



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 284 555**

51 Int. Cl.:

C12N 15/82 (2006.01)

A01H 1/00 (2006.01)

C12N 15/80 (2006.01)

C12N 15/81 (2006.01)

A61K 48/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **00992065 .3**

86 Fecha de presentación : **13.12.2000**

87 Número de publicación de la solicitud: **1238090**

87 Fecha de publicación de la solicitud: **11.09.2002**

54 Título: **Transferencia optimizada de ADN-T y sus vectores.**

30 Prioridad: **16.12.1999 EP 99870264**
10.04.2000 US 195758 P

45 Fecha de publicación de la mención BOPI:
16.11.2007

45 Fecha de la publicación del folleto de la patente:
16.11.2007

73 Titular/es: **CropDesign N.V.**
Technologiepark 3
9052 Zwijnaarde-Gent, BE

72 Inventor/es: **Depicker, Anna;**
Mironov, Vladimir;
Terras, Franky;
Broekaert, Willem;
De Buck, Sylvie y
De Wilde, Chris

74 Agente: **Arias Sanz, Juan**

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

Transferencia optimizada de ADN-T y sus vectores.

5 **Campo de la invención**

La presente invención se refiere al campo de la biología molecular. Más particularmente, describe procedimientos para la incorporación de ADN exógeno al genoma de células eucariotas tales como células de plantas. Los aspectos concretos de la presente invención residen en el diseño de las repeticiones en el extremo que flanquean al ADN-T transferido mediante agrobacterias de manera que la trans-lectura hacia el extremo izquierdo del ADN-T se reduce significativamente y/o que el ADN del esqueleto del vector integrado se puede eliminar fácilmente. Así, se incrementa la frecuencia de obtención de células eucariotas transgénicas que contienen solamente el ADN-T.

Antecedentes de la invención

Se han obtenido variedades de plantas mejoradas mediante cruzamientos “clásicos” desde que el hombre cambió la existencia nómada por un asentamiento permanente. En la historia más reciente, los científicos comenzaron a desentrañar el comportamiento del material genético durante los cruzamientos y los cultivadores de plantas pudieron beneficiarse y aún se benefician, del conocimiento contenido en las leyes de Mendel que predicen la distribución de un rasgo genético dado en la descendencia de un cruzamiento. Con la llegada de la biología molecular de plantas, los cultivadores de plantas pueden insertar con una precisión cada vez mayor, nuevos genes quiméricos en el genoma de una planta. Actualmente están disponibles una variedad de técnicas para mediar en la transformación genética de plantas incluyendo los agrolísticos, microinyección, electroporación, transferencia directa de genes, y bombardeo con partículas recubiertas de ADN. Un sistema de transformación de plantas preferido y usado ampliamente hace uso de la bacteria del suelo *Agrobacterium* (Zupan y Zambryski 1995, Gelvin 1998a, Gheysen y col. 1998). Actualmente, el *Agrobacterium* no sólo se usa para transformar plantas sino también para transformar levaduras, mohos y hongos filamentosos (Bundock y col. 1995, de Groot y col. 1998, Gouka y col. 1999, WO 98/45455). Además se ha demostrado que los componentes de la producción del ADN-T y la maquinaria de transferencia del *Agrobacterium* son útiles para importar ADN a los núcleos de células de mamífero, abriendo perspectivas para el uso de estos componentes en terapia génica (Ziemienowicz y col. 1999). El *Agrobacterium* transfiere al núcleo de la célula eucariota todo el ADN localizado en el ADN-T. Este ADN-T es parte del plásmido tipo natural Ti (en el caso de *Agrobacterium tumefaciens*) o del plásmido tipo natural Ri (en el caso de *A. rhizogenes*). El ADN-T tipo natural porta los genes que provocan, después de la integración en el genoma de la planta, tumores de agalla del cuello o el síndrome de la raíz pilosa en caso de infección con *A. tumefaciens* o *A. rhizogenes*, respectivamente. También están localizados en los plásmidos Ti o Ri tipo natural los genes *vir* (genes de virulencia) que se activan mediante compuestos fenólicos de la planta. Los productos de los genes *vir* son responsables de la transferencia del ADN-T al genoma eucariota. Para fines de transformación, el ADN-T se desactiva (es decir, se eliminan todos los genes que provocan enfermedades) y los genes *vir* se suministran en *trans* sobre un plásmido ayudante (el ADN-T que engloba el gen(es) heterólogo está entonces localizado sobre un segundo vector de transformación binaria de la planta) o en *cis* en el caso de un vector de transformación de la planta co-integrado. Los genes heterólogos de interés se clonan entre medias de las dos secuencias centrales de extremos imperfectos del ADN-T de 22 pb (en el caso de plásmidos Ti de la octopina) o de 25 pb (en el caso de plásmidos Ti de la nopalina) que constituyen el extremo derecho (ED) y el extremo izquierdo (EI), que son los únicos elementos en *cis* necesarios para dirigir el procesamiento del ADN-T. Las secuencias centrales terminales del ED y del EI están organizadas como repeticiones imperfectas.

Las proteínas VirD1 y VirD2 producen un corte de cadena sencilla entre la tercera y cuarta base en la cadena inferior de cada repetición en el extremo (Yanofsky y col. 1986). Mayores niveles de VirD1 y VirD2 potencian la producción de complejos de ADN-T dentro del *Agrobacterium* y dan como resultado una mayor eficacia en la transformación de la planta (Wang y col. 1990).

Durante muchos años se ha creído que solamente el ADN entre las repeticiones, el ADN-T, y no el ADN del vector externo al ADN-T se transfería a la célula eucariota. No obstante, una caracterización reciente y más detallada de los insertos de ADN en plantas transgénicas demuestran que también las secuencias del esqueleto del vector se integran muy frecuentemente en el genoma de la planta (Martineau y col. 1994, Ramanathan y Veluthambi 1995, Cluster y col. 1996, van der Graaff y col. 1996, Kononov y col. 1997, Wenck y col. 1997, Wolters y col. 1998).

Los autores de la presente invención han descubierto previamente que la frecuencia de integración de las secuencias del vector no está influenciada por las especies de plantas o el procedimiento de transformación usado. Esto es coherente con la visión de que la transferencia de secuencias del esqueleto del vector es la consecuencia de la trans-lectura más allá del EI, un proceso que se produce dentro de las células de *Agrobacterium* y está determinado de la forma más probable por factores dentro de estas células. No obstante, se debe advertir que otros han informado de la integración del esqueleto del vector en el 33% de los transformantes de *Arabidopsis* obtenidos mediante transformación radicular y hasta el 62% de transformantes obtenidos mediante infiltración sobre vacío (Wenck y col. 1997). Esto implica que el procedimiento de transformación usado podría ser otro factor que influye en la frecuencia de la integración del esqueleto del vector. Se ha informado de que la integración de secuencias del esqueleto del vector se produce en numerosas especies de plantas incluyendo *Petunia* (Virts y Gelvin 1985, Cluster y col. 1996), *Arabidopsis* (van der Graaff y col. 1996, Wenck y col. 1997), tabaco (Ramanathan y Veluthambi 1995, Kononov y col. 1997, Wenck y col. 1997), y patata (Wolters y col. 1998). La integración del esqueleto del vector aparentemen-

te es independiente del tipo de cepa de *Agrobacterium* usada para la transformación de la planta (Kononov y col. 1997).

Los inventores han analizado previamente diferentes series de transformantes respecto a la presencia de secuen-
 5 cias del esqueleto del vector mediante el uso de reacciones de PCR específicas y análisis de transferencia de ADN
 en gel. Se evaluaron tres procedimientos de transformación diferentes en dos especies de plantas diferentes, a saber,
 transformación radicular y folicular de *Arabidopsis thaliana* y transformación protoplastática y folicular de *Nicotiana*
tabacum. Por último, se evaluó la influencia del tipo replicón, los replicones ColE1 y pVS1. Los resultados demostra-
 10 ron que ni las especies de plantas ni el tipo de explante usado para la transformación, el tipo de replicón o la selección
 tienen una influencia fundamental sobre la frecuencia con la cual se produce la integración de las secuencias del vec-
 tor. En el pasado, se postuló que esta transferencia del ADN del vector que no pertenece al ADN-T podría ser el
 resultado de la trans-lectura en el extremo izquierdo, que evitaría la terminación normal de la transferencia del ADN-
 T. Alternativamente, la transferencia de ADN podría comenzar en el extremo izquierdo y proseguir hacia el extremo
 derecho (Ramanathan y Veluthambi, 1995; van der Graaff y col., 1996). No obstante, en las plantas transgénicas des-
 15 critas anteriormente se observó que muchas contenían secuencias del esqueleto del vector unidas al extremo izquierdo
 así como uniones del vector con el extremo derecho del ADN-T. Las transferencias de ADN en gel indican que en
 la mayoría de estas plantas se integra la secuencia completa del vector. Por tanto, se postuló que la integración en
 el genoma de la planta de secuencias completas del esqueleto del vector puede ser el resultado de una transferencia
 conjugativa iniciada en el extremo derecho y seguida posteriormente con la copia en los extremos izquierdo y derecho,
 20 denominada trans-lectura. Este modelo implica que el extremo izquierdo no es reconocido con frecuencia como sitio
 de iniciación para la transferencia de ADN y que el extremo derecho no es reconocido frecuentemente como sitio de
 terminación para la transferencia de ADN. Estas observaciones concuerdan con los resultados de trabajos previos que
 demuestran que la región del extremo derecho es intrínsecamente más activa que la región del extremo izquierdo en la
 promoción de la transformación del ADN-T (Jen y Chilton 1986a,b, Caplan y col. 1985). A partir de todos los datos
 25 disponibles, se puede concluir que la formación de la cadena T comienza mucho más frecuentemente en el extremo
 derecho que en la región del extremo izquierdo.

En el futuro, será de suma importancia prevenir o curar la integración del esqueleto del vector como consecuencia
 de la transformación mediada por *Agrobacterium*. En primer lugar, las autoridades reguladoras están demandando que
 30 las plantas transgénicas puestas en circulación en el mercado ordinario estén exentas de secuencias del esqueleto del
 vector. Tales secuencias del esqueleto pueden portar orígenes de replicación bacterianos, genes bacterianos de resis-
 tencia a antibióticos y posiblemente una serie de otros genes (exógenos). La misma preocupación rigurosa también será
 expresada por los consumidores que se están informando cada vez más de tales riesgos potenciales asociados a la bio-
 tecnología de plantas. En segundo lugar, también desde un punto de vista científico es deseable no tener la integración
 35 del esqueleto del vector en el genoma de plantas. Tales secuencias pueden influir en la expresión transgénica (Iglesias
 y col. 1997, Matzke y Matzke 1998, Jakowitsch y col. 1999). También es probable que la integración del esqueleto del
 vector interfiera con experimentos de marcaje del ADN-T. Los marcadores pueden ser considerablemente más largos
 de lo esperado (Martineau y col. 1994) y la integración del esqueleto del vector también podría ser la explicación para
 el hecho de que en un gran porcentaje de ADN-T marcado en plantas *Arabidopsis* el ADN-T no se co-segrega con un
 40 fenotipo mutante (Errampali y col. 1991, Feldmann 1991, Koncz y col. 1992, van Lijsebettens y col. 1991).

Aunque el fenómeno de la integración ocasional del esqueleto del vector ya se encontró en 1982 por Ooms y
 col., no fue hasta 1995 cuando se sugirió una primera solución por Ramanathan y Veluthambi (1995): "... se pueden
 45 construir nuevos vectores binarios con señales de "detención de la transferencia" adyacentes al extremo izquierdo".
 Desde entonces, se ha intentado entender el mecanismo de integración del esqueleto del vector. Una posible razón fue
 descrita por Wenck y col. (1997): "... el corte ineficaz puede ser debido a bajas cantidades de proteínas de virulencia,
 principalmente VirD2". No obstante, sólo muy recientemente se han descrito procedimientos que previenen la trans-
 lectura en los extremos del ADN-T (WO 99/01563) por Hanson y col. (1999). Estos procedimientos se basan en la
 50 inclusión de secuencias fuera de los extremos. Estas secuencias son genes que codifican para compuestos tóxicos o
 secuencias capaces de interactuar con proteínas que se unen al ADN o secuencias que están enriquecidas en nucleótidos
 G+C. Un inconveniente importante de los procedimientos descritos en el documento WO 99/01563 de Hanson y col.
 (1999) es que se previene la regeneración de los transformantes de plantas que portan más que la región de ADN-T
 en su genoma. Es imaginable que tales procedimientos mermen la eficacia de la transformación global, es decir, se
 55 obtendrá un menor número de transformantes a partir de un experimento de transformación dado. De hecho Hanson
 y col. (1999) informaron de que la eficacia de transformación del tabaco cayó hasta un 30%. No obstante, Hanson
 y col. (1999) describen su aproximación como una herramienta útil para la eliminación de secuencias no ADN-T de
 individuos transgénicos.

La presente invención describe una solución al problema técnico de la integración no deseada del esqueleto del
 60 vector y proporciona ventajas sobre procedimientos existentes.

Resumen de la invención

La invención describe vectores de transformación que comprenden un ADN-T con extremos de ADN-T flanquean-
 65 tes. Los vectores de ADN-T se caracterizan porque están modificados de manera que permiten la transformación
 genética de una célula eucariota sólo con el ADN-T y no con secuencias del esqueleto del vector. Esto se realiza
 previniendo la transferencia de las secuencias del esqueleto del vector al genoma de una célula eucariota o curando
 de secuencias del esqueleto del vector transferidas al genoma de una célula eucariota. Dichos vectores de ADN-T

ES 2 284 555 T3

modificados permiten un procesamiento eficaz del extremo izquierdo mediante el complejo de corte que implica al menos a la VirD1 y a la VirD2 o permiten la escisión de secuencias del esqueleto del vector transferidas.

La presente solicitud describe vectores de ADN-T optimizados que incluyen una modificación del extremo derecho del ADN-T que comprende una única secuencia central del extremo derecho flanqueada por una región externa del extremo derecho y/o una modificación, incluyendo la multiplicación, del extremo izquierdo del ADN-T diseñados como:

- a) una única secuencia central del extremo izquierdo flanqueada por una región externa natural del extremo izquierdo y una región proximal en el extremo izquierdo intra-ADN-T con una longitud de 10 a 100 pb, preferentemente 20-100 pb, y que está enriquecida en el número de residuos de A y T, estando el porcentaje de residuos de AT preferentemente entre el 60 y el 85%, estando más preferentemente entre el 64 y el 80%, estando de la forma más preferida en el 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 78 ó 79%, siempre que dicha región proximal en el extremo izquierdo intra-ADN-T no sea la región interna del extremo izquierdo de tipo octopina natural correspondiente, o
- b) una única secuencia central del extremo izquierdo flanqueada solamente por una región interna natural del extremo izquierdo, o
- c) una única secuencia central del extremo izquierdo; o
- d) una repetición en tándem de varias secuencias centrales del extremo izquierdo flanqueada solamente por una región externa natural del extremo izquierdo, con la repetición en tándem que contiene preferentemente 2-3 secuencias centrales del extremo izquierdo pero que no excluye un número de copias superior de la repetición imperfecta central del extremo izquierdo y con dichas secuencias centrales del extremo izquierdo repetidas en dicho tándem que están separadas por una secuencia de al menos 10-20 pb que opcionalmente porta codones de terminación en los tres marcos de lectura y en ambas direcciones; o
- e) una región en el extremo izquierdo de tipo nopalina integral adyacente a y aguas abajo o aguas arriba de la región en el extremo izquierdo de tipo octopina integral.

La aplicación describe adicionalmente vectores de transformación de ADN-T optimizados con secuencias de ADN adicionales fuera de las repeticiones centrales del extremo de ADN-T que permiten la eliminación después de la transformación de secuencias del esqueleto del vector integrado. Dichas secuencias de ADN adicionales modifican, incluyendo la multiplicación, las regiones del extremo de ADN-T o sus partes y comprenden:

- f) sitios de recombinación organizados como repeticiones aguas abajo de la secuencia central del extremo izquierdo y aguas arriba de la secuencia central del extremo derecho; o
- g) dichas secuencias de ADN de (f) modificadas adicionalmente mediante la adición de una segunda copia de una región en el extremo izquierdo situada aguas arriba de y preferentemente adyacente a la única región externa del extremo derecho y dicho sitio de recombinación aguas arriba de la secuencia central de dicha segunda región en el extremo izquierdo; o
- h) dichos sitios de recombinación de (f) con un gen que codifica una recombinasa, aguas abajo de dicho sitio de recombinación aguas abajo de la secuencia central del extremo izquierdo, y preferentemente, cuando está presente, adyacente a y aguas abajo de la región externa del extremo izquierdo; o
- i) una secuencia de ADN localizada aguas abajo de la región en el extremo izquierdo, dicha secuencia de ADN que comprende un gen que codifica una recombinasa flanqueado por repeticiones de sitios de recombinación como se ha definido en (h) y que comprende adicionalmente una segunda copia de una región en el extremo izquierdo aguas abajo de dicha recombinasa; o
- j) la secuencia de ADN de (i) con sitios de recombinación adicionales organizados como repeticiones aguas abajo de la segunda secuencia central del extremo izquierdo y aguas arriba de la única secuencia central del extremo derecho.

En cualquiera de dichas modificaciones (f), (g), (h), (i) o (j), dichos sitios de recombinación están localizados adyacentes a y aguas abajo y/o aguas arriba de las secuencias centrales del extremo izquierdo y/o derecho o están separados de las secuencias centrales del extremo izquierdo y/o derecho por una secuencia de al menos 10-20 pb de longitud que opcionalmente portan codones de terminación en los tres marcos de lectura y en ambas direcciones.

En cualquiera de dichas modificaciones (f), (g), (h), (i) o (j), dichos sitios de recombinación son sitios de recombinación específicos de sitio dispuestos como repeticiones directas o secuencias terminales de transposones dispuestas como repeticiones invertidas y dicho gen que codifica una recombinasa es un gen que codifica una recombinasa específico de sitio o un gen de transposasa, respectivamente.

Una primera forma de realización de la invención se refiere a un vector de transformación de ADN-T que comprende ADN-T con extremos de ADN-T modificados flanqueantes a izquierda y derecha, caracterizado porque el extremo derecho de ADN-T modificado consta de una única secuencia central del extremo derecho flanqueada por una región externa del extremo derecho, y el extremo izquierdo de ADN-T modificado consta de:

- (i) una única secuencia central del extremo izquierdo, una región externa del extremo izquierdo y una región proximal en el extremo izquierdo de la región interna con una longitud preferible de 10 a 100 pb, caracterizada adicionalmente porque dicha región proximal en el extremo izquierdo de la región interna está enriquecida en el número de residuos de A y T, estando el porcentaje de residuos de AT entre el 60 y el 85%,
- (ii) una repetición de los extremos izquierdos del ADN-T en el que dicha repetición comprende al menos un extremo de ADN-T de tipo nopalina y al menos un extremo izquierdo de ADN-T de tipo octopina, caracterizada adicionalmente porque dicho extremo izquierdo de ADN-T modificado comprende al menos una secuencia central del extremo izquierdo, o
- (iii) una región externa del extremo izquierdo y una repetición de al menos dos secuencias centrales del extremo izquierdo separadas por una secuencia de al menos 10-20 pares de bases.

Según la presente invención, en el vector anteriormente mencionado, dicha secuencia de al menos 10-20 pares de bases que separa dichas secuencias centrales del extremo izquierdo puede portar codones de terminación en los tres marcos de lectura y en ambas direcciones.

Además, el extremo izquierdo de ADN-T modificado puede constar de una repetición de extremos izquierdos de ADN-T que comprende al menos una región en el extremo izquierdo de tipo nopalina integral adyacente a y aguas abajo o aguas arriba de la región en el extremo izquierdo de tipo octopina integral.

Una segunda forma de realización de la invención se refiere a un vector de la invención como se ha descrito en la primera forma de realización, pero en la que dichos extremos de ADN-T modificados además pueden comprender sitios de recombinación aguas abajo de la secuencia central del extremo izquierdo y sitios de recombinación aguas arriba de la secuencia central del extremo derecho, dichos sitios de recombinación que están organizados como repeticiones.

