đoạn xen 1 bp
SEQ ID NO: 107	nucleotit	alen với đoạn xóa 130 bp
SEQ ID NO: 108	nucleotit	alen với đoạn xóa 132 bp
SEQ ID NO: 109	nucleotit	alen với đoạn xóa 1467 bp

SEQ ID NO: 110	nucleotit	alen với đoạn xóa 1930 bp trong exon 6, đoạn xóa 129 bp trong exon 7, và đoạn xen 12 bp
SEQ ID NO: 111	nucleotit	alen với đoạn xóa 28 bp
SEQ ID NO: 112	nucleotit	alen với đoạn xóa 1387 bp
SEQ ID NO: 113	nucleotit	alen với đoạn xóa 1382 bp và với đoạn xen 11 bp
SEQ ID NO: 114	nucleotit	alen với đoạn xóa 1720 bp
SEQ ID NO: 115	nucleotit	Trình tự xen của SEQ. 99
SEQ ID NO: 116	nucleotit	Trình tự xen của SEQ.110
SEQ ID NO: 117	nucleotit	Trình tự xen của SEQ.113
SEQ ID NO: 118	nucleotit	Trình tự hoán đổi miền
SEQ ID NO: 119	nucleotit	alen với đoạn xóa 452 bp
SEQ ID NO: 120	peptit	CD163 SRCR 5 của lợn
SEQ ID NO: 121	peptit	Chất tương tự CD163L1 SRCR 8 người

YÊU CẦU BẢO HỘ

1. Động vật là lợn hoặc thế hệ con của nó chứa trình tự nhiễm sắc thể sửa đổi ở ít nhất một alen của gen mã hóa cho protein CD163, trong đó trình tự nhiễm sắc thể sửa đổi này ở ít nhất một alen bao gồm đoạn xóa và đoạn xóa này cấu thành từ đoạn xóa 11 cặp bazơ từ nucleotit 3.137 đến nucleotit 3.147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

2. Động vật hoặc thế hệ con theo điểm 1, trong đó việc sửa đổi làm giảm tính nhạy của động vật hoặc thế hệ con, với việc nhiễm trùng bởi virut, khi so sánh với tính nhạy của động vật hoặc thế hệ con mà không bao gồm trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163 với việc nhiễm trùng bởi virut.

3. Động vật hoặc thế hệ con theo điểm 2, trong đó virut bao gồm virut gây hội chứng hô hấp và sinh sản ở lợn (PRRSV).

4. Động vật hoặc thế hệ con theo điểm 3, trong đó việc sửa đổi làm giảm tính nhạy của động vật hoặc thế hệ con với virut PRRSV Typ 1, PRRSV Typ 2, hoặc với cả virut PRRSV Typ 1 và Typ 2.

5. Động vật hoặc thế hệ con theo điểm 4, trong đó việc sửa đổi làm giảm tính nhạy của động vật hoặc thế hệ con với thể phân lập PRRSV được chọn từ nhóm bao gồm NVSL 97-7895, KS06-72109, P129, VR2332, CO90, AZ25, MLV-ResPRRS, KS62-06274, KS483 (SD23983), CO84, SD13-15, Lelystad, 03-1059, 03-1060, SD01-08, 4353PZ, và tổ hợp bất kỳ của chúng.

6. Động vật hoặc thế hệ con theo điểm 1, trong đó động vật hoặc thế hệ con này bao gồm động vật hoặc thế hệ con chỉnh sửa gen.

7. Động vật hoặc thế hệ con theo điểm 6, trong đó động vật đã được chỉnh sửa gen sử dụng endonucleaza tìm nhà được thiết kế, trong đó endonucleaza tìm nhà bao gồm hệ trình tự sắp xếp kiểu lặp đi lặp lại xen khoảng trống (CRISPR) /Cas, men nucleaza tác động giống chất kích hoạt phiên mã (TALEN), nucleaza ngón tay kẽm (ZFN), protein dung hợp recombinaza, meganucleaza, hoặc tổ hợp của chúng.

8. Động vật hoặc thế hệ con theo điểm 7, trong đó động vật đã được chỉnh sửa gen sử dụng hệ CRISPR/Cas.

9. Động vật hoặc thế hệ con theo điểm 8, trong đó hệ CRISPR/Cas bao gồm hệ CRISPR/Cas9.

10. Động vật hoặc thế hệ con theo điểm 8, trong đó động vật hoặc thế hệ con chỉnh sửa gen nêu trên thể hiện tính kháng tăng đối với PRRSV khi so sánh với động vật không được chỉnh sửa.

11. Động vật hoặc thế hệ con theo điểm 1, trong đó đoạn xóa này dẫn đến sự sản xuất ra hoặc hoạt tính của protein CD163 bị giảm, khi so sánh với sự sản xuất ra hoặc hoạt tính của protein CD163 ở động vật hoặc thế hệ con mà không có đoạn xóa này.

12. Động vật hoặc thế hệ con theo điểm 1, trong đó đoạn xóa này dẫn đến sự tạo ra protein CD163 cơ bản không có chức năng bởi động vật hoặc thế hệ con.

13. Động vật là lợn hoặc thế hệ con theo điểm 1, trong đó động vật hoặc thế hệ con bao gồm trình tự nhiễm sắc thể có ít nhất 80% mức tương đồng trình tự so với SEQ ID NO: 47 ở các vùng của trình tự nhiễm sắc thể nêu trên ngoài đoạn xóa.

14. Động vật hoặc thế hệ con theo điểm 1, chứa trình tự nhiễm sắc thể bao gồm SEQ ID NO: 102.

15. Động vật hoặc thế hệ con theo điểm 1, trong đó động vật hoặc thế hệ con là động vật non hoặc động vật trưởng thành.

16. Động vật hoặc thế hệ con theo điểm 1, trong đó động vật hoặc thế hệ con là phôi.

17. Động vật hoặc thế hệ con theo điểm 1, trong đó động vật hoặc thế hệ con là dị hợp tử về trình tự nhiễm sắc thể sửa đổi.

18. Động vật hoặc thế hệ con theo điểm 1, trong đó động vật hoặc thế hệ con là đồng hợp tử về trình tự nhiễm sắc thể sửa đổi.

19. Động vật hoặc thế hệ con theo điểm 1, trong đó động vật hoặc thế hệ con bao gồm: đoạn xóa 11 cặp bazơ từ nucleotit 3.137 đến nucleotit 3.147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và một đoạn xen vào alen khác của gen mã hóa cho protein CD163, đoạn xen này cấu thành từ đoạn xen 7 cặp bazơ giữa nucleotit 3.148 và nucleotit 3.149 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

20. Động vật hoặc thế hệ con theo điểm 19, trong đó đoạn xen 7 cặp bazơ giữa nucleotit 3.148 và nucleotit 3.149 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 cấu thành từ trình tự TACTACT (SEQ ID NO: 115).

21. Động vật hoặc thế hệ con theo điểm 1, trong đó động vật hoặc thế hệ con bao gồm: đoạn xóa 11 cặp bazơ từ nucleotit 3.137 đến nucleotit 3.147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn

xóa trong alen khác của gen mã hóa cho protein CD163, đoạn xóa này cấu thành từ đoạn xóa 1280 cặp bazơ từ nucleotit 2.818 đến nucleotit 4.097 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

22. Động vật hoặc thế hệ con theo điểm 1, trong đó động vật hoặc thế hệ con bao gồm: đoạn xóa 11 cặp bazơ từ nucleotit 3.137 đến nucleotit 3.147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và các đoạn xóa trong alen khác của gen mã hóa cho protein CD163, các đoạn xóa cấu thành từ: đoạn xóa 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2.417 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 12 cặp bazơ bắt đầu tại nucleotit 488, và đoạn xóa 129 cặp bazơ trong exon 7 từ nucleotit 3.044 đến nucleotit 3.172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

23. Động vật hoặc thế hệ con theo điểm 22, trong đó đoạn xen 12 cặp bazơ bắt đầu từ nucleotit 488 cấu thành từ trình tự TGTGGAGAATTC (SEQ ID NO: 116).

24. Động vật hoặc thế hệ con theo điểm 1, trong đó động vật hoặc thế hệ con bao gồm: đoạn xóa 11 cặp bazơ từ nucleotit 3.137 đến nucleotit 3.147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn xóa trong alen khác của gen mã hóa cho protein CD163, đoạn xóa này cấu thành từ đoạn xóa 1467 cặp bazơ từ nucleotit 2.431 đến nucleotit 3.897 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

25. Động vật hoặc thế hệ con theo điểm 1, trong đó động vật hoặc thế hệ con bao gồm: đoạn xóa 11 cặp bazơ từ nucleotit 3.137 đến nucleotit 3.147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và một đoạn xen và đoạn xóa trong alen khác của gen mã hóa cho protein CD163, đoạn xen này cấu thành từ đoạn xen 2 cặp bazơ giữa nucleotit 3.149 và 3.150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 và đoạn xóa này cấu thành từ 377 cặp bazơ đoạn xóa từ nucleotit 2.573 đến nucleotit 2.949 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

26. Động vật hoặc thế hệ con theo điểm 25, trong đó đoạn xen 2 cặp bazơ giữa nucleotit 3.149 và 3.150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 cấu thành từ dinucleotit AG.

27. Tế bào lợn chứa trình tự nhiễm sắc thể sửa đổi ở ít nhất một alen của gen mã hóa cho protein CD163, trong đó trình tự nhiễm sắc thể sửa đổi này ở ít nhất một alen bao

gồm đoạn xóa và đoạn xóa này cấu thành từ đoạn xóa 11 cặp bazơ từ nucleotit 3.137 đến nucleotit 3.147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

28. Tế bào theo điểm 27, trong đó việc sửa đổi làm giảm tính nhạy của tế bào với việc nhiễm trùng bởi virut, khi so sánh với tính nhạy của tế bào mà không bao gồm trình tự nhiễm sắc thể sửa đổi ở gen mã hóa cho protein CD163 với việc nhiễm trùng bởi virut.

29. Tế bào theo điểm 28, trong đó virut bao gồm virut gây hội chứng hô hấp và sinh sản ở lợn (PRRSV).

30. Tế bào theo điểm 29, trong đó việc sửa đổi làm giảm tính nhạy của tế bào với virut PRRSV Typ 1, PRRSV Typ 2, hoặc với cả virut PRRSV Typ 1 và Typ 2.

31. Tế bào theo điểm 30, trong đó việc sửa đổi làm giảm tính nhạy của tế bào với thể phân lập PRRSV được chọn từ nhóm bao gồm NVSL 97-7895, KS06-72109, P129, VR2332, CO90, AZ25, MLV-ResPRRS, KS62-06274, KS483 (SD23983), CO84, SD13-15, Lelystad, 03-1059, 03-1060, SD01-08, 4353PZ, và tổ hợp bất kỳ của chúng.

32. Tế bào theo điểm 27, trong đó tế bào bao gồm tế bào chỉnh sửa gen.

33. Tế bào theo điểm 32, trong đó tế bào đã được chỉnh sửa gen sử dụng endonucleaza tìm nhà được thiết kế, trong đó endonucleaza tìm nhà bao gồm hệ trình tự sắp xếp kiểu lặp đi lặp lại xen khoảng trống (CRISPR) /Cas, men nucleaza tác động giống chất kích hoạt phiên mã (TALEN), nucleaza ngón tay kẽm (ZFN), protein dung hợp recombinaza, meganucleaza, hoặc tổ hợp của chúng.

34. Tế bào theo điểm 33, trong đó tế bào đã được chỉnh sửa gen sử dụng hệ CRISPR/Cas.

35. Tế bào theo điểm 34, trong đó hệ CRISPR/Cas bao gồm hệ CRISPR/Cas9.

36. Tế bào theo điểm 34, trong đó tế bào chỉnh sửa gen nêu trên thể hiện tính kháng tăng đối với PRRSV khi so sánh với động vật không được chỉnh sửa.

37. Tế bào theo điểm 27, trong đó đoạn xóa này dẫn đến sự sản xuất ra hoặc hoạt tính của protein CD163 bị giảm, khi so sánh với sự sản xuất ra hoặc hoạt tính của protein CD163 ở tế bào mà không có đoạn xóa.

38. Tế bào theo điểm 27, trong đó đoạn xóa này dẫn đến sự tạo ra protein CD163 cơ bản không có chức năng bởi tế bào.

39. Tế bào theo điểm 27, trong đó tế bào bao gồm trình tự nhiễm sắc thể có ít nhất 80% mức tương đồng trình tự so với SEQ ID NO: 47 ở các vùng của trình tự nhiễm sắc thể nêu trên ngoài đoạn xóa.

40. Tế bào theo điểm 27, trong đó tế bào này là tinh trùng.

41. Tế bào theo điểm 27, trong đó tế bào này là tế bào trứng.

42. Tế bào theo điểm 41, trong đó tế bào trứng là trứng đã thụ tinh.

43. Tế bào theo điểm 27, trong đó tế bào này là tế bào sinh dưỡng.

44. Tế bào theo điểm 43, trong đó tế bào sinh dưỡng bao gồm nguyên bào sợi.

45. Tế bào theo điểm 44, trong đó nguyên bào sợi bao gồm nguyên bào sợi thai.

46. Tế bào theo điểm 27, trong đó tế bào này chứa trình tự nhiễm sắc thể bao gồm SEQ ID NO: 102.

47. Tế bào theo điểm 27, trong đó tế bào này bao gồm tế bào có nguồn gốc từ động vật non hoặc tế bào có nguồn gốc động vật trưởng thành.

48. Tế bào theo điểm 27, trong đó tế bào này bao gồm tế bào phôi.

49. Tế bào theo điểm 27, trong đó tế bào là dị hợp tử về trình tự nhiễm sắc thể sửa đổi.

50. Tế bào theo điểm 48, trong đó phôi hoặc tế bào phôi là dị hợp tử về trình tự nhiễm sắc thể sửa đổi.

51. Tế bào theo điểm 27, trong đó tế bào là đồng hợp tử về trình tự nhiễm sắc thể sửa đổi.

52. Tế bào theo điểm 27, trong đó tế bào này bao gồm: đoạn xóa 11 cặp bazơ từ nucleotit 3.137 đến nucleotit 3.147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và một đoạn xen vào alen khác của gen mã hóa cho protein CD163, đoạn xen này cấu thành từ đoạn xen 7 cặp bazơ giữa nucleotit 3.148 và nucleotit 3.149 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

53. Tế bào theo điểm 52, trong đó đoạn xen 7 cặp bazơ giữa nucleotit 3.148 và nucleotit 3.149 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 cấu thành từ trình tự TACTACT (SEQ ID NO: 115).

54. Tế bào theo điểm 27, trong đó tế bào này bao gồm: đoạn xóa 11 cặp bazơ từ nucleotit 3.137 đến nucleotit 3.147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn xóa trong alen khác của gen mã hóa cho protein CD163, đoạn xóa này cấu thành từ đoạn xóa 1280 cặp bazơ từ nucleotit 2.818 đến nucleotit 4.097 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

55. Tế bào theo điểm 27, trong đó tế bào này bao gồm: đoạn xóa 11 cặp bazơ từ nucleotit 3.137 đến nucleotit 3.147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và các đoạn xóa trong alen khác của gen mã

hóa cho protein CD163, các đoạn xóa cấu thành từ: đoạn xóa 1930 cặp bazơ từ nucleotit 488 đến nucleotit 2.417 khi so sánh với trình tự tham chiếu SEQ ID NO: 47, trong đó trình tự khuyết được thay thế bằng đoạn xen 12 cặp bazơ bắt đầu tại nucleotit 488, và đoạn xóa 129 cặp bazơ trong exon 7 từ nucleotit 3.044 đến nucleotit 3.172 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

56. Tế bào theo điểm 55, trong đó đoạn xen 12 cặp bazơ bắt đầu từ nucleotit 488 cấu thành từ trình tự TGTGGAGAATTC (SEQ ID NO: 116).

57. Tế bào theo điểm 27, trong đó tế bào này bao gồm: đoạn xóa 11 cặp bazơ từ nucleotit 3.137 đến nucleotit 3.147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và đoạn xóa trong alen khác của gen mã hóa cho protein CD163, đoạn xóa này cấu thành từ đoạn xóa 1467 cặp bazơ từ nucleotit 2.431 đến nucleotit 3.897 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong alen khác của gen mã hóa cho protein CD163.

58. Tế bào theo điểm 27, trong đó tế bào này bao gồm: đoạn xóa 11 cặp bazơ từ nucleotit 3.137 đến nucleotit 3.147 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 trong một alen của gen mã hóa cho protein CD163; và một đoạn xen và đoạn xóa trong alen khác của gen mã hóa cho protein CD163, đoạn xen này cấu thành từ đoạn xen 2 cặp bazơ giữa nucleotit 3.149 và 3.150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 và đoạn xóa này cấu thành từ 377 cặp bazơ đoạn xóa từ nucleotit 2.573 đến nucleotit 2.949 khi so sánh với trình tự tham chiếu SEQ ID NO: 47.

59. Tế bào theo điểm 58, trong đó đoạn xen 2 cặp bazơ giữa nucleotit 3.149 và 3.150 khi so sánh với trình tự tham chiếu SEQ ID NO: 47 cấu thành từ dinucleotit AG.

Fig. 1

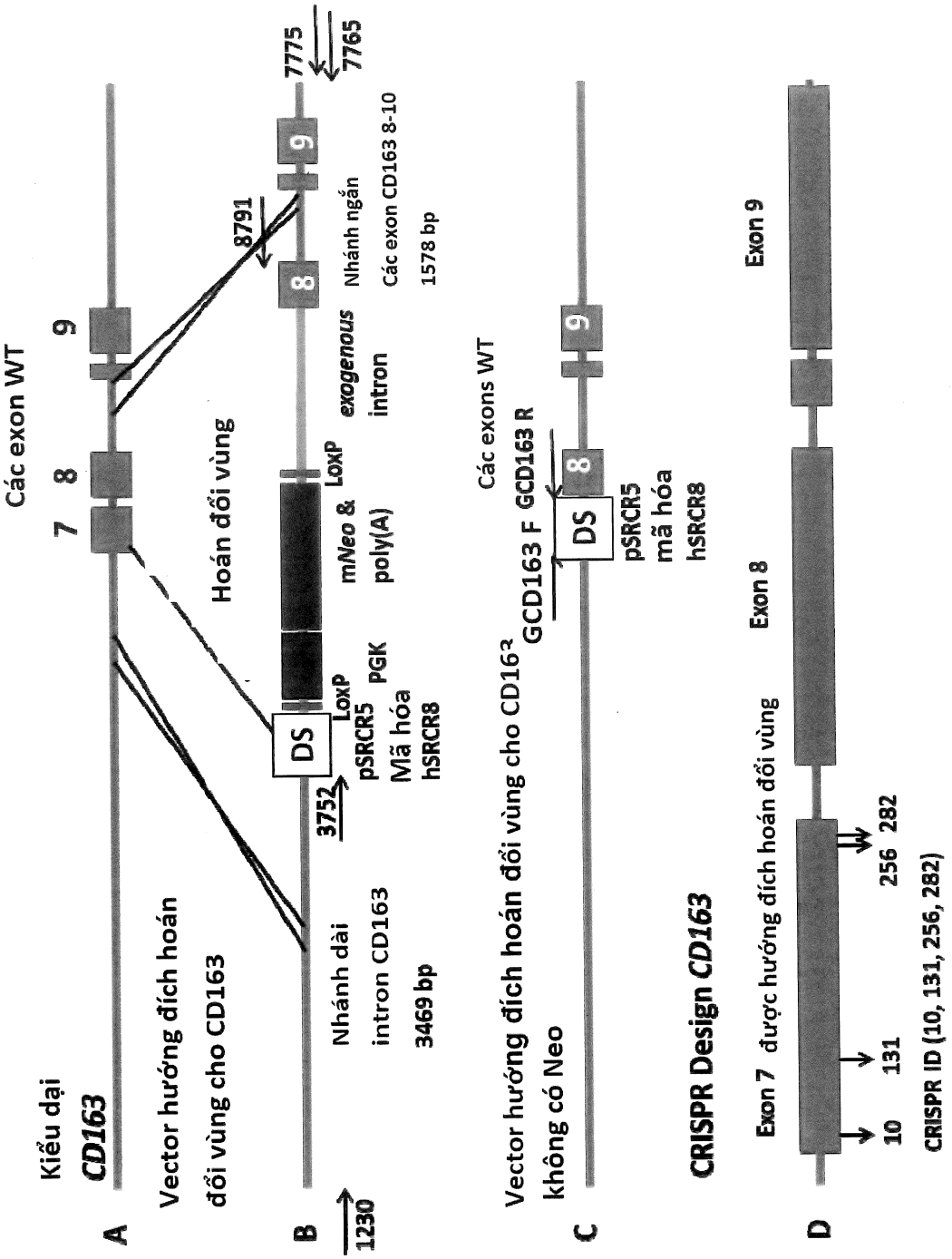

A
Kiểu dại
CD163
Các exon WT
7 8 9
Vector hướng đích hoán đổi vùng cho CD163
B
Nhánh dài intron CD163
3469 bp
1230
3752
DS
LoxP
pSRCR5
Mã hóa
hSRCR8
PGK
mNeo & poly(A)
LoxP
exogenous intron
8791
8 9
7775
7765
Nhánh ngắn
Các exon CD163 8-10
1578 bp
Hoán đổi vùng
C
Vector hướng đích hoán đổi vùng cho CD163
không có Neo
Các exons WT
GCD163 F
GCD163 R
DS
8 9
pSRCR5
mã hóa
hSRCR8
D
CRISPR Design CD163
Exon 7
được hướng đích hoán đổi vùng
Exon 8
Exon 9
10 131 256 282
CRISPR ID (10, 131, 256, 282)

Fig. 2

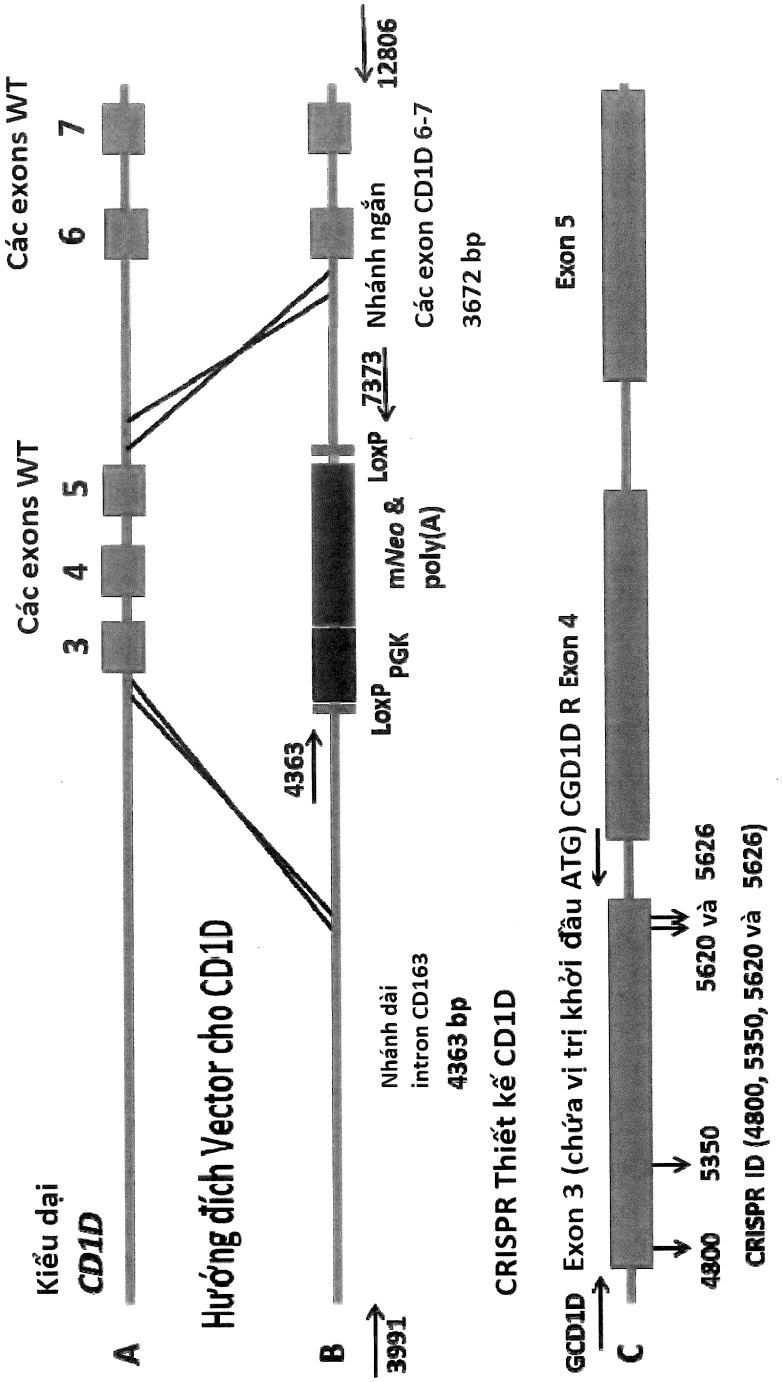

Các exons WT
3 4 5
Các exons WT
6 7
Kiểu dại
CD1D
A
Hướng đích Vector cho CD1D
4363
LoxP
PGK
mNeo &
poly(A)
LoxP
7373
12806
Nhánh ngắn
Các exon CD1D 6-7
3672 bp
Exon 5
B
3991
Nhánh dài
intron CD163
4363 bp
CRISPR Thiết kế CD1D
GCD1D
Exon 3 (chứa vị trí khởi đầu ATG) CGD1D R Exon 4
C
4800
5350
5620 và 5626
CRISPR ID (4800, 5350, 5620 và 5626)

Fig. 3

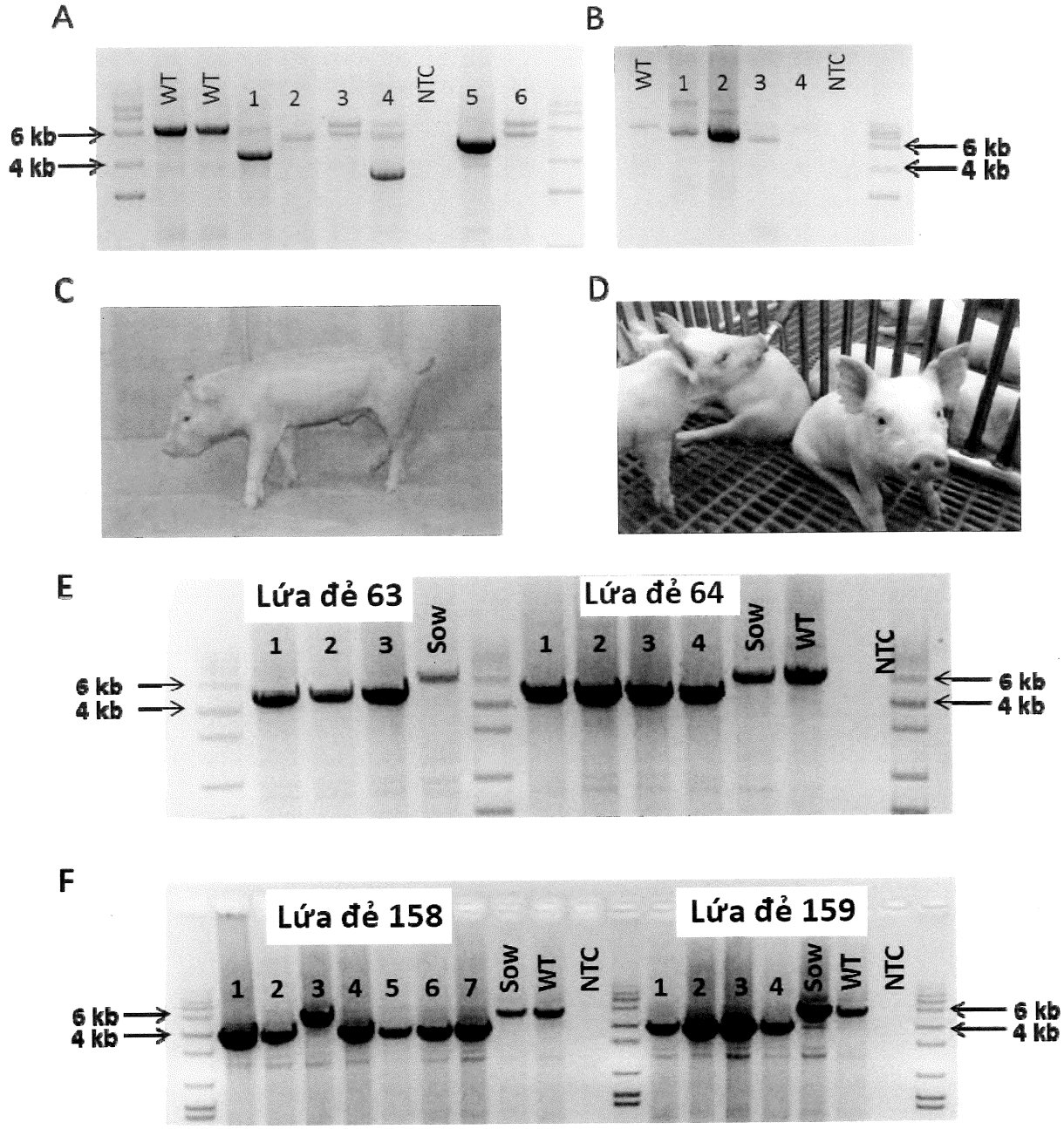
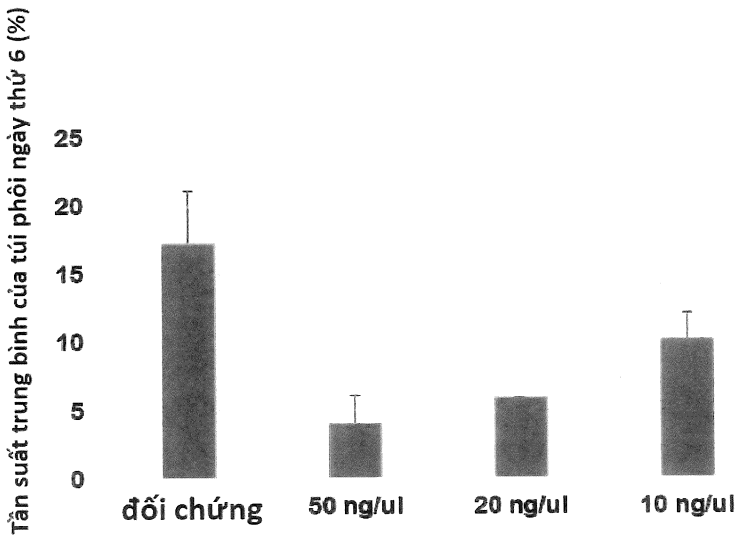
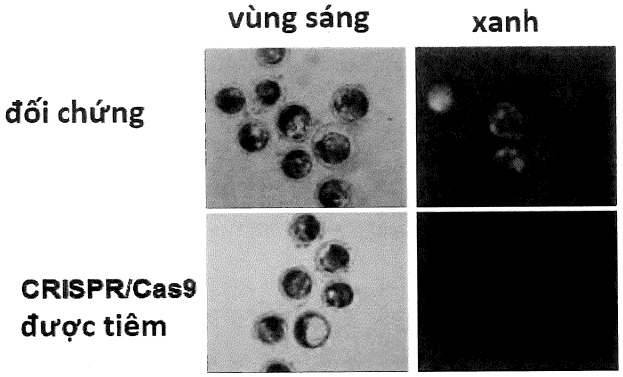

A
WT
WT
1
2
3
4
NTC
5
6
6 kb
4 kb
B
WT
1
2
3
4
NTC
6 kb
4 kb
C
D
E
Lứa đẻ 63
Lứa đẻ 64
1
2
3
Sow
1
2
3
4
Sow
WT
NTC
6 kb
4 kb
F
Lứa đẻ 158
Lứa đẻ 159
1
2
3
4
5
6
7
Sow
WT
NTC
1
2
3
4
Sow
WT
NTC
6 kb
4 kb

Fig. 4

A



B



C

	SEQ ID NO	
WT	16	GGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATC
#1	17	GGTCGCCACCATGG**CCA**TGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCC(+3 bp)
#2	18	GGTCGCCACCATGG------TGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGT(-6 bp)
#3	19	GGTCGCCACCATGGTTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCAT(+1 bp)

Fig. 5

A

	SEQ ID NO	
WT	20	TGCAGGGAACTACAGTGCGGCACTGTGGTTCCCTCCTGGGGG
#1-1	21	TGCAGGGAACTACAGTGCGGCACTG(+16 bp)TGGTTCCCTCCTGGGGG
#1-4	22	TGCAGGGAACTACAGTGCGGC--TGTGGTTCCCTCCTGGGGG
#2-2	23	TGCAGGGAACTACAGTGCGG-AACTACTGTGGTTCCCTCCTGGGGG

B

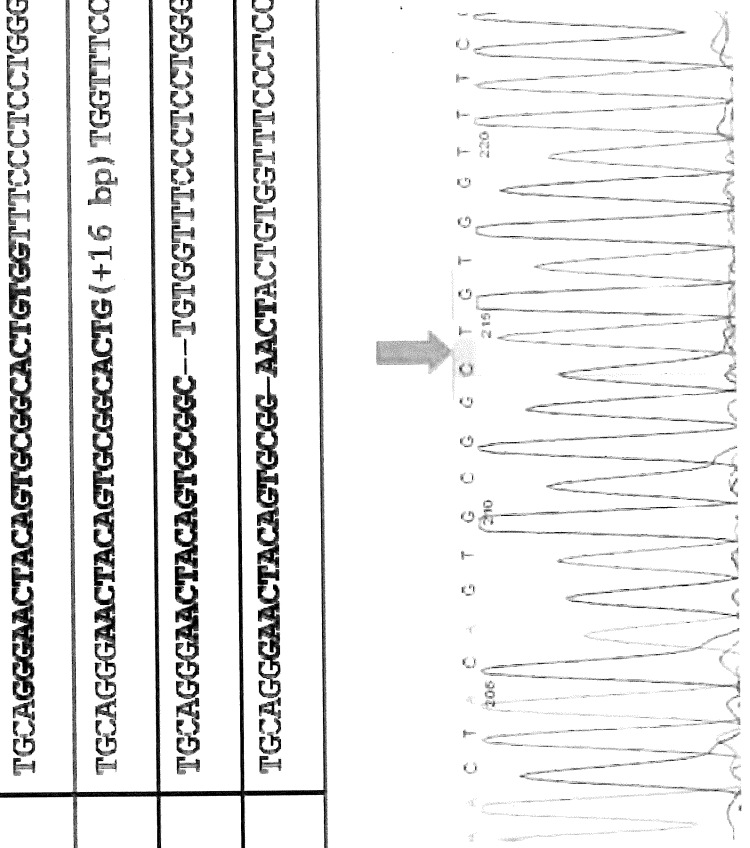
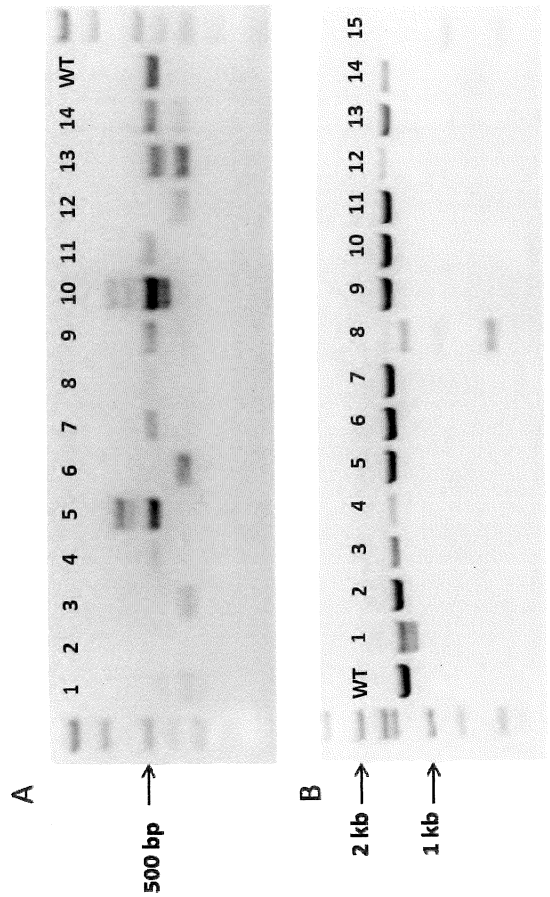


