



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 362 903**

51 Int. Cl.:  
**C12Q 1/68** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **03735170 .7**

96 Fecha de presentación : **02.07.2003**

97 Número de publicación de la solicitud: **1539996**

97 Fecha de publicación de la solicitud: **15.06.2005**

54 Título: **Procedimiento para producir plantas que presentan una eficacia de transpiración mejorada y plantas producidas a partir del mismo.**

30 Prioridad: **02.07.2002 AU PS3339**

45 Fecha de publicación de la mención BOPI:  
**14.07.2011**

45 Fecha de la publicación del folleto de la patente:  
**14.07.2011**

73 Titular/es: **The Australian National University  
Acton  
Canberra, ACT 2601, AU**

72 Inventor/es: **Masle, Josette;  
Farquhar, Graham, Douglas y  
Gilmore, Scott, Robert**

74 Agente: **Curell Aguilá, Marcelino**

ES 2 362 903 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

Procedimiento para producir plantas que presentan una eficacia de transpiración mejorada y plantas producidas a partir del mismo.

5

### Campo de la invención

La presente invención se refiere al campo del cultivo de plantas y a la producción de plantas modificadas mediante ingeniería genética. Más específicamente, la invención descrita en la presente memoria proporciona unos genes que pueden mejorar la eficacia de transpiración de una planta cuando se expresan en la misma. Estos genes son particularmente útiles para la producción de plantas que presentan una eficacia de transpiración mejorada, mediante enfoques tanto de cultivo de plantas tradicional como de ingeniería genética. La invención se extiende además a las plantas producidas mediante los procedimientos descritos en la presente memoria.

### 15 Antecedentes de la invención

#### 1. General

La presente memoria contiene información sobre secuencias de aminoácidos y nucleótidos preparada utilizando PatentIn versión 3.1, presentada en la presente memoria tras las reivindicaciones. Cada secuencia de nucleótidos se identifica en la lista de secuencias mediante el indicador numérico <210> seguido por el identificador de secuencia (por ejemplo <210>1, <210>2, etc.). La longitud y el tipo de secuencia (ADN, proteína (PRT), etc.), y el organismo fuente para cada secuencia de nucleótidos, se indican mediante la información proporcionada en los campos de indicadores numéricos <211>, <212> y <213>, respectivamente. Las secuencias de nucleótidos a las que se hace referencia en la memoria descriptiva se definen mediante la expresión "SEC ID n°:", seguido del identificador de secuencia (por ejemplo SEC ID n°: 1 se refiere a la secuencia en el listado de secuencias designada como <400>1).

La designación de los residuos de nucleótido a los que se hace referencia en la presente memoria es la recomendada por la Comisión de Nomenclatura Bioquímica de la IUPAC-IUB, en el que A representa adenina, C representa citosina, G representa guanina, T representa timina, Y representa un residuo de pirimidina, R representa un residuo de purina, M representa adenina o citosina, K representa guanina o timina, S representa guanina o citosina, W representa adenina o timina, H representa un nucleótido distinto de guanina, B representa un nucleótido distinto de adenina, V representa un nucleótido distinto de timina, D representa un nucleótido distinto de citosina y N representa cualquier residuo de nucleótido. Tal como se utiliza en la presente memoria la expresión "derivado de" se considerará que indica que un número entero especificado se obtiene de una fuente particular aunque no necesariamente de manera directa de esta fuente.

En la presente memoria, a menos que el contexto requiera contrario, la palabra "comprender", o variaciones tales como "comprende" o "que comprende", debe apreciarse que implican la inclusión de una etapa o elemento o número entero o grupo de etapas o elementos o números enteros señalado pero no la exclusión de cualquier otra etapa o elemento o número entero o grupo de elementos o números enteros.

Los expertos en la materia apreciarán que pueden introducirse en la invención descrita variaciones y modificaciones distintas de las descritas específicamente. Debe apreciarse que la invención comprende todas estas variaciones y modificaciones. La invención comprende asimismo todas las etapas, características, composiciones y compuestos a los que se hace referencia o indicados en la presente memoria, individual o colectivamente, y cualquiera y todas las combinaciones o dos o más cualesquiera de dichas etapas o características.

La presente invención no estará limitada en su alcance por las formas de realización específicas descritas en la presente memoria, proporcionadas únicamente a título de ejemplo. Los productos, composiciones y procedimientos funcionalmente están comprendidos en el alcance de la invención, tal como se describe en la presente memoria.

#### 2. Descripción de la técnica relacionada

Es conocido que prácticamente todas las plantas requieren una cierta cantidad de agua para un crecimiento y desarrollo adecuados, porque la fijación y asimilación fotosintética de CO<sub>2</sub> por parte de las plantas requiere agua. Una cantidad significativa del agua absorbida por las plantas a partir del suelo vuelve a la atmósfera por medio de la transpiración de las plantas.

La eficacia de transpiración es una medida de la cantidad de materia seca producida por una planta por unidad de agua transpirada, o, dicho de otro modo, la ganancia de carbono con respecto a la pérdida de agua a través de la transpiración.

Para las plantas que presentan una baja eficacia de transpiración, o cuando el suministro de agua es reducido, la pérdida de agua a través de la transpiración puede limitar los procesos metabólicos clave asociados con el crecimiento y desarrollo de las plantas. Por ejemplo, durante una sequía, o cuando las plantas que presentan una

baja eficacia de transpiración se hacen crecer en ambientes áridos o semiáridos, la productividad de la planta determinada mediante la producción de materia seca o tasa fotosintética, se reduce considerablemente. Por lo tanto, la producción de plantas que presentan una eficacia de utilización de agua o eficacia de transpiración mejorada es sumamente deseable para su adaptación a condiciones áridas o semiáridas, o para mejorar su resistencia a la sequía.

La mejora de la eficacia de utilización de agua o eficacia de transpiración por parte de las plantas es también sumamente deseable teniendo en cuenta el cambio climático global y la creciente presión sobre los recursos de agua en el mundo. Es conocido que la utilización ineficaz del agua para agricultura afecta de manera adversa al suministro de agua navegable, agua potable y agua para uso industrial o recreativo. Por lo tanto, la producción de plantas que presentan una eficacia de transpiración mejorada es sumamente deseable para reducir la presión sobre estos recursos de agua. También es deseable para aumentar la productividad de las plantas en condiciones de riego abundante.

Mejorando la eficacia de transpiración, se mejoran las tasas de ganancia de carbono por unidad de agua transpirada, estimulando de ese modo el crecimiento de la planta en condiciones de riego abundante, o alternativamente, en condiciones de sequía suave o severa. Este se consigue mejorando la ganancia de carbono más que la tasa de transpiración, o reduciendo la cantidad de agua perdida a cualquiera tasa particular de fijación de carbono. Los expertos en la materia también consideran que para una tasa de crecimiento dada, las plantas que presentan una eficacia de transpiración mejorada secan los suelos más lentamente, y utilizan menos agua, que las plantas casi isogénicas menos eficaces.

Se han descrito varios pretratamientos químicos así como medioambientales para mejorar la capacidad de las plántulas para sobrevivir a la sequía, o bien reduciendo la transpiración o bien reduciendo la cantidad de agua que se pierde realmente a la atmósfera.

Los tratamientos medioambientales conocidos implican ampliamente la utilización de barreras físicas. Aunque es conocido que colocar una barrera física sobre los estomas de las plantas reduce la pérdida de agua por medio de transpiración, el procedimiento no siempre es deseable o practicable para cultivos hechos crecer en campo. Por ejemplo, las barreras físicas sobre los estomas de la planta pueden inhibir ciertos procesos de intercambio de gases de la planta. Es más deseable mejorar la eficacia de transpiración o eficacia de utilización de agua real de la planta a través de la manipulación de la función vegetal intrínseca.

Los agentes químicos son normalmente los denominados agentes "antitranspirantes" o "antidesecante", que se aplican ambos a las hojas. Los antitranspirantes son normalmente antitranspirantes metabólicos o películas.

Estos productos forman una película sobre las hojas, o bien bloqueando de este modo los poros de los estomas o bien recubriendo las células epidérmicas de las hojas con una película impermeable al agua. Los antitranspirantes en película típicos incluyen ceras, emulsiones de cera en aceite, alcoholes superiores, siliconas, plásticos, látex y resinas. Por ejemplo, Elmore, patente US nº 4.645.682 dio a conocer un antitranspirante que consiste en una cera de pasta acuosa; Cushman *et al.*, las patentes US nº 3.791.839 y nº 3.847.641 también dieron a conocer emulsiones de cera para controlar la transpiración en plantas; y Petrucco *et al.*, patente US nº 3.826.671, dio a conocer una composición polimérica que se supone que es eficaz para controlar la transpiración en plantas.

Los antitranspirantes metabólicos generalmente cierran los estomas, reduciendo de ese modo la tasa de transpiración. Los antitranspirantes metabólicos típicos incluyen ácidos succínicos, acetato fenilmercúrico, hidroxisulfonatos, el herbicida atrazina, azida sódica y fenilhidrazonas, así como cianuro de carbono.

Los compuestos que presentan una actividad reguladora del crecimiento de la planta también han demostrado ser útiles para reducir la transpiración. Por ejemplo, Bliesner *et al.*, patente US nº 4.671.816, dieron a conocer un compuesto de acetileno, que se supone que presenta utilidad para regular el crecimiento de la planta, mientras que Kuznetsov *et al.* (patente rusa nº SU 1.282.492; solicitud de patente rusa nº SU 1.253.559-AI) y Smirnov *et al.* (patente rusa nº SU 1.098.934) dieron a conocer la utilización de derivados de 2-metil-5-hidroxibenzimidazol, y las sales de cloruro o bromuro de los mismos, como reguladores del crecimiento antitranspirantes. Vichnevetskaia (documento USSN 5.589.437 expedida el 31 de diciembre de 1996) también describe derivados de hidroxibenzimidazol para mejorar la resistencia a la sequía de plantas reduciendo la transpiración. Schulz *et al.*, patente US nº 4.943.315, también dieron a conocer formulaciones que comprenden un compuesto de acetileno y uno de fenilbencilurea, para reducir la transpiración en plantas y/o para evitar deterioros a las plantas provocados por el calor y condiciones secas. El ácido abscísico también ha demostrado reducir o suprimir la transpiración en plantas (por ejemplo Helv. Chim. Acta, 71, 931, 1988; J. Org. Chem., 54, 681, 1989; y publicación de patente japonesa nº 184.966/1991).

Los antitranspirantes metabólicos son de producción costosa y a menudo presentan efectos fitotóxicos o inhiben el crecimiento de la planta (Kozlowski (1979), En: Tree Growth and Environmental Stresses (Univ. de Washington Press, Seattle y Londres)), y prácticamente no se utilizan.

- Estudios recientes han examinado procedimientos alternativos para mejorar la eficacia de transpiración, particularmente enfoques de cultivo para seleccionar líneas que crecen de manera más eficaz en condiciones de sequía suave. La discriminación de isótopos de carbono se ha utilizado para identificar ecotipos de *Arabidopsis* con eficacias de transpiración contrastadas (Masle *et al.*, En: Stable isotopes and plant carbon-water relations, Acad. Press, Physiol. Ser., págs. 371-386, 1993) y para ayudar al cultivo convencional de nuevas variedades de plantas en varias especies (Hall *et al.*, Plant Breeding Reviews 4, 81-113, 1994) incluyendo el arroz (Farquhar *et al.*, En: Breaking the Yield Barrier, ed KG Cassman, IRRI, 95, 101) y lo más recientemente el trigo (Rebetzke *et al.* Crop Science 42:739-745, 2002).
- No se ha identificado ningún gen individual que pueda mejorar la eficacia de transpiración cuando se expresa en una planta. La eficacia de transpiración puede muy bien ser multigénica. Como consecuencia, los genes y las rutas de señalización que regulan los componentes fotosintéticos y/o de los estomas del mecanismo de eficacia de transpiración en plantas no se han identificado ni caracterizado.
- Además, a pesar de que es conocido que el efecto de expresión por regulación por disminución del gen *Rubisco*, o la mutación en genes implicados en el ácido abscísico (por ejemplo *aba*, *abi*), modifican la eficacia de transpiración en cierta medida a través del cierre de los estomas, la consecuencia de tales modificaciones no es necesariamente específica, dando como resultados efectos pleiotrópicos.
- El ecotipo de *Arabidopsis thaliana* Landsberg *erecta* (*L-er1*) es uno de los ecotipos más populares y se utiliza ampliamente para estudios tanto moleculares como genéticos. Aloja la mutación *er1*, que confiere una inflorescencia compacta, frutos cerrados y pecíolos cortos. Hay varios alelos mutantes de *erecta*. La caracterización fenotípica de los alelos mutantes sugiere un papel para el gen *ER* de tipo natural en la regulación de la morfogénesis de la planta, particularmente las formas de los órganos que se originan del meristemo apical del brote. Torii *et al.*, The Plant Cell 8, 735, 1996, mostraron que el gen *ER* codifica para una supuesta proteína cinasa receptora que comprende un dominio catalítico de proteína cinasa citoplasmático, una región transmembrana y un dominio extracelular que consiste en repeticiones ricas en leucina, que se cree que interaccionan con otras macromoléculas.
- El nº de registro de la base de datos Genbank AY106598 da a conocer una secuencia de nucleótidos que es una secuencia parcial de un gen *ERECTA* de *Zea mays*. El documento US 2002/040489 da a conocer secuencias expresadas de *Arabidopsis thaliana*, que incluyen una secuencia de nucleótidos que es una secuencia parcial de un ARNm de *ERECTA*. Thumma *et al.*, Journal of Experimental Botany 52(355):203-214 (2001) da a conocer un procedimiento de análisis de loci de carácter cuantitativo para identificar marcadores genéticos asociados con la eficacia de transpiración en *Stylosanthes scabra*. El documento WO 01/02541 da a conocer un procedimiento para aumentar o reducir la transcripción en una planta modulando el gen de la proteína cinasa activada por ABA (AAPK), que se expresa en células oclusivas de la planta. Torii *et al.*, The Plant Cell 8:735-746 (1996) dan a conocer una secuencia *ERECTA* de *Arabidopsis* y mutantes de esta secuencia que presentan una altura de tallo alterada.

### Sumario de la invención

- En el trabajo que conduce a la presente invención, se pretendió esclarecer los determinantes genéticos específicos de la eficacia de transpiración de las plantas. En las plantas, el desarrollo de marcadores genéticos moleculares, tales como, por ejemplo, marcadores genéticos que se correlacionan con una región del genoma de una planta de cultivo, tal como, por ejemplo, una región del genoma del arroz, genoma del maíz, genoma de la cebada, genoma del sorgo o genoma del trigo, o una región del genoma del tomate o de cualquier *Brassicaceae*, ayuda en la producción de plantas que presentan una eficacia de transpiración mejorada (Edwards *et al.*, Genetics 116, 113 - 125, 1987; Paterson *et al.*, Nature 335, 721-726, 1988).
- Se identificó un locus que está vinculado a la variación genética en la eficacia de transpiración en plantas. Para esclarecer un locus asociado con la eficacia de transpiración de plantas, los inventores establecieron condiciones experimentales y procedimientos de muestreo para determinar la contribución a la eficacia de transpiración total de los factores que influyen en este fenotipo, y, más particularmente, la contribución genética a la variación total en la eficacia de transpiración. Los factores que influyen en la eficacia de transpiración incluyen, por ejemplo, el genotipo de la planta, el entorno (por ejemplo temperatura, luz, humedad, capa límite alrededor de las hojas, condiciones de crecimiento de la raíz), el desarrollo (por ejemplo edad y/o fase y/o postura de las plantas que modifican el intercambio de gases y/o el metabolismo del carbono) y factores específicos de la semilla (Masle *et al.* 1993, citado anteriormente). Los exámenes desarrollados por los inventores también se utilizaron para investigar las poblaciones de tipo natural y mutantes para variaciones en la eficacia de transpiración y para identificar ecotipos que presentan eficacias de transpiración diferentes incluyendo las líneas parentales que se habían utilizado por Lister y Dean (1993). Entonces se determinaron las eficacias de transpiración de los elementos de la población de mapeo de línea endogámica recombinante (RIL) de Lister y Dean (1993), y se realizaron análisis de ligamiento frente a marcadores genéticos para determinar las regiones cromosómicas que están vinculadas a la variación genética en la eficacia de transpiración, identificando de ese modo un locus que condiciona la eficacia de transpiración. Pruebas de complementación, en el que las plantas se transformaron con un alelo de tipo natural en este locus confirmaron la funcionalidad del alelo en la determinación de un fenotipo de eficacia de transpiración.

Por lo tanto la presente invención proporciona procedimientos y utilidades tal como se definen en las reivindicaciones adjuntas. En la presente memoria se describe un locus asociado con la eficacia de transpiración de *A. thaliana*, tal como, por ejemplo el locus *ERECTA* en el cromosoma 2 de *A. thaliana*, o una sonda de hibridación que se correlaciona con la región entre aproximadamente 46 cM y aproximadamente 50,7 cM en el cromosoma 2 de *A. thaliana*. Se identificaron alelos *ERECTA* adicionales o alelos *erecta* en *A. thaliana*, arroz, sorgo, trigo y maíz que están relacionados estructuralmente con este alelo *ERECTA* o *erecta* de *A. thaliana* primario. Basándose en el gran número de alelos *ERECTA/erecta* descritos en la presente memoria, pueden identificarse homólogos del locus *ERECTA* de *A. thaliana* de otras especies vegetales que las mostradas específicamente a título de ejemplo, utilizando los procedimientos descritos en la presente memoria.

En la presente memoria se describe un marcador genético o locus asociado con la variación genética en la eficacia de transpiración de una planta, en el que dicho locus comprende una secuencia de nucleótidos ligada genéticamente a un locus *ERECTA* en el genoma de la planta. El locus o marcador genético es útil para determinar la eficacia de transpiración de una planta.

Tal como se utilizan en la presente memoria, las expresiones “ligado genéticamente” y “se correlaciona con” deben considerarse que hacen referencia a una proximidad genética suficiente entre un ácido nucleico ligado que comprende un gen, alelo, marcador u otra secuencia de nucleótidos y ácido nucleico que comprende todo o parte de un locus *ERECTA* para permitir que dicho ácido nucleico ligado sea útil para determinar la presencia de un alelo particular de dicho locus *ERECTA* en el genoma de una planta. Los expertos en la materia apreciarán que para que un ácido nucleico ligado de este tipo se utilice de esta manera, debe estar suficientemente cerca de dicho locus para no estar en desequilibrio de ligamiento o para presentar una alta frecuencia de recombinación entre dicho ácido nucleico ligado y dicho locus. Preferentemente, el ácido nucleico ligado y el locus están separados menos de aproximadamente 25 cM, más preferentemente separados menos de aproximadamente 10cM, incluso más preferentemente separados menos de aproximadamente 5cM, todavía incluso más preferentemente separados menos de aproximadamente 3cM y todavía incluso más preferentemente separados menos de aproximadamente 1cM.

En la presente memoria se describe un ácido nucleico aislado asociado con la variación genética en la eficacia de transpiración de una planta, comprendiendo dicho ácido nucleico una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

- (a) la secuencia de un gen genómico *ERECTA* o la 5'-UTR o 3'-UTR o región que codifica para proteínas o una región de intrón de la misma;
- (b) la secuencia de una variante alélica de (a) o la 5'-UTR o 3'-UTR o región que codifica para proteínas o una región de intrón de dicha variante alélica;
- (c) la secuencia de un fragmento de (a) o (b) que se hibrida específicamente con ácido nucleico (por ejemplo, ARN o ADN) de una planta en condiciones de hibridación de por lo menos baja rigurosidad; y
- (d) una secuencia que es complementaria a (a) o (b) o (c).

En la presente memoria se describe un gen *ERECTA* aislado de trigo que comprende una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

- (i) la secuencia expuesta en SEC ID nº: 19;
- (ii) una secuencia que codifica para la secuencia de aminoácidos expuesta en SEC ID nº: 20; y
- (iii) una secuencia que es complementaria a (i) o (ii).

En la presente memoria se describe un gen *ERECTA* aislado de maíz que comprende una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

- (i) la secuencia expuesta en SEC ID nº: 44;
- (ii) una secuencia que codifica para la secuencia de aminoácidos expuesta en SEC ID nº: 45; y
- (iii) una secuencia que es complementaria a (i) o (ii).

En la presente memoria se describe un gen *ERECTA* aislado de arroz que comprende una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

- (i) la secuencia expuesta en SEC ID nº: 3;

(ii) una secuencia que codifica para la secuencia de aminoácidos expuesta en SEC ID nº: 4; y

(iii) una secuencia que es complementaria a (i) o (ii).

5 En la presente memoria se describe un gen *ERECTA* aislado de *A. thaliana* que comprende una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

(i) la secuencia expuesta en SEC ID nº: 1;

10 (ii) una secuencia que codifica para la secuencia de aminoácidos expuesta en SEC ID nº: 2; y

(iv) una secuencia que es complementaria a (i) o (ii).

15 En la presente memoria se describe un gen *ERECTA* aislado de *A. thaliana* que comprende una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

(i) la secuencia expuesta en SEC ID nº: 7;

20 (ii) una secuencia que codifica para la secuencia de aminoácidos expuesta en SEC ID nº: 8; y

(v) una secuencia que es complementaria a (i) o (ii).

25 En la presente memoria se describe un gen *ERECTA* aislado de *A. thaliana* que comprende una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

(i) la secuencia expuesta en SEC ID nº: 9;

(ii) una secuencia que codifica para la secuencia de aminoácidos expuesta en SEC ID nº: 10; y

30 (vi) una secuencia que es complementaria a (i) o (ii).

En la presente memoria se describe un gen *ERECTA* aislado de sorgo que comprende una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

35 (i) la secuencia expuesta en SEC ID nº: 5;

(ii) una secuencia que codifica para la secuencia de aminoácidos expuesta en SEC ID nº: 6; y

(vii) una secuencia que es complementaria a (i) o (ii).

40 A pesar de que un gen genómico o gen estructural *ERECTA* o *erecta* o la región que codifica para proteínas del mismo es particularmente útil para fines de producción y/o mapeo, este aspecto no debe limitarse al gen genómico o gen estructural *ERECTA* o *erecta* o la región que codifica para proteínas del mismo. Tal como se muestra a título de ejemplo en la presente memoria, el locus *ERECTA* de *A. thaliana* primario puede determinarse utilizando cualquier ácido nucleico ligado que se correlaciona con una región en el cromosoma a una distancia genética de hasta aproximadamente 3 cM desde el alelo *ERECTA* o *erecta*. El experto en la materia podría fácilmente utilizar sondas similares para identificar el ligamiento a un alelo *ERECTA* o *erecta* en cualquier otra especie vegetal, basándose en la enseñanza proporcionada en la presente memoria de que el alelo *ERECTA* o *erecta* está vinculado al fenotipo de eficacia de transpiración de plantas.

50 Preferentemente, todo o parte del locus asociado con el fenotipo de eficacia de transpiración en una planta (es decir, ácido nucleico ligado genéticamente al gen genómico o estructural *ERECTA* o *erecta*) se proporciona como ácido nucleico recombinante o aislado, tal como, por ejemplo, en forma de un constructo génico (por ejemplo un cósmido o plásmido recombinante), para facilitar el examen del germoplasma.

55 El locus *ERECTA* o un gen que está ligado al locus *ERECTA* es particularmente útil en un programa de producción, para predecir la eficacia de transpiración de una planta, o alternativamente, como marcador de producción selectivo para seleccionar plantas que presentan una eficacia de transpiración mejorada. Una vez mapeado, la selección asistida por marcadores (MAS) se utiliza para introducir el locus *ERECTA* o marcadores ligados al mismo en una amplia variedad de poblaciones. La MAS presenta la ventaja de reducir el tamaño de la población de producción requerido, y la necesidad de pruebas recurrentes continuas de la progenie, y el tiempo requerido para desarrollar una línea superior.

60 Por lo tanto, un aspecto de la presente invención proporciona un procedimiento para seleccionar una planta que presenta una eficacia de transpiración mejorada, que comprende detectar un marcador genético para la eficacia de transpiración, marcador que comprende un locus *ERECTA* en el genoma de la planta y seleccionar una planta que

comprende o expresa el marcador genético, en el que el marcador genético comprende una secuencia de nucleótidos que presenta por lo menos aproximadamente una identidad de secuencia global del 55% con por lo menos aproximadamente 20 nucleótidos de longitud de cualquiera de las SEC ID n°: 1, 3, 5, 7, 9, 11 a 19 ó 21 a 44 o una secuencia complementaria a las mismas, incluyendo una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

(a) una secuencia que presenta una identidad de por lo menos aproximadamente 55% con una secuencia seleccionada de entre el grupo constituido por SEC ID n°: 1, SEC ID n°: 3, SEC ID n°: 5, SEC ID n°: 7, SEC ID n°: 9, SEC ID n°: 11, SEC ID n°: 12, SEC ID n°: 13, SEC ID n°: 14, SEC ID n°: 15, SEC ID n°: 16, SEC ID n°: 17, SEC ID n°: 18, SEC ID n°: 19 SEC ID n°: 21, SEC ID n°: 22, SEC ID n°: 23, SEC ID n°: 24, SEC ID n°: 25, SEC ID n°: 26, SEC ID n°: 27, SEC ID n°: 28, SEC ID n°: 29, SEC ID n°: 30, SEC ID n°: 31, SEC ID n°: 32 SEC ID n°: 33, SEC ID n°: 34, SEC ID n°: 35, SEC ID n°: 36, SEC ID n°: 37, SEC ID n°: 38; SEC ID n°: 39, SEC ID n°: 40, SEC ID n°: 41, SEC ID n°: 42, SEC ID n°: 43, SEC ID n°: 44;

(b) una secuencia que codifica para una secuencia de aminoácidos que presenta una identidad de por lo menos aproximadamente 55% con una secuencia de aminoácidos seleccionada de entre el grupo constituido por SEC ID n°: 2, SEC ID n°: 4, SEC ID n°: 6, SEC ID n°: 8, SEC ID n°: 10, SEC ID n°: 12, SEC ID n°: 20 y SEC ID n°: 45; y

(c) una secuencia complementaria a (a) o (b).

En una forma de realización, el procedimiento comprende:

presentar una eficacia de transpiración mejorada, que comprende:

(a) examinar líneas endogámicas recombinantes o casi isogénicas o mutantes de plantas para segregar alelos en un locus *ERECTA*; y

(b) identificar un marcador polimórfico ligado a dicho locus *ERECTA*.

Los datos mostrados a título de ejemplo en la presente memoria para *A. thaliana* o arroz pueden extrapolarse claramente a las demás especies vegetales. Por ejemplo, la evidencia proporcionada en la presente memoria para el papel del alelo *ERECTA* de *A. thaliana* en la determinación del fenotipo de eficacia de transpiración en esas especies vegetales ha permitido esclarecer un rango amplio de alelos homólogos *ERECTA* en otras especies vegetales, en particular trigo, arroz, sorgo y maíz, que son también probables para determinar el fenotipo de eficacia de transpiración en esas plantas. También se describe en la presente memoria un procedimiento para seleccionar una planta que presenta una eficacia de transpiración mejorada, que comprende seleccionar una planta que comprende o expresa un homólogo funcionalmente equivalente de una región que codifica para proteínas del gen *ERECTA* de *A. thaliana*, maíz, trigo, sorgo o arroz.

También se describe en la presente memoria un procedimiento para seleccionar una planta que presenta una eficacia de transpiración mejorada, que comprende:

(a) identificar un locus en el cromosoma 2 de *Arabidopsis* (46-50,7 cM) o el cromosoma 6 de arroz asociado con una variación genética en la eficacia de transpiración en una planta;

(b) identificar un ácido nucleico en una especie vegetal diferente que comprende una secuencia de nucleótidos que presenta una identidad de por lo menos aproximadamente 55% con la secuencia del locus en (a); y

(c) seleccionar una planta que comprende o expresa el ácido nucleico identificado en (b).

También se describe en la presente memoria un procedimiento para seleccionar una planta que presenta una eficacia de transpiración mejorada, que comprende:

(a) identificar un locus en el cromosoma 2 de *Arabidopsis* (46-50,7 cM) o el cromosoma 6 de arroz asociado con una variación genética en la eficacia de transpiración en una planta;

(b) determinar la secuencia de nucleótidos del locus identificado;

(c) identificar un ácido nucleico de una especie vegetal distinta de *A. thaliana* o arroz que comprende una secuencia de nucleótidos que presenta una identidad de por lo menos aproximadamente 55% con la secuencia del locus en (a); y

(d) seleccionar una planta que comprende o expresa el ácido nucleico identificado en (b).

Preferentemente, la planta seleccionada según cualquiera o más de las formas de realización anteriores es

*Arabidopsis thaliana*, arroz, sorgo, trigo o maíz, sin embargo no se excluyen otras especies.

Preferentemente, el procedimiento de selección objeto comprende vincular el fenotipo de eficacia de transpiración de la planta a la expresión del marcador en la planta, o alternativamente, vincular un polimorfismo estructural en el ADN a un fenotipo de eficacia de transpiración en la planta, por ejemplo, mediante un proceso que comprende detectar un polimorfismo de la longitud de los fragmentos de restricción (RFLP), polimorfismo de la longitud de los fragmentos amplificados (AFLP), polimorfismo en la conformación de cadenas sencillas (SSCP) o análisis de microsatélites. Tal como conocerá el experto en la materia, una sonda de ácido nucleico o cebador de por lo menos aproximadamente 20 nucleótidos de longitud de cualquiera de las SEC ID n°: 1, 3, 5, 7, 9, 11 a 19 ó 21 a 44 o una secuencia complementaria a las mismas puede hibridarse con el ADN genómico de la planta, y la hibridación detectada utilizando medios de detección, identificando de ese modo el polimorfismo.

Resulta claramente preferido que la planta seleccionada presente una eficacia de transpiración mejorada en comparación con una planta casi isogénica que no comprende ni expresa el marcador genético.

Tal como se muestra a título de ejemplo en la presente memoria, se identificaron asimismo genes o alelos específicos que están ligados al locus *ERECTA* de *A. thaliana* y arroz, y determinaron las eficacias de transpiración de esas plantas. Más particularmente, se determinan las eficacias de transpiración de líneas casi isogénicas, que portan, cada una, una mutación dentro de un locus *ERECTA*, y una correlación entre el fenotipo de eficacia de transpiración y la expresión de *ERECTA* o el número de copias génicas, proporcionando de ese modo la contribución genética de genes o alelos en el locus *ERECTA* a la eficacia de transpiración. Este análisis permite una evaluación de la contribución genética de alelos particulares a la eficacia de transpiración, determinando de ese modo las variantes alélicas que están vinculadas a una eficacia de transpiración particular. Por tanto, el esclarecimiento del locus *ERECTA* para la eficacia de transpiración en plantas facilita el mapeo fino y la determinación de variantes alélicas que modulan la eficacia de transpiración. Los procedimientos descritos en la presente memoria pueden aplicarse a una evaluación de la contribución de alelos específicos al fenotipo de eficacia de transpiración para cualquier especie vegetal que sea susceptible de mutagénesis tal como, por ejemplo, mediante mutagénesis por transposones, irradiación o medios químicos. Tal como conocerá el experto en la materia muchas especies de cultivo, tales como, maíz, trigo y arroz, son susceptibles de tal mutagénesis.

También se describe en la presente memoria un procedimiento para identificar un gen que determine la eficacia de transpiración de una planta que comprende:

- (a) identificar un locus asociado con una variación genética en la eficacia de transpiración en una planta;
- (b) identificar un gen o alelo que está ligado a dicho locus, en el que dicho gen o alelo es un gen o alelo candidato para determinar la eficacia de transpiración de una planta; y
- (c) determinar las eficacias de transpiración de un grupo de plantas, en el que no todos los elementos de dicho grupo comprende o expresa dicho gen o alelo, y en el que la variación en la eficacia de transpiración entre los elementos de dicho grupo indica que dicho gen está implicado en la determinación de la eficacia de transpiración.

El procedimiento puede comprender:

- (a) identificar un locus asociado con una variación genética en la eficacia de transpiración en una planta;
- (b) identificar múltiples alelos de un gen que está ligado a dicho locus, en el que dicho gen es un gen candidato implicado en la determinación de la eficacia de transpiración de una planta; y
- (c) determinar las eficacias de transpiración de un grupo de plantas, en el que cada elemento de dicho grupo comprende, y preferentemente expresa, por lo menos uno de dichos múltiples alelos, en el que la variación en la eficacia de transpiración entre los elementos de dicho grupo indica que dicho gen está implicado en la determinación de la eficacia de transpiración.

Preferentemente, el gen identificado o alelo identificado mediante el procedimiento descrito en el párrafo anterior es un alelo *ERECTA*, o un alelo *erecta*, de una planta seleccionada de entre el grupo constituido por *A. thaliana*, sorgo, arroz, maíz y trigo, o un homólogo de los mismos.

El gen o alelo identificado, incluyendo cualquier homólogo de una planta distinta de *A. thaliana*, tal como, por ejemplo, el alelo *ERECTA* de tipo natural o un homólogo de los mismos, es útil para la producción de nuevas plantas. Tales plantas se producen, por ejemplo, utilizando técnicas recombinantes, o enfoques de cultivo de plantas tradicionales tales como introgresión.

También se describe en la presente memoria un procedimiento para modular (es decir, mejorar o reducir) la eficacia de transpiración de una planta que comprende expresar ectópicamente en una planta un gen *ERECTA* aislado o una



variante alélica del mismo o la región que codifica para proteínas de dicho gen *ERECTA* o dicha variante alélica. También se describe en la presente memoria un procedimiento para mejorar la eficacia de transpiración de una planta que comprende introgresar en dicha planta un ácido nucleico que comprende una secuencia de nucleótidos que es homóloga a una región que codifica para proteínas de un gen de *A. thaliana* que se correlaciona con el locus *ERECTA* en el cromosoma 2.

Una forma de realización adicional de la invención proporciona un procedimiento para modular la eficacia de transpiración de una planta que comprende introducir y preferentemente expresar en la misma, un gen *ERECTA* aislado o una variante alélica del mismo o la región que codifica para proteínas del mismo en una planta y seleccionar una planta que presenta una eficacia de transpiración diferente en comparación con una planta casi isogénica que no comprende el gen *ERECTA* o variante alélica o región que codifica para proteínas introducido, en el que el gen *ERECTA* o variante alélica o región que codifica para proteínas comprende una secuencia de nucleótidos que presenta por lo menos una identidad de secuencia global del 55% con por lo menos 20 nucleótidos de longitud de cualquiera de las SEC ID nº: 1, 3, 5, 7, 9, 11 a 19 ó 21 a 44 o una secuencia complementaria a las mismas, en el que el gen *ERECTA* o variante alélica o región que codifica para proteínas se introduce en la planta mediante un proceso que comprende transformar material vegetal con un constructo génico que comprende el gen o variante alélica o región que codifica para proteínas del mismo. Preferentemente el gen *ERECTA* o variante alélica o región que codifica para proteínas comprende una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

(a) una secuencia que presenta una identidad de por lo menos aproximadamente 55% con una secuencia seleccionada de entre el grupo constituido por SEC ID nº: 1, SEC ID nº: 3, SEC ID nº: 5, SEC ID nº: 7, SEC ID nº: 9, SEC ID nº: 11, SEC ID nº: 12, SEC ID nº: 13, SEC ID nº: 14, SEC ID nº: 15, SEC ID nº: 16, SEC ID nº: 17, SEC ID nº: 18, SEC ID nº: 19 SEC ID nº: 21, SEC ID nº: 22, SEC ID nº: 23, SEC ID nº: 24, SEC ID nº: 25, SEC ID nº: 26, SEC ID nº: 27, SEC ID nº: 28, SEC ID nº: 29, SEC ID nº: 30, SEC ID nº: 31, SEC ID nº: 32 SEC ID nº: 33, SEC ID nº: 34, SEC ID nº: 35, SEC ID nº: 36, SEC ID nº: 37, (SEC ID nº: 38; SEC ID nº: 39, SEC ID nº: 40, SEC ID nº: 41, SEC ID nº: 42, SEC ID nº: 43 y SEC ID nº: 44; y

(b) una secuencia que codifica para una secuencia de aminoácidos que presenta una identidad de por lo menos aproximadamente 55% con una secuencia de aminoácidos seleccionada de entre el grupo constituido por SEC ID nº: 2, SEC ID nº: 4, SEC ID nº: 6, SEC ID nº: 8, SEC ID nº: 10, SEC ID nº: 12, SEC ID nº: 20 y SEC ID nº: 45.

La planta en la que se introduce el gen, etc. se selecciona preferentemente de entre el grupo constituido por *Arabidopsis thaliana*, arroz, sorgo, trigo y maíz. Tal como resultará evidente a partir de la presente descripción, la eficacia de transpiración se mejora como consecuencia de la expresión ectópica de un alelo *ERECTA* o la región que codifica para proteínas del mismo en la planta. Por el contrario, la eficacia de transpiración se reduce como consecuencia de la expresión reducida de un alelo *ERECTA* en la planta (por ejemplo, mediante la expresión de ARN antisentido o ARNi u otro ARN inhibidor).

Un aspecto adicional de la invención proporciona la utilización de un gen *ERECTA* aislado o una variante alélica del mismo o la región que codifica para proteínas de dicho gen *ERECTA* o dicha variante alélica en la preparación de un constructo génico para modular (es decir, mejorar o reducir) la eficacia de transpiración de una planta, en la que el gen *ERECTA* o variante alélica o región que codifica para proteínas comprende una secuencia de nucleótidos que presenta por lo menos una identidad de secuencia global del 55% con por lo menos 20 nucleótidos de longitud de cualquiera de las SEC ID nº: 1, 3, 5, 7, 9, 11 a 19 ó 21 a 44 o una secuencia complementaria a las mismas. Por ejemplo, la expresión de proteína *ERECTA* en la planta puede modificarse mediante la expresión epitópica de un alelo *ERECTA* en la planta, o alternativamente, reduciendo la expresión de *ERECTA* endógena utilizando un ARN inhibidor (por ejemplo, antisentido o ARNi).

También se describe en la presente memoria una planta que presenta una eficacia de transpiración mejorada, produciéndose dicha planta mediante un procedimiento descrito en la presente memoria.

Las plantas que presentan una eficacia de transpiración mejorada muestran niveles aumentados de crecimiento en condiciones de crecimiento normales, aumentando de ese modo su biomasa. Por lo tanto, también se describe en la presente memoria un procedimiento para aumentar la biomasa de una planta que comprende mejorar el nivel de expresión de un gen *ERECTA* o variante alélica del mismo o región que codifica para proteínas del mismo en dicha planta.

El procedimiento puede incluir adicionalmente la etapa de seleccionar una planta que presenta una biomasa aumentada en comparación con una planta sin modificar. Procedimientos para determinar la biomasa de una planta son muy conocidos para los expertos en la materia y/o se describen en la presente memoria.

El nivel de expresión puede mejorarse mediante la modificación genética de una secuencia control, por ejemplo una secuencia promotora, asociada con el gen *ERECTA* o variante alélica del mismo.

El nivel de expresión puede mejorarse introduciendo (por ejemplo, mediante cultivo clásico, introgresión o medios

recombinantes) un gen *ERECTA* o variante alélica del mismo o la región que codifica para proteínas del mismo en una planta. Preferentemente, el gen *ERECTA* o variante alélica o región que codifica para proteínas comprende una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

5 una secuencia que presenta una identidad de por lo menos aproximadamente 55% con una secuencia seleccionada de entre el grupo constituido por SEC ID nº: 1, SEC ID nº: 3, SEC ID nº: 5, SEC ID nº: 7, SEC ID nº: 9, SEC ID nº: 11, SEC ID nº: 12, SEC ID nº: 13, SEC ID nº: 14, SEC ID nº: 15, SEC ID nº: 16, SEC ID nº: 17, SEC ID nº: 18, SEC ID nº: 19, SEC ID nº: 21, SEC ID nº: 22, SEC ID nº: 23, SEC ID nº: 24, SEC ID nº: 25, SEC ID nº: 26, SEC ID nº: 27, SEC ID nº: 28, SEC ID nº: 29, SEC ID nº: 30, SEC ID nº: 31, SEC ID nº: 32, SEC ID nº: 33, SEC ID nº: 34, SEC ID nº: 35, SEC ID nº: 36, SEC ID nº: 37, SEC ID nº: 38; SEC ID nº: 39, SEC ID nº: 40, SEC ID nº: 41, SEC ID nº: 42, SEC ID nº: 43 y SEC ID nº: 44; y

15 una secuencia que codifica para una secuencia de aminoácidos que presenta una identidad de por lo menos aproximadamente 55% con una secuencia de aminoácidos seleccionada de entre el grupo constituido por SEC ID nº: 2, SEC ID nº: 4, SEC ID nº: 6, SEC ID nº: 8, SEC ID nº: 10, SEC ID nº: 12, SEC ID nº: 20 y SEC ID nº: 45.

La planta en la que se introduce el gen, etc. se selecciona preferentemente de entre el grupo constituido por *Arabidopsis thaliana*, arroz, sorgo, trigo y maíz.

20 Un aspecto adicional de la presente invención proporciona un procedimiento para aumentar la resistencia de una planta a un estrés medioambiental que comprende mejorar el nivel de expresión de un gen *ERECTA* o variante alélica del mismo o región que codifica para proteínas del mismo en dicha planta, en el que el gen *ERECTA* o variante alélica o región que codifica para proteínas comprende una secuencia de nucleótidos que presenta por lo menos una identidad de secuencia global del 55% con por lo menos 20 nucleótidos de longitud de cualquiera de las SEC ID nº: 1, 3, 5, 7, 9, 11 a 19 ó 21 a 44 o una secuencia complementaria a las mismas.

25 Tal como se utiliza en la presente memoria la expresión "estrés ambiental" debe considerarse en su contexto más amplio como que significa una o más condiciones medioambientales que reducen la capacidad de una planta para crecer, sobrevivir y/o producir semilla/grano. En una forma de realización, un estrés ambiental que afecta a la capacidad de una planta para crecer, sobrevivir y/o producir semilla/grano es una condición seleccionada de entre el grupo constituido por niveles de CO<sub>2</sub> aumentados o reducidos, una temperatura aumentada o reducida, precipitaciones aumentadas o reducidas, una humedad aumentada o reducida, niveles de sal aumentados en el suelo, compactación y consistencia del suelo aumentadas y sequía.

35 En una forma de realización, el procedimiento incluye además la etapa de seleccionar una planta que presenta una resistencia alterada a un estrés medioambiental en comparación con cuando se selecciona una planta sin modificar. Procedimientos para determinar la resistencia de una planta a un estrés ambiental son muy conocidos para los expertos en la materia y/o se describen en la presente memoria.

40 En una forma de realización, el nivel de expresión se mejora mediante la modificación genética de una secuencia control, por ejemplo una secuencia promotora, asociada con el gen *ERECTA* o variante alélica del mismo.

45 En otra forma de realización, el nivel de expresión se mejora introduciendo un gen *ERECTA* o variante alélica del mismo o la región que codifica para proteínas del mismo en una planta. Preferentemente, el gen *ERECTA* o variante alélica o región que codifica para proteínas comprende una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

50 una secuencia que presenta una identidad de por lo menos aproximadamente 55% con una secuencia seleccionada de entre el grupo constituido por SEC ID nº: 1, SEC ID nº: 3, SEC ID nº: 5, SEC ID nº: 7, SEC ID nº: 9, SEC ID nº: 11, SEC ID nº: 12, SEC ID nº: 13, SEC ID nº: 14, SEC ID nº: 15, SEC ID nº: 16, SEC ID nº: 17, SEC ID nº: 18, SEC ID nº: 19, SEC ID nº: 21, SEC ID nº: 22, SEC ID nº: 23, SEC ID nº: 24, SEC ID nº: 25, SEC ID nº: 26, SEC ID nº: 27, SEC ID nº: 28, SEC ID nº: 29, SEC ID nº: 30, SEC ID nº: 31, SEC ID nº: 32, SEC ID nº: 33, SEC ID nº: 34, SEC ID nº: 35, SEC ID nº: 36, SEC ID nº: 37, SEC ID nº: 38; SEC ID nº: 39, SEC ID nº: 40, SEC ID nº: 41, SEC ID nº: 42, SEC ID nº: 43 y SEC ID nº: 44; y

55 una secuencia que codifica para una secuencia de aminoácidos que presenta una identidad de por lo menos aproximadamente 55% con una secuencia de aminoácidos seleccionada de entre el grupo constituido por SEC ID nº: 2, SEC ID nº: 4, SEC ID nº: 6, SEC ID nº: 8, SEC ID nº: 10, SEC ID nº: 12, SEC ID nº: 20 y SEC ID nº: 45.

60 La planta en la que se introduce el gen, etc. se selecciona preferentemente de entre el grupo constituido por *Arabidopsis thaliana*, arroz, sorgo, trigo y maíz.

También se describe en la presente memoria una planta que presenta una resistencia aumentada a un estrés ambiental, en el que dicha planta se produce mediante un procedimiento descrito en la presente memoria.

65 Tanto la temperatura como la humedad disponible han demostrado influir drásticamente en la polinización y el

desarrollo de grano/semilla, procesos conocidos como cuajado de las semillas y llenado de granos. Por lo tanto, un procedimiento que produce una planta que es resistente a un estrés ambiental, es decir, una planta que presenta una eficacia de transpiración aumentada, da como resultado un llenado de granos más eficaz o aumentado y un mayor número de semillas. Dado que *ERECTA* se expresa durante la floración o el desarrollo de las vainas este gen o una variante alélica del mismo es útil para aumentar el llenado de granos en una planta.

Por lo tanto, un aspecto adicional de la presente invención proporciona un procedimiento para aumentar el peso de semilla o de grano en una planta que comprende mejorar el nivel de expresión de un gen *ERECTA* o variante alélica del mismo o región que codifica para proteínas del mismo en dicha planta, en el que el gen *ERECTA* o variante alélica o región que codifica para proteínas comprende una secuencia de nucleótidos que presenta por lo menos una identidad de secuencia global del 55% con por lo menos 20 nucleótidos de longitud de cualquiera de las SEC ID nº: 1, 3, 5, 7, 9, 11 a 19 ó 21 a 44 o una secuencia complementaria a las mismas.

En una forma de realización, el procedimiento incluye además la etapa de seleccionar una planta que presenta un peso de semilla o de grano aumentado en comparación con cuando se selecciona una planta sin modificar. Procedimientos para determinar el peso de semilla o de grano son muy conocidos para los expertos en la materia y/o se describen en la presente memoria.

En una forma de realización, el nivel de expresión se mejora mediante la modificación genética de una secuencia control, por ejemplo una secuencia promotora, asociada con el gen *ERECTA* o variante alélica del mismo

En otra realización, el nivel de expresión se mejora introduciendo un gen *ERECTA* o variante alélica del mismo o la región que codifica para proteínas del mismo en una planta. Preferentemente, el gen *ERECTA* o variante alélica o región que codifica para proteínas comprende una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

una secuencia que presenta una identidad de por lo menos aproximadamente 55% con una secuencia seleccionada de entre el grupo constituido por SEC ID nº: 1, SEC ID nº: 3, SEC ID nº: 5, SEC ID nº: 7, SEC ID nº: 9, SEC ID nº: 11, SEC ID nº: 12, SEC ID nº: 13, SEC ID nº: 14, SEC ID nº: 15, SEC ID nº: 16, SEC ID nº: 17, SEC ID nº: 18, SEC ID nº: 19, SEC ID nº: 21, SEC ID nº: 22, SEC ID nº: 23, SEC ID nº: 24, SEC ID nº: 25, SEC ID nº: 26, SEC ID nº: 27, SEC ID nº: 28, SEC ID nº: 29, SEC ID nº: 30, SEC ID nº: 31, SEC ID nº: 32, SEC ID nº: 33, SEC ID nº: 34, SEC ID nº: 35, SEC ID nº: 36, SEC ID nº: 37, SEC ID nº: 38; SEC ID nº: 39, SEC ID nº: 40, SEC ID nº: 41, SEC ID nº: 42, SEC ID nº: 43 y SEC ID nº: 44; y

una secuencia que codifica para una secuencia de aminoácidos que presenta una identidad de por lo menos aproximadamente 55% con una secuencia de aminoácidos seleccionada de entre el grupo constituido por SEC ID nº: 2, SEC ID nº: 4, SEC ID nº: 6, SEC ID nº: 8, SEC ID nº: 10, SEC ID nº: 12, SEC ID nº: 20 y SEC ID nº: 45.

La planta en la que se introduce el gen, etc. se selecciona preferentemente de entre el grupo constituido por *Arabidopsis thaliana*, arroz, sorgo, trigo y maíz.

También se describe en la presente memoria una planta que presenta un peso de semilla o de grano aumentado, en el que dicha planta se produce mediante un procedimiento descrito en la presente memoria.

También se describe en la presente memoria un procedimiento para modular el número de semillas producidas mediante una planta que comprende mejorar el nivel de expresión de un gen *ERECTA* o variante alélica del mismo en dicha planta.

El procedimiento puede incluir adicionalmente la etapa de seleccionar una planta que presenta un número aumentado de semillas en comparación con cuando se selecciona una planta sin modificar. Procedimientos para determinar el número de semillas o de granos son muy conocidos para los expertos en la materia y/o se describen en la presente memoria.

El nivel de expresión puede mejorarse mediante la modificación genética de una secuencia control, por ejemplo una secuencia promotora, asociada con el gen *ERECTA* o variante alélica del mismo

El nivel de expresión puede mejorarse introduciendo (por ejemplo, mediante cultivo clásico, introgresión o medios recombinantes) un gen *ERECTA* o variante alélica del mismo o la región que codifica para proteínas del mismo en una planta. Preferentemente, el gen *ERECTA* o variante alélica o región que codifica para proteínas comprende una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

una secuencia que presenta una identidad de por lo menos aproximadamente 55% con una secuencia seleccionada de entre el grupo constituido por SEC ID nº: 1, SEC ID nº: 3, SEC ID nº: 5, SEC ID nº: 7, SEC ID nº: 9, SEC ID nº: 11, SEC ID nº: 12, SEC ID nº: 13, SEC ID nº: 14, SEC ID nº: 15, SEC ID nº: 16, SEC ID nº: 17, SEC ID nº: 18, SEC ID nº: 19, SEC ID nº: 21, SEC ID nº: 22, SEC ID nº: 23, SEC ID nº: 24, SEC ID nº: 25, SEC ID nº: 26, SEC ID nº: 27, SEC ID nº: 28, SEC ID nº: 29, SEC ID nº: 30, SEC ID nº: 31, SEC ID nº: 32, SEC ID nº: 33, SEC ID nº: 34, SEC ID nº:

35, SEC ID nº: 36, SEC ID nº: 37, SEC ID nº: 38; SEC ID nº: 39, SEC ID nº: 40, SEC ID nº: 41, SEC ID nº: 42, SEC ID nº: 43 y SEC ID nº: 44; y

una secuencia que codifica para una secuencia de aminoácidos que presenta una identidad de por lo menos aproximadamente 55% con una secuencia de aminoácidos seleccionada de entre el grupo constituido por SEC ID nº: 2, SEC ID nº: 4, SEC ID nº: 6, SEC ID nº: 8, SEC ID nº: 10, SEC ID nº: 12, SEC ID nº: 20 y SEC ID nº: 45.

La planta en la que se introduce el gen, etc. se selecciona preferentemente de entre el grupo constituido por *Arabidopsis thaliana*, arroz, sorgo, trigo y maíz.

También se describe en la presente memoria una planta que presenta un número aumentado de semillas, en el que dicha planta se produce mediante un procedimiento descrito en la presente memoria.

### Breve descripción de los dibujos

La figura 1a es una representación gráfica que muestra las tasas de asimilación de CO<sub>2</sub> ( $\square$ mol de C m<sup>2</sup> s<sup>-1</sup>) de varios genotipos de *A. thaliana*. Las mediciones se completaron en hojas en roseta durante las fases de floración prematura y floración. Las plantas se hicieron crecer en suelo fertilizado. Los genotipos de plantas se indican en el eje de las x, y las tasas de asimilación de CO<sub>2</sub> se indican en las ordenadas. Col indica un contexto genético del ecotipo Columbia. Ld indica un contexto genético del ecotipo Landsberg. Las plantas que expresan alelos *ERECTA* de tipo natural estaban o bien en un contexto Col (Col4-*ER*) o bien en un contexto Ld (Ld-*ER*). Las plantas que eran homocigóticas para un alelo *er* mutante estaban en o bien un contexto Ld (Ld-*er1*) o bien un contexto Col (Col-*er105* o Col-*er2* (línea 3401 en NASC, también denominada Col-*er106* por Torii y colaboradores (véase Lease *et al.* 2001, New Phytologist, 151:133-143)). Las plantas designadas como F1 (Col-*ER* x Ld-*er*) eran *ER/er1* heterocigóticas. Los datos indican que, en un contexto Col, la mutación *er105* conduce a una tasa de asimilación de CO<sub>2</sub> reducida, mientras que la mutación *er1* mejora la tasa de asimilación de CO<sub>2</sub> en un contexto Ld.

