



(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)



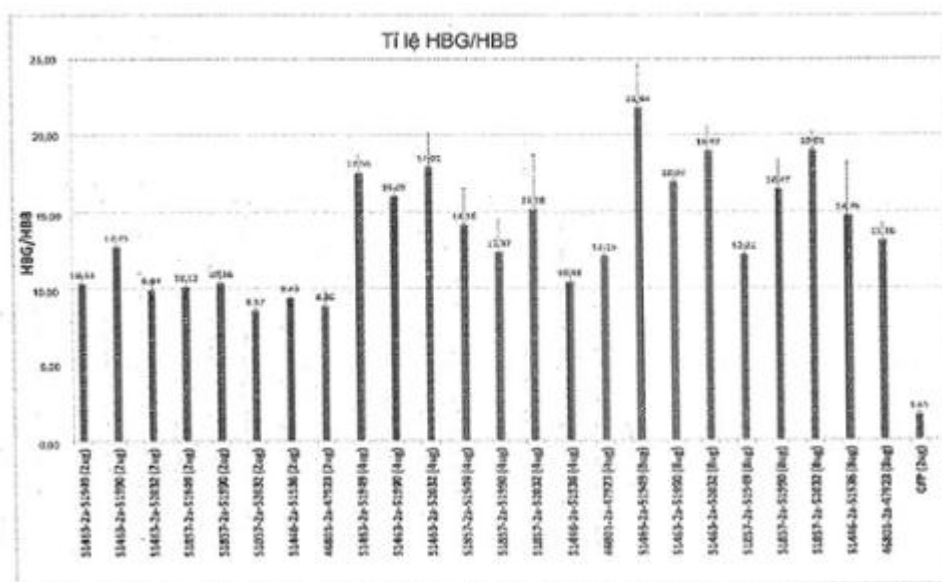
1-0039479

(51)^{2010.01} C12N 9/22; A61K 48/00; C07K 14/705; (13) B
C12N 5/0789; A61K 38/00; C07K 14/46

(21) 1-2017-04476 (22) 12/05/2016
(86) PCT/US2016/032049 12/05/2016 (87) WO 2016/183298 17/11/2016
(30) 62/160,396 12/05/2015 US; 62/303,595 04/03/2016 US
(45) 25/04/2024 433 (43) 26/03/2018 360A
(73) SANGAMO THERAPEUTICS, INC. (US)
501 Canal Blvd., Suite A100, Richmond, CA 94804, United States of America
(72) REIK, Andreas (US).
(74) Công ty TNHH Quốc tế D & N (D&N INTERNATIONAL CO.,LTD.)

(54) PROTEIN NGÓN TAY KẼM, PROTEIN DUNG HỢP CHỨA PROTEIN NGÓN TAY KẼM NÀY VÀ DƯỢC PHẨM CHỨA PROTEIN DUNG HỢP NÀY

(57) Sáng chế thuộc lĩnh vực thao tác di truyền bộ gen, cụ thể là biến đổi hướng đích bộ gen của tế bào tạo máu. Cụ thể là, sáng chế đề cập đến protein ngón tay kẽm, protein dung hợp chứa protein ngón tay kẽm này và dược phẩm chứa protein dung hợp này để sử dụng trong liệu pháp gen và kỹ thuật di truyền bộ gen.



Lĩnh vực kỹ thuật được đề cập

Sáng chế thuộc lĩnh vực thao tác di truyền bộ gen, cụ thể là biến đổi hướng đích bộ gen của tế bào tạo máu.

Tình trạng kỹ thuật của sáng chế

Khi xem xét các nỗ lực trong việc giải trình tự bộ gen đã cho thấy rằng bộ gen người chứa từ 20.000 đến 25.000 gen, nhưng ít hơn 2000 yếu tố điều hòa phiên mã, rõ ràng là một số yếu tố phải tương tác để điều khiển sự biểu hiện gen ở tất cả các biểu hiện đặc hiệu mô, tiến triển, theo thời gian khác nhau của nó. Sự biểu hiện của các gen được điều khiển bởi hỗn hợp có độ phức tạp cao của các yếu tố điều hòa phiên mã chung và đặc hiệu và sự biểu hiện cũng có thể được điều khiển bởi các vùng ADN tác động cis. Các vùng ADN này bao gồm cả vùng ADN cục bộ như trình tự khởi động lõi và các vị trí liên kết yếu tố phiên mã kèm theo của nó cũng như các vùng xa như trình tự tăng cường, trình tự bất hoạt, trình tự cách ly và các vùng kiểm soát locut (LCR) (xem Matson và đồng tác giả (2006) *Ann Rev Genome Hum Genet* 7: 29-50).

Các trình tự tăng cường được xác định lần đầu tiên trong bộ gen virus SV40, và sau đó được phát hiện trong locut chuỗi nặng globulin miễn dịch người. Giờ đây được biết là đóng vai trò điều hòa sự biểu hiện của nhiều gen, các trình tự tăng cường có thể ảnh hưởng chủ yếu đến mô hình không gian và thời gian của sự biểu hiện gen. Cũng được phát hiện ra là các trình tự tăng cường hoạt động theo cách không phụ thuộc vào khoảng cách từ trình tự khởi động lõi của gen, và không phụ thuộc vào hướng của trình tự cụ thể bất kỳ đối với trình tự khởi động này. Các trình tự tăng cường có thể chiếm vài trăm kilobazơ ngược dòng hoặc xuôi dòng của vùng trình tự khởi động lõi, tại đó chúng có thể nằm trong trình tự intron, hoặc thậm chí vượt quá đầu 3' của gen.

Nhiều phương pháp và chế phẩm khác nhau để cắt hướng đích ADN bộ gen đã được mô tả. Các hiện tượng cắt hướng đích này có thể được sử dụng, ví dụ, để gây đột biến hướng đích, gây đột biến mất đoạn hướng đích cho các trình tự ADN tế bào, và tạo thuận lợi cho sự tái tổ hợp hướng đích ở locut nhiễm sắc thể xác định trước. Xem, ví dụ, Patent Mỹ số 9.255.250; 9.200.266; 9.045.763; 9.005.973; 9.150.847; 8.956.828; 8.945.868;

8.703.489; 8.586.526; 6.534.261; 6.599.692; 6.503.717; 6.689.558; 7.067.317; 7.262.054; 7.888.121; 7.972.854; 7.914.796; 7.951.925; 8.110.379; 8.409.861; Công bố patent Mỹ 20030232410; 20050208489; 20050026157; 20050064474; 20060063231; 20080159996; 201000218264; 20120017290; 20110265198; 20130137104; 20130122591; 20130177983; 20130196373; 20150056705 và 20150335708, các phần mô tả này được đưa vào đây bằng cách viện dẫn toàn bộ.

Các phương pháp này thường bao gồm việc sử dụng các hệ phân cắt được thao tác di truyền để gây đứt gãy sợi kép (DSB) hoặc gây ra điểm đứt trong trình tự ADN đích sao cho việc sửa lại điểm đứt gãy do quy trình sinh lỗi như nối đầu cùng không tương đồng (NHEJ) hoặc sửa bằng cách sử dụng khuôn sửa (sửa chữa dựa theo sự tương đồng hoặc HDR) có thể dẫn đến sự bất hoạt gen hoặc sự gắn xen trình tự cần quan tâm (tích hợp hướng đích). Kỹ thuật này cũng có thể được sử dụng để đưa các thay đổi đặc hiệu vị trí vào trình tự bộ gen bằng cách sử dụng oligonucleotit cho, bao gồm đưa vào các đột biến mất đoạn đặc hiệu của vùng gen, hoặc đưa vào các đột biến điểm đặc hiệu hoặc các biến đổi cục bộ (còn được biết là chỉnh sửa gen). Sự cắt có thể xảy ra bằng cách sử dụng các nucleaza đặc hiệu như các nucleaza ngón tay kềm được thao tác di truyền (ZFN), nucleaza tác động giống yếu tố hoạt hóa phiên mã (TALEN), hoặc sử dụng hệ CRISPR/Cas với crARN/tracrARN được thao tác di truyền ('ARN hướng dẫn duy nhất') để hướng dẫn sự cắt đặc hiệu. Ngoài ra, các nucleaza đích đang được phát triển dựa trên hệ thống Argonaute (ví dụ, từ *T. thermophilus*, còn được gọi là 'TtAgo', xem Swarts và đồng tác giả (2014) *Nature* 507(7491): 258-261), mà cũng có thể có khả năng được sử dụng trong việc chỉnh sửa gen và liệu pháp gen.

Các tế bào máu đỏ (RBC), hoặc hồng cầu, là thành phần tế bào chính của máu. Trong thực tế, các RBC chiếm một phần tư số lượng tế bào ở người. Các RBC trưởng thành không có nhân và nhiều bào quan khác ở người, và chứa đầy hemoglobin, metaloprotein có chức năng mang oxy đến các mô cũng như mang cacbon dioxit ra khỏi mô và trở về phổi để đào thải. Protein này chiếm xấp xỉ 97% khối lượng khô của RBC và nó làm tăng khả năng vận chuyển oxy của máu khoảng bảy mươi lần. Hemoglobin là heterotetrame chứa hai chuỗi globin giống alpha (α) và hai chuỗi globin giống beta (β) và 4 nhóm hem. Ở người trưởng thành, tetrame $\alpha_2\beta_2$ được gọi là Hemoglobin A (HbA) hoặc hemoglobin người trưởng thành. Cụ thể là, các chuỗi alpha và beta globin được tổng hợp theo tỉ lệ xấp

xi 1:1 và tỉ lệ này có thể là quan trọng đối với tính ổn định của hemoglobin và RBC. Ở bào thai đang phát triển, dạng khác của hemoglobin, hemoglobin phôi (HbF), được sản sinh mà có ái lực liên kết với oxy cao hơn so với Hemoglobin A sao cho oxy có thể được vận chuyển đến cơ thể của em bé qua dòng máu của mẹ. Có hai gen mã hóa globin phôi mà rất giống nhau về trình tự và được gọi là HPG1 (còn được gọi là Ggamma) và HPG2 (Agamma). Protein hemoglobin phôi còn chứa hai chuỗi α globin, nhưng thay vì các chuỗi β globin người trưởng thành, nó có hai chuỗi gamma (γ) globin phôi (*nghĩa là*, hemoglobin phôi là $\alpha_2\gamma_2$). Ở xấp xỉ 30 tuần thai, sự tổng hợp gamma globin ở thai bắt đầu giảm trong khi sự sản sinh beta globin tăng lên. Đến xấp xỉ 10 tháng tuổi, hemoglobin của trẻ sơ sinh gần như toàn bộ là $\alpha_2\beta_2$ mặc dù một số HbF vẫn còn đến giai đoạn trưởng thành (xấp xỉ 1-3% tổng hemoglobin). Việc điều hòa sự chuyển từ sản sinh gamma globin sang beta globin khá là phức tạp, và chủ yếu bao gồm sự điều hòa giảm phiên mã gamma globin kèm theo điều hòa tăng phiên mã beta globin.

Các khuyết tật di truyền trong trình tự mã hóa các chuỗi hemoglobin có thể chịu trách nhiệm về nhiều bệnh được gọi là các bệnh huyết sắc tố, bao gồm bệnh thiếu máu hồng cầu hình liềm và bệnh tan máu bẩm sinh. Ở phần lớn bệnh nhân mắc các bệnh huyết sắc tố, gen mã hóa gamma globin vẫn có mặt, nhưng sự biểu hiện là tương đối thấp do sự ức chế gen bình thường xảy ra gần lúc sản sinh như được mô tả trên đây.

Ước tính là cứ 5000 người ở Mỹ thì có 1 người mắc bệnh hồng cầu hình liềm (SCD), hầu hết là người có dòng dõi Châu Phi hạ Sahara. Có thể có lợi cho người mang đột biến hồng cầu hình liềm dị hợp tử để bảo vệ chống lại bệnh sốt rét, do đó tính trạng này có thể được chọn lọc tích cực theo thời gian, sao cho ước tính là ở Châu Phi hạ Sahara, một phần ba dân số có tính trạng hồng cầu hình liềm. Bệnh hồng cầu hình liềm bị gây ra bởi đột biến ở gen β globin, dẫn đến valin được thay thế cho axit glutamic ở axit amin #6 (GAG thành GTG ở cấp độ ADN), tại đó hemoglobin tạo thành được gọi là “hemoglobinS” hoặc “HbS”. Trong các điều kiện oxy thấp hơn, sự dịch chuyển cấu hình ở dạng deoxy của HbS làm lộ ra một mảng kỵ nước trên protein này giữa các chuỗi xoắn E và F. Các gốc kỵ nước của valin ở vị trí 6 của chuỗi beta trong hemoglobin có thể liên hợp với mảng kỵ nước này, làm cho các phân tử HbS kết tụ và tạo ra kết tủa sợi. Các khối kết tụ này lại gây ra sự bất thường hoặc ‘tạo hình liềm’ của các RBC, gây mất tính linh động của các tế bào này. Các RBC có hình liềm không còn khả năng đi qua mạng mao mạch và có thể gây ra cơn tắc mạch ở các

bệnh nhân mắc bệnh hồng cầu hình liềm. Ngoài ra, các RBC hình liềm dễ vỡ hơn so với các RBC bình thường, và có xu hướng tan máu, thậm chí dẫn đến thiếu máu ở bệnh nhân.

Việc điều trị và quản lý các bệnh nhân mắc bệnh hồng cầu hình liềm là vấn đề kéo dài cả đời bao gồm điều trị kháng sinh, quản lý đau và truyền trong các giai đoạn cấp tính. Một phương pháp tiếp cận là sử dụng hydroxyure, mà gây tác dụng một phần do sự tăng sản xuất gamma globin. Tuy nhiên, tác dụng phụ lâu dài của liệu pháp hydroxyure mãn tính vẫn chưa được biết, và việc điều trị gây ra các tác dụng phụ không mong muốn và có thể có hiệu lực khác nhau giữa các bệnh nhân khác nhau. Mặc dù các phương pháp điều trị bệnh hồng cầu hình liềm đã tăng hiệu lực, tuổi thọ dự tính của các bệnh nhân vẫn chỉ là giữa hoặc cuối độ tuổi 50 và các bệnh mắc kèm theo bệnh này có tác động sâu sắc đến chất lượng sống của bệnh nhân.

Bệnh tan máu bẩm sinh cũng là các bệnh liên quan đến hemoglobin và thường bao gồm sự biểu hiện giảm các chuỗi globin. Điều này có thể xảy ra do các đột biến trong vùng điều hòa của các gen này hoặc từ đột biến trong trình tự mã hóa globin mà gây ra sự biểu hiện giảm hoặc các mức nồng độ giảm của protein globin chức năng. Bệnh tan máu bẩm sinh alpha chủ yếu xảy ra ở người gốc Tây Phi và Nam Á, và có thể gây kháng bệnh sốt rét. Bệnh tan máu bẩm sinh beta chủ yếu xảy ra ở người gốc Địa Trung Hải, thường là từ Hy Lạp và các vùng ven biển Thổ Nhĩ Kỳ và Ý. Ở bệnh tan máu bẩm sinh thể nhẹ, chỉ một trong số các alen β globin mang đột biến. Các cá thể sẽ mắc bệnh thiếu máu tiểu hồng cầu, và dấu hiệu phát hiện thường bao gồm thể tích trung bình một hồng cầu thấp hơn mức bình thường ($<80\text{fL}$). Các alen của đối tượng mắc bệnh tan máu bẩm sinh thể nhẹ là β^{+}/β hoặc β^0/β (trong đó ' β^{+} ' chỉ các alen cho phép tạo thành một số lượng chuỗi β , ' β ' chỉ các alen β globin kiểu dại, và ' β^0 ' chỉ các đột biến globin β chứa một số dạng đột biến mất đoạn). Đối tượng mắc bệnh tan máu bẩm sinh thể trung gian thường có thể quản lý cuộc sống bình thường nhưng đôi khi có thể cần truyền máu, đặc biệt là tại các thời điểm đau yếu hoặc có thai, tùy theo mức độ nghiêm trọng của tình trạng thiếu máu của họ. Các alen của các bệnh nhân này có thể là β^{+}/β^{+} hoặc β^0/β^{+} . Bệnh tan máu bẩm sinh thể nặng xảy ra khi cả hai alen đều có các đột biến tan máu bẩm sinh. Đây là tình trạng thiếu máu giảm sắc và tiểu hồng cầu nghiêm trọng. Nếu không được điều trị, nó gây thiếu máu, to lách và biến dạng xương nghiêm trọng. Tình trạng này tiến tới tử vong trước tuổi 20. Phương pháp điều trị bao gồm truyền máu định kỳ; phẫu thuật cắt bỏ lách đối với lách to và tạo chelat lượng sắt

quá tải do truyền máu. Phương pháp ghép tủy xương cũng đang được sử dụng để điều trị cho người mắc bệnh tan máu bẩm sinh nghiêm trọng nếu có thể xác định được người hiến tặng thích hợp, nhưng quy trình này có thể có các rủi ro cao.

Một phương pháp được đề xuất để điều trị cả bệnh SCD và bệnh tan máu bẩm sinh beta là làm tăng sự biểu hiện của gamma globin với mục đích để HbF thay thế chức năng hemoglobin bất thường ở người trưởng thành. Như đã đề cập trên đây, phương pháp điều trị cho các bệnh nhân SCD bằng hydroxyure được cho là thành công một phần do tác dụng làm tăng sự biểu hiện gamma globin của nó. Nhóm các hợp chất thứ nhất được phát hiện ra là tác động đến hoạt tính phục hồi gamma globin là các thuốc gây độc tế bào. Khả năng gây ra sự tổng hợp *de novo* của gamma-globin bởi việc sử dụng được lý được chỉ ra lần đầu tiên bằng cách sử dụng 5-azacytidin ở các động vật thí nghiệm (DeSimone (1982) *Proc Nat'l Acad Sci USA* 79(14):4428-31). Các nghiên cứu tiếp theo xác nhận khả năng của 5-azacytidin trong việc làm tăng HbF ở bệnh nhân mắc bệnh tan máu bẩm sinh β và bệnh hồng cầu hình liềm (Ley, và đồng tác giả., (1982) *N. Engl. J. Medicine*, 307: 1469-1475, và Ley, và đồng tác giả., (1983) *Blood* 62: 370-380). Thêm vào đó, các axit béo mạch ngắn (ví dụ, butyrat và các dẫn xuất) đã được chỉ ra trong các hệ thống thí nghiệm là làm tăng HbF (Constantoulakis và đồng tác giả., (1988) *Blood* 72(6):1961-1967). Ngoài ra, có một phần dân số mắc phải tình trạng được gọi là 'Tồn tại dai dẳng về mặt di truyền của hemoglobin bào thai' (Hereditary Persistence of Fetal Hemoglobin: HPFH), trong đó các lượng gia tăng của HbF vẫn tồn tại trong giai đoạn trưởng thành (10-40% trong các dị hợp tử HPFH (xem Thein và đồng tác giả (2009) *Hum. Mol. Genet* 18 (R2): R216-R223). Đây là tình trạng hiếm gặp, nhưng khi vắng mặt các bất thường beta globin kèm theo bất kỳ, tình trạng này không liên quan đến các biểu hiện lâm sàng có ý nghĩa bất kỳ, ngay cả khi 100% hemoglobin của đối tượng này là HbF. Khi các cá thể mắc bệnh tan máu bẩm sinh beta cũng mắc đồng thời HPFH, sự biểu hiện của HbF có thể làm giảm mức độ nghiêm trọng của bệnh này. Ngoài ra, mức độ nghiêm trọng của diễn biến tự nhiên của bệnh hồng cầu hình liềm có thể thay đổi đáng kể giữa các bệnh nhân, và sự thay đổi này, phần nào, có thể do thực tế là một số cá thể mắc bệnh nhẹ hơn biểu hiện mức HbF cao hơn.

Một phương pháp làm tăng sự biểu hiện của HbF bao gồm xác định các gen mà sản phẩm của chúng đóng vai trò trong việc điều hòa sự biểu hiện gamma globin. Một trong số các gen này là BCL11A, được xác định đầu tiên vì vai trò của nó trong việc phát triển các

tế bào lympho. BCL11A mã hóa protein ngón tay kẽm mà được cho là tham gia vào việc điều hòa sự biểu hiện gamma globin đặc hiệu giai đoạn phát triển. BCL11A được biểu hiện trong các tế bào tiền thân của hồng cầu người trưởng thành và sự điều hòa giảm sự biểu hiện của nó dẫn đến sự tăng biểu hiện gamma globin. Ngoài ra, có thể là sự nổi mARN BCL11A được điều hòa phát triển. Trong các tế bào phôi, có thể là các biến thể mARN BCL11A ngắn hơn, còn được gọi là BCL11A-S và BCL11A-XS được biểu hiện chủ yếu, trong khi ở các tế bào của người trưởng thành, các biến thể mARN BCL11A-L và BCL11A-XL dài hơn được biểu hiện chủ yếu. Xem, Sankaran và đồng tác giả (2008) *Science* 322 p. 1839. Protein BCL11A có thể tương tác với locut beta globin để biến đổi cấu hình của nó và nhờ đó sự biểu hiện của nó ở các giai đoạn phát triển khác nhau. Việc sử dụng ARN ức chế gắn đích gen BCL11A đã được đề xuất (xem, ví dụ., Công bố patent Mỹ 20110182867) nhưng công nghệ này có một số nhược điểm tiềm tàng, cụ thể là sự bất hoạt hoàn toàn có thể không đạt được, sự phân phối các ARN này có thể là khó khăn và các ARN này phải có mặt liên tục, đòi hỏi nhiều lần điều trị để sống được.

Sự hướng đích của các trình tự tăng cường BCL11A cung cấp cơ chế để làm tăng HbF. Xem, ví dụ, Công bố patent Mỹ số 20150132269. Các nghiên cứu tương quan toàn bộ nhiễm sắc thể đã xác định được tập hợp các biến đổi di truyền ở BCL11A mà liên quan đến các mức HbF tăng. Các biến đổi này là tập hợp của các SNP được tìm thấy trong các vùng không mã hóa của BCL11A mà hoạt động như là vùng tăng cường giới hạn dòng, đặc hiệu giai đoạn. Nghiên cứu thêm đã chỉ ra rằng trình tự tăng cường BCL11A này là cần thiết trong các tế bào hồng cầu để biểu hiện BCL11A, nhưng không cần thiết cho sự biểu hiện của nó ở các tế bào B (xem Bauer và đồng tác giả, (2013) *Science* 343:253-257). Vùng tăng cường này được tìm thấy trong intron 2 của gen BCL11A, và ba vùng quá mẫn ADNseI (thường biểu thị trạng thái sợi nhiễm sắc liên quan đến khả năng điều hòa) trong intron 2 được xác định. Ba vùng này được xác định là “+62”, “+58” và “+55” theo khoảng cách kilobazơ từ vị trí khởi đầu phiên mã của BCL11A. Các vùng tăng cường này dài xấp xỉ là 350 (+55); 550 (+58); và 350 (+62) nucleotit (Bauer 2013, *ibid*).

Do đó, vẫn có nhu cầu về các phương pháp và chế phẩm khác để thay đổi sự biểu hiện gen BCL11A, ví dụ, để điều trị bệnh huyết sắc tố như bệnh hồng cầu hình liềm và bệnh tan máu bẩm sinh beta.

Bản chất kỹ thuật của sáng chế

Sáng chế mô tả các chế phẩm và phương pháp để sử dụng trong liệu pháp gen và thao tác di truyền bộ gen. Cụ thể là, các phương pháp và chế phẩm được mô tả liên quan đến việc bất hoạt gen BCL11A (ví dụ, bằng cách triệt tiêu hoàn toàn hoặc một phần sự biểu hiện của nó), ví dụ, gen đóng vai trò là trình tự điều hòa một hoặc nhiều gen bổ sung. Cụ thể là, sáng chế mô tả các phương pháp và các chế phẩm để gây cản trở chức năng tăng cường trong gen BCL11A để làm giảm hoặc bất hoạt hoạt tính của nó trong dòng tế bào cụ thể. Ngoài ra, sáng chế đề xuất các phương pháp và chế phẩm để gây cản trở các chức năng tăng cường BCL11A trong đó các trình tự tăng cường không nằm trong gen BCL11A. Tác dụng điều hòa giảm gen BCL11A trong các trường hợp này lại gây ra sự biểu hiện tăng của gamma globin.

Theo một số khía cạnh, sáng chế bao gồm protein ngón tay kẽm xuất hiện không tự nhiên bao gồm protein ngón tay kẽm (ZFP) chứa 4, 5 hoặc 6 ngón tay, mỗi ngón tay chứa vùng xoắn nhận biết mà nhận biết vị trí cận đích trong đó các vùng xoắn nhận biết này chứa các trình tự theo trật tự được thể hiện trong một hàng duy nhất của Bảng 1. Trong một số phương án, ZFP này chứa các vùng xoắn nhận biết như được thể hiện trong Bảng 1 cho các protein được ký hiệu như sau: 51446, 51463, 51484, 51856, 51857 hoặc 51862 (mà liên kết với vị trí đích được thể hiện trong SEQ ID NO:1) và 51536, 51949, 51990, 51993, 51979, 51982, 52015, 52032 (mà liên kết với vị trí đích được thể hiện trong SEQ ID NO:12). Do đó, trong một số phương án, được đề xuất ở đây là protein ngón tay kẽm bao gồm các vùng xoắn nhận biết sau đây:

- (i) F1: STGNLTN (SEQ ID NO:7);
 F2: TSGSLTR (SEQ ID NO:5);
 F3: DQSNLRA (SEQ ID NO:2); và
 F4: AQCCLFH (SEQ ID NO:6); hoặc
- (ii) F1: DQSNLRA (SEQ ID NO:2);
 F2: RPYTLRL (SEQ ID NO:3);
 F3: SRGALKT (SEQ ID NO:8);
 F4: TSGSLTR (SEQ ID NO:5);
 F5: DQSNLRA (SEQ ID NO:2); và

- F6: AQCCLFH (SEQ ID NO:6);
- (iii) F1: DQSNLRA (SEQ ID NO:2);
- F2: RNFSLTMT (SEQ ID NO:9);
- F3: SNGNLRN (SEQ ID NO:10) hoặc STGNLTN (SEQ ID NO:7) hoặc SSYNLAN (SEQ ID NO:11);
- F4: TSGSLTR (SEQ ID NO:5);
- F5: DQSNLRA (SEQ ID NO:2); và
- F6: AQCCLFH (SEQ ID NO:6); hoặc
- (iv) F1: RSDHLTQ (SEQ ID NO:13);
- F2: QSGHLAR (SEQ ID NO:14);
- F3: QKGTGE (SEQ ID NO:15);
- F4: RHRDLR (SEQ ID NO:18); và
- F5: RRDNLHS (SEQ ID NO:17); hoặc
- (v) F1: RNDHRTT (SEQ ID NO:19);
- F2: QKAHLIR (SEQ ID NO:20);
- F3: QKGTGE (SEQ ID NO:15);
- F4: RGRDLR (SEQ ID NO:21) hoặc LKRTLKR (SEQ ID NO:25); và
- F5: RRDNLHS (SEQ ID NO:17); hoặc
- (vi) F1: RSDHLTQ (SEQ ID NO:13);
- F2: QRAHLTR (SEQ ID NO:22);
- F3: QKGTGE (SEQ ID NO:15) hoặc QSGTRNH (SEQ ID NO:24);
- F4: HRNTLVR (SEQ ID NO:23); và
- F5: RRDNLHS (SEQ ID NO:17); hoặc
- (vii) F1: RSDHLTQ (SEQ ID NO:13);
- F2: QKAHLIR (SEQ ID NO:20);
- F3: QKGTGE (SEQ ID NO:15) hoặc QSGTRNH (SEQ ID NO:24);

- F4: RGRDLSR (SEQ ID NO:21); và
 F5: RRDNLHS (SEQ ID NO:17); hoặc
 (viii) F1: F1: RSDHLTQ (SEQ ID NO:13);
 F2: QSGHLAR (SEQ ID NO:14);
 F3: QSGTRNH (SEQ ID NO:24);
 F4: QSSDLSR (SEQ ID NO:16); và
 F5: RRDNLHS (SEQ ID NO:17).

Trong một số phương án, các protein ngón tay kẽm như được mô tả ở đây được dung hợp với vùng chức năng (*ví dụ*, vùng hoạt hóa phiên mã, vùng ức chế phiên mã, vùng phân cắt (để tạo ra nucleaza ngón tay kẽm), v.v...). Các nucleaza ngón tay kẽm có thể được sử dụng trong các cặp đime hóa để phân cắt tại hoặc gần một hoặc cả hai vị trí đích cho các ZFN của cặp này, *ví dụ*, “các phần tử kết hợp bên trái” của Bảng 1 (*ví dụ*, 51446, 51463, 51484, 51856, 51857, hoặc 51862) có thể tạo ra các đime với các “phần tử kết hợp bên phải” của Bảng 1 (*ví dụ*, 51536, 51949, 51990, 51993, 51979, 51982, 52015, hoặc 52032) để phân cắt các trình tự tăng cường BCL11A.

Theo một khía cạnh khác, sáng chế bao gồm phương pháp phân phối ít nhất một nucleaza (*ví dụ*, nucleaza liên kết với trình tự tăng cường BCL11A) cho tế bào gốc người hoặc tế bào tiền thân của người (HSC/PC) với mục đích thao tác di truyền bộ gen. Trong một số phương án, nucleaza này bao gồm protein ngón tay kẽm (ZFP) chứa 4, 5 hoặc 6 ngón tay, mỗi ngón tay chứa vùng xoắn nhận biết mà nhận biết vị trí cận đích trong đó các vùng xoắn nhận biết này chứa các trình tự theo trật tự được thể hiện trong một hàng duy nhất của Bảng 1. (Các) nucleaza như được mô tả ở đây có thể chứa thêm cầu nối (*ví dụ*, giữa vùng liên kết ADN và vùng phân cắt), *ví dụ*, cầu nối như được thể hiện trong các SEQ ID NO:26-29 và công bố Patent Mỹ số 20150132269.

Trong một số phương án, nucleaza này được phân phối dưới dạng peptit, trong khi ở các phương án khác nó được phân phối dưới dạng axit nucleic mã hóa ít nhất một nucleaza. Trong một số phương án, nhiều hơn một nucleaza được sử dụng. Trong một số phương án được ưu tiên hơn, axit nucleic mã hóa nucleaza là mARN, và trong một số trường hợp, mARN này được bảo vệ. Theo một số khía cạnh, mARN này có thể được biến đổi hóa học (*Xem ví dụ* Kormann và đồng tác giả, (2011) *Nature Biotechnology* 29(2):154-

157). Theo các khía cạnh khác, mARN này có thể chứa mã ARCA (xem Patent Mỹ số 7.074.596 và 8.153.773). Trong các phương án khác nữa, mARN này có thể chứa hỗn hợp của nucleotit biến đổi và chưa biến đổi (xem Công bố patent Mỹ 2012/0195936). Trong một phương án được ưu tiên, axit nucleic mã hóa (các) nucleaza được phân phối cho HSC/PC bằng kỹ thuật xung điện. Trong một số phương án, nucleaza phân cắt ra tại hoặc gần vị trí liên kết của yếu tố phiên mã. Theo một số khía cạnh, yếu tố phiên mã này là GATA-1.

Theo các khía cạnh khác, sáng chế bao gồm tế bào hoặc dòng tế bào, trong đó trình tự tăng cường BCL11A nội sinh được biến đổi di truyền bằng nucleaza như được mô tả ở đây (ví dụ, được thể hiện trong Bảng 1), ví dụ, so với trình tự kiểu đại của tế bào này. Các tế bào hoặc dòng tế bào được biến đổi bằng nucleaza như được mô tả ở đây có thể phân biệt được về cấu trúc và/hoặc chức năng với tế bào kiểu đại và tế bào được biến đổi khác (qua trung gian nucleaza). Tế bào và dòng tế bào đã biến đổi di truyền này có thể là dị hợp tử hoặc đồng hợp tử đối với biến đổi này. Các biến đổi này có thể bao gồm đột biến gắn xen (ví dụ, gắn xen chuyển gen), đột biến mất đoạn và/hoặc tổ hợp của chúng. Trong một số phương án được ưu tiên hơn, đột biến gắn xen, mất đoạn và/hoặc tổ hợp của chúng phá vỡ vị trí liên kết yếu tố phiên mã. Trong một số phương án, biến đổi này là ở hoặc ở gần (các) vị trí phân cắt và/hoặc liên kết (các) nucleaza, ví dụ, trong khoảng 1-300 (hoặc giá trị bất kỳ nằm trong khoảng này) cặp bazơ ngược dòng hoặc xuôi dòng (các) vị trí phân cắt, tốt hơn, nếu trong khoảng 1-100 cặp bazơ (hoặc giá trị bất kỳ nằm trong khoảng này) ở một trong hai bên của (các) vị trí phân cắt và/hoặc liên kết được thể hiện trong Bảng 1, thậm chí tốt hơn nữa nếu trong khoảng 1 đến 50 cặp bazơ (hoặc giá trị bất kỳ nằm trong khoảng này) trên một trong hai bên của (các) vị trí phân cắt và/hoặc liên kết. Biến đổi này cũng có thể bao gồm các biến đổi của một hoặc nhiều nucleotit ở vị trí phân cắt và/hoặc một hoặc nhiều vị trí liên kết. Trong một số phương án, một hoặc nhiều vị trí đích nucleaza không được biến đổi. Trong các phương án khác, ít nhất một trong số các vị trí đích cho (các) nucleaza được biến đổi. Trong một số phương án, biến đổi này là ở hoặc ở gần vùng “+58” của trình tự tăng cường BCL11A, ví dụ, tại hoặc gần vị trí liên kết nucleaza được thể hiện trong bất kỳ trong số SEQ ID NO:1 và SEQ ID NO:12. Tế bào hoặc dòng tế bào bất kỳ có thể được biến đổi bởi các nucleaza như được mô tả ở đây, ví dụ tế bào gốc (tế bào gốc tạo máu như tế bào gốc tạo máu CD34+) hoặc tế bào tiền thân của tế bào hồng cầu (RBC). Cũng được mô tả là các tế bào hoặc dòng tế bào thu được sau khi biến đổi bằng

nucleaza như được mô tả ở đây, ví dụ các tế bào hoặc dòng tế bào có nguồn gốc từ tế bào hoặc dòng tế bào được biến đổi bằng nucleaza. Các tế bào được biệt hóa một phần hoặc hoàn toàn có nguồn gốc từ các tế bào gốc đã biến đổi như được mô tả ở đây cũng được đề xuất (ví dụ, các RBC hoặc các tế bào tiền thân RBC). Các tế bào có nguồn gốc từ các tế bào được biến đổi bằng nucleaza có thể được nuôi cấy (và/hoặc được biệt hóa) trên *in vitro* (nuôi cấy) hoặc có thể biệt hóa bên trong cá thể sống, ví dụ sau khi sử dụng *ex vivo* tế bào gốc được biến đổi bằng nucleaza. Bất kỳ trong số các tế bào hoặc dòng tế bào được biến đổi di truyền đã mô tả ở đây có thể thể hiện sự biểu hiện gamma globin tăng. Các chế phẩm như các được phẩm chứa các tế bào biến đổi di truyền như được mô tả ở đây cũng được đề xuất.

Theo các khía cạnh khác, sáng chế bao gồm phương pháp phân phối axit nucleic cho cho tế bào đích để tạo ra tế bào được biến đổi di truyền, trong đó axit nucleic cho này được tích hợp vào tế bào. Axit nucleic cho này có thể được phân phối trước, sau hoặc cùng với axit nucleic mã hóa (các) nucleaza của Bảng 1. Axit nucleic cho này có thể chứa trình tự ngoại sinh (gen chuyển) cần được tích hợp vào bộ gen của tế bào, ví dụ, locut nội sinh. Trong một số phương án, axit nucleic cho này có thể chứa gen có chiều dài đầy đủ hoặc một đoạn của nó được gắn bên sườn bởi các vùng tương đồng có vị trí phân cắt hướng đích. Trong một số phương án, axit nucleic cho này thiếu các vùng tương đồng và được tích hợp vào locut đích bằng cơ chế độc lập với sự tương đồng (*nghĩa là*, NHEJ). Axit nucleic cho này có thể bao gồm trình tự axit nucleic bất kỳ, ví dụ axit nucleic mã, khi được sử dụng làm cơ chất để sửa chữa dựa theo sự tương đồng chỗ đứt gãy sợi kép do nucleaza, gây ra sự mất đoạn đặc hiệu trình tự cho cần được tạo ra ở locut nhiễm sắc thể nội sinh (ví dụ, vùng tăng cường BCL11A) hoặc, ngoài ra (hoặc thêm vào đó), các dạng alen mới (ví dụ, các đột biến điểm cắt bỏ vị trí liên kết yếu tố phiên mã) của locut nội sinh cần được tạo ra. Theo một số khía cạnh, axit nucleic cho này là oligonucleotit, trong đó sự tích hợp này dẫn đến hiện tượng chỉnh sửa gen, hoặc sự mất đoạn hướng đích.

Theo các khía cạnh khác, nucleaza và/hoặc trình tự cho này được phân phối bằng các phương pháp chuyển gen virus và/hoặc không phải virus. Trong các phương án được ưu tiên, trình tự cho này được phân phối cho tế bào bằng virus liên hợp adeno (AAV). Trong một số trường hợp, AAV này chứa các LTR mà thuộc typ huyết thanh khác loại so với typ huyết thanh capsit.

