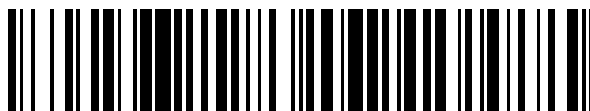


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 772 123**

51 Int. Cl.:

C12N 9/16 (2006.01)

C12N 15/82 (2006.01)

C12P 7/64 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **24.07.2015 PCT/US2015/042044**

87 Fecha y número de publicación internacional: **28.01.2016 WO16014968**

96 Fecha de presentación y número de la solicitud europea: **24.07.2015 E 15747911 (4)**

97 Fecha y número de publicación de la concesión europea: **20.11.2019 EP 3172320**

54 Título: **Variantes de tioesterasas y métodos de uso**

30 Prioridad:

24.07.2014 US 201462028641 P

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

07.07.2020

73 Titular/es:

**CORBION BIOTECH, INC. (100.0%)
One Tower Place, Suite 600
South San Francisco, CA 94080, US**

72 Inventor/es:

**BHAT, RIYAZ;
DAVIS, DAVID;
FRANKLIN, SCOTT y
MOSELEY, JEFFREY L.**

74 Agente/Representante:

TOMAS GIL, Tesifonte Enrique

ES 2 772 123 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

Variantes de tioesterasas y métodos de uso

CAMPO

- 5 [0001] La presente invención se refiere a variantes de acil-ACP tioesterasas, como se reivindica, y su uso en células productoras de aceite, por ejemplo, para aumentar la actividad enzimática hacia determinados sustratos de acil-ACP y para promover una mayor producción de aceites con perfiles de ácidos grasos deseados.

ANTECEDENTES

- 10 [0002] Hoy en día, las grasas y los ácidos grasos vienen principalmente de fuentes vegetales y animales, con la excepción notable de la producción comercial de ácidos grasos omega-3 por fermentación de microbios para su uso en una fórmula para bebés y suplementos nutricionales. Sin embargo, el progreso se está realizando hacia la producción comercial de aceites personalizados utilizando microalgas recombinantes. Véanse las publicaciones PCT WO2008/151149, WO2010/063032, WO2011/150410, WO2011/150411, y la solicitud de patente internacional PCT/US12/23696.

- 15 [0003] Un método para producir un perfil de ácidos grasos deseado en un organismo oleaginoso es introducir un transgén de acil-ACP tioesterasa; por ejemplo, un transgén a partir de una planta que produce un ácido graso deseado.

- 20 [0004] Al finalizar la biosíntesis de ácidos grasos, la acil-proteína portadora de acilos (ACP) tioesterasa (TE) determina funcionalmente la longitud e identidad del producto final de ácido graso (Salas et al., (2002) Archives of Biochemistry and Biophysics 403: 25-34). Según los alineamientos de secuencias de aminoácidos, se ha demostrado que las TEs de la planta se agrupan en dos familias, FatAs, que muestran una marcada preferencia por 18:1-ACP con actividad menor hacia 18:0- e 16:0-ACPs; y FatBs, que hidrolizan principalmente acil-ACP saturadas con longitudes de cadena que varían entre 8-16 carbonos (Voelker, en Genetic Engineering, Volumen 18. Editado por: Setlow JK. Nueva York, Plenum Press; 1996:111-133; Ginalska, et al., Nucl Acids Res (2003) 31:3291-3292; y Jones, et al., (1995) Plant Cell 7: 359-371). Las TEs FatB tienen un dominio de 18 aminoácidos hidrofóbico conservado (Facciotti y Yuan (1998) European Journal of Lipid Science and Technology 100:167-172) y una tríada catalítica de Asn-His-Cys en el dominio catalítico del extremo C-terminal (Blatti, et al., PLoS ONE (2012) 7(9): e42949. doi:10.1371 y Mayer y Shanklin, BMC Plant Biology (2007) 7:1-11). Mayer y Shanklin, BMC Plant Biology (2007) 7:1-11, identifican un dominio catalítico de acil-ACP tioesterasa conservado del extremo C-terminal caracterizado por un pliegue de tipo *hot-dog* en el extremo C-terminal que comprende la tríada catalítica Cys-His-Asn.

