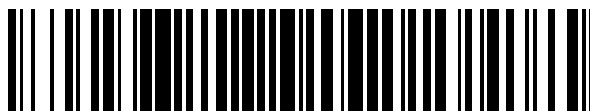


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 389 129**

51 Int. Cl.:
C12Q 1/68 (2006.01)
C12N 1/21 (2006.01)
C12N 15/10 (2006.01)
C12N 15/63 (2006.01)
C12N 15/66 (2006.01)
C40B 40/06 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **08806498 .5**
96 Fecha de presentación: **03.10.2008**
97 Número de publicación de la solicitud: **2205761**
97 Fecha de publicación de la solicitud: **14.07.2010**

54 Título: **Señuelos de factores de transcripción, composiciones y métodos**

30 Prioridad:
03.10.2007 GB 0719367

45 Fecha de publicación de la mención BOPI:
23.10.2012

45 Fecha de la publicación del folleto de la patente:
23.10.2012

73 Titular/es:
**PROCARTA BIOSYSTEMS LIMITED (100.0%)
NORWICH BIOINCUBATOR NORWICH
RESEARCH PARK
NORWICH NR4 7UH, GB**

72 Inventor/es:
MCARTHUR, MICHAEL

74 Agente/Representante:
DE ELZABURU MÁRQUEZ, Alberto

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).

ES 2 389 129 T3

DESCRIPCIÓN

Señuelos de factores de transcripción, composiciones y métodos

Campo de la invención

5 La invención proporciona métodos y composiciones para identificar secuencias reguladoras en *cis*. La invención también se refiere a métodos y composiciones para modular específicamente fenotipos celulares, tales como fenotipos procariotas, mediante la alteración de la asociación de factores de transcripción a elementos reguladores en *cis in vivo* con una alteración simultánea en el patrón de la expresión génica.

Antecedentes de la invención

10 La expresión génica es un determinante principal del fenotipo de una célula y, a su vez, las interacciones ADN-proteína determinan en gran medida los patrones de la expresión génica. La modificación de la expresión génica para influir en el fenotipo es un objetivo principal de la biología y la medicina, ya sea en microbios industriales, modelos experimentales, patógenos o para abordar una enfermedad humana. En este contexto, las secuencias reguladoras en *cis* en el genoma, el componente de ADN de la maquinaria transcripcional, son dianas atractivas para intervenir. En comparación con abordar proteínas, el trabajo con el ADN es mucho más fácil: la secuenciación
15 está muy automatizada y es relativamente económica, y el ADN puede manipularse fácilmente y amplificarse o sintetizarse fácilmente. Las terapias basadas en ADN también han surgido como una nueva clase excitante de agentes terapéuticos. En comparación con los compuestos farmacéuticos tradicionales, productos o moléculas pequeñas naturales, el ADN es un tipo atractivo de agente terapéutico ya que puede:

- *diseñarse por un proceso racional*, lo más sencillamente examinando datos de secuencia;
- 20 • *fabricarse de manera económica a escala*, por síntesis química de oligonucleótidos o replicación biológica;
- *predecirse que tiene una toxicidad baja*, ya que el ADN es un compuesto "natural" y no induce, en y por sí mismo, típicamente respuestas inmunogénicas, y la especificidad puede controlarse por secuencia de la terapia basada en ADN;
- 25 • *reducir en gran medida el gasto de I+D*, ya que todas las etapas del desarrollo de fármacos convencionales (identificación de diana, descubrimiento de compuesto líder, química médica) están truncadas.

Norris *et al* (Gene Therapy (2000) 7, p. 723-725 discuten el uso de un sistema de administración de un agente letal del bacteriófago PI para el tratamiento de tratamiento bacteriano resistente a múltiples fármacos.

30 El reto se convierte en cómo identificar los elementos reguladores en *cis* clave. Las tecnologías y experiencia que se han desarrollado en paralelo con los proyectos de secuenciación del genoma, tales como análisis de expresión génica paralela masiva usando micromatrices de ADN y el uso de bioinformática para registrar las bases de datos de los genomas, no son suficientes para identificar todos los elementos reguladores en *cis* ni para atribuir una función a los que ya se conocen. En 2003, el National Institute of Health en los EEUU lanzó el proyecto ENCODE para catalogar el 1% de los elementos reguladores en *cis* en el genoma humano (Science (2004) 306: 636-640; Nature (2007) 447: 799-816) y para desarrollar tecnologías de alto rendimiento como plataformas de descubrimiento. Los
35 procedimientos desarrollados incluyeron el uso de inmunoprecipitación de cromatina, el sondeo de sensibilidad hipersensibilidad a digestión *in vivo* con ADNaseI (con micromatrices de ADN, PCR cuantitativa de alto rendimiento y bibliotecas genómicas) y el desarrollo de nuevos algoritmos para la detección bioinformática. Aunque estas técnicas tienen el potencial de acelerar mucho la velocidad de descubrimiento de elementos reguladores en *cis*, no darán lugar necesariamente a su caracterización funcional: el resultado del proyecto es un catálogo exhaustivo de estos
40 elementos. Aun así, a partir de dicho trabajo, es probable que la especificidad de tejido de la mayoría de los elementos se conozca y su distancia de un gen y clasificación según qué tipo de factor que actúa en *trans* se une a ellos, esto no será suficiente para determinar cuál es la función biológica real del elemento. Una desventaja adicional común a todos estos procedimientos es que se basan en el genoma del organismo que se está secuenciando. Además, la estrategia usando hipersensibilidad a la digestión con ADNaseI tiene la desventaja adicional de que es
45 específica para las células eucariotas, ya que, en este contexto, la ADNaseI es una sonda de la estructura de la cromatina y no existe una estructura comparable en los procariotas.

Además, hasta ahora no hay medios para cribar rápidamente un gran número de secuencias para secuencias reguladoras en *cis* potenciales.

Resumen de la invención

50 Los inventores abordaron estos problemas en la técnica proporcionando nuevos métodos para identificar y caracterizar secuencias reguladoras en *cis* tanto en organismos procariotas como eucariotas.

Una manera de identificar y caracterizar la función de elementos reguladores en *cis* individuales es usar señuelos de factores de transcripción (TFD). Los oligonucleótidos señuelo se diseñan para mimetizar los sitios de unión de los

factores de transcripción y evitar que estos últimos se unan a sus dianas genómicas afines, con una modificación consecuente de la expresión génica. Como tales, representan una herramienta sencilla y genérica para manipular las interacciones ADN-proteína que regulan genes específicos y que consecuentemente determinan los fenotipos. Su utilidad se ha demostrado principalmente en sistemas eucariotas, en los que un estímulo para su desarrollo ha sido su potencial para funcionar como nuevas clases de agentes terapéuticos (Mann y Dzau (2000) *J. Clin. Investigation* **106**: 1071-1075). Para este objetivo, se han usado oligonucleótidos señuelo para demostrar que el factor de transcripción EF2 reprime la proliferación de músculo liso en ratas (Morishita *et al.* (1995) *Proc. Natl. Acad. Sci. USA* **95**: 5855-5859); para bloquear la proliferación mediada por STAT3 de carcinomas (Leong *et al.* (2003) *Proc. Natl. Acad. Sci. USA* **100**: 4138-4143); y para mostrar que el direccionamiento del elemento de respuesta AMPc puede controlar la proliferación del cáncer *in vivo* (Park *et al.* (1999) *J. Biol. Chem.* **274**: 1573-1580).

Sin embargo, no se ha desarrollado ningún sistema para usar TFD, a gran escala, que sea capaz de cuestionar cada aparición de una secuencia en un fragmento genómico o genoma completo. Tal y como está actualmente, esto requeriría la síntesis de grandes números de oligonucleótidos señuelo y un programa experimental costoso que implica su transfección y cribado para cambio fenotípico.

La invención descrita aquí aborda esta necesidad mediante el desarrollo de un sistema basado en una biblioteca portada en plásmidos capaz de ensayar sistemáticamente grandes números de secuencias para determinar si actúan en el genoma como reguladores en *cis* y asociarlas con un efecto fenotípico específico. Nos referimos a este sistema en la presente memoria como "n[snare]".

Además, una vez que se han encontrado las secuencias reguladoras en *cis* relevantes, estas secuencias pueden usarse para crear TFD que pueden usarse para modificar la expresión génica y obtener control sobre el fenotipo. Por varias razones, aunque los señuelos se desarrollaron para uso en células de mamíferos, son mucho más adecuados para uso en bacterias. Conseguir que los señuelos funcionen en eucariotas puede ser problemático ya que pueden degradarse rápidamente en el suero y extractos nucleares (Chu y Orgel (1992) *Nucl. Acids. Res.* **20**: 5857-5858), la captación celular del señuelo y su transición a través de la membrana nuclear puede ser ineficaz (Griesenbach *et al.* (2002) *Gene Therapy* **9**: 1109-1115) y algunos tratamientos pueden desencadenar efectos no específicos o tóxicos. El uso de señuelos en procariotas debería evitar muchos de estos problemas y como tales podría demostrarse que son una herramienta eficaz para la identificación rápida de secuencias reguladoras que actúan en *cis*, tales como sitios de unión de factores de transcripción que controlan tanto genes específicos como redes reguladoras mayores. Una demostración exitosa de la estrategia fue el uso de un señuelo rico en AT para alterar la expresión de genes que responden a CO₂ en *Cyanobacterium* (Onizuka *et al.* (2003) *FEBS Lett.* **542**: 42-46). En este sistema, los oligonucleótidos complementarios con núcleos modificados (que contienen fósforotioato para ralentizar la degradación por nucleasas) se hibridaron para formar oligonucleótidos señuelo bicatenarios que incorporaban sitios de unión para un factor de transcripción previamente identificados. Éstos se añadieron directamente al medio a partir del cual entraron eficazmente en las células. Esta publicación, sin embargo, es el único ejemplo que conocemos del intento exitoso de esta estrategia para modificar un rasgo procariota particular. De acuerdo con esto, permanece una necesidad en la técnica para extender la metodología de los señuelos en el campo de los factores de transcripción procariotas para alterar de esta manera un amplio rango de fenotipos procariotas. Además, permanece una necesidad en la técnica para métodos de alto rendimiento para identificar factores reguladores que actúan en *cis*, ya sea en sistemas procariotas como eucariotas. Esta descripción de patente proporciona soluciones a las limitaciones y necesidades indicadas anteriormente en la técnica.

En un aspecto particular, la invención proporciona nuevos medios para incrementar la susceptibilidad de las células, por ejemplo, células bacterianas, a antibióticos.

De acuerdo con esto, en un aspecto, la invención proporciona el uso de un polinucleótido señuelo en un método para modular la resistencia a antibióticos de una célula, comprendiendo el método:

(a) proporcionar un polinucleótido señuelo que comprende un sitio de unión para un factor de transcripción (una secuencia señuelo);

e

(b) introducir el polinucleótido en la célula, en el que la célula comprende un gen o genes unidos operativamente a una secuencia reguladora en *cis* que comprende un sitio de unión para el factor de transcripción;

en el que la introducción del polinucleótido reduce la unión del factor de transcripción a la secuencia reguladora en *cis* en la célula y causa una alteración en la expresión del gen o genes unidos operativamente en la célula, modulando de esta manera la resistencia a antibióticos de la célula.

La invención también proporciona:

- un polinucleótido señuelo que comprende un sitio de unión para un factor de transcripción, en el que el factor de transcripción es un regulador de la expresión de uno o más genes de resistencia a antibióticos en un procariota o eucariota y en el que el sitio de unión en el polinucleótido señuelo no está unido operativamente a un gen.

- un plásmido que comprende una o más copias de una secuencia monomérica, en el que la secuencia monomérica comprende una secuencia snare que comprende un sitio de unión de un factor de transcripción y en el que el sitio de unión no está unido operativamente a un gen.
- 5 - una biblioteca de plásmidos que comprende dos o más plásmidos de la invención en el que las secuencias snare en la biblioteca conjuntamente comprenden todas o sustancialmente todas las secuencias reguladoras en *cis* en el ADN genómico o en un fragmento del ADN genómico de un procariota o eucariota;
- una biblioteca de plásmidos que comprende dos o más plásmidos de la invención en el que la secuencia snare en cada plásmido comprende una secuencia de nucleótidos aleatorizada con una longitud de *n* nucleótidos, en el que sustancialmente todas las secuencias de nucleótidos con una longitud de *n* nucleótidos están representadas en la biblioteca;
- 10 - un método para preparar un plásmido que comprende dos o más copias de una secuencia monomérica, comprendiendo el método:
 - (1) proporcionar un oligonucleótido circular que comprende:
 - (i) una secuencia de ensayo de interés; y
 - 15 (ii) un sitio de unión para un cebador adecuado para uso en amplificación por círculo rodante; en el que la secuencia monomérica comprende (i) y (ii);
 - (2) realizar la amplificación por círculo rodante usando el oligonucleótido circular como molde, proporcionando de esta manera un polinucleótido que comprende repeticiones de la secuencia monomérica; y
 - (3) clonar el polinucleótido en un vector plasmídico;
- 20 - un método para preparar una biblioteca de plásmidos a partir de una muestra de ADN genómico en el que cada plásmido comprende una o más copias de una secuencia monomérica y en el que cada secuencia monomérica comprende una secuencia snare que se obtiene del ADN genómico; comprendiendo el método:
 - (1) proporcionar una muestra de ADN genómico bicatenario;
 - (2) fragmentar el ADN genómico;
 - 25 (3) ligar un adaptador a cada extremo de los fragmentos de ADN de (2), en el que cada adaptador comprende:
 - (iii) un medio para la inmovilización del fragmento de ADN;
 - (iv) un sitio de reconocimiento para una primera enzima de restricción que corta a una distancia del sitio de reconocimiento; y
 - (v) un sitio de reconocimiento y corte para una segunda enzima de restricción;
 - 30 (4) opcionalmente eliminar el adaptador no ligado;
 - (5) digerir el fragmento que presenta los adaptadores con la primera enzima de restricción, produciendo de esta manera dos fragmentos con adaptador, en el que cada fragmento con adaptador comprende:
 - (vi) un adaptador; y
 - (vii) un fragmento de ADN que comprende: una cadena más corta; y una cadena más larga, en el que la cadena más larga comprende la secuencia snare;
 - 35 (6) inmovilizar los fragmentos con adaptador producidos en (5);
 - (7) desnaturalizar los fragmentos para proporcionar fragmentos monocatenarios;
 - (8) recrear el sitio de reconocimiento para la segunda enzima de restricción ligando un oligonucleótido complementario al adaptador y digerir el fragmento con adaptador con la segunda enzima de restricción, produciendo de esta manera un fragmento adaptador-snare que comprende:
 - 40 (viii) un fragmento adaptador; y
 - (ix) una secuencia snare monocatenaria; y
 - (9) liberar el fragmento adaptador-snare producido en (8) de la inmovilización; y
 - (10) clonar el fragmento adaptador-snare en un vector plasmídico;

- un método para preparar una biblioteca de plásmidos en el que cada plásmido comprende dos o más copias de una secuencia monomérica, y en el que cada secuencia monomérica comprende una secuencia de nucleótidos aleatorizada con una longitud de n nucleótidos, comprendiendo el método:

(1) proporcionar un oligonucleótido circular que comprende:

5 (i) una secuencia aleatorizada con una longitud de n nucleótidos; y

(ii) un sitio de unión para un cebador adecuado para uso en amplificación por círculo rodante;

en el que la secuencia monomérica comprende (i) y (ii);

(2) realizar la amplificación por círculo rodante usando el oligonucleótido circular como molde, proporcionando de esta manera un polinucleótido que comprende repeticiones de la secuencia monomérica; y

10 (3) clonar el polinucleótido en un vector plasmídico;

- un plásmido o biblioteca de plásmidos preparada según un método de la invención;

15 - una célula que comprende un polinucleótido señuelo exógeno, comprendiendo el polinucleótido un sitio de unión para un factor de transcripción (una secuencia señuelo) que no está unido operativamente a un gen; en el que la célula comprende un gen o genes unidos operativamente a una secuencia reguladora en *cis* que comprende un sitio de unión para el factor de transcripción; y en el que el polinucleótido señuelo causa una alteración en la resistencia a antibióticos de la célula.

- una célula o células huésped que comprenden un plásmido o biblioteca de plásmidos de la invención;

- un método para identificar los límites de uno o más sitio(s) de unión de proteínas en un complejo proteína-ADN, comprendiendo el método:

20 1. proporcionar un complejo proteína-ADN;

2. realizar una digestión con

a. una enzima que tiene una capacidad de mellado del ADN no específica; y

b. una exonucleasa 5'-3'; y

25 3. determinar la posición de las deleciones 5' generadas en cada cadena de ADN en (2) respecto a un punto fijo conocido en la cadena de ADN;

- un método para identificar un regulador que actúa en *cis* de la expresión génica de un gen procariota o eucariota que comprende:

1. proporcionar una biblioteca de plásmidos de la invención;

30 2. introducir la biblioteca de plásmidos en una o unas células huésped en el que la célula huésped comprende la secuencia reguladora en *cis* de interés unida operativamente a un gen o genes, cuya expresión puede determinarse directamente o indirectamente; y

3. determinar directamente o indirectamente la expresión del gen (o genes) en presencia y ausencia de la biblioteca de plásmidos;

- un método para alterar la expresión de un gen o genes en una célula procariota, comprendiendo el método:

35 (a) proporcionar un polinucleótido que comprende un sitio de unión para un factor de transcripción procariota en el que el sitio de unión no está unido operativamente a un gen en el polinucleótido; e

(b) introducir el polinucleótido en la célula;

40 en el que la célula comprende el gen o genes unidos operativamente a una secuencia reguladora en *cis* que comprende el sitio de unión de transcripción o que compete con el sitio de unión del factor de transcripción para la unión del factor de transcripción;

y en el que:

(i) el polinucleótido comprende un plásmido o una biblioteca de plásmidos según la invención;

(ii) el polinucleótido comprende más de una copia del sitio de unión;

(iii) el polinucleótido comprende múltiples repeticiones directas del sitio de unión;

(iv) el polinucleótido comprende secuencia adicional al sitio de unión;

(v) el polinucleótido comprende al menos un elemento de estructura secundaria;

(vi) el polinucleótido comprende ADN bicatenario circular;

5 - un método para alterar la expresión de un gen o genes en una célula procariota o eucariota, comprendiendo el método:

(a) proporcionar un polinucleótido que comprende un sitio de unión para un factor de transcripción procariota o eucariota en el que el sitio de unión no está unido operativamente a un gen en el polinucleótido; e

(b) introducir el polinucleótido en la célula;

10 en el que la célula comprende el gen o genes unidos operativamente a una secuencia reguladora en *cis* que comprende el sitio de unión de transcripción o que compite con el sitio de unión del factor de transcripción para la unión del factor de transcripción;

y en el que la secuencia reguladora en *cis* es una identificada por el método de la invención;

- un método para alterar la expresión de un gen o genes en una célula eucariota, comprendiendo el método:

15 (a) proporcionar un polinucleótido que comprende un sitio de unión para un factor de transcripción eucariota en el que el sitio de unión no está unido operativamente a un gen en el polinucleótido; e

(b) introducir el polinucleótido en la célula;

en el que la célula comprende el gen o genes unidos operativamente a una secuencia reguladora en *cis* que comprende el sitio de unión de transcripción o que compite con el sitio de unión del factor de transcripción para la unión del factor de transcripción;

20 y en el que el polinucleótido comprende un plásmido o biblioteca de plásmidos según la invención;

- una célula preparada según un método de la invención;

- un polinucleótido señuelo para uso en el tratamiento de una infección bacteriana, en el que el polinucleótido comprende un sitio de unión para un factor de transcripción y el sitio de unión no está unido operativamente a un gen;

25 - uso de un polinucleótido señuelo para tratar una infección bacteriana, en el que el polinucleótido comprende un sitio de unión para un factor de transcripción y el sitio de unión no está unido operativamente a un gen;

- uso de un polinucleótido señuelo para la fabricación de un medicamento para tratar una infección bacteriana, en el que el polinucleótido comprende un sitio de unión para un factor de transcripción y el sitio de unión no está unido operativamente a un gen.

30 Así, en una realización de la invención de esta descripción de patente, proporcionamos un método que combina la estrategia de señuelo con un protocolo sencillo de obtención de impronta *in vivo* para identificar rápidamente restos reguladores que actúan en *cis* candidatos. La validación funcional de estas secuencias se consiguió incorporándolas en oligonucleótidos señuelo en forma de mancuerna, cuyo formato circular suprime la degradación por *exo* y endonucleasas [Ahn *et al.* (2003) *Biochemical Biophysical Res. Comm.* **310**: 1048-1053] y ensayando su efecto en el fenotipo *in vivo*.

35 En otra realización según esta invención, demostramos que, usando oligonucleótidos señuelo identificados mediante la aplicación del método de obtención de impronta, podemos alterar significativamente el nivel de producción de antibiótico en streptomycetes.

40 En una realización adicional según esta invención, demostramos la capacidad de utilizar señuelos procariotas en una estrategia terapéutica por la que bacterias patógenas que son resistentes a vancomicina se vuelven, una vez más, susceptibles al antibiótico en la ventana terapéutica.

45 Así, los inventores proporcionan un sistema de cribado genérico mediante el cual los elementos reguladores que actúan en *cis* se identifican en cualquier genoma, y las secuencias obtenidas de dichos elementos identificados de esta manera se utilizan para alterar fenotipos seleccionados de una manera análoga a la demostrada según la ilustrada por realizaciones adicionales de la invención.

Los inventores también han adaptado y extendido la técnica del oligonucleótido señuelo para uso en procariotas y demuestran la capacidad de incrementar la producción de antibióticos.

Además, los inventores han demostrado que los oligodesoxinucleótidos señuelo pueden utilizarse para superar ventajosamente la resistencia a antibióticos para tratar patógenos.

Otras realizaciones, utilidades y detalles de esta invención pueden apreciarse por una revisión de la descripción completa, incluyendo los ejemplos específicos proporcionados en la presente memoria y las reivindicaciones adjuntas a esta descripción, incluyendo los equivalentes de éstas.

Descripción breve de las figuras

Figura 1. Construcción y ensayo de plásmidos n[snare]. Este protocolo también puede usarse para producir bibliotecas n[snare] “genéricas”, preparadas a partir de oligonucleótidos aleatorios o bibliotecas específicas de especie, a partir de genomas fragmentados.

Figura 2. Proporciona una representación gráfica de cómo los plásmidos n[snare] pueden usarse para influir en la expresión génica. El plásmido n[snare] se introduce en la célula diana (por ejemplo, procariota), por medios estándar, y su propagación estable se asegura seleccionando su marcador, habitualmente un gen que codifica resistencia a un antibiótico. Cuando se introduce así en una célula, el plásmido n[snare] es capaz de influir en la expresión del gen diana liberando el factor de transcripción “B” del promotor genómico para liberar la represión transcripcional del gen en 3' (mostrado como caja horizontal “D”, parte superior derecha).

Figura 3. Demostración de la estrategia n[snare]. Los plásmidos n[snare] son capaces de modificar el control de la expresión génica de una manera predecible con un cambio simultáneo en el fenotipo. Estos plásmidos son más baratos de producir y más fáciles de transformar en las células que los señuelos, proporcionan un efecto más sostenido y crucialmente, permiten una estrategia de biblioteca para descubrir elementos reguladores clave.

Figura 4. Parte del proceso para crear bibliotecas de plásmidos n[snare] personalizadas: Convertir fragmentos genómicos en nucleótidos monocatenarios cortos. Las bibliotecas de plásmidos n[snare] se crean a partir de grandes partes de ADN de manera tal que cada secuencia reguladora en *cis* en este ADN debería estar representada en la biblioteca.

Figura 5. Las bibliotecas de plásmidos n[snare] pueden usarse para obtener control de los fenotipos diana, tal como la producción de antibiótico. Después de la introducción de los miembros de la biblioteca por conjugación, se seleccionan clones tomando como base la producción incrementada de cinamicina (derecha).

Figura 6. Detalles adicionales para la aplicación de n[snare] para combatir la resistencia a antibiótico. El esquema en esta figura muestra cómo una biblioteca de plásmidos n[snare] se usa para identificar secuencias reguladoras en *cis* que controlan la resistencia a antibióticos cuando no se conoce nada acerca de la ruta genética del mecanismo.

Figura 7. Construcciones informadoras desarrolladas para identificar elementos reguladores en *cis* en una biblioteca de plásmidos n[snare]. En esta figura, el promotor diana del gen *cin7* de la agrupación biosintética de la cinamicina de la cepa productora de *S. cinnamoneous* se usa para dirigir la expresión del gen *neo* que codifica resistencia al antibiótico kanamicina.

Figura 8. Detección de los reguladores en *cis* negativos de la producción de cinamicina usando un sistema informador genérico y una biblioteca n[snare] personalizada. Una adaptación sencilla hace que el sistema informador sea capaz de detectar la regulación positiva creando un casete que consiste en un gen informador químico que incorpora la secuencia codificadora del gen de la glucosa quinasa (*glkA*) dirigido por el promotor diana.

Figura 9. Hibridación de la biblioteca para identificar clones candidatos. Para distinguir miembros “de fondo” que forman una parte considerable de la señal total de los plásmidos n[snare] capaces genuinamente de interferir con la transcripción, se desarrolló una estrategia de hibridación de la biblioteca usando métodos estándar.

Figura 10. Cribados de hibridación para detectar candidatos a partir de repeticiones independientes del procedimiento para detectar reguladores en *cis* capaces de regular la producción de cinamicina usando un sistema informador genérico y una biblioteca n[snare] personalizada. En este ejemplo, se ensayaron dos bibliotecas independientes de *S. cinnamoneous* con cepas informadoras y 96 de los clones más resistentes a kanamicina de cada una se usó para hacer un filtro que contiene los plásmidos n[snare] para cada clon (F1 y F2) y dos conjuntos de sondas (P1 y P2). Todos los clones identificados por hibridación cruzada (F1 frente a P2 y F2 frente a P1) se consideraron candidatos fuertes y se priorizaron aquellos comunes a ambos experimentos de hibridación cruzada (con círculo en rojo).

Figura 11. El análisis de digestión por restricción de la biblioteca n[snare] genómica de *Escherichia coli* K12 confirma que tiene las propiedades de secuencia esperadas. Esta mezcla de plásmidos se digirió con *EcoRI* para regenerar los fragmentos usados para crear las bibliotecas n[snare] y éstas se sometieron a digestión con *MmeI* para confirmar que había funcionado la biología molecular de la construcción: como se esperaba, los insertos tenían un gran tamaño y éstos se convirtieron en un monómero de 30 pb después de la digestión completa con *MmeI*.

Figura 12. Las bibliotecas n[snare] se usan para obtener por ingeniería fenotipos sin conocimiento previo de la red genética. Este proceso se repitió cuatro veces y la viabilidad celular se midió cada vez.

Figura 13. Vista esquemática global del protocolo para el procedimiento de mapeo *in vivo* con exonucleasa T7/ADNasal. La nueva combinación de exonucleasa T7 y ADNasal en un protocolo de obtención de impronta permite la detección de todos los límites de los complejos ADN-proteína en una región promotora, y no sólo los más cercanos a un sitio de restricción elegido.

Figura 14. El mapeo con exonucleasa T7/ADNasal de restos reguladores candidatos en el promotor de *actII-orf4*. (A) Las células de *S. coelicolor* M145 se recogieron a partir de un cultivo crecido en medio rico (R5) en un punto de tiempo (indicado por una flecha) que precede la producción visible de actinorhodina. (B) Los límites se mapearon en ambas cadenas, como se describe en a Figura 13, y sus posiciones se determinaron después del análisis de tamaño por PAGE al 12% no desnaturalizante, seguida de detección quimioluminiscente de los productos marcados con DIG. (C) La secuencia del promotor *actII-orf4* que muestra las posiciones de los elementos reguladores en *cis* posibles (respecto a los cebadores usados en el protocolo de mapeo).

Figura 15. Ensayos en placa que demuestran que los oligonucleótidos señuelo pueden influir en la producción de antibióticos en *S. coelicolor*. Se saturaron discos de filtro con disoluciones de señuelos o, como control, tampón (como se muestra) y se aplicaron a una capa de *S. coelicolor* M145 cubierta con medio SNA.

Figura 16. Captación y estabilidad del oligonucleótido señuelo. Se transfectaron células que crecen activamente con una disolución que contiene señuelo y se estimaron la captación del oligonucleótido y su estabilidad por PCR cuantitativa en tiempo real (crt-PCR).

Figura 17. Incremento mediado por señuelo en la producción de actinorhodina en cultivos crecidos en medio líquido R5. *S. coelicolor* M145 se creció durante 20 horas antes de la transfección (indicado con flechas) con (A) un control sin señuelo o con (B) el señuelo A24.5.

Figura 18. Incremento mediado por señuelo en la producción de actinorhodina en cultivos crecidos en medio líquido SMM. La comparación de los datos obtenidos con (A) un control transfectado simulado y (B) un cultivo tratado con señuelo A24.5 revelaron que el oligonucleótido señuelo causaba un incremento pronunciado en la producción de actinorhodina.

Figura 19. La delección de SCO5812 da lugar a una producción disminuida de actinorhodina en agar R5 y producción aumentada de undecilprodigiosina en agar SMMS. M145 (lado izquierdo de las placas) y M145 Δ SCO5812 (lado derecho de las placas) se sembraron en estrías en (A) medio agar R5 o (B) medio agar SMMS y se incubaron durante 72 h y 96 h, respectivamente.

Figura 20 proporciona la aplicación de la estrategia del señuelo del factor de transcripción (TFD) para combatir la resistencia a antibióticos. El esquema demuestra cómo se usan los TFD para contrarrestar los mecanismos de resistencia conocidos en bacterias patógenas. La resistencia al antibiótico prescrito vancomicina se usa como un ejemplo.

Figura 21. La estructura del gen de vancomicina. De Hong et al. 2004. *Molecular Microbiology* 52: 1107-1121.

Figura 22. Evidencia de ciclación del señuelo de oligonucleótido vanH5. El oligonucleótido se resuspendió a una concentración final de 100 pmoles/ul en un tampón de ADN ligasa T4 (como suministra el fabricante de la enzima, New England Biolabs) y 400 U de ADN ligasa T4 y se incubó a 16 grados Centígrados durante varios tiempos.

Figura 23. Muestra el efecto de incubar una bacteria resistente a vancomicina con o sin niveles crecientes del señuelo vanH5. La cepa M600 de *S. coelicolor* se creció en medio líquido MMCGT (*Molecular Microbiology* 52: 1107-1121), el crecimiento se midió registrando la absorbancia del cultivo a 430 nm (Densidad Celular) y se representó como una función del tiempo de incubación.

Figura 24. Muestra el producto de amplificación obtenido usando los cebadores en SEQ ID NOS: 24 y 25 y un sustrato vector apropiado (Ejemplo 7.2).

Figuras 25 y 26. Muestran las curvas de crecimiento para *E. faecium* crecido en presencia de vancomicina después de tratamiento con la secuencia señuelo del factor de transcripción VAN o un control negativo (CON), como se describe en los Ejemplos 8(a) y 8(b), respectivamente.

Figura 27. Una copia de las Tablas 1 y 2 de Poole (2005) J. Antimicrobial Chemother. 56, 22-24 que listan la resistencia mediada por eflujo a antibióticos no fluorquinolona y antibióticos fluorquinolona, respectivamente. Los números de referencia en la columna de la derecha se refieren a los proporcionados en el artículo.

Descripción breve de las secuencias

SEQ ID NO: 1 – un oligonucleótido señuelo que contiene el sitio de unión AfsR (Ejemplo 1).

SEQ ID NO: 2 – un polinucleótido de unión R-T7, que contiene un complemento parcial de un cebador usado comúnmente, T7 (Ejemplo 1).

SEQ ID NO: 3 – un oligonucleótido que contiene el sitio de unión AfsR y uso para crear una secuencia señuelo ciclada en el Ejemplo 1.

5 SEQ ID NOS: 4 y 5 – secuencias de cada cadena de una molécula adaptadora bicatenaria (Ejemplo 2) SEQ ID NO: 6 – un oligonucleótido complementario Bbv (Ejemplo 2).

SEQ ID NO: 7 – un oligonucleótido que comprende un sitio *NotI* y una secuencia de nucleótido aleatoria, en el que cada nucleótido aleatorio se representa como “n” (Ejemplo 4).

10 SEQ ID NOS: 8 – 12 – oligonucleótidos señuelo diseñados tomando como base las secuencias reguladoras en *cis* A24.1, A24.2, A24.3, A24.4 y A24.5, respectivamente (Ejemplo 5).

SEQ ID NO: 13 – un oligonucleótido señuelo diseñado tomando como base una secuencia A24.5 desordenada (Ejemplo 5).

SEQ ID NOS: 14 – 20 – cebadores de PCR usados en PCR cuantitativa como en el Ejemplo 5.

15 SEQ ID NO: 21 – un oligonucleótido señuelo vanH5 que contiene un sitio de unión para VanR fosforilado (Ejemplo 6).

SEQ ID NOS: 22 y 23 – cebadores oligonucleotídicos usados para la amplificación de una secuencia diana del vector pGEMT-Easy como en el Ejemplo 7.1.

SEQ ID NOS: 24 y 25 – cebadores oligonucleotídicos usados para la amplificación de una secuencia diana del vector pGEMT-Easy en la producción de señuelos en forma de mancuerna como en el Ejemplo 7.2.

20 SEQ ID NO: 26 - la secuencia señuelo VAN usada en los Ejemplos 8(a) y 8(b) y que contiene el elemento regulador que controla la inducción de la resistencia de tipo VanA en *E. faecium*.

Descripción detallada de la invención

25 Como se describe con más detalle en la presente memoria, los presentes inventores han ideado métodos y composiciones para identificar, caracterizar y tomar como diana secuencias reguladoras en *cis* en procariotas y eucariotas.

30 Una secuencia o elemento regulador en *cis* se refiere generalmente a una secuencia de nucleótidos que ocurre en 5' (5') o en 3' (3') de un gen o genes y que funciona para modular la expresión del gen o genes. Típicamente, una secuencia reguladora en *cis* comprende un sitio de unión para una proteína (factor de transcripción) que regula la transcripción del gen o genes dados. La unión de la proteína a la secuencia resulta directamente o indirectamente en la modulación de la expresión del gen o genes. Por ejemplo, la proteína unida puede interactuar con otra proteína unida a una región cercana que se necesita para la transcripción y el anclaje de la proteína en la posición correcta, o puede inhibir la unión de otra proteína que es necesaria para la transcripción. Típicamente, la secuencia o elemento regulador en *cis* ocurre en la región promotora de un gen, pero no es raro en procariotas que las secuencias reguladoras en *cis* estén situadas cientos de pares de bases en 5' o en 3' de los genes en los que influyen. En eucariotas, las secuencias reguladoras en *cis* pueden actuar a grandes distancias para influir en la expresión de un gen, típicamente del orden de 1-2 kb, pero se sabe que hay secuencias que actúan más de 100 kb a 1 Mb.

35 Una secuencia reguladora en *cis* puede ser represora (inhibe o reduce la transcripción del gen o genes cuando se une un factor de transcripción) o activadora (activa o incrementa la transcripción del gen o genes cuando se une un factor de transcripción). Así, un factor de transcripción que se une a una secuencia reguladora en *cis* puede ser un efector negativo (proteína represora) o un efector positivo (activador).

40 La expresión de determinados genes también se modificará por efectos indirectos cuando, por ejemplo, el impacto de una secuencia reguladora en *cis* está en un gen separado, que, a su vez, incluye en la expresión del gen diana. Esto puede ocurrir, por ejemplo, causando cambios en una red reguladora de que la forma parte el gen diana, o debido a un efecto más global, tal como una respuesta a choque, que causa modificación de la expresión del gen diana.

45 La modulación de la unión de proteínas a secuencias reguladas en *cis* puede proporcionar un medio útil para modular la expresión génica en una célula. Una manera de hacer esto es proporcionar una secuencia de nucleótidos señuelo. La secuencia señuelo comprende un sitio de unión de un factor de transcripción que comprende o compite con la secuencia reguladora en *cis* nativa o endógena en la célula para unión al factor de transcripción afín. Mediante la reducción de la unión del factor de transcripción a la secuencia nativa, el señuelo altera la expresión del gen o genes cuya expresión está regulada normalmente por la secuencia reguladora en *cis* en la célula. Dicha alteración puede proporcionar una alteración útil del fenotipo, por ejemplo, en la producción de metabolitos o resistencia a antibióticos de una célula procariota. La función del señuelo tal y como se usa en la presente memoria

se refiere a la capacidad de una secuencia para competir con una secuencia reguladora en *cis* para unión a un factor de transcripción afín de esta manera.

Los inventores han desarrollado nuevos medios para identificar secuencias que compiten con sitios de unión de factores de transcripción celulares para unión del factor de transcripción, para alterar la expresión génica celular y que pueden proporcionar, por lo tanto, nuevas secuencias reguladoras en *cis* y secuencias señuelo.

plásmidos n[snare]

La estrategia descrita en la presente memoria es hacer adaptaciones importantes de la aplicación conocida en eucariotas de la tecnología de oligonucleótidos "señuelo", que se ha usado para modular la expresión *in vivo*. La lógica de los oligonucleótidos señuelo es sencilla y relevante para todos los sistemas biológicos: reafirmar el control genético alterando la unión de los factores de transcripción a sus sitios afines. Hemos mejorado esta tecnología proporcionando una herramienta genérica y de alto rendimiento para el análisis de la regulación genética. Una primera modificación implica hacer que la estrategia sea portada en plásmido desarrollando un método para clonar homopolímeros de secuencias señuelo, para crear plásmidos "n[snare]" con muchas copias. Éstos se introducen fácilmente en todas las bacterias y se mantienen por selección positiva, superando las deficiencias principales de la estrategia de señuelo, de que puede ser difícil introducir los oligonucleótidos en la célula y que pueden ser sensibles a exonucleasas, proporcionando una vida media relativamente corta.

Una visión global del protocolo de biología molecular usado para crear un plásmido n[snare] se proporciona en la Figura 1. Una visión global de cómo los plásmidos n[snare] influyen en la regulación génica se proporciona en la Figura 2. Una demostración de la eficacia del plásmido n[snare] se proporciona en la Figura 3 en la que se usa para suprimir la producción de antibiótico en *S. coelicolor* por la introducción de una secuencia reguladora en *cis* conocida de un regulador pleyotrópico positivo de la síntesis de antibiótico, AfsR.

1. Diseño de plásmidos n[snare]

Las ventajas de crear versiones portadas en plásmido de oligonucleótidos señuelo en lugar de señuelos por sí mismos para uso en la identificación de la actividad biológica de reguladores en *cis* incluyen:

1. *Coste de fabricación*: El protocolo de biología molecular descrito más adelante es sencillo y robusto y puede aplicarse a oligonucleótidos sintetizados o fragmentos de ADN de diversas fuentes. Como los plásmidos se auto-repican no existe la necesidad de producir grandes cantidades de los plásmidos. Comparativamente esto es favorable con una necesidad potencial de crear cantidades sustanciales de oligonucleótidos señuelo con modificaciones potencialmente costosas del núcleo de los nucleótidos;

2. *Resistencia a la degradación*: En comparación con los señuelos, los plásmidos muestran poca susceptibilidad a degradación *in vivo* por nucleasas;

3. *Se mantiene la concentración de plásmido*: Como los plásmidos se auto-repican (y pueden tener mecanismos para controlar sus concentraciones intracelulares o número de copias) y pueden someterse a selección positiva si es necesario, la vida media de un plásmido n[snare] es teóricamente indefinida;

4. *Rango amplio de huéspedes*: En el Ejemplo 1 se usa un "plásmido lanzadera" que puede transformarse y propagarse fácilmente en *E. coli* (para facilitar la manipulación genética) y *S. coelicolor*. Los expertos en la técnica conocen muchos plásmidos que permiten la transformación de un rango amplio de huéspedes bacterianos;

5. *Pueden ensayarse combinaciones de secuencias reguladoras en cis usando distintos plásmidos n[snare]*: Potencialmente pueden usarse plásmidos compatibles para determinar si un tratamiento simultáneo con dos o más plásmidos n[snare] tiene o no efectos sinérgicos.

En un aspecto, por lo tanto, la invención proporciona un plásmido que es adecuado para ensayar una posible función señuelo de una secuencia reguladora en *cis* conocida o posible o para cribar nuevas secuencias reguladoras en *cis* (que pueden actuar como secuencias señuelo). El plásmido se designa un plásmido n[snare].

El principio de uso de un plásmido n[snare] se ilustra en la Figura 2. El plásmido comprende una secuencia "snare" (mostrada en múltiples copias en la Figura). Si la "snare" comprende un sitio de unión de factor de transcripción que compite con una secuencia reguladora en *cis* celular para unión a un factor de transcripción, la introducción del plásmido n[snare] en la célula resultará en una liberación del factor de transcripción de la secuencia reguladora en *cis* celular y en la snare. Esto puede detectarse como un cambio en la expresión de un gen o genes cuya expresión está regulada en la célula por esa secuencia reguladora en *cis* o por la alteración del fenotipo de la célula. Un cambio en la expresión génica o resultado fenotípico indica así que la snare comprende una secuencia con función señuelo y que el plásmido puede usarse para identificar o confirmar la función de una secuencia reguladora en *cis*.

En general, un plásmido n[snare] comprende un vector plasmídico y una secuencia inserto (comprendiendo el inserto la snare). La incorporación de la secuencia snare en un plásmido aborda los problemas en la técnica de la

degradación del señuelo y permite el mantenimiento estable del señuelo (y cualquier efecto en la expresión génica) en la célula.

La secuencia inserto comprende una o más copias de una secuencia monomérica (que comprende la snare). Así, el inserto puede comprender (por ejemplo) de 1 a 200 secuencias monoméricas. Típicamente, hay dos o más copias, por ejemplo, 2-200 copias, por ejemplo, al menos 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180 ó 190 copias. Por ejemplo, pueden ser 5-200, 5-150, 5-100, 10-150, 10-50, 5-50, 5-40, 10-40, 5-30 copias. Por ejemplo, puede haber 30 copias de la secuencia monomérica. El plásmido comprende típicamente un homopolímero del monómero. Típicamente, hay múltiples copias del monómero, por ejemplo, múltiples repeticiones directas de la secuencia monomérica. El proporcionar múltiples copias del monómero (y así de la snare) incrementa el poder de titulación del señuelo.

La secuencia monomérica comprende la secuencia snare. La snare comprende una secuencia de nucleótidos que se va a ensayar para o usarse para función reguladora en *cis* o señuelo como se describe en la presente memoria.

Por ejemplo, una snare puede comprender una secuencia reguladora en *cis* conocida o posible, una secuencia señuelo, un fragmento de secuencia que se va a ensayar para función señuelo, por ejemplo, un fragmento genómico, tal como un fragmento de promotor, o una secuencia de nucleótidos aleatorizada, o una combinación de secuencias reguladoras en *cis* (por ejemplo, 2, 3, 4 o más secuencias reguladoras en *cis*). Pueden prepararse bibliotecas de plásmidos n[snare], que comprenden secuencias snare derivadas de y que abarcan, sustancialmente toda la secuencia (o secuencia reguladora en *cis*) de un genoma o fragmento genómico. También pueden prepararse bibliotecas de plásmidos n[snare], en las que las secuencias snare comprenden secuencias de nucleótidos aleatorizados con una longitud dada (con una longitud de "n" nucleótidos). Preferiblemente, todas o sustancialmente todas las posibles secuencias con longitud n están representadas en la biblioteca. Estas bibliotecas pueden ser útiles para cribar para identificar secuencias que comprenden y/o compiten con secuencia reguladora en *cis* celular para unión de factores de transcripción y, por lo tanto, para identificar nuevas secuencias reguladoras en *cis* y señuelo.

Típicamente, un sitio de unión de un factor de transcripción en una snare no está unido operativamente a un gen, por ejemplo, en la snare o plásmido snare. En este sentido, el sitio de unión se aísla de su gen o genes afines. Un sitio de unión en una snare también puede aislarse de otros elementos en su promotor afín. En un caso, la secuencia monomérica no comprende un gen.

Un monómero puede comprender secuencia adicional además de la snare. Habitualmente, dicha secuencia adicional deriva del método usado para producir la snare y/o el inserto del plásmido. Por ejemplo, un monómero puede comprender una secuencia adaptadora, tal como la secuencia adaptadora que resulta típicamente cuando se producen snare "personalizadas" para una biblioteca de n[snare] personalizada según los métodos de la presente memoria. Una secuencia adaptadora puede comprender, por ejemplo, sitio de reconocimiento y/o corte para una o más enzimas de restricción.

Un monómero puede comprender una secuencia de nucleótidos que proporciona un sitio de unión para un cebador, por ejemplo, un cebador usado en la producción del monómero o de un inserto que comprende múltiples monómeros.

Por ejemplo, cuando un inserto de plásmido se prepara usando un método de amplificación por círculo rodante como se describe en la presente memoria, un monómero comprende típicamente un segmento que corresponde al sitio de unión para el cebador usado en la replicación por círculo rodante, por ejemplo un cebador T7.

Un monómero que comprende una secuencia snare aleatorizada también comprende típicamente una región o regiones de secuencia constante. Por ejemplo, una secuencia snare aleatorizada de n nucleótidos puede estar flanqueada por regiones de secuencia constante. Alternativamente, un núcleo central de secuencia constante puede estar flanqueado por regiones de secuencia de nucleótidos aleatorizados.

