

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 702 075**

51 Int. Cl.:

C12N 9/02 (2006.01)

C12P 23/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **27.08.2013 PCT/IB2013/056583**

87 Fecha y número de publicación internacional: **26.06.2014 WO14096990**

96 Fecha de presentación y número de la solicitud europea: **27.08.2013 E 13779632 (2)**

97 Fecha y número de publicación de la concesión europea: **19.09.2018 EP 2935570**

54 Título: **Caroteno hidroxilasa y su uso para producir carotenoides**

30 Prioridad:

20.12.2012 EP 12198352

03.06.2013 US 201361830273 P

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

27.02.2019

73 Titular/es:

DSM IP ASSETS B.V. (100.0%)

Het Overloon, 1

6411 TE Heerlen, NL

72 Inventor/es:

FARRELL, CHRISTOPHER;

MAYORGA, MARIA y

CHEVREUX, BASTIEN

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 702 075 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

Caroteno hidroxilasa y su uso para producir carotenoides

5 La presente invención se refiere a una nueva hidroxilasa, a secuencias de ácidos nucleicos que codifican, por lo tanto, construcciones de expresión y vectores que comprenden esta secuencia, a microorganismos transformados con ella, a procedimientos para la hidroxilación microbiológica de isoprenoides, por ejemplo procedimientos para convertir beta-caroteno en zeaxantina o beta-criptoxantina y cantaxantina en astaxantina.

Los carotenoides son pigmentos orgánicos que varían en color de amarillo a rojo que son producidos de forma natural por determinados organismos, incluidos los organismos fotosintéticos (p. ej., plantas, algas, cianobacterias) y algunos hongos.

10 Carotenoides, tales como luteína, zeaxantina o astaxantina, son aditivos importantes en la dieta humana y del ganado como sustancias pigmentantes y precursoras de derivados de la vitamina A. Además, los carotenoides tienen una acción que fomenta la salud, tal como potenciar la respuesta inmune y, debido a sus propiedades antioxidantes, una acción que previene el cáncer, lo que hace interesante su uso como nutracéuticos. Por lo tanto, es de gran importancia un procedimiento económico para preparar carotenoides y alimentos con un mayor contenido en carotenoides.

15 Procedimientos particularmente económicos para preparar carotenoides son procedimientos biotecnológicos que hacen uso de proteínas y genes de biosíntesis de biosíntesis de carotenoides a partir de organismos productores de carotenoides.

20 Las [beta]-caroteno hidroxilasas procarióticas que catalizan la conversión enzimática de [beta]-caroteno en zeaxantina a través de [beta]-criptoxantina, y los genes que codifican estas proteínas se conocen de la bacteria *Erwinia uredovora* (Misawa et al., *J. Bacteriol.* 1990, 6704-6712; documento EP 393690 B1), *Erwinia herbicola* (documento WO 9113078), *Agrobacterium aurantiacum* (Misawa et al., *J. Bacteriol.* 1995, 6575-6584; documento EP 735 137 A1), *Alcaligenes* sp. PC-1 (documento EP 735 137 A1), *Flavobacterium* sp. cepa R1534 (Pasamontes et al., *Gene.* 1997, 185: 35-41; documento EP 747483 A2) y de la cianobacteria *Synechocystis* sp. PCC6803 (Masamoto et al., *Plant Cell Physiol.* 1998, 39(5): 560-564).

También se sabe que las [beta]-caroteno hidroxilasas procarióticas de *Agrobacterium aurantiacum*, *Alcaligenes* y *Erwinia uredovora* son adicionalmente capaces de convertir cantaxantina a través de adonirrubina en astaxantina (Misawa et al., *J. Bacteriol.* 1995, 6575-6584; Fraser et al., *J. Biol. Chem.* 1997, 272: 6128-6135).

30 A partir de fuentes eucarióticas se sabe que tres [beta]-caroteno hidroxilasas de plantas catalizan la conversión enzimática de [beta]-caroteno en zeaxantina a través de [beta]-criptoxantina. Los ADNc correspondientes se han aislado de *Arabidopsis thaliana* (Cunningham et al., *J. Biol. Chem.* 1996, 271:24349-24352, documento WO 9736998) y de *Capsicum annuum* L. (Bouvier et al., *Biochim. Biophys. Acta.* 1998, 1391: 320-328).

35 Los genes de origen eucariótico tienen la ventaja sobre los genes procarióticos de que se expresan mejor en organismos transgénicos superiores, tales como las plantas. Sin embargo, todavía existe la necesidad de mejorar y aumentar la productividad de los carotenoides para un procedimiento económico para preparar derivados de carotenoides o productos alimenticios con un mayor contenido de carotenoides mediante la incorporación de ácidos nucleicos eucarióticos en los organismos.

40 Además, las [beta]-caroteno hidroxilasas eucarióticas apropiadas en la técnica anterior tienen la desventaja de que tienen solo un estrecho intervalo de sustratos, de modo que hay una acumulación de productos metabólicos que no pueden ser convertidos por las hidroxilasas y pueden ejercer un efecto inhibitorio sobre las hidroxilasas.

Sumario de la invención

Es un objeto de la presente invención remediar las deficiencias descritas de la técnica anterior y proporcionar una caroteno hidroxilasa eucariótica con propiedades mejoradas.

45 Los autores de la invención han encontrado que el objeto anterior se logra sorprendentemente mediante una proteína que tiene una actividad enzimática para la hidroxilación de isoprenoides, por ejemplo, [beta]-caroteno o cantaxantina, que comprende la secuencia de aminoácidos SEQ ID NO: 2.

Por lo tanto, la presente invención se refiere a una proteína o polipéptido, que comprende la secuencia de aminoácidos SEQ ID NO: 2.

En particular, la hidroxilasa de acuerdo con la invención, si se expresa en una cepa seleccionada o capaz de producir un carotenoide específico tal como, por ejemplo, zeaxantina o astaxantina, aumenta el nivel de estos carotenoides en comparación con cepas no transformadas con el gen que codifica la proteína o el polipéptido de acuerdo con la invención.

- 5 La presente invención también se refiere al polinucleótido aislado que codifica el polipéptido de la presente invención, construcciones de ácidos nucleicos, vectores de expresión recombinantes y células huésped recombinantes que comprenden el polinucleótido, y a métodos para producir el polipéptido.

La presente invención también proporciona sistemas mejorados para la producción biológica de carotenoides. En un ejemplo preferido, la invención proporciona hongos oleaginosos (que incluyen, por ejemplo, levadura) que producen uno o más carotenoides. La presente invención también proporciona métodos para construir levaduras y hongos de este tipo, métodos para usar levaduras y hongos de este tipo para producir carotenoides, y métodos para preparar composiciones que contienen carotenoides, tales como aditivos para alimentos o piensos, o complementos nutricionales, utilizando carotenoides producidos en levaduras u hongos oleaginosos de este tipo. En particular, la presente invención proporciona sistemas y métodos para generar polinucleótidos que contienen hongos y levaduras que codifican los polipéptidos de la presente invención.

10
15

Descripción general del Listado de Secuencias

SEQ ID NO: 1 es la secuencia de ADN no optimizada que codifica la caroteno hidroxilasa de baja frecuencia de *Haematococcus pluvialis*.

SEQ ID NO: 2 es la secuencia de aminoácidos que se deduce de SEQ ID NO:1.

- 20 SEQ ID NO: 3 es la secuencia de ADN que codifica la caroteno hidroxilasa de baja frecuencia de *Haematococcus pluvialis* optimizada para la expresión en *Yarrowia lipolytica*.

SEQ ID NO:4 es la secuencia de ADN que codifica la caroteno hidroxilasa de *Cronobacter pulveris* (anteriormente *Enterobacter pulveris*) (*Ep*) optimizada para la expresión en *Yarrowia lipolytica*.

- 25 SEQ ID NO:5 es la secuencia de ADN que codifica la caroteno hidroxilasa de la bacteria DC404 (*Dc*) de *Enterobacteriaceae*, optimizada para la expresión en *Yarrowia lipolytica*.

Definiciones

Polipéptido aislado: la expresión "polipéptido aislado" significa un polipéptido que está modificado por la mano del hombre en relación con ese polipéptido tal como se encuentra en la naturaleza. En un aspecto, el polipéptido es al menos 1% puro, *p. e.*, al menos 5% puro, al menos 10% puro, al menos 20% puro, al menos 40% puro, al menos 60% puro, al menos 80% puro y al menos 90% puro, según se determina por SDS-PAGE.

