

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 743 512**

51 Int. Cl.:

C12N 15/82 (2006.01)

C12N 15/10 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **06.03.2015** **E 15158086 (7)**

97 Fecha y número de publicación de la concesión europea: **14.08.2019** **EP 3064587**

54 Título: **Biblioteca de promotores sintéticos para la expresión génica coordinada en células u organismos eucariotas**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
19.02.2020

73 Titular/es:

**LEIBNIZ-INSTITUT FÜR PFLANZENBIOCHEMIE
(IPB) (100.0%)
Weinberg 3
06120 Halle (Saale), DE**

72 Inventor/es:

**MARILLONNET, SYLVESTRE y
TISSIER, ALAIN**

74 Agente/Representante:

UNGRÍA LÓPEZ, Javier

ES 2 743 512 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

Biblioteca de promotores sintéticos para la expresión génica coordinada en células u organismos eucariotas

5 **Campo de la invención**

La invención se refiere a la utilización de un kit o sistema para expresar dos o más secuencias de ácido nucleico de interés en un organismo eucariótico, célula o tejido de un organismo eucariótico, que comprende una biblioteca de promotores y un efector que se puede unir a sitios de unión al efector en los promotores de las bibliotecas. La invención se refiere adicionalmente a células eucarióticas u organismos eucarióticos, en particular células vegetales o plantas, que comprenden dos o más construcciones de ácido nucleico de interés, comprendiendo cada construcción de ácido nucleico de interés un promotor seleccionado de la biblioteca. La invención también se refiere a un proceso de expresión de dos o más secuencias de ácido nucleico de interés en células eucarióticas, en células o tejidos de un organismo eucariótico.

15 **Antecedentes de la invención**

Uno de los objetivos de la biología sintética consiste en establecer redes transcripcionales artificiales para la expresión de nuevas rutas de señalización o metabólicas. Si bien esto se realizó satisfactoriamente en microorganismos unicelulares (Brophy y Voigt, 2014), el desarrollo de tales redes transcripcionales en eucariotas superiores, en particular en plantas superiores, presenta dificultades adicionales debido a la falta de suficientes datos disponibles, a la multicelularidad y a la estructura de la cromatina. Un requisito previo para la construcción de tales redes es la disponibilidad de una gama de promotores transcripcionales de diferentes intensidades que pueden activarse mediante factores de transcripción ortogonales. Dado que tales promotores no están disponibles en la naturaleza, se deben diseñar, sintetizar, ensamblar en unidades de transcripción funcionales y se deben someter a prueba.

Se han adoptado varios enfoques para la modificación por ingeniería genética del promotor, incluyendo la mutagénesis aleatoria, la mutagénesis de saturación de regiones mal conservadas y la modificación por ingeniería genética de promotores híbridos (Blazeck y Alper, 2013). Una propiedad deseable adicional de los promotores sintéticos es su activación por factores de transcripción ortogonales (TF). Esto es particularmente relevante para organismos multicelulares tales como las plantas superiores, por ejemplo, si el objetivo es expresar una ruta de biosíntesis exclusivamente en un tejido o en condiciones específicas, tal como el estrés biótico o abiótico. La mayoría de las bibliotecas de promotores descritas hasta ahora en levadura se basan en secuencias de activación aguas arriba (UAS) de TF endógenos (Blazeck y Alper, 2013). Si el TF en el que se basa el diseño de estos promotores sintéticos se expresa constitutivamente, o solamente en tejidos o condiciones específicas, esto restringirá considerablemente su uso potencial. Para el TF constitutivo, esto evitaría la activación en tejidos específicos, y por el contrario para TF específicos de tejido, restringiría el uso de los promotores a un tejido específico, evitando su aplicación general para otros tejidos o afecciones. Un conjunto de promotores cuya expresión puede ser inducida por un TF ortogonal evitaría estos problemas. En este caso, la localización y el patrón de expresión estarían determinados por el TF ortogonal, que puede ponerse bajo el control de un promotor específico de tejido, por ejemplo.

Tornoe et al., *Gene* 297 (2002) 21-32 describen la generación de una biblioteca de promotores de mamífero sintéticos mediante la modificación de secuencias que separan los sitios de unión del factor de transcripción. También se discute, entre otras cosas, que un número teórico ilimitado de genes puede ser expresado a niveles individuales constantes por promotores aleatorizados, y que los promotores basales se pueden combinar con secuencias que confieren p. ej. una susceptibilidad de inducción dependiente de la fase de crecimiento, seguida de aleatorización de secuencias promotoras.

Boch et al., *Science* (2009), vol. 326, núm. 5959 describen los efectores TAL-Tipo III. Geißler et al., *PLOS ONE*, (2011), vol. 6, núm. 5, e19509 describe activadores transcripcionales de genes humanos con especificidad de ADN programable. Farzadfard et al., *ACS SYNTHETIC BIOLOGY* (2013), vol. 2, núm. 10, 604-613 describen factores de transcripción eucarióticos sintonizables y multifuncionales basados en CRISPR/Cas. De Lange et al., *THE PLANT JOURNAL*, (2014), vol. 78, núm. 5, 753-771 describen los efectores TAL como herramientas para la biología sintética.

Partiendo de la técnica anterior, un objeto de la invención es proporcionar medios para la expresión coordinada de más de una secuencia de ácido nucleico de interés, en particular en organismos eucarióticos o células de los mismos. Otro objeto es proporcionar medios para la expresión ortogonal de secuencias de ácido nucleico de interés a los niveles de expresión deseados. Un objeto adicional de la invención es proporcionar unidades estructurales para construir redes transcripcionales artificiales mediante la introducción de bucles de retroalimentación positiva o negativa, o cascadas transcripcionales.

Compendio de la invención

Para resolver estos objetos, la invención proporciona los contenidos de las reivindicaciones 1 a 13.

5 La presente invención utiliza, en una realización, los Efectores de Tipo Activador de la Transcripción ("TALE" o "Efectores TAL") recientemente descubiertos (Boch y Bonas, 2010, Boch et al., 2009). La estructura modular de su sitio de unión al ADN permite el diseño de factores de transcripción verdaderamente ortogonales al seleccionar promotores de la invención que tienen un sitio de unión al efector que no existe en el genoma, minimizando así la activación del gen fuera de la diana. Desde su descubrimiento, los TALE se han utilizado tanto para la activación de la transcripción como para la edición del genoma, esto último fusionándolos con nucleasas (de Lange et al., 2014). Se dispone de herramientas para la construcción de TALE de diseño personalizado (Weber et al., 2011).

15 La invención describe el diseño, la construcción y la prueba de una biblioteca de promotores sintéticos que se pueden activar mediante efectores tales como un factor de transcripción de diseño TALE ("dTALE"). Basándose en una arquitectura notablemente simple que comprende el sitio de unión dTALE-ADN, una caja TATA y secuencias degeneradas que flanquean estas regiones constantes en uno o ambos lados, estos promotores muestran propiedades que los convierten en herramientas ideales para el desarrollo de redes transcripcionales complejas o para aplicaciones ingeniería metabólica. Primero, como se muestra en los ejemplos para el caso del gen informador GUS, muestran un intervalo utilizable de niveles de expresión en ensayos transitorios que varía de 5 a 90% del promotor 35S, que es uno de los promotores de plantas más fuertes disponible. Se han observado grandes variaciones en los niveles de expresión de genes individuales dentro de una ruta de biosíntesis en las plantas. Por ejemplo, los genes que codifican las citocromo P450 monooxigenasas de la ruta de la partenólida en la matricaria (*Tanacetum parthenium*) se pueden expresar a niveles varios cientos de veces más altos que el gen que codifica la sintasa de germacreno A, que representa la primera etapa comprometida de la ruta (Liu et al., 2014). Por lo tanto, el intervalo de niveles de expresión suministrados por la biblioteca de promotores es particularmente útil para ajustar la expresión de genes de rutas individuales en proyectos de ingeniería metabólica. Esto puede servir para optimizar el flujo de la ruta, pero también para evitar la producción excesiva de ciertas enzimas, que pueden causar una sobrecarga metabólica de las células anfitrionas. En segundo lugar, en ausencia del efector, se detectaron niveles de expresión de fondo nulos o bajos, lo que demuestra la naturaleza ortogonal del sistema efector/promotor.

20 La biblioteca de promotores descrita en la presente memoria se puede utilizar para la ingeniería metabólica de rutas biosintéticas complejas que requieren la expresión simultánea y sintonizable de múltiples genes. Esto es particularmente relevante en plantas, donde los tejidos o condiciones específicas pueden ser sometidos a prueba y comparados por su rendimiento en la producción de clases específicas de compuestos. Además, estos promotores sintéticos proporcionan la base para el diseño y la construcción de redes reguladoras complejas que contienen cascadas transcripcionales, bucles de retroalimentación positiva y negativa. Por lo tanto, su potencial alcanza más allá del campo de la ingeniería metabólica.

Breve descripción de los dibujos

40 **Figura 1. Diseño y procedimiento de clonación de los STAP (promotores activados por TALE).** A. secuencia general de los STAP. Un sitio de unión dTALE constante (EBE002) de 19 bases se resalta sombreando en color azul y la caja TATA en color verde. El triplete atg al final del promotor corresponde al codón de inicio de la traducción. Las secuencias degeneradas aguas arriba y aguas abajo de la región constante se indican con (n)₁₉ y (n)₄₃ respectivamente. B. Procedimiento de clonación de los STAP. La biblioteca se sintetizó utilizando dos oligonucleótidos degenerados solapantes sobre la región constante y se clonaron en un vector Golden Gate de nivel -1. La posterior clonación en un vector de nivel 0 permitió la eliminación de productos aberrantes (por ejemplo, dímeros) y falsos positivos. Los vectores de nivel 0 se pueden utilizar a continuación para el ensamblaje en unidades de transcripción.

45 **Figura 2. Prueba de los STAP en un ensayo transitorio con un gen informador GUS.** A. Descripción general de las construcciones utilizadas para el ensayo transitorio. El dTALE que se une a la secuencia constante EBE002 de los STAP está bajo el control del promotor constitutivo y moderado *Act 2*. Las repeticiones de dTALE están indicadas por líneas verticales estrechas y marcadas como "dominio de unión a ADN". Los STAP controlan la expresión de un transgén GUS con intrones (representados por barras de color blanco dentro de la secuencia codificante sombreada) para evitar la expresión en *Agrobacterium* que se utiliza para la infiltración de hojas de *N. benthamiana*. B. Distribución de los valores de actividad de GUS para los STAP con o sin dTALE, expresados en porcentaje de expresión 35S:GUS.

50 **Figura 3. Prueba de STAP seleccionados con GFP.** Se sometieron a prueba ocho STAP diferentes que mostraban diferentes niveles de expresión con el gen informador GUS con GFP. Se infiltraron hojas de *N. benthamiana* con cepas de *Agrobacterium* que portaban diferentes construcciones STAP:GFP con (lado izquierdo) o sin (lado derecho) el dTALE. Se tomaron fotografías de las hojas con luz normal (primera y tercera columna) o bajo iluminación UV (segunda y cuarta columna) para visualizar GFP.

60 **Figura 4. Uso de STAP para la ingeniería metabólica de un diterpeno vegetal.** A. descripción general de la ruta de los cembratrienoles (CBToles), incluyendo la ruta del metileritritol-fosfato (MEP), destacando los

genes que se expresaban de manera anormalmente alta en *N. benthamiana*. Pyr, piruvato; GA3P, gliceraldehído-3-fosfato; DXP, 1-desoxililulosa 5-fosfato; IPP, isopentenil difosfato; DMAPP, dimetilalil difosfato; GGPP, geranylgeranyl difosfato; SIDXS2, 1-desoxililulosa 5-fosfato sintasa de tomate; NtGGPPS2, GGPP sintasa de tabaco; CBTS2a, cembratrienol sintasa de tabaco. B. Descripción general de las construcciones utilizadas para la expresión de manera anormalmente alta de SIDXS2, NtGGPPS2 y CBTS2a utilizando STAP. Los números de los STAP corresponden a los números en las Figuras 2 y 3. C. Comparación de la expresión dirigida por TALE y la expresión dirigida por p35S en hojas de *N. benthamiana* agroinfiltradas con p19 y CBTS2a, DXS2 y GGPPS2 en diferentes combinaciones. Se infiltraron dos hojas en tres plantas individuales para cada combinación de construcción (n = 6). Las cantidades medias de CBTol se expresan como equivalentes del patrón interno (IS, esclareol) \pm error típico. Los grupos indicados por letras diferentes (a1/a2 y b1/b2) difieren significativamente entre sí con respecto a sus valores de CBTol (P < 0,05, prueba t de Student y ANOVA de dos vías). No se detectó CBTol en las hojas que expresaban el vector T-ADN vacío como control.

Figura 5. Expresión dirigida por TALE y 35S de los genes CBTS2a, DXS2 y GGPPS2 en hojas de *N. benthamiana* agroinfiltradas. Los niveles de transcripción se cuantificaron por medio de RT-PCR en tres réplicas biológicas seleccionadas de las muestras utilizadas en la Figura 4 (rep1-3) de hojas agroinfiltradas que expresaban simultáneamente p19, CBTS2a, DXS2 y GGPPS2 dirigida por STAP (TALE) o el promotor 35S (35S). Se proporcionan los valores medios de expresión de tres réplicas técnicas \pm DT y las cantidades correspondientes de CBTol para cada muestra.

Figura 6. Correlación entre la expresión génica y el nivel de CBTol. Los datos de expresión de la Figura 5 se trazaron frente a los niveles de CBTol. Se muestran las líneas de tendencia y sus parámetros correspondientes. Los puntos de color rojo y verde corresponden a muestras con los STAP y el promotor 35S respectivamente.

25 Descripción detallada de la invención

La biblioteca de múltiples promotores es un conjunto de elementos múltiples únicos. En la biblioteca o conjunto de promotores múltiples utilizados en la invención, cada promotor comprende, en la dirección 5' a 3', (i) un sitio de unión al efector que comprende al menos 10 bases contiguas, en donde las secuencias de bases de los sitios de unión al efector de los múltiples promotores tienen una identidad de secuencia de al menos 75%, (ii) una caja TATA, y (iii) un segmento de secuencia aguas abajo que comprende al menos 20 bases contiguas. En una realización alternativa, cada promotor comprende, en dirección de 5' a 3', (I) un segmento de secuencia aguas arriba de al menos 10 bases contiguas, (II) un sitio de unión al efector que comprende al menos 10 bases contiguas, en donde las secuencias de bases de los sitios de unión al efector de los múltiples promotores tienen una identidad de secuencia de al menos 75% y (III) una caja TATA. Los promotores utilizados en la invención también se denominan en la presente memoria "STAPS" que significa "promotores sintéticos activados por TALE". La caja TATA es un elemento conocido del centro de los promotores eucarióticos. La caja TATA tiene la secuencia de ADN central 5'-TATAA-3' o una variante y puede estar seguida por tres o más pares de bases A-T. La caja TATA se puede definir por comprender la secuencia 5'-TATAW-3', preferiblemente 5'-TATAWW-3', o 5'-TATAWWW-3', en donde W representa A o T. En los ejemplos a continuación, se utiliza una caja TATA de secuencia 5'-TATATAA-3'.

Los promotores de la biblioteca tienen adicionalmente el sitio de unión al efector del apartado (i). La secuencia de ácido nucleico o bases del sitio de unión al efector permite la unión de efectores a los promotores tales como los efectores derivados de efectores de TAL, es decir, efectores que tienen unidades repetidas de un efector de TAL. El sitio de unión al efector comprende al menos 10 bases contiguas. Preferiblemente, comprende al menos 15, más preferiblemente al menos 17 bases contiguas. La longitud mínima de la secuencia de bases del sitio de unión al efector se puede elegir de modo que sea única entre las secuencias de promotores del organismo o célula eucarióticos en los que los promotores se utilizan para expresar secuencias de ácido nucleico de interés. De esta manera, la expresión de una secuencia de ácido nucleico de interés se puede volver ortogonal, es decir, no interfiere en la expresión de genes endógenos. No existe un límite superior estricto para la longitud del sitio de unión al efector. Sin embargo, no es necesario que la longitud sea más larga de lo necesario para la unión específica de un efector dirigido al sitio de unión al efector. Por lo tanto, el sitio de unión al efector generalmente tiene una longitud de 30 o menos bases contiguas, preferiblemente menos de 25 bases contiguas, más preferiblemente menos de 20 bases contiguas. Las longitudes preferidas del sitio de unión al efector son de 10 a 25 bases contiguas, preferiblemente de 14 a 20 bases contiguas.

Los sitios de unión al efector de los múltiples promotores de la biblioteca o conjunto de promotores tienen la misma secuencia de bases o una secuencia de bases muy similar a lo largo de la longitud descrita anteriormente, de modo que el mismo efector se puede unir a todos los múltiples promotores de la biblioteca. Para este propósito, las secuencias de bases de los sitios de unión al efector de los múltiples promotores tienen una identidad de secuencia de al menos 75%, preferiblemente al menos 85%, más preferiblemente al menos 90, más preferiblemente al menos 95% e incluso más preferiblemente de al menos 98%. Cualquier par de dos sitios de unión al efector de los múltiples promotores de la biblioteca puede tener las identidades de secuencia mínimas definidas en la oración anterior. Esto significa que no hay un sitio de unión al efector en la biblioteca que tenga una identidad de secuencia más pequeña

que la dada en la penúltima oración a un sitio de unión al efector de cualquier otro promotor en la biblioteca. Al calcular las identidades de secuencia entre dos secuencias de bases de sitios de unión al efector, se permite hasta un hueco, lo que significa que se puede descuidar hasta un hueco en el cálculo. En una realización preferida, no se permiten huecos, lo que significa que no se descuida ningún hueco en el cálculo. En una realización, las secuencias de bases de los sitios de unión al efector son idénticas entre los promotores en una biblioteca.

Las realizaciones preferidas del sitio de unión al efector son las siguientes: el sitio de unión al efector de los promotores múltiples es un segmento de secuencia de 10 a 30 bases contiguas y las secuencias de bases de los sitios de unión al efector de los promotores múltiples tienen una identidad de secuencia de al menos 85%, preferiblemente al menos 90%, más preferiblemente al menos 95% dentro de tal segmento; el sitio de unión al efector de los promotores múltiples es un segmento de secuencia de 13 a 27 bases contiguas y las secuencias de bases de los sitios de unión al efector de los promotores múltiples tienen una identidad de secuencia de al menos 85%, preferiblemente al menos 90%, más preferiblemente al menos 95% dentro de tal segmento; el sitio de unión al efector de los promotores múltiples es un segmento de secuencia de 16 a 24 bases contiguas y las secuencias de bases de los sitios de unión al efector de los promotores múltiples tienen una identidad de secuencia de al menos 85%, preferiblemente al menos 90%, más preferiblemente al menos 95% dentro de tal segmento.

Como alternativa a la definición de la alta similitud de los sitios de unión al efector entre los promotores de la biblioteca a modo de porcentaje de identidad de secuencia, la similitud mínima de los sitios de unión al efector se puede definir como sigue. En una realización, los sitios de unión al efector de los promotores de la biblioteca comprenden al menos 10 bases contiguas y las secuencias de bases de los sitios de unión al efector de los promotores múltiples son las mismas en una longitud de 10 bases contiguas, excepto que una o dos bases, preferiblemente una base, puede ser intercambiada. Alternativamente, se puede eliminar o añadir una base y se puede intercambiar una base.

