



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 338 952**

51 Int. Cl.:

C12N 15/11 (2006.01)

C12N 5/10 (2006.01)

C12N 15/66 (2006.01)

A01K 67/027 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **95938418 .1**

96 Fecha de presentación : **06.11.1995**

97 Número de publicación de la solicitud: **0791058**

97 Fecha de publicación de la solicitud: **27.08.1997**

54

Título: **Uso de una secuencia de nucleótidos que codifica la enzima I-Sce I.**

30

Prioridad: **07.11.1994 US 336241**

73

Titular/es: **INSTITUT PASTEUR**
25-28, rue du Docteur Roux
75724 Paris Cédex 15, FR
Université Pierre et Marie Curie (Paris 6)

45

Fecha de publicación de la mención BOPI:
13.05.2010

72

Inventor/es: **Choulika, André;**
Perrin, Arnaud;
Dujon, Bernard y
Nicolas, Jean-François

45

Fecha de la publicación del folleto de la patente:
13.05.2010

74

Agente: **Carvajal y Urquijo, Isabel**

ES 2 338 952 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Uso de una secuencia de nucleótidos que codifica la enzima I-Sce I.

5 Antecedentes de la invención

Esta solicitud divulga una secuencia de nucleótidos que codifica la endonucleasa de restricción I-SceI. Esta solicitud también describe vectores que contienen la secuencia de nucleótidos, células transformadas con los vectores, animales transgénicos con base en los vectores, y líneas de células derivadas de células en los animales. Esta solicitud también describe el uso de I-Sce I para el mapeo de genomas de eucariotas y para recombinación genética *in vivo* dirigida al sitio.

La habilidad para introducir genes en la línea germinal de mamíferos es de gran interés en biología. La propensión de células de mamífero para tomar ADN añadido en forma endógena y para expresar genes incluidos en el ADN se conoce desde hace muchos años. Los resultados de la manipulación genética son heredados por la descendencia de estos animales. Las células de estos descendientes heredan el gen introducido como parte de su carga genética. Se dice que tales animales son transgénicos.

Los mamíferos transgénicos han suministrado una forma de estudiar la regulación de los genes durante la embriogénesis y en la diferenciación, para estudiar la acción de los genes, y para estudiar la intrincada interacción de las células en el sistema inmunológico. El animal completo es el sistema último de análisis para genes manipulados, que dirigen un proceso biológico complejo.

Los animales transgénicos pueden proporcionar un ensayo general para analizar funcionalmente minuciosamente secuencias de ADN responsables de la regulación específica o de desarrollo del tejido de una variedad de genes. Además, los animales transgénicos proporcionan vehículos útiles para expresar proteínas recombinantes y para generar modelos animales precisos de trastornos genéticos humanos.

Para una discusión general de clonación y expresión génica en animales y en células de animales, ver Old and Primrose, "Principles of Gene Manipulation," Blackwell Scientific Publications, Londres (1989), página 255 y subsiguientes.

Las líneas transgénicas, que tienen una predisposición a enfermedades específicas y trastornos genéticos, son de gran valor en la investigación de los eventos que conducen a estos estados. Es bien conocido que la eficacia del tratamiento de un trastorno genético puede depender de la identificación del defecto genético que es la causa principal del trastorno. El descubrimiento de tratamientos efectivos se puede acelerar suministrando un modelo animal que conducirá a la enfermedad o al trastorno, lo cual permitirá el estudio de la eficacia, seguridad y modo de acción de los protocolos de tratamiento, tal como la recombinación genética.

Uno de los tejidos clave para la comprensión de la recombinación genética es la naturaleza de la etapa de iniciación. Los estudios de recombinación homóloga en bacterias y en hongos han dado lugar a la propuesta de dos tipos de mecanismos de iniciación. En el primer modelo, un corte de la cadena monocatenaria inicia la asimilación de la cadena y la migración de la rama (Meselson y Radding, 1975). Alternativamente, puede ocurrir un rompimiento bicatenario, seguido por un mecanismo de reparación que utiliza una secuencia homóloga no escindida como molde (Resnick y Martin, 1976). Este último modelo ha ganado soporte a partir del hecho de que se incrementa dramáticamente la transformación integradora en levadura cuando el plásmido transformante se linealiza en la región de homología cromosómica (Orr-Weaver, Szostak y Rothstein, 1981) y a partir de la observación directa de un rompimiento bicatenario durante la interconversión de tipo apareamiento (Strathern y colaboradores, 1982). Recientemente, los rompimientos bicatenarios también han sido caracterizados durante recombinación meiótica normal de levadura (Sun y colaboradores, 1989; Alani, Padmore y Kleckner, 1990).

Se han caracterizado diferentes actividades de endonucleasa bicatenaria en levadura: las endonucleasas codificadas por HO y por el intrón están asociadas con funciones de recombinación homóloga, mientras que otras aún tienen funciones genéticas desconocidas (Endo-SceI, Endo-SceII) (Shibata y colaboradores, 1984; Morishima y colaboradores, 1990). La endonucleasa específica del sitio HO inicia la interconversión de tipo apareamiento provocando un rompimiento bicatenario cerca de la unión YZ de MAT (Kostriken y colaboradores, 1983). El rompimiento es posteriormente reparado utilizando las secuencias intactas HML o HMR y resultando en conversión ectópica del gen. El sitio de reconocimiento HO es una secuencia no simétrica degenerada de 24 pb (Nickoloff, Chen, y Heffron 1986; Nickoloff, Singer y Heffron 1990). Esta secuencia ha sido utilizada como un "recombinador" en construcciones artificiales para promover recombinación mitótica y meiótica intra e intermolecular (Nickoloff, Chen y Heffron, 1986; Kolodkin, Klar y Stahl 1986; Ray y colaboradores, 1988; Rudin y Haber, 1988; Rudin, Sugarman, y Haber 1989).

Las endonucleasas específicas de dos sitios, I-Sce I (Jacquier y Dujon, 1985) y I-SceII (Delahodde y colaboradores, 1989; Wenzlau y colaboradores, 1989), que son responsables por la movilidad del intrón en mitocondria, inician una conversión génica que se parece a la conversión inducida por HO (ver Dujon, 1989 para revisión). I-Sce I, que es codificada por el intrón opcional Sc LSU.1 del gen de ARNr 21S, inicia un rompimiento bicatenario en el sitio de inserción del intrón (Macreadie y colaboradores, 1985; Dujon y colaboradores, 1985; Colleaux y colaboradores, 1986). El sitio de reconocimiento de I-Sce I se extiende sobre una secuencia no simétrica de 18 pb (Colleaux y colaboradores,

1988). Aunque las dos proteínas no están obviamente relacionadas por su estructura (HO tiene una longitud de 586 aminoácidos mientras que I-Sce I tiene una longitud de 235 aminoácidos), ambas generan cortes escalonados de 4 pb con proyecciones 3'OH dentro de sus respectivos sitios de reconocimiento. Se ha encontrado que una endonucleasa codificada en el intrón mitocondrial, transcrita en el núcleo y traducida en el citoplasma, genera un rompimiento bicatenario en un sitio nuclear. Los eventos de reparación inducidos por I-Sce I son idénticos a aquellos iniciados por HO.

En resumen, existe la necesidad en el arte de reactivos y métodos para proporcionar modelos de animales transgénicos de enfermedades humanas y trastornos genéticos. Los reactivos se pueden basar en la enzima de restricción I-Sce I y en el gen que codifica esta enzima. En particular, existe la necesidad de reactivos y de métodos para reemplazar un gen natural con otro gen que sea capaz de aliviar la enfermedad o el trastorno genético.

La presente invención cumple con esta necesidad suministrando métodos, animales transgénicos no humanos y células recombinantes como las descritas en el conjunto de las reivindicaciones.

Resumen de la invención

Por lo tanto, esta invención ayuda a satisfacer estas necesidades en el estado del arte. Específicamente, esta solicitud describe un ADN aislado que codifica a la enzima I-SceI. El ADN tiene la siguiente secuencia de nucleótidos:

```

1  ATG CAT ATG AAA AAC ATC AAA AAA AAC CAG GTA ATG AAC CTG GGT CCG AAC TCT AAA CTG
1  M  H  K  K  N  I  K  K  N  Q  V  M  N  L  G  P  N  S  K  L

61  CTG AAA GAA TAC AAA TCC CAG CTG ATC GAA CTG AAC ATC GAA CAG TTC GAA GCA GGT ATC
21  L  K  E  Y  K  S  Q  L  I  E  L  N  I  E  Q  F  E  A  G  I

121 GGT CTG ATC CTG GGT GAT GCT TAC ATC CGT TCT CGT GAT GAA GGT AAA ACC TAC TGT ATG
41  G  L  I  L  G  D  A  Y  I  R  S  R  D  E  G  K  T  Y  C  M

181 CAG TTC GAG TGG AAA AAC AAA GCA TAC ATG GAC CAC GTA TGT CTG CTG TAC GAT CAG TGG
61  Q  F  E  W  K  N  K  A  Y  M  D  H  V  C  L  L  Y  D  Q  W

241 GTA CTG TCC CCG CCG CAC AAA AAA GAA CGT GTT AAC CAC CTG GGT AAC CTG GTA ATC ACC
81  V  L  S  P  P  H  K  K  E  R  V  N  H  L  G  N  L  V  I  T

301 TGG GGC GCC CAG ACT TTC AAA CAC CAA GCT TTC AAC AAA CTG GCT AAC CTG TTC ATC GTT
101 W  G  A  Q  T  F  K  H  Q  A  F  N  K  L  A  N  L  F  I  V

361 AAC AAC AAA AAA ACC ATC CCG AAC AAC CTG GTT GAA AAC TAC CTG ACC CCG ATG TCT CTG
121 N  N  K  K  T  I  P  N  N  L  V  E  N  Y  L  T  P  M  S  L

421 GCA TAC TGG TTC ATG GAT GAT GGT GGT AAA TGG GAT TAC AAC AAA AAC TCT ACC AAC AAA
141 A  Y  W  F  M  D  D  G  G  K  W  D  Y  N  K  N  S  T  N  K

481 TCG ATC GTA CTG AAC ACC CAG TCT TTC ACT TTC GAA GAA GTA GAA TAC CTG GTT AAG GGT
161 S  I  V  L  N  T  Q  S  F  T  F  E  E  V  E  Y  L  V  K  G

541 CTG CGT AAC AAA TTC CAA CTG AAC TGT TAC GTA AAA ATC AAC AAA AAC AAA CCG ATC ATC
181 L  R  N  K  F  Q  L  N  C  Y  V  K  I  N  K  N  K  P  I  I

601 TAC ATC GAT TCT ATG TCT TAC CTG ATC TTC TAC AAC CTG ATC AAA CCG TAC CTG ATC CCG
201 Y  I  D  S  M  S  Y  L  I  F  Y  N  L  I  K  P  Y  L  I  P

661 CAG ATG ATG TAC AAA CTG CCG AAC ACT ATC TCC TCC GAA ACT TTC CTG AAA TAA
221 Q  M  M  Y  K  L  P  N  T  I  S  S  E  T  F  L  K  *

```

Esta solicitud también describe una secuencia de ADN que comprende un promotor operativamente enlazado a la secuencia de ADN como se define en este documento, que codifica la enzima I-SceI.

Esta solicitud divulga además a un ARN aislado complementario a la secuencia de ADN como se define en este documento que codifica a la enzima I-Sce I y a las otras secuencias de ADN descritas aquí.

En otra modalidad, se proporciona un vector. El vector comprende un plásmido, bacteriófago, o vector cósmido que contiene la secuencia de ADN como se define aquí, que codifica a la enzima I-SceI.

Además, esta solicitud se relaciona con células eucariotas o de *E. coli* transformadas con un vector como se define en este documento.

También, esta solicitud se relaciona con animales transgénicos que contienen la secuencia de ADN que codifica la enzima I-Sce I y líneas de células cultivadas a partir de los animales transgénicos.

Esta invención se relaciona con un organismo transgénico en el cual se ha insertado al menos un sitio de restricción para la enzima I-Sce I en un cromosoma del organismo.

La solicitud se relaciona también con un método de mapear genéticamente un genoma eucariota utilizando la enzima I-SceI.

Esta solicitud también se relaciona con un método para recombinación *in vivo* dirigida al sitio en un organismo utilizando la enzima I-SceI.

Breve descripción de los dibujos

Se describirá más completamente esta invención con referencia a los dibujos en los cuales: La Figura 1 describe el código universal equivalente del gen mitocondrial para I-SceI.

La Figura 2 describe la secuencia de nucleótidos de la invención que codifica a la enzima I-Sce I y la secuencia de aminoácidos de la enzima natural I-SceI.

1 y 2: Estos aminoácidos son absolutamente necesarios para producir actividad catalítica. Otras sustituciones son posibles tales como supresiones de los 10 primeros aminoácidos.

La Figura 3 describe la secuencia de reconocimiento de I-Sce I e indica posibles mutaciones de bases en el sitio de reconocimiento y el efecto de tales mutaciones sobre la rigurosidad del reconocimiento.

La Figura 4 es la secuencia de nucleótidos y la secuencia deducida de aminoácidos de una región del plásmido pSCM525. La secuencia de nucleótidos de la invención que codifica la enzima I-Sce I está incluida en la caja.

La Figura 5 describe variaciones alrededor de la secuencia de aminoácidos de la enzima I-SceI.

La Figura 6 muestra al intrón del Grupo I que codifica endonucleasas y endonucleasas relacionadas.

La Figura 7 describe vectores de expresión de levadura que contienen al gen sintético para I-SceI.

La Figura 8 describe el vector de expresión en mamíferos PRSV de I-SceI.

La Figura 9 es un mapa de restricción del plásmido pAF100. (Ver también YEAST, 6: 521-534, 1990, que se invoca y se incorpora en el presente documento como referencia).

Las Figuras 10A y 10B muestran la secuencia de nucleótidos y los sitios de restricción de regiones del plásmido pAF100.

La Figura 11 describe un vector de inserción pTSM ω , pTKM ω , y pTTc ω que contienen el sitio I-Sce I para *E. coli* y otras bacterias.

La Figura 12 describe un vector de inserción pTYW6 que contiene el sitio de I-Sce I para levadura.

La Figura 13 describe un vector de inserción PMLV LTR SAPLZ que contiene el sitio de I-Sce I para células de mamífero.

La Figura 14 describe un conjunto de siete cepas de levadura transgénica escindidas por I-SceI. Cromosomas de FY1679 (control) y de siete cepas de levadura transgénica con sitios de I-Sce I insertados en diferentes posiciones a lo largo del cromosoma XI fueron tratados con I-SceI. Se sometió a electroforesis el ADN sobre gel de agarosa al 1% (SeaKem) en amortiguador TBE 0,25X a 130 V y 12°C sobre un aparato Rotaphor (Biometra) durante 70 h utilizando tiempos de reducción de pulso de 100 s a 40 s. (A) Se tiñó el ADN con bromuro de etidio (0,2 μ g/ml) y se lo transfirió a una membrana Hybond N (Amersham) para hibridación. (B) Se utilizó el cósmido pUKG040 marcado con ³²P que hibrida con el fragmento más corto del conjunto como sonda. Se indican las posiciones del cromosoma XI y los cromosomas más cortos.

La Figura 15 describe la razón de la estrategia de fragmentación cromosómica anidada para mapeo genético. (A) Las posiciones de los sitios de I-Sce I se colocan sobre el mapa, independientemente de la orientación izquierda/derecha (los fragmentos más cortos se colocan arbitrariamente sobre la izquierda). El tamaño de los fragmentos medidos a partir de PFGE (Fig. 14A) se indican en kb (obsérvese que la suma de los dos tamaños de fragmentos varía ligeramente debido al límite de precisión de cada medición). (B) La hibridación con la sonda que hibrida al fragmento más corto del conjunto determina la orientación de cada fragmento (ver la Fig. 14B). Los fragmentos que hibridan con la sonda (líneas gruesas continuas) han sido colocados arbitrariamente a la izquierda. (C) Las cepas de levadura transgénica han sido ordenadas con tamaños crecientes de fragmentos de hibridación de cromosomas. (D) Mapa deducido de I-Sce I con el tamaño máximo y mínimo de los intervalos indicado en kb (las variaciones en algunos intervalos son debidas a limitaciones de las mediciones de PFGE). (E) Los subfragmentos de cromosoma se utilizan como sondas para asignar cada clon cósmido a un intervalo dado del mapa o a través de un sitio dado de I-SceI.

La Figura 16 describe el mapeo de los sitios de I-Sce I de cepas de levadura transgénica por hibridación con sondas del extremo izquierdo y del extremo derecho del cromosoma XI. Se digirieron los cromosomas de FY1679 (control)

y las siete cepas de levadura transgénica con I-SceI. Se colocaron las cepas transgénicas en orden como se explica en la Fig. 15. Las condiciones de la electroforesis fueron como en la Fig. 14. Se utilizaron los cósmidos pUKG040 y pUKG066 marcados con ^{32}P como sondas del extremo izquierdo y del extremo derecho, respectivamente.

La Figura 17 describe el mapeo de una colección de cósmidos utilizando los fragmentos cromosómicos anidados como sondas. Se digirieron los ADN del cósmido con *EcoRI* y se los sometió a electroforesis sobre gel de agarosa al 0,9% (SeaKem) a 1,5 V/cm durante 14 h, se los tiñó con bromuro de etidio y se los transfirió a una membrana Hybond N. Se colocaron los cósmidos en orden a partir de hibridaciones previas para ayudar a visualizar la estrategia. Se llevaron a cabo las hibridaciones en forma serial sobre tres membranas idénticas utilizando fragmentos de cromosoma anidados del extremo izquierdo purificados sobre PFGE (ver la Fig. 16) como sondas. A: coloración con bromuro de etidio (la escala es la "escala de 1 kb" de BRL), B: membrana #1, sonda: Tel izquierdo hasta el sitio A302, C: membrana #1, sonda: Tel izquierdo hasta el sitio M57, D: membrana #2, sonda: Tel izquierdo hasta el sitio H81, E: membrana #2, sonda: Tel izquierdo hasta el sitio T62, F: membrana #3, sonda: Tel izquierdo hasta el sitio G41, G: membrana #3, sonda: Tel izquierdo hasta el sitio D304, H: membrana #3, sonda: cromosoma XI entero.

La Figura 18 describe un mapa del cromosoma XI de levadura como según lo determinado a partir de la estrategia de fragmentación cromosómica anidada. Se divide el cromosoma en ocho intervalos (con los tamaños indicados en kb, ver la Fig. 15D) separados por siete sitios de I-Sce I (E40, A302...). Los clones de cósmido que caen ya sea dentro de los intervalos o a través de un sitio dado de I-Sce I están enlistados debajo de los intervalos o debajo de los límites de los intervalos, respectivamente. Los clones de cósmido que hibridan con genes seleccionados utilizados como sondas se indican por medio de letras (a-i). Ellas localizan al gen con respecto al mapa de I-Sce I y permiten la comparación con el mapa genético (parte superior).

La Figura 19 describe diagramas de experimentos exitosos de recombinación homóloga dirigida al sitio, llevados a cabo en levadura.

Figura 20. Diseño experimental para la detección de HR inducida por I-SceI. a) Mapas del retrovirus tk-*PhleoLacZ* de 7,5 kb (G-MtkPL) y del retrovirus *PhleoLacZ* de 6,0 kb (G-MPL), SA es un sitio aceptor de empalme. Las secuencias de G-MtkPL (del virus G-MtkPL) contienen al gen de fusión *PhleoLacZ* para selección positiva de células infectadas (en medio que contiene fleomicina) y el gen tk para selección negativa (en medio que contiene ganciclovir). Las secuencias de G-MPL (del virus G-MPL) contienen únicamente secuencias de *PhleoLacZ*. b) Mapas de estructuras provirales después de la integración retroviral de G-MtkPL y G-MPL. La LTR de *PhleoLacZ* para I-Sce I se duplica, colocando las secuencias de *PhleoLacZ* para I-Sce I en la LTR 5'. Se transcribe el vector viral (que actúa como una trampa del promotor) (flecha) por medio de un promotor de flanqueo celular, P. c) I-Sce I crea dos rompimientos bicatenarios (DSB) en ADN del huésped liberando el segmento central y dejando los extremos rotos del cromosoma que pueden aparearse con el plásmido donante, pVRneo (d). e) Locus recombinante esperado después de HR.

Figura 21. A. Esquema de pG-MPL. SD y SA son sitios aceptores de empalme y del donante de empalme. La estructura de los transcritos empalmados de 4,2 kb y no empalmados de 5,8 kb (genómicos) se muestra más adelante. La barra pesada es la sonda *LacZ* marcada en forma radioactiva con ^{32}P (P). B. Análisis de transferencias tipo Northern de ARN de un clon productor ψ -2 transformado con pG MLP utilizando ARN poliadenilado. Obsérvese que el ARNm empalmado y el genómico se producen en el mismo nivel alto.

Figura 22. A. Introducción de sitios de reconocimiento duplicados de I-Sce I en el genoma de células de mamífero por medio de integración de retrovirus. Esquema de los provirus G-MPL y G-MtkPL que ilustran las posiciones de las dos LTR y los sitios de restricción pertinentes. Se indican los tamaños de los fragmentos de *Bcl I* y de los fragmentos I-Sce I. La barra pesada es la sonda *LacZ* marcada en forma radioactiva con ^{32}P (P). B. Análisis de transferencias tipo Southern de ADN celular de células de fibroblastos NIH3T3 infectadas por G-MtkPL y células multipotentes PCC7-S infectadas por G-MPL. *Bcl I* digiere demostrando duplicación de *PhleoLacZ* mediada por LTR; I-Sce I digiere demostrando duplicación fiel de sitios de I-Sce I.

Figura 23. Verificación de la recombinación por Southern. A.: Tamaños de fragmento esperados en pares de kilobases (kb) de provirus en el locus recombinante. 1) el locus proviral de los progenitores. La barra pesada (P) es la sonda marcada en forma radioactiva con ^{32}P utilizada para hibridación. 2) un recombinante derivado después de escisión en los dos sitios de I-Sce I seguido por la reparación del vacío utilizando pVR neo (recombinación homóloga de sitio doble, DsHR). 3) un evento de recombinación iniciado por la escisión en los sitios de I-Sce I en la LTR izquierda (recombinación homóloga de un solo sitio, SsHR). B.: análisis tipo Southern del ADN de los clones NIH3T3/GMtkPL 1 y 2, los clones PCC7-S/G-MPL 3 y 4 y de los transformantes derivados de la cotransfección con pCMV (I-Sce I+) y con pVRneo (1a, 1b, 2a, 3a, 3b y 4a). La digestión con *Kpn I* del ADN progenitor genera un fragmento de 4,2 kb que contiene *LacZ*. Los recombinantes 1a y 3a son ejemplos de los Recombinantes de DsHR 1b, 2a, 3b y 4a son ejemplos de SsHR.

Figura 24. Verificación de la recombinación por medio de los análisis de transferencias tipo Northern. A.: Estructura esperada y tamaños (en kb) del ARN de células del clon 3 PCC7-S/G-MPL antes (parte superior) y después (parte inferior). I-Sce I indujo HR con pVRneo. 1. Las barras pesadas P1 y P2 son sondas marcadas en forma radiactiva con ^{32}P . B.: Análisis de transferencias tipo Northern del clon 3 recombinante PCC7-S/G-MPL (ARN total). El carril 3 es de células progenitoras, el carril 3a de células recombinantes. Los dos primeros carriles fueron investigados con *LacZ* P1, los últimos dos carriles fueron investigados con P2 nuevo. Las células del clon 3 PCC7-S/G-MPL progenitor expresan

un ARN de LacZ de 7,0 kb como se esperaba de la captura de un promotor celular que conduce a la expresión de un ARN de fusión viral celular. El clon recombinante no expresa este ARN de LacZ pero expresa un ARN nuevo de 5,0 kb, correspondiente al tamaño esperado para un reemplazo preciso de PhleoLacZ por el gen nuevo.

Figura 25. Tipos de eventos de recombinación inducidos por los DSB de I-Sce I, a) Dibujo esquemático de la estructura del sustrato de recombinación. El G-MtkPL tiene dos LTR de provirus, conteniendo cada uno un sitio de reconocimiento de I-Sce I y un gen de PhleoLacZ. Las LTR se separan por medio de secuencias virales que contienen al gen tk. El fenotipo de las células que contienen G-MtkPL es Phleo^R, GIs^S, β -Gal $\pm$ b) Formas posibles de recombinación intracromosómica. 1) La I-Sce I endonucleasa corta el sitio de I-Sce I en la LTR 5'. La parte 5' de U3 de la LTR 5' puede aparearse y recombinarse con su secuencia homóloga en la LTR 3' (por medio de SSA). 2) La I-Sce I endonucleasa corta el sitio de I-Sce I en la LTR 3'. La parte 3' de U3 de la LTR 3' puede aparearse y recombinarse con su secuencia homóloga en la LTR 5' (por medio de SSA). 3) La I-Sce I endonucleasa corta sitios de I-Sce I en las dos LTR. Los dos extremos libres pueden relegar (por medio de un mecanismo de unión terminal). El producto resultante de la recombinación en cada uno de los tres modelos es una LTR solitaria (ver el lado derecho). Ninguna modificación se producirá en las secuencias celulares que flanquean el sitio de integración. c) La I-Sce I endonucleasa corta los sitios de I-Sce I en las dos LTR. Los dos extremos libres se pueden reparar (por medio de un mecanismo de reparación de vacíos) utilizando el cromosoma homólogo. Al lado derecho, el producto resultante de la recombinación es la supresión del locus de integración proviral.

Figura 26. Análisis de transferencias tipo Southern de ADN de los recombinantes NIH3T3/G-MtkPL 1 y 2, y PhleoLacZ⁻ derivados de transfecciones con pCMV(I-Sce I+) seleccionados en medio que contiene Ganciclovir. a) Tamaños esperados de los fragmentos en pares de kilobases (kpb) de provirus progenitor después de digestión con *Pst* I endonucleasa. La digestión con *Pst* I del ADN progenitor de NIH3T3/G-MtkPL 1 genera dos fragmentos de 10 kpb y del NIH3T3/G-MtkPL 2 progenitor dos fragmentos de 7 kpb y 9 kpb. b) Análisis de transferencias tipo Southern del ADN digerido por *Pst* I de NIH3T3/G-MtkPL 1, y recombinantes derivados de transfección con pCMV(I-Sce I+) (1,1 a 1,5). c) Análisis de transferencias tipo Southern del ADN digerido por *Pst* I de NIH3T3/G-MtkPL 2, y recombinantes derivados de la transfección con pCMV(I-Sce I+) (2,1 a 2,6).

La barra pesada es la sonda LacZ marcada en forma radioactiva con ³²P (P).

Figura 27. Análisis de transferencias tipo Southern del ADN de los recombinantes de NIH3T3/G-MtkPL 1 y 2, y PhleoLacZ⁺ derivados de las transfecciones con pCMV(I-Sce I+) y pCMV(I-Sce I-) y selección en medio que contiene Fleciclovir y Ganciclovir. a) Tamaños esperados de los fragmentos en kpb de provirus progenitor después de digestión con endonucleasa *Pst* I o *Bcl* I. La digestión con *Pst* I del ADN progenitor de NIH3T3/G-MtkPL 1 genera dos fragmentos de 10 kpb. La digestión *Bcl* I del ADN progenitor de NIH3T3/G-MtkPL 2 genera tres fragmentos de 9,2 kpb, 7,2 kpb y 6,0 kpb. a2) Tamaños esperados de los fragmentos en kpb de recombinantes después de digestión con endonucleasa *Pst* I o *Bcl* I. La digestión con *Pst* I del ADN del recombinante derivado de NIH3T3/G-MtkPL 1 genera un fragmento de 13,6 kpb. La digestión con *Bcl* I del ADN de los recombinantes derivados de NIH3T3/G-MtkPL 2 genera dos fragmentos de 9,2 kpb y 6,0 kpb. b) Análisis de transferencias tipo Southern del ADN de NIH3T3/G-MtkPL 1, y de los recombinantes derivados de la transfección con pCMV(I-Sce I-) y pCMV(I-Sce I+) (1c, 1d). c) Análisis de transferencias tipo Southern del ADN de NIH3T3/G-MtkPL 2, y de los transformantes derivados de la transfección con pCMV (I-Sce I-) (2a, 2b) y pCMV(I-Sce I+) (2c a 2h). La barra pesada es la sonda LacZ marcada en forma radioactiva con ³²P (P).

Figura 28. La Figura 28 es un diagrama que ilustra la pérdida de heterozigosidad por la inserción o la presencia de un sitio de I-Sce I, la expresión de la enzima I-Sce I, la escisión en el sitio, y la reparación del rompimiento bicatenario en el sitio con el cromátido correspondiente.

Figura 29. La Figura 29 es un diagrama que ilustra la activación condicional de un gen. Un sitio de I-Sce I está integrado entre repeticiones en tándem, y se expresa la enzima I-Sce I. La enzima escinde el ADN bicatenario en el sitio de I-Sce I. Se repara el rompimiento bicatenario por medio de apareamiento monocatenario, produciendo un gen activo.

Figura 30. La Figura 30 es un diagrama que ilustra una etapa de reordenamiento de un gen por medio de la integración de un sitio de I-Sce I o por medio del uso de un sitio de I-Sce I presente en el gen. Se introduce en la célula un plásmido que tiene ya sea un sitio de I-Sce I dentro de un gen inactivo, o dos sitios de I-Sce I en cualquier extremo de un gen activo sin un promotor. La célula contiene una forma inactiva del gen correspondiente. La enzima I-Sce I corta al plásmido en los sitios de I-Sce I, y la recombinación entre el cromosoma y el plásmido produce un gen activo que reemplaza al gen inactivo.

Figura 31. La Figura 31 es un diagrama que ilustra la duplicación de un locus. Se insertan un sitio de I-Sce I y una parte distal del locus dentro del gen por medio de un reemplazo clásico de genes. Se escinde el sitio de I-Sce I por medio de la enzima I-Sce I, y se repara el rompimiento por medio de secuencias homólogas. Esto resulta en la duplicación del locus entero.

Figura 32. La Figura 32 es un diagrama que ilustra la supresión de un locus. Se añaden dos sitios de I-Sce I para flanquear el locus que va a ser suprimido. Se expresa la enzima I-Sce I, y se escinden los sitios. Se recombinan los extremos restantes, suprimiendo el locus entre los dos sitios de I-Sce I.

Figura 33. La Figura 33 es un diagrama del plásmido pG-MtkDPAPL que muestra los sitios de restricción. El plásmido se construye por supresión de la región de poliadenilación del gen tk del plásmido pGMtkPL.

Descripción detallada de las modalidades preferidas

El gen mitocondrial genuino (ref. 8) no puede ser expresado en *E. coli*, levadura u otros organismos debido a las peculiaridades del código genético mitocondrial. Se ha construido un "equivalente del código universal" por medio de mutagénesis dirigida al sitio *in vitro*. Su secuencia aparece en la Fig. 1. Obsérvese que todos los codones no universales (excepto dos CTN) han sido reemplazados junto con algunos codones extremadamente raros en *E. coli*.

El equivalente del código universal ha sido expresado exitosamente en *E. coli* y determina la síntesis de una enzima activa. Sin embargo, los niveles de expresión permanecieron bajos debido al gran número de codones que son extremadamente raros en *E. coli*. La expresión del "equivalente del código universal" ha sido detectada en levadura.

Para optimizar la expresión del gen en sistemas heterólogos, se ha diseñado un gen sintético para codificar una proteína con la secuencia genuina de aminoácidos de I-Sce I utilizando, para cada codón, aquel más frecuentemente utilizado en *E. coli*. La secuencia del gen sintético se muestra en la Fig. 2. Se construyó el gen sintético *in vitro* a partir de ocho oligonucleótidos sintéticos con superposiciones parciales. Se diseñaron los oligonucleótidos para permitir cebado mutuo para la síntesis de una segunda hebra por medio de la polimerasa de Klenow cuando se aparean por pares. Se ligaron luego los pares alargados en los plásmidos. Los sitios de restricción apropiadamente colocados dentro de la secuencia diseñada permitieron un ensamblaje final del gen sintético por medio de ligación *in vitro*. El gen sintético ha sido exitosamente expresado tanto en *E. coli* como en levadura.

1. Secuencia del Gen para I-SceI

Esta solicitud se relaciona con una secuencia aislada de ADN que codifica a la enzima I-SceI. La enzima I-Sce I es una endonucleasa. Las propiedades de la enzima (ref. 14) son las siguientes:

I-Sce I es una endonucleasa bicatenaria que escinde el ADN dentro de su sitio de reconocimiento.

I-Sce I genera un corte escalonado de 4 pb con proyecciones OH 3'.

Sustrato: Actúa únicamente sobre ADN bicatenario. El ADN del sustrato puede ser relajado o superenrollado negativamente.

Cationes: La actividad enzimática requiera de Mg^{++} (8 mM es óptimo). Mn^{++} puede reemplazar a Mg^{++} , pero esto reduce la rigurosidad del reconocimiento.

Condiciones óptimas para la actividad: pH alto (9 a 10), temperatura 20-40°C, sin cationes monovalentes.

Estabilidad de la enzima: I-Sce I es inestable a temperatura ambiente. El complejo enzima-sustrato es más estable que la enzima sola (la presencia de sitios de reconocimiento estabiliza la enzima).

La enzima I-Sce I tiene un sitio de reconocimiento conocido (ref. 14). El sitio de reconocimiento de I-Sce I es una secuencia no simétrica que se extiende sobre 18 pb como se determina por medio de análisis mutacional sistemático. La secuencia se lee: (las flechas indican los cortes)

5' TAGGGATAACAGGGTAAT 3'

3' ATCCCTATTGTCCCATTA 5'

El sitio de reconocimiento corresponde, en parte, al exón secuencia arriba y, en parte, al exón secuencia abajo del intrón más la forma del gen.

El sitio de reconocimiento está parcialmente degenerado: las sustituciones de bases individuales dentro de la secuencia de 18 pb de longitud resulta o bien en una insensibilidad completa o en una insensibilidad reducida a la enzima, dependiendo de la posición y de la naturaleza de la sustitución.

- 1 - mutantes del sitio
- 2 - el genoma total de la levadura (*Saccharomyces cerevisiae*, la complejidad del genoma es $1,4 \times 10^7$ pb)

Los resultados son:

- 1 - Mutantes del sitio: Como se muestra en la Fig. 3, existe un cambio general de rigurosidad, es decir, los mutante severamente afectados en Mg^{++} se ven parcialmente afectados en Mn^{++} , los mutantes parcialmente afectados en Mg^{++} no se ven afectados en Mn^{++} .

- 2 - Levadura: En condiciones de magnesio, no se observa escisión en levadura normal. En la misma condición, el ADN de las levaduras transgénicas se escinde completamente en el sitio de I-Sce I insertado en forma artificial y no se puede detectar otro sitio de escisión. Si se reemplaza magnesio por manganeso, se revelan cinco sitios adicionales de escisión en el genoma completo de la levadura, ninguno de los cuales es escindido completamente. Por lo tanto, en manganeso la enzima revela un promedio de 1 sitio por aproximadamente 3 millones de pares de bases ($5/1,4 \times 10^7$ pb).

Definición del sitio de reconocimiento: las bases importantes están indicadas en la Fig. 3. Ellas corresponden a bases para las cuales existen mutantes severamente afectados. Obsérvese sin embargo que:

- 1 - No se han determinado todas las posibles mutaciones en cada posición; por lo tanto una base que no corresponde a un mutante severamente afectado puede aún ser importante si otro mutante fue examinado en esta misma posición.
- 2 - No existe un límite claro de corte entre una base muy importante (todos los mutantes son severamente afectados) y una base moderadamente importante (algunos de los mutantes son severamente afectados). Existe un continuo entre sustratos excelentes y sustratos pobres para la enzima.

La frecuencia esperada de sitios naturales de I-Sce I en una secuencia aleatoria de ADN es, por lo tanto, igual a $(0,25)^{-18}$ o $(1,5 \times 10^{-11})$. En otras palabras, uno debe esperar un sitio natural para el equivalente de aproximadamente 20 genomas humanos, pero la frecuencia de sitios degenerados es más difícil de predecir.

I-Sce I pertenece a una subfamilia “degenerada” de la familia de los dodecapéptidos. Los aminoácidos conservados de los motivos dodecapéptido son necesarios para la actividad. En particular, los residuos aspárticos en las posiciones 9 de los dos dodecapéptidos no pueden ser reemplazados, incluso con residuos glutámicos. Es probable que los dodecapéptidos formen el sitio catalítico o parte de él.

Consecuente con que el sitio de reconocimiento no sea simétrico, es probable que la actividad endonucleolítica de I-Sce I requiera de dos etapas sucesivas de reconocimiento: enlazamiento de la enzima a la mitad secuencia abajo del sitio (correspondiente al exón secuencia abajo). El primer enlazamiento es fuerte, el segundo es más débil, pero los dos son necesarios para la escisión del ADN. *in vitro*, la enzima puede enlazar solo al exón secuencia abajo así como a la secuencia de unión intrón-exón, pero sin producir escisión.

Los motivos dodecapéptido evolucionariamente conservados de I-Sce I codificado por el intrón son esenciales para la actividad de la endonucleasa. Se ha propuesto que el papel de estos motivos es la de posicionar adecuadamente los aminoácidos con respecto a los dominios de reconocimiento de la secuencia de ADN de la enzima para la catálisis de la hidrólisis del enlace fosfodiéster (ref. P3).

La secuencia de nucleótidos como se define aquí, que codifica a la enzima natural I-Sce I es mostrada en la Fig. 2. La secuencia de nucleótidos del gen se derivó por secuenciación del didesoxinucleótido. Las secuencias de las bases de los nucleótidos están escritas en la dirección $5' \rightarrow 3'$. Cada una de las letras mostrada es una designación convencional para los siguientes nucleótidos:

- A Adenina
- G Guanina
- T Timina
- C Citosina

Se prefiere que la secuencia de ADN que codifica a la enzima I-Sce I esté en forma pura. Por ejemplo, la secuencia puede estar libre de proteínas derivadas de sangre humana, proteínas de suero humano, proteínas virales, secuencias de nucleótidos que codifican estas proteínas, tejido humano, componentes de tejido humano, o combinaciones de estas sustancias. Además, se prefiere que la secuencia de ADN como la definida en este documento esté libre de proteínas extrañas y de lípidos, y de microorganismos accidentales, tales como bacterias y virus. La secuencia de ADN esencialmente purificada y aislada que codifica I-Sce I es especialmente útil para la preparación de vectores de expresión.

ES 2 338 952 T3

El plásmido pSCM525 es un derivado de pUC12, que contiene una secuencia artificial que codifica la secuencia de ADN de la invención. La secuencia de nucleótidos y la secuencia deducida de aminoácidos de una región del plásmido pSCM525 se muestran en la Fig. 4. La secuencia de nucleótidos que codifica I-Sce I está encerrada en el cajón. El gen artificial es una pieza *Bam*HI-*Sal*I de la secuencia de ADN de 723 pares de bases, químicamente sintetizada y ensamblada. Se coloca bajo el control del promotor *tac*. La secuencia de ADN del gen artificial difiere de la secuencia natural de codificación o su equivalente del código universal descrito en Cell (1986), Vol. 44, páginas 521-533. Sin embargo, el producto de la traducción del gen artificial es idéntico en secuencia a la endonucleasa omega genuina excepto por la adición de un Met-His en el terminal N. Se entiende que esta endonucleasa modificada está dentro del alcance de la presente invención.