Fig. 6

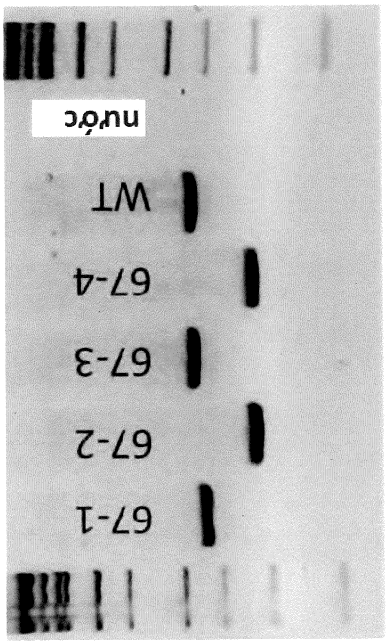


C

	CD163	*SEQ ID NO*	*eGFP*	*SEQ ID NO*
WT	GAAACCCAGGCTGGTTGGAGGGGACATTCCC	24	GGTCGCCACCATGGTGAGCAAGGGCGAGGAG	28
#1	GAAACCCAGGCTGG-------GGACATTCCC	25	GGTCGCCACCATGGCTGAGCAAGGGCGAGGAG GGTCGCCACCATGGTGAG--AGGGCGAGGAG	29 30
#2	-(Δ83 bp)----------AGGGGACATTCCC	26	GGTCGCCACCATGGTTGAGCAAGGGCGAGGAG	31
#3	GAAACCCA----------TTCCC	27	GGTCGCCACCATGGTGAGCAAGGGCGAGGAG+ (17 bp)	32

Fig. 7

A



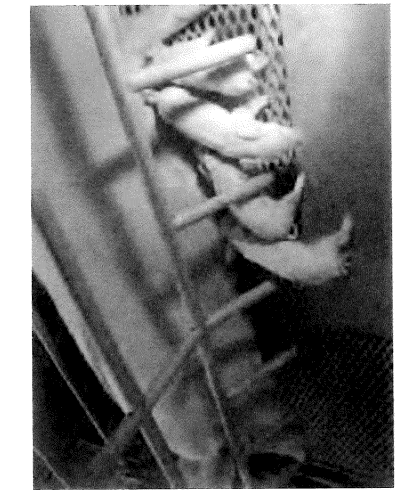
B

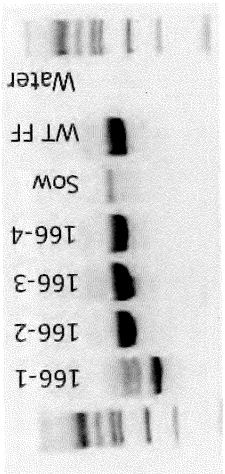


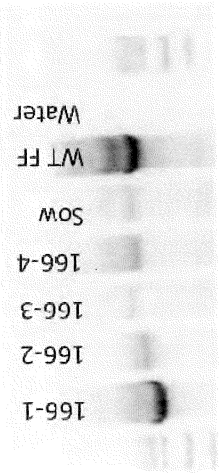
C

	SEQ ID NO	
WT	33	TGCTGTGCAGGGAACTACAGTGCGGCACTGTGGTTTCCCTCCTGGGGGG
#67-1	34	TGCTGTGCAGGGAACT----------CTGTGGTTTCCCTCCTGGGGGG
#67-2	35	-(Δ124 bp)----------CTGTGGTTTCCCTCCTGGGGGG
	36	-(Δ123 bp)----------ACTGTGGTTTCCCTCCTGGGGGG
#67-3	37	TGCTGTGCAGGGAACTACAGTGCGGCAACTGTGGTTTCCCTCCTGGGGGG
#67-4	38	-(Δ130 bp)----------TCCTGGGGGG
	39	-(Δ132 bp)----------CTGGGGGG

Fig. 8







D

	SEQ ID NO	
WT	40	AGAGAGCAGAGCCAGCGACTCGCCCAGCGACatgGGGTACCTGCCGTTTGTG
#166-1		---------- (-632 bp deletion)
	41	AGAGAGCAGAGCCAGCGACTCGCCCAGCGAGat---------- (-27 bp deletion)
	42	AGAGAGCAGAGCCAGCGACTCGCCCAGCGA---------- (-24 bp deletion)
#166-2	43	AGAGCCAGCCTCGCCCAGCAGGGGTACCatgGGGTACCTGCCGTTTGTGT (AG mismatch+Δ6 bp) ---------- (-1598 bp deletion)
#166-3	44	AGAGAGCAGAGCCAGCGACTCGCCCAGCGAGCagtGGGTACCTGCCGTTTGTG (+Δ1 bp)
	45	AGAGAGCAGAGCCAGCGACTCGCCCAGCGATCagtGGGTACCTGCCGTTTGTG (+Δ1 bp)
#166-4	46	AGAGAGCAGAGCCAGCGACTCGCCCAGCGAACatgGGGTACCTGCCGTTTGTG (+Δ1 bp)

Fig. 9

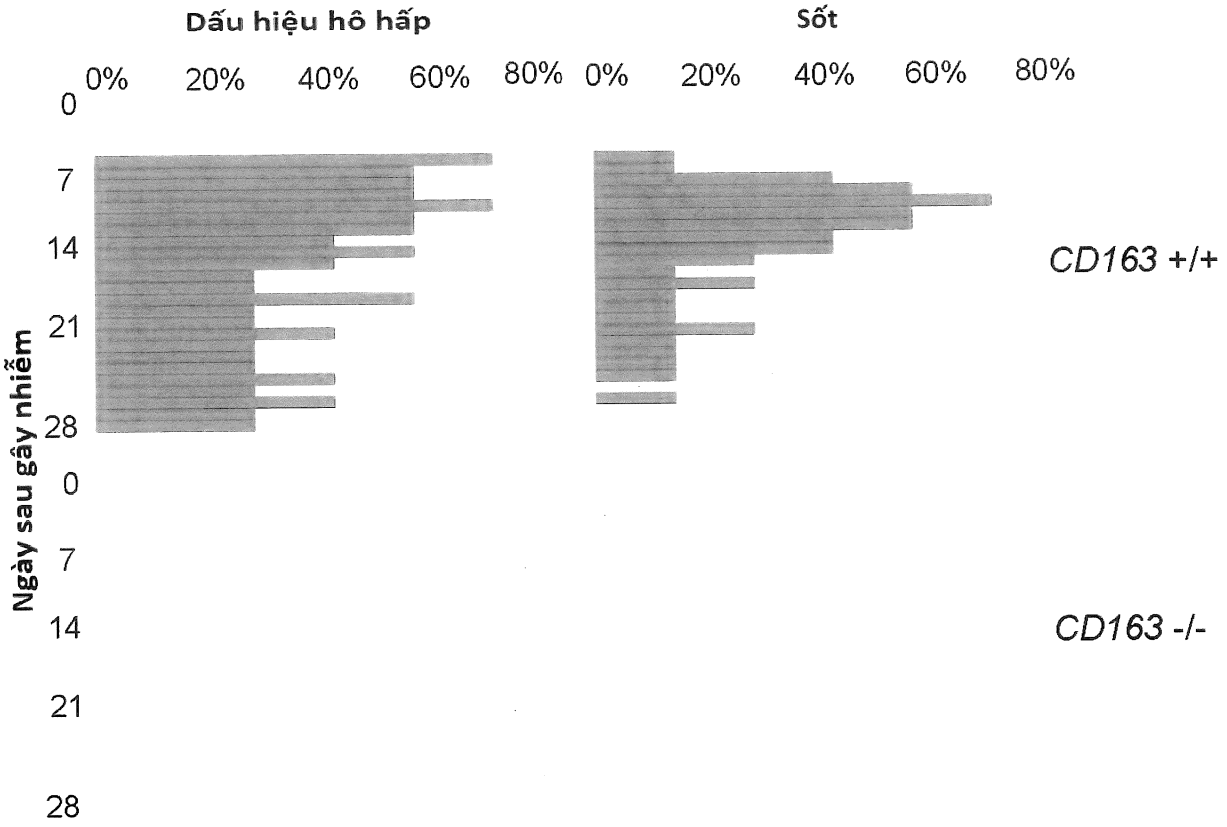

Dấu hiệu hô hấp
Sốt
0%
20%
40%
60%
80%
0
7
14
21
28
Ngày sau gây nhiễm
CD163 +/+
CD163 -/-

Fig. 10

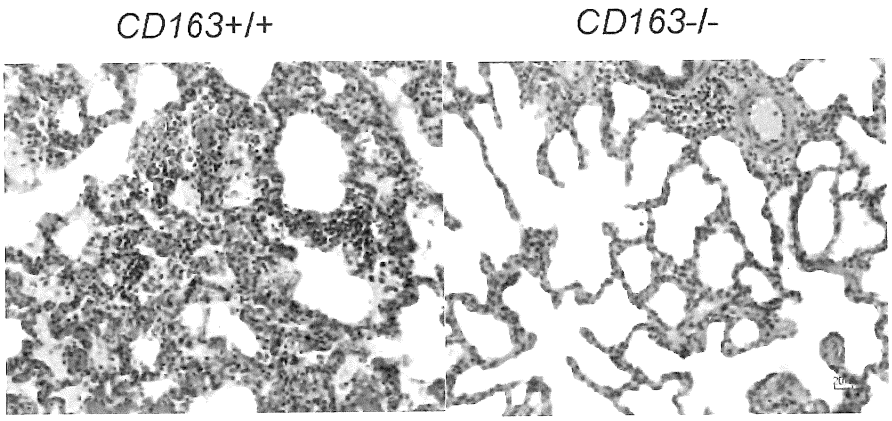


Fig. 11

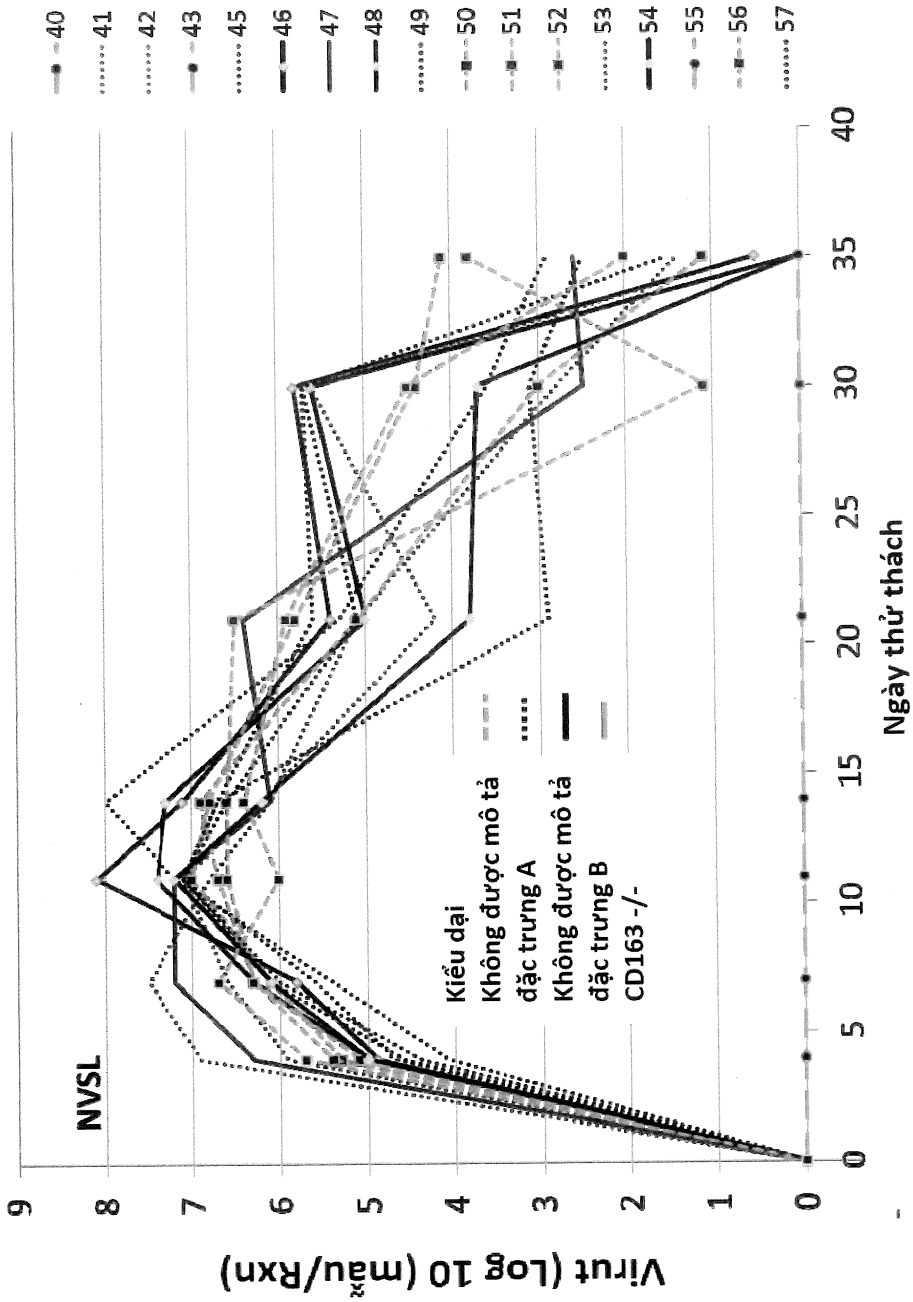

40
41
42
43
45
46
47
48
49
50
51
52
53
54
55
56
57
NVSL
Kiểu dại
Không được mô tả đặc trưng A
Không được mô tả đặc trưng B
CD163 -/-
Virut (Log 10 (mẫu/Rxn)
Ngày thử thách
0
5
10
15
20
25
30
35
40
1
2
3
4
6
7
8
9

Fig. 12

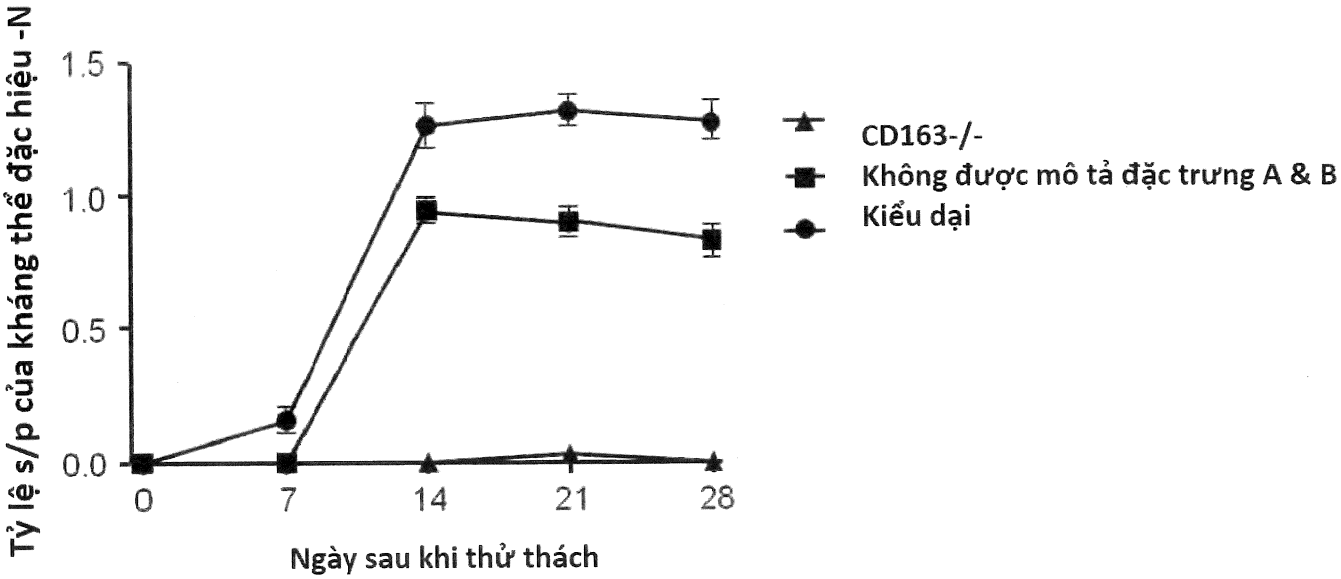

Tỷ lệ s/p của kháng thể đặc hiệu -N
1.5
1.0
0.5
0.0
0
7
14
21
28
Ngày sau khi thử thách
CD163-/-
Không được mô tả đặc trưng A & B
Kiểu dại

FIG. 13

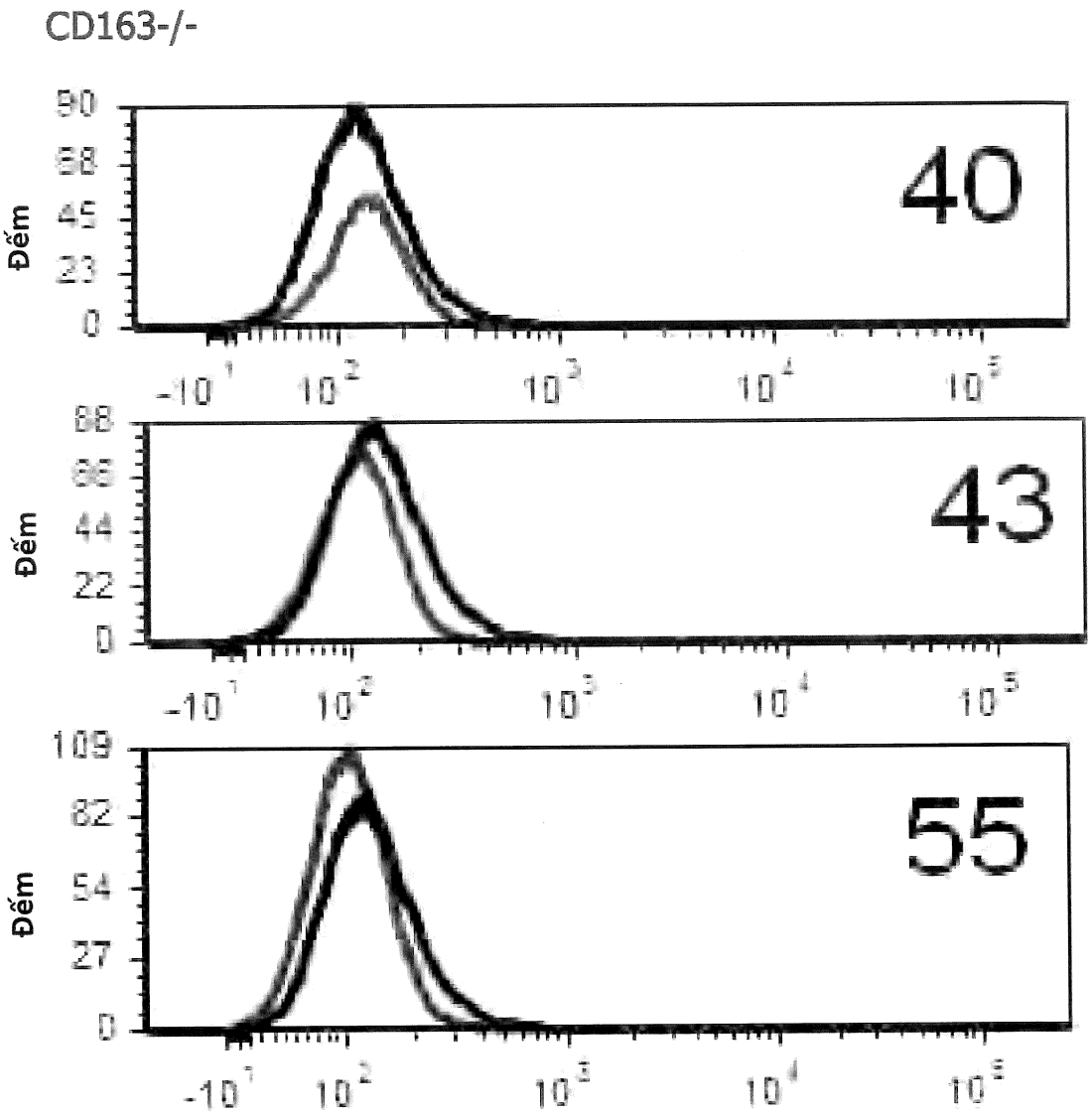


FIG. 13 tiếp

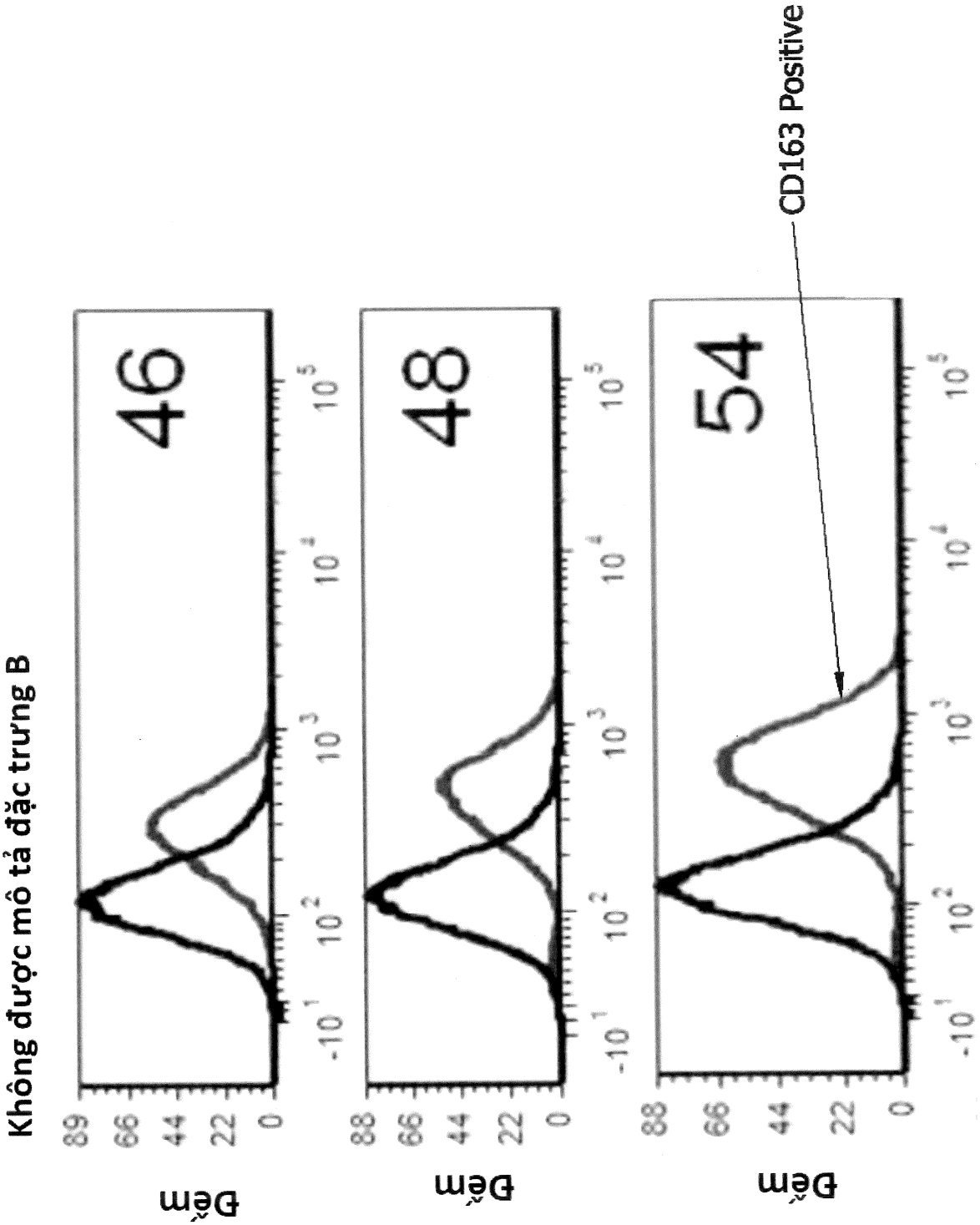

Không được mô tả đặc trưng B
CD163 Positive
46
48
54
Đếm
Đếm
Đếm

FIG. 13 tiếp

Không được mô tả đặc trưng A

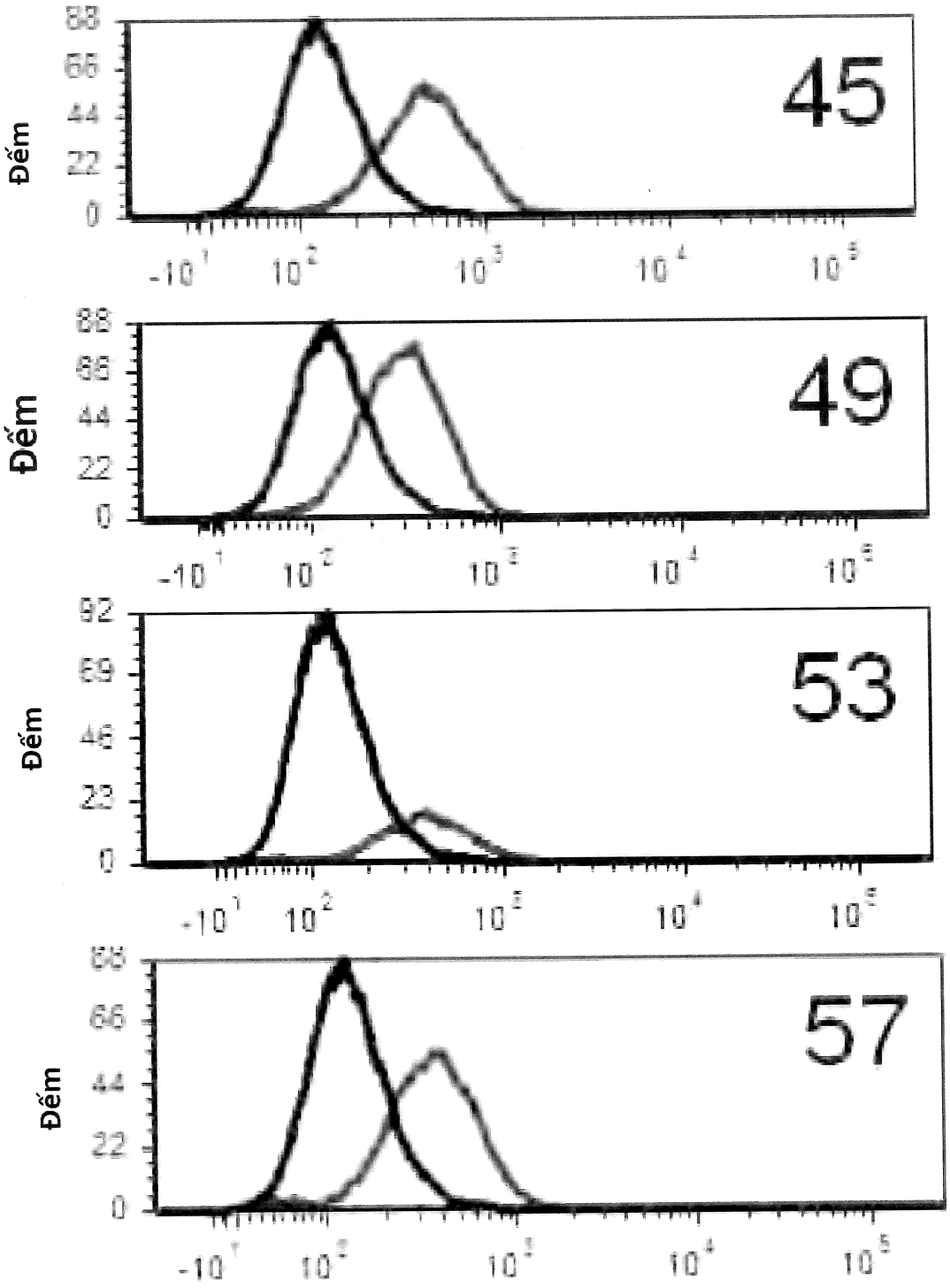


FIG. 13 tiếp.

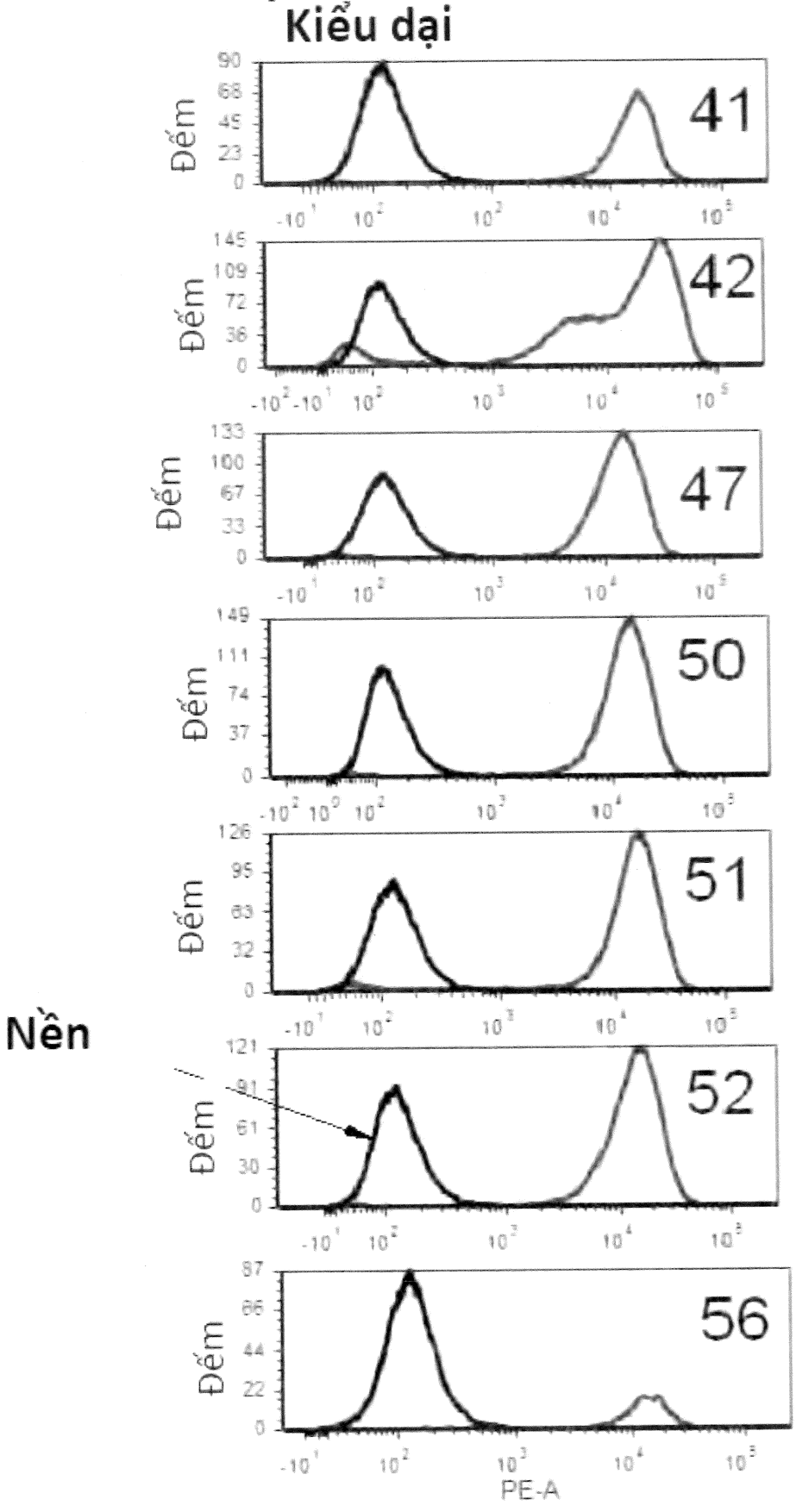


Fig. 14

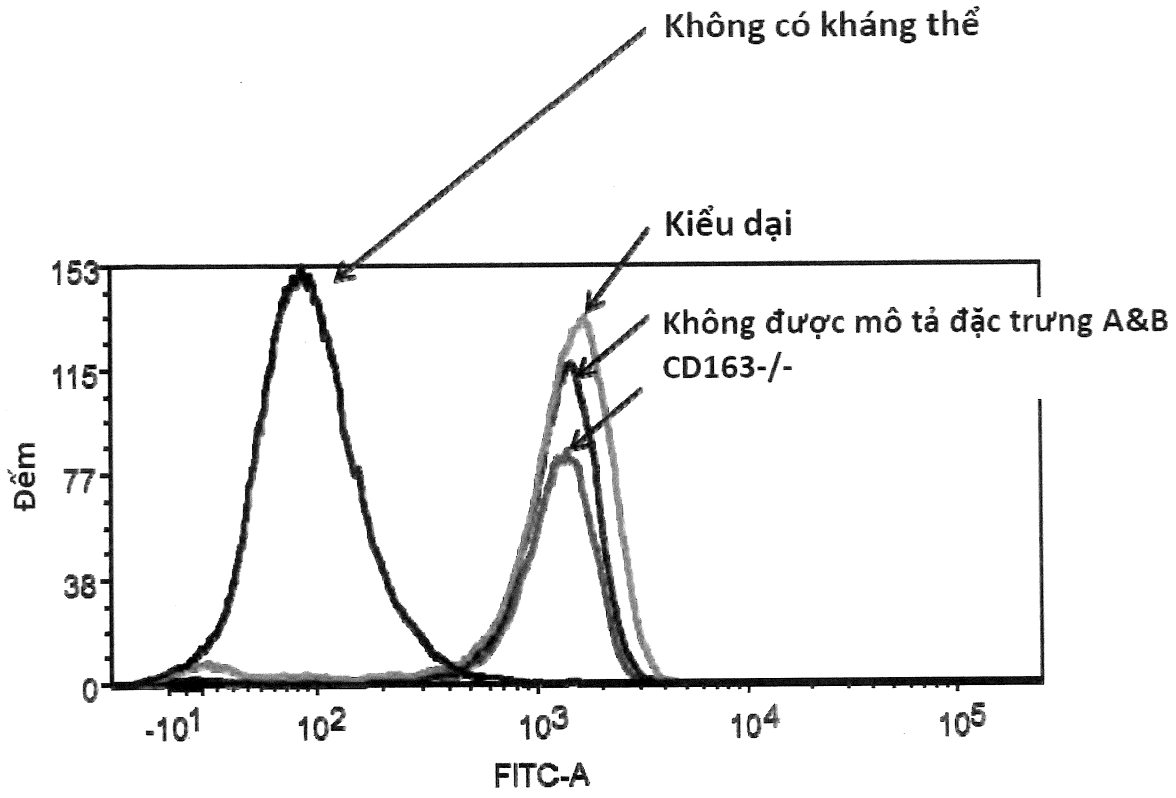

Không có kháng thể
Kiểu dại
Không được mô tả đặc trưng A&B
CD163-/-
153
115
77
38
0
Đếm
-10^1
10^2
10^3
10^4
10^5
FITC-A