30

Polipéptido sustancialmente puro: la expresión "polipéptido sustancialmente puro" significa una preparación que contiene a lo sumo 10%, a lo sumo 8%, a lo sumo 6%, a lo sumo 5%, a lo sumo 4%, a lo sumo 3%, a lo sumo 2%, a lo sumo 1% y a lo sumo 0,5% en peso de otro material polipeptídico con el que está asociado de forma nativa o recombinante. Preferiblemente, el polipéptido es al menos 92% puro, *p. e.*, al menos 94% puro, al menos 95% puro, al menos 96% puro, al menos 97% puro, al menos 98% puro, al menos 99% puro, al menos 99,5% puro y 100% puro en peso del material polipeptídico total presente en la preparación. Los polipéptidos de la presente invención están preferiblemente en una forma sustancialmente pura. Esto se puede lograr, por ejemplo, preparando el polipéptido mediante métodos recombinantes bien conocidos o mediante métodos de purificación clásicos.

35

Identidad de secuencia: la relación entre dos secuencias de aminoácidos o entre dos secuencias de nucleótidos se describe mediante el parámetro "identidad de secuencia".

40

Para los fines de la presente invención, el grado de identidad de secuencia entre dos secuencias de aminoácidos se determina utilizando el algoritmo de Needleman-Wunsch (Needleman y Wunsch, *J. Mol. Biol.* 1970, 48: 443-453) tal como se implementa en el programa de Needle del paquete EMBOSS (EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., *Trends Genet.* 2000, 16: 276-277), preferiblemente la versión 3.0.0 o posterior. Los parámetros opcionales utilizados son la penalización por apertura de hueco de 10, la penalización por extensión de hueco de 0,5 y la matriz de sustitución EBLOSUM62 (versión EMBOSS de BLOSUM62). La salida de Needle marcada como "identidad más larga" (obtenida utilizando la opción -nobrief) se utiliza como el porcentaje de identidad y se calcula de la siguiente manera:

45

(Residuos Idénticos x 100)/(Longitud de Alineamiento - Número Total de Huecos en el Alineamiento)

Para los fines de la presente invención, el grado de identidad de secuencia entre dos secuencias de desoxirribonucleótidos se determina utilizando el algoritmo de Needleman-Wunsch (Needleman y Wunsch, 1970, *supra*) según se implementa en el programa de Needle del paquete EMBOSS (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000, *supra*), preferiblemente versión 3.0.0 o posterior. Los parámetros opcionales utilizados son la penalización por apertura de hueco de 10, la penalización por extensión de hueco de 0,5 y la matriz de sustitución EDNAFULL (versión EMBOSS de NCBI NUC4.4). La salida de Needle marcada como "identidad más larga" (obtenida utilizando la opción -nobrief) se utiliza como el porcentaje de identidad y se calcula de la siguiente manera:

10 (Desoxirribonucleótidos Idénticos x 100)/(Longitud de Alineamiento - Número Total de Huecos en el Alineamiento)

Fragmento: el término "fragmento" significa un polipéptido que tiene uno o más (varios) aminoácidos suprimidos del extremo amino y/o carboxilo de un polipéptido maduro; en donde el fragmento tiene actividad hidroxilasa.

15 **Variante alélica:** la expresión "variante alélica" significa cualquiera de dos o más formas alternativas de un gen que ocupa el mismo locus cromosómico. La variación alélica surge de forma natural a través de la mutación, y puede resultar en un polimorfismo dentro de las poblaciones. Las mutaciones genéticas pueden ser silenciosas (sin cambios en el polipéptido codificado) o pueden codificar polipéptidos que tienen secuencias de aminoácidos alteradas. Una variante alélica de un polipéptido es un polipéptido codificado por una variante alélica de un gen.

20 **Polinucleótido aislado:** la expresión "polinucleótido aislado" significa un polinucleótido que está modificado por la mano del hombre en relación con ese polinucleótido tal como se encuentra en la naturaleza. En un aspecto, el polinucleótido aislado es al menos 1% puro, *p. ej.*, al menos 5% puro, al menos 10% puro, al menos 20% puro, al menos 40% puro, al menos 60% puro, al menos 80% puro, al menos 90% puro y al menos 95% puro, según se determina por electroforesis en agarosa. Los polinucleótidos pueden ser de origen genómico, ADNc, ARN, semisintético, sintético, o cualquier combinación de los mismos.

25 **Polinucleótido sustancialmente puro:** la expresión "polinucleótido sustancialmente puro" significa una preparación de polinucleótido exenta de otros nucleótidos extraños o no deseados y en una forma adecuada para su uso dentro de los sistemas de producción de polipéptidos genéticamente modificados. Por lo tanto, un polinucleótido sustancialmente puro contiene como a lo sumo 10%, a lo sumo 8%, a lo sumo 6%, a lo sumo 5%, a lo sumo 4%, a lo sumo 3%, a lo sumo 2%, a lo sumo 1% y a lo sumo 0,5% en peso de otro material polinucleotídico con el que está asociado de forma nativa o recombinante. Un polinucleótido sustancialmente puro puede, sin embargo, incluir regiones 5' y 3' no traducidas que se producen de forma natural, tales como promotores y terminadores. Preferiblemente, el polinucleótido es al menos 90% puro, *p. ej.*, al menos 92% puro, al menos 94% puro, al menos 95% puro, al menos 96% puro, al menos 97% puro, al menos 98% puro, al menos 99% puro y al menos 99,5% puro en peso. Los polinucleótidos de la presente invención están preferiblemente en una forma sustancialmente pura.

35 **Secuencia codificante:** la expresión "secuencia codificante" significa un polinucleótido, que especifica directamente la secuencia de aminoácidos de un polipéptido. Los límites de la secuencia codificante están generalmente determinados por un marco de lectura abierto, que habitualmente comienza con el codón de inicio de ATG o codones de inicio alternativos como GTG y TTG y termina con un codón de parada, tal como TAA, TAG y TGA. La secuencia codificante puede ser un polinucleótido de ADN, ADNc, sintético o recombinante.

40 **ADNc:** el término "ADNc" significa una molécula de ADN que se puede preparar por transcripción inversa de una molécula de ARNm madura, cortada y empalmada, obtenida de una célula eucariótica. El ADNc carece de secuencias intrónicas que pueden estar presentes en el ADN genómico correspondiente. La transcripción de ARN primaria inicial es un precursor del ARNm que se procesa a través de una serie de etapas, incluido el corte y empalme, antes de aparecer como ARNm cortado y empalmado maduro.

45 **Construcción de ácido nucleico:** la expresión "construcción de ácido nucleico" significa una molécula de ácido nucleico, ya sea de cadena sencilla o doble, que se aísla de un gen que se produce de forma natural o se modifica para que contenga segmentos de ácidos nucleicos de una manera que de otro modo no existiría en la naturaleza o que sea sintético. La expresión construcción de ácido nucleico es sinónima de la expresión "casete de expresión" cuando la construcción de ácido nucleico contiene las secuencias de control requeridas para la expresión de una secuencia codificante de la presente invención.

50 **Secuencias de control:** la expresión "secuencias de control" significa todos los componentes necesarios para la expresión de un polinucleótido que codifica un polipéptido de la presente invención. Cada una de las secuencias de control puede ser nativa o extraña al polinucleótido que codifica el polipéptido o nativa o extraña entre sí. Dichas

5 secuencias de control incluyen, pero no se limitan a una secuencia conductora, una secuencia de poliadenilación, una secuencia propeptídica, un promotor, una secuencia de péptidos señal y un terminador de la transcripción. Como mínimo, las secuencias de control incluyen un promotor y señales de parada de la transcripción y traducción. Las secuencias de control pueden proporcionarse con enlazadores con el fin de introducir sitios de restricción específicos que faciliten el ligamiento de las secuencias de control con la región codificante del polinucleótido que codifica un polipéptido.

Enlazada operativamente: la expresión "enlazada operativamente" significa una configuración en la cual una secuencia de control se coloca en una posición apropiada con respecto a la secuencia codificante de un polinucleótido de modo que la secuencia de control dirige la expresión de la secuencia codificante.

10 **Expresión:** el término "expresión" incluye cualquier etapa implicada en la producción del polipéptido que incluye, pero no se limita a la transcripción, la modificación post-transcripcional, la traducción, la modificación post-traducciona y la secreción.

15 **Vector de expresión:** la expresión "vector de expresión" significa una molécula de ADN lineal o circular que comprende un polinucleótido que codifica un polipéptido y que está enlazada operativamente a nucleótidos adicionales que proporcionan su expresión.

20 **Célula huésped:** la expresión "célula huésped" significa cualquier tipo de célula que sea susceptible de transformación, transfección, transducción y similares, con una construcción de ácido nucleico o un vector de expresión que comprenda un polinucleótido de la presente invención. La expresión "célula huésped" abarca cualquier progenie de una célula progenitora que no sea idéntica a la célula progenitora debido a mutaciones que se producen durante la replicación.