El plásmido pSCM525 puede ser utilizado para transformar cualquier cepa adecuada de *E. coli* y las células transformadas se hacen resistentes a la ampicilina. La síntesis de la endonucleasa omega se logra por medio de la adición de I.P.T.G. o un equivalente del sistema operón lactosa.

Un plásmido identificado como pSCM525 que contiene la enzima I-Sce I fue depositado en la cepa TG1 de *E. coli* con la Collection Nationale de Cultures de Microorganismes (C. N. C. M.) del Instituto Pasteur de Paris, Francia el 22 de noviembre de 1990, bajo el número de acceso del depósito de la colección de cultivos No. I-1014. La secuencia de nucleótidos que codifica a la enzima I-Sce I está por lo tanto disponible en este depósito.

Se puede preparar también el gen que codifica a la enzima I-Sce I por medio de la formación de enlaces fosfato 3' → 5' entre unidades de nucleósidos utilizando técnicas convencionales de síntesis química. Por ejemplo, se pueden emplear las técnicas conocidas del fosfodiéster, fosfotriéster y fosfito triéster, así como modificaciones conocidas de estas aproximaciones. Los desoxirribonucleótidos se pueden preparar con máquinas automáticas de síntesis, tal como aquellas basadas en la aproximación de fosforoamidita. Se pueden obtener también oligo y polirribonucleótidos con la ayuda de ARN ligasa utilizando técnicas convencionales.

Esta solicitud también divulga variantes de la secuencia de ADN como la definida aquí que exhibe sustancialmente las mismas propiedades que la secuencia definida en este documento. Por esto se entiende que no se requiera que las secuencias de ADN sean idénticas a la secuencia divulgada aquí. Las variaciones se pueden atribuir a sustituciones, supresiones, o inserciones o mutaciones locales de bases individuales o múltiples que involucran uno o más nucleótidos que no se apartan sustancialmente de las propiedades de la secuencia de ADN, en la medida que codifiquen una enzima que tenga las propiedades de escisión de la enzima I-SceI.

La Fig. 5 describe algunas de las variaciones que se pueden elaborar alrededor de la secuencia de aminoácidos de I-SceI. Se ha demostrado que se pueden cambiar las siguientes posiciones sin afectar la actividad de la enzima:

posiciones -1 y -2 no son naturales. Los dos aminoácidos se añaden debido a las estrategias de clonación.

posiciones 1 a 10: pueden ser suprimidas.

posición 36: G es tolerado.

posición 40: M o V son tolerados.

posición 41: S o N son tolerados.

posición 43: A es tolerado.

posición 46: V o N son tolerados.

posición 91: A es tolerado.

posiciones 123 y 156: L es tolerado.

posición 223: A y S son tolerados.

Se entiende que las enzimas que contienen estas modificaciones están dentro del alcance de la presente solicitud.

Los cambios a la secuencia de aminoácidos en la Fig. 5 que se ha demostrado que afectan la actividad de la enzima son los siguientes:

posición 19: L hasta S

posición 38: I hasta S o N

posición 39: G hasta D o R

posición 40: L hasta Q

posición 42: L hasta R

posición 44: D hasta E, G o H

posición 45: A hasta E o D

posición 46: Y hasta D

posición 47: I hasta R o N

posición 80: L hasta S

posición 144: D hasta E

posición 145: D hasta E

posición 146: G hasta E

posición 147: G hasta S

Se entenderá también que la presente solicitud está destinada a abarcar fragmentos de la secuencia de ADN divulgados aquí en forma purificada, donde los fragmentos son capaces de codificar a la I-Sce I enzimáticamente activa.

La secuencia de ADN que codifica para la enzima I-Sce I se puede amplificar en la bien conocida reacción en cadena de la polimerasa (PCR), que es útil para la amplificación de todas las regiones o de regiones específicas del gen. Ver, por ejemplo, S. Kwok y colaboradores, J. Virol., 61: 1690-1694 (1987); la patente estadounidense No. 4.683.202; y la patente estadounidense No. 4.683.195. Más particularmente, los pares iniciadores de ADN de una secuencia conocida posicionada 10-300 pares de bases más allá que son complementarios a las hebras más y menos del ADN que va a ser amplificado se pueden preparar por medio de técnicas bien conocidas para la síntesis de oligonucleótidos. Se puede extender y modificar un extremo de cada iniciador para crear sitios de restricción de la endonucleasa cuando se aparea el iniciador al ADN. La mezcla de reacción PCR puede contener el ADN, los pares iniciadores de ADN, cuatro desoxirribonucleósidos trifosfato, MgCl₂, ADN polimerasa, y amortiguadores convencionales. Se puede amplificar el ADN durante una cierta cantidad de ciclos. Generalmente es posible incrementar la sensibilidad de detección utilizando una multiplicidad de ciclos, consistiendo cada ciclo de un corto período de desnaturalización del ADN a una temperatura elevada, enfriando la mezcla de reacción, y polimerización con la ADN polimerasa. Las secuencias amplificadas se pueden detectar por medio del uso de una técnica llamada restricción oligomérica (OR). Ver, R. K. Saiki y colaboradores, Bio/Technology 3: 1008-1012 (1985).

La enzima I-Sce I es una de un número de endonucleasas con propiedades similares. A continuación se da un listado de enzimas relacionadas y sus fuentes.

Las endonucleasas codificadas por un intrón del Grupo I y las enzimas relacionadas se enlistan a continuación con referencias. Los sitios de reconocimiento se muestran en la Fig. 6.

Enzima	Codificada por	Ref.
I-SceI	Intrón Sc LSU-1	Este trabajo
I-SceII	Intrón Sc cox1-4	Sargueil y colaboradores, NAR (1990) 18, 5659 - 5665
I-SceIII	Intrón Sc cox1-3	Sargueil y colaboradores, MGG (1991) 225, 340 - 341
I-SceIV	Intrón Sc cox1-5a	Seraphin y colaboradores (1992) en prensa
I-CeuI	Intrón Ce LSU-5	Marshall, Lemieux Gene (1991) 104, 241 - 245
I-CreI	Intrón Cr LSU-1	Rochaix (no publicado)
I-PpoI	Intrón Pp LSU-3	Muscarella y colaboradores, MCB (1990) 10, 3386 -3396
I-TevI	Intrón T4 td-1	Chu y colaboradores, PNAS (1990) 87, 3574 - 3578 y Bell-Pedersen y colaboradores NAR (1990) 18, 3763 - 3770
I-TevII	Intrón T4 sunY	Bell-Pedersen y colaboradores NAR (1990) 18, 3763-3770
I-TevIII	Intrón RB3 nrdB-1	Eddy, Gold, Genes Dev. (1991) 5, 1032 - 1041
HO	Gen de levadura HO	Nickoloff y colaboradores, MCB (1990) 10, 1174 - 1179
SceI Endo	Mitógeno de levadura RF3	Kawasaki y colaboradores, JBC (1991) 266, 5342 - 5347

ES 2 338 952 T3

Las nuevas enzimas putativas (evidencia genética, pero no hay ninguna actividad todavía) son I-*CsmI* del intrón 1 del citocromo b de mitocondria de *Clamidomonas smithii* (ref. 15), I-*PantI* del intrón 3 del citocromo b de mitocondria de *Podospora anserina* (Jill Salvo), y probablemente enzimas codificadas por intrones Nc nd1*1 y Nc cob*! de *Neurospora crassa*.

Las I-endonucleasas se pueden clasificar de la siguiente manera:

Clase I: Dos motivos dodecapéptido, un corte escalonado de 4 pb con proyecciones OH 3', un corte interno al sitio de reconocimiento

Subclase "I- <i>SceI</i> "	Otras subclases
I- <i>SceI</i>	I- <i>SceII</i>
I- <i>SceIV</i>	I- <i>SceIII</i>
I- <i>CsmI</i>	I- <i>CeuI</i> (únicamente un motivo dodecapéptido)
I- <i>PanI</i>	I- <i>CreI</i> (únicamente un motivo dodecapéptido)
	HO
	TFP1-408 (homólogo de HO)
	<i>SceI</i> Endo

Clase II: motivo GIY-(N₁₀₋₁₁) YIG, corte escalonado de 2 pb con proyecciones OH 3', corte externo al sitio de reconocimiento:

I-*TevI*

Clase III: motivos estructurales no típicos, corte escalonado de 4 pb con proyecciones OH 3', corte interno al sitio de reconocimiento:

I-*PpoI*

Clase IV: motivos estructurales no típicos, corte escalonado de 2 pb con proyecciones OH 3', corte externo al sitio de reconocimiento:

I-*TevII*

Clase V: motivos estructurales no típicos, corte escalonado de 2 pb con proyecciones OH 3':

I-*TevIII*.

2. Sondas de Nucleótido que Contienen 1 gen para I-*SceI*

La secuencia de ADN codifica para la enzima I-*SceI* y divulgada en este documento puede ser utilizada también como sonda para la detección de una secuencia de nucleótido en un material biológico, tal como un tejido o fluidos biológicos. La sonda puede ser marcada con un átomo o radical inorgánico, más comúnmente utilizando un radionúclido, sino también quizás con un metal pesado. Los marcadores radioactivos incluyen ³²P, ³H, ¹⁴C, o similares. Se puede emplear cualquier marcador radioactivo, que establece una señal adecuada y tiene una vida media suficiente. Otros marcadores incluyen ligandos que pueden servir como un miembro específico de enlazamiento con un anticuerpo marcado, fluorescentes, quimioluminescentes, enzimas, anticuerpos que pueden servir como un miembro par de enlazamiento específico para un ligando marcado, y similares. La escogencia del marcador será gobernada por el efecto del marcador sobre la velocidad de hibridación y el enlazamiento de la sonda con el ADN o ARN. Será necesario que el marcador proporcione suficiente sensibilidad para detectar la cantidad de ADN o ARN disponible para hibridación.

Cuando la secuencia de nucleótidos divulgada en la solicitud es utilizada como una sonda para hibridación con un gen, se fija preferiblemente la secuencia de nucleótido a un sólido insoluble en agua, soporte poroso, tal como papel de nitrocelulosa. La hibridación se puede llevar a cabo utilizando polinucleótidos marcados como se define aquí y reactivos de hibridación convencionales. La técnica particular de hibridación no es esencial.

La cantidad de sonda marcada presente en la solución de hibridación variará ampliamente, dependiendo de la naturaleza del marcador, la cantidad de sonda marcadora que puede enlazarse razonablemente con el soporte, y la

rigurosidad de la hibridación. Generalmente, se empleará un exceso sustancial de la sonda sobre estequiométrica para mejorar la velocidad de enlazamiento de la sonda con el ADN fijo.

Se pueden emplear diferentes grados de rigurosidad de hibridación. Entre más severas las condiciones, mayor la complementariedad que se requiere para la hibridación entre la sonda y el polinucleótido para la formación de dúplex. La severidad se puede controlar por medio de la temperatura, la concentración de la sonda, la longitud de la sonda, la fuerza iónica, y similares. Convenientemente, la rigurosidad de la hibridación se varía cambiando la polaridad de la solución reactiva. Las temperaturas que van a ser empleadas se pueden determinar empíricamente o se pueden determinar a partir de formulas bien conocidas desarrolladas para este propósito.

3. *Secuencias de Nucleótidos que Contienen la Secuencia de Nucleótidos que Codifica I-SceI*

Esta solicitud también se relaciona con la secuencia de ADN como la definida aquí que codifica la enzima I-SceI, en donde la secuencia de nucleótidos está enlazada a otros ácidos nucleicos. Se puede obtener el ácido nucleico a partir de cualquier fuente, por ejemplo, de plásmidos, de ADN o ARN clonado, o de ADN o ARN natural de cualquier fuente, incluidos organismos procariotas y eucariotas. Se puede extraer ADN o ARN de un material biológico, tal como fluidos biológicos o tejido, por medio de una variedad de técnicas incluidas aquellas descritas por Maniatis y colaboradores, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York (1982). Generalmente se obtendrá el ácido nucleico a partir de una bacteria, levadura, virus, o un organismo superior, tal como una planta o animal. El ácido nucleico puede ser una fracción de una mezcla más compleja, tal como una porción de un gen contenido en ADN humano completo o una porción de una secuencia de ácido nucleico de un microorganismo particular. El ácido nucleico puede ser una fracción de una molécula mayor o el ácido nucleico puede constituir un gen entero o reunión de genes. El ADN puede estar en forma monocatenaria o bicatenaria. Si el fragmento está en forma monocatenaria, puede ser convertido en forma bicatenaria utilizando ADN polimerasa de acuerdo con técnicas convencionales.

La secuencia de ADN como se define aquí puede estar enlazada a un gen estructural. Como se lo utiliza aquí, el término "gen estructural" se refiere a una secuencia de ADN que codifica a través de su molde o ARNm mensajero una secuencia de aminoácidos característica de una proteína específica o polipéptido. La secuencia de nucleótidos como la definida aquí puede funcionar con una secuencia de control de la expresión, esto es, una secuencia de ADN que controla y regula la expresión del gen cuando está operativamente enlazado al gen.

4. *Vectores que Contienen la Secuencia de Nucleótidos divulgada en la solicitud*

Esta solicitud también se relaciona con vectores para clonación y expresión que contienen la secuencia de ADN como la divulgada en la solicitud que codifica para la enzima I-SceI.

Más particularmente, la secuencia de ADN que codifica la enzima puede estar ligada a un vehículo para clonación de la secuencia. Las etapas principales involucradas en clonación génica incluyen procedimientos para separación de ADN que contiene al gen de interés de procariotas o eucariotas, el corte del fragmento resultante de ADN y el ADN de un vehículo de clonación en sitios específicos, mezclando los dos fragmentos de ADN juntos, y ligando los fragmentos para producir una molécula de ADN recombinante. La molécula recombinante puede ser luego transferida dentro de una célula huésped, y se les permite a las células replicarse para producir células idénticas que contienen clones de la secuencia original de ADN.

El vehículo empleado puede ser cualquier molécula de ADN bicatenario capaz de transportar la secuencia de nucleótidos como se divulga en la solicitud dentro de una célula huésped y capaz de replicarse dentro de la célula. Más particularmente, el vehículo debe contener al menos una secuencia de ADN que puede actuar como el origen de replicación en la célula huésped. Además, el vehículo puede contener dos o más sitios para inserción de la secuencia de ADN que codifica al gen como se define aquí. Estos sitios corresponderán ordinariamente a sitios de enzima de restricción en los cuales se puede formar extremos cohesivos, y que son complementarios con los extremos cohesivos sobre la secuencia promotora que va a ser ligada al vehículo. En general, esto se puede llevar a cabo con un vehículo plásmido, bacteriófago, o cósmido que tenga estas características.

La secuencia de nucleótidos como la divulgada en la solicitud puede tener extremos cohesivos compatibles con cualquier combinación de sitios en el vehículo. Alternativamente, la secuencia puede tener uno o más extremos romos que pueden estar ligados a los extremos romos correspondientes en los sitios de clonación del vehículo. La secuencia de nucleótidos que se va a ligar puede ser procesada adicionalmente, si se desea, por medio de una supresión sucesiva con exonucleasa, tal como con la enzima Bal 31. En el evento de que la secuencia de nucleótidos como la descrita en la solicitud no contenga una combinación deseada de extremos cohesivos, se puede modificar la secuencia por medio de la adición de un enlazador, un adaptador, u homopolímero formador de colas.

Se prefiere que los plásmidos utilizados para la clonación de secuencias de nucleótidos como se describe en la solicitud transporte uno o más genes responsables de una característica útil, tal como un marcador seleccionable, exhibida por la célula huésped. En una estrategia preferida, se escogen los plásmidos que tienen genes para resistencia a dos medicamentos diferentes. Por ejemplo, inserción de la secuencia de ADN en un gen para que un antibiótico inactivo al gen y destruya la resistencia al medicamento. El segundo gen para resistencia a un medicamento no es

afectado cuando las células se transforman con los recombinantes, y las colonias que contienen al gen de interés se pueden seleccionar por la resistencia al segundo medicamento y la susceptibilidad al primer medicamento. Los marcadores antibióticos preferidos son genes que imparten resistencia al cloranfenicol, ampicilina, o tetraciclina a la célula huésped.

Se pueden utilizar una variedad de enzimas de restricción para cortar al vehículo. La identidad de la enzima de restricción generalmente dependerá de la identidad de los extremos sobre la secuencia de ADN que van a ser ligados y los sitios de restricción en el vehículo. La enzima de restricción es emparejada con los sitios de restricción en el vehículo, que a su vez es emparejado con los extremos sobre el fragmento de ácido nucleico que está siendo ligado.

La reacción de ligación puede ser establecida utilizando técnicas bien conocidas y reactivos convencionales. La ligación se lleva a cabo con una ADN ligasa que cataliza la formación de enlaces fosfodiéster entre fosfato 5' adyacente y los grupos hidroxilo 3' libres en dobles de ADN. La ADN ligasa se puede derivar de una variedad de microorganismos. Las ADN ligasas preferidas son enzimas de *E. coli* y del bacteriófago T4. La ADN ligasa de T4 puede ligar fragmentos de ADN con extremos romos o pegajosos, tal como aquellos generados por digestión con enzima de restricción. Se puede utilizar ADN ligasa de *E. coli* para catalizar la formación de enlaces fosfodiéster entre los terminales de las moléculas dobles de ADN que contienen extremos cohesivos.

Se puede llevar a cabo la clonación en células procariotas o eucariotas. El huésped para replicación del vehículo de clonación será desde luego uno que sea compatible con el vehículo y en el cual pueda replicarse el vehículo. Cuando se emplea un plásmido, éste se puede derivar de bacterias o de algunos otros organismos o el plásmido se puede preparar en forma sintética. Se puede emplear el plásmido que se pueda replicar independientemente del cromosoma de la célula huésped o de un plásmido integrador (episoma). El plásmido puede hacer uso de las enzimas replicadoras de ADN de la célula huésped con el propósito de replicarse o el plásmido puede transportar genes que codifiquen para las enzimas requeridas para replicación del plásmido. Se pueden emplear una cantidad de plásmidos diferentes para llevar a cabo lo anterior.

La secuencia de ADN como se define en este documento que codifica la enzima I-Sce I puede ser ligada también a un vehículo para formar un vector de expresión. El vehículo empleado en este caso es uno en el cual es posible expresar el gen operativamente enlazado a un promotor en una célula huésped apropiada. Es preferible emplear un vehículo conocido para uso en genes que se expresan en *E. coli*, levadura, o células de mamífero. Estos vehículos incluyen, por ejemplo, los siguientes vectores de expresión de *E. coli*:

pSCM525, que es un vector de expresión de *E. coli* derivado de pUC12 por inserción de un promotor tac y del gen sintético para I-SceI. La expresión es inducida por IPTG.

pGEX ω 6, que es un vector de expresión de *E. coli* derivado de pGEX en el cual el gen sintético de pSCM525 para I-Sce I se fusiona con el gen de la glutatión S transferasa, produciendo una proteína híbrida. La proteína híbrida posee la actividad de endonucleasa.

pDIC73, que es un vector de expresión de *E. coli* derivado de pET-3C por inserción del gen sintético para I-Sce I (fragmento *NdeI-BamHI* de pSCM525) bajo el control del promotor T7. Este vector es utilizado en la cepa BL21 (DE3) que expresa la ARN polimerasa de T7 bajo inducción de IPTG.

pSCM351, que es un vector de expresión de *E. coli* derivado de pUR291 en el cual el gen sintético para I-Sce I se fusiona con el gen Lac Z, produciendo una proteína híbrida.

pSCM353, que es un vector de expresión de *E. coli* derivado de pEX1 en el cual el gen sintético para I-Sce I se fusiona con el gen Cro/Lac Z, produciendo una proteína híbrida. Los ejemplos de vectores de expresión de levadura son:

pPEX7, que es un vector de expresión de levadura derivado de pRP51Bam O (un derivado LEU2d de pLG-SD5) por inserción del gen sintético bajo el control del promotor de galactosa. La expresión es inducida por galactosa.

pPEX408, que es un vector de expresión de levadura derivado de pLG-SO5 por inserción del gen sintético bajo el control del promotor de galactosa. La expresión es inducida por galactosa.

Varios vectores de expresión de levadura están descritos en la Fig. 7. Los vectores de expresión típicos de mamífero son:

pRSV I-SceI, que es un derivado de pRSV en el cual el gen sintético (fragmento *BamHI-PstI* de pSCM525) está bajo el control del promotor de LTR del Virus del sarcoma de Rous. Este vector de expresión está descrito en la Fig. 8.

También se pueden emplear vectores para expresión en Ovario de Hámster Chino (CHO).

5. Células Transformadas con Vectores

Los vectores como los divulgados en la solicitud se pueden insertar en organismos huésped utilizando técnicas convencionales. Por ejemplo, se pueden insertar los vectores por medio de transformación, transfección, electroporación, microinyección, o por medio de liposomas (lipofección).

Se puede llevar a cabo la clonación en células procariota o eucariotas. El huésped para replicación del vehículo de clonación será desde luego uno que sea compatible con el vehículo y en el cual pueda replicarse el vehículo. La clonación se lleva a cabo preferiblemente en células bacterianas o de levadura, aunque se pueden emplear también células de origen fúngico, animal y de planta. Las células huésped preferidas para llevar a cabo el trabajo de clonación son células bacterianas, tales como *E. coli*. El uso de células de *E. coli* es particularmente preferido debido a que la mayor parte de los vehículos de clonación, tal como plásmidos bacterianos y bacteriófagos, se replican en estas células.

En una modalidad preferida, se inserta un vector de expresión que contiene la secuencia de ADN que codifica la secuencia de nucleótidos como la definida aquí operativamente enlazada a un promotor dentro de una célula de mamífero utilizando técnicas convencionales.

La solicitud también se relaciona con un método para inducir al menos un rompimiento bicatenario dirigido al sitio en ADN de una célula, comprendiendo dicho método:

- (a) el suministro de células que contienen ADN bicatenario, en donde dicho ADN comprende al menos un sitio de restricción de I-SceI;
- (b) la transfección de dichas células con al menos un plásmido que comprende ADN que codifica la I-Sce I meganucleasa; y
- (c) seleccionar células en las cuales se ha inducido al menos un rompimiento bicatenario.

En una modalidad preferida de este método, se selecciona la célula del grupo que consiste de una célula de mamífero, una célula de levadura, y una célula de planta. En otra modalidad del método, la célula es una célula NIH3T3 que contiene al virus G-MtkPL. En particular, el plásmido es pCMV(I-SceI+). También se divulga dentro de la solicitud una célula eucariota recombinante transformada por medio de este método.

La solicitud también describe un método para inducir una recombinación homóloga entre ADN cromosómico de una célula y ADN exógeno añadido a dicha célula, comprendiendo dicho método:

- (a) suministrar células que contengan ADN cromosómico, en donde dicho ADN comprende al menos un sitio de restricción de I-SceI;
- (b) la transfección de dichas células con un plásmido que contiene ADN exógeno, y con un plásmido que contiene ADN que codifica la I-Sce I meganucleasa; y
- (c) seleccionar células en las cuales dicho ADN exógeno es insertado dentro de dicho ADN cromosómico.

En una modalidad preferida de este método, se selecciona la célula del grupo que consiste de una célula de mamífero, una célula de levadura, y una célula de planta. En otra modalidad del método, la célula es una célula NIH3T3 que contiene al virus G-MtkPL. En particular, el plásmido es pCMV(I-Sce I+).

La solicitud también se relaciona con un método para inducir una recombinación homóloga entre ADN cromosómico de una célula y ADN exógeno añadido a dicha célula, comprendiendo dicho método:

- (a) suministrar células que contengan ADN cromosómico;
- (b) insertar al menos un sitio de restricción de I-Sce I en dicho ADN cromosómico;
- (c) la transfección de dichas células con un primer plásmido que contiene ADN exógeno, y con un segundo plásmido que contiene ADN que codifica la I-Sce I meganucleasa; y
- (d) seleccionar células en las cuales dicho ADN exógeno es insertado dentro de dicho ADN cromosómico.

En una modalidad preferida de este método, se selecciona la célula del grupo que consiste de una célula de mamífero, una célula de levadura, y una célula de planta. En otra modalidad del método, el primer plásmido de pVRneo. En otra modalidad, el segundo plásmido es pCMV(I-Sce I+).

La solicitud también divulga un método para inducir al menos un rompimiento dirigido al sitio en ADN de una célula e insertar ADN que codifica un polipéptido, comprendiendo dicho método:

- (a) suministrar células que contengan ADN bicatenario, en donde dichas células puedan ser transformadas por medio de un ADN que contenga un sitio de restricción de I-Sce I y ADN que codifique dicho polipéptido;
- (b) añadir enzima I-Sce I o transformar dicha célula con ADN que codifique a la enzima I-Sce I;
- (c) la transfección de dichas células con dicho ADN que codifica dicho polipéptido o con un vector que contenga dicho ADN; y
- (d) seleccionar células transfectadas con dicho ADN o dicho vector, en donde dichas células expresen dicho polipéptido.

También se divulgan dentro de la solicitud una célula eucariota recombinante transformada por este método así como un animal transgénico que contiene una célula transformada por este método. En una modalidad particular, el polipéptido es un antígeno extraño a la célula. En una modalidad particular, la célula eucariota recombinante es una línea celular de mamífero o es levadura.

La solicitud también se relaciona con un método para expresar un polipéptido en un animal transgénico, dicho método comprendiendo la transformación de células madre embrionarias con un ADN que contiene un sitio de restricción de I-Sce I y ADN que codifica dicho polipéptido; y detectar la expresión de dicho polipéptido en un animal transgénico que resulta de dichas células madre embrionarias transformadas.

La solicitud también se relaciona con una célula madre recombinante que expresa un polipéptido, en donde dicha célula madre es transformada por un ADN que contiene un sitio de restricción de I-Sce I y ADN que codifica dicho polipéptido por medio de:

- (a) la adición de enzima I-Sce I a dicha célula o transformando dicha célula con un vector que contiene al gen que codifica para la enzima I-Sce I;
- (b) la transfección de dichas células con dicho ADN que codifica dicho polipéptido; y
- (c) la selección de células transfectadas con dicho ADN, en donde dichas células expresan dicho polipéptido. Finalmente, la solicitud también describe un método para inducir al menos un rompimiento dirigido al sitio en ADN de células e insertar ADN que codifique un polipéptido, en donde dichas células expresan al menos un producto de proteína, comprendiendo dicho método:
 - (a) suministrar células que contengan ADN bicatenario, en donde dichas células puedan ser transformadas por medio de un ADN que contenga un sitio de restricción de I-Sce I y ADN que codifique dicho polipéptido;
 - (b) añadir enzima I-Sce I a dichas células o transformar dichas células con ADN que codifique a la enzima I-Sce I;
 - (c) la transfección de dichas células con dicho ADN que codifica dicho polipéptido o con un vector que contenga dicho ADN; y
 - (d) seleccionar células transfectadas con dicho ADN o dicho vector, en donde dichas células expresan dicho polipéptido y no expresan dicho producto proteico.

La solicitud también describe una célula recombinante transformada por medio de dicho método.

Aplicación de I-Sce I para mapeo a gran escala

1. Ocurrencia de sitios naturales en diferentes genomas

Utilizando la enzima I-Sce I purificada, se ha examinado la ocurrencia de sitios naturales o degenerados sobre los genomas completos de diferentes especies. No se encontró un sitio natural en *Saccharomyces cerevisiae*, *Bacillus anthracis*, *Borrelia burgdorferi*, *Leptospira biflexa* ni en *L. interrogans*. Se encontró un sitio degenerado sobre el ADN del fago T7.

2. Inserción de sitios artificiales

Dada la ausencia de sitios naturales de I-SceI, se pueden introducir sitios artificiales por medio de transformación o de transfección. Se deben distinguir dos casos: integración dirigida al sitio por medio de recombinación homóloga e

integración aleatoria por medio de recombinación no homóloga, movimiento de transposón o infección retroviral. Lo primero es fácil en el caso de levadura y de unas pocas especies bacterianas, más difícil para eucariotas superiores. Lo segundo es posible en todos los sistemas.

3. Vectores de inserción

Se pueden distinguir dos tipos:

- 1 - casetes específicos del sitio que introducen el sitio de I-Sce I junto con un marcador seleccionable.

Para levadura: todos son derivados de pAF100 (Thierry y colaboradores. (1990) YEAST 6. 521-534) que contienen los siguientes genes marcadores:

pAF101: URA3 (insertado en el sitio de *HindIII*)
 pAF103: NeoR (insertado en el sitio de *BglII*)
 pAF104: HIS3 (insertado en el sitio de *BglII*)
 pAF105: KanR (insertado en el sitio de *BglII*)
 pAF106: KanR (insertado en el sitio de *BglII*)
 pAF107: LYS2 (insertado entre *HindIII* y *EcoR V*)

En la Fig. 9 se muestra un mapa de restricción del plásmido pAF100. La secuencia de nucleótidos y los sitios de restricción de regiones de plásmidos pAF100 se muestran en las Figs. 10A y 10B. Se encuentran disponibles muchas cepas de levadura transgénica con el sitio de I-Sce I en lugares diferentes y conocidos a lo largo de los cromosomas.

- 2 - Vectores derivados de elementos transposables o de retrovirus.

Para *E. coli* y otras bacterias: derivados de Tn5 mini que contienen el sitio de I-Sce I y

pTsm ω Str^R
 pTKm ω Kan^R (Ver Fig. 11)
 pTTc ω Tet^R

Para levadura: pTy ω 6 es un derivado de pD123 en el cual se ha insertado el sitio de I-Sce I en la LTR del elemento Ty (Fig. 12).

pMLV LTR SAPLZ: que contiene el sitio de I-Sce I en la LTR de MLV y Phleo-LacZ (Fig. 13). Se cultiva primero este vector en células Ψ 2 (derivadas de 3T3, de R. Mulligan). Se encuentran disponibles dos líneas de células transgénicas con el sitio de I-Sce I en ubicaciones no determinadas en el genoma: 1009 (células nerviosas pluripotentes, J. F. Nicolas) y D3 (células ES capaces de generar animales transgénicos).

4. Estrategia de fragmentación cromosómica anidada

La estrategia de fragmentación cromosómica anidada para mapear genéticamente un genoma de eucariota explota las propiedades únicas de la endonucleasa de restricción I-SceI, tal como un sitio de reconocimiento largo de 18 pb. La ausencia de sitios de reconocimiento naturales de I-Sce I en la mayoría de los genomas de eucariotas es explotada también en esta estrategia de mapeo.

Primero, se insertan artificialmente uno o más sitios de reconocimiento de I-Sce I en diferentes posiciones en un genoma, por medio de recombinación homóloga utilizando casetes específicos que contienen marcadores seleccionables o por medio de inserción aleatoria, como se discutió más arriba. Se escinde luego completamente el genoma de la cepa transgénica resultante en el(los) sitio(s) de I-Sce I artificialmente insertado(s) después de incubación con la enzima de restricción I-SceI. La escisión produce fragmentos cromosómicos anidados.

Los fragmentos cromosómicos se purifican luego y se separan por medio de electroforesis en gel de campo pulsado (PFG), permitiéndole a uno "mapear" la posición del sitio insertado en el cromosoma. Si se escinde al ADN total con la enzima de restricción, cada sitio introducido artificialmente de I-Sce I proporciona un "hito molecular" único en el genoma. De este modo, se puede crear un conjunto de cepas transgénicas, cada una transportando un sitio único de I-SceI, que define intervalos genómicos físicos entre los hitos. En consecuencia, se puede mapear un genoma completo, un cromosoma o cualquier segmento de interés utilizando sitios de restricción artificialmente introducidos de I-SceI.

Los fragmentos cromosómicos anidados se pueden transferir a una membrana sólida e hibridarse hasta una sonda marcada que contiene ADN complementario al ADN de los fragmentos. Con base en los patrones de las bandas de hibridación que se observan, se puede mapear el genoma eucariota. Se utiliza el conjunto de cepas transgénicas con “hitos” apropiados como referencia para mapear cualquier gen nuevo o clon por medio de hibridación directa.

Ejemplo 1

Aplicación de la Estrategia de Fragmentación Cromosómica Anidada al Mapeo del Cromosoma XI de Levadura

Esta estrategia ha sido aplicada al mapeo del cromosoma XI de levadura de *Saccharomyces cerevisiae*. Se insertó el sitio de I-Sce I en 7 diferentes posiciones a lo largo del cromosoma XI de la cepa diploide FY1679, definiendo por lo tanto ocho intervalos físicos en ese cromosoma. Se insertaron los sitios de un casete URA3-1-I-SceI por medio de recombinación homóloga. Se insertaron dos sitios dentro de genes genéticamente definidos, TIF1 y FAS1, los otros fueron insertados en posiciones desconocidas en el cromosoma de cinco cósmidos que no se superponen de nuestra biblioteca, tomados al azar. El ADN embebido de agarosa de cada una de las siete cepas transgénicas fue luego digerido con I-Sce I y analizado por medio de electroforesis en gel de campo pulsado (Fig. 14A). Se deduce primero la posición del sitio de I-Sce I de cada cepa transgénica en el cromosoma XI a partir de los tamaños de los fragmentos sin consideración de la orientación izquierda/derecha de los fragmentos. Se determinó la orientación de la siguiente manera. La mayoría de los telómeros próximos al sitio de I-Sce I de este conjunto de cepas están en la E40 transgénica debido a que el fragmento de 50 kb es el más corto de todos los fragmentos (Fig. 15A). Por lo tanto, el clon cósmido pUKG040, que fue utilizado para insertar el sitio de I-Sce I en la E40 transgénica, es utilizado ahora como sonda contra todos los fragmentos de cromosoma (Fig. 14B). Como se esperaba, pUKG040 activa los dos fragmentos de la cepa E40 (50 kb y 630 kb, respectivamente). El fragmento grande está cerca al cromosoma XI entero y muestra una señal de hibridación débil debido al hecho de que el inserto de pUKG040, que tiene 38 kb de largo, contiene menos de 4 kb dentro del fragmento largo del cromosoma. Obsérvese que el cromosoma XI completo permanece visible después de digestión con I-SceI, debido al hecho de que las cepas transgénicas son diploides en los cuales se inserta el sitio de I-Sce I únicamente en uno de los dos homólogos. Ahora, la sonda pUKG040 hibrida únicamente hasta un fragmento de todas las otras cepas transgénicas permitiendo una orientación izquierda/derecha no ambigua de sitios de I-Sce I (Ver la Fig. 15B). No es visible una hibridación cruzada significativa entre el vector cósmido y el subfragmento de cromosoma que contiene la inserción del sitio de I-SceI. Se pueden ordenar ahora las cepas transgénicas de tal manera que los sitios de I-Sce I se localicen a distancias cada vez mayores del extremo hibridante del cromosoma (Fig. 15C) y se puede deducir el mapa de I-Sce I (Fig. 15D). La precisión del mapeo depende de la resolución de PFGE y de una calibración óptima. Obsérvese que la orientación izquierda/derecha real del cromosoma con respecto al mapa genético no se conoce en esta etapa. Para ayudar a visualizar nuestra estrategia y para obtener mediciones más precisas de los tamaños del intervalo entre los sitios de I-SceI, se hizo una nueva electroforesis en gel de campo pulsado con las mismas cepas transgénicas colocadas ahora en orden (Fig. 16). Después de la transferencia, se hibridaron exitosamente los fragmentos con los cósmidos pUKG040 y pUKG066 que activan, respectivamente, todos los fragmentos de los extremos opuestos del cromosoma (el clon pUKG066 define el extremo derecho de cromosoma como se define a partir del mapa genético debido a que contiene el gen SIR1. Se observa un progreso regular paso a paso de los tamaños de los fragmentos del cromosoma. Se observa alguna hibridación cruzada entre la sonda pUKG066 y el cromosoma III, probablemente debido a algunas secuencias repetitivas de ADN.

Todos los fragmentos del cromosoma, tomados en conjunto, definen ahora los intervalos físicos como se indica en la Fig. 15d. El mapa obtenido de I-Sce I tiene una resolución promedio de 80 kb.

Ejemplo 2

Aplicación de la Estrategia de Fragmentación Cromosómica Anidada para el Mapeo de los Clones del Cromosoma Artificial de Levadura (YAC)

Se puede aplicar esta estrategia al mapeo de YAC con dos posibilidades.

- 1 - inserción del sitio de I-Sce I dentro del gen de interés utilizando recombinación homóloga en levadura. Esto permite el mapeo de ese gen n el inserto de YAC por medio de digestión con I-Sce I *in vitro*. Se ha hecho esto y funciona.
- 2 - integración al azar de sitios de I-Sce I a lo largo del inserto de YAC por medio de recombinación homóloga en levadura utilizando secuencias altamente repetitivas (por ejemplo, B2 en ratón o Alu en humano). Se utilizan luego las cepas transgénicas como se describe en la ref. P1 para clasificar bibliotecas o genes del mapa.

Se ha extendido ahora el procedimiento a YAC que contiene 450 kb de ADN de Ratón. Con este propósito, se ha insertado una secuencia repetida de ADN de ratón (llamada B2) en un plásmido que contiene el sitio de I-Sce I y un marcador seleccionable de levadura (LYS2). La transformación de las células de levadura que contienen al YAC recombinante con 1 plásmido linealizado dentro de la secuencia B2 resultó en la integración del sitio de I-

Sce I en cinco diferentes ubicaciones distribuidas a lo largo del inserto del ADN de ratón. La escisión en los sitios insertados de I-Sce I utilizando la enzima ha sido exitosa, produciendo fragmentos anidados que pueden ser purificados después de electroforesis. Las etapas posteriores del protocolo comparan exactamente el procedimiento descrito en el Ejemplo 1.

Ejemplo 3

Aplicación de Fragmentos Cromosómicos Anidados a la Clasificación Directa de Bibliotecas de Cósmido

Los fragmentos cromosómicos anidados se pueden purificar a partir de PFG preparativo y se pueden utilizar como sondas contra clones de una sub-biblioteca específica para el cromosoma XI. Esta sub-biblioteca está compuesta de 138 clones cósmidos (correspondiente a una cobertura de ocho veces) que han sido previamente clasificados a partir de nuestras completas bibliotecas genómicas de levadura por medio de hibridación de colonias con el cromosoma XI purificado por PFG. Esta colección de clones no ordenados ha sido secuencialmente hibridada con fragmentos de cromosoma tomados en orden creciente de tamaño del extremo izquierdo del cromosoma. La localización de cada clon cósmido sobre el mapa de I-Sce I se podría determinar sin ambigüedades a partir de tales hibridaciones. Para verificar adicionalmente los resultados y para proporcionar un mapa más preciso se ha digerido un subconjunto de todos los clones cósmidos, ahora colocados en orden, con EcoRI, sometidos a electroforesis e hibridados con las series anidadas de fragmentos de cromosoma en orden creciente de tamaños del extremo izquierdo del cromosoma. En la Figura 17 se dan los resultados.

