

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 537 198**

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: **08.01.2009 E 09700649 (8)**

97 Fecha y número de publicación de la concesión europea: **04.03.2015 EP 2240589**

54 Título: **Sistemas de expresión de proteínas**

30 Prioridad:

08.01.2008 GB 0800272

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

03.06.2015

73 Titular/es:

**PLANT BIOSCIENCE LIMITED (100.0%)
NORWICH RESEARCH PARK COLNEY LANE
NORWICH NORFOLK NR4 7UH, GB**

72 Inventor/es:

**SAINSBURY, FRANK y
LOMONOSSOFF, GEORGE PETER**

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

ES 2 537 198 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

Sistemas de expresión de proteínas

5 **Campo de la invención**

La presente invención se refiere de forma general a métodos y materiales, y especialmente secuencias derivadas de virus, para reforzar la expresión génica en plantas y otras células eucariotas, por ejemplo, de genes heterólogos que codifican proteínas de interés.

10

Antecedentes de la invención

Comovirus (CPMV)

15 Los comovirus son virus de ARN que un genoma bipartido. Los segmentos del genoma de ARN como vírico se denominan ARN-1 y ARN-2. ARN-1 codifica las proteínas replicasa y proteasa de Vpg (Lomonossoff y Shanks, 1983). El virus necesita la replicasa para la replicación del genoma vírico. El ARN-2 del comovirus *virus del mosaico del caupí* (CPMV) codifica una proteína de 58K y 48K, así como dos proteínas de la cubierta vírica L y S.

20 El inicio de traducción del ARN-2 de todos los comovirus se produce en dos puntos de inicio diferentes situados en el mismo marco de lectura del triplete, lo que da como resultado dos carboxiproteínas coterminales. Este fenómeno de doble inicio se produce como resultado de un 'escape del barrido ribosómico (*leaky scanning*) de los ribosomas durante la traducción.

25 Los codones de inicio del extremo 5' (AUG) del ARN-2 de CPMV se encuentran en las posiciones 115, 161, 512 y 524. Los codones de inicio en las posiciones 161 y 512 están en el mismo marco de lectura del triplete. El inicio en el codón de inicio de la posición 161 da como resultado la síntesis de una poliproteína de 105K mientras que el inicio en el codón de inicio de la posición 512 dirige la síntesis de una poliproteína de 95K. Como la síntesis de ambas poliproteínas finaliza en el mismo codón de detención en la posición 3299, las proteínas 105K y 95K son carboxi coterminales. El codón AUG en la posición 524 puede servir como iniciador si el AUG en 512 se borra. Sin embargo, en presencia del AUG 512, no realiza esta función y simplemente codifica el aminoácido metionina (Holness *et al.*, 1989; Wellink *et al.*, 1993). El codón de inicio en posición 115 no es fundamental para la replicación del virus (Wellink *et al.*, 1993).

35 Las proteínas de 105K y 95K codificadas por el genoma ARN-2 de CPMV son productos de traducción primaria que posteriormente se escinden mediante la actividad proteolítica codificada por ARN-1 para producir la proteína de 58K o bien la proteína de 48K, dependiendo de si la poliproteína que se está procesando es de 105K o 95K, y las dos proteínas víricas de la cubierta, L y S. El inicio de traducción en el codón de inicio en la posición 512 en CPMV es más eficaz que el inicio en la posición 161, lo que da como resultado la producción de más poliproteína de 95K que de poliproteína de 105K.

40

El codón de inicio en la posición 115 en el ARN-2 de CPMV se encuentra en dirección 5' de los sitios de inicio en las posiciones 161 y 512 y se encuentra en un marco de lectura diferente. Puesto que este codón de inicio está en fase con un codón de detención en la posición 175, el inicio en este sitio podría dar como resultado la producción de un péptido de 20 aminoácidos. Sin embargo, hasta el momento no se ha detectado la producción de dicho péptido.

45

Necesidad de mantener el marco entre AUG

50 Los experimentos de mutagénesis han mostrado que el mantenimiento del marco entre los sitios de inicio en las posiciones 161 y 512 del ARN-2 de CPMV es esencial para una replicación eficaz de ARN-2 mediante la replicasa codificada por ARN-1 (Holness *et al.*, 1989; van Bokhoven *et al.*, 1993; Rohll *et al.*, 1993; Wellink *et al.*, 1993). Este requisito restringe la longitud de las secuencias que se pueden insertar en dirección 5' del codón de inicio 512 en vectores de expresión basados en ARN-2 de CPMV (véase más adelante), haciendo que la clonación de genes extraños en dichos vectores sea más difícil de lo que sería ideal. Por ejemplo, impide el uso de polienlazadores, ya que el uso de estos frecuentemente altera el marco de lectura abierto (ORF) entre estos sitios de inicio.

55

Vectores CPMV

60 CPMV ha servido como base para el desarrollo de sistemas vectores adecuados para la producción de polipéptidos heterólogos en plantas (Liu *et al.*, 2005; Sainsbury *et al.*, 2007). Estos sistemas están basados en la modificación del ARN-2 pero difieren en si se usan versiones de longitud completa o borradas. En ambos casos, sin embargo, la replicación del ARN-2 modificado se consigue mediante la inoculación simultánea con ARN-1. Los sistemas de expresión basados en una versión del ARN-2 de longitud completa implican la fusión de la proteína extraña con el extremo C de las poliproteínas derivadas del ARN-2. La liberación del polipéptido del extremo N está mediada por la acción de la secuencia del péptido catalítico 2A procedente del virus de la enfermedad de manos y pies (Gopinath *et al.*, 2000). Las moléculas de ARN-2 resultantes pueden diseminarse en una planta y también entre plantas. Esta

65

estrategia se ha utilizado para expresar numerosas proteínas recombinantes, tal como el antígeno principal de la hepatitis B (HBcAg) y proteínas inmunes pequeñas (SIP), en plantas de caupí (Mechtcheriakova *et al.*, 2006; Monger *et al.*, 2006; Alamillo *et al.*, 2006). Aunque ha tenido éxito, el uso de un vector vírico de longitud completa tiene desventajas en términos limitaciones de tamaño de las secuencias insertadas y preocupaciones sobre bioconfinamiento.

Para resolver esto, se ha desarrollado recientemente una versión borrada del ARN-2 de CPMV (Cañizares *et al.*, 2006). En este sistema, se han eliminado la región del ARN-2 que codifica la proteína del movimiento y ambas proteínas de la cubierta. Sin embargo, las moléculas borradas siguen teniendo las secuencias de acción en *cis* necesarias para su replicación mediante la replicasa codificada por ARN-1 y de esta forma se mantienen niveles elevados de amplificación génica sin la posibilidad concomitante de que el virus modificado contamine el medio ambiente. Con la inclusión de un supresor del silenciamiento génico, tal como HcPro de PVY, (Brigneti *et al.*, 1998) en el inóculo además del ARN-1, el vector CPMV borrado se puede usar como un sistema de expresión transitoria (documento WO/2007/135480) Bipartite System, Method And Composition For The Constitutive And Inducible Expression Of High Levels Of Foreign Proteins In Plants; también Sainsbury *et al.*, 2009). Sin embargo, a diferencia de la situación con un vector basado en el ARN-2 de longitud completa, la replicación queda restringida a las hojas inoculadas. Estos vectores de CPMV se han utilizado para expresar complejos multicadena que consisten en un solo tipo de polipéptido.

También se ha demostrado que múltiples copias de vectores basados bien en las versiones de longitud completa o en versiones borradas del ARN-2 de CPMV son adecuadas para la producción de proteínas heteroméricas en plantas (Sainsbury *et al.*, 2008). Se ha utilizado la infiltración simultánea de dos construcciones de ARN-2 de longitud completa que incluyen diferentes genes marcadores en *Nicotiana benthamiana* en presencia de ARN-1 para mostrar que dos proteínas extrañas se pueden expresar eficazmente dentro de la misma célula en tejido inoculado. Además, las proteínas se puede localizar simultáneamente en los mismos compartimentos subcelulares, que es un prerrequisito esencial de la formación de heterómeros.

También se ha investigado la adecuabilidad de los diferentes vectores de ARN-2 de CPMV para la expresión de proteínas heteroméricas en plantas. La inserción de las cadenas ligera y pesada de una IgG en las versiones de ARN-2 tanto de longitud completa como borrada demostró que ambos enfoques conducían a la acumulación de moléculas de IgG de longitud completa en el tejido inoculado, pero que los niveles eran considerablemente mayores cuando se usaban vectores de ARN-2 borrados. La capacidad de las construcciones de ARN-2 de longitud completa para diseminarse sistémicamente parece, por tanto, ser irrelevante para la producción de proteínas heteroméricas, y el uso de versiones de ARN-2 borradas es claramente ventajoso, especialmente porque también ofrecen la ventaja del bioconfinamiento.

Por tanto, los sistemas vectores conocidos basados en CPMV representan herramientas útiles para la expresión de un gen heterólogo que codifica una proteína de interés en plantas. Sin embargo, sigue existiendo necesidad en la técnica de sistemas vectores óptimos para mejorar, por ejemplo, el rendimiento de las proteínas heterólogas expresadas y la facilidad de uso del vector.

Sumario de la invención

Los presentes inventores han descubierto sorprendentemente que la mutación del codón de inicio en la posición 161 de un vector de ARN-2 de CPMV aumenta notablemente los niveles de expresión de una proteína codificada por un gen insertado después del codón de inicio en la posición 512. Los niveles de expresión de proteína aumentaron aproximadamente 20-30 veces en comparación con la expresión de la misma proteína a partir de un vector de ARN-2 de CPMV que solamente difiere en que el codón de inicio en la posición 161 estaba intacto (Sainsbury y Lomonossoff, 2008). La presente invención permite la producción de elevados niveles de proteínas extrañas sin necesidad de replicación vírica.

Los inventores han descubierto también que la mutación del codón de inicio en la posición 161 niega la necesidad de mantener el marco entre la posición del codón de inicio en la posición 161 mutado y el codón de inicio en la posición 512, permitiendo de este modo la inserción de secuencias de cualquier longitud después del codón de inicio en la posición 161 mutado. Esto es especialmente ventajoso ya que permite insertar polienlazadores de cualquier longitud en vectores de ARN-2 tras el codón de inicio mutado, que se pueden usar después para facilitar la clonación de un gen de interés en el vector.

Además, los inventores han descubierto que, a pesar del aumento en la expresión de la proteína, las plantas transformadas con un vector de ARN-2 de CPMV que comprenden un codón de inicio en la posición 161 mutado tenían un aspecto más sano, es decir, mostraron menos necrosis, que plantas transformadas con vectores de ARN-2 de CPMV conocidos. La salud de las plantas es un factor importante en la expresión de proteínas a partir de plantas, ya que las plantas sanas sobreviven durante periodos de tiempo más prolongados. Además, la salud de la planta también es importante para la purificación de proteínas a partir de las plantas, ya que los taninos liberados como resultado de la necrosis interfieren con la purificación de proteínas (Sainsbury y Lomonossoff, 2008).

Así, la presente invención se refiere a sistemas y métodos mejorados de producción de proteínas, basados en secuencias modificadas del virus bipartido.

5 Así, en los diferentes aspectos de la invención tal como se define en las reivindicaciones, se utiliza una secuencia potenciadora de la expresión, secuencia que se deriva de (o comparte homología con) el segmento del genoma de ARN-2 de un virus de ARN bipartido, tal como un comovirus, en el que se ha mutado un sitio de inicio diana.

10 La presente invención proporciona además procesos como se define en las reivindicaciones para aumentar la expresión o potenciar la actividad traduccional de una secuencia derivada de un segmento derivado del genoma de ARN-2 de un virus de ARN bipartido, donde dichos procesos comprenden la mutación de un sitio de inicio en el mismo.

Algunas definiciones y realizaciones particulares de la invención se describirán ahora con más detalle.

15 Secuencias "potenciadoras" (o elementos potenciadores), tal como se cita en el presente documento, son secuencias que se derivan (o comparten homología con) el segmento del genoma de ARN-2 de un virus de ARN bipartido, tal como un comovirus, en el que se ha mutado un sitio de inicio diana. Dichas secuencias pueden potenciar la expresión en dirección 3' de un ORF heterólogo al que están unidas. Sin limitación, se cree que dichas secuencias, cuando están presentes en el ARN transcrito, pueden potenciar la traducción de un ORF heterólogo al que están unidas.

20

25 Un "sitio de inicio diana" tal como se cita en el presente documento, es el sitio de inicio (codón de inicio) de un segmento del genoma de ARN-2 de un virus de ARN bipartido de tipo salvaje (por ejemplo, un comovirus) del que se deriva el potenciador de secuencia en cuestión, que sirve como sitios de inicio para la producción (traducción) de la más larga de las dos carboxiproteínas coterminales codificadas por el segmento del genoma de ARN-2 de tipo salvaje.

30 Tal como se ha descrito anteriormente, de la más larga de las dos carboxiproteínas coterminales codificadas por el ARN-2 de CPMV, la proteína de 105K, se inicia en el sitio de inicio en la posición 161 del segmento del genoma de ARN-2 de CPMV de tipo salvaje. Por tanto, las secuencias potenciadoras del sitio de inicio diana derivadas del segmento del genoma de ARN-2 de CPMV es el sitio de inicio en la posición 161 del ARN-2 de CPMV de tipo salvaje.

35 Las mutaciones alrededor del codón de inicio en la posición 161 pueden tener el mismo efecto (o un efecto similar) a la propia mutación del codón de inicio en la posición 161, por ejemplo, perturbando el contexto alrededor de este codón de inicio puede significar que el codón de inicio se deriva con mayor frecuencia.

40 Un sitio de inicio diana se puede 'mutar' indirectamente mediante la mutación de uno o más nucleótidos en dirección 5' y/o en dirección 3' del sitio de inicio diana, pero reteniendo el sitio de inicio diana de tipo salvaje, en el que el efecto de la mutación de dichos nucleótidos es el mismo, o similar, al efecto observado cuando se muta el propio sitio de inicio diana.

45 Puesto que los sitios de inicio diana sirven como el de unión para la producción de la más larga de las dos carboxiproteínas coterminales codificadas por el segmento del genoma de ARN-2 de tipo salvaje, resulta que los sitios de inicio diana están en el marco (en fase) con un segundo sitio de inicio en el mismo segmento del genoma de ARN-2 de tipo salvaje, que sirve como sitio de inicio para la producción de la más corta de las dos carboxiproteínas coterminales codificadas por el ARN-2 de tipo salvaje. Dos sitios de inicio están en marco si están en el mismo marco de lectura del triplete.

50 El sitio de inicio diana de las secuencias potenciadoras derivadas del segmento del genoma de ARN-2 de CPMV de tipo salvaje, es decir, el sitio de inicio en la posición 161, está en marco con respecto al sitio de inicio en la posición 512, que sirve como sitio de inicio para la producción de la más corta de las dos carboxiproteínas coterminales codificadas por el ARN-2 de CPMV (la proteína de 95K) del segmento del genoma de ARN-2 de CPMV de tipo salvaje.

55

60 Por tanto, un sitio de inicio diana se ubica antes (en 5') de un segundo sitio de inicio en el segmento del genoma de ARN-2 de tipo salvaje del que se deriva la secuencia potenciadora, que sirve como el sitio de inicio para la producción de la más corta de las dos carboxiproteínas coterminales codificadas por el segmento del genoma de ARN-2 de tipo salvaje. Además, un sitio de inicio diana también se puede ubicar se ubica después (en 3') de un tercer sitio de inicio en el segmento del genoma de ARN-2 de tipo salvaje del que se deriva la secuencia potenciadora. En CPMV, el sitio de inicio diana, es decir, el sitio de inicio en la posición 161, está situado en dirección 5' de un segundo sitio de inicio en posición 512 que sirve como sitio de inicio para la producción de la proteína de 95K y en dirección 3' de un tercer sitio de inicio en la posición 115.

65 Un sitio de inicio diana de una secuencia potenciadora derivada del segmento del genoma de ARN-2 de un virus de ARN bipartido es por tanto el primer o el segundo sitio de inicio para la producción de dos carboxiproteínas

coterminal codificadas por el ARN-2 de tipo salvaje. 'Primero' en este contexto se refiere al sitio de inicio situado más cerca del extremo 5' del segmento del genoma de ARN-2 de tipo salvaje.

5 Se puede mutar más de un sitio de unión en la secuencia, si se desea. Por ejemplo, el 'tercer' sitio de inicio en (o correspondiente a) la posición 115 también se puede borrar o alterar. Se ha demostrado que la eliminación de AUG 115 además de la eliminación de AUG 161, potencia adicionalmente la expresión (Sainsbury y Lomonosoff, 2008).

10 Las secuencias potenciadoras de la presente invención se basan en secuencias modificadas procedentes de los segmentos de genoma del ARN-2 de los virus de ARN bipartidos de la familia *Comoviridae*. Todos los géneros de la familia *Comoviridae* parecen codificar dos carboxiproteínas coterminal. Los géneros de la familia *Comoviridae* incluyen Comovirus, Nepovirus, Fabavirus, Cheravirus y Sadwavirus. Los comovirus incluyen *virus del mosaico del caupí* (CPMV), *virus grave del mosaico del caupí* (CPSMV), *virus del mosaico de la calabaza* (SqMV), *virus del moteado del trébol rojo* (RCMV), *virus del moteado de la maceta de judías* (BPMV). Preferentemente, el virus bipartido (o comovirus) es CPMV.

15 Las secuencias de los segmentos del genoma de ARN-2 de estos comovirus y varias cepas específicas están disponibles de la base de datos de NCBI con los números de acceso relacionados entre paréntesis: ARN-2 de *virus del mosaico del caupí* (NC_003550), ARN-2 de *virus grave del mosaico del caupí* (NC_003544), ARN-2 de *virus del mosaico de la calabaza* (NC_003800), ARN-2 de *virus del mosaico de la calabaza* cepa Kimble (AF059533), ARN-2 de *virus del mosaico de la calabaza* cepa Arizona (AF059532), ARN-2 de *virus del moteado del trébol rojo* (NC_003738), ARN-2 de *virus del moteado de la maceta de judías* (NC_003495), ARN-2 de *virus del moteado de la maceta de judías* cepa K-Hopkins1 (AF394609), ARN-2 de *virus del moteado de la maceta de judías* cepa K-Hancock1 (AF394607), *virus del moteado de la patata andina* (APMoV: L16239) y *virus del mosaico del rábano* (RaMV; AB295644). También están disponibles secuencias de ARN-2 disponibles del *virus del mosaico de la judía rugosa* (BRMV; AF263548) y un posible miembro del género Comovirus, *virus del moteado del nabo* (EF191015). También están disponibles numerosas secuencias de otros géneros de la familia *Comoviridae*.

30 Hasta la fecha, todos los comovirus que se han investigado han mostrado tener dos codones de inicio alternativos para la expresión de dos policarboxiproteínas coterminal a partir de sus segmentos del genoma de ARN-2. En particular, los segmentos del genoma de ARN-2 de CPMV, CPSMV, BPMV, SqMV y RCMV son conocidos por comprender dos sitios de inicio alternativos para la expresión de dos policarboxiproteínas coterminal.

35 Los sitios de inicio diana en otros comovirus, que son equivalentes al sitio de inicio en la posición 161 del segmento de ARN-2 de CPMV de tipo salvaje (es decir, correspondiente al mismo) se pueden identificar por métodos conocidos en la materia. Por ejemplo, se pueden identificar sitios de inicio diana mediante una alineación de secuencias entre la secuencia del segmento del genoma de ARN-2 de tipo salvaje de CPMV y la secuencia del segmento del genoma de ARN-2 de otro comovirus. Dichas alineaciones de secuencia se pueden usar después para identificar un sitio de inicio diana en la secuencia del segmento del genoma de ARN-2 comovirica mediante la identificación de un sitio de inicio que, al menos en la alineación, está cerca, o en la misma posición que, el sitio de inicio diana en la posición 161 del ARN-2 de CPMV de tipo salvaje.

45 Los sitios de inicio diana en otros comovirus también se pueden identificar determinando el codón de inicio que sirve como sitio de inicio para la síntesis de la más larga de las dos carboxiproteínas coterminal codificadas por el segmento del genoma de ARN-2 del comovirus de tipo salvaje. Este enfoque también se puede usar junto con una alineación como se ha descrito anteriormente, es decir, este enfoque se puede usar para confirmar que un sitio de inicio de un comovirus identificado mediante una alineación con un ARN-2 de CPMV es un sitio de inicio diana.

50 Por supuesto, los métodos anteriores también se pueden usar para identificar sitios de inicio en otros comoviral segmentos del genoma de ARN-2 del comovirus que sean equivalentes al sitio de inicio en la posición 512 del segmento del genoma de ARN-2 de CPMV de tipo salvaje. Sin embargo, en lugar de identificar el codón de inicio que sirve como sitio de inicio para la síntesis de la más larga de las dos carboxiproteínas coterminal codificadas por el segmento del genoma de ARN-2 del comovirus de tipo salvaje, se identifica el codón de inicio que sirve como sitio de inicio para la síntesis de la más corta de las dos carboxiproteínas coterminal codificadas por el segmento del genoma de ARN-2 del comovirus de tipo salvaje.

55 Una vez que se han identificado los dos sitios de inicio del ARN-2 del comovirus que podrían ser equivalentes a los sitios de inicio en las posiciones 161 y 512 del ARN-2 de CPMV, la identificación del sitio de inicio diana se puede confirmar comprobando que los dos sitios de inicio están en el mismo marco, es decir, en el mismo marco de lectura del triplete, ya que solo pueden servir como sitios de inicio para la producción de dos carboxiproteínas coterminal si ese es el caso.

60 En una realización de la invención, la secuencia potenciadora comprende los nucleótidos 1 a 512 del segmento del genoma de ARN-2 de CPMV (véase la Tabla 1), en la que el sitio de inicio diana en la posición 161 se ha mutado. En otra realización de la invención, la secuencia potenciadora comprende una secuencia equivalente de otro comovirus, en la que el sitio de inicio diana equivalente al codón de inicio en la posición 161 se ha mutado. El sitio de inicio diana se puede mutar por sustitución, delección o inserción. Preferentemente, el sitio de inicio diana se muta

con una mutación puntual.

5 En realizaciones alternativas de la invención, la secuencia potenciadora comprende los nucleótidos 10 a 512, 20 a 512, 30 a 512, 40 a 512, 50 a 512, 100 a 512, 150 a 512, 1 a 514, 10 a 514, 20 a 514, 30 a 514, 40 a 514, 50 a 514, 100 a 514, 150 a 514, 1 a 511, 10 a 511, 20 a 511, 30 a 511, 40 a 511, 50 a 511, 100 a 511, 150 a 511, 1 a 509, 10 a 509, 20 a 509, 30 a 509, 40 a 509, 50 a 509, 100 a 509, 150 a 509, 1 a 507, 10 a 507, 20 a 507, 30 a 507, 40 a 507, 50 a 507, 100 a 507, o 150 a 507 de un de una secuencia del segmento del genoma de ARN-2 del comovirus con un sitio de inicio diana mutado. En otras realizaciones de la invención, la secuencia potenciadora comprende los nucleótidos 10 a 512, 20 a 512, 30 a 512, 40 a 512, 50 a 512, 100 a 512, 150 a 512, 1 a 514, 10 a 514, 20 a 514, 30 a 514, 40 a 514, 50 a 514, 100 a 514, 150 a 514, 1 a 511, 10 a 511, 20 a 511, 30 a 511, 40 a 511, 50 a 511, 100 a 511, 150 a 511, a 509, 10 a 509, 20 a 509, 30 a 509, 40 a 509, 50 a 509, 100 a 509, 150 a 509, 1 a 507, 10 a 507, 20 a 507, 30 a 507, 40 a 507, 50 a 507, 100 a 507, o 150 a 507 de la secuencia del segmento del genoma de ARN-2 de CPMV mostrada en la Tabla 1, en el que el sitio de inicio diana en la posición 161 del segmento del genoma del ARN-2 de CPMV de tipo salvaje se ha mutado.

15 En algunas realizaciones de la invención, la secuencia potenciadora comprende los nucleótidos 1 a 500, 1 a 490, 1 a 480, 1 a 470, 1 a 460, 1 a 450, 1 a 400, 1 a 350, 1 a 300, 1 a 250, o 1 a 200 de un de una secuencia del segmento del genoma de ARN-2 del comovirus con un sitio de inicio diana mutado.

20 En realizaciones alternativas de la invención, la secuencia potenciadora comprende los nucleótidos 1 a 500, 1 a 490, 1 a 480, 1 a 470, 1 a 460, 1 a 450, 1 a 400, 1 a 350, 1 a 300, 1 a 250, o 1 a 200, del segmento del genoma de ARN-2 de CPMV mostrada en la Tabla 1, en el que el sitio de inicio diana en la posición 161 del segmento del genoma del ARN-2 de CPMV de tipo salvaje se ha mutado.

25 Las secuencias potenciadoras comprenden al menos 200, al menos 300, al menos 350, al menos 400, al menos 450, al menos 460, al menos 470, al menos 480, al menos 490 o al menos 500 nucleótidos de una secuencia del segmento del genoma de ARN-2 del comovirus con un sitio de inicio diana mutado son también realizaciones de la invención.

30 Además, las secuencias potenciadoras comprenden al menos 200, al menos 300, al menos 350, al menos 400, al menos 450, al menos 460, al menos 470, al menos 480, al menos 490 o al menos 500 nucleótidos de la secuencia del segmento del genoma de ARN-2 de CPMV mostrada en la Tabla 1, en el que el sitio de inicio diana en la posición 161 del segmento del genoma del ARN-2 de CPMV de tipo salvaje se ha mutado, son también realizaciones de la invención.

35 Las realizaciones alternativas de la invención son secuencias potenciadoras que tienen al menos un 99 %, 98 %, 97 %, 96 %, 95 %, 90 %, 85 %, 80 %, 75 %, 70 %, de identidad con el segmento del genoma de ARN-2 de CPMV mostrada en la Tabla 1, en el que el sitio de inicio diana en la posición 161 del segmento del genoma del ARN-2 de CPMV de tipo salvaje se ha mutado.

40 Los términos "porcentaje de similitud", "porcentaje de identidad" y "porcentaje de homología", cuando se refieren a una secuencia particular, se utilizan como se define en el programa informático GCG de la Universidad de Wisconsin. De este modo, las secuencias potenciadoras pueden hibridarse específicamente con la secuencia complementaria del segmento del genoma de ARN-2 de CPMV mostrada en la Tabla 1, con la condición de que el sitio de inicio diana corresponde a la posición 161 del segmento del genoma de ARN-2 de CPMV de tipo salvaje se haya mutado.

