



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 362 100**

51 Int. Cl.:  
**C12N 9/22** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **05784046 .4**

96 Fecha de presentación : **29.07.2005**

97 Número de publicación de la solicitud: **1771555**

97 Fecha de publicación de la solicitud: **11.04.2007**

54 Título: **Método para producir endonucleasas altamente sensibles, nuevas preparaciones de endonucleasas y usos de las mismas.**

30 Prioridad: **30.07.2004 PCT/EP2004/009166**  
**30.07.2004 PCT/EP2004/009159**

45 Fecha de publicación de la mención BOPI:  
**28.06.2011**

45 Fecha de la publicación del folleto de la patente:  
**28.06.2011**

73 Titular/es: **GENOPLANTE-VALOR**  
**28 rue du Docteur Finlay**  
**75015 Paris, FR**  
**Institut National de la Recherche Agronomique**  
**(INRA)**

72 Inventor/es: **Bendahmane, Abdelhafid;**  
**Sturbois, Bénédicte;**  
**Triques-Rostaing, Karine y**  
**Caboche, Michel**

74 Agente: **Martín Santos, Victoria Sofia**

ES 2 362 100 T3

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

## DESCRIPCIÓN

Método para producir endonucleasas altamente sensibles, nuevas preparaciones de endonucleasas y usos de las mismas.

- 5
- La invención se refiere a la identificación y preparación de endonucleasas específicas de emparejamiento erróneo que tienen una actividad y sensibilidad elevada y una amplia especificidad de sustrato.
- 10 A comienzos del último siglo, el descubrimiento de la posibilidad de inducir mutaciones en el ADN por radiaciones o productos químicos ha aportado una considerable expectativa para comprender la función de genes *in vivo*. Desde entonces, la mutagénesis y la variación de secuencias naturales se han usado ampliamente para identificar nuevas funciones, genes correspondientes a una función específica así como sitios activos dentro de una proteína específica.
- 15 Un aspecto crítico para la implementación de esta estrategia, en particular en el caso de mutaciones puntuales, es la elección de métodos de detección de mutaciones que se diseñan para seleccionar grandes tramos de ADN sin reducir la sensibilidad o especificidad del diagnóstico, proporcionando al mismo tiempo información sobre la localización de la mutación. Entre las herramientas y métodos más usados están los basados en el ADN defectuosamente emparejado que podría crearse *in vitro* por desnaturalización y renaturalización de dos moléculas
- 20 de ADN. Los emparejamientos erróneos se detectan en estas moléculas heteroduplex usando productos químicos como aglutinantes o moléculas de surcos que pueden escindir específicamente el ADN monocatenario en el sitio de emparejamiento erróneo. Como alternativa, se han usado endonucleasas específicas monocatenarias para escindir en ADN en el sitio de emparejamiento erróneo. La mayoría de las endonucleasas descritas en este documento pertenecen ahora a la clase de nucleasas S1/P1.
- 25 Se sabe que nucleasas tales como S1, P1 y la nucleasa de frijol chino, pertenecientes a la misma familia denominada: "familia de nucleasas S1/P1" o "familia de nucleasas S1" escinden el ADN en regiones de una sola cadena. Sin embargo, estas nucleasas tienen un pH óptimo ácido en el intervalo de 4,0-5,0.
- 30 Esto es desventajoso para detectar emparejamientos erróneos, ya que valores bajos de pH favorecen la despurinización del ADN y desestabilizan los duplex de ADN, lo que conduce a una degradación no específica de ADN y reduce la sensibilidad y especificidad en la detección.
- Hace algunos años, en extractos de diversas plantas, OLEYKOWSKI et al. (Nucleic Acids Res, 26, 4597-4602, 1998)
- 35 detectaron una actividad endonucleasa de emparejamiento erróneo que tenía un pH óptimo neutro (aproximadamente pH 8) y que realizaba una escisión en una sola cadena en el lado 3' de un sitio de emparejamiento erróneo. Estos autores describieron que esta actividad endonucleasa de emparejamiento erróneo se asociaba con manoxil glucoproteínas en extractos de brotes de alfalfa, espárragos, apio y tomate. La enzima procedente del apio, denominada CEL I, se purificó de los tallos del apio por etapas sucesivas de precipitación en
- 40 sulfato de amonio, unión a una columna de agarosa-concavalina A y elución mediante  $\alpha$ -D-manosa, unión a una columna de fosfocelulosa y elución mediante un gradiente lineal de KCl, y fraccionamiento por cromatografía de exclusión de tamaño. La preparación de CEL I obtenida de esta manera contenía diversas bandas de proteínas de 34-39 kDa.
- 45 En YANG et al., (Biochemistry, 39, 3533-3541, 2000) y en la solicitud PCT documento WO 01/62974 se describe una purificación de CEL I mejorada mediante el uso de alfametil-manosido en los tampones de purificación para superar la agregación de CEL I con lecitinas endógenas. Estos documentos también describen la clonación del ADNc de CEL I. Basándose en datos de secuencias, CEL I se asignó a una subfamilia de la familia de nucleasas S1/P1 y se identificaron diversos posibles homólogos codificados por los genes BFN 1 de *Arabidopsis* (nucleótido de GenBank
- 50 AY040016), ZEN 1 de *Zinnia* (nucleótido de GenBank AB003131), y DSA6 de lirio (nucleótido GenBank AF 082031).
- Se ha demostrado que la actividad endonucleasa de CEL I es altamente específica para emparejamientos erróneos por sustitución de bases y para emparejamientos erróneos resultantes de sucesos de inserción/delección y que es independiente del contexto de la secuencia flanqueante. Por tanto esta es útil como un reactivo de detección de
- 55 mutaciones en diversos métodos que implican selección mutacional. El sistema de detección de emparejamiento erróneo de CEL I es un ensayo sencillo que requiere amplificación de la secuencia diana por PCR, desnaturalización e hibridación para permitir la formación de heteroduplex entre el alelo de tipo silvestre y el alelo mutante, escisión enzimática del emparejamiento erróneo y análisis del producto por electroforesis en gel. Esto es ventajoso sobre otros sistemas de detección de emparejamiento erróneo conocidos, como HPLC desnaturalizante, debido a su
- 60 especificidad y sensibilidad para la detección de emparejamientos erróneos en grandes tramos de ADN.
- A modo de ejemplo, OLEYKOWSKI et al y YANG et al (en publicaciones citadas anteriormente) indican su uso para detectar modificaciones de secuencias en el gen *BRCA 1* humano y SOKUENKO et al (Nucl. Acids Res., 29 e111, 2001) describen su uso para detectar mutaciones y polimorfismos en grandes regiones de ADN genómico. CEL I también se usa para selección de alto rendimiento en TILLING (Selección de lesiones inducidas localmente en

genomas), en la que después de la mutagénesis química se realiza la selección para mutaciones puntuales, o para la detección de polimorfismos en poblaciones naturales, denominado también "Ecotilling" (COMAI et al, Plant Journal, 37, 778-786, 2004). Esto se ha usado en plantas (COLBERT et al., Plant Physiology 126, 480-484, 2001; TILL et al., Genome Research 13, 524-530, 2003; PERRY et al., Plant Physiology 131, 866-871, 2003) así como en animales tales como el pez cebra (WIENHOLDS et al., Genome res., 13, 2700-2707, 2003). Las solicitud PCT WO 03/066809 propone el uso de CEL I en un método para reclasificar variaciones de secuencia entre polinucleótidos relacionados denominado "Reclasificación Genética por Resolución de Emparejamiento erróneo" (GRAMMR).

Sin embargo CEL I tiene la desventaja de tener una eficacia de escisión que varía de un emparejamiento erróneo a otro: en el caso de un bucle de ADN con una sola inserción de nucleótidos, OLEYKOWSKI et al. (Nucleic Acids Res. 1998 Oct 15; 26 (20): 4567-602) indican que la preferencia de sustrato de CEL I es G>A>C>T; en el caso de emparejamientos erróneos por sustituciones de bases la preferencia de sustrato de CEL I es C/C ≥ C/A ~C/T ≥ G/G >A/C ~A/A ~T/C > T/G ~G/T ~G/A ~A/G > T/T. Esta eficacia es significativa en los emparejamientos erróneos C/A, C/C, C/T, G/G. Se observa una disminución de la actividad para los otros y es casi ineficaz en el emparejamiento erróneo T/T. Esta variación en la eficacia de escisión puede dar como resultado una menor precisión en la detección de algunas mutaciones cuando se necesita la detección de un alelo en un grupo de ADN.

Otro inconveniente limitante del uso de CEL I es el bajo rendimiento de los métodos de purificación disponibles. OLEYKOWSKI et al., comenzando con 7 Kg. de tallo de apio que contenía aproximadamente 350 g de proteína obtuvieron 3 ml de CEL I a 0,1 µg/µl; el procedimiento de purificación descrito por YANG et al. y en el documento PCT WO 01/62974 se obtuvieron 5µg de CEL I purificada con una actividad específica de 3,1 x 10<sup>7</sup> de proteína CEL I unidades/mg comenzando a partir de 10<sup>5</sup> Kg. de tallo de apio.

Para obtener grandes cantidades de CEL I, se ha propuesto producirla por tecnología de ADN recombinante. La solicitud PCT WO 03/066809 propone una amplia lista de vectores y células hospedadoras posiblemente adecuados que incluyen casi cualquiera de los sistemas de expresión procariotas o eucariotas conocidos; sin embargo, el único sistema de expresión realmente descrito en este documento es un vector basado en tobamovirus. La construcción resultante de la clonación en dicho vector del ADNc de CEL I fusionado a una secuencia que codifica una etiqueta de 6-Histidinas se ha usado para infectar plantas de tabaco. CEL I recombinante se recuperó del líquido intracelular de las plantas infectadas y se purificó por cromatografía de afinidad de metales sobre resina de NTA-níquel. El documento PCT WO 03/066809 no dice nada acerca del rendimiento de la enzima purificada. Esto indica que su actividad en una reacción GRAMMR es similar a la de la enzima natural purificada a partir del apio. Uno de los inconvenientes de este sistema es que los virus tienden a recombinar, perdiendo parcial o completamente el gen expresado. Esto conduce a la producción de formas truncadas en la adición de CEL I de longitud completa, disminuyendo la actividad específica de la enzima.

La solicitud PCT WO2004/035771 se refiere a un método para la producción de CEL I en levaduras. Para este efecto, se construyó un gen sintético que codifica CEL I modificando la secuencia de ADN natural de CEL I de acuerdo con el uso de codón en levadura. Este documento indica que la CEL I recombinante producida por este gen sintético puede reconocer todas las combinaciones de emparejamiento erróneo posibles y explica con ejemplos el reconocimiento y la escisión del emparejamiento erróneo A/A. Por otro lado no dice nada acerca de la preferencia de emparejamiento erróneo de dicha CEL I recombinante.

De lo anterior parece que los métodos actualmente disponibles para la producción de CEL I recombinante no proporcionan ninguna mejora significativa sobre la producción de CEL I natural procedente del apio. Además no tratan el problema de la preferencia de sustrato del CEL I natural.

Además, sería deseable tener otras endonucleasas que pudieran, como CEL I, escindir emparejamientos erróneos de pares de bases únicos en moldes de ADN heteroduplex a pH neutro, pero que tuvieran una preferencia de emparejamiento erróneo diferente o preferiblemente que escindieran igual de bien todos los emparejamientos erróneos. Sin embargo, hasta ahora dichas endonucleasas no se han identificado.

La existencia de actividades endonucleasas de tipo CEL I se ha descrito en muchas plantas (consúltese, por ejemplo OLEYKOWSKI et al., 1998 indicado anteriormente). Sin embargo, las enzimas responsables de estas actividades no se han caracterizado, sus propiedades químicas, tales como preferencia de sustrato, no se han estudiado y sus secuencias no se han identificado. Por otro lado, se han identificado por ordenador homólogos estructurales de CEL I (YANG et al. citado anteriormente; TILL et al. Nucleic Acids Res., 32, 2632-41, 2004). Se han descrito tres de estos (BFN1 de *Arabidopsis*, ZEN1 de *Zinnia* y DSA6 de lirio) como implicados en la senectud de las plantas (PEREZ-AMADOR et al., Plant Physiol. 122, 169-180 2000). Sin embargo, no se han purificado y permanecen sin caracterizarse bioquímicamente: en particular, no se ha ensayado su eficacia para reconocer emparejamientos erróneos en ADN heteroduplex *in vitro*. Procedente de la espinaca, se ha purificado una endonucleasa denominada SP que, basándose en su actividad, se ha asignado a la familia S1/P1 y que tiene un pH óptimo neutro. Al igual que CEL I, esta enzima escinde emparejamientos erróneos por inserciones/deleciones y por sustituciones de bases; sin embargo no reconoce a aquellos que contienen un resto de guanina (OLEYKOWSKI et al. Biochemistry, 38, 2200-5, 1999).

Buscando métodos alternativos para obtener CEL I recombinante, los autores de la invención han tratado de expresarla en plantas mediante expresión transitoria mediada por *Agrobacterium*.

- 5 Estos han expresado la CEL I recombinante en hojas de tabaco agroinfiltradas y la han purificado de los extractos de hojas por precipitación con sulfato de amonio. Sorprendentemente, han descubierto que, obtuvieron un rendimiento muy alto de CEL I recombinante con una alta especificidad y además, que dicha preparación de CEL I recombinante reconoce los emparejamientos erróneos con una especificidad más amplia y una mayor sensibilidad que en las preparaciones de CEL I conocidas en la técnica anterior, permitiendo una clara detección incluso de  
10 emparejamientos erróneos, tales como T/T, que se consideran mal reconocidos por CEL I.

A la vista de estos resultados, los autores de la invención han tenido la idea de usar este método para explorar endonucleasas, identificadas por ordenador como pertenecientes a la familia S1/P1, ensayando su actividad *in vitro*.

- 15 Por lo tanto, la presente invención proporciona un método sencillo y rápido para obtener grandes cantidades de endonucleasas, en particular nucleasas S1/P1, a partir de una pequeña cantidad de material de partida.

En este documento, una endonucleasa específica de emparejamiento erróneo, se define como una endonucleasa que es capaz de escindir específicamente todos los emparejamiento erróneos por sustitución de bases (es decir A/A,  
20 G/G, C/C, T/T, A/G, A/C, G/T, C/T, G/A, C/A, T/C, T/G), así como por inserciones/deleciones de uno o más nucleótidos.