Para una sonda dada, se pueden distinguir dos casos: clones cósmidos en los cuales todos los fragmentos de EcoRI hibridan con la sonda y los clones cósmidos en los cuales únicamente algunos de los fragmentos de EcoRI hibridan (es decir, compara pEKG100 con pEKG098 en la Fig. 17b). La primera categoría corresponde a clones en los cuales el inserto está completamente incluido en uno de los dos fragmentos de cromosoma, el segundo a clones en los cuales el inserto traslapa un sitio de I-SceI. Obsérvese que, para los clones de la serie pEKG, el fragmento de EcoRI de 8 kb está enteramente compuesto de secuencias de vector (pWE15) que no hibridan con los fragmentos del cromosoma. En el caso en el cual los fragmentos del cromosoma poseen el vector de integración, se observa una hibridación cruzada débil con el cósmido (Fig. 17e).

El examen de la Fig. 17 muestra que los clones cósmidos pueden ser ordenados sin ambigüedades con respecto al mapa de I-Sce I (Fig. 13E), cayendo cada clon ya sea en un intervalo definido o a través de un sitio de I-SceI. Además, los clones de la segunda categoría nos permiten colocar algunos fragmentos de EcoRI sobre los mapas de I-SceI, mientras que otros permanecen desordenados. El juego completo de clones cósmidos específicos del cromosoma XI, que cubren completamente ocho veces el equivalente de los cromosomas, han sido clasificados con respecto al mapa de I-SceI, como se muestra en la Fig. 18.

5. Mapeo de restricción parcial utilizando I-Sce I

En esta modalidad, la digestión completa del ADN en el sitio artificialmente insertado de I-Sce I es seguida por digestión parcial con endonucleasas bacterianas de restricción de escogencia. Los fragmentos de restricción son luego separados por medio de electroforesis y transferidos. La marcación indirecta del extremo se logra utilizando los sitios de la mitad izquierda o derecha de I-SceI. Esta técnica ha sido exitosa con cromosomas de levadura y se deben aplicar sin dificultad para YAC.

Se ha hecho mapeo de restricción parcial sobre ADN de levadura y sobre ADN de células de mamífero utilizando la enzima comercial I-SceI. El ADN de células que contienen un sitio insertado artificialmente de I-Sce I es escindido primero completamente por medio de I-SceI. Se trata luego el ADN bajo condiciones de escisión parcial con endonucleasas de restricción bacterianas de interés (por ejemplo, BamBI) y se lo somete a electroforesis junto con marcadores de calibración por tamaño. Se transfiere el ADN a una membrana y se hibrida sucesivamente utilizando las secuencias cortas que flanquean los sitios de I-Sce I sobre cualquier lado (estas secuencias son conocidas porque hacen parte del vector original de inserción que fue utilizado para introducir el sitio de I-SceI). Una autorradiografía (u otro sistema equivalente de detección utilizando sondas no radioactivas) permite la visualización de escalas, que representan directamente la sucesión de los sitios de endonucleasas de restricción bacterianas del sitio de I-SceI. El tamaño de cada banda de la escala es utilizado para calcular la distancia física entre los sitios sucesivos de endonucleasas de restricción bacterianas.

Aplicación de I-Sce I para Recombinación Dirigida al Sitio In Vivo

1. Expresión de I-Sce I en levadura

El gen sintético para I-Sce I ha sido colocado bajo el control de un promotor inducible de galactosa sobre los plásmidos multicopia pPEX7 y pPEX408. La expresión es correcta e induce efectos sobre el sitio como se indica más abajo. Se puede construir una levadura transgénica con el gen sintético para I-Sce I insertado en un cromosoma bajo el control de un promotor inducible.

2. Efectos de los rompimientos bicatenarios específicos del sitio en levadura (refs. 18 y P4)

Efectos sobre los sitios de I-Sce I a cargo del plásmido:

- 5 En la Ref. 18 se describen en detalle los efectos intramoleculares. La recombinación intermolecular (plásmido con cromosoma) se puede predecir.

Efectos sobre sitios integrados en el cromosoma de I-Sce I

- 10 En una célula haploide, un rompimiento sencillo dentro de un cromosoma en un sitio artificial de I-Sce I resulta en la detención de la división celular seguido por la muerte (únicamente un pequeño % de supervivientes). La presencia de una secuencia intacta homóloga al sitio de corte resulta en reparación y 100% de supervivencia celular. En una célula diploide, un rompimiento sencillo dentro de un cromosoma en un sitio artificial de I-Sce I resulta en reparación utilizando el cromosoma homólogo y 100% de supervivencia de las células. En ambos casos, la reparación del rompimiento bicatenario inducido resulta en pérdida de heterozigosidad con supresión de las secuencias no homólogas que flaquean el corte e inserción de las secuencias no homólogas de la molécula donante de ADN.

3. Aplicación para recombinación *in vivo* de los YAC en Levadura

- 20 La construcción de un vector de YAC con el sitio de restricción de I-Sce I próximo al sitio de clonación debe permitirle a uno inducir una recombinación homóloga con otro YAC si los insertos están parcialmente superpuestos. Esto es útil para la construcción de cóntigos.

4. Prospectos de otros organismos

- 25 La inserción de un sitio de restricción de I-Sce I ha sido hecha para bacterias (*E. coli*, *Yersinia enterocolitica*, *Y. pestis*, *Y. pseudotuberculosis*), y células de ratón. La escisión en el sitio artificial de I-Sce I *in vitro* ha sido exitosa con ADN de las células de ratón transgénico. La expresión de I-Sce I a partir del gen sintético en células de mamífero o de planta debe ser exitosa.

- 30 El sitio de I-Sce I ha sido introducido en células de ratón y en células de bacteria de la siguiente manera:

- 1 - Células de ratón:

- 35 - a - Las células de ratón (Ψ2) fueron transfectadas con el ADN del vector pMLV LTR SAPLZ que contiene el sitio de I-Sce I utilizando una técnica estándar de transfección de fosfato de calcio.
- b - Se seleccionaron las células transfectadas en medio DMEM que contenía fleomicina con suero fetal de ternera al 5% y se desarrollaron bajo atmósfera de CO₂ al 12%, 100% de humedad a 37°C hasta que formaron colonias.
- 40 - c - Se subclonaron las colonias resistentes a la fleomicina una vez en el mismo medio.
- d - Se escogió el clon MLOP014, que produjo un título de 10⁵ partículas de virus por ml. Este clon fue depositado en el C. N. C. M. el 5 de mayo de 1992 bajo el número de acceso de la colección de cultivos No. I-1207.
- 45 - e - Se utilizó el sobrenadante de este clon para infectar otras células de ratón (1009) esparciendo 10⁵ partículas virales sobre 10⁵ células en medio DMEM con suero fetal de ternera al 10% y 5 mg/ml de "polybrain". Se reemplazó el medio 6 horas después de la infección por el mismo medio fresco.
- 50 - f - 24 horas después de la infección, se seleccionaron células resistentes a la fleomicina en el mismo medio anterior.
- 55 - g - Se subclonaron colonias resistentes a la fleomicina una vez en el mismo medio.
- h - Se tomó un clon y se lo analizó. Se purificó el ADN con procedimientos estándar y se digirió con I-Sce I bajo condiciones óptimas.

60 - 2 - Células bacterianas:

- Se construyeron transposones Tn5 Mini que contenían el sitio de reconocimiento de I-Sce I en *E. coli* por medio de procedimientos estándar de ADN recombinante. Los transposones Tn 5 mini se mantienen sobre un plásmido conjugado. La conjugación bacteriana entre *E. coli* y *Yersinia* se utiliza para integrar al transposón Tn 5 mini en *Yersinia*. Se seleccionan células de *Yersinia* resistentes a Kanamicina, Estreptomicina o tetraciclina (vectores pTKM-ω, pTSM-ω y pTTc-ω, respectivamente).

Se pueden intentar varias estrategias para la inserción específica en el sitio de un fragmento de ADN de un plásmido dentro de un cromosoma. Esto hará posible insertar transgenes en sitios predeterminados sin etapas laboriosas de selección. Las estrategias son:

- 5 - 1 - Construcción de una célula transgénica en la cual el sitio de reconocimiento de I-Sce I es insertado en una única ubicación en un cromosoma. La transformación conjunta de la célula transgénica con el vector de expresión y un plásmido que contiene al gen de interés y un segmento homólogo a la secuencia en la cual se inserta el sitio de I-SceI.
- 10 - 2 - Inserción del sitio de reconocimiento de I-Sce I cerca o dentro del gen de interés transportado en un plásmido. Transformación conjunta de una célula normal con el vector de expresión que transporta al gen sintético para I-Sce I y al plásmido que contiene el sitio de reconocimiento de I-SceI.
- 15 - 3 - Construcción de una línea estable de células transgénicas en la cual se ha integrado el gen para I-Sce I en el genoma bajo el control de un promotor celular inducible o constitutivo. Transformación de la línea de células por medio de un plásmido que contiene al sitio de I-Sce I cerca o dentro del gen de interés.

Recombinación homóloga dirigida al sitio: los diagramas de los experimentos exitosos llevados a cabo en levadura se muestran en la Fig. 19.

Publicaciones citadas en la solicitud

1. B. **Dujon**, Sequence of the intron and flanking exons of the mitochondrial 21 S rRNA gene of yeast strains having different alleles at the w and RIB 1 loci. *Cell* (1980) 20, 185-187.
2. F. **Michel**, A. **Jacquier** and B. **Dujon**, Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure. *Biochimie*, 1982, 64, 867-881.
3. F. **Michel** and B. **Dujon**, Conservation of RNA secondary structures in two intron families including mitochondrial-, chloroplast-, and nuclear-encoded members. *The EMBO Journal*, 1983, 2, 33-38.
4. A. **Jacquier** and B. **Dujon**, The intron of the mitochondrial 21S rRNA gene: distribution in different yeast species and sequence comparison between *Kluyveromyces thermotolerans* and *Saccharomyces cerevisiae*. *Mol. Gen. Gent.* (1983) 192, 487-499.
5. B. **Dujon** and A. **Jacquier**, Organization of the mitochondrial 21S rRNA gene in *Saccharomyces cerevisiae*: mutants of the peptidyl transferase centre and nature of the omega locus in "Mitochondria 1983", Editors R. J. Schweyen, K. Wolf, F. Kaudewitz, *Walter de Gruyter et Co.*, Berlin, New York (1983), 389-403.
6. A. **Jacquier** and B. **Dujon**, An intron encoded protein is active in a gene conversion process that spreads an intrón into a mitochondrial gene. *Cell* (1985) 41, 383-394.
7. B. **Dujon**, G. **Cottarel**, L. **Colleaux**, M. **Betermier**, A. **Jacquier**, L. **D'Auriol**, F. **Galibert**, Mechanism of integration of an intron within a mitochondrial gene: a double strand break and the transposase function of an intron encoded protein as revealed by *in vivo* and *in vitro* assays. "In Achievements and perspectives of Mitochondrial Research". Vol. II, Biogenesis, E. Quagliariello y colaboradores. Eds. Elsevier, Amsterdam (1985) pages 215-225.
8. L. **Colleaux**, L. **D'Auriol**, M. **Betermier**, G. **Cottarel**, A. **Jacquier**, F. **Galibert**, and B. **Dujon**, A universal code equivalent of a yeast mitochondrial intron reading frame is expressed into *Escherichia coli* as a specific double strand endonuclease. *Cell* (1986) 44, 521-533.
9. B. **Dujon**, L. **Colleaux**, A. **Jacquier**, F. **Michel** and C. **Monteilhet**, Mitochondrial introns as mobile genetic elements: the role of intron-encoded proteins. In "Extrachromosomal elements in lower eucaryotes", Reed B y colaboradores. Eds. (1986) *Plenum Pub. Corp.* 5-27.
10. F. **Michel** and B. **Dujon**, Genetic Exchanges between Bacteriophage T4 and Filamentous Fungi? *Cell* (1986) 46, 323.
11. L. **Colleaux**, L. **D'Auriol**, F. **Galibert** and B. **Dujon**, Recognition and cleavage site of the intron encoded omega transposase. *PNAS* (1988), 85, 6022-6026.
12. B. **Dujon**, Group I introns as mobile genetic elements, facts and mechanistic speculations: A Review. *Gene* (1989), 82, 91-114.
13. B. **Dujon**, M. **Belfort**, R.A. **Butow**, C. **Jacq**, C. **Lemieux**, P.S. **Perlman**, V.M. **Vogt**, Mobile introns: definition of terms and recommended nomenclature. *Gene* (1989), 82, 115-118.

14. C. **Monteilhet**, A. **Perrin**, A. **Thierry**, L. **Colleaux**, B. **Dujon**, Purification and Characterization of the *in vitro* activity of I-SceI, a novel and highly specific endonuclease encoded by a group I intron. *Nucleic Acid Research* (1990), 18, 1407-1413.

15. L. **Colleaux**, M-R. **Michel-Wolwertz**, R.F. **Matagne**, B. **Dujon**-The apocytochrome b gene of *Chlamydomonas smithii* contains a mobile intron related to both *Saccharomyces* and *Neurospora* introns. *Mol. Gen. Genet.* (1990) 223, 288-296.

16. B. **Dujon** Des introns autonomes et mobiles. *Annales de l'Institut Pasteur/Actualites* (1990) 1. 181-194.

17. A. **Thierry**, A. **Perrin**, J. **Boyer**, C. **Fairhead**, B. **Dujon**, B. **Prey**, G. **Schmitz**. Cleavage of yeast and bacteriophage 17 genomes at a single site using the rare cutter endonuclease I-Sce. *I Nuc. Ac. Res.* (1991) 19, 189-190.

18. A. **Plessis**, A. **Perrin**, J.E. **Haber**, B. **Dujon**, Site specific recombination determined by I-SceI, a mitochondrial intron-encoded endonuclease expressed in the yeast nucleus. *GENETICS* (1992) 130, 451-460.

Resúmenes

A1. A. **Jacquier**, B. **Dujon**. Intron recombinational insertion at the DNA level: Nature of a specific receptor site and direct role of an intron encoded protein. *Cold Spring Harbor Symposium* 1984.

A2. L. **Colleaux**, L. **D'Auriol**, M. **Demariaux**, B. **Dujon**, F. **Galibert**, and A. **Jacquier**, Construction of a universal code equivalent from a mitochondrial intron encoded transposase gene using oligonucleotide directed multiple mutagenesis. Colloque International de DNRS "oligonucleotids et Genetique Moléculaire" Aussois (Savoie) 8-12 January 1985.

A3. L. **Colleaux**, **D'Auriol**, M. **Demariaux**, B. **Dujon**, F. **Galibert**, and A. **Jacquier**, Expression in *E. coli* of a universal code equivalent of a yeast mitochondrial intron reading frame involved in the integration of an intron within a gene. *Cold Spring Harbor Meeting* on "Molecular Biology of Yeast", August 13-19, 1985.

A4. B. **Dujon**, G. **Cottarel**, L. **Colleaux**, M. **Demariaux**, A. **Jacquier**, L. **D'Auriol**, and F. **Galibert**, Mechanism of integration of an intron within a mitochondrial gene: a double strand break and the "transposase" function of an intron encoded protein as revealed by *in vivo* and *in vitro* assays. International symposium on "Achievements and Perspectives in Mitochondrial Research", Selva de Fasano (Brindisi, Italy) 26 September 1985.

A5. L. **Colleaux**, G. **Cottarel**, M. **Betermier**, A. **Jacquier**, B. **Dujon**, L. **D'auriol**, and F. **Galibert**, Mise en évidence de l'activité endonuclease double brin d'unc protéin codée par un intron mitochondrial de levure. Forum sur la Biologie Moléculaire de la levure, Bonbannes, France 2-4 October 1985.

A6. B. **Dujon**, L. **Colleaux**, F. **Michel** and A. **Jacquier**, Mitochondrial introns as mobile genetic elements. In "Extrachromosomal elements in lower eucaryotes", Urbana, Illinois, 1-5 June 1986.

A7. L. **Colleaux** and B. **Dujon**, Activity of a mitochondrial intron encoded transposase. Yeast Genetics and Molecular Biology Meeting, Urbana, Illinois 3-6 June 1986.

A8. L. **Colleaux** and B. **Dujon**, The role of a mitochondrial intron encoded protein. XIIIth International Conference on Yeast Genetics and Molecular Biology, Banff, Alberta (Canada) 31 August-5 September 1986.

A9. L. **Colleaux**, L. **D'Aurio**, F. **Galibert** and B. **Dujon**, Recognition and cleavage specificity of an intron encoded transposase. 1987 Meeting on *Yeast Genetics and Molecular Biology*. San Francisco, CA 16-21 June 1987.

A10. A. **Perrin**, C. **Monteilhet**, L. **Colleaux** and B. **Dujon**, Biochemical activity of an intron encoded transposase of *Saccharomyces cerevisiae*. Cold Spring Harbor Meeting on "Molecular Biology of Mitochondria and chloroplasts" 25-30 August 1987 *Cold Spring Harbor*, N.Y.

A11. B. **Dujon**, A. **Jacquier**, L. **Colleaux**, C. **Monteilhet**, A. **Perrin**, "Les Introns autoépissables et leurs protéins" Colloque "Biologie Moléculaire de la levure: expression génétique chez *Saccharomyces*" organise par la Société française de Microbiologie 18 January 1988 Institut Pasteur, Paris.

A12. L. **Colleaux**, L. **D'Auriol**, C. **Monteilhet**, F. **Galibert** and B. **Dujon**, Characterization of the biochemical activity of an intron encoded transposase. 14th International Conference on Yeast Genetics and Molecular Biology. Espoo, Finland, 7-13 August 1988.

A13. B. **Dujon**, A group I intron as a mobile genetic element, Albany Conference sur "RNA: catalysis, splicing, evolution", Albany, New York, 22-25 September 1988.

- A14. B. Dujon, L. Colleaux, C. Monteilhet, A. Perrin, L. D'Auriol, F. Galibert, Group I introns as mobile genetic elements: the role of intron encoded proteins and the nature of the target site. 14th Annual EMBO Symposium "Organelle genomes and the nucleus" Heidelberg, 26-29 September 1988.
- 5 A15. L. Colleaux, R. Matagne, B. Dujon, A new mobile mitochondrial intron provides evidence for genetic exchange between Neurospora and Chlamydomonas species. *Cold Spring Harbor*, May 1989.
- A16. L. Colleaux, M.R. Michel-Wolwertz, R.F. Matagne, B. Dujon, The apoxytochrome b gene of *Chlamydomonas smithii* contains a mobile intron related to both Saccharomyces and Neurospora introns. Fourth International Conference on Cell and Molecular Biology of Chlamydomonas. Madison, WI, April 1990.
- 10 A17. B. Dujon, L. Colleaux, E. Luzi, C. Monteilhet, A. Perrin, A. Plessis, I. Stroke, A. Thierry, Mobile Introns, EMBO Workshop on "Molecular Mechanisms of transposition and its control", Roscoff (France) June 1990.
- 15 A18. A. Perrin, C. Monteilhet, A. Thierry, E. Luzi, I. Stroke, L. Colleaux, B. Dujon. I-SceI, a novel double strand site specific endonuclease, encoded by a mobile group I intron in Yeast. Workshop on "RecA and Related Proteins" Saclay, France 17-21 September 1990.
- A19. A. Plessis, A. Perrin, B. Dujon, Site specific recombination induced by double strand endonucleases, HO and I-Sce I in yeast. Workshop on "RecA and Related Proteins" Saclay, France 17-21 September 1990.
- 20 A20. B. Dujon, The genetic propagation of introns 20th FEBS Meeting, Budapest, Hungary, August 1990.
- A21. E. Luzi, B. Dujon, Analysis of the intron encoded site specific endonuclease I-Sce I by mutagenesis, Third European Congress on Cell Biology, Florence, Italy, September 1990.
- 25 A22. B. Dujon, Self splicing introns as contagious genetic elements. Journees Franco-Beiges de Pont a Mousson. October 1990.
- 30 A23. B. Frey, H. Dubler, G. Schmitz, A. Thierry, A. Perrin, J. Boyer, C. Fairhead, B. Dujon, Specific cleavage of the yeast genome at a single site using the rare cutter endonuclease I-Sce I Human Genome, Frankfurt, Germany, November 1990.
- A24. B. Dujon, A. Perrin, I. Stroke, E. Luzi, L. Colleaux, A. Plessis, A. Thierry, The genetic mobility of group I introns at the DNA level. Keystone Symposia Meeting on "Molecular Evolution of Introns and Other RNA elements", Taos, New Mexico, 2-8 February 1991.
- 35 A25. B. Dujon, J. Boyer, C. Fairhead, A. Perrin, A. Thierry, Cartographie chez la levure. Reunion "Strategies d'etablissement des cartes geniques" Toulouse 30-31 Mai 1991.
- 40 A26. B. Dujon, A. Thierry, Nested chromosomal fragmentation using the meganuclease I-SceI: a new method for the rapid mapping of the yeast genome. Elounda, Crete 15-17 Mail 1991.
- A27. A. Thierry, L. Gaillon, F. Galibert, B. Dujon. The chromosome XI library: what has been accomplished, what is left. Brugge meeting 22-24 September 1991.
- 45 A28. B. Dujon, A. Thierry, Nested chromosomal fragmentation using the meganuclease I-SceI: a new method for the rapid physical mapping of the eukaryotic genomes. *Cold Spring Harbor* 6-10 May 1992.
- 50 A29. A. Thierry, L. Gaillon, F. Galibert, B. Dujon. Yeast chromosome XI: construction of a cosmid contig. a high resolution map and sequencing progress. *Cold Spring Harbor* 6-10 May 1992.

En preparación

- 55 P1. A. Thierry and B. Dujon, Nested Chromosomal Fragmentation Using the Meganuclease I-SceI: Application to the physical mapping of a yeast chromosome and the direct sorting of cosmid libraries. Probably Submission to GENOMICS or *EMBO J*.
- 60 P2. A. Thierry, L. Colleaux and B. Dujon: Construction and Expression of a synthetic gene coding for the meganuclease I-SceI. Possible submission: NAR, *EMBO J*.
- P3. I. Stroke, V. Pelicic and B. Dujon: The evolutionarily conserved dodecapeptide motifs of intron-encoded I-Sce I are essential for endonuclease function. Submission to *EMBO J*.
- 65 P4. C. Fairhead and B. Dujon: Consequences of a double strand break induced *in vivo* in yeast at specific artificial sites, using the meganuclease I-SceI. Possible submission to GENETICS, *NATURE*.

P5 A. Perrin, and B. Dujon: Asymetrical recognition by the I-Sce I endonuclease on exon and intron sequences reveals a new step in intron mobility. Possible submission: *NATURE*.

Inducción de recombinación homóloga en cromosomas de mamífero utilizando el sistema de I-Sce I de *Saccharomyces cerevisiae*

Ejemplo 4

10 Introducción

La recombinación homóloga (HR) entre ADN cromosómico y exógeno es la base de los métodos para la introducción de cambios genéticos en el genoma (5B, 20B). Se han determinado los parámetros del mecanismo de recombinación por medio del estudio de secuencias de plásmidos introducidas en células (1B, 4B, 10B, 12B) y en un sistema *in vitro* (8B). La HR es ineficiente en células de mamífero pero es promovida por medio de rompimientos bicatenarios en ADN.

Hasta ahora, no ha sido posible escindir eficientemente un objetivo cromosómico específico, limitando así nuestra comprensión de la recombinación y su explotación. Entre las endonucleasas, la endonucleasa mitocondrial I-Sce I de *Saccharomyces cerevisiae* (6B) tiene características que pueden ser explotadas como una herramienta para la escisión de un objetivo cromosómico específico y, por lo tanto, la manipulación del cromosoma en organismos vivos. La proteína I-Sce I es una endonucleasa responsable de la búsqueda del intrón en mitocondria de levadura, un mecanismo no recíproco por medio del cual se inserta una secuencia predeterminada en un sitio predeterminado. Se ha establecido que la I-Sce I endonucleasa puede catalizar la recombinación en el núcleo de levadura por medio de la iniciación de un rompimiento bicatenario (17B). El sitio de reconocimiento de la I-Sce I endonucleasa tiene un largo de 18 pb, por lo tanto, la proteína I-Sce I es una endonucleasa de restricción de corte muy rara en genomas (22B). Además, como la proteína I-Sce I no es una recombinasa, su potencial para someter a ingeniería genética al cromosoma es mayor que aquella de sistemas con requerimiento de sitios objetivo tanto en moléculas huésped como donantes (9B).

Nosotros demostramos aquí que la I-Sce I endonucleasa de levadura puede inducir eficientemente rompimientos bicatenarios en objetivos cromosómicos en células de mamífero y que los rompimientos se pueden reparar utilizando una molécula donante que comparte homología con las regiones que flanquean el rompimiento. La enzima cataliza la recombinación con una alta eficiencia. Esto demuestra que la recombinación entre ADN cromosómico y ADN exógeno se puede presentar en células de mamífero por medio de la ruta de reparación del rompimiento bicatenario (21B).

Materiales y métodos

40 Construcción del plásmido

pG-MPL fue obtenido en cuatro etapas: (I) inserción del fragmento *Bgl* II-*Sma* I de 0,3 kb (tratado con enzima de Klenow) del gen env (25B) del Virus de Leucemia de Múrido de Moloney (MoMuLV) que contiene SA entre los sitios *Nhe* I y *Xba* I (tratado con enzima de Klenow), en la secuencia U3 de la LTR 3' de MoMuLV, en un plásmido intermedio. (II) inserción en esta LTR modificada con adaptadores de enlazamiento del fragmento *Nco* I-*Xho* I de 3,5 kb que contiene al gen de fusión PhleoLacZ (15B) (de pUT65 del laboratorio Cayla) en el sitio *Xba* I próximo a SA. (III) inserción de esta LTR 3' (que contiene SA y PhleoLacZ), recuperada por digestión doble con *Sal* I-*EcoR* I en el plásmido p5'LTR (un plásmido que contiene la LTR 5' para el nucleótido número 563 de MoMuLV (26B) entre los sitios *Xho* I y *EcoR* I, e (VI) inserción de un sitio de reconocimiento sintético de I-Sce I dentro del sitio *Nco* I en la LTR 3' (entre SA y PhleoLacZ).

pG-MtkPI fue obtenido por la inserción (antisentido al genoma retroviral) del gen tk de 1,6 kb con su promotor con los adaptadores de enlazamiento en el sitio *Pst* I de pG-MPL. pVRneo fue obtenido en dos etapas (I) inserción dentro de pSP65 (de Promega) linealizado por digestión doble con *Pst* I-*EcoR* I del fragmento *Pst* I hasta *EcoR* I de 4,5 kb de pG-MPL que contiene la LTR 3' con el SA y PhleoLacZ, (II) inserción del fragmento *Bgl* II-*BamH* I de 2,0 kb (tratado con enzima de Klenow) que contiene neoPoliA de pRSVneo dentro del sitio de restricción de *Nco* I (tratado con enzima de Klenow) de pSP65 que contiene parte de la LTR 3' de G-MPL (entre SA y PhleoLacZ).

pCMV(I-Sce I+) fue obtenido en dos etapas: (I) inserción del fragmento que contiene *BamH* I-*Sal* I, I-Sce I de 0,73 kb (de pSCM525, A. Thierry, obsequio personal) dentro del plásmido pCMV1 (F. Meyer, obsequio personal) escindido en los sitios de *BamH* I y de *Sal* I, (II) inserción de un fragmento de 1,6 kb (número de nucleótidos 3204 a 1988 en SV40) que contiene la señal de poliadenilación de SV40 dentro del sitio de *Pst* I de pCMV1.

pCMV(I-Sce I-) contiene el ORF de I-Sce I en orientación inversa en el plásmido pCMV(I-Sce I+). Ha sido obtenido por inserción del fragmento del ORF de *BamH* I-*Pst* I-*I-Sce* I (tratado con enzima de Klenow) dentro del vector PoliA de pCMV linealizado por medio de digestión doble con *Nsi* I y *Sal* I y tratado con enzima de Klenow.

ES 2 338 952 T3

Se han descrito los plásmidos pG-MPL, pG-MtkPL, pG-MtkDPAPL. Además de los plásmidos descritos anteriormente, se puede construir cualquier tipo de vector plásmido que contenga diferentes promotores, genes, sitio de poliA, sitio de I-Sce I.

Cultivo y selección de células

3T3, PCC7 S, Ψ 2 se referencian en (7B) y (13B). Medio de selección celular: se añadió ganciclovir (14B, 23B) en el medio para cultivo de tejidos en una concentración de 2 μ M. Se mantuvo la selección con ganciclovir sobre las células durante 6 días. Se añadió G418 en el medio apropiado en una concentración de 1 mg/ml para PCC7-S y de 400 μ g/ml para 3T3. Se mantuvo la selección durante todo el cultivo celular. Se utilizó fleomicina en una concentración de 10 μ g/ml.

Líneas celulares

- la línea de células Ψ fue transfectada con plásmidos que contienen un vector recombinante proviral que contiene el sitio de reconocimiento de I-Sce I: *pG-MPL*, *pG-MtkPL*, *pG-Mtk_{ΔPA}PL*

- la línea de células fibroblásticas NIH 3T3 se infecta con:

G-MPL. Se recuperaron múltiples clones (más de 30). La presencia de 1 a 14 integraciones provirales y la multiplicidad de los diferentes puntos de integración fueron verificadas por medio de análisis molecular.

G-MtkPL. Se recuperaron 4 clones (3 de ellos tienen una integración proviral normal y 1 de ellos tiene una recombinación entre las dos LTR que presenta únicamente un sitio de reconocimiento de I-Sce I).

- la línea de células PCC7-S de *carcinoma embrionario* se infecta con:

G-MPL. Se recuperaron 14 clones, integración proviral normal.

- la línea de células madre embrionarias D3 se infectan con:

G-MPL. Se recuperaron 4 clones (3 tienen integración proviral normal, 1 tiene 4 integraciones provirales).

Células de ratón "Preparadas"

La inserción del retrovirus (integración proviral) induce duplicación de la LTR que contiene el sitio de I-Sce I. La célula es heterocigótica para el sitio.

Transfección, infección, coloración de la célula y análisis de transferencias de ácidos nucleico

Estos procedimientos se llevaron a cabo como se describe en (2B, 3B).

Resultados

Para detectar la HR de I-Sce I hemos diseñado el sistema experimental mostrado en la Fig. 20. Se construyeron retrovirus recombinantes defectuosos (24B) con el sitio de reconocimiento de I-Sce I y un gen de fusión *PhleoLacZ* (15B) insertado en su LTR 3' (Fig. 20a). La integración retroviral resulta en dos sitios de I-Sce I distantes entre sí 5,8 kb ó 7,2 kb dentro del genoma de la célula (Fig. 20b). Nosotros sugerimos la hipótesis de que los rompimientos bicatenarios (DSB) inducidos por I-Sce I en estos sitios (Fig. 20c) podrían iniciar la HR con un plásmido donante (pVRneo, Fig. 20d) que contiene secuencias homólogas a las regiones de flanco de los DSB y que las secuencias no homólogas, transportadas por el plásmido donante, podrían ser copiadas durante esta recombinación (Fig. 20e).

Introducción de sitios de reconocimiento duplicados de I-Sce I dentro del genoma de células de mamífero por integración de retrovirus

Más específicamente, se utilizaron dos secuencias provirales en estos estudios. Las secuencias provirales de G-MtkPL (del virus G-MtkPL) contienen el gen de fusión *PhleoLacZ* para selección positiva de células transducidas (en medio que contiene fleomicina) y el gen tk para selección negativa (en medio que contiene ganciclovir). Las secuencias provirales de G-MPL (del virus G-MPL) contienen únicamente las secuencias de *PhleoLacZ*. G-MtkPL y G-MPL son retrovirus recombinantes defectuosos (16B) construidos a partir de un provirus de leucemia de múrido de Moloney sin potenciador. El vector viral funciona como una trampa del promotor y por lo tanto es activado por medio de promotores celulares de flanco.

Se generaron líneas celulares que producen virus por medio de la transfección de pG-MtkPL o de G-MPL dentro de la línea de células en paquete Ψ -2 (13B). El análisis de transferencias tipo Northern de transcriptos virales muestra (Fig. 21) que la línea Ψ -2-G-MPL expresa transcriptos de 4,2 y 5,8 kb que hibrida con sondas LacZ. Estos transcriptos probablemente inician en la LTR 5' y terminan en la LTR 3'. El transcripto de 4,5 kb corresponde al mensaje empalmado y los transcriptos de 5,8 kb al mensaje genómico no empalmado (Fig. 21.A). Esto verificó la funcionalidad de la LTR 5' y del donante y aceptor de empalme en el virus. Se han obtenido resultados similares con Ψ -2G-MtkPL. Se preparó un virus a partir del medio de cultivo de líneas de células Ψ -2.

Se infectaron a continuación fibroblastos NIH3T3 y líneas de células multipotentes de ratón PCC7-S (7B) por medio de G-MtkPL y G-MPL, respectivamente, y se aislaron los clones. El análisis de transferencias tipo Southern del ADN preparado a partir de los clones demostró duplicación mediada por LTR de secuencias de *PhleoLacZ* para I-Sce I (Fig. 22.a). La digestión con *Bcl* I generó los fragmentos esperados de 5,8 kb (G-MPL) o de 7,2 kb (G-MtkPL). La presencia de dos fragmentos adicionales correspondientes a sitios de *Bcl* I en el ADN cromosómico de flanco demuestra un único objetivo proviral en cada clon aislado. Sus diferentes tamaños entre los clones indican integración de retrovirus en distintos loci. Esos sitios de reconocimiento de I-Sce I que han sido fielmente duplicados fueron demostrados por medio de digestiones con I-Sce I que generaron fragmentos de 5,8 kb (G-MPL) o 7,2 kb (G-MtkPL) (Fig. 22.b).

Inducción por I-Sce I de recombinación que conduce a intercambio de ADN

El fenotipo conferido a las células NIH3T3 por el virus G-MtkPL es $\text{fleo}^R \beta\text{-gal}^+ \text{gls}^S$ y a las células PCC7-S por GMPL es $\text{fleo}^R \beta\text{-gal}^+$ (Fig. 20b). Para permitir una selección directa de eventos de recombinación inducidos por I-Sce I construimos el plásmido donante pVRneo. En pVRneo el gen neo está flanqueado por 300 pb homólogos a secuencias 5' con el rompimiento cromosómico izquierdo y 2,5 kb homólogos a las secuencias 3' con el rompimiento derecho (Fig. 20d). Se posicionó una señal de poliadenilación 3' al gen neo para interrumpir el mensaje de *PhleoLacZ* después de la recombinación. Si ocurre una recombinación inducida entre el provirus y el plásmido, el fenotipo resultante será neo^R y debido a la presencia de una señal de poliadenilación en el plásmido donante, el gen de *PhleoLacZ* no se debe expresar, resultando en un fenotipo $\text{fleo}^S \beta\text{-gal}^-$.

Con G-MtkPL y G-MtkDPQPL, es posible seleccionar simultáneamente la brecha por medio de selección negativa con el gen tk (con ganciclovir) y para el intercambio del plásmido donante con selección positiva con el gen neo (con genética). Con G-MPL únicamente se puede aplicar selección positiva en medio que contiene genética. Por lo tanto, esperamos seleccionar tanto para la HR como para un evento de integración del plásmido donante cerca de un promotor endógeno activo. Estos dos eventos se pueden distinguir como una HR inducida que resulta en un fenotipo $\text{neo}^R \beta\text{-gal}^-$ y en una integración aleatoria del plásmido donante que resulta en un fenotipo $\text{neo}^R \beta\text{-gal}^+$.

Se transfectaron conjuntamente luego dos clones NIH3T3/G-MtkPL diferentes y tres PCC7S/G-MPL diferentes con un vector de expresión para I-Sce I, pCMV(I-Sce I+), y el plásmido donante, pVRneo. La expresión transitoria de I-Sce I puede resultar en los DSB en sitios de I-Sce I, promoviendo por lo tanto HR con pVRneo. El control es la transfección conjunta con un plásmido que no expresa I-Sce I, pCMV(I-Sce I-), y pVRneo.

Se seleccionaron los clones NIH3T3/G-MtkPL ya sea por la pérdida de secuencias provirales y la adquisición del fenotipo neo^R (con ganciclovir y genética) o únicamente por el fenotipo neo^R (Tabla 1). En el primer caso, se recuperaron colonias $\text{neo}^R \text{gls}^R$ con una frecuencia de 10^{-4} en la serie experimental, y no se recuperaron colonias en la serie de control. Además, todas las colonias $\text{neo}^R \text{gls}^R$ eran $\beta\text{-gal}^-$, de conformidad con las suyas derivadas de HR en el sitio proviral. En el segundo caso, se recuperaron colonias neo^R con una frecuencia de 10^{-3} en la serie experimental, y con una frecuencia 10 a 100 veces menor en la serie de control. Además, se encontró que 90% de las colonias neo^R eran $\beta\text{-gal}^-$ (en series con pCMV(I-Sce I+)). Esto muestra que la expresión de I-Sce I induce HR entre pVR neo y el sitio proviral y que la HR dirigida al sitio es diez veces más frecuente que la integración al azar de pVR neo cerca de un promotor celular, y al menos 500 veces más frecuente que HR espontánea.

TABLA 1

Recombinación homóloga inducida con I-Sce I

Selección	G418 + Gls				G418			
	+		-		+		-	
Expresión de I-Sce I								
Fenotipo β -gal	+	-	+	-	+	-	+	-
(A) Línea celular								
NIH 3T3/G-MtkPL								
Clon 1	0	66	0	0	69	581	93	0
Clon 2	0	120	0	0	15	742	30	0
PCC7-S/G-MPL								
Clon 3					54	777	7	0
Clon 4					2	91	1	0
Clon 5					7	338	3	0
(B) Evento molecular								
RI	0					816		
DsHR	15				0	19	0	
SsHR	0				0	4	0	
Del	0					01	0	

Tabla 1: Efecto de la escisión bicatenaria mediada por I-Sce I. A. 10^6 células de los clones 1 y 2 de NIH3T3/G-MtkPL y $5 \cdot 10^6$ células de los clones 3 a 5 de PCC7-S/G-MPL fueron transfectados conjuntamente con pVRneo y o bien pCMV(I-Sce I+) o pCMV(I-Sce I-). Se seleccionaron las células en el medio indicado: Geneticina (G418) o geneticina + ganciclovir (G418_Gls). El fenotipo de la expresión de β -gal se determinó por medio de coloración histoquímica con X-gal. Si ocurre una recombinación inducida entre el provirus y pVRneo, las células adquieren un fenotipo neo^R β -gal⁻. B. Análisis molecular de una muestra de clones recombinantes. RI: integración al azar de pVRneo, estructura proviral progenitora.

DsHR: HR de doble sitio. SsHR: HR de un solo sitio. Del: supresión del provirus (ver también la Fig. 20 y la 23).

Verificación de la recombinación por medio del análisis de transferencias tipo Southern y Northern

Se ha examinado la estructura molecular de los recombinantes neo^R por medio de análisis de transferencias tipo Southern (Fig. 23 y Tabla 1). La HR en los sitios de I-Sce I predice que la digestión de ADN recombinante genera un fragmento de *LacZ* de 6,4 kb en vez del fragmento progenitor de 4,2 kb. Todos los 15 recombinantes neo^R gls^R β -gal⁻ a partir de células NIH3T3 exhibieron el fragmento *Kpn* I de 6,4 kb. Por lo tanto, el procedimiento de selección doble conduce únicamente a los recombinantes esperados creados por medio de reemplazo del gen (Recombinantes Homólogos de Sitio Doble, DsHR).