45 La expresión "hibrida específicamente" se refiere a la asociación entre dos moléculas de ácido nucleico monocatenarias con una complementariedad de secuencia suficiente para permitir dicha hibridación en condiciones predeterminadas generalmente utilizadas en la técnica (a veces denominada "sustancialmente complementaria"). En particular, el término se refiere a la hibridación de un oligonucleótido con una secuencia sustancialmente complementaria incluida en una molécula monocatenaria de ADN o ARN de la invención, para la exclusión sustancial de la hibridación del oligonucleótido con ácidos nucleicos monocatenarios de secuencia no complementaria. "Complementario" se refiere a la asociación de tipo salvaje de secuencias de ácido nucleico mediante emparejamiento de bases (parejas A-G-T con la secuencia complementaria T-C-A). La complementariedad entre dos moléculas monocatenarias puede ser parcial, si solamente parte de las parejas de ácidos nucleicos son complementarias; o completa, si todos los pares de bases son complementarios. El grado de complementariedad afecta a la eficacia y la intensidad de la hibridación y de las condiciones de amplificación.

50 Un sitio de inicio diana de una secuencia potenciadora de la invención se puede mutar por delección, inserción o sustitución, de forma que ya no funcione como un sitio de inicio de traducción. Por ejemplo, se puede realizar una mutación puntual en la posición del sitio de inicio diana de la secuencia potenciadora. Como alternativa, el sitio de inicio diana de la secuencia potenciadora se puede borrar bien parcialmente o en su totalidad. Por ejemplo, se puede realizar una delección que amplíe el sitio de inicio diana de la secuencia potenciadora. Las delecciones que amplían el sitio de inicio puede tener un máximo de 5, máximo de 10, máximo de 15, máximo de 20, máximo de 25, máximo de 30, máximo de 35, máximo de 40, máximo de 45, o un máximo de 50 nucleótidos de longitud, cuando se

compara con la secuencia del segmento del genoma de ARN-2 de tipo salvaje a partir de la que se deriva la secuencia potenciadora.

5 Sin desear quedar ligados a teoría alguna, se cree que la mutación del codón de inicio en la posición 161 de CPMV conduce a la inactivación de un supresor traduccional, que da como resultado un inicio potenciado de la traducción a partir de los codones de inicio situados en la dirección 3' del supresor traduccional inactivado.

10 Por tanto, la presente divulgación proporciona una secuencia potenciadora derivada de un segmento del genoma de ARN-2 de un virus de ARN bipartido, en la que la secuencia potenciadora comprende una secuencia supresora de la traducción inactivada.

15 La presente invención proporciona además un proceso como se define en las reivindicaciones para aumentar la expresión o potenciar la actividad traduccional de una secuencia derivada de un segmento derivado del genoma de ARN-2 de un virus de ARN bipartido, proceso que comprende la inactivación de una secuencia supresora de la traducción en el anterior.

20 Como ya se ha mencionado anteriormente, se cree que la mutación del sitio de inicio en la posición 161 en el segmento del genoma de ARN-2 de CPMV conduce a la inactivación de un supresor de la traducción normalmente presente en el ARN-2 de CPMV.

25 Una secuencia supresora de la traducción, tal como se cita en el presente documento, es una secuencia del segmento del genoma de ARN-2 del virus de ARN bipartido de tipo salvaje (por ejemplo, un comovirus) del que se deriva el potenciador de secuencia en cuestión, que comprende, o consiste de, el sitio de inicio para la producción (traducción) de la más larga de las dos carboxiproteínas coterminales codificadas por el segmento del genoma de ARN-2 de tipo salvaje.

30 Las secuencias supresoras de la traducción en las secuencias potenciadoras derivadas del segmento del genoma de ARN-2 de CPMV, son secuencias que comprenden, o consisten de, el sitio de inicio diana anteriormente descrito. Por tanto, las secuencias supresoras de la traducción comprenden, o consisten de, un sitio de inicio diana tal como se ha definido anteriormente, y se puede inactivar mediante mutagénesis como se ha descrito anteriormente.

35 Las secuencias potenciadoras anteriormente definidas se pueden usar en varios aspectos y realizaciones de la invención de la siguiente forma.

Los aspectos de la presente invención tal como se definen en las reivindicaciones pueden utilizar un ácido nucleico aislado que consiste de, o consiste esencialmente de, una secuencia potenciadora de la expresión como se ha descrito anteriormente.

40 "Ácido nucleico" o una "molécula de ácido nucleico" tal como se usa en el presente documento se refiere a cualquier molécula de ADN o ARN, tanto monocatenaria como bicatenaria y, en caso de ser monocatenaria, la molécula de su secuencia complementaria en forma tanto lineal como circular. En la descripción de las moléculas de ácido nucleico, una secuencia o estructura de una molécula de ácido particular se puede describir en el presente documento de acuerdo con la convención normal de proporcionar la secuencia en la dirección de 5' a 3'. Con referencia a los ácidos nucleicos de la invención, a veces se usa la expresión "ácido nucleico aislado". Esta expresión, cuando se aplica al ADN, se refiere a una molécula de ADN que está separada de las secuencias de la que es inmediatamente continua en el genoma de tipo salvaje del organismo a partir del que se origina.

45 Por ejemplo, un "ácido nucleico aislado" puede comprender una molécula de ADN insertada en un vector, tal como un plásmido o un vector vírico, o integrado en el ADN genómico de una célula procarionta o eucariota u organismo hospedador.

50 Cuando se aplica al ARN, la expresión "ácido nucleico aislado" se refiere principalmente a una molécula de ARN codificada mediante una molécula de ADN aislada tal como se ha definido anteriormente. Como alternativa, la expresión puede referirse a una molécula de ARN que está lo suficientemente separada de otros ácidos nucleicos que se asociarían con su estado de tipo salvaje (es decir, en células o tejidos). Un "ácido nucleico aislado" (tanto de ADN como de ARN) puede representar adicionalmente una molécula producida directamente por medios biológicos o sintéticos y separados de otros componentes presentes durante su producción.

55 El ácido nucleico puede consistir, por tanto, esencialmente de una porción, o fragmento, del segmento del genoma de ARN-2 del virus de ARN bipartido del que se deriva el potenciador. Por ejemplo, en una realización, el ácido nucleico no comprende al menos una porción de la región de codificación del segmento del genoma de ARN-2 del que se deriva. La región de codificación puede ser la región del segmento del genoma de ARN-2 que codifica la más corta de las dos carboxiproteínas coterminales. El ácido nucleico puede consistir, o consiste esencialmente de la porción de un segmento del genoma de ARN-2 de un virus de ARN bipartido que se extiende desde el extremo 5' del segmento del genoma de ARN-2 de tipo salvaje al sitio de inicio a partir del que se inicia la producción (traducción) de la más corta de las dos carboxiproteínas coterminales codificadas por el segmento del genoma de ARN-2 de tipo

salvaje.

5 La expresión "consiste esencialmente de" cuando se refiere a un nucleótido o aminoácido particular significa una secuencia que tiene las propiedades de una SEC ID N° dada. Por ejemplo, cuando se usa en referencia a una secuencia de aminoácidos, la expresión incluye la secuencia por sí misma y las modificaciones moleculares que no afectan las características básicas y novedosas de la secuencia. Por ejemplo, cuando se usa en referencia a un ácido nucleico, la expresión incluye la secuencia por sí misma y los cambios menores y/o extensiones que no afectan la función potenciadora de la secuencia, o proporcionar mayor funcionalidad (adicional).

10 La invención se refiere además a sistemas de expresión génica tal como se ha definido en las reivindicaciones que comprende una secuencia potenciadora de la invención.

15 Por tanto, en otro aspecto, la presente invención proporciona un sistema de expresión génica define en las reivindicaciones que comprende una secuencia potenciadora como se ha descrito anteriormente.

El sistema de expresión génica puede comprender también un gen que codifica una proteína de interés insertada en dirección 3' de la secuencia potenciadora. Las secuencias insertadas que codifican una proteína de interés pueden tener cualquier tamaño.

20 En un aspecto adicional, por tanto, la presente invención proporciona un sistema de expresión génica como se define en las reivindicaciones que comprende:

(a) una secuencia potenciadora como se ha descrito anteriormente; y (b) un gen heterólogo que codifica una proteína de interés, en el que el gen está situado en dirección 3' de la secuencia potenciadora.

25 El gen y la proteína de interés son heterólogos, es decir, no codificados mediante el virus de ARN bipartido de tipo salvaje a partir del que se deriva la secuencia potenciadora.

30 Los sistemas de expresión génica se pueden usar para expresar una proteína de interés en un organismo hospedador. En este caso, la proteína de interés también puede ser heteróloga respecto al organismo hospedador en cuestión, es decir, introducirse dentro de las células en cuestión (por ejemplo, una planta o uno de sus antepasados) mediante ingeniería genética, es decir, mediante intervención humana. Un gen heterólogo de un organismo puede sustituir un gen endógeno equivalente, es decir, uno que realice normalmente una función igual o similar, o la secuencia insertada puede ser adicional al gen endógeno u otra secuencia.

35 Los expertos en la materia entenderán que la expresión de un gen de interés requerirá la presencia de un sitio de inicio (AUG) situado en dirección 5' del gen a expresar. Dichos sitios de inicio se pueden proporcionar tanto como parte de una secuencia potenciadora o como parte de un gen que codifica una proteína de interés.

40 El organismo hospedador puede ser una planta. Sin embargo, como los mecanismos traduccionales están bien conservados en los eucariotas, también se pueden usar los sistemas de expresión génica para expresar una proteína de interés en organismos hospedadores eucariotas diferentes a plantas, por ejemplo en células de insecto como vectores baculovirus modificados, o en levaduras o en células de mamífero.

45 Los sistemas de expresión génica pueden estar operativamente unidos a secuencias del promotor y del terminador.

Por tanto, los sistemas de expresión génica pueden comprender además una secuencia de terminación y el gen que codifica una proteína de interés se puede situar entre la secuencia potenciadora y la secuencia de terminación, es decir, en dirección 3' (3') de la secuencia potenciadora y en dirección 5' (5') de la secuencia de terminación.

50 También se proporciona en el presente documento un casete de expresión que comprende:

- (i) un promotor, unido de manera operativa a
- (ii) una secuencia potenciadora como se ha descrito anteriormente
- (iii) un gen de interés que se desea expresar
- (iv) una secuencia de terminación.

60 Preferiblemente el promotor usado para impulsar el gen de interés será un promotor fuerte. Los ejemplos de promotores para su uso en plantas incluyen:

- (1) p35S: Odell *et al.*, 1985
- (2) promotor del virus del mosaico de la vena de la mandioca, pCAS, Verdaguer *et al.*, 1996
- (3) Promoter de la subunidad pequeña de la ribulosa bifosfato carboxilasa, pRbcS: Outchkourov *et al.*, 2003.

65 Otros promotores fuertes incluyen pUbi (para monocotiledóneas y dicotiledóneas) y pActin.

En una realización preferida, el promotor es un promotor inducible.

El término "inducible" tal como se aplica a un promotor es bien comprendido por los expertos en la materia. Esencialmente, la expresión bajo el control de un promotor inducible está "conectado" o aumentado en respuesta a un estímulo aplicado. La naturaleza del estímulo varía entre promotores. Algunos promotores inducibles ocasionan niveles de expresión bajos o indetectables (o sin expresión) en ausencia del estímulo adecuado. Otros promotores inducibles ocasionan una expresión constitutiva detectable en ausencia del estímulo. Cualquiera que sea el nivel de expresión en ausencia de estímulo, la expresión de cualquier promotor inducible aumenta en la presencia del estímulo correcto.

La secuencia de terminación (terminador) puede ser una secuencia de terminación derivada del segmento del genoma de ARN-2 de un virus de ARN bipartido, por ejemplo un comovirus. En una realización, la secuencia de terminación se puede derivar del mismo virus de ARN bipartido del que se deriva la secuencia potenciadora. La secuencia de terminación puede comprender un codón de detención. La secuencia de terminación también puede ir seguida de señales de poliadenilación.

Los sistemas de expresión génica de la invención también pueden comprender una región no traducida (UTR, por sus siglas en inglés). La UTR se puede situar en la dirección 5' de una secuencia terminadora presente en el casete de expresión génica, construcción de expresión génica o sistema de expresión génica. Cuando los sistemas de expresión génica comprenden un gen que codifica una proteína de interés, la UTR puede estar situada en dirección 3' de dicho gen.

Por tanto, la UTR puede estar situada entre un gen que codifica una proteína de interés y una secuencia de terminación. La UTR se puede derivar de un virus de ARN bipartido, por ejemplo, a partir del segmento del genoma de ARN-2 de un virus de ARN bipartido. La UTR puede ser la UTR a 3' del mismo segmento del genoma de ARN-2 del que se deriva la secuencia potenciadora presente en el casete de expresión génica, construcción de expresión génica o sistema de expresión génica. Preferentemente, la UTR es la UTR a 3' de un segmento del genoma de ARN-2 del comovirus, por ejemplo, la UTR a 3' de un segmento del genoma de ARN-2 del CPMV.

Tal como se ha descrito anteriormente, se había demostrado anteriormente que para conseguir una replicación eficaz del ARN-2 de CPMV mediante la replicasa de CPMV codificada por ARN-1 era fundamental que se mantuviera el marco entre los sitios de inicio en las posiciones 161 y 512 del ARN-2, es decir, que los dos sitios de inicio permanecieran en el mismo marco de lectura del triplete (Holness *et al.*, 1989; van Bokhoven *et al.*, 1993; Rohll *et al.*, 1993). Este requisito limita la longitud de las secuencias que se pueden insertar en dirección 5' del sitio de inicio en la posición 512 en los vectores de expresión basados en CPMV. En particular, se impide el uso de polienlazadores, ya que el uso de estos frecuentemente altera el marco de lectura abierto (ORF) entre estos sitios de inicio.

Los presentes inventores han demostrado que el mantenimiento del marco de lectura entre los sitios de inicio en las posiciones 161 y 512 en el ARN-2 de CPMV es también un requisito para conseguir un inicio eficaz de traducción en el sitio de inicio en la posición 512, es decir, es un requisito para conseguir una expresión eficaz de la más corta de las dos carboxiproteínas coterminales codificadas mediante CPMV (la proteína de 95K).

Sin embargo, los presentes inventores han demostrado también que la mutación del sitio de inicio en la posición 161 en el ARN-2 de CPMV permite la inserción de secuencias en dirección 5' del sitio de inicio en la posición 512, que alteran el marco entre el codón de inicio mutado y el sitio de inicio en la posición 512, sin ningún efecto negativo sobre el nivel de expresión de la proteína de 95K. En consecuencia, la mutación del sitio de inicio en la posición 161 significa que ya no hay ninguna restricción en la longitud de las secuencias que se pueden insertar en dirección 5' del sitio de inicio en la posición 512.

Aunque también se requiere el mantenimiento del marco de lectura entre los sitios de inicio que codifican las dos carboxiproteínas coterminales en otros virus bipartidos, este requisito también se puede superar mutando la AUG que sirve como sitio de inicio para producciones de la más larga de las dos carboxiproteínas coterminales codificadas mediante el segmento del genoma de ARN-2 del virus. Por tanto, en otro aspecto, la presente invención proporciona una construcción de expresión génica de acuerdo con las reivindicaciones que comprende:

- (a) una secuencia potenciadora como se ha descrito anteriormente; y
- (b) una secuencia heteróloga para facilitar la inserción de un gen que codifica una proteína de interés en el sistema de expresión génica, en la que la secuencia heteróloga está situada en dirección 3' del sitio de inicio diana mutado en la secuencia potenciadora.

La secuencia heteróloga se puede localizar en dirección 5' del codón de inicio a partir del que se inicia la producción de la más corta de las dos carboxiproteínas coterminales en el segmento del genoma de ARN-2 de tipo salvaje del que se deriva la secuencia potenciadora del sistema de expresión génica. Como alternativa, la secuencia heteróloga se puede proporcionar alrededor del sitio del codón de inicio, o sustituir el codón de inicio, a partir del que se inicia la producción de la más corta de las dos carboxiproteínas coterminales en el segmento del genoma de ARN-2 de tipo

salvaje del que se deriva la secuencia potenciadora del sistema de expresión génica. En un sistema de expresión génica con una secuencia potenciadora derivada del ARN-2 de CPMV, la secuencia heteróloga se puede proporcionar en dirección 5' de, alrededor del sitio de, o sustituir, el codón de inicio que está en posición 512 del segmento del genoma de ARN-2 de CPMV de tipo salvaje.

5 La secuencia heteróloga puede ser un polienlazador o un sitio de clonación múltiple, es decir, una secuencia que facilita la clonación de un gen que codifica una proteína de interés en el sistema de expresión génica.

10 Por ejemplo, como se describe a partir de ahora en el presente documento, los presentes inventores han proporcionado construcciones que incluyen un polienlazador entre la secuencia líder en 5' y las UTR en 3' de un casete de expresión basado en CPMV. Como se describe a continuación, cualquier polienlazador puede codificar opcionalmente uno o más conjuntos de restos histidina para permitir la fusión de las etiquetas His de los extremos N o C para facilitar la purificación de la proteína.

15 Preferentemente, las construcciones de expresión anteriores están presentes en un vector, y preferentemente comprenden secuencias frontera que permiten la transferencia y la integración del casete de expresión en el genoma del organismo.

20 Preferentemente, la construcción es un vector binario vegetal. Preferentemente, el vector binario de transformación está basado en pPZP (Hajdukiewicz, *et al.* 1994). Otros ejemplos de construcciones incluyen pBin19 (véase Frisch, D. A., L. W. Harris-Haller, *et al.* (1995). "Complete Sequence of the binary vector Bin 19." *Plant Molecular Biology* 27: 405-409).

25 Como se describe en el presente documento, la invención se puede realizar moviendo un casete de expresión con los componentes necesarios a un casete de expresión pBin existente, o en otras realizaciones, se puede utilizar un vector de expresión pBin de clonación directa.

30 Por ejemplo, como se describe a partir de ahora en el presente documento, los presentes inventores tienen vectores binarios modulares diseñados para (pero no restringidos a) su uso con las secuencias potenciadoras descritas en el presente documento. Se basan en mejoras del vector pBINPLUS, mediante el que se ha demostrado que es posible reducir de forma importante el tamaño del vector sin comprometer el rendimiento en términos de replicación y transferencia de ADN-T. Además, los elementos del sistema potenciador (ilustrados mediante el denominado sistema "CPMV-HT") se han incorporado al vector resultante de forma modular de forma que se pueden expresar múltiples proteínas a partir de un único ADN-T. Estas mejoras han conducido a la creación de un vector versátil con elevado nivel de expresión que permite una eficaz clonación directa de genes extraños.

35 Estos ejemplos representan vectores binarios vegetales preferidos. Preferentemente, incluyen el origen de replicación ColEI, aunque los plásmidos que incluyen otros orígenes de replicación también proporcionan un alto número de copias (tales como los plásmidos basados en pRi, Lee y Gelvin, 2008) son también preferidos, especialmente para sistemas de expresión transitoria.

40 Si se desea, se pueden incluir en la construcción marcadores genéticos seleccionables, como los que transmiten fenotipos seleccionables tales como la resistencia a antibióticos o a herbicidas (por ejemplo, kanamicina, higromicina, fosfinotricina, clorsulfurón, metotrexato, gentamicina, espectinomicina, imidazolinonas y glifosato).

45 Los más preferidos son los vectores pEAQ descritos a continuación que permiten la versión de clonación directa mediante el uso de un polienlazador entre la secuencia líder en 5' y las UTR en 3' de un casete de expresión que incluye un potenciador traduccional de la invención, colocado en un ADN-T que también contiene un supresor del silenciamiento génico y un casete NPTII. El polienlazador también codifica uno o dos conjuntos de 6 restos histidina para permitir la fusión de las etiquetas His de los extremos N o C para facilitar la purificación de la proteína.

50 Una ventaja de los vectores derivados de pEAQ es que cada componente de una proteína multicadena tal como una IgG se puede administrar automáticamente a una célula infectada.

55 La presente invención también proporciona métodos para expresar proteínas de acuerdo con las reivindicaciones, por ejemplo, proteínas heterólogas, en organismos hospedadores tales como plantas, levaduras, células de insecto o de mamífero, usando un sistema de expresión génica de la invención.

60 La presente invención proporciona además un método de acuerdo con las reivindicaciones para potenciar la traducción de una proteína heteróloga de interés a partir de un gen o un ORF que codifica la misma que está unido operativamente a una secuencia derivada de un ARN2 como se ha descrito anteriormente, comprendiendo el método mutar un sitio de inicio diana de la secuencia derivada de ARN2.

65 Las secuencias potenciadoras descritas en el presente documento también se pueden usar con sistemas de expresión bipartidos como se describe en el documento WO/2007/135480. La invención por tanto se refiere también a sistemas de expresión génica basados en segmentos génicos de ARN-2 truncados, que comprenden opcionalmente una segunda construcción génica que codifica un supresor del silenciamiento génico unido

operativamente a secuencias promotoras y de terminación.

Así, también se describe en el presente documento un sistema de expresión génica que comprende:

- 5 (a) una primera construcción génica que comprende un ARN-2 truncado de un genoma de un virus bipartido que incluye al menos un gen extraño que codifica una proteína heteróloga de interés unida operativamente a secuencias promotoras y de terminación, en la que la construcción génica comprende un sitio de inicio diana mutado en dirección 5' del gen extraño; y opcionalmente,
- 10 (b) una segunda construcción génica que comprende ARN-1 de dicho genoma de virus bipartido unido operativamente a secuencias promotoras y de terminación; y opcionalmente,
- (c) una tercera construcción génica, que incorpora opcionalmente en dicha primera construcción génica, dicha segunda construcción génica o a la vez, que comprende un supresor del silenciamiento génico unido operativamente a secuencias promotoras y de terminación.

15 Se prefiere la presencia de un supresor del silenciamiento génico en un sistema de expresión génica (incluyendo cualquiera de los descritos anteriormente) de la invención, pero no es esencial. Por tanto, un sistema de expresión génica, como se ha definido anteriormente, comprende preferentemente una tercera construcción génica, que incorpora opcionalmente en dicha primera construcción génica, dicha segunda construcción génica o a la vez, que comprende un supresor del silenciamiento génico unido operativamente a secuencias promotoras y de terminación.

20 Así, también se describe en el presente documento un método para expresar una proteína en una planta que comprende las etapas de:

- (a) introducir una construcción de expresión génica de la invención en una célula vegetal; y opcionalmente,
- 25 (b) introducir una segunda construcción génica que comprende ARN-1 de dicho genoma de virus bipartido unido operativamente a secuencias promotoras y de terminación en la célula vegetal; y opcionalmente,
- (c) introducir una tercera construcción génica, que incorpora opcionalmente en dicha primera construcción génica, dicha segunda construcción génica o a la vez, que comprende un supresor del silenciamiento génico unido operativamente a secuencias promotoras y de terminación en la célula vegetal.

30 Preferentemente, un método para expresar una proteína en una planta, como se ha definido anteriormente, comprende la etapa de introducir una tercera construcción génica, que incorpora opcionalmente en dicha primera construcción génica, dicha segunda construcción génica o a la vez, que comprende un supresor del silenciamiento génico unido operativamente a secuencias promotoras y de terminación en la célula vegetal.

35 La presente invención proporciona también métodos que comprenden la introducción de dicha construcción en una célula vegetal.

40 Los presentes inventores han mostrado niveles de expresión muy elevados incorporando tanto un gen de interés como un supresor del silenciamiento en el mismo ADN-T que el potenciador de la traducción. Por tanto, las realizaciones preferidas pueden utilizar los tres componentes presentes en el mismo ADN-T.

45 Adicionalmente, debe entenderse que el ARN-1 no es necesario para un nivel de expresión elevado en los sistemas descritos en el presente documento, y de este modo, el sistema "CPMV-HT" descrito en el presente documento no se encuentra bajo la acción del ARN-1.

Así, en un aspecto adicional, la presente invención se refiere a un sistema de expresión génica de acuerdo con las reivindicaciones que comprende:

- 50 (a) una primera construcción génica que comprende un ARN-2 truncado de un genoma de un virus bipartido que incluye al menos un gen extraño que codifica una proteína heteróloga de interés unida operativamente a secuencias promotoras y de terminación, en la que la construcción génica comprende un sitio de inicio diana mutado en dirección 5' del gen extraño; y opcionalmente,
- 55 (b) una segunda construcción génica que incorpora opcionalmente en dicha primera construcción génica, un supresor del silenciamiento génico unido operativamente a secuencias promotoras y de terminación.

Por tanto, en otro aspecto, la presente invención proporciona un método para expresar una proteína en una planta que comprende las etapas de:

- 60 (a) introducir una construcción de expresión génica de la invención en una célula vegetal; y opcionalmente,
- (b) introducir una segunda construcción génica que incorpora opcionalmente en dicha primera construcción génica, que comprende un supresor del silenciamiento génico unido operativamente a secuencias promotoras y de terminación en la célula vegetal.

65 Los supresores del silenciamiento génico útiles en estos aspectos son conocidos en la materia y se describen en el documento WO/2007/135480. Incluyen HcPro del virus de la patata Y, He-Pro de TEV, P19 de TBSV, rgsCam, la

proteína 82 de FHV, la proteína pequeña de la cubierta de CPMV, y la proteína de la cubierta de TCV. De manera más preferida, el ARN-2 del sistema está truncado de forma que no se produce un virus infeccioso.

5 Un supresor preferido cuando se producen plantas transgénicas estables es el supresor P19 que incorpora una mutación R43W.

Se describe además una célula hospedadora que contiene una construcción heteróloga de acuerdo con la presente invención.

10 Los vectores de expresión génica de la invención se pueden incorporar de forma estable o transitoria a las células vegetales.

Para producción a pequeña escala, se usa la agroinfiltración mecánica de las hojas con construcciones de la invención. El escalado se consigue mediante, por ejemplo, el uso de infiltración a vacío.

15 En otras realizaciones, se puede incorporar de forma estable un vector de expresión de la invención en el genoma de la planta transgénica o de la célula vegetal.

20 También se describe en el presente documento la etapa de regenerar una planta a partir de una célula vegetal transformada.