Según la presente invención el vector de la segunda forma de realización además puede comprender una segunda copia de una región en el extremo izquierdo aguas arriba de la única región externa del extremo derecho y en la que dicho sitio de recombinación está situado aguas arriba de la secuencia central de dicha segunda región en el extremo izquierdo.

Una tercera forma de realización de la invención se refiere a un vector de la invención como se ha descrito en la segunda forma de realización, que además puede comprender un gen que codifica una recombinasa localizado aguas abajo de dicho sitio de recombinación aguas abajo de dicha secuencia central del extremo izquierdo. Dicho gen que codifica una recombinasa puede estar situado adyacente a y aguas abajo de la región externa del extremo izquierdo. Según la presente invención, dicho gen que codifica una recombinasa puede estar flanqueado por repeticiones de sitios de recombinación y dicho vector además puede comprender una segunda copia de una región en el extremo izquierdo localizada aguas abajo de dicho gen que codifica una recombinasa flanqueado por dichos sitios de recombinación. El vector además puede comprender sitios de recombinación adicionales organizados como repeticiones aguas abajo de la segunda secuencia central del extremo izquierdo y aguas arriba de la única secuencia central del extremo derecho. En todos los vectores de la tercera forma de realización de la presente invención, los sitios de recombinación pueden estar localizados adyacentes a y aguas abajo y/o aguas arriba de las secuencias centrales del extremo izquierdo y/o derecho o pueden estar separados aguas abajo y/o aguas arriba de las secuencias centrales del extremo izquierdo y/o derecho por una secuencia de al menos 10-20 pb de longitud. Según la presente invención, la secuencia de al menos 10-20 pb de longitud puede portar codones de terminación en los tres marcos de lectura y en ambas direcciones.

Una cuarta forma de realización de la invención se refiere a un vector de la invención como se ha descrito en la tercera forma de realización, en el que dichos sitios de recombinación organizados como repeticiones se pueden definir como sitios de recombinación específicos de sitio organizados como repeticiones directas o como secuencias terminales de transposones organizadas como repeticiones invertidas; y en las que dicho gen que codifica una recombinasa se puede definir como un gen que codifica una recombinasa específico de sitio o un gen que codifica una transposasa, respectivamente.

Otra forma de realización de la invención incluye vectores de transformación en los que se aplica cualquiera de dichas modificaciones como se ha definido en (i) a (iii), o se aplican en combinación. Estos vectores de transformación comprenden vectores usados en la transformación mediada por *Agrobacterium* incluyendo vectores de transformación binarios, vectores de transformación de tipo co-integrado, vectores de transformación de tipo súper-binario, vectores de transformación derivados de Ri así como vectores que portan ADN-T usados en transformación agrolística o terapia génica.

La presente solicitud describe un procedimiento para obtener células eucariotas transgénicas transformadas solamente con ADN-T evitando la transferencia de secuencias del esqueleto del vector usando dichos vectores de transformación optimizados que contienen cualquiera de dichas modificaciones (a) a (i).

5 En la presente solicitud además se describe un procedimiento para la obtención de células eucariotas transgénicas transformadas solamente con el ADN-T permitiendo la curación de dichas células transformadas que contienen secuencias del esqueleto del vector usando dichos vectores de transformación optimizados que contienen dicha secuencia de ADN de (f) o (g) en combinación con el suministro de una recombinasa específica de sitio o transposasa, o de dichas secuencias de ADN (h), (i), o (j) eventualmente en combinación con el suministro de una recombinasa específica de sitio o una transposasa. Con curación se quiere decir la eliminación de las secuencias del esqueleto del vector de transformación que posiblemente se originan en dicho vector sin suprimir el evento de transformación del ADN-T.

15 También es parte de la invención un procedimiento para obtener plantas, levaduras, mohos u hongos filamentosos transgénicos que consiste en la transformación mediada por *Agrobacterium* de dichas plantas, levaduras, mohos u hongos filamentosos con un vector de transformación escogido del grupo de:

- a) un vector de transformación según la primera forma de realización de la invención; o
- b) un vector de transformación definido según la segunda forma de realización de la invención en combinación con el suministro de una recombinasa o transposasa, para curar dichas células transformadas resultantes de secuencias del esqueleto del vector integradas que posiblemente se originan en dicho vector; o
- c) un vector de transformación definido en la tercera forma de realización de la invención para curar dichas células transformadas resultantes de secuencias del esqueleto del vector integradas que posiblemente se originan en dicho vector, opcionalmente en combinación con el suministro de una recombinasa o transposasa; o
- d) un vector de transformación según la cuarta forma de realización de la invención.

30 Otro procedimiento de la invención es un procedimiento para prevenir la integración de secuencias del esqueleto del vector en una célula transformada mediada por *Agrobacterium* que comprende el uso de un vector según cualquiera de las formas de realización uno a cuatro de la invención.

35 Dicho procedimiento puede comprender adicionalmente el incremento de la producción por *Agrobacterium* del complejo de corte que al menos implica a VirD1 y VirD2. Dichos componentes del complejo de corte están codificados por el operón *virD* (de tipo octopina o tipo nopalina). Al menos una copia extra del locus VirD puede estar integrada en entidades de ADN cromosómico y/o extracromosómico contenidas y mantenidas dentro de una cepa de *Agrobacterium*. Dicha copia extra del locus VirD así se puede seleccionar entre el locus VirD de tipo octopina o el locus VirD de tipo nopalina. Para reducir la trans-lectura en el extremo izquierdo de ADN-T, dicho procedimiento se puede aplicar solo o junto con vectores de transformación que contienen extremos de ADN-T modificados.

40 En la presente invención además están comprendidas todas las combinaciones de procedimientos y/o modificaciones del vector de ADN-T de la invención para prevenir o curar la integración de secuencias del esqueleto del vector con cualquier otro procedimiento y/o modificaciones del vector de ADN-T aplicadas para prevenir o curar la integración de secuencias del esqueleto del vector.

45 Las células hospedadoras recombinantes que contienen dichos vectores de transformación modificados según la presente invención, como bacterias, preferentemente *Agrobacterium tumefaciens* también constituyen la invención.

50 También están comprendidas en la invención todas las combinaciones del sistema de recombinación específico de sitio o el sistema de recombinación mediado por transposasa de la presente invención para curar la integración de secuencias del esqueleto del vector con el mismo o cualquier otro sistema de recombinación específico de sitio o sistema de recombinación mediado por transposasa usado para cualquier otro propósito.

55 También están comprendidos en la invención los procedimientos de transformación agrolísticos de una célula eucariota que usa un vector de ADN-T de la presente invención.

También están comprendidos en la invención los procedimientos de terapia génica que usan un vector de ADN-T modificado según las alteraciones como las definidas en los vectores como se ha mencionado anteriormente.

60 Las células de plantas o las plantas transgénicas obtenibles mediante un procedimiento de transformación mediado por *Agrobacterium* como se ha definido anteriormente, parte de dicha planta o su progenie también constituyen la presente invención. Las levaduras, mohos u hongos filamentosos transgénicos obtenibles mediante un procedimiento de transformación mediado por *Agrobacterium* como se ha definido anteriormente también constituyen la presente invención.

65 Con los anteriores y otros objetos, ventajas y características de la invención que serán evidentes más adelante, la naturaleza de la invención se puede entender más claramente en referencia a la siguiente descripción detallada de las formas de realización preferidas de la invención y a las reivindicaciones anexas.

Descripción detallada de la invención

La transformación mediada por *Agrobacterium* o la transformación agrolística de plantas, levaduras, mohos u hongos filamentosos se basa en la transferencia de parte de las secuencias del vector de transformación, denominado ADN-T, al núcleo y en la integración de dicho ADN-T en el genoma de dicha eucariota.

Con “*Agrobacterium*” se quiere decir un miembro de las *Agrobacteriaceae*, más preferentemente *Agrobacterium* o *Rhizobacterium* y de la forma más preferida *Agrobacterium tumefaciens*.

Con “ADN-T”, o ADN transferido, se quiere decir que parte del vector de transformación flanqueado por los extremos del ADN-T que es cortado, después de la activación de los genes *vir* de *Agrobacterium*, en los extremos del ADN-T y es transferido como una única cadena de ADN al núcleo de una célula eucariota. La producción de la cadena de ADN-T es el resultado de un evento de procesamiento del ADN iniciado por un corte específico de sitio en una molécula de doble cadena. Existe una familia de reacciones de procesamiento del ADN interrelacionadas en diferentes sistemas procariotas. En cada caso, el sustrato es una molécula de ADN circular súperenrollada y la proteína de corte reconoce una secuencia corta, escinde en un sitio conservado (posiciones conservadas del sitio de corte y región consenso localizada en 3' del corte), y se une covalentemente al extremo 5' de la cadena cortada. Los sistemas de procesamiento de ADN que pertenecen a esta familia incluyen, aparte de la formación de ADN-T en y la transferencia del ADN-T a la célula eucariota por *Agrobacterium*, (1) la iniciación de la replicación de ADN conjugativa y la posterior transferencia de plásmidos de bacterias Gram-negativa, (2) la iniciación de síntesis de más cadena durante la replicación en círculo rodante del fago relacionado con ϕ X174, y (3) la iniciación de la replicación en círculo rodante en ciertos plásmidos de bacterias Gram-positiva (Waters y Guiney 1993). El evento de procesamiento del ADN está catalizado por un complejo relaxosoma oligoproteínico. Una de las proteínas en este complejo, la relaxasa, es la enzima clave que cataliza la escisión del ADN específica de sitio y cadena (Pansegrau y Lanka 1996). En el caso de la transferencia conjugativa de plásmidos, la relaxasa se une covalentemente al 5'-término del ADN a transferir y no es capaz de producir un segundo corte necesario para la transferencia de una única copia del ADN. Se piensa que al menos es necesario un dímero de la relaxasa para la transferencia de una única copia del ADN (Pansegrau y Lanka 1996).

En el caso de *Agrobacterium*, es necesario un mínimo de dos proteínas para el corte y la posterior transferencia del ADN-T. Estas proteínas del relaxosoma son VirD1 y VirD2. VirD1 es una topoisomerasa de tipo I sin especificidad de secuencia (Ghai y Das 1989). La proteína VirD2 es, también en base a la homología de secuencia, la relaxasa de *Agrobacterium* y actúa como endonucleasa de corte (Pansegrau y col. 1994). Durante el proceso de corte, la VirD2 se une covalentemente a ambos extremos 5' del ADN-T procesado en el extremo derecho y al resto del plásmido de ADN-T en el extremo izquierdo (Durrenberger y col. 1989, Young y Nester 1988). Ambos extremos izquierdo y derecho del ADN-T son reconocidos por VirD1 y VirD2. Se han demostrado interacciones fuertes entre VirD1 y VirD2 y entre proteínas VirD2 y se ha sugerido que el desplazamiento de la cadena de ADN-T (mediante la replicación del ADN) se detiene produciendo un segundo corte que depende de la interacción VirD2-VirD2. Esto implica que complejos de relaxosoma separados que incluyen al menos a VirD1 y VirD2 podrían estar involucrados en el procesamiento de cada uno de los dos extremos del ADN-T (Relic y col. 1998).

Con “extremos de ADN-T”, “región en el extremo de ADN-T” o “región terminal” se quiere decir cualquiera del extremo del ADN-T derecho (ED), también denominado “extremo derecho”, o extremo del ADN-T izquierdo (EI) también denominado “extremo izquierdo”. Tal extremo comprende una secuencia central flanqueada por una región interna del extremo como parte del ADN-T que flanquea el extremo y/o una región externa del extremo como parte del extremo que flanquea al esqueleto del vector. Las secuencias centrales comprenden 22 pb en el caso de vectores de tipo octopina y 25 pb en el caso de vectores de tipo nopalina. Las secuencias centrales en la región en el extremo derecho y en la región en el extremo izquierdo forman repeticiones imperfectas. Las secuencias centrales del extremo son indispensables para el reconocimiento y el procesamiento por el complejo de corte de *Agrobacterium* que consta de al menos VirD1 y VirD2. Las secuencias centrales que flanquean un ADN-T son suficientes para promover la transferencia de dicho ADN-T. No obstante, la eficacia de la transformación usando vectores de transformación que portan dicho ADN-T únicamente flanqueado por dichas secuencias centrales es baja. Las regiones interna y externa del extremo son conocidas por modular la eficacia de la transferencia de ADN-T (Wang y col. 1987). Se ha caracterizado un elemento que potencia la transferencia de ADN-T y se encuentra en la región externa del extremo derecho y se denomina multiplicador (Peralta y col. 1986, van Haaren y col. 1987). Las secuencias centrales son como sigue (in: parte de la secuencia de ADN-T; ex: parte de la secuencia del esqueleto del vector; secuencias centrales indicadas en mayúsculas en negrita y subrayado; Gielen y col. 1984, 1999):

ED de tipo octopina: in-tgatgctgact**GGCAGGATATATACCGTTGTAAT**ttgagctcgt-ex (ID SEC N° 1)

EI de tipo octopina: ex-gcggcagcggc**GGCAGGATATATTCAATTGTAAAT**tgcttcgtg-in (ID SEC N° 2)

ED de tipo nopalina: in-tatcagtggt**TGACAGGATATATTGGCGGGTAAAC**ctaagagaa-ex (ID SEC N° 3)

ID de tipo nopalina: ex-ggctggctgg**TGGCAGGATATATTGTGGTGTAAAC**aaattgacg-in (ID SEC N° 4)

Con “región integral del extremo” se quiere decir un extremo de ADN-T de origen natural que comprende la secuencia central del extremo y ambas regiones interna y externa del extremo.

Con “vector de transformación de ADN-T” o “vector de ADN-T” se quiere decir cualquier vector que engloba una secuencia de ADN-T flanqueada por un extremo derecho e izquierdo de ADN-T que consta de al menos las secuencias centrales del extremo derecho e izquierdo, respectivamente, y usado para la transformación de cualquier célula eucariota.

Con “secuencia del esqueleto del vector” o “secuencias del esqueleto del vector” se quiere decir todo el ADN de un ADN-T que contiene un vector que queda fuera de los extremos del ADN-T y, más específicamente, fuera de los sitios de corte de las repeticiones imperfectas centrales del extremo. Con “vector” se quiere decir un vector de transformación o un vector de ADN-T mantenido de manera estable en cultivo bacteriano, por ejemplo, un cultivo de *Escherichia coli* o un cultivo de *Agrobacterium*.

La presente invención incluye construcciones de vectores de ADN-T optimizados de manera que se minimiza o se impide la integración del esqueleto del vector en el genoma de una célula eucariota o de manera que es posible la curación de las secuencias del esqueleto del vector integradas en una célula eucariota. Dichos vectores, cualquiera de sus derivados o cualquier vector que utiliza cualquiera de las modificaciones o cualquier combinación de modificaciones de la presente invención se puede usar como material de partida para la clonación de las secuencias de ADN de interés entre las dos repeticiones del extremo y para aplicaciones posteriores, tales como la transformación de, por ejemplo, una planta cultivada, una levadura o un hongo y tal como terapia génica.

Con “vector de ADN-T optimizado” se quiere decir un vector de ADN-T diseñado para reducir o suprimir la transferencia de secuencias del esqueleto del vector al genoma de una célula eucariota o para permitir la curación de secuencias del esqueleto del vector transferidas al genoma de una célula eucariota.

En los análisis realizados en el marco de la presente invención (véanse Ejemplos 1-3), se comparó la frecuencia de transferencia del esqueleto del vector en plantas transgénicas que se habían transformado con vectores de ADN-T con repeticiones en el extremo en el contexto de la secuencia de octopina natural, y con vectores de ADN-T sin la región en el extremo interna de las repeticiones en el extremo. Se encontró sustancialmente más integración del esqueleto del vector en la serie de plantas transformadas en las que se emplearon los vectores de ADN-T sin las secuencias internas de repetición en el extremo que en la serie de plantas transformadas en las que se usaron regiones completas interna y externa en el extremo. Además se observó que muchas plantas transgénicas contienen secuencias del esqueleto del vector unidas al extremo izquierdo del ADN-T así como uniones del vector con el extremo derecho del ADN-T (véase Ejemplo 2). Los análisis de transferencia de ADN en gel indican que en la mayoría de estas plantas está integrada la secuencia completa del vector (véase Ejemplo 3). Además, la frecuencia de integración de secuencias completas del esqueleto del vector no estaba influenciada por la presencia o ausencia de la región interna del extremo. Estos datos sugieren que la región interna del extremo derecho no contiene lo más probablemente determinantes importantes para el corte eficaz en el extremo izquierdo del ADN-T. Esta conclusión se corrobora adicionalmente por resultados previos obtenidos por Shaw y col. (1984) que indican que la supresión de la región interna del ED no influye en la eficacia de la transformación de la planta. Así lo más probable es que los determinantes para el corte eficaz en el extremo izquierdo del ADN-T residan dentro del propio EI. Se asume que la integración en el genoma de la planta de secuencias completas del esqueleto del vector es el resultado de una transferencia conjugativa iniciada en el extremo derecho y continuada posteriormente con la copia en los extremos izquierdo y derecho, denominado trans-lectura. La transferencia del esqueleto del vector también puede ser la consecuencia del reconocimiento del EI como punto de partida para la transferencia de ADN y continuar la copia en el ED hasta el EI (van der Graaff y col. 1996). Por tanto es importante identificar los elementos del EI involucrados en el corte eficaz en la repetición central del EI. Así, por ejemplo, la supresión parcial o completa de la región interna del extremo izquierdo y/o las regiones exteriores podría poner a la repetición central del extremo izquierdo en un contexto que potencia su afinidad y el reconocimiento por VirD1 y VirD2. Una observación llamativa con respecto a esto procede del análisis de las secuencias del extremo izquierdo (y derecho) de los vectores de transformación usados en experimentos descritos en los Ejemplos 2 y 3. En la Figura 1, se representan los porcentajes de residuos de A y T por 100 pb alcanzados alrededor de (y que no incluyen) las repeticiones centrales del EI y ED para ambos vectores de transformación con y sin el contexto de la región interna del extremo natural. La observación remarcable consiste en el hecho de que el porcentaje de residuos de AT de los 100 pb proximales al EI de la región interna natural (vector de ADN-T K en la Figura 1) es considerablemente superior al porcentaje de residuos de AT de los 100 pb proximales al EI del vector en el que se suprime la región interna natural (vector de ADN-T Hsb en la Figura 1), es decir, el 64% comparado con el 56%, respectivamente. Como se ha mencionado anteriormente, la trans-lectura del EI y la transferencia parcial del esqueleto del vector se produce mucho más frecuentemente cuando las plantas están transformadas con vectores de ADN-T Hsb que cuando se compara con la transformación de plantas con el vector de ADN-T K. Mecánicamente, es posible que este porcentaje inferior de residuos de AT atenúe considerablemente la actividad de corte del relaxosoma unido al EI que incluye al menos a VirD1 y VirD2. Simultáneamente, la maquinaria de replicación del ADN que lleva a cabo el desplazamiento de la cadena de ADN-T sería capaz, a frecuencias relativamente elevadas, de desplazar el relaxosoma del EI antes de que se haya realizado el corte. Una frecuencia incrementada de trans-lectura en el EI sería la consecuencia de ese proceso. Así, el porcentaje de residuos de AT de los 10-100 pb (o preferentemente 20-100 pb) proximales al EI y parte del ADN-T es un elemento candidato del EI que determina la eficacia del corte en la repetición central del EI. Para que la eficacia del corte sea suficientemente elevada, el porcentaje de residuos AT de dicha región proximal del EI tiene que ser suficientemente elevado, es decir, al menos el 60-85% aproximadamente, más preferentemente al menos el 64-80% aproximadamente. Aunque los 100 pb proximales fuera de la repetición central del ED del ADN-T contienen un porcentaje similar de residuos de AT a los 100 pb proximales en el interior de la repetición central del EI del ADN-T, es decir, el 58% comparado con el 56%, respectivamente (véase la Figura 1), el procesamiento correcto y altamente

eficaz en el ED no es problemático. Lo más probable es que esto sea debido a la influencia positiva de la secuencia multiplicadora presente en la región externa del ED (Peralta y col. 1986, van Haaren y col. 1987). No se debe excluir que proteínas auxiliares al complejo del relaxosoma se unan a la secuencia multiplicadora y potencien el proceso de corte en el ED. Esto se asemejaría de nuevo a la situación de transferencia del plásmido conjugativo entre bacterias en las que se han identificado tales proteínas auxiliares (Waters y Guiney 1993). No se debe excluir que proteínas auxiliares, no identificadas hasta el momento, estén implicadas en el procesamiento eficaz del EI.

