



(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)



1-0037040

(51)¹⁹ C12N 15/90

(13) B

(21) 1-2019-07289

(22) 24/05/2018

(86) PCT/IL2018/050573 24/05/2018

(87) WO 2018/216022 29/11/2018

(30) 62/510,921 25/05/2017 US

(45) 25/09/2023 426

(43) 27/04/2020 385ASC

(73) EGGXYT LTD (IL)

Keren HaYessod 36, 9100760 Jerusalem, Israel

(72) OFFEN, Daniel (IL).

(74) Công ty Luật TNHH WINCO (WINCO LAW FIRM)

(54) PHƯƠNG PHÁP VÀ KIT DÙNG ĐỂ XÁC ĐỊNH GIỚI TÍNH CỦA PHÔI CHIM TRONG TRỨNG CHƯA NỞ

(57) Sáng chế đề cập đến phương pháp thụ tinh và xác định và nhận biết giới tính ở đối tượng thuộc loài chim. Cụ thể hơn, sáng chế đề xuất các phương pháp không xâm lấn bằng cách sử dụng động vật thuộc loài chim chuyển gen chứa ít nhất một gen thông báo, cụ thể là protein phát huỳnh quang màu đỏ (RFP), được tích hợp vào ít nhất một nhiễm sắc thể giới tính Z hoặc W. Động vật thuộc loài chim chuyển gen theo sáng chế được sử dụng để xác định giới tính và chọn lọc phôi trong trứng chim chưa nở. Sáng chế còn đề cập đến kit dùng để xác định giới tính của phôi chim trong trứng chưa nở chứa phôi ở trong vỏ trứng toàn vẹn về mặt cấu trúc.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp xác định và nhận biết giới tính ở đối tượng thuộc loài chim. Cụ thể hơn, sáng chế đề xuất phương pháp không xâm lấn và động vật thuộc loài chim chuyển gen để xác định giới tính và chọn lọc phôi trong trứng chim chưa nở.

Tình trạng kỹ thuật của sáng chế

Các tài liệu tham khảo được cho là có liên quan đến tình trạng kỹ thuật của đối tượng được bộc lộ trong bản mô tả này được liệt kê dưới đây:

- Công bố đơn quốc tế số WO 2010/103111
- Công bố đơn quốc tế số WO 2014/0296707
- Bằng độc quyền sáng chế Mỹ số US 6244214
- Đơn sáng chế số 06124456A2
- Công bố đơn sáng chế Mỹ số US2014069336A
- Công bố đơn quốc tế số WO16005539
- Công bố đơn quốc tế số WO 96/39505
- Công bố đơn quốc tế số WO 97/49806
- Quansah, E., Long, J.A., Donovan, D.M., Becker, S.C., Telugu, B., Foster Frey, J.A., Urwin, N. (2014). Sperm-mediated transgenesis in chicken using a PiggyBac transposon system. Poultry Science Association Meeting Abstract. BARC Poster Day.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, 337(6096), 816-821.
- Cong, L., & Zhang, F. (2015). Genome engineering using CRISPR-Cas9 system. *Chromosomal Mutagenesis*, 197-217.
- Bằng độc quyền sáng chế Mỹ số US 6.244.214.
- Công bố đơn quốc tế số WO 2014/0296707.
- Công bố đơn quốc tế số WO 06124456A2;

- Công bố đơn quốc tế số WO16005539.
- Véron N., Qu Z., Kipen PA., Hirst CE., Marcelle C. (2015). CRISPR mediated somatic cell genome engineering in the chicken. *Dev. Biol.* 407(1):68-74. doi: 10.1016/j.ydbio.2015.08.007. Epub 2015 Aug 13.
- Bằng độc quyền sáng chế Canada số CA2264450.
- Niu, Y., B. Shen, Y. Cui, Y. Chen, J. Wang et al., (2014). Generation of genemodified cynomolgus monkey via cas9/rna-mediated gene targeting in one-cell embryos. *Cell*, 156(4): 836–843.
- Hwang, W. Y., Y. Fu, D. Reyon, M. L. Maeder, S. Q. Tsai et al., (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* 31(3): 227-229.
- Nadège, V., Q. Zhengdong, P. A. S. Kipen, C. E. Hirst, M. Christophe et al., (2015). CRISPR mediated somatic cell genome engineering in the chicken. *Dev. Biol.* 407(1): 68-74.
- Bai, Y., L. He, P. Li, K. Xu, S. Shao et al., (2016). Efficient genome editing in chicken DF-1 cells using the CRISPR/Cas9 system. *G3 (Bethesda)* pii: g3.116.027706.
- Doran T. et al., (2016). Sex selection in layer chickens. *ASAP Animal Production*, Adelaide.
- Tizard M. và Doran T. (2014). Precision genome engineering in the chicken: Th gap between science and market place. A presentation at IWRAB II, in Brasilia

Việc ghi nhận các tài liệu tham khảo nêu trên trong bản mô tả này không được suy ra có nghĩa là các tài liệu này có sự liên quan theo cách bất kỳ đến khả năng cấp bằng độc quyền sáng chế cho đối tượng được bộc lộ trong bản mô tả này.

Trong ngành thực phẩm, hàng tỷ con gà non bị tiêu hủy hàng ngày thông qua việc làm nghẹt thở hoặc nghiền. Các con đực bị giết chết bởi vì chúng không hữu dụng đối với việc đẻ trứng hoặc để gây giống lấy thịt và các con cái yếu ớt hoặc không khỏe cũng bị giết chết. Do đó phương pháp xác định giới tính *in-ovo*, hoặc xác định giới tính phôi trước khi nở là rất cần thiết do sự cân nhắc về cả mặt đạo đức và mặt kinh tế.

Cụ thể, sự nhận biết bằng mắt thường đối với trứng gia cầm có khả năng thụ tinh là điều quan trọng để cho phép loại bỏ trứng không có khả năng thụ tinh để tiết kiệm chi phí ấp (bằng cách ngăn cản việc ấp trứng không có khả năng thụ tinh), và làm

giảm các nguy cơ an ninh sinh học có liên quan đến việc tiếp tục ấp trứng không có khả năng thụ tinh dễ bị nhiễm bệnh này kể bên trứng có khả năng thụ tinh.

Sự nhận biết bằng mắt thường về khả năng thụ tinh của trứng ở giai đoạn sớm của phôi, trong khi ở trong trứng chưa nở, bao gồm việc soi bằng nguồn sáng bên ngoài và có thể là khó khăn, gần như là không thể ở giai đoạn phôi sớm. Thử thách còn lớn hơn là nhận biết giới tính của phôi, và hiện nay không có phương pháp thích hợp để phân biệt giữa con đực và con cái trong trứng chưa nở mà được phát hiện là có khả năng thụ tinh. Tuy nhiên, sự nhận biết khả năng thụ tinh ở giai đoạn phôi sớm và sự xác định giới tính của gia cầm là cần thiết cho nghề nuôi gia cầm, nghiên cứu khoa học, và bảo tồn. Sự xác định giới tính ở chim non bằng đặc điểm hình thái là vô cùng khó khăn đối với hầu hết các loài. Có thể xác định giới tính bằng cách xác định giới tính lỗ huyết cá thể bao gồm việc ép phân ra khỏi con non bằng tay, mở nhẹ lỗ hậu môn của con non, cho phép nhìn xem liệu con non có "bướu" nhỏ, sẽ chỉ ra rằng con non là con đực. Tuy nhiên, phương pháp này có nguy cơ cao về gây tổn thương chim và sai lầm trong sự xác định giới tính, cùng với công việc nặng nề được tiến hành thủ công bởi người có tay nghề.

Sự xác định giới tính lỗ huyết hoặc xác định giới tính con non là phương pháp phân biệt giới tính của gà và các con mới nở khác, thường là bởi người có tay nghề được gọi là người xác định giới tính con non hoặc người xác định giới tính gà. Sự xác định giới tính gà được thực hiện chủ yếu bởi trại ấp trứng thương mại lớn, người này phải biết sự khác biệt giữa các giới tính để phân tách chúng thành các nhóm giới tính, và để đưa chúng vào các chương trình khác nhau, có thể bao gồm việc nuôi nhóm mong muốn và việc tiêu hủy nhóm không mong muốn không đáp ứng nhu cầu thương mại. Ví dụ, con đực nở ra từ trứng có nguồn gốc từ dòng thương mại để trứng để gây giống. Con đực đó sẽ không có sản lượng thịt tốt và không đẻ trứng, và do đó sẽ bị tiêu hủy sau khi xác định giới tính. Khi xác định giới tính, giới tính thích hợp tiếp tục tiến trình của nó để phục vụ mục đích trong khi giới tính kia hoặc hầu hết trong số chúng bị tiêu hủy trong những ngày ấp không thích hợp để sản xuất trứng.

Trong trại sản xuất trứng, các con đực là không mong muốn, và các con non có giới tính không mong muốn bị giết chết gần như ngay lập tức để giảm chi phí cho người chăn nuôi. Con non được chuyển xuống băng chuyền, tại đó người xác định giới tính con non tách các con đực ra và thả chúng vào máng trượt tại đó chúng thường bị

nghiên nhỏ khi còn sống trong máy nghiền thịt.

Sự nhận biết và xác định khả năng thụ tinh của trứng và giới tính của phôi trong trứng trước khi chúng nở, có khả năng loại trừ trứng không có khả năng thụ tinh, và loại không mong muốn của phôi khi ở trong trứng, và do đó làm giảm rất nhiều chi phí ấp (bao gồm chi phí năng lượng và hiệu quả song song với sự ô nhiễm không khí và sự tiêu thụ năng lượng). Ngoài ra, sự đau đớn của con non bị chấm dứt và sự ô nhiễm từ việc tiêu hủy bị ngăn chặn. Thiết bị xác định giới tính tự động còn dẫn đến chi phí sản xuất trứng giảm bằng cách loại trừ nhu cầu đối với người xác định giới tính con non, cũng như làm giảm kích thước của trại ấp trứng cần thiết bởi vì ở giai đoạn sớm 50% trứng sẽ được giảm trừ khỏi quá trình, do đó giảm chi phí ấp các quả trứng này, và sau đó cần phải có quy trình giết chết công phu bất kỳ.

Trong tất cả các loại chim thương mại nhằm mục đích gây giống, đẻ trứng, hoặc sản xuất thịt, đều cần phải xác định khả năng thụ tinh và giới tính của phôi. Có lợi nhuận kinh tế lớn; trong việc tiết kiệm năng lượng, giảm nguy cơ an ninh sinh học, loại bỏ chất thải, chi phí nhân công xác định giới tính và sai sót trong việc xác định giới tính, chi phí tiêu hủy và loại bỏ sự tiêu hủy, và phúc lợi động vật.

Công bố đơn quốc tế số WO 2010/103111 mô tả phương pháp can thiệp bao gồm loạt các bước, trong số đó có bước đưa kháng thể đã đánh dấu vào trong trứng, kháng thể này được thiết kế đặc hiệu để phù hợp với kháng nguyên đặc hiệu giới tính trên phôi.

Công bố đơn quốc tế số WO 2014/0296707 mô tả chế phẩm phát sáng được thiết kế để dùng làm chất chỉ thị sinh học để định lượng hoặc đánh giá hiệu quả chủng ngừa được tiêm vào trứng chim. Không có sự xác định giới tính nào được mô tả hoặc thậm chí gợi ý trong tài liệu này. Dụng cụ tiêm *in-ovo* và phương pháp phát hiện được bộc lộ trong Bằng độc quyền sáng chế Mỹ số US 6.244.214.

Công bố đơn quốc tế số WO 06124456A2 bộc lộ phương pháp xâm lấn *in-ovo* để xác định giới tính của phôi chim bằng cách xác định sự có mặt của hợp chất steroid estrogen trong mẫu dịch phôi (ví dụ, dịch niệu nang hoặc máu) từ trứng chim. Việc xác định sự có mặt của hợp chất này được thực hiện bằng cách xác định chất phân tích trong mẫu thu được từ trứng chim nêu trên bằng thử nghiệm miễn dịch cạnh tranh sử dụng kính hiển vi huỳnh quang.

Các phương pháp quang phổ cũng được mô tả, trong số đó Công bố đơn sáng

chế Mỹ số US2014069336A dựa trên sự sàng lọc màu lông phôi chim (trước khi nở) và xác định giới tính của phôi chim, dựa trên màu lông hoặc Công bố đơn quốc tế số WO16005539 bộc lộ thiết bị thu được đáp ứng quang phổ đặc hiệu với tín hiệu ánh sáng tới.

Các phương pháp di truyền khác hướng đến sự xác định giới tính *in ovo* bao gồm việc xác định trình tự ADN của mẫu ADN thu được từ trứng có khả năng thụ tinh để phát hiện hai gen đặc hiệu nằm trên nhiễm sắc thể Z và nhiễm sắc thể W của chim (Công bố đơn quốc tế số WO 96/39505), hoặc việc sử dụng đoạn dò oligonucleotit để lai hóa với trình tự đặc hiệu của nhiễm sắc thể W con cái (Công bố đơn quốc tế số WO 97/49806). Các phương pháp này là phương pháp xâm lấn và do đó không mang lại chiến lược an toàn.

Các đoạn lặp lại đối xứng ngắn cách đều thành cụm (clustered regularly interspaced short palindromic repeats: CRISPR)/hệ thống kết hợp CRISPR (Cas) là hệ thống chỉnh sửa gen trong lĩnh vực này, cho phép thiết kế cấu trúc đơn giản với tỷ lệ thành công cao (M. Jinek, 2012).

Niu và các đồng tác giả (2014) đã tiêm ARN dẫn đường (guide RNA: gARN) và ARN Cas9 vào noãn bào khi để cải biến ba gen đích, và Hwang và các đồng tác giả (2013) đã cải biến các gen *drd3* và *gsk3b* ở phôi cá ngựa vằn để thu được thể đột biến hai locus. Cong và Zhang (2015) đã cải biến hệ thống CRISPR để chỉnh sửa gen bất kỳ ở tế bào sống.

Veron và các đồng tác giả (2015) đã chứng minh được rằng mức độ biểu hiện của tế bào sinh dưỡng ở phôi gà được cải biến bằng cách gây xung điện plasmit gARN CRISPR được định hướng chống lại yếu tố phiên mã PAX7 (Nadège et al. 2015), Bai và các đồng tác giả đã chỉnh sửa PPAR-g, tiểu đơn vị ATP synthaza epsilon (ATP5E).

Quansah, E. và các đồng tác giả, đã bộc lộ sự chuyển gen qua trung gian tinh dịch ở gà bằng cách sử dụng hệ gen nhảy PiggyBac. Cụ thể, họ bộc lộ rằng dạng kết hợp của plasmit aGFP và Lipofectamin LTXTM 9LPX) không có tác dụng nào đối với khả năng sống, tính di động hoặc khả năng thụ tinh của tinh dịch gà.

Bằng độc quyền sáng chế Canada số CA2264450 bộc lộ tế bào đa năng chứa trình tự axit nucleic mã hóa chất chỉ thị protein phát huỳnh quang được tích hợp có chọn lọc vào nhiễm sắc thể giới tính khác loại ở phôi bào và động vật chuyển gen được tạo ra bằng cách sử dụng tế bào đa năng này và, việc sử dụng các tế bào, phôi, và

động vật này. Công bố này đề cập cụ thể đến GFP ở tế bào của động vật có vú.

Doran T. và các đồng tác giả, mô tả chung chiến lược để phân biệt giữa con đực và con cái trước khi nở, bằng cách bổ sung chất chỉ thị sinh học, gen thông báo GFP, vào nhiễm sắc thể giới tính, trong tài liệu ASAP Animal Production 2016, Adelaide. Sự tích hợp gen thông báo GFP vào nhiễm sắc thể giới tính của các con non bằng cách sử dụng công nghệ CRISPR cũng được mô tả trong bản trình bày tại IWRAB^{II}, ở Brasilia, năm 2014. Tuy nhiên, theo tài liệu công bố chung này bộc lộ rằng nhiễm sắc thể đã đánh dấu được quan sát bằng mắt thường *in ovo* bằng cách cho trứng tiếp xúc với ánh sáng UV, do sự tự phát huỳnh quang của trứng, không có tín hiệu có thể phát hiện được mong đợi.

Do đó, các phương pháp hiệu quả và không xâm lấn để nhận biết giới tính trong suốt giai đoạn trứng, trước khi nở ra con non hiện nay không có sẵn. Do đó, có mong muốn từ lâu về phương pháp có khả năng nhận biết giới tính chính xác và an toàn của phôi trong trứng chưa nở.

Bản chất kỹ thuật của sáng chế

Khía cạnh thứ nhất của sáng chế đề cập đến phương pháp xác định giới tính của chim, hoặc phôi chim trong trứng chưa nở, cụ thể là trứng chưa nở có khả năng thụ tinh. Theo một số phương án cụ thể, phương pháp này có thể bao gồm bước:

Trong bước thứ nhất, bước (a), cung cấp hoặc thu được ít nhất một đối tượng hoặc động vật thuộc loài chim chuyển gen chứa ít nhất một gen thông báo ngoại sinh được tích hợp vào ít nhất một vị trí hoặc địa điểm (còn được gọi trong bản mô tả này là locus) ở trong ít nhất một trong hai nhiễm sắc thể giới tính Z và W. Trong bước thứ hai (b), thu được ít nhất một trứng đã thụ tinh từ đối tượng thuộc loài chim chuyển gen hoặc tế bào bất kỳ của chúng.

Bước tiếp theo (c) bao gồm việc xác định trong trứng liệu ít nhất một tín hiệu có thể phát hiện được phát hiện. Theo phương án cụ thể hơn, sự phát hiện ít nhất một tín hiệu có thể phát hiện cho thấy sự biểu hiện của ít nhất một gen thông báo này, nhờ đó chỉ ra sự có mặt của nhiễm sắc thể W hoặc nhiễm sắc thể Z ở phôi chim. Theo một số phương án cụ thể, gen thông báo được tích hợp có thể mã hóa protein phát huỳnh quang màu đỏ (red fluorescent protein: RFP).

Theo khía cạnh thứ hai, sáng chế đề cập đến động vật chuyển gen thuộc loài

chim chứa, trong ít nhất một tế bào của chúng, ít nhất một gen thông báo ngoại sinh được tích hợp vào ít nhất một vị trí hoặc địa điểm trong ít nhất một trong hai nhiễm sắc thể giới tính Z và W. Theo một số phương án cụ thể, gen thông báo được tích hợp có thể mã hóa RFP.

Theo khía cạnh khác nữa, sáng chế đề cập đến tế bào chứa ít nhất một gen thông báo ngoại sinh được tích hợp vào ít nhất một vị trí hoặc locus trong ít nhất một trong hai nhiễm sắc thể giới tính Z và W.

Theo khía cạnh khác nữa, sáng chế đề xuất kit chứa:

(a) ít nhất một protein Cas9 hoặc các mảnh hoặc dẫn xuất bất kỳ của nó, hoặc ít nhất một trình tự axit nucleic thứ nhất chứa ít nhất một trình tự axit nucleic mã hóa ít nhất một protein Cas9 này; và ít nhất một trình tự axit nucleic mã hóa ít nhất một ARN dẫn đường (gARN) hướng đích ít nhất một đoạn tiền đệm nằm ở trong nhiễm sắc thể giới tính Z hoặc W; và (b) ít nhất một trình tự axit nucleic thứ hai chứa ít nhất một gen thông báo này.

Ngoài ra, sáng chế đề xuất phương pháp xác định và phát hiện sự thụ tinh trong trứng chưa nở. Cụ thể hơn, phương pháp này bao gồm bước:

Trong bước thứ nhất (a), cung cấp hoặc thu được ít nhất một đối tượng hoặc động vật thuộc loài chim chuyển gen chứa ít nhất một gen thông báo ngoại sinh được tích hợp vào ít nhất một vị trí hoặc địa điểm ở trong cả hai nhiễm sắc thể giới tính Z và W, trong trường hợp là đối tượng con cái và ở trong cả hai nhiễm sắc thể giới tính Z ở con đực. Trong bước thứ hai (b), thu được ít nhất một trứng đã thụ tinh từ đối tượng thuộc loài chim chuyển gen, hoặc tế bào bất kỳ của chúng. Bước tiếp theo (c) bao gồm việc xác định trong trứng liệu ít nhất một tín hiệu có thể phát hiện được phát hiện. Theo phương án cụ thể hơn, sự phát hiện ít nhất một tín hiệu có thể phát hiện được chỉ ra sự biểu hiện của ít nhất một gen thông báo này, nhờ đó chỉ ra sự có mặt của nhiễm sắc thể W hoặc nhiễm sắc thể Z mẹ được đánh dấu (trong trường hợp là con cái) hoặc nhiễm sắc thể Z bố được đánh dấu ở phôi chim.

Các phương án này và phương án khác của sáng chế trở nên rõ ràng bằng cách tham chiếu đến các hình vẽ sau đây.

Mô tả vắn tắt các hình vẽ

Để hiểu rõ hơn đối tượng sáng chế được bộc lộ trong bản mô tả này và để minh

họa cách mà nó có thể được tiến hành trong thực tiễn, các phương án sẽ được mô tả sau đây chỉ bằng ví dụ không làm giới hạn sáng chế, có tham chiếu đến các hình vẽ kèm theo, trong đó:

Fig.1A-1B thể hiện chế độ thiết lập đầy đủ để đánh giá mức độ phát huỳnh quang của trứng

Fig.1A thể hiện ảnh chụp của thiết bị được sử dụng bằng các phương pháp theo sáng chế để xác định giới tính *in ovo*. Sự thiết lập đầy đủ gồm có nguồn laze (h), bộ giữ nguồn laze (g), bộ đỡ trứng (e), trứng (f), thấu kính (d), bộ lọc (c), bộ đỡ thiết bị phát hiện (b) và thiết bị phát hiện (a). Các thành phần khác được đặt trên giá đỡ rắn (i).

Fig.1B thể hiện hình ảnh minh họa dạng sơ đồ của thiết bị theo sáng chế đã thể hiện trên Fig.1A. Thiết bị này bao gồm nguồn laze (h), bộ giữ nguồn laze (g), bộ đỡ trứng (e), trứng (f), thấu kính (d), bộ lọc (c), bộ đỡ thiết bị phát hiện (b) và thiết bị phát hiện (a). Các thành phần khác được đặt trên giá đỡ rắn (i).

Fig.2 thể hiện cường độ phát huỳnh quang bằng cách sử dụng laze màu xanh lam (473nm) có hoặc không có florescein

Cường độ phát huỳnh quang [P ra] được thu thập trên thiết bị phát hiện khi toàn bộ trứng được chiếu laze màu xanh lam có hoặc không có florescein (fl) 10 μ M hoặc 1mM, và cường độ được thể hiện dưới dạng hàm số của cường độ nguồn sáng [P vào].

Fig.3 thể hiện cường độ phát huỳnh quang bằng cách sử dụng laze màu xanh lục (532nm) có hoặc không có Rodamin

Cường độ phát huỳnh quang [P ra] được thu thập trên thiết bị phát hiện khi toàn bộ trứng được chiếu laze màu xanh lục có hoặc không có Rodamin, và cường độ được thể hiện dưới dạng hàm số của cường độ nguồn sáng [P vào].

Fig.4 thể hiện cường độ phát huỳnh quang bằng cách sử dụng laze màu xanh lục (532nm) có hoặc không có dir

Cường độ phát huỳnh quang [P ra] được thu thập trên thiết bị phát hiện, khi toàn bộ trứng được chiếu laze màu xanh lục có hoặc không có dir, và cường độ được thể hiện dưới dạng hàm số của cường độ nguồn sáng [P vào].

Fig.5 thể hiện cường độ phát huỳnh quang bằng cách sử dụng laze màu đỏ (632,8nm) có hoặc không có dir

Cường độ phát huỳnh quang [P ra] được thu thập trên thiết bị phát hiện, khi toàn bộ trứng được chiếu laze màu đỏ có hoặc không có dir, và cường độ được thể

hiện dưới dạng hàm số của cường độ nguồn sáng [P vào].

Fig.6 thể hiện tế bào biểu hiện GFP

Hình ảnh cung cấp tín hiệu huỳnh quang của các tế bào HEK được chuyển nhiễm bằng vector GFP biểu hiện GFP.

Fig.7 thể hiện tế bào biểu hiện RFP

Hình ảnh cung cấp tín hiệu huỳnh quang của tế bào HEK được chuyển nhiễm bằng vector RFP biểu hiện RFP.

Fig.8 thể hiện phép đo đối chứng

Cường độ phát huỳnh quang bằng cách sử dụng laze màu xanh lục và bộ lọc màu đỏ với các vị trí khác nhau của trứng. Cường độ phát huỳnh quang [P ra] được thu thập trên thiết bị phát hiện được thể hiện dưới dạng hàm số của cường độ nguồn sáng [P vào].

Fig.9 thể hiện phép đo RFP

Các phép đo cường độ phát huỳnh quang của trứng được tiêm với các nồng độ khác nhau của tế bào biểu hiện RFP bằng cách sử dụng laze màu xanh lục và bộ lọc màu đỏ. Cường độ phát huỳnh quang [P ra] được thu thập trên thiết bị phát hiện được thể hiện dưới dạng hàm số của cường độ nguồn sáng [P vào].

Fig.10 thể hiện góc RFP

Phép đo cường độ phát huỳnh quang RFP của trứng được kích thích ở vị trí tối ưu đã xác định của trứng bằng cách sử dụng laze màu xanh lục và bộ lọc màu đỏ. Cường độ phát huỳnh quang [P ra] được thu thập trên thiết bị phát hiện được thể hiện dưới dạng hàm số của cường độ nguồn sáng [P vào].

Fig.11A-11B thể hiện tế bào biểu hiện RFP với PBS hoặc glycerol

Hình vẽ thể hiện cường độ phát huỳnh quang của trứng được tiêm với các nồng độ khác nhau của tế bào biểu hiện RFP với PBS hoặc glycerol bằng cách sử dụng laze màu xanh lục.

Fig.11A thể hiện cường độ phát huỳnh quang của trứng chứa các nồng độ khác nhau của tế bào biểu hiện RFP với nhau trong PBS. **Fig.11B** thể hiện cường độ phát huỳnh quang của trứng chứa các nồng độ khác nhau của tế bào biểu hiện RFP trong glycerol. Cường độ phát huỳnh quang [P ra] được thu thập trên thiết bị phát hiện được thể hiện dưới dạng hàm số của cường độ nguồn sáng [P vào].

Fig.12 thể hiện tế bào biểu hiện GFP với PBS hoặc glycerol

Hình vẽ thể hiện cường độ phát huỳnh quang của trứng được tiêm 30.000 tế bào biểu hiện GFP với PBS hoặc glycerol bằng cách sử dụng laze màu xanh lam. Cường độ phát huỳnh quang [P ra] được thu thập trên thiết bị phát hiện được thể hiện dưới dạng hàm số của cường độ nguồn sáng [P vào].

Fig.13 thể hiện RFP so với GFP

Hình vẽ thể hiện tỷ lệ giữa cường độ của protein phát huỳnh quang từ tế bào biểu hiện GFP hoặc tế bào biểu hiện RFP và cường độ tự phát huỳnh quang.

Thông số R mô tả tỷ lệ giữa sự phát huỳnh quang và sự tự phát huỳnh quang được thể hiện dưới dạng hàm số của cường độ nguồn sáng [P vào].

Fig.14 thể hiện sự kết hợp RFP vào nhiễm sắc thể Z cái

Hình vẽ thể hiện dòng tế bào gà mái được chuyển nhiễm RFP bằng pDsRed với các nhánh trái & phải ChZ ở gà và CMV-hspCas9-H1-gARN.

Fig.15A-15B là hình biểu diễn dạng sơ đồ của trình tự được tích hợp vào nhiễm sắc thể Z của dòng tế bào gà mái

Trình tự in đậm thể hiện nhánh bên trái và nhánh bên phải hai bên của nhiễm sắc thể Z và được thể hiện bằng trình tự SEQ ID NO: 60 và SEQ ID NO: 61 tương ứng; Trình tự (1017bp) được đánh dấu bằng chữ in nghiêng thể hiện nhánh trái, như được thể hiện bằng trình tự SEQ ID NO: 62; Trình tự gạch Chân (629bp) thể hiện gen tăng cường CMV, gen khởi đầu và MCS, như được thể hiện bằng trình tự SEQ ID NO: 63; Trình tự in nghiêng đậm (714bp) thể hiện gen thông báo dsRED2, như được thể hiện bằng trình tự SEQ ID NO: 64; Trình tự không có màu-3332bp thể hiện nội dung plasmit dsRED; Trình tự in nghiêng gạch chân (1026bp) thể hiện nhánh bên phải, như được thể hiện bằng trình tự SEQ ID NO: 65. Phần bên trái của trình tự được thể hiện bằng trình tự SEQ ID NO. 66 (tất cả các phần ở đầu 5' đối với trình tự vector, được thể hiện trên Fig.15A), và phần bên phải của trình tự được thể hiện bằng trình tự SEQ ID NO. 67 (tất cả các phần ở đầu 3' đối với trình tự vector, được thể hiện trên Fig.15B).

Mô tả chi tiết sáng chế

Mỗi ngày, hàng tỷ con non đực bị giết chết thông qua việc làm nghẹt thở hoặc nghiền bởi vì chúng không hữu dụng để đẻ trứng hoặc để gây giống lấy thịt. Khả năng xác định giới tính của phôi trước khi ấp có tầm quan trọng lớn về cả mặt đạo đức và tài chính.

Ở gà- cấu trúc di truyền của nhiễm sắc thể giới tính là ZZ đối với con đực và ZW đối với con cái. Nghĩa là nhiễm sắc thể W xác định giới tính của con cái. Điều này không giống như người, trong đó Y từ bố xác định giới tính đực.

Sáng chế đề xuất phương pháp hiệu quả không xâm lấn để xác định giới tính, bằng cách sử dụng gen thông báo được tích hợp trong nhiễm sắc thể đặc hiệu giới tính của đối tượng thuộc loài chim chuyển gen. Sự biểu hiện của gen thông báo này ở phôi của trứng chưa nở xác định rõ ràng và chính xác giới tính của phôi đó.

Do đó, khía cạnh thứ nhất của sáng chế đề cập đến phương pháp xác định giới tính và tùy ý chọn lọc chim, hoặc phôi chim trong trứng chưa nở, cụ thể là trứng chưa nở đã thụ tinh. Theo một số phương án cụ thể, phương pháp này có thể bao gồm các bước:

Trong bước thứ nhất (a), cung cấp hoặc thu được ít nhất một đối tượng hoặc động vật thuộc loài chim chuyển gen chứa ít nhất một gen thông báo ngoại sinh được tích hợp vào ít nhất một vị trí hoặc địa điểm ở trong ít nhất một trong hai nhiễm sắc thể giới tính Z và W. Trong bước thứ hai (b) thu được ít nhất một trứng đã thụ tinh từ đối tượng, cụ thể là động vật, thuộc loài chim chuyển gen hoặc tế bào bất kỳ của chúng.

Bước tiếp theo (c), bao gồm việc xác định trong trứng có ít nhất một tín hiệu có thể phát hiện được hay không. Theo phương án cụ thể hơn, sự phát hiện ít nhất một tín hiệu có thể phát hiện chỉ ra sự biểu hiện của ít nhất một gen thông báo, nhờ đó chỉ ra sự có mặt của nhiễm sắc thể W hoặc nhiễm sắc thể Z ở phôi chim. Do đó, trong trường hợp gen thông báo đã được tích hợp vào nhiễm sắc thể Z của chim cái chuyển gen, sự nhận biết tín hiệu có thể phát hiện được trong trứng được thử nghiệm chỉ ra rằng phôi được chứa trong đó có nhiễm sắc thể Z mẹ có gen thông báo đã tích hợp trong đó, và nhờ đó phôi được nhận biết là con đực. Theo một số phương án khác nữa, trong trường hợp gen thông báo đã được tích hợp vào nhiễm sắc thể W của chim cái chuyển gen, sự nhận biết tín hiệu có thể phát hiện được trong trứng được thử nghiệm chỉ ra rằng phôi được chứa trong đó mang nhiễm sắc thể W mẹ và do đó được xác định là con cái, nhờ đó cung cấp sự xác định giới tính của phôi.

Cần hiểu rằng chim chuyển gen được đề xuất bởi sáng chế có thể là con cái hoặc con đực, như được mô tả chi tiết hơn dưới đây. Theo một số phương án cụ thể khác nữa, động vật chuyển gen có thể thuộc giới có hai nhiễm sắc thể giới tính khác nhau, cụ thể là đối tượng dị giao tử. Theo một số phương án cụ thể, động vật dị giao tử

có thể là con cái. Theo phương án cụ thể hơn, khi đó đối tượng thuộc loài chim chuyển gen là con cái, trứng được nhận biết bằng phương pháp theo sáng chế được đề bởi chim cái chuyển gen được đề xuất bởi sáng chế. Theo phương án cụ thể hơn, con cái chuyển gen có thể được thụ tinh bởi con đực chuyển gen hoặc theo một số phương án khác, bởi chim đực kiểu đại. Ngoài ra, sự thụ tinh có thể xảy ra bằng cách giao phối hoặc bằng cách cấy tinh chim cái chuyển gen bằng tinh dịch thu được từ chim đực chuyển gen hoặc kiểu đại. Theo một số phương án khác nữa, khi đó chim chuyển gen là con đực, trứng được nhận biết bằng phương pháp theo sáng chế có thể được đề bởi con cái kiểu đại hoặc chuyển gen được giao phối với con đực chuyển gen được đề xuất bởi sáng chế, hoặc được cấy tinh bằng tế bào bất kỳ của chúng, cụ thể là tinh bào có chứa gen thông báo ngoại sinh theo sáng chế được tích hợp vào nhiễm sắc thể giới tính của chúng.

Do đó, sáng chế đề xuất phương pháp phát hiện giới tính của phôi chim ở trong trứng đã thụ tinh chưa nở. Cần hiểu rằng phương pháp theo sáng chế có thể áp dụng cho trứng chưa nở ở giai đoạn phôi bất kỳ của phôi chim.

Cần lưu ý rằng khi được sử dụng trong bản mô tả này, thuật ngữ "**giai đoạn hoặc bước phát triển phôi của phôi chim**" để chỉ giai đoạn của **ngày 1** trong đó đĩa mầm ở giai đoạn bì phôi và khoang phân đoạn có hình dạng của vòng tròn sẫm màu; giai đoạn của **ngày 2** trong đó đường rãnh đầu tiên xuất hiện ở trung tâm của bì phôi và màng noãn hoàng xuất hiện; giai đoạn của **ngày 3** trong đó sự tuần hoàn máu bắt đầu, đầu và thân có thể phân biệt được, cũng như não và cấu trúc tim bắt đầu đập; giai đoạn của **ngày 4** trong đó khoang màng ối phát triển bao quanh phôi và túi niệu xuất hiện; giai đoạn của **ngày 5** trong đó phôi có hình dạng C và các chi kéo dài; giai đoạn của **ngày 6** trong đó các ngón của chi trên và chi dưới trở nên rõ ràng; giai đoạn của **ngày 7** trong đó cổ phân tách rõ đầu khỏi cơ thể, mỏ được hình thành và não dần dần đi vào vùng đầu; giai đoạn của **ngày 8** trong đó sự tạo sắc tố mắt dễ dàng nhìn thấy, cánh và chân được biệt hóa và ống tai ngoài mở; giai đoạn của **ngày 9** trong đó móng xuất hiện và các nang lông đầu tiên mọc ra; giai đoạn của **ngày 10** trong đó lỗ mũi xuất hiện, mí mắt phát triển và răng-trứng xuất hiện; giai đoạn của **ngày 11** trong đó kẽ hở mí mắt có hình dạng elip và phôi có vẻ ngoài của con non; giai đoạn của **ngày 12** trong đó nang lông xung quanh ống tai ngoài và phủ mí mắt trên trong khi đó mí mắt dưới phủ phần lớn giác mạc; giai đoạn của **ngày 13** trong đó túi niệu trở thành

màng đệm túi niệu trong khi móng và vảy chân trở nên rõ ràng; giai đoạn của **các ngày từ 14 đến 16** trong đó toàn bộ cơ thể sinh trưởng nhanh chóng, sự co giãn hoàng tăng nhanh và lòng trắng trứng dần dần biến mất; giai đoạn của **ngày 17** trong đó hệ thận sản sinh urat, mô chỉ đến khoang khí và lòng trắng trứng được tái hấp thụ hoàn toàn; giai đoạn của **ngày 18** trong đó noãn hoàng nội tại hóa và lượng dịch lỏng ối giảm; giai đoạn của **ngày 19** trong đó sự tái hấp thụ noãn hoàng tăng nhanh và mô sẵn sàng chọc thủng màng vỏ trong; giai đoạn của **ngày 20** trong đó noãn hoàng được tái hấp thụ hoàn toàn, rốn đóng lại, con non chọc thủng màng vỏ trong, hô hấp trong khoang khí và sẵn sàng nở; giai đoạn của **ngày 21** trong đó con non chọc thủng vỏ theo đường vòng quanh bằng răng-trứng của nó, tách bản thân nó ra khỏi vỏ trong khoảng thời gian từ 12 đến 18 giờ và để nó khô ráo.

Cụ thể hơn, phương pháp theo sáng chế có thể áp dụng trong việc xác định giới tính của phôi chim *in-ovo*, bên trong trứng, ở mỗi giai đoạn của quá trình phát triển phôi. Cụ thể hơn, từ ngày 1, từ ngày 2, từ ngày 3, từ ngày 4, từ ngày 5, từ ngày 6, từ ngày 7, từ ngày 8, từ ngày 9, từ ngày 10, từ ngày 11, từ ngày 12, từ ngày 13, từ ngày 14, từ ngày 15, từ ngày 16, từ ngày 17, từ ngày 18, từ ngày 19, từ ngày 20 và từ ngày 21. Cụ thể hơn, phương pháp theo sáng chế có thể áp dụng để phát hiện sớm giới tính của phôi, cụ thể là từ ngày 1 đến ngày 10. Theo một số phương án, phương pháp của sáng chế có thể cho phép xác định giới tính của phôi vào ngày một của sự phát triển phôi. Theo một số phương án khác, phương pháp của sáng chế có thể cho phép xác định giới tính của phôi vào ngày hai. Theo một số phương án khác, phương pháp của sáng chế có thể cho phép xác định giới tính của phôi vào ngày ba. Theo các phương án khác, phương pháp của sáng chế có thể cho phép xác định giới tính của phôi vào ngày bốn. Theo một số phương án khác nữa, phương pháp của sáng chế có thể cho phép xác định giới tính của phôi vào ngày năm. Theo các phương án khác, phương pháp của sáng chế có thể cho phép xác định giới tính của phôi vào ngày sáu. Theo các phương án khác, phương pháp của sáng chế có thể cho phép xác định giới tính của phôi vào ngày bảy. Theo các phương án khác, phương pháp của sáng chế có thể cho phép xác định giới tính của phôi vào ngày tám. Theo một số phương án khác, phương pháp của sáng chế có thể cho phép xác định giới tính của phôi vào ngày chín. Theo một số phương án khác nữa, phương pháp của sáng chế có thể cho phép xác định giới tính của phôi vào ngày mười. Theo một số phương án cụ thể, phương pháp của sáng chế có thể

cho phép xác định giới tính của phôi trong khoảng từ ngày 1 đến ngày 5 của sự phát triển phôi.

Như đã lưu ý trên đây, phương pháp theo sáng chế có thể áp dụng cho trứng chưa nở đã thụ tinh. Thuật ngữ "**trứng đã thụ tinh**" trong bản mô tả này dùng để chỉ trứng được đẻ bởi con mái trong đó con mái đã được giao phối bởi con trống trong thời gian hai tuần, cho phép lắng đọng tinh dịch con đực vào cơ cấu hình phễu con cái và sự kiện thụ tinh xảy ra khi giải phóng noãn khỏi buồng trứng. Khi được sử dụng trong bản mô tả này, thuật ngữ "**trứng chưa nở**", để chỉ trứng chứa phôi (còn được gọi trong bản mô tả này là trứng thụ tinh) ở trong vỏ toàn vẹn về mặt cấu trúc (không bị vỡ).

Phương pháp theo sáng chế dựa trên sự xác định tín hiệu có thể phát hiện được được tạo ra bởi gen thông báo đã tích hợp vào locus đặc hiệu của chim cái hoặc chim đực chuyển gen để ra trứng được kiểm tra.

Khi được sử dụng trong bản mô tả này, thuật ngữ "**sự tích hợp ADN/gen ngoại lai hoặc ngoại sinh vào nhiễm sắc thể**", sau đây dùng để chỉ sự cải biến lâu dài của trình tự nucleotit của nhiễm sắc thể sinh vật. Sự cải biến này được truyền tiếp trong quá trình phân chia tế bào và nếu xảy ra ở dòng tế bào mầm, nó cũng được truyền đến đời sau. Trong trường hợp này, gen thông báo đã tích hợp có thể được truyền đến phôi ở trong trứng chưa nở. Khi được sử dụng trong bản mô tả này, thuật ngữ "**ngoại sinh**", dùng để chỉ việc có nguồn gốc từ bên ngoài sinh vật đã được đưa vào sinh vật, ví dụ, bằng cách biến nạp hoặc chuyển nhiễm bằng cụ thể là các vector, virus đã được thao tác hoặc các phương tiện khác bất kỳ. Gen ngoại lai được tích hợp Theo các phương án nhất định, có thể là gen thông báo. Thuật ngữ "**gen thông báo**" liên quan đến gen mã hóa polypeptit, mà sự biểu hiện của nó có thể được phát hiện trong các thử nghiệm đã biết khác nhau và trong đó mức độ của tín hiệu được phát hiện biểu thị sự có mặt của gen thông báo này.

Như đã lưu ý trên đây, có thể tích hợp gen thông báo ngoại sinh vào nhiễm sắc thể giới tính chim Z hoặc W. Khi được sử dụng trong bản mô tả này, thuật ngữ "**nhiễm sắc thể giới tính Z hoặc W**" của chim để chỉ hệ nhiễm sắc thể xác định giới tính của đời sau ở gà trong đó con đực là giới đồng giao tử (ZZ), trong khi con cái là giới dị giao tử (ZW). Sự có mặt của nhiễm sắc thể W trong noãn xác định giới tính của đời sau trong khi nhiễm sắc thể Z được biết là lớn hơn và chứa nhiều gen hơn.

Phương pháp theo sáng chế dựa trên sự phát hiện tín hiệu có thể phát hiện chỉ ra và phản ánh sự có mặt của gen thông báo và nhờ đó chỉ ra sự có mặt của nhiễm sắc thể giới tính cụ thể. "**Tín hiệu có thể phát hiện được**" trong bản mô tả này dùng để chỉ sự thay đổi có thể nhận biết thấy bằng cách quan sát hoặc bằng công cụ. Không làm giới hạn, tín hiệu có thể được phát hiện theo cách trực tiếp. Theo một số phương án, đáp ứng có thể phát hiện được là tín hiệu quang học.

Cần hiểu rõ rằng theo một số phương án cụ thể, ít nhất một đối tượng thuộc loài chim chuyển gen được tạo ra bởi phương pháp theo sáng chế, có thể chứa ít nhất hai gen thông báo khác nhau, mỗi gen thông báo có thể được tích hợp vào ít nhất một vị trí hoặc địa điểm ở một trong hai nhiễm sắc thể giới tính Z hoặc W. Trong trường hợp có ít nhất hai gen thông báo khác nhau, theo một số phương án, mỗi nhiễm sắc thể giới tính có thể được đánh dấu khác nhau. Sự đánh giá tín hiệu có thể phát hiện được được tạo thành, có thể biểu thị giới tính của phôi được kiểm tra.

Theo một số phương án cụ thể khác, gen thông báo được chứa ở trong chim chuyển gen theo sáng chế có thể là ít nhất một gen thông báo phát huỳnh quang. Do đó, theo một số phương án, polypeptit được biểu hiện là protein phát huỳnh quang và theo đó thử nghiệm đo mức độ ánh sáng phát ra khi kích thích, cụ thể là bằng nguồn sáng thích hợp.

Thuật ngữ "**sự phát huỳnh quang**" dùng để chỉ sự phát ánh sáng bởi chất đã được chiếu, được hấp thụ ánh sáng hoặc sự bức xạ điện từ khác. Nó là một dạng của sự phát quang tức là sự phát ánh sáng bởi chất không phải do nhiệt.

"**Protein phát huỳnh quang**" dùng để chỉ protein có huỳnh quang sáng chói khi tiếp xúc với ánh sáng. Protein phát huỳnh quang có bước sóng cụ thể đối với đỉnh của cường độ kích thích chiếu vào và bước sóng đối với đỉnh của cường độ phát huỳnh quang. Sự kích thích dùng để chỉ sự chiếu của protein phát huỳnh quang. Theo một số phương án, sự kích thích có thể được tạo ra bởi laze. Theo một số phương án khác, tín hiệu có thể được phát hiện khi sử dụng bộ lọc quang học thích hợp.

Theo các phương án nhất định, gen thông báo phát huỳnh quang có thể là protein phát huỳnh quang màu đỏ (RFP), protein phát huỳnh quang màu xanh ngọc (CFP), protein phát huỳnh quang màu vàng (YFP), và protein tự phát huỳnh quang bao gồm cả protein phát huỳnh quang màu xanh lam (BFP).

Cần lưu ý rằng theo một số phương án, phương án cụ thể khác, protein phát

huỳnh quang bất kỳ có thể được sử dụng trong sáng chế. Cụ thể hơn, nỗ lực gây đột biến ở protein phát huỳnh quang màu xanh lục của sứa *Aequorea victoria* gốc dẫn đến đoạn dò huỳnh quang mới nằm trong khoảng giới hạn màu từ màu xanh lam đến màu vàng, có thể được sử dụng theo một số phương án trong sáng chế. Protein phát huỳnh quang có bước sóng dài hơn, phát ra ở vùng quang phổ màu cam và màu đỏ, đã được phát triển từ hải quỳ biển, *Discosoma striata*, và rặng san hô thuộc lớp **Anthozoa**. Các loài khác vẫn được khai thác để tạo ra các protein tương tự có sự phát huỳnh quang màu xanh ngọc, màu xanh lục, màu vàng, màu cam, và màu đỏ thẫm.

Theo một số phương án, gen thông báo có thể được sử dụng trong các phương pháp, các đối tượng, cấu trúc, tế bào chim chuyển gen và kit theo sáng chế có thể là protein phát huỳnh quang màu vàng. Cụ thể hơn, họ protein phát huỳnh quang màu vàng được khơi mào sau khi cấu trúc tinh thể của protein phát huỳnh quang màu xanh lục cho thấy rằng gốc threonin 203 (**Thr203**) ở gần nhóm mang màu. Sự đột biến gốc này thành tyrosin được đưa vào để làm ổn định mômen lưỡng cực trạng thái được kích thích của nhóm mang màu và dẫn đến sự dịch chuyển 20-nanomet đến bước sóng dài hơn đối với cả quang phổ kích thích và quang phổ phát xạ. Sự tinh chế thêm dẫn đến sự hiện màu của protein phát huỳnh quang màu vàng được tăng cường (**EYFP**), là một trong số các protein phát huỳnh quang sáng nhất và được sử dụng rộng rãi nhất.

Biến thể **Citrine** của protein phát huỳnh quang màu vàng khá sáng so với EYFP và cũng có thể được sử dụng trong một số phương án của sáng chế. Dẫn xuất khác, có tên **Venus**, là biến thể trưởng thành nhanh nhất và là một trong các biến thể màu vàng sáng nhất được phát triển cho đến nay. Protein rặng san hô, **ZsYellow1**, được tách dòng đầu tiên từ loài *Zoanthus* tự nhiên từ Ấn Độ Dương và Thái Bình Dương, tạo ra sự phát màu vàng thực và cũng được sử dụng theo sáng chế.

Theo một số phương án khác nữa, gen thông báo có thể được sử dụng trong các phương pháp, đối tượng thuộc loài chim chuyển gen, cấu trúc, tế bào và kit theo sáng chế có thể là protein phát huỳnh quang màu xanh lam. Các biến thể màu xanh lam và màu xanh ngọc của protein phát huỳnh quang màu xanh lục thu được từ sự cải biến trực tiếp gốc tyrosin ở vị trí 66 (**Tyr66**) trong chất huỳnh quang tự nhiên. Sự chuyển hóa axit amin này thành histidin thu được sự phát màu xanh lam có bước sóng cực đại bằng 450 nanomet, trong khi sự chuyển hóa thành tryptamin thu được đỉnh huỳnh quang chính xung quanh 480 nanomet cùng với vai đạt đỉnh xung quanh 500 nanomet.

Cụ thể hơn, trong số các protein phát huỳnh quang màu xanh ngọc cải thiện đã được đưa vào, **AmCyan1** và biến thể màu xanh ngọc tăng cường có tên **Cerulean** cho thấy triển vọng nhất. Có nguồn gốc từ rặng san hô, *Anemonia majano*, biến thể protein phát huỳnh quang AmCyan1 cũng có thể được sử dụng cho sáng chế.

Theo một số phương án cụ thể, gen thông báo có thể được sử dụng trong các phương pháp, đối tượng thuộc loài chim chuyển gen, cấu trúc, tế bào và kit theo sáng chế có thể là protein phát huỳnh quang màu đỏ. Protein phát huỳnh quang có nguồn gốc từ san hô đầu tiên được sử dụng rộng rãi có nguồn gốc từ *Discosoma striata* và thường được gọi là **DsRed**. Ngay khi trưởng thành hoàn toàn, quang phổ phát huỳnh quang của DsRed có đỉnh bằng 583 nanomet trong đó quang phổ kích thích có đỉnh chính bằng 558 nanomet và đỉnh phụ khoảng 500 nanomet. Ngoài ra, các biến thể khác có thể được sử dụng trong sáng chế bao gồm, nhưng không giới hạn ở, **DsRed2**, **DsRed-Express** và **RedStar**. Một số protein phát huỳnh quang màu đỏ khác cho thấy triển vọng đáng kể đã được phân lập sinh vật rặng san hô. Một trong các protein này có thể được sử dụng trong sáng chế, có thể là **HcRed1**, được phân lập từ *Heteractis crispa* và hiện có bán trên thị trường. HcRed1 có nguồn gốc ban đầu từ chromoprotein không phát huỳnh quang mà hấp thụ ánh sáng màu đỏ thông qua sự gây đột biến để tạo ra dime bắt buộc phát huỳnh quang yếu có sự hấp thụ cực đại ở 588 nanomet và sự phát xạ cực đại ở 618 nanomet.

Ngoài ra, theo một số phương án, gen thông báo có thể được sử dụng trong các phương pháp, đối tượng thuộc loài chim chuyển gen, cấu trúc, tế bào và kit theo sáng chế có thể là protein phát huỳnh quang bất kỳ, với điều kiện là protein này không phải là GFP, và theo một số phương án cụ thể hơn, không phải là GFP kiểu đại. Tuy nhiên, theo một số phương án, một số thể đột biến và biến thể của GFP tạo ra loài có quang phổ phát màu đỏ với cực đại kích thích và cực đại phát xạ bằng 555 và 585nm, cũng có thể được sử dụng trong sáng chế. Theo một số phương án cụ thể khác nữa, các thể đột biến hoặc biến thể này có thể chứa các gốc Serin 65 và/hoặc Asparagin 68. Theo một số phương án khác nữa, các thể đột biến hoặc biến thể này của GFP có thể chứa sự đột biến của ít nhất một trong số F46L, V163A và I167V. Theo một số phương án khác nữa, biến thể có thể chứa dạng kết hợp bất kỳ của Ser65, Asn68, F46L, V163A và I167V.

Cần hiểu rằng theo một số các phương án khác nữa, gen thông báo có thể được

sử dụng trong các phương pháp, các đối tượng, cấu trúc, tế bào chim chuyển gen và kit theo sáng chế có thể là protein phát huỳnh quang bất kỳ trong số các protein phát huỳnh quang được bộc lộ trong **Bảng 1** dưới đây. Theo một số phương án khác nữa, gen thông báo có thể được sử dụng trong các phương pháp, các đối tượng, cấu trúc, tế bào chim chuyển gen và kit theo sáng chế có thể là protein phát huỳnh quang bất kỳ trong số các protein phát huỳnh quang được bộc lộ trong **Bảng 1**, với điều kiện là protein phát huỳnh quang này không phải là GFP kiểu dại. Theo một số phương án khác nữa, gen thông báo có thể được sử dụng trong các phương pháp, đối tượng thuộc loài chim chuyển gen, cấu trúc, tế bào và kit theo sáng chế có thể là protein phát huỳnh quang bất kỳ trong số các protein phát huỳnh quang được bộc lộ trong **Bảng 1**, với điều kiện là gen thông báo này không phải là protein phát huỳnh quang bất kỳ trong số các **protein phát huỳnh quang màu xanh lục**, được bộc lộ theo Bảng 1, cụ thể hơn là protein bất kỳ trong số EGFP, Emerald, GFP siêu cuộn gấp, Azami Green, mWasabi, TagGFP, TurboGFP, AcGFP, ZsGreen, T-Sapphire.

Bảng 1: các protein phát huỳnh quang:

Protein	Sự kích thích	Sự phát xạ
(Chữ viết tắt)	Cực đại	Cực đại
	(nm)	(nm)
GFP (wt)	395/475	509
Protein phát huỳnh quang màu xanh lục		
EGFP	484	507
Emerald	487	509
GFP siêu cuộn gấp	485	510
Azami màu xanh lục	492	505
mWasabi	493	509
TagGFP	482	505
TurboGFP	482	502
AcGFP	480	505
ZsGreen	493	505

T-Sapphire	399	511
Protein phát huỳnh quang màu xanh lam		
EBFP	383	445
EBFP2	383	448
Azurite	384	450
mTagBFP	399	456
Protein phát huỳnh quang màu xanh ngọc		
ECFP	439	476
mECFP	433	475
Cerulean	433	475
mTurquoise	434	474
CyPet	435	477
AmCyan1	458	489
Midori-Ishi Cyan	472	495
TagCFP	458	480
mTFP1 (Teal)	462	492
Protein phát huỳnh quang màu vàng		
EYFP	514	527
Topaz	514	527
Venus	515	528
mCitrine	516	529
YPet	517	530
TagYFP	508	524
PhiYFP	525	537
ZsYellow1	529	539
mBanana	540	553

Protein phát huỳnh quang màu cam		
Kusabira Orange	548	559
Kusabira Orange2	551	565
mOrange	548	562
mOrange2	549	565
dTomato	554	581
dTomato-Tandem	554	581
TagRFP	555	584
TagRFP-T	555	584
DsRed	558	583
DsRed2	563	582
DsRed-Express (T1)	555	584
DsRed-Monomer	556	586
mTangerine	568	585
Protein phát huỳnh quang màu đỏ		
mRuby	558	605
mApple	568	592
mStrawberry	574	596
AsRed2	576	592
mRFP1	584	607
JRed	584	610
mCherry	587	610
HcRed1	588	618
mRaspberry	598	625
dKeima-Tandem	440	620
HcRed-Tandem	590	637

mPlum	590	649
AQ143	595	655

Như được chứng minh bất ngờ bởi các Ví dụ 1 và 2, các tính chất tự phát huỳnh quang của trứng được khám phá bởi sáng chế chỉ cho phép phát hiện phôi có các tế bào biểu hiện RFP.