25 Los procedimientos y vectores específicos anteriormente utilizados con amplio éxito entre las plantas fueron descritos por Guerineau y Mullineaux (1993) (Plant transformation and expression vectors. En: Plant Molecular Biology Labfax (Croy RRD ed) Oxford, BIOS Scientific Publishers, pp 121-148). Los vectores adecuados pueden incluir vectores derivados de virus de plantas (véase por ejemplo el documento EP-A-194809). Si se desea, se pueden incluir en la construcción marcadores genéticos seleccionables, como los que transmiten fenotipos seleccionables tales como la resistencia a antibióticos o a herbicidas (por ejemplo, kanamicina, higromicina, fosfotricina, clorsulfurón, metotrexato, gentamicina, espectinomina, imidazolinonas y glifosato).

30 El ácido nucleico se puede introducir en las células vegetales mediante cualquier tecnología adecuada, tal como un vector plásmido Ti desmontado transportado por *Agrobacterium* que aprovecha su capacidad de tipo salvaje para transmitir genes (documentos EP-A-270355, EP-A-0116718, NAR 12(22) 8711 -87215 1984; el método de inmersión floral de Clough y Bent, 1998), bombardeo con partículas o microproyectiles (documentos US 5100792, EP-A-444882, EP-A-434616), microinyección (documentos WO 92/09696, WO 94/00583, EP 331083, EP 175966, Green et al., (1987) Plant Tissue and Cell Culture, Academic Press), electroporación (documentos EP 290395, WO 8706614 Gelvin Debeyser), otras formas de captación directa del ADN (documentos DE 4005152, WO 9012096, US 4684611), captación de ADN mediada por liposoma (por ejemplo, Freeman et al. Plant Cell Physiol. 29: 1353 (1984)), o el método de vortización (por ejemplo, Kindle, PNAS U. S.A. 87: 1228 (1990d)). Los métodos físicos para la transformación de células vegetales se han revisado en Oard, 1991, Biotech. Adv. 9: 1-11. Los plásmidos Ti, especialmente los vectores binarios, se describirán con más detalle en lo sucesivo.

35 La transformación mediante *Agrobacterium* es ampliamente utilizada por los expertos en la materia para transformar especies dicotiledóneas. Sin embargo, se ha conseguido un notable éxito en la producción rutinaria de plantas transgénicas fértiles en prácticamente todas las plantas monocotiledóneas económicamente relevantes (véase por ejemplo, Hiei et al. (1994) The Plant Journal 6, 271-282)). El microbombardeo de microproyectiles, la electroporación y la captación directa de ADN son preferidos cuando el *Agrobacterium* en solitario es ineficiente o ineficaz. Como alternativa, se puede emplear una combinación de técnicas diferentes para mejorar la eficacia del proceso de transformación, por ejemplo bombardeo de partículas recubiertas con *Agrobacterium* (documento EP-A-486234) o bombardeo con microproyectiles para inducir heridas seguido por cultivo simultáneo con *Agrobacterium* (documento EP-A-486233).

45 La selección particular de una tecnología de transformación se determinará por su eficacia para transformar determinadas especies vegetales así como la experiencia y la preferencia de la persona que realiza la invención con una metodología de elección concreta.

50 Será evidente para el experto en la materia que la elección concreta de un sistema de transformación para introducir el ácido nucleico en células vegetales no es esencial ni es una limitación de la invención, ni tampoco lo es la selección de la técnica para la regeneración de la planta. En los experimentos realizados por los inventores, el efecto de expresión mejorado se observa en varios diseños de integración del ADN-T.

55 Así, se describe en el presente documento un método para transformar una célula vegetal que implica la introducción de una construcción de la invención en un tejido vegetal (por ejemplo, una célula vegetal) y que ocasiona o permite la recombinación entre el vector y el genoma de la célula vegetal para introducir un ácido nucleico de acuerdo con la presente invención en el genoma. Esto se puede hacer así para realizar una expresión transitoria.

60

65

5 Como alternativa, después de la transformación del tejido vegetal, se puede regenerar una planta, por ejemplo, a partir de células individuales, tejido de callo o discos de hoja, como se conoce en la técnica. Casi cualquier planta se puede regenerar a partir de células, tejidos y órganos de la planta. Las técnicas disponibles se han revisado en Vasil et al., Cell Culture and Somatic Cell Genetics of Plants, Vol I, II y III, Laboratory Procedures and Their Applications, Academic Press, 1984, y en Weissbach y Weissbach, Methods for Plant Molecular Biology, Academic Press, 1989.

10 La generación de plantas transgénicas fértiles se ha conseguido en cereales tales como el arroz, maíz, trigo, avena, y cebada, y también en otras muchas especies de plantas (revisado en Shimamoto, K. (1994) Current Opinion in Biotechnology 5, 158-162.; Vasil, et al. (1992) Bio/Technology 10, 667-674; Vain et al., 1995, Biotechnology Advances 13 (4): 653-671; Vasil, 1996, Nature Biotechnology 14 página 702).

15 Las plantas regeneradas o partes de las mismas se pueden usar para proporcionar clones, semilla, progenie autóloga o híbrida y descendientes (por ejemplo, descendientes F1 y F2), cortes (por ejemplo, partes comestibles), guías de propagación, etc.

La invención proporciona además un organismo transgénico (por ejemplo, que se puede obtener por un método descrito en el presente documento) en el que se ha introducido un vector o casete de expresión, y en el que el gen heterólogo del casete se expresa a un nivel potenciado.

20 También se describe en el presente documento un método para generar la proteína de interés, donde el método comprende las etapas de realizar un método (o usar un organismo) como se ha descrito anteriormente., y opcionalmente recolectar, al menos, un tejido en el que se ha expresado la proteína de interés y aislar la proteína de interés del tejido.

25 Específicamente, la presente invención proporciona por tanto una planta o célula vegetal transgénica transfectada transitoriamente con un vector de expresión de la invención.

En un aspecto adicional, la presente invención proporciona también una planta o célula vegetal transgénica transformada de forma estable con un vector de expresión de la invención.

30 La divulgación también proporciona un propágulo de una planta para dichas plantas, que es cualquier parte que se puede usar en la reproducción o en la propagación, sexual o asexual, incluyendo cortes, semillas y así sucesivamente. También proporciona cualquier parte de estas plantas que incluye las células vegetales o el ADN heterólogo anteriormente descrito.

35 Así, en varios aspectos (y sin limitación), se describe en el presente documento:

40 • Ácidos nucleicos que consisten o consisten esencialmente de una secuencia potenciadora de la invención, donde la secuencia potenciadora puede (por ejemplo) consistir en los nucleótidos 1 a 512 del segmento del genoma de ARN-2 de CPMV, o derivarse del mismo, o de otro segmento del genoma de ARN-2 de un virus de ARN bipartido, en cada caso en donde el sitio de inicio diana correspondiente a la posición 161 del ARN-2 de CPMV se ha mutado).

45 • Los sistemas de expresión génica que comprenden dichas secuencias potenciadoras, por ejemplo en dirección 5' de un ORF que codifica una proteína de interés, o un polienlazador, y opcionalmente un terminador.

• Sistemas de expresión bipartida tal como se describe en el documento WO/2007/135480 modificados de acuerdo con la presente invención para usar las secuencias potenciadoras descritas en el presente documento.

50 • Casetes de expresión que comprenden: (i) un promotor, unido operativamente a (ii) una secuencia potenciadora como se ha descrito anteriormente (iii) un polienlazador o gen de interés que se desea expresar (iv) la UTR en 3' análoga (es decir, de la UTR en 3' del segmento del genoma de ARN-2 de CPMV), (v) una secuencia de terminación.

55 • Métodos de expresión de proteínas, por ejemplo, proteínas análogas, en organismos hospedadores tales como plantas usando sistemas de expresión génica o vectores de la invención.

60 • Las células hospedadoras y los organismos (por ejemplo, plantas o levaduras) que expresan proteínas a partir de los sistemas de expresión génica o vectores de la invención y métodos para producir los mismos.

"Gen" salvo que el contexto indique otra cosa se refiere a cualquier ácido nucleico que codifica información genética para su traducción en un péptido, polipéptido o proteína. De este modo, salvo que el contexto demande otra cosa, se puede usar de forma indistinta con "ORF".

65 Los genes que se puede desear expresar pueden ser transgenes o endógenos.

Los genes de interés incluyen los que codifican rasgos agronómicos, resistencia a insectos, resistencia a enfermedades, resistencia a herbicidas, esterilidad, características del grano, y similares. Los genes pueden estar implicados en el metabolismo del aceite, almidón, hidratos de carbono, nutrientes, etc. Dichos genes o rasgos de interés incluyen, pero sin limitación, rasgos relacionados con el medio ambiente o el estrés, rasgos relacionados con enfermedades, y rasgos que afectan al comportamiento agronómico. Las secuencias diana también incluyen genes responsables de la síntesis de proteínas, péptidos, ácidos grasos, lípidos, ceras, aceites, almidones, azúcares, hidratos de carbono, aromas, olores, toxinas, carotenoides, hormonas, polímeros, flavonoides, proteínas de almacenamiento, ácidos fenólicos, alcaloides, ligninas, taninos, celulosa, glucoproteínas, glucolípidos, etc.

Más preferentemente, los genes diana de monocotiledóneas y/o dicotiledóneas pueden incluir aquellas enzimas codificadoras responsables de la producción de aceite en plantas tales como colza, girasol, soja y maíz; las enzimas implicadas en la síntesis del almidón en plantas tales como patata, maíz, cereales; enzimas que sintetizan, o proteínas que son por sí mismas, medicamentos de tipo salvajeos tales como productos farmacéuticos o veterinarios.

Los ácidos nucleicos heterólogos pueden codificar, *inter alia*, genes de origen bacteriano, fúngico, vegetal o animal. Los polipéptidos se pueden utilizar *in planta* (para modificar las características de la planta por ejemplo, con respecto a la susceptibilidad a plagas, vigor, diferenciación de tejido, fertilidad, valor nutritivo, etc.) o la planta puede ser un intermedio para producir los polipéptidos que se pueden purificar a partir de los anteriores para su uso en otra parte. Dichas proteínas incluyen, pero no se limitan a la proteína del retinoblastoma, p53, angiostatina, y leptina. Del mismo modo, los métodos de la invención pueden usarse para producir proteínas reguladoras de mamífero. Otras secuencias de interés incluyen proteínas, hormonas, factores de crecimiento, citoquinas, albúmina de suero, hemoglobina, colágeno, etc.

Así, el gen o secuencia de nucleótidos diana codifica preferentemente una proteína de interés que es una proteína de resistencia a insectos; una proteína de resistencia a enfermedades; una proteína de resistencia a herbicidas; una proteína de mamífero.

"Vector" se define para incluir, *inter alia*, cualquier vector binario plásmido, cósmido, fago, vírico o de *Agrobacterium* en forma monocatenaria o bicatenaria lineal o circular que puede ser transmisible por sí mismo, o no o movilizable por sí mismo, o no, y que puede transformar un hospedador procarionta o eucariota bien mediante su integración en el genoma celular o existir extracromosómicamente (por ejemplo, plásmido de replicación autónoma con un origen de replicación). Las construcciones utilizadas serán total o parcialmente sintéticas. En particular, son recombinantes porque las secuencias de ácido nucleico no se encuentran juntas en la naturaleza (no se ejecutan de manera contigua), sino que se han unido o combinado artificialmente de otra forma. Salvo que se especifique de otra forma, un vector de acuerdo con la presente invención no tiene que incluir un promotor o cualquier otra secuencia reguladora, especialmente si el vector se va a usar para introducir el ácido nucleico en células para su recombinación en el genoma.

"Vector binario": como bien saben los expertos en la materia, un sistema vector binario incluye (a) secuencias frontera que permiten la transferencia de una secuencia de nucleótidos deseada al genoma de una célula vegetal; (b) la propia secuencia del nucleótido deseado, que por lo general comprende un casete de expresión de (i) un promotor activo en plantas, unido operativamente a (ii) la secuencia diana y/o potenciadora según sea adecuado. La secuencia de nucleótidos deseada se sitúa entre las secuencias frontera y se puede insertar en el genoma de una planta en las condiciones adecuadas. El sistema de vector binario necesitará por lo general otra secuencia (derivada de *A. tumefaciens*) para realizar la integración. Por lo general, esto se puede conseguir mediante la denominada "agroinfiltración" que utiliza una transformación transitoria mediada por *Agrobacterium*. En resumen, esta técnica se basa en la propiedad del *Agrobacterium tumefaciens* para transferir una parte de su ADN ("ADN-T") a una célula hospedadora donde puede quedar integrado en el ADN nuclear. El ADN-T se define por las secuencias frontera izquierda y derecha que tienen una longitud de aproximadamente 21-23 nucleótidos. La infiltración se puede conseguir por ejemplo, mediante jeringa (en hojas) o vacío (plantas completas). En la presente invención, las secuencias frontera están generalmente incluidas alrededor de la secuencia de nucleótidos deseada (el ADN-T) donde el uno o más vectores se introducen en el material vegetal mediante agroinfiltración.

"Casete de expresión" se refiere a una situación en la que un ácido nucleico está bajo el control de, y unido operativamente a, un promotor adecuado u otros elementos adecuados para su transcripción en una célula hospedadora tal como un microbio o célula vegetal.

Un "promotor" es una secuencia de nucleótidos a partir de la cual se puede iniciar la transcripción de un ADN unido operativamente posterior (es decir, en la dirección 3' de la hebra de sentido directo del ADN bicatenario).

"Unido operativamente" significa unido como parte de la misma molécula de ácido nucleico, colocada y orientada adecuadamente para iniciar la transcripción a partir del promotor.

Las especies "vegetales" de interés incluyen, pero sin limitación, maíz (*Zea mays*), *Brassica* sp. (por ejemplo, *B.*

napus, B. rapa, B. juncea), especialmente aquellas especies de Brassica útiles como fuente de semillas oleaginosas, alfalfa (*Medicago sativa*), arroz (*Oryza sativa*), centeno (*Secale cereale*), sorgo (*Sorghum bicolor*, *Sorghum vulgare*), mijo (por ejemplo, mijo perla (*Pennisetum glaucum*)), mijo común (*Panicum miliaceum*), moha (*Setaria italica*), mijo africano, (*Eleusine coracana*)), girasol (*Helianthus annuus*), cártamo (*Carthamus tinctorius*), trigo (*Triticum aestivum*), soja (*Glycine max*), tabaco (*Nicotiana tabacum*), *Nicotiana benthamiana*, patata (*Solanum tuberosum*), cacahuetes (*Arachis hypogaea*), algodón (*Gossypium barbadense*, *Gossypium hirsutum*), batata (*Ipomoea batatas*), mandioca (*Manihot esculenta*), café (*Coffea spp.*), coco (*Cocos nucifera*), piña (*Ananas comosus*), cítricos (*Citrus spp.*), cacao (*Theobroma cacao*), té (*Camellia sinensis*), banana (*Musa spp.*), aguacate (*Persea americana*), higo (*Ficus casica*), guayaba (*Psidium guajava*), mango (*Mangifera indica*), aceituna (*Olea europaea*), papaya (*Carica papaya*), anacardo (*Anacardium occidentale*), macadamia (*Macadamia integrifolia*), almendra (*Prunus amygdalus*), remolacha azucarera (*Beta vulgaris*), caña de azúcar (*Saccharum spp.*), avena, cebada, hortalizas, plantas ornamentales, y coníferas. La persona experta apreciará que el tropismo de los vectores víricos descritos en el presente documento varía. Sin embargo, la determinación de la susceptibilidad de dichos virus está bien comprendida entre las capacidades de la persona experta. Además, puede ser posible alterar dicha especificidad mediante la expresión recombinante de receptores que facilitan la entrada del virus en dicha célula vegetal.

La invención se describirá ahora adicionalmente con referencia a las siguientes Figuras y Ejemplos no limitantes. Los expertos en la materia inventarán otras realizaciones de la invención a la luz de estas.

20 Breve descripción de los dibujos

La **Figura 1** muestra un diagrama esquemático de los vectores de expresión de CPMV 00, 10, 01 y 11. En los vectores de expresión 00, los sitios de inicio en las posiciones 115 y 161 están intactas. En los vectores de expresión 10, los sitios de inicio en la posición 115 se han mutado, pero el sitio de inicio en la posición 161 está intacto. En los vectores de expresión 01, los sitios de inicio en la posición 161 se han mutado, pero el sitio de inicio en la posición 115 está intacto. En los vectores de expresión 11, los sitios de inicio en las posiciones 115 y 161 están ambos mutados. Los vectores de expresión de CPMV 00, 10, 01 y 11 incluyen un sitio de inicio en cualquiera de las posiciones 512 (FSC2-152), 513 (FSC2-513) o 514 (FSC2-514). Las barras se utilizan para indicar los sitios de inicio a partir de los cuales la expresión de la proteína sucede expresamente.

La **Figura 2** muestra el nivel de proteína fluorescente verde (GFP) soluble expresada en plantas transfectadas con los vectores de expresión de CPMV ilustrados esquemáticamente en la Figura 1. En los vectores de expresión FSC2-512, FSC2-513 y FSC2-514, el gen que codifica GFP se insertó después del codón de inicio en la posición 512, 513 y 514, respectivamente. Las hileras de los geles SDS-PAGE se marcaron como 00, 10, 01 y 11, dependiendo de si los sitios de inicio en el vector de CPMV, en las posiciones 115 y 161, estaban intactos o mutados. La hilera marcada '500ng' muestra la posición de una banda correspondiente a 500 ng de proteína GFP e indica de esta forma la posición esperada de la proteína GFP expresada a partir de los vectores de expresión de CPMV. La hilera izquierda de cada gel SDS-PAGE muestra la posición de los marcadores de tamaño de proteína.

La **Figura 3** muestra el nivel de expresión de GFP en hojas de *Nicotiana benthamiana* transfectadas con los mismos vectores de expresión de CPMV usados en el expresión ilustrado en la Figura 2. Las regiones pálidas en las puntas de las hojas corresponden a regiones de la expresión de GFP. También se indican las mutaciones realizadas para inactivar los sitios de inicio en las posiciones 115 y/o 161 de los vectores de expresión 10, 01 y 11.

La **Figura 4** muestra una comparación entre Del-ARN-2 (vector de expresión 00 [FSC2-512] en la Figura 1) y HT (vector de expresión 01 [FSC2-512] en la Figura 1) para la expresión transitoria de la proteína fluorescente verde (GFP), proteína fluorescente roja de *Discosoma* (DsRed), y el antígeno principal de la hepatitis B (HBcAg). delARN-2 o clones HT de cada proteína se infiltraron con el supresor del silenciamiento P19. (A) Tejido 7 días después de la infiltración con las construcciones de delARN-2 se vuelve necrótico cuando se expresan DsRed o HBcAg mientras que este no es el caso para la expresión impulsada por HT. De hecho, el tejido que expresa DsRed mediante HT parece visiblemente rojo en condiciones de luz diurna. (B) Análisis de la expresión de proteínas mediante SDS-PAGE teñidos con Coomassie. Las bandas prominentes correspondientes a la proteínas recombinantes que se indican se confirmaron mediante transferencia western. 1-marcador, 2-tejido no infiltrado, 3-construcción delARN-2, 4- construcción HT, 5-patrón comercial si estaba disponible. Los extractos crudos de aproximadamente 5 mg de tejido infiltrado se cargaron por hilera, así como 2 mg del patrón de GFP y 2 mg del patrón de HBcAg. En ese momento no estaba disponible un patrón para DsRed.

La **Figura 5** muestra una comparación entre Del-ARN-2 (vector de expresión 00 [FSC2-512] en la Figura 1) y HT (vector de expresión 01 [FSC2-512] en la Figura 1) para la expresión transitoria del anticuerpo 2G12 humano dirigido contra el virus de la inmunodeficiencia humana: La cadena pesada de la IgG bien estaba en su forma de tipo salvaje (HL) o bien retenida con ER (HEL) e infiltrada con la cadena ligera y P19. (A) La expresión de 2G12-HEL mediante del-ARN-2 conduce a la necrosis del tejido infiltrado, mientras que esto no sucede con la expresión de HT. (B) Análisis mediante SDS-PAGE de los extractos brutos de tejido infiltrado con las cadenas

pesadas del anticuerpo (delARN-2 o HT) más P19. Para cada muestra, el extracto bruto de 5 mg de tejido infiltrado se cargó junto con 1 mg de patrón de IgG humana. Se observa fácilmente una banda que corresponde a 2G12 tras la tinción de coomassie. (C) La acumulación del anticuerpo 2G12 después de 5 días se midió mediante captura de proteína A y espectroscopia de resonancia de plasmón superficial, y representa la concentración después de su extracción en 2 volúmenes de tampón (PBS, EDTA 5 mM). Por tanto, los autores derivan concentraciones en peso fresco próximas a 150 mg/kg (para HT HEL) sin ninguna optimización de la incubación o la extracción de las plantas. Se midieron tres muestras para cada tratamiento.

La **Figura 6** muestra una microfotografía electrónica de partículas de HBcAg ensambladas, que se expresaron usando el sistema de expresión HT (vector de expresión 01 [FSC2-512] en la Figura 1) descrito en el presente documento. Las partículas de HBcAg ensambladas aparecen como esferas huecas, de aproximadamente 30 nm de diámetro. La savia que contiene las partículas de HBcAg no se concentró antes de adquirir la micrografía electrónica, aunque se eliminaron las sales indeseadas. Por tanto, la microfotografía electrónica representa la concentración de partículas de HBcAg en la savia.

La **Figura 7** muestra el vector pM81-FSC1.

La **Figura 8** muestra el vector pM81-FSC2.

La **Figura 9** muestra una representación esquemática de la construcción de pEAQ. (A) Plásmido inicial basado en pBINPLUS con la secuencia extraña mostrada en gris. (B) Productos de la PCR que contienen los elementos esenciales del vector binario. (C) Plásmido intermedio después de la ligadura en tres partes de los productos de la PCR con la cola personalizada. (D) Plásmido filan después de la amplificación y ligadura posterior de dos fragmentos a partir del intermedio.

La **Figura 10** muestra una representación esquemática del ADN-T de los derivados principales de pEAQ. Los ADN-T incluyen cualquier o ambos casetes P19 y NPTII tal como se indica, dejando posibles sitios de clonación en los sitios de restricción como se indica.

La **Figura 11** muestra niveles de expresión de GFP generados por vectores pEAQ comparados con su plásmido precursor pBB-FSC2-512-HT. El tejido se analizó 6 días después de la infiltración con P19 y el vector indicado, salvo para pEAQexpress, que se infiltró solo para la DO del patrón, o con una dilución de dos veces. (A) Hojas visualizadas con luz UVV light, (B) SDS-PAGE teñido con coomassie al 12 %, y (C) análisis espectrofluorométrico.

La **Figura 12** muestra la capacidad de P19(R43W) para potenciar la expresión de GFP comparado con P19 de tipo salvaje. El tejido se analizó 6 días después de la infiltración con pEAQselectK a una dilución de dos veces (selK -P19), pEAQselectK y P19 (selK); pEAQspecialK (spK), pEAQspecialK a una dilución de dos veces (spK 1:2), pEAQspecialKm (spKm), pEAQspecialKm a una dilución de dos veces (spK 1:2), pEAQexpress (ex), y pEAQexpress a una dilución de dos veces (spK 1:2). (A) Hojas visualizadas bajo luz UV y (B) análisis espectrofluorométrico.

La **Figura 13** muestra la expresión de la IgG de tamaño completo, 2G12, con un único plásmido pEAQ. (A) Representación esquemática de los dos plásmidos derivados de pEAQexpress construidos para expresar 2G12. (B) Esquema de infiltración que indica diluciones y sus respectivas DO para cada combinación de plásmido, y la concentración de extractos de proteínas realizados tras las infiltraciones a cada DO (\pm SD). (C) Análisis por SDS-PAGE teñido con coomassie al 12 %-4 % y (D) detección inmunológica de 2G12 pesada (γ) y (E) detección inmunológica de la región 2G12 Fab (Fab) de la cadena 8 días después de la infiltración. M, marcador con tamaños indicados; C, extracto de control Pat. 2G12 producido por CHO. Para la tinción de coomassie, proteína procedente del equivalente de 3 mg de tejido infiltrado se cargaron en cada hilera con 1 μ g de CHO2G12 y, para transferencia western, el equivalente de 0,75 mg de tejido en cada hilera con 250 ng de CHO2G12. Se indican los productos de ensamblaje/degradación estimados.

La **Figura 14** muestra la clonación y expresión de GFP depEAQ-HT en variantes de tipo salvaje y marcadas con His. (A) Representación esquemática del pEAQ-HT ADN-T con detalle del polienlazador. (B) Análisis espectrofluorométrico de la expresión de GFP. spK = pEAQspecialK-GFP-HT, GFP, HisGFP, y GFPHis = clones pEAQ-HT. (C) Análisis de SDS-PAGE al 12 % y western de la expresión de GFP. C = extracto de control.

La **Figura 15** muestra una construcción de ácido nucleico de la presente invención que es adecuada para su uso en células de insecto como parte de un vector baculovirus.

Ejemplos

Ejemplo 1

1.1 MÉTODOS

Creación del vector FSC2 y sus derivados

5 Ser creó un vector de clonación útil para la expresión de las proteínas de un plásmido basado en pBinP-1-GFP (Canizares *et al.*, 2006) escindiendo la secuencia completa del ARN-2 flanqueado por el promotor 35S del virus del mosaico de la coliflor (CaMV) y el terminador de la nopalina sintasa (nos) de pBinP-S2NT (Liu y Lomonosoff, 2002) e insertando este en el plásmido de mutagénesis pM81W (Liu y Lomonosoff, 2006) como un fragmento *Ascl/Pacl*. El plásmido resultante, pM81W-S2NT, se sometió a un único ciclo de mutagénesis que introdujo simultáneamente
10 cuatro cambios (véase el método en Liu y Lomonosoff, 2006) para dar pM81B-S2NT-1. La mutagénesis eliminó dos sitios *BspHI* de la estructura del vector e introdujo un sitio *BspHI* (T/CATGA) alrededor de AUG 512 y un sitio *Stul* (AGG/CCT) después de UAA 3299, el codón de terminación para la poliproteína codificada por el ARN-2. Posteriormente, el fragmento *BamHI/Ascl* se escindió de pBinP-NS-1 (Liu *et al.*, 2005) y se ligó en pM81B-S2NT-1 que se dirigió de forma similar, dando como resultado pM81-FSC-1. Este vector permite que se escinda el ORF completo del ARN-2 ORF en dirección 3' de AUG 512 mediante digestión con *BspHI* y *Stul* y se sustituyó con cualquier secuencia en los extremos compatibles con *BspHI* y *Stul* (enromado). Es importante el uso del sitio *BspHI* ya que preserva el AUG en 512 y se usó este iniciador para impulsar la traducción del gen insertado. Para expresar el gen extraño en plantas, el plásmido derivado de pM81-FSC-1 se dirigió con *Ascl* y *Pacl* y el fragmento que contenía el casete de expresión que incluía las secuencias extrañas transferidas a pBINPLUS se dirigió de forma
20 similar y los plásmidos resultantes se transformaron finalmente en *A. tumefaciens*.

