

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 546 184**

51 Int. Cl.:

C12N 15/82 (2006.01)

A01H 5/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **27.12.2007** **E 07858186 (5)**

97 Fecha y número de publicación de la concesión europea: **03.06.2015** **EP 2118291**

54 Título: **Secuencia de ADN artificial con función líder optimizada en 5' (5'-UTR) para la expresión mejorada de proteínas heterólogas en plantas**

30 Prioridad:

29.12.2006 IT UD20060280

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

21.09.2015

73 Titular/es:

RODINA HOLDING S.A. (100.0%)

**Via Zurigo 46
6900 Lugano, CH**

72 Inventor/es:

**MARCHETTI, STEFANO;
DE AMICIS, FRANCESCA y
PATTI, TAMARA**

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

Observaciones :

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

ES 2 546 184 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

Secuencia de ADN artificial con función líder optimizada en 5' (5'-UTR) para la expresión mejorada de proteínas heterólogas en plantas

5

Campo de la invención

La presente invención se refiere a una secuencia de ADN artificial para mejorar la expresión de proteínas heterólogas en plantas.

10

Antecedentes de la invención

En el campo de la biotecnología existe una fuerte necesidad de potenciar el nivel de expresión de genes introducidos en los organismos relativos. Este nivel a menudo es insatisfactorio y representa una barrera para la aplicación industrial de innovaciones en la biotecnología de plantas y animales. Existe una cantidad de datos que avalan la importancia de la región líder en la regulación de los niveles de expresión génica, mientras que hay varios elementos estructurales que caracterizan la capacidad de regulación de los mismos.

15

En este caso, la secuencia líder no traducida en 5' (5'-UTR), como se propone en vectores ampliamente difundidos (por ejemplo, pBI121 y derivados, pCAMBIA y derivados), tiene numerosos defectos que la convierten en inadecuada para dirigir niveles adecuados de expresión génica en organismos modificados genéticamente. En particular, cuando se han de aumentar al máximo los rendimientos (por ejemplo, el uso de plantas como fábricas de células para compuestos útiles para el ser humano), es necesario eliminar las restricciones de la producción ejercidas por la secuencia 5'-UTR. Con este fin, se ha propuesto un líder Ω (una secuencia que existe de forma natural en el virus del mosaico del tabaco, TMV) en plantas. No obstante, esto también tiene imperfecciones y redundancias, por lo que caben mejoras.

20

25

Se sabe que la región poli(CAA) en el potenciador de la traducción presente en la líder Ω del TMV (Gallie y Walbot 1992 Nucleic Acids Res., 20, 4631-4638) potencia significativamente los niveles de expresión, es decir, tiene un efecto positivo sobre los niveles de traducción de proteínas heterólogas *in vitro* e *in vivo* (Gallie et al. 1988a Nucleic Acids Res., 16, 883-893, Gallie et al. 1988b Nucleic Acids Res., 16, 8675-8694, Gallie 2002 Nucleic Acids Res., 30, 3401-3411). La región líder Ω , una secuencia de poli(CAA) se asocia con 3 repeticiones de la secuencia ACAATTAC (Gallie et al. 1988a), pero en estudios de detección se ha demostrado que el elemento regulador responsable de potenciar los niveles de expresión puede consistir en una sola copia de la secuencia ACAATTAC en combinación con el motivo (CAA)_n (Gallie y Walbot 1992).

30

35

También se sabe que el sitio de iniciación de la transcripción (*Inr*) del promotor 35S de CaMV (Guilley et al. 1982 Cell, 30, 763-773) favorece una protección eficiente en los extremos del ARNm.

40

Adicionalmente, se sabe que muchos líderes vegetales (Bolle et al. 1996 Plant Mol. Biol. 32, 861-868) tienen una secuencia rica en elementos CT y que las secuencias ricas en TC pueden alterar los niveles de transcripción (Chen et al. 1996 J. Virol., 70, 8411-8421).

45

Se sabe que las secuencias que tienen una longitud de más de 40 nucleótidos estimulan el reconocimiento del primer AUG como auténtico codón inicial de la traducción (Kozak 1989 J. Cell. Biol., 108, 229-241). Por ejemplo, se ha observado que la extensión del líder de 29 a 74 nt produce un incremento del nivel de traducción del ARNm *in vitro* (Kozak 1991, J. Biol. Chem., 266, 19867-19870) e *in vivo* (Gallie y Walbot 1992). No se recomiendan las secuencias líder con un mayor contenido de A/T producen niveles de expresión más altos, ya que la formación de segmentos de ARNm bicatenario, debido al plegamiento de la molécula sobre sí misma. De hecho, es cierto que dichas estructuras secundarias tienen un efecto depresor sobre la eficiencia de la traducción (Pelletier y Sonenberg 1985 Cell, 40, 515-526; Kozak 1986 Proc. Natl. Acad. Sci. USA, 83, 2850-2854). Además, se ha observado que la introducción de porciones de 5'-UTR de origen viral en líderes vegetales se puede reflejar en un incremento de los niveles de expresión de las proteínas indicadoras (Dowson Day et al. 1993 Plant Mol. Biol., 23, 97-109).

50

55

El propósito de la presente invención es obviar los inconvenientes del estado de la técnica y conseguir una secuencia líder que aumente los niveles de expresión de las proteínas recombinantes en plantas.

El solicitante ha concebido, probado y plasmado la presente invención para superar los inconvenientes del estado de la técnica y obtener estos y otros propósitos y ventajas.

60

Sumario de la invención

La presente invención se expone y caracteriza en las reivindicaciones independientes, mientras que las reivindicaciones dependientes describen otras características de la invención o variantes de la idea principal de la invención.

65

De acuerdo con el propósito anterior, una secuencia de ADN artificial que tiene una función líder en 5' (5'-UTR), en lo sucesivo en el presente documento indicada por LL-TCK, se acuerdo con la presente invención comprende de forma simultánea elementos favorables a la expresión génica, tal como elementos de tres nucleótidos repetidos CAA en

combinación con elementos dinucleotídicos repetidos CT.

La secuencia LL-TCK de acuerdo con la presente invención se obtuvo por medio de síntesis artificial y es el fruto del intelecto, ya que no existe en la naturaleza.

5 La secuencia LL-TCK de acuerdo con la presente invención proporciona la combinación de elementos trinucleotídicos CAA con elementos dinucleotídicos CT y una modificación de las secuencias que activan la traducción presente en el líder Ω .

10 La secuencia de acuerdo con la presente invención contiene una región poli(CAA), es decir, un oligonucleótido que consiste en 9 o más copias del elemento CAA, preferentemente, aunque no necesariamente, contiguas entre sí.

La secuencia de acuerdo con la presente invención contiene una región poli(CT), es decir, un oligonucleótido que consiste en 4 o más copias del elemento CT, preferentemente, aunque no necesariamente, contiguas entre sí.

15 La presente invención proporciona que la secuencia contiene una o más copias del octámero ACAATTAC.

Una secuencia obtenida de la combinación de las secuencias con una región poli(CAA) y de aquellas con una región poli(CT) también entran en el campo de la presente divulgación.

20 Una secuencia obtenida de la combinación de las secuencias con una región poli(CAA) y de aquellas con una o más copias del octámero ACAATTACC también entran en el campo de la presente divulgación.

25 Adicionalmente, una secuencia obtenida de la combinación de las secuencias con una región poli(CT) y de aquellas con una o más copias del octámero ACAATTACC también entran en el campo de la presente divulgación.

De acuerdo con la presente invención, se proporciona una secuencia obtenida de la combinación de las secuencias con una región poli(CAA), aquellas con una región poli(CT) y aquellas con una o más copias del octámero ACAATTACC. Adicionalmente, de nuevo de acuerdo con la presente divulgación, es posible proporcionar una secuencia obtenida de la combinación de una o más de las secuencias anteriores con el sitio *Inr* de 35S de CaMV, es decir, el sitio de iniciación de la transcripción del promotor 35S del virus del mosaico de la coliflor.