Do đó, theo các phương án cụ thể hơn, gen thông báo có thể là protein phát huỳnh quang màu đỏ (RFP). Thuật ngữ "**protein phát huỳnh quang màu đỏ**" hoặc "RFP" trong bản mô tả này dùng để chỉ protein phát huỳnh quang phát ra huỳnh quang màu cam, màu đỏ, và màu đỏ xa đã được phân lập từ anthozoa (san hô) và hải quỳ cũng như các biến thể bất kỳ của chúng. Có hai loại chính của protein RFP, DsRed và Kaede. RFP như DsRed có nguồn gốc từ *Discosoma striata* và bao gồm nhưng không giới hạn ở mCherry, zFP538, mKO, mOrange, mRouge, E2-Crimson, mNeptune, TagRFP657, Keima, mKate, mStrawberry, mBanana, mHoneydew, niTangerine, mRaspberry, mPlum, mRFPmars và mRFPruby.

Mặt khác, Kaede là protein phát huỳnh quang tự nhiên được tìm thấy ở san hô đá *Trachyphyllia geoffroyi*, thay đổi không đảo ngược bước sóng phát xạ của nó từ màu xanh lục (518 nm) đến màu đỏ (582 nm) khi bức xạ ở ~400nm. Các thành viên họ Kaede có chứa ví dụ EosFP, dendFP, mcavRFP, và rfloRFP, lần lượt được tìm thấy ở san hô *Lobophyllia hemprichii*, *Dendronephthya*, *Monastrea cavernosa*, và *Ricordea florida*.

Cần hiểu rằng RFP bất kỳ trong số các RFP được mô tả trong bản mô tả này, có nguồn gốc bất kỳ đã biết trong lĩnh vực, có thể được sử dụng cho các phương pháp, động vật chuyển gen, tế bào, kit và thiết bị theo sáng chế.

Theo một số phương án cụ thể khác, RFP có thể được sử dụng bởi phương pháp theo sáng chế có thể là RFP được sử dụng trong pDsRed1-N1, mã hóa protein phát huỳnh quang màu đỏ mới [RFP; Matz, M. V., et al. (1999) Nature Biotech. 17:969–973] đã được tối ưu hóa để biểu hiện ở mức cao trong ở tế bào của động vật có vú (sự kích thích cực đại = 558 nm; sự phát xạ cực đại = 583 nm). RFP được phân lập từ họ hàng hải quỳ biển Ấn Độ-Thái Bình Dương, *Discosoma sp*; trình tự mã hóa DsRed1 chứa 144 sự thay đổi cặp bazơ câm, tương ứng với sự ưu tiên sử dụng bộ ba mã hóa của người để biểu hiện ở mức cao trong tế bào của động vật có vú.

Theo một số phương án, RFP được sử dụng bởi phương pháp và kit theo sáng chế có thể được mã hóa bởi trình tự axit nucleic chứa trình tự như được thể hiện bằng trình tự SEQ ID NO. 20. Theo một số phương án khác nữa, RFP này có thể chứa trình tự axit amin như được thể hiện bằng trình tự SEQ ID NO. 21, hoặc thể tương đồng, biến thể, thể đột biến hoặc dẫn xuất bất kỳ của chúng.

Theo một số phương án cụ thể thay thế khác, RFP được sử dụng bởi sáng chế có thể là RFP được mã hóa bởi trình tự axit nucleic như được bộc lộ bởi GenBank: AF272711.1, có trình tự axit amin như được bộc lộ bởi GenBank: AAG16224.1. Theo một số phương án cụ thể khác nữa, RFP được sử dụng bởi phương pháp và kit theo sáng chế có thể được mã hóa bởi trình tự axit nucleic chứa trình tự như được thể hiện bằng trình tự SEQ ID NO. 22. Theo một số phương án khác nữa, RFP này có thể chứa trình tự axit amin như được thể hiện bằng trình tự SEQ ID NO. 23, hoặc thể tương đồng, thể đột biến hoặc dẫn xuất bất kỳ của chúng.

Cần lưu ý rằng theo một số phương án, gen thông báo được sử dụng bởi phương pháp theo sáng chế có thể là protein phát huỳnh quang bất kỳ, với điều kiện là protein phát huỳnh quang này không phải là protein phát huỳnh quang màu xanh lục (GFP). Mặt khác, theo một số phương án cụ thể, gen thông báo được tích hợp vào nhiễm sắc thể giới tính của động vật chuyển gen được đề xuất bằng phương pháp theo sáng chế có thể là protein phát huỳnh quang bất kỳ với điều kiện là protein phát huỳnh quang này không phải là GFP.

Ngoài ra, theo một số phương án riêng và cụ thể, khi gen thông báo là RFP, cần lưu ý rằng, theo một số phương án, bước sóng kích thích của RFP nằm trong khoảng từ khoảng 500 đến 650 nm trong khi bước sóng phát xạ có thể nằm trong khoảng từ khoảng 550 đến 650 nm.

Theo một số phương án khác, việc phát hiện tín hiệu có thể phát hiện được bằng phương pháp theo sáng chế có thể bao gồm thêm bước đưa trứng này vào nguồn sáng thích hợp. Theo một số phương án khác nữa, sự tiếp xúc trứng với nguồn sáng thích hợp chứa bước sóng nằm trong khoảng từ khoảng 400 đến khoảng 650. Theo một số phương án khác nữa, ánh sáng có bước sóng bất kỳ có thể được sử dụng với điều kiện là nguồn sáng này không phải là ánh sáng cực tím (UV) (bước sóng nằm trong khoảng từ 10 đến 400 nm).

Theo một số phương án cụ thể, bước đưa trứng này vào nguồn sáng là sự kích

thích của protein phát huỳnh quang. Theo một số phương án cụ thể, bước sóng kích thích nằm trong khoảng từ khoảng 500 nm đến khoảng 650 nm. Theo một số phương án khác, bước sóng kích thích bằng khoảng 510, 515, 520, 525, 530, 535, 540, 545, 550, 555, 560, 565, 570, 575, 580, 585, 590, 595, 600, 605, 610, 615, 620, 625, 630, 635, 640 nm.

Theo một số phương án cụ thể hơn, bước sóng kích thích có thể nằm trong khoảng từ khoảng 515 đến khoảng 555. Theo một số phương án khác nữa, bước sóng có thể là 532 nm.

Theo một số phương án cụ thể và không làm giới hạn, nguồn sáng có thể được tạo ra bởi laze.

Khi được sử dụng trong bản mô tả này, thuật ngữ "**laze**" dùng để chỉ bức xạ điện từ có tần số bất kỳ được khuếch đại bằng sự phát xạ kích thích của bức xạ. Laze còn dùng để chỉ thiết bị phát bức xạ điện từ thông qua quá trình được gọi là sự phát xạ kích thích. Ánh sáng laze thường là liên tục trong không gian, có nghĩa là ánh sáng được phát ra theo chùm tia có độ phân kỳ thấp, hẹp, hoặc có thể được chuyển đổi thành chùm tia với sự trợ giúp của các thành phần quang học chẳng hạn như thấu kính.

Theo một số phương án khác, laze có thể là laze màu xanh lam, hoặc laze màu đỏ. Theo phương án cụ thể hơn, nguồn sáng là laze màu xanh lục.

Theo các phương án nhất định, laze có thể được dùng với bộ lọc.

Khi được sử dụng trong bản mô tả này, thuật ngữ "**bộ lọc**" dùng để chỉ "bộ lọc quang học" tương ứng với thiết bị truyền dẫn có chọn lọc ánh sáng có các bước sóng khác nhau, thường được thực hiện dưới dạng các thiết bị nhựa hoặc gương phẳng trong đường dẫn quang được nhuộm màu theo khối hoặc có lớp phủ nhiều.

Theo một số phương án, bộ lọc có thể là bộ lọc màu xanh lục truyền dẫn trên 500 nm, hoặc bộ lọc màu xanh lục đậm truyền dẫn trong khoảng từ 540 nm đến 580 nm, hoặc bộ lọc màu đỏ truyền dẫn trong khoảng từ 590 nm đến 650 nm, hoặc bộ lọc màu đỏ truyền dẫn trên 650 nm hoặc trên 660 nm.

Theo một số phương án cụ thể, nguồn sáng có thể là laze màu xanh lục có bước sóng bằng 532 nm. Theo một số phương án khác, laze này có thể được cung cấp với bộ lọc màu đỏ truyền dẫn trên 650 nm.

Dòng mầm gà phát triển từ quần thể nhỏ của tế bào mầm nguyên thủy, di chuyển đến cung sinh dục từ vị trí ngoài tuyến sinh dục, trong khi trải qua các pha di

chuyển chủ động, cũng như là sự tuần hoàn thụ động trong dòng máu. PGC nằm ở trung tâm của ngoại phôi bì của trứng mới đẻ, chưa được ủ, giai đoạn phát triển được gọi là giai đoạn X. Trong quá trình ủ PGC di chuyển về phía trước và tích lũy ở vòng cung mầm của phôi giai đoạn 10HH (khoảng 33-38 giờ ủ), được cho là vị trí chính cho sự xâm lấn. Sau đó, PGC có thể được tìm thấy trong hệ tuần hoàn từ giai đoạn 12HH đến 17HH (khoảng từ 45 đến 64 giờ ủ), đạt nồng độ đỉnh ở giai đoạn 14HH (khoảng 50-53 giờ ủ). PGC rời khỏi hệ tuần hoàn ở vị trí gần kề với mầm tuyến sinh dục ở trung bì trung gian, tại đó chúng có thể được tìm thấy ngay từ giai đoạn 15HH (55-55 giờ ủ). PGC tiến đến cung sinh dục bằng cách di chuyển chủ động dọc theo màng treo ruột ở lưng, và định cư ở cả hai tuyến sinh dục, tại đó sau này chúng phân biệt thành tinh nguyên bào hoặc noãn nguyên bào.

Phôi nằm ở trên đỉnh túi noãn hoàng. Noãn hoàng xoay tự do trong trứng, và phôi luôn hướng về phía trên noãn hoàng. Trong quá trình ủ, đầu tù của trứng (nơi túi khí nằm) hướng lên trên, và trứng thỉnh thoảng được xoay 90°. Khi phôi phát triển, sự hình thành các mô ngoài phôi co rút phôi ra khỏi vỏ trứng, và phôi khó tiếp cận hơn.

Do đó, theo một số phương án cụ thể, trứng được cho tiếp xúc với nguồn sáng này. Theo một số phương án khác nữa, trứng được đặt ở vị trí tạo thuận lợi cho việc tiếp xúc phôi ở giai đoạn bất kỳ với nguồn sáng. Theo một số phương án khác nữa, vùng chứa mặt trên của noãn hoàng trứng ở giai đoạn X của trứng được kích thích bằng nguồn sáng.

Theo một số phương án khác, bước đưa trứng này vào nguồn sáng được cung cấp bằng hệ thống, dụng cụ hoặc thiết bị có thể chứa nguồn laze, bộ đỡ trứng, thấu kính, bộ lọc, bộ đỡ thiết bị phát hiện và thiết bị phát hiện.

Khi được sử dụng trong bản mô tả này, thuật ngữ "**thiết bị phát hiện**" dùng để chỉ loại thiết bị bất kỳ phát hiện và/hoặc đo ánh sáng. Theo một số phương án, cần lưu ý rằng tín hiệu có thể phát hiện được, cụ thể là tín hiệu huỳnh quang có thể được phát hiện bằng cách sử dụng phương tiện huỳnh quang thích hợp. Theo một số phương án, tín hiệu có thể phát hiện được được tạo ra bởi gen thông báo RFP có thể được phát hiện bằng dụng cụ nhạy cảm ánh sáng chẳng hạn như kính hiển vi quang học cải biến hoặc thiết bị tích điện kép (CCD), thiết bị phát hiện photon có độ nhạy cao.

Theo một số phương án khác nữa, phương pháp theo sáng chế có thể được thực hiện bằng cách sử dụng thiết bị, dụng cụ hoặc hệ thống bất kỳ để cung cấp ít nhất một

nguồn sáng, ít nhất một thiết bị phát hiện, ít nhất một bộ lọc và ít nhất một tay giữ đặt trứng ở vị trí thích hợp tạo thuận lợi cho sự tiếp xúc tế bào biểu hiện gen thông báo theo sáng chế với nguồn sáng này. Theo một số phương án cụ thể và không làm giới hạn, phương pháp theo sáng chế có thể sử dụng thiết bị như được minh họa trên Fig.1. Theo một số phương án khác nữa, khoảng cách giữa bệ đỡ trứng và thấu kính và/hoặc thấu kính và thiết bị phát hiện có thể nằm trong khoảng từ khoảng 1 đến 100 cm, cụ thể là, 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 và 100 cm, cụ thể là trong khoảng từ khoảng 5 đến 25 cm, cụ thể hơn là khoảng 20 cm.

Do đó, cần hiểu rằng theo một số phương án, phương pháp theo sáng chế có thể bao gồm thêm bước cung cấp hệ thống hoặc thiết bị, ví dụ, như được mô tả trong bản mô tả này, được điều chỉnh để xác định sự tồn tại của tín hiệu có thể phát hiện được trong trứng được kiểm tra. Cụ thể là dụng cụ hoặc thiết bị có chứa nguồn laze, bệ đỡ trứng, thấu kính, bộ lọc, bệ đỡ thiết bị phát hiện và thiết bị phát hiện. Như được biểu thị ở trên, theo một số phương án đối với thiết bị này, có thể sử dụng nguồn laze màu xanh lục có bước sóng kích thích bằng khoảng 532nm. Theo một số phương án khác nữa, có thể được sử dụng bộ lọc, cụ thể là bộ lọc màu đỏ, truyền dẫn trên 650nm.

Theo các phương án khác nữa, ít nhất một đối tượng hoặc động vật thuộc loài chim chuyển gen được tạo ra bởi phương pháp theo sáng chế có thể là đối tượng hoặc động vật là chim cái. Theo phương án cụ thể hơn, khi mà ít nhất một gen thông báo được tích hợp vào ít nhất một vị trí của nhiễm sắc thể con cái Z, sự phát hiện tín hiệu có thể phát hiện chỉ ra rằng phôi trong trứng chưa nở là con đực. Cụ thể hơn, phôi chứa nhiễm sắc thể giới tính Z không được đánh dấu của bố và còn chứa nhiễm sắc thể giới tính Z được đánh dấu của mẹ, chứa hai nhiễm sắc thể Z và do đó là con đực.

Theo một số phương án khác nữa, ít nhất một đối tượng hoặc động vật thuộc loài chim chuyển gen được tạo ra bởi phương pháp theo sáng chế có thể là đối tượng hoặc động vật là chim cái. Theo một số phương án cụ thể, khi mà ít nhất một gen thông báo được tích hợp vào ít nhất một vị trí của nhiễm sắc thể con cái W, sự phát hiện tín hiệu có thể phát hiện được, biểu thị rằng phôi trong trứng chưa nở là con cái. Cụ thể hơn, phôi chứa nhiễm sắc thể z không được đánh dấu của bố và nhiễm sắc thể W được đánh dấu của mẹ là phôi dị giao tử chứa các nhiễm sắc thể WZ là con cái.

Theo một số phương án khác nữa, động vật chuyển gen được tạo ra bởi phương pháp theo sáng chế có thể là đối tượng con đực có gen thông báo được tích hợp vào

nhiễm sắc thể Z của chúng. Trong trường hợp này, tín hiệu có thể phát hiện được được xác định trong trứng được thụ tinh bởi con đực chuyển gen này hoặc tinh dịch bất kỳ của chúng, biểu thị rằng phôi mang nhiễm sắc thể Z của bố có chứa gen thông báo chuyển gen, và do đó là con đực. Theo các phương án khác nữa, sự phát hiện tín hiệu có thể phát hiện được trong trứng được đẻ bởi chim cái chuyển gen được thụ tinh bởi chim đực chuyển gen, cả hai mang gen thông báo theo sáng chế được tích hợp vào nhiễm sắc thể Z của chúng, có thể biểu thị trong trường hợp tín hiệu mạnh rằng phôi mang hai bản sao của gen thông báo được tích hợp vào các nhiễm sắc thể Z cái và đực của chúng. Trong trường hợp tín hiệu ít mạnh hơn, ví dụ, tín hiệu có cường độ giảm bằng khoảng 50%. Trứng có thể được xác định là con cái.

Như được thể hiện trước đó trong bản mô tả này, phương pháp theo sáng chế bao gồm bước cung cấp động vật thuộc loài chim chuyển gen. Sự tạo ra động vật thuộc loài chim chuyển gen, đòi hỏi việc sử dụng phương pháp thiết kế di truyền có thể sử dụng hợp chất hoặc thành phần chỉnh sửa gen cụ thể. Các ví dụ không làm giới hạn về thành phần chỉnh sửa gen có thể chứa trong một số phương án, việc sử dụng nucleaza.

Theo một số phương án cụ thể, nucleaza thích hợp trong các phương pháp theo sáng chế có thể là nucleaza protein liên kết ADN được dẫn đường bằng ARN.

Khi được sử dụng trong bản mô tả này, **nucleaza protein liên kết ADN được dẫn đường bằng ARN** là nucleaza được dẫn đường đến vị trí phân cắt của nó (hoặc theo cách khác là, vị trí cho hoạt tính thay thế khác bất kỳ), bởi phân tử ARN. Phân tử ARN này được gọi là ARN dẫn đường.

Do đó, theo một số phương án cụ thể hơn nữa, ít nhất một gen thông báo có thể được tích hợp vào nhiễm sắc thể giới tính của đối tượng hoặc động vật thuộc loài chim chuyển gen được tạo ra bởi phương pháp theo sáng chế bằng cách sử dụng ít nhất một nucleaza được thiết kế theo chương trình (PEN). Khi được sử dụng trong bản mô tả này, thuật ngữ "**nucleaza được thiết kế theo chương trình (PEN)**", dùng để chỉ enzym tổng hợp mà cắt trình tự ADN cụ thể, có nguồn gốc từ nucleaza có trong tự nhiên tham gia vào việc sửa chữa ADN của các tổn thương ADN sợi kép và có khả năng chỉnh sửa hệ gen trực tiếp.

Theo một số phương án cụ thể hơn, PEN được sử dụng bởi phương pháp theo sáng chế có thể là hệ thống bất kỳ trong hệ thống đoạn lặp lại đối xứng ngăn cách đều thành cụm (CRISPR) loại 2 hoặc loại 1.

Cụ thể hơn, hệ thống đoạn lặp lại đối xứng ngắn cách đều thành cụm (Clustered Regularly Interspaced Short Palindromic Repeats - CRISPR) kiểu II là hệ thống miễn dịch vi khuẩn đã được cải biến để thiết kế hệ gen.

Tuy nhiên, cần hiểu rằng các phương pháp thiết kế hệ gen khác, như nucleaza ngón tay kềm (ZFN) hoặc nucleaza tác động như chất hoạt hóa sự phiên mã (TALEN) mà chuyển tiếp khi sử dụng nucleaza protein liên kết ADN tùy biến mà đòi hỏi sự thiết kế và tạo ra cặp nucleaza đặc hiệu đối với mọi đích hệ gen cũng có thể thích hợp trong bản mô tả này.

Hệ thống CRISPR-Cas chia làm hai loại. Hệ thống loại 1 sử dụng phức hợp của nhiều protein Cas để làm thoái biến axit nucleic ngoại lai. Hệ thống loại 2 sử dụng protein Cas lớn đơn lẻ cho cùng mục đích. Cụ thể hơn, loại 1 có thể được chia thành các kiểu I, III, và IV và loại 2 có thể được chia thành các kiểu II, V, và VI.

Khi được sử dụng trong bản mô tả này, mảng CRISPR còn được gọi là SPIDR (Spacer Interspersed Direct Repeats) tạo thành họ locus ADN được mô tả gần đây thường đặc hiệu với các loài vi khuẩn cụ thể. Mảng CRISPR là loại khác biệt của đoạn lặp trình tự ngắn phân cách (SSR) được nhận dạng đầu tiên ở *E. coli*. Trong các năm tiếp theo, các mảng CRISPR tương tự được tìm thấy ở *Mycobacterium tuberculosis*, *Haloferax mediterranei*, *Methanocaldococcus jannaschii*, *Thermotoga maritima* và các vi khuẩn và vi khuẩn cổ khác. Cần hiểu rằng sáng chế dự tính việc sử dụng hệ thống CRISPR bất kỳ trong số các hệ thống CRISPR đã biết, và cụ thể là hệ thống CRISPR trong số các hệ thống CRISPR được bộc lộ trong bản mô tả này. Hệ thống CRISPR-Cas đã tiến hóa ở sinh vật nhân sơ để bảo vệ chống lại sự tấn công thể thực khuẩn và sự sao chép plasmit không mong muốn bằng cách hướng đích ADN hoặc ARN ngoại lai. Hệ thống CRISPR-Cas, hướng đích các phân tử ADN dựa trên trình tự ADN tương đồng ngắn, được gọi là đoạn đệm tồn tại giữa các đoạn lặp. Các đoạn đệm này dẫn đường các protein kết hợp CRISPR (Cas) để ghép nối (và/hoặc bổ sung) các trình tự ở trong ADN ngoại lai, được gọi là đoạn tiền đệm, mà bị phân cắt sau đó. Đoạn đệm có thể được thiết kế hợp lý để hướng đích trình tự ADN bất kỳ. Ngoài ra, thành phần nhận diện này có thể được thiết kế riêng rẽ để nhận diện và hướng đích đích mong muốn bất kỳ. Đối với hệ thống CRISPR, như được nhận biết bởi người có hiểu biết trung bình trong lĩnh vực, cấu trúc của locus CRISPR có trong tự nhiên bao gồm một số trình tự lặp ngắn thường được gọi là "đoạn lặp". Đoạn lặp có mặt trong

cụm và thường được cách quãng đều bằng cách xen kẽ đơn nhất các trình tự được gọi là "đoạn đệm." Thông thường, các đoạn lặp CRISPR thay đổi chiều dài trong khoảng từ khoảng 24 đến 47 cặp bazơ (bp) và đối xứng một phần. Đoạn đệm nằm ở giữa hai đoạn lặp và thông thường mỗi đoạn đệm có các trình tự đơn nhất có chiều dài trong khoảng từ nhỏ hơn hoặc bằng 20 đến lớn hơn hoặc bằng 72 bp. Do đó, Theo các phương án nhất định, đoạn đệm CRISPR được sử dụng trong trình tự mã hóa ít nhất một gARN của phương pháp và kit theo sáng chế có thể chứa trong khoảng từ 10 đến 75 nucleotit (nt) mỗi đoạn. Cụ thể hơn, khoảng 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75 hoặc lớn hơn. Theo một số phương án cụ thể, đoạn đệm chứa trong khoảng từ khoảng 20 đến 25 nucleotit, cụ thể hơn là khoảng 20 nucleobazơ.

Ngoài ít nhất một đoạn lặp và ít nhất một đoạn đệm, locus CRISPR cũng bao gồm trình tự dẫn đầu và tùy ý, trình tự mã hóa ít nhất một tracrARN. Trình tự dẫn đầu thông thường là trình tự giàu AT lên đến 550 bp tiếp giáp trực tiếp đầu 5' của đoạn lặp đầu tiên.

Theo một số phương án cụ thể, PEN được sử dụng bởi phương pháp theo sáng chế có thể là hệ thống CRISPR loại 2. Theo một số phương án cụ thể khác nữa, hệ thống loại 2 có thể là hệ thống CRISPR kiểu II.

Cụ thể hơn, ba kiểu chính của hệ thống CRISPR–Cas được mô tả: kiểu I, kiểu II và kiểu III.

Hệ thống CRISPR–Cas kiểu II bao gồm hệ thống kiểu-'HNH' (giống-*Streptococcus*; còn được gọi là kiểu phụ Nmeni, đối với *Neisseria meningitidis serogroup A str. Z2491*, hoặc CASS4), trong đó Cas9, protein rất lớn, đơn lẻ, dường như là đủ để tạo ra crARN và phân cắt ADN đích, ngoài Cas1 và Cas2 thường gặp. Cas9 chứa ít nhất là hai miền nucleaza, miền nucleaza giống-RuvC ở gần đầu tận cùng amino và miền nucleaza HNH (hoặc giống-McrA) ở chính giữa của protein, nhưng chức năng của các miền này vẫn còn cần được làm sáng tỏ. Tuy nhiên, vì miền nucleaza HNH dư thừa các enzym giới hạn và có hoạt tính endonucleaza chịu trách nhiệm cho việc phân cắt đích.

Hệ thống kiểu II phân cắt tiền-crARN thông qua cơ chế khác thường bao gồm

sự hình thành cặp đôi giữa tracrARN và một phần của đoạn lặp trong tiền-crARN; sự phân cắt đầu tiên trong con đường xử lý tiền-crARN sau đó xảy ra trong vùng lặp này. Ngoài ra, cần lưu ý rằng hệ thống kiểu II chứa ít nhất một trong số các gen *cas9*, *cas1*, *cas2*, *csn2*, và *cas4*. Cần hiểu rõ rằng hệ thống CRISPR-Cas kiểu II bất kỳ có thể thích hợp theo sáng chế, cụ thể là hệ thống bất kỳ trong số các kiểu II-A hoặc B.

Do đó, theo một số phương án thay thế và bổ sung, ít nhất một gen *cas* được sử dụng trong phương pháp và kit theo sáng chế có thể là ít nhất một gen *cas* của hệ thống CRISPR kiểu II (kiểu II-A hoặc kiểu II-B). Theo các phương án cụ thể hơn, ít nhất một gen *cas* của hệ thống CRISPR kiểu II được sử dụng bởi phương pháp và kit theo sáng chế có thể là gen *cas9*. Cần hiểu rõ rằng hệ thống này có thể còn chứa ít nhất một trong số các gen *cas1*, *cas2*, *csn2* và *cas4*.

Sự phân cắt ADN sợi kép (dsADN) bởi Cas9 là dấu hiệu nhận biết của hệ miễn dịch "**CRISPR-Cas kiểu II**". Protein Cas9 kết hợp CRISPR là ADN endonucleaza được dẫn đường bởi ARN sử dụng sự hỗ trợ ARN:ADN để nhận biết vị trí đích trình tự-đặc hiệu sự phân cắt ADN sợi kép (dsADN), tạo ra sự đứt gãy sợi kép (DSB) cần cho HDR dẫn đến sự tích hợp gen thông báo vào trình tự đích đặc hiệu, ví dụ, đích đặc hiệu ở trong các nhiễm sắc thể giới tính loài chim W và Z. Trình tự ADN được hướng đích được đặc hiệu bởi mảng CRISPR, là chuỗi các đoạn đệm nằm trong khoảng từ khoảng 30 đến 40 bp được phân tách bởi các đoạn lặp đôi xứng xứng. Mảng được phiên mã dưới dạng tiền-crARN và được xử lý thành các crARN ngắn hơn kết hợp với phức hợp protein Cas để hướng đích các trình tự ADN hỗ trợ được biết là đoạn tiền đệm. Các đoạn tiền đệm đích này cũng phải có trình tự lân cận bổ sung được biết là motif kề đoạn tiền đệm (PAM) cần thiết cho sự nhận diện đích. Sau khi liên kết, phức hợp protein Cas đóng vai trò là ADN endonucleaza để cắt cả hai sợi ở đích và sự thoái hóa ADN sau đó xảy ra thông qua hoạt tính exonucleaza.

Khi được sử dụng trong bản mô tả này, hệ thống CRISPR kiểu II đòi hỏi sự bao gồm của hai thành phần thiết yếu: ARN "dẫn đường" (gARN) và endonucleaza kết hợp CRISPR không đặc hiệu (Cas9). gARN là ARN tổng hợp ngắn được tạo thành bởi trình tự "giá đỡ" cần thiết cho trình tự "đoạn đệm" hoặc "hướng đích" dài khoảng 20 nucleotit và liên kết Cas9 và xác định đích hệ gen cần được cải biến. Do đó, người ta có thể thay đổi đích hệ gen của Cas9 đơn giản là bằng cách thay đổi trình tự hướng đích có mặt trong gARN. Khi được sử dụng trong bản mô tả này, ARN dẫn đường

(gARN) để chỉ sự dung hợp tổng hợp của crARN và tracrARN vi khuẩn nội sinh, cung cấp cả tính đặc hiệu hướng đích và khả năng đỡ/liên kết đối với Cas9 nucleaza. Nó còn được gọi là “ARN dẫn đường đơn lẻ” hoặc “sgARN”. CRISPR ban đầu được sử dụng để “làm bất hoạt” gen đích ở các loại tế bào và sinh vật khác nhau, nhưng sự cải biến cho enzym Cas9 đã kéo dài việc áp dụng CRISPR để “gắn” gen đích, hoạt hóa hoặc ức chế có chọn lọc gen đích, tinh chế các vùng đặc hiệu của ADN, và thậm chí là chụp ảnh ADN ở tế bào sống bằng cách sử dụng kính hiển vi huỳnh quang. Hơn nữa, sự dễ dàng tạo ra gARN làm cho công nghệ CRISPR là một trong số các công nghệ chỉnh sửa hệ gen có khả năng mở rộng quy mô nhất và gần đây đã được sử dụng để sàng lọc toàn bộ hệ gen.

Đích ở trong hệ gen được chỉnh sửa, cụ thể là locus đích đặc hiệu ở trong nhiễm sắc thể giới tính Z hoặc W, tại đó gen thông báo theo sáng chế được tích hợp, có mặt ở ngay phía trước Motif gần kề đoạn tiền đệm (PAM).

Trình tự PAM tuyệt đối cần thiết cho sự liên kết hướng đích và trình tự chính xác phụ thuộc vào các loài có Cas9 (5' NGG 3' đối với *Streptococcus pyogenes* Cas9). Theo các phương án nhất định, Cas9 từ *S. pyogenes* được sử dụng trong phương pháp và kit theo sáng chế. Tuy nhiên, cần hiểu rằng Cas9 đã biết bất kỳ có thể được sử dụng. Các ví dụ không làm giới hạn đối với Cas9 hữu dụng trong phân bộ tộc này bao gồm nhưng không giới hạn ở *Streptococcus pyogenes* (SP), cũng được chỉ thị trong bản mô tả này dưới dạng SpCas9, *Staphylococcus aureus* (SA), cũng được chỉ thị trong bản mô tả này dưới dạng SaCas9, *Neisseria meningitidis* (NM), cũng được chỉ thị trong bản mô tả này dưới dạng NmCas9, *Streptococcus thermophilus* (ST), cũng được chỉ thị trong bản mô tả này dưới dạng StCas9 và *Treponema denticola* (TD), cũng được chỉ thị trong bản mô tả này dưới dạng TdCas9. Theo một số phương án cụ thể, Cas9 của *Streptococcus pyogenes* M1 GAS, cụ thể là Cas9 có protein id: AAK33936.1, có thể được sử dụng trong phương pháp và kit theo sáng chế. Theo một số phương án, protein Cas9 có thể được mã hóa bởi trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 24. Theo các phương án cụ thể và không làm giới hạn khác, protein Cas9 có thể chứa trình tự axit amin như được thể hiện bằng trình tự SEQ ID NO. 25, hoặc các dẫn xuất, thể đột biến hoặc biến thể bất kỳ của chúng. Ngay khi được biểu hiện, protein Cas9 và gARN, tạo ra phức hợp riboprotein thông qua sự tương tác giữa miền “giá đỡ” gARN và đường rãnh tích điện dương lộ bề mặt trên

Cas9. Cas9 trải qua sự thay đổi cấu hình khi liên kết gARN mà dịch chuyển phân tử từ cấu hình không liên kết ADN, không có hoạt tính, thành cấu hình liên kết ADN có hoạt tính. Điều quan trọng là trình tự “đoạn đệm” của gARN vẫn tự do để tương tác với ADN đích. Phức hợp Cas9-gARN liên kết trình tự hệ gen bất kỳ với PAM, nhưng mức độ mà tại đó đoạn đệm gARN ghép nối ADN đích xác định liệu Cas9 có cắt hay không. Ngay khi phức hợp Cas9-gARN liên kết đích ADN giả định, trình tự “mầm” ở đầu 3' của trình tự hướng đích gARN bắt đầu ủ với ADN đích. Nếu trình tự ADN mầm và trình tự ADN đích khớp lại, gARN tiếp tục ủ với ADN đích theo hướng từ đầu 3' đến đầu 5'.

Cas9 chỉ phân cắt đích nếu đủ độ đồng nhất tồn tại giữa đoạn đệm gARN và trình tự đích. Ngoài ra, Cas9 nucleaza có hai miền endonucleaza chức năng: RuvC và HNH. Cas9 trải qua sự thay đổi cấu hình thứ hai khi liên kết đích mà định vị miền nucleaza để phân cắt sợi đối ngược của ADN đích. Kết quả cuối cùng của sự phân cắt ADN qua trung gian Cas9 là sự đứt gãy sợi kép (DSB) ở trong ADN đích xảy ra trong khoảng từ khoảng 3 đến 4 nucleotit ngược dòng của trình tự PAM.

Sau đó, DSB thu được có thể được sửa chữa bằng một trong hai con đường sửa chữa thông thường, con đường nối đầu không-tương đồng (NHEJ) hiệu quả nhưng dễ bị lỗi và con đường sửa chữa theo hướng tương đồng (HDR) ít hiệu quả hơn nhưng có độ chính xác cao. Theo một số phương án, sự chèn đoạn dẫn đến sự tích hợp đặc hiệu gen thông báo theo sáng chế với locus đích đặc hiệu trong nhiễm sắc thể giới tính W hoặc Z, là kết quả của việc sửa chữa DSB gây ra bởi Cas9. Theo một số phương án cụ thể, gen thông báo theo sáng chế được tích hợp, hoặc được gắn locus đích bằng HDR.

Khi được sử dụng trong bản mô tả này, thuật ngữ "**sửa chữa theo hướng tương đồng (HDR)**", để chỉ cơ chế ở tế bào để sửa chữa các tổn thương ADN sợi kép. Dạng thông thường nhất của HDR là dạng tái tổ hợp tương đồng. Cơ chế sửa chữa HDR chỉ có thể được tế bào sử dụng khi có mảnh tương đồng của ADN có mặt trong nhân, chủ yếu là ở pha G2 và S của chu trình tế bào. Khi mảnh ADN tương đồng không có mặt, quá trình khác được gọi là nối đầu không-tương đồng (NHEJ) có thể diễn ra thay thế. Các chiến lược nucleaza được thiết kế theo chương trình (PEN) để chỉnh sửa hệ gen, dựa trên sự hoạt hóa tế bào của cơ chế HDR sau khi phân cắt ADN sợi kép đặc hiệu.

Như đã bàn luận trước đây, Cas9 tạo ra sự đứt gãy sợi kép (double strand break:

DSB) thông qua hoạt tính được kết hợp của hai miền nucleaza, RuvC và HNH. Các gốc axit amin chính xác ở trong mỗi miền nucleaza quan trọng đối với hoạt tính endonucleaza là đã biết (D10A đối với HNH và H840A đối với RuvC trong *S. pyogenes* Cas9) và phiên bản cải biến của enzym Cas9 chỉ chứa một miền xúc tác có hoạt tính (được gọi là “Cas9 nickaza”) đã được tạo ra. Cas9 nickaza vẫn liên kết ADN dựa trên tính đặc hiệu gARN, nhưng nickaza chỉ có khả năng cắt một trong các sợi ADN, dẫn đến “điểm đứt”, hoặc sự đứt gãy sợi đơn, thay vì DSB. Các điểm đứt ADN nhanh chóng được sửa chữa bằng HDR (sửa chữa theo hướng tương đồng) bằng cách sử dụng sợi ADN bổ trợ nguyên vẹn làm khuôn. Do đó, cần có hai sợi đối diện hướng đích nickaza để tạo ra DSB trong ADN đích (thường được gọi là hệ thống CRISPR “điểm đứt kép” hoặc “nickaza kép”). Yêu cầu này làm tăng đột ngột tính đặc hiệu đích, bởi vì không chắc là hai điểm đứt ngoài-đích sẽ được tạo ra ở độ gần đủ để gây ra DSB. Do đó, cần hiểu rằng, sáng chế còn bao gồm việc sử dụng phương pháp nickaza kép để tạo ra DSB do điểm đứt kép gây ra để tăng tính đặc hiệu và giảm tác động ngoài-đích.

Do đó, Theo các phương án nhất định, ít nhất một gen thông báo có thể được tích hợp vào nhiễm sắc thể giới tính của đối tượng, cụ thể là động vật, thuộc loài chim chuyển gen bằng cách sửa chữa theo hướng tương đồng (HDR) qua trung gian ít nhất một hệ thống CRISPR/endonucleaza 9 kết hợp CRISPR (Cas9).

Như được chỉ ra ở trên, hệ thống CRISPR kiểu II chứa ít nhất hai thành phần nucleaza, cụ thể là Cas9 và ARN dẫn đường. Do đó, để kết hợp và tích hợp gen thông báo được sử dụng bởi sáng chế, vào nhiễm sắc thể giới tính của đối tượng thuộc loài chim chuyển gen, ngoài Cas9 hoặc trình tự axit nucleic bất kỳ mã hóa Cas9, sáng chế còn đề xuất ARN dẫn đường hoặc trình tự axit nucleic bất kỳ mã hóa gARN này hướng đích Cas9 đến vị trí đích trong nhiễm sắc thể giới tính cụ thể. Theo một số phương án khác, gARN của kit theo sáng chế có thể chứa ít nhất một ARN CRISPR (crARN) và ít nhất một crARN hoạt hóa-trans (tracrARN).

Theo một số phương án thay thế, kit theo sáng chế có thể chứa trình tự axit nucleic mã hóa ít nhất một gARN. Trình tự axit nucleic này có thể chứa mảng CRISPR chứa ít nhất một trình tự đoạn đệm hướng đích và do đó giống với ít nhất một đoạn tiền đệm trong trình tự ADN của hệ gen đích. Cần lưu ý rằng trình tự axit nucleic còn chứa trình tự mã hóa ít nhất một tracrARN.

Ngoài ra, theo một số phương án, ít nhất một gen thông báo có thể được tích hợp vào nhiễm sắc thể giới tính của đối tượng, cụ thể là động vật, thuộc loài chim chuyển gen bằng cách tiếp xúc hoặc đồng-chuyển nhiễm ít nhất một tế bào của đối tượng, hoặc động vật, thuộc loài chim hoặc ít nhất một tế bào được đưa vào đối tượng, hoặc động vật, thuộc loài chim, có: (a) các thành phần của hệ thống CRISPR Cas9, cụ thể là ít nhất một protein Cas9 hoặc ít nhất một trình tự axit nucleic thứ nhất chứa ít nhất một trình tự axit nucleic mã hóa ít nhất một protein Cas9 này; và ít nhất một ARN dẫn đường (gARN) hướng đích đoạn tiền đệm trong nhiễm sắc thể giới tính Z hoặc W, hoặc ít nhất một trình tự axit nucleic mã hóa ít nhất một gARN này; và (b) ít nhất một trình tự axit nucleic thứ hai chứa ít nhất một gen thông báo. Cần lưu ý rằng theo một số phương án, trong trường hợp các thành phần của hệ thống CRISPR Cas9 trong (a) được cung cấp ở dạng trình tự axit nucleic, chúng có thể được cung cấp trong một phân tử axit nucleic có chứa các thành phần khác nhau (được đề cập đến trong bản mô tả này dưới dạng "trình tự axit nucleic thứ nhất", hoặc trong hai hoặc nhiều phân tử axit nucleic, mỗi phân tử có chứa một trong số các thành phần được biểu thị trên đây, cụ thể là các trình tự mã hóa gARN và Cas9.

Do đó, để tạo ra động vật thuộc loài chim chuyển gen được sử dụng bởi phương pháp theo sáng chế, cần cung cấp ít nhất là hai phân tử axit nucleic.

Ngoài ra, "**gARN**" hoặc "**ARN hướng đích**" là ARN, khi được phiên mã từ một phần của hệ thống CRISPR mã hóa nó, chứa ít nhất một phân đoạn của trình tự ARN giống với (ngoại trừ việc thay thế T cho U trong trường hợp của ARN) hoặc hỗ trợ cho (và do đó "hướng đích") trình tự ADN ở ADN của hệ gen đích, được đề cập đến trong bản mô tả này dưới dạng đoạn tiền đệm. Hệ thống CRISPR theo phần bộc lộ sáng chế có thể tùy ý mã hóa nhiều hơn một ARN hướng đích, và ARN hướng đích được định hướng đến một hoặc nhiều trình tự đích ở ADN hệ gen. Khi được sử dụng trong bản mô tả này, thuật ngữ "**đoạn tiền đệm**" dùng để chỉ trình tự đích trong nhiễm sắc thể đích. Các đoạn tiền đệm này chứa trình tự axit nucleic có độ bổ sung đủ cho ARN hướng đích được mã hóa bởi đoạn đệm CRISPR được chứa trong trình tự axit nucleic mã hóa gARN của phương pháp và kit theo sáng chế.

Phương pháp theo sáng chế, cũng như kit được mô tả dưới đây, cung cấp trình tự axit nucleic. Khi được sử dụng trong bản mô tả này, thuật ngữ "**axit nucleic hoặc phân tử axit nucleic**" có thể thay thế lẫn nhau với thuật ngữ "**(các) polynucleotit**" và

thường dùng để chỉ polyribonucleotit hoặc poly-deoxyribonucleotit bất kỳ, có thể là ARN hoặc ADN không cải biến hoặc ARN hoặc ADN cải biến hoặc dạng kết hợp bất kỳ của chúng. "Axit nucleic" bao gồm, mà không giới hạn ở, axit nucleic sợi đơn và sợi kép. Khi được sử dụng trong bản mô tả này, thuật ngữ "(các) axit nucleic" cũng bao gồm ADN hoặc ARN như được mô tả trên chứa một hoặc nhiều bazơ cải biến. Khi được sử dụng trong bản mô tả này, thuật ngữ "**oligonucleotit**" được xác định là phân tử được tạo bởi hai hoặc nhiều deoxyribonucleotit và/hoặc ribonucleotit, và tốt hơn là nhiều hơn ba. Kích thước chính xác của nó phụ thuộc vào nhiều yếu tố mà đến lượt nó, phụ thuộc vào chức năng và việc sử dụng cuối cùng của oligonucleotit. Oligonucleotit có thể có chiều dài nằm trong khoảng từ khoảng 8 đến khoảng 1.000 nucleotit. Cụ thể hơn, các/phân tử oligonucleotit được sử dụng bởi kit theo sáng chế có thể có chiều dài bất kỳ trong số 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 bazơ hoặc nhiều hơn.

Phân tử axit nucleic có thể được tạo bởi các monome là nucleotit có trong tự nhiên (chẳng hạn như ADN và ARN), hoặc chất tương tự của nucleotit có trong tự nhiên (ví dụ, dạng alpha-đồng phân đối ảnh của nucleotit có trong tự nhiên), hoặc nucleotit cải biến hoặc dạng kết hợp bất kỳ của chúng. Trong bản mô tả này, thuật ngữ này cũng bao hàm cADN, tức là ADN bổ trợ hoặc bản sao được tạo ra từ khuôn ARN bởi tác dụng của enzym phiên mã ngược (ADN polymeraza phụ thuộc ARN).

Về việc này, "**polynucleotit được phân lập**" là phân tử axit nucleic được phân tách từ hệ gen của sinh vật. Ví dụ, phân tử ADN mã hóa gen thông báo được sử dụng trong phương pháp và kit theo sáng chế hoặc dẫn xuất hoặc thể tương đồng bất kỳ của chúng, cũng như trình tự mã hóa CRISPR/Cas9 và gARN của phương pháp và kit theo sáng chế, đã được phân tách từ ADN hệ gen của tế bào là phân tử ADN được phân lập. Một ví dụ khác về phân tử axit nucleic được phân lập là phân tử axit nucleic được tổng hợp hóa học mà không được tích hợp trong hệ gen của sinh vật. Phân tử axit nucleic đã được phân lập từ loài cụ thể nhỏ hơn so với phân tử ADN hoàn chỉnh của nhiễm sắc thể từ loài đó.

Theo một số phương án, trình tự axit nucleic được sử dụng bởi phương pháp và kit (được mô tả dưới đây) theo sáng chế, cụ thể là trình tự axit nucleic chứa trình tự mã hóa Cas9 và gARN, hoặc theo cách khác là gen thông báo theo sáng chế, có thể được

cung cấp được tạo cấu trúc trong vector, cắt xét hoặc các phương tiện khác bất kỳ. Do đó sáng chế còn đề cập đến cấu trúc ADN tái tổ hợp chứa polynucleotit theo sáng chế, và tùy ý, các thành phần bổ sung khác chẳng hạn như gen khởi đầu, thành phần điều hòa và thành phần kiểm soát, tín hiệu dịch mã, tín hiệu biểu hiện và các tín hiệu khác, được liên kết khi hoạt động với trình tự axit nucleic theo sáng chế. Ví dụ không làm giới hạn đối với vector được đề xuất bởi sáng chế có thể là vector bất kỳ chứa trình tự axit nucleic mã hóa Cas9 và gARN đặc hiệu cần cho sự tích hợp gen thông báo vào nhiễm sắc thể giới tính đích. Theo một số phương án cụ thể, sáng chế đề xuất các cấu trúc ADN hoặc vector chứa trình tự axit nucleic mã hóa Cas9 như được thể hiện bằng trình tự SEQ ID NO. 25, và gARN7 như được thể hiện bằng trình tự SEQ ID NO. 26, hướng đích gen thông báo đến nhiễm sắc thể giới tính Z. Theo một số phương án khác nữa, sáng chế còn đề xuất vector hoặc cấu trúc ADN bất kỳ chứa trình tự axit nucleic mã hóa gen thông báo theo sáng chế, cụ thể là RFP, được kèm hai bên bởi nhánh bên trái và nhánh bên phải cần thiết cho sự tái tổ hợp và sự tích hợp vào locus đích trong nhiễm sắc thể giới tính, cụ thể là nhiễm sắc thể giới tính Z. Ví dụ không làm giới hạn đối với vector này có thể là vector chứa trình tự axit nucleic như được bộc lộ trên Fig.15. Theo một số phương án khác nữa, trình tự axit nucleic bao gồm trình tự axit nucleic như được thể hiện bởi ít nhất một trong số trình tự SEQ ID NO. 66 và/hoặc SEQ ID NO. 67.

Ngoài ra, sáng chế còn bao gồm tế bào chủ bất kỳ chứa và/hoặc biểu hiện vector và cấu trúc ADN theo sáng chế, cụ thể là vector bất kỳ trong số các vector được mô tả theo sáng chế.

Khi được sử dụng trong bản mô tả này, các thuật ngữ “**ADN tái tổ hợp**”, “**trình tự axit nucleic tái tổ hợp**” hoặc “**gen tái tổ hợp**” dùng để chỉ axit nucleic chứa khung đọc mở mã hóa một trong các hệ thống CRISPR theo sáng chế, cụ thể là CRISPR/Cas9 kiểu II, cùng với gARN theo sáng chế hướng đích Cas9 đến đoạn tiền đệm tương ứng ở locus đặc hiệu trong nhiễm sắc thể loài chim Z và/hoặc W. Theo một số phương án khác nữa, khi được sử dụng trong bản mô tả này, ADN tái tổ hợp còn dùng để chỉ axit nucleic chứa khung đọc mở mã hóa gen thông báo theo sáng chế, cụ thể là gen chuyển.

Khi được đề cập trong bản mô tả này, thuật ngữ “**gen**” hoặc “**gen chuyển**” có nghĩa là axit nucleic, có trong tự nhiên hoặc tổng hợp, mã hóa sản phẩm protein. Thuật ngữ “axit nucleic” được dự định có nghĩa là các mảng mạch thẳng, mạch vòng và lần

lượt trong tự nhiên và/hoặc tổng hợp của nucleotit và nucleosit, ví dụ, cADN, ADN hệ gen (gADN), mARN, và ARN, oligonucleotit, oligonucleosit, và dẫn xuất của chúng.

Cụm từ "**được liên kết khi hoạt động**" được dự định có nghĩa là được gắn theo phương thức mà cho phép phiên mã gen chuyển. Như đã lưu ý ở trên, gen chuyển được sử dụng bởi phương pháp theo sáng chế được tạo ra bằng cách cung cấp trình tự axit nucleic mã hóa gen chuyển được tích hợp vào nhiễm sắc thể giới tính Z hoặc W, cụ thể là gen thông báo, và hệ thống chỉnh sửa gen (ví dụ, hệ thống CRISP/Cas9). Thuật ngữ "**mã hóa**" được dự định có nghĩa là axit nucleic đối tượng có thể được phiên mã và dịch mã thành polypeptit mong muốn hoặc protein đối tượng trong hệ thống biểu hiện thích hợp, ví dụ, khi axit nucleic đối tượng được liên kết với trình tự kiểm soát thích hợp chẳng hạn như gen khởi đầu và thành phần gen tăng cường ở vectơ thích hợp (ví dụ, vectơ biểu hiện) và khi vectơ này được đưa vào hệ thống hoặc tế bào thích hợp.

Cần hiểu rằng theo một số phương án, ít nhất một trong trình tự axit nucleic thứ nhất và trình tự axit nucleic thứ hai được cung cấp và sử dụng bởi phương pháp và kit theo sáng chế có thể được tạo cấu trúc và được chứa trong vectơ. Khi được sử dụng trong bản mô tả này, thuật ngữ "**vectơ**" hoặc "**phương tiện truyền**" bao hàm các vectơ chẳng hạn như plasmit, phagemit, virus, mảnh ADN có thể tích hợp, và các phương tiện truyền khác, mà có khả năng tích hợp các mảnh ADN vào hệ gen của vật chủ, hoặc theo cách khác là, có khả năng biểu hiện thành phần di truyền không được tích hợp. Các vectơ thường là cấu trúc ADN hoặc ARN tự sao chép chứa trình tự axit nucleic mong muốn, và thành phần kiểm soát di truyền được liên kết khi hoạt động được nhận diện trong tế bào chủ thích hợp và thực hiện sự dịch mã đoạn đệm mong muốn. Nói chung, thành phần kiểm soát di truyền có thể bao gồm hệ thống gen khởi đầu sinh vật nhân sơ hoặc hệ thống kiểm soát biểu hiện gen khởi đầu sinh vật nhân chuẩn. Hệ thống này thường bao gồm gen khởi đầu phiên mã, gen tăng cường phiên mã để làm gia tăng mức độ biểu hiện của ARN. Vectơ thường chứa điểm bắt đầu sao chép để cho phép vectơ sao chép độc lập với tế bào chủ. Theo một số phương án thay thế khác, vectơ biểu hiện được sử dụng bởi sáng chế có thể chứa thành phần cần thiết cho sự tích hợp gen thông báo mong muốn theo sáng chế vào nhiễm sắc thể đặc hiệu giới tính loài chim W và/hoặc Z.

Theo đó, thuật ngữ "**thành phần kiểm soát và điều hòa**" bao gồm gen khởi

đầu, gen kết thúc và các thành phần kiểm soát sự biểu hiện khác. Các thành phần điều hòa này được mô tả trong tài liệu Goeddel; [Goeddel., et al., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, Calif. (1990)]. Ví dụ, trình tự bất kỳ trong số nhiều trình tự kiểm soát biểu hiện để kiểm soát sự biểu hiện của trình tự ADN được liên kết khi hoạt động với nó có thể được sử dụng trong các vectơ này để biểu hiện trình tự ADN mã hóa protein mong muốn bất kỳ bằng cách sử dụng phương pháp theo sáng chế này.

Vectơ có thể bao gồm thêm các vị trí giới hạn thích hợp, tính kháng kháng sinh hoặc các dấu ấn khác để chọn lọc tế bào chứa vectơ. Plasmid là dạng được sử dụng phổ biến nhất của vectơ nhưng các dạng khác của vectơ dùng làm chức năng tương đương và đã được biết, hoặc trở nên được biết đến trong lĩnh vực phù hợp để sử dụng trong bản mô tả này. Xem tài liệu, ví dụ, Pouwels et al., Cloning Vectors: a Laboratory Manual (1985 and supplements), Elsevier, N.Y.; và Rodriquez, et al. (eds.) Vectors: a Survey of Molecular Cloning Vectors and their Uses, Buttersworth, Boston, Mass (1988), các tài liệu này được kết hợp vào đây để tham khảo.

Để tạo ra động vật thuộc loài chim chuyển gen được sử dụng bởi phương pháp theo sáng chế, tế bào chim chứa gen thông báo được tích hợp vào locus đặc hiệu trong nhiễm sắc thể giới tính Z hoặc W của chúng cần được tạo ra. Tế bào này có thể được tạo ra theo một số phương án, bằng cách tiếp xúc, đưa hoặc đồng-chuyển nhiễm tế bào của đối tượng thuộc loài chim hoặc tế bào bất kỳ được đưa vào đối tượng này, với trình tự axit nucleic thứ nhất và trình tự axit nucleic thứ hai được tạo ra bởi phương pháp và kit theo sáng chế hoặc với cấu trúc, vectơ, phương tiện truyền, bất kỳ chứa chúng. Theo một số phương án, trình tự axit nucleic mã hóa các thành phần chỉnh sửa gen, và trình tự axit nucleic mã hóa gen thông báo được hướng đích, cụ thể là RFP. Khi được sử dụng trong bản mô tả này, thuật ngữ "**sự chuyển nhiễm**" có nghĩa là quá trình chèn vật liệu di truyền, chẳng hạn như ADN và ARN sợi kép, vào tế bào của động vật có vú. Sự chèn đoạn ADN vào tế bào cho phép biểu hiện, hoặc tạo ra, các protein bằng cách sử dụng bộ máy của bản thân tế bào. Do đó, khi được sử dụng trong bản mô tả này, sự đồng-chuyển nhiễm để chỉ sự chuyển nhiễm đồng thời của ít nhất là hai phân tử axit nucleic khác nhau hoặc vectơ bất kỳ chứa chúng vào mỗi tế bào đơn lẻ. Ngoài ra, trình tự axit nucleic được chuyển nhiễm có thể được biểu hiện tạm thời trong khoảng thời gian ngắn, hoặc trở nên được tích hợp vào ADN hệ gen, mà tại đó

sự thay đổi được truyền từ tế bào này sang tế bào khác khi nó phân chia.

Do đó, sáng chế đề xuất phương pháp xác định giới tính *in-ovo* của phôi chim *in-ovo* dựa trên sự biểu hiện của gen thông báo, cụ thể là protein phát huỳnh quang màu đỏ (RFP) được tích hợp trong các/nhiễm sắc thể giới tính của chúng (cụ thể là Z và/hoặc W). “**Sự biểu hiện**” thường dùng để chỉ quá trình mà nhờ đó thông tin được mã hóa bởi gen được biến đổi thành các cấu trúc có mặt và hoạt động trong tế bào. Do đó, theo sáng chế “sự biểu hiện” của gen thông báo, cụ thể là có thể dùng để chỉ sự phiên mã thành polynucleotit, sự dịch mã thành protein, hoặc thậm chí là sự cải biến sau dịch mã của protein.

