

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 694 349**

51 Int. Cl.:

C12N 15/90 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **08.12.2014 PCT/NL2014/050839**

87 Fecha y número de publicación internacional: **11.06.2015 WO15084178**

96 Fecha de presentación y número de la solicitud europea: **08.12.2014 E 14815085 (7)**

97 Fecha y número de publicación de la concesión europea: **08.08.2018 EP 3077521**

54 Título: **Sistema novedoso de alteración de genomas para microorganismos**

30 Prioridad:

06.12.2013 NL 2011912

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

20.12.2018

73 Titular/es:

**HEINEKEN SUPPLY CHAIN B.V. (100.0%)
Tweede Weteringplantsoen 21
1017 ZD Amsterdam, NL**

72 Inventor/es:

**DARAN, JEAN-MARC GEORGES;
GEERTMAN, JAN-MAARTEN ANTON y
BOLAT, IRINA**

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 694 349 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Sistema novedoso de alteración de genomas para microorganismos

5 Campo

La invención se refiere a los campos de la biología molecular y la ingeniería genética de microorganismos, especialmente de levaduras.

10 Introducción

La recombinación homóloga en microorganismos como la levadura se basa en un mecanismo de reparación de rotura de doble cadena, que se une a los fragmentos de ADN. Cuando se detecta una rotura de ADN de doble cadena, una exonucleasa degrada ambos extremos 5', después de lo cual se produce la invasión de la hebra de la plantilla homóloga. El mecanismo de síntesis de ADN repara las dos hebras y la ligadura del ADN completa el proceso sin ninguna eliminación [Storici et al., (2003). PNAS USA 100: 14994-9; Haber, (2000). Trends Genet 16: 259-264]. Aunque la reparación de la recombinación homóloga se producirá con tan solo 30 pb de homología, es mucho más eficiente con 200-400 pb [Sugawara et al., (2000). Mol Cell Biol 20: 5300-5309].

20 Un método alternativo para reparar una rotura de doble cadena se basa en la unión del extremo no homólogo, donde el heterodímero de las llamadas proteínas Ku agarra los extremos del cromosoma roto, lo que promueve la unión de proteínas adicionales. Estas proteínas adicionales procesan los extremos del ADN y los ligan, lo que generalmente crea una eliminación de varios nucleótidos [Storici et al., (2003). PNAS USA 100: 14994-9].

25 La ruta de reparación por recombinación homóloga se usó con éxito para construir un plásmido a partir de dos fragmentos de ADN co-transformados, que contenían regiones homólogas [Ma et al., (1987). Gene, 58: 201-16.26]. Los microorganismos, y especialmente las especies de *S. cerevisiae*, son organismos manejables para desarrollar nuevas técnicas [Kumar y Snyder, (2001). Nat Rev Genet 2: 302-312], en el que la alteración genética se realiza con ADN bicatenario o con ADN de cadena sencilla [Orr-Weaver et al., (1981). PNAS USA 78: 6354-6358; Moerschell et al., (1988). PNAS USA 85: 524-528]. *S. cerevisiae* puede captar y ensamblar al menos 38 oligonucleótidos de cadena simple y un vector de doble cadena lineal en un evento de transformación con superposiciones entre oligonucleótidos de tan solo 20 pares de bases y con una longitud de 200 nucleótidos [Gibson, (2009). Nucleic Acids Res 37: 6984-6990].

35 Una de las herramientas más poderosas en la caracterización funcional de productos genéticos desconocidos es la eliminación completa de genes en un cromosoma. La selección de genes se ha establecido con fragmentos de PCR flanqueados por secuencias homólogas tan cortas como 35-40 pb que permiten la transformación directa debido a la alta eficacia de la recombinación homóloga en *S. cerevisiae* [Baudin et al., (1993). Nucleic Acids Res 21: 3329-3330; Klinner y Schafer, (2004). FEMS Microbiol Rev 28: 201-223]. Sin embargo, en *S. pastorianus* que elabora la lager, la eficacia de la recombinación homóloga es baja debido a la compleja genética. Por lo tanto, las cepas de elaboración de cerveza lager requieren flancos superpuestos homólogos más largos (>400 pb) para tener una reparación o inserción de un casete de eliminación en el ADN genómico.

45 Para modificar genes múltiples, también es necesario reciclar marcadores. Los sistemas anteriores para la eliminación de genes y la posterior escisión del marcador contenían un marcador flanqueado por repeticiones directas de la secuencia de HisG bacteriana [Akada et al., (2002). Yeast 19: 393-402] o por dos sitios objetivo de una recombinasa específica del sitio [McNabb et al., (1997). Biotechniques, 22: 1134-1139; Storici et al., (1999). Levadura 15: 271-283; Gueldener et al., (2002). Nucleic Acids Res 30: e23; Iwaki y Takegawa, (2004). Biosci Biotechnol Biochem 68: 545-550]. Con este método, una copia de las repeticiones permanece en el cromosoma seleccionado. Sin embargo, cuando están presentes múltiples secuencias residuales en el genoma, el porcentaje de integraciones correctas en transformaciones de casete contra-seleccionables sucesivas disminuye dramáticamente [Davidson y Schiestl, (2000). Curr Genet 38: 188-190]. Múltiples sitios de destino pueden provocar reordenamientos cromosómicos. Por ejemplo, cuando cuatro conjuntos de repeticiones de loxP se ubicaron simultáneamente en el genoma, se produjeron reordenamientos cromosómicos a una frecuencia del 50% mediante la expresión de la recombinasa CRE. Esto significa que la orientación sucesiva en el mismo microorganismo requiere más detección en cada ronda para identificar un no-caut correcto [Delneri et al., (2000). Gene 55: 127-135].

60 Para superar este problema, se usó un sistema para la eliminación de genes sin interrupciones en el que se utilizó un casete amplificado por PCR, que contiene un marcador URA3 unido a una secuencia de 40 pares de bases duplicada derivada del locus objetivo, para la interrupción de HIS3 y el reciclado de marcadores sin ninguna cicatriz genómica Akada et al., (2006). Yeast 23: 399-405]. Las colonias se seleccionaron por duplicado utilizando un ácido 5-fluoroorótico para identificar las colonias que habían perdido URA3 por recombinación entre las secuencias de 40 pares de bases duplicadas. Esto dio como resultado la eliminación de HIS3 sin secuencias extrañas residuales [Akada et al., (2006). Yeast 23: 399-405]. También se demostró que un largo tramo de 966 pares de bases era necesario para la orientación correcta del gen. El reemplazo de este estiramiento por una secuencia homóloga corta de 40 pb no generó transformantes [Akada et al., (2006). Yeast 23: 399-405].

La eficacia de la orientación depende en parte de la presencia de una secuencia homóloga larga en una construcción de orientación [Davidson y Schiestl, (2000). *Curr Genet* 38: 188-190]. Sin embargo, la recombinación homóloga en algunos microorganismos, como por ejemplo la mayoría de las cepas de la levadura cervecera *Saccharomyces pastorianus*, es difícil de lograr, incluso en presencia de largas secuencias homólogas en la construcción dirigida (Murakami et al., 2012). *Yeast*, 29: 155-165.

Cuando la recombinación homóloga es menos eficiente, aumenta la posibilidad de obtener falsos positivos. Los falsos positivos por lo general son el resultado de eventos cruzados simples aleatorios.

La presente invención supera el problema de la orientación eficiente al proporcionar un grupo de construcciones de orientación, en las que la expresión correcta de un marcador de selección depende de un evento de recombinación entre las construcciones de orientación. Se encontró que la aparición de un evento de recombinación entre las construcciones de direccionamiento se mejora notablemente después de la integración de las construcciones de direccionamiento en el lugar de orientación correcto. Por lo tanto, el sistema objetivo de la presente invención, que comprende un conjunto de construcciones de orientación, aumenta considerablemente el porcentaje de construcciones correctamente integradas en microorganismos que expresan el marcador de selección, en comparación con un sistema de selección de un solo vector. La división del marcador en dos límites limita la aparición de falsos positivos debido a los eventos de un solo cruce. El enfoque de marcador dividido mejora la proporción de positivos verdaderos sobre falsos positivos (Nielsen et al., 2006). *Fungal Gen Biol* 43: 54-64).

La invención proporciona un grupo de construcciones dirigidas, que comprenden una primera construcción que comprende una primera región de homología con un genoma objetivo de un microorganismo, un sitio de reconocimiento para una endonucleasa y una primera parte de un marcador de selección, y una segunda construcción que comprende una segunda parte del marcador de selección, una copia del sitio de reconocimiento de endonucleasas y una segunda región de homología con el genoma objetivo del microorganismo, por lo que la primera y la segunda regiones de homología con el genoma objetivo comprenden cada una al menos 20 pares de bases (pb), dicho fragmento comprende preferiblemente entre 50 y 600 pb, por lo que una secuencia codificante que codifica la endonucleasa y que está acoplada a un promotor inducible está presente en la primera o segunda construcción; y mediante el cual una parte de la primera región de homología con el genoma objetivo en la primera construcción se duplica entre la copia del sitio de reconocimiento de endonucleasas y la segunda región de homología con el genoma objetivo en la segunda construcción; o una parte de la segunda región de homología con el genoma objetivo en la segunda construcción se duplica entre la primera región de homología con el genoma objetivo y el sitio de reconocimiento de endonucleasa en la primera construcción, comprendiendo dicha región duplicada entre 20 y 200 pb.

Dichas primera y segunda regiones de homología con el genoma objetivo comprenden cada una al menos 20 pares de bases (pb). En principio, no hay límite superior para la longitud de dichas primera y segunda regiones de homología. Sin embargo, por razones prácticas tales como la facilidad y eficiencia de generar las construcciones primera y segunda, dichas regiones de homología primera y segunda comprenden preferiblemente entre 20 pb y 100 kb, más preferiblemente entre 40 pb y 10 kb, más preferiblemente entre 50 pb y 5 kb, más preferido entre 100 bp y 1 kb.

Dicha región duplicada de homología con el genoma objetivo en la primera y segunda construcción de direccionamiento está preferiblemente entre 20 y 500 pb, preferiblemente entre 20 y 200 pb, preferiblemente entre 40 y 100 pb, preferiblemente alrededor de 80 pb. Dicha región duplicada de homología con el genoma objetivo en la primera y la segunda construcción de direccionamiento permite la eliminación íntegra del marcador del genoma objetivo mediante recombinación homóloga.

La primera construcción comprende preferiblemente, en este orden, una primera región de homología con un gen objetivo de un microorganismo, un sitio de reconocimiento para una endonucleasa y una primera parte de un marcador de selección. La segunda construcción comprende preferiblemente, en este orden, una segunda parte del marcador de selección, una secuencia de codificación que codifica la endonucleasa y que está acoplada a un promotor inducible, una copia del sitio de reconocimiento de endonucleasas, una copia de una parte de la primera región de homología con el gen objetivo que está presente en la primera construcción, y una segunda región de homología con el gen objetivo del microorganismo. Esta configuración se representa en la FIG. 1.

El término construcción, como se usa en este documento, se refiere a un segmento de ácido nucleico construido artificialmente. Una construcción preferida es un vector, preferiblemente un vector que contiene genes de resistencia bacteriana para el crecimiento en bacterias. Una construcción más preferida es un plásmido, un ADN de doble cadena lineal o circular que es capaz de replicarse en bacterias independientemente del ADN cromosómico.

El gen objetivo puede ser cualquier gen de un microorganismo, preferiblemente de una levadura, cuya secuencia genómica deba alterarse. El término gen, como se usa en el presente documento, se refiere a una parte del genoma del microorganismo que comprende partes intrónicas y exónicas de un gen, la región promotora de dicho gen y secuencias genómicas que median la expresión de dicho gen, como por ejemplo secuencias potenciadoras.

El experto entenderá que las construcciones de direccionamiento pueden usarse preferiblemente para alterar un gen de un microorganismo. Por lo tanto, la divulgación proporciona además un conjunto de construcciones de direccionamiento

que comprenden una primera construcción que comprende una primera región de homología con un gen objetivo de un microorganismo, un sitio de reconocimiento para una endonucleasa y una primera parte de un marcador de selección y una segunda construcción que comprende una segunda parte del marcador de selección, una copia del sitio de reconocimiento de endonucleasas y una segunda región de homología con el gen objetivo del microorganismo, por lo que un fragmento de la primera parte del marcador de selección se superpone con un fragmento que está presente en el segunda parte del marcador de selección, que permite la recombinación entre la primera y la segunda parte del marcador de selección; por lo que una secuencia codificante que codifica la endonucleasa y que está acoplada a un promotor inducible está presente en la primera o segunda construcción; y mediante el cual una parte de la primera región de homología con el gen objetivo en la primera construcción se duplica entre la copia del sitio de reconocimiento de endonucleasa y la segunda región de homología con el gen objetivo en la segunda construcción; o una parte de la segunda región de homología con el gen objetivo en la segunda construcción se duplica entre la primera región de homología con el gen objetivo y el sitio de reconocimiento de endonucleasa en la primera construcción. Dicha región duplicada de homología con el gen objetivo en la primera y segunda construcción de direccionamiento está preferiblemente entre 20 y 200 pb, preferiblemente entre 40 y 100 pb, preferiblemente alrededor de 80 pb.

El término alteración de la secuencia genómica incluye el reemplazo de uno o más nucleótidos, la inserción de uno o más nucleótidos y/o la eliminación de uno o más nucleótidos en cualquier lugar dentro de un genoma, preferiblemente dentro de un gen.

Por ejemplo, si la primera y la segunda región de homología con un gen objetivo comprenden secuencias genómicas adyacentes del gen, la sustitución de uno o más nucleótidos en la primera región de homología y/o en la segunda región de homología dará lugar a una alteración del gen que sigue a la orientación homóloga con el conjunto de construcciones de direccionamiento de acuerdo con la invención. Dicho reemplazo de uno o más nucleótidos está preferiblemente en la región de homología con el gen objetivo que está presente en la primera y en la segunda construcción.

Dicha alteración de la secuencia genómica es preferiblemente una eliminación de uno o más nucleótidos, preferiblemente en cualquier lugar dentro del gen. Por ejemplo, si la primera y la segunda región de homología con un gen objetivo comprenden secuencias genómicas del gen que están separadas en el genoma del organismo, una alteración del gen después del direccionamiento homólogo con el conjunto de construcciones de direccionamiento de acuerdo con la invención da como resultado una eliminación de la región que estaba ubicada entre la primera y la segunda región de homología en el cromosoma parental.