Un objeto de la presente invención es por tanto un método para producir una endonucleasa recombinante específica de emparejamiento erróneo, en las que dicho método comprende:

- 25 - expresar dicha endonucleasa recombinante en células de una planta hospedadora completa o de un organismo separado de la misma transformado transitoriamente por agroinfiltración con una cepa de *Agrobacterium* que contiene un vector de expresión que comprende un polinucleótido que codifica dicha endonucleasa;  
- aislar dicha endonucleasa recombinante de dichas células de la planta hospedadora.
- 30 Dichas células de plantas pueden formar parte de una suspensión de células o de un cultivo tisular u orgánico. En este caso, la enzima puede recogerse del sobrenadante y/o de las células o tejidos u órganos cultivados. Preferiblemente, formarán parte de una planta completa o de un órgano separado de la misma; en este caso, la transformación transitoria con la cepa *Agrobacterium* se realizará por agroinfiltración.
- 35 La agroinfiltración es un método de expresión transitoria basado en el suministro de *Agrobacteria*, que contiene un gen de interés, dentro del tejido de la planta intacta.

Una construcción de ADN que comprende un gen de interés se clona en un vector binario y se transfiere a una cepa de *Agrobacterium* seleccionada y la *Agrobacteria* transformada se cultiva hasta una fase logarítmica o hasta  
40 saturación y se recoge de la misma manera que para una transformación mediada por *Agrobacterium* convencional. Clásicamente, la agroinfiltración se realiza aplicando una suspensión de las células de *Agrobacterium* transformadas por inyección en un órgano (generalmente las hojas) de la planta seleccionada usando una jeringa sin aguja o por infiltración al vacío durante escasos minutos. Después de la liberación del vacío, el órgano o toda la planta se coloca en una cámara de cultivo. La proteína de interés expresada se extrae del órgano infiltrado, normalmente de uno a  
45 cuatro días después de la infiltración. Los protocolos de agroinfiltración se describen en diversas publicaciones, por ejemplo KAPILA et al., (Plant Sci. 122, 101-108, 1997; BENDAHMANE et al., (Plant Cell, 11, 781-792, 1999); SCOFIELD et al., (Science., 274, 2063-5, 1996); TANG et al., (Science., 274, 2060-3, 1996); MARILLONET et al., (Proc Natl Acad Sci U S A, 101, 9852-7, 2004); WROBLEWSKI et al., (Plant Biotech. J., 3, pp. 259-273, 2005).

50 Los protocolos clásicos usados para la expresión transitoria mediada por *Agrobacterium*, y en particular para la agroinfiltración, pueden usarse en la realización práctica de la presente invención. Se dispone de una gran selección de cepas de *Agrobacterium*, de vectores binarios y de elementos reguladores que controlan la expresión del gen de interés y un experto en la materia puede seleccionar entre estas la más apropiada, por ejemplo de acuerdo con la planta hospedadora que se pretende usar.

55 En los experimentos descritos en los siguientes Ejemplos, los autores de la invención han usado un vector binario derivado de pBIN19, pBIN61, y una cepa de *Agrobacterium* C58C1 que aloja el plásmido pCH32 hipervirulento y el ADNc de las secuencias codificantes genómicas se ha expresado bajo el promotor del VMCa 35S. Sin embargo, pueden usarse otros vectores binarios, otras cepas de *Agrobacterium* y otros promotores constitutivos o inducibles  
60 con el mismo resultado.

Ventajosamente, la agroinfiltración se realizará en las hojas, que opcionalmente pueden separarse de dicha planta inmediatamente antes o inmediatamente después de la infiltración.

Las plantas hospedadoras que pueden usarse en el método de la invención incluyen cualquier planta que sea compatible con la transformación de *Agrobacterium*. Las plantas preferidas incluyen en particular las del género *Nicotiana*, en particular *Nicotiana benthamiana* y *Nicotiana tabacum*.

- 5 De acuerdo con una realización preferida de la invención, dicha endonucleasa se aísla de un órgano de planta agroinfiltrada, en particular una hoja agroinfiltrada mediante un proceso que comprende las siguientes etapas:
- extraer el contenido celular del órgano agroinfiltrado que expresa dicha endonucleasa;
  - añadir a dicho extracto sulfato de amonio a una concentración final del 30% o más y separar el precipitado de proteína del sobrenadante;
- 10 - añadir a dicho sobrenadante sulfato de amonio a una concentración final del 80% o más y recuperar el precipitado de proteína que contiene la endonucleasa.

Dicho precipitado de proteína se resuspende en un tampón apropiado, por ejemplo, un tampón que comprende Tris HCl (pH 8), PMSF y glicerol al 10%. Este puede usarse directamente o conservarse a -80°C hasta su uso.

- 15 Como alternativa, el extracto total obtenido después de la primera etapa indicada anteriormente puede usarse como tal, sin realizar etapas de precipitación con sulfato de amonio.

Opcionalmente, las endonucleasas producidas por el método de la invención pueden purificarse adicionalmente mediante cualquier método apropiado conocido en mismo, tal como purificación por afinidad en columna en la que CEL I se etiqueta con una etiqueta que tiene una afinidad hacia un componente específico en la columna. De acuerdo con una realización preferida, CEL I de la invención se proporciona con una etiqueta de 6-Histidina y se purifica por cromatografía de afinidad de metales sobre níquel-NTA.

25 Un método para ensayar si una endonucleasa candidata es una endonucleasa específica de emparejamiento erróneo comprende:

- a) producir dicha endonucleasa candidata en forma recombinante por el método de la invención, como se ha definido anteriormente;
- b) ensayar dicha endonucleasa recombinante para determinar su capacidad de degradar ADN monocatenario;
- c) ensayar dicha endonucleasa recombinante para determinar su capacidad de escindir un fragmento de ADN heteroduplex del ensayo en un sitio de emparejamiento erróneo previamente definido;
- d) ensayar dicha endonucleasa recombinante para determinar su capacidad de escindir fragmentos de ADN heteroduplex portadores de todos los tipos de emparejamientos erróneos (es decir, emparejamientos erróneos por sustitución de bases (A/A, G/G, C/C, T/T, A/G, A/C, G/T, C/T, G/A, C/A, T/C, T/G, así como emparejamientos erróneos por inserción o delección).

Si una endonucleasa supera los ensayos de las etapas b), c) y d) (es decir, si puede degradar ADN monocatenario, para escindir un fragmento de ADN heteroduplex del ensayo en un sitio de emparejamiento erróneo previamente definido y de escindir fragmentos de ADN heteroduplex portadores de todos los tipos de emparejamientos erróneos) se considera como una endonucleasa específica de emparejamientos erróneos.

Un procedimiento para seleccionar endonucleasas específicas de emparejamientos erróneos comprende:

- a) producir endonucleasas candidatas en forma recombinante mediante el método de la invención, como se ha definido anteriormente;
- b) ensayar dichas endonucleasas recombinantes para determinar su capacidad de degradar ADN monocatenario;
- c) ensayar las endonucleasas recombinantes capaces de degradar ADN monocatenario y ensayarlas para determinar su capacidad de escindir un fragmento de ADN heteroduplex del ensayo en un sitio de emparejamiento erróneo conocido y bien caracterizado;
- d) seleccionar las endonucleasas recombinantes capaces de escindir un fragmento de ADN heteroduplex del ensayo en un sitio de emparejamiento erróneo conocido y bien caracterizado, y ensayarlas para determinar su capacidad de escindir fragmentos de ADN heteroduplex portadores de todos los tipos de emparejamiento erróneo.
- e) seleccionar las endonucleasas recombinantes que superan los ensayos de las etapas b), c) y d).

Los métodos definidos anteriormente pueden comprender adicionalmente una etapa que consiste en ensayar dicha endonucleasa (o endonucleasas) para determinar su sensibilidad ensayando su capacidad para detectar un alelo mutante en un grupo de ADN, en presencia del alelo de tipo silvestre en exceso, y seleccionar las endonucleasas que son capaces de detectar dicho alelo mutante en presencia de del alelo de tipo silvestre en exceso al menos 9 veces (es decir, un alelo mutante en un grupo de 10), preferiblemente en presencia del alelo de tipo silvestre en exceso al menos 14 veces, aun más preferiblemente en presencia de al menos un exceso de 19 veces y para aumentar la preferencia, un exceso de 29 veces, 39 veces, 49 veces, 59 veces .

Preferiblemente, los ensayos definidos anteriormente se realizan en una mezcla de reacción que tiene un pH de 7 a 8, ventajosamente de 7,4 a 7,8 y que contiene de 5 a 20 mM, ventajosamente MgCl<sub>2</sub> 10 mM. Ventajosamente, dicha mezcla de reacción también comprende de 0,5 mM a 2 mM y preferiblemente 1 mM de DTT. Los autores de la invención también han observado que la adición de PEG-8000 del 2% al 10% (p/v) y en particular al 5% de la mezcla de reacción final, aumenta la actividad global de las endonucleasas.

Las endonucleasas candidatas que pueden ensayarse mediante los procedimientos definidos anteriormente pueden encontrarse entre las de la familia S1/P1. En la base de datos PFAM (BATEMAN et al., Nucleic Acids Res. 32, D138-41, 2004) esta familia se denomina PFAM 02265. Se puede, por ejemplo, usar el perfil HMM (Modelo oculto de Márkov) construido a partir de PFAM 02265 como una sonda para explorar los bancos de datos de secuencias de ADN disponibles, usando el programa informático HMMER a identificar endonucleasas candidatas en diferentes plantas.

Como alternativa, el código InterPro IPR003154 (correspondiente a las endonucleasas S1/P1) puede usarse para la exploración del contenido de bases de datos, por ejemplo usando la siguiente dirección:

<http://www.ebi.ac.uk/interpro/ISpy?ipr=IPR003154>

La exploración de las bases de datos Trembl/Swissprot con este código identifica 43 proteínas (Tabla 1).

Tabla 1: resultado de la exploración de la base de datos Trembl/Swissprot con el código IPR003154.

Número de acceso Trembl o Swissprot	Descripción	Longitud de secuencia	Bit (Blastp)
P24504	Nucleasa PA3 (EC 3.1.3.6)	270	127
P24289	Nucleasa P1( EC 3.1.30.1)	270	127
P24021	Nucleasa S1 (EC 3.1.30.1)	267	107
Q00235	Precursor de nucleasa S1.	287	109
Q9P356	Nucleasa Le1.	310	111
Q7S8Q5	Proteína hipotética.	306	114
Q87OT1	Nucleasa probable S1.	324	119
Q8NIH8	Nucleasa Le3.	298	129
Q25267	3'-nucleotidasa/nucleasa.	477	
Q9GNZ4	Precursor de 3'-nucleotidasa/nucleasa.	378	
Q9NJ13	3'-nucleotidasa/nucleasa.	377	
Q9NJY3	Nucleasa específica monocatenaria.	315	
Q86GJ3	Nucleasa P4.	316	
Q8T4M4	Nucleasa de clase I.	316	
065424	Supuesta nucleasa bifuncional	362	217
065425	Supuesta nucleasa bifuncional	454	164
080326	Precursor de endonucleasa.	303	480
081656	Proteína 6 asociada a senectud.	298	461
081958	Endonucleasa.	288	289
Q93WW9	Endonucleasa de tipo Si (Fragmento).	1361	239
Q9ARD4	Supuesta nucleasa.	289	
Q9C9G4	Supuesta nucleasa bifuncional	290	300
Q9FTRO	Supuesta nucleasa bifuncional.	310	
Q9FTR1	Supuesta nucleasa bifuncional.	311	180
Q9LGA5	EST D48949(5i554i).	3081	293

Q9LL59	Endonucleasa de emparejamiento erróneo CEL I.	296	627
Q9SXG1	Nucleasa I.	290	291
Q9ZR87	Nucleasa bifuncional.	328	308
Q9ZR88	Nucleasa bifuncional (Fragmento).	280	305
Q9ZR89	Nucleasa bifuncional bfnl.	305	445
Q7XND5	Proteína OSJNBbOO34I13.4.	252	277
Q7XPN4	Proteína OSJNBaOO6ODO6.10.	290	275
Q8LA68	Endonucleasa, supuesta.	296	281
Q8LCL6	Supuesta nucleasa bifuncional	290	299
Q8LDW6	Supuesta nucleasa bifuncional	294	274
068530	Homol. de S1 de endonucleasa.	309	
Q8XRE8	Supuesta proteína peptídica de señal.	337	
Q989R8	Endonucleasa.	278	
Q7P202	Probable endonucleasa.	274	
Q8F378	Nucleasa Si (EC 3.1.30.1).	306	
Q8P5Y5	Endonucleasa.	270	
Q8PHA3	Endonucleasa.	271	
Q9SXA6	Nucleasa bifuncional bfnl.	305	448

Este análisis se completa preferiblemente realizando un Blast en las bases de datos (blastp o tblastn), usando una secuencia de proteínas de referencia (por ejemplo la secuencia de CEL I) y seleccionando los mejores aciertos.

5

Los siguientes ejemplos describirán con más detalle los ensayos de cinco endonucleasas candidatas de *Arabidopsis thaliana*.

Los autores de la presente invención descubrieron que una de estas endonucleasas, representada en la lista de secuencias adjunta bajo la SEQ ID N° 2 (la secuencia de ADN correspondiente se representa bajo la SEQ ID N° 1) y que corresponde a BFN1 de *Arabidopsis thaliana*, tiene una especificidad diferente y una sensibilidad mucho mayor que CEL I.

Este descubrimiento de las propiedades de BFN1 como una endonucleasa específica de emparejamiento erróneo, que hasta ahora se desconocía, permite proponer su uso como un reactivo de detección de mutaciones, para detectar emparejamientos erróneos resultantes de sustituciones de bases, así como de inserciones/delecciones de uno más nucleótidos.

La BFN1, así como cualquiera de las endonucleasas específicas de emparejamiento erróneo que pueden identificarse de acuerdo con el método de la invención puede obtenerse en grandes cantidades por el método de producción de endonucleasas.

La presente invención también incluye preparaciones de endonucleasas recombinantes que pueden obtenerse mediante el método de producción de acuerdo con la invención. Estas son en particular preparaciones de CEL I recombinante y preparaciones de BFN 1 recombinante.

