(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 699 652**

(51) Int. CI.:

C12N 7/00 (2006.01)
C12N 15/70 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **06.07.2012 PCT/EP2012/063244**

(87) Fecha y número de publicación internacional: **10.01.2013 WO13004817**

(96) Fecha de presentación y número de la solicitud europea: **06.07.2012 E 12732669 (2)**

(97) Fecha y número de publicación de la concesión europea: **12.09.2018 EP 2729563**

(54) Título: **Fago genéticamente modificado y uso del mismo**

(30) Prioridad:

07.07.2011 EP 11173033
07.07.2011 US 201113178402

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
12.02.2019

(73) Titular/es:

DELPHI GENETICS (100.0%)
Rue A. de Saint-Exupéry 5
6041 Gosselies, BE

(72) Inventor/es:

SZPIRER, CÉDRIC

(74) Agente/Representante:

SÁEZ MAESO, Ana

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

Fago genéticamente modificado y uso del mismo

Campo de la invención

La presente invención se refiere al campo de producción de biomoléculas de interés en sistemas biológicos. Específicamente, la presente invención se refiere a un fago genéticamente modificado integrado en el genoma de una célula anfitriona bacteriana, y uso del mismo con el fin de evitar la contaminación de fago del caldo de cultivo y/o lisis bacteriana.

Antecedente de la invención

Las células bacterianas son sistemas de elección para la producción de biomoléculas, especialmente de proteínas objetivo. Entre otras ventajas, los sistemas bacterianos son fáciles de utilizar y permiten la producción rápida de grandes cantidades de proteínas en un volumen limitado de cultivo.

Uno de los sistemas bacterianos utilizado amplia y habitualmente es el sistema de expresión del bacteriófago T7. Este sistema bacteriano fue descrito en el documento US 4,952,496. En este sistema, el gen que codifica la proteína objetivo se coloca bajo el control de un promotor T7 y se transforma en un anfitrión bacteriano, usualmente E. coli, que comprende un fago lisógeno DE3 lambda integrado. El fago lisógeno lambda DE3 lleva el gen que codifica la polimerasa de ARN T7 bajo el control de un promotor lacUV5. Cuando se cultiva en un medio que contiene IPTG, se induce la expresión de la polimerasa de ARN T7, y permite la expresión de la proteína objetivo. El documento US 2003/003472 divulga, una cepa de E. coli: 5HT71, producida por lisogenización de HB101 con λDE3.

Debido a la integración del gen de la polimerasa de ARN T7 (gen 1 de T7) dentro de la secuencia del gen Int, el fago DE3 Lambda debe ser defectuoso en su capacidad para ingresar en la fase lítica. Sin embargo, la lisis bacteriana se observa durante algunas producciones de proteínas y en ausencia de cualquier otro fago, lo que sugiere que el fago DE3 puede recuperar sus propiedades líticas. La lisis bacteriana y aún más, la presencia de fagos infecciosos en el caldo de cultivo es altamente problemático porque (i) compromete el uso de proteínas objetivo producidas para algunas aplicaciones, tales como, por ejemplo, aplicaciones farmacéuticas, (ii) el proceso de descontaminación con el fin de eliminar cualquier rastro de fagos requiere el cierre de las líneas de producción y la renovación completa de los lotes de cultivo y (iii) reduce dramáticamente el rendimiento de la proteína recombinante.

Los fagos λ modificados se describieron en la técnica. Sternberg et al., 1979 (Virology, 96: 129-142) se refieren a vectores P1-λ híbridos, en los que se elimina el gen int, y que comprenden una supresión del 6% que elimina el ADN entre los genes P y Q lambda. Padukone et al., 1990 (Biotechnology pogress, 6: 277-282) describen vectores derivados de bacteriófagos λ para la expresión de genes clonados estables, en los que dichos vectores comprenden una inactivación del gen S. Miao et al., 1993 (Biotechnology progress, 9: 153-159) describe la infección de células de E. coli con un fago lambda (λCE6), en el que se suprime el gen int y en el que el gen S contiene la mutación de lisis Sam7. Lin et al., 1998 (Biotechnology and bioengineering, 57: 529-535) se refieren a un bacteriófago λ que comprende una mutación en el gen Q y al uso del mismo para producir proteínas recombinantes en el sistema de E. coli. Dorgai et al, 1998 (J.Mol. Biol. 277: 1059-1070) estudian la especificidad de recombinación de la integrasa de fagos lambda y HK022 en secuencias de ácidos nucleicos att y secuencias de proteínas Int. En particular, Dorgai et al, 1998, describen una bacteria lisogénica RW721 que lleva un profago Sam7 λ y se eliminó para att, Int e imm. Christensen et al, 2001 (Mol. Biotechnology, 17: 219-224) describen vectores de expresión del bacteriófago λ que comprenden un sistema de expresión derivado del bacteriófago T7 con mutación en el gen Q. El documento US 2002/090678 se refiere a un método para producir biomoléculas, en las que las células bacterianas de la cepa productora están infectadas con un bacteriófago λ que comprende mutación en uno de los genes N, Q y/o R de fase tardía. El documento WO 2006/091483 se refiere a sistemas de vectores de expresión para uso en la expresión de genes exógenos bajo el control del promotor de la polimerasa T7. Dicho vector de expresión comprende genes Q, S, E o R inactivados. Sin embargo, el uso de estos fagos λ modificados todavía se pueden asociar con la lisis de las células anfitrionas y la posterior contaminación del medio.

Se han descrito métodos alternativos para el uso de un fago:
El documento WO03/050240 describe un sistema de expresión para producir una proteína objetivo en una célula anfitriona que comprende un gen que codifica la polimerasa de ARN T7 integrada utilizando recombinación homóloga. Sin embargo, el sistema del documento WO03/050240 es difícil de implementar, debido al tamaño del gen de la polimerasa de ARN T7 y debido al hecho de que la recombinación homóloga no es fácil de hacer en todas las cepas de E. coli (como lo mencionaron Phue et al., Biotechnology and Bioengineering, 101, 831-836, 2008). En consecuencia, el número de células transformadas que llevan la integración de la polimerasa de ARN T7 permanece muy bajo o no se obtienen estas células. Más aún, utilizando una recombinación homóloga, es necesario utilizar un marcador selectivo para seleccionar las bacterias que contienen la integración del gen de la polimerasa de ARN T7. Este marcador no será utilizable para otra etapa de selección y podría no ser deseado para el uso final de la cepa. Por ejemplo, los marcadores selectivos que se utilizan a menudo son genes de resistencia a los antibióticos, pero se

recomienda evitar estos genes en producciones biofarmacéuticas. Por lo tanto, se requiere una etapa adicional para eliminar el gen de resistencia a los antibióticos de la cepa y no siempre es posible hacerlo.

El documento WO2008/139153 describe otro sistema de expresión, en el que la célula anfitriona se transforma con un plásmido que comprende un casete de expresión para la polimerasa de ARN de T7. Sin embargo, debido al uso de un plásmido, las células anfitrionas se deben mantener en condiciones selectivas para asegurarse de que aún comprenden el plásmido. Adicionalmente, los plásmidos se someten con frecuencia a recombinación, lo que altera el sistema de expresión.

Por lo tanto, subsiste la necesidad de un método novedoso para producir una biomolécula de interés, en el que, cuando se utiliza un fago, el cultivo no se contamine por fagos infecciosos o se lisan involuntariamente durante el crecimiento o la producción de proteínas.

Por lo tanto la presente invención proporciona un fago genéticamente modificado que no recupera sus propiedades líticas durante el cultivo, permitiendo de esta manera la producción de una biomolécula de interés sin contaminación por fagos.

Resumen

Un objeto de la invención es una célula anfitriona bacteriana que comprende un fago genéticamente modificado en el que

- se inserta un sistema de expresión, dicho sistema de expresión comprende una secuencia de ácidos nucleicos que codifica una proteína, dicha proteína induce adicionalmente la expresión de una biomolécula de interés, y
- se suprimen los genes int, S, R, Rz, Xis,

en la que se integra fago genéticamente modificado en el genoma de dicha célula anfitriona bacteriana, y
en la que el fago es el fago lambda, el fago 434, el fago phi80, el fago phi81, el fago HK97, o el fago P21.

En una realización, el gen Q se inactiva adicionalmente.

En una realización, el promotor comprendido en el sistema de expresión es inducible.

En otra realización de la invención, el sistema de expresión es el sistema de expresión T7.

En una realización, dicho fago es uno de los fagos P11 o P13, en el que

- P11 comprende la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, R, Rz, Xis y Int, y
- P13 comprende la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, R, Rz, Q, Xis y Int.

En una realización de la invención, la célula anfitriona es una enterobacteria, preferiblemente E. Coli, más preferiblemente BL21.

En otra realización de la invención, la célula anfitriona como se describió en este documento anteriormente comprende adicionalmente la inactivación de por lo menos uno de los genes tonA, galK, araB, araA, lon, ompT, rcsA, hsdR, mrr, endA y recA.

En otra realización de la invención, la célula anfitriona como se describió en este documento anteriormente comprende la inserción del gen ccdB.

Otro objeto de la invención es un kit que comprende la célula anfitriona como se describió en este documento anteriormente y un plásmido que comprende el gen ccdA.

Otro objeto de la invención es un proceso para preparar una célula anfitriona como se describió en este documento anteriormente, que comprende infectar una célula anfitriona con un fago genéticamente modificado como se describió en este documento anteriormente, y utilizar un fago ayudante para complementación.

Otro objeto de la invención es un proceso para producir una biomolécula de interés, que comprende

- cultivar una célula anfitriona que comprende como se describió en este documento, en la que dicha célula anfitriona comprende la secuencia de ácidos nucleicos de la biomolécula de interés,
- expresar dicha biomolécula de interés.

Descripción detallada

La presente invención se refiere a una célula anfitriona bacteriana que comprende un fago genéticamente modificado, en la que se limita la capacidad del fago para recuperar sus propiedades líticas.

El alcance de la presente invención se determina por las reivindicaciones. Las siguientes realizaciones son realizaciones que caen dentro de las restricciones de las reivindicaciones.

Los inventores se centraron en la modificación genética de un fago. Sorprendentemente, los inventores mostraron que no era posible suprimir todas las secuencias víricas del fago integrado, ya que la viabilidad de las bacterias infectadas se vio gravemente comprometida en ese caso. En particular, los inventores mostraron que la supresión de un fragmento de ADN que comprende las secuencias de codificación del gen ral y el gen N conduce a la muerte de la célula anfitriona (véanse los EJEMPLOS). Este resultado fue sorprendente porque no se sabe que los genes ral y N estén involucrados en el estado lisogénico; el gen N solo se describe como esencial para el crecimiento lítico. Por el contrario, en el estado lisogénico, un represor del fago, llamado CI o C2, bloquea la expresión de N y ral.

De esta manera la presente invención se refiere a una célula anfitriona bacteriana que comprende un fago genéticamente modificado en la que

- se inserta un sistema de expresión, y
- se suprimen los genes Int, Xis, R, S, y Rz y opcionalmente en el que se inactiva el gen Q.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes R.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactiva el gen Q.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes Int.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes Xis.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S y Int.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S y Xis.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S y R.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S y Rz.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S y Q se inactivan.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes R y Rz.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes R y Int.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes R y Xis.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes R y Q.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes Q y Rz.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes Q y Int.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes Q y Xis.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, Int y Xis.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, Int y Rz.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, Int y R.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, Xis y Rz.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, Xis y R.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, R y Rz.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, R y Q.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, Q y Rz.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, Q y Int.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, Q y Xis.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes R, Int y Xis.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes R, Int y Rz.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes R, Xis y Rz.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes R, Q y Rz.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes R, Xis y Q.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes R, Int y Q.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes Q, Int y Rz.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes Q, Int y Xis.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes Q, Xis y Rz.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, Int, Xis y Rz.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, Int, Xis y R.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, Int, R y Rz.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, R, Xis y Rz.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes R, Rz, Xis y Int.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes Q, R, Rz y S.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes Q, R, Rz y Int.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes Q, R, S y Int.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes Q, Rz, S y Int.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes Q, R, Rz y Xis.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes Q, R, S y Xis.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes Q, Rz, S y Xis.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes Q, R, Int y Xis.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes Q, Rz, Int y Xis.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes Q, S, Int y Xis.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, R, Rz, Xis y Int.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, R, Q, Rz y Xis.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, R, Q, Rz y Int.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, R, Q, Xis y Int.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, Q, Rz, Xis y Int.

De acuerdo con la divulgación, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes R, Q, Rz, Xis y Int.

En una realización, el fago genéticamente modificado comprende un sistema de expresión y se inactivan los genes S, R, Q, Rz, Xis y Int.

Ejemplos de fagos que se pueden utilizar en la invención incluyen, pero no se limitan a, los fagos lambda (λ), fago similar a lambda y lambdoides. El fago Lambda, también conocido como colifago lambda, es un virus que infecta a Escherichia coli. Lambda es un bacteriófago temperado. Los fagos similares a Lambda forman una familia de

bacteriófagos y virus arqueales que se caracterizan por largas colas no contráctiles. Los fagos lambdoides son parientes naturales del fago lambda. La mayoría de ellos crecen en E. coli, pero algunos provienen de otras células anfitrionas, tales como, por ejemplo, Salmonella typhimurium. Estos fagos pueden tener el mismo orden genético que lambda.

Los ejemplos de fagos de tipo lambda y lambdoides que se podrían utilizar en la presente invención incluyen, pero no se limitan a, colifago 434, phi80, phi81, HK97, P21 y P22.

En una realización, el fago es lambda (Fago lambda de enterobacteria, número de acceso NC_001416). La organización del genoma del fago lambda se muestra en la tabla 1 a continuación.

Tabla 1

Inicio	Fin	Nombre	Descripción
191	736	nu1	Proteína de empaque de ADN
711	2636	A	Proteína de empaque de ADN
2633	2839	W	proteína de unión de cabeza y cola
2836	4437	B	componente de cápside
4418	5737	C	componente de cápside
5132	5737	nu3	proteína de ensamble de cápside
5747	6079	D	proteína de estabilización de ADN de cabeza
6135	7160	E	componente de cápside
7202	7600	Fi	Proteína de empaque de ADN
7612	7965	Fii	proteína de unión de cabeza y cola
7977	8555	Z	componente de cola
8552	8947	U	componente de cola
8955	9695	V	componente de cola
9711	10133	G	componente de cola
10115	10549	T	componente de cola
10542	13103	H	componente de cola
13100	13429	M	componente de cola
13429	14127	L	componente de cola
14276	14875	K	componente de cola
14773	15444	I	componente de cola
15505	18903	J	cola: proteína de especificidad del anfitrión
18965	19585	lom	membrana de anfitrión externa
19650	20855	orf-401	Proteína de fibra de cola
20767	20147	orf206b	proteína hipotética
21029	21973	orf-314	Fibra de cola
21973	22557	orf-194	Proteína de ensamble de fibra putativa
23918	22686	ea47	ea47
25399	24509	ea31	ea31
26973	25396	ea59	ea59
28882	27812	int	proteína de integración
29078	28860	xis	Excisionasa
29285	29118	hipotética	proteína hipotética
29655	29374	ea8.5	ea8.5
30395	39847	ea22	ea22
31024	30839	orf61	proteína hipotética
31196	31005	orf63	proteína hipotética
31351	31169	orf60a	proteína hipotética
32028	31348	exo	exonucleasa
32810	32025	bet	bet
33232	32816	gam	proteína Gam de inhibidor de nucleasa de anfitrión
33330	33187	kil	proteína de muerte de anfitrión
33463	33299	cIII	proteína antiterminativa
35582	33494	ea10	Proteína de unión a ADN de cadena sencilla putativa
35582	33930	ral	proteína de alivio de restricción
34357	34271	orf28	proteína hipotética
34482	35036	lambdap48	Proteína B de exclusión de superinfección
35582	34560	N	regulador de gen temprano
36259	35825	rexb	proteína de exclusión
37114	36275	rexa	proteína de exclusión
37940	37227	cI	represor
38023	38135	cro	antirrepresor
38360	38653	cII	activador transcripcional

Inicio	Fin	Nombre	Descripción
38686	39585	O	Proteína de replicación de ADN
39582	40283	P	Proteína de replicación de ADN
40280	40570	ren	proteína de exclusión ren
40644	41084	NinB	NinB
41081	41953	NinC	Proteína NinC
41950	42123	NinD	Proteína NinD
42090	42272	NinE	Proteína NinE
42269	42439	NinF	Proteína NinF
42429	43043	NinG	Proteína NinG
43040	43246	NinH	Proteína NinH
43224	43889	NinI	Proteína NinI
43886	44509	Q	regulador de gen tardío
44621	44815	orf-64	proteína hipotética
45186	45509	S	Proteína de lisis celular
45493	45969	R	endolisina
45966	46427	Rz	proteína de lisis celular
46186	46368	Rz1	Proteína Rz1
46752	46459	bor	Precursor de proteína Bor
47575	47042	lambdap78	proteína de envoltura putativa
47738	47944	lambdap79	proteína hipotética

En la presente invención, la posición de los residuos dentro de la secuencia del fago lambda se refiere a NC_001416.

En otra realización, el fago es lambda DE3 (número de acceso EU078592). El fago Lambda DE3 es un fago lambda D69 modificado, que comprende el gen que codifica la polimerasa de ARN T7 bajo el control de un promotor lacUV5. La lista de los genes llevados por la secuencia de Lambda DE3 y su posición se muestran en la Tabla 2 a continuación.

Tabla 2

Inicio	Fin	Nombre	Descripción
341	1423	lacI	represor de operón de lactosa
1546	1995	lacZ	Fragmento de terminal N de beta-galactosidasa
2026	4677	1	Polimerasa de ARN dirigida a ADN T7
5804	5586	xis	excisionasa
6011	5844	hipotética	Proteína hipotética
6381	6100	ea8.5	ea8.5
7121	6573	ea22	ea22
7750	7565	hipotética	Proteína hipotética
7922	7731	hipotética	Proteína hipotética
8077	7895	hipotética	Proteína hipotética
8754	8074	exo	exonucleasa
9536	8751	bet	Bet
9958	9542	gam	proteína Gam de inhibidor de nucleasa de anfitrión
10056	9913	kil	proteína de muerte de anfitrión
10189	10025	cIII	proteína antiterminativa
10630	10262	ea10	proteína de unión a ADN de cadena sencilla putativa
11013	10813	ral	proteína de alivio de restricción
11083	10997	hipotética	Proteína hipotética
11391	11092	N	proteína N de regulación probable (regulador de gen temprano)
12356	11706	C2	Proteína C2 represora
12437	12622	cro	Proteína cro reguladora (Antirrepresor)
12738	13037	cII	proteína antiterminativa
13070	13969	O	Proteína de replicación de ADN
13966	14667	P	Proteína de replicación de ADN
14664	15467	ren	Proteína de exclusión ren
15464	16087	Q	regulador de gen tardío
16199	16393	hipotética	Proteína hipotética
16764	17087	S	proteína de lisis celular
17071	17547	R	proteína de lisis celular
17544	18005	Rz	proteína de lisis celular
18330	18037	Bor	Precursor de proteína Bor
19153	18620	putativa	proteína de envoltura putativa
19316	19522	hipotética	Proteína hipotética

Inicio	Fin	Nombre	Descripción
20270	20815	nu1	Proteína de empaque de ADN
20790	22715	A	Proteína de empaque de ADN
22712	22918	W	proteína de unión de cabeza y cola
22915	24516	B	componente de cápside
24497	25816	C	componente de cápside
25826	26158	D	proteína de estabilización de ADN de cabeza
26214	27239	E	componente de cápside
27281	27679	Fi	Proteína de empaque de ADN
27691	28044	Fii	proteína de unión de cabeza y cola
28056	28634	Z	componente de cola
28631	29026	U	componente de cola
29034	29774	V	componente de cola
29790	30212	G	componente de cola
30194	30628	T	componente de cola
30621	33182	H	componente de cola
33179	33508	M	componente de cola
33508	34206	L	componente de cola
34356	34955	K	componente de cola
34853	35524	I	componente de cola
35585	38983	J	cola: proteína de especificidad del anfitrión
39045	39665	lom	Membrana de anfitrión externa
39730	40935	cola	proteína de fibra de cola
41109	41372	cola	fibra de cola
42175	41237	ea59	ea59

En la presente invención, la posición de los residuos dentro de la secuencia del fago Lambda DE3 se refiere a EU078592.

Como se utiliza en este documento, un "sistema de expresión" se refiere a una molécula de ADN lineal o circular compuesta por un fragmento que codifica una secuencia de ácidos nucleicos unida operativamente a un fragmento adicional para la transcripción del sistema.

El fragmento adicional incluye un promotor y una secuencia de codón de parada. El sistema de expresión puede contener adicionalmente uno o más orígenes de replicación, uno o más marcadores de selección y una secuencia que codifica un sitio de unión a ribosoma.

"Unido operativamente" significa que los fragmentos están dispuestos para funcionar como se supone que funcionan, por ejemplo, una vez que la transcripción comienza en el promotor, pasa a través del fragmento codificado para detener el codón.

"Promotor" en el significado de la presente invención es un elemento de control de la expresión que permite la unión de la polimerasa de ARN y el inicio de la transcripción.

En una realización de la invención, la secuencia de ácidos nucleicos está bajo el control de un promotor "fuerte". Un promotor fuerte se caracteriza por una alta afinidad de unión de la secuencia promotora a una polimerasa de ARN, usualmente la polimerasa de ARN correspondiente de origen natural, por una parte, y la tasa de formación de ARNm por esa polimerasa de ARN por otra parte.

En una realización preferida, la secuencia de ácidos nucleicos está bajo el control de un "promotor inducible". Un "promotor inducible" es un promotor que puede estar regulado por factores externos, por ejemplo la presencia de una molécula inductora (también denominada "inductor") o la ausencia de una molécula represora, o factores físicos como el aumento o la disminución de la temperatura, la osmolalidad o el valor de pH. Los diferentes promotores y los respectivos principios de inducción fueron revisados por Makrides et al. (Microbiological Reviews, 1996, (60)3: 512-538). Ejemplos de promotores inducibles que se pueden utilizar en la presente invención incluyen, pero no se limitan a, el promotor tac o trc, el promotor lac o lacUV5 (todos inducibles por lactosa o su análogo IPTG (isopropiltiol-βD-galactosida)), el promotor araBAD estrechamente regulable (PBAD; Guzman et al., 1995, inducible por arabinosa), el promotor trp (inducible por la adición de ácido β-indol acrílico o la inanición del triptófano, reprimible por la adición de triptófano), el promotor lambda pL (λ) (inducción por un aumento de la temperatura), el promotor phoA (inducible por inanición con fosfato), el PprpB (inducción con propionato) u otros promotores adecuados para la expresión de proteínas recombinantes, todos los cuales utilizan polimerasa de ARN de E. coli.

Entre los promotores inducibles se encuentran aquellos que muestran un comportamiento de expresión "permeable". Dichos promotores (llamados "promotores permeables") son, en principio, inducibles, pero muestran no obstante también una expresión basal sin ser inducidos externamente. Los promotores inducibles que muestran una expresión permeable en condiciones no inducidas se pueden comportar de manera similar a los promotores

constitutivos (es decir, son activos de manera constante y continua o se pueden activar o potenciar como resultado de ciertas condiciones de cultivo). Los promotores permeables pueden ser particularmente útiles para procesos de cultivo operados continuamente. Ejemplos de promotores permeables son el promotor T7 y el promotor trp.

En una realización de la invención, el promotor también puede ser constitutivo, es decir, un promotor que controla la expresión sin la necesidad de inducción por un lado, o la posibilidad de represión por otro lado. Por lo tanto, hay una expresión continua y constante en un cierto nivel. Como ejemplo, el promotor de HCD constitutivo fuerte (Poo et al., Biotechnology Letters, 2002, 24:1185-1189; Jeong et al., Protein expression and purification, 2004, 36:150-156) puede aplicarse para la expresión constitutiva.

En una realización, el sistema de expresión comprende una secuencia de ácidos nucleicos que codifica una proteína que induce la expresión de la biomolécula de interés. De forma ventajosa, la expresión de la biomolécula de interés se induce en condiciones particulares, tales como, por ejemplo, bajo selección.

Ejemplos de dichas secuencias de ácidos nucleicos incluyen, pero no se limitan a, el gen que codifica la polimerasa de ARN T7, gen T7 1. En este caso, la expresión de la polimerasa de ARN T7 induce la expresión de la biomolécula de interés colocada bajo el control de un promotor T7.

Preferiblemente, el sistema de expresión es el sistema de expresión T7. El sistema de expresión T7 se describió en el documento US4,952,496, que se incorpora aquí como referencia. El sistema de expresión T7 comprende un fragmento de ADN del fago T7, que contiene la secuencia de codificación completa para la polimerasa de ARN T7 (es decir, el gen 1 de T7). Se eliminó cualquier promotor activo natural del gen 1 de T7 y se insertó un promotor lacUV5 inducible antes de la secuencia de codificación. El promotor lacUV5 se induce mediante la adición de IPTG al medio de cultivo.

De acuerdo con otra realización, el sistema de expresión comprende la secuencia de ácidos nucleicos de la biomolécula de interés.

Con respecto a la biomolécula de interés, no existen limitaciones. Puede ser, en principio, cualquier secuencia de aminoácidos, secuencias de ácidos nucleicos, tales como, por ejemplo, ADN o ARN. Ejemplos de secuencias de aminoácidos incluyen polipéptidos, proteínas o péptidos que se producen en una escala de fabricación, por ejemplo una biomolécula industrial o una biomolécula terapéutica.

Ejemplos de biomoléculas que se pueden producir por el método de la invención son, sin limitación, enzimas, proteínas reguladoras, receptores, péptidos, por ejemplo hormonas peptídicas, citoquinas, proteínas de membrana o transporte.

Las biomoléculas de interés también pueden ser antígenos utilizados para la vacunación, vacunas, proteínas de unión a antígenos, proteínas inmunoestimulantes, alérgenos, anticuerpos de longitud completa o fragmentos de anticuerpos o derivados. Los derivados de anticuerpos se pueden seleccionar del grupo de anticuerpos de cadena sencilla (scFv), fragmentos Fab, fragmentos Fv, anticuerpos de dominio único (VH o fragmento VL), anticuerpos de dominio como anticuerpos de dominio único de camélidos (VHH, nanocuerpos) u otros formatos de anticuerpos como se describe por ejemplo en Andersen and Reilly (Current Opinion in Biotechnology, 2004, 15: 456-462) o Holliger and Hudson (Nature Biotechnology, 2005 (23) 9: 1126-1136).