Para mejorar la facilidad de la clonación, ampliar la selección de enzimas de restricción aplicables, e investigar el efecto del marco de lectura sobre la expresión del gen extraño, se sustituyó el ORF completo del ARN-2 ORF con un polienlazador corto. Una combinación de inserción de oligonucleótidos y de mutagénesis dirigida al sitio dio como
25 resultado pM81-FSC-2, que permite la clonación con *NruI* (TCG/CGA) y tanto *XhoI* (C/TCGAG) como *Stul*. La adenina terminal del sitio *NruI* se encuentra en la posición 512 preservando por tanto el AUG que se encuentra aquí. Las modificaciones alteraron los nucleótidos inmediatamente en 5' del AUG en 512, sin embargo, se mantuvo un buen contexto. La clonación de GFP en pM81-FSC-2 de tal manera que se inició su traducción a partir de AUG en 512, 513, 514, o 515 proporcionó el pM81-FSC-1 derivado de las construcciones pM81-FSC2-512, pM81-FSC2-513,
30 pM81-FSC2-514, y pM81-FSC2-515. Estos plásmidos basados en pM81 son los vectores de clonación que contienen los casetes de expresión que se transfirieron a continuación en el vector binario para producir los vectores de expresión FSC2-512, FSC2-513 y FSC2-514 usados en los experimentos que se muestran en las Figuras 2 y 3. En la Tabla 3 se muestran las diferencias entre la secuencia del segmento de tipo salvaje del genoma de ARN-2 de CPMV y los vectores pM81-FSC1 y pM81-FSC2. Los nucleótidos alterados en los vectores comparados con la
35 secuencia CPMV de tipo salvaje se muestran en mayúscula.

La transformación transitoria mediada por *Agrobacteria* tras la movilización en pBINPLUS (como se ha reseñado anteriormente para pM81-FSC-1) mostró que se obtuvieron niveles menores de proteínas cuando no se mantuvo la continuidad del marco entre AUG 161 y AUG en la dirección 3'. Se observó una disminución significativa en la
40 cantidad de GFP traducido desde las posiciones +1 y +2 con respecto a los AUG 161 y 512, mientras que la traducción desde la posición +3 (esto es, desde 515 y posterior en el marco) fue tan eficaz como la traducción desde un AUG en 512. Para mostrar que esto no fue debido a contextos debilitados de los AUG en 513 y (en una extensión menor) 514, se creó FSC2-515+ para iniciar desde la posición +3 pero con el mismo pobre contexto que FSC2-513. La expresión a partir FSC2-515+ fue tan elevada como la conseguida a partir de FSC2-512 o 515, indicando que el contexto inferior no explica la reducción en la expresión de FSC2-513 y 514.
45

Dado que no se sabe que los mecanismos conocidos por los cuales la traducción puede escapar de la orden del primer AUG requieran la continuidad del marco, es intrigante que la traducción eficaz procedente de vector basado en ARN-2 eliminado dependa de la continuidad del marco entre AUG 161 y AUG en la dirección 3'. A fin de comprender, y esperemos que de superar este fenómeno, se crearon una serie de mutantes con modificaciones en la secuencia 5' de ARN-2. Se usaron parejas de complementos de oligonucleótidos (véase la Tabla 2) en la mutagénesis dirigida al emplazamiento de pM81-FSC2-512, 513, y 514. Las mutaciones eliminaron cualquiera de AUG 115 (el codón de inicio para el uORF), AUG 161 (sin cambiar la secuencia de aminoácidos del uORF), o ambos sitios de inicio en la dirección 5'. Se realizaron mutaciones dobles sometiendo a mutagenización los mutantes A115G con los oligos U162C (Tabla 2).
50
55

La expresión transitoria de estos transcritos mutantes se realizó como se ha descrito para las construcciones pM81-FSC-2 anteriores. El análisis de la expresión de GFP procedente de estos mutantes utilizando SDS-PAGE teñido con coomassie (Figura 2) o luz UV para visualizar las hojas completas (Figurea3) muestra un fuerte aumento de la expresión en ausencia del AUG en 161. Además, la eliminación de AUG 161 solo o de los AUG 115 y 161 alivia la dependencia de la continuidad en marco entre AUG 161 y AUG en la dirección 3'. Por el contrario, la eliminación de solo AUG 115 parece potenciar esta dependencia así como inhibir generalmente la traducción. En conclusión, el uORF parece funcionar para regular por defecto la traducción a partir de AUG 161, que es generalmente inhibidora y confiere dependencia de la continuidad del marco.
60
65

Imagen al microscopio electrónico de la savia que contiene partículas de HBcAg

La savia usada para la micrografía electrónica de las partículas de HBcAg ensambladas en la Figura 6 se preparó como sigue. El tejido de la hoja se extrajo en 2 volúmenes de tampón Tris/NaCl y se intercambió para TE sin concentración en una columna MWCO de 100 kDa. La concentración final de HBcAg fue aproximadamente de 0,2 mg/ml tal como se considera por comparación con el patrón en un gel SDS-PAGE teñido con coomassie.

1.2 RESULTADOS

(1) Efecto de alterar las fases relativas de los sitios de inicio en la posición 161 (AUG161) y 512 (AUG512).

Para conseguir esto, nucleótidos extra se insertaron inmediatamente en la dirección 5' de AUG512 (FSC2-512) para movilizar el AUG a la posición 513, 514 y 515 (FSC2-513, FSC2-514 y FSC2-515) (Figura 1). Colocar AUG512 fuera de fase con AUG161 (FSC2-513 y FSC2-514) proporciona menos expresión de GFP tal como se determinó mediante fluorescencia (Figura 3) y geles teñidos con Coomassie (Figura 2). Al restaurar la fase (FSC2-515) se devolvió la expresión a los niveles observados con la situación de tipo salvaje (FSC2-512). La conclusión es que cuando AUG161 está presente, el inicio en la dirección 3' de AUG es más eficaz cuando está en fase con AUG en la posición 161.

(2) La eliminación del sitio de inicio en la posición 115 (AUG115) se acopló con la alteración de las fases relativas de los sitios de inicio en la posición 161 (AUG161) y 512 (AUG512).

La eliminación de AUG115 tiene poco o ningún efecto cuando se impulsa la expresión de GFP a partir de AUG512 es decir, cuando este segundo AUG está en fase con AUG161 (véanse las bandas marcadas 10 en las Figuras 2 y 3). Sin embargo, la delección de AUG115 cuando el segundo AUG está fuera de fase con AUG161 (513, 514) no da como resultado, virtualmente, la expresión de GFP (véanse las bandas marcadas 10 en las Figuras 2 y 3). Conclusión: AUG115 está algo implicado en la capacidad de los ribosomas de derivar AUG161 y alcanzar AUG512. Sin embargo, esto requiere que el AUG en la dirección 3' esté en la fase correcta.

(3) Efecto de la eliminación del sitio de inicio en la posición 161 (AUG161)

El efecto de esta mutación es muy drástico, y los niveles de expresión de GFP alcanzan 20-30 veces la cantidad que se encuentra cuando AUG161 está presente (véanse las bandas marcadas 01 en las Figuras 2 y 3). Además, ya no parece que sea importante que AUG512 esté en fase, se piensa que en ausencia de AUG161, la idea de fase no significa mucho. Además de que la presencia o ausencia de AUG115 no marca diferencia (véanse las bandas marcadas 11 en las Figuras 2 y 3).

Cuando se usan construcciones de delARN-2 (vector de expresión 00 [FSC2-512] en la Figura 1) para la expresión de DsRed y HBcAg, en 5 días, los parches infiltrados tienen menos presión de turgencia y se vuelven cloróticos (pálidos). En 7 días, el tejido aparece gris y está completamente muerto. Cuando se usa HT (vector de expresión 01 [FSC2-512] en la Figura 1), el tejido sigue estando turgente después de 7 días y el único signo de estrés es una ligera clorosis del tejido que expresa HBcAg. Cuando la cadena pesada de la IgG 2G12 IgG se expresa mediante delARN-2 y se retienen en el ER, la clorosis es evidente después de 7 días, mientras que esto no se observó para HT. El nivel de necrosis observado en plantas cuando se usa HT para expresar una proteína heteróloga es de esta manera mucho menor, a pesar del mayor nivel de expresión de proteína heteróloga conseguido, que cuando se usa delARN-2 para expresar la proteína heteróloga.

DISCUSIÓN

Se pueden expresar niveles muy elevados de expresión de genes extraños a partir de las construcciones de delARN-2 eliminando AUG161. De hecho, usando GFP, los inventores estiman los niveles como un 25-30 % de la proteína soluble total (TSP) o aproximadamente 1 gramo expresado en proteínas por Kg de hojas. Este es un nivel muy elevado y la solución que los inventores usan es extremadamente sencilla. El hecho de que los inventores ya no necesiten preservar un marco de lectura significa que se pueden producir vectores respetuosos con el usuario con polienlazadores.

Ejemplo 2

2.1 ANTECEDENTES

Tal como se describe en el Ejemplo 1, para investigar las características necesarias para una expresión eficaz de la región 5' no traducida (UTR) de ARN-2 de CPMC, Los presentes inventores resolvieron el papel de dos codones AUG que se encuentran en la secuencia líder en 5' en la dirección 5' del sitio de inicio principal. Los inventores demostraron que la delección de un codón de inicio en marco (161) en la dirección 5' del sitio de inicio de traducción principal (512) condujo a un aumento masivo en la acumulación de proteína extraña.

Utilizando este sistema, los inventores han mostrado que 6 d después de la infiltración, numerosas proteínas no relacionadas, incluyendo una IgG de tamaño completo con una partícula de tipo virus con autoensamblaje, se expresaron en >10 % y 20 % de la proteína extraíble, respectivamente. Por tanto, este sistema proporciona un vehículo ideal para una expresión de alto nivel que no se basa en la replicación vírica de los transcritos.

5 Este nuevo sistema (como se ha ilustrado por el vector de expresión 01 [FSC-512] en la Figura 1) se ha denominado "CPMV-HT" para el sistema de expresión de la proteína con el virus del mosaico del tabaco hipertraducible.

El sistema HT-CPMV muestra aumentos drásticos en los niveles de proteínas y de esta manera, es un método excelente para la expresión rápida de alto nivel de proteínas extrañas en plantas.

10 Se ha desarrollado una matriz de crecimiento de vectores binarios para la transformación de plantas en los últimos 25 años (Hellens et al., 2000b; Veluthambi et al., 2003; Lee y Gelvin, 2008). El principal objetivo de estos desarrollos se ha centrado sobre todo en mejorar la integración estable, por ejemplo, ampliando la gama de hospedadores de Agrobacteria (Hiel et al., 1994), la creación de una serie de vectores que permite una elección de marcadores seleccionables, casetes de expresión y proteínas de fusión (ilustrados por la gama pCAMBIA de vectores binarios de código abierto; http://www.cambia.org/daisy/bioforge_legacy/3725.html), o desarrollando sistemas para minimizar la integración de ADN extraño y la transformación exenta de marcadores (por ejemplo pCLEAN; Thole et al., 2007).

20 Se han diseñado también vectores binarios para replicar un número bajo de copias para reducir la frecuencia de múltiples acontecimientos de integración del mismo, ya que esto conduce al silenciamiento del gen (Johansen y Carrington, 2001).

25 Sin embargo, para la expresión transitoria, no se requiere estrictamente asegurar una eficaz integración en el núcleo hospedador y la presencia de un marcador para la selección de plantas. Además, tras la agro-infiltración, cada célula se inundó con moléculas de ADN-T, que se piensa que son transcripcionalmente competentes en el núcleo incluso sin la integración del genoma (Janssen y Gardner, 1989; Narasimhulu et al., 1996). Esto sugiere que la expresión transitoria podría beneficiarse de un mayor número de copias de plásmidos binarios.

30 Otra área de mejora de vectores binarios ha sido la reducción en tamaño de la estructura del vector. Dos ejemplos destacados que siguen demostrando los beneficios de los plásmidos más pequeños son pPZP (Hajdukiewicz et al., 1994) y pGREEN (Helens et al., 2000a). Además de mejorar la eficacia de los procedimientos de clonación y transformación bacteriana, estos vectores han proporcionado plantillas para los sistemas de expresión que se basan en múltiples casetes en un único ADN-T (Tzira et al., 2005; Thole et al., 2007). El presente ejemplo describe refinados no obvios de este vector que facilitan su uso práctico permitiendo que se realice la clonación en una única etapa, en lugar de la necesaria subclonación de los casetes de expresión entre el vector de clonación (por ejemplo, pM81-FSC2) y los sistemas de expresión (por ejemplo, PBIN-PLUS). Más específicamente, los resultados en el presente documento mostraron que era posible reducir drásticamente el tamaño de pBINPLUS sin comprometer el rendimiento en términos de replicación y transferencia del ADNT. Además, los elementos del sistema CPMV-HT se han incorporado al vector resultante de forma modular de forma que se pueden expresar múltiples proteínas a partir de un único ADN-T. Estas mejoras han conducido a la creación de vectores de expresión de vectores versátiles de alto nivel de expresión que permiten una clonación eficaz directa de genes extraños.

2.2 MATERIALES Y MÉTODOS

45 pBD-FSC2-512-U162C (HT) contiene el casete FSC2-512-U162C (véase el Ejemplo 1) insertado en los sitios PacI/AscI de pBINPLUS (van Engelen et al., 1995). Los segmentos esenciales de este plásmido (véase a continuación) se amplificaron con la polimerasa de alta fidelidad PHUSION (New England Biolabs) utilizando oligonucleótidos que codifican sitios de enzimas de restricción únicos para la religadura (Tabla 4.1). La región del ADN-T se amplificó con un cebador de sentido directo homólogo para la secuencia en dirección 5' de un único sitio AhdI (pBD-LB-F) y un cebador de sentido contrario que incluyó un sitio Apal (pBD-RB-Apal-R). Una región que incluía el origen de replicación ColEI, el gen NPTIII, y el locus TrfA se amplificó con un cebador de sentido directo que incluía un sitio Apal (pBD-ColEI-Apal-F), y un cebador de sentido contrario que incluía un sitio SpeI (pBD-TrfA-SpeI-R). El origen de replicación RK2 (OriV) se amplificó con un cebador de sentido directo que incluyó un sitio SpeI (pBD-oriVSpeI-F) y un cebador de sentido contrario que incluía un sitio AhdI (pBD-oriV-AhdI-R). Tras la purificación, los productos se digirieron de acuerdo con sitios de restricción únicos codificados en sus términos y se mezclaron en tres puntos de ligadura. Esto dio como resultado el plásmido pEAQbeta, para el cual se verificaron las uniones por ligadura mediante secuenciación. Se detectó una delección de aproximadamente 1,2 kb procedente del ADN-T que había eliminado una porción del terminador nos del casete CPMV-GFP-HT. Por tanto, una porción del terminador que incluye la frontera derecha de pBD-FSC2-GFP-HT se reamplificó con los cebadores pMini>pMicroBIN-F2 y pBD-RB-Apal-R, así como la estructura de pEAQbeta, que incluía la frontera derecha, usando los cebadores pBD-ColEI-Apal-F y pMini>pMicroBIN-R (Tabla 4.1). Los productos purificados se digirieron con Apal y FseI y se ligaron para dar pEAQ (Figura 9).

65 El gen P19 flanqueado por el promotor 35S y el terminador 35S se amplificó a partir de pBIN61-P19 (Voinnet et al., 2003) usando cualquiera de 35SP19-PacI-F y 35SP19-AscI-R, o 35SP19-FseI-F y 35S-P19-FseI-R como cebadores (Tabla 4.1). El gen NPTII flanqueado por el promotor y el terminador de nos se amplificó a partir de pBD-FSC2-

GFPHT usando los cebadores pBD-NPTII-Fsel-F y pBD-NPTII-Fsel-R (Tabla 4.1). Tras la prolongación de A, se ligaron los casetes amplificados en pGEM-T easy (Promega). El casete P19 se escindió de pGEM-T easy con Fsel y se ligó en pEAQ-GFP-HT digerido con Fsel para dar pEAQexpress-GFP-HT. El casete NPTII escindido con Fsel se ligó en pEAQ-GFP-HT digerido con Fsel en ambas direcciones para dar pEAQselectK-GFP-HT y pEAQselectK(rev)-GFP-HT. Se escindió también el casete NPTII con PacI/Ascl y se ligó en los sitios AsiSI/Mlul de pEAQselectK-GFP-HT para dar pEAQspecialK-GFP-HT. El P19 en pGEM-T se sometió a mutagénesis dirigida al sitio mediante el método QUICKCHANGE (Stratagene) para efectuar la conversión de la Arginina43 a un resto triptófano usando los cebadores P19-R43W-F y P19-R43W-R. El casete P19 mutante se liberó con la digestión de PacI/Ascl y se insertó en los sitios AsiSI/Mlul de pEAQselectK-GFP-HT para dar pEAQspecialKm-GFP-HT.

Los oligonucleótidos que codifican las hebras de sentido directo y de sentido contrario de un polienlazador corto (Tabla 4.1) se hibridaron dejando la mitad de un sitio Nrul en la dirección 3' en el extremo 5' y un saliente que se corresponde con el de XhoI en el extremo 3'. Los oligos hibridados se ligaron con pM81-FSC2-A115G-U162C digerido con Nrul/XhoI (véase anteriormente) para dar pM81-FSC2-POW. Se eliminó el sitio Nrul del casete P19 en pGEM-T mediante mutagénesis dirigida al sitio (QUICKCHANGE; Stratagene) con los cebadores P19-ΔNrul-F y P19-ΔNrul-R, y se reinsertó en los sitios AsiSI/Mlul de pEAQselectK-GFP-HT para dar pEAQspecialKΔNrul-GFP-HT que no mostró reducción en la expresión en comparación con pEAQspecialK-GFP-HT (no se muestran los datos). A continuación se liberó el fragmento PacI/Ascl de pM81-FSC2-POW y se insertó en pEAQspecialKΔNrul-GFP-HT de forma similar sustituyendo por tanto el casete de expresión GFP HT y dando como resultado pEAQ-HT. GFP se amplificó a partir de pBD-FSC2-GFP-HT con un conjunto de cuatro cebadores (Tabla 4.1) en tres combinaciones para la inserción en pEAQ-HT: GFP-Agel-F y GFP-XhoI-R; GFP-Agel-F y GFP-XmaI-R; y GFP-XmaI-F y GFP-XhoI-R. Los productos de la PCR se digirieron con las enzimas especificadas en sus cebadores y se insertaron en pEAQ-HT adecuadamente digerido para dar pEAQ-HT-GFP, pEAQ-HT-GFPHis, y pEAQ-HT-HisGFP.

Tabla 4.1. Se muestran los oligonucleótidos usados para la amplificación y la mutagénesis. los sitios de enzimas de restricción, o partes de los mismos, en el recuadro inferior, y se subrayan las mutaciones en negrita.

Oligo	Secuencia	Función
pBD-LB-F	GCCACTCAGCTTCCTCAGC GGCTTT	cebador de sentido directo para la amplificación de la región 6338-12085 de pBD-FSC2-GFP-HT
pBD-RB-Apal-R	TATTAgggcccCCGGCGCCAG ATCTGGGGAACCCTGTGG GACTTAgggcccGTCCATTTCCGCGCAGACGATGACGTCA	Cebador de sentido contrario para la amplificación de la región 6338-12085 de pBD-FSC2-GFP-HT con un sitio Apal
pBD-ColEI-Apal-F	CT GCATTAaactagtCGCTGGCTG	Cebador de sentido directo para la amplificación de la región 1704-5155 de pBD-FSC2-GFP-HT con un sitio Apal
pBD-TrfA-SpeI-R	CTGAACCCCCAGCCGGAAC TGACC	Cebador de sentido contrario para la amplificación de la región 1704-5155 de pBD-FSC2-GFP-HT con un sitio SpeI
pBD-oriV-SpeI-F	GTAGCactagtGTACATCACC GACGAGCAAGGC	Cebador de sentido directo para la amplificación de la región 14373-670 de pBD-FSC2-GFP-HT con un sitio SpeI
pBD-oriV-AhdI-R	CAGTAgacagggtgtcTCGCGG CCGAGGGGCGCAGCCC ggccggccacgcgtTATCTGCAG	Cebador de sentido contrario para la amplificación de la región 14373-670 de pBD-FSC2-GFP-HT con un sitio AhdI
pMini>pMicr oBIN-F2	AgcgatcgcGAATTGTGAGCG GATAACAATTTACACAGGA AACAGCTATGACC	cebador de sentido directo para la amplificación de la región 2969-85 de pEAQbeta con sitios FesI-Mlul-AsiSI

pMini>pMicroBIN-R	gcgatcgcTCTGCAGATAacgcg tggccggccCTCACTGGTAAA AGAAAAACCACCCAGTAC ATTAAAAACGTCC	cebador de sentido contrario para la amplificación de la región 969-85 de pEAQbeta con sitios AsiSI-MluI-FseI
35SP19-PacI-F	ttaattaaGAATTCGAGCTCGG TACCCCTACTCC	cebador de sentido directo para la amplificación del casete 35S-P19 con el sitio PacI
35SP19-AscI-R	ggcgcgccATCTTTTATCTTTA GAGTTAAGAACTCTTTG	Cebador de sentido contrario para la amplificación del casete 35S-P19 con el sitio AscI
35SP19-FseI-F	ggccggccGAATTCGAGCTCG GTACCCCC	El cebador de sentido directo para la amplificación del casete 35S-P19 con el sitio FseI
35SP19-FseI-R	ggccggccATCTTTTATCTTTA GAGTTAAG	Cebador de sentido contrario para la amplificación del casete 35S-P19 con un sitio FseI
pBD-NPTII-FseI-F	ggccggccTACAGTATGAGCG GAGAATTAAGGGAGTCACG ggccggccTACAGTCCCGATC	Cebador de sentido directo para la amplificación del casete NPTII de pBD-FSC2-GFP-HT con el sitio FseI
pBD-NPTII-FseI-R	TAGTAACATAGATGACACC GCGC	Cebador de sentido contrario para la amplificación del casete NPTII a partir de pBD-FSC2-GFP-HT con el sitio FseI
P19-R43W-F	CGAGTTGGACTGAGTGGTG GCTACATAACGATGAG CTCATCGTTATGTAGCCACC	Cebador de sentido directo para la mutagénesis de la arginina
P19-R43W-R	ACTCAGTCCAACCTCG CCGTTTCTGGAGGGTCTCG	cebador de sentido contrario para la mutagénesis de la arginina 43 de P19 a un resto triptófano
P19-ΔNrul-F	AACTCTTCAGCATC GATGCTGAAGAGTTCGAGA	Cebador de sentido directo para la mutagénesis silenciosa del sitio de restricción Nrul en P19
P19-ANrul-R	CCCTCCAGAAACGG cgaccggtATGCATCACCATCA	Cebador de sentido contrario para la mutagénesis silenciosa del sitio de restricción Nrul en P19
POW-F	CCATCATcccgggCATCACCA TCACCATCACTAGc tcgagCTAGTGATGGTGATGG	oligo de sentido directo para un polienlazador, POW
POW-R	TGATGcccgggATGATGGTGA TGGTGATGCATaccggttcg	oligo de sentido directo para un polienlazador, POW
GFP-AgeI-F	atcggaccggtatgactagcaaaggag aagaac	Oligo de sentido directo para la amplificación de GFP con un sitio AgeI
GFP-XmaI-F	atccgaccgggactagcaaaggaga agaacttttcac	Oligo de sentido directo para la amplificación GFP con un sitio XmaI y sin codón de inicio

GFP-Xmal-R	atccgacccgggttgtatagttcatccat gcc	Oligo de sentido contrario para la amplificación de GFP con un sitio Xmal y sin codón de terminación
GFP-XhoI-R	cgatcctcgagttattgtatagttcatcca tgcc	Oligo de sentido contrario para la amplificación de GFP con un sitio XhoI

2.3.1 pBINPLUS contiene al menos 7,4kb de secuencia extraña

5 La expresión de CPMV-HT permite la producción de niveles extremadamente elevados de proteínas recombinantes. Sin embargo se ha deseado mejorar adicionalmente el sistema y su uso para una transformación transitoria.

10 El primer área de mejora se refiere al hecho de que los plásmidos pequeños son más eficaces que los más grandes en reacciones de ligadura y procedimientos de transformación bacteriana. Las comparaciones con las estructuras de vectores binarios más pequeños indicó que pBINPLUS contiene igualmente cantidades significativas de secuencias extrañas. Se ha determinado que son esenciales cuatro elementos de pBINPLUS para una adecuada función como un vector binario: el ADN-T, el RK2 (OriV) es un origen de replicación de una amplia gama de hospedadores, el gen NPTIII que confiere resistencia a la kanamicina (Trieu-Cuot y Courvalin, 1983), y TrfA de RK2 que promueve la replicación (Figura 9). El análisis bioinformático de las secciones restantes de la estructura las muestra como artefactos de la construcción de pBIN19, que se basan en la presencia de sitios de restricción adecuados en los plásmidos progenitores (Bevan, 1984). Se confirmaron estas observaciones por un informe sobre la secuenciación completa de pBIN19 (Frisch et al., 1995). pBINPLUS incluye el origen de replicación ColEI no esencial para un número mayor de copias en E. coli. Se pueden encontrar aproximadamente 2,6kb de ADN superfluo en el ADN-T. Esto incluye el marcador seleccionable NPTII para la transformación de plantas que no se requiere para la expresión transitoria. En su conjunto, la cantidad total de secuencia extraña con pBINPLUS parece estar en exceso de 7,2kb.