La figura 1b es una representación gráfica que muestra la conductancia estomática (mol de H<sub>2</sub>O m<sup>2</sup> s<sup>-1</sup>) de varios genotipos de *A. thaliana* (las mismas plantas que en la figura 1a). Los genotipos de plantas se indican en el eje de las x y son los mismos que los descritos en la leyenda de la figura 1a. Las conductancias estomáticas se indican en las ordenadas. Los datos indican que, en un contexto Col, la mutación *er2/er106* mejora significativamente la conductancia estomática, mientras que la mutación *er1* mejora significativamente la conductancia estomática en un contexto Ld.

La figura 1c es una representación gráfica que muestra la eficacia de transpiración de (mmol de C mol de H<sub>2</sub>O m<sup>2</sup> s<sup>-1</sup>) de varios genotipos de *A. thaliana*, tal como se determina mediante la razón de la tasa de asimilación de CO<sub>2</sub> con respecto a la conductancia estomática. Los genotipos de plantas se indican en el eje de las x y son los mismos que los descritos en la leyenda de la figura 1a. La eficacia de transpiración se indica en las ordenadas. Los datos indican que la eficacia de transpiración se mejora en plantas que expresan un alelo *ER* de tipo natural en relación con un alelo *er* mutante, en ambos contextos Ld y Col. La menor eficacia de transpiración se observó para plantas que son homocigóticas para el alelo *er105* (es decir Col-*er105*), lo que concuerda con el hecho de que este alelo inhibe la expresión de *ERECTA*. A partir de los datos en las figuras 1a-1c, resulta evidente que la menor eficacia de transpiración de plantas que expresan el alelo *er105* se debe en gran parte a una tasa de fijación de CO<sub>2</sub> reducida, mientras que para ambos alelos *er2/er106* y *er1*, la eficacia de transpiración reducida se debe en gran parte a una conductancia estomática mejorada. La eficacia de transpiración de la planta heterocigótica F1 era intermedia entre las eficacias de transpiración de sus progenitores, lo que sugiera una codominancia de estos alelos. Sin embargo, la F1, presenta una eficacia de transpiración más próxima a la del progenitor donador de polen, Ld-*er1*.

La figura 2a es una representación gráfica que muestra las densidades estomáticas (número de estomas mm<sup>-2</sup> de hoja) para varios genotipos de *A. thaliana* en tres experimentos independientes. Los contextos genéticos de las plantas se indican en el eje de las x (Col, Columbia; Ld, Landsberg), y las densidades estomáticas se indican en las ordenadas. Los genotipos de las plantas se indican en la parte superior de cada barra, tal como sigue: las plantas que expresan alelos *ERECTA* de tipo natural en un contexto Col fueron Col4ER o Col1ER (barras rayadas); las plantas que expresan alelos *ERECTA* de tipo natural en un contexto Ld fueron *ER* (barras vacías); las plantas que expresan alelos *erecta* mutantes en un contexto Col eran o bien *er105* o bien *er2/106* (recuadros Col rellenos); y las plantas que expresan el alelo *er1* mutante en un contexto Ld eran *er1* (recuadros Ld rellenos). Las columnas denominadas a, b son datos de dos experimentos en los que plantas se hicieron crecer en suelo en ausencia de fertilizante. El conjunto de columnas a la derecha de la figura son de un tercer experimento en el que las mismas plantas se hicieron crecer en suelo que comprende fertilizante. Los datos indican que, en un contexto Col, la mutación *er105* y la mutación *er2/106* mejoran la densidad estomática, que en parte es responsable de las conductancias estomáticas mejoradas y las eficacias de transpiración reducidas de plantas que expresan estos alelos (figuras 1b y 1c). El efecto general de estos alelos no depende del estado nutritivo del suelo. Por el contrario, el alelo *er1* sólo mejoró la densidad estomática de las plantas Ld cuando no había fertilizante, lo que sugiere que en este ecotipo la apertura estomática mejorada era responsable de las conductancias estomáticas mejoradas y las eficacias de transpiración reducidas medidas en el mutante *er1* con un suministro de nutrientes suficiente (figuras 1b, 1c). Por tanto, la mutación *er1* afecta tanto a la apertura estomática como a la densidad estomática pero las

contribuciones relativas de estos efectos a la conductancia estomática mejorada por unidad de área de hoja dependen de factores medioambientales y el estado nutritivo de las plantas, y del contexto genético.

La figura 2b es una representación gráfica que muestra el tamaño de célula epidérmica (área superficial,  $\mu\text{m}^2$ ) para varios genotipos de *A. thaliana* en tres experimentos independientes. Los contextos genéticos y genotipos de las plantas se indican en el eje de las x y en la parte superior de cada columna, respectivamente, como en la leyenda de la figura 2a. Las ordenadas indican el tamaño de célula epidérmica. Las columnas denominadas a,b son datos de dos experimentos en los que las plantas se hicieron crecer en suelo en ausencia de fertilizante. El conjunto de columnas a la derecha de la figura son de un tercer experimento en el que las mismas plantas se hicieron crecer en suelo que comprende fertilizante. Los datos indican que, en un contexto Col, la mutación *er105* y la mutación *er2/er106* reducen significativamente el tamaño de célula epidérmica, es decir aumentan el número de células epidérmicas por unidad de área de hoja. Esto revela que el gen *ER* presenta efectos sobre la histogénesis de las hojas que, más allá de sus consecuencias sobre las densidades estomáticas, también pueden afectar directamente a la capacidad de las hojas para realizar fotosíntesis y por tanto la eficacia de transpiración (figuras 1b y 1c). Los efectos generales de estos alelos no dependen del estado nutritivo del suelo. Por el contrario, en un contexto Ld, el alelo *er1* redujo el tamaño de célula epidérmica sólo cuando no había fertilizante.

La figura 2c es una representación gráfica que muestra el índice estomático para varios genotipos de *A. thaliana* en tres experimentos independientes. Los contextos genéticos y genotipos de las plantas se indican en el eje de las x y en la parte superior de cada columna, respectivamente, como en la leyenda de la figura 2a. Las ordenadas indican el índice estomático, tal como se determina a partir de la razón de la densidad estomática con respecto a la densidad de células epidérmicas. Las columnas denominadas a,b son datos de dos experimentos en los que las plantas se hicieron crecer en suelo en ausencia de fertilizante. El conjunto de columnas a la derecha de la figura son de un tercer experimento en el que las mismas plantas se hicieron crecer en suelo que comprende fertilizante. Los datos indican que las mutaciones *er* sometidas a prueba no modifican significativamente el índice estomático en el contexto Col (porque los aumentos en la densidad estomática están correlacionados con aumentos en los números de células epidérmicas en las plantas mutantes Col) pero sí en el contexto Landsberg. Por lo tanto, el gen *ER* parece modificar directamente el desarrollo estomático per se. Por tanto, tomadas conjuntamente las figuras 2ac muestran que el gen *ERECTA* presenta dos tipos de efectos sobre la conductancia estomática de las hojas: a) de desarrollo, b) biofísicos y/o bioquímicos. La expresión de estos efectos y el impacto sobre la tasa de transpiración varían con el contexto genético, lo que sugiere interacciones con otros genes que son polimórficos entre los ecotipos Col y Ld, y también con el estado nutritivo.

La figura 3 es una representación gráfica que muestra la composición de isótopos de carbono (eje de las y; en por mil, para rosetas vegetativas) para 7 series experimentales diferentes (números 1-7) llevadas a cabo en condiciones de cámara de crecimiento y condiciones de invernadero. Para cada serie, la barra izquierda muestra el valor medio de la composición de isótopos de carbono para las líneas que portan el alelo *ERECTA*, mientras que la barra derecha muestra el valor medio entre las líneas con el alelo *erecta*. En todos los casos, los valores de composición isotópica de  $\delta^{13}\text{C}$  para las líneas *er* son más negativos que para las líneas *ER*, lo que indica menores eficacias de transpiración.

La figura 4a es una representación gráfica que muestra el número de copias de gen *ERECTA* y los niveles de expresión en plantas de *A. thaliana* T2 transgénicas para un transgén *ER*. Estas líneas se generaron transformando el mutante Col-*er2/106* con el gen *ER* de tipo natural bajo el promotor 35S. Se averiguó la transformación eficaz y se cuantificaron los niveles de expresión de *ERECTA* en varios transformantes independientes utilizando PCR cuantitativa en tiempo real (ABI PRISM 7700, Sequence Detection System User Bulletin #2. 1997). El número de copias (eje de las y) se indica como una función de la línea vegetal, tras la normalización de *ERECTA* en relación con el número de copias de un gen control (gen de ARN ribosómico 18S). Se mostró independientemente que la expresión del gen de ARNr 18S no se veía afectada por los cambios en la expresión de *ER*. La línea 143 es un control nulo (sin inserto). Las líneas 145, 165, 169 y 279 son líneas transformadas que portan el alelo *ERECTA*. Todas las líneas transgénicas *ER*, excepto la línea 145, muestran un número de copias de ARNm aumentado: desde 4 hasta 9,5 veces aumentado en comparación con el control nulo.

La figura 4b es una representación gráfica que muestra el número de copias de gen *ERECTA* y los niveles de expresión en plantas de *A. thaliana* T2 transgénicas homocigóticas para un transgén *ER*, y generadas mediante la transformación del mutante Col-*er105*. Se averiguó la transformación eficaz y se cuantificaron los niveles de expresión de *ERECTA* en varios transformantes independientes utilizando PCR cuantitativa en tiempo real (ABI PRISM 7700, Sequence Detection System User Bulletin #2. 1997). El número de copias (eje de las y) se indica como una función de la línea vegetal, tras la normalización de *ERECTA* en relación con el número de copias de un gen control (gen de ARN ribosómico 18S). Se mostró independientemente que la expresión del gen de ARNr 18S no se veía afectada por los cambios en la expresión de *ER*. La línea 18 es una línea de control nulo (sin inserto de *ER*, es decir similar a Col-*er105*). Las líneas 8, 19, 29 y 61 son líneas transgénicas que portan el alelo *ERECTA*. Todas las líneas transgénicas *ER* muestran un número de copias de ARNm aumentado: desde 10 hasta 170 veces aumentado en comparación con el control nulo.

La figura 4c es una representación gráfica que muestra el número de copias de gen *ERECTA* y los niveles de

expresión en los ecotipos de *ER* Col y Ld y en una línea transgénica Ld-*ER* (3-7 K) generada mediante la transformación del ecotipo Ld-*er1* (NW20) con el gen de tipo natural *ER* bajo el control del promotor 35S. Se averiguó la transformación eficaz y se cuantificaron los niveles de expresión de *ERECTA* en varios transformantes independientes utilizando PCR cuantitativa en tiempo real (ABI PRISM 7700, Sequence Detection System User Bulletin #2. 1997). El número de copias (eje de las y) se indica como una función de la línea vegetal, tras la normalización de *ERECTA* en relación con el número de copias de un gen control (gen de ARN ribosómico 18S). Se mostró independientemente que la expresión del gen de ARNr 18S no se veía afectada por los cambios en la expresión de *ER*. Las líneas 933, 1093 y 3176 son los ecotipos Col-4, Col-0 y Col-1d de *ERECTA* Columbia no transformados.

La línea 105c es una línea Col-*er105* (deficiente para *ER*), utilizada para generar las líneas transgénicas mostradas en la figura 4b. Las líneas marcadas con 2c y 3401 en el eje de las x describen Col- *er2/106* (2 lotes de semillas, utilizadas para generar las líneas transgénicas mostradas en la figura 4a). La línea NW20 es Ld-*er1*. La línea 3-7K es un transformante Ld-*ER*, obtenido de la transformación de Lder1 con el alelo *ERECTA*. La línea 3177 es el ecotipo Ld-*ER*, casi isogénico con NW20.

La figura 5a es una representación gráfica de un primer experimento que muestra el número de copias del producto de transcripción de ARNm del gen *ERECTA* de arroz en diversos órganos/partes de la planta, cv Nipponbare. L= limbos de hojas maduras; YL= hojas jóvenes en expansión, todavía encerradas en vainas de hojas más viejas; R= raíz madura; YR= raíz joven; SH= vainas; INF= panícula joven no desplegada todavía encerrada en vainas; 07: panículas jóvenes. Se determinaron los números de copias de ARNm *ERECTA* de arroz mediante PCR cuantitativa en tiempo real, con ARNm 18S como gen de control interno para la normalización de los resultados. Los valores en el eje de las y describen las veces que aumenta el ARNm *ERECTA* de arroz en diversas partes en comparación con la muestra L (hojas maduras) fijada a un valor de 1 para la normalización. Los datos muestran un patrón de expresión similar que el gen *ERECTA* en *Arabidopsis* (véase Torii *et al.* 1996) es decir la expresión preferencial en tejidos meristemáticos jóvenes, especialmente en órganos reproductores.

La figura 5b es una representación gráfica de un segundo experimento que muestra el número de copias del producto de transcripción de ARNm del gen *ERECTA* de arroz en diversos órganos/partes de la planta. L= limbos de hojas maduras; YL= hojas jóvenes en expansión, todavía encerradas en vainas de hojas más viejas; R= raíz madura; YR= raíz joven; SH= vainas; INF= panícula joven no desplegada todavía encerrada en vainas; 07: panículas jóvenes. Se determinaron los números de copias de ARNm *ERECTA* de arroz mediante PCR cuantitativa en tiempo real, con ARNm 18S como gen de control interno para la normalización de los resultados. Los valores en el eje de las y describen las veces que aumenta el ARNm *ERECTA* de arroz en diversas partes en comparación con la muestra L (hojas maduras) fijada a un valor de 1 para la normalización. Los datos confirman los mostrados en la figura 5a.

La figura 6 es una representación gráfica que muestra la eficacia de transpiración de las hojas (mmol de C mol de  $H_2O^{-1}$ , figura 6a), calculada a partir de las mediciones directas de la tasa de asimilación de  $CO_2$  de las hojas ( $\square$ mol de C  $m^{-2} s^{-1}$ , figura 6b) y la conductancia estomática (mol de  $H_2O m^{-2} s^{-1}$ , figura 6c) mediante técnicas de intercambio de gases, con 350 ppm de  $CO_2$  (es decir igual que el  $[CO_2]$  ambiente durante el crecimiento de la plántula; barra izquierda en cada par de barras) y 500 ppm de  $CO_2$  (barra derecha en cada par de barras), para Lder1, y dos líneas Ld-*ER*: línea T2(+*ER*), una línea transgénica T2 homocigótica para un transgén *ER* en el contexto Ld-*er1* y la línea 3177, un ecotipo *ER* casi isogénico con Ld-*er1* (información del NASC Stock Centre). Los genotipos se muestran en la parte inferior de la figura. La temperatura de la hoja durante las mediciones se controló a 22°C, el déficit de presión de vapor de la hoja con respecto al aire a aproximadamente 8 mb.

La figura 7 es una representación gráfica que muestra la eficacia de transpiración de las hojas (mmol de C mol de  $H_2O^{-1}$ , figura 7a), calculada a partir de las mediciones directas de la tasa de asimilación de  $CO_2$  de las hojas ( $\square$ mol de C  $m^{-2} s^{-1}$ , figura 7b) y la conductancia estomática (mol de  $H_2O m^{-2} s^{-1}$ , figura 7c) mediante técnicas de intercambio de gases, con 350 ppm de  $CO_2$  (es decir igual que el  $[CO_2]$  ambiente durante el crecimiento de la plántula; barra izquierda en cada par de barras) y 500 ppm de  $CO_2$  (barra derecha en cada par de barras), para 4 genotipos: Col4 (*ER*) (par izquierdo), Ld (*er1*) (par derecho) y su progenie  $F_1$  (dos pares centrales). Los genotipos se muestran en la parte inferior de la figura.

La figura 8 es una representación gráfica que muestra la conductancia estomática y la anatomía epidérmica a 350 ppm de  $CO_2$  en los genotipos descritos en las figuras 6 y 7 y mostrados en la parte inferior de la figura. La inserción del transgén *ER* (línea T2+*ER*) provocó una disminución en la conductancia estomática en comparación con la línea Ld-*er1* (figura 8a), que en parte se debía a una disminución en la densidad estomática (véase la figura 8c). Estos dos efectos indican de nuevo complementación. Juntas, la figura 8b y 8c muestran que la disminución en la densidad estomática es relativamente más importante que la de en la densidad de células epidérmicas, lo que indica un efecto del transgén sobre el desarrollo de la epidermis.

La figura 9 es una representación gráfica que muestra una comparación de la densidad estomática y el área de células epidérmicas en una gama de líneas Col-*er* que portan mutaciones en el gen *ER* (barras 1 a 8 figura 9a; barras 1 a 7 en la figura 9b, mutantes *er105*, *er106*, 108, 111, 114, 116, 117, tal como se describen en Lease *et al.*

2001; un obsequio de Dr Keiko Torii) y en ecotipos de tipo natural Col-*ER* (barras 9-11 ó 8-10 en las figuras 9a y 9b, respectivamente: Col0, ecotipo de contexto para estos mutantes; Col1 Col4 (línea parental ColER para el análisis QTL de las RIL de Lister y Dean), dos líneas Ld\_er1 (NW20 y CS20, barras 12&13 y 11&12 en las figuras 9a y 9b respectivamente, dos líneas muy similares según NASC; NW20 es la otra línea parental para las RIL de Lister y Dean) y finalmente la línea T2+*ER*, una línea Ld-*ER* transgénica que porta el gen de tipo natural *ER* en el contexto Ld-*er1* (barra derecha de extremo en la figura).

La figura 10 es una representación gráfica que muestra la composición isotópica de carbono (por mil, eje de las y) en una gama de líneas (numeradas de 1 a 19 en el eje de las x): mutantes Col-*er* (línea 1-14);

el ecotipo de contexto Col0 (línea 15); líneas Ld-*er1* (líneas 16 y 17); un ecotipo Ld-*ER* casi isogénico con Ld-*er1* (línea 18, línea 3177 en NASC), y una línea Ld-*ER* T2 transgénica (línea numerada con 19) obtenida mediante la transformación del mutante Ld-*er1* con un constructo que porta el alelo *ER* de tipo natural. Los datos muestran que el alelo *ER* proporciona menos valores negativos lo que indica una eficacia de transpiración aumentada.

La figura 11 es una representación gráfica que muestra mediciones directas de la eficacia de transpiración en mutantes Col-*er* transformados con el transgén *ER*, tanto con alta y baja humedad del aire, tal como sucede durante eventos de alta temperatura que provocan o asociados con sequía. La eficacia de transpiración se midió mediante técnicas de intercambio de gases en hojas maduras de rosetas de *Arabidopsis* vegetativas, como una función de la diferencia de presión de vapor (vpd) de las hojas con respecto al aire, es decir la humedad del aire alrededor de las hojas. Cuanto mayor sea la vpd, más seco será el aire. Los círculos opacos describen mediciones para 5 líneas T2 transgénicas independientes homocigóticas para un transgén *ER*; estas líneas se generaron transformando el mutante Col-*er105* (cuadrados vacíos) con un constructo que porta el alelo *ER* bajo el control del promotor 35S. Los datos para las líneas nulas (es decir líneas que pasaron por transgénesis pero que no portan el transgén *ER*) se representan mediante cuadrados opacos. Esta figura demuestra complementación, por toda la gama de humedad sometida a prueba, siendo las eficacias de transpiración en las líneas *ER* T2 mayores que aquellas en el mutante Col-*er105* complementado, y similares a las medidas en el ecotipo Col0-*ER* (triángulos vacíos; ecotipo de contexto para Col-*er105*).

La figura 12 es una representación gráfica de una alineación de secuencias aisladas con toda la región codificante del ortólogo de trigo de *ERECTA*. La posición de cada una de las secuencias aisladas se muestra en relación con el ortólogo de trigo de *ERECTA*. Las secuencias se representan mediante o bien un SEC ID n°: o bien un número de registro de genes.

La figura 13 es una representación gráfica de una alineación de secuencias aisladas con toda la región codificante del ortólogo de maíz de *ERECTA*. La posición de cada una de las secuencias aisladas se muestra en relación con el ortólogo de maíz de *ERECTA*. Las secuencias se representan mediante o bien un SEC ID n°: O bien un número de registro de genes.

La figura 14 es una representación gráfica de una alineación de secuencias por pares de las proteínas *ERECTA* aisladas a partir de *Arabidopsis* (SEC ID n°: 2), maíz (SEC ID n°: 45), arroz (SEC ID n°: 3), sorgo (SEC ID n°: 5) y trigo (SEC ID n°: 20). La alineación se realizó utilizando la herramienta de alineación de múltiples secuencias CLUSTALW. Los residuos que se conservan entre todas las especies se indican mediante asteriscos (\*). La conservación de los grupos STA NEQK NHQK NDEQ QHRK MILV MILF HY o FYW se indica mediante “:”. La conservación de los grupos CSA ATV SAG STNK STPA SGND SNDEQK NDEQHK NEQHRK FVLIM HFY se indica mediante “.”. Los huecos se indican mediante “-”.

La figura 15 es una representación gráfica de un árbol filogenético que indica la relación entre cada una de las proteínas *ERECTA* aisladas a partir de *Arabidopsis* (SEC ID n°: 2), maíz (SEC ID n°: 45), arroz (SEC ID n°: 3), sorgo (SEC ID n°: 5) y trigo (SEC ID n°: 20).

### Descripción detallada de la invención

#### *Loci para la eficacia de transpiración y su identificación*

En la presente memoria se describe un locus asociado con la variación genética en la eficacia de transpiración de una planta, en el que dicho locus comprende una secuencia de nucleótidos ligada genéticamente a un locus *ERECTA* en el genoma de la planta.

Tal como se utiliza en la presente memoria, el término “locus” se considerará que significa la ubicación de uno o más genes en el genoma de una planta que afecta a una característica cuantitativa de la planta, en particular la eficacia de transpiración de una planta. En el presente contexto, una “característica cuantitativa” es un fenotipo de la planta para el que la variación fenotípica entre diferentes genotipos es continua y no puede separarse en clases diferenciadas, independientemente del número de genes que determinan o controlan el fenotipo, o la magnitud de los efectos genéticos que presenta un único gen en la determinación del fenotipo, o la magnitud de los efectos genéticos de genes que interaccionan.

“Asociado con la variación genética en la eficacia de transpiración de una planta” significa que un locus comprende uno o más genes que se expresan para determinar o regular la eficacia de transpiración de una planta, independientemente de la tasa real de transpiración lograda por la planta en una condición ambiental especificada.

5 Preferentemente, el locus está ligado a o comprende un alelo *ERECTA* o alelo *erecta*, o una parte que codifica para proteínas del mismo.

10 Tal como se utiliza en la presente memoria, el término “*ERECTA*” se considerará que hace referencia a un alelo de tipo natural que comprende los siguientes dominios, dominios GTIGYIDPEYARTS, GAAQGLAYLHHDC, y TNLSEKXIIGYGASSTVYKC, en los que X significa Y o H, o dominios con más de un 94% de identidad con estos dominios. Según el conocimiento de los inventores ninguna otra proteína comprende estos dominios. Los alelos *ERECTA* preferidos comprenden una secuencia de nucleótidos que presenta por lo menos aproximadamente una identidad de secuencia global del 55% con la región que codifica para proteínas de uno cualquiera de los alelos *ERECTA* mostrados a título de ejemplo descritos en la presente memoria, particularmente una cualquiera de las SEC ID n°: 1, 3, 5, 7, 9, 11, 13 ó 15. Preferentemente, el porcentaje de identidad con una cualquiera de dichas SEC ID n°: es por lo menos aproximadamente del 59-61%, o del 70% o del 80%, y más preferentemente por lo menos aproximadamente del 90%, y todavía más preferentemente por lo menos aproximadamente del 95% o del 99%.

20 Los alelos *ERECTA* preferidos se derivan de, o están presentes en, el genoma de una planta que es intolerante a la desecación o sequía, o poco adaptada al crecimiento en ambientes secos o áridos, o que presenta un vigor o crecimiento reducido durante periodos de precipitaciones reducidas o sequía, o del genoma de una planta con un aumento de la tasa de crecimiento o la duración del crecimiento o el reparto de C al brote y partes recogidas en condiciones de riego abundante.

25 Más preferentemente, un alelo *ERECTA* se deriva de, o está presente en, el genoma de una planta del género *Brassica*, una planta de cultivo extensivo, una gramínea perenne (por ejemplo de la subfamilia *Pooideae*, o la tribu *Poeae*) o un árbol. Incluso más preferentemente, un alelo *ERECTA* está presente en o se deriva del genoma de una planta seleccionada de entre el grupo constituido por cebada, trigo, centeno, sorgo, arroz, maíz, *Phalaris aquatica*, *Dactylus glomerata*, *Lolium perenne*, *Festuca arundinacea*, algodón, tomate, semilla de soja, colza, álamo y pino.

30 El término “*erecta*” se considerará que significa cualquier variante alélica del alelo *ERECTA* de tipo natural que modifica la eficacia de transpiración de una planta

35 Los alelos *erecta* preferidos incluyen los siguientes alelos *erecta* de *A. thaliana* derivados de las líneas *erecta* (*er*) Columbia (Col) y Landsberg.

Alelos <i>erecta</i> <sup>1</sup>	Posición genómica	Lesión	Dominio afectado
Ler <i>er</i> -1	2249	T→A	PK
Col <i>er</i> -101	6565	T→A	PK
Col <i>er</i> -102/106	6565	T→A	PK
Col <i>er</i> -103	846	G→A	LRR10
Col <i>er</i> -105	inserto de ADN foráneo entre +5 y +1056	inserción	alelo nulo
Col <i>er</i> -108	5649	G→A	
Col <i>er</i> -111	5749	G→A	región no traducida entre dominios LRR y transmembrana
Col <i>er</i> -113	3274	C→T	
Col <i>er</i> -114	6807	G→A	PK
Col <i>er</i> -115	3796	C→T	
Col <i>er</i> -116	6974	G→A	PK
Col <i>er</i> -117	5203	G→A	LRR18

<sup>1</sup> alelos descritos por Lease *et al.* 2001, *New Phytologist*, 151: 133-143, excepto para Ler *er*-1, Col *er*-103 y Col *er*-105 que se describieron en Torii *et al.*, 1996, *The Plant Cell* 8:73 5-746

40 En la presente memoria se describe un alelo *ERECTA* derivado de, o presente en, el genoma de una planta que es intolerante a la desecación o sequía, o está poco adaptada para el crecimiento en ambientes secos o áridos, o que padece crecimiento o vigor reducido durante periodos de precipitaciones reducidas o sequía, o del genoma de una planta con un aumento de la tasa de crecimiento o duración del crecimiento o el reparto de C al brote y partes recogidas en condiciones de riego abundante.

45 Más preferentemente, un alelo *ERECTA* se deriva de, o está presente en, el genoma de una planta del género *Brassica*, una planta de cultivo extensivo, una gramínea perenne (por ejemplo de la subfamilia *Pooideae*, o la tribu *Poeae*) o un árbol. Incluso más preferentemente, un alelo *ERECTA* está presente en o se deriva del genoma de una planta seleccionada de entre el grupo constituido por cebada, trigo, centeno, sorgo, arroz, maíz, *Phalaris aquatica*, *Dactylus glomerata*, *Lolium perenne*, *Festuca arundinacea*, algodón, tomate, semilla de soja, colza, álamo y pino.

50



Para los fines de nomenclatura, la secuencia de nucleótidos de la región que codifica para proteínas *ERECTA* de *Arabidopsis thaliana* y la región no traducida en 5' (UTR) y 3'-UTR, se proporciona en la presente memoria como SEC ID nº: 1. La secuencia de aminoácidos del polipéptido codificado mediante SEC ID nº: 1 se expone en la presente memoria como SEC ID nº: 2.

5 Un alelo *ERECTA* particularmente preferido de arroz (*Oryza sativa*) se deriva del cromosoma 6 de esta especie vegetal. Para los fines de nomenclatura, la región que codifica para proteínas del gen *ERECTA* de arroz se proporciona en la presente memoria como SEC ID nº: 3. La secuencia de aminoácidos del polipéptido codificado mediante SEC ID nº: 3 se expone en la presente memoria como SEC ID nº: 4.

10 Un gen *ERECTA* particularmente preferido derivado del genoma de *Sorghum bicolor*, se proporciona en la presente memoria como SEC ID nº: 5. La secuencia de aminoácidos del polipéptido codificado mediante SEC ID nº: 5 se expone en la presente memoria como SEC ID nº: 6.

15 Un gen *ERECTA* a título de ejemplo adicional derivado de *A. thaliana* se proporciona en la presente memoria como SEC ID nº: 7. La secuencia de aminoácidos del polipéptido codificado mediante SEC ID nº: 7 se expone en la presente memoria como SEC ID nº: 8.

20 Un gen *ERECTA* a título de ejemplo adicional derivado de *A. thaliana* se proporciona en la presente memoria como SEC ID nº: 9. La secuencia de aminoácidos del polipéptido codificado mediante SEC ID nº: 9 se expone en la presente memoria como SEC ID nº: 10.

Fragmentos de un gen *ERECTA* a título de ejemplo derivado del genoma de trigo se proporcionan en la presente memoria como las SEC ID nº: 11 a 18.

25 Un gen *ERECTA* a título de ejemplo derivado del genoma de trigo se proporciona en la presente memoria como SEC ID nº: 19. La secuencia de aminoácidos del polipéptido codificado mediante SEC ID nº: 19 se expone en la presente memoria como SEC ID nº: 20.

30 Fragmentos de un gen *ERECTA* a título de ejemplo derivado del genoma de maíz se proporcionan en la presente memoria como las SEC ID nº: 21 a 43.

Un gen *ERECTA* a título de ejemplo derivado del genoma de maíz se proporciona en la presente memoria como SEC ID nº: 44. La secuencia de aminoácidos del polipéptido codificado mediante SEC ID nº: 44 se expone en la presente memoria como SEC ID nº: 44.

40 Múltiples genes pueden estar ligados genéticamente o correlacionarse con el locus *ERECTA* especificado en el cromosoma 2. Sin restringirse a ninguna teoría o modo de acción, tales múltiples genes ligados pueden interaccionar, tal como, por ejemplo, mediante interacción epistática, para determinar el fenotipo de eficacia de transpiración.

Diferentes alelos de cualquier gen pueden ligarse al locus *ERECTA*, en el que dicho alelo se expresa para determinar el fenotipo de eficacia de transpiración. Tales alelos pueden identificarse detectando un fenotipo de eficacia de transpiración particular que está vinculado a la expresión del alelo particular. Alternativamente, o además, los diferentes alelos ligados a un locus se identifican detectando un polimorfismo estructural en el ADN (por ejemplo un polimorfismo de la longitud de los fragmentos de restricción (RFLP), polimorfismo de la longitud de los fragmentos amplificados (AFLP), polimorfismo en la conformación de cadenas sencillas (SSCP) y similares), que está vinculado a un fenotipo de eficacia de transpiración particular.

50 Los genes y/o alelos que interaccionan pueden estar ligados genéticamente a un locus *ERECTA* y se expresan para determinar un fenotipo de eficacia de transpiración. Tales genes y/o alelos que interaccionan ligados se correlacionarán con un locus *ERECTA* y estarán asociados con la eficacia de transpiración de esa planta. Preferentemente, tales genes y/o alelos que interaccionan comprenden una parte que codifica para proteínas de un gen situado dentro del locus *ERECTA* del genoma que está asociado con la eficacia de transpiración de esa planta.

55 En la presente memoria se describen homólogos y/o ortólogos de los alelos mostrados a título de ejemplo. Los expertos en la materia apreciarán que los términos "homólogo" y "ortólogo" se refieren a unidades equivalentes funcionales. En el presente contexto, un homólogo u ortólogo de un gen que se correlaciona con un locus *ERECTA* se considerará que significa cualquier gen de una especie vegetal que es funcionalmente equivalente a un gen que se correlaciona con un locus *ERECTA* mostrado a título de ejemplo, y comprende una región que codifica para proteínas en su genoma de planta nativa que comparte un grado de similitud o identidad estructural con una región que codifica para proteínas del gen *ERECTA* mostrado a título de ejemplo.

65 Preferentemente, un gen homólogo u ortólogo de una planta distinta de *A. thaliana* estará asociado con la eficacia de transpiración de dicha planta y estará ligado a una región que codifica para proteínas en su genoma de planta nativa que comprende una secuencia de nucleótidos que presenta por lo menos aproximadamente una identidad de

secuencia global del 55% con una región que codifica para proteínas ligada con el locus *ERECTA*. Incluso más preferentemente, el porcentaje de identidad será de por lo menos aproximadamente el 59-61% o el 70% o el 80%, todavía más preferentemente por lo menos aproximadamente el 90%, e incluso todavía más preferentemente por lo menos aproximadamente el 95%.

5 Al determinar si dos secuencias de nucleótidos se encuentran o no dentro de una limitación de porcentaje de identidad particular mencionada en la presente memoria, los expertos en la materia apreciarán que es necesario realizar una comparación lado a lado o una alineación múltiple de secuencias.

10 En tales comparaciones o alineaciones, pueden surgir diferencias en la ubicación de los residuos no idénticos, dependiendo del algoritmo utilizado para realizar la alineación. En el presente contexto, se considerará que la referencia a un porcentaje de identidad entre dos o más secuencias de nucleótidos se refiere al número de residuos idénticos entre dichas secuencias tal como se determina utilizando cualquier algoritmo convencional conocido por los expertos en la materia. Por ejemplo, pueden alinearse las secuencias de nucleótidos y calcularse su identidad  
15 utilizando el programa BESTFIT u otro programa apropiado de Computer Genetics Group, Inc., University Research Park, Madison, Wisconsin, Estados Unidos de America (Devereaux *et al.*, Nucl. Acids Res. 12, 387-395, 1984). En la determinación del porcentaje de identidad de secuencias de nucleótidos utilizando un programa conocido en la técnica o descrito en la presente memoria, es preferible que se utilicen parámetros por defecto.

20 Alternativamente, o además, un alelo *ERECTA* o *erecta* homólogo u ortólogo estará asociado con la eficacia de transpiración de una planta y ligado a una región que codifica para proteínas en su genoma de planta nativa que comprende una secuencia de nucleótidos que codifica para un polipéptido que presenta por lo menos aproximadamente una identidad de secuencia global del 55% con un polipéptido codificado por una región que  
25 codifica para proteínas ligada al locus *ERECTA*. Preferentemente, el porcentaje de identidad a nivel de aminoácidos será de por lo menos aproximadamente el 59-61% o el 70% o el 80%, más preferentemente por lo menos aproximadamente el 90%, y todavía más preferentemente por lo menos aproximadamente el 95%.

Al determinar si dos secuencias de aminoácidos se encuentran o no dentro de estos límites de porcentajes, los expertos en la materia apreciarán que es necesario realizar una comparación lado a lado o una alineación múltiples  
30 de secuencias. En tales comparaciones o alineaciones, surgirán diferencias en la ubicación de los residuos no idénticos, dependiendo del algoritmo utilizado para realizar la alineación. En el presente contexto, se considerará que la referencia a un porcentaje de identidad o similitud entre dos o más secuencias de aminoácidos se refiere al número de residuos idénticos y similares respectivamente, entre dichas secuencias tal como se determina utilizando cualquier algoritmo convencional conocido por los expertos en la materia. Por ejemplo, pueden calcularse las  
35 identidades o similitudes de secuencias de aminoácidos utilizando el programa GAP y/o alinearse utilizando el programa PILEUP de Computer Genetics Group, Inc., University Research Park, Madison, Wisconsin, Estados Unidos de America (Devereaux *et al.*, 1984, citado anteriormente). El programa GAP utiliza el algoritmo de Needleman y Wunsch, J. Mol. Biol. 48, 443-453, 1970, para maximizar el número de residuos idénticos/similares y para minimizar el número y la longitud de los huecos de secuencias en la alineación. Alternativamente o además,  
40 cuando se comparan más de dos secuencias de aminoácidos, se utiliza el programa ClustalW de Thompson *et al.*, Nucl. Acids Res. 22, 4673-4680, 1994. En la determinación del porcentaje de identidad de secuencias de aminoácidos utilizando un programa conocido en la técnica o descrito en la presente memoria, es preferible que se utilicen parámetros por defecto.

45 Alternativamente, o además, un alelo *ERECTA* o *erecta* homólogo u ortólogo estará asociado con la eficacia de transpiración de una planta y ligado a una región que codifica para proteínas en su genoma de planta nativa que se hibrida con un ácido nucleico que comprende una secuencia complementaria a una región que codifica para proteínas ligada a un locus *ERECTA*, tal como, por ejemplo, de *A. thaliana*, arroz, sorgo, maíz, trigo o arroz. Preferentemente, tales homólogos u ortólogos se identificarán mediante hibridación en condiciones de por lo menos  
50 baja rigurosidad, y más preferentemente en condiciones de hibridación de por lo menos rigurosidad moderada o alta rigurosidad.

Para definir el nivel de rigurosidad, una baja rigurosidad se define en la presente memoria como una hibridación o un lavado llevado a cabo en un tampón 6xSSC, SDS al 0,1% (p/v) a 28°C o alternativamente, tal como se muestra a  
55 título de ejemplo en la presente memoria. Generalmente, la rigurosidad se aumenta reduciendo la concentración de sal en el tampón de hibridación o lavado, tal como, por ejemplo, reduciendo la concentración de SSC. Alternativamente, o además, la rigurosidad se aumenta, aumentando la concentración de detergente (por ejemplo SDS). Alternativamente, o además, la rigurosidad se aumenta, aumentando la temperatura de la hibridación o el lavado. Por ejemplo, una rigurosidad moderada puede realizarse utilizando tampón de 0,2xSSC a 2xSSC, SDS al  
60 0,1% (p/v), a una temperatura de aproximadamente 42°C a aproximadamente 65°C. De manera similar, una alta rigurosidad puede realizarse utilizando tampón de 0,1xSSC a 0,2xSSC, SDS al 0,1% (p/v), a una temperatura de por lo menos 55°C. Las condiciones para realizar reacciones de hibridación de ácido nucleico y el posterior lavado de membrana, se entienden bien por un experto habitual en la materia. Sólo para fines de aclaración adicional, se encuentra una referencia a los parámetros que afectan a la hibridación entre moléculas de ácido nucleico en  
65 Ausubel *et al.*, En: Current Protocols in Molecular Biology, Greene/Wiley, Nueva York, EE.UU., 1992.

Se han descrito varios procedimientos de mapeo para determinar loci útiles y estimar sus efectos (por ejemplo Edwards *et al.*, *Genetics* 116, 113-125, 1987; Haley y Knott, *Heredity* 69, 315-324, 1992; Jiang y Zeng, *Genetics* 140, 1111-1127, 1995; Lander y Botstein, *Genetics* 121, 185-199, 1989; Jansen y Stam, *Genetics* 136, 1447-1455, 1994; Utz y Melchinger, En: *Biometrics in Plant Breeding: Applications of Molecular Markers*. Proc. Ninth Meeting of the EUCARPIA Section Biometrics in Plant Breeding, 6-8 de julio de 1994, Wageningen, Países Bajos, (J.W. van Ooijen y J. Jansen, eds), págs. 195-204, 1994; Zeng, *Genetics* 136, 1457-1468, 1994). En el presente contexto, estos procedimientos se aplican para identificar el/los componente(s) principal(es) de la varianza genética total que contribuye(n) a la variación en la eficacia de transpiración de una planta, tal como se determina, por ejemplo, mediante la medición de la discriminación de isótopos de carbono ( $\Delta$ ). Más particularmente, la segregación de marcadores conocidos se utiliza para mapear y/o caracterizar un locus subyacente asociado con la eficacia de transpiración. El procedimiento de locus implica buscar asociaciones entre los marcadores moleculares segregantes y la eficacia de transpiración en una población segregante de plantas, para identificar el ligamiento del marcador con el locus.

Para descubrir un ligamiento marcador/locus, se requiere una población segregante. Pueden utilizarse poblaciones experimentales, tales como, por ejemplo, una generación F<sub>2</sub>, una población de retrocruzamiento (BC), líneas endogámicas recombinantes (RIL) o línea doble haploide (DHL), como población de mapeo. El análisis de segregantes en masa, para la rápida detección de marcadores en regiones genómicas específicas utilizando poblaciones segregantes, se describe por Michelmoore *et al.*, *Proc. Natl Acad. Sci. (EE.UU.)* 88, 9828-9832, 1991. En el caso de las poblaciones de mapeo F<sub>2</sub>, las plantas F<sub>2</sub> se utilizan para determinar el genotipo, y las familias F<sub>2</sub> para determinar el fenotipo. Los expertos en la material conocerán líneas endogámicas recombinantes se producen mediante un descendiente de semilla única. Las líneas endogámicas recombinantes, tales como, por ejemplo, las RIL F<sub>9</sub> de *A. thaliana* (por ejemplo Lister y Dean, *Plant J.*, 4, 745-750, 1993). Se utilizan líneas casi isogénicas (NIL) para mapeo fino, y para determinar el efecto de un locus particular sobre la eficacia de transpiración. Una ventaja de las líneas endogámicas recombinantes y las líneas doble haploides es que son poblaciones permanentes, y como consecuencia, proporcionan la replicación de la contribución de un locus particular al fenotipo de eficacia de transpiración.

Tal como para procedimientos estadísticos, se utiliza un análisis de marcadores individuales (análisis puntual) para detectar un locus en las proximidades de un marcador genético individual. Las eficacias de transpiración medias de una población de plantas que se segrega para un marcador particular, se comparan según la clase de marcador. La diferencia entre dos eficacias de transpiración medias proporciona una estimación del efecto fenotípico de sustituir un alelo por otro alelo en el locus. Para determinar si el efecto fenotípico inferido es significativamente diferente de cero o no, se utiliza una prueba estadística sencilla, tal como la prueba de la t o la prueba de la F. Un valor significativo indica que un locus está ubicado en las proximidades del marcador. El análisis de un solo punto no requiere un mapa de ligamiento molecular completo. Cuanto más alejado esté el locus del marcador, menos probable será detectarlo estadísticamente, como consecuencia de la recombinación entre el marcador y el gen.

En el enfoque de Anova, prueba de la t o GLM, la asociación entre el genotipo del marcador y el fenotipo de eficacia de transpiración comprende:

- (i) clasificar la progenie de una población segregante de plantas mediante el genotipo del marcador, tal como por ejemplo, utilizando RFLP, AFLP, SSCP o análisis de microsátélites, estableciendo de ese modo clases de plantas;
- (ii) comparar las eficacias de transpiración medias de clases de plantas en la población segregante, utilizando una prueba de la t, GLM o ANOVA; y
- (iii) determinar la significación de las diferencias en la media en (ii), una diferencia significativa indicando que el marcador está ligado al locus para la eficacia de transpiración.

Tal como apreciarán los expertos en la materia, la diferencia entre las medias de las clases proporciona una estimación del efecto del locus en la determinación de la eficacia de transpiración de una clase.

En el enfoque de regresión, la asociación entre el genotipo y el fenotipo del marcador está determinada por un proceso que comprende:

- (i) asignar códigos numéricos a genotipos de marcador; y
- (ii) determinar el valor de regresión r para la eficacia de transpiración frente a los códigos, un valor significativo para r indicando que el marcador está ligado al locus para la eficacia de transpiración, y proporcionando la pendiente de regresión una estimación del efecto de un locus particular sobre la eficacia de transpiración.

Para el mapeo por intervalos de QTL, puede utilizarse el algoritmo Mapmaker desarrollado por Lincoln *et al.*, *Constructing genetic linkage maps with MAPMAKER/EXP version 3.0: A tutorial and reference manual*. Whitehead Institute for Biomedical Research, Cambridge, MA, EE.UU., 1993. El principio detrás del mapeo por intervalos es

para someter a prueba un modelo para determinar la presencia de un QTL en muchas posiciones entre dos loci de marcador mapeados. Este modelo es un ajuste de un presunto QTL con respecto a la eficacia de transpiración, en el que la idoneidad del ajuste se somete a prueba determinando la probabilidad máxima de que un QTL para la eficacia de transpiración se encuentre entre dos marcadores segregantes. Por ejemplo, en el caso de un QTL ubicado entre dos marcadores segregantes, los genotipos de marcador de 2 loci de una progenie segregante contendrán cada uno mezclas de los genotipos de QTL. Por lo tanto, es posible buscar los parámetros de loci que se aproximan de la mejor manera a la distribución en la eficacia de transpiración para cada clase de marcador. Los modelos se evaluaron calculando la probabilidad de las distribuciones observadas con y sin ajustar un efecto de QTL. La posición de mapeo de un QTL se determina como la probabilidad máxima a partir de la distribución de los valores de probabilidad (puntuaciones LOD: razón de la probabilidad de que el efecto se produzca mediante ligación:probabilidad de que el efecto se produzca por azar), calculada para cada locus.

El mapeo por intervalos mediante regresión (Haley y Knott., *Heredity* 69, 315-324, 1992) es una simplificación del procedimiento de probabilidad máxima citado anteriormente en el que se realiza una regresión o un análisis de QTL básico en genotipos de marcador codificado, excepto porque los fenotipos se someten a regresión basándose en la probabilidad de un genotipo de QTL tal como se determina a partir del ligamiento entre la eficacia de transpiración y los marcadores flanqueantes más próximos. En la mayoría de los casos, el mapeo por regresión proporciona estimaciones de la posición de QTL y el efecto que son casi idénticas con las proporcionadas mediante el procedimiento de probabilidad máxima. La aproximación sólo se desvía en lugares en los que hay huecos grandes, o faltan muchos genotipos.

En el procedimiento de mapeo por intervalos compuestos (CIM) (Jansen y Stam, *Genetics* 136, 1447-1455, 1994; Utz y Melchinger, 1994, citado anteriormente; Zeng, *Genetics* 136, 1457-1468, 1994), el análisis se realiza de la manera habitual, excepto porque se tiene en cuenta la varianza de otros QTL incluyendo la regresión parcial que proporciona más potencia y precisión que un mapeo por intervalos simples, porque los efectos de otros QTL no están presentes como varianza residual. El CIM puede eliminar el sesgo que puede provocarse mediante los QTL que están ligados a la posición que está sometiéndose a prueba.

Se utiliza software disponible públicamente para mapear un locus para la eficacia de transpiración. Tal software incluye, por ejemplo, lo siguiente:

- (i) MapMaker/QTL (<ftp://genome.wi.mit.edu/pub/mapmaker3/>), para analizar datos de retrocruzamiento o F2 utilizando mapeo por intervalos convencional;
- (ii) MQTL, para mapeo por intervalos compuestos en múltiples entornos o para realizar mapeo por intervalos simples utilizando una progenie homocigótica (por ejemplo doble haploides o líneas endogámicas recombinantes);
- (iii) PLABQTL (Utz y Melchinger, PLABlocus Version 1.0. A computer program to map QTL, Institut für Pflanzenzüchtung, Saatgutforschung und Populationsgenetik, Universität Hohenheim, 70593 Stuttgart, Alemania, 1995; <http://www.uni-hohenheim.de/~ipspwww/soft.html>) para mapeo por intervalos compuestos y mapeo por intervalos simples de un locus en poblaciones de mapeo derivadas de un cruzamiento biparental mediante autofecundación, o en doble haploides;
- (iv) QTL Cartographer (<http://statgen.mcsu.edu/qtlcart/cartographer.html>) para regresión de marcadores individuales, mapeo por intervalos o mapeo por intervalos compuestos, utilizando poblaciones de retrocruzamiento o F2;
- (v) MapQTL (<http://www.cpro.dlo.nl/cbw/>); Qgene para realizar o bien una regresión de marcadores individuales o bien una regresión por intervalos para mapear loci; y
- (vi) SAS para detectar un locus identificando asociaciones entre el genotipo de marcador y la eficacia de transpiración mediante un enfoque de análisis de marcadores individuales tal como ANOVA, prueba de la t, GLM o REG.

En una forma de realización particularmente preferida, QTL cartographer o MQTL se utiliza para identificar un locus asociado con la eficacia de transpiración de plantas.

Los expertos en la materia también apreciarán que es posible detectar múltiples alelos o genes que interaccionan para un rasgo particular, tal como, por ejemplo, utilizando enfoques de mapeo por intervalos compuestos. Para alcanzar este extremo, el mapeo por intervalos compuestos puede repetirse para buscar loci adicionales. Alternativamente, o además, dos o más regiones distintas del genoma pueden nombrarse como loci candidatos, y construirse una matriz de relación de gametos para cada locus candidato, y realizarse una regresión de 2 locus para cada par de loci, determinando el mejor ajuste para los efectos que interaccionan entre los dos loci o alelos en esos loci, incluyendo cualquier efecto aditivo o de dominancia. El algoritmo descrito por Carlborg *et al.*, *Genetics* (2000) puede utilizarse para mapeo simultáneo. En el presente contexto, se realiza un análisis de este tipo con referencia a la segregación de los fenotipos de eficacia de transpiración en la población segregante.

*Utilización del locus ERECTA para mejorar la eficacia de transpiración de plantas*

Como resulta conocido por los expertos en la materia, un único locus, si está presente en el genoma de una planta, puede presentar una influencia significativa sobre el fenotipo de la planta. Por ejemplo, Grandillo *et al.*, Theor. Appl. Genet. 99, 978-987, 1999, mostraron que para tomate una selección realizada a partir de un total de 28 loci que determinaban el peso y el tamaño de la fruta explicaba el 20% de la varianza fenotípica total en este rasgo.

Por lo tanto, en la presente memoria se describe un procedimiento para seleccionar una planta que presenta una eficacia de transpiración mejorada, que comprende:

- (a) identificar un locus asociado con una variación genética en la eficacia de transpiración en una planta; y
- (b) seleccionar una planta que comprende o expresa un gen que se correlaciona con el locus.

Mediante "eficacia de transpiración mejorada" quiere decirse que la planta pierde menos agua por unidad de materia seca producida, o alternativamente, produce una cantidad aumentada de materia seca por unidad de agua transpirada, o alternativamente, fija una cantidad aumentada de carbono por unidad de agua transpirada, en relación con una planta equivalente. Mediante "planta equivalente" quiere decirse una planta que presenta un contexto genético similar o casi idéntico, tal como, por ejemplo, una planta casi isogénica, una hermana o una parental.

Según este procedimiento, un locus se identifica mediante medios de mapeo de locus convencionales y/o mediante búsqueda de homología para genes que se correlacionan con el locus *ERECTA* en el cromosoma 2 del genoma de *A. thaliana*, tal como, por ejemplo, buscando alelos *ERECTA* o alelos *erecta* a partir de una variedad de plantas, tales como, por ejemplo, arroz, trigo, sorgo y maíz, tal como se describe anteriormente en la presente memoria.

Preferentemente, para seleccionar una planta que comprende o expresa el gen apropiado, se utiliza la selección asistida por marcadores (MAS). Tal como sabrán los expertos en la materia, una vez identificado un locus particular, los marcadores genéticos o físicos que están ligados al locus pueden identificarse y utilizarse fácilmente para confirmar la presencia del locus en poblaciones en cultivo. Para un locus que está flanqueado por dos marcadores estrechamente ligados que se recombinan sólo a una baja frecuencia, la presencia de los marcadores flanqueantes es indicativa de la presencia del locus.

Para la selección asistida por marcadores, se prefiere que el marcador sea un marcador genético (por ejemplo un gen o alelo), o un marcador físico (por ejemplo vellosoidad en las hojas o forma de la vaina), o un marcador molecular tal como, por ejemplo, un polimorfismo de la longitud de los fragmentos de restricción (RFLP), una restricción (RAPD), polimorfismo de la longitud de los fragmentos amplificados (AFLP) o una repetición de secuencia corta (SSR) tal como un microsatélite, o SNP. También se encuentra dentro del alcance de la invención utilizar cualquier sonda de hibridación o cebador de amplificación que comprenda por lo menos aproximadamente 10 nucleótidos de longitud derivado de una región cromosómica que está ligada en el genoma de una planta a un locus *ERECTA*, como marcador para seleccionar plantas. Los expertos en la materia podrán determinar fácilmente tales sondas o cebadores basándose en la divulgación en la presente memoria, particularmente para aquellos genomas de plantas que pueden presentar una secuencia cromosómica suficiente en la región de interés en el genoma (por ejemplo *A. thaliana*, arroz, algodón, cebada, trigo, sorgo, maíz, tomate, etc.).

Para los marcadores flanqueantes que no están estrechamente ligados, de modo que hay una gran distancia de recombinación entre los mismos, la presencia del gen apropiado se evalúa identificando aquellas plantas que presentan ambos marcadores flanqueantes y seleccionando después de entre aquellas plantas que presentan una eficacia de transpiración mejorada. Naturalmente, cuanto mayor sea la distancia entre dos marcadores, mayor será la población de plantas requerida para identificar una planta que presenta ambos marcadores, el locus que interviene y un gen dentro de dicho locus. Los expertos en la materia podrán determinar fácilmente el tamaño de la población requerido para identificar una planta que presenta una eficacia de transpiración particular, basándose en las unidades de recombinación (cM) entre dos marcadores.