En otra realización, las secuencias de bases de los sitios de unión al efector de los promotores múltiples son las mismas en una longitud de 15 bases contiguas, excepto que se pueden intercambiar hasta tres bases, preferiblemente hasta dos bases, más preferiblemente hasta una base. Alternativamente, se puede eliminar o añadir una base y se pueden intercambiar una o dos bases, preferiblemente una base.

En otra realización, las secuencias de bases de los sitios de unión al efector de los promotores múltiples son las mismas en una longitud de 20 bases contiguas, excepto que hasta cuatro bases, preferiblemente hasta tres bases, más preferiblemente hasta dos bases, más preferiblemente una base, pueden ser intercambiadas. Alternativamente, se puede eliminar o añadir una base y se pueden intercambiar hasta tres bases, preferiblemente hasta dos bases, más preferiblemente una base.

Se pueden preparar diferentes bibliotecas en donde las secuencias de bases de los sitios de unión al efector tienen diferencias más grandes o tienen más diferencias de las indicadas anteriormente. Sin embargo, dentro de una biblioteca dada, las secuencias de bases de los sitios de unión al efector de los promotores múltiples son iguales o muy similares a lo largo de las longitudes que se definieron anteriormente, de modo que el mismo efector puede unirse a todos los promotores múltiples de la biblioteca. Sin embargo, no es necesario que un efector dado se una a todos los sitios de unión al efector de los promotores de la biblioteca con la misma fuerza.

Si los sitios de unión al efector descritos anteriormente de una biblioteca de promotores se consideran un primer sitio de unión al efector, uno o más promotores de la biblioteca pueden tener un segundo sitio de unión al efector o incluso un sitio de unión al efector adicional aguas arriba de la caja TATA para permitir la unión de un segundo efector a los promotores. Un primer efector es capaz de unirse al primer sitio de unión al efector. Un segundo (o más) efector es capaz de unirse a un segundo (o más) sitio de unión al efector. El primer efector puede ser un activador de la transcripción capaz de unirse al primer sitio de unión al efector. El segundo efector también puede ser un activador de la transcripción o puede ser un represor de la transcripción que se une al segundo sitio de unión al efector. Generalmente, las propiedades de un segundo sitio de unión al efector o un sitio de unión al efector adicional en términos de longitud e identidades mínimas de secuencia de bases corresponden a las indicadas anteriormente para el primer sitio de unión al efector. En el caso de dos o más sitios de unión al efector, los dos o más sitios de unión al efector en un promotor dado pueden tener la misma secuencia de bases. Preferiblemente, sin embargo, los dos o más sitios de unión al efector difieren en la secuencia de ácido nucleico para permitir la unión de diferentes efectores. Para este propósito, las secuencias de bases de múltiples sitios de unión al efector en el mismo promotor generalmente tienen una identidad de secuencia de a lo sumo 80%, preferiblemente a lo sumo 60%. Es posible que todos los promotores de una biblioteca dada tengan un segundo sitio de unión al efector o un sitio de unión al efector adicional. En otra realización, algunos pero no todos los promotores de una biblioteca dada tienen un segundo o más sitios de unión al efector.

Los promotores de una biblioteca comprenden adicionalmente un segmento de secuencia que comprende al menos 20 bases contiguas aguas abajo de la caja TATA. Los promotores múltiples difieren en los segmentos de secuencia aguas abajo, lo que constituye la biblioteca utilizada en la invención. La longitud del segmento de secuencia aguas

abajo es preferiblemente de al menos 25, más preferiblemente al menos 30 nucleótidos (contiguos). No existen límites estrictos para el límite superior de la longitud del segmento de secuencia aguas abajo. Generalmente no hay ninguna ventaja por el hecho de que el segmento de secuencia aguas abajo sea más largo que 150 nucleótidos contiguos. Por lo tanto, la longitud es preferiblemente 150 nucleótidos contiguos o menos, preferiblemente 100 nucleótidos contiguos o menos. En una realización, hay como máximo 40 nucleótidos en los segmentos de secuencia aguas abajo de la biblioteca. Los intervalos posibles de la longitud de los segmentos de secuencia aguas abajo de los promotores de la biblioteca son los siguientes: de 25 a 150 bases, de 30 a 100 bases, de 40 a 90 bases, o de 40 a 70 bases. Los promotores en la biblioteca pueden tener segmentos de secuencia aguas abajo de la misma longitud, es decir, número de pares de bases. Sin embargo, no es necesario que el número de pares de bases sea el mismo entre los diferentes promotores de la biblioteca. Los promotores de la biblioteca difieren al menos en los segmentos de secuencia aguas abajo. El término "difieren en los segmentos de secuencia aguas abajo" significa que los segmentos de secuencia aguas abajo difieren en longitud y/o en secuencia de bases. Los segmentos de secuencia aguas abajo de diferentes promotores de la biblioteca difieren en al menos 1, preferiblemente al menos 2, más preferiblemente al menos 5 bases sobre las longitudes de segmentos de secuencia aguas abajo indicadas anteriormente. Dado que puede haber muchos promotores en la biblioteca, esto todavía permite una amplia variedad de segmentos de secuencia aguas abajo dentro de una biblioteca.

En una realización, los promotores de la biblioteca tienen segmentos de secuencia aguas abajo de longitudes idénticas, y diferentes promotores de la biblioteca difieren en al menos 1, preferiblemente en al menos 2, más preferiblemente en al menos 5 bases en los segmentos de secuencia aguas abajo.

Los promotores de la biblioteca de promotores pueden tener un segmento de secuencia aguas arriba ubicado aguas arriba de dicho sitio o sitios de unión al efector, comprendiendo dicho segmento de secuencia aguas arriba al menos 10 nucleótidos contiguos, en donde los promotores múltiples difieren, adicional o alternativamente a las diferencias en la secuencia aguas abajo segmentos, en los segmentos de secuencia aguas arriba. La longitud del segmento de secuencia aguas arriba es preferiblemente de al menos 20 nucleótidos, más preferiblemente a lo sumo 30 nucleótidos. Generalmente no hay ninguna ventaja por el hecho de tener un segmento de secuencia aguas arriba más largo que 100 nucleótidos contiguos. Por lo tanto, la longitud es preferiblemente de 100 nucleótidos contiguos o menos, preferiblemente 50 nucleótidos contiguos o menos, más preferiblemente 30 nucleótidos contiguos o menos. La variedad de secuencia proporcionada por el segmento de secuencia aguas arriba opcional proporciona más variedad adicional entre la biblioteca de promotores. Los segmentos de secuencia aguas arriba difieren en longitud y/o en secuencia de bases. Los segmentos de secuencia aguas arriba de diferentes promotores pueden diferir en al menos 1, preferiblemente en al menos 2, más preferiblemente en al menos 4 bases. Dado que puede haber muchos promotores en la biblioteca, esto permite una amplia variedad de segmentos de secuencia aguas arriba dentro de una biblioteca. En una realización, los promotores de la biblioteca tienen segmentos de secuencia aguas arriba de longitudes idénticas, y diferentes promotores de la biblioteca difieren en al menos 1, preferiblemente en al menos 2, más preferiblemente en al menos 4 bases. En una realización preferida, los promotores de la biblioteca tienen segmentos de secuencia aguas arriba de longitudes idénticas y segmentos de secuencia aguas abajo de longitudes idénticas, en donde las longitudes posibles son las definidas anteriormente para los segmentos de secuencia aguas arriba y aguas abajo.

Los promotores de la biblioteca pueden tener un tramo de secuencia entre el sitio de unión al efector y la caja TATA. Los sitios de unión al efector pueden estar separados de la caja TATA por un segmento de secuencia de ácido nucleico de a lo sumo 500, preferiblemente a lo sumo 300, más preferiblemente a lo sumo 100 bases, e incluso más preferiblemente a lo sumo 50 o incluso a lo sumo 20 bases. Sin embargo, también es posible que no haya una secuencia o haya una secuencia muy corta entre el sitio de unión al efector y la caja TATA, tal como un nucleótido como en los siguientes ejemplos. Entre los miembros de la biblioteca, no es necesario que todos tengan un tramo entre el sitio de unión al efector y la caja TATA del mismo número de nucleótidos.

Los promotores de la biblioteca pueden tener un tramo de secuencia entre la caja TATA y los segmentos de secuencia aguas abajo. En una realización, los segmentos de secuencia aguas abajo siguen directamente a la caja TATA, en donde la caja TATA se toma para terminar con la última A o T siguiendo, en dirección de 5' a 3', la secuencia TATA de la caja TATA. Esto significa que las diferencias entre los segmentos de secuencia aguas abajo se determinan comenzando con la base que sigue a la última A o T de la caja TATA.

Una biblioteca de promotores utilizados en la invención comprende múltiples promotores, es decir, al menos 2 promotores. Una biblioteca comprende preferiblemente al menos 3, más preferiblemente al menos 5, incluso más preferiblemente al menos 10 promotores, incluso más preferiblemente al menos 20 promotores que difieren en la secuencia de bases de dicho segmento de secuencia aguas abajo y opcionalmente de dicho segmento de secuencia aguas arriba.

Algunos o todos los promotores de la biblioteca pueden comprender o codificar una región 5' no traducida (5'-UTR) y/o un codón de inicio aguas abajo de dicho segmento de secuencia aguas abajo. Los segmentos de secuencia aguas abajo y una 5'-UTR pueden solaparse o no solaparse. Un ejemplo de una región 5' no traducida de uso

común es el potenciador Omega del Virus del Mosaico del Tabaco (Gallie et al., Nucleic Acids Res. 24 de abril de 1987; 15(8):3257-73).

5 La biblioteca utilizada en la invención puede ser producida por medio de síntesis *de novo* de los promotores mediante métodos de síntesis de ácido nucleico. Otro enfoque para la síntesis de la biblioteca se describe en los ejemplos y se ilustra en la Fig. 1. En este enfoque, se sintetizan oligonucleótidos degenerados solapantes que se solapan en la región de los sitios de unión al efector. La degeneración del oligonucleótido introduce la variabilidad en el segmento de secuencia aguas abajo y, opcionalmente, aguas arriba. La variación de secuencia del segmento de secuencia aguas abajo de los promotores y, opcionalmente, el segmento de secuencia aguas arriba es proporcionada por porciones de los oligonucleótidos que no se solapan. Las mezclas de oligonucleótidos directos e inversos se clonan en un vector aceptor, preferiblemente mediante clonación haciendo uso de sitios de restricción de tipo II tales como "Golden Gate Cloning". Los clones obtenidos después de la transformación en bacterias se pueden seleccionar y los promotores se pueden secuenciar.

15 En general, la biblioteca utilizada en la invención se almacena en plásmidos tales como los vectores de nivel -1 o nivel 0 representados en la Fig. 1. Los plásmidos o vectores generalmente contienen un marcador seleccionable y se pueden mantener en bacterias. La biblioteca puede estar en una forma en donde los elementos de la biblioteca se almacenan en recipientes separados. Esto permite, después de haber identificado un promotor adecuado, identificar el recipiente específico que contiene el promotor identificado. En otra realización, los promotores múltiples de la biblioteca están presentes en un recipiente, es decir, están mezclados. En tal realización, la biblioteca aún se puede escrutar y los clones que contienen un promotor seleccionado se pueden identificar utilizando métodos de clonación y secuenciación generalmente conocidos (véanse los ejemplos a continuación).

25 En la presente memoria, las secuencias de ácido nucleico (o bases) se proporciona en la dirección 5' a 3' y se refieren a la hebra codificante a menos que se especifique de manera diferente. El significado de "hebra codificante" es como se define en la Figura 11.1 en GENES IX, Lewin, Benjamin, Jones y Bartlett Publishers. Las longitudes de las secuencias de bases o nucleótidos se proporcionan en números de bases contiguas. Los mismos valores numéricos se aplican cuando las longitudes se proporcionan con respecto a pares de bases o nucleótidos.

30 La invención también proporciona el uso de un kit o sistema que comprende (A) una biblioteca de múltiples promotores según cualquiera de las realizaciones descritas en la presente memoria y (B) un efector capaz de unirse a los sitios de unión al efector de todos los miembros de la biblioteca. Alternativamente, se utiliza una molécula de ácido nucleico que codifica el efector como elemento (B) del sistema o kit. En una primera realización (descrita adicionalmente más adelante), el efector comprende un dominio de repetición que comprende unidades de repetición derivadas de o un efector de tipo activador de la transcripción (TAL), siendo dicho efector capaz de unirse al (primer) sitio de unión al efector de los promotores. En una segunda realización (descrita adicionalmente más adelante), el efector comprende un ARN (tal como ARNg) que tiene complementariedad con el sitio de unión al efector de los promotores, por lo que el efector puede comprender adicionalmente una fusión de CRISPR/Cas9-factor de transcripción (TF) modificada por ingeniería genética.

40 La biblioteca del apartado (A) del kit o sistema puede estar presente en un recipiente, es decir, en un estado mezclado. Alternativamente, la biblioteca puede estar presente en el kit o sistema de modo que cada elemento (promotor) de la biblioteca esté contenido en un recipiente separado. En un efector o una molécula de ácido nucleico que lo codifica o un componente del mismo está generalmente contenido en un recipiente separado de los recipientes que contienen elementos de la biblioteca. La información proporcionada en la presente memoria sobre la biblioteca y el efector también se relaciona con la biblioteca y el efector, respectivamente, del kit, sistema, organismos, células y procedimientos.

50 Una aplicación principal de la invención es la expresión de dos o más, preferiblemente tres o más, más preferiblemente cuatro o más, secuencias de ácido nucleico de interés en un organismo eucariótico o célula o tejido del mismo. Preferiblemente, la expresión de las dos o más secuencias de ácido nucleico de interés se coordina en el sentido de que los niveles de expresión de las dos o más secuencias de ácido nucleico de interés están predeterminados mediante la selección de diferentes promotores apropiados de una biblioteca de promotores para lograr niveles de expresión deseados para cada una de las dos o más secuencias de ácido nucleico de interés. Adicionalmente, la expresión puede ser ortogonal ya que es independiente de los factores de transcripción endógenos en las células u organismo anfitriones. Para este propósito, las al menos dos secuencias de ácido nucleico de interés se pueden colocar bajo el control de diferentes promotores de la biblioteca de promotores para obtener una construcción de ácido nucleico de interés para cada una de las secuencias de ácido nucleico de interés.

60 Para la expresión de dos o más secuencias de ácido nucleico de interés en un organismo eucariótico o célula o tejido del mismo, la invención proporciona una célula u organismo eucarióticos que excluyen a los seres humanos y un proceso de expresión de dos o más secuencias de ácido nucleico de interés en células eucarióticas, en células o tejido de un organismo eucariótico. Si se van a expresar dos secuencias de ácido nucleico de interés, los promotores de la biblioteca de promotores se utilizan para la secuencia de ácido nucleico de interés que difiere en sus

segmentos de secuencia aguas abajo y/o los segmentos de secuencia aguas arriba. Si se van a expresar tres o más secuencias de ácido nucleico de interés, los promotores utilizados para al menos dos de las tres o más secuencias de ácido nucleico de interés difieren en sus segmentos de secuencia aguas abajo y/o los segmentos de secuencia aguas arriba. También es posible que las tres o más secuencias de ácido nucleico de interés se expresen utilizando promotores de la biblioteca que difieran en sus segmentos de secuencia aguas abajo y/o sus segmentos de secuencia aguas arriba.

Para expresar dos o más secuencias de ácido nucleico de interés, se pueden preparar construcciones de ácido nucleico de interés, en donde cada construcción de ácido nucleico de interés comprende un promotor como se describe en la presente memoria y, aguas abajo del mismo, una secuencia de ácido nucleico de interés que va a ser expresada. Por lo tanto, también se describe un conjunto de dos o más construcciones de ácido nucleico de interés, comprendiendo cada construcción una secuencia de ácido nucleico de interés diferente. Al menos dos de dichas dos o más construcciones de ácido nucleico de interés tienen promotores que difieren en la secuencia de bases de sus segmentos de secuencia aguas abajo y/o de los segmentos de secuencia aguas arriba. Una construcción de ácido nucleico de interés puede comprender adicionalmente otros elementos reguladores para expresar las secuencias de ácido nucleico de interés, tales como un terminador aguas abajo de las secuencias de ácido nucleico de interés. El uso de terminadores para la expresión de genes eucarióticos es parte del conocimiento general del experto.

La célula, tejido u organismos eucarióticos de la invención se pueden producir introduciendo las dos o más construcciones de ácido nucleico de interés simultáneamente, p. ej. por co-transformación, o consecutivamente. Las construcciones de ácido nucleico de interés utilizadas para la transformación o transfección son generalmente parte de un vector de transformación que contiene un marcador seleccionable que se puede utilizar para seleccionar células transformadas. Para la producción de organismos o células transformados de forma estable, se realizan procedimientos de selección para seleccionar células u organismos que tengan las construcciones de ácido nucleico introducidas de forma estable en el genoma de las células o el organismo. Para la expresión transitoria de las secuencias de ácido nucleico de interés, generalmente no se realiza la selección de células u organismos transformados. En el proceso de expresión de dos o más secuencias de ácido nucleico de interés en células eucarióticas, en células o tejidos de un organismo eucariótico, o en un organismo eucariótico, se pueden proporcionar células eucarióticas, tejido de un organismo eucariótico o un organismo eucariótico con las dos o más construcciones de ácido nucleico de interés como se describió anteriormente. La información proporcionada en la presente memoria sobre los elementos de los promotores también se aplica a los promotores utilizados en la célula u organismo eucarióticos y a los promotores utilizados en el proceso de expresión.

Los métodos de transformación o transfección de células u organismos eucariotas son parte del conocimiento general.

El efector utilizado en la invención, o un componente del mismo, es capaz de unirse a los sitios de unión al efector y de activar la transcripción de los promotores. El efector utilizado en la invención puede ser un efector que no existe naturalmente en el organismo eucariótico, o en las células del mismo, donde deben expresarse las secuencias de ácido nucleico de interés. Por consiguiente, los sitios de unión al efector de los promotores pueden ser sitios de unión que no son reconocidos por efectores endógenos tales como factores de transcripción o represores presentes en el organismo eucariótico, o células de los mismos, donde se deben expresar las secuencias de ácido nucleico de interés. La célula u organismo eucarióticos comprenden adicionalmente o están provistos de una construcción de ácido nucleico que expresa efector que codifica un efector, en donde el efector se puede unir a los sitios de unión al efector de los promotores y activar la transcripción de dichos promotores. La activación de la expresión de una o más construcciones de ácido nucleico de interés se puede combinar con la represión de la expresión de una o más construcciones de ácido nucleico de interés adicionales.