La longitud de un monómero puede ser, por ejemplo, hasta 1.000 nucleótidos, por ejemplo, hasta 900, 800, 700, 600, 500, 400, 300, 200, 100, 90, 80, 70, 60, 50, 40, 30, 20, 15 ó 10 nucleótidos. Típicamente, la longitud del monómero está en el intervalo, por ejemplo, de 10-100, 10-50, 20-75, 30-60, 30-50, 35-55 tal como 35-54 nucleótidos. Por ejemplo, la longitud de un monómero puede ser 30, 40 ó 50 nucleótidos.

Una parte snare de un monómero puede variar típicamente en tamaño de 10-30 nucleótidos, por ejemplo, 10-25, 10-20, 15-20, tal como 15, 16, 17, 18, 19 ó 20, por ejemplo, 19 nucleótidos.

Una secuencia adaptadora puede comprender, por ejemplo, 5-30 nucleótidos, por ejemplo, 5-25, 5-20, 5-15 ó 5-10 nucleótidos tal como 10, 11, 12, 13, 14 ó 15 nucleótidos.

Típicamente, un inserto en el plásmido n[snare] comprende una o más copias de un monómero como se describe en la presente memoria. Cuando el inserto comprende múltiples repeticiones de un monómero, éstas pueden ser repeticiones en tándem.

Un inserto en el vector plasmídico puede comprender, por ejemplo, aproximadamente 1,5 kb, por ejemplo, 1-2 kb, por ejemplo 1,1, 1,2, 1,3, 1,4, 1,5, 1,6, 1,7 ó 1,8 kb. Sin embargo, puede usarse cualquier tamaño de inserto adecuado, que permita el mantenimiento estable del plásmido en una célula huésped adecuada y uso eficaz en los presentes métodos.

5 Típicamente, las secuencias de todos los monómeros en un único plásmido son las mismas.

El vector plasmídico para uso en el plásmido n[snare] puede ser adecuado para uso en un huésped procariota o eucariota. Por ejemplo, el vector puede ser para uso en un procariota tal como un huésped bacteriano, por ejemplo, actinomicete. Por ejemplo, el vector plasmídico puede ser adecuado para uso en una cepa de *Streptomyces* o *E. coli*, por ejemplo una o más de *Streptomyces coelicolor*, por ejemplo, A3(2) (o cepa M145 o M600), *E. coli*. *Streptomyces lividans* o *Streptomyces cinnamomeus*. Los huéspedes adecuados se describen adicionalmente en la presente memoria.

10

Típicamente, el vector es un vector de amplio rango de huéspedes y/o lanzadera y puede por lo tanto mantenerse y propagarse en más de un huésped. El plásmido puede ser un plásmido conjugativo. Esto permite una transferencia fácil de una célula a otra por conjugación.

15 Preferiblemente, el plásmido es auto-replicativo. Típicamente, el plásmido es un plásmido de alto número de copias. Por ejemplo, el plásmido puede mantenerse a, por ejemplo, 20-100 copias por célula, por ejemplo 20, 30, 40, 50, 60, 70, 80, 90 ó 95 ó 100 copias por célula. El alto número de copias incrementa el poder de titulación de una secuencia señuelo en la snare.

20 Típicamente, un plásmido n[snare] comprende un origen de replicación. Los orígenes adecuados se conocen en la técnica. Típicamente, el plásmido comprende además uno o más genes marcadores detectables, por ejemplo, uno o más genes que codifican resistencia a antibióticos, por ejemplo, el gen *aac* que codifica resistencia a apramicina. La expresión del gen o genes marcadores permite el cribado para el mantenimiento del plásmido en una célula huésped.

25 Los ejemplos de vectores plasmídicos adecuados incluyen por ejemplo, pIJ86. Los vectores adecuados se conocen en la técnica.

Un monómero o secuencia snare puede comprender una secuencia reguladora en *cis* o secuencia señuelo descrita en la presente memoria y/o identificada según los métodos descritos en la presente memoria.

30 Por ejemplo, un sitio de unión de factor de transcripción en un plásmido puede comprender o competir para unión del factor de transcripción con, una secuencia reguladora en *cis* que regula la expresión de un gen o genes que tienen un papel en la producción de metabolitos. El metabolito puede ser, por ejemplo, un antibiótico, enzima o compuesto farmacéutico. Un antibiótico puede ser, por ejemplo, actinorrodina, undecilprodigiosina o cinamicina.

35 Por ejemplo, una snare puede comprender un sitio de unión para la proteína AfsR (un regulador pleyotrópico de la síntesis de antibiótico en *S. coelicolor*). Un sitio de unión de AfsR se muestra en negrita en SEQ ID NO: 1 (Ejemplo 1). Un monómero puede comprender así SEQ ID NO: 1. Así, un monómero puede ser la versión bicatenaria de SEQ ID NO: 1 como en el Ejemplo 1. En un ejemplo, puede haber 30 repeticiones del monómero y/o el plásmido puede comprender el vector lanzadera pIJ86. Así, en un aspecto, la invención se refiere a un plásmidos o plásmidos n[snare] preparados según el método en el Ejemplo 1.

40 En otro ejemplo, una snare puede comprender una secuencia reguladora en *cis* en el promotor *actII-orf4* de *S. coelicolor*, por ejemplo, una secuencia reguladora represora en *cis* o una secuencia que compite con dicha secuencia para unión del factor de transcripción, como se describe en la presente memoria. Por ejemplo, una secuencia reguladora en *cis* puede comprender la secuencia A24.1, A24.2, A24.3, A24.4 o A24.5 identificada en la presente memoria. En un aspecto, un monómero puede comprender una de SEQ ID NOs: 8-12 descritas en la presente memoria o SEQ ID NO: 13.

45 Un sitio de unión de factor de transcripción en un plásmido puede comprender, o competir para unión del factor de transcripción con, una secuencia reguladora en *cis* que regula la expresión de un gen o genes que tienen un papel en la determinación de la resistencia a antibióticos, tal como uno cualquiera o más de los genes listados en la presente memoria.

50 Por ejemplo, una snare puede comprender un sitio de unión para el factor de transcripción VanR, por ejemplo, un sitio de unión VanR localizado en el promotor *vanH*, por ejemplo, de *Enterococcus faecium* o *S. coelicolor* o una secuencia que compite con dicho sitio. Un ejemplo de un sitio de unión VanR de 30 pb se muestra (en mayúsculas) en SEQ ID NO: 21 y un ejemplo adicional se proporciona en SEQ ID NO: 26 que contiene el sitio de unión VanR para Enterococci. En un ejemplo, un monómero puede comprender SEQ ID NO: 21 o SEQ ID NO: 26 o una variante de ésta que compite para unión del factor de transcripción VanR con un sitio de unión VanR nativo en una célula de interés. Por ejemplo, una variante puede comprender el sitio de unión VanR nativo en otra especie o cepa.

Un sitio de unión de un factor de transcripción en un plásmido puede comprender, o competir para unión del factor de transcripción con, una secuencia reguladora en *cis* que regula la expresión de un gen o genes que tienen un papel en la determinación de la tolerancia a disolventes, por ejemplo, tolerancia a butanol.

Una snare puede comprender una secuencia derivada o aislada de un genoma o fragmento genómico. Un fragmento genómico puede comprender un gen o genes que codifican una función o fenotipo particular de interés, por ejemplo, producción de un metabolito o metabolitos tal como un antibiótico (por ejemplo, cinamicina, actinorhodina, undecilprodigiosina), tolerancia a un disolvente o disolventes particulares, por ejemplo, butanol o toxina(s), resistencia a un o unos antibióticos particulares, o cualquier otra función o fenotipo de interés. Una snare puede comprender, por ejemplo, una secuencia derivada o aislada de la región promotora de dicho gen o genes o secuencias circundantes, por ejemplo, secuencias a una distancia mayor de un gen o genes que el promotor, por ejemplo, hasta 200 pb o más.

Una snare puede comprender una secuencia que compite con dicha secuencia para unión del factor de transcripción, como se describe en la presente memoria. Por ejemplo, una snare puede comprender una secuencia reguladora en *cis* o secuencia señuelo derivada de un fragmento del genoma de *S. cinamomeus* que comprende la agrupación de genes biosintéticos de cinamicina. Los métodos para preparar dichas snare se describen en la presente memoria. Una snare puede comprender una secuencia reguladora en *cis* del promotor *cin7* de *S. cinamomeus* o una secuencia que compite con dicha secuencia para unión del factor de transcripción, como se describe en la presente memoria.

Una snare puede comprender una secuencia derivada o aislada del genoma o un fragmento genómico de un procarionta que presenta tolerancia a disolvente, por ejemplo, *E. coli* K12, que tiene tolerancia a butanol. En particular, una snare puede comprender una secuencia reguladora en *cis* derivada del promotor de un gen que codifica tolerancia a disolvente o cuya expresión está asociada con tolerancia a disolvente, o una secuencia que compite con dicha secuencia reguladora en *cis* para unión del factor de transcripción. Los métodos para preparar dichas snare se describen en la presente memoria.

Una snare puede comprender una secuencia (por ejemplo, una secuencia reguladora en *cis* o secuencia señuelo) derivada de o aislada del genoma o un fragmento genómico de un procarionta que presenta resistencia a antibiótico (típicamente el fragmento comprende el o los genes que codifican la resistencia). El fragmento genómico puede comprender un gen o genes que codifican resistencia a antibiótico. Una snare puede comprender una secuencia reguladora en *cis* del promotor de un gen o genes que codifican resistencia a antibiótico o una secuencia señuelo que compite con dicha secuencia para unión del factor de transcripción. Cualquier resistencia a antibiótico de interés puede ser diana. Por ejemplo, resistencia a antibiótico a: la clase de antibióticos conocida como aminoglicósidos (tal como kanamicina y gentamicina); los glicopéptidos (tal como vancomicina); las beta-lactamas que incluyen las penicilinas (tal como ampicilina, carbenicilina y penicilina), los inhibidores de beta-lactamasa y combinaciones de (tal como piperacilina y tazobactam), las cefalosporinas (tal como cefepime), los carbapenemos (tal como meropenem), las monobactamas (tal como Aztreonam); los antibióticos polipeptídicos (tal como polimixina B); las quinolonas (tal como levaquin); las fluorquinolonas (tal como ciprofloxacina); las sulfonamidas (tal como Bactrim); las tetraciclinas (tal como tetraciclina); los macrólidos y ketólidos (tal como azitromicina); las oxazolidinonas (tal como linezolid); los nitroimidazoles (tal como metronidazol); los nitrofuranos (tal como nitrofurantoína); las estreptograminas (tal como dalfopritina); los lipopéptidos cíclicos (tal como daptomicina); las lincosamidas (tal como clindamicina) y de forma diversa, cloranfenicol, rifampicina, isoniazid, etambutol, telavancina, teicoplanina, oritavancina, dalbavancina, trimetoprim/sulfametoxazol, fosfomicina, nitrofurantoína y tigeciclina y Zyvox.

Por ejemplo, resistencia a antibiótico a: la clase de antibióticos conocida como aminoglicósidos (tal como una kanamicina); los carbapenemos (tal como meropenem); las cefalosporinas (tal como cefepime); los glicopéptidos (tal como vancomicina); las penicilinas (tal como ampicilina, carbenicilina y penicilina); los antibióticos polipeptídicos (tal como polimixina B); las quinolonas (tal como levaquin); las sulfonamidas (tal como Bactrim); las tetraciclinas (tal como tetraciclina); y de forma diversa, cloranfenicol, rifampicina, Zyvox y daptomicina.

bibliotecas de plásmidos n[snare]

Habiendo descrito la creación de plásmidos n[snare] capaces de establecer la actividad funcional de elementos reguladores en *cis* únicos, aquí generalizamos los métodos por los que estas secuencias pueden ser descubiertas y caracterizadas, incluyendo mediante un método usando bibliotecas de plásmidos n[snare], una metodología compatible con cribado de alto rendimiento.

El objetivo global de esta metodología es desarrollar un nuevo método para identificar rápidamente reguladores que actúan en *cis* de genes procariontas así como eucariotas y haciendo esto crear herramientas capaces de modificar la expresión *in vivo*. El método permite la disección de redes reguladoras a gran escala. Los avances anticipados en nuestra comprensión de los mecanismos de la regulación genética serán paralelos al avance que rinde el análisis transcriptómico con micromatrices.

Este trabajo identifica modificadores genéticos y determina rutas de una manera diferente a las tecnologías existentes tales como creación de inactivaciones génicas o mutagénesis insercional. Estas estrategias identifican

generalmente reguladores que actúan en *trans* (tales como los factores de transcripción). Mediante la identificación de factores que actúan en *cis*, esta tecnología supera los problemas asociados con estas estrategias tradicionales: redundancia – debida a la complejidad inherente de la regulación transcripcional, lo que significa que más de un factor de transcripción puede unirse al mismo sitio; efectos polares - la mutagénesis puede influir en la expresión de genes localizados 3' del sitio de la mutación en la misma unidad de transcripción; alto coste y tiempo de respuesta– la generación y validación de una única inactivación puede durar muchas semanas a años (dependiendo de la especie) y realizar mutagénesis de saturación es un procedimiento largo y costoso; especificidad de especie – la estrategia descrita en la presente memoria tiene utilidad potencial en todos los sistemas experimentales y se prevé que una biblioteca “señuelo” universal puede aplicarse al mapeo de la regulación en la mayoría de las especies.

En el Ejemplo 1, demostramos una mejora sustancial en la tecnología de señuelo, indicando una modificación que hace que la estrategia sea portada en plásmido mediante el desarrollo de un método para clonar homopolímeros de secuencias señuelo, para crear plásmidos “n[snare]” de alto número de copias. Éstos se introducen fácilmente en muchas especies bacterianas y se mantienen por selección positiva, superando las deficiencias principales de la estrategia señuelo, de que puede ser difícil introducir los oligonucleótidos en la célula y pueden ser sensibles a exonucleasas, proporcionando una vida media relativamente corta. En una segunda modificación y mejora, creamos bibliotecas n[snare] que consisten en fragmentos del genoma completo, detallados en el Ejemplo 2, o secuencias derivadas de oligonucleótidos aleatorizados, descritos en el Ejemplo 4. Estas estrategias permiten la recogida exhaustiva de elementos reguladores para ser cribados en paralelo, lo que es una estrategia mucho más potente que el uso de señuelos para propósitos de cribado, que está limitada a ensayar secuencialmente secuencias definidas. Más adelante, en el Ejemplo 3, se usa un sistema informador, usando la regulación de la producción de antibiótico como un modelo, en una realización ejemplar para seleccionar reguladores positivos y negativos en la biblioteca.

La Figura 4 es un esquema que muestra parte del proceso de cómo se construyen bibliotecas de plásmido n[snare] personalizadas.

La Figura 5 muestra una estrategia experimental usando una biblioteca n[snare] para detectar elementos reguladores en *cis* capaces de regular al alza la producción del antibiótico cinamicina de *Streptomyces cinamomeus*.

La Figura 6 proporciona detalles adicionales para la aplicación de n[snare] para combatir la resistencia a antibióticos.

bibliotecas de plásmidos n[snare] como herramientas de descubrimiento

La expresión génica, y simultáneamente el efecto fenotípico, están controlados en gran medida por las interacciones ADN-proteína. Los esfuerzos para entender y controlar la expresión génica generalmente ponen más énfasis en la manipulación del componente proteína de la maquinaria transcripcional, los factores que actúan en *trans* como los factores de transcripción, que en las secuencias de ADN (los elementos reguladores en *cis*) a los que se unen. El tomar como diana los elementos reguladores en *cis*, a diferencia de los factores que actúan en *trans*, tiene las ventajas siguientes: es más fácil transfectar oligonucleótidos en células que hacer deleciones genéticas de las proteínas diana; el uso de oligonucleótidos es más adaptable al análisis de alto rendimiento; el uso de elementos que actúan en *cis* supera el problema de redundancia en las redes reguladoras, en las que la delección de un factor de transcripción particular se compensa por la actividad de otra unión a la misma secuencia reguladora en *cis*. Un cuello de botella potencial con la estrategia señuelo es, sin embargo, la identificación de secuencias candidatas.

Para este propósito, se demuestra la utilidad de vectores n[snare] creando bibliotecas a partir de un fragmento genómico y ensayando para observar si los miembros de la biblioteca resultante pueden ejercer control sobre el producto sintetizado por los genes contenidos en ese fragmento. Además, la biología molecular de la creación de la biblioteca se ensaya; para confirmar que las bibliotecas son lo suficientemente complejas y tienen la estructura molecular esperada de repeticiones directas de la misma secuencia. Estos éxitos se desarrollan más para crear un método de cribado potente con aplicaciones amplias y genéricas.

Como se ha descrito anteriormente, por lo tanto, la invención también se refiere a bibliotecas de plásmidos n[snare]. Las bibliotecas de plásmidos pueden usarse según los métodos de la invención para cribar nuevas secuencias reguladoras en *cis* y secuencias señuelo en procariotas o eucariotas. El uso de bibliotecas n[snare] para cribar permite un cribado de alto rendimiento de múltiples secuencias para función señuelo.

Las secuencias snare de los plásmidos en la biblioteca pueden derivar de un fragmento genómico o un ADN genómico. La biblioteca puede comprender secuencias snare que representan o abarcan sustancialmente todo el fragmento genómico o secuencia genómica. Este tipo de biblioteca de plásmidos se refiere en la presente memoria como una biblioteca n[snare] personalizada. Alternativamente, las secuencias snare de los plásmidos en la biblioteca pueden comprender una secuencia aleatorizada con una longitud de “n” nucleótidos. La biblioteca puede comprender sustancialmente cada secuencia de nucleótidos posible con una longitud “n” como se describe en la presente memoria. Dicha biblioteca se describe en la presente memoria como una biblioteca n[snare] universal.

Bibliotecas de plásmidos n[snare] personalizadas

Típicamente, las snare derivan de un fragmento genómico o ADN genómico (procariota o eucariota). Por ejemplo, el genoma o un fragmento genómico de un procariota tal como una bacteria, por ejemplo, *Streptomyces*, tal como *S. cinnamomeus*, *S. lividans* o *S. coelicolor* o *E. coli*.

- 5 El genoma o fragmento genómico del que se deriva la biblioteca puede comprender un gen o genes que codifican una función o fenotipo particular de interés, por ejemplo, producción de un metabolito o metabolitos tal como un antibiótico (por ejemplo, cinamicina, actinorhodina, undecilprodigiosina), tolerancia a un disolvente o disolventes o toxina o toxinas particulares, resistencia a un antibiótico o antibióticos particulares o cualquier otra función o fenotipo de interés. Las snare pueden comprender secuencias derivadas o aisladas de la región promotora de dicho gen o genes. Los genes y funciones/fenotipos adecuados se han descrito anteriormente en relación con las secuencias snare.

- 15 Así, el ADN genómico o fragmento puede comprender un gen o genes que tienen un papel en la producción de metabolitos, por ejemplo, en el que el metabolito es un antibiótico, enzima o compuesto farmacéutico, o un gen o genes que tienen un papel en la determinación de la resistencia a antibióticos o un gen o genes que tienen un papel en la determinación de la tolerancia a disolventes.

Un fragmento genómico que comprende un gen o genes que codifican una función o fenotipo particular puede haberse identificado detectando la transferencia horizontal de los genes que codifican la función o fenotipo en una célula heteróloga.

- 20 Un fragmento genómico que comprende un gen o genes que codifican una función o fenotipo particular puede haberse identificado por análisis bioinformático. Alternativamente, puede haberse identificado por un cribado funcional de una colección de dichos fragmentos.

Los métodos para preparar bibliotecas n[snare] personalizadas se describen en la presente memoria.

Bibliotecas universales

- 25 En una biblioteca universal, una snare comprende típicamente una secuencia de nucleótidos aleatorizados con una longitud de n nucleótidos. La biblioteca puede comprender snare que representan todas o sustancialmente todas las permutaciones de secuencias con una longitud de n nucleótidos. En general, n puede variar de 5-50, por ejemplo 10-50, por ejemplo, 20-40, por ejemplo 25-35, por ejemplo 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 ó 20 nucleótidos.

- 30 Un monómero en dicha biblioteca puede comprender además secuencia constante. Así, la secuencia aleatorizada (región variable) puede estar flanqueada por una región constante en uno o los dos lados. Alternativamente, el monómero puede comprender una secuencia central definida flanqueada por regiones variables. Puede introducirse un sesgo de nucleótido en la región variable. Por ejemplo, la región variable puede comprender un sesgo GC alto cuando es apropiado, por ejemplo, cuando la biblioteca se va a usar para el cribado de secuencias reguladoras en *cis* en un organismo en el que el ADN presenta un sesgo GC alto.

- 35 En un ejemplo, la snare comprende una secuencia de nueve nucleótidos aleatorizados (n = 9). El monómero en los plásmidos n[snare] puede comprender la secuencia de oligonucleótidos de SEQ ID NO: 7. Los métodos para producir una biblioteca n[snare] universal se describen en la presente memoria.

Métodos para preparar plásmidos n[snare]

- 40 En general, un plásmido n[snare] como se describe en la presente memoria puede prepararse por un método que comprende:

- proporcionar un polinucleótido que comprende una o más copias de una secuencia monomérica; y
- clonar el polinucleótido en un vector plasmídico adecuado.

La composición del monómero es como se describe en la presente memoria para el plásmido n[snare]. Los vectores plasmídicos adecuados también se han descrito.

- 45 El polinucleótido que comprende una o más copias de una secuencia monomérica puede proporcionarse por un método que comprende:

(1) proporcionar un oligonucleótido circular que comprende:

- (i) una secuencia de ensayo de interés (B en la Figura 1); y
- (ii) un sitio de unión para un cebador adecuado para uso en amplificación por círculo rodante (A en la Figura 1);

- 50 en el que la secuencia monomérica comprende (i) y (ii); y

(2) realizar la amplificación por círculo rodante usando el oligonucleótido circular como molde, proporcionando de esta manera un polinucleótido que comprende repeticiones de la secuencia monomérica.

La etapa (1) del método puede comprender además amplificar los productos de la amplificación por círculo rodante por PCR y aislar los fragmentos de polinucleótido del tamaño requerido, por ejemplo, los fragmentos que comprenden 30-50 repeticiones de la secuencia monomérica. Esto puede hacerse, por ejemplo, por análisis de PAGE.

Un oligonucleótido circular puede prepararse por un método que comprende:

- proporcionar un oligonucleótido lineal monocatenario que comprende: la secuencia de ensayo de interés (i) y el sitio de unión para un cebador adecuado para uso en la amplificación por círculo rodante (ii);

- circularizar el oligonucleótido, por ejemplo, usando ligasa Taq, típicamente en presencia de un oligonucleótido de unión universal;

- opcionalmente digerir el ADN lineal restante con una exonucleasa; y

- recuperar los oligonucleótidos circulares monoméricos, por ejemplo usando PAGE.

Los cebadores adecuados para uso en la amplificación por círculo rodante se conocen en la técnica. Por ejemplo, puede usarse un cebador T7.

Los métodos para realizar la amplificación por círculo rodante se conocen en la técnica. Por ejemplo, puede usarse la polimerasa *BstI*.

En un ejemplo, la amplificación por PCR de los productos de la amplificación por círculo rodante se realiza usando el mismo cebador que se usó para la amplificación por círculo rodante, por ejemplo el cebador T7.

En general, la secuencia de ensayo (i) en el monómero comprende una secuencia snare como se describe en la presente memoria.

Como se describe en la presente memoria, las secuencias snare pueden aislarse a partir de ADN genómico, por ejemplo a partir de un genoma completo, o de un fragmento genómico. Una secuencia snare, una vez aislada, puede usarse para formar un plásmido n[snare] por el método anterior. Así, puede prepararse una biblioteca de plásmidos n[snare] derivada de un genoma o fragmento genómico (una biblioteca personalizada como se describe en la presente memoria).

Un protocolo para la preparación de secuencias snare a partir de un genoma o fragmento genómico se ilustra en la Figura 4. Típicamente un método comprende:

(1) proporcionar una muestra de ADN genómico bicatenario;

(2) fragmentar el ADN genómico;

(3) ligar un adaptador a cada extremo de los fragmentos de ADN de (b), en el que cada adaptador comprende:

(iii) un medio para la inmovilización del fragmento de ADN;

(iv) un sitio de reconocimiento para una primera enzima de restricción que corta a una distancia de (por ejemplo, en 3' (3') del sitio de reconocimiento; y

(v) un sitio de reconocimiento y corte para una segunda enzima;

(4) opcionalmente eliminar el adaptador no ligado;

(5) digerir el fragmento que presenta los adaptadores con la primera enzima de restricción, produciendo de esta manera dos fragmentos con adaptador, en el que cada fragmento con adaptador comprende:

(vi) un adaptador; y

(vii) un fragmento de ADN que comprende: una cadena más corta; y una cadena más larga (por ejemplo con una protuberancia 3'), en el que la cadena más larga comprende la secuencia snare;

(6) inmovilizar los fragmentos con adaptador producidos en (5);

(7) desnaturalizar los fragmentos para proporcionar fragmentos monocatenarios;

(8) recrear el sitio de reconocimiento para la segunda enzima de restricción ligando un oligonucleótido complementario al adaptador y digiriendo el fragmento con adaptador con la segunda enzima de restricción, produciendo de esta manera un fragmento adaptador-snare que comprende:

(viii) un fragmento adaptador; y

5 (ix) una secuencia snare monocatenaria; y

(9) liberar el fragmento adaptador-snare de la inmovilización.

10 El fragmento adaptador-snare producido en (8) comprende la secuencia snare monocatenaria y un fragmento del adaptador que permanece después de la digestión con la segunda enzima de restricción (correspondiente a la parte del adaptador entre el sitio de reconocimiento para la segunda enzima y el extremo del adaptador que está unido a la snare). El monómero en el plásmido n[snare] eventual (y la secuencia de ensayo (i) anterior) comprenderán el fragmento adaptador-snare.

15 La muestra de ADN genómico puede derivar de una célula procariota o eucariota. La muestra de ADN genómico puede comprender un fragmento genómico o un genoma completo. La muestra puede ser de una célula que presenta un fenotipo particular de interés, por ejemplo, producción de un metabolito particular, resistencia a un antibiótico particular, tal como una cepa nativa (por ejemplo, un patógeno) o aislado clínico. Por ejemplo, la muestra puede ser de *S. cinnameus* que produce cinamicina, de *S. coelicolor* que produce actinorhodina, undecilprodigiosina, de *E. coli* K12 que presenta tolerancia al butanol, o de *Enterococcus faecium* o *S. coelicolor* que presenta resistencia a vancomicina. La muestra puede ser de un modelo bacteriano, por ejemplo que ha adquirido un fenotipo o función particular por transferencia génica horizontal.

20 Típicamente, los fragmentos producidos en la etapa (2) tienen una longitud de aproximadamente 500 pb, pero pueden variar, por ejemplo, de 100-1.000 pb, tal como 200-900, 300-800, 400-600 pb, tal como 150, 250, 350, 450, 550, 650, 750, 850, 950 pb. Puede usarse cualquier método de fragmentación adecuado para producir los fragmentos. Preferiblemente, el método produce fragmentos aleatorizados o no sesgados. Por ejemplo, puede usarse la sonicación.

25 Los fragmentos producidos en la etapa (2) pueden comprender diferentes tipos de extremos – por ejemplo, extremos romos o protuberancias 5' ó 3' de diferentes longitudes. Por lo tanto, los fragmentos producidos en la etapa de fragmentación (k) pueden tratarse además para producir una población de fragmentos que son homogéneos en que cada fragmento tiene la misma protuberancia 3' dNTP – dA o dT. Esto puede hacerse, por ejemplo, tratando con polimerasa Taq y dNTP adecuados para reparar los extremos del fragmento si es necesario y añadir un 3' dNTP, por ejemplo protuberancia dA.

30 Los adaptadores pueden unirse a los fragmentos por cualquier medio adecuado. Por ejemplo, un adaptador puede comprender una protuberancia 5' dNTP que es complementaria a la protuberancia 3' dNTP de cada uno de los fragmentos. Así, por ejemplo, si los fragmentos producidos en la etapa (b) comprenden una protuberancia 3' dA, un adaptador puede comprender una protuberancia 5' dT.

35 En el método mostrado anteriormente, los adaptadores comprenden un medio para inmovilizar el fragmento de ADN. La inmovilización no es esencial para el presente método pero tiene la ventaja de que reduce el ruido experimental. Un adaptador puede comprender cualquier medio adecuado para inmovilizar los fragmentos. Por ejemplo, un adaptador puede comprender típicamente en el extremo distal al fragmento de ADN) un miembro de una pareja de moléculas de unión, en el que las moléculas de unión en la pareja se unen entre sí, y en el que el otro miembro de la pareja puede estar comprendido en una matriz de inmovilización adecuada. Por ejemplo, las moléculas de unión en la pareja pueden ser biotina y estreptavidina. Pueden usarse adaptadores biotinilados y el fragmento con adaptador captarse en una matriz de estreptavidina.

45 La primera enzima de restricción corta a una distancia de su sitio de reconocimiento. Por ejemplo, la enzima puede cortar en 3' (3') de su sitio de reconocimiento. La distancia a la que corta determina la longitud de la secuencia snare. Las longitudes adecuadas para la secuencia snare se describen en la presente memoria. Los ejemplos de enzimas que pueden usarse son *MmeI*. Otros ejemplos incluyen *GsuI*, *BpmI* e isoesquizómeros de éstas. Los ejemplos adicionales incluyen miembros de la familia de las enzimas de restricción conocidas como tipo II que cortan fuera de su sitio de reconocimiento en un lado. En la actual descripción, se usan las enzimas que cortan en el lado 3' y se prefieren las que cortan a una distancia mayor, típicamente 15-25 nt (revisado en *Gene* (1991) **100**: 13-26). La naturaleza del corte bicatenario introducido no es crucial pero en los métodos descritos aquí se usa una protuberancia 3'. El método podría adaptarse para usar una protuberancia 5' y un corte romo. La naturaleza de la snare podría alterarse de acuerdo con esto. Por ejemplo, si se usara una enzima que proporciona un producto asimétrico con una protuberancia 5' más larga, sería ventajoso convertir la protuberancia 5' en la snare. También dependiendo de cómo está situado el sitio de reconocimiento (que es típicamente asimétrico) en el adaptador, podrían usarse enzimas que introdujeran cortes 5' de su sitio de reconocimiento.

La segunda enzima de restricción corta en la secuencia del adaptador. Los ejemplos de enzimas que pueden usarse incluyen Nt.Bbv.CI. Ésta se conoce como una endonucleasa de mellado y se usa en esta solicitud ya que introduce

una mella monocatenaria en la cadena superior para permitir la recuperación de la snare. Sin embargo, el método puede adaptarse para usar la mayoría de las enzimas de restricción disponibles comúnmente. Los ejemplos de enzimas que podrían usarse incluyen aquellas que generan protuberancias 3' ó 5' o que generan un corte con extremo romo.

- 5 Una vez preparada, una biblioteca de plásmidos n[snare] puede ensayarse para sesgo en su construcción. Típicamente, esto se hace aislando los plásmidos, digiriendo para liberar el inserto y digiriendo adicionalmente (por ejemplo, con la primera enzima de restricción) para aislar las secuencias monoméricas.

- 10 La creación de bibliotecas de plásmidos n[snare] personalizadas se describe en los presentes Ejemplos. Una biblioteca se prepara a partir de una agrupación génica biosintética de cinamicina de *S. cinnamoneous* en el Ejemplo 2 y una biblioteca se prepara a partir de ADN genómico de *E. coli* K12 en el Ejemplo 4. La invención se refiere a las bibliotecas como se preparan en estos Ejemplos.

- 15 Como se describe en la presente memoria, una secuencia snare puede comprender una secuencia de nucleótidos aleatorizada. Típicamente, la snare comprende una secuencia de nucleótidos aleatorizada (o variable) con una longitud de "n" nucleótidos. En general, n puede variar de 5-50, por ejemplo 10-50, por ejemplo, 20-40, por ejemplo 25-35, por ejemplo 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 ó 20 nucleótidos. Puede prepararse una biblioteca de plásmidos en la que las snare comprenden secuencia aleatorizada con una longitud n. Dicha biblioteca es una biblioteca n[snare] universal.

- 20 Puede sintetizarse un oligonucleótido que comprende la secuencia aleatorizada (y opcionalmente una secuencia constante como se describe) usando los métodos conocidos en la técnica. Típicamente, en la preparación de snare para una biblioteca n[snare] universal, se preparan oligonucleótidos que comprenden todas o sustancialmente todas las secuencias de nucleótidos posibles con una longitud n.

Puede introducirse un sesgo de nucleótidos en la región aleatorizada (o variable) del oligonucleótido. Por ejemplo, puede introducirse un sesgo GC si es apropiado.

- 25 La secuencia monomérica o de ensayo de interés (B en la Figura 1) en el método anterior puede comprender un oligonucleótido preparado de esta manera. Un plásmido n[snare] puede prepararse como se describe.

El Ejemplo 4 describe la preparación de una biblioteca n[snare] universal en la que n= 9. La invención se refiere a la biblioteca preparada según el método en el Ejemplo 4.

- 30 En un aspecto, la invención se refiere a un plásmido n[snare] que comprende un monómero y/o snare que comprende una secuencia reguladora en *cis* o secuencia señuelo descrita en la presente memoria y/o identificada según los métodos descritos en la presente memoria, incluyendo cualquiera de los plásmidos o bibliotecas de plásmidos n[snare] descritas en la presente memoria y/o preparadas según los métodos de la presente memoria, incluyendo los Ejemplos.

Células huésped

- 35 La invención se refiere a una célula o células huésped que comprenden un plásmido o biblioteca de plásmidos n[snare] como se describe en la presente memoria. La invención se refiere además a una célula o células huésped que comprenden un polinucleótido señuelo (molécula señuelo) como se describe en la presente memoria. En particular, la invención se refiere a dicha(s) célula(s) huésped que presenta expresión génica y/o fenotipo alterado, debido a la presencia del plásmido, o polinucleótido señuelo, por ejemplo producción incrementada de un metabolito o producto de fermentación tal como un antibiótico o enzima, resistencia incrementada a un disolvente, sensibilidad incrementada a uno o más antibióticos, comparado con la célula en ausencia del plásmido/molécula señuelo.

Típicamente, el plásmido o biblioteca de plásmidos o polinucleótido se ha introducido en la(s) célula(s) por un medio adecuado, por ejemplo, transformación, transfección o conjugación.

- 45 Una célula huésped para uso en los presentes métodos puede ser procariota o eucariota. Por ejemplo, puede usarse un procariota tal como una célula bacteriana. Un huésped puede ser, por ejemplo, un actinomicete tal como una especie de streptomyces, por ejemplo *S. coelicolor*, por ejemplo *S. coelicolor* A3 (2), (cepa M145 o M600), *S. lividans*, *S. cinnamoneous* o *E. coli*. Otros ejemplos pueden incluir otras especies de bacterias gram positivas, tales como las del grupo Actinobacteria, por ejemplo, bacterias del género *Mycobacterium*, tal como la bacteria patógena *Mycobacterium tuberculosis*, *M. bovis*, *M. africanum* y *M. microti*; *M. leprae*. Un ejemplo adicional de un género de bacterias gram positivas es *Clostridium*, que incluye bacterias patógenas tales como *Clostridium difficile* (un patógeno humano), *C. botulinum*, *C. perfringens* y *C. tetani*, así como bacterias de uso industrial potencial tal como *C. acetylbutylicum*, *C. thermocellum* y *C. ljungdahlii*. Otros géneros de bacterias gram positivas pueden incluir *Bacillus*, *Listeria*, *Staphylococcus*, *Clostridium*, *Corynebacterium*, *Streptococcus* y *Enterococcus*. Las células huésped también pueden ser bacterias gram negativas, que incluyen el género Enterobacteriaceae que incluye patógenos humanos, tales como *Salmonella* y *Escherichia coli*. Otros ejemplos de géneros de bacterias gram negativas pueden incluir *Pseudomonas*, *Bordetella*, *Borrelia*, *Brucella*, *Campylobacter*, *Francisella*, *Haemophilus*, *Klebsiella*, *Neisseria*, *Proteobacteria*, *Rickettsia*, *Vibrio*, *Yersinia*, *Moraxella*, *Helicobacter*, *Stenotrophomonas*,

Bdellovibrio, bacterias de ácido acético, *Legionella*, las cianobacterias, espiroquetas, bacterias verdes del azufre y verdes no del azufre, y muchas otras. Los patógenos gram negativos importantes incluyen las especies de cocci que causan una enfermedad de transmisión sexual (*Neisseria gonorrhoeae*), una meningitis (*Neisseria meningitidis*) y síntomas respiratorios (*Moraxella catarrhalis*). Otros bacilos gram negativos médicamente relevantes incluyen una multitud de especies. Algunas de ellas causan principalmente problemas respiratorios (*Haemophilus influenzae*, *Klebsiella pneumoniae*, *Legionella pneumophila*, *Pseudomonas aeruginosa*), principalmente problemas urinarios (*Escherichia coli*, *Proteus mirabilis*, *Enterobacter cloacae*, *Serratia marcescens*) y principalmente problemas gastrointestinales (*Helicobacter pylori*, *Salmonella enteritidis*, *Salmonella typhi*). Las células huésped también podrían incluir eucariotas tales como levaduras (ejemplos de las cuales pueden incluir: las usadas para producción industrial tales como el género *Saccharomyces* (tal como *S. cerevisiae*), *Schizosaccharomyces* (*S. pombe*) y el género de levaduras metilotróficas *Pichia* (tal como *P. pastoris*) y *Candida*, también *Hansenula polymorpha*; levaduras patógenas tales como el género *Candida* (tal como *C. albicans* y *C. tropicalis*), también el género *Cryptococci* (tal como *C. neoformans*), hongos (que pueden incluir: hongos patógenos tales como el género *Candida*, *Aspergillus* (tal como *A. fumigatus* y *A. flavus*), *Crptococcus*, *Histoplasma* (tal como *H. capsulatum*), *Pneumocystis* (tal como *P. jirovecii*) y *Stachybotrys* (tal como *S. chartarum*); los hongos usados en producción industrial tal como el género *Aspergillus* (en particular *A. niger* y *A. oryzae*) y miembros del género *Neurospora*, tal como *N. crassa*, células de plantas y células de mamíferos, células de aves, bien en cultivo celular o parte de un tejido.

En general, una célula huésped es compatible con el vector plasmídico. Por ejemplo, la célula es compatible con el origen de replicación del plásmido. La célula es típicamente una en la que el plásmido puede mantenerse y replicarse de manera estable. La célula huésped también es típicamente compatible con cualquier gen marcador detectable en el plásmido de manera que el gen puede expresarse en la célula y la célula puede cribarse de esta manera para la introducción y mantenimiento exitosos del plásmido.

Una célula huésped puede considerarse simplemente para manipulación de un plásmido o polinucleótido. Preferiblemente, dicha cepa es una cepa de laboratorio que puede manipularse y mantenerse fácilmente en condiciones de laboratorio, por ejemplo *E. coli*.

Alternativamente o adicionalmente, una célula huésped puede usarse para cribar para secuencias reguladoras en *cis* usando una biblioteca de plásmidos *n[sname]* según la invención y/o para ensayar para función señuelo de una posible secuencia reguladora en *cis* usando un plásmido *n[sname]* como se describe en la presente memoria. La célula huésped también puede ser una en la que se desea alterar la expresión génica usando un plásmido *n[sname]* u otra molécula señuelo como se describe en la presente memoria.

En general, una célula huésped comprende la secuencia reguladora en *cis* de interés, es decir, la secuencia reguladora en *cis* que se está cribando o con la que se pretende que compita la secuencia señuelo introducida en la célula. Típicamente, en la célula la secuencia reguladora en *cis* está unida operativamente a un gen o genes, cuya expresión puede detectarse, directamente o indirectamente. Típicamente, la célula comprende un promotor que contiene la secuencia reguladora en *cis*, unido operativamente al gen o genes. Por unido operativamente se quiere decir que la secuencia reguladora en *cis* y/o el promotor está unido al gen o genes de manera tal que la secuencia y/o promotor puede funcionar (en condiciones apropiadas, por ejemplo presencia del factor o factores de transcripción requeridos) para regular la expresión del gen o genes. Así, cuando se une un factor de transcripción afín, la secuencia reguladora en *cis* funciona para regular (reprimir o activar) la expresión del gen o genes. El promotor puede ser por ejemplo uno que regula un gen o genes que codifican un fenotipo de interés, tal como cualquiera de los genes o fenotipos descritos en la presente memoria.

El funcionamiento de la secuencia reguladora en *cis* en la célula puede determinarse monitorizando la expresión del gen o genes unidos. Esto puede hacerse monitorizando la expresión del gen directamente o monitorizando la expresión de un fenotipo particular que está asociado con la expresión el gen o genes. Por ejemplo, el cribado para función puede comprender el cribado para la producción de un metabolito o producto de fermentación dado, por ejemplo, un antibiótico tal como cinamicina, actinorhodina y/o undecilprodigiosina; para resistencia a un antibiótico particular, por ejemplo vancomicina; o para tolerancia a un disolvente tal como butanol. Los métodos para el cribado se describen en la presente memoria.

En un aspecto, el plásmido causa un cambio en la expresión génica y/o fenotipo de una célula huésped, por ejemplo un incremento en la síntesis de antibiótico, una disminución en la resistencia a antibiótico o un incremento en la tolerancia a disolvente en la célula.

Una célula huésped puede comprender una secuencia reguladora en *cis* (por ejemplo, un promotor que contiene la secuencia) unida operativamente a su gen o genes nativos (afines), es decir, unida al gen o genes cuya expresión regula la secuencia (o promotor) en su estado nativo. Así, por ejemplo, cuando una secuencia reguladora en *cis* de interés (o su promotor nativo) regula un gen o genes que codifican la producción de un metabolito (por ejemplo, antibiótico o enzima), una célula huésped puede comprender una célula productora nativa. Cuando una secuencia reguladora en *cis* de interés (o su promotor nativo) regula un gen o genes que codifican resistencia a antibiótico, una célula huésped puede comprender una cepa resistente a antibiótico nativa, por ejemplo un patógeno o un aislado

clínico. Así, la secuencia reguladora en *cis* en la célula puede estar unida operativamente a un gen o genes, cuya expresión está regulada por esa secuencia en su contexto natural.

Por ejemplo, como se describe en los presentes Ejemplos, *S. coelicolor* es un productor nativo de actinorhodina y undecilprodigiosina. Se ha indicado que la producción de estos antibióticos requiere la unión de la proteína AsfR a un sitio de unión en 5' del gen *asfS*. *S. coelicolor* puede usarse por lo tanto como una célula huésped para cribar snare para la presencia de una secuencia que comprende el sitio de unión AsfR (o cuando se introduce una secuencia señuelo que compite con dicha secuencia). La competición con el sitio de unión AsfR celular nativo resultará en una producción reducida de antibiótico.

También como se describe en los Ejemplos, *S. cinnamoneous* produce el antibiótico cinamicina. *S. cinnamoneous* puede usarse como un huésped para cribar snare para la presencia de secuencias reguladoras en *cis* en los promotores de genes en la agrupación biosintética de cinamicina de *S. cinnamoneous* (o cuando se introduce un señuelo que compite con dicha secuencia). La competición con las secuencias reguladoras en *cis* nativas puede cribarse monitorizando la producción de cinamicina en cepas transformadas, como se describe en la presente memoria.

También como se describe, *E. coli* K12 comprende genes que codifican tolerancia a butanol. Las células de *E. coli* K12 pueden usarse para cribar snare para la presencia de secuencias reguladoras en *cis* en los promotores de genes que codifican la tolerancia (o cuando se usa un señuelo que compite con dicha secuencia). La presencia de una secuencia competidora en las células transformadas puede ensayarse determinando la viabilidad de las células en presencia de varias concentraciones de butanol.

Alternativamente, una célula huésped puede comprender un modelo, por ejemplo un modelo bacteriano, de un fenotipo celular dado, tal como producción de metabolito o resistencia a antibiótico. Dicho modelo comprende la secuencia reguladora en *cis* de interés (por ejemplo, un promotor que contiene la secuencia) unida operativamente al gen o genes nativos pero en una célula heteróloga. En otras palabras, la secuencia reguladora en *cis* (y su promotor nativo) y el gen o genes unidos operativamente no están presentes endógenamente en la célula.

Típicamente, una célula huésped modelo comprende una especie o cepa bacteriana que se manipula más fácilmente en condiciones de laboratorio que una cepa nativa, por ejemplo, no es patogénica o es más susceptible a transferencia génica. En general, en ausencia de un plásmido n[snare] o molécula señuelo descrita en la presente memoria, la expresión del gen o genes está regulada por la secuencia reguladora en *cis* (por ejemplo, en su promotor). Así, la célula comprende los componentes necesarios para que la secuencia reguladora en *cis* funcione de la manera normal, por ejemplo, el factor de transcripción se une a la secuencia y reprime o activa la expresión del gen o genes.

Una célula huésped modelo puede comprender un plásmido que presenta la secuencia reguladora en *cis* y el gen o genes, que se ha introducido en la célula o la secuencia reguladora en *cis* y el gen o genes pueden estar integrados en el genoma del huésped. Un huésped modelo puede comprender un fragmento genómico que comprende la secuencia reguladora en *cis* y el gen o genes, en el que el fragmento se ha adquirido de una célula nativa, por ejemplo por transferencia horizontal.