La primera construcción del vector de ADN-T modificada para prevenir la transferencia del esqueleto del vector descrita anteriormente incluye una región proximal del EI intra-ADN-T enriquecida en residuos de A y T. No obstante, tal aproximación no será practicable para todas las construcciones de ADN-T. Así, en la Figura 2, se representan esquemáticamente una serie de construcciones del vector de ADN-T optimizadas adicionales. Los resultados de la transformación de plantas con la primera de las construcciones de la Figura 2 se describen en los Ejemplos 1-3. Las otras construcciones sustentan varias aproximaciones para obtener vectores de ADN-T optimizados. Las optimizaciones residen en modificaciones del diseño del extremo izquierdo enfocadas a incrementar la eficacia del corte del EI y que dan como resultado que la integración del esqueleto del vector en el genoma de una célula eucariota está muy minimizada o ausente. En dichas construcciones del vector de ADN-T optimizadas, la región del ED es constante, es decir, comprende la región externa del extremo derecho pero se ha suprimido la región interna del extremo derecho. Dicha región del extremo derecho contiene la secuencia central del ED de 25 pb y los 146 pb aguas arriba de la región externa del extremo derecho derivada del pTiAch5 (Gielen y col. 1984). Esta región externa contiene la secuencia multiplicadora. Es claro que las regiones del extremo derecho en su contexto natural, es decir, que comprenden la región interna del extremo derecho también se pueden usar en la presente invención. Los extremos derechos de los cuales se han eliminado las regiones internas del extremo tienen la ventaja, no obstante, de que la transferencia a la planta de ADN "exógeno" sin una función clara y como parte del ADN-T está limitada adicionalmente. Además, se ha descrito anteriormente que la ausencia de la región interna del ED es poco probable que influya en la eficacia del procesamiento del EI. Las regiones del EI en dichos vectores de ADN-T optimizados ejemplares generalmente proceden de plásmidos Ti de tipo octopina y todas contienen una secuencia central natural del extremo izquierdo. Los extremos de ADN-T de tipo nopalina normalmente proceden de plásmidos Ti tales como el pTiC58 (Gielen y col. 1999). En la Figura 2 y en el Ejemplo 4 se describen más detalles sobre las regiones terminales ejemplares usadas en los vectores de ADN-T optimizados ejemplares.

La presente solicitud describe vectores de ADN-T optimizados como se ejemplifican en la Figura 2 y que incluyen una modificación del extremo derecho del ADN-T que comprende una única secuencia central del extremo derecho flanqueada por una región externa del extremo derecho y/o una modificación, incluyendo la multiplicación, del extremo izquierdo del ADN-T diseñada como:

- a) una única secuencia central del extremo izquierdo flanqueada por una región externa natural del extremo izquierdo y una región proximal en el extremo izquierdo intra-ADN-T con una longitud de 10 a 100 pb, preferentemente 20-100 pb, y que está enriquecida en el número de residuos de A y T, estando el porcentaje de residuos de AT preferentemente entre el 60 y el 85%, estando más preferentemente entre el 64 y el 80%, estando de la forma más preferida en el 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 78 ó 79%, siempre que dicha región proximal en el extremo izquierdo intra-ADN-T no sea la región interna del extremo izquierdo de tipo octopina natural correspondiente, o
- b) una única secuencia central del extremo izquierdo flanqueada solamente por una región interna natural del extremo izquierdo, o
- c) una única secuencia central del extremo izquierdo; o
- d) una repetición en tándem de varias secuencias centrales del extremo izquierdo flanqueada solamente por una región externa natural del extremo izquierdo, con la repetición en tándem de contiene preferentemente 2-3 secuencias centrales del extremo izquierdo pero que no excluye un mayor número de copias de la repetición imperfecta central del extremo izquierdo y con dichas secuencias centrales del extremo izquierdo repetidas en dicho tándem que están separadas por una secuencia de al menos 10-20 pb que opcionalmente porta codones de terminación en los tres marcos de lectura y en ambas direcciones; o
- e) una región en el extremo izquierdo de tipo nopalina integral adyacente a y aguas abajo o aguas arriba de la región en el extremo izquierdo de tipo octopina integral.

La incorporación de regiones en el extremo izquierdo modificadas que presentan un corte eficaz en la secuencia central del EI en vectores de transformación usados habitualmente prevendrá la trans-lectura del EI. Una ventaja adicional de la disposición en tándem de las secuencias centrales del EI o de los extremos izquierdos integrales radica en el hecho de que la trans-lectura ocasional o la transferencia de ADN que comienza en la primera copia de dicha secuencia central del EI o en la región en el EI integral se detendrá en una copia adyacente de dicha secuencia central del EI o región en el EI integral en dicho tándem. En ambos casos, se reduce la frecuencia de transferencia de secuencias del esqueleto del vector y la integración de dichas secuencias en un genoma eucariota.

Con "aguas abajo" y "aguas arriba" se quiere decir cualquier secuencia localizada 5' y 3', respectivamente, de cualquier otra secuencia. Como referencia, la cadena de ADN-T se considera que comienza en su extremo 5' en el ED y termina en su extremo 3' en el EI.

Se prevé una optimización adicional de los vectores de transformación del ADN-T una vez que se conozca si las regiones internas y/o externas del extremo izquierdo son necesarias para evitar la integración del esqueleto del vector. Se pueden realizar varios derivados de supresión de dichas regiones internas y/o externas para determinar qué secuencia(s) específica en dichas regiones internas y/o externas del extremo son importantes en la determinación de la terminación de la transferencia de ADN-T. Una vez que se conozca esta secuencia(s), se puede incluir en más copias alrededor de la repetición del EI que da como resultado un corte más eficaz. También son objeto de la presente invención tales extremos izquierdos capaces de optimizar adicionalmente los vectores de transformación de ADN-T y dichos vectores que contienen dichos extremos izquierdos modificados adicionalmente.

La presente invención expresa la incorporación de cualquier modificación descrita por la presente invención en cualquier vector de ADN-T que comprende vectores de transformación binarios, vectores de transformación súper-binarios, vectores de transformación co-integrados, vectores de transformación derivados de Ri así como en vectores que portan ADN-T usados en transformación agrolística o terapia génica.

Con “vector de transformación binario” se quiere decir un vector de transformación de ADN-T que comprende:

a) una región de ADN-T que comprende al menos un gen de interés y/o al menos un marcador seleccionable activo en la célula eucariota a transformar; y

b) una región del esqueleto del vector que comprende al menos orígenes de replicación activos en *E. coli* y *Agrobacterium* y marcadores para la selección en *E. coli* y *Agrobacterium*.

Alternativamente, la replicación del vector de transformación binario en *Agrobacterium* depende de la presencia de un plásmido ayudante aparte. El vector binario pGreen y el plásmido ayudante pSoup forman un ejemplo de tal sistema como se describe en, por ejemplo, Hellens y col. (2000), Plant Mol. Biol. 42, 819-832, o como está disponible en el sitio de internet <http://www.pgreen.ac.uk>.

Los extremos de ADN-T de un vector de transformación binario pueden proceder de plásmidos Ti de tipo octopina o de tipo nopalina o de ambos. El ADN-T de un vector binario sólo se transfiere a una célula eucariota junto con un plásmido ayudante.

Con “plásmido ayudante” se quiere decir un plásmido que se mantiene de manera estable en *Agrobacterium* y que al menos porta el grupo de genes *vir* necesarios para permitir la transferencia del ADN-T.

Dicho grupo de genes *vir* puede proceder de plásmidos Ti de tipo octopina o de tipo nopalina o de ambos. Con “vector de transformación súper-binario” se quiere decir un vector de transformación binario que porta adicionalmente en la región del esqueleto del vector una región *vir* del plásmido Ti pTiBo542 de la cepa A281 súper-virulenta de *A. tumefaciens* (EP0604662, EP0687730). Los vectores de transformación súper-binarios se usan junto con un plásmido ayudante.

Con “vector de transformación co-integrado” se quiere decir un vector de ADN-T que comprende al menos:

a) una región de ADN-T que comprende al menos un gen de interés y/o al menos un marcador seleccionable activo en plantas; y

b) una región del esqueleto del vector que comprende al menos orígenes de replicación activos en *Escherichia coli* y *Agrobacterium*, y marcadores para la selección en *E. coli* y *Agrobacterium*.

Los extremos de ADN-T y dicho grupo de genes *vir* de dicho vector ADN-T pueden proceder de plásmidos Ti de tipo octopina o de tipo nopalina o de ambos.

Con “vector de transformación en plantas derivado de Ri” se quiere decir un vector de transformación binario en el que los extremos de ADN-T proceden de un plásmido Ti y dicho vector de transformación binario que se usa junto con un plásmido Ri “ayudante” que porta el grupo de genes *vir* necesarios.

Con “agrolísticos”, “transformación agrolística” o “transferencia agrolística” se quiere decir en el presente documento un procedimiento de transformación que combina las características de la transformación mediada por *Agrobacterium* y de la administración biolística de ADN. Como tal, un ADN-T que contiene un plásmido diana se co-administra con ADN/ARN permitiendo en la planta una producción de VirD1 y VirD2 con o sin VirE2 (Hansen y Chilton 1996; Hansen y col. 1997; WO 97/12046).

Se reconocen dos posibles mecanismos que dan lugar a la transferencia e integración de secuencias del esqueleto del vector en un genoma eucariota. En un primer mecanismo, la transferencia de ADN-T comienza legítimamente en el extremo derecho del ADN-T pero no se detiene en el extremo izquierdo del ADN-T y es seguida por la copia continuada de secuencias del esqueleto del vector. Este mecanismo es conocido como trans-lectura del extremo izquierdo. En un segundo mecanismo, la transferencia de ADN se inicia ilegítimamente en el extremo izquierdo del ADN-T con la copia continuada hasta que se alcanza de nuevo el extremo izquierdo del ADN-T. Independientemente de qué mecanismo prevalezca, el análisis de plantas transgénicas indica que muchas de ellas tienen las secuencias completas

del esqueleto del vector integradas en su genoma. La presente invención describe adicionalmente una aproximación para curar células transformadas que contienen secuencias del esqueleto del vector por recombinación que implica a una recombinasa y sitios de recombinación.

5 Con “curar” se quiere decir en el presente documento la eliminación de secuencias del esqueleto del vector de transformación sin suprimir el evento de integración del ADN-T. Otra herramienta para la eliminación de secuencias no ADN-T de individuos transgénicos ha sido descrita por Hanson y col. (1999). No obstante, esta herramienta es diferente y no está relacionada con el procedimiento de curación descrito en la presente invención. Como se ha descrito anteriormente, la eliminación de secuencias no ADN-T según los procedimientos de Hanson y col. (1999) o WO 10 99/01563 tiene el inconveniente de reducir la eficacia de la transformación (Hanson y col. 1999).

Así, en otra forma de realización de la invención, la curación de células transformadas se obtiene mediante un evento de recombinación como procedimiento para la eliminación de las secuencias del esqueleto del vector integradas ilegítimamente. Como resultado, se obtienen células transformadas que retienen el ADN-T en su genoma. Al 15 contrario que los procedimientos de Hanson y col. (1999) o WO 99/01563, el procedimiento de curación de la presente invención rescata ADN-T contenido en células transgénicas que previamente portaban secuencias del esqueleto del vector integradas ilegítimamente. Así las eficacias de transformación apenas se ven afectadas.

Con “evento de recombinación” se quiere decir un evento de recombinación específico de sitio o un evento de 20 recombinación llevado a cabo por “salto” de un transposón.

Con “recombinasa” se quiere decir una recombinasa específica de sitio o una transposasa.

Con “sitio de recombinación” se quiere decir sitios de recombinación específicos de sitio o secuencias terminales 25 de transposones.

Con “evento de recombinación específico de sitio” se quiere decir un evento catalizado por un sistema que consta generalmente de tres elementos: un par de secuencias de ADN (las secuencias o sitios de recombinación específicos de sitio) y una enzima específica (la recombinasa específica de sitio). La recombinasa específica de sitio cataliza una 30 reacción de recombinación solamente entre dos secuencias de recombinación específicas de sitio que dependen de la orientación de las secuencias de recombinación específicas de sitio. Las secuencias intermedias entre dos sitios de recombinación específicos de sitio se invertirán en presencia de la recombinasa específica de sitio cuando las secuencias de recombinación específicas de sitio estén orientadas en direcciones opuestas la una en relación a la otra (es decir, repeticiones invertidas). Si las secuencias de recombinación específicas de sitio están orientadas en la 35 misma dirección la una en relación a la otra (es decir, repeticiones directas), entonces cualquier secuencia intermedia será suprimida tras la interacción con la recombinasa específica de sitio. Así, si las secuencias de recombinación específicas de sitio están presentes como repeticiones directas en ambos extremos de las secuencias del esqueleto del vector integradas en un genoma eucariota, tal integración de dichas secuencias se puede eliminar posteriormente mediante la interacción de las secuencias de recombinación específicas de sitio con la recombinasa específica de sitio 40 correspondiente.

Se pueden usar una serie de sistemas recombinasa específicos de sitio diferentes incluyendo, pero no limitado a, el sistema Cre/lox del bacteriófago P1, el sistema FLP/FRT de levadura, la recombinasa Gin del fago Mu, la recombinasa Pin de *E. coli*, la PinB, PinD y PinF de *Shigella*, y el sistema R/RS de *Zygosaccharomyces rouxii*. Las recombinasas 45 generalmente son integrasas, resolvasas o flipasas. También se pueden usar recombinasas específicas dobles junto con repeticiones directas o indirectas de dos sitios de recombinación específicos de sitio diferentes correspondientes a la recombinasa específica doble (WO 99/25840). Los sistemas recombinasa específicos de sitio preferidos son el Cre/lox del bacteriófago P1 y los sistemas FLP/FRT de levadura y R/RS de *Z. rouxii*. En estos sistemas una recombinasa (Cre, FLP o R, respectivamente) interactúa específicamente con su respectiva secuencia de recombinación específica de sitio 50 (lox, FRT o RS, respectivamente) para invertir o eliminar las secuencias intermedias. Las secuencias de recombinación específicas de sitio para cada uno de estos dos sistemas son relativamente cortas (34 pb para lox y 47 pb para FRT). Algunos de estos sistemas ya se han usado con una eficacia elevada en plantas tales como el tabaco (Dale y col. 1990, Onouchi y col. 1991, Sugita y col. 2000) y *Arabidopsis* (Osborne y col. 1995, Onouchi y col. 1995). Los sistemas de recombinación específica de sitio tienen muchas aplicaciones en biología molecular de plantas incluyendo 55 procedimientos para el control de la recombinación homologada (por ejemplo, US5527695), para la inserción dirigida, el apilamiento de genes, etc. (WO 99/25821) y para la resolución de patrones de integración de ADN-T complejos o para la escisión de un marcador seleccionable (WO 99/23202). En estas aplicaciones los sitios de recombinación específicos de sitio normalmente son parte del ADN integrado en el genoma eucariota como ADN-T o mediante recombinación homologada, y así no están presentes en la secuencia del esqueleto del vector.

60 Aunque las secuencias de recombinación específicas de sitio deben estar unidas a los extremos del ADN a escindir o a invertir, el gen que codifica la recombinasa específica de sitio puede estar localizado en otra parte. Por ejemplo, el gen que codifica una recombinasa podría estar ya presente en el ADN de la eucariota o se podría suministrar mediante un fragmento de ADN introducido después directamente en las células, por cruce o por polinización cruzada. 65 Alternativamente, se podría introducir una proteína recombinasa sustancialmente purificada directamente en la célula eucariota, por ejemplo, mediante micro-inyección o bombardeo de partículas. Normalmente, la región que codifica la recombinasa específica de sitio estará unida de manera operable a secuencias reguladoras que permiten la expresión de la recombinasa específica de sitio en la célula eucariota.

Como parte de la presente invención, los sitios de recombinación específicos de sitio se introducen en el vector de ADN-T de manera que se encuentran fuera del ADN-T pero permiten la escisión post-transformacional de las secuencias del esqueleto del vector transferidas o de sus partes que posiblemente se originan en dicho vector de ADN-T mediante la acción de una recombinasa específica de sitio.

Con “evento de recombinación llevado a cabo por salto del transposón” o “recombinación mediada por transposasa” se quiere decir un evento de recombinación catalizado por un sistema constituido por tres elementos: un par de secuencias de ADN (las secuencias terminales del transposón) y una enzima específica (la transposasa). La transposasa cataliza una reacción de recombinación solamente entre dos secuencias terminales del transposón que están dispuestas como repeticiones invertidas.

Se pueden usar una serie de sistemas transposón/transposasa diferentes incluyendo, pero no limitado a, el sistema Ds/Ac, el sistema Spm y el sistema Mu. Estos sistemas tienen su origen en el maíz pero se ha demostrado que al menos el sistema Ds/Ac y el sistema Spm también funcionan en otras plantas (Fedoroff y col. 1993, Schlappi y col. 1993, van Sluys y col. 1987). Se prefieren los transposones de tipo Ds y Spm que están delimitados por secuencias terminales de 11 pb y 13 pb, respectivamente.

Aunque las secuencias terminales del transposón deben estar unidas a los extremos del ADN a escindir, el gen que codifica la transposasa puede estar localizado en otra parte. Por ejemplo, el gen que codifica una recombinasa podría estar ya presente en el ADN de la eucariota o se podría suministrar mediante un fragmento de ADN introducido después directamente en las células, por cruce o por polinización cruzada. Alternativamente, se podría introducir una proteína transposasa sustancialmente purificada directamente en las células, por ejemplo, mediante micro-inyección o bombardeo de partículas.

Como parte de la presente invención, las secuencias terminales del transposón se introducen en el vector de ADN-T de manera que se encuentran fuera del ADN-T y transforman al esqueleto del vector o a una de sus partes en una entidad de tipo transposón que se puede mover mediante la acción de una transposasa.

Como transposones, y así en la presente invención el esqueleto del vector o una de sus partes, a menudo se reintegran en otro locus del genoma del hospedador, podría ser necesaria la segregación de la progenie de los hospedadores en los que la transposasa se deja actuar para separar los hospedadores transformados que solamente contienen el ADN-T y los hospedadores transformados que solamente contienen el esqueleto del vector o una de sus partes.

La presente solicitud además describe vectores de transformación de ADN-T optimizados con secuencias de ADN adicionales fuera de las repeticiones centrales del extremo de ADN-T como se ejemplifica en la Figura 3 y que permiten la eliminación post-transformacional de secuencias integradas del esqueleto del vector. Dichas secuencias de ADN adicionales modifican, incluyendo multiplican, las regiones terminales de ADN-T o sus partes y comprenden:

- f) sitios de recombinación organizados como repeticiones aguas abajo de la secuencia central del extremo izquierdo y aguas arriba de la secuencia central del extremo derecho; o
- g) dichas secuencias de ADN de (f) modificadas adicionalmente mediante la adición de una segunda copia de una región en el extremo izquierdo situada aguas arriba de y preferentemente adyacente a la única región externa del extremo derecho y dicho sitio de recombinación aguas arriba de la secuencia central de dicha segunda región en el extremo izquierdo; o
- h) dichos sitios de recombinación de (f) con un gen que codifica una recombinasa, aguas abajo de dicho sitio de recombinación aguas abajo de la secuencia central del extremo izquierdo, y preferentemente, cuando está presente, adyacente a y aguas abajo de la región externa del extremo izquierdo; o
- i) una secuencia de ADN localizada aguas abajo de la región en el extremo izquierdo, dicha secuencia de ADN que comprende un gen que codifica una recombinasa flanqueado por repeticiones de sitios de recombinación como se ha definido en (h) y que comprende adicionalmente una segunda copia de una región en el extremo izquierdo aguas abajo de dicha recombinasa; o
- j) la secuencia de ADN de (i) con sitios de recombinación adicionales organizados como repeticiones aguas abajo de la segunda secuencia central del extremo izquierdo y aguas arriba de la única secuencia central del extremo derecho.