En una primera realización de efectores utilizables en la invención, el efector se basa en los Efectores de Tipo Activador de la Transcripción (TALE) recientemente descubiertos (Boch y Bonas, 2010, Boch et al., 2009, documento WO 2012/104729). Los TALE son factores de virulencia de patógenos de plantas del género *Xanthomonas* que se translocan a través de un sistema de secreción de tipo III dentro de la célula vegetal. A continuación, los TALE se importan al núcleo, donde se unen a secuencias de ADN específicas y activan transcripcionalmente la expresión génica. La unión al ADN está mediada por una región central repetitiva, formada, generalmente, hasta por 33 repeticiones en tándem de un motivo de 33 a 35 residuos de aminoácidos, correspondiendo cada repetición a un par de bases de ADN de la secuencia diana. Las secuencias de aminoácidos de las repeticiones son casi idénticas, excepto las posiciones de aminoácidos 12 y 13, los denominados díresiduos variables de repetición (RVD). Las repeticiones con diferentes RVD muestran diferentes preferencias de pares de bases de ADN, y los RVD consecutivos en un TALE corresponden directamente a la secuencia de ADN en el lado de unión, lo que da como resultado un código simple de una repetición a un par de bases (Weber et al., 2011). Las repeticiones en tándem también se denominan "módulos" o "unidades" en la presente memoria. La estructura modular del sitio de unión al ADN de los TALE permite el diseño de factores de transcripción ortogonales seleccionando una secuencia de ADN que no exista en el genoma, minimizando así la activación del gen fuera de la diana. Los efectores TAL naturales tales como AvrBs3 bien estudiado son proteínas monoméricas que comprenden

activadores de otros efectores TAL. Adicionalmente, se pueden utilizar dominios activadores de otros activadores transcripcionales, tales como el dominio activador transcripcional de la proteína Virion 16 (VP16) del virus del herpes simple tipo 1 (Cress et al., Science 4 de enero de 1991: vol. 251 núm. 4989 pág. 87-90) o Gal4 de levadura (Ma et al., Cell, Volumen 48, Apartado 5, 13 de marzo de 1987, Páginas 847-853). Alternativamente, el dominio activador puede ser reemplazado por un dominio represor, por lo que el efector se puede utilizar para reprimir la transcripción del promotor. La transformación de efectores en represores transcripcionales se realizó con éxito en varios organismos eucarióticos, incluyendo las plantas (Blount et al., 2012, Mahfouz et al., 2012). Los ejemplos de dominios represores transcripcionales incluyen el dominio SRDX vegetal (Hiratsu et al., 2003, Plant Journal, 34:733-739) y los dominios SID y KRAB humanos (le Cong et al., 2012, Nature Communications, doi: 10.1038/ncomms1962). Weber et al., 2011 describen métodos para producir efectores basados en TALE, los vectores que los codifican y su uso.

Existen varias formas posibles de proporcionar células eucarióticas, tejido de un organismo eucariótico o un organismo eucariótico con el efector. En un ejemplo, se genera una construcción de ácido nucleico que codifica expresamente el efector (también denominado en la presente memoria "construcción de ácido nucleico que expresa el efector"). En tal construcción, el efector codificado está bajo el control de la expresión de un promotor que puede o no ser un promotor de la biblioteca de la biblioteca de múltiples promotores. En una realización básica, el promotor de la construcción de ácido nucleico que expresa el efector no es un promotor de la biblioteca de promotores descritos en la presente memoria, pero aún puede ser un promotor regulado o específico de tejido. En tal realización, la expresión de múltiples secuencias de ácido nucleico de interés se puede controlar controlando la expresión del efector. Alternativamente, el promotor es un promotor constitutivo en las células u organismos utilizados para la expresión. La construcción de ácido nucleico que expresa el efector se puede proporcionar a continuación a las células o a las células del organismo de modo que se pueda expresar en ellas.

En otro ejemplo, el efector se introduce en dichas células, en células de dicho tejido o en células de dicho organismo eucariótico como una proteína. Las proteínas pueden ser transformadas en células a partir de composiciones libres de células como se describe en el documento WO 2004/046360. Los ejemplos específicos consisten en la inyección de una solución que contiene la proteína efectora en el tejido de un organismo. Para este propósito, el efector se puede fusionar con una secuencia de translocación de membrana que permita la introducción directa de la proteína efectora en las células.

En un ejemplo adicional, el efector se introduce en las células de un organismo eucariótico o en un organismo eucariótico por cruzamiento de un organismo eucariótico que contiene las construcciones de ácido nucleico de interés integradas en el genoma, tal como el genoma nuclear, con un organismo eucariótico de la misma especie que contienen en su genoma una construcción de ácido nucleico que expresa el efector que codifica el efector como se describió anteriormente. Preferiblemente, los organismos parentales contienen las construcciones de ácido nucleico respectivas homocigóticamente de modo que toda la progenie contenga una copia de las construcciones deseadas.

En una segunda realización de los efectores, se utilizan efectores basados en el sistema CRISPR/Cas9 (Fineran y Dy, 2014, Hsu et al., 2014, van der Oost et al., 2014). La enzima CRISPR/Cas9 es una nucleasa cuya especificidad de secuencia está determinada por un ARN guía (ARNg) complementario a una secuencia diana (van der Oost et al., 2014). Como tal, CRISPR/Cas9 es adecuado para la edición del genoma, con la posibilidad de elegir como diana varios loci coexpresando o incluso concatenando múltiples ARNg, que se pueden clonar fácilmente debido a su pequeño tamaño (Hsu et al., 2014). En este sentido, la edición del genoma CRISPR/Cas9 es más sencilla de implementar que la edición basada en TALE, porque para cada locus que se va a elegir como diana, se deben diseñar y expresar dos nuevos dTALE (de Lange et al., 2014). El sistema CRISPR/Cas también se modificó para generar TF (CRISPR/Cas-TF), cuya especificidad está determinada por los ARNg (Nissim et al., 2014; documentos WO 2014/197748 y WO 2015/006747). Para lograr esto, los ARNg fueron expresados por los promotores de la ARN pol II, requiriendo la edición del ARN por la enzima Csy4 (Qi et al., 2012). El hecho de que la especificidad de CRISPR/Cas-TF se determine a nivel de ARN, y no a nivel de proteína como con los TALE, constituye una limitación para el diseño de redes reguladoras más complejas con etapas de activación o represión. Debido a que todas las proteínas CRISPR/Cas reconocerán los ARNg y su ADN diana independientemente de si están modificados para que sean activadores o represores transcripcionales, la coexpresión de ambos tipos de CRISPR/Cas-TF en una sola célula podría dar como resultado una expresión impredecible de genes diana. Asimismo, una sola proteína CRISPR/Cas-TF que controla toda la red no puede actuar simultáneamente como represor y activador sobre diferentes dianas. Por lo tanto, para introducir una etapa reguladora negativa en una red reguladora CRISPR/Cas, Nissim et al. recurrieron a un miARN (Nissim et al., 2014). Por el contrario, se pueden coexpresar distintos dTALE con propiedades represoras o activadoras, lo que permite una mayor flexibilidad y posibilidades para el diseño de redes reguladoras (de Lange et al., 2014).

En esta segunda realización, el efector comprende al menos dos componentes. El efector comprende un ARN que tiene complementariedad con un sitio de unión al efector de los promotores. El efector comprende preferiblemente un ARN guía (ARNg), más preferiblemente un ARN guía única (ARNgu) y una fusión de CRISPR/Cas-factor de transcripción. El ARNgu es una fusión funcional modificada por ingeniería genética de ARNcr y ARNtracr (Jinek et

al., 2012). La célula u organismo eucarióticos comprenden o están provistos de una construcción de ácido nucleico que codifica el ARNg o el ARNgu. La célula u organismo eucarióticos comprenden o están provistos adicionalmente de una construcción de ácido nucleico que codifica un CRISPR/Cas-TF. El CRISPR/Cas es preferiblemente CRISPR/Cas9.

CRISPR/Cas9-TF es una proteína de fusión que comprende un dominio de una proteína CRISPR/Cas9 modificada por ingeniería genética (denominado dCas9 en la técnica tal como en Maeder, ML. et al., 2013) que está modificada por ingeniería genética para que sea catalíticamente inactiva, y un dominio de un TF. Se proporcionan ejemplos de proteínas Cas9 modificadas por ingeniería genética catalíticamente inactivas en el documento WO 2015/006747. El TF puede ser un represor o un activador. Se han utilizado fusiones dCas9 a dominios TF para activar genes informadores en *E. coli* y células humanas (Bikard, D., 2013; Gilbert LA., 2013). Los documentos WO 2014/197748 y WO 2015/006747 describen ejemplos de fusiones CRISPR/Cas9-TF. Los métodos para expresar construcciones que codifican efectores CRISPR/Cas9 se conocen en la técnica anterior citada en la presente memoria. El TF es preferiblemente heterólogo para las células u organismo eucarióticos. Un ejemplo de una fusión CRISPR/Cas9-TF es una proteína de fusión del dominio de activación transcripcional dCas9-VP64 descrita por Maeder, ML. et al., 2013. La proteína de fusión del dominio de activación transcripcional dCas9-VP64 puede ser dirigida por uno o múltiples ARNg para aumentar la expresión de genes humanos endógenos específicos.

Las construcciones de ácido nucleico que codifican el ARN de efector y la fusión CRISPR-/Cas-TF están bajo control transcripcional de los promotores en las construcciones para coexpresarlas en las células o en el organismo eucarióticos. Cuando el ARNg se transcribe en las células, el ARNg dirige la unión específica de la secuencia de un complejo CRISPR/Cas9/ARNg a la secuencia diana, es decir, un sitio de unión al efector de un promotor utilizado en la invención. El complejo comprende la fusión CRISPR/Cas9-TF complejada con el ARNg o ARNgu que puede hibridar con el sitio de unión al efector.

Los organismos eucarióticos, o células o tejidos de los mismos, considerados en la invención para la expresión pueden ser plantas, animales u hongos. Entre estos, se prefieren las plantas. Entre los animales, se prefieren mamíferos tales como los roedores, pero los seres humanos están excluidos de la protección. Sin embargo, la invención se puede utilizar para expresar proteínas en células humanas (y otras animales). Entre los hongos, se prefieren las levaduras.

En una realización básica, el efector actúa como activador de la expresión en construcciones de ácido nucleico de interés que contienen promotores como se describe en la presente memoria. Por lo tanto, cuando comienza la expresión del efector en células eucarióticas, o cuando existe un efector introducido directamente en los núcleos celulares, el efector se puede unir a los sitios de unión al efector de construcciones de ácido nucleico de interés y activar la expresión de las secuencias de ácido nucleico de interés contenidas en las construcciones.

Los promotores utilizados en la invención se pueden utilizar para la ingeniería metabólica de rutas biosintéticas complejas, que incluyen la expresión simultánea y sintonizable de múltiples genes. Esto es particularmente relevante en plantas, donde los tejidos o condiciones específicos se pueden someter a prueba y comparar para determinar el rendimiento en la producción de clases específicas de compuestos. Además, los promotores proporcionan la base para el diseño y la construcción de redes reguladoras complejas que contienen cascadas transcripcionales, circuitos de retroalimentación positiva y negativa. Por lo tanto, su potencial alcanza más allá del campo de la ingeniería metabólica. La retroalimentación positiva se puede realizar simplemente añadiendo a uno o más de los promotores un sitio de unión al efector para un segundo efector. La retroalimentación negativa se puede lograr de manera similar reemplazando el dominio activador de un segundo efector por un dominio represor. La transformación de efectores en represores transcripcionales se realizó con éxito en varios organismos eucarióticos, incluyendo las plantas (Blount et al., 2012, Mahfouz et al., 2012). Por lo tanto, la biblioteca de promotores proporciona el material de partida para diseñar circuitos reguladores complejos que podrían incluir cascadas transcripcionales, operadores booleanos, así como circuitos de retroalimentación positiva y negativa.

Por lo tanto, en la célula, tejido u organismo eucarióticos, y en el proceso de expresión de dos o más secuencias de ácido nucleico de interés, una o más de dichas construcciones de ácido nucleico de interés pueden tener un segundo sitio de unión al efector o un sitio de unión al efector adicional aguas arriba de la caja TATA. El segundo sitio de unión al efector o sitio de unión al efector adicional se describió anteriormente en el contexto de la biblioteca. El segundo sitio de unión al efector o sitio de unión al efector adicional se puede ubicar entre el primer sitio de unión al efector y la caja TATA o, preferiblemente, está/están ubicados aguas arriba del primer sitio de unión al efector. Una o más de las construcciones de ácido nucleico de interés pueden contener un ácido nucleico de interés que codifica un segundo efector capaz de unirse al segundo sitio de unión al efector o sitio de unión al efector adicional de una o más construcciones de ácido nucleico de interés para la activación (cuando el segundo efector es o contiene un activador) o represión (cuando el segundo efector es o contiene un represor) de la transcripción de secuencias de ácido nucleico de interés en una o más construcciones de ácido nucleico de interés diferentes que contienen el segundo sitio de unión al efector o sitio de unión al efector adicional.

Dado el número cada vez mayor de secuencias del genoma que están disponibles, incluso de plantas superiores, es posible diseñar efectores que no interfieran o interfieran mínimamente en la expresión de genes endógenos y, por el contrario, que no estén unidos por TF endógenos. Por lo tanto, las combinaciones efector/promotor utilizadas en la invención se pueden convertir en un sistema verdaderamente ortogonal que no dependa ni interfiera en las redes de activación de la transcripción endógena. La relevancia de este sistema para la ingeniería metabólica es alta. Los experimentos de los autores de la presente invención en el ensayo de expresión transitoria para la producción de diterpeno en *N. benthamiana* demostraron que el sistema efector/promotor puede coexpresar varios genes a niveles comparables al promotor CaMV 35S fuerte. Si bien el uso del mismo promotor para coexpresar múltiples genes en un ensayo transitorio funciona correctamente, en plantas transgénicas estables esto puede conducir a consecuencias indeseables tales como el silenciamiento génico (Peremarti et al., 2010). Por lo tanto, una aplicación particularmente relevante de estos promotores será en plantas transgénicas estables, donde se puede realizar una expresión coordinada específica de tejido de genes de la ruta sin requerir la identificación de promotores endógenos múltiples que confieran el mismo perfil de expresión y con diferentes niveles de expresión. Por lo tanto, simplemente cambiando el promotor que impulsa la expresión del efector activador y manteniendo idéntico el resto de la construcción de interés, será posible con un mínimo esfuerzo de clonación someter a prueba diferentes tejidos en paralelo para determinar el rendimiento en la producción de un compuesto de interés. Por ejemplo, se podrían comparar tejidos tales como el endospermo de la semilla, la epidermis, los tricomas glandulares o los pelos radiculares, para los cuales se encuentran disponibles promotores específicos (Potenza et al., 2004, Tissier, 2012).

Por consiguiente, también se describe un proceso de selección de múltiples promotores de una biblioteca de múltiples promotores para expresar una secuencia de ácido nucleico dada de interés en un organismo eucariótico preseleccionado o células del mismo, que comprende preparación de múltiples construcciones de ácido nucleico de interés, cada una de las cuales contiene la misma secuencia de ácido nucleico de interés y, aguas arriba y unidos operativamente a las mismas, diferentes promotores de la biblioteca de promotores como se define en el apartado 1, introducción de las múltiples construcciones de ácido nucleico de interés en el organismo eucariótico preseleccionado o sus células. posterior expresión de cada uno de dichos constructos de ácido nucleico múltiple de interés en el organismo eucariótico preseleccionado o células del mismo, y selección del promotor de la construcción de ácido nucleico de interés que muestra la expresión adecuada o deseada de dicha secuencia de ácido nucleico de interés en el organismo eucariótico preseleccionado o células del mismo.

Las etapas del proceso de dicho proceso de escrutinio se pueden realizar basándose en el conocimiento general del experto y las enseñanzas proporcionadas en la presente memoria. Una vez que se identifican los promotores adecuados, se pueden utilizar para expresar la secuencia de ácido nucleico de interés utilizando la construcción que contiene el promotor seleccionado.

Ejemplos

En los siguientes Ejemplos, se describen el diseño y la construcción de bibliotecas de promotores activados por TALE sintéticos (STAP). Estos se ensamblaron en vectores compatibles Golden Gate para facilitar su uso aguas abajo (Engler et al., 2014). Estos STAP se evaluaron con genes informadores (GUS y GFP) en ensayos transitorios en *Nicotiana benthamiana*, a continuación, se sometieron a prueba en una configuración de ingeniería metabólica para la producción de un diterpeno vegetal.

Construcción de una biblioteca de promotores sintéticos activados por TALE (STAP)

Para generar una biblioteca de secuencias promotoras como se describe en la Figura 1, se elaboraron y dos cebadores degenerados, Talpro7 y Talpro8 (véase las secuencias a continuación), que contenían el sitio de unión TALE-ADN (EBE-002 o sitio de unión dTALE-2, descrito en (Weber et al., 2011)) y las secuencias degeneradas a cada lado del sitio de unión al ADN y se reasociaron en una construcción pL2-dTALE-2 (Weber et al., 2011) con la diferencia de que el promotor 35S en pL2-dTALE-2 se reemplazó por el promotor Act2 de Arabidopsis en pICH74043; la secuencia de nucleótidos de pICH74043 se presenta en SEQ ID NO: 52) y una construcción 35S:GFP como referencia interna (construcción pAGM4731, (Engler et al., 2014)). Como control, las construcciones pAGT615-n también se infiltraron sin TALE. Se utilizó una construcción con 35S:GUS (pICH75181, (Engler et al., 2014)) co-infiltrada con 35S:GFP como control positivo y referencia externa. Los autores de la presente invención infiltraron las construcciones informadoras con y sin TALE y un control positivo (35S:GUS) en diferentes sectores de la misma hoja para eliminar la variación de hoja a hoja. Las cepas de Agrobacterium se cultivaron en medio LB, se resuspendieron en medio de infiltración (MES 10 mM, MgSO₄ 10 mM) a una DO₆₀₀ final de 0,3 para la infiltración. Las plantas se volvieron a colocar a continuación en el invernadero durante cinco días antes del análisis. Para los ensayos cuantitativos de GUS, los discos de hoja de 0,9 cm de diámetro dentro de las áreas infiltradas se cortaron y se transfirieron a microtubos que contenían tres cuentas de acero de 0,2 mm y se congelaron inmediatamente en N₂ líquido. Los tejidos foliares congelados se homogeneizaron a continuación dos veces utilizando un molino mezclador

Retsch MM400 a una frecuencia de 30 Hz durante 1 minuto. El polvo de la hoja se resuspendió después en 1 mL de solución salina tamponada con fosfato (pH 7,4), se sometió a agitación vorticial y se filtró a través de un filtro de PTFE de 0,2 µm en una placa de 96 pocillos (Chromafil, Macherey-Nagel, Alemania). La actividad de GUS se midió utilizando el sustrato fluorogénico 4-metilumbeliferil-β-D-glucopiranosido (MUG) empelando procedimientos publicados a los 10, 30 y 60 minutos (Jefferson et al., 1987). Paralelamente, se utilizaron 200 µL de los extractos para medir la fluorescencia de GFP. La fluorescencia de GFP de cada muestra se utilizó como factor de normalización. La fuerza de los promotores individuales se expresó a continuación como un porcentaje de la expresión con el promotor 35S, basándose en el control de la misma hoja.

10 **Ensayos de GFP**

Los promotores seleccionados de la serie pAGT582-n se clonaron en un vector de nivel 1 (pICH47732) frente a GFP (vector pICH41531, (Engler et al., 2014)) y el terminador ocs (pICH41432). Las construcciones se introdujeron en *Agrobacterium* GV3101 (pMP90) y se expresaron en *N. benthamiana* con o sin TALE-EBE002 (véase más arriba). La construcción 35S:GFP (pAGM4731, véase más arriba) se utilizó como control positivo. Las plantas se incubaron durante tres a cinco días y se tomaron fotos bajo una lámpara UV.