Dicho microorganismo es preferiblemente un microorganismo aneuploide, preferiblemente una levadura aneuploide. El término aneuploidía, como se usa en este documento, se refiere a la presencia de un número anormal de cromosomas dentro de una célula o un organismo que difiere del número normal de cromosomas para ese organismo. Un microorganismo aneuploide puede tener uno o más cromosomas extra o faltantes. El término microorganismo aneuploide incluye un microorganismo poliploide. En los hongos, se sabe que la aneuploidía confiere resistencia a los fármacos antifúngicos y permite una rápida evolución adaptativa [Calo et al., (2013). PLoS Pathog 9(3): e1003181].

Dicho microorganismo, preferiblemente microorganismo aneuploide, preferiblemente es un Ascomycota, preferiblemente una especie Saccharomycotina, preferiblemente un *Saccharomyces sensu stricto* (*Saccharomyces paradoxus*, *S. mikatae*, *S. bayanus*, *S. eubayanus*, *S. kudriavzevii*, *S. paradoxus*, *S. arboricolus*), *Kazachstania*, *Naumovozyma*, *Nakaseomyces*, *Vanderwaltozyma*, *Zygosaccharomyces*, *Lachancea*, *Kluyveromyces*, *Eremothecium*, *Torulaspora*, *Ogataea*, *Debaryomyces*, *Clavispora*, *Candida*, *Komagataella*, y/o *Yarrowia*. Un organismo preferido es la levadura de cerveza *Lager Saccharomyces pastorianus*.

Se supone que *S. pastorianus* es un híbrido de *S. cerevisiae* y *S. eubayanus*. El tamaño del genoma de *S. pastorianus* es hasta un 60% mayor que el de *S. cerevisiae*, e incluye grandes partes de los dos genomas. *S. cerevisiae* contiene un grupo haploide de 16 cromosomas, que varían en tamaño desde 200 hasta 2,200 kb. El tamaño del genoma de *S. pastorianus* es de 24-50 Mb. Adicionalmente, las especies de *Saccharomyces* aneuploides reportadas son *S. monacensis* y *S. uvarum*. Albertin y Marullo (2012) proporcionan una descripción general de los hongos poliploides. Proc R Soc B 279: 2497-2509.

Dicho marcador de selección es preferiblemente un marcador de selección auxotrópico o un marcador de selección dominante, que son conocidos por un experto. Los marcadores auxotróficos preferidos incluyen URA3, KIURA3; CaURA3; HIS3; his5; LEU2; KILEU2; LYS2; TRP1; ADE1; ADE2; y MET15. Los marcadores dominantes preferidos incluyen KanMX; Sh ble; hph CUP1; SFA1; dehH1; PDR3-9; AUR1-C; nat pat; ARO4-OFP; SMR1; FZF1-4; y DsdA. En la Tabla 1 se proporciona una descripción general de los marcadores preferidos que se utilizan habitualmente en los organismos de levadura.

Dicha primera construcción comprende preferiblemente una primera parte, preferiblemente los primeros dos tercios o la primera mitad, de una región que codifica el marcador de selección. Por ejemplo, URA3, también denominado YEL021W, codifica la enzima orotidina-5'-fosfato (OMP) descarboxilasa que cataliza la sexta etapa enzimática en la biosíntesis de novo de pirimidinas, que convierten la OMP en monofosfato de uridina (UMP). La proteína codificada tiene 267 aminoácidos, que está codificada por una secuencia de ácido nucleico de 801 pares de bases (pb). Dicha primera construcción comprende preferiblemente entre 200 y 600 pb de la región de codificación de URA3, más preferiblemente

entre 300 y 500 pb. La segunda construcción comprende preferiblemente entre 200 y 600 pb de la región de codificación de URA3, más preferiblemente entre 300 y 500 pb.

5 La región de superposición entre la primera y la segunda parte del marcador de selección está preferiblemente entre aproximadamente 20 pb y 800 pb, preferiblemente entre aproximadamente 50 pb y aproximadamente 600 pb, preferiblemente alrededor de 200 pb.

10 Un marcador de selección preferido es URA3. URA3 codifica orotidina 5-fosfato descarboxilasa (ODCasa), que es una enzima que cataliza una reacción involucrada en la síntesis de ribonucleótidos de pirimidina en el ARN de levadura. La pérdida de la actividad de ODCasa conduce a una falta de crecimiento celular a menos que se agregue uracilo o uridina a los medios. Cuando está presente un gen URA3 funcional, los microorganismos auxotróficos pueden crecer en ausencia de uracilo y/o uridina. En contraste, la adición de ácido 5-fluoroorótico en presencia de un gen URA3 funcional resulta en la formación de un compuesto tóxico, provocando la muerte de los microorganismos. Por lo tanto, URA3 permite tanto la selección positiva como la negativa.

15 Un marcador de selección preferido adicional es proporcionado por una secuencia de nucleótidos que codifica agmatina ureohidrolasa (agmatinasa) (EC.3.5.3.11) o guanidino-ácido hidrolasa (guanidinobutirasa; EC.3.5.3.7). Los microorganismos, preferiblemente de la familia Saccharomycetaceae, incluidas las cepas de *S. cerevisiae*, no pueden crecer en guanidinobutirato y/o agmatina como única fuente de nitrógeno. Tanto la guanidino-ácido hidrolasa como la agmatinasa catalizan la formación de urea, una fuente de nitrógeno comúnmente asimilada por microorganismos como la *S. cerevisiae*. Por lo tanto, la agmatinasa y la guanidinobutirasa presentan las características esenciales de un marcador seleccionable de "ganancia de función" potencial en microorganismos como *S. cerevisiae*, cuando se cultivan en guanidinobutirato y/o agmatina como única fuente de nitrógeno. Un gen de guanidinobutirasa preferido codifica una proteína que comprende la secuencia de aminoácidos de GenBank XP_456325.1, o una parte enzimáticamente activa de la misma. Un gen de agmatina ureohidrolasa preferido codifica una proteína que comprende la secuencia de aminoácidos de GenBank AAC75974.1, o una parte enzimáticamente activa de la misma.

20 Dicho marcador de selección está acoplado a un promotor que dirige la expresión del marcador de selección en el microorganismo, y un terminador que media la formación eficaz del extremo 3' del ARNm. Dicho promotor es preferiblemente un promotor de levadura, preferiblemente un promotor de levadura seleccionado de un gen glicolítico PGI1, PFK1, PFK2, FBA1, TPI1, TDH1, TDH3, PGK1, GPM1, ENO1, ENO2 y de promotor ACT1, TEF1, AgTEF2, PMA1. Dicho promotor también puede emplearse para expresar un marcador de selección dominante. Los terminadores de varios genes son conocidos por los expertos y se han empleado, por ejemplo, en vectores de expresión, incluidos los terminadores de los genes CYC1, TRP1, ADH1, MF1, FLP y D (Romanos et al., 1992). *Yeast* 8: 423-488).

30 La primera o segunda construcción de direccionamiento comprende una secuencia de codificación que codifica una endonucleasa y que está acoplada a un promotor inducible. La endonucleasa es preferiblemente una endonucleasa de corte raro tal como, por ejemplo, Pacl (secuencia de reconocimiento de objetivo 5'-TTAATTAA); Ascl (secuencia de reconocimiento de objetivos 5'-GGCGCGCC) y AsiSI (secuencia de reconocimiento de objetivos 5'-GCGATCGC). Pacl, Ascl y AsiSI están disponibles en New England Biolabs. La endonucleasa más preferiblemente es una endonucleasa de direccionamiento específico. El término endonucleasa de direccionamiento específico se refiere a una endonucleasa que se codifica como genes independientes dentro de los intrones, como una fusión con una proteína huésped, o como una inteína de auto-empalme. En la Tabla 2 se proporciona una lista preferida de endonucleasas de direccionamiento específico. Ejemplos adicionales de nucleasas de direccionamiento específico son I-Dirl, I-Njal, I-NanI, I-NitI, F-TevI, F-TevII, F-CphI, PI-Mgal, I-CsmI, que son conocidas por los expertos. En Benjamin K (solicitud de patente US2012/052582) se proporcionan ejemplos adicionales de nucleasas de direccionamiento específico. Una nucleasa de direccionamiento específico preferida es PI-PspI (New England Biolabs; secuencia de reconocimiento 5'-TGCAACAGCTATTATGGGTATTATGGGT) o PI-SceI (New England Biolabs; secuencia de reconocimiento 5'-ATCTATGTCGGGTGCGGAGAAAGAGGTAAT). Se conocen las secuencias de codificación de la mayoría de las endonucleasas de direccionamiento específico. Por ejemplo, la secuencia de codificación de PI-SceI y de PI-PspI está disponible en las bases de datos públicas (número de acceso de GenBank Z74233.1 y número de acceso de Genbank U00707.1, respectivamente). El experto en la materia entenderá que una secuencia que difiere de la secuencia disponible públicamente para una nucleasa todavía puede codificar la nucleasa. Por ejemplo, el término región codificante de PI-PspI incluye una secuencia que se desvía de la secuencia disponible públicamente, por ejemplo, por optimización de codones, pero que aún expresa una endonucleasa activa que reconoce y digiere la secuencia de reconocimiento objetivo indicada.

60 Dicha endonucleasa está bajo el control de un promotor inducible. El término promotor inducible, como se usa en el presente documento, se refiere a un promotor cuya expresión puede regularse. Los promotores inducibles son conocidos por los expertos. Ejemplos de promotores inducibles que se han empleado en levaduras son el promotor GAL1 y el promotor GAL10, ambos inducibles por galactosa, el promotor SUC2, que es inducible por sacarosa, el promotor MAL12, inducible por maltosa; el promotor CUP1, que es inducible por cobre, y los promotores tetO7 y tetO2, que son inducibles por tetraciclina [Gari et al., (1997). *Yeast* 13: 837-48; Yen et al., 2003). *Yeast* 20 1255-62]. Un promotor inducible preferido es el promotor GAL1.

65

- Un sitio de reconocimiento que comprende la secuencia de reconocimiento del objetivo para la endonucleasa, se ubica adyacente a (detrás) la primera región de homología con un gen objetivo de un microorganismo en la primera construcción. Una copia de este sitio de reconocimiento está ubicada adyacente a (delante de) la segunda región de homología con el gen objetivo del microorganismo en la segunda construcción. El experto entenderá que cuando una parte de la primera región de homología con el gen objetivo en la primera construcción se duplica entre la copia del sitio de reconocimiento de endonucleasa y la segunda región de homología con el gen objetivo en la segunda construcción, dicha copia del sitio de reconocimiento se encuentra adyacente a (frente a) la duplicación de la primera región de homología con el gen objetivo en la segunda construcción. Alternativamente, el sitio de reconocimiento está ubicado adyacente a (detrás) la parte duplicada de la segunda región de homología con el gen objetivo en la primera construcción cuando una parte de la segunda región de homología con el gen objetivo en la segunda construcción se duplica en la primera construcción. El marcador de selección, incluidas las secuencias del promotor y del terminador, y la región codificante de la endonucleasa, incluido el promotor inducible, se encuentran entre el sitio de reconocimiento en la primera construcción y la copia de este sitio de reconocimiento en la segunda construcción.
- La invención proporciona además un método para alterar un genoma, preferiblemente un gen objetivo, en un microorganismo, que comprende proporcionar el grupo de construcciones de direccionamiento de acuerdo con la invención a dicho microorganismo, y seleccionar un microorganismo en el que el genoma ha sido alterado. Dicha selección de un microorganismo en el que el genoma ha sido alterado se realiza preferiblemente mediante la selección de un microorganismo que exprese funcionalmente un marcador de selección recombinado.
- Como se indicó aquí anteriormente, la aparición de un evento de recombinación entre las construcciones de direccionamiento se mejora notablemente después de la integración de las construcciones de direccionamiento en el locus de direccionamiento correcto. Por lo tanto, la presencia de un marcador de selección funcionalmente recombinante es altamente indicativa de la presencia de construcciones de direccionamiento correctamente integradas en el genoma objetivo y, por lo tanto, de un genoma alterado en el microorganismo.
- Como se indicó aquí anteriormente, los términos alterar, alteración y alterado se refieren a la sustitución de uno o más nucleótidos, la inserción de uno o más nucleótidos y/o la eliminación de uno o más nucleótidos en cualquier lugar dentro del gen objetivo.
- Se puede realizar un reemplazo de uno o más nucleótidos alterando uno o más nucleótidos en la primera región de homología y/o en la segunda región de homología. Cuando la primera región de homología y la segunda región de homología cubren regiones adyacentes del genoma, preferiblemente un gen objetivo, la integración de los vectores de direccionamiento dará como resultado una alteración del genoma. Cuando está presente, dicha sustitución de uno o más nucleótidos se realiza preferiblemente alterando uno o más nucleótidos en la región de superposición de la homología con el genoma que está presente en la primera y en la segunda construcción.
- Dicha alteración de una secuencia genómica es preferiblemente una eliminación de uno o más nucleótidos en cualquier lugar dentro de un genoma, preferiblemente dentro de un gen. Por ejemplo, si la primera y la segunda región de homología con un genoma objetivo comprenden secuencias genómicas que se separan en el genoma del organismo, una alteración del genoma después del direccionamiento homólogo con el grupo de construcciones de direccionamiento de acuerdo con la invención resultará en una supresión de la región que estaba ubicada entre la primera y la segunda región de homología en el cromosoma parental.
- La invención proporciona además un método para producir un microorganismo que comprende un genoma alterado, preferiblemente un gen alterado, el método que comprende proporcionar el conjunto de construcciones de direccionamiento de acuerdo con la invención a dicho microorganismo, y seleccionar un microorganismo en el que el genoma ha sido alterado y que expresa funcionalmente un marcador de selección recombinada.
- El método para producir un microorganismo que comprende un genoma alterado comprende preferiblemente inducir el promotor inducible para la expresión de la endonucleasa, eliminando así el marcador de selección y la región codificante de la endonucleasa, incluido el promotor inducible, del genoma objetivo.
- La divulgación proporciona además un microorganismo, que comprende una alteración genómica que se produce mediante los métodos de la invención. Cuando están presentes, las regiones duplicadas de homología con el genoma objetivo en la primera y segunda construcción de direccionamiento aseguran la eliminación sin interrupciones del marcador del genoma objetivo mediante recombinación homóloga. El microorganismo resultante comprende solo la alteración o alteraciones que estaban presentes en la primera y/o segunda construcción de direccionamiento, o que fueron inducidas por recombinación de las construcciones de direccionamiento en el genoma de direccionamiento, como una inserción en el genoma de direccionamiento o una eliminación del genoma de direccionamiento.
- La invención proporciona además un microorganismo, que comprende una alteración genómica, preferiblemente una alteración de un gen objetivo, la alteración que comprende una inserción de un marcador de selección recombinante funcionalmente y una secuencia de codificación para una endonucleasa que está acoplada a un promotor inducible, por lo que el genoma objetivo comprende una copia de una secuencia de reconocimiento para la endonucleasa en ambos sitios de la inserción.