Sự chèn đoạn và sự tích hợp đặc hiệu của gen thông báo vào locus đích đặc hiệu trong nhiễm sắc thể giới tính của đối tượng thuộc loài chim chuyển gen, có thể bao gồm trong một số phương án với điều kiện là hệ thống CRISPR/Cas9 bao gồm gARN đặc hiệu và trình tự axit nucleic của gen thông báo được tích hợp. Để tạo điều kiện thuận lợi và cho phép sự tích hợp, gen thông báo được kèm hai bên theo một số phương án, với các trình tự có thể là thể tương đồng với các trình tự kèm hai bên vị trí tích hợp được hướng đích và nhờ đó cho phép sự tái tổ hợp. Do đó, theo một số phương án cụ thể khác nữa, ít nhất một gen thông báo trong trình tự axit nucleic thứ hai có thể được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng. Cần hiểu rằng, theo một số phương án, cần có các nhánh này và do đó tạo điều kiện thuận lợi cho HDR của gen thông báo ở vị trí tích hợp.

Theo phương án cụ thể hơn, gen thông báo ở trình tự axit nucleic thứ hai được sử dụng bởi phương pháp theo sáng chế, có thể được kèm hai bên với hai nhánh tương đồng hoặc có tính tương đồng hoặc đồng nhất bằng khoảng 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% hoặc 100% đối với toàn thể, toàn bộ, hoặc hoàn toàn ít nhất một trình tự axit nucleic được chứa trong locus đích trong nhiễm sắc thể giới tính Z hoặc W, mà đóng vai trò là vị trí tích hợp để tạo điều kiện thuận lợi cho sự tích hợp đặc hiệu thông qua HDR. Theo các phương án nhất định, trình tự đích còn được gọi trong bản mô tả này là ít nhất một "đoạn tiền đệm" được nhận diện bởi trình tự "đoạn đệm" mà là một phần của gARN được sử dụng bởi sáng chế, và được cung cấp bởi trình tự axit nucleic thứ nhất.

Khi được sử dụng trong bản mô tả này, thuật ngữ "**nhánh tương đồng**" để chỉ

các khuôn HDR được đưa vào vector hoặc virus đặc hiệu, được sử dụng để tạo ra sự đột biến hoặc sự chèn đoạn đặc hiệu của thành phần mới vào gen, để có lượng nhất định của thể tương đồng xung quanh trình tự đích cần được cải biến (phụ thuộc vào PEN nào được sử dụng). Theo một số phương án cụ thể khác nữa, trong đó CRISPR được sử dụng dưới dạng PEN, các trình tự nhánh (bên trái, ngược dòng và bên phải, xuôi dòng) có thể chứa từ khoảng 10 đến 5000 bp, cụ thể là từ khoảng 50 đến 1000 bp, trong khoảng từ 100 đến 500, cụ thể là, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 bp.

Theo một số phương án khác nữa, trình tự hướng đích trong gARN được mã hóa bởi trình tự axit nucleic thứ nhất được tạo ra bởi phương pháp và kit theo sáng chế, còn được gọi trong bản mô tả này là trình tự "đoạn đệm", có mức độ đồng nhất hoặc tương đồng bằng khoảng 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% hoặc 100% đối với toàn bộ, hoàn toàn hoặc toàn thể ít nhất một trình tự axit nucleic được chứa trong locus đích trong nhiễm sắc thể giới tính Z hoặc W, được đề cập đến trong bản mô tả này dưới dạng "đoạn tiền đệm".

Theo một số phương án, ít nhất một gen thông báo ở trình tự axit nucleic thứ hai có thể được liên kết khi hoạt động với thành phần bất kỳ trong số gen khởi đầu đặc hiệu giới tính, gen khởi đầu đặc hiệu phôi (ví dụ, gen khởi đầu α -Globin) và gen khởi đầu có thể cảm ứng (ví dụ, gen khởi đầu cảm ứng ánh sáng có nguồn gốc từ gen SSU đậu nành, hoặc có nguồn gốc từ gen khởi đầu CHS chalcone synthase rau mùi tây, hoặc phiên bản thiết kế của EL222, protein ánh sáng-oxy-điện áp của vi khuẩn hoạt hóa sự biểu hiện khi được chiếu ánh sáng màu xanh lam). Theo một số phương án cụ thể hơn nữa, gen thông báo có thể chịu sự kiểm soát của gen khởi đầu phôi, nhờ đó làm giới hạn sự biểu hiện của gen thông báo chuyển gen đối với giai đoạn phôi, với sự không biểu hiện ở con non trưởng thành. Theo phương án này, gen chuyển thông báo có thể được sử dụng và chỉ được biểu hiện ở giai đoạn phôi, nhằm mục đích chẩn đoán.

Cụ thể hơn, khi được sử dụng trong bản mô tả này, "**gen khởi đầu**" dùng để chỉ vùng cụ thể của ADN có khả năng kiểm soát sự biểu hiện của gen được đặt xuôi dòng.

Do đó, "**gen khởi đầu đặc hiệu giới tính ở con non**" trong bản mô tả này dùng để chỉ gen khởi đầu hoạt hóa sự biểu hiện của gen, chỉ ở giới tính con non riêng biệt

(tức là con đực hoặc con cái). Ngoài ra, "**gen khởi đầu đặc hiệu sự phát triển ở con non**" dùng để chỉ gen khởi đầu hoạt hóa sự biểu hiện của gen, chỉ ở các giai đoạn riêng biệt của sự phát triển con non.

Theo một số phương án cụ thể, ít nhất một gen thông báo có thể được chèn vào và nhờ đó được tích hợp vào ít nhất một vùng không mã hóa của nhiễm sắc thể giới tính đích. Phương pháp này tránh được sự làm gián đoạn gen có thể cần thiết cho sự phát triển và trưởng thành của phôi chưa nở.

Khi được sử dụng trong bản mô tả này, "**vùng không mã hóa**" dùng để chỉ các thành phần của ADN của sinh vật không mã hóa các trình tự protein. Một số vùng ADN không mã hóa được phiên mã thành phân tử ARN không mã hóa chức năng, các chức năng khác của vùng ADN không mã hóa bao gồm sự điều hòa phiên mã và dịch mã của trình tự mã hóa protein, vùng gắn giá đỡ, điểm mở đầu sao chép ADN, đoạn trung tâm và đoạn cuối. Phần không chức năng theo giả thiết (hoặc ADN có chức năng chưa biết) thường được gọi là "ADN phế thải".

Theo một số phương án cụ thể khác, ít nhất một gen thông báo được sử dụng bởi phương pháp theo sáng chế để tạo ra động vật thuộc loài chim chuyển gen, có thể được tích hợp vào ít nhất một vị trí ở nhiễm sắc thể giới tính Z. Theo phương án cụ thể hơn, locus đặc hiệu ở nhiễm sắc thể Z có thể là vùng bất kỳ trong số các vùng 9156874-9161874, như được thể hiện bằng trình tự SEQ ID NO:15, 27764943-27769943, như được thể hiện bằng trình tự SEQ ID NO:16, 42172748-42177748, như được thể hiện bằng trình tự SEQ ID NO:17, 63363656-63368656, như được thể hiện bằng trình tự SEQ ID NO:18 và 78777477-78782477, như được thể hiện bằng trình tự SEQ ID NO:19 của nhiễm sắc thể Z của gà mái.

Theo một số phương án cụ thể, ít nhất một gen thông báo có thể được tích hợp vào ít nhất một vị trí ở locus nhiễm sắc thể giới tính Z 42172748-42177748, cụ thể là như được thể hiện bằng trình tự SEQ ID NO. 17. Theo một số phương án cụ thể khác nữa, trình tự gARN phù hợp cho sự tích hợp vào locus đặc hiệu trong nhiễm sắc thể Z có thể bao gồm, nhưng không chỉ giới hạn ở, gARN7 được định hướng đến đoạn tiền đệm nằm ở trong locus nhiễm sắc thể Z chrZ_42174515_-1. Theo một số phương án cụ thể, gARN có thể là gARN được ký hiệu là gARN7. Theo phương án cụ thể hơn, gARN7 này có thể chứa trình tự axit nucleic GTAATACAGAGCTAAACCAG, cũng như được thể hiện bằng trình tự SEQ ID NO:26. Theo một số phương án khác nữa, để

tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, trình tự mã hóa gen thông báo, cụ thể là RFP, được đề xuất bởi sáng chế, có thể được kèm hai bên bởi nhánh bên trái ở đầu 5' của chúng, và nhánh bên phải ở đầu 3' của chúng. Theo một số phương án cụ thể khác nữa, nhánh bên trái có thể chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 31, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 32, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN7 trong trình tự SEQ ID NO:26.

Theo một số phương án cụ thể khác nữa, ít nhất một gARN cần thiết để hướng đích gen thông báo đến vị trí đặc hiệu này trong nhiễm sắc thể Z có thể chứa trình tự axit nucleic như được thể hiện bởi gARN bất kỳ trong số gARN3: ACAGACCTATGATATGT, như được thể hiện bằng trình tự SEQ ID NO. 11; gARN4: CGATTATCACTCACAAG, như được thể hiện bằng trình tự SEQ ID NO. 12; gARN5: CTGGTTAGCATGGGGAC, như được thể hiện bằng trình tự SEQ ID NO. 13; gARN6: GTAAAGAGTCAGATACA, như được thể hiện bằng trình tự SEQ ID NO. 14.

Theo một số phương án khác nữa, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 41 và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 42 có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN3 trong trình tự SEQ ID NO: 11. Theo các phương án khác, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 43, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 44, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN4 trong trình tự SEQ ID NO:12. Theo các phương án khác nữa, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 45, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 46, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN5 trong trình tự SEQ ID NO:13. Theo một số phương án khác, để tích

hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 47, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 48, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN6 trong trình tự SEQ ID NO:14.

Các ví dụ không làm giới hạn khác đối với trình tự gARN phù hợp cho sự tích hợp vào locus đặc hiệu trong nhiễm sắc thể Z có thể bao gồm, nhưng không chỉ giới hạn ở, gARN8 của locus nhiễm sắc thể Z chrZ_9157091_1, chứa trình tự axit nucleic ACAGACCTATGATATGTGAG, cũng như được thể hiện bằng trình tự SEQ ID NO:27, gARN9 của locus nhiễm sắc thể Z chrZ_27767602_-1, chứa trình tự axit nucleic GAGCTTGTGAGTGATAATCG, cũng như được thể hiện bằng trình tự SEQ ID NO:28, gARN10 của locus nhiễm sắc thể Z chrZ_78779927_1, chứa trình tự axit nucleic GTAAAGAGTCAGATACACAG, cũng như được thể hiện bằng trình tự SEQ ID NO: 29, và gARN11 của locus nhiễm sắc thể Z chrZ_63364946_-1, chứa trình tự axit nucleic CAGTGGGTACTGAAGCTGTG cũng như được thể hiện bằng trình tự SEQ ID NO: 30.

Theo các phương án khác, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 33, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 34, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN8 trong trình tự SEQ ID NO:27. Theo các phương án khác nữa, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 35, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 36, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN9 trong trình tự SEQ ID NO:28. Theo các phương án khác, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 37, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 38, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN10 trong trình tự SEQ ID NO:29. Theo phương án khác nữa, để tích hợp gen

thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 39, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 40, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN11 trong trình tự SEQ ID NO:30.

Theo một số phương án thay thế khác, ít nhất một gen thông báo có thể được tích hợp vào ít nhất một vị trí ở nhiễm sắc thể giới tính W. Theo phương án cụ thể hơn, locus đặc hiệu ở nhiễm sắc thể W có thể là vị trí 1022859-1024215. Theo một số phương án cụ thể, locus đích có thể chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 3.

Theo phương án cụ thể hơn, ít nhất một gARN cần thiết để hướng đích gen thông báo đến vị trí đặc hiệu này trong nhiễm sắc thể W có thể chứa trình tự axit nucleic như được thể hiện bởi trình tự bất kỳ trong số các trình tự SEQ ID NO. 1 và 2, các gARN này được ký hiệu trong bản mô tả này lần lượt là gARN1 và gARN2.

Theo một số phương án cụ thể hơn nữa, gARN được sử dụng bởi phương pháp theo sáng chế để tạo ra chim cái chuyển gen có thể chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 1 (gARN1). Trong trường hợp này, ít nhất một gen thông báo được chứa trong trình tự axit nucleic thứ hai này có thể được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng chứa trình tự axit amin như được thể hiện lần lượt bằng trình tự SEQ ID NO. 4 và 5, để tạo điều kiện thuận lợi cho sự tích hợp của chúng vào locus đặc hiệu này ở nhiễm sắc thể W. Cần hiểu rằng các nhánh này cũng được đề cập trong bản mô tả này lần lượt là nhánh bên trái và nhánh bên phải.

Theo một số phương án thay thế khác, gARN được sử dụng để tạo ra chim cái chuyển gen theo sáng chế có thể chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 2 (gARN2). Trong trường hợp này, ít nhất một gen thông báo được chứa trong trình tự axit nucleic thứ hai được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng chứa trình tự axit amin như được thể hiện lần lượt bằng trình tự SEQ ID NO. 6 và 7. Cần hiểu rằng các nhánh này cũng được đề cập đến trong bản mô tả này lần lượt là nhánh bên trái và nhánh bên phải.

Khi các locus di truyền của tế bào hợp tử của vật chủ thuộc loài chim, đã được hướng đích và/hoặc được chuyển nhiễm bằng trình tự ngoại sinh, cụ thể là gen thông

báo được sử dụng bởi sáng chế, tốt hơn là có thể sử dụng các tế bào này để tạo ra động vật chuyển gen. Đối với quy trình này, sau khi đưa cấu trúc hướng đích vào tế bào gốc phôi (ES), các tế bào này có thể được đàn mỏng trên lớp tiếp liệu trong môi trường thích hợp, ví dụ, DMEM có bổ sung với các thành phần sinh trưởng và xytokin, huyết thanh bào thai bò và chất kháng sinh. Tế bào gốc phôi có thể có locus được hướng đích đơn lẻ (dị hợp tử) hoặc cả hai locus được hướng đích (đồng hợp tử). Tế bào chứa cấu trúc có thể được phát hiện bằng cách sử dụng môi trường có tính chọn lọc và sau khi đủ thời gian cho khuẩn lạc sinh trưởng, khuẩn lạc có thể được nhặt ra và phân tích về sự xuất hiện của gen hướng đích. Theo một số phương án cụ thể, PCR có thể được dùng để xác nhận sự tích hợp trình tự ngoại sinh mong muốn vào locus đích, bằng cách sử dụng các đoạn môi ở trong và ở ngoài trình tự cấu trúc. Khuẩn lạc có gen hướng đích có thể được sử dụng sau đó để tiêm vào phôi chim. Sau đó, tế bào ES có thể được trypsin hóa và các tế bào cải biến có thể được tiêm thông qua lỗ hở được tạo ra ở bên cạnh trứng. Sau khi bịt kín trứng, có thể ủ trứng trong các điều kiện thích hợp cho đến khi nở. Chim mới nở có thể được kiểm tra về sự có mặt của trình tự cấu trúc đích, ví dụ, bằng cách kiểm tra mẫu sinh học của chúng, ví dụ, mẫu máu. Sau khi chim đã đạt đến độ trưởng thành, chúng được phối giống và con cái của chúng có thể được kiểm tra để xác định liệu trình tự được tích hợp ngoại sinh có được truyền qua dòng mầm hay không.

Chim thể khảm được tạo ra có nguồn gốc một phần từ các tế bào gốc phôi được cải biến hoặc tế bào hợp tử, có khả năng truyền sự cải biến di truyền qua dòng mầm. Việc giao phối chủng chim chứa trình tự ngoại sinh, cụ thể là gen thông báo được sử dụng bởi sáng chế, hoặc các phần của chúng, với chủng mà trong đó locus kiểu đại của chim, hoặc các phần của chúng, được phục hồi, dẫn đến các con cái thể hiện giới tính có thể nhận thấy *in-ovo*.

Ngoài ra, chim chuyển gen cũng có thể được tạo ra theo các phương pháp khác, một số trong đó được bàn luận dưới đây. Trong số các tế bào chim phù hợp cho sự biến nạp để tạo ra động vật chuyển gen là tế bào mầm nguyên thủy (PGC), tinh bào và tế bào hợp tử (bao gồm cả tế bào gốc phôi). Tinh bào có thể được biến nạp với cấu trúc ADN bằng phương pháp thích hợp bất kỳ, bao gồm gây xung điện, bắn phá vi hạt, chuyển nhiễm liposom và dạng tương tự. Tinh dịch có thể được sử dụng để cấy tinh nhân tạo ở loài chim. Con cái của chim đã được cấy tinh có thể được kiểm tra trình tự

ngoại sinh như đã mô tả trên đây.

Theo cách khác, tế bào mầm nguyên thủy có thể được phân lập từ trứng chim, được chuyển nhiễm bằng gen thông báo ngoại sinh theo sáng chế bằng phương pháp thích hợp bất kỳ, và được chuyển nhiễm hoặc chèn vào phôi mới, mà tại đó chúng có thể trở nên được tích hợp vào tuyến sinh dục đang phát triển. Chim đã nở và con cái của chúng có thể được kiểm tra về trình tự gen thông báo ngoại sinh như được mô tả theo sáng chế.

Theo một phương án khác nữa, tế bào bì phôi phân tán được phân lập từ trứng có thể được chuyển nhiễm bằng các phương tiện thích hợp với trình tự gen thông báo ngoại sinh, hoặc các phần của chúng, được tích hợp vào nhiễm sắc thể đặc hiệu giới tính Z hoặc W, sau đó là tiêm vào khoang dưới mầm của trứng nguyên vẹn. Đối tượng thuộc loài chim đã nở và con cái của chúng có thể được kiểm tra về gen thông báo ngoại sinh như đã mô tả trên đây.

Tế bào mầm nguyên thủy (PGC) ở gà là tế bào tiền thân đối với trứng và tinh trùng. Do đó, theo một số khía cạnh của chúng, sáng chế đề xuất sự tạo ra gà chuyển gen thông qua hệ thống truyền dòng mầm bằng cách sử dụng PGC được đồng-chuyển nhiễm bằng cấu trúc gen thông báo và bằng cấu trúc gARN CRISPR/Cas9 định hướng sự tích hợp gen thông báo vào nhiễm sắc thể đặc hiệu giới tính W và/hoặc Z. PGC được phân loại và chuyển vào dòng máu của phôi thể nhận 2,5-ngày để truyền dòng mầm.

PGC là tiền thân của các giao tử và có hoạt tính di chuyển đơn nhất trong quá trình phát sinh phôi sớm ở chim, khi so sánh với động vật có vú. Ở phôi chim, PGC đầu tiên phát sinh từ ngoại phôi bì và di chuyển đến hạ phôi bì của vòng cung mầm ở giai đoạn 4 (HH), khoảng 18–19 giờ sau khi ủ. Giữa các giai đoạn 10 và 12 (HH), PGC di chuyển từ vòng cung mầm vào dòng máu và di chuyển thông qua hệ tuần hoàn cho đến khi chúng tiến đến cung sinh dục và định cư ở tuyến sinh dục đang phát triển. Điều này trái ngược với động vật có vú trong đó PGC di chuyển thông qua mô phôi để tiến đến tuyến sinh dục đang phát triển. Chính đặc điểm đặc biệt này của sự di chuyển PGC ở loài chim thông qua dòng máu đã tạo điều kiện thuận lợi để cải tiến lớn về sự cải biến di truyền của gà.

Do đó, theo một số phương án cụ thể, **"sự tạo ra động vật thuộc loài chim chuyển gen"** dùng để chỉ phương pháp nhiều bước liên quan đến các kỹ thuật thiết kế

di truyền để tạo ra gà có sự cải biến hệ gen trong đó a) tế bào mầm nguyên thủy (PGC) được phân lập từ máu của phôi con non hai ngày tuổi; b) cấu trúc gen chuyển được kết hợp vào PGC được nuôi cấy, ví dụ, bằng cách sử dụng hệ lentivirut, vector nhẩy Piggybac, các kỹ thuật TALENS hoặc CRISPR/Cas9; (c) PGC chuyển gen được xác định và tiêm vào hệ tuần hoàn của phôi và di chuyển đến tuyến sinh dục đang phát triển; d) phôi thể nhận được ủ ở nhiệt độ 37°C cho đến khi nở; (d) con đực đã nở được nuôi đến độ trưởng thành sinh dục và được lai với con mái kiểu đại; (e) đời sau được sàng lọc để nhận biết các con có nguồn gốc từ PGC chuyển gen.

Do đó, theo khía cạnh thứ hai, sáng chế đề cập đến động vật hoặc đối tượng chuyển gen thuộc loài chim chứa, trong ít nhất một tế bào của chúng, ít nhất một gen thông báo ngoại sinh được tích hợp vào ít nhất một vị trí hoặc địa điểm (còn được gọi trong bản mô tả này là locus) ở trong ít nhất một trong hai nhiễm sắc thể giới tính Z và W.

Thuật ngữ "**chim**" đề chỉ loài bất kỳ có nguồn gốc từ chim đặc trưng bởi lông vũ, hàm nhô ra không có răng, để trứng có vỏ cứng, tốc độ trao đổi chất cao, tim bốn ngăn, và khối lượng nhẹ nhưng bộ xương khỏe. Loài chim bao gồm, nhưng không chỉ giới hạn ở, gà, chim cút, gà tây, vịt, *Gallinacea sp*, ngỗng, gà lôi và các loài gà khác. Thuật ngữ "con mái" bao gồm tất cả các con thuộc giới cái của các loài chim. "**Chim chuyển gen**" thường dùng để chỉ chim có trình tự ADN khác loại, hoặc một hoặc nhiều trình tự ADN bổ trợ thường là nội sinh đối với loài chim (được gọi chung trong bản mô tả này là "gen chuyển") được tích hợp nhiễm sắc thể vào tế bào mầm của chim. Kết quả của sự chuyển nhiễm và tích hợp này là, trình tự được chuyển nhiễm có thể được truyền qua tế bào mầm đến đời sau của chim chuyển gen. Chim chuyển gen (bao gồm cả con cháu của chúng) cũng có gen chuyển được tích hợp vào nhiễm sắc thể giới tính của tế bào sinh dưỡng.

Theo một số phương án cụ thể, ít nhất một động vật chuyển gen theo sáng chế có thể chứa ít nhất hai gen thông báo khác nhau. Trong trường hợp này, mỗi gen thông báo có thể được tích hợp vào ít nhất một vị trí hoặc địa điểm ở một trong hai nhiễm sắc thể giới tính Z và/hoặc W. Theo một số phương án khác nữa, mỗi gen thông báo có thể được tích hợp vào ít nhất là hai, ít nhất là ba, ít nhất là bốn, ít nhất là năm, ít nhất là sáu, ít nhất là bảy, ít nhất là tám, ít nhất là chín, ít nhất là mười hoặc nhiều hơn, các vị trí hoặc địa điểm ở một trong hai nhiễm sắc thể giới tính Z và/hoặc W.

Theo một số phương án khác nữa, gen thông báo được chứa trong động vật chuyển gen theo sáng chế, có thể là ít nhất một gen thông báo phát huỳnh quang. Cần hiểu rằng gen thông báo bất kỳ và cụ thể là gen thông báo phát huỳnh quang bất kỳ được bộc lộ bởi sáng chế, cụ thể là trong Bảng 1 trên đây, kết hợp với các khía cạnh khác của sáng chế cũng có thể thích hợp để kết hợp với đối tượng thuộc loài chim chuyển gen được đề xuất bởi sáng chế.

Theo phương án cụ thể hơn, gen thông báo phát huỳnh quang này có thể chứa hoặc có thể là ít nhất một protein phát huỳnh quang màu đỏ (RFP). Tuy nhiên cần lưu ý rằng, theo một số phương án, gen thông báo được biểu hiện dưới dạng gen chuyển bởi đối tượng thuộc loài chim chuyển gen theo sáng chế, có thể là protein phát huỳnh quang bất kỳ, ví dụ, RFP, YFP, BFP và dạng tương tự, với điều kiện là protein phát huỳnh quang này không phải là protein phát huỳnh quang màu xanh lục (GFP). Mặt khác, theo một số phương án cụ thể, gen thông báo được tích hợp vào nhiễm sắc thể giới tính của động vật chuyển gen được đề xuất bởi sáng chế có thể là protein phát huỳnh quang bất kỳ với điều kiện là protein phát huỳnh quang này không phải là GFP.

Theo các phương án nhất định, ít nhất một động vật thuộc loài chim chuyển gen được đề xuất bởi sáng chế, có thể là con cái. Theo phương án cụ thể hơn, ít nhất một gen thông báo ở chim cái chuyển gen này có thể được tích hợp vào ít nhất một vị trí của nhiễm sắc thể con cái Z.

Theo một số phương án thay thế khác, ít nhất một động vật thuộc loài chim chuyển gen có thể là con cái, có ít nhất một gen thông báo được tích hợp vào ít nhất một vị trí của nhiễm sắc thể con cái W.

Theo một số phương án thay thế khác nữa, động vật chuyển gen được đề xuất bởi sáng chế có thể là đối tượng con đực chuyển gen mang gen thông báo gen chuyển được tích hợp vào ít nhất một trong hai nhiễm sắc thể giới tính Z của nó. Theo một số phương án khác, động vật chuyển gen được đề xuất bởi sáng chế có thể là đối tượng con đực chuyển gen mang gen thông báo gen chuyển được tích hợp vào một trong hai nhiễm sắc thể giới tính Z của nó. Theo các phương án thay thế khác, động vật chuyển gen được đề xuất bởi sáng chế có thể là đối tượng con đực chuyển gen mang gen thông báo gen chuyển được tích hợp vào hai trong hai nhiễm sắc thể giới tính Z của nó.

Theo một số phương án cụ thể, ít nhất một gen thông báo có thể được tích hợp vào nhiễm sắc thể giới tính của động vật chuyển gen theo sáng chế bằng cách sử dụng

ít nhất một PEN. Như đã lưu ý ở trên liên quan đến phương pháp theo sáng chế, hệ thống thiết kế di truyền bất kỳ hoặc PEN bất kỳ có thể được sử dụng, cụ thể là hệ thống được dẫn đường hướng bởi ARN bất kỳ, ví dụ, một trong số CRISPR, TALEN, ZFN và dạng tương tự.

Cụ thể hơn, PEN này có thể có mặt trong các phương án nhất định, hệ thống CRISPR có thể được sử dụng. Theo phương án cụ thể hơn, hệ thống CRISPR loại 2 có thể được sử dụng trong việc tạo ra động vật chuyển gen theo sáng chế. Theo một số phương án cụ thể khác nữa, hệ thống CRISPR kiểu II có thể được sử dụng.

Theo một số phương án cụ thể hơn nữa, ít nhất một gen thông báo có thể được tích hợp vào nhiễm sắc thể giới tính của động vật thuộc loài chim chuyển gen theo sáng chế bằng HDR qua trung gian ít nhất một hệ thống CRISPR/Cas9.

Theo phương án cụ thể hơn, ít nhất một gen thông báo có thể được tích hợp vào nhiễm sắc thể giới tính của động vật thuộc loài chim chuyển gen theo sáng chế bằng cách cho tiếp xúc hoặc đồng-chuyển nhiễm ít nhất một tế bào của động vật/đối tượng thuộc loài chim này, hoặc ít nhất một tế bào mà được đưa vào động vật thuộc loài chim này với ít nhất hai trình tự axit nucleic. Cụ thể hơn, tế bào này có thể được cho tiếp xúc hoặc đồng-chuyển nhiễm với (a) ít nhất một thành phần của hệ thống chỉnh sửa gen, cụ thể là ít nhất một protein Cas9 hoặc ít nhất một trình tự axit nucleic thứ nhất chứa ít nhất một trình tự axit nucleic mã hóa ít nhất một protein Cas9 này; và ít nhất một gARN hướng đích ít nhất một đoạn tiền đệm trong ít nhất một nhiễm sắc thể giới tính, Z hoặc W, hoặc ít nhất một trình tự axit nucleic mã hóa ít nhất một gARN này, nhờ đó tạo ra sự tích hợp qua trung gian CRISPR; và (b) ít nhất một trình tự axit nucleic thứ hai chứa ít nhất một gen thông báo hoặc các mảnh bất kỳ của chúng.

Theo phương án cụ thể hơn, ít nhất một gen thông báo ở trình tự axit nucleic thứ hai có thể được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng. Các nhánh này có độ tương đồng với vị trí đích tích hợp trong nhiễm sắc thể giới tính đích, nhờ đó tạo điều kiện thuận lợi cho HDR ở vị trí tích hợp.

Theo một số phương án cụ thể hơn nữa, ít nhất một gen thông báo ở trình tự axit nucleic thứ hai có thể được liên kết khi hoạt động với thành phần bất kỳ trong số gen khởi đầu đặc hiệu giới tính, gen khởi đầu đặc hiệu phôi và gen khởi đầu có thể cảm ứng. Gen khởi đầu này có thể có mặt trong một số phương án, làm giới hạn sự biểu hiện của gen thông báo theo sáng chế cho giới tính mong muốn riêng biệt (trong

trường hợp là gen khởi đầu đặc hiệu giới tính), giai đoạn phôi riêng biệt (gen khởi đầu đặc hiệu phôi) hoặc điều kiện riêng biệt (điều kiện cảm ứng).

Theo một số phương án cụ thể khác nữa, ít nhất một gen thông báo được chứa trong động vật thuộc loài chim chuyển gen theo sáng chế có thể được tích hợp vào ít nhất một vùng không mã hóa của một trong hai nhiễm sắc thể giới tính của nó.

Theo một số phương án thay thế khác nữa, động vật thuộc loài chim chuyển gen theo sáng chế có thể chứa ít nhất một gen thông báo được tích hợp vào ít nhất một vị trí ở nhiễm sắc thể giới tính Z. Theo một số phương án cụ thể và không làm giới hạn, động vật thuộc loài chim chuyển gen này có thể là con cái mang gen thông báo chuyển gen được tích hợp vào nhiễm sắc thể Z. Theo phương án cụ thể hơn, locus đặc hiệu ở nhiễm sắc thể Z có thể là vùng bất kỳ trong số các vùng 9156874-9161874, như được thể hiện bằng trình tự SEQ ID NO:15, 27764943-27769943, như được thể hiện bằng trình tự SEQ ID NO:16, 42172748-42177748, như được thể hiện bằng trình tự SEQ ID NO:17, 63363656-63368656, như được thể hiện bằng trình tự SEQ ID NO:18 và 78777477-78782477, như được thể hiện bằng trình tự SEQ ID NO:19 của nhiễm sắc thể Z của gà mái.

Theo một số phương án cụ thể, ít nhất một gen thông báo có thể được tích hợp vào ít nhất một vị trí ở locus nhiễm sắc thể giới tính Z 42172748-42177748, cụ thể là như được thể hiện bằng trình tự SEQ ID NO. 17. Theo một số phương án cụ thể khác nữa, trình tự gARN phù hợp cho sự tích hợp vào locus đặc hiệu trong nhiễm sắc thể Z có thể bao gồm, nhưng không chỉ giới hạn ở, gARN7 được định hướng đến đoạn tiền đệm nằm ở trong locus nhiễm sắc thể Z chrZ_42174515_-1. Theo một số phương án cụ thể, gARN có thể là gARN được ký hiệu là gARN7. Theo phương án cụ thể hơn, gARN7 này có thể chứa trình tự axit nucleic GTAATACAGAGCTAAACCAG, cũng như được thể hiện bằng trình tự SEQ ID NO:26. Theo một số phương án khác nữa, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, trình tự mã hóa gen thông báo, cụ thể là RFP, được đề xuất bởi sáng chế, có thể được kèm hai bên bởi nhánh bên trái ở đầu 5' của chúng, và nhánh bên phải ở đầu 3' của chúng. Theo một số phương án cụ thể khác nữa, nhánh bên trái có thể chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 31, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 32, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi

gARN7 trong trình tự SEQ ID NO:26.

Theo một số phương án cụ thể hơn nữa, ít nhất một gARN cần thiết để hướng đích gen thông báo đến vị trí đặc hiệu này trong nhiễm sắc thể Z ở đối tượng thuộc loài chim chuyển gen, có thể chứa trình tự axit nucleic như được thể hiện bởi gARN bất kỳ trong số gARN3: ACAGACCTATGATATGT, như được thể hiện bằng trình tự SEQ ID NO. 11; gARN4: CGATTATCACTCACAAG, như được thể hiện bằng trình tự SEQ ID NO. 12; gARN5: CTGGTTAGCATGGGGAC, như được thể hiện bằng trình tự SEQ ID NO. 13; gARN6: GTAAAGAGTCAGATACA, như được thể hiện bằng trình tự SEQ ID NO. 14.

Theo một số phương án khác nữa, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 41, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 42, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN3 trong trình tự SEQ ID NO:11. Theo các phương án khác, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 43, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 44, có thể được sử dụng để tích hợp gen thông báo ở đối tượng thuộc loài chim chuyển gen theo sáng chế vào locus đặc hiệu được định hướng bởi gARN4 trong trình tự SEQ ID NO:12. Theo các phương án khác nữa, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z ở đối tượng thuộc loài chim chuyển gen, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 45, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 46, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN5 trong trình tự SEQ ID NO:13. Theo một số phương án khác, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 47, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 48, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN6 trong trình tự SEQ ID NO:14.

Các ví dụ không làm giới hạn khác đối với trình tự gARN phù hợp cho sự tích

hợp vào locus đặc hiệu trong nhiễm sắc thể Z của đối tượng thuộc loài chim chuyển gen theo sáng chế có thể bao gồm, nhưng không chỉ giới hạn ở, gARN8 của locus nhiễm sắc thể Z chrZ_9157091_1, chứa trình tự axit nucleic ACAGACCTATGATATGTGAG, cũng như được thể hiện bằng trình tự SEQ ID NO:27, gARN9 của locus nhiễm sắc thể Z chrZ_27767602_-1, chứa trình tự axit nucleic GAGCTTGTGAGTGATAATCG, cũng như được thể hiện bằng trình tự SEQ ID NO:28, gARN10 của locus nhiễm sắc thể Z chrZ_78779927_1, chứa trình tự axit nucleic GTAAAGAGTCAGATACACAG, cũng như được thể hiện bằng trình tự SEQ ID NO: 29, và gARN11 của locus nhiễm sắc thể Z chrZ_63364946_-1, chứa trình tự axit nucleic CAGTGGGTACTGAAGCTGTG cũng như được thể hiện bằng trình tự SEQ ID NO: 30.

Theo các phương án khác, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z của đối tượng thuộc loài chim chuyển gen theo sáng chế, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 33, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 34, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN8 trong trình tự SEQ ID NO:27. Theo các phương án khác nữa, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 35, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 36, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN9 trong trình tự SEQ ID NO:28. Theo các phương án khác, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 37, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 38, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN10 trong trình tự SEQ ID NO:29. Theo phương án khác nữa, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 39, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 40, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN11

trong trình tự SEQ ID NO:30.

Theo một số phương án thay thế khác nữa, ít nhất một gen thông báo có thể được tích hợp vào ít nhất một vị trí ở nhiễm sắc thể giới tính W của đối tượng thuộc loài chim chuyển gen theo sáng chế. Theo một số phương án cụ thể, vị trí tích hợp có thể được định vị ở locus 1022859-1024215 ở nhiễm sắc thể W, cụ thể là khoảng giới hạn galGal5_dna của nhiễm sắc thể W:1022859-1024215. Theo một số phương án cụ thể khác nữa, các locus này chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 3.

Để tích hợp đặc hiệu gen thông báo theo sáng chế ở vị trí bất kỳ trong các locus được mô tả ở trên, gARN đặc hiệu có thể là cần thiết. Do đó, theo một số phương án cụ thể và không làm giới hạn, gARN thích hợp được sử dụng để tạo ra động vật thuộc loài chim chuyển gen theo sáng chế có thể chứa trình tự axit nucleic như được chỉ ra bởi trình tự bất kỳ trong số các trình tự SEQ ID NO. 1 và 2. Theo một số phương án cụ thể, các gARN này được đề cập trong bản mô tả này lần lượt là gARN1 và gARN2.

Theo một số phương án cụ thể, động vật thuộc loài chim chuyển gen được đề xuất bởi sáng chế được tạo ra bằng cách sử dụng gARN1 chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 1. Để có thể tích hợp gen thông báo theo sáng chế ở vị trí đặc hiệu này, theo các phương án nhất định, gen thông báo cần được tích hợp phải mang các nhánh cụ thể để tạo điều kiện thuận lợi cho sự tích hợp chúng ở đích vị trí tích hợp được định hướng bởi gARN được sử dụng. Do đó, theo một số phương án cụ thể, ít nhất một gen thông báo có thể được chứa trong trình tự axit nucleic thứ hai, tại đó gen thông báo này được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng chứa trình tự axit amin như được chỉ ra lần lượt bởi các trình tự SEQ ID NO. 4 và 5.

Theo một số phương án thay thế khác, động vật thuộc loài chim chuyển gen được đề xuất bởi sáng chế có thể được tạo ra bằng cách sử dụng gARN2 chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 2. Trong trường hợp này, để có thể tích hợp gen thông báo theo sáng chế ở vị trí đặc hiệu được nhận diện bởi gARN2 này, ít nhất một gen thông báo được chứa trong trình tự axit nucleic thứ hai có thể theo các phương án riêng biệt, được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng chứa trình tự axit amin như được thể hiện lần lượt bởi các trình tự SEQ ID NO. 6 và 7.

Cũng cần hiểu rằng sáng chế còn bao gồm đối tượng bất kỳ trong số các đối tượng thuộc loài chim chuyển gen như được mô tả trong bản mô tả này để sử dụng trong phương pháp bất kỳ trong số các phương pháp theo sáng chế, cụ thể là phương pháp xác định giới tính *in ovo* như được bàn luận trong bản mô tả này.

Theo một số phương án khác nữa, sáng chế còn bao gồm việc sử dụng trứng bất kỳ được đẻ bởi đối tượng bất kỳ trong số các đối tượng thuộc loài chim chuyển gen được đề xuất bởi sáng chế, cụ thể là như được bộc lộ trên đây, để sử dụng trong phương pháp bất kỳ trong số các phương pháp theo sáng chế. Theo một khía cạnh khác nữa, sáng chế đề cập đến tế bào chứa ít nhất một gen thông báo ngoại sinh được tích hợp vào ít nhất một vị trí hoặc địa điểm trong ít nhất một trong hai nhiễm sắc thể giới tính Z và W.

Theo một số phương án cụ thể, tế bào được đề xuất bởi sáng chế có thể là tế bào chim.

Theo một số phương án cụ thể, tế bào chim được đề xuất bởi sáng chế có thể là tế bào mầm nguyên thủy (PGC).

Thuật ngữ "**tế bào mầm**" dùng để chỉ tế bào phôi mà khi hợp nhất với một tế bào mầm khác phát triển thành giao tử. Khi được sử dụng trong bản mô tả này, thuật ngữ "**tế bào mầm nguyên thủy (PGC)**" dùng để chỉ tế bào gốc dòng mầm có vai trò làm tế bào tổ tiên của các giao tử và phát sinh thành tế bào gốc phôi đa năng. Tế bào trong phôi hình thành phôi vị được truyền tín hiệu để trở thành PGC trong quá trình phát sinh phôi, di chuyển vào cung sinh dục trở thành tuyến sinh dục, và phân biệt hóa thành các giao tử trưởng thành.

Theo một số phương án, gen thông báo được tích hợp trong ít nhất một nhiễm sắc thể giới tính của tế bào theo sáng chế có thể là protein phát huỳnh quang, cụ thể là RFP, YFP, BFP, và dạng tương tự hoặc protein bất kỳ trong số các protein phát huỳnh quang được bộc lộ theo Bảng 1. Theo một số phương án khác nữa, gen thông báo được tích hợp trong tế bào theo sáng chế có thể là protein phát huỳnh quang bất kỳ, với điều kiện là protein phát huỳnh quang này không phải là protein phát huỳnh quang màu xanh lục (GFP). Mặt khác, theo một số phương án cụ thể gen thông báo được tích hợp vào nhiễm sắc thể giới tính của tế bào chuyển gen được đề xuất bởi sáng chế có thể là protein phát huỳnh quang bất kỳ với điều kiện là protein phát huỳnh quang này không phải là GFP.

Theo một số phương án cụ thể hơn nữa, ít nhất một gen thông báo có thể được tích hợp vào nhiễm sắc thể giới tính của đối tượng hoặc động vật thuộc loài chim chuyển gen được đề xuất bởi tế bào theo sáng chế bằng cách sử dụng ít nhất một nucleaza được thiết kế theo chương trình (PEN). Khi được sử dụng trong bản mô tả này, thuật ngữ "**nucleaza được thiết kế theo chương trình (PEN)**" dùng để chỉ enzym tổng hợp để cắt ngắn trình tự ADN cụ thể, có nguồn gốc từ nucleaza có trong tự nhiên tham gia vào việc sửa chữa ADN của các tổn thương ADN sợi kép và có khả năng chỉnh sửa hệ gen trực tiếp. Theo phương án cụ thể hơn, PEN có thể là đoạn lặp lại đối xứng ngắn cách đều thành cụm (CRISPR) loại 2, cụ thể là hệ thống kiểu II.

Theo một số phương án khác nữa, tế bào được đề xuất bởi sáng chế có thể chứa ít nhất một gen thông báo được tích hợp vào nhiễm sắc thể giới tính của tế bào. Theo phương án cụ thể hơn, sự tích hợp đặc hiệu gen thông báo này có thể được cho phép bằng cách cho tiếp xúc hoặc đồng-chuyển nhiễm tế bào với: (a) ít nhất một hệ thống chỉnh sửa gen, chứa ít nhất một protein Cas9 hoặc ít nhất một trình tự axit nucleic thứ nhất chứa ít nhất một trình tự axit nucleic mã hóa ít nhất một protein Cas9 này; và ít nhất một gARN hướng đích ít nhất một đoạn tiền đệm trong ít nhất một nhiễm sắc thể giới tính Z và/hoặc W, hoặc ít nhất một trình tự axit nucleic mã hóa ít nhất một gARN này; và (b) ít nhất một trình tự axit nucleic thứ hai chứa ít nhất một gen thông báo này.

Theo các phương án nhất định, ít nhất một gen thông báo trong trình tự axit nucleic thứ hai được đồng-chuyển nhiễm vào tế bào theo sáng chế, có thể được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng đối với HDR ở vị trí tích hợp.

Theo một số phương án cụ thể khác, tế bào được đề xuất bởi sáng chế có thể được tạo ra bằng cách tích hợp ít nhất một gen thông báo theo sáng chế vào nhiễm sắc thể Z của tế bào. Theo các phương án nhất định, để tạo ra tế bào theo sáng chế, ít nhất một gen thông báo có thể được tích hợp vào ít nhất một vị trí ở nhiễm sắc thể giới tính Z. Theo phương án cụ thể hơn, locus đặc hiệu ở nhiễm sắc thể Z có thể là vùng bất kỳ trong số các vùng 9156874-9161874, như được thể hiện bằng trình tự SEQ ID NO:15, 27764943-27769943, như được thể hiện bằng trình tự SEQ ID NO:16, 42172748-42177748, như được thể hiện bằng trình tự SEQ ID NO:17, 63363656-63368656, như được thể hiện bằng trình tự SEQ ID NO:18 và 78777477-78782477, như được thể hiện bằng trình tự SEQ ID NO:19 của nhiễm sắc thể Z của gà mái.

Theo một số phương án thay thế khác nữa, tế bào được đề xuất bởi sáng chế có thể được tạo ra bằng cách sử dụng gARN tạo điều kiện thuận lợi cho sự tích hợp vào ít nhất một vị trí ở locus nhiễm sắc thể giới tính Z 42172748-42177748, cụ thể là SEQ ID NO. 17. Theo một số phương án cụ thể khác nữa, trình tự gARN phù hợp cho sự tích hợp vào locus đặc hiệu trong nhiễm sắc thể Z, có thể bao gồm nhưng không chỉ giới hạn ở, gARN7 của locus nhiễm sắc thể Z chrZ_42174515_-1, chứa trình tự axit nucleic GTAATACAGAGCTAAACCAG, cũng như được thể hiện bằng trình tự SEQ ID NO:26. Theo một số phương án khác nữa, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 31, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 32, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN7 trong trình tự SEQ ID NO:26.

Theo phương án cụ thể hơn, ít nhất một gARN cần thiết để hướng đích gen thông báo đến vị trí đặc hiệu này trong nhiễm sắc thể Z của tế bào theo sáng chế có thể chứa trình tự axit nucleic như được chỉ ra bởi gARN bất kỳ trong số gARN3: ACAGACCTATGATATGT, như được thể hiện bằng trình tự SEQ ID NO. 11; gARN4: CGATTATCACTCACAAG, như được thể hiện bằng trình tự SEQ ID NO. 12; gARN5: CTGGTTAGCATGGGGAC, như được thể hiện bằng trình tự SEQ ID NO. 13; gARN6: GTAAAGAGTCAGATACA, như được thể hiện bằng trình tự SEQ ID NO. 14.

Theo một số phương án khác nữa, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z của tế bào theo sáng chế, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 41, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 42, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN3 trong trình tự SEQ ID NO:11. Theo các phương án khác, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 43, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 44, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN4 trong trình tự SEQ ID NO:12. Theo các phương án khác

nữa, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z của tế bào theo sáng chế, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 45, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 46, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN5 trong trình tự SEQ ID NO:13. Theo một số phương án khác, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 47, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 48, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN6 trong trình tự SEQ ID NO:14.

Các ví dụ không làm giới hạn khác đối với trình tự gARN phù hợp cho sự tích hợp vào locus đặc hiệu trong nhiễm sắc thể Z có thể bao gồm, nhưng không chỉ giới hạn ở, gARN8 của locus nhiễm sắc thể Z chrZ_9157091_1, chứa trình tự axit nucleic ACAGACCTATGATATGTGAG, cũng như được thể hiện bằng trình tự SEQ ID NO:27, gARN9 của locus nhiễm sắc thể Z chrZ_27767602_-1, chứa trình tự axit nucleic GAGCTTGTGAGTGATAATCG, cũng như được thể hiện bằng trình tự SEQ ID NO:28, gARN10 của locus nhiễm sắc thể Z chrZ_78779927_1, chứa trình tự axit nucleic GTAAAGAGTCAGATACACAG, cũng như được thể hiện bằng trình tự SEQ ID NO: 29, và gARN11 của locus nhiễm sắc thể Z chrZ_63364946_-1, chứa trình tự axit nucleic CAGTGGGTACTGAAGCTGTG cũng như được thể hiện bằng trình tự SEQ ID NO: 30.

Theo các phương án khác, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z của tế bào theo sáng chế, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 33, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 34, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN8 trong trình tự SEQ ID NO:27. Theo các phương án khác nữa, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z của tế bào theo sáng chế, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 35, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 36, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào

locus đặc hiệu được định hướng bởi gARN9 trong trình tự SEQ ID NO:28. Theo các phương án khác, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 37, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 38, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN10 trong trình tự SEQ ID NO:29. Theo phương án khác nữa, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 39, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 40, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN11 trong trình tự SEQ ID NO:30.

Theo một số phương án cụ thể, để hướng đích sự tích hợp gen thông báo đến nhiễm sắc thể W ở tế bào được đề xuất bởi sáng chế, gARN đặc hiệu được sử dụng. Theo các phương án cụ thể khác, gARN có thể chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 1 được đề cập đến trong bản mô tả này dưới dạng gARN1. Trong trường hợp này, ít nhất một gen thông báo được chứa trong trình tự axit nucleic thứ hai này có thể được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng chứa trình tự axit amin như được thể hiện lần lượt bằng các trình tự SEQ ID NO. 4 và 5.

Theo một số phương án thay thế khác nữa, tế bào được đề xuất bởi sáng chế có thể được tạo ra bằng cách sử dụng gARN được đề cập trong bản mô tả này là gARN2. Theo các phương án nhất định, gARN2 có thể chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 2. Theo các phương án cụ thể này, ít nhất một gen thông báo được chứa trong trình tự axit nucleic thứ hai có thể được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng chứa trình tự axit amin như được thể hiện lần lượt bằng các trình tự SEQ ID NO. 6 và 7.

Ngoài ra, sáng chế đề xuất phương pháp để phát hiện sự thụ tinh của trứng chưa nở. Theo một số phương án cụ thể, phương pháp này có thể bao gồm bước (a), cung cấp hoặc thu được ít nhất một đối tượng hoặc động vật thuộc loài chim chuyển gen chứa ít nhất một gen thông báo ngoại sinh được tích hợp vào ít nhất một vị trí hoặc địa điểm ở trong cả hai nhiễm sắc thể giới tính Z và W, trong trường hợp động vật chuyển

gen được cung cấp là đối tượng con cái và ở trong cả hai nhiễm sắc thể giới tính Z trong trường hợp động vật chuyển gen được cung cấp là con đực. Trong bước thứ hai (b) thu được ít nhất một trứng đã thụ tinh từ đối tượng, cụ thể là động vật, thuộc loài chim chuyển gen hoặc tế bào bất kỳ của chúng. Bước tiếp theo (c) bao gồm việc xác định trong trứng liệu ít nhất một tín hiệu có thể phát hiện được phát hiện thấy. Theo phương án cụ thể hơn, sự phát hiện ít nhất một tín hiệu có thể phát hiện chỉ ra sự biểu hiện của ít nhất một gen thông báo này, nhờ đó chỉ ra có mặt của nhiễm sắc thể W hoặc nhiễm sắc thể Z mẹ được đánh dấu (trong trường hợp là con cái) hoặc nhiễm sắc thể Z bố được đánh dấu ở phôi chim.

Theo một số phương án khác nữa, động vật chuyển gen được tạo ra bởi phương pháp theo sáng chế có thể là đối tượng con đực có gen thông báo được tích hợp vào cả hai nhiễm sắc thể Z của chúng. Trong trường hợp này, tín hiệu có thể phát hiện được được xác định trong trứng được thụ tinh bởi con đực chuyển gen này hoặc tinh dịch bất kỳ của chúng, biểu thị rằng phôi mang nhiễm sắc thể Z ở bố có chứa gen thông báo chuyển gen, và do đó trứng được thụ tinh. Theo các phương án khác nữa, động vật chuyển gen được tạo ra bởi phương pháp theo sáng chế có thể là đối tượng con cái có gen thông báo được tích hợp vào cả hai, các nhiễm sắc thể W và Z của chúng. Trong trường hợp này, tín hiệu có thể phát hiện được được xác định trong trứng được thực hiện bởi con cái chuyển gen này, biểu thị rằng phôi mang nhiễm sắc thể W hoặc Z ở mẹ có chứa gen thông báo chuyển gen, và do đó trứng được thụ tinh. Cần hiểu rõ theo các phương án khác, để phát hiện trứng đã thụ tinh, gen thông báo, cụ thể là gen RFP có thể được chèn vào hai bản sao của nhiễm sắc thể bất kỳ, để tạo ra đối tượng thuộc loài chim chuyển gen đồng hợp tử, cụ thể là đối tượng con đực. Trứng đã thụ tinh được phát hiện là mang gen thông báo RFP.

Theo khía cạnh khác nữa, sáng chế đề xuất kit chứa:

(a) ít nhất một thành phần của ít nhất một hệ thống chỉnh sửa gen, cụ thể là hệ thống CRISPR/Cas9 có thể chứa ít nhất một protein Cas9 hoặc ít nhất một trình tự axit nucleic thứ nhất chứa ít nhất một trình tự axit nucleic mã hóa ít nhất một protein Cas9 này; và ít nhất một gARN hướng đích ít nhất một đoạn tiền đệm trong ít nhất một nhiễm sắc thể giới tính Z và/hoặc W, hoặc ít nhất một trình tự axit nucleic mã hóa ít nhất một gARN này; và (b) ít nhất một trình tự axit nucleic thứ hai chứa ít nhất một gen thông báo này.

Theo một số phương án khác, gARN được sử dụng bởi phương pháp cũng như bởi kit theo sáng chế (được mô tả dưới đây) có thể chứa ít nhất một ARN CRISPR (crARN) và ít nhất một crARN hoạt hóa trans (tracrARN).

Theo một số phương án thay thế, kit theo sáng chế có thể chứa trình tự axit nucleic mã hóa ít nhất một gARN. Trình tự axit nucleic này có thể bao gồm mảng CRISPR chứa ít nhất một trình tự đoạn đệm hướng đích và do đó tương đồng hoặc đồng nhất với ít nhất một đoạn tiền đệm ở trình tự ADN của hệ gen đích. Cần lưu ý rằng trình tự axit nucleic có thể còn chứa trình tự mã hóa ít nhất một tracrARN.

Theo một số phương án, mảng CRISPR theo sáng chế chứa ít nhất một đoạn đệm và ít nhất một đoạn lặp. Theo một phương án khác, sáng chế còn bao gồm sự lựa chọn của việc cung cấp tiền-crARN có thể được xử lý thành một số sản phẩm gARN cuối để có thể hướng đích đến các đích giống hoặc khác nhau.

Theo một số phương án cụ thể hơn nữa, crARN được chứa trong gARN theo sáng chế có thể là trình tự axit ribonucleic sợi đơn (ssARN) hỗ trợ cho trình tự ADN của hệ gen đích. Theo một số phương án cụ thể, trình tự ADN của hệ gen đích có thể được định vị ngay phía trước trình tự motif gần kề đoạn tiền đệm (PAM) và thêm nữa.

Như được chỉ ra trong bản mô tả này, gARN của kit và phương pháp theo sáng chế có thể hỗ trợ, ít nhất một phần, cho ADN của hệ gen đích, trong trường hợp này, locus đích trong nhiễm sắc thể giới tính Z hoặc W, được đề cập trong bản mô tả này dưới dạng "đoạn tiền đệm". Theo các phương án nhất định, "**sự hỗ trợ**" dùng để chỉ mối quan hệ giữa hai cấu trúc mà mỗi cấu trúc này tuân theo nguyên tắc ổ-và-khóa. Sự hỗ trợ tự nhiên là nguyên tắc cơ bản của sự sao chép và phiên mã ADN vì nó là tính chất được dùng chung giữa hai trình tự ADN hoặc ARN, sao cho khi chúng được đóng thẳng hàng đối song song với nhau, các bazơ nucleotit ở mỗi vị trí trong trình tự sẽ có tính hỗ trợ (ví dụ, A và T hoặc U, C và G).

Như được biểu thị ở trên, trình tự ADN của hệ gen được hướng đích bởi gARN của kit theo sáng chế được định vị ngay phía trước trình tự PAM. Theo một số phương án, trình tự PAM này có thể là trình tự axit nucleic NGG.

Theo các phương án nhất định, trình tự PAM được đề cập bởi sáng chế có thể chứa N, tức là nucleotit bất kỳ, cụ thể là một trong số Adenin (A), Guanin (G), Xytosin (C) hoặc Thymin (T). Theo một số phương án khác nữa, trình tự PAM theo sáng chế gồm có A, G, C, hoặc T và hai Guanin.

Theo một phương án, polynucleotit mã hóa gARN theo sáng chế có thể chứa ít nhất một đoạn đệm và tùy ý, ít nhất một đoạn lặp. Theo một số phương án khác nữa, ADN mã hóa gARN theo sáng chế có thể chứa ít nhất là 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 đoạn đệm hoặc nhiều hơn, cụ thể là 110, 120, 130, 140, 150, 160, 170, 180, 190, 200 đoạn đệm hoặc nhiều hơn. Theo một số phương án, mỗi đoạn đệm được định vị giữa hai đoạn lặp. Cần hiểu thêm rằng đoạn đệm của trình tự axit nucleic mã hóa gARN theo sáng chế có thể là các đoạn đệm giống hoặc khác nhau. Theo nhiều phương án, các đoạn đệm này có thể hướng đích ADN của hệ gen đích giống hoặc khác nhau. Theo một số phương án khác nữa, đoạn đệm này có thể hướng đích ít nhất là 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 trình tự ADN của hệ gen đích hoặc nhiều hơn. Các trình tự đích này có thể có nguồn gốc từ locus đơn lẻ hoặc theo cách khác là, từ một số locus đích.