25 **Variante:** el término "variante" significa un polipéptido que tiene actividad hidroxilasa que comprende una alteración, es decir, una sustitución, inserción y/o deleción de uno o más (varios) residuos aminoácidos en una o más (varias) posiciones. Una sustitución significa un reemplazo de un aminoácido que ocupa una posición con un aminoácido diferente; una deleción significa la separación de un aminoácido que ocupa una posición; y una inserción significa añadir 1-3 aminoácidos adyacentes a un aminoácido que ocupa una posición.

Descripción Detallada de la Invención

Caroteno hidroxilasa significa de aquí en adelante una proteína o un polipéptido de acuerdo con la invención, es decir, una proteína que tiene, por ejemplo, una actividad enzimática para convertir [beta]-caroteno en zeaxantina o cantaxantina en astaxantina, que comprende la secuencia de aminoácidos SEQ ID NO: 2.

30 La secuencia de aminoácidos representada en SEQ ID NO: 2 se deriva de la traducción de la secuencia de ADNc representada en SEQ ID NO: 1. La secuencia optimizada de SEQ ID NO: 1 se muestra en SEQ ID NO: 3.

35 Sustitución significa el reemplazo de uno o más aminoácidos por uno o más aminoácidos. Los reemplazos son preferiblemente aquellos llamados conservativos, en los cuales el aminoácido reemplazado tiene una propiedad similar al aminoácido original, por ejemplo, el reemplazo de Glu por Asp, Gin por Asn, Val por lie, Leu por lie, Ser por Thr.

La deleción es el reemplazo de un aminoácido por un enlace directo. Posiciones preferidas para las deleciones son los extremos del polipéptido y los enlaces entre los dominios de proteínas individuales.

Las inserciones son introducciones de aminoácidos en la cadena polipeptídica, siendo formalmente el reemplazo de un enlace directo por uno o más aminoácidos.

40 La homología entre dos proteínas significa la identidad de los aminoácidos a lo largo de toda la longitud de cada una de las proteínas, que se calcula por comparación con ayuda del programa de ordenador GAP (UWGCG, Universidad de Wisconsin, Genetic Computer Group, algoritmo del programa de Needleman y Wunsch, *J. Mol. Biol.* 1970, 48: 443-453), estableciendo los siguientes parámetros:

45

Peso del Hueco:	12
Peso Longitud:	4
Emparejamiento Medio:	2.912
Emparejamiento Erróneo Medio:	-2.003

- Preferiblemente, las construcciones de acuerdo con la invención abarcan un promotor 5' aguas arriba de la secuencia codificante en cuestión y una secuencia de terminador 3' aguas abajo y, opcionalmente, otros elementos reguladores habituales y, en cada caso, están operativamente enlazados con la secuencia codificante. Debe entenderse que el enlace operativo significa la disposición secuencial del promotor, la secuencia codificante, el terminador y, si es apropiado, otros elementos reguladores de tal manera que cada uno de los elementos reguladores pueda cumplir su función prevista en la expresión de la secuencia codificante. Ejemplos de secuencias operativamente enlazables son secuencias de fijación de objetivo, o bien potenciadores de la traducción, potenciadores, señales de poliadenilación y similares. Elementos reguladores adicionales abarcan marcadores seleccionables, señales de amplificación, orígenes de replicación y similares.
- Además de las secuencias reguladoras artificiales, la secuencia reguladora natural todavía puede estar presente aguas arriba del gen estructural real. Si se desea, esta regulación natural puede desactivarse por modificación genética, y la expresión de los genes puede potenciarse o disminuirse. Sin embargo, la construcción del gen también puede ser más simple en la construcción, es decir, no se insertan señales reguladoras adicionales aguas arriba del gen estructural y no se separa el promotor natural con su regulación. En cambio, la secuencia reguladora natural se muta de tal manera que la regulación ya no tiene lugar y la expresión del gen aumenta o disminuye. Una o más copias de las secuencias de ácido nucleico pueden estar presentes en la construcción génica.
- Ejemplos de promotores adecuados son: promotor *cos*, *tac*, *trp*, *tet*, *trp-tet*, *lpp*, *lac*, *lpp-lac*, *lacIq*, *T7*, *T5*, *T3*, *gal*, *trc*, *ara*, *SP6*, *I-PR* o *I-PL*, que se emplean ventajosamente en bacterias Gram-negativas; y promotores Gram-positivos *amy* y *SP02*, los promotores de levadura *ADC1*, *MFa*, *Ac*, *P-60*, *CYC1*, *GAPDH*, *TEF1* o los promotores de plantas *CaMV/35S*, *SSU*, *OCS*, *lib4*, *usp*, *STLS1*, *B33*, *nos* o el promotor de ubiquitina o de faseolina. Se da particular preferencia al uso de promotores inducibles, por ejemplo promotores inducibles por la luz y, en particular, por la temperatura, tales como el promotor *PrP1*.
- En principio, se pueden utilizar todos los promotores naturales con sus secuencias reguladoras. Además, los promotores sintéticos también pueden utilizarse de una manera ventajosa.
- Las secuencias reguladoras arriba mencionadas están destinadas a permitir la expresión fijada como objetivo de las secuencias de ácido nucleico y la expresión de proteínas. Dependiendo del organismo huésped, esto puede significar, por ejemplo, que el gen se expresa o sobre-expresa solo después de que haya tenido lugar la inducción, o que se expresa o sobre-expresa de forma inmediata y/o constitutiva.
- Las secuencias o los factores reguladores pueden tener preferiblemente un efecto positivo sobre la expresión y de esta manera aumentar o reducir esta última. Por lo tanto, una mejora de los elementos reguladores puede tener lugar ventajosamente en el nivel transcripcional utilizando señales de transcripción fuertes tales como promotores y/o "potenciadores". Además, la traducción también puede potenciarse mejorando, por ejemplo, la estabilidad del ARNm.
- Un casete de expresión se genera fusionando un promotor adecuado con una secuencia adecuada de nucleótidos hidroxilasa y una señal de terminador o señal de poliadenilación. Para este fin, se utilizan las técnicas habituales de recombinación y clonación tal como se describen, por ejemplo, en T. Maniatis, E.F. Fritsch y J. Sambrook, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1989)) y en T.J. Silhavy, M.L. Berman y L.W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1984) y en Ausubel, F.M. et al., *Current Protocols in Molecular Biology*, Greene Publishing Assoc. y Wiley Interscience (1987).
- Para la expresión en un organismo huésped adecuado, la construcción de ácido nucleico recombinante o la construcción génica se inserta ventajosamente en un vector específico para el huésped que permite una expresión génica óptima en el huésped. Los vectores son bien conocidos por el experto y se pueden encontrar, por ejemplo, en "Cloning Vectors" (Pouwels P.H. et al., Ed., Elsevier, Amsterdam-Nueva York-Oxford, 1985). Debe entenderse que los vectores significan no solo plásmidos, sino todos los demás vectores conocidos por los expertos, tales como, por ejemplo, fagos, virus, tales como SV40, CMV, baculovirus y adenovirus, transposones, elementos IS, plásmidos, cósmidos y ADN lineal o circular. Estos vectores pueden replicarse de forma autónoma en el organismo huésped o cromosómicamente.
- Los vectores de acuerdo con la invención permiten la generación de microorganismos recombinantes que se transforman, por ejemplo, con al menos un vector de acuerdo con la invención y que pueden emplearse para producir los mutantes. Las construcciones recombinantes arriba descritas de acuerdo con la invención se introducen ventajosamente en un organismo huésped adecuado y se expresan. Se prefiere utilizar los métodos habituales de clonación y transfección conocidos por los expertos con el fin de lograr la expresión de los ácidos nucleicos

mencionados anteriormente en el sistema de expresión en cuestión. Sistemas adecuados se describen, por ejemplo, en los protocolos actuales en biología molecular, F. Ausubel et al., eds., Wiley Interscience, Nueva York 1997.