30 La secuencia LL-TCK de acuerdo con la presente invención es, por tanto, capaz de incrementar los niveles de expresión de las proteínas heterólogas en plantas transgénicas. De acuerdo con una solución ventajosa de la presente invención, la nueva secuencia LL-TCK se sintetizó para crear una combinación de los elementos siguientes de acuerdo con un patrón original, único de su tipo:

- (1) sitio de iniciación de la transcripción (*Inr*) del promotor 35S del CaMV para una protección eficiente de los extremos del ARNm;
- (2) región poli(CAA) similar al potenciador de la transcripción presente en el líder Ω del TMV;
- (3) una secuencia rica en elementos CT, como muchas líderes vegetales.

35 Adicionalmente, la secuencia LL-TCK tiene una longitud de más de 40 nucleótidos con el fin de estimular el reconocimiento del primer AUG como el auténtico codón de iniciación de la traducción (Kozak 1989) y un contenido global de G+C de menos del 40 %.

De acuerdo con la solución concreta de la presente invención, la secuencia LL-TCK es la que se muestra en la SEC ID N° 1 (5'-3').

50 Es posible prever que las mutaciones pequeñas en la secuencia LL-TCK no alteren su eficacia y, por esta razón, la presente divulgación también se refiere a secuencias líder derivadas de la presente secuencia, por ejemplo tras delección o duplicación de un triplete CAA, sustitución o delección de una única base, etc.

55 La innovación de LL-TCK consiste en el hecho de que se une en una única líder un elemento poli(CAA) modificado, un octámero de la líder Ω del TMV y una secuencia rica en CT de origen vegetal.

La secuencia artificial LL-TCK de acuerdo con la presente invención proporciona la presencia de un único octámero ACAATTAC asociado con 9 repeticiones de CAA localizadas en la posición 5' con respecto al octámero.

60 Dado que el triplete ATT dentro del elemento ACAATTAC puede representar un sitio de iniciación de la traducción no canónico (Tyc et al. 1984 Eur. J. Biochem., 140, 503-511, Schmitz et al. 1996 Nucleic Acids Res., 24, 257-263), en la secuencia líder LL-TCK, este triplete se ha introducido en marco con un codón de terminación.

65 Adicionalmente, en la secuencia líder artificial LL-TCK se ha añadido un elemento (CT)₄ en el extremo 3' del elemento regulador obtenido de la unión del octámero ACAATTAC con la secuencia poli(CAA). La combinación de estos dos elementos, el efecto positivo de cada uno de los cuales sobre la expresión génica se conoce, nunca se ha encontrado en la naturaleza ni previamente preparada por el ser humano.

La secuencia líder LL-TCK, que combina estos dos elementos, produce una potenciación tanto del nivel de la

traducción como del nivel de la transcripción del gen afectado.

Este efecto se ha demostrado comparando los niveles de expresión del gen *uidA* obtenidos en plantas de tabaco (*Nicotiana tabacum*) transformadas con las construcciones 35S-LL-TCK::*uidA* (pSTART) y 35S::*uidA* (pBI121 con la secuencia líder original). El vector pSTART se obtuvo sustituyendo la secuencia líder en pBI121 (Clontech) por LL-TCK. En particular, el objeto de la sustitución y manipulación fue la secuencia pBI121 comprendida entre la región *Inr* (ACACG) y el sitio de restricción *Xba* I (TCTAGA). La secuencia de nucleótidos que flanquea el codón inicial de la traducción de *uidA* se mantuvo sin alterar en las dos construcciones para prevenir la variabilidad en el reconocimiento del codón AUG.

La elección de usar la enzima beta-glucuronidasa (GUS) codificada por el gen *uidA* como proteína indicadora vino determinada por el hecho de que en el tabaco no se puede observar actividad de tipo GUS nativo y el nivel de expresión del transgén *uidA* se puede medir por medio de un ensayo fluorimétrico (Jefferson et al. 1987 EMBO J., 6, 3901-3907) caracterizado por una sensibilidad, precisión, velocidad y facilidad de ejecución considerables.

Las lecturas fluorimétricas relacionadas con la actividad GUS enzimática, medida como ha descrito Jefferson (1987) en plantas regeneradas tras la transformación (generación T₁) han mostrado cómo la presencia de la secuencia líder LL-TCK produce un incremento considerable del nivel de expresión del gen *uidA* (hasta 15 veces) en comparación con la construcción original.

El análisis de la varianza permitió establecer que las diferencias halladas entre las dos poblaciones de tabaco consideradas (transformadas con pSTART y pBI121) son estadísticamente significativas, como lo son las diferencias entre los que mejor expresan de los dos grupos.

Con el fin de excluir los efectos que derivan de las variaciones epigenéticas, el análisis se repitió en las progenies T₂ obtenidas de la autofertilización de los mejores transformantes primarios. También en este caso, las plantas transformadas con pSTART mostraban niveles de expresión para el gen *uidA* que eran mucho más altos que los obtenidos con pBI121. En particular, se observó un incremento igual a 8,6 veces en la actividad, considerando todas las plantas en su totalidad, e igual a 12,5 veces considerando solo lo que expresan por encima de la media.

Con el fin de determinar el efecto LL-TCK sobre la transcripción del gen *uidA*, se seleccionaron las plantas T₂ (10 plantas para pBI121 y 13 plantas para pSTART) caracterizadas por niveles de Gus intermedios para el análisis de los niveles de transcritos por medio de RT-PCR en tiempo real. Comenzando con el ARN total extraído de cada planta, se sintetizó el ADNc usado como molde en la RT-PCR en tiempo real. Se usaron dos pares de cebadores (uno específico del gen *uidA* y uno para el gen endógeno del ARN 18S) y la mezcla maestra para PCR SYBR-Green (Applied Biosystems). La correcta cuantificación se hizo posible preparando 2 líneas de calibración (una para el transgén y una para el gen endógeno) por medio de diluciones en serie de los plásmidos control. El nivel de transcripción del gen *uidA* se calculó después en términos relativos para cada muestra, por medio de la proporción en porcentaje entre la cantidad de ARNm detectado para el transgén y la correspondiente cantidad de ARN 18S ribosómico.

Este análisis permitió verificar en las plantas pSTART un nivel promedio del transcrito para el gen *uidA* 1,7 veces más alto que el hallado en las plantas pBI121.

Para 7 pares de plantas pSTART y pBI121 caracterizadas por valores de transcritos similares, se calculó el IET (índice de eficiencia traduccional). El IET es equivalente a la proporción entre el valor de la proteína GUS medido on el ensayo fluorimétrico y el valor relativo normalizado del ARNm determinado mediante la RT-PCR en tiempo real. Comparando el IET, está claro que la nueva secuencia líder LL-TCK no solo tiene un efecto sobre los niveles de ARNm, sino que también produce un incremento de la eficiencia de la traducción del ARNm.

La secuencia LL-TCK permite incrementar el nivel de expresión de una proteína heteróloga actuando sobre el nivel del contenido de ARNm en relación con el gen afectado y también sobre el nivel de la cantidad final de proteína presente.

El efecto de LL-TCK se estudió en el tabaco, usando el promotor constitutivo de 35S del CaMV y el gen *uidA* que codifica la enzima beta-glucuronidasa (GUS), pero otros usos son posibles, en combinación con otros promotores y otros genes.