Las preparaciones de CEL I recombinante de la invención tienen una especificidad de emparejamiento erróneo diferente y una mayor sensibilidad que las preparaciones de CEL I de la técnica anterior. Las preparaciones de CEL I recombinante de la invención, tienen la siguiente preferencia de emparejamiento erróneo: T/G ~ A/G ~ G/G ~ G/T ~ T/T ~ G/A ≥ A/A ~ C/C ≥ T/C ~ C/T > A/C ~ C/A, mientras que la preferencia de emparejamiento erróneo de las preparaciones de CEL I de la técnica anterior es C/C ≥ C/A ~ C/T ≥ G/G > A/C ~ A/A ~ T/C > T/G ~ G/T ~ G/A ~ A/G > T/T. Las preparaciones de CEL I recombinante de la invención pueden reconocer un alelo mutante en presencia de

23 veces más del alelo de tipo silvestre mientras que las preparaciones de CEL I de la técnica anterior no reconocen eficazmente ningún alelo mutante cuando se diluye 8 veces en una dilución.

Las preparaciones de BFN 1 recombinante de la invención también tienen una especificidad de emparejamiento erróneo diferente a la de las preparaciones de CEL I de la técnica anterior y las preparaciones de CEL I recombinante de la invención. Las preparaciones de BFN 1 recombinante de la invención tienen las siguientes preferencias de emparejamiento erróneo: G/G ~ G/A ~ A/G ~ G/T ~ T/G > T/T ~ A/A ~ C/C ~ T/C > C/T ~ A/C ~ C/A. En la siguiente tabla 2, se resumen las preferencias de emparejamiento erróneo para cada una de estas enzimas.

10 Tabla 2: Comparación de preferencias de reconocimiento de emparejamiento erróneo.

Proteína	Reconocido más eficazmente		Bien reconocido			Reconocido con menor eficacia			Muy débilmente reconocido
CEL I de apio (técnica anterior)	C/C C/A	C/T	G/G	A/C		A/A T/C	G/A T/G	G/T A/G	T/T
CEL I recombinante de acuerdo con la invención *	T/G G/G	A/G G/T	T/T C/C	G/A	A/A	C/T C/A	A/C	T/C	
BFN 1 (ENDO 1)	G/G A/G T/G	G/A G/T	A/A C/C	T/T	T/C	C/T	A/C	C/A*	

\*= BFN1 (ENDO 1) reconoce estos emparejamientos erróneos de modo más eficaz que CEL I recombinante.  
 \*\* N.I.: el cálculo de la eficacia de escisión solo es semicuantitativo; por lo tanto, pueden producirse ligeras variaciones en la clasificación de los emparejamientos erróneos, basándose en su reconocimiento por las endonucleasas, dependiendo del investigador y/o del experimento.

Adicionalmente, las preparaciones de BFN 1 recombinante de la invención tienen una mayor sensibilidad tanto en las preparaciones de CEL I de la técnica anterior como en las preparaciones de CEL I recombinante de la invención. Las preparaciones de BFN 1 recombinante de la invención pueden reconocer un alelo mutante en presencia de 59 veces más el alelo de tipo silvestre.

Las preparaciones de endonucleasas BFN 1 o CEL I recombinantes de la invención pueden usarse, como un reactivo para la detección de mutación, en cualquier método que implique la exploración de emparejamientos erróneos como se ha mencionado anteriormente. Son particularmente ventajosas en genotipificación, en TILLING, TILLING de alto rendimiento, Ecotilling, GRAMMR, etc.

Los métodos para usar las preparaciones de endonucleasas de la invención implican básicamente las mismas etapas a las de la técnica anterior, es decir, amplificación por PCR de la secuencia diana, desnaturalización del producto de amplificación e hibridación para permitir la formación de heteroduplex entre el alelo de tipo silvestre y el mutante, escisión de los heteroduplex por la endonucleasa y análisis de los productos de escisión. Ventajosamente, cuando se realizan estos métodos, puede usarse una mezcla de endonucleasas diferentes para la escisión de los heteroduplex.

Debido a su elevada sensibilidad, las preparaciones de endonucleasas de la invención también permiten realizar métodos de alto rendimiento para identificar mutaciones en una muestra. Por ejemplo, las preparaciones de endonucleasas de acuerdo con la invención pueden usarse para realizar métodos tales como que se describen en el documento WO 01/75167, con un número de muestras mucho mayor dado que es posible agrupar muchas muestras juntas para el análisis.

De acuerdo con la invención, las preparaciones de endonucleasas descritas anteriormente pueden usarse para explorar una o más mutaciones en un gen diana, en una gran cantidad de muestras a partir de un organismo o una línea celular derivada del mismo, realizando las siguientes etapas:

- a) amplificar dicho gen diana o una parte del mismo de cada individuo de dicha población,
- b) clasificar dichos productos de amplificación en una matriz bi- o tri- dimensional, que comprende líneas, filas (matriz 2-D (bidimensional)) y columnas (matriz 3-d (tridimensional)),
- c) agrupar dichos productos de amplificación de manera que se obtengan diferentes grupos, presentado cada grupo una fila, una línea o una columna de dicha matriz,
- d) añadir a cada grupo un producto de amplificación de referencia obtenido a partir de un gen no mutado e incubar dichos grupos en condiciones que permitan la formación de heteroduplex,
- e) incubar cada grupo con una preparación de endonucleasa de acuerdo con la invención, y
- f) detectar la presencia de heteroduplex en dichos grupos incubados.

Como alternativa, en el método anterior puede realizarse clasificando primero las muestras en la matriz, después agrupándolas (con la adición de la referencia) y realizando la amplificación, incubación y etapas de detección.

Dependiendo del número de muestras a explorar, la matriz puede ser bi- o tri-dimensional. Por ejemplo, si se van a explorar 576 muestras, puede usarse una matriz 24 x 24. Si se van a explorar 13.824 muestras, debe emplearse más apropiadamente una matriz tridimensional (24x24x24). Solamente serían necesarias 72 reacciones para explorar esta población y los genes mutados se individualizarían por los grupos en la columna, fila y línea a los que pertenece.

El producto de amplificación de referencia corresponde a un producto de amplificación obtenido a partir de un gen de referencia (en comparación con las mutaciones que se exploran), amplificado con los mismos cebadores que los del gen diana en la población. Es importante añadir un producto de amplificación de referencia. De hecho, aunque es bastante improbable que todas las muestras contengan la misma mutación exacta en comparación con el gen de referencia, esto asegurará que los heteroduplex puedan formarse si el grupo contiene un gen diana que contiene una mutación.

Además, la siguiente descripción adicional ilustrará la presente invención, que se refiere a ejemplos que ilustran la preparación y propiedades de la endonucleasa CEL-I recombinante, el ensayo de cinco endonucleasas candidatas de *Arabidopsis thaliana*, y la identificación de BFN 1, (en lo sucesivo en este documento denominada también ENDO 1) como una endonucleasa específica de emparejamiento erróneo. Debe entenderse sin embargo que estos ejemplos se proporcionan únicamente para ilustrar la invención y no constituyen de ninguna manera ninguna limitación de la misma.

### Leyendas de las figuras

**Figura 1:** detección de mutación puntual en gel de agarosa. Se habían incubado heteroduplex (de ADN mutante y de tipo silvestre) con diferentes diluciones de CEL I recombinante (D104 a D1000), o sin proteína (-).

**Figura 2:** detección de mutación puntual en gel de acrilamida. WT+Mut: antes de la PCR se habían mezclado el ADN de tipo silvestre y el mutante, generando por lo tanto heteroduplex. WT: para la PCR solo se ha usado ADN de tipo silvestre, generando por lo tanto solamente homoduplex. Mut: para la PCR solo se ha usado ADN mutante, generando por lo tanto solamente homoduplex. D100, D500 y D1000: diluciones de la proteína CEL I recombinante.

En la parte superior del gel se indica el tamaño de los homoduplex (661 pb). La fila de 405 pb muestra el fragmento marcado con el fluorocromo FAM. La fila de 256 pb muestra el fragmento marcado con el fluorocromo ROX.

**Figura 3:** detección de mutación puntual en ADN genómico en gel de agarosa. Los productos de la PCR se obtuvieron como se describe más adelante en el ejemplo 4, con los cebadores 4-960 y 4-721 (SEC ID Nos: 3 y 4). Después se digirieron con diferentes diluciones de preparación de CEL I recombinante obtenida por precipitación con sulfato de amonio (como se describe en el ejemplo 2) y se analizó en gel de agarosa. El tamaño de los productos de la PCR para los alelos de tipo silvestre y mutante de 1-13 de rms, es de aproximadamente 500 pb (exactamente 481 pb). Como resultado de la detección de la mutación en el ADN heteroduplex, se obtuvieron dos bandas de aproximadamente 200 y 300 pb. Las dos bandas pueden observarse en un gel de agarosa incluso a una dilución de 1/1000 de la proteína producida en tabaco.

1: D100; 2: D500; 3: D1000; 4: sin enzima.

A: Terese; B: rms 1,13; C: T+rms 1,13

**Figura 4:** análisis de actividad de Cel I recombinante en todos los tipos de emparejamientos erróneos. Se creó, mediante PCR, una serie de mutantes en base al gen Rx y se obtuvieron heteroduplex mezclando los productos de amplificación correspondientes con estos mutantes diferentes. Para la PCR, se usaron oligonucleótidos marcados con fluoroforo IRD700 e IRD800 (MWS®) y se realizó la detección de emparejamiento erróneo en LICOR4300 (LICOR®). La figura muestra los canales (A) con IRD colorante 700 y (B) con IRD colorante 800 de un proceso. El canal A muestra el fragmento de 204 pb con el fragmento del extremo 5' marcado con colorante IRD-700 y el canal B muestra el fragmento de 438 pb con el fragmento del extremo 3' con el colorante IRD-800.

**Figura 5:** sensibilidad de la proteína Cel I recombinante producida en tabaco. 1: Guisante Wt (de tipo silvestre); 2: Le1; Wt+ Le1; 4 Wt+ Le1+2ADN; 5 Wt+ Le1+4ADN; 6Wt+ Le1+6ADN; 7 Wt+ Le1+8ADN; 8 Wt+ Le1+10ADN.

**Figura 6:** escisión de heteroduplex de ADN en el sitio de emparejamiento erróneo C-A/T-G por cinco endonucleasas candidatas. (Ho)= homoduplex de ADN; (Ht)= heteroduplex de ADN.

**Figura 7:** análisis detallado del reconocimiento específico de todos los tipos de emparejamiento erróneo con ENDO1 y ENDO5. Se usó el mismo protocolo que el de la figura 4.

**Figura 8:** ensayo para medir la sensibilidad de detección de ENDO1. Como molde se usó una mezcla de ADN mutante (Le1) y de tipo silvestre (Torstag) para determinar la actividad de ENDO1, en las siguientes proporciones: 1 mutante por 2, 4, 6, 8, 10, 15, 20, 25, 30, 35, 40, 45 y 60 de tipo silvestre (de izquierda a derecha). Las dos últimas líneas de cada panel son homoduplex (solamente mutante y solamente de tipo silvestre). Panel izquierdo: canal IRD700, tamaño del fragmento detectado=338 pb. Panel derecho: canal IRD800, tamaño del fragmento detectado=300 pb.

**Figura 9:** comparación de detección de emparejamiento erróneo por Cell y ENDO1. Las líneas 1, 5, 8 y 10 corresponden a homoduplex. Los fragmentos resultantes de la escisión por la endonucleasa son de 405 pb (marcados en azul) y de 256 pb (marcados en rojo, menos visible en blanco y negro). ENDO1 reconoce emparejamientos erróneos más eficazmente que Cell. Además, el fondo (actividad no específica) es mucho menor con ENDO1.

**Figura 10:** detección de una mutación puntual conocida en gel de acrilamida con ENDO1 en dos diluciones diferentes (D1000 y D5000). M5 y M12 son dos plásmidos que contienen el gen Rx; uno contiene la forma Wt (de



Se recogieron hojas de tabaco agroinfiltradas y se pesaron. Se homogeneizaron dos gramos de hojas agroinfiltradas en 7 ml de tampón que contenía Tris-HCl 0,1M, pH 8, PMSF 200µM, βmercaptoetanol 0,125mM y 10% de glicerol y después se centrifugó a 3000 g durante 25 minutos para sedimentar los residuos celulares. Al sobrenadante se le añadió sulfato de amonio al 100% hasta la concentración final del 30% y después las muestras se incubaron en hielo durante 1 hora y se centrifugaron a 3000g durante 30 minutos a 4° C para sedimentar las proteínas que precipitan al 30% de sulfato de amonio. Al sobrenadante se le añadió sulfato de amonio al 100% para obtener una concentración final del 80% y las muestras se incubaron de nuevo en hielo durante 1 hora y se centrifugaron como se ha indicado anteriormente para sedimentar las proteínas que precipitan al 80% con sulfato de amonio. El sobrenadante que contiene las proteínas precipitadas al 80% de sulfato de amonio se resuspendió en 600µl de tampón de homogeneización. La concentración de proteínas se determinó por medio del kit de ensayo Reactivo Coomassie Plus (Pierce). El sedimento homogeneizado contenía 14µg de proteína µl. Por tanto, se recogieron 8400µg de proteínas a partir de 2 gramos de hojas de tabaco agroinfiltradas. El sedimento homogeneizado se diluyó a 1µg/µl en un tampón que contenía Tris HCl 50mM pH 8, glicerol al 10% y PMSF 100µM, se dividió en alícuotas y se almacenó a -80° C.

#### Purificación de la proteína marcada con His-6 por cromatografía de afinidad con Ni-NTA:

Se recogieron 5 gramos de hojas de tabaco agroinfiltradas y se homogeneizaron en 15ml de tampón helado que contenía Na-fosfato (100mM), Tris HCl pH 8 (10mM), NaCl (200mM), metabisulfito de sodio (0,2%), PMSF (1mM), βMercaptoetanol (10mM). Después de la homogeneización, la muestra se centrifugó a 6000 g (rotor Beckman, JA 20) durante 10 minutos a 4° C. Se añadió imidazol (10mM) al sobrenadante y con NaOH se ajustó el pH a 9. La solución se centrifugó a 42.000 g (rotor Beckman, JA 20) durante 60 minutos a 4° C. El sobrenadante se mezcló con 1 ml de agarosa en Ni-NTA (Quiagen) se pre-equilibró con tampón de homogeneización + imidazol 10mM pH9 (tampón B). La mezcla se homogeneizó durante 2 horas a 4° C para permitir que la proteína se uniese a perlas de agarosa en Ni-NTA. Las perlas se compactaron en una columna de polipropileno de 1ml (Quiagen) y la resina se lavó con 20 ml del tampón B. La proteína se eluyó con 5 ml (5x1 ml) de tampón B + imidazol 250 mM, pH 9. Las alícuotas de las fracciones se guardaron para permitir la actividad de la enzima durante la purificación. Para impedir cualquier inhibición de la actividad enzimática por la elevada concentración de imidazol, las fracciones eluidas se dializaron frente a 4 litros de tampón que contenía Tris-HCl pH 8, 0,1M, PMSF 100µM y ZnCl<sub>2</sub> 2µM durante una noche a 4° C. Por lo tanto, se recuperaron 1000 µg de proteínas. El sedimento homogeneizado se diluyó a 3µg/µl en un tampón que contenía Tris-HCl 50 mM, pH 8, glicerol al 10% y PMSF 100µM, se dividió en alícuotas y se conservó a -80° C. (dilución D 10.000).