Las biomoléculas de interés en la presente invención también se pueden ejemplificar por proteínas (antígeno vírico), por ejemplo, proteína de cubierta, proteína del núcleo, proteasa, transcriptasa inversa, integrasa, y así sucesivamente, codificadas en el genoma de un virus patógeno, por ejemplo, virus de hepatitis B, virus de hepatitis C, I-HV, influenza, y así sucesivamente; factores de crecimiento tales como el factor de crecimiento derivado de plaquetas (PDGF), factor de crecimiento de células madre (SCF), factor de crecimiento de hepatocitos (HGF), factor de crecimiento transformante (TGF), factor de crecimiento nervioso (NGF), factor de crecimiento epidérmico (EGF), factor de crecimiento de fibroblastos (FGF), factor de crecimiento similar a insulina (IGF), y así sucesivamente; citoquinas tales como factor de necrosis tumoral, interferón, interleucina, y así sucesivamente; factores hematopoyéticos tales como eritropoyetina, factor estimulante de colonias de granulocitos, factor estimulante de colonias de macrófagos de granulocitos, factor estimulante de colonias de macrófagos, trombopoyetina, y así sucesivamente; hormonas peptídicas tales como la hormona de liberación de hormona luteinizante (LB-RH), hormona liberadora de tirotropina (TRH), insulina, somatostatina, hormona del crecimiento, prolactina, hormona adrenocorticotrópica (ACTH), hormona estimulante de melanocitos (MSH), hormona estimulante de la tiroides (TSH), hormona luteinizante (LU), hormona estimulante de folículo (FSH), vasopresina, oxitoxina, calcitonina, hormona paratiroidea (PTH), glucagón, gastrina, secretina, pancreozimina, colecistoquinina, angiotensina, placenta lacto8en humana, gonototropina coriónica humana (HCG), ceruleina, motilina, y así sucesivamente; péptidos analgésicos tales como encefalina, endorfina, dinorfina, kiotorfina, y así sucesivamente; enzimas tales como superóxido dismutasa (SOD), uroquinasa, activador del plasminógeno de tejido (TPA), asparaginasa, calicreína, y así sucesivamente; neurotransmisores peptídicos tales como bombesina, neutrotensina, bradiquinina, sustancia P, péptido amiloide de Alzheimer (AD), SOD1, presenilina 1 y 2, renina, Dsunucleina, amiloide A, amiloide P, activina, anti-HER-2, bombesina, encefininasa, inhibidores de proteína, enzimas terapéuticas, antitripsina D 1-, inhibidor de

tripsina de mamíferos, inhibidor de tripsina pancreática de mamíferos, calcitonina, factor de hipertrofia cardíaca, cardiotrofinas (tales como cardiotrofina-1), proteínas CD (tales como CD-3, CD-4, CD-8) y CD-19), CFTR, CTNF, DNasa, gonadotropina coriónica humana, péptido asociado a gonadotropina de ratón, citoquinas, transtirretina, amilina, lipoproteínas, linfoquinas, lisozima, una hormona del crecimiento (que incluyen hormona del crecimiento humano), hormona de crecimiento bovino, factor de liberación de hormona del crecimiento, hormona paratiroidea, hormona estimulante de la tiroides, factores de crecimiento, factor de crecimiento neurotrófico derivado del cerebro, factor de crecimiento epidérmico (EGF), factor de crecimiento de fibroblastos (tales como D FGF y D FGF), factor de crecimiento similar a insulina I y II, des (1-3)-IGF-I (IGF-I de cerebro), proteínas de unión al factor de crecimiento similar a insulina, factor de crecimiento nervioso (tal como NGF-D), factor de crecimiento derivado de plaquetas (PDGF), factor de crecimiento endotelial vascular (VEGF), receptores para hormonas de crecimiento o factores de crecimiento, factor de crecimiento transformante (TGF) (como TGF-D, TGF-D 1, TGF-D2, TGF-D3, TGF-D4 o TGF-D5), factores neurotróficos (tales como neurotrofina-3, -4, -5, o -6), gelsolina, glucagón, calicreínas, sustancia inhibidora de mulleriana, factores neurotróficos, p53, proteína A o D, prorelaxina, cadena A de relaxina, cadena B de relaxina, factores reumatoides, rodopsina, una albúmina sérica (tal como albúmina de suero humano), inhibina, insulina, cadenas de insulina, cadena A de insulina, cadena D de insulina, receptor de insulina, proinsulina, hormona luteinizante, integrina, interleucinas (IL) (tales como IL-1 a IL-10, IL-12), IL-13), eritropoyetina, trombopoyetina, fibrilina, hormona estimulante del folículo, factores de coagulación (tales como factor VIIIC, factor IX, factor de tejido y factor de von Willebrands), factores anticoagulantes (tales como proteína C, factor natriurético auricular, surfactante pulmonar), un activador del plasminógeno (tales como el activador del plasminógeno de tejido humano o uroquinasa), trombina, factor de necrosis tumoral D o D, deshidrogenasa de cetoácido D, adrenesinas, proteínas morfogenéticas óseas (BMP), colágeno, factores estimulantes de colonias (CSF) (tales como M-CSF, GM-CSF y G-CSF), factor de aceleración de desintegración, receptores de asentamiento, interferones (tales como interferón-alfa, -gamma y -beta), queratina, factores osteoinductivos, PRNP, proteínas reguladoras, superóxido dismutasa, proteínas de membrana de superficie, proteínas de transporte, receptores de células T, antígenos tales como inmunotoxinas gpl20(HIb), péptido natriurético auricular, proteína exocrina de vesícula seminal, D 2-microglobulina, PrP, precalcitonina, ataxina 1, ataxina 2, ataxina 3, ataxina 6, ataxina 7, huntingtina, receptor de andrógenos, proteína de unión a CREB, gpl 20, p300, CREB, API, ras, NFAT, jun, fos, proteína asociada a atrofia palidolusiana dentaorubral, una proteína microbiana (por ejemplo, proteína de unión a maltosa, transportador ABC, glutatión S transferasa, tiorredoxina, D-lactamasa), proteína fluorescente verde, proteína fluorescente roja, una enzima tal como el superóxido dismutasa, asparaginasa, arginasa, arginina desaminasa, adenosina desaminasa, ribonucleasa, catalasa, uricasa, bilirrubina oxidasa, tripsina, papaína, fosfatasa alcalina, beta-glucoronidasa, nucleósido de purina fosforilasa o batroxobina, un opioide, por ejemplo endorfinas, encefalinas u opioides no naturales, una hormona o neuropéptido, por ejemplo calcitonina, glucagón, gastrinas, hormona adrenocorticotrópica (ACTH), colecistoquininas, hormona lutenizante, hormona de liberación de gonadotropina, gonadotropina coriónica, factor de liberación de corticotropina, vasopresina, oxitocina, hormonas antidiuréticas, hormona estimulante de la tiroides, hormona de liberación de tirotropina, relaxina, prolactina, péptido YY, neuropéptido Y, polipéptido pancreático, leptina, CART (transcripción regulada de cocaína y anfetamina), un péptido relacionado con CART, perilipina, melano-cortinas (hormonas estimulantes de melanocitos) tales como MC-4, hormonas concentradas de melanina, péptidos natriuréticos, adrenomedulina, endotelina, secretina, amilina, péptido intestinal vasoactivo (VIP), polipéptido activador de adenilato ciclasa de pituaria (PACAP), bombesina, péptidos similares a la bombesina, timosina, proteína de unión a heparina, CD4 soluble, factor de liberación hipotalámico y análogos de melanotoninas o funcionales de los mismos. En otra realización de la invención, la proteína objetivo puede ser una enzima de procesamiento tal como proteasas (por ejemplo, enterocinasa, caspasas de tripsina como serina proteasas), lipasa, fosfatasa, glicosil hidrolasas (por ejemplo, manosidasas, xilosidasas, fucosidasas), quinasa, mono o dioxidasa, peroxidasa, transglutaminasa, carboxipeptidasa, amidasa, esterasa y fosfatasa.

Las fuentes preferidas para dichos polipéptidos de mamíferos incluyen fuentes humanas, bovinas, equinas, porcinas, de lupinos y roedores, siendo particularmente preferidas las proteínas humanas.

La biomolécula de interés de la presente invención también abarca variantes de la proteína mencionada anteriormente. Estas variantes abarcan, por ejemplo, una proteína que tiene la misma actividad que la proteína mencionada anteriormente y que comprende una secuencia de aminoácidos con, en la secuencia de aminoácidos de la proteína mencionada, uno o más aminoácidos suprimidos, sustituidos, insertados y/o agregados. Dicha proteína puede ser ejemplificada por una proteína que tiene la misma actividad que la proteína mencionada anteriormente y que comprende una secuencia de aminoácidos con, en la secuencia de aminoácidos de la proteína mencionada anteriormente, uno o más aminoácidos suprimidos, sustituidos, insertados y/o agregados. Se pueden llevar a cabo dos o más tipos diferentes de modificaciones seleccionadas de supresión, sustitución, inserción y adición.

La biomolécula de interés de la presente invención también abarca "péptidos parciales" de la proteína mencionada anteriormente. Un péptido parcial de la proteína se puede ejemplificar mediante un péptido parcial que comprende una secuencia de aminoácidos en la que una porción de la secuencia de aminoácidos de la proteína mencionada se ejecuta ininterrumpidamente, en la que el péptido parcial tiene preferiblemente la misma actividad que dicha proteína. Dicho péptido parcial se puede ejemplificar mediante un polipéptido que tiene una secuencia de aminoácidos que comprende por lo menos 20 y preferiblemente por lo menos 50 de los residuos de aminoácidos en la secuencia de aminoácidos de la proteína mencionada anteriormente. Este polipéptido contiene preferiblemente la

secuencia de aminoácidos que corresponde a la región que está involucrada con la actividad de la proteína mencionada anteriormente. Adicionalmente, el péptido parcial utilizado en la presente invención también puede ser un péptido parcial dado por una modificación de este polipéptido en el que 1 o una pluralidad de residuos de aminoácidos (por ejemplo, aproximadamente 1 a 20, más preferiblemente aproximadamente 1 a 10, e incluso más preferiblemente, aproximadamente 1 a 5) se suprime de, se sustituye en, se inserta en, y/o se agrega a su secuencia de aminoácidos. El péptido parcial utilizado en la presente invención también se puede utilizar como un antígeno para la producción de anticuerpos.

En una realización de la invención, la biomolécula de interés se selecciona del grupo que comprende hormona de crecimiento humana, insulina humana, hormona estimulante del folículo, factor VIII, eritropoyetina, factor estimulante de colonias de granulocitos, alfa-glactosidasa A, alfa-L-iduronidasa, N-acetilgalactosamina-4-sulfatasa, Dornase alfa, activador del plasminógeno de tejido, glucocerebrosidasa, interferón, factor de crecimiento similar a insulina 1, somatotropina bovina, somatotropina porcina, quimiosina bovina y proteína de envoltura del virus de la hepatitis B.

La biomolécula de interés también abarca polipéptidos modificados o proteínas que se han sometido a modificaciones postraduccionales y posteriores a la exportación en el periplasma, tal como ciclación, glicosilación, fosforilación, metilación, oxidación, deshidratación y división proteolítica.

En una realización, la biomolécula de interés es una enzima para metabolizar una biomolécula en el medio extracelular (denominada en este documento "biomolécula extracelular"). En una realización, la biomolécula extracelular comprende un polisacárido o un lípido. En una realización de la invención, el polisacárido comprende alginato, pectina, celulosa, celobiosa, laminarina o una mezcla de los mismos. En una realización de la invención, el lípido comprende un ácido graso, un glicolípido, un lípido betaína, un glicerolípido, un fosfolípido, un glicerolfosfolípido, un esfingolípido, un lípido de esterol, un lípido de prenol, un sacarolípido, un policetido o una mezcla de los mismos. En una realización de la invención, la biomolécula de interés es una enzima que convierte el polisacárido en un monosacárido, un oligosacárido, o ambos.

En una realización de la invención, la biomolécula de interés es una enzima que convierte el lípido en un ácido graso, un monosacárido, o ambos. En una realización de la invención, el monosacárido u oligosacárido es oligoalginato, manuronato, guluronato, manitol, a-ceto-ácido, 4-desoxi-L-eritro-hexoselulosa uronato (DEHU), 2-ceto-3-desoxi D-gluconato (KDG), glucosa, glucuronato, galacturonato, galactosa, xilosa, arabinosa o manosa. En una realización de la invención, el ácido graso es 14:0, trans-14, 16:0, 16:ln-7, trans-16, 16:2n-6, 18:0, 18:ln-9, 18:2n-6, 18:3n-6, 18:3n-3, 18:4n-3, 20:0, 20:2n-6, 20:3n-6, 20:4n-3,20:4n-6, o 20:5n-3.

En una realización de la invención, la biomolécula de interés es una enzima que convierte la biomolécula extracelular en un producto químico básico. En una realización de la invención, el producto químico básico es etanol, butanol o biodiesel. En una realización de la invención, el biodiesel es un ácido graso, un éster de ácido graso o un terpenoide.

Como se utiliza en el presente documento, el término "inactivado" se refiere a la interrupción o la supresión de la expresión de un gen a niveles transcripcionales o de traducción. Preferiblemente, el término "inactivado" se refiere a un gen cuya transcripción se suprime.

De acuerdo con la invención, la inactivación de un gen puede deberse a la mutación del gen o a la inserción del sistema de expresión dentro de la secuencia de codificación del gen. En el significado de la presente invención, el término "mutación" se refiere a un cambio estable en la secuencia genética. Ejemplos de mutación que podrían conducir a la inactivación de un gen en la presente invención incluyen, pero no se limitan a, mutaciones puntuales, inserciones, supresiones y amplificación o duplicación de genes.

Preferiblemente, la mutación es una supresión. El término "supresión" como se utiliza en este documento significa la pérdida o ausencia de un gen, preferiblemente la pérdida o ausencia total de un gen. Más preferiblemente, la supresión comienza en o antes del codón de inicio del gen suprimido, y termina en o después del codón de parada del gen suprimido.

En una realización de la invención, e se inactiva l gen S. El gen S codifica una proteína holina (S105) y una proteína anti-holina (S107). La proteína holina desencadena la formación de agujeros en la membrana. Se requiere la holina para liberación de la endolisina codificada por el gen R. En el lado opuesto, la proteína antiholina inhibe la formación de huecos S105. De acuerdo con una realización, el gen S tiene la secuencia SEQ ID NO: 1. En otra realización, la secuencia del gen S presenta una identidad de secuencia de por lo menos 70%, por lo menos 75%, por lo menos 80%, por lo menos 85%, por lo menos 90%, por lo menos 91%, por lo menos 92%, por lo menos 93%, por lo menos 94%, por lo menos 95%, 5 por lo menos 96%, por lo menos 97%, por lo menos 98%, e incluso más preferiblemente de por lo menos 99% con la SEQ ID NO: 1.

El gen S puede contener modificaciones de secuencia conservadoras que se refieren a modificaciones de aminoácidos que no afectan ni alteran significativamente la función de la proteína S. Dichas modificaciones conservadoras incluyen sustituciones, adiciones y supresiones de aminoácidos. Se pueden introducir modificaciones

en la secuencia del gen S mediante técnicas estándar conocidas en el arte, tales como la mutagénesis dirigida al sitio y la mutagénesis mediada por PCR. Las sustituciones conservadoras de aminoácidos son normalmente aquellas en las que un residuo de aminoácido se reemplaza con un residuo de aminoácido que tiene una cadena lateral con propiedades fisicoquímicas similares. La secuencia modificada de la proteína S puede comprender una, dos, tres, cuatro o más inserciones, supresiones o sustituciones de aminoácidos. Cuando se hacen las sustituciones, las sustituciones preferidas serán modificaciones conservadoras. Las familias de residuos de aminoácidos que tienen cadenas laterales similares se han definido en la técnica. Estas familias incluyen aminoácidos con cadenas laterales básicas (por ejemplo, lisina, arginina, histidina), cadenas laterales ácidas (por ejemplo, ácido aspártico, ácido glutámico), cadenas laterales polares sin carga (por ejemplo, glicina, asparagina, glutamina, serina, treonina, tirosina, cisteína, triptófano), cadenas laterales no polares (por ejemplo, alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina), cadenas laterales beta-ramificadas (por ejemplo, treonina, valina, isoleucina) y cadenas laterales aromáticas (por ejemplo, tirosina, fenilalanina, triptófano, histidina). Por lo tanto, uno o más residuos de aminoácidos dentro de la secuencia de la proteína S se pueden reemplazar con otros residuos de aminoácidos de la misma familia de cadenas laterales y la proteína S modificada se puede probar para la función retenida (es decir, las propiedades establecidas aquí) mediante comparación con la proteína S codificada por la secuencia SEQ ID NO: 1.

El término "identidad" o "idéntico", cuando se utiliza en una relación entre las secuencias de dos o más polipéptidos, se refiere al grado de relación de secuencia entre los polipéptidos, según lo determinado por el número de coincidencias entre cadenas de dos o más residuos de aminoácidos. La "identidad" mide el porcentaje de coincidencias idénticas entre la menor de dos o más secuencias con alineaciones de espacios (si las hay) abordadas por un modelo matemático o programa de computadora en particular (es decir, "algoritmos"). La identidad de los polipéptidos relacionados se puede calcular fácilmente mediante métodos conocidos. Dichos métodos incluyen, pero no se limitan a, aquellos descritos en Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics 5 and Genome Projects, Smith, D. W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part 1, Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; Sequence Análisis Primer, Gribskov, M. and Devereux, J., eds., M. Stockton Press, New York, 1991; and Carillo et al., SIAM J. Applied Math. 48,1073 (1988). Los métodos preferidos para determinar la identidad se diseñan para proporcionar la mayor coincidencia entre las secuencias probadas. Los métodos para determinar la identidad se describen en programas de ordenador disponibles públicamente. Los métodos de programas de ordenador preferidos para determinar la identidad entre dos secuencias incluyen el paquete de programas GCG, que incluye GAP (Devereux et al., Nucl. Acid. Res. \2, 387 (1984); Genetics Computer Group, University de Wisconsin, Madison, Wis.), BLASTP, BLASTN, y FASTA (Altschul et al., J. Mol. Biol. 215, 403-410 (1990)). El programa BLASTX está disponible públicamente en el Centro Nacional de Información Biotecnológica (NCBI) y otras fuentes (Manual BLAST, Altschul y otros. NCB/NLM/NIH Bethesda, Md. 20894; Altschul et al., supra). El conocido algoritmo Smith Waterman también se puede utilizar para determinar la identidad.

En otra realización, se inactiva el gen R. La proteína R es una endolisina: esta transglicosilasa degrada la mureina de la pared celular de la célula anfitriona. De acuerdo con una realización, el gen R tiene la secuencia SEQ ID NO: 2. En otra realización, la secuencia del gen R presenta una identidad de secuencia de por lo menos 70%, por lo menos 75%, por lo menos 80%, por lo menos 85%, por lo menos 90%, por lo menos 91%, por lo menos 92%, por lo menos 93%, por lo menos 94%, por lo menos 95%, 5 por lo menos 96%, por lo menos 97%, por lo menos 98%, e incluso más preferiblemente de por lo menos 99% con la SEQ ID NO: 2.

El gen R puede contener modificaciones de secuencia conservadoras como se describió aquí anteriormente que se refieren a modificaciones de aminoácidos que no afectan ni alteran significativamente la función de la proteína R. La secuencia modificada de la proteína R puede comprender una, dos, tres, cuatro o más inserciones, supresiones o sustituciones de aminoácidos. Por lo tanto, uno o más residuos de aminoácidos dentro de la secuencia de la proteína R se pueden reemplazar con otros residuos de aminoácidos de la misma familia de cadenas laterales y la proteína R modificada se puede probar para la función retenida (es decir, las propiedades establecidas aquí) mediante comparación con la proteína R codificada por la secuencia SEQ ID NO: 2.

En otra realización, se inactiva el gen Rz. La proteína Rz pertenece a la familia de espanina. Esta proteína puede estar involucrada en la ruptura de la membrana externa de la célula anfitriona durante la fase lítica. De acuerdo con una realización, el gen Rz tiene la secuencia SEQ ID NO: 3. En otra realización, la secuencia del gen Rz presenta una identidad de secuencia de por lo menos 70%, por lo menos 75%, por lo menos 80%, por lo menos 85%, por lo menos 90%, por lo menos 91%, por lo menos 92%, por lo menos 93%, por lo menos 94%, por lo menos 95%, 5 por lo menos 96%, por lo menos 97%, por lo menos 98%, e incluso más preferiblemente de por lo menos 99% con la SEQ ID NO: 3.

El gen Rz puede contener modificaciones de secuencias conservadoras como se describió en este documento anteriormente que se refieren a modificaciones de aminoácidos que no afectan o alteran significativamente la función de la proteína Rz. La secuencia modificada de la proteína Rz puede comprender una, dos, tres, cuatro o más inserciones, supresiones o sustituciones de aminoácidos. De esta manera, uno o más residuos de aminoácidos dentro de la secuencia de la proteína Rz se puede reemplazar con otros residuos de aminoácidos desde la misma familia cadena lateral y la proteína Rz modificada se puede probar para función retenida (es decir, las propiedades

establecidas en este documento) mediante comparación con la proteína Rz codificada mediante la secuencia SEQ ID NO: 3.

En otra realización, el gen Q se inactiva. La proteína Q es un regulador de gen tardío: la proteína Q es un antiterminador de Síntesis de ARN desde el promotor utilizado para transcripción de la región "tardía" completa de ADN de lambda. De acuerdo con una realización, el gen Q tiene la secuencia SEQ ID NO: 35. En otra realización, la secuencia del gen Q presenta una identidad de secuencia de por lo menos 70%, por lo menos 75%, por lo menos 80%, por lo menos 85%, por lo menos 90%, por lo menos 91%, por lo menos 92%, por lo menos 93%, por lo menos 94%, por lo menos 95%, 5 por lo menos 96%, por lo menos 97%, por lo menos 98%, e incluso más preferiblemente de por lo menos 99% con la SEQ ID NO: 35.

El gen Q puede contener modificaciones de secuencias conservadoras como se describió en este documento anteriormente que se refieren a modificaciones de aminoácidos que no afectan o alteran significativamente la función de la proteína Q. La secuencia modificada de la proteína Q puede comprender una, dos, tres, cuatro o más inserciones, supresiones o sustituciones de aminoácidos. De esta manera, uno o más residuos de aminoácidos dentro de la secuencia de la proteína Q se puede reemplazar con otros residuos de aminoácidos desde la misma familia cadena lateral y la proteína Q modificada se puede probar para función retenida (es decir, las propiedades establecidas en este documento) mediante comparación con la proteína Q codificada por la secuencia SEQ ID NO: 35.

En otra realización, se suprime un fragmento de ácido nucleico que comprende la secuencia de codificación del gen S. De acuerdo con la invención, dicho fragmento de ácido nucleico no comprende las secuencias de codificación del gen ral (SEQ ID NO: 4) y/o del gen N (SEQ ID NO: 5). Por lo tanto, de acuerdo con una realización, cuando no se suprime el fago es lambda, la región entre residuos 33930 y 35582.

De acuerdo con otra realización, cuando no se suprime el fago es lambda(DE3), la región entre residuos 10813 y 11391.

Preferiblemente, dicho fragmento de ácido nucleico no comprende la secuencia de codificación de N y C2 (SEQ ID NO: 6), así como también las secuencias reguladoras del promotor de C2. Para asegurarse de que las secuencias reguladoras de C2 no se suprimen, el fragmento que se va a suprimir puede estar en el codón de inicio ATG del siguiente gen: el gen Cro.

Más preferiblemente, dicho fragmento de ácido nucleico no comprende la secuencia de codificación de ral, N y C2 (SEQ ID NO: 6), así como también las secuencias reguladoras del promotor de C2.

El gen C2 codifica una proteína represora, lo cual es importante para mantener el estado lisogénico. Este gen también se llama cI en la secuencia de varios otros fagos lambdoides. En el fago lambda, la secuencia de codificación de los genes ral, N y cI es desde los residuos 33930 a 38040 o desde los residuos 33930 a 38022. En el fago Lambda DE3, la secuencia de codificación de los genes ral, N, C2 es desde los residuos 10813 a 12436.

En una realización, la longitud de dicho fragmento de ácido nucleico suprimido es desde 30 kb hasta 300 b, preferiblemente desde 5 kb hasta 500 b. En otra realización, la longitud de dicho fragmento de ácido nucleico suprimido es 30 kb, 25 kb, 20 kb, 15 kb, 10 kb, 9 kb, 8 kb, 7 kb, 6 kb, 5 kb, 4 kb, 3 kb, 2kb, 1 kb, 750 b, 500b.

De acuerdo con una realización en la que el fago es lambda, el fragmento de ácido nucleico suprimido inicia en una posición que varía desde la posición 35582 hasta la posición 45186, preferiblemente desde 38041 hasta 45186 o desde 38023 hasta 45186 y finaliza en una posición que varía desde la posición 45510 hasta la posición 48502

De acuerdo con otra realización en la que el fago es lambda DE3, el fragmento de ácido nucleico suprimido inicia en una posición que varía desde la posición 11392 hasta la posición 16764, preferiblemente desde 12437 hasta la posición 16764 y finaliza en una posición que varía desde 17088 y 42925.

En otra realización, el fragmento de ácido nucleico que se va a suprimir comprende por lo menos la secuencia de codificación del gen S y Rs. De acuerdo con una realización, cuando el fago es lambda, el fragmento de ácido nucleico inicia en una posición que varía desde la posición 35582 hasta la posición 45186, preferiblemente desde 38041 hasta 45186 o desde 38023 hasta 45186, y finaliza en una posición que varía desde la posición 45970 hasta la posición 48502. En otra realización, cuando el fago es lambda DE3, el fragmento de ácido nucleico inicia en una posición que varía desde la posición 11392 hasta la posición 16764, preferiblemente desde 12437 hasta la posición 16764, y finaliza en una posición que varía desde 17548 y 42925.

En otra realización, el fragmento de ácido nucleico que se va a suprimir comprende por lo menos la secuencia de codificación de los genes S, R y Rz. De acuerdo con una realización, cuando el fago es lambda, el fragmento de ácido nucleico inicia en una posición que varía desde la posición 35582 hasta la posición 45186, preferiblemente desde 38041 hasta 45186 o desde 38023 hasta 45186. De acuerdo con esta realización, el fragmento puede

finalizar en una posición que varía desde la posición 46428 hasta la posición 46458. Incluso de acuerdo con esta realización, el fragmento de ácido nucleico puede finalizar en una posición que varía desde 46428 hasta 48502.

De acuerdo con otra realización, cuando el fago es lambda DE3, el fragmento de ácido nucleico inicia en una posición que varía desde la posición 11392 hasta la posición 16764, preferiblemente desde 12437 hasta la posición 16764. De acuerdo con esta realización, el fragmento puede finalizar en una posición que varía desde la posición 18006 y 18036. Incluso de acuerdo con esta realización, el fragmento de ácido nucleico puede finalizar en una posición que varía desde la posición 18006 hasta 24496. Incluso de acuerdo con esta realización, el fragmento de ácido nucleico puede finalizar en una posición que varía desde la posición 18006 hasta 42925.

De acuerdo con otra realización, el fragmento de ácido nucleico que se va a suprimir comprende por lo menos la secuencia de codificación del gen R.