Aunque en los ejemplos en el presente documento indicaron que la secuencia líder LL-TCK en combinación con el promotor 35S del CaMV para la potenciación de la expresión de *uidA* en plantas de tabaco, dicha secuencia líder se usó on éxito también en tabaco y patata aguas abajo del promotor *rbcS1* inducible por luz (nº de acceso en GenBank AY163904) y en arroz aguas abajo del promotor de *glub4* dependiente de fase específico de endospermas (nº de acceso en GenBank AY427571). Los genes usados en estos experimentos fueron los que codifican el anticuerpo BCL1 murino, la beta-glucosidasa humana y un polímero sintético similar a la elastina. Dado que no se registró ninguna pérdida de funcionalidad en los experimentos llevados a cabo con genes no relacionados caracterizados por una longitud diferente, la composición de bases y la estructura bajo el control de los promotores con una actividad transcripcional dispar y expresados en especies dicotiledóneas así como en monocotiledóneas, se puede afirmar que la utilidad de la secuencia líder LL-TCK o 5'-UTR de composición similar es general, es decir no está limitada a determinados promotores y/o secuencias de codificación y no está limitada a determinadas especies huésped. Por

tanto, las formas de realización preferidas de la presente invención están comprendidas en un intervalo de aplicaciones biotecnológicas, incluida la resistencia a tensiones bióticas/abióticas y herbicidas, la producción de biocombustibles, bioplásticos, biopolímeros sintéticos y enzimas industriales, la preparación molecular de productos biofarmacéuticos (por ejemplo, anticuerpos y sus fragmentos, vacunas, enzimas humanas, citoquinas y factores de crecimiento), la mejora de los alimentos, los piensos y la calidad de la fibra, el desarrollo de sistemas génicos indicadores y marcadores.

Adicionalmente, entra dentro del campo de la presente invención construir, dentro de los vectores de expresión de plantas, 5'-UTR en las que los siguientes elementos están presentes de forma simultánea. El sitio *Inr* de CaMV 35S, poli(CAA)_n, octámero ACAATTAC, poli(CT)_n, en la que n es cualquier número mayor que o igual a 2.

Todas las posibles combinaciones de los elementos que constituyen la secuencia líder 5'-UTR como se ha expresado anteriormente, o en las variantes relativas, con independencia de su posición 5'-3' relativa, entran dentro del campo de la presente divulgación.

Adicionalmente, la presente invención también se refiere a las secuencias complementarias a las descritas anteriormente, o las variantes relativas.

De acuerdo con una variante, la secuencia de acuerdo con la presente invención tiene una longitud comprendida entre 20 y 200 nucleótidos, preferentemente entre 40 y 150 nucleótidos.

De acuerdo con una variante, la secuencia de acuerdo con la presente invención tiene un contenido de G+C de menos del 60 %, preferentemente menos del 50 %.

Uno o más cebadores e amplificación también entran dentro del campo de la presente divulgación, que comprende una secuencia de nucleótidos seleccionada de un grupo que comprende las secuencias de nucleótidos mostradas en las SEC ID N° 2-7 o una complementaria de las mismas.

De acuerdo con otra forma de ejecución de la presente invención, la secuencia de acuerdo con la presente invención se puede obtener mediante:

- a) síntesis artificial;
- b) procesos naturales o inducidos de recombinación o mutación dentro de secuencias naturales o artificiales.

Una característica de la presente divulgación también se refiere a un procedimiento para la síntesis artificial de una secuencia como se ha descrito anteriormente, usando uno o más de dichos cebadores de amplificación.

Las secuencias líder naturales 5'-UTR que se pueden descubrir y que, para un experto en la técnica parecen variantes no importantes, siempre que tengan una funcionalidad similar, de la secuencia de acuerdo con la presente invención, también entran dentro del campo de la presente divulgación.

Las secuencias que derivan de procesos de mutación de la secuencia de acuerdo con la presente invención que parece, para un experto en la técnica, que generan variantes no importantes, siempre que sean funcionalmente similares, de la secuencia de acuerdo con la presente invención, también forman parte de la presente divulgación; las mutaciones se refieren de forma independiente a deleciones, inserciones, transiciones, transversiones de uno o más nucleótidos en la secuencia de acuerdo con la presente invención o en la secuencia complementaria a la misma.

La presente invención también se refiere a las cepas bacterianas portadoras de plásmidos que contienen la secuencia de acuerdo con la presente invención, con particular referencia a las especies *Escherichia coli*, *Agrobacterium tumefaciens* y *Agrobacterium rhizogenes*.

La presente invención también se refiere a cepas bacterianas modificadas genéticamente que contienen la secuencia de acuerdo con la presente invención, con independencia del tipo de organismo huésped.

Adicionalmente, las células vegetales transformadas con vectores de expresión que contienen la secuencia de acuerdo con la presente invención bajo el control de un promotor constitutivo también entran dentro del campo de la presente invención.

Lo siguiente también está cubierto por la presente invención:

- células vegetales transformadas con los vectores de expresión que contienen la secuencia de acuerdo con la presente invención bajo el control de un promotor específico de tejido y, en particular, específico de semillas;
- células vegetales transformadas con los vectores de expresión que contienen la secuencia de acuerdo con la presente invención bajo el control de un promotor inducible;
- células vegetales transformadas con los vectores de expresión que contienen la secuencia de acuerdo con la presente invención bajo el control de un promotor con actividad transcripcional dependiente de fase;
- células vegetales transformadas con los vectores de expresión que contienen la secuencia de acuerdo con la presente invención bajo el control de un promotor activo en el cloroplasto;

- células vegetales transformadas con los vectores de expresión que contienen la secuencia de acuerdo con la presente invención bajo el control de un promotor activo en las mitocondrias;

5 La presente divulgación también comprende plantas caracterizadas por la expresión transitoria de cualquier proteína cuyo ARN mensajero contenga la secuencia de acuerdo con la presente invención, entendiéndose que la expresión transitoria significa la producción de dicha proteína por medio de vectores virales, agroinfiltración, electroporación, liberación de partículas.

10 La presente invención también se refiere a plantas dicotiledóneas, con referencia concreta, aunque no exclusivamente, a las especies pertenecientes a las familias de Solanaceae, Papilionaceae y Cruciferae, transformadas de forma estable con vectores de expresión que contienen la secuencia de acuerdo con la presente invención, y también a progenies de dichas plantas dicotiledóneas.

15 La presente invención también se refiere a plantas monocotiledóneas, con referencia concreta, aunque no exclusivamente, a las especies pertenecientes a la familia de Graminaceae (Poaceae), transformadas de forma estable con vectores de expresión que contienen la secuencia de acuerdo con la presente invención, y también a progenies de dichas plantas monocotiledóneas.

20 La presente invención tiene una aplicación industrial ventajosa, ya que también se refiere al uso de la secuencia de acuerdo con la presente invención para una u otra de las actividades siguientes:

- la producción biotecnología de moléculas;
- la síntesis de proteínas recombinantes;
- la síntesis de proteínas recombinantes destinadas a inducir resistencia a patógenos víricos, bacterianos o fúngicos;
- la síntesis de proteínas recombinantes destinadas a inducir resistencia a herbicidas;
- la síntesis de proteínas recombinantes destinadas a obtener una composición alterada en ácidos grasos en el material bruto y productos que derivan de los mismos;
- la síntesis de proteínas recombinantes destinadas a obtener un valor nutricional alterado del material bruto y productos que derivan de los mismos;
- la síntesis de proteínas recombinantes destinadas a la producción de combustibles, gomas y bioplásticos;
- la síntesis de enzimas industriales y proteínas comerciales;
- la síntesis de proteínas farmacéuticas;
- la síntesis de vacunas administradas por vía oral, destinadas a seres humanos y a animales;
- la síntesis de vacunas inyectables, destinadas a seres humanos y a animales;
- la síntesis de vacunas inyectables específicas del paciente, preferentemente específicas de idiotipo, para su uso en el tratamiento de tumores del sistema linfático;
- la síntesis de proteínas implicadas en la producción de metabolitos secundarios;
- la síntesis de proteínas usadas directamente o indirectamente como factores para identificar y/o seleccionar células transformadas.