35

#### **EJEMPLO 2: DEGRADACIÓN DE ADN ESPECÍFICO MONOCATENARIO**

La actividad de la CEL I recombinante sobre la degradación de ADN monocatenario se realizó como se ha descrito anteriormente (SUNG SC, LAKKOWSKI M Sr.(1962) J Biol Chem. Febrero 1962; 237: 506-11). Se incubaron 30 microgramos de Dnasa, Rnasa y proteasa sin ADN genómico de guisante con 2µg de extracto de proteína en un tampón que contenía Tris-HCl 50mM (pH 7,6), MgCl<sub>2</sub> 10mM, DTT 1mM y PEG-8000 al 5%. Para detener la reacción se añadió un volumen igual de LaCl 3 20 mM en HCl 0,2N. Las muestras se centrifugaron a 21.000 g durante 40 minutos y se midió la absorbancia a 260 nm del sobrenadante usando un espectrofotómetro para determinar la cantidad de ADN que se había vuelto soluble en ácido.

#### **EJEMPLO 3: ESCISIÓN DE ADN HETERODUPLEX Y DETECCIÓN DE UNA MUTACIÓN PUNTUAL CONOCIDA EN UN GEN DE ENSAYO EN AGAROSA Y GELES DE ACRILAMIDA POR LA ENDONUCLEASA PRODUCIDA EN TABACO**

50

Para ensayar si la endonucleasa CEL I producida en tabaco puede reconocer mutaciones puntuales sencillas, se ensayó la actividad de la preparación de CEL I recombinante obtenida por precipitación con sulfato de amonio en ADN heteroduplex a partir de dos clones que difieren en una mutación puntual sencilla: cambio de C-G a A-T (Bendhamane et al., Plant Cell, 11, 781-792, 1999). La PCR se realizó en los dos clones, usando dos oligonucleótidos R21 y R22 (R21 5' GAC ATA TGG ACT ACA GAA GCT TGG G 3' SEC ID N° 8; R22 5' GTT CAC GGG TCA CAT CAT GCA TTC C 3' SEC ID N° 9). La amplificación por PCR y la reconstitución de ADN heteroduplex se realizó usando el siguiente programa: desnaturalización durante 2 min. a 95 °C seguido de 7 ciclos con 20 s a 94C, Tf (55°C) +3 °C a Tf -4°C durante 5 s, -1 °C por ciclo, gradiente a 72 °C a 0.5 °C/s y una extensión a 72 °C durante 1 min, después 44 ciclos con 20 s a 94 °C, Tf -5 °C a 30 °C, gradiente a 72 °C a 0.5 °C/s, y una extensión a 72 °C durante 1 min, una extensión final a 72°C durante 5 min y una etapa de desnaturalización a 94 °C durante 10 min seguido de un aumento de 40°C durante 20 s y -0.3 °C por ciclo.

Los productos de la PCR (una mezcla de ADN de tipo silvestre y mutante o solo ADN de tipo silvestre o mutante) se incubaron con la preparación de CEL I (solución madre a 1µg/µl diluido a 1/100, 1/500 ó 1/1000) de la siguiente manera: por ejemplo, se incubaron 10 µl de producto PCR (500 ng) con 2,5 µl del tampón de reacción (Hepes

65

10mm, MgSO<sub>4</sub> 10mM, Triton X100 al 0,002%, KCl 10mm) y 2,5µl de la preparación de CEL I diluida en un volumen total de 25µl, durante 30 minutos a 37 °C. La reacción se detuvo por 5µl de EDTA 500mM y los productos de digestión se analizaron en un gel de agarosa al 3%.

5 Como se muestra en la Figura 1, la detección de la mutación puntual en el gen de ensayo, Rx, en ADN heteroduplex se reveló por la aparición de dos bandas de aproximadamente 200 pb y 400 pb a diluciones 1/100, 1/500 y 1/1000. Cuando no se añade enzima, estas bandas no aparecen.

10 Los productos de la PCR se obtuvieron usando cebadores marcados con fluorescencia y se digirieron del mismo modo que antes cuando se analizaron en gel de acrilamida sobre un secuenciador ABI377 (Figura 2).

15 Como se muestra en la Figura 2, las dos bandas de 256 y 405 pb aparecen solo cuando se han mezclado los ADN de tipo silvestre Wt y mutante Mut entre sí debido a la formación de heteroduplex entre los dos ADN. Esto aparece incluso más claro cuando se usan diferentes longitudes de onda para la visualización del gel. En particular, la banda de 256 pb, marcada con el fluorocromo ROX (rojo) y no muy claramente individualizada sobre el negro y blanco en la figura 2, está claramente presente cuando se usa la mezcla Wt + Mut como un molde para la PCR y ausente en los otros casos.

20 Estos resultados demuestran que la proteína CEL I producida en plantas puede reconocer emparejamientos erróneos en ADN.

**EJEMPLO 4: DETECCIÓN DE MUTACIONES PUNTUALES EN ADN GENÓMICO DE GUISANTE POR LA CEL I RECOMBINANTE**

25 Para ensayar si CEL I recombinante purificada de tabaco podría usarse para detectar mutaciones puntuales sencillas en guisante, se usaron diferentes mutantes *rms* y *le* de guisante, caracterizados previamente por Catherine Rameau (MORUS et al Plant Physiol 2001, 126 :1205-1213. ; RAMEAU et al, Plant Physiol 2002, 115 :458-467) como un ensayo se usaron; rms1,11 que contiene G---->A en diferentes posiciones en la secuencia del gen; rms 1-12 que contiene G---->A; rms 1-13 que contiene G----> A. Todas se mutaron a G---->A en el mismo gen rms1 pero en diferentes posiciones.

30 Para amplificar los alelos de tipo silvestre y mutante de los locus de *rms* (rms1-13, rms1-10, rms1-12) y del locus *le*, se usaron cincuenta nanogramos de ADN genómico de guisante como molde. Los cebadores usados en esta amplificación por PCR se resumen en la Tabla 4.

35

**TABLA 4:**

Nombre del mutante del guisante	Nombre del cebador	Secuencia 5'-> 3'
<i>rms</i> 1-10	4m118	5'TTGTTGGACTTCACTTTGAGC 3' SEC ID N°10
	4m984	5' CACAACAATCAGCAATGACAGC 3' SEC ID N°11
<i>rms</i> 1-12	4-347	5' GTGATTGCTCCACCTCCGCCACC 3' SEC ID N°12
	4-134	5'TACAGCGATTGATATAATATAAAAATTATCC3'SEC ID N°13
<i>rms</i> 1-13	4-960	5' GTGTTTGTCCAGTAATAGTGTGTCAGCATA 3' SEC ID N°3
	4-721	5' AGGAACCTGAGAAAAGACTCGCCAGC 3' SEC ID N°4
<i>le</i> 1	le 2462	5' TGATATTGTCGTGCAATATGATGAAAC 3' SEC ID N°14
	le 3082	5'ATACCTATTTAGCCCACTTGGACAC 3' SEC ID N°15

40 El programa de la amplificación de la PCR fue de 94 °C, 1 min, (94 °C 15s, 55 °C 15s, 74 °C 1min, X35) 74 °C 7min, 8 °C. Los productos de la PCR se analizaron en gel de agarosa y se digirieron, como se ha descrito en el Ejemplo 2 anterior, con diferentes diluciones de preparación de CEL I recombinante obtenida por precipitación con sulfato de amonio. Como se muestra en la Figura 3, el tamaño del producto de la PCR en el mutante *rms* 1-13 y el mutante de tipo silvestre, es de aproximadamente 500 pb (exactamente 481 pb). Como resultado de la detección de la mutación en ADN heteroduplex, se obtuvieron dos bandas de aproximadamente 200 y 300 pb. Las dos bandas pueden observarse en un gel de agarosa incluso a una dilución de 1/1000 de la proteína producida en tabaco. Este resultado demuestra que la proteína producida en plantas puede en primer lugar, reconocer una mutación puntual presente en ADN genómico y en segundo lugar, es muy activa ya que los productos de digestión pueden observarse incluso si la proteína se diluye a 1/1000.

**50 EJEMPLO 5: EFICACIA DE ESCISIÓN DE EMPAREJAMIENTO ERÓNEO EN DIFERENTES EMPAREJAMIENTOS ERRÓNEOS POR LA CEL I RECOMBINANTE**

55 Para ensayar si la CEL I recombinante preparada de acuerdo con la invención escinde preferencialmente un tipo emparejamiento erróneo, como la CEL I purificada de apio, se creó una serie de mutantes basados en el gen Rx. Para este fin, se diseñaron diferentes cebadores (denominados Rx-A, T, G o C) conteniendo cada uno una mutación

puntual diferente. Cada uno de estos cebadores permite introducir, a una posición determinada del gen Rx, una de las 4 bases (A, T, G o C). Se crearon heteroduplex mezclando los productos de amplificación obtenidos con los diferentes cebadores.

5 La mezcla de la PCR (en volumen total de 50 $\mu$ l) contenía el molde de ADN (50 ng), dNTP (0,2 mM), 5 $\mu$ l de tampón PCR Pfu (10X, Stratagene), cebador Rx 21 (0.4 $\mu$ M), cebador Rx-A o Rx-T o Rx-G o Rx-C (0.4 $\mu$ M), Pfu (5U, Stratagene, 2.5 Unidades/ $\mu$ l). El programa usado para la amplificación PCR fue 94 °C 1min, (94 °C 15s, 55 °C 15s, 74 °C 2min, X35) 74 °C 7 min, 8 °C durante una noche. Para la PCR se usaron oligonucleótidos marcados con fluoroforo IRD700 e IRD800 (MWS®) para permitir la detección de emparejamientos erróneos en LICOR4300  
10 (LICOR®).

Los productos de la PCR se analizaron en gel de agarosa y se clonaron en pGEM 3Zf. Todos los clones se habían secuenciado para garantizar que se había insertado la mutación correcta. Para reconstituir todos los tipos de emparejamiento erróneo, se usaron las combinaciones de estas construcciones como molde en la amplificación por  
15 PCR.

El producto PCR de emparejamiento erróneo incubó como se describe en el Ejemplo 2 anterior, con una preparación de CEL I recombinante obtenida de hojas de tabaco por precipitación con sulfato de amonio como se ha descrito en el Ejemplo 1 anterior.  
20

Los productos de las digestiones se analizaron en gel de acrilamida. Se usó un gel de acrilamida desnaturalizante al 6,5% y las condiciones de electroforesis fueron: 1500V, 40W, 40mA, 45 °C con una velocidad de barrido de 1.

Los resultados se muestran en la Figura 4 (Gel de Licor).  
25

Estos resultados demuestran que las preparaciones de CEL I de la invención reconocen todos los tipos de emparejamiento erróneo y particularmente emparejamientos erróneos indicados en la técnica anterior que se reconocían débilmente por CEL I purificada directamente del apio. Además, como se muestra en la Tabla 2 anterior, la preparación de CEL I de la invención reconoce emparejamientos erróneos, débilmente reconocidos por la  
30 preparación de CEL I de la técnica anterior, como T/T, G/A, A/G, G/T, T/G con especificidad muy elevada, siendo la preferencia de emparejamientos erróneos de esta enzima la siguiente: T/G ~ A/G ~ G/G ~ G/T > T/T ~ G/A ~ A/A ~ C/C > T/C ~ C/T ~ A/C ~ C/A.

Esta actividad específica también se potenció cuando el tampón de reacción de CEL I se complementó con PEG al 5% (datos no mostrados).  
35

#### EJEMPLO 6: SENSIBILIDAD DE LA CEL I RECOMBINANTE

Para verificar que la CEL I recombinante de la invención puede usarse para genotipado de alto rendimiento, se ensayó para ver si podía reconocer mutaciones en un grupo de individuos.  
40

Se usó ADN genómico con un SNP (polimorfismo de un único nucleótido) (C--> T), correspondiente al mutante del guisante enano Le1, y al ADN genómico de la variedad de cultivo Torstag del guisante de tipo silvestre para amplificar el locus Le1. Este ADN genómico se había usado como control para la formación de homoduplex y heteroduplex.  
45

Para crear ADN heteroduplex, ADN genómico derivado de la planta del guisante homocigoto para el locus *le1* se diluyó con ADN genómico derivado de planta del guisante homocigoto para el locus *le1* en diferentes proporciones y se usaron 30 ng para la amplificación por PCR usando los cebadores Le2462 marcados con el fluorocromo TET (5'-TGATATTGTCGTGCAATATGATGAAAC-  
50 3' SEC ID N° 14) y Le3082 marcado con el fluorocromo ROX (MWG®) (5'ATACCTATTTAGCCCCACTTGGACAC-3' SEC ID N° 15). Las reacciones de la PCR se realizaron de la siguiente manera: 94 °C 1min, (94 °C 15s, 55 °C 15s, 74 °C 1 min ) X35, 74 °C 7 min. En el ensayo de detección de emparejamientos erróneos, se usaron, como molde, los heteroduplex de ADN reconstituidos a partir de los productos de la PCR, como se ha escrito anteriormente.  
55

En la Figura 5 se presentan los resultados de este experimento en grupo usando la proteína producida en tabaco. Como se esperaba, la proteína recombinante de la invención puede reconocer en cualquier momento el PSN de *Le1* de un heteroduplex que se ha formado entre una cadena de tipo silvestre y una cadena de ADN mutante. Como resultado de la digestión, se obtuvieron dos bandas de tamaño esperado (300 pb y 338 pb) y también una tercera  
60 banda correspondiente al homoduplex no digerido (638pb). La banda de 338pb, que está marcada en rojo, aparece claramente sobre el gel (pero es menos visible en blanco y negro). De manera más interesante, la proteína puede detectar la mutación incluso cuando se usó un grupo de 25 ADN genómicos diferentes (adición de *Le1* y de tipo silvestre con 23 ADN genómicos desconocidos).