De acuerdo con una realización, cuando el fago es lambda, el fragmento de ácido nucleico inicia en una posición que varía desde la posición 35582 hasta la posición 45493, preferiblemente que varía desde la posición 38041 o 38022 hasta la posición 45493. De acuerdo con esta realización, el fragmento de ácido nucleico que se va a suprimir puede finalizar en una posición que varía desde 45970 hasta la posición 48502.

De acuerdo con otra realización, cuando el fago es lambda DE3, el fragmento de ácido nucleico inicia en una posición que varía desde la posición 11392 hasta la posición 17071, preferiblemente que varía desde la posición 12437 hasta la posición 17071. De acuerdo con esta realización, el fragmento de ácido nucleico que se va a suprimir puede finalizar en una posición que varía desde 17548 hasta la posición 24496. Incluso de acuerdo con esta realización, el fragmento de ácido nucleico que se va a suprimir puede finalizar en una posición que varía desde 17548 hasta la posición 42925.

De acuerdo con otra realización, el fragmento de ácido nucleico que se va a suprimir comprende por lo menos la secuencia de codificación de los genes R y Rz.

De acuerdo con una realización, cuando el fago es lambda, el fragmento de ácido nucleico inicia en una posición que varía desde la posición 35582 hasta la posición 45493, preferiblemente que varía desde la posición 38041 o 38022 hasta la posición 45493. De acuerdo con esta realización, el fragmento de ácido nucleico que se va a suprimir puede finalizar en una posición que varía desde 46428 hasta la posición 46458. Incluso de acuerdo con esta realización, el fragmento de ácido nucleico que se va a suprimir puede finalizar en una posición que varía desde 46428 hasta la posición 48502.

De acuerdo con otra realización, cuando el fago es lambda DE3, el fragmento de ácido nucleico inicia en una posición que varía desde la posición 11392 hasta la posición 17071, preferiblemente que varía desde la posición 12437 hasta la posición 17071. De acuerdo con esta realización, el fragmento de ácido nucleico que se va a suprimir puede finalizar en una posición que varía desde 18006 hasta la posición 18330. Incluso de acuerdo con esta realización, el fragmento de ácido nucleico que se va a suprimir puede finalizar en una posición que varía desde 18006 hasta la posición 24496. Incluso de acuerdo con esta realización, el fragmento de ácido nucleico que se va a suprimir puede finalizar en una posición que varía desde 18006 hasta la posición 42925.

De acuerdo con otra realización, el fragmento de ácido nucleico que se va a suprimir comprende por lo menos la secuencia de codificación del gen Q.

De acuerdo con una realización en la que el fago es lambda, el fragmento de ácido nucleico suprimido inicia en una posición que varía desde la posición 35582 hasta la posición 43886, preferiblemente desde 38041 hasta 43886 o desde 38023 hasta 43886 y finaliza en una posición que varía desde la posición 44510 hasta la posición 48502.

De acuerdo con otra realización en la que el fago es lambda DE3, el fragmento de ácido nucleico suprimido inicia en una posición que varía desde la posición 11392 hasta la posición 15464, preferiblemente desde 12437 hasta la posición 15464 y finaliza en una posición que varía desde 16088 y 42925.

De acuerdo con otra realización, el fragmento de ácido nucleico que se va a suprimir comprende por lo menos la secuencia de codificación de los genes Q y S.

De acuerdo con una realización en la que el fago es lambda, el fragmento de ácido nucleico suprimido inicia en una posición que varía desde la posición 35582 hasta la posición 43886, preferiblemente desde 38041 hasta 43886 o desde 38023 hasta 43886 y finaliza en una posición que varía desde la posición 45510 hasta la posición 48502.

De acuerdo con otra realización en la que el fago es lambda DE3, el fragmento de ácido nucleico suprimido inicia en una posición que varía desde la posición 11392 hasta la posición 15464, preferiblemente desde 12437 hasta la posición 15464 y finaliza en una posición que varía desde 17088 y 42925.

De acuerdo con otra realización, el fragmento de ácido nucleico que se va a suprimir comprende por lo menos la secuencia de codificación del gen Q, S y Rs.

De acuerdo con una realización en la que el fago es lambda, el fragmento de ácido nucleico suprimido inicia en una posición que varía desde la posición 35582 hasta la posición 43886, preferiblemente desde 38041 hasta 43886 o desde 38023 hasta 43886 y finaliza en una posición que varía desde la posición 45970 hasta la posición 48502.

De acuerdo con otra realización en la que el fago es lambda DE3, el fragmento de ácido nucleico suprimido inicia en una posición que varía desde la posición 11392 hasta la posición 15464, preferiblemente desde 12437 hasta la posición 15464 y finaliza en una posición que varía desde 17548 y 42925.

De acuerdo con otra realización, el fragmento de ácido nucleico que se va a suprimir comprende por lo menos la secuencia de codificación de los genes Q, S, R y Rz.

De acuerdo con una realización en la que el fago es lambda, el fragmento de ácido nucleico suprimido inicia en una posición que varía desde la posición 35582 hasta la posición 43886, preferiblemente desde 38041 hasta 43886 o desde 38023 hasta 43886 y finaliza en una posición que varía desde la posición 46428 hasta la posición 48502.

De acuerdo con otra realización en la que el fago es lambda DE3, el fragmento de ácido nucleico suprimido inicia en una posición que varía desde la posición 11392 hasta la posición 15464, preferiblemente desde 12437 hasta la posición 15464 y finaliza en una posición que varía desde 18006 y 42925.

En una realización, se inactiva el gen Int. La proteína Int maneja la inserción y la división del genoma de fago en el genoma del anfitrión. De acuerdo con una realización, el gen Int tiene la secuencia SEQ ID NO: 7. En otra realización, la secuencia del gen Int presenta una identidad de secuencia de por lo menos 70%, por lo menos 75%, por lo menos 80%, por lo menos 85%, por lo menos 90%, por lo menos 91%, por lo menos 92%, por lo menos 93%, por lo menos 94%, por lo menos 95%, 5 por lo menos 96%, por lo menos 97%, por lo menos 98%, e incluso más preferiblemente de por lo menos 99% con la SEQ ID NO: 7.

El gen Int puede contener modificaciones de secuencias conservadoras como se describió en este documento anteriormente que se refieren a modificaciones de aminoácidos que no afectan o alteran significativamente la función de la proteína Int. La secuencia modificada de la proteína Int puede comprender una, dos, tres, cuatro o más inserciones, supresiones o sustituciones de aminoácidos. De esta manera, uno o más residuos de aminoácidos dentro de la secuencia de la proteína Int se puede reemplazar con otros residuos de aminoácidos de la misma familia cadena lateral y la proteína Int modificada se puede probar para función retenida (es decir, las propiedades establecidas en este documento) mediante comparación con la proteína Int codificada por la secuencia SEQ ID NO: 7.

En una realización, se inactiva el gen Xis. La proteína Xis es una excisionasa, que está involucrada en el proceso de división del fago lambda ADN fuera del cromosoma anfitrión bacteriano. De acuerdo con una realización, el gen Xis tiene la secuencia SEQ ID NO: 8. En otra realización, la secuencia del gen Xis presenta una identidad de secuencia de por lo menos 70%, por lo menos 75%, por lo menos 80%, por lo menos 85%, por lo menos 90%, por lo menos 91%, por lo menos 92%, por lo menos 93%, por lo menos 94%, por lo menos 95%, 5 por lo menos 96%, por lo menos 97%, por lo menos 98%, e incluso más preferiblemente de por lo menos 99% con la SEQ ID NO: 8.

El gen Xis puede contener modificaciones de secuencias conservadoras como se describió en este documento anteriormente que se refieren a modificaciones de aminoácidos que no afectan o alteran significativamente la función de la proteína Xis. La secuencia modificada de la proteína Xis puede comprender una, dos, tres, cuatro o más inserciones, supresiones o sustituciones de aminoácidos. De esta manera, uno o más residuos de aminoácidos dentro de la secuencia de la proteína Xis se puede reemplazar con otros residuos de aminoácidos de la misma familia cadena lateral y la proteína Xis modificada se puede probar para función retenida (es decir, las propiedades establecidas en este documento) mediante comparación con la proteína Xis codificada por la secuencia SEQ ID NO: 8.

En una realización, se suprime un fragmento de ácido nucleico que comprende la secuencia de codificación del gen Int. De acuerdo con una realización, el fago es lambda y el fragmento de ácido nucleico que se va a suprimir comprende los residuos desde la posición 27812 hasta 28882. De acuerdo con otra realización, el fago es lambda y el fragmento de ácido nucleico que se va a suprimir inicia en una posición que varía desde la posición 1 hasta la posición 27812; y finaliza en una posición que varía desde 28882 hasta 33929.

En una realización, se suprime un fragmento de ácido nucleico que comprende la secuencia de codificación del gen Xis.

De acuerdo con una realización, el fago es lambda y el fragmento de ácido nucleico suprimido que se va a suprimir comprende los residuos desde la posición 28860 hasta 29078. De acuerdo con otra realización, el fago es lambda y

el fragmento inicia en una posición que varía desde la posición 1 hasta la posición 28860, y finaliza en una posición que varía desde la posición 29078 hasta la posición 33929.

De acuerdo con otra realización, el fago es DE3 y el ácido nucleico suprimido comprende los residuos desde la posición 5586 hasta 5804. De acuerdo con otra realización, el fragmento que se va a suprimir inicia en una posición que varía desde la posición 1 hasta la posición 5586, y finaliza en una posición que varía desde la posición 5804 hasta la posición 10812.

En una realización, se suprime un fragmento de ácido nucleico que comprende las secuencias de codificación de los genes Xis e Int. De acuerdo con una realización, el fago es lambda y el fragmento de ácido nucleico que se va a suprimir comprende los residuos desde la posición 27812 hasta 29078. De acuerdo con otra realización, el fago es lambda y el fragmento de ácido nucleico que se va a suprimir inicia en una posición que varía desde la posición 1 hasta la posición 27812; y finaliza en una posición que varía desde 29078 hasta 33929.

En una realización de la invención, el fago genéticamente modificado de la invención se inactiva adicionalmente para el gen kil. De acuerdo con una realización, el gen kil tiene la secuencia SEQ ID NO: 36. En otra realización, el gen kil presenta una identidad de secuencia de por lo menos 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, más preferiblemente de por lo menos 99% con la SEQ ID NO: 36. Preferiblemente, se suprime la secuencia de codificación del gen kil. De acuerdo con una realización, el fago es lambda y el fragmento de ácido nucleico suprimido comprende los residuos desde la posición 33187 hasta 33331.

De acuerdo con otra realización, el fago es DE3 y el ácido nucleico suprimido comprende los residuos desde la posición 9913 hasta 10057.

De acuerdo con una realización, la secuencia attP no se suprime de la secuencia del fago de la invención. La secuencia attP es la SEQ ID NO: 9. La secuencia attP se ubica desde la posición 27586 y 27817 en la secuencia del fago lambda, y desde la posición 42788 hasta 42925 y desde 1 hasta 94 en el genoma del fago Lambda DE3 (el genoma del fago es circular, de esta manera la secuencia attP es continua).

En una realización de la invención, el fago genéticamente modificado comprende la secuencia attP y la secuencia del gen C2. En otra realización, el fago genéticamente modificado consiste de la secuencia attP y la secuencia del gen C2.

De acuerdo con una realización preferida, un fago genéticamente modificado de la invención P11 tiene la secuencia SEQ ID NO: 10 y es (DE3) ΔS-C, Δxis-ea10 (DE3 se refiere al fago lambda DE3 en el que el gen de polimerasa de ARN T7 se ha integrado dentro de la secuencia de gen Int). Dicho fago P11 modificado corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, R, Rz, Xis y Int.

De acuerdo con otra realización, un fago genéticamente modificado de la invención P12 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes Int y S (un ejemplo de P12 es la secuencia DE3 ΔS, Δxis-ea10).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P13 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, R, Rz, Q, Xis y Int (un ejemplo de P13 es la secuencia DE3 ΔS-C, Δxis-ea10, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P14 corresponde a la secuencia NC_0014l6 en la que se suprimen las secuencias de codificación de los genes R y Int (un ejemplo de P14 es la secuencia DE3 ΔR).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P15 corresponde a la secuencia NC_0014l6 en la que se suprimen las secuencias de codificación de los genes Q y Int (un ejemplo de P15 es la secuencia DE3 ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P16 corresponde a la secuencia NC_0014l6 en la que se suprimen las secuencias de codificación de los genes S y Xis (un ejemplo de P16 es la secuencia NC_001416 ΔS, Δxis-ea10).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P17 corresponde a la secuencia NC_0014l6 en la que se suprimen las secuencias de codificación de los genes R y Xis (un ejemplo de P17 es la secuencia NC_001416 ΔR, Δxis-ea10).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P18 corresponde a la secuencia NC_0014l6 en la que se suprimen las secuencias de codificación de los genes Q y Xis (un ejemplo de P18 es la secuencia NC_001416 Δxis-ea10, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención p19 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes R, Rz y Int (un ejemplo de P19 es la secuencia DE3 ΔR, ΔRz).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P20 corresponde a la secuencia NC_0014l6 en la que se suprimen las secuencias de codificación de los genes S, Rz y Int (un ejemplo de P20 es la secuencia DE3 ΔS, ΔRz).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P21 corresponde a la secuencia NC_0014l6 en la que se suprimen las secuencias de codificación de los genes Rz, Q y Int (un ejemplo de P21 es la secuencia DE3 ΔRz, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P22 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, Q y Int (un ejemplo de P22 es la secuencia DE3 ΔS, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P23 corresponde a la secuencia NC_0014l6 en la que se suprimen las secuencias de codificación de los genes R, Q, y Int (un ejemplo de P23 es la secuencia DE3 ΔR, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P24 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, R y Int (un ejemplo de P24 es la secuencia DE3 ΔS, ΔR).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P25 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes R, Rz y Xis (un ejemplo de P25 es la secuencia NC_001416 ΔR, Δxis-ea10, ΔRz).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P26 corresponde a la secuencia NC_0014l6 en la que se suprimen las secuencias de codificación de los genes S, Rz y Xis (un ejemplo de P26 es la secuencia NC_001416 ΔS, Δxis-ea10, ΔRz).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P27 corresponde a la secuencia NC_0014l6 en la que se suprimen las secuencias de codificación de los genes Q, Rz y Xis (un ejemplo de P27 es la secuencia NC_0014l6 ΔRz, Δxis-ea10, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P28 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, Q y Xis (un ejemplo de P28 es la secuencia NC_001416 ΔS, Δxis-ea10, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P29 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes R, Q y Xis (un ejemplo de P29 es la secuencia NC_001416 ΔR, Δxis-ea10, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P30 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes R, S y Xis (un ejemplo de P30 es la secuencia NC_001416 ΔS, Δxis-ea10, ΔR).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P31 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes R, Xis y Int (un ejemplo de P31 es la secuencia DE3 ΔR, Δxis-ea10).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P32 corresponde a la secuencia NC_0014l6 en la que se suprimen las secuencias de codificación de los genes S, Xis y Int (un ejemplo de P32 es la secuencia DE3 ΔS, Δxis-ea10).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P33 corresponde a la secuencia NC_0014l6 en la que se suprimen las secuencias de codificación de los genes Q, Xis y Int (un ejemplo de P33 es la secuencia DE3 Δxis-ea10, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P34 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, R, Xis y Int (un ejemplo de P34 es la secuencia DE3 ΔS, Δxis-ea10, ΔR).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P35 corresponde a la secuencia NC_0014l6 en la que se suprimen las secuencias de codificación de los genes R, Q, Xis y Int (un ejemplo de P35 es la secuencia DE3 ΔR, Δxis-ea10, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P36 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, Q, Xis y Int (un ejemplo de P36 es la secuencia DE3 ΔS, Δxis-ea10, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P37 corresponde a la secuencia NC_0014l6 en la que se suprimen las secuencias de codificación de los genes R, Rz, Xis y Int (un ejemplo de P37 es la secuencia DE3 ΔR, Δxis-ea10, ΔRz).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P38 corresponde a la secuencia NC_0014l6 en la que se suprimen las secuencias de codificación de los genes S, Rz, Xis y Int (un ejemplo de P38 es la secuencia DE3 ΔS, Δxis-ea10, ΔRz).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P39 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes Rz, Q, Xis y Int (un ejemplo de P39 es la secuencia DE3 ΔRz, Δxis-ea10, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P40 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, R, Q y Int (un ejemplo de P40 es la secuencia DE3 ΔS, ΔR, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P41 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, R, Rz y Int (un ejemplo de P41 es la secuencia DE3 ΔS-C).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P42 corresponde a la secuencia NC_0014l6 en la que se suprimen las secuencias de codificación de los genes R, Rz, Q y Int (un ejemplo de P42 es la secuencia DE3 ΔR, ΔRz, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P43 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, Rz, Q y Int (un ejemplo de P43 es la secuencia DE3 ΔS, ΔRz, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P44 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, R, Q y Xis (un ejemplo de P44 es la secuencia NC_001416 ΔS, ΔR, Δxis-ea10, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P45 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, R, Rz y Xis (un ejemplo de P45 es la secuencia NC_001416 ΔS-C, Δxis-ea10).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P46 corresponde a la secuencia NC_0014l6 en la que se suprimen las secuencias de codificación de los genes R, Rz, Q y Xis (un ejemplo de P46 es la secuencia NC_0014l6 ΔR, ΔRz, Δxis-ea10, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P47 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, Rz, Q y Xis (un ejemplo de P47 es la secuencia NC_0014l6 ΔS, ΔRz, Δxis-ea10, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P48 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, R, Q, Xis y Int (un ejemplo de P48 es la secuencia DE3 ΔS, ΔR, Δxis-ea10, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P49 corresponde a la secuencia NC_0014l6 en la que se suprimen las secuencias de codificación de los genes S, Rz, Q, Xis y Int (un ejemplo de P49 es la secuencia DE3 ΔS, ΔRz, Δxis-ea10, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P50 corresponde a la secuencia NC_0014l6 en la que se suprimen las secuencias de codificación de los genes R, Rz, Q, Xis y Int (un ejemplo de P50 es la secuencia DE3 ΔR, ΔRz, Δxis-ea10, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P51 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, R, Q, Rz y Int (un ejemplo de P51 es la secuencia DE3 ΔS-C, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P52 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, R, Rz, Q y Xis (un ejemplo de P52 es la secuencia NC_001416 ΔS-C, Δxis-ea10, ΔQ).

De acuerdo con otra realización, un fago genéticamente modificado de la invención P53 corresponde a la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes Q, Xis y Int (un ejemplo de P53 es la secuencia DE3 Δxis-ea10, ΔQ).

La presente invención también se refiere a un proceso para producir el fago modificado de la invención, en la que dicho proceso comprende por lo menos dos etapas de supresión de genes.

En una realización, el proceso de la invención se lleva a cabo con un fago integrado en el genoma de una célula anfitriona.

En una realización, la célula anfitriona es un microorganismo, preferiblemente un procariota, más preferiblemente una bacteria, más preferiblemente una bacteria gram negativa. Ventajosamente, la célula anfitriona es una bacteria de la familia Enterobacteriacea de acuerdo con la taxonomía aplicable actual. Si la taxonomía cambiara, el experto en la técnica sabría cómo adaptar los cambios en la taxonomía para deducir las cepas que se podrían utilizar en la presente invención. Ejemplos de bacterias de la familia Enterobacteriacea incluyen, pero no se limitan a, bacterias pertenecientes a los géneros Escherichia, Enterobacter, Erwinia, Klebsiella, Pantoea, Photorhabdus, Providencia, Salmonella, Serratia, Shigella, Morganella y Yersinia. De acuerdo con una realización preferida, la célula anfitriona pertenece al género Escherichia, y más preferiblemente la célula anfitriona es Escherichia coli (E. coli).

Los métodos para eliminar genes de un fago integrado son bien conocidos por los expertos en la técnica. Ejemplos de dichos métodos incluyen, pero no se limitan a, la recombinación homóloga (también llamada recombinación) utilizando los genes codificados Red lambda: exo, bet y gam. También es posible utilizar los genes recE y recT correspondientes del profago Rac. Los genes exo y recE codifican una exonucleasa 5'-3' que produce salientes de 3'. Los genes bet y recT codifican una proteína de emparejamiento o también llamada proteína de hibridación que se une a las salientes 3' y media su hibridación y recombinación homóloga entre dos secuencias de ADN complementarias. El gen gam codifica un inhibidor de la exonucleasa RecBCD de E. coli y, por lo tanto, protege los fragmentos de ADN lineales de interés. El método de recombinación fue bien descrito por varios investigadores, incluidos Datsenko and Wanner (PNAS 97-12, 6640-6645, 2000) y Stewart et al (WO0104288). El principio del método es generar (mediante amplificación por PCR, por ejemplo) un fragmento de ADN que contenga el fragmento a integrar y dos brazos de recombinación. Estos brazos son homólogos a las regiones adyacentes al gen que se va a inactivar. Se utilizarán para apuntar a la inserción del fragmento de interés. Es posible crear este tipo de fragmento de ADN mediante PCR utilizando cebadores que contienen brazos homólogos de 20 a 60 nucleótidos.

La presente invención también se refiere a una célula anfitriona que comprende el fago genéticamente modificado de la invención. En una realización preferida, la célula anfitriona de la invención comprende el fago genéticamente modificado integrado en su genoma.

En una realización, la célula anfitriona es un microorganismo, preferiblemente un procariota, más preferiblemente una bacteria, más preferiblemente una bacteria gram negativa. Ventajosamente, la célula anfitriona es una bacteria de la familia Enterobacteriacea de acuerdo con la taxonomía aplicable actual. Si la taxonomía cambiara, el experto en la materia sabría cómo adaptar los cambios en la taxonomía para deducir las cepas que se podrían utilizar en la presente invención. Ejemplos de bacterias de la familia Enterobacteriacea incluyen, pero no se limitan a, bacterias pertenecientes a los géneros Escherichia, Enterobacter, Erwinia, Klebsiella, Pantoea, Photorhabdus, Providencia, Salmonella, Serratia, Shigella, Morganella y Yersinia. De acuerdo con una realización preferida, la célula anfitriona pertenece al género Escherichia, y más preferiblemente la célula anfitriona es Escherichia coli (E. coli). Ejemplos de cepas de E. coli que se podrían utilizar en la presente invención incluyen, pero no se limitan a, cepas derivadas de E. coli K-12, E. coli B o E. coli W, tales como, por ejemplo, MG1655, W3110, DG1, DG2, Top10, DH10B, DH5alpha, HMS174, BL21, BL21(DE3), HMS174(DE3), BL21(DE3) pLysS, BL21(DE3) pLysE.

En una realización, se pueden inactivar los genes de las células anfitrionas.

En una realización, se inactiva el gen tonA (también conocido como fhuA, SEQ ID NO: 11). La proteína TonA/FhuA es un receptor para los fagos T1, T5 y Phi80.

En una realización, se inactiva el gen galK (SEQ ID NO: 12). La supresión de teste gen permite el uso de la selección positiva/negativa de galK para la supresión de los genes mediante un método con base en la recombinación homóloga.

En una realización, se inactiva el gen araB (SEQ ID NO: 13). En otra realización, se inactiva el gen araA (SEQ ID NO: 14). La inactivación de araB y/o araA se recomienda para el uso del promotor PBAD (inducible por arabinosa) dentro de la célula anfitriona.

En una realización, se inactivan el gen lon (SEQ ID NO: 15) y/o el gen ompT (SEQ ID NO: 16). La proteína Lon es una proteasa dependiente de ATP. La proteína OmpT es una proteasa de membrana externa. Preferiblemente, se inactivan los genes lon y ompT.

En una realización, se inactiva el gen rcsA (SEQ ID NO: 17). La proteína RcsA es un regulador positivo de la síntesis de la cápsula, que se degrada por la proteasa Lon. En una realización, se inactivan el gen hsdR (SEQ ID NO: 18) y/o el gen mrr (SEQ ID NO: 19). Las proteínas HsdR y Mrr son enzimas de restricción con diferente especificidad. Preferiblemente, los genes hsdR y mrr ambos son inactivados.

En una realización, se inactivan el gen endA (SEQ ID NO: 20) y/o el gen recA (SEQ ID NO: 21). EndA es una endonucleasa específica a ADN. RecA es una proteína de recombinación con actividad de proteasa y nucleasa. Preferiblemente, se inactivan ambos genes endA y recA.

En una realización de la invención, por lo menos se inactivan uno de los genes tonA, galK, araB, araA, lon, ompT, rcsA, hsdR, mrr, endA y recA. Preferiblemente, se suprimen inactivados.

En una realización preferida, se inactivan los genes tonA, galK, araB, lon, ompT, rcsA, hsdR, mrr, endA y recA. Preferiblemente, se suprimen los genes tonA, galK, araB, lon, ompT, rcsA, hsdR, mrr, endA y recA.

En una realización de la invención, se transforma la célula anfitriona de la invención con una secuencia de ácidos nucleicos que codifica una molécula tóxica, como se describe en US2004/0115811. La presencia de la secuencia de ácidos nucleicos que codifica la molécula tóxica permitirá la selección de clones recombinantes que tienen integrados un gen de interés y una secuencia de nucleótidos que codifica una proteína de antídoto funcional a una molécula tóxica, en la que dichos clones recombinantes son los que sobreviven luego de su integración en una célula anfitriona que comprende en su genoma una secuencia de nucleótidos que codifica dicha molécula tóxica.

De acuerdo con una realización preferida de la presente invención, la proteína antídoto y la molécula tóxica son, respectivamente, una proteína antiveneno y una proteína venenosa. Dichas proteínas antivenenosas o venenosas podrían ser proteínas de tipo silvestre o proteínas modificadas que son natural o artificialmente venenosas y afectan una o más funciones vitales de una célula (preferiblemente, una célula procariota) y pueden conducir a la destrucción de la célula.

La proteína antídoto y la molécula tóxica se seleccionan preferiblemente del grupo que consiste en proteínas CcdA/CcdB, proteínas Kis/Kid, proteínas Phd/Doc, proteínas RelB/relE, proteínas PasB (o PasC)/PasA, proteínas mazF/mazE como se describe en el documento US2004/0115811, o cualquier otro par de moléculas antivenenosas/venenosas que sean o no de origen plásmido. La molécula tóxica también puede ser una proteína de toxina que es natural o artificialmente tóxica y afecta a una o más funciones vitales de una célula (procariota). La proteína codificada por el gen sacB (de Bacillus amylolique-faciens), la proteína GpE, la proteína GATA-1 y la proteína Crp son otros ejemplos de dichas moléculas tóxicas. El gen sacB codifica la sacarasa de levano que cataliza la hidrólisis de la sacarosa en productos que son tóxicos para E. coli (Pierce et al. Proc. Natl Acad. Sci., Vol. 89, N[deg.]6 (1992) p. 2056-2060). La proteína GpE codifica los genes E del bacteriófago [phi] X174 que incluye seis sitios de restricción únicos y codifica gpE y que causa lisis de las células de E. Coli (Heinrich et al., Gene, Vol. 42 (3) (1986) p. 345- 349). Se ha descrito la proteína GATA-1 por Trudel et al. (Biotechniques 1996, vol. 20 (4), p. 684-693). La proteína Crp ha sido descrita por Schlieper et al. (Anal. Biochem. 1998, Vol. 257 (2), pág. 203-209).