Los 25 recombinantes β -gal⁻ generados a partir de la selección individual caen dentro de cuatro clases: (a) DsHR inducidos por I-Sce I como anteriormente (19 clones); (b) integración de pVRneo en la LTR izquierda como se prueba por la presencia de un fragmento de *Kpn* I de 4,2 (correspondiente a *PhleoLacZ* en la LTR restante), además de fragmento de 6,4 kb (Fig. 23, Tabla 1, Recombinantes Homólogos de un Solo Sitio, SsHR; 3 recombinantes β -gal⁻ independientes del clon 3). Estos clones corresponden a I-Sce I-IHR en DSB izquierdo únicamente o (menos probable) a cruzamiento doble entre LTR y pVRneo; (c) integraciones al azar de pVRneo (Tabla 1, Integraciones al Azar, IR) y HR simultánea (Tabla 1, Supresión, Del) (1 recombinante β -gal⁻); y (d) Integración al azar de pVRneo y supresión simultánea de provirus (1 recombinante β -gal⁻). Nosotros sugerimos que estas cuatro clases corresponden a reparación de los DSB con el cromosoma homólogo. Como se esperaba, todos los recombinantes β -gal⁺ de selección con geneticina únicamente, corresponden a integraciones aleatorias de pVRneo, ya sea que se originen a partir de la serie experimental (ocho clones analizados) o a partir de la serie de control (seis clones analizados).

Nosotros obtuvimos evidencia adicional de que la recombinación había ocurrido en el sitio de I-Sce I de PCC7-S/G-MPL 1 por medio del análisis de los ARN producidos en las células progenitoras y en el recombinante (Fig. 24). Las células progenitoras PCC7-S/G-MPL 1 expresan un ARN de *LacZ* de 7,0 kb indicativo del atrapamiento de un promotor celular que conduce a expresión de un ARN de fusión celular-viral. El clon recombinante no expresa este ARN de *LacZ* pero expresa un ARN neo de 5,0 kb. El tamaño del ARN neo que corresponde al tamaño exacto esperado para un intercambio preciso de *PhleoLacZ* por el gen neo y utiliza el mismo sitio de empalme celular y viral (ARN viral de *PhleoLacZ* en la LTR es de 3,7 kb y el ARN neo en pVRneo es de 1,7 kb).

Discusión

Los resultados presentados aquí demuestran que se pueden inducir rompimientos bicatenarios por medio del sistema de I-Sce I de *Saccharomyces cerevisiae* en células de mamífero, y que los rompimientos en la secuencia cromosómica objetivo inducen recombinación específica del sitio con entrada de ADN donante plasmídico.

Para operar en células de mamífero, el sistema requiere que la actividad del tipo de la I-Sce I endógena esté ausente de las células de mamífero y que la proteína I-Sce I sea neutra para células de mamífero. Es poco probable que actividad como la de la I-Sce I endógena opere activamente en células de mamífero ya que la introducción de sitios de reconocimiento de I-Sce I no parece conducir a reordenamiento o mutación en las secuencias de entrada de ADN. Por ejemplo, todos los clones de NIH3T3 y de PCC7-S infectados con un retrovirus que contiene al sitio de restricción de I-Sce I propagaron en forma estable al virus. Para analizar la toxicidad del producto génico del I-Sce I, se introdujo un plásmido que exprese I-Sce I en la línea de células NIH3T3 (datos no mostrados). Se encontró un porcentaje muy alto de transferencias conjuntas de un gen funcional para I-Sce I, lo que sugiere que no hay selección contra este gen. Se demostró la funcionalidad del gen para I-Sce I por medio del análisis de la transcripción, por medio de la detección de inmunofluorescencia del producto génico y de la función biológica (Chouluka y colaboradores, en preparación).

Luego analizamos si la endonucleasa escindiría un sitio de reconocimiento ubicado sobre un cromosoma. Esto se logró colocando dos sitios de reconocimiento de I-Sce I separados por 5,8 ó 7,2 kb sobre un cromosoma en cada LTR de estructuras provirales y por medio del análisis de los productos de una reacción de recombinación con un vector objetivo en presencia del producto génico I-Sce I. Nuestros resultados indican que en presencia de I-Sce I, el vector donante se recombina muy eficientemente con secuencias dentro de las dos LTR para producir un gen neo funcional. Esto sugiere que I-Sce I indujo muy eficientemente rompimientos bicatenarios en ambos sitios de I-Sce I. Además, ya que se obtuvieron rompimientos bicatenarios con al menos cinco inserciones provirales distintas, la habilidad de la proteína I-Sce I para digerir un sitio de reconocimiento de I-Sce I no es altamente dependiente de las estructuras de los alrededores.

La demostración de la habilidad de la I-Sce I meganucleasa para tener función biológica sobre sitios cromosómicos en células de mamífero allana la ruta para una cantidad de manipulaciones del genoma en organismos vivos. En comparación con las recombinasas específicas para el sitio (9B, 18B), el sistema de I-Sce I no es reversible. Las recombinasas específicas para el sitio localizan no únicamente los sitios para cortar el ADN, sino también para reunir poniendo nuevamente juntos las dos parejas. En contraste, el único requerimiento con el sistema de I-Sce I es la homología de la molécula donante con la región que flanquea el rompimiento inducido por la proteína I-Sce I.

Los resultados indican por primera vez que los rompimientos de ADN bicatenario in objetivos cromosómicos estimulan la HR con ADN introducido en células de mamífero. Ya que nosotros utilizamos una combinación de rompimientos bicatenarios (DSB) en ADN del destinatario cromosómico y ADN donante superenrollado, exploramos la estimulación por medio de la I-Sce I endonucleasa de recombinación por medio de la ruta de reparación del rompimiento bicatenario (21B). Por lo tanto, el rompimiento inducido es probablemente reparado por medio de un evento de conversión del gen que involucra la participación concertada de ambos extremos rotos los cuales, después de la creación de la región monocatenaria por digestión exonucleolítica de 5' a 3', invaden y copian ADN de la copia donante. Sin embargo, una cantidad de estudios de recombinación en células de mamífero y en levadura (10B, 11B, 19B) sugieren que existe una ruta alternativa de recombinación llamada de apareamiento monocatenario (SSA). En la ruta SSA, los rompimientos bicatenarios son sustratos en la acción de una exonucleasa que expone ADN monocatenario complementario homólogo sobre el ADN donante y receptor. El apareamiento de la cadena complementaria es luego seguido por un proceso de reparación que genera recombinantes. El sistema de I-Sce I puede ser utilizado para evaluar la importancia relativa de las dos rutas.

Ejemplo 5

Este ejemplo describe el uso de la I-Sce I meganucleasa (involucrada en la búsqueda del intrón de mitocondria de la levadura *Saccharomyces cerevisiae*) (6B, 28B) para inducir DSB y recombinación mediata en células de mamífero. I-Sce I es una endonucleasa de restricción de corte muy rara, con un sitio de reconocimiento con una longitud de 18 pb (29B, 22B). *in vivo*, la I-Sce I endonucleasa puede inducir recombinación en un núcleo modificado de levadura por medio de la iniciación de un DSB específico que conduce a la reparación de un brecha por parte de la célula (30B, 17B, 21B). Por lo tanto, se puede utilizar potencialmente esta aproximación como un medio para introducir DSB específicos en ADN cromosómico objetivo con el fin de manipular cromosomas en células vivas. La recombinación mediata por I-Sce I es superior al sistema de la recombinasa [11] para la modificación por ingeniería genética del cromosoma ya que esto último requiere de la presencia de sitios objetivo tanto sobre moléculas de ADN donantes como hospederas, conduciendo a una reacción que es reversible.

La expresión de la I-Sce I endonucleasa incluye eventos de recombinación. Por lo tanto, la actividad de I-Sce I puede provocar rompimientos bicatenarios dirigidos al sitio (DSB) en un cromosoma de mamífero. Al menos dos tipos de eventos se presentan en la reparación de los DSB, uno conducente a recombinación homóloga intracromosómica y el otro a la supresión del transgén. Estas recombinaciones mediadas por I-Sce I se presentan con una frecuencia significativamente mayor que lo normal.

Material y métodos

Construcción del plásmido

Se obtuvo pG-MtkPL en cinco etapas: (I) inserción del fragmento *Bgl* II-*Sma* I de 0,3 kpb (tratado con enzima de Klenow) del gen *env* (25B) del Virus de Leucemia de Múrido de Moloney (MoMuLV) que contiene un aceptor de empalme (SA) entre los sitios *Nhe* I y *Xba* I (tratado con enzima de Klenow), en la secuencia U3 de la LTR 3' de MoMuLV, en un plásmido intermedio. (II) Inserción en esta LTR modificada de un fragmento *Nco* I-*Xho* I de 3,5 kpb que contiene al gen de fusión *PhleoLacZ* [13] (de pUT65; Cayla Laboratory, Zone Commerciale du Gros, Toulouse, Francia) en el sitio de *Ikba* I próximo a SA. (III) Inserción de esta LTR 3' (que contiene SA y *PhleoLacZ*), recubierta por digestión doble con *Sal* I-*EcoR* I en el plásmido p5'LTR (un plásmido que contiene la LTR 5' hasta el nucleótido nº 563 de MoMuLV [12]) entre el sitio *Xho* I y el sitio *EcoR* I. (IV) Inserción de un sitio de reconocimiento sintético de *I-Sce* I dentro del sitio *Nco* I en la LTR 3' (entre SA y *PhleoLacZ*), e (V) inserción (antisentido al genoma retroviral) del gen tk de 1,6 kpb con su promotor con adaptadores de enlazamiento en el sitio *Pst* I de pG-MPL.

Se obtuvo pCMV(I-*Sce* I+) en dos etapas: (I) inserción del fragmento que contiene *I-Sce* I de *BamH* I-*Sal* I de 0,73 kpb (de pSCM525, donado por A. Thierry) dentro del plásmido phCMV1 (donado por F. Meyer) escindido con *BamH* I y *Sal* I, (II) inserción de un fragmento de 1,6 kpb (nucleótido nº 3204 hasta 1988 en SV40) que contiene la señal de poliadenilación de SV40 en el sitio de *Pst* I de phCMV1.

pCMV(I-*Sce* I-) contiene al ORF de *I-Sce* I en orientación inversa en el plásmido pCMV(I-*Sce* I+). Fue obtenido por inserción del fragmento del ORF de *I-Sce* I *BamH* I-*Pst* I (tratado con enzima de Klenow) dentro del vector phCMV PoliA linealizado por digestión doble con *Nsi* I y *Sal* I y tratado con enzima de Klenow.

Cell cultura y selección

T3 y Ψ 2 se referencian en (7B) y (13B). Medio de selección celular: se añadió ganciclovir (14B, 23B) dentro del medio para cultivo de tejidos en una concentración de 2 μ M. Se mantuvo la selección con ganciclovir durante 6 días. Se utilizó fleomicina en una concentración de 10 μ g/ml. Se llevaron a cabo selecciones dobles en las mismas condiciones.

Transfección, infección, coloración de las células y análisis de las transferencias de ácidos nucleicos

Estos protocolos fueron llevados a cabo como se describe en (2B, 3B).

Líneas de células que producen Virus

La línea de células que produce virus se genera por transfección de pG-MtkPL dentro de la línea de células de empaquetamiento Ψ -2. Se prepare el virus a partir del medio de cultivo filtrado de líneas de células Ψ -2 transfectadas. Se infectaron fibroblastos de NIH3T3 por medio de G-MtkPL, y se aislaron los clones en un medio que contiene Fleomicina.

Resultados

Para evaluar la actividad de la *I-Sce* I endonucleasa en células de mamífero, se utilizaron células NIH3T3 que contienen al provirus G-MtkPL. El provirus G-MtkPL (Fig. 25a) contiene al gen tk (en lugar de los genes virales *gag*, *pol* y *env*), para selección negativa en medio que contiene ganciclovir y, en las dos LTR, un sitio de reconocimiento de *I-Sce* I y al gen de fusión *PhleoLacZ*. El gen *PhleoLacZ* puede ser utilizado para selección positiva de células transducidas en medio que contiene fleomicina.

Nosotros formulamos la hipótesis de que la expresión de la *I-Sce* I endonucleasa en estas células induciría rompimientos bicatenarios (DSB) en los sitios de reconocimiento de *I-Sce* I que serían reparados por medio de uno de los siguientes mecanismos (ilustrados en la Fig. 25): a) si la *I-Sce* I endonucleasa induce un corte únicamente en una de las dos LTR (Fig. 1-b 1 y 2), las secuencias que son homólogas entre las dos LTR podrían aparearse y recombinarse conduciendo a una recombinación homóloga intracromosómica (es decir, por medio de apareamiento monocatenario (SSA) (12B, 10B) o cruzamiento); b) Si la *I-Sce* I endonucleasa induce un corte en cada una de las dos LTR, se pueden ligar nuevamente los dos extremos libres (mecanismo de unión de los extremos (31B) que conduce a una recombinación intracromosómica (Fig. 25-b 3); o alternativamente c) la brecha creada por los dos DSB se puede reparar por medio de un mecanismo para reparación de la brecha utilizando secuencias ya sea sobre el cromosoma homólogo o sobre otros segmentos cromosómicos, conduciendo a la pérdida de las secuencias provirales (32B) (Fig. 25-c).

El fenotipo conferido a las células NIH3T3 por el provirus G-MtkPL es $\text{Phleo}^R\beta\text{-Gal}^+ \text{Gls}^-$. En una primera serie de experimentos, buscamos la recombinación por medio de la selección por la pérdida del gen tk. Se transfectaron NIH3T3/G-MtkPL 1 y 2 (dos clones independientes con un sitio de integración proviral diferente) con el vector de expresión pCMV(I-*Sce* I+) de *I-Sce* I o con el plásmido de control pCMV(I-*Sce*-) que no expresa la *I-Sce* I endonucleasa. Se propagaron luego las células en medio que contiene Ganciclovir para selección por la pérdida de actividad de tk. Los clones GlsR resultantes fueron también analizados por la actividad de β -galactosidasa por medio de coloración histoquímica (con X-gal) (Tabla 1).

TABLA 1

Número y naturaleza de los clones resistentes a GlS				
Expresión de I-Sce I	pCMV(I-SceI+)		pCMV(I-SceI-)	
Actividad de β -Gal	+	-	+	-
NIH3T3/G-MtkPL 1	11	154	0	0
NIH3T3/G-MtkPL 2	16	196	2	0

Tabla 1: Efecto de la expresión de I-Sce I sobre la frecuencia de recombinación. Se transfectaron 1×10^6 células de NIH3T3/G-MtkPL 1 y 2×10^6 células de NIH3T3/G-MtkPL 1 ya sea con pCMV(I-Sce I+) o con pCMV(I-Sce I-). Se cultivaron las células en medio que contenía ganciclovir. Se determinó el fenotipo de β -Galactosidasa de los clones GlS^R por medio de coloración histoquímica con X-Gal.

En la serie de control transfectada con pCMV(I-SceI-), se encontraron clones GlS^R resistentes con una frecuencia baja (2 clones por 3×10^{-6} células tratadas) y los dos eran β -Gal⁺. En la serie experimental transfectada con pCMV (I-SceI+), la expresión del gen para I-Sce I incrementó la frecuencia de los clones GlS^R 100 veces. Estos clones eran o bien β -Gal⁻ (93%) o β -Gal⁺ (7%). Se analizaron cinco clones β -Gal⁻ de las NIH3T3/G-MtkPL 1 y seis de las NIH3T3/G-MtkPL 2 por medio de transferencias tipo Southern utilizando *Pst* I (Fig. 26). En el ADN progenitor, la *Pst* I endonucleasa corta dos veces en el gen tk del provirus (Fig. 26a). Los tamaños de los dos fragmentos que contenían *PhleoLacZ* se determinan por medio de la posición de los sitios de *Pst* I en el ADN celular de flaqueo. En NIH3T3/G-MtkPL 1, estos dos fragmentos de *PhleoLacZ* son de 10 kpb de largo y en NIH3T3/G-MtkPL 2 son de 7 y 9 kpb de largo. Los cinco clones resistentes GlS^R β -Gal⁻ de NIH3T3/G-MtkPL 1 y los seis clones del NIH3T3/G-MtkPL 2 mostraron todos la ausencia del gen tk y de las dos secuencias de *PhleoLacZ* (Fig. 26b y c).

En la serie experimental el número de clones GlS^R β -Gal⁺ se incrementa aproximadamente 10 veces por la expresión de I-Sce I en comparación con la serie de control. Estas no fueron analizadas adicionalmente.

Con el propósito de incrementar el número de clones GlS^R β -Gal⁺ recuperados, en un segundo grupo de experimentos, las células fueron cultivadas en un medio que contenía tanto Ganciclovir como Fleomicina. Ganciclovir selecciona las células que tienen pérdida de actividad de tk y la Fleomicina para las células que mantuvieron al gen para *PhleoLacZ*. Transfectamos NIH3T3/G-MtkPLs 1 y 2 con pCMV(I-SceI+) o pCMV(I-SceI-) (Tabla 2).

TABLA 2

Número de clones resistentes a Phleo y GlS		
Expresión de I-Sce I	pCMV(I-SceI+)	pCMV(I-SceI-)
NIH3T3/G-MtkPL 1	74	2
NIH3T3/G-MtkPL 2	207	9

Tabla 2: Efecto de la expresión de I-Sce I sobre la frecuencia de la recombinación intracromosómica. Se transfectaron 2×10^6 células de NIH3T3/GMtkPL 1 y 9×10^6 células de NIH3T3/G-MtkPL 2 ya sea con pCMV(I-Sce I+) o con pCMV(I-Sce I-). Se cultivaron las células en medio que contenía Fleomicina y ganciclovir.

En la serie de control, la frecuencia de recuperación de clones resistentes a Phleo^R GlS^R fue de 1×10^{-6} . Este resultado refleja células que han perdido espontáneamente la actividad de tk, mientras se mantiene aún la actividad del gen para *PhleoLacZ*. En la serie experimental, esta frecuencia se elevó aproximadamente 20 a 30 veces, de acuerdo con el primer grupo de experimentos (Tabla 1).

Se analizó la estructura molecular de los clones Phleo^R β -Gal⁺GlS^R por medio de transferencias tipo Southern (Fig. 27). Se analizaron cuatro clones de NIH3T3/G-MtkPL I, dos de la serie experimental y dos de la de control. Se digirió su ADN con *Pst* I endonucleasa. Si había ocurrido un evento intracromosómico, esperamos un fragmento único de *Pst* I de 13,6 kpb (que es la suma de los tres fragmentos de *Pst* I del ADN progenitor menos el fragmento de I-Sce I, ver la Fig. 27a). Todos los cuatro clones resistentes a Phleo^RGlS^R exhibieron este fragmento de *Pst* I de 13,6 kpb, sugiriendo una recombinación intramolecular fiel (Fig. 27b).

Se analizó el ADN de ocho clones de células NIH3T3/G-MtkPL 2 por medio de transferencias tipo Southern utilizando digestión con *Bcl* I (seis de la serie experimental y dos de la de control). La digestión con *Bcl* I del ADN progenitor resulta en un fragmento de 7,2 k pb que contiene las secuencias provirales y en dos fragmentos de flaqueo de 6 kpb y 9,2 kpb. Una recombinación intracromosómica debe resultar en la pérdida del fragmento de 7,2 kpb dejando las otras dos bandas de 6 kpb y 9,2 kpb sin cambio (Fig. 27a). Los ocho clones (2,7 hasta 2,16) mostraron la

desaparición del tk que contiene un fragmento de 7,2 kpb indicativa de una recombinación intracromosómica entre las dos LTR (Fig. 27c).

Discusión

Los resultados presentados aquí demuestran que la I-Sce I endonucleasa de levadura induce recombinación cromosómica en células de mamífero. Esto sugiere fuertemente que I-Sce I es capaz de cortar *in vivo* un cromosoma en un punto predeterminado.

Los rompimientos bicatenarios en secuencias genómicas de diferentes especies estimulan recombinación (21B, 19B). En la levadura diploide, un DSB cromosómico puede conducir al uso del locus homoalélico como una matriz de reparación. Esto resulta en un evento de conversión del gen, convirtiéndose entonces el locus en homocigoto (30B). Los DSB cromosómicos también se pueden reparar por medio del uso de secuencias homólogas de un locus ectópico como matriz (32B). Este resultado se observa en un nivel significativo como consecuencia de un mecanismo de reparación de la brecha de DSB. Si los DSB se presentan entre dos secuencias cromosómicas repetidas en forma directa, el mecanismo de recombinación utiliza la ruta de apareamiento monocatenario (SSA) (11B, 10B). La ruta SSA involucra tres etapas: 1) una exonucleólisis iniciada en el punto del rompimiento dejando los ADN monocatenarios 3' que sobresalen; 2) un apareamiento de los dos ADN monocatenarios por medio de sus secuencias homólogas, 3) una reparación del ADN por medio de genes reparadores complejos y mutadores que resuelven las secuencias no homólogas (33B). Un caso especial se refiere a la levadura haploide para la cual se ha demostrado que los DSB inducidos por HO o I-Sce I endonucleasas en un cromosoma conduce a la reparación del rompimiento por medio de uniones de los extremos (34B). Esto se presenta, pero con una eficiencia baja (30B, 35B).

Nuestros resultados muestran que la presencia de dos sitios de I-Sce I en un objetivo proviral y la expresión de la I-Sce I endonucleasa conduce a un incremento en la supresión de un gen para timidina quinasa con una frecuencia al menos 100 veces superior que aquella que se presenta espontáneamente. Dos tipos de clones a los que se les ha suprimido tk surgen de la recombinación mediada por I-Sce I: clones que han conservado (7%) y clones que han perdido (93%) las secuencias de *PhleoLacZ*.

La generación de células *tk⁻PhleoLacZ⁺* es probablemente la consecuencia de recombinación intracromosómica. Hay estudios que han mostrado que en un provirus recombinante con un sitio de reconocimiento de I-Sce I en las LTR, la I-Sce I endonucleasa conduce en un 20% de los casos a la escisión únicamente de un sitio de I-Sce I proviral y en un 80% a la escisión de los dos sitios de I-Sce I provirales. Si únicamente se corta uno de los dos sitios de I-Sce I por medio de la endonucleasa, puede presentarse una recombinación intracromosómica por medio de la ruta SSA. Si se cortan los dos sitios de I-Sce I, se pueden generar las células *tk⁻PhleoLacZ⁺* por medio de unión por los extremos, permitiendo recombinación intracromosómica (ver Figura 1). Aunque, en la levadura diploide, esta ruta no es favorable (se repara el rompimiento utilizando secuencias cromosómicas homólogas) (2B), sigue siendo posible que esta ruta sea utilizada en células de mamífero.

La generación de células *tk⁻/PhleoLacZ⁻* es probablemente una consecuencia ya sea de un evento de conversión de gens ectópicos y/o homoalélico (36B). El aislamiento y el análisis molecular detallado de los sitios de integración proviral suministrarán información sobre la frecuencia relativa de cada uno de estos eventos para la resolución de los DSB cromosómicos por parte de la célula. Esta información cuantitativa es importante ya que, en células de mamífero, la alta redundancia de secuencias genómicas aumenta la posibilidad de una reparación de los DSB por medio de secuencias homólogas ectópicas. La recombinación ectópica para la reparación de los DSB puede estar involucrada en la configuración y diversidad del genoma en evolución [29].

La habilidad para digerir específicamente un cromosoma en un sitio genómico predeterminado tiene varias aplicaciones potenciales para la manipulación del genoma.

El protocolo de reemplazo génico descrito aquí se puede variar de la siguiente manera:

Variedad de vectores donantes

Tamaño y secuencia de regiones de flaqueo del sitio de I-Sce-I en el plásmido donante (elaborado con 300 pb a la izquierda y 2,5 kb a la derecha): Existen diferentes construcciones con diferentes tamaños de las regiones de flaqueo hasta un total de 11 kb a la izquierda y a la derecha del sitio de I-Sce I. Las secuencias dependen de la construcción (LTR, gen). Se puede utilizar cualquier secuencia que contenga entre 300 pb hasta 11 kb.

- Insertos (se han construido neo, phleo, phleo-LacZ y Pytk-neo). Resistencia antibiótica: neomicina, fleomicina; gen reportero (LacZ); gen de la timidina quinasa HSV1: sensibilidad al ganciclovir. Es posible insertar cualquier clase de secuencia génica de hasta de 10 kb o reemplazarla. El gen se puede expresar bajo un promotor constitutivo o inducible del retrovirus, o por medio de una trampa génica y recombinación homóloga (es decir, Insulina, Hbs, ILs y diferentes proteínas).

Se pueden utilizar diferentes métodos para expresar la enzima I-Sce I: transfección transitoria (plásmido) o inyección directa de proteína (en núcleos embrionarios); transfección estable (diferentes promotores tales como: CMV,

ES 2 338 952 T3

RSV y MoMuLV); retrovirus recombinantes defectuosos (integración de ORF en un cromosoma bajo el promotor MoMuLV); y episomas.

Variación del rango de huéspedes para integrar el sitio de I-Sce I:

Se pueden producir retrovirus recombinantes que portan el sitio de I-Sce I (es decir, pG-MPL, pG-MtkPL, pG-MtkDPAPL) en diferentes empaquetamientos de líneas de células (anfotrópicas o xenotrópicas).

Construcción de líneas de células estables que expresan I-Sce I y protección de la célula contra infección retroviral

La línea de células estables que expresan I-Sce I se protegen contra la infección por medio de un vector retroviral que contiene el sitio de I-Sce I (es decir, la línea de células NIH 3T3 que produce I-Sce I endonucleasa bajo el control del promotor CMV es resistente a la infección por medio de pG-MPL o pGMtkPL o I-Sce I bajo el promotor MoMuLV en células Ψ 2).

Construcción de líneas de células y animales transgénicos que contienen el sitio de I-Sce I

La inserción del sitio de I-Sce I se lleva a cabo por medio de un reemplazo clásico de genes en el locus deseado y en la posición apropiada. Es posible entonces seleccionar la expresión de genes diferentes en la misma ubicación en la célula (inserción del gen donante en el sitio de I-Sce I artificialmente insertado) o en un animal transgénico. El efecto de múltiples medicamentos, ligandos, proteína médica, etc., se puede analizar en una forma específica para el tejido. El gen será consistentemente insertado en la misma posición en el cromosoma.

Para células de ratón “no preparadas”, y para todas las células eucariotas, se lleva cabo un procedimiento en una etapa de reemplazo/integración del gen de la siguiente manera:

- Vectores (diferentes plásmidos donantes) con sitio de I-Sce I: un sitio dentro del gen (o flanco) o dos sitios que flanquean al gen donante.

- Método para expresar la enzima

Expresión transitoria: ORF sobre el mismo plásmido o sobre otro (cotransfección).

Los detalles específicos relacionados con los métodos utilizados están descritos más arriba. Los siguientes detalles adicionales permiten la construcción de lo siguiente:

una línea de células capaz de producir un título alto de una variedad de partículas retrovirales infectivas;

un plásmido que contiene un retrovirus deficiente con sitios de I-Sce I, gen reportero-selector, LTR activas y otras secuencias retrovirales esenciales; un plásmido que contiene secuencias homólogas para flanquear regiones de sitios de I-Sce I en los retrovirus anteriores modificados por ingeniería genética y que contienen un sitio de clonación múltiple; y un vector que permite la expresión de la I-Sce I endonucleasa y adaptado a las aplicaciones específicas.

Se utilizó la línea de células Ψ2 de fibroblastos de ratón para producir vectores retrovirales recombinantes defectuosos ectópicos que contienen sitios de I-Sce I. También se encuentran disponibles líneas de células que producen plásmidos como pG-MPL, pG-MtkPL, pG-MtkDPAPL. Además, cualquiera de las células, como las líneas de células anfotrópicas de ratón (tales como PA12) o líneas de células xenotrópicas, que producen partículas infecciosas de título alto, pueden ser utilizadas para la producción de retrovirus recombinantes que portan el sitio de I-Sce I (es decir, pG-MPL, pG-MtkPL, pGMtk DPAPL) en diferentes líneas de células de empaquetamiento (anfotrópicas, ectrópicas o xenotrópicas).

Se puede utilizar una variedad de plásmidos que contienen I-Sce I en construcción retroviral, incluyendo pG-MPL, pG-MtkPL, y pG-MtkDPAPL. Se pueden construir otras clases de vectores plásmidos que contienen diferentes promotores, genes, sitio poliA, y sitio de I-Sce I. Se pueden construir una variedad de plásmidos que contienen secuencias homólogas a regiones de flanco de I-Sce I. El tamaño y la secuencia de las regiones de flanco del sitio I-Sce I en el plásmido donante se preparan de tal manera que 300 kb están a la izquierda y 2,5 kb están a la derecha. Se pueden utilizar otras construcciones con diferentes tamaños de regiones de flanco aproximadamente hasta de 11 kb a la izquierda y a la derecha del sitio de reconocimiento de I-Sce I.

Se han construido insertos que contienen neomicina, fleomicina y phleo-LacZ. Se pueden insertar otras secuencias tales como de resistencia a un medicamento o genes reporteros, incluyendo el gen de LacZ, HSV1 o de timidina quinasa (sensibilidad al ganciclovir), insulina, CFTR, IL2 y diferentes proteínas. Normalmente es posible insertar cualquier tipo de secuencia hasta de 12 kb, en donde el tamaño depende de la capacidad de encapsulación del virus). El gen se puede expresar bajo un promotor inducible o constitutivo del retrovirus, o por medio de una trampa génica después de recombinación homóloga.

ES 2 338 952 T3

Se pueden construir una variedad de plásmidos que contienen I-Sce I, que producen la endonucleasa. Se pueden introducir vectores de expresión tales como pCMVI-SceI(+) o construcciones similares que contienen el ORF, en células por medio de transfección transitoria, electroporación o lipofección. También se puede introducir directamente la proteína dentro de la célula por medio de inyección de liposomas.

Se pueden producir una variedad de líneas de células con sitios integrados de I-Sce I. Preferiblemente, la inserción del retrovirus (integración proviral) induce duplicación de la LTR que contiene el sitio de I-Sce I. La célula será hemicigota para el sitio. Las líneas apropiadas de células incluyen:

1. Línea de células Fibroblásticas de Ratón, NIH 3T3 con integración proviral de 1 a 14 de G-MPL. Se recuperaron múltiples clones (más de 30). La presencia y multiplicidad de las diferentes integraciones genómicas (no caracterizadas) fueron verificadas por medio de análisis molecular.
2. Línea de células Fibroblásticas de Ratón, NIH 3T3 con 1 copia de G-MtkPL integrado en el genoma. 4 clones fueron cubiertos.
3. Línea de células de Carcinoma Embrionario de Ratón, PCC7-S con 1 a 4 copias de integración proviral de G-MPL en el genoma. 14 clones fueron cubiertos.
4. Línea de células de Carcinoma Embrionario de Ratón, PCC4 con 1 copia de G-MtkPL integrado en el genoma.
5. Línea de células Madre Embrionarias de Ratón D3 con 1 a 4 copias de G-MPL en una variedad de localizaciones genómicas (no caracterizadas). 4 clones fueron cubiertos.

Pueden elaborarse construcciones de otras líneas de células y animales transgénicos que contienen al sitio de I-Sce I por inserción del sitio de I-Sce I por medio de reemplazo clásico del gen en el locus deseado y en la posición apropiada. Cualquier tipo de líneas de células vegetales o animales, podría ser utilizada *a priori* para integrar sitios de I-Sce I en una variedad de localizaciones genómicas con líneas de células adaptadas.

Se puede utilizar la invención de las siguientes formas:

1. Inserción génica específica del sitio

Los métodos permiten la producción de un número ilimitado de líneas de células en las cuales se pueden insertar diferentes genes o mutantes de un gen dado en la localización predeterminada definida por la integración previa del sitio de I-Sce I. Tales líneas de células son útiles por lo tanto para procedimientos de selección, para fenotipos, ligandos, medicamentos y para expresión reproducible a un nivel muy alto de vectores retrovirales recombinantes si la línea de células es una línea de células transcomplementaria para producción de retrovirus.

Se pueden utilizar las células de ratón anteriores o equivalentes de otros vertebrados, incluido el hombre. Cualquier célula de una planta que pueda ser mantenida en cultivo puede ser utilizada también independientemente de si tiene la habilidad para regenerarse o no, o de si dan lugar o no a plantas fértiles. Los métodos pueden ser utilizados también con animales transgénicos.

2. Expresión génica específica del sitio

Se pueden utilizar también líneas de células similares para producir proteínas, metabolitos u otros compuestos de interés biológico o biotecnológico utilizando un transgén, una variedad de promotores, reguladores y/o de genes estructurales. El gen será siempre insertado en la misma posición en el cromosoma. En animales transgénicos, hace posible analizar el efecto de múltiples fármacos, ligandos, o proteínas médicas en una forma específica del tejido.

3. Inserción del sitio de reconocimiento de I-Sce I en el locus de CFTR utilizando secuencias homólogas que flanquean el gen de CFTR en el ADN genómico. Se puede insertar el sitio de I-Sce I por reemplazo génico espontáneo por medio de cruzamiento doble (Le Mouellic y colaboradores, PNAS, 1990, Vol. 87, 4712-4716).

4. Aplicaciones biomédicas

A. En terapia génica, se pueden infectar las células de un paciente con un retrovirus que contenga I-Sce I, seleccionado para integración del retrovirus defectuoso y luego transformado conjuntamente con el vector para producción de I-Sce I y la secuencia del donante. Los ejemplos de células apropiadas incluyen tejido hematopoyético, hepatocitos, células de la piel, células endoteliales de vasos sanguíneos o cualquier célula madre, con tal de que dichas células no sean células embrionarias humanas.

Los retrovirus que contienen I-Sce I incluyen pG-MPL, pG-MtkPL o cualquier clase de vector retroviral que contenga al menos un sitio de I-Sce I.

Los vectores que producen I-Sce I incluyen pCMVI-Sce I(+) o cualquier plásmido que permita la expresión transitoria de I-Sce I endonucleasa.

Las secuencias donantes incluyen (a) Secuencias genómicas que contienen al gen de IL2 completo; (b) Secuencia genómicas que contienen al gen para pre-ProInsulina; (c) Un fragmento grande de los vertebrados, incluidos los humanos, secuencia genómica que contiene elementos que actúan en forma cis para expresión génica. Se reintroducen luego células modificadas en el paciente de acuerdo con protocolos establecidos para terapia génica.

B. Inserción de un promotor (es decir, CMV) con el sitio de I-Sce I, en una célula madre (es decir, linfóide). Se puede insertar secuencia abajo una molécula para reparación de una brecha que contiene un enlazador (sitio de multiclonación) entre el promotor CMV y la secuencia. La inserción de un gen (es decir, un gen para IL-2), presente en los plásmidos donantes, se puede hacer en forma eficiente por medio de la expresión de la I-Sce I meganucleasa (es decir, cotransfección con un vector de expresión de I-Sce I meganucleasa). La inserción directa del gen para IL-2 bajo el promotor CMV conduce a la selección directa de una célula madre que sobreexpresa IL-2.

Para la construcción de líneas de células transgénicas, se utiliza infección retroviral en sistemas actualmente disponibles. Se puede utilizar otro método para introducir sitios de I-Sce I dentro de genomas, incluyendo microinyección de ADN, transfección inducida por fosfato de Ca, electroporación, lipofección, protoplasto o fusión celular, y conjugación célula-bacteria.

La pérdida de heterozigosidad se demuestra de la siguiente manera: Se introduce el sitio de I-Sce I en un locus (con o sin secuencias foráneas), creando una inserción heterocigota en la célula. En ausencia de ADN para reparación, el rompimiento bicatenario inducido se extenderá por medio de exonucleasas no específicas, y se reparará la brecha por medio de la secuencia intacta de la cromátida hermana, de tal manera que la célula se vuelva homocigótica en este locus.

Los ejemplos específicos de terapia génica incluyen inmunomodulación (es decir, cambiar el rango o expresión de los genes para IL); reemplazo de genes defectuosos; y excreción de proteínas (es decir, expresión de diferentes proteínas secretoras en organelos).

Es posible activar un gen específico *in vivo* por medio de recombinación inducida por I-Sce I. El sitio de escisión de I-Sce I se introduce entre una duplicación de un gen en repeticiones en tándem, creando una pérdida de función. La expresión de la I-Sce I endonucleasa induce la escisión entre las dos copias. Se estimula la reparación por recombinación y resulta en un gen funcional.

Macrorreordenamientos genéticos dirigidos al sitio de cromosomas en líneas de células o en organismos

Se puede inducir translocación específica de cromosomas o supresión por escisión con I-Sce I. Se puede obtener inserción del locus por integración de uno en una ubicación específica en el cromosoma por medio de "reemplazo génico clásico". La escisión de la secuencia de reconocimiento por medio de la I-Sce I endonucleasa puede ser reparada por medio de translocaciones no letales o por medio de supresión seguida por unión de los extremos. También se podría obtener una supresión de un fragmento de cromosoma por inserción de dos o más sitios de I-Sce I en regiones de flanco de un locus (ver la figura 32). Se puede reparar la escisión por medio de recombinación y resulta en la supresión de la región completa entre los dos sitios (ver la figura 32).