- 5 Organismos huéspedes adecuados son, en principio, todos los organismos que permiten la expresión de los ácidos nucleicos de acuerdo con la invención, sus variantes alélicas y sus equivalentes o derivados funcionales. Organismos iniciales preferidos son aquellos capaces de sintetizar de forma natural carotenoides. Sin embargo, también son adecuados organismos iniciales capaces de sintetizar carotenoides debido a la introducción de genes de la biosíntesis de carotenoides. Los organismos iniciales significan organismos procarióticos o eucarióticos, tales como, por ejemplo, microorganismos o plantas. Microorganismos preferidos son bacterias, levaduras, algas u hongos.
- 10 Por lo tanto, la invención se refiere, además, a un procedimiento para preparar los organismos modificados genéticamente que se describen más adelante, en donde el gen caroteno hidroxilasa de acuerdo con la invención se introduce en el genoma del organismo inicial. Por organismos iniciales se entienden los organismos antes de la modificación genética de acuerdo con la invención.
- 15 El gen caroteno hidroxilasa de acuerdo con la invención puede, en principio, introducirse por todos los métodos conocidos por el experto en los organismos iniciales descritos más adelante, que se modifican con ello genéticamente.
- Se introducen ventajosamente en los organismos iniciales o células de los mismos mediante transformación, transfección, electroporación, utilizando la llamada pistola de partículas, o por microinyección.
- 20 El experto puede encontrar métodos apropiados para microorganismos en los libros de texto de Sambrook, J. et al. (1989) *Molecular Cloning: A Laboratory manual*, Cold Spring Harbor Laboratory Press, por F.M. Ausubel et al. (1994) *Current protocols in molecular biology*, John Wiley and Sons, por D.M. Glover et al., *DNA Cloning Vol. 1*, (1995), IRL Press (ISBN 019-963476-9), por Kaiser et al. (1994) *Methods in Yeast Genetics*, Cold Spring Harbor Laboratory Press, o Guthrie et al. *Guide to Yeast Genetics and Molecular Biology*, *Methods in Enzymology*, 1994, Academic Press.
- 25 Ejemplos de métodos ventajosos que se pueden mencionar son aquellos tales como la introducción del ADN por recombinación homóloga o heteróloga, por ejemplo, utilizando el gen *URA3*, específicamente el gen *URA3* de *Ashbya*, tal como se describe en la solicitud alemana DE 19801120.2, y/o por el método REMI (= "integración mediada por enzimas de restricción") que se describe más adelante.
- 30 La técnica REMI se basa en la co-transformación de una construcción de ADN lineal que se ha cortado en ambos extremos con la misma endonucleasa de restricción, junto con la endonucleasa de restricción que se utilizó para esta restricción de la construcción de ADN, en un organismo. La endonucleasa de restricción corta entonces el ADN genómico del organismo en el que se ha introducido la construcción de ADN junto con la enzima de restricción. Esto conduce a una activación de los mecanismos de reparación propios de la célula. Estos mecanismos de reparación reparan las roturas de la cadena en el ADN genómico que han sido provocadas por la endonucleasa, y durante esto
- 35 también incorporan con cierta frecuencia la construcción de ADN co-transformado en el genoma. Por lo general, los sitios de escisión de restricción se conservan en ambos extremos del ADN durante esto.
- Esta técnica fue descrita por Bölker et al. (*Mol. Gen. Genet.*, 1995, 248: 547-552) para la mutagénesis de inserción de hongos. El método fue utilizado por Von Schiestl y Petes (*Proc. Natl. Acad. Sci. USA*, 1991, 88: 7585-7589) para averiguar si existe recombinación heteróloga en *Saccharomyces*. El método ha sido descrito por Brown et al. (*Mol. Gen. Genet.* 1996, 251: 75-80) para la transformación estable y la expresión regulada de un gen informador inducible.
- 40 Es posible usar el método REMI para colocar los fragmentos de ácido nucleico de acuerdo con la invención o el gen caroteno hidroxilasa mencionado anteriormente en sitios transcripcionalmente activos en el genoma.
- Es posible y ventajoso clonar los ácidos nucleicos junto con al menos un gen informador en una construcción de ADN, que se introduce en el genoma. Este gen informador debería facilitar la detectabilidad mediante un ensayo de crecimiento, fluorescencia, quimioluminiscencia o bioluminiscencia o mediante una medición fotométrica. Ejemplos que se pueden mencionar de genes informadores son genes de resistencia a antibióticos, genes hidrolasa, genes de proteínas fluorescentes, genes de bioluminiscencia, genes glucosidasa, el gen de la luciferasa, gen de la [beta]-galactosidasa, gen gfp, gen de la lipasa, gen de la esterasa, gen de la peroxidasa, gen de la [beta]-lactamasa, gen de la acetil-, fosfo- o adenil-transferasa. Estos genes hacen posible medir y cuantificar fácilmente la actividad de
- 50 transcripción y, por lo tanto, la expresión de los genes. Esto significa que es posible identificar sitios en el genoma que tengan una productividad que difiera en hasta un factor de 2.

Si se pretende introducir en el organismo una pluralidad de genes, tales como, por ejemplo, genes adicionales de la biosíntesis de carotenoides, todos pueden introducirse junto con un gen informador en un solo vector, o cada uno de los genes individuales con un gen informador se puede introducir en el organismo, en un vector en cada caso, siendo posible introducir los diversos vectores al mismo tiempo o sucesivamente. También es posible insertar fragmentos de genes que codifiquen las actividades respectivas utilizando las técnicas REMI.

Enzimas de restricción adecuadas en principio para integrar el gen de la caroteno hidroxilasa o construcciones de ácidos nucleicos de acuerdo con la invención en el genoma de organismos iniciales son todas conocidas por el experto. Enzimas de restricción que reconocen solo 4 pares de bases como el sitio de escisión de restricción son menos preferidos, porque se cortan con demasiada frecuencia en el genoma o en el vector a integrarse, y las enzimas preferidas reconocen 6, 7, 8 o más pares de bases como sitio de escisión, tales como *Bam*HI, *Eco*RI, *Bg*II, *Sph*I, *Spe*I, *Xba*I, *Xho*I, *Nco*I, *Sa*I, *Cla*I, *Kpn*I, *Hind*III, *Sac*I, *Pst*I, *Bpn*I, *Not*I, *Srf*I o *Sfi*I, por mencionar solo unas pocas de las posibles enzimas. Es ventajoso que las enzimas utilizadas ya no tengan sitios de escisión en el ADN que se ha de introducir; esto aumenta la eficiencia de la integración. En general, en la mezcla REMI se utilizan de 5 a 500 U, preferiblemente de 10 a 250, de manera particularmente preferible de 10 a 100 U de las enzimas. Las enzimas se emplean ventajosamente en una solución acuosa que contiene sustancias para la estabilización osmótica tales como azúcares, tales como sacarosa, trehalosa o glucosa, polioles, tales como glicerol o polietilenglicol, un tampón con un tamponamiento ventajoso en el intervalo de pH de 5 a 9, preferiblemente de 6 a 8, de manera particularmente preferible de 7 a 8, tales como tris, MOPS, HEPES, MES o PIPES y/o sustancias para estabilizar los ácidos nucleicos, tales como sales inorgánicas u orgánicas de Mg, Cu, Co, Fe, Mn o Mo. También es posible, caso de ser apropiado, que estén presentes otras sustancias, tales como EDTA, EDDA, DTT, [beta]-mercaptoetanol o inhibidores de nucleasa. Sin embargo, también es posible realizar la técnica REMI sin estas adiciones.

El procedimiento se lleva a cabo a una temperatura en el intervalo de 5 a 80°C, preferiblemente de 10 a 60°C, en particular preferiblemente de 20 a 40°C. Otros métodos conocidos para desestabilizar las membranas celulares son adecuados para el procedimiento, tales como, por ejemplo, la electroporación, la fusión con vesículas cargadas o la desestabilización con diversas sales de metales alcalinos o alcalinotérreos, tales como las sales de litio, rubidio o calcio, siendo preferidas las sales de litio.

La invención se refiere, además, a un organismo modificado genéticamente de forma correspondiente, aumentándose la expresión del gen de la caroteno hidroxilasa de acuerdo con la invención en comparación con un organismo de tipo salvaje en el caso de que el organismo inicial contenga un gen de la caroteno hidroxilasa, o sea provocado, en el caso en el que el organismo inicial no contenga un gen de la caroteno hidroxilasa, por la modificación genética.

Un organismo modificado genéticamente significa un organismo en el que se ha insertado el gen de caroteno hidroxilasa o la construcción de ácido nucleico de acuerdo con la invención, preferiblemente por uno de los métodos arriba descritos.

El organismo modificado genéticamente contiene al menos un gen de la caroteno hidroxilasa de acuerdo con la invención o al menos una construcción de ácido nucleico de acuerdo con la invención. Dependiendo del organismo inicial, el ácido nucleico puede estar presente dentro o fuera del cromosoma.

El metabolismo de los carotenoides en los organismos modificados genéticamente se altera preferiblemente en comparación con el tipo salvaje.