La eficacia de transpiración se determina mediante cualquier medio conocido por el experto en la materia. Preferentemente, la eficacia de transpiración se determina midiendo la acumulación de materia seca en la planta mediante medios gravimétricos, o midiendo la pérdida de agua, o la razón de la tasa de asimilación de CO<sub>2</sub> con respecto a la conductancia estomática.

La eficacia de transpiración puede determinarse directamente, midiendo la razón de carbono fijado (tasa de asimilación de carbono) con respecto a la pérdida de agua (tasa de transpiración).

La eficacia de transpiración puede determinarse indirectamente a partir del valor de la discriminación de isótopos de carbono ( $\Delta$ ). Farquhar *et al.*, Aust. J. Plant Physiol. 9, 121-137, 1982, mostraron que la discriminación de isótopos de carbono ( $\Delta$ ; una medida del grado en que la razón de <sup>13</sup>C/<sup>12</sup>C de la materia orgánica es menor que la de CO<sub>2</sub> en el aire fuente), es una medida indirecta eficaz de la eficacia de transpiración. La discriminación, ( $\Delta$ ), es aproximadamente la razón isotópica de carbono en el CO<sub>2</sub> fuente menos la de carbono orgánico de la planta. En un

experimento particular, el CO<sub>2</sub> fuente es común para todos los genotipos. La determinación de la eficacia de transpiración de esta manera se basa en la constancia de la razón de <sup>13</sup>C:<sup>12</sup>C atmosférico (aproximadamente 1,1:98,9) y el hallazgo de que las enzimas ribulosa bifsosfato carboxilasa (Rubisco) se discriminan frente a la utilización de <sup>13</sup>C. Por tanto, en plantas C<sub>3</sub> el <sup>13</sup>CO<sub>2</sub> se asimila de manera menos eficaz que el <sup>12</sup>CO<sub>2</sub>, y el <sup>13</sup>CO<sub>2</sub> que queda tiende a difundir de vuelta a través de los estomas dentro y fuera de la hoja. Sin embargo, cuando los estomas pasan a estar casi cerrados, la retrodifusión relativa del <sup>13</sup>CO<sub>2</sub> es más difícil de conseguir y las concentraciones intracelulares de <sup>13</sup>CO<sub>2</sub> aumentan, aumentando de ese modo la proporción de este isótopo que se incorpora en 3-fosfoglicerato y posteriormente en materia seca. Como consecuencia, la discriminación de isótopos de carbono (Δ) es la mayor cuando la tasa de asimilación de CO<sub>2</sub> global durante la fotosíntesis (A) es pequeña, y la conductancia estomática (g<sub>w</sub>) con respecto al vapor de agua es grande. Esta relación se representa mediante el siguiente algoritmo:

$$\Delta (\%) = 27 - 36A / (g_w \times C_a)$$

en el que C<sub>a</sub> es la concentración de CO<sub>2</sub> ambiente (es decir [<sup>12</sup>CO<sub>2</sub>+<sup>13</sup>CO<sub>2</sub>]). La discriminación, Δ, es aproximadamente la composición isotópica del CO<sub>2</sub> fuente menos la del carbono orgánico de la planta.

Para una planta C<sub>3</sub> que presenta un valor en el intervalo de aproximadamente el 4,5 % a aproximadamente el 6,7 % para el término 36A/(g<sub>w</sub> x C<sub>a</sub>), un cambio del 1 % en la discriminación de isótopos de carbono (Δ) corresponde a un cambio en la eficacia de transpiración en el intervalo de aproximadamente el 22% a aproximadamente el 15%, respectivamente.

La relación negativa entre la discriminación de isótopos de carbono (Δ) y la eficacia de transpiración se ha establecido para muchas especies de plantas C<sub>3</sub>, incluyendo trigo (Farquhar y Richards, Aust. J. Plant Physiol. 11, 539-552, 1984; Farquhar *et al.*, Ann. Rev. Plant Physiol. 40,388-397, 1989), *Stylosanthes* (Thumma *et al.*, Proc. 9th Aust. Agronomy Conf., Wagga Wagga New South Wales, Australia, 1998), algodón, cebada y arroz. Por lo tanto, un valor de la discriminación de isótopos de carbono (Δ) menor para una planta de prueba en relación con una planta equivalente es indicativo de una eficacia de transpiración mejorada.

En especies C<sub>4</sub>, como el maíz, los coeficientes en la ecuación anterior son diferentes (Farquhar 1983, Australian Journal of Plant Physiology, 10:205-226; Henderson *et al.*,1992, Aust. J. Plant Physiol. 19: 263-285):

$$\Delta (\%) = 1 + 5A / (g_w \times C_a).$$

Una diferencia del 1% en Δ corresponde a una diferencia de aproximadamente el 38% en la eficacia de transpiración. La relación entre Δ y la eficacia de transpiración es positiva. El <sup>13</sup>C se acumula preferentemente en bicarbonato, el sustrato para la carboxilación de PEP, y así la discriminación frente a <sup>13</sup>C es la menor cuando A es pequeña y g<sub>w</sub> es grande. Sin embargo, a medida que aumenta la pérdida de CO<sub>2</sub> a partir de la vaina fascicular, las plantas C<sub>4</sub> se comportan más como las plantas C<sub>3</sub>.

Alternativamente, o además, la eficacia de transpiración se determina mediante otro indicador, tal como, por ejemplo, la temperatura de las hojas, el contenido en ceniza, el contenido en minerales, o el peso específico de la hoja (materia seca por unidad de área de hoja). Por ejemplo, el peso específico de la hoja se correlaciona positivamente con la eficacia de transpiración en cacahuetes y otras especies (Virgona *et al.*, Aust. J. Plant Physiol., 17, 207-214, 1990; Wright *et al.*, Crop Sci 34, 92-97, 1994). Por lo tanto, un mayor peso específico de la hoja o una mayor tasa de ganancia de carbono para una planta de prueba en relación con una planta equivalente es indicativo de una eficacia de transpiración mejorada de la planta de prueba.

La presencia del locus puede establecerse hibridando una sonda o cebador que está ligado a un locus *ERECTA*, tal como, por ejemplo, una sonda o cebador que se hibrida con la región del cromosoma 2 identificada de *A. thaliana* o la región del cromosoma 6 identificada de arroz.

Preferentemente, la presencia del locus se establece hibridando una sonda o cebador derivado de cualquiera o más de las SEC ID nº: 1, 3, 5, 7, 9, 11 a 19 ó 21 a 44 o de un gen homólogo en otra planta, o una secuencia complementaria a una secuencia de este tipo, con el ADN genómico de la planta, y detectando la hibridación utilizando medios de detección. En una forma de realización, la detección de la hibridación se realiza preferentemente marcando una sonda con una molécula indicadora que puede producir una señal identificable, antes de la hibridación, y detectando después la señal tras la hibridación. Las moléculas indicadoras preferidas incluyen nucleótido trifosfatos marcados radiactivamente y moléculas biotiniladas. Preferentemente, los variantes de los genes mostrados a título de ejemplo en la presente memoria, incluyendo los equivalentes genómicos, se aíslan mediante hibridación con una rigurosidad moderada o más preferentemente, en condiciones de alta rigurosidad, con la sonda.

Alternativamente, o además, la hibridación puede detectarse utilizando cualquier formato de la reacción en cadena

de la polimerasa (PCR), incluyendo AFLP. Para la PCR, dos moléculas cebadoras de ácido nucleico no complementarias que comprenden por lo menos aproximadamente 20 nucleótidos de longitud, y más preferentemente por lo menos 30 nucleótidos de longitud se hibridan con diferentes hebras de una molécula molde de ácido nucleico, y copias de la molécula de ácido nucleico específica del molde se amplifican enzimáticamente. Varios formatos de PCR se describen en McPherson *et al.*, En: PCR A Practical Approach, IRL Press, Oxford University Press, Oxford, Reino Unido, 1991.

Para mejorar la eficacia de transpiración de una planta en el que el locus es polimórfico, tal como, por ejemplo, un alelo, el procedimiento citado anteriormente se modifica para incluir la detección del/de los alelo(s) específico(s) vinculados a la mejora deseada. En la presente memoria se da a conocer un procedimiento para seleccionar una planta que presenta una eficacia de transpiración mejorada, que comprende:

(d) identificar un locus asociado con una variación genética en la eficacia de transpiración en una planta;

(e) identificar un marcador polimórfico dentro de dicho locus que está vinculado a una eficacia de transpiración mejorada; y

(f) seleccionar una planta que comprende o expresa el marcador.

Se utilizan medios convencionales conocidos para el experto en la materia para identificar un marcador dentro del locus que está vinculado a una eficacia de transpiración mejorada. Generalmente se utiliza una población de plantas que se segrega para el marcador polimórfico, en la que el fenotipo de eficacia de transpiración de plantas se correlaciona o asocia entonces con la presencia de una forma alélica particular del marcador. Tal como se muestra a título de ejemplo en la presente memoria, se examinan líneas endogámicas recombinantes o casi isogénicas de plantas para segregar alelos en el locus *ERECTA* y para correlacionar una eficacia de transpiración mejorada con la presencia del alelo *ERECTA* en contraposición a un alelo *erecta*. Alternativamente, se introducen mutaciones en un alelo *ERECTA* tal como, por ejemplo, mediante mutagénesis por transposones, mutagénesis química o irradiación de material vegetal, y se establecen y examinan líneas mutantes de plantas para segregar alelos en el locus *ERECTA* que se correlacionan con la variación genética en la eficacia de transpiración.

Los marcadores adecuados incluyen uno cualquiera o más de los marcadores descritos en la presente memoria como que son adecuados para MAS.

Preferentemente, la selección de plantas según estas formas de realización incluye la etapa adicional de introducir el locus o marcador polimórfico en una planta, tal como, por ejemplo, mediante enfoques de cultivo convencional o mediante medios recombinantes. Esto puede llevarse a cabo al mismo tiempo, o antes, seleccionando el locus o marcador polimórfico.

Los medios recombinantes incluyen generalmente introducir un constructo génico que comprende el locus o marcador en una célula vegetal, seleccionar tejido transformado y regenerar una planta completa a partir del explante de tejido transformado. Los medios para introducir ADN recombinante en tejido o células vegetales incluyen, pero no se limitan a, la transformación utilizando  $\text{CaCl}_2$  y variaciones de la misma, en particular el procedimiento descrito por Hanahan (1983), captación directa de ADN en protoplastos (Krens *et al.*, Nature 296, 72-74, 1982; Paszkowski *et al.*, EMBO J. 3, 2717-2722, 1984), captación mediada por PEG respecto a protoplastos (Armstrong *et al.*, Plant Cell Rep. 9, 335-339, 1990), bombardeo con micropartículas, electroporación (Fromm *et al.*, Proc. Natl. Acad. Sci. (EE.UU.), 82, 5824-5828, 1985), microinyección de ADN (Crossway *et al.*, Mol. Gen. Genet. 202, 179-185, 1986), bombardeo con micropartículas de células o explantes tisulares (Christou *et al.*, Plant Physiol. 87, 671-674, 1988; Sanford, Part. Sci. Technol. 5, 27-37, 1988), infiltración a vacío de tejido con ácido nucleico, o en el caso de plantas, la transferencia mediada por ADN-T desde *Agrobacterium* al tejido vegetal tal como se describe esencialmente por An *et al.*, EMBO J. 4, 277-284, 1985; Herrera-Estrella *et al.*, Herrera-Estrella *et al.*, Nature 303, 209-213, 1983; Herrera-Estrella *et al.*, EMBO J. 2, 987-995, 1983; o Herrera-Estrella *et al.*, En: Plant Genetic Engineering, Cambridge University Press, N.Y., págs. 63-93, 1985.

Para el bombardeo con micropartículas de células, una micropartícula se propulsa al interior de una célula para producir una célula transformada. Puede utilizarse cualquier aparato y metodología de transformación celular balística adecuada en la realización de la presente invención. Procedimientos y aparatos a título de ejemplo se dan a conocer por Stomp *et al.* (patente U.S. nº 5.122.466) y Sanford y Wolf (patente U.S. nº 4,945,050). Cuando se utilizan procedimientos de transformación balística, el constructo génico puede incorporar un plásmido que puede replicarse en la célula que va a transformarse.

Los ejemplos de micropartículas adecuadas para su utilización en tales sistemas incluyen esferas de oro de 1 a 5 micras. El constructo de ADN puede depositarse sobre la micropartícula mediante cualquier técnica adecuada, tal como mediante precipitación.

Una planta completa puede regenerarse a partir de la célula transformada o transfectada, según procedimientos muy conocidos en la materia. El tejido vegetal que puede experimentar propagación clónica posterior, ya sea mediante

organogénesis o embriogénesis, puede transformarse con un constructo génico de la presente invención y una planta completa regenerarse a partir del mismo. El tejido particular seleccionado variará dependiendo de los sistemas de propagación clónica disponibles para, y más adecuados para, la especie particular que esté transformándose. Las dianas de tejido a título de ejemplo incluyen discos foliares, polen, embriones, cotiledones, hipocótilos, megagametofitos, tejido de callo, tejido meristemático existente (por ejemplo, meristemo apical, yemas axilares y meristemas de raíz), y tejido de meristemo inducido (por ejemplo, meristemo de cotiledones y meristemo de hipocótilos).

El término "organogénesis", tal como se utiliza en la presente memoria, hace referencia a un proceso mediante el que los brotes y las raíces se desarrollan secuencialmente a partir de centros meristemáticos.

El término "embriogénesis", tal como se utiliza en la presente memoria, hace referencia a un proceso mediante el que los brotes y las raíces se desarrollan conjuntamente de un modo coordinado (no secuencialmente), ya sea a partir de células somáticas o gametos.

Las plantas transformadas generadas pueden propagarse mediante una variedad de medios, tal como mediante propagación clónica o técnicas de cultivo clásicas. Por ejemplo, una planta transformada de primera generación (o T1) puede autofecundarse para proporcionar un transformante de segunda generación (o T2) homocigótico, y las plantas T2 propagarse adicionalmente a través de técnicas de cultivo clásicas.

Los organismos transformados generados contemplados en la presente memoria pueden adoptar una variedad de formas. Por ejemplo, pueden ser quimeras de células transformadas y células no transformadas; transformantes clonales (por ejemplo, todas las células transformadas para contener el casete de expresión); injertos de tejidos transformados y no transformados (por ejemplo, en plantas, un patrón transformado injertado en una púa no transformada).

Alternativamente, las plantas transformadas se producen mediante un procedimiento de transformación *in planta* utilizando *Agrobacterium tumefaciens*, tal como, por ejemplo, el procedimiento descrito por Bechtold *et al.*, CR Acad. Sci. (Paris, Sciences de la vie/Life Sciences) 316, 1194-1199, 1993 o Clough *et al.*, Plant J 16: 735-74, 1998, en el que se aplica *A. tumefaciens* al exterior de la yema floral en desarrollo y el ADN de vector binario se introduce entonces en la macrospora y/o microspora en desarrollo y/o la semilla en desarrollo, para producir una semilla transformada. Los expertos en la materia apreciarán que la selección de tejido para su utilización en un procedimiento de este tipo puede variar, sin embargo es preferible utilizar generalmente material vegetal en la fase de formación de cigotos para procedimientos de transformación *in planta*.

#### *Identificación de genes para determinar la eficacia de transpiración de una planta*

Tal como se muestra a título de ejemplo en la presente memoria, se identificaron asimismo genes o alelos específicos que están ligados al locus *ERECTA* de *A. thaliana*, y arroz y determinan las eficacias de transpiración de esas plantas. Más particularmente, las eficacias de transpiración de líneas casi isogénicas, que portan, cada una, una mutación dentro de un locus *ERECTA*, y se determinan una correlación entre el fenotipo de eficacia de transpiración y la expresión de *ERECTA* o el número de copias génicas, proporcionando de ese modo la contribución genética de genes o alelos en el locus *ERECTA* a la eficacia de transpiración. Este análisis permite una evaluación de la contribución genética de alelos particulares a la eficacia de transpiración, determinando de ese modo variantes alélicas que están vinculadas a una eficacia de transpiración particular. Por tanto, el esclarecimiento del locus *ERECTA* para la eficacia de transpiración en plantas facilita el mapeo fino y la determinación de las variantes alélicas que modulan la eficacia de transpiración. Los procedimientos descritos en la presente memoria pueden aplicarse a una evaluación de la contribución de alelos específicos al fenotipo de eficacia de transpiración para cualquier especie vegetal que sea susceptible de mutagénesis tal como, por ejemplo, mediante mutagénesis por transposones, irradiación o medios químicos conocidos por los expertos en la materia para plantas en mutación.

Por lo tanto, en la presente memoria se describe un procedimiento para identificar un gen que determine la eficacia de transpiración de una planta que comprende:

- (a) identificar un locus asociado con una variación genética en la eficacia de transpiración en una planta;
- (b) identificar un gen o alelo que está ligado a dicho locus, en el que dicho gen o alelo es un gen o alelo candidato para determinar la eficacia de transpiración de una planta; y
- (c) determinar las eficacias de transpiración de un grupo de plantas, en el que no todos los elementos de dicho grupo comprenden o expresan dicho gen o alelo, y en el que la variación en la eficacia de transpiración entre los elementos de dicho grupo indica que dicho gen está implicado en la determinación de la eficacia de transpiración.

El procedimiento puede comprender:



- (a) identificar un locus asociado con una variación genética en la eficacia de transpiración en una planta;
- 5 (b) identificar múltiples alelos de un gen que está ligado a dicho locus, en el que dicho gen es un gen candidato implicado para determinar la eficacia de transpiración de una planta; y
- 10 (c) determinar las eficacias de transpiración de un grupo de plantas, en el que cada elemento de dicho grupo comprende, y preferentemente expresa, por lo menos uno de dichos múltiples alelos, la variación en la eficacia de transpiración entre los elementos de dicho grupo indicando que dicho gen está implicado en la determinación de la eficacia de transpiración.

En el presente contexto, se considerará que la expresión "plantas casi isogénicas" significa una población de plantas que presentan identidad con respecto a una proporción sustancial de sus genomas, a pesar de la presencia de diferencias suficientemente escasas para permitir la contribución de un alelo o gen distinto a la eficacia de transpiración de una planta que va a determinarse mediante una comparación de los fenotipos de eficacia de transpiración de la población. Tal como conocerá el experto en la materia, las líneas endogámicas recombinantes, las líneas producidas mediante introgresión de un gen o transposón seguido de varias generaciones de retrocruzamiento, o hermanas, son líneas casi isogénicas adecuadas para el presente fin.

20 Preferentemente, el gen identificado o alelo identificado mediante el procedimiento descrito en el párrafo anterior se selecciona de entre el grupo constituido por alelo *ERECTA*, alelo *erecta* y homólogos de *ERECTA*, en el que dichos homólogos son de especies vegetales diferentes de *A. thaliana*.

25 El gen o alelo identificado puede comprender una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

30 (d) una secuencia que presenta una identidad de por lo menos aproximadamente 55% con una secuencia seleccionada de entre el grupo constituido por SEC ID nº: 1, SEC ID nº: 3, SEC ID nº: 5, SEC ID nº: 7, SEC ID nº: 9, SEC ID nº: 11, SEC ID nº: 12, SEC ID nº: 13, SEC ID nº: 14, SEC ID nº: 15, SEC ID nº: 16, SEC ID nº: 17, SEC ID nº: 18, SEC ID nº: 19 SEC ID nº: 21, SEC ID nº: 22, SEC ID nº: 23, SEC ID nº: 24, SEC ID nº: 25, SEC ID nº: 26, SEC ID nº: 27, SEC ID nº: 28, SEC ID nº: 29, SEC ID nº: 30, SEC ID nº: 31, SEC ID nº: 32 SEC ID nº: 33, SEC ID nº: 34, SEC ID nº: 35, SEC ID nº: 36, SEC ID nº: 37, SEC ID nº: 38; SEC ID nº: 39, SEC ID nº: 40, SEC ID nº: 41, SEC ID nº: 42, SEC ID nº: 43 y SEC ID nº: 44;

35 (e) una secuencia que codifica para una secuencia de aminoácidos que presenta una identidad de por lo menos aproximadamente 55% con una secuencia de aminoácidos seleccionada de entre el grupo constituido por SEC ID nº: 2, SEC ID nº: 4, SEC ID nº: 6, SEC ID nº: 8, SEC ID nº: 10, SEC ID nº: 12, SEC ID nº: 20 y SEC ID nº: 45; y

40 (f) una secuencia complementaria a (a) o (b).

45 Preferentemente, el porcentaje de identidad es de por lo menos aproximadamente el 59-61% o el 70% o el 80%, más preferentemente por lo menos aproximadamente 90%, e incluso más preferentemente por lo menos aproximadamente el 95% o el 99%. El gen o alelo identificado puede comprender una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

50 (a) una secuencia seleccionada de entre el grupo constituido por SEC ID nº: 1, SEC ID nº: 3, SEC ID nº: 5, SEC ID nº: 7, SEC ID nº: 9, SEC ID nº: 11, SEC ID nº: 12, SEC ID nº: 13, SEC ID nº: 14, SEC ID nº: 15, SEC ID nº: 16, SEC ID nº: 17, SEC ID nº: 18, SEC ID nº: 19 SEC ID nº: 21, SEC ID nº: 22, SEC ID nº: 23, SEC ID nº: 24, SEC ID nº: 25, SEC ID nº: 26, SEC ID nº: 27, SEC ID nº: 28, SEC ID nº: 29, SEC ID nº: 30, SEC ID nº: 31, SEC ID nº: 32 SEC ID nº: 33, SEC ID nº: 34, SEC ID nº: 35, SEC ID nº: 36, SEC ID nº: 37, SEC ID nº: 38; SEC ID nº: 39, SEC ID nº: 40, SEC ID nº: 41, SEC ID nº: 42, SEC ID nº: 43 y SEC ID nº: 44;

55 (b) una secuencia que codifica para una secuencia de aminoácidos seleccionada de entre el grupo constituido por SEC ID nº: 2, SEC ID nº: 4, SEC ID nº: 6, SEC ID nº: 8, SEC ID nº: 10, SEC ID nº: 12, SEC ID nº: 20 y SEC ID nº: 45; y

(c) una secuencia complementaria a (a) o (b).

60 *Mejora de la eficacia de transpiración utilizando genes aislados*

El gen o alelos identificados, incluyendo cualquier homólogo de una planta distinta de *A. thaliana*, tal como, por ejemplo, el alelo *ERECTA* de tipo natural, o un homólogo de los mismos, son útiles para la producción de nuevas plantas. Tales plantas se producen, por ejemplo, utilizando técnicas recombinantes.

65 Por lo tanto, se describe en la presente memoria un procedimiento para mejorar la eficacia de transpiración de una

planta que comprende expresar ectópicamente en una planta un gen *ERECTA* aislado o una variante alélica del mismo o la región que codifica para proteínas del mismo.

5 Preferentemente, el gen *ERECTA* o variante alélica comprende una secuencia de nucleótidos que es homóloga a una región que codifica para proteínas de un gen que está ligado al locus *ERECTA* de *A. thaliana* en el cromosoma 2.

El gen aislado puede comprender una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

10 (a) una secuencia que presenta una identidad de por lo menos aproximadamente 55% con una secuencia seleccionada de entre el grupo constituido por SEC ID nº: 1, SEC ID nº: 3, SEC ID nº: 5, SEC ID nº: 7, SEC ID nº: 9, SEC ID nº: 11, SEC ID nº: 12, SEC ID nº: 13, SEC ID nº: 14, SEC ID nº: 15, SEC ID nº: 16, SEC ID nº: 17, SEC ID nº: 18, SEC ID nº: 19 SEC ID nº: 21, SEC ID nº: 22, SEC ID nº: 23, SEC ID nº: 24, SEC ID nº: 25, SEC ID nº: 26, SEC ID nº: 27, SEC ID nº: 28, SEC ID nº: 29, SEC ID nº: 30, SEC ID nº: 31, SEC ID nº: 32 SEC ID nº: 33,  
15 SEC ID nº: 34, SEC ID nº: 35, SEC ID nº: 36, SEC ID nº: 37, SEC ID nº: 38; SEC ID nº: 39, SEC ID nº: 40, SEC ID nº: 41, SEC ID nº: 42, SEC ID nº: 43 y SEC ID nº: 44;

(b) una secuencia que codifica para una secuencia de aminoácidos que presenta una identidad de por lo menos aproximadamente 55% con una secuencia de aminoácidos seleccionada de entre el grupo constituido por SEC ID nº: 2, SEC ID nº: 4, SEC ID nº: 6, SEC ID nº: 8, SEC ID nº: 10, SEC ID nº: 12, SEC ID nº: 20 y SEC ID nº: 45;  
20 y

(c) una secuencia complementaria a (a) o (b).

25 Preferentemente, el porcentaje de identidad es de por lo menos aproximadamente el 59-61% o el 70% o el 80%, más preferentemente por lo menos aproximadamente el 90%, e incluso más preferentemente por lo menos aproximadamente el 95% o el 99%.

30 El gen o alelo aislado puede comprender una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

(a) una secuencia seleccionada de entre el grupo constituido por SEC ID nº: 1, SEC ID nº: 3, SEC ID nº: 5, SEC ID nº: 7, SEC ID nº: 9, SEC ID nº: 11, SEC ID nº: 12, SEC ID nº: 13, SEC ID nº: 14, SEC ID nº: 15, SEC ID nº: 16, SEC ID nº: 17, SEC ID nº: 18, SEC ID nº: 19 SEC ID nº: 21, SEC ID nº: 22, SEC ID nº: 23, SEC ID nº: 24, SEC ID nº: 25, SEC ID nº: 26, SEC ID nº: 27, SEC ID nº: 28, SEC ID nº: 29, SEC ID nº: 30, SEC ID nº: 31, SEC ID nº: 32 SEC ID nº: 33, SEC ID nº: 34, SEC ID nº: 35, SEC ID nº: 36, SEC ID nº: 37, SEC ID nº: 38; SEC ID nº: 39, SEC ID nº: 40, SEC ID nº: 41, SEC ID nº: 42, SEC ID nº: 43 y SEC ID nº: 44;

(b) una secuencia que codifica para una secuencia de aminoácidos seleccionada de entre el grupo constituido por SEC ID nº: 2, SEC ID nº: 4, SEC ID nº: 6, SEC ID nº: 8, SEC ID nº: 10, SEC ID nº: 12, SEC ID nº: 20 y SEC ID nº: 45; y

(c) una secuencia complementaria a (a) o (b).

45 Para expresar ectópicamente el gen aislado en una planta, la parte que codifica para proteínas del gen se sitúa generalmente en conexión operativa con una secuencia promotora que puede operar en la planta, que puede ser el promotor endógeno o alternativamente, un promotor heterólogo, y una secuencia de terminación de la transcripción, que también puede ser la secuencia endógena o una heteróloga en relación con el gen de interés. El promotor y la parte que codifica para proteínas y la secuencia de terminación de la transcripción se proporcionan generalmente en forma de un constructo génico, para facilitar la introducción y el mantenimiento del gen en una planta en la que debe expresarse ectópicamente. Se han descrito numerosos vectores adecuados para introducir genes en plantas y están fácilmente disponibles. Éstos pueden adaptarse para expresar un gen aislado en una planta para mejorar la eficacia de transpiración en la misma.

55 La referencia en la presente memoria a un "promotor" debe considerarse en su contexto más amplio e incluye las secuencias reguladoras transcripcionales de un gen genómico eucariota clásico, incluyendo la caja TATA que se requiere para una iniciación de la transcripción precisa, con o sin una secuencia de caja CCAAT y elementos reguladores adicionales (es decir secuencias de activación en el sentido de 5', potenciadores y silenciadores) que alteran la expresión génica en respuesta a estímulos de desarrollo y/o externos, o de una manera específica del tejido. En el presente contexto, el término "promotor" se utiliza también para describir una molécula sintética o de fusión, o derivado que confiere, activa o potencia la expresión de dicha molécula sentido en una célula. Los promotores preferidos pueden contener copias adicionales de uno o más elementos reguladores específicos, para potenciar adicionalmente la expresión y/o para alterar la expresión espacial y/o la expresión temporal de una molécula de ácido nucleico a la que está conectado de manera operativa. Por ejemplo, los elementos reguladores sensibles al cobre pueden situarse adyacentes a una secuencia promotora heteróloga que desencadena la expresión de una molécula de ácido nucleico para conferir la expresión inducible de cobre en la misma.

Situar una molécula de ácido nucleico bajo el control regulador de una secuencia promotora significar colocar dicha molécula de modo que la expresión esté controlada por la secuencia promotora. Un promotor se coloca habitualmente, pero no necesariamente, en el sentido de 5' o en 5' de la parte que codifica para proteínas del gen que regula. Además, los elementos reguladores que comprenden un promotor se colocan habitualmente a como máximo 2 kb del sitio de inicio de la transcripción de las secuencias de nucleótidos que codifican para proteínas estructurales, o un gen quimérico que comprende las mismas. En la construcción de combinaciones de promotor heterólogo / gen estructural se prefiere generalmente colocar el promotor a una distancia del sitio de inicio de la transcripción génica que sea aproximadamente igual que la distancia entre ese promotor y el gen que controla en su entorno natural, es decir, el gen del que se deriva el promotor. Tal como se conoce en la técnica, puede darse cabida a cierta variación en esta distancia sin pérdida de la función promotora. De manera similar, la colocación preferida de un elemento de secuencia reguladora con respecto a un gen heterólogo que va a situarse bajo su control se define mediante la colocación del elemento en su entorno natural, es decir, los genes de los que se deriva. De nuevo, tal como se conoce en la técnica, también puede producirse cierta variación en esta distancia.

Los promotores adecuados para su utilización en los constructos génicos de la presente invención incluyen aquellos promotores derivados de los genes de virus, levaduras, mohos, bacterias, insectos, aves, mamíferos y plantas que pueden funcionar en células vegetales, incluyendo plantas monocotiledóneas o dicotiledóneas, o tejidos u órganos derivados de tales células. El promotor puede regular la expresión génica de manera constitutiva, o de manera diferencial con respecto al tejido en el que se produce la expresión o, con respecto a la fase de desarrollo en la que se produce la expresión, o en response a estímulos externos tales como estreses fisiológicos, patógenos o iones metálicos, entre otros.

Los ejemplos de promotores útiles al realizar esta realización incluyen el promotor CaMV 35S, el promotor de actina de arroz, el promotor de actina de arroz ligado al intrón de actina de arroz (PAR-IAR) (McElroy *et al*, Mol and Gen Genetics, 231(1), 150-160, 1991), el promotor NOS, el promotor de octopina sintasa (OCS), el promotor del gen de la SSU de *Arabidopsis thaliana*, el promotor específico de semilla napin, PcSVMV, promotores que pueden inducir la expresión bajo estrés hídrico, tal como se describe por, por ejemplo, Kasuga *et al*, Nature Biotechnology, 17, 287-291, 1999), el promotor SCSV, el promotor SCBV, el promotor 35s (Kay *et al*, Science 236, 4805, 1987) y similares. Además de los promotores específicos identificados en la presente memoria, son útiles los promotores celulares para los denominados genes de mantenimiento, incluyendo los promotores de actina o los promotores de genes que codifican para histonas.

El término "terminador" se refiere a una secuencia de ADN en el extremo de una unidad transcripcional que señala la terminación de la transcripción. Los terminadores son secuencias de ADN no traducidas en 3' que contienen una señal de poliadenilación, que facilitan la adición de una secuencia de poliadenilación al extremo 3' de un transcrito primario. Terminadores activos en células derivadas de virus, levaduras, mohos, bacterias, insectos, aves, mamíferos y plantas se conocen y se describen en la bibliografía. Pueden aislarse a partir de bacterias, hongos, virus, animales y/o plantas.

Los ejemplos de terminadores particularmente adecuados para su utilización en los constructos génicos de la presente invención incluyen el terminador del gen de la nopalina sintasa (NOS) de *Agrobacterium tumefaciens*, el terminador del gen 35S del virus del mosaico de la coliflor (CaMV), el terminador del gen de la zeína de *Zea mays*, las secuencias de terminador del gen de la subunidad pequeña (SSU) de Rubisco y los terminadores de secuencia del gen del virus del enanismo del trébol subterráneo (SCSV), entre otros.

Los expertos en la materia apreciarán que las secuencias terminadoras y secuencias promotoras adicionales que pueden ser adecuadas para su utilización en realizar la invención. Tales secuencias pueden utilizarse fácilmente sin excesiva experimentación.

Preferentemente, el constructo génico comprende además un origen de la secuencia de replicación para su replicación en un tipo de célula específico, por ejemplo una célula bacteriana, cuando se requiere mantener dicho constructo génico como un elemento genético episomal (por ejemplo molécula de cósmido o plásmido) en dicha célula. Orígenes de replicación preferidos incluyen, pero no se limitan a, los orígenes de replicación *f1-ori* y *colE1*.

Preferentemente, el constructo génico comprende además un gen o genes marcadores seleccionables que son funcionales en una célula en la que se introduce dicho constructo génico.

Tal como se utiliza en la presente memoria, el término "gen marcador seleccionable" incluye cualquier gen que confiera un fenotipo a una célula en la que se expresa para facilitar la identificación y/o selección de células que se transfectan o transforman con un constructo génico descrito en la presente memoria o un derivado del mismo.

Los genes marcadores seleccionables adecuados contemplados en la presente memoria incluyen el de resistencia a la ampicilina (Amp<sup>r</sup>), gen de resistencia a la tetraciclina (Tc<sup>r</sup>), gen de resistencia a la kanamicina bacteriana (Kan<sup>r</sup>), gen de resistencia a la fosfotricina, gen de la neomicina fosfotransferasa (*npII*), gen de resistencia a la higromicina, gen de resistencia a la gentamicina (*gent*), gen de la  $\beta$ -glucuronidasa (GUS), gen de la cloranfenicol

acetiltransferasa (CAT), y gen de la luciferasa, gen de la proteína fluorescente verde (EGFP y variantes), entre otros.

También se describe en la presente memoria la utilización de un gen aislado que comprende una secuencia de nucleótidos que es homóloga a una región que codifica para proteínas de un gen de *A. thaliana* que se coloca entre aproximadamente 46 cM y aproximadamente 50,74 cM en el cromosoma 2 en la preparación de un constructo génico para mejorar la eficacia de transpiración de una planta.

También se describe en la presente memoria una planta que presenta una eficacia de transpiración mejorada, produciéndose dicha planta mediante un procedimiento descrito en la presente memoria.

Claramente, los genes *ERECTA*, variantes alélicas y regiones que codifican para proteínas descritos en la presente memoria son útiles en la determinación de otras proteínas que están implicadas en el proceso de transpiración en plantas. Por ejemplo, un gen *ERECTA*, variante alélica del mismo o región que codifica para proteínas del mismo puede utilizarse en un ensayo de "n" híbridos progresivo para determinar si dicho péptido puede unirse a una proteína o péptido de interés. Los procedimientos de "n" híbridos progresivos son muy conocidos en la técnica, y se describen por ejemplo, por Vidal y Legrain Nucl. Acid Res 27(4), 919-929 (1999) y las referencias en el mismo, e incluyen el de dos híbridos de levadura, dos híbridos bacterianos, dos híbridos de mamífero, (dos) híbridos de PolIII, el sistema Tribid, el sistema de sensor de proteínas discontinuas a base de ubiquitina y el sistema de reclutamiento SOS.

Al adaptar un ensayo de dos híbridos progresivo convencional a la presente invención, una proteína *ERECTA* se expresa como proteína de fusión con un dominio de unión a ADN de, por ejemplo, la proteína GAL4 de levadura. Los procedimientos de construcción de constructos de expresión para la expresión de tales proteínas de fusión son muy conocidos en la técnica, y se describen, por ejemplo, en Sambrook *et al* (En: Molecular Cloning: A laboratory Manual, Cold Spring Harbour, Nueva York, segunda edición, 1989). Una segunda proteína de fusión también se expresa en toda la levadura, comprendiendo dicha proteína de fusión, por ejemplo, una proteína que se cree que interacciona con una proteína *ERECTA*, por ejemplo el dominio de activación GAL4. Estos dos constructos se expresan entonces en una célula de levadura en la que, una molécula indicadora (por ejemplo, *tef*, *Amp<sup>r</sup>*, *Rif*, *bsdf*, *zeof*, *Kan<sup>r</sup>*, *gfp*, *cobA*, *LacZ*, *TRP1*, *LYS2*, *HIS3*, *HIS5*, *LEU2*, *URA3*, *ADE2*, *MET13*, *MET15*) bajo el control de un promotor mínimo situado en conexión operativa con un sitio de unión a GAL 4. Si las proteínas no interaccionan, no se expresa una molécula indicadora. Sin embargo, si dichas proteínas si interaccionan, se expresa dicha molécula indicadora. Por lo tanto se identifica una proteína, polipéptido, péptido que puede unirse específicamente a una proteína diana.

Un procedimiento de "n" híbridos progresivo puede modificarse para facilitar el examen de alto rendimiento de una biblioteca de péptido, polipéptidos y/o proteínas con el fin de determinar aquellos que interaccionan con una proteína *ERECTA*. Los procedimientos para examinar bibliotecas de proteínas son muy conocidos en la técnica y se describen, por ejemplo, en Scopes (En: Protein Purification: Principles and Practice, tercera edición, Springer Verlag, 1994). Las proteínas identificadas mediante este procedimiento están implicadas potencialmente en el proceso de transpiración en plantas.

La presente invención se describe adicionalmente haciendo referencia a los ejemplos no limitativos siguientes.

### Ejemplo 1

Discriminación de  $^{12}\text{C}/^{13}\text{C}$  como marcador para examinar una variación genética en la eficacia de transpiración.

Se establecieron las condiciones experimentales y los procedimientos de toma de muestras para permitir el control de muchos factores, distintos del genético, que influyen en la eficacia de transpiración al nivel de plantas y hojas individuales. Estos factores se encuentran dentro de varias categorías: (a) características del microambiente de la plántula: temperatura, luz, humedad, capa límite alrededor de las hojas, condiciones de crecimiento de la raíz; (b) efectos morfológicos y de desarrollo que modifican el intercambio de gases y metabolismo de C y por tanto la firma isotópica de carbono (por ejemplo edad, fase, postura); y (c) efectos en la semilla.

Se desarrollaron técnicas de espectrómetro de masas de alta resolución para medir las razones de isótopos de C en tejidos completos o compuestos de carbono tales como azúcares solubles, es decir una medida de la eficacia de transpiración integrada a lo largo de la vida de la planta o a lo largo de un día, respectivamente, y también para medir la eficacia de transpiración instantánea durante el intercambio de gases.

Esto significa:

- Precisión analítica de 0,1 por mil en la medición de la composición isotópica de carbono de hoja. La discriminación, ( $\Delta$ ), es aproximadamente la razón de isótopos de carbono en el  $\text{CO}_2$  fuente menos la del carbono orgánico de la planta. En un experimento particular, el  $\text{CO}_2$  fuente es común para todos los genotipos.
- Precisión biológica de 0,1 por mil, que es la variación entre las plántulas replicadas, cultivadas en suelo, o bien

en cámaras de crecimiento o bien en invernaderos con control de temperatura, humedad y CO<sub>2</sub> (que corresponde a aproximadamente el 1,5% de variación en la eficacia de transpiración).

- La capacidad de hacer crecer y examinar grandes lotes de plántulas en invernaderos o cámaras de crecimiento (hasta 1500), en condiciones de crecimiento de la raíz y la hoja estandarizadas, hasta un tamaño de roseta de varios cm en el plazo de 2-3 semanas que permite mediciones individuales, en la misma planta, de razones de isótopos y también de las propiedades subyacentes (por ejemplo medición de temperatura de la hoja *in situ* mediante termometría infrarroja como una medida de la conductancia estomática; fluorescencia de clorofila; expansión de las hojas).

## Ejemplo 2

Variación genética natural en la eficacia de transpiración en *Arabidopsis thaliana*

Se examinaron ecotipos de *A. thaliana* para determinar la  $\Delta$  de las hojas en condiciones de invernadero. Hubo una gran dispersión de valores (que corresponden a aproximadamente el 30% de la variación genética en la eficacia de transpiración). Sin embargo, se observaron grandes efectos medioambientales. Se seleccionaron algunos ecotipos contrastados en los dos extremos del intervalo de valores de  $\Delta$  y se comparan en diversas condiciones de irradiancia (150 a 500  $\mu\text{E m}^{-2}\text{s}^{-1}$ ), espectro de luz (razones de rojo/rojo lejano) y humedad del aire (del 60 al 90%) mientras que las raíces se regaban siempre de manera abundante. La magnitud de las diferencias genéticas en la eficacia de transpiración se veía muy influida por las condiciones ambientales. Esto se debía en parte a variaciones entre ecotipos en la dependencia de fotosíntesis en caso de déficit de presión de vapor y luz. Las diferencias genéticas fueron máximas en una combinación de humedad baja y luz alta, en cámaras de crecimiento.

Los ecotipos *Columbia (Col)* y *Landsberg erecta (Ld-er)* presentan valores de discriminación de isótopos de carbono extremos, presentando *Col* siempre valores de  $\Delta$  más pequeños que *Ld-er* es decir composiciones isotópicas  $\delta^{13}\text{C}$  menos negativas, y por tanto una mayor eficacia de transpiración.

## Ejemplo 3

Identificación de un locus asociado con la eficacia de transpiración en *A. thaliana*

Se realizó un análisis de carácter cuantitativo (QTL) de las líneas endogámicas recombinantes de Lister y Dean (1993) (a las que se hace referencia posteriormente como las RIL) para identificar y mapear un locus asociado con la discriminación de isótopos de carbono ( $\Delta$ ). Las RIL eran de un cruzamiento entre Col-4 y Ler-0. El análisis mostró la importancia de los genes alrededor del locus *ER* en chr2 y un papel para los genes distintos de *ERECTA* en conferir la eficacia de transpiración en *A. thaliana*.

Más particularmente, se generaron 300 líneas de mapeo RI entre ecotipos Col y Ler, disponibles en el Arabidopsis Stock Centre, a partir de un cruzamiento entre los ecotipos de *Arabidopsis Columbia (Col4)* y *Landsberg erecta (Ler-0)* que porta *er1* (Lister y Dean, 1993), utilizando Columbia como el progenitor macho. Se utilizó un subconjunto de 100 de estas líneas, que se escogieron como las mapeadas de la manera más densa y fiable en el presente análisis.

Se multiplicaron las semillas en un invernadero en un intento por minimizar los efectos en la semilla de confusión en nuestras comparaciones. Se obtuvieron grandes números de semillas para la mayoría de las líneas excepto para algunas, que incluyen la parental Col4, que tuvo que volver a pedirse debido a la baja viabilidad de las semillas de la muestra original enviada por el Stock Centre. Se utilizaron las semillas recogidas en estas series de propagación en todos los experimentos hasta la fecha.

Se analizaron los loci utilizando dos programas, QTL cartographer y MQTL. Estos programas calculan las estadísticas de un rasgo en cada posición del marcador, utilizando una gama de procedimientos [regresión lineal (LR), regresión escalonada (SR) y enfoques de probabilidad (mapeo por intervalos simples (SIM) que trata valores en marcadores individuales como valores independientes, y mapeo por intervalos compuestos (CIM) que permite las interacciones entre marcadores y el locus asociado)]. Por naturaleza cada uno de estos procedimientos presenta algunos sesgos y suposiciones arraigadas, por esto la importancia de analizar los datos con más de un programa. Sólo los resultados que fueron constantes entre los dos programas, y robustos con respecto a las adiciones o deleciones al conjunto de marcadores de contexto que se utilizaron para el mapeo por intervalos compuestos se indican a continuación.

Se realizó un análisis QTL inicial en paralelo a la multiplicación de semillas en un subconjunto de 40 líneas para las que se enviaron suficientes semillas. Una vez que todas las semillas se multiplicaron esto se repitió en el conjunto completo de 100 líneas. Estos dos análisis indicaron la existencia de un locus para discriminación de isótopos de carbono ( $\Delta$ ), que se correlaciona con la región que incluye el locus *ERECTA* en el cromosoma 2, en aproximadamente 46-51 cM (tabla 1, serie 1 y 2).

Dada la complejidad y naturaleza integrativa de  $\Delta$  como rasgo fisiológico, no se esperaba un número tan pequeño de loci asociados con el rasgo. Por tanto se diseñaron experimentos posteriores para someter a prueba estos resultados y evaluar su estabilidad por toda la gama de condiciones ambientales conocidas por sus efectos sobre la expresión genética relacionada con  $\Delta$  (véase anteriormente). Se repitió el análisis QTL en varios conjuntos de datos completamente independientes obtenidos en condiciones sumamente controladas en invernaderos o cámaras de crecimiento, en los que se varió o bien la humedad del aire, el fotoperiodo o bien la irradiancia (cantidad, patrón diurno, variación de día a día). Dependiendo del experimento, se incluyeron las 100 líneas endogámicas recombinantes o sólo el subconjunto de líneas con entrecruzamientos en el cromosoma 2. Estos experimentos confirmaron que una variación genética en  $\Delta$  podría atribuirse mayoritariamente a una parte del cromosoma 2 (tabla 1) entre aproximadamente 46-50,7 cM.

Cuando las RIL se clasificaron gráficamente según la discriminación de isótopos de carbono y su genotipo en el marcador *ER* (50,64 cM) y sus proximidades (genotipo *Ld-er1* o genotipo *Col-ER*), las líneas que eran *Ld-er* en el marcador *ERECTA* se clasificaron mayoritariamente en el extremo superior de los valores de discriminación isotópica de carbono. Por el contrario, las líneas que presentan un genotipo de marcador *Col-ERECTA* se clasificaron mayoritariamente en el extremo inferior de los valores de discriminación de isótopos de carbono (datos disponibles a petición). En el medio del intervalo de valores de discriminación de isótopos de carbono, hubo algo de solapamiento entre los dos conjuntos de líneas. Algunas líneas estuvieron siempre en un extremo (en los 18 experimentos realizados), mientras que la clasificación de otras líneas fue más inestable. Estos datos indican un locus para la eficacia de transpiración, tal como se determina mediante el valor de discriminación de isótopos de carbono, en las proximidades del locus *ERECTA* en el cromosoma 2 (tabla 1). Este locus implica de la manera más probable el gen *ER*. Dependiendo de las posiciones de los entrecruzamientos entre *Ld-er* y *Col*, la recombinación entre *ERECTA* y uno o más de los otros genes influye en el fenotipo de eficacia de transpiración de la progenie.

#### 25 Ejemplo 4

Protocolo de transformación para maíz

##### 30 Transformación biolística

Un procedimiento adecuado para la transformación de maíz se basa en la utilización de una pistola de partículas idéntica a la descrita por J. Finer (1992, Plant Cell Report, 11:323-328). Las células diana son células no diferenciadas que se dividen rápidamente que mantienen una capacidad para regenerarse en plantas completas. Este tipo de células componen el callo embriogénico (denominado tipo II) del maíz. Se obtienen estos callos a partir de embriones inmaduros del genotipo Hill según el procedimiento y en el medio descritos por Armstrong (Maize Handbook; 1994 M. Freeling, V. Walbot Eds; págs. 665-671).

Se disponen estos fragmentos de los callos que presentan una superficie de desde 10 hasta 20 mm<sup>2</sup>, 4 horas antes del bombardeo, colocando 16 fragmentos por placa en el centro de una placa Petri que contiene un medio de cultivo idéntico al medio de iniciación de los callos, complementado con 0,2 M de manitol + 0,2 M de sorbitol. Los plásmidos que contienen las secuencias *ERECTA* que van a introducirse, se purifican en una columna QiagenR siguiendo las instrucciones del fabricante.

Entonces, se precipitan en partículas de tungsteno (M10) siguiendo el protocolo descrito por Klein *et al*, Nature, 327, 70-73, (1987). Las partículas así recubiertas se envían hacia las células diana mediante la pistola según el protocolo descrito por Finer *et al*, Plant Cell Report, 11:323-328, 1992. Entonces se sellan las placas bombardeadas de callos mediante ScellofraisR, entonces se cultivan en la oscuridad a 27°C.

El primer trasplante tiene lugar 24 horas después, después cada dos semanas durante 3 meses en medio idéntico al medio de iniciación complementado con un agente selectivo. Tras 3 meses o a veces antes, pueden obtenerse callos cuyo crecimiento no se inhiba mediante el agente selectivo, que consisten habitual y principalmente en células resultan de la división de una célula que presenta integrado en su patrimonio genético una o varias copias del gen de selección. La frecuencia de obtención de tales callos es de aproximadamente 0,8 callos por placa bombardeada.

Se identifican, individualizan, amplifican estos callos, después se cultivan para regenerar plántulas, modificando el equilibrio osmótico y hormonal de las células según el procedimiento descrito por Vain y colaboradores (1989, Plant Cell tissue and organ Culture 18:143-151). Estas plantas se aclimatan entonces en invernadero, en el que pueden cruzarse para obtener híbridos o fertilizarse a sí mismas.

Preferentemente, puede utilizarse un protocolo similar, cuyo principio se describe en Methods of Molecular Biology: Plant gene transfer and expression protocols (1995, vol. 49, págs. 113-123), y en el que los embriones inmaduros de genotipo Hill se bombardean directamente con partículas de oro recubiertas con plásmidos *ERECTA* para su introducción, preparadas según el protocolo descrito por Barcelo y Lazzeri (1995, Methods of Molecular Biology, 49:113-123).

Las etapas de transformación, selección de los eventos, maduración y regeneración son similares a aquellas

descritas en el protocolo previo.

#### Transformación con *Agrobacterium*

5 Otra técnica de transformación útil dentro del marco de la invención utiliza *Agrobacterium tumefaciens*, según el protocolo descrito por Ishida y colaboradores (1996, Nature Biotechnology 14: 754-750), en particular partiendo de embriones inmaduros tomados 10 días después de la fertilización. Todos los medios utilizados se describen en la referencia citada. La transformación empieza con una fase de cocultivo en el que los embriones inmaduros de las plantas de maíz se ponen en contacto durante por lo menos 5 minutos con LBA 4404 de *Agrobacterium tumefaciens* que contiene los vectores superbinarios.

15 El plásmido superbinario es el resultado de una recombinación homóloga entre un vector intermediario que porta el ADN-T, y que contiene el gen de interés y/o el gen marcador de selección, y el vector pSB1 del Japan Tobacco (documento EP 672 752) que contiene: los genes *virB* y *virG* del plásmido pTiBo542 presente en la cepa supervirulenta A281 de *Agrobacterium tumefaciens* (ATCC 37349) y una región homóloga encontrada en el vector intermediario, permitiendo una recombinación homóloga.

20 Se colocan los embriones entonces en medio LSAs durante 3 días en la oscuridad y a 25°C. Se hace una primera selección en los callos transformados: se transfieren los callos embriogénicos a medio LSD5 que contiene fosfotricina (5 mg/l) y cefotaxima (250 mg/l) (eliminación o limitación de la contaminación mediante *Agrobacterium tumefaciens*).

25 Se realiza esta etapa durante 2 semanas en la oscuridad y a 25°C. Se realiza la segunda etapa de selección mediante la transferencia de los embriones que se desarrollaron en el medio LSD5, al medio LSD10 (fosfotricina, 10 mg/l) en presencia de cefotaxima, durante 3 semanas en las mismas condiciones que anteriormente. La tercera fase de selección consiste en cortar los callos de tipo I (fragmentos de desde 1 hasta 2 mm) y en transferirlos durante 3 semanas en la oscuridad y a 25°C a medio LSD 10 en presencia de cefotaxima. Se realiza la regeneración de plántulas cortando los callos de tipo I que proliferaron y transfiriéndolos al medio LSZ en presencia de fosfotricina (5 mg/l) y de cefotaxima durante 2 semanas a 22°C y con luz continua.

30 Las plántulas que se regeneraron se transfieren a medio RM + G2 que contiene Augmentin (100 mg/l) durante 2 semanas a 22°C y con iluminación continua durante la etapa de desarrollo. Las plantas obtenidas se transfieren entonces al fitotrón con el objetivo de aclimatarlas.

#### 35 **Ejemplo 5**

Detectar la expresión de la proteína *ERECTA*

##### Extracción de *ERECTA* de hojas y semillas de maíz.

40 Se recogen las hojas e inmediatamente se congelan en nitrógeno líquido. Se realiza la molienda en un mortero limpiado en etanol al 100% y enfriado en hielo. Se extrae un disco foliar de 18 mm de diámetro en 200 µl de tampón de extracción: Tris-HCl pH 8,0, glicerol al 20%, MgCl<sub>2</sub> 10 mM, EDTA 1 mM, PVP insoluble al 2% (p/v), arena de Fontainebleau e inhibidores de proteasa: leupeptina 2 mg/l, quimostatina 2 mg/l, PMSF 1 mM y E64 1 mg/l. Se centrifuga entonces el material molido a 4°C durante 15 minutos a 20000 g para eliminar fragmentos.