Theo một số khía cạnh, các đột biến mất đoạn chứa các vùng nằm trong các vùng quá miễn ADNseI của trình tự tăng cường (ví dụ, vùng +58 của trình tự tăng cường BCL11A) được tạo ra bằng cách sử dụng một hoặc nhiều nucleaza như được thể hiện trong Bảng 1. Các đột biến mất đoạn này có thể chứa từ khoảng 1 nucleotit đến khoảng 551 nucleotit. Do đó, các đột biến mất đoạn này có thể chứa, 1, 5, 10, 15, 20, 25, 30, 40, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550 nucleotit, hoặc giá trị bất kỳ nằm trong khoảng này. Trong một số phương án, các đột biến mất đoạn này chứa các vùng liên kết cho một hoặc nhiều yếu tố phiên mã. Trong một số phương án được ưu tiên hơn, các đột biến mất đoạn này chứa vị trí liên kết GATA-1, hoặc vị trí liên kết cho GATA-1 kết hợp với các yếu tố khác.

Trong một số phương án, vùng liên kết ADN của Bảng 1 được dung hợp với vùng chức năng. Một số khía cạnh bao gồm thể dung hợp của các vùng liên kết ADN với các vùng có khả năng điều hòa sự biểu hiện của gen. Trong một số phương án, các protein dung hợp này chứa vùng liên kết ADN của Bảng 1 được dung hợp với vùng điều hòa biểu hiện gen, tại đó vùng điều hòa này ức chế sự biểu hiện gen.

Trong một số phương án, các tế bào HSC/PC được cho tiếp xúc với các nucleaza và/hoặc protein liên kết ADN theo sáng chế (nghĩa là, các ZFP như được thể hiện trong Bảng 1). Trong một số phương án, các nucleaza và/hoặc protein liên kết ADN này được phân phối dưới dạng các axit nucleic và trong các phương án khác, chúng được phân phối dưới dạng các protein. Trong một số phương án, axit nucleic là các mARN mã hóa các nucleaza và/hoặc protein liên kết ADN, và trong các phương án khác nữa, các mARN này có thể được bảo vệ. Trong một số phương án, mARN này có thể được biến đổi hóa học, có thể chứa mũ ARCA và/hoặc có thể chứa hỗn hợp của các nucleotit biến đổi hoặc chưa biến đổi. Các tế bào hoặc dòng tế bào có nguồn gốc từ các tế bào này cũng được đề xuất, bao gồm các tế bào biệt hóa một phần hoặc hoàn toàn.

Theo một số khía cạnh, HSC/PC được cho tiếp xúc với các nucleaza và/hoặc protein liên kết ADN theo sáng chế *ex vivo*, sau khi gạn tách HSC/PC từ đối tượng, hoặc sau khi tinh chế từ tủy xương thu được. Trong một số phương án, các nucleaza được mô tả ở đây gây ra các biến đổi bên trong vùng tăng cường BCL11A, ví dụ tạo ra tế bào được biến đổi di truyền mà khác biệt về cấu trúc và/hoặc chức năng với tế bào kiểu đại và/hoặc tế bào được biến đổi khác (ví dụ, được biến đổi bởi nucleaza). Trong các phương án khác nữa,

HSC/PC chứa các biến đổi vùng tăng cường BCL11A được đưa lại vào đối tượng. Trong một số trường hợp, HSC/PC chứa các biến đổi vùng tăng cường BCL11A được nuôi cấy trước khi đưa vào. Theo các khía cạnh khác, các HSC/PC đã biến đổi di truyền được đưa vào đối tượng trong cấy ghép tủy xương, trong đó HSC/PC ghép vào, biệt hóa và trưởng thành *in vivo*. Trong một số phương án, HSC/PC được phân lập từ đối tượng này sau khi huy động do G-CSF và/hoặc plerixafor, và trong các phương án khác, các tế bào này được phân lập từ tủy xương người hoặc dây rốn người. Theo một số khía cạnh, đối tượng này được điều trị cho quy trình phá tủy nhẹ trước khi đưa mô ghép chứa HSC/PC biến đổi vào, trong khi theo các khía cạnh khác, đối tượng này được điều trị bằng liệu pháp hỗ trợ phá tủy mạnh. Trong một số phương án, các phương pháp và chế phẩm theo sáng chế được sử dụng để điều trị hoặc phòng ngừa bệnh huyết sắc tố. Theo một số khía cạnh, bệnh huyết sắc tố này là bệnh tan máu bẩm sinh beta, trong khi theo các khía cạnh khác, bệnh huyết sắc tố này là bệnh hồng cầu hình liềm.

Trong một số phương án, HSC/PC được tiếp xúc thêm với phân tử cho. Trong một số phương án, phân tử cho này được phân phối bởi vector virus. Phân tử cho này có thể chứa một hoặc nhiều trình tự mã hóa polypeptit chức năng (ví dụ, cADN hoặc một đoạn của nó), có hoặc không có trình tự khởi động. Các trình tự bổ sung (các trình tự mã hóa hoặc không mã hóa) có thể được bao gồm khi phân tử cho được sử dụng để bất hoạt, bao gồm nhưng không giới hạn ở, các trình tự mã hóa peptit 2A, vị trí SA, IRES, v.v...

Theo một khía cạnh, các phương pháp và chế phẩm theo sáng chế bao gồm các phương pháp cho HSC/PC tiếp xúc trên *in vivo*. Các nucleaza và/hoặc protein liên kết ADN được phân phối cho HSC/PC *in situ* bằng các phương pháp đã biết trong lĩnh vực này. Trong một số phương án, các nucleaza và/hoặc protein liên kết ADN theo sáng chế chứa hạt virus mà được sử dụng cho đối tượng cần được điều trị, trong khi đó trong các phương án khác, các nucleaza và/hoặc protein liên kết ADN chứa hạt nano (ví dụ, liposom). Trong một số phương án, các hạt virus và/hoặc hạt nano được phân phối cho cơ quan (ví dụ, tủy xương) có HSC/PC trong đó.

Theo một khía cạnh khác, được mô tả ở đây là các phương pháp tích hợp axit nucleic cho vào bộ gen của tế bào bằng cơ chế độc lập tương đồng. Các phương pháp này bao gồm tạo ra đứt gãy sợi kép (DSB) trong bộ gen của tế bào và phân cắt phân tử cho bằng cách sử dụng nucleaza như được mô tả ở đây, sao cho axit nucleic cho này được tích

hợp ở vị trí của DSB. Trong một số phương án, axit nucleic cho này được tích hợp bằng các phương pháp phụ thuộc không tương đồng (ví dụ, NHEJ). Như đã chú ý trên đây, khi phân cắt *in vivo*, các trình tự cho có thể được tích hợp theo cách hướng đích vào bộ gen của tế bào ở vị trí của DSB. Trình tự cho này có thể bao gồm một hoặc nhiều vị trí đích giống nhau cho một hoặc nhiều nucleaza được sử dụng để tạo ra DSB. Do đó, trình tự cho này có thể được phân cắt bởi một hoặc nhiều nucleaza giống nhau được sử dụng để phân cắt gen nội sinh mà sự tích hợp vào đó được mong muốn. Trong một số phương án, trình tự cho này chứa các vị trí đích nucleaza khác nhau từ các nucleaza được sử dụng để cảm ứng DSB. DSB trong bộ gen của tế bào đích có thể được tạo ra bằng cơ chế bất kỳ. Trong một số phương án, DSB này được tạo ra bằng một hoặc nhiều nucleaza ngón tay kềm (ZFN), các protein dung hợp chứa vùng liên kết ngón tay kềm, mà được thao tác di truyền để liên kết trình tự trong vùng cần quan tâm, và vùng phân cắt hoặc nửa vùng phân cắt.

Theo một khía cạnh, trình tự cho này có thể mã hóa protein điều hòa cần quan tâm (ví dụ, ZFP TF, TALE TF hoặc CRISPR/Cas TF) mà liên kết với và/hoặc điều hòa sự biểu hiện của gen cần quan tâm. Trong một phương án, các protein điều hòa này liên kết với trình tự ADN và ngăn chặn sự liên kết của các yếu tố điều hòa khác. Trong một phương án khác, sự liên kết của protein điều hòa này có thể điều biến (nghĩa là, cảm ứng hoặc ức chế) sự biểu hiện của ADN đích.

Trong một số phương án, động vật và/hoặc tế bào HSC/PC chuyển gen chứa gen chuyển mã hóa gen người. Trong một số trường hợp, động vật chuyển gen có sự bất hoạt ở locut nội sinh tương ứng với gen chuyển ngoại sinh, nhờ đó cho phép phát triển hệ *in vivo*, trong đó protein người có thể được nghiên cứu khi phân lập. Các mô hình chuyển gen này có thể được sử dụng với mục đích sàng lọc để xác định các phân tử nhỏ hoặc các phân tử sinh học lớn hoặc các thực thể khác mà có thể tương tác với hoặc biến đổi protein người cần quan tâm. Theo một số khía cạnh, gen chuyển này được tích hợp vào locut đã chọn (ví dụ, cảng an toàn) trong tế bào gốc (ví dụ, tế bào gốc phôi, tế bào gốc vạn năng được cảm ứng, tế bào gốc tạo máu, v.v...) hoặc phôi động vật thu được bằng phương pháp bất kỳ trong các phương pháp được mô tả ở đây, và sau đó phôi này được cấy sao cho động vật sống được sinh ra. Động vật này sau đó được nuôi đến khi thành thực sinh dục và cho sinh sản trong đó ít nhất một số động vật con mang trình tự gen nội sinh đã được chỉnh sửa hoặc gen chuyển đã tích hợp.

Theo khía cạnh khác, được đề xuất ở đây là phương pháp biến đổi sự biểu hiện gen (ví dụ, BCL11A và/hoặc gen globin) trong tế bào, phương pháp này bao gồm: đưa vào tế bào một hoặc nhiều nucleaza như được mô tả ở đây (được thể hiện trong Bảng 1), trong các điều kiện sao cho một hoặc nhiều protein được biểu hiện và sự biểu hiện của gen này được biến đổi. Trong một số phương án, sự biểu hiện của gen globin (ví dụ, gamma globin hoặc beta globin) được biến đổi (ví dụ, được tăng lên). Bất kỳ trong số các phương pháp được mô tả ở đây có thể bao gồm thêm bước tích hợp trình tự cho (ví dụ, gen chuyển hoặc một đoạn của nó dưới sự điều khiển của trình tự khởi động ngoại sinh hoặc nội sinh) vào bộ gen của tế bào, ví dụ tích hợp trình tự cho tại hoặc gần vị trí phân cắt nucleaza trong gen BCL11A. Trình tự cho này được đưa vào tế bào bằng cách sử dụng vector virus, dưới dạng oligonucleotit và/hoặc trên plasmit. Tế bào mà trong đó sự biểu hiện gen được biến đổi có thể là, ví dụ, tế bào tiền thân của tế bào hồng cầu (RBC) và/hoặc tế bào gốc tạo máu (ví dụ, tế bào CD34+).

Trong các phương án khác, được đề xuất ở đây là phương pháp sản sinh tế bào biến đổi gen chứa biến đổi gen bên trong trình tự tăng cường BCL11A nội sinh (biến đổi đối với trình tự nucleotit của trình tự tăng cường BCL11A), phương pháp này bao gồm các bước: a) cho tế bào tiếp xúc với polynucleotit (ví dụ, ADN hoặc mARN) mã hóa nucleaza ngón tay kềm chứa 4, 5, hoặc 6 vùng ngón tay kềm, trong đó mỗi vùng ngón tay kềm này chứa vùng xoắn nhận biết theo trật tự được thể hiện trong một hàng duy nhất của Bảng 1; b) đưa tế bào này vào các điều kiện có lợi cho việc biểu hiện protein ngón tay kềm từ polynucleotit này; và c) biến đổi trình tự tăng cường BCL11A nội sinh có protein ngón tay kềm được biểu hiện đủ để tạo ra tế bào biến đổi di truyền. Trong một số phương án, các tế bào này được kích thích bằng ít nhất một xytokin (ví dụ, trước bước (a)). Polynucleotit này có thể được cho tiếp xúc với tế bào này bằng cách sử dụng phương pháp thích hợp bất kỳ, bao gồm nhưng không giới hạn ở, bằng cách chuyển nạp, bằng cách sử dụng vector không phải virus, bằng cách sử dụng vector virus, bằng phương pháp hóa học hoặc bằng cách phơi nhiễm với điện trường (ví dụ, kỹ thuật xung điện).

Các tế bào chứa một hoặc tổ hợp của các biến đổi gen được mô tả ở đây cũng được đề xuất, bao gồm các tế bào có nguồn gốc từ các tế bào được sản sinh bằng các phương pháp được mô tả ở đây.

Cũng được đề xuất là phương pháp điều trị cho bệnh nhân cần tăng sự biểu hiện gen globin, phương pháp này bao gồm sử dụng cho bệnh nhân được phẩm (tế bào, protein và/hoặc polynucleotit được biến đổi di truyền) như được mô tả ở đây với lượng đủ để làm tăng sự biểu hiện gen globin ở bệnh nhân. Trong một số phương án, bệnh nhân này được biết là mắc, nghi ngờ mắc, hoặc có nguy cơ phát triển bệnh tan máu bẩm sinh hoặc bệnh hồng cầu hình liềm.

Kit, chứa axit nucleic, protein và/hoặc tế bào được biến đổi di truyền theo sáng chế, cũng được đề xuất. Kit này có thể chứa các axit nucleic mã hóa các nucleaza, (ví dụ, các phân tử ARN hoặc các gen mã hóa ZFN, TALEN hoặc hệ CRISPR/Cas có trong vector biểu hiện thích hợp), hoặc các phần phân ước của các protein nucleaza, các phân tử cho, các gen cải biến gốc thích hợp, các tế bào, các gen đệm, và/hoặc các lệnh (ví dụ, để thực hiện các phương pháp theo sáng chế) và tương tự. Sáng chế bao gồm, nhưng không giới hạn ở, tế bào được biến đổi di truyền (ví dụ, tế bào gốc như tế bào gốc tạo máu (CD34+) hoặc tế bào tiền thân RBC) chứa ít nhất một biến đổi gen được tạo ra bởi nucleaza (ví dụ, như được thể hiện thành một hàng duy nhất của Bảng 1), trong đó biến đổi gen này nằm trong trình tự tăng cường BCL11A nội sinh, và trong đó biến đổi gen này được chọn từ nhóm bao gồm đột biến gắn xen, đột biến mất đoạn và tổ hợp của chúng và bao gồm biến đổi tại hoặc gần bất kỳ trong số SEQ ID NO:1 và SEQ ID NO:12. Trong một số phương án, tế bào này là tế bào biệt hóa được biến đổi di truyền có nguồn gốc từ tế bào gốc như được mô tả ở đây (ví dụ, RBC có nguồn gốc từ tế bào gốc tạo máu hoặc tế bào tiền thân RBC).

Nucleaza này có thể chứa ít nhất một nucleaza ngón tay kẽm (ZFN) (ví dụ, như được thể hiện trong Bảng 1) và/hoặc ít nhất một TALEN và (các) nucleaza này có thể được đưa vào tế bào ở dạng protein và/hoặc dưới dạng polynucleotit mã hóa (các) nucleaza. Trong một số phương án, biến đổi gen này bao gồm đột biến gắn xen mà bao gồm sự tích hợp polynucleotit cho mã hóa gen chuyển. Cũng được đề xuất là được phẩm chứa một hoặc nhiều tế bào được biến đổi di truyền như được mô tả ở đây.

Cũng được đề xuất là protein liên kết ADN chứa protein ngón tay kẽm có 4, 5 hoặc 6 vùng ngón tay kẽm chứa vùng xoắn nhận biết, trong đó các protein ngón tay kẽm này chứa các vùng xoắn nhận biết theo trật tự được thể hiện trong một hàng duy nhất của Bảng 1. Cũng được đề xuất là protein TALE chứa nhiều trình tự lặp mà liên kết với trình tự chứa một phần (ví dụ, ít nhất là 4, 5, 6 hoặc nhiều) cặp bazơ của vị trí đích được thể hiện trong

Bảng 1. Protein dung hợp chứa protein ngón tay kẽm hoặc protein TALE như được mô tả ở đây và vùng cắt hoặc nửa vùng cắt kiểu đại hoặc được thao tác di truyền cũng được đề xuất, cũng như các polynucleotit mã hóa các protein (ZFP, TALE, ZFN, TALEN) như được mô tả ở đây. Các tế bào (ví dụ, tế bào gốc đã phân lập như tế bào gốc tạo máu (CD34+)) chứa một hoặc nhiều polynucleotit và/hoặc protein như được mô tả ở đây cũng được đề xuất. Cũng được đề xuất là các kit chứa một hoặc nhiều protein, polynucleotit và/hoặc tế bào như được mô tả ở đây.

Phương pháp biến đổi sự biểu hiện gen globin trong tế bào (ví dụ, tế bào tiền thân RBC và/hoặc tế bào gốc tạo máu) cũng được mô tả, phương pháp này bao gồm: đưa vào tế bào một hoặc nhiều polynucleotit mã hóa một hoặc nhiều nucleaza như được mô tả ở đây, trong các điều kiện sao cho một hoặc nhiều protein được biểu hiện và sự biểu hiện của gen globin (ví dụ, gamma và/hoặc beta globin) được biến đổi (ví dụ, được làm tăng). Trong một số phương án, các phương pháp này bao gồm thêm bước tích hợp trình tự cho vào bộ gen của tế bào, ví dụ bằng cách sử dụng vector virus, dưới dạng oligonucleotit hoặc trên plasmid. Trình tự cho này có thể chứa gen chuyển dưới sự điều khiển của trình tự khởi động nội sinh hoặc ngoại sinh.

Cũng được đề xuất là phương pháp sản sinh tế bào biến đổi gen chứa biến đổi gen nằm trong trình tự tăng cường BCL11A nội sinh (ví dụ, vị trí đích như được thể hiện trong Bảng 1), phương pháp này bao gồm các bước: (a) cho tế bào tiếp xúc với polynucleotit mã hóa protein dung hợp chứa nucleaza ngón tay kẽm có 4, 5, hoặc 6 vùng ngón tay kẽm trong đó mỗi vùng ngón tay kẽm này chứa vùng xoắn nhận biết theo trật tự được thể hiện trong một hàng duy nhất của Bảng 1; (b) đưa tế bào này vào các điều kiện có lợi cho việc biểu hiện protein dung hợp từ polynucleotit; và (c) biến đổi trình tự tăng cường BCL11A nội sinh có protein dung hợp đã biểu hiện đủ để tạo ra tế bào biến đổi gen. Trong một số phương án, phương pháp này bao gồm thêm bước kích thích tế bào bằng ít nhất một xytokin. (Các) polynucleotit có thể được phân phối vào trong tế bào, ví dụ bằng cách sử dụng hệ phân phối không phải virus, hệ phân phối virus, và/hoặc phương tiện phân phối và có thể bao gồm bước đưa các tế bào này vào điện trường.

Các phương pháp điều trị cho bệnh nhân cần tăng sự biểu hiện gen globin (ví dụ, bệnh nhân được biết là mắc, bị nghi ngờ mắc, hoặc có nguy cơ phát triển bệnh globin như bệnh tan máu bẩm sinh (ví dụ, bệnh tan máu bẩm sinh β) hoặc bệnh hồng cầu hình liềm

cũng được đề xuất, phương pháp này bao gồm sử dụng cho bệnh nhân này được phẩm như được mô tả ở đây (ví dụ, protein, polynucleotit và/hoặc tế bào) với lượng đủ để làm tăng sự biểu hiện gen globin ở bệnh nhân.

Các khía cạnh này và các khía cạnh khác sẽ là rõ ràng với người có kiến thức trung bình trong lĩnh vực này nhờ toàn bộ phần mô tả.

Mô tả vắn tắt các hình vẽ

Fig.1 là đồ thị mô tả tỉ lệ tương đối của sự biểu hiện gamma globin người (HBG) với sự biểu hiện beta globin người (HBB) trong các tế bào hồng cầu thu được từ các tế bào CD34+ đã được chỉnh sửa với các cặp ZFN đặc hiệu BCL11a được chỉ ra.

Fig.2A và 2B là các đồ thị mô tả hoạt tính của hai cặp ZFN đặc hiệu BCL11a trong tế bào CD34+ được phân lập từ máu ngoại vi (PB). Các tế bào được chuyển nhiễm bằng cách sử dụng thiết bị xung điện BTX. Tỉ lệ % đột biến gắn xen hoặc làm mất bazơ phát hiện được (phép đo hoạt tính NHEJ phát hiện được) cho mỗi điều kiện được chỉ ra dưới các đồ thị của Fig.2A (giới hạn lượng mARN đầu vào từ 0,5 đến 4 μ g) và Fig.2B (giới hạn lượng mARN đầu vào từ 2,0 đến 8,0 μ g).

Fig.3 là đồ thị mô tả sự biểu hiện của gamma globin người (HBG) dưới dạng tỉ lệ tương đối của HBG với beta globin người (HBB) sau khi biệt hóa hồng cầu của tế bào PB CD34+ đã chỉnh sửa được chỉ ra trong Fig.2B. Mẫu mARN duy nhất, trong đó các ZFN được mã hóa trên cùng một phân tử mARN nhưng được tách rời bởi trình tự peptit tự cắt 2a (được gọi là “2a”), được so sánh với việc sử dụng hai mARN trong đó mỗi mARN mã hóa một phần tử trong cặp ZFN (được ký hiệu là “sep” cho sự riêng biệt).

Fig.4A và Fig.4B là các đồ thị mô tả hoạt tính của hai cặp ZFN đặc hiệu BCL11a trong các tế bào CD34+ được phân lập từ tủy xương (BM). Các tế bào được chuyển nhiễm bằng cách sử dụng thiết bị xung điện BTX. Tỉ lệ % các đột biến gắn xen hoặc làm mất bazơ được phát hiện (phép đo hoạt tính NHEJ phát hiện được) cho mỗi điều kiện được thể hiện dưới các đồ thị cho Fig.4A (giới hạn lượng mARN đầu vào từ 2,0 đến 8,0 μ g). Fig.4B mô tả hoạt tính của một cặp ZFN, trong đó các ZFN này được cung cấp dưới dạng một mẫu mARN duy nhất có trình tự peptit tự cắt 2a phân tách các trình tự mã hóa mỗi ZFN hoặc khi hai ZFN được cung cấp trên các mARN riêng biệt.

Fig.5A và 5B mô tả hoạt tính của các cặp ZFN trong các tế bào CD34+ có nguồn gốc từ PB bằng cách sử dụng thiết bị xung điện Maxcyte. Tỷ lệ % các đột biến gắn xen và làm mất bazơ phát hiện được (phép đo hoạt tính NHEJ phát hiện được) cho mỗi điều kiện được thể hiện dưới các đồ thị cho Fig.5A và Fig.5B. Fig.5A mô tả sự so sánh giữa hai cặp ZFN, và Fig.5B mô tả hoạt tính của các cặp ZFN khi các ZFN được cung cấp dưới dạng một mẫu mARN duy nhất với trình tự peptit tự cắt 2a phân tách các trình tự mã hóa mỗi ZFN hoặc khi hai ZFN được cung cấp trên các mARN riêng biệt.

Fig.6 mô tả đồ thị thể hiện hoạt tính quy mô lớn của cặp A (SBS51446/51536) và cặp B (SBS51857/51949) trong các tế bào CD34+ có nguồn gốc tủy xương. Các mARN mã hóa các cặp ZFN được cung cấp dưới dạng các mARN duy nhất, trong đó các trình tự mã hóa mỗi nửa của cặp ZFN được phân tách bởi trình tự tự cắt 2a, hoặc dưới dạng các mARN riêng biệt mã hóa mỗi ZFN. Hoạt tính được thể hiện dưới dạng tỷ lệ % các đột biến mất đoạn và thêm bazơ phát hiện được.

Fig.7 thể hiện đồ thị mô tả lượng biểu hiện HBG và HBB tương đối được phát hiện sau 14 ngày biệt hóa sau khi chỉnh sửa gen quy mô lớn được thể hiện trong Fig.6. Như trước đây, các mẫu được thử nghiệm dưới dạng các mARN duy nhất mã hóa cả hai ZFN, hoặc dưới dạng các mARN riêng biệt. Lượng đột biến mất đoạn hoặc thêm bazơ được phát hiện vào ngày 0 sau khi biệt hóa được thể hiện ở phần dưới cùng, và chứng minh là dấu vết hoạt tính gây đột biến mất đoạn hoặc thêm bazơ với lượng HGB được biểu hiện.

Fig.8 là đồ thị mô tả tỷ lệ phần trăm của các đột biến mất đoạn và thêm bazơ phát hiện được khi chỉnh sửa quy mô lớn các tế bào CD34+ từ tủy xương đã xử lý với cặp B, dưới dạng các mARN duy nhất hoặc các mARN riêng biệt như được mô tả trên đây.

Fig.9A và Fig.9B là các đồ thị mô tả tỷ lệ phần trăm các đột biến mất đoạn hoặc thêm bazơ khi chỉnh sửa quy mô lớn các tế bào CD34+ từ tủy xương được xử lý bằng cặp B, dưới dạng các mARN duy nhất hoặc các mARN riêng biệt như được mô tả trên đây.

Fig.10A và 10B là các bảng mô tả các kết quả của phân tích sai mục tiêu cho Cặp A (Fig.10A) và Cặp B (Fig.10B).

Fig.11A và 11B là các đồ thị thể hiện phân tích qPCR thời gian thực RT sự biệt hóa *in vitro* trong thí nghiệm 1 (Fig.11A) và thí nghiệm 2 (Fig.11B) sử dụng tế bào bệnh nhân

và tế bào kiểu dại (wt) được xử lý với SB ZFN mARN. Các đồ thị này thể hiện tỉ lệ tương đối của mARN gamma globin so với mARN alpha globin.

Fig.12A và 12B là các đồ thị thể hiện tỉ lệ của gamma globin so với alpha globin trong thí nghiệm 1 (Fig.12A) và thí nghiệm 2 (Fig.12B). Đối với các giá trị gamma globin, các giá trị của pic Agamma và Ggamma và, nếu có thể, pic Agamma T được cộng lại.

Fig.13 thể hiện đồ thị về tỉ lệ protein tương tự gamma/beta được lập đồ thị theo trạng thái alen trong các khuẩn lạc đơn lẻ được phân tích. Số liệu này được sắp xếp theo lớp kiểu gen (“+” đối với alen không biến đổi, “-” đối với alen được chỉnh sửa; “+/+” đối với kiểu dại; “+/-” đối với biến đổi monoalen; và “-/-” đối với biến đổi bialen).

Mô tả chi tiết sáng chế

Được mô tả ở đây là các chế phẩm và phương pháp để thao tác di truyền bộ gen để điều biến sự biểu hiện BCL11A và/hoặc gamma globin và để điều trị và/hoặc phòng ngừa các bệnh huyết sắc tố. Cụ thể là, các nucleaza chứa các ZFP có các vùng xoắn nhận biết như được thể hiện trong một hàng duy nhất của Bảng 1 thu được một cách hữu hiệu trong HSC/PC và gây ra sự thay đổi về sự biểu hiện gamma globin tương đối trong quá trình tạo hồng cầu sau đó. Việc điều biến sự biểu hiện BCL11A và gamma globin này là đặc biệt hữu ích để điều trị các bệnh huyết sắc tố (ví dụ, bệnh tan máu bẩm sinh beta, bệnh hồng cầu hình liềm) trong đó có sự biểu hiện không đủ của beta globin hoặc sự biểu hiện của dạng đột biến của beta-globin. Bằng cách sử dụng các phương pháp và chế phẩm theo sáng chế, các biến chứng và di chứng liên quan đến bệnh bị gây ra bởi beta globin bất thường có thể được cải thiện bằng cách biến đổi sự biểu hiện của gamma globin trong các tế bào tiền thân của hồng cầu.

Tổng quan

Việc thực hành các phương pháp này, cũng như việc bào chế và sử dụng các chế phẩm được mô tả ở đây sử dụng, trừ khi có quy định khác, các kỹ thuật thông thường trong sinh học phân tử, hóa sinh, phân tích và cấu trúc sợi nhiễm sắc, hóa học sử dụng máy điện toán, nuôi cấy tế bào, ADN tái tổ hợp và các lĩnh vực liên quan như nằm trong kiến thức trung bình thuộc lĩnh vực này. Các kỹ thuật này được giải thích đầy đủ trong tài liệu. Xem, ví dụ, Sambrook và đồng tác giả. MOLECULAR CLONING: A LABORATORY MANUAL, Xuất bản lần thứ hai, Cold Spring Harbor Laboratory Press, 1989 và xuất bản

lần thứ ba, 2001; Ausubel và đồng tác giả., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, 1987 và cập nhật định kỳ; seri METHODS IN ENZYMOLOGY, Academic Press, San Diego; Wolffe, CHROMATIN STRUCTURE AND FUNCTION, xuất bản lần thứ ba, Academic Press, San Diego, 1998; METHODS IN ENZYMOLOGY, Vol. 304, “Chromatin” (P.M. Wassarman và A. P. Wolffe, eds.), Academic Press, San Diego, 1999; và METHODS IN MOLECULAR BIOLOGY, Vol. 119, “Chromatin Protocols” (P.B. Becker, ed.) Humana Press, Totowa, 1999.

Định nghĩa

Các thuật ngữ “axit nucleic”, “polynucleotit”, và “oligonucleotit” được sử dụng thay thế lẫn nhau và dùng để chỉ polyme deoxyribonucleotit hoặc ribonucleotit, ở cấu hình thẳng hoặc vòng, và ở dạng chuỗi đơn hoặc kép. Với các mục đích của sáng chế, các thuật ngữ này không được hiểu là làm giới hạn độ dài của polyme. Các thuật ngữ này có thể bao gồm các phân tử tương tự đã biết chứa nucleotit tự nhiên, cũng như nucleotit được biến đổi về gốc bazơ, đường và/hoặc gốc phosphat (ví dụ, khung chính phosphorothioat). Nói chung, phân tử tương tự của một nucleotit cụ thể có cùng tính đặc hiệu tạo cặp bazơ; *nghĩa là*, phân tử tương tự của A sẽ tạo cặp bazơ với T.

Các thuật ngữ “polypeptit”, “peptit” và “protein” được sử dụng thay thế lẫn nhau để đề cập đến polyme chứa các gốc axit amin. Thuật ngữ này cũng áp dụng cho các polyme axit amin, trong đó một hoặc nhiều axit amin là các phân tử tương tự hóa học hoặc dẫn xuất biến đổi của các axit amin tự nhiên tương ứng.

“Liên kết” dùng để chỉ tương tác không đồng hóa trị, đặc hiệu trình tự giữa các đại phân tử (ví dụ, giữa protein và axit nucleic). Không phải tất cả các thành phần của tương tác liên kết đều cần phải đặc hiệu trình tự (ví dụ, tiếp xúc với các gốc phosphat trong khung chính ADN), miễn là tương tác này về mặt tổng thể là đặc hiệu trình tự. Các tương tác này thường được đặc trưng bởi hằng số phân ly (K_d) là 10^{-6} M^{-1} hoặc thấp hơn. “Ái lực” dùng để chỉ cường độ liên kết: ái lực liên kết càng tăng khi K_d càng thấp.

“Protein liên kết” là protein có thể liên kết với một phân tử khác. Protein liên kết có thể liên kết với, ví dụ, phân tử ADN (protein liên kết ADN), phân tử ARN (protein liên kết ARN) và/hoặc phân tử protein (protein liên kết protein). Trong trường hợp protein liên kết protein, nó có thể liên kết với chính nó (để tạo ra homodime, homotrime, v.v...) và/hoặc nó có thể liên kết với một hoặc nhiều phân tử của một hoặc nhiều protein khác. Protein liên

kết có thể có nhiều hơn một kiểu hoạt tính liên kết. Ví dụ, các protein ngón tay kẽm có hoạt tính liên kết ADN, liên kết ARN và liên kết protein.

“Protein liên kết ADN ngón tay kẽm” (hoặc vùng liên kết ADN ngón tay kẽm) là protein, hoặc vùng nằm trong một protein lớn hơn, mà liên kết với ADN theo cách đặc hiệu trình tự qua một hoặc nhiều ngón tay kẽm, mà là các vùng có trình tự axit amin nằm trong vùng liên kết có cấu trúc được ổn định nhờ phối trí của ion kẽm. Thuật ngữ protein liên kết ADN ngón tay kẽm thường được viết tắt là protein ngón tay kẽm hoặc ZFP.

“Vùng liên kết TALE ADN” hoặc “TALE” là polypeptit chứa một hoặc nhiều vùng/đơn vị lặp TALE. Vùng lặp tham gia vào liên kết của TALE với trình tự ADN đích cùng gốc của nó. “Đơn vị lặp” đơn (còn được gọi là “trình tự lặp”) thường dài 33-35 axit amin và thể hiện ít nhất là một số tương đồng trình tự với các trình tự lặp TALE khác trong protein TALE tự nhiên.

Ngón tay kẽm và vùng liên kết TALE có thể được “thao tác di truyền” để liên kết với trình tự nucleotit xác định trước, ví dụ bằng cách thao tác di truyền (biến đổi một hoặc nhiều axit amin) vùng xoắn nhận biết của ngón tay kẽm tự nhiên hoặc protein TALE. Do đó, protein liên kết ADN được thao tác di truyền (ngón tay kẽm hoặc TALE) là các protein không xuất hiện tự nhiên. Ví dụ không giới hạn về các phương pháp thao tác di truyền protein liên kết ADN là thiết kế và chọn lọc. Protein liên kết ADN được thiết kế là protein không xuất hiện trong tự nhiên mà có thiết kế/thành phần chủ yếu do các tiêu chuẩn hợp lý. Các tiêu chuẩn hợp lý để thiết kế bao gồm áp dụng các nguyên tắc thể và các thuật toán vi tính để xử lý thông tin trong cơ sở dữ liệu lưu trữ thông tin của các thiết kế ZFP và/hoặc TALE và dữ liệu liên kết hiện hành. Xem, ví dụ. Patent Mỹ số 6.140.081; 6.453.242; 6.534.261 và 8.585.526; cũng xem WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 và WO 03/016496.

Protein ngón tay kẽm hoặc TALE “được chọn lọc” là protein không được tìm thấy trong tự nhiên mà sự sản sinh ra nó chủ yếu là do quy trình thực nghiệm như biểu hiện thể thực khuẩn, bắt giữ tương tác hoặc chọn lọc lai. Xem ví dụ. Patent Mỹ số 5.789.538; 5.925.523; 6.007.988; 6.013.453; 6.200.759; 8.586.526; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970 WO 01/88197, WO 02/099084.

“TtAgo” là protein Argonaute của tế bào tiền nhân được cho là tham gia vào việc bất hoạt gen. TtAgo thu được từ vi khuẩn *Thermus thermophilus*. Xem, ví dụ, Swarts và đồng tác giả, *ibid*, G. Sheng và đồng tác giả., (2013) *Proc. Natl. Acad. Sci. U.S.A.* 111, 652). “Hệ TtAgo” là tất cả các thành phần cần thiết bao gồm, ví dụ, các ADN hướng dẫn để phân cắt bằng enzym TtAgo.

“Tái tổ hợp” dùng để chỉ quá trình trao đổi thông tin di truyền giữa hai polynucleotit, bao gồm nhưng không giới hạn ở, bất giữ thể cho bằng cách nối đầu cùng không tương đồng (NHEJ) và tái tổ hợp tương đồng. Với các mục đích của phần mô tả, “tái tổ hợp tương đồng (HR)” dùng để chỉ dạng đặc hiệu của quá trình trao đổi xảy ra, ví dụ, trong quá trình sửa chữa các đoạn đứt gãy sợi kép trong tế bào bằng các cơ chế sửa chữa dựa theo sự tương đồng. Quá trình này đòi hỏi sự tương đồng trình tự nucleotit, sử dụng phân tử “cho” để sửa khuôn của phân tử “đích” (nghĩa là, phân tử bị đứt gãy sợi kép), và được gọi khác nhau là “chuyển hóa gen không trao đổi chéo” hoặc “chuyển hóa gen bó ngắn” vì nó dẫn đến sự truyền thông tin di truyền từ thể cho đến đích. Không mong muốn bị ràng buộc bởi lý thuyết bất kỳ, sự truyền này có thể bao gồm sửa chữa lệch đôi của ADN sợi kép dị hợp mà tạo ra giữa đích đứt gãy và trình tự cho, và/hoặc “ủ sợi phụ thuộc quá trình tổng hợp” trong đó trình tự cho này được sử dụng để tái tổng hợp thông tin di truyền mà sẽ trở thành một phần của đích, và/hoặc các quá trình liên quan. HR đặc hiệu thường gây ra sự biến đổi của trình tự của phân tử đích sao cho một phần hoặc toàn bộ trình tự của polynucleotit cho được kết hợp vào polynucleotit đích.

Trong các phương pháp của sáng chế, một hoặc nhiều nucleaza hướng đích như được mô tả ở đây tạo ra vị trí đứt gãy sợi kép (DSB) trong trình tự đích (ví dụ, sợi nhiễm sắc tế bào) ở vị trí xác định trước. DSB có thể gây ra đột biến mất đoạn và/hoặc đột biến gắn xen bằng cơ chế sửa chữa dựa vào sự tương đồng hoặc bằng cơ chế sửa chữa dựa vào sự không tương đồng. Các đột biến mất đoạn có thể chứa số lượng cặp bazơ bất kỳ. Tương tự, các đột biến gắn xen có thể chứa số lượng cặp bazơ bất kỳ bao gồm, ví dụ, sự tích hợp polynucleotit “cho”, tùy ý có sự tương đồng với trình tự nucleotit trong vùng đứt gãy. Trình tự cho này có thể được tích hợp vật lý hoặc, ngoài ra, polynucleotit cho này được sử dụng làm khuôn để sửa vị trí đứt gãy bằng kỹ thuật tái tổ hợp tương đồng, gây ra sự đưa vào toàn bộ hoặc một phần của trình tự nucleotit như trong trình tự cho vào sợi nhiễm sắc tế bào. Do đó, trình tự thứ nhất trong sợi nhiễm sắc tế bào có thể được biến đổi và, trong một số

phương án, có thể được chuyển thành trình tự có mặt trong polynucleotit cho. Do đó, việc sử dụng các thuật ngữ “thay thế” hoặc “sự thay thế” có thể được hiểu là thể hiện sự thay thế một trình tự nucleotit bằng một trình tự khác, (nghĩa là, thay thế một trình tự theo nghĩa thông tin), và không nhất thiết đòi hỏi phải thay thế vật lý hoặc hóa học một polynucleotit bằng một polynucleotit khác.