Las proteínas antídotos contra dicha molécula tóxica son cualquier proteína capaz de reducir o suprimir el efecto de la molécula tóxica correspondiente sobre una célula (preferiblemente una célula procariota), cuando dicha célula tóxica es producida por dicha célula.

De acuerdo con una realización preferida, la célula anfitriona de la invención comprende una secuencia de ácidos nucleicos que codifica la proteína CcdB. El gen ccdB tiene la secuencia SEQ ID NO: 22.

La presente invención también se refiere a un kit que comprende una célula anfitriona como se describió anteriormente, en la que se inserta el gen que codifica una proteína venenosa, y un plásmido que lleva la secuencia de ácidos nucleicos del gen que codifica la proteína antivenenosa. La expresión de la proteína antivenenosa en la célula anfitriona se requiere para mantener la viabilidad de la célula anfitriona. En una realización preferida, la proteína venenosa se codifica por el gen ccdB, y la proteína antivenenosa se codifica por el gen ccdA (SEQ ID NO: 23).

De acuerdo con una realización de la invención, el plásmido del kit adicionalmente puede contener, o se puede modificar para contener adicionalmente, la secuencia de ácidos nucleicos de la biomolécula de interés.

La presente invención también se refiere a un proceso para preparar una célula anfitriona como se describió anteriormente.

En una realización, el proceso de la invención comprende una etapa de infección de la célula anfitriona por un fago genéticamente modificado de acuerdo con la invención. En una realización de la invención, dicha etapa de infección incluye el uso de un fago ayudante. En el significado de la presente invención, el término "fago auxiliar" se refiere a un fago utilizado para complementar una eliminación o una inactivación de otro fago. El fago auxiliar proporcionará las funciones faltantes a otro fago para poder infectar bacterias o para preparar la reserva de fago. Usualmente, el fago auxiliar no puede formar un lisógeno por si mismo porque es cl menos (no tiene represor y, por lo tanto, es virulento).

El proceso para infectar una célula anfitriona con un fago que utiliza un fago auxiliar es bien conocido en la técnica. La primera etapa es la preparación de los lisados y la segunda es la lisogenización. Brevemente, los lisados bacterianos del fago auxiliar se preparan utilizando métodos estándar como se describe en "Molecular cloning: a laboratory manual", Sambrook et al (2001, ISBN 978-087969577-4) o en "Large- and Small-Scale Preparation of Bacteriophage lambda lysate and DNA", Su et al, BioTechniques 25:44-46 (Julio 1998). La preparación del fago de interés se realiza utilizando el mismo principio, después de la inducción del fago (la mayoría de las veces se utiliza radiación UV o cualquier situación en la que un lisógeno sufra daños en el ADN o la respuesta SOS del anfitrión o la producción de Cro) con el fin iniciar el ciclo lítico y utilizar un fago auxiliar para proporcionar las funciones que faltan. Una alternativa al fago auxiliar es el uso de un plásmido que codifica las funciones faltantes.

A continuación, los lisados de fagos se mezclan con las bacterias seleccionadas y se colocan sobre placas LB con el fin de obtener lisógenos (como se describe en el kit de lisogenización DE3 de lambda de Novagen, User Protocol TB031 o se describe un método alternativo en Studier and Moffat, Journal of Molecular Biology, 1986, 189: 113-130). Se puede utilizar un fago de selección para seleccionar específicamente bacterias que contienen el fago de interés. Este fago de selección es un fago virulento que tiene la misma inmunidad que el fago de interés. En consecuencia, el fago de selección no puede formar placas o matar bacterias lisógenas para el fago de interés porque este fago produce el represor cl (también llamado C2 en el fago lambda DE3).

La presente invención también se refiere, por lo tanto, a un kit que comprende el fago modificado de la invención, como se describió anteriormente, y un fago auxiliar. Ejemplos de fagos auxiliares incluyen, pero no se limitan a, el fago auxiliar B10 o cualquier otro fago lambdoide con una inmunidad diferente a la del fago de interés (ejemplo: fago con inmunidad 434, 80,...). Algunos fagos auxiliares y sus manipulaciones fueron bien descritos en la literatura, incluyendo en Haldimann and Wanner (Journal of Bacteriology, 183, 6384-6393, 2001).

En una realización, el proceso para preparar una célula anfitriona de la invención comprende adicionalmente una etapa de supresión, en la que se suprimen las secuencias de ácido nucleico de la célula anfitriona. Los métodos para suprimir la secuencia de un gen son bien conocidos por los expertos en la técnica. El método más eficiente es el método de recombinación homóloga mediado por los genes codificados en Red lambda o los genes recE y recT del profago Rac. Como se describió anteriormente, este método fue bien descrito por varios investigadores, incluidos Datsenko and Wanner (PNAS 97-12, 6640-6645, 2000) y Stewart et al (WO0104288). Los productos de PCR se generan utilizando cebadores con extensiones de 20 a 60 nt que son homólogas a las regiones adyacentes al gen que se va a inactivar. Dado que solo una pequeña cantidad de bacterias recombinará efectivamente el fragmento de interés, es necesario tener un marcador de selección fuerte para seleccionarlo. Se pueden utilizar marcadores de antibióticos para seleccionar los recombinantes: los cebadores modificados se utilizan para amplificar un gen de resistencia a antibióticos. Después de la transformación y activación de los genes de recombinación, las bacterias recombinantes se seleccionan en un medio que contiene el antibiótico apropiado. En este caso, el gen objetivo se reemplaza por un gen de resistencia a antibióticos. Con el fin utilizar la misma estrategia para la próxima supresión, es necesario eliminar este gen de resistencia a los antibióticos durante una segunda etapa. Como se describe en Datsenko and Wanner, es posible utilizar un gen de resistencia a antibióticos que se flanquea por sitios FRT (objetivo de reconocimiento de FLP). Los genes de resistencia se eliminan al utilizar un plásmido auxiliar que codifica la recombinasa FLP. El gen de resistencia a los antibióticos se elimina con esta recombinasa específica del sitio, pero este método deja rastros: un sitio de recombinación específico del sitio aún está presente después de la eliminación del gen de resistencia a los antibióticos.

Para evitar la presencia de este sitio, más preferiblemente, el método de la invención utiliza galK como un gen marcador. El principio de la selección de galK se describe en Warming et al. (Nucleic acid research, 2005, 33(4)). Este método utiliza galK como marcador de selección positiva (crecimiento sobre medio mínimo que contiene galactosa) durante la primera recombinación (inserción). La eliminación de este marcador se realiza durante una segunda etapa de recombinación homóloga. Durante esta etapa, galK se utiliza como marcador de selección negativa en medio mínimo que contiene 2-desoxi-galactosa (DOG). El producto del gen galK, la galactoquinasa, cataliza la primera etapa en la ruta de degradación de la galactosa. La galactoquinasa también cataliza eficientemente la fosforilación del análogo de galactosa DOG. El producto de esta reacción no se puede metabolizar adicionalmente, lo que lleva a la acumulación de una molécula tóxica (2-desoxi-galactosa-1-fosfato). La ventaja de este método es evitar la presencia de un sitio de recombinación específico después de la supresión del gen objetivo y la eliminación del marcador selectivo.

La presente invención también se refiere a un proceso para producir una biomolécula de interés, que comprende

- cultivar una célula anfitriona que comprende el fago genéticamente modificado de acuerdo con la invención y la secuencia de ácidos nucleicos de la biomolécula de interés,
- recuperar la biomolécula de interés.

En una realización de la invención, la secuencia de ácidos nucleicos de la biomolécula de interés está comprendida dentro del sistema de expresión del fago genéticamente modificado. De acuerdo con esta realización, la producción de la biomolécula de interés es directa, es decir, resulta de la expresión del gen del sistema de expresión, por ejemplo mediante cultivo en un medio en el que se induce el promotor comprendido en el sistema de expresión.

En otra realización de la invención, el sistema de expresión del fago genéticamente modificado comprende la secuencia de ácidos nucleicos de la polimerasa de ARN T7 bajo el control de un promotor lac, preferiblemente el promotor lacUV5. De acuerdo con esta realización, el proceso para producir la biomolécula de interés comprende la transformación de la célula anfitriona con un plásmido que comprende la secuencia de ácidos nucleicos de la biomolécula de interés bajo el control de un promotor T7. La expresión del promotor T7 está bajo el control de polimerasa de ARN T7, que es rigurosamente específica para el promotor T7, es decir, el promotor T7 solo puede ser utilizado por la polimerasa de ARN del bacteriófago T7. Cuando se agrega IPTG al medio de cultivo, la polimerasa de ARN de T7 se expresa por transcripción del promotor lac y permite la expresión de la biomolécula de interés.

En el significado de la presente invención, el término "promotor T7" incluye promotores que están presentes en el genoma del bacteriófago T7, así como secuencias de consenso y variantes de dichos promotores con la capacidad de mediar la transcripción por la polimerasa de ARN T7. El bacteriófago T7 contiene diecisiete secuencias promotoras diferentes, todas las cuales comprenden una secuencia de nucleótidos altamente conservada.

De acuerdo con una realización preferida, el plásmido que comprende la secuencia de ácidos nucleicos de la biomolécula de interés también comprende la secuencia de ácidos nucleicos de ccdA, y la célula anfitriona comprende la secuencia de ccdB integrada en su genoma. Por lo tanto, solo se propagan clones recombinantes que contienen el plásmido.

De acuerdo con una realización, se secreta la biomolécula de interés por la célula anfitriona en el caldo de fermentación. De acuerdo con esta realización, la biomolécula de interés se puede recuperar fácilmente del caldo de fermentación.

De acuerdo con otra realización, la biomolécula de interés no es secretada por la célula anfitriona en el caldo de fermentación. Los métodos para recuperar una biomolécula intracelular de interés son bien conocidos en la técnica. Ejemplos de dicho método incluyen, pero no se limitan a, el uso de ácido tricloroacético (TCA) o tampón de craqueo que contiene dodecilsulfato de sodio (SDS) con el fin de recuperar proteínas citoplásmicas totales en condiciones de desnaturalización o el uso de sonicación, prensa francesa o equivalente para interrumpir bacterias bajo presión para recuperar proteínas citoplásmicas totales en condiciones nativas (no desnaturalizantes). A continuación, la proteína de interés se puede purificar utilizando métodos específicos que incluyen, pero no se limitan a, el uso de columnas de afinidad o de intercambio iónico.

Breve descripción de los dibujos

Figura 1: Imagen de un gel SDS-Page coloreado con tinción con azul de Coomassie, que muestra la producción de la proteína de interés. 1 y 6: tamaño; 2: sedimento [pScherryl] M11DE3 antes de inducción; 3: sedimento [pScherryl] M11DE3 después de inducción; 4: sedimento [pScherryl] HMS174DE3 antes de inducción; 5: sedimento [pScherryl] HMS174DE3 después de inducción. La matriz es para identificar la proteína de interés.

Ejemplos

La presente invención se ilustra adicionalmente mediante los siguientes ejemplos.

Ejemplo 1: Depresión de los genes xis, exo, bet, gam, kit, cIII, N y ral

a) Supresión de la región del ADN de xis y reemplazo por galK

Primero, el gen galK se amplificó mediante reacción en cadena de polimerasa (PCR) sobre el plásmido pGalK utilizando los cebadores de inicio (5' GGGGGTAAATCCCGGCGCTCATGACTTCGCCTTCTTCCCAGAATTCCTGTTGAC AATTA 3', SEQ ID NO: 24) y de parada XisgalK (5' GTTCTGATTATTGGAAATCTTCTTTGCCCTCCAGTGTGAGCAGCACTGTCCTGCTCCTTG 3', SEQ ID NO: 25). Estos cebadores contienen en el extremo 5' una secuencia de 40 bases idéntica al objetivo de ADN (en cursiva). Estas secuencias de 40 bases son los brazos de recombinación. Los extremos 3' se diseñaron

para amplificar el gen galK y su promotor constitutivo. El fragmento de ADN amplificado por PCR (1315 pb) se dirigió a los genes xis, exo, bet, gam, kil y cIII (fragmento de ADN de 5133 pb): estos genes se reemplazaron durante la recombinación homóloga por el gen galK y su promotor. Las bacterias electrocompetentes que llevan el profago T7 (DE3) y el plásmido pKD46 se prepararon de acuerdo con Datsenko y Wanner (PNAS 97-12, 6640-6645, 2000). A continuación, el fragmento galK amplificado se electroporó en estas bacterias de acuerdo con los procedimientos estándar (se utilizaron 200 ng de fragmento de ADN para cada electroporación). Se agregó medio SOC y se incubaron bacterias durante 1 hora a 37ºC. A continuación, las bacterias se lavaron (se centrifugaron, se retiraron los medios, se agregaron medios frescos y se volvieron a suspender) dos veces con medio mínimo M9 (Sambrook et al (2001, ISBN 978-087969577-4)) y se colocaron en placas bacterianas que contenían medio M9 mínimo y galactosa al 1%. Las placas se incubaron a 37ºC durante 1 o 2 días.

La siguiente etapa fue el cribado bacteriano: el cribado por PCR se realizó directamente sobre colonias utilizando los cebadores Xis1 (5'GTCTTCAAGTGGAGCATCAG3', SEQ ID NO: 26) y Xis4 (5'ACCAGGACTATCCGTATGAC3', SEQ ID NO: 27). Una amplificación de un fragmento de ADN de 5774 pb corresponde a un cromosoma no modificado (colonia no recombinante) y, por el contrario, una amplificación de un fragmento de ADN de 1955 pb corresponde a un cromosoma recombinante. Las bacterias que permitían la amplificación del fragmento de ADN de 1955 pb se seleccionaron y se sembraron dos veces en placas selectivas (medio M9 mínimo suplementado con 1% de galactosa) para purificarlo y eliminar posibles copias no modificadas del cromosoma. El cribado por PCR se realizó una vez más al final de la etapa de purificación y se seleccionaron 3 bacterias que permiten la amplificación del fragmento de ADN de 1955 pb. Los fragmentos de ADN amplificados correspondientes a estas bacterias se secuenciaron utilizando los mismos cebadores con el fin de confirmar la recombinación del ADN y la supresión de la región del ADN Xis (genes xis, exo, bet, gam, kil y cIII).

b) Eliminación de GalK

Un fragmento de ADN que contiene grandes brazos de recombinación (de 350 pb y 343 pb) se construyó mediante PCR para eliminar los genes galK y N y ral. El primer brazo se amplificó por PCR sobre colonias bacterianas que contenían el profago T7 (DE3) utilizando los cebadores Xis1 y Xis2 (5'CCAAACGGAACAGATGAAGAAGGCGAAGTCATGAG3', SEQ ID NO: 28). Se amplificó un fragmento de ADN de 365 bases. El segundo brazo también se realizó por PCR sobre la misma bacteria utilizando los cebadores Xis3 (5' GACTTCGCCTTCTTCATCTGTTCCGTTTGGCTTCC3', SEQ ID NO: 29) y Xis6 (5'GTAATGGAAAGCTGGTAGTCG3', SEQ ID NO: 30). Se amplificó un ADN de fragmento de 358 bases. Ambos brazos de recombinación se purificaron después de la electroforesis en gel de agarosa. Los cebadores Xis2 y Xis3 fueron diseñados para generar fragmentos de ADN que contienen una secuencia idéntica de 30 pares de bases. Esta secuencia se utilizó para unir ambos brazos de recombinación en una tercera PCR utilizando los cebadores Xisl y Xis6. Se generó un fragmento de ADN de 693 pb. Este fragmento de ADN se electroporó en bacterias seleccionadas anteriormente y que llevan el plásmido pKD46 (preparado como se describe en Datsenko and Wanner). Se agregó medio SOC y se incubaron bacterias a 37ºC durante 1 hora. A continuación, las bacterias se lavaron dos veces con medio M9 y se colocaron en placas selectivas que contenían medio M9 suplementado con glicerol al 0.2% y DOG al 1%. Las placas se incubaron durante 2 días a 37ºC. Se cribaron varias colonias por PCR utilizando los cebadores Xis5 (5 CAGCCGTAAGTCTTGATCTC3', SEQ ID NO: 31) y Xis7 (5'CAGCAGGCATGATCCAAGAG3', SEQ ID NO: 32). Siempre se obtuvo una amplificación de 3246 pb correspondiente al cromosoma de ADN no modificado en lugar de una amplificación de 1122 pb correspondiente al cromosoma modificado. El experimento se reprodujo completamente e independientemente tres veces sin éxito: no se obtuvo ninguna bacteria que comprenda la supresión deseada.

En consecuencia, decidimos eliminar solo el fragmento GalK y dejar los genes N y ral. El fragmento de ADN que contiene los brazos de recombinación se generó como se describió anteriormente utilizando Xis, 1 Xis2b (5' TTTGCCCTCCAGTGTGAAGAAGGCGAAGTCATGAG3', SEQ ID NO: 33) para el primer brazo de recombinación (365bp) y Xis8 (5' CTCATGACTTCGCCTTCTTCACACTGGAGGGCAAAGAAG, SEQ ID NO: 34) y Xis4 para el segundo brazo de recombinación. La PCR de unión se realizó utilizando cebadores Xisl y Xis4 y generó un fragmento de ADN de 714 pb. Este fragmento se electroporó como se describió anteriormente en las bacterias seleccionadas y que llevan el plásmido pKD46. Las bacterias se colocaron en placas en las mismas placas selectivas que contenían DOG y se incubaron durante dos días a 37ºC. El cribado por PCR se realizó utilizando cebadores Xis5 y Xis6. Las bacterias que muestran una amplificación de 1770 pb (en lugar de una amplificación de 3010 pb correspondiente al cromosoma no modificado) se seleccionaron y purificaron. El fragmento de ADN se secuenció utilizando los cebadores Xis5 y Xis6 y mostró la eliminación correcta del gen galK. Esta recombinación se realizó una sola vez, ya que las bacterias se obtuvieron inmediatamente.

Estos resultados muestran que no fue posible eliminar galK asociado a los genes N y ral mediante recombinación homóloga. Sin embargo, utilizando exactamente el mismo procedimiento, pudimos solo eliminar galK.

En conclusión, demostramos que la eliminación de los genes N y ral conduce a la muerte de bacterias, lo cual fue inesperado.

Ejemplo 2: Expresión de proteínas utilizando MG1655

Con el fin de probar la eficiencia de producción de proteínas de la cepa construida de acuerdo con la invención, se realizaron pruebas de expresión a pequeña escala utilizando bacterias llamadas M11 (DE3). El genotipo de esta cepa es MG1655 ΔgalK, Δrcs A, Δlon, ΔhsdR-mmr, ΔfhuA, ΔendA, ΔrecA, ΔaraB, ΔompT, λ-DES (T7pol, Δxis-ea10, ΔS-C). Dado que MG1655 es una Escherichia coli K-12, se comparó con otra cepa K-12 utilizada como estándar en el campo de la producción de proteínas y llamada HSM174 (DE3) (genotipo: recA1, hsdR, λ-DE3, (Rif R)). Ambas cepas se transformaron con ADN de plásmido pSCherryl (Delphi Genetics, Bélgica). Este plásmido codifica una proteína llamada "cherry" fácilmente detectable (por los ojos, color rojo) bajo el control del promotor T7.

Protocolo para una expresión a pequeña escala utilizando IPTG:

1) Se inocularon dos matraces Erlenmeyer que contenían 10 ml de medio LB, cada uno con una colonia única de HMS174 (DE3) y M11 (DE3) que llevaban el plásmido pSCherryl como se incubaron a 30ºC durante la noche.
2) Dos nuevos matraces que contenían medio fresco se inocularon con 1 ml de los cultivos durante la noche y se incubaron con agitación a 37ºC hasta que la OD600 alcanzó 0.6.
3) Se tomó una muestra (1 ml) de cada matraz y se centrifugó. El medio se descartó y el sedimento se mantuvo en hielo. Las muestras fueron los controles no inducidos. Para inducir la expresión de proteínas en el cultivo restante, se agregó IPTG (isopropil β-D-1-tiogalactopiranósido, 90 µl de una solución madre de 100 mM fresca) para alcanzar una concentración final de 1 mM en ambos matraces. La incubación de ambos matraces se continuó durante 4 horas.

4) Al final del período de inducción, se midió la densidad óptica a 600 nm para cada cultivo (1.09 para M11 (DE3) y 1.13 para HMS174 (DE3)). Una muestra de 1 ml de cada matraz se centrifuga a la velocidad máxima (13000 g) durante 10 minutos a 4ºC. Se observó que el sedimento era rojo de acuerdo con la expresión de la proteína Cherry. El sobrenadante se desechó y se agregaron 100 µl de agua para resuspender las bacterias. También se agregaron 100 µl de tampón de "craqueo" (DTT 100 mM, SDS al 2%, Tris-HCl 80 mM, pH 6.8, azul de bromofenol al 0.006%, glicerol al 15%) para lisar las bacterias. Las muestras no inducidas se trataron con el mismo protocolo, excepto que solo se utilizaron 60 µl de agua y 60 µl de tampón de craqueo (de acuerdo con la densidad óptica de las muestras).
5) Las muestras se calentaron a 70ºC-100ºC (10 min.) Para resuspender todas las proteínas y desnaturalizar las proteínas.
6) Se cargaron 10 µl de cada muestra en gel de SDS-PAGE al 12% y se migró durante 2 horas a 100 voltios.
7) Después de la migración, las proteínas se colorearon con tinción con azul de Coomassie.

Como se muestra en la figura 1, ambas cepas fueron capaces de producir la proteína de interés (proteína Cherry, indicada por la matriz), pero la producción es de aproximadamente 5 a 10 veces mayor al utilizar M11 (DE3) que al utilizar HMS174 (DE3). El experimento se realizó dos veces con exactamente los mismos resultados.

Ejemplo 3: Expresión de proteínas utilizando BL21 (DE3)

Las cepas BL21 (DE3) suprimidas de Axis-ea10 y/o ΔS-C se construyeron de acuerdo con el ejemplo 1. La supresión ΔS-C se insertó en el cromosoma de la bacteria (en lambda DE3 suprimida del gen int y codificaba la polimerasa de ARN de T7) sola o en combinación con la supresión Axis-ea10. Se construyeron de esta manera dos derivados BL21 (DE3): BL21 (DE3) ΔS-C y BL21 (DE3) Δxis-ea10, ΔS-C. Las supresiones se confirmaron mediante amplificación por PCR de la región modificada y esta región se secuenció. Esta etapa de secuenciación confirmó la presencia de las supresiones correspondientes de acuerdo con las secuencias teóricas. A continuación, las bacterias construidas se probaron para determinar la expresión de proteínas utilizando dos plásmidos diferentes que codifican la proteína cherry (pSCherryl y pSSC-Cherryl, Delphi Genetics) de acuerdo con el ejemplo 2. Los resultados mostraron claramente una sobreexpresión del gen cherry que codifica la proteína cherry. Esta proteína fue detectada fácilmente por los ojos y en el análisis SDS-PAGE.

Por lo tanto, llegamos a la conclusión de que estas bacterias son capaces de sobreproducir proteínas sin riesgo de lisis bacteriana no deseada debido al fago lambda DE3 porque este fago está por lo menos suprimido de los genes int, S, R y Rz que codifican las funciones requeridas para la división de fagos y lisis bacteriana.

Dado que la lisis bacteriana durante la proteína depende de varios parámetros, que incluyen las condiciones de crecimiento, la proteína sobreexpresada,... diseñamos una cepa modelo para mostrar que las cepas construidas son resistentes a la lisis debido a las supresiones de genes específicos. En el fago lambda, se requiere tener una expresión constitutiva del represor CI con el fin de reprimir el desarrollo lítico y mantener el estado lisogénico del profago. Un mutante (CI^{857}) de este represor es bien conocido debido a su capacidad para ser termosensible: a 30ºC, el represor mutado es completamente activo y mantiene el estado lisogénico. Sin embargo, a una temperatura más alta (37ºC a 42ºC), el CI^{857} no es eficiente e induce el estado lítico del fago lambda. Las cepas que llevan esta mutación CI^{857} y las supresiones de acuerdo con la invención se construyeron: BL21 (DE3) ΔS-C correspondiente a una cepa BL21 que comprende P41; BL21 (DE3) Axis-ea10 y ΔS-C correspondientes a una cepa BL21 que comprende P11; BL21 (DE3) Δxis-ea10, ΔS-C y ΔQ correspondientes a una cepa BL21 que comprende P13; BL21 (DE3) ΔQ correspondiente a una cepa BL21 que comprende P15 y BL21 (DE3) Δxis-ea10 y ΔQ correspondiente a una cepa BL21 que comprende P53.

Al cambiar la temperatura a 42ºC, se induce el estado lítico. No se observó lisis para las cepas de la invención. Adicionalmente, observamos que el rendimiento de producción de la proteína de interés (proteína cherry) se mejoró para las cepas de la invención en comparación con el control.

