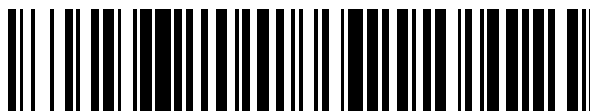


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 529 346**

51 Int. Cl.:

**C12N 15/70** (2006.01)

**C12N 15/61** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **15.09.2004 E 04767054 (2)**

97 Fecha y número de publicación de la concesión europea: **05.11.2014 EP 1664305**

54 Título: **Sistema de selección que contiene un marcador de selección que no es de resistencia a antibióticos**

30 Prioridad:

**15.09.2003 FI 20031319**

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

**19.02.2015**

73 Titular/es:

**FIT BIOTECH OY (100.0%)  
Biokatu 8  
33520 Tampere , FI**

72 Inventor/es:

**TENSON, TANEL;  
LAHT, SILJA;  
ADOJAAN, MAARJA;  
MÄNNIK, ANDRES;  
TOOTS, URVE y  
USTAV, MART**

74 Agente/Representante:

**DE ELZABURU MÁRQUEZ, Alberto**

**ES 2 529 346 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

Sistema de selección que contiene un marcador de selección que no es de resistencia a antibióticos

**Campo de la invención**

5 La presente invención se refiere a un nuevo sistema de selección, que se basa en el uso de un gen *araD* o una forma mutada de un gen *araD* introduciendo un Codón de parada en el Codón 8 del gen *araD* como marcador de selección y al uso de una cepa bacteriana carente del gen *araD*. La presente invención se refiere adicionalmente a nuevos vectores que contienen un gen *araD*, una forma mutada de un gen *araD* introduciendo de Codón de parada en el Codón 8 del gen *araD*, en donde a partir de dicho gen *araD* de dicho vector se expresa una L-ribulosa-5-fosfato 4-epimerasa activa y a nuevas cepas bacterianas carentes de un gen *araD*. La presente invención se refiere adicionalmente a un método de selección de las células transformadas con un plásmido, que contiene el gen de interés.

**Antecedentes de la invención**

15 Un requisito esencial para la modificación genética eficaz de bacterias y otras células propagadas en cultivos celulares es la capacidad para seleccionar las células con una alteración genotípica específica. La estrategia de selección más común en la tecnología del ADN recombinante es incluir un marcador de selección en el vector de clonación o plásmido. Un marcador de selección puede ser un gen clonado o una secuencia de ADN, que permite la separación de las células anfitrionas que contienen el marcador de selección de las que no lo contiene. El marcador de selección junto con un medio de selección adecuado mantiene el vector de clonación en las células. Por otra parte, puesto que la replicación de los plásmidos es una carga energética para el anfitrión bacteriano, en un medio de cultivo las bacterias, que han perdido el plásmido, tendrían una ventaja de crecimiento sobre las células con el plásmido.

25 Para la mayoría de los propósitos, un gen de resistencia a antibióticos es un marcador de selección comúnmente utilizado. Sin embargo, para la producción de agentes terapéuticos recombinantes, donde el objetivo es generar un producto, tal como una vacuna de ADN, con un alto rendimiento para su administración a pacientes, el uso de genes de resistencia a antibióticos presenta problemas: la propagación de patógenos resistentes a los antibióticos es un grave problema mundial [Levy, S. B., J. Antimicrob. Chemother. 49 (2002) 25-30]. Por lo tanto los genes de resistencia a antibióticos no pueden tener un amplio uso en la industria farmacéutica, y, por ejemplo, de acuerdo con las regulaciones de la Administración de Alimentos y Fármacos de los Estados Unidos, no están permitidos genes de resistencia a antibióticos en las vacunas de ADN experimentales que entran en la tercera fase.

30 Alternativamente, se han sugerido sistemas de selección libres de antibióticos. Tales sistemas de selección libres de antibióticos incluyen sistemas bacterianos de toxina-antitoxina [Engelberg-Kulka, H. y Glaser, G., Annu Rev Microbiol 53 (1999) 43-70], genes responsables de la resistencia contra metales pesados, tales como el telurio [Silver, S. y Phung, L. T., Annu Rev Microbiol 50 (1996) 753-789], y sistemas en los que el plásmido codifica un gen que complementa una auxotrofia del anfitrión [Wang, M. D., et al., J. Bacteriol. 169 (1987) 5610-5614].

35 La Solicitud de Patente de los Estados Unidos Núm. 2000/0014476 A1 describe generalmente, entre otras cosas, el uso de un marcador de selección distinto de antibiótico, que puede ser un gen cuyo producto es necesario para el metabolismo de la célula bajo ciertas condiciones de cultivo, tal como un gen de catabolismo, lo que hace posible que la célula asimile una cierta sustancia presente en el medio de cultivo (fuente de carbono o nitrógeno específica) etc. No se proporcionan ejemplos específicos de tales genes adecuados. Este enfoque no es necesariamente aplicable a la producción comercial, ya que la eliminación de un componente esencial, tal como un aminoácido o una fuente de carbono, del medio de crecimiento reduce el rendimiento, lo cual no es deseable. Además, la manipulación del medio de crecimiento en términos de omisión de un nutriente esencial puede aumentar considerablemente el coste del medio de crecimiento, ya que las mezclas nutrientes disponibles comercialmente deben ser reemplazadas por nutrientes individuales.

45 Se ha construido un sistema de selección eficaz sobre la base de los genes *araD/araC* [Ariza, R. R., et al., Carcinogenesis 14 (1993) 303-305]. Sin embargo, este sistema de selección se ha utilizado en los estudios sobre los mecanismos de mutagénesis pero no se ha utilizado antes como marcador de selección para el mantenimiento del plásmido. Ariza *et al.* utilizaron una cepa donde el gen *araC* contiene un codón de terminación y el gen *araD* está inactivado. Se introdujo en el plásmido un producto del gen *supF*, que codifica un supresor de ARNt. En presencia de ARNt supresor activo, se produjo el producto enzimáticamente activo a partir de *araC* causando la detención del crecimiento celular (puesto que *araD* era inactivo). Este sistema permite estudiar la supresión de mutaciones por ARNt de *supF*: en el caso de que *supF* sea inactivado por mutación, las células pueden crecer sobre arabinosa. Por lo tanto, este sistema de selección se basa en el gen *araC* y no en el gen *araD*. *araD* no se introdujo en un plásmido, ni el sistema se diseñó o caracterizó para fines de producción del plásmido.

55 Para fines terapéuticos comerciales sería ventajoso utilizar un gen, que no fuera esencial para el crecimiento del anfitrión, pero cuya manipulación aún afecte al crecimiento en circunstancias seleccionadas. Además, en vista del

5 uso terapéutico, sería ventajoso utilizar un gen, cuya delección conduzca a la acumulación de compuestos, que son tóxicos para la célula anfitriona pero no tóxicos para los mamíferos, incluyendo seres humanos. También sería ventajoso utilizar genes más pequeños, lo que a su vez permitiría la construcción de plásmidos más pequeños para los cuales el consumo de energía para la replicación sería menor y por lo tanto se mejorarían la tasa de crecimiento de cultivo bacteriano y el rendimiento de plásmidos.

**Breve descripción de la invención**

El objeto de la presente invención es proporcionar un nuevo sistema de selección libre de antibióticos, que evita los problemas de los sistemas de selección previamente descritos para su uso en la producción de productos terapéuticos recombinantes.

10 Otro objeto de la invención es proporcionar un nuevo sistema de selección libre de antibióticos, que se puede utilizar de forma segura en la producción de productos terapéuticos recombinantes en términos del medio ambiente y la seguridad del paciente.

15 Un objeto adicional de la invención es proporcionar un nuevo sistema de selección libre de antibióticos, que se puede utilizar rentablemente en la producción de productos terapéuticos recombinantes utilizando medios de crecimiento convencionales.

Un objeto adicional más de la invención es proporcionar un nuevo sistema de selección libre de antibióticos, que proporciona un aumento de la tasa de crecimiento y rendimiento mejorado.

20 Otro objeto adicional de la presente invención es proporcionar un vector novedoso que contiene un marcador de selección, que es no tóxico para el medio ambiente y para los seres humanos y que es capaz de un mantenimiento a largo plazo en el anfitrión.

Otro objeto adicional de la presente invención es proporcionar una nueva célula anfitriona que contiene un defecto génico, que no es peligrosa para el medio ambiente.

Otro objeto más de la presente invención es proporcionar un método para la selección de células que portan un gen de interés para la producción de productos terapéuticos recombinantes.

25 Se encontró sorprendentemente que los objetos de la presente invención se cumplen mediante el uso del gen *araD* o una forma mutada de un gen *araD* que introduce un codón de parada en el codón 8 del gen *araD* como marcador de selección y el uso de un anfitrión bacteriano específico carente del gen *araD*.

30 Por consiguiente, la presente invención proporciona un nuevo sistema de selección libre de antibióticos que comprende una célula bacteriana carente de un gen *araD* al que se ha añadido un vector que porta un gen *araD* como marcador de selección en donde a partir de dicho gen *araD* de dicho vector se expresa una L-ribulosa-5-fosfato 4-epimerasa activa. Una realización de la presente invención se refiere a un sistema de selección en donde el gen *araD* es el gen *araD* o la L-ribulosa-5-fosfato 4-epimerasa (EC 5.1.3.4.). Otra realización de la presente invención se refiere a un sistema de selección en donde el gen *araD* es mutado mediante la introducción de un codón de parada en el codón 8 del gen *araD*.

35 La presente invención proporciona adicionalmente nuevos vectores, que contienen un gen *araD* o una forma mutada de un gen *araD* mediante la introducción de un codón de parada en el codón 8 del gen *araD* como marcador de selección en donde a partir de dicho gen *araD* de dicho vector se expresa una L-ribulosa-5-fosfato 4-epimerasa activa.

La presente invención proporciona adicionalmente nuevas cepas bacterianas, que carecen del gen *araD*.

40 La presente invención proporciona adicionalmente un método de selección de células transformadas con un plásmido, que contiene 1) el gen *araD* o una forma mutada de un gen *araD* mediante la introducción de un codón de parada en el codón 8 del gen *araD*, como marcador de selección y 2) el gen de interés, comprendiendo el método insertar dicho plásmido en la célula anfitriona carente de *araD* y cultivar las células en un medio de crecimiento que contiene arabinosa.

**45 Dibujos**

La Figura 1 muestra el uso de arabinosa como fuente de carbono por las células de *E. coli* (Lin, 1987).

50 La Figura 2 muestra el mapa de S6wtd1 EGFP. Las secuencias codificantes para d1EGFP, E2 y marcador de resistencia a kanamicina aminoglucósido-3'-O-fosfotransferasa (*kana*) se indican con flechas. Las características adicionales se indican mediante casillas compactas: 10E2BS - diez sitios de unión a E2 de BPV con alta afinidad; CMV-tk - promotor inmediato temprano de citomegalovirus humano y secuencia líder del gen Th de HSV; intrón - intrón del gen de la beta-globina conejo con sitios SD y SA optimizados; tkpa - señal de poliadenilación del gen Tk

de HSV; RSV LTR - repetición terminal larga del virus del sarcoma de Rous; bgh pA - señal poliadenilación del gen de la hormona de crecimiento bovina; pUCori - origen de replicación bacteriano derivado del plásmido pUC18.

La Figura 3 muestra el mapa de S6wtd1EGFP*kana/araD*1. Las secuencias codificantes para d1EGFP, E2, marcador de resistencia a kanamicina aminoglucósido-3'-O-fosfotransferasa (*kana*) y L-ribulosa-5-fosfato 4-epimerasa (*araD*) se indican mediante flechas. Las características adicionales se indican mediante casillas compactas: 10E2BS - diez sitios de unión a E2 de BPV con alta afinidad; CMV-tk - promotor inmediato temprano de citomegalovirus humano y secuencia líder del gen Th de HSV; intrón - intrón del gen de la beta-globina de conejo con sitios SD y SA optimizados; tkpa - señal de poliadenilación del gen Tk de HSV; RSV LTR - repetición terminal larga del virus del sarcoma de Rous; bgh pA - señal de poliadenilación del gen de la hormona de crecimiento bovina; pUCori - origen de replicación bacteriano derivado del plásmido pUC18.

La Figura 4 muestra el mapa de S6wtd1EGFP*kana/araD*2. Las secuencias codificantes para d1EGFP, E2, marcador de resistencia a kanamicina aminoglucósido-3'-O-fosfotransferasa (*kana*) y L-ribulosa-5-fosfato 4-epimerasa (*araD*) se indican mediante flechas. Las características adicionales se indican mediante casillas compactas: 10E2BS - diez sitios de unión a E2 de BPV con alta afinidad; CMV-tk - promotor inmediato temprano de citomegalovirus humano y secuencia líder del gen Th de HSV; intrón - intrón del gen de la beta-globina de conejo con sitios SD y SA optimizados; tkpa - señal de poliadenilación del gen Tk de HSV; RSV LTR - repetición terminal larga del virus del sarcoma de Rous; bgh pA - señal de poliadenilación del gen de la hormona de crecimiento bovina; pUCori - origen de replicación bacteriano derivado del plásmido pUC18.

La Figura 5 muestra el mapa de S6wtd1EGFP/*araD*1. Las secuencias codificantes para d1EGFP, E2 y L-ribulosa-5-fosfato 4-epimerasa (*araD*) se indican mediante flechas. Las características adicionales se indican mediante casillas compactas: 10E2BS - diez sitios de unión a E2 de BPV con alta afinidad; CMV-tk - promotor inmediato temprano de citomegalovirus humano y secuencia líder del gen Th de HSV; intrón - intrón del gen de la beta-globina de conejo con sitios SD y SA optimizados; tkpa - señal de poliadenilación del gen Tk de HSV; RSV LTR - repetición terminal larga del virus del sarcoma de Rous; BGH pA - señal de poliadenilación del gen de la hormona de crecimiento bovina; pUCori - origen de replicación bacteriano derivado del plásmido pUC18.

La Figura 6 muestra el mapa de S6wtd1EGFP/*araD*2. Las secuencias codificantes para d1EGFP, E2 y L-ribulosa-5-fosfato 4-epimerasa (*araD*) se indican mediante flechas. Las características adicionales se indican mediante casillas compactas: 10E2BS - diez sitios de unión a E2 de BPV con alta afinidad; CMV-tk - promotor inmediato temprano de citomegalovirus humano y secuencia líder del gen Th de HSV; intrón - intrón del gen de la beta-globina de conejo con sitios SD y SA optimizados; tkpa - señal de poliadenilación del gen Tk de HSV; RSV LTR - repetición terminal larga del virus del sarcoma de Rous; bgh pA - señal de poliadenilación del gen de la hormona de crecimiento bovina; pUCori - origen de replicación bacteriano derivado del plásmido pUC18.

Las Figuras 7A y 7B muestran el análisis electroforético del ADN plasmídico de S6wtd1EGFP/*araD*1 (7A) y S6wtd1EGFP/*araD*2 (7B) extraído de la cepa de *E.coli* AG1delta *araD* cultivada en diferentes medios.

La Figura 8 muestra el análisis de patrón de restricción del ADN plasmídico de S6wtd1EGFP/*araD*1 y S6wtd1EGFP/*araD*2 extraído de la cepa de *E.coli* AG1delta *araD*.

La Figura 9 muestra el análisis electroforético de S6wtd1EGFP/*araD*2 en el análisis de estabilidad.

Las Figuras 10A y 10B muestran el análisis de patrón de restricción de S6wtd1EGFP/*araD*2 en el análisis de estabilidad.

La Figura 11 muestra los parámetros de crecimiento de la fermentación de alimentación discontinua de AG1Δ*araD* S6wtd1EGFP/*araD*2 medidos y registrados durante la fermentación. Las abreviaturas son las siguientes: sPump = velocidad de alimentación; pO<sub>2</sub> = concentración de oxígeno; Temp = Temperatura de crecimiento; mys = tasa de crecimiento deseada; DO = densidad óptica a 600 nm.

La Figura 12 muestra el esquema de la lisis y purificación de AG1Δ*araD* S6wtd1EGFP/*araD*2.

La Figura 13 muestra la secuencia del locus *araD* del clon # 13.

La Figura 14 muestra el mapa del plásmido p3hCG.

La Figura 15 muestra el mapa del plásmido *paraDMgB*.

La Figura 16 muestra el mapa del plásmido p3*araD*1hCG.

La Figura 17 muestra el mapa del plásmido p3*araD*2hCG.

La Figura 18 muestra los resultados del análisis de sensibilidad a L-arabinosa de cepas de *E. coli* con *araD* desorganizado.

La Figura 19 muestra los resultados del análisis de sensibilidad a L-arabinosa en M9 y medio de extracto de levadura con diferentes concentraciones de glucosa y arabinosa.

La Figura 20 muestra el mapa del plásmido p2 MG C # 11.

La Figura 21 muestra el mapa del plásmido *paraD* MG C # 145.