Referencias

1. **Bernstein, N., Pennell, N., Ottaway, C. A. y Shulman, M.J.** 1992. Gene replacement with one-sided homologous recombination. *Mol. Cell Biol.* 12: 360-367.
2. **Bonnerot, C., Legouy, E., Choulaka, A. y Nicolas, J.-F.** 1992. Capture of a cellular transcriptional unit by retrovirus: mode of provirus activation in embryonal carcinoma cells. *J. Virol.* 66: 4982-4991.
3. **Bonnerot, C., y Nicolas, J.-F.** 1993. Application of LacZ gene fusions to post-implantation development. En "Methods in Enzymology: Guide to techniques in mouse development". 451-469. **Wassarman, P. M., DePamphilis, M. L.**
4. **Brenner, D. A., Smogocki, A. y Camerini-Otero, R. D.** 1986. Double-strand gap repair results in homologous recombination in mouse L cells. *Proc. Natl. Acad. Sci. USA.* 83: 1762-1766.
5. **Capecchi, M. R.** 1989. Altering the genome by homologous recombination. *Science.* 244: 1288-1292.
6. **Jacquier, A. y Dujon, B.** 1985. An intron encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. *Cell.* 41: 383-394

7. **Jakob, H. y Nicolas, J. F.** 1987. Mouse Teratocarcinoma Cells. En "Methods in Enzymology: Cell lines for genetic analysis". 66-81. Gottesman, M.H., ed., *Academic Press*.
8. **Jessberger, R. y Berg, P.** 1991. Repair of deletions and double-strand gaps by homologous recombination in a mammalian *in vitro* system. *Mol. Cell Biol.* 11: 445-457.
9. **Kilby, N. J., Snaith, M. R. y Murray, J. A. H.** 1993. Site-specific recombinases: tools for genome engineering. *Reviews.* 9: 413-421
10. **Lin, F. L. M., Sperle, K. y Sternberg N.** 1990. Repair of double-stranded DNA breaks by homologous DNA fragments during transfer of DNA into mouse L cells. *Mol. Cell Biol.* 10:113-119.
11. **Lin, F. L. M., Sperle, K. y Sternberg N.** 1990. Intermolecular recombination between DNAs introduced into mouse L cells is mediated by a nonconservative pathway that leads to crossover products. *Mol. Cell Biol.* 10: 103-112.
12. **Lin, F. L. M., Sperle, K. y Sternberg N.** 1990. Intermolecular recombination between DNAs introduced into mouse L cells is mediated by a nonconservative pathway that leads to crossover products. *Mol. Cell Biol.* 10: 103-112.
13. **Mann, R., Mulligan, R. C. y Baltimore, D.** 1983. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell.* 33: 153-160.
14. **Mansour, S. L., Thomas K. R. and Capecchi, M. R.** 1988. Disruption of the proto-oncogene int-2 in mouse embryoderived stem cells: a general strategy for targeting mutations to non-selectable gene. *Nature.* 336: 348-352.
15. **Mulsant, P., Gatignol, A., Dolens, M. y Tiraby, G.** 1988. *Somat. Cell. Mol. Genet.* 14: 243-252.
16. **Nicolas, J. F. and Rubenstein, J.** 1987. Retroviral vectors. Boston London Durban Singapore Sydney Toronto Wellington, Butterworths.
17. **Plessis, A., Perrin, A., Haber, J. E. y Dujon, B.** 1992. Site specific recombination determined by I-Sce I, a mitochondrial group I intron-encoded endonuclease expressed in the yeast nucleus. *Genetics* 130: 451-460
18. **Sauer, B. y Henderson, N.** 1988. Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Prac. Natl. Acad. Sci. USA.* 85: 5166-5170.
19. **Seidman, M. M.** 1987. Intermolecular homologous recombination between transfected sequences in mammalian cells is primarily nonconservative. *Mol. Cell Biol.* 7: 3561-3565.
20. **Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A. y Kucherlapati, R. S.** 1985. Insertion of DNA sequences into the human chromosomal B-globin locus by homologous recombination. *Nature.* 317: 230-234.
21. **Szostak, J. W., Orr-Weaver, T. L. y Rothstein, R. J.** 1983. The double-strand break repair model for recombination. *Cell.* 33: 25-35.
22. **Thierry, A., Perrin, A., Boyer, J., Fairhead, C., Dujon, B., Frey, B. y Schmitz, G.** 1991. Cleavage of yeast and bacteriophage T7 genomes at a single site using the rare cutter endonuclease I-Sce I. *Nucleic Acids Res.* 19: 189-90
23. **Tybulewicz, V. L. J., Crawford, C. E., Jackson, P. K., Bronson, R. T. y Mulligan, R. C.** 1991. Neonatal Lethality and Lymphopenia in Mice with a Homozygous Disruption of the c-abl Proto-Oncogene. *Cell* 65: 1153-1163
24. **Varmus, H. y Brown, P.** 1989. Retroviruses
25. **Weiss, R., Teich, N., Varmus, H. y Coffin, J.** 1985. RNA tumor viruses. Molecular Biology of tumor viruses. Second Edition. 2) Supplements and appendixes. *Cold Spring Harbor Laboratory.* 1-1222.
26. **Weiss, R., Teich, N., Varmus, H. y Coffin, J.** 1985. RNA tumor viruses. Molecular Biology of tumor viruses. Second Edition. 2) Supplements and appendixes. *Cold Spring Harbor Laboratory.* 1-1222.
27. **Phillips J. y Morgan W.** 1994. Illegitimate recombination induced by DNA double-strand breaks in mammalian chromosomes. *Molecular and Cellular Biology* 14: 5794-5803.
28. **Dujon B.** 1989. Group I introns are mobile genetic elements: facts and mechanistic speculations-a review. *Gene* 82:91-114.
29. **Colleaux L., D'Aurio L., Galibert F. y Dujon B.** 1988. Recognition and cleavage site of the intron-encoded omega transposase. *Proc Natl Acad Sci USA* 85: 6022-6.

30. Fairchild C. y Dujon B. Consequences of unique double-stranded breaks in yeast chromosomes: death or homozygosis. *Molecular general genetics* 240: 170-180.

31. Pfeiffer P., Thode S., Hancke J. y Vielmetter W. 1994. Mechanism of overlap information in nonhomologous DNA end joining. *Molecular and Cellular Biology* 14: 888-895.

32. Mezard C. y Nicholas A. 1994. Homologous, homeologous, and illegitimate repair of double-strand breaks during transformation of a wild-type strain and a rad52 Mutant strain of *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* 14: 1278-1292.

33. Feaver W. J., Svejstrup J. Q., Bradwell L., Bradwell A. J., Buratowski S., Gulyas K., Donahue T. F., Friedberg E. C. y Kornberg R. D. 1993. Dual Roles of a Multiprotein Complex from *S. cerevisiae* in transcription and DNA Repair. *Cell* 75: 1379-1387.

34. Kramer K., Brock J., Bloom K., Moore K. y Haber J. 1994. Two different types of double-strand breaks in *Saccharomyces cerevisiae* are repaired by similar RAD52 independent, nonhomologous recombination events. *Molecular and Cellular Biology* 14: 1293-1301.

35. Weiffenbach B. y Haber J. 1981. Homothallic mating type switching generates lethal chromosomes breaks in rad52 strains of *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* 1: 522-534.

36. Nassif N., Penney J., Pal S., Engels W. y Gloor G. 1994. Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Molecular and cellular biology* 14: 1643-1625.

37. Charlesworth B., Sniegowski P. y Stephan W. 1994. The evolutionary dynamics of repetitive DNA in eucaryotes. *Nature* 371: 215-220.

Referencias citadas en la descripción

Este listado de referencias citado por el solicitante es únicamente para conveniencia del lector. No forma parte del documento europeo de la patente. Aunque se ha tenido gran cuidado en la recopilación, no se pueden excluir los errores o las omisiones y la OEP rechaza toda responsabilidad en este sentido.

Documentos de patente citados en la descripción

- US 4683202 A [0043]
- US 4683195 A [0043]

Literatura citada en la descripción que no es de patente

- Old; Primrose. Principles of Gene Manipulation. *Blackwell Scientific Publications*, 1989, 255 [0005]
- YEAST, 1990, vol. 6, 521-534 [0021]
- Cell, 1986, vol. 44, 521-533 [0035]
- S. Kwok y colaboradores. *J. Virol.*, 1987, vol. 61, 1690-1694 [0043]
- R. K. Saiki y colaboradores. *Bio/Technology*, 1985, vol. 3, 1008-1012 [0043]
- Sargueil y colaboradores. *NAR*, 1990, vol. 18, 5659-5665 [0045]
- Sargueil y colaboradores. *MGG*, 1991, vol. 225, 340-341 [0045]
- Marshall, Lemieux. *Gene*, 1991, vol. 104, 241-245 [0045]
- Muscarella y colaboradores. *MCB*, 1990, vol. 10, 3386-3396 [0045]
- Chu y colaboradores. *PNAS*, 1990, vol. 87, 3574-3578 [0045]
- Pedersen y colaboradores. *NAR*, 1990, vol. 18, 3763-3770 [0045]
- Bell-Pedersen y colaboradores. *NAR*, 1990, vol. 18, 3763-3770 [0045]
- Eddy, Gold. *Genes Dev.*, 1991, vol. 5, 1032-1041 [0045]
- Nickoloff y colaboradores. *MCB*, 1990, vol. 10, 1174-1179 [0045]

- **Kawasaki** y colaboradores. *JBC*, 1991, vol. 266, 5342-5347 [0045]
- **Maniatis** y colaboradores. Molecular Cloning: A Laboratory Manual. *Cold Spring Harbor Laboratory*, 1982 [0052]
- **Thierry** y colaboradores. *YEAST*, 1990, vol. 6, 521-534 [0080]
- **B. Dujon**. Sequence of the intron and flanking exons of the mitochondrial 21 S rRNA gene of yeast strains having different alleles at the w and RIB 1 loci. *Cell*, 1980, vol. 20, 185-187 [0107]
- **F. Michel; A. Jacquier; B. Dujon**. Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure. *Biochimie*, 1982, vol. 64, 867-881 [0107]
- **F. Michel; B. Dujon**. Conservation of RNA secondary structures in two intron families including mitochondrial-, chloroplast-, and nuclear-encoded members. The *EMBO Journal*, 1983, vol. 2, 33-38 [0107]
- **A. Jacquier; B. Dujon**. The intron of the mitochondrial 21S rRNA gene: distribution in different yeast species and sequence comparison between *Kluyveromyces thermotolerans* and *Saccharomyces cerevisiae*. *Mol. Gen. Genet.*, 1983, vol. 192, 487-499 [0107]
- Organization of the mitochondrial 21S rRNA gene in *Saccharomyces cerevisiae*: mutants of the peptidyl transferase centre and nature of the omega locus. **B. Dujon; A. Jacquier**. Mitochondria 1983. *Walter de Gruyter et Co*, 1983, 389-403 [0107]
- **A. Jacquier; B. Dujon**. An intron encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. *Cell*, 1985, vol. 41, 383-394 [0107]
- Mechanism of integration of an intron within a mitochondrial gene: a double strand break and the transposase function of an intron encoded protein as revealed by *in vivo* and *in vitro* assays. **B. Dujon; G. Cottarel; L. Colleaux; M. Betermier; A. Jacquier; L. D'Auriol; F. Galibert** y colaboradores. Achievements and perspectives of Mitochondrial Research. *Elsevier*, 1985, vol. II, 215-225 [0107]
- **L. Colleaux; L. D'Auriol; M. Betermier; G. Cottarel; A. Jacquier; F. Galibert; B. Dujon**. A universal code equivalent of a yeast mitochondrial intron reading frame is expressed into *Escherichia coli* as a specific double strand endonuclease. *Cell*, 1986, vol. 44, 521-533 [0107]
- Mitochondrial introns as mobile genetic elements: the role of intron-encoded proteins. **B. Dujon; L. Colleaux; A. Jacquier; F. Michel; C. Monteilhet** y colaboradores. Extrachromosomal elements in lower eucaryotes. *Plenum Pub. Corp*, 1986, 5-27 [0107]
- **L. Colleaux; L. D'Auriol; F. Galibert; B. Dujon**. Recognition and cleavage site of the intron encoded omega transposase. *PNAS*, 1988, vol. 85, 6022-6026 [0107]
- **B. Dujon**. Group I introns as mobile genetic elements, facts and mechanistic speculations: A Review. *Gene*, 1989, vol. 82, 91-114 [0107]
- **B. Dujon; M. Belfort; R. A. Butow; C. Jacq; C. Lemieux; P.S. Perlman; V. M. Vogt**. Mobile introns: definition of terms and recommended nomenclature. *Gene*, 1989, vol. 82, 115-118 [0107]
- **C. Monteilhet; A. Perrin; A. Thierry; L. Colleaux; B. Dujon**. Purification and Characterization of the *in vitro* activity of I-SceI, a novel and highly specific endonuclease encoded by a group I intron. *Nucleic Acid Research*, 1990, vol. 18, 1407-1413 [0107]
- **L. Colleaux; M-R. Michel-Wolwertz; R. F. Matagne; B. Dujon**. The apocytochrome b gene of *Chlamydomonas smithii* contains a mobile intron related to both *Saccharomyces* and *Neurospora* introns. *Mol. Gen. Genet.*, 1990, vol. 223, 288-296 [0107]
- Des introns autonomes et mobiles. **B. Dujon**. *Annales de l'Institut Pasteur/Actualites*. 1990, vol. 1, 181-194 [0107]
- **A. Thierry; A. Perrin; J. Boyer; C. Fairhead; B. Dujon; B. Prey; G. Schmitz**. Cleavage of yeast and bacteriophage 17 genomes at a single site using the rare cutter endonuclease I-Sce. *I Nuc. Ac. Res.*, 1991, vol. 19, 189-190 [0107]
- **A. Plessis; A. Perrin; J. E. Haber; B. Dujon**. Site specific recombination determined by I-SceI, a mitochondrial intron-encoded endonuclease expressed in the yeast nucleus. *GENETICS*, 1992, vol. 130, 451-460 [0107]

• A. **Jacquier**; B. **Dujon**. Intron recombinational insertion at the DNA level: Nature of a specific receptor site and direct role of an intron encoded protein. *Cold Spring Harbor Symposium*, 1984 [0107]

• I. **Colleaux**; L. **D'Auriol**; M. **Demariaux**; B. **Dujon**; F. **Galibert**; A. **Jacquier**. Construction of a universal code equivalent from a mitochondrial intron encoded transposase gene using oligonucleotide directed multiple mutagenesis. Colloque International de DNRS "oligonucleotids et Genetique Moleculaire", 08 January 1985 [0108]

• L. **Colleaux**; D'Auriol; M. **Demariaux**; B. **Dujon**; F. **Galibert**; A. **Jacquier**. Expression in *E. coli* of a universal code equivalent of a yeast mitochondrial intron reading frame involved in the integration of an intron within a gene. *Cold Spring Harbor Meeting on "Molecular Biology of Yeast"*, 13 August 1985 [0108]

• B. **Dujon**; G. **Cottarel**; L. **Colleaux**; M. **Demariaux**; A. **Jacquier**; L. **D'Auriol**; F. **Galibert**.

Mechanism of integration of an intron within a mitochondrial gene: a double strand break and the "transposase" function of an intron encoded protein as revealed by *in vivo* and *in vitro* assays. International symposium on "Achievements and Perspectives in Mitochondrial Research", 26 September 1985 [0108]

• L. **Colleaux**; G. **Cottarel**; M. **Betermier**; A. **Jacquier**; B. **Dujon**; L. **D'auriol**; F. **Galibert**. Mise en evidence de l'activite endonuclease double brin d'unc protein codee par un intron mitochondrial de levure. Forum sur la Biologie Moleculaire de la levure, 02 October 1985 [0108]

• B. **Dujon**; L. **Colleaux**; F. **Michel**; A. **Jacquier**. Mitochondrial introns as mobile genetic elements. Extrachromosomal elements in lower eucaryotes, 01 June 1986 [0108]

• L. **Colleaux**; B. **Dujon**. Activity of a mitochondrial intron encoded transposase. Yeast Genetics and Molecular Biology Meeting, 03 June 1986 [0108]

• L. **Colleaux**; B. **Dujon**. The role of a mitochondrial intron encoded protein. XIIIth International Conference on Yeast Genetics and Molecular Biology, 31 August 1986 [0108]

• L. **Colleaux**; L. **D'Aurio**; F. **Galibert**; B. **Dujon**. Recognition and cleavage specificity of an intron encoded transposase. 1987 Meeting on Yeast Genetics and Molecular Biology, 16 June 1987 [0108]

• Biochemical activity of an intron encoded transposase of *Saccharomyces cerevisiae*. A. **Perrin**; C. **Monteilhet**; L. **Colleaux**; B. **Dujon**. Cold Spring Harbor Meeting on "Molecular Biology of Mitochondria and chloroplasts". *Cold Spring Harbor*, 25 August 1987 [0108]

• B. **Dujon**; A. **Jacquier**; L. **Colleaux**; C. **Monteilhet**; A. **Perrin**. "Les Introns autoeppissables et leurs proteins" Colloque. Biologie Moleculaire de la levure: expression genetique chez Saccharomyces, 18 January 1988 [0108]

• L. **Colleaux**; L. **D'Auriol**; C. **Monteilhet**; F. **Galibert**; B. **Dujon**. Characterization of the biochemical activity of an intron encoded transposase. 14th International Conference on Yeast Genetics and Molecular Biology, 07 August 1988 [0108]

• B. **Dujon**. A group I intron as a mobile genetic element, Albany Conference. RNA: catalysis, splicing, evolution, 22 September 1988 [0108]

• B. **Dujon**; L. **Colleaux**; C. **Monteilhet**; A. **Perrin**; L. **D'Auriol**; F. **Galibert**. Group I introns as mobile genetic elements: the role of intron encoded proteins and the nature of the target site. 14th Annual *EMBO* Symposium. "Organelle genomes and the nucleus" Heidelberg, 26 September 1988 [0108]

• L. **Colleaux**; R. **Matagne**; B. **Dujon**. A new mobile mitochondrial intron provides evidence for genetic exchange between *Neurospora* and *Chlamydomonas* species. *Cold Spring Harbor*, May 1989 [0108]

• L. **Colleaux**; M. R. **Michel-Wolwertz**; R. F. **Matagne**; B. **Dujon**. The apoxytochrome b gene of *Chlamydomonas smithii* contains a mobile intron related to both *Saccharomyces* and *Neurospora* introns. Fourth International Conference on Cell and Molecular Biology of *Chlamydomonas*, April 1990 [0108]

• B. **Dujon**; L. **Colleaux**; E. **Luzi**; C. **Monteilhet**; A. **Perrin**; A. **Plessis**; I. **Stroke**; A. **Thierry**. Mobile Introns. *EMBO Workshop on "Molecular Mechanisms of transposition and its control"*, June 1990 [0108]

• A. **Perrin**; C. **Monteilhet**; A. **Thierry**; E. **Luzi**; I. **Stroke**; L. **Colleaux**; B. **Dujon**. I-SceI, a novel double strand site specific endonuclease, encoded by a mobile group I intron in Yeast. Workshop on "RecA and Related Proteins", 17 September 1990 [0108]

• A. **Plessis**; A. **Perrin**; B. **Dujon**. Site specific recombination induced by double strand endonucleases, HO and I-Sce I in yeast. Workshop on "RecA and Related Proteins", 17 September 1990 [0108]

- **B. Dujon**. The genetic propagation of introns 20th *FEBS* Meeting, August 1990 [0108]
- **E. Luzi; B. Dujon**. Analysis of the intron encoded site specific endonuclease I-Sce I by mutagenesis. Third European Congress on Cell Biology, September 1990 [0108]
- **B. Dujon**. Self splicing introns as contagious genetic elements. Journees Franco-Beiges de Pont a Mousson, October 1990 [0108]
- **B. Frey; H. Dubler; G. Schmitz; A. Thierry; A. Perrin; J. Boyer; C. Fairhead; B. Dujon**. Specific cleavage of the yeast genome at a single site using the rare cutter endonuclease I-Sce I Human Genome, November 1990 [0108]
- **B. Dujon; A. Perrin; I. Stroke; E. Luzi; L. Colleaux; A. Plessis; A. Thierry**. The genetic mobility of group I introns at the DNA level. Keystone Symposia Meeting. Molecular Evolution of Introns and Other RNA elements, 02 February 1991 [0108]
- **B. Dujon; J. Boyer; C. Fairhead; A. Perrin; A. Thierry**. Cartographie chez la levure. Reunion "Strategies d'établissement des cartes geniques" Toulouse, 30 May 1991 [0108]
- **B. Dujon; A. Thierry**. Nested chromosomal fragmentation using the meganuclease I-SceI: a new method for the rapid mapping of the yeast genome, 17 May 1991 [0108]
- **A. Thierry; L. Gaillon; F. Galibert; B. Dujon**. The chromosome XI library: what has been accomplished, what is left. Brugge meeting, 22 September 1991 [0108]
- **B. Dujon; A. Thierry**. Nested chromosomal fragmentation using the meganuclease I-SceI: a new method for the rapid physical mapping of the eukaryotic genomes. *Cold Spring Harbor*, 06 May 1992 [0108]
- **A. Thierry; L. Gaillon; F. Galibert; B. Dujon**. Yeast chromosome XI: construction of a cosmid contig. a high resolution map and sequencing progress. *Cold Spring Harbor*, 06 May 1992 [0108]
- **A. Thierry; B. Dujon**. Nested Chromosomal Fragmentation Using the Meganuclease I-SceI: Application to the physical mapping of a yeast chromosome and the direct sorting of cosmid libraries. Probably Submission to GENOMICS or *EMBO J.* [0109]
- **A. Thierry; L. Colleaux; B. Dujon**. Construction and Expression of a synthetic gene coding for the meganuclease I-SceI. Possible submission: NAR, *EMBO J.* [0109]
- **I. Stroke; V. Pelicic; B. Dujon**. The evolutionarily conserved dodecapeptide motifs of intron-incoded I-Sce I are essential for endonuclease function. Submission to *EMBO J.* [0109]
- **C. Fairhead; B. Dujon**. Consequences of a double strand break induced *in vivo* in yeast at specific artificial sites, using the meganuclease I-SceI. Possible submission to *GENETICS, NATURE* [0109]
- **A. Perrin; B. Dujon**. Asymmetrical recognition by the I-Sce I endonuclease on exon and intron sequences reveals a new step in intron mobility. Possible submission: *NATURE* [0109]
- **Bernstein, N.; Pennell, N.; Ottaway, C. A.; Shulman, M. J.** Gene replacement with one-sided homologous recombination. *Mol. Cell Biol.*, 1992, vol. 12, 360-367 [0185]
- **Bonnerot, C.; Legouy, E.; Choulika, A.; Nicolas, J.-F.** Capture of a cellular transcriptional unit by retrovirus: mode of provirus activation in embryonal carcinoma cells. *J. Virol.*, 1992, vol. 66, 4982-4991 [0185]
- Application of LacZ gene fusions to post-implantation development. **Bonnerot, C.; Nicolas, J.-F.** Methods in Enzymology: Guide to techniques in mouse development. 1993, 451-469 [0185]
- **Brenner, D.A.; Smogocki, A.; Camerini-Otero, R.D.** Double-strand gap repair results in homologous recombination in mouse L cells. *Proc. Natl. Acad. Sci. USA.*, 1986, vol. 83, 1762-1766 [0185]
- **Capecchi, M.R.** Altering the genome by homologous recombination. *Science*, 1989, vol. 244, 1288-1292 [0185]
- **Jacquier, A.; Dujon, B.** An intron encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. *Cell*, 1985, vol. 41, 383-394 [0185]
- Mouse Tertocarcinoma Cells. **Jakob, H.; Nicolas, J. F.** Methods in Enzymology: Cell lines for genetic analysis. *Academic Press*, 1987, 66-81 [0185]
- **Jessberger, R.; Berg, P.** Repair of deletions and double-strand gaps by homologous recombination in a mammalian *in vitro* system. *Mol. Cell Biol.*, 1991, vol. 11, 445-457 [0185]

• **Kilby, N. J.; Snaith, M. R.; Murray, J. A. H.** Site-specific recombinases: tools for genome engineering. *Reviews*, 1993, vol. 9, 413-421 [0185]

• **Lin, F. L. M.; Sperle, K.; Sternberg N.** Repair of double-stranded DNA breaks by homologous DNA fragments during transfer of DNA into mouse L cells. *Mol. Cell Biol.*, 1990, vol. 10, 113-119 [0185]

• **Lin, F. L. M.; Sperle, K.; Sternberg N.** Intermolecular recombination between DNAs introduced into mouse L cells is mediated by a nonconservative pathway that leads to crossover products. *Mol. Cell Biol.*, 1990, vol. 10, 103-112 [0185]

• **Lin, F. L. M.; Sperle, K.; Sternberg N.** Intermolecular recombination between DNAs introduced into mouse L cells is mediated by a nonconservative pathway that leads to crossover products. *Mol. Cell. Biol.*, 1990, vol. 10, 103-112 [0185]

• **Mann, R.; Mulligan, R. C.; Baltimore, D.** Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell*, 1983, vol. 33, 153-160 [0185]

• **Mansour, S. L.; Thomas K. R.; Capecchi, M. R.** Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable gene. *Nature*, 1988, vol. 336, 348-352 [0185]

• **Mulsant, P.; Gatignol, A.; Dolens, M.; Tiraby, G. Somat.** *Cell. Mol. Genet.*, 1988, vol. 14, 243-252 [0185]

• **Nicolas, J. F.; Rubenstein, J.** *Retroviral vectors*, 1987 [0185]

• **Plessis, A.; Perrin, A.; Haber, J. E.; Dujon, B.** Site specific recombination determined by I-Sce I, a mitochondrial group I intron-encoded endonuclease expressed in the yeast nucleus. *Genetics*, 1992, vol. 130, 451-460 [0185]

• **Sauer, B.; Henderson, N.** Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proc. Natl. Acad. Sci. USA*, 1988, vol. 85, 5166-5170 [0185]

• **Seidman, M.M.** Intermolecular homologous recombination between transfected sequences in mammalian cells is primarily nonconservative. *Mol. Cell. Biol.*, 1987, vol. 7, 3561-3565 [0185]

• **Smithies, O.; Gregg, R. G.; Boggs, S. S.; Koralewski, M. A.; Kucherlapati, R. S.** Insertion of DNA sequences into the human chromosomal B-globin locus by homologous recombination. *Nature*, 1985, vol. 317, 230-234 [0185]

• **Szostak, J. W.; Orr-Weaver, T. L.; Rothstein, R. J.** The double-strand break repair model for recombination. *Cell*, 1983, vol. 33, 25-35 [0185]

• **Thierry, A.; Perrin, A.; Boyer, J.; Fairhead, C.; Dujon, B.; Frey, B.; Schmitz, G.** Cleavage of yeast and bacteriophage T7 genomes at a single site using the rare cutter endonuclease I-Sce I. *Nucleic Acids Res.*, 1991, vol. 19, 189-90 [0185]

• **Tybulewicz, V. L. J.; Crawford, C. E.; Jackson, P. K.; Bronson, R. T.; Mulligan, R. C.** Neonatal Lethality and Lymphopenia in Mice with a Homozygous Disruption of the c-abl Proto-Oncogene. *Cell*, 1991, vol. 65, 1153-1163 [0185]

• **Varmus, H.; Brown, P.** *Retroviruses*, 1989 [0185]

• RNA tumor viruses. Molecular Biology of tumor viruses. **Weiss, R.; Teich, N.; Varmus, H.; Coffin, J.** Supplements and appendixes. *Cold Spring Harbor Laboratory*, 1985, 1-1222 [0185]

• **Phillips J.; Morgan W.** Illegitimate recombination induced by DNA double-strand breaks in mammalian chromosomes. *Molecular and Cellular Biology*, 1994, vol. 14, 5794-5803 [0185]

• **Dujon B.** Group I introns are mobile genetic elements: facts and mechanistic speculations-a review. *Gene*, 1989, vol. 82, 91-114 [0185]

• **Colleaux L.; D'Aurio L.; Galibert F.; Dujon B.** Recognition and cleavage site of the intron-encoded omega transposase. *Proc Natl Acad Sci USA*, 1988, vol. 85, 6022-6 [0185]

• **Fairchild C.; Dujon B.** Consequences of unique double-stranded breaks in yeast chromosomes: death or homozygosis. *Molecular general genetics*, vol. 240, 170-180 [0185]

• **Pfeiffer P.; Thode S.; Hancke J.; Vielmetter W.** Mechanism of overlap information in nonhomologous DNA end joining. *Molecular and Cellular Biology*, 1994, vol. 14, 888-895 [0185]

• **Mezard C.; Nicholas A.** Homologous, homeologous, and illegitimate repair of double-strand breaks during transformation of a wild-type strain and a rad52 Mutant strain of *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, 1994, vol. 14, 1278-1292 [0185]

5 • **Feaver W. J.; Svejstrup J. Q.; Bradwell L.; Bradwell A. J.; Buratowski S.; Gulyas K.; Donahue T. F.; Friedberg E. C.; Kornberg R. D.** Dual Roles of a Multiprotein Complex from *S. cerevisiae* in transcription and DNA Repair. *Cell*, 1993, vol. 75, 1379-1387 [0185]

10 • **Kramer K.; Brock J.; Bloom K.; Moore K.; Haber J.** Two different types of double-strand breaks in *Saccharomyces cerevisiae* are repaired by similar RAD52 independent, nonhomologous recombination events. *Molecular and Cellular Biology*, 1994, vol. 14, 1293-1301 [0185]

15 • **Weiffenbach B.; Haber J.** Homothallelic mating type switching generates lethal chromosomes breaks in rad52 strains of *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, 1981, vol. 1, 522-534 [0185]

• **Nassif N.; Penney J.; Pal S.; Engels W.; Gloor G.** Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Molecular and cellular biology*, 1994, vol. 14, 1643-1625 [0185]

20 • **Charlesworth B.; Sniegowski P.; Stephan W.** The evolutionary dynamics of repetitive DNA in eucaryotes. *Nature*, 1994, vol. 371, 215-220 [0185].

REIVINDICACIONES

1. Un método para inducir al menos un rompimiento bicatenario dirigido al sitio en ADN cromosómico de células animales, comprendiendo dicho método

- (a) suministrar células animales que contengan ADN cromosómico bicatenario;
- (b) insertar al menos un sitio de restricción de I-Sce I en dicho ADN cromosómico;
- (c) la transfección de dichas células con al menos un plásmido que contiene ADN que codifica la I-Sce I meganucleasa; y
- (d) seleccionar células en las cuales se ha inducido al menos un rompimiento bicatenario.

2. El método de la reivindicación 1, en donde dichas células animales son células de mamífero.

3. El método de la reivindicación 2, en donde dicha célula es una célula NIH3T3 que contiene el virus G-MtkPL.

4. El método de la reivindicación 1, en donde dicho plásmido es pCMV (I-Sce I+)

5. Un método para inducir recombinación homóloga entre ADN cromosómico de células animales y ADN exógeno añadido a dicha célula, comprendiendo dicho método

- (a) suministrar células animales que contengan ADN cromosómico, en donde dicho ADN comprende al menos un sitio de restricción de I-SceI;
- (b) la transfección de dichas células con un plásmido que contiene ADN exógeno, y con un plásmido que contiene ADN que codifica la I-Sce I meganucleasa; y
- (c) seleccionar células en las cuales dicho ADN exógeno es insertado dentro de dicho ADN cromosómico.

6. El método de la reivindicación 5, en donde dichas células animales son células de mamífero.

7. El método de la reivindicación 6, en donde dicha célula es una célula NIH3T3 que contiene al virus G-MtkPL.

8. El método de la reivindicación 5, en donde el plásmido que contiene al ADN que codifica a la I-Sce I meganucleasa es pCMV (I-Sce I+).

9. Un método para inducir recombinación homóloga entre ADN cromosómico de una célula animal y ADN exógeno añadido a dicha célula, comprendiendo dicho método

- (a) suministrar células animales que contengan ADN cromosómico;
- (b) insertar al menos un sitio de restricción de I-Sce I en dicho ADN cromosómico;
- (c) la transfección de dichas células con un primer plásmido que contiene ADN exógeno, y con un segundo plásmido que contiene ADN que codifica a la I-Sce I meganucleasa; y
- (d) seleccionar células en las cuales dicho ADN exógeno es insertado dentro de dicho ADN cromosómico.

10. El método de la reivindicación 9, en donde dichas células animales son células de mamífero.

11. El método de la reivindicación 9, en donde dicho primer plásmido es pVRneo.

12. El método de la reivindicación 9, en donde dicho segundo plásmido es pCMV (I-Sce I+).

13. Un método para inducir al menos un rompimiento dirigido al sitio en ADN cromosómico de células animales e insertar el ADN que codifica un polipéptido, comprendiendo dicho método

- (a) suministrar células animales que contengan ADN cromosómico bicatenario, y transformar dichas células con un ADN que contenga un sitio de restricción de I-Sce I;
- (b) añadir enzima I-Sce I o transformar dicha célula con ADN que codifique a la enzima I-Sce I;

(c) la transfección de dichas células con un ADN que codifique un polipéptido o con un vector que contenga dicho ADN; y

(d) seleccionar células en las cuales dicho ADN que codifica dicho polipéptido sea insertado en ADN cromosómico, en donde dichas células expresen dicho polipéptido.

14. Una célula recombinante obtenible por medio del método de cualquiera de las reivindicaciones 1 y 13, con tal de que dicha célula no sea ni una célula germinal humana ni una célula embrionaria humana.

15. Un animal transgénico no humano que comprende una célula transformada por medio del método de cualquiera de las reivindicaciones 1 y 13.

16. Un método para expresar un polipéptido en un animal transgénico no humano, comprendiendo dicho método

(a) transformar células madre embrionarias de un animal no humano con un ADN que contiene un sitio de restricción de I-Sce I;

(b) añadir enzima I-Sce I a dicha célula o transformar dicha célula con un vector que contiene al gen que codifica para la enzima I-Sce I;

(c) transfectar dichas células con un ADN que codifica dicho polipéptido, y

(d) detectar la expresión de dicho polipéptido en un animal transgénico no humano que resulta de dichas células madre embrionarias transformadas en las cuales dicho ADN que codifica dicho polipéptido es insertado dentro de dicho ADN cromosómico.

17. Una célula madre recombinante que expresa un polipéptido, en donde dicha célula madre puede obtenerse por medio de

a. el suministro de células madre, en donde dichas células son transformadas por medio de un ADN que contiene un sitio de restricción de I-Sce I;

b. la adición de enzima I-Sce I a dicha célula o transformar dicha célula con un vector que contiene al gen que codifica para la enzima I-Sce I;

c. la transfección de dichas células con dicho ADN que codifica dicho polipéptido, y

d. la selección de células en las cuales dicho ADN que codifica dicho polipéptido es insertado dentro de dicho ADN cromosómico de dichas células, en donde dichas células expresan dicho polipéptido.

con tal de que dicha célula no sea una célula embrionaria humana.

18. Una célula recombinante como la reivindicada en la reivindicación 14, en donde dicho polipéptido es un antígeno foráneo a la célula.

19. Una célula recombinante como la reivindicada en la reivindicación 14, en donde dicha célula es una línea de células de mamífero.

20. Una célula recombinante como la reivindicada en la reivindicación 14, en donde dicha célula es una célula humana del tejido hematopoyético, un hepatocito, una célula de piel, una célula endotelial de vaso sanguíneo o una célula madre.

```

AAAAATAAAATCAT ATG AAA AAT ATT AAA AAA AAT CAA GTA ATG AAT CTG GGT CCT AAT TCT
      M   K   N   I   K   K   N   Q   V   H   N   L   G   P   N   S

AAA TTA TTA AAA GAA TAT AAA TCA CAA TTA ATT GAA TTA AAT ATT GAA CAA TTT GAA GCA
K   L   L   K   E   Y   K   S   Q   L   I   E   L   N   I   E   Q   F   E   A

GGT ATT GGT TTA ATT TTA GGA GAT GCT TAT ATT CGT AGT CGT GAT GAA GGT AAA ACT TAT
G   I   G   L   I   L   G   D   A   Y   I   R   S   R   D   E   G   K   T   Y

TGT ATG CAA TTT GAG TGG AAA AAT AAG GCA TAC ATG GAT CAT GTA TGT TTA TTA TAT CAT
C   H   Q   F   E   W   K   N   K   A   Y   H   D   H   V   C   L   L   Y   D

CAA TGG GTA TTA TCA CCT CCT CAT AAA AAA GAA AGA GTT AAT CAT TTA GGT AAT TTA GTA
Q   W   V   L   S   P   P   H   K   K   E   R   V   N   H   L   G   N   L   V

ATT ACC TGG GGA GCT CAA ACT TTT AAA CAT CAA GCT TTT AAT AAA TTA GCT AAC TTA TTT
I   T   W   G   A   Q   T   F   K   H   Q   A   F   N   K   L   A   N   L   F

ATT GTA AAT AAT AAA AAA CTT ATT CCT AAT AAT TTA GTT GAA AAT TAT TTA ACA CCT ATG
I   V   N   N   K   K   L   I   P   N   N   L   V   E   N   Y   L   T   P   H

AGT CTG GCA TAT TGG TTT ATG GAT GAT GGA GGT AAA TGG GAT TAT AAT AAA AAT TCT CTT
S   L   A   Y   W   F   H   D   Q   G   G   K   W   D   Y   N   K   N   S   L

AAT AAA AGT ATT GTA TTA AAT ACA CAA AGT TTT ACT TTT GAA GAA GTA GAA TAT TTA GTT
N   K   S   I   V   L   N   T   Q   S   F   T   F   E   E   V   E   Y   L   V

AAA GGT TTA AGA AAT AAA TTT CAA TTA AAT TGT TAT GTT AAA ATT AAT AAA AAT AAA CCA
K   G   L   R   N   K   F   Q   L   N   C   Y   V   K   I   N   K   N   K   P

ATT ATT TAT ATT GAT TCT ATG AGT TAT CTG ATT TTT TAT AAT TTA ATT AAA CCT TAT TTA
I   I   Y   I   D   S   H   S   Y   L   I   F   Y   N   L   I   K   P   Y   L

ATT CCT CAA ATG ATG TAT AAA CTG CCT AAT ACT ATT TCA TCC GAA ACT TTT TTA AAA TAA
I   P   Q   H   H   Y   K   L   P   N   T   I   S   S   E   T   F   L   K   .