Para terminar, la proteína producida en la planta tiene una sensibilidad elevada, lo que permite la identificación de un PSN conocido con secuencias de ADN amplificadas procedentes de ADN genómico de al menos 24 individuos.

**EJEMPLO 7: INVESTIGACIÓN BIOINFORMÁTICA PARA DIFERENTES ENDONUCLEASAS EN ARABIDOPSIS.**

5 Se realizó un análisis de una familia de genes de *Arabidopsis* que codifican endonucleasas, usando el perfil de proteína PF02265 de la base de datos PFAM. Este perfil HMM se usó como una sonda para dirigir todas las 27117 proteínas predichas en el genoma de *Arabidopsis* y la totalidad del genoma se tradujo en las 6 desfasas del marco de lectura diferentes para evitar el efecto de la anotación automáticamente estructural. En este análisis se  
10 identificaron 5 genes candidatos At1g11190, At1g68290, At4g21590, At4g21585 y At4g21600.

**EJEMPLO 8: CLONACION Y EXPRESION DE GENES CANDIDATOS DE ARABIDOPSIS EN HOJAS DE TABACO**

15 El ADNc de cada gen candidato se amplificó por PCR y se insertó entre el promotor 35S y el terminador transcripcional del VMCa en el vector binario pBin61 para crear pBIN35S-ENDO1, -ENDO2, -ENDO3, -ENDO4 y -ENDO5 que corresponden a At1g11190, At1g68290, At4g21585, At4g21590 y At4g21600, respectivamente (Tabla 5). Estas construcciones se transformaron en la cepa C58C1 de *Agrobacterium* portadora del plásmido auxiliar de virulencia pCH32 (HAMILTON CM, et al. (1996) Proc Natl Acad Sci U S A., 93(18):9975-9). El plásmido pCH32  
20 expresa VirG y VirE y se usó para potenciar la transferencia de ADN-T. Las células de *Agrobacterium* se inocularon en 2 ml de medio con caldo L (SAMBROOK et al. 1989) complementado con 50µg/ml de kanamicina y 5µg/ml de tetraciclina y se cultivaron a 28°C durante una noche. Las células se precipitaron y se resuspendieron hasta una concentración final de DO<sub>600</sub> de 0,5 en una solución que contenía MgCl<sub>2</sub> 10mm, MES 10mM, pH 5,6 y 150µm de acetosiringona. Los cultivos se incubaron a temperatura ambiente durante 2 h. antes de la agroinfiltración en hojas  
25 de *Nicotiana benthamiana*. Las plantas de *Nicotiana benthamiana* agroinfiltradas se incubaron durante al menos 24 horas a 24°C, 16 horas de luz, 60% de humedad. Para ensayar la eficacia de la agroinfiltración, las plantas también se agroinfiltraron con una construcción que expresaba la proteína fluorescente verde (GFP). Antes de recoger las hojas, se midió la intensidad de la expresión de la GFP para cada hoja usando una lámpara UV. Solo se recogían las hojas de la planta que expresaban la GFP. Las proteínas candidatas se extrajeron por precipitación con sulfato de amonio como se describe en el Ejemplo 1.  
30

**Tabla 5: Secuencias de los cebadores usados para la clonación de la endonucleasa candidata con o sin etiqueta Histidina**

Proteínas candidatas	Genes AGI	Cebadores
ENDO5	<b>At4g21600</b>	Directo: AAGGATCCGAAAGCTCTGTGTTTCAGA <b>SEC ID Nº :16</b> Inverso: GGAGTTGTTACGTGGTTCTCAAGGATC <b>SEC ID Nº :17</b>
ENDO4	At4g21585	Directo:CTGGATCCCTGTTTTTAACITTTGGAAAG <b>SEC ID Nº :18</b> Inverso: GGATGTTCAAGTATTCTCCTGGATC <b>SEC ID Nº :19</b>
ENDO3	At4g21590	Directo:AAGGATCCATTCGACAAACTTTGTAAC <b>SEC ID Nº :20</b> Inverso: AGAGTGGTCTTGGGAATATTTATCTCAG <b>SEC ID Nº :21</b>
ENDO2	At1g68290	Directo:ACGGATCCCATTTCAAAGAACTCTGA <b>SEC ID Nº :22</b> Inverso: GACCAATCATTATGCTGTAACITTCAG <b>SEC ID Nº :23</b>
ENDO1 <b>SEC ID Nº :2</b>	At1g11190	Directo: CAGGATCCAAGTTTCAAACITTTGAAG <b>SEC ID Nº :24</b> Inverso: CGGTATGTCCGGTTTGGTTCAAGTGG <b>SEC ID Nº :25</b>

**35 EJEMPLO 9: CARACTERIZACION BIOQUÍMICA DE LAS PROTEINAS CANDIDATAS**

Como un ensayo simple para observar si la proteína candidata es activa, se incubo algún plásmido superenrollado con diferentes diluciones de extracto de proteína (precipitado de sulfato de amonio aL 80%). Como resultado de la presencia de una endonucleasa en el extracto de proteína, la estructura súper enrollada debe relajarse y aparecerán  
40 algunas nuevas bandas de ADN cuando se procesa el medio de incubación sobre un gel de agarosa. Usando las diferentes soluciones pueden compararse las diferentes endonucleasas y observar cual es la más activa. El control

fue siempre endonucleasa CEL I recombinante preparada por precipitación con sulfato de amonio como se describe en el Ejemplo 1 anterior.

5 Para explorar las proteínas candidatas que pueden escindir el ADN heteroduplex en el sitio de emparejamiento erróneo, también se usó un sistema de caracterización rápido basado en tres etapas consecutivas.

10 En la primera etapa las proteínas candidatas se ensayaron para determinar su capacidad de degradar el ADN monocatenario. Esta condición se ensayó porque el ADN en el sitio de emparejamiento erróneo en el ADN heteroduplex es monocatenario. Por tanto, una endonucleasa que no puede digerir el ADN monocatenario se supone que no puede escindir el ADN en el sitio de emparejamiento erróneo. En segundo lugar, la proteína candidata debe escindir un ADN heteroduplex de ensayo en un sitio de emparejamiento erróneo conocido y bien caracterizado. En tercer lugar, las proteínas que pasan el ensayo 1 y 2 se evaluaron para determinar su eficacia para escindir los fragmentos de ADN heteroduplex que llevan todos los tipos de emparejamientos erróneos. Esto se realizó usando herramientas de ADN que comprenden un grupo de construcciones de plásmidos bien caracterizadas que contienen en posición específica del inserto, cada uno de los cuatro posibles nucleótidos.

#### **Capacidad de degradar un ADN monocatenario**

20 La actividad de las proteínas candidatas sobre la degradación de un ADN monocatenario se realizó como se describe en el Ejemplo 2. En este análisis las cinco proteínas candidatas al completo mostraron actividad nucleasa y se clasificaron desde la más activa a la menos activa de la siguiente manera: ENDO1, ENDO5, ENDO2, ENDO3 y ENDO4.

#### **Escisión de ADN heteroduplex en sitios de emparejamiento erróneo C-A/T-G**

25 Para ensayar si las proteínas candidatas producidas en tabaco podían reconocer mutaciones puntuales sencillas, se ensayó en la actividad de los extractos de proteínas en el ADN heteroduplex creado como se ha descrito en el Ejemplo 3. Los productos de la PCR que contenían cada uno o ambos de los alelos se incubaron con una dilución 1/1000 de los extractos de proteína derivados de hojas agroinfiltradas bien con el gen de endonucleasa candidato o con la GFP como control, de la siguiente manera: se incubaron 500 ng de productos de la PCR con extracto de proteína de tabaco en un volumen final de 25µl que contenía Tris-HCl 50mM (pH 7,6), MgCl<sub>2</sub> 10mM, DTT 1mM y 5% de PEG-8000 durante 30 minutos a 37° C. Las reacciones se detuvieron con EDTA a la concentración final de 80mM y se analizaron en un gel de agarosa al 3%.

30 Si los dos oligonucleótidos estaban marcados con fluorescencia, los productos de digestión se analizaban en gel de acrilamida, en un secuenciador de ADN 377Abi. En este experimento, se predijo que si el extracto de proteína contenía una endonucleasa específica de emparejamiento erróneo, el ADN heteroduplex podría escindir el sitio de emparejamiento erróneo, liberando por lo tanto dos bandas de 256 pb y 405 pb.

40 Los resultados se muestran en la figura 6 (obsérvese que la banda de 256 pb, marcada en rojo, es claramente visible cuando la enzima está en presencia de heteroduplex y ausente de otra manera; esta banda se diferencia menos en blanco y negro).

45 A partir de este análisis bioquímico, se llega a la conclusión de que las cinco endonucleasas muestran actividades de escisión específica de emparejamiento erróneo. Además, las tres enzimas, ENDO 1, ENDO 5, ENDO 2, que tienen mayor actividad específica nucleasa para el ADN monocatenario, escinden también con mayor eficacia el ADN heteroduplex en el sitio de emparejamiento erróneo C-A/T-G. Además, entre estas tres endonucleasas, ENDO 1 y ENDO 5 eran las más activas. Por tanto, estas dos nucleasas se seleccionaron para una caracterización más precisa.

#### **Eficacia de escisión de emparejamiento erróneo por la endonucleasa candidata en diferentes emparejamientos erróneos**

50 El principal objetivo de este trabajo es identificar una endonucleasa que escinda emparejamientos erróneos que CEL-I no reconozca eficazmente. Esto se realizó usando herramientas de ADN que comprenden un grupo de construcciones de plásmidos bien caracterizadas que contienen en posición específica del inserto cada uno de los cuatro posibles nucleótidos. Para reconstruir todos los tipos de emparejamiento erróneo, como molde se usó la combinación de estas construcciones en la amplificación por PCR.

55 El producto de la PCR de emparejamiento erróneo se incubó como se ha indicado anteriormente (véase el ejemplo 5) con la endonucleasa candidata y se analizó sobre la secuenciadora LICOR.

60 Los resultados se muestran en la Figura 7.

65 En este análisis, CEL I al igual que ENDO 5 reconoce débilmente el tipo de emparejamiento erróneo T/T. Por otro lado, ENDO 1 reconoce casi todos los tipos de emparejamiento erróneo con elevada eficacia. A partir de este análisis, se puede llegar a la conclusión de que ENDO 1 puede reconocer emparejamientos erróneos no detectados por CEL- 1.

**EJEMPLO 10: SENSIBILIDAD DE LA ENDO 1 PARA DETECTAR UN ALELO MUTANTE EN UN GRUPO DE ADN.**

5 La sensibilidad de ENDO 1 se evaluó como se describe en el Ejemplo 3, con diluciones de 2, 4, 6, 8, 10, 15, 20, 25, 30, 35, 40, 50 o 60 veces con ADN portador del alelo de tipo silvestre.

En la Figura 8, se presentan los resultados de este experimento colectivo, que muestra la detección de un alelo a partir de un mutante homocigoto entre 60 alelos a partir de un tipo silvestre homocigoto. Se realizó la PCR de los dos organismos y los amplicones se cuantificaron en gel de agarosa y mezclaron juntos siguiendo una proporción decreciente de 1:1 a 1:60 de los amplicones de la PCR. Después, se ensayó la actividad de ENDO 1 en estas diferentes proporciones de mutantes: tipo silvestre.

En experimentos idénticos, CEL I purificada de apio puede detectar un alelo en un máximo de 16, con baja sensibilidad y solo se obtiene una sensibilidad correcta cuando la dilución es 8 veces inferior o igual.

En resumen, las endonucleasas de acuerdo con la invención y particularmente ENDO 1 tienen una interferencia de referencia mucho menor que CEL I, pueden usarse a diluciones muy altas en comparación con CEL I y tienen mejor especificidad y actividad que la endonucleasa CEL I.

***Detección de mutación puntual en gel de acrilamida.***

Se creó una serie de mutantes basados en el gen Rx. Se diseñaron diferentes plásmidos conteniendo cada uno un tipo de emparejamiento erróneo para demostrar la especificidad de ENDO I. La mezcla de PCR contenía oligonucleótidos específicos para los plásmidos. Para la PCR se usaron oligonucleótidos marcados con los fluoróforos ROX y FAM (MWG®) para permitir la detección de emparejamientos erróneos en ABI377 MVG®.

Como se muestra en la figura 9, ENDO I no escinde cuando solo hay homoduplex, por ejemplo línea 1 o línea 5 y las únicas bandas que pueden observarse en el gel son homoduplex en la parte superior del gel (aproximadamente 600 pb).

En cualquier momento se obtiene un emparejamiento erróneo en las muestras de este documento (resultante de la formación de heteroduplex entre dos cadenas de ADN procedentes de 2 plásmidos diferentes, como por ejemplo en las líneas 2, 3, 4 ó 6, 7 y 9) la ENDO I reconoce y escinde en el sitio de emparejamiento erróneo, dando como resultado la aparición de dos bandas correspondientes a los dos productos, cada uno marcado con un fluoróforo. Para la digestión con ENDO I, puede observarse una disminución en la referencia y solo pueden detectarse algunos emparejamientos erróneos con ENDO I (línea 9 por ejemplo).

***Detección de una mutación puntual conocida en gel de acrilamida de ENDO 1 en dos diluciones diferentes D 1000 y D 5000.***

Se usaron dos plásmidos que contenían una forma silvestre y una forma mutada del gen Rx. Estos difieren solamente en una mutación puntual conocida. La mezcla de la PCR contiene oligonucleótidos específicos para los plásmidos. Se usaron oligonucleótidos marcados con los fluoróforos ROX y FAM (MWG®) en la PCR para permitir la detección de emparejamiento erróneo en ABI377 MVG®.

Como se muestra en la Figura 10, en cualquier momento existe un heteroduplex en las muestras de este documento (M5+M12), ENDO 1 puede reconocer y escindir en el sitio de emparejamiento erróneo, dando como resultado la aparición de dos bandas marcadas con un fluoróforo. Cuando solamente están presentes los homoduplex M5 o M12, ENDO 1 no escinde y las únicas bandas que pueden observarse en el gel son homoduplex en la parte superior del gel (aproximadamente 600 pb). Cuando se usa la D 5000 para la ENDO 1, puede observarse una mejora de la referencia. Por lo tanto, ENDO 1 puede usar una dilución muy alta en comparación con cualquiera de las nucleasas conocidas en la técnica.