La invención proporciona además un método para producir un microorganismo que comprende un genoma alterado, el método comprende proporcionar un microorganismo que comprende una alteración del genoma, preferiblemente de un gen objetivo, la alteración comprende una inserción de un marcador de selección recombinante funcionalmente y una secuencia de codificación para una endonucleasa que está acoplada a un promotor inducible, por lo que el genoma objetivo comprende una copia de una secuencia de reconocimiento para la endonucleasa en ambos sitios de la inserción, e induce al promotor inducible a eliminar las secuencias de ácido nucleico entre las secuencias de reconocimiento de la endonucleasa. Nuevamente, cuando están presentes, las regiones duplicadas de homología con el gen objetivo en la primera y la segunda construcción de direccionamiento aseguran la eliminación sin interrupción del marcador del genoma objetivo mediante recombinación homóloga al proporcionar al ADN genómico una pequeña pieza homóloga para volver a conectar las cadenas de ADN rotas eficientemente. El microorganismo resultante comprende solo la alteración o alteraciones que estaban presentes en la primera y/o segunda construcción de direccionamiento, o que fueron inducidas por recombinación de las construcciones de direccionamiento en el genoma de direccionamiento, como una inserción en el genoma de direccionamiento o una eliminación del genoma, preferiblemente una inserción en un gen objetivo o una eliminación del gen objetivo o una eliminación de dentro del gen objetivo.

Tabla 1

Marcadores auxotróficos			
Gen marcador	Modo de acción	Reciclabe/método	Referencia
URA3	Repara deficiencia de uracilo	Sí/Selección negativa con 5-FOA	Alani E, Cao L & Kleckner N (1987) A method for gene disruption that allows repeated use of URA3 selection in the construction of multiply disrupted yeast strains. <i>Genetics</i> 116: 541-545. Langlerouault F & Jacobs E (1995) A method for performing precise alterations in the yeast genome using a recyclable selectable marker. <i>Nucleic Acids Res</i> 23: 3079-3081.
KIURA3	Repara deficiencia de uracilo	Sí/Selección negativa con 5-FOA	Shuster JR, Moyer D & Irvine B (1987) Sequence of the <i>Kluyveromyces lactis</i> URA3 gene. <i>Nucleic Acids Res</i> 15: 8573-8573.
CaURA3	Repara deficiencia de uracilo	Sí/Selección negativa con 5-FOA	Losberger C & Ernst JF (1989) Sequence and transcript analysis of the <i>C. albicans</i> URA3 gene encoding Orotidine-5'Phosphate Decarboxylase. <i>Curr Genet</i> 16: 153-157.
HIS3	Repara deficiencia de histidina	No/--	Wach A, Brachat A, Alberti-Segui C, Rebischung C & Philippsen P (1997) Heterologous HIS3 marker and GFP reporter modules for PCR-targeting in <i>Saccharomyces cerevisiae</i> . <i>Yeast</i> 13: 1065-1075.
his5	Repara deficiencia de histidina	No/--	Wach A, Brachat A, Alberti-Segui C, Rebischung C & Philippsen P (1997) Heterologous HIS3 marker and GFP reporter modules for PCR-targeting in <i>Saccharomyces cerevisiae</i> . <i>Yeast</i> 13: 1065-1075.
LEU2	Repara deficiencia de leucina	No/--	Brachmann CB, Davies A, Cost GJ, Caputo E, Li JC, Hieter P & Boeke JD (1998) Designer deletion strains derived from <i>Saccharomyces cerevisiae</i> S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. <i>Yeast</i> 14: 115-132.
KILEU2	Repara deficiencia de leucina	No/--	Zhang YP, Chen XJ, Li YY & Fukuhara H (1992) LEU2 gene homolog in <i>Kluyveromyces lactis</i> . <i>Yeast</i> 8: 801-804.
LYS2	Repara deficiencia de lisina	Sí/Selección negativa con alfaaminoadipato	Chattoo BB, Sherman F, Azubalis DA, Fjellstedt TA, Mehnert D & Ogur M (1979) Selection of <i>lys2</i> mutants of the yeast <i>Saccharomyces cerevisiae</i> by the utilization of alphaaminoadipate. <i>Genetics</i> 93: 51-65.

ES 2 694 349 T3

TPP1	Repara deficiencia de triptofano	No/--	Brachmann CB, Davies A, Cost GJ, Caputo E, Li JC, Hieter P & Boeke JD (1998) Designer deletion strains derived from <i>Saccharomyces cerevisiae</i> S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. <i>Yeast</i> 14: 115-132.
ADE1	Repara deficiencia de adenina	No/--	Nakayashiki T, Ebihara K, Bannai H & Nakamura Y (2001) Yeast [PSI+] "prions" that are cross-transmissible and susceptible beyond a species
ADE2	Repara deficiencia de adenina	No/--	Brachmann CB, Davies A, Cost GJ, Caputo E, Li JC, Hieter P & Boeke JD (1998) Designer deletion strains derived from <i>Saccharomyces cerevisiae</i> S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. <i>Yeast</i> 14: 115-132.
MET15	Repara deficiencia de metionina	Si/Selección negativa con metil-mercurio	Singh A & Sherman F (1974) Association of methionine requirement with methyl mercury resistant mutants of yeast. <i>Nature</i> 247: 227-229.
Marcadores dominantes			
Gen marcador	Modo de acción	Reciclabe/método	Referencia
KanMX	Resistencia a G418	No/--	Wach A, Brachat A, Pohlmann R & Philippsen P (1994) New heterologous modules for classical or PCR-based gene disruptions in <i>Saccharomyces cerevisiae</i> . <i>Yeast</i> 10: 1793-1808.
ble	Resistencia a fleomicina	No/--	Gatignol A, Baron M & Tiraby G (1987) Phleomycin resistance encoded by the ble gene from transposon Tn5 as a dominant selectable marker in <i>Saccharomyces cerevisiae</i> . <i>Mol Genet</i> 207: 342-348.
Sh ble	Resistencia a Zeocina	No/--	Drocourt D, Calmels T, Reynes JP, Baron M & Tiraby G (1990) Cassettes of the <i>Streptoalloteichus hindustanus</i> ble gene for transformation of lower and higher eukaryotes to phleomycin resistance. <i>Nucleic Acids Res</i> 18: 4009-4009.
hph	Resistencia a higromicina	No/--	Gritz L & Davies J (1983) Plasmid encoded Higromicina-B resistance the sequence of Higromicina-B phosphotransferase gene and its expression in <i>Escherichia coli</i> and <i>Saccharomyces cerevisiae</i> . <i>Gene</i> 25: 179-188.
Cat	Resistencia a cloramfenicol	No/--	Hadfield C, Cashmore AM & Meacock PA (1986) An efficient cloramfenicol resistance marker for <i>Saccharomyces cerevisiae</i> and <i>Escherichia coli</i> . <i>Gene</i> 45: 149-158.
CUP1	Resistencia a Cu ²⁺	No/--	Henderson RCA, Cox BS & Tubb R (1985) The transformation of brewing yeasts with a plasmid containing the gene for copper resistance. <i>Curr Genet</i> 9: 133-138.
SFA1	Resistencia a formaldehido	No/--	Van den Berg MA & Steensma HY (1997) Expression cassettes for formaldehido and fluoroacetato resistance, two marcadores dominantes in <i>Saccharomyces cerevisiae</i> . <i>Yeast</i> 13: 551-559.

ES 2 694 349 T3

deh1	Resistencia a fluoroacetato	No/--	Van den Berg MA & Steensma HY (1997) Expression cassettes for formaldehido and fluoroacetato resistance, two marcadores dominantes in <i>Saccharomyces cerevisiae</i> . <i>Yeast</i> 13: 551-559.
PDR3-9	Resistencia a múltiples fármacos	No/--	Lackova D & Subik J (1999) Use of mutated PDR3 gene as a dominant selectable marker in transformation of prototrophic yeast strains. <i>Folia Microbiol</i> 44: 171-176.
AUR1-C	Resistencia a aureobasidina	No/--	Hashida-Okado T, Ogawa A, Kato I & Takesako K (1998) Transformation system for prototrophic industrial yeasts using the AUR1 gene as a dominant selection marker. <i>FEBS Lett</i> 425: 117-122.
nat	Resistencia a nourseotricina	No/--	Goldstein AL & McCusker JH (1999) Three new dominant drug resistance cassettes for gene disruption in <i>Saccharomyces cerevisiae</i> . <i>Yeast</i> 15: 1541-1553.
CYH2	Resistencia a cicloheximida	No/--	Delpozo L, Abarca D, Claros MG & Jimenez A (1991) Cicloheximida resistance as a yeast cloning marker. <i>Curr Genet</i> 19: 353-358.
pat	Resistencia a bialafos	No/--	Goldstein AL & McCusker JH (1999) Three new dominant drug resistance cassettes for gene disruption in <i>Saccharomyces cerevisiae</i> . <i>Yeast</i> 15: 1541-1553.
ARO4OFP	Resistencia a o-Fluoro-DL-fenilalanina	No/--	Cebollero E & Gonzalez R (2004) Comparison of two alternative dominant selectable markers for wine yeast transformation. <i>Appl Environ Microb</i> 70: 7018-7023.
SMR1	Resistencia a metil sulfometurón	No/--	Xie Q & Jimenez A (1996) Molecular cloning of a novel allele of SMR1 which determines sulfometuron methyl resistance in <i>Saccharomyces cerevisiae</i> . <i>FEMS Microbiol Lett</i> 137: 165-168.
<i>FZF1-4</i>	Aumento de tolerancia a sulfito	No/--	Cebollero E & Gonzalez R (2004) Comparison of two alternative dominant selectable markers for wine yeast transformation. <i>Appl Environ Microb</i> 70: 7018-7023.
<i>DsdA</i>	Resistencia a D-Serina	No/--	Vorachek-Warren MK & McCusker JH (2004) <i>DsdA</i> (D-serine deaminase): a new heterologous MX cassette for gene disruption and selection in <i>Saccharomyces cerevisiae</i> . <i>Yeast</i> 21: 163-171.
<i>amdS</i>	Permitir el crecimiento en acetamida	Sí/ Fluoroacetamida	Selten G, Swinkels B & Van Gorcom R (2000) Selection Gen marcador free recombinant strains: A method for obtaining them and the use of these strains. Pat.no.US6051431. Swinkels B, Selten G, Bakhuis J, Bovenberg R & Vollebregt A (1997) The use of homologous <i>amdS</i> genes as selectable markers. Patent application WO97/06261.

Tabla 2

Enzima	Secuencia de reconocimiento	Corte	SF	Fuente	D	SCL
I-Anil	5' TTGAGGAGGTTTCTCTGTAAT AA 3' 3' AACTCCTCCAAGAGACATTTA TT 5'	5' ---TTGAGGAGGTTTC TCTGTAATAA--- 3' 3' ---AACTCCTCC AAAGAGACATTTATT--- 5'	HI	Aspergillus nidulans	E	mfo
I-Ceul	5' TAACTATAACGGTCCTAAGTA GCCA 3' 3' ATTGATATTGCCAGGATTCCAT CGCT 5'	5' ---TAACTATAACGGTCCTAA GGTAGCGA--- 3 3' ---ATTGATATTGCCAG GATTCCATCGCT--- 5'	HI	Chlamydomonas eugametos	E	cloro
I-Chul	5' GAAGGTTTGGCACCTCGATGTC GGCTCATC 3' 3' CTTCCAACCGTGGAGCTACAG CCGAGTAG 5'	5' ---GAAGGTTTGGCACCTCG ATGTCGGCTCATC--- 3' 3' ---CTTCCAACCGTG GAGCTACAGCCGAGTAG--- 5'	HI	Chlamydomonas humicola	E	cloro

Enzima	Secuencia de reconocimiento	Corte	SF	Fuente	D	SCL
I-Cpal	5' CGATCCTAAGGTAGCGAAATTC A 3' 3' GCTAGGATTCCATCGCTTTAAG T 5'	5' ---CGATCCTAAGGTAGCGAA ATTCA--- 3' 3' ---GCTAGGATTCCATC GCTTTAAGT--- 5'	HI	Chlamydomonas pallidostigmata	E	cloro
I-Cpall	5' CCCGGCTAACTCTGTGCCAG 3' 3' GGGCCGATTGAGACACGGTC 5'	5' ---CCCGGCTAACTCTGTGCCAG-- 3' 5' ---GGCCGATTGAGACACGGTC--- 3'	HI	Chlamydomonas pallidostigmata	E	cloro
I-Crel	5' CTGGGTTCAAAAACGTGCTGAG ACAGTTTGG 3' 3' GACCCAAAGTTTGCAGCACTCT GTCAAACC 5'	5' ---CTGGGTTCAAAAACGTGCTGTA GACAGTTTGG--- 3' 3' ---GACCCAAAGTTTGCAG CACTCTGTCAAACC--- 5'	HI	Chlamydomonas <u>reinhardtii</u>	E	cloro
Dmol	5' ATGCCTTGCCGGGTAAGTTCCG GCGGCAT 3' 3' TACGGAACGGCCCATCAAGG CCGGCGTA 5'	5' --- ATGCCTTGCCGGGTAA GTTCCGG CGGCAT--- 3' 3' --- TACGGAACGGCC CATTCAAGGCC GGCGTA--- 5'		<u>Desulfurococcus</u> <u>mobilis</u>	A	Chrm

Enzima	Secuencia de reconocimiento	Corte	SF	Fuente	D	SCL
<u>H-DreI</u>	5' CAAAACGTCGTAAGTTCCGGC GCG 3' 3' GTTTTCAGCATTCAAGGCCG GC 5'	5' --- CAAAACGTCGTAA GTTCCGGCGC G--- 3' 3' --- GTTTTCAG CATTCAAGGCCGG C--- 5'	Hi	Escherichia coli	B	
<u>H-HmuI</u>	5' AGTAATGAGCCTAACGCTCAGC AA 3' 3' TCATTACTCGGATTGCGAGTCG TT 5'	:* 3--- TCATTACTCGGATTGC GAGTCGTT --- 5'	Hi II	<u>Bacillus subtilis</u> SPO 1	B	fago
<u>H-HmuII</u>	5' AGTAATGAGCCTAACGCTCAAC AA 3' 3' TCATTACTCGGATTGCGGAGTT GTT 5'	3' --- TCATTACTCGGATTGCGGAGTTGTTN NNNN--- 5'	Hi II	fago SP82 <u>Bacillus subtilis</u>	B	fago