Listado de secuencias

<110> DELPHI GENETICS
Szpirer, CÃ©dric

<120> Fago genéticamente modificado y uso del mismo

<130> CV-153/PCT

<160> 36

<170> BiSSAP 1.0

<210> 1
<211> 324
<212> ADN
<213> Fago lambda de enterobacteria

<220>
<221> fuente
<222> 1..324
<223> /mol_tipo="ADN"/organismo="Fago lambda de enterobacteria"

<400> 1

atgaagatgc cagaaaaaca tgacctgttg gccgccattc tcgcggcaaa ggaacaaggc 60

atcggggcaa tccttgcgtt tgcaatggcg taccttcgcg gcagatataa tggcggtgcg 120

tttacaaaaa cagtaatcga cgcaacgatg tgcgccatta tcgcctggtt cattcgtgac 180

cttctcgact tcgccggact aagtagcaat ctcgcttata taacgagcgt gtttatcggc 240

tacatcggta ctgactcgat tggttcgctt atcaaacgct tcgctgctaa aaaagccgga 300

gtagaagatg gtagaaatca ataa 324

<210> 2
<211> 477
<212> ADN
<213> Fago lambda de enterobacteria

<220>
<221> fuente
<222> 1..477
<223> /mol_tipo="ADN"/organismo="Fago lambda de enterobacteria"

<400> 2

atggtagaaa tcaataatca acgtaaggcg ttcctcgata tgctggcgtg gtcggaggga 60

actgataacg gacgtcagaa aaccagaaat catggttatg acgtcattgt aggcggagag 120

ctattcactg attactccga tcaccctcgc aaacttgtca cgctaaaccc aaaactcaaa 180

tcaacaggcg ccggacgcta ccagcttctt tcccgttggt gggatgccta ccgcaagcag 240

cttggcctga aagacttctc tccgaaaagt caggacgctg tggcattgca gcagattaag 300

gagcgtggcg ctttacctat gattgatcgt ggtgatatcc gtcaggcaat cgaccgttgc 360

agcaatatct gggcttcact gccgggcgct ggttatggtc agttcgagca taaggctgac 420

agcctgattg caaaattcaa agaagcgggc ggaacggtca gagagattga tgtatga 477

<210> 3
<211> 462
<212> ADN
<213> Fago lambda de enterobacteria

<220>
<221> fuente
<222> 1..462
<223> /mol_tipo="ADN"/organismo="Fago lambda de enterobacteria"

<400> 3
atgagcagag tcaccgcgat tatctccgct ctggttatct gcatcatcgt ctgcctgtca 60

tgggctgtta atcattaccg tgataacgcc attacctaca aagcccagcg cgacaaaaat 120

gccagagaac tgaagctggc gaacgcggca attactgaca tgcagatgcg tcagcgtgat 180

gttgctgcgc tcgatgcaaa atacacgaag gagttagctg atgctaaagc tgaaaatgat 240

gctctgcgtg atgatgttgc cgctggtcgt cgtcggttgc acatcaaagc agtctgtcag 300

tcagtgcgtg aagccaccac cgcctccggc gtggataatg cagcctcccc ccgactggca 360

gacaccgctg aacgggatta tttcaccctc agagagaggc tgatcactat gcaaaaacaa 420

ctggaaggaa cccagaagta tattaatgag cagtgcagat ag 462

<210> 4
<211> 201
<212> ADN
<213> Fago lambda de enterobacteria

<220>
<221> fuente
<222> 1..201
<223> /mol_tipo="ADN"/organismo="Fago lambda de enterobacteria"

<400> 4
atgactacga ctattgataa aaatcaatgg tgtggacaat tcaagcgatg caatggatgc 60

aagctgcaat cggaatgcat ggttaagcct gaagaaatgt ttcctgtaat ggaagatggg 120

aaatatgtcg ataaatgggc aatacgaacg acggcaatga ttgccagaga acttggtaaa 180

cagaacaaca aagctgcctg a 201

<210> 5
<211> 300
<212> ADN
<213> Fago lambda de enterobacteria

<220>
<221> fuente
<222> 1..300
<223> /mol_tipo="ADN"/organismo="Fago lambda de enterobacteria"

<400> 5

atggtaacca ttgtctggaa agaatccaaa ggtacggcaa aaagccgcta caaagctcgc 60

agagcagaac ttattgccga gcgacgcagt aatgaagcac tggcgcgaaa aattgcgcta 120

aagctctctg gttgcgtcag agcagataaa gcagcatcac tcggaagcct tcgctgcaag 180

aaggcagaag aagtcgagcg taaacagaac cgtatttact acagcaagcc acgcagtgaa 240

atgggtgtga cttgtgttgg tcgccagaaa attaaattag gcagcaaacc acttatttga 300

<210> 6
<211> 651
<212> ADN
<213> Fago lambda de enterobacteria

<220>
<221> fuente
<222> 1..651
<223> /mol_tipo="ADN"/organismo="Fago lambda de enterobacteria"

<400> 6

atgaatacac aactgatggg tgagcgtatt cgcgctcgca gaaaagaact caagattagg 60

caggctgccc ttggcaagat ggttggcgtg tctaatgttg ctatttccca atgggagcga 120

tctgaaactg agcccaatgg cgaaaaccta ttggccttag ccaaggcttt gcagtgctcc 180

cctgattacc tgttgaaagg agaggatagt ctttcaaaca ttgcctatca cagcaggcat 240

gatccaagag gttcgtatcc tctaattagt tgggtaagcg caggatgttg gatggaagct 300

gtagagccat atcataggcg tgcaatagat aactggtacg acacaacggt agattgttct 360

gaagactctt tttggctcga cgttaaaggc gattcaatga ctgccccggc aggactgagt 420

attcctgagg ggatgattat tctcgtcgac ccagaagtcg aaccacgtaa tggaaagctg 480

gtagtcgcca aacttgaagg agaaaacgag gcgacattca aaaagttagt tattgatgcc 540

ggtagaaaat tcctgaaacc actcaatcca caatacccaa tgattgaaat caatgggaac 600

tgtaaaatca ttggcgttgt cgttgatgcc aagctagcaa accttcctta a 651

<210> 7
<211> 1071
<212> ADN
<213> Fago lambda de enterobacteria

<220>
<221> fuente
<222> 1..1071

<223> /mol_tipo="ADN"/organismo="Fago lambda de enterobacteria"

<400> 7

atgggaagaa ggcgaagtca tgagcgccgg gatttacccc ctaaccttta tataagaaac 60

aatggatatt actgctacag ggacccaagg acgggtaaag agtttggatt aggcagagac 120

aggcgaatcg caatcactga agctatacag gccaacattg agttattttc aggacacaaa 180

cacaagcctc tgacagcgag aatcaacagt gataattccg ttacgttaca ttcatggctt 240

gatcgctacg aaaaaatcct ggccagcaga ggaatcaagc agaagacact cataaattac 300

atgagcaaaa ttaaagcaat aaggaggggt ctgcctgatg ctccacttga agacatcacc 360

acaaaagaaa ttgcggcaat gctcaatgga tacatagacg agggcaaggc ggcgtcagcc 420

aagttaatca gatcaacact gagcgatgca ttccgagagg caatagctga aggccatata 480

acaacaaacc atgtcgctgc cactcgcgca gcaaaatcag aggtaaggag atcaagactt 540

acggctgacg aatacctgaa aatttatcaa gcagcagaat catcaccatg ttggctcaga 600

cttgcaatgg aactggctgt tgttaccggg caacgagttg gtgatttatg cgaaatgaag 660

tggtctgata tcgtagatgg atatctttat gtcgagcaaa gcaaaacagg cgtaaaaatt 720

gccatcccaa cagcattgca tattgatgct ctcggaatat caatgaagga aacacttgat 780

aaatgcaaag agattcttgg cggagaaacc ataattgcat ctactcgtcg cgaaccgctt 840

tcatccggca cagtatcaag gtattttatg cgcgcacgaa aagcatcagg tctttccttc 900

gaaggggatc cgcctacctt tcacgagttg cgcagtttgt ctgcaagact ctatgagaag 960

cagataagcg ataagtttgc tcaacatctt ctcgggcata agtcggacac catggcatca 1020

cagtatcgtg atgacagagg cagggagtgg gacaaaattg aaatcaaata a 1071

<210> 8
<211> 219
<212> ADN
<213> Fago lambda de enterobacteria

<220>
<221> fuente
<222> 1..219
<223> /mol_tipo="ADN"/organismo="Fago lambda de enterobacteria"

<400> 8

atgtacttga cacttcagga gtggaacgca cgccagcgac gtccaagaag ccttgaaaca 60

gttcgtcgat gggttcggga atgcaggata ttcccacctc cggttaagga tggaagagag 120

tatctgttcc acgaatcagc ggtaaaggtt gacttaaatc gaccagtaac aggtggcctt 180

ttgaagagga tcagaaatgg gaagaaggcg aagtcatga 219

<210> 9
<211> 232
<212> ADN
<213> Fago lambda de enterobacteria

<220>

<221> fuente
<222> 1..232
<223> /mol_tipo="ADN”/organismo="Fago lambda de enterobacteria"

<400> 9

tacaggtcac taataccatc taagtagttg attcatagtg actgcatatg ttgtgtttta 60

cagtattatg tagtctgttt tttatgcaaa atctaattta atatattgat atttatatca 120

ttttacgttt ctcgttcagc ttttttatac taagttggca ttataaaaaa gcattgctta 180

tcaatttgtt gcaacgaaca ggtcactatc agtcaaaata aaatcattat tt 232

<210> 10
<211> 29865
<212> ADN
<213> secuencias artificiales

<220>
<221> fuente
<222> 1..29865
<223> /mol_tipo="ADN" /nota="Fago genéticamente modificado”/organismo="secuencias artificiales"

<400> 10

gctttttat actaagttgg cattataaaa aagcattgct tatcaatttg ttgcaacgaa 60

caggtcacta tcagtcaaaa taaaatcatt atttgatttc aattttgtcc cactccctgc 120

ctctgtcatc acgatactgt gatgccatgg tgtccgactt atgcccgaga agatgttgag 180

caaacttatc gcttatctgc ttctcataga gtcttgcaga caaactgcgc aactcgtgaa 240

aggtaggcgg atccagatcc cggacaccat cgaatggcgc aaaacctttc gcggtatggc 300

atgatagcgc ccggaagaga gtcaattcag ggtggtgaat gtgaaaccag taacgttata 360

cgatgtcgca gagtatgccg gtgtctctta tcagaccgtt tcccgcgtgg tgaaccaggc 420

cagccacgtt tctgcgaaaa cgcgggaaaa agtggaagcg gcgatggcgg agctgaatta 480

cattcccaac cgcgtggcac aacaactggc gggcaaacag tcgttgctga ttggcgttgc 540

cacctccagt ctggccctgc acgcgccgtc gcaaattgtc gcggcgatta aatctcgcgc 600

cgatcaactg ggtgccagcg tggtggtgtc gatggtagaa cgaagcggcg tcgaagcctg 660

taaagcggcg gtgcacaatc ttctcgcgca acgcgtcagt gggctgatca ttaactatcc 720

gctggatgac caggatgcca ttgctgtgga agctgcctgc actaatgttc cggcgttatt 780

tcttgatgtc tctgaccaga cacccatcaa cagtattatt ttctcccatg aagacggtac 840

gcgactgggc gtggagcatc tggtcgcatt gggtcaccag caaatcgcgc tgttagcggg 900

cccattaagt tctgtctcgg cgcgtctgcg tctggctggc tggcataaat atctcactcg 960

caatcaaatt cagccgatag cggaacggga aggcgactgg agtgccatgt ccggttttca 1020

acaaaccatg caaatgctga atgagggcat cgttcccact gcgatgctgg ttgccaacga 1080

tcagatggcg ctgggcgcaa tgcgcgccat taccgagtcc gggctgcgcg ttggtgcgga 1140

tatctcggta gtgggatacg acgataccga agacagctca tgttatatcc cgccgttaac 1200

caccatcaaa caggattttc gcctgctggg gcaaaccagc gtggaccgct tgctgcaact 1260

ctctcagggc caggcggtga agggcaatca gctgttgccc gtctcactgg tgaaaagaaa 1320

aaccaccctg gcgcccaata cgcaaaccgc ctctccccgc gcgttggccg attcattaat 1380

gcagctggca cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaatg 1440

taagttagct cactcattag gcaccccagg ctttacactt tatgcttccg gctcgtataa 1500

tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac catgattacg 1560

gattcactgg ccgtcgtttt acaacgtcgt gactgggaaa accctggcgt tacccaactt 1620

aatcgccttg cagcacatcc ccctttcgcc agctggcgta atagcgaaga ggcccgcacc 1680

gatcgccctt cccaacagtt gcgcagcctg aatggcgaat ggcgctttgc ctggtttccg 1740

gcaccagaag cggtgccgga aagctggctg gagtgcgatc ttcctgaggc cgatactgtc 1800

gtcgtccccc caaactggca gatgcacggt tacgatgcgc ccatctacac caacgtgacc 1860

tatcccatta cggtcaatcc gccgtttgtt cccacggaga atccgacggg ttgttactcg 1920

ctcacattta atgttgatga aagctggcta caggaaggcc agacgcgaat tatttttgat 1980

ggcgtcggga tctgatccgg atttactaac tggaagaggc actaaatgaa cacgattaac 2040

atcgctaaga acgacttctc tgacatcgaa ctggctgcta tcccgttcaa cactctggct 2100

gaccattacg gtgagcgttt agctcgcgaa cagttggccc ttgagcatga gtcttacgag 2160

atgggtgaag cacgcttccg caagatgttt gagcgtcaac ttaaagctgg tgaggttgcg 2220

gataacgctg ccgccaagcc tctcatcact accctactcc ctaagatgat tgcacgcatc 2280

aacgactggt ttgaggaagt gaaagctaag cgcggcaagc gcccgacagc cttccagttc 2340

ctgcaagaaa tcaagccgga agccgtagcg tacatcacca ttaagaccac tctggcttgc 2400

ctaaccagtg ctgacaatac aaccgttcag gctgtagcaa gcgcaatcgg tcgggccatt 2460

gaggacgagg ctcgcttcgg tcgtatccgt gaccttgaag ctaagcactt caagaaaaac 2520

gttgaggaac aactcaacaa gcgcgtaggg cacgtctaca agaaagcatt tatgcaagtt 2580

gtcgaggctg acatgctctc taagggtcta ctcggtggcg aggcgtggtc ttcgtggcat 2640

aaggaagact ctattcatgt aggagtacgc tgcatcgaga tgctcattga gtcaaccgga 2700

atggttagct tacaccgcca aaatgctggc gtagtaggtc aagactctga gactatcgaa 2760

ctcgcacctg aatacgctga ggctatcgca acccgtgcag gtgcgctggc tggcatctct 2820

ccgatgttcc aaccttgcgt agttcctcct aagccgtgga ctggcattac tggtggtggc 2880

tattgggcta acggtcgtcg tcctctggcg ctggtgcgta ctcacagtaa gaaagcactg 2940

atgcgctacg aagacgttta catgcctgag gtgtacaaag cgattaacat tgcgcaaaac 3000

accgcatgga aaatcaacaa gaaagtccta gcggtcgcca acgtaatcac caagtggaag 3060

cattgtccgg tcgaggacat ccctgcgatt gagcgtgaag aactcccgat gaaaccggaa 3120

gacatcgaca tgaatcctga ggctctcacc gcgtggaaac gtgctgccgc tgctgtgtac 3180

cgcaaggaca aggctcgcaa gtctcgccgt atcagccttg agttcatgct tgagcaagcc 3240

aataagtttg ctaaccataa ggccatctgg ttcccttaca acatggactg gcgcggtcgt 3300

gtttacgctg tgtcaatgtt caacccgcaa ggtaacgata tgaccaaagg actgcttacg 3360

ctggcgaaag gtaaaccaat cggtaaggaa ggttactact ggctgaaaat ccacggtgca 3420

aactgtgcgg gtgtcgataa ggttccgttc cctgagcgca tcaagttcat tgaggaaaac 3480

cacgagaaca tcatggcttg cgctaagtct ccactggaga acacttggtg ggctgagcaa 3540

gattctccgt tctgcttcct tgcgttctgc tttgagtacg ctggggtaca gcaccacggc 3600

ctgagctata actgctccct tccgctggcg tttgacgggt cttgctctgg catccagcac 3660

ttctccgcga tgctccgaga tgaggtaggt ggtcgcgcgg ttaacttgct tcctagtgaa 3720

accgttcagg acatctacgg gattgttgct aagaaagtca acgagattct acaagcagac 3780

gcaatcaatg ggaccgataa cgaagtagtt accgtgaccg atgagaacac tggtgaaatc 3840

tctgagaaag tcaagctggg cactaaggca ctggctggtc aatggctggc ttacggtgtt 3900

actcgcagtg tgactaagcg ttcagtcatg acgctggctt acgggtccaa agagttcggc 3960

ttccgtcaac aagtgctgga agataccatt cagccagcta ttgattccgg caagggtctg 4020

atgttcactc agccgaatca ggctgctgga tacatggcta agctgatttg ggaatctgtg 4080

agcgtgacgg tggtagctgc ggttgaagca atgaactggc ttaagtctgc tgctaagctg 4140

ctggctgctg aggtcaaaga taagaagact ggagagattc ttcgcaagcg ttgcgctgtg 4200

cattgggtaa ctcctgatgg tttccctgtg tggcaggaat acaagaagcc tattcagacg 4260

cgcttgaacc tgatgttcct cggtcagttc cgcttacagc ctaccattaa caccaacaaa 4320

gatagcgaga ttgatgcaca caaacaggag tctggtatcg ctcctaactt tgtacacagc 4380

caagacggta gccaccttcg taagactgta gtgtgggcac acgagaagta cggaatcgaa 4440

tcttttgcac tgattcacga ctccttcggt accattccgg ctgacgctgc gaacctgttc 4500

aaagcagtgc gcgaaactat ggttgacaca tatgagtctt gtgatgtact ggctgatttc 4560

tacgaccagt tcgctgacca gttgcacgag tctcaattgg acaaaatgcc agcacttccg 4620

gctaaaggta acttgaacct ccgtgacatc ttagagtcgg acttcgcgtt cgcgtaacgc 4680

caaatcaata cgactccgga tcccccttcga aggaaagacc tgatgctttt cgtgcgcgca 4740

taaaatacct tgatactgtg ccggatgaaa gcggttcgcg acgagtagat gcaattatgg 4800

tttctccgcc aagaatctct ttgcatttat caagtgtttc cttcattgat attccgagag 4860

catcaatatg caatgctgtt gggatggcaa tttttacgcc tgttttgctt tgctcgacat 4920

aaagatatcc atctacgata tcagaccact tcatttcgca taaatcacca actcgttgcc 4980

cggtaacaac agccagttcc attgcaagtc tgagccaaca tggtgatgat tctgctgctt 5040

gataaatttt caggtattcg tcagccgtaa gtcttgatct ccttacctct gattttgctg 5100

cgcgagtggc agcgacatgg tttgttgtta tatggccttc agctattgcc tctcggaatg 5160

catcgctcag tgttgatctg attaacttgg ctgacgccgc cttgccctcg tctatgtatc 5220

cattgagcat tgccgcaatt tcttttgtgg tgatgtcttc aagtggagca tcaggcagac 5280

ccctccttat tgctttaatt ttgctcatgt aatttatgag tgtcttctgc ttgattcctc 5340

tgctggccag gatttttcg tagcgatcaa gccatgaatg taacgtaacg gaattatcac 5400

tgttgattct cgctgtcaga ggcttgtgtt tgtgtcctga aaataactca atgttggcct 5460

gtatagcttc agtgattgcg attcgcctgt ctctgcctaa tccaaactct ttacccgtcc 5520

ttgggtccct gtagcagtaa tatccattgt ttcttatata aaggttaggg ggtaaatccc 5580

ggcgctcatg acttcgcctt cttcacactg gagggcaaag aagatttcca ataatcagaa 5640

caagtcggct cctgtttagt tacgagcgac attgctccgt gtattcactc gttggaatga 5700

atacacagtg cagtgtttat tctgttattt atgccaaaaa taaaggccac tatcaggcag 5760

ctttgttgtt ctgtttacca agttctctgg caatcattgc cgtcgttcgt attgcccatt 5820

tatcgacata tttcccatct tccattacag gaaacatttc ttcaggctta accatgcatt 5880

ccgattgcag cttgcatcca ttgcatcgct tgaattgtcc acaccattga tttttatcaa 5940

tagtcgtagt catacggata gtcctggtat tgttccatca catcctgagg atgctcttcg 6000

aactcttcaa attcttcttc catatatcac ctcaaataag tggtttgctg cctaatttaa 6060

ttttctggcg accaacacaa gtcacaccca tttcactgcg tggcttgctg tagtaaatac 6120

ggttctgttt acgctcgact tcttctgcct tcttgcagcg aaggcttccg agtgatgctg 6180

ctttatctgc tctgacgcaa ccagagagct ttagcgcaat ttttcgcgcc agtgcttcat 6240

tactgcgtcg ctcggcaata agttctgctc tgcgagcttt gtagcggctt tttgccgtac 6300

ctttggattc tttccagaca atggttacca tgatggtctc ctttaagtgg ctttggcgca 6360

tgacgcgtcg aggtgcttat cttctcgatc gctgtcttgt agctgcaatt cgcgccatcc 6420

ccaaaaccac tcaagttctg gtctcaacgg ttaggttgag agtccgtcga tgttaaagag 6480

cctgccaatc tgttccgttt ggcttccagc gtcctgctga tggcttaaat ttaagacttc 6540

ttaatttatt ggtcaagtgc atttttgaag aaaacttaat tttatgggcg tgaatttagt 6600

ttgtctttga tttttaacgg gaaataaaaa aggggcgaaa gccccttaag gaaggtttgc 6660

tagcttggca tcaacgacaa cgccaatgat tttacagttc ccattgattt caatcattgg 6720

gtattgtgga ttgagtggtt tcaggaattt tctaccggca tcaataacta acttttttgaa 6780

tgtcgcctcg ttttctcctt caagtttggc gactaccagc tttccattac gtggttcgac 6840

ttctgggtcg acgagaataa tcatcccctc aggaatactc agtcctgccg gggcagtcat 6900

tgaatcgcct ttaacgtcga gccaaaaaga gtcttcagaa caatctaccg ttgtgtcgta 6960

ccagttatct attgcacgcc tatgatatgg ctctacagct tccatccaac atcctgcgct 7020

tacccaacta attagaggat acgaacctct tggatcatgc ctgctgtgat aggcaatgtt 7080

tgaaagacta tcctctcctt tcaacaggta atcaggggag cactgcaaag ccttggctaa 7140

ggccaatagg ttttcgccat tgggctcagt ttcagatcgc tcccattggg aaatagcaac 7200

attagacacg ccaaccatct tgccaagggc agcctgccta atcttgagtt cttttctgcg 7260

agcgcgaata cgctcaccca tcagttgtgt attcatagtt aagacatctt aaataaactt 7320

gacttaagat tcctttggtg gataatttaa gtgttcttta atttcggagc gagtctatgt 7380

acaaaaaaga tgttattgac cacttcggaa cccagcgtgc tgttgctaaa gcactaggca 7440

ttagcgatgc agcagtctct cagtggaaag aagttatccc agagaaagac gcctatcgat 7500

tggaaatcgt tacagctggc gccctgaagt atcaagaaag tgcttaccgc caagcggcat 7560

aagcaaattg ctctttaaca gttctggcct ttcacctcta accgggtgag caaacatcag 7620

cggcaaatcc attgggtgtg ccgctataac tcaatatcaa tataggtaaa ttaacaaatg 7680

gcacaagcaa gctacagcaa gccaacacag cgagaaattg atcgcgctga aactgattta 7740

ctcatcaacc tgtcaacgct tacccagcgc ggtctggcaa agatgattgg ctgtcatgaa 7800

tcgaagataa gcagaacgga ctggagattt attgcttcgg tcttgtgtgc tttcggaatg 7860

gcatcagaca tcagtccgat tagcagggct tttaagtatg cgcttgatgg actcacaaag 7920

aaaaaacgcc cggtgtgcaa gaccgagcgt tctgaacaaa tccagatgga gttctgaggt 7980

cattactgga tctatcaaca ggagtcatta tgacaaatac agcaaaaata ctcaacttcg 8040

gcagaggtaa ctttgccgga caggagcgta atgtggcaga tctcgatgat ggttacgcca 8100

gactatcaaa tatgctgctt gaggcttatt cgggcgcaga tctgaccaag cgacagttta 8160

aagtgctgct tgccattctg cgtaaaacct atgggtggaa taaaccaatg gacagaatca 8220

ccgattctca acttagcgag attacaaagt tacctgtcaa acggtgcaat gaagccaagt 8280

tagaactcgt cagaatgaat attatcaagc agcaaggcgg catgtttgga ccaaataaaa 8340

acatctcaga atggtgtatc cctcaaaacg agggaaaatc ccctaaaacg agggataaaa 8400

catccctcaa attgggggat tgctatccct caaaacaggg ggacacaaaa gacactatta 8460

caaaagaaaa aagaaaagat tattcgtcag agaattctgg cgaatcctct gaccagccag 8520

aaaacgacct ttctgtggtg aaaccggatg ctgcaattca gagcggcagc aagtggggga 8580

cagcagaaga cctgaccgcc gcagagtgga tgtttgacat ggtgaagact atcgcaccat 8640

cagccagaaa accgaatttt gctgggtggg ctaacgatat ccgcctgatg cgtgaacgtg 8700

acggacgtaa ccaccgcgac atgtgtgtgc tgttccgctg ggcatgccag gacaacttct 8760

ggtccggtaa cgtgctgagc ccggccaaac tccgcgataa gtggacccaa ctcgaaatca 8820

accgtaacaa gcaacaggca ggcgtgacag ccagcaaacc aaaactcgac ctgacaaaca 8880

cagactggat ttacggggtg gatctatgaa aaacatcgcc gcacagatgg ttaactttga 8940

ccgtgagcag atgcgtcgga tcgccaacaa catgccggaa cagtacgacg aaaagccgca 9000

ggtacagcag gtagcgcaga tcatcaacgg tgtgttcagc cagttactgg caactttccc 9060

ggcgagcctg gctaaccgtg accagaacga agtgaacgaa atccgtcgcc agtgggttct 9120

ggcttttcgg gaaaacggga tcaccacgat ggaacaggtt aacgcaggaa tgcgcgtagc 9180

```
ccgtcggcag aatcgaccat ttctgccatc acccgggcag tttgttgcat ggtgccggga     9240
agaagcatcc gttaccgccg gactgccaaa cgtcagcgag ctggttgata tggtttacga     9300
gtattgccgg aagcgaggcc tgtatccgga tgcggagtct tatccgtgga aatcaaacgc     9360
gcactactgg ctggttacca acctgtatca gaacatgcgg gccaatgcgc ttactgatgc     9420
ggaattacgc cgtaaggccg cagatgagct tgtccatatg actgcgagaa ttaaccgtgg     9480
tgaggcgatc cctgaaccag taaaacaact tcctgtcatg ggcggtagac ctctaaatcg     9540
tgcacaggct ctggcgaaga tcgcagaaat caaagctaag ttcggactga aaggagcaag     9600
tgtatgacgg gcaaagaggc aattattcat tacctgggga cgcataatag cttctgtgcg     9660
ccggacgttg ccgcgctaac aggcgcaaca gtaaccagca taaatcaggc cgcggctaaa     9720
atggcacggg caggtcttct ggttatcgaa ggtaaggtct ggcgaacggt gtattaccgg     9780
tttgctacca gggaagaacg ggaaggaaag atgagcacga acctgatgaa caaactggat     9840
acgattggat tcgacaacaa aaaagacctg cttatctcgg tgggcgattt ggttgatcgt     9900
ggtgcagaga acgttgaatg cctggaatta atcacattcc cctggttcag agctgtacgt     9960
ggaaaccatg agcaaatgat gattgatggc ttatcagagc gtggaaacgt taatcactgg    10020
ctgcttaatg gcggtggctg gttctttaat ctcgattacg acaaagaaat tctggctaaa    10080
gctcttgccc ataaagcaga tgaacttccg ttaatcatcg aactggtgag caaagataaa    10140
aaatatgtta tctgccacgc cgattatccc tttgacgaat acgagtttgg aaagccagtt    10200
gatcatcagc aggtaatctg gaaccgcgaa cgaatcagca actcacaaaa cgggatcgtg    10260
aaagaaatca aaggcgcgga cacgttcatc tttggtcata cgccagcagt gaaaccactc    10320
aagtttgcca accaaatgta tatcgatacc ggcgcagtgt tctgcggaaa cctaacattg    10380
attcaggtac agggagaagg cgcatgagac tcgaaagcgt agctaaattt cattcgccaa    10440
aaagcccgat gatgagcgac tcaccacggg ccacggcttc tgactctctt tccggtactg    10500
atgtgatggc tgctatgggg atggcgcaat cacaagccgg attcggtatg gctgcattct    10560
gcggtaagca cgaactcagc cagaacgaca aacaaaaggc tatcaactat ctgatgcaat    10620
ttgcacacaa ggtatcgggg aaataccgtg gtgtggcata tcttgaagga aatactaagg    10680
caaaggtact gcaagtgctc gcaacattcg cttatgcgga ttattgccgt agtgccgcga    10740
cgccgggggc aagatgcaga gattgccatg gtacaggccg tgcggttgat attgccaaaa    10800
cagagctgtg ggggagagtt gtcgagaaag agtgcggaag atgcaaaggc gtcggctatt    10860
caaggatgcc agcaagcgca gcatatcgcg ctgtgacgat gctaatccca aaccttaccc    10920
aacccacctg gtcacgcact gttaagccgc tgtatgacgc tctggtggtg caatgccaca    10980
aagaagagtc aatcgcagac aacattttga atgcggtcac acgttagcag catgattgcc    11040
acggatggca acatattaac ggcatgatat tgacttattg aataaaattg ggtaaatttg    11100
actcaacgat gggttaattc gctcgttgtg gtagtgagat gaaaagaggc ggcgcttact    11160
```