En cualquiera de dichas modificaciones (f), (g), (h), (i) o (j), dichos sitios de recombinación están localizados adyacentes a y aguas abajo y/o aguas arriba de las secuencias centrales del extremo izquierdo y/o derecho o están separadas de las secuencias centrales del extremo izquierdo y/o derecho por una secuencia de al menos 10-20 pb de longitud que opcionalmente porta codones de terminación en los tres marcos de lectura y en ambas direcciones.

En cualquiera de dichas modificaciones (f), (g), (h), (i) o (j), dichos sitios de recombinación son sitios de recombinación específicos de sitio dispuestos como repeticiones directas o secuencias terminales de transposones dispuestas como repeticiones invertidas y dicho gen que codifica una recombinasa es un gen que codifica una recombinasa específico de sitio o un gen de transposasa, respectivamente.

Es obvio que dichas modificaciones (f) o (g) en dichos vectores de transformación de ADN-T se deben considerar modificaciones simples, incluyendo la multiplicación, de las regiones terminales de ADN-T o de sus partes.

Se ha descrito la introducción de secuencias de ADN adicionales aguas abajo de la repetición central del extremo izquierdo del ADN-T para prevenir la trans-lectura en el extremo izquierdo (WO 99/01563) o para eliminar secuencias de no ADN-T (Hanson y col. 1999) y estas secuencias previenen el desarrollo de transformantes que tienen secuencias integradas del esqueleto del vector. Además de dichas modificaciones (f), dicha modificación (g) del vector de ADN-T tiene la ventaja de que la trans-lectura en la primera secuencia central del EI se puede detener en la segunda secuencia central del EI, previniendo así la duplicación del fragmento de ADN-T. Las secuencias del esqueleto del vector insertadas se eliminan posteriormente mediante la acción de una recombinasa apropiada. Además de dichas modificaciones (f), dicha modificación (h) introduce en un vector de ADN-T una secuencia de un gen que codifica una recombinasa aguas abajo de la secuencia central del extremo izquierdo para permitir la resolución de la secuencia del esqueleto del vector integrada que posiblemente se origina en dicho vector de ADN-T en los sitios de recombinación introducidos en dichos extremos del ADN-T modificados. Así, y al contrario que en el documento WO 99/01563, los transformantes que tienen secuencias del esqueleto del vector integradas se pueden rescatar después de la curación, es decir, después de la eliminación de las secuencias del esqueleto del vector flanqueadas por los sitios de recombinación introducidos en dichos extremos del ADN-T modificados. Simultáneamente, el gen que codifica una recombinasa se escinde.

Dichas modificaciones (i) además cambian el diseño de la región en el extremo izquierdo, es decir, se añade una copia adicional de una región en el extremo izquierdo. No obstante, ambas copias no están dispuestas en tándem como se ha descrito anteriormente en la primera aproximación para prevenir la integración del esqueleto del vector, sino que están separadas por el gen que codifica una recombinasa y las secuencias centrales del extremo se encuentran en su contexto en el extremo más amplio. Dichas modificaciones (i) así se deben considerar como una modificación de la construcción del vector de ADN-T descrita anteriormente en la que la única región en el EI contiene una disposición en tándem de la secuencia central del EI. La presencia de una segunda copia de una región en el extremo izquierdo del ADN-T tiene la ventaja de que se puede detener la transferencia de ADN ilegítima que comienza en o la trans-lectura en la primera secuencia central del extremo izquierdo en la segunda secuencia central del extremo izquierdo. Así, sólo se transfiere parte de la secuencia del esqueleto del vector al núcleo eucariota. Puesto que esta región del esqueleto del vector está flanqueada por sitios de recombinación, se puede eliminar fácilmente por medio de la acción de una recombinasa. No obstante, aún se puede producir la transferencia de ADN ilegítima que comienza en o la trans-lectura en la segunda copia de la secuencia central del extremo izquierdo.

Por tanto, dicha modificación (j) además añade un par de sitios de recombinación aguas abajo de la segunda copia de la secuencia central del extremo izquierdo y aguas arriba de la única región en el extremo derecho. Las secuencias del esqueleto del vector integradas de nuevo se pueden eliminar fácilmente mediante la acción de una recombinasa en los sitios de recombinación. En el caso de la transferencia de ADN ilegítima que comienza en la segunda copia de la secuencia central del extremo izquierdo, la recombinasa no se suministrará mediante el esqueleto del vector de transformación de ADN-T. Así se debe suministrar de otra parte, por ejemplo, por cruce sexual con una planta que ya contiene el gen que codifica una recombinasa en su genoma.

Cualquiera de dichas modificaciones (f), (g), (h), (i) o (j) también es aplicable para curar células transgénicas de secuencias del esqueleto del vector integradas en el genoma de una eucariota independientemente del ADN-T, es decir, no unidas físicamente al ADN-T.

En la presente solicitud además se describe un procedimiento que usa dichos vectores de transformación del ADN-T que contienen dichas secuencias de ADN de (f) o (g) en combinación con el suministro de una recombinasa, o que usa dichos vectores de transformación de ADN-T que contienen dichas secuencias de ADN (h), (i) o (j) para la curación de células transformadas que contienen secuencias del esqueleto de dichas secuencias del esqueleto del vector opcionalmente en combinación con el suministro de una recombinasa.

Las recombinasas específicas de sitio o transposasas introducidas en el genoma del hospedador para permitir la recombinación se pueden eliminar posteriormente mediante la segregación de la progenie del hospedador transformado en el que se ha permitido que se produzca la recombinación. La segregación de dicha progenie también es una manera de separar hospedadores transformados que contienen solamente el ADN-T y hospedadores transformados que contienen solamente el esqueleto del vector o sus partes.

La presente invención además expresa la incorporación de cualquiera de dichas secuencias de ADN adicionales en cualquier vector de ADN-T que comprende vectores de transformación binarios, vectores de transformación súper-binarios, vectores de transformación co-integrados, vectores de transformación derivados de Ri así como en vectores que portan ADN-T usados en transformación agrolística o terapia génica.

En una forma de realización de la presente invención, el gen que codifica una recombinasa se suministra a las plantas transgénicas que contienen una secuencia del esqueleto del vector flanqueada por sitios de recombinación mediante cruce sexual con una planta que contiene el gen que codifica una recombinasa en su genoma. Dicha recombinasa puede estar unida de manera operable a un promotor constitutivo o inducible. El gen que codifica una recombinasa alternativamente puede estar bajo el control de una única subunidad de promotores específicos de la ARN polimerasa de bacteriófago, tales como un promotor específico de T7 o T3, siempre que las células del hospedador también

comprendan la ARN polimerasa correspondiente en una forma activa. Aún otro procedimiento alternativo para la expresión de la recombinasa consiste en la unión de manera operable del marco de lectura abierto de la recombinasa con una secuencia de activación aguas arriba disparada mediante un factor de transcripción transactivante tal como GAL4 o sus derivados (US5801027, WO 97/30164, WO 98/59062) o el represor Lac (EP0823480), siempre que la célula hospedadora se suministre de una forma apropiada con el factor de transcripción.

En otra forma de realización de la invención, el gen que codifica una recombinasa se suministra en el esqueleto del vector de transformación y el promotor de dicho gen que codifica una recombinasa preferentemente es un promotor inducible. Tales promotores son conocidos por aquellos familiarizados con la técnica e incluyen, por ejemplo, un promotor sensible al choque térmico, un promotor inducible por glucocorticoides o un promotor inducible por otro compuesto químico (por ejemplo, como se describe en los documentos EP0332104 y WO 90/08826).

En una forma de realización preferida, se impide la expresión del gen que codifica una recombinasa en hospedadores bacterianos incluyendo una secuencia(s) intrón en la región codificante de dicho gen que codifica una recombinasa.

Es conocido en la técnica que niveles incrementados de VirD1 y VirD2 en *Agrobacterium* dan lugar a mayores niveles de transformación de la planta (Wang y col. 1990). VirD1 y VirD2 también se han integrado en un procedimiento para la integración mejorada de ADN exógeno administrado a células eucariotas por medio de la transformación de dicha célula eucariota con genes *virD1* y/o *virD2* quiméricos (WO 97/12046). No obstante, hasta ahora no se ha demostrado que la producción incrementada de VirD1 y/o VirD2 pueda prevenir la integración de secuencias del esqueleto del vector.

Así, en otra forma de realización de la invención, la integración de la secuencia del esqueleto del vector se previene potenciando la eficacia del corte en la secuencia central del extremo izquierdo de un vector de ADN-T incrementando la producción del complejo de corte de ADN-T que implica al menos a las endonucleasas VirD1 y VirD2.

En una forma de realización preferida de la invención, se integran copias adicionales del locus *virD* de tipo topina en el genoma de *Agrobacterium* y/o en un plásmido ayudante y/o en un vector de transformación binario y/o en un vector de transformación súper-binario y/o en un vector de transformación co-integrado y/o en un vector de transformación de plantas derivado de Ri. En otra forma de realización de la invención, como se ha mencionado se pueden suministrar copias adicionales del locus *virD* de tipo nopalina.

La presente solicitud así presenta tres procedimientos para prevenir o curar la integración de secuencias del esqueleto del vector de transformación y el uso de:

1) cualquiera de los vectores de ADN-T (a) a (e) modificados en la región en el extremo izquierdo del ADN-T de manera que se previene o se detiene la trans-lectura o se previene o se detiene la transferencia de ADN que comienza en este extremo, o

2) cualquiera de los vectores de ADN-T (f) a (j) que permite la resolución de secuencias del esqueleto del vector de transformación integradas en el genoma de células transgénicas por medio de recombinación; o

3) al menos una copia extra del locus *virD* en *Agrobacterium* para incrementar los niveles de VirD1 y VirD2 como constituyentes mínimos del complejo de corte en el extremo de ADN-T.

Es claro para aquellos familiarizados con la técnica que cualquiera de estos procedimientos o sus partes se puede usar solo, en combinaciones de dos, o como una combinación de los tres. Los diseños de un vector de ADN-T ejemplar que reflejan algunas de tales combinaciones se dan en la Figura 3.

Las combinaciones preferidas de los procedimientos de la invención (1) o (2) con el procedimiento de la invención (3) consisten en un vector de transformación que porta una región en el EI modificada que contiene una disposición en tándem de la secuencia central del EI o múltiples copias de la región en el EI y la producción potenciada de VirD1 y VirD2 por *Agrobacterium*. De hecho se puede esperar que el tándem de las secuencias centrales del EI o las múltiples copias de la región en el EI titulen las proteínas VirD1 y VirD2 disponibles en una cepa de *Agrobacterium* usada normalmente para la transformación. Como resultado, estas proteínas ya no estarán disponibles durante más tiempo para establecer el corte en la única secuencia central del EI.

La presente solicitud describe el uso de cualquiera de dichos procedimientos y/o modificaciones del vector de ADN-T para prevenir o curar la integración del esqueleto del vector de transformación por separado o en cualquier combinación. Las combinaciones preferidas son aquellas en las que niveles potenciados de proteínas VirD producidas por *Agrobacterium* se dejan actuar sobre vectores de ADN-T que albergan múltiples copias de la secuencia central del EI o múltiples copias de las regiones en el EI.

También es claro que se puede combinar cualquiera de dichos procedimientos y/o modificaciones del vector de ADN-T para prevenir o curar la integración del esqueleto del vector con cualquier otro procedimiento que previene o cura la integración del esqueleto del vector. Tales procedimientos y modificaciones del vector de ADN-T incluyen la adición de genes o secuencias de ADN aguas abajo del extremo izquierdo, por ejemplo, como se describe en el documento WO 99/01563. Tales genes/secuencias de ADN incluyen genes que codifican compuestos citotóxicos,

genes de mantenimiento antisentido, secuencias que impiden el desenrollamiento del ADN detrás de la región del extremo izquierdo, por ejemplo secuencias con un elevado contenido en GC o secuencias con la caja *vir* que interactúan con las proteínas de unión al ADN.

5 Así otra forma de realización de la invención comprende el uso de cualquiera de los procedimientos y/o modificaciones del vector de ADN-T de la presente invención para prevenir o curar la integración de secuencias del esqueleto del vector de transformación en combinación con cualquier otro procedimiento y/o la modificación del vector de transformación de ADN-T para prevenir o curar la integración del esqueleto del vector de transformación.

10 Además será claro para alguien experto en la materia que cualquiera de los procedimientos y/o modificaciones del vector de ADN-T de la presente invención para curar la integración de secuencias del esqueleto del vector de transformación se puede combinar con cualquier otro procedimiento y/o modificación del vector de ADN-T usando o incorporando cualquier sistema de recombinación, por ejemplo, para escindir marcadores seleccionables activos en plantas. De hecho dicho procedimiento de curación se puede combinar con, por ejemplo, la escisión de un marcador
15 seleccionable flanqueando el marcador seleccionable ejemplar con el mismo o con diferentes sitios de recombinación que aquellos que flanquean la secuencia del esqueleto del vector. En dicho procedimiento de curación, la secuencia del esqueleto del vector también puede estar flanqueada por dos sitios de recombinación específicos de sitio diferentes y la curación se puede llevar a cabo mediante una recombinasa específica doble con especificidades que se corresponden con los sitios de recombinación específicos de sitio usados.

20 Una forma de realización adicional de la presente invención comprende así el uso de los procedimientos y/o modificaciones del vector de ADN-T de la presente invención para curar la integración de secuencias del esqueleto del vector de transformación en combinación con procedimientos y/o modificaciones del vector de ADN-T usando o incorporando cualquier sistema de recombinación para fines distintos de la curación de la integración de la secuencia
25 del esqueleto del vector.

Otra forma de realización de la presente invención comprende la curación de secuencias del esqueleto del vector de transformación integradas que implica la modificación de los extremos del ADN-T añadiendo cualquier par de sitios de recombinación específicos de sitio mutuamente diferentes y el uso junto con al menos una recombinasa específica
30 doble con especificidades correspondientes al menos a los sitios de recombinación específicos de sitio usados.

En otra forma de realización de la invención, cualquiera de dichas construcciones del vector de ADN-T según la invención se moviliza a una cepa de *Agrobacterium*. Las cepas resultantes también constituyen la invención.

35 En todavía otra forma de realización de la invención, cualquiera de dichos vectores de ADN-T según la invención se usa en una transformación mediada por *Agrobacterium* o agrolística, tales procedimientos de transformación que son conocidos por alguien experto en la materia. Dichos vectores de ADN-T según la invención se pueden usar para transformar varios tejidos de plantas que comprenden raíces, protoplastos, hojas, etc. de diferentes especies de plantas monocotiledóneas y dicotiledóneas tales como, pero no limitado a, *Arabidopsis*, tabaco, petunia, tomate, patata, judías,
40 arroz y trigo. Más en general, puede ser cualquier planta monocotiledónea o dicotiledónea, que pertenece preferentemente a una especie de plantas de interés en agricultura, cultivo de madera, horticultura o a una especie de plantas aplicada a la producción de productos farmacéuticos, bioquímicos, anticuerpos o perfumes. Tales plantas incluyen plantas cultivadas, plantas radicales, plantas que producen aceites, plantas que producen madera, plantas agrícolas, plantas que producen frutos, pasto o legumbres de forraje, plantas de compañía o plantas de horticultura. Tales plan-
45 tas además incluyen albaricoque, alcachofa, espárrago, manzana, plátano, cebada, brócoli, coles de Bruselas, repollo, canola, zanahoria, mandioca, coliflor, apio, cereza, escarola, berzas, algodón, abeto Douglas, abeto (especies *Abies* y *Picea*), lino, ajo, uvas, col rizada, lenteja, maíz, roble, avena, aceite de colza, kimbombó, cebolla, pera, pimienta, álamo, centeno, sorgo, soja, calabaza, remolacha azucarera, caña de azúcar, girasol.

50 Dichos vectores de ADN-T según la invención también se pueden usar en combinación con el procedimiento de transformación de plantas por inmersión de las flores (Clough y Bent 1998) o con el procedimiento de transformación de brotes apicales en plantas (WO 99/14348). Dichos vectores de ADN-T según la invención también se pueden usar para transformar otras células eucariotas incluyendo levaduras, mohos y hongos filamentosos y sus conidias, hifas o protoplastos derivados de ellos. La transformación mediada por *Agrobacterium* de dichas células eucariotas es conocida por aquellos expertos en la materia (por ejemplo, WO 98/45455). Una característica clave de las células eucariotas transformadas mediante transferencia mediada por *Agrobacterium* o transferencia agrolística de cualquiera de los vec-
55 tores de transformación de ADN-T de la invención es una frecuencia significativamente reducida de secuencias del esqueleto del vector integradas sin afectar mucho a la frecuencia de integración de secuencias de ADN-T.

60 Una forma de realización preferida de la invención comprende células de plantas transgénicas obtenidas mediante cualquier procedimiento de transferencia mediada por *Agrobacterium* o transferencia agrolística de cualquiera de los vectores de transformación de ADN-T de la invención. También están comprendidas las plantas transgénicas obtenidas o regeneradas mediante cualquier procedimiento después de cualquier procedimiento de transferencia mediada por *Agrobacterium* o transferencia agrolística de cualquiera de los vectores de transformación de ADN-T de la invención.
65 Tales plantas se caracterizan porque contienen secuencias de ADN-T, pero no secuencias del esqueleto del vector, en su genoma. Además está comprendida la descendencia de dichas plantas transgénicas así como las células, protoplastos, callos, tejidos, órganos, semillas, frutas, polen, huevos, cigotos, embriones zigóticos o somáticos derivados de ellas o derivados de dichas células de la planta transgénica.

El potencial para el uso de vectores que portan ADN-T según la presente invención en terapia génica se puede explicar como sigue. Se ha descrito que complejos reconstituidos *in vitro* que consta de VirD2-ADNss-VirE2 son capaces de transferir el ADNss intacto a núcleos de mamífero (Ziemienowicz y col. 1999).

5 La VirD2 protege al ADNss frente a la degradación exonucleolítica debido a que está covalentemente unida al extremo 5' del ADNss (Durrenberger y col. 1989, Young y Nester 1988) y, por medio de dos secuencias de localización nucleares (NLSs), dirige el ADNss al núcleo de la célula de la planta (Narasimhulu y col. 1996, Rossi y col. 1993, Shurvinton y col. 1992). La VirD2 además contiene un dominio "omega" importante para la integración eficaz del ADN-T en el genoma del hospedador (Narasimhulu y col. 1996, Mysore y col. 1998, Tinland y col. 1995). Las NLSs de VirD2 se ha demostrado que son activas en células animales incluyendo células HeLa humanas, embriones de *Drosophila* y oocitos de *Xenopus* (Guralnick y col. 1996, Ziemienowicz y col. 1999). La VirD1 permanece localizada en el citoplasma de las células de mamífero pero se puede importar al núcleo mediante un mecanismo de remolque que implica a la VirD2 (Ziemienowicz y col. 1999).

15 La VirE2 protege al ADNss frente a la degradación (endo)nucleolítica y preserva la integridad del ADN-T (Gelvin 1998b, Rossi y col. 1996). La proteína VirE2 además está activamente involucrada en la dirección del ADN-T al núcleo de la planta (Gelvin 1998b). Las NLSs de VirE2 pueden ser específicas de la planta pero el reposicionamiento de un único aminoácido dentro de las NLSs de VirE2 dirige estas proteínas modificadas al núcleo de células animales (Guralnick y col. 1996). Las proteínas VirE2 interactúan entre ellas pero la interacción de VirE1 con VirE2 es mucho más fuerte y VirE1 inhibe la auto-interacción de VirE2. En células de *Agrobacterium*, VirE1 previene la agregación de VirE2. La VirE1 así parece funcionar como una chaperona molecular de VirE2 (Deng y col. 1999).

25 Las características de estas proteínas Vir las hacen candidatos ideales para su aplicación en experimentos de terapia génica como indica Ziemienowicz y col. (1999). Uno de los problemas principales en terapia génica de hecho consiste en la dificultad de la administración del ADN a través de las barreras intracelulares que incluyen la degradación nucleolítica y la captación nuclear.

