

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 949 396**

51 Int. Cl.:

A01H 1/08 (2006.01)
A01H 5/00 (2008.01)
A01H 5/06 (2008.01)
A01H 5/10 (2008.01)
C07K 14/415 (2006.01)
C12N 15/82 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **28.08.2014** E 14182719 (6)

97 Fecha y número de publicación de la concesión europea: **24.05.2023** EP 2989889

54 Título: **Generación de plantas haploides**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
28.09.2023

73 Titular/es:

KWS SAAT SE & CO. KGAA (100.0%)
Grimsehlstraße 31
37574 Einbeck, DE

72 Inventor/es:

HOUBEN, ANDREAS;
KARIMI-ASHIYANI, RAHELEH;
ISHII, TAKAYOSHI;
STEIN, NILS y
KUMLEHN, JOCHEN

74 Agente/Representante:

SÁEZ MAESO, Ana

Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 949 396 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

Generación de plantas haploides

- 5 La presente invención se refiere a plantas no transgénicas y transgénicas, preferiblemente plantas de cultivo, que comprenden una secuencia de nucleótidos que codifica una proteína centrómero histona H3 (CENH3) que comprende un dominio CATD, en la que la secuencia de nucleótidos comprende una mutación que provoca una sustitución de aminoácidos en (a) posición 92 de SEQ ID NO: 34; (b) posición 106 de SEQ ID NO: 44; o (c) posición 4 de SEQ ID NO: 1, en la que la sustitución de aminoácidos confiere la actividad biológica de un inductor de haploides. Además, la presente invención proporciona métodos para generar las plantas de la presente invención.
- 10 La generación y uso de haploides es uno de los medios biotecnológicos más poderosos para mejorar las plantas cultivadas. La ventaja de los haploides para los mejoradores es que la homocigosidad puede lograrse ya en la primera generación después de la dihaploidización, creando plantas doblemente haploides, sin necesidad de varias generaciones de retrocruzamiento requeridas para obtener un alto grado de homocigosidad. Además, el valor de los haploides en la investigación y el mejoramiento de plantas radica en el hecho de que las células fundadoras de los
- 15 haploides dobles son productos de la meiosis, de modo que las poblaciones resultantes constituyen grupos de diversos individuos recombinantes y al mismo tiempo fijados genéticamente. De este modo, la generación de haploides duplicados proporciona no solo una variabilidad genética perfectamente útil para seleccionar con respecto a la mejora de cultivos, sino que también es un medio valioso para producir poblaciones de mapeo, endogamia recombinante, así como mutantes homocigóticos instantáneos y líneas transgénicas.
- 20 Lermontova and Schubert (First Edition. Edited by Jiming Jiang and James A. Birchler. 2013 John Wiley & Sons, Inc. Publicado en 2013 por John Wiley & Sons, Inc.) discuten la regulación de la ocurrencia, la conservación funcional, las interacciones y la aplicación potencial de la variante centromérica de la histona H3 y sus peculiaridades en las plantas.
- 25 Los haploides se pueden obtener por enfoques *in vitro* o *en vivo*. Sin embargo, muchas especies y genotipos son recalcitrantes a estos procedimientos. Alternativamente, los cambios sustanciales de la variante histona H3 específica del centrómero (CENH3, también llamada CENP-A), al intercambiar sus regiones N-terminales y fusionarlas con GFP ("GFP-tailswap" CENH3), crean líneas inductoras haploides en la planta *Arabidopsis thaliana* el modelo. (Ravi and Chan, Nature, 464 (2010), 615-618). Las proteínas CENH3 son variantes de las proteínas histonas H3 que son miembros del complejo cinetocoro de centrómeros activos. Con estas líneas inductoras de haploides "GFP-tailswap", la haploidización se produjo en la progenie cuando se cruzó una planta inductora de haploides con una planta de tipo salvaje. Curiosamente, la línea inductora haploide se mantuvo estable tras la autofecundación, lo que sugiere que una
- 30 competencia entre el centrómero modificado y el de tipo salvaje en el embrión híbrido en desarrollo da como resultado la inactivación del centrómero del padre inductor y, en consecuencia, la eliminación del cromosoma uniparental. Como resultado, los cromosomas que contienen la proteína CENH3 alterada se pierden durante el desarrollo embrionario temprano, lo que produce una progenie haploide que contiene solo los cromosomas del progenitor de tipo salvaje.
- 35 De este modo, pueden obtenerse plantas haploides cruzando plantas transgénicas "GFP-tailswap" como inductoras de haploides con plantas de tipo salvaje. Sin embargo, como se describió anteriormente, esta técnica requiere cambios sustanciales de la proteína CENH3 y las plantas comprenden un transgén heterólogo, que es económicamente problemático debido a la creciente renuencia del público hacia los cultivos modificados genéticamente.
- 40 Por lo tanto, un objeto de la presente invención es superar los problemas antes mencionados y, en particular, proporcionar plantas inductoras de haploides alternativas que no comprendan modificaciones sustanciales de su proteína CENH3 y/o que no estén modificadas genéticamente.
- 45 Este problema se resuelve mediante el objeto de las reivindicaciones independientes, en particular una planta que tiene actividad biológica de un inductor de haploides y que comprende una secuencia de nucleótidos que codifica una proteína histona H3 centrómero (CENH3) que comprende un dominio CATD, en la que la secuencia de nucleótidos comprende una mutación provocando una sustitución de aminoácido en (a) posición 92 de SEQ ID NO: 34; (b) posición 106 de SEQ ID NO: 44; o (c) posición 4 de SEQ ID NO: 1, en la que la sustitución de aminoácidos confiere la actividad biológica de un inductor de haploides. El dominio CATD de la proteína CENH3 corresponde a la secuencia de aminoácidos desde la posición 113 hasta la posición 155 como se establece en SEQ ID No. 38 derivada de *Arabidopsis thaliana* y/o el dominio CATD de la proteína CENH3 está codificado por una secuencia de nucleótidos correspondiente
- 50 a los nucleótidos desde la posición 337 a la posición 465 como se establece en SEQ ID No. 37 derivada de *Arabidopsis thaliana*. Las secuencias de *A. thaliana* sirven solo como referencias y no limitan la invención a la secuencia de *A. thaliana* particular. Debido al alto nivel de conservación, un experto en la técnica puede encontrar la secuencia de nucleótidos y la secuencia de aminoácidos correspondientes a las secuencias de *A. thaliana* en cualquier otro material vegetal o especie vegetal.
- 55 La mutación que causa una sustitución de un aminoácido se encuentra dentro de la hélice $\alpha 2$ del dominio CATD. La hélice $\alpha 2$ corresponde a la secuencia de aminoácidos desde la posición 127 hasta la posición 155 como se establece en SEQ ID No. 38 derivada de *Arabidopsis thaliana* y/o la hélice $\alpha 2$ está codificada por una secuencia de nucleótidos correspondiente a los nucleótidos desde la posición 379 hasta la posición 465 como se establece en SEQ ID No. 37

derivada de *Arabidopsis thaliana*. Las secuencias de *A. thaliana* sirven solo como referencias y no limitan la invención a las secuencias de *A. thaliana* particular. Debido al alto nivel de conservación, los expertos en la técnica pueden encontrar la secuencia de nucleótidos y la secuencia de aminoácidos correspondientes a las secuencias de *A. thaliana* en cualquier otro material vegetal o especie vegetal.

5 Las proteínas CENH3 son variantes de las proteínas histonas H3 que son miembros del complejo cinetocoro de centrómeros activos, es decir, la estructura de la proteína en los cromosomas donde se unen las fibras del huso durante la división celular. Básicamente, las proteínas CENH3 se caracterizan por un dominio de cola variable, que no forma una estructura secundaria rígida, y un dominio de pliegue de histona conservado que consiste en tres regiones α -helicoidales, denominadas $\alpha 1$ a 3, que están conectadas por secciones de bucle. Dentro del dominio de plegamiento de histonas se encuentra el dominio CATD altamente conservado (dominio diana CENP-A), que está formado por partes de la hélice $\alpha 1$, la hélice $\alpha 2$ completa y el bucle de conexión 1. El dominio CATD conservado es necesario para que los acompañantes carguen CENH3 y, de este modo, es vital para la localización del cinetocoro y la función del centrómero.

15 Los presentes inventores descubrieron sorprendentemente que se puede obtener una planta que posee la capacidad de producir descendencia haploide, es decir, un inductor haploide, sustituyendo un solo aminoácido dentro del dominio CATD conservado en la hélice $\alpha 2$ de la proteína CENH3. Ventajosamente, esto puede lograrse mediante métodos transgénicos, así como no transgénicos. Se prefieren los métodos no transgénicos debido a los enormes costes de la desregulación de los organismos genéticamente modificados (GMO), así como al creciente rechazo público de los organismos genéticamente modificados (GMO) o las plantas generadas por medio de GMO, en particular cultivos para el consumo humano, y el amplio mercado procedimientos de autorización que incluyan evaluaciones de seguridad rigurosas de tales GMO.

20 La presente invención proporciona una planta que tiene actividad biológica de un inductor haploide y que comprende una secuencia de nucleótidos que codifica una proteína centrómero histona H3 (CENH3) que comprende un dominio CATD, en la que la secuencia de nucleótidos comprende una mutación que provoca una sustitución de aminoácidos en (a) posición 92 de SEQ ID NO: 34; (b) posición 106 de SEQ ID NO: 44; o (c) posición 4 de SEQ ID NO: 1, en la que la sustitución de aminoácidos confiere la actividad biológica de un inductor de haploides.