50 En primer lugar se reducen los granos a polvo en una trituradora de perlas (Retsch). Se extraen las proteínas suspendiendo 100 µl de polvo en 400 µl del tampón descrito anteriormente en hielo. Esta mezcla se agita por vórtex y se centrifuga a 4°C durante 15 minutos a 20000 g para eliminar fragmentos.

Se miden entonces los niveles de proteína *ERECTA* utilizando técnicas conocidas por los expertos en la materia, y descritas, por ejemplo, en Scopes (En: Protein purification: principles and practice, tercera edición, Springer Verlag, 1994).

#### 55 **Ejemplo 6**

Determinación de un rol para el gen *ERECTA* en la regulación de la eficacia de transpiración

60 Se compararon ecotipos Col y Ler con líneas mutantes casi isogénicas para el gen *erecta*, para examinar un posible papel del gen *ERECTA* en la determinación de la discriminación de isótopos de carbono (Δ).

65 Las plantas que expresan el gen *ERECTA* de tipo natural (SEC ID nº: 1), o un alelo mutante *erecta* en el contexto Columbia (por ejemplo Col-er1, Col-er2, Col-er101 a -er105; o Col-er106 a -er123) y en el contexto Landsberg (*Ld-er1*) se han descrito públicamente. Estaban disponibles tres de estos mutantes para su comparación con las líneas isogénica o casi isogénica (tabla 2).

Se incluyeron sistemáticamente *Col4* (ER) y *Ld-er1*, las líneas parentales para las RIL de Lister y Dean en la comparación. Cuando era posible, se incluyeron también otros “ecotipos” Col, (por ejemplo *Col10*, *Col1*, *Col3-7*), para evaluar su similitud con respecto a la discriminación de isótopos de carbono, especialmente en comparación con el ecotipo parental de RIL *Col4*.

Se describen los resultados de estas comparaciones en la tabla 3. Los datos indican las diferencias en los valores de discriminación de isótopos de carbono entre líneas *er* y *ER* para 15 series experimentales diferentes que corresponden al crecimiento con luz de baja a alta (100 a 800  $\mu\text{E m}^{-2} \text{s}^{-1}$ ), humedad de baja a alta (del 40 al 85%), días de cortos a largos (8, 10, 24 h), temperaturas de normales a altas (de 22/20°C a 28/20°C).

Como se esperaba, la dispersión de los valores de discriminación de isótopos de carbono entre líneas variaba con las condiciones ambientales. Las líneas que portan mutaciones *er* presentan un valor mayor de discriminación de isótopos de carbono global que aquéllas que presentan en el gen *ER* de tipo natural (véase la tabla 3, columna 1), indicativo de una eficacia de utilización de agua menor. Hay habitualmente poca diferencia en la discriminación isotópica de C entre las diversas líneas Col, (véanse los promedios similares obtenidos para las columnas 2, 3 y 4 en la tabla 3, en la que *er105* se compara con 3 ecotipos Col diferentes *Col10*, *Col4* y 3176 o *Col1*). Cuando estuvo presente, el mutante *er105* siempre presenta el mayor valor de discriminación de isótopos de carbono de todas las líneas, incluyendo *er1* y *er2* (columnas 2-4 en comparación con columnas 5-6 en la tabla 3, o columna 8 en comparación con la columna 9 en la tabla 3). El valor medido en el mutante *er105* es siempre significativamente mayor que en la línea isogénica *ER* (columna 4 en la tabla 3). El valor medido en *er1* (línea parental Landsberg *NW20*) es habitualmente también mayor que el de las líneas *ER* 3177 (casi isogénicas, columna 6 de la tabla 3), y en un menor grado *Col4* (línea parental Columbia, columna 7 de la tabla 3). Estas observaciones dan evidencia directa de que el gen *ERECTA* desempeña un papel significativo en la determinación de diferencias genéticas en la discriminación isotópica de carbono en *Arabidopsis*.

Esta conclusión se confirma independientemente mediante las mediciones de intercambio de gases en la hoja que permiten la medición directa de la eficacia de transpiración (razón de la fijación de  $\text{CO}_2$  neta con respecto a la pérdida de agua; columna 4 en la tabla 4; figuras 1a-1c, 2a-2c). Mediciones en hojas maduras revelan que las líneas *ER* se caracterizan por una mayor razón de asimilación de  $\text{CO}_2$  con respecto a la pérdida de agua que las líneas que portan mutaciones *er*. Esto resulta más evidente cuando se compara el par *Col1/er105* con una eficacia de transpiración un 21% mayor (razón A/E) en *Col1* que *er105*, o el par *Col1/er2* con una eficacia de transpiración un 16% mayor en *Col1*. Concordando con las mediciones de discriminación de isótopos de carbono, el efecto *er/ER* es relativamente más pequeño en el contexto *Ld* (una razón A/E un 9% mayor en *Ld-ER* (3177) que en la línea *Ld-er1* (*NSW20*)).

También concuerda con las mediciones de discriminación de carbono la diferencia del 20% en la eficacia de transpiración entre las dos líneas parentales RIL (4,06 y 3,38 mmol C/mol  $\text{H}_2\text{O}$  en *Col4-ER* y *Ld-er1*, respectivamente).

Que de los 3 mutantes *erecta* examinados, *er105* presente la mayor discriminación de carbono más extrema y los fenotipos de eficacia de transpiración, sugiere que la mutación *er105* afecta a una parte más crucial del gen *ERECTA* que *er2* o *er1*. Esto concuerda con los datos publicados sobre el mutante *er105*. Esta mutación corresponde a la inserción de un “inserto foráneo” grande en el gen *ERECTA*. La inserción inhibe la transcripción del gen y provoca el fenotipo *erecta* más fuerte de todos los mutantes *erecta* aislados en Col (con respecto a agrupamiento de inflorescencias y forma y ancho de silicua). Alternativamente, o además, los datos indican que las mutaciones *erecta* presentan un efecto más fuerte sobre los valores de discriminación de isótopos de carbono en un contexto genético Columbia que en un contexto Landsberg (comparación de efectos fenotípicos de *er105* y *er1*), lo que implica que otros genes, polimórficos entre los ecotipos Landsberg y Columbia, interaccionan con *ERECTA* en la determinación de la eficacia de transpiración. Esto podría explicar también la diferencia mayor en la eficacia de transpiración entre líneas *er/ER* en contexto Col que en un contexto *Ld* (véase anteriormente, tabla 4). Alternativamente, o además, los datos indican que la mutación *erecta* no es la única mutación presente en el mutante *er105*. Por ejemplo, las semillas Col mutagenizadas pueden haber portado la mutación *gl1*, inducida mediante la irradiación de neutrones rápida, que también contribuye al fenotipo observado.

Una comparación de perfiles de transcrito en líneas isogénicas *er/ER* (en contexto tanto Col como *Ld*) permite la determinación de la participación de genes adicionales a *ERECTA* y el efecto del medio ambiente sobre su expresión.

### Ejemplo 7

Detección QTL centrada en el marcador *ERECTA* y el locus del gen *ERECTA* en el cromosoma 2 de *Arabidopsis thaliana*

#### 1. Procedimientos

Numerosas series que utilizan las líneas endogámicas recombinantes de Lister y Dean (1993) entre *Col-4* y *Ler-0* se



hicieron crecer en un invernadero de temperatura controlada (20/20°C) o dentro de cámaras de crecimiento (21°C y niveles de luz que oscilan desde 100 hasta 500  $\mu\text{E m}^{-2} \text{s}^{-1}$  de irradiancia y un 50-70% de humedad relativa). Las series incluyeron una variedad de las 100 RIL así como subconjuntos de estas 100 junto con líneas parentales Col-4 y líneas parentales Ler-0 (NW20). Se replicaron las RIL individuales dentro de las series. Las semillas o bien se trataron en frío en papel filtro húmedo durante 2-4 días, se trataron en frío y se plantaron directamente en el suelo; o bien se sembraron en placas en agar, tratadas en frío durante 2-4 días, que se hicieron crecer en agar durante aproximadamente 11-15 días antes de ser transferidas al suelo. Las plantas se regaron de manera abundante y se hicieron crecer durante 4-5 semanas antes de la cosecha. Se recogieron muestras (toda o parte de la roseta) y se secaron en un horno a 80°C antes de molerse y analizarse para determinar su composición isotópica de C. El valor utilizado para el análisis QTL para una línea individual era el promedio de las plantas replicadas de esa línea dentro de una serie.

## 2. Selección de marcadores

El conjunto convencional de 64 marcadores para las líneas recombinantes de Lister y Dean se descargaron de la página web de NASC. Se añadieron marcadores adicionales a este conjunto de datos cuando se determinó en primer lugar la significación para obtener mapeos a escala más fina en las regiones de interés. Un total de 121 marcadores se utilizaron en los 5 cromosomas.

## 3. Análisis

Se analizaron las series utilizando mapeo por intervalos simples (SIM) (Lander y Botstein 1989) y mapeo por intervalos compuestos (CIM) (Zeng 1993 y 1994). Se utilizaron dos programas para analizar los datos, QTL Cartographer versión 1.14 (Basten *et al.* 1999) y MQTL versión 0.98 (Tinker y Mathers 1995). Los dos programas difieren en cómo tratan los marcadores de contexto para mapeo por intervalos compuestos (CIM). En MQTL los marcadores de contexto se escogen al azar y se colocan en el archivo de entrada del mapeo. En QTL Cartographer los marcadores de contexto no se escogen al azar sino que más bien se escogen del análisis de regresión escalonada que selecciona los “mejores” marcadores de contexto. El establecimiento o la selección de estos marcadores también presenta una influencia en el nivel de significación estadística. Tinker argumenta que no es posible encontrar un umbral apropiado para el control del error estadístico cuando se seleccionan los marcadores de contexto basándose en los datos. Por tanto se utilizaron los dos programas y se han concentrado en los QTL que estuvieron presentes en ambos conjuntos de análisis.

### a) QTL cartographer

Se utilizó el Qstat para determinar si los datos presentaban una distribución normal (si no, entonces se tomaron medidas para arreglar la distribución). Se realizaron la regresión lineal (LR) y regresión escalonada (SR) utilizando los ajustes por defecto (la regresión escalonada se utilizó de manera progresiva con eliminación regresiva) al 5% de significación. Se realizaron la CIM e intervalos simples utilizando la función Zmap.qtl. Se analizaron los datos en todos los cromosomas con una velocidad de paso de 2 cM. Para el modelo 6 (CIM) el número de parámetros de contexto se dejó por defecto en 5 junto con el tamaño de ventana que se dejó por defecto en 10 cM. Se realizaron mil permutaciones con el CIM (Churchill y Doerge 1994). Entonces se ejecutó Eqtl para determinar los QTL significativos.

### b) MQTL

El mismo conjunto de marcadores utilizado en QTL Cartographer se utilizó en MQTL. Se escogieron los marcadores de contexto al azar para CIM. El número de marcadores escogido fue de aproximadamente la mitad del número de RIL utilizadas en el conjunto. Se seleccionó el ajuste por defecto de una velocidad de paso de 5 cM, se realizaron 3000 permutaciones para determinar los niveles de significación con el error de tipo 1 fijado al 5%.

Los QTL que estuvieron presentes en ambos programas y de conjuntos de marcadores de contexto variados del MQTL se consideraron genuinos. Esto, asociado con análisis QTL repetido en experimentos independientes, condujo a un locus repetible significativo que rodea al gen *ERECTA* en el cromosoma 2 (tabla 5).

Los datos en la tabla 5 indican que hay un QTL principal con una puntuación de LOD significativa en el nivel de probabilidad del 5% y, para la mayoría de series, de incluso el 1%, en el cromosoma 2 de *Arabidopsis thaliana*. En todos los casos, ese intervalo se encuentra por encima del marcador *ER* en el cromosoma 2. Dependiendo de la serie experimental, este QTL explica del 18 al 64% (véase la columna  $R^2$ ) de la varianza genética total en la eficacia de transpiración.

Los datos en la figura 3 indican un efecto aditivo positivo del QTL identificado basado en el valor medio de la composición de isótopos de carbono en plantas que portan el alelo Col-4 *ERECTA*.

**Ejemplo 8**

Prueba de complementación: transformación de líneas de *A. thaliana* que portan mutaciones *erecta* con el gen *ERECTA* de tipo natural bajo el control del promotor 35S.

1. Procedimientos

Se transformaron dos líneas *erecta* Columbia utilizando un vector binario cedido generosamente por Dr Keiko Torii. Se construyó ese plásmido utilizando el vector de plásmido pPZP222 (véanse detalles de este vector en Hajdukiewicz *et al.* Plant Mol Biol 25, 989-994, 1994). Los vectores pPZP portan genes quiméricos en un casete de expresión 35S de CaMV que les confiere resistencia a la kanamicina o gentamicina en plantas. El marcador seleccionable de la planta (gen de resistencia a la gentamicina para el vector pPZP222) se clona al lado del LB. Los sitios de clonación para el gen de interés (*ER* en nuestro caso) están entre el marcador de la planta y las secuencias RB. Esto garantiza que ese gen se transfiera en primer lugar a la planta, seguido del gen *gent*. Por tanto, la resistencia a la gentamicina se obtendrá sólo si el gen *ER* también está presente.

Se transfirió el vector binario a una cepa desarmada AGL1 de *Agrobacterium tumefaciens* mediante apareamientos triparentales convencionales (Ditta *et al.*, 1980, PNAS 77,7347-7351) utilizando la cepa cooperadora pRK2013 de *E. coli*.

Se transformaron las plantas *Arabidopsis* utilizando el procedimiento de inmersión floral convencional para la transformación mediante cepas desarmadas de *A. tumefaciens* (Clough y Bent, 1998, The Plant Journal 16, 735-743).

Se transformaron dos líneas *erecta* Columbia, para las que se tienen numerosos datos que muestran constantemente más valores isotópicos negativos en aquellas líneas (es decir eficacia de transpiración inferior) que en plantas de tipo natural Col *ER* casi isogénicas. Estas dos líneas fueron tal como sigue:

1. *er* 105, un mutante desactivado debido a la inserción de un fragmento grande de ADN en el gen *ERECTA* y

2. línea Col-*er2* (identificador 3401 NASC), el mismo que *er106* (Lease *et al.* 2001).

Se examinaron las plántulas en placas MS en 100 µg/ml de sulfato de gentamicina. Se transfirieron transformantes putativos al suelo y se examinó su progenie de nuevo para determinar la resistencia a la gentamicina, para la confirmación e identificación de líneas homocigóticas y la recolección de semillas T3.

Se obtuvieron muchas líneas transformantes independientes y entre aquellas había varias líneas homocigóticas *ER*, que se seleccionaron para su análisis posterior (véase la tabla 6).

También se obtuvo una línea *ER* Landsberg homocigótica transgénica estable transformando el ecotipo Ld-*er1* (NW20) con el mismo constructo que se describió anteriormente cedido por Dr Keiko Torii (línea T3-7K en la tabla 6 o "T2+ER" en las figuras 6-9).

2. Resultados:

El análisis inicial de varios transformantes ER en el contexto Col-*er105*, Col-*er106/er2* y Ld-*er1* (tal como se muestra en la tabla 6 anterior):

Se averiguó la transformación eficaz y se cuantificaron los niveles de expresión de ER en varios transformantes T2 independientes utilizando PCR cuantitativa en tiempo real (ABI PRISM 7700, Sequence Detection System User Bulletin #2. 1997). Básicamente esa técnica permitió cuantificar el número de copias del gen ER en líneas de interés, tras la normalización con respecto al número de copias de un gen control, en las mismas plantas (mismo conjunto de ARN). El gen de ARN ribosómico 18S se utilizó como gen control tras comprobar que su expresión no se veía afectada por los cambios en la expresión de ER.

Se muestran los resultados en las figuras 4a, 4b y 4c, en las que el eje de las y en cada figura describe el número de copias de ARNm de *erecta* (normalizado al de ARNm 18S) en las líneas *ER* de tipo natural, mutantes *er* y transgénicos ER en contextos tanto Columbia como Landsberg.

Todas las líneas transgénicas ER, excepto la línea 145 (figura 4a) mostraron un número de copias de ARNm aumentado: un aumento de desde 4 hasta 170 veces en comparación con los controles nulos. De manera interesante, todas las líneas, incluso aquellas con niveles de ARNm enormemente aumentados parecen "normales", sanas y de tamaño similar.

El análisis fenotípico inicial muestra complementación del "fenotipo de eficacia de transpiración". Dicho de otro modo, las líneas transgénicas *ER* muestran valores de composición isotópica de carbono menos negativos que el

control *er* nulo y las líneas nulas tal como se muestra en la tabla 7. Esos valores convergen hacia los valores medidos para los ecotipos ER de tipo natural. Por tanto, en un contexto Columbia, los transgénicos ER muestran valores del -30,6 al -31,2 por mil en promedio en comparación con valores del -31,7 al -32,2 por mil en los transgénicos nulos (tabla 7), y el -30,9 por mil en el tipo natural ER de Col0 (ecotipo de contexto para mutante *er*-105). Las composiciones isotópicas de carbono menos negativas en transgénicos ER son indicativas de una mayor eficacia de transpiración en estas plantas, como se esperaba.

Se confirman los datos presentados en la tabla 7 mediante la medición directa de la eficacia de transpiración de la hoja (razón A/E de la tasa de asimilación de CO<sub>2</sub> por unidad de área de hoja con respecto a la tasa de transpiración) utilizando técnicas de intercambio de gases. También se determinan la densidad estomática, capacidad fotosintética de la hoja y tasa de crecimiento para analizar las causas subyacentes de la reversión del fenotipo de eficacia de transpiración (anatomía y desarrollo de la hoja, propiedades bioquímicas de las hojas, características estomáticas).

### Ejemplo 9

Especificidad de tejido en la expresión del gen *ERECTA* en arroz de tipo natural *Oryza sativa* (cv Nipponbare):

Se identificó un ortólogo del alelo *ERECTA* de *A. thaliana* (SEC ID nº: 1) en arroz *in silico* mediante búsquedas de homologías de la base de datos de proteínas del NCBI utilizando el programa BLAST en condiciones convencionales. La secuencia de entrada fue SEC ID nº: 2. Se presenta la secuencia de nucleótidos del ortólogo del arroz en SEC ID nº: 3, comprendiendo la proteína codificada la secuencia de aminoácidos expuesta en SEC ID nº: 4.

Se determinó el número de copias de ARNm del gen *ERECTA* del arroz para diversos órganos/partes de la planta, tal como se indica en las figuras 5a y 5b. Se determinaron los números de copias de ARNm *ERECTA* mediante PCR cuantitativa en tiempo real, utilizando ARNm 18S como gen de control interno para la normalización de los datos. El patrón de expresión de *ERECTA* en arroz era similar al patrón de expresión génica en *A. thaliana*, observándose la mayor expresión en tejidos meristemáticos jóvenes, hojas jóvenes e incluso más, las inflorescencias. No se encuentra ninguna o muy baja expresión en raíces, en lo que respecta a *A. thaliana*.

Estas similitudes en la especificidad de tejido entre el arroz y *Arabidopsis* indican que el ortólogo del arroz proporcionado en la presente memoria como SEC ID nº: 3 es un ortólogo verdadero del alelo *ERECTA* de *A. thaliana* expuesto en SEC ID nº: 1, con función similar.

### Ejemplo 10

Demostración de un papel funcional para el gen *ERECTA* del arroz en la modulación de la eficacia de transpiración

Para determinar un papel funcional para el gen *ERECTA* del arroz (SEC ID nº: 3), se analizan líneas de plantas de arroz que portan inserciones de transposones que afectan a la expresión de ese gen. Se identificaron nueve de tales mutantes en la colección disponible públicamente de mutantes insercionales de transposones TOS 17 en el Japanese NIAS Institute. Se describe el retrotransposón TOS 17 en detalle por Hirochika, Current Opinion in Plant Biology, 4, 118-122, 2001 y por Hirochika Plant Mol Biol 35, 231-240, 1997. Se identificaron las nueve líneas mutantes a través de la página web URL <http://tos.nias.affrc.go.jp/~miyao/pub/tos17/>, y presentan los números de registro NG0578 (mutante A), ND3052 (mutante B), ND4028 (mutante C), NC0661 (mutante D), NE1049 (mutante E), NF8517 (mutante F), NE8025 (mutante G), NE3033 (mutante H) y NF8002 (mutante I).

Se pidieron nueve mutantes insercionales de transposones de NIAS, que portan el inserto de retrotransposón estable TOS17 en diversas partes del gen *ERECTA* en el contexto Nipponbare, el genotipo utilizado para el secuenciamiento del genoma del arroz: NG0578, ND3052, ND4028, NC0661, NE1049, NF8517, NE8025, NE3033 y NF8002.

Las inserciones de transposones en estas nueve líneas afectan la región transmembrana de la proteína (mutantes I, D, E) o los dominios de repeticiones ricas en leucina (LRR) (mutante H y G) en LRR 7 y LRR 18, respectivamente. En el mutante B, el TOS 17 altera la secuencia de codificación sólo en el sentido de 5' de las secuencias que codifican para el dominio I de proteína cinasa. En los mutantes C y F, la inserción de TOS17 altera la secuencia que codifica para el dominio VIa de la proteína *ERECTA*. En el mutante A, la inserción de TOS7 es en una secuencia que codifica para una región entre los dominios IX y X. La información de secuencia en estos mutantes está disponible públicamente de la página web del NIAS.

Utilizando una semilla mutante para líneas A-I recibidas de NIAS, se hicieron crecer las plantas para la amplificación de semilla y análisis. Excepto en dos mutantes en los que varias plantas murieron, las plantas parecían sanas, el buen crecimiento indicando que, como en *Arabidopsis*, existe gran potencial para alterar el gen *ERECTA* hacia una eficacia de transpiración alterada sin afectar adversamente al crecimiento y/o a la producción.

Basándose en la información de secuencia para cada mutante A-I, se diseñaron cebadores para amplificar los alelos *erecta* mutantes a partir del material de plántula derivado de 20 semillas. Se realizó la amplificación en condiciones

convencionales, para identificar para cada mutante, plantas que son homocigóticas, heterocigóticas o nulas en el locus *ERECTA*. Se identificaron los mutantes B y E TOS17 homocigóticos, y las líneas heterocigóticas y las líneas nulas en todas las líneas A-I.

5 En paralelo a la ganancia de información sobre si las líneas mutantes eran o no mutantes homocigóticas o heterocigóticas o nulas, se extraen partes de la planta específicas para el análisis de las consecuencias de las mutaciones en la expresión del gen *ERECTA* (niveles y especificidad de tejido de la expresión), utilizando PCR cuantitativa en tiempo real tal como se describe en la presente memoria. Adicionalmente, el fenotipo de eficacia de transpiración de cada línea mutante se determina midiendo la composición isotópica C y O y el contenido en ceniza de las muestras de planta.

10 Los resultados iniciales sobre la composición isotópica de  $^{13}\text{C}$  de limbos maduros de plántulas de arroz revela una variación significativa entre las líneas mutantes (del -32,8 al -34,2 por mil) y, en por lo menos 4 mutantes, desviaciones significativas con respecto a los valores de tipo natural, hacia valores más negativos, lo que sugiere que las mutaciones *erecta* sí afectan a la eficacia de transpiración en arroz, como en *Arabidopsis*.

15 Se aplican unos procedimientos similares a los anteriores para analizar la progenie de las plantas mutantes, para facilitar el análisis de los efectos de las mutaciones *erecta* en una gama de condiciones, que incluyen inundación (como es la práctica más común para Nipponbare), estrés hídrico tal como por secado del suelo (condiciones de crecimiento del arroz de secano) o humedad del aire baja (periodos de calor). Se observan diferencias en la morfología, anatomía y dominancia apical de la planta en cada condición ambiental. Los parámetros que se caracterizan incluyen patrones de labrado, la anatomía de las hojas y meristemos, tasa de crecimiento y desarrollo.

20 Las comparaciones entre mutantes A-I se utilizan adicionalmente para caracterizar el papel de los diferentes dominios de proteína a la hora de conferir diferentes fenotipos observados para cada línea en diferentes condiciones de crecimiento agrícola y/o ambientales. Es interesante que, entre los 4 mutantes que presentan composiciones isotópicas de C mucho más inferiores que el tipo natural, tres son aquéllos mutantes en los que el inserto TOS17 afecta a la región transmembrana.

### 30 **Ejemplo 11**

Efecto de silenciamiento de la expresión del gen *ERECTA* sobre la eficacia de transpiración

35 Para confirmar el papel del gen *ERECTA* a la hora de conferir el fenotipo de eficacia de transpiración a una planta, la expresión del alelo *ERECTA* de tipo natural se reduce o inhibe utilizando procedimientos convencionales de biología molecular vegetal, tales como, por ejemplo, inhibición antisentido de la expresión de *ERECTA* o la expresión de ARN de interferencia inhibidor (ARNi) que tiene como objetivo la expresión de *ERECTA* a nivel de ARN. Todos los procedimientos de este tipo se llevarán a cabo fácilmente por el experto en la materia utilizando las secuencias de nucleótidos dadas a conocer de los genes *ERECTA* proporcionados en la presente memoria o secuencias complementarias a las mismas.

Para la transformación del arroz y *Arabidopsis*, se preparan transgenes en vectores binarios, no tumorigenos, desarmados que portan ADN-T en los límites izquierdo y derecho y un marcador seleccionable operativo en *E. coli*.

45 Los vectores binarios utilizados para transferencia de ADN incluyen vectores seleccionados de entre el grupo constituido por:

1. pPZP222 (Hajdukiewicz *et al.*, 1994, Plant Mol Biol 25, 989-994);
- 50 2. PBI 121 (Clonetech) (Ueda *et al.* 1999, Protoplasma 206, 201-206);
3. pOCA18 (Olszewski *et al.* 1988, Nucl. Acid Res, 16 10765-10782);
4. pGreen y pSoup o variantes de los mismos (Hellens *et al.*, 2000, Plant Mol Biol 42, 819-832) y
- 55 5. vectores binarios desarrollados en las estructuras principales de los vectores pCAMBIA descritos en la página web de CAMBIA.

60 El material de partida para todos estos vectores era la estructura principal desarrollada por Hajdukiewicz *et al.*, 1994. Las series de vectores pPZP comprenden (i) un origen de replicación de intervalo de huésped amplio del plásmido pVSI de *Pseudomonas*, que es estable en ausencia de selección; (ii) el origen de replicación pBR322 (de tipo pMB9) para permitir preparaciones de ADN de alto rendimiento en *E. coli*; (iii) límites izquierdo (LB) y derecho (RB) del ADN-T, incluyendo la secuencia *overdrive*; y (iv) un casete de expresión de promotor CaMV35S. Aunque las series de vectores pPZP también sirvieron como las estructuras principales para las series pCAMBIA, se han modificado mucho para aplicaciones particulares.

Los vectores que contienen en su ADN-T diversas combinaciones de los siguientes componentes se prefieren particularmente:

- 5 1. casete del gen de resistencia hptII para conferir resistencia a la higromicina en material vegetal transformado, en el que la expresión de hptII está operativamente bajo el control de Ubi 1 o el promotor 3 5 S;
2. un casete de gen indicador que comprende ácido nucleico que codifica para los indicadores EGFP (proteína fluorescente verde mejorada) y/o beta-glucuronidasa (GUS y GUSPlus);
- 10 3. casete transactivador Gal4/VP16; y
4. uno o más casetes de expresión de genes de plantas que comprenden ADNc de genes *ERECTA* o bien de longitud completa o bien parciales en la orientación sentido o antisentido, o que pueden expresar ARNi que comprende secuencias derivadas del gen *ERECTA*, incluyendo cualquiera de los fragmentos genómicos de ADN
- 15 de la planta.

Se transfieren los vectores binarios a una cepa desarmada AGL1 de *Agrobacterium tumefaciens* mediante apareamientos triparentales convencionales (Ditta *et al*, 1980, Proc. Natl Acad. Sci. 77,7347-7351) utilizando la cepa cooperadora pRK2013 de *E coli*. Las plantas *A. thaliana* se transforman utilizando el procedimiento de inmersión floral convencional para la transformación mediante cepas desarmadas de *A tumefaciens* (Clough y Bent, 1998, The Plant Journal 16, 735-743). El arroz se transforma generando callos embriogénicos a partir de embriones cortados y sometiendo los callos embriogénicos a una transformación mediada por *Agrobacterium tumefaciens* según procedimientos publicados (por ejemplo Wang *et al* 1997, J Gen and Breed, 51 325-334, 1997). Se analizan las plantas transformadas para confirmar que aquellas líneas que expresan constructos de ARNi o antisentido presentan una expresión reducida de la proteína *ERECTA* funcional y se asemejan de manera más próxima al fenotipo *erecta* que las plantas de tipo natural o las plantas que expresan ectópicamente un gen *ERECTA* de tipo natural en la orientación sentido.

#### Ejemplo 12

30 Identificación de un ortólogo de sorgo de *ERECTA* de *A. thaliana*

Se identificó un ortólogo del alelo *ERECTA* de *A. thaliana* (SEC ID nº: 1) en sorgo *in silico* mediante búsqueda de homologías de la base de datos de proteínas del NCBI utilizando el programa BLAST en condiciones convencionales. La secuencia de entrada era SEC ID nº: 2. La secuencia de nucleótidos del ortólogo de sorgo se presenta en SEC ID nº: 5, comprendiendo la proteína codificada la secuencia de aminoácidos expuesta en SEC ID nº: 6.

#### Ejemplo 13

40 Identificación de homólogos de *ERECTA* de *A. thaliana*

Se identificaron dos homólogos del alelo *ERECTA* de *A. thaliana* (SEC ID nº: 1) *in silico* mediante búsqueda de homologías de la base de datos de proteínas del NCBI utilizando el programa BLAST en condiciones convencionales. La secuencia de entrada era SEC ID nº: 2. Las secuencias de nucleótidos de los homólogos de *ERECTA* de *A. thaliana* se presentan en las SEC ID nº: 7 y 9, comprendiendo las proteínas codificadas las secuencias de aminoácidos expuestas en las SEC ID nº: 8 y 10, respectivamente.

Los mutantes insercionales de ADN-T para estos dos genes homólogos, ambos en chr5, se han identificado en la colección mutante del Salk Institute (dirección web: signal.salk.edu/cg:-bin/tdnaexpress). Se pidieron varios de estos mutantes: Salk\_007643 y Salk\_026292 para el gen At5g07180; Salk\_045045 y Salk\_081669 para el gen At5g62230. Se diseñaron pares de cebadores con el fin de determinar número de copias de inserto y la homocigocidad/heterocigocidad en las plántulas que se hicieron crecer a partir de las semillas que se recibieron. Se identificaron las líneas homocigóticas con 1 inserto y están en caracterización con el fin de comparar los patrones de expresión (localización de tejido y niveles de ARNm) de los dos genes y del gen *ERECTA* en una gama de condiciones ambientales y determinar si los tres genes está relacionados funcionalmente.

#### Ejemplo 14

60 Identificación de ortólogos de trigo de *ERECTA* de *A. thaliana*

Se identificó inicialmente un secuencia de ADNc parcial de ortólogos del alelo *ERECTA* de *A. thaliana* (SEC ID nº: 1) en trigo *in silico* mediante búsqueda de homologías de la base de datos de proteínas del NCBI utilizando el programa BLAST en condiciones convencionales. Fue necesario, sin embargo, realizar búsquedas adicionales de bases de datos privadas con el fin de vincular las secuencias parciales identificadas en la base de datos del NCBI. La corrección de secuencias parciales ubicadas en la base de datos del NCBI fue también necesaria con el fin de

generar un cóntigo que correspondiera al ortólogo *ERECTA* de trigo.

La secuencia de entrada es la secuencia de aminoácidos de *A. thaliana* (SEC ID nº: 2) o de arroz (SEC ID nº: 4) o una secuencia de nucleótidos que codifica para la misma. Se presentan las secuencias de nucleótidos del ortólogo de trigo en las SEC ID nº: 11-19, comprendiendo las proteínas codificadas las secuencias de aminoácidos expuestas en SEC ID nº: 20.

Las secuencias expuestas en las SEC ID nº: 11 a 18 son secuencias de ADNc parciales. Se aísla la secuencia correspondiente del ortólogo *ERECTA* de trigo (SEC ID nº: 19) mediante un examen de hibridación de ácido nucleico convencional de una biblioteca de ADNc de trigo.

Para confirmar el papel de los ortólogos *ERECTA* de trigo en la eficacia de transpiración, se utilizan conjuntos de datos de expresión para estudios *in silico* de la expresión del gen *ERECTA* en una gama de tejidos de plantas de trigo que se hacen crecer en una gama de condiciones ambientales, proporcionando de ese modo indicaciones de especificidades de tejido en patrones de expresión y los datos preliminares sobre los tipos de ambientes en los que el ortólogo *ERECTA* es más probable que desempeñe un papel fisiológico en relación con el agua utilizada en esta especie. En estos estudios, los ácidos nucleicos que comprenden la secuencia expuesta en SEC ID nº: 11 a 19, o una secuencia complementaria a la misma, se utilizan para producir sondas de hibridación y/o cebadores de amplificación.

Adicionalmente, se introduce un gen *ERECTA* (SEC ID nº: 11 a 19) en la orientación sentido o antisentido en el trigo, produciendo de ese modo líneas de expresión transformadas. Los constructos génicos son específicamente para silenciar la expresión del gen *ERECTA* utilizando tecnología de ARNi, o alternativamente, para expresar ectópicamente el marco de lectura abierto completo del gen.

Basándose en una función similar, también se introduce el marco de lectura abierto del gen *ERECTA* de *A. thaliana* (es decir, SEC ID nº: 1) en el material vegetal del trigo en la orientación sentido, expresando de ese modo ectópicamente *ERECTA* de *A. thaliana* en trigo.

Se introducen los constructos génicos en el trigo siguiendo uno cualquiera de varios procedimientos convencionales, tales como, por ejemplo, utilizando la transformación mediada por *A. tumefaciens* tal como se describe en los documentos publicados AU 738153 o EP 856.060-A1 o CA 2.230.216 para Monsanto Company, o utilizando procedimientos de transformación biolística publicados tal como se describe por Pellegrineschi *et al.*, Genome 45(2), 421-30, 2002. Por lo tanto, la transformación genética se utiliza fácilmente para generar líneas de trigo con expresión alterada de un gen *ERECTA*. Aproximadamente se producen de 30 a 40 transformantes diferentes, dependiendo de la eficacia del ARNi en reducir la expresión de *ERECTA* en trigo.

Los transformantes primarios (T0) se caracterizan para determinar el número y loci en los que se insertan los transgenes. Se generan entonces las progenies segregantes T1 y T2 a partir de transformantes T0 seleccionados, y se analizan para determinar la razón de segregación y confirmar el número de loci que presentan los transgenes insertados. Se seleccionan aquellas líneas T1 y/o T2 que presentan inserciones transgénicas únicas y se utilizan para generar y multiplicar la semilla para estudios fisiológicos.

Se determina la eficacia de utilización de agua en las líneas T1 y/o T2 a través de (a) mediciones gravimétricas de aumentos de biomasa y agua transpirada; (b) discriminación isotópica de <sup>13</sup>C en tejidos de la planta, (es decir, determinando  $\Delta$ ; y (c) contenido en ceniza del tejido de la planta.

Se analiza también el desarrollo de las hojas y los meristemos, especialmente con respecto a la diferenciación y anatomía de la epidermis, los complejos estomáticos y el tejido mesófilo y examinando las propiedades de intercambio de gases de la hoja. Esto se realiza utilizando microscopía, técnicas de obtención de imágenes *in situ* y mediciones en línea concurrentes de discriminación isotópica de C ( $\Delta$ ) y de CO<sub>2</sub> y los flujos de agua que entran y salen de las hojas. La información sobre la regulación de genes y la red de genes en los que el ortólogo *ERECTA* opera en sus efectos sobre la eficacia de transpiración, se determina mediante el análisis de transcriptoma de un conjunto restringido de las líneas transgénicas con expresión de *ERECTA* alterada.

Tal como se describe en la presente memoria para *A. thaliana* y arroz, las correlaciones entre las mediciones fisiológicas y el número de copias o nivel de expresión de genes confirman el papel del ortólogo en conferir el fenotipo de eficacia de transpiración en trigo.

## 60 Ejemplo 15

Identificación de un ortólogo de maíz de *ERECTA* de *A. thaliana*

Se identificó inicialmente un secuencia de ADNc parcial de ortólogos del alelo *ERECTA* de *A. thaliana* (SEC ID nº: 1) en maíz *in silico* mediante búsqueda de homologías de la base de datos de proteínas del NCBI utilizando el programa BLAST en condiciones convencionales. Fue necesario, sin embargo, realizar búsquedas adicionales de

bases de datos privadas con el fin de vincular las secuencias parciales identificadas en la base de datos del NCBI. La corrección de secuencias parciales ubicadas en la base de datos del NCBI fue también necesaria con el fin de generar un cóntigo correspondiente al ortólogo *ERECTA* de maíz.

5 La secuencia de entrada fue SEC ID nº: 2. Se presenta la secuencia de nucleótidos de un ortólogo de maíz en las SEC ID nº: 21 a 44, comprendiendo la proteína codificada la secuencia de aminoácidos expuesta en SEC ID nº: 45.

10 Las secuencias expuestas en las SEC ID nº: 21 a 43 son secuencias de ADNc parciales. Se aisló la secuencia correspondiente del ortólogo de maíz (SEC ID nº: 44) mediante un examen de hibridación de ácido nucleico convencional de una biblioteca de ADNc de trigo.

15 Para confirmar el papel del ortólogo *ERECTA* de trigo en la eficacia de transpiración, se utilizan conjuntos de datos de expresión para estudios *in silico* de la expresión del gen *ERECTA* en una gama de tejidos de plantas de maíz que se hacen crecer en una gama de condiciones ambientales, proporcionando de ese modo indicaciones de especificidades de tejido en patrones de expresión y los datos preliminares sobre los tipos de ambientes en los que el ortólogo *ERECTA* es más probable que desempeñe un papel fisiológico en relación con el agua utilizada en estas especies. En estos estudios, el ácido nucleico que comprende la secuencia expuesta en SEC ID nº: 15, o una secuencia complementaria a la misma, se utiliza para producir sondas de hibridación y/o cebadores de amplificación.

20 Adicionalmente, se busca en las colecciones de mutantes de maíz etiquetadas con transposones para seleccionar aquéllos que presentan inserciones que afectan a la expresión del gen *ERECTA* y el nivel de expresión y/o número de copias del ortólogo *ERECTA* se correlaciona con la eficacia de transpiración en la gama de condiciones de crecimiento ambientales, esencialmente tal como se describe en la presente memoria para *A. thaliana* y arroz.

25 Adicionalmente, se introduce un gen *ERECTA* en la orientación sentido o antisentido en el maíz, produciendo de ese modo líneas de expresión transformadas. Los constructos génicos son específicamente para silenciar la expresión del gen *ERECTA* utilizando tecnología de ARN, o alternativamente, para expresar ectópicamente el marco de lectura abierto completo del gen.

30 Basándose en una función similar, también se introduce el marco de lectura abierto del gen *ERECTA* de *A. thaliana* (es decir, SEC ID nº: 1) en el material vegetal de maíz en la orientación sentido, expresando de ese modo ectópicamente *ERECTA* de *A. thaliana* en maíz.

35 Se introducen los constructos génicos en el maíz siguiendo uno cualquiera de varios procedimientos convencionales, tales como, por ejemplo, cualquiera de los procedimientos descritos por Gordon-Kamm *et al.*, Plant Cell 2(7), 603-618, 1990; la patente US nº 5.177.010 concedida a la University of Toledo; la patente US nº 5.981.840 concedida a Pioneer Hi-Bred; o la solicitud US publicada nº 20020002711 A1 (Goldman y Graves);. Por lo tanto, la transformación genética se utiliza para generar líneas de maíz con expresión alterada de un gen *ERECTA*.

40 Se producen de aproximadamente 30 a 40 transformantes diferentes, dependiendo de la eficacia del ARNi en reducir la expresión de *ERECTA*.

45 Los transformantes primarios (T0) se caracterizan para determinar el número y loci en los que se insertan los transgenes. Se generan entonces las progenies segregantes T1 y T2 a partir de transformantes T0 seleccionados, y se analizan para determinar la razón de segregación y confirmar el número de loci que presentan los transgenes insertados. Se seleccionan aquellas líneas T1 y/o T2 que presentan inserciones transgénicas únicas y se utilizan para generar y multiplicar la semilla para estudios fisiológicos.

50 Se determina la eficacia de utilización de agua en las líneas T1 y/o T2 a través de (a) mediciones gravimétricas de aumentos de biomasa y agua transpirada; (b) discriminación isotópica de <sup>13</sup>C en tejidos de la planta, (es decir, determinando  $\Delta$ ); y (c) contenido en ceniza del tejido de la planta.

55 Se analiza también el desarrollo de las hojas y los meristemos, especialmente con respecto a la diferenciación y anatomía de la epidermis, los complejos estomáticos y el tejido mesófilo y examinando las propiedades de intercambio de gases de la hoja. Esto se realiza utilizando microscopía, técnicas de obtención de imágenes *in situ* y mediciones en línea concurrentes de  $\Delta$  y de CO<sub>2</sub> y los flujos de agua que entran y salen de las hojas. La información sobre la regulación de genes y la red de genes en los que el ortólogo *ERECTA* opera en sus efectos sobre la eficacia de transpiración, se determina mediante el análisis de transcriptoma de un conjunto restringido de las líneas transgénicas con expresión de *ERECTA* alterada.

60 Tal como se describe en la presente memoria para *A. thaliana* y arroz, las correlaciones entre las mediciones fisiológicas y el número de copias o nivel de expresión de genes confirman el papel del ortólogo en conferir el fenotipo de eficacia de transpiración en maíz.

**Ejemplo 16**

Mecanismo de una eficacia de transpiración mejorada y herencia de *ERECTA* en *Arabidopsis* (contextos Landsberg y Columbia)

5 Se realizaron mediciones directas de la eficacia de transpiración (razón de la tasa de asimilación de CO<sub>2</sub> con respecto a la tasa de transpiración) en contextos tanto Landsberg como Columbia. Para confirmar el papel de *ERECTA* en una gama más amplia de condiciones relevantes de manera agronómica, se determinaron las eficacias de transpiración de plantas transformadas que portan un alelo *ERECTA* en respuesta a las condiciones ambientales variables (es decir, agua en el suelo y contenido en iones, humedad atmosférica y niveles de CO<sub>2</sub>) y se compararon con la respuesta de plantas de tipo natural (por ejemplo, ER). Los resultados de estos experimentos se presentan en las figuras 6-11, y las tablas 8 y 9.

15 Los datos en las figuras 6 muestran que la eficacia de transpiración mejorada obtenida insertando un transgén que porta el alelo ER de tipo natural en el mutante *Ld-er1* (línea T2+ER) se debe principalmente a una conductancia estomática disminuida. El fenotipo de la línea transgénica (T2+ER en los gráficos) es similar a la de un ecotipo *Ld-ER* casi isogénico a un *Ld-er1* obtenido del Stock Centre (línea 3177 en los gráficos). La eficacia de transpiración aumentada en ER transgénico, en comparación con los niveles observados en línea ER del tipo natural se observa tanto en niveles de CO<sub>2</sub> ambientales actuales como en niveles de CO<sub>2</sub> aumentados que están dentro de los límites que está previsto que se produzcan a nivel mundial durante las próximas dos décadas.

25 Los datos en la figura 8 muestran que la conductancia estomática reducida en la línea transgénica T2 ER en comparación con la línea *Ld-er1* está, por lo menos en gran parte, provocada por una densidad estomática reducida (reducción en el número de estomas por unidad de área en más de la mitad, hasta niveles similares a aquellos observados en el *Ld-ER* de tipo natural). Esta reducción en la densidad estomática es relativamente más alta que la que en la densidad de células epidérmicas cuya área superficial aumenta en sólo aproximadamente el 10%. Por tanto se deduce que el transgén ER ha afectado al desarrollo estomático, específicamente, y provocó una reducción en el índice estomático. Estos datos muestran complementación con respecto a los procesos que llevan a la variación en la eficacia de transpiración.

30 También se realizaron cruces recíprocos entre las dos líneas parentales NW20 (*Ld-er1*) y Col4 (*Col-ER*). La notación F2 (*Col\*Ld*) se refiere a las plantas F1 en las que Col era el receptor de polen *Ld*, mientras que la notación F1 (*Ld\*Col*) indica lo contrario (ovario *Ld* que recibe polen Col). Se ha realizado un análisis inicial de estos dos tipos de plantas F1 para determinar: las propiedades fotosintéticas y de intercambio de gases, la eficacia de transpiración (figura 7) y la composición isotópica de C (tabla 9), la tasa de desarrollo y la forma de la roseta, la anatomía de la epidermis de la hoja (figura 8), la fecha de floración, la inflorescencia y la forma de la vaina. Concordando con nuestros experimentos de análisis de complementación, los datos muestran que el gen *ERECTA* afecta a todos estos fenotipos y no sólo a la inflorescencia y la forma de la vaina.

40 Los datos también muestran una herencia compleja del gen *ERECTA*, de manera que el gen es dominante, sin efecto recíproco sobre la forma de la vaina (vainas más largas, tallos más largos y pedicelos en todas las plantas F<sub>1</sub>, similar a la parental *Col-ER*). Sin embargo, para otros rasgos, los resultados indican efectos maternos: por tanto los valores de eficacia de transpiración (véase la figura 7a) y composición de isótopos de carbono de rosetas en plantas F<sub>1</sub> (tabla 9) son intermedios entre los valores parentales, pero diferentes entre los dos conjuntos de plantas F<sub>1</sub>: los valores para las plantas F<sub>1</sub> (*Col\*Ld*) están más próximos a los valores Col, mientras que aquéllos para las plantas F<sub>1</sub> (*Ld\*Col*) están más próximos a los valores para el parental *Ld*.

50 Los datos en la figura 8 indican que la conductancia estomática (transpiración por unidad de área de hoja, figura 8a) muestra valores próximos al parental *Ld-er1* en todas las plantas F<sub>1</sub>, a pesar de que las densidades estomáticas están próximas a la del parental *Col-ER* (figura 8c). Esto muestra que el gen ER no sólo afecta al desarrollo de la epidermis sino también a la apertura estomática (dinámica de los estomas) y que aunque el efecto de ER sobre densidad estomática parece ser dominante, el efecto sobre la apertura estomática no lo es.

55 Los datos en la figura 9 muestran el efecto de diversas mutaciones *er* (en el contexto Col, mutantes obtenidos del Stock Centre o Dr Torii) en el número de estomas por unidad de área de hoja. Las densidades estomáticas para todos a excepción de dos de esos mutantes son mayores que las de las hojas de tipo natural *ColER*, y confirman el efecto del gen *ERECTA* sobre ese parámetro.

60 Los datos en la figura 10 muestran que se confirma una eficacia de transpiración mejorada en la línea transgénica ER en comparación con *Ld-er1* nulo (sin inserción del transgén) mediante los valores de la composición isotópica de C menos negativos medidos en el material de la hoja (comparar valores para líneas NW20 y CS20 (*Ld-er1*; líneas 16 y 17 en el eje de las x) y una línea *Ld-ER* T2 transgénica, homocigótica para el transgén ER (línea 19 en la figura). Los valores isotópicos de C medidos en la línea transgénica ER son similares a aquéllos en los ecotipos *Ld-ER* casi isogénicos (línea 18 en la figura 10). Esto demuestra la complementación en este rasgo fenotípico, y valida una vez más la utilización de la composición isotópica de C como un indicador cuantitativo (sustituto) de la eficacia de transpiración.



Los datos en la figura 10 también las composiciones isotópicas de C de una gama de mutantes Col-er, incluyendo aquéllos que analizados en la figura 9 para determinar las densidades estomáticas. La mayoría de los mutantes muestran valores isotópicos de C más negativos que el ecotipo COL-ER. Estos concuerda con las densidades estomáticas aumentadas descritas en la figura 9 y con todas las otras comparaciones de composiciones isotópicas de C o mediciones directas de eficacias de transpiración en las líneas er/ER y de nuevo indicativo del efecto positivo del alelo ER sobre la eficacia de transpiración.

Algunos mutantes en la figura 10 destacan, por ejemplo Col-er105, o línea 3140 (una línea de NASC que porta mutaciones er1 y gl1-1). Como la información genética para estos mutantes está disponible (naturaleza y posición de las mutaciones) estos mutantes proporcionan información funcional muy útil funcional sobre el/los dominio(s) de proteína de la proteína *ERECTA* que son esenciales para conferir el fenotipo de eficacia de transpiración y los procesos subyacentes.

Se realizaron también mediciones directas de la eficacia de transpiración (razón de la tasa de asimilación de CO<sub>2</sub> con respecto a la tasa de transpiración) en varios transformantes T2 generados en un contexto Columbia (es decir transformación del mutante er-105 y er-2/106 anterior). Los resultados de estas mediciones se muestran en la figura 11. Estos datos muestran que el fenotipo puede complementarse en un contexto Columbia, tal como se determina midiendo la eficacia de transpiración, tasas de asimilación de CO<sub>2</sub> y transpiración. Se observa complementación en condiciones tanto de humedad alta como de humedad baja, de ahí la demostración de que el gen *ERECTA* desempeña un papel en el control de la eficacia de transpiración en condiciones tanto de riego abundante como de sequía, y que la sobreexpresión de ese gen tiene el potencial para aumentar el crecimiento y la resistencia a la sequía y los estreses relacionados con la sequía.

Más particularmente, los datos en la figura 11 demuestran el papel del gen *ERECTA* sobre la eficacia de transpiración en una gama de humedades, incluyendo humedades bajas tales como las que prevalecen en áreas cálidas y secas:

- El mutante er-105 portador de una mutación inactiva de *ERECTA* (casi ningún transcrito ER) (cuadros negros abiertos) en contexto Col0 presenta una eficacia de transpiración inferior a la del Col0 casi isogénico de tipo natural (triángulos abiertos).
- Este mutante se transformó con un transgén ER bajo el promotor 35S y se produjeron varias líneas T2 homocigóticas ER (círculos opacos). Aquellas líneas (5 transformantes independientes se incluyen en el gráfico) presentan eficacias de transpiración muy aumentadas (del +40 al 70%) en comparación con las líneas nulas (cuadrados opacos) y similares a aquellas medidas para la línea ER-Col0 de tipo natural, en toda la gama de déficits de presión de vapor en las hojas con respecto al aire sometidos a prueba en los experimentos.

Adicionalmente, las líneas nulas que no son portadoras de ninguna inserción de transgén pero que fueron sometidas a transformación y selección en antibióticos muestran valores similares como el mutante er-105 de partida lo que demuestra que estas manipulaciones por sí mismas no presentan efectos de confusión detectables sobre la eficacia de transpiración.