5 La Figura 22 muestra el fragmento genómico de *E. coli* que contiene el gen *sgbE*.

La Figura 23 muestra el fragmento genómico de *E. coli* que contiene el Gen *ulaF*.

### Descripción detallada de la invención

10 La presente invención se basa en un esfuerzo por encontrar un sistema de selección libre de antibióticos alternativo, que pueda ser utilizado en la producción de productos terapéuticos recombinantes para administrarlos *in vivo*, especialmente en la producción de vacunas de ADN. Sorprendentemente, se encontró que el gen *araD* implicado en la ruta de las pentosas fosfato tanto de organismos procarióticos como eucarióticos, tales como mamíferos, incluyendo seres humanos, se puede utilizar satisfactoriamente como marcador de selección en una célula anfitriona auxótrofa para el plásmido. El uso de la auxotrofia tiene la ventaja de no implicar un uso o generación de sustancias tóxicas que posteriormente podrían contaminar la preparación de plásmido.

15 El gen *araD* codifica una enzima que es responsable de la epimerización de la ribulosa-5-fosfato a xilulosa-5-fosfato (Fig. 1) y por lo tanto permite el uso de arabinosa en la ruta de las pentosas fosfato [Engelsberg, E., et al., J. Bacteriol. 84: (1962) 137-146]. Si *araD* se inactiva, la ribulosa-5-fosfato se acumula en la célula bacteriana conduciendo a la detención del crecimiento.

20 Si la copia cromosómica de *araD* se inactiva en la célula anfitriona y se inserta en el plásmido una copia intacta del gen *araD* o una forma mutada del gen *araD* mediante la introducción de un codón de parada en el codón 8 del gen *araD*, la ventaja de crecimiento de las células que contienen el plásmido en medio que contiene L-arabinosa se logra como resultado de dos efectos. En primer lugar, las células que contienen el plásmido pueden utilizar arabinosa como fuente de carbono, y en segundo lugar, no se acumula ribulosa-5-fosfato tóxica. Esto permite el uso de medios de crecimiento ricos con suplementos de arabinosa. En los medios de crecimiento ricos las células de *E. coli* crecen rápido y el rendimiento de plásmido es alto. Se pueden utilizar componentes convencionales de bajo coste de los medios de crecimiento bacteriano, tales como extracto de levadura, como fuente de aminoácidos. Las trazas de la ribulosa-5-fosfato que teóricamente podrían contaminar la preparación de plásmidos no son un problema, cuando la preparación se administra *in vivo*, puesto que la ribulosa-5-fosfato puede ser metabolizada por las células humanas eficazmente y no es tóxica.

30 El uso de una forma mutada del gen *araD* mediante la introducción de un codón de parada en el codón 8 del gen *araD* ofrece ventajas concretas. Los sistemas de selección de la invención que comprenden una célula bacteriana carente de un gen *araD* en la que un vector que porta tal forma mutada del gen *araD* como marcador de selección producen una concentración óptima del producto del gen *araD* L-ribulosa-5-fosfato 4-epimerasa para proporcionar un rápido crecimiento no inhibido de las bacterias. Se obtienen ventajas similares mediante el uso de sistemas de selección que contienen un vector que porta un gen *araD* intacto pero que comprenden deleciones o mutaciones en otros lugares en el locus del gen *araD*.

35 El sistema de selección de la invención comprende 1) un vector que porta un gen *araD* o una forma mutada del gen *araD* mediante la introducción de un codón de parada en el codón 8 del gen *araD* como marcador de selección y 2) una cepa bacteriana específica carente del gen *araD* a la que se ha añadido el vector. Cuando el anfitrión específico carente del gen *araD* se cultiva en presencia de arabinosa, las únicas células supervivientes son las que contienen el vector, que contiene un gen *araD* o una forma mutada del gen *araD* mediante la introducción de un codón de parada en el codón 8 del gen *araD*.

40 En el sistema de selección de la invención se puede emplear cualquier vector de expresión comúnmente utilizado en la producción de productos terapéuticos, por medio de lo cual se inserta en el vector el gen *araD* o una forma mutada del gen *araD* mediante la introducción de un codón de parada en el codón 8 del gen *araD* utilizando métodos generalmente conocidos en la técnica. En el presente contexto, el gen *araD* comprende preferiblemente la secuencia identificada mediante el SEQ ID NO. 1, mediante el SEQ ID NO. 19, o una secuencia hibridable a la misma. Sin embargo, también se contemplan genes *araD* aplicables cualesquiera. En el presente contexto, el término "un fragmento catalíticamente activo del gen *araD*" es cualquier fragmento del gen que codifica un polipéptido o una proteína susceptible de epimerización de L-ribulosa-5-fosfato a D-xilulosa-5-fosfato. En una realización específica de la invención, el gen *araD* se inserta en el vector susceptible de un mantenimiento a largo plazo y por lo tanto es susceptible de proporcionar una expresión estable del antígeno o los antígenos deseados.

50 En otra realización específica de la invención, una forma mutada de un gen *araD* mediante la introducción de un codón de parada en el codón 8 del gen *araD* se inserta en el vector susceptible de un mantenimiento a largo plazo y

por lo tanto susceptible de proporcionar una expresión estable del antígeno o los antígenos deseados.

En una realización específicamente preferida de la invención, el vector utilizado es un vector de expresión que comprende:

- 5 (a) una secuencia de ADN que codifica una proteína de anclaje nuclear conectada operativamente a un promotor heterólogo, comprendiendo dicha proteína de anclaje nuclear (i) un dominio de unión a ADN que se une a una secuencia específica de ADN, y (ii) un dominio funcional que se une a un componente nuclear, o un equivalente funcional del mismo; y
- (b) una secuencia de ADN multimerizada que forma un sitio de unión para la proteína de anclaje nuclear, en donde dicho vector carece de un origen de la replicación del virus del papiloma, y
- 10 (c) un gen *araD* o una forma mutada de un gen *araD* mediante la introducción de un codón de parada en el codón 8 del gen *araD*.

Tales vectores se han descrito en detalle en la solicitud de patente internacional WO02/090558, que se incorpora a la presente memoria como referencia.

15 Lo más preferiblemente, el vector utilizado en el método de selección de la presente invención es un vector de expresión que comprende:

- (a) la proteína E2 del Virus del Papiloma Bovino tipo 1 (BPV), y
- (b) múltiples sitios de unión de la proteína E2 de BPV incorporados al vector como una agrupación, donde los sitios pueden ser estructuras de cabeza a cola o pueden ser incluidos en el vector mediante posicionamiento espaciado, en donde dicho vector carece de un origen de la replicación del virus del papiloma, y
- 20 (c) el gen *araD*.

En el sistema de selección de la invención, en principio, podría emplearse cualquier anfitrión conocido carente del gen *araD* y adecuado para su uso en la producción de productos terapéuticos. En el presente contexto, el término "carente" denota un anfitrión, en el que el gen *araD* es totalmente suprimido o inactivado mediante cualquier método conocido.

En una realización preferida de la invención, se utiliza una cepa de *Escherichia coli*, preferiblemente las cepas de *E. coli* DH5alfa-T1, AG1 o JM109 asequibles comercialmente, de las cuales se ha suprimido el gen *araD* con métodos generalmente conocidos, tales como los descritos a continuación en los ejemplos. En otra realización preferida de la invención, se utiliza una cepa *E. coli*, preferiblemente la cepa de *E. coli* DH5alfa-T1, AG1 o JM109, en la que se han realizado delecciones combinadas para el agotamiento de los otros genes que codifican proteínas con actividad L-ribulosa-5-fosfato 4-epimerasa. Alternativamente, se pueden emplear cepas de *E. coli* disponibles comercialmente, preferiblemente cepas de *E. coli* DH5alfa-T1, AG1 o JM109, en las que el gen *araD* y/u otros genes que codifican proteínas con actividad L-ribulosa-5-fosfato 4-epimerasa se han inactivado mediante cualquier método conocido. En el método para la selección de células que portan un gen de interés para la producción de productos terapéuticos recombinantes, el gen de interés se inserta en las células anfitrionas carentes de un gen *araD* y/u otros genes que codifican proteínas con actividad L-ribulosa-5-fosfato 4-epimerasa utilizando un método bien conocido en la técnica y las células se cultivan en un medio de crecimiento que contiene arabinosa en un medio de cultivo y en las condiciones adecuadas para el anfitrión en cuestión.

Se puede utilizar cualquier medio de crecimiento adecuado para el cultivo de células de *E. coli*. Para la producción comercial, el medio de crecimiento será naturalmente optimizado en términos de rendimiento. Los ejemplos de los medios de crecimiento adecuados son los medios de crecimiento disponibles comercialmente, tales como M9 y LB (disponible de varios fabricantes, tales como Fermentas, Lituania). La cantidad de arabinosa añadida en el medio de crecimiento no es crítica pero, naturalmente, debe estar presente arabinosa en una cantidad que sea suficiente para el periodo de cultivo total. Se ha encontrado que una cantidad tan baja como 0,1% es suficiente para la selección. Típicamente la arabinosa se añade al medio en una cantidad de aproximadamente 0,1% a aproximadamente 2,0%, preferiblemente en una cantidad de aproximadamente 0,2% a aproximadamente 1,0%, más preferiblemente 0,2% a aproximadamente 0,5%. Sin embargo se observa el efecto de la L-arabinosa a concentraciones tan bajas como 0,01% y la L-arabinosa se puede añadir hasta 5% en el medio de crecimiento. En una realización especial, cuando se utiliza L-arabinosa como agente de selección y como fuente de carbono limitada, 0,2% de L-arabinosa es una cantidad adecuada para añadirla al medio de crecimiento.

El sistema de selección de la invención es adecuado para su uso en cualquier sistema de expresión. Es especialmente adecuado para su uso en la expresión de productos terapéuticos recombinantes, tales como vacunas de ADN, destinadas a ser utilizadas *in vivo*, ya que se evitan los problemas asociados con el uso de genes de

resistencia a antibióticos. Asimismo, el sistema de selección de la invención es adecuado para su uso en la producción de proteínas recombinantes.

5 La posible contaminación de arabinosa en el producto final resultante del proceso de preparación es intrascendente, ya que la arabinosa es el azúcar comestible contenido en los alimentos de forma natural y como aditivo y por tanto no es tóxico para los mamíferos incluyendo los seres humanos.

10 Además, el gen *araD* tiene un tamaño más pequeño que los genes de resistencia a antibióticos comúnmente utilizados contra, por ejemplo, ampicilina y tetraciclina y tiene un tamaño similar a los genes de resistencia a kanamicina y cloranfenicol. Esto proporciona una ventaja adicional, ya que permite la construcción de plásmidos pequeños para los que el consumo de energía para la replicación es menor que para grandes plásmidos. De este modo, se incrementan tanto la tasa de crecimiento de cultivo bacteriano como el rendimiento de plásmido.

La presente invención puede entenderse mejor mediante la referencia a los siguientes Ejemplos no limitantes, que se proporcionan como ilustrativos de la invención. Los siguientes ejemplos se presentan con el fin de ilustrar más completamente las realizaciones preferidas de la invención. No se deben considerar de ninguna manera, sin embargo, limitantes del amplio alcance de la invención.

15 Ejemplo 1

Clonación de plásmidos de selección de *araD*

20 Para la clonación de constructos de selección de *araD* se utilizó el plásmido S6wtd1EGFP (Figura 2). Tiene un origen de replicación pMB1 y un marcador de resistencia a kanamicina como elementos funcionales de la cadena principal del plásmido. La resistencia a la kanamicina en este plásmido es conferida por el gen que deriva del transposón Tn903 de *E. coli*.

El gen *araD* se amplificó utilizando la reacción en cadena de la polimerasa (PCR) a partir del cromosoma *DH5α* de *E. coli* de acuerdo con el procedimiento convencional. El producto de la PCR se clonó en plásmidos seleccionados en dos orientaciones diferentes con los pares de cebadores s6araDL1 + s6araDR1 o s6araDL1 + s6araDR1, generando productos denominados *araD1* y *araD2*, respectivamente:

25 s6araDL1:

CGCCATGGTTCTCATGTTTGACAGCTTATCATCGATAAGCTTTAATGCGGTAGTTTAGCACGAAGGAGT  
CAACATG (SEQ ID NO. 2);

s6araDR1:

30 CGCCATGGACTAGTAAAAAAGCCCGCTCATTAGGCGGGCTGTCATTACTGCCCGTAATATGC (SEQ  
ID NO. 3);

s6araDL2:

CGCCATGGACTAGTTCTCATGTTTGACAGCTTATCATCGATAAGCTTTAATGCGGTAGTTTAGCACGAA  
GGAGTCAACATG (SEQ ID NO. 4);

35 s6araDR2:

CGCCATGGAAAAAAGCCCGCTCATTAGGCGGGCTGTCATTACTGCCCGTAATATGC (SEQ ID NO. 5);

Los cebadores se diseñaron de manera que el promotor P2 del plásmido pBR322 (utilizado para accionar el gen de resistencia a tetraciclina en pBR322) y la secuencia de terminación del operón *trp* de *E. coli* se añadieron durante la PCR aguas arriba y aguas abajo de la secuencia codificante de *araD*, respectivamente.

40 Los productos de la PCR de 814 y 815 pb se clonaron en el vector pUC18 linealizado con HincII (Fermentas, Lituania) y las secuencias correctas se verificaron mediante secuenciación utilizando cebadores de secuenciación universales

M13F22: GCCAGGGTTTTCCAGTCACGA (SEQ ID NO. 6) y

M13R24: GAGCGGATAACAATTTACACAGG (SEQ ID NO. 7) y cebadores específicos de *araD*

45 *araD* F311: CCAACTCACGGCTGCTCTATC (SEQ ID NO. 8),

*araD* F614: AATGCCGAAGATGCGGTGCATAAC (SEQ ID NO. 9),

*araD* R700: TAACTGCGGCGCTAACTGAC (SEQ ID NO. 10), y

*araD* R421: GGTGGCTGGAATCGACTGAC (SEQ ID NO. 11).

Las mutaciones en secuencias amplificadas fueron reparadas mediante recombinación de diferentes clones.

Para clonar *araD* en S6wtd1EGFP, el vector se linealizó mediante digestión parcial con la enzima de restricción *PagI* (posición 4761) (Fermentas, Lituania) y los extremos 5' del ADN se desfosforilaron con Fosfatasa Alcalina de Intestino de Ternera (CIAP; Fermentas, Lituania). Los fragmentos *araD1* y *araD2* se separaron por corte de pUC18 con *NcoI* (Fermentas, Lituania) y se ligaron a S6wtd1EGFP/PAGL.

Ambas mezclas de ligación se transformaron en células competentes de *E. coli DH5α* y se sembraron en placas que contenían medio LB que contenía 50 µg/ml de kanamicina y se incubaron a 37°C durante la noche. Las colonias se analizaron primero con PCR de colonias, después de lo cual el ADN se aisló y se digirió con diferentes enzimas de restricción.

La clonación dio como resultado los plásmidos S6wtd1EGFP*kana/araD1*, S6wtd1EGFP*kana/araD2*, que se muestran en las Figuras 3 y 4.

Para eliminar el gen marcador de resistencia a la kanamicina de los plásmidos, S6wtd1EGFP*kana/araD1* y S6wtd1EGFP*kana/araD2* se digirieron con la endonucleasa de restricción *BcuI* (Fermentas, Lituania) y se autoligó un fragmento del vector de 6473 pb.

Las mezclas de ligación se transformaron en una cepa de *E. coli AG1 ΔaraD* (véase el Ejemplo 3) y se cultivaron en placas que contenían medio M9 con un suplemento de L-arabinosa al 2% y se incubaron a 37°C durante 36 horas. Las colonias se analizaron primero con PCR de colonias, después de lo cual el ADN se aisló y se digirió con diferentes enzimas de restricción. La clonación dio como resultado los plásmidos S6wtd1EGFP/*araD1*, S6wtd1EGFP/*araD2*, respectivamente, mostrados en las Figuras 5 y 6.

Las colonias bacterianas que contenían S6wtd1EGFP/*araD1* y S6wtd1EGFP/*araD2* se cultivaron en dos medios diferentes: LB con un suplemento de L-arabinosa al 2,5% y M9 con un suplemento de L-arabinosa al 0,2% a 37°C con sacudimiento vigoroso. Las células se cosecharon y el ADN del plásmido se extrajo de las células utilizando el kit QIAprep Spin Miniprep (Qiagen) y se analizaron mediante electroforesis en gel de agarosa (Figuras 7A y 7B, respectivamente).