25 En una realización, la presente invención se refiere a una planta que comprende una secuencia de nucleótidos que codifica una proteína centrómero histona H3 (CENH3) que comprende un dominio CATD, en la que la parte de la secuencia de nucleótidos que codifica el dominio CATD comprende al menos una mutación y en la que la al menos una mutación provoca una sustitución de aminoácidos en la hélice $\alpha 2$ que a) está codificada por una secuencia de nucleótidos que corresponde a los nucleótidos de la posición 379 a la posición 465 como se establece en SEQ ID No. 37 derivada de *Arabidopsis thaliana*, que corresponde a la secuencia de aminoácidos desde la posición 127 hasta la posición 155 como se establece en SEQ ID No. 38 derivada de *arabidopsis thaliana*, o tener la secuencia consenso de SEQ ID No. 1, y b) estar posicionada dentro del dominio CATD de la proteína CENH3 como se definió anteriormente. De este modo, la al menos, una sustitución de un aminoácido está ubicada en la hélice $\alpha 2$ del dominio CATD. La hélice $\alpha 2$ no mutada del dominio CATD está altamente conservada entre las especies de plantas y tiene una longitud de 29 aminoácidos que comienza en la posición 1 y termina en la posición 29. En la presente invención, cualquier posición de aminoácido dada con respecto a la hélice $\alpha 2$ o la secuencia consenso descrita a continuación se refiere a este sistema de numeración. Preferiblemente, la hélice $\alpha 2$ no mutada presenta la secuencia de aminoácidos que se indica en la tabla 1.

Tabla 1: Aminoácidos especificados en la hélice $\alpha 2$ del dominio CATD

Posición dentro de la hélice $\alpha 2$	Aminoácido (s)
1	X
2	X
3	A
4	L o V
5	X
6	A, C o S
7	I, L, M o V

ES 2 949 396 T3

8	H o Q
9	E
10	A o S
11	A, S o T
12	E
13	X
14	X
15	I, L o V
16	I o v
17	X
18	I, L, M o V
19	F o L
20	X
21	X
22	X
23	X
24	X
25	C
26	A, S o T
27	I, L o V
28	H
29	A o S

Más preferiblemente, la hélice $\alpha 2$ tiene la secuencia consenso de SEQ ID No. 1, que es

XXALXXXQEXXEXXXXXXFXXXXXCAIHX.

1 10 20 29

5 Como se indicó anteriormente, la hélice $\alpha 2$ comprende aminoácidos no específicos [marcados como X] y aminoácidos específicos [marcados como código de una letra].

Un aminoácido no específico como se indica en la tabla 1 o en SEQ ID No. 1 es un aminoácido que, aunque se especifica en un grupo de especies de plantas en particular, en un género de plantas en particular o en una especie de plantas en particular, no se conserva en una gama mayor de especies vegetales. De este modo, un aminoácido no específico de SEQ ID No. 1 o como se proporciona en la tabla 1 está en un grupo de especies de plantas particulares, en un género de plantas particular o en una especie de planta particular un aminoácido específico bien definido que, sin embargo, posiblemente no se encuentra en el mismo lugar en otra especie de planta. De este modo, una sustitución de aminoácido de un aminoácido no específico de SEQ ID No. 1 o como se indica en la tabla 1 significa que en una planta, es decir, en una especie de planta específica, el aminoácido específico pero no conservado es sustituido por otro aminoácido que ocurre naturalmente en ese lugar en este grupo de especies vegetales particulares, en este género vegetal particular o en esta especie vegetal particular en la proteína CENH3 nativa codificada endógenamente de dicha especie vegetal.

Los aminoácidos específicos proporcionados en la tabla 1 y, en particular, los aminoácidos específicos de SEQ ID No. 1 son aquellos que se encuentran en una amplia gama de especies de plantas y que, de este modo, están bien conservados.

En una realización preferida, se ha compilado la secuencia de consenso SEQ ID No. 1 a partir de secuencias de hélice $\alpha 2$ derivadas de especies seleccionadas del grupo que consiste en *Hordeum vulgare*, *Hordeum bulbosom*, *Sorghum bicolor*, *Saccharum officinarum*, *Zea mays*, *Setaria italica*, *Oryza minuta*, *Oryza sativa*, *Oryza australiensis*, *Oryza alta*, *Triticum aestivum*, *Secale cereale*, *Malus domestica*, *Brachypodium distachyon*, *Hordeum marinum*, *Aegilops tauschii*, *Daucus glochidiatus*, *Beta vulgaris*, *Daucus pusillus*, *Daucus muricatus*, *Daucus carota*, *Eucalyptus grandis*, *Nicotiana sylvestris*, *Nicotiana tomentosiformis*, *Nicotiana tabacum*, *Solanum lycopersicum*, *Solanum tuberosum*, *Coffea canephora*, *Vitis vinifera*, *Erythrante guttata*, *Genlisea aurea*, *Cucumis sativus*, *Morus notabilis*, *Arabidopsis arenosa*, *Arabidopsis lyrata*, *Arabidopsis thaliana*, *Crucihimalaya himalaica*, *Crucihimalaya wallichii*, *Cardamine flexuosa*, *Lepidium virginicum*, *Capsella bursa pastoris*, *Olmarabidopsis pumila*, *Arabis hirsute*, *Brassica napus*, *Brassica oleracia*, *Brassica rapa*, *Raphanus sativus*, *Brassica juncea*, *Brassica nigra*, *Eruca vesicaria subsp. sativa*, *Citrus sinensis*, *Jatropha curcas*, *Populus trichocarpa*, *Medicago truncatula*, *Cicer yamashitae*, *Cicer bijugum*, *Cicer arietinum*, *Cicer reticulatum*, *Cicer judaicum*, *Ajanus cajanifolius*, *Cajanus scarabaeoides*, *Phaseolus vulgaris*, *Glycine max*, *Astragalus sinicus*, *Lotus japonicus*, *Torenia fournieri*, *Allium cepa*, *Allium fistulosum*, *Allium sativum*, y *Allium tuberosum*.

En una realización, la mutación provoca una sustitución de un aminoácido específico en la posición 4 de SEQ ID No. 1.

En el contexto de la presente invención, una mutación es preferiblemente una mutación o sustitución puntual no sinónima en la secuencia de ADN que codifica la proteína CENH3 que da como resultado un cambio en el aminoácido. Esto también se llama una mutación sin sentido invertida. Además, el cambio de aminoácido o la sustitución de aminoácido puede ser conservador, es decir, un cambio a un aminoácido con propiedades fisicoquímicas similares, semiconservador, por ejemplo, negativo a un aminoácido cargado positivamente, o radical, es decir, un cambio a un aminoácido muy diferente.

En una realización preferida de la presente invención, la presente planta que tiene actividad biológica de un inductor de haploides es homocigótica con respecto a la al menos una mutación. En una realización adicional de la presente invención, la presente planta que tiene actividad biológica de un inductor de haploides es heterocigota con respecto a la al menos una mutación.

La planta según la presente invención tiene la actividad biológica de un inductor de haploides. Esto significa que el cruce entre la planta según la presente invención y una planta de tipo salvaje o una planta que expresa la proteína CENH3 de tipo salvaje produce al menos 0.1 %, 0.2 %, 0.3 %, 0.4 %, 0.5 %, 0.6 %, 0.7 %, 0.8 %, 0.9 %, preferiblemente al menos 1 %, preferiblemente al menos 2 %, preferiblemente al menos 3 %, preferiblemente al menos 4 %, preferiblemente al menos 5 %, preferiblemente al menos 6 %, preferiblemente al menos 7 %, preferiblemente al menos 8 %, preferiblemente al menos el 9 %, lo más preferiblemente al menos el 10 %, al menos el 15 %, al menos el 20 % o más progenie haploide. Así, una planta de tipo salvaje es preferiblemente una planta de la misma especie que no comprende al menos una mutación de la planta según la presente invención dentro del gen CENH3 endógeno correspondiente, es decir, la planta es capaz de expresar la proteína CENH3 nativa, y una planta que expresa CENH3 de tipo salvaje es preferiblemente una planta de la misma especie que comprende i) una secuencia de nucleótidos que codifica la proteína CENH3 sin la al menos una mutación de la planta según la presente invención y es capaz de expresar dicha proteína CENH3 nativa o ii) una secuencia de nucleótidos que codifica una proteína CENH3 de otra especie vegetal que muestra una funcionalidad comparable a la CENH3 nativa, por ejemplo, tal proteína CENH3 derivada de otra especie vegetal puede introducirse como un transgén.

De este modo, la presente invención proporciona de forma más ventajosa medios y métodos para generar líneas inductoras de haploides en una amplia gama de especies de eudicotiledóneas, dicotiledóneas y monocotiledóneas. La presente invención también permite el intercambio de citoplasma materno y crear, por ejemplo, plantas de esterilidad masculina citoplasmática con un genotipo deseado en una única etapa de procesamiento. La presente invención es ventajosa en la medida en que se puede generar una mutación de un único aminoácido mediante mutagénesis o cualquier otro enfoque no basado en GMO.