Fig. 15

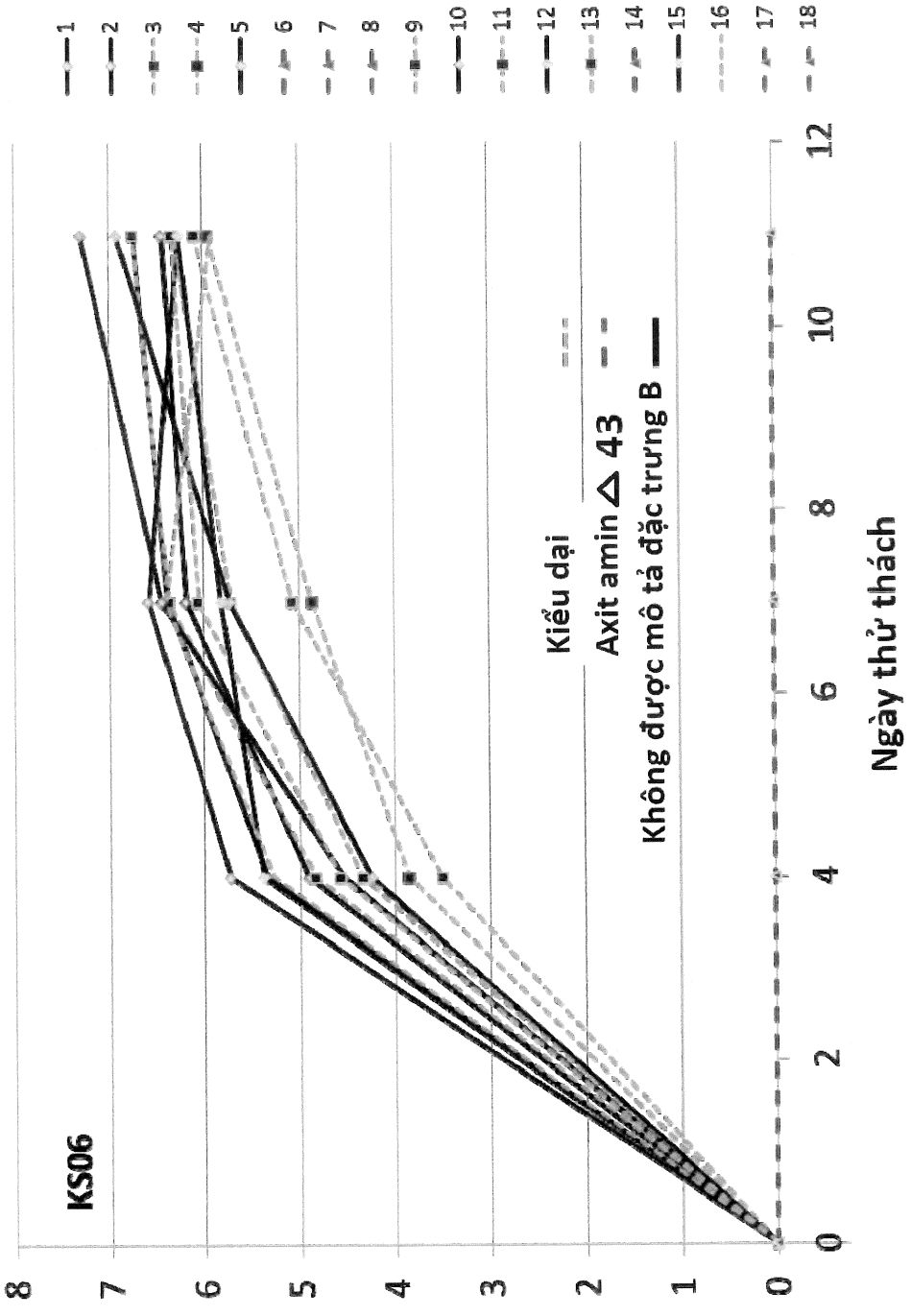

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
KS06
Kiểu dại
Axit amin Δ 43
Không được mô tả đặc trưng B
Ngày thử thách
0
2
4
6
8
10
12
0
1
2
3
4
5
6
7
8

FIG. 16

SEQ ID NO. 47

```
LOCUS       WT_CD163_Referen        4990 bp ds-DNA     linear
DEFINITION  Reference CD163 gene:  3000 bp upstream of exon 7 to the last base of
exon 10
ACCESSION
VERSION
SOURCE      .
  ORGANISM  . pig
COMMENT
COMMENT     ApEinfo:methylated:1
FEATURES             Location/Qualifiers
    misc_feature     1..3000
                     /label=intron 6

    misc_feature     3001..3315
                     /label=exon 7

    misc_feature     3316..3412
                     /label=intron 7

     misc_feature    3413..3727
                     /label=exon 8

     misc_feature    3728..4501
                     /label=intron 8

     misc_feature    4502..4594
                     /label=exon 9

     misc_feature    4595..4676
                     /label=intron 9

     misc_feature    4676..4989
                     /label=exon 10

ORIGIN
        1 tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg
       61 agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga
      121 ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat
      181 tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata
      241 ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct
      301 ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct
      361 attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct
      421 tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg
      481 gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt
      541 gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat
      601 cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac
      661 tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa
      721 atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc
      781 caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa
      841 gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta
      901 tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag
      961 aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg
     1021 tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat
     1081 actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag
     1141 agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg
     1201 aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctcctttgt
     1261 cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt
     1321 ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc
     1381 tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact
     1441 cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg
     1501 atcctgcgtt gctgtggctg tggctgtggt gtaggccagc agcaaacagc tctgattaga
     1561 cccctagcct ggaaacctcc atatgccaca ggtgcagccc taaaaagaca aaaaaagaga
```

FIG. 16 tiếp

```
1621 aaagacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gaaccccccag
1681 aggtatttat ttgtttttgc ctttttcac tgactgttct ttgtttgttt gtttgagact
1741 gatctagaag actagagatt acaagaaata tggatttggc tcactctaag aaactgcttt
1801 cattccaagg tttgggtcta tccaaaagtg gaatagaatc atatgaatac tagtttatga
1861 gtatttagtg agaggaattt caagctcaaa taatgattca gcaagattaa attaaggagg
1921 gaattttcct tgtggctgag tgggttaagg acccaatgtt gtctctgtga ggatgtaggt
1981 tccatcctgg gctttgctca ttaggttaag gatctggcat tgctgcagct cagacccagt
2041 gctgccctgg ttgtggctta ggccaaagct gcagctccaa ttcaatctct ggcctgggaa
2101 cctccatgtg ctacaaggtg cggccttaaa aggaaaaaaa aaaaattaaa tcaaggactc
2161 aagagtcttt cattatttgt gttgtggaag ctatatttgt tttaaagtct tagttgtgtt
2221 tagaaagcaa gatgttcttc aactcaaatt tgggagggaa cttgtttcat acatttttaa
2281 tggataagtg gcaaaatttt catgctgagg tgatctatag tgttgtaatg cagaatatag
2341 tcagatcttg aacattttag gaagttggtg agggccaatt gtgtatctgt gccatgctga
2401 taagaatgtc aagggatcac aagaattcgt gttatttgac agcagtcatc tttaaaaggc
2461 atttgagaaa gtccaatttc aaatgcattt cctttcttta aaagataaat tgaagaaaat
2521 aagtctttat ttcccaagta aattgaattg cctctcagtc tgttaaaaga aactcttacc
2581 ttgatgattg cgctcttaac ctggcaaaga ttgtctttaa aatctgagct ccatgtcttc
2641 tgctttattt ctggtgtgcc tttgactcca gattacagta aatggaggac tgagtatagg
2701 gctaaaaagt agagagaatg gatgcatatt atctgtggtc tccaatgtga tgaatgaagt
2761 aggcaaatac tcaaaggaaa gagaaagcat gctccaagaa ttatgggttc cagaaggcaa
2821 agtcccagaa ttgtctccag ggaaggacag ggaggtctag aatcggctaa gcccactgta
2881 ggcagaaaaa ccaagaggca tgaatggctt ccctttctca cttttcactc tctggcttac
2941 tcctatcatg aaggaaaata ttggaatcat attctccctc accgaaatgc tatttttcag
3001 CCCACAGGAA ACCCAGGCTG GTTGGAGGGG ACATTCCCTG CTCTGGTCGT GTTGAAGTAC
3061 AACATGGAGA CACGTGGGGC ACCGTCTGTG ATTCTGACTT CTCTCTGGAG GCGGCCAGCG
3121 TGCTGTGCAG GGAACTACAG TGCGGCACTG TGGTTTCCCT CCTGGGGGGA GCTCACTTTG
3181 GAGAAGGAAG TGGACAGATC TGGGCTGAAG AATTCCAGTG TGAGGGGCAC GAGTCCCACC
3241 TTTCACTCTG CCCAGTAGCA CCCCGCCCTG ACGGGACATG TAGCCACAGC AGGGACGTCG
3301 GCGTAGTCTG CTCAAgtgag acccagggaa tgtgttcact ttgttcccat gccatgaaga
3361 gggtagggtt aggtagtcac agacatcttt ttaaagccct gtctccttcc agGATACACA
3421 CAAATCCGCT TGGTGAATGG CAAGACCCCA TGTGAAGGAA GAGTGGAGCT CAACATTCTT
3481 GGGTCCTGGG GGTCCCTCTG CAACTCTCAC TGGGACATGG AAGATGCCCA TGTTTTATGC
3541 CAGCAGCTTA AATGTGGAGT TGCCCTTTCT ATCCCGGGAG GAGCACCTTT TGGGAAAGGA
3601 AGTGAGCAGG TCTGGAGGCA CATGTTTCAC TGCACTGGGA CTGAGAAGCA CATGGGAGAT
3661 TGTTCCGTCA CTGCTCTGGG CGCATCACTC TGTTCTTCAG GGCAAGTGGC CTCTGTAATC
3721 TGCTCAGgta agagaataag ggcagccagt gatgagccac tcatgacggt gccttaagag
3781 tgggtgtacc taggagttcc cattgtggct cagtggtaac aaactcgact ggtatccatg
3841 agggtatggg tttgatccct ggccttgctc aatgggttaa ggatccagca ttgctgtgag
3901 ctgtggtata ggttgcagac tctgctcagg tcccatgttg ctgtgattgt ggtgtaggct
3961 gactgctgca gcttcaattt gacccctagc ccgggaattt ccataggcca cacgtgcagc
4021 actaaggaag gaaaaaaaGa aaaaaaaaaa aaaagagtgg gtgtgcctat agtgaagaac
4081 agatgtaaaa gggaagtgaa agggattccc ccattctgag ggattgtgag aagtgtgcca
4141 gaatattaac ttcatttgac ttgttacagg gaaagtaaac ttgactttca cggacctcct
4201 agttacctgg tgcttactat atgtcttctc agagtacctg attcattccc agcctggttg
4261 acccatcccc ctatctctat ggctatgttt atccagagca catctatcta acactccagc
4321 tgatcttcct gacacagctg tggcaaccct ggatccttta accaactgtg ccaggctgga
4381 gatcaaacct aagcctctgc agcaacccaa gctgctgcag tcagattttt aaccccctgt
4441 gccactgtgg gtatctccga tattttgtat cttctgtgac tgagtggttt gctgtttgca
4501 gGGAACCAGA GTCAGACACT ATCCCCGTGC AATTCATCAT CCTCGGACCC ATCAAGCTCT
4561 ATTATTTCAG AAGAAAATGG TGTTGCCTGC ATAGgtgaga atcagtgacc aacctatgaa
4621 aatgatctca atcctctgaa atgcatttta ttcatgtttt atttcctctt tgcagGGAGT
4681 GGTCAACTTC GCCTGGTCGA TGGAGGTGGT CGTTGTGCTG GGAGAGTAGA GGTCTATCAT
4741 GAGGGCTCCT GGGGCACCAT CTGTGATGAC AGCTGGGACC TGAATGATGC CCATGTGGTG
4801 TGCAAACAGC TgAGCTGTGG ATGGGCCATT AATGCCACTG GTTCTGCTCA TTTTGGGGAA
4861 GGAACAGGGC CCATTTGGCT GGATGAGATA AACTGTAATG GAAAAGAATC TCATATTTGG
4921 CAATGCCACT CACATGGTTG GGGGCGGCAC AATTGCAGGC ATAAGGAGGA TGCAGGAGTC
4981 ATCTGCTCGG
```

Fig. 17

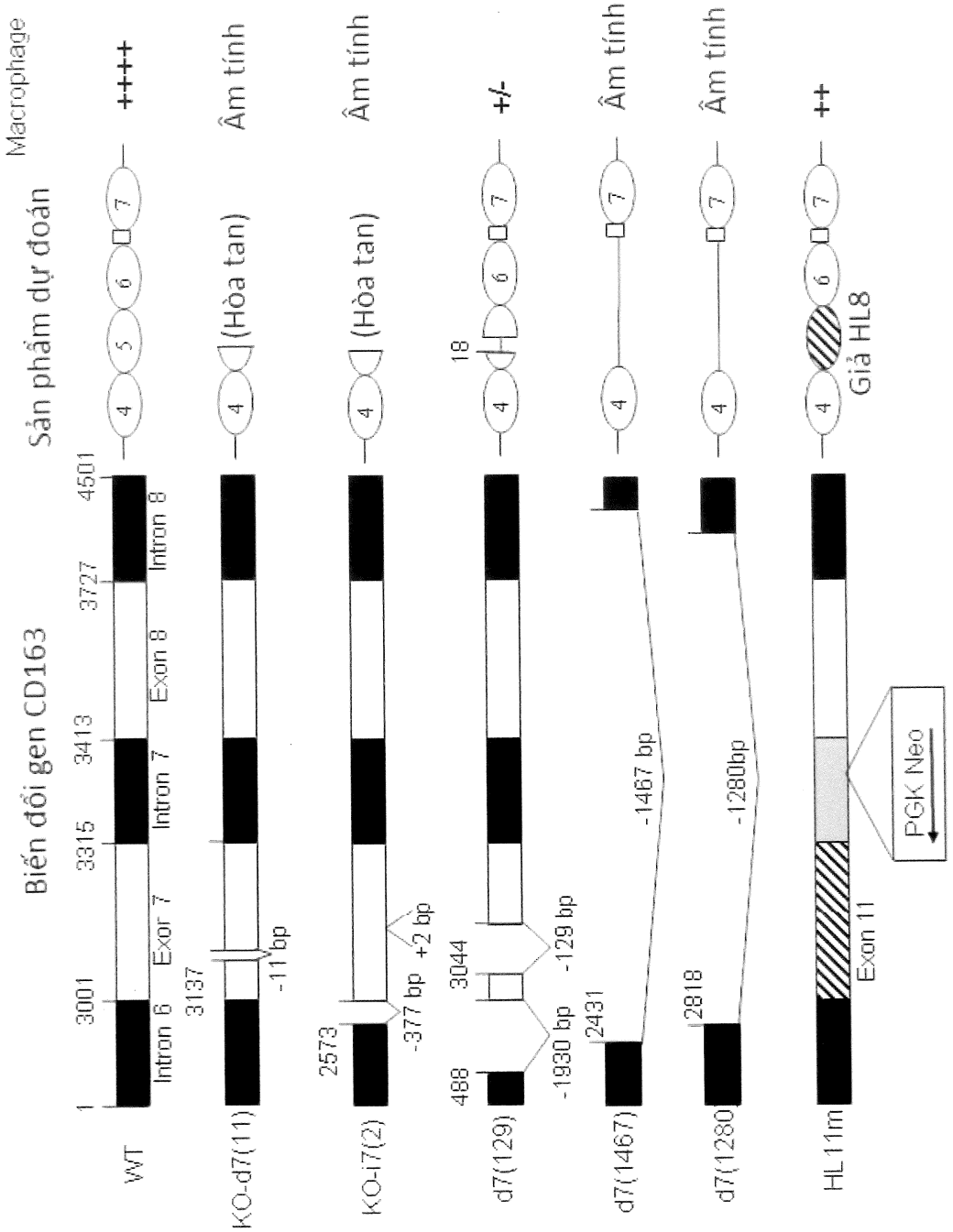

Macrophage
Sản phẩm dự đoán
Biến đổi gen CD163
WT
KO-d7(11)
KO-i7(2)
d7(129)
d7(1467)
d7(1280
HL11m
++++
Âm tính
(Hòa tan)
+/-
++
Giả HL8
Intron 6
Exor 7
Intron 7
Exon 8
Intron 8
Exon 11
PGK Neo
-11bp
-377 bp
+2 bp
-1930 bp
-129 bp
-1467 bp
-1280bp

Fig. 18

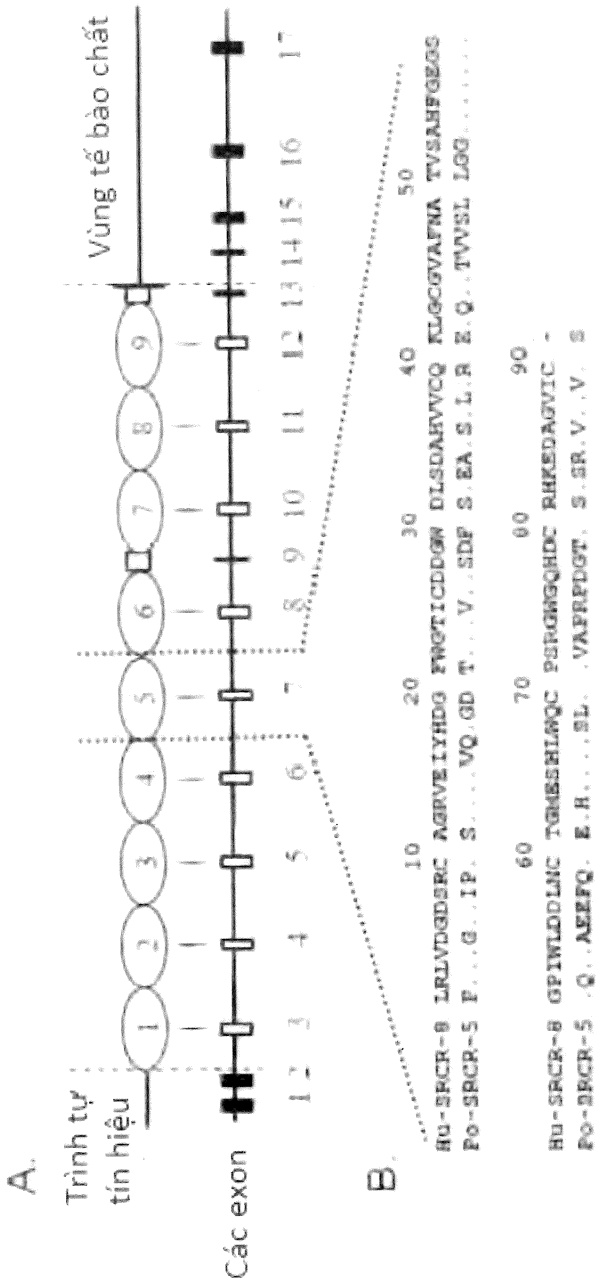

A.
Trình tự tín hiệu
Các exon
Vùng tế bào chất
B.
Hu-SRCR-8 LRLVDGDSRC AGRVEIYHDG EWGTICDDGW DLSDAHVVCQ QLGCGVAFNA TVSAHPGSGS
Po-SRCR-5 F...G..IP. S....VQ.GD T...V..SDF S.EA.S.L.R E.Q..TVVSL LGG.....
Hu-SRCR-8 GPIWLDDLNC TGHESHLAQC PSRGWGQHDC RHKSDAGVIC -
Po-SRCR-5 .Q..AEEFQ. E.H....SL. .VAPRPDGT. S.SR.V..V. S

Fig. 19

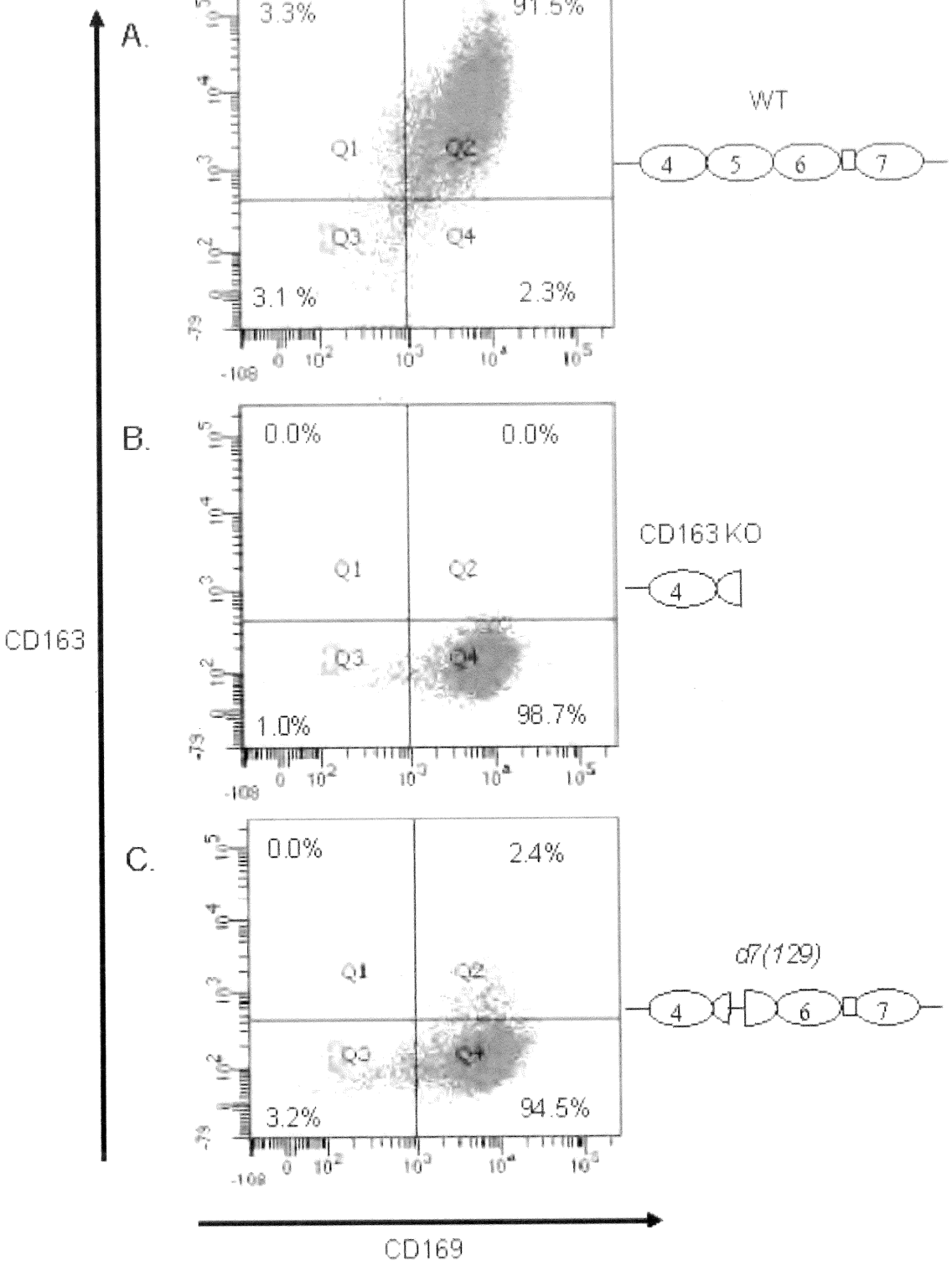

A.
3.3%
91.5%
Q1
Q2
Q3
Q4
3.1 %
2.3%
WT
4
5
6
7
B.
0.0%
0.0%
Q1
Q2
Q3
Q4
1.0%
98.7%
CD163 KO
4
C.
0.0%
2.4%
Q1
Q2
Q3
Q4
3.2%
94.5%
d7(129)
4
6
7
CD163
CD169

FIG. 19 TIẾP

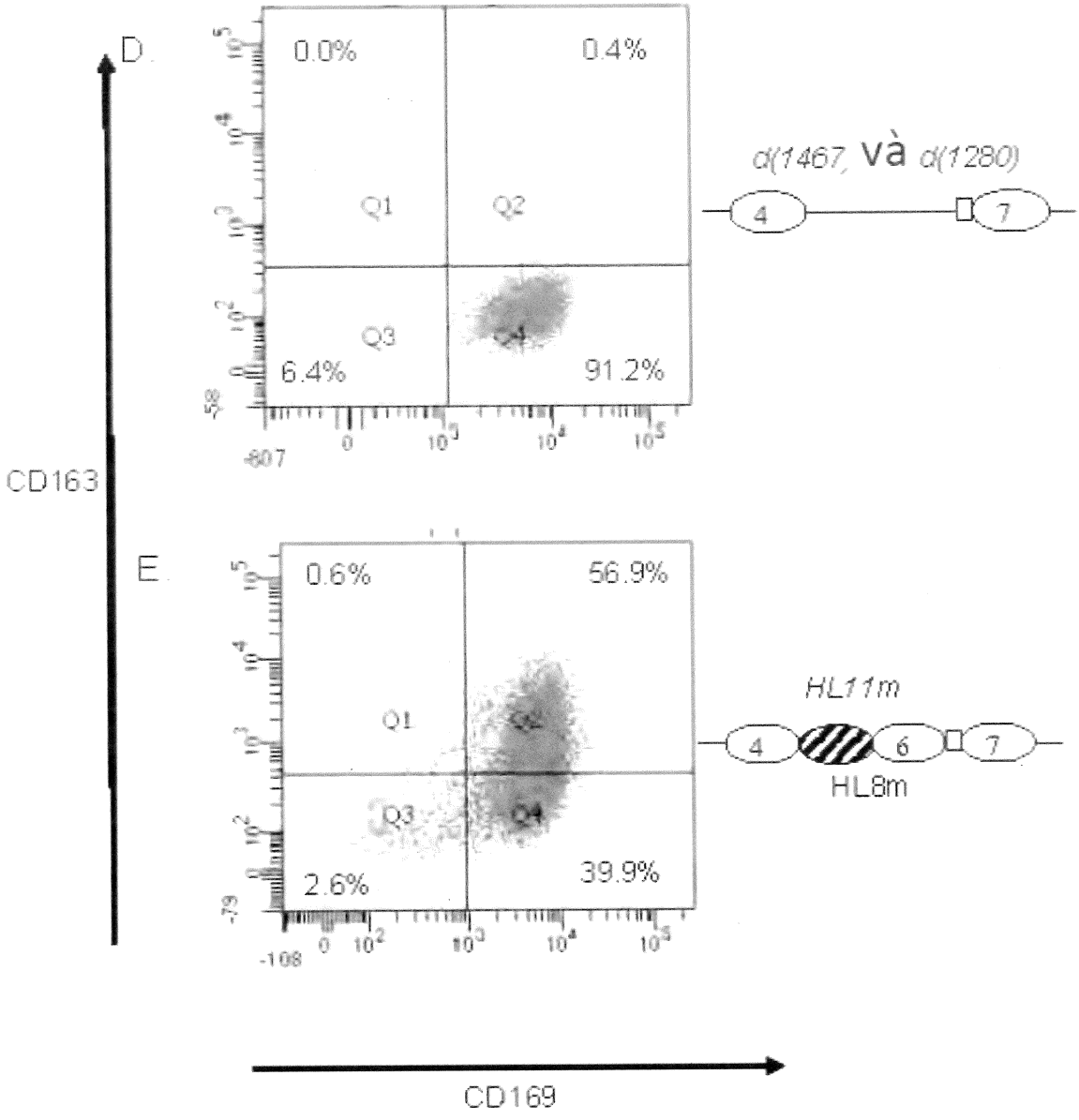

D
0.0%
0.4%
Q1
Q2
Q3
Q4
6.4%
91.2%
d(1467, và d(1280)
4
7
CD163
E
0.6%
56.9%
Q1
Q2
Q3
Q4
2.6%
39.9%
HL11m
4
6
7
HL8m
CD169

FIG. 20

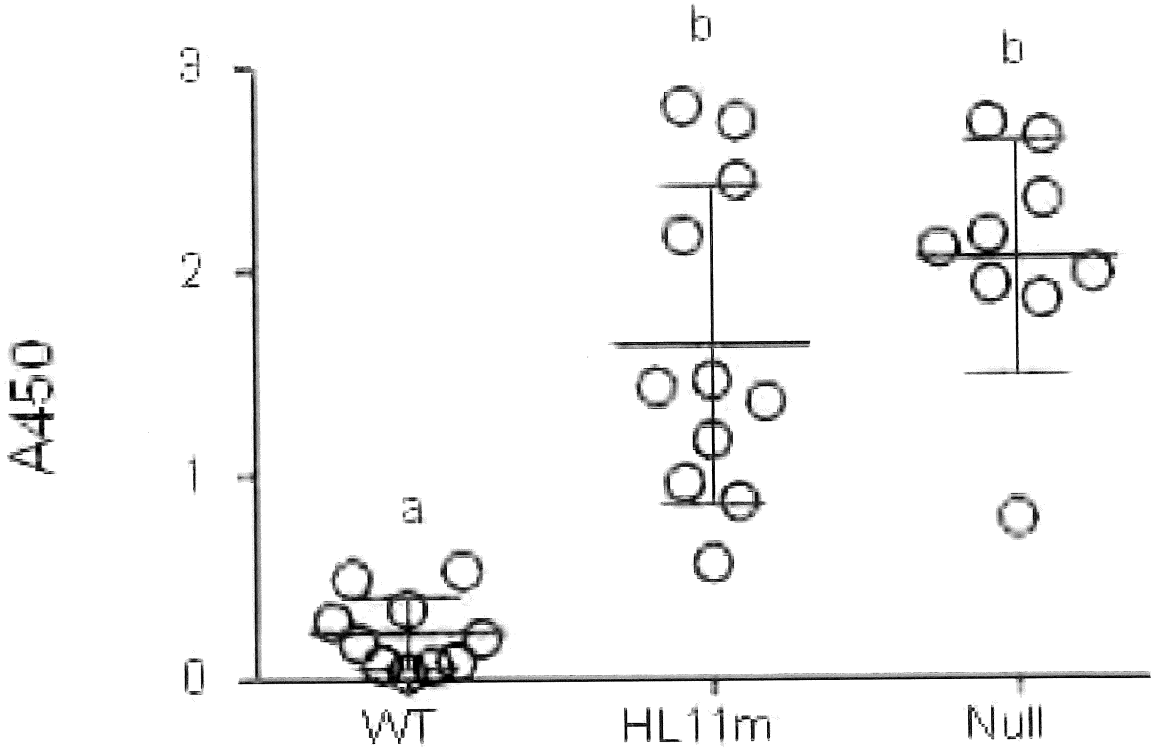


FIG. 21

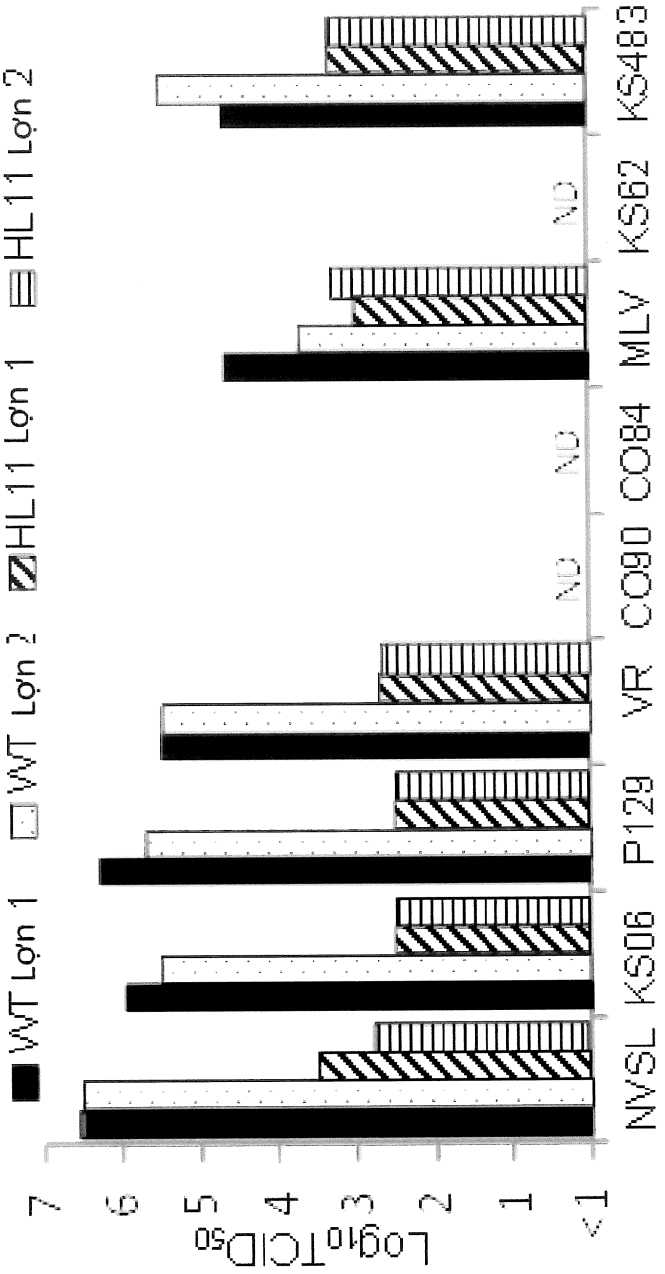
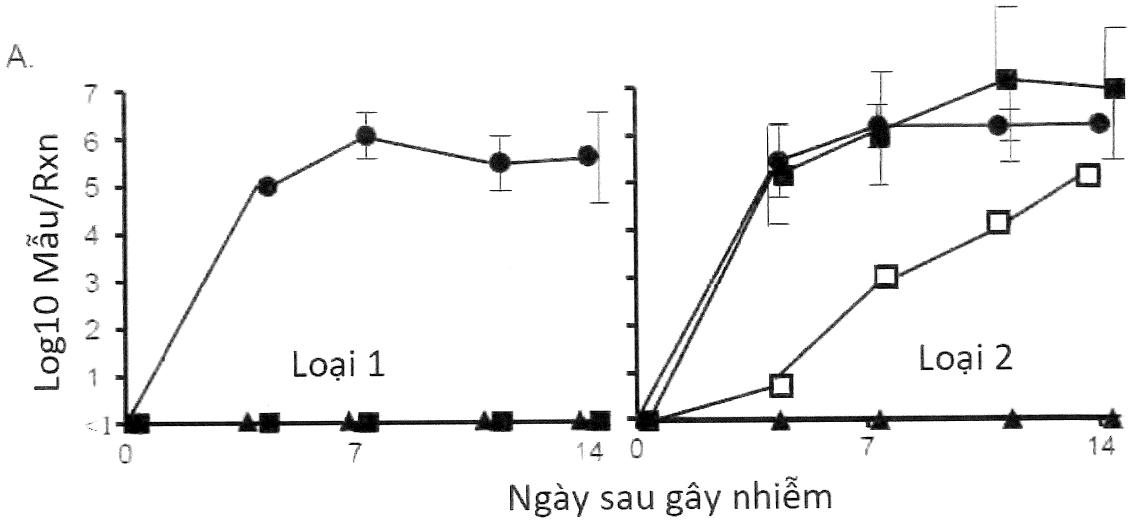

WT Lợn 1
WT Lợn 2
HL11 Lợn 1
HL11 Lợn 2
Log10TCID50
7
6
5
4
3
2
1
<1
ND
ND
ND
NVSL
KS06
P129
VR
CO90
CO84
MLV
KS62
KS483

FIG. 22


A.
Log10 Mẫu/Rxn
7
6
5
4
3
2
1
<1
Loại 1
Loại 2
0
7
14
0
7
14
Ngày sau gây nhiễm

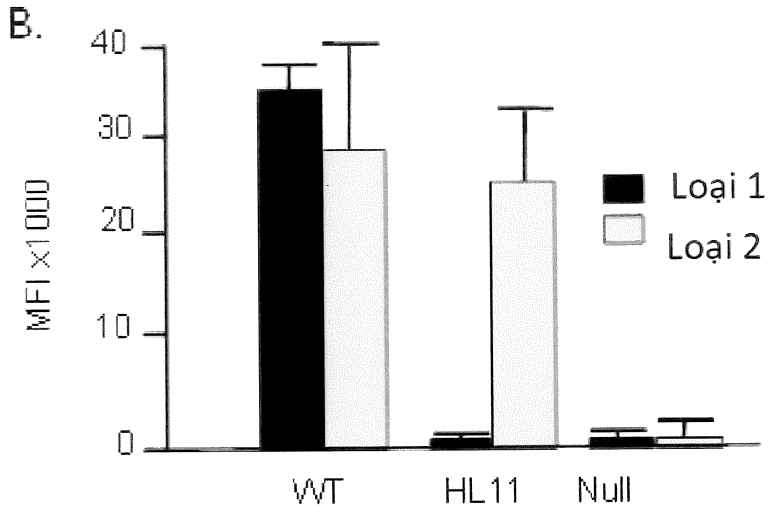
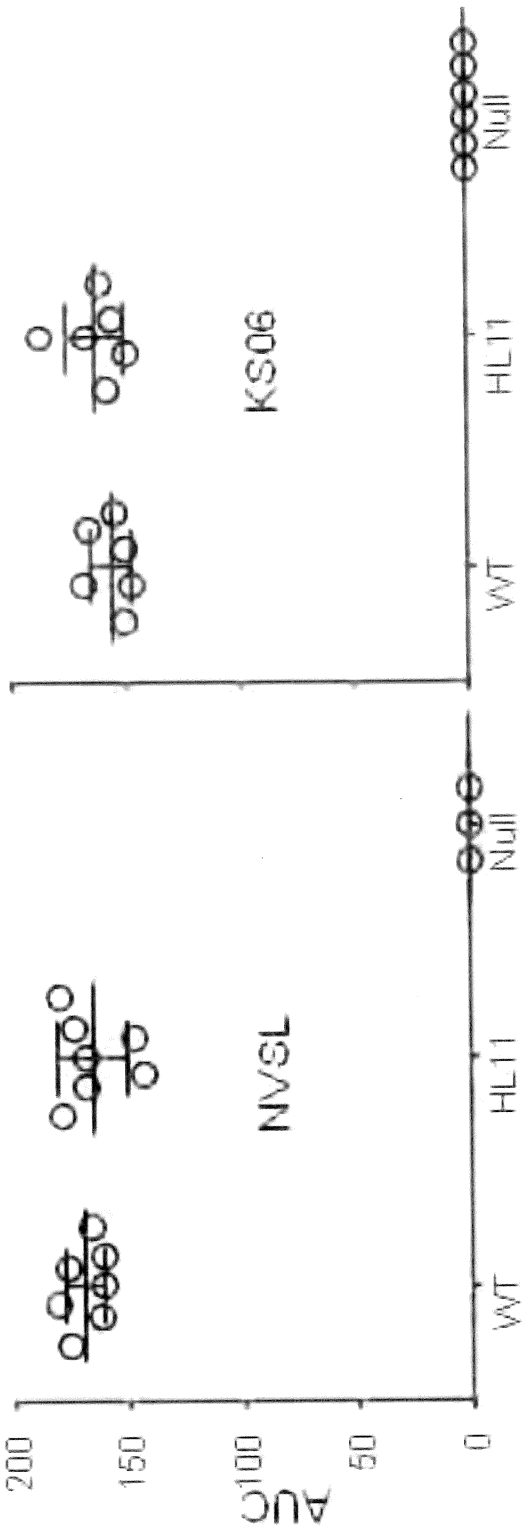