Las muestras de ADN plasmídico de los cultivos en los medios LB y M9 se analizaron mediante electroforesis en gel de agarosa antes y después de la digestión con la endonucleasa de restricción *PagI* (Fermentas, Lituania), (Figura 8). Los tamaños pronosticados de los fragmentos obtenidos en la digestión con *PagI* tenían 3954 y 2519 pb para S6wtd1EGFP/*araD1* y 4315 y 2157 pb para S6wtd1EGFP/*araD2*. El ADN de *lambda* digerido con *Eco91I* (M15 en la Figura 8C) y el ADN de *lambda* digerido con *EcoRI/HindIII* (Fermentas, Lituania) (M3 en la Figura 8C) se utilizaron como marcadores de peso molecular. Todos los clones bacterianos analizados contenían el plásmido correcto en el análisis con enzimas de restricción, pero el rendimiento de ADN fue muy bajo cuando los plásmidos se cultivaron en medio LB. Dos de los clones bacterianos analizados de cuatro clones S6wtd1EGFP/*araD2* (Núm. 13 y Núm. 14 en la Figura 8B) tenían mayor tasa de crecimiento cuando se cultivaron en medio M9 con un suplemento de L-arabinosa al 0,2% (Figuras 7 y 8), lo que dio como resultado un mayor rendimiento de plásmido por cultivo.

Se realizó un análisis adicional de estos dos clones con la mejora del crecimiento. Estos dos plásmidos tenían la misma estructura que los otros plásmidos según se juzgó mediante el análisis de restricción. Los plásmidos se extrajeron de las bacterias y se caracterizaron adicionalmente mediante secuenciación del locus del gen *araD*. La secuencia de locus *araD* del clon Núm. 13 (SEQ ID NO. 18; SEQ ID NO. 19) indicó que la secuencia codificante del gen *araD* porta un codón de parada en lugar de un codón para Glutamina en la posición 8 de la L-ribulosa-5-fosfato 4-epimerasa. Esta mutación resultó de la sustitución de la Citidina en el codón 8 de la L-ribulosa-5-fosfato 4-epimerasa (secuencia codificante de *araD*) (5'-CAG-3') por Timidina, dando como resultado un codón de parada (5'-TAG-3'). El plásmido que porta una mutación en tal gen *araD* proporcionó efectivamente la capacidad de crecer en el medio selectivo en presencia de L-arabinosa, aunque la secuencia codificante contiene el codón de PARADA. Se ha demostrado que el codón de PARADA UAG es leído eficazmente por los ribosomas de *Escherichia coli*, cuando tal PARADA está al principio de la secuencia codificante [para la referencia, véase la revisión de Murgola, E. J., Annu. Rev. Genet. 19 (1985) 57-80]. Sin vincularse a la teoría, los autores de la presente invención plantearon la hipótesis de que el alto rendimiento del plásmido, que es una indicación de un rápido crecimiento no inhibido de las bacterias, requiere una concentración óptima del producto del gen *araD* L-ribulosa-5-fosfato 4-epimerasa.

El análisis de la secuencia del locus *araD* Núm. 14 del clon indicó que la secuencia codificante *araD* es perfecta como se pronosticó. Sin embargo, se observaron reordenamientos de secuencia cerca del promotor *araD* que abarca los sitios de unión de la proteína E2 (véase la Figura 13, SEQ ID NO. 18). Estos datos sugieren además que tales reordenamientos cerca del promotor pueden dar como resultado la regulación a la baja de la actividad del promotor, por lo tanto el nivel del producto *araD*.

Ejemplo 2



Clonación de plásmidos de selección de *araD* mutado

Para la clonación de los constructos de selección de *araD* mutado el plásmido p3hCG (Figura 14) que porta resistencia a la kanamicina [transposón Tn5 derivado del gen marcador de resistencia a kanamicina (neo)] se escindió con las endonucleasas de restricción BcuI y HindIII, los extremos se rellenaron utilizando el Fragmento de Klenow (Fermentas, Lituania) y el fragmento con el tamaño de 4647 pb se purificó a partir del gel después de la electroforesis en gel de agarosa. El origen de replicación pMB1 y la secuencia *araD* que porta la mutación C a T, que da como resultado un codón de parada en la posición 8 de la secuencia codificante del gen *araD*, se escindió del plásmido *paraDMgB* (Figura 15) con las endonucleasas de restricción BcuI y Eco52I, los extremos se rellenaron utilizando el Fragmento de Klenow (Fermentas, Lituania), y el extremo 5' del ADN se desfosforiló con fosfatasa alcalina de intestino de ternera (CIAP; Fermentas, Lituania). El fragmento con el tamaño de 1532 pb se purificó a partir del gel después de la electroforesis en gel de agarosa y se ligó con el fragmento de 4647 pb obtenido anteriormente. La cepa de *Escherichia coli* AG1 carente de *araD* se transformó con esta mezcla de ligación y se cultivó en placas de agar que contenían medio M9 selectivo con extracto de levadura al 0,5%, L-arabinosa al 2% y 25 µg/ml de kanamicina. Las colonias se inspeccionaron 24 horas después del cultivo en placa y mostraron que el tamaño de las colonias era uniforme. Los plásmidos se extrajeron de las bacterias y se caracterizaron adicionalmente secuenciando el locus del gen *araD*.

La clonación dio como resultado plásmidos p3*araD*1hCG y p3*araD*2hCG, que se muestran en las Figuras 16 y 17, respectivamente. De acuerdo con el análisis de la secuencia, las bacterias contenían plásmidos no reordenados con la mutación C a T en el codón 8 (p3*araD*1hCG; Figura 16; p3*araD*2hCG, Figura 17).

Cuando este experimento se repitió con la secuencia de tipo salvaje y las placas transformadas se inspeccionaron 24 horas después de la transformación, el resultado fue diferente. Se observaron dos tipos de colonias: en primer lugar, las colonias de gran tamaño, y las pequeñas colonias, que tenían un retraso en el crecimiento. El análisis de la secuencia de estos plásmidos indicó que la secuencia codificante del gen *araD* porta un codón de parada en lugar de un codón para glutamina (plásmido Núm. 3A, *araD*2) o la mutación se había producido en la secuencia de Shine-Dalgarno en el sitio de unión ribosomal (AGGAG fue reemplazado por AGTAG) (plásmido Núm. 2A, *araD*2). El plásmido Núm. 7 (*araD*1) tenía la secuencia correcta en todas las regiones del locus del gen *araD*, sin embargo, las bacterias crecieron muy lentamente y dieron como resultado un rendimiento de plásmido 10 veces menor cuando se cultivaron en medios líquidos.

Ejemplo 3

Construcción de cepas de *Escherichia coli*  $\Delta$ *araD* sensibles a arabinosa.

Se utilizaron tres cepas de *E. coli*, DH5alfa T1, AG1 y JM109, para la construcción de mutantes  $\Delta$ *araD*. El gen *araD* en el genoma de *E. coli* se desorganizó utilizando el método descrito por Datsenko y Wanner [PNAS 97 (2000) 6640-6645]. Este método explota un sistema de recombinación Red del fago  $\lambda$ . En pocas palabras, la estrategia de este sistema es reemplazar una secuencia cromosómica con un gen de resistencia a antibióticos seleccionable que se genera por medio de PCR utilizando cebadores con extensiones de homología. Esto se logra mediante la recombinación mediada por Red en estas homologías flanqueantes.

Para la transformación de pKD46 (Datsenko y Wanner, *supra*), que codifica el sistema de recombinación del fago  $\lambda$ , *E. coli*, las células se hicieron químicamente competentes utilizando disoluciones RF1 y RF2:

RF1 100 ml

RbCl 1	1,2 g
MnCl <sub>2</sub> · 4H <sub>2</sub> O	0,99 g
KAc 1 M pH 7,5	3 ml
CaCl <sub>2</sub> · 2H <sub>2</sub> O	0,15 g
Glicerol	15 g
pH 5,8	(Añadir CH <sub>3</sub> COOH)

40

RF2 100 ml

## ES 2 529 346 T3

MOPS 0,5 M	2 ml
RbCl	0,12 g
CaCl <sub>2</sub> · 2H <sub>2</sub> O	1,1 g
Glicerol	15 g
pH 6,8	(Añadir NaOH)

5 Las células se cultivaron en 2 ml de medio LB a DO<sub>600</sub> de 0,2-0,5. El cultivo se centrifugó y el sedimento se resuspendió en 1 ml de RF1. La mezcla se mantuvo en hielo durante 10 min y se centrifugó. El sedimento se suspendió en 100 µl de RF2 y la suspensión se mantuvo en hielo durante 30-45 min. Se añadieron aproximadamente 50 ng de pKD43 y las células se mantuvieron en hielo durante 30 min adicionales seguido de choque térmico de 5 min a 37°C. Después de la incubación durante 10 min en hielo se añadieron 900 µl de medio SOB a las células transformadas y la mezcla se incubó a 37°C durante una hora. Las células se cultivaron en medio LB que contenía ampicilina (100 µg/ml). Las colonias se recogieron de las placas de transformación y se cultivaron en 2 ml del mismo medio a DO<sub>600</sub> de aproximadamente 1 y se elaboraron provisiones de partida de glicerol (2 ml de cultivo + 0,6 ml de glicerol al 50%). Las provisiones de partida se almacenaron a -80°C.

10 Para la desorganización del gen *araD* se generó un producto de PCR lineal que contenía un gen de resistencia a la kanamicina. El plásmido pKD13 (Datsenko y Wanner, PNAS vol. 97, núm. 12, Junio de 2000) se utilizó como molde para la PCR. Los cebadores utilizados fueron *ara* (pr1) y *ara* (pr4):

*ara* (pr1)

15 5'-CTCAAACGCCAGGTATTAGAAGCCAACCTGGCGCTGCCAAAACACGTGTAG GCTGGAGCTGCTTC-3'  
(SEQ ID NO. 12)

*ara* (pr4)

5'-GGCCCGATCACAAAGACGCCGCGCTCGCGATCAACGGCGCATTCCGGGGAT CCGTCGACC-3'  
(SEQ ID NO. 13)

20 Estos cebadores tienen las secuencias del complemento con pKD13 para la hibridación en la PCR y con el gen *araD* para la recombinación homóloga.

La mezcla de reacción de PCR fue la siguiente: tampón nativo PFU (5 µl), dNTP 10 mM (5 µl), cebador *ara* (pr1) 10 µM (1 µl), cebador *ara* (pr4) 10 µM (1 µl), 100 ng de pKD13 (2 µl), DMSO (4 µl), PFU 2,5 U (1 µl), y agua mQ hasta 50 µl.

25 El procedimiento de PCR fue el siguiente: desnaturalización 45 s, 96°C, hibridación 45 s, 50°C, síntesis 2 min 30 s, 72°C, 25 ciclos. El producto de PCR obtenido fue de 1,4 kb.

30 Se realizaron cinco reacciones de forma simultánea; el ADN se purificó a partir de gel de agarosa al 2% utilizando el kit de purificación Ultrapure (MoBio Laboratories Inc.) y se eluyó con 60 µl de agua. El ADN se concentró con precipitación con etanol y se disolvió en 5 µl de agua. La concentración final fue de 0,6 µg/µl. Se utilizó una alícuota de 1,5 µl en una electroporación.

35 El producto de PCR se sometió a electroporación en células de *E. coli* DH5alfa T1 pKD46, AG1 pKD46 (Datsenko y Wanner, *supra*), y JM109 pKD46. En primer lugar, a 200 ml de medio YENB que contenía 10 mM de L-arabinosa para la inducción del sistema de recombinación y 100 µg/ml de ampicilina se les inoculó un cultivo de una noche de células de *E. coli* DH5alfa T1 pKD46, AG1 pKD46, y JM109 pKD46. Los cultivos se hicieron crecer a 30°C a DO<sub>600</sub> 0,8 (DH5alfa T1 y JM109) y 0,6 (AG1). Las bacterias se recogieron mediante centrifugación a 4.000 g durante 10 min a 4°C, se lavaron dos veces con 20 ml de agua estéril y una vez con 20 ml de agua estéril que contenía glicerol al 10%. Las células se suspendieron en 300 µl de agua que contenía glicerol al 10%. Se utilizaron 40 µl de células competentes en una electroporación.

40 La electroporación se realizó con BioRad *E. coli* Pulser utilizando cubetas de 0,2 cm y 2,5 kV. El producto purificado mediante PCR (1,5 µl) se añadió a las células competentes, se mantuvo en hielo durante 1 min, e inmediatamente después de la electroporación, se añadieron 2 ml de medio SOB caliente a las células y la mezcla se incubó a 37°C durante 1 hora. Las células se cultivaron en placa en medio LB que contenía kanamicina (25 µg/ml). Se utilizaron

100 pg de plásmido grande resistente a la kanamicina (GTU-MultiHIV clado C) como control positivo, no se añadió plásmido al control negativo. La eficacia de transformación fue de  $10^6$  para AG1 y  $10^7$  para JM109 para el control positivo. No hubo colonias en la placa de control negativo, se obtuvieron 215 colonias en la placa JM109 + producto de PCR, 70 colonias en la placa AG1 + producto de PCR y 50 colonias en la placa DH5alfa T1 + producto de PCR.

5 Ejemplo 4

Ensayo de las cepas de *E. coli* DH5alfa T1  $\Delta$ araD, AG1 $\Delta$ araD y JM109 $\Delta$ araD

Las colonias obtenidas a partir de la electroporación como se describe en el Ejemplo 2 se sometieron a ensayo para determinar la presencia del gen de resistencia a kanamicina mediante PCR de colonias utilizando cebadores *araVlisF* (5' CGGCACGAAGGAGTCAACAT 3'; SEQ ID NO. 14) *AraVlisR* y (5' TGATAGAGCAGCCGGTGAGT 3'; SEQ ID NO. 15) que contenían sitios de hibridación sobre el gen *araD* cerca del sitio de inserción. Se esperaba un producto de PCR de 272 pb de las cepas de *E. coli* DH5alfa T1, AG1 y JM109 sin inserción en *araD* y un producto de 1545 pb, si el producto de PCR se había insertado en el gen *araD*. Se verificaron tres colonias de DH5alfa T1  $\Delta$ araD, nueve colonias de AG1 $\Delta$ araD y 14 colonias de JM109 $\Delta$ araD de 15 y cada una proporcionó el producto de 1545 pb. Por lo tanto, se concluyó que estas cepas contenían la inserción del gen de resistencia a kanamicina.

15 Para confirmar la inserción del gen de la kanamicina se realizó otra PCR de colonias utilizando cebadores *kanaSF* (5' TCAGATCCTTGCGGCAAGA 3'; SEQ ID NO. 16) y *araVR* (5' TGTAAATCGACGCCGAAGGT 3'; SEQ ID NO. 17). Estos cebadores producen un producto de 435 pb, si el gen de resistencia a la kanamicina ha sido insertado en el gen *araD*. Se sometieron a ensayo seis colonias de las cepas AG1 $\Delta$ araD y JM109 $\Delta$ araD y tres colonias de las cepas DH5alfa T1  $\Delta$ araD y todas proporcionaron el producto correcto.

20 Seis colonias de AG1 $\Delta$ araD y JM109 $\Delta$ araD, y tres colonias de DH5alfa T1  $\Delta$ araD se cultivaron en placa en medio LB que contenía 25  $\mu$ g/ml de kanamicina y se incubaron a 37°C durante la noche para eliminar el plásmido pKD46, que tenía un origen de replicación sensible a la temperatura. Las células se sometieron a ensayo para determinar la sensibilidad a la ampicilina por réplica de cultivo en placas sobre medio LB y medio LB que contenía ampicilina. Ninguna crecía sobre el medio que contenía ampicilina, y se concluyó que las bacterias no contenían más plásmido pKD46.

25 La sensibilidad a la arabinosa se sometió a ensayo sobre las cepas AG1 $\Delta$ araD y JM109 $\Delta$ araD producidas. Una colonia de AG1 $\Delta$ araD y una colonia de JM109 $\Delta$ araD se inocularon cada una en 2 ml de LB. Los cultivos se hicieron crecer durante 8 horas, se diluyeron 1:100 en medio M9 que contenía glicerol al 0,2%, 25  $\mu$ g/ml de kanamicina, tiamina al 0,01% (prolina al 0,05% para JM109 $\Delta$ araD) y se añadieron diferentes concentraciones de L-arabinosa al medio de crecimiento. Los cultivos se hicieron crecer durante la noche a 37°C en una incubadora con movimiento oscilante y se midió la DO<sub>600</sub> (Tabla 1).

Tabla 1. Ensayos de sensibilidad a la arabinosa.

L-arabinosa%	DO <sub>600</sub> de AG1 $\Delta$ araD	DO <sub>600</sub> de JM109 $\Delta$ araD
0	3,2	1,9
0,1	0,03	0,03
0,2	0,030	0,026
0,5	0,030	0,020
1	0,024	0,025
2	0,017	0,021

35 Como se puede observar en la Tabla 1, una cantidad tan baja como 0,1% de L-arabinosa es suficiente para inhibir el crecimiento de las cepas  $\Delta$ araD de la invención.

La sensibilidad a la arabinosa se sometió a ensayo adicionalmente en AG1 $\Delta$ araD, DH5alfaT1  $\Delta$ araD y JM109 $\Delta$ araD como antes pero utilizando concentraciones más bajas de L-arabinosa. Los resultados se proporcionan en la Figura 18. Como se puede observar en la Figura 18, una cantidad tan baja como 0,0005% de L-arabinosa es suficiente para inhibir el crecimiento de las cepas  $\Delta$ araD de la invención.