De este modo, todo el procesamiento de haploidización mediante la aplicación de una línea inductora de haploides caracterizada por un gen CENH3 endógeno con mutación puntual que codifica una proteína CENH3 con sustituciones de aminoácidos en al menos una de las posiciones proporcionadas por la presente invención no es transgénica en una realización preferida.

5 En el contexto de la presente invención, un gen, alelo o proteína "endógenos" se refiere a una secuencia no recombinante de una planta tal como la secuencia se presenta en la planta respectiva, en particular en la planta de tipo salvaje. El término "mutado" se refiere a una secuencia alterada por humanos. Los ejemplos de mutación no transgénica inducida por humanos incluyen la exposición de una planta a una dosis alta de mutágeno químico, radiológico u otro con el fin de seleccionar mutantes. Alternativamente, las mutaciones transgénicas inducidas por humanos, es decir, alteraciones recombinantes o ingeniería genómica, por ejemplo, mediante nucleasas TALE, nucleasas con dedos de zinc o un sistema CRISPR/Cas, incluyen fusiones, inserciones, deleciones y/o cambios en el ADN o la secuencia de aminoácidos.

15 Una secuencia de polinucleótidos o polipéptidos es "heteróloga o exógena a" un organismo si se origina a partir de una especie extraña o, si es de la misma especie, se modifica a partir de su forma original. "Recombinante" se refiere a una secuencia polinucleotídica o polipeptídica alterada por humanos, es decir, transgénica. Un "transgén" se usa como se entiende el término en la técnica y se refiere a un ácido nucleico, preferiblemente heterólogo, introducido en una célula mediante manipulación molecular humana del genoma de la célula, por ejemplo, por transformación molecular. De este modo, una "planta transgénica" es una planta que comprende un transgén, es decir, es una planta modificada genéticamente. La planta transgénica puede ser la planta inicial en la que se introdujo el transgén, así como su progenie cuyo genoma también contiene el transgén.

20 El término "codificación de secuencia de nucleótidos" se refiere a un ácido nucleico que dirige la expresión de una proteína específica, en particular la proteína CENH3 o partes de la misma. Las secuencias de nucleótidos incluyen tanto la secuencia de la cadena de ADN que se transcribe en ARN como la secuencia de ARN que se traduce en la proteína. Las secuencias de nucleótidos incluyen tanto las secuencias de ácido nucleico de longitud completa como las secuencias de longitud no completa derivadas de las secuencias de longitud completa.

25 El término "gen" se refiere a una secuencia de nucleótidos codificante y secuencias de nucleótidos reguladoras asociadas.

30 El término 'elemento regulador' se refiere a una secuencia, preferiblemente una secuencia de nucleótidos, ubicada aguas arriba (5'), dentro y/o aguas abajo (3') para una secuencia de nucleótidos, preferiblemente una secuencia codificante, cuya transcripción y expresión están controladas por el elemento regulador, potencialmente en conjunción con el aparato biosintético de proteínas de la célula. "Regulación" o "regular" se refiere a la modulación de la expresión génica inducida por elementos de la secuencia de ADN situados principalmente, pero no exclusivamente, aguas arriba (5') desde el inicio de la transcripción del gen de interés. La regulación puede resultar en una respuesta de todo o nada a una estimulación, o puede resultar en variaciones en el nivel de expresión génica.

35 Se dice que un elemento regulador, en particular una secuencia de ADN, como un promotor, está "operablemente unido a" o "asociado con" una secuencia de ADN que codifica un ARN o una proteína, si las dos secuencias están situadas y orientadas de manera que la secuencia de ADN reguladora efectúa la expresión de la secuencia de ADN codificante.

40 Un 'promotor' es una secuencia de ADN que inicia la transcripción de una secuencia de ADN asociada, en particular que se ubica aguas arriba (5') desde el comienzo de la transcripción y está involucrada en el reconocimiento y siendo de la ARN-polimerasa. Dependiendo de la región promotora específica, también puede incluir elementos que actúan como reguladores de la expresión génica, tales como activadores, potenciadores y/o represores.

45 Un elemento regulador '3' (o extremo '3') se refiere a la porción de un gen que comprende un segmento de ADN, excluyendo la secuencia 5' que impulsa el inicio de la transcripción y la porción estructural del gen, que determina el sitio de terminación correcto y contiene una señal de poliadenilación y cualquier otra señal reguladora capaz de efectuar el procesamiento del ARN mensajero (ARNm) o la expresión génica. La señal de poliadenilación generalmente se caracteriza por efectuar la adición de tramos de ácido poliadenílico al extremo 3' del precursor de ARNm. Las señales de poliadenilación a menudo se reconocen por la presencia de homología con la forma canónica 5'-AATAAA-3'.

50 El término "secuencia codificante" se refiere a la porción de un gen que codifica una proteína, un polipéptido o una porción del mismo, y excluye las secuencias reguladoras que impulsan el inicio o la terminación de la transcripción.

El gen, la secuencia codificante o el elemento regulador puede ser uno que normalmente se encuentra en la célula, en cuyo caso se denomina 'autólogo' o 'endógeno', o puede ser uno que normalmente no se encuentra en una ubicación celular, en cuyo caso es denominado 'heterólogo', 'transgénico' o 'transgén'.

55 Un gen, una secuencia codificante o un elemento regulador 'heterólogos' también pueden ser autólogos de la célula, pero, sin embargo, están dispuestos en un orden y/u orientación o en una posición genómica o entorno que normalmente no se encuentra o no ocurre en la célula en la que se transfiere.

El término "vector" se refiere a una construcción de ADN recombinante que puede ser un plásmido, un virus, una secuencia de replicación autónoma, un cromosoma artificial, tal como el cromosoma artificial bacteriano BAC, un fago u otra secuencia de nucleótidos, en la que al menos dos secuencias de nucleótidos, al menos uno de los cuales es una molécula de ácido nucleico de la presente invención, se han unido o recombinado. Un vector puede ser lineal o circular. Un vector puede estar compuesto por un ADN o ARN monocatenario o bicatenario.

El término "expresión" se refiere a la transcripción y/o traducción de un gen endógeno o un transgén en plantas.

'Transformación', 'que transforma' y 'transferencia' se refieren a métodos para transferir moléculas de ácido nucleico, en particular ADN, a las células, incluidos, pero no se limitan a, enfoques biolísticos tales como el bombardeo de partículas, la microinyección, la permeabilización de la membrana celular con diversos tratamientos físicos por ejemplo electroporación, o tratamientos químicos, por ejemplo tratamientos con polietilenglicol o PEG; la fusión de protoplastos o *Agrobacterium tumefaciens* o transformación mediada por rizogenes. Para la inyección y electroporación de ADN en células vegetales no existen requisitos específicos para los plásmidos usados. Pueden usarse plásmidos tales como derivados de pUC. Si se van a regenerar plantas completas a partir de tales células transformadas, se prefiere el uso de un marcador seleccionable. Dependiendo del método para la introducción de los genes deseados en la célula vegetal, pueden ser necesarias más secuencias de ADN; si, por ejemplo, el plásmido Ti o Ri se usa para la transformación de la célula vegetal, al menos el borde derecho, a menudo, sin embargo, los bordes derecho e izquierdo del T-ADN del plásmido Ti y Ri tienen que estar unidos como región flanqueante a los genes que se van a introducir. Preferiblemente, las moléculas de ácido nucleico transferidas se integran de forma estable en el genoma o plasmoma de la planta receptora.

En el contexto de la presente invención, el término "actividad biológica de un inductor de haploides" o "inductor de haploides" o "línea de inductores de haploides" se refiere a una planta o línea de plantas que tiene la capacidad de producir una progenie o descendencia haploide en al menos 0.1 %, al menos 0.2 %, 0.3 %, 0.4 %, 0.5 %, 0.6 %, 0.7 %, 0.8 %, 0.9 %, preferiblemente al menos 1 %, preferiblemente al menos 2 %, preferiblemente al menos 3 %, preferiblemente al menos 4 %, preferiblemente al menos 5 %, preferiblemente al menos 6 %, preferiblemente al menos 7 %, preferiblemente al menos 8 %, preferiblemente al menos 9 %, lo más preferiblemente al menos 10 %, lo más preferiblemente al menos 15 %, lo más preferiblemente al menos 20 % de casos cuando se cruza con una planta de tipo salvaje o una planta que expresa al menos la proteína CENH3 de tipo salvaje. Dado que los cromosomas del inductor haploide se eliminan durante la meiosis, la progenie haploide resultante solo comprende los cromosomas del progenitor de tipo salvaje. Sin embargo, en caso de que el inductor haploide fuera el progenitor del óvulo del cruce, la progenie haploide posee el citoplasma del inductor y los cromosomas del progenitor de tipo salvaje.

El término "planta" según la presente invención incluye plantas enteras o partes de tal planta entera.

Las plantas enteras son preferiblemente plantas con semillas o un cultivo. Las partes de una planta son, por ejemplo, órganos/estructuras vegetativas del brote, por ejemplo, hojas, tallos y tubérculos; raíces, flores y órganos/estructuras florales, por ejemplo, brácteas, sépalos, pétalos, estambres, carpelos, anteras y óvulos; semilla, incluidos el embrión, el endospermo y la cubierta de la semilla; fruto y el ovario maduro; tejido vegetal, por ejemplo, tejido vascular, tejido triturado y similares; y células, por ejemplo, células protectoras, óvulos, tricomas y similares; y descendencia de la misma.