## LISTA DE SECUENCIAS

55

<110> GENOPLANTE-VALOR  
 INSTITUTO NACIONAL DE INVESTIGACION AGRONÓMICA  
 BENDAHMANE, Abdelhafid  
 STURBOIS, Bénédicte  
 TRIQUES, Karine  
 CABOCHE, Michel

60

<120> MÉTODO PARA PRODUCIR ENDONUCLEASAS ALTAMENTE SENSIBLES, NUEVAS PREPARACIONES DE ENDONUCLEASAS Y USOS DE LAS MISMAS.

65

<130> MJP/bv1516-19  
<150> PCT/EP2004/009159  
<151> 30-07-2004  
5  
<150> PCT/EP2004/009166  
<151> 30-07-2004  
10  
<160> 25  
<170> PatentIn versión 3.3  
<210> 1  
<211> 2640  
15  
<212> ADN  
<213> Arabidopsis thaliana  
<220>  
<221> misc\_feature  
20  
<223> Gen At1g11190 de ENDO 1  
<400> 1

aaattcgatg aggttgttat agacaagaga agacatTTTT atacaaaaga gtttatcatt	60
atataagttt caaacTTTga agatatggca tCGGctTTta gatcatccac gaggttgatt	120
cttgtattag gtatactgat tttgtgttcg gtttcttctg tccgaagctg gagcaaagaa	180
ggtcataattc ttactttag aattgctcag gtaattaagt taatgatcta ttgtttgaag	240
caactatTTT ggTTattctt gtcttatata tgtattagtg agatatacct acaaatTTTt	300
aattaggatt gactTTTaaa ttgctatacg ttaccatgcc taacatctca tgtagatgat	360
catgaataca aacatgtcta atggcatatc aaattccaag tTTTTttggT agagatctga	420
gtcatttgac cgttataaga ttcataacaa aagttcgtat gtgtgtgttt ttgtggtgtg	480
accagaatct tttagaagcc ggaccagcac atgtagtaga gaatctgtta ccggattacg	540
tgaaaggaga tttatcagca ttgtgtgtgt ggcctgacca gatccgacat tggTacaagt	600
atcgttggac cagccatctc cattacatcg aactcccga ccaagcctgc tcttacgaat	660
actctagtaa gTcacaaccg agacatTTTc agataacctt aatccgTTTT ctaattatct	720
tgaaccggag ttaaccaaaa aatcaattac aaataccaaa ccggattaaa aacaggggat	780
tgtcatgac aacatggatt gaaggatatg tgtgtggatg gagcaatcca gaatttcacg	840
tctcagcttC agcattacgg tgaaggaaca tctgatcgta gatgtatgtc atcattttca	900
tttatttcat ataatgatga tatccaaagt gtaactgcgt atTTTgtatt ttgatgcata	960
acttaagttt tTaaaattat aatatacct tgttcaatca catagataac atgaccgaag	1020

```

cccttttgtt cttgtctcat ttcattggag atattcatca ggtttattac tcatcatcga 1080
ttcatttcac acctccacac atatagctct atttccatgt taaatattta attaacatgg 1140
tttttttttt tttccttaa aagccgatgc atgtgggatt cacaagtgat gaaggaggaa 1200
acacgataga tttacgttg tacaacaca aatccaatct acatcatgta agcttcttct 1260
tttgtctctt tcaacttta atttcatcat gaaaacaaaa aaaaaattaa cgaaggaaac 1320
aaaatatgta ggtatgggat agagagatca ttctcacggc tctaaaagaa aactacgaca 1380
agaacttggg tcttctcaa gaggatcttg agaagaacat caccaatgta atagacacta 1440
atttattcat attttactat aattttaaga atctttataa tggttatcat atattagggg 1500
ttatggcacg acgatctatc ttcgtggaca gaatgcaacg atcttatcgc ttgtccacac 1560
aagtaagttt taaattactt ggtttaagat tggcttgacg ctcgtttgaa gctagctaca 1620
aattttgata cttttctgg tccaaaaatc ttacaagat actgaaaata aaataatagg 1680
ttttaaactt ttaatttatt tggagttgga taggattaag tttcactaac ttccaattca 1740
aagtcaatta atagtagttt accatgatta gtgggttgac taatgtacca tatatattac 1800
cttatatcac atcttatttc cgatgtgaga tttcttatga aacataatta gactcgaacc 1860
ttttgtgttt cgatatatgt agtgtattca tgatcagaat cttattaagt ttacaactga 1920
aaactaaaat attaacatca taattataga ttcttaagta ggttttgttt ggggtggagaa 1980
aatatccaat ttcgaataac attatataaa atattgaact aattttaatt gtatacgag 2040
gtatgcttca gagagtataa agttagcttg taaatgggga tacaaggcg tcaagtctgg 2100
tgaaacgtta tcaggtacgt tgtttgcttc tcttttttct cgtacgctaa caaaaatatt 2160
taaaaataaa cccgaccaa tgaagtttaa ttaatcggat taatgatttt taatagtcac 2220
tacttttttt gtgtgggata tatgactgtc taatatataa ttttataaga aagctaaagg 2280
atttgtttta taatttccga taaataattt tgcagaagaa tatttcaata caaggttgcc 2340
aatagtgatg aagagaatag ttcagggagg agttagacta gccatgatac taaaccgggt 2400
ttttagtgac gatcatgcta ttgctggtgt tgctgccact tgaaccaaac ccgacatacc 2460
ggggcatcaa agcatttgat taagagatta tttgatacat tcacaaaatt aattaaggct 2520
gatcagacat tcttttcttt tagtagcttt atctatgtga cagctaatgc ctgtggactg 2580
ctttgtttag aagtgtttag cattagatca tatgctaatt caatgttatt aattcatcgt 2640

```

<210> 2

<211> 305

5 <212> PRT

<213> Arabidopsis thaliana

<220>

<221> misc\_feature

10 <223> At1g11190 de la proteína ENDO 1

<400> 2

ES 2 362 100 T3

Met Ala Ser Ala Phe Arg Ser Ser Thr Arg Leu Ile Leu Val Leu Gly  
 1 5 10 15  
 Ile Leu Ile Leu Cys Ser Val Ser Ser Val Arg Ser Trp Ser Lys Glu  
 20 25 30  
 Gly His Ile Leu Thr Cys Arg Ile Ala Gln Asn Leu Leu Glu Ala Gly  
 35 40 45  
 Pro Ala His Val Val Glu Asn Leu Leu Pro Asp Tyr Val Lys Gly Asp  
 50 55 60  
 Leu Ser Ala Leu Cys Val Trp Pro Asp Gln Ile Arg His Trp Tyr Lys  
 65 70 75 80  
 Tyr Arg Trp Thr Ser His Leu His Tyr Ile Asp Thr Pro Asp Gln Ala  
 85 90 95  
 Cys Ser Tyr Glu Tyr Ser Arg Asp Cys His Asp Gln His Gly Leu Lys  
 100 105 110  
 Asp Met Cys Val Asp Gly Ala Ile Gln Asn Phe Thr Ser Gln Leu Gln  
 115 120 125  
 His Tyr Gly Glu Gly Thr Ser Asp Arg Arg Tyr Asn Met Thr Glu Ala  
 130 135 140  
 Leu Leu Phe Leu Ser His Phe Met Gly Asp Ile His Gln Pro Met His  
 145 150 155 160  
 Val Gly Phe Thr Ser Asp Glu Gly Gly Asn Thr Ile Asp Leu Arg Trp  
 165 170 175  
 Tyr Lys His Lys Ser Asn Leu His His Val Trp Asp Arg Glu Ile Ile  
 180 185 190  
 Leu Thr Ala Leu Lys Glu Asn Tyr Asp Lys Asn Leu Asp Leu Leu Gln  
 195 200 205  
 Glu Asp Leu Glu Lys Asn Ile Thr Asn Gly Leu Trp His Asp Asp Leu  
 210 215 220  
 Ser Ser Trp Thr Glu Cys Asn Asp Leu Ile Ala Cys Pro His Lys Tyr  
 225 230 235 240  
 Ala Ser Glu Ser Ile Lys Leu Ala Cys Lys Trp Gly Tyr Lys Gly Val  
 245 250 255  
 Lys Ser Gly Glu Thr Leu Ser Glu Glu Tyr Phe Asn Thr Arg Leu Pro  
 260 265 270

Ile Val Met Lys Arg Ile Val Gln Gly Gly Val Arg Leu Ala Met Ile  
 275 280 285

Leu Asn Arg Val Phe Ser Asp Asp His Ala Ile Ala Gly Val Ala Ala  
 290 295 300

**Thr  
 305**

5 <210> 3  
 <211> 28  
 <212> ADN  
 <213> Secuencia artificial

10 <220>  
 <223> Cebador 4-960

<400> 3  
 ggtttgtcc agtaatagtg tcagcata 28

15 <210> 4  
 <211> 26  
 <212> ADN  
 <213> Secuencia artificial

20 <220>  
 <223> Cebador 4-721

<400> 4  
 aggaacctga gaaaagactc gccagc 26

25 <210> 5  
 <211> 40  
 <212> ADN  
 <213> Secuencia artificial

30 <220>  
 <223> N-terminal de CEL

<400> 5  
 tatcgttcta gagggatga cgcgattata ttctgtgttc 40

35 <210> 6  
 <211> 27  
 <212> ADN  
 <213> Secuencia artificial

40 <220>  
 <223> C-terminal de CEL

<400> 6  
 tatctgaatt catgccaaag aatgatc 27

45 <210> 7  
 <211> 50  
 <212> ADN  
 <213> Secuencia artificial

50 <220>  
 <223> C-terminal con His de 8 aminoácidos de CEL

<400> 7  
 aattcaatgg tgatggtggt gatggtgatg tgccaaagaa tgatctgagg 50

5 <210> 8  
 <211> 25  
 <212> ADN  
 <213> Secuencia artificial

10 <220>  
 <223> Cebador (R21)

<400> 8  
 gacatatgga ctacagaagc ttggg 25

15 <210> 9  
 <211> 25  
 <212> ADN  
 <213> Secuencia artificial

20 <220>  
 <223> Cebador (R22)

<400> 9  
 gttcacgggt cacatcatgc attcc 25

25 <210> 10  
 <211> 22  
 <212> ADN  
 <213> Secuencia artificial

30 <220>  
 <223> Cebador 4m118

35 <400> 10  
 ttggttgac ttcacttga gc 22

<210> 11  
 <211> 22  
 <212> ADN  
 <213> Secuencia artificial

40 <220>  
 <223> Cebador 4m984

45 <400> 11  
 cacaacaatc agcaatgaca gc 22

<210> 12  
 <211> 23  
 <212> ADN  
 <213> Secuencia artificial

50 <220>  
 <223> Cebador 4-347

55 <400> 12  
 gtgattgctc cacctccgcc acc 23

60 <210> 13  
 <211> 30  
 <212> ADN

<213> Secuencia artificial  
 <220>  
 <223> Cebador 4-134  
 5  
 <400> 13  
 tacagcgatt gatataatat aaaattatcc 30  
  
 <210> 14  
 <211> 27  
 <212> ADN  
 <213> Secuencia artificial  
 10  
  
 <220>  
 <223> Cebador le 2462  
 15  
 <400> 14  
 tgatattgtc gtgcaatatg atgaaac 27  
  
 <210> 15  
 <211> 25  
 <212> ADN  
 <213> Secuencia artificial  
 20  
  
 <220>  
 <223> Cebador le 3082  
 25  
 <400> 15  
 atacctattt agcccacttg gacac 25  
 30  
  
 <210> 16  
 <211> 27  
 <212> ADN  
 <213> Secuencia artificial  
 35  
  
 <220>  
 <223> Cebador directo para ENDO5  
 <400> 16  
 aaggatccga aagctctgtg ttccaga 27  
 40  
  
 <210> 17  
 <211> 28  
 <212> ADN  
 <213> Secuencia artificial  
 45  
  
 <220>  
 <223> Cebador inverso para ENDO5  
 50  
 <400> 17  
 ggagttgta cgtgggttct caaggatc 28  
  
 <210> 18  
 <211> 28  
 <212> ADN  
 <213> Secuencia artificial  
 55  
  
 <220>  
 <223> Cebador directo para ENDO4  
 60  
 <400> 18

ctggatccct gttttaact ttgaaag 28

5 <210> 19  
<211> 26  
<212> ADN  
<213> Secuencia artificial

10 <220>  
<223> Cebador inverso para ENDO4

<400> 19  
ggatgtcaa gtgattctcc tggatc 26

15 <210> 20  
<211> 27  
<212> ADN  
<213> Secuencia artificial

20 <220>  
<223> Cebador directo para ENDO3

<400> 20  
aaggatccat tcgacaaact ttgtaac 27

25 <210> 21  
<211> 28  
<212> ADN  
<213> Secuencia artificial

30 <220>  
<223> Cebador inverso para ENDO3

<400> 21  
agagtgtctc tgggaatatt tatctcag 28

35 <210> 22  
<211> 26  
<212> ADN  
<213> Secuencia artificial

40 <220>  
<223> Cebador directo para ENDO2

<400> 22  
acggatccca ttcaaagaa ctctga 26

45 <210> 23  
<211> 26  
<212> ADN  
<213> Secuencia artificial

50 <220>  
<223> Cebador inverso para ENDO2

55 <400> 23  
gaccaatcat tatgctgtaa ctctcag 26

60 <210> 24  
<211> 25  
<212> ADN  
<213> Secuencia artificial

<220>

<223> Cebador directo para ENDO1

<400> 24  
caggatccaa gtttcaaact tgaag 25

5

<210> 25  
<211> 26  
<212> ADN  
<213> Secuencia artificial

10

<220>  
<223> Cebador inverso para ENDO1

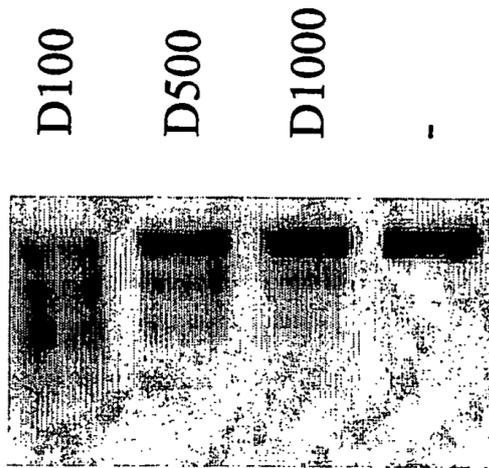
<400> 25  
cggtatgtcg ggtttggttc aagtgg 26