TABLA 1

Análisis QTL de discriminación de isótopos de carbono en las líneas endogámicas recombinantes de Lister y Dean						
SERIE n <sup>o</sup> Condiciones experimentales	locus de chr2 (cM)	Método de análisis QTL	locus de chr4 (cM)	CONCLUSIÓN		
				número de QTL	posición en el mapeo prevista	
Serie 1 (40 líneas) Invernadero						
12 h de duración de día	58,5	SIM y CIM		2	chr2:	58,5-61,02
Irradiancia 150-350 μE m <sup>-2</sup> s <sup>-1</sup>	46,77	SIM y CIM			chr2:	46,77-50,75
Plántulas transferidas de placas con agar	61,02	SIM	108,5			
Datos de la serie 1 pero con la utilización de diferentes marcadores						
	de 56,94 a 58,00	CIM y SIM		1		
	de 46,77 a 50,75	SIM				
	63,02					
Serie 1 con diferente número de líneas						
	de 58,5 a 61,02			2		
	56-61					

TABLA 1

Análisis QTL de discriminación de isótopos de carbono en las líneas endogámicas recombinantes de Lister y Dean						
SERIE n° Condiciones experimentales	locus de chr2 (cM)	Método de análisis QTL	locus de chr4 (cM)	CONCLUSIÓN		
				número de QTL	posición en el mapeo prevista	
Serie 2 Invernadero septiembre de semillas sembradas en suelo						
lote 1	50,75	CIM (QTL cart)		2	chr2:	56,94-61,02
	61,02	MQTL			chr2:	50,75
lote 2	50,75	MQTL		NS		
lotes 3-5						
todos los lotes	58,5	MQTLcart		NS		
Serie 2	56,94-58,5	MQTL				
Serie 3 37 líneas: parentales y líneas con entrecruzamientos en el cromosoma 2 5 condiciones de crecimiento que difieren en humedad, irradiancia, modo de establecimiento (semillas sembradas en suelo o plántulas trasplantadas de agar)						
lote B	61,02-61,06		108 NS			
lote C	56,94-58,00					
lote D	63,02	QTLcar				
	63,02	MQTL				
todos los lotes (condiciones)	58,5			lor 2	chr2:	56,94-58,5
	61,02			3	chr2:	61,02-63,02
Serie 4 mismas líneas que en la serie 3 cámaras de crecimiento 10 h de luz de día						
	50,74				chr2:	50,74
Serie 5 repetición de la serie 1 PERO TODAS las líneas						
	50,74			1	chr2:	50,74
Serie 7 mismas líneas que en la serie 1 pero en cámara de crecimiento y luz más alta						
10 h de duración de día	46,77-50,75	CIM y SIM		1	chr 2:	46,77-5065
Serie 7 470-510 □E m <sup>-2</sup> s <sup>-1</sup> irradiancia						

5

TABLA 2

Contexto	Mutación	Nombre Stock Centre	línea ER isogénica y nombre Stock Centre
Landsberg	<i>er1</i>	CS20 o NW20 <sup>a</sup>	3177 o CS163
Columbia	<i>er2<sup>b</sup>/er106</i>	3401	Col1 o 3176
Columbia	<i>er105<sup>c</sup></i>		Col3 con marcador gl1 o Col0

a, NW20 es un parental *Ler* para las líneas recombinantes de Lister y Dean, que porta la mutación *er1*. Las líneas 3177 o CS163 son las líneas ER isogénicas más próximas.

b, *er2* es un alelo *er* identificado por Rédéi en contexto Col. Coll o 3176 son las líneas casi isogénicas Col más próximas. La *er2* es la misma mutación que la mutación *er-106* notificada posteriormente por Torii y colaboradores (Lease *et al.* 2001)

c, *er105* se aisló a partir de una población de semillas Col sometidas a irradiación de neutrones rápida (Torii *et al.*, 1996).

d, Col4, el parental *Col* para el parental de Lister y Dean se incluyó sistémicamente en todas las comparaciones.

TABLA 3

Comparación de líneas er/ER tanto en contexto Col como Ld para los valores de discriminación de isótopos de carbono (por mil) en material de hoja en una gama condiciones ambientales									
Serie nº	Diferencias en los valores de discriminación de isótopos de carbono medios (por mil)								
	(1) <i>er-ER</i> (todas las líneas)	(2) <i>er105-Col0</i>	(3) <i>er105-Col4</i>	(4) <i>er105-3176</i>	(5) <i>er2-3176</i>	(6) <i>er1-3177</i>	(7) <i>er1-Col4</i> (líneas parentales para las RIL)	(8) <i>er105-Coli</i>	(9) <i>er1-Coli</i>
1	0,13	0,16						0,16	
2	0,89	1,18						1,18	
3	0,26	0,11						0,11	
4	1,12	1,60						1,60	
5	1,03		1,83	1,67	0,92	0,64	0,82	1,75	0,73
6	0,70	1,13	1,01	0,71	0,27	0,74	0,73	0,95	0,73
7	0,70	1,32	1,12	1,23	0,75	0,35	0,05	1,22	0,15
8	0,59	1,16	1,11	1,19	0,54	0,28	0,06	1,16	0,17
9	0,30	1,09	0,77	0,77	0,02	0,00	0,56	0,88	0,28
10	0,56	1,05	0,94	0,87	0,38	0,39	0,33	0,95	0,36
11	0,48		0,40				0,52	0,40	0,52
12	0,36	0,82	1,31	1,08	0,33	0,05	0,07	1,07	0,01
13	0,38		0,90	0,82	0,07	0,60	0,52	0,86	0,56
14	0,65	1,42	0,60			0,58	0,06	1,01	0,32
15	0,82						0,82		0,82
Para todas las series:									
Media	0,60	1,01	1,00	1,04	0,41	0,40	0,41	0,95	0,42
E.E.	0,07	0,14	0,11	0,11	0,10	0,08	0,09	0,12	0,08
Para series comunes:									
Media:	0,58	1,10	1,12	1,04	0,41	0,38	0,39	1,11	0,37
E.E.	0,08	0,05	0,11	0,11	0,10	0,09	0,10	0,10	0,09

TABLA 4

Serie 9 de diciembre de 2001: Mediciones de intercambio de gases en las hojas en líneas de *Arabidopsis er/ER*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Genotipo	E	A	Gw	A/E	pa	pi	pi/pa	1-pi/pa
	(mmol H <sub>2</sub> O/m <sup>2</sup> /s)	( $\square$ mol C/m <sup>2</sup> /s)	(molm <sup>2</sup> /s)	(mmolC/molH <sub>2</sub> O)	( $\square$ bar)	( $\square$ bar)		
Fila (1)								
Ld-R	Media	12,33	0,273	3,67	360	282	0,782	0,218
	E.E.	1,64	0,039	0,14	10	11	0,010	0,010
Fila (2)								
Ld-er	Media	8,73	0,218	3,38	348	280	0,804	0,196
	E.E.	0,31	0,005	0,04	5	4	0,002	0,002
Fila (3)								
Col-ER	Media	13,55	0,291	4,06	350	270	0,772	0,228
	E.E.	1,16	0,040	0,22	4	7	0,020	0,020
Fila (4)								
Col-ER	Media	10,13	0,180	4,55	346	254	0,734	0,266
	E.E.	1,47	0,048	0,24	5	9	0,021	0,021
Fila (5)								
Col-er	Media	8,55	0,198	3,76	356	283	0,795	0,205
	E.E.	0,17	0,005	0,07	11	10	0,006	0,006
Fila (6)								
Col-er	Media	11,90	0,256	3,92	357	279	0,780	0,220
	E.E.	0,56	0,027	0,12	1	6	0,014	0,014
<b>CONCLUSIÓN:</b>								
Comparación Ld-ER/Ld-er	La línea er presenta una menor A/E con menor g y mayor A. La diferencia en A/E se debe a A							
Comparación 933/NW20	NSW20 (er1) presenta una menor A/E con menor g y mayor A							
Comparación Col1/Ld-er1	La diferencia en A/E se debe a A							
Comparación Col1/Col-er105	er105 presenta una A/E MUCHO menor con mayor g y menor A es decir, la diferencia en A/E se debe a A y g							
Comparación Col1/Col-er2	er2 presenta menor A/E con MUCHO mayor g y MAYOR A es decir, la diferencia en A/E se debe a g y se opone o no se debe a A							
NOTA: p <sub>a</sub> y p <sub>i</sub> son las presiones parciales ambiente e intercelular de CO <sub>2</sub> , respectivamente.								

TABLA 5

Serie experimental	LOD	P(0,05)	P(0,01)	Hueco de cM en la posición QTL	LOD-LOD <sub>p0,05</sub>	R <sup>2</sup>	R <sup>2</sup> total
1	5,2861	4,7469	5,9923	39,32-50,65	0,5392	0,19	0,89
4	3,519	3,3561	4,2134	50,63-50,65	0,1629	0,29	0,57
7	9,6489	3,9627	4,9083	50,63-50,65	5,6862	0,53	0,77
9	6,1748	4,2328	5,0278	46,77-50,65	1,942	0,44	0,81
10	Cero Qtl	3,6051	4,3889				
16	11,5132	3,188	3,8896	48,96-50,65	8,3252	0,64	0,65
2 (lote1)	5,5459	3,3264	4,2907	50,63-51,02	2,2195	0,26	0,40

LOD = log<sub>10</sub> de la razón de probabilidades

TABLA 6

5

Tabla resumen de las líneas utilizadas para la caracterización funcional inicial y el análisis de efectos ER:		
Ecotipo de contexto	Transformantes ER homocigóticos T2 estables	Control er nulo
Col-er105	T8; T29; T19; T61	T18
Col-er106/er2/3401	T165; T169; T279; T290	T143
Ld-er1	T3-7K	NW20

TABLA 7

Composición de isótopos de carbono (por mil) de 3 hojas maduras, molidas conjuntamente, recogidas el 4/6/03 es decir 32 días tras la siembra a partir de rosetas todavía vegetativas					
	Transformantes homocigóticos ER T2		Transgénicos nulos (er)	Mutante er de contexto	
Línea:	T46	T29	T18	Col-er105	
	-31,4	-31,2	-32,2		
Línea:	T145	T165	T279	T290	Col er106/er2/3401
	-30,4	-31	-30,5	-30,8	
Promedio:		-30,6			-31,7
Ee:		0,15			0,12
línea	T3-7K			Ld-er1	NW20
	-30,4				-31,3

10

TABLA 8

Línea de alelo erecta	Nombre de la línea	Composición isotópica de C (por mil)		Serie 18 Promedio	Err. est.	
		Serie 14 Promedio	Err. est.			
Mutantes Col_er	102K			-29,4	0,011	
	103K			-29,0	0,10	
	105C	-31,4	0,12	-30,3	0,09	
	105KH	-30,2		-29,3	0,11	
	105KS	-29,8	0,07	-29,5	0,07	
	3401	-30	0,21	-29,4	0,05	
	106C	-30,2	0,04	-29,4	0,12	
	108K	-30,2	0,07	-29,5	0,06	
	111KH	-30,4	0	-29,5	0,11	
	111KS	-30,2	0,11	-29,7	0,12	
	114K	-30,2	0,11	-29,5	0,04	
	116K	-29,7	0,21	-29,0	0,13	
	117K	-29,7	0,18	-29,3	0,08	
	3140			-32,2	0,06	
	Col0_ER	1093	-29,6	0	-29,0	0,10
	Ld_er1	NW20	-30,0	0,25	-28,9	0,15
		CS20	-29,9	0,08		
Ld_ER Ld_er1 transgénico + ER de tipo natural	3177	-29,4		-28,2	0,09	
	3-7K	-29,5	0,07	-28,4	0,10	

TABLA 9

	Composición de isótopos de C (por mil)
Col-ER (línea 933)	-28,1
F1 (Col*Ld)	-28,5
F1 (Ld*COL)	-29,3
Ld-er1 (línea NW20)	-29,9

**Listado de secuencias**

5	<110> The Australian National University	
	<120> PROCEDIMIENTO PARA PRODUCIR PLANTAS QUE PRESENTAN UNA EFICACIA DE TRANSPIRACIÓN MEJORADA Y PLANTAS PRODUCIDAS A PARTIR DEL MISMO	
10	<130> 94948/MRO	
	<150> AU PS3339	
	<151> 2002-07-02	
15	<160> 45	
	<170> PatentIn versión 3.1	
20	<210> 1	
	<211> 3176	
	<212> ADN	
	<213> Alelo <i>ERECTA</i> de <i>Arabidopsis thaliana</i>	
25	<400> 1	
	gtttcttctt catggagact.tgaaagcttt taaagtatat ctaaaaacgc agtcgtttta	60
	agactgtgtg tgagaaatgg ctctgttttag agatattggt cttcttgggt ttctcttctg	120
	cttgagctta gtagctactg tgacttcaga ggaggagca acgttgctgg agattaagaa	180
	gtcattcaaa gatgtgaaca atgttcttta tgactggaca acttcacctt cttcggatta	240
	ttgtgtctgg agaggtgtgt cttgtgaaaa tgtcaccttc aatgttgttg ctcttaattt	300
	gtcagatttg aatcttgatg gagaaatctc acctgctatt ggagatctca agagtctctt	360
	gtcaattgat ctgagagga atcgcttgtc tggacaaatc cctgatgaga ttggtgactg	420
	ttcttctttg caaaacttag acttatcctt caatgaatta agtggtgaca taccgttttc	480
	gatttcgaag ttgaagcaac ttgagcagct gattctgaag aataaccaat tgataggacc	540
	gatcccttca acactttcac agattccaaa cctgaaaatt ctggacttgg cacagaataa	600
	actcagtggg gagataccea gacttattta ctggaatgaa gttcttcagt atcttgggtt	660
	gcgaggaaac aacttagtgc gtaacatttc tccagatttg tgcactga ctggtctttg	720
	gtattttgac gtaagaaaca acagtttgac tggtagtata cctgagacga taggaaattg	780
	cactgccttc caggttttgg acttgccta caatcagcta actggtgaga tcccttttga	840
30	catcgcttc ctgcaagttg caacattatc attgcaaggc aatcaactct ctggaagat	900

# ES 2 362 903 T3

tccatcagtg attggtctca tgcaagccct tgcagtctta gatctaagtg gcaacttggt 960  
 gagtggatct attcctccga ttctcggaaa tcttaacttc accgagaaat tgtatttgca 1020  
 cagtaacaag ctgactgggt caattccacc tgagcttga aacatgtcaa aactccatta 1080  
 cctggaactc aatgataatc atctcacggg tcatatacca ccagagcttg ggaagcttac 1140  
 tgacttgttt gatctgaatg tggccaacaa tgatctggaa ggacctatac ctgatcatct 1200  
 gagctcttgc acaaatctaa acagcttaa tgttcatggg aacaagtta gtggcactat 1260  
 accccgagca ttcaaaagc tagaaagtat gacttacctt aatctgtcca gcaacaatat 1320  
 caaaggtcca atcccgggtg agctatctcg taccgtaac ttagatacat tggatcttcc 1380  
 caacaacaag ataaatggaa tcattccttc ttcccttggg gatttggagc atcttctcaa 1440  
 gatgaacttg agtagaaatc atataactgg ttagttcca ggcgacttgg gaaatctaag 1500  
 aagcatcatg gaaatagatc ttcaaaataa tgatatctct ggccaatc cagaagagct 1560  
 taaccaatta cagaacataa ttttctgag actggaaaat aataacctga ctggtaatgt 1620  
 tggttcatta gccaaactgc tcagctctac tgtattgaat gtatctcata acaacctcgt 1680  
 aggtgatatc cctaagaaca ataactctc aagatttca ccagacagct tcattggcaa 1740  
 tcttggctct tgcggtagtt ggctaaactc accgtgtcat gattctcgtc gaactgtacg 1800  
 agtgtcaatc tctagagcag ctattcttgg aatagctatt gggggacttg tgatccttct 1860  
 5 catggtctta atagcagctt gccgaccgca taatcctct ccttttcttg atggatcaact 1920  
 tgacaacca gtaacttatt cgacaccgaa gctcgtcatc ctcatatga acatggcaact 1980  
 ccacgtttac gaggatatca tgagaatgac agagaatcta agtgagaagt atatcattgg 2040  
 gcacggagca tcaagcactg tacaacaatg tgttttgaag aattgtaaac cggttgcat 2100  
 taagcggctt tactctcaca acccacagtc aatgaaacag tttgaaacag aactcgagat 2160  
 gctaagtagc atcaagcaca gaaatcttgt gagcctacaa gcttattccc tctctcactt 2220  
 ggggagtctt ctgttctatg actatttga aatggtagc ctctgggatc ttcttcatgg 2280  
 ccctacgaag aaaaagactc ttgattgga cacacggctt aagatagcat atggtgcagc 2340  
 acaaggttta gcttatctac accatgactg tagtccaag atcattcaca gagacgtgaa 2400  
 gtcgtccaac attctcttgg acaagactt agaggctcgt ttgacagatt ttggaatagc 2460  
 gaaaagcttg tgtgtgtcaa agtcacatac ttcaacttac gtgatgggca cgataggtta 2520  
 catagacccc gagtatgctc gcacttcacg gctcactgag aaatccgatg tctacagtta 2580  
 tggaaatgac cttcttgagt tgttaacccg aaggaaagcc gttgatgacg aatccaatct 2640  
 ccacatctg ataatgtcaa agacggggaa caatgaagtg atggaaatgg cagatccaga 2700  
 catcacatcg acgtgtaag atctcgggtg ggtgaagaaa gttttccaac tggcactcct 2760  
 atgcacaaa agacagccga atgatcgacc cacaatgac caggtgactc gtgttctcgg 2820  
 cagttttatg ctatcggaac aaccacctgc tgcgactgac acgtcagcga cgctggctgg 2880

# ES 2 362 903 T3

```

ttcgtgctac gtcgatgagt atgcaaatct caagactcct cattctgtca attgctcttc 2940
catgagtgct tctgatgctc aactgtttct tcggtttga caagttattt ctcagaacag 3000
tgagtagttt ttcgtagga ggagaatctt taaaacggta tcttttcggt gcgtaagct 3060
gtagaaaaa ttaatgtctc atgtaaagta ttatgcactg ccttattatt attagacaag 3120
tgtgtggtgt gaatatgtct tcagactggc acttagactt cctataagtt cttgcc 3176

```

<210> 2

5 <211> 976

<212> PRT

<213> Alelo *ERECTA* de *Arabidopsis thaliana*

<400> 2

10

```

Met Ala Leu Phe Arg Asp Ile Val Leu Leu Gly Phe Leu Phe Cys Leu
1           5           10           15

```

```

Ser Ieu Val Ala Thr Val Thr Ser Glu Glu Gly Ala Thr Leu Leu Glu
           20           25           30

```

```

Ile Lys Lys Ser Phe Lys Asp Val Asn Asn Val Leu Tyr Asp Trp Thr
           35           40           45
Thr Ser Pro Ser Ser Asp Tyr Cys Val Trp Arg Gly Val Ser Cys Glu
           50           55           60

```

```

Asn Val Thr Phe Asn Val Val Ala Leu Asn Leu Ser Asp Leu Asn Leu
65           70           75           80

```

```

Asp Gly Glu Ile Ser Pro Ala Ile Gly Asp Leu Lys Ser Leu Leu Ser
           85           90           95
Ile Asp Leu Arg Gly Asn Arg Leu Ser Gly Gln Ile Pro Asp Glu Ile
           100           105           110

```

```

Gly Asp Cys Ser Ser Leu Gln Asn Leu Asp Leu Ser Phe Asn Glu Leu
           115           120           125

```

```

Ser Gly Asp Ile Pro Phe Ser Ile Ser Lys Leu Lys Gln Leu Glu Gln
           130           135           140
Leu Ile Leu Lys Asn Asn Gln Leu Ile Gly Pro Ile Pro Ser Thr Leu
145           150           155           160

```

```

Ser Gln Ile Pro Asn Leu Lys Ile Leu Asp Leu Ala Gln Asn Lys Leu
           165           170           175

```

```

Ser Gly Glu Ile Pro Arg Leu Ile Tyr Trp Asn Glu Val Leu Gln Tyr
           180           185           190

```



Leu Gly Leu Arg Gly Asn Asn Leu Val Gly Asn Ile Ser Pro Asp Leu  
 195 200 205

Cys Gln Leu Thr Gly Leu Trp Tyr Phe Asp Val Arg Asn Asn Ser Leu  
 210 215 220

Thr Gly Ser Ile Pro Glu Thr Ile Gly Asn Cys Thr Ala Phe Gln Val  
 225 230 235 240  
 Leu Asp Leu Ser Tyr Asn Gln Leu Thr Gly Glu Ile Pro Phe Asp Ile  
 245 250 255

Gly Phe Leu Gln Val Ala Thr Leu Ser Leu Gln Gly Asn Gln Leu Ser  
 260 265 270

Gly Lys Ile Pro Ser Val Ile Gly Leu Met Gln Ala Leu Ala Val Leu  
 275 280 285  
 Asp Leu Ser Gly Asn Leu Leu Ser Gly Ser Ile Pro Pro Ile Leu Gly  
 290 295 300

Asn Leu Thr Phe Thr Glu Lys Leu Tyr Leu His Ser Asn Lys Leu Thr  
 305 310 315 320  
 Gly Ser Ile Pro Pro Glu Leu Gly Asn Met Ser Lys Leu His Tyr Leu  
 325 330 335

Glu Leu Asn Asp Asn His Leu Thr Gly His Ile Pro Pro Glu Leu Gly  
 340 345 350

Lys Leu Thr Asp Leu Phe Asp Leu Asn Val Ala Asn Asn Asp Leu Glu  
 355 360 365  
 Gly Pro Ile Pro Asp His Leu Ser Ser Cys Thr Asn Leu Asn Ser Leu  
 370 375 380

Asn Val His Gly Asn Lys Phe Ser Gly Thr Ile Pro Arg Ala Phe Gln  
 385 390 395 400

Lys Leu Glu Ser Met Thr Tyr Leu Asn Leu Ser Ser Asn Asn Ile Lys  
 405 410 415  
 Gly Pro Ile Pro Val Glu Leu Ser Arg Ile Gly Asn Leu Asp Thr Leu  
 420 425 430

Asp Leu Ser Asn Asn Lys Ile Asn Gly Ile Ile Pro Ser Ser Leu Gly  
 435 440 445

Asp Leu Glu His Leu Leu Lys Met Asn Leu Ser Arg Asn His Ile Thr  
 450 455 460

5

Gly Val Val Pro Gly Asp Phe Gly Asn Leu Arg Ser Ile Met Glu Ile  
 465 470 475 480

Asp Leu Ser Asn Asn Asp Ile Ser Gly Pro Ile Pro Glu Glu Leu Asn  
 485 490 495  
 Gln Leu Gln Asn Ile Ile Leu Leu Arg Leu Glu Asn Asn Asn Leu Thr  
 500 505 510

Gly Asn Val Gly Ser Leu Ala Asn Cys Leu Ser Leu Thr Val Leu Asn  
 515 520 525

Val Ser His Asn Asn Leu Val Gly Asp Ile Pro Lys Asn Asn Asn Phe  
 530 535 540  
 Ser Arg Phe Ser Pro Asp Ser Phe Ile Gly Asn Pro Gly Leu Cys Gly  
 545 550 555 560

Ser Trp Leu Asn Ser Pro Cys His Asp Ser Arg Arg Thr Val Arg Val  
 565 570 575

Ser Ile Ser Arg Ala Ala Ile Leu Gly Ile Ala Ile Gly Gly Leu Val  
 580 585 590  
 Ile Leu Leu Met Val Leu Ile Ala Ala Cys Arg Pro His Asn Pro Pro  
 595 600 605

Pro Phe Leu Asp Gly Ser Leu Asp Lys Pro Val Thr Tyr Ser Thr Pro  
 610 615 620

Lys Leu Val Ile Leu His Met Asn Met Ala Leu His Val Tyr Glu Asp  
 625 630 635 640  
 Ile Met Arg Met Thr Glu Asn Leu Ser Glu Lys Tyr Ile Ile Gly His  
 645 650 655

5 Gly Ala Ser Ser Thr Val Tyr Lys Cys Val Leu Lys Asn Cys Lys Pro  
 660 665 670

Val Ala Ile Lys Arg Leu Tyr Ser His Asn Pro Gln Ser Met Lys Gln  
 675 680 685

Phe Glu Thr Glu Leu Glu Met Leu Ser Ser Ile Lys His Arg Asn Leu  
 690 695 700

Val Ser Leu Gln Ala Tyr Ser Leu Ser His Leu Gly Ser Leu Leu Phe  
705 710 715 720

Tyr Asp Tyr Leu Glu Asn Gly Ser Leu Trp Asp Leu Leu His Gly Pro  
725 730 735

Thr Lys Lys Lys Thr Leu Asp Trp Asp Thr Arg Leu Lys Ile Ala Tyr  
740 745 750  
Gly Ala Ala Gln Gly Leu Ala Tyr Leu His His Asp Cys Ser Pro Arg  
755 760 765

Ile Ile His Arg Asp Val Lys Ser Ser Asn Ile Leu Leu Asp Lys Asp  
770 775 780

Leu Glu Ala Arg Leu Thr Asp Phe Gly Ile Ala Lys Ser Leu Cys Val  
785 790 795 800  
Ser Lys Ser His Thr Ser Thr Tyr Val Met Gly Thr Ile Gly Tyr Ile  
805 810 815

Asp Pro Glu Tyr Ala Arg Thr Ser Arg Leu Thr Glu Lys Ser Asp Val  
820 825 830  
Tyr Ser Tyr Gly Ile Val Leu Leu Glu Leu Leu Thr Arg Arg Lys Ala  
835 840 845

Val Asp Asp Glu Ser Asn Leu His His Leu Ile Met Ser Lys Thr Gly  
850 855 860

Asn Asn Glu Val Met Glu Met Ala Asp Pro Asp Ile Thr Ser Thr Cys  
865 870 875 880  
Lys Asp Leu Gly Val Val Lys Lys Val Phe Gln Leu Ala Leu Leu Cys  
885 890 895

Thr Lys Arg Gln Pro Asn Asp Arg Pro Thr Met His Gln Val Thr Arg  
900 905 910

Val Leu Gly Ser Phe Met Leu Ser Glu Gln Pro Pro Ala Ala Thr Asp  
915 920 925  
Thr Ser Ala Thr Leu Ala Gly Ser Cys Tyr Val Asp Glu Tyr Ala Asn  
930 935 940

Leu Lys Thr Pro His Ser Val Asn Cys Ser Ser Met Ser Ala Ser Asp  
945 950 955 960

Ala Gln Leu Phe Leu Arg Phe Gly Gln Val Ile Ser Gln Asn Ser Glu  
965 970 975

<210> 3  
<211> 3000  
<212> ADN  
<213> *ERECTA* de arroz

5

10

<400> 3

	atggcggcgg cgagggcgcc gtggctgtgg tggagggtgg tggagggtgt tgggtgtggc	60
	gtggcggagg cggcctccgg aggaggagga gggggagatg gggaggggaa ggcgctgatg	120
	ggcgtgaagg ccggtttcgg gaacgcggcc aacgcgctcg tcgactggga cggcggcgcc	180
	gaccactgcg cgtggcgcgg cgtcacctgc gacaacgcct ccttcgccgt cctcgccctg	240
5	aacttgtaa atctaaacct aggagggtgag atctcgccgg ccatcggaga gctcaagaat	300
	ctacagttcg ttgatctcaa ggggaacaag ctactggcc aaatcccaga tgagattggg	360
	gactgcatct ccttaaaata ttggatttg tctggcaact tgctgtatgg agacatcccc	420
	ttctccatct ccaagctcaa gcagcttgag gagctgattt tgaagaacaa ccagctcacg	480
	ggaccatcc cttccacatt gtcccaaatt ccaaactca agacattgga cctggcacag	540
	aaccagctta caggcgatg cccaaggctc atatactgga atgaagtct gcaataccta	600
	ggtttgaggg gtaactcact gactggaact ttgtcacctg acatgtgcca actgactggc	660
	ctgtggtact ttgatgtaag gggaaacaat ctcacagga ccattccaga gagcataggg	720
	aactgcacca gctttgagat tctggacatt tcgtataacc aaatctctgg agaaatacct	780
	tacaacatag gctttcttca agtagccaca ctgtcacttc aaggaaatag actgactggg	840
	aaaattccag atgtgattgg cctgatgcaa gctcttgctg ttctagacct gagtgagaac	900
	gagctggtag ggcccattcc ttctatactg ggcaatctat cctatactgg aaaactatat	960
	ttacatggga acaaaacttac tggagtcata ccgccggagc ttgggaacat gagtaaactt	1020
	agctacctac aactgaatga taatgaattg gtgggcacaa ttccagcaga gcttggcaaa	1080
	cttgaagagc tttttgaact aaatcttgcc aacaacaatc ttcaagggtcc tattcctgca	1140
	aacatcagtt cttgcactgc tctaaacaaa ttcaatgitt atggcaataa gctaaatggt	1200
	tctattcctg ctggtttcca gaagttggag agtctgactt acttgaacct atcttcaaac	1260
10	aatttcaaag gcaatattcc ttctgagctt ggtcacatca tcaacttga cacattggat	1320
	ctttctaca atgaattctc tggaccagtt cctgctacca ttggtgatct agagcacctt	1380
	cttgaactga atttgagtaa gaaccatctt gatgggccag ttctgctga gtttggaaac	1440
	ttgagaagcg tccaagtaat tgatatgtcc aacaacaact tatctggtag tctgcccag	1500
	gaacttgac aacttcaaaa ccttgatagc ctgattctta acaacaacaa tttggttggg	1560
	gagatccctg ctcaattggc caactgcttc agcttaaata accttgcatc tcaggaatctt	1620
	gtcatacaac aatttatctg gacatgtccc gatggcaaag aacttctcga aattccaat	1680
	ggaaagcatc ttctaatttc tgattgcaac cagtacataa atcataaatg cagcttcttg	1740
	ggtaatccat tactgcatgt ttactgocaa gattccagct gtggacactc tcatggacaa	1800
	agagttaata ttcaaagac agcaattgct tgcatatct taggctttat catattgctc	1860
	tgcyttctgc tgttggtat atataaaaca aatcaaccac agccacttgt caaaggatcc	1920
	gataagccag tgcaaggacc tccaagcta gttgttctcc agatggacat ggctatccat	1980
	acttacgagg acatcatgag gctgacagag aatttgagcg agaaatacat cattggctat	2040

ggcgcctcaa gcaactgtcta caaatgtgaa ctcaagagcg gcaaggccat tgctgtcaag 2100  
 cggcctttaca gtcagtataa ccatagcctc cgagagtttg aacagaact agagacaatt 2160  
 ggacagcatac ggcacaggaa tcttgttagc ctccatggct tctcgtatc tccacatgga 2220  
 aacttgctct tctatgatta catgaaaaat ggctccttgg gggatcttct ccacggcca 2280  
 tcaaagaaag tgaagctcaa ctgggacaca agactgagga tcgctgctcg agctgcacaa 2340  
 gggctggcct atctccacca tgactgcaac cctcgcataa tccacagaga tgtcaagtcc 2400  
 tccaacatcc tgctcgagca gaacttcgaa gcgcacctct cagatttcgg catagccaaa 2460  
 tgtgtcccct ctgccaagtc ccatgcctcc acttatgtgc taggaacct cggtacatt 2520  
 gatccggagt atgcaggac ttccaggctc aatgagaaat ctgatgtgta cagcttcggc 2580  
 atcgtccttc tggaattgct cacagggaaag aaggccgtcg acaacgaatc gaacttgcct 2640  
 caattgatac tctccaaagc tgatgacaac acagtcatgg aggcagtgga ctcggagggtg 2700  
 tcagtgcagt gcacggacat gggactggtc aggaaggcct tccagctcgc ccttctgtgc 2760  
 accaagaggc acccttcaga cgggccgacc atgcacgagg ttgcaagggt gctgctctcc 2820  
 ctgctgccgg cctccgccat gacaacgccc aagacggtgg actactcccg gttgctggcg 2880  
 tcgacgacga cggcggccga catgcgaggg cagcagctga ccgacatcgg cgacaacagc 2940  
 5 tcctccgacg agcagtggtt cgtcaggttc ggcgaggcca tatccaagca cacaatgtga 3000

<210> 4

<211> 999

<212> PRT

10 <213> *ERECTA* de arroz

<400> 4

Met Ala Ala Ala Arg Ala Pro Trp Leu Trp Trp Trp Val Val Val Val

1 5 10 15

Val Gly Val Ala Val Ala Glu Ala Ala Ser Gly Gly Gly Gly Gly Gly

20 25 30

Asp Gly Glu Gly Lys Ala Leu Met Gly Val Lys Ala Gly Phe Gly Asn

35 40 45

Ala Ala Asn Ala Leu Val Asp Trp Asp Gly Gly Ala Asp His Cys Ala

50 55 60

Trp Arg Gly Val Thr Cys Asp Asn Ala Ser Phe Ala Val Leu Ala Leu

65 70 75 80

Asn Leu Ser Asn Leu Asn Leu Gly Gly Glu Ile Ser Pro Ala Ile Gly

85 90 95

15

Glu Leu Lys Asn Leu Gln Phe Val Asp Leu Lys Gly Asn Lys Leu Thr  
 100 105 110

Gly Gln Ile Pro Asp Glu Ile Gly Asp Cys Ile Ser Leu Lys Tyr Leu  
 115 120 125

Asp Leu Ser Gly Asn Leu Leu Tyr Gly Asp Ile Pro Phe Ser Ile Ser  
 130 135 140  
 Lys Leu Lys Gln Leu Glu Glu Leu Ile Leu Lys Asn Asn Gln Leu Thr  
 145 150 155 160

Gly Pro Ile Pro Ser Thr Leu Ser Gln Ile Pro Asn Leu Lys Thr Leu  
 165 170 175

Asp Leu Ala Gln Asn Gln Leu Thr Gly Asp Ile Pro Arg Leu Ile Tyr  
 180 185 190  
 Trp Asn Glu Val Leu Gln Tyr Leu Gly Leu Arg Gly Asn Ser Leu Thr  
 195 200 205

Gly Thr Leu Ser Pro Asp Met Cys Gln Leu Thr Gly Leu Trp Tyr Phe  
 210 215 220

Asp Val Arg Gly Asn Asn Leu Thr Gly Thr Ile Pro Glu Ser Ile Gly  
 225 230 235 240  
 Asn Cys Thr Ser Phe Glu Ile Leu Asp Ile Ser Tyr Asn Gln Ile Ser  
 245 250 255

Gly Glu Ile Pro Tyr Asn Ile Gly Phe Leu Gln Val Ala Thr Leu Ser  
 260 265 270

Leu Gln Gly Asn Arg Leu Thr Gly Lys Ile Pro Asp Val Ile Gly Leu  
 275 280 285

Met Gln Ala Leu Ala Val Leu Asp Leu Ser Glu Asn Glu Leu Val Gly  
 290 295 300  
 Pro Ile Pro Ser Ile Leu Gly Asn Leu Ser Tyr Thr Gly Lys Leu Tyr  
 305 310 315 320

Leu His Gly Asn Lys Leu Thr Gly Val Ile Pro Pro Glu Leu Gly Asn  
 325 330 335

Met Ser Lys Leu Ser Tyr Leu Gln Leu Asn Asp Asn Glu Leu Val Gly  
 340 345 350

5

Thr Ile Pro Ala Glu Leu Gly Lys Leu Glu Glu Leu Phe Glu Leu Asn  
 355 360 365

Leu Ala Asn Asn Asn Leu Gln Gly Pro Ile Pro Ala Asn Ile Ser Ser  
 370 375 380

Cys Thr Ala Leu Asn Lys Phe Asn Val Tyr Gly Asn Lys Leu Asn Gly  
 385 390 395 400  
 Ser Ile Pro Ala Gly Phe Gln Lys Leu Glu Ser Leu Thr Tyr Leu Asn  
 405 410 415

Leu Ser Ser Asn Asn Phe Lys Gly Asn Ile Pro Ser Glu Leu Gly His  
 420 425 430  
 Ile Ile Asn Leu Asp Thr Leu Asp Leu Ser Tyr Asn Glu Phe Ser Gly  
 435 440 445

Pro Val Pro Ala Thr Ile Gly Asp Leu Glu His Leu Leu Glu Leu Asn  
 450 455 460

Leu Ser Lys Asn His Leu Asp Gly Pro Val Pro Ala Glu Phe Gly Asn  
 465 470 475 480  
 Leu Arg Ser Val Gln Val Ile Asp Met Ser Asn Asn Asn Leu Ser Gly  
 485 490 495

Ser Leu Pro Glu Glu Leu Gly Gln Leu Gln Asn Leu Asp Ser Leu Ile  
 500 505 510

Leu Asn Asn Asn Asn Leu Val Gly Glu Ile Pro Ala Gln Leu Ala Asn  
 515 520 525  
 Cys Phe Ser Leu Asn Asn Leu Ala Phe Gln Glu Phe Val Ile Gln Gln  
 530 535 540

Phe Ile Trp Thr Cys Pro Asp Gly Lys Glu Leu Leu Glu Ile Pro Asn  
 545 550 555 560

Gly Lys His Leu Leu Ile Ser Asp Cys Asn Gln Tyr Ile Asn His Lys  
 565 570 575  
 Cys Ser Phe Leu Gly Asn Pro Leu Leu His Val Tyr Cys Gln Asp Ser  
 580 585 590

Ser Cys Gly His Ser His Gly Gln Arg Val Asn Ile Ser Lys Thr Ala  
 595 600 605

5

Ile Ala Cys Ile Ile Leu Gly Phe Ile Ile Leu Leu Cys Val Leu Leu  
 610 615 620

Leu Ala Ile Tyr Lys Thr Asn Gln Pro Gln Pro Leu Val Lys Gly Ser  
 625 630 635 640

Asp Lys Pro Val Gln Gly Pro Pro Lys Leu Val Val Leu Gln Met Asp  
 645 650 655  
 Met Ala Ile His Thr Tyr Glu Asp Ile Met Arg Leu Thr Glu Asn Leu  
 660 665 670

Ser Glu Lys Tyr Ile Ile Gly Tyr Gly Ala Ser Ser Thr Val Tyr Lys  
 675 680 685

Cys Glu Leu Lys Ser Gly Lys Ala Ile Ala Val Lys Arg Leu Tyr Ser  
 690 695 700  
 Gln Tyr Asn His Ser Leu Arg Glu Phe Glu Thr Glu Leu Glu Thr Ile  
 705 710 715 720

Gly Ser Ile Arg His Arg Asn Leu Val Ser Leu His Gly Phe Ser Leu  
 725 730 735

Ser Pro His Gly Asn Leu Leu Phe Tyr Asp Tyr Met Glu Asn Gly Ser  
 740 745 750  
 Leu Trp Asp Leu Leu His Gly Pro Ser Lys Lys Val Lys Leu Asn Trp  
 755 760 765

Asp Thr Arg Leu Arg Ile Ala Val Gly Ala Ala Gln Gly Leu Ala Tyr  
 770 775 780  
 Leu His His Asp Cys Asn Pro Arg Ile Ile His Arg Asp Val Lys Ser  
 785 790 795 800

Ser Asn Ile Leu Leu Asp Glu Asn Phe Glu Ala His Leu Ser Asp Phe  
 805 810 815

Gly Ile Ala Lys Cys Val Pro Ser Ala Lys Ser His Ala Ser Thr Tyr  
 820 825 830  
 Val Leu Gly Thr Ile Gly Tyr Ile Asp Pro Glu Tyr Ala Arg Thr Ser  
 835 840 845

Arg Leu Asn Glu Lys Ser Asp Val Tyr Ser Phe Gly Ile Val Leu Leu  
 850 855 860

Glu Leu Leu Thr Gly Lys Lys Ala Val Asp Asn Glu Ser Asn Leu His  
 865 870 875 880

5



Gln Leu Ile Leu Ser Lys Ala Asp Asp Asn Thr Val Met Glu Ala Val  
 885 890 895

Asp Ser Glu Val Ser Val Thr Cys Thr Asp Met Gly Leu Val Arg Lys  
 900 905 910

Ala Phe Gln Leu Ala Leu Leu Cys Thr Lys Arg His Pro Ser Asp Arg  
 915 920 925  
 Pro Thr Met His Glu Val Ala Arg Val Leu Leu Ser Leu Leu Pro Ala  
 930 935 940

Ser Ala Met Thr Thr Pro Lys Thr Val Asp Tyr Ser Arg Leu Leu Ala  
 945 950 955 960

Ser Thr Thr Thr Ala Ala Asp Met Arg Gly His Asp Val Thr Asp Ile  
 965 970 975  
 Gly Asp Asn Ser Ser Ser Asp Glu Gln Trp Phe Val Arg Phe Gly Glu  
 980 985 990

Val Ile Ser Lys His Thr Met  
 995

5

<210> 5  
 <211> 2766  
 <212> ADN  
 <213> *ERECTA* de sorgo

10

<400> 5

atgacgacga cggcgcgccg tgctctcgtc gccctcctcc tcgtcgccgt cgccgtcgcc 60  
 gacgatgggg cgacgtggt ggagatcaag aagtccttcc gcaacgtcgg caacgtactg 120  
 tacgattggg cggcgacga ctactgtctc tggcgcggcg tcctgtgcga caacgtcaca 180  
 ttcgccgtcg ctgcgtcaa cctctctggc ctcaaccttg agggcgagat ctctccagcc 240  
 gtcggcagcc tcaagagcct cgtctccatc gatctgaagt caaatgggct atccgggcag 300  
 atccctgatg agattggtga ttgttcatca cttaggacgc tggacttttc tttcaacaac 360  
 ttggatggcg acataccatt ttctatatca aagctgaagc acctggagaa cttgatattg 420  
 aagaacaacc agctgattgg tgcgatccca tcaacattgt cacagctccc aaatttgaag 480  
 attttgatt tggcacaana caaactgact ggggagatc caaggcttat ctactggaat 540  
 gaggttcttc aatatcttga tgtgaagaac aatagcttga ccggggtgat accagacacc 600  
 attgggaact gtacaagttt tcaagtcttg gatttgtctt acaaccgctt tactggacca 660

15

# ES 2 362 903 T3

	atccattca acattggttt cctacaagtg gctacactat ccttgcaagg gaacaagttc	720
	accggtccaa ttccttcagt aattggcttt atgcaggctc tcgctgttct agatctgagt	780
	tacaaccaat tatctggctc tataccatca atactaggca acttgacata cactgagaag	840
	ctgtacatcc aaggcaataa gttaactggg tcgataccac cagagttagg aaatatgtca	900
	acacttcatt acctagaact gaacgataat caacttactg ggtcaattcc accagagctt	960
	ggaaggctaa caggcttgtt tgacctgaac cttgcgaata accacctgga aggaccaatt	1020
	cctgacaacc taagtctatg tgtgaatctc aatagcttca atgcttatgg caacaagtta	1080
	aatgggacca ttcctcgttc gttgcgaaa cttgaaagca tgacctattt aaatctgtca	1140
	tcaaaactca taagtggctc tattcctatt gagttatcaa ggatcaacaa ttggacacg	1200
	ctggatttat cctgtaacat gatgactggc ccaattccat catcaattgg cagcctagag	1260
	catctattga gacttaactt gagcaagaat ggtctagtgg gattcatccc cgcgagttt	1320
	ggtaatttga ggagtgtcat ggagattgat ttatcctata atcaccttgg tggcctgatt	1380
	cctcaagaac ttgaaatgct gcaaaacctg atgttgctaa atgtgtcgtc caataatttg	1440
	gctgggttgg tccctgctga caacaacttc acacggtttt cacctgacag ctttttaggt	1500
	aatcctggac tctgtggata ctggcttggc tcgtcgtgtc gttccactgg ccaccacgag	1560
	aaaccgcta tctcaaaggc tgccataatt ggtgttgctg tgggtggact tgttatcctc	1620
	ttgatgatct tagtagctgt ttgcaggcca catcgtccac ctgcttttaa agatgtcact	1680
	gtaagcaagc cagtgagaaa tgctcccccc aagctgggta tccttcatat gaacatggcc	1740
	cttcatgtat acgatgacat aatgaggatg actgagaact tgagtgagaa atacatcatt	1800
5	ggatacgggg cgcaagtac agtttataaa tgtgtcctaa agaattgcaa accggtggca	1860
	ataaaaaagc tgtatgcca ctaccacag agccttaagg aatttgaac tgagcttgag	1920
	actggttgta gcatcaagca ccggaatcta gtcagccttc aagggtactc attatcacct	1980
	gttgggaacc tcctctttta tgattatag gaatgtggca gcttatggga tgttttacat	2040
	gaaggttcat ccaagaagaa aaaacttgac tgggagactc gcctacggat tgctcttggc	2100
	gcagctcaag gccttgctta ccttcacat gactgcagtc cacggataat tcatcgggat	2160
	gtaaaatcaa agaataact ccttgacaaa gattatgagg cccatcttac agactttgga	2220
	attgctaaga gcttatgtgt ctcaaaaact cacacatcaa cctatgtcat gggaactatt	2280
	ggctacattg atcctgagta cgcccgcact tcccgtctca acgaaaagtc tgatgtctac	2340
	aggctatggc attgttctgc tggagctgct gactggcaag aagccagtgg acaacgaatc	2400
	ctatcgaaga cggaagcaa cgaggtcatg gataccgtgg accctgacat cggggacacc	2460

# ES 2 362 903 T3

```

tgcaaggacc tcggcgaggt gaagaagctg ttccagctgg cgctcctttg caccaagcgg 2520
caaccctcgg accgaccgac gatgcacgag gtggtgcgcg tcttgactg cctggtgaac 2580
cgggacccgc cgccaaagcc gtcggcgcac cagctgccc agccgtccc agccgtgcca 2640
agctacatca acgagtacgt cagcctgcgg ggcaccggcg ctctctctg cgccaactcg 2700
accagcacct cggacgccga gctgttcctc aagttcggcg aggccatctc gcagaacatg 2760

gagtag 2766

```

5 <210> 6  
 <211> 921  
 <212> PRT  
 <213> *ERECTA* de sorgo

10 <400> 6

```

Met Thr Thr Thr Ala Ala Arg Ala Leu Val Ala Leu Leu Leu Val Ala
1           5           10           15

Val Ala Val Ala Asp Asp Gly Ala Thr Leu Val Glu Ile Lys Lys Ser
                20           25           30
Phe Arg Asn Val Gly Asn Val Leu Tyr Asp Trp Ala Gly Asp Asp Tyr
          35           40           45

Cys Ser Trp Arg Gly Val Leu Cys Asp Asn Val Thr Phe Ala Val Ala
          50           55           60

Ala Leu Asn Leu Ser Gly Leu Asn Leu Glu Gly Glu Ile Ser Pro Ala
65           70           75           80
Val Gly Ser Leu Lys Ser Leu Val Ser Ile Asp Leu Lys Ser Asn Gly
          85           90           95

Leu Ser Gly Gln Ile Pro Asp Glu Ile Gly Asp Cys Ser Ser Leu Arg
          100          105          110

Thr Leu Asp Phe Ser Phe Asn Asn Leu Asp Gly Asp Ile Pro Phe Ser
          115          120          125
Ile Ser Lys Leu Lys His Leu Glu Asn Leu Ile Leu Lys Asn Asn Gln
          130          135          140

Leu Ile Gly Ala Ile Pro Ser Thr Leu Ser Gln Leu Pro Asn Leu Lys
145           150           155           160

Ile Leu Asp Leu Ala Gln Asn Lys Leu Thr Gly Glu Ile Pro Arg Leu
          165          170          175
Ile Tyr Trp Asn Glu Val Leu Gln Tyr Leu Asp Val Lys Asn Asn Ser
          180          185          190

15 Leu Thr Gly Val Ile Pro Asp Thr Ile Gly Asn Cys Thr Ser Phe Gln

```



Glu Met Leu Gln Asn Leu Met Leu Leu Asn Val Ser Tyr Asn Asn Leu  
 465 470 475 480

Ala Gly Val Val Pro Ala Asp Asn Asn Phe Thr Arg Phe Ser Pro Asp  
 485 490 495

Ser Phe Leu Gly Asn Pro Gly Leu Cys Gly Tyr Trp Leu Gly Ser Ser  
 500 505 510  
 Cys Arg Ser Thr Gly His His Glu Lys Pro Pro Ile Ser Lys Ala Ala  
 515 520 525

Ile Ile Gly Val Ala Val Gly Gly Leu Val Ile Leu Leu Met Ile Leu  
 530 535 540  
 Val Ala Val Cys Arg Pro His Arg Pro Pro Ala Phe Lys Asp Val Thr  
 545 550 555 560

Val Ser Lys Pro Val Arg Asn Ala Pro Pro Lys Leu Val Ile Leu His  
 565 570 575

Met Asn Met Ala Leu His Val Tyr Asp Asp Ile Met Arg Met Thr Glu  
 580 585 590  
 Asn Leu Ser Glu Lys Tyr Ile Ile Gly Tyr Gly Ala Ser Ser Thr Val  
 595 600 605

Tyr Lys Cys Val Leu Lys Asn Cys Lys Pro Val Ala Ile Lys Lys Leu  
 610 615 620

Tyr Ala His Tyr Pro Gln Ser Leu Lys Glu Phe Glu Thr Glu Leu Glu  
 625 630 635 640  
 Thr Val Gly Ser Ile Lys His Arg Asn Leu Val Ser Leu Gln Gly Tyr  
 645 650 655

Ser Leu Ser Pro Val Gly Asn Leu Leu Phe Tyr Asp Tyr Met Glu Cys  
 660 665 670

Gly Ser Leu Trp Asp Val Leu His Glu Gly Ser Ser Lys Lys Lys Lys  
 675 680 685  
 Leu Asp Trp Glu Thr Arg Leu Arg Ile Ala Leu Gly Ala Ala Gln Gly  
 690 695 700

Leu Ala Tyr Leu His His Asp Cys Ser Pro Arg Ile Ile His Arg Asp  
 705 710 715 720

5



# ES 2 362 903 T3

gagattggaa actgtgttcc tcttgcttat gtggatttct ccaccaatth gttgtttgga 300  
 gacataccgt tttcaatctc taaactcaaa cagctgacct taactcagat tccaaacctt 360  
 aagacccttg acctcgcaag aaaccagctt actggtgaga taccaaggtt actctactgg 420  
 aatgaagttt tacagtatct cggtttacgt ggaatatgt taactgggac attgtctcct 480  
 gatattgtgc agctgacggg tctgtggtac tttgatgtga gaggcaacaa ccttactgga 540  
 actatcccag agagcattgg caattgcaca agctttgaga tcttgatgt atcttataat 600  
 cagattaccg gagttatacc ctacaatatt ggttccctcc aagtagctac tctgtcactt 660  
 caagaaaca agttgactgg cagaattccg gaagtgattg gtctgatgca ggctcttgct 720  
 gtattggatt tgagtgacaa tgaattaact gggcctattc caccaatact tgggaatctg 780  
 tcattcactg gaaaactgta tctccatggc aacaagctca ctggacaaat cccaccggag 840  
 ctaggcaata tgtcacgact cagctatttg caactaaatg ataatgaact agtgggaaag 900  
 atcccacctg agcttgggaa gctggaacaa ttgttcgaac tgaatcttgc gaacaacaat 960  
 cttgtagggc tgattccatc taacattagt tcctgtgctg ccttgaatca attcaatgtt 1020  
 catgggaact tcttgatggg agctgtacca cttgaattcc ggaatcttgg aagcttgact 1080  
 tatctaaatc tttcctcaaa cagtttcaag ggcaaaatac ctgctgagct tggccatatac 1140  
 atcaatcttg atacattgga tctgtctggc aacaatttct caggctcaat tccattaaca 1200  
 5 cttggtgatc ttgagcatct tctcatctta aacttgagca gaaatcatct gaatggcaca 1260  
 ttgcctgcag aattcgggaa cctccgaagc attcagatca tcgatgtgct atttaatttt 1320  
 cttgccggtg ttattccaac tgaacttggc cagttgcaga acataaactc tctgatactg 1380  
 aacaacaaca agattcatgg gaaaatccct gatcagctaa ctaactgctt cagtcttggc 1440  
 aatctgaaca tctccttcaa taatctttct ggaataatcc cacctatgaa gaactttaca 1500  
 cgtttttccc cggccagctt ctttggaat ccatttctct gcgggaactg ggttgatca 1560  
 atctgtggcc catctttacc taagtcacaa gtattcacca gagttgccgt gatttgtatg 1620  
 gttctcgggt tcactactct catatgcatg atattcattg cggtttaca gtcaaagcag 1680  
 cagaaaccag tcttgaaagg ctcttcaaaa caacctgaag ggtcaacgaa gctggtgatt 1740  
 cttcacatgg acatggctat tcacacgttt gatgatata tgagagttac agaaaacctc 1800  
 gatgagaaat acatcattgg atacggtgct tctagcacag tttacaagtg cacctccaaa 1860  
 actcccgcac ctattgcat taagcgaatc tacaatcagt atcccagcaa ctcccgcgag 1920  
 tttgaaacag agctcgagac cattgggagc atcagacaca gaaacatagt aagcttgcac 1980  
 ggatacgcct tatctccctt tggcaacctc ctctctacg actacatgga aaatggctct 2040  
 ctttgggatc tctccatgg gcctgggaag aaggtgaagc ttgactggga aacaaggctg 2100

# ES 2 362 903 T3

```

aagatagctg ttggagctgc gcaaggactt gcatatcttc accatgactg cacacctagg 2160
ataatccatc gagacatcaa gtcacaaac atactccttg atgggaattt cgaagcgcgt 2220
ttgtcagatt ttgggattgc caagagcata ccagccacca aaacttatgc ttcaacctat 2280
gttcttgaa ccattggata tattgacca gaggatgctc gaacttcgcg tctgaacgag 2340
aagcttgata tctacagttt cggattgtc cttcttgagc ttctaaccgg caagaaggct 2400
gtggataacg aggccaaactt gcatcaaatg attctatcaa aggcggatga taacacagta 2460
atggaagctg ttgatgcaga ggtctcagtg acttgcctgg actcaggaca catcaagaaa 2520
acatttcagc tagctctctt gtgcaccaag cgaaatcctt tggagagacc caccatgcag 2580
gaggtctcta gggttctgct ctcaactgtc ccgtctccac ctccaagaa gttaccgtcg 2640
cctgcaaaag tacaggaagg ggaagaacgg cgtgagagcc actcttcaga tacaacaacc 2700
ccacagtgtt ttgttcagtt ccgtgaagat atctccaaa gtagcttata a 2751

```

5 <210> 8  
 <211> 932  
 <212> PRT  
 <213> Homólogo *ERECTA* de *Arabidopsis thaliana*