En cualquier caso, la planta de la presente invención comprende al menos una célula que comprende una secuencia de nucleótidos que codifica una proteína histona H3 del centrómero que comprende un dominio CATD, en la que la secuencia de nucleótidos comprende una mutación que provoca una sustitución de aminoácidos en (a) posición 92 de SEQ ID NO: 34; (b) posición 106 de SEQ ID NO: 44; o (c) posición 4 de SEQ ID NO: 1, en la que la sustitución de aminoácidos confiere la actividad biológica de un inductor de haploides. Lo más preferiblemente, la mayoría o en particular todas las células de la planta de la presente invención comprenden la mutación como se describe en este documento.

Las especies de plantas que se pueden usar en el método de la invención son preferiblemente plantas eudicotiledóneas, dicotiledóneas y monocotiledóneas.

El término "planta" en una realización preferida se refiere únicamente a una planta completa, es decir, una planta que exhibe el fenotipo completo de una planta desarrollada y capaz de reproducirse, una etapa anterior de desarrollo de la misma, por ejemplo, un embrión de planta, o ambos.

En una realización de la presente invención, el término "planta" se refiere a una parte de una planta completa, en particular material vegetal, células vegetales o cultivos de células vegetales.

El término "célula vegetal" describe la unidad estructural y fisiológica de la planta y comprende un protoplasto y una pared celular. La célula vegetal puede estar en forma de una única célula aislada, tal como células protectoras estomáticas o células cultivadas, o como parte de una unidad organizada superior tal como, por ejemplo, un tejido vegetal o un órgano vegetal.

El término "material vegetal" incluye partes de plantas, en particular células vegetales, tejidos vegetales, en particular material de propagación vegetal, preferiblemente hojas, tallos, raíces, radículas emergidas, flores o partes de flores,

pétalos, frutos, polen, tubos polínicos, filamentos de anteras, óvulos, sacos embrionarios, óvulos, ovarios, cigotos, embriones, embriones cigóticos per se, embriones somáticos, secciones de hipocótilo, meristemas apicales, haces vasculares, periciclos, semillas, raíces, esquejes, cultivos de células o tejidos, o cualquier otra parte o producto de una planta.

5 De este modo, la presente invención también proporciona material de propagación vegetal de las plantas de la presente invención. Dicho "material de propagación vegetal" se entiende que es cualquier material vegetal susceptible de ser reproducido sexual o asexualmente in vivo o in vitro. Particularmente preferidos dentro del alcance de la presente invención son protoplastos, células, callos, tejidos, órganos, semillas, embriones, polen, óvulos, cigotos, junto con cualquier otro material de propagación obtenido de plantas transgénicas. Partes de plantas, tales como por ejemplo flores, tallos, frutos, hojas, raíces que se originan en plantas mutadas o su descendencia previamente mutada, preferiblemente transformadas, mediante los métodos de la presente invención y que por lo tanto consisten al menos en parte en células mutadas, son también objeto de la presente invención.

Preferiblemente, la planta según la presente invención se selecciona del grupo que consiste en cebada (*Hordeum vulgare*), sorgo (*Sorghum bicolor*), centeno (*Secale cereale*), Triticale, caña de azúcar (*Saccharum officinarum*), maíz (*zea mays*), mijo cola de zorra (*setaria italiana*), arroz (*Oriza sativa*), arroz pequeño, arroz australiano, arroz alto, trigo (*Triticum aestivum*), trigo duro, cebada bulbosa, falso bromo púrpura (*Brachypodium distachyon*), cebada de mar (*Hordeum marinum*), hierba de cabra (*Aegilops tauschii*), manzana (*Malus domestica*), Beta vulgaris, girasol (*Helianthus annuus*), zanahoria australiana (*Daucus glochidiatus*), zanahoria silvestre americana (*Daucus pusillus*), *Daucus muricatus*, zanahoria (*Daucus carota*), eucalipto (*Eucalyptus grandis*), Erythranthe guttata, Genlisea aurea, tabaco de monte (*Nicotiana glauca*), tabaco (*Nicotiana tabacum*), Nicotiana tomentosiformis, tomate (*Solanum lycopersicum*), patata (*Solanum tuberosum*), café (*Coffea canephora*), vid (*Vitis vinifera*), pepino (*Cucumis sativus*), morera (*Morus notabilis*), berro thale (*Arabidopsis thaliana*), Arabidopsis lyrata, berro de arena (*Arabidopsis arenosa*), Crucihimalaya Himalaya, Crucihimalaya wallichii, berro ondulado (*Cardamine flexuosa*), hierba de la pimienta (*Lepidium virginicum*), bolsa de pastor (*Capsella bursa-pastoris*), Olmarabidopsis pumila, berro peludo (*Arabis hirsuta*), colza (*Brassica napus*), brócoli (*Brassica oleracea*), Brassica rapa, Brassica juncacea, mostaza negra (*Brassica nigra*), rábano (*Raphanus sativus*), Eruca vesicaria sativa, naranja (*Citrus sinensis*), Jatropha curcas, Glycine max y álamo negro (*Populus trichocarpa*).

Particularmente preferida es la planta seleccionada del grupo que consiste en cebada (*Hordeum vulgare*), sorgo (*Sorghum bicolor*), centeno (*Secale cereale*), Triticale, caña de azúcar (*Saccharum officinarum*), maíz (*Zea mays*), arroz (*Oriza sativa*), trigo (*Triticum aestivum*), Trigo duro, Avena sativa, *Hordeum bulbosum*, Beta vulgaris, girasol (*Helianthus annuus*), zanahoria (*Daucus carota*), tabaco (*Nicotiana tabacum*), tomate (*Solanum lycopersicum*), patatas (*Solanum tuberosum*), café (*Coffea canephora*), uva (*Vitis vinifera*), pepino (*Cucumis sativus*), berro de thale (*Arabidopsis thaliana*), colza (*Brassica napus*), brócoli (*Brassica oleracea*), Brassica rapa, Brassica juncacea, mostaza negra (*Brassica nigra*), rábano (*Raphanus sativus*), y Glycine max.

35 La planta según la presente invención contiene en una realización preferida la secuencia de nucleótidos que codifica el CENH3 ya sea como un gen endógeno o como un transgén.

La invención se refiere en una realización preferida a una planta según la presente enseñanza, en la que la al menos una sustitución de aminoácido se introduce en la secuencia de nucleótidos que codifica CENH3 de forma no transgénica o transgénica.

40 De este modo, preferiblemente en una realización, en la que la al menos una mutación se efectúa en el gen CENH3 endógeno, la planta obtenida no es transgénica. Preferiblemente, la mutación se efectúa mediante mutagénesis no transgénica, en particular mutagénesis química, preferiblemente mediante TILLING inducido por EMS (etilmetanosulfonato).

45 De este modo, la presente invención se refiere a una planta, en la que la introducción no transgénica de la al menos una mutación que provoca en el dominio CATD una sustitución de aminoácido que confiere la actividad biológica de un inductor haploide se efectúa mediante mutagénesis química, en particular mediante TILLING.

En otra realización preferida, la al menos una mutación se introduce en la planta en forma de transgén. Preferiblemente, esto se hace transformando un vector que comprende una secuencia de nucleótidos que codifica al menos el dominio CATD de CENH3 que comprende al menos una sustitución de aminoácido como se describe en este documento. Los métodos para la transformación de una planta y la introducción de un transgén en el genoma de una planta son bien conocidos en la técnica anterior.

55 De este modo, en una realización se proporciona una planta, en la que la introducción transgénica de la sustitución de aminoácidos o sustituciones de aminoácidos dentro de la proteína CENH3 se efectúa mediante la transformación de un vector que comprende una secuencia de nucleótidos que codifica al menos la hélice $\alpha 2$ que se encuentra posicionada en el dominio CATD y correspondiente a los nucleótidos desde la posición 379 a la posición 465 de la proteína CENH3 como se establece en SEQ ID No. 38 derivada de *Arabidopsis thaliana* pero que comprende al menos una sustitución de aminoácido en la posición 4 de SEQ ID NO: 1, en la que la sustitución de aminoácidos confiere la actividad biológica de un inductor de haploides. En otra realización, se proporciona una planta, en la que la introducción

de la sustitución de aminoácidos o sustituciones de aminoácidos dentro de la proteína CENH3 se efectúa mediante la transformación de un vector que comprende una secuencia de nucleótidos que codifica al menos el dominio CATD o una proteína CENH3 que comprende el dominio CATD que comprende al menos una sustitución de aminoácido en la posición 4 de SEQ ID NO: 1, en la que la sustitución de aminoácidos confiere la actividad biológica de un inductor de haploides.

Preferiblemente, para la transformación se usan la transformación mediada por *Agrobacterium*, el método de inmersión floral o el bombardeo de partículas.

En la realización preferida, en la que la secuencia de nucleótidos que codifica la proteína CENH3 mutada según la presente invención se transforma en la planta en forma de un transgén y uno o dos alelos del gen CENH3 endógeno se inactivan o bloquean preferiblemente. Otra realización preferida, en la que la secuencia de nucleótidos que codifica la proteína CENH3 mutada según la presente invención se transforma en la planta en forma de un transgén y el transgén se sobreexpresa para que sea más competitivo como proteína CENH3 endógena y se prefiera durante la generación de un complejo cinetocoro.