Khi được sử dụng trong bản mô tả này, thuật ngữ "**đoạn đệm**" dùng để chỉ trình tự đoạn đệm không có tính lặp được thiết kế để hướng đích trình tự đặc hiệu và được định vị giữa nhiều đoạn lặp trực tiếp ngắn (tức là, đoạn lặp CRISPR) của mảng CRISPR. Theo một số phương án cụ thể, đoạn đệm có thể chứa từ khoảng 15 đến khoảng 30 nucleotit, cụ thể là khoảng 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 nucleotit hoặc nhiều hơn. Cụ thể hơn là khoảng 20-25 nucleotit.

ARN dẫn đường hoặc ARN hướng đích được mã hóa bởi hệ thống CRISPR theo sáng chế có thể chứa ARN CRISPR (crARN) và ARN hoạt hóa trans (tracrARN). Trình tự của ARN hướng đích được mã hóa bởi đoạn đệm CRISPR không bị giới hạn cụ thể, ngoài việc yêu cầu nó cần được định hướng (tức là, có phân đoạn giống hoặc bổ sung với) trình tự đích ở ADN hệ gen của chim, cụ thể là ở locus đích trong nhiễm sắc thể giới tính Z hoặc W, mà còn được gọi trong bản mô tả này là "**đoạn tiền đệm**". Các đoạn tiền đệm này chứa trình tự axit nucleic có tính hỗ trợ đủ cho ARN hướng

đích được mã hóa bởi đoạn đệm CRISPR được chứa trong trình tự axit nucleic mã hóa gARN của phương pháp và kit theo sáng chế.

Theo một số phương án, crARN chứa hoặc gồm có 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39 hoặc 40 nt của trình tự đoạn đệm (hướng đích) sau đó là 19-36 nt của trình tự đoạn lặp. Theo các phương án riêng biệt và không làm giới hạn, đoạn đệm hướng đích có thể chứa hoặc gồm có phân đoạn hướng đích trình tự bất kỳ trong số các trình tự ADN hệ gen mà tại đó trình tự đoạn đệm đại diện được chỉ ra trong bản mô tả này.

Cần lưu ý rằng theo một số phương án cụ thể, các đoạn đệm của hệ thống CRISPR theo sáng chế có thể mã hóa ARN dẫn đường hướng đích (gARN). "gARN" hoặc "ARN hướng đích" là ARN, khi được phiên mã từ một phần của hệ thống CRISPR mã hóa nó, chứa ít nhất một phân đoạn của trình tự ARN tương đồng với (ngoại trừ việc thay thế T cho U trong trường hợp của ARN) hoặc bổ sung với (và do đó "hướng đích") trình tự ADN ở ADN của hệ gen đích. Hệ thống CRISPR theo sáng chế có thể tùy ý mã hóa nhiều hơn một ARN hướng đích, và ARN hướng đích được định hướng đến một hoặc nhiều trình tự đích ở ADN hệ gen.

Theo một số phương án, protein Cas9 được cung cấp bởi kit theo sáng chế có thể được mã hóa bởi trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 24. Theo các phương án cụ thể và không giới hạn khác, protein Cas9 có thể chứa trình tự axit amin như được thể hiện bằng trình tự SEQ ID NO. 25, hoặc các dẫn xuất, thể đột biến hoặc biến thể bất kỳ của chúng.

Cần lưu ý rằng "trình tự axit amin" hoặc "trình tự peptit" là thứ tự trong đó các gốc axit amin được liên kết bằng liên kết peptit, nằm trong chuỗi trong peptit và protein. Trình tự thường được thông báo từ đầu tận cùng-N chứa nhóm amin tự do đến đầu tận cùng-C chứa amit. Trình tự axit amin thường được gọi là trình tự peptit, protein nếu nó thể hiện cấu trúc bậc một của protein, tuy nhiên phải phân biệt giữa các thuật ngữ "trình tự axit amin" hoặc "trình tự peptit" và "protein", bởi vì protein được xác định là trình tự axit amin được cuộn gấp thành cấu hình ba chiều riêng biệt và thường đã cải biến sau dịch mã, chẳng hạn như phosphoryl hóa, axetyl hóa, glycosyl hóa, manosyl hóa, amit hóa, carboxyl hóa, sự tạo thành liên kết sulfhydryl, sự phân cắt và quá trình tương tự.

Thuật ngữ "mảnh hoặc peptit" có nghĩa là phân đoạn của các phân tử Cas9 hoặc

RFP (được mô tả dưới đây). "Mảnh" của phân tử, chẳng hạn như trình tự bất kỳ trong số các trình tự axit amin theo sáng chế, có nghĩa là dùng để chỉ tập hợp con axit amin bất kỳ của các phân tử Cas9 hoặc RFP. Nó cũng có thể bao gồm "biến thể" hoặc "dẫn xuất" của chúng. "Peptit" có nghĩa là dùng để chỉ tập hợp con axit amin cụ thể có hoạt tính chức năng. Thuật ngữ "chức năng" có nghĩa là có cùng chức năng sinh học, ví dụ, có khả năng thực hiện việc chỉnh sửa gen như Cas 9 hoặc là tạo ra tín hiệu có thể phát hiện được như RFP.

Cần hiểu rõ sáng chế bao gồm biến thể hoặc dẫn xuất bất kỳ của các phân tử RFP hoặc Cas9 theo sáng chế và các polypeptit bất kỳ về cơ bản là tương đồng hoặc đồng nhất. Thuật ngữ "*dẫn xuất*" được sử dụng để xác định các trình tự axit amin (polypeptit), có sự chèn đoạn, làm khuyết đoạn, thay thế và cải biến đối với trình tự axit amin (polypeptit) mà không làm thay đổi hoạt tính của các polypeptit ban đầu. Về vấn đề này, dẫn xuất hoặc mảnh của các phân tử Cas9 hoặc RFP theo sáng chế có thể là dẫn xuất hoặc mảnh bất kỳ của các phân tử Cas9 hoặc RFP, cụ thể là như được thể hiện bằng trình tự SEQ ID NO. 25, 21, 23, mà không làm giảm hoặc làm thay đổi hoạt tính của các phân tử Cas9 hoặc RFP. Thuật ngữ "dẫn xuất" còn được gọi là thể tương đồng, biến thể và thể tương tự của chúng. Thể đồng đẳng hoặc thể tương đồng của protein có độ đồng nhất hoặc tương đồng trình tự đối với các protein mong muốn theo sáng chế, cụ thể là có thể có chung ít nhất là 50%, ít nhất là 60% và cụ thể là 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, hoặc cao hơn, cụ thể là khi so sánh với toàn bộ trình tự của protein mong muốn theo sáng chế, ví dụ, protein bất kỳ trong số các protein chứa trình tự axit amin như được thể hiện bằng trình tự SEQ ID NO. 25, 21, 23. Cụ thể là thể tương đồng chứa hoặc bao gồm trình tự axit amin có mức độ giống ít nhất là 50%, ít nhất là 60% và cụ thể là 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, hoặc cao hơn so với trình tự SEQ ID NO. 21, 23, 25 cụ thể là toàn bộ trình tự như được thể hiện bằng các trình tự SEQ ID NO. 21, 23, 25.

Theo một số phương án, dẫn xuất dùng để chỉ các polypeptit, khác với các polypeptit được xác định cụ thể theo sáng chế bởi sự chèn đoạn, làm khuyết đoạn hoặc thay thế các gốc axit amin. Cần hiểu rằng khi được sử dụng trong bản mô tả này, các thuật ngữ "**sự chèn đoạn**", "**làm khuyết đoạn**" hoặc "**thay thế đoạn**", có nghĩa tương ứng là sự bổ sung, làm khuyết đoạn hoặc thay thế bất kỳ, của các gốc axit amin đối với

polypeptit được bộc lộ bởi sáng chế, trong khoảng từ 1 đến 50 gốc axit amin, trong khoảng từ 20 đến 1 gốc axit amin, và cụ thể là trong khoảng từ 1 đến 10 gốc axit amin. Cụ thể hơn, sự chèn đoạn, làm khuyết đoạn hoặc thay thế đoạn có thể là 1, 2, 3, 4, 5, 6, 7, 8, 9, hoặc 10 axit amin. Cần lưu ý rằng sự chèn đoạn, làm khuyết đoạn hoặc thay thế đoạn được bao gồm bởi sáng chế có thể xảy ra ở vị trí bất kỳ của peptit cải biến, cũng như là ở đầu tận cùng bất kỳ trong số các đầu tận cùng N' hoặc C' của chúng.

Đối với trình tự axit amin, người có hiểu biết trung bình trong lĩnh vực này sẽ biết rằng sự thay thế, làm khuyết đoạn hoặc bổ sung axit nucleic, peptit, polypeptit, hoặc trình tự protein làm thay đổi, bổ sung hoặc làm khuyết axit amin đơn lẻ hoặc tỷ lệ phần trăm nhỏ của axit amin ở trình tự mã hóa là “biến thể cải biến bảo toàn” trong đó sự thay đổi dẫn đến sự thể axit amin bằng axit amin tương tự về mặt hóa học. Các bằng thể bảo toàn cung cấp các axit amin tương tự về mặt chức năng đã được biết rõ trong lĩnh vực này. Các biến thể được cải biến bảo toàn này là ngoài ra và không loại trừ biến thể đa hình, thể tương đồng giữa các loài, và các alen theo sáng chế.

Ví dụ, sự thể có thể được tạo ra trong đó axit amin béo (G, A, I, L, hoặc V) được thể bằng thành viên khác của nhóm, hoặc sự thể chẳng hạn như sự thể của một gốc phân cực cho thành viên khác, chẳng hạn như arginin cho lysin, glutamic cho axit aspartic, hoặc glutamin cho asparagin. Mỗi nhóm trong tám nhóm sau đây chứa các axit amin làm ví dụ khác là sự thể bảo toàn đối với nhau:

- 1) Alanin (A), Glyxin (G);
- 2) Axit aspartic (D), Axit glutamic (E);
- 3) Asparagin (N), Glutamin (Q);
- 4) Arginin (R), Lysin (K);
- 5) Isoleuxin (I), Leuxin (L), Methionin (M), Valin (V);
- 6) Phenylalanin (F), Tyrosin (Y), Tryptophan (W);
- 7) Serin (S), Threonin (T); và
- 8) Xystein (C), Methionin (M).

Do đó, theo một số phương án, sáng chế bao gồm các phân tử Cas9 hoặc RFP hoặc dẫn xuất bất kỳ của chúng, cụ thể là dẫn xuất chứa ít nhất là 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 sự thay thế bảo toàn hoặc nhiều hơn đối với trình tự axit amin như được thể hiện bằng trình tự bất kỳ trong số các trình tự SEQ ID NO. 21, 23, 25. Cụ thể hơn, “sự thể” axit amin là kết quả của việc thay thế một axit amin bằng một axit amin khác có các

tính chất về mặt cấu trúc và/hoặc hóa học tương tự, tức là sự thay thế axit amin bảo toàn. Sự thế axit amin có thể được thực hiện trên cơ sở tính tương tự về độ phân cực, điện tích, tính hòa tan, tính kỵ nước, tính ưa nước, và/hoặc bản chất lưỡng tính của các gốc tham gia. Ví dụ, axit amin “**kỵ nước**” không phân cực được chọn từ nhóm bao gồm Valin (V), Isoleuxin (I), Leuxin (L), Methionin (M), Phenylalanin (F), Tryptophan (W), Xystein (C), Alanin (A), Tyrosin (Y), Histidin (H), Threonin (T), Serin (S), Prolin (P), Glyxin (G), Arginin (R) và Lysin (K); axit amin “**phân cực**” được chọn từ nhóm bao gồm Arginin (R), Lysin (K), axit aspartic (D), axit glutamic (E), Asparagin (N), Glutamin (Q); axit amin “**mang điện tích dương**” được chọn từ nhóm bao gồm Arginin (R), Lysin (K) và Histidin (H) và trong đó axit amin “**có tính axit**” được chọn từ nhóm bao gồm axit aspartic (D), Asparagin (N), axit glutamic (E) và Glutamin (Q).

Biến thể của polypeptit theo sáng chế có thể có tính tương tự hoặc tính đồng nhất trình tự ít nhất là 80%, thông thường tính tương tự hoặc tính đồng nhất trình tự ít nhất là 85%, tính tương tự hoặc tính đồng nhất trình tự là 90%, hoặc tính tương tự hoặc tính đồng nhất trình tự ít nhất là 95%, 96%, 97%, 98%, hoặc 99% ở cấp độ axit amin, với protein quan tâm, chẳng hạn như các polypeptit khác nhau theo sáng chế.

Theo một số phương án, ít nhất một gen thông báo ở trình tự axit nucleic thứ hai được chứa trong kit theo sáng chế, có thể được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng đối với HDR ở vị trí tích hợp.

Theo một số phương án cụ thể khác nữa, ít nhất một gen thông báo ở trình tự axit nucleic thứ hai của kit theo sáng chế có thể được liên kết khi hoạt động với gen bất kỳ trong số gen khởi đầu đặc hiệu giới tính, gen khởi đầu đặc hiệu phôi và gen khởi đầu có thể cảm ứng.

Theo một số phương án thay thế khác nữa, ít nhất một gen thông báo có thể được tích hợp vào ít nhất một vị trí ở nhiễm sắc thể giới tính Z. Theo phương án cụ thể hơn, locus đặc hiệu ở nhiễm sắc thể Z có thể là vùng bất kỳ trong số các vùng 9156874-9161874, như được thể hiện bằng trình tự SEQ ID NO:15, 27764943-27769943, như được thể hiện bằng trình tự SEQ ID NO:16, 42172748-42177748, như được thể hiện bằng trình tự SEQ ID NO:17, 63363656-63368656, như được thể hiện bằng trình tự SEQ ID NO:18 và 78777477-78782477, như được thể hiện bằng trình tự SEQ ID NO:19 của nhiễm sắc thể Z của gà mái.

Theo một số phương án cụ thể, ít nhất một gen thông báo có thể được tích hợp vào ít nhất một vị trí ở locus nhiễm sắc thể giới tính Z 42172748-42177748. Các ví dụ không làm giới hạn về trình tự axit nucleic thứ nhất của kit theo sáng chế có thể chứa gARN, ít nhất một trong số gARN7 của locus nhiễm sắc thể Z chrZ_42174515_-1, chứa trình tự axit nucleic GTAATACAGAGCTAAACCAG, cũng như được thể hiện bằng trình tự SEQ ID NO:26. Cụ thể hơn, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 31, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 32, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN7 trong trình tự SEQ ID NO:26.

Do đó, theo các phương án cụ thể hơn, trình tự axit nucleic thứ nhất của kit theo sáng chế có thể chứa gARN, ít nhất một trong số gARN3: ACAGACCTATGATATGT, như được thể hiện bằng trình tự SEQ ID NO. 11; gARN4: CGATTATCACTCACAAG, như được thể hiện bằng trình tự SEQ ID NO. 12; gARN5: CTGGTTAGCATGGGGAC, như được thể hiện bằng trình tự SEQ ID NO. 13; gARN6: GTAAAGAGTCAGATACA, như được thể hiện bằng trình tự SEQ ID NO. 14.

Theo các phương án khác, ít nhất một gen thông báo được chứa trong trình tự axit nucleic thứ hai của kit theo sáng chế, có thể được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng chứa trình tự axit amin như được thể hiện bằng trình tự bất kỳ trong số các trình tự SEQ ID NO. từ 41 đến 48. Cụ thể hơn, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 41, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 42, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN3 trong trình tự SEQ ID NO:11. Theo các phương án khác, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 43, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 44, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN4 trong trình tự SEQ ID NO:12. Theo các phương án khác

nữa, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 45, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 46, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN5 trong trình tự SEQ ID NO:13. Theo một số phương án khác, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 47, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 48, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN6 trong trình tự SEQ ID NO:14.

Các ví dụ không làm giới hạn khác về trình tự axit nucleic thứ nhất của kit theo sáng chế có thể chứa gARN, ít nhất một trong số gARN8 của locus nhiễm sắc thể Z chrZ_9157091_1, chứa trình tự axit nucleic ACAGACCTATGATATGTGAG, cũng như được thể hiện bằng trình tự SEQ ID NO:27, gARN9 của locus nhiễm sắc thể Z chrZ_27767602_-1, chứa trình tự axit nucleic GAGCTTGTGAGTGATAATCG, cũng như được thể hiện bằng trình tự SEQ ID NO:28, gARN10 của locus nhiễm sắc thể Z chrZ_78779927_1, chứa trình tự axit nucleic GTAAAGAGTCAGATACACAG, cũng như được thể hiện bằng trình tự SEQ ID NO: 29, và gARN11 của locus nhiễm sắc thể Z chrZ_63364946_-1, chứa trình tự axit nucleic CAGTGGGTACTGAAGCTGTG cũng như được thể hiện bằng trình tự SEQ ID NO: 30.

Theo các phương án khác, ít nhất một gen thông báo được chứa trong trình tự axit nucleic thứ hai của kit theo sáng chế, có thể được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng chứa trình tự axit amin như được thể hiện bằng trình tự bất kỳ trong số các trình tự SEQ ID NO. từ 31 đến 40. Cụ thể hơn, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 33, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 34, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN8 trong trình tự SEQ ID NO:27. Theo các phương án khác nữa, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO.

35, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 36, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN9 trong trình tự SEQ ID NO:28. Theo các phương án khác, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 37, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 38, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN10 trong trình tự SEQ ID NO:29. Theo phương án khác nữa, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 39, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 40, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN11 trong trình tự SEQ ID NO:30.

Theo các phương án nhất định, ít nhất một gen thông báo được cung cấp bởi kit theo sáng chế có thể được tích hợp vào ít nhất một vùng không mã hóa của nhiễm sắc thể giới tính, cụ thể là vào nhiễm sắc thể W. Trong trường hợp này, trình tự axit nucleic thứ nhất của kit theo sáng chế có thể mã hóa ít nhất một gARN chứa trình tự axit nucleic như được thể hiện bằng trình tự bất kỳ trong số các trình tự SEQ ID NO. 1 và 2, còn được đề cập đến trong bản mô tả này lần lượt là gARN1 và gARN2.

Theo một số phương án cụ thể, trình tự axit nucleic thứ nhất của kit theo sáng chế có thể chứa gARN, là gARN1. Theo một số phương án, gARN1 này có thể chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 1. Trong trường hợp này, gen thông báo được chứa trong trình tự axit nucleic thứ hai này của kit theo sáng chế, có thể được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng chứa trình tự axit amin như được thể hiện bằng các trình tự SEQ ID NO. 4 và 5.

Theo một số phương án thay thế khác nữa, kit theo sáng chế có thể chứa trong trình tự axit nucleic thứ nhất của chúng, trình tự mã hóa gARN2. Theo một số phương án cụ thể, trình tự này mã hóa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 2. Theo một số phương án khác nữa, ít nhất một gen thông báo được chứa trong trình tự axit nucleic thứ hai của kit theo sáng chế, có thể được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng chứa trình tự axit amin như được

thể hiện bằng các trình tự SEQ ID NO. 6 và 7 tương ứng.

Theo một số phương án, gen thông báo được chứa trong trình tự axit nucleic thứ hai của kit theo sáng chế có thể ít nhất một gen thông báo phát huỳnh quang. Ngoài ra, theo một số phương án, gen thông báo được cung cấp bởi kit theo sáng chế có thể là protein phát huỳnh quang bất kỳ, cụ thể là RFP, YFP và protein phát huỳnh quang bất kỳ trong số các protein phát huỳnh quang được bộc lộ bởi sáng chế hoặc protein phát huỳnh quang bất kỳ trong số các protein phát huỳnh quang được bộc lộ trong Bảng 1. Theo một số phương án, gen thông báo được sử dụng bởi kit theo sáng chế có thể là protein phát huỳnh quang bất kỳ, với điều kiện là protein phát huỳnh quang này không phải là protein phát huỳnh quang màu xanh lục (GFP). Mặt khác, theo một số phương án cụ thể, gen thông báo được tích hợp vào nhiễm sắc thể giới tính của động vật chuyển gen được cung cấp bởi kit theo sáng chế có thể là protein phát huỳnh quang bất kỳ với điều kiện là protein phát huỳnh quang này không phải là GFP. Theo một số phương án khác nữa, gen thông báo có thể là RFP. Theo phương án cụ thể hơn, RFP này có thể chứa trình tự axit amin như được thể hiện bằng trình tự SEQ ID NO. 21, hoặc thể tương đồng, biến thể, thể đột biến hoặc dẫn xuất bất kỳ của chúng. Theo một số phương án cụ thể thay thế khác, RFP được sử dụng bởi sáng chế có thể là RFP có thể chứa trình tự axit amin như được thể hiện bằng trình tự SEQ ID NO. 23, hoặc thể tương đồng, thể đột biến hoặc dẫn xuất bất kỳ của chúng.

Theo một số phương án khác nữa, kit theo sáng chế có thể phù hợp để sử dụng trong việc tạo ra động vật thuộc loài chim chuyển gen chứa ít nhất một gen thông báo ngoại sinh được tích hợp vào ít nhất một vị trí hoặc địa điểm ở trong ít nhất một trong hai nhiễm sắc thể giới tính Z và W.

Theo một số phương án, phương pháp theo sáng chế có thể sử dụng kit bất kỳ trong số các kit theo sáng chế như được mô tả trong bản mô tả này.

Ngoài ra, cần hiểu rằng kit theo sáng chế có thể còn có chứa chất phản ứng, chất đệm, môi trường hoặc nguyên liệu bất kỳ cần thiết cho sự tạo ra động vật thuộc loài chim chuyển gen theo sáng chế. Kit theo sáng chế có thể còn bao gồm hướng dẫn sử dụng cũng như đồ chứa đối với các thành phần khác nhau của chúng.

Theo một số phương án, kit theo sáng chế có thể được sử dụng trong việc tạo ra đối tượng bất kỳ trong các đối tượng thuộc loài chim chuyển gen theo sáng chế hoặc tế bào bất kỳ của chúng theo sáng chế. Theo một số phương án khác nữa, đối tượng và tế

bào chim chuyển gen được tạo ra bằng cách sử dụng kit theo sáng chế, có thể được sử dụng theo phương pháp bất kỳ trong các phương pháp theo sáng chế, bằng cách sử dụng hệ thống và thiết bị bất kỳ trong các hệ thống và thiết bị như được bộc lộ trong bản mô tả này. Ngoài ra, theo một số phương án, kit theo sáng chế có thể còn có chứa dụng cụ, hệ thống, và/hoặc thiết bị bất kỳ được làm thích ứng để xác định sự có mặt của gen thông báo trong phôi *in ovo* được kiểm tra bằng phương pháp theo sáng chế. Theo một số phương án cụ thể, dụng cụ hoặc thiết bị này có thể là thiết bị bất kỳ trong số các thiết bị được bộc lộ bởi sáng chế, cụ thể là như được thể hiện trên Fig.1.

Theo khía cạnh khác nữa của chúng, sáng chế đề xuất hệ thống, dụng cụ hoặc thiết bị có thể được làm đáp ứng theo cách đặc hiệu để phát hiện gen thông báo trong phôi *in ovo*. Theo một số phương án, thiết bị theo sáng chế có thể chứa nguồn laze, bộ đỡ trứng, thấu kính, bộ lọc, bộ đỡ thiết bị phát hiện và thiết bị phát hiện. Thiết bị theo sáng chế có thể tùy ý, theo một số phương án, có chứa giá đỡ rắn giữ tất cả các thành phần của hệ thống theo sáng chế, ví dụ, như được thể hiện trên Fig.1. Khi được sử dụng trong bản mô tả này, thuật ngữ "**thiết bị phát hiện**" dùng để chỉ loại thiết bị bất kỳ phát hiện và/hoặc đo ánh sáng.

Theo một số phương án, cần lưu ý rằng tín hiệu có thể phát hiện được, cụ thể là tín hiệu huỳnh quang có thể được phát hiện bằng thiết bị theo sáng chế bằng cách sử dụng phương tiện huỳnh quang thích hợp. Theo một số phương án, tín hiệu có thể phát hiện được được tạo ra bởi gen thông báo RFP có thể được phát hiện bằng dụng cụ nhạy với ánh sáng chẳng hạn như kính hiển vi quang học được cải biến hoặc thiết bị tích điện kép (CCD), thiết bị phát hiện photon có độ nhạy cao.

Theo một số phương án khác nữa, thiết bị, dụng cụ hoặc hệ thống theo sáng chế có thể chứa ít nhất một nguồn sáng, ít nhất một thiết bị phát hiện, ít nhất một bộ lọc và ít nhất một tay giữ đặt trứng ở vị trí thích hợp tạo thuận lợi cho sự tiếp xúc tế bào biểu hiện gen thông báo theo sáng chế với nguồn sáng này. Theo một số phương án, trứng có thể là trứng chim chưa nở đã được thụ tinh. Theo một số phương án khác nữa, trứng có thể được đẻ bởi đối tượng bất kỳ trong các đối tượng thuộc loài chim chuyển gen theo sáng chế. Theo một số phương án cụ thể và không làm giới hạn, thiết bị theo sáng chế có thể là như được minh họa trên Fig.1. Theo một số phương án khác nữa, khoảng cách giữa bộ đỡ trứng và thấu kính và/hoặc thấu kính và thiết bị phát hiện có thể nằm trong khoảng từ khoảng 1 đến 100 cm, cụ thể là 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50,

55, 60, 65, 70, 75, 80, 85, 90, 95 và 100 cm, cụ thể là nằm trong khoảng từ khoảng 5 đến 25 cm, cụ thể hơn là khoảng 20 cm.

Do đó, theo một số phương án, cần hiểu rằng hệ thống hoặc thiết bị theo sáng chế có thể được làm thích ứng cho việc thực hiện phương pháp bất kỳ trong số các phương pháp được bộc lộ bởi sáng chế.

Theo một số phương án cụ thể, thiết bị theo sáng chế có thể chứa nguồn sáng thích hợp cho phép phát hiện gen thông báo trong trứng được kiểm tra. Theo một số phương án khác nữa, nguồn sáng có thể có bước sóng nằm trong khoảng từ khoảng 400 đến khoảng 650. Theo một số phương án khác nữa, ánh sáng có bước sóng bất kỳ có thể được sử dụng với điều kiện là nguồn sáng này không phải là ánh sáng cực-tím (UV) (bước sóng nằm trong khoảng từ 10 đến 400 nm).

Theo một số phương án cụ thể, nguồn sáng thích hợp phù hợp để kích thích protein phát huỳnh quang. Theo một số phương án cụ thể, bước sóng kích thích có thể nằm trong khoảng từ khoảng 500 nm đến khoảng 650 nm. Theo một số phương án khác, bước sóng kích thích bằng khoảng 510, 515, 520, 525, 530, 535, 540, 545, 550, 555, 560, 565, 570, 575, 580, 585, 590, 595, 600, 605, 610, 615, 620, 625, 630, 635, 640 nm.

Theo một số phương án cụ thể hơn, bước sóng kích thích có thể nằm trong khoảng từ khoảng 515 đến khoảng 555. Theo một số phương án khác nữa, bước sóng có thể là 532 nm. Theo một số phương án cụ thể và không làm giới hạn, nguồn sáng có thể được tạo ra bởi laze. Theo một số phương án khác, laze của thiết bị theo sáng chế có thể là laze màu xanh lam, hoặc laze màu đỏ. Theo phương án cụ thể hơn, nguồn sáng là laze màu xanh lục.

Theo các phương án nhất định, laze có thể được dùng với bộ lọc. Theo một số phương án, bộ lọc có thể là bộ lọc màu xanh lục truyền dẫn trên 500 nm, hoặc bộ lọc màu xanh lục đậm truyền dẫn trong khoảng từ 540 nm đến 580 nm, hoặc bộ lọc màu đỏ truyền dẫn trong khoảng từ 590 nm đến 650 nm, hoặc bộ lọc màu đỏ truyền dẫn trên 650 nm hoặc trên 660 nm.

Theo một số phương án cụ thể, nguồn sáng có thể là laze màu xanh lục có bước sóng bằng 532 nm. Theo một số phương án khác, laze này có thể được cung cấp với bộ lọc màu đỏ truyền dẫn trên 650 nm.

Do đó cần hiểu thêm rằng sáng chế còn bao gồm thiết bị như được mô tả trong

bản mô tả này để sử dụng trong phương pháp bất kỳ trong số các phương pháp theo sáng chế, cụ thể là phương pháp xác định giới tính *in ovo* như được bàn luận trong bản mô tả này, bằng cách sử dụng đối tượng bất kỳ trong số các đối tượng thuộc loài chim chuyển gen theo sáng chế.

Theo các phương án nhất định, cần hiểu rằng (các) oligonucleotit hoặc (các) polynucleotit được sử dụng bởi (các) kit và (các) phương pháp theo sáng chế là các phân tử được phân lập và/hoặc được tinh chế. Khi được sử dụng trong bản mô tả này, thuật ngữ "**được phân lập**" hoặc "**được tinh chế**" khi được sử dụng liên quan đến axit nucleic có nghĩa là trình tự có trong tự nhiên đã được loại bỏ ra khỏi môi trường tế bào (ví dụ, ở nhiễm sắc thể) thông thường của nó hoặc được tổng hợp trong môi trường không tự nhiên (ví dụ, được tổng hợp theo cách nhân tạo). Do đó, trình tự "được phân lập" hoặc "được tinh chế" có thể ở trong dung dịch không có tế bào hoặc được đặt trong môi trường tế bào khác. Thuật ngữ "được tinh chế" không ngụ ý rằng trình tự chỉ có nucleotit, mà nghĩa là về cơ bản nó (độ tinh khiết khoảng 90-95%) không có vật liệu không phải nucleotit kết hợp tự nhiên với nó, và do đó được phân biệt với nhiễm sắc thể đã phân lập.

Tất cả các thuật ngữ khoa học và kỹ thuật được sử dụng trong bản mô tả này có nghĩa như thường được sử dụng trong lĩnh vực này, trừ khi có chỉ dẫn cụ thể theo cách khác. Các định nghĩa được cung cấp trong bản mô tả này là để tạo điều kiện thuận lợi cho việc hiểu các thuật ngữ nhất định được sử dụng thường xuyên trong bản mô tả này và không có nghĩa làm giới hạn phạm vi của sáng chế.

Trước khi các khía cạnh và phương án cụ thể của sáng chế được mô tả chi tiết trong tài liệu, cần hiểu rằng sáng chế này không bị giới hạn ở các phương pháp cụ thể, và các điều kiện thí nghiệm được mô tả, vì các phương pháp và điều kiện này có thể thay đổi. Cũng cần hiểu rằng các thuật ngữ được sử dụng ở đây là chỉ nhằm mục đích mô tả các phương án cụ thể, và không nhằm giới hạn sáng chế, vì phạm vi của sáng chế sẽ chỉ được giới hạn bởi yêu cầu bảo hộ kèm theo.

Khi được sử dụng trong bản mô tả này và phần yêu cầu bảo hộ kèm theo, các danh từ số ít bao gồm cả nghĩa số nhiều, trừ khi ngữ cảnh chỉ ra điều ngược lại một cách rõ ràng. Do đó ví dụ, việc đề cập đến "phương pháp" bao gồm một hoặc nhiều phương pháp, và/hoặc các bước của kiểu được mô tả trong bản mô tả này và/hoặc sẽ trở nên rõ ràng đối với người có hiểu biết trung bình trong lĩnh vực khi đọc phần bộc

lộ này, v.v..

Trừ khi có chỉ dẫn khác, tất cả các thuật ngữ kỹ thuật và khoa học dùng trong bản mô tả này có cùng nghĩa như thường được hiểu bởi người có hiểu biết trung bình trong lĩnh vực mà sáng chế này thuộc về. Mặc dù phương pháp hoặc nguyên liệu bất kỳ tương tự hoặc tương đương với phương pháp hoặc nguyên liệu được mô tả ở đây có thể được sử dụng trong thực hành hoặc thử nghiệm sáng chế, phương pháp và nguyên liệu được ưu tiên được mô tả dưới đây.

Trong toàn bộ bản mô tả này và phần yêu cầu bảo hộ kèm theo, trừ khi ngữ cảnh có yêu cầu khác, từ “chứa”, và các biến thể của nó, sẽ được hiểu theo ngụ ý là bao gồm số nguyên hoặc bước hoặc nhóm các số nguyên hoặc các bước được chỉ ra mà không loại trừ số nguyên hoặc bước hoặc nhóm các số nguyên hoặc các bước khác bất kỳ. Cụ thể hơn, các thuật ngữ "chứa", "có chứa", “bao gồm”, "bao gồm cả", "có" và các từ cùng gốc của chúng có nghĩa là "bao gồm, nhưng không chỉ giới hạn ở". Thuật ngữ này bao hàm các thuật ngữ "gồm có" và "về cơ bản gồm có". Cụm từ "về cơ bản gồm có" có nghĩa là chế phẩm hoặc phương pháp có thể bao gồm các thành phần và/hoặc bước bổ sung, nhưng chỉ khi các thành phần và/hoặc bước bổ sung không làm thay đổi đáng kể các đặc tính mới và cơ bản của chế phẩm hoặc phương pháp được yêu cầu bảo hộ.

Khi được sử dụng trong bản mô tả này, thuật ngữ "khoảng" biểu thị các giá trị có thể chênh lệch lên đến 1%, cụ thể hơn là 5%, cụ thể hơn là 10%, cụ thể hơn là 15%, và trong một số trường hợp cao hơn hoặc thấp hơn lên đến 20% so với giá trị được đề cập đến, khoảng giới hạn chênh lệch bao gồm giá trị số nguyên, và nếu thích hợp, giá trị không phải số nguyên cũng tạo thành khoảng giới hạn liên tục. Khi được sử dụng trong bản mô tả này, thuật ngữ "khoảng" dùng để chỉ $\pm 10\%$.

Cần lưu ý rằng các phương án khác nhau theo sáng chế này có thể được thể hiện theo định dạng khoảng giới hạn. Cần hiểu rằng việc mô tả ở dạng khoảng giá trị chỉ để cho thuận tiện và ngắn gọn và không được hiểu là giới hạn cứng nhắc trên phạm vi của sáng chế. Theo đó, việc mô tả khoảng giá trị cần được cho là đã bộc lộ một cách cụ thể tất cả các khoảng giá trị con có thể có cũng như các giá trị bằng số cụ thể nằm trong khoảng giá trị đó. Ví dụ, việc mô tả khoảng giới hạn chẳng hạn như nằm trong khoảng từ 1 đến 6 cần được cho là có các khoảng giới hạn con được bộc lộ cụ thể chẳng hạn như trong khoảng từ 1 đến 3, từ 1 đến 4, từ 1 đến 5, từ 2 đến 4, từ 2 đến 6,

từ 3 đến 6, v.v., cũng như các chữ số riêng lẻ trong khoảng giới hạn đó, ví dụ, 1, 2, 3, 4, 5, và 6. Điều này áp dụng bất kể độ rộng của khoảng giá trị. Bất cứ khi nào khoảng giới hạn số học được biểu thị trong bản mô tả này, có nghĩa là bao gồm số được nêu ra bất kỳ (phân số hoặc nguyên) trong khoảng giới hạn được chỉ ra. Các cụm từ "nằm trong khoảng/khoảng giới hạn từ" số biểu thị thứ nhất đến số biểu thị thứ hai và "nằm trong khoảng/khoảng giới hạn từ" số biểu thị thứ nhất "đến" số biểu thị thứ hai được sử dụng thay thế lẫn nhau trong bản mô tả này và có nghĩa là bao gồm số biểu thị thứ nhất và số biểu thị thứ hai và tất cả các số dạng phân số và số nguyên ở giữa chúng.

Các ví dụ sau đây được đưa ra để cung cấp cho người có hiểu biết trung bình trong lĩnh vực với phần bộc lộ và mô tả đầy đủ về cách tạo ra và sử dụng phương pháp và được phẩm theo sáng chế, và không được dự định để làm giới hạn phạm vi mà các tác giả sáng chế cho là sáng chế của họ. Các nỗ lực đã được thực hiện để đảm bảo độ chính xác đối với các chữ số được sử dụng (ví dụ, lượng, nhiệt độ, v.v.) nhưng cần được tính đến sai số và độ lệch của thử nghiệm. Trừ khi có chỉ dẫn khác, các phần được tính theo phần trọng lượng, trọng lượng phân tử là trọng lượng phân tử trung bình, nhiệt độ tính theo độ C, và áp suất bằng hoặc gần bằng áp suất khí quyển.

Các ví dụ là đại diện của các kỹ thuật được dùng bởi tác giả sáng chế trong khi tiến hành các khía cạnh của sáng chế. Cần hiểu rằng mặc dù các kỹ thuật này là ví dụ của các phương án ưu tiên để thực hành sáng chế, người có hiểu biết trung bình trong lĩnh vực, theo phần bộc lộ sáng chế, sẽ hiểu rằng có thể tiến hành nhiều cải biến mà vẫn nằm trong phạm vi được dự định của sáng chế.

Cần hiểu rằng để cho rõ ràng, các dấu hiệu nhất định của sáng chế được mô tả trong ngữ cảnh của các phương án riêng rẽ, cũng có thể được cung cấp ở dạng kết hợp trong phương án đơn lẻ. Ngược lại, để cho ngắn gọn, các dấu hiệu khác nhau của sáng chế được mô tả trong ngữ cảnh phương án đơn lẻ, cũng có thể được cung cấp riêng rẽ hoặc ở dạng kết hợp nhỏ thích hợp bất kỳ hoặc nếu thích hợp theo phương án được mô tả khác của sáng chế. Các dấu hiệu nhất định được mô tả trong ngữ cảnh các phương án khác nhau không được cho là các dấu hiệu thiết yếu của các phương án đó, trừ khi phương án không có hiệu lực khi không có các thành phần đó.

Các phương án và khía cạnh khác nhau của sáng chế như đã mô tả trên đây trong bản mô tả này và như được yêu cầu bảo hộ trong các mục yêu cầu dưới đây được chứng minh bằng thí nghiệm trong các ví dụ sau đây.

Mặc dù được bộc lộ và được mô tả, cần hiểu rằng sáng chế này không bị giới hạn ở các ví dụ, phương pháp, bước và chế phẩm cụ thể được bộc lộ trong bản mô tả này vì các bước phương pháp và chế phẩm này có thể thay đổi đến mức độ nào đó. Cũng cần hiểu rằng thuật ngữ được sử dụng trong bản mô tả này được sử dụng chỉ nhằm mục đích mô tả các phương án cụ thể và không được dự định làm giới hạn vì phạm vi của sáng chế chỉ bị giới hạn bởi các yêu cầu bảo hộ kèm theo và các dạng tương đương của chúng.

Cần lưu ý rằng khi được sử dụng trong bản mô tả này và phần yêu cầu bảo hộ kèm theo, các dạng số ít bao gồm cả nghĩa số nhiều, trừ khi ngữ cảnh chỉ ra rõ ràng là có nghĩa khác.

Ví dụ thực hiện sáng chế

Không cần giải thích thêm, chắc chắn là người có hiểu biết trung bình trong lĩnh vực có thể sử dụng sáng chế đến mức độ đầy đủ nhất bằng cách sử dụng phần mô tả nêu trên. Do đó, các phương án cụ thể được ưu tiên dưới đây được hiểu là chỉ để minh họa, và không làm giới hạn sáng chế được yêu cầu bảo hộ theo cách bất kỳ.

Quy trình sinh học phân tử tiêu chuẩn đã biết trong lĩnh vực không được mô tả cụ thể trong bản mô tả này như trong tài liệu: Sambrook et al., *Molecular cloning: A laboratory manual*, Cold Springs Harbor Laboratory, New-York (1989, 1992), và trong tài liệu Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, Maryland (1988).

Chất phản ứng

Động vật:

Gà White Leghorn thương mại thu được từ Hendrix ISA, và Minnesota.

Gà dòng chỉ thị từ Trung tâm nghiên cứu thực phẩm-nông nghiệp Thái Bình Dương (Pacific Agri-Food Research Centre), Agassiz, British Columbia, Canada.

Các thử nghiệm trên động vật được thực hiện nghiêm ngặt tuân theo quy trình được phê chuẩn IACUC và dưới sự giám sát của Ủy ban Crystal Bioscience IACUC để đảm bảo rằng không có động vật nào bị bệnh hoặc chết trong suốt quá trình thử nghiệm.

Vecto:

Hệ thống vectơ PrecisionX™ Cas9 SmartNuclease được đặt hàng từ System

Bioscience Inc., số danh mục CAS8/9xx series.

pDsRed1-N1 từ Clontech#6921-1.

Plasmid GFP- FUGW-H1-GFP-neomycin (Addgen), số danh mục 37632.

Dòng tế bào: Tế bào con cái là của *Gallus gallus*, dòng tế bào con cái nguyên bào lymphô Con A - C1 – VICK, được đặt hàng từ Mã số ATCC®: CRL-12135™. Đối với thử nghiệm nuôi cấy tế bào gà, dòng tế bào con cái nguyên bào sợi gà UMNSAH/DF-1 (ATCC® CRL-12203™) từ ATCC được sử dụng. Tế bào con đực là của *Gallus gallus*, dòng tế bào con đực biểu mô LMH/2A được đặt hàng từ Mã số ATCC®: CRL-2118™.

Dòng tế bào mầm nguyên thủy (PGC).

Dòng tế bào thận phôi người 293 (HEK-293) (ATTC).

Chất phản ứng và vật liệu cần thiết để tạo ra đối tượng thuộc loài chim chuyển gen

Phân lập PGC

1. Thiết bị nhỏ pipet vi điện cực (Shutter Instrument Co.).
2. Thiết bị vi nghiền (NARISHIGE).
3. Micropipet thủy tinh đường kính nhỏ (25µm).
4. Pipet kiểm soát miệng (Sigma-Aldrich).
5. 1x dung dịch muối cân bằng Hank (HBSS) mà không có CaCl₂ hoặc MgCl₂ (Hyclone, Logan, UT).
6. 1x nước muối đệm phosphat Dulbecco (PBS) mà không có Ca₂ hoặc Mg₂ (Hyclone).
7. Trypsin/axit etylenđiamintetraaxetic (EDTA) (Gibco).
8. Phôi gà trong các giai đoạn Hamburger và Hamilton (HH) 14–17.
9. Chất đệm phân loại tế bào được hoạt hóa từ tính (MACS). 1x PBS được bổ sung albumin huyết thanh bò (BSA) 0,5% và EDTA 2mM.
10. Cột MiniMACS (Miltenyi Biotec).
11. Kháng thể kháng nguyên phôi đặc hiệu giai đoạn 1 (SSEA-1) (Santa Cruz Biotechnology, Santa Cruz, CA).
12. Vi hạt kháng IgM chuột (Miltenyi Biotec).

Nuôi cấy PGC

1. PBS (Hyclone).

2. HBSS (Hyclone).

3. Môi trường nuôi cấy PGC (50 ml). 3,75ml huyết thanh bào thai bò (FBS, Hyclone), 1,25ml huyết thanh gà (Sigma-Aldrich, St. Louis, MO), 500µl chất bổ sung GlutaMAX-I (100x, Invitrogen), 500µl nucleosit (100x, Millipore), 500µl kháng sinh/kháng nấm (ABAM; 100x, Invitrogen), 500µl chất bổ sung selen-transferrin-insulin (100x, Gibco), 500µl axit amin không thiết yếu (NEAA; 100x, Invitrogen), 50µl β-mercaptoetanol (1000x, Sigma-Aldrich), 10µl thành phần sinh trưởng nguyên bào sợi cơ bản người 50 ng/µl (bFGF, Sigma-Aldrich), và bổ sung môi trường Eagle được cải biến Dulbecco làm bất hoạt (DMEM; Invitrogen) đến 50ml. Bảo quản ở nhiệt độ 4°C.

4. Accutase (Millipore).

Sản xuất thể khảm dòng mầm và chuyển gen

Chuyển gen vào PGC gà

1. Lipofectamine 2000 (Invitrogen).

2. Môi trường MEM tối ưu (Invitrogen).

3. Môi trường chuyển nhiễm (môi trường nuôi cấy PGC mà không có ABAM; 50ml): 3,75ml FBS (Hyclone), 1,25ml huyết thanh gà (Sigma-Aldrich), 500µl chất bổ sung GlutaMAX-I (100x, Invitrogen), 500µl nucleosit (100x, Millipore), 500µl chất bổ sung selen-transferrin-insulin (100x, Gibco), 500µl NEAA (100x, Invitrogen), 50µl β-mercaptoetanol (1000 x, Sigma-Aldrich), và 10µl bFGF 50 ng/µl (Sigma-Aldrich), và bổ sung DMEM làm bất hoạt (Invitrogen) đến 50ml. Bảo quản ở nhiệt độ 4°C.

4. PBS mà không có Ca₂ hoặc Mg₂.

5. HBSS mà không có Ca₂ hoặc Mg₂.

6. Điều chế ADN plasmit ở 1 µg/µl.

Sự tăng sinh *in vitro* và sự chọn lọc PGC chuyển gen

1. HBSS mà không có Ca₂ hoặc Mg₂ (Hyclone).

2. Máy đếm tế bào máu.

3. Thuốc nhuộm xanh trypan, 0,4% (Invitrogen).

4. Kính hiển vi huỳnh quang đảo ngược (Leica Microsystems).

5. Kháng sinh chọn lọc Geneticin® (Gibco).

Cấy PGC vào trứng thể nhận

1. Micropipet bằng thủy tinh đường kính nhỏ (25µm) (được tạo ra với thiết bị

nhỏ pipet vi điện cực (Shutter Instrument Co.) và thiết bị vi nghiền (NARISHIGE).

2. Pipet kiểm soát miệng.
3. HBSS mà không có CaCl_2 hoặc MgCl_2 (Hyclone).
4. Phôi gà thể nhận ở giai đoạn HH 14–17.
5. PBS mà không có Ca_2 hoặc Mg_2 (Hyclone).
6. Dung dịch rửa pipet. Hydroperoxit 0,1% (Sigma-Aldrich) trong nước cất thanh trùng.
7. Kẹp.
8. Kính hiển vi soi nổi có thiết bị chiếu.
9. Keo dán nhiệt nóng chảy và súng keo.
10. Parafilm.
11. PGC thể cho được tạo ra ở 3×10^3 tế bào/ μl trong HBSS.

Thử nghiệm chéo

1. Gà mái Ogye Hàn Quốc (KO) kiểu đại.
2. Ống tiêm 1ml.
3. PBS mà không có Ca_2 hoặc Mg_2 (Hyclone).

Con cái có nguồn gốc từ thể cho và sự xác nhận gà chuyển gen

1. Kit DNeasy Blood & Tissue (Qiagen).
2. Bộ đoạn mồi đặc hiệu PGC thể cho.
3. Bộ đoạn mồi đặc hiệu gen chuyển.
4. Máy PCR.
5. Chất phản ứng PCR: chất đệm PCR, dNTP, Taq polymeraza.
6. Agarosa.
7. Đệm TAE.
8. Máy điện di.

Quy trình thử nghiệm

Thử nghiệm phát huỳnh quang

Thiết lập hoàn chỉnh của thử nghiệm bao gồm nguồn laze, bộ giữ hoặc bộ đỡ có khả năng cố định trứng ở vị trí thích hợp cho phép tiếp xúc các tế bào biểu hiện gen thông báo trong trứng, với nguồn laze này, thấu kính, bộ lọc, bộ đỡ thiết bị phát hiện và thiết bị phát hiện như được minh họa trên Fig.1. Khoảng cách giữa mẫu và thấu kính và giữa thấu kính và thiết bị phát hiện là khoảng 20cm. Mặt phẳng tiêu cự của

thấu kính là khoảng 10cm và bộ lọc được đặt ở đó.

Các điều kiện tối ưu được phát hiện bằng cách đặt bộ lọc trước thiết bị phát hiện và bằng cách bao phủ tất cả khu vực xung quang thiết bị phát hiện để đảm bảo rằng ánh sáng tán xạ sẽ không đi vòng qua bộ lọc.

Ba loại laser được sử dụng với nguồn ban đầu I_0 , và bộ lọc như chỉ ra trong **Bảng 2** sau đây:

Bảng 2: nguồn sáng được sử dụng:

Loại laser (Bước sóng)	Nguồn [W]	Bộ lọc
Laser màu xanh lam (473±1nm)	$I_0 = 17,3mW$	trên 500nm
Laser màu xanh lục (532nm)	$I_0 = 3,9mW$	540 nm-580nm
Laser màu đỏ (632,8nm)	$I_0 = 3,3mW$	trên 660nm

Tiêm tế bào biểu hiện GFP hoặc RFP đã chuyển nhiễm vào trứng

Tế bào HEK được duy trì trong môi trường DMEM, có bổ sung L-glutamin 1%, huyết thanh bào thai bò (FBS) 10%, và PenStrep (penicillin + streptomycin) 1% (Biological industries, Israel). Các tế bào được cho sinh trưởng trong tủ áp được làm ấm ở nhiệt độ 37°C với CO₂ 5%. Môi trường được làm mới lại 2-3 lần một tuần, và số đếm tế bào được duy trì trong khoảng từ 1×10^5 đến 3×10^5 tế bào/ml. Tế bào được cấy chuyển và tái tạo huyền phù trong môi trường mới như sau:

1. Dịch nổi bề mặt được loại bỏ ra khỏi bình phản ứng để tế bào bám dính vào bề mặt bình này.
2. Các tế bào không bám dính được tái tạo huyền phù trong 5ml môi trường mới.

Sau khi cấy chuyển lần thứ 2, tế bào được chuyển nhiễm với 2 µg ADN của plasmit bằng cách sử dụng chất phản ứng chuyển nhiễm ADN *In Vitro* PolyJet™ (SignaGen Laboratories, MD, Mỹ) theo quy trình của nhà sản xuất. Hai plasmit được chuyển nhiễm vào tế bào một cách riêng rẽ; plasmit GFP và plasmit RFP, như chỉ ra ở trên.

Sau đó, tế bào đã chuyển nhiễm (RFP) được tạo ra ở nồng độ bằng 10^6 tế bào/ml. Tiếp đó, tổng thể tích 100 μ l tế bào được tiêm vào trứng gà tươi ở trung tâm, theo **Bảng 3** và **Bảng 4**.

Bảng 3: Tiêm RFP

Số lượng trứng	Tình trạng trứng	Nước muối (ul)	Dòng tế bào (ul) có 1M tế bào/ul	Dòng tế bào sau khi chuyển nhiễm GFP hoặc RFP
1.	Còn nguyên	0	0	-----
2.	Lỗ nhỏ tại đỉnh	100	0	0,1ml nước muối
11.	Lỗ nhỏ tại đỉnh	97	3	3.000 tế bào + RFP
12.	Lỗ nhỏ tại đỉnh	97	3	3.000 tế bào + RFP
13.	Lỗ nhỏ tại đỉnh	90	10	10.000 tế bào + RFP
14.	Lỗ nhỏ tại đỉnh	90	10	10.000 tế bào + RFP
15.	Lỗ nhỏ tại đỉnh	70	30	30.000 tế bào + RFP
16.	Lỗ nhỏ tại đỉnh	70	30	30.000 tế bào + RFP
17.	Lỗ nhỏ tại đỉnh	0	100	100.000 tế bào + RFP
18.	Lỗ nhỏ tại đỉnh	0	100	100.000 tế bào + RFP

Trong thử nghiệm thứ hai, các tế bào được sinh trưởng và chuyển nhiễm như đã mô tả trên đây. Sau đó tế bào được chuyển vào PBS hoặc glyxerol 50% (để kiểm soát sự phân tán) và tiêm vào trứng gà tươi ở trung tâm. Các trứng đã được tiêm khác nhau được bọc lộ trong **Bảng 4**.

Bảng 4: Tiêm RFP hoặc GFP

Số lượng trứng	Tình trạng trứng	PBS (ul) hoặc glyxerol 50%	Dòng tế bào (ul) có 1M tế bào/ul	Dòng tế bào sau sự chuyển nhiễm GFP hoặc RFP
1.	Còn nguyên	0		Đối chứng-KHÔNG chuyển nhiễm
2.	Lỗ nhỏ tại đỉnh	100	0	0,2ml PBS
3.	Lỗ nhỏ tại đỉnh	100	0	0,2ml glyxerol 50%
4.	Lỗ nhỏ tại đỉnh	99	1	1000 tế bào trong PBS
5.	Lỗ nhỏ tại đỉnh	99	1	1000 tế bào trong Gly 50%
6.	Lỗ nhỏ tại đỉnh	97	3	3000 tế bào trong PBS
7.	Lỗ nhỏ tại đỉnh	97	3	3000 tế bào trong Gly 50%
8.	Lỗ nhỏ tại đỉnh	90	10	10000 tế bào trong PBS
9.	Lỗ nhỏ tại đỉnh	90	10	10000 tế bào trong Gly 50%
10.	Lỗ nhỏ tại đỉnh	70	30	30000 tế bào trong PBS
11.	Lỗ nhỏ tại đỉnh	70	30	30000 tế bào trong Gly 50%
12.	Lỗ nhỏ tại đỉnh	99	1	1000 tế bào trong PBS
13.	Lỗ nhỏ tại đỉnh	99	1	1000 tế bào trong Gly 50%
14.	Lỗ nhỏ tại đỉnh	97	3	3000 tế bào trong PBS
15.	Lỗ nhỏ tại đỉnh	97	3	3000 tế bào trong Gly 50%

16.	Lỗ nhỏ tại đỉnh	90	10	10000 tế bào trong PBS
17.	Lỗ nhỏ tại đỉnh	90	10	10000 tế bào trong Gly 50%
18.	Lỗ nhỏ tại đỉnh	70	30	30000 tế bào trong PBS
19.	Lỗ nhỏ tại đỉnh	70	30	30000 tế bào trong Gly 50%

Dòng tế bào gà giống đực

Tế bào giống đực, cụ thể là *Gallus gallus*, gà (gan; được gây cảm ứng hóa học; được chuyển nhiễm bằng gen thụ thể estrogen gà), được sinh trưởng trong môi trường MB 752/1 của Waymouth có bổ sung L-glutamin 1%, huyết thanh bào thai bò 10% (FBS), và PenStrep 1% (penixilin + streptomycin) (Biological industries, Israel). Bình nuôi cấy (T25) được phủ bằng gelatin 0,1%). Môi trường được làm mới lại mỗi 3 ngày. Tế bào được sinh trưởng trong tủ áp được làm ẩm ở nhiệt độ 37°C với CO₂ 5%.

Dòng tế bào gà giống cái

Sử dụng dòng tế bào giống cái nguyên bào sợi gà **UMNSAH/DF-1 (ATCC[®] CRL-12203[™])** từ ATCC. Chi tiết hơn về tế bào này được thể hiện trong Bảng 5:

Bảng 5: Dòng tế bào gà giống cái

Sinh vật	<i>Gallus gallus</i> , gà
Loại tế bào	nguyên bào sợi tự biến nạp
Dạng sản phẩm	được đông lạnh
Hình thái học	nguyên bào sợi
Tính chất nuôi cấy	bám dính
Tuổi	10 ngày áp
Giới tính	giống cái

Tế bào giống cái được sinh trưởng trong môi trường DMEM có bổ sung L-glutamin 1%, huyết thanh bào thai bò 10% (FBS), và PenStrep 1% (penixilin + streptomycin) (Biological industries, Israel). Các tế bào được sinh trưởng trong tủ áp

được làm ẩm ở nhiệt độ 37°C với CO₂ 5%. Môi trường được làm mới lại 2-3 lần một tuần.