Enzima	Secuencia de reconocimiento	Corte	SF	Fuente	D	SCL
<u>H-Ial</u>	5' CACATCCATAAACCATATCATTT TT 3' 3' GTGTAGGTATTGGTATAGTAAA AA 5'	5'---CACATCCATAA CCATATCATTTTT---3' 3' --- GTGTAGGTATTGGTATAGTAA AA A---5'	HI II	<u>Lactococcus lactis</u>	B	chrm
<u>I-Msol</u>	5' CTGGGTTCAAAAACGTCGTGAG ACAGTTTGG 3' 3' GACCCAAAGTTTTCAGCACTCT GTCAAACC 5'	5' --- CTGGGTTCAAAAACGTCGTGA GAC AGTTTGG---3' 3' --- GACCCAAAGTTTTCAG CACTCTGT CAAACC---5'		<u>Monomastix sp.</u>	E	
<u>PI-Pful</u>	5' GAAGATGGGAGGAGGACCGG ACTCAACTT 3' 3' CTTCTACCCCTCCTCCCTGGCCT GAGTTGAA 5'	5' --- GAAGATGGGAGGAGGG ACCGGAC TCAACTT---3' 3' --- CTTCTACCCCTCC TCCCTGGCCTGA GTTGAA---5'		<u>Pyrococcus furiosus Vc1</u>	A	
<u>PI-Pkoll</u>	5' CAGTACTACGGTTAC3' 3' GTCATGATGCCAATG5'	5'---CAGTACTACG GTTAC---3' 3'---GTCATG ATGCCAATG-5'		<u>Pyrococcus kodakaraensis KOD 1</u>	A	

Enzima	Secuencia de reconocimiento	Corte	SF	Fuente	D	SCL
<u>I-PorI</u>	5' GCGAGCCCGTAAGGGTGTGTA CGGG 3'	5' GGGAGCCCGTAAGGGT GTGTACG GG---3' 3' --- CGCTCGGGCATTCCCACACATGC CC---5'	HI II	<u>Pyrobaculum organotrophum</u>	A	chr m
<u>I-PpoI</u>	5' TAACTATGACTCTCTTAAGGTA GCCAAAT 3'	5' TAACTATGACTCTCTTAA GGTAGC CAAAT---3' 3' --- ATTGATACTGAGAG AATTCCATCG GTTTA---5'	HI V	<u>Physarum polycephalum</u>	E	nucle ar
<u>PI-Pspl</u>	5' TGGCAAACAGCTATTATGGGTA TTATGGGT 3'	5' TGGCAAACAGCTATTAT GGGTATT ATGGGT---3' 3' --- ACCGTTTGTGCGATAATACCCAT AATACCCA	HI	<u>Pyrococcus sp.</u>	A	chr m
<u>I-Scal</u>	5'	5'	HI	<u>Saccharomyces capensis</u>	E	mito

Enzima	Secuencia de reconocimiento	Corte	SF	Fuente	D	SCL
	TGTCACATGAGGTGCACTAGT TATTAC 3' ACAGTGTAAGTCCACGTGATCA ATAATG 5'	TGTCACATGAGGTGCACT AGTTA TTAC--- 3' --- ACAGTGTAAGTCCAC GTGATCAAT AATG--- 5'				
<u>I-SceI</u>	5' AGTTACGCTAGGGATAACAGG GTAATATAG 3' TCAATGCGATCCCTATTGTCCC ATTATATC 5'	5' --- AGTTACGCTAGGGATAA CAGGGT AATATAG--- 3' --- TCAATGCGATCCC TATTGTCCCAT TATATC--- 5'	HI	Saccharom yces cerevisiae	E	mito
<u>Pi-SceI</u>	5' ATCTATGTCGGGTGCGGAGAA AGAGGTAATGAAATGGCA 3' TAGATACAGCCACGCCCTCTTT CTCCATTACTTTACCGT 5'	5' --- ATCTATGTCGGGTGC GGAGAAAG AGGTAATGAAATGGCA--- 3' --- TAGATACAGCC CACGCCCTCTTTCT CCATTACTTTACCGT--- 5'	HI	Saccharom yces cerevisiae	E	
<u>I-SceI</u>	5' TTTTGATTCTTTGGTCACCCCTG AAGTATA 5'	5' --- TTTTGATTCTTTGGTCACCC TGAA GTATA--- 3'	HI	Saccharom yces cerevisiae	E	mito

Enzima	Secuencia de reconocimiento	Corte	SF	Fuente	D	SCL
	3' AAAATAAGAAACCAGTGGGA CTTCATAT	3' AAAATAAGAAACCAG TGGGACT TCATAT--- 5'				
<u>I-SceII</u>	5' ATTGGAGGTTTGGTAACTATT TATTACC 3' TAACCTCCAAAACCATTGATAA ATAATGG	5' --- ATTGGAGGTTTGGTAAC TATTTA TTACC--- 3' --- TAACCTCCAAAACC ATTGATAAAT AATGG--- 5'	HI	<u>Saccharom yces cerevisiae</u>	E	<u>mito</u>
<u>I-SceIV</u>	5' TCTTTTCTCTTGATTAGCCCTA ATCTACG 3' AGAAAAGAGAACTAATCGGGA TTAGATGC	5' --- TCTTTTCTCTTGATTA GCCCTAAT CTACG--- 3' --- AGAAAAGAGAAC TAATCGGGATT AGATGC--- 5'	HI	<u>Saccharomyces cerevisiae</u>	E	<u>mito</u>
<u>I-SceV</u>	5' AATAATTTTCTTCTTAGTAATG CC 3' TTATTAAGAAGAATCATTACGG	5' ---AATAATTTTCT TCTTAGTAATGCC-- 3' 3' --- TTATTAAGAAGAATCATTACGG --- 5'	HI II	<u>Saccharomyces cerevisiae</u>	E	<u>mito</u>

Enzima	Secuencia de reconocimiento	Corte	SF	Fuente	D	SCL
<u>I-SceVI</u>	5' GTTATTAAATGTTTAGTAGTT GG 3' CAATAAATTACAAAATCATCAA CC	5' --- GTTATTTAATG TTTTAGTAGTTGG --- 3' 3' --- CAATAAATTACAAAATCATCA ACC ---	HI II	<u>Saccharom yces cerevisiae</u>	<u>E</u>	<u>mito</u>
<u>I-SceVII</u>	5' TGTCACATGAGGGTGCACTAGT TATTAC 3' ACAGTGTAACCTCCACGTTGATCA ATAATG	desconocido	HI	<u>Saccharomyces cerevisiae</u>	<u>E</u>	<u>mito</u>
<u>I-Ssp6803</u>	5' GTCGGGCTCATAACCCGAA 3' CAGCCCGAGTATTGGGCTT	5' ---GTCGGGCT CATAACCCGAA--- 3' 3' ---CAGCCCGAGTA TTGGGCTT-5'		<u>Synechocystis sp. PCC 6803</u>	<u>B</u>	
<u>I-TevI</u>	5' AGTGGTATCAACGCTCAGTAGA TG 3' TCACCATAGT	5' AGTGGTATCAAC GCTCAGTAGAT G--- 3' TCACCATAGT TGGGAGTCATCTAC-5'	HI I	<u>fago T4</u> <u>Escherichia coli</u>	<u>B</u>	<u>fago</u>
	TGCGAGTCATCTAC					

Enzima	Secuencia de reconocimiento	Corte	SF	Fuente	D	SCL
<u>I-TevII</u>	5' GCTTATGAGTATGAAGTGAACA CGTTATTC 3' CGAATACTCATACTTCACTTGT GCAATAAG	5' --- GCTTATGAGTATGAAGTGAACAGT TATTC--- 3' 3' --- CGAATACTCATACTTCACTTGT C AATAAG--- 5'	HI I	fase T4 <u>Escherichia coli</u>	B	fase
<u>I-TevIII</u>	5' TATGTATCTTTTGGGTGTACCT TTAACTTC 3' ATACATAGAAAACGCACATGGA AATTGAAG	5' --- T ATGTATCTTTTGGGTGTACCTTT AACTTC--- 3' 3' --- AT ACATAGAAAACGCACATGGAA AATTGAAG--- 5'	HI II	fase RB3 <u>Escherichia coli</u>	B	fase
<u>PI-TIII</u>	5' TAYGCNGAYACNGACGGYTTY T 3' ATRCGNCTRTGNCTGCCTAARA	5' --- TAYGCNGAYACNGACGG YTTYT-3' 3' ---ATRCGNCTRTGNC TGCCTAARA-5'	HI	<u>Thermococcus</u> <u>litoralis</u>	A	chrm

Enzima	Secuencia de reconocimiento	Corte	SF	Fuente	D	SCL
<u>PI-TIII</u>	5' AAATGCTTGC AAACAGCTATT ACGGCTAT 3' TTTAACGAACGTTGTGATAA TGCCGATA	Desconocido**	HI	<i>Thermococcus</i> <i>litoralis</i>	A	<u>chrm</u>
<u>I-Tsp061I</u>	5' CTTCAGTATGCCCGAAAC 3' GAAGTCATACGGGCTTTG	5' ---CTTCAGTAT GCCCGAAAC--- 3' 3' ---GAAGT CATACGGGCTTTG-5'		<u><i>Thermoproteus</i> sp. IC-061</u>	A	
<u>I-Vdi141I</u>	5' CCTGACTCTCTTAAGGTAGCCA AA 3' GGACTGAGAGAATTCCATCGG TTT	5' --- CCTGACTCTCTTAA GG TAGCCAAA --- 3' 3' --- GGACTGAG AGAATTCCATCGGTT T--- 5'		Vulcanisaeta distributa IC-141	A	

Tabla 2. Descripción general del direccionamiento específico de endonucleasas y sus secuencias objetivo.

Abreviaturas: SF (familia estructural); HI: familia LAGLIDADG; HII: familia GIY-YIG; HIII: familia H-N-H; VIH: familia de caja de His-Cys.

D: Dominio biológico de la fuente: A: archaea; B: bacterias; E: eukarya. SCL: Ubicación del subcelular: cloro: cloroplasto; chrm: cromosómico; mito: mitocondrial; nuclear: extracromosómico nuclear; fago: bacteriófago.

Leyenda de las figuras

Figura 1

Vector 1 y 2 con todas las partes esenciales para el cassette de eliminación estándar. La superposición de 400 bases en el marcador de selección amdS (indicado por una cruz) está diseñada para recombinarse debido a la homología.

Figura 2

El vector de clonación TOPO Blunt no direccional y pUC19 se utilizan para la construcción del plásmido con el vector 1 y 2.

Figura 3

3A: Eliminación de genes dirigida utilizando el casete amdS-I-SceI con flancos homólogos de 500 pb para modificar ScARO10.

3B: Eliminación del marcador sin interrupciones utilizando el GAL1p de I-SceI para obtener la proteína I-SceI de endonucleasa activa, que restringe el ADN genómico en el casete que introdujo los sitios de reconocimiento SceI.

Figura 4

Los vectores 1 y 2 comprenden fragmentos superpuestos del marcador de selección KIGBU 1, que codifican una guanidinobutirasa.

EJEMPLOS

Ejemplo 1

Materiales y métodos

Cepas y condiciones de cultivo.

Todas las cepas de *Saccharomyces* utilizadas en este estudio se enumeran en la Tabla I. Para el crecimiento en medios líquidos, se colocaron matraces de agitación en el agitador de la incubadora a 200 rpm y 30°C. Para el crecimiento en placas, todas las cepas excepto CBS1483 se cultivaron a 30°C CBS1483 se cultivó a 20°C. Las cepas se cultivaron en diferentes medios: en condiciones no selectivas, la levadura se cultivó en medio complejo Pextrona de levadura (YPD) que contenía 10 g/l de extracto de levadura, 20 g/l de peptona y 22 g/l glucosa 1 hidrato, pH 6), Medios sintéticos (SM) que contienen 5.0 g/l (NH₄)₂SO₄, 3.0 g/l KH₂PO₄, 0.5 g/l MgSO₄·7H₂O, 1 ml/l de solución de elementos traza y 1 ml/l de una solución de vitamina (Verduyn et al., 1992) se utilizaron [Verduyn et al., (1990). *J General Microbiology* 136: 395-403].

Cuando se usó amdSYM (AgTEF2-amdS-AgTEF2ter) como marcador, (NH₄)₂SO₄ se reemplazó por 0.6 g/l de acetamida como fuente de nitrógeno y 6.6 g/l K₂SO₄ para compensar el suministro de sulfato (SM-Ac). Las células recicladas sin marcador se seleccionaron en SM que contenía 2.3 g/l de fluoroacetamida (SM-Fac). SM, SM-Ac y SM-Fac se complementaron con 20 mg/l de adenina y 15 mg/l de L-canavanina sulfato cuando fue necesario. En todos los experimentos, se utilizaron 20 g/l de glucosa como fuente de carbono. El pH en todos los medios se ajustó a 6.0 con KOH. Los medios sólidos se prepararon añadiendo agar al 2% a los medios descritos anteriormente.

Para la inducción del GAL1p, la fuente de carbono glucosa fue reemplazada por la fuente de carbono alternativa galactosa con una concentración de 20 g/l. Cuando se usó medio líquido de galactosa, se estimuló el crecimiento con la adición de glucosa al 0.0125%. Para fines de almacenamiento, las levaduras que habían ensamblado correctamente los fragmentos de transformación en el ADN genómico se cultivaron en un cultivo líquido YPD o MM. Las levaduras con un plásmido correctamente ensamblado se cultivaron en circunstancias selectivas en un cultivo MM. Después de alcanzar la cantidad suficiente de OD₆₆₀, se agregó glicerol para obtener una concentración del 30% y las células se almacenaron a -80°C.

Técnicas de ADN

Reacción en cadena polimerizada (PCR)

Se realizaron dos métodos de PCR diferentes. Para fines de secuenciación y clonación, se aplicó la ADN polimerasa de alta fidelidad de Phusion (Finnzymes, Vantaa, Finlandia) para obtener una lectura más precisa, ya que tiene una actividad de corrección de la exonucleasa de 3' a 5'. Para el procedimiento de selección (PCR de colonias de levadura) se usó la mezcla maestra de PCR DreamTaq (Fermentas GmbH, St. Leon-Rot, Alemania), que carece de la actividad de corrección de la exonucleasa de 3' a 5'.

Tabla I. Especies de *Saccharomyces* utilizadas en el Ejemplo 1.