10 <400> 8

```

Met Ala Ile Lys Ala Ser Phe Ser Asn Val Ala Asn Met Leu Leu Asp
1           5           10           15

Trp Asp Asp Val His Asn His Asp Phe Cys Ser Trp Arg Gly Val Phe
           20           25           30

Cys Asp Asn Val Ser Leu Asn Val Val Ser Leu Asn Leu Ser Asn Leu
           35           40           45
Asn Leu Gly Gly Glu Ile Ser Ser Ala Leu Gly Asp Leu Met Asn Leu
           50           55           60

Gln Ser Ile Asp Leu Gln Gly Asn Lys Leu Gly Gly Gln Ile Pro Asp
65           70           75           80

Glu Ile Gly Asn Cys Val Ser Leu Ala Tyr Val Asp Phe Ser Thr Asn
           85           90           95
Leu Leu Phe Gly Asp Ile Pro Phe Ser Ile Ser Lys Leu Lys Gln Leu
           100          105          110

Glu Phe Leu Asn Leu Lys Asn Asn Gln Leu Thr Gly Pro Ile Pro Ala
           115          120          125

Thr Leu Thr Gln Ile Pro Asn Leu Lys Thr Leu Asp Leu Ala Arg Asr
           130          135          140

```



Gln Leu Thr Gly Glu Ile Pro Arg Leu Leu Tyr Trp Asn Glu Val Leu  
 145                                    150                                    155                                    160

Gln Tyr Leu Gly Leu Arg Gly Asn Met Leu Thr Gly Thr Leu Ser Pro  
    165                                    170                                    175

Asp Met Cys Gln Leu Thr Gly Leu Trp Tyr Phe Asp Val Arg Gly Asn  
    180                                    185                                    190  
 Asn Leu Thr Gly Thr Ile Pro Glu Ser Ile Gly Asn Cys Thr Ser Phe  
    195                                    200                                    205

Glu Ile Leu Asp Val Ser Tyr Asn Gln Ile Thr Gly Val Ile Pro Tyr  
    210                                    215                                    220  
 Asn Ile Gly Phe Leu Gln Val Ala Thr Leu Ser Leu Gln Gly Asn Lys  
    225                                    230                                    235                                    240

Leu Thr Gly Arg Ile Pro Glu Val Ile Gly Leu Met Gln Ala Leu Ala  
    245                                    250                                    255

Val Leu Asp Leu Ser Asp Asn Glu Leu Thr Gly Pro Ile Pro Pro Ile  
    260                                    265                                    270  
 Leu Gly Asn Leu Ser Phe Thr Gly Lys Leu Tyr Leu His Gly Asn Lys  
    275                                    280                                    285

Leu Thr Gly Gln Ile Pro Pro Glu Leu Gly Asn Met Ser Arg Leu Ser  
    290                                    295                                    300

Tyr Leu Gln Leu Asn Asp Asn Glu Leu Val Gly Lys Ile Pro Pro Glu  
    305                                    310                                    315                                    320  
 Leu Gly Lys Leu Glu Gln Leu Phe Glu Leu Asn Leu Ala Asn Asn Asn  
    325                                    330                                    335

Leu Val Gly Leu Ile Pro Ser Asn Ile Ser Ser Cys Ala Ala Leu Asn  
    340                                    345                                    350

Gln Phe Asn Val His Gly Asn Phe Leu Ser Gly Ala Val Pro Leu Glu  
    355                                    360                                    365  
 Phe Arg Asn Leu Gly Ser Leu Thr Tyr Leu Asn Leu Ser Ser Asn Ser  
    370                                    375                                    380

Phe Lys Gly Lys Ile Pro Ala Glu Leu Gly His Ile Ile Asn Leu Asp  
    385                                    390                                    395                                    400

5

Thr Leu Asp Leu Ser Gly Asn Asn Phe Ser Gly Ser Ile Pro Leu Thr  
 405 410 415

Leu Gly Asp Leu Glu His Leu Leu Ile Leu Asn Leu Ser Arg Asn His  
 420 425 430

Leu Asn Gly Thr Leu Pro Ala Glu Phe Gly Asn Leu Arg Ser Ile Gln  
 435 440 445  
 Ile Ile Asp Val Ser Phe Asn Phe Leu Ala Gly Val Ile Pro Thr Glu  
 450 455 460

Leu Gly Gln Leu Gln Asn Ile Asn Ser Leu Ile Leu Asn Asn Asn Lys  
 465 470 475 480

Ile His Gly Lys Ile Pro Asp Gln Leu Thr Asn Cys Phe Ser Leu Ala  
 485 490 495  
 Asn Leu Asn Ile Ser Phe Asn Asn Leu Ser Gly Ile Ile Pro Pro Met  
 500 505 510

Lys Asn Phe Thr Arg Phe Ser Pro Ala Ser Phe Phe Gly Asn Pro Phe  
 515 520 525

Leu Cys Gly Asn Trp Val Gly Ser Ile Cys Gly Pro Ser Leu Pro Lys  
 530 535 540  
 Ser Gln Val Phe Thr Arg Val Ala Val Ile Cys Met Val Leu Gly Phe  
 545 550 555 560

Ile Thr Leu Ile Cys Met Ile Phe Ile Ala Val Tyr Lys Ser Lys Gln  
 565 570 575

Gln Lys Pro Val Leu Lys Gly Ser Ser Lys Gln Pro Glu Gly Ser Thr  
 580 585 590

Lys Leu Val Ile Leu His Met Asp Met Ala Ile His Thr Phe Asp Asp  
 595 600 605  
 Ile Met Arg Val Thr Glu Asn Leu Asp Glu Lys Tyr Ile Ile Gly Tyr  
 610 615 620

Gly Ala Ser Ser Thr Val Tyr Lys Cys Thr Ser Lys Thr Ser Arg Pro  
 625 630 635 640

Ile Ala Ile Lys Arg Ile Tyr Asn Gln Tyr Pro Ser Asn Phe Arg Glu  
 645 650 655

5

Phe Glu Thr Glu Leu Glu Thr Ile Gly Ser Ile Arg His Arg Asn Ile  
 660 665 670

Val Ser Leu His Gly Tyr Ala Leu Ser Pro Phe Gly Asn Leu Leu Phe  
 675 680 685

Tyr Asp Tyr Met Glu Asn Gly Ser Leu Trp Asp Leu Leu His Gly Pro  
 690 695 700  
 Gly Lys Lys Val Lys Leu Asp Trp Glu Thr Arg Leu Lys Ile Ala Val  
 705 710 715 720

Gly Ala Ala Gln Gly Leu Ala Tyr Leu His His Asp Cys Thr Pro Arg  
 725 730 735  
 Ile Ile His Arg Asp Ile Lys Ser Ser Asn Ile Leu Leu Asp Gly Asn  
 740 745 750

Phe Glu Ala Arg Leu Ser Asp Phe Gly Ile Ala Lys Ser Ile Pro Ala  
 755 760 765  
 Thr Lys Thr Tyr Ala Ser Thr Tyr Val Leu Gly Thr Ile Gly Tyr Ile  
 770 775 780

Asp Pro Glu Tyr Ala Arg Thr Ser Arg Leu Asn Glu Lys Ser Asp Ile  
 785 790 795 800

Tyr Ser Phe Gly Ile Val Leu Leu Glu Leu Leu Thr Gly Lys Lys Ala  
 805 810 815  
 Val Asp Asn Glu Ala Asn Leu His Gln Met Ile Leu Ser Lys Ala Asp  
 820 825 830

Asp Asn Thr Val Met Glu Ala Val Asp Ala Glu Val Ser Val Thr Cys  
 835 840 845

Met Asp Ser Gly His Ile Lys Lys Thr Phe Gln Leu Ala Leu Leu Cys  
 850 855 860  
 Thr Lys Arg Asn Pro Leu Glu Arg Pro Thr Met Gln Glu Val Ser Arg  
 865 870 875 880

Val Leu Leu Ser Leu Val Pro Ser Pro Pro Pro Lys Lys Leu Pro Ser  
 885 890 895

Pro Ala Lys Val Gln Glu Gly Glu Glu Arg Arg Glu Ser His Ser Ser  
 900 905 910  
 Asp Thr Thr Pro Gln Trp Phe Val Gln Phe Arg Glu Asp Ile Ser  
 915 920 925

Lys Ser Ser Leu  
 930

<210> 9

5

<211> 2901  
 <212> ADN  
 <213> Homólogo *ERECTA* de *Arabidopsis thaliana*

5 <400> 9

atgaaggaga agatgcagcg aatgggttta tctttagcaa tggtagggtt tatgggtttt 60  
 ggtggtgctt cggctatgaa caacgaaggg aaagctctga tggcgataaa aggctctttc 120  
 agcaacttag tgaatatgct tttggattgg gacgatgttc acaacagtga cttgtgttct 180  
 tggcgagggtg ttttctgcga caacgttagc tactccgttg tctctctgaa tttgtccagt 240  
 ctgaatcttg gaggggagat atctccagct attggagacc tacggaattt gcaatcaata 300  
 gacttgcaag gtaataaact agcagggtcaa attccagatg agattggaaa ctgtgcttct 360  
 cttgtttatc tggattgtc cgagaatctg ttatatggag acataccttt ctcaatctct 420  
 aaactcaagc agcttgaac tctgaatctg aagaacaatc agctcacagg tcctgtacca 480  
 gcaaccttaa cccagattcc aaaccttaag agacttgatc ttgctggcaa tcatctaacg 540  
 ggtgagatat cgagattgct ttactggaat gaagttttgc agtatcttgg attacgaggg 600  
 aatattgttg ctggaacgtt atcttctgat atgtgtcagc taaccgggtt gtggtacttt 660  
 gatgtgagag gaaataatct aactggaacc atcccggaga gcatcggaaa ttgcacaagc 720  
 tttcaaatcc tggacatata ttataatcag ataacaggag agattcctta caatatcggc 780

10

ttcctccaag ttgctactct gtcacttcaa ggaacacagat tgacgggtag aattccagaa 840  
 gttattggtc taatgcagcg tcttctgttt ttggatttga gtgacaatga gcttgttggg 900  
 cctatcccac cgatacttgg caatctctca tttaccggaa agttgtatct ccatggcaat 960  
 atgctcactg gtccaatccc ctctgagctt gggaaatgtt cacgtctcag ctatttgag 1020  
 ctaaacgaca ataaactagt gggaaactatt ccacctgagc ttggaaagct ggagcaattg 1080  
 tttgaaactga atcttgccaa caaccgttta gtagggccca taccatccaa cattagtcca 1140  
 tgtgcagcct tgaatcaatt caatgttcat gggaaacctct tgagtggatc tattccactg 1200  
 gcgtttcgca atctcgggag cttgacttat ctgaatcttt cgtcgaacaa tttcaaggga 1260  
 aaaataccag ttgagcttgg acatataatc aatcttgaca aactagatct gtctggcaat 1320  
 aacttctcag ggtctatacc attaacgctt ggcgatcttg aacaccttct catattaat 1380  
 cttagcagaa accatcttag tggacaatta cctgcagagt ttgggaacct tcgaagcatt 1440  
 cagatgattg atgtatcatt caatctgctc tccggagtta ttccaactga acttggccaa 1500  
 ttgcagaatt taaactcttt aatattgaac aacaacaagc ttcattggaa aattccagat 1560  
 cagcttacga actgcttcac tcttgtcaat ctgaatgtct ccttcaacaa tctctcggg 1620  
 atagtccac caatgaaaaa cttctcagct tttgtccag ccagctttgt tggaaatcca 1680  
 tatctttgtg gaaactgggt tggatctatt tgtggtcctt taccgaaatc tcgagtattc 1740  
 tccagagggtg ctttgcactg cattgttctt ggcgtcatca ctctcctatg tatgattttc 1800

cttgcagttt acaaatcaat gcagcagaag aagattctac aaggctcctc aaaacaagct 1860  
 gaaggggtaa ccaactagat gattctccac atggacatgg caattcatac atttgatgat 1920  
 atcatgagag tgactgagaa tcttaacgaa aagtttataa ttggatatgg tgcttctagc 1980  
 acggtataca aatgtgcatt aaaaagttcc cgacctattg ccattaagcg actctacaat 2040  
 cagtatccgc ataacttgcg ggaatttgag acagaacttg agaccattgg gagcattagg 2100  
 cacagaaaca tagtcagctt gcatggatat gccttgctc ctactggcaa ccttcttttc 2160  
 tatgactaca tggaaaatgg atcactttgg gaccttctc atgggtcatt gaagaaagtg 2220  
 aagcttgatt gggagacaag gttgaagata gcggttgag ctgcacaagg actagcctat 2280  
 cttcaccacg attgtactcc tcgaatcatt caccgtgaca tcaagtcac gaacatactt 2340  
 cttgatgaga atttogaagc acatttatct gatttcggga ttgctaagag cataccagct 2400  
 agcaaaacc atgcctcgac ttatgtttg ggaacaattg gttatataga cccagagtat 2460  
 gctcgtactt cacgaatcaa tgagaaatcc gatataata gcttcggtat tgttcttctt 2520  
 gagcttctca ctgggaagaa agcagtgat aacgaagcta acttgcata actgatattg 2580  
 tcaaaggctg atgataatac tgtgatggaa gcagttgatc cagaggttac tgtgacttgt 2640  
 atggacttgg gacatatcag gaagacattt cagctggctc tcttatgcac aaagcgaaac 2700  
 ccttagaga gaccacaat gcttgaagtc tctagggttc tgctctctct tgtccatct 2760  
 5 ctgcaagtag caaagaagct accttctctt gatcactcaa ccaaaaagct gcagcaagag 2820  
 aatgaagtta ggaatcctga tgcagaagca tctcaatggt ttgttcagtt cctgaagtc 2880  
 atctcaaaa gtagcatata a 2901

<210> 10

<211> 966

10 <212> PRT

<213> Homólogo *ERECTA* de *Arabidopsis thaliana*

<400> 10

Met Lys Glu Lys Met Gln Arg Met Val Leu Ser Leu Ala Met Val Gly  
 1 5 10 15

15

Phe Met Val Phe Gly Val Ala Ser Ala Met Asn Asn Glu Gly Lys Ala  
 20 25 30

Leu Met Ala Ile Lys Gly Ser Phe Ser Asn Leu Val Asn Met Leu Leu  
 35 40 45

Asp Trp Asp Asp Val His Asn Ser Asp Leu Cys Ser Trp Arg Gly Val  
 50 55 60

Phe Cys Asp Asn Val Ser Tyr Ser Val Val Ser Leu Asn Leu Ser Ser  
 65 70 75 80

Leu Asn Leu Gly Gly Glu Ile Ser Pro Ala Ile Gly Asp Leu Arg Asn  
 85 90 95

Leu Gln Ser Ile Asp Leu Gln Gly Asn Lys Leu Ala Gly Gln Ile Pro  
 100 105 110

Asp Glu Ile Gly Asn Cys Ala Ser Leu Val Tyr Leu Asp Leu Ser Glu  
 115 120 125  
 Asn Leu Leu Tyr Gly Asp Ile Pro Phe Ser Ile Ser Lys Leu Lys Gln  
 130 135 140

Leu Glu Thr Leu Asn Leu Lys Asn Asn Gln Leu Thr Gly Pro Val Pro  
 145 150 155 160

Ala Thr Leu Thr Gln Ile Pro Asn Leu Lys Arg Leu Asp Leu Ala Gly  
 165 170 175  
 Asn His Leu Thr Gly Glu Ile Ser Arg Leu Leu Tyr Trp Asn Glu Val  
 180 185 190

Leu Gln Tyr Leu Gly Leu Arg Gly Asn Met Leu Thr Gly Thr Leu Ser  
 195 200 205  
 Ser Asp Met Cys Gln Leu Thr Gly Leu Trp Tyr Phe Asp Val Arg Gly  
 210 215 220

Asn Asn Leu Thr Gly Thr Ile Pro Glu Ser Ile Gly Asn Cys Thr Ser  
 225 230 235 240

Phe Gln Ile Leu Asp Ile Ser Tyr Asn Gln Ile Thr Gly Glu Ile Pro  
 245 250 255  
 Tyr Asn Ile Gly Phe Leu Gln Val Ala Thr Leu Ser Leu Gln Gly Asn  
 260 265 270

Arg Leu Thr Gly Arg Ile Pro Glu Val Ile Gly Leu Met Gln Ala Leu  
 275 280 285

Ala Val Leu Asp Leu Ser Asp Asn Glu Leu Val Gly Pro Ile Pro Pro  
 290 295 300  
 Ile Leu Gly Asn Leu Ser Phe Thr Gly Lys Leu Tyr Leu His Gly Asn  
 305 310 315 320

Met Leu Thr Gly Pro Ile Pro Ser Glu Leu Gly Asn Met Ser Arg Leu  
 325 330 335

Ser Tyr Leu Gln Leu Asn Asp Asn Lys Leu Val Gly Thr Ile Pro Pro  
 340 345 350

5

Glu Leu Gly Lys Leu Glu Gln Leu Phe Glu Leu Asn Leu Ala Asn Asn  
 355 360 365

Arg Leu Val Gly Pro Ile Pro Ser Asn Ile Ser Ser Cys Ala Ala Leu  
 370 375 380  
 Asn Gln Phe Asn Val His Gly Asn Leu Leu Ser Gly Ser Ile Pro Leu  
 385 390 395 400

Ala Phe Arg Asn Leu Gly Ser Leu Thr Tyr Leu Asn Leu Ser Ser Asn  
 405 410 415

Asn Phe Lys Gly Lys Ile Pro Val Glu Leu Gly His Ile Ile Asn Leu  
 420 425 430  
 Asp Lys Leu Asp Leu Ser Gly Asn Asn Phe Ser Gly Ser Ile Pro Leu  
 435 440 445

Thr Leu Gly Asp Leu Glu His Leu Leu Ile Leu Asn Leu Ser Arg Asn  
 450 455 460

His Leu Ser Gly Gln Leu Pro Ala Glu Phe Gly Asn Leu Arg Ser Ile  
 465 470 475 480  
 Gln Met Ile Asp Val Ser Phe Asn Leu Leu Ser Gly Val Ile Pro Thr  
 485 490 495

Glu Leu Gly Gln Leu Gln Asn Leu Asn Ser Leu Ile Leu Asn Asn Asn  
 500 505 510

Lys Leu His Gly Lys Ile Pro Asp Gln Leu Thr Asn Cys Phe Thr Leu  
 515 520 525  
 Val Asn Leu Asn Val Ser Phe Asn Asn Leu Ser Gly Ile Val Pro Pro  
 530 535 540

5 Met Lys Asn Phe Ser Arg Phe Ala Pro Ala Ser Phe Val Gly Asn Pro  
 545 550 555 560

Tyr Leu Cys Gly Asn Trp Val Gly Ser Ile Cys Gly Pro Leu Pro Lys  
 565 570 575

Ser Arg Val Phe Ser Arg Gly Ala Leu Ile Cys Ile Val Leu Gly Val  
 580 585 590  
 Ile Thr Leu Leu Cys Met Ile Phe Leu Ala Val Tyr Lys Ser Met Gln  
 595 600 605

Gln Lys Lys Ile Leu Gln Gly Ser Ser Lys Gln Ala Glu Gly Leu Thr  
 610 615 620

Lys Leu Val Ile Leu His Met Asp Met Ala Ile His Thr Phe Asp Asp  
 625 630 635 640

Ile Met Arg Val Thr Glu Asn Leu Asn Glu Lys Phe Ile Ile Gly Tyr  
 645 650 655

Gly Ala Ser Ser Thr Val Tyr Lys Cys Ala Leu Lys Ser Ser Arg Pro  
 660 665 670

Ile Ala Ile Lys Arg Leu Tyr Asn Gln Tyr Pro His Asn Leu Arg Glu  
 675 680 685  
 Phe Glu Thr Glu Leu Glu Thr Ile Gly Ser Ile Arg His Arg Asn Ile  
 690 695 700

Val Ser Leu His Gly Tyr Ala Leu Ser Pro Thr Gly Asn Leu Leu Phe  
 705 710 715 720  
 Tyr Asp Tyr Met Glu Asn Gly Ser Leu Trp Asp Leu Leu His Gly Ser  
 725 730 735

Leu Lys Lys Val Lys Leu Asp Trp Glu Thr Arg Leu Lys Ile Ala Val  
 740 745 750

Gly Ala Ala Gln Gly Leu Ala Tyr Leu His His Asp Cys Thr Pro Arg  
 755 760 765  
 Ile Ile His Arg Asp Ile Lys Ser Ser Asn Ile Leu Leu Asp Glu Asn  
 770 775 780

Phe Glu Ala His Leu Ser Asp Phe Gly Ile Ala Lys Ser Ile Pro Ala  
 785 790 795 800

Ser Lys Thr His Ala Ser Thr Tyr Val Leu Gly Thr Ile Gly Tyr Ile  
 805 810 815  
 Asp Pro Glu Tyr Ala Arg Thr Ser Arg Ile Asn Glu Lys Ser Asp Ile  
 820 825 830

Tyr Ser Phe Gly Ile Val Leu Leu Glu Leu Leu Thr Gly Lys Lys Ala  
 835 840 845

Val Asp Asn Glu Ala Asn Leu His Gln Leu Ile Leu Ser Lys Ala Asp  
 850 855 860

5

Asp Asn Thr Val Met Glu Ala Val Asp Pro Glu Val Thr Val Thr Cys  
 865 870 875 880

Met Asp Leu Gly His Ile Arg Lys Thr Phe Gln Leu Ala Leu Leu Cys  
 885 890 895



Thr Lys Arg Asn Pro Leu Glu Arg Pro Thr Met Leu Glu Val Ser Arg  
 900 905 910

Val Ieu Leu Ser Leu Val Pro Ser Leu Gln Val Ala Lys Lys Leu Pro  
 915 920 925

Ser Ieu Asp His Ser Thr Lys Lys Leu Gln Gln Glu Asn Glu Val Arg  
 930 935 940  
 Asn Pro Asp Ala Glu Ala Ser Gln Trp Phe Val Gln Phe Arg Glu Val  
 945 950 955 960

Ile Ser Lys Ser Ser Ile  
 965

5 <210> 11  
 <211> 636  
 <212> ADN  
 <213> *ERECTA* de trigo parcial

10 <400> 11  
 atgaattctg caatacctag gcttgagggg taactcactg actggaacct tgtcacctga 60  
 catgtgcaa ctcactggcc tgtggtactt tgatgtgagg ggcaacaatc taacaggaac 120  
 aattccacag agcatagggg actgcactag ctttgagatt ctggacattt catatacaa 180  
 aatctctgga gaaatacctt acaacatagg tttccttcaa gtagctacac tgtcacttca 240  
 aggaaataga ctgactggga aaattccaga agtgattggc ctcatgcaag ctcttgctgt 300  
 tcttgatctg agcgaaaacg aactagtagg ggccattcct ccgatactcg gcaacctgtc 360  
 ctacactggc aaactatatt tgcatggcaa taaacttact ggtgaagtac ccccggaact 420  
 tgggaacatg actaaactta gctacctgca actgaatgac aatgaattag tgggcgcaat 480  
 tccagctgag cttgggaaac ttgaagagct attcgaatta aatcttgcca acaacaatct 540  
 tgagggtcct attcctacaa acatcagttc ttgcactgca ctaaacaat tcaatgttta 600  
 cggcaataga ttgaacgggt ctatccctgc tggttt 636

15 <210> 12  
 <211> 466  
 <212> ADN  
 <213> *ERECTA* de trigo parcial

20 <400> 12  
 ttcaatgttt atggcaatag attgaacggt tctatccctg ctggtttcca gaatttgag 60  
 agtttgacta acttgaattt atcctcaaac aattttaaag gccatatccc atctgaactt 120  
 ggtcatatca tcaatttggg cacactggat ctttctaca atgaactctc tggaccagtt 180  
 cctgctacta ttggtgatct tgagcatcct cttcaactaa atttgagcaa aaacctctt 240

# ES 2 362 903 T3

```

agcgggtcag tgcctgctga gttcggaaac ttgagaagca tccaagtaat tgatttatcc 300

aacaacgcca tgtctgggta tctccctgaa gaactaggcc aacttcagaa ccttgatagt 360

ttgattctta acaacaatat tttggtcgga gagatccctg ctcagttggc taactgcttc 420

agcttaaaca tcttgaactt gtcacataac aacttttctg gacatg 466

<210> 13
<211> 372
5 <212> ADN
   <213> ERECTA de trigo parcial

<220>
<221> misc_feature
10 <222> (62)..(62)
   <223> no determinado

<400> 13

15 ttgtcagcct tctggcttct cactctctcc caatggaaac ctgctcttct acgattacat 60
   gngaaaacgg ttccttggg gatccttctcc atggtccatc aaagaaagtg aagcttgact 120

   gggacacccg actgaggatc gcggtcggcg cggcacaagg gctggcctat ctgcaccatg 180

   actgcaatc gcggatagtc cacagggagc tcaagtcctc caacatcctg ctgcagcagc 240
   actttgaagc gcatctctcg gacttcggca tcgccaaatg cgtcccggca gccaaagacc 300

   atgcgtccac atatgtgcta ggaaccatcg gctacatcga tccagagtac gcccgagcgt 360

   cgaggtgaa cg 372

<210> 14
20 <211> 357
   <212> ADN
   <213> ERECTA de trigo parcial

<400> 14

25 ttgactaact tgaatttatc ctcaacaat tttaaaggcc atatcccatc tgaacttgg 60

   catatcatca atttggacac actggatctt tcctacaatg aactctctgg accagttcct 120

   gctactattg gtgatcttga gcatcttctt caactaaatt tgagcaaaaa ccatcttagc 180
   gggtcagtgc ctgctgagtt cggaaacttg agaagcatcc aagtaattga tttatccaac 240

   aacgccatgt ctggttatct ccctgaagaa ctaggccaac ttcagaacct tgatagtttg 300

   attcttaaca acaatatttt ggtcggggag atccctgctc agttggctaa ctgcttc 357

<210> 15
30 <211> 314
   <212> ADN
   <213> ERECTA de trigo parcial

<400> 15

35 cacactggat ctttcttaca atgaattctc tggaccagac cctgctacta ttggtgatct 60

   tgagcatggt cttcagatta aatttgagca aaaacatct tactgggcca atgcctgctg 120

```

	agttctgaaa cttgagaagc atccaagtaa ttgatttata caacaacgcc atgtctgggt	180
	atctccctga agaactacgc caacttcaga atcttgatag tttgatgctt aacaacaata	240
	ctttgggttg ggagatccct gctcatctgg ctaactgctt caacttaaac atcttgaact	300
	tgccatataa caac	314
	<210> 16	
	<211> 549	
5	<212> ADN	
	<213> <i>ERECTA</i> de trigo parcial	
	<400> 16	
	catcatcggc tacggcgctt caagtaccgt gtataaatgt gtgctcaaga gtggcaaggc	60
	cattgctgtg aagcggctct acagccaata caacatggc gcccgtagt ttgagacaga	120
	gctggagaca gtcggtagca tccggcacag gaatctgtc agccttcag gcttctcact	180
10	ctctccaaat ggaaacctgc tcttctacga ttacatggaa aatggttcct tgtgggatct	240
	tctccacggc ccatcaaaga aggtgaaact tgactgggac acccgactga gaatcgccgt	300
	cgggtcggca caagggctgg catatcttca ccatgactgc aaccctcgga tagtccacag	360
	ggacgtcaag tcctccaaca tcttctcga cgagcacttt gaagcgcac tctcggactt	420
	cggcatcggc aaatgcgtcc cagctgcaa gacccacgag tccacctatg tgctaggaac	480
	catcggctac atcgatccag agtacgccc gacgtcccag ctgaacgaga aatctgatgt	540
	gtacagctt	549
	<210> 17	
15	<211> 615	
	<212> ADN	
	<213> <i>ERECTA</i> de trigo parcial	
	<400> 17	
20	ccacgcgtcg atcatcaatt gggacacacg ggatctttcc tacaatgaat tctccgggcc	60
	agttcctgct actattggtg atctggagca tcttcttcaa ctaaatttga gcaaaaacca	120
	tcttagtggg tctgtgctg ctgagttcgg aaacttgaga agcatccaag taattgattt	180
	atccaacaac gccatttctg gttatctccc tgaagaacta ggccaacttc agaaccttga	240
	tagtttgatt cttaaacaaca atactttggt tggggagatc cctgctcagt tggctaactg	300
	cttcagctta aacatcttga acttgcata taacaacttt tctggacatg tcccattcgc	360
	taagaacttc tcaaagtcc cgggggaaag cttcttggga aatccgatgc tgagcgttca	420
	ctgcaaagac tccagctgtg gcaactctca tggatcaaaa gtgaatactc ggacagcgat	480
	tgcttgcatc atctcgggct tctgcatatt gctctgtgtt ctgctattgg gcaatatata	540
	aaacaaagcg accacagcca cctatcaaag catctgataa accagggcaa ggacctccaa	600
	agatagtact cctcc	615
25	<210> 18	
	<211> 719	

<212> ADN  
 <213> *ERECTA* de trigo parcial

<400> 18

5  
 cgttcactgc aaagactcca gctgtggcaa ctctcatgga tcaaaagtga atattcggac 60  
 ggcgattgct tgcacatcct cgggcttcgt catactgcta tgtgttctgc tattggcaat 120  
 atataaaaca aagcgaccac agccacctat caaagcatct gataaaccag tgcaaggacc 180  
 tccaaagata gtactcctcc aaatggacat ggctatccat acctatgatg atattatgag 240  
 gctgacagag aatttgagcg agaatacat catcggctac ggcgcttcaa gtaccgtgta 300  
 taaatgtgtg ctcaagagtg gcaaggccat tgctgtgaag cggctctaca gccaatataa 360  
 ccatggcgcc cgtgagtttg agacagagct ggagacagtc ggtagcatcc ggcacaggaa 420  
 tcttgtcagc ettcatggct tctcactctc tccaaatgga aacctgctct tctacgatta 480  
 catggaaaat ggttctctgt gggatcttct ccacggctca tcaagaagg tgaaacttga 540  
 ctgggacacc cgactgagaa tcgccgtcgg tgcggcaca gggctggcat atcttcacca 600  
 tgactgcaac cctcggatag tccacagga cgtcaagtcc tccaacatcc tgctcgacga 660  
 gcactttgaa gcgcatctct cggacttcgg catcgcccaa tgcgtccca gctgccaag 719

10 <210> 19  
 <211> 1346  
 <212> ADN  
 <213> *ERECTA* de trigo

15 <400> 19

ttcaatgttt atggcaatag attgaacggt tctatccctg ctggtttcca gaatttgag 60  
 agtttgacta acttgaattt atcctcaaac aattttaaag gccatatccc atctgaactt 120  
 ggtcatatca tcaatttga cacactggat ctttctaca atgaactctc tggaccagtt 180  
 cctgactacta ttggatctct tgagcatctt cttcaactaa atttgagcaa aaaccatctt 240  
 agcgggtcag tgctcgtgta gttcggaaac ttgagaagca tccaagtaat tgatttatcc 300  
 aacaacgcca tgtctgggta tctccctgaa gaactaggcc aacttcagaa ccttgatagt 360  
 ttgattctta acaacaatat tttggtcggg gagatccctg ctcagttggc taactgcttc 420  
 agcttaaaca tcttgaactt gtcacataac aactttctg gacatgtccc attcgctaag 480  
 aacttctcaa agttccccgg ggaaagcttc ttgggaaatc cgatgctgag cgttcactgc 540  
 aaagactcca gctgtggcaa ctctcatgga tcaaaagtga atactcggac agcgattgct 600  
 tgcacatcct cgggcttcgt catactgcta tgtgttctgc tattggcaat atataaaaca 660  
 aagcgaccac agccacctat caaagcatct gataaaccag ggcaaggacc tccaaagata 720  
 gtactcctcc aaatggacat ggctatccat acctatgatg atattatgag gctgacagag 780  
 20 aatttgagcg agaaatacat catcggctac ggcgcttcaa gtaccgtgta taaatgtgtg 840

ctcaagagtg gcaaggccat tgctgtgaag cggctctaca gccaatacaa ccatggcgcc 900  
 cgtgagtttg agacagagct ggagacagtc ggtagcatcc ggcacaggaa tcttgtcagc 960  
 cttcatggct tctcactctc tccaaatgga aacctgctct tctacgatta catggaaaat 1020  
 ggttccttgt gggatcttct ccacggtcca tcaaagaagg tgaaacttga ctgggacacc 1080  
 cgactgagaa tcgccgtcgg tgcggcacia gggctggcat atcttcacca tgactgcaac 1140  
 cctcggatag tccacagga cgtaagtcc tccaacatcc tgctcgacga gcactttgaa 1200  
 gcgcatctct cggacttcgg catcgcaaaa tgcgtcccag ctgccaagac ccacgcgtcc 1260  
 acatatgtgc taggaacctc cggctacatc gatccagagt acgccggac gtcccagctg 1320  
 aacgagaaat ctgatgtgta cagctt 1346

5 <210> 20  
 <211> 448  
 <212> PRT  
 <213> *ERECTA* de trigo

10 <400> 20

Phe Asn Val Tyr Gly Asn Arg Leu Asn Gly Ser Ile Pro Ala Gly Phe  
 1 5 10 15  
 Gln Asn Leu Glu Ser Leu Thr Asn Leu Asn Leu Ser Ser Asn Asn Phe  
 20 25 30  
 Lys Gly His Ile Pro Ser Glu Leu Gly His Ile Ile Asn Leu Asp Thr  
 35 40 45  
 Leu Asp Leu Ser Tyr Asn Glu Leu Ser Gly Pro Val Pro Ala Thr Ile  
 50 55 60  
 Gly Asp Leu Glu His Leu Leu Gln Leu Asn Leu Ser Lys Asn His Leu  
 65 70 75 80  
 Ser Gly Ser Val Pro Ala Glu Phe Gly Asn Leu Arg Ser Ile Gln Val  
 85 90 95  
 Ile Asp Leu Ser Asn Asn Ala Met Ser Gly Tyr Leu Pro Glu Glu Leu  
 100 105 110  
 Gly Gln Leu Gln Asn Leu Asp Ser Leu Ile Leu Asn Asn Asn Ile Leu  
 115 120 125  
 Val Gly Glu Ile Pro Ala Gln Leu Ala Asn Cys Phe Ser Leu Asn Ile  
 130 135 140  
 Leu Asn Leu Ser His Asn Asn Phe Ser Gly His Val Pro Phe Ala Lys  
 145 150 155 160

15

Asn Phe Ser Lys Phe Pro Gly Glu Ser Phe Leu Gly Asn Pro Met Leu  
 165 170 175

Ser Val His Cys Lys Asp Ser Ser Cys Gly Asn Ser His Gly Ser Lys  
 180 185 190

Val Asn Thr Arg Thr Ala Ile Ala Cys Ile Ile Ser Gly Phe Val Ile  
 195 200 205

Leu Leu Cys Val Leu Leu Leu Ala Ile Tyr Lys Thr Lys Arg Pro Gln  
 210 215 220

Pro Pro Ile Lys Ala Ser Asp Lys Pro Gly Gln Gly Pro Pro Lys Ile  
 225 230 235 240  
 Val Leu Leu Gln Met Asp Met Ala Ile His Thr Tyr Asp Asp Ile Met  
 245 250 255

Arg Leu Thr Glu Asn Leu Ser Glu Lys Tyr Ile Ile Gly Tyr Gly Ala  
 260 265 270

Ser Ser Thr Val Tyr Lys Cys Val Leu Lys Ser Gly Lys Ala Ile Ala  
 275 280 285  
 Val Lys Arg Leu Tyr Ser Gln Tyr Asn His Gly Ala Arg Glu Phe Glu  
 290 295 300

Thr Glu Leu Glu Thr Val Gly Ser Ile Arg His Arg Asn Leu Val Ser  
 305 310 315 320

Leu His Gly Phe Ser Leu Ser Pro Asn Gly Asn Leu Leu Phe Tyr Asp  
 325 330 335  
 Tyr Met Glu Asn Gly Ser Leu Trp Asp Leu Leu His Gly Pro Ser Lys  
 340 345 350

Lys Val Lys Leu Asp Trp Asp Thr Arg Leu Arg Ile Ala Val Gly Ala  
 355 360 365

Ala Gln Gly Leu Ala Tyr Leu His His Asp Cys Asn Pro Arg Ile Val  
 370 375 380  
 His Arg Asp Val Lys Ser Ser Asn Ile Leu Leu Asp Glu His Phe Glu  
 385 390 395 400

Ala His Leu Ser Asp Phe Gly Ile Ala Lys Cys Val Pro Ala Ala Lys  
 405 410 415

Thr His Ala Ser Thr Tyr Val Leu Gly Thr Ile Gly Tyr Ile Asp Pro  
 420 425 430

Glu Tyr Ala Arg Thr Ser Gln Leu Asn Glu Lys Ser Asp Val Tyr Ser  
 435 440 445

5

<210> 21  
 <211> 1273  
 <212> ADN  
 <213> *ERECTA* de maíz parcial

5

<400> 21

cctggactct gtggatattg gcttggttct tcatgtcgtt cactggcca cggagacaaa 60  
 ccgccaatct caaaggctgc cataattggt gttgctgtgg gtggacttgt taccctcctg 120  
 atgatcttag tagctgtatg caggccacac catccacctg cttttaaaga tgccactgta 180  
 agcaagccag tgagcaatgg tccacccaag ctggtgatcc ttcataatgaa catggctctt 240  
 catgtccttg atgatataat gaggatgact gagaacttga gtgagaaata catcattgga 300  
 tacggggcat caagtacagt ttataaatgt gttctaaaga attgcaaacc agtggcaata 360  
 aaaaagctgt atgccacta ccctgcagag ccttaaggaa ttgaaactg agctcgagac 420  
 tgttgtagc atcaaacacc ggaatctagt cagccttgcc aagggtactc gttgtcacct 480  
 gtgggaacc tcctcttcta tgattatag gagagtggca gcttatggga tgttttcat 540  
 gaaggctcat ccaagaagaa caaacttgac tgggtgactc gcctacggat cgctcttgg 600  
 gcagctcaag gcctcgctta ccttaccat gactgcagcc cacgaataat tcaccgggac 660  
 10 gtaaaatcaa agaataact cctcgacaaa gattatgagg cccatcttac agacttcggc 720  
 atcgctaaga gcttatgtgt ctggaagact cacacgtcaa cctacgtcat gggcactatt 780  
 ggttacattg atcccagta cgcccgcacc tcccgcctca acgagaagtc tgatgtctac 840  
 agctacggca tcgttctgct ggagctgctg accggcaaga agccagtggga caacgagtgc 900  
 aatctccatc acttgatcct atcgaagacg gcgagcaacg aggtcatgga gacggtggac 960  
 cccgacgtgg gagacacctg caaggacctg ggcgaggtga agaagctggt ccagctggcg 1020  
 ctctctgca ccaagcggca gccctcggac cggccgacga tgcacgaggt ggtgcgcgtc 1080  
 ctgactgccc tgggtgaacc ggagccgccc ccgacggccg agcagcagca gcagaagggc 1140  
 gcacgcgcac caccagctgc cgccgcagcc gtcgcccggc gcctacgtcg acgagtacgt 1200  
 cagcctgccc ggcactggcg ccctctctg cgccaactcg tccagcacct cggacgccga 1260  
 gctgttcctc aag 1273

15

<210> 22  
 <211> 100  
 <212> ADN  
 <213> *ERECTA* de maíz parcial

20

<400> 22

cacaaaatgt cagtcaact actccccctg caatcggcct cactcaaggc gcctcaccga 60  
 acgtctacgt cttcccctac accatgttct gcgagatggc 100

25

<210> 23  
 <211> 599  
 <212> ADN  
 <213> *ERECTA* de maíz parcial

<220>  
 <221> misc\_feature  
 <222> (529)..(529)  
 5 <223> no determinado

<400> 23

tttttttttt tttttttttt tttttgagga ggaagctccg ctgctcttgc gttgctcac 60  
 atgacttttt acagctaaca acaccctagc tactgagtcc catgttaatc tcctgctgctg 120

cgccccacaa aatgtcagtc aaactactcc ctgcaatcgg cctcaactcaa ggcgcctcac 180

cgaacgtcta cgtcttcccc tacaccatgt tctgcgagat ggcctcgccg aacttgagga 240

10 acagctcggc gtccgaggtg ctggacgagt tggcgagga gagggcgccg gtgccccgca 300  
 ggctgacgta ctgctcgacg taggccggcg gcgacggctg cggcggcagc tgggtggtgctg 360

cgtgctcctt ctgctgctgc tgctgctgct gcggcgcccg ctccgggttc accaggcagt 420

caaggacgcg caccacctcg tgcacgtcgc gccggtccga gggctgccgc ttggtgcaga 480  
 ggagcgcag ctggaacagc ttcttcacct cgccaggctc cttgcaggng tctcccacgt 540

cggggtccac cgtctccatg acctcgttgc tcgccgttct cgataggatc aaggatgga 599

<210> 24  
 15 <211> 436  
 <212> ADN  
 <213> *ERECTA* de maíz parcial

<400> 24

20 tttttttttt ttttttttga agaagctccg ctgctctcgc gttgctcac atgacttttt 60

acagataaca ccaccctagc tactgagtcc catgttaatc tcctgctgctg cgttccacaa 120

aatgtcagcc aaactactcc ctgcaatcgg cctcaactcaa ggcgcctcac cgaacgtcta 180

cgtcttcccc tacaccatgt tctgcgagat ggcctcgccg aacttgagga acagctcggc 240  
 gtccgaggtg ctggacgagt tggcgagga gagggcgccg gtgccccgca ggctgacgta 300

ctgctcgacg taggccggcg gggacggctg cggcggcagc tgggtggtgctg cgtgctcctt 360

ctgctgctgc tgctgctgct gcggcgcccg ctccgggttc accaggcagt caaggacgcg 420

caccacctcg ggcac 436

<210> 25  
 25 <211> 509  
 <212> ADN  
 <213> *ERECTA* de maíz parcial

<400> 25

30 ccaaagaaaa atggagaggg gggataaaga agatgaggaa gaagctccgc tgctcttgcg 60

ttgctcaca tgacttttta cagtaacaa caccctagct actgagtccc atgttaatct 120

cctgctgctgc gtccccacaa atgtcagtca aactactccc tgcaatcggc ctcaactcaag 180

gcgcctcacc gaacgtctac gtcttcccc acaccatgtt ctgctgagatg gcctcgccga 240



acttgaggaa cagctcggcg tccgaggtgc tggacgagtt ggcgaggag agggcgccgg 300

tgccccgag gctgacgtac tcgtcgacgt aggccggcgg cgacggctgc ggcgagct 360

ggtggtgccc gtgcgcttc tgctgctgct gctgcggctg cggcgccggc tccgggttca 420

ccaggcagtc aaggacggc accacctcgt gcatcgtcgg ccggtccgag ggctgcccgt 480  
 tggcagag gagcggcagc tggaacagc 509

5 <210> 26  
 <211> 318  
 <212> ADN  
 <213> *ERECTA* de maíz parcial

10 <400> 26

gatggatcaa tacagcctcc tagtaagtta gaccaccaa gaaaaatggg gaggggggat 60

aaagaagagg aagaagctcc gctgctcttg cgtcacatga ctttttttac agctaacaac 120

accctagcta ctgagtccca tgtaatctc ctgcgctgcg tcccacaaaa tgtcagtcaa 180  
 actactcccc ctgcaatcgg cctcactcaa ggcgctcac cgaacgtcta cgtcttcccc 240

tacaccatgt tctgcgagat ggcctcggcg aacttgagga acagctcggc gtccgaggtg 300

ctggacgagt tggcgagc 318

15 <210> 27  
 <211> 103  
 <212> ADN  
 <213> *ERECTA* de maíz parcial

20 <400> 27

agcaagccag tgagcaatgg tccaccaag ctggtgatcc ttcatatgaa catggctctt 60

catgtctttg atgatataat gaggatgact gagaacttga gtg 103

25 <210> 28  
 <211> 458  
 <212> ADN  
 <213> *ERECTA* de maíz parcial

<400> 28

ataattcacc gggacgtaaa atcaaagaat atactcctcg acaaagatta tgaggcgcat 60

cttacagact tcggcatcgc taagagctta tegtgtctcga agactcacac gtcaacctac 120

gtcatgggca ctattggtac acttgatcct gagtacgccc gcacctccc cctcaacgag 180

30 aagtctgatg tctacagcta cggcatcgtt ctgctggagc tgctgaccgg caagaagcca 240  
 gtggacaacg agtgcaatct ccatcacttg atcctatcga agacggcgag ccaacgaggt 300

catggagacg gtggaccccc acgtgggaga cacctgcaag gacctgggag aggtgaagaa 360

gctgttccag ctggcgtccc tctgcaccaa gcggcagccc tcggaccggc cgacgatgca 420

cgaggtggty cgcgtccttg actgcctggt gaacccgg 458

35 <210> 29  
 <211> 593  
 <212> ADN

<213> *ERECTA* de maíz parcial

<400> 29

5 tttttttttt tttttttttt ttttttgagg aagaagctcc gctgctcttg cgttgctgca 60  
 catgactttt tacagctaac aacaccctag ctactgagtc ccatgttaat ctctgctgct 120  
 gcgtcccaca aaatgtcagt caaactactc cctgcaatcg gcctcatttt tttgttgctc 180  
 tcaccgaacg tctacgtctt cccctacacc atgttctgcg agatggcctc gccgaacttg 240  
 aggaacagct cggcgtccga ggtgctggac gagttggcgc aggaaagggc gccggtgccc 300  
 cgcaggctga cgtactcgtc gacgtaggcc ggcggcgacg gctgcggcgg cagctggtgg 360  
 tgcgcgtgcg ccttctgctg ctgctgctgc ggctgcggcg gcggctccgg gttcaccagg 420  
 cagtcaagga cgcgcaccac ctctgtcctc gtcggccggg ccgagggctg ccgcttggtg 480  
 cagaggagcg ccagctggaa cagcttcttc acctcgccca ggtccttgca ggtgtctccc 540  
 acgtcggggg ccaccggctc catgacctcg ttgctcgccg tcttcgatag gat 593

<210> 30

<211> 206

<212> ADN

<213> *ERECTA* de maíz parcial

<400> 30

15 tcacaaaaga tcatcaagca gaggaacggg agagatgatg atggatcaat acagcctcct 60  
 agtaagttag accacaaaaga aaaatgggga ggggggataa agaagaggaa gaagctccgc 120  
 tgctcttgcg tcacatgact ttttttacag ctaacaacac cctagctact gagtcccatg 180  
 ttaatctcct gcgctgctc ccacaa 206

<210> 31

<211> 534

<212> ADN

<213> *ERECTA* de maíz parcial

<400> 31

25 caagcagagg aacgggagag atgatgatgg atcaatacag cctcctagta agttagacca 60  
 caaagaaaaa tggggagggg ggataaagaa gaggaagaag ctccgctgct cttgctcac 120  
 atgacttttt ttacagctaa caacacccta gctactgagt cccatgttaa tctcctgctc 180  
 tgcgtcccac aaaatgtcag tcaaactact ccccctgcaa tcggcctcac tcaaggcggc 240  
 tcaccgaacg tctacgtctt cccctacacc atgttctgcg agatggcctc gccgaacttg 300  
 aggaacagct cggcgtccga ggtgctggac gagttggcgc aggagagggc gccggtgccc 360  
 cgcaggctga cgtactcgtc gacgtaggcc ggcggcgacg gctgcggcgg cagctggtgg 420  
 tgcgcgtgcg ccttctgctg ctgctgctgc ggctgcggcg gcggctccgg gttcaccagg 480  
 cagtcaagga cgcgcaccac ctctgtcctc gtcggccggg ccgagggctg ccgc 534

<210> 32

<211> 527

30

<212> ADN  
<213> *ERECTA* de maíz parcial

<400> 32

5  
 gaaagtcaca agatcataag gaagaggaac gggagagatg atgatggatc aatacagcct 60  
 cctagtaagt tagaccacca aagaaaaatg gagagggggg ataaagaaga tgaggaagaa 120  
 gctccgctgc tcttgcggtg cgtcacatga ctttttacag ctaacaacac cctagctact 180  
 gagtcccatg ttaatctcct gcgctgcgtc ccacaaaatg tcagtcaaac tactccctgc 240  
 aatcggcctc actcaaggcg cctcaccgaa cgtctacgtc ttcccctaca ccatgttctg 300  
 cgagatggcc tcgcccgaact tgaggaacag ctcggcgtcc gaggtgctgg acgagttggc 360  
 gcaggaagg gcgccggtgc cccgcaggct gacgtactcg tcgacgtagg ccggcggcga 420  
 cggctgcggc ggcagctggt ggtgcgctg cgccttctgc tgetgctgct gcgctgcgg 480  
 cggcggctcc gggttcacca ggcagtcaag gacgcgcacc acctcgt 527

10 <210> 33  
 <211> 412  
 <212> ADN  
 <213> *ERECTA* de maíz parcial

15 <400> 33

cttgcggtgc gtcacatgac tttttacagc taacaacacc ctactactg agtcccatgt 60  
 taatctctg cgctgcgtcc cacaaaatg cagtcaaact actccctgca atcggcctca 120  
 ctcagggggc ctcaccgaac gtctactctt tcccctacac caggttctgc gagatggcct 180  
 cgccgaactt gaggaacagc tcggcgtccg aggggctgga cgagttggcg caggaaaggg 240  
 cgccggggcc ccgaggctg acgtactcgt cgacgtaggc cggcggcgac ggctgcggcg 300  
 gcagctgggg gtgcgctgtc gccttctgct gctgctgctg cggttgcggc gggcggctccg 360  
 ggttcaccag gcagtcaagg acgcgcacca cctcgggcat cgtcggccgg tc 412

20 <210> 34  
 <211> 533  
 <212> ADN  
 <213> *ERECTA* de maíz parcial

25 <400> 34

tcgagttttt tttttttttt ttttgatgat ggatcaatac agcctcctag taagttagac 60  
 caccaaagaa aaatggagag gggggataaa gaagatgagg aagaagctcc gctgctcttg 120  
 cgttgcgtca catgactttt tacagctaac aacaccctag ctactgagtc ccatgttaat 180  
 ctctgcgct gcgtoccaca aaatgtcagt caaactactc cctgcaatcg gcctcactca 240  
 aggcgctca ccgaactgct acgtcttccc ctacaccatg ttctgcgaga tggcctcggc 300  
 gaacttgagg aacagctcgg cgtccgaggt gctggacgag ttggcgagg agagggcgcc 360  
 ggtgccccgc aggtgacgt actcgtcgac gtaggcccgc ggcgacggct gcggcggcag 420  
 ctggtggtgc cgtgcgcct tctgctgctg ctgctgcggc tgcggcggcg gctccgggtt 480

	caccaggcag tcaaggacgc gcaccacctc gtgcatcgtc ggccgggccg agg	533
	<210> 35	
	<211> 191	
5	<212> ADN	
	<213> <i>ERECTA</i> de maíz parcial	
	<400> 35	
	agcctcctag taagttagac caccaaagaa aaatggagag gggggataaa gaagatgagg	60
	aagaagctcc gctgctcttg cgttgctgca catgactttt tacagctaaa caacacccta	120
	gctactgagt cccatggtaa tctcctgctc tgcgtccac aaaatgtcag tcaaactact	180
10	ccctgcaatc g	191
	<210> 36	
	<211> 683	
15	<212> ADN	
	<213> <i>ERECTA</i> de maíz parcial	
	<400> 36	
	gacggtggga acctcctctt ttatgcttta tggagagtgg cagcttatgg gatgttttac	60
	atgaaggctc atccaagaag aacaaacttg actgggtgac tgcctacgg atcgtctctg	120
	gtgcagctca aggcctcgtc taccttcacc atgactgcag cccacgaata attcaccggg	180
	acgtaaaatc aaagaatata ctctcgcaca aagattatga ggcgcatctt acagacttcg	240
	gcatcgctaa gagcttatgt gtctcgaaga ctcacacgctc aacctacgctc atgggcacta	300
	ttggttacat tgatctgag tacgcccga cctcccgcct caacgagaag tctgatgtct	360
	acagctacgg catcgttctg ctggagctgc tgaccggcaa gaagccagtg gacaacgagt	420
20	gcaatctcca tcacttgatc ctatcgaaga cggcgagcaa cgaggctatg gagacggtgg	480
	accccagctg gggagacacc tgcaaggacc tgggcgaggt gaagaagctg ttccagctgg	540
	cgctcctctg caccaagcgg cagccctcgg accggccgac gatgcacgag gtggtgcgcg	600
	tccttgactg cctggtgaac ccggagccgc cgccgcagcc gcagcagcag cagcagaagg	660
	cgcacgcgca ccaccagctg ccg	683
25	<210> 37	
	<211> 610	
	<212> ADN	
	<213> <i>ERECTA</i> de maíz parcial	
30	<400> 37	
	cttcggcatc gctaagagct tatgtgtctc gaagactcac acgtcaacct acgtcatggg	60
	cactattggt tacattgatc ctgagtacgc ccgcacctcc cgcctcaacg agaagtctga	120
	tgtctacagc tacggcatcg ttctgtctga gctgctgacc ggcaagaagc cagtggacaa	180
	cgagtgcaat ctccatcact tgatcctatc gaagacggcg agcaacgagg tcatggagac	240