Una célula modelo puede haber adquirido genes que confieren un fenotipo de interés dado, por ejemplo, producción de metabolito, producción de antibiótico, resistencia a antibiótico, tolerancia a disolvente, por ejemplo, una cepa modelo puede haber adquirido un fenotipo particular por transferencia horizontal de genes que codifican ese fenotipo, por ejemplo, producción de antibiótico y/o resistencia a antibiótico. Dicha cepa modelo puede prepararse sin conocimiento previo de los genes codificadores. Esto se hace por transferencia horizontal de ADN de la cepa nativa (que presenta el fenotipo de interés) a la cepa modelo y cribando para obtención del fenotipo.

Cuando las células van a ser huésped de una biblioteca n[snare] personalizada, las células pueden comprender, por ejemplo, el ADN genómico o fragmento genómico del que se ha derivado la biblioteca (bien endógenamente o por adquisición génica).

Las células huésped adecuadas se conocen en la técnica y se describen en la presente memoria en los presentes Ejemplos. Por ejemplo, el Ejemplo 6 describe la cepa M600 de *S. coelicolor*, que es un modelo bacteriano de resistencia a vancomicina.

Células informadoras

Una célula huésped puede comprender una secuencia reguladora en *cis* de interés (o un promotor de interés que comprende secuencias reguladoras en *cis*) unida operativamente a un gen o genes cuya expresión no está regulada por esa secuencia o promotor en su contexto nativo (un gen no nativo). Dicho gen se refiere como un gen informador. Típicamente, el gen informador está unido operativamente a una secuencia promotora que comprende la secuencia reguladora en *cis* de interés.

La función de la secuencia reguladora en *cis* puede determinarse determinando la expresión del gen informador. Típicamente, el gen informador es uno cuya expresión puede detectarse y cribarse fácilmente. Por ejemplo, el gen

informador puede codificar un compuesto fluorescente, o resistencia a antibiótico. Los ejemplos de genes informadores son luciferasa, o el gen *neo* que codifica resistencia a kanamicina. Dichas células informadoras son particularmente útiles cuando las secuencias reguladoras en *cis* (o promotores) controlan nativamente la expresión del gen o genes que no tienen un fenotipo fácilmente evaluable.

En una realización, el gen informador está regulado negativamente por la unión de un factor de transcripción (represor transcripcional) a la secuencia reguladora en *cis* de interés. Dicha célula huésped es adecuada para cribar secuencias reguladoras en *cis* que regulan negativamente (reprimen) la expresión de su gen o genes afines. En una realización alternativa, el gen informador está regulado positivamente por la unión de un factor de transcripción (activador transcripcional) a la secuencia reguladora en *cis*. Dicha célula huésped es útil para cribar secuencias reguladoras en *cis* que regulan positivamente la expresión de su gen o genes afines.

En el sistema informador puede usarse cualquier cepa huésped adecuada. Las cepas adecuadas son conocidas en la técnica y se describen en la presente memoria. Cuando es apropiado, se elegirán cepas que no contienen copia genómica del gen informador. De manera similar, las cepas informadoras pueden obtenerse por ingeniería por la delección dirigida de genes usados como informadores del genoma antes de la introducción del gen informador verdadero.

El gen informador y la secuencia reguladora en *cis* (por ejemplo, en un promotor) pueden estar contenidos en un plásmido. Alternativamente, el gen informador y la secuencia reguladora en *cis* (o promotor) pueden estar integrados en el genoma del huésped. Por ejemplo, el sistema informador puede haberse introducido en el huésped en un plásmido integrativo, por ejemplo, pSET152. Cuando el sistema está contenido en un plásmido, típicamente el plásmido comprende un gen marcador detectable que permite la selección positiva para el plásmido en las células huésped. Por ejemplo, el gen marcador detectable puede codificar resistencia a antibiótico, por ejemplo, el gen *aac* que codifica resistencia a apramicina.

Los inventores han desarrollado un sistema de célula informadora genérico que puede usarse para cribar para la presencia de una secuencia reguladora en *cis*, por ejemplo, en una biblioteca de plásmidos *n[snare]* o plásmido *n[snare]* como se describe en al presente memoria. La célula y métodos de cribado permiten el cribado para la presencia de la secuencia reguladora en *cis* tomando como base la viabilidad celular. La expresión del gen informador es determinativa para la viabilidad o supervivencia celular en condiciones de cultivo apropiadas. En las condiciones del ensayo, las células huésped sólo son viables si se ha introducido en la célula una secuencia que compite con la secuencia reguladora en *cis* de interés (por ejemplo, en el promotor de interés) para unión del factor de transcripción. El sistema se refiere como un sistema de cribado "muerto o vivo". El sistema puede automatizarse y tiene la ventaja de que puede usarse para cribar células cultivadas en medio líquido.

Cuando la secuencia reguladora en *cis* de interés funciona para reprimir la expresión de un gen unido operativamente, el gen informador típicamente codifica un producto que es necesario para la supervivencia celular en condiciones de cultivo apropiadas. Por ejemplo, el gen informador puede codificar resistencia a antibiótico de manera que cuando la célula huésped se cultiva en presencia del antibiótico, la supervivencia celular sólo es posible en presencia de una secuencia competidora que inhibe la unión del factor de transcripción a la secuencia reguladora en *cis* unida al informador y permite la expresión del gen informador de antibiótico. Los genes de resistencia a antibiótico adecuados incluyen, por ejemplo, el gen resistente a kanamicina *neo*. Esto permite la detección de las secuencias reguladoras en *cis* que están unidas por represores transcripcionales. Otros genes adecuados se conocen en la técnica.

En este caso, una célula huésped típicamente comprende otros componentes que son necesarios para que funcione la secuencia reguladora en *cis*. La célula comprende típicamente (por ejemplo, expresa) el o los factores de transcripción necesarios para unirse a la secuencia reguladora en *cis*. La célula huésped puede comprender ya la secuencia reguladora en *cis*, por ejemplo, un promotor que comprende la secuencia reguladora en *cis*. La célula huésped puede comprender un fragmento genómico del que se deriva la secuencia reguladora en *cis* de interés.

Cuando la secuencia reguladora en *cis* funciona para activar la expresión de un gen unido operativamente, el gen informador codifica un producto que es letal para la célula en condiciones de cultivo apropiadas. Dicho gen se refiere habitualmente como un gen informador suicida. Las células sólo son viables (en condiciones de cultivo apropiadas) en presencia de una secuencia competidora que inhiba o evite que el factor de transcripción se una a la secuencia reguladora en *cis* unida al informador. Un ejemplo de un gen informador suicida es, por ejemplo, el gen de la glucosa quinasa (*glkA*). Este gen convierte un metabolito (2 desoxi glucosa (DOG)) en una toxina. Así, la expresión del gen *glkA* es letal para la célula cuando se cultiva en presencia de DOG. Esto permite la detección de secuencias reguladoras en *cis* que están unidas por activadores transcripcionales. Otros genes adecuados se conocen en la técnica.

Típicamente, en este sistema el huésped no expresa de otra manera el gen informador pero comprende los componentes necesarios para que la secuencia reguladora en *cis* funcione. La célula comprende típicamente, por ejemplo expresa, un o unos factores de transcripción necesarios para unirse a la o las secuencias reguladoras en *cis*. En dicho sistema, cualquier actividad en la secuencia reguladora en *cis* será debida a factores de transcripción positivos (activadores) derivados del genoma de la célula huésped. Mediante la búsqueda en este contexto, es

posible encontrar reguladores al alza de la expresión heteróloga del cromosoma del huésped. Mediante la liberación de estos activadores de la secuencia reguladora en *cis* unida al informador y en una secuencia señuelo, por ejemplo un plásmido n[snare], se evita la expresión del gen informador, por ejemplo *glkA*, y las células son viables incluso cuando se crecen en presencia del metabolito DOG.

El sistema informador muerto o vivo puede adaptarse para uso con cualquier secuencia reguladora en *cis* adecuada, tal como cualquiera de las descritas y/o identificadas en la presente memoria. Por ejemplo, el promotor *cin7*, derivado de la agrupación biosintética de cinamicina de *S. cinnamoneous* puede usarse como se describe en el Ejemplo 3. El promotor *cin7* puede estar unido operativamente al gen *neo* que codifica resistencia a kanamicina como se describe en los Ejemplos. Éste puede estar comprendido en un plásmido integrativo tal como el núcleo pSET152 usado en los ejemplos. El plásmido también porta un gen marcador adecuado (el gen *aac* que codifica resistencia a apramicina). El plásmido puede introducirse en o integrarse en una cepa huésped *S. lividans* (1326) que además porta la ruta biosintética de cinamicina. Dicha célula informadora es adecuada para cribar, por ejemplo, una biblioteca de plásmidos n[snare] derivada del genoma de *S. cinnamoneous* o fragmento genómico que comprende la agrupación biosintética de cinamicina como se describe en la presente memoria.

Alternativamente, el promotor *cin7* puede estar unido operativamente al gen informador *glkA* y de la misma manera integrarse en un plásmido integrativo. En este caso, una cepa huésped adecuada sería la cepa TK24 de *S. lividans* que carece de la agrupación biosintética de cinamicina y también carece del gen *glkA*.

En un aspecto, la invención se refiere a las células informadoras descritas y preparadas según los presentes Ejemplos.

Métodos para identificar y caracterizar secuencias reguladoras en *cis*

El principio del uso de los presentes plásmidos n[snare] se ilustra en la Figura 2. Si la secuencia snare en el plásmido comprende un sitio de unión de factor de transcripción que comprende y/o compite para la unión del factor de transcripción con, una secuencia reguladora en *cis* en la célula en la que el plásmido se introduce, esto puede detectarse como un cambio en la expresión del gen que está regulado por esa secuencia reguladora en *cis* en la célula. Así, los plásmidos n[snare] pueden usarse para identificar secuencias reguladoras en *cis* y señuelo. El plásmido también puede usarse con una snare que comprende una secuencia señuelo conocida, para interrumpir la expresión génica en la célula.

Así, los plásmidos n[snare] de la invención pueden usarse para varios propósitos diferentes. Como anteriormente, puede usarse una biblioteca de plásmidos n[snare] para cribar posibles secuencias reguladoras en *cis*, por ejemplo en un fragmento de ADN genómico usando una biblioteca n[snare] personalizada. Un plásmido n[snare] puede usarse para ensayar una secuencia particular para función señuelo. En cualquiera de estos métodos, una secuencia snare se identifica como que comprende una secuencia reguladora en *cis* o una secuencia que compite con una secuencia reguladora en *cis* para unión del factor de transcripción, si la snare en el plásmido n[snare] es capaz de titular el factor de transcripción de la secuencia reguladora en *cis* presente en el huésped (como se determina por una alteración en la expresión génica o fenotipo de la célula huésped). Así, las secuencias identificadas usando los métodos también pueden actuar como secuencias señuelo, y en un aspecto, los métodos pueden considerarse como métodos para identificar secuencias y moléculas señuelo. Un plásmido n[snare] que comprende una secuencia señuelo puede usarse para modular la expresión génica y/o el fenotipo en una célula huésped.

Así, en un aspecto, la invención proporciona métodos para identificar y/o caracterizar secuencias reguladoras en *cis* y señuelo. En general dicho método comprende:

1. proporcionar un plásmido n[snare] o una biblioteca de plásmidos n[snare] como se describe en la presente memoria;
2. introducir el plásmido o biblioteca de plásmidos n[snare] en una o unas células huésped en el que la célula huésped comprende la secuencia reguladora en *cis* de interés unida operativamente a un gen o genes, cuya expresión puede determinarse directamente o indirectamente; y
3. determinar la expresión del gen (o genes), o alteración del fenotipo, en presencia y ausencia del plásmido o la biblioteca de plásmidos n[snare].

Un plásmido n[snare] o biblioteca de plásmidos n[snare] para usarse en el método puede ser como se describe en la presente memoria. La snare en cada plásmido comprende la secuencia que se va a ensayar para función competidora.

En general, una célula huésped comprende un elemento regulador en *cis* de interés (por ejemplo, un promotor que comprende un elemento regulador en *cis*) unido operativamente a un gen o genes, cuya expresión puede monitorizarse o determinarse directamente o indirectamente. Por ejemplo, la expresión puede monitorizarse determinando la expresión del gen o genes particulares o monitorizando un fenotipo particular codificado por el gen o genes.

Las células huésped adecuadas para uso en los presentes métodos se han descrito en la presente memoria e incluyen, por ejemplo, cepas nativas, aislados clínicos, modelos de laboratorio y células informadoras. Puede usarse el sistema de cribado informador “muerto o vivo” descrito en la presente memoria.

En los Ejemplos 1 y 2 en la presente memoria se demuestra que la estrategia de oligonucleótido señuelo puede adaptarse a un formato de plásmido, incluyendo la confirmación de que se puede conseguir el uso de una biblioteca para identificar reguladores en *cis* de una secuencia definida. Una posible limitación de esta iteración de la estrategia basada en una biblioteca n[snare] es que puede no ser adecuada óptimamente para reguladores de cambios fenotípicos sutiles o complicados. En muchas circunstancias, la alteración fenotípica puede no ser detectable fácilmente y puede necesitar un sistema complicado de evaluación, tal como cuantificar la producción de metabolitos por análisis cromatográfico. Habría una ventaja obvia para la creación de un sistema informador en el que la evaluación del fenotipo estuviera estandarizada, lo que permitiría acelerar el cribado de la biblioteca n[snare] de manera que pueda realizarse con alto rendimiento.

Una posible limitación técnica de la estrategia adoptada en el Ejemplo 2, concretamente cribado de todos los miembros de una biblioteca para cambio de fenotipo, es que la necesidad de plaquar o cultivar físicamente todos los miembros de bibliotecas completas supone una limitación práctica en cómo pueden cribarse muchas colonias y por lo tanto la cantidad de ADN que puede evaluarse. No es factible evaluar fragmentos genómicos grandes que generarían bibliotecas con >100.000 miembros, lo que equivale a aproximadamente 25 kb de secuencia. Como se describe más adelante, los dos genes informadores usados en un sistema desarrollado por los inventores confieren viabilidad a las células que tienen un plásmido n[snare] que contiene un elemento regulador en *cis* candidato. Por lo tanto, la selección es por viabilidad lo que permite un enriquecimiento rápido de secuencias candidatas en cultivo líquido simplemente dejando morir a la mayoría de las células que no contienen un plásmido candidato.

Por lo tanto, los inventores han adaptado la estrategia de biblioteca n[snare] descrita en la presente memoria para abordar estos problemas desarrollando sistemas de ensayo basados en informador genérico. En esta iteración, los reguladores negativos y positivos se detectan por su efecto en genes informadores que codifican resistencia a antibiótico (Figura 7) o sensibilidad a metabolito (informadores suicida) como se muestra en la Figura 8. Los promotores diana se introducen en los casetes informadores para permitir la detección de regulación negativa (represores transcripcionales) o regulación positiva (activadores transcripcionales). Esta estrategia resuelve el problema de la necesidad de un fenotipo evaluable ya que como en circunstancias normales cada informador causa la muerte celular, las células sólo pueden crecer cuando un plásmido n[snare] de la biblioteca libera esto liberando el factor de transcripción. Por lo tanto, la selección de ambos informadores se basa en la supervivencia celular, lo que acelera en gran medida el proceso de cribado. Otra característica del sistema es que los cribados muerto-o-vivo pueden automatizarse fácilmente tomando como base los cribados rápidos, exhaustivos o de alto rendimiento. También es posible identificar secuencias reguladoras que controlan promotores que en su contexto natural no tienen un fenotipo que se pueda evaluar fácilmente.

El plásmido o biblioteca de plásmidos se introduce en células huésped en los presentes métodos por cualquier medio adecuado. Los medios adecuados se conocen en la técnica. Si el plásmido n[snare] es conjugativo, el plásmido puede introducirse por conjugación. Pueden usarse otros medios tales como transfección/transformación.

Típicamente, las células huésped se monitorizan para propagación estable del plásmido seleccionando para la expresión de un gen marcador detectable en el plásmido como se describe en la presente memoria, por ejemplo, un gen de resistencia a antibiótico. Las células se cultivan en condiciones que seleccionan células que expresan el marcador particular, por ejemplo, en presencia del antibiótico.

Preferiblemente, el método comprende el uso de uno o más controles adecuados. Por ejemplo, dichos controles incluyen células huésped no tratadas con el plásmido, células huésped en las que se ha introducido un vector plasmídico vacío y células huésped en las que se ha introducido un plásmido n[snare] que comprende una posible secuencia reguladora en *cis* desordenada.

Una vez los plásmidos se han introducido en las células, las células se criban para la expresión del gen o genes unidos operativamente a la secuencia regulador en *cis* (o promotor) de interés. La expresión en presencia de la secuencia snare (en el plásmido n[snare]) se compara con la expresión en ausencia de la snare. Los clones que producen una alteración en la expresión del gen o genes de ensayo (por ejemplo, una alteración del fenotipo celular) se seleccionan ya que probablemente comprenderán una secuencia reguladora en *cis*. El ADN de estos clones se aísla y la posible secuencia reguladora en *cis* (o secuencia señuelo) se aísla.

El cribado para una alteración en la expresión génica puede realizarse por cualquier método adecuado. El cribado puede comprender detectar o medir el producto de expresión del gen o genes, por ejemplo, ensayando para la función del producto génico o puede comprender determinar un cambio en un fenotipo de la célula huésped asociado con la expresión del gen o genes.

Por ejemplo, los cambios en la expresión de un gen o genes que codifican la producción de un metabolito, por ejemplo un antibiótico, puede monitorizarse determinando la expresión del metabolito dado, por ejemplo, la cantidad (o presencia o ausencia) del metabolito o función del metabolito.

Por ejemplo, la actinorhodina y undecilprodigiosina son antibióticos pigmentados. La actinorhodina es azul y la undecilprodigiosina es roja. Así, la expresión de estos antibióticos puede monitorizarse fácilmente por técnicas colorimétricas.

5 La producción de algunos metabolitos o antibióticos puede monitorizarse usando cepas bacterianas indicadoras y ensayos en placa adecuados. Por ejemplo, la producción de cinamicina puede detectarse por un ensayo en placa usando la cepa indicadora *Bacillus subtilis*, y medios adecuados, por ejemplo, agar sólido R2YE. Las células se cultivan típicamente en placas de agar que se han sembrado con la cepa indicadora. La expresión del antibiótico se determina por el grado de reacción, por ejemplo, muerte de la cepa indicadora, por ejemplo, por el diámetro de un halo en una placa de agar.

10 La expresión del gen o genes de resistencia a antibiótico o fenotipo, por ejemplo, resistencia a vancomicina, resistencia a kanamicina, puede determinarse cultivando células huésped en presencia del antibiótico y determinando la sensibilidad al antibiótico. Típicamente, las células también se cultivan en ausencia del antibiótico como un control.

15 De manera similar, los genes que codifican tolerancia a disolvente, por ejemplo, tolerancia a butanol, pueden monitorizarse cultivando las células en presencia del disolvente (de nuevo, típicamente se lleva a cabo un control en el que las células se cultivan en ausencia del disolvente).

20 En algunos casos, el cribado para la expresión del gen o genes relevantes comprende determinar la viabilidad de las células huésped en condiciones de cultivo adecuadas. Por ejemplo, cuando el gen o genes codifican resistencia a antibiótico, el cribado puede comprender cultivar las células en presencia del antibiótico y determinar si las células son viables. En un cribado que comprende las células huésped informadoras "muerto o vivo", en las condiciones de cultivo dadas, las células sólo son viables si el plásmido n[snare] introducido en las células comprende una secuencia que puede competir con la secuencia reguladora en *cis* de interés. Así, el cribado de las células comprende cultivar las células en condiciones en las que la expresión del gen o genes informadores (el gen o genes unidos operativamente a la secuencia reguladora en *cis* de interés) es determinativa para la viabilidad de las células huésped y aislar las células viables.

25 Los presentes métodos pueden comprender cultivar las células huésped en medios líquidos, como se describe en la presente memoria, por ejemplo, si el fenotipo celular que se está determinado es la viabilidad celular. Esto puede tener la ventaja de que sólo hay que analizar un pequeño número de células. El cribado puede comprender el uso de técnicas de sustracción de ADN como se describe en la presente memoria. Esto puede ser particularmente útil cuando la expresión del gen o genes relevantes afecta la viabilidad celular (en condiciones de cultivo adecuadas).

30 Típicamente, una etapa de sustracción de ADN comprende sustraer una población de ADN de las células con un fenotipo dado y de las células sin el fenotipo.

35 Por ejemplo, en algunos casos, la introducción de una secuencia que comprende una secuencia reguladora en *cis* competidora, resultará en células huésped que se vuelven no viables en condiciones de cultivo adecuadas. Este puede ser el caso si, por ejemplo, la secuencia reguladora en *cis* de interés activa la expresión de un gen o genes de resistencia a antibiótico o reprime la expresión de un gen letal (producto de expresión letal en condiciones de cultivo particulares). Cuando una secuencia reguladora en *cis* competidora se introduce en las células, por ejemplo en una snare de un plásmido n[snare], la secuencia competidora libera el factor de transcripción y resulta en la expresión reducida del gen de resistencia a antibiótico o la expresión del gen de letalidad.

40 Las células transformadas se cultivan típicamente (a) en condiciones en las que las células con expresión interrumpida de la secuencia reguladora en *cis* serán no viables; y (b) en condiciones en las que las células serán viables. Las poblaciones de ADN se aíslan a partir de los dos cultivos (típicamente después del aislamiento y/o amplificación del ADN introducido, por ejemplo, el plásmido n[snare] o inserto del plásmido) y se sustraen (por ejemplo, por hibridación). Es posible entonces determinar el ADN que falta de las células en el cultivo (a) y que comprende probablemente la secuencia reguladora en *cis* competidora.

45 Por ejemplo, si el fenotipo que se está investigando es resistencia a antibiótico, una secuencia reguladora en *cis* de interés introducida en una célula huésped, por ejemplo, en una snare de un plásmido n[snare], puede restaurar la sensibilidad al antibiótico. Las células transformadas se cultivan (a) en presencia de antibiótico; y (b) en ausencia de antibiótico. Las células que contienen la secuencia reguladora en *cis* de interés morirán en presencia del antibiótico. 50 El ADN se aísla de ambos cultivos y el ácido nucleico que comprende las secuencias snare se aísla, por ejemplo, por amplificación con PCR. Mediante la sustracción de las poblaciones de secuencias snare, es posible aislar aquellas que faltan de las muestras tratadas con antibiótico.

En un aspecto, la población enriquecida de secuencias snare puede reclonarse y repetirse el proceso de selección.

55 Una vez las células que presentan expresión génica o fenotipo alterado se han aislado, el ADN de estas células se aísla. Típicamente, el ADN que se ha introducido en las células, por ejemplo, el ADN plasmídico n[snare] y/o el inserto del plásmido, se aísla. La snare que causó la expresión alterada y que comprende la probable secuencia reguladora en *cis* (o señuelo) puede determinarse. Esto puede hacerse, por ejemplo, por amplificación con PCR.

Cuando el presente método comprende el uso de una biblioteca de plásmidos n[snare], el método puede comprender además una o más etapas adicionales.

Por ejemplo, el método puede comprender una etapa de hibridación de la biblioteca. Debido a la complejidad de las bibliotecas usadas en estas clases de procedimientos de enriquecimiento es raro generar muestras que consistan completamente en las secuencias deseadas. Generalmente, hay un fondo de falsos negativos en la muestra enriquecida que hay rechazar y no arrastrar para análisis adicional. La estrategia de hibridación de la biblioteca se diseña para hacer esto detectando plásmidos n[snare] que son comunes a todas las repeticiones independientes del proceso de cribado, siendo la lógica que si los plásmidos que portan la misma secuencia se detectan en repeticiones independientes con apariciones por encima del fondo, éstos deberían ser los que se arrastren para análisis adicional. Una visión global del proceso se proporciona en la Figura 9.

Típicamente, cuando se usa la estrategia de hibridación de la biblioteca, se realizan x repeticiones independientes del método de cribado anterior (que comprende las etapas (1), (2) y (3)). Por ejemplo, x puede ser 2, 3, 4, 5, 6, 7 u 8.

En cada repetición, los clones que producen una alteración en la expresión génica/fenotipo se seleccionan como se describe. El ADN plasmídico se aísla de los clones y se marca para crear una muestra de sonda combinada (P). Por ejemplo, el ADN plasmídico puede marcarse por PCR de cebado aleatorio, por ejemplo, usando una molécula dUTP marcada con DIG como en los presentes Ejemplos.

Los clones también se plaquean en medios adecuados a una concentración que permite distinguir a las colonias individuales. Típicamente, en esta etapa, se toma una muestra de cada colonia y se cultiva adicionalmente para proporcionar una fuente de ADN plasmídico si es necesario en una etapa posterior.

El ADN total se extrae de cada colonia y se inmoviliza en una matriz adecuada, por ejemplo, una membrana de nilón, por ejemplo por hibridación. El ADN de cada colonia está en una posición localizable en la matriz. Cada matriz se hibrida separadamente con cada uno de los conjuntos de sondas. Las muestras que son comunes a más de una repetición se seleccionan típicamente para análisis adicional.

Así, por ejemplo, si $x = 4$, se preparan cuatro conjuntos de sondas y cuatro matrices, por ejemplo, 4 membranas de nilón. Cada uno de los 4 conjuntos de sondas, P1, P2, P3 y P4 se hibrida con cada uno de los 4 filtros, F1, F2, F3 y F4. La hibridación de P1 con F1 produce un cromatograma en el que cada muestra se detecta. Las muestras F2, F3 o F4 que hibridan con P1 serán candidatos potenciales para análisis adicional. Cuantas más muestras de sonda hibriden con una colonia, más fuerte es esa colonia como candidata.

La auto-hibridación (por ejemplo, P1 con F1) también puede usarse para detectar falsos positivos en un cribado.

El método de la invención puede comprender además repetir el proceso de selección una o más veces, de manera que la selección es iterativa. Así, los clones seleccionados de un primer cribado pueden usarse para transformar las células huésped de nuevo y el cribado se repite.

La estrategia de biblioteca n[snare] descrita en la presente memoria podría usarse para identificar secuencias reguladoras en *cis* que influyen en los fenotipos bacterianos con consecuencias médicas importantes, siendo un ejemplo la prevención de la inducción de mecanismos de resistencia a antibiótico en bacterias patógenas que infectan a los seres humanos (Figura 6). El esquema en esta figura muestra cómo una biblioteca de plásmidos n[snare] se usa para identificar secuencias reguladoras en *cis* que controlan la resistencia a antibiótico cuando no se conoce nada acerca de la ruta genética del mecanismo. En este ejemplo, los genes que confieren resistencia a antibiótico se trasladan a un huésped bacteriano conveniente, tal como *S. coelicolor* o *E. coli*, por transferencia génica horizontal para crear un modelo bacteriano conveniente. Esto puede hacerse sin conocimiento previo de cuáles son los genes cribando para la obtención de resistencia a antibiótico. Un mecanismo similar habrá trasladado los mismos genes de resistencia en las cepas patógenas detectadas en clínica. Las bibliotecas de plásmidos n[snare] se crean a partir del fragmento genómico que porta los genes de resistencia, el genoma completo del aislado naturalmente resistente o clínico, o de oligonucleótidos aleatorios para crear bibliotecas universales que contienen posiblemente cada posible secuencia reguladora en *cis*. Dichas bibliotecas se introducen en el modelo bacteriano y el transformante se criba para susceptibilidad incrementada al antibiótico diana. Esto se consigue bien por evaluación directa del fenotipo o usando técnicas de sustracción de ADN establecidas. Dependiendo de la complejidad del sistema, el proceso se realiza de forma única o iterativa para identificar una única secuencia reguladora en *cis* o mezcla de dichas secuencias, que se usa entonces para sintetizar o fabricar TFD correspondientes. Éstos se validan en aislados resistentes o clínicos, antes de proceder a un modelo animal apropiado, tal como un modelo de ratón, en el que la eficacia de los señuelos se ensaya tratando un animal infectado con una cepa patógena que es resistente al tratamiento con antibiótico solo.

Como se ha indicado, la biología del genoma, el patrón de la expresión génica y tiempo de la replicación, está controlado principalmente por interacciones ADN-proteína, y el mapeo de estas interacciones puede ser un pre-requisito para los intentos de controlar los genes. El análisis con micromatrices puede producir un informe de genomas completos para identificar qué genes se expresan, pero hasta la fecha no existe ninguna tecnología para identificar, en una escala similar, las proteínas (reguladores en *trans*) y sus sitios de unión afines (elementos reguladores en *cis*) que determinan el patrón de la expresión. El Ejemplo 4 de la presente memoria y la realización

de la invención descrita en la presente memoria abordan esta deficiencia y desarrollan nuestra tecnología actual para formar una herramienta genérica capaz de identificar rápidamente los elementos reguladores en *cis* a lo largo del genoma, para delinear redes genéticas y confirmar el control sobre ellas. El método según este aspecto de la invención identifica elementos reguladores a lo largo del genoma que controlan la expresión de genes diana. Como un sistema rápido y genérico para delinear redes reguladoras, la herramienta tiene el potencial de producir datos de estándar de oro para apoyar el direccionamiento hacia biología de sistemas y la utilidad en la definición de regiones para ingeniería genética basada en conocimiento y acelerar el mapeo de la variación genética que causa enfermedades.

En algunos casos, existe un conocimiento *a priori* de la maquinaria genética subyacente al fenotipo diana, por ejemplo, promotores conocidos en la agrupación de la cinammina. Más comúnmente, se conoce poca información acerca de las redes genéticas y su regulación subyacente a los fenotipos diana. Por ejemplo, muchos mecanismos de resistencia a antibiótico entre las infecciones patógenas que aparecen en clínica no se entienden al nivel genético y este hecho presenta una barrera importante para el desarrollo de tratamientos para resolver este problema. Asimismo, en el contexto de la biotecnología industrial, una técnica que podría alterar favorablemente características complejas de bacterias sin la necesidad de conocimiento previo de determinantes genéticos sería valiosa. Un ejemplo de esto, sería modificar por ingeniería una bacteria para que sea tolerante a disolvente para mejorar su eficacia en la fermentación de azúcares en el proceso de producción de biocombustible. Los ejemplos de aplicaciones clínicas e industriales de la presente tecnología se proporcionan en la presente memoria.

Para demostrar la eficacia de esta estrategia para detectar elementos reguladores en *cis* sin conocimiento previo de una ruta genética, los inventores usaron bibliotecas n[snare] para incrementar la tolerancia a butanol en *Escherichia coli* (Ejemplo 4). Una de las bibliotecas se derivó del genoma completo de *E. coli* usando técnicas similares a las descritas en el Ejemplo 2. La segunda biblioteca fue una biblioteca "n[snare] universal" que consiste posiblemente en copias directas de cada posible secuencia de 9 nucleótidos, o con cualquier longitud deseada. La biblioteca n[snare] universal se hace usando métodos similares a los descritos en el Ejemplo 2 para la creación de un único plásmido n[snare] AfsR con la excepción de que en lugar de tener un oligonucleótido secuenciado que contiene una sección central con una secuencia reguladora en *cis* definida, se inserta una secuencia aleatoria en su lugar (y en esta etapa la longitud de la secuencia puede controlarse). El procedimiento para usar dichas bibliotecas n[snare] para detectar nuevos reguladores en *cis* de tolerancia a butanol es similar a la descrita en el Ejemplo 3 en el sentido de que el fenotipo diana se evalúa midiendo la viabilidad frente a concentraciones crecientes de disolvente.

Las bibliotecas de plásmidos n[snare] como se describen y/o preparan según los métodos de la presente memoria, también pueden usarse para identificar secuencias reguladoras en *cis* cribando la biblioteca directamente para unión a uno o más factores de transcripción.

Los presentes métodos para identificar y caracterizar elementos reguladores en *cis* pueden comprender además el uso de un procedimiento de mapeo usando una combinación de obtención de impronta y exonucleasa como se describe en la presente memoria.

Métodos para mapear límites de secuencias reguladoras en cis

En un aspecto, la invención proporciona un método para mapear los límites de los sitios de unión a proteínas en el ADN. Este método puede usarse para mapear más precisamente los límites de dichos sitios de unión.

El método puede usarse para mapear los límites de secuencias reguladoras en *cis*. El método puede usarse en combinación con el método descrito en la presente memoria para identificar y caracterizar secuencias reguladoras en *cis*, por ejemplo, un método usando un plásmido n[snare] o una biblioteca de plásmidos como se describe en la presente memoria. Proporcionando una información de secuencia más precisa acerca de los elementos reguladores en *cis* y así de los sitios de unión para reguladores proteicos tales como factores de transcripción, el método es útil para diseñar oligonucleótidos señuelo que, cuando se introducen en una célula adecuada, competirán con una secuencia reguladora en *cis* en la célula para unión del regulador proteico.

El presente método usa una combinación de una enzima o agente químico con capacidad de mellado de ADN no específica (por ejemplo, ADNasa I) y una enzima exonucleasa 5'-3' (por ejemplo, exonucleasa T7) en un protocolo de obtención de impronta modificado. Las alternativas a la ADNasa I incluirían el tratamiento de las células con permanganato de potasio o radicales libres generados por hidróxidos. Otro ejemplo de una exonucleasa adecuada que podría usarse sería la exonucleasa lambda. El protocolo se diseña para mapear los límites 5' del complejo proteína-ADN en cada cadena del ADN y definir por lo tanto la región protegida del ADN en el complejo, que comprende el elemento regulador en *cis*. La ADNasa I introduce mellas en el ADN alrededor de los complejos. Estas sirven como sustratos para la actividad exonucleasa 5'-3' de la enzima exonucleasa 5'-3'. La acción combinada de las dos enzimas demarca los límites 5' del complejo proteína-ADN en cada cadena de ADN. Así, el presente método tiene la ventaja de que puede detectar todos los límites de los complejos ADN-proteína en una región dada, por ejemplo, una región promotora, no sólo los más cercanos a los sitios de restricción.

El principio del presente método de mapeo se ilustra en la Figura 13. Típicamente, el método comprende:

1. proporcionar un complejo proteína-ADN;

2. realizar una digestión con

a. una enzima que tiene capacidad de mellado de ADN no específica, por ejemplo, ADNasaI,

b. una exonucleasa 5'-3' (exonucleasa T7);

5 c. opcionalmente, una enzima de restricción que corta a una distancia corta en 5' de la posición probable del complejo ADN-proteína; en el caso del gen *actIIorf4* se determinó que ésta es *SacI*; y

3. determinar la posición de las deleciones 5' generadas en cada cadena de ADN en (2) respecto a un punto fijo conocido en la cadena de ADN.

10 El complejo proteína-ADN puede ser cualquier complejo de interés. En un ejemplo, el método se usa para mapear los límites de los complejos proteína-ADN en una región promotora de ADN, por ejemplo, de complejos proteína-ADN en secuencias reguladoras en *cis*. En un ejemplo, el complejo proteína-ADN puede ser un complejo factor de transcripción-ADN que regula la expresión de uno o más genes de resistencia a antibiótico en un procarionte o eucariota. En otro ejemplo, el método puede usarse para mapear sitios de unión proteína-ADN en el promotor *actIIorf4* de *S. coelicolor* como en los presentes Ejemplos.

15 El mapeo puede realizarse durante un estado transcripcional de interés, por ejemplo, durante la represión si el elemento regulador en *cis* de interés está implicado en la represión, es decir, está unido por un represor transcripcional, o durante la expresión activa desde un promotor si la secuencia reguladora en *cis* de interés está unida por un factor de transcripción.

20 Así, por ejemplo, en un aspecto, el mapeo puede realizarse *in vivo* (por ejemplo, usando células recién recogidas) en una etapa en la que se sabe que la expresión génica en las células está reprimida o activada. Típicamente, las células se cultivan y aíslan en la etapa transcripcional apropiada. Las células pueden permeabilizarse con detergentes que permiten que las enzimas entren en las células. Así, si los complejos proteína-ADN que se van a mapear funcionan para reprimir la expresión de un gen o genes dados o fenotipo, los complejos proteína-ADN se ensayan en un punto en el que la expresión del gen o genes o fenotipo está reprimida. Al contrario, si los complejos proteína-ADN que se van a mapear funcionan para activar la expresión de un gen o genes dados o fenotipo, los complejos proteína-ADN se ensayan en un punto en el que la expresión del gen o genes o fenotipo es activa. Típicamente, las células pueden monitorizarse para la expresión del gen o genes dados o fenotipo y aislarse en la etapa apropiada.

25 Por ejemplo, si la secuencia reguladora en *cis* de interés (en el complejo proteína-ADN) tiene un papel en la represión de la producción de antibiótico y se sabe que la producción de antibiótico ocurre tardíamente en el crecimiento celular, las células productoras se cultivan y recogen antes de esta etapa tardía del crecimiento.

30 Las células pueden ser procariotas o eucariotas. En un ejemplo, las células son procariotas, por ejemplo, células bacterianas. Por ejemplo, un actinomicete tal como una especie de *streptomyces*, por ejemplo, puede usarse *S. coelicolor*, por ejemplo, *S. coelicolor* A3 (2), (cepa M145 o M600), *S. lividans*, *S. cinnamomeus* o *E. coli*.

35 Típicamente, la cantidad de una enzima no específica (ADNasaI) y exonucleasa (por ejemplo, T7) que se va a usar en el método se determina empíricamente, como se describe en la presente memoria en los presentes Ejemplos.

El uso de una enzima de restricción (c) que corta en un sitio de restricción en 5' de la posición probable del complejo ADN-proteína produce un extremo 5' estándar para todos los complejos, con el que pueden hibridarse oligonucleótidos en las etapas posteriores de captura y amplificación.

40 Una vez ha terminado la digestión en (2), el ácido nucleico digerido se recupera. La digestión crea deleciones 5' en cada cadena de ADN. Mediante la determinación de la posición de estas deleciones 5' en cada cadena, por ejemplo, respecto a un punto fijo conocido en la cadena, es posible mapear de manera precisa el sitio de unión de la proteína en el ADN.

45 Una manera para determinar la posición de las deleciones 5' en cada cadena es como sigue. Una vez recuperado, el ADN digerido se desnaturaliza bien por calentamiento o tratamiento con disoluciones básicas, tales como 1M NaOH y se hibrida con una cadena de ADN complementario que comprende el sitio o sitios de unión de interés. Por ejemplo, el ADN complementario puede comprender un fragmento de un promotor (que contiene la o las secuencias reguladoras en *cis* de interés). Puede usarse un fragmento de PCR del promotor. En general, la cadena de ADN comprende un conector en un extremo. Puede realizarse una reacción de amplificación, por ejemplo, PCR, usando un cebador marcado que se une al promotor o conector.

50 Los tamaños de los productos de amplificación marcados se determinan, por ejemplo, por PAGE, u otros métodos tales como electroforesis capilar. En general, aunque los límites precisos del complejo proteína-ADN no se conocen, la posición aproximada, y por lo tanto la secuencia de la región de ADN que comprende el sitio de unión de la proteína de interés puede determinarse a partir de estos datos. Por ejemplo, como la secuencia de la región promotora se conocerá, una comparación del tamaño de los fragmentos marcados, y la posición del sitio de unión del

cebador con la secuencia de la región de ADN que comprende el o los sitios de unión de la proteína, hará posible determinar la posición precisa en la secuencia de ADN del límite 5' del complejo de unión proteína-ADN.

Los límites en la cadena de ADN opuesta se mapean de la misma manera, pero usando una cadena de ADN con un conector en el extremo opuesto.

- 5 Típicamente, la cadena de ADN complementaria se inmoviliza. Por ejemplo, el conector puede permitir la inmovilización en una matriz sólida. Por ejemplo, el conector puede estar biotinilado para inmovilización en una matriz de estreptavidina. La inmovilización tiene la ventaja de que las cadenas de ADN digeridas de interés pueden aislarse fácilmente a partir de la muestra de ADN digerida total.

- 10 En un aspecto, la invención se refiere a un método para identificar un regulador que actúa en *cis* de la expresión génica de un gen procariota o eucariota, que comprende uno o ambos de:

(a) realizar un mapeo de las secuencias de ácido nucleico protegidas;

- 15 (b) proporcionar una biblioteca de moléculas n[snare] en la que dicha biblioteca contiene secuencias que representan todas las posibles secuencias reguladoras del genoma de dicho procariota o dicho eucariota, y bien (i) identificar factores que se unen a dicha biblioteca o (ii) introducir dicha biblioteca de moléculas n[snare] en un organismo que puede indicar la activación o supresión diferencial de un gen diana comparado con cuando dichas moléculas n[snare] no se introducen. El método puede realizarse *in vivo*. Cuando el método es *in vivo*, puede realizarse con una bacteria y puede comprender poner en contacto la bacteria con cantidades eficaces de ADNasa I y exonucleasa T7 de manera que las secuencias reguladoras protegidas por los factores de transcripción permanezcan intactas mientras que el resto del genoma de dicha bacteria se destruye.

- 20 Una vez identificadas según los métodos de la presente memoria, las secuencias reguladoras en *cis* pueden usarse en ensayos de cribado para identificar factores de transcripción.

Métodos para modular la expresión génica y/o el fenotipo

- 25 Los plásmidos n[snare] y los métodos de la presente memoria pueden usarse para identificar y caracterizar secuencias que compiten con una secuencia reguladora en *cis* dada en una célula para unión al factor de transcripción afín.

- 30 Dichas secuencias pueden usarse para preparar secuencias señuelo. Una secuencia señuelo mimetiza el sitio de unión nativo (secuencia reguladora en *cis*) para una proteína reguladora (por ejemplo, factor de transcripción). Cuando se introduce en células huésped adecuadas que comprenden la secuencia reguladora en *cis* (por un método descrito en la presente memoria o de otra manera), la secuencia señuelo compete con la secuencia reguladora en *cis* en la célula para unirse al factor de transcripción afín.

Cuando dicha competición se produce, existe una alteración simultánea de la expresión de un gen o genes cuya expresión está regulada por la secuencia reguladora en *cis*. Esto puede causar una modulación del fenotipo celular, por ejemplo, producción de antibiótico, resistencia a antibiótico, tolerancia a disolvente, como se describe en la presente memoria.

- 35 Se apreciará que un plásmido n[snare] que causa una alteración en la expresión génica o fenotipo según los métodos descritos en la presente memoria puede usarse como una molécula señuelo en los presentes métodos. Una secuencia snare identificada como competidora con una secuencia reguladora en *cis* para unión al factor de transcripción según los métodos descritos en la presente memoria puede usarse como una secuencia señuelo

- 40 De acuerdo con esto, en un aspecto, la invención se refiere a métodos para modular la expresión génica y o fenotipo en una célula, que comprenden el uso de secuencias señuelo.

En general, un método para modular la expresión de un gen o genes según la invención comprende:

(a) proporcionar un polinucleótido que comprende un sitio de unión para un factor de transcripción (una secuencia señuelo); e

- 45 (b) introducir el polinucleótido en una célula, en el que la célula comprende el gen o genes unidos operativamente a una secuencia reguladora en *cis* que comprende el sitio de unión para el factor de transcripción o que compete con el sitio de unión del factor de transcripción para unión del factor de transcripción.

Generalmente, la secuencia señuelo (sitio de unión del factor de transcripción) en el polinucleótido no está unida operativamente a un gen. El sitio de unión del factor de transcripción puede aislarse de cualesquiera otros elementos de un promotor afín.

- 50 El polinucleótido que comprende la secuencia señuelo puede referirse como un polinucleótido señuelo.

El polinucleótido señuelo puede comprender un vector plasmídico. Por ejemplo, el polinucleótido señuelo puede comprender un plásmido n[snare] como se describe en la presente memoria y/o prepararse según un método descrito en la presente memoria.

5 El polinucleótido señuelo puede comprender más de una copia de la secuencia señuelo. El polinucleótido puede comprender una molécula multimérica que comprende múltiples copias de la secuencia señuelo. Por ejemplo de 1 a 1.000 copias. Típicamente, hay dos o más copias, por ejemplo, 2-1.000 copias, por ejemplo, al menos 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800 ó 900 copias. Por ejemplo puede haber 10-200, 10-150, 20-120, 20-100, 30-100, 30-80, 30-50, 30-40 copias. Por ejemplo, puede haber 30 copias de la secuencia señuelo. Típicamente, hay múltiples copias del señuelo, por ejemplo, múltiples repeticiones directas de la secuencia señuelo.

10 El polinucleótido señuelo puede comprender una secuencia adicional además de la secuencia señuelo. Típicamente, la secuencia adicional resulta en una resistencia incrementada a la degradación de la secuencia señuelo debido a la acción de exo y/o endonucleasas. El polinucleótido señuelo puede comprender al menos un elemento de estructura secundaria. Típicamente, esta estructura secundaria resulta en resistencia incrementada a la degradación de la secuencia señuelo debido a la acción de exo y/o endonucleasas. El polinucleótido señuelo puede comprender bases o azúcares modificados para conferir resistencia mayor a nucleasa. El polinucleótido señuelo puede comprender nucleótidos 2' OH o aminas en el extremo del polinucleótido para reducir o inhibir la actividad exonucleasa. En un aspecto, el polinucleótido señuelo puede comprender un oligonucleótido lineal. El polinucleótido señuelo puede comprender ADN bicatenario circular (una estructura denominada en forma de mancuerna). En un aspecto, el polinucleótido señuelo puede comprender una modificación de colesterol en uno o en cada uno de los extremos 5' de la molécula.