Cepa	Genotipo o descripción	Fuente o referencia
CEN.PK113-7D	MATa MAL2-8c SUC2	van Dijken et al (2000), Entian & Kotter (2007)
CEN.PK122	Diploide	van Dijken et al (2000), Entian & Kotter (2007)
PRG410	Cepa de cerveza Ale <i>Saccharomyces cerevisiae</i>	Cortesía del Dr. JM Geertman (Heineken Supply Chain, Zoeterwoude Países Bajos)
CMBS33	Cepa de cerveza Lager <i>Saccharomyces pastorianus</i>	KU Leuvena
CBS1483	Cepa de cerveza Lager <i>Saccharomyces pastorianus</i>	The Centraalbureau voor Schimmelcultures (CBS) Fungal Biodiversity Centre, (http://www.cbs.knaw.nl/index.php)
CBS12357	<i>Saccharomyces eubayanus</i> sp.nov	The Centraalbureau voor Schimmelcultures (CBS) Fungal Biodiversity Centre, (http://www.cbs.knaw.nl/index.php) Libkind et al. (2011)
IMK486	MATa MAL2-8c SUC2 ar010Δ-AgTEF2pr-amdSAgTEF2ter-GAL1p-ISCEI-CYC 1ter	Este estudio
IMK487	MATa MAL2-8c SUC2 ar010Δ-AgTEF2pr-amdSAgTEF2ter-GAL1p-ISCEI-CYC 1ter	Este estudio
IMK488	PRG410 aro10Δ-AgTEF2pr-amdSAgTEF2ter-GAL 1p-ISCEI-CYC1ter	Este estudio
IMK489	CMBS33 aro10Δ-AgTEF2pr-amdSAgTEF2ter-GAL 1p-ISCEI-CYC1ter	Este estudio
IMK490	CBS1483 aro10Δ-AgTEF2pr-amdSAgTEF2ter-GAL 1p-ISCEI-CYC1ter	Este estudio

^a K.U. Leuven, Centro para la Colección de Malteado y Elaboración de Cerveza, Centro de Malteado y Elaboración de Cerveza, Lovaina, Bélgica

Referencias: van Dijken et al., (2000). *Enzyme Microb Technol* 26: 706-714; Entian and Kotter, (2007). 25 yeast genetic strain and plasmid collections. Academic Press, Amsterdam, the Netherlands. In: *Methods in Microbiology* (Stansfield I & Stark J, eds) 36: 629-666 The Centraalbureau voor Schimmelcultures (CBS) Fungal Biodiversity Centre, (<http://www.cbs.knaw.nl/index.php>), Libkind et al., (2011). *PNAS USA* 108: 14539-14544

Para generar una plantilla de ADN adecuada al realizar la PCR de la colonia de levadura, una pequeña fracción de la colonia única se resuspendió en 15 µl de NaOH 0.02 N y la suspensión de la colonia de levadura se hirvió durante 10 minutos a 100°C. Después de qué 2 µl de esto la suspensión celular se utilizó como plantilla para la PCR que se realizó con la mezcla maestra de PCR DreamTaq (Fermentas GmbH, St. Leon-Rot, Alemania) siguiendo las recomendaciones del fabricante.

La PCR de colonias de alta fidelidad se utilizó para confirmar la inserción del casete de eliminación génica en el ADN genómico. Las células de un cultivo de levadura líquida (200 µl) se centrifugaron, se resuspendieron en 100 µl de LiAc 0.2 M con 1% de solución de SDS y se incubaron durante 5 minutos a 70°C. La adición de etanol al 96% y se sometió a vórtice a fondo, fue seguido de centrifugado durante 3 minutos a 15,000 g. El sedimento se lavó con etanol al 70% y se volvió a centrifugar. Se eliminó el etanol y se secó el sedimento a una temperatura máxima de 30°C, después de lo cual se agregaron 100 µl de tampón TE para disolver el sedimento. Los residuos celulares se centrifugaron durante 30 segundos a 15,000 g y se usaron 5 µl del sobrenadante transferido para la PCR de colonia de alta fidelidad.

Todos los cebadores utilizados en la PCR se describen en la Tabla II.

Tabla II: Secuencias de los cebadores utilizados

Cebadores de integración de casete	
KanA r	CGCACGTCAAGACTGTCAAG
KanB f	TCGTATGTGAATGCTGGTCC
KanA f	CTTGACAGTCTTGACGTGCG
KanB r	CGACCAGCATTCACATACGA
Fw 5'ScARO10/amdS check	ACAAGTTGACGCGACTTCTGTAAAG
Rv 3'ScARO10/amdS check	CAACTGGACAAAGAACTCTGTGGTAG
FK072	CTCGAGTCATGTAATTAGTTATG
ScARO10-Fw interno	GGTGTGGCCAAAGTCCATAG
ScARO10-Rv interno	CCTGTTTCACAAAACGACAACATC
Nombre de cebador	
Cebadores TOPO Blunt	
M13f	GTAAACGACGGCCAG
M13r	CAGGAAACAGCTATGA
Fw V1 Ndel	GCGCATATCGTCCAGCAGAACCGTCAGCA
Rv V1 SclI SacI	GCGGAGCTATTACCCTGTTATCCCTAAGGCTAGAGATCTGTTTAGCTTGCC
Fw V2 Ndel	GCGCATATGAGATTGCCATACGCTAAGATGG
Rv V2 sobrepone IScel	GCGAGGCACGGAGGAGAGTCTTCCGTCGGAGGGCTGTCCGCCCGCTCGGGCGG
Fw V2 IScel	CTTCTAATCCGTTATTAAGGGTCTCTCGAGAGCTCC
Rv V2 SclI EcoRI	GCGACGGATTAGAACGCCCGCGAG
Fw 5' ScARO10 KpnI	GCGGAATTCATTACCCTGTTATCCCTACAAATTAAGCCCTTCGAGCGTCC
Fw 5'ScARO10 NotI	GCGGTACCATGGCACCTGTTACAATTGAAAAGTTCC
Rv 5' ScARO10 SacI	GCGGCGGCCGCGATGGCACCTGTTACAATTGAAAAGTTCC
Fw 3' ScARO10 BamHI 80bp	GCGGAGCTCCTACCGAGCAAGCGACTCTATCTT
	GCGGGATCCAGCTACATGATTCAAATTTTAAAGGGCCAAATCATAAAGTATAT
	CATGATATGGTAAAAGATAGAGTCGTTGCTCGGTACATTTGGCATGGCCCTT
	CCTG
Rv 3' ScARO10 PstI	GCGCTGCAGCTATTTTTTATTTCTTTTAAAGTCCCGCTGC

Restricción

Las enzimas de restricción utilizadas en este proyecto fueron suministradas por Fermentas. La mezcla de digestión se preparó en un volumen total de 20 μ l o 30 μ l dependiendo de la cantidad final deseada. Una mezcla general contenía 1 μ l de cada enzima de restricción (1 unidad FastDigest/ μ l), 2 o 3 μ l de 10 x de tampón FastDigest y agua desmineralizada libre de nucleasas. Se utilizaron aproximadamente 50-200 ng de ADN en una mezcla total de 20 μ l y 100 ng-2 μ g en una mezcla total de 30 μ l. La mezcla de digestión se incubó a una temperatura de 37°C.

Electroforesis en gel

Para analizar diferentes técnicas de ADN, las moléculas de ácido nucleico se cargaron en gel de agarosa al 1% (teñidas con SybrSafe al 0.001% por adelantado) en tampón TAE 1x (acetato de Tris 40 mM y EDTA 1 mM). Todas las moléculas de ADN se separaron mediante la aplicación de un campo eléctrico de 100 V durante 30 min. Se tomaron imágenes de gel exponiéndolas a luz UV.

Recuperación de Gel

Para purificar el ADN del gel, se utilizó un kit de extracción de gel (Sigma-Aldrich, Zwijndrecht, Países Bajos). El fragmento de ADN de interés se escindió del gel de agarosa con una cuchilla. El ADN se extrajo siguiendo las recomendaciones del fabricante.

Ligadura

Los extremos pegajosos o romos de los fragmentos de ADN se unieron utilizando T4 ADN ligasa (Life Technologies Europe BV, Bleiswijk, Países Bajos). La mezcla de ligación se preparó de acuerdo con el protocolo estándar de los fabricantes y se incubó a 16°C durante 20 horas. Los plásmidos generados después de la ligadura estaban listos para transformarse en *E. coli*.

Transformación de *E. coli*

Se añadió un pequeño volumen de plásmidos (5 μ l) a 50 μ l de células químicamente competentes One Shot TOP10 (Life Technologies) y se siguió el protocolo para la transformación química de *E. coli* del proveedor. Las células transformadas se incubaron a 37°C durante una hora con agitación, lo que permitió que las células se recuperaran y aumentaran la resistencia al antibiótico usado. La suspensión celular se colocó en placas selectivas precalentadas y se incubó durante la noche a 37°C. Las células con plásmidos resultantes de la clonación de Blunt End TOPO no direccional (Life Technologies) se colocaron en placas de agar LB (10 g de triptona, 5 g de levadura), extracto, 10 g de NaCl, 15 g/l de agar) que contiene 50 μ g/ml de kanamicina (125 μ l) o 100 μ g/ml de ampicilina (125 μ l). Las células con plásmidos, que resultaron del corte y la ligadura en pUC19, se rayaron en LB con 100 μ g/ml de ampicilina (250 μ l). Después de la incubación durante la noche, las colonias individuales se volvieron a rayar en nuevas placas selectivas precalentadas para crecer nuevamente durante la noche. Después de volver a rayar, la selección de colonias únicas de *E. coli* fue posible sin fondo. Se inocularon colonias individuales en 5 ml de medio LB con el antibiótico apropiado para el crecimiento durante la noche. Antes del aislamiento de los plásmidos, se agregaron 200 μ l de solución de glicerol a 800 μ l del cultivo, después de lo cual la muestra se almacenó en un congelador a -80°C.

Aislamiento de plásmidos

Para aislar los plásmidos, se hizo crecer un cultivo de *E. coli* con el plásmido requerido durante la noche en medio LB. Después de sedimentar una muestra de 1-5 ml, se siguió el protocolo GenElute Plasmid Miniprep (Sigma-Aldrich). En la última etapa, se utilizaron 50 μ l en lugar de 100 μ l de agua desmineralizada sin nucleasas para eluir el plásmido purificado de la columna.

Transformación de Levadura

Para la transformación de levadura, la densidad óptica (OD) se determinó por la absorbancia medida con un espectrofotómetro a una longitud de onda de 660 nm. Con un OD₆₆₀ entre 0.6-0.8 (2×10^7 células/ml para 10 transformaciones) o corregido por la cantidad de células requeridas para una transformación individual, la transformación se realizó de acuerdo con el método de ADN-polietilenglicol del vehículo de acetato de litio de cadena sencilla [Gietz y Woods, (2002). *Methods in Enzymology* 350: 87-96]. Las células de levadura se rayaron en las correspondientes placas selectivas.

Construcción de plásmidos

La parte central del casete completo de eliminación de genes, que se llamará amdS-ISceI, consistió en el gen amdS optimizado con codón y el I-SCEI bajo el promotor de galactosa 1 (GAL1p) flanqueado por dos sitios de reconocimiento SceI. El casete de eliminación amdS-ISceI se separó en dos fragmentos, que contenían una superposición de 400 pares

de bases dentro del gen *amdS* para hacer una etapa de recombinación homóloga una parte esencial en el ensamblaje in vivo del casete. Las secuencias de los cebadores se proporcionan en la Tabla II.

El primer fragmento, denominado vector 1, se construyó con el cebador inverso Rv V1 *Scel* Sad, que se une al inicio del casete *amdSYM* en el promotor *AgTEF2* y que incluye un sitio de reconocimiento *Scel* y el sitio de restricción Sad, y con el cebador directo Fw V1 *NdeI*, que se unió al final de las 400 bases previstas se superpuso dentro de *amdS* e incluyó un sitio de restricción *NdeI* en su extremo 5' (Figura 1). *pUGamdS* se usó como plantilla para la construcción del vector 1.

Para la construcción del plásmido, el vector de fragmento de PCR 1 descrito anteriormente se clonó en *pCR4Blunt-TOPO* produciendo el plásmido *pUD266*.

La última parte de *amdS* se amplificó a partir de *pUGamdS* [Solis-Escalante, D. et al. (2013). *FEMS Yeast Research*, 13: 126-39] utilizando el cebador directo Fw V2 *NdeI*, que se unió al principio del pareado de 400 pares de bases previsto e incluyó un sitio de restricción *NdeI* en el producto de la PCR, y el cebador inverso Rv V2-overlap-*ISceI*, para crear una superposición de 60 pares de bases con el promotor *GAL1* que controlaba la expresión de *I-SCEI*. Este fragmento fue nombrado vector 2A. El casete de expresión *GAL1p-I-SCEI* se amplificó a partir de *pUDC073* (Kuijpers et al., 2013). *FEMS Yeast Research*. 13: 769-81) con el cebador directo Fw V2 *ISceI* y el cebador inverso Rv V2 *Scel EcoRI*, que incluía un sitio *Scel* en el otro lado del casete y un sitio *EcoRI*. Este fragmento fue nombrado vector 2B. Los productos de PCR, vector 2A y 2B se fusionaron mediante PCR con Fw V2 *NdeI* y Rv V2 *Scel EcoRI*. La PCR se realizó mezclando cantidades molares iguales de ambos fragmentos dando como resultado un fragmento de fusión, denominado vector 2 (Figura 1). El fragmento de PCR vector 2 se clonó en *pCR4Blunt-TOPO* produciendo *pUD267*. Los plásmidos se verificaron mediante análisis de restricción para verificar la orientación del fragmento clonado.

Casete de eliminación de genes

Para la construcción de los fragmentos de eliminación del gen final, 500 secuencias de pares de bases de los extremos 5' y 3' de *ScARO10*, un gen que codifica una fenil piruvato descarboxilasa involucrada en la vía de Ehrlich [Vuralhan. et al., (2005). *Appl Environ Microbiol* 71 3276-3284] se amplificaron a partir de ADN genómico de CBS 1483.

Tabla V Plásmidos de construcción de vectores estándar

Plásmido	Fragmento	Plantilla	Cebador directo	Cebador inverso
<i>pUD266</i> (Topo Blunt + V1)	<i>Scel-pamdS</i>	<i>pUGamdS</i>	Fw V1 <i>NdeI</i>	Rv V1 <i>Scel SacI</i>
<i>pUD267</i> (Topo Blunt + V2)	<i>amdSt-ISceI-Scel</i>	<i>pUD073</i>	Fw V2 <i>NdeI</i>	Rv V2 <i>Scel EcoRI</i>
<i>pUD268</i>	<i>pUD266</i> + 5' <i>ScARO10</i>	CBS1483	Fw 5' <i>ScARO10 NotI</i>	Rv 5' <i>ScARO10 SacI</i>
<i>pUD269</i>	<i>pUD267</i> + 3' <i>ScARO10</i>	CBS1483	Fw 3' <i>ScARO10 BamHI</i> 80bp	Rv 3' <i>ScARO10 PstI</i>

En los plásmidos *pUD268* y *pUD269*, los dos fragmentos completos para la eliminación de genes se ubicaron entre los cebadores directo e inverso M13, que se usaron para amplificar el casete de eliminación *amdS-ISceI*.