Los vectores virales son uno de los vehículos principales usados por los científicos en terapia génica para conseguir una secuencia de ADN expresada en el hospedador apropiado. Los vectores retrovirales (incluyendo el VIH y MMLV) se emplean en el 63% de los protocolos de terapia génica aprobados por el Recombinant DNA Advisory Committee (una división del NIH) mientras que los vectores adenovirales se usan en el 16% de esos protocolos. Otros vectores virales incluyen aquellos basados en AAV, HSV y *Vaccinia*. Los vectores virales forman un área de nuevo desarrollo continuo en terapia génica.

35 Se puede prever que los genes para las proteínas de *Agrobacterium* VirD1, VirD2 y VirE2 (y eventualmente VirE1) se incorporen en un vector viral de tal forma (es decir, incluyendo el uso del codón adaptado y las secuencias reguladoras apropiadas) que se puedan expresar en células animales, preferentemente de manera transitoria, preferentemente sin la integración de dichos genes *vir* en el genoma del hospedador. Si un ADN-T que contiene el gen de interés para fines de terapia génica, está presente en el mismo vector viral o se co-administra en un vector viral aparte u otro tipo de vector viral, dicho ADN-T se podría transferir eficazmente a continuación al genoma en los núcleos de células animales. Un requerimiento estricto para la aprobación de tal estrategia de terapia génica sería, no obstante, que sólo el ADN-T, y no cualquier otro ADN externo, se transfiera al núcleo. Cualquier modificación presentada en la presente invención para prevenir o curar la transferencia de secuencias de ADN que no pertenecen al ADN-T sería así de gran valor para incrementar la aceptabilidad de estrategias de terapia génica que incluyen el uso de proteínas de *Agrobacterium* para proporcionar y transferir el ADN de interés. Por tanto, los vectores que portan el ADN-T modificados según cualquiera de las modificaciones o cualquier combinación de modificaciones de la presente invención y que se usan con fines de terapia génica están incluidos en la presente invención.

Además es claro para los facultativos expertos que al menos los vectores de ADN-T según la presente invención y que comprenden EIs modificados que dan como resultado un corte muy eficaz en este EI se pueden utilizar como fuente de, por ejemplo, complejos de VirD2-ADN-Tss-VirE2 que se pueden usar en otras técnicas de terapia génica incluyendo micro-inyección, electroporación, administración mediada por un vehículo e inyección balística de ADN. Dichas complejos se podrían enriquecer en, por ejemplo, cultivo de *Agrobacterium* inducido por acetosiringona o un cultivo de *E. coli* que expresa al menos las proteínas VirD1, VirD2 y VirE2 por medio de, por ejemplo, cromatografía de afinidad usando un anticuerpo que reconoce la proteína VirD2.

La invención, que hasta ahora se ha descrito de manera general, se puede entender más claramente en referencia a los siguientes ejemplos, que se incluyen meramente con fines ilustrativos de ciertos aspectos y formas de realización de la presente invención y no se pretende que limiten la invención. Los contenidos de todas las referencias mencionadas en este texto se incorporan en el presente documento por referencia.

Descripción de las figuras

Figura 1. Representación esquemática del porcentaje de residuos de A y T (indicado por encima y por debajo de la línea horizontal que indica el vector) por bloques de 100 pb alrededor de (y que no incluyen) las repeticiones imperfectas centrales del extremo izquierdo y derecho de los vectores de ADN-T K y Hsb descritos y usados en los Ejemplos 1-3.

Figura 2. Representación esquemática de construcciones de ADN-T diseñadas para valorar la eficacia del corte en el extremo izquierdo modificado junto con un extremo derecho que carece de la región interna del extremo. Gus: secuencia que codifica la β -glucuronidasa; *nptII*: secuencia que codifica la neomicina fosfotransferasa II; 3' ocs: terminador en 3' del gen de la octopina sintasa; P nos: promotor del gen de la nopalina sintasa; 3' nos: terminador en 3' del gen de la nopalina sintasa; P35S: promotor constitutivo del ARN 35S del virus del mosaico de la coliflor.

Figura 3. Representación esquemática de construcciones de ADN-T diseñadas para permitir la curación de secuencias del esqueleto del vector integradas en el genoma de una eucariota. EI: extremo izquierdo, ED: extremo derecho.

Figura 4. Representación esquemática de construcciones de ADN-T ejemplares, EI: extremo izquierdo, ED: extremo derecho; EImod: extremo izquierdo modificado (véase la Figura 2 para posibles modificaciones).

Figura 5. Análisis de transferencia de ADN en gel.

(A) Representación esquemática de la digestión y la sonda usada en el análisis de transferencia de ADN en gel (no representado a escala). Cada muestra de ADN genómico de la planta y los plásmidos usados como controles positivos se digirieron con *KpnI* y *SacI*, ambos que cortan en el esqueleto del vector fuera del EI y el ED. Como sonda se usó la secuencia del vector del plásmido K, mostrado como una barra sombreada, después de la restricción con *KpnI* y *SacI*.

(B) Análisis de transferencia de ADN en gel. Los vectores de ADN-T de los cuales al menos 1000 pb del esqueleto del vector se encontraron mediante análisis de PCR de las plantas transgénicas se indican en la parte superior de cada ruta. Se observa un fragmento de 8476 pb (flecha 1) cuando se integra toda la secuencia del esqueleto del vector del plásmido K (K1 y K2 portan un gen estructural diferente en su ADN-T pero además son idénticos), mientras que se detectan fragmentos de 7066 pb (flecha 2) cuando se integra toda la secuencia del esqueleto del vector del plásmido Hsb. Los plásmidos de ADN-T K y Hsb se describen en el Ejemplo 1.

Las muestras de ADN genómico en los carriles 1, 2 y 5 proceden de plantas en experimentos distintos a aquellos descritos.

Carril C. El ADN genómico de *A. thaliana* sin transformar.

Carriles de plásmidos: Contienen ADN puro de los plásmidos indicados digeridos con *KpnI* y *SacI*.

Ejemplos

Ejemplo 1

Transformación de A. thaliana (L.) Heynh y *Nicotiana tabacum* (L.)

Se usaron los plásmidos pK2L610 (vector de ADN-T "K"; De Buck y col. 1998), pSingle gus (vector de ADN-T "Ksbl"; véase a continuación) y pHSB610 (vector de ADN-T "Hsb"; De Buck y col. 1999) para la transformación de la planta. El vector pSingle gus ("Ksbl") es similar al vector pHSB610 ("Hsb") excepto porque el fragmento Pnos-hpt-3'nos de pHSB610 ("Hsb") se intercambia por un casete de Pnos-*nptII*-3'nos en pSingle gus ("Ksbl"). Como marcador seleccionable para la transformación de la planta, los vectores K y Ksbl contienen el gen de la neomicina fosfotransferasa II (*nptII*) y el vector Hsb contiene el gen de la higromicina fosfotransferasa (*hpt*). La *A. thaliana* se co-transformó con los plásmidos K y Hsb según el procedimiento de transformación de la raíz de *Arabidopsis* mediada por *Agrobacterium* (De Buck y col. 1999, Valvekens y col. 1988). El tabaco (*N. tabacum*) se transformó con el plásmido Ksbl según el procedimiento de transformación del disco de la hoja mediada por *Agrobacterium* (Horsch y col. 1985). Las plantas transgénicas se regeneraron en un medio que contiene las hormonas de crecimiento de la planta apropiadas y el agente selectivo apropiado. En la serie 1, se obtuvieron 18 plantas de *A. thaliana* transgénicas se co-transformadas con los vectores de ADN-T K y Hsb. En la serie 2, se obtuvieron 36 plantas de *N. tabacum* transformadas con el vector de ADN-T Ksbl.

Ejemplo 2

Integración de secuencias del esqueleto del vector valorada por PCR

Para el análisis por PCR, el ADN se aisló a partir del material de la hoja de *N. tabacum* como se describe por Jones y col. (1985) y a partir de *A. thaliana* según De Neve y col. (1997). Para seleccionar las plantas transgénicas en la integración de secuencias del vector, se realizaron diferentes reacciones de PCR. El ADN (100 ng) se incubó con 500 ng de cada cebador en tampón de incubación de Taq polimerasa 1x (Roche Diagnostics, Bruselas, Bélgica). Se añadieron dos unidades y media de Taq polimerasa hasta un volumen final de 100 μ l. Las muestras se calentaron a 94°C durante 5 minutos antes de la PCR. La desnaturalización se produjo a 94°C durante 1 minuto. La hibridación se produjo durante 2 minutos a 57°C, y la reacción de extensión fue a 72°C durante 5 minutos, mientras que se realizaron 30 ciclos. Las combinaciones del cebador se escogieron de manera que la presencia de al menos 100 pb ("ED100" y "EI100" en las Tablas 1 y 2) o al menos 1000 pb ("ED1000" y "EI1000" en las Tablas 1 y 2) del esqueleto del vector dan como resultado un fragmento de PCR diagnóstico de tamaño conocido. Para asegurarse de que el producto amplificado por PCR no procede de la contaminación con células de *Agrobacterium* que aún estarían presentes en el

tejido de la planta, se llevó a cabo una reacción de PCR sobre cada ADN con dos cebadores específicos para el gen cromosómico de *A. tumefaciens* *picA* (Yusibov y col. 1994).

La serie 1 de plantas de *A. thaliana* contiene 18 plantas transgénicas co-transformadas con los vectores de ADN-T K y Hsb. El vector K contiene extremos en el contexto de la octopina natural con regiones terminales internas y externas de 150 pb y 255 (ED) - 300 (EI) pb, respectivamente, mientras que el vector Hsb contiene solamente las regiones terminales externas. La selección para la integración de secuencias del esqueleto del vector reveló que secuencias del vector procedentes de los ADN-T K y Hsb se encontraron el 33% (6/18) y en el 61% (11/18) de las plantas transgénicas, respectivamente. Sólo en un transformante las secuencias del vector no estaban unidas al ADN-T. No se detectaron diferencias importantes en las frecuencias del vector unido a la región ED del Hsb (6/18) y la región ED (4/18) del plásmido K (véase Tabla 1). No obstante, se integró mucho más vector en la región EI del ADN-T Hsb (11/18) que del ADN-T K. En todos los casos, en los que un ADN-T integrado contiene secuencias del vector en el ED, también se pudieron detectar secuencias del vector en el EI.

La serie 2 de plantas del tabaco contiene 36 plantas transgénicas transformadas con un vector Ksb. Este vector también contiene ED y EI sin regiones terminales internas y por tanto es comparable al vector Hsb. En el 53% (19/36) de los transformantes analizados (véase Tabla 2) se encontraron secuencias del vector; éstas estaban unidas solamente al EI (8/19), solamente al extremo derecho del ADN-T (1/19), o a ambos extremos (10/19).

En otra serie de transformantes del tabaco que contiene el ADN-T Ksb (datos no mostrados), el 81% de los transformantes (21/26) contenía secuencias del vector unidas al EI y el 61% (16/26) unidas al ED (datos no mostrados). No obstante, sorprendentemente, de estos 16 transformantes que contienen las secuencias del esqueleto del ED, 15 de ellos también contienen secuencias del vector del EI integradas (datos no mostrados).

Tomados juntos, en las tres series diferentes de transformantes en las que se usó un vector de ADN-T sin secuencias terminales internas, más del 50% de los transformantes contenían ADN del esqueleto del vector. Muy pocos transformantes contenían ADN del vector unido solamente al extremo derecho del ADN-T. En general, cuando estaba presente el ADN del vector unido al extremo derecho, también se pudo detectar ADN del vector unido al extremo izquierdo del ADN-T. Esto implica que especialmente la trans-lectura en la repetición en el extremo izquierdo, debido al corte ineficaz, es responsable de la integración de secuencias del esqueleto del vector. Es posible que la supresión de la región terminal interna, un trozo de ADN-T presente en el plásmido Ti original a partir del cual procede el vector, provoque el corte ineficaz de la repetición del EI, se da como resultado la trans-lectura más allá del EI y la transferencia de secuencias del vector localizadas aguas abajo. Consecuentemente, Horsch y Klee (1986) observaron que la velocidad global de transferencia a plantas de plásmidos que solamente contienen la repetición de 25 pb no es tan elevada como con los fragmentos más largos del extremo procedente de Ti, que sugiere que secuencias adicionales que rodean las repeticiones en el extremo desempeñan algún papel en la transferencia. No obstante, estos autores no han abordado la cuestión de la posible integración de secuencias del esqueleto del vector en plantas transgénicas usando dicho vector.

Ejemplo 3

Integración de secuencias del esqueleto del vector valorada por transferencia de ADN en gel

El análisis de transferencia de ADN en gel se llevó a cabo esencialmente según Maniatis y col. (1982) sobre 0,6 µg aproximadamente de ADN genómico. Los sitios de restricción usados (*KpnI*-*SacI*) se indican en la Figura 5A. Como sonda, se usó la secuencia del vector del plásmido K después de la restricción con *KpnI* y *SacI*. Se usaron el módulo de marcaje de cebado aleatorio "Gene images" no radioactivo y el módulo de detección "Gene images" CDP Star (Amersham, Aylesburg, RU) para la hibridación y la detección, respectivamente.

Como se resume en el Ejemplo 2, se observó la unión de ambas regiones en el ED y EI a al menos un fragmento del esqueleto del vector de 1000 pb de uno o ambos plásmidos en diferentes transformantes de la serie 1 (véase Tabla 1). Para determinar si el plásmido de ADN-T binario completo (ADN-T + secuencias del esqueleto del vector) se integra en el genoma de estos transformantes, se llevó a cabo un análisis de transferencia de ADN en gel con la secuencia del vector K como sonda. Cuando está presente toda la secuencia del vector, la restricción del ADN genómico con *KpnI* y *SacI* daría como resultado la detección de un fragmento de 8467 pb aproximadamente para K (flecha 1 en la Figura 5B) y de 7066 pb para Hsb (flecha 2 en la Figura 5B). El análisis de transferencia de ADN en gel dado en la Figura 3B muestra que en 10 de las 11 plantas analizadas, toda la secuencia del esqueleto del vector se integró en el genoma de la planta. Las muestras en las carriles 1, 2 y 5 proceden de plantas en experimentos distintos a aquellos descritos. Estos resultados sugieren que secuencias completas del vector se integran frecuentemente en el genoma de la planta.

Ejemplo 4

Construcción de vectores de ADN-T optimizados con extremos izquierdos modificados

Se construyeron como sigue vectores de ADN-T binarios optimizados ejemplares con modificaciones de los tipos indicados en la Figura 2.

En este ejemplo, todos los vectores de ADN-T binarios optimizados descritos en este ejemplo proceden del pTHW 136. El vector binario pTHW 136 contiene secuencias centrales del extremo derecho e izquierdo de tipo octopina

ES 2 284 555 T3

que forman una repetición imperfecta, cada una flanqueada por regiones externas terminales de tipo octopina que se originan en pTi15955 (224 pb en el caso de la región externa del ED y 268 pb en el caso de la región externa del EI). Sobre el ADN-T del pTHW 136 está localizado un casete de expresión del intrón GUS (promotor de 35S - marco de lectura abierta de GUS interrumpido por un intrón - terminador de 35S) y un gen marcador seleccionable *nptII* (promotor de nos - marco de lectura abierta *nptII* - terminador de ocs).

Para crear vectores de ADN-T binarios ejemplares de los tipos indicados en la Figura 2, se eliminó el casete de expresión del intrón GUS de pTHW 136 por digestión con *XbaI* y *HindIII*, seguido del relleno de los salientes con la polimerasa Klenow y de la religación de los extremos romos resultantes. El vector obtenido mediante este procedimiento se denominó posteriormente p0130.

En una etapa posterior, el gen del marcador seleccionable *nptII* se eliminó del p0130 por digestión con *BamHI* y la religación posterior de los salientes compatibles del vector. Este procedimiento dio un vector de ADN-T binario denominado pCDVIB. El vector pCDVIB así contiene una región en el EI de tipo octopina que se expande a una secuencia central flanqueada por una región externa terminal y así es una construcción del tipo B como se indica en la Figura 2.

Para crear una construcción del tipo E indicada en la Figura 2, se diseñaron dos oligonucleótidos largos, prmCDVIB1 F y prmCDVIB1 R, con las siguientes secuencias:

-prmCDVIB1 F:

5' (1)

CATGGAGCGGCGGCAGGATATATTCAATTGTAAATGGCTAGCGGCGGCAGG

ATATATTCAATTGTAAATGGCTG (84) 3' (ID SEC N° 5); sitio *NcoI* parcial en negrita y subrayado, secuencias centrales del extremo izquierdo en cursiva y subrayado, el residuo de G en el extremo 3' parte de un sitio *BamHI*; y

-prmCDVIB1 R:

5' (1)

GATCCAGCCATTTACAATTGAATATATCCTGCCGCCGCTAGCCATTTACAATTGAATA
TATCCTGCCGCCGCTC

(84) 3' (ID SEC N° 6); sitio *BamHI* parcial en negrita y subrayado, secuencias centrales del extremo izquierdo en cursiva y subrayado, el residuo de G en el extremo 3' parte de un sitio *NcoI*.

Los oligonucleótidos prmCDVIB1 F y prmCDVIB1 R son complementarios a lo largo de una distancia de 80 nucleótidos y dan, después de la hibridación, un fragmento de ADNds que tiene en su extremo 5' (en relación a prmCDVIB1F) un saliente 5' *NcoI* ("CATG") y en su extremo 3' (en relación a prmCDVIB1R) un saliente 5' *BamHI* ("GATC"). Dicho fragmento de ADNds además contiene un tándem de dos secuencias centrales del EI de tipo octopina correspondientes a los nucleótidos 12-33 y 46-67 de prmCOVIB1F. Ambas secuencias centrales del EI se extienden aguas arriba y aguas abajo mediante 6 pares de bases adicionales que se originan en el EI natural de pTi15955 y son idénticas a pTiAch5 (Gielen y col. 1984). Las repeticiones en el tándem de secuencias centrales del EI así están separadas por 12 pb. Dicho fragmento de ADNds con las secuencias centrales del EI dispuestas en tándem se insertaron aguas arriba de y adyacentes a la secuencia central del EI de p0130 digerido con *NcoI* y *BamHI*, dando pCDVIB1. El pCDVIB1 así contiene una región en el extremo izquierdo de tipo octopina que se extiende a una región externa del EI y tres secuencias centrales del EI dispuestas en tándem.

Para crear una construcción del tipo A indicada en la Figura 2, se diseñaron dos oligonucleótidos largos, prmCDVIB2F y prmCDVtB2R, con las siguientes secuencias:

ES 2 284 555 T3

-prmCDVIB2F:

5' (1)

CATGGCCGGGAAATCTACATGGATCAGCAATGAGTATGATGGTCAATATGGA

GAAAAAGAAAGAGTAATTACCAATTTTTTTTCAATTCAAAAATGTAGATGTC

CG (116) 3' (ID SEC N° 7); sitio *NcoI* parcial en negrita y subrayado, el residuo de G en el extremo 3' parte de un sitio *BamHI*; y

-pCDVIB2R:

5' (1)

GATCCGGACATCTACATTTTTGAATTGAAAAAAATTGGTAATTACTCTTTCTT

TTTCTCCATATTGACCATCATACTCATTGCTGATCCATGTAGATTTCCCGGC (116)

3' (ID SEC N° 8); sitio *BamHI* parcial en negrita y subrayado, el residuo de C en el extremo 3' parte de un sitio *NcoI*.

Los oligonucleótidos prmCDVIB2F y prmCDVIB2R son complementarios a lo largo de una distancia de 112 nucleótidos y dan, después de la hibridación, un fragmento de ADNds que tiene en su extremo 5' (en relación a prmCDVIB1F) un saliente 5' *NcoI* ("CATG") y en su extremo 3' (en relación a prmCDVIB1F) un saliente 5' *BamHI* ("GATC"). Dicho fragmento de ADNds además contiene una región interna de 112 pb de pTi15955. Este fragmento se insertó aguas arriba de y adyacente a la secuencia central del EI de p0130 digerido con *NcoI* y *BamHI*, dando pCDVIB2. El pCDVIB2 así contiene una región en el extremo izquierdo de tipo octopina que se extiende a una región externa del EI embebida en las regiones interna y externa del EI.