# ES 2 362 903 T3

```

ggtaggacccc gacgtgggag acacctgcaa ggacctgggc gaggtgaaga agctgttcca 300
gctggcgctc ctctgcacca agcggcagcc ctccggaccgg ccgacgatgc acgaggtggt 360
gcgcgtcctt gactgcctgg tgaaccggga gccgccggc cagccgagc agcagcagca 420
gaaggcgcac gcgcaccacc agctgccgcc gcagccgtcg ccgccggcct acgtcgacga 480
gtacgtcagc ctgcggggca ccggcgccct ctctgcgcc aactcgtcca gcacctcgga 540
cgccgagctg ttctcaagt tcggcgaggc catctcgag aacatggtgt aggggaagac 600
gtagacgttc 610
5 <210> 38
  <211> 208
  <212> ADN
  <213> ERECTA de maíz parcial
10 <220>
  <221> misc_feature
  <222> (138)..(138)
  <223> no determinado
15 <400> 38
  gcaagccagt gagcaatggt ccaccaagc tgggatcctt catatgaaca tggctcttca 60
  tgtctttgat gatataatga ggatgactga gaacttgagt gaaaaataca tcattggata 120
  cggggcatca agtactgntt ataaatggtt tctaaagaat tgcaaacagc tggcaataaa 180
  aaagctgtat gccactacc ctgagagc 208
20 <210> 39
  <211> 634
  <212> ADN
  <213> ERECTA de maíz parcial
  <400> 39
  gaccgggacg taaaatcaaa gaatatactc ctgcacaaag attatgaggc gcatcttaca 60
  gacttcggca tcgctaagag cttatgtgtc tcgaagactc acacgtcaac ctacgtcatg 120
  ggcactattg gttacattga tcctgagtac gcccgcacct cccgcctcaa cgagaagtct 180
25 gatgtctaca gctacggcat cgttctgctg gagctgctga ccggcaagaa gccagtggac 240
  aacgagtgca atctccatca cttgatccta tcgaagacgg cgagcaacga ggtcatggag 300
  acggtggacc ccgacgtggg agacacctgc aaggacctgg gcgaggtgaa gaagctgttc 360
  cagctggcgc tcctctgcac caagcggcag ccctcggacc ggccgacgat gcacgaggtg 420
  gtgcgcgtcc ttgactgcct ggtgaaccgg gagccggcgc cgcagccgca gcagcagcag 480
  cagaaggcgc acgcgcacca ccagctgccg ccgcagccgt cgcgccggc ctacgtcgac 540
  gagtacgtca gcctcggggg caccggcgc ctctcctgcg ccaactcgtc cagcacctcg 600
  gacgccgagc tgttcctcaa gttcggcgag gcca 634
30 <210> 40
  <211> 558

```

<212> ADN  
 <213> *ERECTA* de maíz parcial

<400> 40

5  
 acttgatgcc ccgatatcaa tgatgtatct ctcactcaag ttctcagtc tccctattat 60  
 atcatcaaag acatgaagag ccatgttcat atgaaggatc accagcttgg gtggaccatt 120  
 gtcactggc ttgcttacag tggcatcttt aaaagcaggt ggatgggtg gcctgcatac 180  
 agctactaag atcatcagga ggataacaag tccaccaca gcaacacaa ttatggcagc 240  
 ctttgagatt ggcggtttgt ctccgtggcc agtggaaacga catgaagaac caagccaata 300  
 tccacagagt ccaggattac ctaaaaagct gtcattgtgaa aaccgtgtga agttgttctc 360  
 agtagggaca gcaccagcca aattattgta tgacacattt aagatattga ggctgaagca 420  
 gttcatcaga gaagagacat cgccagttat attgtgtttt tccagttta gcaacatcag 480  
 gttttgcagc attccaagtt cttgaggaat cagaccacca agatgattat aggataaatc 540  
 aatctccatg acacttct 558

10 <210> 41  
 <211> 429  
 <212> ADN  
 <213> *ERECTA* de maíz parcial

15 <400> 41

tacttgatgc cccgatatca atgatgtatt tctcactcaa gttctcagtc atcctcatta 60  
 tatcatcaaa gacatgaaga gccatgttca tatgaaggat caccagcttg ggtggaccat 120  
 tgctcactgg cttgcttaca gtggcatctt taaaagcagg tggatgggtg ggcctgcata 180  
 cagctactaa gatcatcagg aggataacaa gtccaccac agcaacacca attatggcag 240  
 cctttgagat tggcggtttg tctcgtggcc cagtggaaac acatgaagaa ccaagccaat 300  
 atccacagag tccaggatta ctaaaaagc tgcattgtga aaaccgtgtg aagttgttgt 360  
 cagtagggac agcaccagcc aaattattgt atgacacatt taagatattg aggctgaagc 420  
 agttcatca 429

20 <210> 42  
 <211> 556  
 <212> ADN  
 <213> *ERECTA* de maíz parcial

25 <400> 42

acatgcaagt caacaggta actggatcga taccaccaga gctaggaat atgtcaacac 60  
 ttcatcactt agaactgaat gataatcaac ttactgggtc aattccacca gagcttggaa 120  
 ggctaacagc cttgtttgac ctgaacctg cgaataacca ccttgaagga ccaattcctg 180  
 acaacctaag ttcattgtgt aatctcaata gcttcaatgc ttatggcaac aagttaaatg 240

	gaaccattcc tcgttcgctg cggaaacttg aaagcatgac ctatttaa	300
	atccataag tggctctatt cctattgagc tatcaaggat caacaattg	360
	acttatcctg taacatgatg acgggtccaa ttccatcacc cattggcaac	420
	tattgaggct taacttgagc aagaatgac tagttggatt catccctg	480
	atctgagaag tgctatggag attgatttat cctataatca tcttgg	540
	aagaacttgg aatgct	556
5	<210> 43 <211> 683 <212> ADN <213> <i>ERECTA</i> de maíz parcial	
10	<400> 43	
	gacgttggga acctcctctt ttatgcttta tggagagtgg cagcttat	60
	atgaaggctc atccaagaag aacaaacttg actgggtgac tgcctac	120
	gtgcagctca aggcctcgtc tacctcacc atgactgcag cccacgaata	180
	acgtaaaaac aaagaatata ctctcgaca aagattatga ggcgcatctt	240
	gcatcgctaa gagcttatgt gtctcgaaga ctcacacgct aacctac	300
	ttggttacat tgatcctgag tacgcccga cctcccgcct caacgagaag	360
	acagctacgg catcgttctg ctggagctgc tgaccggcaa gaagccagt	420
	gcaatctcca tcaactgatc ctatcgaaga cggcgagcaa cgaggtcat	480
	acccccgact gggagacacc tgcaaggacc tgggcgaggt gaagaagct	540
	cgctcctctg caccaagcgg cagccctcgg accggccgac gatgcacgag	600
	tccttgactg cctggtgaac ccgagaccgc cgccgagcc gcagcagcag	660
	cgcacgcgca ccaccagctg ccg	683
15	<210> 44 <211> 2315 <212> ADN <213> <i>ERECTA</i> de maíz	
20	<400> 44	
	acatgcaagt caacaggtta actggatcga taccaccaga gctaggaaat	60
	ttcattacct agaactgaat gataatcaac ttactgggtc aattccacca	120
	ggctaacagg cttgtttgac ctgaacctg cgaataacca ccttgaagga	180
	acaacctaag ttcattgatg aatctcaata gttcaatgc ttatggcaac	240
	gaaccattcc tcgttcgctg cggaaacttg aaagcatgac ctatttaa	300
	atccataag tggctctatt cctattgagc tatcaaggat caacaattg	360
	acttatcctg taacatgatg acgggtccaa ttccatcacc cattggcaac	420
	tattgaggct taacttgagc aagaatgac tagttggatt catccctg	480

# ES 2 362 903 T3

```

atttgagaag tgtcatggag attgatttat cctataatca tcttggtggt ctgattcctc 540
aagaacttgg aatgctgcaa aacctgatgt tgctaaaact ggaaaacaac aatataactg 600
gcgatgtctc ttctctgatg aactgcttca gcctcaatat cttaaatgtg tcatacaata 660
atttggctgg tgctgtccct actgacaaca acttcacacg gttttcacat gacagctttt 720
taggtaatcc tggactctgt ggatattggc ttggttcttc atgtcgttcc actggccacc 780
gagacaaacc gccaatctca aaggctgcca taattggtgt tgctgtgggt ggacttgta 840
tcctcctgat gatcttagta gctgtatgca ggccacacca tccacctgct tttaaagatg 900
ccactgtaag caagccagtg agcaatggtc cacccaagct ggtgatcctt catatgaaca 960
tggctcttca tgtctttgat gatataatga ggatgactga gaacttgagt gaaaaataca 1020
tcattggata cggggcatca agtactgttt ataatgtgt tctaaagaat tgcaaaccag 1080
tggcaataaa aaagctgtat gccactacc tgacagcct taaggaattt gaaactgagc 1140
tcgagactgt tggtagcatc aaacaccgga atctagtcat cctgcaaggg tactcgttgt 1200
cacctgttgg gaacctcctc ttttatgctt atatggagag tggcagctta tgggatgttt 1260
tacatgaagg ctcatccaag aagaacaaac ttgactgggt gactcgctta cggatcgctc 1320
ttggtgcagc tcaaggcctc gcttaccttc accatgactg cagcccacga ataattcacc 1380
gggacgtaaa atcaaagaat atactcctcg acaaagatta tgaggcgcac cttacagact 1440
tcggcatcgc taagagctta tgtgtctcga agactcacac gtcaacctac gtcattggca 1500
5 ctattggta cattgatcct gactacgcc gcacctccg cctcaacgag aagtctgatg 1560
tctacagcta cggcatcgtt ctgtgggagc tgctgaccgg caagaagcca tgggacaacg 1620
agtcaatct ccactacttg atcctatcga agacggcgag caacgaggtc atggagacgg 1680
tggaccccg cgtgggagac acctgcaagg acctgggcca ggtgaagaag ctgttccagc 1740
tggcgtcctc ctgcaccaag cggcagccct cggaccggcc gacgatgcac gagtggtgac 1800
gctccttga ctgcctgtg aaccggagc cgcggccgca gccgagcag cagcagcaga 1860
aggcgcacgc gcaccaccag ctgccgccc agccgtcgcc gccggcctac gtcgacgagt 1920
acgtcagcct cgggggaccc ggcgccctct cctgcgcaa ctgctccagc acctcggagc 1980
ccgagctgtt cctcaagttc ggcgaggcca tctcgagaa catggtgtag ggaagacgt 2040
agacgttcgg tgaggccct tgagtgaggc cgattgcagg gagtagtttg actgacattt 2100
tgtgggacgc agcgcaggag attaacatgg gactcagtag ctagggtgtt gttagctgta 2160
aaaagtcatg tgacgcaacg caagagcagc ggagcttctt cctcatcttc tttatcccc 2220
ctctccattt ttctttggtg gtctaactta ctaggaggct gtattgatcc atcatcatct 2280
ctcccgttcc tcttccttat gatcttgtga ctttc 2315

```

10

```

<210> 45
<211> 675
<212> PRT
<213> ERECTA de maíz

```

15



<400> 45

Met Gln Val Asn Arg Leu Thr Gly Ser Ile Pro Pro Glu Leu Gly Asn  
 1                    5                    10                    15

Met Ser Thr Leu His Tyr Leu Glu Leu Asn Asp Asn Gln Leu Thr Gly  
                   20                    25                    30

Ser Ile Pro Pro Glu Leu Gly Arg Leu Thr Gly Leu Phe Asp Leu Asn  
                   35                    40                    45

Leu Ala Asn Asn His Leu Glu Gly Pro Ile Pro Asp Asn Leu Ser Ser  
                   50                    55                    60

Cys Val Asn Leu Asn Ser Phe Asn Ala Tyr Gly Asn Lys Leu Asn Gly  
 65                    70                    75                    80  
 Thr Ile Pro Arg Ser Leu Arg Lys Leu Glu Ser Met Thr Tyr Leu Asn  
                   85                    90                    95

Leu Ser Ser Asn Phe Ile Ser Gly Ser Ile Pro Ile Glu Leu Ser Arg  
                   100                    105                    110

5

Ile Asn Asn Leu Asp Thr Leu Asp Leu Ser Cys Asn Met Met Thr Gly  
                   115                    120                    125  
 Pro Ile Pro Ser Ser Ile Gly Asn Leu Glu His Leu Leu Arg Leu Asn  
                   130                    135                    140

Leu Ser Lys Asn Asp Leu Val Gly Phe Ile Pro Ala Glu Phe Gly Asn  
 145                    150                    155                    160

Leu Arg Ser Val Met Glu Ile Asp Leu Ser Tyr Asn His Leu Gly Gly  
                   165                    170                    175  
 Leu Ile Pro Gln Glu Leu Gly Met Leu Gln Asn Leu Met Leu Leu Lys  
                   180                    185                    190

Leu Glu Asn Asn Asn Ile Thr Gly Asp Val Ser Ser Leu Met Asn Cys  
                   195                    200                    205

Phe Ser Leu Asn Ile Leu Asn Val Ser Tyr Asn Asn Leu Ala Gly Ala  
                   210                    215                    220  
 Val Pro Thr Asp Asn Asn Phe Thr Arg Phe Ser His Asp Ser Phe Leu  
 225                    230                    235                    240

Gly Asn Pro Gly Leu Cys Gly Tyr Trp Leu Gly Ser Ser Cys Arg Ser  
                   245                    250                    255

Thr Gly His Arg Asp Lys Pro Pro Ile Ser Lys Ala Ala Ile Ile Gly  
 260 265 270

Val Ala Val Gly Gly Leu Val Ile Leu Leu Met Ile Leu Val Ala Val  
 275 280 285

Cys Arg Pro His His Pro Pro Ala Phe Lys Asp Ala Thr Val Ser Lys  
 290 295 300  
 Pro Val Ser Asn Gly Pro Pro Lys Leu Val Ile Leu His Met Asn Met  
 305 310 315 320

Ala Leu His Val Phe Asp Asp Ile Met Arg Met Thr Glu Asn Leu Ser  
 325 330 335

Glu Lys Tyr Ile Ile Gly Tyr Gly Ala Ser Ser Thr Val Tyr Lys Cys  
 340 345 350

Val Leu Lys Asn Cys Lys Pro Val Ala Ile Lys Lys Leu Tyr Ala His  
 355 360 365

Tyr Leu Gln Ser Leu Lys Glu Phe Glu Thr Glu Leu Glu Thr Val Gly  
 370 375 380

Ser Ile Lys His Arg Asn Leu Val Ser Leu Gln Gly Tyr Ser Leu Ser  
 385 390 395 400  
 Pro Val Gly Asn Leu Leu Phe Tyr Ala Tyr Met Glu Ser Gly Ser Leu  
 405 410 415

Trp Asp Val Leu His Glu Gly Ser Ser Lys Lys Asn Lys Leu Asp Trp  
 420 425 430

Val Thr Arg Leu Arg Ile Ala Leu Gly Ala Ala Gln Gly Leu Ala Tyr  
 435 440 445

5

Leu His His Asp Cys Ser Pro Arg Ile Ile His Arg Asp Val Lys Ser  
 450 455 460  
 Lys Asn Ile Leu Leu Asp Lys Asp Tyr Glu Ala His Leu Thr Asp Phe  
 465 470 475 480

Gly Ile Ala Lys Ser Leu Cys Val Ser Lys Thr His Thr Ser Thr Tyr  
 485 490 495

Val Met Gly Thr Ile Gly Tyr Ile Asp Pro Glu Tyr Ala Arg Thr Ser  
 500 505 510

Arg Leu Asn Glu Lys Ser Asp Val Tyr Ser Tyr Gly Ile Val Leu Leu  
 515 520 525

Glu Leu Leu Thr Gly Lys Lys Pro Val Asp Asn Glu Cys Asn Leu His  
 530 535 540

His Leu Ile Leu Ser Lys Thr Ala Ser Asn Glu Val Met Glu Thr Val  
 545 550 555 560  
 Asp Pro Asp Val Gly Asp Thr Cys Lys Asp Leu Gly Glu Val Lys Lys  
 565 570 575

Leu Phe Gln Leu Ala Leu Leu Cys Thr Lys Arg Gln Pro Ser Asp Arg  
 580 585 590  
 Pro Thr Met His Glu Val Val Arg Val Leu Asp Cys Leu Val Asn Pro  
 595 600 605

Glu Pro Pro Pro Gln Pro Gln Gln Gln Gln Lys Ala His Ala His  
 610 615 620

His Gln Leu Pro Pro Gln Pro Ser Pro Pro Ala Tyr Val Asp Glu Tyr  
 625 630 635 640  
 Val Ser Leu Arg Gly Thr Gly Ala Leu Ser Cys Ala Asn Ser Ser Ser  
 645 650 655

Thr Ser Asp Ala Glu Leu Phe Leu Lys Phe Gly Glu Ala Ile Ser Gln  
 660 665 670

Asn Met Val  
 675

## REIVINDICACIONES

1. Procedimiento para seleccionar una planta que presenta una eficacia de transpiración mejorada, que comprende detectar un marcador genético para la eficacia de transpiración, comprendiendo dicho marcador un locus *ERECTA* en el genoma de la planta y seleccionar una planta que comprende o expresa el marcador genético, en el que el marcador genético comprende una secuencia de nucleótidos que presenta una identidad de secuencia global de por lo menos 55% con por lo menos 20 nucleótidos de longitud de cualquiera de las SEC ID nº: 1, 3, 5, 7, 9, 11 a 19 ó 21 a 44 o una secuencia complementaria a las mismas.
2. Procedimiento según la reivindicación 1, en el que el marcador genético comprende una secuencia de nucleótidos seleccionada de entre el grupo constituido por:
- (a) una secuencia que presenta una identidad de por lo menos 55% con una secuencia seleccionada de entre el grupo constituido por SEC ID nº: 1, SEC ID nº:3, SEC ID nº:5, SEC ID nº: 7, SEC ID nº: 9, SEC ID nº: 11, SEC ID nº:12, SEC ID nº: 13, SEC ID nº:14, SEC ID nº:15, SEC ID nº: 16, SEC ID nº:17, SEC ID nº:18, SEC ID nº:19, SEC ID nº:21, SEC ID nº:22, SEC ID nº:23, SEC ID nº:24, SEC ID nº:25, SEC ID nº:26, SEC ID nº:27, SEC ID nº:28, SEC ID nº:29, SEC ID nº: 30, SEC ID nº: 31, SEC ID nº:32, SEC ID nº:33, SEC ID nº: 34, SEC ID nº: 35, SEC ID nº:36, SEC ID nº: 37, SEC ID nº:38; SEC ID nº:39, SEC ID nº: 40, SEC ID nº: 41, SEC ID nº:42, SEC ID nº: 43 y SEC ID nº:44;
- (b) una secuencia que codifica para una secuencia de aminoácidos que presenta una identidad de por lo menos 55% con una secuencia de aminoácidos seleccionada de entre el grupo constituido por SEC ID nº: 2, SEC ID nº: 4, SEC ID nº: 6, SEC ID nº: 8, SEC ID nº: 10, SEC ID nº: 12, SEC ID nº: 20 y SEC ID nº: 45; y
- (c) una secuencia complementaria a (a) o (b).
3. Procedimiento según la reivindicación 1 ó 2, en el que la planta se selecciona de entre el grupo constituido por *Arabidopsis thaliana*, arroz, sorgo, trigo y maíz.
4. Procedimiento según cualquiera de las reivindicaciones 1 a 3, que comprende vincular el fenotipo de eficacia de transpiración de la planta a la expresión del marcador en la planta.
5. Procedimiento según cualquiera de las reivindicaciones 1 a 3, que comprende vincular un polimorfismo estructural en el ADN a un fenotipo de eficacia de transpiración en la planta.
6. Procedimiento según la reivindicación 5, en el que el polimorfismo está determinado por un proceso que comprende detectar un polimorfismo de la longitud de los fragmentos de restricción (RFLP), polimorfismo de la longitud de los fragmentos amplificados (AFLP), polimorfismo en la conformación de cadenas sencillas (SSCP) o análisis de microsatélites.
7. Procedimiento según cualquiera de las reivindicaciones 1 a 6, que comprende hibridar una sonda o cebador de por lo menos aproximadamente 20 nucleótidos de longitud de cualquiera de las SEC ID nº: 1, 3, 5, 7, 9, 11 a 19 ó 21 a 44 o una secuencia complementaria a las mismas con el ADN genómico de la planta, y detectar la hibridación utilizando unos medios de detección.
8. Procedimiento según cualquiera de las reivindicaciones 1 a 7, en el que la planta seleccionada presenta una eficacia de transpiración mejorada en comparación con una planta casi isogénica que no comprende o no expresa el marcador genético.
9. Procedimiento según la reivindicación 1, en el que la detección del marcador genético comprende:
- (a) cribar las líneas endogámicas recombinantes o casi isogénicas o mutantes de plantas para segregar los alelos en un locus *ERECTA*; e
- (b) identificar un marcador polimórfico ligado a dicho locus *ERECTA*.
10. Procedimiento para modular la eficacia de transpiración de una planta que comprende introducir un gen *ERECTA* aislado o una variante alélica del mismo o la región que codifica para proteínas del mismo en una planta y seleccionar una planta que presenta una eficacia de transpiración diferente en comparación con una planta casi isogénica que no comprende el gen *ERECTA* o la variante alélica o la región que codifica para proteínas introducidos, en el que el gen *ERECTA* o la variante alélica o la región que codifica para proteínas comprende una secuencia de nucleótidos que presenta una identidad de secuencia global de por lo menos 55% con por lo menos 20 nucleótidos de longitud de cualquiera de las SEC ID nº: 1, 3, 5, 7, 9, 11 a 19 ó 21 a 44 o una secuencia complementaria a las mismas, en el que el gen *ERECTA* o la variante alélica o la región que codifica para proteínas se introduce en la planta mediante un proceso que comprende transformar material vegetal con un constructo génico que comprende el gen o la variante alélica o la región que codifica para proteínas del mismo.

11. Procedimiento según la reivindicación 10, en el que el gen *ERECTA* o la variante alélica o la región que codifica para proteínas comprende una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

5 (c) una secuencia que presenta una identidad de por lo menos 55% con una secuencia seleccionada de entre el grupo constituido por SEC ID nº: 1, SEC ID nº:3, SEC ID nº:5, SEC ID nº: 7, SEC ID nº: 9, SEC ID nº: 11, SEC ID nº:12, SEC ID nº: 13, SEC ID nº:14, SEC ID nº:15, SEC ID nº: 16, SEC ID nº:17, SEC ID nº:18, SEC ID nº:19, SEC ID nº:21, SEC ID nº:22, SEC ID nº:23, SEC ID nº:24, SEC ID nº:25, SEC ID nº:26, SEC ID nº:27, SEC ID nº:28, SEC ID nº:29, SEC ID nº: 30, SEC ID nº: 31, SEC ID nº:32, SEC ID nº:33, SEC ID nº: 34, SEC ID nº: 35, SEC ID nº:36,  
10 SEC ID nº: 37, SEC ID nº:38; SEC ID nº:39, SEC ID nº: 40, SEC ID nº: 41, SEC ID nº:42, SEC ID nº: 43 y SEC ID nº:44; y

(d) una secuencia que codifica para una secuencia de aminoácidos que presenta una identidad de por lo menos 55% con una secuencia de aminoácidos seleccionada de entre el grupo constituido por SEC ID nº: 2, SEC ID nº: 4, SEC ID nº: 6, SEC ID nº: 8, SEC ID nº: 10, SEC ID nº: 12, SEC ID nº: 20 y SEC ID nº: 45.

12. Procedimiento según la reivindicación 10 u 11, en el que la planta se selecciona de entre el grupo constituido por *Arabidopsis thaliana*, arroz, sorgo, trigo y maíz.

20 13. Procedimiento según cualquiera de las reivindicaciones 10 a 12, que comprende además expresar el gen o variante alélica o región que codifica para proteínas introducido en la planta.

14. Procedimiento según cualquiera de las reivindicaciones 10 a 13, en el que la eficacia de transpiración se mejora en la planta.

25 15. Procedimiento según la reivindicación 14, en el que la eficacia de transpiración se mejora como consecuencia de la expresión ectópica de un alelo *ERECTA* o la región que codifica para proteínas del mismo en la planta.

30 16. Procedimiento según cualquiera de las reivindicaciones 10 a 13, en el que la eficacia de transpiración se reduce en la planta.

17. Procedimiento según la reivindicación 16, en el que la eficacia de transpiración se reduce como consecuencia de la expresión reducida de un alelo *ERECTA* en la planta.

35 18. Utilización de un gen *ERECTA* aislado o variante alélica o región que codifica para proteínas del mismo, que comprende una secuencia de nucleótidos que presenta por lo menos una identidad de secuencia global del 55% con por lo menos 20 nucleótidos de longitud de cualquiera de las SEC ID nº: 1, 3, 5, 7, 9, 11 a 19 ó 21 a 44 o una secuencia complementaria a las mismas, en la preparación de un constructo genético para modular la eficacia de transpiración de una planta.

40 19. Utilización según la reivindicación 18, en el que el gen *ERECTA* o la variante alélica o la región que codifica para proteínas comprende una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

45 (a) una secuencia que presenta una identidad de por lo menos 55% con una secuencia seleccionada de entre el grupo constituido por SEC ID nº: 1, SEC ID nº: 3, SEC ID nº: 5, SEC ID nº: 7, SEC ID nº: 9, SEC ID nº: 11, SEC ID nº: 12, SEC ID nº: 13, SEC ID nº: 14, SEC ID nº: 15, SEC ID nº: 16, SEC ID nº: 17, SEC ID nº: 18, SEC ID nº: 19, SEC ID nº: 21, SEC ID nº: 22, SEC ID nº: 23, SEC ID nº: 24, SEC ID nº: 25, SEC ID nº: 26, SEC ID nº: 27, SEC ID nº: 28, SEC ID nº: 29, SEC ID nº: 30, SEC ID nº: 31, SEC ID nº: 32, SEC ID nº: 33, SEC ID nº: 34, SEC ID nº: 35, SEC ID nº: 36, SEC ID nº: 37, SEC ID nº: 38; SEC ID nº: 39, SEC ID nº: 40, SEC ID nº: 41, SEC ID nº: 42, SEC ID nº: 43 y SEC ID nº: 44; y

50 (b) una secuencia que codifica para una secuencia de aminoácidos que presenta una identidad de por lo menos 55% con una secuencia de aminoácidos seleccionada de entre el grupo constituido por SEC ID nº: 2, SEC ID nº: 4, SEC ID nº: 6, SEC ID nº: 8, SEC ID nº: 10, SEC ID nº: 12, SEC ID nº: 20 y SEC ID nº: 45.

55 20. Procedimiento para aumentar la resistencia de una planta a un estrés medioambiental que comprende mejorar el nivel de expresión de un gen *ERECTA* o una variante alélica del mismo o una región que codifica para proteínas del mismo en dicha planta, en el que el estrés medioambiental se selecciona de entre precipitaciones elevadas, precipitaciones reducidas o sequía, y en el que el gen *ERECTA* o la variante alélica del mismo o la región que  
60 codifica para proteínas del mismo comprende una secuencia de nucleótidos que presenta una identidad de secuencia global de por lo menos 55% con por lo menos 20 nucleótidos de longitud de cualquiera de las SEC ID nº: 1, 3, 5, 7, 9, 11 a 19 ó 21 a 44 o una secuencia complementaria a las mismas.

65 21. Procedimiento para aumentar el peso de semilla o de grano en una planta que comprende mejorar el nivel de expresión en dicha planta de un gen *ERECTA* o variante alélica del mismo o región que codifica para proteínas del mismo que comprende una secuencia de nucleótidos que presenta una identidad de secuencia global de por lo

menos 55% con por lo menos 20 nucleótidos de longitud de cualquiera de las SEC ID nº: 1, 3, 5, 7, 9, 11 a 19 ó 21 a 44 o una secuencia complementaria a las mismas.

5 22. Procedimiento según la reivindicación 20 ó 21, en el que el nivel de expresión se mejora introduciendo un gen *ERECTA* o una variante alélica del mismo o la región que codifica para proteínas del mismo en una planta

23. Procedimiento según la reivindicación 22, en el que el gen *ERECTA* o la variante alélica o región que codifica para proteínas comprende una secuencia de nucleótidos seleccionada de entre el grupo constituido por:

10 (a) una secuencia que presenta una identidad de por lo menos 55% con una secuencia seleccionada de entre el grupo constituido por SEC ID nº: 1, SEC ID nº: 3, SEC ID nº: 5, SEC ID nº: 7, SEC ID nº: 9, SEC ID nº: 11, SEC ID nº:12, SEC ID nº: 13, SEC ID nº: 14, SEC ID nº: 15, SEC ID nº: 16, SEC ID nº: 17, SEC ID nº: 18, SEC ID nº: 19, SEC ID nº: 21, SEC ID nº: 22, SEC ID nº: 23, SEC ID nº: 24, SEC ID nº: 25, SEC ID nº: 26, SEC ID nº: 27, SEC ID nº: 28, SEC ID nº: 29, SEC ID nº: 30, SEC ID nº: 31, SEC ID nº: 32, SEC ID nº: 33, SEC ID nº: 34, SEC ID nº: 35, SEC ID nº:36, SEC ID nº: 37, SEC ID nº: 38; SEC ID nº: 39, SEC ID nº: 40, SEC ID nº: 41, SEC ID nº: 42, SEC ID nº: 43 y SEC ID nº:44; y

15  
20 (b) una secuencia que codifica para una secuencia de aminoácidos que presenta una identidad de por lo menos 55% con una secuencia de aminoácidos seleccionada de entre el grupo constituido por SEC ID nº: 2, SEC ID nº: 4, SEC ID nº: 6, SEC ID nº: 8, SEC ID nº: 10, SEC ID nº: 12, SEC ID nº: 20 y SEC ID nº: 45.