Trong một phương pháp bất kỳ trong số các phương pháp được mô tả ở đây, các cặp protein ngón tay kềm bổ sung hoặc TALEN có thể được sử dụng để cắt sợi kép bổ sung của vị trí đích bổ sung bên trong tế bào này.

Bất kỳ trong số các phương pháp được mô tả ở đây có thể được sử dụng để gắn xen trình tự cho có kích thước bất kỳ và/hoặc bất hoạt một phần hoặc hoàn toàn một hoặc nhiều trình tự đích trong một tế bào bằng cách tích hợp đích trình tự cho mà phá vỡ sự biểu hiện của (các) gen cần quan tâm. Các dòng tế bào có các gen được bất hoạt một phần hoặc hoàn toàn cũng được đề xuất.

Trong một phương pháp bất kỳ trong các phương pháp được mô tả ở đây, trình tự nucleotit ngoại sinh (“trình tự cho” hoặc “gen chuyển”) có thể chứa các trình tự mà tương đồng, nhưng không đồng nhất, với các trình tự bộ gen trong vùng cần quan tâm, do đó kích thích sự tái tổ hợp tương đồng để gắn xen trình tự không đồng nhất vào vùng cần quan tâm. Do đó, trong một số phương án, các phần của trình tự cho mà tương đồng với các trình tự trong vùng cần quan tâm thể hiện sự đồng nhất trình tự khoảng từ 80 đến 99% (hoặc số nguyên bất kỳ trong khoảng này) với trình tự bộ gen bị thay thế. Trong các phương án khác, sự tương đồng giữa trình tự cho và trình tự bộ gen là cao hơn 99%, ví dụ nếu chỉ 1 nucleotit khác nhau giữa trình tự cho và trình tự bộ gen trong số hơn 100 cặp bazơ liên kề. Trong một số trường hợp, phần không tương đồng của trình tự cho có thể chứa các trình tự không có mặt trong vùng cần quan tâm, sao cho các trình tự mới được đưa vào vùng cần quan tâm. Trong các trường hợp này, trình tự không tương đồng thường được gắn bên sườn bởi các trình tự gồm 50-1.000 cặp bazơ (hoặc giá trị nguyên bất kỳ trong khoảng này) hoặc số lượng cặp bazơ bất kỳ lớn hơn 1.000, mà tương đồng hoặc đồng nhất với các trình tự trong vùng cần quan tâm. Trong các phương án khác, trình tự cho này là không tương đồng với trình tự thứ nhất, và được gắn xen vào bộ gen bằng cơ chế tái tổ hợp không tương đồng.

“Phân cắt” dùng để chỉ sự phá vỡ khung chính đồng hóa trị của phân tử ADN. Sự phân cắt có thể được khởi động bằng nhiều phương pháp khác nhau bao gồm, nhưng không

giới hạn ở, quy trình thủy phân bằng enzym hoặc hóa học liên kết phosphodiester. Cả sự phân cắt sợi đơn và phân cắt sợi kép đều là có thể, và sự phân cắt sợi kép có thể xảy ra như là kết quả của hai hiện tượng phân cắt sợi đơn riêng biệt. Sự phân cắt ADN có thể tạo thành các đầu bằng hoặc các đầu so le. Trong một số phương án, các polypeptit dung hợp được sử dụng để phân cắt ADN sợi kép hướng đích.

“Nửa vùng phân cắt” là trình tự polypeptit mà, nối với polypeptit thứ hai (giống hoặc khác nhau) tạo ra phức hợp có hoạt tính phân cắt (tốt hơn, nếu là hoạt tính phân cắt sợi kép). Thuật ngữ “nửa vùng phân cắt thứ nhất và thứ hai”; “nửa vùng phân cắt + và -” và “nửa vùng phân cắt trái và phải” được sử dụng thay thế lẫn nhau để đề cập đến cặp các nửa vùng phân cắt đime hóa.

“Nửa vùng phân cắt được thao tác di truyền” là nửa vùng phân cắt được biến đổi để tạo ra các heterodime bắt buộc với nửa vùng phân cắt khác (ví dụ, nửa vùng phân cắt được thao tác di truyền khác). Cũng xem, Công bố Patent Mỹ số 2005/0064474, 20070218528, 20080131962 và 20110201055, được đưa vào đây bằng cách viện dẫn toàn bộ.

Thuật ngữ “trình tự” dùng để chỉ trình tự nucleotit có độ dài bất kỳ, có thể là ADN hoặc ARN; có thể là thẳng, vòng hoặc có nhánh và có thể là sợi đơn hoặc sợi kép. Thuật ngữ “trình tự cho” dùng để chỉ trình tự nucleotit được gắn xen vào bộ gen. Trình tự cho có thể có độ dài bất kỳ, ví dụ độ dài từ 2 đến 100.000.000 nucleotit (hoặc trị số nguyên bất kỳ trong khoảng này hoặc trên khoảng này), tốt hơn, nếu dài từ 100 đến 100.000 nucleotit (hoặc giá trị nguyên bất kỳ trong khoảng này), tốt hơn nữa, nếu dài khoảng từ 2000 đến 20.000 nucleotit (hoặc giá trị bất kỳ nằm trong khoảng này) và thậm chí tốt hơn nữa, nếu khoảng từ 5 đến 15 kb (hoặc giá trị bất kỳ nằm trong khoảng này).

“Sợi nhiễm sắc” là cấu trúc nucleoprotein chứa bộ gen tế bào. Sợi nhiễm sắc tế bào chứa axit nucleic, chủ yếu là ADN, và protein, bao gồm các protein nhiễm sắc thể histon và không histon. Phần lớn sợi nhiễm sắc tế bào nhân chuẩn tồn tại ở dạng các thể nhân, trong đó lõi thể nhân chứa xấp xỉ 150 cặp bazơ của ADN liên hợp với octame chứa mỗi hai trong số các histon H2A, H2B, H3 và H4; và ADN cầu nối (có độ dài thay đổi phụ thuộc vào sinh vật) kéo dài giữa các lõi thể nhân. Phân tử của histon H1 thường được liên hợp với ADN cầu nối này. Với mục đích của phần mô tả này, thuật ngữ “sợi nhiễm sắc” có nghĩa là bao gồm tất cả các loại nucleoprotein tế bào, cả tế bào tiền nhân và tế bào nhân

chuẩn. Sợi nhiễm sắc tế bào bao gồm cả sợi nhiễm sắc của nhiễm sắc thể và sợi nhiễm sắc của thể bổ sung.

“Nhiễm sắc thể” là phức hợp chất nhiễm sắc chứa toàn bộ hoặc một phần của bộ gen tế bào. Bộ gen của tế bào thường được đặc trưng bởi nhiễm sắc thể đồ, mà là tập hợp của tất cả các nhiễm sắc thể chứa bộ gen của tế bào. Bộ gen của tế bào có thể bao gồm một hoặc nhiều nhiễm sắc thể.

“Thể bổ sung” là axit nucleic sao chép, phức hợp nucleoprotein hoặc cấu trúc khác chứa axit nucleic mà không phải là một phần của nhiễm sắc thể đồ của tế bào. Ví dụ về các thể bổ sung bao gồm các plasmid và một số bộ gen virus.

“Vùng tiếp cận được” là vị trí trong sợi nhiễm sắc tế bào mà trong đó vị trí đích có mặt trong axit nucleic có thể được liên kết bởi phân tử ngoại sinh nhận biết vị trí đích. Không mong muốn bị ràng buộc bởi lý thuyết bất kỳ, tin rằng vùng tiếp cận được là một vùng mà không được gói vào trong cấu trúc thể nhân. Cấu trúc riêng của vùng tiếp cận được thường có thể được phát hiện bởi độ nhạy của nó với các đoạn dò hóa học và enzym, ví dụ, các nucleaza.

“Vị trí đích” hoặc “trình tự đích” là trình tự axit nucleic xác định phần axit nucleic mà phân tử liên kết sẽ liên kết vào, tạo ra các điều kiện đủ để liên kết tồn tại.

Phân tử “ngoại sinh” là phân tử mà bình thường không có mặt trong tế bào, nhưng có thể được đưa vào tế bào nhờ một hoặc nhiều phương pháp di truyền, hóa sinh hoặc các phương pháp khác. “Sự có mặt bình thường trong tế bào” được xác định cho một giai đoạn phát triển cụ thể và các điều kiện môi trường cụ thể của tế bào này. Do đó, ví dụ, phân tử mà chỉ có mặt trong quá trình phát triển phôi của cơ là phân tử ngoại sinh đối với tế bào cơ của người trưởng thành. Tương tự, phân tử được tạo ra bằng cách sốc nhiệt là phân tử ngoại sinh đối với tế bào không được sốc nhiệt. Phân tử ngoại sinh có thể bao gồm, ví dụ, dạng hoạt động của phân tử nội sinh bị lỗi hoặc dạng bị lỗi của phân tử nội sinh hoạt động bình thường.

Phân tử ngoại sinh có thể là, ngoài các loại phân tử khác, phân tử nhỏ, như được tạo ra bằng quy trình hóa học tổ hợp, hoặc đại phân tử như protein, axit nucleic, hydrat cacbon, lipid, glycoprotein, lipoprotein, polysaccharit, dẫn xuất được biến đổi bất kỳ của các phân tử trên đây, hoặc phức hợp bất kỳ bao gồm một hoặc nhiều phân tử trên đây. Các axit nucleic

bao gồm ADN và ARN, có thể là sợi đơn hoặc sợi kép; có thể là thẳng, có nhánh hoặc vòng; và có thể có độ dài bất kỳ. Các axit nucleic bao gồm các phân tử có khả năng tạo ra sợi kép, cũng như các axit nucleic tạo ra bộ ba. Xem, ví dụ, Patent Mỹ số 5.176.996 và 5.422.251. Các protein bao gồm, nhưng không giới hạn ở, protein liên kết ADN, yếu tố phiên mã, yếu tố tái cấu trúc sợi nhiễm sắc, protein liên kết ADN được methyl hóa, polymeraza, methylaza, demethylaza, axetylaza, deaxetylaza, kinaza, phosphataza, integraza, recombinaza, ligaza, topoisomeraza, gyraza và helicaza.

Phân tử ngoại sinh có thể cùng loại phân tử với phân tử nội sinh, ví dụ, protein hoặc axit nucleic ngoại sinh. Ví dụ, axit nucleic ngoại sinh có thể bao gồm bộ gen virus gây nhiễm, plasmid hoặc thể bổ sung được đưa vào trong tế bào, hoặc nhiễm sắc thể mà bình thường không có mặt trong tế bào này. Các phương pháp đưa các phân tử ngoại sinh vào trong tế bào đã được biết bởi người có kiến thức trung bình trong lĩnh vực này và bao gồm, nhưng không giới hạn ở, kỹ thuật chuyển qua trung gian lipid (nghĩa là, liposom, bao gồm các lipid trung tính và cation hóa), kỹ thuật xung điện, tiêm trực tiếp, dung hợp tế bào, bắn phá bằng hạt, đồng kết tủa canxi phosphat, kỹ thuật chuyển qua trung gian DEAE-dextran và chuyển qua trung gian vector virus. Phân tử ngoại sinh cũng có thể cùng loại phân tử với phân tử nội sinh nhưng thu được từ các loài khác với loài mà tế bào này thu được từ đó. Ví dụ, trình tự axit nucleic người có thể được đưa vào dòng tế bào thu được đầu tiên từ chuột hoặc hamster. Phương pháp đưa phân tử ngoại sinh vào tế bào thực vật đã được biết đối với người có kiến thức trung bình trong lĩnh vực này và bao gồm, nhưng không giới hạn ở, phương pháp biến nạp thể nguyên sinh, silicon cacbua (ví dụ, WHISKERS™), biến nạp qua trung gian *Agrobacterium*, chuyển qua trung gian lipid (nghĩa là, liposom, bao gồm các lipid trung tính và cation hóa), kỹ thuật xung điện, tiêm trực tiếp, dung hợp tế bào, kỹ thuật bắn phá bằng hạt (ví dụ, sử dụng “súng gen”), đồng kết tủa canxi phosphat, chuyển qua trung gian DEAE-dextran và chuyển qua trung gian virus.

Ngược lại, phân tử “nội sinh” là phân tử mà bình thường có mặt trong một tế bào cụ thể ở một giai đoạn phát triển cụ thể trong các điều kiện môi trường cụ thể. Ví dụ, axit nucleic nội sinh có thể bao gồm nhiễm sắc thể, bộ gen của ty thể, lục lạp hoặc bào quan khác, hoặc axit nucleic thể bổ sung tự nhiên. Các phân tử nội sinh khác có thể bao gồm protein, ví dụ, yếu tố phiên mã và enzym.

Như được sử dụng ở đây, thuật ngữ “sản phẩm của axit nucleic ngoại sinh” bao gồm các sản phẩm polynucleotit và polypeptit, ví dụ, các sản phẩm phiên mã (polynucleotit như ARN) và các sản phẩm dịch mã (polypeptit).

Phân tử “dung hợp” là phân tử trong đó hai hoặc nhiều phân tử tiểu đơn vị được liên kết, tốt hơn, nếu là đồng hóa trị. Các phân tử tiểu đơn vị này có thể là cùng loại phân tử hóa học, hoặc có thể là các loại phân tử hóa học khác nhau. Ví dụ về loại phân tử dung hợp thứ nhất bao gồm, nhưng không giới hạn ở, các protein dung hợp (ví dụ, thể dung hợp giữa vùng liên kết ZFP hoặc TALE ADN và một hoặc nhiều vùng hoạt hóa) và các axit nucleic dung hợp (ví dụ, axit nucleic mã hóa protein dung hợp được mô tả trên đây). Ví dụ về loại phân tử dung hợp thứ hai bao gồm, nhưng không giới hạn ở, thể dung hợp giữa axit nucleic tạo bộ ba và polypeptit, và thể dung hợp giữa phân tử liên kết rãnh phụ và axit nucleic.

Sự biểu hiện của protein dung hợp trong tế bào có thể do sự phân phối protein dung hợp này vào tế bào hoặc bằng cách phân phối polynucleotit mã hóa protein dung hợp này vào tế bào, trong đó polynucleotit này được sao chép, và bản sao này được dịch mã, để tạo ra protein dung hợp. Nói chéo, phân cắt polypeptit và nối polypeptit cũng có thể tham gia vào sự biểu hiện protein trong tế bào. Các phương pháp phân phối polynucleotit và polypeptit vào tế bào được trình bày trong các đoạn khác của phần mô tả này.

“Gen” với các mục đích của phần mô tả này, bao gồm vùng ADN mã hóa sản phẩm gen (xem dưới đây), cũng như tất cả các vùng ADN điều hòa sự sản sinh sản phẩm gen này, dù các trình tự điều hòa này có liên kết với trình tự mã hóa và/hoặc được sao chép hay không. Theo đó, gen bao gồm, nhưng không nhất thiết giới hạn ở, các trình tự khởi động, trình tự kết thúc, các trình tự điều hòa dịch mã như các vị trí liên kết ribosom và các vị trí đi vào trong ribosom, trình tự tăng cường, trình tự bất hoạt, trình tự cách ly, phân tử biên, gốc sao chép, vị trí gắn chất nền và các vùng điều khiển locut.

“Biểu hiện gen” dùng để chỉ sự chuyển thông tin, chứa trong gen, thành sản phẩm gen. Sản phẩm gen có thể là sản phẩm phiên mã trực tiếp của gen (ví dụ, mARN, tARN, rARN, ARN đối nghĩa, ribozym, ARN cấu trúc hoặc loại ARN bất kỳ khác) hoặc protein được tạo ra nhờ sự dịch mã mARN. Các sản phẩm gen cũng bao gồm các ARN được biến đổi, bằng các quá trình như gắn mũ, polyadenyl hóa, metyl hóa, và chỉnh sửa, và các protein đã biến đổi nhờ, ví dụ, sự metyl hóa, axetyl hóa, phosphoryl hóa, ubiquitin hóa, ADP-ribosyl hóa, myristyl hóa, và glycosyl hóa.

“Điều biến” sự biểu hiện gen đề cập đến sự thay đổi về hoạt tính của gen. Việc điều biến sự biểu hiện có thể bao gồm, nhưng không giới hạn ở, hoạt hóa gen và ức chế gen. Việc chỉnh sửa bộ gen (ví dụ, phân cắt, biến đổi, bất hoạt, gây đột biến ngẫu nhiên) có thể được sử dụng để điều biến sự biểu hiện. Sự bất hoạt gen đề cập đến sự giảm biểu hiện gen bất kỳ so với tế bào không chứa ZFP, TALE hoặc hệ CRISPR/Cas như được mô tả ở đây. Do đó, sự bất hoạt gen có thể là một phần hoặc hoàn toàn.

mARN “được bảo vệ” là mARN mà trong đó mARN này được biến đổi theo một số cách để làm tăng tính ổn định hoặc sự dịch mã của mARN này. Ví dụ về các phương pháp bảo vệ bao gồm sử dụng kỹ thuật thay thế đến 25% các gốc cytosin và uridin bằng 2-thiouridin (s2U) và 5-methylcytidin (m5C). mARN tạo thành thể hiện khả năng gây miễn dịch thấp hơn và ổn định hơn so với dạng chưa được biến đổi của nó. (xem Karikó và đồng tác giả. ((2012), *Molecular Therapy*, Tập 16, số 11, trang 1833 – 1844). Các thay đổi khác bao gồm sự bổ sung thêm mũ ARCA, làm tăng khả năng dịch mã của mARN được tạo ra trên *in vitro* (xem Patent Mỹ US7074596).

“Vùng cần quan tâm” là vùng bất kỳ của sợi nhiễm sắc tế bào, như, ví dụ, gen hoặc trình tự không mã hóa nằm trong hoặc liền kề với gen, mà mong muốn liên kết với phân tử ngoại sinh trong đó. Sự liên kết có thể là nhằm mục đích phân cắt ADN đích và/hoặc tái tổ hợp đích. Vùng cần quan tâm có thể có mặt trong nhiễm sắc thể, thể bổ sung, bộ gen của cơ quan tế bào (ví dụ, ti thể, lục lạp), hoặc ví dụ, bộ gen của virus gây nhiễm. Vùng cần quan tâm có thể nằm trong vùng mã hóa của gen, nằm trong vùng không mã hóa được sao chép như, ví dụ, các trình tự dẫn, các trình tự kéo sau hoặc intron, hoặc trong vùng không được sao chép, ngược dòng hoặc xuôi dòng vùng mã hóa. Vùng cần quan tâm có thể ngắn bằng một cặp nucleotit duy nhất hoặc dài đến 2.000 cặp nucleotit, hoặc số cặp nucleotit nguyên bất kỳ.

Tế bào “nhân chuẩn” bao gồm, nhưng không giới hạn ở, các tế bào nấm (như nấm men), tế bào thực vật, tế bào động vật và tế bào của người (ví dụ, tế bào T).

Các thuật ngữ “liên kết hoạt động” và “được liên kết hoạt động” (hoặc “được liên kết hoạt động”) được sử dụng thay thế lẫn nhau về vị trí liền kề của hai hoặc nhiều thành phần (như các đơn vị trình tự), trong đó các thành phần này được sắp xếp sao cho cả hai thành phần đều hoạt động bình thường và cho phép khả năng là ít nhất một trong số các thành phần này có thể thực hiện chức năng mà được gây ra bởi ít nhất một trong số các

thành phần còn lại. Bằng cách minh họa, trình tự điều hòa phiên mã, như trình tự khởi động, được liên kết hoạt động với trình tự mã hóa nếu trình tự điều hòa phiên mã điều khiển mức phiên mã của trình tự mã hóa tương ứng với sự có mặt hoặc vắng mặt của một hoặc nhiều yếu tố điều hòa phiên mã. Trình tự điều hòa phiên mã thường được liên kết hoạt động trong cấu hình cis với trình tự mã hóa, nhưng không cần phải liên kết trực tiếp với nó. Ví dụ, trình tự tăng cường là trình tự điều hòa phiên mã mà được liên kết hoạt động với trình tự mã hóa, ngay cả khi chúng không liên kết.

Đối với polypeptit dung hợp, thuật ngữ “được liên kết hoạt động” có thể đề cập đến trường hợp là mỗi một trong số các thành phần này, khi liên kết với thành phần còn lại, thực hiện chức năng giống như khi nó không được liên kết như vậy. Ví dụ, đối với polypeptit dung hợp, trong đó vùng liên kết ZFP, TALE hoặc Cas ADN được dung hợp với vùng hoạt hóa, vùng liên kết ZFP, TALE hoặc Cas ADN này và vùng hoạt hóa này là liên kết hoạt động nếu, trong polypeptit dung hợp này, phần vùng liên kết ZFP, TALE hoặc Cas ADN có thể liên kết với vị trí đích của nó và/hoặc vị trí liên kết của nó, trong khi vùng hoạt hóa này có thể điều hòa tăng sự biểu hiện gen. Khi polypeptit dung hợp, trong đó vùng liên kết ZFP, TALE hoặc Cas ADN được dung hợp với vùng phân cắt, vùng liên kết ZFP, TALE hoặc Cas ADN và vùng phân cắt này là liên kết hoạt động nếu, trong polypeptit dung hợp này, phần vùng liên kết ZFP, TALE hoặc Cas ADN có thể liên kết với vị trí đích của nó và/hoặc vị trí liên kết của nó, trong khi vùng phân cắt có thể phân cắt ADN ở vùng lân cận của vị trí đích.

“Đoạn chức năng” của protein, polypeptit hoặc axit nucleic là protein, polypeptit hoặc axit nucleic có trình tự không tương đồng với protein, polypeptit hoặc axit nucleic có chiều dài đầy đủ, nhưng vẫn giữ được chức năng giống như protein, polypeptit hoặc axit nucleic có chiều dài đầy đủ này. Đoạn chức năng có thể có số lượng gốc nhiều hơn, ít hơn, hoặc bằng số lượng gốc của phân tử tự nhiên tương ứng, và/hoặc có thể chứa một hoặc nhiều axit amin thế hoặc nucleotit thế. Các phương pháp để xác định chức năng của axit nucleic (ví dụ, chức năng mã hóa, khả năng lai với axit nucleic khác) đã được biết rõ trong lĩnh vực này. Tương tự, các phương pháp để xác định chức năng protein là đã được biết rõ. Ví dụ, chức năng liên kết ADN của polypeptit có thể được xác định, ví dụ, bằng các thử nghiệm liên kết bộ lọc, phương pháp điện di dịch chuyển-di động, hoặc các thử nghiệm kết tủa miễn dịch. Sự phân cắt ADN có thể được thử nghiệm bằng kỹ thuật điện di gel. Xem

Ausubel và đồng tác giả., trên đây. Khả năng tương tác của protein với một protein khác có thể được xác định, ví dụ, bằng thử nghiệm đồng kết tủa miễn dịch, thử nghiệm lai đôi hoặc thử nghiệm bổ trợ, cả di truyền và hóa sinh.

“Vecto” có khả năng chuyển các trình tự gen sang tế bào đích. Thông thường, “cấu trúc vectơ”, “vector biểu hiện” và “vector chuyển gen” có nghĩa là cấu trúc axit nucleic bất kỳ có khả năng điều khiển sự biểu hiện của gen cần quan tâm và có thể chuyển các trình tự gen sang tế bào đích. Do đó, thuật ngữ này bao gồm các phương tiện tách dòng, và biểu hiện cũng như các vector tích hợp.

Thuật ngữ “đối tượng” và “bệnh nhân” được sử dụng thay thế lẫn nhau và đề cập đến các động vật có vú như bệnh nhân là người và các động vật linh trưởng không phải người, cũng như các động vật thí nghiệm như thỏ, chó, mèo, chuột hoang, chuột và các động vật khác. Theo đó, thuật ngữ “đối tượng” hoặc “bệnh nhân” như được sử dụng ở đây có nghĩa là bệnh nhân hoặc đối tượng là động vật có vú bất kỳ mà các tế bào gốc theo sáng chế có thể được sử dụng cho. Đối tượng theo sáng chế bao gồm các đối tượng bị phơi nhiễm với một hoặc nhiều độc tố hóa học, bao gồm, ví dụ, độc tố thần kinh.

“Tính gốc” đề cập đến khả năng tương đối của tế bào bất kỳ trong việc hoạt động theo kiểu tương tự tế bào gốc, *nghĩa là*, mức độ toàn năng, đa năng hoặc vài tiềm năng và mức độ tự tái sinh vô hạn hoặc được mở rộng mà tế bào gốc cụ thể bất kỳ có thể có.

Nucleaza

Được mô tả ở đây là các chế phẩm, cụ thể là các nucleaza, mà hữu ích để phân cắt *in vivo* phân tử cho mang gen chuyển và các nucleaza để phân cắt bộ gen của tế bào sao cho gen chuyển này được tích hợp vào bộ gen theo cách được gắn đích. Trong một số phương án, một hoặc nhiều nucleaza này xuất hiện tự nhiên. Trong các phương án khác, một hoặc nhiều nucleaza này không xuất hiện tự nhiên, nghĩa là, được thao tác di truyền trong vùng liên kết ADN và/hoặc vùng phân cắt. Ví dụ, vùng liên kết ADN của nucleaza tự nhiên có thể được biến đổi để liên kết với vị trí đích chọn lọc (ví dụ, meganucleaza mà được thao tác di truyền để liên kết với vị trí khác với vị trí liên kết cùng gốc). Trong các phương án khác, nucleaza này chứa các vùng phân cắt và liên kết ADN khác loại (ví dụ, các nucleaza ngón tay kẽm; protein liên kết ADN vùng hiệu ứng TAL; vùng liên kết ADN meganucleaza có các vùng phân cắt khác loại).

Vùng liên kết ADN

Trong một số phương án, vùng liên kết ADN của một hoặc nhiều nucleaza được sử dụng để phân cắt *in vivo* và/hoặc phân cắt hướng đích bộ gen của tế bào chứa protein ngón tay kẽm. Tốt hơn, nếu protein ngón tay kẽm là không xuất hiện tự nhiên, nghĩa là nó được thao tác di truyền để liên kết với vị trí đích đã chọn. Xem, ví dụ, Beerli và đồng tác giả. (2002) *Nature Biotechnol.* 20:135-141; Pabo và đồng tác giả. (2001) *Ann. Rev. Biochem.* 70:313-340; Isalan và đồng tác giả. (2001) *Nature Biotechnol.* 19:656-660; Segal và đồng tác giả. (2001) *Curr. Opin. Biotechnol.* 12:632-637; Choo và đồng tác giả. (2000) *Curr. Opin. Struct. Biol.* 10:411-416; Patent Mỹ số 6.453.242; 6.534.261; 6.599.692; 6.503.717; 6.689.558; 7.030.215; 6.794.136; 7.067.317; 7.262.054; 7.070.934; 7.361.635; 7.253.273; và Công bố Patent Mỹ số 2005/0064474; 2007/0218528; 2005/0267061, tất cả các tài liệu này được đưa vào đây bằng cách viện dẫn toàn bộ.

Vùng liên kết ngón tay kẽm được thao tác di truyền có thể có tính đặc hiệu liên kết mới, so với protein ngón tay kẽm xuất hiện tự nhiên. Các phương pháp thao tác di truyền bao gồm, nhưng không giới hạn ở, thiết kế hợp lý và các loại chọn lọc khác nhau. Thiết kế hợp lý bao gồm, ví dụ, sử dụng cơ sở dữ liệu chứa các trình tự nucleotit bộ ba (hoặc bộ bốn) và các trình tự axit amin ngón tay kẽm đơn lẻ, trong đó mỗi trình tự nucleotit bộ ba hoặc bộ bốn này được liên hợp với một hoặc nhiều trình tự axit amin của các ngón tay kẽm mà liên kết với một trình tự bộ ba hoặc bộ bốn cụ thể. Xem, ví dụ, Patent Mỹ đồng sở hữu 6.453.242 và 6.534.261, được đưa vào đây bằng cách viện dẫn toàn bộ.

Các phương pháp chọn lọc điển hình bao gồm các hệ biểu hiện thể thực khuẩn và hệ lai đôi, được mô tả trong Patent Mỹ số 5.789.538; 5.925.523; 6.007.988; 6.013.453; 6.410.248; 6.140.466; 6.200.759; và 6.242.568; cũng như WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 và GB 2.338.237. Ngoài ra, việc tăng cường tính đặc hiệu liên kết cho các vùng liên kết ngón tay kẽm được mô tả, ví dụ, trong WO 02/077227 đồng sở hữu.

Ngoài ra, như được mô tả trong các tài liệu tham khảo này và các tài liệu khác, các vùng ngón tay kẽm và/hoặc các protein ngón tay kẽm có nhiều ngón có thể được liên kết với nhau bằng cách sử dụng các trình tự cầu nối thích hợp bất kỳ, bao gồm ví dụ, các cầu nối dài 5 hoặc nhiều axit amin. Xem, Patent Mỹ số 6.479.626; 6.903.185; và 7.153.949 về các trình tự cầu nối điển hình dài 6 hoặc nhiều axit amin. Các protein được mô tả ở đây có

thể bao gồm tổ hợp bất kỳ của các cầu nối thích hợp giữa các ngón tay kẽm đơn lẻ của protein này.

Chọn lọc vị trí đích; các ZFP và các phương pháp để thiết kế và cấu trúc các protein dung hợp (và polynucleotit mã hóa protein này) đã được biết đối với người có kiến thức trung bình trong lĩnh vực này và được mô tả chi tiết trong Patent Mỹ số 6.140.081; 5.789.538; 6.453.242; 6.534.261; 5.925.523; 6.007.988; 6.013.453; 6.200.759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970 WO 01/88197; WO 02/099084; WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 và WO 03/016496.

Gần như bất kỳ cầu nối (đoạn đệm) nào cũng có thể được sử dụng giữa một hoặc nhiều thành phần của vùng liên kết ADN (ví dụ, các ngón tay kẽm), giữa một hoặc nhiều vùng liên kết ADN và/hoặc giữa vùng liên kết ADN này và vùng chức năng (ví dụ, nucleaza). Ví dụ không giới hạn về các trình tự cầu nối thích hợp bao gồm các Patent Mỹ số 8.772.453; 7.888.121; 6.479.626; 6.903.185; và 7.153.949; và Công bố Patent Mỹ số 20090305419; 20150064789 và 20150132269. Do đó, các protein được mô tả ở đây có thể bao gồm tổ hợp bất kỳ của các cầu nối thích hợp giữa các thành phần liên kết ADN đơn lẻ và/hoặc giữa vùng liên kết ADN và vùng chức năng của các thành phần được mô tả ở đây.

B. Vùng phân cắt

Vùng phân cắt thích hợp bất kỳ có thể được liên kết hoạt động với các vùng liên kết ADN như được mô tả ở đây để tạo ra nucleaza. Vùng phân cắt này có thể là khác loại với vùng liên kết ADN này, ví dụ vùng liên kết ADN ngón tay kẽm và vùng phân cắt từ nucleaza. Các vùng phân cắt khác loại có thể thu được từ endonucleaza hoặc exonucleaza bất kỳ. Các endonucleaza điển hình mà vùng phân cắt có thể thu được từ đó bao gồm, nhưng không giới hạn ở, endonucleaza giới hạn và endonucleaza hướng đích. Xem, ví dụ, 2002-2003 Catalogue, New England Biolabs, Beverly, MA; và Belfort và đồng tác giả. (1997) *Nucleic Acids Res.* 25:3379-3388. Các enzym phân cắt ADN khác đã được biết (ví dụ, Nucleaza S1; nucleaza đậu xanh; DNaza I của tụy; nucleaza vi cầu khuẩn; endonucleaza nấm men HO; cũng xem Linn và đồng tác giả. (eds.) *Nucleases*, Cold Spring Harbor Laboratory Press, 1993). Một hoặc nhiều enzym này (hoặc đoạn chức năng của chúng) có thể được sử dụng làm nguồn chứa các vùng phân cắt và nửa vùng phân cắt.

Tương tự, nửa vùng phân cắt có thể thu được từ nucleaza bất kỳ hoặc một phần của nó, như được chỉ ra trên đây, mà cần đime hóa cho hoạt tính phân cắt. Nói chung, hai protein dung hợp là cần thiết cho sự phân cắt nếu các protein dung hợp này chứa các nửa vùng phân cắt. Ngoài ra, một protein duy nhất chứa hai nửa vùng phân cắt có thể được sử dụng. Hai nửa vùng phân cắt có thể thu được từ cùng một endonucleaza (hoặc các đoạn chức năng của chúng), hoặc mỗi nửa vùng phân cắt có thể thu được từ một endonucleaza khác (hoặc các đoạn chức năng của chúng). Ngoài ra, các vị trí đích cho hai protein dung hợp tốt hơn, nếu được bố trí với nhau, sao cho sự liên kết của hai protein dung hợp này với các vị trí đích tương ứng của chúng đặt các nửa vùng phân cắt theo hướng không gian với nhau cho phép các nửa vùng phân cắt này tạo ra vùng phân cắt hoạt động, ví dụ, bằng cách đime hóa. Do đó, trong một số phương án, các mép gần của các vị trí đích được phân tách bởi 5-8 nucleotit hoặc bởi 15-18 nucleotit. Tuy nhiên, số nguyên nucleotit hoặc cặp nucleotit bất kỳ có thể xen vào giữa hai vị trí đích (ví dụ, từ 2 đến 50 cặp nucleotit hoặc nhiều hơn). Nói chung, vị trí phân cắt nằm giữa các vị trí đích.

Endonucleaza giới hạn (enzyme giới hạn) xuất hiện ở nhiều loài và có khả năng liên kết đặc hiệu trình tự với ADN (ở vị trí nhận biết), và cắt ADN tại hoặc gần vị trí liên kết. Một số enzym giới hạn (ví dụ, Typ IIS) cắt ADN ở các vị trí được loại bỏ khỏi vị trí nhận biết và có các vùng phân cắt và liên kết có thể tách rời. Ví dụ, enzym Typ IIS Fok I xúc tác cho sự phân cắt sợi kép của ADN, ở vị trí 9 nucleotit từ vị trí nhận biết của nó trên một sợi và 13 nucleotit từ vị trí nhận biết của nó trên sợi còn lại. Xem, ví dụ, Patent Mỹ số 5.356.802; 5.436.150 và 5.487.994; cũng như Li và đồng tác giả. (1992) Proc. Natl. Acad. Sci. USA 89:4275-4279; Li và đồng tác giả. (1993) Proc. Natl. Acad. Sci. USA 90:2764-2768; Kim và đồng tác giả. (1994a) Proc. Natl. Acad. Sci. USA 91:883-887; Kim và đồng tác giả. (1994b) J. Biol. Chem. 269:31,978-31,982. Do đó, trong một phương án, các protein dung hợp chứa vùng phân cắt (hoặc nửa vùng phân cắt) của ít nhất một enzym giới hạn Typ IIS và một hoặc nhiều vùng liên kết ngón tay kẽm, mà có thể hoặc không thể được thao tác di truyền.

Enzym giới hạn Typ IIS điển hình, mà có vùng phân cắt tách được khỏi vùng liên kết, là FokI. Enzym cụ thể này hoạt động như một đime. Bitinaite và đồng tác giả. (1998) Proc. Natl. Acad. Sci. USA 95:10,570-10,575. Theo đó, với mục đích của phần mô tả này, phần enzym Fok I được sử dụng trong các protein dung hợp được mô tả này được coi là

nửa vùng phân cắt. Do đó, đối với sự phân cắt sợi kép hướng đích và/hoặc thay thế hướng đích các trình tự tế bào sử dụng các thể dung hợp ngón tay kẽm-Fok I, hai protein dung hợp, mỗi protein chứa nửa vùng phân cắt FokI, có thể được sử dụng để khôi phục lại vùng phân cắt hoạt động xúc tác. Ngoài ra, một phân tử polypeptit duy nhất chứa vùng liên kết ngón tay kẽm và hai nửa vùng phân cắt Fok I cũng có thể được sử dụng. Các tham số cho sự phân cắt hướng đích và biến đổi trình tự đích sử dụng các thể dung hợp ngón tay kẽm-Fok I được đưa ra trong các đoạn khác của phần mô tả này.

Vùng phân cắt hoặc nửa vùng phân cắt có thể là phần bất kỳ của protein giữ được hoạt tính phân cắt, hoặc giữ được khả năng multimer hóa (ví dụ, dimer hóa) để tạo ra vùng phân cắt chức năng.

Các enzym giới hạn Typ IIS điển hình được mô tả trong Patent Mỹ số 7.888.121 được đưa toàn bộ vào đây. Các enzym giới hạn khác cũng chứa các vùng phân cắt và liên kết có thể tách rời, và các enzym này được dự tính bởi sáng chế. Xem, ví dụ, Roberts và đồng tác giả. (2003) *Nucleic acids Res.* 31:418-420.

Trong một số phương án, vùng phân cắt này chứa một hoặc nhiều nửa vùng phân cắt được thao tác di truyền (còn được gọi là đột biến vùng dimer hóa) mà làm giảm tối đa hoặc ngăn cản sự homodimer hóa, như được mô tả, ví dụ, trong Xem, ví dụ, Patent Mỹ số 7.914.796; 8.034.598 và 8.623.618, phần mô tả của tất cả các tài liệu này được đưa vào đây bằng các viện dẫn toàn bộ. Các gốc axit amin ở các vị trí 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537, và 538 của FokI đều là đích để tác động đến sự dimer hóa các nửa vùng phân cắt FokI.

Các nửa vùng phân cắt được thao tác di truyền điển hình của FokI mà tạo ra các heterodimer bắt buộc bao gồm một cặp trong đó nửa vùng phân cắt thứ nhất chứa các đột biến ở các gốc axit amin ở các vị trí 490 và 538 của FokI và nửa vùng phân cắt thứ hai chứa các đột biến ở các gốc axit amin 486 và 499.