Breve descripción de las figuras

45 Estas y otras características de la presente invención se harán evidentes a partir de la siguiente descripción de algunas formas de realización preferenciales, proporcionadas como ejemplo no restrictivo, con referencia a las figuras adjuntas en la misma:

- La fig. 1 es una comparación entre las secuencias líder en pBI121 y en pSTART, en la que el sitio de iniciación de la transcripción está subrayado. Dado que las secuencias entre el sitio *Eco* RV y el sitio de iniciación de la transcripción y entre el sitio *Xba* I y el triplete ATG de *uidA* son idénticas en pSTART y en pBI121, se han omitido parcialmente (puntos);
- La fig. 2 muestra los niveles de expresión de beta-glucuronidasa (GUS) en las plantas transgénicas T₁;
- La fig. 3 muestra los niveles de expresión de beta-glucuronidasa (GUS) en las plantas transgénicas TT₂; las plantas se agrupan en cuatro grupos, representando cada uno plantas hermanas que descienden de los mejores transformantes T₁, diferencia mínima significativa (P = 0,01) = 4,7 U / mg de proteína total;
- La fig. 4a muestra los niveles de transcripción relativos de *uidA* (*gusA*) según se determina mediante RT-PCR en tiempo real en plantas T₂ obtenidas con pSTART y pBI121 caracterizados por niveles de expresión de beta-glucuronidasa intermedios. • Se identifican los siete pares de plantas con niveles de transcripción similares;
- La fig. 4b muestra los valores del índice de eficiencia de traducción (IET) para las plantas T₂ con los niveles de transcripción similares. El IET se calculó del siguiente modo: para cada transformante, la concentración de beta-glucuronidasa (GUS) [U / mg de proteína total] se dividió por el nivel relativo normalizado de ARNm; el IET más alto se consideró igual a 1,00 y los valores registrados para cada planta transgénica se expresaron en consecuencia;
- La fig. 5 es un diagrama de la superposición de los cebadores inversos y directos para la síntesis de LL-TCK por medio de PCR recursiva.

Descripción detallada de la presente invención

A) Síntesis de la secuencia líder artificial LL-TCK mostrada en la SEC ID N°: 1.

ES 2 546 184 T3

A.1) La síntesis de la secuencia LL-TCK o las 5'-UTR de composición similar puede lograrse más convenientemente mediante síntesis artificial, haciendo uso de los servicios especializados disponibles en el mercado. Debido a la longitud limitada de la secuencia, es especialmente útil añadir a cada lado del líder una región flanqueante que termina con un sitio de restricción ya presente dentro de la secuencia del promotor (región 5' flanqueante) y la secuencia de codificación (región flanqueante 3'). Es obvio para un experto en la técnica que estas regiones flanqueantes reproducen con precisión la secuencia aguas arriba del sitio (*Inr*) iniciador y la secuencia de codificación, respectivamente, a menos que una modificación del promotor y / o la secuencia de codificación se planee de forma concurrente.

A.2) Otro procedimiento para obtener dicha secuencia líder es la PCR recursiva (Podromou y Pearl 1992 Protein Eng., 5, 827-829, Wheeler et al. 1996 Gene, 169, 251-255, Prytulla et al. 1996 FEBS Letters, 399, 283-289).

Una vez que la secuencia de LL-TCK o una 5'-UTR de composición similar se obtiene mediante cualquiera de los métodos, las variantes líder pueden producirse fácilmente mediante PCR o cualquier otro procedimiento desarrollado para la mutagénesis aleatoria o *in situ*.

En este ejemplo, se notifica la descripción de la síntesis de LL-TCK mediante PCR recursiva para su inserción en pBI121 (nº de acceso en GenBank AF485783), en particular entre el promotor 35S del CaMV y la secuencia de codificación de *uidA*.

Cinco oligonucleótidos sintéticos se usaron como cebadores, que tiene una longitud comprendida entre 42 y 54 nt y un grado de superposición parcial igual a 24 nt, y un cebador inverso terminal de 19 nt, mostrado respectivamente en las secuencias SEC ID Nº 2, 3, 4, 5, 6 y 7.

Todos los cebadores se escriben en la dirección 5'-3'. Las secuencias SEC ID Nº 2, 4, 6 son cebadores directos, mientras que las secuencias SEC ID Nº 3, 5 y 7 son cebadores inversos. Los cebadores directo e inverso se solapan entre sí de acuerdo con el diagrama de la fig. 5.

Para facilitar la manipulación y la inserción de LL-TCK en la secuencia del vector de que se trata, se añadió una porción terminal que comienza un sitio *Eco* RV en 5', mientras que un solo sitio *Xba* I se añadió en el extremo 3'.

Por lo tanto, los cebadores se diseñaron para proporcionar la reconstrucción de la porción 3' terminal del promotor 35S (desde el sitio *Eco* RV a la región *Inr*) a fin de facilitar la posterior inserción en el vector pBI121 (Clontech).

El cebador inverso externo introduce el sitio *Xba* I en el terminal 3', mientras que en 5' se utiliza el sitio *Eco* RV.

En pBI121, estos sitios caen dentro del promotor 35S del CaMV y en la proximidad con la señal de iniciación de la traducción de *uidA*, respectivamente. Por lo tanto, se proporcionó la síntesis de la secuencia deseada y la clonación para reemplazar el fragmento [*Eco* RV - *Xba* I].

Los cebadores que comprenden las secuencias nucleotídicas mostradas en las SEC ID Nº 2-7 se diseñaron y crearon para confirmar la secuencia del promotor en la región entre el sitio *Eco* RV y el sitio iniciador de 35S del CaMV, para sintetizar la secuencia líder LL-TCK y proporcionar un gancho molecular al extremo 3' para la clonación.

La síntesis de LL-TCK se llevó a cabo mediante una sola PCR, usando una mezcla de reacción de PCR en la que la concentración de los cebadores externos, SEC ID Nº 2 y 7 (correspondiente a los dos extremos del segmento sintetizado) era 100 veces mayor que la de los cebadores internos SEC ID Nº 3, 4, 5 y 6.

Con el fin de lograr una mayor fidelidad de la síntesis de ADN, se utilizó un ADN polimerasa de lectura de prueba en combinación con una reducción del 50 % de la concentración de dNTP.

La mezcla de la reacción PCR es la siguiente:

10X ufp tampón que contiene Mg²⁺ 15 mM: 10 microl

Cebador, SEC ID Nº 2 [10 microM]: 2 microl

Cebador, SEC ID Nº 3 [0,1 microM]: 2 microl

Cebador, SEC ID Nº 4 [0,1 microM]: 2 microl

Cebador, SEC ID Nº 5 [0,1 microM]: 2 microl

Cebador, SEC ID Nº 6 [0,1 microM]: 2 microl

Cebador, SEC ID Nº 7 [10 microM]: 2 microl

Ufp ADN polimerasa [3 U/microl]: 0,8 microl

dNTP [2,5 miliM cada uno]: 4 microl

Agua hasta un volumen final de 100 microl.

En particular, para la síntesis y la amplificación de de ADN se utilizó Ufp DE la polimerasa Taq (Promega) y el ciclo siguiente: 1 x (95 °C durante 5 min); 40 x (95 °C durante 15 s; 48 °C durante 30 s; 72 °C durante 20 s); 1 x (72 °C durante 7 min).

El producto de la PCR se purificó mediante precipitación en etanol, se sometió a electroforesis en un gel de agarosa al 1 % en tampón TAE, se recuperó del gel con la ayuda de un kit comercial, se añadió A en la cola con AmpliTaq

ES 2 546 184 T3

Gold™, y se ligó en pGEM-T (Promega) para la secuenciación en ambas cadenas.

La mezcla de ligación se utilizó para transformar células competentes de *Escherichia coli*, cepa JM101. La ausencia de cualquier apareamiento erróneo entre la secuencia clonada y la diseñada se verificó mediante secuenciación en la doble cadena. B) Construcción de un vector de expresión vegetal que alberga la secuencia LL-TCK.

La posible adición de regiones flanqueantes o ganchos moleculares a la secuencia LL-TCK o 5'-UTR de composición similar ofrece una amplia gama de soluciones de clonación en vectores de expresión de diferentes tipos. En este ejemplo, se describe el método utilizado para clonar el fragmento [*Eco* RV - *Xba* I] del Ejemplo 1 en sustitución del fragmento [[*Eco* RV-*Xba* I] de pBI121 (nº de acceso en GenBank AF485783.). Dado que pBI121 tiene múltiples sitios *Eco* RV más allá del interior del promotor 35S del CaMV, el segundo promotor se escindió del pBI121 (Clontech) (Jefferson et al. 1987) haciendo uso de las enzimas de restricción *Hind* III y *Xba* I.