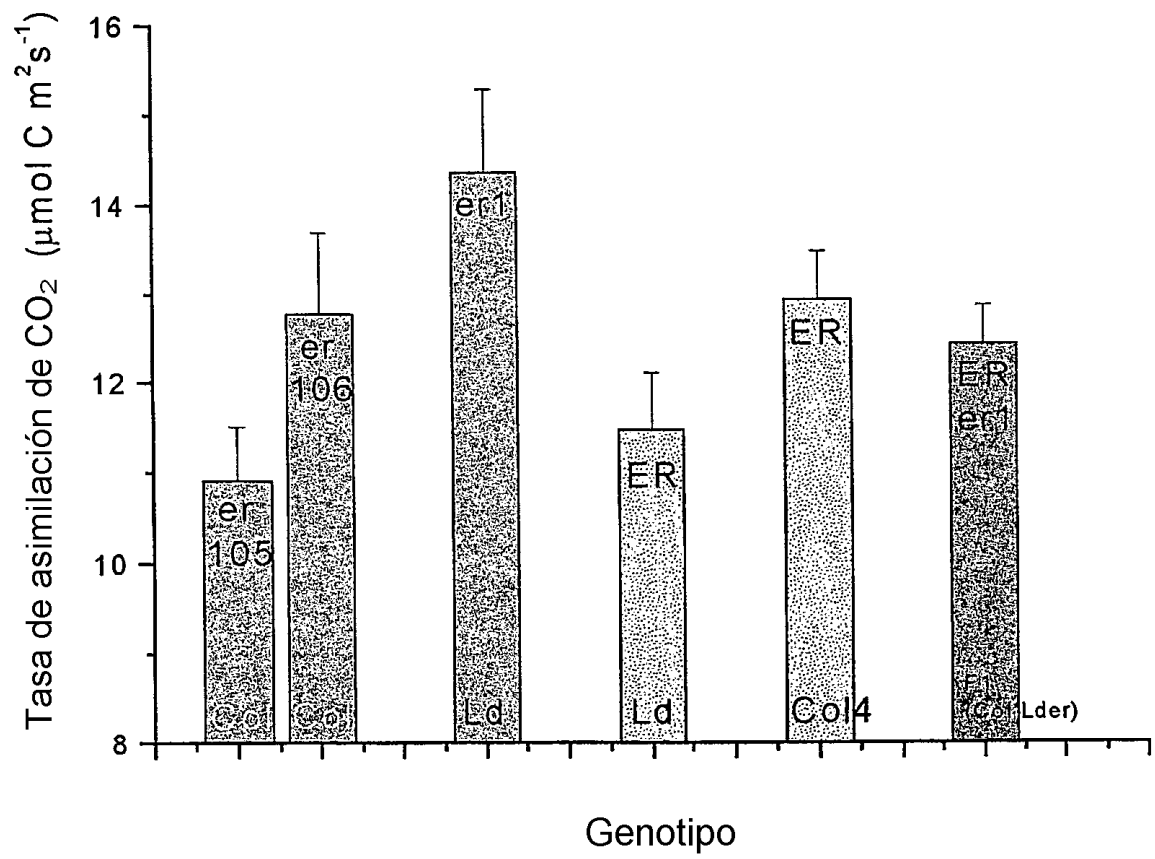


Figura 1a

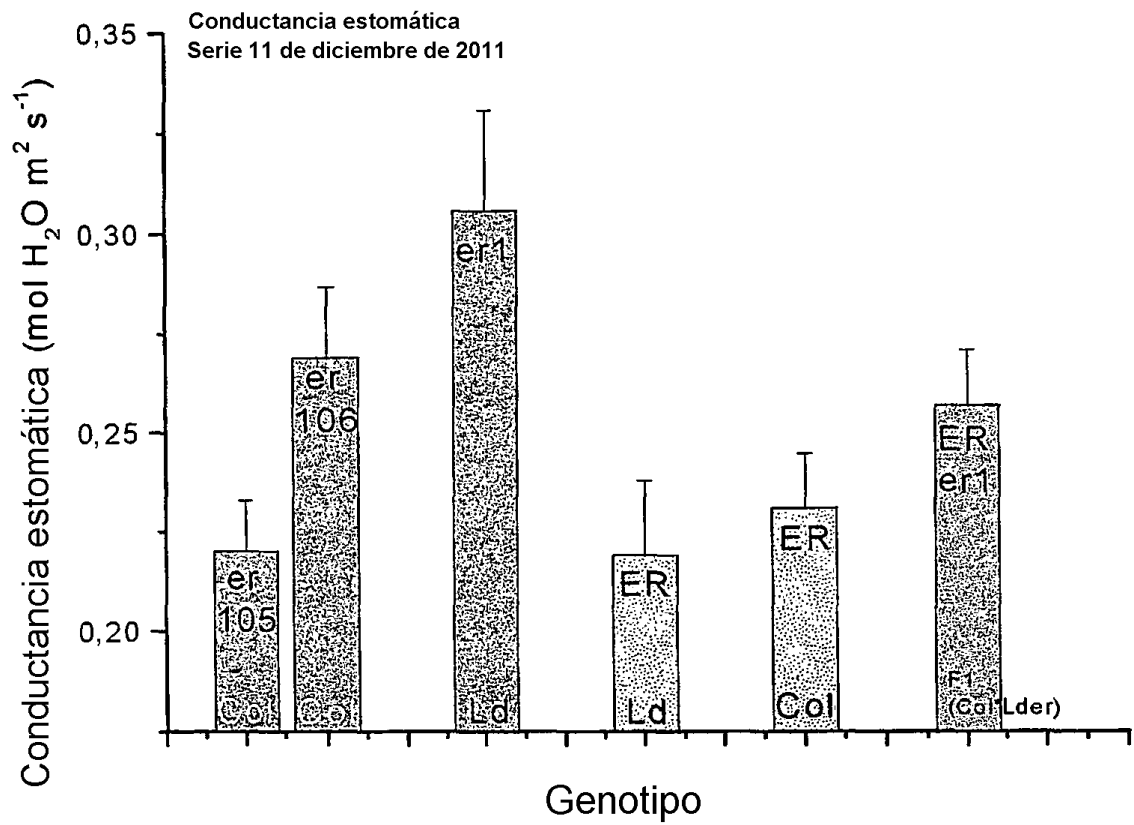


Figura 1b



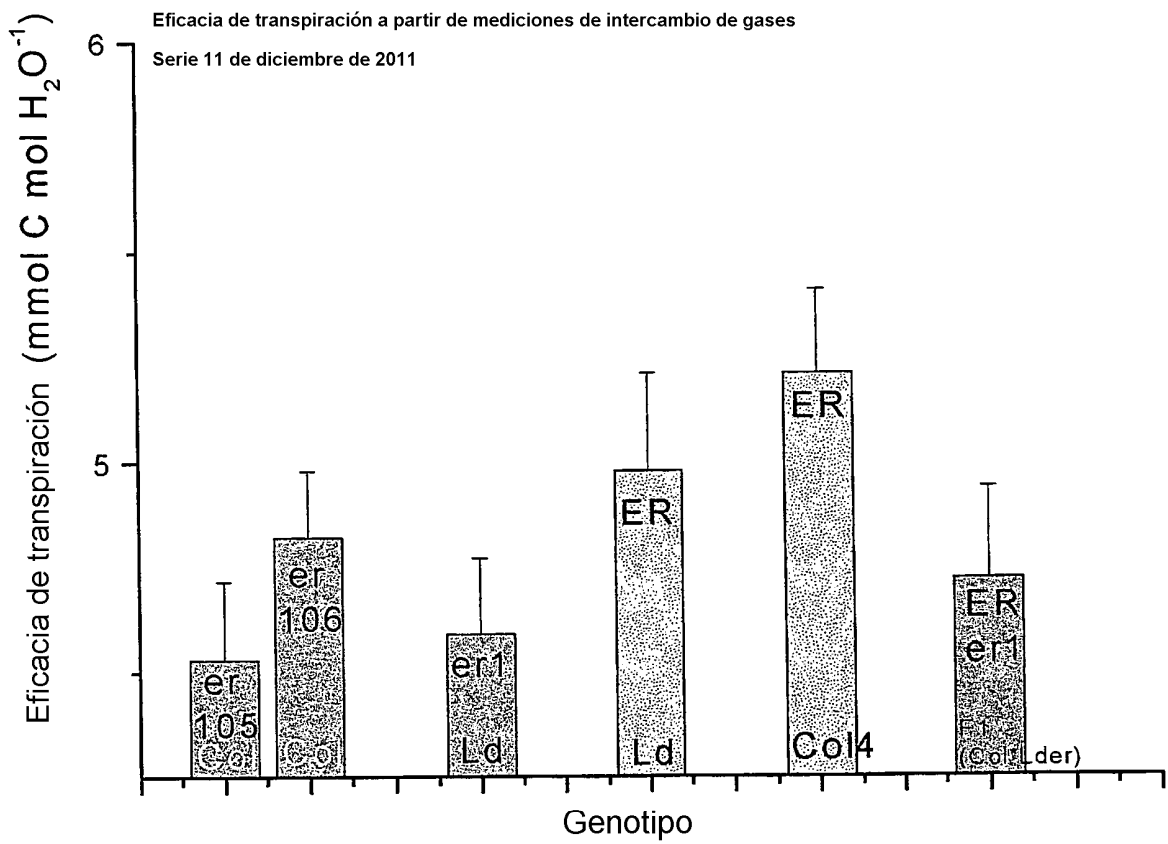


Figura 1c

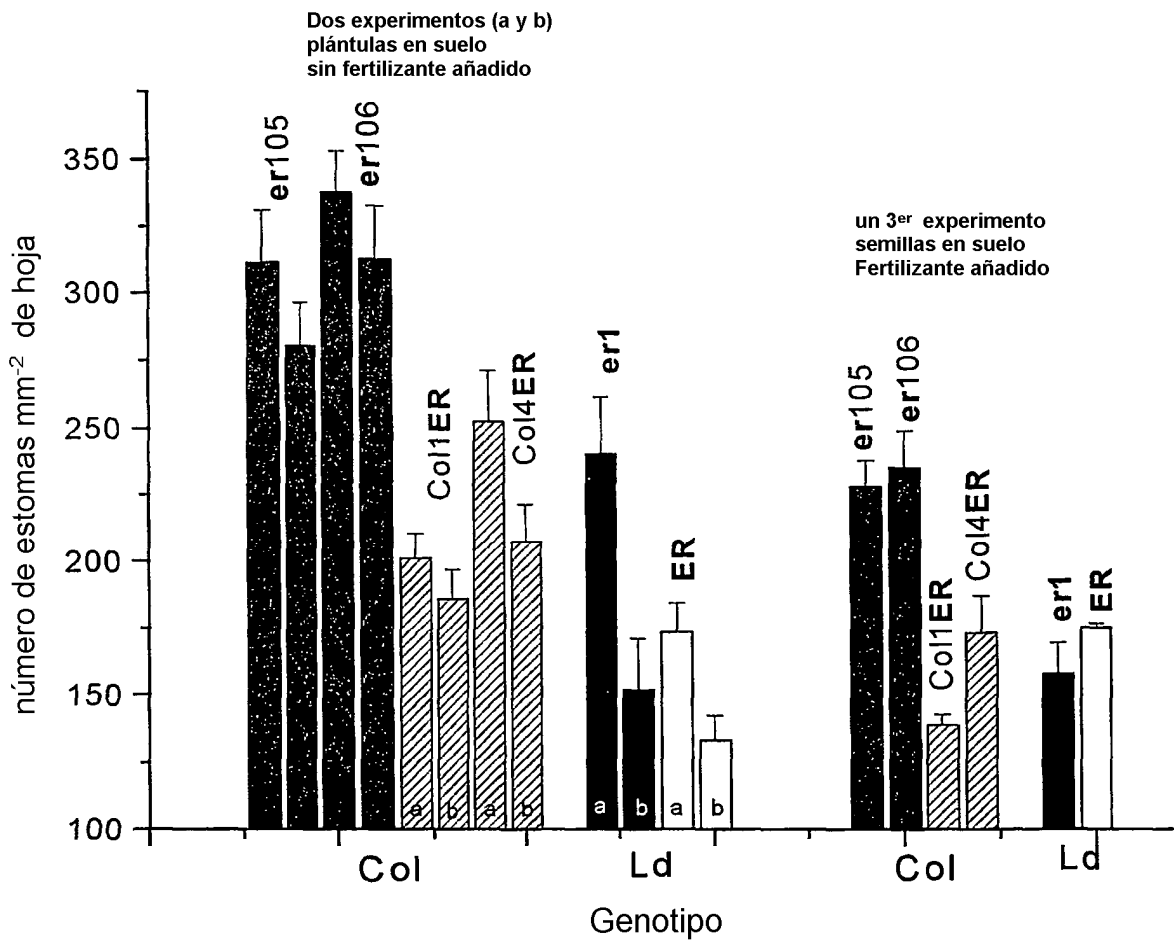


Figura 2a

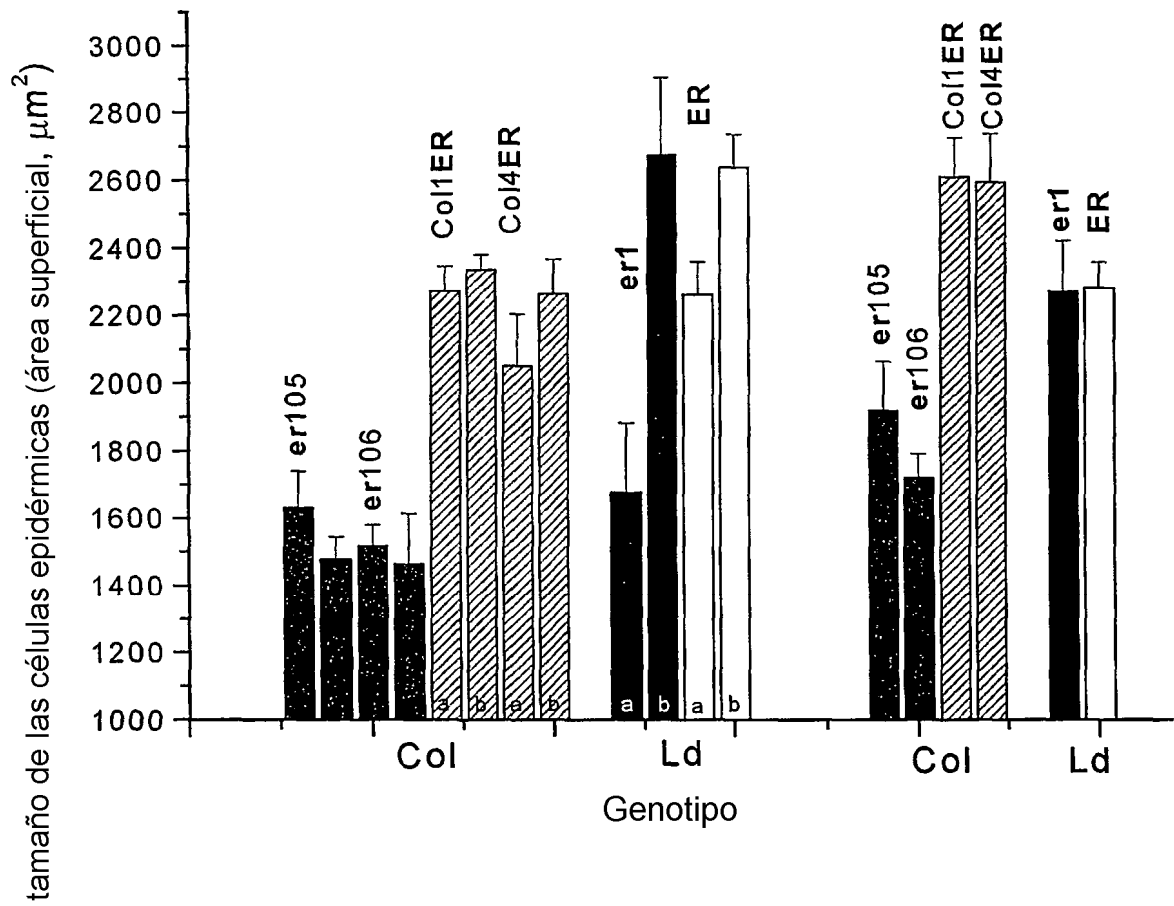


Figura 2b

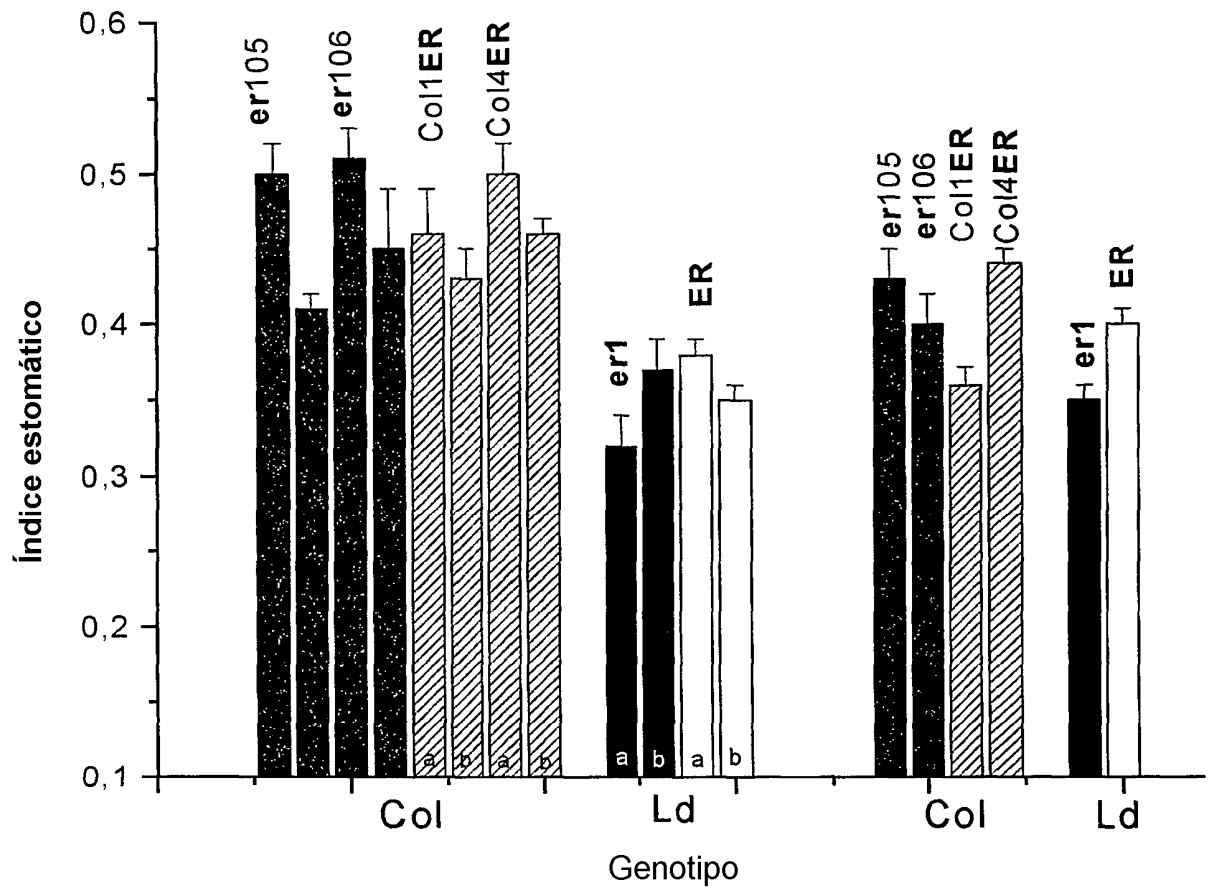


Figura 2c

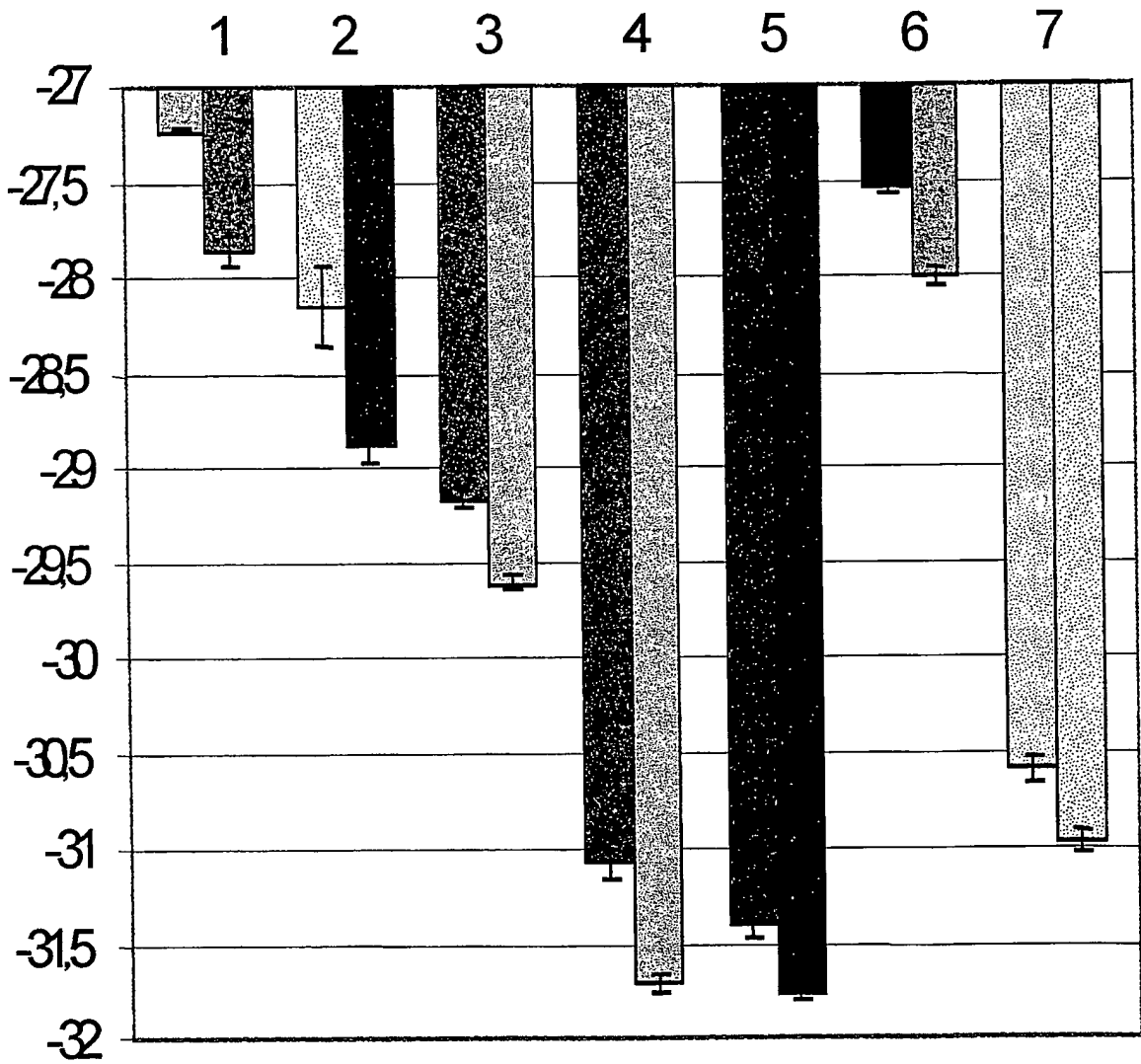


Figura 3

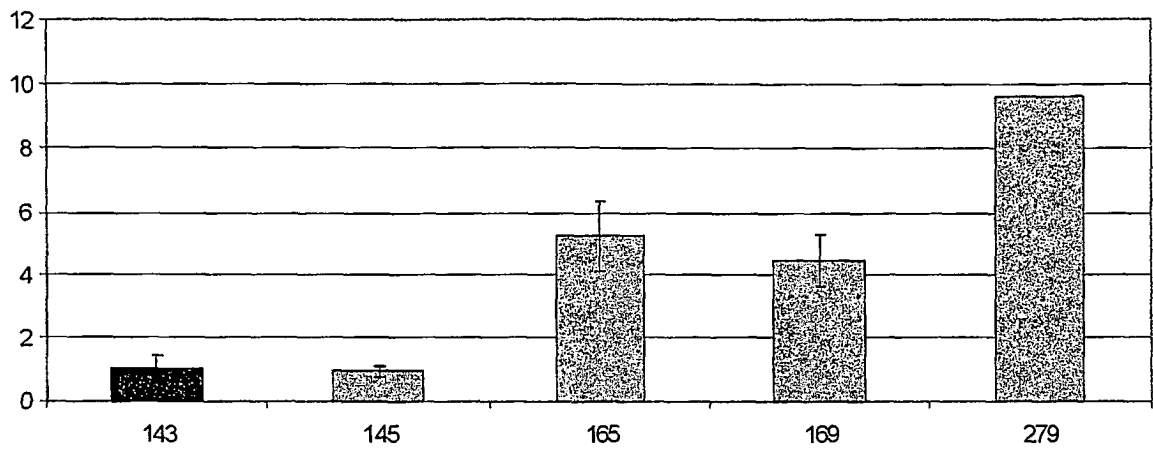


Figura 4a

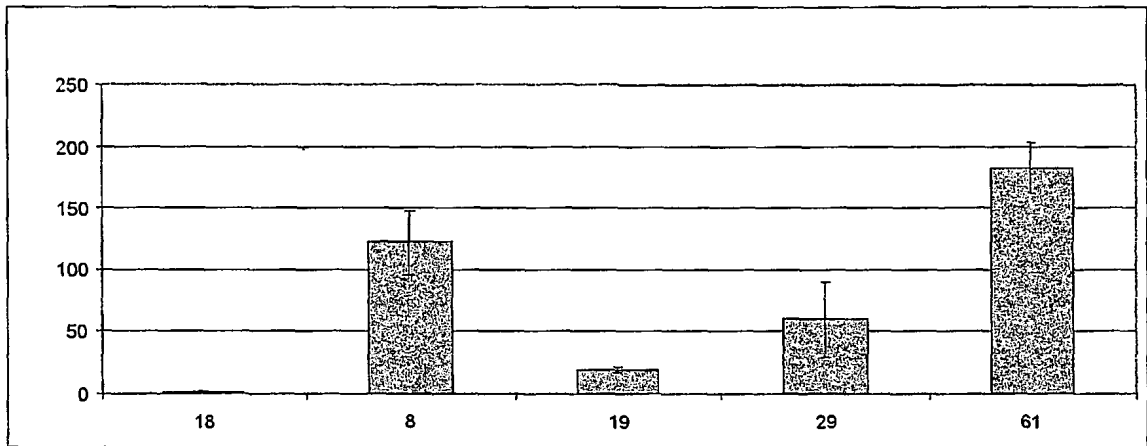


Figura 4b

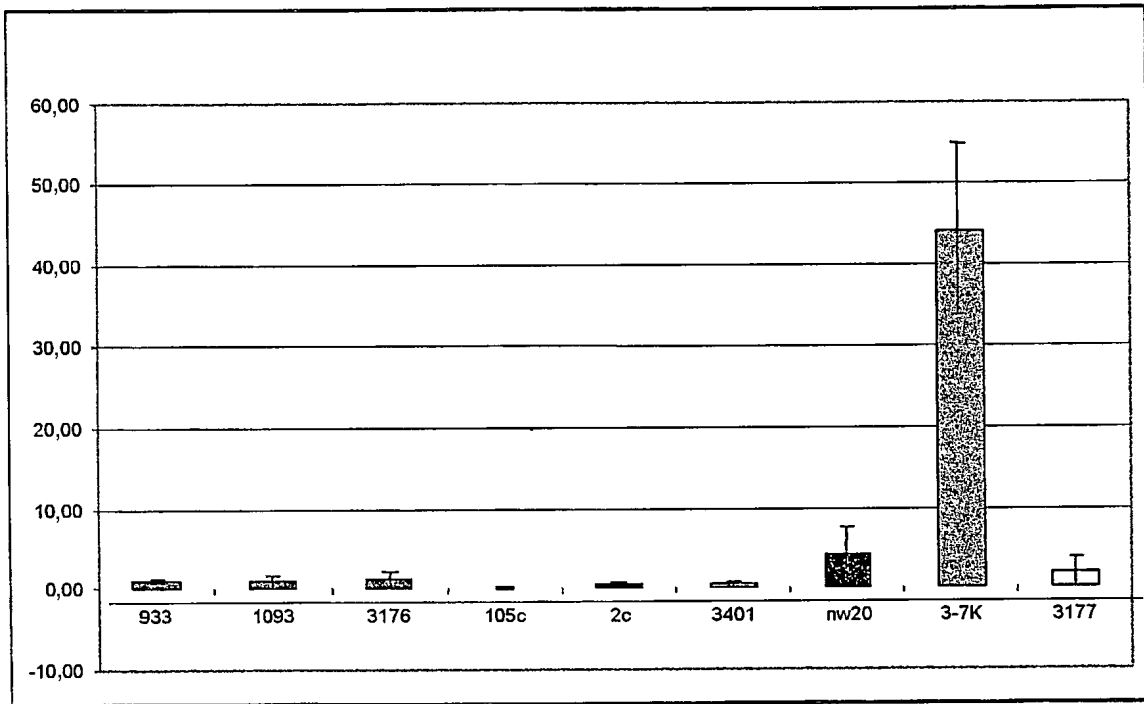


Figura 4c



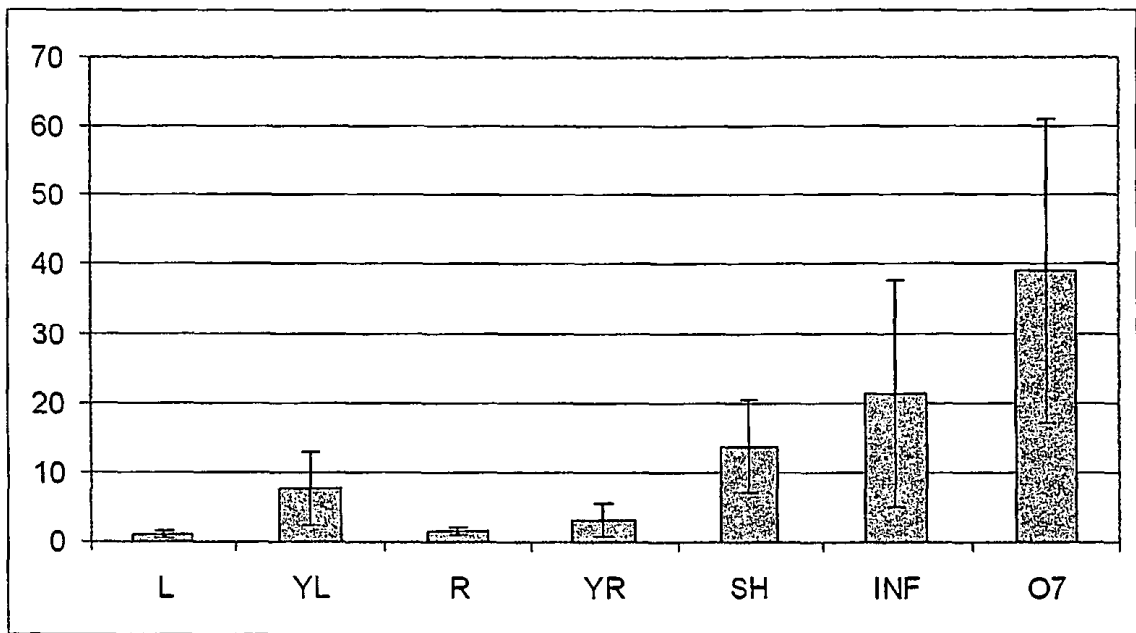


Figura 5a

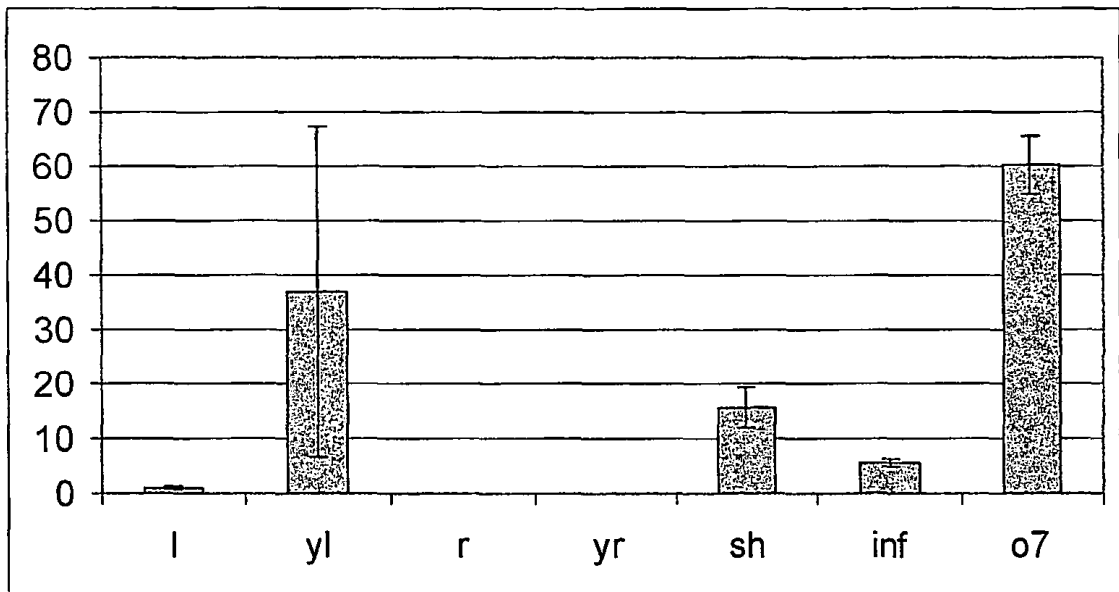


Figura 5b

Figura 6a: Eficacia de transpiración a 350 y 500ppm de CO<sub>2</sub> (mmolC mol H<sub>2</sub>O<sup>-1</sup>)

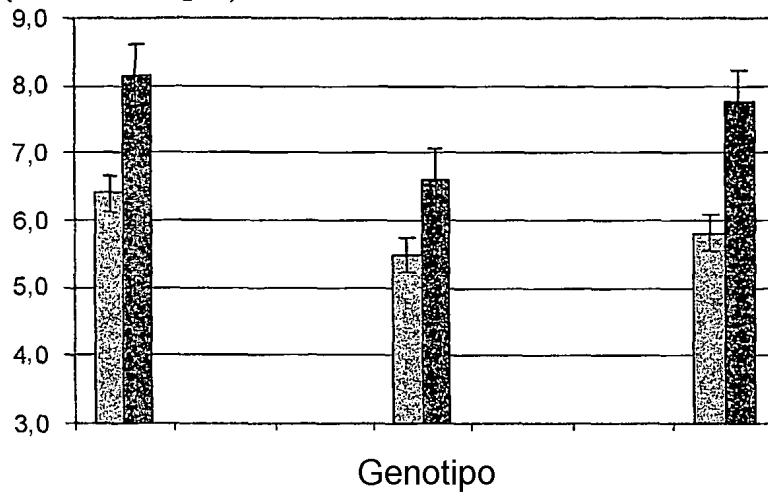


Figura 6b: Tasa de asimilación de CO<sub>2</sub> (μmol C m<sup>-2</sup>s<sup>-1</sup>)

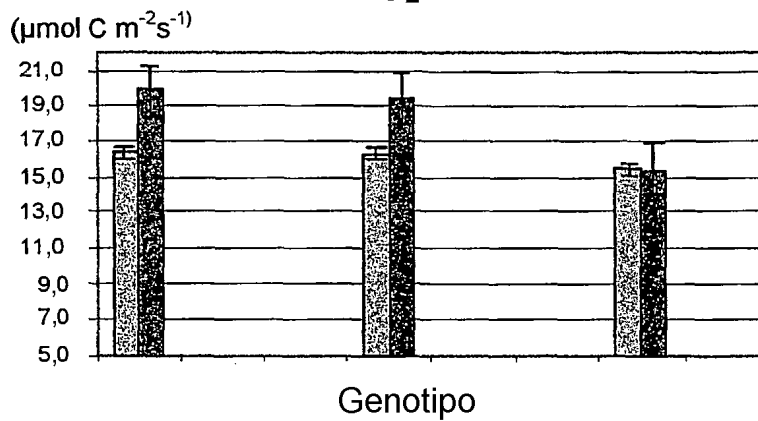
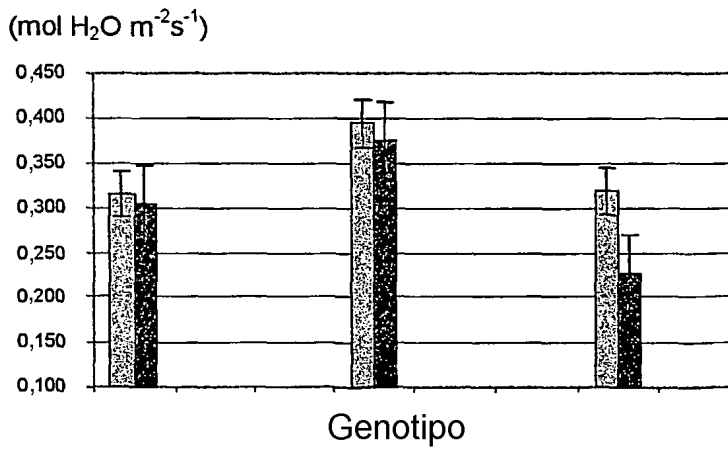


Figura 6c: Conductancia estomática (mol H<sub>2</sub>O m<sup>-2</sup>s<sup>-1</sup>)



T2(+ER)  
Ld\_ER

Ld -er1

Ld\_ER  
(línea 3177)

Figura 7a: Eficacia de transpiración ( $\text{mmol C mol H}_2\text{O}^{-1}$ ) a 350 y 500ppm  $\text{CO}_2$

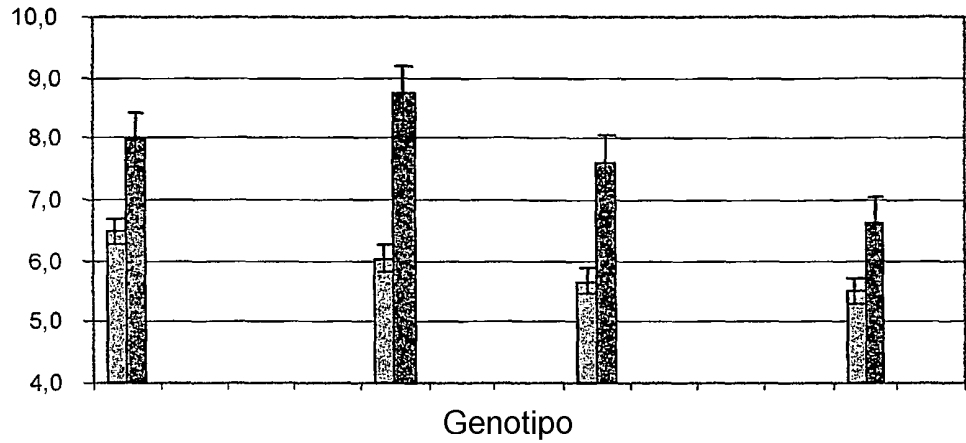


Figura 7b: Tasa de asimilación de  $\text{CO}_2$  ( $\mu\text{mol C m}^{-2}\text{s}^{-1}$ )

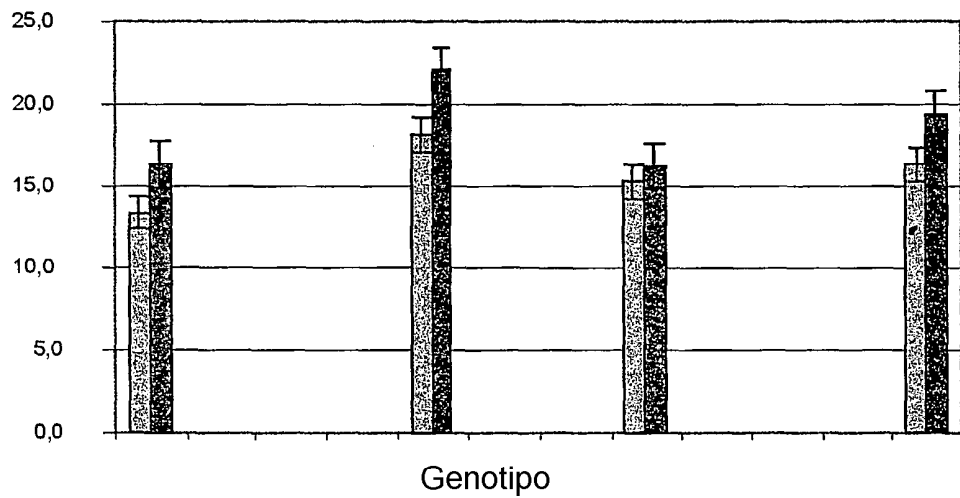


Figura 7c: Conductancia estomática ( $\text{mol H}_2\text{O m}^{-2}\text{s}^{-1}$ )

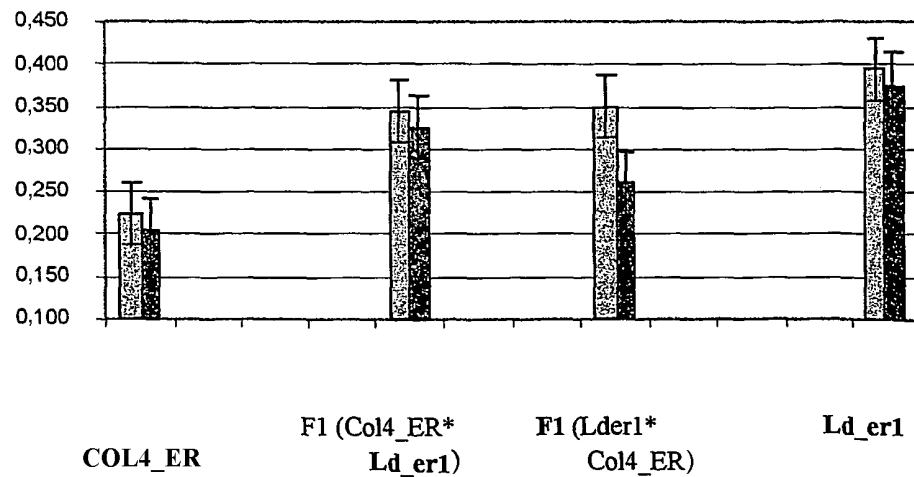


Figura 8a Conductancia estomática ( $\text{mol H}_2\text{O m}^{-2} \text{s}^{-1}$ )

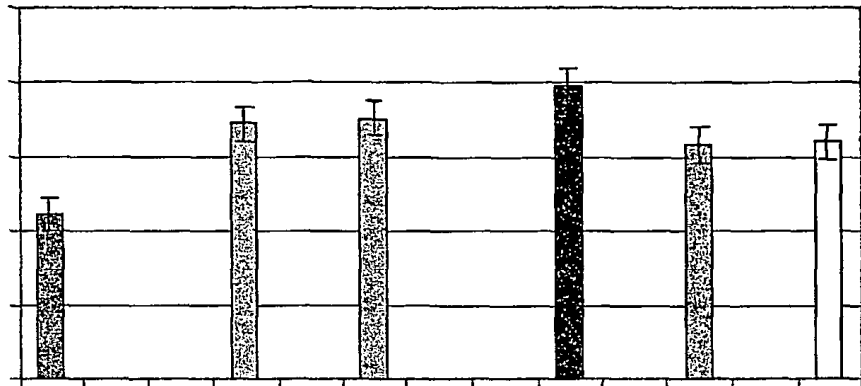


Figura 8b Área de células epidérmicas ( $\mu\text{m}^2$ )

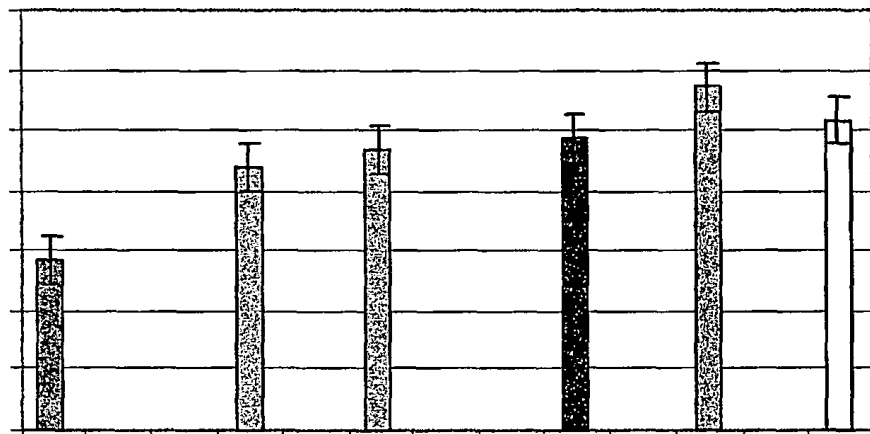


Figura 8c Número de estomas por  $\text{mm}^2$

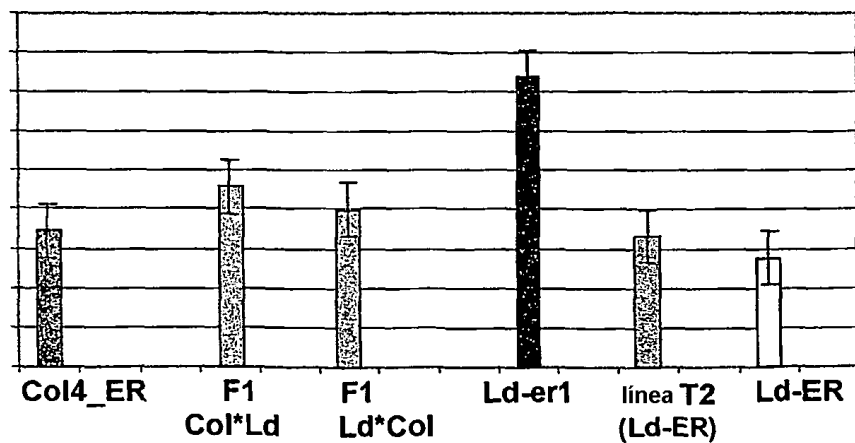


Figura 9a

Número de estomas por mm<sup>2</sup>

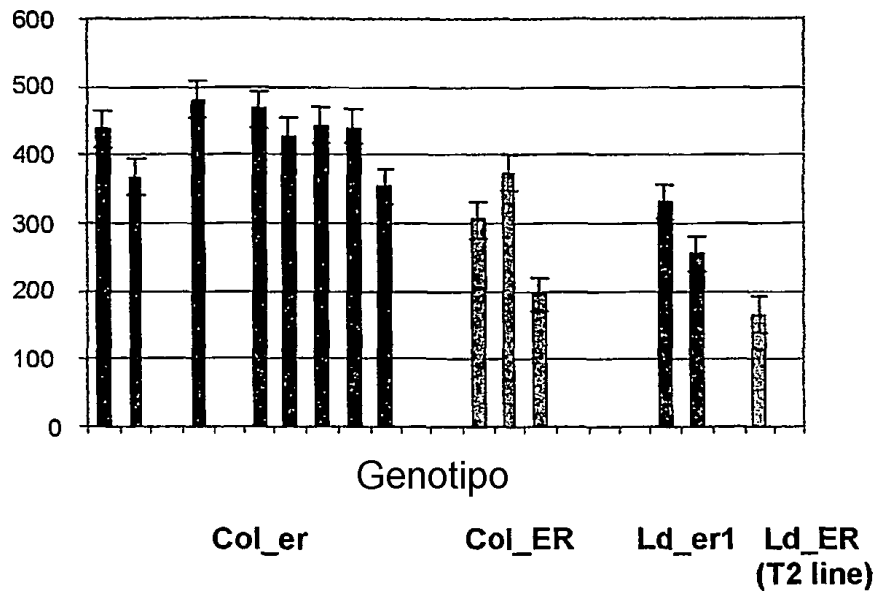
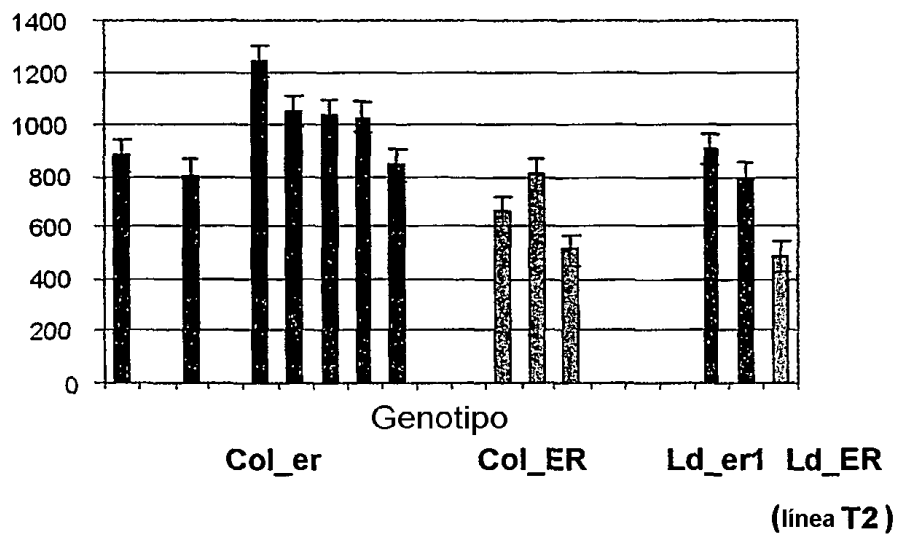


Figura 9b

Área de células epidérmicas (µm<sup>2</sup>)



**Composición de isótopos de carbono (por mil)  
para una gama de mutantes er y ecotipos ER o  
líneas transgénicas**

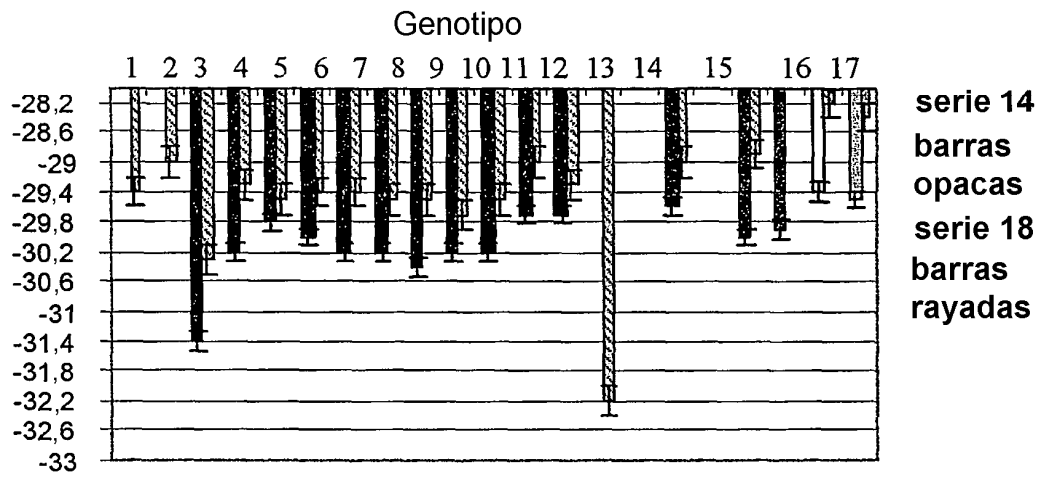


Figura 10

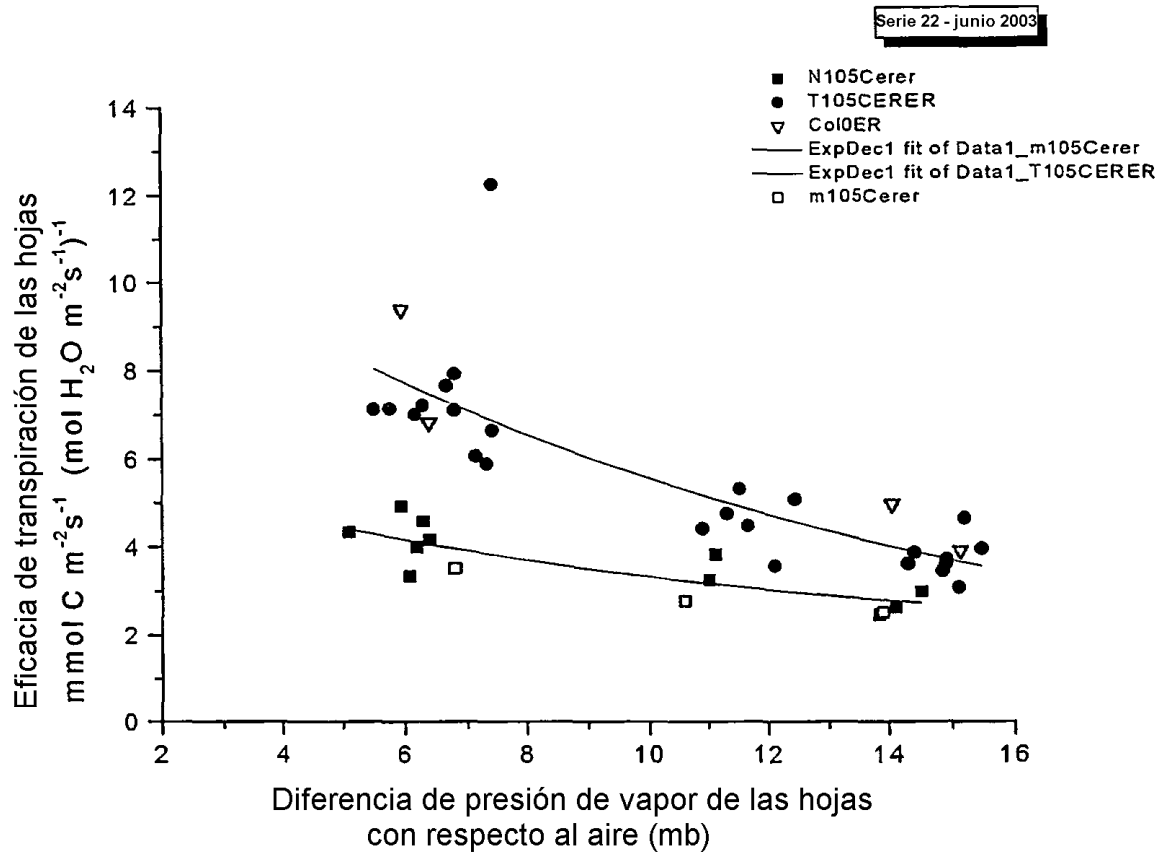


Figura 11



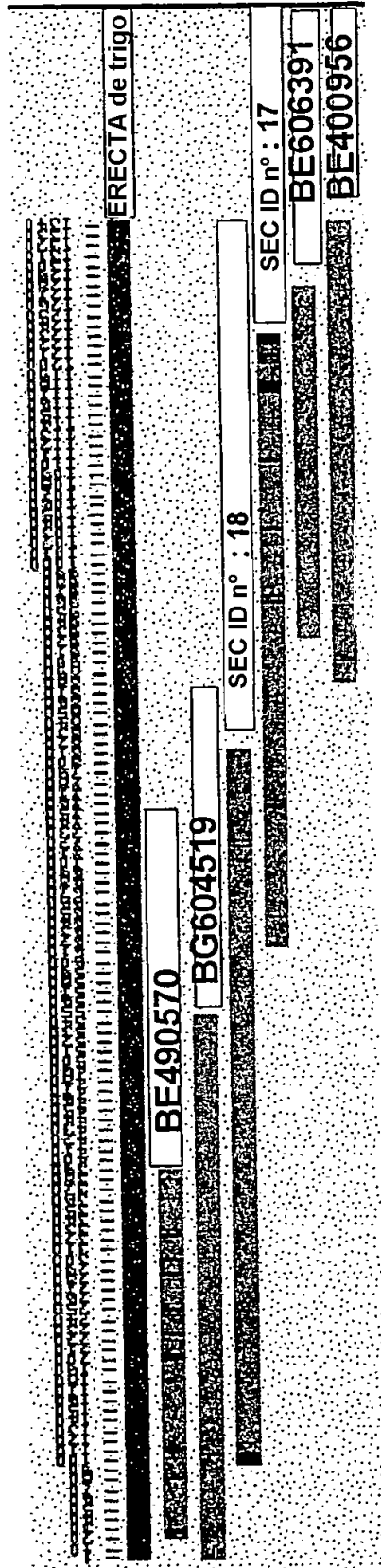


Figura 12

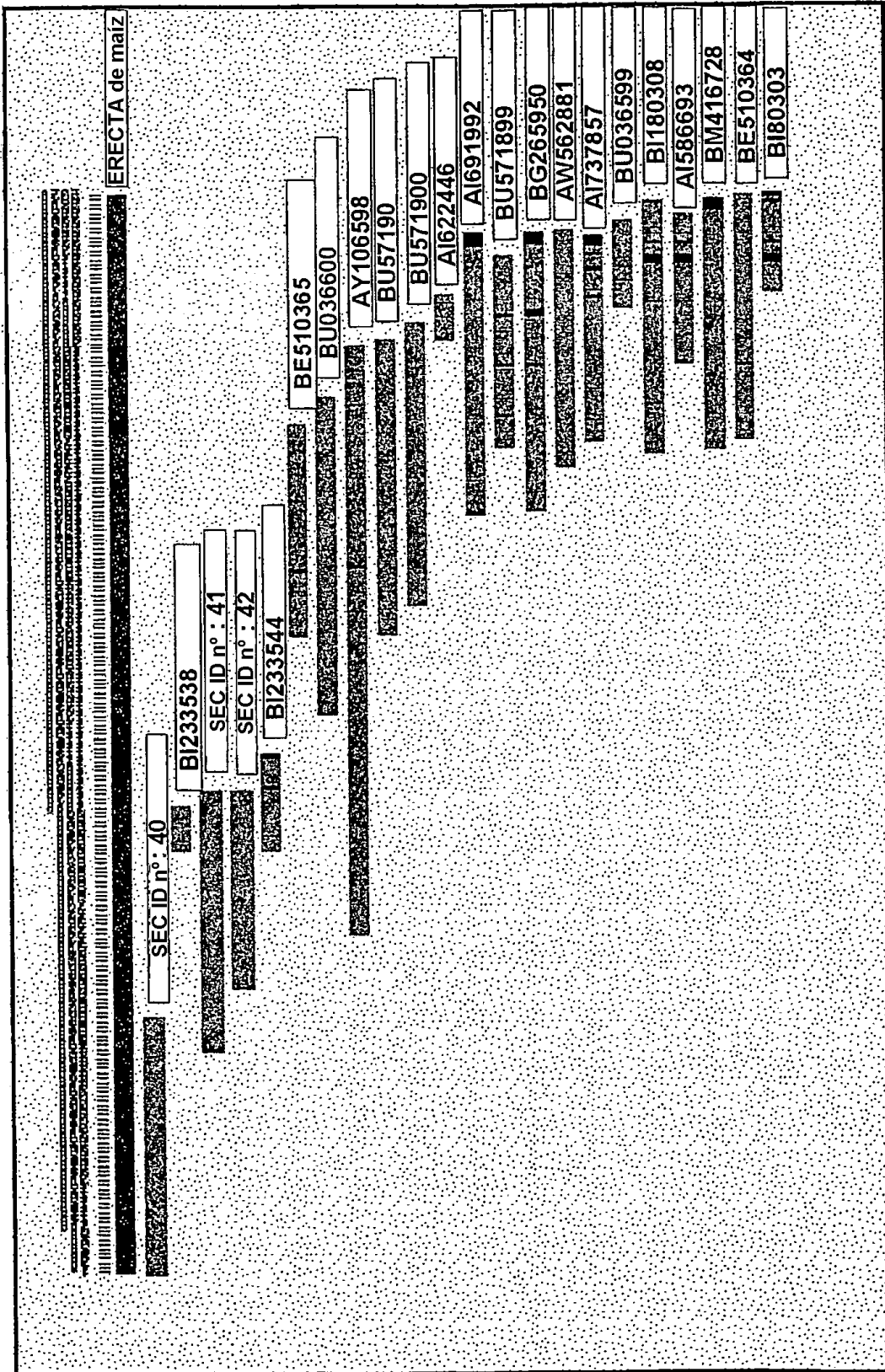


Figura 12

```

Erecta_sorgo          -----MTTAAARALVALLLVAVAVA-----DDGATLVEIKKSERIVGNLYDHWAG--
Erecta_mai3          MALFRDIVLGGFLFCLSLVATVYS-----EGGATLLEIKKSEKDKAIVLYDWTTFSP
Erecta_Arabidopsis:  MAARAARWAAWAVVVVGVAVAEASGGGGGGGEGEAKALGVVKAAGFGRANALVWQDC--
Erecta_trigo          -----

Erecta_sorgo          -DDYCSWRGVLCDNVTFAVAALNLSGLLGEIEI SPAVGSILKSLVSDIKLSGLSGQIPDE
Erecta_mai3          SSDYCVWRGVCENVTFFVVALNLSDLHLGDEI SPAIGDLKSLLSIDLRGHLSGQIPEDE
Erecta_Arabidopsis:  GADHCARWGVTCDAVSFAVLAALNLSHLGGDEI SPAIGELKNLQFVLDLKGHLTGQIPEDE
Erecta_trigo          -----

Erecta_sorgo          IGDCSSLATLDFSENNLGDIFFSISKLHLENLIKNNQGLGAI PSTLSQLPNLKIIDL
Erecta_mai3          IGDCSSLQMLDSFNELSGDIFFSISKLQLEQLIKNNQGLGAI PSTLSQIPHLAKIDL
Erecta_Arabidopsis:  IGDCISLKYLDLSGNLLVGDIFFSISKLQLEE LIKNNQGLTGAI PSTLSQIPHLAKIDL
Erecta_trigo          -----

Erecta_sorgo          AQNKLAGEIFRILIYVNEVLQYL-----DVKNNSLGTGVIEDT
Erecta_mai3          AQNKLAGEIFRILIYVNEVLQYLGRLGNLGVNISPDLQQLTGLWYEDVRRNSLGTGVIET
Erecta_Arabidopsis:  AQNQLTGDI FRLIYVNEVLQYLGRLGNSLTGTLSPEMCQLTGLWYEDVRRNSLGTGPES
Erecta_trigo          -----

Erecta_sorgo          IGNCYSFQVLDLSYNRE TGPIEFNIGLQVATLSLQGNKKTGPIPSVIGLHQAIVLDS
Erecta_mai3          IGNCYAFQVLDLSYNQLTGEIFFDIGLQVATLSLQGNLKSGLNIPVIGLHQAIVLDS
Erecta_Arabidopsis:  IGNCYSFEI LDISTYQISGDEI PYMIGLQVATLSLQGNLTKNIPVIGLHQAIVLDS
Erecta_trigo          -----

Erecta_sorgo          YNQLSGPISILGNLTYTEKLYIQGNKLTGSIPELQGSMTIHYLELNQNLGTSIPEEL
Erecta_mai3          YNQLSGPISILGNLTYTEKLYIQGNKLTGSIPELQGSMTIHYLELNQNLGTSIPEEL
Erecta_Arabidopsis:  GNL LSGSIPPI LGNLTYTEKLYLHNSKLTGSIPELQGNLKSGLNIPVIGLHQAIVLDS
Erecta_trigo          ENE LVGPIPSILGNLSTYTKLYLHNSKLTGSIPELQGNLKSGLNIPVIGLHQAIVLDS

Erecta_sorgo          GRLTGFLDLNANNHLEGPIPNLSSCVNINSENAVGNKLANGTIPRSIKLESMTYLNLS
Erecta_mai3          GRLTGFLDLNANNHLEGPIPNLSSCVNINSENAVGNKLANGTIPRSIKLESMTYLNLS
Erecta_Arabidopsis:  GRLTDLFLDLNANNHLEGPIPNLSSCVNINSENAVGNKLANGTIPRSIKLESMTYLNLS
Erecta_trigo          GKLEELFELANLANNLGGPI PANISSCTALNKENVYGNLKSGLNIPVIGLHQAIVLDS
                             -----FNHVNRLNLSIPAGEQLLES LTNHLS
                             :*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

Erecta_sorgo          SHEISGSIPIELSRINNLDLTLSCNMTGPISSIGSLEHLRLNLSKNGLVGSIPEAF
Erecta_mai3          SNEISGSIPIELSRINNLDLTLSCNMTGPISSIGSLEHLRLNLSKNDLVGSIPEAF
Erecta_Arabidopsis:  SNNIKGPIVEISRIGNLDTLTLSCNNGKGLPSSLDLEHLRLNLSKNNHIVGVPGDF
Erecta_trigo          SNNFKHPISELGHI INLDLTLSYNEFSGEPATIGDLLEHLRLNLSKNNHLOGVPPEAF
                             SNNFKHPISELGHI INLDLTLSYNEFSGEPATIGDLLEHLRLNLSKNNHLOGVPPEAF
                             ** * : * * * * : * * * * * * : * * : * : * * : * * * * : * * : *

Erecta_sorgo          GNLRSVMEIDLSTYNHGLGGLIQELEMQLMILL
Erecta_mai3          GNLRSVMEIDLSTYNHGLGGLIQELEMQLMILLKLENNITGDV-SSLAHNCFSLAH---IL
Erecta_Arabidopsis:  GNLRSVMEIDLSTYNHGLGGLIQELEMQLMILLKLENNITGDV-SSLAHNCFSLAH---VL
Erecta_trigo          GNLRSVQVTIMSHRNLSSGLPELQGLQMLDLSLLIRNNVIGI EFAQLRKFSLANLAFQ
                             GNLRSVQVTIMSHRNLSSGLPELQGLQMLDLSLLIRNNVIGI EFAQLRKFSLANLAFQ
                             -----IL
                             ***** : * * : * : * * : * * : * : * * * * : * : * * * * :

Erecta_sorgo          NVSYNLAGVPAEINNE TRFSEPD-----SFLGNEPLGCGYWLSSCRSTG
Erecta_mai3          NVSYNLAGVPAEINNE TRFSEPD-----SFLGNEPLGCGYWLSSCRSTG
Erecta_Arabidopsis:  NVSHNLVGDIPKNNNE SRFSEPD-----SFLGNEPLGCGYWLSSPCDHR
Erecta_trigo          EFVIQGFITWC PDGKLELEIPNGHLLS DCCNYYINHCFS LGHPLHVYCDSSCGHS-
                             NLSHNNESHVFEARNEKFE PGE-----SFLGNEPLSHVCKDSSCGNS-
                             1. : : : * : : : . : * * * * * * * * : * * :

Erecta_sorgo          HHEKFPISKAATIGVAVGGVILMLLWAVCRPHRPAE KDVTVSKEVRNAPPRKLVILHM
Erecta_mai3          HRDKEFPISKAATIGVAVGGVILMLLWAVCRPHRPAE KDATVSKEVRSHGPKRVLILHM
Erecta_Arabidopsis:  RTVRSISRAATIGVAVGGVILMLLWAVCRPHRPAE LDGSLDRFVYVSTPELILHM
Erecta_trigo          HSGKVNISKTAIACILGSLLELILLGILLKRYTRKRPQPLVKS--DKVQ--GEPKILVLIQM
                             HGSKVNMT-RTAIACISSEVILLGILLKRYTRKRPQPLVKS--DKVQ--GEPKILVLIQM
                             : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *

Erecta_sorgo          NMAHVYDDIMRMTENLSEKVIIGYASSTVYKCVLWCKEPAVKKLVAHYI PQLSKEFE
Erecta_mai3          NMAHVYDDIMRMTENLSEKVIIGYASSTVYKCVLWCKEPAVKKLVAHYI PQLSKEFE
Erecta_Arabidopsis:  NMAHVYDDIMRMTENLSEKVIIGYASSTVYKCVLWCKEPAVKKLVAHYI PQLSKEFE
Erecta_trigo          DMAIHTYDDIMRMTENLSEKVIIGYASSTVYKCVLWCKEPAVKKLVAHYI PQLSKEFE
                             DMAIHTYDDIMRMTENLSEKVIIGYASSTVYKCVLWCKEPAVKKLVAHYI PQLSKEFE
                             : * * : * : * * * : * : * * : * : * * : * : * : * : * : * : *

Erecta_sorgo          ELETGVSINRNLVSLQGYSLSPVGNLLEFYDMCEGSLWDLHEGSSSKKSLWETRLRI
Erecta_mai3          ELETGVSINRNLVSLQGYSLSPVGNLLEFYDMCEGSLWDLHEGSSSKKSLWETRLRI
Erecta_Arabidopsis:  ELETGVSINRNLVSLQGYSLSPVGNLLEFYDMCEGSLWDLHEGSSSKKSLWETRLRI
Erecta_trigo          ELETGVSINRNLVSLQGYSLSPVGNLLEFYDMCEGSLWDLHEGSSSKKSLWETRLRI
                             ELETGVSINRNLVSLQGYSLSPVGNLLEFYDMCEGSLWDLHEGSSSKKSLWETRLRI
                             ***** : * : * * * * : * * * * * * : * : * * * * :

Erecta_sorgo          ALGAAQGLAYLHDCSPRI IHRDVKSNKNI LLDKDYEAHLDGFIKSLCVSKHTITSTVM
Erecta_mai3          ALGAAQGLAYLHDCSPRI IHRDVKSNKNI LLDKDYEAHLDGFIKSLCVSKHTITSTVM
Erecta_Arabidopsis:  AVGAAQGLAYLHDCSPRI IHRDVKSNKNI LLDKDYEAHLDGFIKSLCVSKHTITSTVM
Erecta_trigo          AVGAAQGLAYLHDCSPRI IHRDVKSNKNI LLDKDYEAHLDGFIKSLCVSKHTITSTVM
                             AVGAAQGLAYLHDCSPRI IHRDVKSNKNI LLDKDYEAHLDGFIKSLCVSKHTITSTVM
                             ***** : * * : * * * * : * : * * * * :

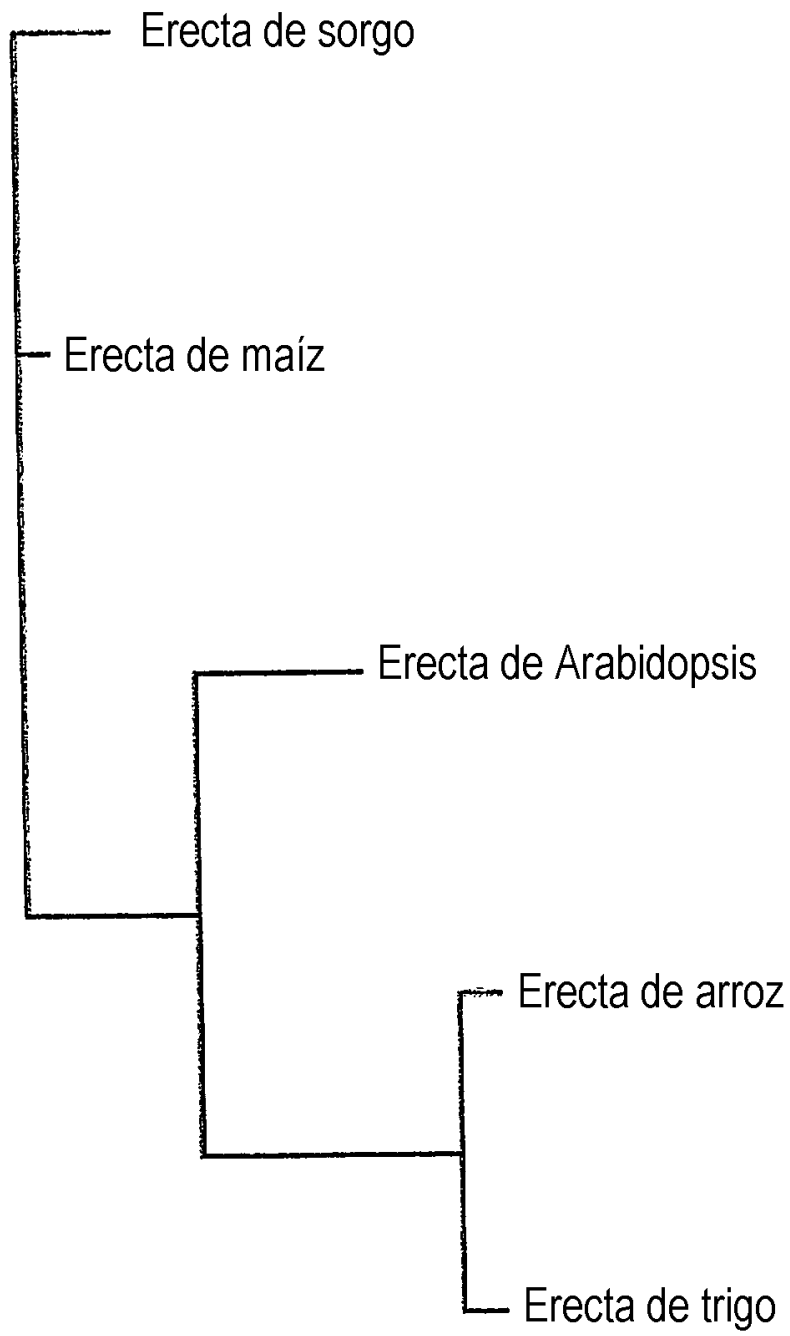
Erecta_sorgo          GTIGYIDFVARTSRLNEKSDVYR----LWHSAG--AADWQASGQRLSKTASNEVND
Erecta_mai3          GTIGYIDFVARTSRLNEKSDVYR----LWHSAG--AADWQASGQRLSKTASNEVND
Erecta_Arabidopsis:  GTIGYIDFVARTSRLNEKSDVYR----LWHSAG--AADWQASGQRLSKTASNEVND
Erecta_trigo          GTIGYIDFVARTSRLNEKSDVYR----LWHSAG--AADWQASGQRLSKTASNEVND
                             GTIGYIDFVARTSRLNEKSDVYR----LWHSAG--AADWQASGQRLSKTASNEVND
                             ***** : * * : * * * * : * : * * * * :

Erecta_sorgo          TVDFDIDGDKDLGKLVLLAALLCTKQPSDRPTHEVVRVLDCLVNEPPEKFP---
Erecta_mai3          TVDFDIDGDKDLGKLVLLAALLCTKQPSDRPTHEVVRVLDCLVNEPPEKFP---
Erecta_Arabidopsis:  MADFDITSDKDLGKLVLLAALLCTKQPSDRPTHEVVRVLDCLVNEPPEKFP---
Erecta_trigo          AVDFEVSFTCTEMGLVRAFAQ LALLCTKQPSDRPTHEVVRVLDCLVNEPPEKFP---
                             AVDFEVSFTCTEMGLVRAFAQ LALLCTKQPSDRPTHEVVRVLDCLVNEPPEKFP---
                             ***** : * : * * * * : * : * * * * :

Erecta_sorgo          ----SAHQLEPQSPAVPSYIIEYVSLRGTGALSCANSTSTDAEILNFGSAIQQNE
Erecta_mai3          QKANAHHQLP--PQSPAVPSYIIEYVSLRGTGALSCANSTSTDAEILNFGSAIQQNE
Erecta_Arabidopsis:  ----DTS--ATLASSTYVDEYRHLKTFHSVICS--SMSASDAQLELRFQVTSQNSG
Erecta_trigo          -----NV--DYSLLASTTADNRGHEVDIIG--DNSSSEDCQVRFGEVSKHTM
                             -----

```

Figura 13



**FIGURA 14**