Do đó, trong một phương án, đột biến ở vị trí 490 thay thế Glu (E) bằng Lys (K); đột biến ở vị trí 538 thay thế Iso (I) bằng Lys (K); đột biến ở vị trí 486 thay thế Gln (Q) bằng Glu (E); và đột biến ở vị trí 499 thay thế Iso (I) bằng Lys (K). Cụ thể là, các nửa vùng phân cắt được thao tác di truyền đã mô tả ở đây được tạo ra bằng cách gây đột biến các vị trí 490 (E→K) và 538 (I→K) trong một nửa vùng phân cắt để tạo ra nửa vùng phân cắt được thao tác di truyền được ký hiệu là “E490K:I538K” và bằng cách gây đột biến các vị

trí 486 (Q→E) và 499 (I→L) trong một nửa vùng phân cắt khác để tạo ra nửa vùng phân cắt được thao tác di truyền được ký hiệu là “Q486E:I499L”. Các nửa vùng phân cắt được thao tác di truyền đã mô tả ở đây là các đột biến heterodime bắt buộc, trong đó sự phân cắt bất thường được làm giảm tối đa hoặc được loại bỏ. *Xem*, ví dụ, Công bố Patent Mỹ số 2008/0131962, phần mô tả của tài liệu này được đưa vào đây bằng cách viện dẫn toàn bộ cho tất cả các mục đích. Trong một số phương án, nửa vùng phân cắt được thao tác di truyền chứa các đột biến ở các vị trí 486, 499 và 496 (được đánh số so với FokI kiểu đại), ví dụ, các đột biến thay thế gốc Gln (Q) kiểu đại ở vị trí 486 bằng gốc Glu (E), gốc Iso (I) kiểu đại ở vị trí 499 bằng gốc Leu (L) và gốc Asn (N) kiểu đại ở vị trí 496 bằng gốc Asp (D) hoặc Glu (E) (còn được gọi lần lượt là vùng “ELD” và “ELE”). Trong các phương án khác, nửa vùng phân cắt được thao tác di truyền chứa các đột biến ở các vị trí 490, 538 và 537 (được đánh số so với FokI kiểu đại), ví dụ các đột biến thay thế gốc Glu (E) kiểu đại ở vị trí 490 bằng gốc Lys (K), gốc Iso (I) kiểu đại ở vị trí 538 bằng gốc Lys (K), và gốc His (H) kiểu đại ở vị trí 537 bằng gốc Lys (K) hoặc gốc Arg (R) (còn được gọi lần lượt là vùng “KKK” và “KKR”). Trong các phương án khác, nửa vùng phân cắt được thao tác di truyền chứa các đột biến ở các vị trí 490 và 537 (được đánh số so với FokI kiểu đại), ví dụ các đột biến thay thế gốc Glu (E) kiểu đại ở vị trí 490 bằng gốc Lys (K) và gốc His (H) kiểu đại ở vị trí 537 bằng gốc Lys (K) hoặc gốc Arg (R) (còn được gọi lần lượt là vùng “KIK” và “KIR”). *Xem*, ví dụ, Patent Mỹ số 7.914.796; 8.034.598 và 8.623.618. Trong các phương án khác, nửa vùng phân cắt được thao tác di truyền chứa các đột biến “Sharkey” và/hoặc “Sharkey” “ (xem Guo và đồng tác giả, (2010) *J. Mol. Biol.* 400(1):96-107).

Nửa vùng phân cắt được thao tác di truyền được mô tả ở đây có thể được tạo ra bằng cách sử dụng phương pháp thích hợp bất kỳ, ví dụ, bằng kỹ thuật gây đột biến điểm các nửa vùng phân cắt kiểu đại (Fok I) như được mô tả trong Công bố Patent Mỹ số 7.888.121; 7.914.796; 8.034.598 và 8.623.618.

Ngoài ra, các nucleaza có thể được lắp ráp in vivo ở các vị trí đích axit nucleic bằng cách sử dụng công nghệ được gọi là “enzym tách” (xem, ví dụ Công bố Patent Mỹ số 20090068164). Các thành phần của enzym tách này có thể được biểu hiện trên các cấu trúc biểu hiện riêng biệt, hoặc có thể được liên kết trong một khung đọc mở mà tại đó các thành phần đơn lẻ được phân tách, ví dụ, bởi peptit tự cắt 2A hoặc trình tự IRES. Các thành phần

có thể là các vùng liên kết ngón tay kẽm đơn lẻ hoặc các vùng của vùng liên kết axit nucleic meganucleaza.

Các nucleaza có thể được sàng lọc về hoạt tính trước khi sử dụng, ví dụ trong hệ nhiễm sắc thể dựa trên nấm men như được mô tả trong WO 2009/042163 và 20090068164. Sự biểu hiện của nucleaza này có thể nằm dưới sự điều khiển của trình tự khởi động cơ định hoặc trình tự khởi động cảm ứng được, ví dụ trình tự khởi động galactokinaza mà được hoạt hóa (được khử ức chế) khi có mặt raffinosa và/hoặc galactosa và bị ức chế khi có mặt glucoza.

(Các) nucleaza như được mô tả ở đây có thể thực hiện một hoặc nhiều sự cắt sợi kép và/hoặc sợi đơn ở vị trí đích. Trong một số phương án, nucleaza này chứa vùng phân cắt bất hoạt xúc tác (ví dụ, FokI và/hoặc protein Cas). Xem, ví dụ, Patent Mỹ số 9.200.266; 8.703.489 và Guillinger và đồng tác giả. (2014) *Nature Biotech.* 32(6):577-582. Vùng phân cắt bất hoạt xúc tác có thể, kết hợp với vùng hoạt động xúc tác, hoạt động như là nickaza (enzym gây đứt) để cắt sợi đơn. Do đó, hai nickaza có thể được sử dụng kết hợp để thực hiện sự cắt sợi kép trong vùng đặc hiệu. Các nickaza khác cũng được biết trong lĩnh vực này, ví dụ, McCaffery và đồng tác giả. (2016) *Nucleic acids Res.* 44(2):e11. doi: 10.1093/nar/gkv878. Công bố điện tử ngày 19, tháng 10, năm 2015.

Vị trí đích

Như được mô tả chi tiết trên đây, các vùng ADN có thể được thao tác di truyền để liên kết với trình tự bất kỳ đã chọn. Vùng liên kết ADN được thao tác di truyền này có thể có tính đặc hiệu liên kết mới, so với vùng liên kết ADN tự nhiên. Trong một số phương án, các vùng liên kết ADN này liên kết với trình tự bên trong trình tự tăng cường BCL11A, ví dụ vị trí đích (thường là 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 hoặc nhiều cặp bazơ) nằm giữa exon 2 và exon 3 của BCL11A, bao gồm các vùng liên kết ADN mà liên kết với trình tự bên trong vị trí quá mã ADNseI trong trình tự tăng cường BCL11A (ví dụ, +58) như được thể hiện trong Bảng 1. Các phương pháp thao tác di truyền bao gồm, nhưng không giới hạn ở, thiết kế hợp lý và nhiều loại chọn lọc khác nhau. Thiết kế hợp lý bao gồm, ví dụ, sử dụng các cơ sở dữ liệu chứa các trình tự nucleotit bộ ba (hoặc bộ bốn) và các trình tự axit amin ngón tay kẽm đơn lẻ, trong đó mỗi trình tự nucleotit bộ ba hoặc bộ bốn được liên hợp với một hoặc nhiều trình tự axit amin của các ngón tay kẽm mà liên kết trình tự bộ ba hoặc bộ bốn cụ thể này. Xem, ví dụ, Patent Mỹ đồng sở hữu 6.453.242 và

6.534.261, được đưa vào đây bằng cách viện dẫn toàn bộ. Thiết kế hợp lý của các vùng hiệu ứng TAL cũng có thể được thực hiện. Xem, ví dụ, Công bố Mỹ số 20110301073.

Các phương pháp chọn lọc điển hình có thể áp dụng cho các vùng liên kết ADN, bao gồm hệ biểu hiện thể thực khuẩn và hệ lai đôi, được mô tả trong Patent Mỹ 5.789.538; 5.925.523; 6.007.988; 6.013.453; 6.410.248; 6.140.466; 6.200.759; và 6.242.568; cũng như WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 và GB 2.338.237. Ngoài ra, sự tăng cường tính đặc hiệu liên kết cho vùng liên kết ngón tay kẽm được mô tả, ví dụ, trong WO 02/077227 được đồng sở hữu.

Sự chọn lọc các vị trí đích; nucleaza và các phương pháp để thiết kế và cấu trúc các protein dung hợp (và polynucleotit mã hóa các protein này) đã được biết đối với người có kiến thức trung bình trong lĩnh vực này và được mô tả chi tiết trong Công bố đơn Patent Mỹ số 20050064474 và 20060188987, được đưa vào đây bằng cách viện dẫn toàn bộ.

Thêm vào đó, như được mô tả trong các tài liệu này và các tài liệu khác, các vùng liên kết ADN (ví dụ, các protein ngón tay kẽm có nhiều ngón) và/hoặc các thể dung hợp của (các) vùng liên kết ADN này và (các) vùng chức năng có thể được liên kết với nhau bằng cách sử dụng các trình tự cầu nối thích hợp bất kỳ, bao gồm ví dụ, các cầu nối chứa 5 hoặc nhiều axit amin. Các Patent Mỹ số 8.772.453; 7.888.121 (ví dụ, cầu nối “ZC”); 6.479.626; 6.903.185; và 7.153.949; Công bố Mỹ số 20090305419) và 20150064789. Các protein được mô tả ở đây có thể bao gồm tổ hợp bất kỳ của các cầu nối thích hợp giữa các vùng liên kết ADN đơn lẻ của protein này. Xem, Patent Mỹ số 8.586.526.

Trình tự cho

Trong một số phương án, sáng chế đề cập đến sự tích hợp hướng đích nhờ nucleaza của một trình tự ngoại sinh vào bộ gen của tế bào bằng cách sử dụng phân tử liên kết vùng tăng cường BCL11A được mô tả ở đây. Như đã chú ý trên đây, sự gắn xen trình tự ngoại sinh (còn được gọi là “trình tự cho” hoặc “thể cho” hoặc “gen chuyển”), ví dụ để gây mất đoạn vùng xác định và/hoặc sửa gen đột biến hoặc để biểu hiện tăng gen kiểu đại. Rõ ràng là trình tự cho này thường không giống với trình tự bộ gen mà nó được đưa vào. Trình tự cho có thể chứa trình tự không tương đồng được gắn hai bên sườn là hai vùng tương đồng cho phép HDR hiệu quả ở vị trí cần quan tâm hoặc có thể được tích hợp bằng cơ chế sửa đổi dựa theo sự không tương đồng. Thêm vào đó, các trình tự cho có thể chứa phân tử vector chứa các trình tự không tương đồng với vùng cần quan tâm trong sợi nhiễm sắc của tế bào.

Phân tử cho có thể chứa một số vùng không liên tục tương đồng với sợi nhiễm sắc của tế bào, và ví dụ, dẫn đến sự mất đoạn của vùng tăng cường BCL11A (hoặc một đoạn của nó) khi được sử dụng làm cơ chất để sửa DBS bị gây ra bởi một trong số các nucleaza được mô tả ở đây. Ngoài ra, để gắn xen hướng đích các trình tự bình thường không có mặt ở vùng cần quan tâm, các trình tự này có thể có mặt trong phân tử axit nucleic cho và được gắn bên sườn bởi các vùng tương đồng với trình tự trong vùng cần quan tâm.

Polynucleotit để gắn xen còn có thể được gọi là polynucleotit “ngoại sinh”, polynucleotit hoặc phân tử “cho” hoặc “gen chuyển”. Polynucleotit cho có thể là ADN hoặc ARN, sợi đơn và/hoặc sợi kép và có thể được đưa vào tế bào ở dạng thẳng hoặc dạng vòng. *Xem, ví dụ, Công bố Patent Mỹ số 20100047805 và 20110207221.* (Các) trình tự cho tốt hơn, nếu nằm trong ADN MC, mà có thể được đưa vào tế bào ở dạng vòng hoặc dạng thẳng. Nếu được đưa vào ở dạng thẳng, các đầu của trình tự cho này có thể được bảo vệ (ví dụ, khỏi sự thoái biến phân cắt ngoài nhân) bằng các phương pháp đã biết đối với người có kiến thức trung bình trong lĩnh vực này. Ví dụ, một hoặc nhiều gốc dideoxynucleotit được bổ sung vào đầu cùng 3' của phân tử mạch thẳng và/hoặc các oligonucleotit tự bổ sung được nối với một hoặc cả hai đầu. *Xem, ví dụ, Chang và đồng tác giả. (1987) Proc. Natl. Acad. Sci. USA 84:4959-4963; Nehls và đồng tác giả. (1996) Science 272:886-889.* Các phương pháp khác để bảo vệ polynucleotit ngoại sinh khỏi sự thoái biến bao gồm, nhưng không giới hạn ở, bổ sung (các) nhóm amino đầu cùng và sử dụng các môi liên kết biến đổi giữa các nucleotit như, ví dụ, các gốc phosphorothioat, phosphoramidat, và O-metyl riboza hoặc deoxyriboza. Nếu được đưa vào ở dạng sợi kép, phân tử cho này có thể bao gồm một hoặc nhiều vị trí đích nucleaza, ví dụ, vị trí đích nucleaza gắn bên sườn gen chuyển cần được tích hợp vào bộ gen tế bào. *Xem, ví dụ, Công bố Patent Mỹ số 20130326645.*

Polynucleotit có thể được đưa vào tế bào như một phần của phân tử vector có các trình tự bổ sung như, ví dụ, các gốc sao chép, trình tự khởi động và gen mã hóa tính kháng kháng sinh. Ngoài ra, polynucleotit cho có thể được đưa vào dưới dạng axit nucleic trần, dưới dạng axit nucleic được tạo phức với hợp chất như liposom hoặc poloxame, hoặc có thể được phân phối bằng các virus (ví dụ, adenovirus, AAV, herpesvirus, retrovirus, lentivirus và lentivirus khuyết tật integrase (integrase defective lentivirus: IDLV)).

Trong một số phương án, phân tử cho sợi kép bao gồm các trình tự (ví dụ, trình tự mã hóa, còn được gọi là gen chuyển) dài hơn 1 kb, ví dụ từ 2 đến 200 kb, từ 2 đến 10kb (hoặc giá trị bất kỳ nằm trong khoảng này). Phân tử cho sợi kép cũng bao gồm, ví dụ, ít nhất một vị trí đích nucleaza. Trong một số phương án, phân tử cho bao gồm ít nhất 2 vị trí đích, ví dụ cho cặp ZFN hoặc TALEN. Thông thường, các vị trí đích nucleaza nằm ngoài các trình tự gen chuyển, ví dụ, 5' và/hoặc 3' đến các trình tự gen chuyển, để phân cắt gen chuyển này. (Các) vị trí phân cắt nucleaza có thể là cho (các) nucleaza bất kỳ. Trong một số phương án, vị trí đích nucleaza nằm trong trình tự cho sợi kép là cho (các) nucleaza giống nhau được sử dụng để phân cắt đích nội sinh mà trình tự cho đã phân cắt này được tích hợp vào đó bằng các phương pháp độc lập với sự tương đồng.

Trình tự cho này thường được gắn xen do đó sự biểu hiện của nó được điều khiển bởi trình tự khởi động nội sinh ở vị trí tích hợp, cụ thể là trình tự khởi động điều khiển sự biểu hiện của gen nội sinh mà trình tự cho được gắn xen vào đó (ví dụ, globin, AAVS1, v.v...). Tuy nhiên, rõ ràng là trình tự cho này có thể chứa trình tự khởi động và/hoặc yếu tố tăng cường, ví dụ trình tự khởi động cơ định hoặc trình tự khởi động cảm ứng được hoặc trình tự khởi động đặc hiệu mô.

Phân tử cho này có thể được gắn xen vào gen nội sinh sao cho toàn bộ, một số hoặc không có gì của gen nội sinh này được biểu hiện. Trong các phương án khác, gen chuyển này (ví dụ, có hoặc không có trình tự mã hóa globin) được tích hợp vào locut nội sinh bất kỳ, ví dụ locut cảng an toàn. Xem, ví dụ, Công bố patent Mỹ 20080299580; 20080159996 và 201000218264.

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

Các gen chuyển được mang trên trình tự cho được mô tả ở đây có thể được phân lập từ plasmid, tế bào hoặc các nguồn khác bằng cách sử dụng các kỹ thuật chuẩn đã biết trong lĩnh vực này như PCR. Các trình tự cho để sử dụng có thể bao gồm các loại cấu trúc hình học không gian khác nhau, bao gồm dạng vòng siêu xoắn, vòng tháo xoắn, dạng thẳng và tương tự. Ngoài ra, chúng có thể được tổng hợp hóa học bằng cách sử dụng các kỹ thuật tổng hợp oligonucleotit chuẩn. Thêm vào đó, các trình tự cho có thể được metyl hóa hoặc

thiếu sự metyl hóa. Các trình tự cho có thể là ở dạng nhiễm sắc thể nhân tạo vi khuẩn hoặc nấm men (BAC hoặc YAC).

Polynucleotit cho sợi kép được mô tả ở đây có thể bao gồm một hoặc nhiều bazo và/hoặc các khung chính không tự nhiên. Cụ thể là, sự gắn xen của phân tử cho với các xytosin được metyl hóa có thể được thực hiện bằng cách sử dụng các phương pháp được mô tả ở đây để đạt được trạng thái ngừng phiên mã trong vùng cần quan tâm.

Polynucleotit (cho) ngoại sinh có thể chứa trình tự cần quan tâm bất kỳ (trình tự ngoại sinh). Các trình tự ngoại sinh điển hình bao gồm, nhưng không giới hạn ở trình tự mã hóa polypeptit (ví dụ, cADN), các trình tự khởi động, các trình tự tăng cường, nhãn epitop, gen đánh dấu, vị trí nhận biết enzym phân cắt và các loại cấu trúc biểu hiện khác nhau bất kỳ. Các gen đánh dấu bao gồm, nhưng không giới hạn ở, các trình tự mã hóa protein mà gây ra tính kháng kháng sinh (ví dụ, tính kháng ampicilin, tính kháng neomycin, tính kháng G418, tính kháng puromycin), các trình tự mã hóa protein có màu hoặc protein huỳnh quang hoặc protein phát quang (ví dụ, protein huỳnh quang xanh lá cây, protein huỳnh quang xanh lá cây tăng cường, protein huỳnh quang đỏ, luciferaza), và các protein gây ra sự sinh trưởng tế bào tăng và/hoặc sự khuếch đại gen (ví dụ, dihydrofolat reductaza). Nhãn epitop bao gồm, ví dụ, một hoặc nhiều bản sao của FLAG, His, myc, Tap, HA hoặc trình tự axit amin phát hiện được bất kỳ.

Trong một phương án được ưu tiên, trình tự ngoại sinh (gen chuyển) chứa polynucleotit mã hóa polypeptit bất kỳ mà sự biểu hiện của nó trong tế bào được mong muốn, bao gồm, nhưng không giới hạn ở các kháng thể, kháng nguyên, enzym, thụ thể (bề mặt tế bào hoặc nhân), hormon, lymphokin, xytokin, polypeptit chỉ thị, yếu tố tăng trưởng, và các đoạn chức năng của bất kỳ trong số các phân tử trên đây. Trình tự mã hóa có thể là, ví dụ, các cADN.

Ví dụ, trình tự ngoại sinh có thể chứa trình tự mã hóa polypeptit mà thiếu hoặc không hoạt động ở đối tượng mắc bệnh di truyền, bao gồm nhưng không giới hạn ở bất kỳ trong số các bệnh di truyền sau đây: chứng loạn sản sụn, chứng mù màu, chứng thiếu axit maltaza, chứng thiếu adenosin deaminaza (OMIM No.102700), chứng loạn dưỡng não chất trắng thượng thận, hội chứng aicardi, bệnh thiếu alpha-1 antitrypsin, bệnh tan máu bẩm sinh alpha, hội chứng không nhạy cảm androgen, hội chứng apert, bệnh loạn sản thất phải loạn nhịp, chứng thất điều giãn mạch, hội chứng barth, bệnh tan máu bẩm sinh beta, hội

chứng nevi mụn nước cao su xanh, bệnh loạn dưỡng chất trắng canavan, bệnh u hạt mãn tính (CGD), hội chứng tiếng khóc như mèo kêu, bệnh xơ hóa nang, bệnh u mỡ đau, hội chứng loạn sản ngoại bì, bệnh thiếu máu fanconi, chứng loạn sản xơ hóa xương, hội chứng nhiễm sắc thể X dễ gãy, bệnh rối loạn chuyển hóa galactose, bệnh Gaucher, bệnh gangliosidosis toàn thể (ví dụ, GM1), bệnh thừa sắt, đột biến hemoglobin C ở codon thứ 6 của beta-globin (HbC), bệnh máu khó đông, bệnh Huntington, hội chứng Hurler, bệnh rối loạn chuyển hóa xương, hội chứng Klinefelter, bệnh Krabbes, hội chứng Langer-Giedion, suy giảm kết dính bạch cầu (LAD, OMIM số 116920), loạn dưỡng chất trắng não, hội chứng QT dài, hội chứng Marfan, hội chứng Moebius, bệnh rối loạn chuyển hóa bẩm sinh mucopolysaccharidosis (MPS), hội chứng móng tay-xương bánh chè, bệnh đái tháo nhạt do thận, bệnh u sợi thần kinh, bệnh Neimann-Pick, bệnh xương thủy tinh, bệnh rối loạn chuyển hóa porphyria, hội chứng Prader-Willi, hội chứng già trước tuổi, hội chứng Proteus, bệnh u nguyên bào võng mạc, hội chứng Rett, hội chứng Rubinstein-Taybi, hội chứng Sanfilippo, thiếu hụt miễn dịch tổ hợp nghiêm trọng (SCID), hội chứng Shwachman, bệnh hồng cầu hình liềm (bệnh thiếu máu hồng cầu hình liềm), hội chứng Smith-Magenis, hội chứng Stickler, hội chứng Tay-Sachs, hội chứng giảm tiểu cầu thiếu xương quay (TAR), hội chứng Treacher Collins, hiện tượng thể ba, bệnh xơ cứng củ, hội chứng Turner, rối loạn chu trình chuyển hóa urê, bệnh von Hippel-Landau, hội chứng Waardenburg, hội chứng Williams, bệnh Wilson's, hội chứng Wiskott-Aldrich, hội chứng tăng sinh lympho bào liên kết X (XLP, OMIM số 308240).

Các bệnh điển hình khác mà có thể được điều trị bằng cách tích hợp hướng đích bao gồm bệnh suy giảm miễn dịch mắc phải, bệnh dự trữ lysosom (ví dụ, bệnh Gaucher, GM1, bệnh Fabry và bệnh Tay-Sachs), bệnh rối loạn chuyển hóa bẩm sinh mucopolysaccharidosis (ví dụ, bệnh Hunter, bệnh Hurler), bệnh huyết sắc tố (ví dụ, bệnh hồng cầu hình liềm, HbC, bệnh tan máu bẩm sinh α , bệnh tan máu bẩm sinh β) và bệnh máu khó đông.

Trong một số phương án, các trình tự ngoại sinh này có thể chứa gen đánh dấu (được mô tả trên đây), cho phép lựa chọn tế bào đã tích hợp hướng đích, và trình tự liên kết mã hóa chức năng bổ sung. Ví dụ không giới hạn về các gen đánh dấu bao gồm GFP, (các) gen đánh dấu chọn lọc thuốc và tương tự.

Các trình tự gen khác mà có thể được gắn xen có thể bao gồm, ví dụ, các gen kiểu đại để thay thế các trình tự bị đột biến. Ví dụ, trình tự gen Yếu tố IX kiểu đại có thể được gắn xen vào bộ gen của tế bào gốc mà trong đó bản sao nội sinh của gen này bị gây đột biến. Bản sao kiểu đại có thể được gắn xen ở locut nội sinh, hoặc ngoài ra có thể được gắn đích đến locut cảng an toàn.

Cấu trúc của các catxet biểu hiện, theo phần mô tả này, sử dụng hệ phương pháp đã được biết rõ trong lĩnh vực sinh học phân tử (xem, ví dụ, Ausubel hoặc Maniatis). Trước khi sử dụng catxet biểu hiện để tạo ra động vật chuyển gen, sự đáp ứng của catxet biểu hiện với yếu tố gây căng thẳng kèm theo các yếu tố điều khiển chọn lọc có thể được thử nghiệm bằng cách đưa catxet biểu hiện này vào trong dòng tế bào thích hợp (ví dụ, dòng tế bào sơ cấp, dòng tế bào đã biến nạp, hoặc dòng tế bào bất tử).

Ngoài ra, mặc dù không cần thiết cho sự biểu hiện, các trình tự ngoại sinh cũng có thể bao gồm các trình tự điều hòa dịch mã hoặc phiên mã, ví dụ, các trình tự khởi động, yếu tố tăng cường, trình tự cách ly, vị trí đi vào trong ribosom, các trình tự mã hóa peptit 2A và/hoặc các tín hiệu polyadenyl hóa. Ngoài ra, các yếu tố điều khiển của gen cần quan tâm có thể được liên kết hoạt động với các gen chỉ thị để tạo ra các gen khảm (ví dụ, catxet biểu hiện chỉ thị).

Sự gắn xen hướng đích của trình tự axit nucleic không mã hóa cũng có thể đạt được. Các trình tự mã hóa ARN đối nghĩa, ARNi, shARN và vi ARN (miARN) cũng có thể được sử dụng để gắn xen hướng đích.

Trong các phương án khác, axit nucleic cho này có thể chứa trình tự không mã hóa là các vị trí đích đặc hiệu cho các thiết kế nucleaza khác. Do đó, các nucleaza khác có thể được biểu hiện trong các tế bào sao cho phân tử cho ban đầu bị cắt và biến đổi bằng cách gắn xen một phân tử cho khác cần quan tâm. Theo cách này, sự tích hợp lặp lại của các phân tử cho có thể được tạo ra cho phép xếp chồng đặc tính ở một locut cụ thể cần quan tâm hoặc ở locut cảng an toàn.

Phân phối

Các nucleaza như được mô tả ở đây (Bảng 1), polynucleotit mã hóa các nucleaza này, polynucleotit cho và các chế phẩm chứa các protein và/hoặc polynucleotit được mô tả

ở đây có thể được phân phối in vivo hoặc ex vivo bằng phương pháp thích hợp bất kỳ vào loại tế bào bất kỳ.

Các tế bào thích hợp bao gồm các dòng tế bào và/hoặc tế bào nhân chuẩn (ví dụ, động vật) và tiền nhân. Ví dụ không giới hạn về các tế bào này hoặc dòng tế bào được tạo ra từ các tế bào này bao gồm COS, CHO (ví dụ, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11, CHO-DUKX, CHOK1SV), VERO, MDCK, WI38, V79, B14AF28-G3, BHK, HaK, NS0, SP2/0-Ag14, HeLa, HEK293 (ví dụ, HEK293-F, HEK293-H, HEK293-T), và các tế bào perC6 cũng như các tế bào côn trùng như *Spodoptera fugiperda* (Sf), hoặc các tế bào nấm như *Saccharomyces*, *Pichia* và *Schizosaccharomyces*. Trong một số phương án, dòng tế bào này là CHO, MDCK hoặc dòng tế bào HEK293. Các tế bào thích hợp cũng bao gồm các tế bào gốc, như, bằng cách lấy ví dụ, tế bào gốc phôi, các tế bào gốc đa năng được cảm ứng, tế bào gốc tạo máu, các tế bào gốc thần kinh và các tế bào gốc trung mô.

Phương pháp phân phối nucleaza như được mô tả ở đây được mô tả, ví dụ, trong Patent Mỹ số 6.453.242; 6.503.717; 6.534.261; 6.599.692; 6.607.882; 6.689.558; 6.824.978; 6.933.113; 6.979.539; 7.013.219; và 7.163.824, phần mô tả của tất cả các tài liệu này được đưa vào đây bằng cách viện dẫn toàn bộ.

Các nucleaza và/hoặc các cấu trúc cho như được mô tả ở đây cũng có thể được phân phối bằng cách sử dụng các vector chứa các trình tự mã hóa một hoặc nhiều ZFN, được mô tả ở đây. Các hệ vector bất kỳ có thể được sử dụng bao gồm, nhưng không giới hạn ở, các vector plasmit, vector retrovirus, vector lentivirus, vector adenovirus, vector virus phát ban; vector herpesvirus và vector virus liên hợp adeno, v.v... Xem, Patent Mỹ số 6.534.261; 6.607.882; 6.824.978; 6.933.113; 6.979.539; 7.013.219; và 7.163.824, được đưa vào đây bằng cách viện dẫn toàn bộ. Ngoài ra, rõ ràng là bất kỳ trong số các vector này có thể chứa một hoặc nhiều trình tự cần thiết cho việc điều trị. Do đó, khi một hoặc nhiều nucleaza và cấu trúc cho được đưa vào trong tế bào, các nucleaza và/hoặc polynucleotit cho này có thể được mang trên cùng một vector hoặc trên các vector khác nhau ((các) ADN MC). Khi nhiều vector được sử dụng, mỗi vector có thể chứa trình tự mã hóa một hoặc nhiều nucleaza và/hoặc cấu trúc cho. Các phương pháp chuyển gen dựa trên virus hoặc không phải virus thông thường có thể được sử dụng để đưa các axit nucleic mã hóa các nucleaza và/hoặc các cấu trúc cho vào trong các tế bào (ví dụ, các tế bào động vật có vú) và các mô đích. Các hệ phân phối không phải vector virus bao gồm các plasmit ADN hoặc ARN, ADN MC,

axit nucleic trần, và axit nucleic được tạo phức với phương tiện phân phối như liposom hoặc poloxame. Các vector không phải virus thích hợp bao gồm các vector nanotaxis, bao gồm các vector đã có trên thị trường của InCellArt (Pháp). Hệ phân phối vector virus bao gồm các virus ADN và ARN, mà có các bộ gen thể bổ sung hoặc được tích hợp sau khi phân phối vào tế bào. Để xem lại sự phân phối in vivo các protein liên kết ADN được thao tác di truyền và protein dung hợp chứa các protein liên kết này, xem, ví dụ, Rebar (2004) Expert Opinion Invest. Drugs 13(7):829-839; Rossi và đồng tác giả. (2007) Nature Biotech. 25(12):1444-1454 cũng như các tài liệu tham khảo về sự phân phối gen nói chung như Anderson, Science 256:808-813 (1992); Nabel & Felgner, TIBTECH 11:211-217 (1993); Mitani & Caskey, TIBTECH 11:162-166 (1993); Dillon, TIBTECH 11:167-175 (1993); Miller, Nature 357:455-460 (1992); Van Brunt, Biotechnology 6(10):1149-1154 (1988); Vigne, Restorative Neurology and Neuroscience 8:35-36 (1995); Kremer & Perricaudet, British Medical Bulletin 51(1):31-44 (1995); Haddada và đồng tác giả, trong Current Topics in Microbiology and Immunology Doerfler và Böhm (eds.) (1995); và Yu và đồng tác giả., Gene Therapy 1:13-26 (1994).

Các phương pháp phân phối axit nucleic không phải bằng virus bao gồm kỹ thuật xung điện, phương pháp chuyển nhiễm liposom, vi tiêm, súng bắn gen, virosom, liposom, liposom miễn dịch, polycation hoặc các thể liên hợp lipid:axit nucleic, ADN trần, hạt virus nhân tạo, sự hấp thu ADN tăng cường nhờ hoạt chất. Sự tạo lỗ tạm thời trên màng bởi sóng âm (sonoporation) sử dụng, ví dụ, hệ thống Sonitron 2000 (Rich-Mar) cũng có thể được sử dụng để phân phối các axit nucleic.

Các hệ phân phối axit nucleic điển hình khác bao gồm các hệ phân phối được cung cấp bởi Amaxa Biosystems (Cologne, Đức), Maxcyte, Inc. (Rockville, Maryland), BTX Molecular Delivery Systems (Holliston, MA) và Copernicus Therapeutics Inc., (xem ví dụ US6008336). Chuyển nhiễm liposom được mô tả trong ví dụ, Patent Mỹ số 5.049.386; 4.946.787; và 4.897.355) và các chất phản ứng chuyển nhiễm liposom đã được bán trên thị trường (ví dụ, Transfectam™ và Lipofectin™). Các lipid cation hóa và trung tính thích hợp cho việc chuyển nhiễm liposom nhận biết thụ thể hiệu quả của polynucleotit bao gồm các lipid của Felgner, WO 91/17424, WO 91/16024.

Việc bào chế các phức hợp lipid:axit nucleic, bao gồm các liposom hướng đích như các phức hợp lipid miễn dịch, đã được biết rõ bởi người có kiến thức trung bình trong lĩnh

vực này (xem, ví dụ, Crystal, *Science* 270:404-410 (1995); Blaese và đồng tác giả., *Cancer Gene Ther.* 2:291-297 (1995); Behr và đồng tác giả., *Bioconjugate Chem.* 5:382-389 (1994); Remy và đồng tác giả, *Bioconjugate Chem.* 5:647-654 (1994); Gao và đồng tác giả, *Gene Therapy* 2:710-722 (1995); Ahmad và đồng tác giả, *Cancer Res.* 52:4817-4820 (1992); Patent Mỹ số 4.186.183, 4.217.344, 4.235.871, 4.261.975, 4.485.054, 4.501.728, 4.774.085, 4.837.028, và 4.946.787).

Các phương pháp phân phối khác bao gồm ứng dụng gói các axit nucleic cần được phân phối vào các phương tiện phân phối EnGeneIC (EDV). Các EDV này được phân phối đặc hiệu vào các mô đích bằng cách sử dụng các kháng thể đặc hiệu kép, trong đó một nhánh của kháng thể này có tính đặc hiệu với mô đích và nhánh còn lại có tính đặc hiệu với EDV. Kháng thể này mang các EDV đến bề mặt tế bào đích và sau đó EDV được mang vào trong tế bào bằng quá trình nhập bào. Khi ở trong tế bào, các thành phần bên trong được giải phóng (xem MacDiarmid và đồng tác giả (2009) *Nature Biotechnology* 27(7):643).

Việc sử dụng các hệ dựa trên virus ARN hoặc ADN để phân phối các axit nucleic mã hóa ZFP, TALE và/hoặc các hệ CRISPR/Cas được thao tác di truyền tận dụng được các quá trình tiến hóa cao để gắn đích virus vào các tế bào xác định trong cơ thể và vận chuyển tải lượng virus vào nhân. Các vector virus có thể được sử dụng trực tiếp cho bệnh nhân (in vivo) hoặc chúng có thể được sử dụng để điều trị cho các tế bào in vitro và các tế bào biến đổi được sử dụng cho bệnh nhân (ex vivo). Các hệ dựa trên virus thông thường để phân phối ZFP bao gồm, nhưng không giới hạn ở, các vector retrovirus, lentivirus, adenovirus, vector virus liên hợp adeno, vector virus bệnh đậu mùa và vector virus herpes simplex để chuyển gen. Sự tích hợp vào bộ gen vật chủ là có thể đối với các phương pháp chuyển gen retrovirus, lentivirus, và virus liên hợp adeno, thường gây ra sự biểu hiện kéo dài của gen chuyển đã gắn xen. Thêm vào đó, các hiệu quả tải nạp cao được quan sát thấy ở nhiều loại tế bào và các mô đích khác nhau.

Tính hướng kích thích của retrovirus có thể được biến đổi bằng cách kết hợp các protein vỏ ngoài ngoại lai, nhân giống quần thể đích tiềm năng của các tế bào đích. Các vector lentivirus là các vector retrovirus có khả năng tải nạp hoặc gây nhiễm các tế bào không phân chia và thường tạo ra hiệu giá virus cao. Việc lựa chọn hệ chuyển gen retrovirus phụ thuộc vào mô đích. Vector retrovirus được tạo bởi các đoạn lặp đầu cùng dài hoạt động cis

với dung tích gói lên đến 6-10 kb trình tự ngoại lai. Các LTR hoạt động cis tối thiểu là đủ cho sự sao chép và đóng gói của các vectơ, mà sau đó được sử dụng để tích hợp gen điều trị vào tế bào đích để tạo ra sự biểu hiện gen chuyển bền vững. Các vectơ retrovirus được sử dụng rộng rãi bao gồm các vectơ dựa trên virus bệnh bạch cầu chuột (MuLV), virus bệnh bạch cầu vượn (gibbon ape) (GaLV), virus gây suy giảm miễn dịch ở khỉ (SIV), virus gây suy giảm miễn dịch ở người (HIV), và kết hợp của chúng (xem, ví dụ, Buchscher và đồng tác giả, *J. Virol.* 66:2731-2739 (1992); Johann và đồng tác giả, *J. Virol.* 66:1635-1640 (1992); Sommerfelt và đồng tác giả, *Virol.* 176:58-59 (1990); Wilson và đồng tác giả, *J. Virol.* 63:2374-2378 (1989); Miller và đồng tác giả, *J. Virol.* 65:2220-2224 (1991); PCT/US94/05700).