Construcciones de ingeniería metabólica de diterpeno

20 Las secuencias codificantes de los genes *SIDXS2* (pAGT345), *NtGGPPS2* (pAGT169) y *CBTS2a* (pAGT238) fueron las descritas en (Brückner y Tissier, 2013). Éstas fueron clonadas detrás de los promotores de pAGT582-3, pAGT582-5 y pAGT582-1 respectivamente. Los terminadores utilizados fueron SI vATPase1 (pICH 71431), Ocs (pICH 41432) y CBTS (pAGT168) respectivamente (Brückner y Tissier, 2013, Engler et al., 2014). Estos fragmentos se ensamblaron en los vectores de nivel 1 pICH47751, pICH47742 y pICH47802 respectivamente. Los vectores se transformaron en *Agrobacterium tumefaciens* GV3101:pMP90 y se utilizaron para la infiltración foliar de *N. benthamiana* como se describió anteriormente. Las construcciones correspondientes con el promotor 35S fueron las descritas en (Brückner y Tissier, 2013). La extracción de metabolitos y las mediciones de diterpeno por GC-MS se realizaron como se describe en (Brückner y Tissier, 2013) con las siguientes modificaciones. Los extractos con hexano de 1 ml de los discos de 6 hojas por muestra no se evaporaron y 160 µL de estos extractos se combinaron con 40 µL de patrón interno (IS) (solución de partida de Esclareol 50 µM en hexano), se mezclaron, se centrifugaron y se transfirieron a un vial de GC. Las muestras se midieron en un sistema QP2010 Ultra GC-MS (Shimadzu) en las siguientes condiciones. Se inyectó 1 µL sin división a 250°C con un tiempo de muestreo de 1 min en una columna RXi-5Sil MS de 30 m, 0,25 mm de diámetro (Restek). La elución se realizó con un flujo constante de He de 0,9 ml/min. La temperatura del horno se elevó de 50°C a 300°C (7°C/min), a continuación a 320°C (50°C/min) y se mantuvo durante 2 minutos. Las temperaturas de la interfaz y de la fuente de iones fueron de 300°C y 250°C respectivamente y el voltaje de ionización se ajustó a 70 V.

Análisis cuantitativo de la expresión génica en tejido foliar infiltrado

40 El ARN total se extrajo de los discos foliares infiltrados de los cuales se extrajeron antes los diterpenos mediante lavado con hexano. Los discos foliares se homogeneizaron a un polvo fino utilizando un Molino Mezclador MM-400 (Retsch) y el ARN se extrajo con el kit Spectrum™ Plant Total RNA (Sigma-Aldrich) de acuerdo con las especificaciones del fabricante. Se eliminó cualquier contaminación genómica antes de la síntesis de ADNc utilizando el kit sin ADN (Ambion) de acuerdo con las instrucciones del fabricante. La primera cadena de ADNc se sintetizó a partir de 1 µg de ARN total utilizando el H Minus First Strand cDNA Synthesis Kit (Thermo Scientific) y una mezcla de oligo(dT)18 y cebadores hexaméricos aleatorios. El ADNc se diluyó 10 veces y se utilizó 1 µl en una mezcla de reacción de 10 µl que contenía 1 µl (2 pmoles/µl) de cebadores directos e inversos y 2 µl de my-Budget 5x EvaGreen QPCR Mix II (Bio & Sell). La PCR cuantitativa en tiempo real se realizó en placas de 96 pocillos utilizando el sistema de detección por PCR en tiempo real CFX Connect (Bio-Rad). El perfil térmico de la reacción fue el siguiente: una desnaturalización inicial a 95°C durante 15 minutos, seguida de 40 ciclos a 95°C durante 15 segundos y 58°C durante 30 segundos con adquisición de fluorescencia después de cada ciclo. Para verificar la especificidad del cebador, finalmente se generó una curva de disociación aumentando la temperatura de 65 a 95°C. Los cebadores de la qPCR para la amplificación de los genes *CBTS2a*, *DXS2* y *GGPPS2* fueron diseñados utilizando el programa Primer3 (Untergasser et al., 2012). El factor de elongación eucariótico de *N. benthamiana* 1a se seleccionó como referencia endógena para la normalización de los niveles de transcrito, y se amplificó por PCR utilizando los cebadores descritos en Liu et al., 2012. Para verificar la especificidad de los cebadores *CBTS2a*, *DXS2* y *GGPPS2*, se realizó la qPCR también en ADNc de *N. benthamiana* y no se pudo detectar amplificación en ninguna de las muestras de control (Figura 5). La secuencia de los oligonucleótidos utilizados para la PCR en tiempo real se proporciona en los datos suplementarios. Todas las muestras de ADNc se amplificaron en tres réplicas técnicas de la misma preparación de ARN y se consideró el valor medio. La expresión de transgenes en tejido foliar infiltrado se calculó utilizando el método $2^{-(\Delta\Delta Ct)}$ (Livak y Schmittgen, 2001).

Resultados

Diseño y construcción de una biblioteca de promotores sintéticos activados por TALE

5 La estructura de los STAP se presenta en la Figura 1A. Contienen un dominio de unión a ADN de dieciocho bases de longitud, llamado EBE002, flanqueado por una secuencia degenerada de diecinueve bases de longitud inmediatamente aguas arriba, y una caja TATA de consenso aguas abajo. La caja TATA es seguida por una secuencia degenerada de cuarenta y tres bases de longitud y el codón de inicio ATG. Este diseño se basa en la observación de que la secuencia de los TALE que se une al ADN de patógenos bacterianos que activan genes *en plantas* normalmente se encuentran cerca del sitio de inicio de la transcripción (menos de 100 pb) y esa inserción de una secuencia de unión al ADN en la posición -55 o -40 era suficiente para conferir inducibilidad mediada por TALE (Kay et al., 2009). La biblioteca se clonó utilizando el sistema MoClo basado en la clonación Golden Gate (Werner et al., 2012). El esquema para la construcción de la biblioteca de STAP se presenta en la Figura 1B. La biblioteca se realizó en un vector de nivel -1 del sistema MoClo, produciendo pAGT559-L, y se transfirió en un vector de nivel 0, produciendo pAGT582-L. Los vectores de nivel 0 contienen partes (típicamente promotor, ADNc, terminador) que se ponen en contacto en vectores de nivel 1 para ensamblar unidades de transcripción funcionales. La transferencia de la biblioteca del nivel -1 al nivel 0 asegura que solamente se mantengan los fragmentos que contienen salientes correctos, lo que limita el número de falsos positivos y aumenta la eficiencia de la clonación en las etapas posteriores. Se recogieron aproximadamente 50 clones de la biblioteca de nivel 0 y se verificó la presencia de un fragmento del tamaño esperado. Cuarenta y tres clones fueron seleccionados y secuenciados (véase la Lista de secuencias). Todas las secuencias promotoras contenían el sitio de unión al ADN EBE002.

Ensayos transitorios de los STAP con un gen informador GUS

25 Cada STAP se clonó aguas arriba del gen informador GUS y el terminador de la transcripción ocs en un vector de nivel 1, que también se puede utilizar para la transferencia de ADN-T a *Agrobacterium tumefaciens* (Figura 2A). La actividad transcripcional de estos promotores se evaluó primero en ensayos transitorios en *Nicotiana benthamiana* midiendo la actividad de GUS en extractos de proteínas. El dTALE está bajo el control del promotor Act2 (Figura 2A), que proporciona suficiente expresión para la activación de la transcripción. Para la cuantificación de la expresión, las hojas se infiltraron por triplicado con las construcciones STAP:GUS, con o sin dTALE y con una construcción 35S:GFP. Se midió la fluorescencia de GFP y se utilizó como factor de normalización. Los resultados (Figura 2B) muestran una distribución de intensidades de expresión que oscila de menos de 1% a más de 90% de la del promotor 35S. Dado que para algunos promotores hay una expresión de fondo residual significativa sin TALE, el intervalo útil de expresión en estos ensayos transitorios se estima entre 5 y 90% de la del promotor 35S. Sin embargo, para la mayoría de los STAP, no se detectó una expresión significativa en ausencia de TALE, lo que indica que la actividad observada con TALE no proviene de activadores endógenos, estableciendo la ortogonalidad de los promotores.

Ensayos transitorios de STAP seleccionados con GFP

40 Para confirmar estos resultados con otro gen informador, se clonó una selección de ocho STAP con diferentes intensidades de expresión (STAP números 1, 2, 3, 5, 14, 17, 42 y 44) frente al gen de GFP y se expresaron transitoriamente en *N. benthamiana*. Los resultados (Figura 3) son coherentes con los del ensayo de GUS, lo que indica que la intensidad de expresión para GUS y GFP parece ser independiente del gen transcrito. Para todos los STAP sometidos a prueba, no se puede observar ninguna señal de GFP en ausencia de TALE, lo que confirma la alta especificidad de los promotores sintéticos. Los promotores más fuertes (números 1, 3 y 5) muestran una intensidad de señal cercana a la del promotor 35S, mientras que los promotores más débiles (números 17 y 42) aún producen una señal detectable.

Uso de los STAP para la ingeniería metabólica de diterpeno en *N. benthamiana*

A continuación, los autores de la presente invención buscaron validar la utilidad de estos promotores utilizándolos para la ingeniería metabólica de un diterpenoide de vegetal. Como resultado para el ensayo de ingeniería metabólica de los autores de la presente invención, se utilizó la cembratrienol sintasa (CBTS2a) del tabaco (Ennajdaoui et al., 2010, Wang y Wagner, 2003). Esta diterpeno sintasa produce una mezcla de dos estereoisómeros (α y β) del cembratrienol macrocíclico (CBTol) (Figura 4). En las plantas, los precursores terpenoides isopentenil difosfato (IPP) y dimetilalil difosfato (DMAPP) se producen por dos rutas distintas, la ruta plástica de 2-C-metil-D-eritritol-4-fosfato (MEP) y la ruta citosólica de mevalonato (MEV) (Vranova et al., 2012). Los diterpenos se sintetizan en los plástidos y, por lo tanto, utilizan predominantemente IPP y DMAPP de la ruta MEP. Anteriormente los autores de la presente invención mostraron que la coexpresión de la 1-desoxilulosa 5-fosfato sintasa (DXS), la primera etapa de la ruta MEP, con una geranilgeranil difosfato sintasa (GGPPS) conduce a más del doble de la producción de diterpeno en ensayos transitorios con *N. benthamiana* (Brückner y Tissier, 2013). Estos ensayos se llevaron a cabo con el promotor fuerte del Virus del Mosaico de la Coliflor 35S que controla la expresión de todos los genes. Los genes *DXS2* (tomate), *GGPPS2* (tabaco) y el *CBTS2a* fueron clonados aguas abajo de algunos de los STAP más fuertes

caracterizados anteriormente y expresados de manera transitoria junto con dTALE. Los resultados muestran que tanto con el promotor 35S como con los STAP, la cantidad de CBToles producidos es significativamente mayor cuando todos los genes (*CBTS2a*, *GGPPS2* y *DXS2*) se coexpresan en comparación con las otras combinaciones (*CBTS2a* solo, *CBTS2a* + *DXS2* o *CBTS2a* + *GGPPS2*) (Figura 4C), confirmando resultados anteriores con el promotor 35S (Brückner y Tissier, 2013). Sin embargo, los CBToles producidos con los STAP son aproximadamente tres veces más bajos que con 35S. Una explicación podría ser que la expresión génica es menor, lo que lleva a niveles reducidos de CBToles. Para verificar esto, los autores de la presente invención midieron los niveles de transcrito por medio de RT-PCR cuantitativa de las mismas muestras utilizadas para la cuantificación de CBToles. Sorprendentemente, los niveles de transcrito fueron más altos con los STAP que con el promotor 35S (1,9 veces para *CBTS2a*, 11,5 para *GGPPS2* y 1,5 para *DXS2*) (Figura 5). Así, los niveles más bajos de CBToles no pueden estar causados por una reducción de los niveles de transcrito de los genes. El trazado de los niveles de transcrito del gen frente a la cantidad de CBToles producidos indica una correlación negativa, independientemente de los promotores utilizados (STAP o 35S) (Figura 1 suplementaria). Por lo tanto, este efecto inverso de la expresión génica sobre los niveles de metabolitos no parece ser específico de los STAP.

Estos resultados muestran que los STAP se pueden utilizar para aplicaciones de ingeniería metabólica y pueden suministrar niveles de expresión, que son comparables a los de los promotores más fuertes conocidos.

Discusión

Los autores de la presente invención han validado la utilidad de los STAP para la ingeniería metabólica en un ensayo transitorio. Han confirmado que la coexpresión de *GGPPS2* y *DXS2* con la diterpeno sintasa *CBTS2a* conduce a mayores niveles de CBToles (Figura 4C). Sin embargo, los niveles fueron más bajos que con el promotor 35S, lo que podría deberse a niveles más bajos de expresión génica. Sin embargo, inesperadamente, los niveles de transcrito medidos por RT-PCR cuantitativa muestran que los STAP confieren una expresión significativamente mayor que el promotor 35S, en particular hasta más de diez veces mayor para el gen *GGPPS2*. Se pueden proponer varias hipótesis para explicar esta correlación negativa entre los niveles de transcrito y el rendimiento del producto. La alta expresión de las enzimas podría conducir a una retroalimentación negativa sobre la ruta MEP y, por lo tanto, a una reducción del flujo a través de esta ruta. Otra posibilidad es que los CBToles se produzcan en mayores cantidades con los STAP pero se modifiquen por enzimas de *N. benthamiana* endógenas (p. ej., acil- o glicosil-transferasas, glutatión transferasas), como se ha demostrado para los mono y sesquiterpenoides (Ting et al., 2013). Aquí, sin embargo, habría que recurrir a que cuantos más CBToles se producen, más se modifican en proporción. Otra posibilidad más es que los niveles de transcrito más altos no están acompañados de velocidades de traducción más altas, sino que, por el contrario, esto conduce a una reducción de la actividad de traducción. Someter a prueba estas diferentes hipótesis requerirá extensos experimentos adicionales que deberían llevarse a cabo en el futuro.

Actualmente se describen promotores eucarióticos que contienen dos dominios principales, a saber, el elemento promotor central, que se define como la región donde se recluta la maquinaria de transcripción y se inicia la transcripción, y los elementos aguas arriba, que están unidos por factores de transcripción que pueden activar o reprimir la transcripción (Hahn y Young, 2011). Está bien establecido que la unión a elementos específicos aguas arriba por factores de transcripción desencadena el reclutamiento de la maquinaria de transcripción central y por lo tanto activa la transcripción (Hahn y Young, 2011). Sin embargo, apenas se sabe nada sobre cómo la secuencia del elemento promotor central influye en la actividad transcripcional (Lubliner et al., 2013). En un estudio inicial, el análisis de 95 promotores de levadura condujo a la identificación de la llamada señal de "localizador" de 30-10 pb aguas arriba del sitio de inicio de la transcripción (TSS), que es rico en T y pobre en A, y es "más fuerte" en promotores con alta actividad (Maicas y Friesen, 1990). Más recientemente, una encuesta de 859 promotores centrales de levadura identificó la riqueza de T aguas arriba del TSS como un fuerte predictor de la actividad del promotor, independientemente de si los promotores son constitutivos o inducibles (Lubliner et al., 2013). Hasta donde se ha podido deducir de la bibliografía disponible, no se sabe nada en las plantas. Debido a su secuencia corta y a la presencia de un sitio de unión de ADN idéntico, los STAP constituyen un material apropiado para estudiar la influencia de la composición de bases del promotor central sobre la actividad de transcripción. Además, es posible que la secuencia del gen pueda tener un impacto sobre los niveles de expresión génica.

El diseño de los autores de la presente invención de los STAP se basó en informes que identifican la presencia de sitios de unión a AvrBs3 dentro de los 100 pb del sitio de inicio de la transcripción de genes activados por AvrBs3, un efector de TAL de la bacteria patógena de plantas *Xanthomonas campestris* (Kay et al., 2007, Kay et al., 2009, Romer et al., 2007). Además, Kay y sus colaboradores demostraron que una caja TATA inmediatamente aguas abajo del sitio de unión a TALE conduce a una expresión enormemente aumentada (Kay et al., 2009). Aunque el diseño de los autores de la presente invención dio resultados satisfactorios, las variaciones adicionales podrían proporcionar opciones reguladoras adicionales. Por ejemplo, la adición de múltiples sitios de unión a TALE podría conducir a una mayor expresión, como se muestra en las células de mamíferos (Pérez-Pinera et al., 2013). Sin embargo, dado que los expresores más fuertes de los autores de la presente invención están casi a la par con el fuerte promotor viral 35S, esto puede tener un interés limitado. Más relevante sería la inserción de sitios de unión a TALE para la regulación de retroalimentación positiva o negativa.

Secuencias de ácido nucleico utilizadas y mencionadas en la presente memoria

Secuencias de los promotores. El sitio de unión a ADN de 19 bases de longitud de TALE se indica en negrita:

- 5 SEQ ID NO: 1 >PAGT582_1
GGAGGTTATCACACTTGGTAATTT**CCCCGCATAGCTGAACAT**CTATACAACACTCATC
GCGCGGAGAAACGCGCAACAAATTGGAGGCGCATTAAATG
- SEQ ID NO: 2 >PAGT582_2
GGAGCTTACGTAGCAGCACGGCCT**CCCCGCATAGCTGAACAT**CTATATAAATCTTAGC
CCGACTGCACCGTGAGGAAAAGAGGGTTTATATATAATG
- 10 SEQ ID NO: 3 >PAGT582_3
GGAGGTCGTCAAATGATAAGTACT**CCCCGCATAGCTGAACAT**CTATATAACGCAGCAA
ATGGTTGAAACGTATTAGCAACAGCTAAACCTTCGAATGT
- SEQ ID NO: 4 >PAGT582_4
GGAGGCGTGTGTTTTAGTGAGGT**CCCCGCATAGCTGAACAT**CTATATAAGGTTTTGC
TATTCATTGAAAGCAGTAGTGACTGATTTGTATATAATG
- SEQ ID NO: 5 >PAGT582_5
GGAGGTTCTGTGGGTTTTTTGGGT**CCCCGCATAGCTGAACAT**CTATATAATGAGAAAC
AGGGGCAAATCCCCGAAAACAATAACGCGAAGCAATG
- 15 SEQ ID NO: 6 >PAGT582_7
GGAGATGTGCCATGGTCGGGGGCT**CCCCGCATAGCTGAACAT**CTATATAATCGGAAT
GAATTGCAAATAGGGATGTTATAGGGTTATCTCGGCAATG
- SEQ ID NO: 7 >PAGT582_8
GGAGAGATGATTGTTGTAATATT**CCCCGCATAGCTGAACAT**CTATATAAGAGACGAG
CTACCACGATAAGCGCAGCAAAGACACATACGTTAATG
- 20 SEQ ID NO: 8 >PAGT582_9
GGAGGTGCCTTTTTAAAAGGGTTT**CCCCGCATAGCTGAACAT**CTATATAAGAAACACG
GACGCCGAAACCCCGCCTAGTTGACTCTCCGGAATG
- SEQ ID NO: 9 >PAGT582_10
GAGTGTGAACGTGGTGATGTGGT**CCCCGCATAGCTGAACAT**CTATATAAAACTGTAA
CAGAACTGCTGCTCAACTAGTCTCCAATATATAAATG
- 25 SEQ ID NO: 10 >PAGT582_13
GGAGGTAAGTTCTGGCCTCTGGGT**CCCCGCATAGCTGAACAT**CTATATAATAGGACA
GTAAGATATGCCAGACAATGTAGCACGTAGAACGGAAATG
- SEQ ID NO: 11 >PAGT582_14
GGAGCACCTTGAAAAACAGGTGGT**CCCCGCATAGCTGAACAT**CTATATAAGAAATTAG
TAACTACGCTTTCAAAGACATCTTAATTAATTAGAATG
- 30 SEQ ID NO: 12 >PAGT582_15
GGAGAGTGGGCGCGTTTTGGTAAGT**CCCCGCATAGCTGAACAT**CTATATAAATAAGCT
GCCTACGATTTAGCGCGACACAATTCGACGGTTGGAATG
- SEQ ID NO: 13 >PAGT582_21
GGAGGCCGTGCGGGGGGGGTCTGT**CCCCGCATAGCTGAACAT**CTATATAATACCGGT
TATCCAGCGACAACCGTCAATAACTTAACATTGAACAATG