El fragmento se recuperó del gel de agarosa al 1 % en tampón TAE y se subclonó en pUC18 (Pharmacia, nº de acceso en GenBank L08752), previamente digerido con las mismas enzimas.

Como se ha indicado, este pasaje fue necesario porque pBI121 tiene múltiples sitios *Eco* RV. El vector pUC18 / 35S obtenido se utilizó para hacer la nueva combinación de promotor 35S - líder LL-TCK.

La secuencia de LL-TCK se escindió del vector pGEM-T por medio de digestión con *Eco* RV y *Xba* I (NEB), se separó de la secuencia del vector mediante electroforesis en gel de agarosa y posteriormente se recuperó del gel con la ayuda de un kit comercial. El vector pUC18 / 35S a su vez se digirió con las mismas enzimas, se trató con fosfatasa alcalina (Pharmacia), se sometió a electroforesis y se recuperó del gel como anteriormente. A continuación, se llevó a cabo una reacción de ligación a 4 °C durante 16 horas en presencia de ADN ligasa de T4 (Promega). En particular, 3,5 ng del fragmento [*Eco* RV - *Xba* I] se combinaron con 25 ng del vector en presencia de 1 U de ADN ligasa de T4 en un volumen de 10 microlitros que contiene un tampón de reacción adecuado.

El vector pBI121 (Clontech) se sometió a digestión con *Xba* I y *Hind* III (NEB) para eliminar el promotor 35S del CaMV. El complejo 35S-LL-TCK se escindió a su vez del vector de clonación pUC18 mediante digestión con las mismas enzimas; el marco del vector pBI121 y el inserto 35S-LL-TCK se sometieron a electroforesis en gel y se recuperaron del gel como se ha indicado anteriormente. Por último, se realizó una ligadura de 35S-LL-TCK en el marco pBI121, obteniendo el vector pBI121 / 35S-LL-TCK:: uidA::NOS al que se le asignó el nombre pSTART (fig. 1). C) transformación de plantas con un vector de expresión que contiene la secuencia LL-TCK.

Las plantas transgénicas que albergan la secuencia LL-TCK o las 5'-UTR de composición similar se pueden producir a través de diversos métodos, incluyendo la coinfección con cepas modificadas genéticamente de *Agrobacterium* spp., Infección o transfección con cepas modificadas genéticamente de fitovirus, electroporación, liberación de partículas, microinyección de ADN.

El vector de expresión pSTART se sometió a electroporación en la cepa EHA 105 de *Agrobacterium tumefaciens* y las células de *Agrobacterium* transformadas utilizadas para la transformación del tabaco (*Nicotiana tabacum* L., cv. Xanthi). Brevemente, 2 ml de medio LB suplementado con kanamicina (50 mg / l) se inocularon con células de *Agrobacterium* transformadas. Los cultivos de bacterias se incubaron a 29 °C durante 16 horas. Los discos de hojas (7 mm de diámetro) se obtuvieron con un perforador de corcho de plántulas de 30 días de edad cultivadas axénicamente o de hojas maduras recolectadas a partir de plantas en la etapa tardía de roseta. En este último caso, las hojas de tabaco se enjuagaron con agua destilada, se esterilizaron en superficie en 1 % de hipoclorito sódico durante 5 minutos y en etanol al 95 % durante 30 segundos, y se transfirieron a papel de filtro estéril bajo una campana de flujo laminar.

Los discos de hojas se colocaron en una placa Petri que contiene 15 ml de medio Murashige y Skoog suplementado con 0.1 mg / l de ácido naftaleno acético (NAA), 1 mg / l de 6-benciladenina (BA), 30 g / l de sacarosa y solidificado con 8 g / l de agar. Inmediatamente después de esta transferencia, se vertieron 2 ml del cultivo de *Agrobacterium* antes mencionado en la placa Petri y los discos de hojas se humedecieron uniformemente. Después de la eliminación del medio LB en exceso, la placa se selló y se incubó durante 24 horas a 25 °C con luz (30,5 microE / metro cuadrado / s).

Los discos de hojas se transfirieron después a una nueva placa Petri que contiene 15 ml de medio Murashige y Skoog suplementado con 0.1 mg / l de ácido naftaleno acético (NAA), 1 mg / l de 6-benciladenina (BA), 500 mg / l de cefotaxima, 30 g / l de sacarosa y solidificado con 8 g / l de agar. Se incubaron durante una semana a 28 °C y se iluminaron durante 16 horas / día; al final, se transfirieron a un sustrato idéntico al anterior excepto por la presencia de 200 mg / l de kanamicina. Los explantes se subcultivaron cada 3 semanas; los brotes regenerados se aislaron de tejido de callo y se enraizaron en un medio semisólido de Murashige y Skoog suplementado con 2 mg / l de ácido indol-3-butírico, 500 mg / l de cefotaxima, 200 mg / l de kanamicina, 30 g / l de sacarosa.

Supuestas plantas transgénicas se plantaron en turba y se cultivaron en un invernadero bajo lámparas Powerstar HQL-T (Osram) (fotones 200 μmol / metro cuadrado / s a nivel del dosel) durante 16 horas / día a 25 °C / 19 °C luz / oscuridad.

En este ejemplo, la transformación se confirmó mediante PCR y ensayos de beta-glucuronidasa. Para el ensayo de

ES 2 546 184 T3

PCR, el ADN total se extrajo de acuerdo con Doyle y Doyle (1990) y se usaron los siguientes cebadores:

Directo 5' - ACAATTACGTATTTCTCTCTCTAGA - 3'
Inverso 5' - CGATCGGGGAAATTCGAGCTC - 3'

5 Los cebadores directos e inversos se hibridan al final de la secuencia LL-TCK y a parte del terminador NOS, respectivamente, y no dan lugar a ningún producto de amplificación en las plantas Xanthi no transformadas (en plantas transgénicas, la longitud del amplicón se 1.936 pb, como se esperaba).

10 Se encontró que aproximadamente el 93% de las plantas regeneradas eran transgénicas cuando se formó una mezcla de reacción estándar y se utilizó el siguiente ciclo de temperatura:

1 x (94 °C durante 5 min); 40 x (94 °C durante 1 min 15 s; 60 °C durante 45 s; 72 °C durante 2 min); 1 x (72 °C durante 5 min).

15 La transformación de las plantas se demostró además mediante ensayos histoquímicas GUS (Jefferson et al. 1987) y la determinación fluorimétrica de la actividad de GUS. Los controles consistieron en plantas Xanthi plantadas *in vitro* a partir de discos no infectados. Los métodos utilizados para el ensayo fluorimétrico y los resultados obtenidos se exponen con detalle en el punto D.

20 Se siguió el mismo procedimiento para producir y caracterizar las plantas transgénicas que albergan la secuencia líder original. *Ceteris paribus*, no se observaron efectos de la secuencia de LL-TCK sobre los índices de regeneración y transformación. D) Efecto de la secuencia de LL-TCK sobre los niveles de expresión del transgén.