RESUMEN

- 35 [0005] En un aspecto se proporciona una acil graso-ACP tioesterasa no natural, un gen aislado que codifica la tioesterasa no natural, un casete de expresión que expresa la tioesterasa no natural, o una célula huésped que comprende el casete de expresión, donde la proteína no natural tiene al menos 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad con SEQ ID NO: 45 y comprende alanina (A) en la posición que corresponde con la posición 74 de SEQ ID NO: 45 (G96 de Gm FATA de tipo salvaje). En algunas formas de realización, la proteína no natural también comprende F, K o S en la posición correspondiente a la posición 69 de SEQ ID NO: 45 (L91 de Gm FATA de tipo salvaje), y/o F, A, K o V en la posición correspondiente a la posición 134 de SEQ ID NO: 45 (T156 de Gm FATA de tipo salvaje). En algunas formas de realización, la proteína no natural también comprende A o V en la posición correspondiente a la posición 89 de SEQ ID NO: 45 (S111 de Gm FATA de tipo salvaje) y/o A en la posición correspondiente a la posición 171 de SEQ ID NO: 45 (V193 de Gm FATA de tipo salvaje), y/o A o V en la posición correspondiente a la posición 86 de SEQ ID NO: 45 (G108 de Gm FATA de tipo salvaje).

- 45 [0006] En otro aspecto se proporciona un método para producir un aceite de triglicéridos. El método comprende expresar, en una célula huésped, la proteína anteriormente descrita; cultivar la célula huésped; y aislar el aceite. En algunas formas de realización, la proteína también comprende A o V en la posición correspondiente a la posición 89 de SEQ ID NO: 45 (S111 de Gm FATA de tipo salvaje) y/o un en la posición correspondiente a la posición 171 de SEQ ID NO: 45 (V193 de Gm FATA de tipo salvaje), y/o A o V en la posición correspondiente a la posición 86 de SEQ ID NO: 45 (G108 de Gm FATA de tipo salvaje).

- 50 [0007] En otro aspecto se proporciona un método para aumentar los ácidos grasos C18:0 en un perfil de ácidos grasos de un aceite producido por una célula huésped opcionalmente oleaginosa. El método comprende proporcionar un gen parental que codifique una enzima acil graso-ACP tioesterasa A (FATA), mutando el gen para

codificar una tioesterasa no natural que tiene al menos 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad con SEQ ID NO: 45 y que comprende A en la posición correspondiente a la posición 74 de SEQ ID NO: 45 (G96 de Gm FATA de tipo salvaje); expresar el gen mutado en la célula huésped; y producir el aceite, por lo que el perfil de ácidos grasos del aceite se incrementa en ácidos grasos C18:0 relativos al gen parental. En varias formas de realización, el método implica mutar también el gen para tener F, K o S en la posición correspondiente a la posición 69 de SEQ ID NO: 45 (L91 de Gm FATA de tipo salvaje), y/o F, A, K o V en la posición correspondiente a la posición 134 de SEQ ID NO: 45 (T156 de Gm FATA de tipo salvaje). En varias formas de realización, el método también implica mutar el gen para tener A o V en la posición correspondiente a la posición 89 de SEQ ID NO: 45 (S111 de Gm FATA de tipo salvaje) y/o A en la posición correspondiente a la posición 171 de SEQ ID NO: 45 (V193 de Gm FATA de tipo salvaje), y/o A o V en la posición correspondiente a la posición 86 de SEQ ID NO: 45 (G108 de Gm FATA de tipo salvaje).