20

2.3.2 Construcción de la serie pEAQ

25 Para vigilar los efectos sobre la expresión resultantes de las modificaciones del vector, los inventores escogieron comenzar con el plásmido derivado de pBINPLUS, pBD-FSC2-512-U162C(HT). Tres regiones, consistentes en el ADN-T, el origen de replicación RK2 (OriV), y un segmento que contiene el origen ColEI, NPTIII, y TrfA, se amplificaron mediante la PCR a partir de pBD-FSC2-GFP-HT. La ligadura de estos tres fragmentos dio como resultado el plásmido pEAQbeta (Figura 9), que tiene 4584 pb más pequeños que su plásmido progenitor. Un ciclo adicional de amplificación mediante la PCR de pEAQbeta eliminó 2639 pb de secuencias no esenciales de la región del ADN-T región e insertó tres únicos sitios de restricción, AsiSI, MluI, y FseI. La digestión de AsiSI/MluI es compatible con la inserción de fragmentos de PacI/Ascl, y es por tanto, extremadamente útil para los casetes de clonación múltiple procedentes de todos los vectores de clonación de CPMV anteriores. FseI proporciona un único sitio de reconocimiento de 8 bases para intercambiar diferentes marcadores de selección o silenciar los casetes supresores. El plásmido pEAQGFP-HT resultante es menor que la mitad del tamaño de pBINPLUS y sin el casete de expresión de CPMVHT sería solo de 5137 pb, haciendo de este uno de los vectores binarios conocidos más pequeños (Figura 9). Se secuenció el plásmido pEAQ completo y se descubrió que el origen de la replicación de RK2 estaba en orientación contraria a la del anteriormente notificado (Frisch et al., 1995) y está indicado por tanto en la orientación correcta en pEAQ-GFP-HT.

35

40 Se usó pEAQ-GFP-HT como punto de partida para la inclusión de diversas características adicionales en el ADN-T (Figura 10). El casete NPTII de pBINPLUS se reinsertó en el sitio FseI de pEAQ en las orientaciones directa e inversa con respecto al casete GFP-HT para dar pEAQselectK-GFP-HT y pEAQselectK(rev)-GFP-HT.

40

45 El casete 35S-P19 se insertó en el sitio FseI para dar pEAQexpress-GFP-HT. Finalmente, el casete 35S-P19 se insertó en los sitios MluI/AsiSI de pEAQselectK-GFP-HT para dar pEAQspecialK-GFP-HT. Por tanto, se construyeron una series de pequeños vectores binarios para una expresión transitoria fácil y rápida.

45

2.3.3 La reducción del tamaño no compromete la expresión transitoria de pEAQ

50 La agro-infiltración de la serie pEAQ de vectores muestra que la gran reducción en el tamaño no compromete significativamente los niveles de expresión en los ensayos transitorios. La infiltración simultánea de pEAQ-GFP-HT, y pEAQselectK(rev)-GFP-HT con P19 proporcionada por pBIN61-P19, dio como resultado niveles de expresión no significativamente diferentes a la infiltración simultánea de pBD-FSC2-512-HT y P19. Esto se puede observar con iluminación UV (Figura 11A), SDS-PAGE (Figure 11B), y medidas de espectrofluorescencia de GFP en extractos de proteínas (Figura 11C). De manera interesante, la orientación del casete NPTII en el ADN-T parece afectar el nivel de expresión. pEAQselectK muestra una marcada mejora en comparación con pEAQselectK(rev) idéntico de otra manera, lo que da como resultado una reducción en la acumulación de GFP.

55

Teóricamente, la incorporación de un supresor del casete de silenciamiento en pEAQ no debería afectar su capacidad para mejorar el nivel de expresión transitoria de un gen extraño que se va a expresar desde el mismo ADN-T. De hecho, la infiltración de pEAQexpress-GFP-HT en solitario da como resultado también niveles de expresión similares a, o mejores que, pBD-FSC2-GFP-HT (Figure 7.3). Además, para ensayar la eficacia de pEAQexpress, el cultivo de *Agrobacterium* se diluyó dos veces, de tal manera que la densidad óptica final (DO) fue la de cada cultivo individual de las infiltraciones simultáneas.

Como era de esperar, esto dio como resultado niveles de expresión similarmente elevados y demuestra que la incorporación que incorporaban ambos genes de interés y el supresor del silenciamiento en el mismo ADN-T permite el uso de la mitad de la cantidad de *Agrobacterias* (Figura 11). Por tanto, se puede usar CPMV-HT para expresar elevados niveles de la proteína extraña cuando todos los componentes están presentes en el mismo ADN-T.

2.3.4 El mutante P19 puede suprimir el silenciamiento de un transgén en un ensayo transitorio

Para aprovechar el aumento en la expresión que da como resultado la orientación directa del casete NPTII en el ADN-T, se insertó el casete P19 entre los sitios AsiSI y MluI en pEAQselectK-GFP-HT para dar pEAQspecialK-GFP-HT (Figure 10). La presencia de P19 en el mismo ADN-T como la secuencia de GFP da como resultado similares niveles de expresión de pEAQselectK-GFPHT infiltrado simultáneamente con P19 (Figura 12). Esto es más que la expresión generada por pEAQexpress-GFP-HT, y parece deberse a la presencia del casete NPTII (Figura 12). Por otra parte, la expresión inferior de pEAQexpress podría deberse a la posición y orientación diferentes del casete P19 en el ADN-T. Sin embargo, como con los vectores pEAQexpress, los vectores pEAQspecialK proporcionan un nivel de expresión elevado con suspensiones de *Agrobacteria* a la mitad de la DO final de la usada cuando deben infiltrarse simultáneamente dos cultivos.

Combinar el casete de expresión génica extraño con un casete P19 y un marcador seleccionable hace posible ensayar el rendimiento de CPMV-HT en plantas transgénicas. Sin embargo, la expresión constitutiva de los supresores del silenciamiento de tipo P19 puede dar como resultado fenotipos intensos debido a su interferencia con el silenciamiento de genes endógenos asociados con los procesos de desarrollo (Silhavy y Burgyan, 2004). Se ha propuesto que una mutación recientemente caracterizada de P19 (R43W) tiene una actividad reducida hacia el silenciamiento del gen endógeno y por tanto puede ser un candidato mejor para la supresión de el silenciamiento del transgén en transformantes estables (Scholthof, 2007). Para investigar la factibilidad de una transformación estable con el sistema CPMV-HT, ambos P19 de tipo salvaje y mutante se insertaron en el ADN-T de pEAQselectK-GFP-HT para evaluar transitoriamente las variantes. Como se muestra mediante la iluminación UV de las hojas infiltradas, SDS-PAGE de los extractos de proteínas, y las medidas espectrofluorométricas de los niveles de GFP, el mutante P19 en pEAQspecialKm es aproximadamente la mitad de eficaz en la mejora de la expresión del gen extraño que el P19 de tipo salvaje en pEAQspecialK (Figura 12). Esto representa el primer estudio sobre el efecto de la mutación R43W en P19 sobre la capacidad de suprimir el silenciamiento de un transgén.

Ejemplo 3

Expresión de IgG de nivel elevado a partir de un único plásmido

a fin de aprovechar la naturaleza modular de la serie pEAQ, casetes de expresión de CPMV-HT que contenían la cadena pesada retenida por ER (HE) y la cadena ligera (L) de la IgG dirigida contra VIH humano, 2G12 en los sitios PacI/Ascl y AsiSI/MluI de pEAQexpress. Para determinar si el sitio de inserción influencia los niveles de expresión, las cadenas L y HE se insertaron en ambas posiciones dando como resultado pEAQex-2G12HEL y pEAQex-2G12LHE (Figura 13A). la infiltración de *N. benthamiana* con cultivos individuales de *Agrobacterium* que contenían los plásmidos anteriores dio como resultado la formación de anticuerpos 2G12 ensamblados completamente idénticos en tamaño a los 2G12 producidos mediante mezcla de tres cultivos de *Agrobacterium* que expresaron cada uno los componentes individuales, L, HE y P19 (Figura 13C). la proteína cargada en cada banda representa 1/30 del extracto obtenido a partir de 90 mg de tejido infiltrado o 1/333 de la proteína potencialmente obtenible a partir de 1 g de tejido. La cantidad máxima de la IgG ensamblada producida a partir de la mezcla de las tres cepas corresponde a 1 mg de 2G12 producido por CHO en el gel SDS-PAGE no reducido teñido con coomassie. Esto sugiere un nivel de expresión de 2G12 superior a 325 mg/kg de peso de tejido fresco, lo que concuerda con las concentraciones medidas con SPR. El uso de pEAQex-2G12HEL parece sobrepasar este nivel ya alto de acumulación de anticuerpo.

Una ventaja de los vectores derivados de pEAQ es que cada componente de una proteína multicadena tal como una IgG se puede administrar automáticamente a una célula infectada. Por tanto, deberían mantenerse niveles de expresión elevados a diluciones más altas de las suspensiones de *Agrobacteria* que los cultivos múltiples que se han utilizado. Para ensayar esto en este caso de forma práctica, los cultivos que se resuspendieron inicialmente a DO 1,2, y se mezclaron cuando fue necesario, se sometieron a diluciones en serie de dos veces (Figura 13B). Esto dio como resultado DO finales para cada cultivo individual de la mezcla de tres cultivos de 0,4-0,13, y 0,04. Los cultivos individuales que incluyen construcciones pEAQexpress se infiltraron a DO de 1,2, 0,4, y 0,13. Cuando se utilizaron tres cultivos independientes, el nivel de 2G12 ensamblado disminuyó intensamente durante la dilución en serie. Por el contrario, la expresión de 2G12 derivada de pEAQex-2G12HEL y pEAQex-2G12LHE, se mantuvo en un nivel consistentemente elevado, siendo muy modesta cualquier reducción en la dilución (Figura 13C-E). La falta de

sensibilidad debida a la dilución confirma la eficacia mejorada conseguir al colocar los tres casetes de expresión en el mismo ADN-T. De manera interesante, la cantidad de proteína total extraída del tejido infiltrado fue casi la mitad cuando la DO del infiltrado se redujo de 1,2 a 0,4. Esto sugiere que una fracción significativa de la proteína en los extractos tisulares a los que se había infiltrado la suspensión de mayor DO puede consistir en proteínas derivadas de Agrobacteria o proteínas vegetales producidas en respuesta a las concentraciones más elevadas de Agrobacteria.

La inspección de la Figura 13C sugiere que la posición relativa de un casete dentro del ADN-T puede afectar a los niveles de expresión. La expresión global procedente de pEAQex-2G12LHE fue ligeramente inferior que para pEAQex-2G12HEL. Esto se confirmó mediante transferencia western de las muestras no reducidas, lo que también indicó algunas diferencias en la abundancia de los productos de degradación y de las cadenas de inmunoglobulina no incorporadas (Figura 13C - E). El tejido infiltrado con pEAQex-2G12LHE parece carecer de un producto de degradación específico de la cadena pesada de aproximadamente 70 a 80 kDa (Figura 13D). Asimismo, parece haber mucho menos del intermedio ensamblado HL2, así como más cadena ligera libre (Figura 13E). Como se sabe que la cadena pesada limita el ensamblaje de 2G12 en plantas (Markus Sack, pers. comm., RWTH, Aachen, Alemania), lo que se confirma por la falta de cadena pesada libre discernible en todas las muestras, estos resultados indican que pEAQex-2G12LHE produce menos cadena pesada que pEAQex-2G12HEL. Esto podría deberse a una expresión reducida desde el casete CPMV-HT cercano a la frontera izquierda del ADN-T.

En otros experimentos (no se muestran los datos), el sistema CPMV-HT también se usó con éxito en el formato transitorio en *N. benthamiana* para expresar:

- el virus de la lengua azul (serotipo 10) VP2, VP3, VP5, VP7 y NS1.
- el rotavirus NSP5.
- la calmodulina de *Medicago truncatula* (que después se purificó).
- El ectodominio difícil de expresar del receptor Fc gamma receptor 1 humano (CD64) que se había purificado y demostrado funcional en estudios de unión de anticuerpo.
- Las proteínas de la cubierta de CPMV pequeña (S) y grande (L) se expresaron simultáneamente y se demostró que se ensamblaban en partículas análogas a virus (no se muestran los datos)

Ejemplo 4

Clonación directa en un vector de expresión CPMV-HT

A pesar de la combinación de elementos del sistema en un único plásmido, los vectores descritos anteriormente en el presente documento siguen necesitando un procedimiento de clonación en dos etapas para introducir una secuencia a expresar en el plásmido binario. El presente ejemplo proporciona un plásmido binario en el que se puede introducir directamente un gen de interés. El plásmido incorpora un polienlazador que no solamente permite la inserción directa en el plásmido basado en pEAQ, también permite la fusión a una etiqueta de histidina del extremo C o N si se desea (pEAQ-HT; Figura 14A). El polienlazador se insertó primero como oligonucleótidos hibridados en el pM81-FSC2-512(A115G)(U162C) proporcionando pM81-FSC-POW. Esta construcción puede seguir usándose en el procedimiento convencional de clonación en dos etapas para la generación de construcciones basadas en pEAQ para la expresión de múltiples polipéptidos. Además, el uso de un líder en 5' doblemente mutado puede permitir niveles de expresión incluso más elevados de lo posible con la mutación simple. El casete CPMV-HT se transfirió después a pEAQspecialK mediante los sitios PacI/AscI para proporcionar pEAQ-HT. La inserción de GFP en las tres posiciones del polienlazador de pEAQ-HT dio como resultado un GFP no marcado, y fusiones de la etiqueta His en 5' (HisGFP) y 3' (GFPHis).

Como era de esperar, el GFP no marcado se expresó en un nivel incluso mayor del obtenido con pEAQspecialK-GFP-HT y un exceso de 1,6 g/kg de tejido FW (Figura 14B). Esto se debe probablemente al hecho de que el líder CPMV 5' de pEAQ-HT incluye la mutación extra que elimina AUG 115 que, cuando se elimina junto con AUG 161, potencia adicionalmente la expresión.

La presencia de esta etiqueta His detectada mediante transferencia western confirmó la fusión correcta en ambos extremos N y C de los residuos de aminoácidos codificados mediante el polienlazador. Los tres variantes de GFP fueron detectables, con anticuerpos contra GFP, mientras que solamente HisGFP y GFPHis fueron detectables con anticuerpos contra His (Figura 14C), y la presencia de la etiqueta His reduce la movilidad de la banda GFP en SDS-PAGE en la cantidad esperada. La etiqueta también redujo la cantidad de GFP detectada mediante análisis de fluorescencia (Figura 14B). Este efecto fue más intenso para la etiqueta His del extremo N. La intensidad de las bandas teñidas con coomassie sugiere que esto representa una reducción en la acumulación de GFP marcado (Figura 14C), más que la interferencia con las propiedades fluorógenas de GFP. Sin embargo, los niveles de las proteínas etiquetadas con His siguieron siendo muy elevados superando 0,6 y 1,0 g de GFP por kg de tejido FW.

Discusión de los ejemplos 2-4

Para mejorar la facilidad de uso y el rendimiento del sistema de expresión CPMV-HT, se ha creado un conjunto modular de vectores para conseguir una expresión fácil y rápida en plantas.

La eliminación de más de la mitad de la estructura principal del plásmido del vector binario, pBINPLUS, y parte de la región ADN-T no esencial para la expresión transitoria dio como resultado uno de los plásmidos Ti binarios más pequeños conocidos sin comprometer los niveles de expresión.

Una proporción similar de la estructura principal ya se había retirado de pBIN19 sin pérdida de rendimiento (Xiang et al., 1999). Sin embargo, pBINPLUS tiene dos mejoras significativas respecto de pBIN19 (van Engelen et al., 1995); un mayor número de copias en *E. coli* debido a la adición del origen de replicación ColEI y un ADN-T reorientado que garantiza que el gen de interés está más alejado de la frontera izquierda para evitar que experimente demasiadas deleciones en la planta (Rossi et al., 1996). Aunque el menor tamaño de los plásmidos pEAQ no tiene efectos notables sobre el número de copias, proporcionan rendimientos muy mejorados durante los procedimientos de clonación usando kits comerciales de extracción de plásmidos ya que estos son más eficaces para plásmidos que tengan menos de 10 kb (no se muestran los datos).

La naturaleza modular del vector binario pEAQ añade funcionalidad a la expresión de CPMV-HT al permitir que cualquier gen supresor del silenciamiento y/o marcador, si es necesario, se exprese simultáneamente con uno dos casetes de CPMV-HT. Por ejemplo, la inserción de un segundo casete de HT que incluye una secuencia heteróloga en los sitios AsiSI/MluI de pEAQexpress-GFP-HT permitiría el rastreo de la expresión con la fluorescencia GFP.

Además, la flexibilidad de los vectores simplifica el sistema de expresión transitoria, necesitando solamente la infiltración de una sola construcción de *Agrobacterium*, y mejora la eficacia al reducir la cantidad de infiltrado necesaria en proporción con el número de casetes de expresión presentes en el ADN-T. Al ocupar P19 el sitio FseI, la presencia de dos sitios de clonación para aceptar casetes HT de los vectores de clonación (tales como pM81-FSC2-U162C) también permite una expresión incluso más eficaz de proteínas de multisubunidades como anticuerpos de tamaño completo.

El efecto de P19 sobre la mejora de los niveles de expresión de transgenes está bien caracterizada (Voinnet et al., 2003). Sin embargo, este estudio presenta la primera demostración de su eficacia cuando se administran simultáneamente a cada célula con del mismo ADN-T. Un estudio anterior había informado de la administración simultánea de P19 con un ADN-T independiente dentro del mismo *Agrobacterium* como transgén que contiene ADN-T (Hellens et al., 2005). Sin embargo, no se observaron efectos del P19 hasta 6 días después de la infiltración, que sugiere una transferencia ineficaz del ADN-T. El presente estudio también demuestra el primer uso del P19 de R43W mutante para potenciar la expresión de un transgén. El hallazgo de que el mutante era aproximadamente la mitad de eficaz para potenciar la expresión de GFP que P19 de tipo salvaje concuerda con su reducción en la actividad conocida, que compromete tanto la infectividad de TBSV (Chu et al., 2000), como la capacidad de la proteína para unirse a la clase más pequeña (21 - 22 nts) de los ARN interferentes cortos (Omarov et al., 2006). Sin embargo, es posible que este rasgo convierta potencialmente al mutante R43W en más adecuado para aplicaciones que implican una transformación estable. Los micro ARN asociados con el desarrollo también se encuentran en la clase de tamaño menor (Vaucheret, 2006; Zhang et al., 2006) y, por tanto, es posible que los procesos de desarrollo no se vean tan gravemente afectados por la presencia del P19 mutante de lo que serían mediante la versión de tipo salvaje (Scholthof, 2007). Además, el mutante puede proporcionar una forma de controlar la expresión transitoria de las proteínas citotóxicas extrañas.

La expresión de 2G12 a partir de un único plásmido representa el mayor rendimiento informado de un anticuerpo procedente de tejido vegetal infiltrado con un único cultivo de *Agrobacterium*. Además de utilizar 3 cultivos de *Agrobacterium* para expresión de CPMV-HT, la única forma de conseguir niveles similares con otro sistema implicó la infiltración de 6 cultivos independientes y un enfoque de vector vírico (Giritch et al., 2006). Además, el uso de un solo plásmido permite una reducción en la cantidad de bacterias necesarias para garantizar la administración simultánea de múltiples casetes de expresión, lo que proporcionaría un ahorro de costes significativos a niveles de producción industrial. El proceso de infiltración es también físicamente más sencillo de realizar con cultivos más diluidos debido a que se produce menos cegado de los espacios intercelulares del tejido foliar. Además, la dilución hasta una DO total de 0,4 redujo la cantidad de contaminantes derivados de infiltración de proteínas. El análisis de nueve infiltraciones separadas en cada DO mostró una reducción en las concentraciones de proteína en los extractos de $2,7 \pm 0,2$ a $1,5 \pm 0,1$ mg/ml cuando la DO de los cultivo se redujo de 1,2 a 0,4. Puesto que el uso de pEAQexpress genera tanto 2G12 a DO 0,4 como los tres sistemas de cultivo a una DO infiltrada de 1,2, la proteína diana recombinante debe purificarse a partir de solamente la mitad de la cantidad de proteína contaminante usando pEAQexpress. Esto proporciona una ventaja muy útil e inesperada para el procesamiento en dirección 3'. La expresión de 2G12 a partir de pEAQexpress también indica un efecto de la posición en un casete de expresión incluido en el ADN-T de los vectores pEAQ sobre el nivel de expresión obtenido. El aumento en la acumulación de cadena ligera libre a partir de pEAQex-2G12LHE sugiere que se expresa menos cadena ligera con esta construcción, que parece dar como resultado menos anticuerpo ensamblado. Esto podría ser debido a la disposición de los casetes de expresión en el ADN-T. Como alternativa, una parte de los ADN-Ts son susceptibles a la

degradación nucleolítica en la frontera izquierda (Rossi et al., 1996). La inserción del casete NPTII dentro del ADN-T parece tener un marcado efecto sobre la expresión dependiendo de su orientación. Durante las manipulaciones de clonación resultó que pEAQselectK-GFP-HT alcanzaba un número de copias de plásmido en *E. coli* de aproximadamente 1,5 veces el de pEAQselectK(rev)-GFP-HT (determinado a partir de mediciones del rendimiento de tres preparaciones de plásmido independientes realizadas con el kit QIAprep, QIAGEN). Esto se correlaciona poco con la diferencia en los niveles de expresión observados entre los dos vectores. No se sabe qué contribuye al mayor número de copias, o incluso si la diferencia ya existe cuando los plásmidos se transfieren a *Agrobacterium*. Sin embargo, estas observaciones sugieren que el número de copias del plásmido puede ser importante para conseguir una expresión transitoria mediada por *Agrobacterium* eficaz. A este respecto, el uso del origen RK2 (oriV en la Figura 9) por pBIN19 y sus derivados le convierte en una buena elección para la expresión transitoria, ya que los plásmidos RK2 son conocidos por acumular de 7 a 10 copias en *Agrobacterium* (Veluthambi et al., 1987). Esto es similar al origen pVS1 utilizado por pPZP y aproximadamente 2-5 veces mayor de lo generado por el origen pSa (Lee y Gelvin, 2008), que está presente en el vector binario pGREEN ampliamente usado (Helens et al., 2000). Los plásmidos que contienen orígenes de replicación que proporcionan mayores números de copia tales como los plásmidos basados en pRi (Lee y Gelvin, 2008) pueden estar incluso más adecuados para la expresión transitoria.

Para conseguir un elevado nivel de expresión con vectores pEAQ fácilmente accesibles a laboratorios sin experiencia previa con la expresión basada en CPMV o incluso, la expresión basada en plantas, de forma general, se creó una versión de pEAQ para clonación directa. Esto se consiguió insertando un polienlazador entre la secuencia líder en 5' y las UTR en 3' de un casete de expresión CPMVHT, que se colocó en un ADN-T que también incluía casetes P19 y NPTII. El casete NPTII se incluyó porque su presencia parecía potenciar la expresión (véase anteriormente). El polienlazador también codifica o dos conjuntos de 6 restos histidina para permitir la fusión de las etiquetas His de los extremos N o C para facilitar la purificación de la proteína. Las construcciones resultantes también se aprovechan la segunda mutación en el líder 5' que potencia la expresión con respecto a HT.

Estos casetes de expresión potenciados también se pueden subclonar a partir del vector de clonación pM81-FSC-P0W en cualquier plásmido pEAQ. El uso de pEAQHT conduce a una mayor expresión de GFP comparada con pEAQspecialK, que contiene solamente la mutación única (U162C). Además, el diseño del polienlazador también permitió la expresión de variantes marcadas con His usando un procedimiento de clonación en una etapa. Los vectores binarios modulares presentados aquí están diseñados específicamente, pero no de forma restringida, para su uso con la expresión de CPMV-HT. Un nivel de expresión extremadamente elevado se ha vinculado con una mejor eficacia de clonación y facilidad de uso. El sistema proporciona el método más eficaz y directo para la expresión transitoria de proteínas de valor añadido sin las complicaciones de la amplificación vírica. Permite obtener cantidades de miligramos de proteína recombinante en un plazo de dos semanas desde la identificación de la secuencia en cualquier laboratorio de biología molecular con acceso a instalaciones de crecimiento de plantas. Por tanto, se anticipa que proporcionará una herramienta extremadamente valiosa en escenarios tanto académicos como industriales.

Ejemplo 5

Integración de grapa con plásmidos pEAQ y plantas transgénicas

Aunque la serie del vector pEAQ se diseñó pensando en la expresión transitoria, la inserción del casete NPTII en el ADN-T proporciona un marcador seleccionable para la integración del genoma. Esto permite potencialmente que se usen estos vectores binarios más pequeños y más útiles para transformación estable de plantas y cultivos de células vegetales. Cuando se usa para transformar discos de hojas de *N. benthamiana*, los vectores pEAQ que contienen el casete NPTII dentro del ADN-T fueron capaces de inducir la formación de callos tras selección con la misma eficacia que las construcciones basadas en pBINPLUS. Además, la expresión de GFP fue detectable en estos tejidos con luz UV (no se muestran los datos). Esto demuestra que las moléculas de ADN-T multi-casete de los vectores pEAQ se pueden integrar de forma estable en el genoma vegetal e impulsar la expresión de genes extraños.

También se regeneraron plantas fluorescentes. Las hojas de los transformantes primarios (T_0) fueron fluorescentes bajo luz UV indicando elevados niveles de expresión de GFP. La semilla de las plantas T_0 autofertilizadas fueron viables, y las plántulas T_1 resultantes que incluían el transgén también fueron fluorescentes (no se muestran los resultados).

Ejemplo 6

Uso del sistema HT basado en CPMV con vectores baculovirus

La Figura 15 muestra una construcción adecuada para usar un sistema HT basado en CPMV con vectores baculovirus en células de insecto. Bajo el control del promotor p10, las UTR HyperTrans ARN-2 de CPMV también potenciaron la expresión de GFP en células de insecto usando el sistema de expresión en baculovirus. Se obtuvo una potenciación de aproximadamente 5 veces en los niveles de fluorescencia en las células sf21 infectadas con baculovirus, medidos mediante citometría de flujo, en comparación con una construcción sin el casete CPMV-HT.