Trong các ứng dụng, trong đó sự biểu hiện tạm thời được ưu tiên hơn, các hệ dựa trên adenovirus có thể được sử dụng. Các vectơ dựa trên adenovirus có khả năng tải nạp với hiệu quả rất cao trong nhiều loại tế bào và không đòi hỏi phân chia tế bào. Với các vectơ này, hiệu quả cao và các mức độ biểu hiện cao đã thu được. Vectơ này có thể được tạo ra với lượng lớn trong hệ thống tương đối đơn giản. Vectơ virus liên hợp deno ("AAV") cũng được sử dụng để tải nạp các tế bào với các axit nucleic đích, ví dụ, để sản sinh in vitro các axit nucleic và peptit, và cho các quy trình liệu pháp gen in vivo và ex vivo (xem, ví dụ, West và đồng tác giả, *Virology* 160:38-47 (1987); Patent Mỹ số 4.797.368; WO 93/24641; Kotin, *Human Gene Therapy* 5:793-801 (1994); Muzyczka, *J. Clin. Invest.* 94:1351 (1994). Cấu trúc của các vectơ AAV tái tổ hợp được mô tả trong nhiều công bố, bao gồm Patent Mỹ số 5.173.414; Tratschin và đồng tác giả, *Mol. Cell. Biol.* 5:3251-3260 (1985); Tratschin, và đồng tác giả, *Mol. Cell. Biol.* 4:2072-2081 (1984); Hermonat & Muzyczka, *PNAS* 81:6466-6470 (1984); và Samulski và đồng tác giả, *J. Virol.* 63:03822-3828 (1989).

Hiện nay, có ít nhất sáu phương pháp vectơ virus sẵn có cho việc chuyển gen trong các thử nghiệm lâm sàng, mà sử dụng các phương pháp bao gồm bổ sung cho các vectơ khiếm khuyết bằng các gen được gắn xen vào dòng tế bào trợ giúp để tạo ra chất tải nạp.

pLASN và MFG-S là các ví dụ về các vectơ retrovirus mà được sử dụng trong các thử nghiệm lâm sàng (Dunbar và đồng tác giả, *Blood* 85:3048-305 (1995); Kohn và đồng tác giả, *Nat. Med.* 1:1017-102 (1995); Malech và đồng tác giả, *PNAS* 94:22 12133-12138 (1997)). PA317/pLASN là vectơ điều trị đầu tiên được sử dụng trong thử nghiệm liệu pháp

gen. (Blaese và đồng tác giả, *Science* 270:475-480 (1995)). Hiệu quả tải nạp là 50% hoặc cao hơn được quan sát thấy cho các vector được đóng gói MFG-S. (Ellem và đồng tác giả, *Immunol Immunother.* 44(1):10-20 (1997); Dranoff và đồng tác giả, *Hum. Gene Ther.* 1:111-2 (1997).

Các vector virus liên hợp adeno tái tổ hợp (rAAV) là các hệ phân phối gen triển vọng khác dựa trên virus typ 2 liên hợp adeno parvovirus không gây bệnh và khiếm khuyết. Tất cả các vector đều thu được từ plasmit mà chỉ giữ lại các đoạn lặp đầu cùng đảo ngược AAV 145 bp gắn bên sườn catxet biểu hiện gen chuyển. Sự chuyển gen hiệu quả và sự phân phối gen chuyển bền vững do tích hợp vào các bộ gen của tế bào đã tải nạp là các đặc điểm chính đối với hệ vector này. (Wagner và đồng tác giả, *Lancet* 351:9117 1702-3 (1998), Kearns và đồng tác giả, *Gene Ther.* 9:748-55 (1996)). Các typ huyết thanh AAV khác, bao gồm AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9 và AAVrh.10 và typ huyết thanh AAV mới bất kỳ cũng có thể được sử dụng theo sáng chế.

Vector adenovirus (Ad) tái tổ hợp thiếu khả năng sao chép có thể được tạo ra ở hiệu quả cao và gây nhiễm dễ dàng nhiều loại tế bào khác nhau. Hầu hết vector adenovirus được thao tác di truyền sao cho gen chuyển thay thế các gen Ad E1a, E1b, và/hoặc E3; do đó vector thiếu khả năng sao chép này được nuôi cấy trong các tế bào 293 người mà cung cấp chức năng gen bị mất trong nhân tổ trans. Các vector Ad có thể tải nạp nhiều loại mô in vivo, bao gồm các tế bào biệt hóa, không phân chia như các tế bào được tìm thấy trong gan, thận và cơ. Các vector Ad thông thường có dung tích mang lớn. Ví dụ về việc sử dụng vector Ad trong thử nghiệm lâm sàng bao gồm liệu pháp polynucleotit để gây miễn dịch chống khối u bằng cách tiêm trong cơ (Sternman và đồng tác giả, *Hum. Gene Ther.* 7:1083-9 (1998)). Các ví dụ khác về việc sử dụng các vector adenovirus để chuyển gen trong các thử nghiệm lâm sàng bao gồm Rosenecker và đồng tác giả, *Infection* 24:1 5-10 (1996); Sternman và đồng tác giả, *Hum. Gene Ther.* 9:7 1083-1089 (1998); Welsh và đồng tác giả, *Hum. Gene Ther.* 2:205-18 (1995); Alvarez và đồng tác giả, *Hum. Gene Ther.* 5:597-613 (1997); Topf và đồng tác giả, *Gene Ther.* 5:507-513 (1998); Sternman và đồng tác giả, *Hum. Gene Ther.* 7:1083-1089 (1998).

Các tế bào đóng gói được sử dụng để tạo ra các hạt virus có khả năng gây nhiễm tế bào chủ. Các tế bào này bao gồm các tế bào 293, mà đóng gói adenovirus, và các tế bào ψ 2 hoặc tế bào PA317, mà đóng gói retrovirus. Vector virus được sử dụng trong liệu pháp gen

thường được tạo ra bằng dòng tế bào sản xuất mà đóng gói vector axit nucleic vào hạt virus. Các vector này thường chứa các trình tự virus nhỏ nhất cần thiết cho việc đóng gói và tích hợp tiếp vào vật chủ (nếu có thể áp dụng), các trình tự virus khác đang được thay thế bằng catxet biểu hiện mã hóa protein cần được biểu hiện. Các chức năng virus bị mất được cung cấp trong nhân tố trans bởi dòng tế bào đóng gói. Ví dụ, các vector AAV được sử dụng trong liệu pháp gen thường chỉ có các trình tự lặp đầu cùng đảo ngược (ITR) từ bộ gen AAV mà cần thiết cho việc đóng gói và tích hợp vào bộ gen vật chủ. ADN virus được đóng gói trong dòng tế bào, mà chứa plasmit trợ giúp mã hóa các gen AAV còn lại, cụ thể là rep và cap, nhưng thiếu các trình tự ITR. Dòng tế bào này cũng được gây nhiễm bằng adenovirus dưới dạng trợ giúp. Virus trợ giúp này thúc đẩy sự sao chép của vector AAV và sự biểu hiện của các gen AAV từ plasmit trợ giúp. Plasmit trợ giúp này không được đóng gói trong các lượng đáng kể do thiếu các trình tự ITR. Sự nhiễm adenovirus có thể được giảm bởi, ví dụ, việc xử lý nhiệt mà adenovirus là nhạy cảm với việc xử lý nhiệt hơn so với AAV.

Trong nhiều ứng dụng liệu pháp gen, mong muốn là vector liệu pháp gen được phân phối với mức độ đặc hiệu cao đến loại mô xác định. Theo đó, vector virus có thể được biến đổi để có tính đặc hiệu đối với loại tế bào đã cho bằng cách biểu hiện phối tử dưới dạng protein dung hợp với protein vỏ virus trên bề mặt ngoài của virus này. Phối tử này được chọn để có ái lực với thụ thể được biết là trình diện trên loại tế bào cần quan tâm. Ví dụ, Han và đồng tác giả, *Proc. Natl. Acad. Sci. USA* 92:9747-9751 (1995), báo cáo là virus bệnh bạch cầu chuột Moloney có thể được biến đổi để biểu hiện heregulin người được dung hợp với gp70, và virus tái tổ hợp này gây nhiễm một số tế bào ung thư vú ở người biểu hiện thụ thể yếu tố tăng trưởng biểu mô người. Nguyên tắc này có thể được mở rộng cho các cặp virus-tế bào đích khác, trong đó tế bào đích này biểu hiện thụ thể và virus này biểu hiện protein dung hợp chứa phối tử cho thụ thể bề mặt tế bào này. Ví dụ, thể thực khuẩn sợi có thể được thao tác di truyền để biểu hiện các đoạn kháng thể (ví dụ, FAB hoặc Fv) có ái lực liên kết đặc hiệu với thụ thể tế bào đã chọn bất kỳ thực sự. Mặc dù phần mô tả trên đây áp dụng chủ yếu cho vector virus, các nguyên tắc tương tự có thể được áp dụng cho các vector không phải virus. Các vector này có thể được thao tác di truyền để chứa các trình tự hấp thu đặc hiệu mà ưu tiên hấp thu bởi các tế bào đích đặc hiệu.

Vector liệu pháp gen có thể được phân phối in vivo bằng cách sử dụng cho một bệnh nhân đơn lẻ, thường là bằng cách sử dụng toàn thân (ví dụ, truyền trong tĩnh mạch, trong màng bụng, trong cơ, dưới da, hoặc trong sọ) hoặc sử dụng tại chỗ, như được mô tả dưới đây. Ngoài ra, các vector có thể được phân phối đến các tế bào ex vivo, như các tế bào được cấy bên ngoài từ một bệnh nhân đơn lẻ (ví dụ, các tế bào lympho, lymphocytes, hút tủy xương, sinh thiết mô) hoặc tế bào gốc tạo máu cho toàn năng, sau đó cấy lại các tế bào này vào bệnh nhân, thường là sau khi chọn lọc các tế bào đã kết hợp vector này.

Vector (ví dụ, retrovirut, adenovirut, liposom, v.v...) chứa các nucleaza và/hoặc các cấu trúc cho cũng có thể được sử dụng trực tiếp cho sinh vật để tải nạp các tế bào in vivo. Ngoài ra, ADN trần có thể được sử dụng. Việc sử dụng là bằng đường bất kỳ trong các đường thường được sử dụng để đưa một phân tử vào tiếp xúc với máu hoặc tế bào mô bao gồm, nhưng không giới hạn ở, tiêm, truyền, sử dụng tại chỗ và kỹ thuật xung điện. Các phương pháp sử dụng thích hợp các axit nucleic này là sẵn có và đã được biết rõ bởi người có kiến thức trung bình trong lĩnh vực này, và, mặc dù có nhiều hơn một đường có thể được dùng để sử dụng một chế phẩm cụ thể, một đường cụ thể thường có thể tạo ra phản ứng tức thì hơn và hiệu quả hơn so với một đường khác.

Các vector thích hợp để đưa polynucleotit (ví dụ, các trình tự cho mã hóa nucleaza và/hoặc sợi kép) vào được mô tả ở đây bao gồm các vector lentivirut không tích hợp (IDLV). Xem, ví dụ, Ory và đồng tác giả (1996) Proc. Natl. Acad. Sci. USA 93:11382-11388; Dull và đồng tác giả (1998) J. Virol. 72:8463-8471; Zuffery và đồng tác giả (1998) J. Virol. 72:9873-9880; Follenzi và đồng tác giả (2000) Nature Genetics 25:217-222; Công bố Patent Mỹ số 2009/0117617.

Các chất mang được dùng được quyết định một phần bởi chế phẩm cụ thể đang được sử dụng, cũng như bởi phương pháp cụ thể được sử dụng để sử dụng chế phẩm này. Theo đó, có nhiều dạng bào chế thích hợp khác nhau của các dược phẩm sẵn có, như được mô tả dưới đây (xem, ví dụ, Remington's Pharmaceutical Sciences, xuất bản lần thứ 17, 1989).

Rõ ràng là trình tự mã hóa nucleaza và các cấu trúc cho có thể được phân phối bằng cách sử dụng các hệ giống hoặc khác nhau. Ví dụ, các nucleaza và các trình tự cho có thể được mang bởi cùng một ADN MC. Ngoài ra, polynucleotit cho có thể được mang bởi MC, trong khi đó một hoặc nhiều nucleaza có thể được mang bởi plasmit chuẩn hoặc vector AAV. Ngoài ra, các vector khác nhau có thể được sử dụng bằng các đường giống hoặc khác

nhau (tiêm trong cơ, tiêm tĩnh mạch đùi, tiêm trong tĩnh mạch khác, sử dụng trong màng bụng và/hoặc tiêm trong cơ). Các vector có thể được phân phối đồng thời hoặc theo trật tự lần lượt bất kỳ.

Do đó, phần mô tả này bao gồm phương pháp điều trị in vivo hoặc ex vivo bệnh hoặc tình trạng mà có thể được điều trị bằng cách gắn xen gen chuyển mã hóa protein điều trị. Các chế phẩm này được sử dụng cho bệnh nhân là người với lượng hữu hiệu để đạt được nồng độ mong muốn của polypeptit điều trị trong huyết thanh hoặc cơ quan hoặc tế bào đích. Việc sử dụng có thể bằng phương pháp bất kỳ, trong đó polynucleotit này được phân phối đến tế bào đích mong muốn. Ví dụ, cả phương pháp in vivo và ex vivo đều được đề xuất. Tiêm trong tĩnh mạch vào tĩnh mạch cửa là phương pháp sử dụng được ưu tiên hơn. Các cách sử dụng in vivo khác bao gồm, ví dụ, tiêm trực tiếp vào các thùy của gan hoặc ống mật và tiêm trong tĩnh mạch xa gan, bao gồm qua động mạch gan, tiêm trực tiếp vào nhu mô gan, tiêm qua động mạch gan, và/hoặc tiêm ngược dòng qua các nhánh đường mật. Các cách sử dụng ex vivo bao gồm tải nạp in vitro các tế bào gan bị cắt bỏ hoặc các tế bào khác của gan, sau đó truyền các tế bào gan đã cắt ra, đã tải nạp vào lại mạch cửa, nhu mô gan hoặc nhánh của đường mật của bệnh nhân là người, xem ví dụ, Grossman và đồng tác giả, (1994) *Nature Genetics*, 6:335-341.

Lượng hữu hiệu của (các) nucleaza và trình tự cho cần được sử dụng sẽ thay đổi giữa các bệnh nhân và theo polypeptit điều trị cần quan tâm. Theo đó, các lượng hữu hiệu được quyết định tốt nhất bởi bác sỹ sử dụng các chế phẩm này và các liều thích hợp có thể được xác định dễ dàng bởi người có kiến thức trung bình trong lĩnh vực này. Sau khi cho phép đủ thời gian để tích hợp và biểu hiện (thường là 4-15 ngày, ví dụ), phân tích nồng độ huyết thanh hoặc nồng độ trong mô khác của polypeptit điều trị và so sánh với mức ban đầu trước khi sử dụng sẽ xác định được là lượng đang được sử dụng là quá thấp, trong khoảng đúng hay quá cao. Các chế độ thích hợp để sử dụng lần đầu và sử dụng tiếp cũng thay đổi, nhưng điển hình là sử dụng lần đầu sau đó sử dụng tiếp nếu cần. Các lần sử dụng tiếp theo có thể được thực hiện cách nhau các khoảng thời gian thay đổi, nằm trong khoảng từ hàng ngày đến hàng năm đến mỗi vài năm. Một người có kiến thức trung bình trong lĩnh vực này sẽ hiểu rõ là các kỹ thuật ức chế miễn dịch thích hợp có thể được đề nghị để tránh ức chế hoặc phong bế sự tải nạp bởi sự ức chế miễn dịch của vector phân phối, xem ví dụ, Vilquin và đồng tác giả, (1995) *Human Gene Ther.*, 6:1391-1401.

Các dạng bào chế để sử dụng ex vivo và in vivo bao gồm các huyền phù trong chất lỏng hoặc các chất lỏng được nhũ hóa. Các thành phần hoạt chất thường được trộn với các tá dược mà là dược dụng và tương thích với thành phần hoạt chất này. Các tá dược thích hợp bao gồm, ví dụ, nước, nước muối, dextroza, glyxerol, etanol hoặc các chất tương tự, và hỗn hợp của chúng. Thêm vào đó, chế phẩm này có thể chứa lượng nhỏ các chất phụ gia, như chất làm ẩm hoặc chất nhũ hóa, chất đệm pH, chất ổn định hoặc các chất khác mà tăng cường hiệu quả của dược phẩm này.

Tế bào

Cũng được mô tả ở đây là các tế bào và/hoặc dòng tế bào trong đó trình tự tăng cường BCL11A nội sinh được biến đổi bởi các nucleaza được mô tả ở đây (Bảng 1). Sự biến đổi này có thể là, ví dụ, so với trình tự kiểu đại của tế bào này. Tế bào hoặc dòng tế bào này có thể là dị hợp tử hoặc đồng hợp tử đối với biến đổi này. Các biến đổi đối với trình tự BCL11A có thể bao gồm gắn xen, mất đoạn và/hoặc kết hợp của chúng.

Tốt hơn, nếu biến đổi này nằm tại hoặc gần các vị trí liên kết nucleaza và/hoặc (các) vị trí phân cắt, ví dụ, trong khoảng 1-300 (hoặc giá trị bất kỳ nằm trong khoảng này) cặp bazơ ngược dòng hoặc xuôi dòng (các) vị trí phân cắt, tốt hơn nữa, nếu trong khoảng 1-100 cặp bazơ (hoặc giá trị bất kỳ nằm trong khoảng này) của mỗi bên của (các) vị trí liên kết và/hoặc phân cắt, thậm chí tốt hơn nữa nếu trong khoảng từ 1 đến 50 cặp bazơ (hoặc giá trị bất kỳ nằm trong khoảng này) trên mỗi phía của (các) vị trí liên kết và/hoặc phân cắt. Trong một số phương án, biến đổi này là ở hoặc ở gần vùng “+58” của trình tự tăng cường BCL11A, ví dụ, tại hoặc gần vị trí liên kết nucleaza được thể hiện trong hàng bất kỳ trong cột thứ nhất của bảng 1.

Tế bào hoặc dòng tế bào bất kỳ có thể được biến đổi, ví dụ tế bào gốc, ví dụ tế bào gốc phôi, tế bào gốc đa năng được cảm ứng, tế bào gốc tạo máu, tế bào gốc thần kinh và tế bào gốc trung mô. Ví dụ không giới hạn khác về các tế bào như được mô tả ở đây bao gồm các tế bào T (ví dụ, CD4+, CD3+, CD8+, v.v...); tế bào tua; tế bào B. Thế hệ con của tế bào gốc, bao gồm tế bào được biệt hóa một phần hoặc hoàn toàn, cũng được đề xuất (ví dụ, RBC hoặc tế bào tiền thân RBC). Ví dụ không giới hạn về các dòng tế bào khác bao gồm trình tự BCL11A biến đổi bao gồm COS, CHO (ví dụ, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11, CHO-DUKX, CHOK1SV), VERO, MDCK, WI38, V79, B14AF28-G3, BHK, HaK, NS0, SP2/0-Ag14, HeLa, HEK293 (ví dụ, HEK293-F, HEK293-H, HEK293-

T), và các tế bào perC6 cũng như các tế bào côn trùng như *Spodoptera fugiperda* (Sf), hoặc tế bào nấm như *Saccharomyces*, *Pichia* và *Schizosaccharomyces*.

Các tế bào như được mô tả ở đây là hữu ích trong việc điều trị và/hoặc phòng ngừa rối loạn, ví dụ, bằng các liệu pháp *ex vivo*. Các tế bào được biến đổi bằng nucleaza có thể được nuôi cấy và sau đó được đưa lại vào bệnh nhân bằng cách sử dụng các kỹ thuật chuẩn. Xem, ví dụ, Tebas và đồng tác giả (2014) *New Eng J Med* 370(10):901. Trong trường hợp của các tế bào gốc, sau khi truyền vào đối tượng, sự biệt hóa *in vivo* các tế bào tiền thân này thành các tế bào biểu hiện gen chuyển chức năng cũng xảy ra. Các được phẩm chứa các tế bào như được mô tả ở đây cũng được đề xuất. Thêm vào đó, các tế bào này có thể được bảo quản lạnh trước khi sử dụng cho bệnh nhân.

Bất kỳ trong số các tế bào hoặc dòng tế bào biến đổi được mô tả ở đây có thể thể hiện sự biểu hiện gamma globin tăng. Các chế phẩm như được phẩm chứa tế bào được biến đổi di truyền như được mô tả ở đây cũng được đề xuất

Ứng dụng

Các phương pháp và chế phẩm được mô tả ở đây là để làm biến đổi sự biểu hiện của protein, hoặc sửa trình tự gen bất thường mà mã hóa protein được biểu hiện trong bệnh di truyền, như bệnh hồng cầu hình liềm hoặc bệnh tan máu bẩm sinh. Do đó, các phương pháp và chế phẩm này được đề xuất để điều trị và/hoặc phòng ngừa các bệnh di truyền này. Kỹ thuật chỉnh sửa bộ gen, ví dụ của tế bào gốc, có thể được sử dụng để sửa gen bất thường, gắn xen gen kiểu đại, hoặc thay đổi sự biểu hiện của gen nội sinh. Bằng cách ví dụ không giới hạn, gen kiểu đại, ví dụ mã hóa ít nhất một globin (ví dụ, α và/hoặc β globin), có thể được gắn xen vào tế bào (ví dụ, vào trình tự tăng cường BCL11A nội sinh bằng cách sử dụng một hoặc nhiều nucleaza như được mô tả ở đây) để tạo ra các protein globin khiếm khuyết và/hoặc thiếu hụt trong tế bào và nhờ đó điều trị bệnh di truyền, ví dụ, bệnh huyết sắc tố, gây ra bởi sự biểu hiện globin lỗi. Ngoài ra hoặc thêm vào đó, kỹ thuật chỉnh sửa bộ gen có hoặc không sử dụng trình tự cho thích hợp, có thể chỉnh sửa gen nội sinh lỗi, ví dụ, chỉnh sửa đột biến điểm trong α - hoặc β -hemoglobin, để khôi phục sự biểu hiện của gen này và/hoặc điều trị bệnh di truyền, ví dụ bệnh hồng cầu hình liềm và/hoặc bất hoạt hoặc biến đổi (biểu hiện quá mức hoặc ức chế) gen điều hòa globin trực tiếp hoặc gián tiếp bất kỳ (ví dụ, bất hoạt gen điều hòa γ globin BCL11A hoặc KLF1 điều hòa BCL11A). Cụ

thể là, các phương pháp và chế phẩm theo sáng chế có ứng dụng trong việc điều trị hoặc phòng ngừa các bệnh huyết sắc tố.

Các nucleaza theo sáng chế được gắn đích vào vùng tăng cường BCL11A, được biết là cần thiết cho sự biểu hiện của BCL11A, và do đó điều hòa giảm sự biểu hiện gamma globin. Biến đổi của vùng tăng cường này có thể tạo ra tế bào hồng cầu có sự biểu hiện gamma globin tăng cường, và do đó có thể là hữu ích để điều trị hoặc phòng ngừa bệnh hồng cầu hình liềm hoặc bệnh tan máu bẩm sinh beta.

Phần ví dụ thực hiện sáng chế sau đây đề cập đến các phương án điển hình của sáng chế, trong đó nucleaza này chứa nucleaza ngón tay kẽm (ZFN). Rõ ràng là các phương án này chỉ nhằm mục đích minh họa và các nucleaza khác có thể được sử dụng, ví dụ TtAgo và hệ CRISPR/Cas, endonucleaza hướng đích (meganucleaza) với các vùng liên kết ADN được thao tác di truyền và/hoặc các thể dung hợp của dạng tự nhiên của các vùng liên kết ADN của endonucleaza hướng đích (meganucleaza) được thao tác di truyền và các vùng phân cắt khác loại và/hoặc các thể dung hợp của các meganucleaza và protein TALE.

Ví dụ thực hiện sáng chế

Ví dụ 1: Lắp ráp các nucleaza ngón tay kẽm

ZFN được lắp dựa vào gen BCL11A người và được thử nghiệm bằng các thử nghiệm CEL1 như được mô tả trong Miller và đồng tác giả (2007) *Nat. Biotechnol.* 25:778-785. ZFN đặc hiệu với vùng +58 của vùng tăng cường được tạo ra như đã mô tả. Các nucleaza được thể hiện dưới đây trong Bảng 1:

Bảng 1: Các cặp ZFN đặc hiệu với vùng tăng cường +58 BCL11A

SBS # (vị trí đích, 5'-3')	Thiết kế					
	F1	F2	F3	F4	F5	F6
Phần tử kết hợp bên trái						
46801 aaAGCAACtGTTAGCT TGCACTagacta (SEQ ID NO:1)	DQSNLRA (SEQ ID NO:2)	RPYTLRL (SEQ ID NO:3)	SGYNLEN (SEQ ID NO:4)	TSGSLTR (SEQ ID NO:5)	DQSNLRA (SEQ ID NO:2)	AQCCLFH (SEQ ID NO:6)
51446 aaAGCAACtGTTAGCT tgcactagacta (SEQ ID NO:1)	STGNLTN (SEQ ID NO:7)	TSGSLTR (SEQ ID NO:5)	DQSNLRA (SEQ ID NO:2)	AQCCLFH (SEQ ID NO:6)	KHÔNG XÁC ĐỊNH	KHÔNG XÁC ĐỊNH
51463 aaAGCAACtGTTAGCT tgcactagacta (SEQ ID NO:1)	STGNLTN (SEQ ID NO:7)	TSGSLTR (SEQ ID NO:5)	DQSNLRA (SEQ ID NO:2)	AQCCLFH (SEQ ID NO:6)	KHÔNG XÁC ĐỊNH	KHÔNG XÁC ĐỊNH

51484 aaAGCAACtGTTAGCT TGCACtagacta (SEQ ID NO:1)	DQSNLRA (SEQ ID NO:2)	RPYTLRL (SEQ ID NO:3)	SRGALKT (SEQ ID NO:8)	TSGSLTR (SEQ ID NO:5)	DQSNLRA (SEQ ID NO:2)	AQCCLFH (SEQ ID NO:6)
51856 aaAGCAACtGTTAGCT TGCACtagacta (SEQ ID NO:1)	DQSNLRA (SEQ ID NO:2)	RNFSLTM (SEQ ID NO:9)	SNGNLRN (SEQ ID NO:10)	TSGSLTR (SEQ ID NO:5)	DQSNLRA (SEQ ID NO:2)	AQCCLFH (SEQ ID NO:6)
51857 aaAGCAACtGTTAGCT TGCACtagacta (SEQ ID NO:1)	DQSNLRA (SEQ ID NO:2)	RNFSLTM (SEQ ID NO:9)	STGNLTN (SEQ ID NO:7)	TSGSLTR (SEQ ID NO:5)	DQSNLRA (SEQ ID NO:2)	AQCCLFH (SEQ ID NO:6)
51862 aaAGCAACtGTTAGCT TGCACtagacta (SEQ ID NO:1)	DQSNLRA (SEQ ID NO:2)	RNFSLTM (SEQ ID NO:9)	SSYNLAN (SEQ ID NO:11)	TSGSLTR (SEQ ID NO:5)	DQSNLRA (SEQ ID NO:2)	AQCCLFH (SEQ ID NO:6)
51477 aaAGCAACtGTTAGCT TGCACtagacta (SEQ ID NO:1)	DQSNLRA (SEQ ID NO:2)	RPYTLRL (SEQ ID NO:3)	SSSNLTN (SEQ ID NO:26)	TSGSLTR (SEQ ID NO:5)	DQSNLRA (SEQ ID NO:2)	AQCCLFH (SEQ ID NO:6)
51478 aaAGCAACtGTTAGCT TGCACtagacta (SEQ ID NO:1)	DQSNLRA (SEQ ID NO:2)	RPYTLRL (SEQ ID NO:3)	SSSNLGN (SEQ ID NO:27)	TSGSLTR (SEQ ID NO:5)	DQSNLRA (SEQ ID NO:2)	AQCCLFH (SEQ ID NO:6)
51487 aaAGCAACtGTTAGCT TGCACtagacta (SEQ ID NO:1)	DQSNLRA (SEQ ID NO:2)	RPYTLRL (SEQ ID NO:3)	SRSALRV (SEQ ID NO:28)	TSGSLTR (SEQ ID NO:5)	DQSNLRA (SEQ ID NO:2)	AQCCLFH (SEQ ID NO:6)
Phần tử kết hợp bên phải						
47923 caCAGGCTCCAGGAAG Ggtttggcctct (SEQ ID NO:12)	RSDHLTQ (SEQ ID NO:13)	QSGHLAR (SEQ ID NO:14)	QKGTIGE (SEQ ID NO:15)	QSSDLRS (SEQ ID NO:16)	RRDNLHS (SEQ ID NO:17)	KHÔNG XÁC ĐỊNH
51536 caCAGGCTCCAGGAAG Ggtttggcctct (SEQ ID NO:12)	RSDHLTQ (SEQ ID NO:13)	QSGHLAR (SEQ ID NO:14)	QKGTIGE (SEQ ID NO:15)	RHRDLRS (SEQ ID NO:18)	RRDNLHS (SEQ ID NO:17)	KHÔNG XÁC ĐỊNH
51949 caCAGGCTCCAGGAAG Ggtttggcctct (SEQ ID NO:12)	RNDHRTT (SEQ ID NO:19)	QKAHLIR (SEQ ID NO:20)	QKGTIGE (SEQ ID NO:15)	RGRDLRS (SEQ ID NO:21)	RRDNLHS (SEQ ID NO:17)	KHÔNG XÁC ĐỊNH
51990 caCAGGCTCCAGGAAG Ggtttggcctct (SEQ ID NO:12)	RSDHLTQ (SEQ ID NO:13)	QRAHLTR (SEQ ID NO:22)	QKGTIGE (SEQ ID NO:15)	HRNTLVR (SEQ ID NO:23)	RRDNLHS (SEQ ID NO:17)	KHÔNG XÁC ĐỊNH
51993 caCAGGCTCCAGGAAG Ggtttggcctct (SEQ ID NO:12)	RSDHLTQ (SEQ ID NO:13)	QRAHLTR (SEQ ID NO:22)	QSGTRNH (SEQ ID NO:24)	HRNTLVR (SEQ ID NO:23)	RRDNLHS (SEQ ID NO:17)	KHÔNG XÁC ĐỊNH
51979 caCAGGCTCCAGGAAG Ggtttggcctct (SEQ ID NO:12)	RSDHLTQ (SEQ ID NO:13)	QKAHLIR (SEQ ID NO:20)	QKGTIGE (SEQ ID NO:15)	RGRDLRS (SEQ ID NO:21)	RRDNLHS (SEQ ID NO:17)	KHÔNG XÁC ĐỊNH
51982 caCAGGCTCCAGGAAG Ggtttggcctct (SEQ ID NO:12)	RSDHLTQ (SEQ ID NO:13)	QKAHLIR (SEQ ID NO:20)	QSGTRNH (SEQ ID NO:24)	RGRDLRS (SEQ ID NO:21)	RRDNLHS (SEQ ID NO:17)	KHÔNG XÁC ĐỊNH
52015 caCAGGCTCCAGGAAG Ggtttggcctct (SEQ ID NO:12)	RNDHRTT (SEQ ID NO:19)	QKAHLIR (SEQ ID NO:20)	QKGTIGE (SEQ ID NO:15)	LKRTLKR (SEQ ID NO:25)	RRDNLHS (SEQ ID NO:17)	KHÔNG XÁC ĐỊNH

52032 caCAGGCTCCAGGAAG Ggtttggcctct (SEQ ID NO:12)	RSDHLTQ (SEQ ID NO:13)	QSGHLAR (SEQ ID NO:14)	QSGTRNH (SEQ ID NO:24)	QSSDLSR (SEQ ID NO:16)	RRDNLHS (SEQ ID NO:17)	KHÔNG XÁC ĐỊNH
51541 caCAGGCTCCAGGAAG Ggtttggcctct (SEQ ID NO:12)	RSDHLTQ (SEQ ID NO:13)	QSGHLAR (SEQ ID NO:14)	QKGTIGE (SEQ ID NO:15)	RHRDLSR (SEQ ID NO:18)	RRDNLHS (SEQ ID NO:17)	KHÔNG XÁC ĐỊNH
51519 caCAGGCTCCAGGAAG Ggtttggcctct (SEQ ID NO:12)	RSDHLTQ (SEQ ID NO:13)	QSGHLAR (SEQ ID NO:14)	QSGTRNH (SEQ ID NO:24)	QSSDLSR (SEQ ID NO:16)	RRDNLHS (SEQ ID NO:17)	KHÔNG XÁC ĐỊNH
51534 caCAGGCTCCAGGAAG Ggtttggcctct (SEQ ID NO:12)	RSDHLTQ (SEQ ID NO:13)	QSGHLAR (SEQ ID NO:14)	QKGTIGE (SEQ ID NO:15)	RGRDLSR (SEQ ID NO:21)	RRDNLHS (SEQ ID NO:17)	KHÔNG XÁC ĐỊNH
51535 caCAGGCTCCAGGAAG Ggtttggcctct (SEQ ID NO:12)	RSDHLTQ (SEQ ID NO:13)	QSGHLAR (SEQ ID NO:14)	QKGTIGE (SEQ ID NO:15)	RSRDLTR (SEQ ID NO:29)	RRDNLHS (SEQ ID NO:17)	KHÔNG XÁC ĐỊNH
51556 caCAGGCTCCAGGAAG Ggtttggcctct (SEQ ID NO:12)	RSDHLTQ (SEQ ID NO:13)	QSGHLAR (SEQ ID NO:14)	QKGTIGE (SEQ ID NO:15)	FRQTRAR (SEQ ID NO:30)	RRDNLHS (SEQ ID NO:17)	KHÔNG XÁC ĐỊNH

*51446 và 51463 khác nhau về trình tự cầu nối

Tất cả các ZFN đều được thử nghiệm về chức năng (hoạt tính phân cắt) và được phát hiện là hoạt động.

Ví dụ 2: Hoạt tính của ZFN ở tế bào K562 của người

Nói ngắn gọn, các tế bào K562 của người được nuôi cấy trong RPMI được bổ sung 10% FBS và 200.000 tế bào được chuyển nhiễm bằng nồng độ dưới điểm tối ưu là 25ng mỗi ADN plasmid mã hóa phần tử kết hợp ZFN bên trái và bên phải bằng Amara Nucleofector® theo hướng dẫn của nhà sản xuất (Bảng 2a). Thêm vào đó, các thử nghiệm được thực hiện với 25ng ZFN trái và 5ng ZFN phải (Bảng 2b). Thử nghiệm Cel-I (Surveyor™, Transgenomics) như được mô tả trong Perez và đồng tác giả (2008) Nat. Biotechnol. 26: 808-816 và Guschin và đồng tác giả (2010) Methods Mol Biol. 649:247-56), được sử dụng để phát hiện các biến đổi bị gây ra bởi ZFN của gen đích hai hoặc ba ngày sau khi chuyển nhiễm. Trong thử nghiệm này, sự khuếch đại PCR vị trí đích được tiếp tục bằng việc định lượng các đột biến gắn xen và/hoặc mất đoạn (“các tỉ lệ phần trăm các đột biến mất đoạn và thêm bazơ”) bằng cách giải trình tự. Việc giải trình tự sâu trên Illumina platform (“miSEQ”) được sử dụng theo hướng dẫn của nhà sản xuất để đo hiệu quả chỉnh sửa cũng như bản chất của các alen được tạo ra bằng cách chỉnh sửa. Các kết quả được thể hiện dưới đây trong Bảng 2, trong đó các số thể hiện tỉ lệ phần trăm hoạt tính NHEJ quan sát được:

Bảng 2a: Bảng ma trận trong các tế bào K562 (25ng mỗi ZFN)

	51949	51977	51979	51982	51990	51993	52015	52032		Trung bình
51856	22,2	21,0	27,7	23,2	21,0	23,9	17,0	26,7		22,8
51857	28,5	24,4	29,4	28,1	26,7	23,9	19,2	32,8		26,6
51862	15,9	16,3	15,5	15,3	11,7	20,8	13,9	22,4		16,5
51877	12,0	13,2	12,4	13,9	10,7	11,3	9,1	13,8		12,1
51879	14,2	15,1	14,9	13,8	12,0	14,4	10,8	16,8		14,0
Trung bình	18,6	18,0	20,0	18,8	16,4	18,9	14,0	22,5		
46801: 47923	9,3									

Bảng 2b: Bảng ma trận trong các tế bào K562 (25ng ZFN trái, 5ng ZFN phải)

	51949	51977	51979	51982	51990	51993	52015	52032		Trung bình
51856	10,9	12,6	16,9	13,4	11,0	12,4	8,5	15,5		12,7
51857	15,7	12,8	15,2	14,1	12,5	14,4	11,8	14,5		13,9
51862	13,0	11,4	13,9	10,4	14,3	12,2	12,2	14,3		12,7
51877	8,2	7,2	8,8	6,7	8,2	7,1	7,0	6,9		7,5
51879	8,5	7,9	11,5	8,3	8,4	7,1	4,9	11,2		8,5
Trung bình	11,3	10,4	13,3	10,6	10,9	10,6	8,9	12,5		
46801: 47923	7,3									

Các ZFN cũng được cấu trúc bằng cách sử dụng bốn cầu nối khác nhau giữa vùng liên kết ADN và vùng nucleaza (xem Công bố Patent Mỹ số 20150064789). Các trình tự cầu nối được thử nghiệm được thể hiện dưới đây, trong đó phần 'HTKIH' của trình tự này là đầu cùng carboxy của vùng liên kết ADN và phần 'ELEEK' là đầu cùng amino của vùng nucleaza. Phần được gạch chân là trình tự cầu nối giữa hai vùng này:

Các trình tự cầu nối

L7a: HTKIH LRGSQLVKSKSEAAAR ELEEK (SEQ ID NO:31)

L7c5: HTKIH LRGSISRARPLNPHP ELEEK (SEQ ID NO:32)

L0: HTKIH LRGSISRARPLNPHP ELEEK (SEQ ID NO:33)

L8c4: HTKIH LRGSYAPMPPLALASP ELEEK (SEQ ID NO:34)

Trong các thử nghiệm này, các cầu nối L0 hoặc L8c4 được thử nghiệm trên phân tử kết hợp bên phải kết hợp với các cầu nối L7c5 hoặc L7a trên phân tử kết hợp bên trái. Các

kết hợp của các ZFN được thử nghiệm về hoạt tính phân cắt trong các tế bào K562, và các kết quả (tỉ lệ phần trăm hoạt tính NHEJ) được thể hiện dưới đây trong Bảng 3.