Organismos preferidos son hongos recombinantes y levaduras. En una realización particular, el hongo recombinante es oleaginoso, porque puede acumular lípidos hasta al menos aproximadamente el 20% de su peso de célula seco; y produce al menos un carotenoide seleccionado del grupo que consiste en anteraxantina, adonirrubina, adonixantina, astaxantina, cantaxantina, capsorrubina, β -criptoxantina, α -caroteno, β,ψ -caroteno, δ -caroteno, ϵ -caroteno, equinenona, 3-hidroxi equinenona, 3'-hidroxi equinenona, γ -caroteno, 4-ceto- γ -caroteno, ζ -caroteno, α -criptoxantina, desoxiflexixantina, diatoxantina, 7,8-didehidroastaxantina, didehidrolipopeno, fucoxantina, fucoxantanol, isorenierateno, β -isorenierateno, lactucaxantina, luteína, licopeno, mixobactona, neoxantina, neurosporeno, hidroxineurosporeno, peridinina, fitoeno, rodopina, rodopina glucósido, 4-ceto-rubixantina, sifonaxantina, esferoideno, esferoidenona, espiriloxantina, toruleno, 4-ceto-toruleno, 3-hidroxi-4-ceto-toruleno, uriolida, acetato de uriolida, violaxantina, zeaxantina- β -diglucósido, zeaxantina, un carotenoide C30 y combinaciones de los mismos, y pueden acumular el caroteno producido. a al menos aproximadamente el 1% de su peso de célula seca. Preferiblemente, el hongo recombinante es un miembro de un género seleccionado del grupo que consiste en: *Aspergillus*, *Blakeslea*, *Botrytis*, *Candida*, *Cercospora*, *Cryptococcus*, *Cunninghamella*, *Fusarium* (*Gibberella*), *Cluyveromyces*, *Lipomyces*, *Mortierella*, *Mucor*, *Neurospora*, *Penicillium*, *Phycomyces*, *Pichia* (*Hansenula*), *Puccinia*, *Pythium*, *Rhodospodium*, *Rhodotorula*, *Saccharomyces*, *Sclerotium*, *Trichoderma*, *Trichosporon*,

5 *Xanthophyllomyces* (*Phaffia*) y *Yarrowia*, o es de una especie seleccionada del grupo que consiste en: *Aspergillus terreus*, *Aspergillus nidulans*, *Aspergillus niger*, *Blakeslea trispora*, *Botrytis cinerea*, *Candida japonica*, *Candida pulcherrima*, *Candida revkaufi*, *Candida tropicalis*, *Candida utilis*, *Cercospora nicotianae*, *Cryptococcus curvatus*, *Cunninghamella echinulata*, *Cunninghamella elegans*, *Fusarium fujikuroi* (*Gibberella zeae*), *Kluyveromyces lactis*, *Lipomyces starkeyi*, *Lipomyces lipoferus*, *Mortierella alpina*, *Mortierella ramanniana*, *Mortierella isabellina*, *Mucor circinelloides*, *Neurospora crassa*, *Phycomyces blakesleanus*, *Pichia pastoris*, *Puccinia distincta*, *Pythium irregulare*, *Rhodospodium toruloides*, *Rhodotorula glutinis*, *Rhodotorula graminis*, *Rhodotorula mucilaginoso*, *Rhodotorula pinicola*, *Rhodotorula gracilis*, *Saccharomyces cerevisiae*, *Sclerotium rolfsii*, *Trichoderma reesei*, *Trichosporon cutaneum*, *Trichosporon pullulans*, *Xanthophyllomyces dendrorhous* (*Phaffia rhodozyma*) y *Yarrowia lipolytica*.

10 De estas cepas oleaginosas por naturaleza, algunas también producen de forma natural carotenoides y algunas otras no; estas cepas pueden utilizarse adicionalmente como una célula huésped mediante la introducción de genes de biosíntesis de carotenoides tal como se describe en la Patente de EE.UU. 7 851 199.

15 En otras realizaciones, la presente invención proporciona un método para producir un carotenoide, comprendiendo el método las etapas de cultivar un hongo en condiciones que permiten la producción del carotenoide; y aislar el carotenoide producido.

20 El cultivo del organismo modificado genéticamente de acuerdo con la invención tiene lugar de una manera conocida per se, tal como el cultivo del tipo salvaje apropiado, por ejemplo en el caso de microorganismos en un medio adecuado tal como, por ejemplo, en placas de agar o en cultivo en suspensión, o en el caso de plantas en el suelo o en medios nutrientes adecuados. Por cosecha se entiende, en el caso de los microorganismos, el aislamiento de los microorganismos, y en el caso de las plantas, el corte de la planta o, en su caso, partes particulares de la planta que contienen los carotenoides. Los carotenoides se aíslan de una manera conocida per se, por ejemplo, mediante la ruptura de las células del organismo, la extracción de los carotenoides y la posterior purificación de los carotenoides mediante métodos de separación química o física, tales como extracción o cromatografía.

Los siguientes ejemplos ilustran la invención.

25 **Ejemplos**

La Tabla 1 que figura a continuación describe determinadas cepas de *Yarrowia lipolytica* utilizadas en la siguiente ejemplificación:

Tabla 1: Cepas de *Yarrowia lipolytica*.

ML5252	<i>MATA erg9-4789::ura3 tef1P-{HMG-tr GGS carB carRP}</i> prototrófica	Técnicas de genética molecular clásicas y estándares
ML6804	<i>MATB erg9-4789::ura3 tef1P-{HMG-tr GGS carB carRP crtW}</i> prototrófica	Técnicas de genética molecular clásicas y estándares
ML9335	<i>MATA erg9-4789::URA3 tef1P-{HMG-tr GGS carB carRP crtW Dc-crtZ}</i> prototrófica	Técnicas de genética molecular clásicas y estándares
ML9863	<i>MATB erg9-4789::URA3 tef1P-{HMG-tr GGS carB carRP crtWXa-crtZ Dc-crtZ}</i> prototrófica	Técnicas de genética molecular clásicas y estándares
ML11218	<i>ML9863 crtW-Δ6180</i>	Ruptura fijada como objetivo con casete Hyg ^R ; posterior separación del marcador utilizando el sistema <i>cre-lox</i>
ML11453	<i>ML9335 tef1P-carRPA-/Hyg^R</i>	Transformación no fijada como objetivo; también pueden haberse incorporado copias no marcadas adicionales de actividades pre-existentes.
ML11584	<i>ML9863 tef1P-Ep-crtZ/Nat^R</i>	Transformación no fijada como objetivo; también pueden haberse incorporado copias no marcadas adicionales de actividades pre-existentes
ML11956	<i>MATB erg9-4789::ura3 tef1P-{HMG-tr GGS carB carRP crtW Dc-crtZ}</i>	ML11453 × ML11584

ML12526	ML9335 with <i>tef1P-HMG-tr tef1P-carB</i> <i>3X-tef1P-carRP</i>	Transformaciones no fijadas como objetivo, seguidas de separación de Hyg ^R y Nat ^R utilizando el sistema <i>cre-lox</i>
---------	---	---

5 Las cepas de *Yarrowia* M L5252, M L6804, M L9863 y ML9335 se construyeron mediante la introducción de genes heterólogos bajo el control del promotor *TEF1* endógeno, acoplado con varias generaciones de cruces, comenzando con ML350 y ATCC201249 tal como se describe en la patente de EE.UU. 7.851.199. El gen *GGS* y el gen *HMG* truncado ("*HMG-tr*") se derivaron de secuencias de *Yarrowia* correspondientes a los genes nativos geranilgeranil pirofosfato sintasa e hidroximetilglutaril-CoA reductasa, respectivamente. Los genes *carRP* y *carB* se derivaron de *Mucor circinelloides*, y codifican una fitoeno sintetasa/licopeno ciclasa bifuncional y una fitoeno deshidrogenasa, respectivamente. El gen *crtW* se sintetizó para codificar la caroteno ketolase de *Parvularcula bermudensis* (documento US 2012/0156718).

10 El gen *crtZ* se amplificó a partir de *Xanthobacter autotrophicus* (*Xa*) o se sintetizó para codificar la caroteno hidroxilasa de *Cronobacter pulveris* (antes *Enterobacter pulveris*) (*Ep*) (SEQ ID NO: 4), bacteria de Enterobacteriaceae DC404 (*Dc*) (SEQ ID NO: 5) o *Flavobacterium* sp. R1534 (*Fb*) (Patente de EE.UU. 6.087.152). Estos genes a veces, aunque no siempre, se asocian con marcadores auxotróficos (*URA3*, *LEU2*, *URA2*, *LYS1*, *ADE1*) o un sitio *loxP*, remanente de un marcador Hyg^R (resistencia a la higromicina) o Nat^R (resistencia a la nourseotricina).