Nuôi cấy tế bào PGC

Tuyến sinh dục phôi được lấy ra khỏi phôi White Leghorn (WL) ở thời điểm 6,5 ngày ấp và sau đó tế bào tuyến sinh dục đã lấy ra được xử lý bằng thiết bị phân loại tế bào được hoạt hóa từ tính (MACS) để phân tách PGC trong toàn bộ tế bào tuyến sinh dục phôi. PGC được sinh trưởng trong KO-DMEM (Life Technologies), mà trong đó 40% được xử lý sơ bộ trên tế bào gan chuột buffalo (BRL, ATCC), và được bổ sung 7,5% huyết thanh bào thai bò (Hyclone), 2,5% huyết thanh gà được chiếu xạ, 1X axit amin không thiết yếu, glutamin 2mM, natri pyruvat 1mM, β -mercaptoetanol 0,1mM (tất cả từ Life Technologies), 4ng/ml yếu tố sinh trưởng nguyên bào sợi người tái tổ hợp, 6ng/ml yếu tố tế bào mầm chuột tái tổ hợp (cả hai từ hệ thống R&D) và được sinh trưởng trên lớp tiếp liệu được chiếu xạ của tế bào BRL. Các tế bào được cấy chuyển 3 lần trong mỗi tuần trên lớp tiếp liệu mới.

Sự tách dòng không cắt giới hạn (RF)

Sự chèn đoạn gARN vào vector Cas9-SmartNucleaseTM được thực hiện bằng cách dùng phương pháp không cắt giới hạn [Peleg Y et al., J. Struct. Biol. 172(1):34-44 (2010) 2010]. Đoạn mồi được đặt hàng từ Sigma-Genosys (Rehovot, Israel) và sau đó phản ứng RF được tiến hành bằng cách sử dụng Phusion polymeraza (Thermo Scientific, Hudson, NH, Mỹ). Tiến hành tinh chế plasmit bằng cách sử dụng lần lượt kit MEGAspin và kit tinh chế plasmit ADN DNA-spin, (Intron Biotechnology, Daejeon, Hàn Quốc).

Sự chuyển nhiễm vào nhiễm sắc thể W hoặc Z của gà

Tế bào được đồng-chuyển nhiễm với 2 μ g ADN của mỗi plasmit (xem chi tiết dưới đây) bằng cách sử dụng chất phản ứng chuyển nhiễm ADN *In Vitro* PolyJetTM. (SignaGen Laboratories, MD, Mỹ) theo quy trình của nhà sản xuất. Thực hiện đồng-chuyển nhiễm hai plasmit như sau:

1. pDsRed có các nhánh bên trái và phải nhiễm sắc thể Z gà hoặc các nhánh bên trái và phải nhiễm sắc thể W và CMV-hspCas9-H1-gARN (thử nghiệm).
2. pCMVGluc có các nhánh bên trái & phải nhiễm sắc thể Z gà hoặc các nhánh bên trái & phải nhiễm sắc thể W và CMV-hspCas9-H1-gARN (đối chứng).

Sau đó đàn mỏng tế bào với chất kháng Neomycin và cấy trong đĩa có 48-lỗ với

mật độ 10^5 tế bào/lỗ. Sau 3 ngày, 40 μ g/ml Neomycin được bổ sung để chọn lọc tế bào có sự tích hợp ổn định của gen thông báo.

Phân tích ADN về sự tích hợp vào nhiễm sắc thể Z hoặc W

Để xác nhận sự tích hợp nhiễm sắc thể, sử dụng chiến lược sau đây. Phân lập tế bào và ADN của một nửa trong số chúng được chiết và chạy trên gel aga (1%). ADN hệ gen (để lại ADN plasmid trên gel) được cắt và làm sạch. Chạy thử nghiệm PCR với các đoạn mồi sau đây. Phân tách sản phẩm trên gel và sau đó giải trình tự. Bởi vì đoạn chèn quá dài không thể phát hiện bằng thử nghiệm PCR, thử nghiệm PCR của số trượt được thiết kế để bắt đầu trên sản phẩm PCR ban đầu.

a. Thử nghiệm đối với dòng tế bào có đoạn chèn (bao gồm một phần của nhánh bên trái, gen tăng cường CMV, gen khởi đầu và MCS và một phần của dsRED2):

Đoạn mồi bên trái-TTGGTGTGTGCTAATAGGCAGT như được thể hiện bằng trình tự SEQ ID NO 49; đoạn mồi bên phải-TAGTCCTCGTTGTGGGAGGT như được thể hiện bằng trình tự SEQ ID NO 50, kích thước sản phẩm: 1489bp.

b. Thử nghiệm đối với dòng tế bào có đoạn chèn (bao gồm một phần của nhiễm sắc thể Z kèm hai bên (đối với nhánh), nhánh bên trái, và một phần của gen tăng cường CMV, gen khởi đầu):

Đoạn mồi bên trái- GCATTGATCTGTCCAGTTGC như được thể hiện bằng trình tự SEQ ID NO 51, đoạn mồi bên phải- TACTGCCAAAACCGCATCAC như được thể hiện bằng trình tự SEQ ID NO 52, kích thước sản phẩm: 1462bp.

Tạo ra gà chuyển gen

Phương tiện truyền đã cô đặc (có thể là lentivirus ở độ chuẩn bằng khoảng 10^7 MOI) hoặc ADN plasmid) được tiêm truyền vào 25 phôi trong trứng mới đẻ. Tiến hành tiêm với ba lần tiêm mỗi tuần. Phôi đã tiêm nở 3 tuần sau khi tiêm. Chúng là chim G0. Ngay sau khi nở, ADN được chiết từ mẫu CAM của các con non đã nở và tiến hành phát hiện sự có mặt/không có mặt của ADN vectơ bằng PCR bán định lượng. Lấy mẫu máu của các con non G0 lúc 2-3 tuần tuổi và lặp lại sàng lọc PCR. Nuôi chim G0 đến khi trưởng thành sinh dục, 16-20 tuần đối với giống đực, 20-24 tuần đối với giống cái. Con trống non được thử nghiệm về sự tạo ra tinh dịch từ khoảng 16 tuần.

Con mái được cấy tinh, trứng đã thụ tinh được thu gom hằng ngày. Các con non

G1 nở 3 tuần sau đó và mỗi cánh con non riêng lẻ bị nẹp và mẫu màng đệm túi niêu con non (CAM) được lấy từ vỏ. Chiết tách ADN từ mẫu CAM và tiến hành sàng lọc PCR đối với sự có mặt của gen chuyển, được dự kiến là hàm lượng bản sao đơn lẻ. Lắp lại sàng lọc để xác nhận và xác định giới tính các con non trên ADN từ mẫu máu 2-3 tuần sau đó.

Ở thời điểm vài tuần tuổi, mẫu máu được thu được từ chim G1 để tạo ra ADN hệ gen để phân tích PCR. Sử dụng chim G1 để gây giống G2.

Phân lập PGC

1. Ủ trứng gà mới (EGK giai đoạn X), ví dụ, Ogye Hàn Quốc (KO), ở nhiệt độ 37°C trong thời gian từ 50 đến 54 giờ (HH giai đoạn 14–17) để phân lập PGC máu. Đặt trứng đã ủ theo chiều ngang trên đĩa trứng và lau nhẹ vỏ trứng bằng cách sử dụng bông vệ sinh có etanol 70%.

2. Sử dụng kẹp nhọn, cẩn thận làm vỡ vỏ trứng (đường kính <1 cm) để phân lập tế bào máu toàn phần.

3. Thu được khoảng 2–3 µl tế bào máu toàn phần từ động mạch chủ ở lưng phôi bằng cách sử dụng kim thủy tinh đã mài mỏng 25 µm và pipet miệng. Trộn với 500µl PBS. Để vi tiêm PGC, có thể sử dụng pipet miệng và micropipet thủy tinh có đường kính 20–25 µm. Micropipet thủy tinh được tạo ra với thiết bị nhỏ pipet vi điện cực (Shutter Instrument Co.) và mài ở 25 độ bằng cách sử dụng thiết bị vi mài (NARISHIGE).

4. Chuyển nhiễm tế bào máu toàn phần vào ống vô khuẩn 1,5ml, ly tâm (250 x g, 5 phút), và lấy dịch nổi bề mặt ra.

5. Để phân lập PGC tuyến sinh dục, ủ trứng gà mới ở nhiệt độ 37°C trong thời gian 5,5 ngày (HH 20–26). Chiết tách tuyến sinh dục phôi khỏi phôi ở giai đoạn HH 20–26 bằng kẹp nhọn, và ủ với 500µl trypsin 0,05%/EDTA ở áp suất nhiệt độ 37°C trong thời gian 5 phút.

6. Bổ sung 50µl FBS để bất hoạt, và ly tâm (250 x g, 5 phút) và lấy dịch nổi bề mặt ra.

7. Tái tạo huyền phù tế bào máu toàn phần hoặc tế bào tuyến sinh dục trộn lẫn trong 1ml PBS và dùng kháng thể kháng-SSEA-1 (Santa Cruz Biotechnology, SC-21702) với tỷ lệ pha loãng 1:200, và sau đó ủ hỗn hợp trong thời gian 15 phút ở nhiệt độ trong phòng (RT).

8. Rửa tế bào để loại bỏ kháng thể sơ cấp không liên kết bằng cách bổ sung 5ml chất đệm MACS (BSA 0,5% và EDTA 2mM trong PBS, độ pH= 7,2) trên mỗi 10^7 tổng số tế bào và ly tâm (250 x g, 5 phút).

9. Hút hoàn toàn dịch nổi bề mặt ra và tái tạo huyền phù viên tế bào trong 80µl chất đệm MACS trên mỗi 10^7 tổng số tế bào.

10. Bổ sung 20µl vi hạt của chuột chống kháng IgM-chuột nhắt (Miltenyi Biotec. 130-047-301) trên mỗi 10^7 tổng số tế bào. Đối với số lượng tế bào cao hơn, tăng thêm thể tích chất đệm tương ứng.

11. Ủ tế bào với kháng thể ở nhiệt độ trong khoảng từ 2 đến 8°C trong thời gian 20 phút.

12. Rửa tế bào bằng cách bổ sung 2ml chất đệm MACS trên mỗi 10^7 tổng số tế bào và ly tâm (250 x g, 5 phút).

13. Hút hoàn toàn dịch nổi bề mặt ra và tái tạo huyền phù lên đến 10^8 tổng số tế bào trong 500 µl đệm MACS.

14. Đặt cột trong từ trường của thiết bị phân tách MACS phù hợp.

15. Chuẩn bị cột bằng cách súc rửa bằng 500µl chất đệm MACS.

16. Cho dịch huyền phù tế bào lên cột, và rửa cột bằng 500µl chất đệm ba lần. Bổ sung chất đệm mới khi vật chứa cột rỗng.

17. Lấy cột khỏi thiết bị phân tách và đặt nó lên ống thu gom phù hợp.

18. Bổ sung 1ml chất đệm MACS vào cột và ngay lập tức xối rửa tế bào đã đánh dấu từ tính bằng cách đẩy mạnh trụ trượt vào cột.

19. Chuyển tế bào đã phân lập chứa PGC vào môi trường nuôi cấy PGC đã làm ấm trước (1ml môi trường PGC trên mỗi 1×10^5 PGC đã tinh chế).

Nuôi cấy PGC

1. Chuyển tế bào đã phân lập chứa PGC vào môi trường nuôi cấy PGC đã làm ấm trước (1ml môi trường PGC trên mỗi 1×10^5 PGC đã tinh chế) và ủ ở nhiệt độ 37°C.

2. Sau 7–14 ngày sinh trưởng, hầu hết các PGC được nuôi cấy sơ cấp tạo ra khuẩn lạc và mất khối kết tập khuẩn lạc tế bào.

3. Khuẩn lạc PGC được tạo huyền phù có thể được hút nhẹ ra khỏi môi trường sau khi dùng pipet nhẹ nhàng và ly tâm (200 x g, 5 phút).

4. Làm tan khối kết tụ viên tế bào bằng Accutase (1ml dung dịch Accutase trên

mỗi $\sim 5 \times 10^5$ PGC).

5. Ly tâm (200 x g, 5 phút) và tái tạo huyền phù viên tế bào trong môi trường nuôi cấy PGC.

6. Cây PGC đã được tạo huyền phù trong đĩa 12-lỗ (1×10^5 PGC được nuôi cấy trong 1ml môi trường nuôi cấy PGC trên mỗi lỗ của đĩa 12-lỗ).

7. PGC có thể được cấy chuyển thường xuyên mỗi 3–4 ngày.

Chuyển gen vào PGC gà, tạo ra thể khảm dòng mầm, và sự chuyển gen

Chuyển gen vào PGC gà

1. Trộn 4 μ g plasmit trợ giúp chứa CAGG-PBase (pCyL43B) và 6 μ g gen nhảy piggyBac (pCyL50) chứa gen thông báo (RFP) và tùy ý, chất chỉ thị có chọn lọc (ví dụ, gen kháng neomycin) với 100 μ l Opti-MEM (Gibco) và ủ trong thời gian 5 phút ở RT. Để chuyển nhiễm CRISPR/Cas, trộn 2,5 μ g vectơ plasmit biểu hiện CMVRFP, 2,5 μ g vectơ plasmit biểu hiện Cas9 và ARN dẫn đường với 100 μ l Opti-MEM (Gibco) và ủ trong thời gian 5 phút ở RT.

2. Trộn 10 μ l chất phản ứng Lipofectamine 2000 (Thermo Fisher-Invitrogen) với 100 μ l Opti-MEM và ủ trong thời gian 5 phút ở RT.

3. Trộn plasmit với Opti-MEM và chất phản ứng Lipofectamine 2000 với Opti-MEM và ủ trong thời gian 20 phút ở RT.

4. Trong quá trình ủ, thu hoạch PGC đã được nuôi cấy và ly tâm (200 x g, 5 phút). Lắng gạn dịch nổi bề mặt.

5. Bổ sung 1ml Accutase (Millipore) vào PGC thu được và ủ trong thời gian 10 phút ở nhiệt độ 37°C.

6. Xác định số lượng PGC bằng cách sử dụng máy đếm tế bào máu (Marienfeld) và cấy 5×10^5 PGC trong đĩa có 12-lỗ có 1ml môi trường nuôi cấy PGC mà không có kháng sinh.

7. Cho phức hợp ADN-Lipofectamine vào PGC và ủ trong thời gian 1 ngày ở nhiệt độ 37°C trong tủ áp CO₂.

8. Thu hoạch PGC đã chuyển nhiễm và ly tâm (200 x g, 5 phút). Lấy dịch nổi bề mặt ra.

9. Rửa PGC bằng 1 ml HBSS ba lần và tạo huyền phù chúng với môi trường nuôi cấy PGC có kháng sinh.

Sự tăng sinh *in vitro* và sự chọn lọc PGC chuyển gen

1. Chọn lọc PGC đã chuyển nhiễm trong môi trường nuôi cấy PGC chứa 100 µg/ml G418 (đối với gen kháng neomycin) vào ngày sau khi chuyển nhiễm.
2. Theo dõi sự biểu hiện gen thông báo bằng cách sử dụng kính hiển vi huỳnh quang.
3. Cấy chuyển PGC mỗi 3–4 ngày. Khoảng thời gian chọn lọc hoàn toàn cần thiết lên đến 3 tuần.
4. Nếu không có chất chỉ thị chọn lọc-được chất trong vector plasmit được chuyển nhiễm, PGC biểu hiện protein phát huỳnh quang có thể được phân loại bằng FACS.

Cấy PGC vào trứng thể nhận

1. Ủ trứng nhận lên đến giai đoạn HH 14–17 ở nhiệt độ 37°C trong không khí có độ ẩm tương đối 60–70%.
2. Thu hoạch PGC đã chuyển nhiễm và ly tâm (200 x g, 5 phút) và lấy dịch nổi bề mặt ra.
3. Bổ sung 1ml Accutase (Millipore) vào PGC thu được và ủ trong thời gian 10 phút ở nhiệt độ 37°C.
4. Ly tâm (200 x g, 5 phút). Lắng gạn dịch nổi bề mặt.
5. Tạo huyền phù PGC trong HBSS (Hyclone).
6. Tạo ra lỗ nhỏ trên đầu nhọn của trứng nhận và vi tiêm 2µl phân phân ước chứa nhiều hơn 3.000 PGC bằng micropipet vào động mạch chủ ở lưng của phôi thể nhận.
7. Bịt kín lỗ trứng của phôi thể nhận bằng parafilm bằng cách sử dụng súng keo, và ủ trứng với đầu nhọn xuống cho đến khi nở ở nhiệt độ 37°C trong không khí có độ ẩm tương đối 60–70%.

Theo dõi PGC đã cải biến di truyền trong tuyến sinh dục phôi

1. Ủ trứng nhận lên đến giai đoạn HH 28–30 ở nhiệt độ 37°C trong không khí có độ ẩm tương đối 60–70%.
2. Giải phẫu tuyến sinh dục ở phôi vào ngày 6.
3. Theo dõi sự biểu hiện protein phát huỳnh quang ở tuyến sinh dục phôi bằng cách sử dụng kính hiển vi huỳnh quang.

Thử nghiệm chéo

1. Thu thập tinh dịch hai lần trong 1 tuần từ thể nhận trưởng thành về mặt sinh

dục và gà kiêu đại (WT).

2. Đưa 50 μ l tinh dịch từ thể nhận giống đực trưởng thành và con trống WT vào lần lượt con mái đang đẻ WT và thể nhận giống cái trưởng thành.

3. Thu thập trứng từ con mái đang đẻ WT và thể nhận giống cái trưởng thành ngày sau khi cấy tinh nhân tạo, và ủ trứng với đầu nhọn xuống cho đến khi nở ở nhiệt độ 37 °C trong không khí có độ ẩm tương đối 60–70%.

Ví dụ 1

Phân tích sự phát huỳnh quang bằng cách sử dụng các laze có màu và bộ lọc khác nhau để nhận biết giới tính bằng mắt ở gia cầm

Hệ thống phát huỳnh quang quang học

Để chứng minh tính khả thi của sự nhận biết giới tính bằng mắt của gia cầm *in-ovo*, việc sử dụng huỳnh quang màu xanh lục khi được so sánh với gen thông báo phát huỳnh quang màu xanh lam hoặc huỳnh quang màu đỏ được đánh giá.

Sự tự phát huỳnh quang của toàn bộ trứng hoặc của trứng đã được phân tách thành lòng trắng trứng, noãn hoàng trứng và vỏ được xác định bằng cách sử dụng laze màu xanh lam, laze màu xanh lục hoặc laze màu đỏ. So sánh trứng rồng (vỏ) với toàn bộ trứng cho thấy không có sự khác biệt đáng kể về mức độ tự phát huỳnh quang (dữ liệu không được thể hiện).

Cường độ tán xạ và tự phát huỳnh quang được tạo ra bởi các thành phần và phần khác nhau của trứng được mô tả trong **Bảng 6**.

Điều quan trọng cần lưu ý là độ nhiễu nền (không có laze) là khoảng 4nW, trong khi khi có mặt laze (532nm), nó bằng khoảng 12nW do ánh sáng tán xạ đi vòng qua bộ lọc (bộ lọc truyền thấp truyền trên 660nm).

Bảng 6. Cường độ tán xạ và cường độ tự phát huỳnh quang của các phần của trứng bằng cách sử dụng 3 laze huỳnh quang khác nhau

Laze	Cường độ tán xạ			Cường độ tự phát huỳnh quang		
Thử nghiệm	Vật chứa	Vật chứa +	Vật chứa +	Vật chứa	Vật chứa + lòng trắng trứng	Vật chứa + lòng vàng trứng

	rỗng	lòng trắng trứng	lòng vàng trứng	rỗng		
Laze màu xanh lam (473±1nm)	2mW	1.02mW	1.2mW	16.5nW	$26.12 \pm 6.25nW$	$10.33 \pm 1.5nW$
Laze màu xanh lục (532nm)	370μW	484μW	415μW	13.6nW	14.5nW	16nW
Laze màu đỏ (632,8nm)	300μW	160μW	156μW	13nW	$8.33 \pm 1.86nW$	$10.82 \pm 3.42nW$

Việc sử dụng laze màu xanh lam (473nm) cùng với bộ lọc màu xanh lục (+500nm) hoặc bộ lọc màu đỏ (590-650 nm) có bổ sung hoặc không bổ sung các nồng độ khác nhau của floresxein (1μM, 10mM) vào toàn bộ trứng được thử nghiệm và được tóm tắt trên **Fig.2**. Đường như có sự tự phát huỳnh quang lớn, tuy nhiên cường độ phát huỳnh quang không được phát hiện khi sử dụng nồng độ floresxein cao.

Việc sử dụng laze màu xanh lục (532nm) cùng với bộ lọc màu xanh lục đậm (540-580 nm) hoặc bộ lọc màu đỏ (+660nm) có bổ sung hoặc không bổ sung Rodamin B (10μM) vào toàn bộ trứng được thử nghiệm và được tóm tắt trên **Fig.3**. Đường như ở trên 40mW, sự phát huỳnh quang của thuốc nhuộm có thể nhận thấy đáng kể với bộ lọc màu xanh lục (lớn hơn gần 2 lần). Tuy nhiên, với bộ lọc màu đỏ, có thể quan sát thấy không có sự khác biệt.

Việc sử dụng laze màu xanh lục (532nm) cùng với bộ lọc màu đỏ A (590-650 nm) hoặc bộ lọc màu đỏ B (+660nm) có bổ sung hoặc không bổ sung dir (10 μM) vào toàn bộ trứng được thử nghiệm và được tóm tắt trên **Fig.4**. Đường như ở trên 80mW, sự phát huỳnh quang của thuốc nhuộm có thể nhận thấy bằng cách sử dụng bộ lọc màu đỏ A trong khi bằng cách sử dụng bộ lọc màu đỏ B (+660nm), có thể quan sát thấy chỉ có sự khác biệt nhỏ nhất.

Việc sử dụng laze màu đỏ (632,8nm) cùng với bộ lọc màu đỏ (+660nm) hoặc bộ lọc màu xanh lục (540-580 nm) có bổ sung hoặc không bổ sung dir (10μM) vào toàn bộ trứng được thử nghiệm và được tóm tắt trên **Fig.5**. Hầu như không có sự khác biệt nào được thể hiện khi có hoặc không có thuốc nhuộm huỳnh quang.

Lặp lại tất cả các thí nghiệm bằng cách chỉ sử dụng vỏ trứng. Quan sát thấy các kết quả tương tự (dữ liệu không được thể hiện).

Các quan sát này chứng minh tính khả thi của việc phát hiện RFP thông qua vỏ trứng.

Ví dụ 2

Huỳnh quang của tế bào đã chuyển nhiễm RFP hoặc GFP vào trứng

Để minh họa tính khả thi của việc phát hiện tế bào biểu hiện RFP trong phôi *in-ovo*, các tác giả sáng chế tiếp tục kiểm tra liệu tế bào được chuyển nhiễm với protein phát huỳnh quang, cụ thể là RFP hoặc GFP, có thể được phát hiện khi tiêm chúng vào toàn bộ trứng hay không.

Cụ thể hơn, tế bào HEK được chuyển nhiễm với vector GFP hoặc RFP như được mô tả trong phần Quy trình thí nghiệm. Kiểm tra sự phát huỳnh quang của cả tế bào được chuyển nhiễm GFP và RFP được sau khi kích thích. **Fig.6** (GFP) và **Fig.7** (RFP) chứng minh rõ ràng rằng cả GFP và RFP được biểu hiện bởi tế bào đã được chuyển nhiễm và có thể quan sát thấy tín hiệu phát huỳnh quang.

Sau đó, các tế bào đã được chuyển nhiễm HEK được tiêm vào trứng mới như được mô tả trong phần Quy trình thí nghiệm và cường độ phát huỳnh quang của trứng được thử nghiệm.

Trước tiên, cường độ phát huỳnh quang của trứng khi có hoặc không có tế bào biểu hiện protein phát huỳnh quang màu đỏ RFP được đặc trưng bởi sự kích thích trứng bằng laze màu xanh lục (532nm) và bộ lọc màu đỏ bằng +650nm. **Fig.8** cho thấy rằng cường độ phát huỳnh quang của trứng khi không có RFP được đo với năm vị trí khác nhau của trứng là khá tương tự. **Fig.9** thể hiện cường độ phát huỳnh quang đo được bằng cách sử dụng các nồng độ khác nhau của tế bào biểu hiện RFP khi trứng được kích thích ở trung tâm. Sự tương quan giữa nồng độ RFP và cường độ trên thiết bị phát hiện được quan sát thấy rõ.

Một số thử nghiệm được thực hiện để kiểm tra xem khi có mặt của tế bào biểu hiện RFP, vị trí của trứng, tức là địa điểm chính xác của sự kích thích được tạo ra bởi laze, ảnh hưởng đến cường độ phát huỳnh quang hay không. Dường như là cường độ trên thiết bị phát hiện đã thay đổi lớn đối với vị trí của trứng trong quá trình kích thích. Điều này được mong đợi bởi vì, sau khi tiêm tế bào biểu hiện RFP vào trứng, RFP không được phân bố đồng đều. **Fig.10** thể hiện cường độ tối đa được quan sát thấy ở vị trí cụ thể của trứng trong quá trình kích thích, tương ứng với vị trí mà tại đó nồng độ của tế bào biểu hiện RFP là cao nhất. Hình vẽ này chứng minh rằng lượng tối thiểu

bằng khoảng 3000 tế bào có thể được phát hiện.

Trong phôi, đã biết trong lĩnh vực này rằng tế bào giới tính, PGC nằm ở trung tâm của ngoại phôi bì của trứng mới đẻ ở giai đoạn phát triển X của trứng. Cần lưu ý rằng phôi luôn hướng đến mặt trên của túi noãn hoàng trứng.

Các kết quả được biểu lộ ở trên gợi ý rằng khi kích thích vùng ở sau vỏ trứng một vài millimet và gần noãn hoàng, sự phát RFP tối đa có thể được phát hiện.

Lặp lại các thí nghiệm một tuần sau khi tiêm trứng bằng tế bào biểu hiện RFP. Các kết quả tương tự chứng minh rằng thuốc nhuộm RFP không khuếch tán vào trong trứng.

Sự ảnh hưởng của PBS và glyxerol lên cường độ phát huỳnh quang của trứng bằng cách sử dụng laze màu xanh lục (532 nm) và bộ lọc màu đỏ bằng +650nm được kiểm tra khi có hoặc không có mặt của tế bào biểu hiện RFP. Dường như là glyxerol hoặc PBS đều không ảnh hưởng đến các phép đo cường độ phát huỳnh quang của tế bào biểu hiện RFP trong trứng như được minh họa trên Fig.11A và Fig.11B.

Sau đó, kiểm tra sự ảnh hưởng của PBS và glyxerol đối với cường độ phát huỳnh quang của trứng bằng cách sử dụng laze màu xanh lam (483 nm) và bộ lọc bằng +500 nm khi có hoặc không có mặt của tế bào biểu hiện GFP. Dường như là tế bào biểu hiện GFP không được phát hiện ở nồng độ đã cho bất kỳ (1000, 3000 10.000 và 30.000 tế bào), như được minh họa bởi Fig.12 ví dụ với 30.000 tế bào biểu hiện GFP. Thu được các kết quả như nhau bằng cách sử dụng ba bộ lọc khác nhau: +500 nm, 530-550 nm và 540-580 nm.

Để đánh giá thêm lợi thế bất ngờ của việc sử dụng RFP dưới dạng gen thông báo, huỳnh quang của tế bào biểu hiện RFP và GFP cuối cùng được so sánh như được minh họa trên Fig.13. Tham số R được xác định là biểu thị tỷ lệ giữa cường độ protein phát huỳnh quang (FP) (I_{FP}) và cường độ tự phát huỳnh quang ($I_{\text{tự phát huỳnh quang}}$):

$$R = \frac{I_{FP}}{I_{\text{tự phát huỳnh quang}}}$$

Tham số R được tính đối với các nồng độ khác nhau của tế bào biểu hiện RFP và tế bào biểu hiện GFP. Như được thể hiện trên hình vẽ, dường như là không thể nhận thấy các tế bào biểu hiện GFP ở nồng độ bất kỳ trong số các nồng độ đã cho (1000 và 30.000 tế bào như được thể hiện trên Fig.13).

Mặt khác, tế bào biểu hiện RFP tạo ra giá trị tham số R đáng kể ngay cả ở nồng

độ thấp (1000 tế bào biểu hiện RFP). Ở nồng độ cao của tế bào biểu hiện RFP (30000 tế bào biểu hiện RFP), giá trị R đạt đến 3, và do đó chứng minh rõ ràng rằng gen thông báo RFP có thể dễ dàng được phát hiện trong trứng được kích thích.

Các kết quả phát hiện chính được tóm tắt dưới đây:

1. Có sự tự phát huỳnh quang mạnh màu xanh lam trong khi không có sự tự phát huỳnh quang ở màu xanh lục và màu đỏ.
2. Sự tự phát huỳnh quang chủ yếu có nguồn gốc từ vỏ trứng.
3. Hệ thống có thể phát hiện thuốc nhuộm huỳnh quang màu xanh lục và màu đỏ với nồng độ thuốc nhuộm tương đối thấp và với nguồn kích thích yếu.
4. Huỳnh quang màu xanh lam khó phát hiện (chỉ ở nồng độ thuốc nhuộm mM) do sự tự phát huỳnh quang.
5. RFP có thể nhận thấy một cách đáng kể và dễ dàng bên trong trứng.
6. Loại GFP không thể phát hiện được trong hệ thống này.
7. Vị trí chính xác của tế bào biểu hiện RFP bên trong trứng có ý nghĩa để xác định vùng kích thích.
8. Thuốc nhuộm RFP không khuếch tán vào trứng trong 14 ngày từ khi tiêm.
9. PBS hoặc Glyxerol không ảnh hưởng đến sự phát hiện RFP.

Các kết quả này chứng minh rõ ràng là phôi biểu hiện gen RFP được tích hợp trong nhiễm sắc thể giới tính, có thể được phát hiện ở giai đoạn phát triển sớm, trong đó khoảng 1000 tế bào biểu hiện RFP.

Ví dụ 3

Thiết kế vector ARN dẫn đường

Để kết hợp gen thông báo RFP vào nhiễm sắc thể giới tính W hoặc Z, phương pháp HDR qua trung gian CRISPR/Cas9 được chọn. Sau đó tìm kiếm vị trí gARN thích hợp từ cả hai nhiễm sắc thể giới tính.

Sự tích hợp nhiễm sắc thể Z giống cái

Các vùng 9156874-9161874, như được thể hiện bằng trình tự SEQ ID NO:15, 27764943-27769943, như được thể hiện bằng trình tự SEQ ID NO:16, 42172748-42177748, như được thể hiện bằng trình tự SEQ ID NO:17, 63363656-63368656, như được thể hiện bằng trình tự SEQ ID NO:18 và 78777477-78782477, như được thể hiện bằng trình tự SEQ ID NO:19 của nhiễm sắc thể Z của giống cái gà được phân tích để

thiết kế ARN dẫn đường. Cụ thể hơn, để định hướng sự tích hợp RFP vào nhiễm sắc thể Z, gARN được chỉ định gARN7 của locus nhiễm sắc thể Z chrZ_42174515_-1, chứa trình tự axit nucleic GTAATACAGAGCTAAACCAG, cũng như được thể hiện bằng trình tự SEQ ID NO:26, được tạo ra sau đó.

Thực hiện quy trình theo sách hướng dẫn sử dụng “PrecisionV Cas9 SmartNuclease Vector System”. Cụ thể hơn, hai đoạn môi được thiết kế để tách dòng:

ChZgARN_Xuôi: TGTATGAGACCACGTAATACAGAGCTAAACCAG, như được thể hiện bằng trình tự SEQ ID NO. 53, ChZgARN_Ngược: AAACCTGGTTTAGCTCTGTATTACGTGGTCTCA, như được thể hiện bằng trình tự SEQ ID NO. 54.

Các đoạn môi được ủ để tạo ra cặp đôi trong hỗn hợp phản ứng chứa 5 μ M mỗi đoạn môi. Ủ hỗn hợp ở nhiệt độ 95°C trong thời gian 5 phút và được lấy ra để làm nguội xuống đến RT.

Sau đó cặp đôi được tách dòng vào vector Cas9 đã được tạo mạch thẳng được tạo ra bằng phản ứng gắn được thiết lập như sau: 1 μ l vector đã được tạo mạch thẳng, 3 μ l hỗn hợp oligo được ủ, 1 μ l chất đệm gắn và 0,25 μ l Fast ligaza. Ủ hỗn hợp trong thời gian từ 5 đến 7 phút ở nhiệt độ 25°C.

Biến nạp hỗn hợp vào tế bào khả biến DH5 α và nuôi cấy tế bào trên đĩa LB chứa 50 μ g/ml Kanamycin và ủ qua đêm ở nhiệt độ 37°C.

Ngày hôm sau, nhặt ngẫu nhiên hai khuẩn lạc và cho sinh trưởng trong môi trường LB/Kanamycin qua đêm ở nhiệt độ 37°C kèm theo lắc.

Ngày hôm sau, ADN plasmit được tạo ra bằng cách sử dụng kit Qiagen mini-prep và được giải trình tự bằng cách sử dụng đoạn môi giải trình tự được cung cấp.

Theo một số phương án khác nữa, bốn ARN dẫn đường bổ sung được chọn lọc, tổng hợp và tách dòng riêng rẽ vào vector Cas9 SmartNuclease chứa Cas9 nucleaza kiểu đại (Horizon) theo Quy trình tách dòng không giới hạn: gARN3: ACAGACCTATGATATGT, như được thể hiện bằng trình tự SEQ ID NO. 11; gARN4: CGATTATCACTCACAAG, như được thể hiện bằng trình tự SEQ ID NO. 12; gARN5: CTGGTTAGCATGGGGAC, như được thể hiện bằng trình tự SEQ ID NO. 13; gARN6: GTAAAGAGTCAGATACA, như được thể hiện bằng trình tự SEQ ID NO. 14.

Các ví dụ không làm giới hạn khác đối với trình tự gARN phù hợp cho sự tích

hợp vào locus đặc hiệu trong nhiễm sắc thể Z có thể bao gồm, nhưng không chỉ giới hạn ở, gARN8 của locus nhiễm sắc thể Z chrZ_9157091_1, chứa trình tự axit nucleic ACAGACCTATGATATGTGAG, cũng như được thể hiện bằng trình tự SEQ ID NO:27, gARN9 của locus nhiễm sắc thể Z chrZ_27767602_-1, chứa trình tự axit nucleic GAGCTTGTGAGTGATAATCG, cũng như được thể hiện bằng trình tự SEQ ID NO:28, gARN10 của locus nhiễm sắc thể Z chrZ_78779927_1, chứa trình tự axit nucleic GTAAAGAGTCAGATACACAG, cũng như được thể hiện bằng trình tự SEQ ID NO: 29, và gARN11 của locus nhiễm sắc thể Z chrZ_63364946_-1, chứa trình tự axit nucleic CAGTGGGTACTGAAGCTGTG cũng như được thể hiện bằng trình tự SEQ ID NO: 30.

Các gARN này có một vài vị trí ngoài-dịch được dự đoán, không có thành phần nào trong số chúng ở trong trình tự mã hóa đã biết.

Sự tích hợp nhiễm sắc thể W giống cái

Vùng 1022859-1024215 của nhiễm sắc thể W của gà giống cái, chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 3, được phân tích để thiết kế ARN dẫn đường. Hai ARN dẫn đường được chọn lọc, tổng hợp và tách dòng riêng rẽ vào vector Cas9 SmartNuclease chứa Cas9 nucleaza kiểu đại (Horizon) theo Quy trình tách dòng không giới hạn: gARN1: GCACTAGGAACCAGCAGCAG, như được thể hiện bằng trình tự SEQ ID NO. 1 và gARN2: GTAGCCCCAAGAGGGCTAGG, như được thể hiện bằng trình tự SEQ ID NO. 2.

Các tham số được dự đoán của hai gARN này được thể hiện trong Bảng 7:

Bảng 7

Các tham số gARN	gARN1	gARN2
Bộ thiết kế sgARN	0,506	0,63
Điểm	0,8677	0,5323
Điểm sgARN	94,8	99,9

Ví dụ 4

Thiết kế vector hướng đích RFP

Tính đồng nhất trình tự kèm hai bên của các trình tự kèm hai bên thích hợp được biểu thị trên đây của nhiễm sắc thể giống cái W hoặc ở locus nhiễm sắc thể giống

cái Z, được đưa vào vector biểu hiện RFP, cụ thể là plasmid pDsRed1-N1.

Để, tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 31, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 32, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN7 trong trình tự SEQ ID NO:26.

Cụ thể hơn, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN7, cụ thể là GTAATACAGAGCTAAACCAG, như được thể hiện bằng trình tự SEQ ID NO:26, việc tách dòng ‘nhánh bên trái’ được thực hiện ở phía trước gen tăng cường CMV ở plasmid pDsRed1-N1, bằng cách sử dụng phương pháp RF.

Sử dụng các đoạn mỗi sau đây:

Đoạn mỗi xuôi:

ACGGGGTCATTAGTTCATAGCCCATATATGGTGAACGGATAGGCAGCAAG C, như được thể hiện bằng trình tự SEQ ID NO. 55 (chứa trình tự nhánh bên trái 5' và 30 nucleotit ở vị trí tích hợp ở plasmid).

Đoạn mỗi ngược:

GGCGGGCCATTTACCGTAAGTTATGTAACGTGAGGAAGGGTCTGTTACTGG A, như được thể hiện bằng trình tự SEQ ID NO. 56.

Nhánh bên phải được tách dòng ở phía sau gen kháng Neo bằng cách sử dụng các đoạn mỗi sau đây.

Đoạn mỗi xuôi:

TTCTATCGCCTTCTTGACGAGTTCTTCTGAGTCATGTCGGTGGAGGAGAAA, như được thể hiện bằng trình tự SEQ ID NO. 57 (trình tự kháng Neo 3' được đánh dấu màu xanh lục)

Đoạn mỗi ngược:

GTGATGGCAGGTTGGGCGTCGCTTGGTCGGGCATGGGATGTTAAAGAGAA GCT, như được thể hiện bằng trình tự SEQ ID NO. 58 (Ở trình tự vector màu cam 3' đến gen Neo).

Việc tách dòng được xác nhận bằng cách giải trình tự để đảm bảo sự tích hợp phù hợp.

Lưu ý rằng có sự làm khuyết 193 nucleotit khỏi gen tăng cường CMV do sự lặp

lại 16 nucleotit (CGGTAAATGGCCCGCC, như được thể hiện bằng trình tự SEQ ID NO. 59) ở vùng gen tăng cường kèm hai bên vị trí tích hợp.

Theo một số phương án khác nữa, gen thông báo RFP có thể được tách dòng để sử dụng gARN3, như được thể hiện bằng trình tự SEQ ID NO. 11, gARN4 : như được thể hiện bằng trình tự SEQ ID NO. 12, gARN5, như được thể hiện bằng trình tự SEQ ID NO. 13, gARN6, như được thể hiện bằng trình tự SEQ ID NO. 14.

Để tách dòng bằng cách sử dụng gARN3, "nhánh bên trái" chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 41, và "nhánh bên phải" chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 42 được cung cấp. Để tách dòng bằng cách sử dụng gARN4, "nhánh bên trái" chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 43, và "nhánh bên phải" chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 44 được cung cấp. Để tách dòng bằng cách sử dụng gARN5, "nhánh bên trái" chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 45, và "nhánh bên phải" chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 46 được cung cấp. Để tách dòng bằng cách sử dụng gARN6, "nhánh bên trái" chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 47, và "nhánh bên phải" chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 48 được cung cấp.

Theo các phương án khác, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 33, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 34, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN8 trong trình tự SEQ ID NO:27. Theo các phương án khác nữa, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 35, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 36, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN9 trong trình tự SEQ ID NO:28. Theo các phương án khác, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 37, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 38, có thể được sử

dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN10 trong trình tự SEQ ID NO:29. Theo phương án khác nữa, để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu trong nhiễm sắc thể Z, nhánh bên trái chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 39, và nhánh bên phải chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 40, có thể được sử dụng để tích hợp gen thông báo theo sáng chế vào locus đặc hiệu được định hướng bởi gARN11 trong trình tự SEQ ID NO:30.

Đối với nhiễm sắc thể W giống cái, gen thông báo, cụ thể là RFP có thể được tách dòng để sử dụng đoạn dẫn đường 1 (gARN1), như được thể hiện bằng trình tự SEQ ID NO. 1 hoặc đoạn dẫn đường 2 (gARN2): như được thể hiện bằng trình tự SEQ ID NO. 2. Để tách dòng bằng cách sử dụng gARN1, "nhánh bên trái" chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 4, và "nhánh bên phải" chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 5 được cung cấp. Để tách dòng bằng cách sử dụng gARN2, "nhánh bên trái" chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 6, và "nhánh bên phải" chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 7 được cung cấp.

Ngoài ra, "nhánh bên trái" đối với vùng ở phía trước gen khởi đầu CMV chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 8, và "nhánh bên phải" đối với vùng xuôi dòng sự kháng Neomycin, có thể chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 9, hoặc SEQ ID NO.10 đối với vùng ở phía sau vị trí polyA.

Ví dụ 5

Sự tích hợp gen thông báo RFP vào nhiễm sắc thể Z của dòng tế bào gà giống cái

Dòng tế bào gà giống cái được đồng-chuyển nhiễm bằng hai plasmit tức là pDsRed chứa các nhánh bên trái và phải nhiễm sắc thể Z gà và plasmit chứa CMV-hspCas9-H1-gARN được nêu chi tiết trong phần Quy trình thí nghiệm.

Trình tự gARN được dùng là gARN 7 như được thể hiện bằng trình tự SEQ ID NO: 26. Bảng 8 cung cấp chi tiết hơn về ARN dẫn đường được sử dụng.

Bảng 8. Đặc tính của gARN 7 được sử dụng

đoạn tiền đệm gARN	sợi	vị trí_	nhiễm	id_dẫn đường
--------------------	-----	---------	-------	--------------

		cắt	sắc thể	
GTAATACAGAGCT AAACCAG, SEQ ID NO. 26	-1	42174 515	chrZ	ZNF367_HABP4_ENSGALG0 0000012629_ENSGALG00000 012628_chrZ_42174515_-1

Sau đó, kích thích tế bào đã được chuyển nhiễm bằng huỳnh quang màu đỏ. Như có thể nhìn thấy trên **Fig.14** sự chuyển nhiễm là dương tính. Thu gom các tế bào chiếu sáng, chiết tách ADN hệ gen và thực hiện thử nghiệm PCR trượt (như được mô tả trong phần Quy trình thí nghiệm). Sau đó đưa các mảnh PCR đi giải trình tự SANGER và thu được các trình tự được tích hợp như được minh họa trên **Fig.15**, chứng minh sự tích hợp thành công gen RFP vào nhiễm sắc thể Z.

Tóm lại, các kết quả này chứng minh rằng:

- quy trình sinh trưởng thành công của dòng tế bào gà giống cái đã được thiết lập
- sự chuyển nhiễm thành công gen thông báo RFP và hệ thống CRISPR Cas9 vào dòng tế bào gà giống cái đã được thiết lập
- tín hiệu phát huỳnh quang mạnh được tạo ra trong sự kích thích màu đỏ
- sự tích hợp định hướng thành công gen thông báo RFP vào nhiễm sắc thể Z đã được thiết lập

Ví dụ 6

Sự truyền dẫn dòng mầm của tế bào đã xử lý CRISPR

Việc sử dụng tế bào mầm nguyên thủy (PGC) để tạo ra gà chuyển gen có nhiều lợi thế trong việc tạo mẫu và công nghệ sinh học động vật. Bởi vì phôi gà là dạng dễ trứng, PGC ở giai đoạn sớm có thể dễ dàng tiếp cận được và có thể được sao chép *in vitro* cho ứng dụng thực tiễn, bao gồm sự khôi phục vật liệu di truyền, chỉnh sửa hệ gen, và nghiên cứu chuyển gen. Các nỗ lực đáng kể đã được thực hiện để thiết lập hệ thống nuôi cấy đối với PGC gà từ các nguồn gốc phôi khác nhau, và đã được chứng minh rằng chỉ có PGC gà có thể được duy trì vô hạn định *in vitro* mà không làm mất đặc tính của chúng.

Cụ thể hơn, ở gà, PGC trước hết phân tách ở ngoại phôi bì trong phôi Eyal-Giladi và Kochav (EGK) giai đoạn X và sau đó di chuyển xuống hạ phôi bì của vùng thấu quang, và sau đó đến vòng cung mầm và đi vào dòng máu. Sau khi di chuyển qua

hệ tuần hoàn, chúng đến cung sinh dục.

Các đặc tính khác biệt này trong sự phát triển dòng mầm làm cho có thể sử dụng chúng cho việc tạo ra gà chuyển gen thông qua sự tiêm PGC đã được thao tác di truyền vào mạch máu của trứng thể nhận.

Hai vector đã mô tả trên đây, cụ thể là vector gARN/Cas9 và vector gen thông báo được đồng-chuyển nhiễm vào PGC như được nêu chi tiết trong phần quy trình thí nghiệm. Sau khi nhận biết các dòng ổn định, tế bào được nhân rộng và được xác nhận sự tích hợp RFP bằng PCR. Các dòng đã xác nhận được tiêm vào phôi thể nhận ở gà ở giai đoạn 14–16 (H&H). Phôi đã tiêm được chuyển vào vỏ thay thế và được ủ cho đến khi nở ở nhiệt độ 37°C. Giới tính của các con non được xác định sau khi nở bằng PCR đối với nhiễm sắc thể Z hoặc nhiễm sắc thể W.

Thể khảm giống cái và giống đực được cho sinh trưởng đến khi trưởng thành sinh dục và được lai với gà giống cái và giống đực kiểu dại. Các con non nở ra được đánh giá về sự biểu hiện RFP, và đời sau dòng mầm được xác nhận bằng PCR là mang RFP được hướng đích.

YÊU CẦU BẢO HỘ

1. Phương pháp xác định giới tính của phôi chim trong trứng chưa nở chứa phôi ở trong vỏ trứng toàn vẹn về mặt cấu trúc, phương pháp này bao gồm các bước:

- (a) tạo ra ít nhất một đối tượng thuộc loài chim chuyển gen chứa ít nhất một gen thông báo ngoại sinh được tích hợp vào ít nhất một vị trí hoặc địa điểm ở trong ít nhất một trong hai nhiễm sắc thể giới tính Z và W, trong đó gen thông báo nêu trên là protein phát huỳnh quang màu đỏ (Red Fluorescent Protein: RFP);
- (b) thu được ít nhất một trứng chưa nở chứa phôi nêu trên nằm trong vỏ toàn vẹn về mặt cấu trúc từ đối tượng thuộc loài chim chuyển gen nêu trên; và
- (c) xác định trong phôi ở trong vỏ toàn vẹn về mặt cấu trúc của trứng chưa nở nêu trên có ít nhất một tín hiệu có thể phát hiện được phát hiện hay không, trong đó sự phát hiện ít nhất một tín hiệu có thể phát hiện này chỉ ra sự biểu hiện của ít nhất một gen thông báo trong phôi ở trong vỏ toàn vẹn về mặt cấu trúc của trứng chưa nở nêu trên, nhờ đó chỉ ra sự có mặt của nhiễm sắc thể Z hoặc nhiễm sắc thể W trong phôi này,

nhờ đó xác định giới tính của phôi chim trong trứng chưa nở chứa phôi ở trong vỏ toàn vẹn về mặt cấu trúc.

2. Phương pháp theo điểm 1, trong đó gen thông báo nêu trên mã hóa protein RFP có chiều dài bước sóng kích thích nằm trong khoảng từ 500 đến 650nm và chiều dài bước sóng phát xạ nằm trong khoảng từ 550 đến 650nm.

3. Phương pháp theo điểm 1 hoặc 2, trong đó việc xác định liệu tín hiệu có thể phát hiện được phát hiện hay không bao gồm bước đưa trứng chưa nở chứa phôi ở trong vỏ toàn vẹn về mặt cấu trúc nêu trên vào nguồn sáng.

4. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 3, trong đó ít nhất một đối tượng thuộc loài chim chuyển gen này là đối tượng thuộc loài chim giống cái, và trong đó ít nhất một gen thông báo nêu trên được tích hợp vào ít nhất một vị trí của nhiễm sắc thể giống cái Z, nhờ đó phát hiện tín hiệu có thể phát hiện, điều này chỉ ra

rằng phôi trong trứng chưa nở chứa phôi ở trong vỏ toàn vẹn về mặt cấu trúc nêu trên là giống đực.

5. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 4, trong đó ít nhất một đối tượng thuộc loài chim chuyển gen này là đối tượng thuộc loài chim giống cái, và trong đó ít nhất một gen thông báo nêu trên được tích hợp vào ít nhất một vị trí của nhiễm sắc thể giống cái W, nhờ đó phát hiện tín hiệu có thể phát hiện, điều này chỉ ra rằng phôi trong trứng chưa nở chứa phôi ở trong vỏ toàn vẹn về mặt cấu trúc nêu trên là giống cái.

6. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 5, trong đó ít nhất một gen thông báo được tích hợp vào nhiễm sắc thể giới tính của động vật thuộc loài chim chuyển gen này bằng cách sử dụng ít nhất một nucleaza được thiết kế theo chương trình (programmable engineered nuclease: PEN).

7. Phương pháp theo điểm 6, trong đó PEN là hệ thống đoạn lặp lại đối xứng ngắn cách đều thành cụm (clustered regularly interspaced short palindromic repeat: CRISPR) kiểu II.

8. Phương pháp theo điểm 7, trong đó ít nhất là một gen thông báo được tích hợp vào nhiễm sắc thể giới tính này của đối tượng thuộc loài chim chuyển gen bằng cách sửa chữa theo hướng tương đồng (homology directed repair: HDR) qua trung gian ít nhất một hệ thống CRISPR/endonucleaza 9 kết hợp CRISPR (Cas9).

9. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 8, trong đó ít nhất một gen thông báo được tích hợp vào nhiễm sắc thể giới tính của đối tượng thuộc loài chim chuyển gen nêu trên bằng cách cho tiếp xúc hoặc đồng-chuyển nhiễm ít nhất một tế bào của đối tượng thuộc loài chim này hoặc ít nhất một tế bào được đưa vào động vật thuộc loài chim, với:

- (a) ít nhất một protein Cas9 hoặc ít nhất một trình tự axit nucleic thứ nhất chứa ít nhất một trình tự axit nucleic mã hóa ít nhất một protein Cas9; và ARN dẫn đường (guide RNA: gARN) hướng đích ít nhất một đoạn tiền đệm trong

ít nhất một nhiễm sắc thể giới tính Z và/hoặc W, hoặc ít nhất một trình tự axit nucleic mã hóa ít nhất một gARN; và

(b) ít nhất một trình tự axit nucleic thứ hai chứa ít nhất một gen thông báo nêu trên.

10. Phương pháp theo điểm 9, trong đó ít nhất một gen thông báo trong trình tự axit nucleic thứ hai nêu trên được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng đối với HDR ở vị trí tích hợp.

11. Phương pháp theo điểm bất kỳ trong số các điểm 9 và 10, trong đó ít nhất một gen thông báo trong trình tự axit nucleic thứ hai nêu trên được liên kết khi hoạt động với gen bất kỳ trong số gen khởi đầu đặc hiệu giới tính, gen khởi đầu đặc hiệu phôi và gen khởi đầu có thể cảm ứng.

12. Phương pháp theo điểm bất kỳ trong số các điểm từ 9 đến 11, trong đó ít nhất một gen thông báo nêu trên được tích hợp vào ít nhất một vùng không mã hóa của nhiễm sắc thể giới tính nêu trên.

13. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 12, trong đó ít nhất một gen thông báo nêu trên được tích hợp vào ít nhất một vị trí ở locus nhiễm sắc thể giới tính Z 42172748-42177748.

14. Phương pháp theo điểm bất kỳ trong số các điểm 9 và 13, trong đó ít nhất một gARN chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 26.

15. Phương pháp theo điểm 14, trong đó ít nhất một gen thông báo được chứa trong trình tự axit nucleic thứ hai nêu trên được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng chứa trình tự axit nucleic như được thể hiện tương ứng bằng trình tự SEQ ID NO. 31 và 32.

16. Phương pháp theo điểm 9, trong đó ít nhất một gen thông báo nêu trên được tích hợp vào ít nhất một vị trí ở locus nhiễm sắc thể giới tính W 1022859-1024215, trong đó ít nhất một trong số:

(a) gARN nêu trên chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 1, và ít nhất một gen thông báo được chứa trong trình tự axit nucleic thứ hai nêu trên được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng chứa trình tự axit amin như được thể hiện tương ứng bằng trình tự SEQ ID NO. 4 và 5; và

(b) gARN nêu trên chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 2, và ít nhất một gen thông báo nêu trên được chứa trong trình tự axit nucleic thứ hai được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng chứa trình tự axit amin như được thể hiện tương ứng bằng trình tự SEQ ID NO. 6 và 7.

17. Kit bao gồm:

(a) ít nhất một protein Cas9 hoặc ít nhất một trình tự axit nucleic thứ nhất chứa ít nhất một trình tự axit nucleic mã hóa ít nhất một protein Cas9 nêu trên; và ít nhất một gARN hướng đích ít nhất một đoạn tiền đệm trong ít nhất một nhiễm sắc thể giới tính Z và/hoặc W, hoặc trình tự axit nucleic mã hóa ít nhất một gARN; và

(b) ít nhất một trình tự axit nucleic thứ hai chứa ít nhất một gen thông báo, trong đó gen thông báo này là protein phát huỳnh quang màu đỏ (RFP), trong đó sự biểu hiện của ít nhất một gen thông báo nêu trên là thích hợp để xác định giới tính của phôi chim trong trứng chưa nở chứa phôi ở trong vỏ toàn vẹn về mặt cấu trúc.

18. Kit theo điểm 17, trong đó ít nhất một gen thông báo nêu trên mã hóa protein RFP có chiều dài bước sóng kích thích nằm trong khoảng từ 500 đến 650nm và chiều dài bước sóng phát xạ nằm trong khoảng từ 550 đến 650nm.

19. Kit theo điểm bất kỳ trong số các điểm 17 và 18, trong đó ít nhất một gen thông báo trong trình tự axit nucleic thứ hai nêu trên được liên kết khi hoạt động với gen bất kỳ trong số gen khởi đầu đặc hiệu giới tính, gen khởi đầu đặc hiệu phôi và gen khởi đầu có thể cảm ứng.

20. Kit theo điểm bất kỳ trong số các điểm từ 17 đến 19, trong đó ít nhất một gen thông báo nêu trên được làm thích ứng để tích hợp vào ít nhất một vùng không mã hóa của nhiễm sắc thể giới tính nêu trên.

21. Kit theo điểm 20, trong đó ít nhất một gen thông báo được chứa trong trình tự axit nucleic thứ hai nêu trên được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng chứa trình tự axit amin như được thể hiện tương ứng bằng trình tự SEQ ID NO. 31 và 32, và trong đó ít nhất một gARN chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 26.