accgattccg cctagttggt cacttcgacg tatcgtctgg aactccaacc atcgcaggca 11220

gagaggtctg caaaatgcaa tcccgaaaca gttcgcaggt aatagttaga gcctgcataa 11280

cggtttcggg atttttttata tctgcacaac aggtaagagc attgagtcga taatcgtgaa 11340

gagtcggcga gcctggttag ccagtgctct ttccgttgtg ctgaattaag cgaataccgg 11400

aagcagaacc ggatcaccaa atgcgtacag gcgtcatcgc cgcccagcaa cagcacaacc 11460

caaactgagc cgtagccact gtctgtcctg aattccatgc ttgaacccgc ctatgcgcgg 11520

gttttctttt gtgcgcttgc aggccagctt gggatcagca gcctgacgga tgcggtgtcc 11580

ggcgacagcc tgactgccca ggaggcactc gcgacgctgg cattatccgg tgatgatgac 11640

ggaccacgac aggcccgcag ttatcaggtc atgaacggca tcgccgtgct gccggtgtcc 11700

ggcacgctgg tcagccggac gcgggcgctg cagccgtact cggggatgac cggttacaac 11760

ggcattatcg cccgtctgca acaggctgcc agcgatccga tggtggacgg cattctgctc 11820

gatatggaca cgcccggcgg gatggtggcg ggggcatttg actgcgctga catcatcgcc 11880

cgtgtgcgtg acataaaacc ggtatgggcg cttgccaacg acatgaactg cagtgcaggt 11940

cagttgcttg ccagtgccgc ctcccggcgt ctggtcacgc agaccgcccg gacaggctcc 12000

atcggcgtca tgatggctca cagtaattac ggtgctgcgc tggagaaaca gggtgtggaa 12060

atcacgctga tttacagcgg cagccataag gtggatggca acccctacag ccatcttccg 12120

gatgacgtcc gggagacact gcagtcccgg atggacgcaa cccgccagat gtttgcgcag 12180

aaggtgtcgg catataccgg cctgtccgtg caggttgtgc tggataccga ggctgcagtg 12240

tacagcggtc aggaggccat tgatgccgga ctggctgatg aacttgttaa cagcaccgat 12300

gcgatcaccg tcatgcgtga tgcactggat gcacgtaaat cccgtctctc aggagggcga 12360

atgaccaaag agactcaatc aacaactgtt tcagccactg cttcgcaggc tgacgttact 12420

gacgtggtgc cagcgacgga gggcgagaac gccagcgcgg cgcagccgga cgtgaacgcg 12480

cagatcaccg cagcggttgc ggcagaaaac agccgcatta tggggatact caactgtgag 12540

gaggctcacg gacgcgaaga acaggcacgc gtgctggcag aaacccccgg tatgaccgtg 12600

aaaacggccc gccgcattct ggccgcagca ccacagagtg cacaggcgcg cagtgacact 12660

gcgctggatc gtctgatgca gggggcaccg gcaccgctgg ctgcaggtaa cccggcatct 12720

gatgccgtta acgatttgct gaacacacca gtgtaaggga tgtttatgac gagcaaagaa 12780

acctttaccc attaccagcc gcagggcaac agtgacccgg ctcataccgc aaccgcgccc 12840

ggcggattga gtgcgaaagc gcctgcaatg accccgctga tgctggacac ctccagccgt 12900

aagctggttg cgtgggatgg caccaccgac ggtgctgccg ttggcattct tgcggttgct 12960

gctgaccaga ccagcaccac gctgacgttc tacaagtccg gcacgttccg ttatgaggat 13020

gtgctctggc cggaggctgc cagcgacgag acgaaaaaac ggaccgcgtt tgccggaacg 13080

```
gcaatcagca tcgtttaact ttacccttca tcactaaagg ccgcctgtgc ggctttttt    13140
acgggatttt tttatgtcga tgtacacaac cgcccaactg ctggcggcaa atgagcagaa    13200
atttaagttt gatccgctgt ttctgcgtct ctttttccgt gagagctatc ccttcaccac    13260
ggagaaagtc tatctctcac aaattccggg actggtaaac atggcgctgt acgtttcgcc    13320
gattgtttcc ggtgaggtta tccgttcccg tggcggctcc acctctgaat ttacgccggg    13380
atatgtcaag ccgaagcatg aagtgaatcc gcagatgacc ctgcgtcgcc tgccggatga    13440
agatccgcag aatctggcgg acccggctta ccgccgccgt cgcatcatca tgcagaacat    13500
gcgtgacgaa gagctggcca ttgctcaggt cgaagagatg caggcagttt ctgccgtgct    13560
taagggcaaa tacaccatga ccggtgaagc cttcgatccg gttgaggtgg atatgggccg    13620
cagtgaggag aataacatca cgcagtccgg cggcacggag tggagcaagc gtgacaagtc    13680
cacgtatgac ccgaccgacg atatcgaagc ctacgcgctg aacgccagcg gtgtggtgaa    13740
tatcatcgtg ttcgatccga aaggctgggc gctgttccgt tccttcaaag ccgtcaagga    13800
gaagctggat acccgtcgtg gctctaattc cgagctggag acagcggtga aagacctggg    13860
caaagcggtg tcctataagg ggatgtatgg cgatgtggcc atcgtcgtgt attccggaca    13920
gtacgtggaa aacggcgtca aaaagaactt cctgccggac aacacgatgg tgctggggaa    13980
cactcaggca cgcggtctgc gcacctatgg ctgcattcag gatgcggacg cacagcgcga    14040
aggcattaac gcctctgccc gttacccgaa aaactgggtg accaccggcg atccggcgcg    14100
tgagttcacc atgattcagt cagcaccgct gatgctgctg gctgaccctg atgagttcgt    14160
gtccgtacaa ctggcgtaat catggccctt cggggccatt gtttctctgt ggaggagtcc    14220
atgacgaaag atgaactgat tgcccgtctc cgctcgctgg gtgaacaact gaaccgtgat    14280
gtcagcctga cggggacgaa agaagaactg gcgctccgtg tggcagagct gaaagaggag    14340
cttgatgaca cggatgaaac tgccggtcag gacacccctc tcagccggga aaatgtgctg    14400
accggacatg aaaatgaggt gggatcagcg cagccggata ccgtgattct ggatacgtct    14460
gaactggtca cggtcgtggc actggtgaag ctgcatactg atgcacttca cgccacgcgg    14520
gatgaacctg tggcatttgt gctgccggga acggcgtttc gtgtctctgc cggtgtggca    14580
gccgaaatga cagagcgcgg cctggccaga atgcaataac gggaggcgct gtggctgatt    14640
tcgataacct gttcgatgct gccattgccc gcgccgatga aacgatacgc gggtacatgg    14700
gaacgtcagc caccattaca tccggtgagc agtcaggtgc ggtgatacgt ggtgtttttg    14760
atgaccctga aaatatcagc tatgccggac agggcgtgcg cgttgaaggc tccagcccgt    14820
ccctgtttgt ccggactgat gaggtgcggc agctgcggcg tggagacacg ctgaccatcg    14880
gtgaggaaaa tttctgggta gatcgggttt cgccggatga tggcggaagt tgtcatctct    14940
ggcttggacg gggcgtaccg cctgccgtta accgtcgccg ctgaaagggg gatgtatggc    15000
cataaaaggt cttgagcagg ccgttgaaaa cctcagccgt atcagcaaaa cggcggtgcc    15060
```