B.
MFI x1000
40
30
20
10
0
Loại 1
Loại 2
WT
HL11
Null

FIG. 23



DANH MỤC TRÌNH TỰ

<110> Curators of the University of Missouri
Prather, Randall
Wells, Kevin
Whitworth, Kristin

<120> ĐỘNG VẬT LÀ LỢN HOẶC THẾ HỆ CON CỦA NÓ CHỨA TRÌNH TỰ NHIỄM SẮC THỂ SỬA ĐỔI Ở ÍT NHẤT MỘT ALEN CỦA GEN MÃ HÓA CHO PROTEIN CD163 VÀ TẾ BÀO LỢN CHỨA TRÌNH TỰ NHIỄM SẮC THỂ NÀY

<130> UMO 16001.WO2

<150> US 62/202,145
<151> 2015-08-06

<160> 121

<170> PatentIn version 3.5

<210> 1
<211> 23
<212> ADN
<213> Sus scrofa

<400> 1
ggaaacccag gctggttgga ggg 23

<210> 2
<211> 23
<212> ADN
<213> Sus scrofa

<400> 2
ggaactacag tgcggcactg tgg 23

<210> 3
<211> 23
<212> ADN
<213> Sus scrofa

<400> 3
cagtagcacc ccgccctgac ggg 23

<210> 4
<211> 23
<212> ADN
<213> Sus scrofa

<400> 4
tgtagccaca gcagggacgt cgg 23

<210> 5
<211> 23
<212> ADN
<213> Sus scrofa

<400> 5
ccagcctcgc ccagcgacat ggg 23

<210> 6
<211> 23
<212> ADN
<213> Sus scrofa

<400> 6
ctttcattta tctgaactca ggg 23

<210> 7
<211> 23
<212> ADN
<213> Sus scrofa

<400> 7
ttatctgaac tcagggtccc cgg 23

<210> 8
<211> 23
<212> ADN
<213> Sus scrofa

<400> 8
cagctgcagc atatatttaa ggg 23

<210> 9
<211> 23
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 9
ctcctcgccc ttgctcacca tgg 23

<210> 10
<211> 23
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 10
gaccaggatg ggcaccaccc cgg 23

<210> 11
<211> 23
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 11
ctctccctca ctctaaccta ctt 23

<210> 12
<211> 22
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 12
tatttctctc acatggccag tc 22

<210> 13
<211> 23
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 13
ctctccctca ctctaaccta ctt 23

<210> 14
<211> 22
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 14
gactggccat gtgagagaaa ta 22

<210> 15
<211> 27767
<212> ADN
<213> Sus scrofa

<400> 15
atacaagtgc cttttacaga caatctgcac aagttatttg ttagacatat ttgattatag 60

aattaatatt aaaaggggtt ataacaatca agcattgata atttaattat gtttgcctat 120

tttactttag ttttttgaca taactgtgta actattgcga ttttttttatt cctaatgtaa 180

ttagttcaaa acaaagtgca gaaatttaaa atattcaatt caacaacagt atataagtca 240

atattccccc cttaaatttt tacaaatctt tagggagtgt ttctcaattt ctcaatttct 300

ttggttgttt catgtcccat atggaagaaa acatgggtgt gaaagggaag cttactcttt 360

tgattacttc ccttttctgg ttgactccac ctccattatg aagcctttct gtatttttgt 420

ggaagtgaaa tgatttttag aattcttagt ggttctcttc ttcaggagaa catttctagg 480

taataataca agaagattta aatggcataa aaccttggaa tggacaaact cagaatggtg 540

ctacatgaaa actctggatc tgcaggtaaa atcttctcat ttattctata tttacctttt 600

aaatagagtg tagcaatatt ccgacagtca atcaatctga tttaatagtg attggcatct 660

ggagaagaag taacagggaa aaaggcaata agctttataa ggggaacttt tatcttccat 720

agactcaaaa ttgaagacgt gactagaaga ttgctagatt tggcatcagt tttgtaaaat 780

tgctgaggtg aaattaagta agggatgaaa attaactaaa ttgtgttgag tatgaaacta 840

gtagttgtta gaaaagatag aacatgaagg aatgaatatt gattgaaagt tgatgaccta 900

gaggacattt agactaacac ctctgagtgt caaagtctaa tttatgattt acatcgatgc 960

gttaaactca tttaacattc ttacttttttt ccctcaagc atttaagctg aagtataaca 1020

tttcacatga aagcctggat tataaatgca cagttcagtg acctatctca gaggagtgac 1080

tgccatagca ttttttttgt ctttttgcct tcagagccac agcaacgcgg gatccgaagc 1140

cgcgtctgcg acccacacca cagctcacgg caatgccgga tctttaaccc actgagcgag 1200

gccggggatc gaacccgcag tctcatggtt cctagtagga ttcgttaacc actgcgccac 1260

gacgggaact cctaccatag cattttact tttaagttac tgttggttta gagtaagaag 1320

gagaaatgag agtgatggag cgtttgctat atttggagac aaggtcctat attggaggtt 1380

ctcaaatata aattttgtcg ctttttcctc caatgtattg ttcaactact atttagcagg 1440

ccactgtgcc aggtactggt gaaactggtg aacatgatag atgtaattca ttccctcatg 1500

gaactttcca tctaacaatg tggatcaggt aggcttggag atgagaatgc cagtggttga 1560

ctatgactct gtggctgaag ggagagctac tcacttcgta gtttcatcaa tgtcttttg 1620

gttttccagg ttttaagccc tgctcttgca attcttttcc cttctccaac tttcttctaa 1680

tttctcaccc ctaggatgcc tataaacatg agtattttca aagctacttc actgaggtta 1740

tatgatcctg gtgtgaattt ttcctgcctg acttgccatt tagaaggaag tgtttcctgg 1800

aatttccatt gtggcttggt ggttaaagac cctgcattgt ctctgtgagg atgtgggttc 1860

aatctctggc ctcattcagt gagtgggtta aggatctggt gtcgctgcaa gctgtggcta 1920

agatcccaca ttgccatgtc tgtggtgtag actggcacct ggagctctga tttgaccaca 1980

atcttaggaa cttcagatgt ggccataaaa aggaaaaaaa agttaggaag ggttttctgt 2040

cttgtttgga ccttcgttaa tctcaaacct ttggaaccat ctctcctcca aaacctcctt 2100

tgggtaagac tgtatgtttg ccctctctct tcttttcgca gactttagaa gatgttctgc 2160

ccatttaagt tccttcactt tggctgtagt cgctgttctc agtgcctgct tggtcactag 2220

ttctcttggt gagtactttg acaaatttac ttgtaaccga gcccaactgt gacaagaaac 2280

actgaaaagc aaataattgc tcctgaagtc tagatagcat ctaaaaacat gcttcatggt 2340

ttcaaggatc atatattgaa accccaggga tcctctagag tcgacctgca gcatgcaggg 2400

gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 2460

gggggggggg gggggggggg gggggggggg gggggggggg gtgcataagg aaagactatc 2520

tcaacgtctt attcctcagc ttacattaga tttgaaactc tagtcaccta aaatgcaaat 2580

ctcatttact taccatcaga gatattaatg acctatagaa ttcagcataa ataaagtttc 2640

atgtatggat attagcttat ggttctagtc actgctaatt gaaacctgtg atattgctgt 2700

ttgttttgac tcctatgaaa taacattctc ccattgtacc atggatgggt ccagaaacat 2760

ttctcaaatc ctggcttgaa aaaataaata agtaatctaa agaataataa ttctctactt 2820

gctctttgaa tcttgaccaa ttgctgcatt tacctattgt tacaggagga aaagacaagg 2880

agctgaggct aacgggtggt gaaaacaagt gctctggaag agtggaggtg aaagtgcagg 2940

aggagtgggg aactgtgtgt aataatggct gggacatgga tgtggtctct gttgtttgta 3000

ggcagctggg atgtccaact gctatcaaag ccactggatg ggctaatttt agtgcaggtt 3060

ctggacgcat ttggatggat catgtttctt gtcgagggaa tgagtcagct ctctgggact 3120

gcaaacatga tggatgggga aagcataact gtactcacca acaggatgct ggagtaacct 3180

gctcaggtaa gacatacaca aataagtcaa gcctatacat gaaatgcttt gtgggaaaaa 3240

atgtatagat gagttaaaaa caaaaaggaa ccagttttct ataagtcatc tagtccatgt 3300

ataaaattac ccaatccatt actaaaagac cacttctggt attttacaca tgacaaagcc 3360

catattaaaa aaaaaaaatt cagaagagat tctgaatgct ataataaatg agcaagtgac 3420

tagcttcaat tttatattag gtcattctac cttctacttc tacatgaaaa tatcataatg 3480

tctaagttaa ttccttgtcc cctttcccaa taaagcactg ctttcatgca ctggcctatg 3540

aatcatgaac tttttgccct ttaactgatg atcaacttac caaatcaaga aataaatatt 3600

cttagcactg atccttttt gttgttgttg gaggaagaat gttttgcaaa gtagaattgc 3660

ttttttctgt ttaacagtgc tattcatttc atttacatgg tcgttttaat ttataaaaca 3720

tttcataagt ttcacctcat atgcccttac aataactcag gaagttatat gttagacctt 3780

tctgctgaca aatcccagag tcatgtttct gacccagttc agattccttg gcttcccatt 3840

tctctttgct catgtcattg acctttatgc agccctctta cctcccacct ttctattaca 3900

gaccatctcc tccataggac tggtgttaga aagtactaat ctctacccag gcattgtggt 3960

gcaatgtggg cagcacaggc tggtatctag aaaaatgctg aagtgaattc cagctcagct 4020

gctcgttaat actatcgttt taagtaagct gttcaatcct ttgaaattca ctttctgagc 4080

actcagtgat ataataaatg tagagctact ggtacactgt ctggtatgta ataggtgtta 4140

ccaattaacc ttagtttcct catgggtcac tggttctcat tacctagaca actcatttct 4200

ctttcttcct ctttctcttt ctccattctc ctcctccttc ttcctcttct tcttgtctgt 4260

tattgttata tcattttgct gagaaagtta agaaataaca actctaacct ctacatcgac 4320

cacctagagc aaagttaaaa ataataataa accttgccag actcttacta taattgttgc 4380

tgtctataga gttgactgtt taagttaaga catcagtata tatttttaat ttttgtgttt 4440

ttttttttcat acttttacat gaggatcctt tatataagga tgagttaaac aaacttgatt 4500

tttgaagttt ataccccctga ggctcaactg cataataata gaaagggatc catagcctct 4560

caaggactta actagtttca tgagttttca gaatctgaat ttctgagatt ctccacccca 4620

attaaagctc aagcctcaga acatatatcc ttctcttggt aaattctatt cttatcacat 4680

gcgtaataat aaaaaagaga gatgttggag acagattttt ttcctcacat tctgtctcta 4740

ctgttttcta ggtgtttgat tctgtgttat ttaacctcag tttgcttatc tgtgaagtag 4800

ggattatggt aataacatat aatgctttat gttgtaaaga ctaaagaaga tagcatatgt 4860

aacacatttg gaacagggaa tgcatatttt gattgtgagc tcttattatt attaccattc 4920

agccctaata aaaatcttgg taagtggaag gctttggatt tcagaacttt taaaatctaa 4980

ttacttttttc aaaaaagaac ttcttagggt ttttttttttt taaccacaaa gtgtttctat 5040

tttttaggtg tcccaaaatt tcgttccaaa tatctttttc tcagatattt tagtcctcat 5100

agaacaccta gggatagtgg atagagaaaa ttttctttat taaaaagctg ttctttgcta 5160

aaaattgtag caggtacttt tgggaggggg gaaaactttg attcagaaac tgctaagaca 5220

tggagtgttt tgactaattt ttcctcaatt tttaatgttt tttataccat agggtacttt 5280

tgcaaactat tatgcatact tatatatttt tacttttttc ctgtctttta acttccaaat 5340

tcaacttcag acaattattc atgcactaaa ctgtttgtag taagaaagat taaaattaaa 5400

aaattaacca ttcaacaaat gactggtttg ccatttttac tactttgttg tatgaacaat 5460

tttttttttct acaaatgaat actttgagtc tgatttatcc attcctacat aaaagttttt 5520

actatatctt agtattggaa ggaaacaaaa caaaacacaa tgtaaatttt aatctataaa 5580

ttttgggggg gtaaatatac atagatgaaa gtcttaacca ttaattagag tcaaaagatt 5640

aaaattctcc aatatgtgaa cttaggctgc atccaaaatg aagcatcatt tttaaggaca 5700

gcatcaaaag tgaccagagg aattttactt tctttctttt ttttttttttt gaattttagt 5760

ttctaaactc acttctgaat aaatacaact tctaaattct cgtcttttct ctactctaga 5820

tggatctgat ttagagatga ggctggtgaa tggaggaaac cggtgcttag gaagaataga 5880

agtcaaattt caaggaacgt ggggaacagt gtgtgatgat cacttcaaca taaatcatgc 5940

ttctgtggtt tgtaaacaac ttgaatgtgg aagtgctgtc agtttctctg gttcagctaa 6000

ttttggagaa ggttctggac caatctggtt tgatgatctt gtatgcaatg gaaatgagtc 6060

agctctctgg aactgcaaac atgaaggatg gggaaagcac aattgcgatc atgctgagga 6120

tgctggagtg atttgcttaa gtaaggactg acctgggttt gttctgttct ccatgagagg 6180

gcaaaaaaag gggagtaaaa gtcttaaaag ctcaaactgt taaaaacata atgatgattg 6240

cttcttttat catcttatta ttatctaatt tcaggtcgaa attctagtac ctgtgcagtt 6300

ttttaccttа actgaaatta agataaatag gatagggagg aaggatgagc agtgacattt 6360

aggtccaagt catgaggtta gaaggaaatg ttcagagaat agcccattcc ctcagccctc 6420

aaagaaagaa agaaagaaaa agaaaaaaaa aaagaaagct taactagaaa attttgttct 6480

ctggatgttt tagaggcaaa ccatcccttt atcattccta cctacaaagc cttctcttaa 6540

tcacattacc cacccttтcc tactatagtc agggggggggg gggggggggg gggggggggg 6600

gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 6660

gggggggggg gtgaaaaaag aaccaaacaa tttcaacaaa aaaccaaaca attccaacaa 6720

aattggtcca ataagcaaac ctctagataa atttcagtgc cctggatgtt ttgttaggaa 6780

ctcttcctac aatgcgtgct ttccattctg aaaagtccta tctacttgcc tgatccactt 6840

ctccttccat cctaaacgat tttcagtggt agtatattac tgttgtctct gtctctactt 6900

atatatcttc ccctttтcac tcactcctct caggtacagc tcttcagttt gcccttattc 6960

ttgtttcctt gtcaatgact tgttttgtgt ccctcttaca gatggagcag acctgaaact 7020

gagagtggta gatggagtca ctgaatgttc aggaagattg gaagtgaaat tccaaggaga 7080

atggggaaca atctgtgatg atggctggga tagtgatgat gccgctgtgg catgtaagca 7140

actgggaggt ccaactgctg tcactgccat tgtcgagtta acgccagtga gggactggac 7200

acattggctc acacacatac agccatgaca cgatctgctc tatggtccga tgattaaagg 7260

gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 7320

gggggggggg gggggggggg gggggggggg ggggggggag aagagctggt ggacatttct 7380

ggaaaggaac caaaacccgg aagggccttg ttcttcagga tttgggatgg attggggagg 7440

gagaaaattg tttctaatat ttcttggtgg gaattctttt acagttgtga caaatctttc 7500

acatattctt catttgagta gtttggaggg ttgtctgact gttttctata ataaatgtcc 7560

caagtgctat gaggtaccac atttcaaatt ctaattctac ctgaagctcc aaaaagacaa 7620

aatgttatag gtcttttctt tatatctaat ttgcttatgg tttttagcca ttgacaattt 7680

ttttttctta actcttgaaa ctataaccct atttctaacc aaattcatgt tctatactgg 7740

ctcttcaaaa acccaggaga tgggaaagcc agaatctcca gtgtttcagc ttctgggaag 7800

gagcaagttt ttaaaaatac cctctgggag ctaaattcca catgtatcta tggcctaagt 7860

gtatgtttat tttgcagatg gatcagatct ggaactgaga cttaaaggtg gaggcagcca 7920

ctgtgctggg acagtggagg tggaaattca gaaactggta ggaaaagtgt gtgatagaag 7980

ctggggactg aaagaagctg atgtggtttg caggcagctg ggatgtggat ctgcactcaa 8040

aacatcatat caagtttatt ccaaaaccaa ggcaacaaac acatggctgt ttgtaagcag 8100

ctgtaatgga aatgaaactt ctctttggga ctgcaagaat tggcagtggg gtggacttag 8160

ttgtgatcac tatgacgaag ccaaaattac ctgctcaggt aagaatttca atcaatgtgt 8220

taggaaattg cattctactt tcttttacat gtagctgtcc agttttccca gcaccacttg 8280

ttgaagagac tgtcttttct tcatcatata gtcctacatc ctttgtcata aattaattga 8340

ccataggtgt gtgggtttat atctgggctc tctattctgt tcctttgatc tatatgtctg 8400

tttttatgcc agcaccatgc tgttttgatt actatagctt tgtagtatca tctgaagtca 8460

ggaaacatga ttcctccagc tttgttcttc tttctcaaga ttgttttgtc tattcagagt 8520

ttatgttccc atgcagattt aatttttaaa tttatttaat ttttatttt tatttttaat 8580

ttaaattaat ttaaattttt tatttcccaa cgtacagcca agggggccag ggtaaccttt 8640

acatgtatac attaaaaatt tcaggttttt cccccaccca tttctttctg ttggcaagta 8700

aatttttgaa caaagtttcc caatgctttt taaggggaat tcccttgggg gggggggggg 8760

gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 8820

gggggggggg gggggggggg gggggggggg gggggggggg agacgaaatt gactatattt 8880

tctttgttgg gaatctttta cagttgtgac aaatctttca catattcttc atttgagtag 8940

tttggagggt tgtctgactg ttttctataa taaatgtccc aagtgctatg aggtaccaca 9000

tttcaaattc taattctacc tgaagctcca aaaagacaaa atgttatagg tcttttcttt 9060

atatctaatt tgcttatggt ttttagccat tgacaatttt tttttcttaa ctcttgaaac 9120

tataacccta tttctaacca aattcatgtt ctatactggc tcttcaaaaa cccaggagat 9180

gggaaagcca gaatctccag tgtttcagct tctgggaagg agcaagtttt taaaaatacc 9240

ctctgggagc taaattccac atgtatctat ggcctaagtg tatgtttatt ttgcagatgg 9300

atcagatctg gaactgagac ttaaaggtgg aggcagccac tgtgctggga cagtggaggt 9360

ggaaattcag aaactggtag gaaaagtgtg tgatagaagc tggggactga aagaagctga 9420

tgtggtttgc aggcagctgg gatgtggatc tgcactcaaa acatcatatc aagtttattc 9480

caaaaccaag gcaacaaaca catggctgtt tgtaagcagc tgtaatggaa atgaaacttc 9540

tctttgggac tgcaagaatt ggcagtgggg tggacttagt tgtgatcact atgacgaaac 9600

caaaattacc tgctcaggta agaatttcaa tcaatgtgtt aggaaaattg cattctactt 9660

tcttttacat gtagctgtcc agttttccca gcaccacttg ttgaaaaaac tgtctttttc 9720

ttcatcatat agtcctacat cccttggcca taaattaatt gaccataagg ggtgtgggtt 9780

taatatccgg ggctcctcaa ttcgggtccc ttggatccta aaagccggtt ttataacccg 9840

acacatggcc tgtttttgac taaataaaac ctttggaaaa caatcccgaa ggtcgaggaa 9900

catggaatcc ccccaacaaa ggaccttctt tccccaaaaa tgcggctcag ccaactcaaa 9960

aagattttat gaatcacaaa ccgcacatta tcttcctaaa attactattc ctatgtttta 10020

atttgcaaag tcattccgat atagttggcg cagagtaact catttagata tccaccccac 10080

cagttcctca ctcaagtaag gggggggggg gggggggggg gggggggggg gggggggggg 10140

gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggc 10200

ccccatgtga gattttgtgt gtcctttaag agtggagtct ctatttccca ctgctctctg 10260

gttctcccca aagtaagccc tgctggcttt caaaacttct gggagcttgc cttcttggta 10320

taggactcct gggctaggga gtctaatgtt tggcttagac cccttactgc ttgggaagaa 10380

tctctgcaac tgtaatgaat tatcttccta tttgtgggtt gctgaggata tggtcttaac 10440

tgttctgtgt tctacccctc ctatccatct tgttgtggtt ccttctttat atctttagtt 10500

gtagaaaagt ttttcttatc aacagttgct ctgtaaattg taacttgggt gtacacctag 10560

taggaggtga gctcagggtc ttcctactct gccatcttgg ccatgtcctc taaacatttt 10620

ggtgtatttc actgcaacct ttttaaaaat ctcaaaagtg agctgtgatt ggctagtctt 10680

gtggataatc tctagcattt gatgctaatc atatttatac aaatactttg ttgaaaagtg 10740

atgccttttt aactattatt aaaaaacgta ttgacataac tattgctatt atactgaaaa 10800

gaaagacctt agagaaaata gcataagagc aaaaccatta aacatggaga catctagtca 10860

tagggtggaa attttatgtg gtccatatcc cctaaccagt ggctttacac caggcacatc 10920

ctaactaaga tctgctccca agtgtcttcc ctgatgcttt aaattgtgtt acatggaaac 10980

tatcctttga tgaagaaatg caacctttta aaatacaaca ttgaaacttt tgtgctttaa 11040

ttttgctttt caacattttt tctttttaaa agaagaaatt tatttgtttt tttaaatttt 11100

aatggccacg gcatatggaa gttctcaggc cagggataga attcaagcca caggtgcgac 11160

ccatgccaca actgctgcaa caccagatcc tttaacccac tgcaccaggc cagggattga 11220

agccttgcct tactgacaat ctgagccact tcagtcagat aaagaaattt cttcattaag 11280

cagagtattc acatggttta aacttcaaaa tattaaagtg taaactcttt ccccaccact 11340

gtccccagct caccaactct acttaccaca gacaactgat gtggttaggg tatttaaata 11400

gtaaatccaa gaaaatataa acaaatccgt atatataggt ttcaccccat tttattatcc 11460

taatgttgca tatcatataa actatactgt cccttgggta ttcacttagt aaaatatttt 11520

gatcataatt tcctatcagt atttaaagag ctttctgaaa ttatttctgt ataacatttc 11580

ttttctcatc ggtaggggggg gggggggggg gggggggggg gggggggggg gggggggggg 11640

gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg ggggaatggg 11700

aagaaaaaac caccatggtt aatttttttt atccctctac acccgggaaa attacccttg 11760

gggccacact tttctataga aaggggatta tttaaaaggg tctgaaaaag aatttttttt 11820

tcgaaagggg aaatatttgg cctaacttag tcacataagc catgttctct ggcaagttag 11880

gtaacataca tttttgtcat tgggggcaac aaaaacaatt ttccttttgg acctttgggg 11940

actccgcatt ggttagggaa ggggaagtat attggaattc ggaaaattcc ttccaaatta 12000

aaaaaggttt gttattttca tattaaccta tttcatatta attagcatga attccagcgc 12060

cattaaaagg gaaaacacct ggagtggtaa gaaaaaagtt ttttttctc ttttttttt 12120

ttttttttta atggccacat ctgtggcatg tgaagttccc aggctagggg tcgaatagga 12180

gctacagctg ccagcttgca ccacagccac aacaatgcca gagccaagcc tcatctgcga 12240

cctataccac aactcatggc aatgctggtt ccttaacccc ctgagtgagg cctggggtca 12300

aacccacatc ctcatggata ctaaccggct ttgttaccgc tgagccatga gggaaactcc 12360

ctttttctca ttgaaaataa gtcaaataga taagcagctt aaggctgttt gggtgattct 12420

gtggtccagt aattatcaaa tcctactgga caagaataga gaatgtgcaa atgagggaac 12480

gtgttggtga gatcaggctc tgcccactga gctatcctct gtcatgggcc ctgtgctgtt 12540

ctcagagctg tacttcctag ggcattgttc tcatttcaat tctgagttca gtgtggagag 12600

tatacgtgtg tgggggctgc acgcttttca caacccactt tctgctgata ctgatttagg 12660

gatccttgga ttgctttaca gttgagtcat cattaactag tgtcacttgc cttcaaagtc 12720

agcaaaataa ttgtctccaa actagtaggc ttctagtgta tttgctttaa tccaatgcca 12780

tgtgaaagta acatggtcaa agaataagtt atataccttg acctaccctg tgaccaggct 12840

cttcctctta atttattgac cactgcctta aggtcatttg aaaccatggg tttgggagga 12900

aggcaaggcc taaatcccgt ctttgttgga aggctcactg tccttgtctt tagagcatca 12960

ttttttttta aactggggta cagtttattt acagtgttgt gtcaatttct gctgtacagc 13020

acagtgaccc agtcatacac atacatacat tctttttctc atactatctt caattttatt 13080

ttctgctaag tctgccattt tatcatcacc tcagtttgaa ggacaggata tttagagttt 13140

gtttttttt tccccccaat cctgcaattt ctaaattata agactctcaa ttagccgtat 13200

ataacagctg caggcacagg atgtctccct cacaaaattg gtatttttcc ttccatttct 13260

tcttgcagtt tggctatttc ttgtctgagt tcatctctct ttttaagtgt taaaaagggc 13320

aaggaggatt catgctatgt caacattatg atttttttctt ttctatactt gataagagta 13380

tacttttccc aaatgtcatc caacttttca gcatcagttt ggacatggtt ttcttttcaa 13440

ggtggtattt ctctaatgtc acttgaataa caagactcgt tagttctcca ggctacaata 13500

tcctagtctg agtatattct gcatgttaat tctattcagc cacatccata atttaggttt 13560

tattcctgga acacctcact ttttttttt tttttggtct ttttatagcc ataaccatgg 13620

catatggagg ttcccaggct aggggtctaa tctgagcttt agccactggc ccatgccaca 13680

gccacagcca tgccacatct gagccacatc tgtgaccttt tccacagctc acagaaacac 13740

cagatcccta acccactgag tgaggccagg ggtcaaacct gtaacccttc catggttcct 13800

agtcagattc gttcctctgt accacgatgg gaattcctaa tacctcactt atgataacac 13860

attctgaatt atttaggatt ctattatact gcatgtaata gaaatcccaa atagcaaaat 13920

ttgcaactta aggcaggttc ctgtctttac aaaatcatgt tttcctttgc tatatgtgca 13980

ctttgctttc ctctgtgaat tccctttttt gttatatttc tatagctttt ggaaacactt 14040

ttacttattt gggggggcct agatttttaa ccctctcctt gtttttctag aaatagagtt 14100

tataatttta tttcttcatt tacttgatac tttcaagaga ttcccaggaa aaaaattatg 14160

gaaatactgt ctctgtgcct gccaagttca aactaagaat tgtataatct gttttaattc 14220

ttaagcattt atagatgaca aggctttgtg tctgataggg gccagcgaac tcagtaaaga 14280

gggaagatga gaaagataat ggcaagaatt tatccctgaa gtgtagtttt gacaaaccag 14340

tcacaaagag gtctaagaaa ttttggtcac aaagttgttt tgaatcccag gcattttatt 14400

tgcaatgatt gcatatgttc tggaaaggac atctgaacct aagaaatagt tcatttgcat 14460

tgtgttatat tttactaagg tctgagaaat aatcttgaga tgagaatgaa ctctacttct 14520

tcagagtctg gaaggaataa attatgaaaa tgtattaatg cttctttaaa ccatattgta 14580

tatttatcta ttactaaaca aaaagaagta gctctattta tttatttatt tatttattta 14640

tttatgtctt ttgtctcttt agggccacac ctgtggcata tggaggttcc caggctagag 14700

gtccaattgg agatgtagca gccagcctat gccagagcca ccgcaacacg ggatctgagc 14760

cacgtctgtg acttacacca cagctcacag caacgcctga tcctcaaccc actgagcgag 14820

gccagggatc gaacccatgt cctcatggat gctagttggg ttcgttaact gctgagccat 14880

gatgggaact ccaaattaat tatttcttat atttgttctt catatattca tttctataga 14940

aagaaataaa tacagattca gttaatgatg gcaggtaaaa gcttaactta ttaatcaaag 15000

gagttaatcc aggcacaaaa attcaattca tggctctctg ttaaaattta ggtataggtt 15060

tagcaggaag aaaaggttag tagatgcaga ctattacatt tagaatggat ggacaatgaa 15120

gtcctactat acagcacagg gaactatatc caatctcttg ggatagaata tgatggaaga 15180

caaaatcaga acaagagagt atatatatat gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt 15240