ES 2 743 512 T3

SEQ ID NO: 14 >PAGT582_23

GGAGTTAAGTAGGCTGTTCTTATT**CCCCGCATAGCTGAACAT**CTATATAAATAGCTCAG
TTGCCTTAAACGTAGACAGTCTATTACACTAATAAATG

SEQ ID NO: 15 >PAGT582_24

GGAGGTTTATACTTTCTCTAATTT**CCCCGCATAGCTGAACAT**CTATATAAACTGGA
TTCACTGACCTGACAGTGCTTTCATTGACGCGGAATG

SEQ ID NO: 16 >PAGT582_25

5 GGAGAAGCTCACCTGCATCCTTG**CCCCGCATAGCTGAACAT**CTATATAAAATTTATAC
ATTTAAACCAAACCTATAAAATTTCTGTCCGCGAATG

SEQ ID NO: 17 >PAGT582_29

GGAGTAATGTACAGCTTCGCGGCT**CCCCGCATAGCTGAACAT**CTATATAAACTGTT
CTACATACTCTCGGATAAATAACGCAACTTAATATAATG

SEQ ID NO: 18 >PAGT582_30

10 GGAGTCCCGGATCCTGGTTTCACT**CCCCGCATAGCTGAACAT**CTATATAAAACAAATC
TCGTCAAGCAAACCAATTACGATGGTCCGGGAAGGAATG

SEQ ID NO: 19 >PAGT582_34

GGAGGGGTGGGGGGTTACCGTCT**CCCCGCATAGCTGAACAT**CTATATAAAACCAAT
CGACGACTCATTGGAATGCGTCCTGCTCCAATAGCAATG

SEQ ID NO: 20 >PAGT582_36

GGAGCTGCAATGCTTACTGTTATT**CCCCGCATAGCTGAACAT**CTATATAAAGGTTGAG
AAGTGGCACTTGCAATAATTGCCGGGCCCATGGGAAATG

SEQ ID NO: 21 >PAGT582_37

15 GGAGAGAGGTGGGTTCGCGGGCATT**CCCCGCATAGCTGAACAT**CTATATAAAAAGCAG
TTCATACGGGAACGCAGACTAGTCACGAAATTCGCAATG

SEQ ID NO: 22 >PAGT582_38

GGAGGCGCGTTTGATTGGGGGGT**CCCCGCATAGCTGAACAT**CTATATAAACTGAT
AACTTAGTCAGACTAGCAAATGCTCACAATACAAAGAATG

SEQ ID NO: 23 >PAGT582_39

20 GGAGGAACCGCGGTTTCGGGAGCGT**CCCCGCATAGCTGAACAT**CTATATAACAGGGG
CGGGATATGATATAGAAACACCAGGGTTAGTAACAAGAATG

SEQ ID NO: 24 >PAGT582_40

GGAGCAATGGCAGAGGGGTATGGT**CCCCGCATAGCTGAACAT**CTATATAAAATTGGG
ACCTATGATAATGCTGTAGAACCTCATAAATGATGGAATG

SEQ ID NO: 25 >PAGT582_41

25 GGAGGGTTTGTACCGTGTATTCT**CCCCGCATAGCTGAACAT**CTATATAACGAGCGGT
CGAAAGCTTCAAATTC AATGAGTAAGACCCCAACAATG

SEQ ID NO: 26 >PAGT582_42

GGAGTCTTGCCATACTCAGTTCT**CCCCGCATAGCTGAACAT**CTATATAATGTCTCTC
GAATGACACCGATACTTCGATTTAGCAGGCTGCCAATG

SEQ ID NO: 27 >PAGT582_43

30 GGAGGCATTACGGGCGTAGTTTGT**CCCCGCATAGCTGAACAT**CTATATAATCGACCAG
AGCAACGTTCTCGCCTAGGCGCGCCTTACCGGGAAAATG

ES 2 743 512 T3

SEQ ID NO: 28 >PAGT582_44

GGAGGTATCCCGAGGGCGAATGGT**CCCCGCATAGCTGAACAT**CTATATAAGGCATCT
GGTTAGCAAGCCTCACCAGCCGTGAACAATATAATCAATG

SEQ ID NO: 29 >PAGT582_45

GGAGTCAGAGTGTGGTCGTGTTTT**CCCCGCATAGCTGAACAT**CTATATAAGTTTACAC
CTGCCCGACCTATTCCTTGATGAGCTCGCGCACCGAATG

SEQ ID NO: 30 >PAGT582_47

5 GGAGCACTTAGGTTAACTGCGCGT**CCCCGCATAGCTGAACAT**CTATATAAGGAACAGA
CGGCTAACCCATCAATGGAACAACGAACGGCCATAATG

SEQ ID NO: 31 >PAGT582_48

GGAGTGGTTTTAGCACAGACA**ACTCCCCGCATAGCTGAACAT**CTATATAAGGAACCGA
CTCTCTCCCGCCCCGACACAGGACCCCTCTGGTAATG

SEQ ID NO: 32 >PAGT582_52

10 GGAGCATGCTGGGGTTGGAGGGGT**CCCCGCATAGCTGAACAT**CTATATAACCCGGCC
CCTCAATTAACCCTGCTCAAAGCCAAAGCAACAATCAATG

SEQ ID NO: 33 >PAGT582_53

GGAGTGGGCTGAGGGGGACGGGT**CCCCGCATAGCTGAACAT**CTATATAACCCAACG
TGAGATCACGTTTACAATACTCGTCTTGACTGACACAATG

SEQ ID NO: 34 >PAGT582_54

GGAGGGCATGAGGTTAGCTAGTT**CCCCGCATAGCTGAACAT**CTATATAATGAATGTG
TGAGAAGCTTTGACAAGCTTCGCCAGCGGCCCTAAATGT

SEQ ID NO: 35 >PAGT582_56

15 GGAGATATTCCTCGGCTGTC**ACTCCCCGCATAGCTGAACAT**CTATATAATCAAATTC
TAGGGGCTGCCATGGTTACCCCTTGCTATGTAGCAATG

SEQ ID NO: 36 >PAGT582_58

GAGTGCGGGGAGTTGGTATGGT**TCCCTGCATAGCTGAACAT**CTATATAATAACGGC
CCTGAGTGTGCTAGCAAATGCACCAAACCAACTCAATG

SEQ ID NO: 37 >PAGT582_60

20 GGAGCGATATGTTTAATCGGCATT**CCCCGCATAGCTGAACAT**CTATATAAATCCGCGC
CTTAAATACGGTTGTTTGCCTATTTTGAGCGAGTAAATG

SEQ ID NO: 38 >PAGT582_62

GGAGTATTGGTCCACTTGCTCCGT**CCCCGCATAGCTGAACAT**CTATATAACGCAGAAA
CCTCGCAACTGACGTTCCGATTGCGATAAACACACAATG

SEQ ID NO: 39 >PAGT582_65

25 GGAGTTGTAAACTCCAGGTTAGAT**CCCCGCATAGCTGAACAT**CTATATAAATCGCCCA
TTGCCAGGACACGCAACCCTGCAACAATCAGTACAATG

SEQ ID NO: 40 >PAGT582_69

GGAGCCCTTGCTATGACGATGCCT**CCCCGCATAGCTGAACAT**CTATATAAAAATTCCA
TAACGTACAGAACGAAGAACGGACCCATATACAATAATG

SEQ ID NO: 41 >PAGT582_73

30 GGAGGCGCGGGCAGGGGTTGGGT**CCCCGCATAGCTGAACAT**CTATATAATCACCCA
ACGCCGCATGCAACA**ACTAGTAACAAGGCCATTGAGAATG**

ES 2 743 512 T3

SEQ ID NO: 42 >PAGT582_78

GGAGCTTTTTACGGGGCACCTGTTCCCGCATAGCTGAACATCTATATAACAAAATAA
CAGATTGTCTGTCCAAGAGGCAGACATTCTTCATAATG

Secuencias de los cebadores utilizados para el análisis cuantitativo de la expresión génica:

- 5 SEQ ID NO: 43 NsCBTS2a-1106F, 5'- CGATCCAAAGATGGGACGA-3',
SEQ ID NO: 44 NsCBTS2a-1257R, 5'- CATCTCTTTTTTCGCATAGTAGATGT-3',
SEQ ID NO: 45 NtGGPPS2-85F, 5'- CCTCCACAGAGACATTCTTATAGTTTC-3',
SEQ ID NO: 46 NtGGPPS2-227R, 5'- TCAAAGTCAAACCTTAGGCAAGATG-3',
SEQ ID NO: 47 SIDXS2-556F, 5'- AGCATCTCCGCTGGTCTT-3',
10 SEQ ID NO: 48 SIDXS2-699R, 5'- CGAGTCAAGGAATCCTGCAT-3'
SEQ ID NO: 49: Talpro7: 5'-ttt ggtctc a acat GGAG (n)₁₉ tccccgcatagctgaacatc
SEQ ID NO: 50: Talpro8: 5'-ttt ggtctc a acaa CATT (n)₄₃ ttatata g atgttcagctatgccccg
SEQ ID NO: 51: consenso de repetición del efector TAL
SEQ ID NO: 52: pICH74043 secuencia que incluye el promotor Act2 y la secuencia no traducida, la secuencia
15 condificante del efector TAL (subrayada) y el terminador Ocs

tcgacaaaatttagaacgaactaattatgatctcaaatacattgatacatatctcatctagatctagggtatcattatgtaagaaag
ttttgacgaatatggcagcagacaaaatggctagactcgcgatgtaattggatctcaactcaacattataactataccaaacattagta
gacaaaattaaacaactatTTTTatgatgcaagagtcagcatatgtataaattgattcagaatcgtttgacgagttcggatgtagt
agtagccattatTaatgtacataactaatcgtgaatagtgaaatgatgaaacattgtatcttattgtataaatatccataaacacatc
atgaaagacacttctttcacggtcgaattaattatgatacaattctaatagaaaacgaattaattacggtgaaattgatgaaatc
aattgaacaagccaaccacgacgactaacgttgctggattgactcggtttaagttaaccactaaaaaacggagctgt
catgtaacacgagcagcaggtcacagtcagcatgaagccatcaaagcaaaagaactaatccaagggctgagatgattaat
tagtttaaaaattagttaacacgagggaaaaggctgtcagacgccaggtcacgttatctttacctgtggtcgaaatgattcgtgt
ctgtcgatttaattatTTTTgaaaggccgaaaataaagtgtgaagagataaacccgcctatataaatcatatatttctctccgc
ttgaaactgtataagagctcattttacaacaattaccaacaacaacaacaacaacattacaattacattacaattat
cgataca

atggatccgattcgttctcggacaccttcccgcctcggagctccttctgaccacaaccagatggagtgcaaccaacagct
gataaggggtatctccaccgcccgtggacctctgacggcttacctgaaggccacacatgicccgaccagattgccagt
ccaccagcaccgaagtcagcgtttcagcggcagcttctcagacctctgagacagttgatccctcttttcaacacctact
gtttgactcacttccgcttttgggctcaccatactgaagccgctactggtagtggatgaggtgcagtcaggcttaagaca
gctgatgccccaccctacgatgaggtggcagtaactgctgctaggccaccgagagctaaaccagctccaagacgcag
agcagcacaaccttccgatgcatctcctgctcgaagtcgatcttccactttagggtatagccagcaacaacaggagaaga
tcaaacctaaagtcggaagtacagttgcgcaacacccatgaagcccttgttggctatgggttcactcatgccacatagttgact
atcccaacatcctgctgcttggaaactgttgcggtgaagtaccaggacatgattgctgctttacctgaagcaacacagggc
aatagtcgggttggcaaacagtggtctgagcctagggtctgaagccctcttacagttgcaggagaattgcggggacct
ccttgcaactcagatccggacaattcctgaagattgccaacagctggtggtgcttaccggcagtagaagcagtcagcttggag
aatgcttaactggagacccttgaatctcagccagacaagtggtcctatcgcctcccatgatgtggaaaacaagcact
agaaactgtccaaagattatgctgttcttgcaggcacacggacttaccacaacaagtcgttgcatagcttcccatgatg
gagggaaagcaagcgttagaaacagtcagcggctactccctgtattatgccaggtcatggttaactccacaacaagttg
ggtatagctcagcagatggtgtaaacagggcacttgaaccgtccaagactcctgcccgtctctgccagccacacggc
ctacccccgaacaagtggtggctattgcttgcctatgtaggtaagcaggctttagagacagtcagagactactaccggt
ctatgccaggccatggtttgaccccggaacaggttgttgcctatgctcaataagggcggcaagcaagcgttggaaaccgt
tcaagcattactccctgttctctgcaagcacaatggctaacgcccagcaggttgttcaattgcatcacatgatggagaaa
gcaggccttagaaacggtacaagcactttaccagctccttggcaagcacacggccttacacccgaacaagtggtcctattg
caagtaatataggtggaaaacaagcactggaaaccgtgcaggcgttttggcgttattatgccaagctcagggcctaactct
gaacaggtggttgcgattgctcaaatgttggggtaaacagcagcagtagagactgtgcagcggctttgctgttttgcctcaag
ctcaggttagcaccagagcaggtggtcgtatagctagtaacattggaggtaaacagcagcgttgaaccgtgcaaccgtc
ctgccagttctatgcaagctcatggttgaacccacaacaggtttagcagctcctcaataaaggaggaagcaagcctc
agaaacggtgcagggcctcccgttcttgcaggccatgattgacccggagcaggtggtcgaatcggcagctatg
atggaggtaacgaggccttggaaaccgttcaggcgttactcccgttctatgccagcgcagtgccctgaccctgaacaggtt
gtggcgtatgccagtaacggcgggggaaagcaggcacttgaaccgtacaacgactcctccagctctttgtcaagcccac
ggattgactccagaacaagtagttgctatagcttcaataaaggaggaaagcaggcccttgaacaagctcagcgtctttgcc
gtttgtgcaagcacagcaggtgactcctgaacaggttgcgcttgcctatcaatcgggtgtaagcaagctctcgaaccg
tacagcagctcttgcctgttctatgccaaagcagcaggttgcagccggaaacaggtggtagccatgcaagcaacataggtg
caaacagccttgaacaagctcaaaaggttgccttgccttgcgaagcccagcgtttgacccctcaacaggtggtgctatag
catcacatgatggggagcggcctccttggagcagtcagcggccttgcctgttgcctgaagcgcagtgcttaacaccgg
aacaggtcgtggcaattgcctcaaatatggcggcaacaagcagcgtggaaccgttcagcagccttgcctgttctgtgcca
gctacggctgagcggcccaacaggttgccttgcctcaaatgagggaggaggccagcccttgaagtcgattgtgcaca
gctactcggcccgaccctcgtttagcggctcgaacaatgcatcttgcctcctcggcctgcttgaaggtcggcccaagcttaga
cgagtaaaaaaggtctaccctatgctcggccttaatacagagcagcaatcgtagaatcccagaaacgaacgacccatcg
cgtagccgacacgcctcaagttgtaggggttttaggtttttcagtgctcattcacatccggcacaagccttgcagatgacc
cagtttggtatgcaaggcatggattactcgaactttcagaagagtaggagtgacagagctcgaagccagaagcggaaactc
gccacccgctagccaaaagatggatagatattgcaggcagtggaatgaagcgcgcaaacctatccaacaagcactc
aaacccggatcaagcagcttgcagcgttgcgagattctcgaacgagattggatgcccttctccaatgcacgaaggtg
atcaaaactagggcagtagcaggaagaggtctaggagtagcgtgcaattacggcccctcagcacaacagcttttaggt
caggtgccagaacaagggagcgtttacatctcccattgtctggcgtgtaaaaaagccggcaactatgtagggggg
ttaccggaccagggaccctcctcctgctcctagctccttctagctccttctagctccttctagctccttctagctccttct
ggggcagctgatgacttcccgcattcaacgaagaagaattagcatggtgatgaggttactgccacagtaa
gctgtcctgctttaatgagatagcagaagcctatgatcagatgatttgccttcaattctgttgcaggttgaataaacctga
gcatgtgtagctcagatcctaccgcccgttgcggttcaattctaataatataaccggctactatcgtattttatgaataatctc
cgttcaattactgattgacctactatctatgtacaatataaaatgaaacaataatgtgctgaataggtttatagcagat
ctatgatagagcgcacaataacaacaattgcgtttatttacaataccaatttataaaaaagcggcagaaccggtaaac
ctaaaagactgattacataaatcttcaaatccaagtgccccaggggctagtatctacgacacaccgagcggcgaact
aataacgctcactgaagggaactccggttccccgcccggcgcagtggtgagattcctgaagttgagttggcgtcggctc
taccgaaagttacgggcaccattcaaccggctcagcagcggcggcgggtaaccgacttgcctcccggagattatgcagc
atttttggtgatgtggcccccaatgaagtgcaggtcaaaccttgacagtgacgacaatcgttgggcccgtccagggcga
atttgcgacaacatgctcagggctcagcagac

Referencias

- 5 Bikard D, et al. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.* 2013.
- Blazeck, J. y Alper, H.S. (2013) Promoter engineering: recent advances in controlling transcription at the most fundamental level. *Biotechnol J*, 8, 46-58.
- 10 Blount, B.A., Weenink, T., Vasylechko, S. y Ellis, T. (2012) Rational diversification of a promoter providing fine-tuned expression and orthogonal regulation for synthetic biology. *PLoS One*, 7, e33279.
- Boch, J. y Bonas, U. (2010) *Xanthomonas AvrBs3 Family-Type III Effectors: Discovery and Function.* *Annu. Rev. Phytopathol.*, 48, 419-436.
- 15 Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A. y Bonas, U. (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*, 326, 1509-1512.
- Brophy, J.A.N. y Voigt, C.A. (2014) Principles of genetic circuit design. *Nat. Methods*, 11, 508-520.
- 20 Brückner, K. y Tissier, A. (2013) High-level diterpene production by transient expression in *Nicotiana benthamiana*. *Plant Methods*, 9, 46.
- de Lange, O., Binder, A. y Lahaye, T. (2014) From dead leaf, to new life: TAL effectors as tools for synthetic biology. *The Plant Journal*, 78, 753-771.
- 25 Engler, C., Kandzia, R. y Marillonnet, S. (2008) A one pot, one step, precision cloning method with high throughput capability. *PLoS One*, 3, e3647.
- 30 Engler, C., Youles, M., Gruetzner, R., Ehnert, T.M., Werner, S., Jones, J.D., Patron, N.J. y Marillonnet, S. (2014) A Golden Gate Modular Cloning Toolbox for Plants. *ACS synthetic biology*, **Publicación electrónica antes de impresión.**
- 35 Ennajdaoui, H., Vachon, G., Giacalone, C., Besse, I., Sallaud, C., Herzog, M. y Tissier, A. (2010) Trichome specific expression of the tobacco (*Nicotiana glauca*) cembratrien-ol synthase genes is controlled by both activating and repressing cis-regions. *Plant Mol. Biol.*, 73, 673-685.
- Fineran, P.C. y Dy, R.L. (2014) Gene regulation by engineered CRISPR-Cas systems. *Curr. Opin. Microbiol.*, 18, 83-89.
- 40 Gilbert LA, et al. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell.* 2013.
- Hahn, S. y Young, E.T. (2011) Transcriptional Regulation in *Saccharomyces cerevisiae*: Transcription Factor Regulation and Function, Mechanisms of Initiation, and Roles of Activators and Coactivators. *Genetics*, 189, 705-736.
- 45 Hsu, P.D., Lander, E.S. y Zhang, F. (2014) Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell*, 157, 1262-1278.
- Jefferson, R.A., Kavanagh, T.A. y Bevan, M.W. (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *Embo j*, 6, 3901-3907.
- 50 Jinek, M. et al., (2012) *Science* 337, 816-821.
- Kay, S., Hahn, S., Marois, E., Hause, G. y Bonas, U. (2007) A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science*, 318, 648-651.
- 55 Kay, S., Hahn, S., Marois, E., Wieduwild, R. y Bonas, U. (2009) Detailed analysis of the DNA recognition motifs of the *Xanthomonas* type III effectors AvrBs3 and AvrBs3 Δ rep16. *The Plant Journal*, 59, 859-871.
- 60 Koncz, C. y Schell, J. (1986) The Promoter of TI-DNA Gene 5 Controls the Tissue-Specific Expression of Chimeric Genes Carried by a Novel Type of *Agrobacterium* Binary Vector. *Mol. Gen. Genet.*, 204, 383-396.
- Liu, D., Shi, L., Han, C., Yu, J., Li, D., Zhang, Y. (2012) Validation of reference genes for gene expression studies in virus-infected *Nicotiana benthamiana* using quantitative real-time PCR. *PLoS One*, 7, e46451.