DEFINICIONES

[0008] Una "acil-ACP tioesterasa" o "acil-ACP TE" se refiere, indistintamente, a una enzima que cataliza la escisión de un ácido graso a partir de una proteína portadora de acilos (ACP) durante la síntesis de lípidos. Las acil-proteína portadora de acilos (ACP) tioesterasas (TEs) hidrolizan los enlaces tioéster de acil-ACP, liberando ácidos grasos libres y ACP.

[0009] El término "TE que prefiere acil-ACP" se refiere a la especificidad de sustrato de acil graso-ACP de una TE. Una TE que prefiere acil-ACP libera preferentemente un ácido graso particular de un sustrato de acil-ACP. Por ejemplo, la TE que prefiere acil-ACP puede liberar preferentemente un ácido graso dado sobre todos los demás ácidos grasos en el conjunto de ácidos grasos C8:0, C10:0, C12:0, C14:0, C16:0, C18:0, C18:1 y C18:2. La preferencia de la TE que prefiere acil-ACP se puede detectar como una V_{max} más alta (o una k_{cat} más alta, o una V/K más alta) en comparación con otros sustratos de ácido graso-ACP no preferidos. La preferencia se puede inferir de los cambios en el perfil de ácidos grasos de una célula genéticamente modificada para sobreexpresar la TE que prefiere acil-ACP relativa a una célula de control que no sobreexpresa la TE que prefiere acil-ACP.

[0010] La numeración de un polímero de aminoácidos o polímero de ácido nucleico dado "corresponde a" o es "relativa a" la numeración de un polímero de aminoácidos o polímero de ácido nucleico seleccionado cuando la posición de cualquier componente polimérico dado (por ejemplo, aminoácido, nucleótido, referido también, genéricamente, como un "residuo") se designa por referencia a la misma posición o a una posición equivalente (por ejemplo, en base a un alineamiento óptimo o a una secuencia de consenso) en el polímero de aminoácidos o de ácido nucleico seleccionado, en vez de por la posición numérica real del componente en el polímero dado.

[0011] Una "variante" es un polipéptido que comprende una secuencia que difiere en una o más posición(es) de aminoácidos de las de una secuencia de polipéptido parental (por ejemplo, por sustitución, delección o inserción). Una variante puede comprender una secuencia que difiere de la secuencia de polipéptido parental en hasta el 40 % del número total de residuos de la secuencia de polipéptido parental, como en hasta el 40 %, 35 %, 30 %, 25 %, 20 %, 15 %, 10 %, 9 %, 8 %, 7 %, 6 %, 5 %, 4 %, 3 %, 2 % o 1 % del número total de residuos de la secuencia de polipéptido parental. Por ejemplo, una variante de una secuencia de polipéptido de 400 aminoácidos comprende una secuencia que difiere en hasta el 40 % del número total de residuos de la secuencia de polipéptido parental, es decir, en hasta 160 posiciones de aminoácidos dentro de la secuencia de polipéptido de 400 aminoácidos (tal como en 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 78, 79, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155 o 160 posiciones de aminoácido en la secuencia de referencia).

[0012] "De origen natural", como se aplica a una composición que se puede encontrar en la naturaleza, a diferencia de ser producida artificialmente por el hombre. Por ejemplo, un polipéptido o polinucleótido que está presente en un organismo (incluidos virus, bacterias, protozoos, insectos, plantas o tejido de mamíferos) que se puede aislar a partir de una fuente en la naturaleza y que no ha sido modificado intencionalmente por el hombre en el laboratorio es de origen natural. "De origen no natural" (también denominado "sintético" o "artificial"), como se aplica a un objeto significa que el objeto no se produce de forma natural, es decir, el objeto no se puede encontrar en la naturaleza a diferencia de ser producido artificialmente por el hombre.