15 Tabla 2: Plásmidos

Plásmido	Cadena principal	Inserción	Oligos o fuente
pMB6486	pMB6157 (Hyg ^R <i>tef1P-xprT</i>)	<i>Hp^{LF}-crtZ</i>	Fragmento <i>NheI-MluI</i> sintetizado
pMB6487	pMB6157 (Hyg ^R <i>tef1P-xprT</i>)	<i>Fb-crtZ</i>	Fragmento <i>NheI-MluI</i> sintetizado
pMB6056	pMB6157 (Hyg ^R <i>tef1P-xprT</i>)	<i>Ep-crtZ</i>	Fragmento <i>NheI-MluI</i> sintetizado

20 Todos los procedimientos básicos de biología molecular y manipulación del ADN descritos en esta memoria se realizan generalmente de acuerdo con Sambrook *et al.* o Ausubel *et al.* (J. Sambrook, E.F. Fritsch, T. Maniatis (eds). 1989. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press: Nueva York; F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Struhl (eds.). 1998. *Current Protocols in Molecular Biology*. Wiley: Nueva York).

Ejemplo 1a: Producción de pMB6486 (Hyg^R *tef-Hp^{LF}-crtZ*), que codifica caroteno hidroxilasa de *H. pluvialis* LF.

25 Como resultado de un análisis de la secuencia de ARNm utilizando *Haematococcus pluvialis* (UTEX N° 2505), los autores de la invención encontraron sorprendentemente la nueva β-caroteno hidroxilasa, que se denominó hidroxilasa Hp^{LF}-CrtZ (*H. pluvialis* LF (baja frecuencia).

30 Una secuencia de ORF de caroteno hidroxilasa (Hp^{LF}-CrtZ) optimizada por codones se sintetizó *de novo* en base a la secuencia proteica deducida del gen de *H. pluvialis*, utilizando la preferencia codónica de *Y. lipolytica*, tal como se especifica en SEQ ID NO: 3. Durante la síntesis *de novo*, la secuencia 5'-TGCTAGCCACAAAA, que contiene un sitio de restricción *NheI* y una secuencia de Kozak típica para permitir una traducción eficiente, se añadió inmediatamente aguas arriba del ATG. La secuencia ACGCGT-3', que comprende un sitio de restricción *MluI*, se añadió inmediatamente aguas abajo del codón de parada. Esta secuencia se escindió utilizando *NheI* y *MluI* y se ligó a pMB6157 cortado con *NheI* y *MluI* para producir pMB6486. La proteína Hp^{LF}-CrtZ codificada resultante de pMB6486 se especifica en SEQ ID NO: 2.

35 **Ejemplo 1b: Producción de pMB6056 (Hyg^R *tef-crtZ-Ep*), que codifica *C. pulveris* (anteriormente *E. pulveris*) caroteno hidroxilasa.**

40 Una secuencia de ORF de caroteno hidroxilasa (Ep-CrtZ) optimizada por codones se sintetizó *de novo* en base a la secuencia proteica deducida del gen de *Cronobacter pulveris*, anteriormente conocida como *Enterobacter pulveris*, (N° de Acceso a EMBL CAZ90621.1), utilizando la preferencia codónica de *Y. lipolytica*, tal como se especifica en SEQ ID NO: 4. Durante la síntesis *de novo*, la secuencia 5'-TGCTAGCCACAAAA, que contiene un sitio de restricción *NheI* y una secuencia de Kozak típica para permitir una traducción eficiente, se añadió inmediatamente aguas arriba del ATG. La secuencia ACGCGT-3', que comprende un sitio de restricción *MluI*, se añadió

inmediatamente aguas abajo del codón de parada. Esta secuencia se escindió utilizando *NheI* y *MluI* y se ligó en pMB6157 cortado con *NheI* y *MluI* para producir pMB6056.

Ejemplo 1c: Producción de pMB6487 (*Hyg^Rtef-crtZ-Fb*), que codifica *Flavobacterium sp. R1534* caroteno hidroxilasa.

5 Una caroteno hidroxilasa (Fb-CrtZ) optimizada por codones se sintetizó *de novo* en base a la secuencia proteica de *Flavobacterium sp. R1534* tal como se especifica en la patente de EE.UU. 6.087.152, utilizando la preferencia codónica de *Y. lipolytica*. Durante la síntesis *de novo*, la secuencia 5'-TGCTAGCCACAAAA, que contiene un sitio de restricción *NheI* y una secuencia de Kozak típica para permitir una traducción eficiente, se añadió inmediatamente aguas arriba del ATG. La secuencia ACGCGT-3', que comprende un sitio de restricción *MluI*, se añadió
10 inmediatamente aguas abajo del codón de parada. La secuencia se escindió utilizando *NheI* y *MluI* y se ligó a pMB6157 cortado con *NheI* y *MluI* para producir pMB6487.

Ejemplo 2A: Introducción de tres genes de caroteno hidroxilasa en la cepa ML6804 productora de cantoxantina de *Y. lipolytica* para producir astaxantina

15 Para testar el potencial de hidroxilación de Hp^{LF} CrtZ y compararlo con *Fb* CrtZ y *Ep* CrtZ, la cepa ML6804 se transformó con tres construcciones diferentes:

- 1) un fragmento *HindIII* - *XbaI* de pMB6487 que alberga el gen *crtZ* de *Flavobacterium* bajo el control de *tef1p* y el marcador de resistencia a higromicina *Hyg^R*;
- 2) un fragmento *PvuII* de pMB6056 que alberga el gen *crtZ* de *Cronobacter pulveris* (anteriormente *Enterobacter pulveris*) bajo el control de *tef1p* y el marcador de resistencia a higromicina *Hyg^R*;
- 20 3) un fragmento *HindIII* - *XbaI* de pMB6486 que alberga el gen *crtZ* de "baja frecuencia" de *H. pluvialis* bajo el control de *tef1p* y el marcador de resistencia a la higromicina *Hyg^R*.

Los transformantes se seleccionaron en YPD con 100 mg/L de higromicina después de 3-4 días de crecimiento a 30°C. Diez transformantes de cada una de las construcciones se cultivaron en matraces de agitación durante 4 días en YPD. Todos los transformantes produjeron astaxantina. En la Fig. 1 se muestra un transformante representativo
25 de cada una de las transformaciones junto con la cepa progenitora ML6804. La cepa ML12471, que contiene el gen *crtZ* de *Flavobacterium sp. R1534* produjo 3% de astaxantina y 9% de adonirrubina (como porcentaje del carotenoide total). La cepa ML11622, que contiene el gen *crtZ* de *Cronobacter pulveris* (anteriormente *Enterobacter pulveris*), produjo 7% de astaxantina y 11% de adonirrubina. Y la cepa ML12466, que contiene el gen *crtZ* de *H. pluvialis*^{LF}, produjo 12% de astaxantina y 22% de adonirrubina.

30 **Ejemplo 2B. Introducción de caroteno hidroxilasa de *H. pluvialis LF* en cepas productoras de astaxantina de *Y. lipolytica* para aumentar la pureza de astaxantina**

Las cepas ML9863 y ML11956 se transformaron con un fragmento *HindIII* - *XbaI* de pMB6486 que alberga el promotor *tef1*, el gen *crtZ* de Hp^{LF} y el marcador seleccionable para la resistencia a la higromicina, *Hyg^R*. Se eligieron
35 veinte transformantes de *Hyg^R* de cada una de las cepas que parecieron más oscuras que los progenitores en las placas de transformación (YPD con 100 mg/L de higromicina) después de 3-4 días de crecimiento a 30°C. Los transformantes se cultivaron en matraces de agitación durante 4 días en YPD. En la Fig. 2 se muestran dos transformantes representativos en comparación con sus cepas progenitoras. ML12562 se deriva de ML9863 y ML12566 de ML11956. Como porcentaje de carotenoides totales, ML12562 y ML12566 produjeron 3 veces más astaxantina (40% frente a 13%) y 2 veces más (27% frente a 13%), respectivamente. que sus progenitores (Fig. 2).

40 **Ejemplo 2C. Introducción de la hidroxilasa de *H. pluvialis LF* en una cepa productora de β -caroteno de *Y. lipolytica* para producir β -criptoxantina:**

La cepa ML5252 se transformó con tres construcciones diferentes:

- 1) un fragmento *HindIII* - *XbaI* de pMB6487 que alberga el gen *crtZ* de *Flavobacterium sp. R1534* bajo el control de *tef1p* y el marcador de resistencia a higromicina *Hyg^R*;
- 45 2) un fragmento *PvuII* de pMB6056 que alberga el gen *crtZ* de *Cronobacter pulveris* (anteriormente *Enterobacter pulveris*) bajo el control de *tef1p* y el marcador de resistencia a higromicina *Hyg^R*;
- 3) un fragmento *HindIII* - *XbaI* de pMB6486 que alberga el gen *crtZ* de *H. pluvialis LF* bajo el control de *tef1p* y el marcador de resistencia a la higromicina *Hyg^R*.