En una realización particularmente preferida, el aminoácido leucina en la posición 4 de la hélice $\alpha 2$ está sustituido en la planta según la presente invención. En particular, el aminoácido leucina en la posición 4 se sustituye preferiblemente por isoleucina o fenilalanina.

La presente invención también proporciona una planta obtenible, en particular obtenida, mediante un método según la presente invención y que se caracteriza por tener la actividad biológica de un inductor de haploides.

En una realización preferida de la presente invención, el método de producción de la planta que tiene actividad biológica de un inductor de haploides según la presente invención no es un método esencialmente biológico.

Además, se proporciona un método para generar una planta que tiene actividad biológica de un inductor de haploides según la presente invención, que comprende las etapas de:

i) someter las semillas de una planta a una cantidad suficiente del mutágeno etilmetano sulfonato (EMS) para obtener plantas M1,

ii) permitir una producción suficiente de plantas M2 fértiles,

iii) aislar ADN genómico de plantas M2 y

iv) seleccionar individuos que posean al menos una sustitución de aminoácido en el dominio CATD de CENH3.

La presente invención se relaciona además en una realización preferida con un método para generar una planta que tiene actividad biológica de un inductor de haploides según la presente invención, que comprende las etapas de:

a) transformar una célula vegetal con una secuencia de nucleótidos que codifica al menos el dominio CATD de una proteína CENH3 que comprende una mutación que provoca una sustitución de aminoácidos en (a) posición 92 de SEQ ID NO: 34; (b) posición 106 de SEQ ID NO: 44; o (c) posición 4 de SEQ ID NO: 1 o un vector que comprende una secuencia de nucleótidos que codifica al menos el dominio CATD de una proteína CENH3 que comprende una mutación que causa una sustitución de aminoácidos en (a) posición 92 de SEQ ID NO: 34; (b) posición 106 de SEQ ID NO: 44; o (c) posición 4 de SEQ ID NO: 1, y

b) regenerar una planta que tiene la actividad biológica de un inductor haploide a partir de la célula vegetal.

Se divulga que la planta haploide identificada puede convertirse en una planta haploide doble, preferiblemente mediante tratamiento con colchicina. En particular, los presentes métodos no se basan únicamente en, en particular, no consisten en fenómenos naturales tales como el cruce o la selección, sino que, de hecho, se basan esencialmente en la enseñanza técnica para proporcionar una secuencia de nucleótidos específicamente mutada preparada por la contribución de la humanidad. De este modo, la presente invención introduce una característica estructural específica, a saber, una mutación, en una secuencia de nucleótidos y una planta de la presente divulgación, cuya mutación no está provocada ni asociada con ningún fenómeno natural tal como el cruzamiento o la selección.

Además, se divulga que la planta según la presente invención también se puede usar en un método para restaurar la fertilidad masculina proporcionando un citoplasma normal a un socio de cruzamiento que es CMS. A través de dicho cruce, los cromosomas de la planta CMS se introducen en el citoplasma normal del inductor haploide de la presente invención que no es CMS. Sin embargo, la producción de polen de la planta CMS debe inducirse a través de la temperatura, la luz, la duración del día, etc.

Sin estar ligado a la teoría, en la figura se ofrece un posible modelo de cómo funcionan los presentes métodos, en particular un método de eliminación uniparental de cromosomas, funciona en el *inductor CENH3* × embriones híbridos interespecíficos *CENH3 de tipo salvaje*. (A) Los óvulos probablemente derivados del inductor haploide contienen menos CENH3 o, en comparación con el tipo salvaje, tienen una 'firma requerida de transgeneración de CENH3' desconocida. Una cantidad reducida de CENH3 materno es menos probable según los estudios realizados con un

reportero CENH3-GFP en *A. thaliana* los núcleos de los espermatozoides de las plantas, pero no los óvulos, están marcados por CENH3. Sin embargo, todavía es posible que los CENH3 maternos residuales, que generan una "impresión centromérica", se transmitan a la descendencia. (B) Dentro de unas pocas horas después de la fertilización, también el CENH3 de tipo salvaje paterno se elimina activamente del núcleo del cigoto, y (C) la recarga centromérica de CENH3-GFP en el cigoto ocurre en la etapa de 16 núcleos del desarrollo del endospermo en *A. thaliana*. (D) En los embriones que experimentan haploidización, la recarga centromérica de los cromosomas maternos está alterada o retrasada, lo que causa retrasos en los cromosomas debido a la inactividad del centrómero durante la anafase. Posteriormente, los cromosomas del *inductor haploide* micronucleado se degradarán y (E) se desarrollará un embrión haploide. Los embriones haploides contienen cromosomas derivados del padre en el fondo del citoplasma derivado de la madre.

La presente invención también se refiere a una secuencia de nucleótidos que codifica al menos el dominio CATD de una proteína CENH3 que comprende una mutación que provoca una sustitución de aminoácidos en (a) posición 92 de SEQ ID NO: 34; (b) posición 106 de SEQ ID NO: 44; o (c) posición 4 de SEQ ID NO: 1.

La presente invención también se refiere a un vector, en particular vector viral, construcción o plásmido que comprende dicha secuencia de nucleótidos y, si está presente, secuencias asociadas, preferiblemente como se indica en este documento.

En una realización particularmente preferida de la presente invención, la secuencia de nucleótidos que codifica al menos el dominio CATD de una proteína CENH3 comprende preferiblemente al menos la región codificante completa de CENH3, en particular el gen de CENH3.

En una realización más preferida de la presente invención, la secuencia codificante de CENH3 puede asociarse con elementos reguladores, tales como elementos reguladores 5' y/o 3', más preferiblemente con un promotor, preferiblemente un promotor constitutivo o inducible.

Además, la presente invención proporciona una célula vegetal que comprende dicha secuencia de nucleótidos o un vector que la comprende como un transgén.

En el contexto de la presente invención, el término "que comprende" como se usa en este documento tiene el significado de "que incluye" o "que contiene", lo que significa que, además del elemento mencionado explícitamente, posiblemente estén presentes otros elementos.

En una realización preferida de la presente invención, el término "que comprende", como se usa en este documento, también se entiende que significa "que consiste en excluir de ese modo la presencia de otros elementos además del elemento mencionado explícitamente".

En una realización preferida adicional, el término "que comprende", tal como se usa en este documento, también se entiende que significa "que consiste esencialmente en excluir de ese modo la presencia de otros elementos que proporcionan una contribución significativa a la enseñanza divulgada además del elemento mencionado explícitamente".

Otras realizaciones preferidas de la presente invención son el objeto de las reivindicaciones subordinadas.

La invención se describirá ahora con algo más de detalle por medio de ejemplos no limitativos y una figura.

El protocolo de secuencia muestra:

SEQ ID No. 1: la secuencia consenso de aminoácidos de la hélice $\alpha 2$ de CENH3,

SEQ ID Nos. 2 a 32: secuencias de nucleótidos de cebadores usados en la presente enseñanza,

SEQ ID No. 33: la secuencia de nucleótidos de ADNc de la proteína β -CENH3 de tipo salvaje de *Hordeum vulgare*,

SEQ ID No. 34: la secuencia de aminoácidos de β -CENH3 de *Hordeum vulgare*,

SEQ ID No. 35: la secuencia de ADNc del β -CENH3 de *Hordeum vulgare* (mutante de la línea 4528 de TILLING),

SEQ ID No. 36: la secuencia de aminoácidos de β -CENH3 de *Hordeum vulgare* (mutante de la línea 4528 de TILLING),

SEQ ID No. 37: la secuencia de nucleótidos de la secuencia codificante de tipo salvaje (ADNc) de *A. thaliana* CENH3,

SEQ ID No. 38: la secuencia de aminoácidos del tipo salvaje *A. thaliana* CENH3,

SEQ ID No. 39: la secuencia de nucleótidos de la secuencia codificante (ADNc) de la *A. thaliana* CENH3 mutada (mutante L a I),

SEQ ID No. 40: la secuencia de aminoácidos del *A. thaliana* CENH3 mutado (mutante L a I),

SEQ ID No. 41: la secuencia de nucleótidos de la secuencia codificante (ADNc) de la *A. thaliana* CENH3 mutada (mutante L a F),

SEQ ID No. 42: la secuencia de aminoácidos del *A. thaliana* CENH3 mutado (mutante L a F),

5 SEQ ID No. 43: la secuencia de nucleótidos de la secuencia codificante de tipo salvaje (ADNc) de *Beta vulgaris* CENH3,

SEQ ID No. 44: la secuencia de aminoácidos del tipo salvaje *Beta vulgaris* CENH3,

SEQ ID No. 45: la secuencia de nucleótidos de la secuencia codificante (ADNc) de *Beta vulgaris* CENH3 (mutante L a F),

SEQ ID No. 46: la secuencia de aminoácidos de *Beta vulgaris* CENH3 mutado (mutante L a F),

10 SEQ ID No. 47: la secuencia de nucleótidos de la secuencia codificante (ADNc) de *Beta vulgaris* CENH3 (mutante L a I) y

SEQ ID No. 48: la secuencia de aminoácidos del *Beta vulgaris* CENH3 mutado (mutante L a I).

La figura muestra esquemáticamente un modelo mecánico relacionado con los métodos de la presente invención.