[0013] Un "aceite celular" o una "grasa celular" significará un aceite predominantemente triglicérido obtenido de un organismo, donde el aceite no se ha mezclado con otro aceite natural o sintético o fraccionado para alterar sustancialmente el perfil de ácidos grasos del triglicérido. En relación con un aceite que comprende triglicéridos de una regioespecificidad particular, el aceite celular o la grasa celular no se ha sometido a interesterificación u otro proceso sintético para obtener ese perfil de triglicérido regioespecífico, sino que la regioespecificidad se produce naturalmente por una célula o población de células. Para un aceite celular o una grasa celular producidos por una célula, el perfil de esteroles del aceite se determina generalmente por los esteroides producidos por la célula, no por reconstitución artificial del aceite mediante la adición de esteroides para imitar el aceite celular. En relación

con un aceite celular o una grasa celular, y como se usa generalmente en toda la presente descripción, los términos aceite y grasa se usan indistintamente, excepto donde se indique lo contrario. Por lo tanto, un "aceite" o una "grasa" puede ser líquido, sólido o parcialmente sólido a temperatura ambiente, dependiendo de la constitución de la sustancia y otras condiciones. Aquí, el término "fraccionamiento" significa eliminar material del
 5 aceitede manera que cambia su perfil de ácidos grasos relativo al perfil producido por el organismo, independientemente de cómo se realice. Los términos "aceite celular" y "grasa celular" abarcan tales aceites obtenidos de un organismo, donde el aceite se ha sometido a un procesamiento mínimo, incluyendo refinación, blanqueo y/o desgomado, que no cambia sustancialmente su perfil de triglicéridos. Un aceite celular también puede ser un "aceite celular no interesterificado", lo que significa que el aceite celular no se ha sometido a un
 10 proceso en el que los ácidos grasos se han redistribuido en sus enlaces de acilo con el glicerol y permanecen esencialmente en la misma configuración que cuando se recuperaron del organismo.

[0014] Un "perfil de ácidos grasos" es la distribución de grupos acilo grasos en los triglicéridos del aceite sin referencia a la unión a un esqueleto de glicerol. Los perfiles de ácidos grasos se determinan típicamente mediante conversión a un éster metílico de ácidos grasos (FAME), seguido de un análisis por cromatografía de gases (GC)
 15 con detección de ionización de llama (FID). El perfil de ácidos grasos se puede expresar como uno o más por ciento de un ácido graso en la señal de ácido graso total determinada del área bajo la curva para ese ácido graso. La medición FAME-GC-FID se aproxima a los porcentajes en peso de los ácidos grasos.

[0015] Las "microalgas" son organismos microbianos que contienen un cloroplasto o plástido, y opcionalmente que es capaz de realizar la fotosíntesis, o un organismo microbiano procariótico capaz de realizar la fotosíntesis.
 20 Las microalgas incluyen fotoautótrofos obligados, que no pueden metabolizar una fuente de carbono fija como energía, así como los heterótrofos, que pueden vivir solamente de una fuente de carbono fija. Las microalgas incluyen organismos unicelulares que se separan de las células hermanas poco después de la división celular, tal como *Chlamydomonas*, así como microbios tales como, por ejemplo, *Volvox*, que es un microbio fotosintético multicelular simple de dos tipos de células diferentes. Las microalgas incluyen Chlorophyceae eucariotas, tal como
 25 *Chlorella*, *Dunaliella*, y *Prototheca*. Las microalgas también incluyen otros organismos fotosintéticos microbianos que exhiben la adhesión célula-célula, tal como *Agmenellum*, *Anabaena*, y *Pyrobotrys*. Las microalgas también incluyen microorganismos heterótrofos obligados que han perdido la capacidad de realizar la fotosíntesis, tales como ciertas especies de algas dinoflageladas y especies del género *Prototheca*.

[0016] Una célula "oleaginosa" es una célula no humana capaz de producir al menos 20 % de lípidos por peso de células secas, de forma natural o mediante una mejora de la cepa recombinante o clásica. Un "microbio oleaginoso" o "microorganismo oleaginoso" es un microbio, incluida una microalga que es oleaginosa.