```

FIGURA 1

El gen sintético para I-Sce I

Bam HI
 |
 CCGGATCCATG CAT ATG AAA AAC ATC AAA AAA AAC CAG GTA ATG AAC CTG GGT CCG AAC TCT
 M H M K N I K K N Q Y M N L G P N S
 AAA CTG CTG AAA GAA TAC AAA TCC CAG CTG ATC GAA CTG AAC ATC GAA CAG TTC GAA GCA
 K L L K E Y K S Q L I E L N I E Q F E A
 1 GGT ATC GGT CTG ATC CTG GGT [GAT] GCT TAC ATC CGT TCT CGT GAT GAA GGT AAA ACC TAC
 G I G L I L G [Q] A Y I R S R D E G K T Y
 TGT ATG CAG TTC GAG TGG AAA AAC AAA GCA TAC ATG GAC CAC GTA TGT CTG CTG TAC GAT
 C M Q F E W K N K A Y M D H V C L L Y D
 CAG TGG GTA CTG TCC CCG CCG CAC AAA AAA GAA CGT GTT AAC CAC CTG GGT AAC CTG GTA
 Q W V L S P P H K K E R V N H L G N L V
 ATC ACC TGG GGC GCC CAG ACT TTC AAA CAC CAA GCT TTC AAC AAA CTG GCT AAC CTG TTC
 I T W G A Q T F K H Q A F N K L A N L F
 ATC GTT AAC AAC AAA AAA ACC ATC CCG AAC AAC CTG GTT GAA AAC TAC CTG ACC CCG ATG
 I V N N K K T I P N N L V E N Y L T P M
 2 TCT CTG GCA TAC TGG TTC ATG [GAT GAT GGT] GGT AAA TGG GAT TAC AAC AAA AAC TCT ACC
 S L A Y W F H [D D G] G K W D Y N K N S T
 AAC AAA TCG ATC GTA CTG AAC ACC CAG TCT TTC ACT TTC GAA GAA GTA GAA TAC CTG GTT
 N K S I V L N T Q S F T F E E V E Y L V
 AAG GGT CTG CGT AAC AAA TTC CAA CTG AAC TGT TAC GTA AAA ATC AAC AAA AAC AAA CCG
 K G L R N K F Q L N C Y V K I N K N K P
 ATC ATC TAC ATC GAT TCT ATG TCT TAC CTG ATC TTC TAC AAC CTG ATC AAA CCG TAC CTG
 I I Y I D S M S Y L I F Y N L I K P Y L
 ATC CCG CAG ATG ATG TAC AAA CTG CCG AAC ACT ATC TCC TCC GAA ACT TTC CTG AAA TAA
 I P Q M M Y K L P N T I S S E T F L K -
 TAAGTCGACTGCAGGATCCGGTAAGTAAGTAA
 | | |
 Sall PstI BamHI

FIGURA 2

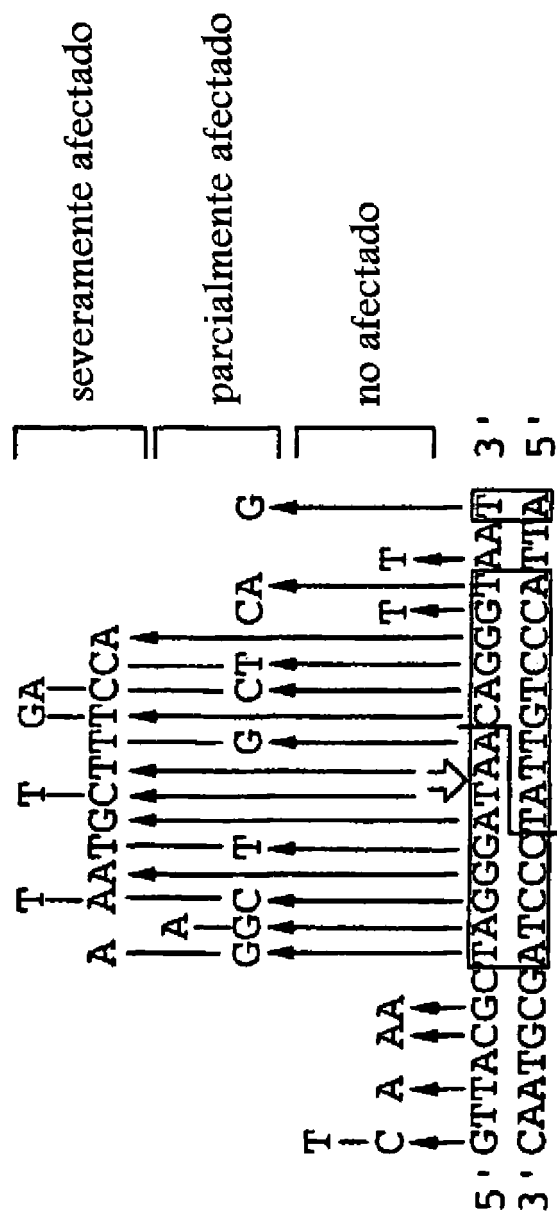


FIGURA 3

1667	CGCGACAGGTATCCGGTAAGCGCGAGGGTCGGAACAGAGAGCGGCAGCGAGCTTCACAGGGGGAACCGCTGGTATCT	1746
1747	TTATAGTCCTGTGCGGTTTCGCCACCTCTGACCTTGAGCGCTCGAATTTTTCG ATG CTC GTC AGG GCG GCG GAG	1818
1819	CCT ATG GAA AAA CGC CAG CAA CGC GGC CTT TTT ACG GTT CTT GGC CTT TTG CTG CCC TTT	1878
1879	TGC TCA CAT GTT TCC TGC GTT ATC CCC TGA TTCTGTGGTAACCGTATTACCGCCTTTGAGTCAGC	1947
1948	TGATACCGCTCGCCGACCGGACCGAGCGCGACGCGAGCTAGTCAGCGGAGGAGAGCGGCCCAATACGCAAC	2027
2028	CGCCTTCGCCGCGGCTTGCGCGAATTCATTA ATG CAG CTG GCA CGA CAG GTT TCC CGA CTG GAA AGC	2094
2095	GGG CAG TGA GCGCAAGCGCAATTA ATG TGA GTTACCTCAGCTCATTAGCGACCCCGAGCTTTACACTTT	2164
2165	CTT CCG GCT CGT ATG TTG TGT GGA ATT GTG AGC GGA TAA CAATTCACACAGGAACAGCT	2228
2229	ACC ATG ATT ACG AAT TCT CAT GTT TGA CAGCTTATCATCGATAAGCTTTA ATG CCG TAG TTTATCAC	2295
2296	AGTTAAATGTCTAAGCGGCTGAGCGACCGGTGT ATG AAA TCT AAC AAT GCG CTC ATC GTC ATC CTC GGC	2363
2364	ACC GTC ACC CTG GAT GCT GTA GGC ATA GGC TTG GTT ATG CCG GTA CTG CCG GGC CTC TTG	2423
2424	CGG GAT ATC CGC CTG ATG CGT GAA CGT GAC GGA CGT AAC CAC CGC GAC ATG TGT GTG CTG	2483
2484	TTC CGC TGG GCA TGC CAG GAC AAC TTC TGG TCC GGT AAC GTG CTG AGC CCG GCC AAG CTT	2543

FIGURA 4A

2544	ACT	CCG	GAT	CCC	GTT	GAC	ATT	TAA	TCATCGCTGTATATA	ATG	TGT	GCA	ATT	GTG	AGC	GGA	2606
2545	T	P	H	P	P	V	D	N	+	M	C	G	I	V	S	G	7
2501	TAA	CAATTTCACACACGAAACACGCTC	ATG	CAT	ATG	AAA	AAC	ATC	AAA	AAA	AAC	CAG	GTA	ATG			2670
2502	R								H	H	R	N	I	K	N	O	12
2671	AAC	CTG	GGT	CCG	AAC	TGT	AAA	CTG	AAA	GAA	TAC	AAA	TCC	CAC	CTG	ATC	2730
2672	L	G	P	N	S	K	L	L	K	E	Y	K	S	Q	L	I	32
2731	ATC	GAA	CAG	TTC	GAA	GCA	GGT	ATC	CTG	ATC	GGT	GAT	GCT	TAC	ATC	CGT	2790
2732	I	E	Q	F	E	A	G	I	G	L	I	L	G	D	A	Y	52
2851	GAT	GAA	GGT	AAA	ACC	TAC	TGT	ATG	CAG	TTC	CAG	TGG	AAA	AAC	AAC	GCA	2850
2852	D	F	G	K	T	Y	C	H	Q	F	E	H	K	N	K	A	72
2911	GTA	TGT	CTG	CTG	TAC	GAT	CAG	TGG	GTA	CTG	TCC	CCG	CAC	AAA	AAA	GAA	2910
2912	V	C	L	L	Y	D	Q	W	V	L	S	P	H	K	K	E	92
2971	CAC	CTG	GGT	AAC	CTG	GTA	ATC	ACC	TGG	GGC	GCC	CAG	ACT	ITC	AAA	CAC	2970
2972	L	G	N	L	V	I	T	W	G	A	Q	T	F	K	H	Q	112
3031	AAA	CTG	GCT	AAC	CTG	TTC	ATC	GTT	AAC	AAC	AAA	ACC	ATC	CCG	AAC	AAC	3030
3032	R	L	A	N	L	F	I	V	N	N	K	K	T	I	P	N	132
3091	AAC	TAC	CTG	ACC	CCG	ATG	TCT	CTG	GCA	TAC	TGG	TTC	ATG	GAT	GAT	GGT	3090
3092	N	Y	L	T	P	M	S	L	A	Y	W	F	H	D	D	G	152
3151	TAC	AAC	AAA	TCT	ACC	AAC	AAA	TGG	ATC	GTA	CTG	AAC	ACC	CAG	TCT	TTC	3150
3152	Y	H	K	N	S	T	N	K	S	I	V	L	N	T	Q	S	172
3211	GAA	GTA	GAA	TAC	CTG	GTT	AAG	GGT	CTG	CGT	AAC	AAA	TTC	GAA	CTG	AAC	3210
3212	E	V	E	Y	L	V	K	G	L	R	N	K	F	Q	L	N	192
3271	ATC	AAC	AAA	AAC	AAA	CCG	ATC	ATC	TAC	ATC	GAT	TCT	ATG	TCT	TAC	CTG	3270
3272	I	N	K	N	K	P	I	I	Y	I	D	S	H	S	Y	L	212
3331	CTG	ATC	AAA	CCG	TAC	CTG	ATC	CCG	CAG	ATG	ATG	TAC	AAA	CTG	CCG	AAC	3330
3332	L	I	K	P	Y	L	I	P	Q	H	H	Y	K	L	P	N	232
3401	GAA	ACT	TTC	CTG	AAA	TAA	TAA	GCGAGGCGAGCCGACGCTTGGCAGCTGGCCGCTGCTTTTACAAACGTCGTGCACT									3404
3402	E	T	F	L	K	+											238

SalI PstI

FIGURA 4B

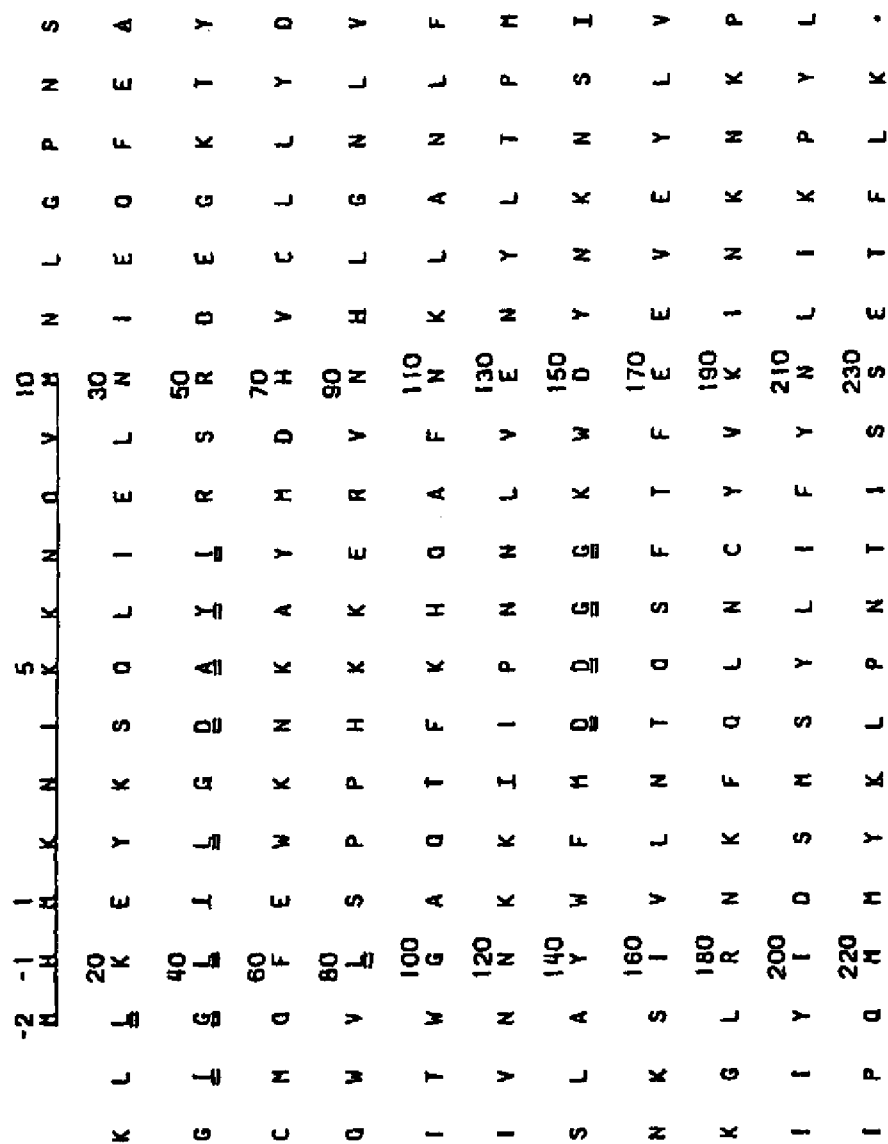


FIGURA 5

Endonucleasa		Secuencia de Reconocimiento	Sitio de escisión	▼ Sitio del Intrón
Familia de Dos Dodecapeptidos (o cortadores de 4 pb)	I-Sce I (<i>Saccharomyces mitochondria</i>)	C G C T A G G G A T A A C A G G G T A A T A T A G C G C G A T C C C T A T T G T C C C A T T A T A T C G		
	I-Sce IV (<i>Saccharomyces mitochondria</i>)	T T C T C A T G A T T A G C T C T A A T C C A T G G A A G A G T A C T A A T C G A G A T T A G G T A C C		
	I-Sce II (<i>Saccharomyces mitochondria</i>)	C T T T G G T C A T C C A G A A G T A T A T A T T T G A A A C C A G T A G G T C T T C A T A T A T A A A		
	I-Ceu I (<i>Chlamydomonas chloroplast</i>)	T A A C G G T C C T A A G G T A G C G A A A T T C A A T T G C C A G G A T T C C A T C G C T T T A A G T		
	I-Ppo I (<i>Physarum nucleus</i>)	T G A C T C T C T T A A G G T A G C C A A T G C C A C T G A G A G A A T T C C A T C G G T T A C G G		
	I-Sce III (<i>Saccharomyces mitochondria</i>)	G G A G G T T T T G G T A A C T A T T T A T T A C C C C T C C A A A A C C A T T G A T A A A T A A T G G		
	I-Cre I (<i>Chlamydomonas chloroplast</i>)	G G G T T C A A A A C G G C G T G A G A C A G T T T C C C A A G T T T T G C A G C A C T C T G T C A A A		
	Endo.Sce I (RF3) (<i>Saccharomyces mitochondria</i>) (no intrónico)	G A T G C T G T A G G C A T A G G C T T G G T T A T C T A C G A C A T C C G T A T C C G A A C C A A T A		
	HO (<i>Saccharomyces nucleus</i>) (no intrónico)	C T T T C C G C A A C A G T A T A A T T T T A T A A G A A A G G C G T T G T C A T A T T A A A A T A T T		
	I-Csm I (<i>Chlamydomonas mitochondria</i>) (Endonucleasa putativa)	A C C A T G G G G T C A A A T G T C T T T C T G G G T G G T A C C C C A G T T T A C A G A A A G A C C C		
	I-Pan I (<i>Podospira mitochondria</i>) (Endonucleasa putativa)	G T G C C T G A A T G A T A T T T A T T A C C T T T C A C G G A C T T A C T A T A A A T A A T G G A A A		
Otras familias Estructurales	(Bacteriófago T4)			
	I-Tev I	C A A C G C T C A G T A G A T G T T T T C T T G G G T C T A C C G T T T A A T G T T G C G A G T C A T C T A C A A A G A A G C C A G A T G G C A A A T T A		
	I-Tev II	C A A G C T T A T G A G T A T G A A G T G A A C A C G T T A T T G T T C G A A T A C T C A T A C T T C A C T T G T G C A A T A A		
	I-Tev III	G C T A T T C G T T T T T A T G T A T C T T T T G C G T G T A G C T T T A A C G A T A A G C A A A A A T A C A T A G A A A A C G C A C A T C G A A A T T		

FIGURA 6

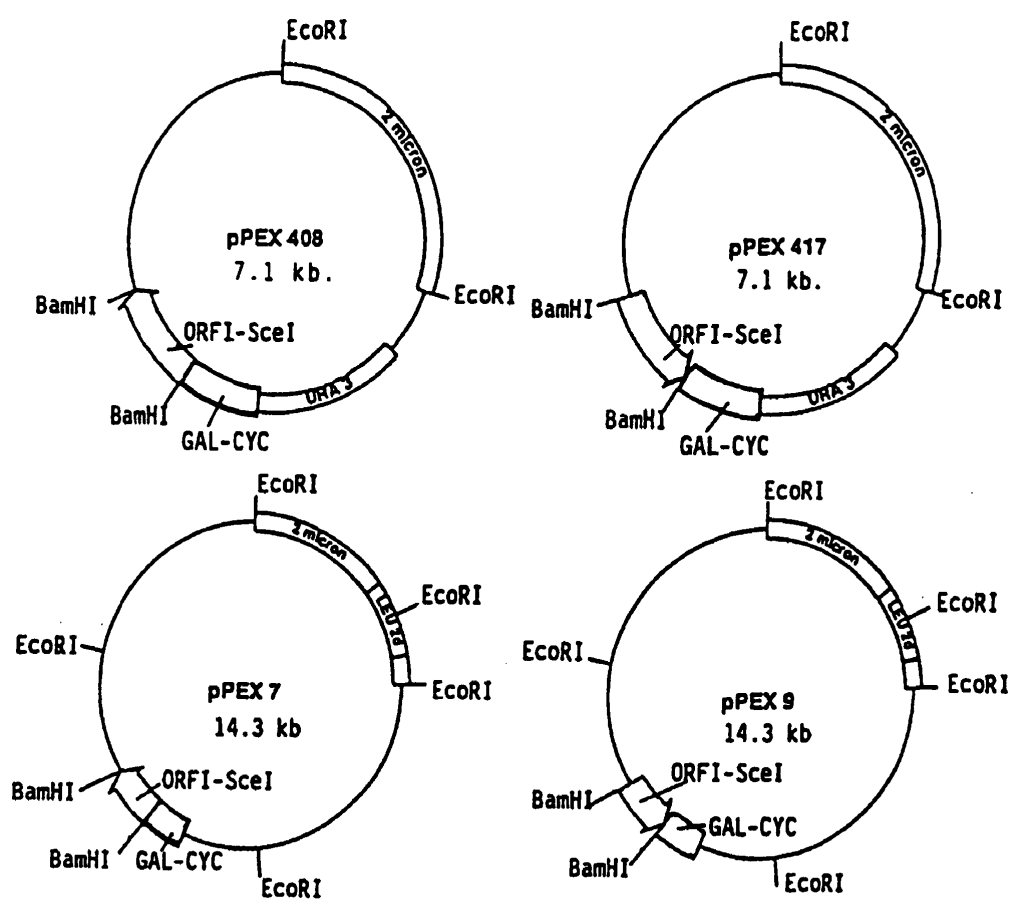


FIGURA 7

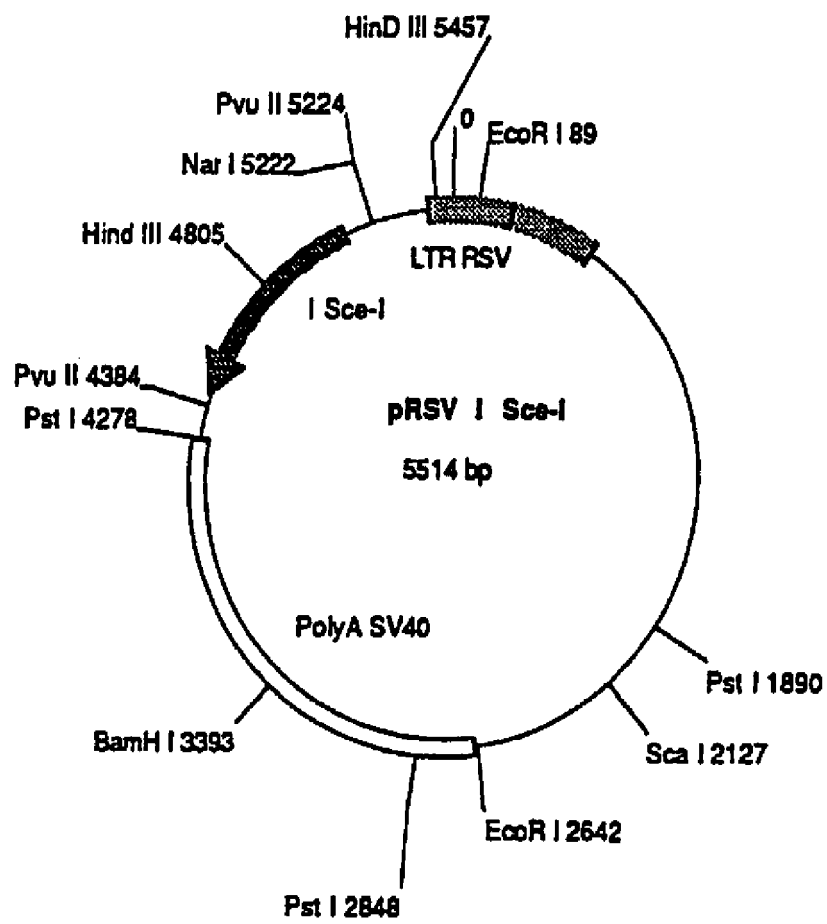


FIGURA 8

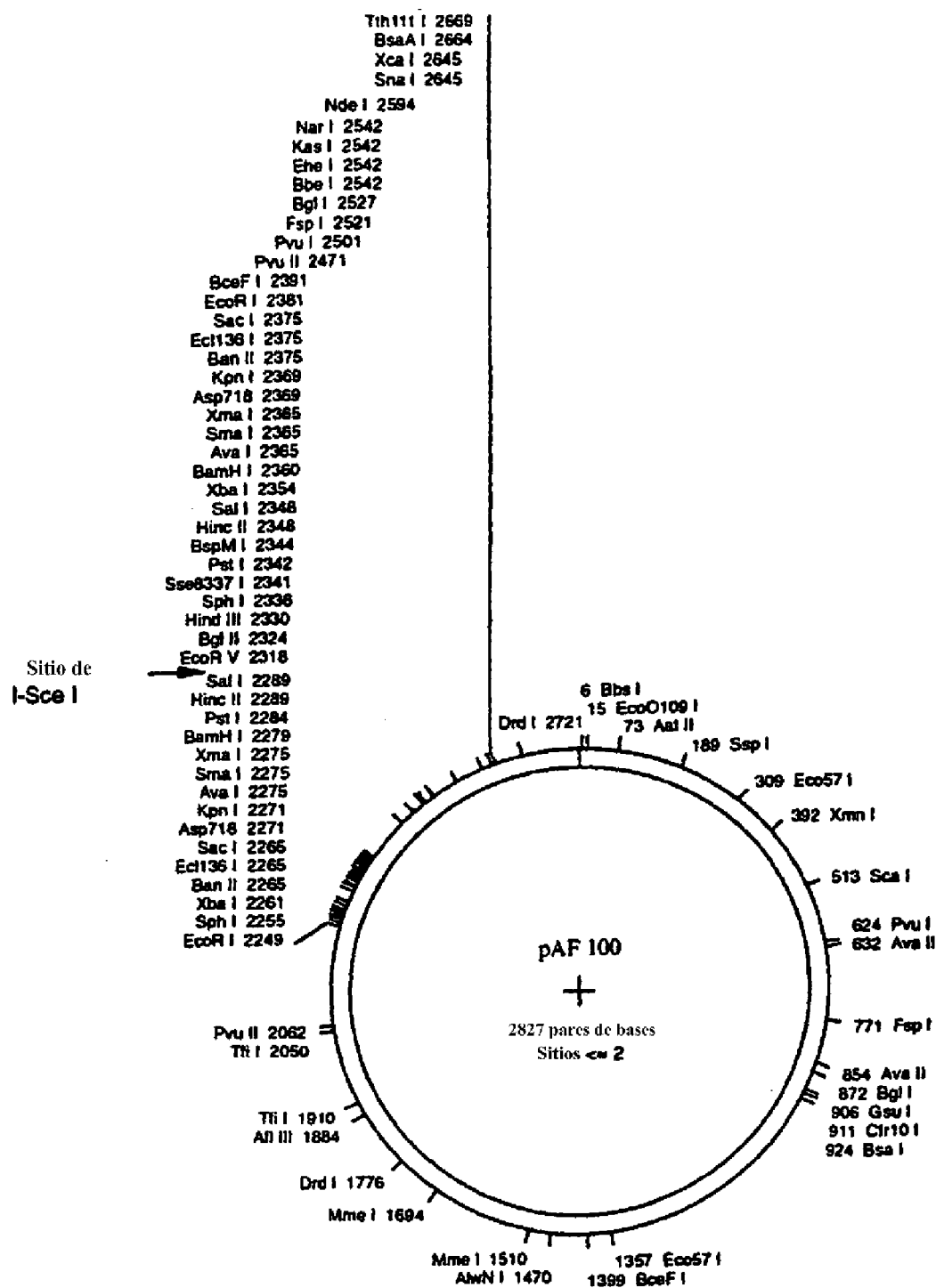


FIGURA 9

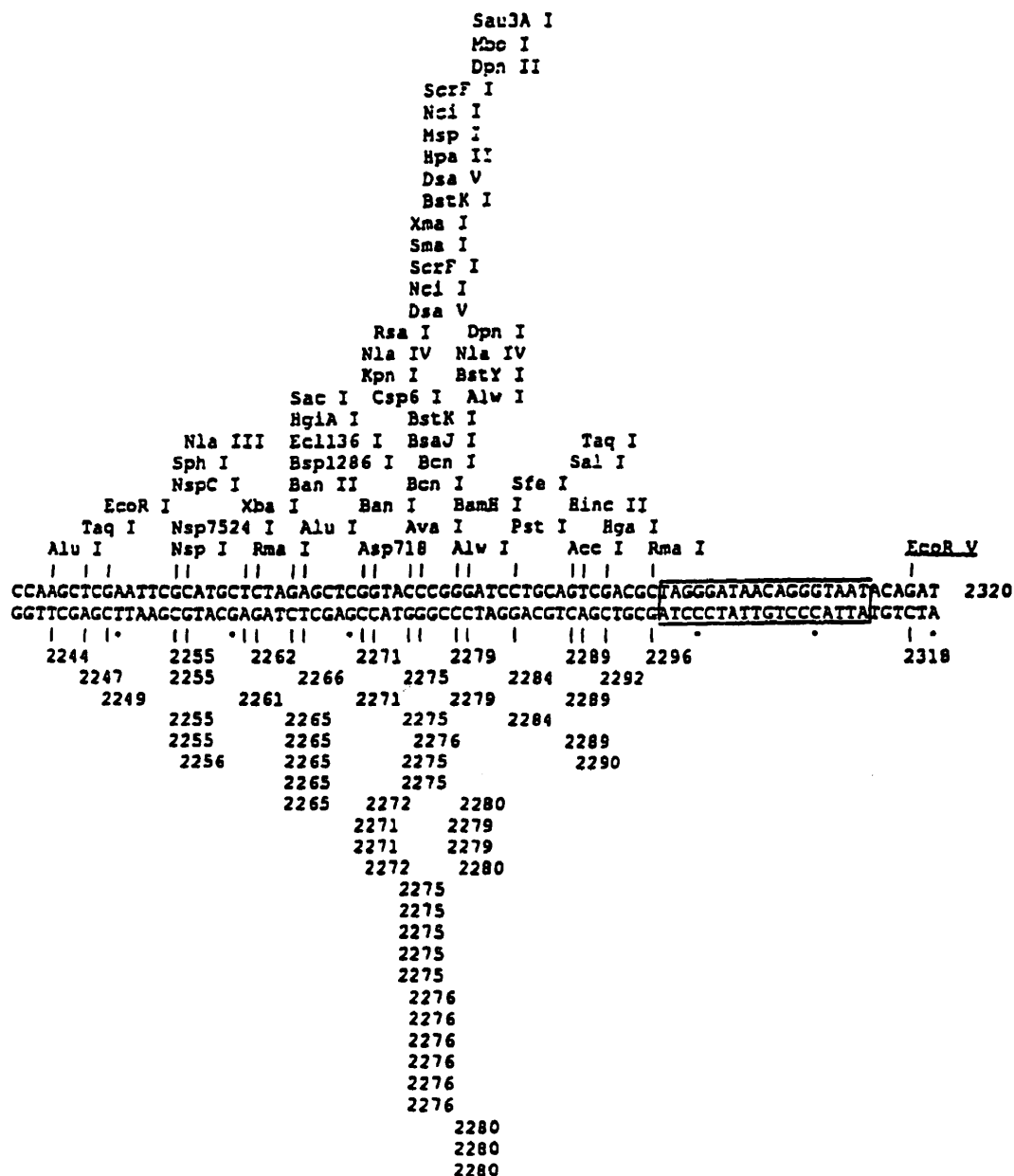


FIGURA 10A

ES 2 338 952 T3

ScrF I
 Nci I
 Msp I
 Bpa II
 Dsa V
 BstK I
 Bcn I
 Xma I
 Sma I
 ScrF I
 Nci I
 Dsa V
 BstK I
 BsaJ I
 Bcn I
 Sau3A I
 Mbo I
 Dpn II
 Dpn I
 Nla IV
 BstY I
 BamH I
 Alw I
 Nla III
 Sph I
 Dpn I
 BstY I
 Bgl II
 Alu I
 Hind III
 Dde I
 Sau3A I
 Mbo I
 Dpn II
 Dpn I
 BstY I
 Bgl II
 Sfe I
 Pst I
 Sse8337 I
 Nla III
 Sph I
 NspC I
 Nsp7524 I
 Nsp I
 Ple I
 Hinf I
 Sal I
 Hinc II
 Acc I
 Rsa I
 Xba I
 Mnl I
 Ava I
 BsaJ I
 Alw I
 Asp718
 Sac I
 HgiA I
 Eco136 I
 Rsa I
 Casp6 I
 Nla IV
 Kpn I
 Ban I
 Alu I
 Bsp1286 I
 Ban II
 Taq I
 Bsr I
 BceF I
 Hae III
 Gdi II
 Eae I
 Bsr I

ATCAGATCTAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTACCGAGCTCGAATTCACCTGCCCGTCGTTT 2400
 TAGTCTAGATTCCGAACGTACGGACGCTCCAGCTGAGATCTCTTAGGGGCCCATGGCTCGAGCTTAAAGTGACCCGGCAGCAAA

2324 2336 2344 2354 2361 2369 2379 2387
 2324 2336 2348 2355 2364 2375 2389
 2325 2336 2348 2358 2365 2375 2389
 2325 2336 2348 2360 2369 2376 2390
 2325 2337 2349 2360 2369 2381 2391
 2325 2341 2360 2369
 2328 2342 2351 2360 2370
 2330 2342 2351 2361 2370
 2331 2361 2375
 2361 2375
 2361 2375
 2365
 2365
 2365
 2365
 2365
 2365
 2365
 2365
 2365
 2365
 2366
 2366
 2366
 2366
 2366
 2366
 2366