Adicionalmente, la sensibilidad a la L-arabinosa se sometió a ensayo en medio M9 y de extracto de levadura con diferentes concentraciones de glucosa y arabinosa (0,2% de glucosa, 0,2% de arabinosa, 2% de arabinosa). Los cultivos se incubaron a 37°C en una incubadora con agitador durante la noche. A continuación se midió la DO<sub>600</sub> para cuantificar la densidad celular. Los resultados se proporcionan en la Figura 19.

- 5 Ambas concentraciones de arabinosa (0,2% y 2%) inhibieron el crecimiento de las cepas  $\Delta araD$  de la invención. Sin embargo, no se inhibió el crecimiento de las cepas con el gen *araD* intacto.

Adicionalmente, se sometió a ensayo el rendimiento de ADN plasmídico de las cepas  $\Delta araD$ . El plásmido S6wtd1EGFP<sub>araD2</sub> preparado en el Ejemplo 1 se transformó en las cepas AG1 $\Delta araD$  y JM109 $\Delta araD$ . Las células competentes se prepararon con las disoluciones RF1 y RF2 como se ha descrito en el Ejemplo 3.

- 10 Las colonias de las placas de transformación se inocularon en 2 ml de medio M9 que contenía extracto de levadura al 0,5 y 25 µg/ml de kanamicina + tiamina al 0,01% + L-arabinosa (2% y 0,2%).

Los cultivos se incubaron a 37°C durante 17 horas. A continuación se midió la DO<sub>600</sub> para cuantificar la densidad de las células y el ADN plasmídico se extrajo con Qiagen Miniprep Kit. Se utilizó un coeficiente de 2,8 (DO<sub>600</sub>/ml) para el aislamiento de miniprep para obtener resultados comparables. Los resultados se muestran en la Tabla 2.

- 15 La concentración de ADN se midió con un espectrofotómetro como la DO a 260 nm. Para el análisis microscópico se aplicó una gota de cultivo bacteriano sobre el portaobjetos de vidrio y se cubrió con un cubreobjetos. El cultivo se inspeccionó visualmente a un aumento de 100x con un objetivo en inmersión en aceite.

Tabla 2. Rendimiento de ADN plasmídico de cepas  $\Delta araD$

Cepa	L-arabinosa (%)	DO <sub>600</sub>	Conc. de ADN plasmídico (µg/µl)	Rendimiento de ADN plasmídico (µg por ml de cultivo)	Apariencia al microscopio
AG1 $\Delta araD$	2	7,6	0,039	5,3	no hay filamentos
AG1 $\Delta araD$	0,2	5,8	0,057	5,9	no hay filamentos
JM109 $\Delta araD$	2	4,9	0,043	3,8	muy pocos filamentos
JM109 $\Delta araD$	0,2	4,3	0,038	2,9	muy pocos filamentos
DH5 $\alpha$ T1 $\Delta araD$	2	6,6	0,017	3,5	no hay filamentos
DH5 $\alpha$ T1 $\Delta araD$	0,2	6,4	0,016	3,4	no hay filamentos

- 20 De acuerdo con estos resultados la L-arabinosa al 0,2% es suficiente para obtener el número de copias del plásmido al mismo nivel que con arabinosa al 2%.

Para este plásmido AG1 $\Delta araD$  parece ser mejor, debido a que el rendimiento de plásmido es algo mayor y las densidades celulares también.

Ejemplo 5.

- 25 Generación de una cepa de Escherichia coli con mutaciones adicionales dentro de los genes que codifican potencialmente L-ribulosa-5-fosfato 4-epimerasa.

El cromosoma de *E. coli* contiene dos secuencias codificantes adicionales para L-ribulosa-5-fosfato 4-epimerasas en diferentes operones. Los genes *ulaF* y *sgbE* de la ruta de degradación de L-ascorbato codifican los genes con actividad epimerasa (Wen Shan Yew, Jhon A. Gerlt, J. Bacteriol. 184 (2002) 302-306. Con el fin de aumentar la rigurosidad de la selección y para evitar o eliminar los posibles mecanismos de adaptación de las cepas de *E. coli* debidos a otros genes con actividad epimerasa, se interrumpieron las secuencias codificantes de los genes *UlaF* y *SgbE* en el genoma de *E. coli*. Tales mecanismos de adaptación podrían ocurrir en la producción de plásmido a

30

largo plazo en condiciones adecuadas.

Los genes *UlaF* y *SgbE* en las cepas de *E. coli* DH5 $\alpha$ T1 $\Delta$ *araD* y AG1 $\Delta$ *araD* se desorganizaron utilizando el Sistema de recombinación Red del fago  $\lambda$  como se ha descrito en el Ejemplo 3.

5 En primer lugar, se eliminó el gen resistente a la kanamicina en las cepas de *E. coli* AG1 $\Delta$ *araD* y DH5 $\alpha$ T1 $\Delta$ *araD*. El plásmido de expresión de la recombinasa FLP pKD20 (Datsenko y Wanner, *supra*) es resistente a la ampicilina y sensible a la temperatura. Los mutantes resistentes a la kanamicina se transformaron con pCP20 (el gen resistente a la kanamicina está flanqueado por FRT), y los transformantes resistentes a la ampicilina se seleccionaron a 30°C (48 horas), después de lo cual las mismas colonias se purificaron no selectivamente a 42°C (24 horas dos veces). A continuación se sometieron a ensayo para determinar la pérdida de resistencias a kanamicina y ampicilina.

10 La inactivación del gen *ulaF* cromosómico (SEQ ID NO. 20) por el sistema de recombinación Red del fago  $\lambda$  se realizó utilizando los cebadores *ulaFylem* y *ulaFalum*:

*ulaFylem*

CAGCAGGTATTTGAAGCCAACATGGAGCTGCCGCGCTACGGGCTGGTGTAGGCTGGAGCTGCTTC  
(SEQ ID NO. 21)

15 *ulaFalum*

AAACGGCTGCGGAATTAGACCAGTTATCTCCCAGGAAGGAAATTAATCCGGGGATCCGTCGACC  
(SEQ ID NO. 22)

20 Se observó una gran cantidad de colonias en ambas placas de transformación. Se sometieron a ensayo quince colonias obtenidas a partir de la electroporación para determinar la presencia del gen de resistencia a la kanamicina mediante PCR de colonias utilizando cebadores *ulaFvalisR* y *ulaFvalisF*:

*ulaFvalisR*

AAACGGCTGCGGAATTAGACC (SEQ ID NO. 23)

*ulaFvalisF*

GCCGTACCTGATTGAGATGTGGAG (SEQ ID NO. 24)

25 Estos cebadores contienen sitios de hibridación sobre el gen *UlaF* cerca del sitio de inserción. Se esperaba un producto de PCR de 864 pb a partir de las cepas de *E. coli* DH5 $\alpha$ T1 $\Delta$ *araD* y AG1 $\Delta$ *araD* sin inserción en *UlaF* y un producto de 1527 pb, si el producto de PCR se había insertado en el gen *UlaF*. Para confirmar la inserción del gen de kanamicina se realizó otra PCR de colonias utilizando los cebadores *ulaFvalisR* (SEQ ID NO 23) y *kanaSF* (SEQ ID NO 16).

30 Estos cebadores producen un producto de 428 pb, si el gen de resistencia a la kanamicina ha sido insertado en el gen *UlaF*. Se sometieron a ensayo cuatro colonias de las cepas AG1 $\Delta$ *araD* $\Delta$ *ulaF* y DH5 $\alpha$ T1 $\Delta$ *araD* $\Delta$ *ulaF* y todas proporcionaron el producto correcto. Se utilizó adicionalmente una colonia de cada cepa.

35 La eliminación del gen resistente a la kanamicina en las cepas de *E. coli* AG1 $\Delta$ *araD* $\Delta$ *ulaF* y DH5 $\alpha$ T1 $\Delta$ *araD* $\Delta$ *ulaF* se realizó como se ha descrito anteriormente. La inactivación del gen cromosómico *sgbE* (SEQ ID NO. 25) por el sistema de recombinación Red del fago  $\lambda$  se realizó como se describe en el Ejemplo 3. Los cebadores utilizados fueron *sgbEalum* y *sgbEylem*:

*sgbEalum*

CGTTACAGCAAGGAACATATCAATTCGTTAGTGCCGGGGCGATGAAGAATCCGGGGATCCGTCGACC  
(SEQ ID NO 26)

40 *sgbEylem*

GCAGGAGGCTGGATTTATATGTTAGAGCAACTGAAAGCCGACGTGGTGTAGGCTGGAGCTGCTTC (SEQ ID NO 27)

45 Se observó una gran cantidad de colonias en ambas placas de transformación. Quince colonias obtenidas a partir de la electroporación se sometieron a ensayo para determinar la presencia del gen de resistencia a kanamicina mediante PCR de colonias utilizando los cebadores *sgbEvalisR* y *sgbEvalisF*:

*sgbEvalisR*

CGGCGTTACAGCAAGGAACATATC (SEQ ID NO. 28)

sgbE<sub>evalisF</sub>

ATTGAAGCGCGTATGCAGGAGG (SEQ ID NO. 29)

5 Se esperaba un producto de PCR de 792 pb a partir de las cepas de *E. coli* DH5 $\alpha$ T1 $\Delta$ araD $\Delta$ ulaF $\Delta$ sgbE y AG1 $\Delta$ araD $\Delta$ ulaF $\Delta$ sgbE sin inserción en *SgbE* y un producto de 1413 pb, si el producto de PCR se había insertado en el gen *SgbE*. Para confirmar la inserción del gen de la kanamicina se realizó otra PCR de colonias utilizando cebadores sgbE<sub>evalisR</sub> (SEQ ID NO. 28) y kanaSF (SEQ ID NO. 16):

Se sometieron a ensayo quince colonias de ambas cepas y cuatro proporcionaron el producto correcto.

10 La sensibilidad a la arabinosa se sometió a ensayo en las cepas de *E. coli* DH5 $\alpha$ T1 $\Delta$ araD $\Delta$ ulaF $\Delta$ sgbE y AG1 $\Delta$ araD $\Delta$ ulaF $\Delta$ sgbE producidas y se comparó con la de las cepas de *E. coli* DH5 $\alpha$ T1 $\Delta$ araD y AG1 $\Delta$ araD. Una colonia de cada cepa se inoculó en 2 ml de medio M9 que contenía extracto de levadura al 0,5%, 25  $\mu$ g/ml de kanamicina, solo glucosa al 0,2% o L-arabinosa al 0,2% o 2%, respectivamente. Los resultados se muestran en la Tabla 3.

Tabla 3. Ensayo de sensibilidad a la arabinosa

Cepa	DO <sub>600</sub> Glc	DO <sub>600</sub> Glc + L-arabinosa al 0,2%	DO <sub>600</sub> Glc + L-arabinosa al 2%
AG1 $\Delta$ araD	7,3	0,82	0,26
DH5 $\alpha$ T1 $\Delta$ araD	7,7	0,95	0,35
AG1 $\Delta$ araD $\Delta$ ulaF $\Delta$ sgbE	8,3	0,82	0,35
DH5 $\alpha$ T1 $\Delta$ araD $\Delta$ ulaF $\Delta$ sgbE	7,5	0,75	0,28

15 Como se puede observar en la Tabla 3, no hubo diferencias esenciales en la sensibilidad a la arabinosa de las cepas de la invención. Del mismo modo, cuando se sometió a ensayo el rendimiento de ADN plasmídico de las cepas  $\Delta$ araD y  $\Delta$ araD $\Delta$ ulaF $\Delta$ sgbE como se describe en el Ejemplo 3 (los resultados no se muestran), no se encontraron diferencias entre las cepas de *E. coli* AG1 $\Delta$ araD y AG1 $\Delta$ araD $\Delta$ ulaF $\Delta$ sgbE o DH5 $\alpha$ T1 $\Delta$ araD y DH5 $\alpha$ T1 $\Delta$ araD $\Delta$ ulaF $\Delta$ sgbE.

#### Ejemplo 6

##### Estabilidad de S6wtD1EGFP/araD2

25 Una característica importante del vector de vacunación es la estabilidad durante la propagación en células bacterianas. Para someter a ensayo la estabilidad de S6wtD1EGFP/araD2 en las bacterias el plásmido se transformó en las cepas de *E. coli* AG1 $\Delta$ araD y JM109 $\Delta$ araD preparadas en el Ejemplo 3 y la integridad del vector se siguió mediante el análisis de ADN plasmídico durante cuatro generaciones.

30 El plásmido S6wtD1EGFP/araD2 se mezcló con células de *E. coli* AG1 $\Delta$ araD y JM109 $\Delta$ araD competentes y se incubaron sobre hielo durante 30 minutos. Con posterioridad, la suspensión de células se sometió a un choque térmico durante 3 minutos a 37°C seguido de un enfriamiento rápido sobre hielo. Se añadió 1 mL de medio LB a la muestra y la mezcla se incubó durante 45 minutos a 37°C con sacudimiento vigoroso. Finalmente, una porción de las células se cultivó en placas de medio M9 que contenían extracto de levadura al 0,5%, L-arabinosa al 2% y 25  $\mu$ g/ml de kanamicina. Al día siguiente, las células de una colonia se transfirieron a la nueva placa que contenía el mismo medio. Este procedimiento se repitió hasta que se hubieron desarrollado cuatro pases de bacterias. Dos colonias de cada pase de ambas cepas bacterianas se utilizaron para inocular 2 ml de medio M9 que contenía extracto de levadura al 0,5%, L-arabinosa al 2% y 25  $\mu$ g/ml de kanamicina incubado durante la noche a 37°C con sacudimiento vigoroso. Las células se cosecharon y el ADN del plásmido se extrajo de las bacterias utilizando el kit QIAprep Spin Miniprep (Qiagen). Las muestras de ADN plasmídico antes (Figura 9) y después de la digestión con endonucleasa de restricción HindIII (Figura 10) (Fermentas, Lituania) se analizaron mediante electroforesis en gel de agarosa en comparación con el ADN S6wtD1EGFP/araD2 original utilizado para la transformación (como control en las Figuras 9 y 10). El ADN de lambda digerido con EcoRI/HindIII (Fermentas, Lituania) se utilizó como un marcador de peso molecular (M3 en la Figura 10).

Las muestras se digirieron con HindIII como se muestra en la Figura 10A para *E. coli* AG1 $\Delta$ araD y en la Figura 10B para la cepa JM109 $\Delta$ araD, se observaron patrones idénticos al ADN plasmídico S6wtd1EGFP/araD2 original. Los tamaños pronosticados de los fragmentos resultantes de la digestión con HindIII son 3274, 1688 y 1510 pb. Se puede concluir que el vector de vacunación S6wtd1EGFP/araD2 es estable cuando se propaga en las cepas de *E. coli* AG1 $\Delta$ araD y JM109 $\Delta$ araD.

Ejemplo 7

Comparación de un sistema de selección de antibióticos con el sistema de selección de L-arabinosa de la invención

En la comparación de un sistema de selección de antibióticos con el sistema de selección de L-arabinosa de la invención se utilizaron los siguientes medios de crecimiento.

10 Para *E. coli* AG1 que porta el plásmido p2 MG C # 11:

Medio 1: medio M9 más extracto de levadura al 0,5%, glucosa al 0,2% y 25  $\mu$ g/ml de kanamicina (medio selectivo);

Medio 2: medio M9 más extracto de levadura al 0,5% y glucosa al 0,2% (medio no selectivo);

15 Medio 3: medio M9 más extracto de levadura al 0,5%, L-arabinosa al 0,2% y 25  $\mu$ g/ml de kanamicina; (medio selectivo); y

Medio 4: medio M9 más extracto de levadura al 0,5% y L-arabinosa al 0,2% (medio no selectivo).

Para *E. coli* AG1 $\Delta$ araD que porta paraD MG C #145:

Medio 5: medio M9 más extracto de levadura al 0,5%, L-arabinosa al 0,2% y 25  $\mu$ g/ml de kanamicina (medio selectivo); y

20 Medio 6: medio M9 más extracto de levadura al 0,5%, glucosa al 0,2% y 25  $\mu$ g/ml de kanamicina (medio no selectivo).

Los plásmidos p2 MG C # 11 (Figura 20) y paraD MG C # 145 (Figura 21) se transformaron en *E. coli* AG1 y en *E. coli* AG1 $\Delta$ araD que portaban la mutación C a T en el codón 8. Las colonias bacterianas transformadas se cultivaron a 37°C durante la noche en una incubadora. A la mañana siguiente las colonias se inocularon en los medios líquidos selectivos y no selectivos como se ha indicado anteriormente. Los cultivos inoculados se hicieron crecer en un aparato de movimiento oscilante en 2 ml del medio respectivo hasta que alcanzaron la fase estacionaria, y la densidad de los cultivos se midió a DO<sub>600</sub>. El plásmido se extrajo de los cultivos y el rendimiento de ADN plasmídico se determinó por medio de la medición del ADN plasmídico a 260 nm. El rendimiento de plásmido se calculó sobre la base de que 50  $\mu$ g producen una densidad óptica de 1 a 260 nm.