22. Kit theo điểm bất kỳ trong số các điểm từ 17 đến 20, trong đó ít nhất một thành phần trong số:

(a) gARN nêu trên chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 1, và trong đó ít nhất một gen thông báo được chứa trong trình tự axit nucleic thứ hai nêu trên được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng chứa trình tự axit amin như được thể hiện tương ứng bằng trình tự SEQ ID NO. 4 và 5; và

(b) gARN nêu trên chứa trình tự axit nucleic như được thể hiện bằng trình tự SEQ ID NO. 2, và ít nhất một gen thông báo nêu trên được chứa trong trình tự axit nucleic thứ hai được kèm hai bên ở đầu 5' và 3' của chúng bằng các nhánh tương đồng chứa trình tự axit amin như được thể hiện tương ứng bằng trình tự SEQ ID NO. 6 và 7.

DANH MỤC TRÌNH TỰ

<110>	EGGXYT LTD.	
<120>	PHƯƠNG PHÁP VÀ KIT DÙNG ĐỂ XÁC ĐỊNH GIỚI TÍNH CỦA PHÔI CHIM TRONG TRỨNGCHỨA NỮ	
<130>	2562455	
<150>	US 62/510,921	
<151>	2017-05-25	
<160>	67	
<170>	PatentIn phiên bản 3.5	
<210>	1	
<211>	20	
<212>	ADN	
<213>	Trình tự nhân tạo	
<220>		
<223>	Đoạn dẫn đường 1 đối với nhiễm sắc thể W	
<400>	1	
	gcactaggaa ccagcagcag	20
<210>	2	
<211>	20	
<212>	ADN	
<213>	Trình tự nhân tạo	
<220>		
<223>	Đoạn dẫn đường 2 đối với nhiễm sắc thể W	
<400>	2	
	gtagcccaa gagggctagg	20
<210>	3	
<211>	1460	
<212>	ADN	
<213>	Trình tự nhân tạo	
<220>		
<223>	Khoảng giới hạn galGal5_ADN =chrW:1022859-1024215	
<400>	3	
	tcgggctaac attagacccg ttgtgtctta atcctgccgt cgggtataca tataactaaa	60
	gaacctcctt tccctcctat aaatcggagc gagacaaggc attctttttt tattattatt	120
	ttggttgctt gcccatcaga tgtaatgaaa gttgcacagg gtagggctaa gaaggtaaag	180
	gagttaaatt ccccgaggag gttagagatt ccctgtattg tggttgcttg cccattagac	240
	gtaacaaaag ttgtgaaggg taggactaag gatccaggta aagaggtag tccccaggg	300
	ttagagtccc tcctcagaag taaatgtaaa acataaccac ggtaattgtc gttgagtgcg	360

cttgtttgat tgttaattgt ttgtgctgtt attgtatgct gattgtttgt gttgttatta 420
 tacaaagagc tcacatgcat cgttaaaatg ggcataggat cacctagggg tgagggaatt 480
 ttgaaaaagt cacttttggg ctgtgtttta agtcattgga aagaaatagc tggatcccct 540
 ggtgggggtt cacagaagga tgatttaagt acgatggatg cctttgaggc aggcactagg 600
 aaccagcagc agtggcagat ttgtgatcat ggaaagaatt ggtgggagca tacggcacca 660
 ggactcagat aagatagcta gctaaggatga tgaattgtag atgggtgtcag ataacgaaag 720
 aagaaaaaca aatgtaataa agtgttaaat gccaatatgt tttttataat gaggcaatat 780
 cctgaatggc attagtgtgg gatcaatttg gccagatga ccctctgatt ttaatggtag 840
 aaaatgataa gagagagagg gagaaaggaa gagagtgtaa aacgtgttgt tcagcatgta 900
 gtattggact gagatgcttc aagcgaaaca agaattgagca ggaggaggat ttagagatgc 960
 tagtagcccc aagagggcta gggggcctca aaatcatgga acccaactcc ctggagtaga 1020
 agggtcaggt aatgagtcca aaagaggagc tataggaacc atgccagtaa cagagaggac 1080
 tcacagccag catagggctg tgctgtaagc cctgttgagg cagttggaaa tgacagtctg 1140
 gtagtagtta aggtaccctt ttaaattaca gatttaaaca tttggaaagc agctgctggc 1200
 agctaccgtg atgaccta acgggtagct aatgcttttg aaatgatgat taaaactcag 1260
 gaactggata ggaaagatat ggaatttatt atgcacatgt tgtttgatag tacagaaaaa 1320
 taaatgattc actagaccac acggaccaa gtggaggatc aggtaatggc aggggttttg 1380
 cttggatgta atgctggaaa ggacaagagt cctttcctgt cctggggagg atttagaggt 1440
 ggtagacct tggggaggag 1460

<210> 4

<211> 592

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Để tách dòng bằng cách sử dụng gARN1, nhiễm sắc thể W, "nhánh bên trái"

<400> 4

tcgggctaac attagacccg ttgtgtctta atcctgccgt cgggtataca tatactaaaa 60
 gaacctcctt tccctcctat aaatcgagc gagacaaggc attctttttt tattattatt 120
 ttggttgccct gccatcaga tgtaatgaaa gttgcacagg gtagggctaa gaaggtaaag 180
 gagttaaatt cccccgggag gttagagatt ccctgtattg tggttgcctg cccattagac 240
 gtaacaaaag ttgtgaagg taggactaag gatccaggta aagaggtag tccccaggg 300
 ttagagtccc tcctcagaag taaatgtaaa acataaccac ggtaattgtc gttgagtgcg 360

cttgtttgat tgttaattgt ttgtgctgtt attgtatgct gattgtttgt gttgttatta 420
 tacaaagagc tcacatgcat cgttaaaatg ggcatggat cacctagggg tgagggaatt 480
 ttgaaaaagt cacttttggg ctgtgtttta agtcattgga aagaaatagc tggatcccct 540
 ggtgggggtg cacagaagga tgatttaagt acgatggatg cctttgaggc ag 592

<210> 5
 <211> 848
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Để tách dòng bằng cách sử dụng gARN1, nhiễm sắc thể W, "nhánh bên phải"

<400> 5
 tggcagattt gtgatcatgg aaagaattgg tgggagcata cggcaccagg actcagataa 60
 gatagctagc taaggtgatg aattgtagat ggtgtcagat aacgaaagaa gaaaaacaaa 120
 tgtaataaag tgttaaagtc caatatgttt ttataatga ggcaatatcc tgaatggcat 180
 tagtgtggga tcaatttggc ccagatgacc ctctgatttt aatggtagaa aatgataaga 240
 gagagaggga gaaaggaaga gagtgtaaaa cgtgttggtc agcatgtagt attggactga 300
 gatgcttcaa gcgaaacaag aatgagcagg aggaggattt agagatgcta gtagcccaa 360
 gagggctagg gggcctcaa atcatggaac ccaactccct ggagtagaag ggtcaggtaa 420
 tgagtccaaa agagggacta taggaacat gccagtaaca gagaggactc acagccagca 480
 tagggctctg ctgtaagccc tgttgaggca gttggaaatg acagtgcgtt agtagttaag 540
 gtaccctttt aaattacaga tttaaactt tggaaagcag ctgctggcag ctaccgtgat 600
 gatcctaaac gggtagctaa tgcttttgaa atgatgatta aaactcagga actggatagg 660
 aaagatatgg aattttattat gcacatgttg tttgatagta cagaaaaata aatgattcac 720
 tagaccacac ggacccaagt ggaggatcag gtaatggcag gggttttgct tggatgtaat 780
 gctggaaagg acaagagtcc tttcctgtcc tggggaggat ttagaggttg ttagaccttg 840
 gggaggag 848

<210> 6
 <211> 962
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Để tách dòng bằng cách sử dụng gARN2, nhiễm sắc thể W, "nhánh bên trái"

<400> 6
 tcgggctaac attagaccgc ttgtgtctta atcctgccgt cgggtataca tataactaaa 60

gaacctcctt tccctcctat aaatcggagc gagacaaggc attctttttt tattattatt 120
 ttggttgctt gcccatcaga tgtaatgaaa gttgcacagg gtagggctaa gaaggtaaag 180
 gagttaaatt cccccgggag gttagagatt ccctgtattg ttggttgctg cccattagac 240
 gtaacaaaag ttgtgaaggg taggactaag gatccaggta aagaggttag tccccaggg 300
 ttagagtccc tcctcagaag taaatgtaaa acataaccac ggtaattgtc gttgagtgcg 360
 cttgtttgat tgttaattgt ttgtgctgtt attgtatgct gattgtttgt gttgttatta 420
 tacaagagc tcacatgcat cgttaaaatg ggcatggat cacctagggg tgagggaatt 480
 ttgaaaaagt cacttttggg ctgtgtttta agtcattgga aagaaatagc tggatcccct 540
 ggtgggggtt cacagaagga tgatttaagt acgatggatg cctttgaggc aggcactagg 600
 aaccagcagc agtggcagat ttgtgatcat ggaaagaatt ggtgggagca tacggcacca 660
 ggactcagat aagatagcta gctaaggtga tgaattgtag atggtgtcag ataacgaaag 720
 aagaaaaaca aatgtaataa agtgttaaat gccaatatgt tttttataat gaggcaatat 780
 cctgaatggc attagtgtgg gatcaatttg gcccagatga ccctctgatt ttaatggtag 840
 aaaatgataa gagagagagg gagaaaggaa gagagtgtaa aacgtgttgt tcagcatgta 900
 gtattggact gagatgcttc aagcgaaaca agaattgagca ggaggaggat ttagagatgc 960
 ta 962

<210> 7

<211> 478

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Để tách dòng bằng cách sử dụng gARN2, nhiễm sắc thể W, "nhánh bên phải"

<400> 7

gggcctcaaa atcatggaac ccaactccct ggagtagaag ggtcaggtaa tgagtccaaa 60
 agaggggacta taggaaccat gccagtaaca gagaggactc acagccagca tagggctgtg 120
 ctgtaagccc tgttgaggca gttggaaatg acagtctggt agtagttaag gtaccctttt 180
 aaattacaga tttaaacatt tggaaagcag ctgctggcag ctaccgtgat gacctaacc 240
 gggtagctaa tgcttttgaa atgatgatta aaactcagga actggatagg aaagatatgg 300
 aattttattat gcacatgttg ttgatagta cagaaaaata aatgattcac tagaccacac 360
 ggacccaagt ggaggatcag gtaatggcag gggttttgct tggatgtaat gctggaaagg 420
 acaagagtcc tttcctgtcc tggggaggat ttagaggtgg ttagaccttg gggaggag 478

<210> 8
 <211> 669
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> "Nhánh bên trái" đối với vùng ở phía trước gen tăng cường CMV

<400> 8
 aggcaaggct tgaccgacaa ttgcatgaag tcgggctaac attagaccgc ttgtgtctta 60
 atcctgccgt cgggtataca tataactaaaa gaacctcctt tccctcctat aaatcggagc 120
 gagacaaggc attctttttt tattattatt ttggttgctt gccatcaga tgtaatgaaa 180
 gttgcacagg gtagggctaa gaaggtaaag gagttaaatt ccccgaggag gttagagatt 240
 ccctgtattg tggttgcctg cccattagac gtaacaaaag ttgtgaaggg taggactaag 300
 gatccaggta aagaggttag tccccagggt ttagagtccc tcctcagaag taaatgtaaa 360
 acataaccac ggtaattgtc gttgagtgcg cttgtttgat tgtaattgt ttgtgctgtt 420
 attgtatgct gattgtttgt gttgttatta tacaagagc tcacatgcat cgttaaaatg 480
 ggcatggat cacctagggg tgagggaatt ttgaaaaagt cacttttggg ctgtgtttta 540
 agtcattgga aagaaatagc tggatcccct ggtgggggtt cacagaagga tgatttaagt 600
 acgatggatg cctttgaggc aggcactagg aaccagcagc gatgtacggg ccagatatac 660
 gcgttgaca 669

<210> 9
 <211> 911
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> "Nhánh bên phải" đối với vùng ở phía sau kháng Neomycin

<400> 9
 ttctatgcc ttcttgacga gttcttctga cagtggcaga tttgtgatca tggaaagaat 60
 tgggtgggagc atacggcacc aggactcaga taagatagct agctaaggat atgaattgta 120
 gatggtgtca gataacgaaa gaagaaaaac aaatgtaata aagtgttaaa tgccaatatg 180
 ttttttataa tgaggcaata tcctgaatgg cattagtgtg ggatcaattt ggcccagatg 240
 accctctgat tttaatggta gaaaatgata agagagagag ggagaaagga agagagtgtg 300
 aaacgtgttg ttcagcatgt agtattggac tgagatgctt caagcgaaac aagaatgagc 360
 aggaggagga tttagagatg ctagtagccc caagagggtc agggggcctc aaaatcatgg 420
 aacccaactc cctggagtag aagggtcagg taatgagtcc aaaagaggga ctataggaac 480
 catgccagta acagagagga ctcacagcca gcatagggtc gtgctgtaag ccctgttgag 540

gcagttggaa atgacagttc ggtagtagtt aaggtaccct tttaaattac agatttaaac 600
 atttgaaag cagctgctgg cagctaccgt gatgaccta aacgggtagc taatgctttt 660
 gaaatgatga ttaaaactca ggaactggat aggaaagata tggaatttat tatgcacatg 720
 ttgtttgata gtacagaaaa ataaatgatt cactagacca cacggaccca agtggaggat 780
 caggtaatgg caggggtttt gcttggatgt aatgctggaa aggacaagag tcctttcctg 840
 tcctggggag gatttagagg tggtagacc ttggggagga gggttcgaaa tgaccgacca 900
 agcgacgccc a 911

<210> 10

<211> 911

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> "Nhánh bên phải" đối với vùng ở phía sau vị trí polyA

<400> 10

atctgtgtgt tggttttttg tgtgtctaga cagtggcaga tttgtgatca tggaaagaat 60
 tgggtgggagc atacggcacc aggactcaga taagatagct agctaagggtg atgaattgta 120
 gatgggtgtca gataacgaaa gaagaaaaac aaatgtaata aagtgttaaa tgccaatatg 180
 ttttttataa tgaggcaata tcctgaatgg cattagtgtg ggatcaattt ggcccagatg 240
 accctctgat tttaatggta gaaaatgata agagagagag ggagaaagga agagagtgtg 300
 aaacgtgttg ttcagcatgt agtattggac tgagatgctt caagcgaaac aagaatgagc 360
 aggaggagga tttagagatg ctagtagccc caagagggtc agggggcctc aaaatcatgg 420
 aaccaactc cctggagtag aagggtcagg taatgagtcc aaaagaggga ctataggaac 480
 catgccagta acagagagga ctcacagcca gcatagggtc gtgctgtaag ccctgttgag 540
 gcagttggaa atgacagttc ggtagtagtt aaggtaccct tttaaattac agatttaaac 600
 atttgaaag cagctgctgg cagctaccgt gatgaccta aacgggtagc taatgctttt 660
 gaaatgatga ttaaaactca ggaactggat aggaaagata tggaatttat tatgcacatg 720
 ttgtttgata gtacagaaaa ataaatgatt cactagacca cacggaccca agtggaggat 780
 caggtaatgg caggggtttt gcttggatgt aatgctggaa aggacaagag tcctttcctg 840
 tcctggggag gatttagagg tggtagacc ttggggagga gcaagtaaga tgcttttctg 900
 tgctgcaata g 911

<210> 11

<211> 17

<212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Đoạn dẫn đường 3 đối với nhiễm sắc thể Z

 <400> 11
 acagacctat gatatgt 17

 <210> 12
 <211> 17
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Đoạn dẫn đường 4 đối với nhiễm sắc thể Z

 <400> 12
 cgattatcac tcacaag 17

 <210> 13
 <211> 17
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Đoạn dẫn đường 5 đối với nhiễm sắc thể Z

 <400> 13
 ctggttagca tggggac 17

 <210> 14
 <211> 17
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Đoạn dẫn đường 6 đối với nhiễm sắc thể Z

 <400> 14
 gtaaagagtc agataca 17

 <210> 15
 <211> 4980
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> ENSGALG0000025945_ENSGALG0000023622|chrZ:9156874-9161874

 <400> 15
 cctgcaggac atgtggacac tctgccaggt actgggggag gatccctcc actgctccct 60
 gggagatggc aaacatctag agagggacac tgcaggatg ctgcatgaac acctccagcc 120
 ttctgcaggc tgtgatgggg tcagggtga aggtcaacac aagcaatgac tgcttgctga 180

caatgtgggc attgccgacc acagacctat gatatgtgag tggtagtgagg gccaggaga 240
 gcaggacagg agtgtggact cggaggcgcg ggcagaggag gtagctcaga gcatgagata 300
 tattcgccag gtgtcagtggt acttctggca gtgcatactg cagagagctt tggttgctgg 360
 ccatctggat gatgacaaaa tctcagccac agaccatgtg ggttggaaga gtcctcagga 420
 gatcaccttt ccacctcctg ctaaagcaca ttccctacag tatgtttcac atgacaccgt 480
 gctggcagat ttttaatatc atcagagaat gacactccat atcctctctg ggcagccttt 540
 ttcagtgtc tgtcacccac aaagtaaagt atttgctcat gttctaattg aacttcctat 600
 gttctagttt gtgaccattg aaccttggtc tgtcactgga cagtgtgaa ggagcctggc 660
 cctatccact tgactccac actttacata ttacaagca tcgataagac cccctcagt 720
 cttctccaga ctagatagcc ccaggctctc caaccattcc ttttatggga gatgcttcag 780
 gcccctcatc acctctgtgg ccctctgctg gactgcctgc agtagttcca ggttttttt 840
 tttgaactgg ggagcccaga acttgacacg gtattccaga tgtggcctca ccagcgaga 900
 ggagagggga aggctcacct ccctcgacct ggtggccatg ttttttttaa cagacctcag 960
 gataccactg gccttcttgg cccaaggga cactgctggc tcatggcca tgggttggt 1020
 tccaggactc tcaggctctt ctccgtagag ctctctcaa gcaggacaac cctcagcctg 1080
 tactgatgtg tgcggtgcgg ttattcctcc ccacgtgaa gactctacat ctgccattgt 1140
 taaatctcat aagcttctc tctgccaac tctccactgt gtacgggcct tagtgaatgg 1200
 cagcacagac ttctggtgtg aggatagcat tcagaacaag ctgctgtacc ttcccagtca 1260
 cagaggtgag ggtgactggc ttgtactttc ccagctttgc cttcttgccc ttttgaaga 1320
 ctggagcgac attggccttc tggaggaaac ggtcctcagg caccctcct gttctccatg 1380
 acctttgaaa gatgatcgag agcaccttgg cagtcacttc tgccagctcc ttcagcacgc 1440
 atgagtgcac ccgcccagtg cccatggatt tgtgttcatt gaattagcct aggtgacctc 1500
 tgatcagatc cttcttgacc aagggaaga cttcctttcc ccaggttctt atctccaggg 1560
 cctgggattc ccaaggggca atcttagcag gaaaggatga agtaaaggag gtgtgcagtt 1620
 acagtgcttt catggtgtcc tccattactc caatgaggtc acccacctca ttcagcagca 1680
 gacatacatt ctcccagtct tcctttttct actgccatac taaagtagtc cttcatgttg 1740
 tccttgactg aaggtcaccc aacttgattt ccagttagcc ttagcattcc ttgtagcatt 1800
 catgcatgtc ctgacaacat tcctattttc ttccaagtg gccagagacc tttccacac 1860
 tccagaggct ttgttcttcc attactgcct ttccatgagc ttcttgctca tctctgcagg 1920
 tctcatgcca ctttggccg atttcttagt cttagagatg caccgatctt gaatttggaa 1980

gaagtgggtgc ttgaatgaca accagctctc ttgggtccccc ttgctttcca aagctctaac 2040
 ccatggaatg cctccatgta tgtctgtgaa gaggtcaaag tcagctctcc tgaagtcgtg 2100
 gttgcaatcc tacaaactgc ctgatttctc ccatgggaga tcccgaactc cactatctca 2160
 cagtcactgc actcaaggct tccccagcc ttcacatccc cagctacagg aagagaagag 2220
 gctcgaggct acttgacgct gtctgcggaa ggctttctcca gcttcctcct cttcatcagg 2280
 cggcctgttg taaacaccca caacagcgtc acccatcata gcctgtccct tagtcctcac 2340
 ccataaactt tctactcatt catcatgcat cccttggcag agctctgtac attccagctg 2400
 ttccctcaca taaagagcaa ctccaccacc tcacaccgac gtctgtctgt cctaaaaagt 2460
 acagagccat ccctgagagc attccagtca tgcgagctgc cccaccacgt ccctctgctg 2520
 atcgcaactga gatcgcggcc ctacgacat ccatggatct atgactcttc ctgtttactc 2580
 cccgcactgt gtgtgatgtt acacaggcac ttcagagaag aaaagagagc acaggatgcc 2640
 gattttccctt ggagggaccc cagggggccgc acgcagcccc cacctctatc agcaggctctg 2700
 agcaggactg gagctgcagg cagagcacag ctacgccctg agacacctcc atgctgcctc 2760
 ctgcacagcc tcggcccccac agcccccgcc aaggagagca cgggtgggggt gccgggagcc 2820
 acgctgcgct ctgacacact gcgtgtcccc tgctcccagc gcgagctgca gggcccgtgc 2880
 aggaacgagc tgggccccaa cgtgaccctc ggccccggtc actgctgaac gacagcgctc 2940
 agctctccag gctgcaaaga gcgagccgtg aaatggagca acacgcggga ttcccgggtg 3000
 tgaattccta aacggattcg aagggtgcca acaggcagcc actgcgtccg cagctgtgct 3060
 tcggtacctg gagacccgcg agcccgtctg ttcccaccag cctccagcag ccggtgctcg 3120
 cggatcaaaa gaccgccgtg aaggccgcag ggccgtccct gggacggagc gccggccgct 3180
 tccgccgcca gccgcggctc cgccctcccc gccgagctga cgctggcagc gcgcagagcg 3240
 cggagcggag ccgcggcgct ttccggcggc agcggccgca cggggctcgc cgagggccgc 3300
 gcagagcgct gtgcccggcg ggccgcgtcg ccgctccggc ggcaccgcga ttccggcacg 3360
 gagcgcaacg agcggccgct gcccggtgc aggacgcggt cggggccaca ccgcggcggg 3420
 gccgcagcct gccttccgcg ggaccgcacg tccggcctcc ccgaatttg cccgcagcc 3480
 cccgggctgt caccgtgctg tcaccgtgcc gttcacggtt tacgggctgg ttccggcccgg 3540
 ttccggcccgg ttccgtgact ggcgggacag agggattcgt ggatccgcgc ctccgggaag 3600
 ggaaacggga ccccgaataa ttaaaaacag cggccacgat ctgaggggtt acaaattgac 3660
 taaaacgtac taaaacaaa cagctgctta acacctcta ttcaaaggc ccatcgacac 3720
 cccttggcac cctcaaagta tccccttctg aggtaatggg taccaagaat gcacggagcc 3780

cctgggccag tcacagcagg gtgcttttgc cagtctgtac cggtagaggcc tatttcggcc 3840
 acaaacacag tcaattcttg ggaacccccca ttcaccccat gaatataagt tgattctgtc 3900
 ccttggtgac tcgagggcgt gacagtgcac tgtgcaccag ggtccaccag ggccctgtat 3960
 tcctgagggt ctgatgtgcc acaccacgaa cccacacggg tcaatacact ccatcatcct 4020
 ttccccccat ggtgggggca ggggtccctct atttgttacc tctgtcgttc ttagagcggg 4080
 tgccatgtcc tgcagcaatg ggagcgatca cctctttggg gattaatctt ggtgggttaat 4140
 ctgccttgta attattttac ctgtgcttga aaatgccct ccagggtgaaa gcaaactgca 4200
 gcctgcattc tgtggccaaa ggaatggaaa agaaggcgtt agcaatgcca gtggtggcat 4260
 accacttggc ttcttttgac tccagttcat actggagttc taacatgtcc agaacagcag 4320
 cgctcagtgg ggggtgtgact tcattctggc catggtagtc tactgtcagc ctccattctg 4380
 cactggcttt atgcactgga gcacagggct tatggggagc ggcgaggga gctgggatgg 4440
 ttcagtctgc agatgaggag gctcagggga gacctattg ctctctgtaa ctccatgaag 4500
 ggaggttgta gtgagctggg ggtcggcttc ttctctcatg tgactagtgg taggactaga 4560
 gggaatggcc tccagttgca gcaggggaga ttcaggctgg atgttaggaa atgggctgga 4620
 gtgggctgcc cggggaggtg gtggagtcac cgaccctgga ggtgctcaag gaatgtttag 4680
 acaccgtgtt tagggacatg gtatagttag aactgttggg gatgggtgga cagttgggct 4740
 ggatgatctt gtaggtcttt ttaaagcttg gtgattcttt gattcgataa gtgcaaagtg 4800
 ttgcatttgg gtcagagcaa tctcacgggt gtgcacaggc tgggagaaga actcctggga 4860
 gaagaagaga gcagccctgc agagaaggac ttgacatgaa agacttaacg tcagccagac 4920
 tccctctgct ctgccctcat gagcctctcc tactctgtaa aggtttgggc cccagcaca 4980

<210> 16

<211> 4980

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> ENSGALG0000015036_ENSGALG0000000438|chrZ:27764943-27769943

<400> 16

aggagtggct tctgaggctg gaaacaaggt gctcccctct tgtcaaacag ccaattccat 60
 ctagctccaa cccatcctcc actggccaaa gctgagccca tcgctgatgg tgatagcatc 120
 tgtaatagca tctcagaaag gctaaaagca ctgcgcagca gctgagagaa aagagtgagg 180
 gaaaatggga cagaaacagc tctgcagccc tcagaggcaa tgagaggag aggcaggagg 240
 tgctccaggc atggagcatt agtttcttgc agtccaggag ttacaggtga tgggtcaggc 300

tctgtgcagc ccagtgatga ctgtttcagg atccatgctg cagcctgtgc cagagccagg	360
ggataggctt ggctgcttag tgcaggacgt ggagagccca tgctgagtgc aagctcccaa	420
aaggctagga cagaacaggt gtagtggagt agatcttgta ctgtggagca gtcttttcct	480
ggaggactaa cctgtgaaga ggaggactgc gctgaagcgg ttcatagaaga atgaaacctt	540
gtgggaggga cccacactg aaaaaggaga aaaccacgat gaggaaggag cagcagagag	600
ctccccagcc ctgtgtgcac tgccaggaga gaaggaagta gaaaacgtgg gaatgtgcag	660
gtgtaaagct gagtctggga aaaaggtggg ggtggaataa aggtattttt agtcctatac	720
ctttctcatc aacctacttt aattgtagat tggccataaa ttaaattgtt tccccaggtc	780
aagtctgttt cttgcttgat atggtggatg atctccctgt cctcatcca atcctgaagc	840
tatttcatca tatattctct ccctgccctg tttcaggaca gggagtgcaa gcagtttggc	900
agccagacga ggttaacca gcaccagagg tttagtcaca tgcaagtga agctattgcc	960
tcaacttaaa tgggagaaaa catttttagca ggttgagtct caccttgttt atccagatta	1020
tacagccatg ccagcacatg ggctgcatgc agcctggccc aaccatcct ctagccaccc	1080
cctgccccat actgaaaatg cagcggctct gccctgccaa gctgccaaagc ctggctccag	1140
gcactcacc attgcttaca aacaacaaaa cgtcagcgtg tgagcgcgcc ctgtgagctg	1200
caactggcat ggggagggca cagtcagcac aaatacatga cctgcaaatt ttcacatgga	1260
atgtatggag ttgcaatcag atactcagct caaaaaattt aatcatggct cttacctgt	1320
ctttttttta acctttctct tagctcagag aggtatacct cacttcacaa tcatgccgta	1380
ttcatgtcat tgctttctgg cagtacgtac gtttgtgaac agctgttttc aagggtgatg	1440
cacaggaaga gtaaaatttc attaaaaaat ctccaaggag cacatggaga gtcactcag	1500
aagcctcacc gctgccactg aacaagactg atgcattagt ttcataaat caaggctatg	1560
tgtcccactg gtttcattat tttgttactc tttctttaaa aaaaaaaaaat tgttttattg	1620
cttatataca tcaactatgt tgttatatat tttccaaagg ccgctccaaa agtaatgcct	1680
cctgtttcat tctgttgcc caccgtgtca gaggcagatg gtggtacggc agtagagatg	1740
gaaccttccc accaatattc cattacatgt tgccatgtga tagacagcag cagatgggca	1800
ctctgacaga atggtatctg acatggaagc gtggatggag caaagtgtg tcaactgaatt	1860
cctccataag gaaatgagca ccactaaca ttactgatg cttgctgaac atttctggag	1920
accaaacagt ggatgtgagc acagtgaggt ggtgggtggt gcatttcagc agtggtgaca	1980
atgacgttgg gtcacctctg ctggtgcagg ttttgatgag tgcagcctgt gaggtcttgt	2040
tcatcaatgg agaaaatgca caactaatgg tggtgactgc tgaaaaatag tgttttgtag	2100

ctgagaat	ttt	tctctgtcaa	acagtgctaa	ttgtgctctt	tgtatctgtt	gtagtttcca	2160
tgggaata	aaa	taggaggcat	tacttgcaga	ctgacctaca	tatatgtggc	cccaggcaag	2220
tcaacagctg	tacgaaaagt	ccatgcttcc	tcattataaa	cggaggaaaa	aaagttgttt		2280
acagctgtaa	tgggattata	aaaggcaa	at	ggggatagta	cagtggtag	aaccagatct	2340
atgaaggaaa	agccctaaga	aaaagagagt	gcacagatac	tcttaacat	ctaataactg		2400
tttcctccat	cctacagctc	agagttaaga	cttacagagg	actctagtag	ttagtaagat		2460
gaatacagagc	taatagtggc	aaaaataatc	ccagtgcctc	aacactgacc	tgggaaaaag		2520
gggcatgtat	agaccttctg	atattgtgat	gctgtgtttg	tacacttatt	atctacattt		2580
tcagaaatta	ggttaaactt	cagaaaactg	aagatctcca	gggctttag	cagaccctga		2640
ccaccagact	gggtcccgat	tatcactcac	aagctctcaa	accgattgtg	gctgtttctc		2700
agaggcaact	ctttgcgctc	tagccctctt	ataacatggg	gcaacttccc	ctgccccacc		2760
ttcccttcct	gtatcttctg	aaaagcttgt	agccctcaat	tgtcacgctc	cagccatatg		2820
aatcatccca	ccacgtccct	gtgttcacaa	ttaccttcta	gtcttcta	at	tgcaccatgg	2880
cttccaattc	ctcctgttta	tcacccacgt	tgcattccatt	gatgtagagg	cacttcagtc		2940
gggctatcag	caatgttacc	acctgtgagg	agccctcttg	agatctttta	ggacaattta		3000
gaggttttcc	ccactgtttc	ctaaaattac	agcattccct	gtcccttctt	agtgatatag		3060
aactccatcc	ccttccacca	tcaaacctag	cttaagctct	ggtaa	atgat	ccagccagtc	3120
tgtttcccaa	aactcttgct	cagttgcatt	ctatctgatg	cctacatgcc	caacaccttc		3180
aaggcctgtc	caagatcata	gaacacaaaag	gcttgatcat	gacaacatcc	atgcagccaa		3240
tcattcacat	gaccatttcc	tctcctcctt	tccagggtccc	aacggccaac	tgagagttgt		3300
gctctgagct	cttcagtgct	cttccgaggg	atgtaaagtc	ccttttaatg	ttctgcattt		3360
tctttgtcac	agccttctga	gacccaactt	gaatgtggag	gagtggagag	taatcatccc		3420
gctctgatta	tcagatacat	tagcctcttc	ttggcacctc	tggcagacag	aagacttccc		3480
tggagagatt	acctggatga	taactggggg	cctcagtgcc	tcacagcagc	aagtccccag		3540
ttactaagac	tctacacctc	tttgggtggcg	ctagttctga	tgtagttctt	cagttgctta		3600
acatgcttgt	tgtttcccaa	gatttgactg	ctatctgcac	cctccctttc	ttccagcccc		3660
agagccttat	atcatgagct	gagaacattc	tgtgttcaaa	atacaagaca	cattttatgca		3720
atttcgtatt	ctctacccat	atttttgtga	tagaaattag	ggctacttat	ttcccatagt		3780
gatttccatg	ctgcctacca	ggttctaagt	caggtcaccc	tgccagcttc	ttccaaacca		3840
tgtctcaggag	tgcatacaat	cacacgcaga	cccaaggaaa	aatgtacca	aacaatatta		3900

ccacgcatgc aaactcctca gtagctctgt cccatatgct acaattaggt ctgtataggg 3960
 ttgttagatc ttgtctactt gattagtgtt ttccagcagg cctacaatcc aacctaaact 4020
 gaacagacat taaagacatg tttcactaat acaaaacaac tccacctagt tcaacagtct 4080
 gctcttttgt gtatgggtacc tgaaactctt ctgttcactg aacttatccc aacatttttc 4140
 ctctaatgc atactttttc tagacgtgat gtgaaagtac agtggttatac cctgggtgctg 4200
 cattaatttg cataaccaca gcacacaaaag cataatcata acacacaaaag gcctcctatg 4260
 aaaacatctg cttaaagtca taggtcactt atttattctc cctctagtta tgtaccagaa 4320
 tgatgtatgt taataaataa tagattatta aatccatcaa aaattaggga agagctctag 4380
 cacaaggca aagcaaagaa gctttgaaga agtgcaagta aaatgaatta tttccaaaca 4440
 actagcttct cactgtccca ctaagaacaa agactattct tatgcactgc tgccttaaa 4500
 attatcagct gctttgtgtt tatttcattg agtaacacat gcaatatttt ttacaagaaa 4560
 taagaaacca cagcagcaga agacgaatct gcttttattg aacgcagaaa tagatggaca 4620
 ctgaaacaac agaagtttct accatgtttt gctgctgtga gagcaaactt gctgttgatc 4680
 aactattagt ttgaaaagt atatgcaact ctgctatcaa ctctttcccc acatcacatc 4740
 tcgtttcaag gagaaccttc ctgtccaggc aactagccag ggattttaat ttaacattt 4800
 gctatctcag ataattttt aatatttcta ggccgtgcac ataaccaca tagccctaac 4860
 tccagatgtt cccagcatca cttagcagta gttccacttg aagcataaat acagcactgg 4920
 cacaacaaat gaagacaatt caagcacgcc ttaacaagc atacagaaca tacagtttta 4980

<210> 17

<211> 4980

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> ENSGALG0000012629_ENSGALG0000012628|chrZ:42172748-42177748

<400> 17

acaaattaaa tgtattttct cttgatacta ttttgattat aagctattaa cttgatttgt 60
 aaaagcaccg atattatggg taatgtatgt atggtaactg ttgagtcaca gcctgaacct 120
 ctgattgacc acctgcggaa aggtctgggt gagccctggg agcacaggca aaaacaattc 180
 agttgtgcaa ctggaaggcc tctcttagac ctcatTTaag ggctgactgt cactggggaa 240
 ggatctgatt ctggagatcc cttattgggt aagctttccc ctgcaaacct agaatttct 300
 gatgcaagtg agcaatcttc ttcccttccc tagcagctct ctgttgacc agtccttcca 360
 tccttgacc tttgttgtaa tgcttttccc atggcattga tctgtccagt tgctacagt 420

tagaaacaag aaagtgtttt ctcaaaacat tatacatgct tactatcaga tcttacttaa	480
cattgataag aagtacacct ttcattacaa cagtgaacgg ataggcagca agcattcaga	540
acaatgaaaa agatctgcgt tctattgtaa agcaaaaaa ggtgatgacg gaacgtttat	600
ttgagataca ggcaccatag tacaagatat caaagaggaa ttaattcaac cctcccaaca	660
acagttggga aaaaaaaagg ataagctgag gaatgcagcc ttctccgatt tttactctgg	720
tgccggaagg aggaggagga gcgggggaaa acccagcggga ctgagactgc atacgcactc	780
tcggaagtac ctctgaggcg ggggtgagagg aagggtggat ctaggctccc cccgctccac	840
acactcacgt gctcgtgtac tggttcaggc tgtgacttgg caactcctct acatctatcc	900
atctagttag actcaacaag accatcgtga ttcaggagtc tgaagatgct taactctctc	960
agaacgtcta gttctcctgg cctggtaaata gctatgtttc tcatactgcc tctctgaaga	1020
atgcttcgaa tgcttctagt gatgctctaa agttctaaac agaaaaatct ggagacagtt	1080
ctggtcttta gatagaaaaa atgccaacat gccaaaggat ggttacatcc ttcaagcaac	1140
cttggtgcat gctgtacaat agactcatgt aataacttag ccgtagtcat cgtatctctt	1200
atttacctgt tcgttattac attttcctgg tactgcttta tatttagtca gttgtccttt	1260
taagacaaat tttttggtgt gtgctaatag gcagttacca aatgttctag agggagggaa	1320
tataatcagt gagtatgtag atgtatatag atgtataacc agtaagtata cagagaagtt	1380
agggtggtct ttgagagca tatagctgga aagattatca catgggaaaa ggtaatcaga	1440
aataaatgga aaaatgctca ccagtgtcat aggctcaca gaaaactccc caaggagcaa	1500
tctccagtaa cagacccttc ctgagtgga gtcagccctt aaatggcatc taggagaggt	1560
gtagccaggc accacccttc ctgtcacata ggtgaattgc cttcagctgt gtcctgcag	1620
ctgactcagt gttcacctca ggtgatcaat cagaggcca gaccatgatt caacatttcc	1680
catgcggagg gagatgaaac aagactgtca tctcatcttc tgaagctgat aggagaatgg	1740
ttacaataag taggtcccct tccactggtt tagctctgta ttactcagtt taaaacttta	1800
tgtagttaat ggagttccat agaatcacat cgctggaaaa tgtatgtgat agctataccc	1860
aaaaagtgac acagcatcac agcctctacc aggtcctctg cctccagaca ggaccttgag	1920
ctgtgtgatc cttttgggat gtttctgttt aagacagggg acacttacac attaggaaac	1980
tggagtcatg tcggtggagg agaaagtggc tgtagaaaag gtaacggaca agaaatccct	2040
cactggctta gctgtaaagg tcagaggaaa atcatcttct gttctccata cacatctcct	2100
gtaattaggc acctgtgccc tttaaaaaag taaagtaatc aaagagtcatt ttggtgaaaa	2160
tagaagccaa acagttacaa aatttagtta atattcaaag acaaaaacaac tctggttcca	2220

aatataaaac ccttttgta aacactgcgg gaaggtgcc aagagccacc tacacaaaga	2280
gatgccacag agtaatttgc gtacccaac aataggatca tatatgggca aacaaaaaga	2340
ccaacaaaga ccccttgcaa taatcccttg attggcagaa gcaagtgcag cgttagattt	2400
agtatgtgtg acaattctga gaataggaag gaatttcact gaaactgtgt gaagatttgt	2460
atgttaaata catataatga gggttattta gctctgggggt gtgcatgcta tgtggagaga	2520
tccccatgtg cccagcactg caatagacta atgtcagctt tttaaaccat catttggttt	2580
gaagagttta ttatgatttt cagaaacaga attatgtatt cagtgtctgt acattcttac	2640
aagcctgctg ttctgtacac atgaaaaagc tacttatggg gcaaatcagt atagagaagc	2700
agttttgaca tacagtggc atgactgggt gatctgctgt cagaaggtaa gacactggta	2760
tacagaattg caaacgtgag atgaagacct caagtgaac tctgtcatca taccacttct	2820
cttgctatta tctattcagt gtttgttcta aatattcagg cagaaggctt ggtttcacat	2880
caggtgttac atttaagtat tctttgccca ttttttgctg tttgtatgtg taacttcaga	2940
ttctcacatt gactgtgtag tgtaaatgtg caaatagatg taacagcttc tctttaacat	3000
cccatgccaa acttttaatc actttcaatc aatgttgga atctgcagta cataaatatt	3060
attttgattc attattcaga gaatcagaga aataaggact tttttagagg tttatgttta	3120
ttttagaaaa tataaataaa tcttgatac atagatttcc aactagtaaa atttttagct	3180
atcgctgtag catttctatg aggagtttta ctttttactg tctcttatga aacttagtgg	3240
aaacaaagct aggtttcata aaacgtctaa tctcatattc atctctctcc aaaatatggg	3300
gaccttcaaa acaagaacag ttgcttatg taaacatttc ttacattgta ttctattaca	3360
ttatcatatt catgaagata ctttagaatt ttgttgctgt tagttctgtt gtgactggct	3420
gcttttctca gcttaagagc tctgcagaag tgatgataat atatattctc tttgctgagt	3480
tacaggaatt tcctgctttc ctctactgta cgttggagcc caggagatca gatccaacta	3540
ccagaatttt atatagctct ttgtgatcag tccttcctta gttgctgtta ttcctattga	3600
tgcaaaccta tttatgggtg attaaatgaa gaagatataa tgaggacagt ctaataatca	3660
aagttttttt ctgcagccag tagtacaggt agggtgacat gccttgcata aacagatcca	3720
gatgaaaggt gtgggtgagg caataagtct tcaggccctg ggctaaggca ggacctgttt	3780
tgtttcatga gatatgccta ggcctctcta tagacagggg gagatgtcag agaattctga	3840
agcatcagca gaggctatct tggttttcag gccctgtgct aagtatgaga ggaagttgag	3900
aacagataat gtgtgcaaca aagttagcaa catgcttttt caaacagag gaaacgatcc	3960
ataaatactg gaacagaaca acctcacaga aatcactgac tgtccctttg ctggccatgg	4020

catcctgaag gttgtgctgg tagcactcat tgggttgctg tagctgataa ggcagtgc	4080
ggacattggg tctgcacctg ctatgaacag gcactggggc ctccaggagc tgctggtatt	4140
agtcattggg atcagaggtg aagaaggcac ttgtagttct cacagctctg gcactgaccg	4200
atatttctgc tgtagttaat gggattattt gcatgtagtc aacagtttac aggctgattt	4260
ggctctgctc cgcaactagc aggacagagg gaccacaga tccatgtcct ggaaggggat	4320
aagggaag gatagagaga tgggcctaaa aaaaaacaga agcaacgatc tgaggataaa	4380
caaagtaatt tactaaataa gatagtggaa tgcatgataa cacactatag tacaacataa	4440
tgcaatgtaa ttagaactgg agctaataaa tcaaataaa agagagtgtc tgaaagccaa	4500
aggccttact ctaatgctgt gatgagatgg caccacaaa agagacgagc cagatgattg	4560
aaagggatgt atggtggtct accaaggcct gttatctgtt tgccctgagc agaaatgata	4620
gcagaacagc aaacagtcct ctggggaatg tagtagttct tctcttctgg gacaggtagc	4680
tggacctaga gcattaactc ttttaacccc agtacactac atgatgttat gatgtggaat	4740
accaataacc aaacatcata aaacatgac atagtcttat gtcgctgtgt tgtaactgga	4800
agctgaggat ctccactcag tatttctgct gaatgaagca gatctttcag ttgaaggaga	4860
ccaacttaat catccacttc caaccctta ctatggacag gttcaccgtt cactagatca	4920
agctgcccag ggccccatcc aacctagcct tgaatgcctc cagggatgag gcatccacca	4980

<210> 18

<211> 4980

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> ENSGALG0000015617_ENSGALG0000027714|chrZ:63363656-63368656

<400> 18

gaaaatctct gccactgatt gttttgcagc ttagttgtct agttattcgt gagatgttga	60
ttcaaacctt tgctcagatc aggcacagaa gcagtgtga actttgtttt tccgtagcct	120
tgctagtgtt gtaactgctg tcttagtgaa aacctgctta tttgcatccc cttcccaact	180
ttgggaaagt tgctgtctgt gattccagaa ttaggactcc ctggttgca acatgactca	240
gattttcaga ccttggtctc agaggcacag tggcttttgg tgagagaaa aatgaggctg	300
agaagaccag ggcatgatg agtgaggaag atgcaacaga acaagataga gaagaggga	360
gagtttttta ggagctccac cagacatact gtagaagcag tgtagatgca cagtaccagt	420
aatgtactcc agccaaccac tcggtgaaga cccatgaaaa cctaagagg tttagtaagg	480
ccaagtgcag ggtgctgtac ttgggtcggg gcaatcccag gtatttatac aaactagggg	540

aagatctcct tgagagcagc cctgcagaga aggacttggg ggtcctgggtg gatgagaagc 600
tggacgtgag ccagcagtgt gcacctgcag cccggaaggc caactgtgtt ctgggctgca 660
ttaaagaagg ggtgaccagc agggagaggg aggtgattgt cccctctac tcagctcttg 720
tgaggcccca tctggagtac tgcattccagg cctggagccc ccagtacagg aaggacgtgg 780
agctcttgga gcggttccag aggaagacca ctaagatgat cagagggtg gagcacctct 840
cctatgagga aaggttgagg gaactgtcct tgttttagctt agagaagaga aggcttcagg 900
gaaacctcag tgcagctttc cagtacttgg aaggagcaca taaacaggaa ggggaatggc 960
tgtttacgag ggtggatagt gacaggataa gggggaatgg ttttaaactg agacgggaga 1020
gttttaggtt agatattagg aggaagtttt tcacacagag ggtggtgacg cactggaaca 1080
ggttgcccaa ggaggttgtg gatgccccat ccctggaggc attcaaggcc aggctggatg 1140
tggctctggg cagcctgctc tgggtggctgg caaccctcca catagcaggg gggttgaaac 1200
tgcgtgatca ttgtagtctt tttcaacca ggccattcta tgattccatg aagactctgt 1260
aaggaccca atgcatcag ctgtccccac agcttcagta cccactgcca ccacattttt 1320
actgagatag gtagtttggg agccttcctc ccatcttgga taggggaggg ttaactgctg 1380
gcatttttgt tctttcctgc aatatgtgtt ttctgcatgc attctctttt caccctaat 1440
tttaatgtgg acggactttg aggatattct gcatctgccc aatatttctt acagcctcac 1500
atctactgat ttatgcacag tatcttatat ataatgtata ttatattatg tatattatgc 1560
actgaatcac atctactgtg tatgcacagt aacttctcct atggtactgt aagcccagaa 1620
atcccagata tgtcactaca acgtgtctgc tatgtattgc ttctgtgaga cacagatgtg 1680
ataatcagag gctgtacaca gtagaaatgc atacatatcc tactgtgaaa tctctgaaat 1740
ctagccttaa tcttgaaca gaaatgaata tgggtgacatc ttgatctgat agaattgggt 1800
gccagtagca cagctgtaat catccatgat atgaatcaaa ccaagcacag gtaaacaggt 1860
gagagaaatc atgaacaatt acatgcaaac ggaccagcta aaatgtgttt gtttgtgtt 1920
tttttttttt caggtgattc ttgattacag taagatcaga agctgctaca ttagcagacc 1980
agccactgca ctcaaggctg tgattcacag cttgcagacc tgacagcagt tctgtggaag 2040
aggcaggctc ctgtcaacca gttaatacaa taaatcagtc tcgtgtacac aaataatgtt 2100
atcctgcacc actgctgggtg ttacactatt tcacccaagt ttatcaccag caaactgagt 2160
cttatcgttc ctactgtgct ttgcttttct tgcttagcaa aacatatcag aaaggttttc 2220
agttaaaaaa aaaaaaaaaa aaaaaggatt atactgcagc catcaaagag ttctgttaaa 2280
gaagttgtgg ttcatattct tttcaagttc tgaagacaat ctgacaatag aggagatgag 2340

ggtttagtgg aggtaaatgg agctgtttct gaaagtaggg caccgtcgtg aggtaatctc 2400
 accactagat gttgctgtaa gcatgatgag gggagaagct ggagcacaaa aggtgtaaac 2460
 ccaatggttg atcatgcgca ctccggacaga ggaaatcttc aaaatgggaa agttagcttt 2520
 gaactgattt ggtgggttcc cattaataat gggggactta ctgccttatt gtctttatgt 2580
 acccatttgg gggatagatg ccatacagaa acaccagagc ttacgggtgtg aattcattgc 2640
 acatcttatt agatcaacaa aatggaataa aagaatacag aagattacta ctccattggg 2700
 catgtggacg ttttacaggc ctggataaat tagatcttaa aaacaaacaa acaaacaaac 2760
 aaaaaaactc cttgcaaaga taatgttatg taatattagt tgcaagaagt aagcaaacac 2820
 acagaactgg gagcagaagc aaagcactaa gttattaaag caagttgcac attttgagtt 2880
 gcattttgcc actggtttta taaacatgtt tagcatgtct ggtcagaatt tgggcaccag 2940
 gatgctttta agatgtctgt ctatggaacc tgtcagtgct caagaataac ttctgttatt 3000
 tggatgctgc accaaagaat tcagaggaag acgagccaag ccagacgtta tcatagtcac 3060
 tagtaaagtg gttctaagcc taattaagac atgtcagaac tatgtgttgt gcaaccaaact 3120
 cctccaaaag agaaatcaga ggtgaacttg tgcaataaat atagaagaca cgtaaactct 3180
 gaggcagtta gctaaccata tgaagccaat catacctgac tgctggacgc aggagactga 3240
 accttacaga ccctggagaa tcaactgtttg gcctagttag gcctgaatgg aatttaccaa 3300
 gattcatgac tttaacatgg atcaggtgca aagaaaagaa ggctcagtta gttcctacag 3360
 gctaccagat ctttttccac ctctgctcag cctggagctg tgggctctac ctgcccctga 3420
 agtgaagtgg catcaactgc aacacttttt gcagaggcaa aaatcatcaa gtcgtgcctc 3480
 tgcattttta ggtgatggat tccaagagat agattccaaa cccaacactg agattcctct 3540
 ggtgatgcag atgactgctt actgggatcc ttctctctct atagcctaact cccatgcagc 3600
 accaaaagt gggcacctga atactatcta tcttctcca tctgacctct gctgctggct 3660
 catcacgcat ggctcctgag cctgacagag ccatggggaa agacatggcc tgaccttttc 3720
 agacagcacg acttctcca gtgtgctgca ctaacgacaa tattaataat aacactgcag 3780
 gaaaacatgc tgttttctag caggctggag gaaaaaaaag gtgttcaagg tgagactcta 3840
 taagatcact ctttccaaac acagcagtta actcttttct cacttgaagc tgaaattcat 3900
 cttcagtagg tgacggttac aagcagctct ccaccggtaa gtgaaagttt gctgggtttt 3960
 tttcaggcac ttcaagacac gaaggtaatt gctgcaacac cttcctagct cttcctgagt 4020
 actcagccca aagctacact cttcagggt actctcctag cacggctctc tgtaattagc 4080
 ctggctttct aacttatgaa ttaacgctgc agctactcaa cagggcagct ttgcagctat 4140

gaactttatg catacacttt ttattctggg tggtacaaat tcattgctaa gaaatcttat 4200
 aaactctatg gaacactttt aaaagcctac atgtgattaa aaacaaaaaa acaaaaaaca 4260
 ctgatatcac atgctatatt ctctagaaaa gagtattttg gaaacaattg taattcacag 4320
 ctcaggagaa agagcaattg ttttcctttg tgctggagct ggaattctat gcagtataaa 4380
 tacagatgtc caaacaaact cattaaatgg tagtcatagc aaatggagga aaaaaaaagg 4440
 gggggggggg gggaactgaa gacactgaag aaattctgct gaaattcatg gagctgcaac 4500
 aagaattgat attgattact tcaccatatt ccgctctaata taactagcac aacactgact 4560
 gttgtgaaaa caagcatgtg atttgacca ataagtgtt cttttagtaa ccaaaaccaa 4620
 acaacaaca acccttctggc tgaccagatg gaggataaat tctaacctga tcattctgac 4680
 acaaatgagg aattatgtag gttggaactg cattccaaag caaagcatgt ttttggctta 4740
 gttgaacaca cagaattcat ttttcttgct gattcttgta actaaggga ggtatgaatt 4800
 ggttccctta aaagaagtcc ccagaagcca tggctctcaa caagtcttct cagattagca 4860
 aggatgaacg actgagaaat caaaacccaa attcagatct aagagccttc aacagatttg 4920
 ctaaataaaa ctagcaaata ctgtagaagt taacttgatt ttgaaataaa acaggatttc 4980

<210> 19

<211> 4980

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> ENSGALG00000008188_ENSGALG00000008197|chrZ:78777477-78782477