Se realizó otra construcción del vector de ADN-T con un EI modificado que combina características de las construcciones de tipo E y F indicadas en la Figura 2. Se obtuvo una región en el EI de tipo nopalina integral como un fragmento *BclI-EcoRI* de 331 pb del vector de ADN-T binario pPZP200. Este fragmento se extiende a la región externa del EI de tipo nopalina, a la secuencia central del EI y a la región interna del EI que se origina en el pTiC58 (Gielen y col. 1999). Dicha región en el EI de tipo nopalina se insertó aguas arriba de y adyacente a la región en el EI de tipo octopina de pCDVIB2 (que contiene secuencias centrales del EI de tipo octopina repetidas en tándem) digerido con *BamHI* (localizado en el 3' del tándem de la secuencia central del EI, en relación a prmCOVIB2F) y *EcoRI*, dando pCDVIB3. El pCDVIB3 así contiene una región en el EI que se extiende a una copia de una región en el EI de tipo octopina integral con tres secuencias centrales dispuestas en tándem y una copia de una región en el EI de tipo nopalina integral.

Todos los vectores de ADN-T ejemplares descritos con una región en el EI modificada, es decir, pCDVIB, pCDVIB1, pCDVIB2 y pCDVIB3 sirven como punto de partida para insertar un gen(es) de interés y/o un gen(es) marcador seleccionable entre el ED y el EI de dichos vectores de ADN-T.

El gen marcador seleccionable neomicina fosfotransferasa (*nptII*) bajo el control del promotor de la nopalina sintasa (nos) (Pnos-*nptII*-3'ocs) y el casete de expresión de la β -glucuronidasa (*gus*) bajo el control del promotor del virus del mosaico de la coliflor 35S (P35S-*gus*-3'nos) se insertaron entre el ED y el EI de dichos vectores de ADN-T.

El casete *nptII*-*gus* procede del plásmido pXD610 (De Loose y col., 1995) como un fragmento *EcoRI*-*AgeI*. Este fragmento se insertó en los sitios *SalI*/*EcoRI* de pCDVIB; pCDVIB1; pCDVIB2 y pCDVIB3. Para obtener extremos cohesivos compatibles, se usaron dos oligonucleótidos adaptadores entre los sitios abiertos *AgeI* y *SalI* con las siguientes secuencias:

* Napod3:

CCGGTGGCTCGAGG (ID SEC N° 9); sitio *AgeI* parcial en negrita, el residuo de G en el extremo 3' parte de un sitio *SalI* y

* Napod4:

TCGACCTCGAGCCA (ID SEC N° 10); sitio *SalI* parcial
subrayado, el residuo de A en el extremo 3' parte de un
sitio *AgeI*.

Los oligonucleótidos Napod3 y Napod4 son complementarios a lo largo de una distancia de 8 nucleótidos y dan, después de la hibridación, un fragmento de ADNds que tiene en su extremo 5' (en relación a Napod3) un saliente 5'*AgeI* ("CCGG") y en su extremo 3' (en relación a Napod3) un saliente *SalI* ("TCGA"). Los vectores obtenidos se denominaron pCDVIB+gusnpt, pCDVIB1+gusnpt, pCDVIB2+gusnpt, pCDVIB3+gusnpt.

Ejemplo 5

Construcción de vectores de ADN-T optimizados con sitios de recombinación integrados

Se construyeron como sigue vectores de ADN-T binarios optimizados ejemplares de los tipos indicados en la Figura 3. En este ejemplo, el vector de ADN-T binario optimizado procede del pCDVIB (véase Ejemplo 4).

Una estrategia para introducir sitios de recombinación comprende:

1. Diseñar pares de cebadores para la amplificación por PCR de dominios del vector aguas arriba del ED y aguas abajo del EI. Dichos dominios tienen que incluir un único sitio de restricción aguas arriba del ED (por ejemplo, el sitio *NarI* en el replicón pVS1 contenido dentro del esqueleto del vector pCDVIB) y aguas abajo del EI (por ejemplo, el sitio *BclI* en el marcador de resistencia Sm/Sp contenido dentro del esqueleto del vector pCDVIB). Los cebadores de la PCR localizados en o cercanos al ED o al EI están diseñados de manera que contienen un único sitio de restricción en 5' seguido del sitio de recombinación (por ejemplo, el sitio FRT) en su orientación correcta (como repeticiones directas fuera de los extremos del ADN-T en el caso de los sitios FRT) y seguido por parte de las regiones externas del extremo. Los productos obtenidos de la PCR se digieren posteriormente con las enzimas apropiadas.

2. Diseñar pares de oligonucleótidos complementarios (similar al Ejemplo 4) dando después de la hibridación moléculas de ADNds que se extienden a la secuencia central del ED o en el EI y con un saliente complementario al único sitio de restricción 5' de dichos cebadores en (1) localizados en o próximos a los extremos y con un segundo saliente complementario a un único sitio de restricción aguas abajo de la secuencia central del ED (por ejemplo, *ScaI* en pCDVIB) o complementario a un único sitio de restricción aguas arriba de la secuencia central del EI (por ejemplo, *NcoI* en pCDVIB).

3. (a) Digerir pCDVIB con *NarI* (en el replicón pVS) y *ScaI* (aguas abajo de la secuencia central del ED) seguido de una ligación trimolecular de 1) el vector resultante con 2) el producto digerido de la PCR que incluye el fragmento pCDVIB suprimido y la introducción de sitios de recombinación y 3) el ADNds basado en los oligonucleótidos hibridados y la reintroducción de la secuencia central del ED.

- (b) Digerir el vector obtenido finalmente en (3a) con *BclI* (en el marcador de resistencia Sm/Sp) y *NcoI* (aguas arriba de la secuencia central del EI) seguido por una ligación trimolecular de 1) el vector resultante con 2) el producto digerido de la PCR que incluye el fragmento pCDVIB suprimido y la introducción de sitios de recombinación y 3) el ADNds basado en los oligonucleótidos hibridados y la reintroducción de la secuencia central del EI.

Los vectores de ADN-T ejemplares descritos con sitios de recombinación introducidos sirven como punto de partida para insertar un gen(es) de interés y/o un gen(es) marcador seleccionable entre el ED y el EI de dicho vector de ADN-T.

Ejemplo 6

Predicción de las frecuencias de transferencia del esqueleto del vector usando cultivos de Agrobacterium inducidos

Se pueden inducir cepas de *Agrobacterium* que contienen cualquier vector de ADN-T para la producción de cadenas de ADN-Tss añadiendo acetosiringona a una concentración de 100 μ M al medio de cultivo (Durrenberger y col. 1989). La producción de cadenas de ADN-Tss se puede analizar usando uno de los siguientes procedimientos (véanse los incluidos en la Figura para una visión esquemática).

1. Se aisló el ADN total de las células de *Agrobacterium* que contienen un vector de ADN-T dado. Los duplicados de transferencias Southern no desnaturalizantes (Tinland y col. 1995) de 5 μ g aproximadamente de este ADN aislado de las diferentes cepas de *Agrobacterium* se hibridaron a dos sondas diferentes: una que reconoce una secuencia de ADN que es parte del ADN-T y una que reconoce una secuencia de ADN que es parte del esqueleto del vector que flanquea el extremo izquierdo. La señal de hibridación de la sonda que reconoce las secuencias del esqueleto del vector está ausente o muy reducida en el caso de un vector de ADN-T a partir del cual se procesa el extremo izquierdo con una eficacia elevada. No obstante, la señal de hibridación de la sonda que reconoce el ADN-T es constante.

2. Se aisló el ADN total de las células de *Agrobacterium* que contienen un vector de ADN-T dado. Se prepararon dos digestiones separadas que contienen cantidades iguales del ADN total, una con una enzima de restricción que corta el ADN-T en o muy próximo al extremo izquierdo, la otra con una enzima de restricción que corta dentro de la secuencia de ADN-T. No obstante, las cadenas de ADN-Tss permanecerán intactas. Se usaron pares de cebadores específicos en una PCR cuantitativa para amplificar una secuencia procedente del ADN-T y una secuencia procedente del vector aguas abajo del extremo izquierdo. De nuevo, se obtuvo mucho menos o ningún producto de amplificación del esqueleto del vector en el caso de un vector de ADN-T a partir del cual se procesa el extremo izquierdo con una eficacia elevada. No obstante, las cantidades del producto de la amplificación específica del ADN-T son constantes.

Para ambos tipos de análisis, se clonó un casete de expresión GFP (proteína verde fluorescente) adicional inmediatamente aguas abajo del extremo izquierdo de los diferentes vectores de ADN-T. Este casete proporciona una secuencia "esqueleto del vector" completamente conocida aguas abajo del extremo izquierdo. El casete de expresión GUS se usa como diana específica del ADN-T. Ambos casetes de expresión contienen un intrón en las regiones codificantes de GFP y GUS, respectivamente, para prevenir la expresión de ambos marcadores en *Agrobacterium*.

En la Figura 6 se representan esquemáticamente ambos montajes experimentales. Por encima de la línea negra continua que indica el vector de ADN-T, se muestra la aproximación de la hibridación y por debajo de la línea continua, se muestra la aproximación de la PCR cuantitativa.

Ejemplo 7

Análisis tempranos de secuencias de transferencia del esqueleto del vector a células de la planta

Se usaron las construcciones apuntadas en el Ejemplo 6 y que contienen un casete de expresión de GFP adicional aguas abajo del extremo izquierdo en un ensayo de expresión transitorio para la valoración de la transferencia del esqueleto del vector a células de la planta. Con este propósito, fragmentos radiculares de *Arabidopsis* se transformaron con *Agrobacterium* (De Buck y col. 1999, Valvekens y col. 1988) que contiene un vector de ADN-T descrito anteriormente. Después de tres días de cultivo conjunto, las raíces se valoraron primero para la expresión transitoria de la GFP por valoración de la GUS expresada transitoriamente. El nivel de expresión de GFP es bajo o cero en el caso de un vector de ADN-T a partir del cual se procesa el extremo izquierdo con una eficacia elevada. No obstante, los niveles de actividades de la GUS son constantes.

Alternativamente, el procedimiento de transcripción inversa seguido de la PCR se continúa con la detección y cuantificación de los niveles de transcritos de GFP y GUS acumulados transitoriamente (Narasimhulu y col. 1996). De nuevo, se obtiene mucho menos o ningún producto de la amplificación específica de la GFP en el caso de un vector de ADN-T a partir del cual se procesa el extremo izquierdo con una eficacia elevada. No obstante, las cantidades del producto de la amplificación específica de la GUS son constantes.

Ejemplo 8

Curación de plantas transformadas de secuencias del esqueleto del vector integradas

Se transformaron plantas (por ejemplo, *Arabidopsis*) (véase Ejemplo 1) con el vector de ADN-T ejemplar del Ejemplo 5 que contiene los sitios de recombinación FRT. Las plantas regeneradas se probaron para la integración del esqueleto del vector según cualquiera de los procedimientos descritos en los Ejemplos 2-3 ó 6-7 y se seleccionaron las plantas transgénicas que contienen el ADN-T y el esqueleto del vector en su genoma.

En la floración, un grupo de plantas transgénicas seleccionadas se sometió a cruce por polinización con polen procedente de otra planta transgénica que expresa constitutivamente la recombinasa específica del sitio FLP mientras que un segundo grupo de plantas transgénicas seleccionadas se sometió a cruce por polinización con polen procedente de una planta natural. Se recogieron las semillas y se sembraron en medio que contiene un agente selectivo. Las plantas supervivientes (debido al marcador seleccionable en el ADN-T) de ambos cruces se analizaron de nuevo para la presencia de ADN-T y para la presencia del esqueleto del vector. En una parte sustancial de la progenie de los cruzamientos que implican al progenitor que expresa la FLP y que contienen el ADN-T, se eliminaron las secuencias del esqueleto del vector. En la progenie de los cruzamientos que implican al progenitor del tipo natural y que contienen el ADN-T, aún están presentes las secuencias del esqueleto del vector.

Ejemplo 9

Clonación del locus virD e influencia de una copia extra del locus sobre la transferencia del esqueleto del vector

Se conocen las secuencias de los loci *virD* de tipo octopina y de tipo nopalina (Yanofsky y col. 1986, Jayaswal y col. 1987; y Wang y col. 1990, respectivamente). Se diseñaron cebadores específicos que permiten la amplificación por PCR de los loci *virD* completos. Estos loci se subclonaron posteriormente en un plásmido que porta el replicón P15a (Chang y Cohen 1987) y un marcador de resistencia a tetraciclina procedente de pAlter (Promega). Los plásmidos se movilaron posteriormente a *Agrobacterium* que contiene un vector de ADN-T. La transferencia de las secuencias del esqueleto del vector mediante dichas cepas de *Agrobacterium* se analizó según cualquiera de los procedimientos descritos en los Ejemplos 2, 3, 5 ó 6.

Ejemplo 10

Determinación de la frecuencia de iniciación de la transferencia en el extremo izquierdo

Las plantas se transformaron con los vectores de ADN-T descritos en la Figura 2. Se preparó el ADN total y se sintetizaron tres grupos de cebadores para llevar a cabo diferentes reacciones de PCR. El primer grupo de cebadores amplifica un fragmento de ADN-T interno, adyacente a la repetición en el extremo izquierdo (fragmento de PCR 1). El segundo grupo de cebadores amplifica un fragmento de ADN-T/vector que se extiende a la región en el extremo izquierdo del ADN-T y la región del vector adyacente a la repetición en el extremo izquierdo (fragmento de PCR 2). Finalmente, el tercer grupo de cebadores amplifica un fragmento del vector adyacente a la repetición del extremo izquierdo (fragmento de PCR 3). Dependiendo de si la transferencia del vector da como resultado la trans-lectura en el extremo izquierdo o la iniciación en el extremo izquierdo, se amplifican diferentes fragmentos por PCR. Cuando sólo está presente el fragmento de PCR 1, la transferencia del ADN-T se inicia en el extremo derecho y se detiene en el extremo izquierdo y no se transfiere ADN del vector. Cuando están presentes los fragmentos de PCR 1, 2 y 3, la transferencia del ADN-T se inicia en el extremo derecho pero no se detiene en el extremo izquierdo. Cuando están presentes los fragmentos de PCR 1 y 3, pero no el fragmento 2, la transferencia del vector se inicia en el extremo izquierdo, independientemente de la transferencia del ADN-T en el extremo derecho. Realizando estas diferentes reacciones de PCR para 50 transformantes aproximadamente se obtiene así la frecuencia de trans-lectura en el extremo izquierdo y la frecuencia de iniciación en la repetición en el extremo izquierdo. La frecuencia de trans-lectura y la iniciación de la transferencia del vector se determinan para diferentes vectores de ADN-T con la repetición en el extremo izquierdo en contextos diferentes.

(Tabla pasa a página siguiente)

Tabla 1. Presencia (+) o ausencia (-) de secuencias del vector unidas al EI y al ED del ADN-T en plantas de *A. thaliana* transgénicas co-transformadas con ambos vectores de ADN-T K y Hsb.

N° de co-trans ^a	EI	EI	ED	ED	EI	EI	ED	ED	EI	EI	ED	ED	Pic	Unl ^b
	100	1000	100	1000	100	1000	100	1000	100	1000	100	1000	A	
Serie 1 (18 plantas)														
3	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	-	-	-	-	+	-	-	-	-	-	-	-	-	+(D)
3	-	-	-	-	+	-	-	-	+	-	-	-	-	+(D)
	-	-	-	-	+	-	-	-	+	-	-	-	-	+(D)
	-	-	-	-	+	-	-	-	+	-	-	-	-	+(D)
1	+	+	-	-	-	-	-	-	-	-	-	-	-	-
	-	-	-	-	+	-	-	-	+	-	+	+	-	-
	+	+	-	-	+	+	+	+	+	-	-	-	-	np ^c
2	-	-	-	-	+	-	-	-	+	+	+	+	-	np
	+	+	+	-	-	-	+	-	-	-	-	-	-	np
	+	+	+	+	-	-	-	-	-	-	-	-	-	np
1	+	+	+	+	+	+	+	+	+	+	+	+	-	np
	+	+	+	+	+	+	+	+	+	+	+	+	-	np
	+	+	+	+	+	+	+	+	+	+	+	+	-	np

ES 2 284 555 T3

^a Indica el número de transformantes con un patrón particular de secuencias del vector presentes en las plantas transformadas.

^b Los fragmentos de repetición en el extremo externo derecho (D) o izquierdo (I) están presentes en cualquiera de los vectores de ADN-T, no unidos a cualquiera de los extremos.

^c No probado

TABLA 2

Presencia (+) o ausencia (-) de secuencias del vector unidas al EI y al ED del ADN-T en plantas de N. tabacum transgénicas transformadas con el vector de ADN-T Ksb

N° de co-trans ^a	EI100	EI1000	ED100	ED1000	PicA	Unl ^b
Serie 2 (36 plantas)	Ksb					
17	-	-	-	-	-	-
3	+	-	-	-	-	-
2	+	+	-	-	-	+(D)
3	+	+	-	-	-	-
1	-	-	+	+	-	-
1	+	-	+	-	-	np ^c
1	+	+	+	-	-	np
8	+	+	+	+	-	np

^a Indica el número de transformantes con un patrón particular de secuencias del vector presentes en las plantas transformadas.

^b Los fragmentos de repetición en el extremo externo derecho (D) o izquierdo (I) están presentes en cualquiera de los vectores de ADN-T, no unidos a cualquiera de los extremos.

^c No probado

Bibliografía

EP0332104 **Montoya, A., Duesing, J., Harms, C., Meins, F., Payne, G., Ryals, J. y Sperisen, C.** 1989. Chemically regulatable DNA sequences and genes and uses thereof.

EP0823480 **Gaelweiler, L., Palme, K., Grosskopf-Kroiler, D., Schell, J. y Moore, I.** 1998. Controlled gene expression in plants.

US5527695 **Hodges T. K. y Lyznik L.** 1996. Controlled modification of eukaryotic genomes.

US5801027 **Ramsay, N., Bennett, M. y May, S.** 1998. Method of using transactivation proteins to control gene expression in transgenic plants.

W090/08826 **Bright, S. W. J., Greenland, A. J., Schuch, W. W. y Bridges, I. G.** 1990. Gene switch.

W097/12046 **Hansen, G. y Chiton, M.-D.** 1997. Improved integration of exogenous DNA delivered to eukaryotic cells.