- Liu, Q., Manzano, D., Tanic, N., Pesic, M., Bankovic, J., Pateraki, I., Ricard, L., Ferrer, A., de Vos, R., van de Krol, S. y Bouwmeester, H. (2014) Elucidation and in planta reconstitution of the parthenolide biosynthetic pathway. *Metab Eng*, 23, 145-153.
- 5 Livak, K. J., Schmittgen, T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, 25, 402-408.
- Lubliner, S., Keren, L. y Segal, E. (2013) Sequence features of yeast and human core promoters that are predictive of maximal promoter activity. *Nucleic Acids Res.*, 41, 5569-5581.
- 10 Mahfouz, M., Li, L., Piatek, M., Fang, X., Mansour, H., Bangarusamy, D. y Zhu, J.-K. (2012) Targeted transcriptional repression using a chimeric TALE-SRDX repressor protein. *Plant Mol. Biol.*, 78, 311-321.
- 15 Maicas, E. y Friesen, J.D. (1990) A sequence pattern that occurs at the transcription initiation region of yeast RNA polymerase II Promoters. *Nucleic Acids Res.*, 18, 3387-3393.
- Nissim, L., Perli, S.D., Fridkin, A., Perez-Pinera, P. y Lu, T.K. (2014) Multiplexed and Programmable Regulation of Gene Networks with an Integrated RNA and CRISPR/Cas Toolkit in Human Cells. *Mol. Cell*, 54, 698-710.
- 20 Peremarti, A., Twyman, R.M., Gomez-Galera, S., Naqvi, S., Farre, G., Sabalza, M., Miralpeix, B., Dashevskaya, S., Yuan, D.W., Ramessar, K., Christou, P., Zhu, C.F., Bassie, L. y Capell, T. (2010) Promoter diversity in multigene transformation. *Plant Mol. Biol.*, 73, 363-378.
- Perez-Pinera, P., Ousterout, D.G., Brunger, J.M., Farin, A.M., Glass, K.A., Guilak, F., Crawford, G.E., Hartemink, A.J. y Gersbach, C.A. (2013) Synergistic and tunable human gene activation by combinations of synthetic transcription factors. *Nat Meth*, 10, 239-242.
- 25 Potenza, C., Aleman, L. y Sengupta-Gopalan, C. (2004) Targeting transgene expression in research, agricultural, and environmental applications: Promoters used in plant transformation. *In Vitro Cell Dev-Pl*, 40, 1-22.
- 30 Qi, L., Haurwitz, R.E., Shao, W.J., Doudna, J.A. y Arkin, A.P. (2012) RNA processing enables predictable programming of gene expression. *Nat. Biotechnol.*, 30, 1002-1006.
- Romer, P., Hahn, S., Jordan, T., Strauss, T., Bonas, U. y Lahaye, T. (2007) Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. *Science*, 318, 645-648.
- 35 Ting, H.M., Wang, B., Rydén, A.M., Woittiez, L., van Herpen, T., Verstappen, F.W., Ruyter-Spira, C., Beekwilder, J., Bouwmeester, H.J. y van der Krol, A. (2013) The metabolite chemotype of *Nicotiana benthamiana* transiently expressing artemisinin biosynthetic pathway genes is a function of CYP71AV1 type and relative gene dosage. *New Phytol.*, 199, 352-366.
- Tissier, A. (2012) Trichome specific expression: promoters and their applications. In *Transgenic Plants, advances and limitations* (Çiftçi, Y.O. ed: InTech, pp. 353-378.
- 45 Untergrasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S.G. (2012) Primer3 - new capabilities and interfaces. *Nucleic Acids Research*, 50, e115.
- van der Oost, J., Westra, E.R., Jackson, R.N. y Wiedenheft, B. (2014) Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nature Reviews Microbiology*, 12, 479-492.
- 50 Vranova, E., Coman, D. y Grissem, W. (2012) Structure and Dynamics of the Isoprenoid Pathway Network. *Mol. Plant.*, 5, 318-333.
- Wang, E. y Wagner, G.J. (2003) Elucidation of the functions of genes central to diterpene metabolism in tobacco trichomes using posttranscriptional gene silencing. *Planta*, 216, 686-691.
- 55 Weber, E., Gruetzner, R., Werner, S., Engler, C. y Marillonnet, S. (2011) Assembly of Designer TAL Effectors by Golden Gate Cloning. *PLoS One*, 6, e19722.
- 60 Werner, S., Engler, C., Weber, E., Gruetzner, R. y Marillonnet, S. (2012) Fast track assembly of multigene constructs using Golden Gate cloning and the MoClo system. *Bioeng Bugs*, 3, 38-43.

LISTA DE SECUENCIAS

	<110> Leibniz-Institut für Pflanzenbiochemie	
5	<120> Biblioteca de promotores sintéticos para la expresión génica coordinada en células u organismos eucarióticos	
	<130> EAD-16714	
10	<160> 52	
	<170> PatentIn versión 3.5	
15	<210> 1	
	<211> 97	
	<212> ADN	
	<213> Secuencia Artificial	
20	<220>	
	<223> secuencia promotora	
	<400> 1	
	ggaggttatc acacttggtgta atttccccgc atagctgaac atctatacaa cactcatcgc	60
	gcgagaaaac gcgcaacaaa ttggaggcgc attaatg	97
25	<210> 2	
	<211> 97	
	<212> ADN	
	<213> Secuencia Artificial	
30	<220>	
	<223> secuencia promotora	
	<400> 2	
	ggagcttacg tagcagcacg gcctccccgc atagctgaac atctatataa atcttagccc	60
	gactgcaccg tgaggaaaag agggtttata tataatg	97
35	<210> 3	
	<211> 98	
	<212> ADN	
	<213> Secuencia Artificial	
40	<220>	
	<223> secuencia promotora	
	<400> 3	
	ggaggtcgtc aatgataag tactccccgc atagctgaac atctatataa cgcagcaaat	60
45	ggttgaaacg tattagcaac agctaaacct tcgaatgt	98
50	<210> 4	
	<211> 97	
	<212> ADN	
	<213> Secuencia Artificial	
50	<220>	
	<223> secuencia promotora	
	<400> 4	
	ggaggcgtgt cgtttttagtg aggtccccgc atagctgaac atctatataa ggttttgcta	60
55	ttcattgaaa gcagtagtga ctgatttgta tataatg	97
	<210> 5	

ES 2 743 512 T3

	<211> 96	
	<212> ADN	
	<213> Secuencia Artificial	
5	<220>	
	<223> secuencia promotora	
	<400> 5	
	ggaggttctg tgggtttttt gggccccgc atagctgaac atctatataa tgagaaacag	60
10	gggcaaatcc cggaaaacaa ataacgcgaa gcaatg	96
	<210> 6	
	<211> 97	
	<212> ADN	
	<213> Secuencia Artificial	
15	<220>	
	<223> secuencia promotora	
	<400> 6	
	ggagatgtgc catggtcggg ggctccccgc atagctgaac atctatataa tcggaatgaa	60
20	ttgcaaatag ggatgttata gggttatctc ggcaatg	97
	<210> 7	
	<211> 97	
	<212> ADN	
	<213> Secuencia Artificial	
25	<220>	
	<223> secuencia promotora	
	<400> 7	
	ggagagatga ttgttcgtaa tattccccgc atagctgaac atctatataa gagacgagct	60
30	accacgataa gcgcagcaaa agacacatac gttaatg	97
	<210> 8	
	<211> 97	
	<212> ADN	
	<213> Secuencia Artificial	
35	<220>	
	<223> secuencia promotora	
	<400> 8	
	ggaggtgcct ttttaaaagg gtttccccgc atagctgaac atctatataa gaaacacgga	60
40	cgccgaaacc cccccgccta gttgactctc cggaatg	97
	<210> 9	
	<211> 96	
	<212> ADN	
	<213> Secuencia Artificial	
45	<220>	
	<223> secuencia promotora	
	<400> 9	
	gagtgatgaac gtggatgatg ggtccccgca tagctgaaca tctatataaa aactgtaaca	60
50	gaactgctgc tcaactagtc tccaactata taaatg	96
55	<210> 10	

ES 2 743 512 T3

	<211> 97	
	<212> ADN	
	<213> Secuencia Artificial	
5	<220>	
	<223> secuencia promotora	
	<400> 10	
	ggaggttaagt tctggcctct gggccccgc atagctgaac atctatataa taggacagta	60
	agatatgccca gacaatgtag cacgtagaac ggaaatg	97
10	<210> 11	
	<211> 96	
	<212> ADN	
	<213> Secuencia Artificial	
15	<220>	
	<223> secuencia promotora	
	<400> 11	
	ggagcacctt ggaaaacagg tggccccgc atagctgaac atctatataa gaaattagta	60
20	actacgcttt caaagacatc ttaattaatt agaatg	96
	<210> 12	
	<211> 97	
	<212> ADN	
25	<213> Secuencia Artificial	
	<220>	
	<223> secuencia promotora	
30	<400> 12	
	ggagagtggg cgcgtttggg aagccccgc atagctgaac atctatataa ataagctgcc	60
	tacgatttag cgcgacacaa ttcgacgggt ggaaatg	97
	<210> 13	
	<211> 97	
35	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
	<223> secuencia promotora	
40	<400> 13	
	ggagggcgtc gggggggggt ctgtccccgc atagctgaac atctatataa taccggttat	60
	ccagcgacaa cgtcaataa cttaacattg aacaatg	97
	<210> 14	
45	<211> 97	
	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
50	<223> secuencia promotora	
	<400> 14	
	ggagttaagt aggctgttct tattccccgc atagctgaac atctatataa atagctcagt	60
	tgccttaaac gtagacagtc tattacacta ataaatg	97
55	<210> 15	

ES 2 743 512 T3

<211> 97
 <212> ADN
 <213> Secuencia Artificial

5

<220>
 <223> secuencia promotora

<400> 15
 ggaggtttat actttctcta atttccccgc atagctgaac atctatataa actggaactt 60
 tcactgacct gacagtgcct tcattgacgc gcgaatg 97

10

<210> 16
 <211> 97
 <212> ADN
 <213> Secuencia Artificial

15

<220>
 <223> secuencia promotora

<400> 16
 ggagaagctc acctgcatcc ttgtccccgc atagctgaac atctatataa aatttataca 60
 tttaaaccaa accctataaa atttctgtcc gcgaatg 97

20

<210> 17
 <211> 97
 <212> ADN
 <213> Secuencia Artificial

25

<220>
 <223> secuencia promotora

30

<400> 17
 ggagtaatgt acagcttcgc ggctccccgc atagctgaac atctatataa taactgttct 60
 acatactctc ggataaataa cgcaacttaa tataatg 97

35

<210> 18
 <211> 97
 <212> ADN
 <213> Secuencia Artificial

40

<220>
 <223> secuencia promotora

<400> 18
 ggagtcccg atcctggttt cactccccgc atagctgaac atctatataa aacaaatctc 60
 gtcaagcaaa ccaattacga tggcccgga aggaatg 97

45

<210> 19
 <211> 97
 <212> ADN
 <213> Secuencia Artificial

50

<220>
 <223> secuencia promotora

<400> 19
 ggaggggtgg gggggttacc gtctccccgc atagctgaac atctatataa aaccaatcga 60
 cgactcattg gaatgcgtcc tgctccaaat agcaatg 97

55

<210> 20

ES 2 743 512 T3

<211> 97
 <212> ADN
 <213> Secuencia Artificial

5

<220>
 <223> secuencia promotora

<400> 20
 ggagctgcaa tgcttactgt tattccccgc atagctgaac atctatataa aggttgagaa 60
 gtggcacttg caataattgc cgggcccatg ggaaatg 97

10

<210> 21
 <211> 97
 <212> ADN
 <213> Secuencia Artificial

15

<220>
 <223> secuencia promotora

<400> 21
 ggagagaggt gggtcgcggg cattccccgc atagctgaac atctatataa aaagcagttc 60
 atacgggaac gcagactagt cacgaaatc cgcaatg 97

20

<210> 22
 <211> 97
 <212> ADN
 <213> Secuencia Artificial

25

<220>
 <223> secuencia promotora

30

<400> 22
 ggagggcgcgt ttgattgggg gggtccccgc atagctgaac atctatataa aactgataac 60
 ttagtcagac tagcaaatgc tcacaataca aagaatg 97

35

<210> 23
 <211> 97
 <212> ADN
 <213> Secuencia Artificial

40

<220>
 <223> secuencia promotora

<400> 23
 ggaggaaccg cggttcggga gcgtccccgc atagctgaac atctatataa caggggcggg 60
 atatgatata gaaacaccag ggtagtaac aagaatg 97

45

<210> 24
 <211> 97
 <212> ADN
 <213> Secuencia Artificial

50

<220>
 <223> secuencia promotora

<400> 24
 ggagcaatgg cagaggggta tgggtccccgc atagctgaac atctatataa aattgggacc 60
 tatgataatg ctgtagaacc tcataaatga tggaatg 97

55

<210> 25

ES 2 743 512 T3

<211> 97
 <212> ADN
 <213> Secuencia Artificial

5

<220>
 <223> secuencia promotora

<400> 25
 ggagggtttg ttaccgtgta ttctccccgc atagctgaac atctatataa cgagcggctcg 60
 aaagcttcaa aattcaatga gtaagacccc aacaatg 97

10

<210> 26
 <211> 97
 <212> ADN
 <213> Secuencia Artificial

15

<220>
 <223> secuencia promotora

<400> 26
 ggagtcttgc ccatactcag ttctccccgc atagctgaac atctatataa tgtctctcga 60
 atgacaccga tacttcgatt tagcaggctg cccaatg 97

20

<210> 27
 <211> 97
 <212> ADN
 <213> Secuencia Artificial

25

<220>
 <223> secuencia promotora

30

<400> 27
 ggaggcatta cgggcgtagt ttgtccccgc atagctgaac atctatataa tcgaccagag 60
 caacgttctc gcctaggcgc gccttaccgg gaaaatg 97

35

<210> 28
 <211> 97
 <212> ADN
 <213> Secuencia Artificial

40

<220>
 <223> secuencia promotora

<400> 28
 ggaggtatcc cgagggcgaa tgggtccccgc atagctgaac atctatataa ggcacatcgg 60
 tagcaagcct caccagccgt gaacaatata atcaatg 97

45

<210> 29
 <211> 97
 <212> ADN
 <213> Secuencia Artificial

50

<220>
 <223> secuencia promotora

<400> 29
 ggagtccagag tgtggctcgtg ttttccccgc atagctgaac atctatataa gtttacacct 60
 gcccgacctt ttccttgatg agctcgcgca ccgaatg 97

55

<210> 30

ES 2 743 512 T3

<211> 97
 <212> ADN
 <213> Secuencia Artificial

5

<220>
 <223> secuencia promotora

<400> 30
 ggagcactta ggtaactgc gcgtccccgc atagctgaac atctatataa ggaacagacg 60
 gctaaccat caatggaaca acgaacggcc cataatg 97

10

<210> 31
 <211> 97
 <212> ADN
 <213> Secuencia Artificial

15

<220>
 <223> secuencia promotora

<400> 31
 ggagtggttt tagcacagac aactccccgc atagctgaac atctatataa ggaaccgact 60
 ctctcccgcc ccgacacag gaccccctct ggtaatg 97

20

<210> 32
 <211> 97
 <212> ADN
 <213> Secuencia Artificial

25

<220>
 <223> secuencia promotora

30

<400> 32
 ggagcatgct ggggttgag ggggtccccgc atagctgaac atctatataa cccggcccct 60
 caattaacc tgctcaaagc caaagcaaca atcaatg 97

35

<210> 33
 <211> 97
 <212> ADN
 <213> Secuencia Artificial

40

<220>
 <223> secuencia promotora

<400> 33
 ggagtgggct gagggggacg gggtccccgc atagctgaac atctatataa cccaacgtga 60
 gatcacgttt acaataactcg tcttgactga cacaatg 97

45

<210> 34
 <211> 98
 <212> ADN
 <213> Secuencia Artificial

50

<220>
 <223> secuencia promotora

<400> 34
 ggagggcatg aggttagcta gtttccccgc atagctgaac atctatataa tgaatgtgtg 60
 agaagctttg acaagcttcg cccagcggcc ctaaattg 98

55

<210> 35

ES 2 743 512 T3

	<211> 97	
	<212> ADN	
	<213> Secuencia Artificial	
5	<220>	
	<223> secuencia promotora	
	<400> 35	
	ggagatattc ctgggctgtc acttccccgc atagctgaac atctatataa tcaaattcct	60
	aggggctgcc atggttaccc cttgctatgt agcaatg	97
10	<210> 36	
	<211> 96	
	<212> ADN	
	<213> Secuencia Artificial	
15	<220>	
	<223> secuencia promotora	
	<400> 36	
	gagtgcgggg agttggtatg ggtccctgca tagctgaaca totatataaa taacggcct	60
20	gagtgtcgct agcaaatgca ccaaaccaac tcaatg	96
	<210> 37	
	<211> 97	
	<212> ADN	
25	<213> Secuencia Artificial	
	<220>	
	<223> secuencia promotora	
30	<400> 37	
	ggagcgatat gtttaatcgg cattccccgc atagctgaac atctatataa atccgcgcct	60
	taaatacggg tgtttgccta ttttgagcga gtaaatg	97
	<210> 38	
	<211> 97	
35	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
	<223> secuencia promotora	
40	<400> 38	
	ggagtattgg tccacttgct ccgtccccgc atagctgaac atctatataa cgcagaaacc	60
	tcgcaactga cgttccgatt gcgataaaca cacaatg	97
	<210> 39	
45	<211> 97	
	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
50	<223> secuencia promotora	
	<400> 39	
	ggagttgtaa actccagggt agatccccgc atagctgaac atctatataa atcgcccatt	60
	gcccaggaca cgcaaccctg caacaatcag tacaatg	97
55	<210> 40	