gtgtgtgtgt gtgtgactgg gtcaccctgc ggcacagcag aaattggcag aacattgtaa 15300

atcaactata ctttaatagg aaaaatactt ttaagggcta aatttccaat attctaacca 15360

tgtacacaga gtaaatgtca taaggatgcc agtctgtgta gagattgatg tgttactagc 15420

agattcatga aataaaggct gaggatgtag tccccaagtc acttctgagt ggaagaattt 15480

ctcctttgtc ctggactcaa atattttagg ataaaggaaa aaagaagata tttatagaag 15540

ggacttgttt tcaagtactt gacaaaattt caccattaaa gagaaatttg tgggagttcc 15600

catcgtggct cagtggaaac aaatccaact aggaaccatg aggttgtggg tttgatccct 15660

ggcctcactc agtgggttaa ggatccggtg ttgccgtgag ctgtggtgta ggttgcagac 15720

acggttctga tcctgcgttg ctgtggctgt ggctgtggtg taggccagca gcaaacagct 15780

ctgattagac ccctagcctg gaaacctcca tatgccacag gtgcagccct aaaaagacaa 15840

aaaaagagaa aagacaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaag 15900

aacccccaga ggtatttatt tgtttttgcc ttttttcact gactgttctt tgtttgtttg 15960

tttgagactg atctagaaga ctagagatta caagaaatat ggatttggct cactctaaga 16020

aactgctttc attccaaggt ttgggtctat ccaaaagtgg aatagaatca tatgaatact 16080

agtttatgag tatttagtga gaggaatttc aagctcaaat aatgattcag caagattaaa 16140

ttaaggaggg aattttcctt gtggctgagt gggttaagga cccaatgttg tctctgtgag 16200

gatgtaggtt ccatcctggg ctttgctcat taggttaagg atctggcatt gctgcagctc 16260

agacccagtg ctgccctggt tgtggcttag gccaaagctg cagctccaat tcaatctctg 16320

gcctgggaac ctccatgtgc tacaaggtgc ggccttaaaa ggaaaaaaaa aaaattaaat 16380

caaggactca agagtctttc attatttgtg ttgtggaagc tatatttgtt ttaaagtctt 16440

agttgtgttt agaaagcaag atgttcttca actcaaattt gggagggaac ttgtttcata 16500

catttttaat ggataagtgg caaaattttc atgctgaggt gatctatagt gttgtaatgc 16560

agaatatagt cagatcttga acattttagg aagttggtga gggccaattg tgtatctgtg 16620

ccatgctgat aagaatgtca agggatcaca agaattcgtg ttatttgaca gcagtcatct 16680

ttaaaaggca tttgagaaag tccaatttca aatgcatttc ctttctttaa aagataaatt 16740

gaagaaaata agtctttatt tcccaagtaa attgaattgc ctctcagtct gttaaaagaa 16800

actcttacct tgatgattgc gctcttaacc tggcaaagat tgtctttaaa atctgagctc 16860

catgtcttct gctttatttc tggtgtgcct ttgactccag attacagtaa atggaggact 16920

gagtataggg ctaaaaagta gagagaatgg atgcatatta tctgtggtct ccaatgtgat 16980

gaatgaagta ggcaaatact caaaggaaag agaaagcatg ctccaagaat tatgggttcc 17040

agaaggcaaa gtcccagaat tgtctccagg gaaggacagg gaggtctaga atcggctaag 17100

cccactgtag gcagaaaaac caagaggcat gaatggcttc cctttctcac ttttcactct 17160

ctggcttact cctatcatga aggaaaatat tggaatcata ttctccctca ccgaaatgct 17220

atttttcagc ccacaggaaa cccaggctgg ttggagggga cattccctgc tctggtcgtg 17280

ttgaagtaca acatggagac acgtggggca ccgtctgtga ttctgacttc tctctggagg 17340

cggccagcgt gctgtgcagg gaactacagt gcggcactgt ggtttccctc ctgggggggag 17400

ctcactttgg agaaggaagt ggacagatct gggctgaaga attccagtgt gaggggcacg 17460

agtcccacct ttcactctgc ccagtagcac cccgccctga cgggacatgt agccacagca 17520

gggacgtcgg cgtagtctgc tcaagtgaga cccagggaat gtgttcactt tgttcccatg 17580

ccatgaagag ggtagggtta ggtagtcaca gacatctttt taaagccctg tctccttcca 17640

ggatacacac aaatccgctt ggtgaatggc aagaccccat gtgaaggaag agtggagctc 17700

aacattcttg ggtcctgggg gtccctctgc aactctcact gggacatgga agatgcccat 17760

gttttatgcc agcagcttaa atgtggagtt gccctttcta tcccgggagg agcacctttt 17820

gggaaaggaa gtgagcaggt ctggaggcac atgtttcact gcactgggac tgagaagcac 17880

atgggagatt gttccgtcac tgctctgggc gcatcactct gttcttcagg gcaagtggcc 17940

tctgtaatct gctcaggtaa gagaataagg gcagccagtg atgagccact catgacggtg 18000

ccttaagagt gggtgtacct aggagttccc attgtggctc agtggtaaca aactcgactg 18060

gtatccatga gggtatgggt ttgatccctg gccttgctca atgggttaag gatccagcat 18120

tgctgtgagc tgtggtatag gttgcagact ctgctcaggt cccatgttgc tgtgattgtg 18180

gtgtaggctg actgctgcag cttcaatttg acccctagcc cgggaatttc cataggccac 18240

acgtgcagca ctaaggaagg aaaaaaagaa aaaaaaaaaa aaagagtggg tgtgcctata 18300

gtgaagaaca gatgtaaaag ggaagtgaaa gggattcccc cattctgagg gattgtgaga 18360

agtgtgccag aatattaact tcatttgact tgttacaggg aaagtaaact tgactttcac 18420

ggacctccta gttacctggt gcttactata tgtcttctca gagtacctga ttcattccca 18480

gcctggttga cccatccccc tatctctatg gctatgttta tccagagcac atctatctaa 18540

cactccagct gatcttcctg acacagctgt ggcaaccctg gatcctttaa ccaactgtgc 18600

caggctggag atcaaaccta agcctctgca gcaacccaag ctgctgcagt cagattttta 18660

acccccctgtg ccactgtggg tatctccgat attttgtatc ttctgtgact gagtggtttg 18720

ctgtttgcag ggaaccagag tcagacacta tccccgtgca attcatcatc ctcggaccca 18780

tcaagctcta ttatttcaga agaaaatggt gttgcctgca taggtgagaa tcagtgacca 18840

acctatgaaa atgatctcaa tcctctgaaa tgcattttat tcatgtttta tttcctcttt 18900

gcagggagtg gtcaacttcg cctggtcgat ggaggtggtc gttgtgctgg gagagtagag 18960

gtctatcatg agggctcctg gggcaccatc tgtgatgaca gctgggacct gaatgatgcc 19020

catgtggtgt gcaaacagct gagctgtgga tgggccatta atgccactgg ttctgctcat 19080

tttggggaag gaacagggcc catttggctg gatgagataa actgtaatgg aaaagaatct 19140

catatttggc aatgccactc acatggttgg gggcggcaca attgcaggca taaggaggat 19200

gcaggagtca tctgctcggg taagttctgc acatcacttc gggttacagt gatttaagaa 19260

acaactaagg tggggcaaag ggtagtgagg catatccatc agagcaaatt ccttgaaata 19320

cggactcaga gggaaccatt gtgagattga ggttcccaga ggtgtggatt taatgaatta 19380

gtgttacctc atgtacaagg tagtatacta ccagaaagat aaaaattcag aagcgagttt 19440

gcagcaaaac tcatagggag aacttctttt ataaataata tgaagctgga tatttagtgc 19500

accacctgat gaccacttta ttaataaata aagagttcct gttgtggcgc agcggaaatg 19560

aatccgacaa ataatcatga gtttgcgggt ttgatccctg acctcgctca gtgggttggg 19620

gatctggtgt tgccatgagc tgtggtgtag gtcgcagatg ctgcttggat cctgctttgc 19680

tgtggctgtg gtataggctt gtggctacag ctccgatttg accgctagcc tgggaacctc 19740

catatgctgc gggggtggcc ctcaaaagaa aaataaataa ataagtaaat aaataagtag 19800

tttaaaaagg acaagaagaa atatatttgg tgttatattc tacagagaca aagataatca 19860

ccatgcccga ttgatttttc aaggcatata aatgagacgt catgggagca aaaatggtca 19920

taatacaatg cccttgtttt gtgtacatgg taagatttta gaaagcattg tgaggtaaaa 19980

aagtgtactc agttataata tattggggaa aacagtacta tgagaagtaa aaaaatctac 20040

atgccggaag ttatttttt aatgtctctt ttagagtcgc acatgcggca tatggaggtt 20100

cccaggctag gggtcgaatc agagctatag ccactggctt atggcacagc cacaacaaca 20160

ctagatctga gccacatcag cgacctatac tatagctcat ggcaatgcca gatccttaac 20220

ctactgagcc aagccatggg tcaaatccag gtcctcacgg atcctaggca aattcatttc 20280

tgctgagcca cgaagggaac tcctcagaag tgattttgat gttactttct tttcatgaca 20340

aatctggtaa agtacataca catagaaact gaagtgtcag aaagggaaat atttcatttt 20400

aaggtaatgt atacaaaaca gtggttttac catctgagta tctcgctaaa ttttaactat 20460

caaggacaat tgccaaaaaa aaaaaaaaaa gagagagaga gagaacagaa tagggttatg 20520

aagctaaaat cacagggtta tgaagctaaa atcacagtaa tttagggaga aaaaaatcca 20580

aagcatgtaa ttgataaaag gttctgagcc tttgtttgag atttagaatt caacttagaa 20640

ataccggtgg tattttaaag cagtccataa gtataaaatc caaggctaaa aaaccagaag 20700

gtatttgtag aacaaatata ttttaataag ctctaccaag tcatccagaa gctattaaag 20760

aattactggt cactgacata gtgtacctgt tttcaaggcc attcttacat cagaataaag 20820

ggagagcacc ctctgaatct tcagaaaaga tgtgaaagtg ctaattctct atttcatccc 20880

agagttcatg tctctgagac tgatcagtga aaacagcaga gagacctgtg cagggcgcct 20940

ggaagttttt tacaacggag cttggggcag cgttggcaag aatagcatgt ctccagccac 21000

agtgggggtg gtatgcaggc agctaggctg tgcagacaga ggggacatca gccctgcatc 21060

ttcagacaag acagtgtcca ggcacatgtg ggtggacaat gttcagtgtc ctaaaggacc 21120

tgacacacta tggcagtgcc catcatctcc atggaagaag agactggcca gcccctcaga 21180

ggagacatgg atcacatgtg ccagtgagta tccattcttt agcgccactg ttatcttctg 21240

atctacctaa gcagaagttt tataatctgt agttaatccc tattctacct ggatgatggg 21300

attcattctg tttaatttgg tgtgcaggta ttcagcatca gtgatcattt tcccaaagac 21360

catcatgctc tgatggtctt ctcaaaagtt ctaatcagtt gcttcctccg tgaacagttg 21420

aggagcagag aatatgtaat tcagaatttg actattgaat catcccattt ttctttcaca 21480

tagtcttttg ttgcactgaa tataaggaga gaagcagtca gaaagatcaa tcctgaatta 21540

tttctccatt ctacatctgt tttaaatttc aaaaaaaaaa attgttatag gtgatttaca 21600

atgtctgtca atttctgctc tacagcaaag tgacccagtt atttacatat acattctgtt 21660

tctcatattt ttaaaccagg agatttctat ctgcctggcg gtttgaggga atttaacatt 21720

atgcatttat gttaacttta ttcacctgat gttttctaag tcatactgag attcttatgc 21780

ccaggatgga atacacctgg tttgctggaa agacatgtgc tttcataaag acgaattttg 21840

gaaaaaatat aaaatttaaa aggcccatta aataagcaaa gttttaagag atttcaaaaa 21900

aaatttcatc tctctctttt cctctttgac ctcttgggca cgttcatctt ctcaaatatg 21960

atcttggtgt ttctgacttt tcagacaaaa taagacttca agaaggaaac actaattgtt 22020

ctggacgtgt ggagatctgg tacggaggtt cctggggcac tgtgtgtgac gactcctggg 22080

accttgaaga tgctcaggtg gtgtgccgac agctgggctg tggctcagct ttggaggcag 22140

gaaaagaggc cgcatttggc caggggactg ggcccatatg gctcaatgaa gtgaagtgca 22200

aggggaatga aacctccttg tgggattgtc ctgccagatc ctggggccac agtgactgtg 22260

gacacaagga ggatgctgct gtgacgtgtt caggtgaggg cagagagtct ggattgagct 22320

tggaagctct ggcagcaaag agagggtggg cggtgacctg cattgggtaa agattggaag 22380

gtccagccta aggatctggt ggtgggggga gacatgatgt ttcagtctga agaatgatga 22440

aaacctgtgt ggttacgcat gggccttcgc cgaggaaagg gacataacta ccatgtatcc 22500

tcctgcagag ggaggaagaa ctaggggatt ctagttttgt gtgggaagga gcagtttact 22560

tggctcagga ggcactaaag gctcagatag gaaacagaga tctgttccat tcttactccc 22620

agaactgatt ctcttctctt ttctcctaca gaaattgcaa agagccgaga atccctacat 22680

gccacaggta tatcaaaaag tttaagaaca tgggacccat tgtctgcatt ttgtggaatc 22740

cctcttatta agacattctg ggtcagaagt tctgaggatt tgacatttac ttcagctatc 22800

tgttatctta cccaagagag ggatggtaac taggaaccca ggtcttttag ctaagacatt 22860

atcacctctt gtgatgttta cttgttctca ggtcgctcat cttttgttgc acttgcaatc 22920

tttggggtca ttctgttggc ctgtctcatc gcattcctca tttggactca gaagcgaaga 22980

cagaggcagc ggctctcagg tctgaacaaa attacggtct ctctaatgtt tctatgggag 23040

aagaagcctc tctggataat aaaacaaaaa aattacattc aagtatcagt tggccagaaa 23100

gagggaacct agaagaggtt taagcagttt ctccgaaaca gggaacaaga attcagagaa 23160

gaaaaggcac attggctgta ctgatgatac ctgcactcgc tatgtatgtt taatggggga 23220

cagtagagaa ttgatagttt agaaggagta tgcttatatg gttctggatg aatcctgtat 23280

ccccccaaac atttattttc tcttactata tacttattac taatttaact cttctgtcaa 23340

gccatgtgct aggttctgaa gatggttcag acttggataa ccaagtgctt ttgttttcat 23400

ggaatttcca gtttagtgga agagataaat atgtaaacaa ataaatgggg gggggggggg 23460

gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 23520

gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 23580

gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 23640

gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 23700

gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 23760

gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 23820

gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 23880

gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 23940

gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg gggggggggg 24000

gggggggggg gggggggggg gggggggggg gggggggggg ggggggggtt ggcgggcccc 24060

cctcgaggtc gacggtatcg ataagcttga tatcgaattc gtgagccaga ggacgagact 24120

agagatggat gatgactacg ttatgcttgc actgctgggg aaaagcacac atagggaggg 24180

aacgttttat tatgacccag tccctaacct atgacctctg ttatcagttt tctcaggagg 24240

agagaattct gtccatcaaa ttcaataccg ggagatgaat tcttgcctga aagcagatga 24300

aacggatatg ctaaatccct caggtccgtg ggttctttga ggggctgtag ccctggggtt 24360

cagatcagca gctgcagttg aggttgaggc atgctacttt gcatagcagt agaaagaaat 24420

ctcaactgta ataggaagct tgggatgcat atgaggaaga aaggcaagaa tgaactacaa 24480

attattctta gggaagataa aaattgcagt catggggaga cctctggctg agagggccgt 24540

gattatttct gacagaggga ttatggagta gaatatgatg gcttggacct tttttcacta 24600

aaacaagtca gtcttctcaa aggtagttta gcttttcata tatctttcac agtttcttcc 24660

attcccattt cctgccattt tcctttctct aacttttatt tattatattt tttcctaaaa 24720

gtttaaattt tctatatctt tatcccttca gaagccatcc ctagtcacag gactagtctc 24780

atttcccatt atgtaatgct tctttctctg tctgttgact tctatttaga accagtgcac 24840

taaatctgcc tttaggaaca tacctctgct aggttgcaag aaatatccca ttccccactc 24900

actctgtgaa gactcaatgc ttctcaatat tccttacctc ctgagaggga cttgcctcac 24960

ttctttaatc caagggactc gatttttgcc aaaactaagt caggaaaacc tacataagac 25020

ataggaaaga cttgctgtgc ttcttaaacc ccactgtttg ttttcctaat tgtgaacagt 25080

atttttaaag ttaacaagag agcttctaag gcacttgagg ggagatctga tttatttccc 25140

agtaattatt ttcttccttt cagaaaattc cactgaataa gatggtttta acggatgtgg 25200

gactaatttt tgtgtctaaa tctcttccta tttctggatg aaaaaaagga gaccactctg 25260

aagtacaatg aaaaggaaaa tgggaattat aacctggtga ggtgagtagg aagaatttat 25320

tcatcattgc tgaaaacagg tacattcctt ttgaaagttg agaactcctc tggtattaga 25380

aaaaaaaaaa gaacgtatat acacatatat ttccatgtct atgtttatgt ttgtaaatcc 25440

atattcagaa tatgcaacaa ctttttataa ctatgacttc agtccatctt ttagttacat 25500

atatattcta aacaacaact attgctaaga gaagctgggt aagtaaatgt gaataaatct 25560

tctaaagata ttacaggaag ttcctgctgc ggctcagtgg gttaaagact tgatgtcttt 25620

gtgaagatga gggctcgagc cctggcctca ctcagtgagt taaggatcta gcattgctgt 25680

aagctgcagc gtaggttgca gatagggctc agatccagtg ttgctgtggc tgtggcctca 25740

gttgcagctc tgattcaacc cttaggcgag gaacttccat atgcagcaaa tgtggccatt 25800

aaaaaaaaaa aaaacattat aggagtcatt tcataaaaga gataagacgt ttctatagtt 25860

atatagtgca tactctggta aagatagtat aggatactat aggaatatag aaagcttgcc 25920

tatgaaaatt tgggaagatt gtggaaaaga catctcaaaa tatggcatag aaaagaatca 25980

tatctttgag gaacagtaag tttttcattc aaaaccgtgt attgaacata cttgtggtga 26040

caagtggtgt cctgagtact aaaaattcag tgataaaaga tgctcttgac aaagacatgg 26100

ctgttgaata gaaggtctca ctgtcaatgt gtgggaatta tggacagcct atgtggacac 26160

agggaataga tgagactcta ggctggaagg ctgcattgag cccaataatg aatggtcctg 26220

tctgatatat ttcatgctca tattttattt tagggactat tggggaggtg gtgggttttg 26280

gaagattaag ctgaggcaag acacaatcag attgcctttt ataatttact ttcaggagga 26340

aagtctaact aaaaaaagaa ttcgatatca agcttatcga taccgtcgac ctcgaggagg 26400

cccgcctgcc cttttggggg gggggggggg gggggggggg gggggggggg gggggggggg 26460

gggggggggg gggggggggg gggggggggg gggggggggg gggggcagca 26520

ccaattttat tattggcggg aataaagaga aaaatgtaat ttcaaagatt gctgttggaa 26580

atgaggggtg tggtagcttt tggagaaagc attctggaga cttctattaa ttttttttt 26640

ttaagtgctt caaagatcct ttgatccaac aattctactc ctaaaaattt cttccataca 26700

gataaagcca tttgtctgta tataacaaat agaagagaat tccttttttgc agccttgtta 26760

gtagtgcccc caaactggaa acaaagtgaa tatcagtcag tggggtagcg gctggaaaaa 26820

ttttagtgca cccaaccaac aaagaaaaac catgcacaaa aattcaataa atatcatctc 26880

acttttgtgt tcatgttatt gaatataatt aaacataatg tttacatcta taaaattatc 26940

atatgtatac atgtaaagaa acattaaaac atttttaaca gactgtaaac ttgaggactg 27000

tgaatgactt ttgattgata atctcaaaca tatggatact attctgatgt aataaataat 27060

gattaaattt tttccctaaa gagtaatcac tactgaatcg ttgcctcaga atcatatgga 27120

ggtgctttta aaaaaggcat ttctgcactg ttgttctctg gaatagaagt aattcttatg 27180

tacactgaag tttgaaaatc attgcattta agtgttctgt tcaggaaagt agtgtgcttt 27240

ttaatatttg tgagtgaatg agtaacacaa tacattatat cacattttaa tgtaattcta 27300

cacatgtgca tatgaagaga aaagtaacat ttttttctat ttatgtcttt agttcagcct 27360

ttaagatacc ttgatgaaga cctggactat tgaatgagca agaatctgcc tcttacactg 27420

aagattacaa tacagtcctc tgtctcctgg tattccaaag actgctgttg aatttctaaa 27480

aaatagattg gtgaatgtga ctactcaaag ttgtatgtaa gactttcaag ggcattaaat 27540

aaaaaagaat attgctgatt cttgttcttg attttctgaa tttctgaatc tcttattggg 27600

cttctaattt aaaaaaaaat atctgggcgc ccgcagatat cgaactcttg ggcagtgtga 27660

ccaaacgaag acatatccaa tcaagcatgc aaatggacca gcccactgta ctagcacgct 27720

gtggcagcca atctgaccga gaaagcagac aaccgcaggg agcaacg 27767

<210> 16
<211> 55
<212> ADN
<213> Sus scrofa

<400> 16
ggtcgccacc atggtgagca agggcgagga gctgttcacc ggggtggtgc ccatc 55

<210> 17
<211> 54
<212> ADN
<213> Sus scrofa

<400> 17
ggtcgccacc atggccatga gcaagggcga ggagctgttc accggggtgg tgcc 54

<210> 18
<211> 48
<212> ADN
<213> Sus scrofa

<400> 18
ggtcgccacc atggtgagca agggcgagga gctgttcacc ggggtggt 48

<210> 19
<211> 55
<212> ADN
<213> Sus scrofa

<400> 19
ggtcgccacc atggttgagc aagggcgagg agctgttcac cggggtggtg cccat 55

<210> 20
<211> 43
<212> ADN
<213> Sus scrofa

<400> 20
tgcagggaac tacagtgcgg cactgtggtt tccctcctgg ggg 43

<210> 21
<211> 60

<212> ADN
<213> Sus scrofa

<400> 21
tgcagggaac tacagtgcgg cactgtaaac cactactact gtggtttccc tcctgggggg 60

<210> 22
<211> 41
<212> ADN
<213> Sus scrofa

<400> 22
tgcagggaac tacagtgcgg ctgtggtttc cctcctgggg g 41

<210> 23
<211> 46
<212> ADN
<213> Sus scrofa

<400> 23
tgcagggaac tacagtgcgg aactactgtg gtttccctcc tggggg 46

<210> 24
<211> 31
<212> ADN
<213> Sus scrofa

<400> 24
gaaacccagg ctggttggag gggacattcc c 31

<210> 25
<211> 24
<212> ADN
<213> Sus scrofa

<400> 25
gaaacccagg ctggggacat tccc 24

<210> 26
<211> 13
<212> ADN
<213> Sus scrofa

<400> 26
aggggacatt ccc 13

<210> 27
<211> 13
<212> ADN
<213> Sus scrofa

<400> 27
gaaacccatt ccc 13

<210> 28
<211> 31
<212> ADN
<213> Sus scrofa

<400> 28
ggtcgccacc atggtgagca agggcgagga g 31

<210> 29
<211> 32
<212> ADN
<213> Sus scrofa

<400> 29
ggtcgccacc atggctgagc aagggcgagg ag 32

<210> 30
<211> 29
<212> ADN
<213> Sus scrofa

<400> 30
ggtcgccacc atggtgagag ggcgaggag 29

<210> 31
<211> 32
<212> ADN
<213> Sus scrofa

<400> 31
ggtcgccacc atggttgagc aagggcgagg ag 32

<210> 32
<211> 48
<212> ADN
<213> Sus scrofa

<400> 32
ggtcgccacc atggtgagca agggcgagga gaacccaggc tggttgga 48

<210> 33
<211> 49
<212> ADN
<213> Sus scrofa

<400> 33
tgctgtgcag ggaactacag tgcggcactg tggtttccct cctgggggg 49

<210> 34
<211> 38
<212> ADN
<213> Sus scrofa

<400> 34
tgctgtgcag ggaactctgt ggtttccctc ctgggggg 38

<210> 35
<211> 22
<212> ADN
<213> Sus scrofa

<400> 35
ctgtggtttc cctcctgggg gg 22

<210> 36
<211> 23
<212> ADN
<213> Sus scrofa

<400> 36
actgtggttt ccctcctggg ggg 23

<210> 37
<211> 50
<212> ADN
<213> Sus scrofa

<400> 37
tgctgtgcag ggaactacag tgcggcaact gtggtttccc tcctgggggg 50

<210> 38
<211> 10
<212> ADN
<213> Sus scrofa

<400> 38
tcctgggggg 10

<210> 39
<211> 8
<212> ADN
<213> Sus scrofa

<400> 39
ctgggggg 8

<210> 40
<211> 52
<212> ADN
<213> Sus scrofa

<400> 40
agagagcaga gccagcgact cgcccagcga catggggtac ctgccgtttg tg 52

<210> 41
<211> 33
<212> ADN
<213> Sus scrofa

<400> 41
agagagcaga gccagcgact cgcccagcga gat 33

<210> 42
<211> 30
<212> ADN
<213> Sus scrofa

<400> 42
agagagcaga gccagcgact cgcccagcga 30

<210> 43
<211> 50
<212> ADN
<213> Sus scrofa

<400> 43
agagccagcc tcgcccagca ggggtaccat ggggtacctg ccgtttgtgt 50

<210> 44
<211> 53
<212> ADN

<213> Sus scrofa

<400> 44
agagagcaga gccagcgact cgcccagcga gcagtgggta cctgccgttt gtg 53

<210> 45
<211> 53
<212> ADN
<213> Sus scrofa

<400> 45
agagagcaga gccagcgact cgcccagcga tcagtgggta cctgccgttt gtg 53

<210> 46
<211> 53
<212> ADN
<213> Sus scrofa

<400> 46
agagagcaga gccagcgact cgcccagcga acatggggta cctgccgttt gtg 53

<210> 47
<211> 4990
<212> ADN
<213> Sus scrofa

<400> 47
tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt 540

gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat 600

cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac 660

tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa 720

atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc 780

caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa 840

gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta 900

tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag 960

aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 1020

tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat 1080

actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag 1140

agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg 1200

aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctccttttgt 1260

cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt 1320

ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc 1380

tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact 1440

cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg 1500

atcctgcgtt gctgtggctg tggctgtggt gtaggccagc agcaaacagc tctgattaga 1560

cccctagcct ggaaacctcc atatgccaca ggtgcagccc taaaaagaca aaaaaagaga 1620

aaagacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gaacccccag 1680

aggtatttat ttgtttttgc ctttttttcac tgactgttct ttgtttgttt gtttgagact 1740

gatctagaag actagagatt acaagaaata tggatttggc tcactctaag aaactgcttt 1800

cattccaagg tttgggtcta tccaaaagtg gaatagaatc atatgaatac tagtttatga 1860

gtatttagtg agaggaattt caagctcaaa taatgattca gcaagattaa attaaggagg 1920

gaattttcct tgtggctgag tgggttaagg acccaatgtt gtctctgtga ggatgtaggt 1980

tccatcctgg gctttgctca ttaggttaag gatctggcat tgctgcagct cagacccagt 2040

gctgccctgg ttgtggctta ggccaaagct gcagctccaa ttcaatctct ggcctgggaa 2100

cctccatgtg ctacaaggtg cggccttaaa aggaaaaaaa aaaaattaaa tcaaggactc 2160

aagagtcttt cattatttgt gttgtggaag ctatatttgt tttaaagtct tagttgtgtt 2220

tagaaagcaa gatgttcttc aactcaaatt tgggagggaa cttgtttcat acatttttaa 2280

tggataagtg gcaaaatttt catgctgagg tgatctatag tgttgtaatg cagaatatag 2340

tcagatcttg aacattttag gaagttggtg agggccaatt gtgtatctgt gccatgctga 2400

taagaatgtc aagggatcac aagaattcgt gttatttgac agcagtcatc tttaaaaggc 2460

atttgagaaa gtccaatttc aaatgcattt ccttttctta aaagataaat tgaagaaaat 2520

aagtctttat ttcccaagta aattgaattg cctctcagtc tgttaaaaga aactcttacc 2580

ttgatgattg cgctcttaac ctggcaaaga ttgtctttaa aatctgagct ccatgtcttc 2640

tgctttattt ctggtgtgcc tttgactcca gattacagta aatggaggac tgagtatagg 2700

gctaaaaagt agagagaatg gatgcatatt atctgtggtc tccaatgtga tgaatgaagt 2760

aggcaaatac tcaaaggaaa gagaaagcat gctccaagaa ttatgggttc cagaaggcaa 2820

agtcccagaa ttgtctccag ggaaggacag ggaggtctag aatcggctaa gcccactgta 2880

ggcagaaaaa ccaagaggca tgaatggctt ccctttctca cttttcactc tctggcttac 2940

tcctatcatg aaggaaaata ttggaatcat attctccctc accgaaatgc tatttttcag 3000

cccacaggaa acccaggctg gttggagggg acattccctg ctctggtcgt gttgaagtac 3060

aacatggaga cacgtggggc accgtctgtg attctgactt ctctctggag gcggccagcg 3120

tgctgtgcag ggaactacag tgcggcactg tggtttccct cctgggggga gctcactttg 3180

gagaaggaag tggacagatc tgggctgaag aattccagtg tgaggggcac gagtcccacc 3240

tttcactctg cccagtagca ccccgccctg acgggacatg tagccacagc agggacgtcg 3300

gcgtagtctg ctcaagtgag acccagggaa tgtgttcact ttgttcccat gccatgaaga 3360

gggtagggtt aggtagtcac agacatcttt ttaaagccct gtctccttcc aggatacaca 3420

caaatccgct tggtgaatgg caagacccca tgtgaaggaa gagtggagct caacattctt 3480

gggtcctggg ggtccctctg caactctcac tgggacatgg aagatgccca tgttttatgc 3540

cagcagctta aatgtggagt tgccctttct atcccgggag gagcaccttt tgggaaagga 3600

agtgagcagg tctggaggca catgtttcac tgcactggga ctgagaagca catgggagat 3660

tgttccgtca ctgctctggg cgcatcactc tgttcttcag ggcaagtggc ctctgtaatc 3720

tgctcaggta agagaataag ggcagccagt gatgagccac tcatgacggt gccttaagag 3780

tgggtgtacc taggagttcc cattgtggct cagtggtaac aaactcgact ggtatccatg 3840

agggtatggg tttgatccct ggccttgctc aatgggttaa ggatccagca ttgctgtgag 3900

ctgtggtata ggttgcagac tctgctcagg tcccatgttg ctgtgattgt ggtgtaggct 3960

gactgctgca gcttcaattt gacccctagc ccgggaattt ccataggcca cacgtgcagc 4020

actaaggaag gaaaaaaaga aaaaaaaaaa aaaagagtgg gtgtgcctat agtgaagaac 4080

agatgtaaaa gggaagtgaa agggattccc ccattctgag ggattgtgag aagtgtgcca 4140

gaatattaac ttcatttgac ttgttacagg gaaagtaaac ttgactttca cggacctcct 4200

agttacctgg tgcttactat atgtcttctc agagtacctg attcattccc agcctggttg 4260

acccatcccc ctatctctat ggctatgttt atccagagca catctatcta acactccagc 4320

tgatcttcct gacacagctg tggcaaccct ggatccttta accaactgtg ccaggctgga 4380

gatcaaacct aagcctctgc agcaacccaa gctgctgcag tcagattttt aacccccctgt 4440

gccactgtgg gtatctccga tattttgtat cttctgtgac tgagtggttt gctgtttgca 4500

gggaaccaga gtcagacact atccccgtgc aattcatcat cctcggaccc atcaagctct 4560

attatttcag aagaaaatgg tgttgcctgc ataggtgaga atcagtgacc aacctatgaa 4620

aatgatctca atcctctgaa atgcatttta ttcatgtttt atttcctctt tgcagggagt 4680

ggtcaacttc gcctggtcga tggaggtggt cgttgtgctg ggagagtaga ggtctatcat 4740

gagggctcct ggggcaccat ctgtgatgac agctgggacc tgaatgatgc ccatgtggtg 4800

tgcaaacagc tgagctgtgg atgggccatt aatgccactg gttctgctca ttttggggaa 4860

ggaacagggc ccatttggct ggatgagata aactgtaatg gaaaagaatc tcatatttgg 4920

caatgccact cacatggttg ggggcggcac aattgcaggc ataaggagga tgcaggagtc 4980

atctgctcgg 4990

<210> 48
<211> 24
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 48

caccggaaac ccaggctggt tgga 24

<210> 49
<211> 24
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 49
aaactccaac cagcctgggt ttcc 24

<210> 50
<211> 24
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 50
caccggaact acagtgcggc actg 24

<210> 51
<211> 24
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 51
aaaccagtgc cgcactgtag ttcc 24

<210> 52
<211> 25
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 52
caccgcagta gcaccccgcc ctgac 25

<210> 53

<211> 25
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 53
aaacgtcagg gcggggtgct actgc 25

<210> 54
<211> 25
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 54
caccgtgtag ccacagcagg gacgt 25

<210> 55
<211> 25
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 55
aaacacgtcc ctgctgtggc tacac 25

<210> 56
<211> 25
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 56
caccgccagc ctcgcccagc gacat 25

<210> 57
<211> 25
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 57
aaacatgtcg ctgggcgagg ctggc 25

<210> 58
<211> 25
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 58
caccgcagct gcagcatata tttaa 25

<210> 59
<211> 25
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 59
aaacttaaat atatgctgca gctgc 25

<210> 60
<211> 25
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 60
caccgctttc atttatctga actca 25

<210> 61
<211> 25
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 61

aaactgagtt cagataaatg aaagc 25

<210> 62
<211> 25
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 62
caccgttatc tgaactcagg gtccc 25

<210> 63
<211> 25
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 63
aaacgggacc ctgagttcag ataac 25

<210> 64
<211> 25
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 64
caccgctcct cgcccttgct cacca 25

<210> 65
<211> 25
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 65
aaactggtga gcaagggcga ggagc 25

<210> 66

<211> 25
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 66
caccggacca ggatgggcac caccc 25

<210> 67
<211> 25
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 67
aaacgggtgg tgcccatcct ggtcc 25

<210> 68
<211> 24
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 68
ttgttggaag gctcactgtc cttg 24

<210> 69
<211> 20
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 69
acaactaagg tggggcaaag 20

<210> 70
<211> 24
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 70
ttgttggaag gctcactgtc cttg 24

<210> 71
<211> 23
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 71
ggagctcaac attcttgggt cct 23

<210> 72
<211> 23
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 72
ggcaaaattt tcatgctgag gtg 23

<210> 73
<211> 23
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 73
gcacatcact tcgggttaca gtg 23

<210> 74
<211> 23
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 74

cccaagtatc ttcagttctg cag 23

<210> 75
<211> 23
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 75
tacaggtagg agagcctgtt ttg 23

<210> 76
<211> 23
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 76
cccaagtatc ttcagttctg cag 23

<210> 77
<211> 23
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 77
ctcaaaagga tgtaaaccct gga 23

<210> 78
<211> 22
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 78
tgttgatgtg gtttgtttgc cc 22

<210> 79

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

<220>
<223> Oligonucleotit tổng hợp

<400> 79
tacaggtagg agagcctgtt ttg 23

<210> 80
<211> 23
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 80
ggaggtctag aatcggctaa gcc 23

<210> 81
<211> 20
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 81
ggctacatgt cccgtcaggg 20

<210> 82
<211> 21
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 82
gcaggccact aggcagatga a 21

<210> 83
<211> 23
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 83
gagctgacac ccaagaagtt cct 23

<210> 84
<211> 22
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 84
ggctctagag cctctgctaa cc 22

<210> 85
<211> 22
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 85
ggacttgaag aagtcgtgct gc 22

<210> 86
<211> 44
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 86
taatacgact cactataggg agaatggact ataaggacca cgac 44

<210> 87
<211> 21
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 87

gcgagctcta ggaattctta c 21

<210> 88
<211> 40
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 88
ttaatacgac tcactatagg ctcctcgccc ttgctcacca 40

<210> 89
<211> 20
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 89
aaaagcaccg actcggtgcc 20

<210> 90
<211> 38
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 90
ttaatacgac tcactatagg aaacccaggc tggttgga 38

<210> 91
<211> 20
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 91
aaaagcaccg actcggtgcc 20

<210> 92

<211> 38
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 92
ttaatacgac tcactatagg aactacagtg cggcactg 38

<210> 93
<211> 20
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 93
aaaagcaccg actcggtgcc 20

<210> 94
<211> 40
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 94
ttaatacgac tcactatagg ccagcctcgc ccagcgacat 40

<210> 95
<211> 20
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 95
aaaagcaccg actcggtgcc 20

<210> 96
<211> 40
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 96
ttaatacgac tcactatagg cagctgcagc atatatttaa 40

<210> 97
<211> 20
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Oligonucleotit tổng hợp