5 Los transformantes se seleccionaron en YPD con 100 mg/L de higromicina después de 3-4 días de crecimiento a 30°C. Diez transformantes de cada una de las construcciones se cultivaron en matraces de agitación durante 4 días en YPD. Todos los transformantes produjeron zeaxantina y β -criptoxantina. Los transformantes representativos se muestran en la Fig. 3 en comparación con la cepa progenitora ML5252. La cepa ML12458, que contiene el gen *crtZ* de *Flavobacterium*, produjo 6% de β -criptoxantina y 6% de zeaxantina (como porcentaje del carotenoide total). La cepa ML10341, que contiene el gen *crtZ* de *Cronobacter pulveris* (anteriormente *Enterobacter pulveris*), produjo un 3% de β -criptoxantina y un 39% de zeaxantina (como porcentaje del carotenoide total). Y la cepa ML12453 que contiene el gen *crtZ* de *H. pluvialis* LF produjo 21% de β -criptoxantina y 4% de zeaxantina (como porcentaje del carotenoide total).

10 **Extracción y cuantificación de la producción de carotenoides por HPLC a partir de células de *Yarrowia lipolytica***

Los ensayos en matraz de agitación y el análisis de carotenoides de las cepas generadas se realizaron de acuerdo con los métodos descritos previamente en la patente de EE.UU. Nº 7.851.199 B2.

LISTADO DE SECUENCIAS

- 15 <110> DSM IP Assets B.V.
 <120> Caroteno Hidroxilasa
 <130> Case 28352 PCT
 <160> 5
 <170> PatenIn versión 3.5
- 20 <210> 1
 <211> 894
 <212> ADN
 <213> Haematococcus pluvialis
- 25 <400> 1

ES 2 702 075 T3

```

atggctttgt cgaagatgca gccaatgggg gtgcaaggac ggcgcatcgc attgcctcgt      60
gacatctcaa ggctcagat gccctctcaa aggcagccat tggcccaggt gctacgcgta      120
gcagccccag acacagaggc ggtggagtcc tcacgggtgc agccgtctgt cgggtgcaggc      180
aatcagcaaa ctgcagacga cgccctcaa cagctcgacc ggtccatcgc tctgcgcgcg      240
gctcaacgca aaagggagca gctgacctac caggctgcag ccatagcagc gtccgtgggc      300
atatcaagcc ttgccatcct tgccacctac acgcgggttt caatgcacat cacatcagat      360
ggcgagatgc cttggagcga cctcttgggc accctctcac tgggtggcagg tggcgcgttt      420
ggcatggaga tgtatgcgcg ttatgcacac aaggccctat ggcatgagtc ccccctgggt      480
tggctgctgc acaagagcca ccacactccc cgcaccggac cctttgaagc caatgactta      540
ttcgccatca tcaatggcct gccagccatg ctgctgtgct attatgggtt ctgggagcct      600
ggcatggtcg ggcctcatg ctttgggtct gggctgggca tcaccctgta cgggatggcg      660
tacatgttca tccatgacgg cctagtccac cggcgtttc ccaccgggcc catctcagac      720
ctgccctaca tgaagcgcct gacagtggcc caccagatcc accacagcgg caagtatgat      780
ggcgcgccct ggggcatggt cctggggcct caggagctgg aggcaatatc aggtgccaca      840
gaggagctgg accgcttggg ggcagacctg gactggtcaa agcgcagggt gtga      894

```

<210> 2

<211> 297

<212> PRT

5 <213> Haematococcus pluvialis

<400> 2

```

Met Ala Leu Ser Lys Met Gln Pro Met Gly Val Gln Gly Arg Arg Ile
1           5           10           15

```

```

Ala Leu Pro Arg Asp Ile Ser Arg Pro Gln Met Pro Ser Gln Arg Gln
          20           25           30

```

ES 2 702 075 T3

Pro Leu Ala Gln Val Leu Arg Val Ala Ala Pro Asp Thr Glu Ala Val
 35 40 45

Glu Ser Ser Arg Val Gln Pro Ser Val Gly Ala Gly Asn Gln Gln Thr
 50 55 60

Ala Asp Asp Ala Leu Gln Gln Leu Asp Arg Ser Ile Ala Leu Arg Arg
 65 70 75 80

Ala Gln Arg Lys Arg Glu Gln Leu Thr Tyr Gln Ala Ala Ala Ile Ala
 85 90 95

Ala Ser Val Gly Ile Ser Ser Leu Ala Ile Leu Ala Thr Tyr Thr Arg
 100 105 110

Phe Ser Met His Ile Thr Ser Asp Gly Glu Met Pro Trp Ser Asp Leu
 115 120 125

Leu Gly Thr Leu Ser Leu Val Ala Gly Gly Ala Phe Gly Met Glu Met
 130 135 140

Tyr Ala Arg Tyr Ala His Lys Ala Leu Trp His Glu Ser Pro Leu Gly
 145 150 155 160

Trp Leu Leu His Lys Ser His His Thr Pro Arg Thr Gly Pro Phe Glu
 165 170 175

Ala Asn Asp Leu Phe Ala Ile Ile Asn Gly Leu Pro Ala Met Leu Leu
 180 185 190

Cys Tyr Tyr Gly Phe Trp Glu Pro Gly Met Val Gly Ala Ser Cys Phe
 195 200 205

Gly Ala Gly Leu Gly Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Ile
 210 215 220

His Asp Gly Leu Val His Arg Arg Phe Pro Thr Gly Pro Ile Ser Asp
 225 230 235 240

Leu Pro Tyr Met Lys Arg Leu Thr Val Ala His Gln Ile His His Ser
 245 250 255

Gly Lys Tyr Asp Gly Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu
 260 265 270

Leu Glu Ala Ile Ser Gly Ala Thr Glu Glu Leu Asp Arg Leu Val Ala
 275 280 285

Asp Leu Asp Trp Ser Lys Arg Arg Val
 290 295

ES 2 702 075 T3

<211> 891
 <212> ADN
 <213> Haematococcus pluvialis

5 <400> 3
 atggctctgt ccaagatgca gcccatgggt gtccagggcc gacgaattgc tctccccga 60
 gacatctccc gaccccagat gccttcccag cgacagcctc tcgcccaggt tctgcgagtt 120
 gctgcccccg acaccgaggc cgtcgagtct tctcgagtcc agccctccgt cggtgccggc 180
 aaccagcaga ccgccgacga cgccctccag cagctcgacc gatccattgc tctgcgacga 240
 gcccagcgaa agcgagagca gctcacctac caggctgctg ccattgccgc ctccgttggc 300
 atctcctctc tcgccatcct cgccacctac acccgattct ccatgcacat cacctccgac 360
 ggcgagatgc cctggtcgca tctcctcggc actctgtctc tagtogccgg tggcgccttt 420
 ggcatggaga tgtacgccc atacgcccac aaggctctgt ggcaagagtc tcccctcggc 480
 tggctcctcc acaagtccca ccacactccc cgaaccggcc ccttcgaggc caacgacctc 540
 tttgccatca tcaacggctc gcccgccatg ctctctgct actacggctt ctgggagccc 600
 ggtatggtcg gtgcttcttg tttcgggtct ggcctcggta tcaccctcta cggtatggcc 660
 tacatgttca tccacgacgg tctgggtccac cgacgattcc ccaactggccc catctccgat 720
 ctgccctaca tgaagogaact caccgtcgc caccagatcc accactccgg caagtacgac 780
 ggtgctccct ggggtatggt cctcggcccc caggagctgg aggccatttc cggtgccacc 840
 gaggagcttg accgactcgt tgccgatctc gactggtcca agcagcagat c 891

<210> 4
 <211> 567
 <212> ADN
 <213> Enterobacter pulveris

10 <400> 4
 atgcttgtgc tgtggaatac gcttatcgtc atcacgacat tttgcctgat ggagattgtg 60
 gcaacgcttg ccataaata catcatgcac ggctggggat ggggctggca ccgttcgcac 120
 caccagccgc ggcacggctg gtttgaggtc aacgatctct acgcagtggc gttcgcactg 180
 ctcgccatcg tgctgatcgc actgggaaca gcaggctggc ggccgctgca atggcttggc 240
 gcaggcatga cgctctacgg tctgttctat ttctcgtgc acgacggctc ggtgcaccgg 300
 agatggccgt ttaactacat tccgcgaga ggctacctta agcggcttta ccaggcacac 360
 15 cggctgcacc acgcagtgaa tggcaaagag ggttgctct ccttcggttt tctctacgca 420
 gctaagccgg aagcaattca ggcagagctg cggagacgtc acggtcggaa acctaaagca 480
 gacgccgcca gagccggcg ggacggtaaa cctgtcgcac agagcgagag ccgagcacia 540
 gacctgccgc cgaaatcggc agagtaa 567