15

## REIVINDICACIONES

1. Un método para producir una endonucleasa recombinante específica de emparejamientos erróneos en el que dicho método comprende:
- expresar dicha endonucleasa recombinante en células de una planta hospedadora completa o de un órgano separado de esta, transformada transitoriamente por agroinfiltración con una cepa de *Agrobacterium* que contiene un vector de expresión que comprende un polinucleótido que codifica dicha endonucleasa;
  - aislar dicha endonucleasa recombinante de dichas células de la planta hospedadora.
2. Un método de acuerdo con la reivindicación 1, en el que dicha agroinfiltración se realiza en una hoja de dicha planta hospedadora.
3. Un método de acuerdo con cualquiera de las reivindicaciones 1 ó 2, en el que dicha planta hospedadora pertenece al género *Nicotiana*.
4. Un método de cualquiera de las reivindicaciones 1 a 3, en el que dicha endonucleasa se aísla del órgano de la planta agroinfiltrada por un proceso que comprende las siguientes etapas:
- extraer el contenido de las células procedentes del órgano agroinfiltrado que expresa dicha endonucleasa;
  - añadir a dicho extracto sulfato de amonio a una concentración final de al menos 30% y separar el precipitado proteico del sobrenadante;
  - añadir a dicho sobrenadante sulfato de amonio a una concentración final de al menos 80% y recuperar el precipitado proteico que contiene la endonucleasa.
5. El método de acuerdo con la reivindicación 4, en el que el sulfato de amonio se añade a una concentración final de 30% en la primera etapa de precipitación y a una concentración final de 80% en la segunda etapa de precipitación.
6. Una preparación de endonucleasa recombinante específica de emparejamientos erróneos que puede obtenerse por el método de cualquiera de las reivindicaciones 1 a 5, en el que dicha endonucleasa se selecciona entre CEL I de *Apium graveolens* y BFN1 de *Arabidopsis thaliana* de SEC ID N°: 2.
7. Una preparación de endonucleasa CEL I recombinante de la reivindicación 6, en la que dicha endonucleasa CEL I recombinante tiene la siguiente preferencia de emparejamientos erróneos: T/T ~ T/G ~ A/G ~ G/G ~ G/A ~ G/T ≥ A/A ~ C/C ≥ T/C ~ C/T > A/C ~ C/A y puede reconocer un alelo mutante en presencia de 23 veces más el alelo de tipo silvestre.
8. Una preparación de endonucleasa BFN 1 recombinante de la reivindicación 6, que la que dicha endonucleasa BFN 1 recombinante tiene la siguiente preferencia de emparejamientos erróneos: G/G ~ G/A ~ A/G ~ G/T ~ T/G > T/T ~ A/A ~ C/C ~ T/C > C/T ~ A/C ~ C/A, y puede reconocer un alelo mutante en presencia de 59 veces más el alelo de tipo silvestre.
9. El uso de una preparación de endonucleasa recombinante de acuerdo con cualquiera de las reivindicaciones 6 a 8, para detectar en un dúplex de ADN, un emparejamiento erróneo resultante de una sustitución de bases o de inserción o deleción de uno o más nucleótidos en una cadena de dicho dúplex.
10. El uso de una preparación de endonucleasa recombinante de acuerdo con cualquiera de las reivindicaciones 6 a 8, en un protocolo de escisión de emparejamiento erróneo por detección de lesiones inducidas localmente en los genomas (TILLING).
11. El uso de una preparación de endonucleasa recombinante de acuerdo con cualquiera de las reivindicaciones 6 a 8, para la identificación de polimorfismos de ADN en poblaciones naturales, por Ecotilling.
12. El uso de una preparación de endonucleasa recombinante de acuerdo con cualquiera de las reivindicaciones 6 a 8, como un reactivo de detección de emparejamientos erróneos.
13. El uso de una preparación de endonucleasa recombinante de acuerdo con cualquiera de las reivindicaciones 6 a 8, en un método de selección de emparejamientos erróneos.
14. El uso de una preparación de endonucleasa recombinante de acuerdo con cualquiera de las reivindicaciones 6 a 8, para seleccionar de manera simultánea una o más mutaciones en un gen diana en una población de cualquier organismo o de una línea celular derivada del mismo realizando las etapas de:

- 5 a) amplificar dicho gen diana o parte del mismo para cada individuo de dicha población,  
b) clasificar dichos productos de amplificación en una matriz bi o tridimensional, que comprende líneas, filas (matriz bidimensional) y columnas (matriz tridimensional),  
c) reagrupar dichos productos de amplificación, para obtener diferentes grupos, representando cada grupo una fila, una línea o una columna de dicha matriz,  
d) añadir a cada grupo un producto de amplificación de referencia obtenido de un gen no mutado e incubar dichos grupos en condiciones que permitan la formación de heteroduplex, e  
e) incubar cada grupo con dicha preparación de endonucleasa, y  
f) detectar la presencia de heteroduplex en dichos grupos incubados.
- 10 **15.** El uso de una preparación de endonucleasa recombinante de acuerdo con cualquiera de las reivindicaciones 6 a 8, para seleccionar de manera simultánea una o más mutaciones en un gen diana en una población de cualquier organismo o de línea celular derivada del mismo, realizando las etapas de:
- 15 a) clasificar cada individuo de dicha población en una matriz bi- o tri- dimensional que comprende líneas, filas (matriz bidimensional) y columnas (matriz tridimensional),  
b) agrupar cada fila, línea y columna para obtener diferentes grupos, representando por lo tanto cada grupo una fila, línea o una columna de dicha matriz,  
c) añadir a cada grupo un producto génico de referencia obtenido de un gen no mutado,
- 20 d) amplificar dicho gen diana o parte del mismo en cada grupo para conseguir grupos de productos amplificados,  
e) incubar dichos grupos de productos amplificados en condiciones que permitan la formación de heteroduplex,  
f) incubar dichos grupos de productos amplificados con dicha preparación de endonucleasa, y
- 25 g) detectar la presencia de heteroduplex en dichos grupos incubados.



**Figura 1**

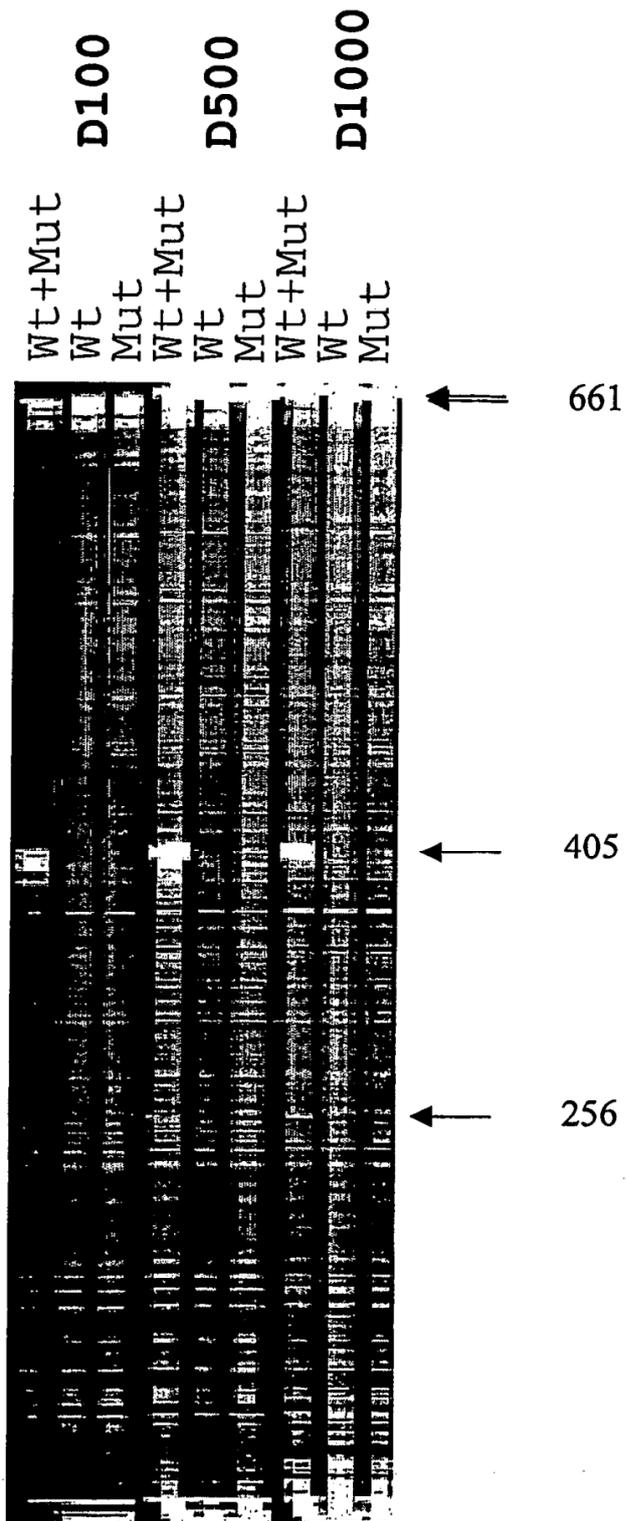
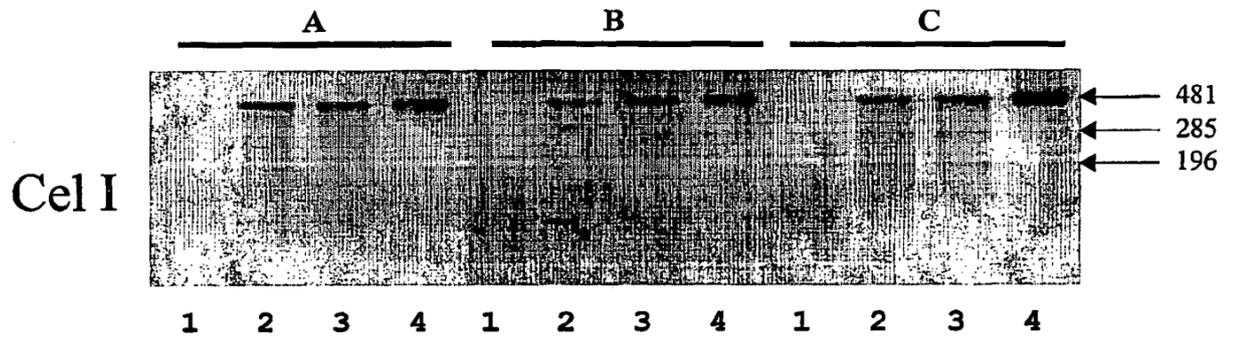
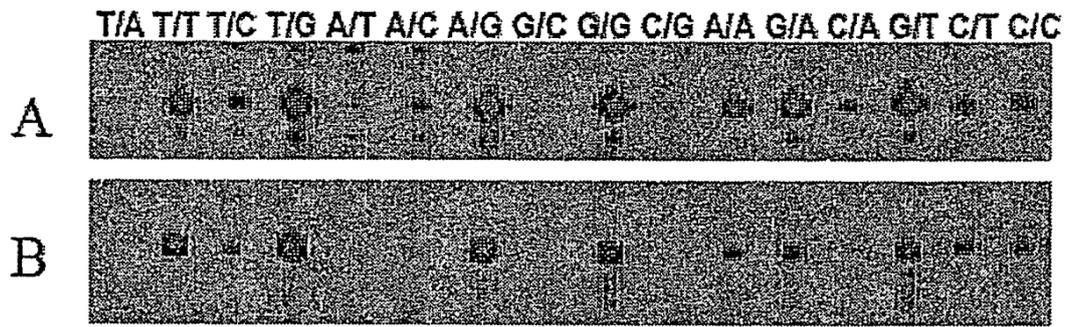


Figura 2

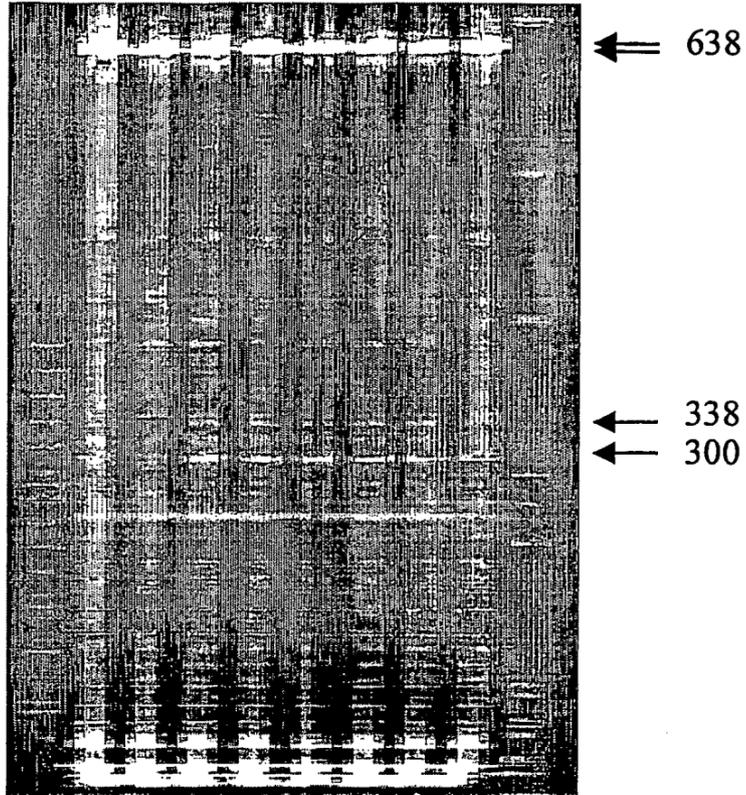


**Figura 3**



**Figura 4**

**1 2 3 4 5 6 7 8**

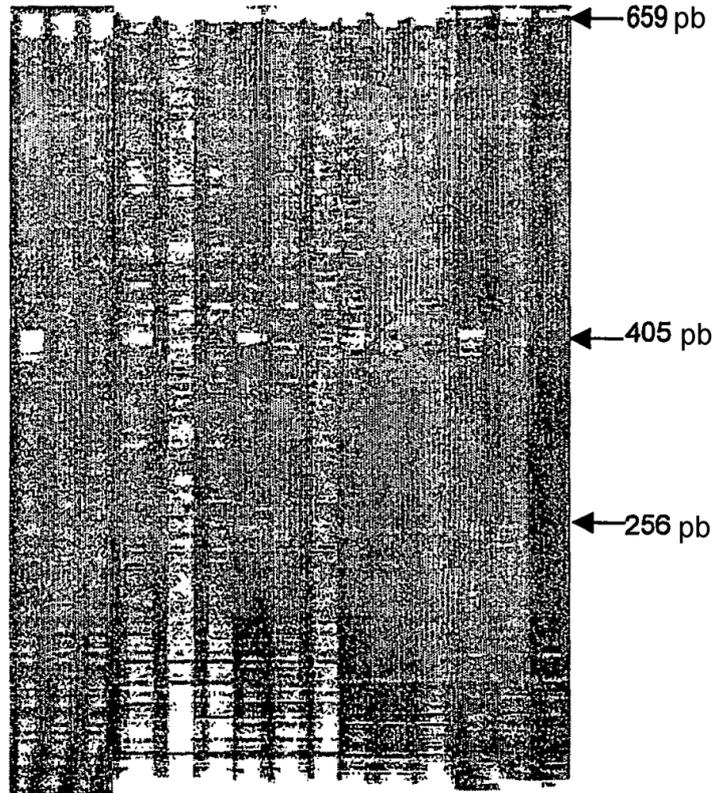


- 1** Guisante Wt (tipo silvestre)
- 2** Le1
- 3** Wt+ Le1
- 4** Wt+ Le1+2ADN
- 5** Wt+ Le1+4ADN
- 6** Wt+ Le1+6ADN
- 7** Wt+ Le1+8ADN
- 8** Wt+ Le1+10ADN