REFERENCIAS

- 5 Alamillo, J.M., Monger, W., Sola, I., Garcia, B., Perrin, Y., Bestagno, M., Burrone, O.R., Plana-Duran, J., Enjuanes, L., Lomonossoff, P.G. y Garcia, J.A. (2006) Use of virus vectors for the expression in plants of active full-length and single chain anticoronavirus antibodies. *Biotechnol. J.* 1, 1103-1111.
- 10 Brigneti, G., Voinnet, O., Li, W.x., Ji, L.H., Ding, S.Wand Baulcombe, D.C. (1998) Viral pathogenicity determinants are suppressors of transgene silencing in *Nicotiana benthamiana*. *EMBO Journal* 17: 6739-6746.
- 15 Cañizares, M.C., Liu, L., Perrin, Y., Tsakiris, E. y Lomonossoff, G.P. (2006). A bipartite system for the constitutive and inducible expression of high levels of foreign proteins in plants. *Plant Biotechnol. J.* 4, 183-193.
- 20 Gopinath, K., Wellink, J., Porta, C., Taylor, K.M., Lomonossoff, G.P. y van Kammen, A. (2000) Engineering cowpea mosaic virus RNA-2 into a vector to express heterologous proteins in plants. *Virology* 267: 159-173.
- 25 Holness, C.L., Lomonossoff, G.P., Evans, D. y Maule, A.J. (1989). Identification of the initiation codons for translation of cowpea mosaic virus middle component RNA using site directed mutagenesis of an infectious cDNA clone. *Virology* 172, 311-320.
- 30 Liu, L. y Lomonossoff, G.P. (2002) Agroinfection as a rapid method for propagating Cowpea mosaic virus-based constructs. *J. Virol. Methods* 105, 343-348.
- 35 Liu, L. y Lomonossoff, G.P. (2006) A site-directed mutagenesis method utilising large double-stranded DNA templates for the simultaneous introduction of multiple changes and sequential multiple rounds of mutation: Application to the study of whole viral genomes. *J. Virol. Methods* 137, 63-71.
- 40 Liu, L., Canizares, M.C., Monger, W., Perrin, Y., Tsakiris, E., Porta, C., Shariat, N., Nicholson, L. y Lomonossoff, G.P. (2005). Cowpea mosaic virus-based systems for the production of antigens and antibodies in plants. *Vaccine* 23, 1788-1792.
- 45 Lomonossoff, G. P. & Shanks, M. (1983). The nucleotide sequence of cowpea mosaic virus B RNA. *EMBO Journal* 2, 2253-2258.
- 50 Mechtcheriakova, I.A., Eldarov, M.A., Nicholson, L., Shanks, M., Skryabin, K.G. y Lomonossoff, G.P. (2006) The use of viral vectors to produce hepatitis B virus core particles in plants. *J. Virol. Methods* 131, 10-15.
- 55 Monger, W., Alamillo, J.M., Sola, I., Perrin, Y., Bestagno, M., Burrone, O.R., Sabella, P., Plana-Duran, J., Enjuanes, L., Garcia, J.A. y Lomonossoff, G.P. (2006) An antibody derivative expressed from viral vectors passively immunizes pigs against transmissible gastroenteritis virus infection when supplied orally in crude plant extracts. *Plant Biotechnol. J.* 4, 623-631.
- 60 Rohll, J.B., Holness, C.L., Lomonossoff, G.P. y Maule, A.J. (1993). 3' terminal nucleotide sequences important for the accumulation of cowpea mosaic virus M-RNA. *Virology* 193, 672-679.
- 65 Sainsbury, F., Lavoie, P-O., D'Aoust, M-A., Vezina, L-P. y Lomonossoff, G.P. (2008). Expression of Multiple Proteins Using Full-Length and Deleted Versions of Cowpea Mosaic Virus RNA-2. *Plant Biotechnology Journal*, 6: 82-92.
- 70 Sainsbury, F., Cañizares, M.C. y Lomonossoff, G.P. (2007) Cowpea mosaic virus-based expression vectors. In: *Virus Expression Vectors* (Hefferon, K. ed), pp. 339-555. Kerala, India: Transworld Research Network.
- 75 Sainsbury, F. y Lomonossoff, G.P. (2008). Extremely high-level and rapid transient protein production in plants without the use of viral replication. *Plant Physiology* 148, 1212-1218.
- 80 Sainsbury, F., Liu, L. y Lomonossoff G.P. (2009) Cowpea mosaic virus-based expression of antigens and antibodies in plants. En: *Methods in Molecular Biology Vol. 483 : Recombinant Pharmaceutical Proteins from Plants* (Faye, L. y Gomord, V. eds), pp25-39, NY: Humana Press.
- 85 van Bokhoven, H., Le Gall, O, Kasteel, D., Verver, J., Wellink, J. y van Kammen, A. (1993). Cis- and Trans-acting Elements in Cowpea Mosaic Virus RNA Replication. *Virology* 195, 377-386.
- 90 Wellink J, Verver J, van Kammen A. (1993). Mutational analysis of AUG codons of cowpea mosaic virus M RNA. *Biochimie.* 75(8):741-7.

Referencias adicionales:

- Lee LY, Gelvin SB (2008) T-DNA binary vectors and systems. *Plant Physiology* 146: 325-332
- 5 Hajdukiewicz P, Svab Z, Maliga P (1994) The small, versatile pPZP family of *Agrobacterium* binary vectors for plant transformation. *Plant Molecular Biology* 25: 989-994
- Hellens R, Mullineaux P, Klee H (2000b) A guide to *Agrobacterium* binary Ti vectors. *Trends in Plant Science* 5: 446-451
- 10 Hiei Y, Ohta S, Komari T, Kumashiro T (1994) Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence-analysis of the boundaries of the T-DNA. *Plant Journal* 6: 271-282
- 15 Veluthambi K, Jayaswal RK, Gelvin SB (1987) Virulence genes-a, gene-G, and gene-D mediate the double-stranded border cleavage of T-DNA from the *Agrobacterium* Ti plasmid. *Proceedings of the National Academy of Sciences of the United States of America* 84: 1881-1885
- Thole V, Worland B, Snape JW, Vain P (2007) The pCLEAN dual binary vector system for *Agrobacterium*-Mediated plant transformation. *Plant Physiology* 145: 1211-1219
- 20 Johansen LK, Carrington JC (2001) Silencing on the spot. Induction and suppression of RNA silencing in the *Agrobacterium*-mediated transient expression system. *Plant Physiology* 126: 930-938
- Janssen BJ, Gardner RC (1989) Localized transient expression of Gus in leaf-disks following cocultivation with *Agrobacterium*. *Plant Molecular Biology* 14: 61-72
- 25 Narasimhulu SB, Deng X, Sarria R, Gelvin SB (1996) Early transcription of *Agrobacterium* T-DNA genes in tobacco and maize. *Plant Cell* 8: 873-886
- 30 Hellens RP, Edwards EA, Leyland NR, Bean S, Mullineaux PM (2000a) pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. *Plant Molecular Biology* 42: 819-832
- Tzfira T, Tian GW, Lacroix B, Vyas S, Li JX, Leitner-Dagan Y, Krichevsky A, Taylor T, Vainstein A, Citovsky V (2005) pSAT vectors: a modular series of plasmids for autofluorescent protein tagging and expression of multiple genes in plants. *Plant Molecular Biology* 57: 503-516
- 35 Voinnet O, Rivas S, Mestre P, Baulcombe D (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant Journal* 33: 949-956
- 40 Frisch DA, Harrishaller LW, Yokubaitis NT, Thomas TL, Hardin SH, Hall TC (1995) Complete sequence of the binary vector BIN-19. *Plant Molecular Biology* 27: 405-409
- Trieu-cuot P, Courvalin P (1983) Nucleotide-sequence of the *Streptococcus-faecalis* plasmid gene encoding the 3'5"-Aminoglycoside phosphotransferase type-III. *Gene* 23: 331-341
- 45 Bevan M (1984) Binary *Agrobacterium* vectors for plant transformation. *Nucleic Acids Research* 12: 8711-8721
- Silhavy D, Burgyan J (2004) Effects and side-effects of viral RNA silencing suppressors on short RNAs. *Trends in Plant Science* 9: 76-83
- 50 Scholthof HB (2007) Heterologous expression of viral RNA interference suppressors: RISC management *Plant Physiology* 145: 1110-1117
- van Engelen FA, Molthoff JW, Conner AJ, Nap JP, Pereira A, Stiekema WJ (1995) pBINPLUS - an improved plant transformation vector based on pBIN19. *Transgenic Research* 4: 288-290
- 55 Rossi L, Hohn B, Tinland B (1996) Integration of complete transferred DNA units is dependent on the activity of virulence E2 protein of *Agrobacterium tumefaciens*. *Proceedings of the National Academy of Sciences of the United States of America* 93: 126-130
- 60 Xiang CB, Han P, Lutziger I, Wang K, Oliver DJ (1999) A mini binary vector series for plant transformation. *Plant Molecular Biology* 40: 711-717
- 65 Hellens RP, Allan AC, Friel EN, Bolitho K, Grafton K, Templeton MD, Karunairetnam S, Gleave AP, Laing WA (2005) Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. *Plant Methods* 1: 13

- Chu M, Desvoyes B, Turina M, Noad R, Scholthof HB (2000) Genetic dissection of tomato bushy stunt virus p19-protein-mediated host-dependent symptom induction and systemic invasion. *Virology* 266: 79-87
- 5 Omarov R, Sparks K, Smith L, Zindovic J, Scholthof HB (2006) Biological relevance of a stable biochemical interaction between the tombusvirus-encoded P19 and short interfering RNAs. *Journal of Virology* 80: 3000-3008
- Vaucheret H (2006) Post-transcriptional small RNA pathways in plants: mechanisms and regulations. *Genes & Development* 20: 759-771
- 10 Zhang BH, Pan XP, Cobb GP, Anderson TA (2006) Plant microRNA: A small regulatory molecule with big impact. *Developmental Biology* 289: 3-16
- Giritch A, Marillonnet S, Engler C, van Eldik G, Botterman J, Klimyuk V, Gleba Y (2006) Rapid high-yield expression of full-size IgG antibodies in plants coinfecting with noncompeting viral vectors. *Proceedings of the National Academy of Sciences of the United States of America* 103: 14701-14706
- 15

Tabla 1

Segmento del genoma completo de ARN-2 de CPMV (nucleótidos de 1 a 3481)	
1	tattaaaatc tfaataggtt ttgataaaag cgaacgtggg gaaacccgaa ccaaaccttc
61	ttctaaatc tctctcatct ctcttaaagc aaacttctct ctgtcttcc ttgcatgagc
121	gatctcaac gttgcagat cgtgctcgg caccagtaca atgttttct tcactgaagc
181	gaaatcaaaag atctcttgt ggacacgtag tgcgcgcca taaataacg tgtactgtc
241	ctattctgt cgggtggtc ttgggaaaag aaagcttgc ggaggctgt gttcagccc
301	atacattact tgttacgatt ctgctgact tcggcgggtg caatactct acttctgct
361	gacgaggtat tttgcctgt acttcttct tcttctct gctgattgt tctataagaa
421	atctagtatt tctttgaaa cagagtttc ccgtggttt cgaacttga gaaagattg
481	taagctctg tataattgc ccaaattga aatgaaagc attatgagcc gtggtattcc
541	ttcaggaatt ttggaggaaa aagctattca gttcaacgt gccaaagaag ggaataaacc
601	ctgaaggat gagattcca agcctgagga tatgatgtg tctcacact ctaaattgaa
661	tgtctcaga aaaatgacc aaaagactgt ggatcttcc aaagcagctg ctggatggg
721	attcatcaat aagcatatgc ttacgggcaa catctggca caaccaaca cagtctgga
781	tattccgctc acaaaggata aaacactgc gatggcagc gattttattc gtaaggagaa
841	tctcaagact tctgcaatc acattggagc aattgagatt attatccaga gctttgctc
901	ccctgaaagt gattgatgg gaggctttt gcttggat tcttacaca ctgatacagc
961	taatgctatt cgtagcatt ttgtgctcc aatgcgggga ggaagaccag tcagagtgg
1021	gaccttcca aatacactgg cacctgtatc atgtgatctg aacaatagat tcaagctat
1081	ttgctcattg ccaaactgtg atattgtcca ggtagccaa gtacagaag tgagtgtaa
1141	ttgtgcagga tgtgctact ccatagagaa atctcacacc ccttccaat tgtatacaga
1201	ggaattgaa aaggaggggt ctgttgtgt agaatacta ggcagacaga cctattgtc
1261	tcagctagc aattacca cagaagaaa acttcgctc ctaagttg acttcatgt
1321	tgaacaacca agtgcctga agttatcaa ttctgcaat cgcactttg tcaagggaga
1381	aagttgaaa tactctatt ttggcaaga agcagaaaac catgcagtc atgctactgt
1441	ggtctctga gaaggggct ctgcccacc caagcaat atctctatt tggacgggt
1501	gctggatcca cgaatggga atgtggctt tcacaaaatg gagcaaaact ttttgccct
1561	ttcttggat gatacaagct cagttcgtg ttcttctgac gacacaaaat tcgcacaaac
1621	tcgagtttg ttgtccaagg ctatggctgg ttgtgatgtg ttattggatg agtatctca
1681	tgatgtggtc aatggacaag attttagagc tactgtcgtc ttttgcgca cccatgtat
1741	aacaggcaaa ataaagggtga cagctaccac caacattct gacaactcg gttgtgtt
1801	gatgtggcc ataaatagtg gtgtgagggg taagtatagt actgatgtt atactatctg
1861	ctctcaagc tccatgacgt ggaaccagg gtgcaaaaag aacttctgt tcaatttaa
1921	tccaaaccct tgtggggtt ctgtgtctg tgagatgata agtcgaagca gagttaggat
1981	gacagttatt tgtttcgg gatggacct atctctacc acagatgtga ttgccaagct
2041	agactgtca atgtcaatg agaaatgga gccaccatt taccacttg ctgattgtca
2101	gaattgtta cccctaatc gttggatggg aaaattgact ttcccagg gtgtgacaag
2161	tgaggttca aggatgcctc ttctatagg aggcgtgct ggtcagactc aagcttct
2221	ggccaatag ccaattcat ggatatcaat gtggagatc tttagagggt aactcact
2281	tgaagttact aaaatgact cccatatac taaagccact gttacattc tcatagctt
2341	tgttaactct agtgatcct ttgtttta tgagatttt cctcatagaa ttgtcaatt
2401	tgctgaggtt gaggaaaaat gtactttgt ttctccca caagatttg tcactgctg
2461	gtcaacacaa gtaaaccca gaaccacact tgaagcagat gttgtccct acctatg
2521	aattattcat gatagtaca caggtacaat ctccggagat ttaattctg ggtcaagct
2581	tgttgcaat aaggatttt ttggtatagg ttctaatccg ggtattgat gttcccgt
2641	gcttggagct atagcacaag gacctgttg tctgaagcc tcagatgtg atagcccatg
2701	tatgatagct agcactcctc ctgctcatt ttcagacgtt acagcagtaa ctttgact
2761	aatcaacggc aaaataactc ctgttggtga tgacaattg aatacgcaca tftaatcc
2821	tccaattatg aatgtctgct gtaactgctg ttggaatct ggaactattc atgtcaact

```

2881 taatgtagg ggtgctggtg taaaagagc agattggat ggcaagctc ttgttacct
2941 gcgccagtc atgaacctg aaagttatga tgcgaggaca ttgtgatc cacaacctg
3001 ttctgcatg ttgaactct ctttgatg catagggccg aatagcggat tgaattgc
3061 cгааagcca tggccaatc agaccactg gtatctgaa tgtgtgcta ccaatcccag
3121 acaaatacag caattgagg tcaacatgcg ctctgatct aattcaggg ttgccggcaa
3181 tctctgatg cccccattc cactgtcaac ggaaactcca ccgtattaa agtttaggt
3241 tcgggatatt gaacgctcca agcgtagtgt tatggttga cacactgcta ctgctgcta
3301 actctggtt cattaattt tcttagttt gaattactg ttattgggtg tgcatttcta
3361 tgttgggtg gcggtttct gtgctcagag tgtgttatt ttatgtaatt taatttctt
3421 gtgagctct gtttagcagg tctccctc agcaaggaca caaaaagatt taaatttat
3481 t
    
```

Los codones de inicio en las posiciones 115, 161, 512 y 524 del segmento del genoma de ARN-2 de CPMV se muestran en negra y subrayado.

5

Tabla 2

<u>Oligonucleótidos utilizados en la mutagénesis de la región en 5' de los clones de pM81-FSC-2</u>		
Oligonucleótido	Secuencia	Mutación
A115G-F	CTTGTCTTTCTTGC<u>GT</u>GAGCGATCTT CAACG	Elimina AUG (→GUG) en 115 eliminando la traducción de uORF
A115G-R	CGTTGAAGATCGCTCA<u>C</u>GCAAGAAAG ACAAG	
U162C-F	GGCACCAGTACAAC<u>G</u>TTTTCTTTCAC TGAAGCG	Elimina AUG (→ACG) en 161 eliminando la traducción de AUG 161 manteniendo a la vez la secuencia de aminoácidos de uORF
U162C-R	CGCTTCAGTGAAAGAAAAC<u>G</u>TTGTAC TGGTGCC	

Los nucleótidos mutantes de los oligonucleótidos usados en la mutagénesis de la región 5' de los clones pM81-FSC-2 se muestran en negra.

TABLA 3

Secuencia CPMV de tipo salvaje de la Tabla 1	tatattctgc ccaaattga <u>aat</u> ggaaagc <u>attat</u> gagcc gtggtattcc
Secuencia mutada en pM81-FSC-1	tatattctgc ccaaattt<u>g</u>T <u>Cat</u>gAaaagc <u>attat</u>gagcc gtggtattcc 509 BspH1
Secuencia mutada en pM81-FSC-2	tatattctgc ccaaatt<u>CGC</u> GACGATCGTA CTCTCGAGGC CT 507 NruI XhoI

10

Las diferencias de nucleótidos entre la secuencia de los vectores pM81-FSC-1 y pM81-FSC-2 y la secuencia CPMV de tipo salvaje de la Tabla 1 se muestran en mayúsculas.

LOCUS pm81-FSC1 7732 pb ADN circular
 10-OCT-2007

CARACTERÍSTICAS	Localización/ Modificadores
5'UTR	342..501 /vntifkey="52" /label=CPMV\RNA2\5'UTR
promotor	27..341 /vntifkey="29" /label=CaMV\35S\promoter
terminador	4669..4921 /vntifkey="43" /label=Nos\Terminator
mat_péptido	3712..4422 /vntifkey="84" /label=GFP
3'UTR	4432..4615 /vntifkey="50" /label=CPMV\RNA2\3'UTR
CDS	complement(5944..6804) /vntifkey="4" /label=AmpR
misc_feature	complement(7391..7546) /vntifkey="21" /label=lacZ_a
promotor	complement(6846..6874) /vntifkey="30" /label=AmpR\promoter
rep_origen	complement(7067..7373) /vntifkey="33" /label=f1_origin
rep_origen	complement(5170..5789) /vntifkey="33" /label=pBR322_origin
mat_péptido	502..1878 /vntifkey="84" /label=CPMV\Movement\Protein
mat_péptido	1879..2999

```

/vntifkey="84"
/label=CPMV\Lg.\Coat\Protein
mat_péptido 3000..3638
/vntifkey="84"
/label=CPMV\Sm.\Coat\Protein
RECUENTO DE BASES 2105 a 1682 c 1770 g 2175 t
ORIGEN

1 ttaattaaga attcgagctc caccgcggaa acctcctcgg attccattgc ccagctatct
61 gtcactttat tgagaagata gtggaaaagg aagggtggctc ctacaaatgc catcattgog
121 ataaaggaaa ggccatcggt gaagatgcct ctgccgacag tggccccaaa gatggacccc
181 caccacagag gagcatcgtg gaaaaagaag acgttccaac cacgtcttca aagcaagtgg
241 attgatgtga tatctccact gacgtaaggg atgacgcaca atcccactat ccttcgcaag
301 acccttcctc tatataagga agttcatttc atttggagag gtattaaaaat cttaataggt
361 tttgataaaa gcgaacgtgg ggaaacccga accaaacctt cttctaaact ctctctcatc
421 tctcttaaaag caaacttctc tcttgtcttt cttgcatgag cgatcttcaa cgttgtcaga
481 tcgtgcttcg gcaccagtac aatgttttct ttcactgaag cgaaatcaa gatctctttg
541 tggacacgta gtgcccgcgc attaaataac gtgtacttgt cctattcttg tgggtgtggt
601 cttgggaaaa gaaagcttgc tggaggctgc tgttcagccc catacattac ttgttacgat
661 tctgtgact ttcggcgggt gcaatatctc tacttctgct tgacgaggta ttgttgctg
721 tacttctttc ttcttcttct tgetgattgg ttctataaga aatctagtat tttctttgaa
781 acagagtttt cccgtggttt tcgaacttgg agaaagattg ttaagcttct gtatattctg
841 cccaaatttg tcatgaaaag cattatgagc cgtgggtattc cttcaggaat tttggaggaa
901 aaagctattc agttcaaagc tgccaaagaa gggaaataaac ccttgaagga tgagattccc
961 aagcctgagg atatgtatgt gtctcacact tctaaatgga atgtgctcag aaaaatgagc
1021 caaaagactg tggatctttc caaagcagct gctgggatgg gattcatcaa taagcatatg
1081 cttacgggca acatcttggc acaaccaaca acagtcttgg atattcccgt cacaaggat
1141 aaaacacttg cgatggccag tgattttatt cgtaaggaga atctcaagac ttctgccatt
1201 cacattggag caattgagat tattatccag agctttgctt cccctgaaag tgatttgatg
1261 ggaggctttt tgcttgtgga ttctttacac actgatacag ctaatgctat tcgtagcatt
1321 tttgttgctc caatgcgggg aggaagaacca gtcagagtgg tgaccttccc aaatacactg
1381 gcacctgtat tatgtgatct gaacaataga ttcaagctca ttgtctcatt gccaaactgt
1441 gatattgtcc agggtagcca agtagcagaa gtgagtgtaa atgttgcagg atgtgctact
1501 tccatagaga aatctcacac cccttcccaa ttgtatacag aggaatttga aaaggagggg
1561 gctgttgttg tagaatactt aggcagacag acctattgtg ctacgcctag caatttacc
1621 acagaagaaa aacttcggtc ccttaagttt gactttcatg ttgaacaacc aagtgtcctg
1681 aagttatcca attcctgcaa tgcgcacttt gtcaaggaa aaagtttgaa atactctatt
1741 tctggcaaaag aagcagaaaa ccatgcagtt catgctactg tggctctctg agaaggggct
1801 tctgcccac ccaagcaata tgatcctatt ttgggacggg tgctggatcc acgaaatggg
1861 aatgtggctt ttccacaaat ggagcaaaac ttgtttgccc tttctttgga tgatacaagc

```

1921 tcagttcgtg gttctttgct tgacacaaaa ttgcacaaa ctcgagtttt gttgtccaag
 1981 gctatggctg gtggtgatgt gttattggat gagtatctct atgatgtggt caatggacaa
 2041 gatttttagag ctactgtcgc ttttttgcgc acccatgtta taacaggcaa aataaagggtg
 2101 acagctacca ccaacatttc tgacaactcg ggttggtggt tgatggtggc cataaatagt
 2161 ggtgtgaggg gtaagtatag tactgatggt tatactatct gctctcaaga ctccatgacg
 2221 tggaaaccag ggtgcaaaaa gaacttctcg ttcacattta atccaaacc ttgtgggat
 2281 tcttggctcg ctgagatgat aagtcgaagc agagttagga tgacagttat ttgtgttcg
 2341 ggatggacct tatctcctac cacagatgtg attgccagc tagactggtc aattgtcaat
 2401 gagaaatgtg agcccacat ttaccacttg gctgattgtc agaattggtt acccctaat
 2461 cgttggatgg gaaaattgac ttttccccag ggtgtgacaa gtgaggttcg aaggatgcct
 2521 ctttctatag gaggcgtgc tgggtgcgact caagctttct tggccaatat gcccaattca
 2581 tggatatcaa tgtggagata ttttagaggt gaacttcact ttgaagttac taaaatgagc
 2641 tctccatata ttaaagccac tgttacattt ctcatagctt ttggtaatct tagtgatgcc
 2701 tttggttttt atgagagttt tctcataga attgttcaat ttgctgaggt tgaggaaaa
 2761 tgtactttgg ttttctccca acaagagttt gtcactgctt ggtcaacaca agtaaacccc
 2821 agaaccacac ttgaagcaga tggttgtccc tacctatatg caattattca tgatagtaca
 2881 acaggtacaa tctccggaga ttttatcttg gggcaagct tgttggcatt aaggattttt
 2941 gtggtatagg ttctaaccg ggtattgatg gttcccgtt gcttggagct atagcacaag
 3001 gacctgtttg tgetgaagcc tcagatgtgt atagcccatg tatgatagct agcactctc
 3061 ctgctccatt ttcagacgtc acagcagtaa acttttgact taatcaacgg caaaataact
 3121 cctgttgggtg atgacaattg gaatacgcac atttataatc ctccaattat gaatgtcttg
 3181 cgtactgctg cttggaatc tggaaactatt catgttcaac ttaatgtag ggggtgctggt
 3241 gtcaaaaagag cagattggga tggtaagtc tttgtttacc tgcgccagtc catgaaccct
 3301 gaaagttatg atgcgcggac atttgtgate tcacaacctg gttctgccat gttgaacttc
 3361 tcttttgata tcatagggcc gaatagcggg tttgaatttg ccgaaagccc atgggccaat
 3421 cagaccacct ggtatcttga atgtgttgc accaatccca gacaaataca gcaatttgag
 3481 gtcaacatgc gcttcgatcc taatttcagg gttgcggca atatcctgat gcccacattt
 3541 cactgtcaa cggaaactcc accgttatta aagtttaggt ttogggatat tgaacgctcc
 3601 aagcgtagtg ttatggttgg acacactgct actgctgcag cgcctgcaaa acagctctta
 3661 aactttgacc tacttaagtt agcaggtgac gttgagtcca accctgggccc cagtaaagga
 3721 gaagaacttt tcaactggag tgtcccaatt cttgttgaat tagatggtga tgttaatggg
 3781 cacaaatfff ctgtcagtgg agaggggtgaa ggtgatgcaa catacggaaa acttaccctt
 3841 aaatttattt gcactactgg aaaactacct gttccatggc caacacttgt cactactttc
 3901 tcttatgggtg ttcaatgctt ttcaagatac ccagatcata tgaaacggca tgactttttc
 3961 aagagtgcc a tgcggagg ttatgtacag gaaagaacta tatttttcaa ggatgacggg
 4021 aactacaaga cacgtgctga agtcaagttt gaaggtgata cccttggtta tagaatcgag
 4081 ttaaaaggta ttgattttaa agaagatgga aacattcttg gacacaaatt ggaatacaac
 4141 tataactcac acaatgtata catcatggca gacaaacaaa agaattggaat caaagttaac
 4201 ttcaaaatta gacacaacat tgaagatgga agcgttcaac tagcagacca ttatcaacaa
 4261 aatactccaa ttggcgatgg cctgtcctt ttaccagaca accattacct gtccacacaa