[0017] Como se usa con respecto a los polipéptidos o polinucleótidos, el término "aislado" se refiere a un polipéptido o polinucleótido que se ha separado de al menos otro componente que está típicamente presente con el polipéptido o polinucleótido. Por lo tanto, un polipéptido de origen natural esta aislado si se ha purificado lejos
 35 de al menos otro componente de origen natural con el polipéptido o polinucleótido. Un polipéptido o polinucleótido recombinante esta aislado si se ha purificado lejos de al menos otro componente presente cuando se produce el polipéptido o polinucleótido.

[0018] Los términos "polipéptido" y "proteína" se usan indistintamente aquí para referirse a un polímero de aminoácidos y, a menos que se limite de otro modo, incluyen aminoácidos atípicos que pueden funcionar de una manera similar a los aminoácidos de origen natural.

[0019] El término "secuencia", como se usa en relación con un polipéptido o polímero de ácido nucleico, se refiere al orden de los monómeros que constituyen el polímero o el subpolímero o el fragmento que tiene esa secuencia.

[0020] Una "subsecuencia" de una secuencia de aminoácidos o nucleótidos es una porción de una secuencia mayor o el péptido o el subpolímero de ácido nucleico o el fragmento caracterizado por la porción de la secuencia mayor.
 45

[0021] Los términos "idéntico" o "porcentaje de identidad," en el contexto de dos o más secuencias de aminoácidos o nucleótidos, se refieren a dos o más secuencias o subsecuencias que son la misma o tienen un porcentaje especificado de residuos de aminoácidos o nucleótidos que son el mismo, cuando se comparan y alinean para una correspondencia máxima, como se mide usando un algoritmo de comparación de secuencias o mediante inspección visual.
 50

[0022] Para la comparación de secuencias para determinar el porcentaje de identidad de nucleótidos o aminoácidos, típicamente una secuencia actúa como una secuencia de referencia, con la cual se comparan las secuencias de prueba. Cuando se usa un algoritmo de comparación de secuencias, las secuencias de prueba y referencia se insertan en un ordenador, se designan las coordenadas de subsecuencia, si es necesario, y se designan los parámetros del programa de algoritmo de secuencia. Posteriormente, el algoritmo de comparación
 55

de secuencias calcula el porcentaje de identidad de secuencia para la(s) secuencia(s) de prueba relativas a la secuencia de referencia, en función de los parámetros del programa designados. El alineamiento óptimo de secuencias para la comparación se puede llevar a cabo utilizando BLAST establecido con los parámetros por defecto.