<400> 97
aaaagcaccg actcggtgcc 20

<210> 98
<211> 3484
<212> ADN
<213> Sus scrofa

<400> 98
tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt 540

gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat 600

cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac 660

tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa 720

atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc 780

caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa 840

gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta 900

tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag 960

aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 1020

tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat 1080

actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag 1140

agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg 1200

aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctcctttgt 1260

cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt 1320

ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc 1380

tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact 1440

cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg 1500

atcctgcgtt gctgtggctg tggcacattc cctgctctgg tcgtgttgaa gtacaacatg 1560

gagacacgtg gggcaccgtc tgtgattctg acttctctct ggaggcggcc agcgtgctgt 1620

gcagggaact acagtgcggc actgtggttt ccctcctggg gggagctcac tttggagaag 1680

gaagtggaca gatctgggct gaagaattcc agtgtgaggg gcacgagtcc cacctttcac 1740

tctgcccagt agcaccccgc cctgacggga catgtagcca cagcagggac gtcggcgtag 1800

tctgctcaag tgagacccag ggaatgtgtt cactttgttc ccatgccatg aagagggtag 1860

ggttaggtag tcacagacat ctttttaaag ccctgtctcc ttccaggata cacacaaatc 1920

cgcttggtga atggcaagac cccatgtgaa ggaagagtgg agctcaacat tcttgggtcc 1980

tgggggtccc tctgcaactc tcactgggac atggaagatg cccatgtttt atgccagcag 2040

cttaaatgtg gagttgccct ttctatcccg ggaggagcac cttttgggaa aggaagtgag 2100

caggtctgga ggcacatgtt tcactgcact gggactgaga agcacatggg agattgttcc 2160

gtcactgctc tgggcgcatc actctgttct tcagggcaag tggcctctgt aatctgctca 2220

ggtaagagaa taagggcagc cagtgatgag ccactcatga cggtgcctta agagtgggtg 2280

tacctaggag ttcccattgt ggctcagtgg taacaaactc gactggtatc catgagggta 2340

tgggtttgat ccctggcctt gctcaatggg ttaaggatcc agcattgctg tgagctgtgg 2400

tataggttgc agactctgct caggtcccat gttgctgtga ttgtggtgta ggctgactgc 2460

tgcagcttca atttgacccc tagcccggga atttccatag gccacacgtg cagcactaag 2520

gaaggaaaaa aagaaaaaaa aaaaaaaaga gtgggtgtgc ctatagtgaa gaacagatgt 2580

aaaagggaag tgaaagggat tcccccattc tgagggattg tgagaagtgt gccagaatat 2640

taacttcatt tgacttgtta cagggaaagt aaacttgact ttcacggacc tcctagttac 2700

ctggtgctta ctatatgtct tctcagagta cctgattcat tcccagcctg gttgacccat 2760

ccccctatct ctatggctat gtttatccag agcacatcta tctaacactc cagctgatct 2820

tcctgacaca gctgtggcaa ccctggatcc tttaaccaac tgtgccaggc tggagatcaa 2880

acctaagcct ctgcagcaac ccaagctgct gcagtcagat ttttaacccc ctgtgccact 2940

gtgggtatct ccgatatttt gtatcttctg tgactgagtg gtttgctgtt tgcagggaac 3000

cagagtcaga cactatcccc gtgcaattca tcatcctcgg acccatcaag ctctattatt 3060

tcagaagaaa atggtgttgc ctgcataggt gagaatcagt gaccaaccta tgaaaatgat 3120

ctcaatcctc tgaaatgcat tttattcatg ttttatttcc tctttgcagg gagtggtcaa 3180

cttcgcctgg tcgatggagg tggtcgttgt gctgggagag tagaggtcta tcatgagggc 3240

tcctggggca ccatctgtga tgacagctgg gacctgaatg atgcccatgt ggtgtgcaaa 3300

cagctgagct gtggatgggc cattaatgcc actggttctg ctcattttgg ggaaggaaca 3360

gggcccattt ggctggatga gataaactgt aatggaaaag aatctcatat ttggcaatgc 3420

cactcacatg gttgggggcg gcacaattgc aggcataagg aggatgcagg agtcatctgc 3480

tcgg 3484

<210> 99
<211> 4997
<212> ADN
<213> Sus scrofa

<400> 99
tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt 540

gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat 600

cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac 660

tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa 720

atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc 780

caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa 840

gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta 900

tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag 960

aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 1020

tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat 1080

actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag 1140

agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg 1200

aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctcctttgt 1260

cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt 1320

ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc 1380

tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact 1440

cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg 1500

atcctgcgtt gctgtggctg tggctgtggt gtaggccagc agcaaacagc tctgattaga 1560

cccctagcct ggaaacctcc atatgccaca ggtgcagccc taaaaagaca aaaaaagaga 1620

aaagacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gaaccccccag 1680

aggtatttat ttgtttttgc ctttttttcac tgactgttct ttgtttgttt gtttgagact 1740

gatctagaag actagagatt acaagaaata tggatttggc tcactctaag aaactgcttt 1800

cattccaagg tttgggtcta tccaaaagtg gaatagaatc atatgaatac tagtttatga 1860

gtatttagtg agaggaattt caagctcaaa taatgattca gcaagattaa attaaggagg 1920

gaattttcct tgtggctgag tgggttaagg acccaatgtt gtctctgtga ggatgtaggt 1980

tccatcctgg gctttgctca ttaggttaag gatctggcat tgctgcagct cagacccagt 2040

gctgccctgg ttgtggctta ggccaaagct gcagctccaa ttcaatctct ggcctgggaa 2100

cctccatgtg ctacaaggtg cggccttaaa aggaaaaaaa aaaaattaaa tcaaggactc 2160

aagagtcttt cattatttgt gttgtggaag ctatatttgt tttaaagtct tagttgtgtt 2220

tagaaagcaa gatgttcttc aactcaaatt tgggagggaa cttgtttcat acatttttaa 2280

tggataagtg gcaaaatttt catgctgagg tgatctatag tgttgtaatg cagaatatag 2340

tcagatcttg aacattttag gaagttggtg agggccaatt gtgtatctgt gccatgctga 2400

taagaatgtc aagggatcac aagaattcgt gttatttgac agcagtcatc tttaaaaggc 2460

atttgagaaa gtccaatttc aaatgcattt ccttttctta aaagataaat tgaagaaaat 2520

aagtctttat ttcccaagta aattgaattg cctctcagtc tgttaaaaga aactcttacc 2580

ttgatgattg cgctcttaac ctggcaaaga ttgtctttaa aatctgagct ccatgtcttc 2640

tgctttattt ctggtgtgcc tttgactcca gattacagta aatggaggac tgagtatagg 2700

gctaaaaagt agagagaatg gatgcatatt atctgtggtc tccaatgtga tgaatgaagt 2760

aggcaaatac tcaaaggaaa gagaaagcat gctccaagaa ttatgggttc cagaaggcaa 2820

agtcccagaa ttgtctccag ggaaggacag ggaggtctag aatcggctaa gcccactgta 2880

ggcagaaaaa ccaagaggca tgaatggctt ccctttctca cttttcactc tctggcttac 2940

tcctatcatg aaggaaaata ttggaatcat attctccctc accgaaatgc tatttttcag 3000

cccacaggaa acccaggctg gttggagggg acattccctg ctctggtcgt gttgaagtac 3060

aacatggaga cacgtggggc accgtctgtg attctgactt ctctctggag gcggccagcg 3120

tgctgtgcag ggaactacag tgcggctact actactgtgg tttccctcct gggggggagct 3180

cactttggag aaggaagtgg acagatctgg gctgaagaat tccagtgtga ggggcacgag 3240

tcccaccttt cactctgccc agtagcaccc cgccctgacg ggacatgtag ccacagcagg 3300

gacgtcggcg tagtctgctc aagtgagacc cagggaatgt gttcactttg ttcccatgcc 3360

atgaagaggg tagggttagg tagtcacaga catcttttta aagccctgtc tccttccagg 3420

atacacacaa atccgcttgg tgaatggcaa gaccccatgt gaaggaagag tggagctcaa 3480

cattcttggg tcctgggggt ccctctgcaa ctctcactgg gacatggaag atgcccatgt 3540

tttatgccag cagcttaaat gtggagttgc cctttctatc ccgggaggag caccttttgg 3600

gaaaggaagt gagcaggtct ggaggcacat gtttcactgc actgggactg agaagcacat 3660

gggagattgt tccgtcactg ctctgggcgc atcactctgt tcttcagggc aagtggcctc 3720

tgtaatctgc tcaggtaaga gaataagggc agccagtgat gagccactca tgacggtgcc 3780

ttaagagtgg gtgtacctag gagttcccat tgtggctcag tggtaacaaa ctcgactggt 3840

atccatgagg gtatgggttt gatccctggc cttgctcaat gggttaagga tccagcattg 3900

ctgtgagctg tggtataggt tgcagactct gctcaggtcc catgttgctg tgattgtggt 3960

gtaggctgac tgctgcagct tcaatttgac ccctagcccg ggaatttcca taggccacac 4020

gtgcagcact aaggaaggaa aaaaagaaaa aaaaaaaaaa agagtgggtg tgcctatagt 4080

gaagaacaga tgtaaaaggg aagtgaaagg gattccccca ttctgaggga ttgtgagaag 4140

tgtgccagaa tattaacttc atttgacttg ttacagggaa agtaaacttg actttcacgg 4200

acctcctagt tacctggtgc ttactatatg tcttctcaga gtacctgatt cattcccagc 4260

ctggttgacc catcccccta tctctatggc tatgtttatc cagagcacat ctatctaaca 4320

ctccagctga tcttcctgac acagctgtgg caaccctgga tcctttaacc aactgtgcca 4380

ggctggagat caaacctaag cctctgcagc aacccaagct gctgcagtca gatttttaac 4440

cccctgtgcc actgtgggta tctccgatat tttgtatctt ctgtgactga gtggtttgct 4500

gtttgcaggg aaccagagtc agacactatc cccgtgcaat tcatcatcct cggacccatc 4560

aagctctatt atttcagaag aaaatggtgt tgcctgcata ggtgagaatc agtgaccaac 4620

ctatgaaaat gatctcaatc ctctgaaatg cattttattc atgttttatt tcctctttgc 4680

agggagtggt caacttcgcc tggtcgatgg aggtggtcgt tgtgctggga gagtagaggt 4740

ctatcatgag ggctcctggg gcaccatctg tgatgacagc tgggacctga atgatgccca 4800

tgtggtgtgc aaacagctga gctgtggatg ggccattaat gccactggtt ctgctcattt 4860

tggggaagga acagggccca tttggctgga tgagataaac tgtaatggaa aagaatctca 4920

tatttggcaa tgccactcac atggttgggg gcggcacaat tgcaggcata aggaggatgc 4980

aggagtcatc tgctcgg 4997

<210> 100
<211> 3710
<212> ADN
<213> Sus scrofa

<400> 100
tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt 540

gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat 600

cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac 660

tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa 720

atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc 780

caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa 840

gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta 900

tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag 960

aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 1020

tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat 1080

actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag 1140

agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg 1200

aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctcctttgt 1260

cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt 1320

ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc 1380

tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact 1440

cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg 1500

atcctgcgtt gctgtggctg tggctgtggt gtaggccagc agcaaacagc tctgattaga 1560

cccctagcct ggaaacctcc atatgccaca ggtgcagccc taaaaagaca aaaaaagaga 1620

aaagacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gaacccccag 1680

aggtatttat ttgtttttgc ctttttttcac tgactgttct ttgtttgttt gtttgagact 1740

gatctagaag actagagatt acaagaaata tggatttggc tcactctaag aaactgcttt 1800

cattccaagg tttgggtcta tccaaaagtg gaatagaatc atatgaatac tagtttatga 1860

gtatttagtg agaggaattt caagctcaaa taatgattca gcaagattaa attaaggagg 1920

gaattttcct tgtggctgag tgggttaagg acccaatgtt gtctctgtga ggatgtaggt 1980

tccatcctgg gctttgctca ttaggttaag gatctggcat tgctgcagct cagacccagt 2040

gctgccctgg ttgtggctta ggccaaagct gcagctccaa ttcaatctct ggcctgggaa 2100

cctccatgtg ctacaaggtg cggccttaaa aggaaaaaaa aaaaattaaa tcaaggactc 2160

aagagtcttt cattatttgt gttgtggaag ctatatttgt tttaaagtct tagttgtgtt 2220

tagaaagcaa gatgttcttc aactcaaatt tgggagggaa cttgtttcat acatttttaa 2280

tggataagtg gcaaaatttt catgctgagg tgatctatag tgttgtaatg cagaatatag 2340

tcagatcttg aacattttag gaagttggtg agggccaatt gtgtatctgt gccatgctga 2400

taagaatgtc aagggatcac aagaattcgt gttatttgac agcagtcatc tttaaaaggc 2460

atttgagaaa gtccaatttc aaatgcattt ccttttcttta aaagataaat tgaagaaaat 2520

aagtctttat ttcccaagta aattgaattg cctctcagtc tgttaaaaga aactcttacc 2580

ttgatgattg cgctcttaac ctggcaaaga ttgtctttaa aatctgagct ccatgtcttc 2640

tgctttattt ctggtgtgcc tttgactcca gattacagta aatggaggac tgagtatagg 2700

gctaaaaagt agagagaatg gatgcatatt atctgtggtc tccaatgtga tgaatgaagt 2760

aggcaaatac tcaaaggaaa gagaaagcat gctccaagaa ttatgggttc cagaagggaa 2820

agggattccc ccattctgag ggattgtgag aagtgtgcca gaatattaac ttcatttgac 2880

ttgttacagg gaaagtaaac ttgactttca cggacctcct agttacctgg tgcttactat 2940

atgtcttctc agagtacctg attcattccc agcctggttg acccatcccc ctatctctat 3000

ggctatgttt atccagagca catctatcta acactccagc tgatcttcct gacacagctg 3060

tggcaaccct ggatccttta accaactgtg ccaggctgga gatcaaacct aagcctctgc 3120

agcaacccaa gctgctgcag tcagattttt aacccccctgt gccactgtgg gtatctccga 3180

tattttgtat cttctgtgac tgagtggttt gctgtttgca gggaaccaga gtcagacact 3240

atccccgtgc aattcatcat cctcggaccc atcaagctct attatttcag aagaaaatgg 3300

tgttgcctgc ataggtgaga atcagtgacc aacctatgaa aatgatctca atcctctgaa 3360

atgcatttta ttcatgtttt atttcctctt tgcagggagt ggtcaacttc gcctggtcga 3420

tggaggtggt cgttgtgctg ggagagtaga ggtctatcat gagggctcct ggggcaccat 3480

ctgtgatgac agctgggacc tgaatgatgc ccatgtggtg tgcaaacagc tgagctgtgg 3540

atgggccatt aatgccactg gttctgctca ttttggggaa ggaacagggc ccatttggct 3600

ggatgagata aactgtaatg gaaaagaatc tcatatttgg caatgccact cacatggttg 3660

ggggcggcac aattgcaggc ataaggagga tgcaggagtc atctgctcgg 3710

<210> 101
<211> 3617
<212> ADN
<213> Sus scrofa

<400> 101
tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt 540

gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat 600

cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac 660

tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa 720

atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc 780

caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa 840

gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta 900

tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag 960

aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 1020

tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat 1080

actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag 1140

agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg 1200

aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctcctttgt 1260

cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt 1320

ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc 1380

tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact 1440

cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg 1500

atcctgcgtt gctgtggctg tggctgtggt gtaggccagc agcaaacagc tctgattaga 1560

cccctagcct ggaaacctcc atatgccaca ggtgcagccc taaaaagaca aaaaaagaga 1620

aaagacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gaacccccag 1680

aggtatttat ttgtttttgc ctttttttcac tgactgttct ttgtttgttt gtttgagact 1740

gatctagaag actagagatt acaagaaata tggatttggc tcactctaag aaactgcttt 1800

cattccaagg tttgggtcta tccaaaagtg gaatagaatc atatgaatac tagtttatga 1860

gtatttagtg agaggaattt caagctcaaa taatgattca gcaagattaa attaaggagg 1920

gaattttcct tgtggctgag tgggttaagg acccaatgtt gtctctgtga ggatgtaggt 1980

tccatcctgg gctttgctca ttaggttaag gatctggcat tgctgcagct cagacccagt 2040

gctgccctgg ttgtggctta ggccaaagct gcagctccaa ttcaatctct ggcctgggaa 2100

cctccatgtg ctacaaggtg cggccttaaa aggaaaaaaa aaaaattaaa tcaaggactc 2160

aagagtcttt cattatttgt gttgtggaag ctatatttgt tttaaagtct tagttgtgtt 2220

tagaaagcaa gatgttcttc aactcaaatt tgggagggaa cttgtttcat acatttttaa 2280

tggataagtg gcaaaatttt catgctgagg tgatctatag tgttgtaatg cagaatatag 2340

tcagatcttg aacattttag gaagttggtg agggccaatt gtgtatctgt gccatgctga 2400

taagaatgtc aagggatcac aagaattcgt gttatttgac agcagtcatc tttaaaaggc 2460

atttgagaaa gtccaatttc aaatgcattt cctttcttta aaagataaat tgaagaaaat 2520

aagtctttat ttcccaagta aattgaattg cctctcagtc tgttaaaaga aactcttacc 2580

ttgatgattg cgctcttaac ctggcaaaga ttgtctttaa aatctgagct ccatgtcttc 2640

tgctttattt ctggtgtgcc tttgactcca gattacagta aatggaggac tgagtatagg 2700

gctaaaaagt agagagaatg gattgaaagg gattccccca ttctgaggga ttgtgagaag 2760

tgtgccagaa tattaacttc atttgacttg ttacagggaa agtaaacttg actttcacgg 2820

acctcctagt tacctggtgc ttactatatg tcttctcaga gtacctgatt cattcccagc 2880

ctggttgacc catcccccta tctctatggc tatgtttatc cagagcacat ctatctaaca 2940

ctccagctga tcttcctgac acagctgtgg caaccctgga tcctttaacc aactgtgcca 3000

ggctggagat caaacctaag cctctgcagc aacccaagct gctgcagtca gatttttaac 3060

cccctgtgcc actgtgggta tctccgatat tttgtatctt ctgtgactga gtggtttgct 3120

gtttgcaggg aaccagagtc agacactatc cccgtgcaat tcatcatcct cggacccatc 3180

aagctctatt atttcagaag aaaatggtgt tgcctgcata ggtgagaatc agtgaccaac 3240

ctatgaaaat gatctcaatc ctctgaaatg cattttattc atgttttatt tcctctttgc 3300

agggagtggt caacttcgcc tggtcgatgg aggtggtcgt tgtgctggga gagtagaggt 3360

ctatcatgag ggctcctggg gcaccatctg tgatgacagc tgggacctga atgatgccca 3420

tgtggtgtgc aaacagctga gctgtggatg ggccattaat gccactggtt ctgctcattt 3480

tggggaagga acagggccca tttggctgga tgagataaac tgtaatggaa aagaatctca 3540

tatttggcaa tgccactcac atggttgggg gcggcacaat tgcaggcata aggaggatgc 3600

aggagtcatc tgctcgg 3617

<210> 102
<211> 4979
<212> ADN
<213> Sus scrofa

<400> 102
tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt 540

gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat 600

cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac 660

tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa 720

atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc 780

caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa 840

gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta 900

tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag 960

aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 1020

tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat 1080

actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag 1140

agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg 1200

aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctcctttgt 1260

cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt 1320

ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc 1380

tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact 1440

cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg 1500

atcctgcgtt gctgtggctg tggctgtggt gtaggccagc agcaaacagc tctgattaga 1560

cccctagcct ggaaacctcc atatgccaca ggtgcagccc taaaaagaca aaaaaagaga 1620

aaagacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gaacccccag 1680

aggtatttat ttgtttttgc ctttttttcac tgactgttct ttgtttgttt gtttgagact 1740

gatctagaag actagagatt acaagaaata tggatttggc tcactctaag aaactgcttt 1800

cattccaagg tttgggtcta tccaaaagtg gaatagaatc atatgaatac tagtttatga 1860

gtatttagtg agaggaattt caagctcaaa taatgattca gcaagattaa attaaggagg 1920

gaattttcct tgtggctgag tgggttaagg acccaatgtt gtctctgtga ggatgtaggt 1980

tccatcctgg gctttgctca ttaggttaag gatctggcat tgctgcagct cagacccagt 2040

gctgccctgg ttgtggctta ggccaaagct gcagctccaa ttcaatctct ggcctgggaa 2100

cctccatgtg ctacaaggtg cggccttaaa aggaaaaaaa aaaaattaaa tcaaggactc 2160

aagagtcttt cattatttgt gttgtggaag ctatatttgt tttaaagtct tagttgtgtt 2220

tagaaagcaa gatgttcttc aactcaaatt tgggagggaa cttgtttcat acatttttaa 2280

tggataagtg gcaaaatttt catgctgagg tgatctatag tgttgtaatg cagaatatag 2340

tcagatcttg aacattttag gaagttggtg agggccaatt gtgtatctgt gccatgctga 2400

taagaatgtc aagggatcac aagaattcgt gttatttgac agcagtcatc tttaaaaggc 2460

atttgagaaa gtccaatttc aaatgcattt cctttcttta aaagataaat tgaagaaaat 2520

aagtctttat ttcccaagta aattgaattg cctctcagtc tgttaaaaga aactcttacc 2580

ttgatgattg cgctcttaac ctggcaaaga ttgtctttaa aatctgagct ccatgtcttc 2640

tgctttattt ctggtgtgcc tttgactcca gattacagta aatggaggac tgagtatagg 2700

gctaaaaagt agagagaatg gatgcatatt atctgtggtc tccaatgtga tgaatgaagt 2760

aggcaaatac tcaaaggaaa gagaaagcat gctccaagaa ttatgggttc cagaaggcaa 2820

agtcccagaa ttgtctccag ggaaggacag ggaggtctag aatcggctaa gcccactgta 2880

ggcagaaaaa ccaagaggca tgaatggctt ccctttctca cttttcactc tctggcttac 2940

tcctatcatg aaggaaaata ttggaatcat attctccctc accgaaatgc tatttttcag 3000

cccacaggaa acccaggctg gttggagggg acattccctg ctctggtcgt gttgaagtac 3060

aacatggaga cacgtggggc accgtctgtg attctgactt ctctctggag gcggccagcg 3120

tgctgtgcag ggaactctgt ggtttccctc ctgggggggag ctcactttgg agaaggaagt 3180

ggacagatct gggctgaaga attccagtgt gaggggcacg agtcccacct ttcactctgc 3240

ccagtagcac cccgccctga cgggacatgt agccacagca gggacgtcgg cgtagtctgc 3300

tcaagtgaga cccagggaat gtgttcactt tgttcccatg ccatgaagag ggtagggtta 3360

ggtagtcaca gacatctttt taaagccctg tctccttcca ggatacacac aaatccgctt 3420

ggtgaatggc aagaccccat gtgaaggaag agtggagctc aacattcttg ggtcctgggg 3480

gtccctctgc aactctcact gggacatgga agatgcccat gttttatgcc agcagcttaa 3540

atgtggagtt gccctttcta tcccgggagg agcacctttt gggaaaggaa gtgagcaggt 3600

ctggaggcac atgtttcact gcactgggac tgagaagcac atgggagatt gttccgtcac 3660

tgctctgggc gcatcactct gttcttcagg gcaagtggcc tctgtaatct gctcaggtaa 3720

gagaataagg gcagccagtg atgagccact catgacggtg ccttaagagt gggtgtacct 3780

aggagttccc attgtggctc agtggtaaca aactcgactg gtatccatga gggtatgggt 3840

ttgatccctg gccttgctca atgggttaag gatccagcat tgctgtgagc tgtggtatag 3900

gttgcagact ctgctcaggt cccatgttgc tgtgattgtg gtgtaggctg actgctgcag 3960

cttcaatttg acccctagcc cgggaatttc cataggccac acgtgcagca ctaaggaagg 4020

aaaaaaagaa aaaaaaaaaa aaagagtggg tgtgcctata gtgaagaaca gatgtaaaag 4080

ggaagtgaaa gggattcccc cattctgagg gattgtgaga agtgtgccag aatattaact 4140

tcatttgact tgttacaggg aaagtaaact tgactttcac ggacctccta gttacctggt 4200

gcttactata tgtcttctca gagtacctga ttcattccca gcctggttga cccatccccc 4260

tatctctatg gctatgttta tccagagcac atctatctaa cactccagct gatcttcctg 4320

acacagctgt ggcaaccctg gatcctttaa ccaactgtgc caggctggag atcaaaccta 4380

agcctctgca gcaacccaag ctgctgcagt cagattttta acccccctgtg ccactgtggg 4440

tatctccgat attttgtatc ttctgtgact gagtggtttg ctgtttgcag ggaaccagag 4500

tcagacacta tccccgtgca attcatcatc ctcggaccca tcaagctcta ttatttcaga 4560

agaaaatggt gttgcctgca taggtgagaa tcagtgacca acctatgaaa atgatctcaa 4620

tcctctgaaa tgcattttat tcatgtttta tttcctcttt gcagggagtg gtcaacttcg 4680

cctggtcgat ggaggtggtc gttgtgctgg gagagtagag gtctatcatg agggctcctg 4740

gggcaccatc tgtgatgaca gctgggacct gaatgatgcc catgtggtgt gcaaacagct 4800

gagctgtgga tgggccatta atgccactgg ttctgctcat tttggggaag gaacagggcc 4860

catttggctg gatgagataa actgtaatgg aaaagaatct catatttggc aatgccactc 4920

acatggttgg gggcggcaca attgcaggca taaggaggat gcaggagtca tctgctcgg 4979

<210> 103
<211> 4615
<212> ADN
<213> Sus scrofa

<400> 103
tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt 540

gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat 600

cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac 660

tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa 720

atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc 780

caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa 840

gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta 900

tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag 960

aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 1020

tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat 1080

actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag 1140

agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg 1200

aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctcctttgt 1260

cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt 1320

ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc 1380

tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact 1440

cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg 1500

atcctgcgtt gctgtggctg tggctgtggt gtaggccagc agcaaacagc tctgattaga 1560

cccctagcct ggaaacctcc atatgccaca ggtgcagccc taaaaagaca aaaaaagaga 1620

aaagacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gaacccccag 1680

aggtatttat ttgtttttgc ctttttttcac tgactgttct ttgtttgttt gtttgagact 1740

gatctagaag actagagatt acaagaaata tggatttggc tcactctaag aaactgcttt 1800

cattccaagg tttgggtcta tccaaaagtg gaatagaatc atatgaatac tagtttatga 1860

gtatttagtg agaggaattt caagctcaaa taatgattca gcaagattaa attaaggagg 1920

gaattttcct tgtggctgag tgggttaagg acccaatgtt gtctctgtga ggatgtaggt 1980

tccatcctgg gctttgctca ttaggttaag gatctggcat tgctgcagct cagacccagt 2040

gctgccctgg ttgtggctta ggccaaagct gcagctccaa ttcaatctct ggcctgggaa 2100

cctccatgtg ctacaaggtg cggccttaaa aggaaaaaaa aaaaattaaa tcaaggactc 2160

aagagtcttt cattatttgt gttgtggaag ctatatttgt tttaaagtct tagttgtgtt 2220

tagaaagcaa gatgttcttc aactcaaatt tgggagggaa cttgtttcat acatttttaa 2280

tggataagtg gcaaaatttt catgctgagg tgatctatag tgttgtaatg cagaatatag 2340

tcagatcttg aacattttag gaagttggtg agggccaatt gtgtatctgt gccatgctga 2400

taagaatgtc aagggatcac aagaattcgt gttatttgac agcagtcatc tttaaaaggc 2460

atttgagaaa gtccaatttc aaatgcattt cctttcttta aaagataaat tgaagaaaat 2520

aagtctttat ttcccaagta aattgaattg cctctcagtc tgttaaaaga aagaaggaaa 2580

atattggaat catattctcc ctcaccgaaa tgctattttt cagcccacag gaaacccagg 2640

ctggttggag gggacattcc ctgctctggt cgtgttgaag tacaacatgg agacacgtgg 2700

ggcaccgtct gtgattctga cttctctctg gaggcggcca gcgtgctgtg cagggaacta 2760

cagtgcggca cagtgtggtt tccctcctgg ggggagctca ctttggagaa ggaagtggac 2820

agatctgggc tgaagaattc cagtgtgagg ggcacgagtc ccacctttca ctctgcccag 2880

tagcaccccg ccctgacggg acatgtagcc acagcaggga cgtcggcgta gtctgctcaa 2940

gtgagaccca gggaatgtgt tcactttgtt cccatgccat gaagagggta gggttaggta 3000

gtcacagaca tcttttaaa gccctgtctc cttccaggat acacacaaat ccgcttggtg 3060

aatggcaaga ccccatgtga aggaagagtg gagctcaaca ttcttgggtc ctgggggtcc 3120

ctctgcaact ctcactggga catggaagat gcccatgttt tatgccagca gcttaaatgt 3180

ggagttgccc tttctatccc gggaggagca ccttttggga aaggaagtga gcaggtctgg 3240

aggcacatgt ttcactgcac tgggactgag aagcacatgg gagattgttc cgtcactgct 3300

ctgggcgcat cactctgttc ttcagggcaa gtggcctctg taatctgctc aggtaagaga 3360

ataagggcag ccagtgatga gccactcatg acggtgcctt aagagtgggt gtacctagga 3420

gttcccattg tggctcagtg gtaacaaact cgactggtat ccatgagggt atgggtttga 3480

tccctggcct tgctcaatgg gttaaggatc cagcattgct gtgagctgtg gtataggttg 3540

cagactctgc tcaggtccca tgttgctgtg attgtggtgt aggctgactg ctgcagcttc 3600

aatttgaccc ctagcccggg aatttccata ggccacacgt gcagcactaa ggaaggaaaa 3660

aaagaaaaaa aaaaaaaaag agtgggtgtg cctatagtga agaacagatg taaaagggaa 3720

gtgaaaggga ttcccccatt ctgagggatt gtgagaagtg tgccagaata ttaacttcat 3780

ttgacttgtt acagggaaag taaacttgac tttcacggac ctcctagtta cctggtgctt 3840

actatatgtc ttctcagagt acctgattca ttcccagcct ggttgaccca tccccctatc 3900

tctatggcta tgtttatcca gagcacatct atctaacact ccagctgatc ttcctgacac 3960

agctgtggca accctggatc ctttaaccaa ctgtgccagg ctggagatca aacctaagcc 4020

tctgcagcaa cccaagctgc tgcagtcaga tttttaaccc cctgtgccac tgtgggtatc 4080

tccgatattt tgtatcttct gtgactgagt ggtttgctgt ttgcagggaa ccagagtcag 4140

acactatccc cgtgcaattc atcatcctcg gacccatcaa gctctattat ttcagaagaa 4200

aatggtgttg cctgcatagg tgagaatcag tgaccaacct atgaaaatga tctcaatcct 4260

ctgaaatgca ttttattcat gttttatttc ctctttgcag ggagtggtca acttcgcctg 4320

gtcgatggag gtggtcgttg tgctgggaga gtagaggtct atcatgaggg ctcctggggc 4380

accatctgtg atgacagctg ggacctgaat gatgcccatg tggtgtgcaa acagctgagc 4440

tgtggatggg ccattaatgc cactggttct gctcattttg gggaaggaac agggcccatt 4500

tggctggatg agataaactg taatggaaaa gaatctcata tttggcaatg ccactcacat 4560

ggttgggggc ggcacaattg caggcataag gaggatgcag gagtcatctg ctcgg 4615

<210> 104
<211> 4866
<212> ADN
<213> Sus scrofa

<400> 104
tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt 540

gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat 600

cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac 660

tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa 720

atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc 780

caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa 840

gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta 900

tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag 960

aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 1020

tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat 1080

actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag 1140

agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg 1200

aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctcctttgt 1260

cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt 1320

ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc 1380

tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact 1440

cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg 1500

atcctgcgtt gctgtggctg tggctgtggt gtaggccagc agcaaacagc tctgattaga 1560

cccctagcct ggaaacctcc atatgccaca ggtgcagccc taaaaagaca aaaaaagaga 1620

aaagacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gaacccccag 1680

aggtatttat ttgtttttgc ctttttttcac tgactgttct ttgtttgttt gtttgagact 1740

gatctagaag actagagatt acaagaaata tggatttggc tcactctaag aaactgcttt 1800

cattccaagg tttgggtcta tccaaaagtg gaatagaatc atatgaatac tagtttatga 1860

gtatttagtg agaggaattt caagctcaaa taatgattca gcaagattaa attaaggagg 1920