25 Como se ha indicado anteriormente, los plásmidos pBI121 y pSTART se usaron individualmente para la transformación mediada por *Agrobacterium* de discos de hojas de tabaco. Dado que en ambos casos, el gen bajo control es *uidA* (también conocido como *gusA*), los niveles de expresión del transgén alcanzables con LL-TCK y el líder pBI121 ampliamente distribuido se pueden comparar directamente mediante la determinación de la actividad de la enzima codificada por *uidA*, beta-glucuronidasa (EC 3.2.1.31). Aproximadamente veinte transformantes primarios (es decir, plantas transgénicas que pertenecen a la primera generación, T₁) de cada población se analizaron para
30 determinar la presencia del transgén mediante PCR y posteriormente se analizaron para determinar la actividad beta-glucuronidasa (fig. 2). Cuando se llegó a la etapa de roseta tardía (30 días tras el endurecimiento), las 3 hojas más jóvenes se recogieron de cada planta para obtener savia bruta mediante prensado (Erich Pollahne); 100 microlitros de la savia bruta se mezclaron con 2 volúmenes de tampón de extracción (Jefferson, 1989) que contienen 12 mg de polivinilpirrolidona de alto peso molecular (PVP). Después de centrifugación durante 15 minutos a 11.500 x
35 g, el sobrenadante se recogió y analizó fluorimétricamente (Fluorómetro Dyna Quant 200; GE Healthcare) por duplicado utilizando 4-metilumbeliferil-beta-D-glucurónido (MUG, Sigma-Aldrich) como sustrato. Se realizaron ensayos para determinar el ruido de fondo debido a la fluorescencia intrínseca de las muestras, inactivación, así como la degradación del sustrato por factores distintos de la enzima recombinante. El nivel de expresión de los transgenes se midió en términos de unidades de beta-glucuronidasa por ml de savia bruta, definiéndose una unidad
40 como la cantidad de enzima liberadora de 4-metilumbeliferona 1 nM (4-MU) min⁻¹ utilizando las mismas condiciones de ensayo que las descritas anteriormente (Jefferson 1989). Los datos se sometieron a una transformación logarítmica para evitar cualquier correlación entre la varianza y media; el análisis de varianza se llevó a cabo después de comprobar la normalidad de la distribución de los datos de registro por medio de la prueba de Kolmogorov-Smirnov y la homogeneidad de las varianzas con la fórmula de Bartlett. Los promedios se compararon
45 con la prueba de Duncan de rango múltiple al nivel de probabilidad, P = 0,05.

50 El análisis de varianza realizado con los datos de fluorescencia mostró la ausencia de cualquier variación significativa entre las hojas jóvenes de la misma planta, mientras que existieron diferencias notables entre las plantas. Específicamente, el líder sintético determinó un aumento altamente significativo (hasta 15 veces) en la expresión de *uidA* (Tabla 1).

Tabla 1: Actividad de beta-glucuronidasa (U / ml de savia bruta) en plantas T₁ escogidas al azar

pSTART	pBI121
96,10	6,24
89,81	5,37
67,06	5,34
57,89	4,52
33,16	4,20
27,62	3,55
15,97	2,95
4,24	2,89
3,51	2,79

ES 2 546 184 T3

1,85	2,34
1,64	2,30
1,46	1,62
1,21	1,17
0,77	1,12
0,64	0,75
0,46	0,47
0,38	0,23
	0,15

Para demostrar que estos resultados no estaban sesgados por la variación epigenética, los análisis se repitieron en la progenie T₂. En particular, las mejores 4 T₁ en cada población se autofecundaron y la semilla resultante sembrada en un medio enriquecido con kanamicina para la selección; 5-7 plantas T₂ resistentes se eligieron al azar dentro de cada progenie y se criaron a la etapa de roseta tardía para la verificación de la presencia del transgén mediante PCR y medición de la actividad beta-glucuronidasa mediante ensayos fluorimétricos; los niveles de enzimas se expresaron de nuevo como unidades por ml de savia bruta, pero también como unidades por mg de proteína total (según lo determinado por el ensayo de Bradford) para dar cuenta de la variación de planta a planta en el metabolismo.

De manera similar a lo observado en la generación T₁, las plantas transgénicas T₂ que albergan la nueva secuencia líder mostraron niveles significativamente más altos de expresión de *uidA* (fig. 3); en comparación con la secuencia líder pBI121, se estimaron un incremento de 8,6 veces y un aumento de 12,5 veces de la actividad teniendo en cuenta toda la población de plantas o los expresores superiores a la media, respectivamente (Tabla 2).

Tabla 2: Actividad de beta-glucuronidasa (U/mg de proteína) en plantas T² obtenidos a partir de los mejores 4 T₁.

pSTART	pBI121
139,07	789
71,27	5,17
45,43	4,44
44,81	3,96
41,67	3,94
32,92	3,78
29,81	3,77
19,79	3,58
18,04	3,55
16,40	3,23
15,10	3,15
13,55	3,09
12,84	2,94
12,32	2,62
10,78	2,45
10,02	1,71
9,00	1,28
8,50	1,22
6,98	1,14
6,57	0,64
5,39	0,52
5,16	0,46
5,01	0,44

3,17	0,31
------	------

E) Efecto de la secuencia de LL-TCK sobre la transcripción y la traducción génica.

5 La combinación de diferentes elementos en LL-TCK o en 5'-UTR de composición similar se refleja en mejoras mensurables de la transcripción, así como la eficiencia de la traducción de un transgén dado. En este ejemplo, dichas mejoras se muestran en plantas de tabaco transgénicas T₂ obtenidas con pBI121 y pSTART como se ha descrito antes. Con el fin de recoger esas pruebas, se analizaron las plantas pertenecientes al grupo de pBI121 o pSTART (10 y 13 T₂, respectivamente) y caracterizadas por niveles intermedios de expresión de *uidA* se analizaron para determinar:

- 10
- i. los niveles medios del transcrito de *uidA*;
 - ii. los niveles medios del transcrito de ARN 18S;
 - iii. la cantidad de beta-glucuronidasa producida en realidad.

15 Para minimizar el error experimental, se recolectó una hoja joven de cada planta y se cortó en dos, usándose una mitad para el aislamiento de ARN, la otra para el ensayo de beta-glucuronidasa.

20 El ARN total se extrajo con RNAGents Total RNA Isolation System (Promega). La primera hebra del ADNc se sintetizó a partir 1 microg de ARN mediante transcriptasa inversa AMV (Promega) en presencia de cebadores aleatorios. La reacción de síntesis de ADNc se diluyó A 1: 5 y 1 microl se utilizó para la PCR en tiempo real (qRT-PCR).

25 Las qRT-PCR se llevaron a cabo haciendo uso de la mezcla maestra SYBR-Green PCR Master Mix (Applied Biosystems) y cebadores específicos a una concentración final 0,3 microM. Todas las reacciones se llevaron a cabo con el sistema de detección iCycler iQ multicolor-PCR en tiempo real (Bio-Rad) y se ejecutaron con el siguiente programa: 1 x (95 °C durante 10 min); 50 x (95 °C durante 15 s; 62 °C durante 30 s; 72 °C durante 30 s); 1 x (72 °C durante 30 s). Para amplificar la transcripción *uidA*, se utilizaron los siguientes cebadores:

30 Directo 5' - TTACGCTGAAGAGATGCTCGAC -3'
Inverso 5'-CCTAAAGAGAGGTTAAAGCCGACAG-3'

Para la secuencia diana de ARN 18S, los cebadores se diseñaron sobre la base del nº de acceso en GenBank: AJ236016:

35 Directo 5' - ACATCCAAGGAAGGCAGCAG-3'
Inverso 5'-GACTCATAGAGCCCGGTATTGTTATT-3'

40 En ambos casos, la longitud del amplicón fue 90 kb. En cada ciclo de PCR, se incluyeron diluciones en serie de los plásmidos control en paralelo con cantidades conocidas del número de copias de entrada con el fin de realizar curvas de calibración estándar. Específicamente, como moldes se usaron diluciones en serie por 10 (de 10⁵ to 10² copias) del plásmido pBI221 portador de *uidA*. Para el mismo propósito se clonó un fragmento de 550 pb del gen del ARN 18S (AJ236016) en pGEM-T Easy (Promega) y se usó en el intervalo de 10⁸ -10⁵ copias.

45 Las cantidades de partida del transcrito de *uidA* y el ARN control se determinaron con el software del sistema de detección en tiempo real ICycler IQ, versión 3.0. Para cada muestra se realizaron al menos 3 estimaciones independientes; en todos los casos, el valor máximo para el coeficiente de variación (CV=DE/media) se calculó al 20 %. La relación en % entre los niveles medios de transcrito de *uidA* y el ARN 18S se calculó para cada muestra. Para normalizar los datos, los niveles más altos del transcrito de *uidA* se consideraron iguales a 1,00 y los valores registrados para cada planta transgénica se expresaron en consecuencia.