Ejemplos

15 Ejemplo 1: Mutagénesis de la cebada α y β CENH3 mediante lesiones locales dirigidas EN genomas (TILLING)

Identificar si una mutación de un solo punto de CENH3 endógeno podría dar como resultado un inductor haploide y una población de cebada diploide inducida por TILLING. (*Hordeum vulgare*) (Gottwald et al., 2 (2009), BMC Res Notes, 258), una especie que codifica dos variantes funcionales de CENH3 (α y β CENH3) (Sanei et al., 108 (2011), Proc Natl Acad Sci USA, E498-505) fue cribada. Suponiendo que la complementación de cualquiera de las variantes de CENH3, una mutación funcional de α CENH3 o β CENH3 aún permitiría la generación de descendencia.

20 Para ello, fue cribada una población de TILLING de 10.279 plantas de cebada diploide tratada con EMS (*Hordeum vulgare*) del cv. *Barke* para identificar los alelos mutantes de α y β CENH3. Combinaciones de cuatro y tres cebadores

Hv_aCENH3_EX1+2+3_F: AGGCAGGGTCTCAATTCCTT (SEQ ID No. 2),

Hv_aCENH3_EX1+2+3_R: GTCCATCATCCATCGTCTT (SEQ ID No. 3),

25 Hv_aCENH3_EX4+5_F: CCCACTTCCTTGTTGTGGAC (SEQ ID No. 4), Hv_aCENH3_EX4+5_R:

GGCGATAAATGTATCTTGCATTC (SEQ ID No. 5), Hv_aCENH3 EX6 F:

TGGTAGCAACCAGAGCTACG (SEQ ID No. 6), Hv_aCENH3_EX6_R:

ACTGGCATGTTTCTTCTGC (SEQ ID No. 7), Hv_aCENH3 EX7 F:

CGGACGGAGGGAGTATTTCT (SEQ ID No. 8), Hv_aCENH3_EX7 R:

30 GGACATGCCCAAAGAAAGTG (SEQ ID No. 9), Hv_bCENH3_EX1+2_F:

GCCAGCGAGTACTCCTACAAG (SEQ ID No. 10), Hv_bCENH3_EX1 R:

TTGAGTTACCAGCCACCACTC (SEQ ID No. 11), Hv_bCENH3_EX3 F:

GTCATGCACTGTGTCTTGCA (SEQ ID No. 12), Hv_bCENH3EX3 R:

TGCTAAGATCGGATAACTGTGG (SEQ ID No. 13), Hv_bCENH3_EX4 F:

35 TGCTCCTGAACAACTGAACC (SEQ ID No. 14), Hv_bCENH3_EX4 R:

GTGGCCGTCAGTACAATCG (SEQ ID No. 15)

se usaron para amplificar todos los exones de las variantes de CENH3 α y β y partes de los intrones correspondientes, respectivamente, mediante PCR con un paso heterodúplex como se describió anteriormente (Gottwald et al., (2009), BMC Res Notes, 258). Los productos de PCR se digirieron con kit de escisión de ADNds y se analizaron con Mutation Discovery Kit y Gel - kit de reactivos ADNds en el sistema AdvanCETM FS96 según las pautas del fabricante (Advanced Analytical, IA, EE. UU.).

40 Extracción de ARN, PCR y RT-PCR cuantitativa en tiempo real

El ARN total se aisló a partir de raíces, hojas usando el método Trizol (Chomczynski and Sacchi, 162 (1987), Anal Biochem, 156-159) de anteras (microscópicamente escalonada entre la meiosis y el desarrollo del polen maduro), carpelo, endospermo y embrión mediante el kit de aislamiento de ARN Picopure (Arcturus) según el fabricante. La ausencia de contaminación por ADN genómico se confirmó mediante PCR usando cebadores GAPDH (véase la tabla 2). 10 µl de mezcla de PCR contenían 1 µl de la plantilla de ADNc, 5 µl de 2x Power SYBR Green PCR Master Mix (Applied Biosystems), 0.33 mM de los cebadores directo e inverso para cada gen (véase la tabla 2). Las reacciones se realizaron en un sistema de PCR en tiempo real rápido Applied Biosystems 7900HT. La PCR se realizó usando las siguientes condiciones: 95 °C, durante 10 min, seguido de 40 ciclos a 95 °C, durante 15 s, a la temperatura de hibridación de 60 °C, durante 60 s. Se realizaron tres réplicas técnicas para cada muestra de ADNc. El sistema PCR rápido en tiempo real y los datos se analizaron con el software SDS v 2.2.2. Los niveles de transcripción de cada gen se normalizaron a GAPDH mediante la siguiente fórmula: $R = 2^{-\Delta(\text{Ct}_{\text{GOI}} - \text{Ct}_{\text{H}})} \times 100$, donde R = cambios relativos, GOI = gen de interés y H = constitutivo (GAPDH). La especificidad y la eficacia de ambos cebadores se determinaron mediante qRT-PCR usando una serie de diluciones de plásmidos de genes α y β CENH3 de cebada de longitud completa clonados. Un valor Ct similar (el ciclo de PCR en el que la señal fluorescente del colorante indicador supera el nivel de fondo) para la misma cantidad de plásmido indica que ambos cebadores pueden amplificar transcritos específicos con la misma eficacia.

Tabla 2

Nombre del cebador	Secuencia (5' a 3')
GAPDH-F	CAATGATAGCTGCACCACCAACTG (SEQ ID No. 21)
GAPDH-R	CTAGCTGCCCTTCCACCCTCTCCA (SEQ ID No. 22)
Hv α -F	AGTCGGTCAATTTTCTCATCCC (SEQ ID No. 23)
Hv α -R	CTCTGTAGCCTCTTGAAGTGC (SEQ ID No. 24)
Hv β -F	GCCATTGTGGAACAAGAAGG (SEQ ID No. 25)
Hv β -R	TAACACGGTGCGAATGAATG (SEQ ID No. 26)
CH3A L130_F_for	phos-GACAGCTGAAGCATTGTTGCTCTTC (SEQ ID No. 27)
CENH3L130_I_for	phos-GACAGCTGAAGCTATTGTTGCTCTTC (SEQ ID No. 28)
CENH3L130_F I_rev	phos-CAACGATTGATTTGGGGAGGG (SEQ ID No. 29)
cenh3-1_mut_for	GGTGCGATTTCTCCAGCAGTAAAATC (SEQ ID No. 30)
cenh3-1_mut_rev	CTGAGAAGATGAAGCACCGCGATAT (SEQ ID No. 31)
cenh3-1_mut2429r	AACTTTTGCCATCCTCGTTTTCTGTT (SEQ ID No. 32)

Solo se identificaron mutaciones puntuales sin sentido invertidas para ambas variantes de cebada CENH3.

La falta de funcionalidad de los CENH3 mutados de mutantes M2 homocigóticos se analizó mediante inmunotinción de los centrómeros con anticuerpos específicos de la variante CENH3. Los cromosomas mitóticos y meióticos de la línea TILLING 4528 de *H. vulgare* de tipo salvaje y homocigota (planta según la presente invención) se sometieron a inmunotinción con anticuerpos específicos para α CENH3 y β CENH3. Las señales de α CENH3 y β CENH3 en los centrómeros se revelaron en todos los mutantes, mientras que solo la línea homocigota TILLING 4528 que contiene una sustitución de leucina por fenilalanina en el aminoácido 92 (SEQ ID No. 36), es decir, el aminoácido en la posición 4 de la secuencia consenso SEQ ID No 1, no mostró señales centroméricas de β CENH3 en células mitóticas, meióticas y en interfase. La sustitución de leucina por fenilalanina en el aminoácido 92 de SEQ ID No. 36 de la secuencia de *H. vulgare* β -CENH3 corresponde a una sustitución de un solo nucleótido de C a T en la posición 274 de la secuencia de ADNc de *H. vulgare* β -CENH3 (SEQ ID NO: 35).

En esta línea, solo se encontraron señales débiles de β CENH3 dispersas fuera de los centrómeros. Se han medido los niveles de transcripción de α y β CENH3 en tipo salvaje (cv Barke) (SEQ ID no. 33 y 34) y la línea TILLING 4528 con β CENH3 mutado. El nivel de expresión relativo de α y β CENH3 se midió en diferentes tejidos usando cebadores específicos (Tabla 2). Se preparó ADNc a partir de ARN total y los niveles de expresión génica se normalizaron al nivel de expresión de gliceril fosfato deshidrogenasa (GRPTA). Obviamente, la carga centromérica de la variante β CENH3 mutada parece estar alterada, mientras que no se encontró un nivel de transcripción diferente entre el tipo salvaje y el β CENH3 mutado. Por consiguiente, los centrómeros compuestos exclusivamente de α CENH3 son suficientes para la función del centrómero mitótico en la cebada, ya que no se encontraron defectos obvios de segregación cromosómica, tales como puentes de anafase, así como cambios de ploidía o valores de ciclo. Además, no se observaron cambios evidentes en el habitus de la planta en las plantas mutantes. En particular, no se pudieron detectar diferencias significativas en el fenotipo, los niveles de ploidía, los valores del ciclo y el fenotipo de crecimiento entre las plantas homocigóticas de la línea TILLING 4528 y la cebada de tipo salvaje.