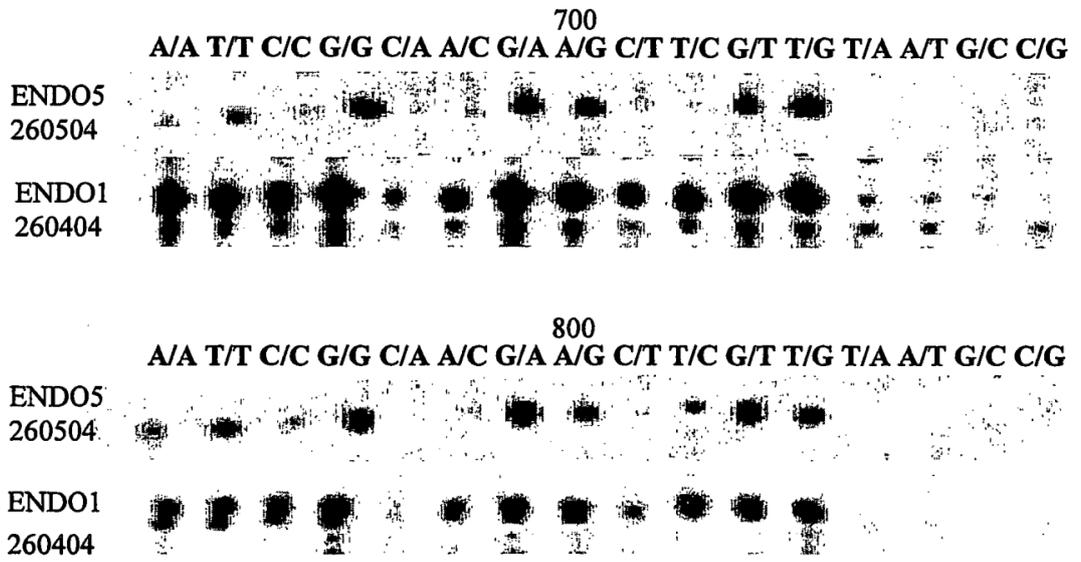
**Figura 5**

ENDO5 ENDO4 ENDO3 ENDO2 ENDO1

Ht Ho Ho Ht Ho Ho Ht Ho Ho Ht Ho Ho Ht Ho Ho



**Figura 6**



**Figura 7**

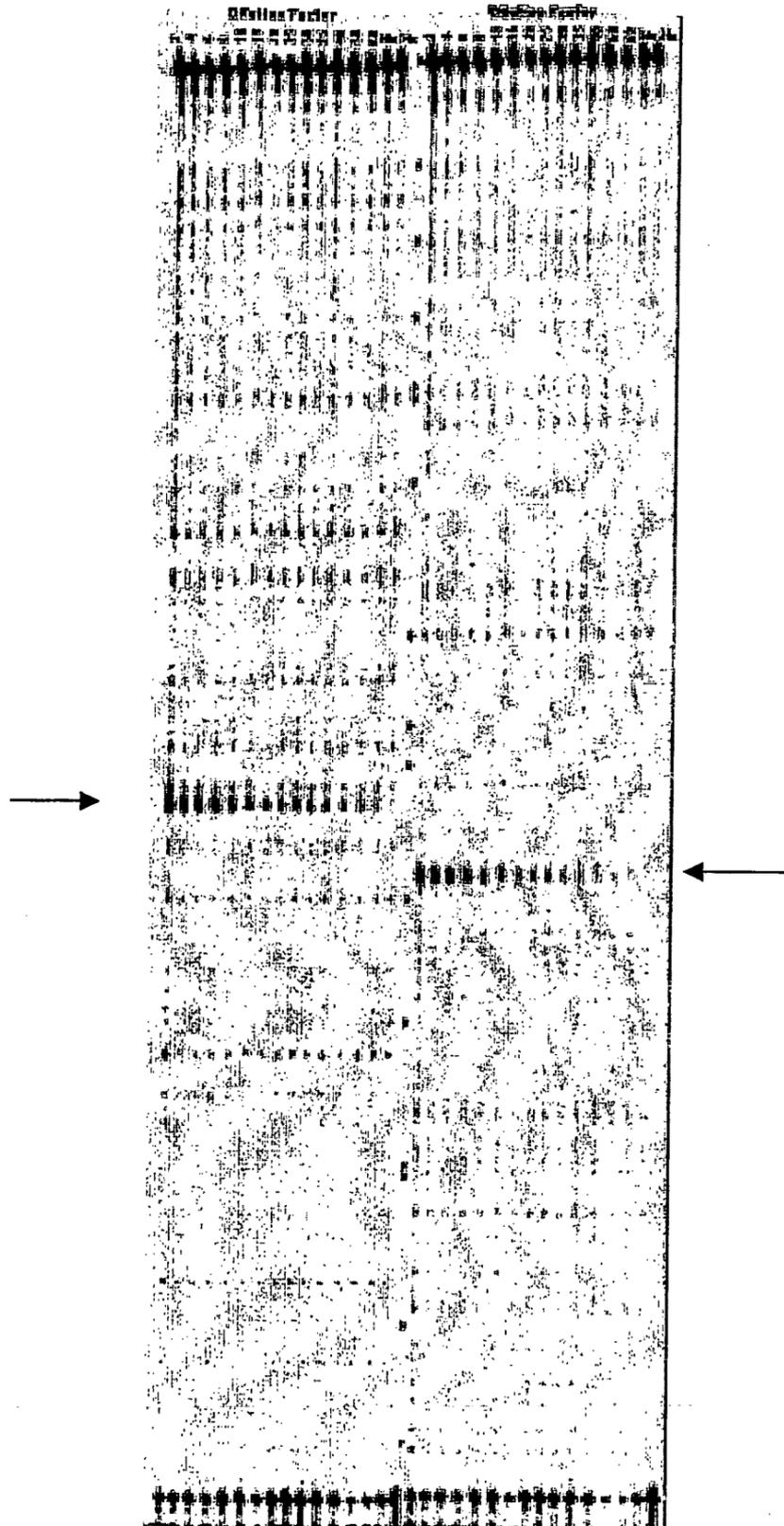


Figura 8

# D1000

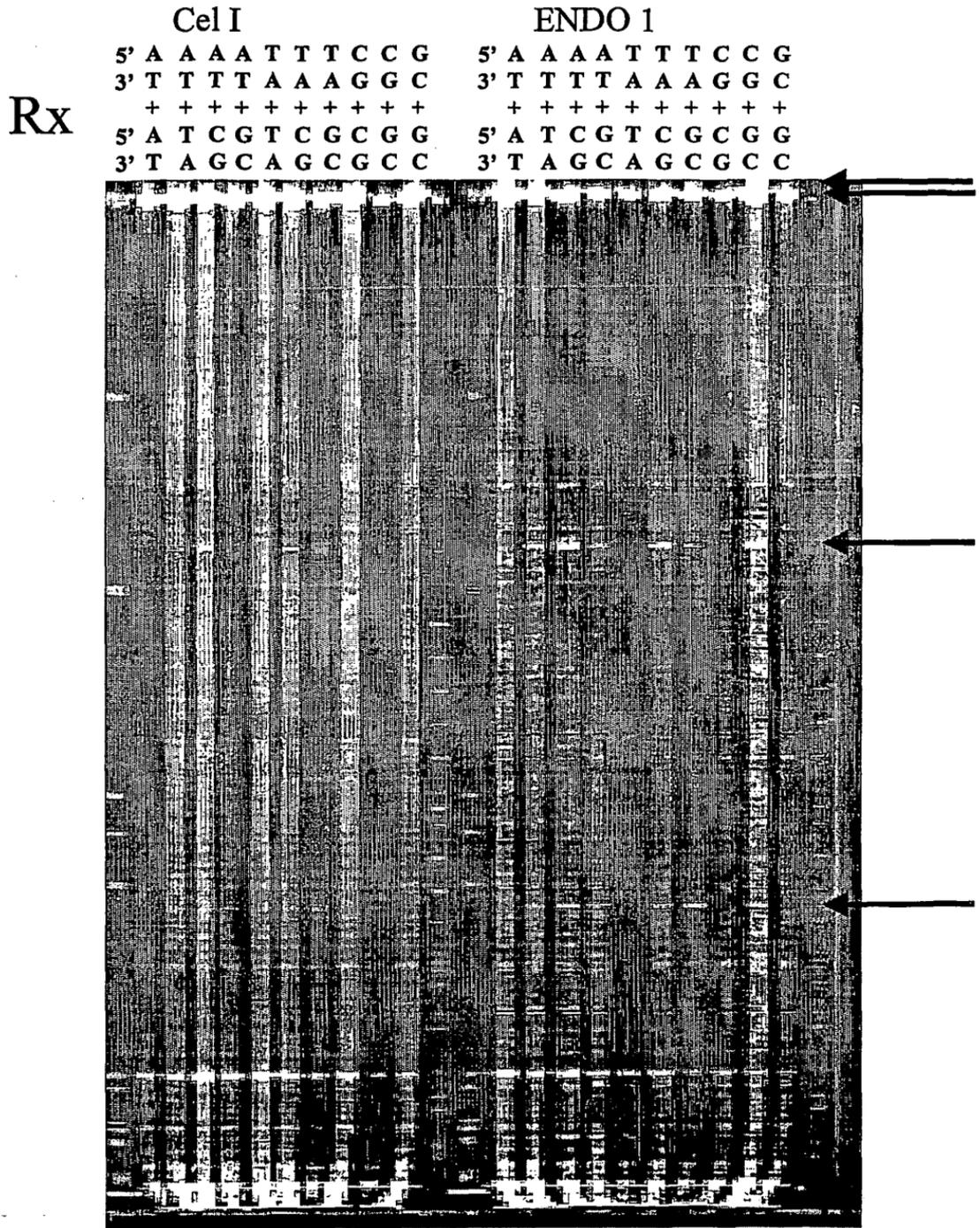


Figura 9

**D1000**

**D5000**

Cell ENDO1

Cell ENDO1

M5M12  
M5  
M12  
M5M12  
M5  
M12

M5M12  
M5  
M12  
M5M12  
M5  
M12

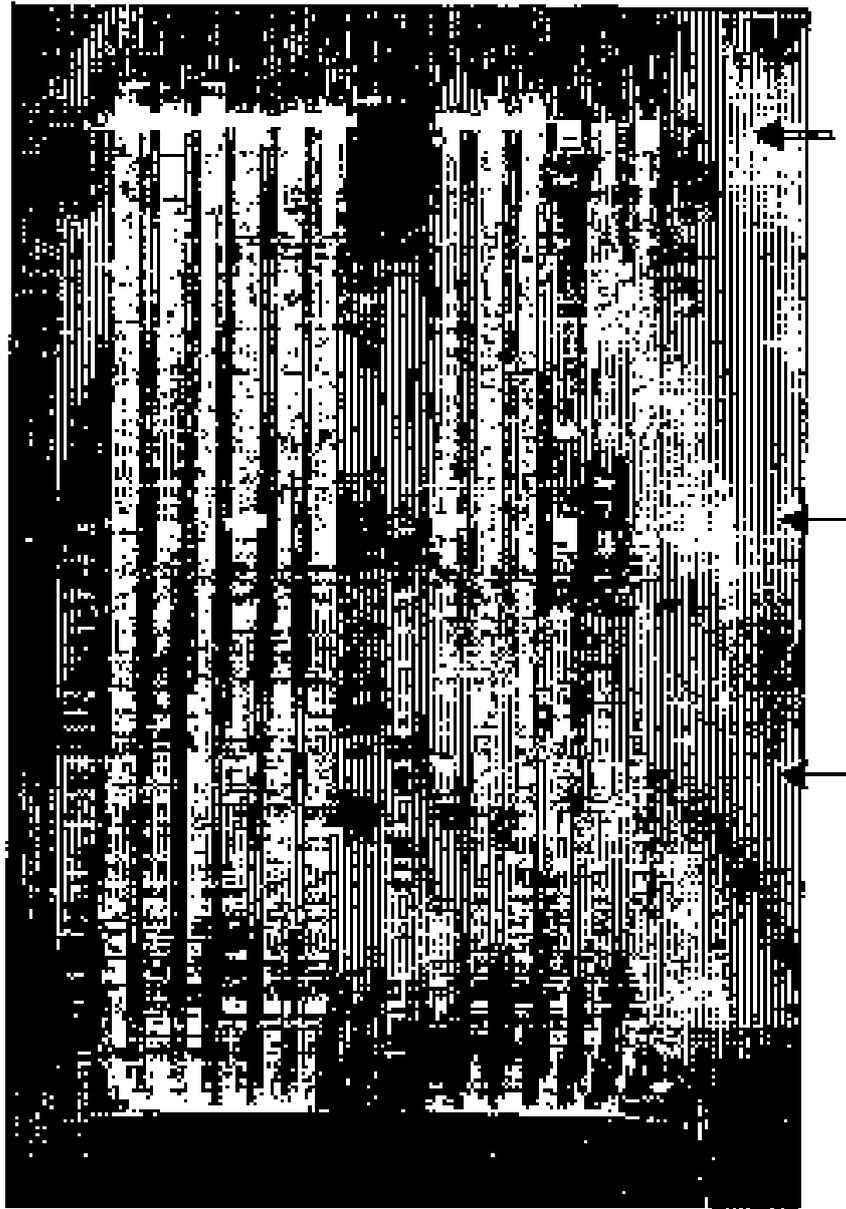


Figura 10