50 Los resultados obtenidos en la qRT-PCR indicaron que la sustitución de la secuencia líder de pBI121 por LL-TCK determina un claro incremento de los niveles medios del transcrito de *uidA*. Específicamente, se halló que la eficiencia de la transcripción era 1,7 veces mayor en las plantas que alojaban la secuencia líder LL-TCK (Tabla 3).

55 Tabla 3: Relación en % entre los transcritos de *uidA* y 18S en varias plantas T₂

pSTART	pBI121
0,01371%	0,00881%
0,01189%	0,00832%
0,01132%	0,00761%
0,00968%	0,00437%
0,00960%	0,00315%
0,00896%	0,00308%

0,00578%	0,00282%
0,00556%	0,00234%
0,00425%	0,00225%
0,00315%	0,00158%
0,00272%	0,00072%
0,00201%	
0,00189%	

Además, después de haber identificado 7 parejas de plantas con niveles de transcripción de *uidA* casi solapantes (fig. 4a), se atribuyó un índice de eficiencia de la traducción (IET) a cada planta (fig. 4b.), lo que hace la relación entre la concentración de beta-glucuronidasa y el nivel relativo estandarizado del transcrito de *uidA*. Mediante la comparación de los valores del IET se encontró que las dos secuencias líder determinaban una eficiencia de la traducción claramente diferente de los transcritos de *uidA*, que fue mayor usando la secuencia líder de acuerdo con la presente invención (Tabla 4).

10 Tabla 4: relación entre la actividad enzimática de GUS y el nivel de transcripción relativo del gen *uidA* medido en algunas plantas T₂

IET	pSTART	pBI121	IET
1,00	299,07	103,06	0,34
0,93	277,07	66,58	0,22
0,71	211,87	58,52	0,20
0,68	203,76	19,82	0,07
0,65	195,24	42,97	0,14
0,65	194,12	47,18	0,16
0,52	154,91	20,86	0,07
0,34	102,01	29,52	0,10
0,30	89,48	8,94	0,03
0,27	80,26	12,34	0,04
0,25	73,69	22,81	0,08
0,20	59,70		
0,17	49,40		

LISTADO DE SECUENCIAS

- 15 <110> UNIVERSITA' DEGLI STUDI DI UDINE
- <120> SECUENCIA DE ADN ARTIFICIAL CON FUNCIÓN LÍDER OPTIMIZADA EN 5' (5'-UTR) PARA LA EXPRESIÓN MEJORADA DE PROTEÍNAS HETERÓLOGAS EN PLANTAS
- 20 <130> T3-2288
- <150> IT UD2006A0000280
- <151> 29-12-2006
- 25 <160> 7
- <170> PatentIn versión 3.3
- 30 <210> 1
- <211> 73
- <212> ADN
- <213> Artificial
- 35 <220> 5'UTR
- <223> expresión mejorada de proteínas heterólogas en plantas

ES 2 546 184 T3

<400> 1

acacgtattt ttacaacaat accaacaaca acaacaacaa acaacattac aattacgtat 60

ttctctctct aga 73

5 <210> 2
<211> 54
<212> ADN
<213> Artificial

10 <220> cebador_unión
<223> elemento cebador directo para la síntesis artificial por medio de PCR recursiva

<400> 2

15 gatatctcca ctgacgtaag ggatgacgca caatcccact atccttcgca agac 54

<210> 3
<211> 54
<212> ADN
<213> Artificial

20 <220> cebador_unión
<223> elemento cebador inverso para la síntesis artificial por medio de PCR recursiva

25 <400> 3

atgaaatgaa ctctctata tagaggaagg gtcttgcgaa ggatagtggg attg 54

<210> 4
<211> 53
<212> ADN
<213> Artificial

30 <220> cebador_unión
<223> elemento cebador directo para la síntesis artificial por medio de PCR recursiva

35 <400> 4

tctatataag gaagtcatt tcattggag agaacacgta ttttacaac aat 53

40 <210> 5
<211> 52
<212> ADN
<213> Artificial

45 <220> cebador_unión
<223> elemento cebador inverso para la síntesis artificial por medio de PCR recursiva

50 <400> 5

taatgtgtt tgttgtgtt gttgttgta ttgttgtaa aatacgtgtt ct 52

<210> 6
<211> 53
<212> ADN
<213> Artificial

55 <220> cebador-unión
<223> elemento cebador directo para la síntesis artificial por medio de PCR recursiva

60 <400> 6

aacaacaaca acaacaaca ttacaattac gtatttctct ctctagagga tcc 53

<210> 7
<211> 19
<212> ADN
<213> Artificial

65

ES 2 546 184 T3

<220> cebador-unión

<223> elemento cebador inverso para la síntesis artificial por medio de PCR recursiva

5

<400> 7

cacccgggga tcctctaga 19

REIVINDICACIONES

- 5 1. ADN artificial de una región líder 5'-UTR adecuado para mejorar la expresión de proteínas heterólogas en plantas, que tiene elementos favorables a la expresión génica, que comprende una única copia del octámero ACAATTAC asociado con al menos una región poli(CAA), que comprende 9 repeticiones de CAA situadas en posición 5' con respecto al octámero y en combinación con al menos una región poli(CT) que comprende un elemento (CT)₄ añadido al extremo 3' del elemento regulador obtenido a partir de la unión del octámero ACAATTAC con la región poli(CAA).
- 10 2. ADN artificial de acuerdo con la reivindicación 1, que comprende una combinación de dicho ADN artificial con el sitio *Inr* de 35S del CaMV, es decir, el sitio de iniciación de la transcripción del promotor 35S del virus del mosaico de la coliflor.
3. ADN artificial de acuerdo con las reivindicaciones 1 o 2, que comprende la secuencia mostrada en la SEC ID N° 1.
- 15 4. ADN artificial de acuerdo con cualquiera de las reivindicaciones anteriores, que comprende secuencias complementarias a las indicadas en una cualquiera de las reivindicaciones anteriores.
5. ADN artificial de acuerdo con cualquiera de las reivindicaciones anteriores, que tiene una longitud comprendida entre 40 y 150 nucleótidos.
- 20 6. ADN Artificial de acuerdo con cualquiera de las reivindicaciones anteriores, que tiene un contenido de G+C de menos del 60 %, preferentemente de menos del 50 %.
- 25 7. ADN artificial de acuerdo con cualquiera de las reivindicaciones anteriores en cuanto a que se puede obtener a partir de:
- a) síntesis artificial; b) procesos de recombinación naturales o inducidos dentro de secuencias naturales o artificiales.
- 30 8. Una construcción dentro de vectores de expresión vegetales de 5'-UTR, que comprende un ADN artificial de acuerdo con cualquiera de las reivindicaciones anteriores presente en la región 5'-UTR.
- 35 9. Cepas bacterianas portadoras de plásmidos que contienen el ADN artificial de acuerdo con cualquiera de las reivindicaciones 1 a 7, con referencia particular a la especie *Escherichia coli*, *Agrobacterium tumefaciens* y *Agrobacterium rhizogenes*.
10. Cepas de virus modificados genéticamente que contienen el ADN artificial de acuerdo con cualquiera de las reivindicaciones 1 a 7, con independencia del organismo huésped.
- 40 11. Células vegetales transformadas de forma estable con vectores de expresión que contienen el ADN artificial de acuerdo con cualquiera de las reivindicaciones 1 a 7 bajo el control de un promotor seleccionado de un grupo que comprende: un promotor constitutivo, un promotor específico de tejido y en particular específico de la semilla, un promotor inducible, un promotor con actividad transcripcional dependiente de la fase, un promotor activo en el cloroplasto, un promotor activo en la mitocondria.
- 45 12. Plantas transformadas de forma estable con vectores de expresión que contienen el ADN artificial de acuerdo con cualquiera de las reivindicaciones 1 a 7, **caracterizadas por** la expresión transitoria de cualquier proteína en la que el ARN mensajero contiene dicha región líder en 5'-UTR, entendiéndose que la expresión transitoria significa la producción de dicha proteína por medio de vectores virales, agroinfiltración, liberación de partículas, electroporación.
- 50 13. Plantas o progenies de la misma, transformadas de manera estable con vectores de expresión que contienen el ADN artificial de acuerdo con cualquiera de las reivindicaciones 1 a 7.
- 55 14. Uso del ADN artificial de acuerdo con cualquiera de las reivindicaciones 1 a 7 para la producción biotecnológica de moléculas o para la síntesis de:
- proteínas recombinantes;
 - proteínas recombinantes destinadas a inducir resistencia a patógenos víricos, bacterianos o fúngicos;
 - proteínas recombinantes destinadas a inducir resistencia a herbicidas;
 - 60 -proteínas recombinantes destinadas a obtener una composición modificada en ácidos grasos en el material bruto y productos que derivan de las mismas;
 - proteínas recombinantes destinadas a obtener un valor nutricional modificado del material bruto y de los productos que derivan de las mismas;
 - proteínas recombinantes destinadas a la producción de combustibles, gomas y bioplásticos;
 - 65 - enzimas industriales y proteínas comerciales;
 - proteínas farmacéuticas;
 - vacunas administradas por vía oral, destinadas para el ser humano o los animales;
 - vacunas inyectables, destinadas para el ser humano o los animales;
 - vacunas inyectables específicas del paciente, preferentemente específicas de idiotipo, para su uso en el