ES 2 743 512 T3

<211> 97
 <212> ADN
 <213> Secuencia Artificial

5

<220>
 <223> secuencia promotora

<400> 40
 ggagcccttg ctatgacgat gcctccccgc atagctgaac atctatataa aaattccata 60
 acgtacagaa cgaagaacgg acctatatac aataatg 97

10

<210> 41
 <211> 97
 <212> ADN
 <213> Secuencia Artificial

15

<220>
 <223> secuencia promotora

<400> 41
 ggagggcggg gcaggggttg ggttccccgc atagctgaac atctatataa tcaccaacg 60
 ccgcatgcaa caactagtaa caaggccatt gagaatg 97

20

<210> 42
 <211> 97
 <212> ADN
 <213> Secuencia Artificial

25

<220>
 <223> secuencia promotora

30

<400> 42
 ggagcttttt acggggcacc tgttccccgc atagctgaac atctatataa caaaataaca 60
 gattgtctgt ccaagaggca gacattcgtt cataatg 97

35

<210> 43
 <211> 19
 <212> ADN
 <213> Secuencia Artificial

40

<220>
 <223> cebador

<400> 43
 cgatccaaag atgggacga 19

45

<210> 44
 <211> 26
 <212> ADN
 <213> Secuencia Artificial

50

<220>
 <223> cebador

<400> 44
 catctctttt ttcgcatagt agatgt 26

55

<210> 45
 <211> 27
 <212> ADN
 <213> Secuencia Artificial

ES 2 743 512 T3

<220>
 <223> cebador

 5 <400> 45
 cctccacaga gacattctta tagtttc 27

 <210> 46
 <211> 24
 <212> ADN
 10 <213> Secuencia Artificial

 <220>
 <223> cebador

 15 <400> 46
 tcaaagtcaa acttaggcaa gatg 24

 <210> 47
 <211> 18
 20 <212> ADN
 <213> Secuencia Artificial

 <220>
 <223> cebador

 25 <400> 47
 agcatctccg ctggtctt 18

 <210> 48
 <211> 20
 30 <212> ADN
 <213> Secuencia Artificial

 <220>
 <223> cebador

 35 <400> 48
 cgagtcaagg aatcctgcat 20

 40 <210> 49
 <211> 57
 <212> ADN
 <213> Secuencia Artificial

 45 <220>
 <223> Talpro7

 <220>
 <221> rasgo_misc
 50 <222> (19)..(37)
 <223> n es a, c, g, or t

 <400> 49
 55 ttggtctca acatggagnn nnnnnnnnnn nnnnnntcc ccgcatagct gaacatc 57

 <210> 50
 <211> 87
 <212> ADN
 60 <213> Secuencia Artificial

 <220>
 <223> Talpro8

 <220>

ES 2 743 512 T3

<221> rasgo_misc
 <222> (19)..(61)
 <223> n es a, c, g, or t

5

<400> 50
 ttgtgtctca acaacattnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 60
 nttatataga tgttcagcta tgcggggg 87

10

<210> 51
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

15

<220>
 <223> efector consenso TAL

20

<220>
 <221> rasgo_misc
 <222> (12)..(13)
 <223> Xaa puede ser cualquier aminoácido de origen natural

<400> 51
 Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Xaa Xaa Gly Gly Lys
 1 5 10 15
 Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala
 20 25 30
 His Gly

25

<210> 52
 <211> 5087
 <212> ADN
 <213> Secuencia Artificial

30

<220>
 <223> pICH74043

<400> 52
 tcgacaaaat ttagaacgaa ctttaattatg atctcaaata cattgatata tatctcatct 60
 agatctaggt tatcattatg taagaaagtt ttgacgaata tggcagcaca aatgggctag 120
 actcgatgta attggtatct caactcaaca ttatacttat accaaacatt agttagacaa 180
 aatttaaca actatTTTTT atgtatgcaa gagtcagcat atgtataatt gattcagaat 240

ES 2 743 512 T3

cgttttgacg agttcggatg tagtagtagc cattatthaa tgtacatact aatcgtgaat 300
 agtgaatatg atgaaacatt gtatcttatt gtataaatat ccataaacac atcatgaaag 360
 acactttctt tcacggctctg aattaattat gatacaattc taatagaaaa cgaattaat 420
 tacgttgaat tgtatgaaat ctaattgaac aagccaacca cgacgacgac taacgttgcc 480
 tggattgact cggtttaagt taaccactaa aaaaacggag ctgtcatgta acacgaggat 540
 cgagcaggtc acagtcatag agccatcaaa gcaaaagaac taatccaagg gctgagatga 600
 ttaattagtt taaaaattag ttaaacgag gggaaaaggct gtctgacagc caggtcacgt 660
 tatctttacc tgtggtcgaa atgattcgtg tctgtcgtt ttaattattt ttttgaaagg 720
 ccgaaaataa agttgtaaga gataaacccg cctatataaa ttcataatatt ttcctctccg 780
 ctttgaatac tgtataagag ctcatthtta caacaattac caacaacaac aaacaacaaa 840
 caacattaca attacattta caattatcga tacaatggat ccgattcgtt ctgggacacc 900
 ttctcccgtc cgagagctcc ttctggacc acaaccagat ggagtgaac caacagctga 960
 taggggtgta tctctccag ccggggacc tcttgacggc ttacctgcaa ggcgcacaat 1020
 gtcccgtacc agattgccc agtccaccagc accaagtcca gcgttttcag cgggcagctt 1080
 ctctgacctc ctgagacagt ttgatccctc tttgttcaac acctcactgt ttgactcact 1140
 tccgcctttt ggggctcacc aactgaagc cgctactggg gagtgggatg aggtgcagtc 1200
 aggttaaga gcagctgat ccccaccacc tacgatgagg gtggcagtaa ctgctgctag 1260
 gccaccgaga gctaaaccag ctccaagacg cagagcagca caacctccg atgcatctcc 1320
 tgcctgctcaa gtcgatcttc gcactttagg gtatagccag caacaacagg agaagatcaa 1380
 acctaaagtc cgaagtacag ttgcgcaaca ccatgaagcc cttgttggtc atgggttcac 1440
 tcatgccac atagttgcac tatcccaaca tcctgctgct cttggaactg ttgcggtgaa 1500
 gtaccaggac atgattgctg ctttacctga agcaacacac gaggcaatag tcggtgttg 1560
 caaacagtggt tctggcgtca gggctctaga agccctcctt acagttgcag gagaattgag 1620
 gggacctccc ttgcagctcg ataccggaca attgctgaag attgccaaac gtggtggtgt 1680
 tacggcagta gaagcagtc atgcttgag gaatgctcta actggagcac ccttgaatct 1740
 cacgccagaa caagtggctc ctatcgcctc ccatgatggg ggaaaacaag cactagaaac 1800
 tgtccaaaaga ttattgctg ttctttgtca ggcacacgga cttacccac aacaagtct 1860
 tgetatagcc tcccatgatg gaggaagca agcgttagaa acagtgacgc ggctactccc 1920
 tgtattatgc caggctcatg gtctaactcc acaacaagtg gtggctatag cctcacagca 1980
 tgggtgtaaa caggcacttg aaaccgtcca aagactcctg ccggtcctct gccaggcaca 2040
 cggcctcacc cccgaacaag tgggtgctat tgcttcgcat gatggaggta agcaggcttt 2100
 agagacagtc cagagactac taccggttct atgccaggcc catggtttga ccccgaaca 2160

ES 2 743 512 T3

gggtgttgct attgctcaa ataagggcgg caagcaagcg ttggaaccg ttcaagcatt 2220
 actccctggt ctctgtcaag cacatgggct aacgcccgag cagggtgttg caattgcatc 2280
 acatgatgga ggaaagcagg ccttagaaac ggtacaggca cttttaccag tcctttgcc 2340
 agcacacggg cttacacccg aacaagtggc cgctattgca agtaatatag gtggaaaaca 2400
 agcactggaa accgtgcagg cgcttttgcc ggtattatgc caagctcacg gcctaactcc 2460
 tgaacaggtg gttgcgattg cctcaaatgg tgggggtaaa caggcactgg agactgtgca 2520
 gcggcttttg cctgttttgt gtcaagctca tggattgaca ccagagcagg tggtcgctat 2580
 agctagtaac attggaggtg aacaagcgct tgaaacctg caacgtctgc tgccagttct 2640
 atgtcaagct catgggttga cccacaaca ggtttagcgg atcgcttcca ataaaggagg 2700
 aaagcaagct ctagaaacgg tgcagaggct cctcccgggt ctttgtcagg cgcatggatt 2760
 gaccccgagg cagggtgtcg caatcgccag tcatgatgga ggtaagcagg ccttgaaac 2820
 cgttcaggcg ttactccgg ttctatgcc ggcgatggc ctgacccctg aacaggttgt 2880
 ggcgatagcc agtaacggcg ggggaaagca ggcacttga accgtacaac gactcctccc 2940
 agtcctttgt caagcccacg gattgactcc agaacaagta gttgctatag cttcgaataa 3000
 gggaggaaag caggcccttg aaacagttca gcgtcttttg ccagtgttgt gtcaagcaca 3060
 cggattgact cctgaacagg ttgtcgccat tgcactaat atcgggtgta agcaagctct 3120
 cgaaaccgta cagcgactct tgctgttct atgccaaagc catggcttga cgccggaaca 3180
 ggtggtagcc atagcaagca acataggtgg caaacaagct cttgaaacag ttcaaaggtt 3240
 gttacctgtg ctttgccaag cccacggttt gaccctcaa cagggtggtg ctatagcatc 3300
 acatgatggg ggacggcctg ctcttgagac agtgcagcgc ctggtgcccg tgttgtgtca 3360
 agcgcatggc ttaacaccgg aacaggtcgt ggcaattgcg tcaaatttg gcggcaaaca 3420
 agcgctggaa accgttcagc gactcttgcc tgttctgtgc caagctcacg gtctgacgcc 3480
 ccaacaggtt gttgccattg cttcaaatgg aggaggagg ccagcccttg agtcgattgt 3540
 cgcacagcta tctcggccc accctgcttt agccgctctg acaaatgatc atcttgtggc 3600
 tctcgcctgc ttaggaggtc gccagcttt agacgcagta aaaaagggtc tacctcatgc 3660
 tccggcctta atcaagagga cgaatcgtag aatcccagaa cgaacgagcc atcgcgtagc 3720
 cgatcacgct caagttgtta gggttttag ttttttcag tgcattcac atccggcaca 3780
 agctttcgat gatgccatga cccagtttg tatgtcaagg catggattac tgcaactttt 3840
 cagaagagta ggagtgacag agctcgaagc cagaagcgga actctgccac ccgctagcca 3900
 aagatgggat aggatattgc agcgagtggt aatgaagcgc gcgaaaccat ctccaacaag 3960
 cactcaaacc ccgatcaag cgagtttgca cgctttcgca gattctctcg aacgagattt 4020

ES 2 743 512 T3

ggatgccct tctccaatgc acgaaggtga tcaaactagc gcgagtagca ggaagaggtc	4080
taggagtgat cgtgcagtta cgggcccctc agcacaacag tcttttgagg tcagggtgcc	4140
agaacaaagg gacgctttac atctcccatt gtcttggcgt gtaaaaaggc cgcgaaactag	4200
tattggaggg ggattaccgg acccagggac ccccactgct gctgatctag ctgcttctag	4260
tacggtaatg cgcgagcaag acgaggatcc atttgctggg gcagctgatg acttccccgc	4320
attcaacgaa gaagaattag catggttgat ggagttactg ccacagtaag cttgtcctgc	4380
tttaatgaga tatgcgagaa gcctatgatc gcatgatatt tgctttcaat tctgttgtgc	4440
acgttgtaaa aaacctgagc atgtgtagct cagatcctta ccgccggttt cggttcattc	4500
taatgaatat atcaccctgt actatcgtat ttttatgaat aatattctcc gttcaattta	4560
ctgattgtac cctactactt atatgtacaa tattaaaatg aaaacaatat attgtgctga	4620
ataggtttat agcgacatct atgatagagc gccacaataa caaacaattg cgttttatta	4680
ttacaaatcc aattttaaaa aaagcggcag aaccgggtcaa acctaaaaga ctgattacat	4740
aaatcttatt caaatattcaa aagtgcctca ggggctagta tctacgacac accgagcggc	4800
gaactaataa cgctcactga agggaactcc ggttccccgc cggcgcgcat gggtgagatt	4860
ccttgaagtt gagtattggc cgtccgctct accgaaagtt acgggcacca ttcaaccggg	4920
tccagcacgg cggccgggta accgacttgc tgccccgaga attatgcagc atttttttgg	4980
tgtatgtggg ccccaaatga agtgcaggtc aaaccttgac agtgacgaca aatcgttggg	5040
cgggtccagg gcgaattttg cgacaacatg tcgaggctca gcaggac	5087

REIVINDICACIONES

- 5 1. Un procedimiento para expresar dos o más secuencias de ácido nucleico de interés en células eucarióticas, en células o tejidos de un organismo eucariótico, o en un organismo eucariótico excluyendo seres humanos, que comprende:
- 10 (a) proporcionar células eucarióticas, tejido de un organismo eucariótico o un organismo eucariótico excluyendo seres humanos que comprenden dos o más construcciones de ácido nucleico de interés, comprendiendo cada construcción de ácido nucleico un promotor y, aguas abajo del mismo, una secuencia de ácido nucleico de interés para ser expresada, en donde cada promotor de dichas construcciones de ácido nucleico de interés se selecciona de una biblioteca de múltiples promotores que comprende cada uno, en la dirección 5' a 3',
- 15 (i) un sitio de unión al efector que comprende al menos 10 bases contiguas, en donde las secuencias de bases de los sitios de unión al efector de los promotores múltiples tienen una identidad de secuencia de al menos 75%,
(ii) una caja TATA, y
(iii) un segmento de secuencia aguas abajo que comprende al menos 20 bases contiguas, en donde los segmentos de secuencia aguas abajo de al menos dos construcciones de ácido nucleico de interés difieren, y
- 20 (b) proporcionar a las células eucarióticas, el tejido de un organismo eucariótico o un organismo eucariótico de la etapa anterior un efector capaz de unirse a los sitios de unión al efector de los promotores y activar la transcripción de dichos promotores.
- 25 2. El procedimiento de acuerdo con la reivindicación 1, en donde la etapa (b) comprende:
- 30 transfectar las células eucarióticas, el tejido de un organismo eucariótico o el organismo eucariótico obtenidos en la etapa (a) con una construcción de ácido nucleico que expresa el efector que codifica dicho efector; o introducir dicho efector en dichas células, en células de dicho tejido, o en células de dicho organismo eucariótico obtenidos en la etapa (a).
- 35 3. Célula u organismo eucarióticos, excluyendo seres humanos, que comprende dos o más construcciones de ácido nucleico de interés, comprendiendo cada construcción de ácido nucleico de interés un promotor y, aguas abajo del mismo, una secuencia de ácido nucleico de interés para ser expresada, comprendiendo cada construcción una secuencia nucleica de interés diferente, en donde los promotores de dichas construcciones de ácido nucleico de interés se seleccionan de una biblioteca de múltiples promotores que comprende cada uno, en dirección 5' a 3',
- 40 (i) un sitio de unión al efector que comprende al menos 10 bases contiguas, en donde las secuencias de bases de los sitios de unión al efector de los promotores múltiples tienen una identidad de secuencia de al menos 75%,
(ii) una caja TATA, y
(iii) un segmento de secuencia aguas abajo que comprende al menos 20 bases contiguas, siempre que al menos dos de dichas dos o más construcciones de ácido nucleico de interés tengan promotores que difieran
- 45 en sus segmentos de secuencia aguas abajo,
- que comprende adicionalmente una o varias construcciones de ácido nucleico que expresa el efector que codifica un efector, en donde dicho efector se puede unir a los sitios de unión al efector de los promotores de dichas dos o más construcciones de ácido nucleico de interés y activar la transcripción de dichos promotores.
- 50 4. La célula u organismo eucarióticos de acuerdo con la reivindicación 3, en donde dicho efector comprende un dominio de repetición que comprende unidades de repetición derivadas de un efector de tipo activador de la transcripción (TAL).
- 55 5. La célula u organismo eucarióticos de acuerdo con la reivindicación 3, en donde dicho efector comprende un ARN que tiene complementariedad con un sitio de unión al efector de dichos promotores, y dicha célula u organismo eucarióticos comprenden una construcción de ácido nucleico que codifica dicho ARN.
- 60 6. La célula u organismo eucarióticos de acuerdo con la reivindicación 5, en donde dicho efector comprende un ARN guía (ARNg) que tiene complementariedad con un sitio de unión al efector, preferiblemente dicho ARN efector es un ARN guía único (ARNgu) que comprende un ARNcr y un ARNtracr.
7. La célula u organismo eucarióticos de acuerdo con una cualquiera de las reivindicaciones 3 a 6, en donde una o más de dichas construcciones de ácido nucleico de interés comprenden un segundo sitio de unión al efector o un

sitio de unión al efector adicional aguas arriba de la caja TATA.

5 8. La célula u organismo eucarióticos de acuerdo con una cualquiera de las reivindicaciones 3 a 7, en donde una o más de dichas construcciones de ácido nucleico de interés contienen una construcción de ácido nucleico de interés que codifica un efector capaz de unirse a un segundo sitio de unión a efector o un sitio de unión al efector adicional para activar o reprimir la transcripción de secuencias de un ácido nucleico de interés en (a) una o varias construcciones de ácido nucleico que contienen el segundo sitio de unión al efector o sitio de unión al efector adicional.

10 9. El uso de un sistema o kit para expresar dos o más secuencias de ácido nucleico de interés en un organismo eucariótico excluyendo seres humanos o células o tejidos de un organismo eucariótico y para la ingeniería metabólica de una ruta biosintética que requiere la expresión simultánea y sintonizable de múltiples genes, que comprende:

15 (A) una biblioteca de múltiples promotores, comprendiendo cada uno de dichos múltiples promotores, en dirección de 5' a 3',

20 (i) un sitio de unión al efector que comprende al menos 10 bases contiguas, en donde las secuencias de bases de los sitios de unión al efector de los promotores múltiples tienen una identidad de secuencia de al menos 75%,

(ii) una caja TATA, y

(iii) un segmento de secuencia aguas abajo que comprende al menos 20 bases contiguas,

25 en donde los promotores múltiples difieren en los segmentos de secuencia aguas abajo, y

(B) siendo un efector capaz de unirse a los sitios de unión al efector de dichos promotores múltiples; o una molécula de ácido nucleico que codifica dicho efector.

30 10. El uso de acuerdo con la reivindicación 9, en donde el sitio de unión al efector de los promotores múltiples es un segmento de secuencia de 10 a 30 bases contiguas, preferiblemente de 13 a 27 bases contiguas, más preferiblemente de 16 a 24 bases contiguas; y/o

en donde las secuencias de bases de los sitios de unión al efector de los promotores múltiples tienen una identidad de secuencia de al menos 85%, preferiblemente al menos 90%, más preferiblemente al menos 95%.