<400> 19

aaaggagttt cttgcctgaa accccttaag attttgtgta ttgttggtgc cctggtttca 60
 gctgggacag gactgatatc cttcacagta tctggatga tgctgtgtt tggttttagg 120
 agaaaaccag tgctgatggc acaccgaggt ttaggttgcg actgagcagt gctgcacaga 180
 gccaaaggatg tttcagctgc ccacactggc cttcctctaa gggggtgtga gggcacatgg 240
 agctggaggg aaccaaaggg gtattccaca ccatatgtca ctgagcaaag acattgctgc 300
 ccaggcagag aatacagtgg tagaaatacg tcacataggt gccatctac cccagaactg 360
 ggccagcgaa gaacatcaga acaacaagca tgcgaatcag gctgctcaga ttgatgtggc 420
 tcaggtagat ctcaactgtc aactgtcacc cataaaatgg atgagacagg agcaaggcct 480
 gcctgtccaa actgaccttg gggccccggg cacagctgct cccgagcaag gagaggcaga 540
 gaattgtgga ggaaacctc ttccacctag ctccaagggt gcacgctcaa tgcttgatg 600
 aagattcaag ttctcaatga agaactgcat tcaaagacat ctttgaggt gaccagact 660

gcctctggga tgaatccaga acaaacagaa tcttcccagc aacacctttg tttatcatac 720
 acgttcagaa atgcagtgcc tgcctggtat ttttttaaac tcatcacaga ggaatctctg 780
 agtggagcag aggagtgttt cctcttgctt ttttcttccc cttttctgta agaaatgcat 840
 atgccagttt ccccctaaat gttttcaaac tccaaattgt ttcctgctgg tgacattctc 900
 ctctccagca ctgcctacac cccagctgcg tctgatagga aaagcaccca gtaccatgct 960
 ggcatggcat ggtcagtgc acccactggg tagcatgggg acaaactgcc agagcagagg 1020
 gctgctgaaa cgccgagtgc caactgacac tgttcagctg acagcctcac gagagtgctg 1080
 cgagggttaa gtgaggcaac aaaatacaag tactcaaaaa tagaatggaa tggaatggaa 1140
 taacagaggg aatggggaac aggctgcagt gagaaggaaa ccccgctgca ccgaccaacc 1200
 cgctcccagt agcctgggtcc ctaccttgcc ccactccagc agctccacga ggtactcgaa 1260
 gtgcccattg ccggccatcc ccagcatggc acagctgtca cctgggagca gcgggggtcc 1320
 ggcaggcatc gccggcaccc acagcctggg ctgggctgag gatgggcttg gtgggcagct 1380
 ggggtctgcc gtggcagctg agtgccgtcg tgcagggtggg ctcgggtgctc tctgcagccc 1440
 acggctcggg acccgtggtg ggcatactgc tgggcaaggg gacggctcgg atgtgcatcc 1500
 cagagcaggg tgcattgatgt gtgcgcagcg gcacacagca gcgcacttgt ccccgaggca 1560
 ccacccggtc tgatgaaatt ccaggttttg ttttaagatga aaatgcatgt gggcatacat 1620
 ttgcacatga gaacacctgc taatgaaaca cacagagttg tatgtgcagt ctccctggct 1680
 ggtggcatgt ctgtgggggc agttaggcag gaagaaaagc tccaccatct gtttctgtgt 1740
 cccgggcttg gtgctgccac tagagatggg atgcacacag catgcctgcg tgctgccgag 1800
 gcagctcagt acctgctgct gggggagagc agccctcagg gagctgctgc aaagaccccg 1860
 tgggatggct ggtgaacagt gctggagaca tggcagatgt cgaccccctg ctccacacca 1920
 ccctccgctg cagctgcctg cacgatgcca tttcagcgtg ccacggcttg tttcttctc 1980
 ttgcagtctc aggttgccgc agcagcagct gcctgaggtc tggtaggcaaa gggcagagca 2040
 cccagccac tgctctccat gagtgaaggc gtagggggca cgccgtgctt cgcttctggt 2100
 atcgggtggc ctgtggagca catcacccct ggacagtga aacatgtcaa aggggtgttt 2160
 ataacagtat agcatatccc tttgaggctg aatttcctga gcttatggat atggtagtaa 2220
 tgctatacca ggctctctct gcaggttgct cactgaaaca atataccctt tctttctcaa 2280
 gaaatgggac tcatgaatag ctaccaggcc tttctgctac tagatactgt agaattacat 2340
 atatcagatg gctcattaat cacatagctt gttattggga cagaggcagg ttttcagcat 2400
 ttcccacact gcttttctgc aaatggatta ggagtaaaga gtcagatata caggggacca 2460

tcactcactt cctaggccat gcttcgtagt ttaactccat gcaaggtatt ttcttgctct	2520
gatctaactc tagatcacta aatggcactt gtgcaggaca cttttccact gtcatttttg	2580
ggggtagaag tgtggtcaga gccagggagg cttgcagggc ccacaccagc tgtggcccaa	2640
gcagctgctg tacaatggct gtatggcagt atcgatgtga acttcctgat aataaacaga	2700
actcagcagc aaataaaccc agattgttct gatcagtaac gagaggctgt gcaaaaagct	2760
gagaaatgta cagcccttcc tgctcaccac tacagccctt catgcagcgg cctgctgcac	2820
tagggcctgc ttggccagag gcagggctgc agcttgagat acgcagcacc cagtaccca	2880
gcaggtgcc atgtgcttct gtcaggccat gaattgcaca agttggata gttttataca	2940
gtgtcatgca gaagcacaac tcttaccttg gcacactgat atggaaagca aagtgtaaag	3000
gaactgtatg tttcttggct tgtgagttcc attggttctc ccaggacagc agtgacagct	3060
ctagattttc tggttgatct cttccatcta agtttttact tctgaatata tctgtgcagc	3120
aaagcctgat gctttttttt ttccccctcc cctgggatcg tgaaatatta ttttccatc	3180
tccataatac agatgatcca ctcaaacatg catcctcagt cgcagtcctc agaagggcct	3240
tttcccacca tacacctaca cgctacacca gactttattg gtcaccatgc cttgagggat	3300
tctgcctaac acaggcacac aggaagggt gataccagta attcattgct gaggagagaa	3360
ggggcacttc ttttatagta gagccacaga atcaccagga ttggaaaaca cctccaagat	3420
catccaatcc aaccaccac ctaccaccag tatttccac taaaccatgc cccttggtac	3480
cacatctaaa agcatcttct gcaccttcac agcttcattg cttttctcag gacgtgctcc	3540
agggccttgg tgtctttctt gtattgagtg gcccaaaatt gaacacaata ctcgagggtgc	3600
agcttcacca aagctgagtg cagaggggtg atcacctcta ctaggccttg ttttaccct	3660
tccccctgca tacctagttt aaagccctt cagcgagccc taccagctcc tgggctagga	3720
tccattttcc cctttaagag aggtgaaacc caccgcagg catcaggcca ggcactgagt	3780
taaactgctc catcagctaa ccacccaat gccctgatgt cccttctgat agtcctcagg	3840
cttctctcag cgacttcac actgccagcc tgagctatcc gtaatggata acaaccagag	3900
ggctgaacca gactaggag tctcctggta atgtccctaa gtaatgaatt catgggggaa	3960
atggatctgt gcagatgcac tgccatctca cagggtcaa ctgtcacaga cagttgtcac	4020
aaaccagagc agcagagaga gcagcaaagg agatatcata aatagtggaa aagggtgcca	4080
gtgtttgtta ttgactgctt cactgtgtgt ttggaaatca caaagaaaac acaaaaaaa	4140
ccccacagaa tacacatttc caagcatgca tgcttcaggt aaggactgcg gcccgttta	4200
attcccctgg aagcgtccag aggcttacac acaaatcagg ccacaacaaa caggttccgt	4260

gcaagcagtt tgtgagttgt atgaaaatag ctttcttcca taaagcgtgg cattgatggg 4320
aactgttact agcagatgaa aatctgccat gcacgacacc cttaatgggtt ggagtgggtc 4380
gactgagttt tgcattcctc attctgatgg atgctggcaa aaccaagatc acccaacacg 4440
taacactgaa tatagctcac caacagcaac gagccacacc agatactgag gcatctccaa 4500
ggctacgtct ggtggaaaga agagacagtc tcaccctgca gtgcaaaagc atcagggctt 4560
ctgcgtaagg ttatttgatc ccaaagatag gcaagttaat attcatcagt ttaatagaca 4620
ggtactcaca ttactggaga tggcatgcc a ggagagcctc tgccaagtgc tctactgccag 4680
tggcagccgg ccagtgaca gcttcttgggt gctgtggcca tctatataac agaaggaagc 4740
aattaagcat cacctatcaa gataagcaag atcattcaaa gagaataagg gtacttgggtt 4800
agagctgcag agcacttcct cacgaatgag tttcttccag ctacatgcag atgttacagg 4860
cagtttgctt gtgttctagg gataacacaa accctctctc atgggctagc tgcctgctca 4920
ggctgtgggtc tgtatTTTTTg cacgtgcact gctagccttt acttcacatt tccctttact 4980

<210> 20

<211> 681

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> pDsRed1

<400> 20

atggtgCGct cctccaagaa cgTcatcaag gagttcatgc gcttcaaggt gcGcatggag 60
ggcaccgtga acggccacga gttcgagatc gagggcgagg gcgagggccg cccctacgag 120
ggccacaaca ccgtgaagct gaaggTgacc aaggggcgGc ccctgcccTt cgCctgggac 180
atcctgtccc cccagttcca gtacggctcc aaggTgtacg tgaagcacc cGccgacatc 240
cccgactaca agaagctgtc cttccccgag ggcttcaagt gggagcgCgt gatgaacttc 300
gaggacggcg gcgtggTgac cgtgaccag gactcctccc tgcaggacgg ctgcttcatc 360
tacaaggtga agttcatcgg cgtgaacttc ccctccgacg gccccgtaat gcagaagaag 420
accatgggct gggaggcctc caccgagcgC ctgtaccccc gcgacggCgt gctgaaggGc 480
gagatccaca aggccctgaa gctgaaggac ggcgGccact acctggTgga gttcaagtcc 540
atctacatgg ccaagaagcc cgtgcagctg cccggctact actacgtgga ctccaagctg 600
gacatcacct cccacaacga ggactacacc atcgtggagc agtacgagcg caccgagggc 660
cgccaccacc tgttctgtgta g 681

<210> 21
 <211> 226
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> pDsRed1

<400> 21

Met Val Arg Ser Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys
 1 5 10 15

Val Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Glu Gly
 20 25 30

Glu Gly Glu Gly Arg Pro Tyr Glu Gly His Asn Thr Val Lys Leu Lys
 35 40 45

Val Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro
 50 55 60

Gln Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile
 65 70 75 80

Pro Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg
 85 90 95

Val Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser
 100 105 110

Ser Leu Gln Asp Gly Cys Phe Ile Tyr Lys Val Lys Phe Ile Gly Val
 115 120 125

Asn Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp
 130 135 140

Glu Ala Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly
 145 150 155 160

Glu Ile His Lys Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Leu Val
 165 170 175

Glu Phe Lys Ser Ile Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly
 180 185 190

Tyr Tyr Tyr Val Asp Ser Lys Leu Asp Ile Thr Ser His Asn Glu Asp
 195 200 205

Tyr Thr Ile Val Glu Gln Tyr Glu Arg Thr Glu Gly Arg His His Leu
 210 215 220

Phe Leu
 225

<210> 22
 <211> 876
 <212> ADN
 <213> Discosoma sp

<220>
 <221> đặc điểm khác
 <222> (1)..(876)
 <223> protein phát huỳnh quang màu đỏ (FP593)

GenBank: AF272711.1

<400> 22
 agtttcagcc agtgacaggg tgagctgcca ggtattctaa caagatgagt tgttccaaga 60
 atgtgatcaa ggagttcatg aggttcaagg ttcgtatgga aggaacggtc aatgggcacg 120
 agtttgaaat aaaaggcgaa ggtgaaggga ggccttacga aggtcactgt tccgtaaagc 180
 ttatggtaac caagggtgga cctttgccat ttgcttttga tattttgtca ccacaatttc 240
 agtatggaag caaggtatat gtcaaacacc ctgccgacat accagactat aaaaagctgt 300
 catttcctga gggatttaaa tgggaaaggg tcatgaactt tgaagacggg ggcgtggtta 360
 ctgtatccca agattccagt ttgaaagacg gctgtttcat ctacgaggtc aagttcattg 420
 ggggtgaactt tccttctgat ggacctgtta tgcagaggag gacacggggc tgggaagcca 480
 gctctgagcg tttgtatcct cgtgatgggg tgctgaaagg agacatccat atggctctga 540
 ggctggaagg aggcggccat tacctcgttg aattcaaaag tatttacatg gtaaagaagc 600
 cttcagtga gttgccaggc tactattatg ttgactccaa actggatatg acgagccaca 660
 acgaagatta cacagtcgtt gagcagtatg aaaaaacca gggacgccac catccgttca 720
 ttaagcctct gcagtgaact cggctcagtc atggattagc ggtaatggcc acaaaaggca 780
 cgatgatcgt tttttaggaa tgcagccaaa aattgaaggt tatgacagta gaaatacaag 840
 caacaggcct tgcttattaa acatgtaatt gaaaaac 876

<210> 23
 <211> 230
 <212> PRT
 <213> Discosoma sp

<220>

<221> đặc điểm khác

<222> (1)..(230)

<223> protein phát huỳnh quang màu đỏ D. GenBank: AAG16224.1

<400> 23

Met Ser Cys Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys Val
 1 5 10 15

Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Lys Gly Glu
 20 25 30

Gly Glu Gly Arg Pro Tyr Glu Gly His Cys Ser Val Lys Leu Met Val
 35 40 45

Thr Lys Gly Gly Pro Leu Pro Phe Ala Phe Asp Ile Leu Ser Pro Gln
 50 55 60

Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro
 65 70 75 80

Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val
 85 90 95

Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Ser Gln Asp Ser Ser
 100 105 110

Leu Lys Asp Gly Cys Phe Ile Tyr Glu Val Lys Phe Ile Gly Val Asn
 115 120 125

Phe Pro Ser Asp Gly Pro Val Met Gln Arg Arg Thr Arg Gly Trp Glu
 130 135 140

Ala Ser Ser Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Asp
 145 150 155 160

Ile His Met Ala Leu Arg Leu Glu Gly Gly Gly His Tyr Leu Val Glu
 165 170 175

Phe Lys Ser Ile Tyr Met Val Lys Lys Pro Ser Val Gln Leu Pro Gly
 180 185 190

Tyr Tyr Tyr Val Asp Ser Lys Leu Asp Met Thr Ser His Asn Glu Asp
 195 200 205

Tyr Thr Val Val Glu Gln Tyr Glu Lys Thr Gln Gly Arg His His Pro
 210 215 220

Phe Ile Lys Pro Leu Gln
 225 230

<210> 24
 <211> 4107
 <212> ADN
 <213> Streptococcus pyogenes

<220>
 <221> đặc điểm khác
 <222> (1)..(4107)
 <223> M1 GAS

Cas9 protein id: AAK33936.1

<400> 24
 atggataaga aatactcaat aggcttagat atcggcacaa atagcgtcgg atgggcgggtg 60
 atcactgatg aatataaggt tccgtctaaa aagttcaagg ttctgggaaa tacagaccgc 120
 cacagtatca aaaaaaatct tataggggct cttttatttg acagtggaga gacagcggaa 180
 gcgactcgtc tcaaacggac agctcgtaga aggtatacac gtcggaagaa tcgtatttgt 240
 tatctacagg agattttttc aaatgagatg gcgaaagtag atgatatgtt ctttcatcga 300
 cttgaagagt ctttttttgt ggaagaagac aagaagcatg aacgtcatcc tatttttgga 360
 aatatagtag atgaagttgc ttatcatgag aaatatccaa ctatctatca tctgcgaaaa 420
 aaattggtag attctactga taaagcggat ttgcgcttaa tctatttggc cttagcgcgt 480
 atgattaagt ttcgttgtca ttttttgatt gagggagatt taaatcctga taatagtgat 540
 gtggacaaac tatttatcca gttggtacaa acctacaatc aattatttga agaaaaccct 600
 attaacgcaa gtggagtaga tgctaaagcg attctttctg cacgattgag taaatcaaga 660
 cgattagaaa atctcattgc tcagctcccc ggtgagaaga aaaatggctt atttgggaat 720
 ctcatgtgctt tgtcattggg tttgaccctt aattttaaat caaatTTTga tttggcagaa 780
 gatgctaaat tacagctttc aaaagatact tacgatgatg atttagataa tttattggcg 840
 caaattggag atcaatatgc tgatttgttt ttggcagcta agaatttatc agatgctatt 900
 ttactttcag atatcctaag agtaaatact gaaataacta aggctcccct atcagcttca 960
 atgattaaac gctacgatga acatcatcaa gacttgactc ttttaaaagc tttagttcga 1020
 caacaacttc cagaaaagta taaagaaatc ttttttgatc aatcaaaaaa cggatatgca 1080
 ggttatattg atggggggagc tagccaagaa gaattttata aatttatcaa accaatttta 1140

gaaaaaatgg atggtactga ggaattattg gtgaaactaa atcgtgaaga tttgctgcgc	1200
aagcaacgga cctttgacaa cggctctatt ccccatcaaa ttcacttggg tgagctgcat	1260
gctattttga gaagacaaga agacttttat ccatttttaa aagacaatcg tgagaagatt	1320
gaaaaaatct tgacttttcg aattccttat tatgttggtc cattggcgcg tggcaatagt	1380
cgttttgcat ggatgactcg gaagtctgaa gaaacaatta ccccatggaa ttttgaagaa	1440
gttgtcgata aaggtgcttc agctcaatca tttattgaac gcatgacaaa ctttgataaa	1500
aatcttccaa atgaaaaagt actacaaaaa catagtttgc tttatgagta ttttacggtt	1560
tataacgaat tgacaaaggt caaatatgtt actgaaggaa tgcgaaaacc agcatttctt	1620
tcaggatgaac agaagaaagc cattgttgat ttactcttca aaacaaatcg aaaagtaacc	1680
gttaagcaat taaaagaaga ttatttcaaa aaaatagaat gttttgatag tgttgaaatt	1740
tcaggagttg aagatagatt taatgcttca ttaggtacct accatgattt gctaaaaatt	1800
attaaagata aagatttttt ggataatgaa gaaaatgaag atatcttaga ggatattgtt	1860
ttaacattga ccttatttga agatagggag atgattgagg aaagacttaa aacatatgct	1920
cacctctttg atgataaggt gatgaaacag cttaaacgtc gccgttatac tggttgggga	1980
cgtttgtctc gaaaattgat taatggtatt agggataagc aatctggcaa aacaatatta	2040
gattttttga aatcagatgg ttttgccaat cgcaatttta tgcagctgat ccatgatgat	2100
agtttgacat ttaaagaaga cattcaaaaa gcacaagtgt ctggacaagg cgatagttta	2160
catgaacata ttgcaaattt agctggtagc cctgctatta aaaaaggtat tttacagact	2220
gtaaaagttg ttgatgaatt ggtcaaagta atggggcggc ataagccaga aaatatcggt	2280
attgaaatgg cacgtgaaaa tcagacaact caaaagggcc agaaaaattc gcgagagcgt	2340
atgaaacgaa tcgaagaagg tatcaaagaa ttaggaagtc agattcttaa agagcatcct	2400
gttgaaaata ctcaattgca aaatgaaaag ctctatctct attatctcca aaatggaaga	2460
gacatgtatg tggaccaaga attagatatt aatcgtttaa gtgattatga tgtcgatcac	2520
attgttccac aaagtttcct taaagacgat tcaatagaca ataaggtcct aacgcgttct	2580
gataaaaatc gtggtaaatc ggataacgtt ccaagtgaag aagtagtcaa aaagatgaaa	2640
aactattgga gacaacttct aaacgccaag ttaatcactc aacgtaagtt tgataattta	2700
acgaaagctg aacgtggagg tttgagtga cttgataaag ctggttttat caaacgcaa	2760
ttggttgaaa ctcgccaaat cactaagcat gtggcacaaa ttttgatag tcgcatgaat	2820
actaaatagc atgaaaatga taaacttatt cgagaggtta aagtgattac cttaaaatct	2880
aaattagttt ctgacttccg aaaagatttc caattctata aagtacgtga gattaacaat	2940

```

taccatcatg cccatgatgc gtatctaaat gccgtcgttg gaactgcttt gattaagaaa 3000
tatccaaaac ttgaatcgga gtttgtctat ggtgattata aagtttatga tgttcgtaaa 3060
atgattgcta agtctgagca agaaataggc aaagcaaccg caaaatattt cttttactct 3120
aatatcatga acttcttcaa aacagaaatt acacttgcaa atggagagat tcgcaaacgc 3180
cctctaatacg aaactaatgg ggaaactgga gaaattgtct gggataaagg gcgagatttt 3240
gccacagtgc gcaaagtatt gtccatgccc caagtcaata ttgtcaagaa aacagaagta 3300
cagacaggcg gattctccaa ggagtcaatt ttacaaaaa gaaattcgga caagcttatt 3360
gctcgtaaaa aagactggga tccaaaaaaa tatggtggtt ttgatagtcc aacggtagct 3420
tattcagtcc tagtggttgc taaggtggaa aaagggaat cgaagaagtt aaaatccgtt 3480
aaagagttac tagggatcac aattatggaa agaagttcct ttgaaaaaaa tccgattgac 3540
tttttagaag ctaaaggata taaggaagtt aaaaaagact taatcattaa actacctaaa 3600
tatagtcttt ttgagttaga aaacggtcgt aaacggatgc tggctagtgc cggagaatta 3660
caaaaaggaa atgagctggc tctgccaagc aaatatgtga attttttata tttagctagt 3720
cattatgaaa agttgaaggg tagtccagaa gataacgaac aaaaacaatt gtttgtggag 3780
cagcataagc attattttaga tgagattatt gagcaaatca gtgaattttc taagcgtggt 3840
attttagcag atgccaattt agataaagtt cttagtgcac ataacaaca tagagacaaa 3900
ccaatacgtg aacaagcaga aaatattatt catttattta cgttgacgaa tcttgagct 3960
cccgctgctt ttaaataatt tgatacaaca attgatcgta aacgatatac gtctacaaaa 4020
gaagtttttag atgccactct tatccatcaa tccatcactg gtctttatga aacacgcatt 4080
gatttgagtc agctaggagg tgactga 4107

```

<210> 25
 <211> 1368
 <212> PRT
 <213> Streptococcus pyogenes

<220>
 <221> đặc điểm khác
 <222> (1)..(1368)
 <223> M1 GAS

Cas9 protein id: AAK33936.1

<400> 25

Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
 1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
 20 25 30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
 35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
 50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
 65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
 85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
 100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
 115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
 130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
 145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
 165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
 180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
 195 200 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
 210 215 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
 225 230 235 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
 245 250 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
 260 265 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
 275 280 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
 290 295 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
 305 310 315 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Lys
 325 330 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
 340 345 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
 355 360 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
 370 375 380

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
 385 390 395 400

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
 405 410 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
 420 425 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
 435 440 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
 450 455 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
 465 470 475 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
 485 490 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
 500 505 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
 515 520 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
 530 535 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
 545 550 555 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
 565 570 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
 580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
 595 600 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
 610 615 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
 625 630 635 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
 645 650 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
 660 665 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
 675 680 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
 690 695 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
 705 710 715 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
 725 730 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
740 745 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
755 760 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
770 775 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785 790 795 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
805 810 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
820 825 830

Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
835 840 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
850 855 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865 870 875 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
885 890 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
900 905 910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
915 920 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
930 935 940

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945 950 955 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
965 970 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
 980 985 990

Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
 995 1000 1005

Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala
 1010 1015 1020

Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe
 1025 1030 1035

Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala
 1040 1045 1050

Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu
 1055 1060 1065

Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val
 1070 1075 1080

Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr
 1085 1090 1095

Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys
 1100 1105 1110

Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro
 1115 1120 1125

Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val
 1130 1135 1140

Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys
 1145 1150 1155

Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser
 1160 1165 1170

Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys
 1175 1180 1185

Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu
 1190 1195 1200

Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly
 1205 1210 1215
 Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val
 1220 1225 1230
 Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser
 1235 1240 1245
 Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys
 1250 1255 1260
 His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys
 1265 1270 1275
 Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala
 1280 1285 1290
 Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn
 1295 1300 1305
 Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala
 1310 1315 1320
 Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser
 1325 1330 1335
 Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr
 1340 1345 1350
 Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp
 1355 1360 1365

<210> 26
 <211> 20
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> gARN7 đối với locus nhiễm sắc thể Z chrZ_42174515_-1

<400> 26
 gtaatacaga gctaaaccag

20

<210> 27
 <211> 20
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> gARN8 đối với locus nhiễm sắc thể Z chrZ_9157091_1

 <400> 27
 acagacctat gatatgtgag 20

 <210> 28
 <211> 20
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> gARN9 đối với locus nhiễm sắc thể Z chrZ_27767602_-1

 <400> 28
 gagcttgtga gtgataatcg 20

 <210> 29
 <211> 20
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> gARN10 đối với locus nhiễm sắc thể Z chrZ_78779927_1

 <400> 29
 gtaaagagtc agatacacag 20

 <210> 30
 <211> 20
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> gARN11 đối với locus nhiễm sắc thể Z chrZ_63364946_-1

 <400> 30
 cagtgggtac tgaagctgtg 20

 <210> 31
 <211> 1012
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Nhánh bên trái đối với sự tích hợp gARN7 vào nhiễm sắc thể Z

 ZNF367_HABP4_ENSGALG00000012629_ENSGALG00000012628_chrZ_42174515_
 -1

 <400> 31
 gtgaacggat aggacagcaag cattcagaac aatgaaaaag atctgcgttc tattgtaaaag 60

 caaacaagg tgatgacgga acgtttatgt gagatacagg caccatagta caagatatca 120

aagaggaatt aattcaaccc tcccaacaac agttgggaaa aaaaaaggat aagctgagga 180
atgcagcctt ctccgatttt tactctggtg ccggaaggag gaggaggagc gggggaaaaac 240
ccagcggact cagactgcat acgcactctc ggaagtacct ctgaggcggg gtgagaggaa 300
gggtggatct aggtccccc cgctccacac actcacgtgc tcgtgtactg gttcaggctg 360
tgacttggca actcctctac atctatccat ctagttagac tcaacaagac catcgtgatt 420
caggagtctg aagatgctta actctctcag aacgtctagt tctcctggcc tggtaaattgc 480
tatgtttctc atactgcctc tctgaagaat gcttcgaatg cttctagtga tgctctaaag 540
ttctaaacag aaaaatctgg agacagtctt ggtctttaga tagaaaaaat gccaacatgc 600
caaaggatgg ttacatcctt caagcaacct tgttgcacgc tgtacaatag actcatgtaa 660
taacttagcc gtagtcacgc tatctcttat ttacctgttc gttattacat tttcctggta 720
ctgctttata ttttagtcagt tgtcctttta agacaaatth tttgggtgtg gctaataggc 780
agttaccaaa tgttctagag ggagggaata taatcagtga gtatgtagat gtatatagat 840
gtataaccag taagtataca gagaagttag ggtgggtctt tgagagcata tagctggaaa 900
gattatcaca tgggaaaagg taatcagaaa taaatggaaa aatgctcacc agtgtcatag 960
gctcacaaga aaactcccca aggagcaatc tccagtaaca gacccttcct ca 1012

<210> 32

<211> 1023

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Nhánh bên phải đối với sự tích hợp gARN7 vào nhiễm sắc thể Z

ZNF367_HABP4_ENSGALG0000012629_ENSGALG0000012628_chrZ_42174515_
-1

<400> 32

gtcatgtcgg tggaggagaa agtggctgta gaaaaggtaa cggacaagaa atccctcact 60
ggcttagctg taaaggctcag aggaaaatca tcttctgttc tccatacaca tctcctgtaa 120
ttaggcacct gtgcccttta aaaaagtaaa gtaatcaaag agtcatttgg tgaaaataga 180
agccaaacag ttacaaaatt tagttaatat tcaaagacaa aacaactctg gttccaaata 240
taaaaccctt tgtgtaaaca ctgcgggaag gtgccaaaga gccacctaca caaagagatg 300
ccacagagta atttgcgtac cccaacaata ggatcatata tgggcaaacc aaaagaccaa 360
caaagacccc ttgcaataat cccttgattg gcagaagcaa gtgcagcgtt agatttagta 420
tgtgtgacaa ttctgagaat aggaaggaat ttactgaaa ctgtgtgaag atttgtatgt 480
taaatacata taatgagggt tatttagctc tgggggtgtgc atgctatgtg gagagatccc 540

catgtgccca gcactgcaat agactaatgt cagcttttta aaccatcatt tggtttgaag 600
 agtttattat gattttcaga aacagaatta tgtattcagt gtctgtacat tcttacaagc 660
 ctgctgttct gtacacatga aaaagctact tatgggtgcaa atcagtatag agaagcagtt 720
 ttgacatata gtggatcatga ctgggtgatc tgctgtcaga aggtaagaca ctggtataca 780
 gaattgcaaa cgtgagatga agacctcaag tgcaactctg tcatcataacc acttctcttg 840
 ctattatcta ttcagtggtt gttctaaata ttcaggcaga aggcttggtt tcacatcagg 900
 tgttacattt aagtattctt tgcccatitt ttgctgtttg tatgtgtaac ttcagattct 960
 cacattgact gtgtagtgta aatttgcaaa tagatgtaac agcttctctt taacatccca 1020
 tgc 1023

<210> 33

<211> 1014

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Nhánh bên trái đối với sự tích hợp gARN8 vào nhiễm sắc thể Z

ENSGALG0000025945_AVD_ENSGALG0000025945_ENSGALG0000023622_chrZ
 _9157091_1

<400> 33

ctctgcatgg tcacaacggc cccaacagcc tcccatcccg atgtcaccaa tcagttcact 60
 attactggtg agcaaccagt ccagcactgc atccccctgg tggggctgtc tgttacgtgg 120
 ctcaggaagt taccctcagt gcattccagg tgctcctgga tcgcctgtag ctaccgtgc 180
 tacttttcca gcagggtgtc ggggtggtaga agtccccag caggacgaga gcctgcgatt 240
 gtgacagctt ctgtagctgg agggagaatg cttcatcaac aggctccgct tgatcagtgt 300
 ccctgtactg ataccacaag gcttcttttg ctgctcccat cttactctc accccatcgt 360
 ggtgttatac gtgcataaca tcatgcagtg cactaagagt taaagagtta atgctccagt 420
 tccgtcgatt gtcaataatt ccagggtgtac ctttctcaga agagaagaac taaaaatccc 480
 ttaagatctt actgttccat tttcattctc cattcagaag gaaaggtaaa actgcagtgt 540
 gtgctcccta gcattagagt aatttctccc ttagctcggt gccaccattt tgtttttagg 600
 cccacctcct taccttttcc ctacaccct cctgtgtctc tgagacgtgg gtccatgggt 660
 ccctcaacca tagtagacac aaactcagat caaaccaagc cagagctttg acaaccctta 720
 agctttcagc ttactcaggg ttgctctgca gggactgctc ttcacactct attccttcct 780
 tgatatagag ggcaatgtct cctctcctcc ttccttacct gtccctcctc atctgcctgt 840

tgctgtcagt ggccacactt cagccatggg aatcatccca ccctgctttt gtgaaggaaa 900
 ctacattatg gttctcaggt agcacagtgg cttccatcac ctctgtttg cttcccaagc 960
 tgtgtgtgtt ggtgtagagg cacttcagct gggcagctga agaactggac gcac 1014

<210> 34
 <211> 1032
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Nhánh bên phải đối với sự tích hợp gARN8 vào nhiễm sắc thể Z

ENSGALG0000025945_AVD_ENSGALG0000025945_ENSGALG0000023622_chrZ
 _9157091_1

<400> 34
 cacctttcca cctcctgcta aagcacattc cctacagtat gtttcacatg acaccgtgct 60
 ggcagatttt taatatcatc agagaatgac actccatata ctctctgggc agcctttttc 120
 agtgctctgt caccacaaaa gtaaagtatt tgctcatgtt ctaatggaac ttcctatgtt 180
 ctagtttgtg accattgaac cttgttctgt cactggacag tgctgaagga gcctggcctt 240
 atccacttga ctccacact ttacatattt acaagcatcg ataagacccc cctcagtctt 300
 ctccagacta gatagcccca ggtctcccaa ccattccttt tatgggagat gcttcaggcc 360
 cctcatcacc tctgtggccc tctgctggac tgctgcagt agttccaggt tttttttttt 420
 gaactgggga gcccagaact tgacacggta ttccagatgt ggcctcacca gcgcagagga 480
 gaggggaagg ctacactccc tcgacctggg ggccatgttt tttttaacag acctcaggat 540
 accactggcc ttcttggccc aagggcacac tgctggctca tggcccatgg gttggcttcc 600
 aggactctca ggtccttctc cgtagagctc ctctcaagca ggacaaccct cagcctgtac 660
 tgatgtgtgc ggtgcggtta ttctcccca cgtgcaagac tctacatctg ccattgttaa 720
 atctcataag cttcctctct gcccaactct ccactgtgta cgggccttag tgaatggcag 780
 cacagacttc tgggtgtgagg atagcattca gaacaagctg ctgtaccttc ccagtcacag 840
 aggtgagggg gactggcttg tactttccca gctttgcctt cttgcccttt tggaagactg 900
 gagcgacatt ggccttcttg aggaaacggg cctcaggcac ccctcctgtt ctccatgacc 960
 tttgaaagat gatcgagagc accttggcag tcacttctgc cagctccttc agcacgcatg 1020
 agtgcatccc gc 1032

<210> 35
 <211> 1025
 <212> ADN
 <213> Trình tự nhân tạo

<220>

<223> Nhánh bên trái đối với sự tích hợp gARN9 vào nhiễm sắc thể Z

RIC1_ERMP1_ENSGALG00000015036_ENSGALG00000000438_chrZ_27767602_-1

<400> 35

```

gcagtacgta cgtttgtgaa cagctgtttt caagggtgat gcacaggaag agtaaaattt      60
cattaacaaaa tctccaagga gcacatggag agctcactca gaagcctcac cgctgccact      120
gaacaagact gatgcattag tttcatgaaa tcaaggctat gtgtcccact ggtttcatta      180
ttttgttact ctttctttta aacaaaaaaaa ttgttttatt gcttatatac atcaactatg      240
ttgttatata ttttcctaaag gccgctccaa aagtaatgcc tcctgtttca ttctgttggc      300
ccaccgtgtc agaggcagat ggtggtacgg cagtagagat ggaaccttcc caccaatatt      360
ccattacatg ttgcatgtg atagacagca gcagatgggc actctgacag aatgggtatct      420
gacatggaag cgtggatgga gcaaagttgt gtcactgaat tcctccataa ggaaatgagc      480
accactaac attcactgat gcttgctgaa ctttctgga gaccaaacag tggatgtgag      540
cacagtgagg tgggtgggtgg tgcatttcag cagtgggtgac aatgacgttg ggtcacctct      600
gctgggtgag gttttgatga gtgcagcctg tgaggtcttg ttcacatg gagaaaatgc      660
acaactaatg gtggtgactg ctgaaaaata gtgttttgta gctgagaatt ttctctgtca      720
aacagtgcta attgtgctct ttgtatctgt ttagtattcc atgggaataa ataggaggca      780
ttacttgag actgacctac atatatgtgg cccagggcaa gtcaacagct gtacgaaaag      840
tccatgcttc ctcattataa acggaggaaa aaaagttgtt tacagctgta atgggattat      900
aaaaggcaaa tggggatagt acagtgggtga gaaccagatc tatgaaggaa aagccctaag      960
aaaaagagag tgcacagata ctcttaacca tctaataact gtttctcca tcctacagct     1020
cagag                                           1025

```

<210> 36

<211> 1040

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Nhánh bên phải đối với sự tích hợp gARN9 vào nhiễm sắc thể Z

RIC1_ERMP1_ENSGALG00000015036_ENSGALG00000000438_chrZ_27767602_-1

<400> 36

```

tggcttccaa ttctcctgt ttatcaccca cgttgcatcc attgatgtag aggcacttca      60
gtcgggctat cagcaatgtt accacctgtg aggagccctc ttgagatctt ttaggacaat      120
ttagagggtt tccccactgt ttctaaaaat tacagcattc cctgtccctt cttagtata      180
tagaactcca tccccttcca ccatcaaacc tagcttaagc tctggtaaat gatccagcca      240

```

```

gtctgcttcc caaaactctt gctcagttgc attctatctg atgcctacat gccaacacc      300
ttcaaggcct gtccaagatc atagaacaca aaggcttgat catgacaaca tccatgcagc      360
caatcattca catgacccat tcctctcttc ctttccaggt cccaacggcc aactgagagt      420
tgtgctctga gctcttcagt gtccttccga gggatgtaaa gtccctttta atgttctgca      480
ttttctttgt cacagccttc tgagacccaa cttgaatgtg gaggagtgga gagtaatcat      540
cccgctctga ttatcagata cattagcctc ttcttggcac ctctggcaga cagaagactt      600
ccctggagag attacctgga tgataactgg gggcctcagt gcctcacagc agcaagtccc      660
cagttactaa gactctacac ctctttgggtg gcgctagttc tgatgtagtt cttcagttgc      720
ttaacatgct tgttgtttcc caagatttga ctgctatctg caccctccct ttcttccagc      780
cccagagcct tatatcatga gctgagaaca ttctgtgttc aaaatacaag acacatttat      840
gcaatttcgt attctctacc catatttttg tgatagaaat tagggctact tatttcccat      900
agtgatttcc atgctgccta ccaggttcta agtcaggtca ccctgccagc ttcttccaaa      960
ccatgctcag gagtgcacaa aatcacacgc agacccaagg aaaaaatgta ccaaacaata     1020
ttaccacgca tgcaaaactcc                                             1040

```

<210> 37

<211> 1024

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Nhánh bên trái đối với sự tích hợp gARN10 vào nhiễm sắc thể Z
 TRIM36_PGGT1B_ENSGALG00000008188_ENSGALG00000008197_chrZ_78779927
 _1Chromosome Z

<400> 37

```

ctaccttgcc ccaactccagc agctccacga ggtactcgaa gtgcccattgt ccggccatcc      60
ccagcatggc acagctgtca cctgggagca gcgggggtcc ggcaggcatc gccggcacc      120
acagcctggg ctgggctgag gatgggcttg gtgggcagct ggggtgctgcc gtggcagctg      180
agtgccgtcg tgcagggtggg ctcggtgctc tctgcagccc acggctcggg acccggtgtg      240
ggcatactgc tgggcaaggg gacggctcgg atgtgcatcc cagagcaggg tgcatgatgt      300
gtgcgagcgc gcacacagca gcgcacttgt ccccgaggca ccacccggtc tgatgaaatt      360
ccaggttttg tttaagatga aaatgcatgt gggcatacat ttgcacatga gaacacctgc      420
taatgaaaca cacagagttg tatgtgcagt ctccctggct ggtggcatgt ctgtgggggc      480
agttaggcag gaagaaaagc tccaccatct gtttctgtgt cccgggcttg gtgctgccac      540
tagagatggg atgcacacag catgcctgcg tgctgccgag gcagctcagt acctgctgct      600

```

gggggagagc agccctcagg gagctgctgc aaagaccccg tgggatggct ggtgaacagt 660
 gctggagaca tggcagatgt cgcaccctg ctccacacca ccctccgctg cagctgcctg 720
 cacgatgcca tttcagcgtg ccacggcttg tttcttcctc ttgcagtctc aggttgccgc 780
 agcagcagct gcctgaggtc tgggtggcaaa gggcagagca cccagcccac tgctctccat 840
 gagtgagcgg gtagggggca cgccgtgctt cgcttctggg atcgggtggc ctgtggagca 900
 catcaccctt ggacagtga aacatgtcaa aggggtgtttt ataacagtat agcatatccc 960
 tttgaggctg aatttcctga gcttatggat atggtagtaa tgctatacca ggctctctct 1020
 gcag 1024

<210> 38

<211> 1022

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Nhánh bên phải đối với sự tích hợp gARN10 vào nhiễm sắc thể Z
 TRIM36_PGGT1B_ENSGALG0000008188_ENSGALG0000008197_chrZ_78779927
 _1Chromosome Z

<400> 38

cagctgctgt acaatggctg tatggcagta tcgatgtgaa cttcctgata ataaacagaa 60
 ctcagcagca aataaaccca gattgttctg atcagtaacg agaggctgtg caaaaagctg 120
 agaaatgtac agcccttcct gctcaccact acagcccttc atgcagcggc ctgctgcact 180
 agggcctgct tggccagagg cagggtgca gcttgagata cgagcaccc agtaccctag 240
 caggctgcca tgtgcttctg tcaggccatg aattgcacaa gttggtatag ttttatacag 300
 tgtcatgcag aagcacaact cttaccttgg cactgata tggaaagcaa agtgtaaagg 360
 aactgtatgt ttcttggtt gtgagttcca ttggttctcc caggacagca gtgacagctc 420
 tagattttct gttgatctc ttccatctaa gtttttactt ctgaatacat ctgtgcagca 480
 aagcctgatg cttttttttt tccccctccc ctgggatcgt gaaatattat tttcccatct 540
 ccataataca gatgatccac tcaaacatgc atcctcagtc gcagtcctca gaagggcctt 600
 ttcccacat acacctacac gctacaccag actttatttg tcaccatgcc ttgagggatt 660
 ctgcctaaca caggcacaca ggaagggctg ataccagtaa ttcattgctg aggagagaag 720
 gggcacttct tttatagtag agccacagaa tcaccaggat tggaaaacac ctccaagatc 780
 atccaatcca accaccacc taccaccagt atttccact aaaccatgcc ctttggtacc 840
 acatctaaaa gcatcttctg caccttcaca gcttcattgc ctttctcagg acgtgctcca 900
 gggccttggg gtcttttctg tattgagtgg cccaaaattg aacacaatac tcgaggtgca 960

gcttcaccaa agctgagtgC agaggggtga tcacctctac taggccttgt tttaccctt 1020
cc 1022

<210> 39
<211> 1003
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Nhánh bên trái đối với sự tích hợp gARN11 vào nhiễm sắc thể Z
RPS23_ATG10_ENSGALG00000015617_ENSGALG00000027714_chrZ_63364946_-
1

<400> 39
gggttgaaac tgcgtgatca ttgtagtctt tttcaaccca ggccattcta tgattccatg 60
aagactctgt aaggacccca atgcatcag ctgtcccccac agcttcagta cccactgcca 120
ccacattttt actgagatag gtagtttggc agccttcctc ccatcttgga taggggaggg 180
ttaactgctg gcatttttgt tctttcctgc aatatgtgtt ttctgcatgc attctctttt 240
cacccaaat ttaaatgtgg acggactttg aggatattct gcatctgccc aatatttctt 300
acagcctcac atctactgat ttatgcacag tatcttatat ataatgtata ttatattatg 360
tatattatgc actgaatcac atctactgtg tatgcacagt aacttctcct atggactgtg 420
aagcccagaa atcccagata tgtcactaca acgtgtctgc tatgtattgc ttctgtgaga 480
cacagatgtg ataatcagag gctgtacaca gtagaaatgc atacatatcc tactgtgaaa 540
tctctgaaat ctagccttaa tcttggaaca gaaatgaata tggtgacatc ttgatctgat 600
agaattgggt gccagtagca cagctgtaat catccatgat atgaatcaaa ccaagcacag 660
gtaaacaggt gagagaaatc atgaacaatt acatgcaaac ggaccagcta aaatgtgttt 720
gtttgttgtt tttttttttt cagggtgattc ttgattacag taagatcaga agctgctaca 780
ttagcagacc agccactgca ctcaaggctg tgattcacag cttgcagacc tgacagcagt 840
tctgtggaag aggcaggctc ctgtcaacca gtttaaatcaa taaatcagtc tcgtgtacac 900
aaataatgtt atcctgcacc actgctggtg ttacactatt tcaccaagt ttatcaccag 960
caaactgagt cttatcggtc ctactgtgct ttgcttttct tgc 1003

<210> 40
<211> 1003
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Nhánh bên phải đối với sự tích hợp gARN11 vào nhiễm sắc thể Z
RPS23_ATG10_ENSGALG00000015617_ENSGALG00000027714_chrZ_63364946_-

1

<400> 40
 cagaaacacc agagcttacg gtgtgaattc attgcacatc ttattagatc aacaaaatgg 60
 aataaaagaa tacagaagat tactactcca ttgggcatgt ggacgtttta caggcctgga 120
 taaattagat cttaaaaaca aacaaacaaa caaacaaaaa aactccttgc aaagataatg 180
 ttatgtaata ttagttgcaa gaagtaagca aacacacaga actgggagca gaagcaaagc 240
 actaagttat taaagcaagt tgcacatctt gagttgcatt ttgccactgg ttttataaac 300
 atgttttagca tgtctgggtca gaatttgggc accaggatgc ttttaagatg tctgtctatg 360
 gaacctgtca gtgctcaaga ataacttctg ttatttggat gctgcaccaa agaattcaga 420
 ggaagacgag ccaagccaga cgttatcata gtcactagta aagtggttct aagcctaatt 480
 aagacatgtc agaactatgt gttgtgcaac caaatcctcc aaaagagaaa tcagaggtga 540
 acttgtgcaa taaatataga agacacgtaa atcctgaggc agttagctaa ccatatgaag 600
 ccaatcatac ctgactgctg gacgcaggag actgaacctt acagaccctg gagaatcact 660
 gtttggccta gttaggcctg aatggaatct accaagattc atgactttaa catggatcag 720
 gtgcaaagaa aagaaggctc agttagttcc tacaggctac cagatctttt tccacctctg 780
 ctcagcctgg agctgtgggc tctacctgcc cctgaagtga agtggcatca actgcaacac 840
 tttttgcaga ggcaaaaatc atcaagtcgt gcctctgcat ttttaaggatg tggattccaa 900
 gagatagatt ccaaacccaa cactgagatt cctctgggtga tgcatgatgac tgcttactgg 960
 gatccttctc tctctatagc ctaatcccat gcagcaccaa aag 1003

<210> 41

<211> 200

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Để tách dòng bằng cách sử dụng gARN3, nhiễm sắc thể Z, "nhánh bên trái"

<400> 41

cctgcaggac atgtggacac tctgccaggt actgggggag gatcccctcc actgctccct 60
 gggagatggc aaacatctag agagggacac tgcagggatg ctgcatgaac acctccagcc 120
 ttctgcaggc tgtgatgggg tcagggtga aggtcaacac aagcaatgac tgcttgctga 180
 caatgtgggc attgccgacc 200

<210> 42

<211> 258

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Để tách dòng bằng cách sử dụng gARN3, nhiễm sắc thể Z, "nhánh bên phải"

<400> 42

tggtgagtgg ggccaggaga gcaggacagg agtgtggact cggaggcgcg ggcagaggag	60
gtagctcaga gcatgagata tattcgccag gtgtcagtgg acttctggca gtgcatactg	120
cagagagctt tggttgctgg ccatctggat gatgacaaaa tctcagccac agaccatgtg	180
ggttgaaga gtcctcagga gatcaccttt ccacctcctg ctaaagcaca ttccctacag	240
tatgtttcac atgacacc	258

<210> 43

<211> 436

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Để tách dòng bằng cách sử dụng gARN4, nhiễm sắc thể Z, "nhánh bên trái"

<400> 43

tcaacagctg tacgaaaagt ccatgcttcc tcattataaa cggaggaaaa aaagttgttt	60
acagctgtaa tgggattata aaaggcaaat ggggatagta cagtgggtgag aaccagatct	120
atgaaggaaa agccctaaga aaaagagagt gcacagatac tcttaaccat ctaataactg	180
tttcctccat cctacagctc agagttaaga cttacagagg actctagtag ttagtaagat	240
gaatacgagc taatagtggc aaaaataatc ccagtgcctc aacactgacc tgggaaaaag	300
gggcatgtat agaccttctg atattgtgat gctgtgtttg tacacttatt atctacattt	360
tcagaaatta ggttaaactt cagaaaactg aagatctcca gggctttagtag cagaccctga	420
ccaccagact ggtccc	436

<210> 44

<211> 521

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Để tách dòng bằng cách sử dụng gARN4, nhiễm sắc thể Z, "nhánh bên phải"

<400> 44

tcaaaccgat tgtggctgtt tctcagaggc aactctttgc gctctagccc ctctataaca	60
tggggcaact tcccctgccc caccttcctt tctgtatct tctgaaaagc ttgtagccct	120
caattgtcac gctccagcca tatgaatcat cccaccacgt ccctgtgttc acaattacct	180
tctagtcttc taattgcacc atggcttcca attcctcctg tttatcacc acgttgcatc	240
cattgatgta gaggcacttc agtcgggcta tcagcaatgt taccacctgt gaggagccct	300

cttgagatct tttaggacaa tttagagggtt ttccccactg tttcctaaaa ttacagcatt 360
 ccctgtccct tcttagtgat atagaactcc atccccctcc accatcaaac ctagcttaag 420
 ctctggtaaa tgatccagcc agtctgcttc caaaaactct tgctcagttg cattctatct 480
 gatgcctaca tgcccaacac cttcaaggcc tgtccaagat c 521

<210> 45
 <211> 505
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Để tách dòng bằng cách sử dụng gARN5, nhiễm sắc thể Z, "nhánh bên trái"

<400> 45
 gcctgtccaa actgaccttg gggccccggg cacagctgct cccgagcaag gagaggcaga 60
 gaattgtgga ggaaaccctc ttccacctag ctccaagggt gcacgctcaa tgcttgatg 120
 aagattcaag ttctcaatga agaactgcat tcaaagacat ctttgagggt gaccagact 180
 gcctctggga tgaatccaga acaaacagaa tcttcccagc aacaccttg tttatcatatc 240
 acgttcagaa atgcagtgcc tgctgggtat ttttttaaac tcatcacaga ggaatctctg 300
 agtgagcag aggagtgttt cctcttgctt ttttcttccc ctttctgta agaaatgcat 360
 atgccagttt cccctaaat gttttcaaac tccaaattgt ttctgctgg tgacattctc 420
 ctctccagca ctgcctacac cccagctgcg tctgatagga aaagcaccca gtaccatgct 480
 ggcatggcat ggtcagtgc accca 505

<210> 46
 <211> 599
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Để tách dòng bằng cách sử dụng gARN5, nhiễm sắc thể Z, "nhánh bên phải"

<400> 46
 ctgccagagc agagggtgc tgaaacgcc agtgccaact gacactgttc agctgacagc 60
 ctcacgagag tgctgcgagg gttaagtgc gcaacaaaat acaagtactc aaaaatagaa 120
 tggaatggaa tggaataaca gagggaatgg ggaacaggct gcagtgagaa ggaaaccccc 180
 gtgcaccgac caaccgctc ccagtagcct ggtccctacc ttgccccact ccagcagctc 240
 cacgaggtac tcgaagtgcc catgtccggc catccccagc atggcacagc tgtcacctgg 300
 gagcagcggg ggtccggcag gcatcgccgg caccacagc ctgggctggg ctgaggatgg 360
 gcttggtggg cagctgggtg ctgccgtggc agctgagtgc cgtcgtgcag gtgggctcgg 420

tgctctctgc agccacggc tcgggacccg tgggtgggcat actgctgggc aaggggacgg 480
 ctcggtatgt catccagag caggggtgcat gatgtgtgcg cagcggcaca cagcagcgca 540
 cttgtccccg aggcaccacc cggctctgatg aaattccagg ttttgtttaa gatgaaaat 599

<210> 47
 <211> 453
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Để tách dòng bằng cách sử dụng gARN6, nhiễm sắc thể Z, "nhánh bên trái"

<400> 47
 ttgcagtctc aggttgccgc agcagcagct gcctgaggctc tgggtggcaaa gggcagagca 60
 cccagcccac tgctctccat gaggtagcgg gtagggggca cgccgtgctt cgcttctggt 120
 atcgggtggc ctgtggagca catcaccctt ggacagtga aacatgtcaa aggggtgtttt 180
 ataacagtat agcatatccc tttagaggctg aatttcctga gcttatggat atggtagtaa 240
 tgctatacca ggctctctct gcaggttgct cactgaaaca atataccctt tctttctcaa 300
 gaaatgggac tcatgaatag ctaccaggcc tttctgctac tagatactgt agaattacat 360
 atatcagatg gctcattaat cacatagctt gttattggga cagaggcagg ttttcagcat 420
 ttcccacact gcttttctgc aaatggatta gga 453

<210> 48
 <211> 439
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Để tách dòng bằng cách sử dụng gARN6, nhiễm sắc thể Z, "nhánh bên phải"

<400> 48
 gggaccatca ctcaattcct aggccatgct tcgtagttta actccatgca aggtattttc 60
 ttgctctgat ctaactctag atcactaaat ggcacttggtg caggacactt ttccactgtc 120
 attttggggg gtagaagtgt ggtcagagcc agggaggctt gcagggcca caccagctgt 180
 ggccaagca gctgctgtac aatggctgta tggcagtatc gatgtgaact tcctgataat 240
 aaacagaact cagcagcaaa taaaccaga ttgttctgat cagtaacgag aggctgtgca 300
 aaaagctgag aaatgtacag cccttcctgc tcaccactac agcccttcat gcagcggcct 360
 gctgcactag ggcctgcttg gccagaggca gggctgcagc ttgagatacg cagcaccag 420
 taccacagca ggctgcat 439

<210> 49
 <211> 22
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> ĐOẠN MỖI BÊN TRÁI

 <400> 49
 ttgggtgtgtg ctaataggca gt 22

 <210> 50
 <211> 20
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> ĐOẠN MỖI BÊN PHẢI

 <400> 50
 tagtcctcgt tgtgggaggt 20

 <210> 51
 <211> 20
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> ĐOẠN MỖI BÊN TRÁI

 <400> 51
 gcattgatct gtccagttgc 20

 <210> 52
 <211> 20
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> ĐOẠN MỖI BÊN PHẢI

 <400> 52
 tactgcaaaa accgcatcac 20

 <210> 53
 <211> 33
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> ChZgARN_ĐOẠN MỖI XUÔI

 <400> 53
 tgtatgagac cacgtaatac agagctaaac cag 33

<210> 54
 <211> 33
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> ChZgaARN_ĐOẠN MỖI NGƯỢC

 <400> 54
 aaacctgggtt tagctctgta ttacgtgggc tca 33

 <210> 55
 <211> 51
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Đoạn mỗi xuôi để tách dòng 'nhánh bên trái' của gARN 7

 <400> 55
 acgggggtcat tagttcatag cccatatatg gtgaacggat aggcagcaag c 51

 <210> 56
 <211> 52
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Đoạn mỗi ngược để tách dòng 'nhánh bên trái' của gARN 7

 <400> 56
 ggcgggcat ttaccgtaag ttatgtaacg tgaggaaggg tctgttactg ga 52

 <210> 57
 <211> 51
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Đoạn mỗi xuôi để tách dòng 'nhánh bên phải' của gARN 7

 <400> 57
 ttctatcgcc ttcttgacga gttcttctga gtcatgtcgg tggaggagaa a 51

 <210> 58
 <211> 53
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Đoạn mỗi ngược để tách dòng 'nhánh bên phải' của gARN 7

 <400> 58
 gtgatggcag gttgggcgtc gcttggtcgg gcatgggatg ttaaagagaa gct 53

<210> 59
 <211> 16
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Sự lặp lại 16 nucleotit

<400> 59
 cggtaaatgg cccgcc 16

<210> 60
 <211> 261
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Nhánh bên trái kèm hai bên của nhiễm sắc thể Z

<400> 60
 tggagatccc ttatttgtga agctttcccc tgcaaacctga gaatcttctg atgcaagtga 60
 gcaatcttct tcccttccct agcagctctc tgttgcacca gtccttccat cttgcacct 120
 ttgttgtaat gcttttccca tggcattgat ctgtccagtt gctacagtgt agaaacaaga 180
 aagtgttttc tcaaaacatt atacatgctt actatcagat ctacttaac attgataaga 240
 agtacacctt tcattacaac a 261

<210> 61
 <211> 245
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Nhánh bên phải kèm hai bên của nhiễm sắc thể Z

<400> 61
 caaactttta atcactttca atcaatgttg gaaatctgca gtacataaat tttattttga 60
 ttcattattc agagaatcag agaaataagg acttttttag aggtttatgt ttattttaga 120
 aaatataaat aaatcttgta tacatagatt tccaactagt aaaattttta gctatcgctg 180
 tagcatttct atgaggagtt ttacttttta ctgtctctta tgaaacttag tggaacaaaa 240
 gctag 245

<210> 62
 <211> 1014
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Nhánh bên trái của nhiễm sắc thể Z

<400> 62

gtgaacggat aggcagcaag cattcagaac aatgaaaaag atctgcgttc tattgtaaaag 60
 caaacaagg tgatgacgga acgtttatct gagatacagg caccatagta caagatatca 120
 aagaggaatt aattcaaccc tcccaacaac agttgggaaa aaaaaaggat aagctgagga 180
 atgcagcctt ctccgatttt tactctggtg ccggaaggag gaggaggagc gggggaaaac 240
 ccagcggact cagactgcat acgcactctc ggaagtacct ctgaggcggg gtgagaggaa 300
 gggtggatct aggtccccc cgctccacac actcacgtgc tcgtgtactg gttcaggctg 360
 tgacttggca actcctctac atctatccat ctagttagac tcaacaagac catcgtgatt 420
 caggagtctg aagatgctta actctctcag aacgtctagt tctcctggcc tggtaaattgc 480
 tatgtttctc atactgcctc tctgaagaat gcttcgaatg cttctagtga tgctctaaag 540
 ttctaaacag aaaaatctgg agacagttct ggtctttaga tagaaaaaat gccaacatgc 600
 caaaggatgg ttacatcctt caagcaacct tgttgcattg tgtacaatag actcatgtaa 660
 taacttagcc gtagtcattg tatctcttat ttacctgttc gttattacat tttcctggta 720
 ctgctttata ttttagtcagt tgcctttta agacaaattt tttggtgtgt gctaattaggc 780
 agttaccaa tgctctagag ggagggaata taatcagtga gtatgtagat gtatatagat 840
 gtataaccag taagtataca gagaagttag ggtgggtctt tgagagcata tagctggaaa 900
 gattatcaca tgggaaaagg taatcagaaa taaatggaaa aatgctcacc agtgtcatag 960
 gctcacaaga aaactcccca aggagcaatc tccagtaaca gacccttcct cagt 1014