30 A continuación, una alícuota de 20  $\mu$ l de los cultivos estacionarios se inoculó en medio de nueva aportación (dilución 100 veces), y los cultivos se hicieron crecer hasta la fase estacionaria (8-12 horas). La densidad de los cultivos se midió a DO<sub>600</sub>, el plásmido se extrajo y se determinó el rendimiento, y de nuevo se inoculó una alícuota en 2  $\mu$ l del medio líquido. Este procedimiento se repitió 7 veces (preparaciones 1 a 7). Los resultados del experimento se proporcionan en la Tabla 5 de más abajo.

35

Tabla 5. Comparación de un sistema de selección de antibióticos con el sistema de selección de L-arabinosa de la invención

Número de medio/número de preparación	DO <sub>600</sub>	Cantidad de ADN plasmídico por 1 ml de cultivo
1/1	6,215	6,35 µg
1/7	3,278	2,3 µg
2/1	6,652	6,15 µg
2/7	5,133	0,65 µg
3/1	7,317	10,9 µg
3/7	3,046	1,6 µg
4/1	6,874	6 µg
4/7	4,634	0,75 µg
5/1	7,271	6,45 µg
5/7	7,014	5,15 µg
6/1	6,131	5,3 µg
6/7	6,031	4,4 µg

5 Se puede concluir a partir de estos datos que un plásmido que porta el gen de resistencia a la kanamicina y que confiere a *E. coli* la resistencia en presencia de kanamicina se pierde en las etapas de dilución consecutiva/crecimiento del cultivo en las condiciones no selectivas así como en las selectivas. El rendimiento del plásmido a partir del cultivo de 1 ml se reduce 3 veces bajo las condiciones selectivas y 10 veces bajo las condiciones no selectivas en la séptima ronda de dilución (preparaciones 1/1 vs. 1/7 y 2/1 vs. 2/7, respectivamente, en la Tabla 5). Se obtiene el mismo resultado básico, cuando la fuente de carbono para *E. coli* que porta un plásmido con resistencia a la kanamicina es L-arabinosa en lugar de glucosa (preparaciones 3/1 vs 3/7 y 4/1 vs 4/7, respectivamente, en la Tabla 5). Sin embargo, cuando el sistema de selección *araD* de la invención se utiliza en el plásmido, el rendimiento de ADN plasmídico es alto tanto en condiciones selectivas (preparación 5/1 vs 5/7 en la Tabla 5) como no selectivas (preparación 6/1 vs 6/7 en Tabla 5). Tanto bajo condiciones selectivas como no selectivas el rendimiento de ADN plasmídico se redujo a lo largo de 7 generaciones en aproximadamente 20%. Esto indica claramente que los plásmidos que portan el sistema de selección *araD* de la invención son mucho más estables y crecen de manera eficiente en las condiciones selectivas, así como no selectivas.

#### Ejemplo 8

##### Fermentación de alimentación discontinua de AG1Δ*araD* S6wtd1EGFP/*araD2*

20 El sistema de selección basado el gen *araD* se sometió a ensayo también en la fermentación de alimentación discontinua para el propósito de la producción de bacterias que contienen plásmidos. Se recogió una sola colonia de la placa AG1Δ*araD* S6wtd1EGFP/*araD2* y se inoculó en 250 ml de medio M9 que contenía 0,5% de extracto de levadura, 0,2% de L-arabinosa y 25 µg/ml de kanamicina y se incubó durante la noche a 37°C con sacudimiento vigoroso. Después de 18 horas, la DO<sub>600</sub> del inóculo fue de 6,4. Se añadieron 160 ml de inóculo al fermentador que contenía 5 l de Fermenter Starting Medium (8 g/l KH<sub>2</sub>PO<sub>4</sub>; 10 g/l de NaCl; 5 g/l NH<sub>4</sub>Cl; 5 g/l de extracto de levadura; 25 g/l de L-arabinosa; 2 g/l MgSO<sub>4</sub>, 25 mg/l de kanamicina y 0,1 g/l de tiamina; pH 6,7 con NH<sub>4</sub>OH). Después de 5,5 horas de crecimiento la alimentación automática se inició con una velocidad de crecimiento determinada de 0,15 h<sup>-1</sup> (permite crecimiento limitado por la fuente de carbono) con medio de alimentación del fermentador (300 g/l de L-arabinosa; 150 g/l de extracto de levadura; 50 mg/l de kanamicina; 0,2 g/l de tiamina). La velocidad de alimentación se controló por ordenador de acuerdo con las fórmulas  $F(t)=myS^*S_{in}/S_F$  donde myS es la tasa de crecimiento



- 5 deseada,  $S_{in}$  es la cantidad de fuente de carbono añadida en el momento y  $S_F$  es la concentración de la fuente de carbono en el medio de alimentación. El crecimiento fue seguido mediante la medición de la  $DO_{600}$  y se tomaron muestras para el ADN del plásmido. Los datos registrados durante la fermentación se representan en la Figura 11. La fermentación se terminó cuando se consumió 1 l de medio de alimentación. La  $DO_{600}$  final fue de 45. La masa bacteriana se recogió mediante centrifugación y se lavó una vez con 2 l de tampón STE. El rendimiento de biomasa bacteriana fue 410 g de peso húmedo. Los datos de contenido de ADN plasmídico se muestran en la Tabla 6.

Tabla 6. Rendimiento de ADN plasmídico durante la fermentación de AG1 $\Delta$ araD S6wtd1EGFP/araD2

Tiempo	$DO_{600}$	Conc. de ADN plasmídico ( $\mu$ g/ $\mu$ l)	Rendimiento de ADN plasmídico ( $\mu$ g por ml de cultivo)
Inóculo	6,4	0,04	4,6
4 h	3,1	0,02	1,1
21 h	28	0,1	50
24 h	37	0,13	87
29 h	45	0,14	113

- 10 Los datos de la Tabla 6 indican que el sistema de selección de L-arabinosa funciona muy bien a altas densidades celulares. Es probablemente debido a que más copias del plásmido en la célula bacteriana proporcionan una ventaja en las condiciones de limitación de L-arabinosa al permitir que la bacteria utilice el azúcar más rápidamente.

## Ejemplo 9

Purificación de AG1 $\Delta$ araD S6wtd1EGFP/araD2

La purificación de AG1 $\Delta$ araD S6wtd1EGFP/araD2 se realizó como sigue (Figura 12):

- 15 a) Preparación de la alimentación

Se preparó un producto lisado claro de acuerdo con el Manual de Purificación de Plásmidos de Qiagen, excepto que no se utilizó ARNasa.

- 20 Se resuspendieron 200 g de pasta de células de *E. coli* en 2000 ml de Tampón de Resuspensión y después se utilizaron volúmenes iguales de P2 y P3 para la lisis y neutralización. Los restos celulares se eliminaron mediante centrifugación a 6000g durante 30 minutos a 4°C. El producto lisado claro se vertió a través de la toalla de papel, se añadió 1/10 de Triton X-114 (Sigma) al 10% y la disolución se dejó en hielo durante 1 hora. (Se ha demostrado que Triton X-114 reduce eficazmente el nivel de endotoxinas en las proteínas, Liu et al., Clinical Biochemistry, 1997) Después de una hora los ácidos nucleicos fueron precipitados con 0,6 volúmenes de isopropanol frío. El sobrenadante se decantó y el precipitado se almacenó durante la noche a -20°C.

- 25 b) Purificación del ADN plasmídico

La purificación de ADN plasmídico se realizó de acuerdo con el procedimiento de purificación del plásmido superenrollado de tres etapas de Amersham Pharmacia, donde se adoptaron algunas modificaciones.

- 30 Etapa 1. El precipitado se disolvió de nuevo en 1500 ml de TE (Tris-Cl 10 mM, EDTA 1 mM; pH 8,0) y se cargó para la eliminación de ARN y el cambio de tampón sobre Sepharose 6 FF (Amersham Pharmacia), previamente equilibrada con tampón A -  $(NH_4)_2SO_4$  2 M, Tris Cl 100 mM, EDTA 10 mM, pH 7,5.

Etapa 2. El volumen vacío se dirigió a la columna PlasmidSelect (Amersham Pharmacia) (equilibrada con Tampón A) y después del lavado y la elución con Tampón B2 (NaCl 1,6 M,  $(NH_4)_2SO_4$  2 M, Tris Cl 100 mM, EDTA 10 mM, pH 7,5), se capturó el ADN plasmídico superenrollado.

- 35 Etapa 3. El plásmido eluido se diluyó con cinco volúmenes de agua destilada, desionizada y se cargó en SOURCE 30Q (Amersham Pharmacia) equilibrada con tampón C1 (NaCl 0,4 M, Tris Cl 100 mM, EDTA 10 mM, pH 7,5). Después del lavado, el plásmido purificado se eluyó con tampón C2 (NaCl 1M, Tris Cl 100 mM, EDTA 10 mM, pH 7,5) y se recogió el pico de elución. Tamaño de la fracción fue de 150 ml y contenía 100 mg de plásmido S6wtd1EGFP/araD2 libre de endotoxinas (<10 UE/mg).

LISTA DE SECUENCIAS

	<110> FIT Biotech Oyj Plc.	
	<120> Nuevo sistema de selección	
5	<130> 2031002	
	<160> 29	
10	<170> PatentIn versión 3.1	
	<210> 1	
	<211> 780	
	<212> ADN	
15	<213> Escherichia coli	
	<400> 1	
	gtttcgtttg attggetgtg gttttataca gtcattactg cccgtaatat gccttcgcgc	60
	catgcttacg cagatagtgt ttatccagca gcgtttgctg catatccggt aactgcggcg	120
	ctaactgacg gcagaatata cccatataag cgacctcttc cagcacgatg gcgttatgca	180
	ccgcactctc gccatttttg ccccatgcaa acggggcctg ggaatggacc agaacgccgg	240
	gcattttgcgc tgcactgata ccctgttttt caaaggtttc tacgatgacg ttaccggttt	300
	cccactcata ttccgcttg atttctgcgt cggtcatttt gcgggtgcag ggaatggtgc	360
	cgtagaaata gtcggcgtgg gtggtgccgg ttgctggaat cgactgacct gcctgcgccc	420
	agatggtggc gtggcgcgag tgcgtatgca caatgccgcc aatggagggg aatgectgat	480
	agagcagccg gtgagttggc gtgtcggagg agggcttttt cgtaccttca accacttca	540
	cggtttcgat gctaaccacg accatatcgt cagcggctcat gacgctgtaa tcgacgccg	600
	aaggtttgat caciaaagacg ccgcgctcgc gatcaacggc gctgacgttg ccccatgtga	660
	gcgtgaccag gttgtgtttt ggcagcgcca ggttgcttc taatacctgg cgtttgagat	720
	cttctaacaat gttgactcct tcgtgccgga tgcgctttgc ttatccggcc taaaaatcg	780
20	<210> 2	
	<211> 76	
	<212> ADN	
	<213> Secuencia Artificial	
25	<220>	
	<223> Cebador	
	<400> 2	
	cgccatgggt ctcattgttg acagcttata atcgataagc tttaatgcgg tagtttagca	60
	cgaaggagtc aacatg	76
30	<210> 3	
	<211> 64	
	<212> ADN	
	<213> Secuencia Artificial	
35	<220>	
	<223> Cebador	
	<400> 3	

ES 2 529 346 T3

cgccatggac tagtaaaaaa aagcccgctc attagggcggg ctgtcattac tgcccgtaat 60

atgc 64

5 <210> 4  
 <211> 81  
 <212> ADN  
 <213> Secuencia Artificial

10 <220>  
 <223> Cebador

<400> 4  
 cgccatggac tagttctcat gtttgacagc ttatcatcga taagctttaa tgcggtagtt 60

tagcacgaag gagtcaacat g 81

15 <210> 5  
 <211> 58  
 <212> ADN  
 <213> Secuencia Artificial

20 <220>  
 <223> Cebador

<400> 5  
 cgccatggaa aaaaagccc gctcattagg cgggctgtca ttactgccg taatatgc 58

25 <210> 6  
 <211> 22  
 <212> ADN  
 <213> Secuencia Artificial

30 <220>  
 <223> Cebador

<400> 6  
 gccagggttt tccagtcac ga 22

35 <210> 7  
 <211> 24  
 <212> ADN  
 <213> Secuencia Artificial

40 <220>  
 <223> Cebador

45 <400> 7  
 gagcggataa caattcaca cagg 24

50 <210> 8  
 <211> 22  
 <212> ADN  
 <213> Secuencia Artificial

<220>  
 <223> Cebador

55 <400> 8  
 ccaactcacc ggctgctcta tc 22

60 <210> 9  
 <211> 24  
 <212> ADN

# ES 2 529 346 T3

```

<213> Secuencia Artificial
<220>
<223> Cebador
5
<400> 9
aatgccgaag atgcggtgca taac          24

<210> 10
<211> 20
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador
15
<400> 10
taactgcggc gctaactgac          20

<210> 11
<211> 20
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador
25
<400> 11
ggttgctgga atcgactgac          20

<210> 12
<211> 66
<212> ADN
<213> Secuencia Artificial
35
<220>
<223> Cebador

<400> 12
ctcaaacgcc caggtattag aagccaacct ggcgctgcc aaacacgtgt aggctggagc          60
40
tgcttc          66

<210> 13
<211> 60
<212> ADN
<213> Secuencia Artificial
45
<220>
<223> Cebador

<400> 13
ggtttgatca caaagacgcc gcgctcgca tcaacggcgc attccgggga tccgtcgacc          60

<210> 14
<211> 20
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador
60
<400> 14
cggcacgaag gagtaacat          20

```

# ES 2 529 346 T3

<210> 15  
 <211> 20  
 <212> ADN  
 5 <213> Secuencia Artificial  
  
 <220>  
 <223> Cebador  
  
 10 <400> 15  
 tgatagagca gccggtgagt 20  
  
 <210> 16  
 <211> 20  
 15 <212> ADN  
 <213> Secuencia Artificial  
  
 <220>  
 <223> Cebador  
 20  
 <400> 16  
 tcagatcctt gccggaaga 20  
  
 <210> 17  
 <211> 20  
 25 <212> ADN  
 <213> Secuencia Artificial  
  
 <220>  
 <223> Cebador  
 30  
 <400> 17  
 tgtaatcgac gccggaaggt 20  
  
 <210> 18  
 <211> 1030  
 <212> ADN  
 <213> Artificial  
 35  
  
 40 <400> 18  
 ggatccgacc ggcaacggta cagatccgac cggcaacggt acagatccga ccggcaacgg 60  
 tcagatccga ccggcaacgg tacagatccg accggcaacg gtacagatcc gaccggcaac 120  
 ggtacagatc cgaccggcaa cggtacagat ccgaccggca acggtacaga tccgaccggc 180  
 aacggtacag atccgaccgg caacggtaca gatcccccta gcgaattgac tagttctcat 240  
 gtttgacagc ttatcatcga taagctttaa tgcggtagtt tagcacgaag gagtcaacat 300  
 gttagaagat ctcaaacgcc aggtattaga agccaacctg gcgctgcaa aacacaacct 360  
 ggtcacgctc acatggggca acgtcagcgc cgttgatcgc gagcgcggcg tctttgtgat 420  
 caaaccttcc ggcgtcgatt acagcgtcat gaccgctgac gatatggtcg tggttagcat 480

ES 2 529 346 T3

cgaaaccggt gaagtgggtg aaggtacgaa aaagccctcc tccgacacgc caactcaccg 540  
gctgctctat caggcattcc cctccattgg cggcattgtg catacgcact cgcgccacgc 600  
caccatctgg gcgcaggcgg gtcagtcgat tccagcaacc ggcaccaccc acgccgacta 660  
tttctacggc accattccct gcacccgcaa aatgaccgac gcagaaatca acggcgaata 720  
tgagtgggaa accggtaacg tcatcgtaga aacctttgaa aaacagggtg tcgatgcagc 780  
gcaaatgccc ggcgttctgg tccattccca cggcccgttt gcatggggca aaaatgccga 840  
agatgcgggtg cataacgccca tcgtgctgga agaggctcgt tatatgggga tattctgccg 900  
tcagttagcg ccgcagttac cggatatgca gcaaacgctg ctggataaac actatctgcg 960  
taagcatggc gcgaaggcat attacgggca gtaatgacag cccgcctaata gagcgggctt 1020  
tttttccat 1030