El polinucleótido señuelo puede comprender una cualquiera o más de las características anteriores en cualquier combinación adecuada.

25 Cuando se introduce en una célula huésped adecuada, una secuencia señuelo en un polinucleótido señuelo es capaz de competir con una secuencia reguladora en *cis* en la célula para unión del factor de transcripción que se une a la secuencia reguladora en *cis* endógena. Las secuencias pueden cribarse para esta función por los métodos descritos en la presente memoria. La secuencia reguladora en *cis* con la que compite la secuencia señuelo puede ser cualquiera de las secuencias reguladoras en *cis* descritas en la presente memoria y/o identificadas según los métodos de la presente memoria.

30 Por ejemplo, la secuencia reguladora en *cis* puede ser una que regula la expresión de un gen o genes que tienen un papel en la producción de metabolitos, por ejemplo, cuando el metabolito es un antibiótico (por ejemplo, actinorhodina, undecilprodigiosina o cinamicina), enzima o compuesto farmacéutico. La producción de metabolitos puede estar incrementada en la célula.

35 La secuencia reguladora en *cis* puede ser una que regula la expresión de un gen o genes que tienen un papel en determinar la resistencia a antibióticos. La resistencia a antibióticos en la célula puede estar disminuida. La secuencia reguladora en *cis* puede ser una que regula la expresión de un gen o genes que tienen un papel en determinar la tolerancia a disolventes (por ejemplo, butanol). La tolerancia a disolventes en la célula puede estar incrementada.

40 Por ejemplo, la secuencia reguladora en *cis* puede comprender una secuencia que codifica un sitio de unión para la proteína AfsR (un regulador pleyotrópico de la síntesis de antibióticos en *S. coelicolor*). Un sitio de unión de AfsR se muestra en negrita en SEQ ID NO: 1 (Ejemplo 1).

En otro ejemplo, una secuencia reguladora en *cis* puede ser una en el promotor *actII-orf4* de *S. coelicolor*, por ejemplo, una secuencia reguladora en *cis* represora. Por ejemplo, una secuencia reguladora en *cis* puede comprender la secuencia A24.1, A24.2, A24.3, A24.4 o A24.5 identificadas en la presente memoria.

45 En un ejemplo adicional, una secuencia reguladora en *cis* puede comprender un sitio de unión para el factor de transcripción VanR, por ejemplo, un sitio de unión de VanR localizado en el promotor *vanH* de, por ejemplo, *Enterococcus faecium* o *S. coelicolor*. Un ejemplo de un sitio de unión de VanR de 30 pb se muestra en SEQ ID NO: 21 y un ejemplo adicional está en SEQ ID NO: 26.

50 Una secuencia reguladora en *cis* puede ser una que funcione para regular la expresión de un gen o genes que codifican una función o fenotipo particular de interés, por ejemplo, la producción de un metabolito o metabolitos tal como un antibiótico (por ejemplo, cinamicina, actinorhodina, undecilprodigiosina), tolerancia a un disolvente(s) o toxina(s) particulares, resistencia a un antibiótico(s) particulares o cualquier otra función o fenotipo de interés. Típicamente, la secuencia estará localizada en la región promotora de dicho gen o genes.

Por ejemplo, una secuencia reguladora en *cis* puede regular la expresión de los genes biosintéticos de cinammina en *S. cinnamomeous*, por ejemplo, una secuencia reguladora en *cis* del promotor cin7.

Una secuencia reguladora en *cis* puede regular la expresión de un gen o genes que codifican tolerancia a disolvente, por ejemplo, en un procarionte, por ejemplo, tolerancia a butanol en *E. coli* K12.

- 5 Una secuencia reguladora en *cis* puede ser una que regula la expresión de un gen o genes procariontes de resistencia a antibióticos, típicamente una secuencia reguladora en *cis* del promotor de un gen o genes que codifican resistencia a antibióticos. Cualquier resistencia a antibióticos de interés puede considerarse. Por ejemplo, resistencia a antibióticos a: la clase de antibióticos conocida como aminoglicósidos (tal como kanamicina); los carbapenemos (tal como meropenem); las cefalosporinas (tal como cefepime); los glicopéptidos (tal como vancomicina y daptomicina);
10 las penicilinas tal como ampicilina, carbenicilina y penicilina; los antibióticos polipeptídicos (tal como polimixina B); las quinolonas (tal como levofloxacina); las sulfonamidas (tal como Bactrim); las tetraciclinas (tal como tetraciclina); y de forma diversa, cloranfenicol, rifampicina y Zyvox.

- 15 En un aspecto, un gen o genes que codifican resistencia a un antibiótico codifican proteínas que proporcionan resistencia a un antibiótico específico o, en algunos casos, a una clase de antibióticos, tal como los antibióticos y clases listados anteriormente. Dichos genes de resistencia a antibióticos pueden codificar proteínas dirigidas a un mecanismo o estructura específico del antibiótico o clase de antibióticos. Habitualmente, dichos genes de resistencia son adquiridos por una bacteria después de la exposición al antibiótico.

En un aspecto, los genes de resistencia a antibiótico incluyen uno o más de los genes o tipos de genes siguientes:

- 20 Los genes del operón vanHAX, regulados por el factor de transcripción VanR, que aparecen en varias bacterias incluyendo *Enterococcus* (por ejemplo, *E. faecalis*, *E. faecium*), *Staphylococcus* (por ejemplo, *S. aureus*) (Courvalin (2006) Clin. Infect. Dis. 42, S25) y se han indicado como apariciones raras en otras bacterias patógenas (Werner (2008) Future Microbiol. 3, 547). Estos genes cuando se expresan proporcionan resistencia de tipo VanA al antibiótico vancomicina. Una secuencia señuelo dirigida a los genes vanHAX comprende un sitio de unión VanR nativo (tal como en SEQ ID NO: 21 o SEQ ID NO: 26) o una variante de un sitio nativo, que compete con el sitio
25 nativo para unión de VanR en la célula de interés.

- Aquellos genes que codifican beta-lactamasas, que causan resistencia a la clase beta-lactama de antibióticos (en particular las penicilinas, cefalosporinas, cefamicinas y carbapenemos). Estos genes tienen una importancia médica particular en infecciones por Gram-negativas que contienen una subclase de beta-lactamasas conocidas como beta-lactamasas de espectro extendido (ESBL), que son evidentes en bacterias incluyendo *C. freundii*, *P. aeruginosa* y cada vez más *K. pneumoniae*, *E. coli* y *Salmonella* sp. Un informe de las beta-lactamasas encontradas actualmente puede encontrarse en Paterson (2005) Clin. Microbiol. Rev. 18, 657.

- Los tipos principales de beta-lactamasas, un ejemplo de un organismo y los antibióticos en los que influyen, y el gen responsable son como sigue: beta-lactamasas TEM, más comúnmente por el gen TEM 1 que influye en la resistencia a ampicilina en la mayor parte de Gram-negativas, incluyendo *E. coli*, *K. pneumoniae*, *H. influenzae* y *N. gonorrhoeae*, aunque se han identificado más de 140 enzimas de tipo TEM (George (2005) N. Eng. J. Med. 352, 380); beta-lactamasas SHV, más comúnmente SHV-5 y SHV-12 que causan, por ejemplo, resistencia a ampicilina en *K. pneumoniae* (Paterson (2003) Antimicrob. Agents Chemother. 47: 3554); beta-lactamasas CTX-M, que causan, por ejemplo, resistencia a cefotaxima y otras beta-lactamasas oximino (tal como ceftazidima, ceftriaxona o cefepime) en *E. coli* y *S. enterica*, se han descrito 40 enzimas CTX-M (Canton (2008) Clin. Microbiol. Inf. 14: 134);
35 beta-lactamasas OXA, que causan, por ejemplo, resistencia a oxacilina y cloxacilina en Enterobacterias, tal como *E. coli* y *K. pneumoniae* y *P. aeruginosa*; beta-lactamasas resistentes a inhibidor, tales como variantes de beta-lactamasas TEM encontradas, por ejemplo, en *E. coli* y *K. pneumoniae*, que son resistentes a ácido clavulánico y otros inhibidores; beta-lactamasas AmpC encontradas en muchas bacterias Gram-negativas incluyendo especies de Enterobacter que proporcionan resistencia de amplio espectro a cefalosporinas; carbapenemasas que codifican resistencia a cefamicinas, cefalosporinas y carbapenemos, este tipo de beta-lactamasa puede dividirse en tipo IMP (presente en bacterias Gram-negativas, particularmente *Pseudomonas* y *Acinetobacter*), VIM (siendo un ejemplo la variante dominante VIM-2 encontrada principalmente en *P. aeruginosa* y que codifica resistencia a todas las beta-lactamasas con la excepción de las monobactamas) y KPC (que codifican resistencia a carbapenem en *K. pneumoniae*).

- 50 Aquellos genes que codifican bombas de eflujo que extruden activamente antibióticos de las bacterias, de los que hay cinco superfamilias principales (Poole (2007) Ann. Med. 39: 162): facilitador principal (MFS), casete de unión a ATP (ABC), pequeña resistencia a múltiples fármacos (SMR), resistencia a la nodulación (RND) y extrusión de multifármacos y compuestos tóxicos (MATE). Aquellos antibióticos influidos por cada uno de estos sistemas de eflujo, los genes responsables y los organismos en los que aparecen comúnmente se proporcionan en la Tabla 1 de Poole (2005) J. Antimicrobial Chemother. 56, 20 y se muestran en la Figura 27. Los genes de resistencia médicamente importantes incluyen: la resistencia a macrólidos está codificada comúnmente por el gen Mef(A), por ejemplo en *Streptococci* (Pozzi (2004) Curr. Drug Targets Infect. Disord. 4, 203) y bacterias Gram-Negativas; MsrD

codifica resistencia a quetólidos, por ejemplo en *S. pneumoniae* (Daly (2004) J. Clin. Microbiol. 42: 3570); la resistencia a cloranfenicol está codificada, por ejemplo, por el sistema de eflujo Cml/CmlAB, como se describe en *P. aeruginosa*, *K. pneumoniae* y *S. enterica* (Schwarz (2004) FEMS Microbiol. Rev. 28: 519); la resistencia a eritromicina está codificada por MexCD-OprJ en *Pseudomonas* sp. (Tauch (2003) Mol. Genet. Genomics 268: 57); las resistencias a tetraciclina están codificadas por la familia Tet comúnmente TetA, TetB, TetC, TetD, TetE, Tet30 y Tet39 en bacterias Gram-negativas (Roberts (1996) FEMS Microbiol. Lett. 19: 1 y Butaye (2003) Int. J. Antimicrob. Agents 22: 205) y TetK y TetL en Gram-positivas (Butaye (2003) Int. J. Antimicrob. Agents 22: 205); resistencia a beta-lactama y aminoglicósido en *H. influenzae* está codificada por AcrAB-TolC (Rosenberg (2000) J. Bac. 182: 1754). Las bombas de eflujo de la superfamilia MATE confieren resistencia a fluoroquinolonas entre otros y aquellos genes responsables han sido identificados en *E. coli* (NorE, Morita [1998] Antimicrob. Agents Chemother. 42: 1178), *N. gonorrhoea* (NorM, Roquette-Loughlin [2003] J. Bac. 185: 1101), *H. influenzae* (HmrM, Xu [2003] Microbiol. Immunol. 47: 937), *P. aeruginosa* (PmpM, He [2004] J. Bac. 186: 262), *C. difficile* (CdeA, Kaatz [2005] Antimicrob. Agents Ther. 49: 1857), *S. aureus* (MepA, Kaatz [2005] Antimicrob. Agents Ther. 49: 1857) y *E. coli* (AcrAB, Nishino [2001] J. Bac. 183: 5803).

Aquellos genes que codifican resistencia a aminoglicósidos (tales como estreptomycin, kanamicina, tobramicina, amikacina). La resistencia está codificada comúnmente por genes que producen enzimas que modifican los antibióticos más comúnmente por N-acetilación (aminoglicósido acetiltransferasas, AAC), adenilación (aminoglicósido nucleotidiltransferasas, ANT) o O-fosforilación (Aminoglicósido fosfotransferasas, APH) (Shakil [2008] J. Biomedical Sci. 15: 5). Los ejemplos de dichos genes se encuentran tanto en bacterias Gram-negativas como Gram-positivas: AAC(6')-Ie APH(2'')-Ia confiere un espectro amplio de resistencia a aminoglicósido en enterococci y staphylococci (Hedge [2001] J. Biol. Chem. 276: 45876); la familia AAC(3) es la mayor (Sunada [2003] J. Antibiot. (Tokio) 52: 809) y, por ejemplo, confiere una amplia resistencia a aminoglicósidos en enterococci (Draker [2004] Biochem. 43: 446); los genes *ant(2'')* y *ant(4'')* codifican resistencia a gentamicina y tobramicina en bacterias gram-negativas mientras que *ant(4')* y *ant(6)* y *ant(9)* hacen lo mismo en bacterias Gram-positivas (Jana [2006] Appl. Microbiol. Biotechnol. 70: 140); en organismos Gram-positivos el gen *aph(3')* está ampliamente extendido y confiere resistencia a un rango amplio de aminoglicósidos, especialmente en staphylococci y enterococci (McKay [1996] Antimicrob. Agents. Chemother. 40: 2648).

Además, el gen *ermB* está implicado en la determinación de la resistencia a eritromicinas, tales como Zitromax(Azitromicina, particularmente en infecciones Streptococcales (Richter [2005] Clin. Infect. Dis. 41: 599).

En un aspecto, los presentes métodos pueden usarse para tener como diana la regulación de la expresión de uno cualquiera o más de los genes o tipos de genes de resistencia anteriores, por ejemplo, los genes que codifican beta-lactamasas, genes que codifican enzimas que modifican aminoglicósidos, para alterar de esta manera la resistencia al antibiótico o antibióticos correspondientes, por ejemplo, en una cepa bacteriana listada anteriormente. Las secuencias señuelo para tomar como diana la regulación de un gen o genes pueden identificarse y ensayarse como se describe en la presente memoria.

Una secuencia señuelo puede comprender una secuencia reguladora en *cis* en sí misma o, por ejemplo, una variante o fragmento de ésta que retiene la función competidora necesaria. La secuencia reguladora en *cis* puede comprender cualquiera de las descritas en la presente memoria y/o identificada según cualquiera de los métodos descritos en la presente memoria. Una secuencia señuelo puede comprender una secuencia snare como se describe en la presente memoria y/o identificada en la presente memoria, o una variante o fragmento de ésta que retiene la función competidora necesaria.

Una secuencia señuelo para tomar como diana un conjunto específico de genes, por ejemplo, genes de resistencia a antibióticos, puede prepararse tomando como base el sitio de unión nativo del factor de transcripción en una célula. Por ejemplo, las secuencias están disponibles en la bibliografía publicada. En un aspecto, una secuencia señuelo de ensayo puede comprender un sitio de unión endógeno o una secuencia consenso o una variante o fragmento de un sitio nativo.

Una secuencia señuelo de ensayo puede evaluarse para función señuelo. Típicamente, la secuencia señuelo se prepara como un polinucleótido señuelo como se describe en la presente memoria. El polinucleótido señuelo se introduce en una célula huésped que comprende un sitio de unión para el factor de transcripción dado, unido operativamente a un gen o genes, cuya expresión puede detectarse directamente o indirectamente. Por ejemplo, el cribado puede comprender ensayar directamente para expresión de un gen regulado, o ensayar para un fenotipo que está ligado causalmente a la expresión del gen. Una secuencia señuelo que causa una alteración en la expresión del gen o genes se dice que tiene función señuelo.

En la presente memoria se describen métodos adecuados para ensayar función señuelo, por ejemplo, usando los plásmidos n[snare], sistemas informadores.

Las secuencias señuelo de ensayo también pueden determinarse por los métodos descritos en la presente memoria usando bibliotecas de plásmidos n[snare] y/o métodos para identificar límites de los sitios de unión de proteínas.

Estos métodos pueden usarse para examinar la base genética para fenotipo en las células. Por ejemplo, para examinar la base genética para resistencia a antibiótico en un aislado clínico.

Un polinucleótido señuelo puede comprender cualquiera de las secuencias señuelo descritas en la presente memoria y también pueden comprender una secuencia adicional como se describe.

5 Un polinucleótido señuelo puede comprender cualquiera de las secuencias señuelo y/o polinucleótidos señuelo descritos en los Ejemplos. Así, por ejemplo, un polinucleótido señuelo que comprende el sitio de unión AfsR puede comprender un plásmido n[snare] del Ejemplo 1 o la snare comprendida en él, o el oligonucleótido de SEQ ID NO: 3, que forma una estructura de señuelo en forma de mancuerna señuelo ciclada. Un polinucleótido señuelo puede comprender un plásmido n[snare] identificado como que causa una regulación al alza en la producción de cinamicina en el Ejemplo 2 ó 3 o una snare comprendida en él. Un polinucleótido señuelo puede comprender un plásmido n[snare] identificado como que causa un incremento en la tolerancia a butanol en el Ejemplo 4 o una snare comprendida en él. Un polinucleótido señuelo puede comprender cualquiera de SEQ ID NO: 8, 9, 10, 11 ó 12 como se describe en el Ejemplo 5. Un polinucleótido señuelo puede comprender el oligonucleótido de SEQ ID NO: 21 que forma un señuelo en forma de mancuerna circular, o SEQ ID NO: 26, o una variante de especie del sitio de unión VanR.

Los polinucleótidos señuelo pueden prepararse por cualquier método adecuado. Por ejemplo, los señuelos en forma de mancuerna pueden prepararse por PCR usando cebadores apropiados, como se describe en el Ejemplo 7.2. Cada cebador contiene generalmente una parte que formará la horquilla de la estructura en forma de mancuerna. Los ejemplos de dichos cebadores se proporcionan en SEQ ID NOS: 24 y 25. La amplificación por PCR usando los cebadores está seguida típicamente de digestión por restricción del producto de amplificación y ligación para formar la estructura en forma de mancuerna circular cerrada.

Alternativamente, las estructuras en forma de mancuerna pueden prepararse por digestión de restricción de un plásmido como se describe en el Ejemplo 7.3. La digestión está seguida de ligación para formar la estructura en forma de mancuerna circular cerrada.

25 Los presentes métodos pueden usarse para alterar la expresión génica en células procariotas o eucariotas. En un aspecto, cuando el método se aplica en células eucariotas, el polinucleótido señuelo comprende un plásmido n[snare] como se describe en la presente memoria y/o la secuencia señuelo compite con una secuencia reguladora en *cis* identificada según los métodos descritos en la presente memoria y/o la secuencia señuelo comprende una secuencia reguladora en *cis* identificada según los métodos descritos en la presente memoria o una variante o fragmento de ésta.

Un polinucleótido señuelo puede introducirse en una célula huésped por cualquier medio adecuado. Por ejemplo, transformación, transfección, conjugación. Por ejemplo, cuando el polinucleótido comprende un plásmido conjugativo, éste puede introducirse por conjugación. Los polinucleótidos señuelo, por ejemplo, estructuras en forma de mancuerna circulares, pueden introducirse en las células por transfección. Las células pueden estar en cultivo líquido y el señuelo añadirse al líquido. Alternativamente, las células pueden cultivarse en medio sólido y los señuelos transfectarse de discos de papel absorbente saturados con señuelo y utilizados para cubrir el medio. Típicamente, el polinucleótido señuelo se añade al medio de cultivo y es captado por las células. Cuando las células se cultivan en medio sólido, los polinucleótidos señuelo pueden añadirse a un disco de filtro y ser captados por las células desde el disco. Puede usarse un tampón de permeabilidad para ayudar en la transfección.

40 En un aspecto, la transfección de polinucleótidos señuelo puede comprender el uso de colesterol. En particular, los métodos pueden usar polinucleótidos señuelo lineales, que presentan una modificación de colesterol en uno o ambos extremos 5'. La modificación se cree que facilita la captación por las células.

Los señuelos pueden marcarse adicionalmente, por ejemplo, en un extremo 5' con un marcador detectable tal como un agente de tinción de fluorescencia, por ejemplo, Cy5. Esto facilitará la monitorización de la captación y mantenimiento en la célula.

Los señuelos marcados con colesterol y/o detectablemente pueden prepararse usando colesterol y/o cebadores marcados detectables, como se describe en el Ejemplo 7.1.

La transfección del polinucleótido señuelo en una célula puede comprender el uso de R9-colesterol, que consiste en una molécula de colesterol unida a una cadena lineal de nueve D-argininas (Kim W.J., et al., Mol. Ther 2006 14: 343-350).

En general, se monitoriza la captación y/o mantenimiento de los polinucleótidos señuelo en las células. Por ejemplo, un señuelo de plásmido puede comprender un marcador detectable, por ejemplo, que codifica resistencia a

antibióticos, que permite la selección positiva para la presencia del plásmido y monitorización de la propagación del plásmido. La presencia de un polinucleótido señuelo también puede determinarse por crt-PCR.

En general, las células en las que se introduce el señuelo son aquellas que comprenden la secuencia reguladora en *cis* (típicamente el promotor que contiene la secuencia) con la que el señuelo compite para unión al factor de transcripción, unida operativamente al gen o genes cuya expresión se va a modular. Las células huésped adecuadas se describen en la presente memoria. Típicamente, cuando el método se usa para alterar el fenotipo de una célula, las células huésped presentan el fenotipo en ausencia del señuelo. Los señuelos para la alteración de un fenotipo médicamente o terapéuticamente relevante, por ejemplo, para incrementar la sensibilidad a antibióticos, pueden cribarse en una serie de células. Por ejemplo, los señuelos pueden ensayarse en primer lugar en un modelo bacteriano del fenotipo, por ejemplo, un modelo de resistencia a antibióticos y validarse adicionalmente, por ejemplo, en un aislado patógeno o clínico. Los señuelos pueden validarse adicionalmente en un modelo animal, como se describe en la presente memoria.

Una vez el polinucleótido señuelo se ha introducido en las células, típicamente las células se criban para la modulación de la expresión génica. El cribado puede ser directo o indirecto. Los métodos de cribado para la modulación de la expresión génica y/o fenotipo en células se describen en la presente memoria respecto a los métodos para identificar secuencias reguladoras en *cis*.

En un aspecto adicional, más de un polinucleótido señuelo puede introducirse en una célula para alterar el fenotipo. Por ejemplo, puede usarse una combinación de plásmidos n[snare], siempre que los plásmidos sean compatibles.

En un aspecto, los métodos descritos en la presente memoria son métodos *in vitro*.

Los métodos para modular el fenotipo descritos en la presente memoria tienen varias aplicaciones, por ejemplo, en industria y en terapéutica.

Los métodos pueden usarse para alterar el fenotipo en células industrialmente importantes, por ejemplo, para incrementar la producción de metabolitos útiles. De acuerdo con esto, en un aspecto, la invención proporciona un método para modular (incrementar o disminuir) la producción de un metabolito en una célula que comprende un método descrito en la presente memoria. Típicamente, en los métodos la secuencia reguladora en *cis* de interés regula la expresión de un gen o genes que codifican una proteína o proteínas necesarias para la producción de un metabolito.

En un aspecto, la secuencia reguladora en *cis* comprende una de A24.1, A24.2, A24.3, A24.4 o A24.5 identificadas en la presente memoria. La invención se refiere a un método para modular la producción de antibióticos que comprende el uso de cualquiera de A24.1, A24.2, A24.3, A24.4 o A24.5 identificadas en la presente memoria, por ejemplo, A24.1, A24.3, A24.5, tal como A24.5 o de moléculas señuelo que compiten con cualquiera de los mencionados anteriormente para unión de un factor de transcripción. La invención se refiere además a un método para modular la producción de antibióticos que comprende el uso de cualquiera de SEQ ID NOS: 8-12.

Los métodos también pueden usarse para que las células, por ejemplo, células bacterianas, se vuelvan más sensibles a los antibióticos. De acuerdo con esto, la invención proporciona además un método para modular (incrementar o disminuir) la sensibilidad a los antibióticos de las células, por ejemplo células bacterianas, que comprende un método descrito en la presente memoria. Típicamente, en los métodos la secuencia reguladora en *cis* de interés regula la expresión de un gen o genes que codifican una proteína o proteínas necesarias para la resistencia a antibióticos.

Así, en un aspecto, la invención proporciona el uso de un polinucleótido señuelo para modular la resistencia a antibióticos de una célula. En general, la modulación se causa por una alteración en la expresión de un gen o genes de resistencia a antibióticos en la célula. El efecto puede ser específico para un antibiótico particular o, en algunos casos, para una clase de antibióticos.

Generalmente, la célula comprende un gen o genes unidos operativamente a una secuencia reguladora en *cis* que comprende un sitio de unión para un factor de transcripción y el polinucleótido señuelo comprende una secuencia señuelo con el mismo sitio de unión o un sitio que compite con el sitio de unión celular para la unión del factor de transcripción. La introducción del polinucleótido en la célula reduce la unión del factor de transcripción a la secuencia reguladora en *cis* en la célula y causa una alteración en la expresión del gen o genes en la célula, modulando de esta manera la resistencia a antibióticos de la célula.

Típicamente, la secuencia reguladora en *cis* diana en la célula regula la expresión de un gen o genes de resistencia a antibióticos como se describe en la presente memoria. El señuelo interrumpe la regulación de la expresión del gen o genes y causa de esta manera una alteración (por ejemplo, una disminución) en la resistencia a antibióticos en la célula.

La célula puede ser un modelo de laboratorio de resistencia o una cepa resistente nativa, por ejemplo, un patógeno o aislado clínico. Las células adecuadas se describen en la presente memoria.

En un aspecto, la invención se refiere a un método para alterar el fenotipo de un procariota, distinto del nivel de expresión de genes de respuesta a CO₂ en *Cyanobacterium*, que comprende proporcionar un señuelo que comprende un exceso de un ácido nucleico que codifica una secuencia a la que se une un factor de transcripción procariota, en el que dicha secuencia se identifica según un método descrito en la presente memoria, de manera que después de contactar dicho señuelo con dicho procariota, el factor de transcripción se inhibe competitivamente de unirse a su sitio de unión afín en el genoma de dicho procariota.

La invención también se refiere a un método para alterar el fenotipo de un procariota, distinto del nivel de expresión de genes de respuesta a CO₂ en *Cyanobacterium*, que comprende proporcionar un señuelo que comprende un exceso de un ácido nucleico que codifica una secuencia a la que se une un factor de transcripción procariota, de manera que después de contactar dicho señuelo con dicho procariota, el factor de transcripción se inhibe competitivamente de unirse a su sitio de unión afín en el genoma de dicho procariota. El señuelo puede comprender una parte de un promotor de un gen procariota. El señuelo puede comprender un sitio de unión de un factor de transcripción del promotor de un gen procariota, por ejemplo, un gen de síntesis o regulación de antibióticos. La unión del señuelo al factor de transcripción puede resultar en susceptibilidad del procariota a un antibiótico al que el procariota es de otra manera resistente. Por ejemplo, el procariota puede ser normalmente resistente al antibiótico vancomicina pero, en presencia de dicho señuelo, el procariota ya no es resistente. El procariota puede ser un patógeno. Alternativamente, la unión del señuelo al factor de transcripción resulta en una producción incrementada por el procariota de un antibiótico.

En un caso, cuando el fenotipo que se va a alterar según los presentes métodos es resistencia a antibióticos, una posible secuencia señuelo se ensaya en primer lugar en una célula modelo de laboratorio o célula informadora como se describe en la presente memoria. Las secuencias señuelo candidatas se ensayan generalmente en una célula resistente nativa, por ejemplo, una célula patógena o un aislado clínico. En general, el método comprende además ensayar secuencias señuelo candidatas en un modelo animal adecuado. Por ejemplo, puede usarse un modelo de ratón. En general, el modelo animal se infecta con una cepa patógena que es resistente al tratamiento con antibióticos sólo.

En algunos casos, cuando una secuencia señuelo (basada en una secuencia de sitio de unión en una especie bacteriana modelo) se ensaya frente a un panel de aislados clínicos, puede encontrarse que hay algunos aislados que son refractarios al tratamiento. En dichos casos esos aislados pueden identificarse y realizarse análisis de secuencia para determinar si hay algunas variaciones en la secuencia reguladora que mimetiza el TFD. Si es así, la aparición de estas variaciones puede estimarse y crearse una mezcla del TFD que contiene secuencias modelo y variantes en proporción aproximada respecto a su incidencia en la clínica.

Los presentes polinucleótidos señuelo tienen varias aplicaciones, incluyendo aplicaciones médicas y veterinarias, así como usos *in vitro*.

Por ejemplo, los polinucleótidos señuelo con una susceptibilidad celular incrementada a uno o más antibióticos pueden usarse para tratar infecciones bacterianas en seres humanos o animales, o en métodos *ex vivo* para matar o inhibir procariotas, por ejemplo, en composiciones de limpieza antibacterianas.

Típicamente en el uso, un señuelo se usará en combinación con el antibiótico o antibióticos para los que hace a la célula más sensible. El antibiótico puede administrarse simultáneamente con, o antes o después del señuelo. El antibiótico y el señuelo pueden administrarse en la misma composición o en composiciones separadas.

Así, por ejemplo, un señuelo dirigido a la regulación VanR del operón vanHAX para disminuir la resistencia a vancomicina se usará típicamente en combinación con el antibiótico vancomicina.

La infección o afección asociada particular con la que se usa el señuelo para tratar depende en general de la célula patógena a la que está dirigido el señuelo.

Por ejemplo, Enterococci resistentes a vancomicina (VRE) tales como *E. faecium* y *E. faecalis* están asociadas con infecciones abdominales, infecciones de la piel, infecciones del tracto urinario, infecciones sanguíneas. Por lo tanto, un señuelo dirigido a la regulación VanR de los genes de resistencia a vancomicina en un VRE puede usarse para el tratamiento de dichas infecciones, típicamente en combinación con vancomicina.

Por ejemplo, las infecciones por Gram-negativas (tales como *E. coli* y *K. pneumoniae*) que portan genes que codifican para beta-lactamasas de Espectro Extendido que muestran una resistencia amplia a penicilinas están asociadas con infecciones graves del tracto urinario y el intestino. En un aspecto, el tratamiento con señuelos

diseñados para prevenir o regular a la baja la expresión de los genes que codifican las beta-lactamasas (tales como la familia CTM-X de genes) en combinación con una penicilina pueden usarse para el tratamiento de dicha infección.

En un aspecto, la resistencia a Zitromax, causada por la expresión de los genes *mefA* o *ermB* en *S. aureus*, *S. pneumoniae* o *C. pneumoniae*, y evidente en infecciones del oído y garganta podría tratarse por aplicación de señuelos dirigidos a la regulación de *mefA* y *ermB* en dichos organismos, típicamente en combinación con el antibiótico.

En un aspecto, las infecciones resistentes a fluoroginolona, tales como *S. pneumonia*, que causan neumonía podrían tratarse con señuelos dirigidos a la regulación de *norA* (el gen que codifica la bomba de eflujo) y administrarse en combinación con el antibiótico.

En un aspecto, la resistencia a macrólidos, tales como claritromicinas (por ejemplo Biaxina) en infecciones por *C. pneumoniae*, que causan neumonía, podrían tratarse por administración del señuelo dirigido a la regulación del gen de resistencia *ermB* en combinación con el antibiótico.

En un aspecto, señuelos dirigidos al gen *ampC*, que codifica resistencia a beta-lactamas (tales como penicilina, por ejemplo, Augmentine), podría usarse como una terapia frente a un amplio rango de infecciones nosocomiales por Gram-positivas y Gram-negativas por inyección con el antibiótico en un contexto hospitalario.

En un aspecto, la invención se refiere a una composición farmacéutica que comprende un polinucleótido señuelo y un vehículo o excipiente fisiológicamente aceptable. La composición puede comprender además uno o más antibióticos como se describe.

Los vehículos o diluyentes aceptables para uso terapéutico son muy conocidos en la técnica farmacéutica, y se describen, por ejemplo, en Remington's Pharmaceutical Sciences, Mack Publishing Co (A.R. Gennaro adit. 1985). La elección del vehículo, excipiente o diluyente farmacéutico puede seleccionarse respecto a la ruta pretendida de administración y práctica farmacéutica estándar. Las composiciones farmacéuticas pueden comprender como, o además de, el vehículo, excipiente o diluyente, cualquier aglutinante, lubricante, agente de suspensión, agente de recubrimiento, agente solubilizante adecuado.

Los señuelos pueden administrarse por cualquier medio adecuado, por ejemplo por inyección intravenosa, aplicación tópica o administración oral. Como anteriormente, la administración puede ser en combinación con una dosis adecuada de antibiótico, administrándose el o los antibióticos al mismo tiempo que el señuelo o separadamente.

Los polinucleótidos señuelo que incrementan la resistencia a antibióticos también pueden usarse en composiciones de limpieza tales como desinfectantes, de nuevo típicamente en combinación con el o los antibióticos para los que hacen a la célula más sensible y/o en combinación con uno o más agentes antibacterianos. En un aspecto, la invención se refiere a dicha composición de limpieza.

La invención proporciona además un kit que comprende un polinucleótido señuelo como se describe en la presente memoria y uno o más antibióticos o agentes antibacterianos, en el que el señuelo y el antibiótico o antibióticos o agente o agentes antibacterianos son para uso combinado para matar o inhibir procariotas tales como bacterias.

En un aspecto, por lo tanto, la invención se refiere además al uso de un polinucleótido señuelo como se describe y/o identifica en la presente memoria, para la fabricación de un medicamento para tratar una infección bacteriana en un sujeto. La invención se refiere además al polinucleótido señuelo para el tratamiento de una infección bacteriana y a un método para tratar a un sujeto con una infección bacteriana que comprende administrar el polinucleótido señuelo como se describe en la presente memoria.

La invención se refiere además a composiciones y medicamentos que comprenden un polinucleótido señuelo como se describe y/o identifica en la presente memoria. Los aspectos adicionales incluyen terapias de combinación en las que un polinucleótido señuelo como se identifica y/o describe en la presente memoria, se administra a un sujeto en combinación con uno o más antibióticos u otras terapias antibacterianas.

En aspectos adicionales, la invención se refiere a secuencias reguladoras en *cis* identificadas según los métodos descritos en la presente memoria, y a su uso, por ejemplo, en secuencias señuelo. En particular, la invención se refiere a secuencias reguladoras en *cis* identificadas en los presentes Ejemplos, por ejemplo, secuencias reguladoras en *cis* que comprenden A24.1, A24.2, A24.3, A24.4 o A24.5 identificadas en la presente memoria y polinucleótidos señuelo que las comprenden. La invención se refiere además a polinucleótidos señuelo como se describe en la presente memoria y/o como se describe en la presente memoria, por ejemplo a SEQ ID NOS: 3, 8-12 y 21.

La invención se refiere además a cualquiera de los plásmidos n[snare] o bibliotecas de plásmidos descritos en la presente memoria y/o preparados según los métodos descritos en la presente memoria, incluyendo los preparados en los presentes Ejemplos.

La invención se refiere además a células huésped que comprenden plásmidos n[snare] y/o polinucleótidos señuelo como se describe en la presente memoria. En particular, la invención se refiere a células huésped que presentan una expresión génica y/o fenotipos alterados. Por ejemplo, la invención se refiere a células huésped que presentan una producción incrementada de un metabolito dado, por ejemplo, antibiótico, susceptibilidad incrementada a un antibiótico o tolerancia incrementada a un disolvente dado, incluyendo los preparadas en los presentes Ejemplos.

Así, en un aspecto, la invención se refiere a células huésped (tales como las descritas en la presente memoria) que comprenden un polinucleótido señuelo como se describe en la presente memoria, incluyendo plásmidos n[snare] o bibliotecas de plásmidos como se describe en la presente memoria. Están particularmente incluidas aquellas células que tienen una susceptibilidad alterada a antibióticos (por ejemplo, susceptibilidad incrementada) y preparadas según los métodos de la presente memoria. Así, en un aspecto, la invención se refiere a patógenos o aislados clínicos que se han hecho más susceptibles a antibióticos usando los polinucleótidos señuelo como se describe en la presente memoria.

Descripción adicional de realizaciones específicas de la invención

En una primera realización según esta invención, proporcionamos un método genérico para identificar elementos reguladores en *cis*. Como en la mejor ciencia, los proyectos de Secuenciación de Genoma plantean más cuestiones de las que responden. Esencialmente, lo que proporcionaron fue una lista de todos los genes en un organismo (el genoma), pero esto planteó la cuestión de cuáles de éstos se usaban (expresaban) en un momento dado, y en qué condiciones. Ésta es una consideración importante, ya que los patrones de expresión proporcionan claves de la función de los genes e identifican aquellos responsables de características importantes (fenotipos). Consideremos las células humanas: todas contienen los mismos genes, pero el patrón de expresión dependerá de (y posiblemente determina) si la célula es del hígado o del cerebro (regulación específica de tejido). Se han desarrollado nuevas tecnologías (tecnología de micromatrices) capaces de describir el patrón de la expresión génica y pueden, por ejemplo, identificar conjuntos de genes responsables de fenotipos médicamente importantes, tales como resistencia a enfermedad. Una de las promesas de los proyectos de Secuenciación de Genoma fue que dicho conocimiento ayudará a acelerar el diseño de nuevos fármacos y estrategias para tratar enfermedades. Sin embargo, las nuevas herramientas necesitarán llevar a la práctica este potencial.

Como ahora tenemos la capacidad de identificar genes que determinan fenotipos, el reto siguiente se convierte en obtener el control de estos genes, y haciendo esto, obtener control sobre la expresión fenotípica. Esto tendría una importancia fundamental, para nuestra comprensión de la biología en general, la identificación de nuevas dianas para fármacos y también para permitirnos mejorar la producción de compuestos médicamente y comercialmente importantes a partir de bacterias industriales (por ejemplo, los denominados productos de fermentación). Los productos de fermentación varían desde alimentos, medicinas, y, cada vez más, enzimas industriales y combustibles. Es un objetivo que merece la pena el obtener estos productos a partir de una fuente sostenible (bacterias) al contrario que basarnos en los métodos de fabricación (petro)-químicos tradicionales y costos medioambientalmente.

Los genes se activan e inactivan por proteínas especializadas, es decir, factores de transcripción. Éstos tienen una doble función: se unen a secuencias de ADN específicas (elementos reguladores), habitualmente cercanos a un gen diana y alteran la probabilidad de que el gen se exprese. Los intentos iniciales se centraron en cambiar las propiedades o abundancia de los factores de transcripción en sí mismos. Nuestra estrategia es tomar como diana los elementos reguladores a los que se unen los factores de transcripción. Una estrategia que hemos desarrollado es introducir un exceso de estos elementos reguladores (denominados "señuelos"), para evitar de manera competitiva que el factor de transcripción se una a la versión genómica y bloquear de esta manera cualquier influencia en el gen en 3' (véase la discusión más adelante respecto a la segunda y tercera realizaciones según esta invención). En esta realización de la invención, construimos bibliotecas de señuelos (que contienen bien todos los elementos reguladores de una especie o una versión universal con todos los posibles elementos reguladores) que se pueden ensayar rápidamente para identificar qué conjuntos de elementos reguladores tienen el efecto deseado. Una parte clave de la estrategia consiste en construir un "sistema informador" genérico, de manera que, en lugar de buscar cambios en fenómenos complejos, tales como rendimiento incrementado de una bacteria industrial, la modificación exitosa se monitoriza por cambios en la expresión de un gen informador introducido que confiere resistencia a antibióticos y que es fácil de detectar. El sistema se diseña de tal manera que el cribado puede realizarse con un alto rendimiento. Así, este aspecto de la invención representa una tecnología nacida a partir de la era genómica en la que puede ser posible producir datos de alta calidad rápidamente y relevantes para el genoma completo. Se prevé que la mezcla de una biblioteca de señuelos universal y un sistema informador genérico mejore rápidamente la administración de productos biotecnológicos y sea una herramienta de investigación incalculable.

Más específicamente, según esta realización de la invención, hemos adaptado la estrategia de oligonucleótidos “señuelo” para controlar la regulación genética en un procariota y para influir en el control de la producción de antibióticos en *Streptomyces coelicolor*. Los señuelos, sin embargo, son sensibles a la degradación y la estrategia no es factible para alto rendimiento. Abordamos estos dos problemas en el método según esta invención, (véanse los Ejemplos 1-4). Generalmente, usando del Mecanismo por Círculo Rodante de amplificación de ADN, un oligonucleótido se convirtió en una serie (aproximadamente treinta) de repeticiones directas bicatenarias. Éstas se clonaron en un vector lanzadera de alto número de copias (capaz de propagarse en *E. coli* y *Streptomyces*) y se introdujo en *S. coelicolor* por conjugación. El plásmido se mantuvo por selección con antibióticos. Hemos encontrado que este sistema se comporta mejor que los señuelos: hubo una mayor represión de la producción de antibióticos y persistió durante la duración del experimento. El método de construcción de la biblioteca se ha adaptado de manera que pueden prepararse bibliotecas genéricas a partir de oligonucleótidos aleatorios sintéticos, o bibliotecas específicas de especie preparadas a partir de fragmentos del genoma completo. Para crear un sistema de cribado universal, se crean informadores para selección negativa y positiva en el huésped común para expresión heteróloga, *Streptomyces lividans*. El promotor diana se clona en 5' de un gen de resistencia a kanamicina o *glkA*, que determina la sensibilidad al metabolito 2-desoxiglucosa (DOG). Después de la transformación de una biblioteca, los transformantes se criban en cultivo líquido para tolerancia incrementada a kanamicina (lo que detecta regulación negativa) o sensibilidad incrementada a DOG (lo que detecta positiva). Como la selección es en cultivo líquido, la gran mayoría de los clones se pierden, dejando sólo un número pequeño y manejable para seguir con el análisis. Se desarrolla un sistema de validación en el que conjuntos de colonias de cribados independientes se ensayan para ver cuáles hibridan de manera cruzada con miembros de otros conjuntos. De esta manera es posible identificar clones que se repiten en la biblioteca y que son buenos candidatos para análisis de secuencia.

En una segunda realización según esta invención, (véase el Ejemplo 5), usamos oligonucleótidos señuelo para estudiar la regulación del antibiótico pigmentado azul actinorhodina en *Streptomyces coelicolor* A3(2). Este organismo contiene (para un procariota) un genoma relativamente grande (8,7 Mb) con un patrón complejo y adaptativo de regulación génica, particularmente respecto a las indicaciones de desarrollo y medioambientales que controlan la producción de antibióticos (Bibb (2005) *Curr. Opin. Microbiol.* **8**: 205-215). *S. coelicolor* también es el organismo modelo para los actinomycetes, y el incremento del nivel de comprensión de la regulación de la producción de antibiótico en esta cepa puede informar acerca de nuevas estrategias para obtener acceso a una amplia variedad de metabolitos secundarios producidos por estos organismos. Muchos de estos compuestos tienen aplicaciones importantes en medicina, por ejemplo como antibióticos, y en agricultura (Berdy (2005) *J. Antibiot.* **58**: 1-26) y los actinomycetes siguen siendo una fuente rentable de nuevos fármacos y enzimas (Ward (2006) *Curr. Opin. Microbiol.* **9**: 279-286).

Quizá como consecuencia de la regulación compleja de la producción de antibióticos, muchos mutantes pleyotrópicos identificados por cribados genéticos son condicionales; por ejemplo, el fenotipo no productor de antibiótico de un mutante anulador *relA* es muy dependiente del medio (Chakraborty (1997) *J. Bacteriol.* **179**: 5854-5861). La aparición de fenotipos nutricionalmente condicionales implica que los cribados genéticos pueden subestimar el número de factores reguladores que influyen en la producción de antibióticos. La inactivación de un factor de transcripción *bona fide* puede perderse si, en las condiciones usadas, la activación de los genes diana puede estar mediada por un factor de transcripción o ruta reguladora alternativa. Una de las ventajas de los señuelos es que identifican y manipulan el componente de secuencia de las interacciones ADN-proteína que controlan la expresión génica, bloqueando potencialmente la interacción de numerosos factores de transcripción en un único sitio. En esta realización según la invención (véase el Ejemplo 5), demostramos la relativa facilidad de tomar como diana secuencias reguladoras y la capacidad de identificar rápidamente nuevos genes reguladores de una manera que es complementaria a los cribados genéticos convencionales.

En una tercera realización según esta invención (véase el Ejemplo 6), demostramos la utilidad de los oligodesoxinucleótidos señuelo para conferir sensibilidad a un antibiótico en una bacteria que de otra manera es resistente a ese antibiótico. Los expertos en la técnica apreciarán la significancia extrema de este aspecto de la invención en una era en la que las denominadas “bacterias multirresistentes” se encuentran en todas partes, incluyendo en nuestros hospitales. Frecuentemente atribuido al uso extendido de antibióticos, el desarrollo de estas denominadas bacterias multirresistentes tiene el potencial de generar infecciones pandémicas. Mediante este aspecto de la invención, sin embargo, los expertos en la técnica apreciarán que mediante la comprensión apropiada de los mecanismos por los que los patógenos consiguen resistencia a antibióticos específicos, pueden usarse los oligonucleótidos señuelo identificados según los métodos de esta invención antes de, al mismo tiempo que, o tanto antes como al mismo tiempo que la administración del antibiótico particular.