Bảng 3: Hoạt tính ZFN làm thay đổi các cầu nối vùng

		L7c5	L7c5	L7c5	L7a	L7a	trung bình
	SBS	51856	51857	51862	46801	51446	
L0	51949	6,4	9,2	7,3	4,8	10,6	7,66
L0	51977	8,8	8,9	5,6	5,0	8,7	7,41
L0	51979	11,4	10,5	10,8	4,5	11,3	9,70
L0	51982	8,3	9,7	9,3	4,9	10,5	8,56
L0	51990	10,3	11,0	7,5	5,3	10,2	8,86
L0	51993	10,7	13,5	7,3	5,1	10,9	9,52
L0	52015	8,5	8,6	7,2	5,7	6,7	7,34
L0	52032	11,5	11,3	7,0	5,0	14,8	9,91
L0	47923	7,6	7,6	5,0	4,0	11,5	7,14
L8c4	52075	10,9	10,9	7,2	5,6	11,7	9,27
L8c4	52103	9,9	8,9	8,5	3,8	10,9	8,37
L8c4	52105	16,7	13,0	16,3	6,3	15,2	13,49
L8c4	52108	9,7	7,5	6,5	4,2	12,0	7,99
L8c4	52116	12,0	15,7	9,8	4,0	13,4	10,99
L8c4	52119	13,1	9,8	8,3	4,5	10,3	9,19
L8c4	52141	15,2	7,2	6,9	3,4	13,2	9,17
L8c4	52158	9,3	10,3	7,4	3,6	12,5	8,64
L8c4	51536	13,3	12,1	11,3	8,0	15,4	12,04
	trung bình- >	10,76	10,32	8,28	4,88	11,65	

Ví dụ 3: Hoạt tính của ZFN trong các tế bào CD34+

ZFN như được mô tả ở đây cũng được thử nghiệm trong các tế bào CD34+ của người. Để tải nạp CD34+, thiết bị BTX ECM830 với cuvet có khoảng cách giữa hai điện cực là 2 mm được sử dụng. Các tế bào CD34+ của người được nuôi trong môi trường x-vivo10 (Lonza) với 1xCC110 (Stem cell Technology) trong các đĩa không được xử lý nuôi cấy mô. Các tế bào được đếm và được thu gom bằng cách ly tâm ở tốc độ 1200 vòng/phút trong 10 phút ở nhiệt độ phòng. Các tế bào được rửa 1-2x với PBS ở nhiệt độ phòng. 200.000 tế bào được sử dụng cho mỗi lần chuyển nhiễm, và chúng được tạo huyền phù lại trong 100µL dung dịch BTexpress. Đối với các thử nghiệm CD34+, các ARN mã hóa các ZFN được sử dụng hơn là ADN. ARN được tạo ra bằng cách sử dụng mMessageMachine T7 Ultra Kit (Ambion). 500ng ARN mã hóa mỗi ZFN được bổ sung cho mỗi chuyển nhiễm và hỗn hợp được chuyển sang cuvet. Ngay sau khi chuyển, hỗn hợp này được chiếu xung

điện ở 250V trong 5 mili giây. Môi trường đã được làm ấm trước được bổ sung vào cuvet và các tế bào cộng môi trường được chuyển sang đĩa 48 lỗ không được xử lý nuôi cấy mô và sau đó được ủ ở 37°C.

Sau hai hoặc ba ngày, sau đó các tế bào được phân tích bộ gen bằng cách sử dụng Illumina MiSeq. Để định lượng tỉ lệ phần trăm các alen được chỉnh sửa, vùng cần quan tâm của bộ gen được khuếch đại PCR bằng cách sử dụng các đoạn mồi mà bổ sung các trình tự tiếp hợp giải trình tự Illumina chuẩn. Nhóm thứ hai gồm 13 vòng PCR được thực hiện để bổ sung các trình tự tiếp hợp mã vạch và cầu vào cả hai đầu. Việc giải trình tự được thực hiện trên Illumina MiSeq theo quy trình của nhà sản xuất để giải trình tự đơn vị siêu sao chép. MiSeq tạo ra hai đoạn đọc từ hai đầu của cùng một trình tự, mà được hợp nhất và được cắt bằng trình tự tiếp hợp bằng cách sử dụng phần mềm sắp xếp thẳng hàng chuẩn. Sau đó các đoạn đọc được tách bởi mẫu qua các cặp trình tự mã vạch bằng cách sử dụng các bản gốc tùy chỉnh. Các trình tự đơn vị siêu sao chép sau đó được so sánh toàn bộ với trình tự tham chiếu bằng cách thực hiện thuật toán Needleman–Wunsch (Needleman and Wunsch (1970). *Jour Mol Bio* 48 (3): 443–53). Các khoảng cách hoặc sự gắn xen trong trình tự thẳng hàng được đếm là % hiện tượng NHEJ, và được so sánh với trình tự mẫu đối chứng chưa được xử lý để xác định tỉ lệ nền đặc hiệu trình tự. Các kết quả được thể hiện dưới đây trong Bảng 4.

Bảng 4: Hoạt tính ZFN trong các tế bào CD34+

	51949	51977	51979	51982	51990	51993	52015	52032	51536	Trung bình
51856									không phát hiện được	29,2
	38,4	11,9	26,9	23,9	37,3	37,0	26,4	32,1		
51857	52,1	17,2	39,8	30,9	51,0	43,9	35,2	53,9	40,0	40,4
51862	18,4	5,2	13,0	12,4	26,6	20,7	13,7	16,2	17,7	16,0
51877	19,6	6,0	8,9	9,6	24,0	18,3	11,6	19,4	13,7	14,6
51879	36,4	10,6	21,8	19,5	39,2	30,7	21,8	34,4	22,7	26,4
51446									Không phát hiện được	45,4
	60,1	22,6	41,2	43,6	57,8	50,4	39,7	47,6		
Trung bình	37,5	12,3	25,3	23,3	39,3	33,5	24,7	33,9	23,5	

Khi lượng mARN đầu vào tăng được sử dụng cho tập hợp các cặp đại diện đã chọn (1µg mỗi ZFN), các lượng bổ sung của sự cắt được quan sát thấy, như được thể hiện trong Bảng 5.

Bảng 5A và 5B: Nồng độ ZFN tăng dần đến hoạt tính tăng

Bảng 5A

0,5ug : 0,5ug	
46801:47923	1,6
51556:51484	17,6
1ug : 1ug	
46801:47923	1,5
51536:51446	49,0
52032:51446	65,7
52032:51857	60,8
51979:51446	60,0
51979:51857	47,5
51536:51857	55,7
GFP	không có trình tự nào

Bảng 5B

	51536	51541	51556	51519	51534	51535
51446	83,72	87,79	74,11	81,52	84,21	82,22
51484	75,21	84,34	64,20	67,88	79,50	74,22
51463						Không phát hiện được
	82,71	85,24	74,00	78,81	85,53	
51477	72,02	85,11	63,74	70,42	80,96	78,30
51478	72,33	82,38	58,17	66,22	75,20	70,90
51487	66,27	83,26	64,11	61,97	68,99	70,42

Các lần chuyển nhiễm trên đây được thực hiện chủ yếu trong các điều kiện lượng mARN đầu vào không bão hòa để cho phép so sánh tốt nhất hoạt tính của các kết hợp ZFN khác nhau. Để thử nghiệm lượng biến đổi tối đa có thể thu được tại hoặc gần lượng mARN đầu vào bão hòa, chúng tôi chuyển nhiễm lượng tăng dần của các cấu trúc 2a hoặc kết hợp hai ZFN trong một mARN vào các tế bào CD34+ bằng cách sử dụng kỹ thuật xung điện BTX, và các kết quả được thể hiện dưới đây trong Bảng 6.

Bảng 6: Biến đổi của trình tự tăng cường BCL11A với lượng mARN đầu vào tăng dần

2A:	2000ng	4000ng	8000ng
51463 2a 51949	59,60	79,83	82,69
51463 2a 51990	67,98	80,17	77,30
51463 2a 52032	58,58	80,36	83,26
51857 2a 51949	61,22	73,89	78,26
51857 2a 51990	60,63	70,55	74,40
51857 2a 52032	55,84	71,17	74,29
euf 51446 2a 51536	56,36	66,69	74,85
euf 46801 2a 47923	51,15	72,20	75,97

Số liệu trong Bảng 6 thể hiện sự biến đổi rất cao của vùng đích ZFN khi lượng mARN đầu vào tăng dần đối với tất cả các kết hợp được thử nghiệm.

Tương tự với các thử nghiệm được thực hiện trong thí nghiệm 2 phân tích hoạt tính với các ZFN chứa các cầu nối biến đổi, tác động của các cầu nối trên hoạt tính cũng được thử nghiệm trong các tế bào CD34⁺. Như trên đây, các lượng mARN thay đổi (500, 1000 hoặc 2000ng cho mỗi mẫu) được sử dụng để phân phối các ZFN trong các thử nghiệm này. Bảng 7 thể hiện tác động của đặc tính cầu nối trên hoạt tính của cặp ZFN.

Bảng 7: Tác động của đặc tính cầu nối trên hoạt tính ZFN trong các tế bào CD34⁺

	500ng	51463	51857	1000ng	51463	51857	2000ng	51463	51857
L0	51949	29,9	31,0	51949	48,8	46,7	51949	44,5	43,5
	51979	26,3	23,8	51979	39,2	40,0	51979	36,3	35,1
	51990	28,6	28,1	51990	39,8	41,8	51990	36,1	36,3
	52032	24,7	25,5	52032	39,6	39,7	52032	37,2	33,6
L8c4	51536	23,0	19,1	51536	36,0	31,0	51536	35,7	30,6
	52075	22,1	21,2	52075	34,7	32,0	52075	35,2	27,5
	52105	21,9	19,1	52105	37,2	32,8	52105	36,2	31,6
	52116	14,7	13,9	52116	27,6	28,8	52116	28,4	28,9
	52158	22,2	21,1	52158	36,3	34,9	52158	41,7	38,2
	500ng	51446		1000ng	51446		2000ng	51446	
	51536	19,8		51536	35,5		51536	28,7	

Ví dụ 4: Sự biệt hóa các tế bào CD34⁺ được chỉnh sửa và sự biểu hiện hemoglobin

Để thử nghiệm tác động trên sự biểu hiện gamma globin tương đối, các mARN mã hóa mẫu đại diện của các cặp ZFN được đưa vào các tế bào CD34⁺ (thu được từ những người hiến tặng tự nguyện khỏe mạnh) bằng kỹ thuật chuyển gen bằng xung điện (nucleofection) BTX theo hướng dẫn của nhà sản xuất. Sau đó các tế bào được biệt hóa thành các hồng cầu. Nói ngắn gọn, các tế bào CD34⁺ được tinh chế bằng cách sử dụng Ficoll-Paque (GE Healthcare) và vi hạt CD34⁺ (Miltenyi Biotec) theo hướng dẫn của nhà sản xuất. Các tế bào CD34⁺ được nuôi cấy trong Iscove's MDM với BIT 95000 (StemCell Technologies) khi có mặt các yếu tố sinh trưởng. Các tế bào được biệt hóa thành dòng hồng cầu sử dụng mô hình nuôi cấy lỏng 3 bước. Trong 6 ngày đầu tiên (pha thứ nhất), các tế

bào CD34⁺ được nuôi cấy bằng SCF (100ng/ml), Flt3-L (100ng/ml), và IL-3 (20ng/ml). Sau đó các tế bào đã nuôi cấy được ủy nhiệm và biệt hóa thành dòng hồng cầu (pha thứ hai) với Epo (2U/ml) và SCF (50ng/ml). Xem, Giarratana và đồng tác giả (2011) *Blood* 118(19):5071-9.

Để phân tích sự biểu hiện gamma globin tương đối, các tỉ lệ của mRNA mã hóa gamma globin, alpha globin và beta globin sau khi xử lý ZFN được xác định ở thời điểm 14 ngày sau khi bắt đầu biệt hóa bằng phân tích RT-PCR. Phân tích này được thực hiện bằng phân tích Taqman® chuẩn, theo quy trình và sử dụng các thử nghiệm đặc hiệu gen được cung cấp bởi nhà sản xuất (Applied Biosystems). Các mức nồng độ tương đối của gamma globin (HBG) được chuẩn hóa bằng mức độ biểu hiện của alpha (HBA) hoặc beta globin (HBB) trong đó tỉ lệ này được so sánh với tỉ lệ gamma/alpha hoặc gamma/beta ở các tế bào đối chứng.

Số liệu được thể hiện dưới đây trong Bảng 8, và chứng minh rằng so với các tế bào được xử lý bằng plasmit mã hóa GFP, có sự gia tăng biểu hiện gamma globin trong các tế bào được xử lý ZFN.

Bảng 8: Thay đổi về sự biểu hiện gamma globin so với alpha hoặc beta globin trong các tế bào CD34⁺ được chỉnh sửa

Cặp ZFN	HBG/HBA	HBG/HBB
46801/47293	2,7	4,4
51446/51536	5,0	8,1
51463/51536	8,0	10,5
51484/51536	3,8	6,4
GFP	1,6	2,8

Sự biệt hóa hồng cầu *in vitro* của các tế bào CD34⁺ được chuyển nhiễm tại hoặc gần các nồng độ mRNA bão hòa sau đó phân tích RT-PCR sự biểu hiện globin như được mô tả trên đây thể hiện sự hoạt hóa gamma-globin bởi tất cả các cặp ZFN đã chọn mà gần đích trình tự tăng cường BCL11A là rất hiệu quả so với mẫu đối chứng được chuyển nhiễm GFP mRNA (Fig.1). Số liệu được thể hiện dưới đây trong Bảng 9.

Bảng 9: Tỷ lệ biểu hiện gamma globin người tăng

	<u>HBG/HBB</u>		
<u>mARN được sử dụng</u>	<u>2µg</u>	<u>4µg</u>	<u>8µg</u>
<u>51463-2a-51949</u>	<u>10,33</u>	<u>17,56</u>	<u>21,84</u>
<u>51463-2a-51990</u>	<u>12,75</u>	<u>16,05</u>	<u>16,96</u>
<u>51463-2a-52032</u>	<u>9,94</u>	<u>17,91</u>	<u>18,97</u>
<u>51857-2a-51949</u>	<u>10,12</u>	<u>14,16</u>	<u>12,21</u>
<u>51857-2a-51990</u>	<u>10,36</u>	<u>12,37</u>	<u>16,47</u>
<u>51857-2a-52032</u>	<u>8,57</u>	<u>15,18</u>	<u>19,01</u>
<u>51446-2a-51536</u>	<u>9,45</u>	<u>10,43</u>	<u>14,76</u>
<u>46801-2a-47923</u>	<u>8,86</u>	<u>12,15</u>	<u>13,10</u>
<u>GFP</u>	<u>1,65</u>		

Ví dụ 5: Hoạt tính của ZFN trong các tế bào CD34+ sử dụng thiết bị xung điện BTX, các mARN riêng biệt và các mARN duy nhất

Hoạt tính của hai cặp ZFN được thử nghiệm trong các tế bào CD34+ người được huy động được phân lập từ máu ngoại vi của người và trong các tế bào CD34+ được phân lập từ tủy xương. Nói ngắn gọn, các tế bào CD34+ được phân lập từ người hiến tặng khỏe mạnh như sau. Các mẫu gan bạch cầu được loại bỏ tiểu cầu bằng cách ly tâm tốc độ chậm và loại bỏ dịch nổi. Sau khi loại bỏ tiểu cầu, các tế bào được đánh dấu bằng vi hạt từ kháng CD34 (Miltenyi Biotec, Đức) và được chọn lọc dương bằng cách sử dụng Hệ thống tách tế bào Miltenyi CliniMACS Plus. Sau khi chọn lọc, phân đoạn dương (được làm giàu CD34+ HSPC) được rửa và được tạo huyền phù lại trong môi trường nuôi cấy (nghĩa là, môi trường X-VIVO 10 được bổ sung 2mM L-glutamin, 100ng/ml mỗi thành phần bao gồm phối tử tyrosin kinaza 3 tương tự FMS (Flt-3L), yếu tố tế bào gốc (SCF), và thrombopoietin (TPO)) ở nồng độ 1×10^6 tế bào/mL, và được chuyển vào các túi nuôi cấy VueLife (Saint-Gobain, Gaithersburg, MD) và được ủ ở 37°C/5% CO₂. Để tinh chế các tế bào CD34+ từ tủy xương, các mẫu thu được được loại bỏ tế bào hồng cầu (RBC) bằng phương pháp kết tủa tinh bột hydroxyetyl. Sau khi loại bỏ RBC, các tế bào được đánh dấu bằng các vi hạt từ kháng CD34 (Miltenyi Biotec, Đức) và được chọn lọc dương bằng cách sử dụng Miltenyi CliniMACS Prodigy. Sau khi chọn lọc, phân đoạn dương (được làm giàu CD34+ HSPC) được rửa và được tạo huyền phù lại trong môi trường nuôi cấy (nghĩa là, môi trường X-

VIVO 10 được bổ sung 2mM L-glutamin, 100ng/ml mỗi thành phần bao gồm phối tử tyrosin kinaza 3 tương tự FMS (Flt-3L), yếu tố tế bào gốc (SCF), và thrombopoietin (TPO)) ở nồng độ 1×10^6 tế bào/mL, và được chuyển vào các túi nuôi cấy VueLife (Saint-Gobain, Gaithersburg, MD) và được ủ ở $37^{\circ}\text{C}/5\% \text{CO}_2$.

Các cặp được sử dụng là 51446/51536 (cặp A) và SBS51857/51949 (cặp B). Các ZFN được thử nghiệm dưới dạng mARN được đưa vào trong tế bào. Để chuyển nhiễm thiết bị BTX được sử dụng. Nói ngắn gọn, 200.000 tế bào mỗi mẫu được tạo huyền phù trong dung dịch xung điện BTXpress (BTX) và được trộn với ARN này. Sau đó hỗn hợp này được tạo xung trong 4 mili giây ở 250 vôn và được đưa vào các điều kiện sốc lạnh (30°C qua đêm) trước khi để các tế bào phục hồi ở 37°C . Phân tích hoạt tính ZFN được thực hiện hai hoặc ba ngày sau khi chuyển nhiễm. Các ZFN được thử nghiệm dưới dạng một loại mARN duy nhất, trong đó trình tự peptit tự cắt 2a được sử dụng giữa hai trình tự mã hóa ZFN, và số liệu này được thể hiện trong Fig.2. Thêm vào đó, các điều kiện giống như vậy được sử dụng để thử nghiệm 1 cặp ZFN, trong đó sự cung cấp các mARN này là dưới dạng hai loại riêng biệt, trong đó mỗi mARN mã hóa một ZFN. Bảng 10a và Bảng 10b dưới đây thể hiện các kết quả về hoạt tính (%NHEJ hoặc tỉ lệ phần trăm đột biến mất đoạn và thêm bazơ) cho phương pháp một mARN duy nhất so với hai mARN (số liệu mô tả là từ vài thử nghiệm).

Bảng 10a: So sánh loại mARN duy nhất với loại mARN đôi (% các đột biến mất đoạn và thêm bazơ)

mARN được sử dụng	2,0 μg	4,0 μg
51857_2a_51949, thử nghiệm #1	56,0	64,7
51857_2a_51949, thử nghiệm #2	61,2	73,9
51857 + 51949	83,1	80,5

Bảng 10b: So sánh loại mARN duy nhất với loại đôi (% các đột biến mất đoạn và thêm bazơ)

mARN được sử dụng	0,5 μg	2,0 μg	4,0 μg
51857_2a_51949, thử nghiệm #3	32,0	51,5	58,7
51857 + 51949, thử nghiệm #3	40,0	58,4	63,1

Chúng tôi cũng thử nghiệm sự biểu hiện của gamma globin người và beta globin người bằng cách sử dụng TaqMan® theo các quy trình chuẩn. Các kết quả sau đó được chuẩn hóa thành tỉ lệ của HBG (gamma) so với HBB (beta), và được mô tả trong Fig.3, lại so sánh hai cặp là ZFN: 51446/51536 (cặp A) và SBS51857/51949 (cặp B). Thêm vào đó, sự biểu hiện của HBG và HBB cũng được đo so sánh sự cung cấp cặp ZFN dưới dạng một mARN duy nhất, trong đó các trình tự mã hóa mỗi ZFN trong cặp này được phân tách bởi trình tự peptit tự cắt 2a, với các điều kiện trong đó mARN mã hóa mỗi ZFN được cung cấp riêng. Số liệu này chứng minh rằng trong các điều kiện này, cặp B, SBS51857/51949, là có hoạt tính cao hơn trong việc phân cắt đích BCL11a và trong việc làm tăng sự biểu hiện HBG.

Các cặp ZFN được thử nghiệm trong các tế bào CD34+ thu được từ tủy xương và cả hai cặp lại được phát hiện là có hoạt tính (xem Fig.4a). Hoạt tính đối với cặp SBS51857/51949 cũng được thử nghiệm, trong đó các ZFN được cung cấp trên một mARN duy nhất với 2a hoặc dưới dạng hai mARN riêng biệt như được mô tả trên đây (Fig.4b). Cặp SBS51857/51949 thể hiện hoạt tính cao hơn so với cặp 51446/51536.

Ví dụ 7: Phân tích tính đặc hiệu

Phân tích bắt giữ không bị chệch:

Thử nghiệm bắt giữ này dựa trên quan sát là việc đưa đồng thời nucleaza và trình tự cho ADN kép vào tế bào đích gây ra sự “bắt giữ” trình tự cho vào một đoạn của vị trí đứt gãy bộ gen tạo thành bằng con đường sửa NHEJ ADN (Orlando và đồng tác giả, (2010) *Nucleic acids Res.* 38(15) e152; Gabriel, R. và đồng tác giả (2011). *Nat Biotechnol.* 29: 816-23.). Chú ý là hiện tượng bắt giữ này không được điều khiển tương đồng (thực vậy, trình tự cho ADN kép này không chứa sự tương đồng bất kỳ với bộ gen người). Chú ý thêm là mARN mã hóa các ZFN không thể bị bắt giữ trong vị trí đứt gãy ADN, chỉ có trình tự cho ADN kép là có thể. Khi bị bắt giữ, bộ gen kép này trình diện nhân thường trực của hiện tượng phân cắt này. Sau khi phân lập ADN bộ gen, các vị trí bắt giữ có thể được xác định bằng cách kéo dài đoạn mỗi từ trình tự cho vào trình tự bộ gen bên sườn, sau đó nối tiếp hợp, PCR, và giải trình tự các mối nối trình tự cho-bộ gen thu được.

Nói ngắn gọn, các nghiên cứu phân tích bắt giữ được thực hiện trong dòng tế bào K562 để tối đa hóa sự phân phối trình tự cho, biểu hiện ZFN, và bắt giữ trình tự cho vào

vị trí DSB; các tế bào này được xử lý bằng xung điện với mARN mã hóa ZFN và trình tự cho kép oligonucleotit. Riêng biệt là, các tế bào CD34⁺ thu được từ BM và PB được xử lý bằng xung điện sử dụng thiết bị Maxcyte như được mô tả trên đây với mARN mã hóa ZFN. Oligonucleotit cho kép đã sử dụng được thể hiện dưới đây (SEQ ID NO: 30 và 31):

5' NNNNAGTAGTGTGTGCCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCNN
NNNAGTGTGGAAAATCTCTAGCAG 3'

3' TCATCACACACGGGCAGACAACACACTGAGACCATTGATCTCTAGGGAGTCTGGGAAAATCAGNNNNNT
CACACCTTTTAGAGATCGTCNNNN 5'

NNNN ở đầu 5' của mỗi sợi thể hiện tetrame sợi đơn, ngẫu nhiên nhô ra, trong khi NNNN được gạch chân thể hiện trình tự kép ngẫu nhiên đóng vai trò là trình tự mã vạch để phân biệt giữa các hiện tượng tích hợp giống nhau khác. Các mẫu lặp ba được tạo ra cho mỗi tổ hợp của oligo và mARN. Vào ngày thứ 7 và 14, ADN bộ gen sau chuyển nhiễm được phân lập (Kit Máu và Mô Qiagen DNeasy), và 1 µg (330000 bộ gen / bản sao) được sử dụng làm đầu vào cho quy trình khuếch đại. Sau đó mẫu được xử lý chủ yếu như đã mô tả (Paruzynski, A. và đồng tác giả (2010). *Nat. Protocols*. 5:1379-1395). Các đơn vị siêu sao chép được tinh chế bằng cách sử dụng Kit tinh chế PCR QIAquick (Qiagen), và được khuếch đại bằng PCR để đưa các trình tự mã vạch và trình tự tiếp hợp vào để giải trình tự sâu trên nền Illumina. Các sản phẩm cuối cùng được định lượng, gộp lại và giải trình tự trên Thiết bị MiSeq (Illumina) bằng cách sử dụng kit giải trình tự 300 chu trình v2 với hai đoạn đọc 150 bp từ hai đầu của cùng một trình tự và các đoạn đọc chỉ số kép 8 bp để phát hiện các trình tự mã vạch trên mỗi đầu của đơn vị siêu sao chép này. Nỗ lực này tạo ra tập hợp các locut xa đích ứng viên mà sau đó được xác định kiểu gen trong các tế bào CD34 thu được từ BM hoặc PB. Các kết quả được xác định cho cặp A (SBS#51446/51536) được thể hiện trong Fig.10a, trong khi đó các kết quả cho cặp B (SBS#51857/51949) được thể hiện trong Fig.10b. Số liệu được tóm tắt thêm dưới đây trong Bảng 11.

Bảng 11: Phân tích xa đích cho cặp A và cặp B, CD34⁺ từ tủy xương

A	250ug/ml	<u>63%</u>	21	9%
B	125ug/ml	<u>63%</u>	9	3%
B	250ug/ml	<u>74%</u>	20	17%

Các nghiên cứu tương tự được thực hiện trên các tế bào CD34⁺ thu được từ PB, và Bảng 12 dưới đây tóm tắt các kết quả được tìm ra.

Bảng 12: Phân tích xa đích cho cặp A và cặp B, CD34+ từ máu ngoại vi

ZFN	Liều	Trúng đích	# xa đích	Hoạt tính tích lũy
St 4	120ug/ml	60%	11	4%
St 5	120ug/ml	56%	10	3%

Ví dụ 6: Hoạt tính của ZFN trong các tế bào CD34+ sử dụng thiết bị xung điện Maxcyte, các mARN riêng biệt và các mARN duy nhất

Tiếp theo, tác giả sáng chế thực hiện các thử nghiệm quy mô lớn hơn bằng cách sử dụng các cặp ZFN trong các tế bào CD34+ sử dụng thiết bị xung điện Maxcyte GT. Các tế bào CD34+ huy động được được phân lập từ máu ngoại vi của người hiến tặng bình thường như được mô tả trên đây (PB) được thử nghiệm như sau: các tế bào (3 triệu mỗi mẫu) được tạo huyền phù lại trong đệm RT Maxcyte EP đến nồng độ cuối cùng là 30×10^6 tế bào/mL. Tế bào được trộn với mARN và được xử lý bằng xung điện bằng cách sử dụng chương trình được xác định bởi nhà sản xuất. Các tế bào được để phục hồi nhanh ở 37°C trong hai mươi phút, sau đó được pha loãng và được đưa vào các điều kiện sốc lạnh (30°C qua đêm) trước khi để cho các tế bào phục hồi ở 37°C. Hoạt tính được phân tích hai đến ba ngày sau đó. Các thử nghiệm được thực hiện với cả cặp A và cặp B như trên đây, trong đó mỗi cặp được đưa lên một mARN duy nhất (Fig.5A). Cặp ZFN B (SBS51857/51949) cũng được thử nghiệm dưới dạng một mARN duy nhất hoặc hai mARN riêng biệt như được mô tả trên đây (Fig.5b). Trong các thử nghiệm này, cặp B thể hiện hoạt tính cao nhất.

Đối với các tế bào thu được từ tủy xương, việc chỉnh sửa thành công đòi hỏi nhiều mARN hơn, nhưng hoạt tính cao được quan sát thấy (xem Fig.6). Các tế bào được phân tích về sự biểu hiện HBG và HBB như được mô tả trên đây và tỉ lệ HBG/HBB được so sánh với tỉ lệ phần trăm hoạt tính gây đột biến mất đoạn hoặc thêm bazơ vào ngày 0 (Fig.7). Có sự phù hợp giữa các mức biểu hiện HBG cao hơn và hoạt tính gây đột biến mất đoạn hoặc thêm bazơ cao hơn.

Nhiều phân tích hơn được thực hiện bằng cách sử dụng quy trình Maxcyte (Fig.8 và 9), trong đó đã phát hiện ra là mặc dù tần số của các đột biến mất đoạn hoặc thêm bazơ không phải luôn giống nhau trong mỗi thử nghiệm, hoạt tính tương đối (ví dụ, cấu trúc 2a SBS51857/51949 cao hơn so với các mARN riêng biệt) được quan sát thấy trong mỗi lần chạy thử nghiệm.

Ví dụ 7: Ghép các tế bào CD34+ người đã chỉnh sửa vào chuột

Như được mô tả trên đây, các tế bào người CD34+ được xử lý bằng các mARN mã hóa các ZFN đặc hiệu yếu tố tăng cường +58 và sau đó được ghép vào chuột NSG. Các tế bào CD34+ thu được từ người tình nguyện khỏe mạnh. Trong một số trường hợp, các chiến lược huy động CD34+ được thực hiện, bằng cách sử dụng G-CSF (Neupogen®) hoặc G-CSF + Plerixafor (Mozobil®) trước khi tách chiết. G-CSF được sử dụng hàng ngày trong bốn ngày trước khi tách chiết theo hướng dẫn của nhà sản xuất, và nếu Plerixafor được sử dụng, nó được sử dụng vào buổi tối cuối cùng trước khi thu hoạch, lại theo hướng dẫn của nhà sản xuất. Việc tách chiết được thực hiện bằng phương pháp chuẩn. Các tế bào CD34+ được làm giàu từ sản phẩm của quy trình chiết tách bạch cầu ra khỏi mẫu máu được làm giàu PBMC được huy động bằng cách sử dụng hệ thống Miltenyi CliniMACs bằng các phương pháp chuẩn và theo hướng dẫn của nhà sản xuất.

Các mARN được gắn mũ và được poly-adenyl hóa mã hóa ZFN được tổng hợp bằng cách sử dụng siêu kit Ambion mMessage mMachine® T7 như được hướng dẫn bởi nhà sản xuất và sau đó được xử lý bằng xung điện vào các tế bào CD34+ bằng cách sử dụng hệ thống Maxcyte GT hoặc máy tạo xung BTX ECM830, đều theo hướng dẫn của nhà sản xuất.

Chuột NOD.Cg-*Prkdc^{scid} Il2rg^{tm1Wjl}*/SzJ được sử dụng để nhận cấy ghép CD34+. Một ngày (16- 24 giờ) trước khi cấy, chuột được chiếu xạ gây chết (300 RAD). Các tế bào CD34+ đã được xử lý ZFN trên đây được cấy vào chuột đã chiếu xạ bằng cách tiêm tĩnh mạch đuôi, trong đó 1 triệu tế bào trong 0,5mL PBS-0,1%BSA được sử dụng cho mỗi chuột.

Đối với thử nghiệm này, các tế bào CD34+ được xử lý bằng xung điện với các mARN mã hóa cặp 46801/47923. Các gen mã hóa ZFN được tách dòng cùng nhau trong khung đọc mở duy nhất được phân tách bởi trình tự mã hóa peptit tự cắt 2A. GFP được sử dụng làm đối chứng. Sau khi cấy vào chuột, các mẫu được lấy ở các thời điểm 4, 8 và 12 tuần sau khi cấy để quan sát mức độ đánh dấu đặc hiệu tế bào người ở các tế bào.

Các tế bào CD34+ đã chỉnh sửa ZFN ghép vào và biệt hóa, và mức độ ghép là tương đương giữa các tế bào đã chỉnh sửa so với các đối chứng chưa được chỉnh sửa.

Ví dụ 8: Hoạt tính của ZFN ở các tế bào thu được từ bệnh nhân

Hoạt tính của ZFN được thử nghiệm trong các tế bào CD34+ người đã huy động được phân lập từ máu ngoại vi của người và trong các tế bào CD34+ được phân lập từ tủy xương. HSPC từ năm đối tượng mắc bệnh tan máu bẩm sinh β (được ký hiệu P11, P18, P04, P08 và P19) được huy động và tinh chế như đã mô tả (Yannaki và đồng tác giả (2012) *Mol Ther* 20(1):230). Trong cả hai thử nghiệm, 200.000 tế bào CD34+ được xử lý bằng xung điện hai ngày sau khi làm rã đông, bằng cách sử dụng thiết bị tạo xung BTX (Holliston, MA, Điện thế = 250 V, Độ dài xung=5 ms) trong 100 μ L dung dịch xung điện BTX Express. Việc chuyển nhiễm sử dụng 4 μ g mARN mã hóa protein huỳnh quang xanh lá cây (GFP) (làm đối chứng về hiệu quả xung điện, và để kiểm soát các tác động không đặc hiệu của chính xung điện, và về các tác động không đặc hiệu của việc đưa mARN vào trong tế bào), hoặc 4 μ g và 8 μ g SB 51857-2a-51949 mARN.

Trong chuyển nhiễm quy mô nhỏ sử dụng thiết bị tạo xung BTX, các lượng SB 51857/51949 mARN này gây ra sự biến đổi gen đích và sự hoạt hóa γ -globin tương đương với các nồng độ 80-120 μ g/ml được sử dụng trong các chuyển nhiễm bằng thiết bị MaxCyte quy mô lớn hơn.

Sau khi xử lý bằng xung điện, việc nuôi cấy qua đêm nhanh ở 30°C được thực hiện. Các tế bào được nuôi cấy thêm 48 giờ nữa ở 37°C, ngay sau đó sự biệt hóa *in vitro* được bắt đầu và phần phân ước tế bào được thu gom để phân tích biến đổi ADN. Sau khi chuyển nhiễm, các tế bào được nuôi cấy trong môi trường X Vivo 10 (Lonza, Walkersville, MD) được bổ sung bằng hỗn hợp xytokin CC 100 (Stem Cell Technologies, Vancouver, Canada).

Biến đổi gen BCL11A được đo bằng cách giải trình tự sâu MiSeq 72 giờ sau khi xử lý bằng xung điện, ở thời điểm khi sự biệt hóa *in vitro* được bắt đầu (do đó d3 sau chuyển nhiễm là d0 của biệt hóa) và ở ngày thứ 14 của sự biệt hóa hồng cầu. Các kết quả được thể hiện trong Bảng 13.

Các mẫu tế bào CD34+ mà tác giả sáng chế thu được từ bệnh nhân được làm đông lạnh hai lần trước khi sử dụng, và do đó một số mẫu này thể hiện khả năng sống sót và sinh trưởng tế bào giảm khi rã đông. Khả năng sống sót thấp sau khi rã đông được quan sát thấy là trùng hợp với sự giảm nhiều hơn về khả năng sống sau khi chuyển nhiễm và với sự biến đổi gen đích thấp hơn, cụ thể là ở các thời điểm sớm, khi ADN từ các tế bào chết không được chuyển nhiễm vẫn có mặt. Như một chỉ báo độc lập chuyển nhiễm SB-ZFN về khả

năng sống của tế bào sau khi chuyển nhiễm, Bảng 13 thể hiện khả năng sống của tế bào cho mỗi nguồn tế bào trong mẫu đối chứng đã được chuyển nhiễm với GFP mARN. Bảng này thể hiện là các mẫu tế bào bệnh nhân P18 trong thí nghiệm 1 có khả năng sống thấp hơn nhiều (38%) so với hai mẫu còn lại (71% và 80%) trong thử nghiệm này và không đạt tới mức biến đổi giống nhau là ~70% alen bị biến đổi. Tương tự trong thử nghiệm thứ hai, vào ngày 3 mẫu tế bào bệnh nhân P08 thể hiện khả năng sống rất yếu (22%) và các mức biến đổi sớm và các mức biến đổi dưới mức tối ưu rất thấp ngay cả sau khi các tế bào khỏe mạnh nhân lên và sinh trưởng nhanh.

Bảng 13: Phân tích biến đổi gen *BCL11A* bằng MiSeq

	Nguồn tế bào	Khả năng sống d3 của đối chứng GFP (%)	SB-mRENH mARN (μg)	Biến đổi gen đích (%)	
				Ngày 3	Ngày 14
Thử nghiệm 1	WT	71	4	70	69
			8	72	69
	P11	80	4	70	71
			8	73	73
	P18	38	4	45	53
			8	51	59
Thử nghiệm 2	WT	89	4	71	72
			8	77	77
	P04	62	4	48	62
			8	53	67
	P08	22	4	19	62
			8	32	68
	P19	68	4	không có dữ liệu	69
			8	không có dữ liệu	73

ND= không có dữ liệu, WT= kiểu dại

Các dữ liệu này chỉ ra là sự phá vỡ trình tự tăng cường BCL11A sau khi xử lý bằng xung điện SB ZFN mARN vào trong các HSPC đã tinh chế, huy động G CSF từ người hiến

tăng khỏe mạnh và từ bệnh nhân mắc bệnh tan máu bẩm sinh với bệnh tan máu bẩm sinh β xảy ra trong hầu hết các mẫu trong khoảng được mong đợi cho các mẫu lâm sàng. Do khả năng sống thấp của một số mẫu, sự phá vỡ yếu tố tăng cường ở P18 và P08 là thấp hơn so với trong các tế bào của người tình nguyện hiến tặng khỏe mạnh, và do đó hai mẫu có khả năng sống thấp được loại bỏ khỏi các phân tích dưới đây. Quan trọng là, các mẫu từ các đối tượng thể hiện khả năng sống sót mạnh mẽ của tế bào (ví dụ, các mẫu P11 và P04) cũng thể hiện các mức biến đổi gen tương đương với các mức được thấy ở các tế bào kiểu đại.