<210> 19  
<211> 696  
5 <212> ADN  
<213> Escherichia coli

<400> 19  
atggttagaag atctcaaacg ccaggtatta gaagccaacc tggcgctgcc aaaacacaac 60  
ctggtcacgc tcacatgggg caacgtcagc gccgttgatc gcgagcggg cgtctttgtg 120  
atcaaaccct cggcgctcga ttacagcgtc atgaccgctg acgatatggt cgtgggttagc 180  
atcgaaccg gtgaagtggg tgaaggtacg aaaaagccct cctccgacac gccaaactcac 240  
cggetgctct atcaggcatt cccctccatt ggcggcattg tgcatacgc ctcgcgccac 300  
gccaccatct gggcgcaggc gggtcagtcg attccagcaa ccggcaccac ccacgccgac 360  
tatttctacg gcaccattcc ctgcaaccgc aaaatgaccg acgcagaaat caacggcga 420  
tatgagtggg aaaccggtaa cgtcatcgtg gaaacctttg aaaaacaggg tatcgatgca 480  
gcgcaaatgc ccggcgttct ggtccattcc cacggcccgt ttgcatgggg caaaaatgcc 540  
gaagatgcgg tgcataacgc catcgtgctg gaagaggctc cttatatggg gatattctgc 600  
cgtcagttag cgcgcagtt accggatatg cagcaaacgc tgctggataa acactatctg 660  
cgtaagcatg gcgcgaaggc atattacggg cagtaa 696

10 <210> 20  
<211> 687  
<212> ADN  
<213> Escherichia coli

15 <400> 20

ES 2 529 346 T3

atgcaaaaagc taaaacagca ggtatattgaa gccaacatgg agctgcccgc ctacgggctg 60  
gtgaccttta cctggggcaa cgtcagcgc atcgaccgcg aacgcgggct ggtggtgatc 120  
aagcccagcg gcgttgccta cgaaaccatg aaagcggccg atatggtggt ggttgatag 180  
agcggcaagg tgggtggaagg ggagtatcgc ccatcttccg acaactgcgac gcatctcgaa 240  
ctctaccgtc gttaccgctc gcttgggtggc attgtccata cccactccac tcatgccacc 300  
gcatggggcg aggcggggct ggcgatcccg gcgttaggca ccacgcacgc cgactacttc 360  
tttgcgaca ttccgtgtac gcgcgggta agcgaagaag aggtgcaggc cgagtatgaa 420  
ctgaacaccg gcaaagtgat ttcgaaacg ctgggcaacg ccgagccgct gcatacgcg 480  
ggaattgtgg tgtatcagca cgggcgcttc gcctggggga aagatgctca cgatgcggtg 540  
cataacgcgg tggatgatgga agaagtggcg aaaatggcgt ggattgccg cggcattaac 600  
ccacaactca atcacatcga cagcttctg atgaataaac acttcatgcy taaacacggt 660  
cctaacgctt attacgggca gaagtag 687

<210> 21  
<211> 65  
5 <212> ADN  
<213> Artificial

<400> 21  
cagcaggat ttgaagccaa catggagctg ccgcgctacg ggctggtgta ggctggagct 60  
gcttc 65

10 <210> 22  
<211> 66  
<212> ADN  
<213> Artificial

15 <400> 22  
aaacggctgc ggaattagac cagttatctc ccgaggaagg aaattaattc cggggatccg 60  
tcgacc 66

20 <210> 23  
<211> 21  
<212> ADN  
<213> Artificial

25 <400> 23  
aaacggctgc ggaattagac c 21

30 <210> 24  
<211> 24  
<212> ADN  
<213> Artificial

<400> 24  
gccgtacctg attgagatgt ggag 24

35 <210> 25  
<211> 696  
<212> ADN  
<213> Escherichia coli

40 <400> 25

ES 2 529 346 T3

	atggttagagc aactgaaagc cgaacgtgctg ggggogaatc tggcgcttcc cgctcaccat	60
	ctggtgacgt tcacctgggg caatgtcagc gcggtagacg aaacgcggca atggatggta	120
	atcaaacctt cggcgcgcga gtacgacgtg atgaccgccc acgatatggt ggtggttgag	180
	atagccagcg gtaaggtggt ggaaggcagc aaaaaacctt cttccgatac accaacgcat	240
	ctggcgctct accgtcgcta tgccgaaatt ggcggtattg tgcatacca ctccgcgccac	300
	gccaccatct ggtcacaggc cgggctggat ctccccgcct ggggcaccac ccacgcgat	360
	tatttttacg gtgccatccc ctgcacgcga cagatgaccg cagaggagat taacggcgaa	420
	tatgaatata agaccggcga agtgatcatt gaaaccttcg aagaacgtgg caggagtccg	480
	gcacaaatcc cggcgggtgct ggtgcattct cacggcccg tccatgggg taaaaagcc	540
	gccgatgccg tgcataacgc cgtagtactc gaagaatgcg cctatatggg tctattctcg	600
	cgccagcttg cggcgcagct ccctgcgatg caaaacgaac tgctggataa gcactacctg	660
	cgtaagcatg gggccaatgc ctattacggg cagtaa	696
	<210> 26	
	<211> 67	
5	<212> ADN	
	<213> Artificial	
	<400> 26	
	cgttacagca agaacatat caattcgtag tgccggggcg atgaagaatt ccggggatcc	60
	gtcgacc	67
10		
	<210> 27	
	<211> 65	
	<212> ADN	
	<213> Artificial	
15		
	<400> 27	
	gcaggaggct ggatttatat gttagagcaa ctgaaagccg acgtggtgta ggctggagct	60
	gcttc	65
20		
	<210> 28	
	<211> 24	
	<212> ADN	
	<213> Artificial	
	<400> 28	
25	cgcggttaca gcaaggaaca tacc	24
	<210> 29	
	<211> 22	
	<212> ADN	
30	<213> Artificial	
	<400> 29	
	attgaagcgc gtatgcagga gg	22
35		



## REIVINDICACIONES

- 5 1. Un sistema de selección libre de antibióticos que comprende una célula bacteriana carente de un gen *araD*, en el que un vector que porta un gen *araD* ha sido añadido como marcador de selección, en donde a partir de dicho gen *araD* de dicho vector se expresa una L-ribulosa-5-fosfato 4-epimerasa activa.
2. Un sistema de selección libre de antibióticos de acuerdo con la reivindicación 1, en donde la célula bacteriana es una célula de *Escherichia coli*.
- 10 3. Un sistema de selección libre de antibióticos de acuerdo con la reivindicación 2, en donde una mutación en el gen *araD* del vector introduce un codón de parada en el codón 8 del gen *araD*, en donde a partir de dicho gen *araD* de dicho vector se expresa una L-ribulosa-5-fosfato 4-epimerasa activa.
4. Un sistema de selección libre de antibióticos de acuerdo con la reivindicación 2, en donde la *E. coli* es una cepa de *E. coli* JM109.
5. Un sistema de selección libre de antibióticos de acuerdo con la reivindicación 2 o 3, en donde la *E. coli* es una cepa de *E. coli* DH5 alfa-T1.
- 15 6. Un sistema de selección libre de antibióticos de acuerdo con la reivindicación 2 o 3, en donde la *E. coli* es una cepa de *E. coli* AG1.
7. Un sistema de selección libre de antibióticos de acuerdo con la reivindicación 2, en donde dicha cepa de *E. coli* es la cepa DH5 alfa-T1 carente del gen *araD* y del gen *ulaF*.
- 20 8. Un sistema de selección libre de antibióticos de acuerdo con la reivindicación 2, en donde dicha cepa de *E. coli* es la cepa AG1 carente del gen *araD* y del gen *sgbE*.
9. Un sistema de selección libre de antibióticos de acuerdo con cualquiera de las reivindicaciones 5-6, en donde dicha cepa de *E. coli* carece del gen *araD*, del gen *ulaF*, y del gen *sgbE*.
10. Un vector que comprende una *araD* gen con una mutación que genera un codón de parada en el codón 8 del gen *araD* como marcador de selección.
- 25 11. Un vector de acuerdo con la reivindicación 10, en donde el vector es un vector de expresión que comprende:
- (a) una secuencia de ADN que codifica una proteína de anclaje nuclear ligada operativamente a un promotor heterólogo, comprendiendo dicha proteína de anclaje nuclear
- (i) un dominio de unión a ADN que se une a una secuencia de ADN específica, y
- 30 (ii) un dominio funcional que se une a un componente nuclear, o un equivalente funcional del mismo; y
- (b) una secuencia de ADN multimerizada que forma un sitio de unión para la proteína de anclaje nuclear, en donde dicho vector carece de un origen de la replicación del virus del papiloma.
12. Un vector de acuerdo con la reivindicación 11, en donde
- (a) la proteína de anclaje nuclear es la proteína E2 del virus del papiloma bovino tipo 1 (BPV), y
- 35 (b) la secuencia de ADN multimerizada son múltiples sitios de unión para la proteína E2 de BPV incorporada al vector como agrupación, donde los sitios pueden ser estructuras de cabeza a cola o pueden estar incluidos en el vector mediante posicionamiento espaciado, en donde dicho vector carece de un origen de replicación del virus del papiloma.
13. Un vector de acuerdo con la reivindicación 12, que comprende adicionalmente una deleción en la secuencia de ADN multimerizada.
- 40 14. Un vector de acuerdo con la reivindicación 12, que comprende adicionalmente una mutación en la secuencia de Shine-Dalgarno del gen *araD*.
15. Un método de selección de células transformadas con un plásmido que contiene un gen *araD* como marcador de selección y el gen de interés, comprendiendo el método insertar el plásmido en la célula anfitriona carente de *araD* y hacer crecer las células en un medio de crecimiento que contiene arabinosa, en donde a partir de dicho gen *araD* de dicho plásmido se expresa una L-ribulosa-5-fosfato 4-epimerasa activa.
- 45

16. Un método de acuerdo con la reivindicación 15, en donde la célula es una célula de *Escherichia coli*.

5 17. Un método de acuerdo con la reivindicación 16, en donde el gen *araD* del plásmido se muta generando un codón de parada en el codón 8 del gen *araD*, en donde a partir de dicho gen *araD* de dicho plásmido se expresa una L-ribulosa-5-fosfato 4-epimerasa activa.

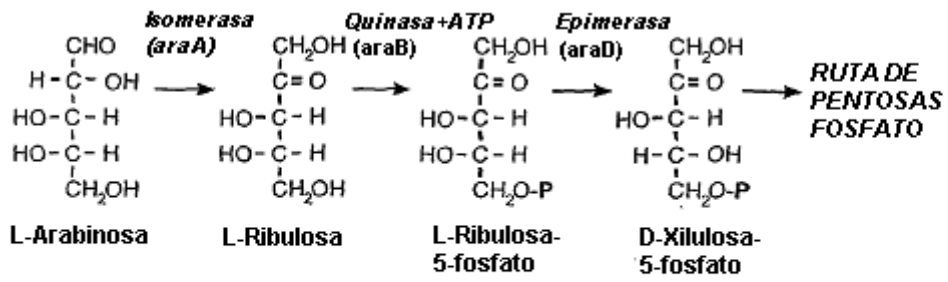


Figura 1

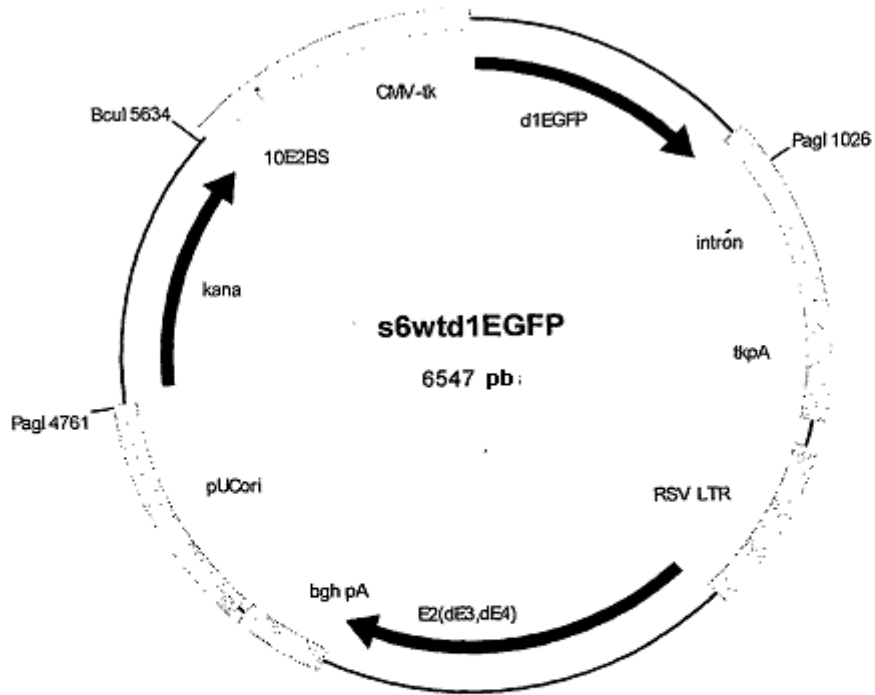


Figura 2

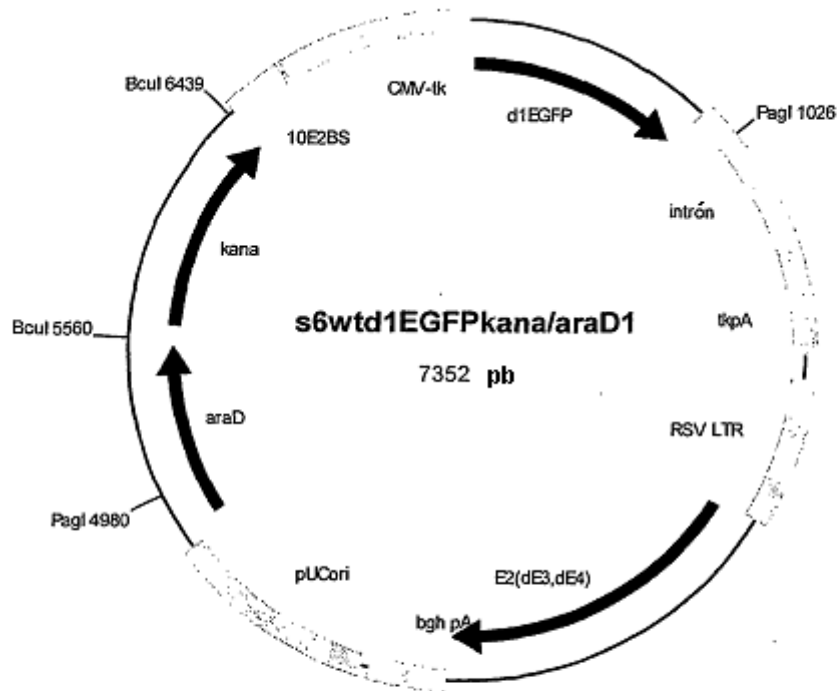


Figura 3

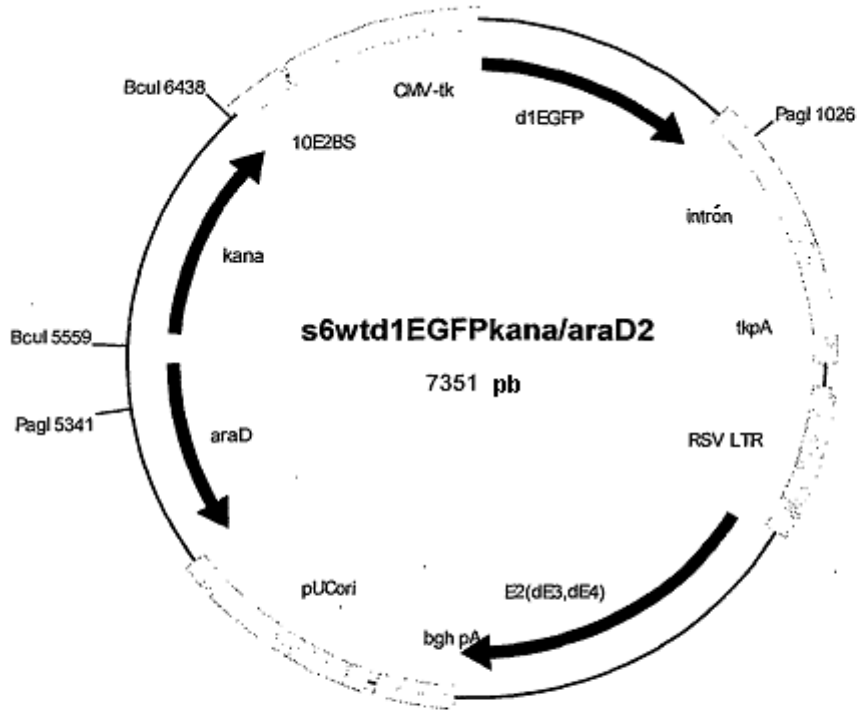
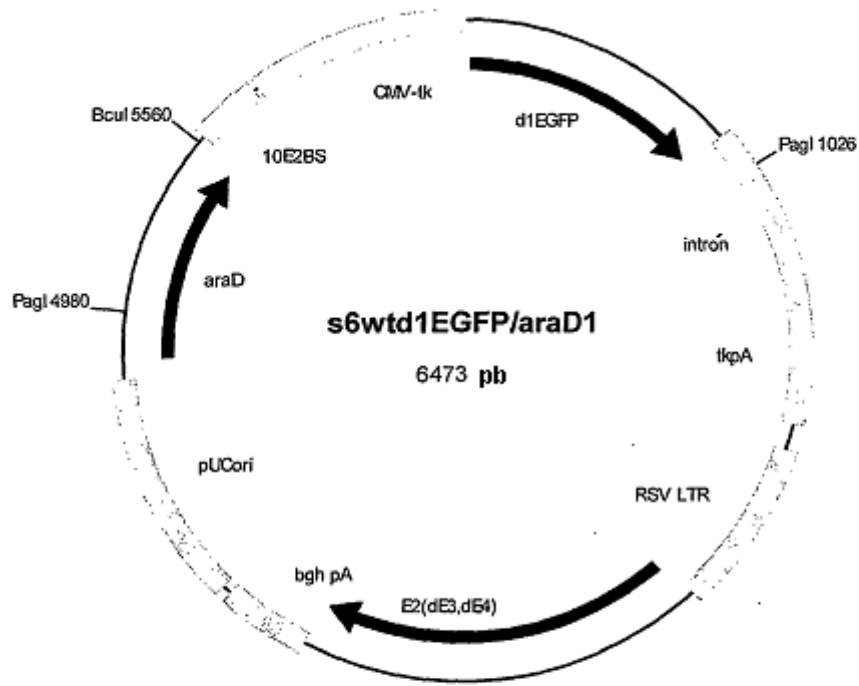