En la transformación de las diferentes cepas, los equivalentes molares para dos fragmentos se aplicaron en la mezcla de transformación de levadura con la cantidad total de 837 ng de ADN: 337 ng para el vector 1 con 5' *ScARO10* y 500 ng para el vector 2 con 3' *ScARO10*. Para *S. pastorianus* CBS1483, la transformación se repitió con 3.37 µg y 5 µg para ambas piezas, respectivamente.

Resultados

Ensamble de recombinación in vivo de *pUDC 114*

Aunque hay muchos datos que se pueden encontrar sobre las eficiencias de transformación en CEN.PK113-7D, no hay datos comparativos para la recombinación homóloga con 60 pares de bases disponibles para las cepas utilizadas en esta investigación. Debido a que la poliploidía y los reordenamientos cromosómicos pueden afectar la capacidad de las células para realizar una recombinación homóloga, tres especies de *S. cerevisiae* con diferentes ploidías se transformaron con los fragmentos superpuestos enumerados en la Tabla II para generar el plásmido *pUDC 114*: el haploide CEN.PK113-7D que genera IMC067, el diploide CEN.PK122 que genera IMC076 y el poliploide PRG410 que genera IMC077. Además, las cepas poliploides de *S. pastorianus* CMBS33 y CBS1483 y el diploide recientemente descubierto *S. eubayanus* CBS12357 [Libkind et al., (2011). *PNAS USA* 108: 14539-14544] se estudiaron para determinar la capacidad de producir

el plásmido pUDC 114, que genera IMC064 para CBS 12357 e IMC066 para CMBS33. La expectativa para CBS 1483 era que el ensamblaje no se produjera debido a la corta longitud de las secuencias de homología.

5 Las cepas IMC067, IMC076, IMC077, IMC066 e IMC064 fueron capaces de recombinar el plásmido pUDC 114 para producir el gen *amdS* y crecer en acetamida como única fuente de nitrógeno. Como se esperaba, el CBS 1483 no pudo recombinar las secuencias superpuestas de 60 pares de bases que forman el plásmido pUDC114. Se calcularon los transformantes por μg de ADN. Los resultados se muestran en la tabla VI.

10 TABLA VI Rendimientos de transformación promedio medidos en transformantes por μg de ADN con error estándar de la media (SEM). La transformación de PRG410 para generar IMC077 se realizó varias veces, pero solo una placa contenía colonias de color naranja positivas*

Cepa	Transformantes por mg de ADN	Error estándar de la media
IMC067 (CEN.PK113-7D)	20563	1515
IMC076 (CEN.PK122)	5952	433
IMC077 (PRG410)*	19	-
IMC064 (CBS 12357)	364	53
IMC066 (CMBS33)	278	85
CBS1483	0	-

15 Se formaron más de 2×10^4 transformantes por μg de ADN en el haploide IMC067, aproximadamente 6×10^3 transformantes por μg de ADN se produjeron en el diploide IMC076 y en el IMC077 solo se crearon 19 transformantes por μg de ADN. Así, al aumentar la ploidía de las diferentes cepas de *S. cerevisiae*, la cantidad de transformantes por μg de ADN disminuye significativamente de acuerdo con los datos presentados en la Tabla VI. En el tipo silvestre diploide *S. eubayanus* CBS 12357, la cantidad de transformantes por μg de ADN también fue 10 veces más baja que en la cepa diploide de laboratorio CEN.PK122 (Tabla VI). La cepa de cerveza lager *S. pastorianus* CMBS33 no fue muy eficiente en comparación con la cepa de laboratorio diploide de *S. cerevisiae* y solo fue ligeramente menos eficiente que la otra cepa parental *S. eubayanus* CBS 12357.

Construcción de los vectores estándar

25 Para generar el marcador de selección bipartito, el marcador *amdS* se dividió en dos partes que compartían una superposición de 400 pares de bases dentro del ORF de *amdS*. La primera parte se obtuvo amplificando el promotor TEF2 de *Ashbya gossypii* y los primeros 1138 nucleótidos del marco de lectura abierto de *amdS* de pUGamdSY. El fragmento resultante de 1569 pares de bases albergaba en su 5' un sitio de restricción de endonucleasa *SceI*. La clonación en el plásmido pCR4Blunt-TOPO de este fragmento dio como resultado pUD266. La segunda parte del sistema de marcador bipartito se construyó en dos pasos: 1) la segunda parte del marcador incluyó los últimos 908 pares de bases del marcador seleccionable *amdS* y el terminador *AgTEF2*. Este casete estaba flanqueado en su extremo 3' por una extensión de 60 pb complementaria al casete *SCEI*. 2) El casete *SCEI* de endonucleasa que llevaba el gen *SCEI* bajo el control del promotor *GAL1* se amplificó desde pUDC073. En su extremo 3' este fragmento albergaba un sitio de restricción *SceI* de endonucleasa adicional. Posteriormente, los dos fragmentos se conectaron entre sí mediante PCR de fusión y el fragmento fusionado resultante se clonó en el vector pCR4Blunt-TOPO para producir pUDC267.

40 Para controlar la dirección de integración, se realizó un análisis de restricción en pUDC266 con doble digestión con *NotI* y *BamHI*, cortando *NotI* dentro del plásmido pCR4Blunt-TOPO y cortando *BamHI* solo dentro del vector 1, lo que dio como resultado un patrón de bandas característico de 438 y 5087 pares de bases. Para la dirección del control de integración en pUD267, el análisis de restricción se realizó con digestión con *NotI* y *HindIII*, lo que dio como resultado un patrón de bandas característico de 815 y 5732 pares de bases.

El fragmento bipartito contenido en pUDC266 se secuenció y no reveló ninguna mutación.

45 Los análisis de restricción y los análisis de secuencia confirmaron la corrección de los plásmidos y no ocurrió ninguna mutación en la secuencia diseñada. Sin embargo, el análisis de pUDC267 y la secuencia de consenso adquirida mostró varios polimorfismos de nucleótido único (SNP). Un SNP en realidad dio un cambio en el primer nucleótido de un codón de *amdS*, cambiando el aminoácido del triptófano en arginina. Aunque esto significó un cambio de un aminoácido hidrófobo a un aminoácido cargado positivamente (por lo tanto, hidrófilo), no se descubrió ningún cambio en la efectividad del crecimiento de la acetamidasa. Los otros SNP e incluso la adición de 7 nucleótidos estaban fuera de las secuencias de *amdS* e *I-SceI*.

Adquirir el casete de eliminación

Aunque la parte central del casete de eliminación génica amdS-ISceI se ligó en los vectores estándar pUD266 y pUD267, las secuencias de recombinación homólogas no se integraron en el casete de eliminación. Para producir los fragmentos de recombinación homóloga, el ADN genómico de CBS 1483 se amplificó con Fw 5'ScARO10 NotI y Rv 5'ScARO10 Sad para la parte en la dirección ascendente del ScARO10 y con Fw 3'ScARO10 BamHI 80 pb y Rv 3'ScARO10 PstI para la parte en la dirección descendente. Los pUDC266 y pUD267 se digirieron con NotI/SacI y PstI/PmeI y se ligaron con el fragmento 5' (NotI/SacI) y el fragmento 3' (PstI preparado), respectivamente. Estas ligaduras generaron dos nuevos vectores pUD268 que llevan el primer elemento de casete seleccionable que se orienta al lado 5' del locus ScARO10 y pUD269 que lleva el segundo elemento de casete seleccionable que se orienta al lado 3' del locus ScARO10.

Eliminación del alelo ScARO10.

Para la prueba del principio de la eliminación de genes y la recuperación de marcadores usando este nuevo método de un casete amdS-ISceI con 500 pares de bases de flancos homólogos, la primera etapa es eliminar con éxito genes en varias cepas. La eliminación del gen objetivo con los dos casetes generados previamente consistió en tres pasos de recombinación esenciales basados en la homología (Figura 3A).

Se requirió que las partes de amdS no funcionales ubicadas dentro de ambos vectores estándar se recombinaran y formaran un gen amdS seleccionable funcional para el crecimiento en acetamida. Las otras dos etapas fueron la recombinación entre los fragmentos 5'ScARO10 y 3'ScARO10 en el casete y esos sitios en el ADN genómico.

Para evaluar si este nuevo método para la eliminación de genes y la recuperación del marcador podría aplicarse a *S. pastorianus*, se probaron varias cepas. Como control, se transformaron tres cepas de *Saccharomyces cerevisiae*, la cepa MATa haploide de laboratorio CEN.PK113-7D. El laboratorio diploide MATa/MATa cepa CEN.PK122 y la cerveza industrial cepa PRG410. A lo largo de estas manchas de levadura también se transformaron dos cepas de *Saccharomyces pastorianus* CBS 1483 y CMBS33.

La eliminación del alelo ScARO10 en cepas de la familia CEN.PK utilizando el enfoque bipartito produjo más de 200 transformantes por cada μg de ADN. En contraste, la transformación de la cepa industrial PRG410 produjo un número mucho menor de transformantes (0.4 transformantes por cada μg de ADN).

Si bien la eliminación con flancos cortos, el casete de eliminación generado por PCR no tuvo éxito en *S. pastorianus* CBS1483, la aplicación del enfoque bipartito permitió la identificación exitosa de mutantes eliminados correctamente. Sin embargo, el número de transformantes se mantuvo bajo (0.5 transformante por cada μg de ADN) (Tabla VII). Una transformación representativa de CBS1483 con 8.37 μg del marcador amdS bipartito dirigido al locus ScARO10 condujo al crecimiento de 2 a 9 transformantes en una placa de medio de acetamida 30 mM.

A partir de una transformación que resultó en nueve transformantes CBS1483, dos colonias positivas se volvieron a rayar para obtener colonias únicas verdaderas y las otras 7 colonias se verificaron mediante PCR de colonias de alta fidelidad con los cebadores internos Fw 5'ScARO10/amdS check y KanB r, generando un fragmento de 2851 pares de bases. Todos ellos fueron positivos para la recombinación homóloga entre la parte 5' ScARO10 y los 400 pares de bases se superponen dentro de amdS, lo que sugiere que los casetes de eliminación se integraron con éxito. Se obtuvo un solo aislado de colonia y se le cambió el nombre a IMK490 (CBS1483 con un locus eliminado Scaro10 Δ :: amdS).

De manera similar, el aislado de colonia individual se obtuvo de CEN.PK113-7D, CEN.PK122, PRG410, CMBS33 y cambió su nombre por IMK386, IMK487, IMK488 e IMK489 respectivamente.

Tabla VII Eficiencia promedio de transformación medida en transformantes por μg de ADN con error estándar de la media (SEM)

Cepa	Anfitrión	Transformantes por mg de ADN	SEM
IMK486	CEN.PK113-7D	215.9	27.4
IMK487	CEN.PK122	212.7	25.9
IMK488	PRG410	0.4	0.2
IMK489	CMBS33	27.9	7.2
IMK490	CBS1483	0.5	0.4

Para estar absolutamente seguros de que el casete de modificación se integró con éxito en el ADN genómico dirigido, la recombinación homóloga de la superposición dentro de amdS también se controló con los cebadores amdS KanA f y KanB r (2164 pb) y la recombinación homóloga del ScARO10 3' con cebadores internos y externos FK072 y Rv

3'ScARO10/amdS check (1081 pb). Debido a que cada cepa contiene más de una copia de ScARO10, excepto para CEN.PK113-7D, la presencia de otras copias de ScARO10 se controló con los cebadores ScARO10-Fw en el interior y ScARO10-Rv en el interior (959 pb).

5 Los resultados mostraron que todas las cepas utilizadas en este proyecto fueron capaces de eliminar al menos una copia de ScARO10. En CEN.PK113-7D scaro10Δ:: amdS, las otras cepas, incluidas las dos colonias verificadas para IMK490, aún tenían al menos una copia del ScARO10 en su genoma. Todas las colonias IMK486-IMK490 se almacenaron y una muestra de cada una se volvió a inocular en YPD para crecer para la segunda etapa, la eliminación del marcador.

10 Eliminación de marcador sin interrupciones

Para verificar que la construcción bipartita integrada podría recuperarse fácilmente del locus objetivo utilizando la endonucleasa inducible SCEI, la cepa IMK490 (CBS1483 con un locus eliminado Scaro10Δ:: amdS) se hizo crecer en un medio de galactosa para inducir el promotor GAL1 que controla la expresión de SCEI. (Figura 3B). La cepa IMK490 se hizo crecer en medio sintético con galactosa durante 48 horas. Tras la inducción, la endonucleasa crea un corte en los sitios Scel que flanquean el casete de eliminación y, mientras tanto, elimina su región de codificación del cromosoma, lo que permite un reciclaje de la construcción de edición del genoma.

20 Las cepas IMK386, IMK487, IMK488 e IMK489 se trataron de manera similar.

La selección de las cepas con marcador seleccionado por duplicado se realizó mediante dos métodos diferentes:

1 - Plato de galactosa

25 Las células IMK490 cultivadas con galactosa se rayaron en placas que contenían galactosa como fuente de carbono y se incubaron a 30°C durante 3 días. Los aislados de colonias individuales se resuspendieron en 100 µl de agua estéril y 5 µl se transfirieron a placas de medios sintéticos que contenían amonio o acetamida como única fuente de nitrógeno. Estas placas se cultivaron durante 2 días a 30°C. De manera similar, se aislaron colonias aisladas individuales de IMK386, IMK487, IMK488 e IMK489 en placas de medios sintéticos que contenían amonio o acetamida como única fuente de nitrógeno.

30 Para IMK486, de las 5 colonias seleccionadas ninguna había reciclado el marcador amdS, para IMK487 solo 1 de los 5 clones había reciclado el marcador amdS, y para IMK489 los cinco clones seleccionados habían perdido el marcador amdS. Para la cepa IMK490 se comprobaron 11 colonias y 10 habían extirpado el marcador seleccionable de amdS. Para determinar si el marcador se eliminó del cromosoma, se realizó una PCR de colonias en el ADN genómico de un clon representativo con los cebadores externos-externos Fw 5'ScARO10/amdS check y Rv 3'ScARO10/amdS check, que produjo un fragmento de 1426 pares de bases. Esto confirmó que el casete amdS-IScel se había eliminado del genoma.