tggtgccgcc gcaatggcca ttaaccgcgt tgcttcatcc gcgatatcgc agtcggcgtc 15120

acaggttgcc cgtgagacaa aggtacgccg gaaactggta aaggaaaggg ccaggctgaa 15180

aagggccacg gtcaaaaatc cgcaggccag aatcaaagtt aaccgggggg atttgcccgt 15240

aatcaagctg ggtaatgcgc gggttgtcct ttcgcgccgc aggcgtcgta aaaaggggca 15300

gcgttcatcc ctgaaaggtg gcggcagcgt gcttgtggtg ggtaaccgtc gtattcccgg 15360

cgcgtttatt cagcaactga aaaatggccg gtggcatgtc atgcagcgtg tggctgggaa 15420

aaaccgttac cccattgatg tggtgaaaat cccgatggcg gtgccgctga ccacggcgtt 15480

taaacaaaat attgagcgga tacggcgtga acgtcttccg aaagagctgg gctatgcgct 15540

gcagcatcaa ctgaggatgg taataaagcg atgaaacata ctgaactccg tgcagccgta 15600

ctggatgcac tggagaagca tgacaccggg gcgacgtttt ttgatggtcg ccccgctgtt 15660

tttgatgagg cggattttcc ggcagttgcc gtttatctca ccggcgctga atacacgggc 15720

gaagagctgg acagcgatac ctggcaggcg gagctgcata tcgaagtttt cctgcctgct 15780

caggtgccgg attcagagct ggatgcgtgg atggagtccc ggatttatcc ggtgatgagc 15840

gatatcccgg cactgtcaga tttgatcacc agtatggtgg ccagcggcta tgactaccgg 15900

cgcgacgatg atgcgggctt gtggagttca gccgatctga cttatgtcat tacctatgaa 15960

atgtgaggac gctatgcctg taccaaatcc tacaatgccg gtgaaaggtg ccgggaccac 16020

cctgtgggtt tataagggga gcggtgaccc ttacgcgaat ccgctttcag acgttgactg 16080

gtcgcgtctg gcaaaagtta aagacctgac gcccggcgaa ctgaccgctg agtcctatga 16140

cgacagctat ctcgatgatg aagatgcaga ctggactgcg accgggcagg ggcagaaatc 16200

tgccggagat accagcttca cgctggcgtg gatgcccgga gagcaggggc agcaggcgct 16260

gctggcgtgg tttaatgaag gcgatacccg tgcctataaa atccgcttcc cgaacggcac 16320

ggtcgatgtg ttccgtggct gggtcagcag tatcggtaag gcggtgacgg cgaaggaagt 16380

gatcacccgc acggtgaaag tcaccaatgt gggacgtccg tcgatggcag aagatcgcag 16440

cacggtaaca gcggcaaccg gcatgaccgt gacgcctgcc agcacctcgg tggtgaaagg 16500

gcagagcacc acgctgaccg tggccttcca gccggagggc gtaaccgaca agagctttcg 16560

tgcggtgtct gcggataaaa caaaagccac cgtgtcggtc agtggtatga ccatcaccgt 16620

gaacggcgtt gctgcaggca aggtcaacat tccggttgta tccggtaatg gtgagtttgc 16680

tgcggttgca gaaattaccg tcaccgccag ttaatccgga gagtcagcga tgttcctgaa 16740

aaccgaatca tttgaacata acggtgtgac cgtcacgctt tctgaactgt cagccctgca 16800

gcgcattgag catctcgccc tgatgaaacg gcaggcagaa caggcggagt cagacagcaa 16860

ccggaagttt actgtggaag acgccatcag aaccggcgcg tttctggtgg cgatgtccct 16920

gtggcataac catccgcaga agacgcagat gccgtccatg aatgaagccg ttaaacagat 16980

tgagcaggaa gtgcttacca cctggcccac ggaggcaatt tctcatgctg aaaacgtggt 17040

gtaccggctg tctggtatgt atgagtttgt ggtgaataat gcccctgaac agacagagga 17100

cgccgggccc gcagagcctg tttctgcggg aaagtgttcg acggtgagct gagttttgcc 17160

ctgaaactgg cgcgtgagat ggggcgaccc gactggcgtg ccatgcttgc cgggatgtca 17220

tccacggagt atgccgactg gcaccgcttt tacagtaccc attattttca tgatgttctg 17280

ctggatatgc acttttccgg gctgacgtac accgtgctca gcctgttttt cagcgatccg 17340

gatatgcatc cgctggattt cagtctgctg aaccggcgcg aggctgacga agagcctgaa 17400

gatgatgtgc tgatgcagaa agcggcaggg cttgccggag gtgtccgctt tggcccggac 17460

gggaatgaag ttatccccgc ttccccggat gtggcggaca tgacggagga tgacgtaatg 17520

ctgatgacag tatcagaagg gatcgcagga ggagtccggt atggctgaac cggtaggcga 17580

tctggtcgtt gatttgagtc tggatgcggc cagatttgac gagcagatgg ccagagtcag 17640

gcgtcatttt tctggtacgg aaagtgatgc gaaaaaaaca gcggcagtcg ttgaacagtc 17700

gctgagccga caggcgctgg ctgcacagaa agcggggatt tccgtcgggc agtataaagc 17760

cgccatgcgt atgctgcctg cacagttcac cgacgtggcc acgcagcttg caggcgggca 17820

aagtccgtgg ctgatcctgc tgcaacaggg ggggcaggtg aaggactcct tcggcgggat 17880

gatccccatg ttcaggggggc ttgccggtgc gatcaccctg ccgatggtgg gggccacctc 17940

gctggcggtg gcgaccggtg cgctggcgta tgcctggtat cagggcaact caaccctgtc 18000

cgatttcaac aaaacgctgg tcctttccgg caatcaggcg ggactgacgg cagatcgtat 18060

gctggtcctg tccagagccg ggcaggcggc agggctgacg tttaaccaga ccagcgagtc 18120

actcagcgca ctggttaagg cgggggtaag cggtgaggct cagattgcgt ccatcagcca 18180

gagtgtggcg cgtttctcct ctgcatccgg cgtggaggtg gacaaggtcg ctgaagcctt 18240

cgggaagctg accacagacc cgacgtcggg gctgacggcg atggctcgcc agttccataa 18300

cgtgtcggcg gagcagattg cgtatgttgc tcagttgcag cgttccggcg atgaagccgg 18360

ggcattgcag gcggcgaacg aggccgcaac gaaagggttt gatgaccaga cccgccgcct 18420

gaaagagaac atgggcacgc tggagacctg ggcagacagg actgcgcggg cattcaaatc 18480

catgtgggat gcggtgctgg atattggtcg tcctgatacc gcgcaggaga tgctgattaa 18540

ggcagaggct gcgtataaga aagcagacga catctggaat ctgcgcaagg atgattattt 18600

tgttaacgat gaagcgcggg cgcgttactg ggatgatcgt gaaaaggccc gtcttgcgct 18660

tgaagccgcc cgaaagaagg ctgagcagca gactcaacag gacaaaaatg cgcagcagca 18720

gagcgatacc gaagcgtcac ggctgaaata taccgaagag gcgcagaagg cttacgaacg 18780

gctgcagacg ccgctggaga aatataccgc ccgtcaggaa gaactgaaca aggcactgaa 18840

agacgggaaa atcctgcagg cggattacaa cacgctgatg gcggcggcga aaaaggatta 18900

tgaagcgacg ctgaaaaagc cgaaacagtc cagcgtgaag gtgtctgcgg gcgatcgtca 18960

ggaagacagt gctcatgctg ccctgctgac gcttcaggca gaactccgga cgctggagaa 19020

gcatgccgga gcaaatgaga aaatcagcca gcagcgccgg gatttgtgga aggcggagag 19080

tcagttcgcg gtactggagg aggcggcgca acgtcgccag ctgtctgcac aggagaaatc 19140

cctgctggcg cataaagatg agacgctgga gtacaaacgc cagctggctg cacttggcga 19200

caaggttacg tatcaggagc gcctgaacgc gctggcgcag caggcggata aattcgcaca 19260

gcagcaacgg gcaaaacggg ccgccattga tgcgaaaagc cgggggctga ctgaccggca 19320

ggcagaacgg gaagccacgg aacagcgcct gaaggaacag tatggcgata atccgctggc 19380

gctgaataac gtcatgtcag agcagaaaaa gacctgggcg gctgaagacc agcttcgcgg 19440

gaactggatg gcaggcctga agtccggctg gagtgagtgg gaagagagcg ccacggacag 19500

tatgtcgcag gtaaaaagtg cagccacgca gacctttgat ggtattgcac agaatatggc 19560

ggcgatgctg accggcagtg agcagaactg gcgcagcttc acccgttccg tgctgtccat 19620

gatgacagaa attctgctta agcaggcaat ggtggggatt gtcgggagta tcggcagcgc 19680

cattggcggg gctgttggtg gcggcgcatc cgcgtcaggc ggtacagcca ttcaggccgc 19740

tgcggcgaaa ttccattttg caaccggagg atttacggga accggcggca aatatgagcc 19800

agcggggatt gttcaccgtg gtgagtttgt cttcacgaag gaggcaacca gccggattgg 19860

cgtggggaat ctttaccggc tgatgcgcgg ctatgccacc ggcggttatg tcggtacacc 19920

gggcagcatg gcagacagcc ggtcgcaggc gtccgggacg tttgagcaga ataaccatgt 19980

ggtgattaac aacgacggca cgaacgggca gataggtccg gctgctctga aggcggtgta 20040

tgacatggcc cgcaagggtg cccgtgatga aattcagaca cagatgcgtg atggtggcct 20100

gttctccgga ggtggacgat gaagaccttc cgctggaaag tgaaacccgg tatggatgtg 20160

gcttcggtcc cttctgtaag aaaggtgcgc tttggtgatg gctattctca gcgagcgcct 20220

gccgggctga atgccaacct gaaaacgtac agcgtgacgc tttctgtccc ccgtgaggag 20280

gccacggtac tggagtcgtt tctggaagag cacgggggct ggaaatcctt tctgtggacg 20340

ccgccttatg agtggcggca gataaaggtg acctgcgcaa aatggtcgtc gcgggtcagt 20400

atgctgcgtg ttgagttcag cgcagagttt gaacaggtgg tgaactgatg caggatatcc 20460

ggcaggaaac actgaatgaa tgcacccgtg cggagcagtc ggccagcgtg gtgctctggg 20520

aaatcgacct gacagaggtc ggtggagaac gttatttttt ctgtaatgag cagaacgaaa 20580

aaggtgagcc ggtcacctgg caggggcgac agtatcagcc gtatcccatt caggggagcg 20640

gttttgaact gaatggcaaa ggcaccagta cgcgccccac gctgacggtt tctaacctgt 20700

acggtatggt caccgggatg gcggaagata tgcagagtct ggtcggcgga acggtggtcc 20760

ggcgtaaggt ttacgcccgt tttctggatg cggtgaactt cgtcaacgga aacagttacg 20820

ccgatccgga gcaggaggtg atcagccgct ggcgcattga gcagtgcagc gaactgagcg 20880

cggtgagtgc ctcctttgta ctgtccacgc cgacggaaac ggatggcgct gtttttccgg 20940

gacgtatcat gctggccaac acctgcacct ggacctatcg cggtgacgag tgcggttata 21000

gcggtccggc tgtcgcggat gaatatgacc agccaacgtc cgatatcacg aaggataaat 21060

gcagcaaatg cctgagcggt tgtaagttcc gcaataacgt cggcaacttt ggcggcttcc 21120

tttccattaa caaactttcg cagtaaatcc catgacacag acagaatcag cgattctggc 21180

gcacgcccgg cgatgtgcgc cagcggagtc gtgcggcttc gtggtaagca cgccggaggg 21240

ggaaagatat ttcccctgcg tgaatatctc cggtgagccg gaggcgtatt tccgtatgtc 21300

gccggaagac tggctgcagg cagaaatgca gggtgagatt gtggcgctgg tccacagcca 21360

ccccggtggt ctgccctggc tgagtgaggc cgaccggcgg ctgcaggtgc agagtgattt 21420

gccgtggtgg ctggtctgcc gggggacgat tcataagttc cgctgtgtgc cgcatctcac 21480

cgggcggcgc tttgagcacg gtgtgacgga ctgttacaca ctgttccggg atgcttatca 21540

tctggcgggg attgagatgc cggactttca tcgtgaggat gactggtggc gtaacggcca 21600

gaatctctat ctggataatc tggaggcgac ggggctgtat caggtgccgt tgtcagcggc 21660

acagccgggc gatgtgctgc tgtgctgttt tggttcatca gtgccgaatc acgccgcaat 21720

ttactgcggc gacggcgagc tgctgcacca tattcctgaa caactgagca aacgagagag 21780

gtacaccgac aaatggcagc gacgcacaca ctccctctgg cgtcaccggg catggcgcgc 21840

atctgccttt acggggattt acaacgattt ggtcgccgca tcgaccttcg tgtgaaaacg 21900

ggggctgaag ccatccgggc actggccaca cagctcccgg cgtttcgtca gaaactgagc 21960

gacggctggt atcaggtacg gattgccggg cgggacgtca gcacgtccgg gttaacggcg 22020

cagttacatg agactctgcc tgatggcgct gtaattcata ttgttcccag agtcgccggg 22080

gccaagtcag gtggcgtatt ccagattgtc ctgggggctg ccgccattgc cggatcattc 22140

tttaccgccg gagccaccct tgcagcatgg ggggcagcca ttggggccgg tggtatgacc 22200

ggcatcctgt tttctctcgg tgccagtatg gtgctcggtg gtgtggcgca gatgctggca 22260

ccgaaagcca gaactccccg tatacagaca acggataacg gtaagcagaa cacctatttc 22320

tcctcactgg ataacatggt tgcccagggc aatgttctgc ctgttctgta cggggaaatg 22380

cgcgtggggt cacgcgtggt ttctcaggag atcagcacgg cagacgaagg ggacggtggt 22440

caggttgtgg tgattggtcg ctgatgcaaa atgttttatg tgaaaccgcc tgcgggcggt 22500

tttgtcattt atggagcgtg aggaatgggt aaaggaagca gtaaggggca taccccgcgc 22560

gaagcgaagg acaacctgaa gtccacgcag ttgctgagtg tgatcgatgc catcagcgaa 22620

gggccgattg aaggtccggt ggatggctta aaaagcgtgc tgctgaacag tacgccggtg 22680

ctggacactg aggggaatac caacatatcc ggtgtcacgg tggtgttccg ggctggtgag 22740

caggagcaga ctccgccgga gggatttgaa tcctccggct ccgagacggt gctgggtacg 22800

gaagtgaaat atgacacgcc gatcacccgc accattacgt ctgcaaacat cgaccgtctg 22860

cgctttacct tcggtgtaca ggcactggtg gaaaccacct caaagggtga caggaatccg 22920

tcggaagtcc gcctgctggt tcagatacaa cgtaacggtg gctgggtgac ggaaaaagac 22980

atcaccatta agggcaaaac cacctcgcag tatctggcct cggtggtgat gggtaacctg 23040

ccgccgcgcc cgtttaatat ccggatgcgc aggatgacgc cggacagcac cacagaccag 23100

ctgcagaaca aaacgctctg gtcgtcatac actgaaatca tcgatgtgaa acagtgctac 23160

ccgaacacgg cactggtcgg cgtgcaggtg gactcggagc agttcggcag ccagcaggtg 23220

agccgtaatt atcatctgcg cgggcgtatt ctgcaggtgc cgtcgaacta taacccgcag 23280

acgcggcaat acagcggtat ctgggacgga acgtttaaac cggcatacag caacaacatg 23340

gcctggtgtc tgtgggatat gctgacccat ccgcgctacg gcatggggaa acgtcttggt 23400

gcggcggatg tggataaatg ggcgctgtat gtcatcggcc agtactgcga ccagtcagtg 23460

ccggacggct ttggcggcac ggagccgcgc atcacctgta atgcgtacct gaccacacag 23520

cgtaaggcgt gggatgtgct cagcgatttc tgctcggcga tgcgctgtat gccggtatgg 23580

aacgggcaga cgctgacgtt cgtgcaggac cgaccgtcgg ataagacgtg gacctataac 23640

cgcagtaatg tggtgatgcc ggatgatggc gcgccgttcc gctacagctt cagcgccctg 23700

aaggaccgcc ataatgccgt tgaggtgaac tggattgacc cgaacaacgg ctgggagacg 23760

gcgacagagc ttgttgaaga tacgcaggcc attgcccgtt acggtcgtaa tgttacgaag 23820

atggatgcct ttggctgtac cagccggggg caggcacacc gcgccgggct gtggctgatt 23880

aaaacagaac tgctggaaac gcagaccgtg gatttcagcg tcggcgcaga agggcttcgc 23940

catgtaccgg gcgatgttat tgaaatctgc gatgatgact atgccggtat cagcaccggt 24000

ggtcgtgtgc tggcggtgaa cagccagacc cggacgctga cgctcgaccg tgaaatcacg 24060

ctgccatcct ccggtaccgc gctgataagc ctggttgacg gaagtggcaa tccggtcagc 24120

gtggaggttc agtccgtcac cgacggcgtg aaggtaaaag tgagccgtgt tcctgacggt 24180

gttgctgaat acagcgtatg ggagctgaag ctgccgacgc tgcgccagcg actgttccgc 24240

tgcgtgagta tccgtgagaa cgacgacggc acgtatgcca tcaccgccgt gcagcatgtg 24300

ccggaaaaag aggccatcgt ggataacggg gcgcactttg acggcgaaca gagtggcacg 24360

gtgaatggtg tcacgccgcc agcggtgcag cacctgaccg cagaagtcac tgcagacagc 24420

ggggaatatc aggtgctggc gcgatgggac acaccgaagg tggtgaaggg cgtgagtttc 24480

ctgctccgtc tgaccgtaac agcggacgac ggcagtgagc ggctggtcag cacggcccgg 24540

acgacggaaa ccacataccg cttcacgcaa ctggcgctgg ggaactacag gctgacagtc 24600

cgggcggtaa atgcgtgggg gcagcagggc gatccggcgt cggtatcgtt ccggattgcc 24660

gcaccggcag caccgtcgag gattgagctg acgccgggct attttcagat aaccgccacg 24720

ccgcatcttg ccgtttatga cccgacggta cagtttgagt tctggttctc ggaaaagcag 24780

attgcggata tcagacaggt tgaaaccagc acgcgttatc ttggtacggc gctgtactgg 24840

atagccgcca gtatcaatat caaaccgggc catgattatt acttttatat ccgcagtgtg 24900

aacaccgttg gcaaatcggc attcgtggag gccgtcggtc gggcgagcga tgatgcggaa 24960

ggttacctgg atttttttcaa aggcaagata accgaatccc atctcggcaa ggagctgctg 25020

gaaaaagtcg agctgacgga ggataacgcc agcagactgg aggagtttc gaaagagtgg 25080

aaggatgcca gtgataagtg gaatgccatg tgggctgtca aaattgagca gaccaaagac 25140

ggcaaacatt atgtcgcggg tattggcctc agcatggagg acacggagga aggcaaactg 25200

agccagtttc tggttgccgc caatcgtatc gcatttattg acccggcaaa cgggaatgaa 25260

acgccgatgt ttgtggcgca gggcaaccag atattcatga acgacgtgtt cctgaagcgc 25320

ctgacggccc ccaccattac cagcggcggc aatcctccgg ccttttccct gacaccggac 25380

ggaaagctga ccgctaaaaa tgcggatatc agtggcagtg tgaatgcgaa ctccgggacg 25440

ctcagtaatg tgacgatagc tgaaaactgt acgataaacg gtacgctgag ggcggaaaaa 25500

atcgtcgggg acattgtaaa ggcggcgagc gcggcttttc cgcgccagcg tgaaagcagt 25560

gtggactggc cgtcaggtac ccgtactgtc accgtgaccg atgaccatcc ttttgatcgc 25620

cagatagtgg tgcttccgct gacgtttcgc ggaagtaagc gtactgtcag cggcaggaca 25680

acgtattcga tgtgttatct gaaagtactg atgaacggtg cggtgattta tgatggcgcg 25740

gcgaacgagg cggtacaggt gttctcccgt attgttgaca tgccagcggg tcggggaaac 25800

gtgatcctga cgttcacgct tacgtccaca cggcattcgg cagatattcc gccgtatacg 25860

tttgccagcg atgtgcaggt tatggtgatt aagaaacagg cgctgggcat cagcgtggtc 25920

tgagtgtgtt acagaggttc gtccgggaac gggcgtttta ttataaaaca gtgagaggtg 25980

aacgatgcgt aatgtgtgta ttgccgttgc tgtctttgcc gcacttgcgg tgacagtcac 26040

tccggcccgt gcggaaggtg gacatggtac gtttacggtg ggctattttc aagtgaaacc 26100

gggtacattg ccgtcgttgt cgggcgggga taccggtgtg agtcatctga aagggattaa 26160

cgtgaagtac cgttatgagc tgacggacag tgtgggggtg atggcttccc tggggttcgc 26220

cgcgtcgaaa aagagcagca cagtgatgac cggggaggat acgtttcact atgagagcct 26280

gcgtggacgt tatgtgagcg tgatggccgg accggtttta caaatcagta agcaggtcag 26340

tgcgtacgcc atggccggag tggctcacag tcggtggtcc ggcagtacaa tggattaccg 26400

taagacggaa atcactcccg ggtatatgaa agagacgacc actgccaggg acgaaagtgc 26460

aatgcggcat acctcagtgg cgtggagtgc aggtatacag attaatccgg cagcgtccgt 26520

cgttgttgat attgcttatg aaggctccgg cagtggcgac tggcgtactg acggattcat 26580

cgttggggtc ggttataaat tctgattagc caggtaacac agtgttatga cagcccgccg 26640

gaaccggtgg gctttttttgt ggggtgaata tggcagtaaa gatttcagga gtcctgaaag 26700

acggcacagg aaaaccggta cagaactgca ccattcagct gaaagccaga cgtaacagca 26760

ccacggtggt ggtgaacacg gtgggctcag agaatccgga tgaagccggg cgttacagca 26820

tggatgtgga gtacggtcag tacagtgtca tcctgcaggt tgacggtttt ccaccatcgc 26880

acgccgggac catcaccgtg tatgaagatt cacaaccggg gacgctgaat gattttctct 26940

gtgccatgac ggaggatgat gcccggccgg aggtgctgcg tcgtcttgaa ctgatggtgg 27000

aagaggtggc gcgtaacgcg tccgtggtgg cacagagtac ggcagacgcg aagaaatcag 27060

ccggcgatgc cagtgcatca gctgctcagg tcgcggccct tgtgactgat gcaactgact 27120

cagcacgcgc cgccagcacg tccgccggac aggctgcatc gtcagctcag gaagcgtcct 27180

ccggcgcaga agcggcatca gcaaaggcca ctgaagcgga aaaaagtgcc gcagccgcag 27240

agtcctcaaa aaacgcggcg gccaccagtg ccggtgcggc gaaaacgtca gaaacgaatg 27300

ctgcagcgtc acaacaatca gccgccacgt ctgcctccac cgcggccacg aaagcgtcag 27360

aggccgccac ttcagcacga gatgcggtgg cctcaaaaga ggcagcaaaa tcatcagaaa 27420

cgaacgcatc atcaagtgcc ggtcgtgcag cttcctcggc aacggcggca gaaaattctg 27480

ccagggcggc aaaaacgtcc gagacgaatg ccaggtcatc tgaaacagca gcggaacgga 27540

gcgcctctgc cgcggcagac gcaaaaacag cggcggcggg gagtgcgtca acggcatcca 27600

cgaaggcgac agaggctgcg ggaagtgcgg tatcagcatc gcagagcaaa agtgcggcag 27660

aagcggcggc aatacgtgca gaaaattcgg caaaacgtgc agaagatata gcttcagctg 27720

tcgcgcttga ggatgcggac acaacgagaa aggggatagt gcagctcagc agtgcaacca 27780

acagcacgtc tgaaacgctt gctgcaacgc caaaggcggt taaggtggta atggatgaaa 27840

cgaacagaaa agcccactgg acagtccggc actgaccgga acgccaacag caccaaccgc 27900

gctcagggga acaaacaata cccagattgc gaacaccgct tttgtactgg ccgcgattgc 27960

agatgttatc gacgcgtcac ctgacgcact gaatacgctg aatgaactgg ccgcagcgct 28020

cgggaatgat ccagattttg ctaccaccat gactaacgcg cttgcgggta aacaaccgaa 28080

gaatgcgaca ctgacggcgc tggcagggct ttccacggcg aaaaataaat taccgtattt 28140

tgcggaaaat gatgccgcca gcctgactga actgactcag gttggcaggg atattctggc 28200

aaaaaattcc gttgcagatg ttcttgaata ccttggggcc ggtgagaatt ctaagcggag 28260

atcgcctagt gattttaaac tattgctggc agcattcttg agtccaatat aaaagtattg 28320

tgtacctttt gctgggtcag gttgttcttt aggaggagta aaaggatcaa atgcactaaa 28380

cgaaactgaa acaagcgatc gaaaatatcc ctttgggatt cttgactcga taagtctatt 28440

attttcagag aaaaaatatt cattgttttc tgggttggtg attgcaccaa tcattccatt 28500

caaaattgtt gttttaccac acccattccg cccgataaaa gcatgaatgt tcgtgctggg 28560

catagaatta accgtcacct caaaaggtat agttaaatca ctgaatccgg gagcactttt 28620

tctattaaat gaaaagtgga aatctgacaa ttctggcaaa ccatttaaca cacgtgcgaa 28680

ctgtccatga atttctgaaa gagttacccc tctaagtaat gaggtgttaa ggacgctttc 28740

attttcaatg tcggctaatc gatttggcca tactactaaa tcctgaatag ctttaagaag 28800

gttatgttta aaaccatcgc ttaatttgct gagattaaca tagtagtcaa tgctttcacc 28860

taaggaaaaa aacatttcag ggagttgact gaatttttta tctattaatg aataagtgct 28920

tacttcttct ttttgaccta caaaaccaat tttaacattt ccgatatcgc atttttcacc 28980

atgctcatca aagacagtaa gataaaacat tgtaacaaag gaatagtcat tccaaccatc 29040

tgctcgtagg aatgccttat ttttttctac tgcaggaata tacccgcctc tttcaataac 29100

actaaactcc aacatatagt aacccttaat tttattaaaa taaccgcaat ttatttggcg 29160

gcaacacagg atctctcttt taagttactc tctattacat acgttttcca tctaaaaatt 29220

agtagtattg aacttaacgg ggcatcgtat tgtagttttc catatttagc tttctgcttc 29280

cttttggata acccactgtt attcatgttg catggtgcac tgtttatacc aacgatatag 29340

tctattaatg catatatagt atcgccgaac gattagctct tcaggcttct gaagaagcgt 29400

ttcaagtact aataagccga tagatagcca cggacttcgt agccattttt cataagtgtt 29460

aacttccgct cctcgctcat aacagacatt cactacagtt atggcggaaa ggtatgcatg 29520

ctgggtgtgg ggaagtcgtg aaagaaaaga agtcagctgc gtcgtttgac atcactgcta 29580

tcttcttact ggttatgcag gtcgtagtgg gtggcacaca aagctttgca ctggattgcg 29640

aggctttgtg cttctctgga gtgcgacagg tttgatgaca aaaaattagc gcaagaagac 29700

aaaaatcacc ttgcgctaat gctctgttac aggtcactaa taccatctaa gtagttgatt 29760

catagtgact gcatatgttg tgttttacag tattatgtag tctgtttttt atgcaaaatc 29820

taatttaata tattgatatt tatatcattt tacgtttctc gttca 29865

<210> 11
<211> 2244
<212> ADN
<213> Escherichia coli

<220>
<221> fuente
<222> 1..2244
<223> /mol_tipo="ADN"/organismo="Escherichia coli"

<400> 11
atggcgcgtt ccaaaactgc tcagccaaaa cactcactgc gtaaaatcgc agttgtagta 60

gccacagcgg ttagcggcat gtctgtttat gcacaggcag cggttgaacc gaaagaagac 120

actatcaccg ttaccgctgc acctgcgccg caagaaagcg catgggggcc tgctgcaact 180

attgcggcgc gacagtctgc taccggcact aaaaccgata cgccgattca aaaagtgcca 240

cagtctattt ctgttgtgac cgccgaagag atggcgctgc atcagccgaa gtcggtaaaa 300

gaagcgctta gctacacgcc gggtgtctct gttggtacgc gtggcgcatc caacacctat 360

gaccacctga tcattcgcgg ctttgcggca gaaggccaaa gccagaataa ctatctgaat 420

ggcctgaagt tgcagggcaa cttctataac gatgcggtca ttgacccgta tatgctggaa 480

cgcgctgaaa ttatgcgtgg cccggtttcc gtgctttacg gtaaaagcag tcctggcggc 540

ctgttgaata tggtcagcaa gcgtccgacc accgaaccgc tgaaagaagt tcagtttaaa 600

gccggtactg acagcctgtt ccagactggt tttgacttta gcgattcgtt ggatgatgac 660

ggtgtttact cttatcgcct gaccggtctt gcgcgttctg ccaatgccca gcagaaaggg 720

tcagaagagc agcgttatgc tattgcaccg gcgttcacct ggcgtccgga tgataaaacc 780

aattttacct tcctttctta cttccagaac gagccggaaa ccggttatta cggctggttg 840

ccgaaagagg gaaccgttga gccgctgccg aacggtaagc gtctgccgac agactttaat 900

gaaggggcga agaacaacac ctattctcgt aatgagaaga tggtcggcta cagcttcgat 960

cacgaattta acgacacctt tactgtgcgt cagaacctgc gctttgctga aaacaaaacc 1020

tcgcaaaaca gcgtttatgg ttacggcgtc tgctccgatc cggcgaatgc ttacagcaaa 1080

cagtgtgcgg cattagcgcc agcggataaa ggccattatc tggcacgtaa atacgtcgtt 1140

gatgatgaga agctgcaaaa cttctccgtt gatacccagt tgcagagcaa gtttgccact 1200

ggcgatatcg accacaccct gctgaccggt gtcgacttta tgcgtatgcg taatgacatc 1260

aacgcctggt ttggttacga cgactctgtg ccactgctca atctgtacaa tccggtgaat 1320

accgatttcg acttcaatgc caaagatccg gcaaactccg gcccttaccg cattctgaat 1380

aaacagaaac aaacgggcgt ttatgttcag gatcaggcgc agtgggataa agtgctggtc 1440

accctaggcg gtcgttatga ctgggcagat caagaatctc ttaaccgcgt tgccgggacg 1500

accgataaac gtgatgacaa acagtttacc tggcgtggtg gtgttaacta cctgtttgat 1560

aatggtgtaa caccttactt cagctatagc gaatcgtttg aaccttcttc gcaagttggg 1620

aaggatggta atattttcgc accgtctaaa ggtaagcagt atgaagtcgg cgtgaaatat 1680

gtaccggaag atcgtccgat tgtagttact ggtgccgtgt ataatctcac taaaaccaac 1740

aacctgatgg cggaccctga gggttccttc ttctcggttg aaggtggcga gatccgcgca 1800

cgtggcgtag aaatcgaagc gaaagcggcg ctgtcggcga gtgttaacgt agtcggttct 1860

tatacttaca ccgatgcgga atacaccacc gatactacct ataaaggcaa tacgcctgca 1920

caggtgccaa aacacatggc ttcgttgtgg gctgactaca ccttctttga cggtccgctt 1980

tcaggtctga cgctgggcac cggtggtcgt tatactggct ccagttatgg tgatccggct 2040

aactccttta aagtgggaag ttatacggtc gtggatgcgt tagtacgtta tgatctggcg 2100

cgagtcggca tggctggctc caacgtggcg ctgcatgtta acaacctgtt cgatcgtgaa 2160

tacgtcgcca gctgctttaa cacttatggc tgcttctggg gcgcagaacg tcaggtcgtt 2220

gcaaccgcaa ccttccgttt ctaa 2244

<210> 12
<211> 1149
<212> ADN
<213> Escherichia coli

<220>
<221> fuente
<222> 1..1149
<223> /mol_tipo="ADN"/organismo="Escherichia coli"

<400> 12

atgagtctga aagaaaaaac acaatctctg tttgccaacg catttggcta ccctgccact 60

cacaccattc aggcgcctgg ccgcgtgaat ttgattggtg aacacaccga ctacaacgac 120

ggtttcgttc tgccctgcgc gattgattat caaaccgtga tcagttgtgc accacgcgat 180

gaccgtaaag ttcgcgtgat ggcagccgat tatgaaaatc agctcgacga gttttccctc 240

gatgcgccca ttgtcgcaca tgaaaactat caatgggcta actacgttcg tggcgtggtg 300

aaacatctgc aactgcgtaa caacagcttc ggcggcgtgg acatggtgat cagcggcaat 360

gtgccgcagg gtgccgggtt aagttcttcc gcttcactgg aagtcgcggt cggaaccgta 420

ttgcagcagc tttatcatct gccgctggac ggcgcacaaa tcgcgcttaa cggtcaggaa 480

gcagaaaacc agtttgtagg ctgtaactgc gggatcatgg atcagctaat ttccgcgctc 540

ggcaagaaag atcatgcctt gctgatcgat tgccgctcac tggggaccaa agcagtttcc 600

atgcccaaag gtgtggctgt cgtcatcatc aacagtaact tcaaacgtac cctggttggc 660

agcgaataca acacccgtcg tgaacagtgc gaaaccggtg cgcgtttctt ccagcagcca 720

gccctgcgtg atgtcaccat tgaagagttc aacgctgttg cgcatgaact ggacccgatc 780

gtggcaaaac gcgtgcgtca tatactgact gaaaacgccc gcaccgttga agctgccagc 840

gcgctggagc aaggcgacct gaaacgtatg ggcgagttga tggcggagtc tcatgcctct 900

atgcgcgatg atttcgaaat caccgtgccg caaattgaca ctctggtaga aatcgtcaaa 960

gctgtgattg gcgacaaagg tggcgtacgc atgaccggcg gcggatttgg cggctgtatc 1020

gtcgcgctga tcccggaaga gctggtgcct gccgtacagc aagctgtcgc tgaacaatat 1080

gaagcaaaaa caggtattaa agagactttt tacgtttgta aaccatcaca aggagcagga 1140

cagtgctga 1149

<210> 13
<211> 1701
<212> ADN
<213> Escherichia coli

<220>
<221> fuente
<222> 1..1701
<223> /mol_tipo="ADN"/organismo="Escherichia coli"

<400> 13

```
atggcgattg caattggcct cgattttggc agtgattctg tgcgagcttt ggcggtggac      60
tgcgctaccg gtgaagagat cgccaccagc gtagagtggt atccccgttg gcagaaaggg     120
caattttgtg atgccccgaa taaccagttc cgtcatcatc cgcgtgacta cattgagtca     180
atggaagcgg cactgaaaac cgtgcttgca gagcttagcg tcgaacagcg cgcagctgtg     240
gtcgggattg gcgttgacag taccggctcg acgcccgcac cgattgatgc cgacggaaac     300
gtgctggcgc tgcgcccgga gtttgccgaa aacccgaacg cgatgttcgt attgtggaaa     360
gaccacactg cggttgaaga agcggaagag attacccgtt tgtgccacgc gccgggcaac     420
gttgactact cccgctacat tggtggtatt tattccagcg aatggttctg ggcaaaaatc     480
ctgcatgtga ctcgccagga cagcgccgtg gcgcaatctg ccgcatcgtg gattgagctg     540
tgcgactggg tgccagctct gctttccggt accacccgcc cgcaggatat tcgtcgcgga     600
cgttgcagcg ccgggcataa atctctgtgg cacgaaagct ggggcggcct gccgccagcc     660
agtttctttg atgagctgga cccgatcctc aatcgccatt tgccttcccc gctgttcact     720
gacacttgga ctgccgatat tccggtgggc accttatgcc cggaatgggc gcagcgtctc     780
ggcctgcctg aaagcgtggt gatttccggc ggcgcgtttg actgccatat gggcgcagtt     840
ggcgcaggcg cacagcctaa cgcactggta aaagttatcg gtacttccac ctgcgacatt     900
ctgattgccg acaaacagag cgttggcgag cgggcagtta aaggtatttg cggtcaggtt     960
gatggcagcg tggtgcctgg atttatcggt ctggaagcag gccaatcggc gtttggtgat    1020
atctacgcct ggtttggtcg cgtactcggc tggccgctgg aacagcttgc cgcccagcat    1080
ccggaactga aaacgcaaat caacgccagc cagaaacaac tgcttccggc gctgaccgaa    1140
gcatgggcca aaaatccgtc tctggatcac ctgccggtgg tgctcgactg gtttaacggc    1200
cgccgcacac cgaacgctaa ccaacgcctg aaaggggtga ttaccgatct taacctcgct    1260
accgacgctc cgctgctgtt cggcggtttg attgctgcca ccgcctttgg cgcacgcgca    1320
atcatggagt gctttaccga tcaggggatc gccgttaata acgtgatggc actgggcggc    1380
atcgcgcgga aaaaccaggt cattatgcag gcctgctgcg acgtgctgaa tcgcccgctg    1440
caaattgttg cctctgacca gtgctgtgcg ctcggtgcgg cgatttttgc tgccgtcgcc    1500
gcgaaagtgc acgcagacat cccatcagct cagcaaaaaa tggccagtgc ggtagagaaa    1560
accctgcaac cgtgcagcga gcaggcacaa cgctttgaac agctttatcg ccgctatcag    1620
caatgggcga tgagcgccga acaacactat cttccaactt ccgccccggc acaggctgcc    1680
caggccgttg cgactctata a                                              1701
```

<210> 14
<211> 1503
<212> ADN
<213> Escherichia coli

<220>
<221> fuente
<222> 1..1503
<223> /mol_tipo="ADN"/organismo="Escherichia coli"

<400> 14

atgacgattt ttgataatta tgaagtgtgg tttgtcattg gcagccagca tctgtatggc 60

ccggaaaccc tgcgtcaggt cacccaacat gccgagcacg tcgttaatgc gctgaatacg 120

gaagcgaaac tgccctgcaa actggtgttg aaaccgctgg gcaccacgcc ggatgaaatc 180

accgctattt gccgcgacgc gaattacgac gatcgttgcg ctggtctggt ggtgtggctg 240

cacaccttct ccccggccaa aatgtggatc aacggcctga ccatgctcaa caaaccgttg 300

ctgcaattcc acacccagtt caacgcggcg ctgccgtggg acagtatcga tatggacttt 360

atgaacctga accagactgc acatggcggt cgcgagttcg gcttcattgg cgcgcgtatg 420

cgtcagcaac atgccgtggt taccggtcac tggcaggata aacaagccca tgagcgtatc 480

ggctcctgga tgcgtcaggc ggtctctaaa caggataccc gtcatctgaa agtctgccga 540

tttggcgata acatgcgtga agtggcggtc accgatggcg ataaagttgc cgcacagatc 600

aagttcggtt tctccgtcaa tacctgggcg gttggcgatc tggtgcaggt ggtgaactcc 660

atcagcgacg gcgatgttaa cgcgctggtc gatgagtacg aaagctgcta caccatgacg 720

cctgccacac aaatccacgg caaaaaacga cagaacgtgc tggaagcggc gcgtattgag 780

ctggggatga agcgtttcct ggaacaaggt ggcttccacg cgttcaccac cacctttgaa 840

gatttgcacg gtctgaaaca gcttcctggt ctggccgtac agcgtctgat gcagcagggt 900

tacggctttg cgggcgaagg cgactggaaa actgccgccc tgcttcgcat catgaaggtg 960

atgtcaaccg gtctgcaggg cggcacctcc tttatggagg actacaccta tcacttcgag 1020

aaaggtaatg acctggtgct cggctcccat atgctggaag tctgcccgtc gatcgccgca 1080

gaagagaaac cgatcctcga cgttcagcat ctcggtattg gtggtaagga cgatcctgcc 1140

cgcctgatct tcaataccca aaccggccca gcgattgtcg ccagcttgat tgatctcggc 1200

gatcgttacc gtctactggt taactgcatc gacacggtga aaacaccgca ctccctgccg 1260

aaactgccgg tggcgaatgc gctgtggaaa gcgcaaccgg atctgccaac tgcttccgaa 1320

gcgtggatcc tcgctggtgg cgcgcaccat accgtcttca gccatgcact gaacctcaac 1380

gatatgcgcc aattcgccga gatgcacgac attgaaatca cggtgattga taacgacaca 1440

cgcctgccag cgtttaaaga cgcgctgcgc tggaacgaag tgtattacgg gtttcgtcgc 1500

taa 1503

<210> 15
<211> 2355
<212> ADN
<213> Escherichia coli

<220>
<221> fuente
<222> 1..2355
<223> /mol_tipo="ADN”/organismo="Escherichia coli"

<400> 15

atgaatcctg agcgttctga acgcattgaa atccccgtat tgccgctgcg cgatgtggtg 60

gtttatccgc acatggtcat ccccttattt gtcgggcggg aaaaatctat ccgttgtctg 120

gaagcggcga tggaccatga taaaaaaatt atgctggtcg cgcagaaaga agcttcaacg 180

gatgagccgg gtgtaaacga tcttttcacc gtcgggaccg tggcctctat attgcagatg 240

ctgaaactgc ctgacggcac cgtcaaagtg ctggtcgagg ggttacagcg cgcgcgtatt 300

tctgcgctct ctgacaatgg cgaacacttt tctgcgaagg cggagtatct ggagtcgccg 360

accattgatg agcgggaaca ggaagtgctg gtgcgtactg caatcagcca gttcgaaggc 420

tacatcaagc tgaacaaaaa aatcccacca gaagtgctga cgtcgctgaa tagcatcgac 480

gatccggcgc gtctggcgga taccattgct gcacatatgc cgctgaaact ggctgacaaa 540

cagtctgttc tggagatgtc cgacgttaac gaacgtctgg aatatctgat ggcaatgatg 600

gaatcggaaa tcgatctgct gcaggttgag aaacgcattc gcaaccgcgt taaaaagcag 660

atggagaaat cccagcgtga gtactatctg aacgagcaaa tgaaagctat tcagaaagaa 720

ctcggtgaaa tggacgacgc gccggacgaa aacgaagccc tgaagcgcaa aatcgacgcg 780

gcgaagatgc cgaaagaggc aaaagagaaa gcggaagcag agttgcagaa gctgaaaatg 840

atgtctccga tgtcggcaga agcgaccgta gtgcgtggtt atatcgactg gatggtacag 900

gtgccgtgga atgcgcgtag caaggtcaaa aaagacctgc gtcaggcgca ggaaatcctt 960

gataccgacc attatggtct ggagcgcgtg aaagatcgaa tccttgagta tcttgcggtt 1020

caaagccgtg tcaacaaaat caagggaccg atcctctgcc tggtagggcc gccgggggta 1080

ggtaaaacct ctcttggtca gtccattgcc aaagccaccg ggcgtaaata tgtccgtatg 1140

gcgctgggcg gcgtgcgtga tgaagcggaa atccgtggtc accgccgtac ttacatcggt 1200

tctatgccgg gtaaactgat ccagaaaatg gcgaaagtgg gcgtgaaaaa cccgctgttc 1260

ctgctcgatg agatcgacaa aatgtcttct gacatgcgtg gcgatccggc ctctgcactg 1320

cttgaagtgc tggatccaga gcagaacgta gcgttcagcg accactacct ggaagtggat 1380

tacgatctca gcgacgtgat gtttgtcgcg acgtcgaact ccatgaacat tccggcaccg 1440

ctgctggatc gtatggaagt gattcgcctc tccggttata ccgaagatga aaaactgaac 1500

atcgccaaac gtcacctgct gccgaagcag attgaacgta atgcactgaa aaaaggtgag 1560

ctgaccgtcg acgatagcgc cattatcggc attattcgtt actacacccg tgaggcgggc 1620

gtgcgtggtc tggagcgtga aatctccaaa ctgtgtcgca aagcggttaa gcagttactg 1680

ctcgataagt cattaaaaca tatcgaaatt aacggcgata acctgcatga ctatctcggt 1740

gttcagcgtt tcgactatgg tcgcgcggat aacgaaaacc gtgtcggtca ggtaaccggt 1800

ctggcgtgga cggaagtggg cggtgacttg ctgaccattg aaaccgcatg tgttccgggt 1860

aaaggcaaac tgacctatac cggttcgctc ggcgaagtga tgcaggagtc cattcaggcg 1920

gcgttaacgg tggttcgtgc gcgtgcggaa aaactgggga tcaaccctga tttttacgaa 1980

aaacgtgaca tccacgtcca cgtaccggaa ggtgcgacgc cgaaagatgg tccgagtgcc 2040

ggtattgcta tgtgcaccgc gctggtttct tgcctgaccg gtaacccggt tcgtgccgat 2100

gtggcaatga ccggtgagat cactctgcgt ggtcaggtac tgccgatcgg tggtttgaaa 2160

gaaaaactcc tggcagcgca tcgcggcggg attaaaacag tgctaattcc gttcgaaaat 2220

aaacgcgatc tggaagagat tcctgacaac gtaattgccg atctggacat tcatcctgtg 2280

aagcgcattg aggaagttct gactctggcg ctgcaaaatg aaccgtctgg tatgcaggtt 2340

gtgactgcaa aatag 2355

<210> 16
<211> 954
<212> ADN
<213> Escherichia coli

<220>
<221> fuente
<222> 1..954
<223> /mol_tipo="ADN"/organismo="Escherichia coli"