Se abordó el problema de si la falta de β CENH3 se compensa con α CENH3 adicional para mantener la función del centrómero en el mutante. Por lo tanto, las señales de inmunotinción de α CENH3 de tipo salvaje (126 centrómeros medidos) y de la línea 4528 (56 centrómeros medidos) se cuantificaron comparativamente mediante mediciones de intensidad de píxeles. Un aumento del 19.8 % de α CENH3 en el mutante indica que el β CENH3 faltante se compensa parcialmente con α CENH3 incorporado adicionalmente. La mutación β CENH3 está ubicada en un dominio diana evolutivamente altamente conservado (CATD) definido por partes de la hélice α 1, bucle 1 y la hélice α 2 del pliegue de la histona. Este dominio es necesario para la carga del centrómero de CENH3 por acompañantes.

20 Inmunotinción indirecta

La inmunotinción indirecta de núcleos y cromosomas se llevó a cabo como se describió anteriormente (Sanei et al., 108 (2011), Proc Natl Acad Sci USA, E498-505). Se detectó CENH3 de cebada con anticuerpos anti- α CENH3 específicos de cobaya y anticuerpos específicos de conejo anti- β CENH3. Se usó un anticuerpo específico de HTR12 de conejo (abcam, ab72001) para la detección de *A. thaliana* CENH3 (AtCENH3). Las señales de epifluorescencia se registraron con una cámara CCD enfriada (ORCA-ER, Hamamatsu). La obtención de imágenes se realizó usando un microscopio Olympus BX61 y una cámara CCD ORCA-ER (Hamamatsu). Para analizar las estructuras de inmunoseñales y cromatina a una resolución óptica de -100 nm (super-resolución) se aplicó Microscopía de Iluminación Estructurada (SIM) usando un objetivo Korr C-Apo 63x/1.2W de un sistema de microscopio Elyra y el software ZEN (Zeiss, Alemania). Las imágenes se capturaron por separado para cada fluorocromo usando filtros de excitación y emisión apropiados. Las imágenes se fusionaron usando el software Adobe Photoshop 6.0. Para determinar la cantidad de α y β CENH3 en los núcleos, se midieron las intensidades de fluorescencia usando el software TINA 2.0 en proyecciones de máxima intensidad generadas a partir de pilas de secciones SIM ópticas a través de las muestras mediante el software ZEN. Se estableció un umbral de intensidad para restar computacionalmente los píxeles de fondo de la imagen. La suma corregida de los valores grises de todas las señales dentro del núcleo se usó para determinar el contenido de CENH3. La representación 3D basada en pilas de imágenes SIM se realizó usando el software ZEN.

Ejemplo 2: *Arabidopsis thaliana*

Para probar si la mutación en el dominio CATD causó la carga alterada del centrómero observada, eYFP se fusionó en el extremo N con la secuencia codificante (CDS) de *A. thaliana* CENH3 (SEQ ID No. 37, proteína: SEQ ID No. 38) con un L/I (leucina/isoleucina) (CDS: SEQ ID No. 39, proteína SEQ ID No. 40) o L/F (leucina/fenilalanina) (CDS: SEQ ID No. 41, proteína SEQ ID No. 42) intercambio de las posiciones correspondientes (L130I o L130F, correspondiente a la posición de aminoácido 92 en β CENH3 de cebada, es decir, la posición de aminoácido 4 de la secuencia consenso SEQ ID No. 1) en *A. thaliana* CENH3. La sustitución de leucina por isoleucina en la posición 130 de *A. thaliana* corresponde a una única sustitución de nucleótido de C por A en la posición 388 de SEQ ID no. 37.

La sustitución de aminoácidos de leucina a fenilalanina en la posición 130 está provocada por dos sustituciones de nucleótidos, a saber, TC por AT en las posiciones 387 y 388 de SEQ ID No. 37.

Doble marcaje de transgénicos *A. thaliana* con el CENH3 de tipo antisalvaje correspondiente y el anti-GFP revelaron una diana significativamente reducida en el centrómero de los CENH3 mutadas.

A continuación, para probar la capacidad inductora de haploides, se usaron construcciones genómicas de CENH3 de *A. thaliana* con un intercambio L130I o L130F para transformar plantas heterocigotas de *A. thaliana* con CENH3 bloqueado (Ravi and Chan, Nature, 464 (2010), 615-618). El genotipado identificó mutantes nulos homocigóticos de CENH3 que se complementaron con construcciones genómicas de CENH3 de tipo salvaje, L130I o L130F. Como se obtuvieron plantas diploides viables que contenían cualquiera de las construcciones, es probable que esta mutación no perjudique la función del centrómero en plantas de *A. thaliana* homocigóticas. Cuando los mutantes nulos de CENH3 que expresan una proteína L130F CENH3 con mutación puntual se cruzaron con el tipo salvaje, los cromosomas del mutante se eliminaron, produciendo una progenie haploide. El análisis de citometría de flujo reveló que el 10.7 % de las plantas F1 eran haploides.

Clonación y generación de transgenes *CENH3*

Para generar fragmentos genómicos de *CENH3* portadores de mutaciones, lo que resulta en fenilalanina 130 (F130) e isoleucina 130 (I130) en lugar de leucina 130 de tipo salvaje (L130), un fragmento genómico de *CENH3* en vector pCAMBIA1300 usado para complementar *cenh3-1/cenh3-1* (mutante nulo de *cenh3*) (Ravi and Chan, Nature, 464 (2010), 615-618; Ravi et al., Genetics, 186 (2010), 461-471) fue subclonado a través de los sitios HindIII y BamHI únicos en pBlueScript II KS (Stratagene, www.stratagene.com). Las mutaciones de *CENH3*, L130I o L130F, se generaron en pBlueScript II KS usando un Kit de mutagénesis dirigida al sitio Phusion® (Finnzymes, www.finnzymes.com) según las instrucciones del fabricante con cambios menores como se describe. Se usaron los siguientes cebadores fosforilados en 5' para la mutagénesis: CH3A+L130_F_for, CENH3L130_I_for y CENH3L130_F I_rev. Los fragmentos genómicos de *CENH3* mutados fueron subclonados a través de los sitios HindIII y BamHI únicos en el pCAMBIA1300 inicial que contiene un marcador de resistencia a la higromicina. Todas las construcciones fueron verificadas por secuenciación. Para los cebadores, consulte la tabla 2 anterior.

Para generar construcciones de fusión *p35S::eYFP-CENH3* que contienen mutaciones dentro de la *CENH3* CDS, dando como resultado L130I o L130F, un plásmido (p35S-BM; Schmidt, www.dna-cloning.com) que contiene una expresión de *p35S::eYFP-CENH3* (Lermontova, 18 (2006), Plant Cell, 615-618) se usó como plantilla para el Phusion® Kit de mutagénesis dirigida al sitio (Finnzymes, www.finnzymes.com). Los cebadores y las estrategias para introducir las mutaciones deseadas fueron los mismos que los anteriores. Casetes de expresión resultantes (35Spro, eYFP-(mutado)*CENH3* y NOS terminador) fueron subclonados a través de un único sitio de restricción *SfiI* en el vector pLH7000 que contiene un marcador de resistencia a fosfinotricina (Schmidt, www.dna-cloning.com) y verificado por secuenciación.

Transformación de plantas, condiciones de cultivo y polinización cruzada.

A. thaliana de tipo salvaje (SEQ ID No. 37 y 38) y plantas heterocigotas *cenh3-1/CENH3* (ambas de la accesión Columbia-0) fueron transformadas por el método de inmersión floral (Clough and Bent, 16 (1998), Plant J, 735-743). Las progenies transgénicas se seleccionaron en medio sólido de Murashige y Skoog que contenía el antibiótico correspondiente. Las plantas se germinaron en placas de Petri en condiciones de día largo (20 °C 16 h de luz/18 °C 8 h de oscuridad), se cultivaron durante 4 semanas en condiciones de día corto (20 °C 8 h de luz/18 °C 16 h de oscuridad) y luego se cambiaron a condiciones de día largo de nuevo. Para cruzar, los brotes cerrados de mutantes *cenh3 A. thaliana* fueron emasculados extrayendo las anteras inmaduras con la ayuda de fórceps. Los estigmas de los brotes emasculados se fertilizaron con el polen amarillento a partir de las anteras maduras de flores recién abiertas de *A. thaliana* de tipo salvaje

Extracción de ADN y genotipado de *A. thaliana*

Las preparaciones de ADN genómico y el genotipado basado en PCR se realizaron usando métodos estándar. El ADN se extrajo según Edwards et al. (1991), Nucleic Acids Res 19, 1349.

Las plantas fueron genotipadas para *cenh3-1* en una reacción de genotipado dCAPS. Los cebadores dCAPS, *cenh3-1_mut_for* y *cenh3-1_mut_rev*, se usaron para amplificar *CENH3*. Los amplicones se digirieron con EcoRV y se resolvieron en un gel. El alelo mutante *cenh3-1* no se corta (215 pb) mientras que el alelo WT *CENH3* se corta (191 y 24 pb). Para los cebadores, consulte la tabla 2. Para genotipar el locus de *CENH3* endógeno para *cenh3-1* en la descendencia de plantas *cenh3-1/CENH3* transformadas con el locus genómico *CENH3* (transgén *CENH3* sin marcar con L130, L130I o L130F), se realizó una reacción de PCR inicial con un cebador fuera del locus del transgén *CENH3*, lo que permite la amplificación específica de locus *CENH3* endógena y no la transgénica. Los cebadores usados fueron *cenh3-1_mut_for* y *cenh3-1_mut2320r/cenh3-1_mut2429r*. Los amplicones se purificaron y se usaron como plantilla para una segunda reacción de genotipado de dCAPS PCR como se describió anteriormente para plantas *cenh3-1*. Para los cebadores, consulte la tabla 2. La presencia de transgén se verificó por PCR.