40 2-Inoculación en medio líquido y contra selección en fluoroacetamida.

El segundo método utilizado para eliminar el marcador consistió en inocular 1 ml de un cultivo de cada cepa que contiene el casete amdS-IScel en medio líquido con 2% de galactosa como fuente principal de carbono y 0.05% de glucosa para mejorar el crecimiento de las diferentes cepas de levadura. Después de crecer durante 4 horas en este medio líquido, se tomaron muestras y se diluyeron 200 veces en agua estéril y se colocaron 100 µl en medio sintético con placas de galactosa y fluoroacetamida. Las colonias que expresan el gen amdS por lo tanto hidrolizarían la fluoroacetamida en amonio y fluoroacetato, que es tóxico. Solo las células que hayan perdido el marcador seleccionable de amdS crecerían.

50 En este proyecto, se investigó la recombinación homóloga mediante el ensamblaje in vivo del plásmido pUDC114 a partir de cuatro fragmentos de ADN de doble cadena que tenían superposiciones de 60 pares de bases. pUDC114 contenía los genes de caroteno crtYBEI, que dieron a las colonias positivas un color naranja como medio para una rápida identificación. Los resultados de la transformación mostraron que al aumentar la ploidía de las diferentes cepas de *S. cerevisiae*, la cantidad de transformantes por µg de ADN disminuyó significativamente, de 2x10⁴ transformantes por µg de ADN en el haploide CEN.PK113-7D, a 6x10³ transformantes por µg de ADN en el diploide CEN.PK122 y finalmente a 19 transformantes por µg de ADN en la cepa de cerveza poliploide PRG410. En el recientemente descubierto diploide de tipo silvestre *S. eubayanus* CBS 12357, la cantidad de transformantes por µg de ADN también fue 10 veces más baja que en la cepa diploide de laboratorio CEN.PK122.

60 La eficacia de la recombinación homóloga en las complejas cepas de elaboración de cerveza lager no se investigó antes. El *S. pastorianus* CMBS33 fue 20 veces menos eficiente que el diploide CEN.PK122, pero solo un poco menos eficiente que la otra cepa parental *S. eubayanus* CBS12357. La eficiencia de transformación y la recombinación homóloga de los fragmentos pUDC114 en la cepa CMBS33 de cerveza lager fue de ± 300 transformantes por µg de ADN, pero cero para CBS1483. Esto prueba que no todas las cepas de elaboración de cerveza lager son idénticas, con diferencias genéticas determinadas por las particularidades del proceso de elaboración de la cerveza de las que se seleccionaron.

65 El sistema de eliminación de genes con la eliminación del marcador amdS por la endonucleasa I-Scel es una forma novedosa de alterar y eliminar genes en cepas de cerveza lager y en cepas de laboratorio de *S. cerevisiae*. La posibilidad

de alterar y/o eliminar genes y la posterior eliminación de marcadores en *S. pastorianus* CBS1483 con el casete amdS-IScel contribuye sustancialmente a la caja de herramientas de los investigadores en la industria cervecera.

Ejemplo 2

5

El ADN genómico de la cepa ATC 8585 de *Kluyveromyces lactis* se preparó como se describe (Burke et al., 2000. Cold Spring Harbor Laboratory. Methods in yeast genetics: a Cold Spring Harbor Laboratory course manual). ORF KLLA0F27995g, que codifica KIGBU1, se amplificó a partir del ADN genómico usando la polimerasa Phusion Hot-Start (Finnzymes) y los cebadores cebador directo GBU1 (5'-CATCCGAACATAAACAACCATGAA GGTTCAGGATTTATATTG) y cebador inverso GBU1 (5'-CAAGAAT CTTTTATTGTCAGTACTGATCAGGCTTGCAAAACAAATTGTTC). Se obtuvo la secuencia codificante del gen GBU1 de *K. lactis*.

10

15

En la Figura 4 se proporciona un conjunto de construcciones de orientación que comprenden el marcador de selección KIGBU1 con todas las partes esenciales para el casete de eliminación estándar. El solapamiento de 400 bases en el marcador de selección KIGBU1 (indicado por una cruz) está diseñado para recombinarse debido a la homología.

REIVINDICACIONES

- 5 1. Un grupo de construcciones de direccionamiento, que comprende una primera construcción que comprende, en este orden, una primera región de homología con un genoma objetivo de un microorganismo, un sitio de reconocimiento para una endonucleasa y una primera parte de un marcador de selección y una segunda construcción que comprende, en este orden, una segunda parte del marcador de selección, una copia del sitio de reconocimiento de endonucleasas y una segunda región de homología con el genoma objetivo del microorganismo, por lo que la primera y la segunda regiones de homología con el genoma objetivo cada una comprende al menos 20 pares de bases (pb);
- 10 un fragmento de la primera parte del marcador de selección se superpone con un fragmento que está presente en la segunda parte del marcador de selección, permitiendo la recombinación entre la primera y la segunda parte del marcador de selección, comprendiendo dicho fragmento preferiblemente entre 50 y 600 pb;
- 15 una secuencia codificante que codifica la endonucleasa y que está acoplada a un promotor inducible está presente en la primera o segunda construcción; y
- una parte de la primera región de homología con el genoma objetivo en la primera construcción se duplica entre la copia del sitio de reconocimiento de endonucleasas y la segunda región de homología con el genoma objetivo en la segunda construcción; o una parte de la segunda región de homología con el genoma objetivo en la segunda construcción se duplica entre la primera región de homología con el genoma objetivo y el sitio de reconocimiento de endonucleasa en la primera construcción, comprendiendo dicha región duplicada entre 20 y 200 pb.
- 20 2. El grupo de construcciones de direccionamiento de acuerdo con la reivindicación 1, en donde el fragmento superpuesto del marcador de selección es de aproximadamente 200 pares de bases (pb).
- 25 3. El grupo de construcciones de direccionamiento de acuerdo con la reivindicación 1 o la reivindicación 2, en el que la región duplicada de homología con el genoma objetivo en la primera y segunda construcciones de direccionamiento es preferiblemente de aproximadamente 80 pb.
- 30 4. El grupo de construcciones de direccionamiento de acuerdo con una cualquiera de las reivindicaciones anteriores, por lo que la endonucleasa es una endonucleasa de direccionamiento específico.
- 35 5. El grupo de construcciones de direccionamiento de acuerdo con una cualquiera de las reivindicaciones anteriores, por lo que el marcador seleccionable es un marcador auxotrófico y/o un marcador dominante.
- 40 6. El grupo de construcciones de direccionamiento de acuerdo con una cualquiera de las reivindicaciones anteriores, por lo que el promotor inducible se selecciona entre el promotor GAL1, el promotor GAL10, el promotor SUC2, el promotor MAL12, CUP1 y un promotor regulable por tetraciclina.
- 45 7. El grupo de construcciones de direccionamiento de acuerdo con una cualquiera de las reivindicaciones anteriores, en donde el microorganismo es un microorganismo aneuploide.
8. El grupo de construcciones de direccionamiento de acuerdo con una cualquiera de las reivindicaciones anteriores, en donde el microorganismo es una Ascomycota, preferiblemente una Saccharomycotina.
- 50 9. El grupo de construcciones de direccionamiento de acuerdo con una cualquiera de las reivindicaciones anteriores, en donde el microorganismo es *Saccharomyces pastorianus*.
10. Un método para alterar un genoma en un microorganismo, que comprende proporcionar el grupo de construcciones de direccionamiento de acuerdo con una cualquiera de las reivindicaciones 1-9 a dicho microorganismo, y seleccionar un microorganismo en el que el genoma ha sido alterado.
- 55 11. Un método para producir un microorganismo que comprende una alteración genómica, el método que comprende proporcionar el conjunto de construcciones de direccionamiento de acuerdo con una cualquiera de las reivindicaciones 1-9 a dicho microorganismo, y seleccionar un microorganismo en el que el genoma ha sido alterado y que se expresa funcionalmente. El marcador de selección recombinada.
- 60 12. El método de acuerdo con la reivindicación 11, que comprende además inducir el promotor inducible para la expresión de la endonucleasa.
- 65 13. Un microorganismo, que comprende una alteración genómica que se produce mediante el método de la reivindicación 11.
14. Un microorganismo, que comprende una alteración genómica, la alteración comprende la inserción de un marcador de selección funcional recombinante y una secuencia de codificación que codifica una endonucleasa y que está acoplada

a un promotor inducible, por lo que el sitio de inserción comprende una copia de una secuencia de reconocimiento para la endonucleasa en ambos sitios de la inserción.

- 5 15. Un método para producir un microorganismo comprende una región cromosómica alterada, el método comprende proporcionar el microorganismo de acuerdo con la reivindicación 14, e inducir al promotor inducible a eliminar las secuencias de ácido nucleico entre las secuencias de reconocimiento de la endonucleasa.