FIGURA 10B

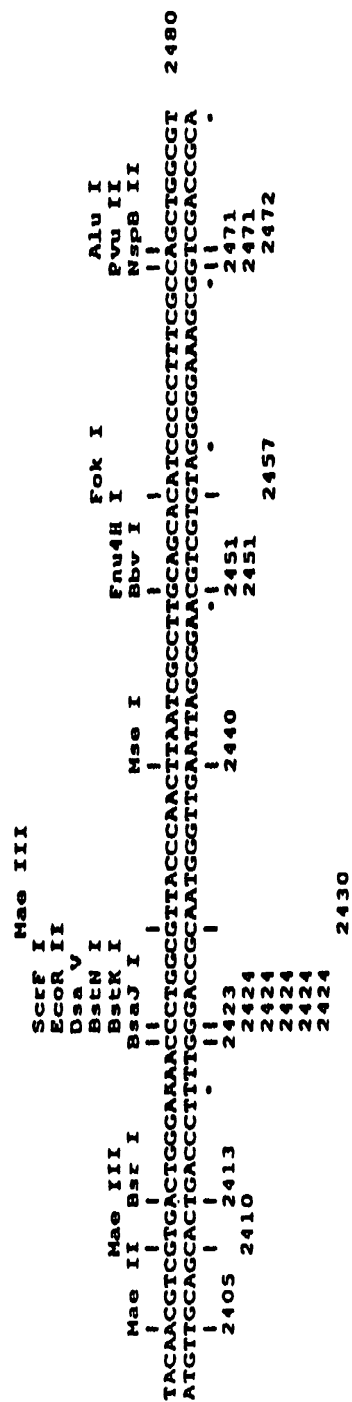


FIGURA 10C

Nombres : pTSm ω
pTKm ω
pTTc ω

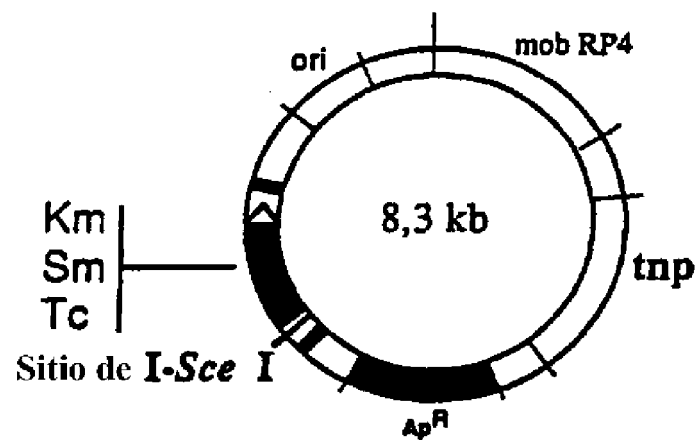


FIGURA 11

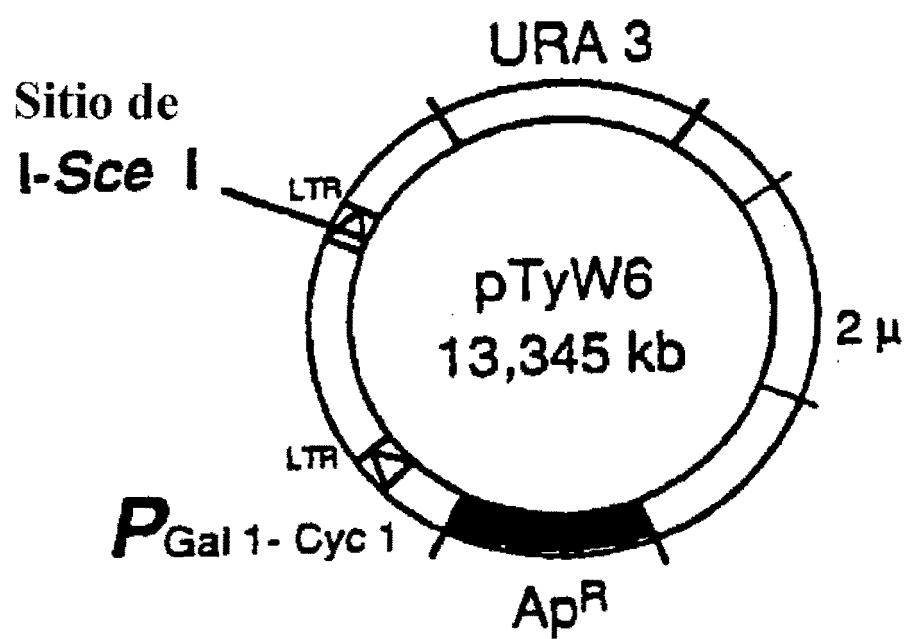


FIGURA 12

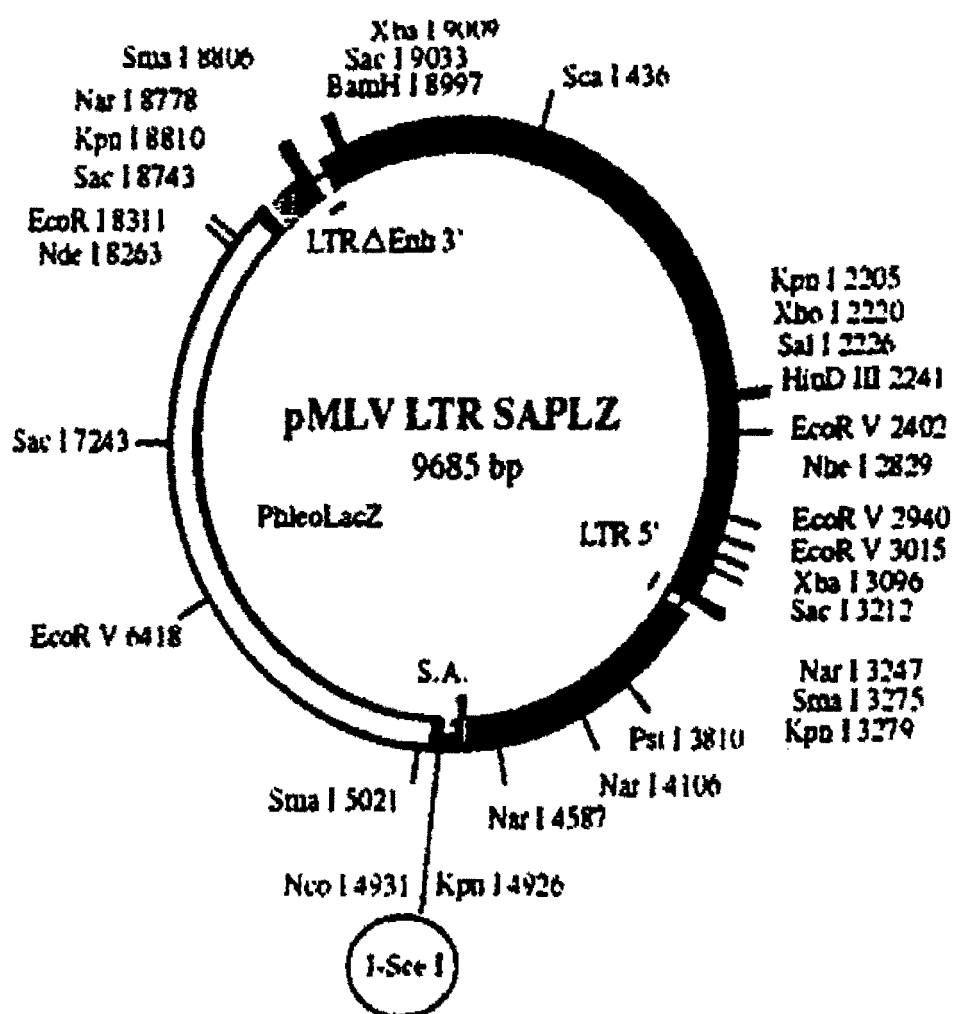


FIGURA 13A

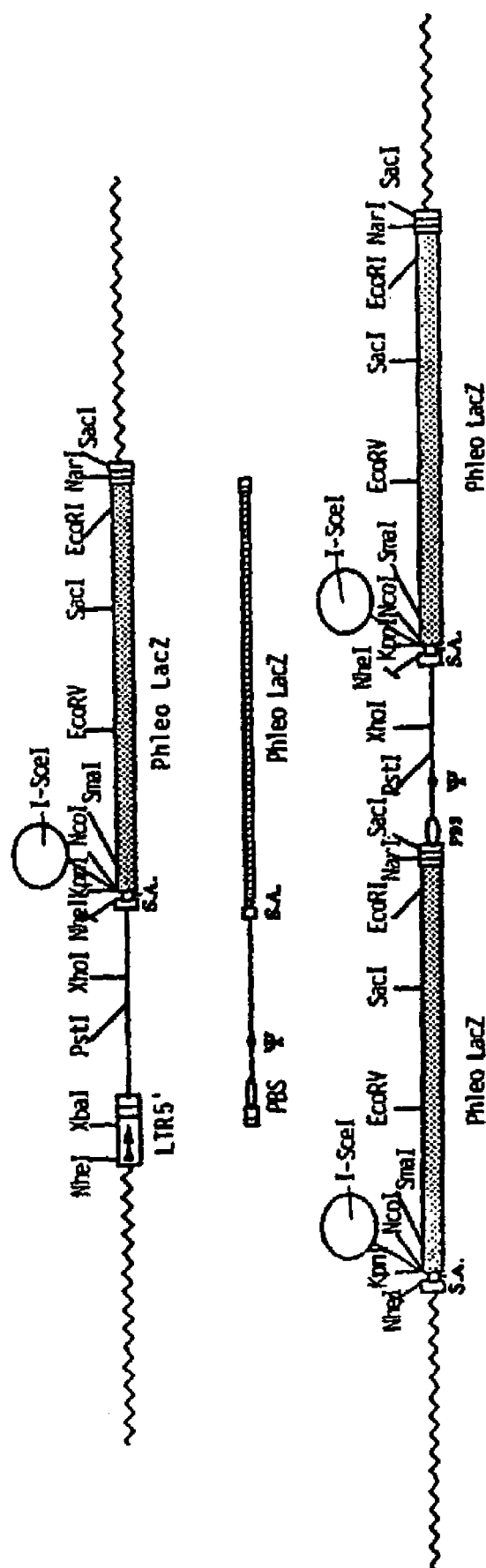


FIGURA 13B

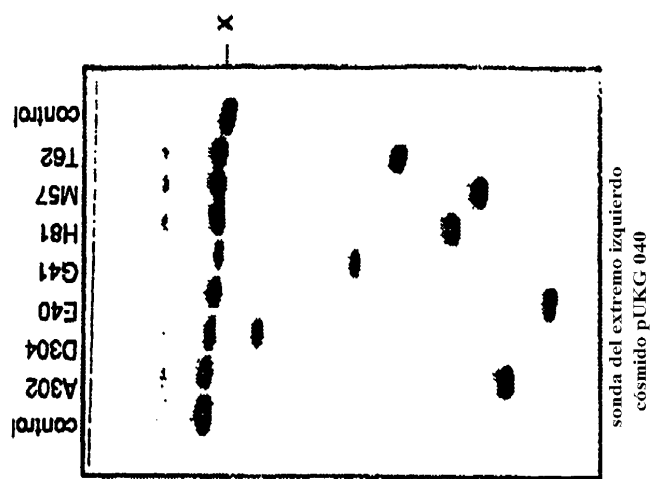


FIGURA 14B

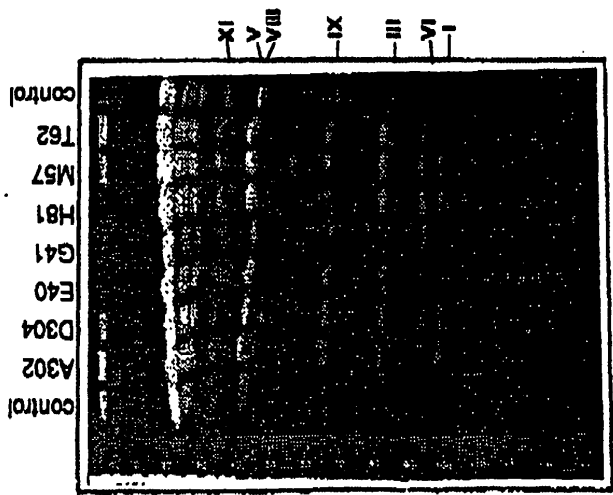


FIGURA 14A

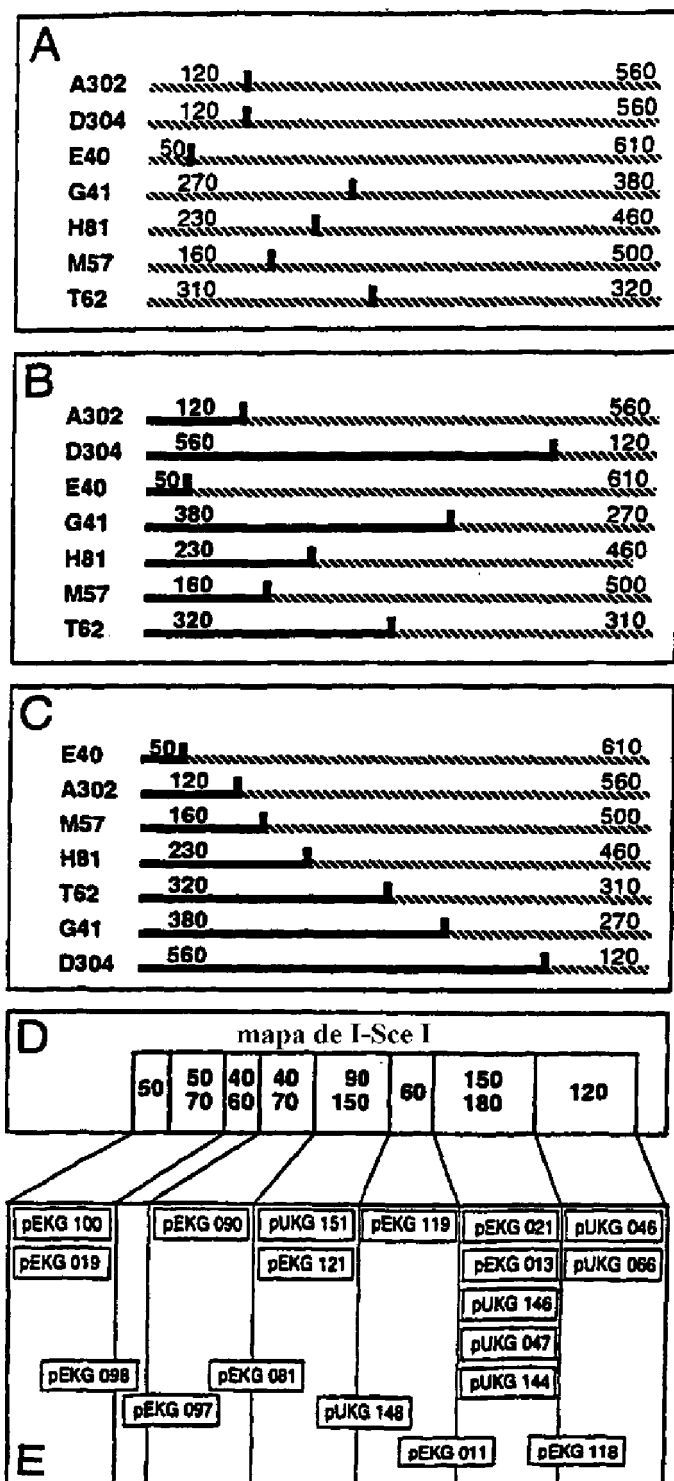


FIGURA 15

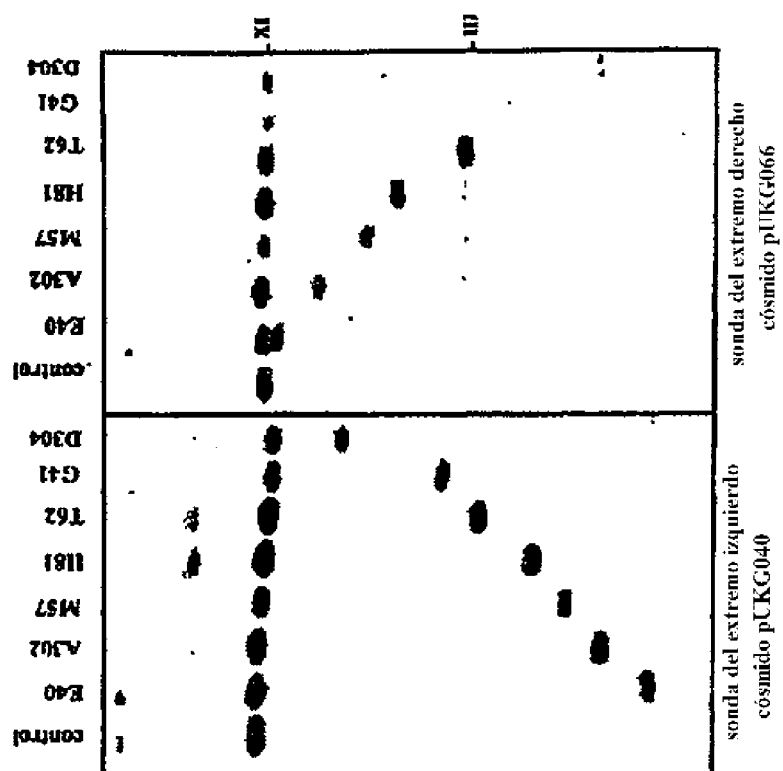


FIGURA 16B

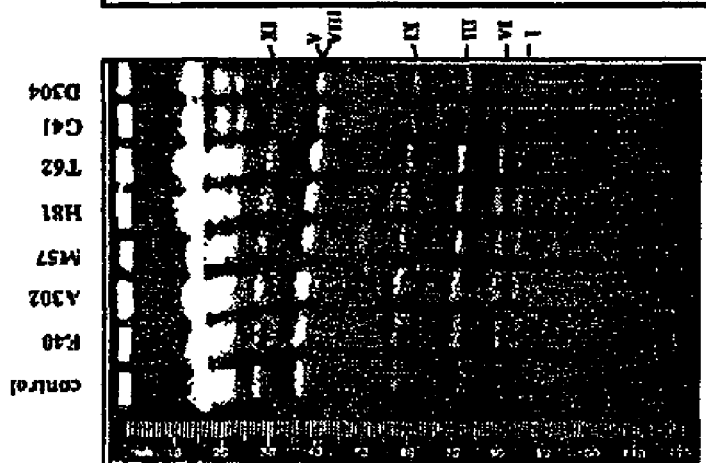


FIGURA 16A

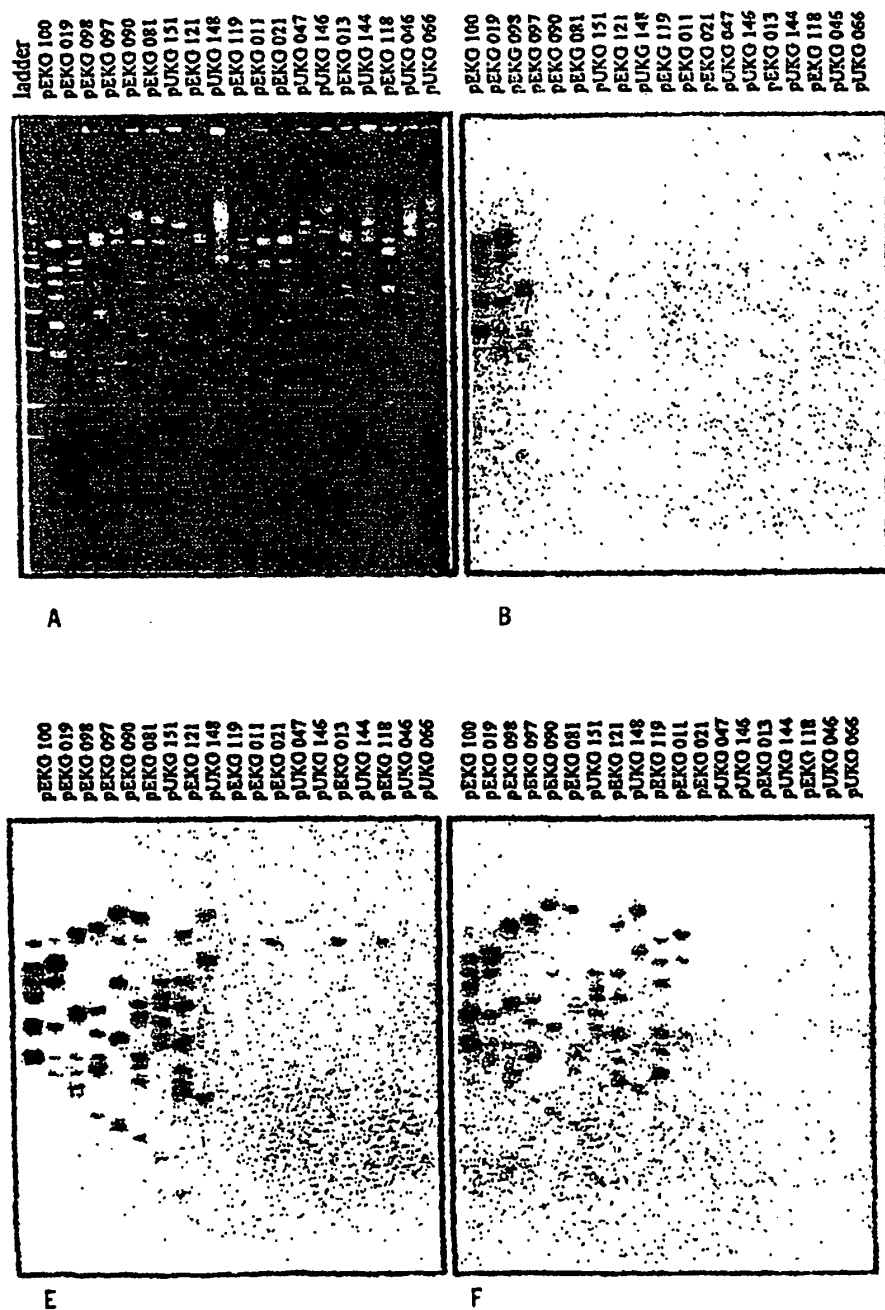


FIGURE 17A

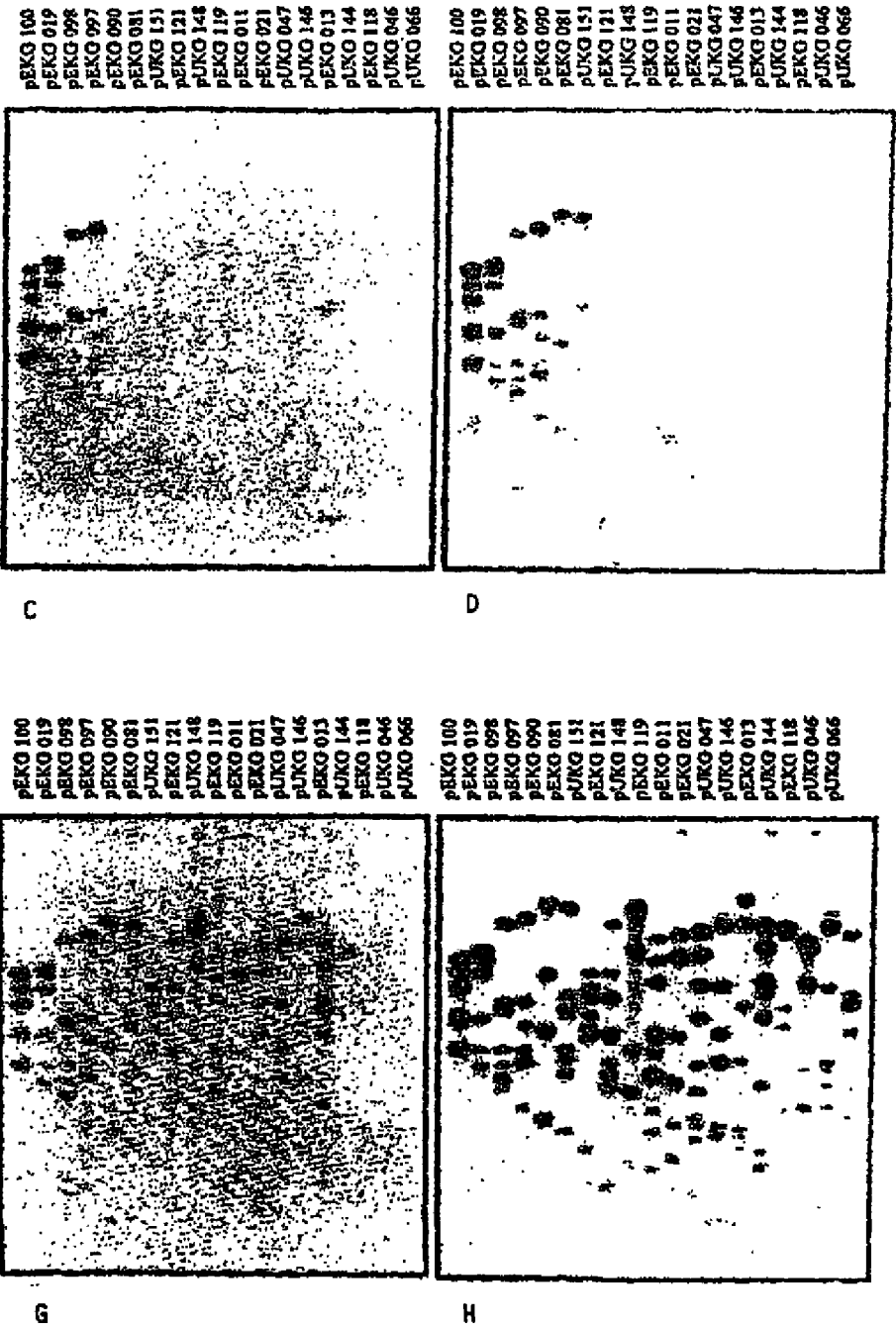
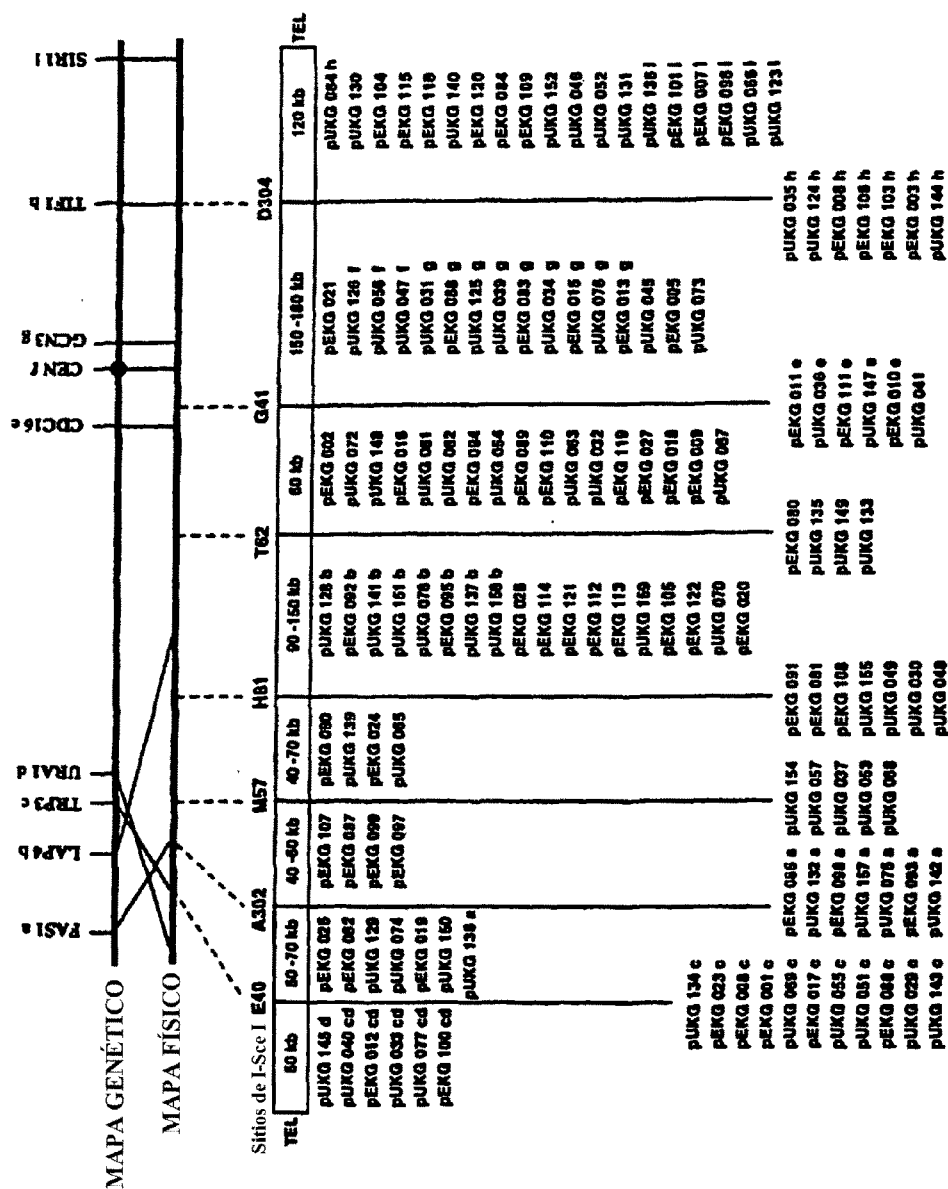


FIGURA 17B

FIGURA 18



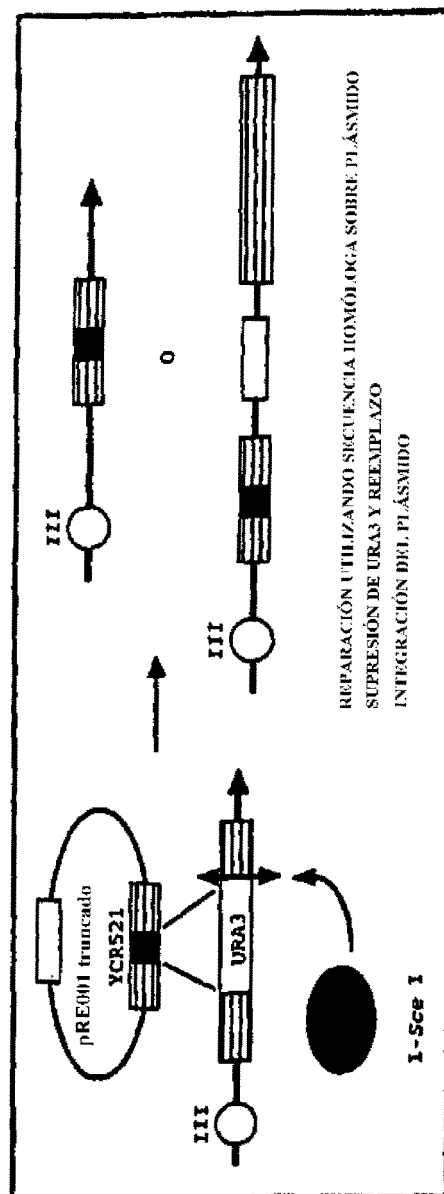
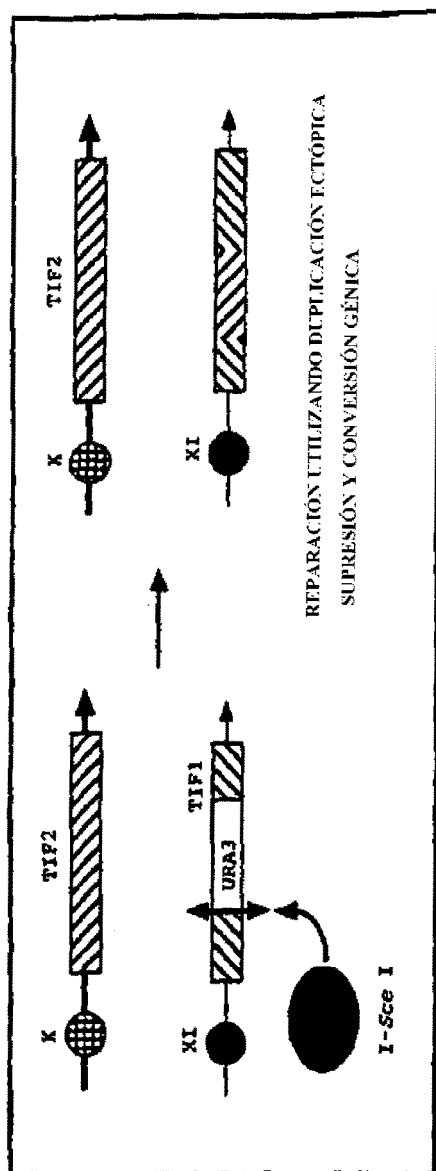


FIGURA 19

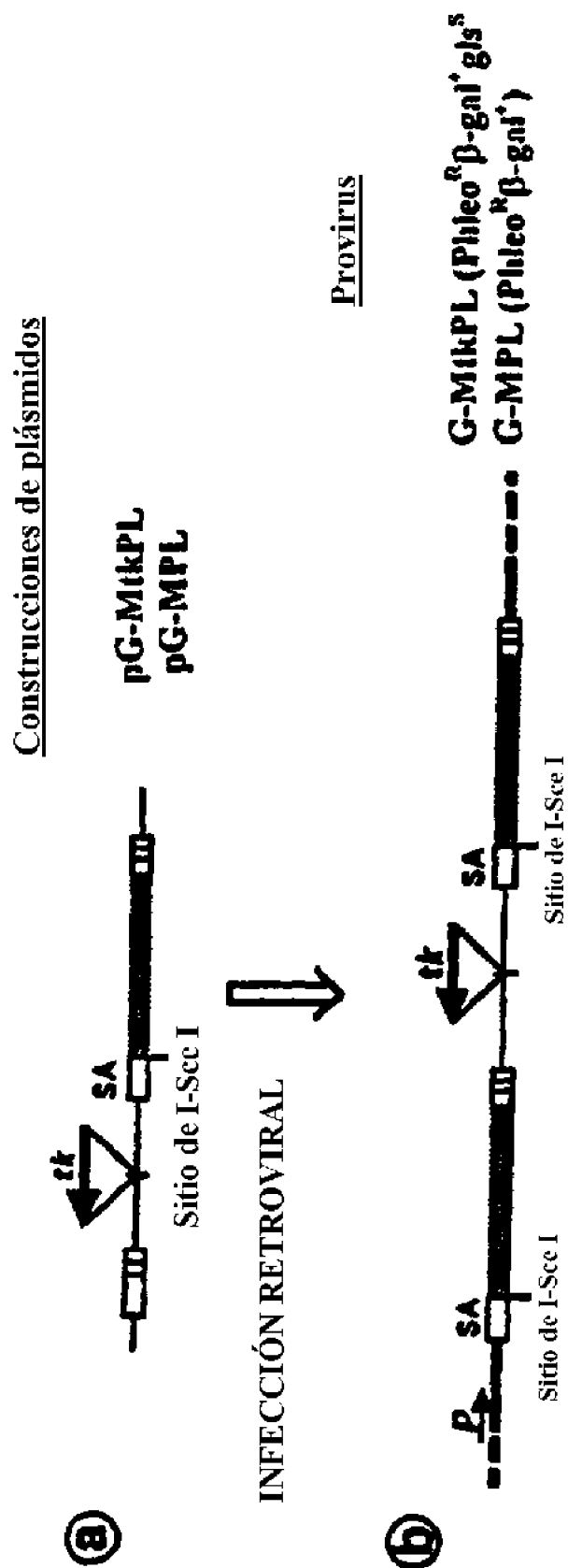
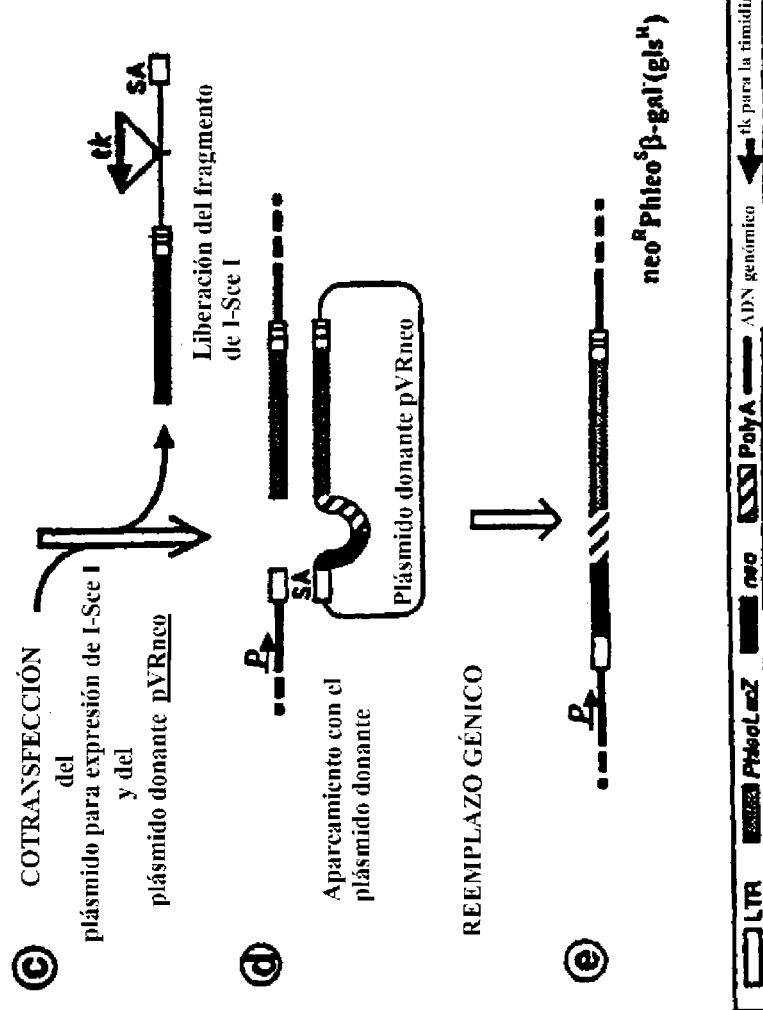


FIGURA 20A

FIGURA 20B



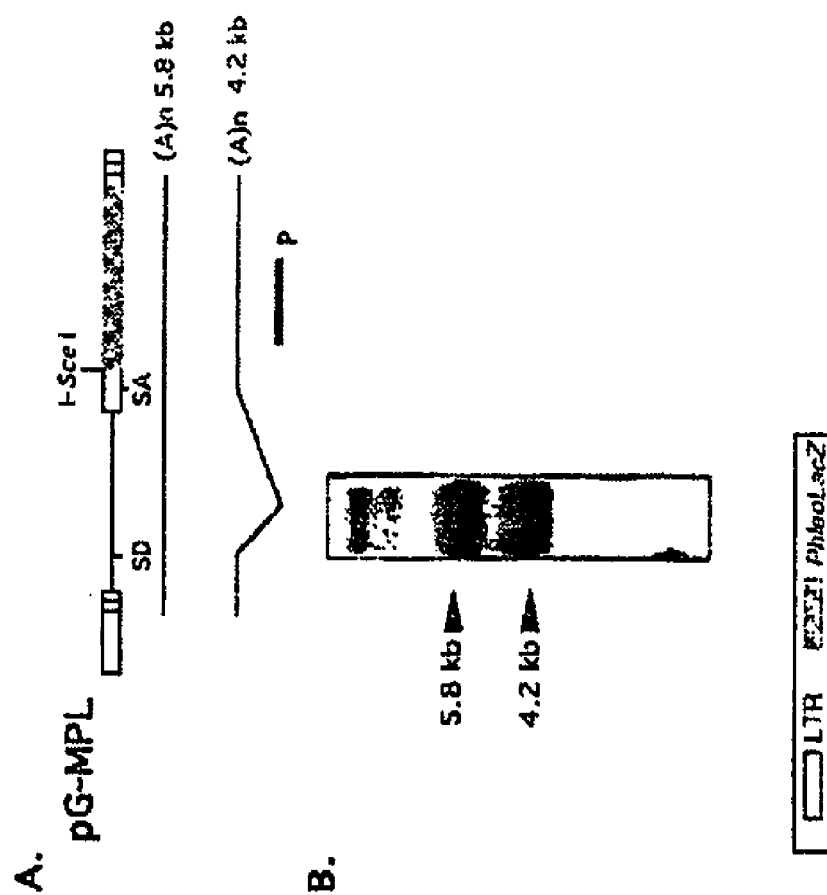


FIGURA 21

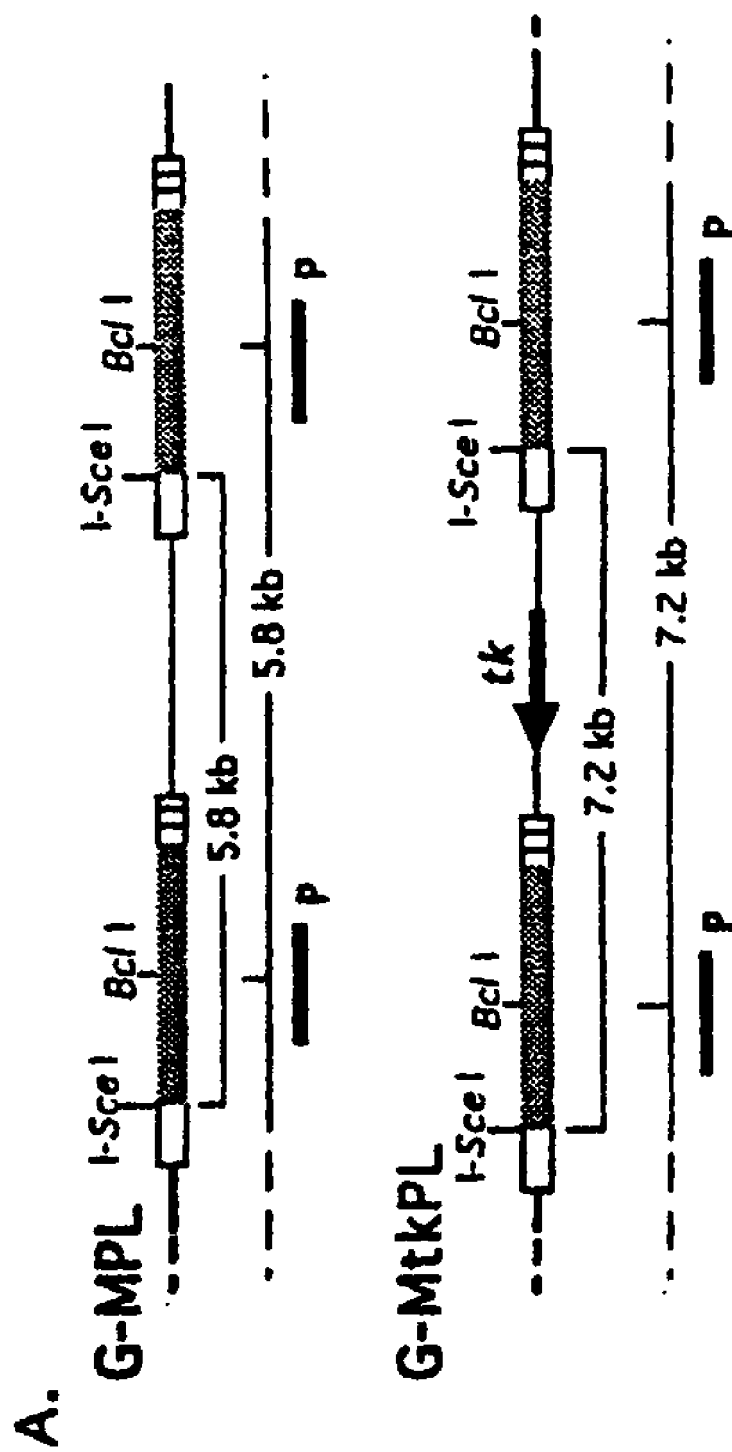
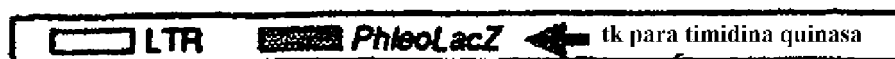
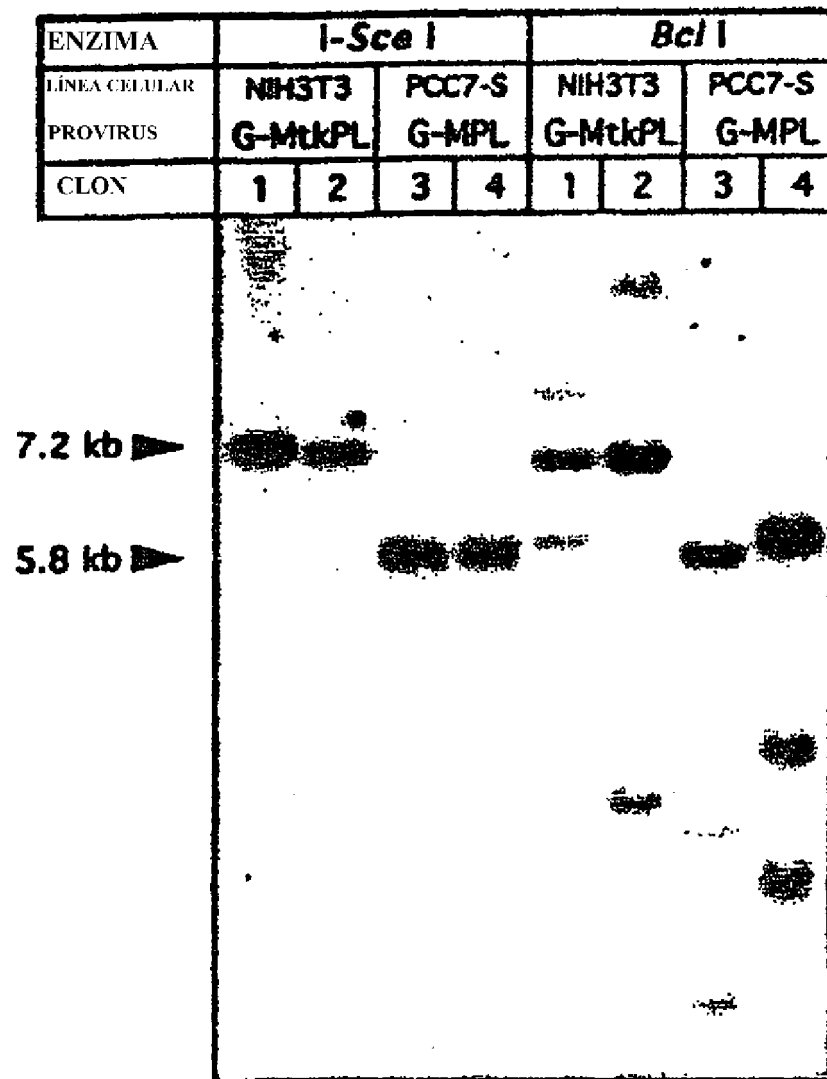


FIGURA 22A

B.