Análisis de citometría de flujo de plantas y semillas.

Para los análisis de ploidía de citometría de flujo de plantas, se cortaron simultáneamente cantidades iguales de material foliar de 5 a 10 individuos en solución reguladora de aislamiento de núcleos (Galbraith et al. (1983), Science 220, 1049-1051) suplementada con RNasa libre de DNasa (50 µg/ml) y yoduro de propidio (50 µg/ml) con una hoja de afeitar afilada. Las suspensiones de núcleos se filtraron a través de un filtro de células de 35 µm en tubos de poliestireno de 5 ml (BD Biosciences) y se midieron en un clasificador de células FACStar^{PLUS} (BD Biosciences) equipado con un láser de iones de argón INNOVA 90C (Coherent). Se midieron y analizaron aproximadamente 10.000 núcleos usando el software CELL Quest ver. 3.3 (BD Biosciences). Los histogramas resultantes se compararon con un histograma de referencia que representaba una planta diploide de tipo salvaje. En los casos en los que se detectó un pico adicional en la posición haploide, las plantas se midieron individualmente de nuevo para identificar a los individuos haploides.

El aislamiento de núcleos de las semillas se realizó como se describe anteriormente usando la solución reguladora de aislamiento de núcleos MA VI (Tris-HCl 100 mM, MgCl₂ 5.3 mM, NaCl 86 mM, citrato de sodio 30.6 mM, Triton X-100 1.45 mM, pH 7.0; suplementado con 50 µg/ml de RNasa libre de DNasa y 50 µg/ml de yoduro de propidio). Las suspensiones de núcleos se midieron en un clasificador de células FACSAria (BD Biosciences) y se analizaron usando el software FACS Diva ver. 6.1.3 (BD Biosciences). De manera similar a lo anterior, las primeras 10 a 20 semillas se

agruparon para identificar líneas con embriones haploides y, en una segunda etapa, semillas individuales se cortaron junto con material foliar de *Raphanus sativus* (Genebank Gatersleben, número de acceso: RA 34) como referencia interna para confirmar la ocurrencia de semillas haploides.

Ejemplo 3: *Beta vulgaris*

- 5 Además, la importancia funcional de la mutación identificada también se analizó en la remolacha azucarera *Beta vulgaris*. Construcciones de indicadores RFP que contienen el ADNc de *Beta vulgaris* CENH3 (SEQ ID No. 43, proteína SEQ ID No. 44) con un L106F (SEQ ID No. 45, proteína SEQ ID No. 46) o L106I (SEQ ID No. 47, proteína SEQ ID No. 48) de intercambio (correspondiente a la posición de aminoácido 92 de la cebada, la posición de aminoácido 4 de la secuencia de consenso SEQ ID No. 1) se generaron y usaron para la transformación estable de la remolacha azucarera y se detectó un centrómero reducido dirigido a CENH3s mutada.

La sustitución de aminoácidos de leucina a fenilalanina en la posición 106 está provocada por dos sustituciones de nucleótidos, a saber, C por T en la posición 316 y G por T en la posición 318 de SEQ ID no. 43.

La sustitución de aminoácidos de leucina a isoleucina en la posición 106 está provocada por dos sustituciones de nucleótidos, a saber, C por A en la posición 316 y G por T en la posición 318 de SEQ ID no. 43.

- 15 Transformación de plantas y condiciones de cultivo.

Hojas de *Beta vulgaris* de tipo salvaje de plantas de 6 - 8 semanas de edad (cultivadas en condiciones de invernadero semicontroladas) se transformaron transitoriamente mediante bombardeo de partículas (300 µg de oro recubierto con 0.5 µg de ADN plasmídico). Las hojas bombardeadas se incubaron durante 48 - 72 h (25 °C, 16 h de luz (350 µmolm⁻²s⁻¹) / 8h oscuridad) antes del análisis microscópico. La transformación estable de callo de *B. vulgaris* se realizó como se describe en Lindsey & Gallois, 1990 (Journal of experimental botany, 41(5), 529-536) (selección vía kanamicina). Después de aprox. 2 meses (24 °C, 16 h luz (55 µmolm⁻²s⁻¹)/8 h de oscuridad) se analizaron microscópicamente las células del callo.

Clonación y generación de transgenes *CENH3*

- 25 Para generar construcción de fusión *35S::RFP-CENH3*, CENH3 se amplificó a partir de ADNc de remolacha azucarera con los siguientes cebadores:

BvCENH3-cds1: GGATCCATGAGAGTTAAACACACTGC (SEQ ID No. 16), BvCENH3-cds2: GGATCCTGTTTCAGTTACCATCCCCTC (SEQ ID No. 17),

clonados en un vector que contiene un casete de expresión de *35Pro*, *RFP* y *terminador 35S*. Para construcciones que contienen mutaciones dentro de la secuencia codificante *CENH3*, lo que da como resultado F106 e I1106 en lugar de L106, el plásmido mencionado anteriormente que contiene el casete de expresión *35S::RFP-CENH3* se usó como plantilla para la mutagénesis basada en cebadores. El sitio *PstI* cercano a la posición de la mutación deseada para dividir CENH3 en dos partes. Mediante mutaciones en los cebadores se integraron las mutaciones deseadas:

BvCENH3_mut_Fw: ATGGATCCATGAGAGTTAAACACACTGC (SEQ ID No. 18),

BvCENH3_L->F_Re: CTCTGCAGCCTCTTGAAGGGCCATAAAAGC (SEQ ID No. 19),

35 BvCENH3_L->I_Re: CTCTGCAGCCTCTTGAAGGGCCATAATAGC (SEQ ID No. 20).

Los casetes de expresión resultantes (*35Spro*, *RFP*-(mutado)*CENH3* y *35S-terminador*) fueron verificados por secuenciación.

Análisis de unión *CENH3* en *B. vulgaris*

40 Para analizar la unión de CENH3 y el CENH3 mutado, se analizó el material de hoja o callo usando un objetivo C-Apo 63x/1.2W Korr de un sistema de microscopio Axio Imager M2 y el software ZEN (Zeiss, Alemania).

REIVINDICACIONES

- 5 1. Planta que tiene actividad biológica de un inductor haploide y que comprende una secuencia de nucleótidos que codifica una proteína centrómero histona H3 (CENH3) que comprende un dominio CATD, en la que la secuencia de nucleótidos comprende una mutación que causa una sustitución de aminoácidos en (a) posición 92 de SEQ ID NO: 34; (b) posición 106 de SEQ ID NO: 44; o (c) posición 4 de SEQ ID NO: 1, en la que la sustitución de aminoácidos confiere la actividad biológica de un inductor de haploides.
2. Planta según la reivindicación 1, en la que la mutación provoca una sustitución de un aminoácido especificado de SEQ ID NO: 1 en la que el aminoácido especificado es L en (a) posición 4 de SEQ ID NO: 1; (b) posición 92 de SEQ ID NO: 34; o (c) posición 106 de SEQ ID NO: 44.
- 10 3. Planta según cualquiera de las reivindicaciones anteriores, en la que el cruce entre la planta y una planta de tipo salvaje o una planta que expresa la proteína CENH3 de tipo salvaje produce al menos un 0.1 % de progenie haploide.
4. Planta según cualquiera de las reivindicaciones anteriores, en la que la secuencia de nucleótidos que comprende la mutación es un gen endógeno o un transgén.
- 15 5. Planta según la reivindicación 4, en la que la secuencia de nucleótidos que comprende la mutación es un transgén y al menos un gen endógeno que codifica una proteína CENH3 está inactivado o bloqueado.
6. Planta según cualquiera de las reivindicaciones anteriores, en la que el aminoácido leucina en (a) posición 4 de SEQ ID NO: 1, (b) posición 92 de SEQ ID NO: 34; o (c) posición 106 de SEQ ID NO: 44 está sustituido, preferiblemente sustituido por isoleucina o fenilalanina.
- 20 7. Parte de la planta según cualquiera de las reivindicaciones anteriores, que es preferiblemente un órgano vegetativo del brote, raíz, flor u órgano floral, semilla, fruto, óvulo, embrión, tejido vegetal o célula, en la que la parte de la planta comprende una secuencia de nucleótidos como se define en la reivindicación 1.
8. Secuencia de nucleótidos que codifica al menos el dominio CATD de una proteína CENH3 que comprende una mutación que provoca una sustitución de aminoácidos en (a) posición 92 de SEQ ID NO: 34; (b) posición 106 de SEQ ID NO: 44; o (c) posición 4 de SEQ ID NO: 1.
- 25 9. Vector que comprende la secuencia de nucleótidos de la reivindicación 8.
10. Célula vegetal o célula huésped que comprende la secuencia de nucleótidos de la reivindicación 8 o el vector de la reivindicación 9 como un transgén.
11. Un método de generación de una planta según las reivindicaciones 1 a 6, que comprende las etapas de:
 - 30 a) transformar una célula vegetal con la secuencia de nucleótidos de la reivindicación 8 o el vector de la reivindicación 9, y
 - b) regenerar una planta que tiene la actividad biológica de un inductor haploide a partir de la célula vegetal.

FIGURA