Figura 4



**Figura 5**

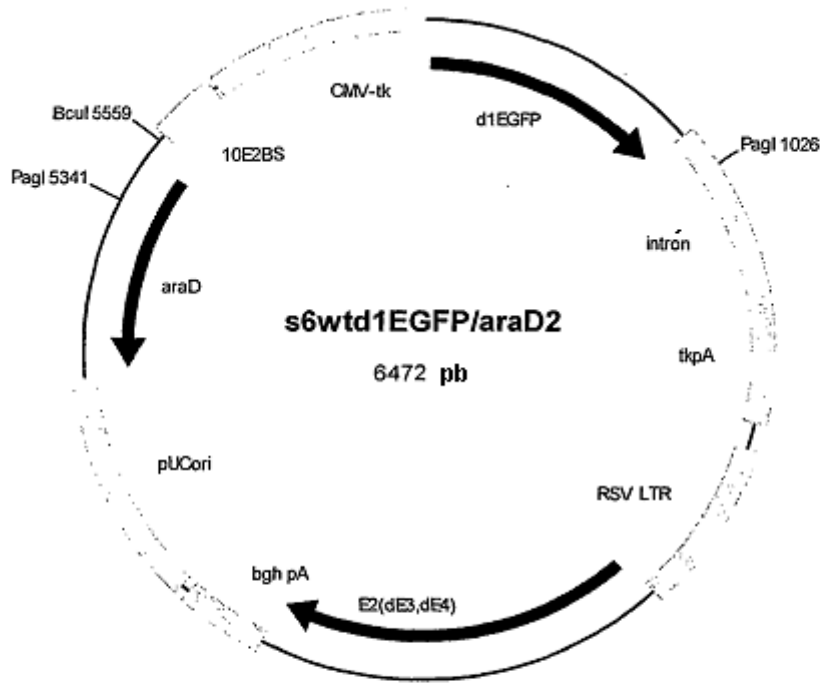


Figura 6



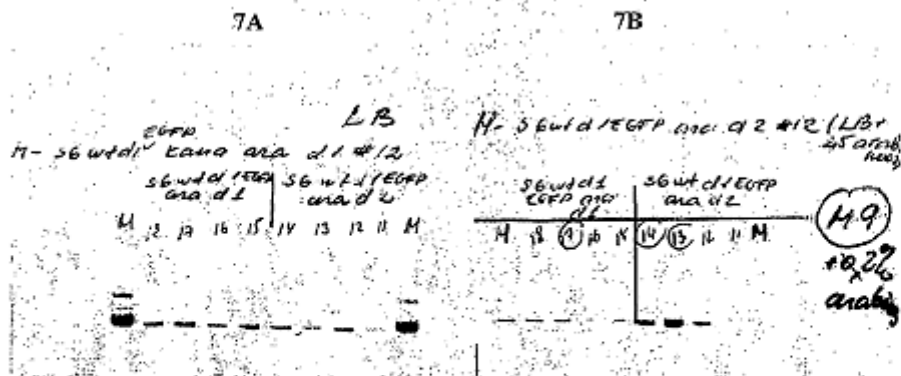


Figura 7

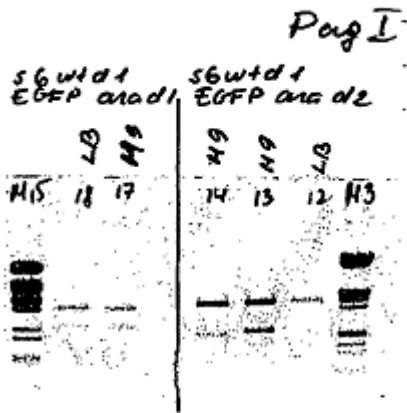


Figura 8

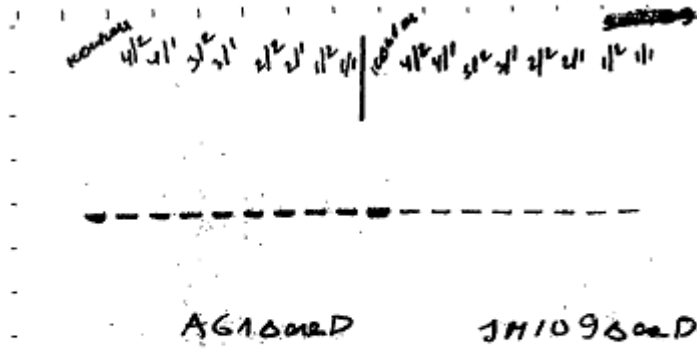


Figura 9

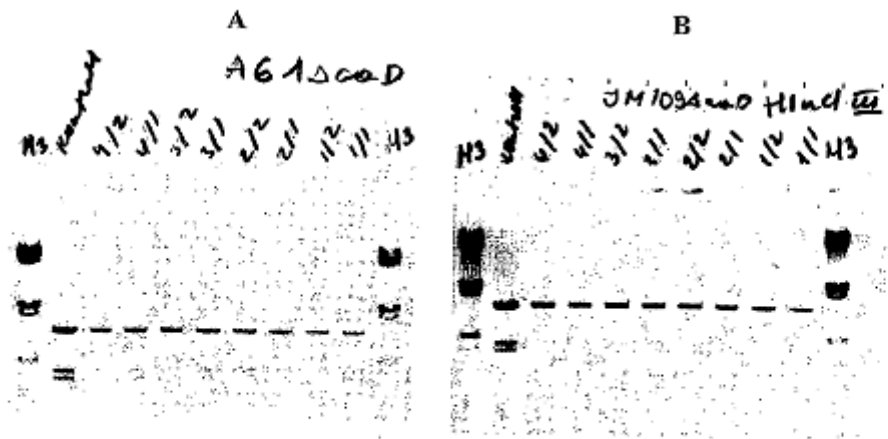


Figura 10

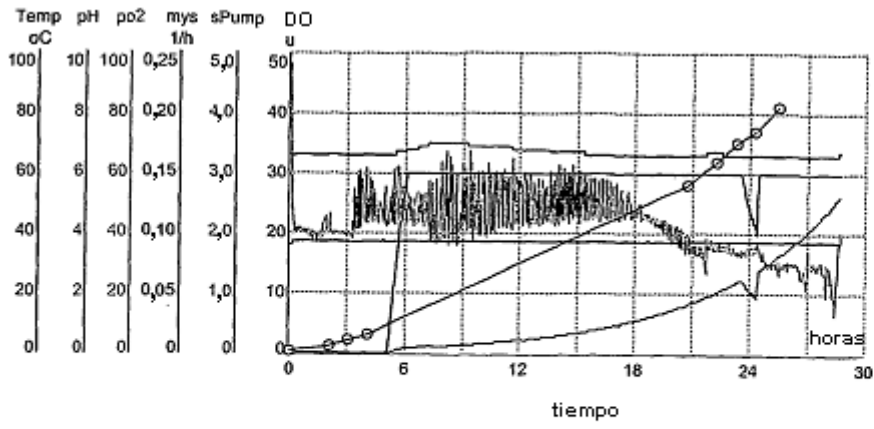
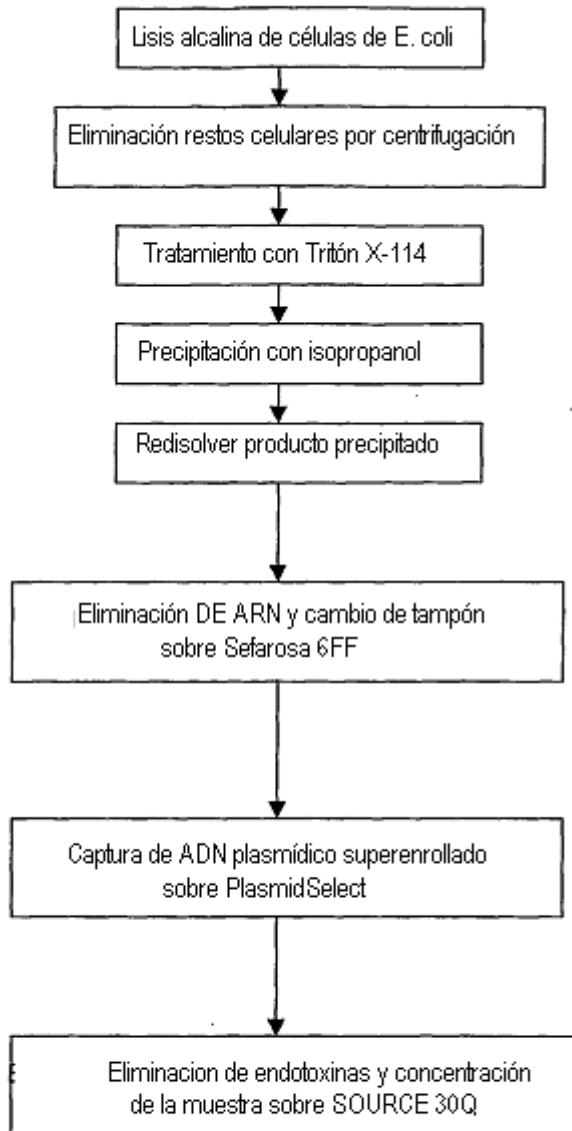


Figura 11



**Figura 12**

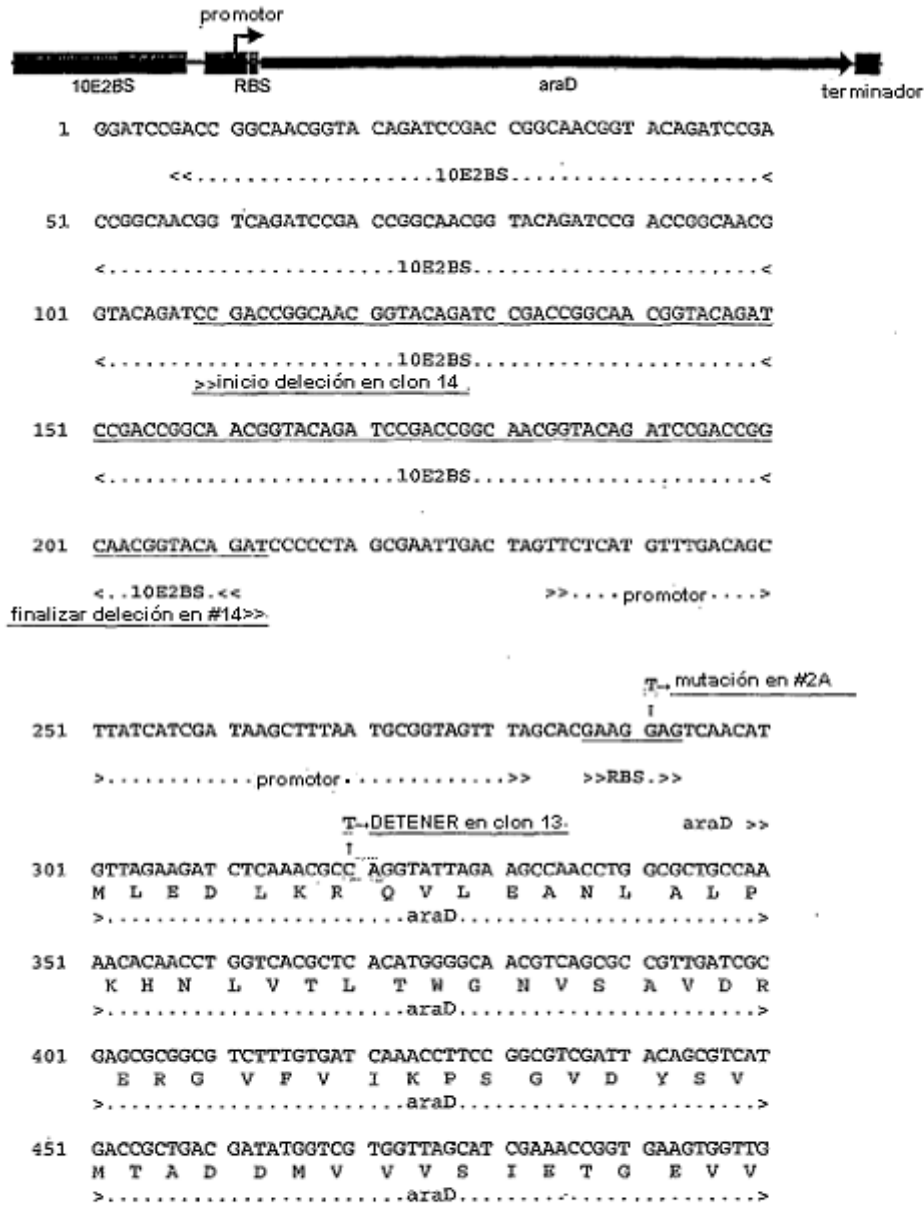


Figura 13

```

501 AAGGTACGAA AAAGCCCTCC TCCGACACGC CAACTCACCG GCTGCTCTAT
    E G T K K P S S D T P T H R L L Y
    >.....araD.....>

551 CAGGCATTCC CCTCCATTGG CGGCATTGTG CATAACGCACT CGCGCCACGC
    Q A F P S I G G I V H T H S R H
    >.....araD.....>

601 CACCATCTGG GCGCAGGCGG GTCAGTCGAT TCCAGCAACC GGCACCACCC
    A T I W A Q A G Q S I P A T G T T
    >.....araD.....>

651 ACGCCGACTA TTTCTACGGC ACCATTCCTT GCACCCGCAA AATGACCGAC
    H A D Y F Y G T I P C T R K M T D
    >.....araD.....>

701 GCAGAAATCA ACGCGAATA TGAGTGGGAA ACCGGTAACG TCATCGTAGA
    A E I N G E Y E W E T G N V I V
    >.....araD.....>

751 AACCTTTGAA AAACAGGGTA TCGATGCAGC GCAAATGCCG GGCCTTCTGG
    E T F E K Q G I D A A Q M P G V L
    >.....araD.....>

801 TCCATTCCCA CGGCCCGTTT GCATGGGGCA AAAATGCCGA AGATGCGGTG
    V H S H G P F A W G K N A E D A V
    >.....araD.....>

851 CATAACGCCA TCGTGCTGGA AGAGTCGCT TATATGGGGA TATTCTGCCG
    H N A I V L E E V A Y M G I P C
    >.....araD.....>

901 TCAGTTAGCG CGCAGTTAC CGGATATGCA GCAAACGCTG CTGGATAAAC
    R Q L A P Q L P D M Q Q T L L D K
    >.....araD.....>

951 ACTATCTGCG TAAGCATGGC GCGAAGGCAT ATTACGGGCA GTAATGACAG
    H Y L R K H G A K A Y Y G Q -
    >.....araD.....>>
                                     terminador' >>

1001 CCCGCCTAAT GAGCGGGCTT TTTTTCAT
    >..... terminador'.....>>

```

Figura 13 (cont.)