ES 2 702 075 T3

<210> 5
<211> 558
<212> ADN
<213> Enterobacteriaceae DC404

5

```
<400> 5
atgcttgcgt tgtggaatac cgggatcgtg ctactgacta tcatcatcat ggaaggggtg      60
gcaacgttcg cacacaagta catcatgcac ggctggggat ggggctggca tcattcgcac      120
cataccccgc gcaaaggggc gtttgagcgt aacgatctct atgcggtggt gtttgcgcta      180
ctggccattg cgctgattta cgcgggcagc gaaggg tact ggccgcttca gtggattggc      240
gcgggaatga cgggctacgg cgtgatctac tttatcgttc acgatggctt tgtccaccag      300
cgctggccgt tccgttacgt gccgcgcgcg ggctatctgc gccgcctcta catggcacac      360
cggctgcata acgcgggtgcg ggggcgcgaa ggggtcgtct ccttcgggtt tatctacgcc      420
ccaccggtgg acaagctgca ggcggtgctg cgcgaaacgta acggcagacc cgccagcgcg      480
ggcgctgcca gaggtgcgga tcgcgcggcg gccagctcgc cttccgggaa gccatcgctt      540
gcttcgcgcc gaaataa                                     558
```

10



REIVINDICACIONES

- 5 1. Un polipéptido que tiene actividad de hidroxilasa de la hidroxilación microbiana de isoprenoides, en donde dicho polipéptido comprende la secuencia de SEQ ID NO: 2, y en donde el polipéptido tiene una actividad enzimática para convertir cantaxantina en adonirrubina y/o astaxantina, en donde el porcentaje de astaxantina es al menos 12% basado en la cantidad total de carotenoides.
2. Un polipéptido de acuerdo con la reivindicación 1, en donde el polipéptido tiene, además, una actividad enzimática para convertir beta-caroteno en beta-criptoxantina y/o zeaxantina, en donde el porcentaje de criptoxantina es al menos 21% basado en la cantidad total de carotenoides.
3. Un ácido nucleico aislado que codifica el polipéptido de la reivindicación 1 o 2.
- 10 4. El ácido nucleico aislado de la reivindicación 3, que consiste en la secuencia representada en SEQ ID NO: 1 y SEQ ID NO: 3.
5. Una construcción de ácido nucleico o un vector de expresión que comprende el polinucleótido de la reivindicación 3 o 4, enlazado operativamente a una o más secuencias de control que dirigen la producción del polipéptido en una célula huésped de expresión.
- 15 6. Un microorganismo transformado, en donde se expresa la construcción de ácido nucleico de la reivindicación 5.
7. El microorganismo transformado de la reivindicación 6, cuyo metabolismo de carotenoides es diferente del de un tipo salvaje.
8. El microorganismo transformado de la reivindicación 7, en donde dicho microorganismo es una cepa oleaginoso.
- 20 9. El microorganismo transformado de la reivindicación 8, en donde dicha cepa oleaginoso es una cepa de *Yarrowia lipolytica*.
10. Un procedimiento para producir el microorganismo transformado de la reivindicación 6, que comprende introducir un ácido nucleico o una construcción de ácido nucleico de la secuencia representada en SEQ ID NO: 1, que está funcionalmente enlazado a una o más señales de regulación.
- 25 11. Un procedimiento para la preparación de carotenoides o derivados carotenoides, que comprende convertir una beta-ionona en una hidrox-beta-ionona y/o una 4-ceto-beta-ionona en un elemento estructural de 3-hidrox-4-ceto-beta-ionona, en presencia del polipéptido de la reivindicación 1 o 2, que comprende las etapas de:
 - (a) cultivar un microorganismo transformado de acuerdo con cualquiera de las reivindicaciones 6 a 9, bajo condiciones que conducen a la producción de carotenoides; y
 - (b) recuperar el carotenoide o derivado de carotenoide.
- 30 12. Una composición que comprende el polipéptido obtenido de *Haematococcus pluvialis* de acuerdo con cualquiera de las reivindicaciones 1 o 2.

Fig. 1

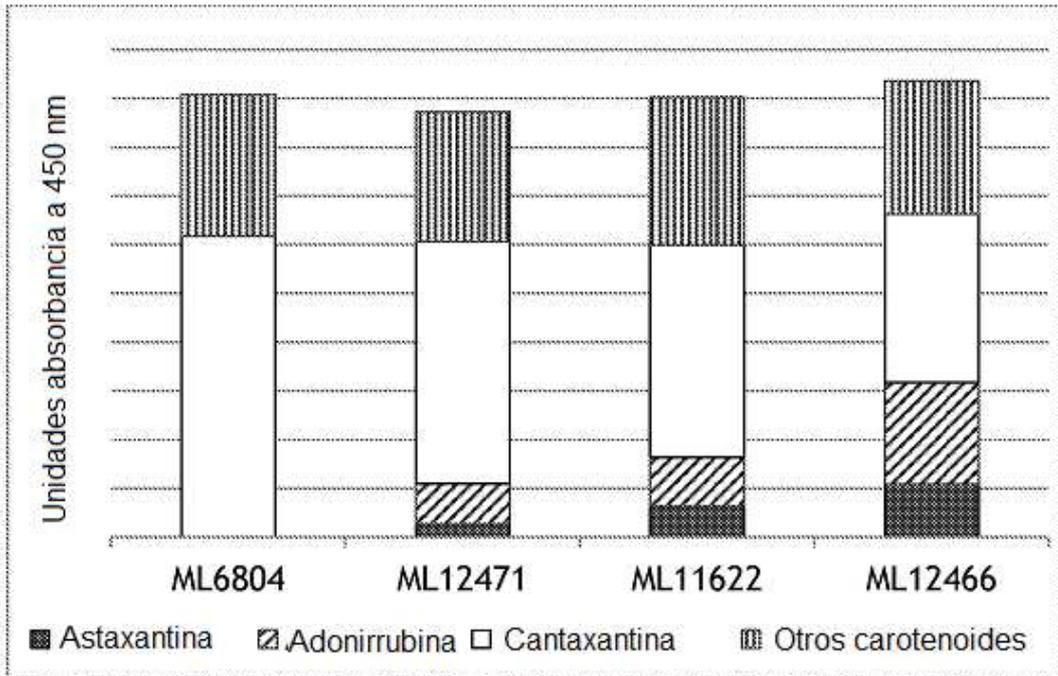


Fig. 2

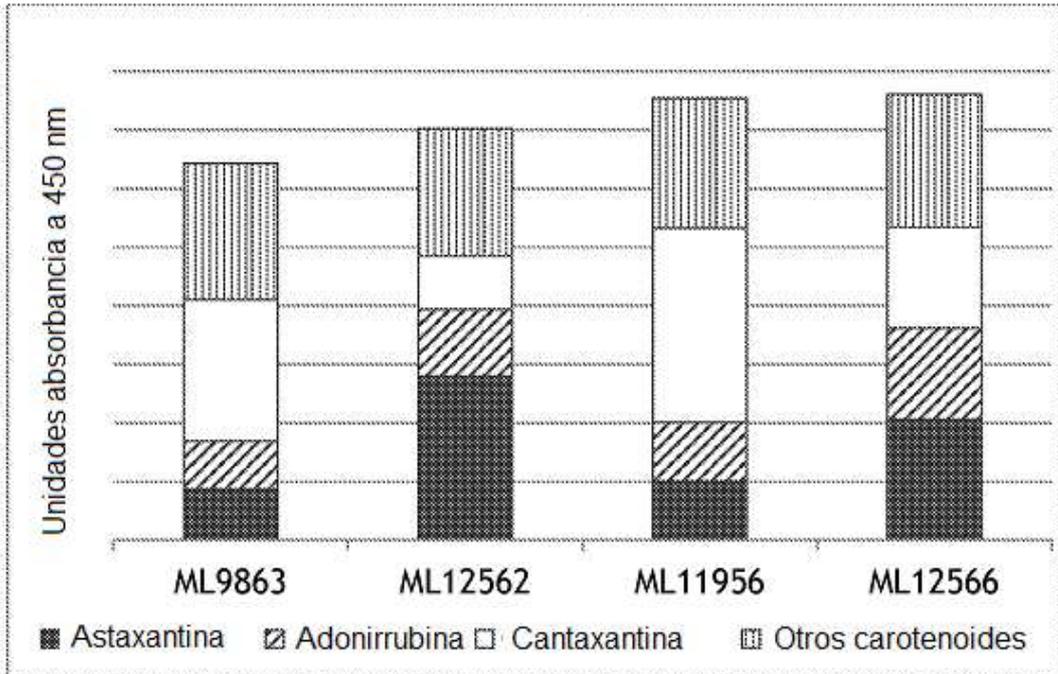


Fig. 3