Muchos genes están controlados a múltiples niveles mediante la interacción de factores reguladores con sitios reguladores en *cis*. Los complejos de nucleoproteína formados pueden hacer esto de una manera específica de tejido, específica del estado de desarrollo o dependiente de estímulo. Brevemente, las interacciones ADN-proteína *in vivo* se determinan fisiológicamente y representan la interacción entre dos componentes dinámicos: las proteínas de unión a ADN que constituyen el entorno que actúa en *trans* y las secuencias reguladoras en *cis*. Aunque nuestro conocimiento de los patrones de expresión génica ha crecido sustancialmente en los últimos años, este crecimiento no ha sido paralelo a un incremento comparable de nuestro conocimiento de los factores reguladores que controlan

genes específicos que influyen en procesos celulares específicos. Prevemos que la aplicación de las técnicas de la presente memoria proporcionará probablemente nuevas visiones sobre las complejidades de la regulación genética procariota y las establecerá como una estrategia complementaria al análisis genético convencional.

Ejemplos

Habiendo descrito generalmente esta invención, incluyendo sus realizaciones preferidas y mejor modo, se proporcionan los ejemplos específicos siguientes, junto con descripción adicional, para permitir y extender completamente la descripción descrita de esta invención. Los expertos en la técnica apreciarán, sin embargo, que esta invención no debe considerarse limitada a la parte específica de los Ejemplos. En lugar de esto, para propósitos de comprensión del alcance de esta invención, debe hacerse referencia a las reivindicaciones adjuntas a esta descripción, incluyendo los equivalentes de éstas.

Aunque en general muchas de las técnicas mencionadas en la presente memoria son muy conocidas en la técnica, puede hacerse referencia en particular a Sambrook et al., 1989, *Molecular Cloning: a laboratory manual*.

Ejemplo 1

Demostración de la creación de versiones portadas en plásmidos de oligonucleótidos señuelo y su utilidad para modificar fenotipos

Para conseguir este trabajo, fue necesario en primer lugar desarrollar un protocolo de biología molecular capaz de crear un plásmido n[snare]. Una característica deseada de dicho vector plasmídico es que debe contener una sección que consiste en copias directas del mismo fragmento de ADN. Habitualmente, el número de estas copias es 30, pero puede variar en número de una a mil y la longitud de la secuencia repetida está en el intervalo de 35-54 pb, pero puede ser mayor, por ejemplo, hasta 1.000 pb. En el ejemplo de construcción proporcionado más adelante, en esta secuencia repetida, hay un segmento (27 pb) que es común a todos los plásmidos n[snare] y es una consecuencia del método de fabricación. También en este ejemplo, el vector plasmídico se eligió: (i) para permitir la transformación y propagación en *E. coli* y *S. coelicolor*; (ii) para contener un gen marcador seleccionable que codifica resistencia al antibiótico apramicina; y (iii) para mantener un número alto de copias en la célula, típicamente cerca de 100 copias por célula.

Método para crear plásmido [snare] que contiene el sitio de unión AfsR

Una visión global del proceso se proporciona en la Figura 1. Se sintetizan oligonucleótidos monocatenarios con una estructura quimérica: (A) un sitio de hibridación para el cebador T7, y (B) una región que contiene la secuencia de la secuencia reguladora en *cis* o de la posible secuencia reguladora en *cis*. El oligonucleótido se circulariza por la acción de la ligasa *Taq*, en presencia de un oligonucleótido de unión universal. Después de digestión con exonucleasa para digerir el ADN lineal, se recuperan círculos monoméricos de un gen de acrilamida preparativo. Éstos se usan posteriormente como moldes en una reacción de amplificación por el Mecanismo por Círculo Rodante usando polimerasa *Bst* y el cebador T7. El mismo cebador se usa de nuevo en una reacción de PCR y los productos se separan en un gel de agarosa. Los fragmentos de alto peso molecular, que contienen típicamente 30-50 repeticiones, se aíslan y se clonan en un vector de PCR antes de subclonarlos en un vector lanzadera de alto número de copias, para formar un plásmido n[snare]. Para determinar la función, el plásmido n[snare] se transforma en la cepa nativa o en una cepa informadora genérica. Este protocolo también pueden usarse para producir bibliotecas n[snare] "genéricas", preparadas a partir de oligonucleótidos aleatorios, o bibliotecas específicas de especie, a partir de genomas fragmentados. Una ilustración de cómo los plásmidos n[snare] influyen en la expresión de genes diana se proporciona en la Figura 2. En la parte superior izquierda de esta figura, un gen (representado por una barra horizontal "A") es transcripcionalmente inactivo debido a la unión de un factor de transcripción represor (círculos "B") a una secuencia reguladora en *cis* (barra horizontal "C") en el promotor del gen. La secuencia de la secuencia reguladora en *cis* que constituye el sitio de unión para el factor de transcripción puede determinarse por varias técnicas incluyendo análisis bioinformático, "obtención de impronta" del promotor o los distintos métodos n[snare] descritos en la presente memoria. Esta secuencia se incorpora en un oligonucleótido señuelo para influir en la expresión del gen, pero se muestra aquí clonada en un plásmido seleccionable en varias (típicamente 30) repeticiones directas, para crear un plásmido n[snare]. El plásmido es habitualmente un plásmido "lanzadera", lo que significa que puede propagarse en *E. coli* (para facilitar la manipulación genética) y el organismo diana; además el plásmido es habitualmente de "alto número de copias", lo que significa típicamente que producirá aproximadamente 100 copias de sí mismo en la célula huésped. El plásmido n[snare] se introduce en el procariota diana, por medios estándar, y su propagación estable se asegura seleccionando su marcador, habitualmente un gen que codifica resistencia a un antibiótico. Cuando se introduce así en una célula el plásmido n[snare] es capaz de influir en la expresión del gen diana liberando el factor de transcripción "B" del promotor genómico para liberar la represión transcripcional del gen en 3' (mostrado como la caja horizontal "D", parte superior derecha).

Se sintetizó un oligonucleótido señuelo que contiene el sitio de unión AfsR. Tenía la secuencia siguiente:

5'-Fosfato-aat acg act cac tat agg **ggc gtt gag cga acg ttt ttc gcg gcc** gc-3' SEQ. ID. 1

en la que el sitio de unión AfsR se muestra por la secuencia en negrita y un sitio de restricción para *NotI* está subrayado. El oligonucleótido se sintetiza con un grupo fosfato en el extremo 5'.

También se sintetizó un nucleótido "de unión", R-T7, denominado así porque contiene un complemento parcial de un cebador usado comúnmente, T7:

5'-ccc tat agt gag tcg tat tgc gg-3' SEQ.ID.2

Ambos cebadores se resuspendieron a concentraciones finales de 250 pmoles/μl en un tampón de reacción que consiste en tampón 1 x ligasa Taq (como lo suministran los fabricantes, New England Biolabs) y 50 U ligasa Taq. La mezcla se calentó hasta 95°C durante 10 minutos antes de dejarlo enfriar hasta 45°C, después de lo cual la reacción se suplementó con 4 U/ml ligasa Taq y la reacción se incubó toda la noche a 45°C.

Se usó 1 μl de la reacción como un molde para la amplificación por círculo rodante (RCA) en una mezcla de incubación que consiste en tampón de reacción 1 x Thermopol (como lo suministran los fabricantes, New England Biolabs) suplementado con 0,2 mM dNTP, 70 nM cebador T7 (el complemento de R-T7) y 120 U/μl de polimerasa Bst (New England Biolabs). La mezcla se incubó toda la noche a 60°C.

El ADN amplificado se recuperó después de una etapa de limpieza en un sistema de columna disponible comercialmente (kit de Purificación de PCR Qiagen) y se usó 1 μl de este ADN en una reacción de PCR estándar que consiste en tampón 1 x polimerasa Taq (suministrado por los fabricantes, Roche Diagnostics) suplementado con 0,2 mM dNTP, 5% DMSO, 25 pmoles cebador T7 y 0,5 U de polimerasa Taq (Roche Diagnostics). La PCR se realizó usando los parámetros de ciclado siguientes: 95°C durante 5 min, 25 repeticiones de la secuencia siguiente: 93°C 15s, 55°C durante 15s, 72°C durante 2 min, seguido de 72°C durante 2 min.

Los productos de la reacción se analizaron por electroforesis en gel 1% agarosa/TBE y se visualizaron por transiluminación del gel teñido con bromuro de etidio.

Una preparación exitosa aparece como una "escalera" de secuencias con una longitud de repetición de aproximadamente 50 pb y con una cantidad suficiente del material migrando con un tamaño aparente de 1,5 kb para permitir la purificación (equivalente a aproximadamente 30 repeticiones del monómero de 50 pb).

En general, cuando se realiza este procedimiento, si el tamaño de los productos es principalmente pequeño (menor de un tamaño de 500 pb) entonces las amplificaciones se repiten inicialmente variando las cantidades de moldes usadas, y en la reacción de PCR, el número de ciclos de la amplificación empleado. En el evento de que esto no produzca productos del tamaño deseado, se sigue una estrategia alternativa en la que se realiza una digestión parcial con *NotI* del producto de la amplificación con *Bst*, los fragmentos se fraccionan por tamaño para aislar un molde de 1,5 kb que se liga a los conectores compatibles con *NotI* con el fin de permitir una amplificación eficaz en una reacción de PCR Mediada por Ligación estándar (LM-PCR).

El fragmento de 1,5 kb se subclona en el vector disponible comercialmente pGEMT-Easy de Promega, que se adapta con protuberancias 3'T para clonar eficazmente los productos de PCR. Usando técnicas de biología molecular estándar los fragmentos se recuperan después de digestión con *EcoRI* del ADN aislado de una alicuota de la biblioteca pGEMT-Easy y se purifican por gel. La colección de los fragmentos *EcoRI* se subclona en un sitio de restricción similar en un vector lanzadera, en nuestro caso pJ86, que es capaz de replicarse en *E. coli* y *S. coelicolor*. El fragmento n[snare] portado en el plásmido pJ86 se introduce en *S. coelicolor* por conjugación usando métodos estándar (Kieser *et al.* (2000) Practical *Streptomyces* Genetics. John Innes Foundation).

Uso del plásmido n[snare] afsR para alterar el fenotipo de producción de antibióticos en S. coelicolor

AfsR es un regulador pleyotrópico de la producción de antibióticos en *S. coelicolor* y las delecciones preparadas por ingeniería de este gen han mostrado que es esencial para la producción de los dos antibióticos pigmentados, actinorhodina (que es azul) y undecilprodigiosina (que es rojo) en *S. coelicolor* (Hong *et al.* (1991) *J. Bacteriology* **173**: 2311-2318). Una secuencia reguladora en *cis* clave que actúa como un sitio de unión para AfsR se encuentra en 5' del gen *afsS* (Lee *et al.* (2002) *Molecular Microbiology* **43**: 1413-1430) y fue esta secuencia la que se usó para ensayar el plásmido n[snare]. Se razonó que se esperaba que un señuelo (preparado en paralelo para propósitos de comparación) o un plásmido n[snare] que contiene este sitio inhibirá la expresión de *afsS* liberando la proteína activadora AfsR del promotor, con una disminución resultante de la producción de los dos antibióticos pigmentados. Por lo tanto, se monitorizaron tres curvas de crecimiento para *S. coelicolor* no tratada que servirán como un control, un cultivo tratado con un TFD ciclado que contiene el sitio AfsR (como se ha descrito anteriormente y mostrado en negrita, usando el oligonucleótido: 5'-Fosfato-ata gcg ttg agc gaa cgt ttt tc **gcg tttt cgc ga aaa acg ttc** gct caa cgc tat ag tttt ct-3' SEQ ID. 3) y una cepa transformada con el plásmido n[snare] AfsR (Figura 3). En esta figura el promotor

de tipo salvaje de *afsS* se muestra parte superior izquierda "A", AfsR (óvalo "B") se une a su sitio afín (rectángulo "C") para activar la expresión del gen regulador. El gráfico "I" en la parte superior derecha muestra la curva de crecimiento del cultivo y la producción de los dos antibióticos pigmentados, actinorhodina (Act) y undecilprodigiosina (Rojo); (la ordenada izquierda = producción de antibióticos, la ordenada derecha = crecimiento bacteriano, la línea sólida muestra medidas de peso seco bacteriano, los diamantes representan producción de actinorhodina (Act) y los cuadrados representan producción de undecilprodigiosina (Rojo)). El gráfico "II" muestra la introducción de un señuelo en forma de mancuerna circular (+ Señuelos; parte media izquierda) que proporciona competición para la unión a AfsR, que reprime la expresión del regulador codificado por *afsS*, con una disminución simultánea de la producción de antibiótico (parte media derecha gráfico II). Los plásmidos *n[snare]* que contienen repeticiones homopoliméricas de la secuencia reguladora en *cis* "C" (parte inferior izquierda) se demuestran más eficaces para suprimir la producción de antibiótico (*n[snare]*) posiblemente debido a la selección positiva para el plásmido, y el efecto represor se observa a lo largo del curso del experimento (gráfico II parte inferior derecha). Los plásmidos *n[snare]* son capaces de modificar el control de la expresión génica de una manera predecible con un cambio simultáneo en el fenotipo. Estos plásmidos son más baratos de producir y más fáciles de introducir en las células que los señuelos, proporcionan un efecto más sostenido y, crucialmente, permiten una estrategia de biblioteca para descubrir elementos reguladores clave.

Por lo tanto, lo que se concluyó fue que, comparado con el cultivo control (Figura 3, fila superior), la producción de antibiótico se suprime tanto en cultivos tratados con señuelo como con *n[snare]*. Sin embargo, con el tratamiento con señuelo, el efecto es en gran medida transitorio recuperándose la producción de antibiótico después de 48 h, mientras que con el tratamiento con *n[snare]*, el efecto supresor persistió más allá del curso del experimento. Los experimentos control usando versiones desordenadas de la secuencia de unión AfsR bien en el formato señuelo o incorporadas en un plásmido *n[snare]* fracasaron para suprimir significativamente la producción de antibiótico. Estas observaciones apoyan la conjetura de que los plásmidos *n[snare]* pueden usarse para identificar y caracterizar secuencias reguladoras en *cis* y en algunos aspectos (estabilidad y grado en el que el fenotipo se modifica) son mejores que la técnica existente de oligonucleótidos señuelo.

Ejemplo 2

Un método para identificar elementos reguladores en *cis* y, por lo tanto, señuelos para modificar el fenotipo en procariotas o eucariotas, usando bibliotecas *n[snare]*

En este ejemplo, la biblioteca *n[snare]* se prepara a partir de fragmentos de ADN cortos aislados de la agrupación biosintética de cinamicina de 17.083 kb (Widdick y Bibb (2003) *Proc. Natl. Acad. Sci. USA* **100**: 4316-4321) y la biblioteca se reintroduce en la cepa productora de *Streptomyces cinnamoneus*. La detección del antibiótico se realiza usando un ensayo en placa en el que se permite que crezcan las colonias de la cepa productora en una placa de agar que se cubre con un agar nutriente blando sembrado con un inóculo de la cepa informadora *Bacillus subtilis*. Un halo de agar nutriente transparente aparece alrededor de las colonias productoras en las que el antibiótico ha matado a la cepa indicadora en el que en otro lugar el agar nutriente aparece turbido debido al crecimiento continuo de la cepa. El diámetro del halo es un indicador de la cantidad de cinamicina producida y, después de la introducción de la biblioteca *n[snare]*, puede usarse como un cribado conveniente para las colonias exconjugantes superproductoras. De esta manera, es posible determinar si los miembros de la biblioteca pueden regular al alza la producción del antibiótico y confirma la utilidad de la estrategia como una herramienta de descubrimiento para escanear fragmentos genómicos considerables para elementos reguladores en *cis* que pueden incorporarse en TFD. La lógica de la estrategia puede expandirse para desarrollar estrategias para bibliotecas *n[snare]* para detectar secuencias candidatas en evaluaciones mucho mayores (incluyendo genomas completos) y las demostraciones de estas estrategias y sus utilidades potenciales se discuten en los Ejemplos posteriores. Los expertos en la técnica apreciarán, sin embargo, que el método no está limitado a las partes específicas de esta ejemplificación y otros fenotipos procariotas o eucariotas pueden explorarse de manera similar usando esta metodología para conseguir alteraciones fenotípicas esencialmente de cualquier rasgo de interés.

1. Construcción de bibliotecas *n[snare]* personalizadas

Un esquema del procedimiento para generar estos fragmentos se muestra en la Figura 4. Las bibliotecas de plásmidos *n[snare]* se crean a partir de grandes partes de ADN de manera tal que cada secuencia reguladora en *cis* en ese ADN debe estar representada en la biblioteca. En esta figura se muestra que el ADN genómico o fracciones de genomas, tales como agrupaciones biosintéticas, se fragmentan por sonicación y los extremos libres se reparan por tratamiento con polimerasa *Taq* y dNTP, para generar extremos con una única protuberancia dA en el extremo 3' del ADN. A éstos se ligan conectores biotinilados que contienen sitios de restricción para *Nt.BbvCI* (caja oscura), *MmeI* (caja clara) y una protuberancia 5' dT complementaria. El adaptador no ligado se elimina de la mezcla de ligación antes de que los fragmentos modificados se digieran con *MmeI*, una enzima que corta 20/18 nucleótidos en 3' de su sitio de reconocimiento, permitiendo la introducción de dT en la protuberancia; las moléculas resultantes consisten en el adaptador biotinilado intacto más una cadena de 19/17 nucleótidos de ADN genómico, a la que nos referimos como la "snare". Estas moléculas se capturan en una matriz cubierta con estreptavidina y se desnaturalizan para dejar la cadena superior (que contiene la cadena de 19 nt) unida a los lechos. Un

oligonucleótido adicional se hibrida ahora diseñado para recrear el sitio *Nt.BbvCI* y éste se digiere posteriormente con esta enzima para liberar una parte monocatenaria de ADN que consiste en una parte del adaptador más la secuencia snare. Esto se usa en etapas posteriores para crear bibliotecas n[snare].

1.1 Preparación de los fragmentos de ADN

- 5 Un fragmento de ADN que contiene las aproximadamente 17 kb de la agrupación biosintética de cinamicina se aisló del cósmido en el que se había clonado previamente. El fragmento se purificó por gel y se sonicó hasta una longitud media de 500 pb como se juzga analizando una alícuota del ADN por electroforesis en un gel de TBE/agarosa. La sonicación produce una población heterogénea de ADN que termina con protuberancias 3' y 5' de varias longitudes y ADN romo en el que había ocurrido una rotura bicatenaria. Con el fin de convertir éstas en una población homogénea que consiste en extremos con una protuberancia 3' dA, el ADN se trató con polimerasa *Taq*.

1.2 Creación de "snare" con una longitud de 18 nucleótidos a partir de un fragmento genómico

El término "snare" se usa para referirse a la parte de ADN monocatenaria corta derivada de los fragmentos de secuencia genómica que se usan para crear los plásmidos n[snare].

- 15 Usando métodos estándar se prepara un adaptador biotinilado de la secuencia siguiente (en la que P significa Fosfato) y se liga a los fragmentos de ADN tratados:

Nt.BbvCI *MmeI*

5'-Biotina-ggt ccg ggc cac ggt ggt cta cga **gcc tca gcc agg tcc gac t**-3' SEQ ID 4

3'- agt cgg tcc agg ctg-P-5' SEQ ID. 5

- 20 Los fragmentos de ADN se separan del conector no incorporado usando el kit de Purificación de PCR Qiagen y se resuspenden en 30 µl de tampón de Elución (10 mM Tris.HCl pH8). Los fragmentos de ADN adaptados se digieren a 37°C durante 16 h en un tampón con un volumen de reacción de 200 µl (1 x Tampón 4 [New England Biolabs] suplementado con 50 µM de S-adenosilmetionina) que contiene 20 Unidades de enzima de restricción *MmeI* (New England Biolabs). Después de la digestión, los fragmentos se precipitan y separan en un gel de acrilamida 12% no desnaturizante y las bandas se analizan. El adaptador libre es evidente como lo es una especie que migra menos de adaptador más la protuberancia de 19/16 nt generados por la digestión asimétrica del fragmento genómico por la enzima de restricción tipo IIS *MmeI*. Esta banda se escinde del gel y los fragmentos se recuperan usando técnicas estándar. Los 19 nt en la cadena superior es la parte del material genético incorporado en los plásmidos n[snare] como candidatos para secuencias señuelo. Como estos fragmentos están biotinilados se capturan en una matriz paramagnética recubierta con estreptavidina, tal como lechos M-280 de Dynal. El ADN en estos lechos se desnaturaliza en 0,5 M NaOH seguido de calentamiento hasta 80°C y el ADN monocatenario resultante se lava en 1 x Tampón 4 (New England Biolabs). 100 pmoles de un oligonucleótido complementario *Bbv* (que contiene el sitio *Nt.BbvCI* que está subrayado) de la secuencia siguiente:

5'-gga cct ggc tga ggc tcg tag acc acc gtg gcc cgg acc-3' SEQ. ID. 6

- 35 se hibridan con el ADN monocatenario capturado con el fin de preparar un sitio de restricción para *Nt.BbvCI*. Antes de la digestión los lechos se lavan concienzudamente para eliminar cualquier oligonucleótido que no se liga. La mezcla se suplementa ahora con 25 Unidades de la enzima y la reacción se incuba durante 4 h a 37°C, después de lo cual la mezcla se calienta hasta 60°C durante 15 min para desnaturalizar el ADN mellado. La parte "snare" se libera de los lechos mientras el resto se retiene (Figura 4). Después de la captura de los lechos en un soporte magnético, el ADN del sobrenadante se recupera y se usa para crear los plásmidos n[snare] usando métodos similares, como se ha descrito previamente.

2. Uso de una biblioteca de plásmidos n[snare] para regular al alza la producción del antibiótico cinamicina en su huésped natural

- 45 La biblioteca n[snare] creada a partir de la agrupación biosintética de cinamicina se introdujo en la cepa productora *S. cinnamoneous* por conjugación y se plaqueó en agar R2YE (Kieser *et al.* (2000) *Practical Streptomyces Genetics*. John Innes Foundation) y se permitió que creciera durante 3 días. Se realizaron experimentos control en paralelo en los que *S. cinnamoneous* no se trató o se conjugó con una cepa donante que contiene un vector vacío (que no contiene ninguna repetición directa). Estos controles permitieron una estimación del tamaño medio y distribución de los tamaños de los "halos" de manera que la medida de los tamaños de los halos en la muestra tratada con n[snare] se usa para determinar los superproductores estadísticamente significativos. Un ejemplo de un incremento en el tamaño del halo en un exconjugante n[snare] se muestra en la Figura 5. Cuando se plaquea en medio sólido R2YE, *S. cinnamoneous* secreta un antibiótico, cinamicina, en el agar que es capaz de matar células de *Bacillus subtilis*

cuando la colonia productora se cubre con un cultivo de esa cepa (lado izquierdo). El diámetro de este halo se usa como una medida de la cantidad de cinamicina producida. La producción se incrementa por introducción de una biblioteca n[snare] que contiene fragmentos derivados de la agrupación biosintética de cinamicina de 17.083 kb. Después de la introducción de los miembros de la biblioteca por conjugación, se seleccionan clones tomando como base la producción incrementada de cinamicina (derecha).

Usando esta estrategia, fue posible identificar plásmidos n[snare] capaces de regular al alza la producción de antibióticos y, por extensión, encontrar las secuencias reguladoras en *cis* clave en la agrupación biosintética de cinamicina. En una aplicación, dichas secuencias en los plásmidos n[snare], o incorporadas en TFD, podrían usarse para manipular la producción de antibióticos, y otros compuestos biológicos industrialmente valiosos. Usando esta estrategia, hemos sido capaces de aumentar el nivel de producción de compuestos derivados de actinomycete identificados previamente (Figura 5). Esto es frecuentemente un obstáculo principal en el desarrollo comercial de un producto natural, y esta tecnología es adecuada para explotar el potencial comercial completo de los actinomycetes activando o aumentar en gran medida la expresión de las denominadas agrupaciones génicas metabólicas secundarias “crípticas” (Zazopoulos (2003) *Nat. Biotech.* **21**: 187-190) o rendimientos incrementados en general para biotecnología industrial. Se prevé así que esta tecnología en plataforma beneficie la salud humana y animal y ejemplifica la explotación colaboradora de la secuenciación genómica y los desarrollos en genómica.

Ejemplo 3

Un método para usar bibliotecas n[snare] para identificar elementos reguladores en *cis* y, por lo tanto, señuelos, para modificar fenotipo usando sistemas informadores preparados por ingeniería.

1. Identificación de reguladores de la producción de cinamicina con bibliotecas n[snare] y cepas informadoras *cin7*

La agrupación biosintética de cinamicina se ha descrito previamente (Widdick y Bibb (2003) *Proc. Natl. Acad. Sci. USA* **100**: 4316-4312) y el papel de algunos de los genes identificados se ha confirmado por análisis bioinformático y genético. El gen *cin7* se sabe que está precedido por el promotor que controla la transcripción del operón *cinMXTH* que codifica la proteína pequeña (*cinA*) que se convierte enzimáticamente para proporcionar el antibiótico cinamicina (Sean O'Rourke y Mervyn Bibb, datos no publicados). Creamos un sistema informador dirigido por el promotor *cin7* y se introduce esta construcción en una cepa de *Streptomyces lividans* que porta la ruta biosintética completa de cinamicina (cepa 1326). Con el fin de demostrar la estrategia, se introdujo la biblioteca n[snare] personalizada de cinamicina (descrita en el Ejemplo 2) en la cepa por conjugación y los exconjugantes se cribaron para supervivencia en concentraciones normalmente letales de kanamicina.

2. Procedimiento para detectar regulación negativa con bibliotecas n[snare] y un sistema basado en informador

El promotor del gen diana (*cin7*) se pone en 5' de un gen que confiere resistencia a kanamicina, *neo*, como se muestra en la Figura 7. En esta figura el promotor diana del gen *cin7* de la agrupación biosintética de cinamicina de la cepa productora de *S. cinnamoneous*, se usa para dirigir la expresión del gen *neo* que codifica resistencia al antibiótico kanamicina. Una secuencia reguladora en *cis* en el promotor *cin7* se muestra como una barra “A”, al que se une un represor transcripcional (óvalo “B”), que vuelve el gen *neo* en 3' inactivo y consecuentemente la cepa es sensible a kanamicina. Típicamente, este gen quimérico se introduce en el genoma de la cepa productora por uso de vectores integrativos. Una biblioteca n[snare] personalizada, en este caso creada a partir del ADN fraccionado del genoma de la cepa informadora *S. cinnamoneous*, se introduce y los transformantes/exconjugantes se criban para resistencia a concentraciones crecientes de kanamicina. Las células que son más resistentes se eligen para análisis adicional.

Por lo tanto, cuando no hay o hay poca transcripción en la agrupación biosintética de cinamicina en la cepa productora heteróloga *S. lividans* 1326, el promotor *cin7* es transcripcionalmente inerte y la célula huésped es susceptible a selección con kanamicina, ocurriendo la muerte celular en las cepas que contienen un plásmido con un gen *neo* sin promotor que codifica resistencia a kanamicina a concentraciones menores de 5 µg/ml (Labes *et al.* (1997) *Microbiology* **143**: 1503-1512). El gen quimérico *cin7-neo* se porta en un plásmido integrativo (basado en el núcleo pSET152 Kieser *et al.* (2000) *Practical Streptomyces Genetics*. John Innes Foundation) que también porta el gen *aacC* que codifica resistencia al antibiótico apramicina. Introducimos el casete en la cepa de *S. lividans* 1326 (que contiene la agrupación biosintética de cinamicina) por métodos muy conocidos en la técnica para crear una nueva cepa (MM14) con el gen quimérico integrado de manera estable en el genoma. La biblioteca n[snare] personalizada de cinamicina preparada como se describe en el Ejemplo 2 se introdujo en MM14 por métodos estándar para generar una suspensión de esporas de la biblioteca. Se usó un número fijo de esporas para inocular medio líquido R3 (Shirahama *et al.* (1981) *Agric. Biol. Chem.* **45**: 1271-1273) y se monitorizó el crecimiento en presencia de una concentración creciente de kanamicina. La viabilidad celular a cada concentración se determinó tomando alícuotas del mismo volumen de los cultivos y extendiéndolas en placas de agar (agar ADN) para determinar las unidades formadoras de colonias/ml de cultivo y a partir de éste, el porcentaje de células viables. Estos datos muestran que en este experimento, la viabilidad a 650 µg/ml de kanamicina fue menor de 0,01%, lo que se estima igual a 1.200 unidades formadoras de colonias en un cultivo de 50 ml. Ésta se consideró como la

concentración óptima de kanamicina ya que enriqueció en gran medida con los que tienen una resistencia incrementada al antibiótico, lo que resulta en un número de células manejable. La viabilidad a 150 µg/ml fue 85%, lo que establece la actividad en la línea base de la construcción *cin7-neo*.

Se realizó una etapa de hibridación de la biblioteca. En este ejemplo, en cuatro replicados independientes, la cepa MM14 se conjugó con la biblioteca personalizada de cinamicina. La concentración óptima de kanamicina para cada repetición se determinó empíricamente. Finalmente, 96 clones de cada repetición que sobrevivieron a la concentración óptima se cultivaron en medio R3 y se dividieron en dos alícuotas. De una, se extrajeron los plásmidos de las células usando procedimientos estándar, teniendo cuidado de evitar contaminación bien por ARN o ADN genómico. Este ADN plasmídico se marcó por PCR con cebado aleatorio usando una molécula de dUTP marcada con DIG según métodos estándar (manual Roche) para crear una muestra de sonda. La segunda alícuota se diluyó y se extendió en placas de agar ADN a una concentración en la que podían verse claramente colonias individuales. Las placas se incubaron a 30°C durante un periodo de 16 h, después del cual las colonias todavía mantenían una consistencia cerosa y todavía no se habían endurecido. 96 de dichas colonias se rasparon con un palillo, que se usó primero para preparar una estría en una segunda placa de ADN (suplementada con 50 µg/ml de apramicina) y se mezcló en 25 µl de dimetilsulfóxido (DMSO), se calentó hasta 95°C durante 10 min y se enfrió para crear una disolución de ADN total derivado del clon. Tanto la estría como la disolución de ADN se numeraron de manera que si posteriormente se encontraba que la muestra de ADN contenía un plásmido n[snare] candidato podría recuperarse fácilmente un exconjugante viable (cepa MM14 más plásmido n[snare]).

Para identificar los plásmidos n[snare] genuinamente capaces de interferir con la transcripción (para distinguir los miembros de "fondo" que forman una parte considerable de la señal total de los plásmidos n[snare] genuinamente capaces de interferir con la transcripción), se desarrolló una estrategia de hibridación de la biblioteca. En este ejemplo, usando métodos estándar, se procesaron muestras de ADN de 96 clones de cada repetición y se hibridaron con una membrana de nilón cada uno con una posición localizable. La estrategia seguida se muestra en la Figura 9. Para cada repetición, se prepararon cuatro de dichos filtros (F1, F2, F3 y F4 derivados respectivamente de cada repetición) y se hibridaron separadamente con diferentes sondas, de manera que eventualmente todos los conjuntos de sondas se hibridan con cada uno de los filtros. La hibridación de P1 y F1 detecta un cromatograma en el que cada muestra se detecta. No se espera que la detección sea homogénea ya que las secuencias n[snare] que aparecen múltiples veces en la biblioteca son proporcionalmente más brillantes. En este ejemplo, las hibridaciones de P1 y P2 con F3 y F4 detectan muestras comunes a todas las bibliotecas. Estas muestras se considerarán candidatos fuertes para continuar el escrutinio.

Los datos ejemplares de dicho cribado se muestran en la Figura 10. En este ejemplo, dos repeticiones independientes de la biblioteca de la cinamicina se ensayaron con cepas informadoras y 96 de los clones más resistentes a kanamicina de cada uno se usaron para preparar un filtro que contiene los plásmidos n[snare] en cada replicado (F1 y F2) y dos conjuntos de sondas correspondientes (P1 y P2). Todos los clones identificados por hibridación cruzada (F1 frente a P2 y F2 frente a P1) se consideraron candidatos fuertes, y se dio prioridad a los comunes a ambos experimentos de hibridación cruzada (con círculo en rojo). La secuencia de tres de los seis clones estudiados identificó una parte del promotor *cin7* y otros dos alrededor de un regulador en *cis* potencial de otro gen en la agrupación de cinamicina, *cinR*, que se ha propuesto que tiene un papel en la regulación de la agrupación. La secuencia final no se identificó y se presumió que derivaba del cromosoma de *S. lividans* o que era un artefacto.

En algunos contextos, la hibridación de P1 con F1, y otros experimentos de auto hibridación, pueden ser capaces de detectar clones falsos positivos debido al enriquecimiento, por ejemplo, de secuencias repetitivas en bibliotecas eucariotas. Esto generaría señales mucho más fuertes debido a su abundancia en la biblioteca, es decir, mayor que los positivos verdaderos que se espera que aparezcan más de una vez en la biblioteca pero no tantas veces como ADN repetitivo (tipos de los cuales pueden representar el 10% del genoma humano en masa).

3. Adaptación de informadores para detectar regulación positiva

Se usa una adaptación sencilla para preparar el sistema informador capaz de detectar regulación positiva, creando un casete que consiste en un gen informador quimérico que incorpora la secuencia codificadora del gen de la glucosa quinasa (*glkA*) dirigido por el promotor diana. Esta enzima convierte el metabolito 2-desoxiglucosa (DOG) en la versión fosforilada, tóxica del compuesto. Por lo tanto, mediante la preparación de un gen informador quimérico que consiste en el promotor *cin7* unido al gen *glkA* e integrando de manera estable éste en una cepa de *S. lividans* que contiene una delección *glkA*⁻ se proporciona dicho sistema informador (Figura 8).

En el contexto de este experimento, se usa la cepa TK24 de *S. lividans*. Ésta no contiene la agrupación biosintética de cinamicina (como tenía la cepa usada para detectar los elementos reguladores en *cis* negativos que controlan la producción) y como tal cualquier actividad del promotor *cin7* será debida a los reguladores positivos derivados del genoma de *S. lividans*. Por lo tanto, la búsqueda de secuencias reguladoras en *cis* positivas en este entorno se espera que encuentre reguladores al alza de la expresión heteróloga del cromosoma del huésped *S. lividans*. La liberación de estos reguladores positivos del promotor *cin7* y en un plásmido n[snare], evita la expresión de *glkA* y las células crecen en presencia de DOG. Hasta la fecha, se ha realizado un trabajo considerable usando cribados

genéticos convencionales para detectar dichos reguladores genéricos de producción heteróloga, que puede estar causada por la sobreexpresión de determinados componentes de la maquinaria transcripcional, tales como factores sigma y componentes de la maquinaria de traducción. Además, se ha identificado una clase de factores de transcripción conocidos como SARP (Butler *et al.* (2003) *Appl. Microbiol. Biotechnol.* **61**: 512-516) y los elementos reguladores en *cis* a los que éstos se unen son dianas potenciales para este cribado. El cribado se realiza de la misma manera que la descrita previamente con ventajas similares: compatibilidad con procesamiento de alto rendimiento, métodos estandarizados y fenotipos fácilmente evaluables.

Ejemplo 4

Un método para usar bibliotecas n[snares] para identificar elementos reguladores en *cis* y, por lo tanto, señuelos para modificar el fenotipo en sistemas genéticos no definidos

1. Creación de bibliotecas n[snares] genómicas

El ADN genómico purificado de la cepa K12 de *E. coli* se sonizó para proporcionar un tamaño medio de 500 bp. Los fragmentos se trataron como se describe en el Ejemplo 2: adaptados con un conector biotinilado que contiene sitios de restricción convenientes, digerido para producir adaptadores más una protuberancia de 19/17 nucleótidos, capturado en una matriz de estreptavidina, antes de restricción y desnaturalización para liberar un fragmento de ADN monocatenario que contiene una parte del adaptador más la protuberancia de 19 nucleótidos. Estos tipos de moléculas se usan para crear bibliotecas de plásmidos n[snares] usando las técnicas descritas en el Ejemplo 2.

Se estimó que una versión de la biblioteca genómica n[snares] de *E. coli* K12 (referida de aquí en adelante como "biblioteca K12") tenía 1.560.000 miembros, lo que, considerando el tamaño del genoma y de los fragmentos usados para construir la biblioteca n[snares], corresponde abarcar hasta el 99,8% del genoma completo usando análisis estándar. Por lo tanto, si no hay ningún sesgo de secuencia en la creación de las bibliotecas, tres versiones independientes de la biblioteca K12 permitirían abarcar completamente el genoma. Con el fin de evaluar si había algún sesgo en la biblioteca, se aislaron plásmidos de una alícuota de dicho cultivo crecido en líquido. Esta mezcla de plásmidos se digirió con *EcoRI* para regenerar los fragmentos usados para crear las bibliotecas n[snares] y éstos se sometieron a digestión con *MmeI* para confirmar que había funcionado la biología molecular de la construcción: como se esperaba los insertos tenían un tamaño grande y éstos se convirtieron en un monómero de 30 pb después de digestión completa con *MmeI* (Figura 11), mientras que la digestión parcial dio lugar a productos intermedios con tamaños que eran múltiplos de la repetición de 30 pb. La secuenciación de la colección de plásmidos confirmó claramente la secuencia del elemento "espaciador" derivado del adaptador y que el espacio se mantenía bien sobre al menos diez copias. Como se esperaba, la parte "snare" de la colección de plásmidos era esencialmente una secuencia aleatoria con poco sesgo de nucleótidos en cualquiera de las 19 posiciones.

2. Creación de bibliotecas n[snares] universales

Se sintetizaron oligonucleótidos con la secuencia siguiente:

5'-Fosfato-aat acg act cac tat agg gnn nnn nnn ngc ggc cgc-3' SEQ. ID. 7

en los que un sitio de restricción para *NotI* se muestra subrayado y un nucleótido aleatorizado se representa por "n". En este ejemplo, el número de nucleótido aleatorizados es 9, pero este número puede variar. El tipo de secuencias codificadas por la región variable también puede modificarse, por ejemplo por inclusión de regiones constantes en la secuencia (es decir, una secuencia núcleo definida flanqueada por secuencias variables) o por la introducción de sesgo de nucleótidos en las regiones variables (es decir, estipulando que "n" debe ser un dGTP o dCTP el 60% del tiempo para crear secuencias aleatorizadas con un contenido GC mayor). El protocolo para crear plásmidos n[snares] con este oligonucleótido que contiene secuencia aleatorizada y, por lo tanto, una biblioteca n[snares] universal es el mismo que el proporcionado en el Ejemplo 2. Nos referimos en la presente memoria a estas bibliotecas n[snares] derivadas de oligonucleótidos aleatorizados como bibliotecas Universales.

Se encontró que la titulación de una versión de la biblioteca excedía 2 millones. Usando análisis estándar, la confianza de que una biblioteca de dicha titulación contuviera todas las posibles combinaciones de secuencias de 9 pb sería $7,5 \times 10^5$. Sin embargo, se piensa que es improbable que dicha biblioteca fuera nunca verdaderamente universal ya que determinadas secuencias estarán ausentes de la biblioteca por muchas razones tales como sesgo de PCR frente a secuencias ricas en GC o secuencias con estructura secundaria extensa y aquellas secuencias de la colección que contienen uno/o sitios *MmeI* o *NotI*. Sin embargo, se apreciará que lo que se ha creado tiene el potencial de ser una colección sustancial y exhaustiva de secuencias reguladoras en *cis* potenciales para uso en la estrategia de biblioteca n[snares].

3. Procedimiento para usar bibliotecas n[snare] para detectar reguladores en cis de fenotipos con red genética no definida

La cepa de *E. coli* K12 se transformó con la biblioteca n[snare] genómica o una versión de la biblioteca universal y se creció en cultivo líquido. Cuando el crecimiento alcanzó la fase logarítmica media como se estimó por lectura de la absorbancia del cultivo (una lectura a A_{600} de 0,4) y se usaron alícuotas para sembrar medio suplementado sin disolvente o con concentraciones de butanol que varían de 0,5% a 5%. Los cultivos se incubaron durante 1 h más y las unidades formadoras de colonias se midieron con el fin de calcular la viabilidad celular. Las células se aislaron de los cultivos en los que la viabilidad había disminuido 10.000 veces (hasta 0,01% o menos) y se lavaron en medio sin disolvente y se usaron para reinocular medio fresco sin disolvente. El ciclo experimental se repitió; este cultivo se creció hasta una densidad de 0,4 y se usó para inocular medio suplementado con concentraciones crecientes de butanol. Este proceso se repitió cuatro veces y la viabilidad celular se midió cada tiempo (Figura 12). La iteración del proceso de selección (que significa que la biblioteca de los miembros existentes después de un paso a través de butanol se extrajeron y usaron para transformar *E. coli* K12) incrementó el recuento de viabilidad como una función del número de repeticiones del procedimiento de selección. Las *E. coli* que no se trataron con la biblioteca K12 n[snare] se designaron "0" (representado por los diamantes llenos), aquellas con 1 tratamiento se representaron con cuadrados llenos, aquellas con 2 con triángulos llenos, aquellas con 3 con círculos vacíos y aquellas con 4 con asteriscos. Lo que fue evidente es que las bacterias transformadas con la biblioteca n[snare] comparado con la transformación control, que consiste en el núcleo del plásmido sin inserto n[snare] o no transformadas, proporcionaron unos recuentos de viabilidad celular sustancialmente mayores cuando se cultivaron en presencia de altas concentraciones de butanol. Además, fue evidente que la iteración del proceso de selección incrementó el recuento de viabilidad como una función del número de repeticiones del procedimiento de selección.

La biblioteca personalizada de *E. coli* K12 se comportó mejor que la biblioteca n[snare] universal, lo que es consistente con que esta biblioteca sea más representativa del genoma de la cepa.

Ejemplo 5

Manipulación y comprensión de la producción de antibiótico en *Streptomyces coelicolor* A3(2) con oligonucleótidos señuelo

Los que se ha descrito hasta ahora es un conjunto de estrategias innovadoras para identificar secuencias reguladoras en cis implicadas en el control de fenotipos. Las tecnologías son aplicables universalmente a todos los organismos. A continuación, se demostró que el conocimiento de estas secuencias podría usarse para crear herramientas capaces de modificar la expresión génica *in vivo*. Estas herramientas son de nuevo funcionales en todos los organismos, pero la atención de los ejemplos siguientes es el uso en procariotas. La estrategia del señuelo del factor de transcripción se adapta para uso en procariotas y se muestra que es una manera potente de movilizar la información generada por los protocolos n[snare] para desarrollar un sistema capaz de modificar la expresión génica. La combinación de las dos estrategias se refiere en la presente memoria como "ingeniería reguladora".

En el ejemplo siguiente las secuencias reguladoras en cis que influyen en la producción de antibiótico en *Streptomyces coelicolor* se determinan por un nuevo procedimiento de mapeo y estas secuencias se usan para diseñar oligonucleótidos señuelo. Los señuelos se ensayan por introducción en *S. coelicolor* y se monitorizan su captación, estabilidad y efecto en la producción de antibiótico. Una descripción completa de los métodos usados se proporciona en la sección 7 de este ejemplo.

1. Mapeo de elementos reguladores en el promotor *actII-orf4*

Las interacciones ADN-proteína que controlan la expresión de *actII-orf4* se estudiaron por un mapeo *in vivo* con ADNasal/exonucleasa T7. Este método se desarrolló para identificar los límites de los elementos reguladores en cis de manera que las secuencias deducidas pudieran usarse para diseñar oligonucleótidos señuelo para estudios funcionales. Un resumen esquemático de la estrategia se proporciona en la Figura 13. Para mapear los límites *in vivo* de los complejos ADN-proteína en localizaciones específicas en el genoma de *S. coelicolor*, se añadieron altas concentraciones de ADNasal y exonucleasa T7 a células recién recogidas. La ADNasal introdujo "mellas" en el ADN alrededor de los complejos que sirvieron como sustratos para la actividad exonucleasa 5' a 3' de la exonucleasa T7. Por lo tanto, las actividades conjuntas de las enzimas demarcan los límites 5' de los complejos. El ADN se recupera de las células tratadas y los fragmentos del promotor diana (*actII-orf4*) se capturan por hibridación en una cadena inmovilizada de un fragmento de PCR del promotor, que incorpora un conector biotinilado. Las posiciones de los límites en la población de fragmentos capturados se mapean realizando una reacción de PCR usando un segundo cebador marcado no radiactivamente con DIG (incorporando dioxigenina, DIG [Roche]) que hibrida con el conector biotinilado. Los tamaños de los productos marcados se determinan por PAGE y detección quimio-luminiscente. Los límites de la cadena opuesta se mapean de una manera similar, usando un producto de PCR con el conector biotinilado en el extremo opuesto. Los fragmentos de PCR marcados se fraccionaron por tamaño por electroforesis en gel de poliacrilamida al 12% y la presencia de los fragmentos y su tamaño se determinó por detección quimiluminiscente. La nueva combinación de exonucleasa T7 y ADNasal en un protocolo de obtención de impronta

permite la detección de todos los límites de los complejos ADN-proteína en una región promotora y no sólo los más cercanos al sitio de restricción elegido.