4321 tctgcccttt cgaaagatec caacgaaaag agagaccaca tggtccttct tgagtttgta
4381 acagctgctg ggattacaca tggcatggat gaactataca aataaaggcc tttactctg
4441 gtttcattaa attttcttta gtttgaattt actgttattc ggtgtgcatt tctatgtttg
4501 gtgagcgggt ttctgtgctc agagtgtgtt tttttatgt aatttaattt ctttgtgagc
4561 tcctgttttag caggtcgtcc cttcagcaag gacacaaaaa gattttaatt ttattaaaaa
4621 aaaaaaaaaa aaagaccggg aattcgatat caagcttatac gacctgcaga tcgttcaaac
4681 atttggcaat aaagtttctt aagattgaat cctgttgccg gtcttgcgat gattatcata
4741 taatttctgt tgaattacgt taagcatgta ataattaaca tgtaatgcat gacgttattt
4801 atgagatggg tttttatgat tagagtcccg caattataca tttaatcgc gatagaaaac
4861 aaaatatagc gcgcaacta ggataaatta tcgcgcgcggt gtcatctat gttactagat
4921 ctctagagtc tcaagcttgg cgcgccagct gcattaatga atcggccaac gcgcggggag
4981 aggcggtttg cgtattgggc gctcttcgc ttcctcctc actgactcgc tgcgctcggg
5041 cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggg tatccacaga
5101 atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg
5161 taaaaaggcc gcgttgctgg cgtttttcca taggctcgc cccctgacg agcatcacia
5221 aatcgcgc tcagtacaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt
5281 tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct
5341 gtccgccttt ctccctcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct
5401 cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc
5461 cgaccgctgc gccttatecg gtaactatcg tcttgagtcc aaccggtaa gacacgactt
5521 atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcgggtg
5581 tacagagttc ttgaagtggg ggcctaacta cggctacact agaagaacag tatttgggat
5641 ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa
5701 acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa
5761 aaaaggatct caagaagatc ctttgatctt ttctacggg tctgacgctc agtggaaacga
5821 aaactcacgt taagggattt tggttatgag attatcaaaa aggatcttca cctagatcct
5881 tttaaattaa aatgaagtt taaatcaat ctaaagtata tatgagtaaa cttggtctga
5941 cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcac
6001 catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg
6061 cccagtgct gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat
6121 aaaccagcca gccggaaggg ccgagcgcag aagtggctct gcaactttat ccgcctccat
6181 ccagtctatt aattggtgcc gggaaagctag agtaagtagt tcgccagtta atagtttgcg
6241 caacgttggt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc
6301 attcagctcc ggttcccaac gatcaaggcg agttacatga tccccatgt tgtgcaaaaa
6361 agcggttagc tcttcggtc ctccgatcgt tgtcagaagt aagttggccg cagtgttatac
6421 actcatgggt atggcagcac tgcataatc tcttactgtc atgccatccg taagatgctt
6481 ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag
6541 ttgctcttgc ccggcgtcaa tacgggataa taccgcgcca catagcagaa ctttaaaagt
6601 gctcatcatt ggaaaacggt cttcggggcg aaaactctca aggatcttac cgctgttgag
6661 atccagttcg atgtaacca ctcgtgcacc caactgatct tcagcatctt ttactttcac

6721 cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc
 6781 gacacggaag tgttgaatac tcatactctt cctttttcaa tattattgaa gcatttatca
 6841 gggttattgt cttatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg
 6901 ggttccgcgc acatttcccc gaaaagtgcc acctaaattg taagcgtaa tattttgta
 6961 aaattcgcgt taaatTTTTg ttaaatacagc tcatttttta accaataggc cgaaatcggc
 7021 aaaatccctt ataaatcaaa agaataagacc gagatagggg tgagtgttgt tccagtttgg
 7081 aacaagagtc cactattaaa gaacgtggac tccaacgtca aagggcgaaa aaccgtctat
 7141 cagggcgatg gccactacg tgaaccatca ccctaatcaa gttttttggg gtcgaggtgc
 7201 cgtaaagcac taaatcggaa ccctaaaggg agccccgat ttagagcttg acggggaaaag
 7261 ccggcgaacg tggcgagaaa ggaagggaag aaagcgaaaag gagcgggcgc tagggcgctg
 7321 gcaagtgtag cggtcacgct gcgcgtaacc accacacccg ccgcgcttaa tgcgccccta
 7381 cagggcgcggt cccattcggc attcaggctg cgcaactgtt ggggaagggcg atcgggtcgg
 7441 gcctcttcgc tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg attaagttgg
 7501 gtaacgacag ggTTTTCCA gtcacgacgt tgtaaaacga cggccagtga gtactttggc
 7561 gtaatcatgg tcatagctgt ttcctgtgtg aaattgttat ccgctcacia ttccacacia
 7621 catacgagcc ggaagcataa agtgtaaagc ctgggggtgcc taatgagtga gctaactcac
 7681 attaattgag ttgcgctcac tgcccgttt ccagtcggga aacctggcgc gc

//

Secuencia de nucleótidos de pM81-FSC-2

LOCUS pM81-FSC2 4173 pb ADN circular
10-OCT-2007

CARACTERÍSTICAS	Localización/ Modificadores
rep_origen	complement (1271..1890) /vntifkey="33" /label=pBR322_origen
rep_origen	complement (3168..3474) /vntifkey="33" /label=f1_origen
promotor	complement (2947..2975) /vntifkey="30" /label=AmpR\promoter
misc_feature	complement (3492..3647) /vntifkey="21" /label=lacZ_a
CDS	complement (2045..2905) /vntifkey="4" /label=AmpR
3'UTR	533..716 /vntifkey="50" /label=CPMV\RNA2\3'UTR
terminador	770..1022 /vntifkey="43" /label=Nos\Terminator
promotor	3859..4173 /vntifkey="29" /label=CaMV\35S\promoter
5'UTR	1..160 /vntifkey="52" /label=CPMV\RNA2\5'UTR
misc_feature	507..532 /vntifkey="21" /label=FSC-2\MCS

RECUESTO DE BASES 1090 a 969 c 982 g 1132 t
ORIGEN

1 tattaaaatc ttaataggtt ttgataaaag cgaacgtggg gaaacccgaa ccaaacttc

61 ttctaaactc tctctcatct ctcttaaagc aaacttctct cttgtcttcc ttgcatgagc
121 gatcttcaac gttgtcagat cgtgcttcgg caccagtaca atgttttctt tcaactgaagc
181 gaaatcaaag atctctttgt ggacacgtag tgcggcgcca ttaaataacg tgtacttgtc
241 ctattcttgt cgggtgtggc ttgggaaaag aaagcttgct ggaggetgct gttcagcccc
301 atacattact tgttacgatt ctgctgactt tcggcgggtg caatatctct acttctgctt
361 gacgaggtat tgttgccctgt acttcttctt tcttcttctt gctgattggt tctataagaa
421 atctagtatt ttctttgaaa cagagtttcc cgtggtttt cgaacttggg gaaagattgt
481 taagcttctg tatattctgc ccaaatctgc gacgacgta ctctcgaggc cttaaactct
541 ggtttcatta aattttcttt agtttgaatt tactgttatt cgggtgtgat ttctatgttt
601 ggtgagcggg tttctgtgct cagagtgtgt ttattttatg taatttaatt tctttgtgag
661 ctctgttta gcaggtcgtc ccttcagcaa ggacacaaaa agattttaat tttattaaaa
721 aaaaaaaaaa aaaagaccgg gaattcgata tcaagcttat cgacctgcag atcgttcaaa
781 catttgcaa taaagtttct taagattgaa tctgttgcc ggtcttgoga tgattatcat
841 ataattctg ttgaattacg ttaagcatgt aataattaac atgtaatgca tgacgttatt
901 tatgagatgg gtttttatga ttagagtccc gcaattatac atttaatacg cgatagaaaa
61 caaaatatag cgcgcaaact aggataaatt atcgcgcgcg gtgtcatcta tgttactaga
1021 tctctagagt ctcaagcttg ggcgcgcagc tgcattaatg aatcggcaa cgcgcgggga
1081 gaggcgggtt gcgtattggg cgtctctccg ctctctcctt cactgactcg ctgcgctcgg
1141 tcgttcggct gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag
1201 aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc
1261 gtaaaaaggc cgcgttgetg gcgttttctc ataggctccg cccccctgac gagcatcaca
1321 aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt
1381 tteccccctg aagctccctc gtgcctctc ctgttccgac cctgccgctt accggatacc
1441 tgtccgectt tctccctcg ggaagcgtgg cgttttctca tagctcagc tgtaggtatc
1501 tcagttcggg gtaggtcgtt cgctccaagc tgggctgtgt gcaagcaacc cccgttcagc
1561 ccgaccgctg cgccttatcc ggtaaactat gtcttgagtc caaccggta agacacgact
1621 tatcgccact ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcgggtg
1681 ctacagagtt cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggtg
1741 tctgcgctct gctgaagcca gttacctcg gaaaaagagt tggtagctct tgatccggca
1801 aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa
1861 aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaaag
1921 aaaactcacg ttaagggatt ttggttatga gattatcaaa aaggatcttc acctagatcc
1981 ttttaaatga aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg
2041 acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat
2101 ccatagttgc ctgactcccc gtcgtgtaga taactacgat acgggagggc ttaccatctg
2161 gccccagtgc tgcaatgata ccgcgagacc cacgctcacc ggtccagat ttatcagcaa
2221 taaaccagcc agccggaagg gccgagcgca gaagtggctc tgcaacttta tccgctcca
2281 tccagtctat taattgttgc cgggaagcta gagtaagtag ttcgccagtt aatagtttgc
2341 gcaacgttgt tgccattgct acaggcacgc tgggtgtcacg ctctcgttt ggtatggctt
2401 cattcagctc cgtttcccaa cgatcaaggc gagttacatg atccccatg ttgtgcaaaa

2461 aagcggtag ctcttcggt cctccgatcg ttgtcagaag taagttggcc gcagtggtat
 2521 cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct
 2581 tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga
 2641 gttgctcttg cccggcgtca atacgggata ataccgcgcc acatagcaga actttaaaag
 2701 tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga
 2761 gatccagttc gatgtaacce actcgtgac ccaactgatc ttcagcatct tttactttca
 2821 ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg
 2881 cgacacggaa atggtgaata ctcatactct tcctttttca atattattga agcatttatc
 2941 agggttattg tcttatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag
 3001 gggttccgcg cacatttccc cgaaaagtgc cacctaaatt gtaagcgtta atattttggt
 3061 aaaattcgcg ttaaattttt gttaaactag ctcatTTTTT aaccaatagg ccgaaatcgg
 3121 caaaatccct tataaatcaa aagaatagac cgagataggg ttgagtgttg ttcagtttg
 3181 gaacaagagt cactattaa agaactgga ctccaactgc aaaggcgaa aaaccgtcta
 3241 tcagggcgat ggcccactac gtgaaccatc accctaatca agttttttgg ggtcagggtg
 3301 ccgtaaagca ctaaactcga accctaaagg gagccccga ttagagctt gacggggaaa
 3361 gccggcgaac gtggcgagaa aggaagggaa gaaagcgaaa ggagcggcg ctagggcgct
 3421 ggcaagtgta gcggtcacgc tgcgcgtaac caccacacc gccgcgctta atgcgcgct
 3481 acagggcgcg tcccattcgc cattcaggct gcgcaactgt tgggaagggc gatcgggtcg
 3541 ggcctcttgc ctattacgce agctggcgaa agggggatgt gctgcaaggc gattaagttg
 3601 ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg agtactttgg
 3661 cgtaatcatg gtcatactg tttcctgtgt gaaattgta tccgctcaca attecacaca
 3721 acatacgagc cggaagcata aagtgtaaag cctgggggtgc ctaatgagtg agctaactca
 3781 cattaattgc gttgcgctca ctgccgctt tccagtcggg aaacctggcc gcttaattaa
 3841 gaattcgagc tccaccgcg aaacctctc ggattccatt gccagctat ctgtcacttt
 3901 attgagaaga tagtgaaaa ggaaggtggc tcctacaaat gccatcattg cgataaagga
 3961 aaggccatcg ttgaagatgc ctctgccgac agtgggtcca aagatggacc cccaccacg
 4021 aggagcatcg tggaaaaaga agacgtcca accacgtctt caaagcaagt ggattgatgt
 4081 gatatctcca ctgacgtaag ggatgacgca caatcccact atccttcgca agacccttcc
 4141 tctatataag gaagttcatt tcatttggag agg

//

REIVINDICACIONES

1. Un sistema de expresión génica que comprende:

- 5 (a) una secuencia potenciadora de la expresión derivada del segmento del genoma de ARN-2 de un virus de ARN bipartido de *Comoviridae*, en el que se ha mutado un sitio de inicio de traducción en el segmento del genoma de ARN-2, en donde el segmento del genoma del ARN-2 de tipo salvaje del virus *Comoviridae* codifica dos carboxiproteínas coterminales a través de dos sitios de inicio de traducción diferentes situados en el mismo marco de lectura del triplete, en donde el sitio de inicio mutado es el primero de estos dos sitios de inicio y
10 corresponde al sitio de inicio en la posición 161 en el segmento de ARN-2 del CPMV de tipo salvaje; en donde la secuencia potenciadora de la expresión bien:
- (i) comprende al menos una secuencia de 200 nucleótidos del segmento del genoma de ARN-2 que incluye el sitio de inicio mutado, o
15 (ii) tiene al menos una identidad del 70 % con los nucleótidos 1 a 507 del segmento del genoma de ARN-2 de la secuencia de CPMV mostrada en la Tabla 1, y en donde el sitio de inicio en la posición 161 del segmento del genoma de ARN-2 de CPMV de tipo salvaje se ha mutado y
- (b) un gen heterólogo que codifica una proteína de interés
20 en donde el gen que codifica la proteína de interés está situado en dirección 3' de la secuencia potenciadora y está unido operativamente a la secuencia promotora y a la secuencia de terminación.

2. Un sistema de expresión génica que comprende una construcción de expresión génica que comprende:

- 25 (a) una secuencia potenciadora de la expresión derivada del segmento del genoma de ARN-2 de un virus de ARN bipartido de *Comoviridae*, en el que se ha mutado un sitio de inicio en el segmento del genoma de ARN-2, en donde el segmento del genoma del ARN-2 de tipo salvaje del virus *Comoviridae* codifica dos carboxiproteínas coterminales a través de dos sitios de inicio de traducción diferentes situados en el mismo marco de lectura del triplete,
30 en donde el sitio de inicio mutado es el primero de estos dos sitios de inicio y corresponde al sitio de inicio en la posición 161 en el segmento de ARN-2 del virus del mosaico del caupí (CPMV) de tipo salvaje, en donde la secuencia potenciadora de la expresión bien:
- (i) comprende al menos una secuencia de 200 nucleótidos del segmento del genoma de ARN-2 que incluye el sitio de inicio mutado, o
35 (ii) tiene al menos una identidad del 70 % con los nucleótidos 1 a 507 del segmento del genoma de ARN-2 de la secuencia de CPMV mostrada en la Tabla 1, y en donde el sitio de inicio en la posición 161 del segmento del genoma de ARN-2 de CPMV de tipo salvaje se ha mutado; y
- (b) una secuencia heteróloga para facilitar la inserción de un gen que codifica una proteína de interés en el sistema de expresión génica, en donde la secuencia heteróloga está situada en dirección 3' del sitio de inicio mutado en la secuencia potenciadora; y opcionalmente,
40 (c) una UTR en 3'.

45 3. Un sistema de expresión génica de acuerdo con la reivindicación 1 o con la reivindicación 2 que comprende o que comprende además una UTR en 3' que se deriva opcionalmente del mismo virus de ARN bipartido.

50 4. Un sistema de expresión génica de acuerdo con una cualquiera de las reivindicaciones 1 a 3, en el que el virus de ARN bipartido es un comovirus, en donde el comovirus es opcionalmente un CPMV.

55 5. Un sistema de expresión génica de acuerdo con la reivindicación 4, en el que la secuencia potenciadora comprende al menos los nucleótidos 10 a 512, 20 a 512, 30 a 512, 40 a 512, 50 a 512, 100 a 512, 150 a 512, 1 a 514, 10 a 514, 20 a 514, 30 a 514, 40 a 514, 50 a 514, 100 a 514, 150 a 514, 1 a 511, 10 a 511, 20 a 511, 30 a 511, 40 a 511, 50 a 511, 100 a 511, 150 a 511, 1 a 509, 10 a 509, 20 a 509, 30 a 509, 40 a 509, 50 a 509, 100 a 509, 150 a 509, 1 a 507, 10 a 507, 20 a 507, 30 a 507, 40 a 507, 50 a 507, 100 a 507 o 150 a 507 de una secuencia del segmento del genoma de ARN-2 del comovirus con dicho sitio de inicio diana mutado.

60 6. Un sistema de expresión génica de acuerdo con la reivindicación 5 en el que la secuencia potenciadora comprende los nucleótidos 10 a 512, 20 a 512, 30 a 512, 40 a 512, 50 a 512, 100 a 512, 150 a 512, 1 a 514, 10 a 514, 20 a 514, 30 a 514, 40 a 514, 50 a 514, 100 a 514, 150 a 514, 1 a 511, 10 a 511, 20 a 511, 30 a 511, 40 a 511, 50 a 511, 100 a 511, 150 a 511, 1 a 509, 10 a 509, 20 a 509, 30 a 509, 40 a 509, 50 a 509, 100 a 509, 150 a 509, 1 a 507, 10 a 507, 20 a 507, 30 a 507, 40 a 507, 50 a 507, 100 a 507 o 150 a 507 de la secuencia del segmento del genoma de ARN-2 de CPMV mostrada en la Tabla 1, en el que el sitio de inicio en la posición 161 del segmento del genoma de ARN-2 de CPMV de tipo salvaje se ha mutado.
65

7. Un sistema de expresión génica de acuerdo con la reivindicación 4, en el que la secuencia potenciadora tiene al menos 99 %, 98 %, 97 %, 96 %, 95 %, 90 %, 85 %, 80 % o 75 % de identidad con la secuencia del segmento del genoma de ARN-2 de CPMV mostrado en la Tabla 1, en donde el sitio de inicio en la posición 161 del segmento del genoma de ARN-2 de CPMV de tipo salvaje se ha mutado.

8. Un sistema de expresión génica de acuerdo con una cualquiera de las reivindicaciones 3 a 7 que comprende:

- (a) un promotor;
- (b) nucleótidos 1 a 507 de la secuencia del segmento del genoma de ARN-2 del virus del mosaico del caupí mostrada en la Tabla 1, en donde el AUG en la posición 161 se ha mutado como se muestra en la Tabla 2, situada en dirección 3' del promotor;
- (c) un gen que codifica una proteína de interés situada en dirección 3' de la secuencia definida en (b);
- (d) nucleótidos 3302 a 3481 de la secuencia del segmento del genoma de ARN-2 del virus del mosaico del caupí mostrada en la Tabla 1, situada en dirección 3' del gen que codifica la proteína de interés; y
- (e) un terminador de la nopalina sintasa situado en dirección 3' de la secuencia definida en (d), o que comprende:

- (a) un promotor;
- (b) una secuencia potenciadora de la expresión que tiene al menos una identidad del 70 % con los nucleótidos 1 a 507 del segmento del genoma de ARN-2 de la secuencia de CPMV del virus del mosaico del caupí mostrada en la Tabla 1, en donde el AUG en la posición 161 se ha mutado, situada en dirección 3' del promotor;
- (c) un gen que codifica una proteína de interés situada en dirección 3' de la secuencia definida en (b);
- (d) nucleótidos 3302 a 3481 de la secuencia del segmento del genoma de ARN-2 del virus del mosaico del caupí mostrada en la Tabla 1, situada en dirección 3' del gen que codifica la proteína de interés; y
- (e) un terminador de la nopalina sintasa situado en dirección 3' de la secuencia definida en (d).

9. Un sistema de expresión génica de acuerdo con una cualquiera de las reivindicaciones 4 a 8, en el que el virus de ARN bipartido es un CPMV, y el AUG en la posición 115 también está mutado.

10. Un sistema de expresión génica de acuerdo con una cualquiera de las reivindicaciones 3 a 7 que comprende:

- (a) una primera construcción génica que comprende una secuencia derivada de un ARN-2 truncado de un genoma de un virus bipartido *Comoviridae* que incluye al menos un gen extraño que codifica una proteína heteróloga de interés unida operativamente a secuencias promotoras y de terminación, en donde la construcción génica comprende el sitio de inicio mutado en dirección 5' del gen extraño, en donde el segmento del genoma del ARN-2 del virus *Comoviridae* codifica dos carboxiproteínas coterminales a través de dos sitios de inicio de traducción diferentes situados en el mismo marco de lectura del triplete, en donde el sitio de inicio mutado es el primero de estos dos sitios de inicio y corresponde al sitio de inicio en la posición 161 en el segmento de ARN-2 del CPMV de tipo salvaje; y opcionalmente,
- (b) una segunda construcción génica incorporada opcionalmente en dicha primera construcción génica que comprende un supresor heterólogo del silenciamiento génico unido operativamente a secuencias promotoras y de terminación.

11. Un sistema de expresión génica de acuerdo con una cualquiera de las reivindicaciones anteriores que es un vector binario de ADN.

12. Un proceso para aumentar la expresión o potenciar la actividad traduccional de una secuencia derivada de un segmento derivado del genoma de ARN-2 de un virus bipartido *Comoviridae*, que comprende la mutación de un sitio de inicio en el mismo, en donde el segmento del genoma del ARN-2 del virus *Comoviridae* codifica dos carboxiproteínas coterminales a través de dos sitios de inicio de traducción diferentes situados en el mismo marco de lectura del triplete, en donde el sitio de inicio mutado es el primero de estos dos sitios de inicio y corresponde al sitio de inicio en la posición 161 en el segmento de ARN-2 del CPMV de tipo salvaje, en donde dicha mutación se realiza mediante:

- (i) una mutación puntual en el sitio de inicio, o
- (ii) una delección de hasta 50 nucleótidos de longitud que incluye dicho sitio de inicio,

en donde dicha mutación potencia la expresión de un ORF heterólogo al que está unida la secuencia.

13. Un método para expresar una proteína heteróloga de interés en un organismo hospedador usando un sistema de expresión génica de acuerdo con una cualquiera de las reivindicaciones 1 a 10, en donde el organismo hospedador es un hospedador eucariota, que es opcionalmente una planta o un insecto.

14. Un método para potenciar la traducción de una proteína heteróloga de interés a partir de un gen o de un marco de lectura abierto (ORF) que codifica el mismo, que está unido operativamente a un segmento de genoma de ARN-2

de una secuencia derivada de un virus bipartido *Comoviridae*,
en donde el segmento del genoma del ARN-2 del virus *Comoviridae* codifica dos carboxiproteínas coterminales a
través de dos sitios de inicio de traducción diferentes situados en el mismo marco de lectura del triplete,
que comprende mutar el primero de estos dos sitios de inicio en la secuencia derivada de ARN-2, en donde el sitio
5 de inicio mutado corresponde al sitio de inicio en la posición 161 en el segmento de ARN-2 del CPMV de tipo
salvaje,
en donde dicha mutación se realiza mediante:

- 10 (i) una mutación puntual en el sitio de inicio, o
- (ii) una delección de hasta 50 nucleótidos de longitud que incluye dicho sitio de inicio.

15. Un método de acuerdo con la reivindicación 13, para expresar una proteína heteróloga en una planta, que
comprende las etapas de:

- 15 (a) introducir una construcción de expresión génica que comprende una primera construcción génica que
comprende una secuencia derivada de un ARN-2 truncado de un genoma de un virus bipartido *Comoviridae* que
incluye al menos un gen extraño que codifica una proteína heteróloga de interés unida operativamente a
secuencias promotoras y de terminación,
20 en donde la construcción génica comprende un sitio de inicio mutado en dirección 5' del gen extraño en una
célula vegetal, en donde el segmento del genoma del ARN-2 del virus *Comoviridae* codifica dos carboxiproteínas
coterminales a través de dos sitios de inicio de traducción diferentes situados en el mismo marco de lectura del
triplete,
en donde el sitio de inicio mutado es el primero de estos dos sitios de inicio y corresponde al sitio de inicio en la
posición 161 en el segmento de ARN-2 del CPMV de tipo salvaje; y opcionalmente,
25 (b) introducir una segunda construcción génica que opcionalmente se incorpora en dicha primera construcción
génica, que comprende un supresor del silenciamiento génico unido operativamente a secuencias promotoras y
de terminación en la célula vegetal.

16. Un organismo hospedador transfectado transitoriamente con, y que comprende, un sistema de expresión génica
30 de acuerdo con una cualquiera de las reivindicaciones 1 a 11, en donde el organismo hospedador es opcionalmente
una planta o una célula vegetal.

17. Un organismo hospedador transgénico transformado de forma estable con, y que comprende, un sistema de
35 expresión génica de acuerdo con una cualquiera de las reivindicaciones 1 a 11.