gaattttcct tgtggctgag tgggttaagg acccaatgtt gtctctgtga ggatgtaggt 1980

tccatcctgg gctttgctca ttaggttaag gatctggcat tgctgcagct cagacccagt 2040

gctgccctgg ttgtggctta ggccaaagct gcagctccaa ttcaatctct ggcctgggaa 2100

cctccatgtg ctacaaggtg cggccttaaa aggaaaaaaa aaaaattaaa tcaaggactc 2160

aagagtcttt cattatttgt gttgtggaag ctatatttgt tttaaagtct tagttgtgtt 2220

tagaaagcaa gatgttcttc aactcaaatt tgggagggaa cttgtttcat acatttttaa 2280

tggataagtg gcaaaatttt catgctgagg tgatctatag tgttgtaatg cagaatatag 2340

tcagatcttg aacattttag gaagttggtg agggccaatt gtgtatctgt gccatgctga 2400

taagaatgtc aagggatcac aagaattcgt gttatttgac agcagtcatc tttaaaaggc 2460

atttgagaaa gtccaatttc aaatgcattt cctttcttta aaagataaat tgaagaaaat 2520

aagtctttat ttcccaagta aattgaattg cctctcagtc tgttaaaaga aactcttacc 2580

ttgatgattg cgctcttaac ctggcaaaga ttgtctttaa aatctgagct ccatgtcttc 2640

tgctttattt ctggtgtgcc tttgactcca gattacagta aatggaggac tgagtatagg 2700

gctaaaaagt agagagaatg gatgcatatt atctgtggtc tccaatgtga tgaatgaagt 2760

aggcaaatac tcaaaggaaa gagaaagcat gctccaagaa ttatgggttc cagaaggcaa 2820

agtcccagaa ttgtctccag ggaaggacag ggaggtctag aatcggctaa gcccactgta 2880

ggcagaaaaa ccaagaggca tgaatggctt ccctttctca cttttcactc tctggcttac 2940

tcctatcatg aaggaaaata ttggaatcat attctccctc accgaaatgc tatttttcag 3000

cccacaggaa acccaggctg gttctgtggt ttccctcctg gggggagctc actttggaga 3060

aggaagtgga cagatctggg ctgaagaatt ccagtgtgag gggcacgagt cccacctttc 3120

actctgccca gtagcacccc gccctgacgg gacatgtagc cacagcaggg acgtcggcgt 3180

agtctgctca agtgagaccc agggaatgtg ttcactttgt tcccatgcca tgaagagggt 3240

agggttaggt agtcacagac atctttttaa agccctgtct ccttccagga tacacacaaa 3300

tccgcttggt gaatggcaag accccatgtg aaggaagagt ggagctcaac attcttgggt 3360

cctgggggtc cctctgcaac tctcactggg acatggaaga tgcccatgtt ttatgccagc 3420

agcttaaatg tggagttgcc ctttctatcc cgggaggagc acctttgggg aaaggaagtg 3480

agcaggtctg gaggcacatg tttcactgca ctgggactga gaagcacatg ggagattgtt 3540

ccgtcactgc tctgggcgca tcactctgtt cttcagggca agtggcctct gtaatctgct 3600

caggtaagag aataagggca gccagtgatg agccactcat gacggtgcct taagagtggg 3660

tgtacctagg agttcccatt gtggctcagt ggtaacaaac tcgactggta tccatgaggg 3720

tatgggtttg atccctggcc ttgctcaatg ggttaaggat ccagcattgc tgtgagctgt 3780

ggtataggtt gcagactctg ctcaggtccc atgttgctgt gattgtggtg taggctgact 3840

gctgcagctt caatttgacc cctagcccgg gaatttccat aggccacacg tgcagcacta 3900

aggaaggaaa aaaagaaaaa aaaaaaaaaa gagtgggtgt gcctatagtg aagaacagat 3960

gtaaaaggga agtgaaaggg attcccccat tctgagggat tgtgagaagt gtgccagaat 4020

attaacttca tttgacttgt tacagggaaa gtaaacttga ctttcacgga cctcctagtt 4080

acctggtgct tactatatgt cttctcagag tacctgattc attcccagcc tggttgaccc 4140

atccccctat ctctatggct atgtttatcc agagcacatc tatctaacac tccagctgat 4200

cttcctgaca cagctgtggc aaccctggat cctttaacca actgtgccag gctggagatc 4260

aaacctaagc ctctgcagca acccaagctg ctgcagtcag atttttaacc ccctgtgcca 4320

ctgtgggtat ctccgatatt ttgtatcttc tgtgactgag tggtttgctg tttgcaggga 4380

accagagtca gacactatcc ccgtgcaatt catcatcctc ggacccatca agctctatta 4440

tttcagaaga aaatggtgtt gcctgcatag gtgagaatca gtgaccaacc tatgaaaatg 4500

atctcaatcc tctgaaatgc attttattca tgttttattt cctctttgca gggagtggtc 4560

aacttcgcct ggtcgatgga ggtggtcgtt gtgctgggag agtagaggtc tatcatgagg 4620

gctcctgggg caccatctgt gatgacagct gggacctgaa tgatgcccat gtggtgtgca 4680

aacagctgag ctgtggatgg gccattaatg ccactggttc tgctcatttt ggggaaggaa 4740

cagggcccat ttggctggat gagataaact gtaatggaaa agaatctcat atttggcaat 4800

gccactcaca tggttggggg cggcacaatt gcaggcataa ggaggatgca ggagtcatct 4860

gctcgg 4866

<210> 105
<211> 4867
<212> ADN
<213> Sus scrofa

<400> 105
tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt 540

gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat 600

cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac 660

tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa 720

atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc 780

caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa 840

gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta 900

tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag 960

aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 1020

tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat 1080

actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag 1140

agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg 1200

aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctcctttgt 1260

cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt 1320

ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc 1380

tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact 1440

cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg 1500

atcctgcgtt gctgtggctg tggctgtggt gtaggccagc agcaaacagc tctgattaga 1560

cccctagcct ggaaacctcc atatgccaca ggtgcagccc taaaaagaca aaaaaagaga 1620

aaagacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gaacccccag 1680

aggtatttat ttgtttttgc ctttttttcac tgactgttct ttgtttgttt gtttgagact 1740

gatctagaag actagagatt acaagaaata tggatttggc tcactctaag aaactgcttt 1800

cattccaagg tttgggtcta tccaaaagtg gaatagaatc atatgaatac tagtttatga 1860

gtatttagtg agaggaattt caagctcaaa taatgattca gcaagattaa attaaggagg 1920

gaattttcct tgtggctgag tgggttaagg acccaatgtt gtctctgtga ggatgtaggt 1980

tccatcctgg gctttgctca ttaggttaag gatctggcat tgctgcagct cagacccagt 2040

gctgccctgg ttgtggctta ggccaaagct gcagctccaa ttcaatctct ggcctgggaa 2100

cctccatgtg ctacaaggtg cggccttaaa aggaaaaaaa aaaaattaaa tcaaggactc 2160

aagagtcttt cattatttgt gttgtggaag ctatatttgt tttaaagtct tagttgtgtt 2220

tagaaagcaa gatgttcttc aactcaaatt tgggagggaa cttgtttcat acatttttaa 2280

tggataagtg gcaaaatttt catgctgagg tgatctatag tgttgtaatg cagaatatag 2340

tcagatcttg aacattttag gaagttggtg agggccaatt gtgtatctgt gccatgctga 2400

taagaatgtc aagggatcac aagaattcgt gttatttgac agcagtcatc tttaaaaggc 2460

atttgagaaa gtccaatttc aaatgcattt cctttcttta aaagataaat tgaagaaaat 2520

aagtctttat ttcccaagta aattgaattg cctctcagtc tgttaaaaga aactcttacc 2580

ttgatgattg cgctcttaac ctggcaaaga ttgtctttaa aatctgagct ccatgtcttc 2640

tgctttattt ctggtgtgcc tttgactcca gattacagta aatggaggac tgagtatagg 2700

gctaaaaagt agagagaatg gatgcatatt atctgtggtc tccaatgtga tgaatgaagt 2760

aggcaaatac tcaaaggaaa gagaaagcat gctccaagaa ttatgggttc cagaaggcaa 2820

agtcccagaa ttgtctccag ggaaggacag ggaggtctag aatcggctaa gcccactgta 2880

ggcagaaaaa ccaagaggca tgaatggctt ccctttctca cttttcactc tctggcttac 2940

tcctatcatg aaggaaaata ttggaatcat attctccctc accgaaatgc tatttttcag 3000

cccacaggaa acccaggctg gttactgtgg tttccctcct ggggggagct cactttggag 3060

aaggaagtgg acagatctgg gctgaagaat tccagtgtga ggggcacgag tcccaccttt 3120

cactctgccc agtagcaccc cgccctgacg ggacatgtag ccacagcagg gacgtcggcg 3180

tagtctgctc aagtgagacc cagggaatgt gttcactttg ttcccatgcc atgaagaggg 3240

tagggttagg tagtcacaga catcttttta aagccctgtc tccttccagg atacacacaa 3300

atccgcttgg tgaatggcaa gaccccatgt gaaggaagag tggagctcaa cattcttggg 3360

tcctgggggt ccctctgcaa ctctcactgg gacatggaag atgcccatgt tttatgccag 3420

cagcttaaat gtggagttgc cctttctatc ccgggaggag caccttttgg gaaaggaagt 3480

gagcaggtct ggaggcacat gtttcactgc actgggactg agaagcacat gggagattgt 3540

tccgtcactg ctctgggcgc atcactctgt tcttcagggc aagtggcctc tgtaatctgc 3600

tcaggtaaga gaataagggc agccagtgat gagccactca tgacggtgcc ttaagagtgg 3660

gtgtacctag gagttcccat tgtggctcag tggtaacaaa ctcgactggt atccatgagg 3720

gtatgggttt gatccctggc cttgctcaat gggttaagga tccagcattg ctgtgagctg 3780

tggtataggt tgcagactct gctcaggtcc catgttgctg tgattgtggt gtaggctgac 3840

tgctgcagct tcaatttgac ccctagcccg ggaatttcca taggccacac gtgcagcact 3900

aaggaaggaa aaaaagaaaa aaaaaaaaaa agagtgggtg tgcctatagt gaagaacaga 3960

tgtaaaaggg aagtgaaagg gattccccca ttctgaggga ttgtgagaag tgtgccagaa 4020

tattaacttc atttgacttg ttacagggaa agtaaacttg actttcacgg acctcctagt 4080

tacctggtgc ttactatatg tcttctcaga gtacctgatt cattcccagc ctggttgacc 4140

catcccccta tctctatggc tatgtttatc cagagcacat ctatctaaca ctccagctga 4200

tcttcctgac acagctgtgg caaccctgga tcctttaacc aactgtgcca ggctggagat 4260

caaacctaag cctctgcagc aacccaagct gctgcagtca gatttttaac cccctgtgcc 4320

actgtgggta tctccgatat tttgtatctt ctgtgactga gtggtttgct gtttgcaggg 4380

aaccagagtc agacactatc cccgtgcaat tcatcatcct cggacccatc aagctctatt 4440

atttcagaag aaaatggtgt tgcctgcata ggtgagaatc agtgaccaac ctatgaaaat 4500

gatctcaatc ctctgaaatg cattttattc atgttttatt tcctctttgc agggagtggt 4560

caacttcgcc tggtcgatgg aggtggtcgt tgtgctggga gagtagaggt ctatcatgag 4620

ggctcctggg gcaccatctg tgatgacagc tgggacctga atgatgccca tgtggtgtgc 4680

aaacagctga gctgtggatg ggccattaat gccactggtt ctgctcattt tggggaagga 4740

acagggccca tttggctgga tgagataaac tgtaatggaa aagaatctca tatttggcaa 4800

tgccactcac atggttgggg gcggcacaat tgcaggcata aggaggatgc aggagtcatc 4860

tgctcgg 4867

<210> 106
<211> 4991
<212> ADN
<213> Sus scrofa

<400> 106
tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt 540

gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat 600

cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac 660

tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa 720

atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc 780

caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa 840

gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta 900

tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag 960

aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 1020

tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat 1080

actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag 1140

agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg 1200

aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctcctttgt 1260

cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt 1320

ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc 1380

tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact 1440

cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg 1500

atcctgcgtt gctgtggctg tggctgtggt gtaggccagc agcaaacagc tctgattaga 1560

cccctagcct ggaaacctcc atatgccaca ggtgcagccc taaaaagaca aaaaaagaga 1620

aaagacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gaacccccag 1680

aggtatttat ttgtttttgc ctttttttcac tgactgttct ttgtttgttt gtttgagact 1740

gatctagaag actagagatt acaagaaata tggatttggc tcactctaag aaactgcttt 1800

cattccaagg tttgggtcta tccaaaagtg gaatagaatc atatgaatac tagtttatga 1860

gtatttagtg agaggaattt caagctcaaa taatgattca gcaagattaa attaaggagg 1920

gaattttcct tgtggctgag tgggttaagg acccaatgtt gtctctgtga ggatgtaggt 1980

tccatcctgg gctttgctca ttaggttaag gatctggcat tgctgcagct cagacccagt 2040

gctgccctgg ttgtggctta ggccaaagct gcagctccaa ttcaatctct ggcctgggaa 2100

cctccatgtg ctacaaggtg cggccttaaa aggaaaaaaa aaaaattaaa tcaaggactc 2160

aagagtcttt cattatttgt gttgtggaag ctatatttgt tttaaagtct tagttgtgtt 2220

tagaaagcaa gatgttcttc aactcaaatt tgggagggaa cttgtttcat acatttttaa 2280

tggataagtg gcaaaatttt catgctgagg tgatctatag tgttgtaatg cagaatatag 2340

tcagatcttg aacattttag gaagttggtg agggccaatt gtgtatctgt gccatgctga 2400

taagaatgtc aagggatcac aagaattcgt gttatttgac agcagtcatc tttaaaaggc 2460

atttgagaaa gtccaatttc aaatgcattt cctttcttta aaagataaat tgaagaaaat 2520

aagtctttat ttcccaagta aattgaattg cctctcagtc tgttaaaaga aactcttacc 2580

ttgatgattg cgctcttaac ctggcaaaga ttgtctttaa aatctgagct ccatgtcttc 2640

tgctttattt ctggtgtgcc tttgactcca gattacagta aatggaggac tgagtatagg 2700

gctaaaaagt agagagaatg gatgcatatt atctgtggtc tccaatgtga tgaatgaagt 2760

aggcaaatac tcaaaggaaa gagaaagcat gctccaagaa ttatgggttc cagaaggcaa 2820

agtcccagaa ttgtctccag ggaaggacag ggaggtctag aatcggctaa gcccactgta 2880

ggcagaaaaa ccaagaggca tgaatggctt ccctttctca cttttcactc tctggcttac 2940

tcctatcatg aaggaaaata ttggaatcat attctccctc accgaaatgc tatttttcag 3000

cccacaggaa acccaggctg gttggagggg acattccctg ctctggtcgt gttgaagtac 3060

aacatggaga cacgtggggc accgtctgtg attctgactt ctctctggag gcggccagcg 3120

tgctgtgcag ggaactacag tgcggcaact gtggtttccc tcctgggggg agctcacttt 3180

ggagaaggaa gtggacagat ctgggctgaa gaattccagt gtgaggggca cgagtcccac 3240

ctttcactct gcccagtagc accccgccct gacgggacat gtagccacag cagggacgtc 3300

ggcgtagtct gctcaagtga gacccaggga atgtgttcac tttgttccca tgccatgaag 3360

agggtagggt taggtagtca cagacatctt tttaaagccc tgtctccttc caggatacac 3420

acaaatccgc ttggtgaatg gcaagacccc atgtgaagga agagtggagc tcaacattct 3480

tgggtcctgg gggtccctct gcaactctca ctgggacatg gaagatgccc atgttttatg 3540

ccagcagctt aaatgtggag ttgccctttc tatcccggga ggagcacctt ttgggaaagg 3600

aagtgagcag gtctggaggc acatgtttca ctgcactggg actgagaagc acatgggaga 3660

ttgttccgtc actgctctgg gcgcatcact ctgttcttca gggcaagtgg cctctgtaat 3720

ctgctcaggt aagagaataa gggcagccag tgatgagcca ctcatgacgg tgccttaaga 3780

gtgggtgtac ctaggagttc ccattgtggc tcagtggtaa caaactcgac tggtatccat 3840

gagggtatgg gtttgatccc tggccttgct caatgggtta aggatccagc attgctgtga 3900

gctgtggtat aggttgcaga ctctgctcag gtcccatgtt gctgtgattg tggtgtaggc 3960

tgactgctgc agcttcaatt tgacccctag cccgggaatt tccataggcc acacgtgcag 4020

cactaaggaa ggaaaaaaag aaaaaaaaaa aaaaagagtg ggtgtgccta tagtgaagaa 4080

cagatgtaaa agggaagtga aagggattcc cccattctga gggattgtga gaagtgtgcc 4140

agaatattaa cttcatttga cttgttacag ggaaagtaaa cttgactttc acggacctcc 4200

tagttacctg gtgcttacta tatgtcttct cagagtacct gattcattcc cagcctggtt 4260

gacccatccc cctatctcta tggctatgtt tatccagagc acatctatct aacactccag 4320

ctgatcttcc tgacacagct gtggcaaccc tggatccttt aaccaactgt gccaggctgg 4380

agatcaaacc taagcctctg cagcaaccca agctgctgca gtcagatttt taacccccctg 4440

tgccactgtg ggtatctccg atattttgta tcttctgtga ctgagtggtt tgctgtttgc 4500

agggaaccag agtcagacac tatccccgtg caattcatca tcctcggacc catcaagctc 4560

tattatttca gaagaaaatg gtgttgcctg cataggtgag aatcagtgac caacctatga 4620

aaatgatctc aatcctctga aatgcatttt attcatgttt tatttcctct ttgcagggag 4680

tggtcaactt cgcctggtcg atggaggtgg tcgttgtgct gggagagtag aggtctatca 4740

tgagggctcc tggggcacca tctgtgatga cagctgggac ctgaatgatg cccatgtggt 4800

gtgcaaacag ctgagctgtg gatgggccat taatgccact ggttctgctc attttgggga 4860

aggaacaggg cccatttggc tggatgagat aaactgtaat ggaaaagaat ctcatatttg 4920

gcaatgccac tcacatggtt gggggcggca caattgcagg cataaggagg atgcaggagt 4980

catctgctcg g 4991

<210> 107
<211> 4860
<212> ADN
<213> Sus scrofa

<400> 107
tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt 540

gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat 600

cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac 660

tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa 720

atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc 780

caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa 840

gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta 900

tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag 960

aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 1020

tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat 1080

actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag 1140

agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg 1200

aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctcctttgt 1260

cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt 1320

ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc 1380

tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact 1440

cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg 1500

atcctgcgtt gctgtggctg tggctgtggt gtaggccagc agcaaacagc tctgattaga 1560

cccctagcct ggaaacctcc atatgccaca ggtgcagccc taaaaagaca aaaaaagaga 1620

aaagacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gaacccccag 1680

aggtatttat ttgtttttgc ctttttttcac tgactgttct ttgtttgttt gtttgagact 1740

gatctagaag actagagatt acaagaaata tggatttggc tcactctaag aaactgcttt 1800

cattccaagg tttgggtcta tccaaaagtg gaatagaatc atatgaatac tagtttatga 1860

gtatttagtg agaggaattt caagctcaaa taatgattca gcaagattaa attaaggagg 1920

gaattttcct tgtggctgag tgggttaagg acccaatgtt gtctctgtga ggatgtaggt 1980

tccatcctgg gctttgctca ttaggttaag gatctggcat tgctgcagct cagacccagt 2040

gctgccctgg ttgtggctta ggccaaagct gcagctccaa ttcaatctct ggcctgggaa 2100

cctccatgtg ctacaaggtg cggccttaaa aggaaaaaaa aaaaattaaa tcaaggactc 2160

aagagtcttt cattatttgt gttgtggaag ctatatttgt tttaaagtct tagttgtgtt 2220

tagaaagcaa gatgttcttc aactcaaatt tgggagggaa cttgtttcat acatttttaa 2280

tggataagtg gcaaaatttt catgctgagg tgatctatag tgttgtaatg cagaatatag 2340

tcagatcttg aacattttag gaagttggtg agggccaatt gtgtatctgt gccatgctga 2400

taagaatgtc aagggatcac aagaattcgt gttatttgac agcagtcatc tttaaaaggc 2460

atttgagaaa gtccaatttc aaatgcattt cctttcttta aaagataaat tgaagaaaat 2520

aagtctttat ttcccaagta aattgaattg cctctcagtc tgttaaaaga aactcttacc 2580

ttgatgattg cgctcttaac ctggcaaaga ttgtctttaa aatctgagct ccatgtcttc 2640

tgctttattt ctggtgtgcc tttgactcca gattacagta aatggaggac tgagtatagg 2700

gctaaaaagt agagagaatg gatgcatatt atctgtggtc tccaatgtga tgaatgaagt 2760

aggcaaatac tcaaaggaaa gagaaagcat gctccaagaa ttatgggttc cagaaggcaa 2820

agtcccagaa ttgtctccag ggaaggacag ggaggtctag aatcggctaa gcccactgta 2880

ggcagaaaaa ccaagaggca tgaatggctt ccctttctca cttttcactc tctggcttac 2940

tcctatcatg aaggaaaata ttggaatcat attctccctc accgaaatgc tatttttcag 3000

cccacaggaa acccaggctg gttggagggt cctgggggga gctcactttg gagaaggaag 3060

tggacagatc tgggctgaag aattccagtg tgaggggcac gagtcccacc tttcactctg 3120

cccagtagca ccccgccctg acgggacatg tagccacagc agggacgtcg gcgtagtctg 3180

ctcaagtgag acccagggaa tgtgttcact ttgttcccat gccatgaaga gggtagggtt 3240

aggtagtcac agacatcttt ttaaagccct gtctccttcc aggatacaca caaatccgct 3300

tggtgaatgg caagacccca tgtgaaggaa gagtggagct caacattctt gggtcctggg 3360

ggtccctctg caactctcac tgggacatgg aagatgccca tgttttatgc cagcagctta 3420

aatgtggagt tgccctttct atcccgggag gagcaccttt tgggaaagga agtgagcagg 3480

tctggaggca catgtttcac tgcactggga ctgagaagca catgggagat tgttccgtca 3540

ctgctctggg cgcatcactc tgttcttcag ggcaagtggc ctctgtaatc tgctcaggta 3600

agagaataag ggcagccagt gatgagccac tcatgacggt gccttaagag tgggtgtacc 3660

taggagttcc cattgtggct cagtggtaac aaactcgact ggtatccatg agggtatggg 3720

tttgatccct ggccttgctc aatgggttaa ggatccagca ttgctgtgag ctgtggtata 3780

ggttgcagac tctgctcagg tcccatgttg ctgtgattgt ggtgtaggct gactgctgca 3840

gcttcaattt gacccctagc ccgggaattt ccataggcca cacgtgcagc actaaggaag 3900

gaaaaaaaga aaaaaaaaaa aaaagagtgg gtgtgcctat agtgaagaac agatgtaaaa 3960

gggaagtgaa agggattccc ccattctgag ggattgtgag aagtgtgcca gaatattaac 4020

ttcatttgac ttgttacagg gaaagtaaac ttgactttca cggacctcct agttacctgg 4080

tgcttactat atgtcttctc agagtacctg attcattccc agcctggttg acccatcccc 4140

ctatctctat ggctatgttt atccagagca catctatcta acactccagc tgatcttcct 4200

gacacagctg tggcaaccct ggatccttta accaactgtg ccaggctgga gatcaaacct 4260

aagcctctgc agcaacccaa gctgctgcag tcagattttt aacccccctgt gccactgtgg 4320

gtatctccga tattttgtat cttctgtgac tgagtggttt gctgtttgca gggaaccaga 4380

gtcagacact atccccgtgc aattcatcat cctcggaccc atcaagctct attatttcag 4440

aagaaaatgg tgttgcctgc ataggtgaga atcagtgacc aacctatgaa aatgatctca 4500

atcctctgaa atgcatttta ttcatgtttt atttcctctt tgcagggagt ggtcaacttc 4560

gcctggtcga tggaggtggt cgttgtgctg ggagagtaga ggtctatcat gagggctcct 4620

ggggcaccat ctgtgatgac agctgggacc tgaatgatgc ccatgtggtg tgcaaacagc 4680

tgagctgtgg atgggccatt aatgccactg gttctgctca ttttggggaa ggaacagggc 4740

ccatttggct ggatgagata aactgtaatg gaaaagaatc tcatatttgg caatgccact 4800

cacatggttg ggggcggcac aattgcaggc ataaggagga tgcaggagtc atctgctcgg 4860

<210> 108
<211> 4858
<212> ADN
<213> Sus scrofa

<400> 108
tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt 540

gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat 600

cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac 660

tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa 720

atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc 780

caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa 840

gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta 900

tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag 960

aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 1020

tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat 1080

actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag 1140

agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg 1200

aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctcctttgt 1260

cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt 1320

ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc 1380

tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact 1440

cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg 1500

atcctgcgtt gctgtggctg tggctgtggt gtaggccagc agcaaacagc tctgattaga 1560

cccctagcct ggaaacctcc atatgccaca ggtgcagccc taaaaagaca aaaaaagaga 1620

aaagacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gaacccccag 1680

aggtatttat ttgtttttgc ctttttttcac tgactgttct ttgtttgttt gtttgagact 1740

gatctagaag actagagatt acaagaaata tggatttggc tcactctaag aaactgcttt 1800

cattccaagg tttgggtcta tccaaaagtg gaatagaatc atatgaatac tagtttatga 1860

gtatttagtg agaggaattt caagctcaaa taatgattca gcaagattaa attaaggagg 1920

gaattttcct tgtggctgag tgggttaagg acccaatgtt gtctctgtga ggatgtaggt 1980

tccatcctgg gctttgctca ttaggttaag gatctggcat tgctgcagct cagacccagt 2040

gctgccctgg ttgtggctta ggccaaagct gcagctccaa ttcaatctct ggcctgggaa 2100

cctccatgtg ctacaaggtg cggccttaaa aggaaaaaaa aaaaattaaa tcaaggactc 2160

aagagtcttt cattatttgt gttgtggaag ctatatttgt tttaaagtct tagttgtgtt 2220

tagaaagcaa gatgttcttc aactcaaatt tgggagggaa cttgtttcat acatttttaa 2280

tggataagtg gcaaaatttt catgctgagg tgatctatag tgttgtaatg cagaatatag 2340

tcagatcttg aacattttag gaagttggtg agggccaatt gtgtatctgt gccatgctga 2400

taagaatgtc aagggatcac aagaattcgt gttatttgac agcagtcatc tttaaaaggc 2460

atttgagaaa gtccaatttc aaatgcattt cctttcttta aaagataaat tgaagaaaat 2520

aagtctttat ttcccaagta aattgaattg cctctcagtc tgttaaaaga aactcttacc 2580

ttgatgattg cgctcttaac ctggcaaaga ttgtctttaa aatctgagct ccatgtcttc 2640

tgctttattt ctggtgtgcc tttgactcca gattacagta aatggaggac tgagtatagg 2700

gctaaaaagt agagagaatg gatgcatatt atctgtggtc tccaatgtga tgaatgaagt 2760

aggcaaatac tcaaaggaaa gagaaagcat gctccaagaa ttatgggttc cagaaggcaa 2820

agtcccagaa ttgtctccag ggaaggacag ggaggtctag aatcggctaa gcccactgta 2880

ggcagaaaaa ccaagaggca tgaatggctt ccctttctca cttttcactc tctggcttac 2940

tcctatcatg aaggaaaata ttggaatcat attctccctc accgaaatgc tatttttcag 3000

cccacaggaa acccaggctg gttggagggc tgggggggagc tcactttgga gaaggaagtg 3060

gacagatctg ggctgaagaa ttccagtgtg aggggcacga gtcccacctt tcactctgcc 3120

cagtagcacc ccgccctgac gggacatgta gccacagcag ggacgtcggc gtagtctgct 3180

caagtgagac ccagggaatg tgttcacttt gttcccatgc catgaagagg gtagggttag 3240

gtagtcacag acatcttttt aaagccctgt ctccttccag gatacacaca aatccgcttg 3300

gtgaatggca agaccccatg tgaaggaaga gtggagctca acattcttgg gtcctggggg 3360

tccctctgca actctcactg ggacatggaa gatgcccatg ttttatgcca gcagcttaaa 3420

tgtggagttg ccctttctat cccgggagga gcacctttg ggaaaggaag tgagcaggtc 3480

tggaggcaca tgtttcactg cactgggact gagaagcaca tgggagattg ttccgtcact 3540

gctctgggcg catcactctg ttcttcaggg caagtggcct ctgtaatctg ctcaggtaag 3600

agaataaggg cagccagtga tgagccactc atgacggtgc cttaagagtg ggtgtaccta 3660

ggagttccca ttgtggctca gtggtaacaa actcgactgg tatccatgag ggtatgggtt 3720

tgatccctgg ccttgctcaa tgggttaagg atccagcatt gctgtgagct gtggtatagg 3780

ttgcagactc tgctcaggtc ccatgttgct gtgattgtgg tgtaggctga ctgctgcagc 3840

ttcaatttga cccctagccc gggaatttcc ataggccaca cgtgcagcac taaggaagga 3900

aaaaaagaaa aaaaaaaaaa aagagtgggt gtgcctatag tgaagaacag atgtaaaagg 3960

gaagtgaaag ggattccccc attctgaggg attgtgagaa gtgtgccaga atattaactt 4020

catttgactt gttacaggga aagtaaactt gactttcacg gacctcctag ttacctggtg 4080

cttactatat gtcttctcag agtacctgat tcattcccag cctggttgac ccatcccccc 4140

atctctatgg ctatgtttat ccagagcaca tctatctaac actccagctg atcttcctga 4200

cacagctgtg gcaaccctgg atcctttaac caactgtgcc aggctggaga tcaaacctaa 4260

gcctctgcag caacccaagc tgctgcagtc agatttttaa ccccctgtgc cactgtgggt 4320

atctccgata ttttgtatct tctgtgactg agtggtttgc tgtttgcagg gaaccagagt 4380

cagacactat ccccgtgcaa ttcatcatcc tcggacccat caagctctat tatttcagaa 4440

gaaaatggtg ttgcctgcat aggtgagaat cagtgaccaa cctatgaaaa tgatctcaat 4500

cctctgaaat gcattttatt catgttttat ttcctctttg cagggagtgg tcaacttcgc 4560

ctggtcgatg gaggtggtcg ttgtgctggg agagtagagg tctatcatga gggctcctgg 4620

ggcaccatct gtgatgacag ctgggacctg aatgatgccc atgtggtgtg caaacagctg 4680

agctgtggat gggccattaa tgccactggt tctgctcatt ttggggaagg aacagggccc 4740

atttggctgg atgagataaa ctgtaatgga aaagaatctc atatttggca atgccactca 4800

catggttggg ggcggcacaa ttgcaggcat aaggaggatg caggagtcat ctgctcgg 4858

<210> 109
<211> 3523
<212> ADN
<213> Sus scrofa

<400> 109
tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt 540

gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat 600

cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac 660

tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa 720

atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc 780

caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa 840

gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta 900

tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag 960

aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 1020

tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat 1080

actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag 1140

agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg 1200

aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctcctttgt 1260

cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt 1320

ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc 1380

tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact 1440

cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg 1500

atcctgcgtt gctgtggctg tggctgtggt gtaggccagc agcaaacagc tctgattaga 1560

cccctagcct ggaaacctcc atatgccaca ggtgcagccc taaaaagaca aaaaaagaga 1620

aaagacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gaacccccag 1680

aggtatttat ttgtttttgc ctttttttcac tgactgttct ttgtttgttt gtttgagact 1740

gatctagaag actagagatt acaagaaata tggatttggc tcactctaag aaactgcttt 1800

cattccaagg tttgggtcta tccaaaagtg gaatagaatc atatgaatac tagtttatga 1860

gtatttagtg agaggaattt caagctcaaa taatgattca gcaagattaa attaaggagg 1920

gaattttcct tgtggctgag tgggttaagg acccaatgtt gtctctgtga ggatgtaggt 1980

tccatcctgg gctttgctca ttaggttaag gatctggcat tgctgcagct cagacccagt 2040

gctgccctgg ttgtggctta ggccaaagct gcagctccaa ttcaatctct ggcctgggaa 2100

cctccatgtg ctacaaggtg cggccttaaa aggaaaaaaa aaaaattaaa tcaaggactc 2160

aagagtcttt cattatttgt gttgtggaag ctatatttgt tttaaagtct tagttgtgtt 2220

tagaaagcaa gatgttcttc aactcaaatt tgggagggaa cttgtttcat acatttttaa 2280

tggataagtg gcaaaatttt catgctgagg tgatctatag tgttgtaatg cagaatatag 2340

tcagatcttg aacattttag gaagttggtg agggccaatt gtgtatctgt gccatgctga 2400

taagaatgtc aagggatcac aagaattcgt gagctgtggt ataggttgca gactctgctc 2460

aggtcccatg ttgctgtgat tgtggtgtag gctgactgct gcagcttcaa tttgacccct 2520

agcccgggaa tttccatagg ccacacgtgc agcactaagg aaggaaaaaa agaaaaaaaa 2580

aaaaaaagag tgggtgtgcc tatagtgaag aacagatgta aaagggaagt gaaagggatt 2640

cccccattct gagggattgt gagaagtgtg ccagaatatt aacttcattt gacttgttac 2700

agggaaagta aacttgactt tcacggacct cctagttacc tggtgcttac tatatgtctt 2760

ctcagagtac ctgattcatt cccagcctgg ttgacccatc cccctatctc tatggctatg 2820

tttatccaga gcacatctat ctaacactcc agctgatctt cctgacacag ctgtggcaac 2880

cctggatcct ttaaccaact gtgccaggct ggagatcaaa cctaagcctc tgcagcaacc 2940

caagctgctg cagtcagatt tttaaccccc tgtgccactg tgggtatctc cgatattttg 3000

tatcttctgt gactgagtgg tttgctgttt gcagggaacc agagtcagac actatccccg 3060

tgcaattcat catcctcgga cccatcaagc tctattattt cagaagaaaa tggtgttgcc 3120

tgcataggtg agaatcagtg accaacctat gaaaatgatc tcaatcctct gaaatgcatt 3180

ttattcatgt tttatttcct ctttgcaggg agtggtcaac ttcgcctggt cgatggaggt 3240

ggtcgttgtg ctgggagagt agaggtctat catgagggct cctggggcac catctgtgat 3300

gacagctggg acctgaatga tgcccatgtg gtgtgcaaac agctgagctg tggatgggcc 3360

attaatgcca ctggttctgc tcattttggg gaaggaacag ggcccatttg gctggatgag 3420

ataaactgta atggaaaaga atctcatatt tggcaatgcc actcacatgg ttgggggcgg 3480

cacaattgca ggcataagga ggatgcagga gtcatctgct cgg 3523

<210> 110
<211> 3603
<212> ADN
<213> Sus scrofa

<400> 110
tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgttgt ggagaattcc acaagaattc gtgttatttg acagcagtca tctttaaaag 540