<400> 16

atgcgggcga aacttctggg aatagtcctg acaaccccta ttgcgatcag ctcttttgct 60

tctaccgaga ctttatcgtt tactcctgac aacataaatg cggacattag tcttggaact 120

ctgagcggaa aaacaaaaga gcgtgtttat ctagccgaag aaggaggccg aaaagtcagt 180

caactcgact ggaaattcaa taacgctgca attattaaag gtgcaattaa ttgggatttg 240

atgccccaga tatctatcgg ggctgctggc tggacaactc tcggcagccg aggtggcaat 300

atggtcgatc aggactggat ggattccagt aaccccggaa cctggacgga tgaaagtaga 360

caccctgata cacaactcaa ttatgccaac gaatttgatc tgaatatcaa aggctggctc 420

ctcaacgaac ccaattaccg cctgggactc atggccggat atcaggaaag ccgttatagc 480

tttacagcca gaggtggttc ctatatctac agttctgagg agggattcag agatgatatc 540

ggctccttcc cgaatggaga aagagcaatc ggctacaaac aacgttttaa aatgccctac 600

attggcttga ctggaagtta tcgttatgaa gattttgaac tcggtggcac atttaaatac 660

agcggctggg tggaatcatc tgataacgat gaacactatg acccgggaaa aagaatcact 720

tatcgcagta aggtcaaaga ccaaaattac tattctgttg cagtcaatgc aggttattac 780

gtcacaccta acgcaaaagt ttatgttgaa ggcgcatgga atcgggttac gaataaaaaa 840

ggtaatactt cactttatga tcacaataat aacacttcag actacagcaa aaatggagca 900

ggtatagaaa actataactt catcactact gctggtctta agtacacatt ttaa 954

<210> 17
<211> 624
<212> ADN
<213> Escherichia coli

<220>
<221> fuente
<222> 1..624
<223> /mol_tipo="ADN"/organismo="Escherichia coli"

<400> 17
atgtcaacga ttattatgga tttatgtagt tacacccgac taggtttaac cgggtatctg 60

ttgagtagag gggttaaaaa aagagaaatc aacgacattg aaaccgttga tgaccttgcc 120

atagcttgtg attcacagcg cccttcagtg gtgtttatta atgaggactg tttcatccac 180

gatgcttcta acagtcagcg tatcaagctc atcattaatc aacatcccaa tacgttattt 240

atcgttttta tggcaattgc caatgttcat tttgatgaat atctattggt cagaaaaaat 300

ttattgatca gttctaaatc gattaaaccg gaatctctcg acgatatcct tggcgatatt 360

ctgaaaaaag agacaacgat aacctcgttt ttaaatatgc cgacgttatc attgagccga 420

accgaatcga gtatgttgcg aatgtggatg gcaggtcagg gaaccattca aatctctgac 480

caaatgaata tcaaagccaa gaccgtttca tcgcataaag gtaatattaa acgtaagatc 540

aaaacgcata ataaacaggt tatctaccat gtcgtccgac tgacggataa tgtgactaat 600

ggtatttttg tcaacatgcg ctaa 624

<210> 18
<211> 3567
<212> ADN
<213> Escherichia coli

<220>
<221> fuente
<222> 1..3567
<223> /mol_tipo="ADN"/organismo="Escherichia coli"

<400> 18
atgttatggg ccttaaatat ttggacaggc ccgcacagca atggattaat aacaatgatg 60

aataaatcca attttgaatt cctgaagggc gtcaacgact tcacttatgc catcgcctgt 120

gcggcggaaa ataactaccc ggatgatccc aacacgacgc tgattaaaat gcgtatgttt 180

ggcgaagcca cagcgaaaca tcttggtctg ttactcaaca tccccccttg tgagaatcaa 240

cacgatctcc tgcgtgaact cggcaaaatc gcctttgttg atgacaacat cctctctgta 300

tttcacaaat tacgccgcat tggtaaccag gcggtgcacg aatatcataa cgatctcaac 360

gatgcccaga tgtgcctgcg actcgggttc cgcctggctg tctggtacta ccgtctggtc 420

actaaagatt atgacttccc ggtgccggtg tttgtgttgc cggaacgtgg tgaaaacctc 480

tatcaccagg aagtgctgac gctaaaacaa cagcttgaac agcaggtgcg agaaaaagcg 540

cagactcagg cagaagtcga agcgcaacag cagaagctgg ttgccctgaa cggctatatc 600

gccattctgg aaggcaaaca gcaggaaacc gaagcgcaaa cccaggctcg ccttgcggca 660

ctggaagcac agctcgccga gaagaacgcg gaactggcaa aacagaccga acaggaacgt 720

aaggcttacc acaaagaaat taccgatcag gccatcaagc gcacactcaa ccttagcgaa 780

gaagagagtc gcttcctgat tgatgcgcaa ctgcgtaaag caggctggca ggccgacagc 840

aaaaccctgc gcttctccaa aggcgcacgt ccggaacccg gcgtcaataa agccattgcc 900

gaatggccga ccggaaaaga tgaaacgggt aatcagggct ttgcggatta tgtgctgttt 960

gtcggcctca aacccatcgc ggtggtagag gcgaaacgta acaatatcga cgttcccgcc 1020

aggctcaatg agtcgtatcg ctacagtaaa tgtttcgata atggcttcct gcgggaaacc 1080

ttgcttgagc actactcacc ggatgaagtg catgaagcag tgccagagta tgaaaccagc 1140

tggcaggaca ccagcggcaa acaacggttt aaaatcccct tctgctactc gaccaacggg 1200

cgcgaatacc gcgcaacaat gaagaccaaa agcggcatct ggtatcgcga cgtgcgtgat 1260

acccgcaata tgtcgaaagc cttacccgag tggcaccgcc cggaagagct gctggaaatg 1320

ctcggcagcg aaccgcaaaa acagaatcag tggtttgccg ataaccctgg catgagcgag 1380

ctgggcctgc gttattatca ggaagatgcc gtccgcgcgg ttgaaaaggc aatcgtcaag 1440

gggcaacaag agatcctgct ggcgatggcg accggtaccg gtaaaacccg tacggcaatc 1500

gccatgatgt tccgcctgat ccagtcccag cgttttaaac gcattctctt ccttgtcgac 1560

cgccgttctc ttggcgaaca ggcgctgggc gcgtttgaag atacgcgtat taacggcgac 1620

accttcaaca gcattttcga cattaaaggg ctgacggata aattcccgga agacagcacc 1680

aaaattcacg ttgccaccgt acagtcgctg gtgaaacgca ccctgcaatc agatgaaccg 1740

atgccggtgg cccgttacga ctgtatcgtc gttgacgaag cgcatcgcgg ctatattctc 1800

gataaagagc agaccgaagg cgaactgcag ttccgcagcc agctggatta cgtctctgcc 1860

taccgtcgca ttctcgatca cttcgatgcg gtaaaaatcg ctctcaccgc caccccggcg 1920

ctacatactg tgcagatttt cggcgagccg gtttaccgtt atacctaccg taccgcggtt 1980

atcgacggtt ttctgatcga ccaggatccg cctattcaga tcatcacccg caacgcgcag 2040

gagggggttt atctctccaa aggcgagcag gtagagcgca tcagcccgca gggagaagtg 2100

atcaatgaca ccctggaaga cgatcaggat tttgaagtcg ccgactttaa ccgtggcctg 2160

gtgatcccgg cgtttaaccg cgccgtctgt aacgaactca ccaattatct tgacccgacc 2220

ggatcgcaaa aaacgctggt cttctgcgtc accaatgccc atgccgatat ggtggtggaa 2280

gagctgcgtg ccgcgttcaa gaaaaagtat ccgcaactgg agcacgacgc gatcatcaag 2340

atcaccggtg atgccgataa agacgcgcgc aaagtgcaga ccatgatcac ccgcttcaat 2400

aaagagcggc tgcccaatat cgtggtaacc gtcgacctgc tgacgaccgg cgtcgatatt 2460

ccgtcgatct gtaatatcgt gttcctgcgt aaagtacgca gccgcattct gtacgaacag 2520

atgaaaggcc gcgccacgcg cttatgcccg gaggtgaata aaaccagctt taagattttt 2580

gactgtgtcg atatctacag cacgctggag agcgtcgaca ccatgcgtcc ggtggtggtg 2640

cgcccgaagg tggaactgca aacgctggtc aatgaaatta ccgattcaga aacctataaa 2700

atcaccgaag cggatggccg cagttttgcc gagcacagcc atgaacaact ggtggcgaag 2760

ctccagcgta tcatcggtct ggccacgttt aaccgtgacc gcagcgaaac gatagataaa 2820

caggtgcgtc gtctggatga gctatgccag gacgcggcgg gcgtgaactt taacggcttc 2880

gcctcgcgcc tgcgggaaaa agggccgcac tggagcgccg aagtctttaa caaactgcct 2940

ggctttatcg cccgtctgga aaagctgaaa acggacatca acaacctgaa tgatgcgccg 3000

atcttcctcg atatcgacga tgaagtggtg agtgtaaaat cgctgtacgg tgattacgac 3060

acgccgcagg atttcctcga agcctttgac tcgctggtgc aacgttcccc gaacgcgcaa 3120

ccggcattgc aggcagttat taatcgcccg cgcgatctca cccgtaaagg gctggtcgag 3180

ctacaggagt ggtttgaccg ccagcacttt gaggaatctt ccctgcgcaa agcatggaaa 3240

gagacgcgca atgaagatat cgccgcccgg ctgattggtc atattcgccg cgctgcggtg 3300

ggcgatgcgc tgaaaccgtt tgaggaacgt gtcgatcacg cgctgacgcg cattaagggc 3360

gaaaacgact ggagcagcga gcaattaagc tggctcgatc gtttagcgca ggcgctgaaa 3420

gagaaagtgg tgctcgacga cgatgtcttc aaaaccggca acttccaccg tcgcggcggg 3480

aaggcgatgc tgcaaagaac ctttgacgat aatctcgata ccctgctggg caaattcagc 3540

gattatatct gggacgagct ggcctga 3567

<210> 19
<211> 915
<212> ADN
<213> Escherichia coli

<220>
<221> fuente
<222> 1..915
<223> /mol_tipo="ADN”/organismo="Escherichia coli"

<400> 19

atgacggttc ctacctatga caaatttatt gaacctgttc tgcgttatct ggcaacaaaa 60

ccggaaggtg cagccgcgcg tgatgttcat gaggctgccg cggatgcatt aggactggat 120

gacagccagc gagcgaaagt cattaccagc ggacaacttg tttataaaaa tcgtgcaggc 180

tgggcgcatg accgtttaaa acgtgccggg ttgtcgcaaa gtttgtcgcg tggcaaatgg 240

tgcctgactc ctgcgggttt tgactgggtt gcgtctcatc cccagccaat gacggagcag 300

gagacgaacc atctggcctt cgcttttgtg aatgtcaaac ttaagtcacg gccggatgcc 360

gtcgatttag atccgaaagc cgactctccc gatcatgaag aacttgcaaa gagcagcccg 420

gacgatcggt tagatcaggc gctaaaagag cttcgtgatg cggtggctga tgaggttctg 480

gaaaacttat tgcaggtttc tccttcgcgc tttgaagtca ttgttctgga tgttttgcat 540

cgcctggggt atggcggcca ccgtgatgat ttgcagcgtg ttggcggtac tggagatggt 600

ggcatcgatg gtgtgatatc gcttgataaa cttggcctgg agaaagttta tgttcaggca 660

aaacgttggc agaatactgt aggcaggcca gaattacagg cattttacgg cgcactggct 720

gggcaaaaag cgaaacgtgg ggtgtttatt accacttctg gatttacttc tcaggcgcgt 780

gactttgccc aatccgtcga gggtatggtg ttggttgatg gggaacgcct ggtgcactta 840

atgatcgaaa acgaagtagg ggtttcttca cgtttgttga aggtgccgaa actggatatg 900

gactattttg agtga 915

<210> 20
<211> 708
<212> ADN
<213> Escherichia coli

<220>
<221> fuente
<222> 1..708
<223> /mol_tipo="ADN”/organismo="Escherichia coli"

<400> 20

```
atgtaccgtt atttgtctat tgctgcggtg gtactgagcg cagcattttc cggcccggcg      60
ttggccgaag gtatcaatag tttttctcag gcgaaagccg cggcggtaaa agtccacgct     120
gacgcgcccg gtacgtttta ttgcggatgt aaaattaact ggcagggcaa aaaaggcgtt     180
gttgatctgc aatcgtgcgg ctatcaggtg cgcaaaaatg aaaaccgcgc cagccgcgta     240
gagtgggaac atgtcgttcc cgcctggcag ttcggtcacc agcgccagtg ctggcaggac     300
ggtggacgta aaaactgcgc taaagatccg gtctatcgca agatggaaag cgatatgcat     360
aacctgcagc cgtcagtcgg tgaggtgaat ggcgatcgcg gcaactttat gtacagccag     420
tggaatggcg gtgaaggcca gtacggtcaa tgcgccatga aggtcgattt caaagaaaaa     480
gctgccgaac caccagcgcg tgcacgcggt gccattgcgc gcacctactt ctatatgcgc     540
gaccaataca acctgacact ctctcgccag caaacgcagc tgttcaacgc atggaacaag     600
atgtatccgg ttaccgactg ggagtgcgag cgcgatgaac gcatcgcgaa ggtgcagggc     660
aatcataacc cgtatgtgca acgcgcttgc caggcgcgaa agagctaa                  708
```

<210> 21
<211> 1062
<212> ADN
<213> Escherichia coli

<220>
<221> fuente
<222> 1..1062
<223> /mol_tipo="ADN”/organismo="Escherichia coli"

<400> 21

atggctatcg acgaaaacaa acagaaagcg ttggcggcag cactgggcca gattgagaaa 60

caatttggta aaggctccat catgcgcctg ggtgaagacc gttccatgga tgtggaaacc 120

atctctaccg gttcgctttc actggatatc gcgcttgggg caggtggtct gccgatgggc 180

cgtatcgtcg aaatctacgg accggaatct tccggtaaaa ccacgctgac gctgcaggtg 240

atcgccgcag cgcagcgtga aggtaaaacc tgtgcgttta tcgatgctga acacgcgctg 300

gacccaatct acgcacgtaa actgggcgtc gatatcgaca acctgctgtg ctcccagccg 360

gacaccggcg agcaggcact ggaaatctgt gacgccctgg cgcgttctgg cgcagtagac 420

gttatcgtcg ttgactccgt ggcggcactg acgccgaaag cggaaatcga aggcgaaatc 480

ggcgactctc acatgggcct tgcggcacgt atgatgagcc aggcgatgcg taagctggcg 540

ggtaacctga agcagtccaa cacgctgctg atcttcatca accagatccg tatgaaaatt 600

ggtgtgatgt tcggtaaccc ggaaaccact accggtggta acgcgctgaa attctacgcc 660

tctgttcgtc tcgacatccg tcgtatcggc gcggtgaaag agggcgaaaa cgtggtgggt 720

agcgaaaccc gcgtgaaagt ggtgaagaac aaaatcgctg cgccgtttaa acaggctgaa 780

ttccagatcc tctacggcga aggtatcaac ttctacggcg aactggttga cctgggcgta 840

aaagagaagc tgatcgagaa agcaggcgcg tggtacagct acaaaggtga gaagatcggt 900

cagggtaaag cgaatgcgac tgcctggctg aaagataacc cggaaaccgc gaaagagatc 960

gagaagaaag tacgtgagtt gctgctgagc aacccgaact caacgccgga tttctctgta 1020

gatgatagcg aaggcgtagc agaaactaac gaagattttt aa 1062

<210> 22
<211> 306
<212> ADN
<213> Escherichia coli

<220>
<221> fuente
<222> 1..306
<223> /mol_tipo="ADN"/organismo="Escherichia coli"

<400> 22
atgcagttta aggtttacac ctataaaaga gagagccgtt atcgtctgtt tgtggatgta 60

cagagtgata ttattgacac gcccgggcga cggatggtga tccccctggc cagtgcacgt 120

ctgctgtcag ataaagtctc ccgtgaactt tacccggtgg tgcatatcgg ggatgaaagc 180

tggcgcatga tgaccaccga tatggccagt gtgccggtct ccgttatcgg ggaagaagtg 240

gctgatctca gccaccgcga aaatgacatc aaaaacgcca ttaacctgat gttctgggga 300

atataa 306

<210> 23
<211> 219
<212> ADN
<213> Escherichia coli

<220>
<221> fuente
<222> 1..219
<223> /mol_tipo="ADN”/organismo="Escherichia coli"

<400> 23
atgaagcagc gtattacagt gacagttgac agcgacagct atcagttgct caaggcatat 60

gatgtcaata tctccggtct ggtaagcaca accatgcaga atgaagcccg tcgtctgcgt 120

gccgaacgct ggaaagcgga aaatcaggaa gggatggctg aggtcgcccg gtttattgaa 180

atgaacggct cttttgctga cgagaacagg gactggtga 219

<210> 24
<211> 59
<212> ADN
<213> secuencias artificiales

<220>
<221> fuente
<222> 1..59
<223> /mol_tipo="ADN" /nota="cebadores XisgalKstart”/organismo="secuencias artificiales"

<400> 24
gggggtaaat cccggcgctc atgacttcgc cttcttccca gaattcctgt tgacaatta 59

<210> 25
<211> 60
<212> ADN
<213> secuencias artificiales

<220>
<221> fuente
<222> 1..60
<223> /mol_tipo="ADN" /nota="parada XisgalK”/organismo="secuencias artificiales"

<400> 25
gttctgatta ttggaaatct tctttgccct ccagtgtgag cagcactgtc ctgctccttg 60

<210> 26
<211> 20
<212> ADN
<213> secuencias artificiales

<220>
<221> fuente
<222> 1..20
<223> /mol_tipo="ADN" /nota="cebador Xis1”/organismo="secuencias artificiales"

<400> 26
gtcttcaagt ggagcatcag 20

<210> 27
<211> 20
<212> ADN
<213> secuencias artificiales

<220>
<221> fuente
<222> 1..20
<223> /mol_tipo="ADN" /nota="cebador Xis4”/organismo="secuencias artificiales"

<400> 27
accaggacta tccgtatgac 20

<210> 28
<211> 35
<212> ADN
<213> secuencias artificiales

<220>
<221> fuente
<222> 1..35
<223> /mol_tipo="ADN" /nota="cebador Xis2”/organismo="secuencias artificiales"

<400> 28
ccaaacggaa cagatgaaga aggcgaagtc atgag 35

<210> 29
<211> 35
<212> ADN
<213> secuencias artificiales

<220>
<221> fuente
<222> 1..35
<223> /mol_tipo="ADN" /nota="cebador Xis3”/organismo="secuencias artificiales"

<400> 29
gacttcgcct tcttcatctg ttccgtttgg cttcc 35

<210> 30
<211> 21
<212> ADN
<213> secuencias artificiales

<220>
<221> fuente
<222> 1..21
<223> /mol_tipo="ADN" /nota="cebador Xis6”/organismo="secuencias artificiales"

<400> 30
gtaatggaaa gctggtagtc g 21

<210> 31
<211> 20
<212> ADN
<213> secuencias artificiales

<220>
<221> fuente
<222> 1..20
<223> /mol_tipo="ADN" /nota="cebador Xis5”/organismo="secuencias artificiales"

<400> 31
cagccgtaag tcttgatctc 20

<210> 32
<211> 20
<212> ADN
<213> secuencias artificiales

<220>
<221> fuente
<222> 1..20
<223> /mol_tipo="ADN" /nota="cebador Xis7”/organismo="secuencias artificiales"

<400> 32
cagcaggcat gatccaagag 20

<210> 33
<211> 35
<212> ADN
<213> secuencias artificiales

<220>
<221> fuente
<222> 1..35
<223> /mol_tipo="ADN" /nota="cebador Xis2b"/organismo="secuencias artificiales"

<400> 33
tttgccctcc agtgtgaaga aggcgaagtc atgag 35

<210> 34
<211> 39
<212> ADN
<213> secuencias artificiales

<220>
<221> fuente
<222> 1..39
<223> /mol_tipo="ADN" /nota="cebador Xis8"/organismo="secuencias artificiales"

<400> 34
ctcatgactt cgccttcttc acactggagg gcaaagaag 39

<210> 35
<211> 624
<212> ADN
<213> Fago lambda de enterobacteria

<220>
<221> fuente
<222> 1..624
<223> /mol_tipo="ADN"/organismo="Fago lambda de enterobacteria"

<400> 35
atgagactcg aaagcgtagc taaatttcat tcgccaaaaa gcccgatgat gagcgactca 60

ccacgggcca cggcttctga ctctctttcc ggtactgatg tgatggctgc tatggggatg 120

gcgcaatcac aagccggatt cggtatggct gcattctgcg gtaagcacga actcagccag 180

aacgacaaac aaaaggctat caactatctg atgcaatttg cacacaaggt atcggggaaa 240

taccgtggtg tggcaaagct tgaaggaaat actaaggcaa aggtactgca agtgctcgca 300

acattcgctt atgcggatta ttgccgtagt gccgcgacgc cgggggcaag atgcagagat 360

tgccatggta caggccgtgc ggttgatatt gccaaaacag agctgtgggg gagagttgtc 420

gagaaagagt gcggaagatg caaaggcgtc ggctattcaa ggatgccagc aagcgcagca 480

tatcgcgctg tgacgatgct aatcccaaac cttacccaac ccacctggtc acgcactgtt 540

aagccgctgt atgacgctct ggtggtgcaa tgccacaaag aagagtcaat cgcagacaac 600

attttgaatg cggtcacacg ttag 624

<210> 36
<211> 144

<212> ADN
<213> Fago lambda de enterobacteria

<220>
<221> fuente
<222> 1..144
<223> /mol_tipo="ADN”/organismo="Fago lambda de enterobacteria"

<400> 36

atggatcaaa cacttatggc tatccagact aaattcacta tcgccacttt tattggcgat 60

gaaaagatgt ttcgtgaagc cgtcgacgct tataaaaaat ggatattaat actgaaactg 120

agatcaagca aaagcattca ctaa 144

REIVINDICACIONES

1. Una célula anfitriona bacteriana que comprende un fago genéticamente modificado en la que

- se inserta un sistema de expresión, dicho sistema de expresión comprende una secuencia de ácidos nucleicos que codifica una proteína, dicha proteína que induce adicionalmente la expresión de una biomolécula de interés, y
- se suprimen los genes int, S, R, Rz, Xis, y

en la que dicho fago genéticamente modificado se integra en el genoma de dicha célula anfitriona bacteriana, y en la que el fago es el fago lambda, el fago 434, el fago phi80, el fago phi81, el fago HK97, o el fago P21.

2. La célula anfitriona bacteriana de acuerdo con la reivindicación 1, en la que se inactiva adicionalmente el gen Q.

3. La célula anfitriona bacteriana de acuerdo con la reivindicación 1 o 2, en la que el promotor comprendido en el sistema de expresión es inducible.

4. La célula anfitriona bacteriana de acuerdo con una cualquiera de las reivindicaciones 1 a 3, en la que el sistema de expresión es el sistema de expresión T7.

5. La célula anfitriona bacteriana de acuerdo con una cualquiera de las reivindicaciones 1 a 4, en la que dicho fago es uno de los fagos P11 o P13, en la que

- P11 comprende la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, R, Rz, Xis y Int, y
- P13 comprende la secuencia NC_001416 en la que se suprimen las secuencias de codificación de los genes S, R, Rz, Q, Xis y Int.

6. La célula anfitriona bacteriana de acuerdo con una cualquiera de las reivindicaciones 1 a 5, en la que la célula anfitriona bacteriana es una enterobacteria, preferiblemente E. coli, más preferiblemente BL21.

7. La célula anfitriona bacteriana de acuerdo con una cualquiera de las reivindicaciones 1 a 6, comprende adicionalmente la inactivación de por lo menos uno de los genes tonA, galK, araB, araA, lon, ompT, rcsA, hsdR, mrr, endA y recA.

8. La célula anfitriona bacteriana de acuerdo con una cualquiera de las reivindicaciones 1 a 7, que comprende la inserción del gen ccdB.

9. Un kit que comprende la célula anfitriona bacteriana de acuerdo con la reivindicación 8 y un plásmido que comprende el gen ccdA.

10. Un proceso para preparar la célula anfitriona bacteriana de acuerdo con una cualquiera de las reivindicaciones 1 a 8, que comprende infectar una célula anfitriona bacteriana con un fago genéticamente modificado en el que

- se inserta un sistema de expresión, dicho sistema de expresión comprende una secuencia de ácidos nucleicos que codifica una proteína, dicha proteína que induce adicionalmente la expresión de una biomolécula de interés,
- se suprimen los genes int, S, R, Rz, Xis, y
- el fago es el fago lambda, el fago 434, el fago phi80, el fago phi81, el fago HK97, o el fago P21, y

utilizar un fago ayudante para complementación.

11. Un proceso para producir una biomolécula de interés, que comprende

- cultivar una célula anfitriona bacteriana de acuerdo con una cualquiera de las reivindicaciones 1 a 8, en la que dicha célula anfitriona comprende la secuencia de ácidos nucleicos de la biomolécula de interés,
- expresar dicha biomolécula de interés.

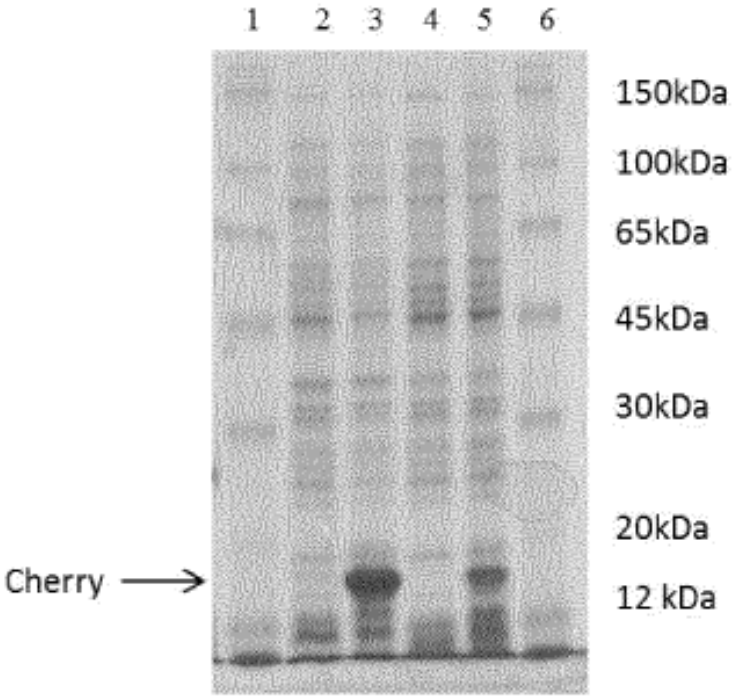

FIG. 1