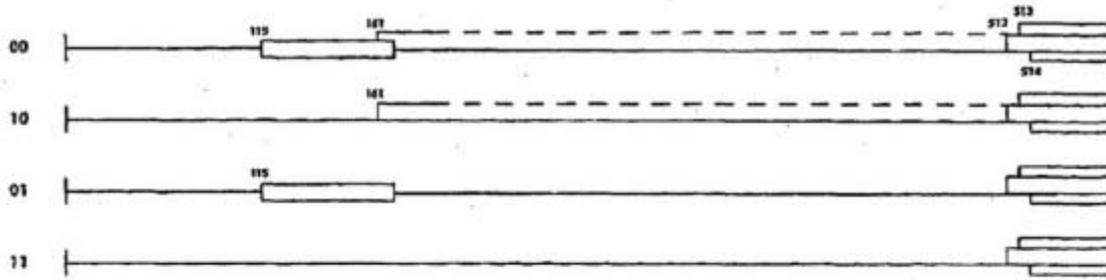


FIGURA 1

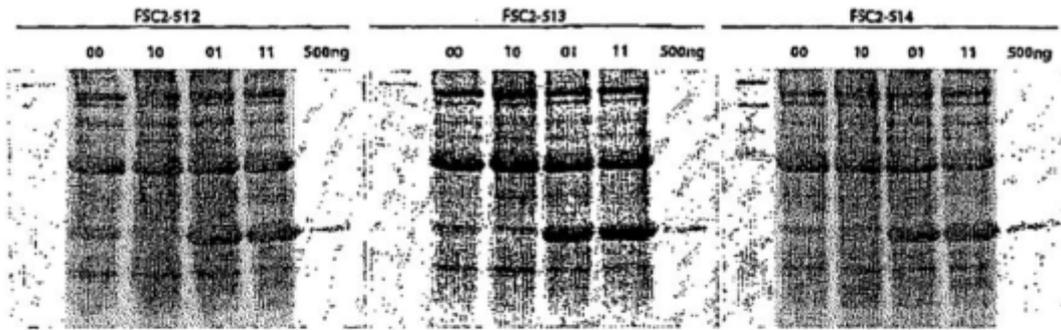


FIGURA 2

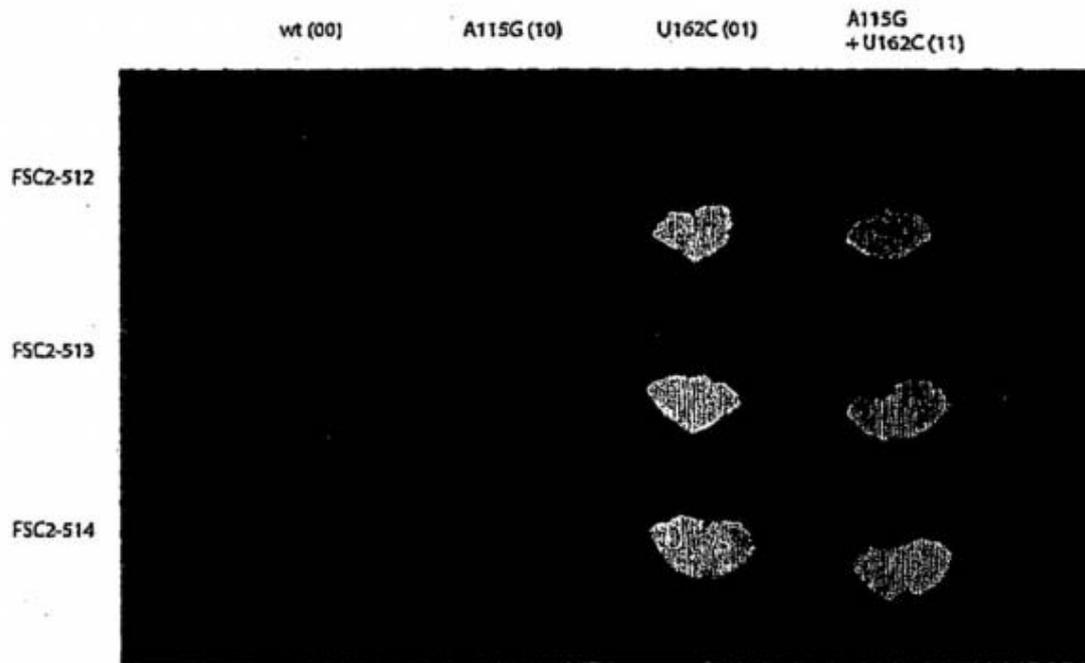


FIGURA 3

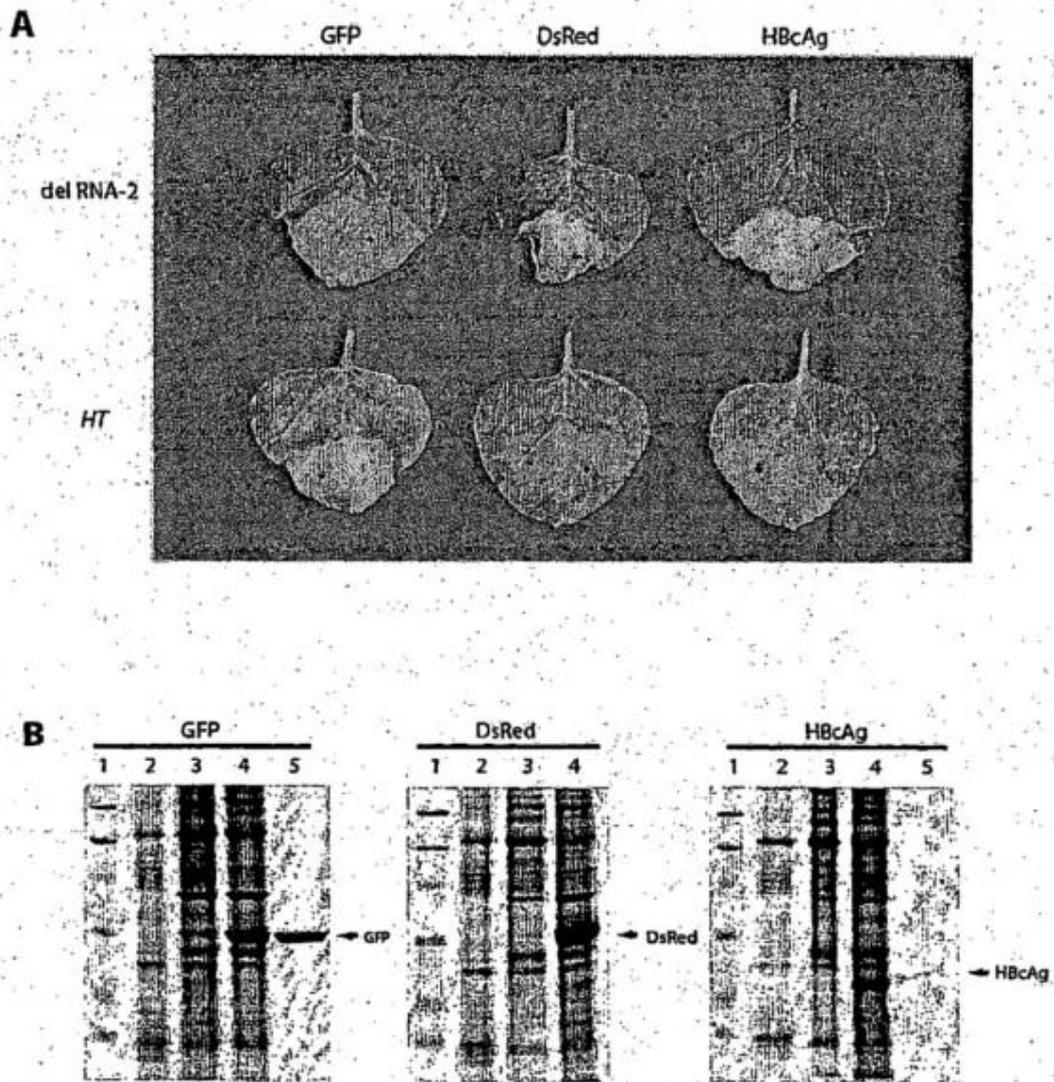


FIGURA 4

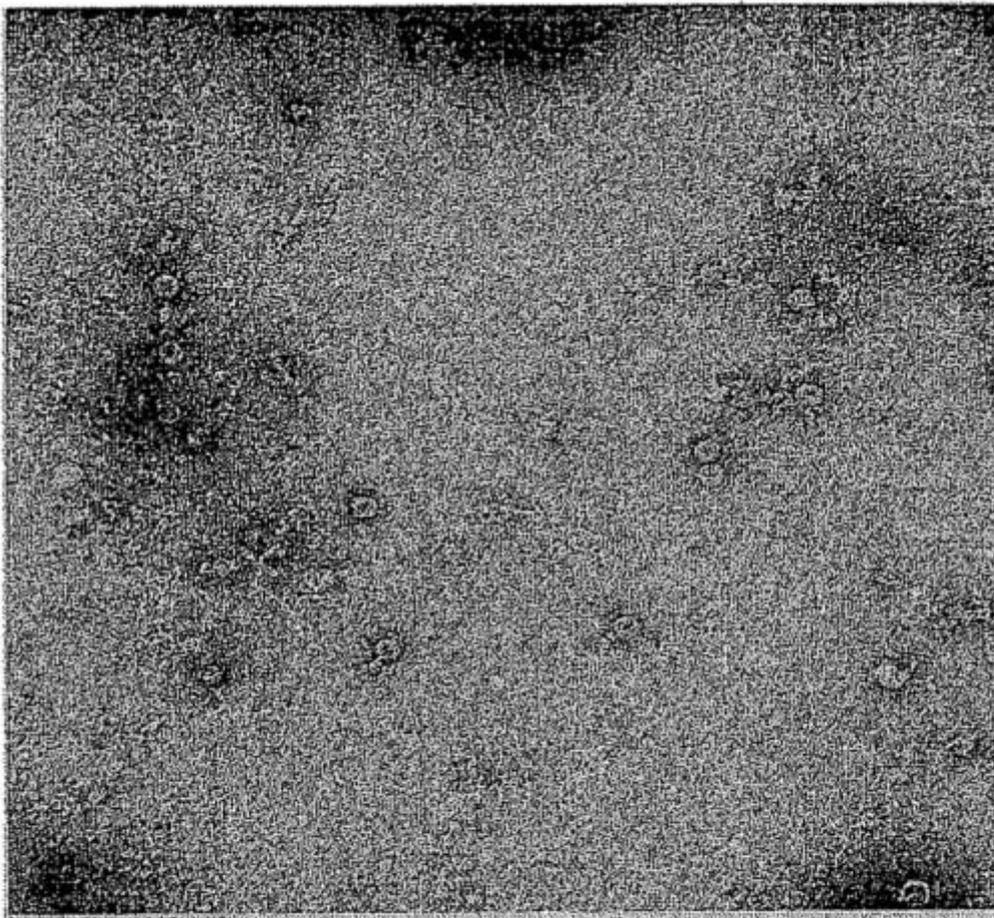


FIGURA 6

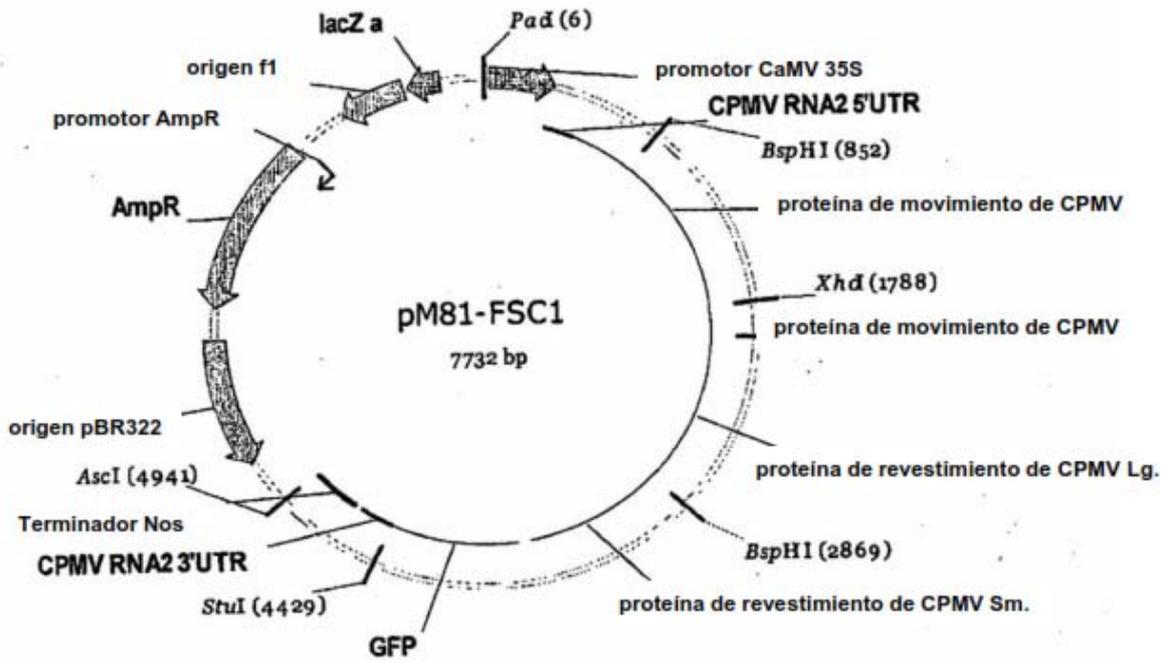


FIGURA 7

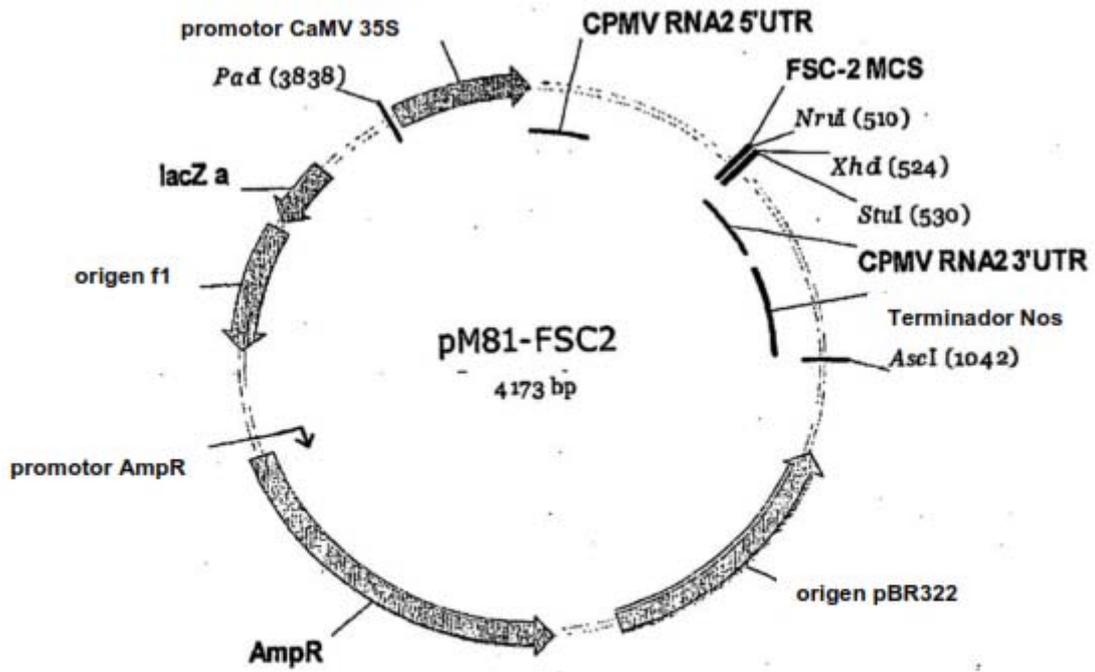


FIGURA 8

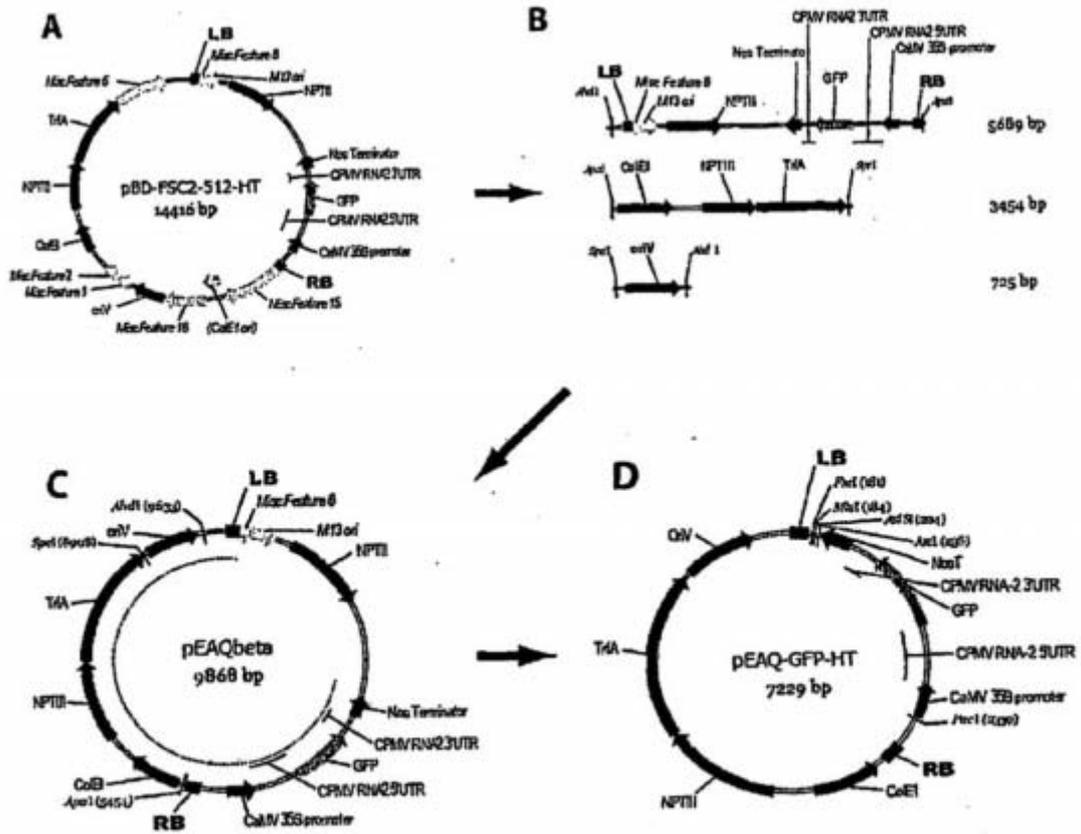


FIGURA 9

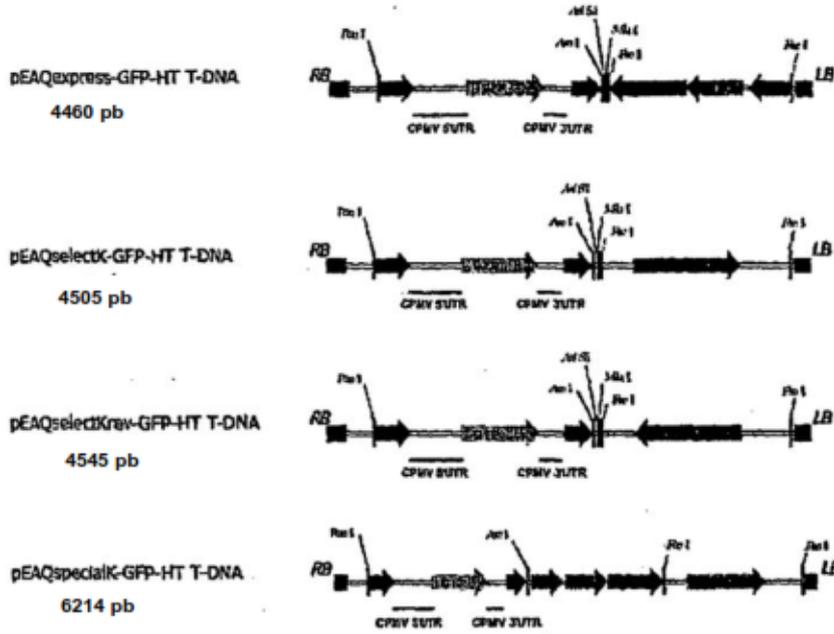


FIGURA 10

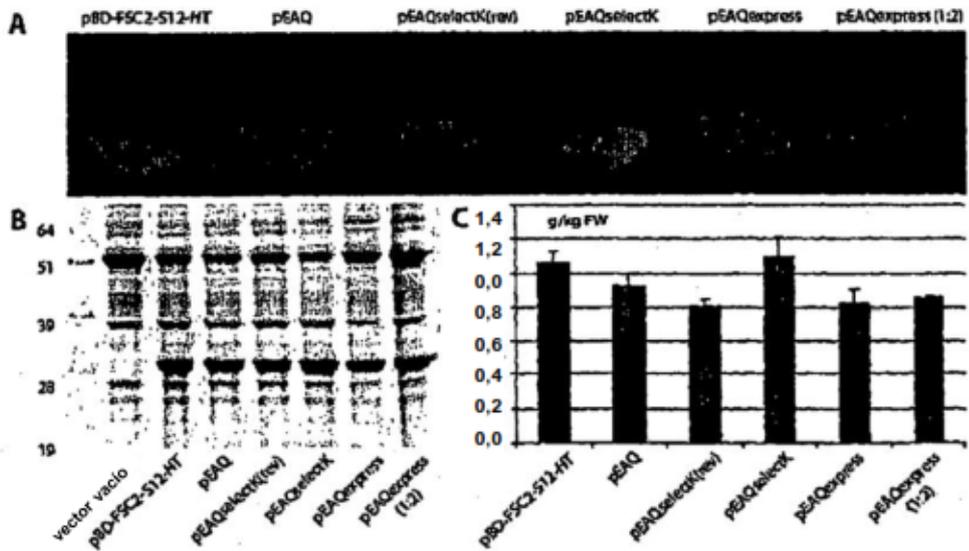


FIGURA 11

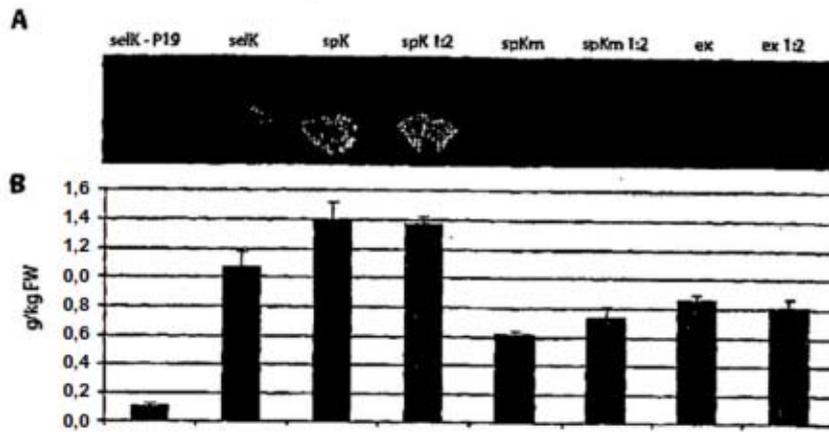


FIGURA 12

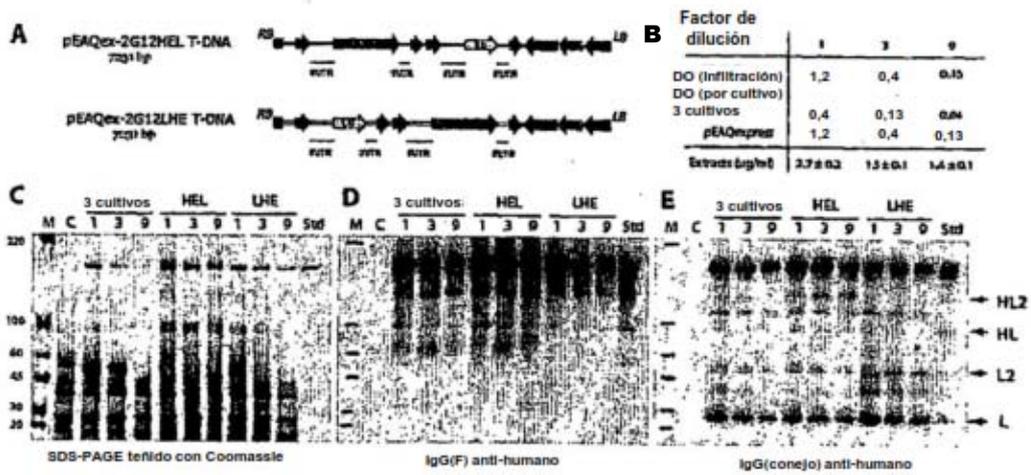


FIGURA 13

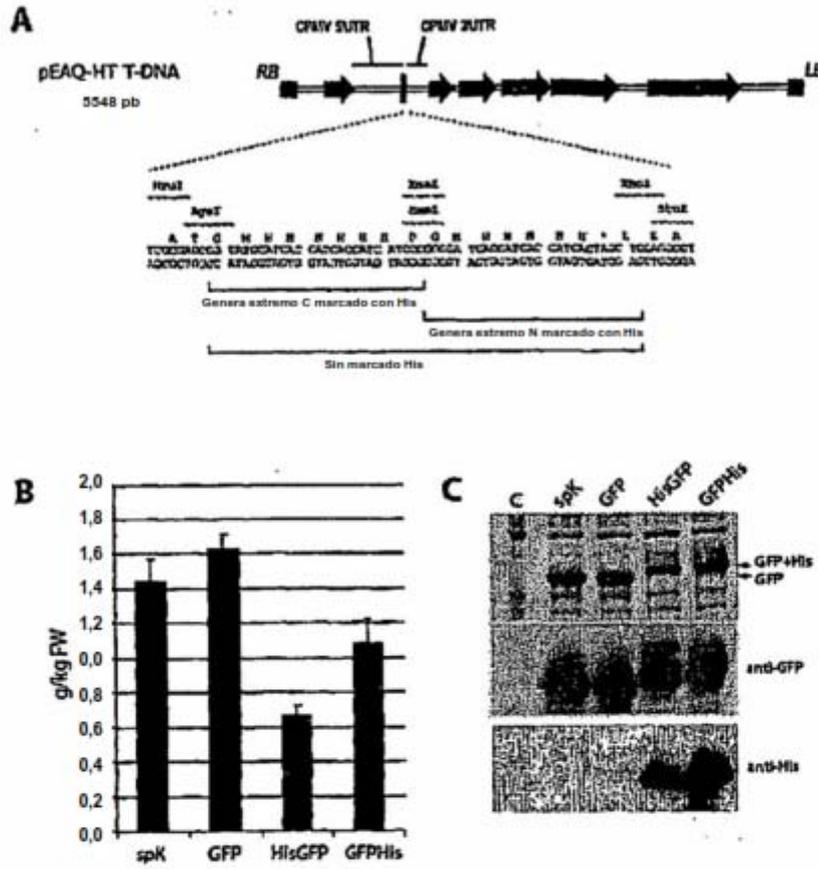


FIGURA 14

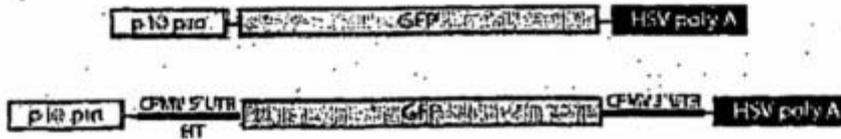


FIGURA 15