<210> 63

<211> 629

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Gen tăng cường CMV

<400> 63

tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 60
 cccgccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 120
 attgacgtca atgggtggag tatttacggt aaactgcca cttggcagta catcaagtgt 180
 atcatatgcc aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt 240
 atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 300
 tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 360
 actcacgggg atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc 420
 aaaatcaacg ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg 480

gtaggcgtgt acggtgggag gtctatataa gcagagctgg tttagtgaac cgtcagatcc 540
 gctagcgcta ccggactcag atctcgagct caagcttcga attctgcagt cgacggtacc 600
 gcgggccccg gatccaccgg tcgccacca 629

<210> 64
 <211> 714
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Gen thông báo dsRED2

<400> 64
 tggcctcctc cgagaacgtc atcaccgagt tcatgcgctt caaggtgctc atggagggca 60
 ccgtgaacgg ccacgagttc gagatcgagg gcgagggcga gggccgcccc tacgagggcc 120
 acaacaccgt gaagctgaag gtgaccaagg gcggccccct gcccttcgcc tgggacatcc 180
 tgtccccca gttccagtac ggctccaagg tgtacgtgaa gcaccccgcc gacatccccg 240
 actacaagaa gctgtccttc cccgagggct tcaagtggga gcgcgtgatg aacttcgagg 300
 acggcggcgt ggcgaccgtg acccaggact cctccctgca ggacggctgc ttcatttaca 360
 aggtgaagtt catcggcgtg aacttcccct ccgacggccc cgtgatgcag aagaagacca 420
 tgggctggga ggctccacc gagcgcctgt acccccgcga cggcgtgctg aaggcgaga 480
 cccacaaggc cctgaagctg aaggacggcg gccactacct ggtggagttc aagtccatct 540
 acatggccaa gaagcccgtg cagctgcccc gctactacta cgtggacgcc aagctggaca 600
 tcacctcca caacgaggac tacaccatcg tggagcagta cgagcgacc gagggccgcc 660
 accacctgtt cctgtagcgg ccgcgactct agatcataat cagccatacc acat 714

<210> 65
 <211> 1026
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Nhánh bên phải của nhiễm sắc thể Z

<400> 65
 gagtcatgtc ggtggaggag aaagtggctg tagaaaaggt aacggacaag aaatccctca 60
 ctggcttagc tgtaaaggct agaggaaaat catcttctgt tctccatata catctcctgt 120
 aattaggcac ctgtgccctt taaaaaagta aagtaatcaa agagtcattt ggtgaaaata 180
 gaagccaaac agttacaaaa tttagttaat attcaaagac aaaacaactc tggttccaaa 240
 tataaaaccc tttgtgtaaa cactgcggga aggtgccaaa gagccaccta cacaaagaga 300

```

tgccacagag taatttgcgt accccaacaa taggatcata tatgggcaaa ccaaaagacc      360
aacaaagacc ccttgcaata atcccttgat tggcagaagc aagtgcagcg ttagatttag      420
tatgtgtgac aattctgaga ataggaagga atttactga aactgtgtga agatttgtat      480
gttaaataca tataatgagg gttatttagc tctgggggtgt gcatgctatg tggagagatc      540
cccatgtgcc cagcactgca atagactaat gtcagctttt taaaccatca tttggtttga      600
agagtttatt atgattttca gaaacagaat tatgtattca gtgtctgtac attcttataa      660
gcctgctgtt ctgtacacat gaaaaagcta cttatgggtgc aaatcagtat agagaagcag      720
ttttgacata cagtgggtcat gactgggtga tctgctgtca gaaggtgaaga cactgggtata      780
cagaattgca aacgtgagat gaagacctca agtgcaactc tgtcatcata ccacttctct      840
tgctattatc tattcagtgt ttgttctaaa tattcaggca gaaggcttgg tttcacatca      900
ggtgttacat ttaagtattc ttgcccatt ttttgcgtgt tgtatgtgta acttcagatt      960
ctcacattga ctgtgtagtg taaatttgca aatagatgta acagcttctc tttaacatcc     1020
catgca                                                                    1026

```

<210> 66

<211> 2618

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Nhánh bên phải của nhiễm sắc thể Z trên Fig.15

<400> 66

```

tggagatccc ttattgggtga agctttcccc tgcaaacctga gaatcttctg atgcaagtga      60
gcaatcttct tcccttccct agcagctctc tgttgacca gtccttccat ccttgcacct      120
ttgttgtaat gcttttccca tggcattgat ctgtccagtt gctacagtgt agaaacaaga      180
aagtgttttc tcaaaacatt atacatgctt actatcagat cttacttaac attgataaga      240
agtacacctt tcattacaac agtgaacgga taggcagcaa gcattcagaa caatgaaaaa      300
gatctgcgtt ctattgtaaa gcaaacaag gtgatgacgg aacgtttatt tgagatacag      360
gcaccatagt acaagatatc aaagaggaat taattcaacc ctccaacaa cagttgggaa      420
aaaaaaagga taagctgagg aatgcagcct tctccgattt ttactctggg gccggaagga      480
ggaggaggag cgggggaaaa ccagcggac tcagactgca tacgcactct cggaagtacc      540
tctgaggcgg ggtgagagga aggggtggatc taggctcccc ccgctccaca cactcacgtg      600
ctcgtgtact ggttcaggct gtgacttggc aactcctcta catctatcca tctagttaga      660
ctcaacaaga ccatcgtgat tcaggagtct gaagatgctt aactctctca gaacgtctag      720

```

ttctcctggc ctggtaaatg ctatgtttct catactgcct ctctgaagaa tgcttcgaat	780
gcttctagtg atgctctaaa gttctaaaca gaaaaatctg gagacagttc tggctcttag	840
atagaaaaaa tgccaacatg ccaaaggatg gttacatcct tcaagcaacc ttgttgcacg	900
ctgtacaata gactcatgta ataacttagc cgtagtcacg gtatctctta ttacctgtt	960
cgttattaca ttttcctggg actgctttat atttagtcag ttgtcctttt aagacaaatt	1020
ttttggtgtg tgctaataagg cagttaccaa atgttctaga gggagggaat ataatcagtg	1080
agtatgtaga tgtatataga tgtataacca gtaagtatac agagaagtta ggggtggtctt	1140
ttgagagcat atagctggaa agattatcac atgggaaaag gtaatcagaa ataaatggaa	1200
aaatgctcac cagtgtcata ggctcacaag aaaactcccc aaggagcaat ctccagtaac	1260
agacccttcc tcagttggag ttccgcgtta cataacttac ggtaaattggc ccgcctggct	1320
gaccgccccaa cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc	1380
caatagggac tttcattga cgtcaatggg tggagtattht acggtaaact gccacttgg	1440
cagtacatca agtgtatcat atgccaagta cggcccttat tgacgtcaat gacggtaaatt	1500
ggccgcctg gcattatgcc cagtacatga ctttatggga ctttctact tggcagtaca	1560
tctacgtatt agtcatcgct attaccatgg tgatgcggtt ttggcagtac atcaatgggc	1620
gtggatagcg gtttgactca cggggatttc caagtctcca cccattgac gtcaatggga	1680
gtttgttttg gcacaaaaat caacgggact ttccaaaatg tcgtaacaac tccgccccat	1740
tgacgcaaatt gggcggtagg cgtgtacggg gggagggtcta tataagcaga gctggtttag	1800
tgaaccgtca gatccgctag cgctaccgga ctcagatctc gagctcaagc ttgaattct	1860
gcagtcgacg gtaccgcggg cccgggatcc accggctgcc accatggcct cctccgagaa	1920
cgatcatcacc gagttcatgc gcttcaaggt gcgcatggag ggcaccgtga acggccacga	1980
gttcgagatc gagggcgagg gcgagggccg cccctacgag ggccacaaca ccgtgaagct	2040
gaaggtgacc aagggcggcc ccctgccctt cgcctgggac atcctgtccc ccagttcca	2100
gtacggctcc aaggtgtacg tgaagcaccg cgccgacatc cccgactaca agaagctgtc	2160
cttccccgag ggcttcaagt gggagcgcgt gatgaacttc gaggacggcg gcgtggcgac	2220
cgtgaccag gactcctccc tgcaggacgg ctgcttcacg tacaaggtga agttcatcgg	2280
cgtgaacttc ccctccgacg gcccgtgat gcagaagaag accatgggct gggaggcctc	2340
caccgagcgc ctgtaccccc gcgacggcgt gctgaagggc gagaccaca aggccctgaa	2400
gctgaaggac ggcgccact acctgggtga gttcaagtcc atctacatgg ccaagaagcc	2460
cgtgcagctg cccggctact actacgtgga cgccaagctg gacatcacct cccacaacga	2520

ggactacacc atctgtggagc agtacgagcg caccgagggc cgccaccacc tgttcctgta 2580
gcgggccgcga ctctagatca taatcagcca taccacat 2618

<210> 67
<211> 1271
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Nhánh bên trái của nhiễm sắc thể Z trên Fig.15

<400> 67
gagtcattgtc ggtggaggag aaagtggctg tagaaaaggt aacggacaag aaatccctca 60
ctggccttagc tgtaaaggctc agaggaaaat catcttctgt tctccataca catctcctgt 120
aattaggcac ctgtgccctt taaaaaagta aagtaatcaa agagtcattt ggtgaaaata 180
gaagccaaac agttacaaaa tttagttaat attcaaagac aaaacaactc tggttccaaa 240
tataaaaccc tttgtgtaaa cactgcgagg aggtgccaaa gagccaccta cacaagaga 300
tgccacagag taatttgcgt accccaacaa taggatcata tatgggcaaa ccaaaagacc 360
aacaagacc ctttgcaata atcccttgat tggcagaagc aagtcagcg ttagatttag 420
tatgtgtgac aattctgaga ataggaagga atttcactga aactgtgtga agatttgtat 480
gttaaataca tataatgagg gttatttagc tctgggggtgt gcatgctatg tggagagatc 540
cccatgtgcc cagcactgca atagactaat gtcagctttt taaaccatca tttggtttga 600
agagtttatt atgattttca gaaacagaat tatgtattca gtgtctgtac attcttaca 660
gcctgctgtt ctgtacacat gaaaaagcta cttatgggtgc aaatcagtat agagaagcag 720
ttttgacata cagtgggtcat gactgggtga tctgctgtca gaaggaaga cactgggtata 780
cagaattgca aacgtgagat gaagacctca agtgcaactc tgtcatcata ccacttctct 840
tgctattatc tattcagtgt ttgttctaaa tattcaggca gaaggcttgg tttcacatca 900
gggtgttacat ttaagtattc ttgcccatt tttgctgtt tgtatgtgta acttcagatt 960
ctcacattga ctgtgtagtg taaatttgca aatagatgta acagcttctc tttaacatcc 1020
catgcacaaa cttttaatca ctttcaatca atgttggaat tctgcagtac ataaatttta 1080
ttttgattca ttattcagag aatcagagaa ataaggactt ttttagaggt ttatgtttat 1140
tttagaaaat ataaataaat cttgtataca tagatttcca actagtaaaa tttttagcta 1200
tcgctgtagc atttctatga ggagttttac tttttactgt ctcttatgaa acttagtgga 1260
aacaagcta g 1271

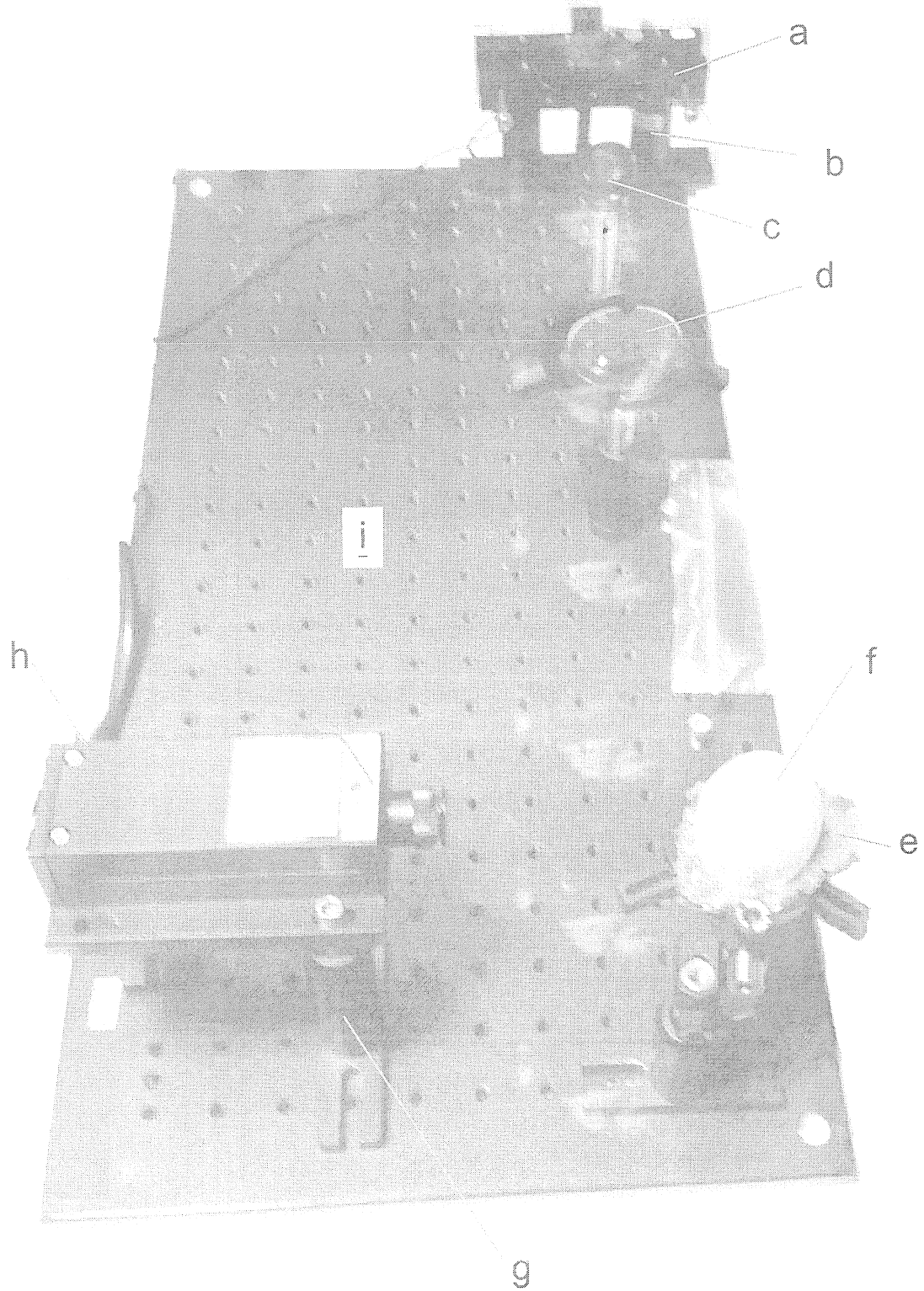


Fig.1A

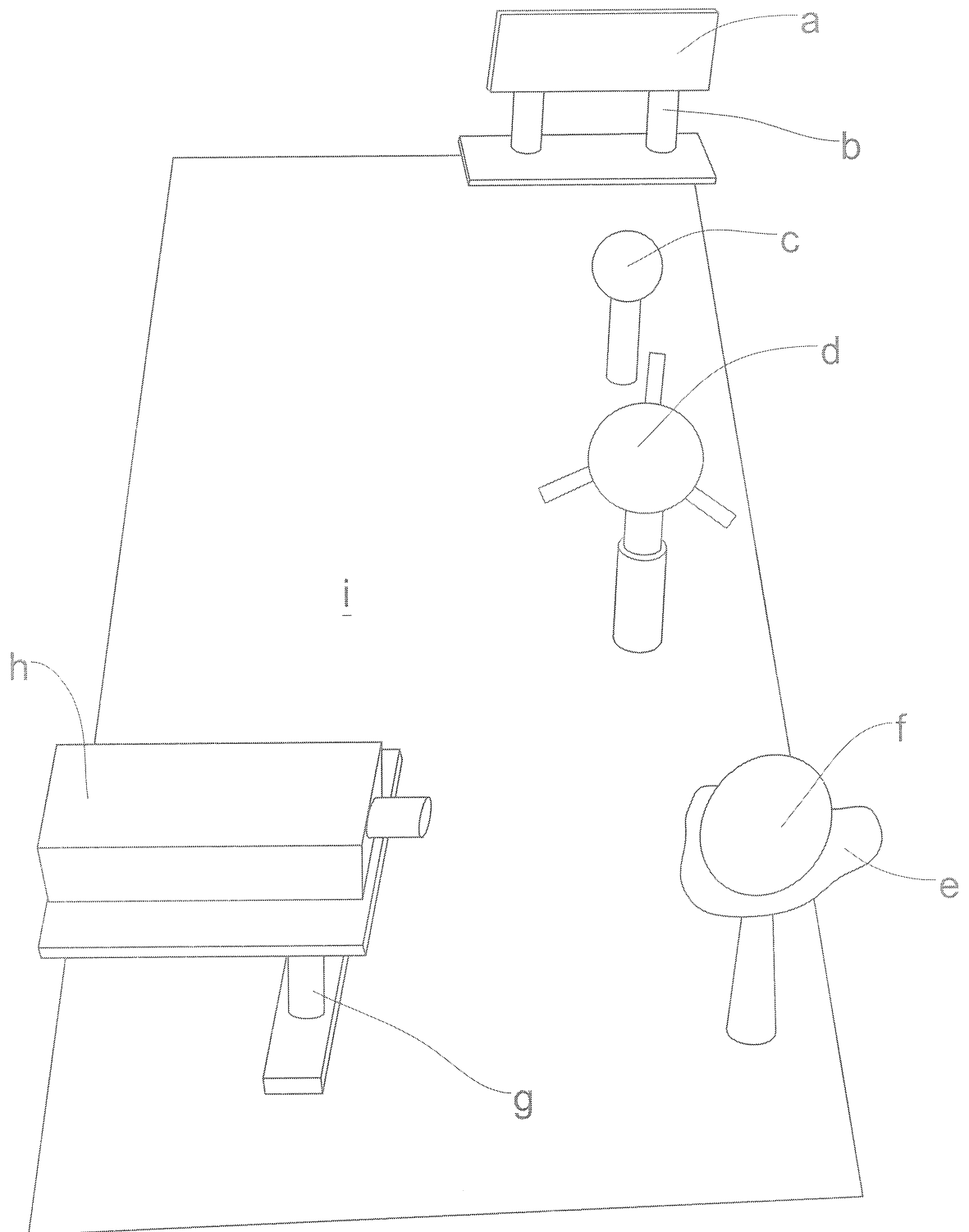


Fig.1B

3/10

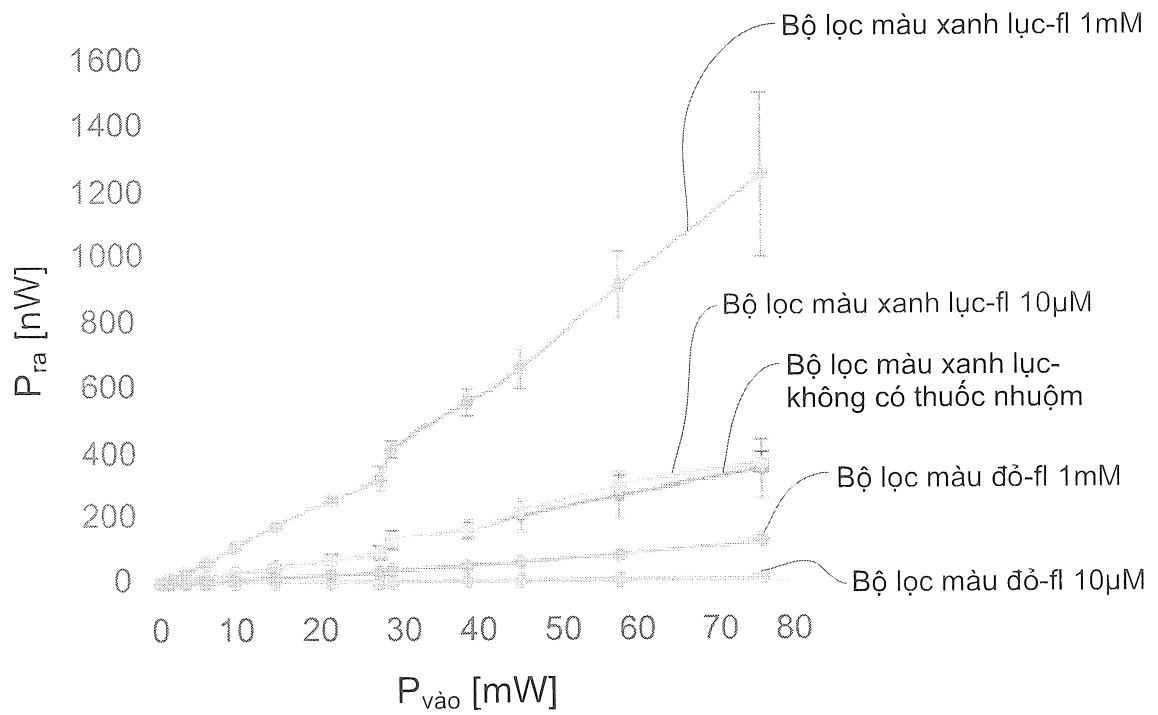


Fig.2

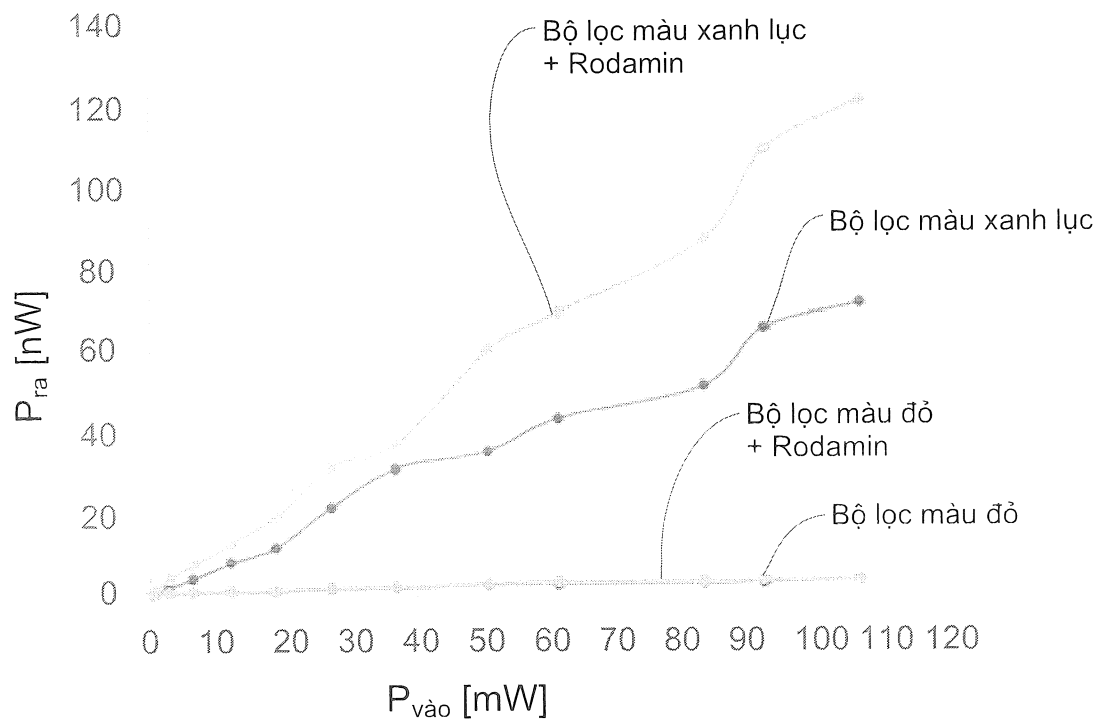


Fig.3

4/10

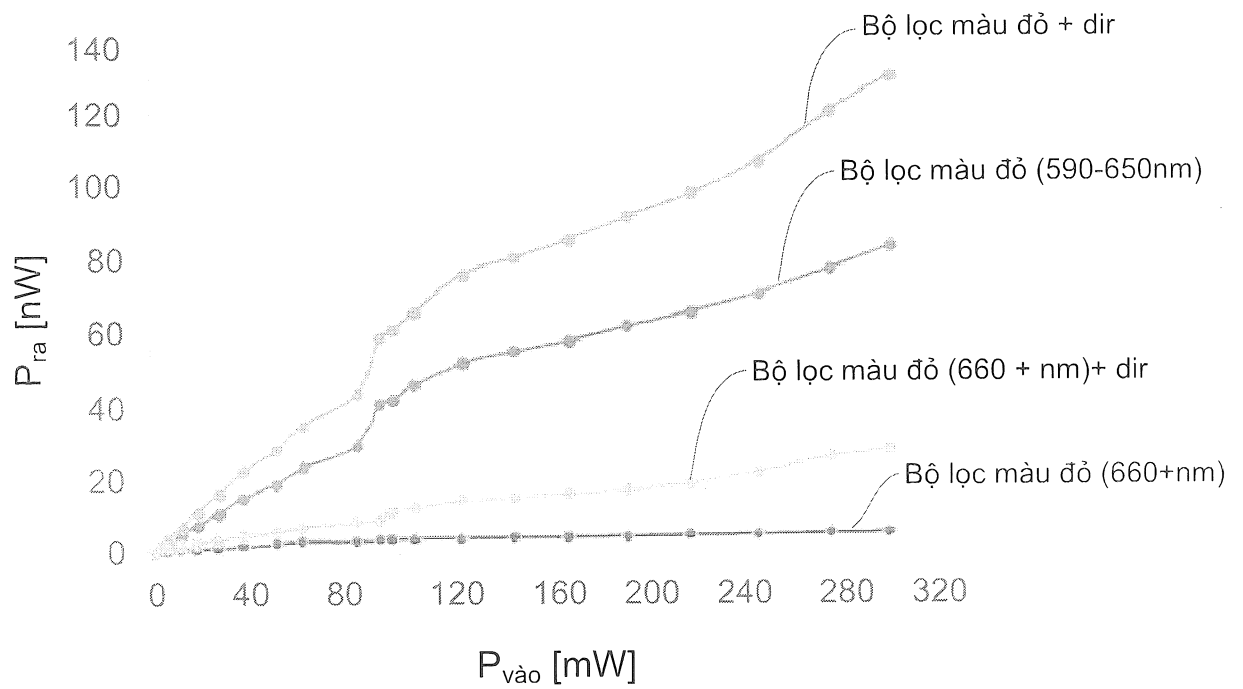


Fig.4

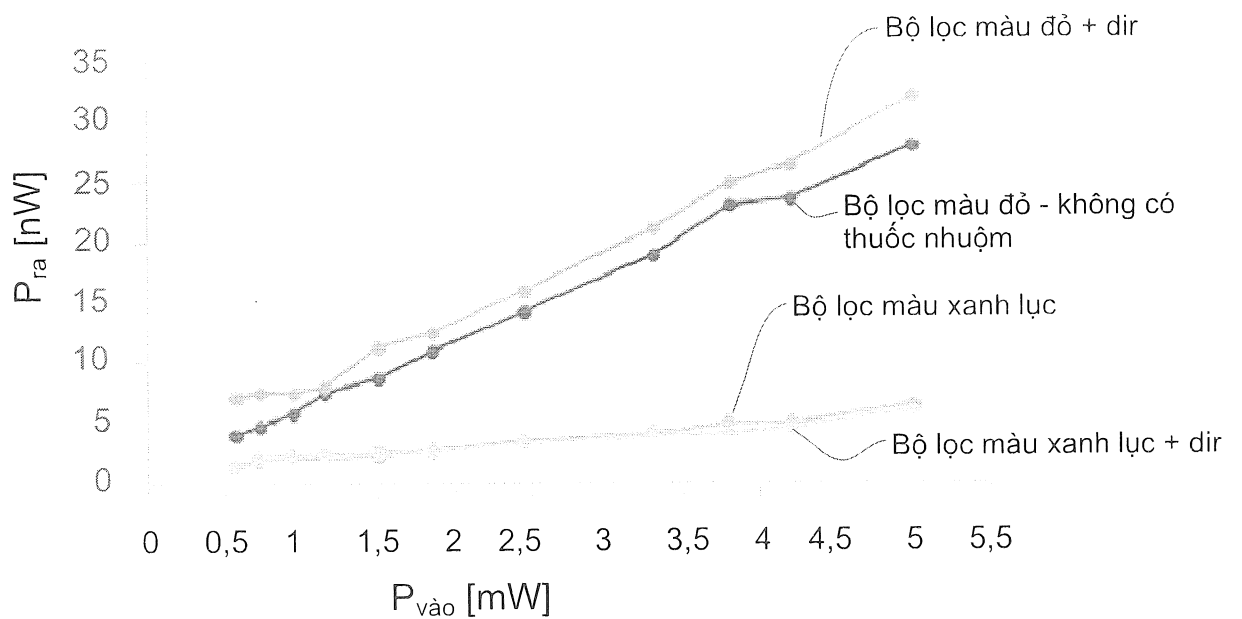


Fig.5

5/10



Fig.6

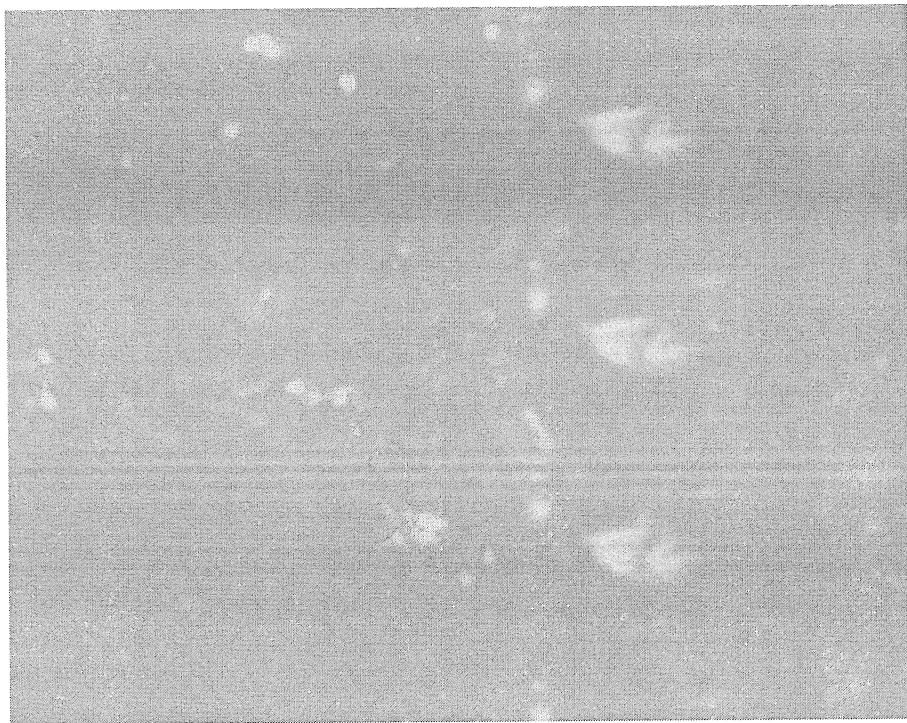


Fig.7

6/10

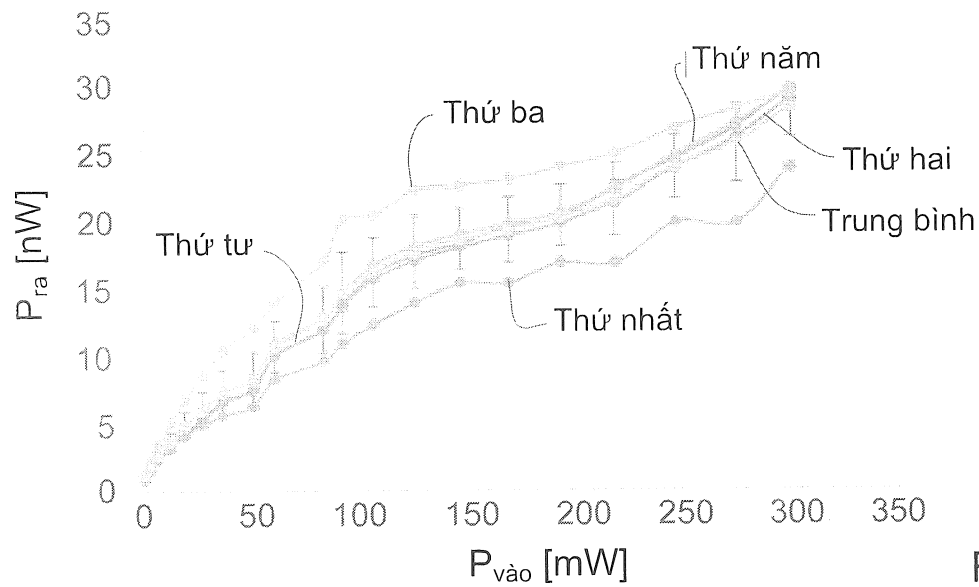


Fig.8

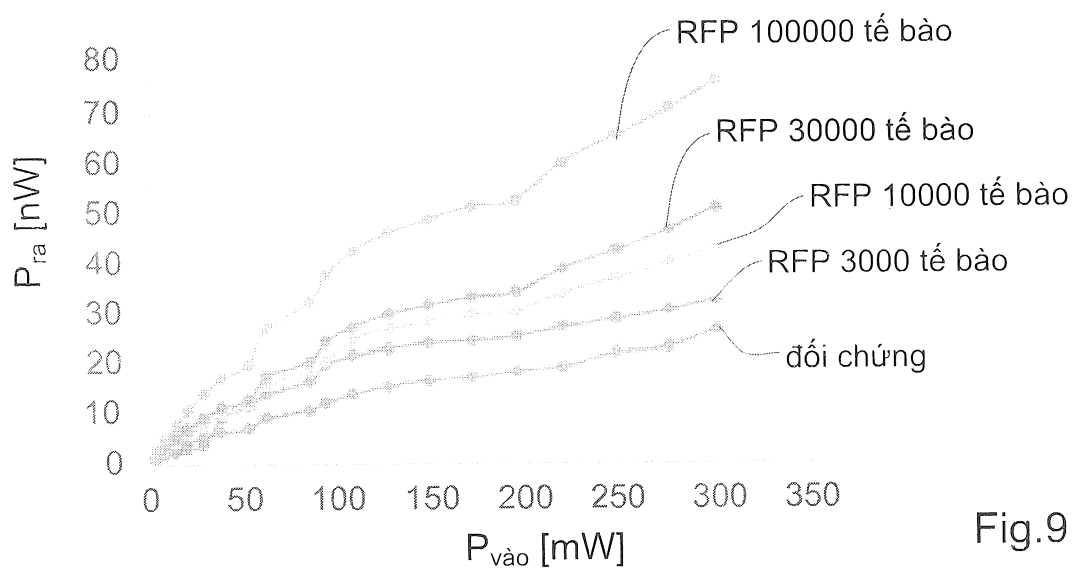


Fig.9

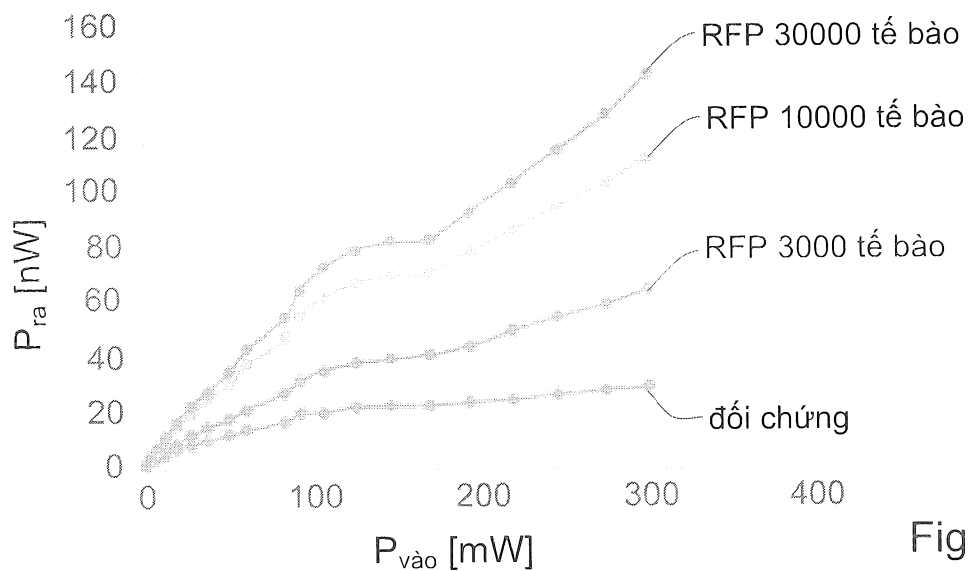


Fig.10

7/10

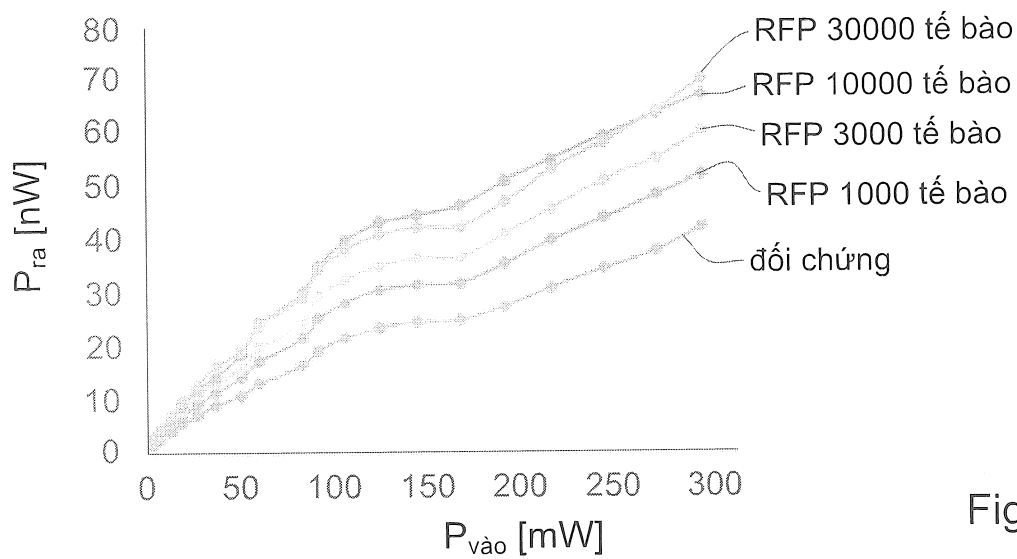


Fig.11A

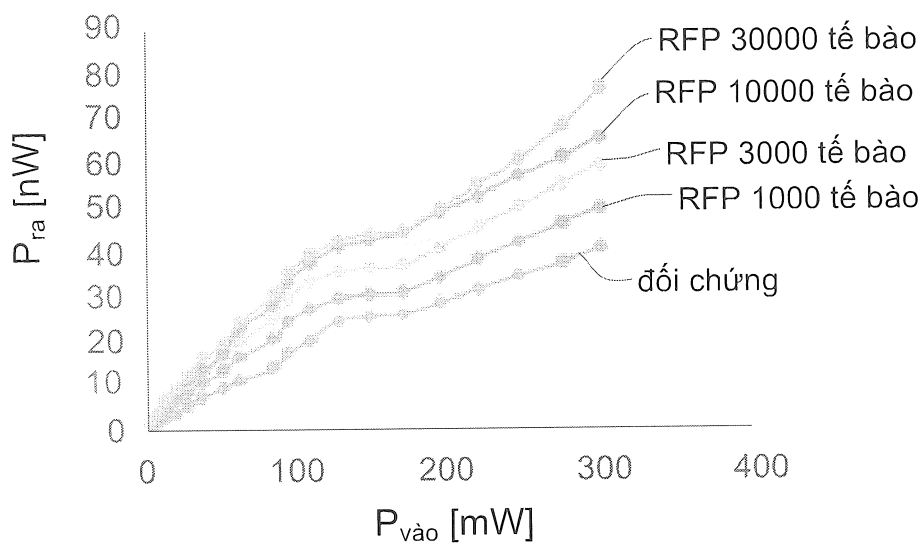


Fig.11B

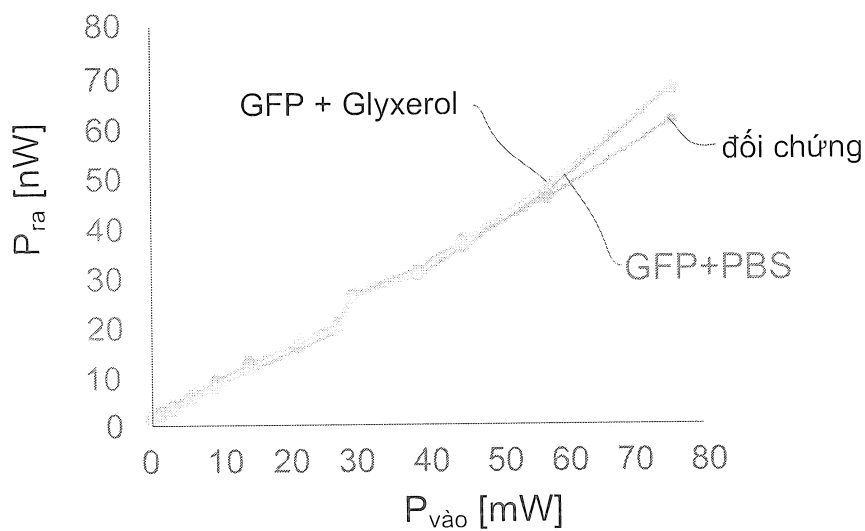


Fig.12

8/10

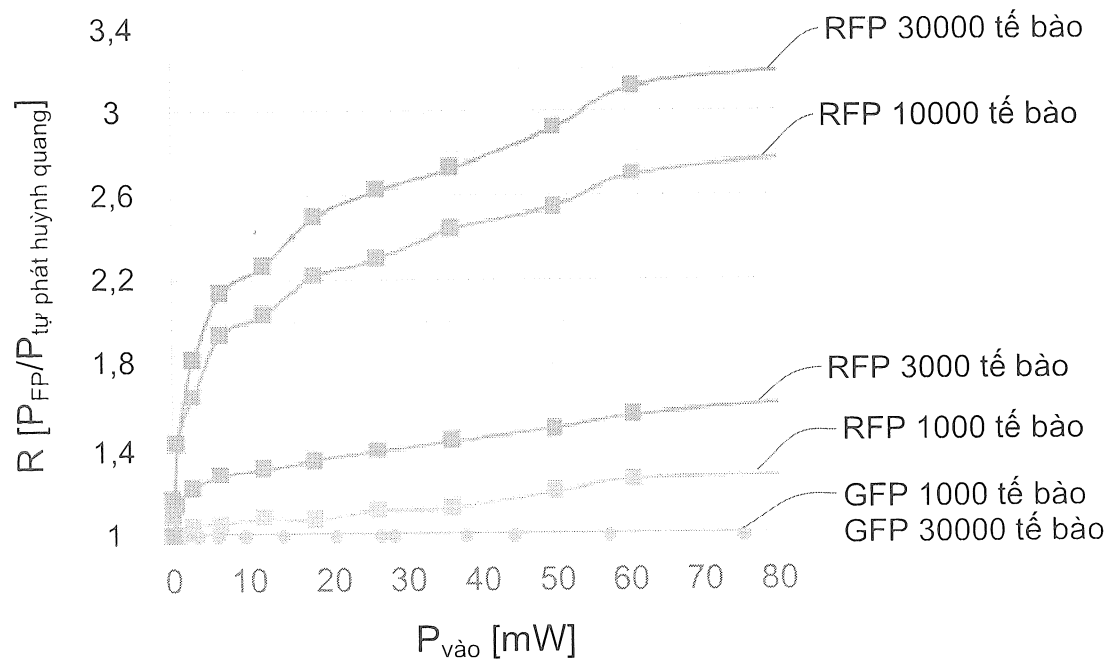


Fig.13

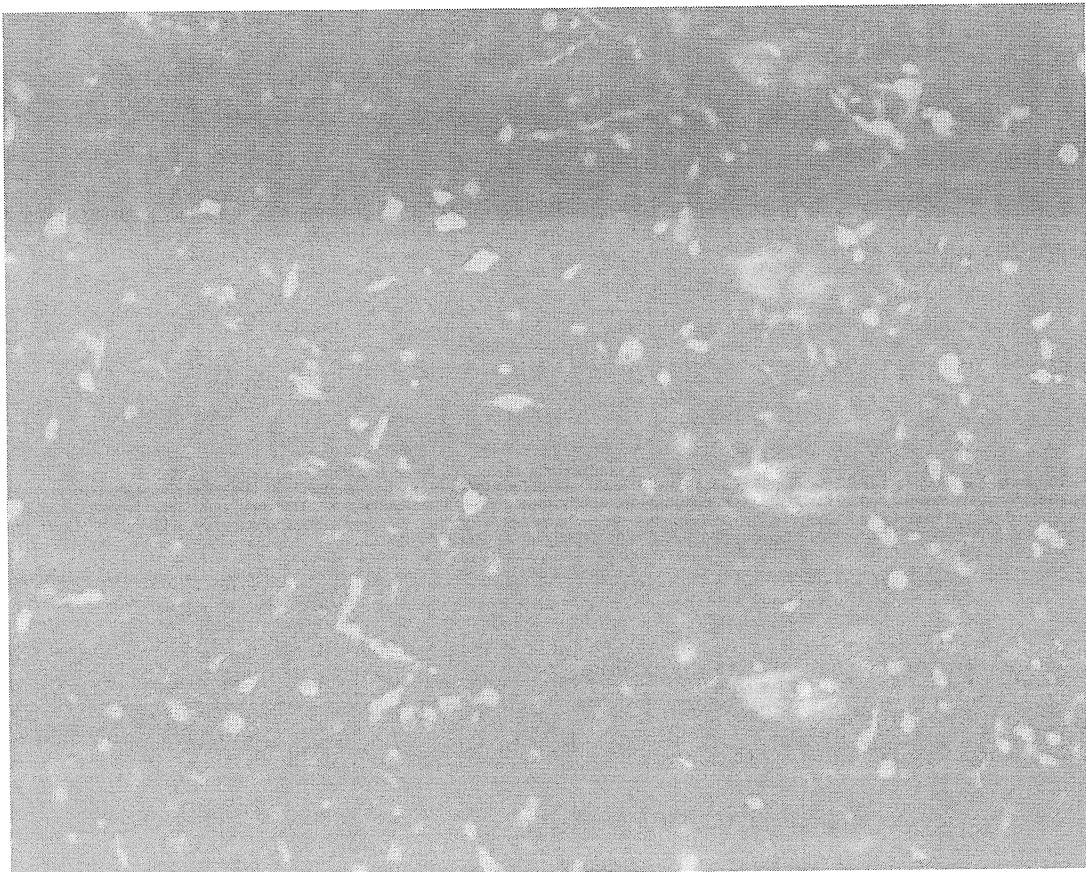


Fig.14

TGGAGATCCCTTATTGGTGAAGCTTTCCCTTGCAAACCTAGATACTTCTGTGATGCAAGTGAGCAATCTTCTCCCTTCCCTAGCAGCTCTCTGTGTCACC
 AGTCCTTCCATCCTTGCACCTTTGTGTAAATGCTTTTCCCATGGCAATTGATCTGTCCAGTTGCTACAGTGTAGAAACAAGAAAGTGTTTTCTCAAAACA
 TTATACATGCTTACTATCAGATCTTACTTAAACATTTGATAAGAGTACACCTTTCAATTACAACAGTGAACGGATAGGCAGCAAGCATTCAGAAACAATGAA
 AAAGATCTGCGTTCTATTGTAAAGCAAAACAAGGTGATGACGGAACGTTTATTGTGAGATACAGGCACCATAGTACAAGATATCAAGAGGAATTAATTC
 AACCTCCCAACAACAGTTGGGAAAAAAGGATAAGCTGAGGAATGCAGCCTTCTCCGATTTTACTCTGTGTCGCGAAGGAGGAGGAGGAGCGGGG
 GAAAAACCGGACTCAGACTGATACGCACCTCTCGGAAGTACCTCTGAGGGGGGTGAGGAAGGTGGATCTAGGCTCCCCCGCTCCACACACT
 CACGTGCTCGTACTGTTCAAGCTGTGACTTGGCAACTCCTCTACATCTATCCATCTAGTTAGACTCAACAAGACCATCGTGATTCAGGAGTCTGTAA
 GATGCTTAACCTCTCTCAGAAAGTCTAGTTCTCTCTGCTGCTTATGCTATGTTTCTCATACTGCTCTCTGAAGAAATGCTTCCGAATGCTTCTAGTGAT
 GCTCTAAAGTTCTAAACAGAAAAATCTGGAGACAGTCTCTGGTCTTTAGATAGAAAAAATGCCAAACATGCCAAAGGATGGTTACATCCTTCAAGCAACCT
 TGTTCATGCTGTACAATAGACTCATGTATAAATCTAGCCGTAGTCACTGCTATCTCTTATTACCTGTCTCGTTATTACATTTTCTCTGCTACTGCTTTAT
 ATTTAGTCAGTTGTCCTTTTAAGACAAAATTTTGTGTGTGCTAATAGGCAGTTACCAAATGTTCTAGAGGGAGGGAATATAATCAGTGAGTATGTAG
 ATGTATATAGATGTATAACCAAGTAAGTATACAGAGAAGTTAGGTGGTCTTTTGGAGAGCATATAGCTGGGAAAGATTATCACATGGGAAAGGTAATCAG
 AAATAAATGGAAAAATGCTCACCAAGTGTATAGGCTCACAGAAAACTCCCCAAGGAGCAATCTCCAGTAACAGACCCCTTCTCAGTTGGAGTTCCTCG
 CGTTACATAA ACTTACGGTAAATGCCCCGCTGGCTGACCGCCAAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCAT
 AGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCA
 TATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCCCTGGCATTATGCCCAGTACATGACCTTAATGGGACTTTTCC
 TACTTGGCAGTACATCTACGTATTAGTCAATCGCTATTACCATGGTGTATGCGGTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTT
 GACTCACGGGATTTC AAGTCTCCACCCCATTTGACGTCAATGGGAGTTCATTTTGGCACCAAAATCAACGGGACITTTCCAAAATGT
 CGTAAACA ACTCCGCCCATTTGACGCAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGAGCTGGTTTAGTGAAACCG
 TCAGATCCGCTAGCGCTACCGGACTCAGATCTCAGAGTCAAGCTTCGAATTTCTGCAGTCGACGGTACCGGGGGCCCCGGGATCCACCG
 GTCGCCACCATGGCCCTCTCCGAGAACGTATCACCGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGCACCCGTGAACGGCCACGAGT
 TCGAGATCGAGGGCGAGGGCGGCCCTACGAGGGCCACAACCCGTGAAGCTGAAGTGAAGCAACCCGCCGACATCCCCGACTACAAGAGCT
 TCGCCCTGGGACATCCTGTCCCCCAGTTTCCAGTACGGCTCCAAGGTGTACGTGAAGCAACCCGCCGACATCCCCGACTACAAGAGCT
 GTCTTCCCCGAGGGCTTCAAGTGGGAGCGGTGATGAACCTTCGAGGACGGGCGTGGCGACCGTGACCCAGGACTCTCTCCCTGCA
 GGACGGCTGCTTCATCTACAAGGTGAAGTTTCATCGGCGTGAACCTTCCCCCTCCGACGGCCCCGTGATGCAGAGAAAGACCATGGGCTGG
 GAGGCCCTCACCGAGCGCTGTACCCCCCGGACGGCGTGTGAAGGGCGAGACCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCAC
 TACCTGGTGGAGTTCAAGTCCATCTACATGGCCAAAGAACCCGTGACGTGCCCCGGTACTACTACGTGGACGCCAAGCTGGACATCA
 CCTCCACACAACGAGGACTACACCATCGTGGAGCAGTACGAGGCCACCGAGGGGCCCAACCTGTTCCTGTAGCGGGCCCGGACTCTA
 GATCATAATCAGCCATACCACAT

Fig. 15A

[tín hiệu SV40 poly(A)-fl ori-gen khởi đầu AmpR-gen khởi đầu sv40-NeoR/KanR-tín hiệu HSVTKpoly(A)-ori]

GAGTCATGTCGGTGGAGGAGAAAGTGGCTGTAGAAAAAGGTAACGGACAAGAAAATCCCTCACTGGCTTAGCTGTAAAGGTCAGAGGAAAAATCATCTTCTG
 TTCTCCATACACATCTCCTGTAATTAGGCACCTGTGCCCTTTAAAAAGTAAGTAATCAAGAGTCAATTTGGTGAAAAATAGAAAGCCAAACAGTTACAA
 AATTAGTTAATAATTCAAAGACAAAACAACCTCTGGTTCCAAATATAAAACCTTTTGTTAAACACTGCGGGAAGGTGCCAAAGAGCCACCTACACAAAG
 AGATGCCACAGAGTAATTTCCGTACCCCAACAATAGGATCATATATATGGGCAAAACCAAGACCCCTTGCAATAATCCCTTGATTGGCAG
 AAGCAAGTGCAGCGTTAGATTAGTATGTGTGACAAATTCTGAGATAGGAAGGAATTTCACTGAAACTGTGTGAAGATTGTATGTTAAATACATATAA
 TGAGGGTTATTTAGCTCTGGGTGTCATGCTATGTGGAGAGATCCCCATGTGCCAGCACTGCAATAGACTAATGTCAGCTTTTAAACCATCATTTG
 GTTTGAAGAGTTTATTATGATTTTCAGAAACAGAAATTATGTATTCAAGTGTCTGTACATTTCTTAAGCCTGCTGTTCTGTACACATGAAAAAGCTACTT
 ATGGTGCAAAATCAGTATAGAGAAGCAGTTTTTGACATACAGTGGTCATGACTGSGTGATCTGCTGTCAAGAGGTAAGACACTGGTATACAGAAATTGCAAA
 CGTGAGATGAAGACCTCAAGTGCAACTCTGTCAATACCACTTCTCTTGCTATTATCTATTTCAGTGTGTTTCTTAAATATTCAGGCAGAAAGGCTTGGT
 TTCACATCAGGTGTTACATTTAAGTATTCTTTTGCCCATTTTGTGCTGTTGTATGTGTAACTTCAGATTCTCACATTGACTGTGTAGTAAATTTGCA
 AATAGATGTAACAGCTTCTCTTTAACATCCCATGCAAACTTTTAATCACATTTGGAAATCTGCAGTACATAAAATTTTATTTGATTC
 ATTATTCAGAGAAATCAGAGAAATAAGGACTTTTTTAGAGGTTTATGTTTTATTTAGAAAAATATAAATCTTGTATACATAGATTTCCAACTAGTAA
 AATTTTAGCTATCGCTGTAGCATTTCTATGAGGAGCTTTTACTTTTACTGCTCTTATGAAACTTAGTGGAACAAAGCTAG

Fig.15B