Các mức nồng độ α và γ globin mRNA phân lập được từ dòng hồng cầu con của CD34+ HSPC từ các đối tượng mắc bệnh tan máu bẩm sinh β thể hiện sự tăng mức (γ) globin thai sau khi xử lý bằng SB ZFN mRNA khi được phân tích bằng RT qPCR (Fig.11). Các mức γ globin mRNA thai được chỉ ra là được chuẩn hóa so với α globin mRNA vì các tế bào bệnh tan máu bẩm sinh biểu hiện β globin mRNA thấp hoặc không biểu hiện.

Dòng hồng cầu con của CD34+ HSPC từ các đối tượng mắc bệnh tan máu bẩm sinh β được xử lý bằng SB ZFN mRNA cho thấy sự tăng đã dự đoán trước về tỉ lệ (γ) globin mRNA thai với α globin mRNA, đạt tới tỉ lệ gamma-globin so với alpha-globin tương tự với các tỉ lệ được thấy ở các tế bào cho kiểu đại, cụ thể là ở các mẫu của bệnh nhân mà thể hiện khả năng sống tốt sau khi rã đông và do đó, mức biến đổi gen đích là tương đương với ở các tế bào kiểu đại.

Sau đó HPLC pha đảo được sử dụng để xác định xem biến đổi về trình tự tăng cường hồng cầu BCL11A có làm tăng hemoglobin thai ở mức độ protein hay không.

Tỉ lệ gamma globin (tổng số pic Agamma và Ggamma)/ alpha globin cho hai thử nghiệm thể hiện sự tăng rõ ràng của globin protein thai được quan sát thấy trong các tế bào hồng cầu (RBC) thu được từ người tình nguyện khỏe mạnh và bệnh nhân mắc bệnh tan máu bẩm sinh khi phá vỡ SB ZFN của trình tự tăng cường BCL11A, ngay cả khi các tỉ lệ gamma/alpha chưa được xử lý trong các tế bào tan máu bẩm sinh đặc biệt là trong các tế bào $\beta 0/\beta 0$ thường là cao hơn so với các tỉ lệ này ở các tế bào kiểu đại (wt). Phân tích các tỉ lệ globin thai so với người trưởng thành ở các tế bào từ bệnh nhân mắc bệnh tan máu bẩm sinh β là phức tạp chủ yếu bởi thực tế là sự tetramer hóa và kết tủa của α globin trong các tế bào chưa biến đổi khử nó từ nguồn gộp HPLC phân tích được.

Do đó, chỉ α globin được tetrame hóa với β globin còn lại (ở các bệnh nhân mắc bệnh tan máu bẩm sinh β^+) và γ globin, hoặc chỉ α globin được tetrame hóa với γ globin (trong các tế bào β^0/β^0) có thể được bộc lộ trong thử nghiệm này. Do đó, nếu các mức γ globin protein tăng, các mức α globin protein được tetrame hóa năng suất cũng có thể tăng và tỉ lệ γ/α globin protein đánh giá thấp sự tăng mức γ globin protein. Một tỉ lệ khác mà hữu ích để kiểm tra trong các tế bào kiểu đại là tỉ lệ γ globin / protein tương tự β (số liệu sau là tổng của hai mức γ globin protein cộng δ globin cộng β globin). Tuy nhiên, ở các tế bào của bệnh nhân mắc bệnh tan máu bẩm sinh, cụ thể là ở các tế bào β^0/β^0 , tỉ lệ này thường cao hơn 90% ngay cả trong các tế bào không được xử lý ZFN và thậm chí sự gia tăng đáng kể về các mức γ globin protein sau khi xử lý ZFN không làm tăng đáng kể tỉ lệ này.

Ví dụ 9: Phân tích sự phân bố alen biến đổi và tác động lên globin thai trong các tế bào CD34+ kiểu đại được xử lý ZFN

Tác giả sáng chế cũng đánh giá các tế bào hồng cầu thu được từ các tế bào CD34+ được xử lý bằng SB ZFN, ở mức độ một tế bào duy nhất về hai điểm cuối: (1) sự phân bố các alen của vùng tăng cường hồng cầu BCL11A được đánh giá bằng hoạt động của các ZFN giữa các tế bào đơn lẻ, và (2) tác động của sự phân bố của các alen đơn lẻ này (kiểu đại và được biến đổi di truyền) trên mức nồng độ globin thai (như được đo theo tỉ lệ γ/β globin mRNA). Do đó, các tế bào CD34+ đơn lẻ được tìm thấy trong SB ZFN HSPC được phân loại và được biệt hóa *in vitro*. Các khuẩn lạc của các tế bào được hemoglobin hóa thu được từ tế bào đơn lẻ tạo thành được thu gom riêng; ADN bộ gen (gADN) cho mỗi khuẩn lạc này được giải trình tự ở vùng đích SB ZFN trong trình tự tăng cường hồng cầu BCL11A để xác định locut này có bị phá vỡ hay không, và nếu có, dạng alen chính xác nào của locut này được tạo ra. Ngoài ra, ARN tổng được phân lập từ cùng một khuẩn lạc, và các mức độ biểu hiện globin trong các dòng đơn lẻ này được phân tích bằng phản ứng chuỗi polymeraza định lượng phiên mã đảo thời gian thực (RT qPCR).

Phương pháp

Đối với các nghiên cứu tế bào đơn lẻ, các tế bào SB ZFN HSPC đã chuyển nhiễm được rửa đông ở 37°C, được bổ sung vào 10mL X VIVO ở nhiệt độ phòng, và được ly tâm ở tốc độ 450 x g trong 5 phút ở nhiệt độ phòng. Các pelet tế bào được tạo huyền phù lại ở nồng độ $1 \times 10^6/\text{mL}$ trong môi trường X VIVO được bổ sung Flt3L, TPO, và SCF

(100ng/ml mỗi thành phần), penixilin (100U/ml), và streptomycin (100µg/mL). Sau khi nuôi cấy qua đêm trong đĩa 24 lỗ được xử lý không phải nuôi cấy mô ở 37°C, 5% CO₂ trong tủ ủ được làm ẩm, các tế bào được thu gom, ly tâm, và được tạo huyền phù lại trong nước muối đệm phosphat (PBS) được bổ sung bằng albumin huyết thanh bò 0,5% ở nồng độ 2×10^6 /mL. Các tế bào được phân loại vào môi trường nuôi cấy hồng cầu Bước 1 (200µL/lỗ) trong các đĩa 96 lỗ đáy chữ U được xử lý không phải TC ở mật độ 2 tế bào/lỗ sử dụng FACS Aria III. Môi trường nuôi cấy hồng cầu Bước 1 bao gồm Môi trường Dulbecco biến đổi của Iscove chứa Glutamax (Iscove's Modified Dulbecco's Medium: IMDM) được bổ sung 100U/ml penixilin, 100µg/mL streptomycin, 5% huyết tương AB+ người, 330µg/mL holo transferrin người, 20µg/mL insulin người, 2U/ml heparin, 3U/ml erythropoietin tái tổ hợp của người, 100ng/ml SCF, 5ng/ml IL 3, và 1µM/mL hydrocortison. Sau 7 ngày nuôi cấy ở 37°C, 5% CO₂ trong môi trường nuôi cấy hồng cầu Bước 1, 150µL môi trường mỗi lỗ được lấy ra và thay thế bằng 100µL môi trường nuôi cấy hồng cầu Bước 2, mà tương tự với môi trường Bước 1 nhưng không bổ sung IL 3 và hydrocortison. Sau 4 ngày nuôi cấy thêm, 100µL môi trường mỗi lỗ được lấy ra và được thay thế bằng 100µL môi trường Bước 3, mà tương tự với môi trường Bước 2 nhưng không có SCF.

Vào ngày 14 sau khi biệt hóa, 10µL huyền phù tế bào mỗi lỗ được thu hoạch để giải trình tự sâu. Ngoài ra, 100µL môi trường mỗi lỗ được lấy ra và được thay thế bằng 100µL môi trường Bước 3 mới. Các tế bào còn lại được nuôi cấy trong 3 ngày nữa khi 50µL huyền phù tế bào mỗi lỗ được thu gom, được nhuộm bằng thể tích tương đương của NucRed (2 giọt/mL trong PBS 0,5% BSA), và được đánh số trên Guava easyCyte về tỉ lệ tế bào và tỉ lệ chết nhân. Các tế bào còn lại được ly tâm, rửa một lần bằng PBS, và được dung giải trong 20µL nước sử dụng cho sắc ký lỏng hiệu năng cao (HPLC). Mảnh tế bào được lấy ra bằng cách ly tâm (10,000 x g, 15 phút, 4°C). Dịch máu tan được lưu trữ ở 70°C đến khi sẵn sàng cho việc phân tích chuỗi globin bằng sắc ký lỏng siêu cao áp (UPLC) pha đảo.

Hiệu quả biến đổi gen được đánh giá bằng cách giải trình tự ADN sâu. Vùng cần quan tâm (chứa vị trí liên kết ZFN nằm trong vùng tăng cường hồng cầu BCL11A) được khuếch đại phản ứng chuỗi polymeraza (PCR) và mức độ biến đổi được xác định bằng cách giải trình tự sâu từ hai đầu của cùng một trình tự trên Illumina MiSeq. Để tạo ra các thư viện tương thích với nền giải trình tự Illumina MiSeq, các trình tự gá lắp, trình tự mã vạch,

và phân tử liên kết tế bào dòng (trình tự ADN ngắn) được gắn vào các đơn vị siêu sao chép đặc hiệu đích bằng cách sử dụng hai tập hợp đoạn mồi dung hợp trong các phản ứng chuỗi polymeraza liên tiếp. Các đoạn mồi sau đây được sử dụng cho MiSeq Adaptor PCR (các phần được gạch chân là các trình tự đặc hiệu vùng tăng cường hồng cầu BCL11A): PRJIYLFN f:

5' ACACGACGCTCTCCGATCTNNNNAGTCCTCTTCTACCCACCC (SEQ ID NO:35) và
PRJIYLFN r:

5' GACGTGTGCTCTCCGATCTCTACTCTTAGACATAACACACC (SEQ ID NO:36).

Các môi trường nuôi cấy hồng cầu thu được từ từng tế bào đơn lẻ được thu hoạch, và gADN được chiết bằng cách sử dụng QuickExtract™ (Epicentre) cho phân tích xác định kiểu gen bằng cách sử dụng kỹ thuật giải trình tự sâu trên nền Illumina.

Đối với phân tích chuỗi globin bằng UPLC pha đảo, 5µL dịch tan máu được tiêm lên Cột Waters Acquity UPLC Protein BEH C4 (300A, 1,7 microm, 2,1mm X 100mm). Dịch rửa giải thu được ở nhiệt độ phòng với tốc độ dòng là 0,2mL/phút bằng cách sử dụng gradien tuyến tính 18 phút từ 38% đến 42,5% axetonitril trong nước với axit trifloaxetic không đổi ở 0,1%. Việc rửa giải được tiếp tục ở 220nm. Tỷ lệ phần trăm diện tích cho các chuỗi globin đặc hiệu, γ , β hoặc α , thể hiện lượng của mỗi chuỗi globin đặc hiệu, được định lượng bằng cách sử dụng phần mềm Agilent OpenLAB.

Kết quả

Đối với nghiên cứu này, 120 hoặc 80µg/mL SB ZFN được chuyển nhiễm vào các tế bào CD34+ để tạo ra các tế bào SB ZFN HSPC và các mẫu tế bào được lấy ở thời điểm ba ngày sau khi chuyển nhiễm và được phân tích về các mức phá vỡ vùng tăng cường hồng cầu BCL11A bằng cách giải trình tự sâu. Nghiên cứu này cho thấy xấp xỉ 67% alen BCL11A bị biến đổi trong mẫu đã chuyển nhiễm SB ZFN đối với thử nghiệm 1 và 58% đối với thử nghiệm 2 (Bảng 14).

Để thực hiện phân tích tế bào đơn lẻ này, từng tế bào trong SB ZFN HSPC từ hai thử nghiệm khác nhau (1 và 2), được phân loại và được đưa vào các đĩa 96 lỗ và trải qua sự biệt hóa hồng cầu *in vitro*, và từng môi trường nuôi cấy tế bào đơn lẻ lần lượt được phân tích bằng cách giải trình tự ADN thông lượng cao cũng như phân tích sự biểu hiện globin bằng UPLC.

Các kết quả xác định kiểu gen MiSeq cho các môi trường nuôi cấy tế bào đơn lẻ được tóm tắt trong Bảng 14. Đối với mỗi lô tế bào HSPC được xử lý bằng SB ZFN, từ 200 đến 300 môi trường nuôi cấy hồng cầu tế bào đơn lẻ riêng được phân tích. Trong số tất cả các dòng có kiểu hình rõ ràng (các dòng hỗn hợp được loại ra, 205 dòng cho thử nghiệm 1 và 265 dòng cho thử nghiệm 2), 28 hoặc 36% là các dòng kiểu đại (+/+), 14 hoặc 11% là các dòng dị hợp tử (+/-), và 58 hoặc 52% là các dòng đồng hợp tử (-/-), lần lượt đối với thử nghiệm 1 và thử nghiệm 2. Trong số tất cả các alen trong các môi trường nuôi cấy hồng cầu tế bào đơn lẻ thu được từ các tế bào được xử lý bằng SB ZFN, 65% hoặc 58%, lần lượt cho các thử nghiệm 1 và 2, được phá vỡ ở locut tăng cường hồng cầu BCL11A. Trong số tất cả các môi trường nuôi cấy hồng cầu tế bào đơn lẻ mang alen biến đổi bất kỳ, 81% (thử nghiệm 1) hoặc 82% (thử nghiệm 2) các dòng tế bào biến đổi có cả hai alen BCL11A bị phá vỡ.

Bảng 14: Xác định kiểu gen của môi trường nuôi cấy hồng cầu tế bào đơn lẻ thu được từ HSPC được chuyển nhiễm SB ZFN

Lô SB-ZFN-HSPC	Thử nghiệm số	
	#1	#2
Tổng số dòng được kiểm tra (Tổng số alen)	205 (410)	265 (530)
Các dòng kiểu đại (+/+)	57 (28%)	96 (36%)
Các dòng dị hợp tử (Được biến đổi monoalen) (+/-)	28 (14%)	30 (11%)
Các dòng đồng hợp tử (Được biến đổi hai alen) (-/-)	120 (58%)	139 (52%)
Được biến đổi thực (Biến đổi dị hợp tử [+/-] + biến đổi hai alen [-/-])	148 (72%)	169 (63%)
Tỉ lệ tế bào biến đổi mà được biến đổi hai alen [-/-]	81%	82%
<i>Biến đổi trình tự tăng cường hồng cầu BCL11A</i> trong nguồn gộp trước khi nuôi cấy tế bào đơn lẻ (% trong tổng số alen)	67%	58%
% trong tổng số alen <i>Biến đổi trình tự tăng cường hồng cầu BCL11A</i> thực từ số liệu tế bào đơn lẻ (% trong tổng số alen)	65%	58%

Số liệu thể hiện là nguồn gộp của SB ZFN HSPC mang 58-67% biến đổi trình tự tăng cường hồng cầu BCL11A hướng đích được tạo bởi 28-36% tế bào kiểu đại, 11-14% tế bào mang biến đổi monoalen, và 52-58% tế bào mang biến đổi hai alen của locut đích.

Trong số các dòng có kiểu gen rõ ràng, 152 và 172 dòng (thử nghiệm 1 và 2, tương ứng) được biệt hóa thành công thành các tế bào hồng cầu *in vitro*, như được chỉ ra bởi tốc độ loại bỏ nhân đo được bằng cách nhuộm NucRed. Các môi trường nuôi cấy hồng cầu tế bào đơn lẻ này sau đó được phân tích UPLC pha đảo để đo mức độ biểu hiện chuỗi globin. Các khuẩn lạc khác biệt về mức độ trưởng thành hồng cầu của chúng; một số thay đổi về tỉ lệ γ/β globin được mong đợi ngay cả trong các khuẩn lạc mang cùng một kiểu gen trình tự tăng cường hồng cầu BCL11A. Ngoài ra, tỉ lệ alen bị phá vỡ của trình tự tăng cường hồng cầu BCL11A có thể giữ được một phần hoặc toàn bộ chức năng, có thể là do vẫn giữ được vị trí liên kết GATA 1 (Vierstra và đồng tác giả (2015) Nat Methods. 12(10): 927-30; Canver và đồng tác giả (2015) Nature 527(7577): 192-7).

Như có thể thấy trong số liệu (Fig.13), các kết quả cho thấy mối tương quan rõ ràng giữa kiểu gen của khuẩn lạc đối với locut tăng cường hồng cầu BCL11A và tỉ lệ γ/β globin của nó. Cụ thể là, các khuẩn lạc mang biến đổi hai alen (đồng hợp tử) của BCL11A có tỉ lệ γ/β được chuẩn hóa trung bình là 35% và 39% cho thử nghiệm 1 và 2, tương ứng, các khuẩn lạc có biến đổi monoalen (dị hợp tử) có tỉ lệ γ/β được chuẩn hóa trung bình là 19% và 13% cho thử nghiệm 1 và 2, tương ứng, và các khuẩn lạc kiểu đại là 14% và 13% cho thử nghiệm 1 và 2, tương ứng. Trong hai test t có đuôi với hiệu chỉnh Welch, giá trị P của so sánh “kiểu đồng hợp tử so với kiểu đại” là $<0,0001$ đối với cả hai thử nghiệm 1 và 2; giá trị P của so sánh “dị hợp tử với đồng hợp tử” là $<0,0001$ đối với thử nghiệm 1 và 0,0025 đối với thử nghiệm 2; và giá trị P của so sánh “dị hợp tử với kiểu” là 0,02 và 0,0002 cho cả hai thử nghiệm 1 và 2, tương ứng. Trong đồ thị dưới đây (Fig.13), tỉ lệ γ/β và γ/α globin được lập đồ thị cho tất cả các khuẩn lạc được thử nghiệm, được phân loại theo lớp xác định kiểu gen của các alen trình tự tăng cường hồng cầu BCL11A.

Tất cả các patent, đơn yêu cầu cấp patent và các công bố được đề cập ở đây được đưa vào đây bằng cách viện dẫn toàn bộ.

Mặc dù phần mô tả được cung cấp ở mức độ chi tiết bằng cách minh họa và lấy ví dụ nhằm mục đích rõ ràng và dễ hiểu, sẽ được biết rõ bởi người có kiến thức trung bình trong lĩnh vực này là nhiều thay đổi và biến đổi khác nhau có thể được thực hành mà không

vượt ra khỏi phạm vi hoặc tinh thần của sáng chế. Theo đó, phần mô tả trên đây và các ví dụ không nên được hiểu là làm giới hạn sáng chế.

YÊU CẦU BẢO HỘ

1. Protein ngón tay kẽm chứa 4, 5 hoặc 6 ngón tay được ký hiệu là F1 đến F4, F1 đến F5 hoặc F1 đến F6, mỗi ngón tay chứa vùng xoắn nhận biết mà nhận biết vị trí đích phụ, trong đó protein này được chọn từ một trong các nhóm gồm có:

(i) protein chứa các vùng xoắn nhận biết như sau:

F1: STGNLTN (SEQ ID NO:7);

F2: TSGSLTR (SEQ ID NO:5);

F3: DQSNLRA (SEQ ID NO:2); và

F4: AQCCLFH (SEQ ID NO:6);

(ii) protein chứa các vùng xoắn nhận biết như sau:

F1: DQSNLRA (SEQ ID NO:2);

F2: RPYTLRL (SEQ ID NO:3);

F3: SRGALKT (SEQ ID NO:8);

F4: TSGSLTR (SEQ ID NO:5);

F5: DQSNLRA (SEQ ID NO:2); và

F6: AQCCLFH (SEQ ID NO:6);

(iii) protein chứa các vùng xoắn nhận biết như sau:

F1: DQSNLRA (SEQ ID NO:2);

F2: RNFSLTM (SEQ ID NO:9);

F3: SNGNLRN (SEQ ID NO:10) hoặc SSYNLAN (SEQ ID NO:11);

F4: TSGSLTR (SEQ ID NO:5);

F5: DQSNLRA (SEQ ID NO:2); và

F6: AQCCLFH (SEQ ID NO:6);

(iv) protein chứa các vùng xoắn nhận biết như sau:

F1: RSDHLTQ (SEQ ID NO:13);

F2: QSGHLAR (SEQ ID NO:14);

F3: QKGTLEGE (SEQ ID NO:15);

F4: RHRDLRS (SEQ ID NO:18); và

F5: RRDNLHS (SEQ ID NO:17);

(v) protein chứa các vùng xoắn nhận biết như sau:

F1: RNDHRTT (SEQ ID NO:19);

F2: QKAHLIR (SEQ ID NO:20);

F3: QKGTLEGE (SEQ ID NO:15);

F4: LKRTLKR (SEQ ID NO:25); và

F5: RRDNLHS (SEQ ID NO:17);

(vi) protein chứa các vùng xoắn nhận biết như sau:

F1: RSDHLTQ (SEQ ID NO:13);

F2: QRAHLTR (SEQ ID NO:22);

F3: QKGTLEGE (SEQ ID NO:15) hoặc QSGTRNH (SEQ ID NO:24);

F4: HRNTLVR (SEQ ID NO:23); và

F5: RRDNLHS (SEQ ID NO:17);

(vii) protein chứa các vùng xoắn nhận biết như sau:

F1: RSDHLTQ (SEQ ID NO:13);

F2: QKAHLIR (SEQ ID NO:20);

F3: QKGTLEGE (SEQ ID NO:15) hoặc QSGTRNH (SEQ ID NO:24);

F4: RGRDLRS (SEQ ID NO:21); và

F5: RRDNLHS (SEQ ID NO:17); và

(viii) protein chứa các vùng xoắn nhận biết như sau:

F1: RSDHLTQ (SEQ ID NO:13);

F2: QSGHLAR (SEQ ID NO:14);

F3: QSGTRNH (SEQ ID NO:24);

F4: QSSDLSR (SEQ ID NO:16); và

F5: RRDNLHS (SEQ ID NO:17).

2. Protein dung hợp chứa protein ngón tay kềm theo điểm 1 và vùng chức năng.
3. Protein dung hợp theo điểm 2, trong đó vùng chức năng này là vùng hoạt hóa phiên mã, vùng ức chế phiên mã, hoặc vùng phân cắt.
4. Polynucleotit mã hóa protein ngón tay kềm theo điểm 1.
5. Tế bào phân lập chứa protein dung hợp theo điểm 2 hoặc điểm 3 hoặc polynucleotit theo điểm 4, trong đó protein dung hợp này được biểu hiện trong tế bào.
6. Tế bào phân lập theo điểm 5, trong đó tế bào này là tế bào gốc tạo máu hoặc tế bào tiền thân của tế bào hồng cầu.
7. Tế bào phân lập theo điểm 6, trong đó tế bào này là tế bào của người.
8. Tế bào phân lập theo điểm bất kỳ trong các điểm từ 5 đến 7, trong đó bộ gen của tế bào được biến đổi bởi protein dung hợp này.
9. Tế bào phân lập theo điểm 8, trong đó sự biến đổi bộ gen được chọn từ nhóm bao gồm đột biến gắn xen, đột biến mất đoạn và tổ hợp của chúng.
10. Tế bào phân lập theo điểm 8 hoặc 9, trong đó sự biến đổi bộ gen là nằm trong vùng +58 của trình tự tăng cường BCL11A.
11. Tế bào hoặc dòng tế bào con được biến đổi di truyền, được phân lập, được sản sinh ra từ tế bào theo điểm bất kỳ trong số các điểm từ 5 đến 10.
12. Tế bào được biệt hóa một phần hoặc hoàn toàn, được biến đổi di truyền, được phân lập có nguồn gốc từ tế bào hoặc dòng tế bào con theo điểm 11.
13. Tế bào phân lập theo điểm bất kỳ trong số các điểm từ 5 đến 10, trong đó vùng chức năng là vùng phân cắt và hơn nữa trong đó tế bào này thể hiện sự biểu hiện của gamma và/hoặc beta globin tăng so với tế bào không có biến đổi bộ gen.
14. Dược phẩm chứa protein dung hợp theo điểm 2 hoặc điểm 3, hoặc polynucleotit theo điểm 4 hoặc tế bào theo điểm bất kỳ trong các điểm từ 5 đến 13.
15. Phương pháp sản sinh ra tế bào được biến đổi di truyền chứa biến đổi bộ gen nằm trong trình tự tăng cường BCL11A nội sinh nằm trong intron 2 của gen BCL11A, phương pháp

này bao gồm các bước:

a) cho tế bào tiếp xúc với polynucleotit mã hóa protein dung hợp theo điểm 2 hoặc điểm 3, trong đó protein dung hợp này chứa vùng phân cắt;

b) đưa tế bào này vào các điều kiện có lợi cho việc biểu hiện protein dung hợp trong đó protein được biểu hiện biến đổi trình tự tăng cường BCL11A nội sinh để tạo ra tế bào được biến đổi di truyền.

16. Phương pháp theo điểm 15, phương pháp này còn bao gồm bước kích thích tế bào bằng ít nhất một xytokin.

17. Kit chứa protein theo điểm bất kỳ trong số các điểm từ 1 đến 3, polynucleotit theo điểm 4 và/hoặc tế bào theo điểm bất kỳ trong các điểm từ 5 đến 13.

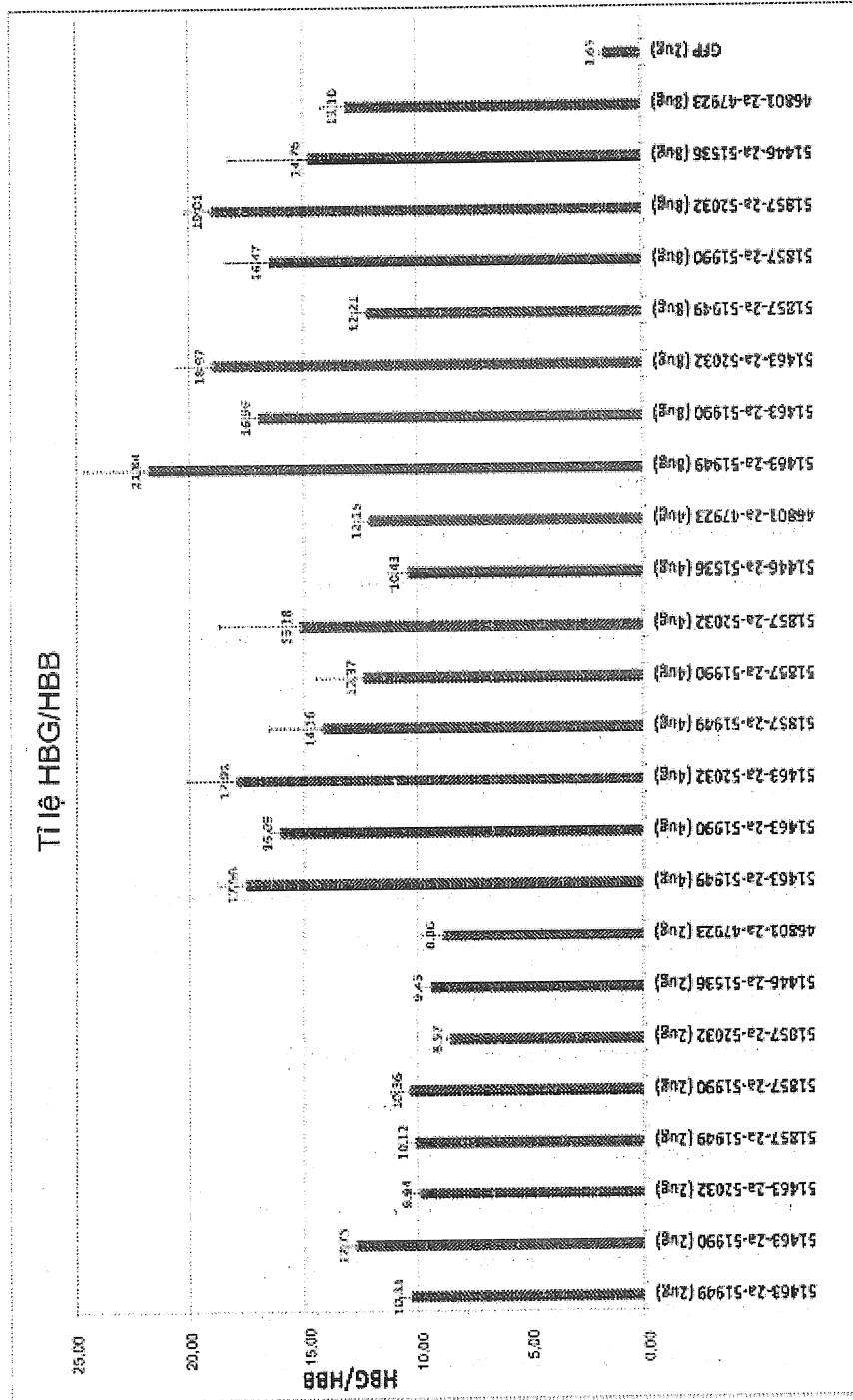


Fig. 1

Fig. 2A

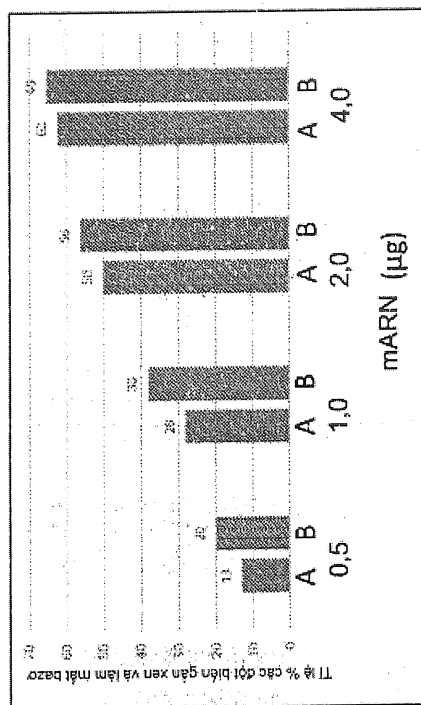
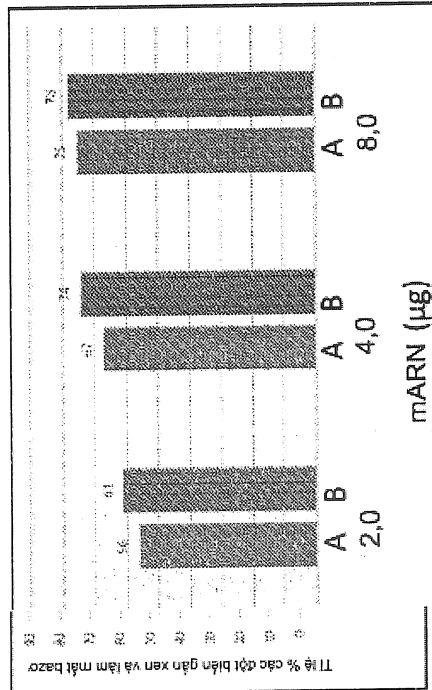


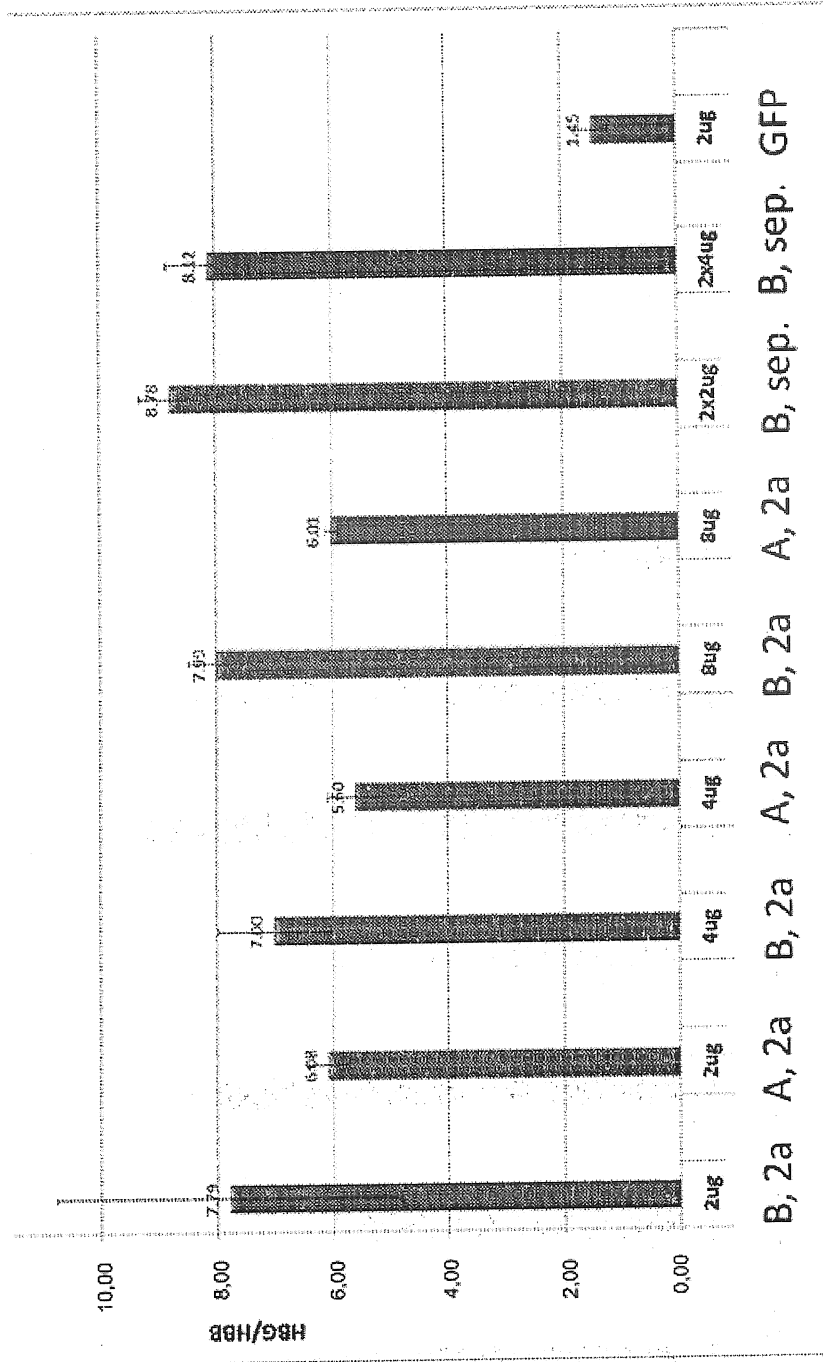
Fig. 2B



Tỉ lệ % các đột biến gắn xen và làm mất bazơ				
mARN được sử dụng/phản ứng	0,5	1,0	2,0	4,0
A: 51446_2a_51536	12,8	28,2	50,1	62,2
B: 51857_2a_51949	20,5	38,4	56,0	64,7

Tỉ lệ % các đột biến gắn xen và làm mất bazơ			
mARN được sử dụng/phản ứng	2,0	4,0	8,0
A: 51446_2a_51536	56,4	66,7	74,9
B: 51857_2a_51949	61,2	73,9	78,3

Fig. 3



B, 2a A, 2a B, 2a A, 2a B, 2a A, 2a B, sep. B, sep. B, sep. GFP

Fig. 4A

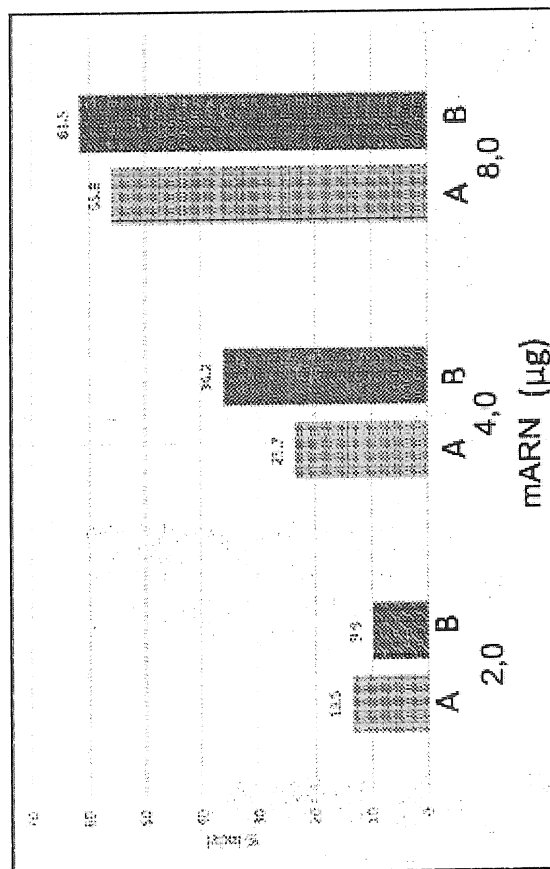
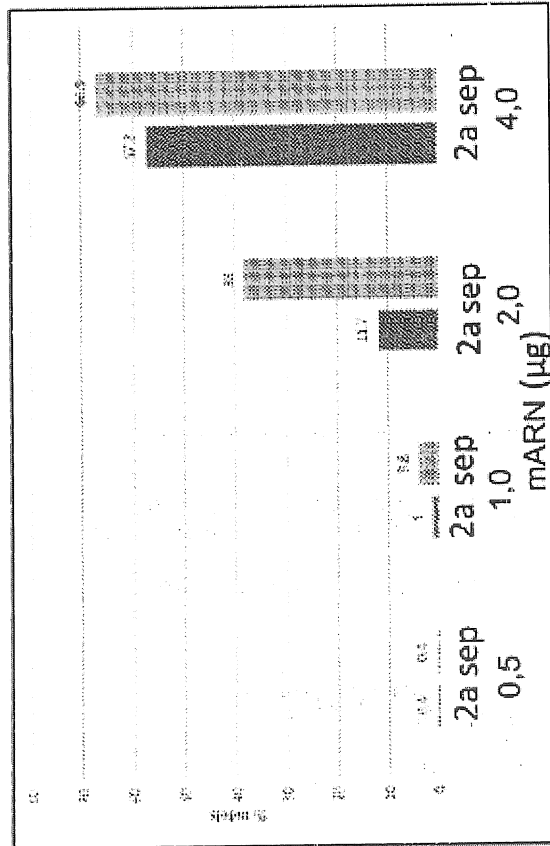