- 5 [0023] Como se usa con referencia a los polipéptidos, el término "de tipo salvaje" se refiere a cualquier polipéptido que tiene una secuencia de aminoácidos presente en un polipéptido de un organismo de origen natural, independientemente de la fuente de la molécula; es decir, el término "de tipo salvaje" se refiere a características de secuencia, independientemente de si la molécula se purifica a partir de una fuente natural; se expresa recombinantemente, seguido de purificación; o se sintetiza.
- 10 [0024] El término "mutación" significará un cambio en una secuencia o subsecuencia de proteína, polipéptido o péptido producida al alterar uno o más nucleótidos en un nucleótido que codifica la proteína, polipéptido o péptido, independientemente de la alteración que se obtenga. Por ejemplo, una mutación se puede producir de forma aleatoria, por mutación por PCR, por síntesis de un gen completo o por cualquier otro método.
- 15 [0025] El término "vector" se utiliza en este caso para describir una construcción de ADN que contiene un polinucleótido. Tal vector se puede propagar de manera estable o transitoria en una célula huésped. El vector puede ser, por ejemplo, un plásmido, un vector viral o simplemente un inserto genómico potencial. Una vez introducido en un huésped adecuado, el vector puede replicarse y funcionar independientemente del genoma huésped o puede, en algunos casos, integrarse en el genoma huésped.
- 20 [0026] Como se utiliza en este caso, los términos "vector de expresión" o "construcción de expresión" o "casete de expresión" se refieren a una construcción de ácido nucleico, generada de forma recombinante o sintética, con una serie de elementos de ácido nucleico especificados que permiten la transcripción de un ácido nucleico particular en una célula huésped. El vector de expresión puede ser parte de un fragmento de plásmido, virus o ácido nucleico. Un "casete de expresión" incluye un ácido nucleico codificante (CDS) para ser transcrito operativamente unido a un promotor y a una 3'UTR. Opcionalmente, y en los ejemplos a continuación, el promotor
- 25 de un casete de expresión es un promotor heterólogo.
- [0027] El "gen exógeno" se refiere a un ácido nucleico transformado en una célula. El gen exógeno puede ser de una especie diferente (y así heteróloga) o de las mismas especies (y así homólogas), relativo a la célula que se está transformando. En el caso de un gen homólogo, este ocupa una ubicación diferente en el genoma de la célula, relativo a la copia endógena del gen. El gen exógeno puede estar presente en más de una copia en la
- 30 célula. El gen exógeno se puede mantener en una célula como una inserción en el genoma o como una molécula episómica.
- [0028] Un "promotor inducible" es uno que media la transcripción de un gen enlazado operativamente en respuesta a un estímulo particular.
- [0029] Como se utiliza en este caso, la frase "en enlace operable" se refiere a un enlace funcional entre dos
- 35 secuencias, tal como una secuencia de control (típicamente un promotor) y la secuencia enlazada. Un promotor está en enlace operable con un gen exógeno si puede mediar la transcripción del gen.
- [0030] Un "promotor" se define como un conjunto de secuencias de control de ácido nucleico que dirigen la transcripción de un ácido nucleico. Como se utiliza en este caso, un promotor incluye secuencias de ácido nucleico necesarias cerca del sitio de inicio de la transcripción, tal como, en el caso de un promotor de tipo polimerasa II,
- 40 un elemento TATA. Un promotor también incluye opcionalmente elementos intensificadores o represores distales, que se pueden situar a varios miles de pares de bases desde el sitio de inicio de la transcripción.
- [0031] Como se utiliza en este caso, el término "recombinante", cuando se usa con referencia, por ejemplo, a una célula, o un ácido nucleico, una proteína, o un vector, indica que la célula, el ácido nucleico, la proteína o el vector,
- 45 se ha modificado mediante la introducción de un ácido nucleico o una proteína exógenos o la alteración de un ácido nucleico o proteína nativos, o que la célula se deriva de una célula así modificada. Por lo tanto, las células recombinantes expresan genes que no se encuentran dentro de la forma nativa (no recombinante) de la célula o expresan genes nativos que de otro modo se expresan de manera anormal, se sobreexpresan, se subexpresan o no se expresan en absoluto. El "ácido nucleico recombinante", como se utiliza en este caso, se refiere a moléculas de ácido nucleico que se sintetizan inicialmente a través del uso de métodos de laboratorio, por lo que, de esta
- 50 manera, se crean secuencias de ácido nucleico que normalmente no se encuentran en la naturaleza. Mediante el uso de métodos de laboratorio se obtienen moléculas de ácido nucleico recombinante en enlace operable con secuencias diferentes (por ejemplo, promotor, secuencia de dirección, etc.). Por lo tanto, un ácido nucleico aislado, en forma lineal, o un vector de expresión formado *in vitro* al ligar moléculas de ADN que no se unen normalmente, se consideran recombinantes para los fines de esta invención. Se entiende que, una vez que se produce un ácido
- 55 nucleico recombinante y se reintroduce en una célula huésped o un organismo, se replicará de manera no

recombinante, es decir, usando la maquinaria celular *in vivo* de la célula huésped en vez de las manipulaciones *in vitro*; sin embargo, tales ácidos nucleicos, una vez producidos de manera recombinante, aunque posteriormente replicados de manera no recombinante, todavía se consideran recombinantes para los fines de este documento. De forma similar, una "proteína recombinante" es una proteína realizada usando técnicas recombinantes, es decir, a través de la expresión de un ácido nucleico recombinante, como se representa anteriormente.