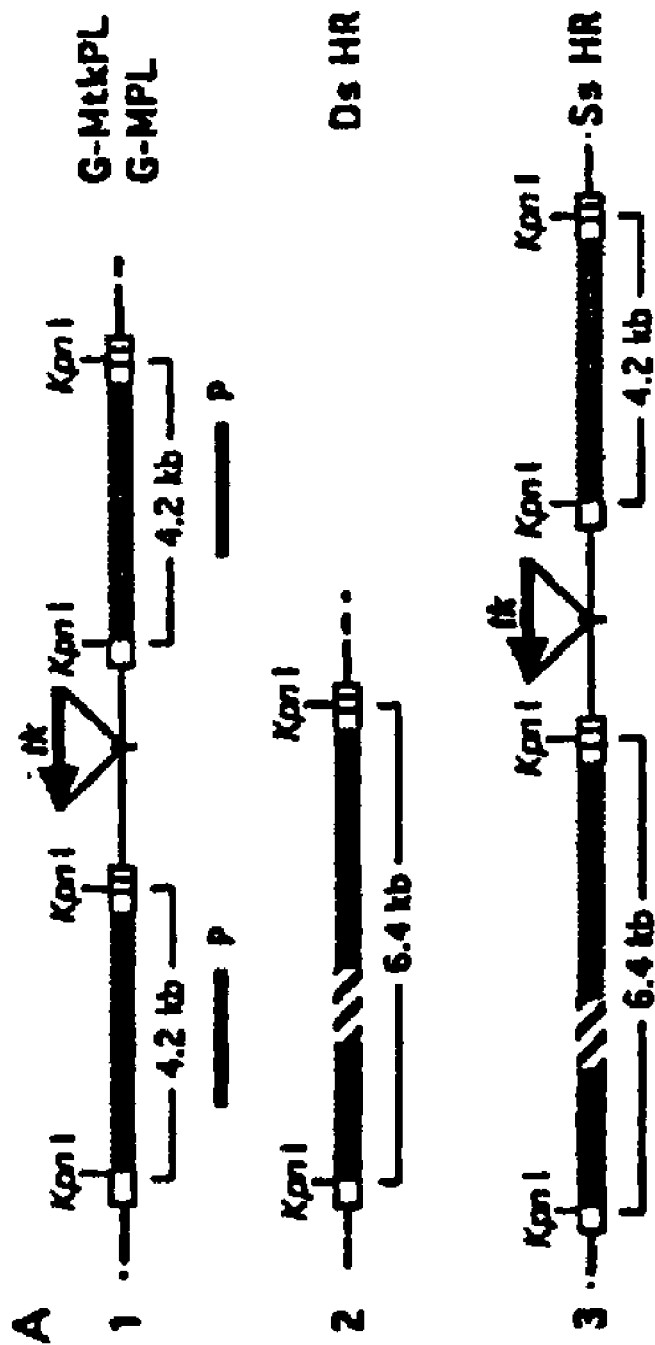


FIGURA 23A

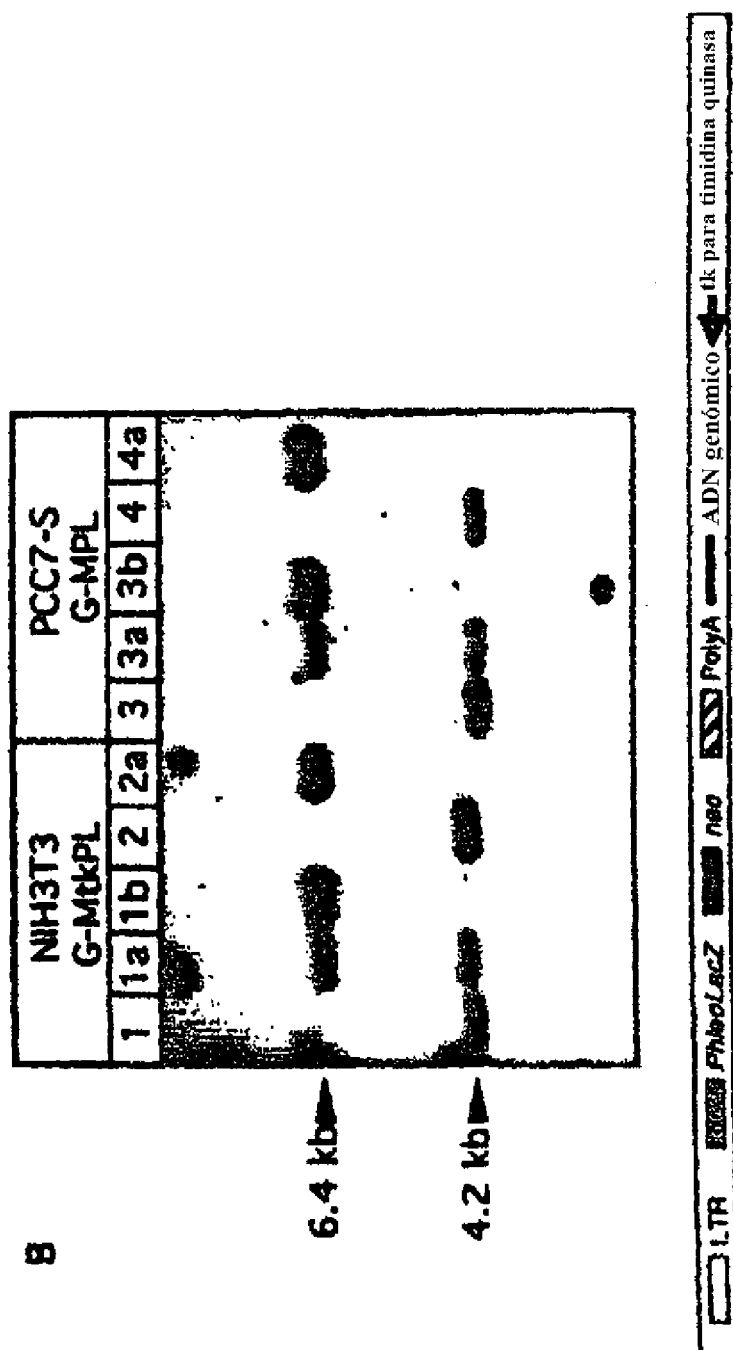


FIGURA 23B

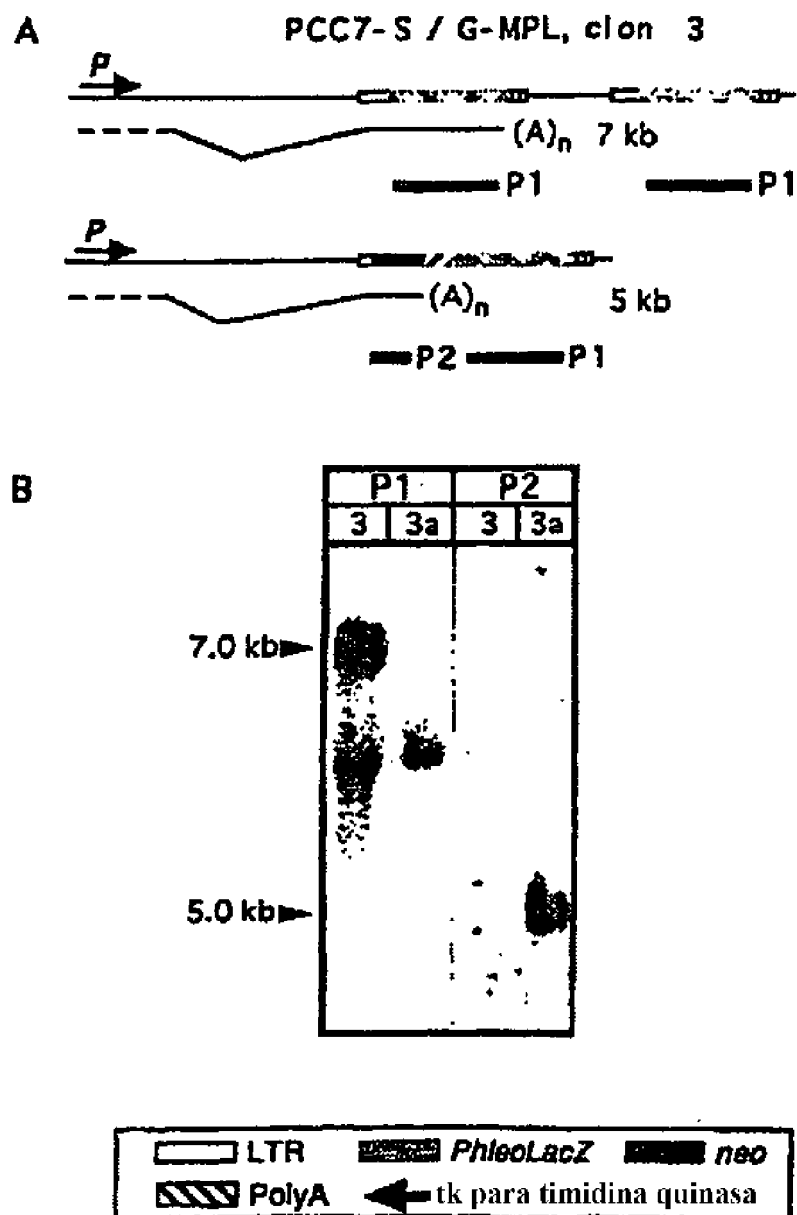


FIGURA 24

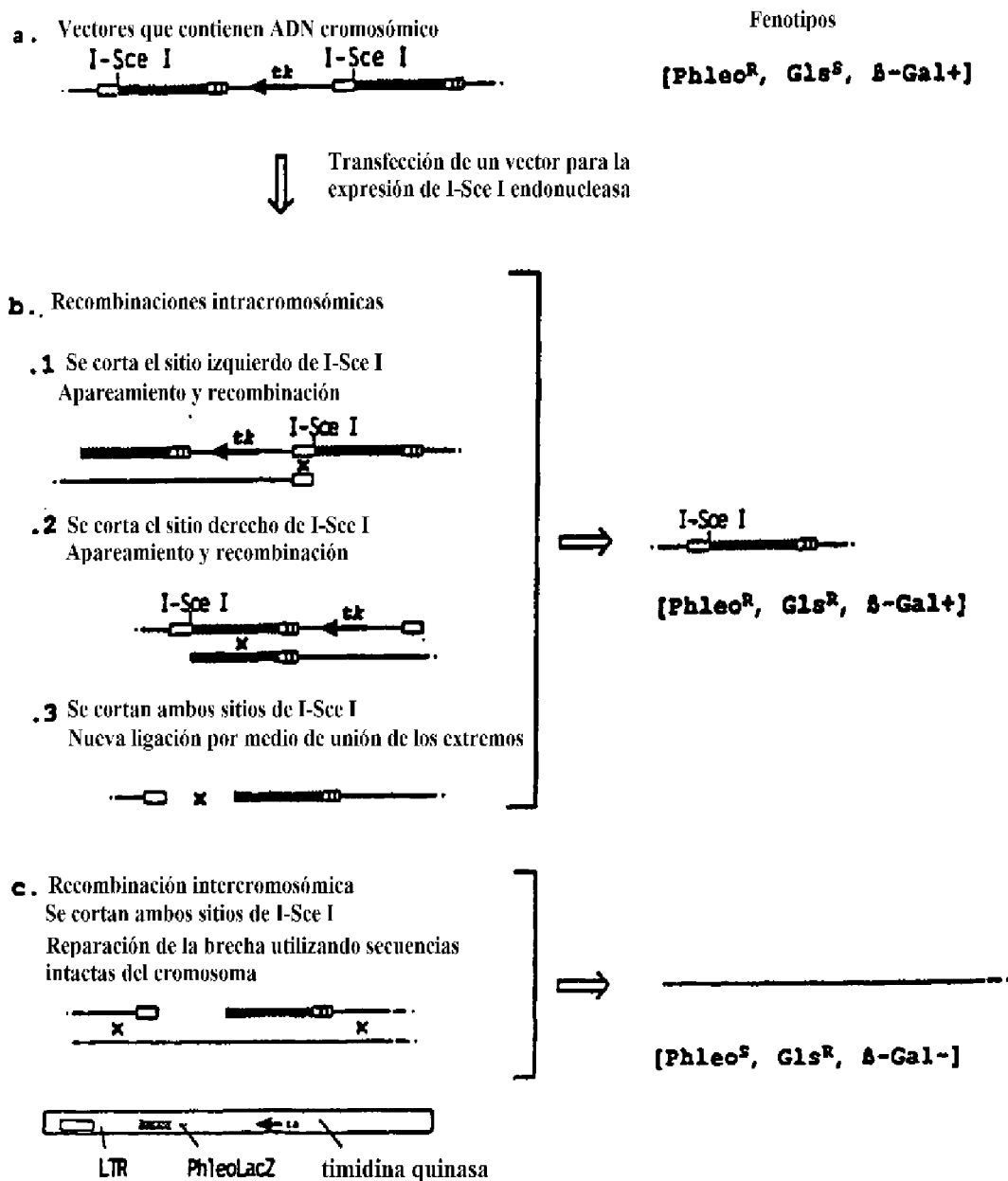


FIGURA 25

a. ADN progenitor, G-MtkPL

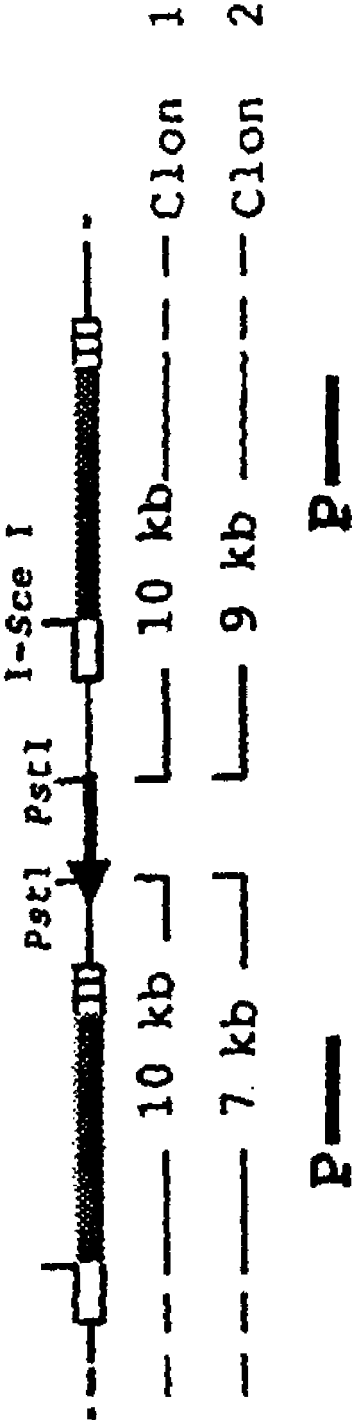


FIGURA 26A

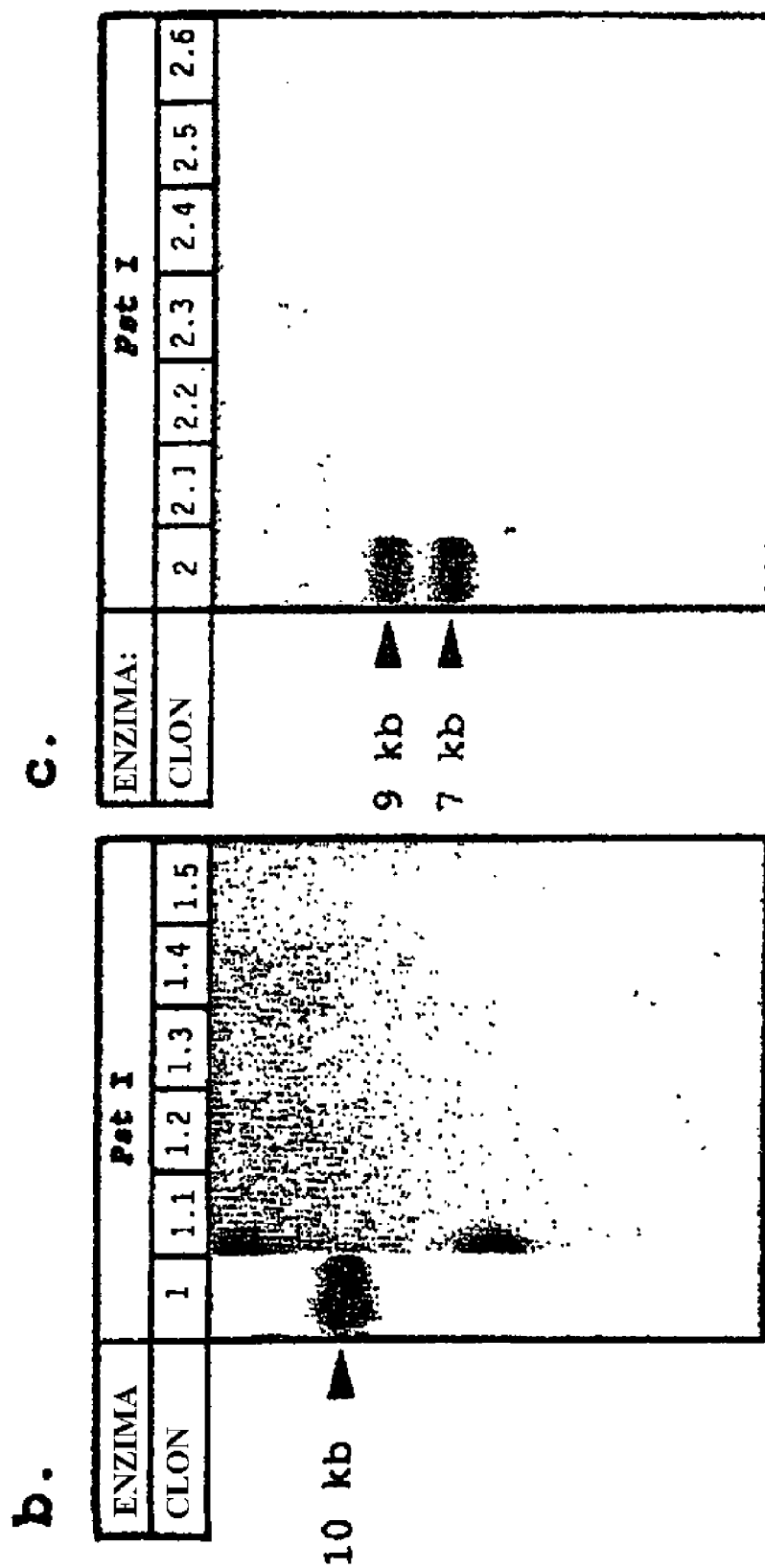


FIGURA 26B

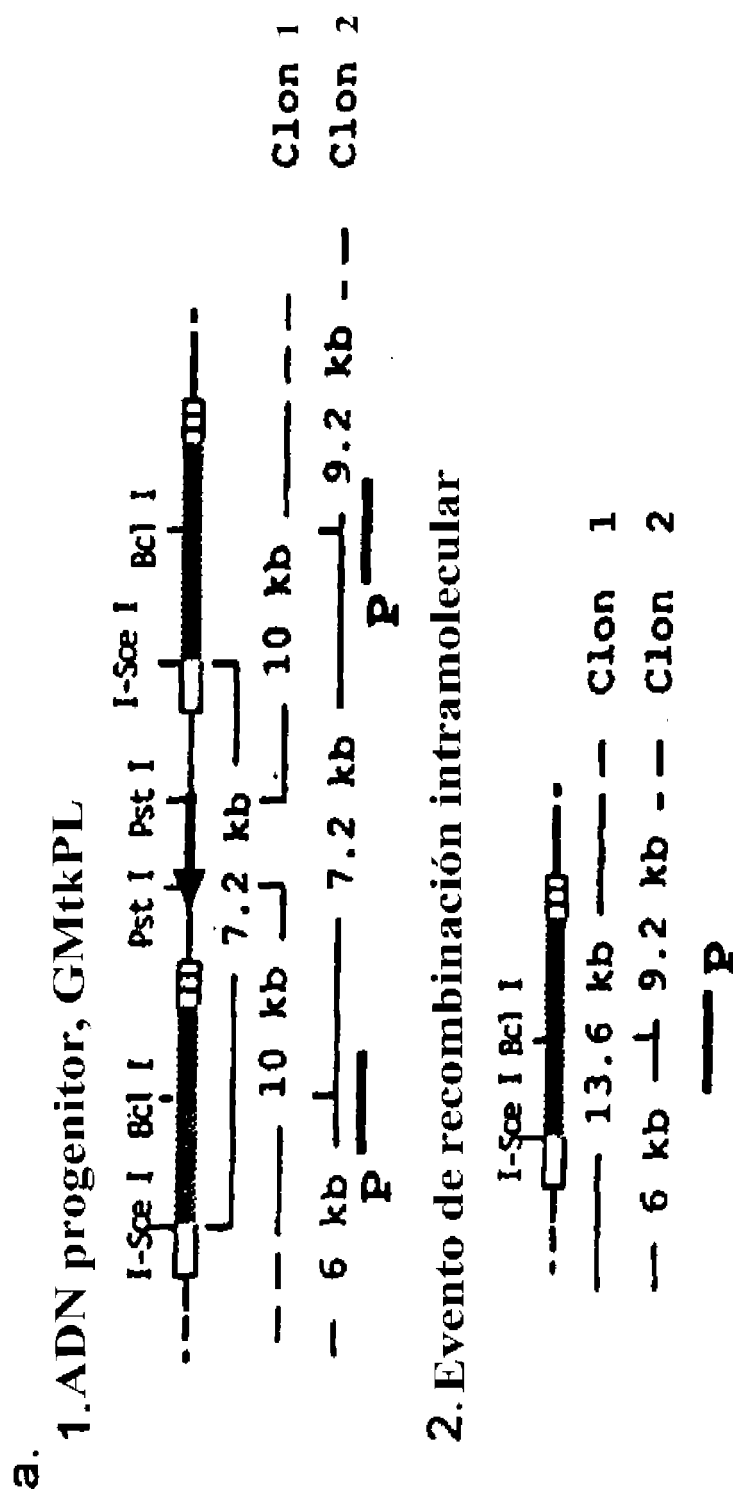
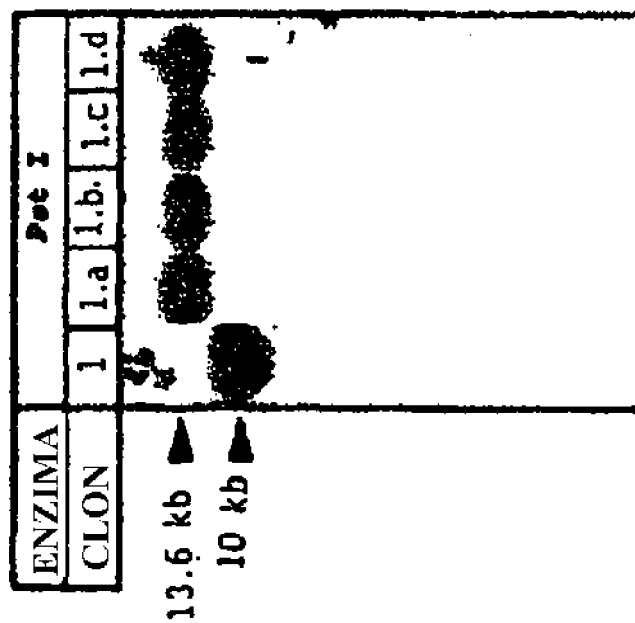


FIGURA 27A

b.



c.

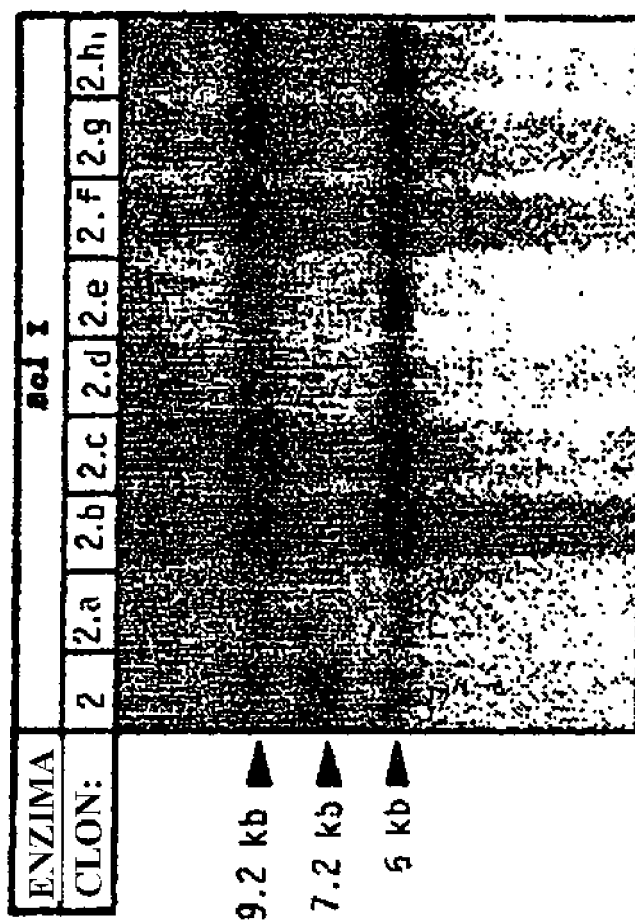


FIGURA 27B

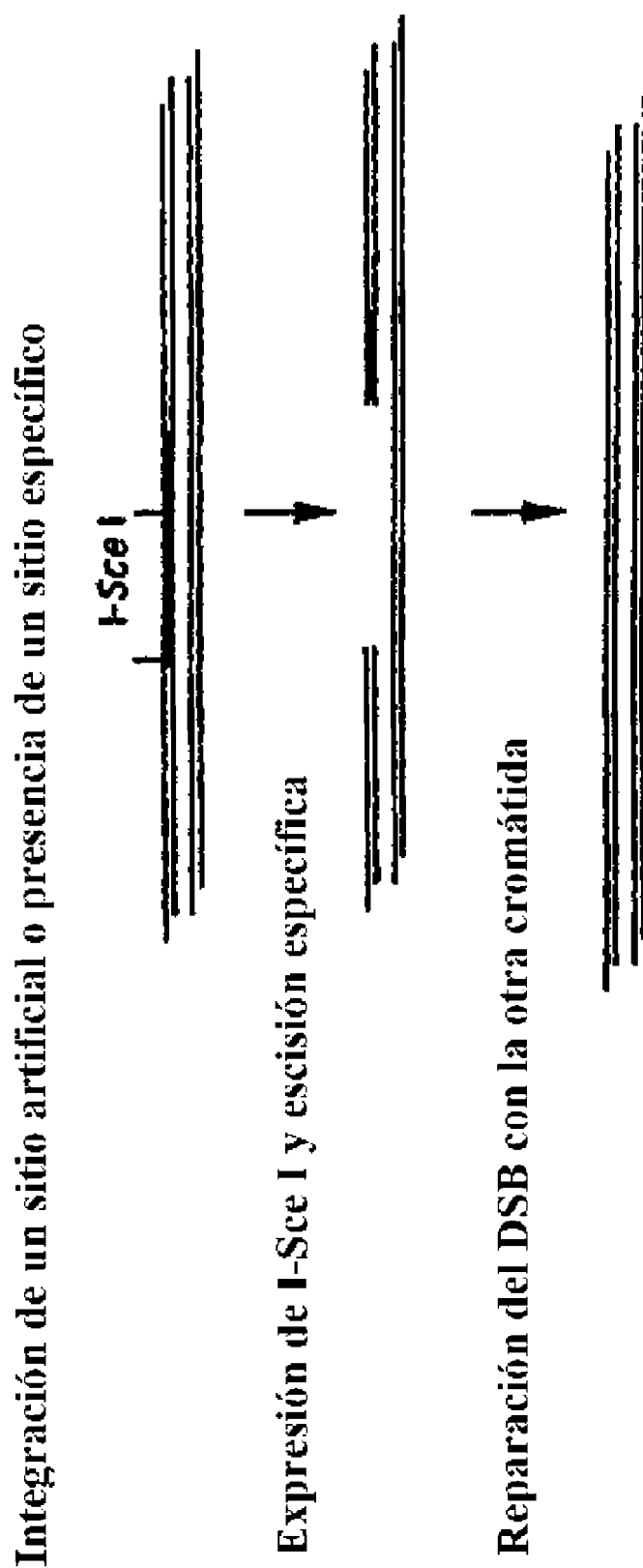


FIGURA 28

ACTIVACIÓN CONDICIONAL (Repetición en tándem)

Integración de un sitio artificial entre repeticiones en tándem

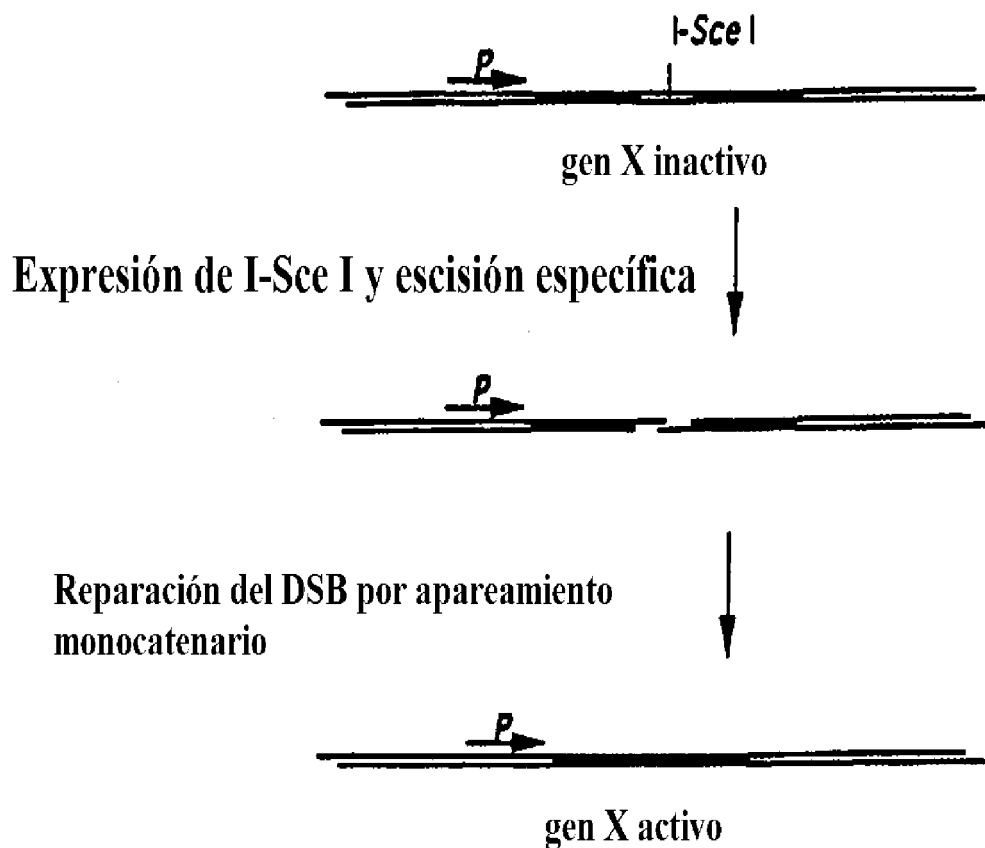


FIGURA 29

Integración de un sitio artificial o presencia de un sitio activo

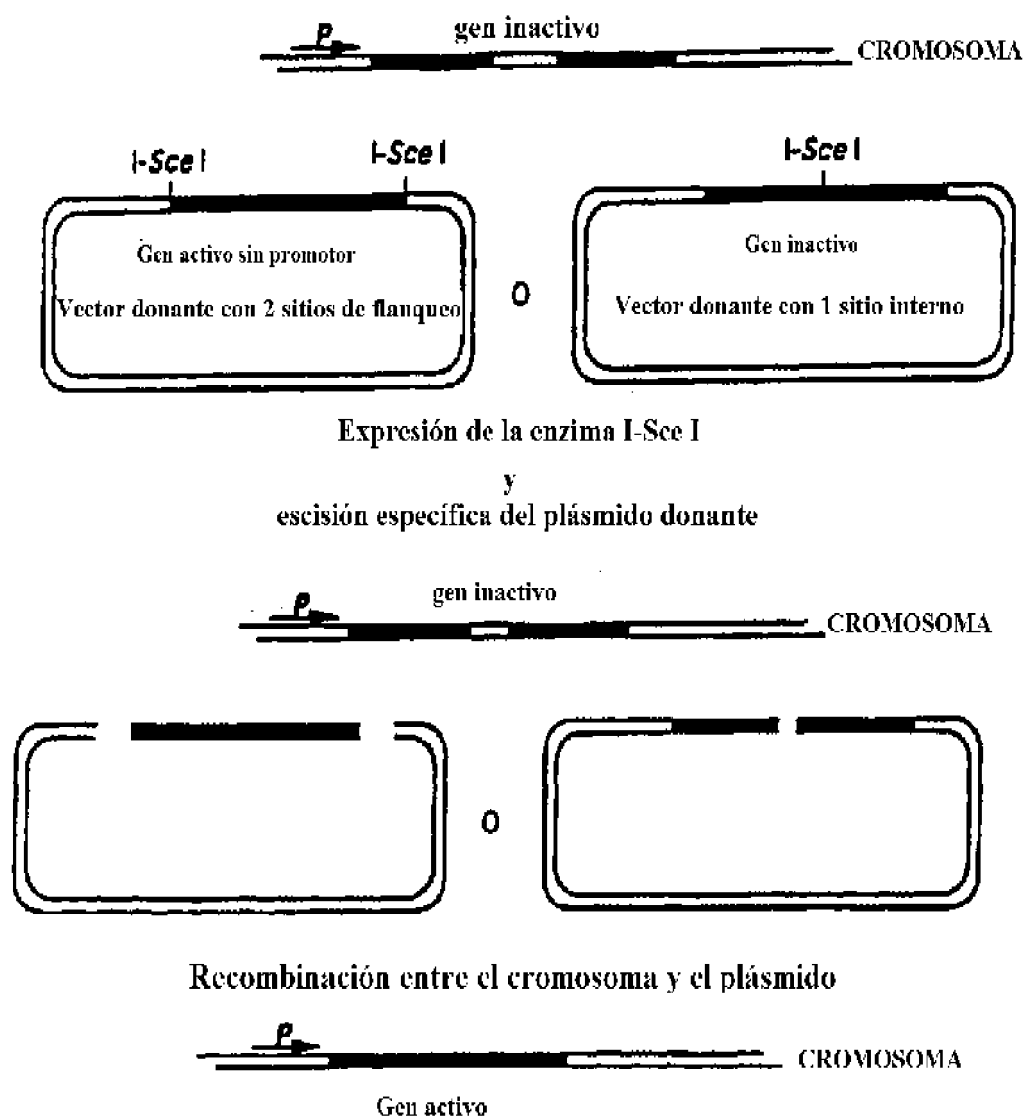
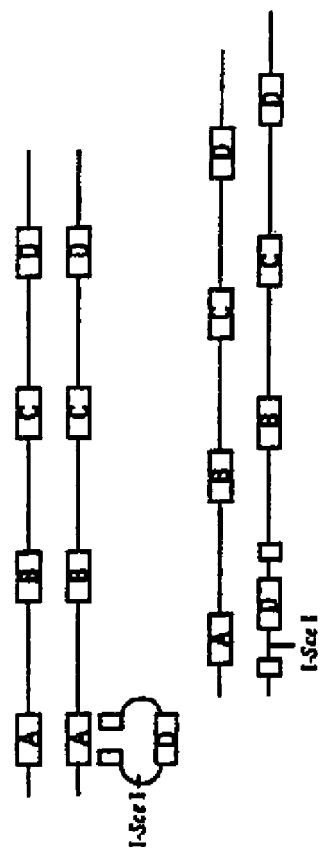


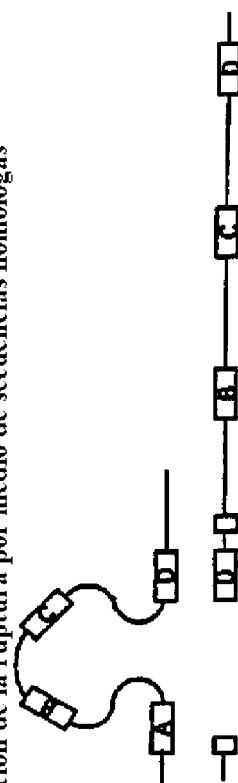
FIGURA 30

1 Inserción del sitio de I-Sce I por medio de reemplazo clásico de genes



2 Escisión específica por medio de la enzima I-Sce I y

reparación de la ruptura por medio de secuencias homólogas



3 Duplicación de la totalidad del locus

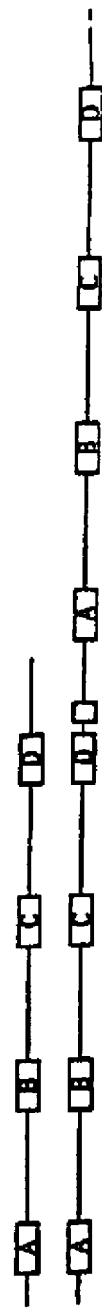
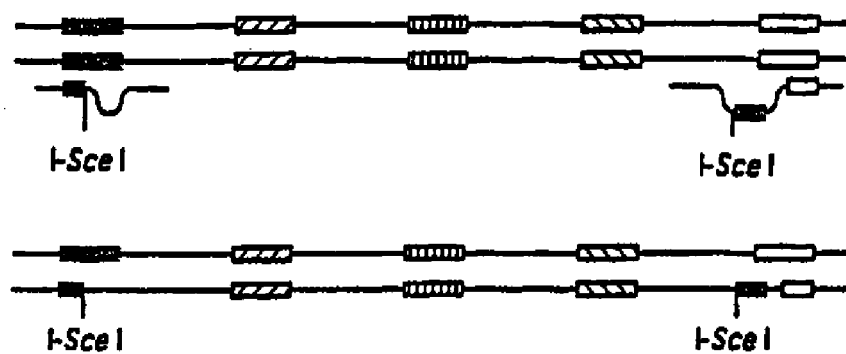
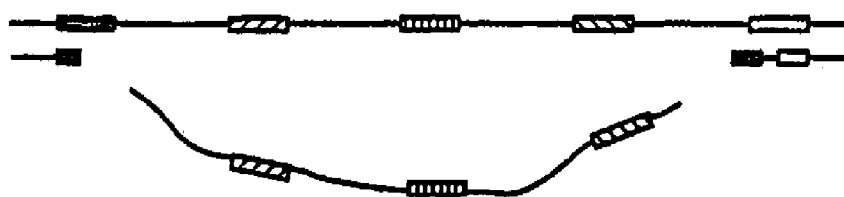


FIGURA 31

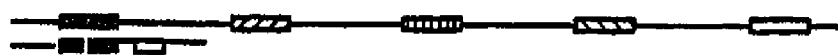
1 Inserción de dos sitios de I-Sce I que flanquean el locus



2 Expresión de la enzima y escisión



3 Recombinación entre los dos extremos



4 supresión del locus

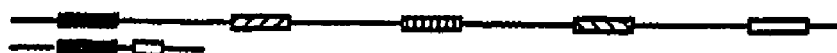


FIGURA 32

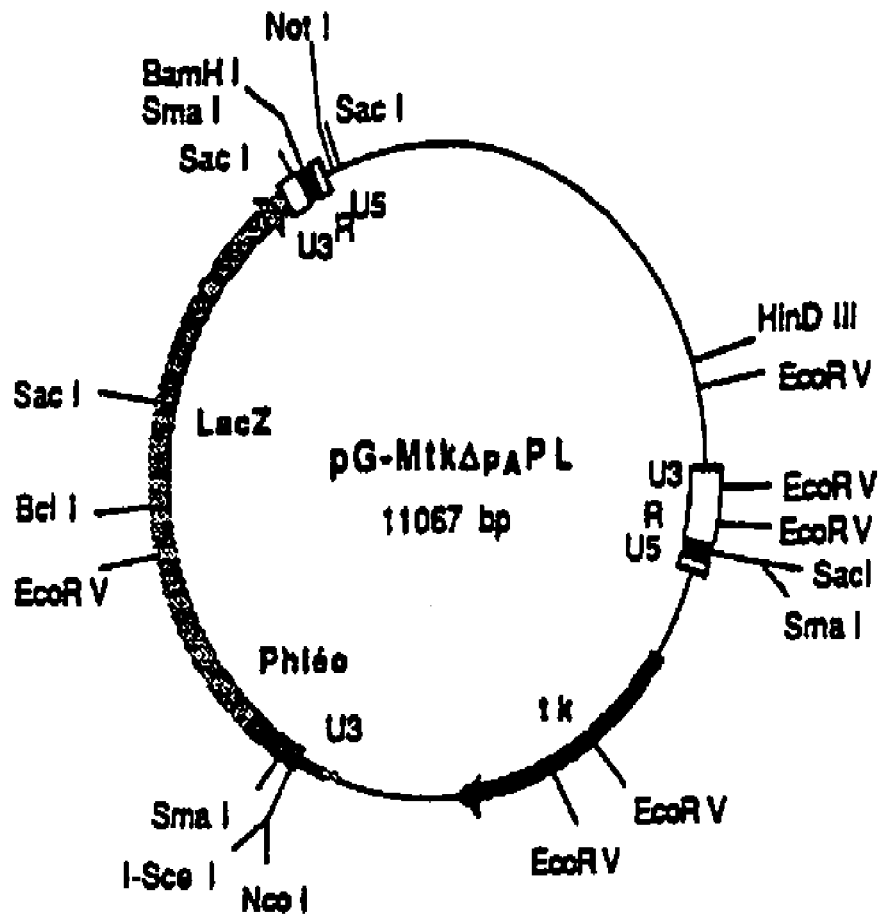


FIGURA 33