Como nuestro objetivo era descubrir elementos reguladores implicados en la represión de la expresión de *actII-orf4*, mapeamos el promotor transcripcionalmente regulado a la baja en un medio líquido rico (R5). En estas condiciones, los represores pueden ocupar el promotor y evitar la expresión. Se obtuvieron curvas de crecimiento midiendo la densidad celular (A_{430} del cultivo) y se determinaron la producción de los dos antibióticos pigmentados (la actinorhodina azul y la undecilprodigiosina roja [Kieser *et al.* (2000) Practical *Streptomyces* Genetics. John Innes Foundation]) y la actividad transcripcional de *actII-orf4*. La última se sabe que está inducida durante las etapas tardías del crecimiento y así las muestras para mapeo se recogieron antes del inicio visualmente detectable de la producción de actinorhodina (indicada con una flecha en la Figura 14A). Las células de *S. coelicolor* M145 se recogieron de un cultivo crecido en medio rico (R5) en un punto de tiempo (indicado con una flecha) anterior a la producción visible de actinorhodina. El crecimiento celular (diamante) y la producción de actinorhodina (triángulo) y undecilprodigiosina (círculo) se monitorizaron todo el rato. Las cantidades de las enzimas requeridas para la digestión se determinaron empíricamente, necesitando la concentración de ADNasa una optimización cuidadosa. Por ejemplo, un exceso de ADNasa resultó en la pérdida de la claridad de la señal debido al corte en los complejos ADN-proteína, mientras que muy poca enzima fue ineficaz para mapear. 250 U de exonucleasa T7 por reacción funcionó bien en la mayoría de los casos; fue necesario añadir un exceso de la enzima ya que no es altamente procesadora, ni lo fue la digestión realizada en un tampón óptimo. Los límites se mapearon en ambas cadenas, como se describe en la Figura 13, y sus posiciones se determinaron después del análisis de tamaño por PAGE no desnaturizante al 12%, seguido de detección quimio-luminiscente de los productos marcados con DIG (Figura 14B) y por comparación con una escalera de tamaño estándar fue posible definir los límites de los complejos ADN-proteína en la región promotora transcripcionalmente silenciosa *actII-orf4* (Figura 14C). La secuencia del promotor *actII-orf4* muestra las posiciones de los posibles elementos reguladores en *cis* (respecto a los cebadores usados en el protocolo de mapeo). Las áreas en cajas indican las secuencias codificadoras del gen en 5' (*actII-orf3*) y de *actII-orf4* mismo. La secuencia en mayúsculas marca los elementos reguladores candidatos con sus nombres mostrados encima. Las secuencias subrayadas indican las cajas -35 y -10 para el promotor *actII-orf4*, el asterisco muestra la posición del sitio de inicio transcripcional y las flechas convergentes indican la repetición invertida presente en A24.4.

Nos referimos a las secuencias definidas por estas regiones como elementos reguladores y en total se observaron cinco. Los elementos reguladores se marcaron A24.1 a A24.5 (yendo hacia el sitio de inicio transcripcional) y estas secuencias se usaron para diseñar los oligonucleótidos señuelo usados en los estudios funcionales posteriores.

2. Cribado rápido para función señuelo en placas de agar

Desarrollamos un ensayo rápido basado en placas de agar para ensayar si los señuelos tenían algún efecto en la producción de antibiótico. Se usó agar R2YE ya que estimula la producción de los dos antibióticos pigmentados a niveles que son fácilmente detectables a simple vista. Las placas se inocularon con una suspensión diluida de esporas y se incubaron durante 24 h a 30°C de manera que se formaron capas confluentes de *S. coelicolor*. En esta etapa, empezó la producción de undecilprodigiosina (que es roja) pero no la de actinorhodina (que es azul). Las placas se cubrieron con un capa fina de SNA en agarosa al 0,5% y antes de eso, se pusieron en el medio discos pequeños de papel Whatman (discos de Ensayo de Antibiótico) y se saturaron con 15 µl de una disolución de 10 pmoles/µl de oligonucleótido señuelo o disolución control, como se indica en la Figura 15. Los discos de filtro se saturaron con disoluciones de señuelos o, como control, tampón (como se muestra) y se aplicaron a una capa de *S. coelicolor* M145 cubierta con medio SNA (Agar Nutritivo Blando: Agar Difco al 0,5% p/v). Al continuar creciendo las bacterias, las incidencias de la producción aumentada de antibióticos pudieron reconocerse por la acumulación temprana de los antibióticos pigmentados en y alrededor de los discos (evidente a las 48 horas después de la adición de los señuelos). Los controles negativos (tampón solo o una versión "desordenada" de señuelo A24.5) no mostraron la producción precoz. Todas las muestras se prepararon en un tampón que contenía 0,5% (v/v) de los dos detergentes no iónicos NP-40 y Tritón X-100, que se encontró que mejoraban la eficacia de la transfección. Los controles consistieron en el tampón solo o un oligonucleótido señuelo desordenado, en el que la secuencia del señuelo A24.5 era aleatorizada. La placa se incubó más a 30°C y se inspeccionó regularmente para determinar si había algún efecto en la cantidad o tiempo de la producción de antibióticos. Se observó un halo morado/rojo alrededor del disco empapado con A24.5 y, en menor grado, A24.3 y la mezcla de señuelos A24.1 a A24.4 (Figura 15, 48 h; en puntos de tiempo más tempranos la actinorhodina aparece roja ya que la molécula actúa como un indicador de pH, aunque los halos eran visibles claramente por encima del fondo de undecilprodigiosina secretada). Para confirmar que la síntesis de actinorhodina había empezado, los discos a 48 horas se recuperaron, los pigmentos se extrajeron y cuantificaron espectrofotométricamente; el señuelo A24.5 activó fuertemente la producción temprana de actinorhodina, mientras que A24.3 y la mezcla de señuelos también lo hizo en menor medida (datos no mostrados). No se observó una producción temprana de actinorhodina alrededor de los discos control. A las 96 horas, la capa completa de *S. coelicolor* M145 había producido grandes cantidades de ambos antibióticos (no mostrado) y no eran aparentes zonas de inhibición del crecimiento alrededor de ninguno de los discos. La capacidad de algunos de los señuelos para aumentar la producción de actinorhodina en puntos de tiempo tempranos, potencialmente por interferencia con la unión de represores a sus sitios identificados en la región

promotora *actII-orf4*, demostró la eficacia de la estrategia y nos estimuló para realizar estudios funcionales más extensos introduciendo los señuelos en células crecidas en cultivo líquido.

3. Captación y estabilidad de señuelos en cultivos líquidos

Los siguientes puntos a abordar fueron si un señuelo podría introducirse eficazmente o no en un micelio en cultivo líquido, y si se podía, ¿durante cuánto tiempo persistiría? Seis cultivos líquidos de *S. coelicolor* M145 se prepararon en medio SMM (Kieser *et al.* (2000) Practical *Streptomyces* Genetics. John Innes Foundation) y se crecieron hasta fase exponencial media. Las células se recogieron por centrifugación suave y se resuspendieron en un décimo del volumen original de un tampón de permeabilización (que contiene las mismas concentraciones de detergentes como se usan en los ensayos de placa) antes de la transfección con cantidades varias de señuelo A24.1 (0 mM, 5 mM, 10 mM, 20 mM, 50 mM y 100 mM). Después de una incubación breve, las células se resuspendieron en el medio retenido y se continuó la incubación. La captación del oligonucleótido señuelo se midió por crt-PCR, usando cebadores diseñados para amplificar un fragmento pequeño (40 pb) que abarca la unión introducida en la formación del señuelo en forma de mancuerna. La referencia a una curva estándar se usó para calcular el número absoluto de copias de señuelo presentes en las células (después de centrifugación y dos etapas de lavado) y esto se corrigió por referencia a un control genómico. Para evaluar la estabilidad, se tomaron alícuotas en diferentes puntos de tiempo (0 a 72 horas). La velocidad de captación se saturó por encima de concentraciones de 20 mM. La captación óptima se consiguió con una transfección 20 mM, 2 horas después de lo cual el 45% del señuelo había entrado en las células (Figura 16). Las células que crecían activamente se transfectaron con una disolución que contiene señuelo y la captación del oligonucleótido y su estabilidad se estimaron por PCR en tiempo real cuantitativa (crt-PCR). Después de la transfección con 20 mM A24.1, las células se recogieron y lavaron antes de usar crt-PCR para estimar las copias de señuelo que permanecen en la célula como una función de tiempo. Los datos representan tres determinaciones independientes. El señuelo podría detectarse intracelularmente 72 horas después de la adición, con 22% (correspondiente a la mitad que había entrado en las células) persistiendo después de 36 horas. Como los cultivos se transfectaron cerca de la fase estacionaria, se asumió que los señuelos se estaban degradando lentamente por endonucleasas endógenas a diferencia de perderse debido a la dilución. Así, el señuelo fue capaz de entrar en el micelio y persistir durante un período prolongado, lo que sugiere que esta estrategia podría usarse para alterar la expresión génica en cultivos líquidos.

4. Uso de oligonucleótidos señuelo para controlar la producción de antibiótico

Los cinco señuelos, correspondientes a los elementos reguladores identificados, se usaron para transfectar cultivos líquidos de *S. coelicolor* M145 crecida en medio líquido R5 o SMM. La transfección se calculó en el tiempo para preceder a la expresión de *actII-orf4*, proporcionando a los señuelos la oportunidad de interactuar con sus factores de transcripción afines e influir potencialmente en la producción de actinorhodina. Los experimentos control fueron similares a los usados en los ensayos en placa y consistieron en un procedimiento de transfección simulado o en la introducción de tres señuelos desordenados (basados en las secuencias de los señuelos A24.1, A24.3 y A24.5). En cada experimento, el crecimiento del cultivo se midió, así como la producción de los antibióticos pigmentados y la expresión de *actII-orf4*, a intervalos fijos después de la transfección. Ninguno de los señuelos alteró el crecimiento de las células pero varios tuvieron un efecto en la producción de actinorhodina y, en algunos casos, de undecilprodigiosina. En medio R5, cuando la producción de antibiótico es relativamente alta, el señuelo A24.5 reguló al alza la producción de actinorhodina, incrementando el rendimiento un 95% en el punto de tiempo de 96 horas (Figura 17A y B). *S. coelicolor* M145 se creció durante 20 horas antes de la transfección (indicada con flechas) con (A) un control sin señuelo o con (B) el señuelo A24.5. El crecimiento celular (diamantes) y la cantidad de undecilprodigiosina producida (círculos) fue similar en los dos cultivos y la única variación observada fue en la acumulación de actinorhodina (triángulos), que se estimó después de tratamiento con el señuelo A24.5. Los datos representan la media de tres determinaciones independientes y las barras muestran el error estándar. Los señuelos A24.1 y A24.3 tuvieron efectos más suaves, causando una regulación al alza de los dos antibióticos pigmentados. En todos los tratamientos en los que ocurrió un incremento en la producción de actinorhodina, hubo un incremento correspondiente en el nivel absoluto de la expresión de *actII-orf4* (determinado por crt-PCR; datos no mostrados). Por lo tanto, los resultados fueron muy consistentes con los observados en las placas, lo que sugiere que el efecto de los señuelos es específico y predecible. Los señuelos también se ensayaron en medio SMM, un medio mínimo que apoya menos producción de antibiótico que R5. La transfección con el señuelo A24.5 dio lugar a una duplicación de la producción de actinorhodina (Figura 18A y B). La comparación de los datos obtenidos con (A) un control transfectado simulado y (B) un cultivo tratado con el señuelo A24.5 reveló que el oligonucleótido señuelo causaba un incremento pronunciado en la producción de actinorhodina. Los datos se presentan en la Figura 17 y son la media de tres determinaciones independientes; las barras muestran el error estándar. En este medio, los incrementos en la producción de actinorhodina también se observaron para los señuelos A24.1 y A24.3 (datos no mostrados). Por lo tanto, los señuelos actuaban como herramientas convenientes para validar la identificación de elementos reguladores en *cis* en el promotor *actII-orf4*, permitiéndonos influir en el inicio de la producción de antibióticos. Como se ha indicado anteriormente, varios de los señuelos también regulaban al alza la producción de undecilprodigiosina y las posibles razones de esta co-regulación se abordan más adelante.

5. Descubrimiento de un nuevo modificador de la producción de actinorhodina

La combinación de una nueva técnica de mapeo y oligonucleótidos señuelo se ha usado para validar tres elementos reguladores que controlan la expresión del gen regulador de actinorhodina *actII-orf4*. A continuación, nos preguntamos si estos restos reguladores ocurrían en las regiones promotoras de otros genes cualesquiera; dichos genes también podrían estar implicados en la regulación de la producción de antibióticos. La realización de una búsqueda BLAST con las cinco secuencias señuelo identificó uno de dichos genes. SCO5812 contenía equivalencias fuertes con A24.1 (uno de 10 de 14) y A24.3 (uno de 11 de 18) El gen en sí mismo es un homólogo potencial de la ribonucleasa HII, y como tal puede tener una función similar a un mutante pleiotrópico previamente identificado en la producción de antibióticos, *absB* (SCO5572; Adamidis y Champness (1992) *J. Bacteriol.* **174**: 4622-4628), que es un homólogo de la ARNasaIII y se piensa que está implicado en el procesamiento de los transcritos (Chang *et al.* (2005) *J. Biol. Chem.* **280**: 33213-33219). Para establecer el papel, si existe, de SCO5812 en la producción de antibióticos, deletamos el gen de M145 y comparamos la producción con la cepa parental en medio R5 y SMMS (Figura 19). En esta figura, M145 (lado izquierdo de las placas) y M145 Δ SCO5812 (lado derecho de las placas) se sembraron en estrías en (A) medio agar R5 o (B) medio agar SMMS y se incubaron durante 72 h y 96 h, respectivamente. Mientras que la delección de SCO5812 redujo dramáticamente la producción de actinorhodina en agar R5, aumentó la producción de undecilprodigiosina en SMMS. Es interesante indicar que las células transfectadas con A24.1 o A24.3 mostraron una regulación al alza ligera de la producción de undecilprodigiosina, lo que sugiere que los factores de transcripción que se unen a estos sitios pueden influir en la expresión de las dos agrupaciones de genes biosintéticos de antibiótico. También uno de los señuelos ensayados, A24.4, que contiene una repetición invertida de 6 pb (Figura 14), se encontró que tenía una equivalencia (8 de 12) en la región promotora de *redD*, el gen activador específico de ruta de la agrupación biosintética de undecilprodigiosina (Takano *et al.* (1992) *Mol. Microbiol.* **6**: 2797-2804). Este resto es un sitio de unión predicho para la proteína nucleóide IHF (en la base de la homología de secuencia respecto a sitios identificados previamente), surgiendo la posibilidad de que el sitio tiene un papel arquitectónico en el complejo de nucleoproteína en lugar de influir directamente en la transcripción.

6. Observaciones concluyentes

Hemos usado técnicas nuevas para identificar y validar un sitio de unión para un represor de la producción de antibióticos en *S. coelicolor*. La validación se realizó con oligonucleótidos señuelo: la transfección de copias de estos elementos reguladores dio lugar a la desrepresión del gen *diana* y a una producción incrementada de actinorhodina. La intención de nuestro trabajo fue demostrar que los oligonucleótidos señuelo pueden ser una herramienta valiosa en procariotas para delinear rápidamente redes genéticas. ¿Ha tenido éxito? En dos sentidos sí. Tres de los cinco señuelos ensayados mostraron la actividad esperada. El señuelo con el efecto más fuerte, A24.5, se ha mostrado recientemente que está unido por un regulador transcripcional semejante a TetR, AtrA, cuando el promotor es activo (Uguru *et al.* (2005) *Mol. Microbiol.* **58**: 131-150). Nuestro propio trabajo, usando purificación por afinidad para identificar represores de *actII-orf4*, ha identificado de manera similar factores de transcripción semejantes a TetR unidos a este sitio; nuestra asunción es que como el señuelo A24.5 se añade antes de la transcripción de *actII-orf4*, permite la expresión disminuyendo la unión de estos posibles represores.

7. Materiales y Métodos

7.1 Cepas y condiciones de crecimiento

Una referencia estándar para técnicas generales referentes al manejo de streptomycetes es Kieser *et al.* (2000) *Practical Streptomyces Genetics*. John Innes Foundation. Se germinaron esporas de *S. coelicolor* A3(2) cepa M145 de manera sincronizada por tratamiento con calor y se crecieron en medio SMM o R5 a 30°C con agitación continua. El crecimiento del cultivo y la producción de actinorhodina se midieron como se ha descrito previamente. Se usaron R2YE y SNA para ensayos en placa con agar de la producción de antibióticos inducida por señuelo. La delección de SCO5812 se consiguió por direccionamiento con PCR (Gust *et al.* (2003) *Proc. Natl. Acad. Sci. USA* **100**: 1541-1546).

7.2 Mapeo in vivo con exonucleasa T7

Antes del tratamiento con el reactivo de obtención de impronta, los cultivos de fragmentos miceliales de *S. coelicolor* se suplementaron con 0,5 mM CaCl_2 y 50 unidades de ADNasaI (para eliminar ADN extracelular) y se incubaron durante 15 min adicionales a 30°C. El micelio se recogió por centrifugación a baja velocidad y se lavó concienzudamente en tampón TES (13) suplementado con 5 mM EDTA. Para la obtención de impronta con ADNasaI, las células se lavaron en TES suplementado con 0,5% (v/v) NP-40 y 0,5% (v/v) Tritón X-100 y se incubaron a 30°C durante 15 min. Las células se lavaron en tampón de Digestión con ADNasaI (DDB) precalentado hasta 30°C y se digirieron con varias cantidades de ADNasaI durante 5 minutos. Las reacciones se pararon por adición de un volumen igual de tampón de PARADA (50 mM Tris.HCl, 5 mM EDTA, 0,2% (p/v) SDS, 10 mg/ml Proteinasa K, pH8) y se incubaron toda la noche a 55°C. Los ácidos nucleicos se precipitaron con etanol después de

dos ciclos de extracción con fenol-cloroformo. Las muestras se digirieron con ARNasaA antes de la reprecipitación y cuantificación.

7.3 Estudios con oligonucleótidos señuelo

Las secuencias de todos los oligonucleótidos usados se listan en la Tabla 1. Los oligonucleótidos señuelo se sintetizaron (Invitrogen) y se ligaron con CircLigasa (Epicentre) para crear estructuras en forma de mancuerna (Mann y Dzau (2000) *J. Clin. Investigation* **106**: 1071-1075). Se mezclaron 100 pmoles de cada oligonucleótido señuelo en un volumen de reacción de 20 µl de tampón 1 x CircLigasa suplementado con 50 µM ATP, 2,5 mM MnCl₂, 500 U CircLigasa y se incubó durante 1 hora a 60°C. La mezcla se trató con un exceso de exonucleasa I para eliminar el ADN lineal y los círculos covalentes restantes se precipitaron. Cada preparación se analizó por electroforesis en gel no desnaturante al 12% para comprobar que la mayoría de los productos eran círculos covalentes monovalentes. La muestra se resuspendió a 200 pmoles/µl en tampón TE (10 mM Tris-HCl, pH 8,0, 1 mM EDTA). 10 µl de cada señuelo se pusieron en discos de antibiótico de 3 mm que se presionaron suavemente en una capa fina de agar SNA que se había vertido en una capa confluyente de 24 horas de *S. coelicolor* M145 crecida en una placa de agar R2YE. La inducción de la producción de antibiótico resultaría en un incremento localizado de la pigmentación alrededor del disco. Alternativamente, los oligonucleótidos señuelo en forma de mancuerna se usaron para transfectar micelio de cultivos que crecen exponencialmente de *S. coelicolor* M145. La transfección implicó lavar las células en volúmenes iguales de tampón TES suplementado con 0,5% (v/v) NP-40 y 0,5% (v/v) Tritón X-100, resuspenderlas en 1/10 de su volumen original en TES más detergentes y suplementado con el oligonucleótido señuelo. Las células se incubaron a 30°C durante 15 min con mezclado suave, se diluyeron en el medio de cultivo retenido y la incubación continuó. La velocidad de crecimiento se ralentizó temporalmente después de este tratamiento pero se recuperó pronto. Se tomaron muestras posteriormente para evaluar los niveles de producción de los dos antibióticos pigmentados, actinorhodina y undecilprodigiosina (13). Se tomaron muestras simultáneamente para análisis de ARN.

7.4 Expresión génica y análisis del número de copias de los señuelos

Los análisis de expresión se realizaron por PCR cuantitativa en tiempo real (crt-PCR). La preparación del ADNc se realizó como se ha descrito previamente (Ryding *et al.* (2002) *J. Bacteriol.* **184**: 794-805). Brevemente, 1 µg de ARN se desnaturizó con calor (se incubó a 70°C durante 10 min y se puso en hielo) y se mezcló con 0,5 U Transcriptasa Inversa AVM (Amersham), 1 mM dNTP, 25 pmoles de cebador personalizado en 20 µl del tampón recomendado por el suministrador. La reacción se incubó a tres temperaturas sucesivas, 45°C, 50°C y 55°C, cada una durante 30 min. 4 µl de esta reacción se usaron como un molde en una reacción de crt-PCR de 20 µl preparada en la mezcla de reacción SYBR Greener de Invitrogen suplementada con 10% (v/v) DMSO y 25 pmoles de los cebadores directos e inversos personalizados para el gen diana (*actII-orf4*) y una referencia interna (SCO4742, una proteína hipotética conservada que se encontró que mostraba poca variación en la expresión después de análisis en micromatrices [<http://www.biomedcentral.com/1471-2164/8/261/abstract>]) (Tabla 1). La amplificación y análisis se realizaron en una máquina Chromo4 de BioRad. La determinación del número de copias de los señuelos se realizó de manera similar usando detección SYBR-verde. Los cebadores se diseñaron (Tabla 1) para detectar oligonucleótidos señuelo circularizados y su número de copias se determinó por sonicación del micelio y comparación con cantidades conocidas de señuelo tratado con exonucleasa; estos valores se corrigieron para el número de genomas en la muestra por referencia al número de copias de SCO4742 presentes.

7.5 Tabla 1. Cebadores oligonucleotídicos

Las secuencias reguladoras en *cis* en cada uno de los oligonucleótidos señuelo están subrayadas.

Nombre No.	Secuencia	SEQ.	ID.
Oligonucleótidos señuelo			
A24.1	5'- P-atatcactctcgatgctggcggttttcgcccacatcgagagtgatagttttct- 3'	8	
A24.2	5'- P-atatcagggaatgccagatgcgttttcgcatctggcattcctgatagttttct- 3'	9	
A24.3	5'- P-atactgcctctcggttaagcggttttcgcttaccgagaggcagtatagttttct- 3'	10	
A24.4	5'- P-atatgcagctcgcgtcacgcgttttcgctgcagcgagctgcatatagttttct- 3'	11	
A24.5	5'- P-ataaatctgttgagtagggcggttttcgcccactcaacagattttagttttct- 3'	12	
A24.5 desordenado	5'- P-atagcaatttagatggtggcggttttcgcccacatctaaattgctatagttttct- 3'	13	
PCR cuantitativa para el número de copias del señuelo			
A24.1 copy f	5'- cgacatcgagagtgatagtttt- 3'	14	
A24.1 copy r	5'- gccgacatcgagagtgatat- 3'	15	
PCR cuantitativa para los datos de expresión			
A24 personalizado	5'- tgtcgccccaggagacggag- 3'	16	
qA24f	5'- cttaaatcctcgaaggcgaccag -3'	17	
qA24r	5'- tcctcgagccggttctcctcg- 3'	18	
q4742f	5'- gctgggtggacatcggtct- 3'	19	
q4742r	5'- gccggtactgtgctctc- 3'	20	

Ejemplo 6

Demostración de que los oligodesoxinucleótidos señuelo pueden usarse en el contexto terapéutico para superar la resistencia a antibióticos en patógenos

- 5 La estrategia global adoptada para identificar un señuelo capaz de alterar la susceptibilidad a vancomicina, u otros mecanismos de resistencia a antibióticos que se ha caracterizado a nivel genético, se presenta en la Figura 20. Este esquema demuestra cómo se usan los TFD para contrarrestar mecanismos de resistencia conocidos en bacterias patógenas. La resistencia al antibiótico prescrito vancomicina se usa como ejemplo. Esto ocurre en cepas de *Enterococcus faecium*, un patógeno humano, y *Streptomyces coelicolor*. En las dos bacterias, la resistencia está
- 10 codificada por el operón *vanHAX* que se induce por la proteína VanR. Mediante el diseño de señuelos que interfieren con la unión de VanR a su sitio afín en el genoma es posible evitar la inducción de *vanHAX* y de esta manera volver a las bacterias susceptibles a vancomicina. En este ejemplo, los señuelos se validan funcionalmente en modelos bacterianos (no patógenos) que contienen mecanismos de resistencia similares a los patógenos o los mecanismos reales trasladados al modelo por transferencia génica horizontal. Posteriormente, el señuelo se ensaya en modelos
- 15 patógenos, tales como aislados clínicos, y finalmente se trasladan a modelos animales, tales como ratones infectados con *Enterococcus* resistente a vancomicina (VRE) o *S. aureus* resistente a vancomicina. Debe apreciarse que la estrategia podría aplicarse a un amplio rango de fenotipos bacterianos que están controlados por dichos interruptores genéticos. Éstos incluyen, pero no estarían limitados al descubrimiento de otros señuelos o combinaciones de señuelos que podrían volver a las bacterias susceptibles a otros antibióticos tales como penicilina, cloranfenicol, tetraciclina, daptomicina, etc
- 20

La estructura del operón de resistencia a vancomicina en *Streptomyces coelicolor* se muestra en la Figura 21 (De Hong et al. 2004 Molecular Microbiology 52: 1107-1121). Los operones de vancomicina consisten en cuatro operones, como se indica con las flechas. El operón *vanRS* codifica un sistema regulador con dos componentes que actúa para detectar la presencia de vancomicina (directamente o indirectamente) e inducir la expresión del operón

25 *vanHAX* que codifica el mecanismo de resistencia a vancomicina. El operón *vanHAX* es inducido por la versión fosforilada de VanR uniéndose a su promotor, VanS a su vez es la quinasa que fosforila VanR después de la detección de vancomicina. Se espera que nuestro señuelo interrumpa la unión de VanR fosforilado al sitio en el promotor *vanH*.

30 Debe indicarse que la fuente evolutiva de los genes de resistencia en *Enterococcus faecalis*, que es la causa de algunas de las infecciones resistentes a vancomicina observadas en clínica, se piensa que es un actinomicete que produce vancomicina.

La secuencia del señuelo vanH5 usado se muestra más adelante, con el sitio de unión para VanR fosforilado en mayúsculas:

5'-P- ata tctatatgaa gcgacgtggt cgatgagcgg cagcg tttt cgctg CGG CTC ATC GAC
CAC GTC GCT TCA TAT AGA tatag tttt ct- 3' SEQ ID. NO. 21

El oligonucleótido se preparó como un señuelo en forma de mancuerna circular como se muestra en la Figura 22. El oligonucleótido se resuspendió a una concentración final de 100 pmoles/μl en un tampón de ADN ligasa T4 (como suministra el fabricante de la enzima, New England Biolabs) y 400 U de la ADN ligasa T4 y se incubó a 16 grados Centígrados durante varios tiempos. Durante la incubación el oligonucleótido formó una estructura en horquilla que, debido a la actividad de la ligasa, puede convertirse en una molécula de ADN "circular" monocatenaria unida covalentemente. El grado de esta ciclación depende del tiempo de incubación con la ligasa; la incubación más larga (16 h) dio lugar a una conversión casi completa. Por lo tanto, después de la incubación del oligonucleótido con ADN ligasa T4 durante 0,1 2, 4, 6 y 16 h (a-f, respectivamente), se tomaron alícuotas de cada reacción, se trataron con calor para inactivar la enzima y el ADN se recuperó por precipitación con etanol antes de analizarlo por electroforesis en gel de poliacrilamida al 6% y visualización con tinción de bromuro de etidio. Los carriles a-f muestran la conversión de moléculas de ADN monocatenarias lineales con migración más lenta en un dúplex circular cerrado en forma de mancuerna circular que migra a una mayor velocidad (carril f).

El señuelo está en una configuración "en forma de mancuerna circular" (Ahn JD, Kim CD, Magae J, Kim YH, Kim HJ, Park KK, Hong S, Park KG, Lee IK, Chang YC (2003) Biochemical Biophysical Res Comm 310: 1048-1053), aunque ésta no se requiere. Pueden usarse otras formas de oligodesoxinucleótido, incluyendo pero no limitado a, por ejemplo (a) oligodesoxinucleótidos bicatenarios más largos que comprenden esta secuencia, de manera que las exonucleasas necesitarían reducir sustancialmente la longitud del oligonucleótido antes de la pérdida de la función señuelo; (b) oligodesoxinucleótidos que comprenden bases o azúcares modificados para conferir una mayor resistencia a nucleasas; (c) nucleótidos 2'OH o aminos unidas al extremo de los oligonucleótidos, que bloquearían la actividad exonucleasa (que es todo lo que hacen las estructuras en forma de mancuerna); (d) moléculas de ADN bicatenarias circulares pequeñas; o (e) moléculas multiméricas que comprenden múltiples copias de la secuencia señuelo activa.

Esencialmente, lo que se requiere es que el ODN usado tenga una región bicatenaria que incorpora la secuencia diana en el promotor vanH flanqueada por pequeñas estructuras en horquilla u otras secuencias. Para la estructura en forma de mancuerna circular, la molécula se sintetiza como un oligonucleótido lineal y se cicla por incubación con ADN ligasa T4.

En la Figura 23 se proporciona un diagrama que muestra la introducción de un señuelo denominado vanH5 en un cultivo líquido de la cepa de *Streptomyces coelicolor* M600. La cepa de *S. coelicolor* M600 se creció en medio líquido MMCGT (Molecular Microbiology 52: 1107-1121), el crecimiento se midió registrando la absorbancia del cultivo a 430 nm (Densidad Celular) y se representó como una función del tiempo de incubación. Se monitorizó el crecimiento de cuatro cultivos: (1) sin nada añadido al medio (diamantes "M600"); (2) el medio suplementado con una concentración sub-lethal de 20 μg/ml del antibiótico vancomicina a 0 h (cuadrados "más 20 μg/ml VAN"); como en (2) pero con el oligonucleótido señuelo H5 añadido a una concentración final de 64 pM seguido de 20 h de incubación (triángulos "VAN más 64 pM H5"); como en (2) pero con el oligonucleótido señuelo H5 añadido a una concentración final de 256 pM seguido de 20 h de incubación (círculos "VAN más 256 pM H5").

A partir de estos datos, es evidente que M600 crece de manera comparable en presencia o ausencia de 20 μg/ml de vancomicina (añadida o no añadida al inicio del proceso de cultivo), confirmando que la cepa tiene resistencia a vancomicina. En presencia de 20 μg/ml de vancomicina, la adición del señuelo ciclado vanH5 a una concentración final de 64 pM tiene un pequeño efecto detectable, mientras que la adición de 256 pM causa que las células paren de crecer durante un periodo de aproximadamente 18 h. Esto puede referirse coloquialmente como una "ventana terapéutica", en la que los señuelos han entrado en las células, han interferido con la unión del regulador vanR al promotor vanH y evitado de esta manera la inducción del mecanismo de resistencia. Después de 18 h, nuestra evidencia de otros sistemas sugiere, que ocurre la degradación de los señuelos (por endonucleasas en la célula), lo que puede influir en su eficacia.

Respecto a la razón por la que el señuelo vanH parece tener un efecto bacteriostático, creemos que esto se debe a que el mecanismo de resistencia es una modificación de la pared de la célula bacteriana que evita la acción de la vancomicina. Cuando la vancomicina se añade al cultivo, esto es en el inicio del experimento. El señuelo se añade a las 20 horas. En este momento, las bacterias habrán estado creciendo con una pared celular modificada. Después de la adición del señuelo, desde ese punto hasta que el señuelo se disipa, la capacidad de modificar la pared celular bacteriana está bloqueada. Como resultado, cualquier crecimiento bacteriano nuevo se vuelve sensible a vancomicina, aunque las células más viejas son capaces de persistir. Después del agotamiento del señuelo, las células persistentes con resistencia a vancomicina residual son capaces de proliferar de nuevo. La adición del

señuelo antes de la vancomicina tendría poco efecto, ya que el operón de resistencia no está activado. Prevemos, sin embargo, que la administración de los ODN señuelo al mismo tiempo induce un efecto bactericida.

Así, este ejemplo demuestra la sensibilidad de *S. coelicolor* a vancomicina en presencia de los TFD y no en su ausencia. Esto sería aceptado por los expertos en la técnica como un suplente aceptable para demostrar este efecto en un patógeno, tal como *E. faecalis*. El operón *van* de *E. faecium* se induce por el mismo mecanismo (unión de vanR a vanH), incluso aunque la homología a nivel de secuencia entre los dos sistemas no es muy alta. Los expertos en la técnica serán capaces de modificar fácilmente la secuencia señuelo para uso en el patógeno.

El señuelo vanH5 es un ejemplo de oligodesoxinucleótidos que pueden unirse a un factor de transcripción procariota de manera que la unión del factor de transcripción a su diana afín en el genoma del procariota disminuye o desaparece.

2. Diseño de TFD para combatir la resistencia a antibióticos en patógenos

El principio de cómo los TFD pueden usarse para restaurar la sensibilidad a vancomicina se ha discutido anteriormente. Para situaciones en las que los genes responsables se desconocen, se usan las bibliotecas n[snare] universales o bibliotecas personalizadas creadas a partir de ADN genómico del patógeno proporcionadas por esta invención. Como se muestra en la Figura 6, la biblioteca puede introducirse en las bacterias patogénicas en condiciones de laboratorio o los elementos genéticos que determinan la resistencia pueden clonarse e introducirse en una cepa de laboratorio no patogénica para actuar como un sistema modelo, que puede transformarse. Las células se cultivarán en ausencia del antibiótico elegido y la biblioteca se introduce en las células por transformación en cultivo líquido, la muestra se divide y el antibiótico se añade a la mitad de la muestra y la incubación continúa. Las poblaciones de células se recuperan y los TFD concatamerizados se amplifican de los plásmidos por PCR. Estas dos poblaciones se sustraen para aislar los TFD que no están en la muestra tratada con antibiótico, y por lo tanto los que confieren sensibilidad. Esta población enriquecida de TFD se vuelve a clonar y el proceso de selección se repite hasta que está lo suficientemente enriquecido en TFD capaces de volver a la célula sensible a antibiótico.

Las dianas potenciales para dicha estrategia incluirían la investigación de los mecanismos de resistencia de muchos de los antibióticos clínicamente prescritos y los futuros para los que la resistencia a antibióticos empieza a limitar su eficacia. Los ejemplos de los antibióticos actuales que se considerarían para investigación usando las metodologías n[snare] y tratamiento posterior con TFD para vencer la resistencia a antibióticos incluirían: los de la clase de antibióticos conocida como aminoglicósidos (tal como kanamicina); de los carbapenemos (tal como meropenem); las cefalosporinas (tal como cefepime); los glicopéptidos (tal como vancomicina y daptomicina); las penicilinas (tal como ampicilina, carbenicilina y penicilina); los antibióticos polipeptídicos (tal como polymixina B); las quinolonas (tal como levaquin); las sulfonamidas (tal como Bactrim); las tetraciclinas (tal como tetraciclina); y de forma diversa, cloranfenicol, rifampicina y Zyvox.

Ejemplo 7

7.1 TFD marcados con colesterol en presencia de Estreptolisina-O

Usando cebadores oligonucleotídicos, uno de los cuales tiene una modificación 5' con colesterol y el otro una modificación similar en su extremo 5' o alguna otra (tal como un agente de tinción fluorescente, tal como Cy5, de manera que la captación del TFD puede medirse fácilmente) se prepara un TFD por PCR. Si el TFD se ha clonado previamente en un vector (pGEMT-Easy) los cebadores se diseñan para hibridar con las secuencias del vector que flanquean inmediatamente el inserto, por ejemplo:

SEQ ID NO:22 Chol_Tef: 5' Colesterol-TEG-ggc cgc cat ggc ggc cgc ggg aat tc

SEQ ID NO:23 Cy5_Ter: 5' Cy5- AGG CGG CCG CGA ATT CAC TAG TG.

Si la secuencia que se va a usar para un TFD no se ha clonado puede sintetizarse directamente (si es lo suficientemente corta) e hibridarse para formar el TFD o amplificarse directamente a partir del ADN genómico usando cebadores diseñados para hibridar en el TFD.

El producto de PCR se precipita con etanol y se resuspende en tampón TE a una concentración de 500-1.000 ng/μl. Típicamente, los ensayos de sensibilidad a antibiótico se realizan usando placas de 96 pocillos, conteniendo cada pocillo 200 μl de caldo. Por ejemplo, en el caso de *Enterococcus faecium* este caldo es medio BHI (de Becton Dickinson) suplementado con 0,2 U/ml de Estreptolisina-O (Sigma) y 5 μg/ml de antibiótico vancomicina y se inocula con una cepa resistente a vancomicina de *E. faecium*.

7.2 Preparación de Estructuras en Forma de Mancuerna por PCR

Los señuelos en forma de mancuerna son ADN monocatenario cerrado covalentemente caracterizados por un centro bicatenario, que contiene el sitio de unión para el factor diana, flanqueado por estructuras en horquilla. La horquilla

estabiliza el señuelo evitando la acción de exonucleasas que de otra manera degradarían el polinucleótido señuelo. Por lo tanto, los señuelos En forma de mancuerna (DB) se llaman así por su forma característica.

Los DB se preparan por PCR usando como molde un plásmido derivado de pGEMTEasy que contiene el sitio de unión diana, como se describe en 7.1. Los cebadores usados en la amplificación son:

(SEQ ID NO: 24) DBTEf: 5' P- CTTGG TTTT CCAAG AGAAGAGC ccg cca tgg cgg ccg

cgg gaa ttc

(SEQ ID NO: 25) DBTEr: P- CCG TCT TTT TGA CGG CGA AGA GCA GGC GGC CGC

GAA TTC ACT AGT GA.

La parte de los cebadores que formará las horquillas está subrayada. La amplificación con el vector apropiado proporciona el producto de ADN mostrado en la Figura 24, en el que la parte del DB que se unirá al factor de transcripción se proporciona por "NNN NNN". Las secuencias en negrita representan un sitio de unión para la enzima de restricción de mellado Nt.BspQ1. En la segunda parte de la Figura 24, se muestra la consecuencia de digerir el producto de PCR con Nt.BspQ1; esto expone las estructuras en horquilla como regiones monocatenarias que formarán una horquilla y pueden ligarse posteriormente por tratamiento con ADN ligasa T4 para formar un círculo cerrado covalentemente y DB.

7.3 Preparación de oligonucleótidos en forma de mancuerna por digestión por restricción del plásmido

Alternativamente, pueden prepararse DB clonando el producto de PCR como mostrado en la Figura 24 en un vector de clonación de PCR adecuado, tal como pGEMTEasy, lo que confirma su identidad y preparación del plásmido. El plásmido puede digerirse para liberar el inserto que se digiere adicionalmente con Nt.BspQ1 para liberar el fragmento mostrado en la segunda parte de la Figura 24. Éste puede tratarse de manera similar con ADN ligasa T4 con el fin de cerrar covalentemente la molécula de ADN y formar un DB.

La ventaja de esta estrategia es que es más factible, tanto prácticamente como económicamente, para aumentar de escala si el DB se requiere en grandes cantidades.

7.4 Transfección con el agente R9-colesterol.

El R9-colesterol se ha descrito por sus propiedades de ayudar en la transfección de moléculas de siARN (u otra terapia basada en ácido nucleico) y semejantes en células eucariotas (Patente US: 20070207966). Aquí describimos su utilidad para transfectar varias bacterias con TFD.

El R9-colesterol, que consiste en una molécula de colesterol unida a una cadena lineal de nueve D-argininas, se sintetizó como se ha descrito previamente (Kim W.J. et al., Mol. Ther 2006 14: 343-350). Los TFD se mezclaron con cantidades crecientes de R9-colesterol en un tampón basado en TE suplementado con 5% glucosa. La mezcla se incubó a temperatura ambiente durante 1 hora y se usó directamente en transfecciones o se analizó por electroforesis en gel de agarosa. Típicamente, se usó la cantidad mínima de R9-colesterol que causaba que el complejo con ADN no corriera en el gel; es decir, la carga del núcleo de ácido nucleico se había neutralizado por la unión de poli-arginina. La molécula de colesterol ayuda a que el TFD se asocie con la membrana bacteriana y de esta manera entre en la célula.

Los conjugados TFD/R9-colesterol se mezclaron a varias concentraciones en 200 µl de cultivo en una placa de 96 pocillos. Por ejemplo, en el caso de *Enterococcus faecium* este caldo es medio BHI (de Becton Dickinson) suplementado con 0,2 U/ml de Estreptolisina-O (Sigma) y 5 µg/ml de antibiótico vancomicina y se inoculó con una cepa resistente a vancomicina de *E. faecium*.

Ejemplo 8

Uso de una secuencia señuelo Van para sensibilizar *Enterococcus faecium* a vancomicina

8(a) Uso de TFD marcados con colesterol en presencia de Estreptolisina-O

Los polinucleótidos señuelo marcados con Colesterol/Cy5 se prepararon como en el ejemplo 7.1. Los ensayos se realizaron usando placas de 96 pocillos, conteniendo cada pocillo 200 µl de caldo que consiste en medio BHI (de Becton Dickinson) suplementado con 0,2 U/ml de Estreptolisina-O (Sigma) y 5 µg/ml de antibiótico vancomicina y se inoculó con una cepa resistente a vancomicina de *E. faecium*.

Se añadió 1 µl de varias concentraciones de un señuelo marcado con Colesterol/Cy5 a cada pocillo y se monitorizó su efecto en el crecimiento bacteriano de *E. faecium* midiendo la absorbancia del caldo a intervalos durante la incubación. Las placas se incubaron a 37°C con agitación y se tomaron lecturas de absorbancia (a 450 nm) usando un lector de placas. Se ensayaron dos señuelos: VAN contiene el elemento regulador que controla la inducción de resistencia a antibiótico de tipo VanA; CON es una secuencia señuelo que no aparece en el genoma de *E. faecium* y se usó como control negativo. La secuencia de la secuencia señuelo VAN (la parte clonada en el vector pGEMTEasy para crear un plásmido usado en la etapa de amplificación por PCR con los cebadores Chol_Tef y Cy5_Ter) fue:

VAN TFD 5' AAA AAA GAA TCA TCA TCT TAA GAA ATT CTT AGT CAT TTA 3'
(SEQ ID NO:26)

- 10 Las curvas de crecimiento resultantes se muestran en la Figura 25. Todos los puntos de datos se realizaron en triplicado. Es evidente que el tratamiento con el TFD VAN a concentraciones tan bajas como 40 nM resensibilizó la cepa de *E. faecium* a vancomicina, mientras que el control negativo CON no tuvo efecto. La captación de los TFD se confirmó con cPCR y microscopía de fluorescencia.

8(b) Uso de transfección con el agente R9-colesterol

- 15 La disposición usó una placa de 96 pocillos, conteniendo cada pocillo 200 µl de cultivo, que consiste en medio BHI (de Becton Dickinson) suplementado con 0,2 U/ml de Estreptolisina-O (Sigma) y 5 µg/ml de antibiótico vancomicina y se inoculó con una cepa resistente a vancomicina de *E. faecium*.

- 20 Se añadió 1 µl de varias concentraciones de los conjugados TFD/R9-colesterol a cada pocillo y se monitorizó su efecto en el crecimiento bacteriano midiendo la absorbancia del caldo a intervalos durante la incubación. Las placas se incubaron a 37°C con agitación y se tomaron lecturas de absorbancia (a 595 nm) usando un lector de placas. Se ensayaron dos TFD en forma de mancuerna: VAN contiene el elemento regulador que controla la inducción de resistencia a antibiótico de tipo VanA (que contiene la misma secuencia que Seq. 26); CON es una secuencia señuelo que no aparece en el genoma de *E. faecium* y se usó como control negativo (que contiene la misma secuencia que CON en el Ejemplo 8(a)).

- 25 Las curvas de crecimiento resultantes se muestran en la Figura 26. Todos los puntos de datos se realizaron en triplicado. Es evidente que el tratamiento con el TFD VAN a concentraciones tan bajas como 40 nM resensibilizó la cepa de *E. faecium* a vancomicina, mientras que el control negativo no tuvo impacto negativo en el crecimiento celular.