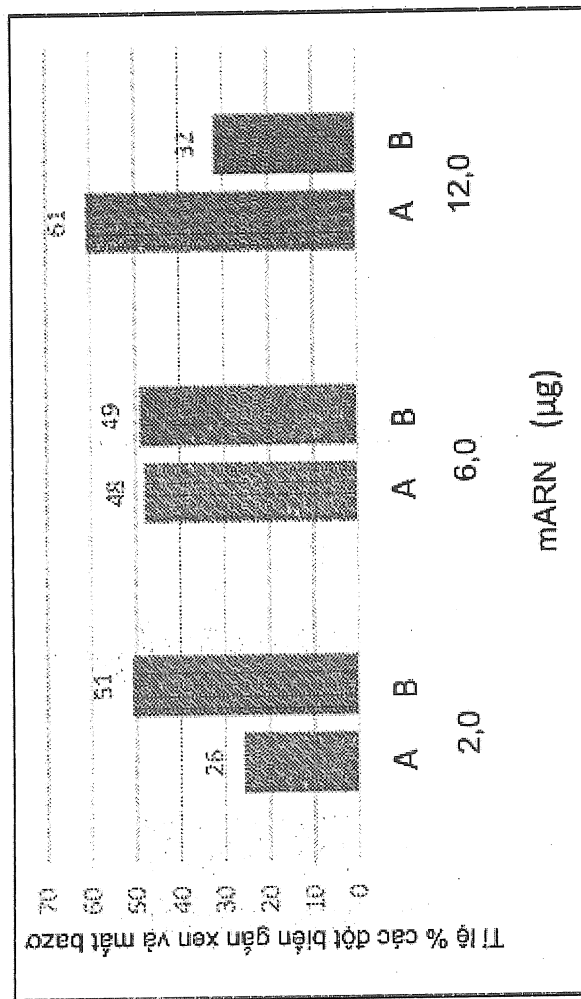
Fig. 4B



Tỉ lệ % các đột biến gắn xen và làm mất bazơ			
mARN được sử dụng/phản ứng	2,0	4,0	8,0
A: 51446_2a_51536	13,5	23,7	55,8
B: 51857_2a_51949	9,9	36,3	61,5

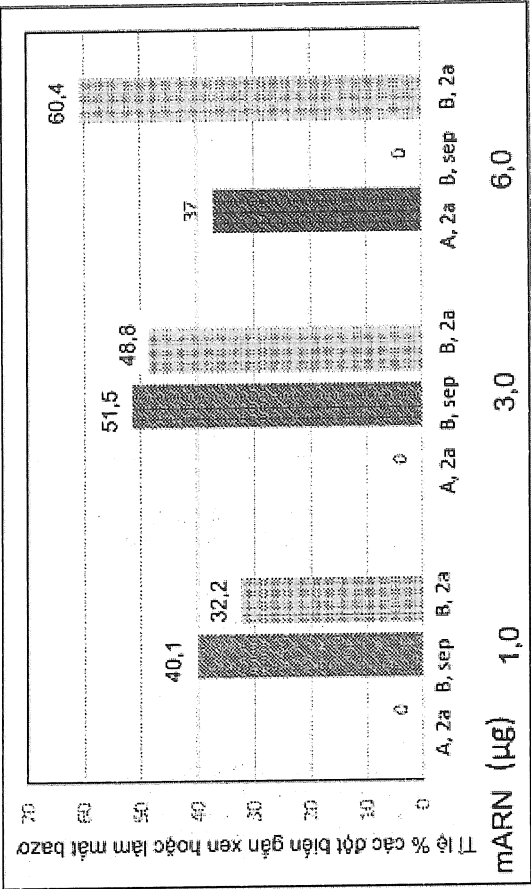
Tỉ lệ % các đột biến gắn xen và làm mất bazơ			
mARN được sử dụng/phản ứng	0,5	2,0	4,0
B: 51857_2a_51949	0,4	1,0	11,7
B: 51857_51949 sep	0,4	3,8	38,0

Fig.5A



Tỉ lệ % các đột biến gắn xen và mất bazơ			
mARN được sử dụng/phản ứng	2,0	6,0	12,0
A: 51446_2a_51536	25,9	48,1	60,7
B: 51857_2a_51949	50,8	49,4	32,6

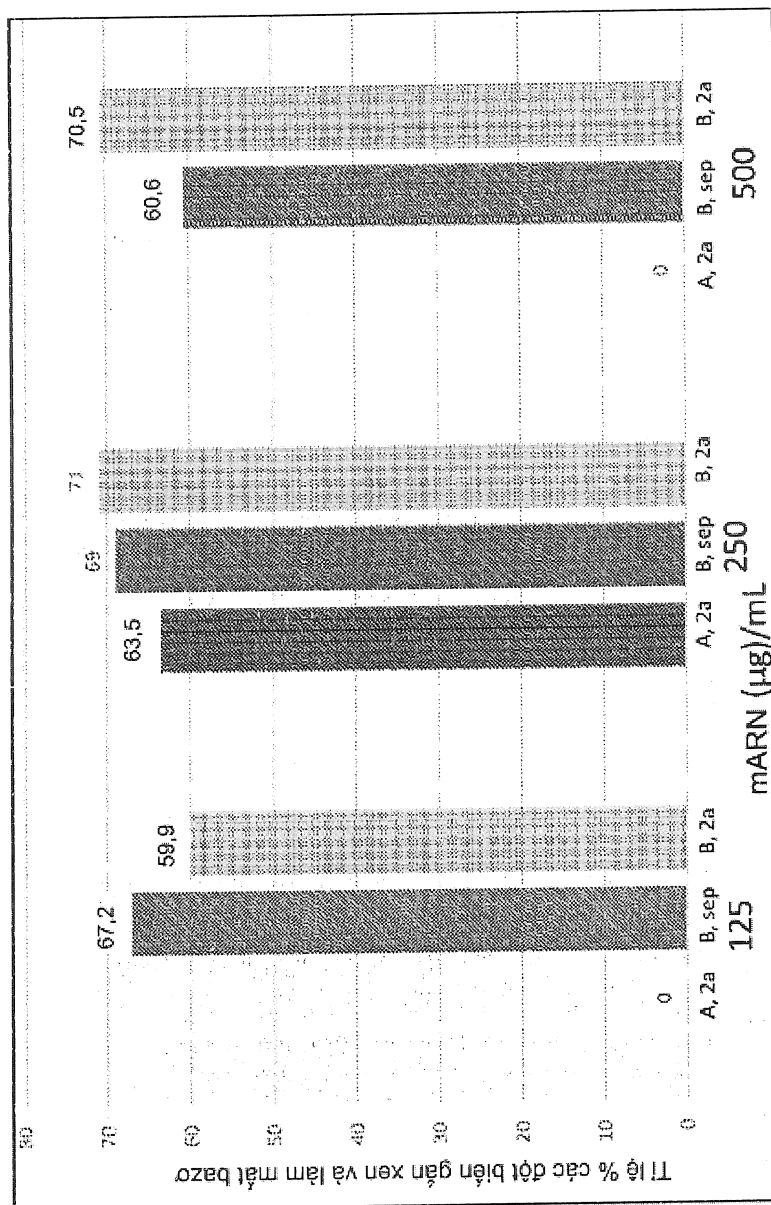
Fig. 5B



Tỉ lệ % các đột biến gần xen và làm mất bazơ			
mARN được sử dụng/phản ứng	1,0	3,0	6,0
A: 51446_2a_51536	nt	nt	37
B: 51857 + 51949, sep	40,1	51,5	nt
B: 51857_2a_51949	32,2	48,8	60,4

Nt=không được thử nghiệm

Fig. 6



Tỉ lệ % các đột biến gắn xen và làm mất bazơ			
mARN được sử dụng/ml	125	250	500
A: 51446_2a_51536	nt	63,5	nt
B: 51857 + 51949, sep	67,2	69,0	60,6
B: 51857_2a_51949	59,9	71	70,5

Fig. 7

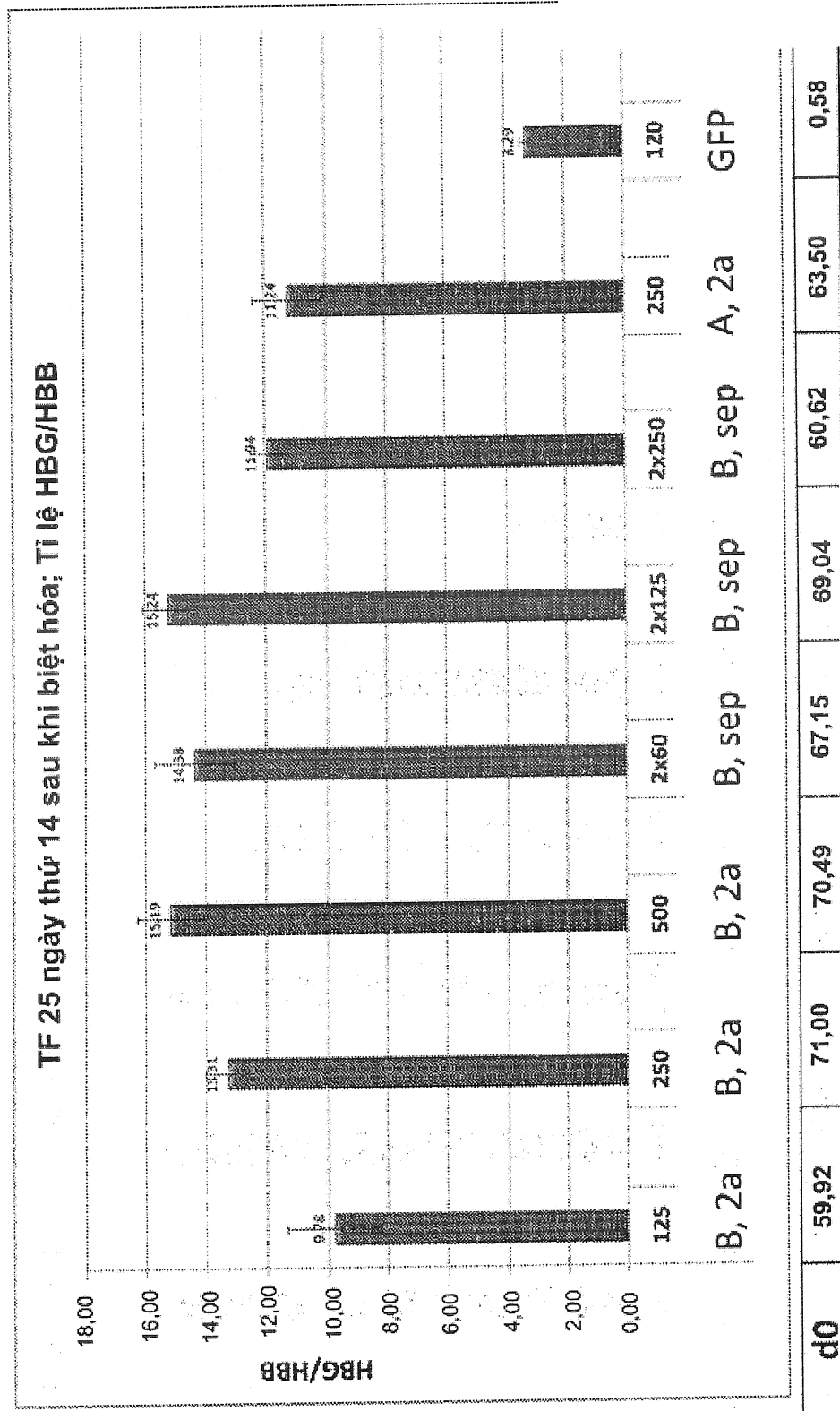
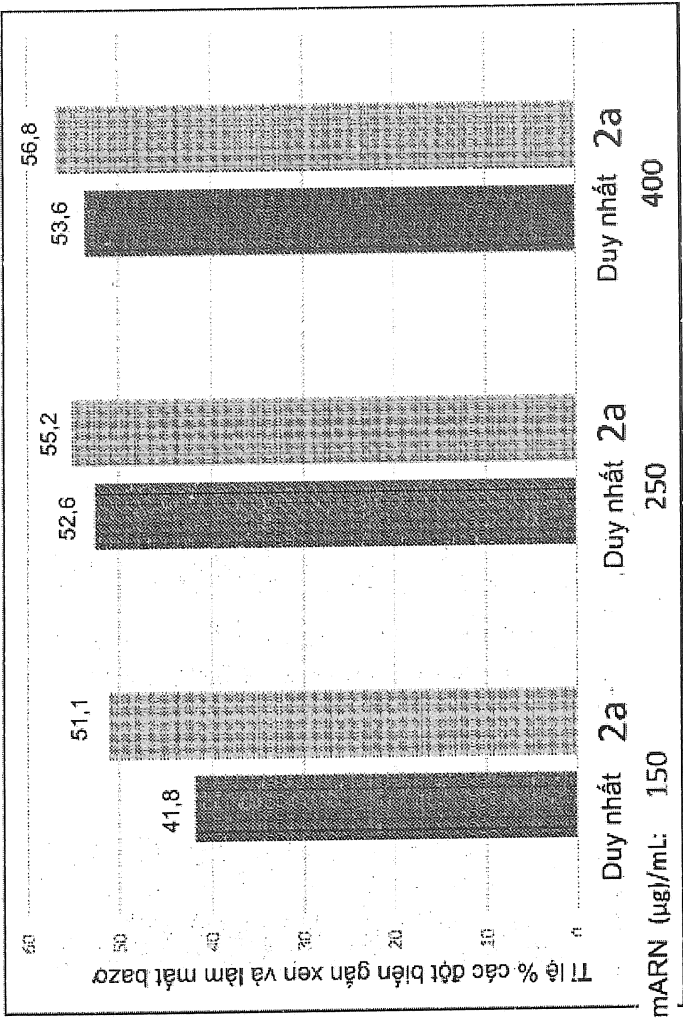
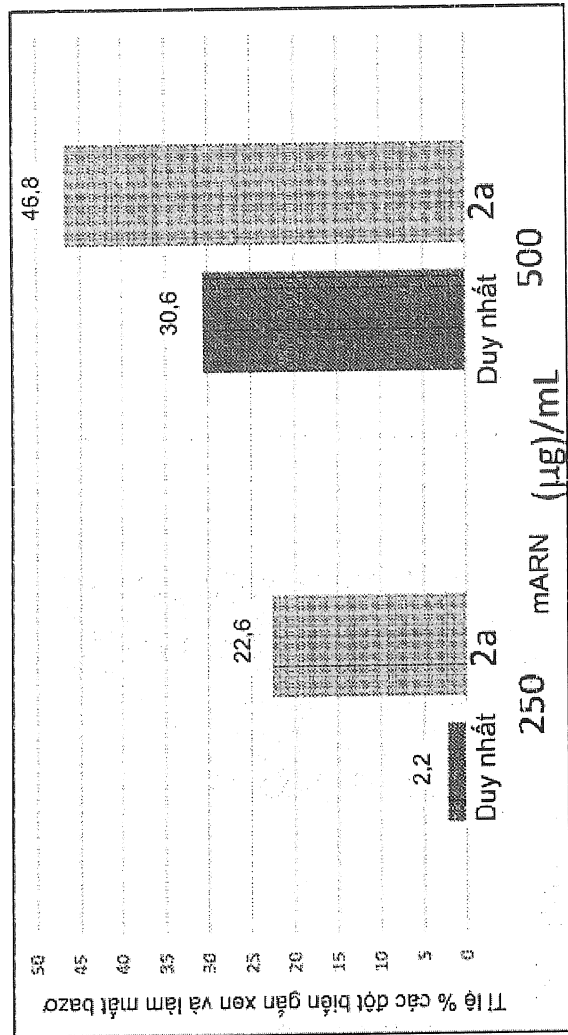


Fig. 8

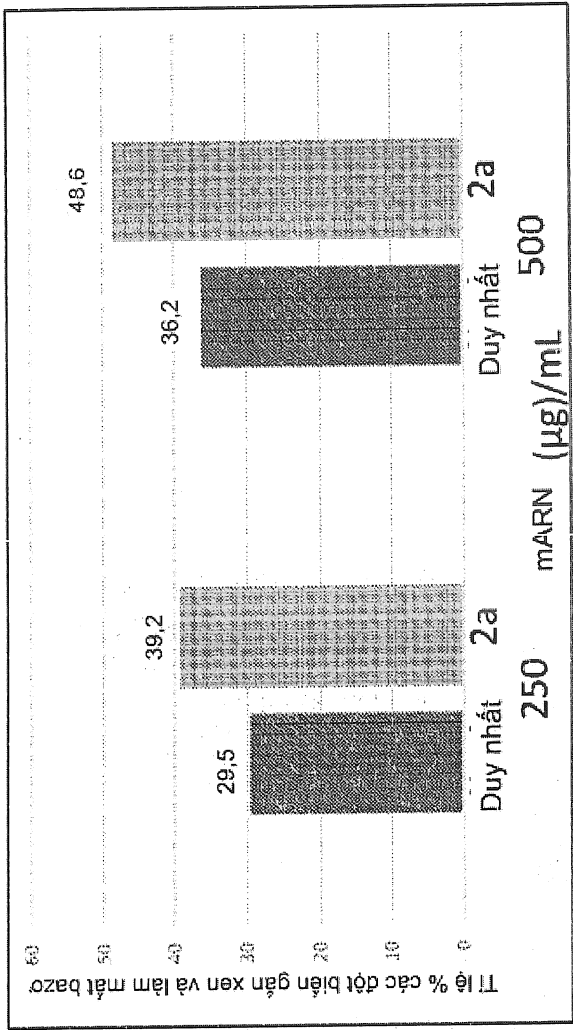


Tỉ lệ % các đột biến gắn xen và làm mất bazơ			
mARN được sử dụng/mL	150	250	400
B: 51857 + 51949, sep	41,8	52,6	53,6
B: 51857_2a_51949	51,0	55,2	56,8



Tỉ lệ % các đột biến gắn xen và làm mất bazơ		
mARN được sử dụng/mL	250	500
B: 51857 + 51949, sep	2,2	30,6
B: 51857_2a_51949	22,6	46,8

Fig. 9A



Tỉ lệ % các đột biến gắn xen và làm mất bazơ		
mARN được sử dụng/mL	250	500
B: 51857 + 51949, sep	29,5	36,2
B: 51857_2a_51949	39,2	48,5

Fig. 9B

Fig. 10A

Đơn vị siêu sao chép_vị trí		Tần số	Tần số đột biến xen và gây mất bazơ được xử lý	Giá trị bonferroni_p	thuộc gen	mã hóa	gen	vị trí
chr2:60722338-60722476		63,261%	0,00000		y	n	LRRC1	intron
chr6:53670690-53670868		2,312%	0,00000		y	n	LHX1	intron
chr17:35295932-35296095		1,184%	0,00000		n	n	n/a	giữa các gen
chr8:140226152-140226311		0,711%	0,00000		n	n	MMP27	ngược dòng
chr11:102556535-102556674		0,344%	0,00000		y	n	COMMD6 exon 1	ngược dòng
chr13:76123418-76123581		0,674%	0,00000		n	n	DSCC1	ngược dòng
chr8:120868614-120868752		0,303%	0,00000		n	n	ACOT11	ngược dòng
chr1:54972397-54972573		0,569%	0,00000		y	n	INPP5A	intron
chr10:134468242-134468405		0,425%	0,00000		n	n	ZMAT3	xuôi dòng
chr3:178716601-178716740		0,117%	0,00029		n	n	MIR5093	giữa các gen
chr16:85350785-85350924		0,156%	0,00045		y	n	DUSP14	intron
chr17:35861457-35861634		0,221%	0,02113		n	n	PRR15	giữa các gen
chr7:29662136-29662313		0,179%	0,08897		n	n	GIPR	giữa các gen
chr19:46189232-46189371		0,171%	0,09034		n	n	NEUROG3	xuôi dòng
chr10:71330866-71331005		0,095%	0,24740		n	n	COMMD1	giữa các gen
chr2:62391864-62392042		0,147%	0,54640		y	n	CDC25C	intron
chr5:137673092-137673271		0,315%	1,00000		y	n	KLHL29	intron
chr2:23925597-23925776		0,132%	1,00000		y	n	EPB41L3	intron
chr18:5447651-5447812		0,128%	1,00000		y	n	RAPGEF4	exon
chr2:173883408-173883572		0,106%	1,00000		y	y	CACTIN	intron
chr19:3625953-3626129		0,094%	1,00000		y	n	GLT8D2	intron
chr12:104442523-104442688		0,085%	1,00000		y	n		

Fig. 10B

đơn vị siêu sao chép_vị trí	tần số đột biến gần xen và gây mất bazo_được xử lý	bonferroni_p	giá trị	thuộc gen	mã hóa	gen	vị trí
chr2:60722337-60722476	74,09%		0,00000	n	n	DSCC1	ngược dòng
chr8:120368614-120868752	10,36%		0,00000	n	n	COMMD1	giữa các gen
chr2:62391864-62392042	1,13%		0,00000	y	n	LRRC1	intron
chr6:53670690-53670868	0,65%		0,00000	y	n	KLHL29	intron
chr2:23925597-23925775	0,58%		0,00000	y	n	INPP5A	intron
chr10:134468242-134468405	0,52%		0,00000	n	n	ACOT11	ngược dòng
chr1:54972393-54972568	0,51%		0,00000	n	n	NEUROG3	xuôi dòng
chr10:71330866-71331005	0,49%		0,00000	y	n	CTNBL1	intron
chr20:36335775-36335954	0,38%		0,00000	y	n	MOSPD1	intron
chrX:134048679-134048858	0,35%		0,00000	n	n		giữa các gen
chr14:67888671-67888848	0,30%		0,00000	y	n	COMMD6	exon 1
chr13:76123402-76123581	0,29%		0,00000	n	n	MMP27	ngược dòng
chr11:102556535-102556674	0,27%		0,00000				
MULTIPLE_TIED_LOCI	0,26%		0,00000				
MULTIPLE_TIED_LOCI	0,22%		0,00000				
chr11:33915930-33916106	0,22%		0,00000	n	n	LMO2	ngược dòng
chrX:30298492-302989671	0,20%		0,00372	n	n	MAGEB1	giữa các gen
chr17:27181433-27181606	0,18%		0,00206	n	n	ERAL1	ngược dòng
chr1:29508603-29508766	0,16%		1,00000	n	n	SRSF4	ngược dòng
chr12:104442523-104442688	0,11%		0,00077	y	n	GLT8D2	intron
chr19:45903865-45910028	0,11%		1,00000	y	n	CD3EAP	exon 1

Fig. 11A

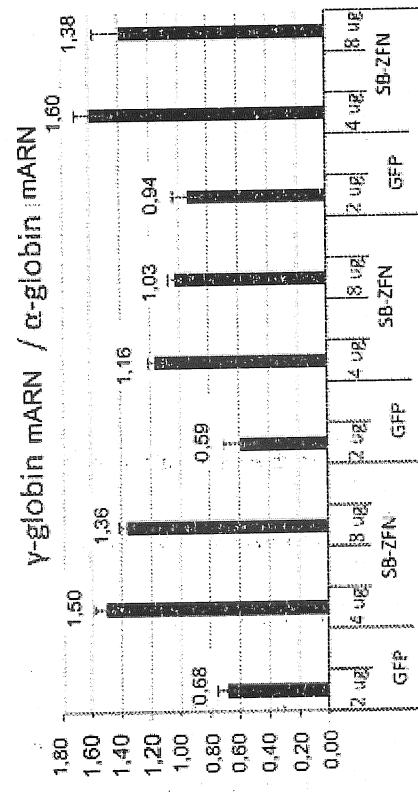
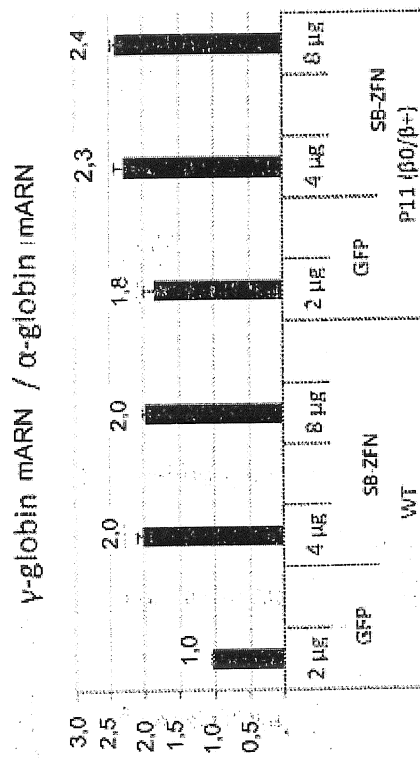


Fig. 11B

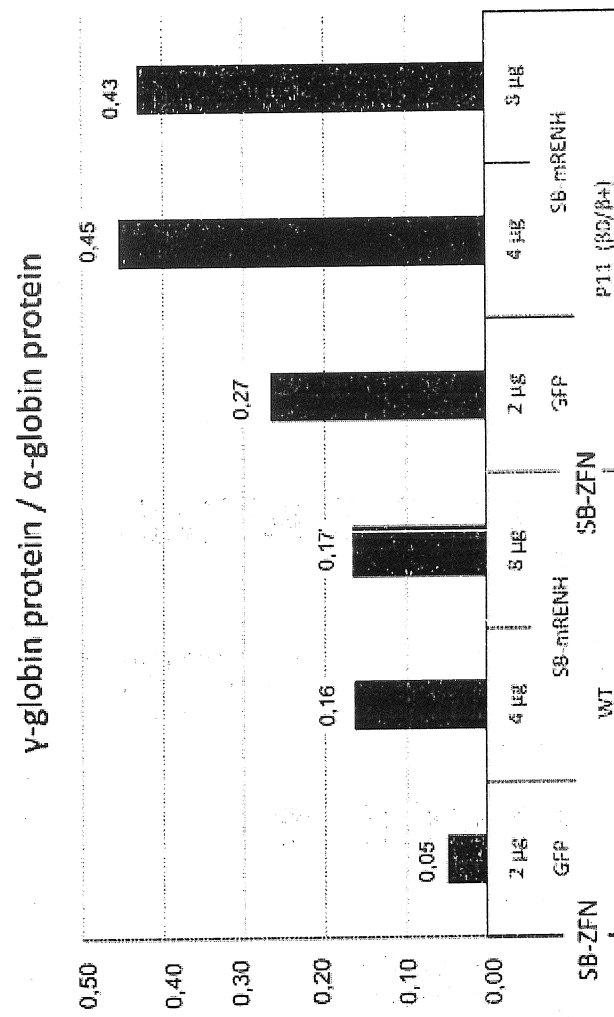


Fig. 12A

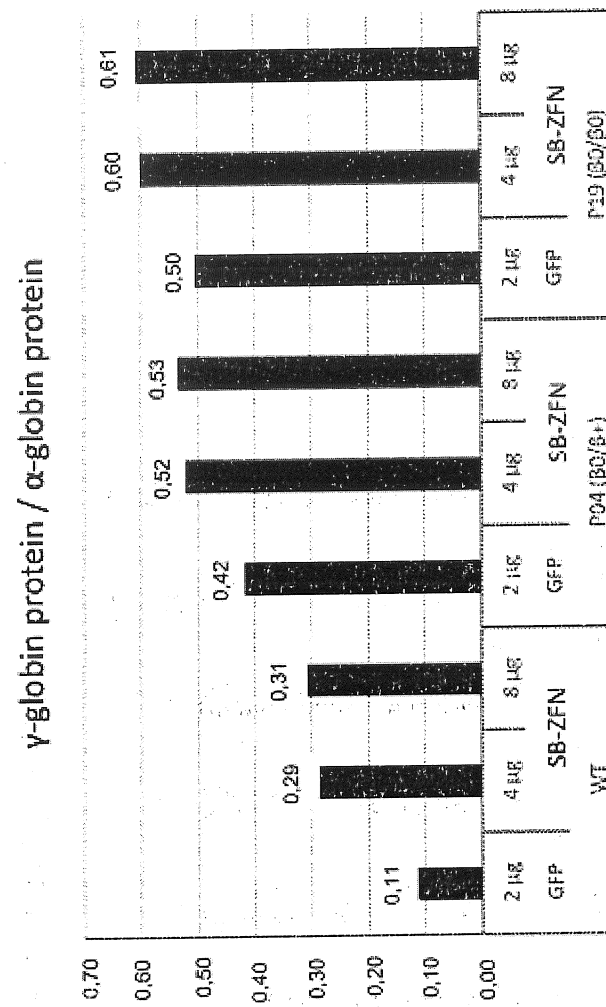
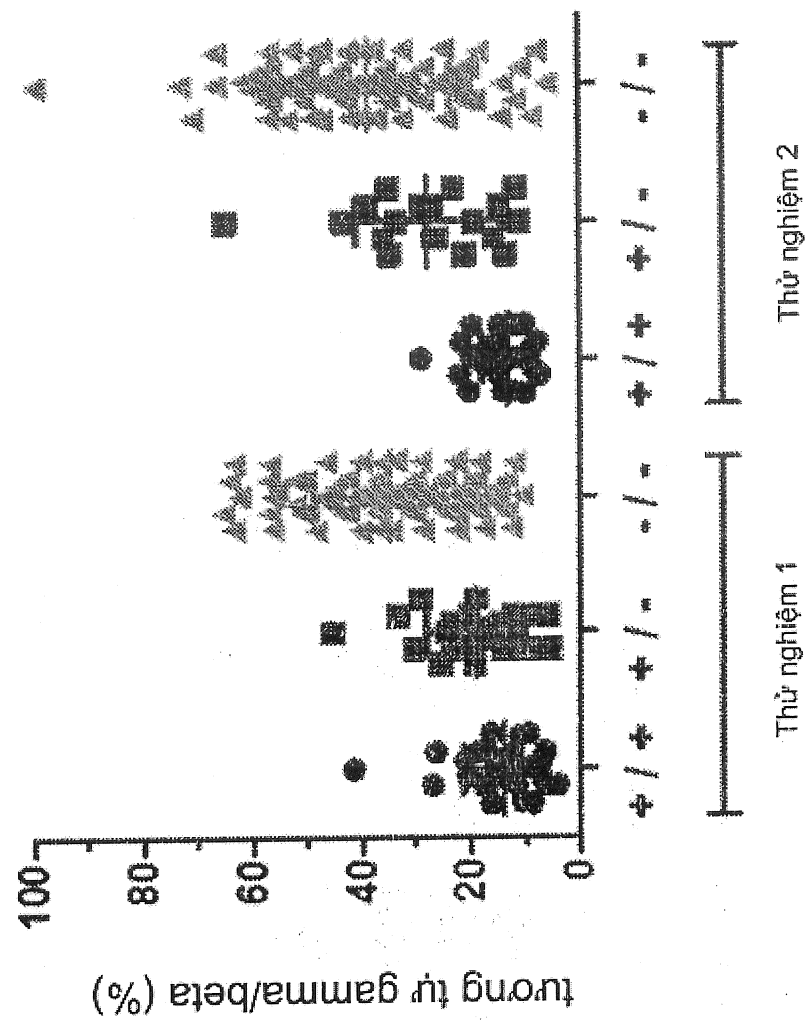


Fig. 12B

Fig. 13



DANH MỤC TRÌNH TỰ

<110> SANGAMO BIOSCIENCES, INC.

<120> PROTEIN NGÓN TAY KẼM, PROTEIN DUNG HỢP CHỨA PROTEIN NGÓN TAY KẼM VÀ ĐƯỢC PHẨM CHỨA PROTEIN DUNG HỢP NÀY

<130> 8328-0132.40

<140> PCT/US2016/032049

<141> Ngày 12-05-2016

<150> 62/303.595

<151> Ngày 04-03-2016

<150> 62/160.396

<151> Ngày 12-05-2015

<160> 38

<170> PatentIn phiên bản 3.5

<210> 1

<211> 28

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Mô tả Trình tự nhân tạo: Oligonucleotit tổng hợp

<400> 1

aaagcaactg ttagcttgca ctagacta

28

<210> 2

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 2

Asp Gln Ser Asn Leu Arg Ala

1

5

<210> 3

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 3

Arg Pro Tyr Thr Leu Arg Leu

1

5

<210> 4

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 4

Ser Gly Tyr Asn Leu Glu Asn

1

5

<210> 5

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 5

Thr Ser Gly Ser Leu Thr Arg

1

5

<210> 6

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 6

Ala Gln Cys Cys Leu Phe His

1

5

<210> 7

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 7

Ser Thr Gly Asn Leu Thr Asn

1

5

<210> 8

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 8

Ser Arg Gly Ala Leu Lys Thr

1

5

<210> 9
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 9
 Arg Asn Phe Ser Leu Thr Met
 1 5

<210> 10
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 10
 Ser Asn Gly Asn Leu Arg Asn
 1 5

<210> 11
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 11
 Ser Ser Tyr Asn Leu Ala Asn
 1 5

<210> 12
 <211> 28
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Mô tả trình tự nhân tạo: Oligonucleotit tổng hợp

<400> 12
 cacaggctcc aggaagggtt tggcctct

28

<210> 13
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 13
Arg Ser Asp His Leu Thr Gln
1 5

<210> 14
<211> 7
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 14
Gln Ser Gly His Leu Ala Arg
1 5

<210> 15
<211> 7
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 15
Gln Lys Gly Thr Leu Gly Glu
1 5

<210> 16
<211> 7
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 16
Gln Ser Ser Asp Leu Ser Arg
1 5

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

<220>
<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 17
Arg Arg Asp Asn Leu His Ser
1 5

<210> 18
<211> 7

<212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 18
 Arg His Arg Asp Leu Ser Arg
 1 5

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

<220>
 <223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 19
 Arg Asn Asp His Arg Thr Thr
 1 5

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

<220>
 <223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 20
 Gln Lys Ala His Leu Ile Arg
 1 5

<210> 21
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 21
 Arg Gly Arg Asp Leu Ser Arg
 1 5

<210> 22
 <211> 7
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 22

Gln Arg Ala His Leu Thr Arg
1 5

<210> 23

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 23

His Arg Asn Thr Leu Val Arg
1 5

<210> 24

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 24

Gln Ser Gly Thr Arg Asn His
1 5

<210> 25

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 25

Leu Lys Arg Thr Leu Lys Arg
1 5

<210> 26

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 26

Ser Ser Ser Asn Leu Thr Asn
1 5

<210> 27

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 27

Ser Ser Ser Asn Leu Gly Asn
1 5

<210> 28

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 28

Ser Arg Ser Ala Leu Arg Val
1 5

<210> 29

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 29

Arg Ser Arg Asp Leu Thr Arg
1 5

<210> 30

<211> 7

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 30

Phe Arg Gln Thr Arg Ala Arg
1 5

<210> 31

<211> 26

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 31

His	Thr	Lys	Ile	His	Leu	Arg	Gly	Ser	Gln	Leu	Val	Lys	Ser	Lys	Ser
1				5				10						15	

Glu	Ala	Ala	Ala	Arg	Glu	Leu	Glu	Glu	Lys
			20				25		

<210> 32

<211> 25

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 32

His	Thr	Lys	Ile	His	Leu	Arg	Gly	Ser	Ile	Ser	Arg	Ala	Arg	Pro	Leu
1				5				10						15	

Asn	Pro	His	Pro	Glu	Leu	Glu	Glu	Lys
			20				25	

<210> 33

<211> 25

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 33

His	Thr	Lys	Ile	His	Leu	Arg	Gly	Ser	Ile	Ser	Arg	Ala	Arg	Pro	Leu
1				5				10						15	

Asn	Pro	His	Pro	Glu	Leu	Glu	Glu	Lys
			20				25	

<210> 34

<211> 26

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Peptit tổng hợp

<400> 34

His	Thr	Lys	Ile	His	Leu	Arg	Gly	Ser	Tyr	Ala	Pro	Met	Pro	Pro	Leu
1				5				10						15	

Ala	Leu	Ala	Ser	Pro	Glu	Leu	Glu	Glu	Lys
			20				25		

<210> 35
 <211> 43
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Mô tả trình tự nhân tạo: Đoạn mỗi tổng hợp

<220>
 <221> bazo_biến đổi
 <222> (21)..(24)
 <223> a, c, t, g, chưa biết hoặc khác

<400> 35
 acacgacgct cttccgatct nnnnagtcct cttctacccc acc

43

<210> 36
 <211> 42
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Mô tả trình tự nhân tạo: Đoạn mỗi tổng hợp

<400> 36
 gacgtgtgct cttccgatct ctactcttag acataacaca cc

42

<210> 37
 <211> 93
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Mô tả trình tự nhân tạo: Oligonucleotit tổng hợp

<220>
 <221> bazo_biến đổi
 <222> (1)..(4)
 <223> a, c, t, g, chưa biết hoặc khác

<220>
 <221> bazo_biến đổi
 <222> (68)..(72)
 <223> a, c, t, g, chưa biết hoặc khác

<400> 37
 nnnnagtagt gtgtgcccgt ctgttggtgtg actctggtaa cttagatcc ctcagaccct

60

tttagtcnnn nnagtgtgga aaatctctag cag

93

<210> 38
 <211> 93
 <212> ADN
 <213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: Oligonucleotit tổng hợp

<220>

<221> bazơ biến đổi

<222> (1)..(4)

<223> a, c, t, g, chưa biết hoặc khác

<220>

<221> bazơ biến đổi

<222> (26)..(30)

<223> a, c, t, g, chưa biết hoặc khác

<400> 38

nnnctgcta gagattttcc acactnnnnn gactaaaagg gtctgaggga tctctagtta 60

ccagagtcac acaacagacg ggcacacact act 93