35 11. El uso de un sistema o kit para expresar dos o más secuencias de ácido nucleico de interés en un organismo eucariótico, excluyendo seres humanos o células o tejidos de un organismo eucariótico, en donde la expresión se coordina ya que los niveles de expresión de las dos o más secuencias de ácido nucleico de interés está predeterminada por la selección de diferentes promotores de una biblioteca de promotores para lograr los niveles de expresión deseados para cada una de las dos o más secuencias de ácido nucleico de interés, que comprende:

40 (A) una biblioteca de múltiples promotores, comprendiendo cada uno de dichos promotores múltiples, en dirección 5' a 3',

45 (i) un sitio de unión al efector que comprende al menos 10 bases contiguas,

(ii) una caja TATA, y

(iii) un segmento de secuencia aguas abajo que comprende al menos 20 bases contiguas,

50 en donde los promotores múltiples difieren en los segmentos de secuencia aguas abajo; y

en donde las secuencias de bases de los sitios de unión al efector de los múltiples promotores son las mismas en una longitud de

10 bases contiguas, excepto que se pueden intercambiar una o dos bases, o se puede eliminar o añadir una base y se puede intercambiar una base, o

15 bases contiguas, excepto que se pueden intercambiar hasta tres bases, o se puede eliminar o añadir una base y se pueden intercambiar una o dos bases, o

55 20 bases contiguas, excepto que se pueden intercambiar hasta cuatro bases, o se puede eliminar o añadir una base y se pueden intercambiar hasta tres bases y

(B) un efector que es capaz de unirse a los sitios de unión al efector de dichos promotores múltiples; o una molécula de ácido nucleico que codifica dicho efector.

60 12. El uso de acuerdo con una cualquiera de las reivindicaciones 9 a 11, comprendiendo dicha biblioteca de promotores adicionalmente un segmento de secuencia aguas arriba ubicado aguas arriba de dicho sitio de unión al efector, comprendiendo dicho segmento de secuencia aguas arriba comprende al menos 10 bases contiguas, en donde dichos promotores múltiples pueden diferir en dichos segmento secuencia aguas arriba.

13. El uso de acuerdo con una cualquiera de las reivindicaciones 9 a 12, en donde dicha biblioteca está presente en un recipiente, o cada promotor de dicha biblioteca está contenido en un recipiente separado.

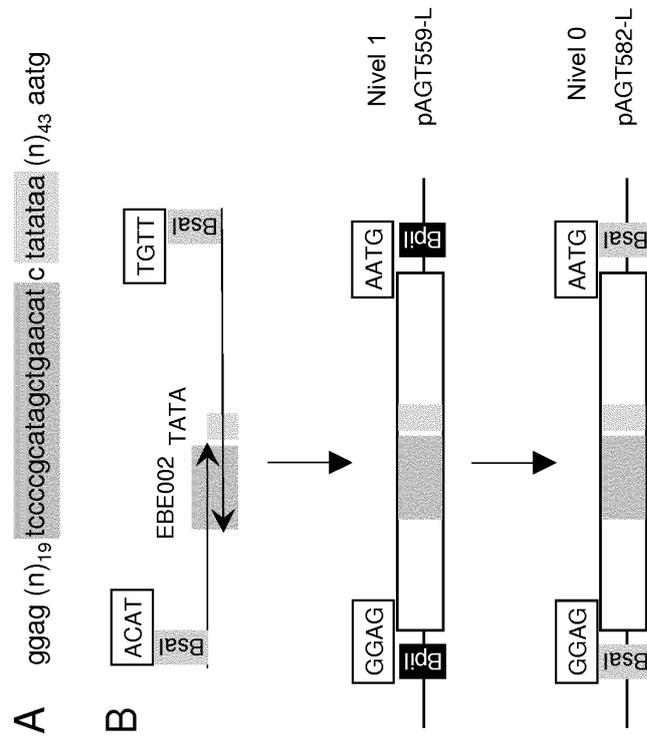


Fig. 1

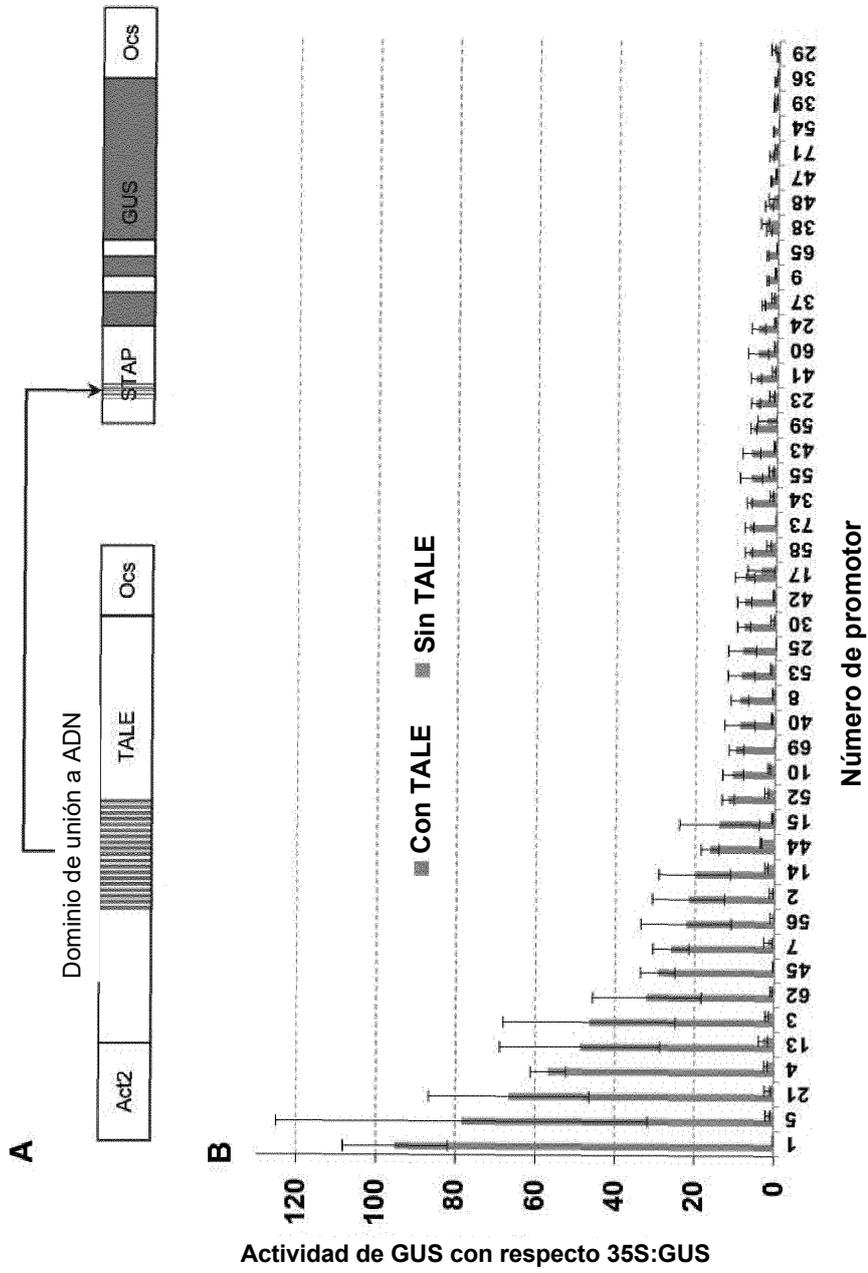


Fig. 2

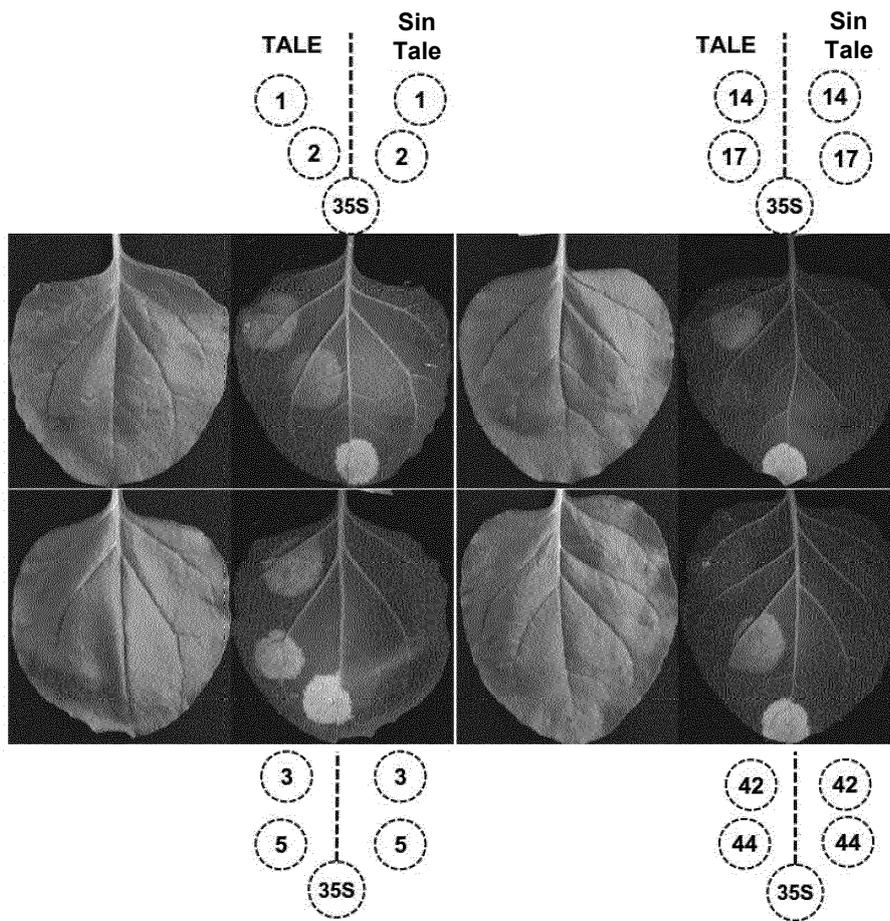


Fig. 3

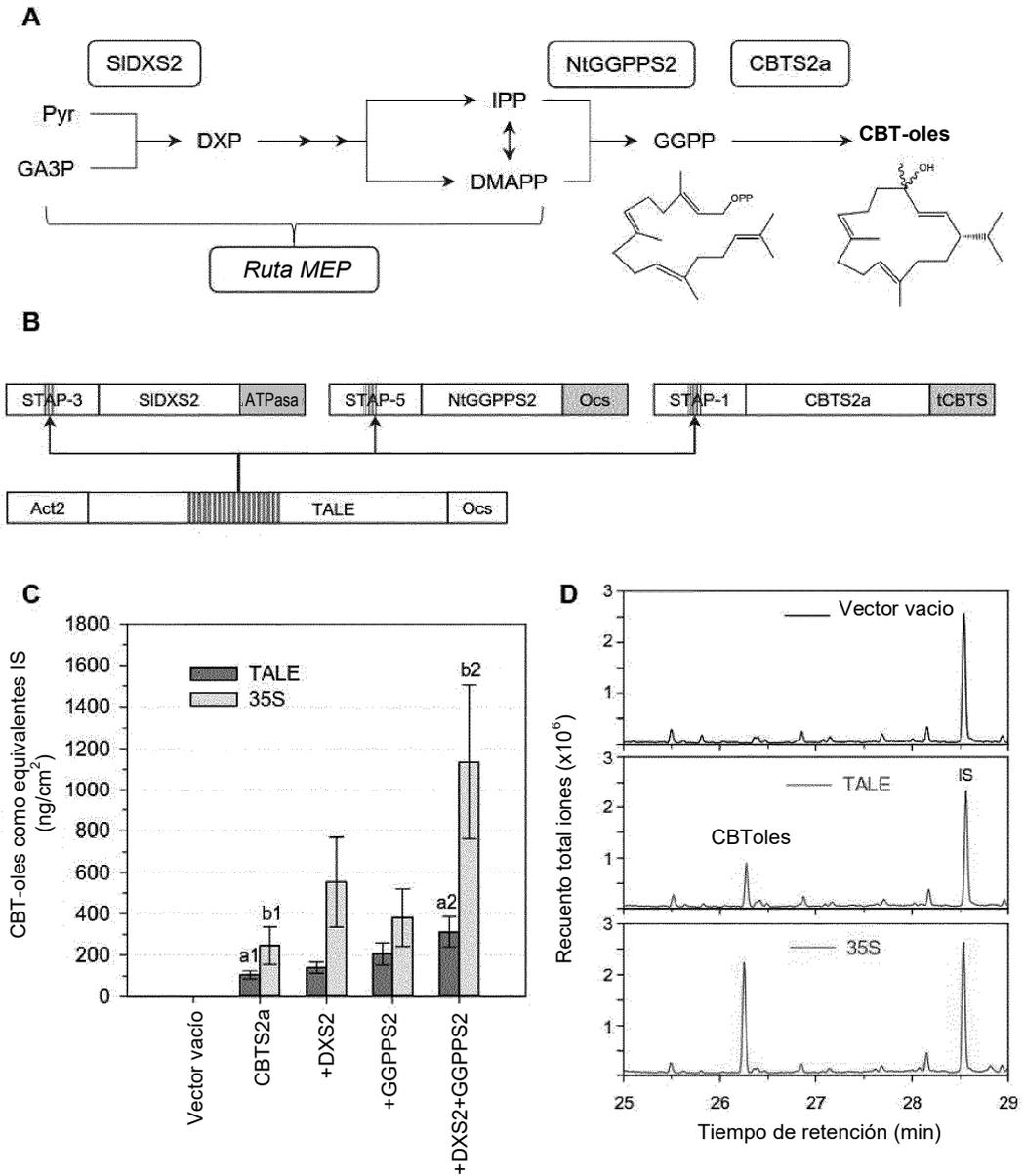


Fig. 4

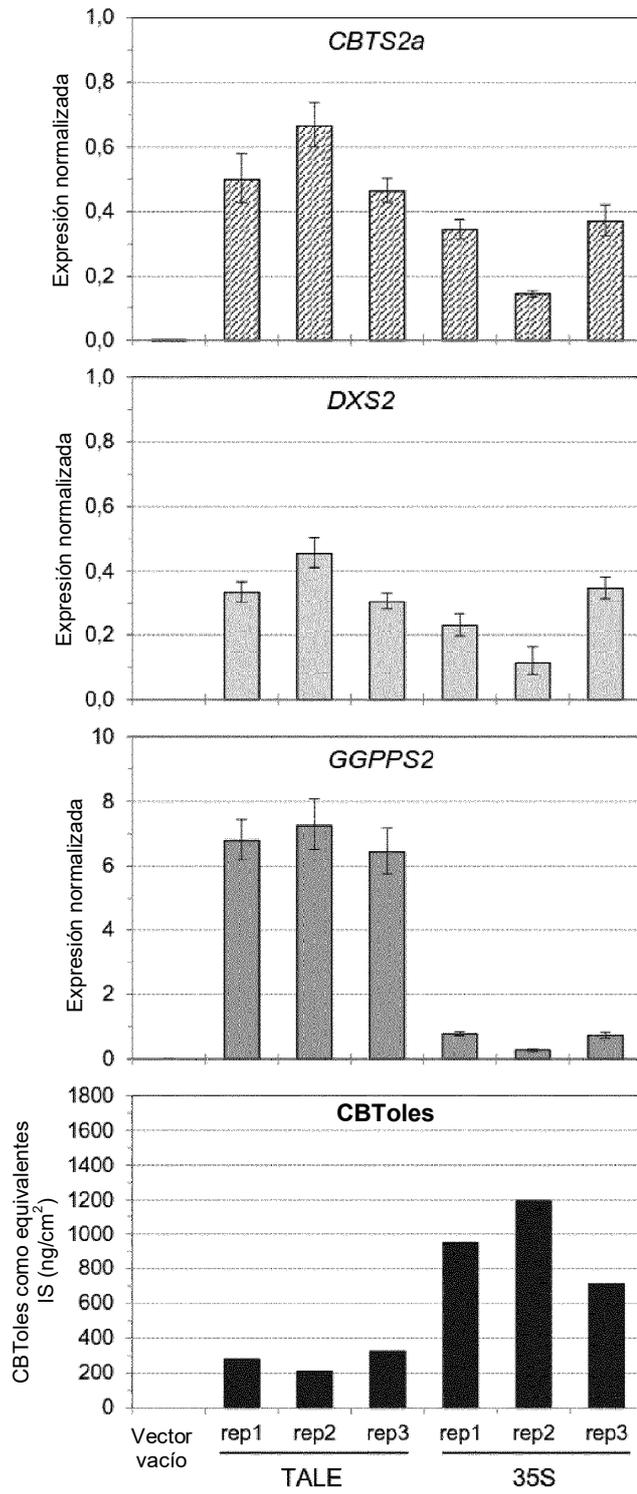


Fig. 5

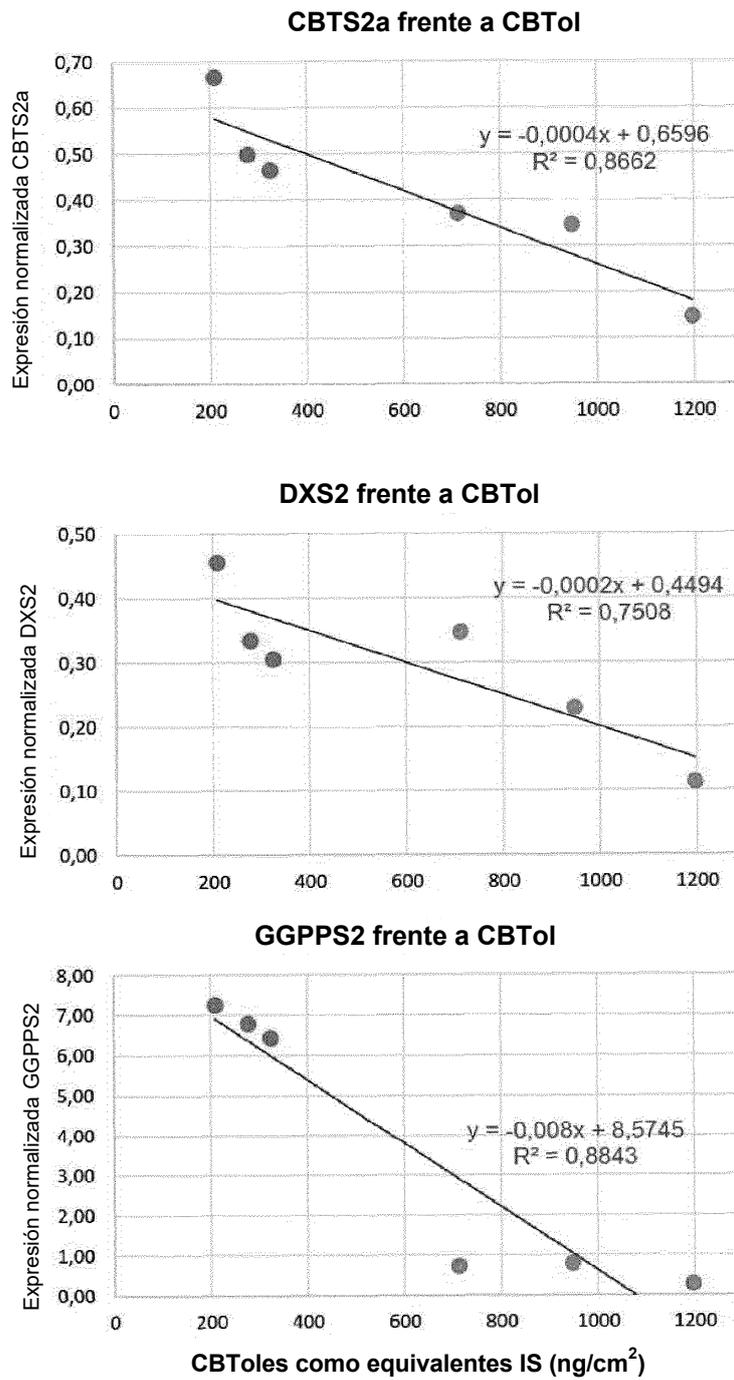


Fig. 6