REIVINDICACIONES

1. Un polinucleótido señuelo para uso en el tratamiento de infecciones bacterianas, en el que el polinucleótido comprende un sitio de unión para un factor de transcripción y el sitio de unión no está unido operativamente a un gen.
2. Un polinucleótido señuelo según la reivindicación 1, en el que el factor de transcripción es un regulador de la expresión de uno o más genes de resistencia a antibiótico en un procariota o eucariota.
3. Un polinucleótido señuelo según la reivindicación 2, en el que el uno o más genes resistentes a antibiótico son resistentes a un antibiótico seleccionado de uno o más: los aminoglicósidos; los carbapenemos; las cefalosporinas; los glicopéptidos; las penicilinas; los antibióticos polipeptídicos; las quinolinas; las sulfonamidas; o las tetraciclinas.
4. Un polinucleótido señuelo según una cualquiera de las reivindicaciones 1 a 3, en el que el tratamiento de la infección bacteriana comprende el uso de uno o más antibióticos.
5. Un polinucleótido señuelo según una cualquiera de las reivindicaciones 1 a 4, en el que la infección bacteriana es:
 - (a) una bacteria gram positiva de un género seleccionado de: Actinomyces; Streptomyces; Mycobacteria; Clostridium; Bacillus; Listeria; Staphylococcus; Streptococcus; Enterococcus;
 - o
 - (b) una bacteria gram negativa de un género seleccionado de: Enterobacteriaceae; Pseudomonas; Moraxella; Helicobacter; Stenotrophomonas; Bdellovibrio, bacterias de ácido acético; Legionella; Cyanobacteria; Spirochaetes; Bacterias verdes del azufre; y Bacterias verdes no del azufre,
- preferiblemente la infección bacteriana es un patógeno seleccionado de: Mycobacterium tuberculosis; Mycobacterium bovis; Mycobacterium africanum; Mycobacterium microti; Mycobacterium leprae; Clostridium difficile; Clostridium botulinum; Clostridium perfringens; Clostridium tetani; Salmonella sp.; Escherichia coli; Enterococcus faecium; Enterococcus faecalis; Neisseria gonorrhoeae; Neisseria meningitidis; Moraxella catarrhalis; Hemophilus influenza; Klebsiella pneumoniae; Legionella pneumophila; Pseudomonas aeruginosa; Proteus mirabilis; Enterobacter cloacae; Serratia marcescens; Helicobacter pylori; Salmonella enteritidis; y Salmonella typhi.
6. Un polinucleótido señuelo según una cualquiera de las reivindicaciones 1 a 5, en el que la infección bacteriana es un patógeno resistente a vancomicina o un modelo no patogénico de resistencia a vancomicina.
7. Un polinucleótido señuelo según una cualquiera de las reivindicaciones 1 a 6, en el que el polinucleótido señuelo comprende: ADN bicatenario circular o un oligonucleótido lineal, preferiblemente una estructura en forma de mancuerna circular; y/o al menos un elemento de estructura secundaria; y/o más de una copia del sitio de unión del factor de transcripción; y/o secuencia adicional al sitio o sitios de unión; y/o bases o azúcares modificados para incrementar la resistencia a nucleasas del polinucleótido; y/o un plásmido o biblioteca de plásmidos.
8. Un polinucleótido señuelo según una cualquiera de las reivindicaciones 1 a 7, en el que el sitio de unión del factor de transcripción en el polinucleótido señuelo comprende la secuencia de SEQ ID NO: 21 o SEQ ID NO: 26.

Figura 1

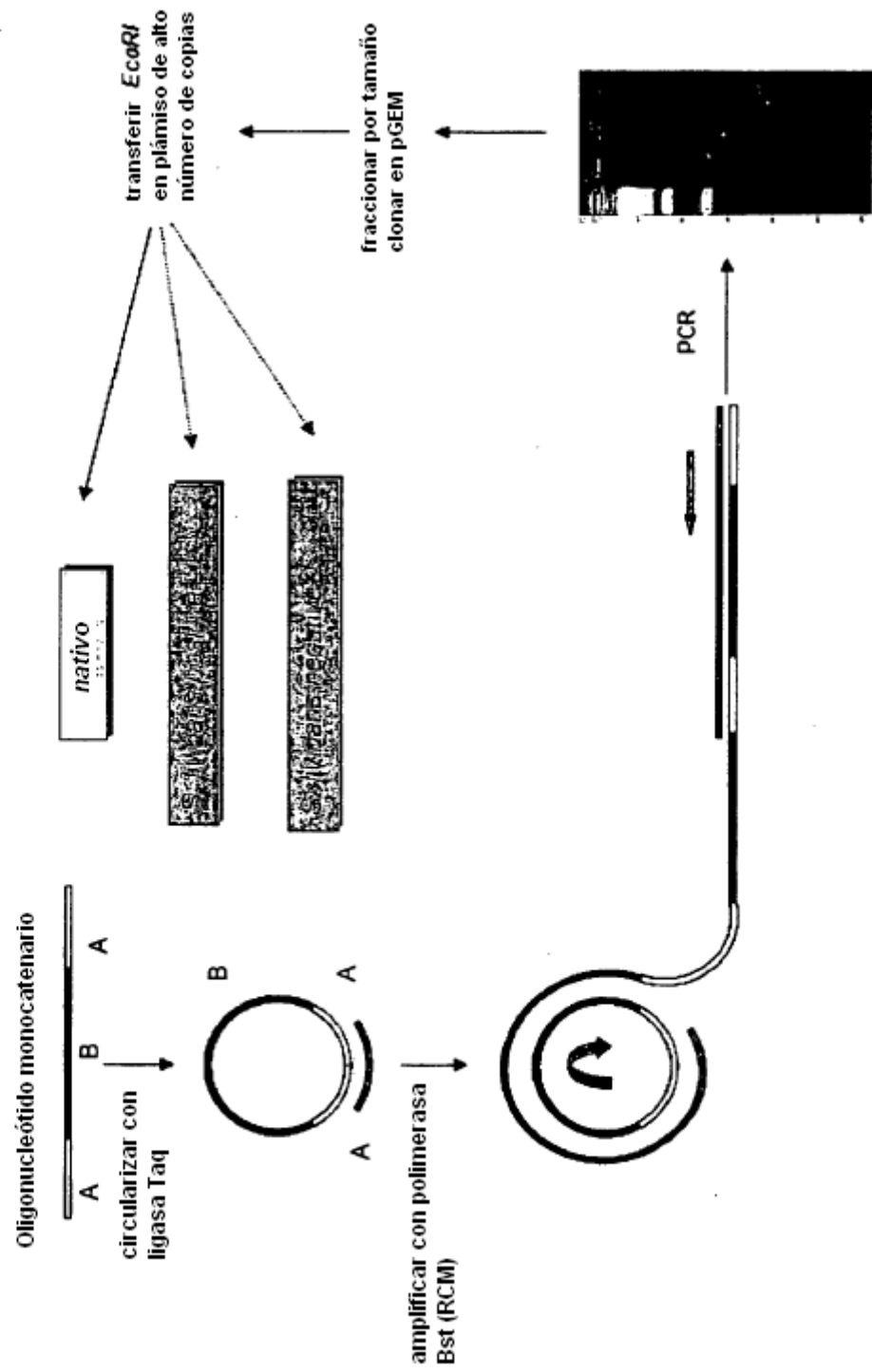


Figura 2

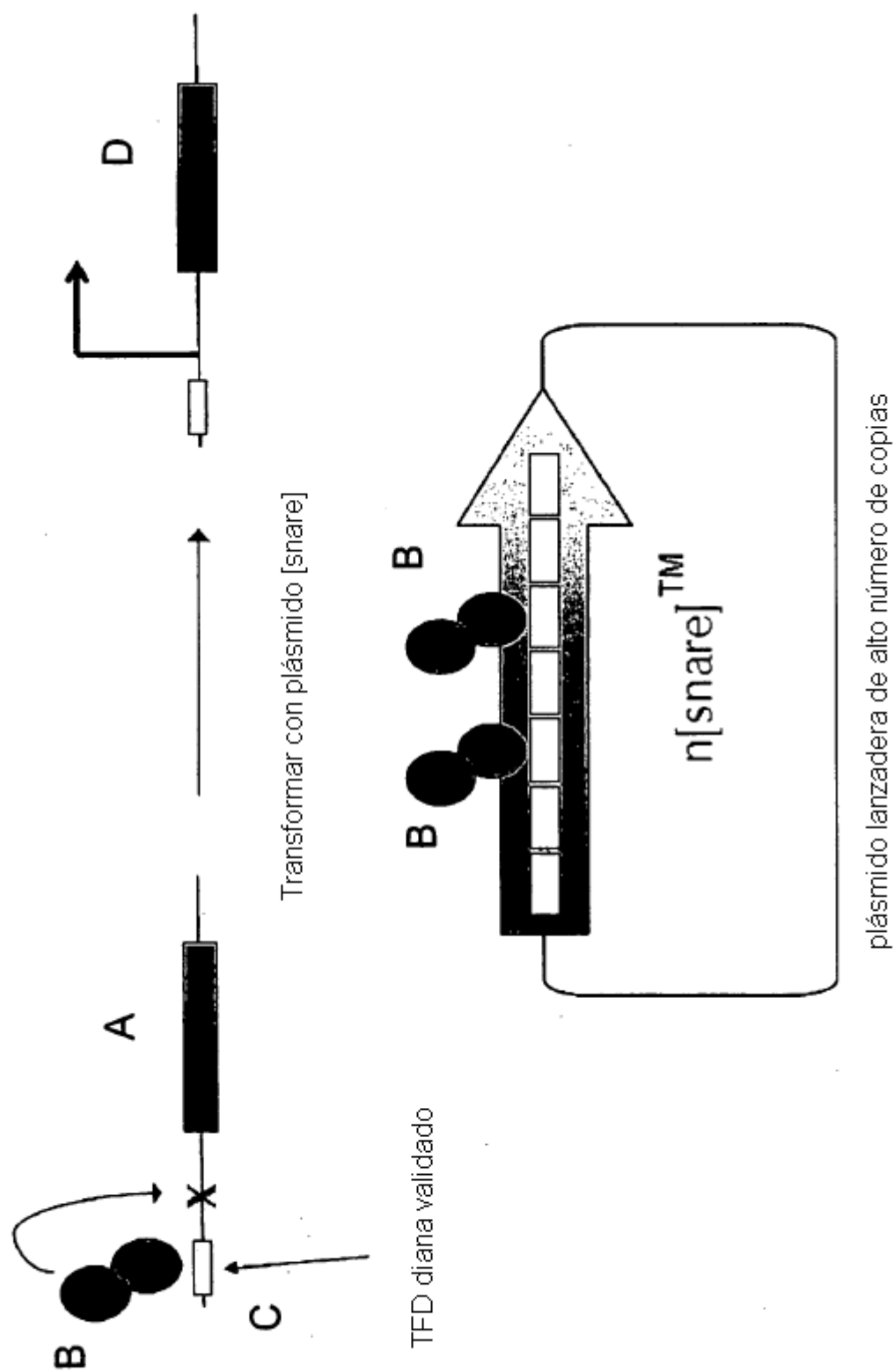


Figura 3

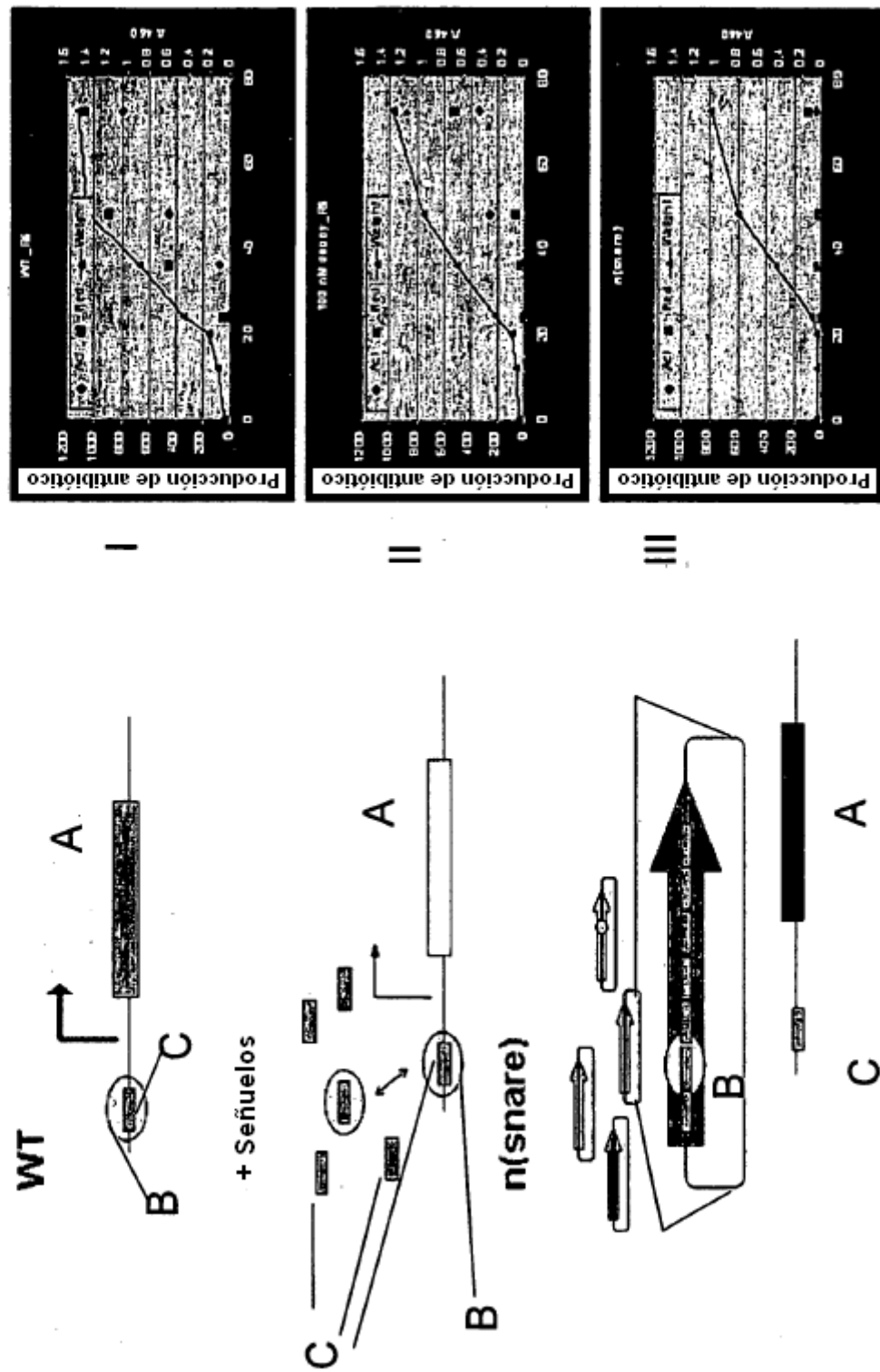


Figura 4

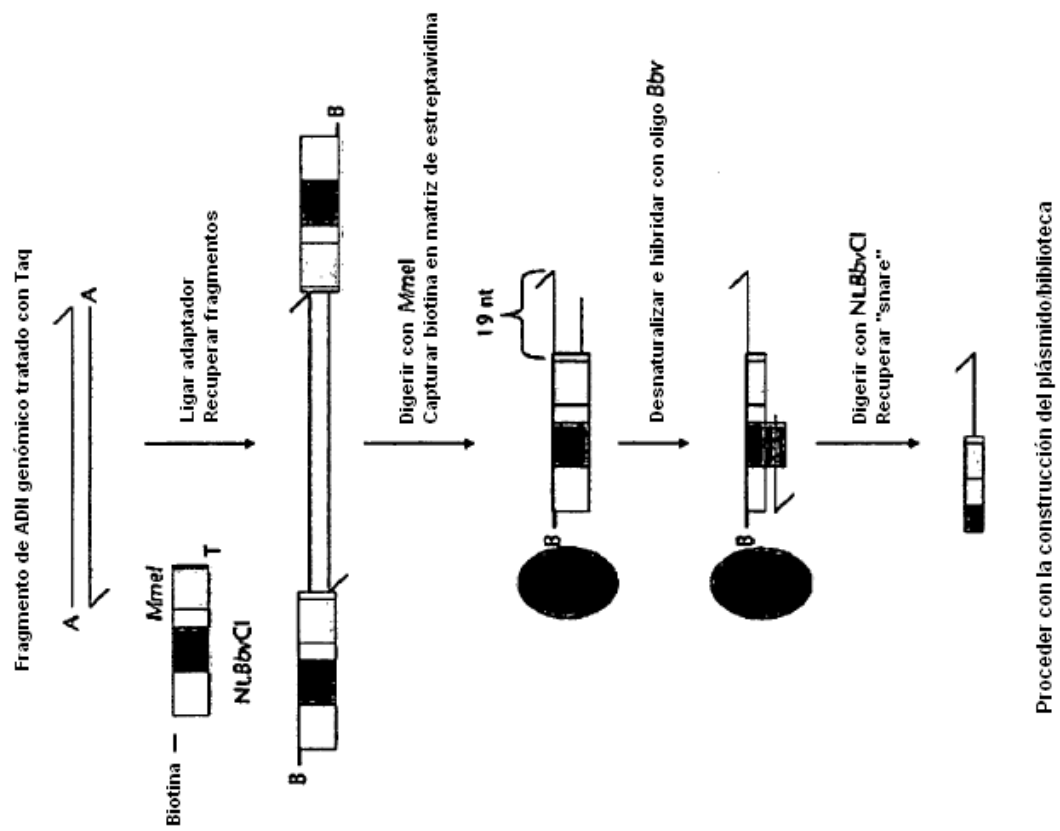
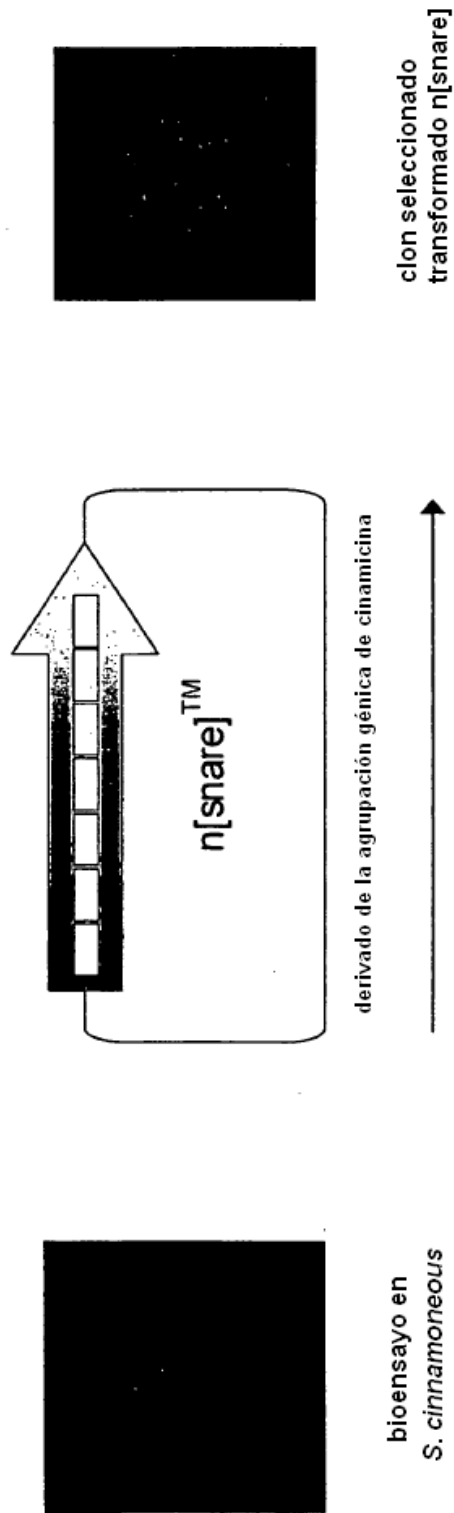


Figura 5

Las bibliotecas n[snare] pueden crearse a partir de fragmentos genómicos, genomas completos u oligonucleótidos aleatorios para maximizar el número de los TFD diana que contienen



Bibliotecas n[snare] para regular al alza la producción de antibióticos

Figura 6

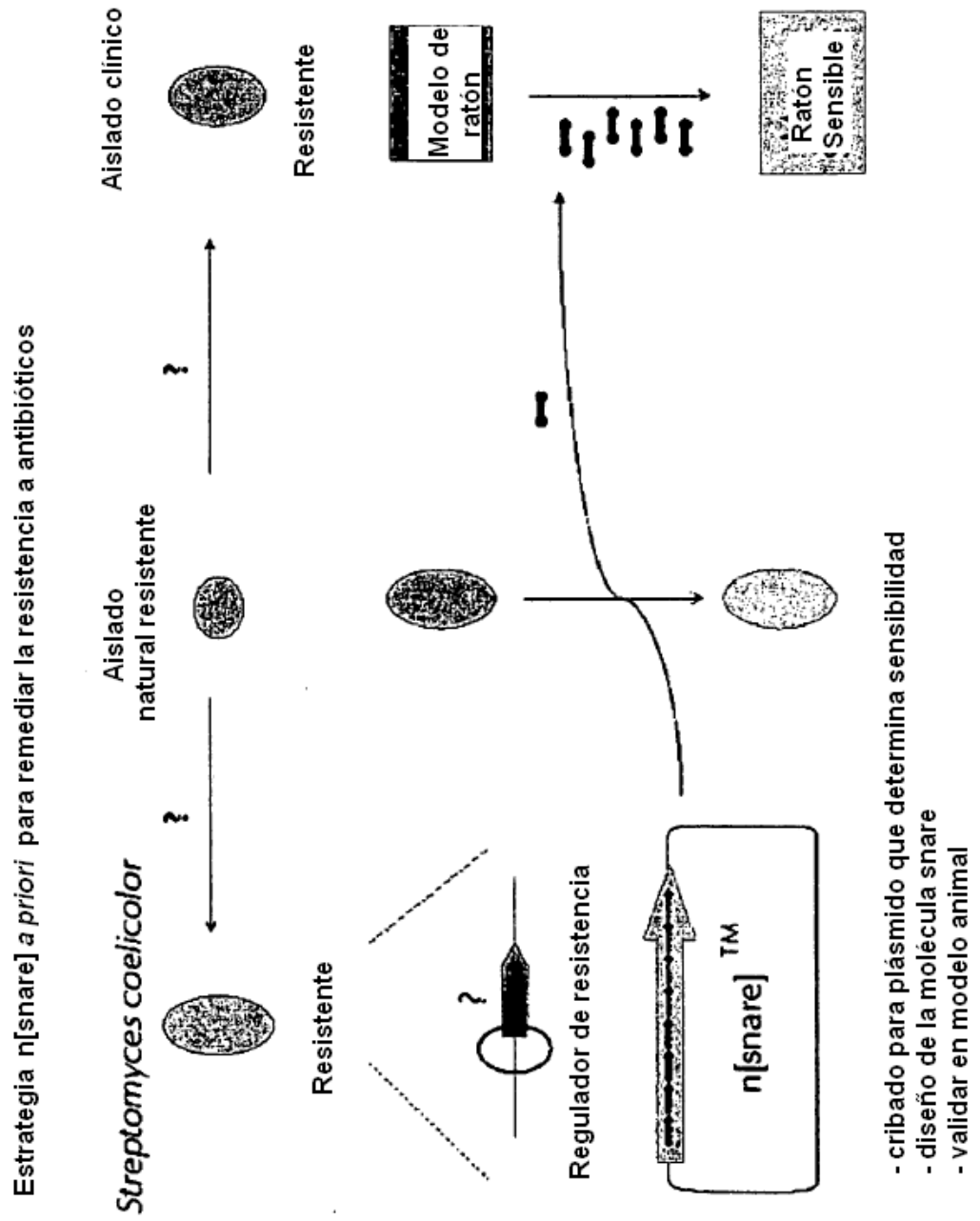


Figura 7

Sistemas informadores para adaptar la estrategia de biblioteca n[snare] para seleccionar reguladores negativos

Sensible a Kanamicina

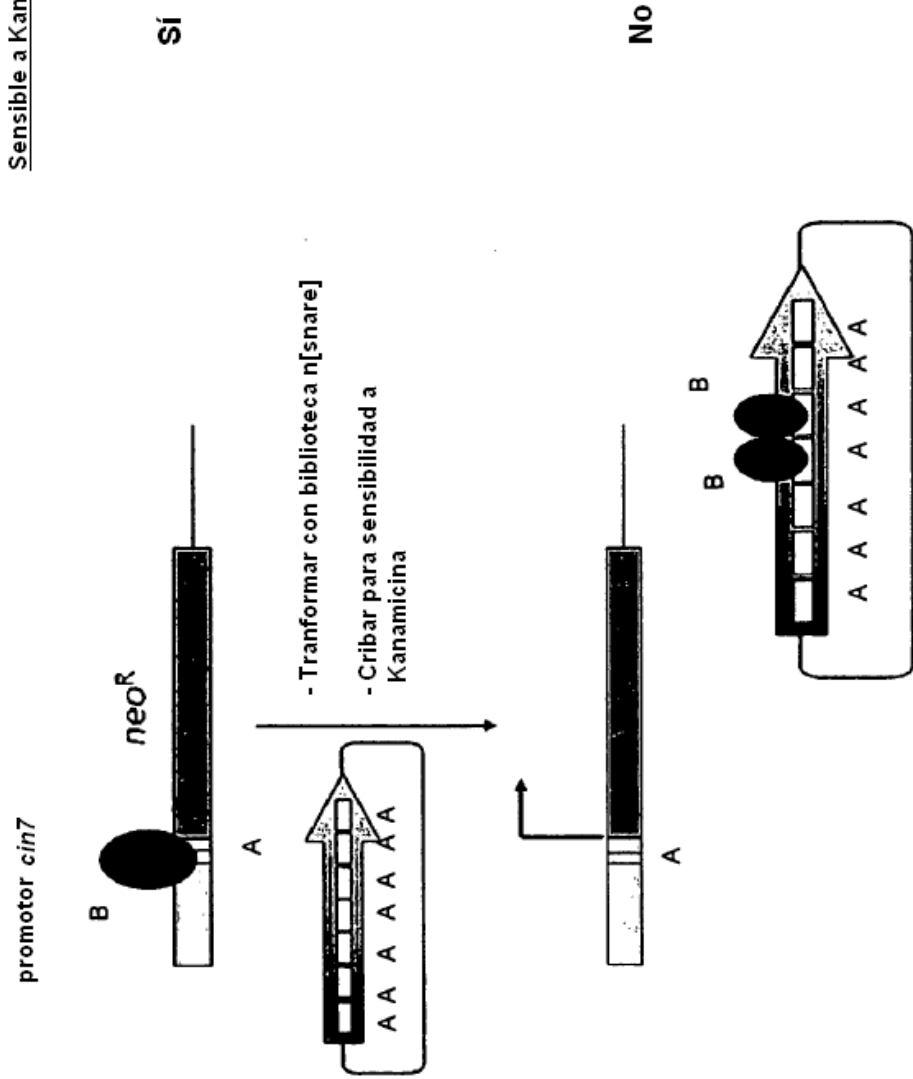
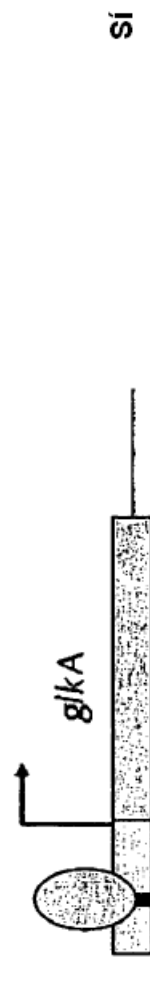


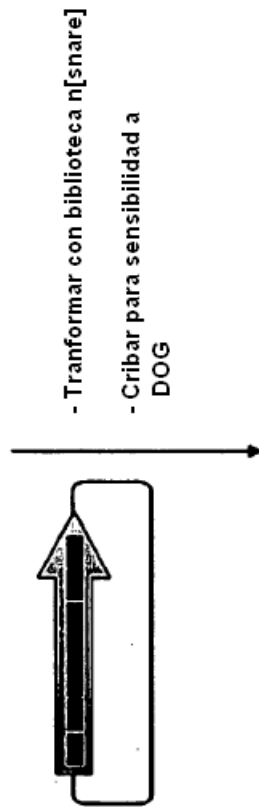
Figura 8

Sistemas informadores para adaptar la estrategia de biblioteca n[snare] para seleccionar reguladores positivos

Sensible a DOG



Si



No

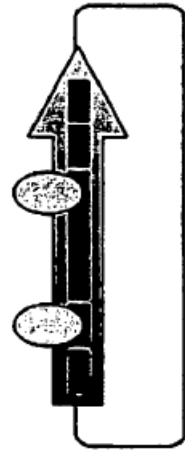


Figura 9

- Seleccionar 96 colonias super-resistentes
- Disponer en 96 pocillos
- Realizar 3 repeticiones independientes
- Hibridar secuencialmente con sondas creadas de muestras combinadas de una única repetición

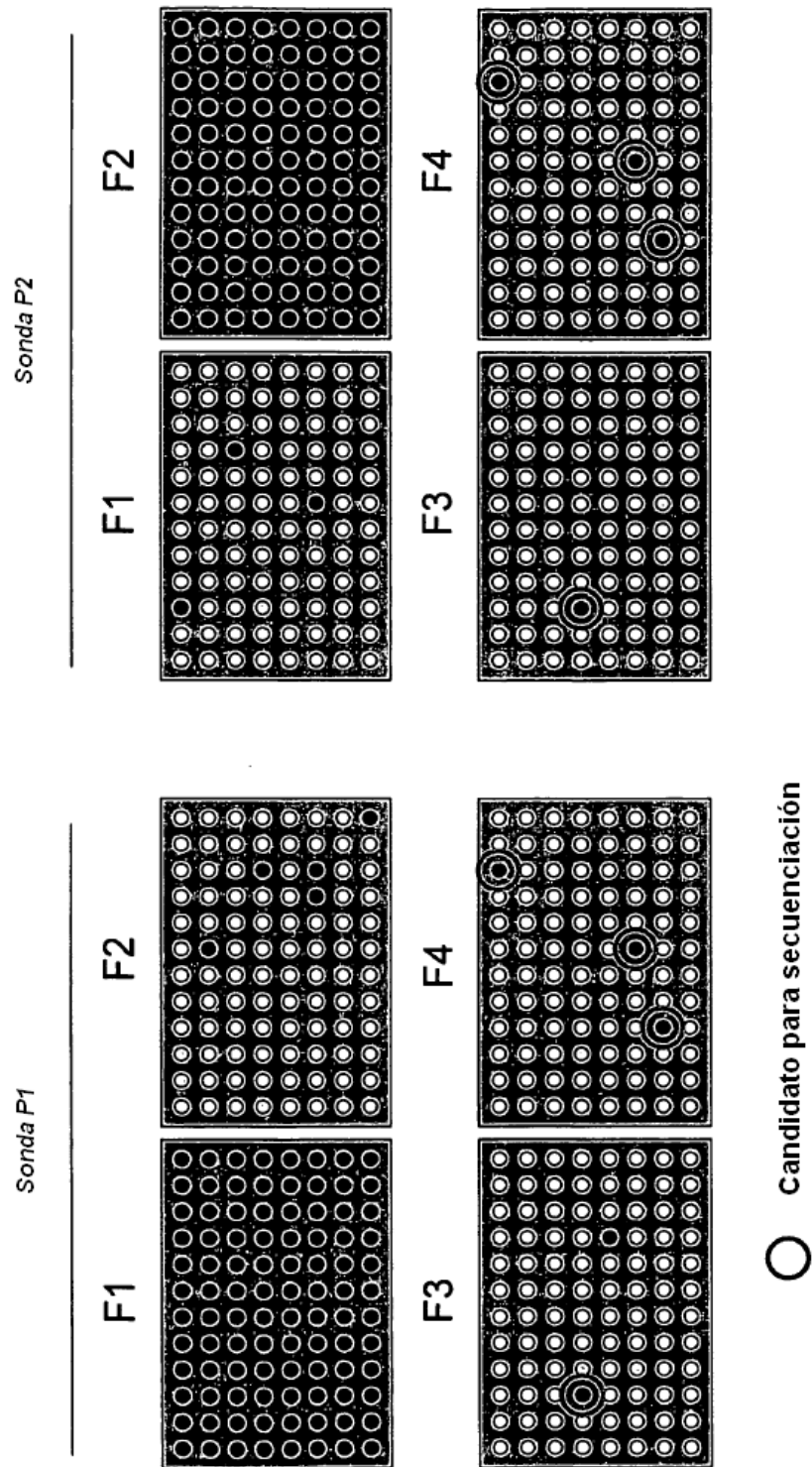


Figura 10

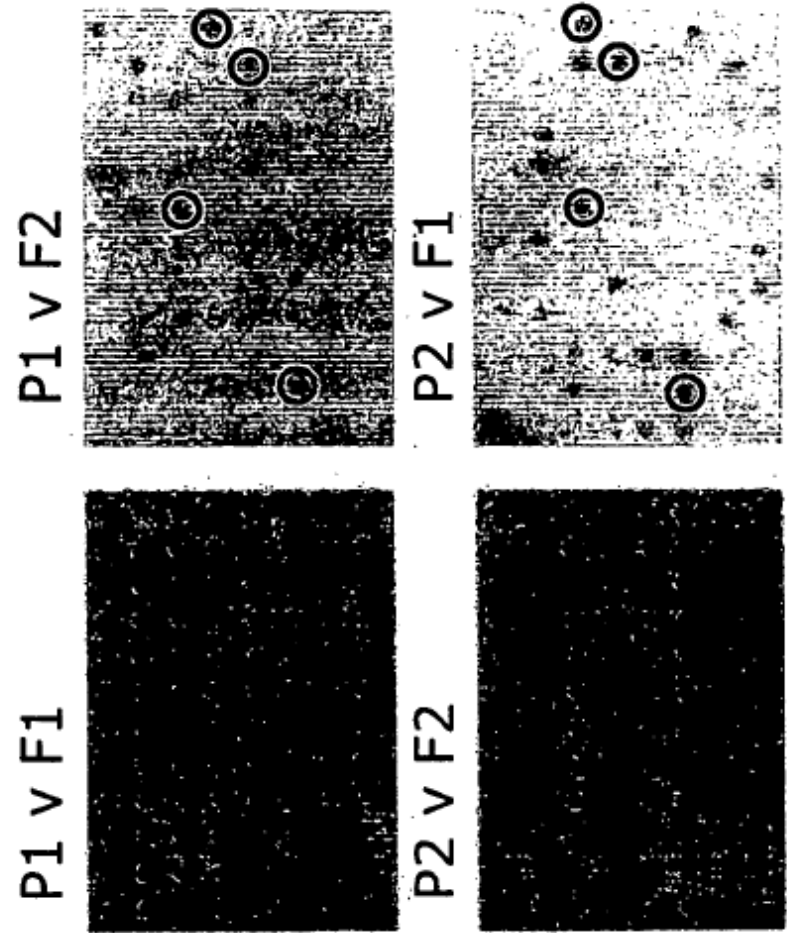




Figura 11

Figura 12

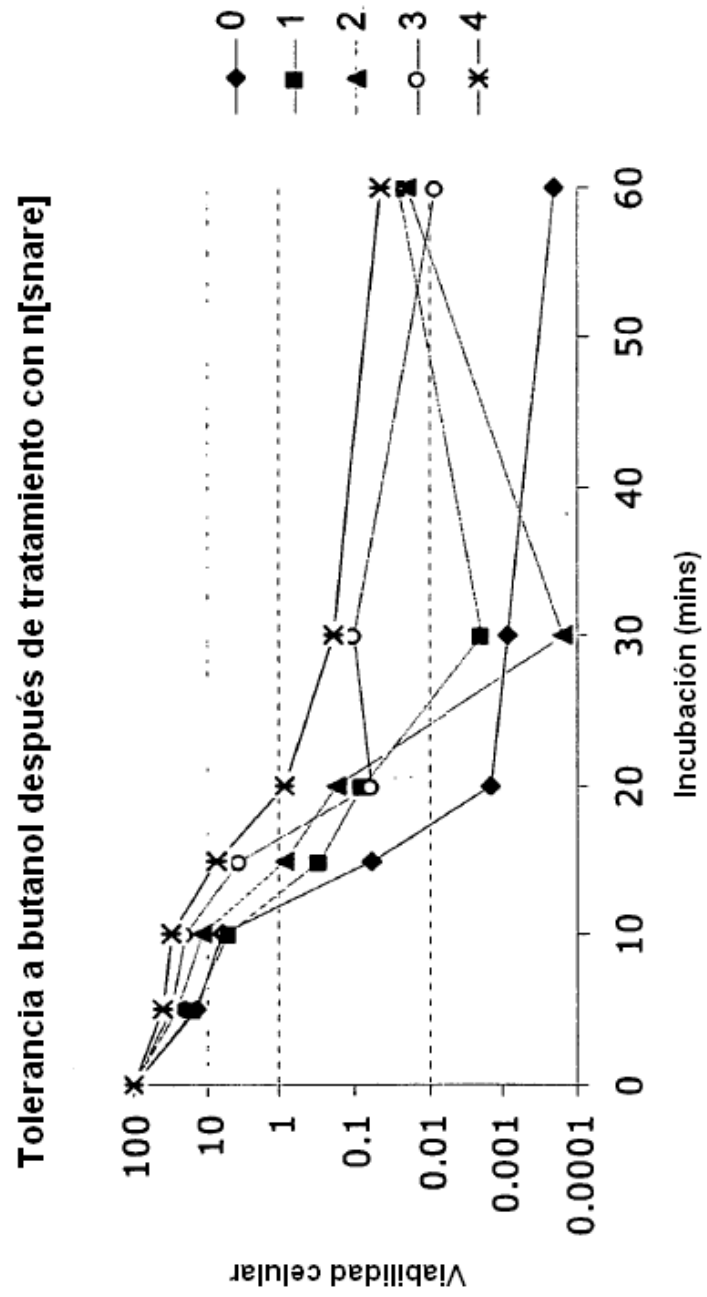
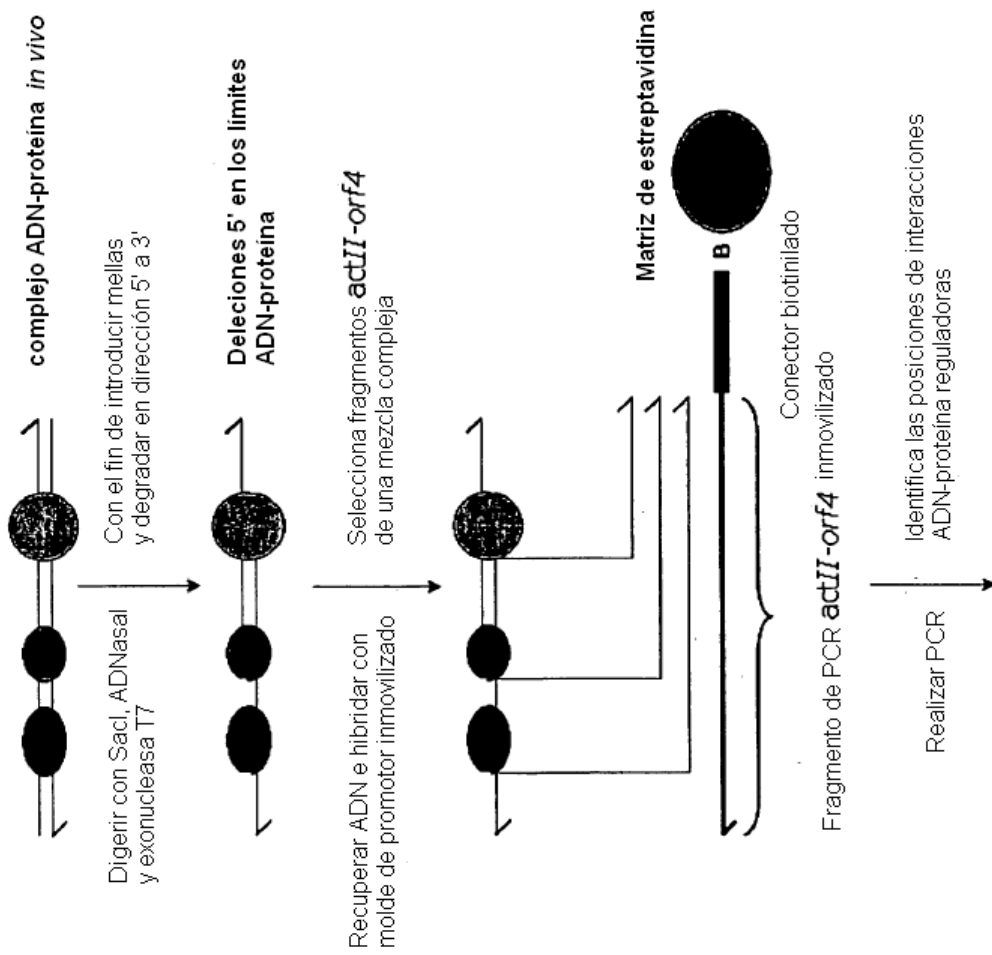


Figura 13



Analizar los fragmentos por PAGE para mapear las posiciones de los límites ADN-proteína

Figura 14

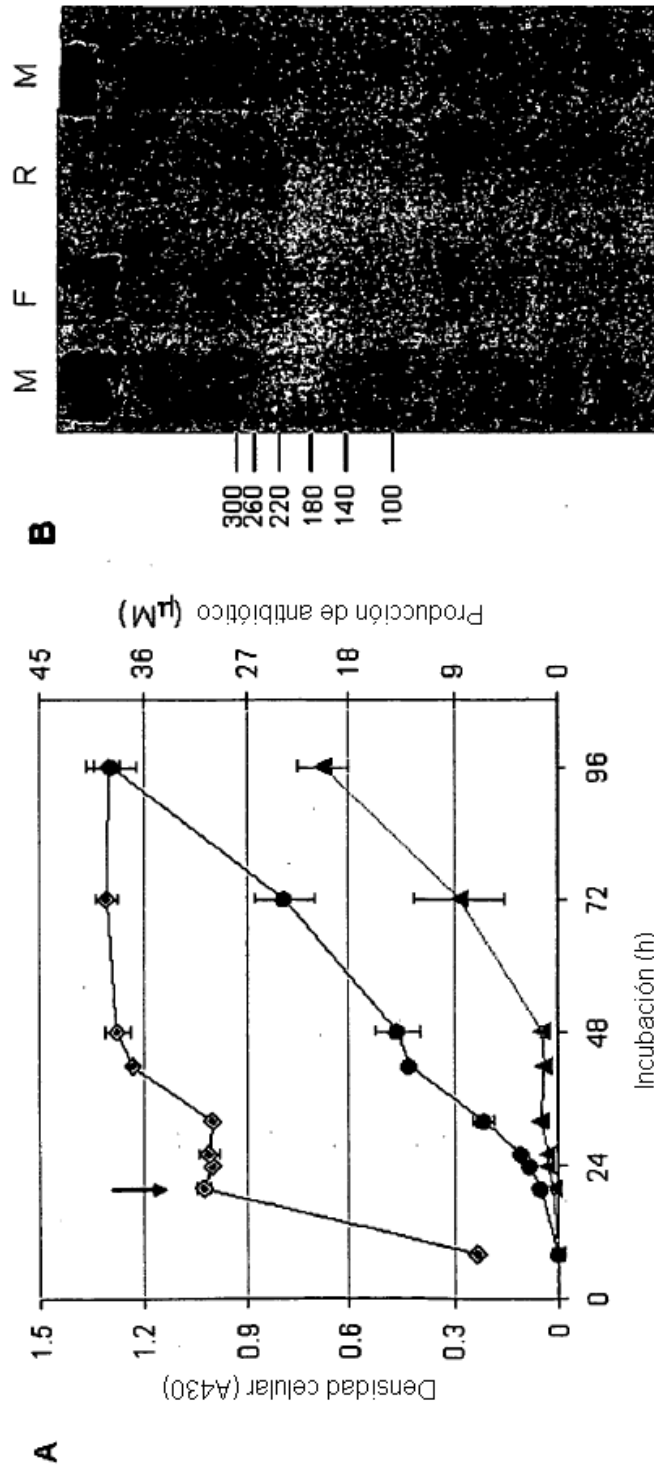
[illegible]