ES 2 546 184 T3

tratamiento de tumores del sistema linfático;

- proteínas implicadas en la producción de metabolitos secundarios;

- proteínas utilizables directamente o indirectamente como factores para identificar y/o seleccionar células transformadas.

5

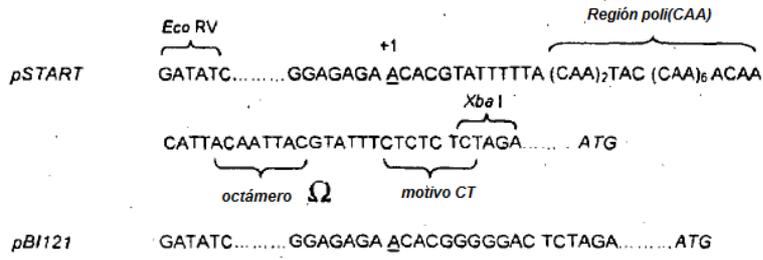


Fig. 1

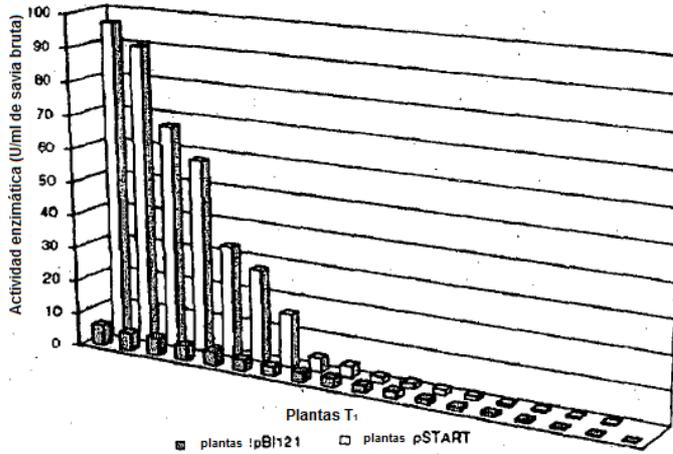


Fig. 2

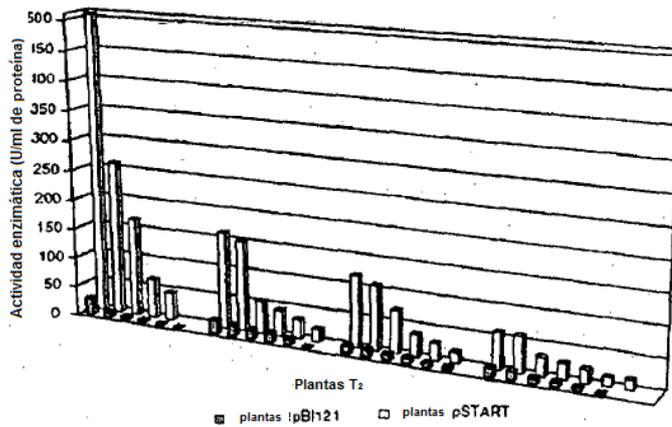


Fig. 3

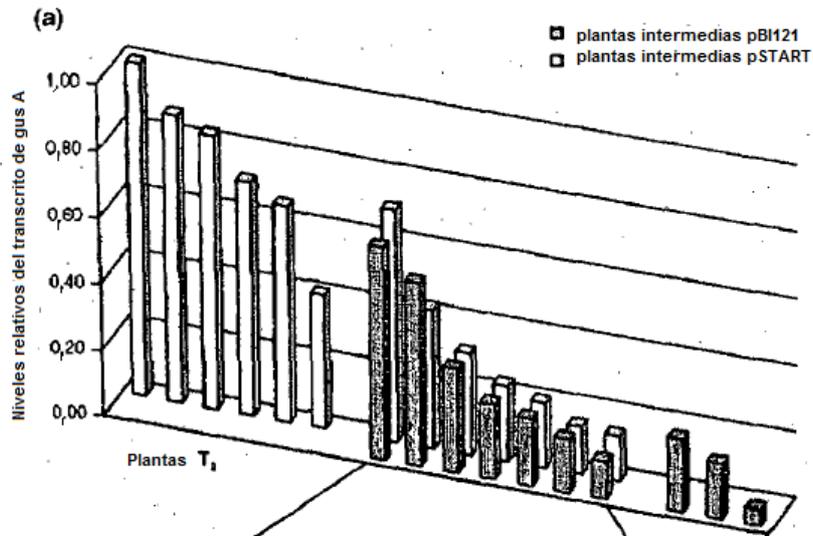


Fig. 4a

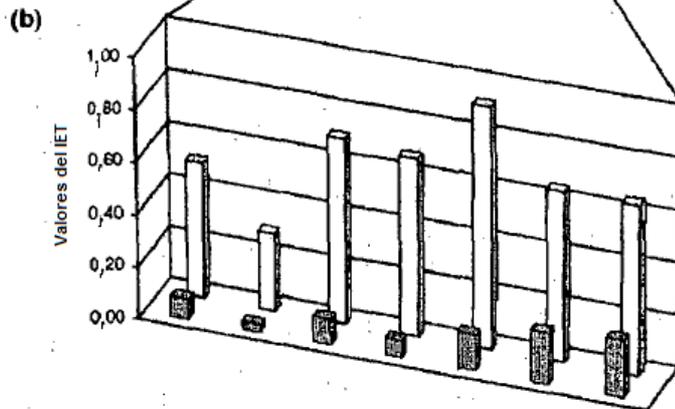


Fig. 4b

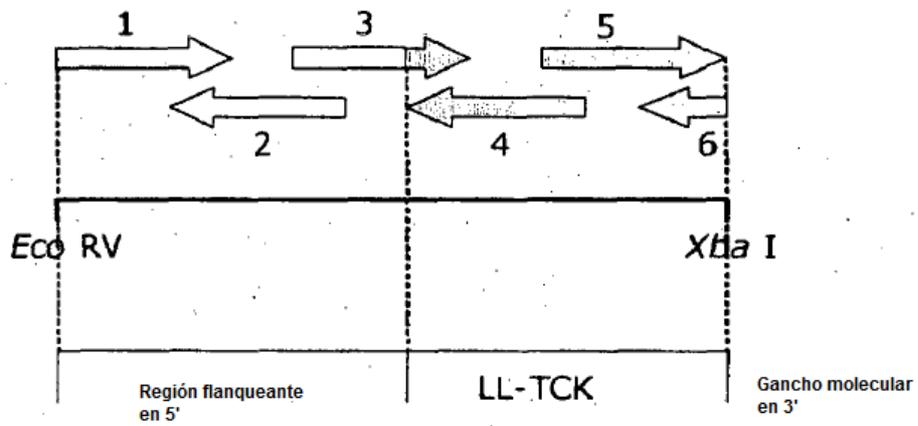


Fig. 5