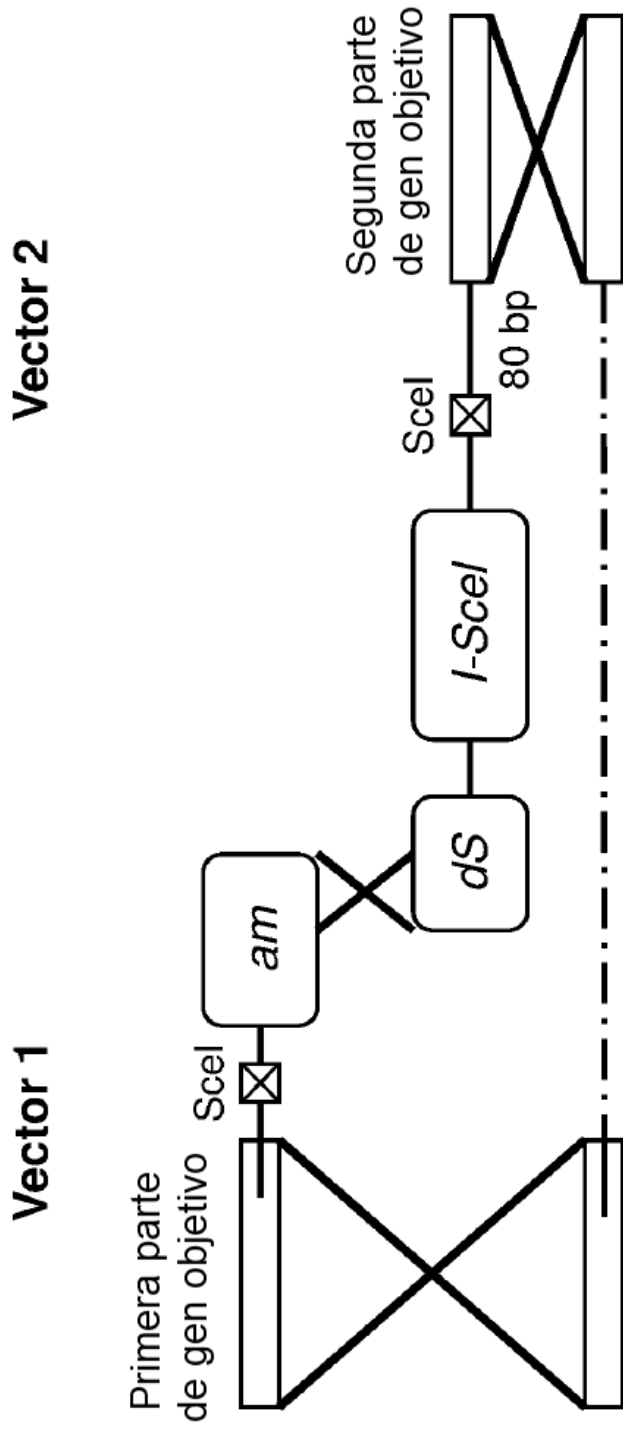


FIG. 1

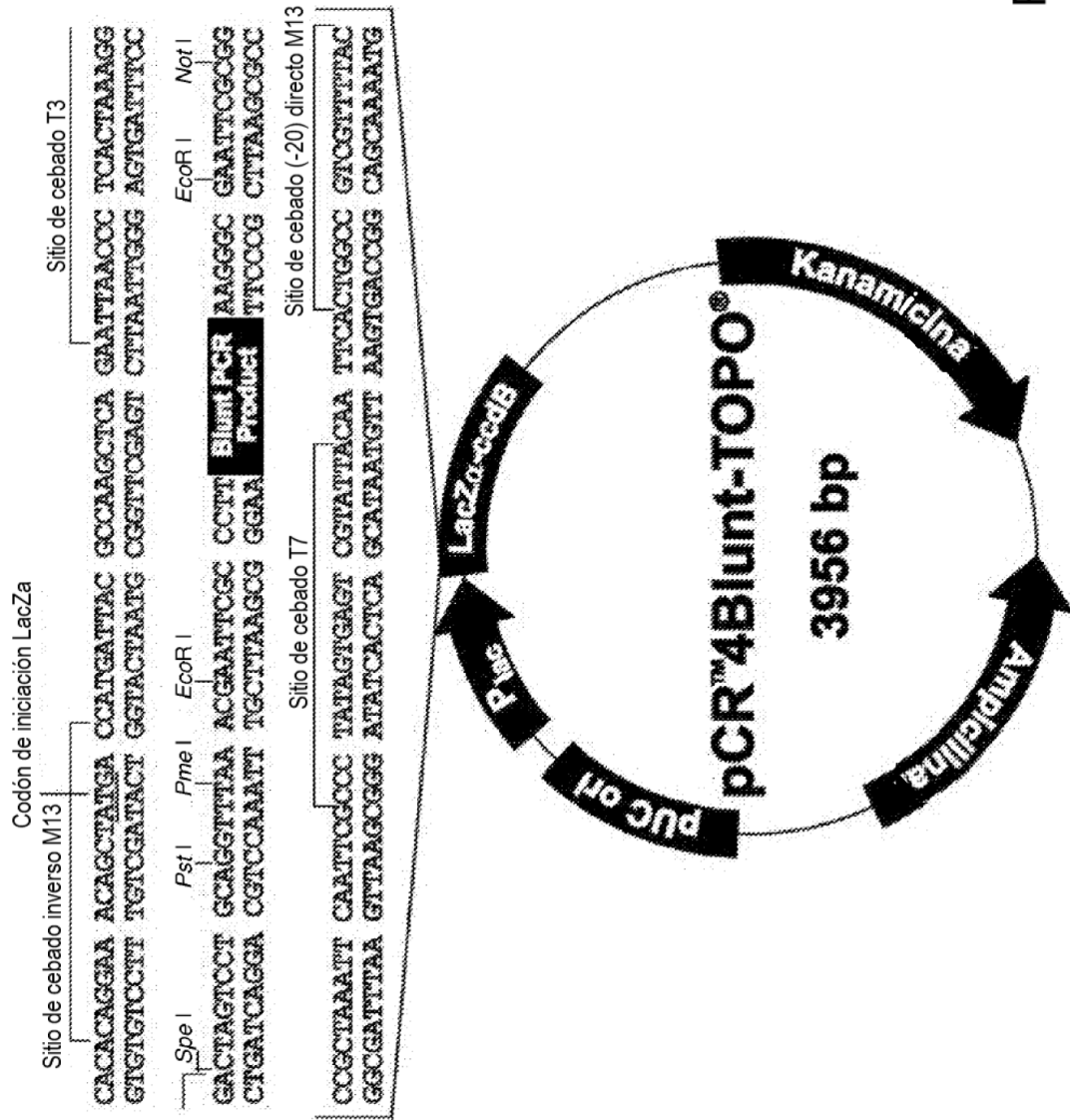


FIG. 2

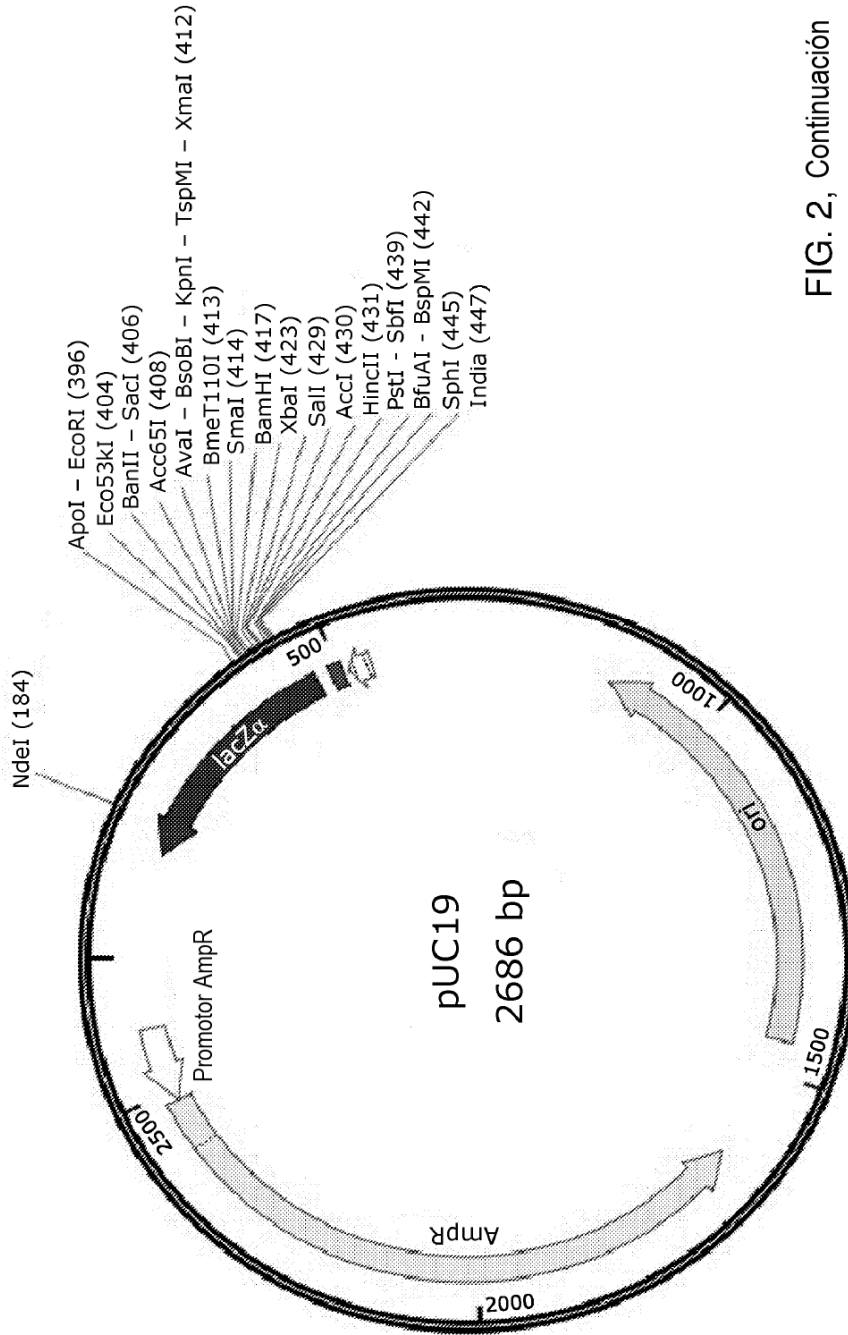


FIG. 2, Continuación

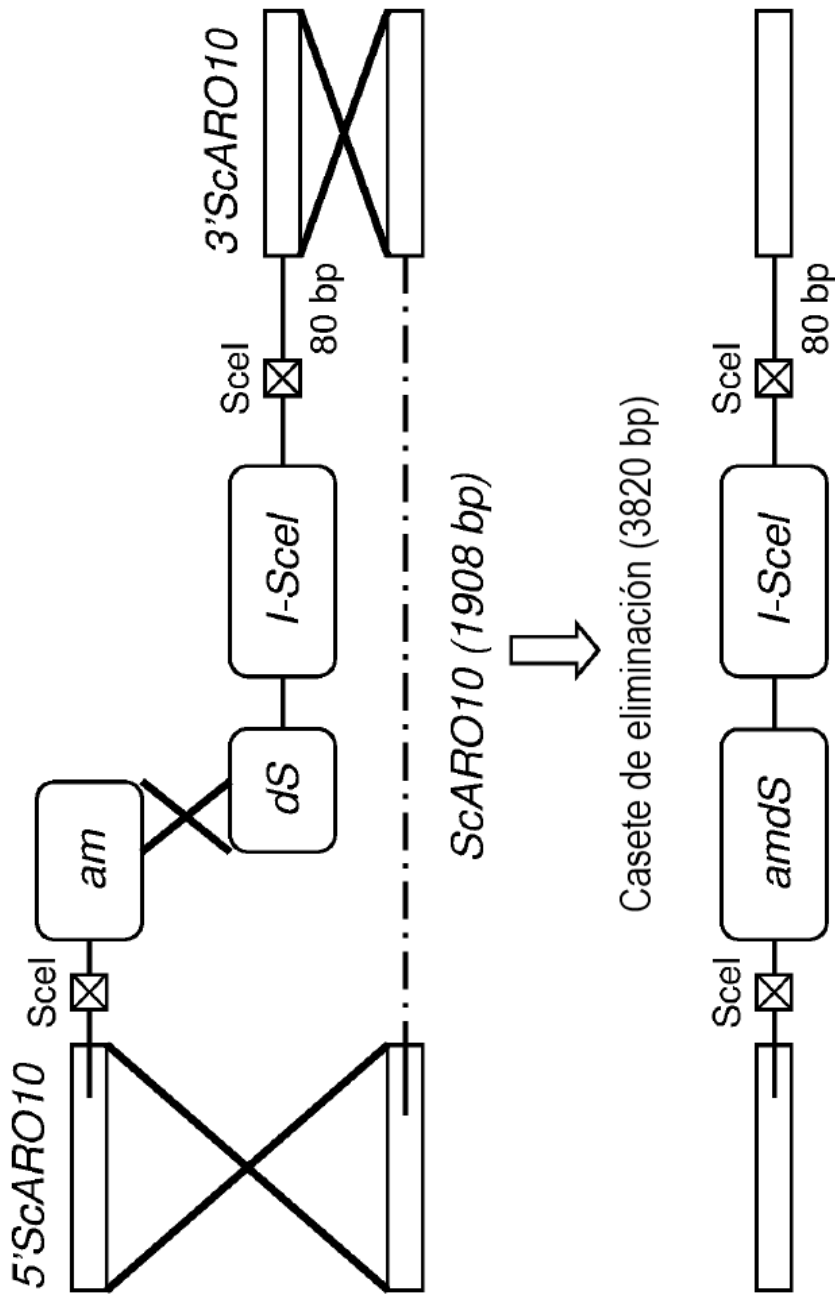


FIG. 3A

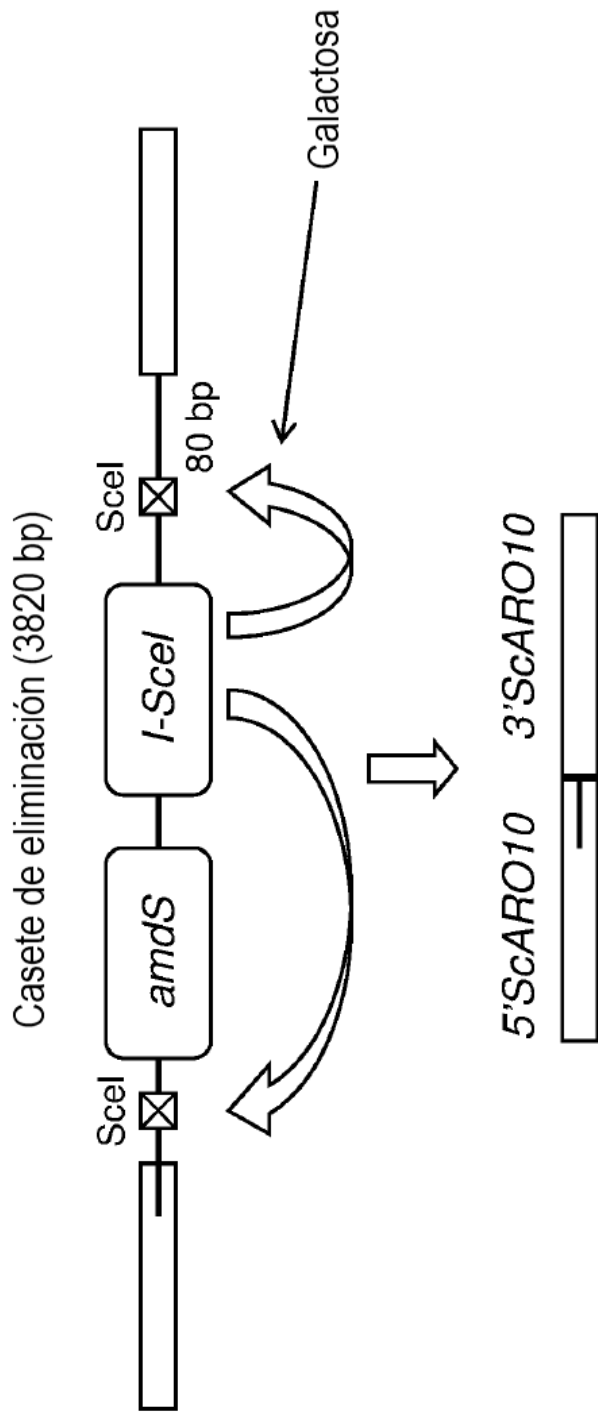


FIG. 3B

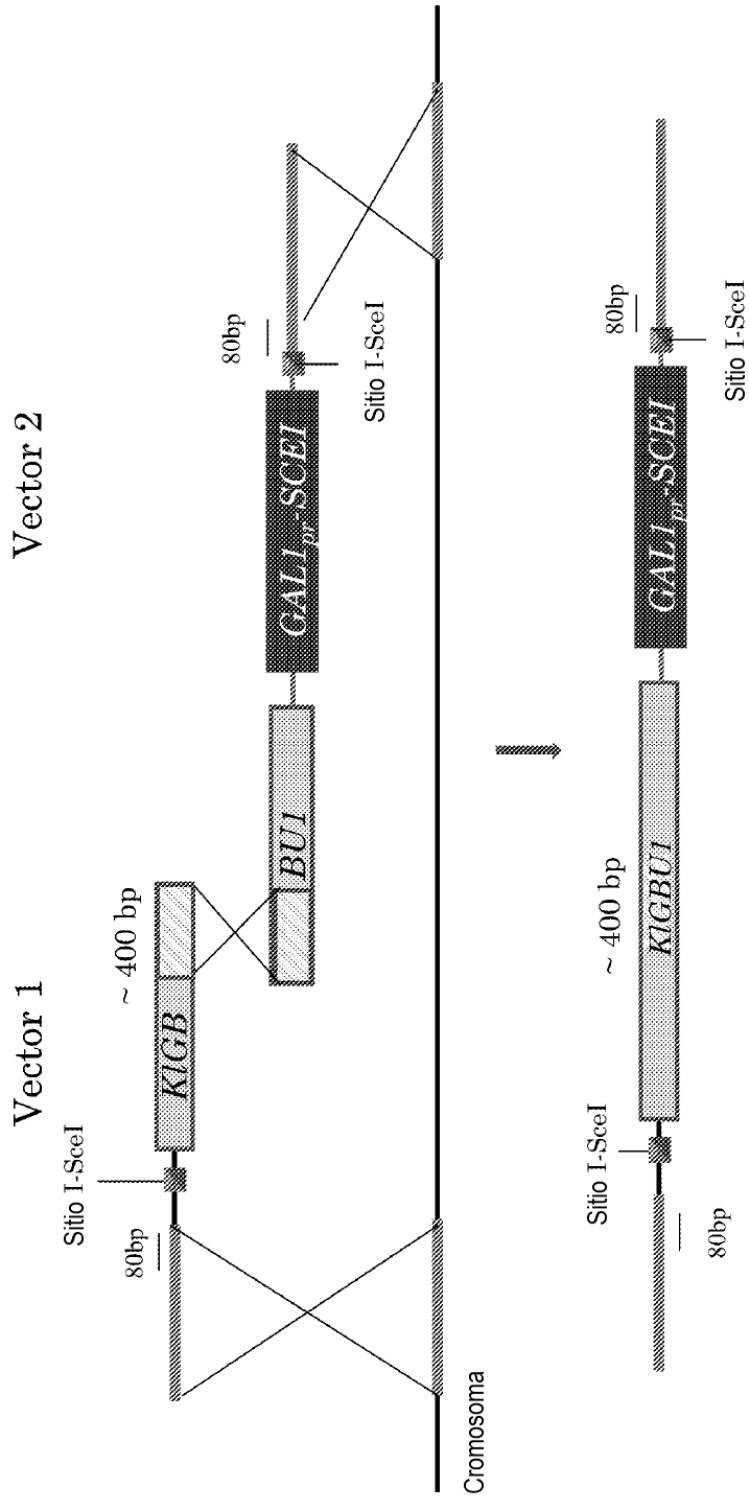


FIG. 4