Figura 15

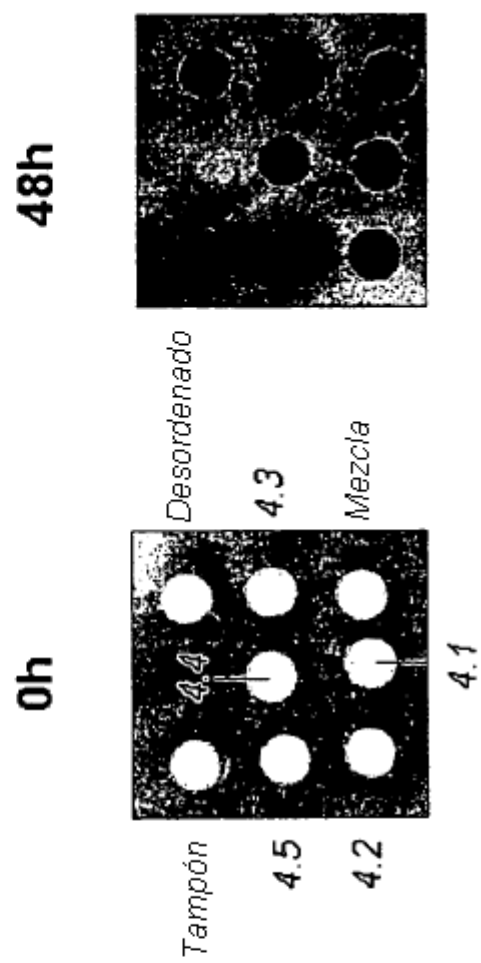


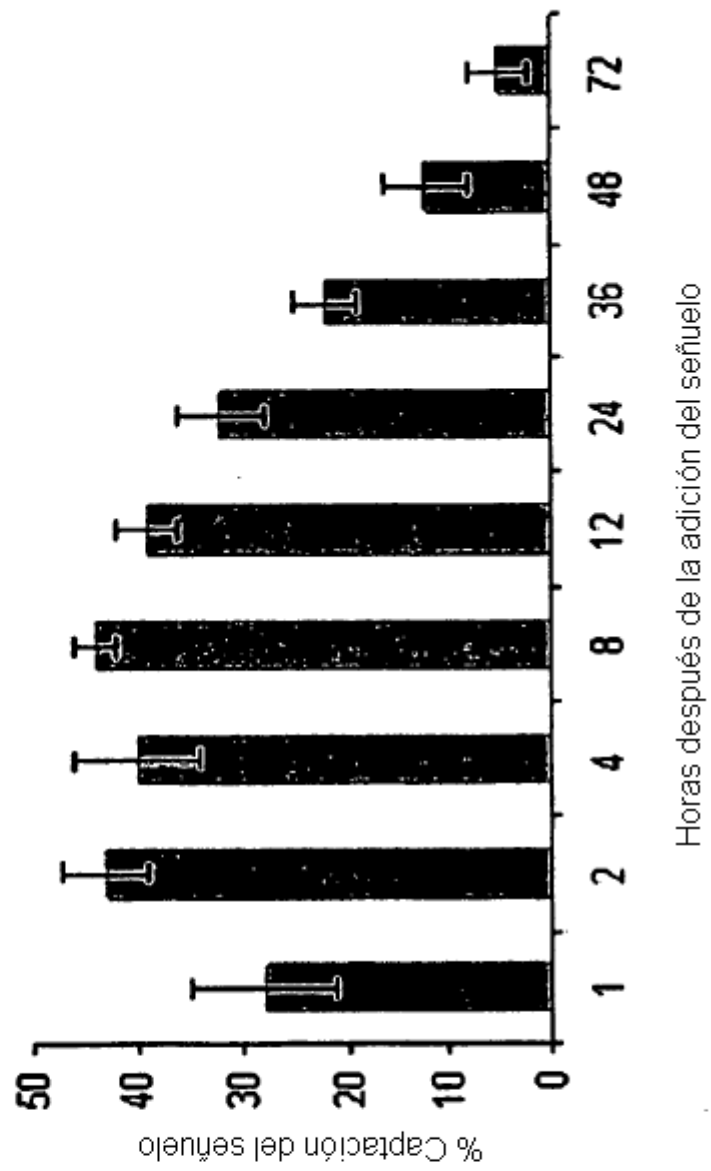
Figura 16

Figura 17

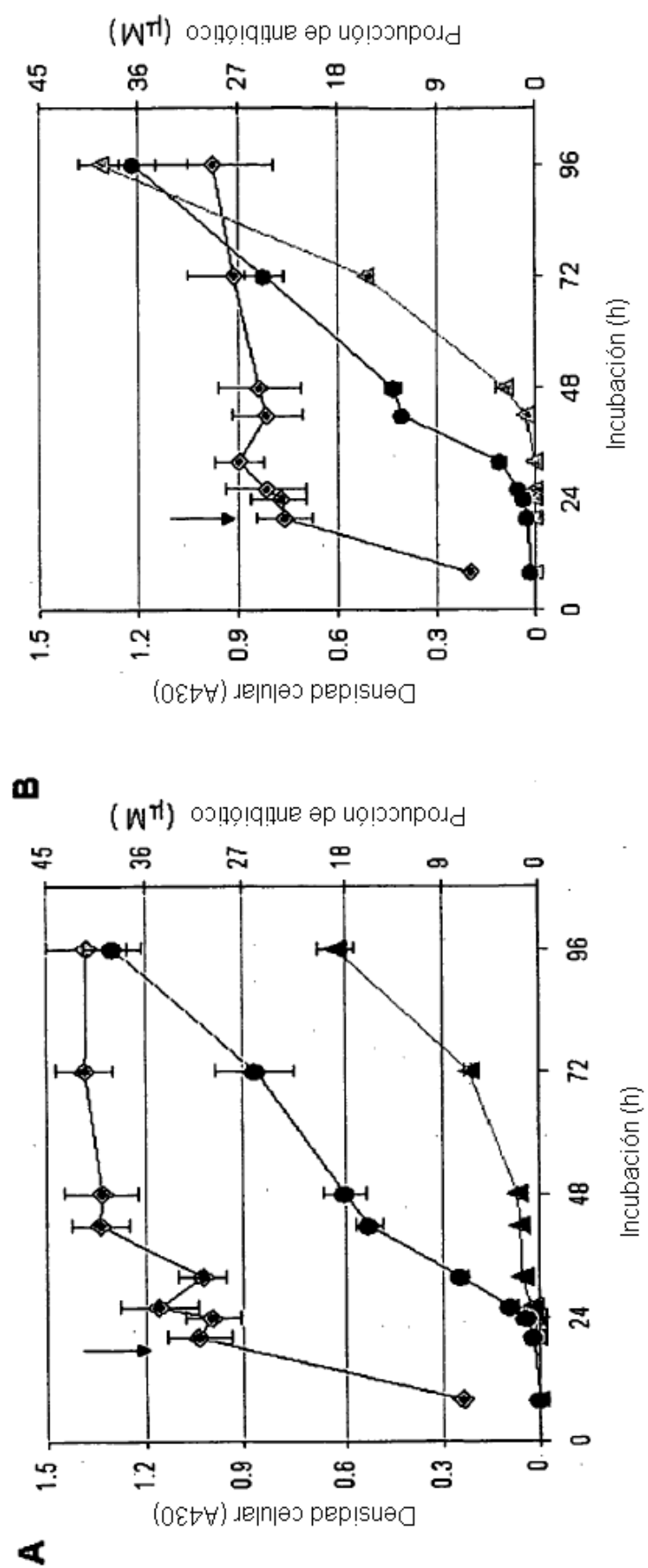


Figura 18

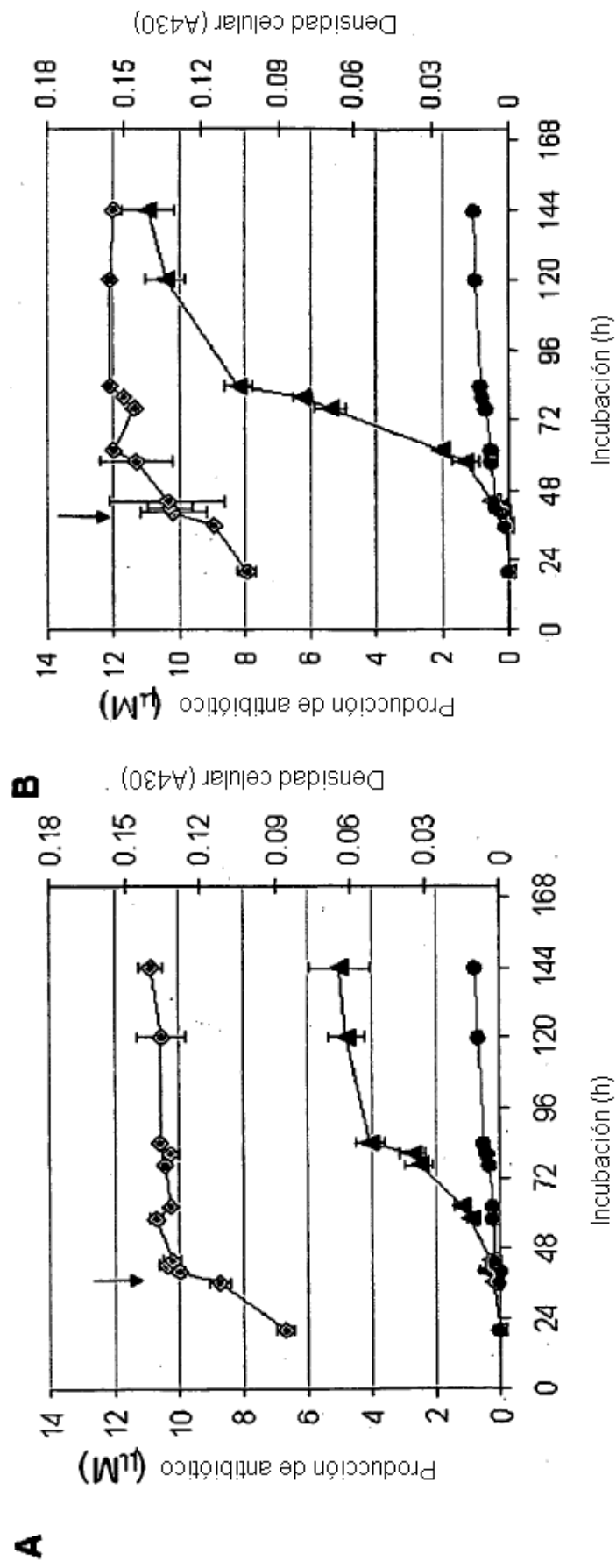


Figura 19

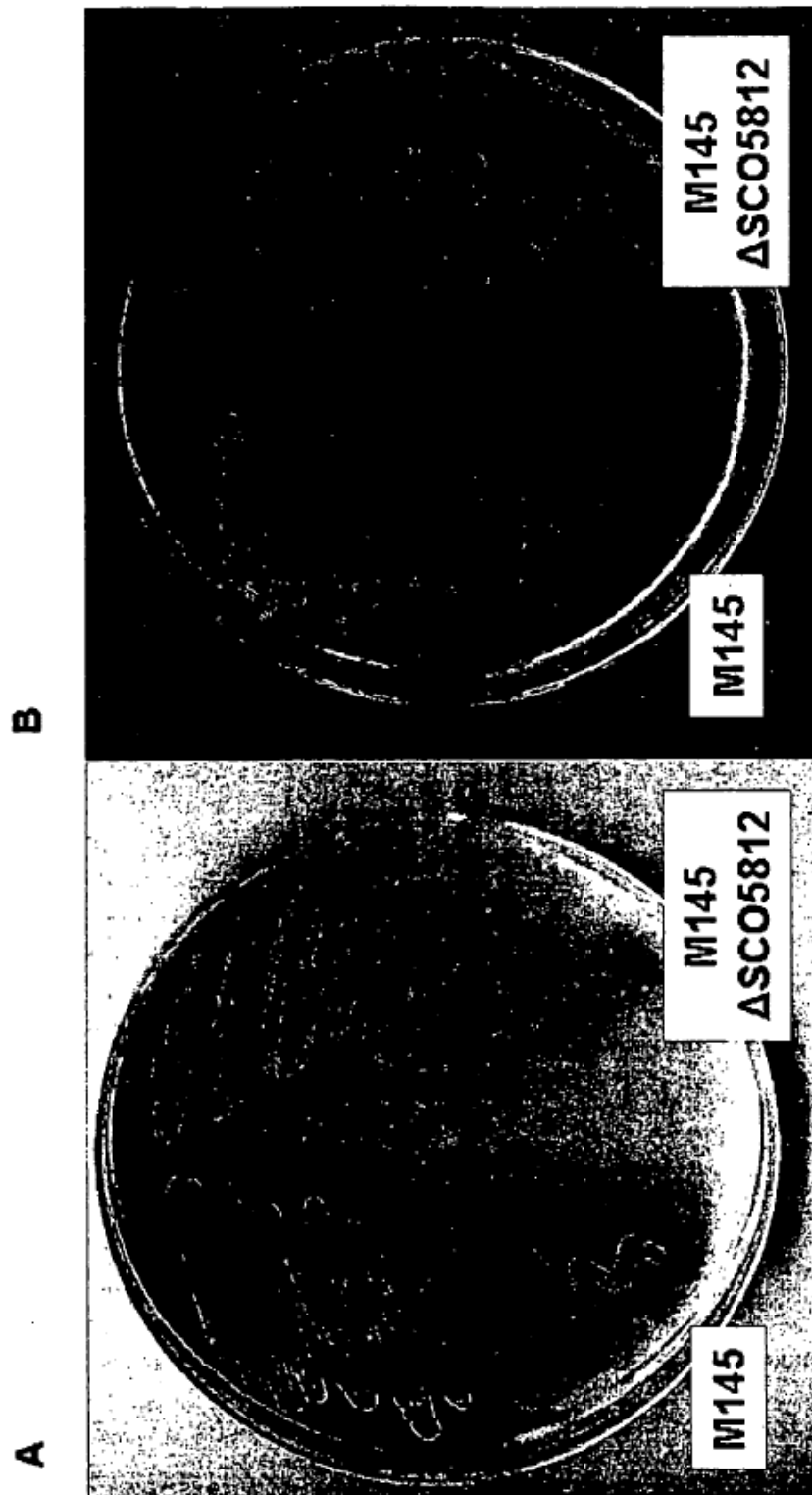


Figura 20

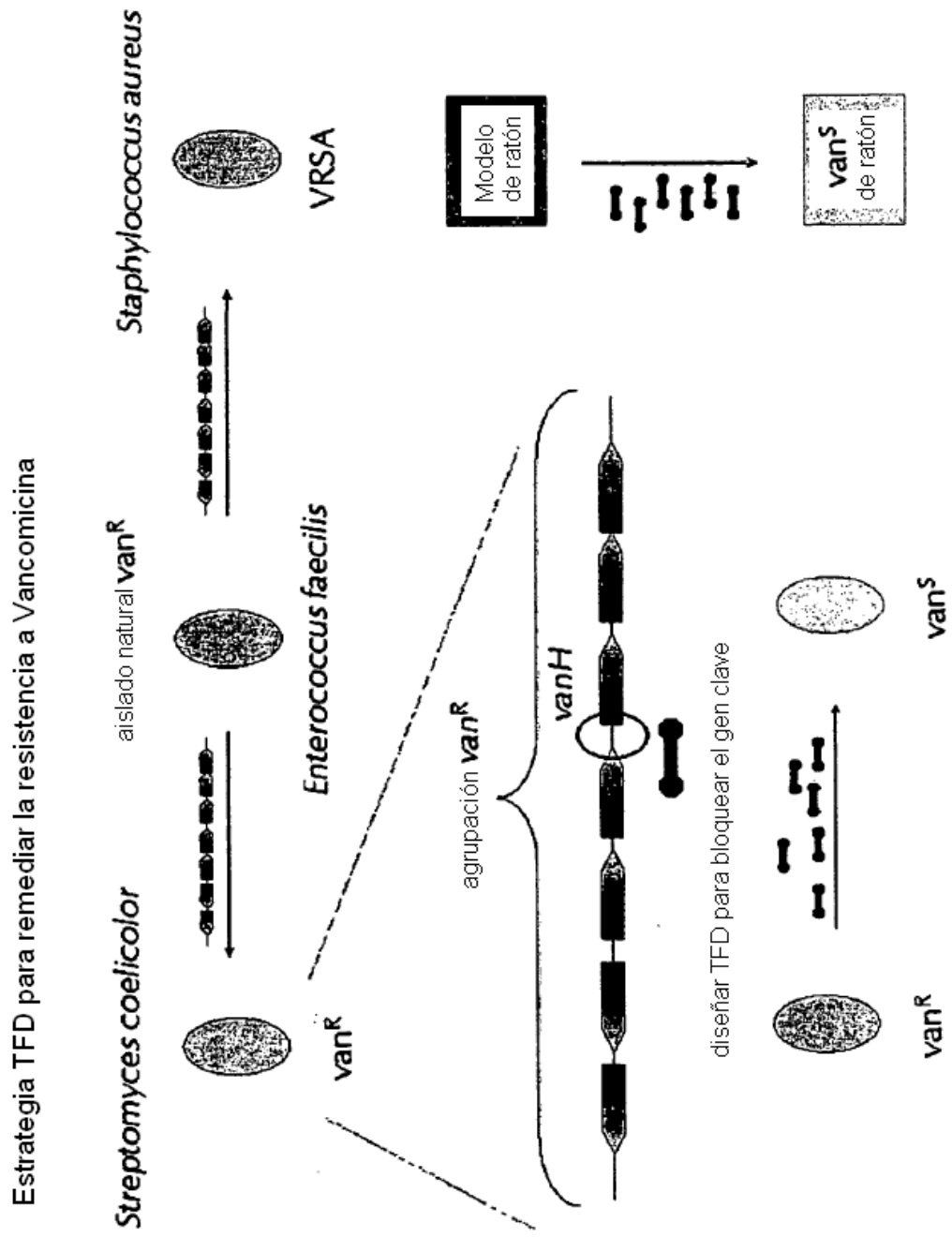
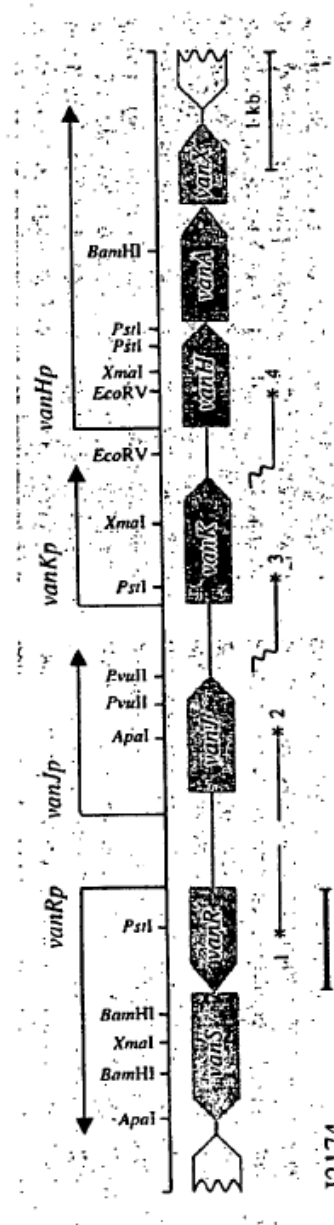


Figura 21



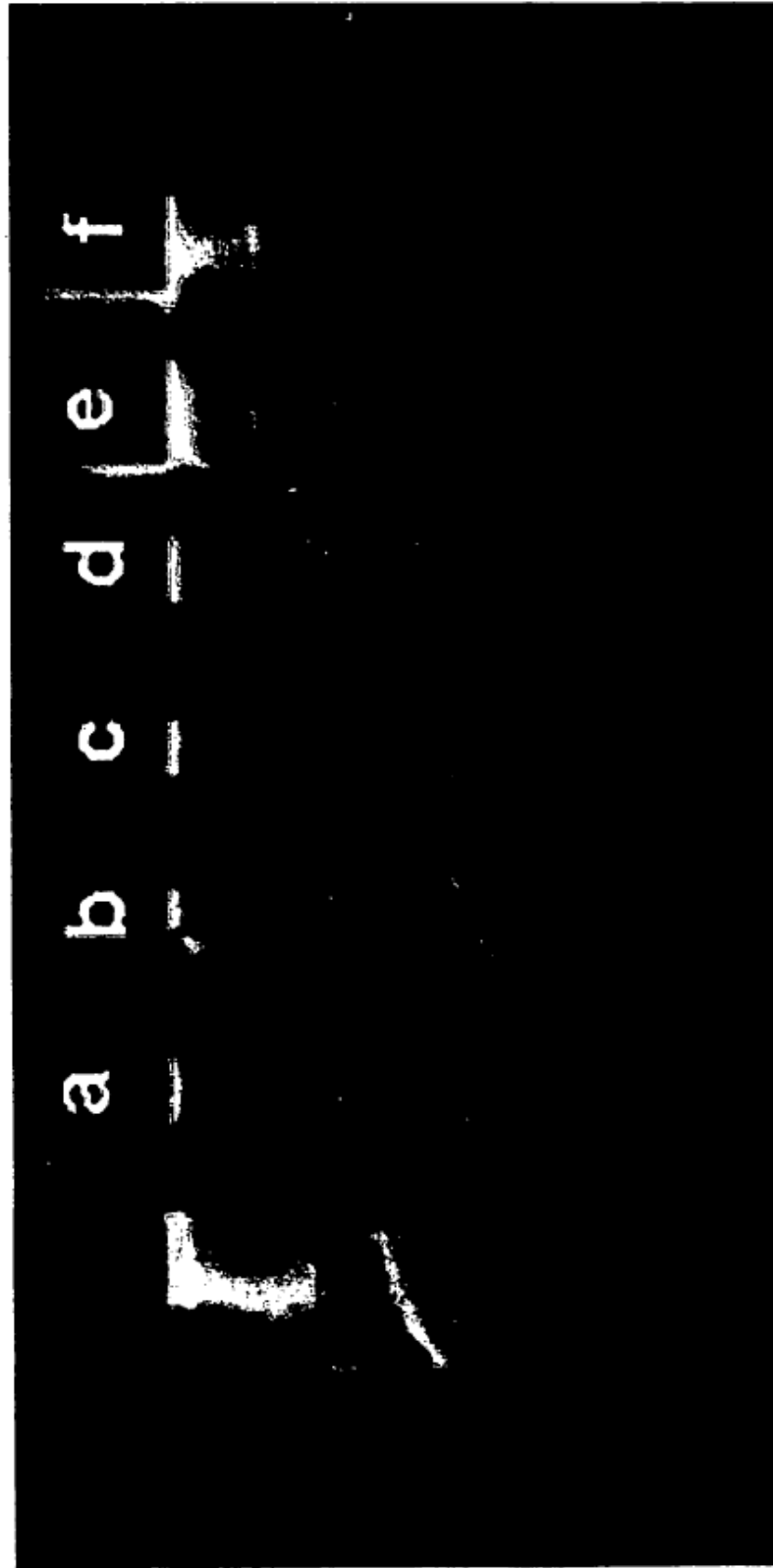


Figura 22

Figura 23

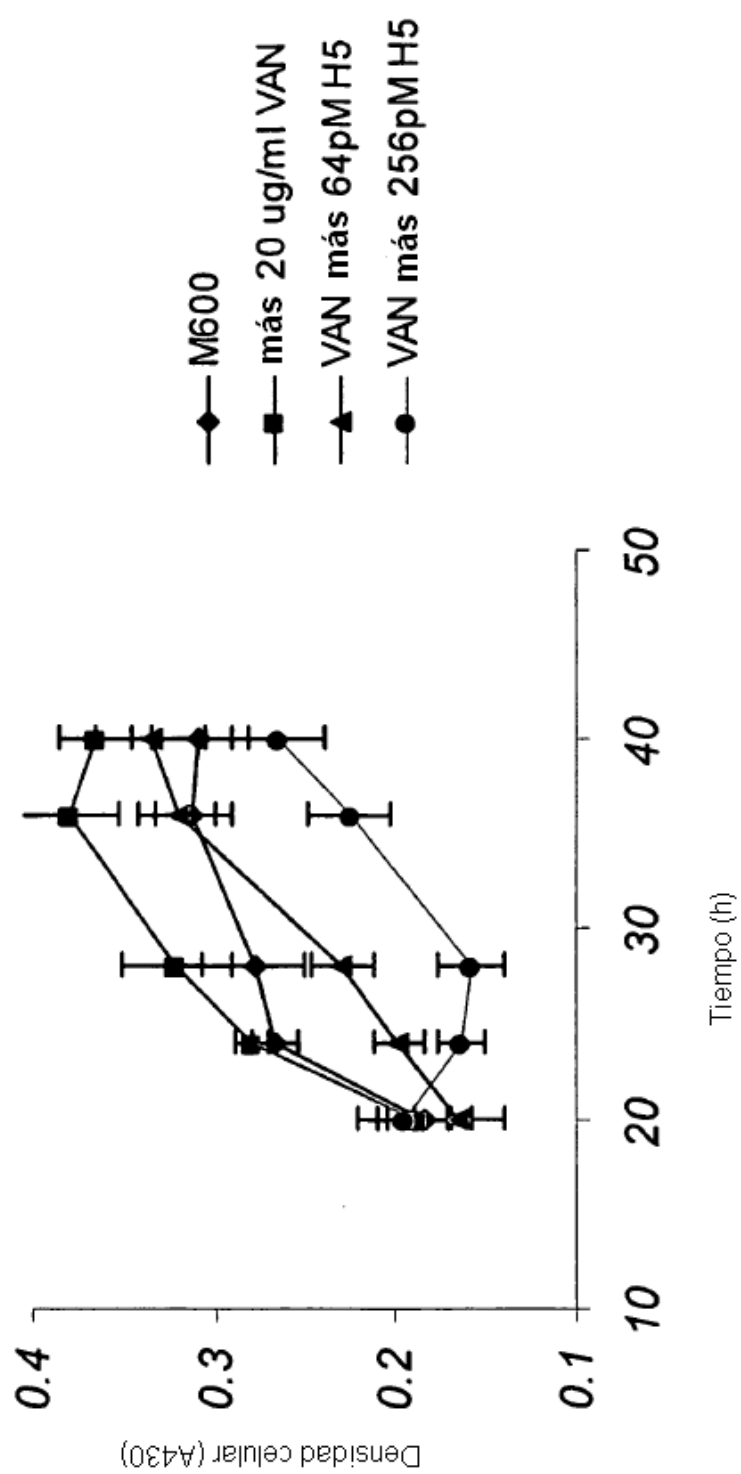


Figura 24

```

CTTGG TTTT CCAAG AGAAGAGC ccg cca tgg cgg ccg cgg gaa ttc NNN NNN TCA CTA GTG AAT TCG CGG CGG CCT GCTCTTCG CCG TCA AAA AGA CGG a
GAACC AAAA GGTC TCTTCTCG ggc ggt acc gcc ggc gcc ctt aag NNN NNN AGT GAT CAC TTA AGC GCC GGC GGA CGTGAAGC GGC AGT TTT TCT GCC

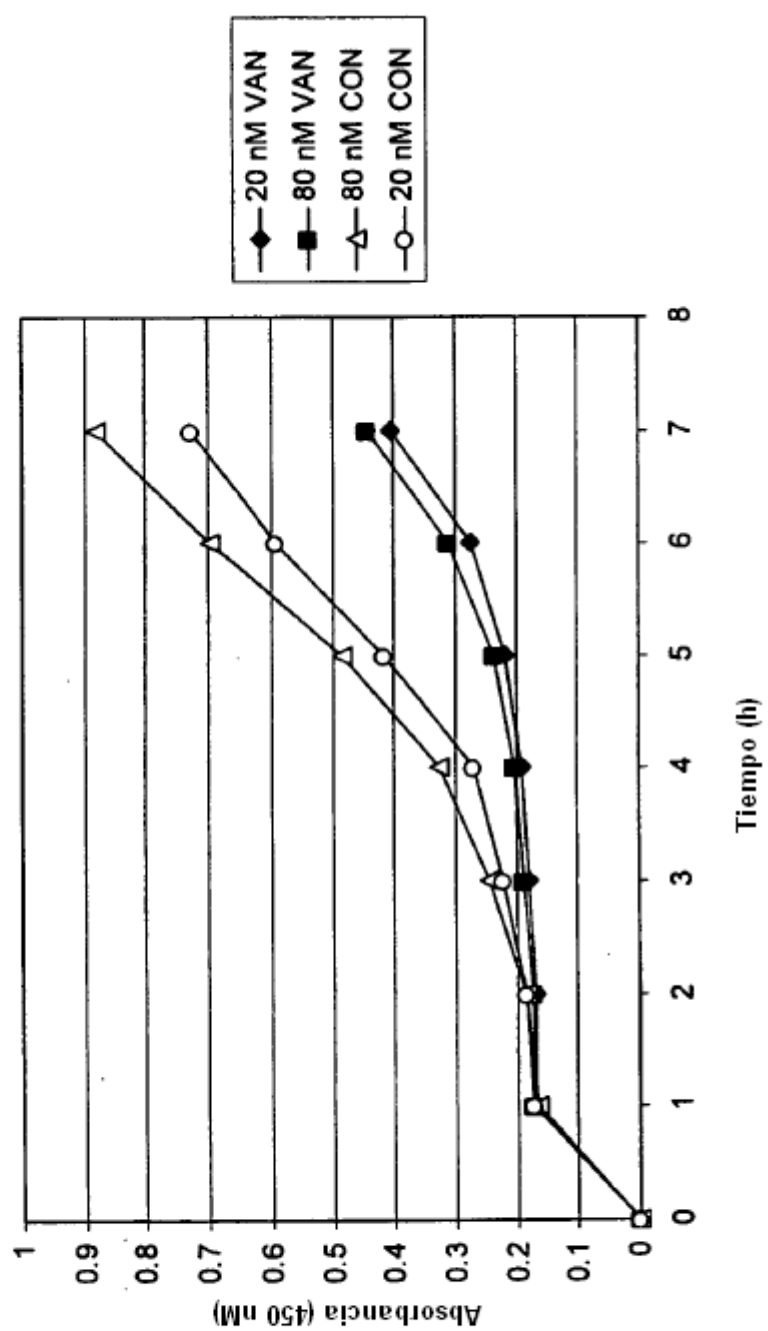
T CON Nt.BspQ1 (NEB: R0644L)

CTTGG TTTT CCAAG AGAAGAGC ccg cca tgg cgg ccg cgg gaa ttc NNN NNN TCA CTA GTG AAT TCG CGG CGG CCT GCTCTTCG
TCTTCTCG ggc ggt acc gcc ggc gcc ctt aag NNN NNN AGT GAT CAC TTA AGC GCC GGC GGA CGTGAAGC GGC AGT TTT TCT GCC

E CON ADN POLIMERASA T4 :

```

Figura 25



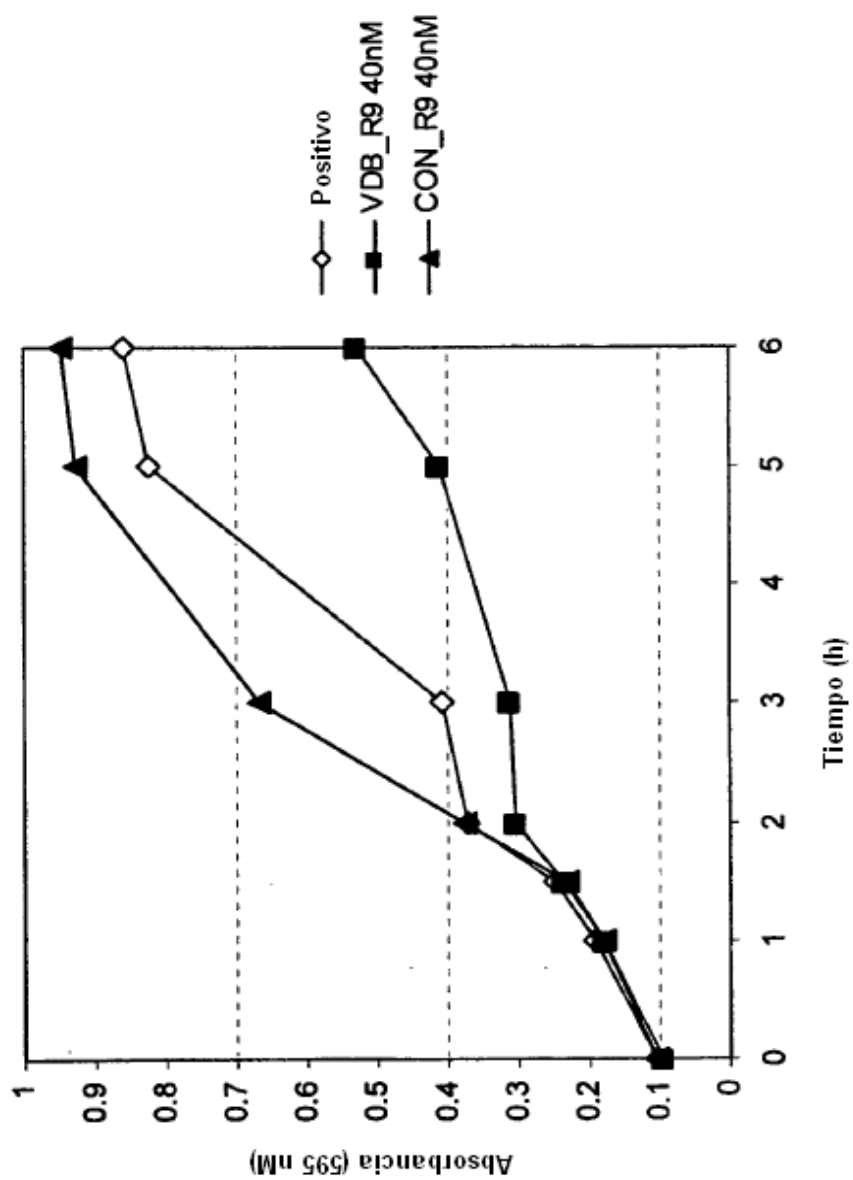


Figura 26

Figura 27

Tabla 1. Resistencia mediada por eflujo a antimicrobianos no fluoroquinolona

Antimicrobiano	Sistema de Eflujo	Familia de bombas	Localización génica	Organismo(s)	Referencia ^a
Cloranfenicol y cl & florfenicol					
Cloranfenicol	Cml, CmlA, CmlB	MF	mayoría plásmido; algo cromosoma	<i>P. aeruginosa</i> , <i>E. aerogenes</i> , <i>K. pneumoniae</i> , <i>S. enterica</i> serovar Typhimurium	48
	Cml, Cmlv, Cmr, Cmx, CmrA	?	plásmido y cromosoma	<i>Streptomyces</i> spp., <i>Corynebacterium</i> spp., <i>Rhodococcus</i> spp.	48
	MdfA ^b	MF	cromosoma	<i>E. coli</i>	74
	MexEF-OprN ^c , variedad de bombas RND de tres componentes ^d	RND	cromosoma	<i>P. aeruginosa</i> , varias bacterias Gram-negativas	74, 79
	OqxAB	RND	plásmido	<i>E. coli</i>	84
Cloranfenicol, florfenicol	Flo, FloR, pp-Flo	MF	plásmido y cromosoma	<i>E. coli</i> , <i>K. pneumoniae</i> , <i>V. cholerae</i> , <i>S. enterica</i> serovar Typhimurium	48
	FexA	?	plásmido	<i>S. lentus</i>	70
Macrólidos, lincosamidas, estreptograminas y quetólidos					
Macrólidos	Mef(A)	MF	cromosoma	<i>Streptococcus</i> spp., <i>Corynebacterium</i> spp., <i>Enterococcus</i> spp., <i>Micrococcus</i> spp., <i>Staphylococcus</i> spp., <i>Acinetobacter</i> spp., <i>Bacteroides</i> spp., <i>Neisseria</i> spp., varias <i>Enterobacteriaceae</i> y <i>Pseudomonadaceae</i> commensals	146, 176
Macrólidos, estreptograminas tipo B	Msr(A)	ABC	plásmido	<i>Staphylococcus</i> spp.	173, 175
	Msr(C)	ABC	cromosoma	<i>E. faecium</i>	146, 180
Macrólidos, quetólidos	Msr(D)	ABC	cromosoma	<i>S. pneumoniae</i>	149
Macrólidos, lincosamidas, estreptograminas tipo A	MdeA	MF	cromosoma	<i>S. aureus</i> , <i>S. hemolyticus</i> , <i>B. cereus</i> , <i>B. subtilis</i>	192
Estreptograminas Tipo A	Vga(A/B)	ABC	plásmido	<i>S. aureus</i>	173
Lincosamidas, estreptograminas	Lsa	ABC	cromosoma	<i>E. faecalis</i>	173, 186
Clindamicina	Lsa(B)	ABC	plásmido	<i>S. sciuri</i>	187
Lincosamidas	LmrB	MF	cromosoma?	<i>B. subtilis</i>	189, 190
Lincosamidas	LmrB	MF	cromosoma	<i>C. glutamicum</i>	191
Eritromicina	Cmc	MF	cromosoma	<i>C. difficile</i>	172
Macrólidos, lincosamidas, estreptograminas tipo B	?	?	?	<i>S. pyogenes</i>	188
Macrólidos	MacAB-TolC	ABC	cromosoma	<i>E. coli</i>	521
Macrólidos	MtrCDE	RND	cromosoma	<i>N. gonorrhoeae</i>	168, 197, 198
Macrólidos, lincosamidas, quetólidos	AcrAB-TolC, Mex ^e ; bombas RND en varias bacterias Gram-negativas ^h	RND	cromosoma	<i>E. coli</i> , <i>E. aerogenes</i> , <i>P. aeruginosa</i> ; otras bacterias Gram-negativas	74, 195, 196
Eritromicina	MdfA ^m	MF	cromosoma	<i>E. coli</i>	74
Eritromicina	?	?	?	<i>Campylobacter</i> spp.	202, 480

Figura 27 (Cont.)

Tabla 1. (continuación)

Antimicrobiano	Sistema de Eflujo	Familia de bombas	Localización génica	Organismo(s)	Referencia ^a
Eritromicina, roxitromicina	MexCD-OprJ	RND	plásmido	<i>Pseudomonas</i> spp.	201
Tetraciclinas y gliciliclinas					
Tetraciclinas	Tet(A), Tet(B), Tet(C), Tet(D), Tet(E), Tet(G), Tet(H), Tet(J), Tet(Y), Tet(Z), Tet(30), Tet(39)	MF	Plásmido	Bacterias Gram-negativas	47, 96, 522
	Tet(K), Tet(L)	MF	plásmido	Bacterias Gram-positivas	47, 522
	Tet38	MF	cromosoma	<i>S. aureus</i>	94
	Tet(V)	MF	cromosoma	<i>M. tuberculosis</i> , <i>M. fortuitum</i>	95
	Rv1258/Tap	MF	cromosoma	<i>M. tuberculosis</i> , <i>M. fortuitum</i>	129, 130
	P55/Rv1410	MF	cromosoma	<i>M. tuberculosis</i> , <i>M. bovis</i>	131
	MdfA ^m	MF	cromosoma	<i>E. coli</i>	74
	MexAB-OprM ⁱ , varias bombas de tres componentes de la familia RND				74
Gliciliclinas	AcrAB-TolC, MexXY-OprM, MexAB-OprM, MexCD-OprJ	RND	cromosoma	<i>P. mirabilis</i> , <i>E. coli</i> , <i>K. pneumoniae</i> , <i>M. morganii</i> , <i>P. aeruginosa</i>	460-464a
	MepA	MATE	cromosoma	<i>S. aureus</i>	464a
β-Lactamas					
	AcrAB-TolC ^j	RND	cromosoma	<i>H. influenzae</i>	331
	MexAB-OprM ^k , varias bombas de tres componentes de la familia RND ^l	RND	cromosoma	<i>P. aeruginosa</i> , varias bacterias Gram-negativas	74, 250
Aminoglicósidos					
	LmrA ^m	ABC	cromosoma	<i>L. lactis</i>	85
	AcrAD-TolC	RND	cromosoma	<i>E. coli</i>	74, 523
	BpeAB-OprB	RND	cromosoma	<i>B. pseudomallei</i>	292
	AmrAB-OprA	RND	cromosoma	<i>B. pseudomallei</i>	524
	MexXY-OprM	RND	cromosoma	<i>P. aeruginosa</i>	74, 306, 307
	MexAB-OprM ⁿ	RND	cromosoma	<i>P. aeruginosa</i>	74
	EmrE ⁿ	RND	cromosoma	<i>P. aeruginosa</i>	74
	MdfA ^m	MF	cromosoma	<i>E. coli</i>	74
	LmrA ^m	ABC	cromosoma	<i>L. lactis</i>	85
Oxazolidinonas					
	AcrAB-TolC, AcrEF-TolC	RND	cromosoma	<i>E. coli</i>	196

^aCuando se han descrito sistemas de eflujo en artículos previos de revisión, éstos se resaltan aquí, de otra manera se citan los artículos originales.

^bEste sistema de eflujo multifármaco se identificó originalmente como el exportador de cloranfenicol Cmr/CmlA.

^cCepas mutantes de *P. aeruginosa* que sobreexpresan MexEF-OprN se seleccionan fácilmente *in vitro* usando cloranfenicol.⁷⁸⁻⁸⁰

^dCuando se ensayaron, los sistemas de eflujo tipo RND que estimulan resistencia a cloranfenicol también estimulan resistencia a florfenicol (por ejemplo, AcrAB-TolC⁷⁵ y AcrEF-TolC⁸⁶ en *S. enterica* serovar Thyphimurium).

^eEl eflujo se ha demostrado aunque no se ha identificado un sistema de eflujo.

^fEl gen o genes de eflujo no se han identificado aunque se ha confirmado su movilidad.

^gCepas mutantes de *P. aeruginosa* que sobreexpresan MexCD-OprJ pueden seleccionarse *in vitro* usando eritromicina (Poole, resultados no publicados).

^hLa exportación de macrólidos, lincosamidas y estreptograminas por sistemas de eflujo multifármaco de tipo RND de tres componentes puede explicar la falta de susceptibilidad de muchas bacterias Gram-negativas a estos agentes.

ⁱCepas mutantes de *P. aeruginosa* que sobreproducen MexAB-OprM pueden seleccionarse *in vitro* usando tetraciclina.^{78, 132, 133}

^jAsociado con resistencia a ampicilina en cepas clínicas.³³¹

^kImplicado en resistencia a meropenem³²² y ticarcilina^{323, 324} en aislados clínicos.

^lAunque muchos sistemas de eflujo multifármaco de tipo RND se adaptan a β-lactamas (véase el texto), pocos están implicados como determinantes principales de resistencia, en aislados de laboratorio o clínicos.

^mEl gen clonado estimula incrementos muy modestos en MIC para el antimicrobiano indicado.

ⁿMuestran una contribución modesta a la resistencia intrínseca a aminoglicósido y sólo se observa en medios con baja fuerza iónica.

Figura 27 (Cont.)**Tabla 2.** Resistencia mediada por eflujo a fluoroquinolonas^a

Sistema de eflujo	Familia de bombas	Antimicrobiano ^b	Organismo(s)	Referencia
Gram-positivas				
NorA	MF	NOR, CIP	<i>S. aureus</i>	209, 525
NorB	MF	NOR, CIP, MOX, SPR	<i>S. aureus</i>	94
?	?	NOR, CIP, MOX, GAT, SPR	<i>S. aureus</i>	212
PmrA	MF	NOR, CIP	<i>S. pneumoniae</i>	209, 526
?	?	NOR, CIP, MOX	<i>S. pneumoniae</i>	213
EmcA	MF	NOR, CIP	<i>E. faecalis</i>	235, 235a
Lde	MF	NOR, CIP	<i>L. monocytogenes</i>	233
EfrAB	ABC	NOR, CIP	<i>E. faecalis</i>	352
?	?	FQ	<i>B. anthracis</i>	237, 527
Bmr	MF	FQ	<i>B. subtilis</i>	209
Blt	MF	FQ	<i>B. subtilis</i>	209
Bmr3	MF	FQ	<i>B. subtilis</i>	528
MD1, MD2 ^c	ABC	CIP	<i>M. hominis</i>	240
LfrA	MF	FQ (CIP, NOR?)	<i>M. smegmatis</i>	242, 529
EfpA	MF	CIP, NOR	<i>M. smegmatis</i>	242
Rv1634	MF	FQ	<i>M. tuberculosis</i>	129
Rv1258c	MF	OFL	<i>M. tuberculosis</i>	245
Rv2686c-Rv2687c-Rv2688c ^e	ABC	FQ	<i>M. tuberculosis</i>	243
LmrA	ABC	CIP, OFL	<i>L. lactis</i>	85
Mmr	SMR	CIP, NOR	<i>M. smegmatis</i>	242
Gram-negativas				
AcrAB-TolC, AcrEF-TolC	RND	FQ	<i>E. coli</i>	74
MexAB-OprM, MexCD-OprJ	RND	FQ	<i>P. aeruginosa</i>	74, 250
MexEF-OprN, MexXY-OprM ^h				
AcrAB-TolC	RND	FQ	<i>Enterobacter</i> spp.	74, 258, 530
AcrAB-TolC	RND	FQ	<i>Klebsiella</i> spp.	74, 76, 263, 531
AcrAB-TolC	RND	FQ	<i>S. enterica</i> (Typhimurium, Enteritidis)	75, 270-272, 532
AcrEF-TolC	RND	FQ	<i>S. enterica</i> serovar Typhimurium	86
SmeABC, SmeDEF	RND	FQ	<i>S. maltophilia</i>	74, 277, 289, 533, 534
CmeABC, CmeDEF	RND	FQ	<i>C. jejuni</i>	74, 256, 257, 535
SdeAB	RND	FQ	<i>S. marcescens</i>	194, 276, 535a
SdeXY	RND	NOR	<i>S. marcescens</i>	74
MtrCDE	RND	FQ	<i>N. gonorrhoeae</i>	210, 280
CeoAB-OpcM	RND	FQ	<i>B. cepacia</i> (cenocepacia)	82
AcrAB	RND	CIP	<i>P. mirabilis</i>	74
AdeABC	RND	FQ	<i>A. baumannii</i>	74
VcaM	ABC	CIP, NOR	<i>V. cholerae</i>	290
Orf12-Orf11-Orf10(plásmido) ⁱ	ABC	NAL, NOR	<i>Pseudomonas</i> spp.	193
MdfA	MF	FQ	<i>E. coli</i>	210, 291
?	?	FQ	<i>A. salmonicida</i>	255
?	?	FQ	<i>C. freundii</i>	536, 537
?	?	OFL	<i>P. vulgaris</i>	538
?	?	CIP, NOR	<i>B. fragilis</i>	209, 539

^aExcluyendo la familia MATE de exportadores multifármaco.^bNOR, norfloxacin; CIP, ciprofloxacina; MOX, moxifloxacina; GAT, gatifloxacina; SPR, esparfloxacina; OFL, ofloxacina; NAL, ácido nalidixico; FQ, fluoroquinolonas^cResistencia mediada por eflujo observada pero el determinante específico no se ha indentificado; no NorA.^dResistencia mediada por eflujo observada pero el determinante específico no se ha indentificado; no PmrA.^eResistencia mediada por eflujo observada pero el determinante específico no se ha indentificado.^fSistema de eflujo de tipo ABC asociado a la membrana citoplásmica ensamblado de 2 subunidades.^gSistema de eflujo de tipo ABC asociado a la membrana citoplásmica ensamblado de 3 subunidades.^hSe adapta a fluoroquinolonas aunque no hay indicaciones de fluoroquinolonas seleccionando mutantes que sobreproducen MexXY *in vitro* o *in vivo*.ⁱCodifica un componente ABC-MFP de un probable sistema de exportación multifármaco ABC-MFP-OMF.