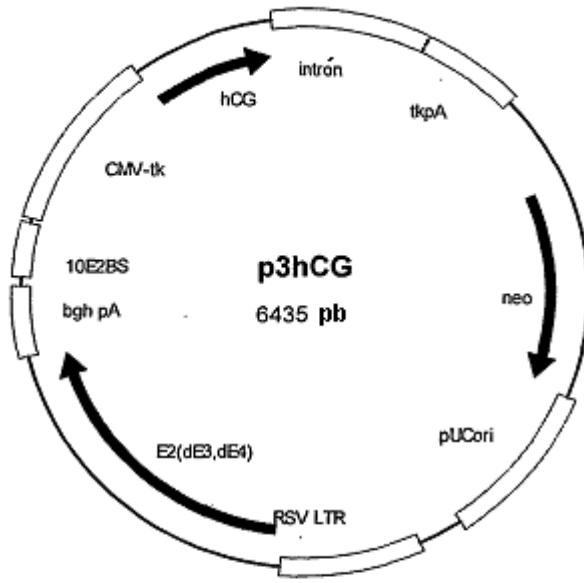


Figura 14

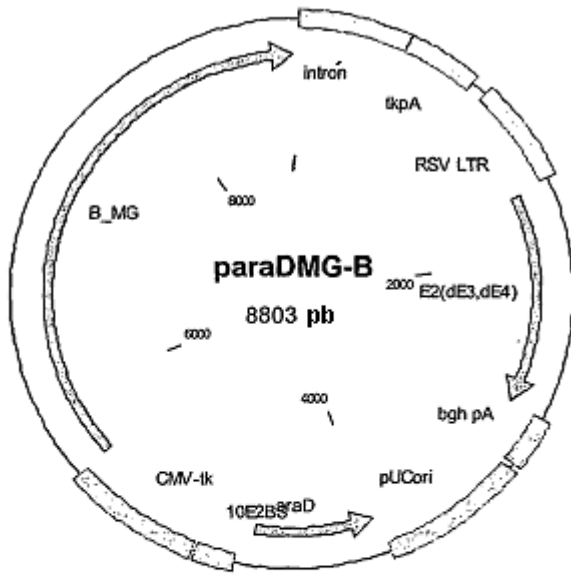


Figura 15

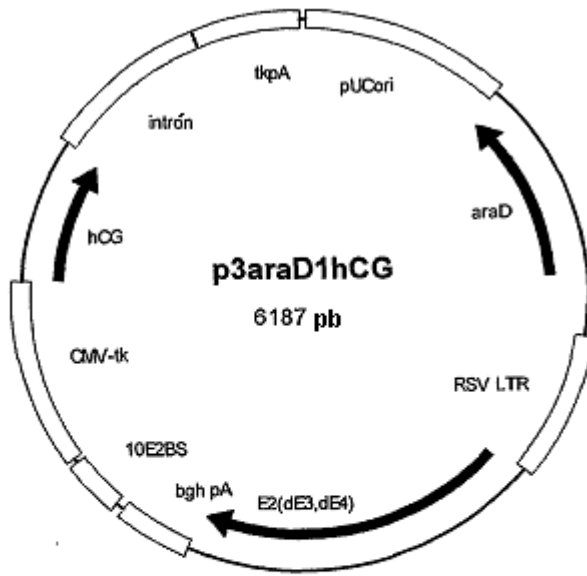


Figura 16

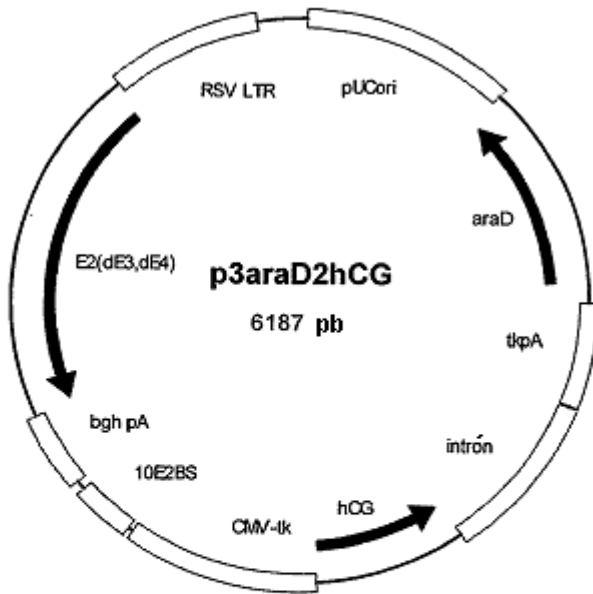


Figura 17

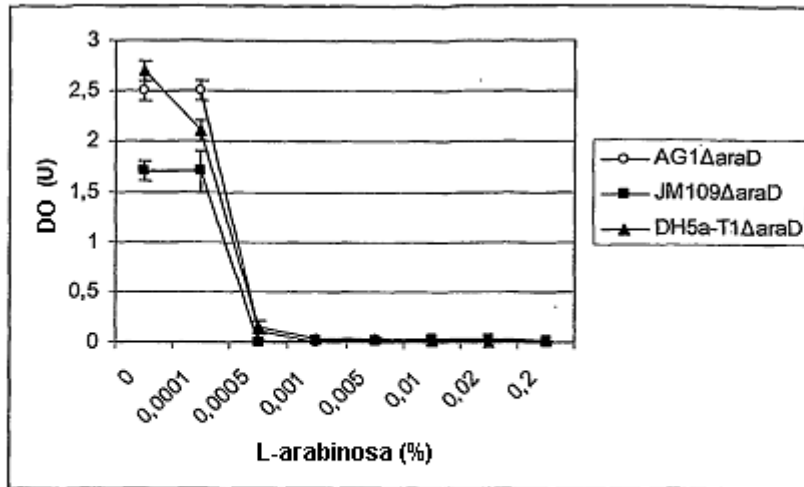
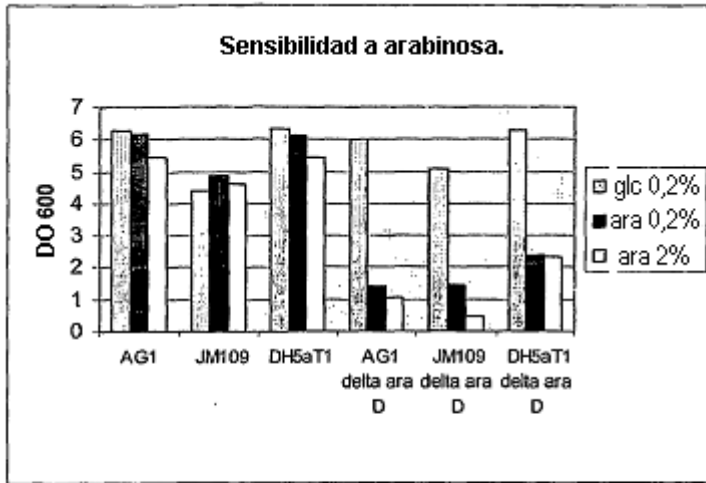


Figura 18



**Figura 19**

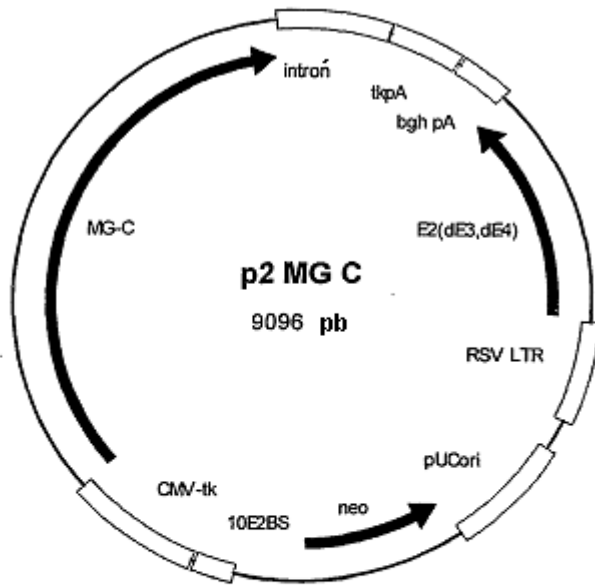
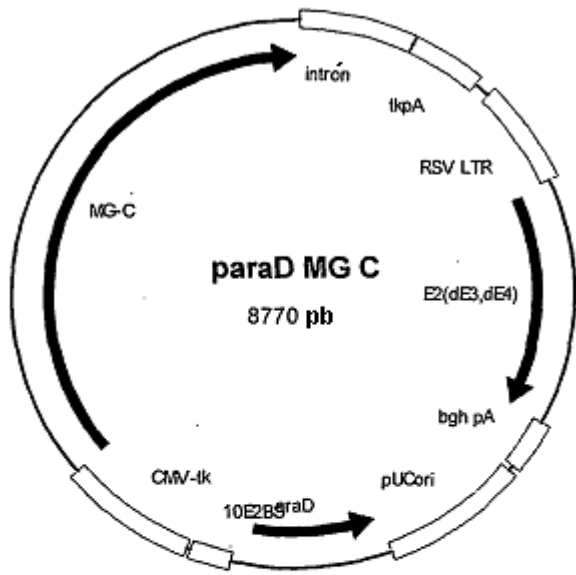
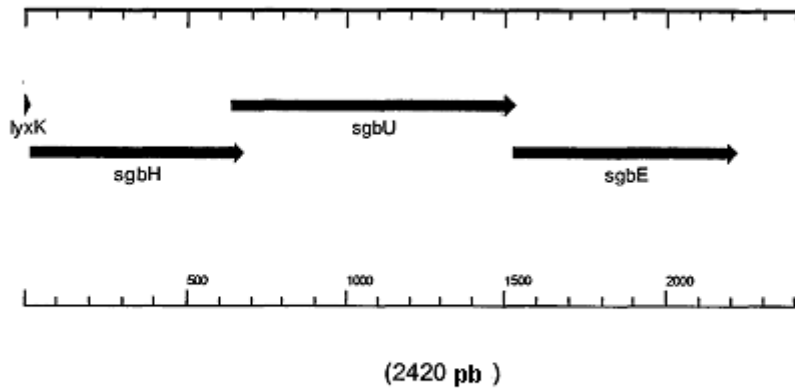


Figura 20



**Figura 21**





```

1400  ATAAACTGAA CTATCGCGGT TCTTTCCTGA TTGAGATGTG GACCGAAAAA
      >.....sgbU.....>
1450  GCCAAGAGC CGGTGCTGGA GATTATTCAG GCGCGGCGTT GGATTGAAGC
      >.....sgbU.....>
1500  GCGTATGCAG GAGGCTGGAT TTATATGTTA GAGCAACTGA AAGCCGACGT
      M L E Q L K A D
      >.....sgbU.....>>
      >>.....sgbE.....>
1550  GCTGGCGGCG AATCTGGCGC TTCCCGCTCA CCATCTGGTG ACGTTCACCT
      V L A A N L A L P A H H L V T F T
      >.....sgbE.....>
    
```

Figura 22

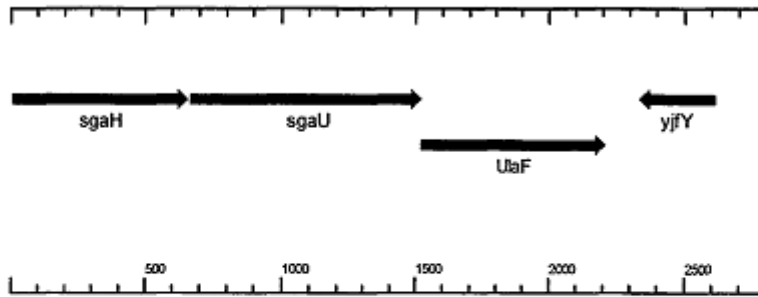
```

1600 GGGGCAATGT CAGCGCGGTA GACGAAACGC GGCAATGGAT GGTAATCAAA
    W G N V S A V D E T R Q W M V I K
    >.....sgbE.....>
1650 CCTTCCGGCG TCGAGTACGA CGTGATGACC GCCGACGATA TGGTGGTGGT
    P S G V E Y D V M T A D D M V V
    >.....sgbE.....>
1700 TGAGATAGCC AGCCGTAAGG TGGTOGAAGG CAGCAAAAAA CCCTCTCCG
    V E I A S G K V V E G S K K P S S
    >.....sgbE.....>
1750 ATACACCAAC GCATCTGCGC CTCTACCGTC GCTATGCCGA AATTGGCGGT
    D T P T H L A L Y R R Y A E I G G
    >.....sgbE.....>
1800 ATTGTGCATA CCCACTCGCG CCACGCCACC ATCTGGTCAC AGGCCGGGCT
    I V H T H S R H A T I W S Q A G
    >.....sgbE.....>
1850 GGATCTCCCC GCCTGGGGCA CCACCCACGC CGATTATTTT TACGGTGCCA
    L D L P A W G T T H A D Y F Y G A
    >.....sgbE.....>
1900 TCCCCTGCAC GCGACAGATG ACCGCAGAGG AGATTAACGG CGAATATGAA
    I P C T R Q M T A E E I N G E Y E
    >.....sgbE.....>
1950 TATCAGACCG GCGAAGTGAT CATTGAAACC TTCGAAGAAC GTGGCAGGAG
    Y Q T G E V I I E T F E E R G R
    >.....sgbE.....>
2000 TCCGGCACAA ATCCCGGGCG TGCTGGTGCA TTCTCACGGC CCGTTCGCAT
    S P A Q I P A V L V H S H G P F A
    >.....sgbE.....>
2050 GGGGTAAAAA CGCCGCCGAT GCCGTGCATA ACGCCGTAGT ACTCGAAGAA
    W G K N A A D A V H N A V V L E E
    >.....sgbE.....>
2100 TGCGCCTATA TGGGTCTATT CTCGCGCCAG CTTGCGCCGC AGCTCCCTGC
    C A Y M G L F S R Q L A P Q L P
    >.....sgbE.....>
2150 GATGCAAAAC GAACTGCTGG ATAAGCACTA CCTGCGTAAG CATGGGGCCA
    A M Q N E L L D K H Y L R K H G A
    >.....sgbE.....>
2200 ATGCTATTA CGGGCAGTAA TCCCTCACGC CGGGGCTTCA TCGCCCCGGC
    N A Y Y G Q -
    >.....sgbE.....>>
2250 ACTACGAATT GATATGTTCC TTGCTGTAAC GCCGCTTCCA CGCTGCTGGC

2300 G

```

Figura 22 (cont.)



(2792 pb)

```

1400  CGGGCCGTAC CTGATTGAGA TGTGGAGCGA AACGGCGGAA GACCCGGCGG
      >.....sgaU.....>
1450  CAGAAGTGGC GAAAGCGCGT GATTGGGTGA AAGCGCGCAT GGCGAAAGCG
      >.....sgaU.....>
1500  GGCATGGTGG AGGCGGCATA ATGCAAAGC TAAACAGCA GGTATTTGAA
      >.....sgaU.....>>
      >>.....UlaF.....>
1550  GCCAACATGG AGCTGCCGCG CTACGGGCTG GTGACCTTTA CCTGGGGCAA
      A N M E L P R Y G L V T F T W G
      >.....UlaF.....>
    
```

Figura 23

```

1600 CGTCAGCGCT ATCGACCGCG AACCGGGGCT GGTGGTGATC AAGCCCAGCG
    N V S A I D R E R G L V V I K P S
    >.....UlaF.....>
1650 GCGTTGCCTA CGAAACCATG AAAGCGGCCG ATATGGTGGT GGTGATATG
    G V A Y E T M K A A D M V V V D M
    >.....UlaF.....>
1700 AGCGGCAAGG TGGTGAAGG GGAGTATCGC CCATCTCCG ACACTGCGAC
    S G K V V E G E Y R P S S D T A
    >.....UlaF.....>
1750 GCATCTCGAA CTCTACCGTC GTTACCGTC GCTGGTGGC ATTGTCCATA
    T H L E L Y R R Y P S L G G I V H
    >.....UlaF.....>
1800 CCCACTCCAC TCATGCCACC GCATGGGCGC AGGCGGGGCT GCGGATCCG
    T H S T H A T A W A Q A G L A I P
    >.....UlaF.....>
1850 GCGTTAGGCA CCACGCACGC CGACTACTTC TTTGGCGACA TTCCGTGTAC
    A L G T T H A D Y F F G D I P C
    >.....UlaF.....>
1900 GCGCGGGTTA AGCGAAGAAG AGGTGCAGG CGAGTATGAA CTGAACACCG
    T R G L S E E E V Q G E Y E L N T
    >.....UlaF.....>
1950 GCAAAGTGAT TATCGAAACG CTGGGCAACG CCGAGCCGCT GCATACGCCG
    G K V I I E T L G N A E P L H T P
    >.....UlaF.....>
2000 GGAATTGTGG TGTATCAGCA CGGGCCGTTT GCCTGGGGGA AAGATGCTCA
    G I V V Y Q H G P F A W G K D A
    >.....UlaF.....>
2050 CGATGCGGTG CATAACGCGG TGGTGATGGA AGAAGTGGCG AAAATGGCGT
    H D A V H N A V V M E E V A K M A
    >.....UlaF.....>
2100 GGATTGCCCG CGGCATTAAC CCACAAC TCA ATCACATCGA CAGCTTCTG
    W I A R G I N P Q L N H I D S F L
    >.....UlaF.....>
2150 ATGAATAAAC ACTTCATGCG TAAACACGCT CCTAACGCTT ATTACGGGCA
    M N K H F M R K H G P N A Y Y G
    >.....UlaF.....>
2200 GAAGTAGAAC ACGCGCTGCG GAAATTCCT TCCTGGGAG ATNACTGGTC
    Q K -
    >....>> UlaF
2250 TAATTCGGCA GCCGTTTTTC AAAAAAAGC CCCCTGCGAA GGGGGCAAAG

2300 C

```

Figura 23 (cont.)