- W097/30164 **Haseloff, J. P. y Hodge, S.** 1997. Improvements in or relating to gene expression.
- W098/45455 **Beijersbergen, A. G., de Groot, M. J. A., Gouka, R. J., Hooykaas, P. J. y Bundock, P.** 1998. *Agrobacterium*-mediated transformation of moulds, in particular those belonging to the genus *Aspergillus*.
- 5 W098/59062 **Liu, Z.-B. y Odell, J. T.** 1998. Specific gene activation by chimeric GAL4 transcription factors in stable transgenic plants.
- 10 W099/01563 **Stuiver, M. H., Ponstein, A. S., Ohl, S. A., Goddijn, O. J. M., Simons, L. H., Dekker, B. M. M., Hoekstra, S. y Tigelaar, H.** 1999. Plasmids for plant transformation and method for using the same.
- W099/14348 **Lefebvre, D., Ko, K., Ko, Z. y Trembla, L.** 1999. In planta transformation of plants.
- 15 W099/23202 **Ow, D., y Srivastava, V.** 1999. Resolution of complex integration patterns to obtain single copy transgenes.
- W099/25821 **Baszczynski, C. L., Bowen, B. A., Peterson, D. J. y Tagliani, L. A.** 1999. Compositions and methods for genetic modification of plants.
- 20 W099/25840 **Baszczynski C. L., Lyznik L., Gordon-Kamm W. J., Rao A. G., Tagliani, L. A. y Guan, X.** 1999. A novel method for the integration of foreign DNA into eukaryotic genomes.
- 25 **Bundock, P., den Dulk-Ras, A., Beijersbergen, A. y Hooykaas, P. J.** 1995. Trans- kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. *EMBO J* 14: 3206-3214.
- Caplan, A. B., van Montagu, M. y Schell, J.** 1985. Genetic analysis of integration mediated by single T-DNA borders. *J Bacteriol* 161, 655-664.
- 30 **Chang, A. C. y Cohen, S. N.** 1987. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J Bacteriol* 134: 1141-1156.
- Clough, S. J. y Bent, A. F.** 1998. Floral dip: a simplified method for *Agrobacterium*- mediated transformation of *Arabidopsis thaliana*. *Plant J* 16: 735-743.
- 35 **Cluster, P. D., O'Dell, M., Metzlaiff, M. y Flavell, R. B.** 1996. Details of T-DNA structural organization from a transgenic Petunia population exhibiting co-suppression. *Plant Mol Biol* 32: 1197-1203.
- Dale, E. C. y Ow, D. W.** 1990. Intra-and intermolecular site-specific recombination in plant cells mediated by bacteriophage P1 recombinase. *Gene* 91: 79-85.
- 40 De **Buck, S., Jacobs, A., van Montagu, M. y Depicker, A.** 1998. *Agrobacterium tumefaciens* transformation and cotransformation frequencies of *Arabidopsis thaliana* root explants and tobacco protoplasts. *Mol Plant-Microbe Interact* 11: 449-457.
- 45 De **Buck, S., Jacobs, A., van Montagu, M. y Depicker, A.** 1999. The DNA sequences of T-DNA junctions suggest that complex T-DNA loci are formed by a recombination process resembling T-DNA integration. *Plant J* (in press).
- de **Groot, M. J., Bundock, P., Hooykaas, P. J. y Beijersbergen, A. G.** 1998. *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. *Natl Biotechnol* 16: 839-842.
- 50 De **Neve, M., De Buck, S., Jacobs, A., van Montagu, M. y Depicker A.** 1997. T-DNA integration patterns in co-transformed plant cells suggest that T-DNA repeats originate from ligation of separate T-DNAs. *Plant J* 11: 15-29.
- 55 **Deng, W., Chen, L., Peng, W. T., Liang, X., Sekiguchi, S., Gordon, M. P., Comai, L. y Nester, E. W.** 1999. VirE1 is a specific molecular chaperone for the exported single- stranded-DNA-binding protein VirE2 in *Agrobacterium*. *Mol Microbiol* 31: 1795-1807.
- Durrenberger, F., Cramer, A., Hohn, B. y Koukolikova-Nicola, Z.** 1989. Covalently bound VirD2 protein of *Agrobacterium tumefaciens* protects the T-DNA from exonucleolytic degradation. *Proc Natl Acad Sci USA* 86: 9154-9158.
- 60 **Errampali, D., Patton, D., Castle, L., Mickelson, K., Hansen, K., Schnall, J., Feldmann, K. y Meinke, D.** 1991. Embryogenic lethals and T-DNA insertional mutagenesis in *Arabidopsis*. *Plant Cell* 3: 149-157.
- 65 **Fedoroff, N. V. y Smith, D. L.** 1993. A versatile system for detecting transposition in *Arabidopsis*. *Plant J* 3: 273-289.
- Feldmann, K. A.** 1991. T-DNA insertion mutagenesis in *Arabidopsis* a mutational spectrum. *Plant J* 1: 71-82.

Gelvin, S. B. 1998a. The introduction and expression of transgenes in plants. *Curr Opin Biotechnol* 9: 227-232.

Gelvin, S. B. 1998b. *Agrobacterium* VirE2 proteins can form a complex with T strands in the plant cytoplasm. *J Bacteriol* 180: 4300-4302.

Ghai, J. y Das, A. 1989. The *virD* operon of *Agrobacterium tumefaciens* Ti plasmid encodes a DNA-relaxing enzyme. *Proc Natl Acad Sci USA* 86: 3109-3113.

Gheysen, G., Angenon, G. y van Montagu, M. 1998. *Agrobacterium*-mediated plant transformation: a scientifically intriguing story with significant applications. In K. Lindsey (Ed.), *Transgenic Plant Research*. Harwood Academic Publishers, Amsterdam, pp. 1-33.

Gielen, J., De Beuckeleer, M., Seurinck, J., Deboeckl, F., De Greve, H., Lemmers, M., van Montagu, M. y Schell, J. 1984. The complete nucleotide sequence of the TL-DNA of the *Agrobacterium tumefaciens* plasmid pTiAch5. *EMBO J* 3: 835-846.

Gielen, J., Terryn, N., Villaroel, R. y van Montagu, M. 1999. Complete nucleotide sequence of the T-DNA region of the plant tumour-inducing *Agrobacterium tumefaciens* Ti plasmid pTiC58. *J Exper Bot* 50: 1421-1422.

Gouka, R. J., Gerk, C., Hooykaas, P. J., Bundock, P., Musters, W., Verrips, C. T. y de Groot, M. J. 1999. Transformation of *Aspergillus awamori* by *Agrobacterium tumefaciens*-mediated homologous recombination. *Natl Biotechnol* 17, 598-601.

Guralnick, B., Thomsen, G. y Citovsky, V. 1996. Transport of DNA into the nuclei of *Xenopus* oocytes by a modified VirE2 protein of *Agrobacterium*. *Plant Cell* 8: 363- 373.

Hansen, G. y Chiton, M.-D. 1996. "Agrolistic" transformation of plant cells: integration of T-strands generated in planta. *Proc Natl Acad Sci USA* 93: 14978-14983.

Hansen, G., Shilito, R. D. y Chiton, M.-D. 1997. T-strand integration in maize protoplasts after codelivery of a T-DNA substrate and virulence genes. *Proc Natl Acad Sci USA* 94: 11726-11730.

Hanson, B., Engler, D., Moy, Y., Newman, B., Ralston, E. y Gutterson, N. 1999. A simple method to enrich an *Agrobacterium*-transformed population for plants containing only T-DNA sequences. *Plant J* 19: 727-734.

Horsch, R. B., Fry, J. E., Hoffmann, N. L., Eichholtz, D., Rogers, S. G. y Fraley, R. T. 1985. A simple and general method for transferring genes into plants. *Science* 227: 1229-1231.

Horsch, R. B. y Klee, H. J. 1986. Rapid assay of foreign gene expression in leaf discs transformed by *Agrobacterium tumefaciens*: role of T-DNA borders in the transfer process. *Proc Natl Acad Sci USA* 83: 4428-4432.

Iglesias, V. A., Moscone, E. A., Papp, I., Neuhuber, F., Michalowski, S., Phelan, T., Spiker, S. Matzke, M. y Matzke, A. J. M. 1997. Molecular and cytogenetic analyses of stably and unstably expressed transgene loci in tobacco. *Plant Cell* 9: 1251-1264.

Jayaswal, R. K., Veluthambi, K., Gelvin, S. B. y Slightom, J. L. 1987. Double-stranded cleavage of T-DNA and generation of single-stranded T-DNA molecules in *Escherichia coli* by a *virD*-encoded border-specific endonuclease from *Agrobacterium tumefaciens*. *J Bacteriol* 169: 5035-5045.

Jakowitsch, J., Papp, I., Moscone, E. A., van der Winden, J., Matzke, M. y Matzke, A. J. M. 1999. Molecular and cytogenetic characterization of a transgene locus that induces silencing and methylation of homologous promoters in trans. *Plant J* 17: 131-140.

Jen, G. C. y Chiton, M. D. 1986a. Activity of T-DNA borders in plant cell transformation by mini-T plasmids. *J Bacteriol* 166: 491-499.

Jen, G. C. y Chiton, M. D. 1986b. The right border region of pTiT37 T-DNA is intrinsically more active than the left border region in promoting T-DNA transformation. *Proc Natl Acad Sci USA* 83: 3895-3899.

Jones, J. D. G., Dunsmuir, P., y Bedbrook, J. (1985). High level expression of introduced chimaeric genes in regenerated transformed plants. *EMBO J.* 4, 2411-2418.

Koncz, C., Nemeth, K., Redei, G. P. y Schell, J. 1992. T-DNA insertional mutagenesis in *Arabidopsis*. *Plant Mol. Biol* 20: 963-976.

- Kononov, M. E., Bassuner, B. y Gelvin, S. B. 1997. Integration of T-DNA binary vector “backbone” sequences into the tobacco genome: evidence for multiple complex patterns of integration. *Plant J* 11: 945-957.
- Maniatis, T., Fritsch, E. F. y Sambrook, J. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, *Cold Spring Harbor*.
- Martineau, B., Voelker, T. A. y Sanders, R. A. 1994. On defining T-DNA. *Plant Cell* 6: 1032-1033 (Letter to the editor).
- Matzke, A. J. M. y Matzke, M. A. 1998. Position effects and epigenetic silencing of plant transgenes. *Curr Opin Plant Biol* 1: 142-148.
- Mysore, K. S., Bassuner, B., Deng, X. B., Darbinian, N. S., Motchoulski, A., Ream, W. y Gelvin, S. B. 1998. Role of the *Agrobacterium tumefaciens* VirD2 protein in T-DNA transfer and integration. *Mol Plant Microbe Interact* 11: 668-683.
- Narasimhulu, S. B., Deng, X. B., Sarria, R. y Gelvin, S. B. 1996. Early transcription of *Agrobacterium* T-DNA genes in tobacco and maize. *Plant Cell* 8: 873-886.
- Onouchi, H., Nishihama, R., Kudo, M., Machida, Y., y Machida, C. (1995). Visualization of site-specific recombination catalyzed by a recombinase from *Zygosaccharomyces rouxii* in *Arabidopsis thaliana*. *Mol. Gen. Genet.* 247, 653-660.
- Onouchi, H., Yokoi, K., Machida, C., Matsuzaki, H., Oshima, Y., Matsuoka, K., Nakamura, K., y Machida, Y. (1991). Operation of an efficient site-specific recombination system of *Zygosaccharomyces rouxii* in tobacco cells. *Nucleic Acids Res.* 19, 6373-6378.
- Ooms, G., Bakker, A., Molendijk, L., Wullems, G. J., Gordon, M. P., Nester, E. W. y Schilperoort, R. A. 1982. T-DNA organization in homogeneous and heterogeneous octopine-type crown gall tissues of *Nicotiana tabacum*. *Cell* 30: 589-597.
- Osborne, B. I., Wirtz, U. y Baker, B. 1995. A system for insertional mutagenesis and chromosomal rearrangement using the Ds transposon and Cre-lox. *Plant J* 7: 687-701.
- Pansegrau, W. y Lanka, E. 1996. Mechanisms of initiation and termination reactions in conjugative DNA processing. *J Biol Chem* 271: 13068-13076.
- Peralta, E. G., Hellmiss, R. y Ream, W. 1986. Overdrive, a T-DNA transmission enhancer on the *A. tumefaciens* tumour-inducing plasmid. *EMBO J* 5, 1137-1142.
- Ramanathan, V. y Veluthambi, K. 1995. Transfer of non-T-DNA portions of the *Agrobacterium tumefaciens* Ti plasmid pTiA6 from the left terminus of T-DNA. *Plant Mol Biol* 28: 1149-1154.
- Relic, B., Andjelkovic, M., Rossi, L., Nagamine, Y. y Hohn, B. 1998. Interaction of the DNA modifying proteins VirD1 and VirD2 of *Agrobacterium tumefaciens*: analysis by subcellular localization in mammalian cells. *Proc Natl Acad Sci USA* 95: 9105-9110.
- Rossi, L., Hohn, B. y Tinland, B. 1993. The VirD2 protein of *Agrobacterium tumefaciens* carries nuclear localization signals important for transfer of T-DNA to plant. *Mol Gen Genet* 239: 345-353.
- Rossi, L., Hohn, B. y Tinland, B. 1996. Integration of complete transferred DNA units is dependent on the activity of virulence E2 protein of *Agrobacterium tumefaciens*. *Proc Natl Acad Sci USA* 93: 126-130.
- Schlappi, M., Smith, D. y Fedoroff, N. 1993. TnpA trans-activates methylated maize Suppressor-mutator transposable elements in transgenic tobacco. *Genetics* 133: 1009-1021.
- Shaw, C. H., Watson, M. D., Carter, G. H. y Shaw, C. H. 1984. The right hand copy of the opaline Ti-plasmid 25 bp repeat is required for tumour formation. *Nucleic Acids Res* 12: 6031-6041.
- Shurvinton, C. E., Hodges, L., Ream, W. 1992. A nuclear localization signal and the C- terminal omega sequence in the *Agrobacterium tumefaciens* VirD2 endonuclease are important for tumor formation. *Proc Natl Acad Sci USA* 89: 11837-11841.
- Sugita, K., Kasahara, T., Matsunaga, E., y Ebinuma, H. (2000). Technical advance: A transformation vector for the production of marker-free transgenic plants containing a single copy transgene at high frequency [In Process Citation]. *Plant J.* 22, 461-469.

Tinland, B., Schoumacher, F., Gloeckler, V., Bravo-Angel, A. M. y Hohn, B. 1995. The *Agrobacterium tumefaciens* virulence D2 protein is responsible for precise integration of T-DNA into the plant genome. *EMBO J* 14: 3585-3595.

Valvekens, D., van Montagu, D. y van Lijsebettens, M. 1988. *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* root explants by using kanamycin selection. *Proc Natl Acad Sci USA* 85: 5536-5540.

Van der Graaff, E., den Dulk-Ras, A. y Hooykaas, P. J. J. 1996. Deviating T-DNA transfer from *Agrobacterium tumefaciens* to plants. *Plant Mol Biol* 31: 677-681.

Van Haaren, M. J. J., Sedee, N. J. A., Schilperoort, R. A. y Hooykaas, P. J. J. 1987. Overdrive is a T-region transfer enhancer which stimulates T-strand production in *Agrobacterium tumefaciens*. *Nucl Acids Res* 15: 8983-8997.

Van Lijsebettens, M., Vanderhaegen, R., y van Montagu, M. 1991. Insertional mutagenesis in *Arabidopsis thaliana*: isolation of a T-DNA-linked mutation that alters leaf morphology. *Theor Appl Genet* 81: 277-284.

Van Sluys, M. A., Tempe, J. y Fedoroff, N. 1987. Studies on the introduction and mobility of the maize Activator element in *Arabidopsis thaliana* and *Daucus carota*. *EMBO J* 6: 3881-3889.

Virts, E. L. y Gelvin, S. B. 1985. Analysis of transfer of tumor-inducing plasmids from *Agrobacterium tumefaciens* to Petunia protoplasts. *J. Bacteriol* 162: 1030-1038.

Wang, K., Genetello, C., van Montagu, M. y Zambryski, P. C. 1987. Sequence context of the T-DNA border repeat element determines its relative activity during T-DNA transfer to plant cells. *Mol Gen Genet* 210: 338-346.

Wang, K., Herrera-Estrella, A. y van Montagu, M. 1990. Overexpression of *virD1* and *virD2* genes in *Agrobacterium tumefaciens* enhances T-complex formation and plant transformation. *J Bacteriol* 172: 4432-4440.

Waters, V. L. y Guiney, D. G. 1993. Processes at the nick region link conjugation, T-DNA transfer and rolling circle replication. *Mol Microbiol* 9: 1123-1130.

Wenck, A., Czako, M., Kanevski, I. y Marton, L. 1997. Frequent collinear long transfer of DNA inclusive of the whole binary vector during *Agrobacterium*-mediate transformation. *Plant Mol Biol* 34: 913-922.

Wolters, A.-M. A., Trindade, L. M., Jacobsen, E. y Visser, R. G. F. 1998. Fluorescence *in situ* hybridization on extended DNA fibres as a tool to analyse complex T-DNA loci in potato. *Plant J* 13: 837-847.

Yanofsky, M. F., Porter, S. G., Young, C., Alright, L. M., Gordon, M. P. y Nester, E. W. 1986. The *virD* operon of *Agrobacterium tumefaciens* encodes a site-specific endonuclease. *Cell* 47: 471-477.

Young, C. y Nester, E. W. 1988. Association of the *virD2* protein with the 5' end of T strands in *Agrobacterium tumefaciens*. *J Bacteriol* 170: 3367-3374.

Yusibov, V. M., Steck, T. R., Gupta, V. y Gelvin, S. B. 1994. Association of single-stranded transferred DNA from *Agrobacterium tumefaciens* with tobacco cells. *Proc Natl Acad Sci USA* 91: 2994-2998.

Ziemienowicz, A., Gorlich, D., Lanka, E., Hohn, B., y Rossi, L. (1999). Import of DNA into mammalian nuclei by proteins originating from a plant pathogenic bacterium. *Proc. Natl. Acad. Sci. U. S. A* 96, 3729-3733.

Zupan, J. R. y Zambryski, P. 1995. Transfer of T-DNA from *Agrobacterium* to the plant cell. *Plant Physiol* 107: 1041-1047.

REIVINDICACIONES

1. Un vector de transformación de ADN-T que comprende ADN-T con extremos de ADN-T flanqueantes a izquierda y derecha modificados, **caracterizado** porque el extremo derecho de ADN-T modificado consiste en una única secuencia central del extremo derecho flanqueada por una región externa del extremo derecho, y el extremo izquierdo de ADN-T modificado consiste en (i) una única secuencia central del extremo izquierdo, una región externa del extremo izquierdo y una región proximal en el extremo izquierdo de la región interna con una longitud preferible de 10 a 100 pb, **caracterizado** adicionalmente porque dicha región proximal en extremo izquierdo de la región interna está enriquecida en número de residuos de A y T, estando el porcentaje de residuos de AT entre el 60 y el 85%, (ii) una repetición de las extremos izquierdos del ADN-T en la que dicha repetición comprende al menos un extremo de ADN-T de tipo nopalina y al menos un extremo izquierdo de ADN-T de tipo octopina, **caracterizado** adicionalmente porque dicho extremo izquierdo de ADN-T modificado comprende al menos una secuencia central del extremo izquierdo, o (iii) una región externa del extremo izquierdo y una repetición de al menos dos secuencias centrales del extremo izquierdo separadas por una secuencia de al menos 10-20 pares de bases.
2. El vector según la reivindicación 1 en el que dicha secuencia de al menos 10-20 pares de bases que separa dichas secuencias centrales del extremo izquierdo porta codones de terminación en los tres marcos de lectura y en ambas direcciones.
3. El vector según la reivindicación 1 en el que dicho extremo izquierdo de ADN-T modificado consiste en una repetición de extremos izquierdos de ADN-T que comprende al menos una región en el extremo izquierdo de tipo nopalina integral adyacente a y aguas abajo o aguas arriba de la región en el extremo izquierdo de tipo octopina integral.
4. El vector según cualquiera de las reivindicaciones 1 a 3, **caracterizado** porque dichos extremos de ADN-T modificados además comprenden sitios de recombinación aguas abajo de la secuencia central del extremo izquierdo y sitios de recombinación aguas arriba de la secuencia central del extremo derecho, estando dichos sitios de recombinación que organizados como repeticiones.
5. El vector según la reivindicación 4 que comprende adicionalmente una segunda copia de una región en el extremo izquierdo aguas arriba de la única región externa del extremo derecho y en el que dicho sitio de recombinación está situado aguas arriba de la secuencia central de dicha segunda región en el extremo izquierdo.
6. El vector según cualquiera de las reivindicaciones 4 ó 5 que comprende adicionalmente un gen que codifica una recombinasa localizado aguas abajo de dicho sitio de recombinación aguas abajo de dicha secuencia central del extremo izquierdo.
7. El vector según la reivindicación 6 en el que dicho gen que codifica una recombinasa está situado adyacente a y aguas abajo de la región externa del extremo izquierdo.
8. El vector según la reivindicación 6 ó 7 que comprende dicho gen que codifica una recombinasa flanqueado por repeticiones de sitios de recombinación y comprendiendo adicionalmente dicho vector una segunda copia de una región en el extremo izquierdo localizada aguas abajo de dicho gen que codifica una recombinasa flanqueado por dichos sitios de recombinación.
9. El vector según la reivindicación 8 que comprende adicionalmente sitios de recombinación organizados como repeticiones aguas abajo de la segunda secuencia central del extremo izquierdo y aguas arriba de la única secuencia central del extremo derecho.
10. El vector según cualquiera de las reivindicaciones 4 a 9 **caracterizado** porque dichos sitios de recombinación están localizados adyacentes a y aguas abajo y/o aguas arriba de las secuencias centrales del extremo izquierdo y/o derecho o están separados aguas abajo y/o aguas arriba de las secuencias centrales del extremo izquierdo y/o derecho por una secuencia de al menos 10-20 pb de longitud.
11. El vector según la reivindicación 10 en el que dicha secuencia de al menos 10-20 pb de longitud porta codones de terminación en los tres marcos de lectura y en ambas direcciones.
12. El vector según cualquiera de las reivindicaciones 4 a 11 en el que dichos sitios de recombinación organizados como repeticiones se definen como sitios de recombinación específicos de sitio organizados como repeticiones directas o como secuencias terminales de transposones organizadas como repeticiones invertidas; y en el que dicho gen que codifica una recombinasa se define como un gen que codifica una recombinasa específica de sitio o un gen que codifica una transposasa, respectivamente.
13. Un vector para la transformación mediada por *Agrobacterium*, para la transformación agrolística o para fines de terapia génica con extremos de ADN-T modificados como se ha definido en cualquiera de las reivindicaciones 1 a 12, en el que dicho vector se escoge entre vectores de transformación binarios, vectores de transformación súper-binarios, vectores de transformación co-integrados, vectores de transformación derivados de Ri y vectores que portan ADN-T usados en transformación agrolística o terapia génica.