[0032] Un "péptido de tránsito" es una secuencia de aminoácidos que dirige el tráfico de un polipéptido fusionado a la secuencia señal. En relación con las células plásticas que expresan el polipéptido, el péptido de tránsito puede dirigir el tráfico del polipéptido hacia el plástido (es decir, un péptido dirigido al plástido).

[0033] El término "polinucleótido" se refiere a un polímero desoxirribonucleótido o ribonucleótido y, a menos que se limite de otro modo, incluye análogos conocidos de nucleótidos naturales que pueden funcionar de manera similar a los nucleótidos de origen natural. El término "polinucleótido" se refiere a cualquier forma de ADN o ARN, que incluye, por ejemplo, ADN genómico; ADN complementario (ADNc), que es una representación de ADN de un ARNm, obtenida normalmente por transcripción inversa de ARN mensajero (ARNm) o amplificación; moléculas de ADN producidas sintéticamente o por amplificación; y ARNm. El término "polinucleótido" abarca moléculas de ácido nucleico bicatenarias, así como moléculas monocatenarias. En los polinucleótidos bicatenarios, las cadenas de polinucleótido no necesitan ser coextensivas (es decir, un polinucleótido bicatenario no necesita ser bicatenario a lo largo de toda la longitud de ambas cadenas).

[0034] El término "célula huésped" se refiere a una célula capaz de mantener un vector de forma transitoria o estable. Las células huésped incluyen, sin limitación, células bacterianas, células de levadura, células de insecto, células de algas (por ejemplo, células de microalgas), células vegetales y células de mamíferos. Otras células huésped conocidas en la técnica, o que se vuelven conocidas, también son adecuadas para usarse en la invención.

[0035] Como se usa en este caso, el término "complementario" se refiere a la capacidad para el emparejamiento preciso entre dos nucleótidos. Por ejemplo, si un nucleótido en una posición dada de una molécula de ácido nucleico es capaz de hibridarse con un nucleótido de otra molécula de ácido nucleico, entonces las dos moléculas de ácido nucleico se consideran complementarias entre sí en esa posición. El término "sustancialmente complementario" describe secuencias que son suficientemente complementarias entre sí para permitir la hibridación específica bajo condiciones de hibridación rigurosas. En varias formas de realización, las variantes de genes que codifican las variantes de genes FATB descritas a continuación se pueden reemplazar por un gen sustancialmente complementario que tiene actividad adecuada.

[0036] La frase "condiciones de hibridación rigurosas" generalmente se refiere a una temperatura aproximadamente 5°C inferior a la temperatura de fusión (T_m) para una secuencia específica con una fuerza iónica y un pH definidos. Las condiciones rigurosas ejemplares adecuadas para realizar la hibridación específica de la mayoría de las secuencias están a una temperatura de al menos aproximadamente 60°C y a una concentración de sal de aproximadamente 0,2 molar a pH 7,0.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0037]

La figura 1 ilustra un alineamiento de secuencias de la *Cuphea hookeriana* FATB2, en comparación con la *Cuphea avigera* FATB1 que ilustra las dos diferencias de aminoácidos entre estas tioesterasas dentro de su dominio de especificidad N-terminal;

las figuras 2A-B ilustran (A) un alineamiento de secuencias de tioesterasas FATB aisladas de genomas de *Cuphea*. La posición de la Metionina conservada relativa a la Triada Catalítica (Cys, His y Asn) y el dominio de especificidad N-terminal se destaca; y (B) una comparación de secuencias de las Cpal FATB1, Ch FATB2 y Ca FATB1 que rodea la metionina destacada. La Ca FATB1 es única debido a la presencia de una lisina en vez de la metionina; y

las figuras 3A-E ilustran histogramas de perfiles de ácidos grasos C8-C14 de aceite de microalgas con valores medios y de las medianas para múltiples transformantes de tipo salvaje y la variante de la posición 228 de *Cuphea hookeriana* FATB2 (ChFATB2), *Cuphea avigera* FATB1 (CaFATB1) que parten de predicciones basadas en datos previos de un modelo de *E. coli*.