gcatttgaga aagtccaatt tcaaatgcat ttcctttctt taaaagataa attgaagaaa 600

ataagtcttt atttcccaag taaattgaat tgcctctcag tctgttaaaa gaaactctta 660

tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 720

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 780

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 840

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 900

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 960

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 1020

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 1080

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 1140

gagatgttgt ggagaattcc acaagaattc gtgttatttg acagcagtca tctttaaaag 1200

gcatttgaga aagtccaatt tcaaatgcat ttcctttctt taaaagataa attgaagaaa 1260

ataagtcttt atttcccaag taaattgaat tgcctctcag tctgttaaaa gaaactctta 1320

ccttgatgat tgcgctctta acctggcaaa gattgtcttt aaaatctgag ctccatgtct 1380

tctgctttat ttctggtgtg cctttgactc cagattacag taaatggagg actgagtata 1440

gggctaaaaa gtagagagaa tggatgcata ttatctgtgg tctccaatgt gatgaatgaa 1500

gtaggcaaat actcaaagga aagagaaagc atgctccaag aattatgggt tccagaaggc 1560

aaagtcccag aattgtctcc agggaaggac agggaggtct agaatcggct aagcccactg 1620

taggcagaaa aaccaagagg catgaatggc ttccctttct cacttttcac tctctggctt 1680

actcctatca tgaaggaaaa tattggaatc atattctccc tcaccgaaat gctatttttc 1740

agcccacagg aaacccaggc tggttggagg ggacattccc tgctctcact ttggagaagg 1800

aagtggacag atctgggctg aagaattcca gtgtgagggg cacgagtccc acctttcact 1860

ctgcccagta gcaccccgcc ctgacgggac atgtagccac agcagggacg tcggcgtagt 1920

ctgctcaagt gagacccagg gaatgtgttc actttgttcc catgccatga agagggtagg 1980

gttaggtagt cacagacatc tttttaaagc cctgtctcct tccaggatac acacaaatcc 2040

gcttggtgaa tggcaagacc ccatgtgaag gaagagtgga gctcaacatt cttgggtcct 2100

gggggtccct ctgcaactct cactgggaca tggaagatgc ccatgtttta tgccagcagc 2160

ttaaatgtgg agttgcccct tctatcccgg gaggagcacc ttttgggaaa ggaagtgagc 2220

aggtctggag gcacatgttt cactgcactg ggactgagaa gcacatggga gattgttccg 2280

tcactgctct gggcgcatca ctctgttctt cagggcaagt ggcctctgta atctgctcag 2340

gtaagagaat aagggcagcc agtgatgagc cactcatgac ggtgccttaa gagtgggtgt 2400

acctaggagt tcccattgtg gctcagtggt aacaaactcg actggtatcc atgagggtat 2460

gggtttgatc cctggccttg ctcaatgggt taaggatcca gcattgctgt gagctgtggt 2520

ataggttgca gactctgctc aggtcccatg ttgctgtgat tgtggtgtag gctgactgct 2580

gcagcttcaa tttgacccct agcccgggaa tttccatagg ccacacgtgc agcactaagg 2640

aaggaaaaaa agaaaaaaaa aaaaaaagag tgggtgtgcc tatagtgaag aacagatgta 2700

aaagggaagt gaaagggatt cccccattct gagggattgt gagaagtgtg ccagaatatt 2760

aacttcattt gacttgttac agggaaagta aacttgactt tcacggacct cctagttacc 2820

tggtgcttac tatatgtctt ctcagagtac ctgattcatt cccagcctgg ttgacccatc 2880

cccctatctc tatggctatg tttatccaga gcacatctat ctaacactcc agctgatctt 2940

cctgacacag ctgtggcaac cctggatcct ttaaccaact gtgccaggct ggagatcaaa 3000

cctaagcctc tgcagcaacc caagctgctg cagtcagatt tttaaccccc tgtgccactg 3060

tgggtatctc cgatattttg tatcttctgt gactgagtgg tttgctgttt gcagggaacc 3120

agagtcagac actatccccg tgcaattcat catcctcgga cccatcaagc tctattattt 3180

cagaagaaaa tggtgttgcc tgcataggtg agaatcagtg accaacctat gaaaatgatc 3240

tcaatcctct gaaatgcatt ttattcatgt tttatttcct ctttgcaggg agtggtcaac 3300

ttcgcctggt cgatggaggt ggtcgttgtg ctgggagagt agaggtctat catgagggct 3360

cctggggcac catctgtgat gacagctggg acctgaatga tgcccatgtg gtgtgcaaac 3420

agctgagctg tggatgggcc attaatgcca ctggttctgc tcattttggg gaaggaacag 3480

ggcccatttg gctggatgag ataaactgta atggaaaaga atctcatatt tggcaatgcc 3540

actcacatgg ttgggggcgg cacaattgca ggcataagga ggatgcagga gtcatctgct 3600

cgg 3603

<210> 111
<211> 4962
<212> ADN
<213> Sus scrofa

<400> 111

tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt 540

gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat 600

cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac 660

tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa 720

atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc 780

caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa 840

gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta 900

tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag 960

aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 1020

tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat 1080

actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag 1140

agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg 1200

aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctcctttgt 1260

cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt 1320

ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc 1380

tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact 1440

cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg 1500

atcctgcgtt gctgtggctg tggctgtggt gtaggccagc agcaaacagc tctgattaga 1560

cccctagcct ggaaacctcc atatgccaca ggtgcagccc taaaaagaca aaaaaagaga 1620

aaagacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gaacccccag 1680

aggtatttat ttgtttttgc ctttttttcac tgactgttct ttgtttgttt gtttgagact 1740

gatctagaag actagagatt acaagaaata tggatttggc tcactctaag aaactgcttt 1800

cattccaagg tttgggtcta tccaaaagtg gaatagaatc atatgaatac tagtttatga 1860

gtatttagtg agaggaattt caagctcaaa taatgattca gcaagattaa attaaggagg 1920

gaattttcct tgtggctgag tgggttaagg acccaatgtt gtctctgtga ggatgtaggt 1980

tccatcctgg gctttgctca ttaggttaag gatctggcat tgctgcagct cagacccagt 2040

gctgccctgg ttgtggctta ggccaaagct gcagctccaa ttcaatctct ggcctgggaa 2100

cctccatgtg ctacaaggtg cggccttaaa aggaaaaaaa aaaaattaaa tcaaggactc 2160

aagagtcttt cattatttgt gttgtggaag ctatatttgt tttaaagtct tagttgtgtt 2220

tagaaagcaa gatgttcttc aactcaaatt tgggagggaa cttgtttcat acatttttaa 2280

tggataagtg gcaaaatttt catgctgagg tgatctatag tgttgtaatg cagaatatag 2340

tcagatcttg aacattttag gaagttggtg agggccaatt gtgtatctgt gccatgctga 2400

taagaatgtc aagggatcac aagaattcgt gttatttgac agcagtcatc tttaaaaggc 2460

atttgagaaa gtccaatttc aaatgcattt cctttcttta aaagataaat tgaagaaaat 2520

aagtctttat ttcccaagta aattgaattg cctctcagtc tgttaaaaga aactcttacc 2580

ttgatgattg cgctcttaac ctggcaaaga ttgtctttaa aatctgagct ccatgtcttc 2640

tgctttattt ctggtgtgcc tttgactcca gattacagta aatggaggac tgagtatagg 2700

gctaaaaagt agagagaatg gatgcatatt atctgtggtc tccaatgtga tgaatgaagt 2760

aggcaaatac tcaaaggaaa gagaaagcat gctccaagaa ttatgggttc cagaaggcaa 2820

agtcccagaa ttgtctccag ggaaggacag ggaggtctag aatcggctaa gcccactgta 2880

ggcagaaaaa ccaagaggca tgaatggctt ccctttctca cttttcactc tctggcttac 2940

tcctatcatg aaggaaaata ttggaatcat attctccctc accgaaatgc tatttttcag 3000

cccacaggaa acccaggctg gttggagggg acattccctg ctctggtcgt gttgaagtac 3060

aacatggaga cacgtggggc accgtctgtg attctgactt ctctctggag gcggccagcg 3120

tgctgtgcag ggaactacag tgcgtcactt tggagaagga agtggacaga tctgggctga 3180

agaattccag tgtgaggggc acgagtccca cctttcactc tgcccagtag caccccgccc 3240

tgacgggaca tgtagccaca gcagggacgt cggcgtagtc tgctcaagtg agacccaggg 3300

aatgtgttca ctttgttccc atgccatgaa gagggtaggg ttaggtagtc acagacatct 3360

ttttaaagcc ctgtctcctt ccaggataca cacaaatccg cttggtgaat ggcaagaccc 3420

catgtgaagg aagagtggag ctcaacattc ttgggtcctg ggggtccctc tgcaactctc 3480

actgggacat ggaagatgcc catgttttat gccagcagct taaatgtgga gttgcccttt 3540

ctatcccggg aggagcacct tttgggaaag gaagtgagca ggtctggagg cacatgtttc 3600

actgcactgg gactgagaag cacatgggag attgttccgt cactgctctg ggcgcatcac 3660

tctgttcttc agggcaagtg gcctctgtaa tctgctcagg taagagaata agggcagcca 3720

gtgatgagcc actcatgacg gtgccttaag agtgggtgta cctaggagtt cccattgtgg 3780

ctcagtggta acaaactcga ctggtatcca tgagggtatg ggtttgatcc ctggccttgc 3840

tcaatgggtt aaggatccag cattgctgtg agctgtggta taggttgcag actctgctca 3900

ggtcccatgt tgctgtgatt gtggtgtagg ctgactgctg cagcttcaat ttgaccccta 3960

gcccgggaat ttccataggc cacacgtgca gcactaagga aggaaaaaaa gaaaaaaaaa 4020

aaaaaagagt gggtgtgcct atagtgaaga acagatgtaa aagggaagtg aaagggattc 4080

ccccattctg agggattgtg agaagtgtgc cagaatatta acttcatttg acttgttaca 4140

gggaaagtaa acttgacttt cacggacctc ctagttacct ggtgcttact atatgtcttc 4200

tcagagtacc tgattcattc ccagcctggt tgacccatcc ccctatctct atggctatgt 4260

ttatccagag cacatctatc taacactcca gctgatcttc ctgacacagc tgtggcaacc 4320

ctggatcctt taaccaactg tgccaggctg gagatcaaac ctaagcctct gcagcaaccc 4380

aagctgctgc agtcagattt ttaacccccc gtgccactgt gggtatctcc gatattttgt 4440

atcttctgtg actgagtggt ttgctgtttg cagggaacca gagtcagaca ctatccccgt 4500

gcaattcatc atcctcggac ccatcaagct ctattatttc agaagaaaat ggtgttgcct 4560

gcataggtga gaatcagtga ccaacctatg aaaatgatct caatcctctg aaatgcattt 4620

tattcatgtt ttatttcctc tttgcaggga gtggtcaact tcgcctggtc gatggaggtg 4680

gtcgttgtgc tgggagagta gaggtctatc atgagggctc ctggggcacc atctgtgatg 4740

acagctggga cctgaatgat gcccatgtgg tgtgcaaaca gctgagctgt ggatgggcca 4800

ttaatgccac tggttctgct cattttgggg aaggaacagg gcccatttgg ctggatgaga 4860

taaactgtaa tggaaaagaa tctcatattt ggcaatgcca ctcacatggt tgggggcggc 4920

acaattgcag gcataaggag gatgcaggag tcatctgctc gg 4962

<210> 112
<211> 3603
<212> ADN
<213> Sus scrofa

<400> 112
tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt 540

gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat 600

cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac 660

tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa 720

atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc 780

caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa 840

gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta 900

tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag 960

aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 1020

tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat 1080

actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag 1140

agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg 1200

aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctcctttgt 1260

cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt 1320

ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc 1380

tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact 1440

cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg 1500

atcctgcgtt gctgtggctg tggctgtggt gtaggccagc agcaaacagc tctgattaga 1560

cccctagcct ggaaacctcc atatgccaca ggtgcagccc taaaaagaca aaaaaagaga 1620

aaagacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gaacccccag 1680

aggtatttat ttgtttttgc ctttttttcac tgactgttct ttgtttgttt gtttgagact 1740

gatctagaag actagagatt acaagaaata tggatttggc tcactctaag aaactgcttt 1800

cattccaagg tttgggtcta tccaaaagtg gaatagaatc atatgaatac tagtttatga 1860

gtatttagtg agaggaattt caagctcaaa taatgattca gcaagattaa attaaggagg 1920

gaattttcct tgtggctgag tgggttaagg acccaatgtt gtctctgtga ggatgtaggt 1980

tccatcctgg gctttgctca ttaggttaag gatctggcat tgctgcagct cagacccagt 2040

gctgccctgg ttgtggctta ggccaaagct gcagctccaa ttcaatctct ggcctgggaa 2100

cctccatgtg ctacaaggtg cggccttaaa aggaaaaaaa aaaaattaaa tcaaggactc 2160

aagagtcttt cattatttgt gttgtggaag ctatatttgt tttaaagtct tagttgtgtt 2220

tagaaagcaa gatgttcttc aactcaaatt tgggagggaa cttgtttcat acatttttaa 2280

tggataagtg gcaaaatttt catgctgagg tgatctatag tgttgtaatg cagaatatag 2340

tcagatcttg aacattttag gaagttggtg agggccaatt gtgtatctgt gccatgctga 2400

taagaatgtc aagggatcac aagaattcgt gttatttgac agcagtcatc tttaaaaggc 2460

atttgagaaa gtccaatttc aaatgcattt cctttcttta aaagataaat tgaagaaaat 2520

aagtctttat ttcccaagta aattgaattg cctctcagtc tgttaaaaga aactcttacc 2580

ttgatgattg cgctcttaac ctggcaaaga ttgtctttaa aatctgagct ccatgtcttc 2640

tgctttattt ctggtgtgcc tttgactcca gattacagta aatggaggac tgagtatagg 2700

gctaaaaagt agagagaatg gatgcatatt atctgtggtc tccaatgtga tgaatgaagt 2760

aggcaaatac tcaaaggaaa gagaaagcat gctccaagaa ttatgggttc cagaaggcaa 2820

agtcccagaa ttgtctccag ggaaggacag ggaggtctag aatcggctaa gcccactgta 2880

ggcagaaaaa ccaagaggca tgaatggctt ccctttctca cttttcactc tctggcttac 2940

tcctatcatg aaggaaaata ttggaatcat attctccctc accgaaatgc tatttttcag 3000

cccacaggaa acccaggctg gttggagggg acattccctg ctctggtcgt gttgaagtac 3060

aacatggaga cacgtggggc accgtctgtg attctgactt ctctctggag gcggccagcg 3120

tgctgtgcag ggaactacag tgcgattcat catcctcgga cccatcaagc tctattattt 3180

cagaagaaaa tggtgttgcc tgcataggtg agaatcagtg accaacctat gaaaatgatc 3240

tcaatcctct gaaatgcatt ttattcatgt tttatttcct ctttgcaggg agtggtcaac 3300

ttcgcctggt cgatggaggt ggtcgttgtg ctgggagagt agaggtctat catgagggct 3360

cctggggcac catctgtgat gacagctggg acctgaatga tgcccatgtg gtgtgcaaac 3420

agctgagctg tggatgggcc attaatgcca ctggttctgc tcattttggg gaaggaacag 3480

ggcccatttg gctggatgag ataaactgta atggaaaaga atctcatatt tggcaatgcc 3540

actcacatgg ttgggggcgg cacaattgca ggcataagga ggatgcagga gtcatctgct 3600

cgg 3603

<210> 113
<211> 3619
<212> ADN
<213> Sus scrofa

<400> 113
tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt 540

gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat 600

cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac 660

tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa 720

atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc 780

caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa 840

gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta 900

tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag 960

aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 1020

tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat 1080

actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag 1140

agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg 1200

aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctcctttgt 1260

cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt 1320

ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc 1380

tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact 1440

cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg 1500

atcctgcgtt gctgtggctg tggctgtggt gtaggccagc agcaaacagc tctgattaga 1560

cccctagcct ggaaacctcc atatgccaca ggtgcagccc taaaaagaca aaaaaagaga 1620

aaagacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gaacccccag 1680

aggtatttat ttgtttttgc ctttttttcac tgactgttct ttgtttgttt gtttgagact 1740

gatctagaag actagagatt acaagaaata tggatttggc tcactctaag aaactgcttt 1800

cattccaagg tttgggtcta tccaaaagtg gaatagaatc atatgaatac tagtttatga 1860

gtatttagtg agaggaattt caagctcaaa taatgattca gcaagattaa attaaggagg 1920

gaattttcct tgtggctgag tgggttaagg acccaatgtt gtctctgtga ggatgtaggt 1980

tccatcctgg gctttgctca ttaggttaag gatctggcat tgctgcagct cagacccagt 2040

gctgccctgg ttgtggctta ggccaaagct gcagctccaa ttcaatctct ggcctgggaa 2100

cctccatgtg ctacaaggtg cggccttaaa aggaaaaaaa aaaaattaaa tcaaggactc 2160

aagagtcttt cattatttgt gttgtggaag ctatatttgt tttaaagtct tagttgtgtt 2220

tagaaagcaa gatgttcttc aactcaaatt tgggagggaa cttgtttcat acatttttaa 2280

tggataagtg gcaaaatttt catgctgagg tgatctatag tgttgtaatg cagaatatag 2340

tcagatcttg aacattttag gaagttggtg agggccaatt gtgtatctgt gccatgctga 2400

taagaatgtc aagggatcac aagaattcgt gttatttgac agcagtcatc tttaaaaggc 2460

atttgagaaa gtccaatttc aaatgcattt cctttcttta aaagataaat tgaagaaaat 2520

aagtctttat ttcccaagta aattgaattg cctctcagtc tgttaaaaga aactcttacc 2580

ttgatgattg cgctcttaac ctggcaaaga ttgtctttaa aatctgagct ccatgtcttc 2640

tgctttattt ctggtgtgcc tttgactcca gattacagta aatggaggac tgagtatagg 2700

gctaaaaagt agagagaatg gatgcatatt atctgtggtc tccaatgtga tgaatgaagt 2760

aggcaaatac tcaaaggaaa gagaaagcat gctccaagaa ttatgggttc cagaaggcaa 2820

agtcccagaa ttgtctccag ggaaggacag ggaggtctag aatcggctaa gcccactgta 2880

ggcagaaaaa ccaagaggca tgaatggctt ccctttctca cttttcactc tctggcttac 2940

tcctatcatg aaggaaaata ttggaatcat attctccctc accgaaatgc tatttttcag 3000

cccacaggaa acccaggctg gttggagggg acattccctg ctctggtcgt gttgaagtac 3060

aacatggaga cacgtggggc accgtctgtg attctgactt ctctctggag gcagccagcg 3120

tgctttgcag ggaaccagag tcagacacta tccccgtgca attcatcatc ctcggaccca 3180

tcaagctcta ttatttcaga agaaaatggt gttgcctgca taggtgagaa tcagtgacca 3240

acctatgaaa atgatctcaa tcctctgaaa tgcattttat tcatgtttta tttcctcttt 3300

gcagggagtg gtcaacttcg cctggtcgat ggaggtggtc gttgtgctgg gagagtagag 3360

gtctatcatg agggctcctg gggcaccatc tgtgatgaca gctgggacct gaatgatgcc 3420

catgtggtgt gcaaacagct gagctgtgga tgggccatta atgccactgg ttctgctcat 3480

tttggggaag gaacagggcc catttggctg gatgagataa actgtaatgg aaaagaatct 3540

catatttggc aatgccactc acatggttgg gggcggcaca attgcaggca taaggaggat 3600

gcaggagtca tctgctcgg 3619

<210> 114
<211> 3270
<212> ADN
<213> Sus scrofa

<400> 114
tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt 540

gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat 600

cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac 660

tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa 720

atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc 780

caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa 840

gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta 900

tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag 960

aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 1020

tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat 1080

actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag 1140

agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg 1200

aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctcctttgt 1260

cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt 1320

ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc 1380

tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact 1440

cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg 1500

atcctgcgtt gctgtggctg tggctgtggt gtaggccagc agcaaacagc tctgattaga 1560

cccctagcct ggaaacctcc atatgccaca ggtgcagccc taaaaagaca aaaaaagaga 1620

aaagacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gaacccccag 1680

aggtatttat ttgtttttgc ctttttttcac tgactgttct ttgtttgttt gtttgagact 1740

gatctagaag actagagatt acaagaaata tggatttggc tcactctaag aaactgcttt 1800

cattccaagg tttgggtcta tccaaaagtg gaatagaatc atatgaatac tagtttatga 1860

gtatttagtg agaggaattt caagctcaaa taatgattca gcaagattaa attaaggagg 1920

gaattttcct tgtggctgag tgggttaagg acccaatgtt gtctctgtga ggatgtaggt 1980

tccatcctgg gctttgctca ttaggttaag gatctggcat tgctgcagct cagacccagt 2040

gctgccctgg ttgtggctta ggccaaagct gcagctccaa ttcaatctct ggcctgggaa 2100

cctccatgtg ctacaaggtg cggccttaaa aggaaaaaaa aaaaattaaa tcaaggactc 2160

aagagtcttt cattatttgt gttgtggaag ctatatttgt tttaaagtct tagttgtgtt 2220

tagaaagcaa gatgttcttc aactcaaatt tgggagggaa cttgtttcat acatttttaa 2280

tggataagtg gcaaaatttt catgctgagg tgatctatag tgttgtaatg cagaatatag 2340

tcagatcttg aacattttag gaagttggtg agggccaatt gtgtatctgt gccatgctga 2400

taagaatgtc aagggatcac aagaattcgt gttatttgac ttgttacagg gaaagtaaac 2460

ttgactttca cggacctcct agttacctgg tgcttactat atgtcttctc agagtacctg 2520

attcattccc agcctggttg acccatcccc ctatctctat ggctatgttt atccagagca 2580

catctatcta acactccagc tgatcttcct gacacagctg tggcaaccct ggatccttta 2640

accaactgtg ccaggctgga gatcaaacct aagcctctgc agcaacccaa gctgctgcag 2700

tcagattttt aacccccctgt gccactgtgg gtatctccga tattttgtat cttctgtgac 2760

tgagtggttt gctgtttgca gggaaccaga gtcagacact atccccgtgc aattcatcat 2820

cctcggaccc atcaagctct attatttcag aagaaaatgg tgttgcctgc ataggtgaga 2880

atcagtgacc aacctatgaa aatgatctca atcctctgaa atgcatttta ttcatgtttt 2940

atttcctctt tgcagggagt ggtcaacttc gcctggtcga tggaggtggt cgttgtgctg 3000

ggagagtaga ggtctatcat gagggctcct ggggcaccat ctgtgatgac agctgggacc 3060

tgaatgatgc ccatgtggtg tgcaaacagc tgagctgtgg atgggccatt aatgccactg 3120

gttctgctca ttttggggaa ggaacagggc ccatttggct ggatgagata aactgtaatg 3180

gaaaagaatc tcatatttgg caatgccact cacatggttg ggggcggcac aattgcaggc 3240

ataaggagga tgcaggagtc atctgctcgg 3270

<210> 115
<211> 7
<212> ADN
<213> Sus scrofa

<400> 115
tactact 7

<210> 116
<211> 12
<212> ADN
<213> Sus scrofa

<400> 116
tgtggagaat tc 12

<210> 117
<211> 11
<212> ADN
<213> Sus scrofa

<400> 117
agccagcgtg c 11

<210> 118
<211> 8532
<212> ADN

<213> Sus scrofa

<400> 118

tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt 540

gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat 600

cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac 660

tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa 720

atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc 780

caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa 840

gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta 900

tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag 960

aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 1020

tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat 1080

actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag 1140

agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg 1200

aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctcctttgt 1260

cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt 1320

ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc 1380

tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact 1440

cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg 1500

atcctgcgtt gctgtggctg tggctgtggt gtaggccagc agcaaacagc tctgattaga 1560

cccctagcct ggaaacctcc atatgccaca ggtgcagccc taaaaagaca aaaaaagaga 1620

aaagacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gaacccccag 1680

aggtatttat ttgtttttgc ctttttttcac tgactgttct ttgtttgttt gtttgagact 1740

gatctagaag actagagatt acaagaaata tggatttggc tcactctaag aaactgcttt 1800

cattccaagg tttgggtcta tccaaaagtg gaatagaatc atatgaatac tagtttatga 1860

gtatttagtg agaggaattt caagctcaaa taatgattca gcaagattaa attaaggagg 1920

gaattttcct tgtggctgag tgggttaagg acccaatgtt gtctctgtga ggatgtaggt 1980

tccatcctgg gctttgctca ttaggttaag gatctggcat tgctgcagct cagacccagt 2040

gctgccctgg ttgtggctta ggccaaagct gcagctccaa ttcaatctct ggcctgggaa 2100

cctccatgtg ctacaaggtg cggccttaaa aggaaaaaaa aaaaattaaa tcaaggactc 2160

aagagtcttt cattatttgt gttgtggaag ctatatttgt tttaaagtct tagttgtgtt 2220

tagaaagcaa gatgttcttc aactcaaatt tgggagggaa cttgtttcat acatttttaa 2280

tggataagtg gcaaaatttt catgctgagg tgatctatag tgttgtaatg cagaatatag 2340

tcagatcttg aacattttag gaagttggtg agggccaatt gtgtatctgt gccatgctga 2400

taagaatgtc aagggatcac aagaattcgt gttatttgac agcagtcatc tttaaaaggc 2460

atttgagaaa gtccaatttc aaatgcattt cctttcttta aaagataaat tgaagaaaat 2520

aagtctttat ttcccaagta aattgaattg cctctcagtc tgttaaaaga aactcttacc 2580

ttgatgattg cgctcttaac ctggcaaaga ttgtctttaa aatctgagct ccatgtcttc 2640

tgctttattt ctggtgtgcc tttgactcca gattacagta aatggaggac tgagtatagg 2700

gctaaaaagt agagagaatg gatgcatatt atctgtggtc tccaatgtga tgaatgaagt 2760

aggcaaatac tcaaaggaaa gagaaagcat gctccaagaa ttatgggttc cagaaggcaa 2820

agtcccagaa ttgtctccag ggaaggacag ggaggtctag aatcggctaa gcccactgta 2880

ggcagaaaaa ccaagaggca tgaatggctt ccctttctca cttttcactc tctggcttac 2940

tcctatcatg aaggaaaata ttggaatcat attctccctc accgaaatgc tatttttcag 3000

cccacaggaa acccaggctg gttggagctg acatgccctg ctctggtcgt gttgaagtaa 3060

aacatgcaga cacgtggggc tccgtctgtg attctgactt ctctctgcat gcggccaacg 3120

tgctgtgcag ggaactaaat tgcggcgatg cgatttccct ctcggtggga gatcactttg 3180

gaaaaggaaa tggactgacc tgggctgaaa aattccagtg tgaggggagc gagacccacc 3240

ttgcactctg cccaatagta caacaccctg aagacacatg tatccacagc agggaagtcg 3300

gcgtagtctg ctcaagtaag agtttactga aaataacact cttaaaatct tgttatgttt 3360

ttattcataa tgtgaatgag tagtagtgga aaataactac cagtttccta agcttataac 3420

ttcgtatagc atacattata cgaagttata agcctgcagg aattctaccg ggtaggggag 3480

gcgcttttcc caaggcagtc tggagcatgc gctttagcag ccccgctggg cacttggcgc 3540

tacacaagtg gcctctggcc tcgcacacat tccacatcca ccggtaggcg ccaaccggct 3600

ccgttctttg gtggcccctt cgcgccacct tctactcctc ccctagtcag gaagttcccc 3660

cccgccccgc agctcgcgtc gtgcaggacg tgacaaatgg aagtagcacg tctcactagt 3720

ctcgtgcaga tggacagcac cgctgagcaa tggaagcggg taggcctttg gggcagcggc 3780

caatagcagc tttgctcctt cgctttctgg gctcagaggc tgggaagggg tgggtccggg 3840

ggcgggctca ggggcgggct caggggcggg gcgggcgccc gaaggtcctc cggaggcccg 3900

gcattctgca cgcttcaaaa gcgcacgtct gccgcgctgt tctcctcttc ctcatctccg 3960

ggcctttcga cctgcagcca ccatggccat gattgaacag gatggcctgc atgcaggttc 4020

tccagctgcc tgggtggaga gactgtttgg ctatgactgg gcacagcaga ccattggttg 4080

ctctgatgca gcagtgttca gactgtcagc ccagggcagg ccagtcctgt ttgtgaagac 4140

agacctcagt ggggctctca atgagctgca ggatgaggct gccagactct cctggctggc 4200

aacaactggg gtcccctgtg cagctgtcct ggatgtggtc acagaagctg gaagggactg 4260

gctcctgctg ggtgaggtgc ctgggcagga cctcctgtcc tctcacctgg ctccagctga 4320

gaaagtgtca atcatggctg atgccatgag aagactccac accctggacc cagccacctg 4380

ccccctttgac caccaggcca agcacaggat tgagagggcc agaaccagga tggaggctgg 4440

cctggtggac caggatgacc tggatgaaga acaccagggc ctggcccctg ctgaactgtt 4500

tgccaggctc aaggcatcca tgccagatgg tgaggacctg gtggtgactc atggggatgc 4560

ctgcctgccc aacatcatgg tggaaaatgg aaggttctct ggcttcattg actgtggcag 4620

gctgggagtg gctgacaggt accaggacat tgccctggca accagggaca ttgcagaaga 4680

gctgggggga gaatgggcag acaggttcct ggtgctctat ggcattgcag cccctgactc 4740

ccagagaatt gccttctaca gactgctgga tgagttcttc taaatcgatt ataatcagcc 4800

ataccacatt tgtagaggtt ttacttgctt taaaaaacct cccacacctc cccctgaacc 4860

tgaaacataa aatgaatgca atgttgttgt taaacttgtt tattgcagct tataatggtt 4920

acaaataaag caatagcatc acaaatttca caaataaagc atttttttca ctgcattcta 4980

gttgtggttt gtccaaactc atcaatgtgg aaaataacta ccagtttcct aagcttataa 5040

cttcgtatag catacattat acgaagttat aagctagaca aaagtatctt accccaatgg 5100

tagccctgta cccaataaaa gtaggtgttc agtttcatat cctatgaaat accctcttga 5160

tacttttact ttgcatgagg atttagaaga aaaaagtttt actataatcc ttaacttagg 5220

aaattctttt gaattggaaa tgaaacacaa attgcttttc attgatatgc catatgatta 5280

tatgaataaa acatgaaatc ttcatattgg attctagtat atacccaagt aaatattttt 5340

tccctagaag agtgccaagt gtgttaaaac cttttggttt aataaagcag aaaaaaataa 5400

actctaaaaa tcataattaa aaatgaaatg cttttattta tagcaattaa ctacaacatg 5460

tttagactta catactatta aatataatat atttaagatc ccctcatgat aaatatgttc 5520

attattttgt aggctgttga tgcactaata tgtatgtaga ttactttgtg aattgccctt 5580

aataaaattt aaaactttag gctagtaaac ctgtaacact caacttagtt ctgaactatc 5640

tcactattct tttgcaagaa tttacttagg taatgccaac taatttattc caaggccaaa 5700

aagatgacaa tgtcttatat attataaaaa ctaataaaaa ccattttaaa acctagtata 5760

aatttaaagg tacttgctct tctggttcat ctcttctttt gtttacttct gctttcaaaa 5820

acttatttat tgtgaccata ttctttactt ccatttattg ttataattta taagatacta 5880

tacttgcaag caataaatgt tatcttttta gcttttaaat ggtctcattt gaaaagaata 5940

tataattagt aagtcatagc tactttaaat aaaaacttat tctttaagag attaaacact 6000

tctccaagtg atctgttttt ctttaattaa aacgttatta actcccaaaa tgatgttatt 6060

gtttttttat aatcttaaat accaataatt accaggtcta ttttgatttt gatacaggat 6120

aaaaactact attaattact taagaatgtg ttctttttta tatgtaccat tttcatgatc 6180

aaagttggtg atatgactga ggttttgatt attattaaac agatagttaa tatgatatat 6240

tcctcatttt tccaaatgaa aggaaaaatg tcttatatgg aggaaaagat tggggcaggg 6300

ggattagtaa attattactt aaatatctga ataggaggat ttttcaatga aaggataaag 6360

gaagaatgat tgtatcatct gaatctttcc ctccctttcc tggagtttgt cctttcaacc 6420

cagtatacct accactccct tcatcaccta ctttcccatt acagtcccta tgtgttgggt 6480

ggtaactatt ttgttttggt gttaatatcc aagtttccct taataacacc tagtgaatgg 6540

aggaaggatg agcataccta cccatcagac atatttagcc accatattta atcaacaagc 6600

atgaagaaag gaagctagcc tctccccttc ctttcctcct gcctctctct ctcttctctg 6660

tcctcgctcc ctttcttccc atcaatattt tcagagcacc tcttatgcgc caggcattgg 6720

gatactcaaa ctggaggaaa caagaaaaaa aaaaaaaaaa ggcgaagacc tcagggaaat 6780

ttatattgct gctatatttt tttgagccta gtgtaaatta aaattcctta atgctgtgcc 6840

ttttaaaaac acaaataagc aaaatagttt atttcttcaa cagttaaatc cttagggtag 6900

gaaagtgatt caggatctat tgctactatt aactcttctt tcattttcac acaggataca 6960

cacaaatccg cttggtgaat ggcaagaccc catgtgaagg aagagtggag ctcaacattc 7020

ttgggtcctg ggggtccctc tgcaactctc actgggacat ggaagatgcc catgttttat 7080

gccagcagct taaatgtgga gttgcccttt ctatcccggg aggagcacct tttgggaaag 7140

gaagtgagca ggtctggagg cacatgtttc actgcactgg gactgagaag cacatgggag 7200

attgttccgt cactgctctg ggcgcatcac tctgttcttc agggcaagtg gcctctgtaa 7260

tctgctcagg taagagaata agggcagcca gtgatgagcc actcatgacg gtgccttaag 7320

agtgggtgta cctaggagtt cccattgtgg ctcagtggta acaaactcga ctggtatcca 7380

tgagggtatg ggtttgatcc ctggccttgc tcaatgggtt aaggatccag cattgctgtg 7440

agctgtggta taggttgcag actctgctca ggtcccatgt tgctgtgatt gtggtgtagg 7500

ctgactgctg cagcttcaat ttgaccccta gcccgggaat ttccataggc cacacgtgca 7560

gcactaagga aggaaaaaaa gaaaaaaaaa aaaaaagagt gggtgtgcct atagtgaaga 7620

acagatgtaa aagggaagtg aaagggattc ccccattctg agggattgtg agaagtgtgc 7680

cagaatatta acttcatttg acttgttaca gggaaagtaa acttgacttt cacggacctc 7740

ctagttacct ggtgcttact atatgtcttc tcagagtacc tgattcattc ccagcctggt 7800

tgacccatcc ccctatctct atggctatgt ttatccagag cacatctatc taacactcca 7860

gctgatcttc ctgacacagc tgtggcaacc ctggatcctt taaccaactg tgccaggctg 7920

gagatcaaac ctaagcctct gcagcaaccc aagctgctgc agtcagattt ttaacccccc 7980

gtgccactgt gggtatctcc gatattttgt atcttctgtg actgagtggt ttgctgtttg 8040

cagggaacca gagtcagaca ctatccccgt gcaattcatc atcctcggac ccatcaagct 8100

ctattatttc agaagaaaat ggtgttgcct gcataggtga gaatcagtga ccaacctatg 8160

aaaatgatct caatcctctg aaatgcattt tattcatgtt ttatttcctc tttgcaggga 8220

gtggtcaact tcgcctggtc gatggaggtg gtcgttgtgc tgggagagta gaggtctatc 8280

atgagggctc ctggggcacc atctgtgatg acagctggga cctgaatgat gcccatgtgg 8340

tgtgcaaaca gctgagctgt ggatgggcca ttaatgccac tggttctgct cattttgggg 8400

aaggaacagg gcccatttgg ctggatgaga taaactgtaa tggaaaagaa tctcatattt 8460

ggcaatgcca ctcacatggt tgggggcggc acaattgcag gcataaggag gatgcaggag 8520

tcatctgctc gg 8532

<210> 119
<211> 4538
<212> ADN
<213> Sus scrofa

<400> 119
tatagatgac aaggctttgt gtctgatagg ggccagcgaa ctcagtaaag agggaagatg 60

agaaagataa tggcaagaat ttatccctga agtgtagttt tgacaaacca gtcacaaaga 120

ggtctaagaa attttggtca caaagttgtt ttgaatccca ggcattttat ttgcaatgat 180

tgcatatgtt ctggaaagga catctgaacc taagaaatag ttcatttgca ttgtgttata 240

ttttactaag gtctgagaaa taatcttgag atgagaatga actctacttc ttcagagtct 300

ggaaggaata aattatgaaa atgtattaat gcttctttaa accatattgt atatttatct 360

attactaaac aaaaagaagt agctctattt atttatttat ttatttattt atttatgtct 420

tttgtctctt tagggccaca cctgtggcat atggaggttc ccaggctaga ggtccaattg 480

gagatgtagc agccagccta tgccagagcc accgcaacac gggatctgag ccacgtctgt 540

gacttacacc acagctcaca gcaacgcctg atcctcaacc cactgagcga ggccagggat 600

cgaacccatg tcctcatgga tgctagttgg gttcgttaac tgctgagcca tgatgggaac 660

tccaaattaa ttatttctta tatttgttct tcatatattc atttctatag aaagaaataa 720

atacagattc agttaatgat ggcaggtaaa agcttaactt attaatcaaa ggagttaatc 780

caggcacaaa aattcaattc atggctctct gttaaaattt aggtataggt ttagcaggaa 840

gaaaaggtta gtagatgcag actattacat ttagaatgga tggacaatga agtcctacta 900

tacagcacag ggaactatat ccaatctctt gggatagaat atgatggaag acaaaatcag 960

aacaagagag tatatatata tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 1020

tgtgtgactg ggtcaccctg cggcacagca gaaattggca gaacattgta aatcaactat 1080

actttaatag gaaaaatact tttaagggct aaatttccaa tattctaacc atgtacacag 1140

agtaaatgtc ataaggatgc cagtctgtgt agagattgat gtgttactag cagattcatg 1200

aaataaaggc tgaggatgta gtccccaagt cacttctgag tggaagaatt tctcctttgt 1260

cctggactca aatattttag gataaaggaa aaaagaagat atttatagaa gggacttgtt 1320

ttcaagtact tgacaaaatt tcaccattaa agagaaattt gtgggagttc ccatcgtggc 1380

tcagtggaaa caaatccaac taggaaccat gaggttgtgg gtttgatccc tggcctcact 1440

cagtgggtta aggatccggt gttgccgtga gctgtggtgt aggttgcaga cacggttctg 1500

atcctgcgtt gctgtggctg tggctgtggt gtaggccagc agcaaacagc tctgattaga 1560

cccctagcct ggaaacctcc atatgccaca ggtgcagccc taaaaagaca aaaaaagaga 1620

aaagacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gaacccccag 1680

aggtatttat ttgtttttgc ctttttttcac tgactgttct ttgtttgttt gtttgagact 1740

gatctagaag actagagatt acaagaaata tggatttggc tcactctaag aaactgcttt 1800

cattccaagg tttgggtcta tccaaaagtg gaatagaatc atatgaatac tagtttatga 1860

gtatttagtg agaggaattt caagctcaaa taatgattca gcaagattaa attaaggagg 1920

gaattttcct tgtggctgag tgggttaagg acccaatgtt gtctctgtga ggatgtaggt 1980

tccatcctgg gctttgctca ttaggttaag gatctggcat tgctgcagct cagacccagt 2040

gctgccctgg ttgtggctta ggccaaagct gcagctccaa ttcaatctct ggcctgggaa 2100

cctccatgtg ctacaaggtg cggccttaaa aggaaaaaaa aaaaattaaa tcaaggactc 2160

aagagtcttt cattatttgt gttgtggaag ctatatttgt tttaaagtct tagttgtgtt 2220

tagaaagcaa gatgttcttc aactcaaatt tgggagggaa cttgtttcat acatttttaa 2280

tggataagtg gcaaaatttt catgctgagg tgatctatag tgttgtaatg cagaatatag 2340

tcagatcttg aacattttag gaagttggtg agggccaatt gtgtatctgt gccatgctga 2400

taagaatgtc aagggatcac aagaattcgt gttatttgac agcagtcatc tttaaaaggc 2460

atttgagaaa gtccaatttc aaatgcattt ccttttctta aaagataaat tgaagaaaat 2520

aagtctttat ttcccaagta aattgaattg cctctcagtc tgttaaaaga aactcttacc 2580

ttgatgattg cgctcttaac ctggcaaaga ttgtctttaa aatctgagct ccatgtcttc 2640

tgctttattt ctggtgtgcc tttgactcca gattacagta aatggaggac tgagtatagg 2700

gctaaaaagt agagagaatg gatgcatatt atctgtggtc tccaatgtga tgaatgaagt 2760

aggcaaatac tcaaaggaaa gagaaagcat gctccaagaa ttatgggttc cagaaggcaa 2820

agtcccagaa ttgtctccag ggaaggacag ggaggtctag aatcggctaa gcccactgta 2880

ggcagaaaaa ccaagaggca tgaatggctt ccctttctca cttttcactc tctggcttac 2940

tcctatcatg aaggaaaata ttggaatcat attctccctc accgaaatgc tatttttcag 3000

cccacaggaa acccagctca acattcttgg gtcctggggg tccctctgca actctcactg 3060

ggacatggaa gatgcccatg ttttatgcca gcagcttaaa tgtggagttg ccctttctat 3120

cccgggagga gcacctttg ggaaaggaag tgagcaggtc tggaggcaca tgtttcactg 3180

cactgggact gagaagcaca tgggagattg ttccgtcact gctctgggcg catcactctg 3240

ttcttcaggg caagtggcct ctgtaatctg ctcaggtaag agaataaggg cagccagtga 3300

tgagccactc atgacggtgc cttaagagtg ggtgtaccta ggagttccca ttgtggctca 3360

gtggtaacaa actcgactgg tatccatgag ggtatgggtt tgatccctgg ccttgctcaa 3420

tgggttaagg atccagcatt gctgtgagct gtggtatagg ttgcagactc tgctcaggtc 3480

ccatgttgct gtgattgtgg tgtaggctga ctgctgcagc ttcaatttga cccctagccc 3540

gggaatttcc ataggccaca cgtgcagcac taaggaagga aaaaaagaaa aaaaaaaaaa 3600

aagagtgggt gtgcctatag tgaagaacag atgtaaaagg gaagtgaaag ggattccccc 3660

attctgaggg attgtgagaa gtgtgccaga atattaactt catttgactt gttacaggga 3720

aagtaaactt gactttcacg gacctcctag ttacctggtg cttactatat gtcttctcag 3780

agtacctgat tcattcccag cctggttgac ccatcccccc atctctatgg ctatgtttat 3840

ccagagcaca tctatctaac actccagctg atcttcctga cacagctgtg gcaaccctgg 3900

atcctttaac caactgtgcc aggctggaga tcaaacctaa gcctctgcag caacccaagc 3960

tgctgcagtc agatttttaa ccccctgtgc cactgtgggt atctccgata ttttgtatct 4020

tctgtgactg agtggtttgc tgtttgcagg gaaccagagt cagacactat ccccgtgcaa 4080

ttcatcatcc tcggacccat caagctctat tatttcagaa gaaaatggtg ttgcctgcat 4140

aggtgagaat cagtgaccaa cctatgaaaa tgatctcaat cctctgaaat gcattttatt 4200

catgttttat ttcctctttg cagggagtgg tcaacttcgc ctggtcgatg gaggtggtcg 4260

ttgtgctggg agagtagagg tctatcatga gggctcctgg ggcaccatct gtgatgacag 4320

ctgggacctg aatgatgccc atgtggtgtg caaacagctg agctgtggat gggccattaa 4380

tgccactggt tctgctcatt ttggggaagg aacagggccc atttggctgg atgagataaa 4440

ctgtaatgga aaagaatctc atatttggca atgccactca catggttggg ggcggcacaa 4500

ttgcaggcat aaggaggatg caggagtcat ctgctcgg 4538

<210> 120
<211> 101
<212> PRT
<213> Sus scrofa

<400> 120

Pro Arg Leu Val Gly Gly Asp Ile Pro Cys Ser Gly Arg Val Glu Val
1 5 10 15

Gln His Gly Asp Thr Trp Gly Thr Val Cys Asp Ser Asp Phe Ser Leu
20 25 30

Glu Ala Ala Ser Val Leu Cys Arg Glu Leu Gln Cys Gly Thr Val Val
35 40 45

Ser Leu Leu Gly Gly Ala His Phe Gly Glu Gly Ser Gly Gln Ile Trp
50 55 60

Ala Glu Glu Phe Gln Cys Glu Gly His Glu Ser His Leu Ser Leu Cys
65 70 75 80

Pro Val Ala Pro Arg Pro Asp Gly Thr Cys Ser His Ser Arg Asp Val
85 90 95

Gly Val Val Cys Ser
100

<210> 121
<211> 100
<212> PRT
<213> Người hiện đại

<400> 121

Leu Arg Leu Val Asp Gly Asp Ser Arg Cys Ala Gly Arg Val Glu Ile
1 5 10 15

Tyr His Asp Gly Phe Trp Gly Thr Ile Cys Asp Asp Gly Trp Asp Leu
20 25 30

Ser Asp Ala His Val Val Cys Gln Lys Leu Gly Cys Gly Val Ala Phe
35 40 45

Asn Ala Thr Val Ser Ala His Phe Gly Glu Gly Ser Gly Pro Ile Trp
50 55 60

Leu Asp Asp Leu Asn Cys Thr Gly Met Glu Ser His Leu Trp Gln Cys
65 70 75 80

Pro Ser Arg Gly Trp Gly Gln His Asp Cys Arg His Lys Glu Asp Ala
85 90 95

Gly Val Ile Cys
100