ES 2 284 555 T3

14. Un procedimiento para la obtención de plantas, levaduras, mohos u hongos filamentosos transgénicos que consiste en una transformación mediada por *Agrobacterium* de dichas plantas, levaduras, mohos u hongos filamentosos con un vector de transformación escogido entre el grupo de:

a) un vector de transformación según cualquiera de las reivindicaciones 1 a 3; o

b) un vector de transformación definido en cualquiera de las reivindicaciones 4 ó 5 en combinación con el suministro de una recombinasa o una transposasa, para curar dichas células transformadas resultantes de secuencias del esqueleto del vector integradas que posiblemente se originan en dicho vector; o

c) un vector de transformación definido en cualquiera de las reivindicaciones 6 a 12 para curar dichas células transformadas resultantes de secuencias del esqueleto del vector integradas que posiblemente se originan en dicho vector, opcionalmente en combinación con el suministro de una recombinasa o transposasa; o

d) un vector de transformación según la reivindicación 13.

15. Un procedimiento para prevenir la integración de secuencias del esqueleto del vector en una célula transformada mediada por *Agrobacterium* que comprende el uso de un vector según cualquiera de las reivindicaciones 1 a 13.

16. El procedimiento según la reivindicación 15 que comprende adicionalmente el incremento de la producción de las proteínas VirD1 y/o VirD2 comprendidas dentro del complejo de corte de ADN-T.

17. El procedimiento de la reivindicación 16 que supone la integración de al menos una copia adicional del locus *virD* en el genoma o en una entidad extracromosómica de *Agrobacterium*.

18. El procedimiento de la reivindicación 17 en el que dicha copia adicional del locus *virD* se selecciona entre el locus *virD* de tipo octopina y el locus *virD* de tipo nopalina.

19. Un procedimiento de transformación agrolística de una célula eucariota usando un vector de ADN-T de cualquiera de las reivindicaciones 1 a 13.

20. Una célula hospedadora recombinante que contiene un vector de cualquiera de las reivindicaciones 1 a 13.

21. La célula hospedadora de la reivindicación 20 identificada como *Agrobacterium tumefaciens*.

22. Una planta o una célula de una planta transgénica obtenible mediante un procedimiento de transformación mediada por *Agrobacterium* según cualquiera de las reivindicaciones 14 a 19, una de sus partes o de su progenie que comprende el ADN-T resultante del vector de cualquiera de las reivindicaciones 1 a 13, en el que dicho ADN-T comprende una región proximal en el extremo izquierdo de la región interna con una longitud preferible de 10 a 100 pb, **caracterizada** adicionalmente porque dicha región proximal en el extremo izquierdo de la región interna está enriquecida en el número de residuos de A y T, estando el porcentaje de residuos de AT entre el 60 y el 85%.

23. Una levadura, moho u hongo filamentoso trasgénico que comprende el ADN-T resultante del vector de cualquiera de las reivindicaciones 1 a 13 obtenible mediante un procedimiento de transformación mediada por *Agrobacterium* según cualquiera de las reivindicaciones 14 a 19, en el que dicho ADN-T comprende una región proximal en el extremo izquierdo de la región interna con una longitud preferible de 10 a 100 pb, **caracterizada** adicionalmente porque dicha región proximal en el extremo izquierdo de la región interna está enriquecida en el número de residuos de A y T, estando el porcentaje de residuos de AT entre el 60 y el 85%.

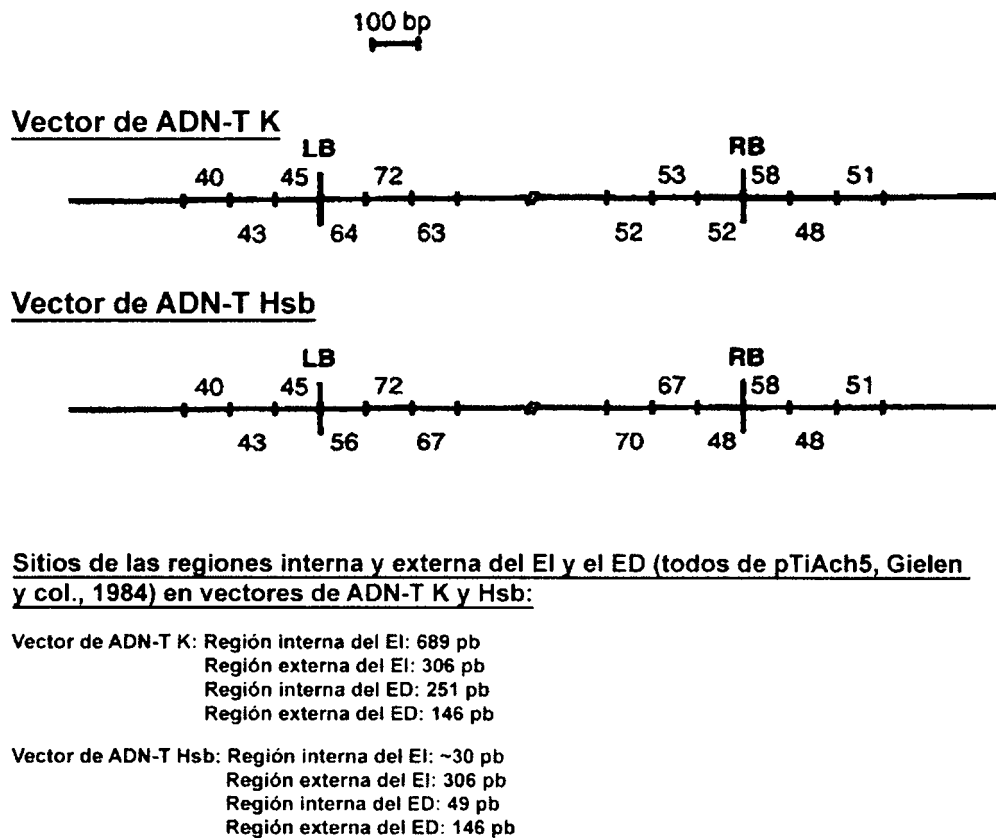


FIGURA 1

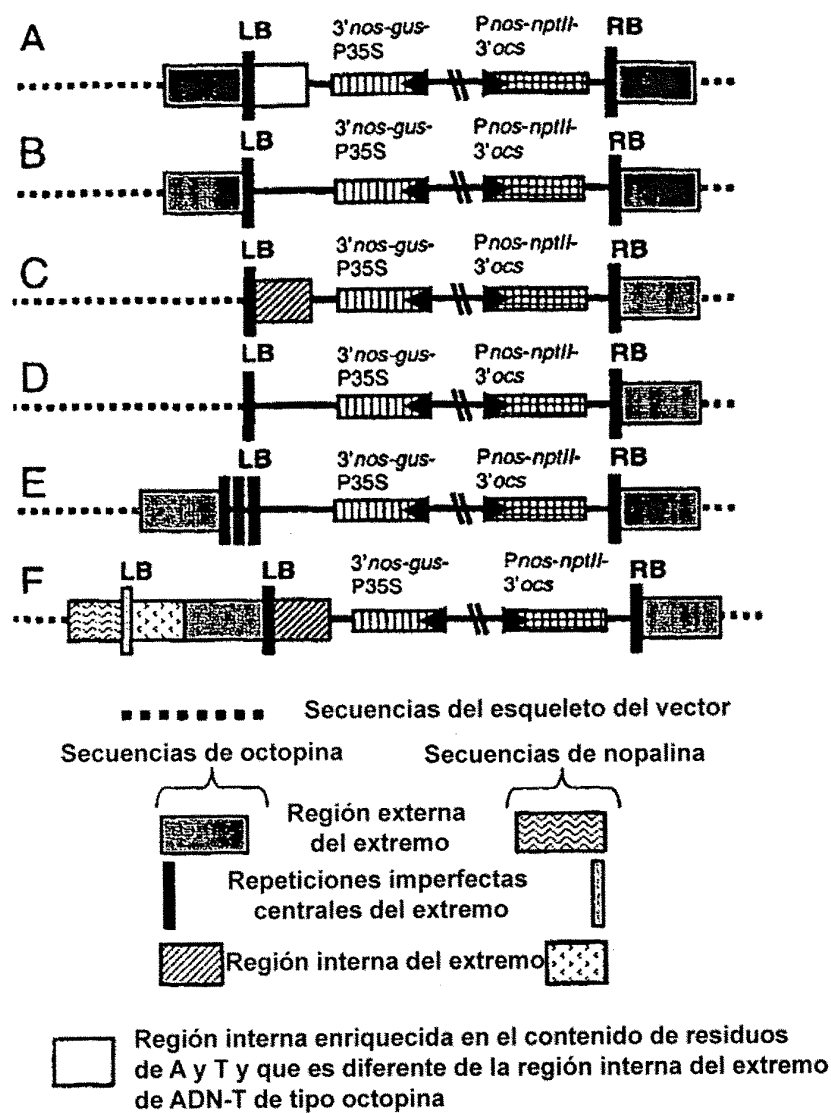


FIGURA 2

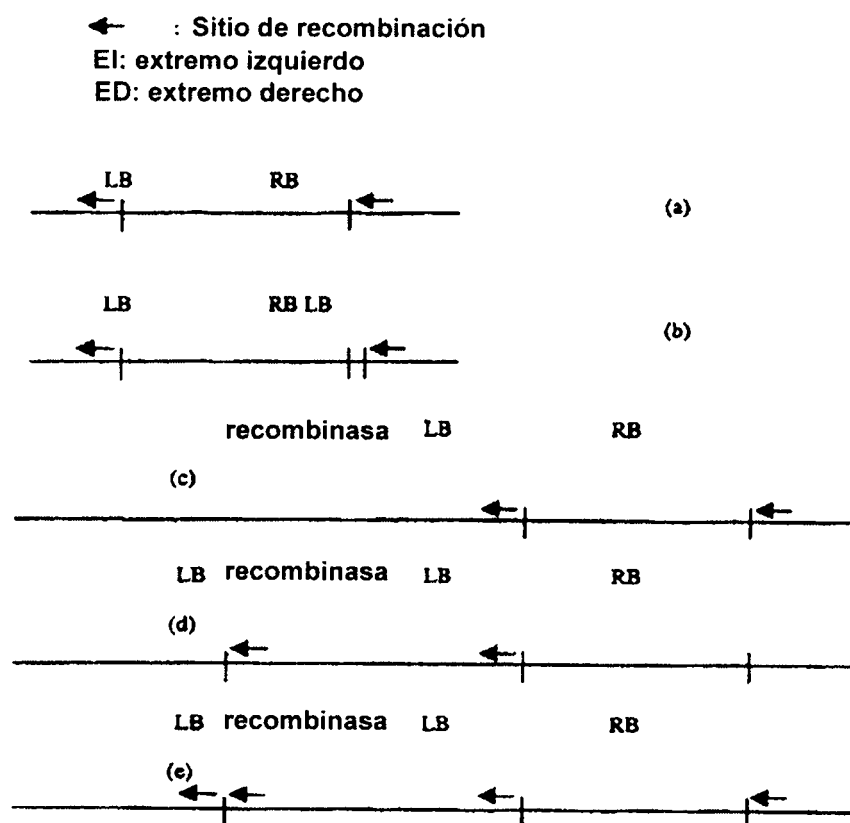


FIG. 3

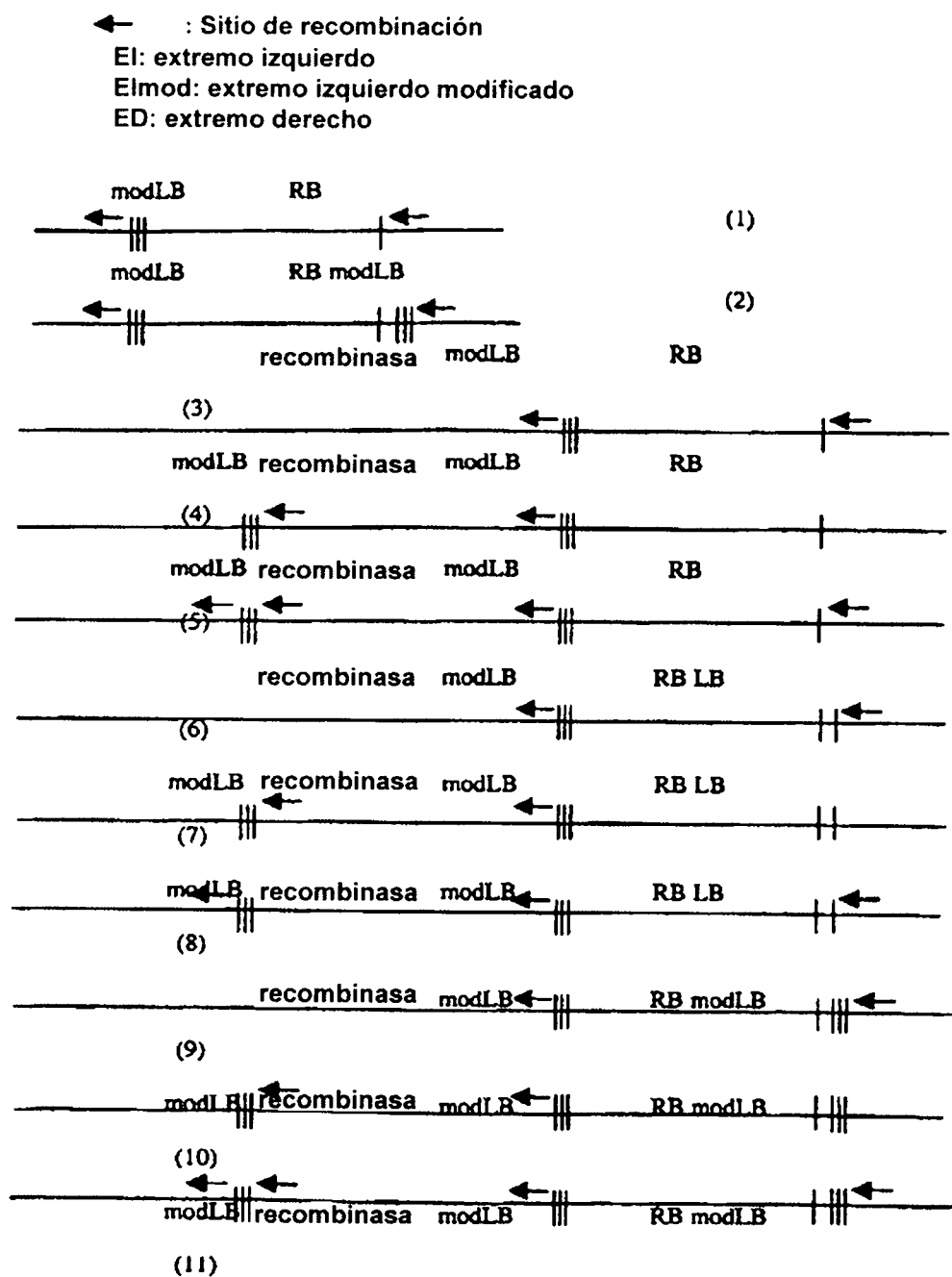


FIG. 4

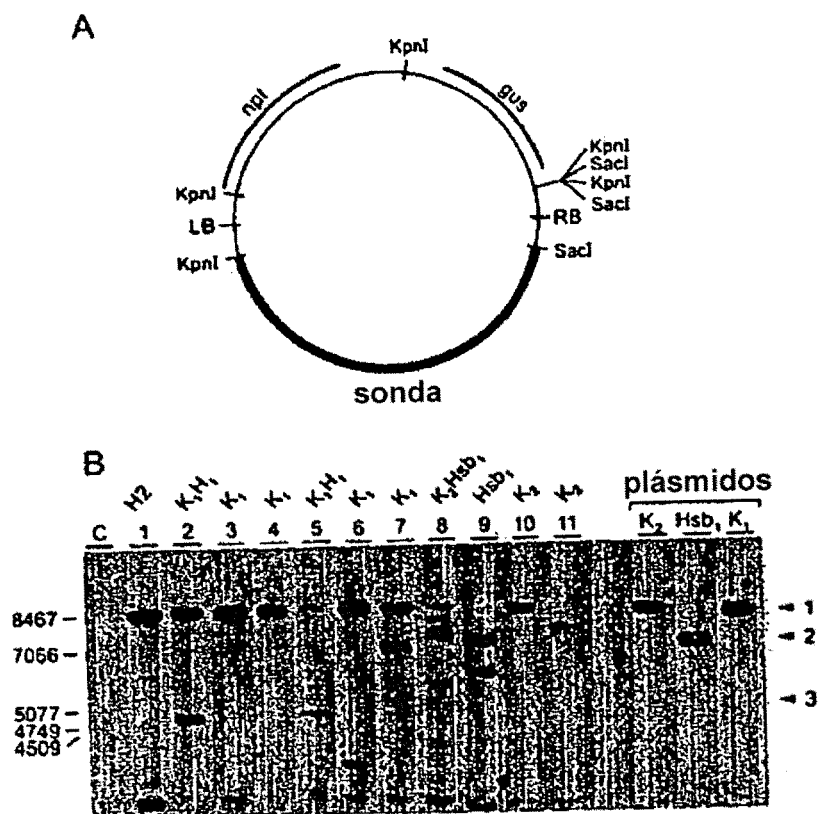


FIG. 5

(1) hibridación

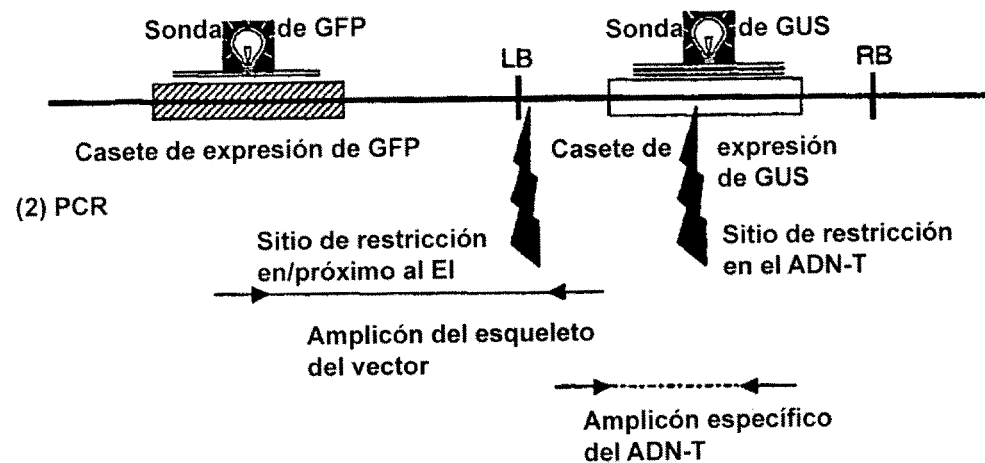


FIG. 6