DESCRIPCIÓN DETALLADA

Introducción

[0038] Las variantes de acil-ACP tioesterasas FatB descritas aquí permiten el control sobre la especificidad de sustrato de acil-ACP tioesterasa. Como resultado, las células huésped que expresan las acil-ACP tioesterasas producen aceites con perfiles de ácidos grasos alterados. Las células huésped que expresan las variantes de acil-ACP tioesterasas producen aceites celulares ricos en triglicéridos con perfiles de ácidos grasos caracterizados por ácidos grasos de cadena media elevados, tales como los ácidos grasos C8:0, C10:0, C12:0 y C14:0. Un ejemplo específico de la divulgación incluye proporcionar un gen de acil-ACP tioesterasa FATB, mutar el gen para alterar los aminoácidos en el producto del gen en las posiciones correspondientes a H163 y/o L186 del gen de referencia de la *Cuphea hookeriana* FATB2 (SEQ ID NO: 1). Opcionalmente, el mutante H163 y/o L186 se combina con una mutación en M228.

[0039] Como se describe con más detalle en el ejemplo 1, al expresar tales variantes de genes FATB2, integrados de forma estable en el núcleo de células plásticas oleaginosas, produjimos cepas que excedían las cepas de control que expresan ChFATB2 de tipo salvaje en términos de C8:0, C8:10 o la suma de la producción de C8:0 y C10:0, incluidas las cepas que produjeron aceites con perfiles de ácidos grasos donde la producción de C8 y C10 excedía el 9, 11, 14 o 18 % del perfil. En este último caso, el nivel de C8+C10 (es decir, la suma de la producción de C8:0 y C10:0 en el perfil de ácidos grasos, como se determina por FAME-GC con detección FID) fue más del doble, relativo a aproximadamente el 8 % de C8+C10 de la cepa ChFATb2 de tipo salvaje. Las variantes específicas con la producción mejorada de C8 + C10 incluyen aquellas con P, K o A en la posición 186; Y o F en la posición 163, o combinaciones de las mismas, tales como 186P/163Y, 186P/163F, 186K/163Y, 186K/163F, 186A/163Y o 186A/163F. De los mutantes dobles encontramos que la variante H163Y/L186P producía un aceite que tiene concentraciones particularmente altas de C8+C10. Usando variantes simples o dobles, los porcentajes de perfil de ácidos grasos C8:0 se pueden aumentar en un 50, 60, 70, 80, 100 % o más relativos a una cepa de control que expresa ChFATB2 de tipo salvaje; por ejemplo, hasta más del 2; 2,5; 3 o 3,5 % del perfil de ácidos grasos vs. 1,5 % para el control (véase el ejemplo 1).

[0040] Los mutantes dobles enumerados anteriormente también se pueden combinar con una tercera mutación que corresponde a 230 de *Cuphea palustris* FATB1. Para muchos genes FATB, tal como *Cuphea hookeriana* FATB2 y *Cuphea avigera* FATB1, este residuo corresponde al residuo 228. Por ejemplo, una mutación M228K en *Cuphea hookeriana* FATB2 expresada en una microalga eucariótica oleaginosa aumentó la proporción C8/C10 en el perfil de ácidos grasos del aceite de aproximadamente 0,25 a aproximadamente 1,0. Las mutaciones en esta posición a Iso, Val, Phe, y Leu, Ala, o Thr en combinación con los mutantes simples o dobles en las posiciones 186 y 163 mencionados anteriormente también pueden ser ventajosas.

[0041] Aunque *Cuphea hookeriana* FATB2 se usó como un sistema modelo, los métodos para realizar las mutaciones mencionadas anteriormente, los métodos para expresarlas en una célula oleaginosa y los métodos para producir aceite con estas variantes se pueden aplicar para cualquier gen de acil-ACP tioesterasa, incluidos los que tienen al menos 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99% de identidad con SEQ ID NO:1, o el fragmento de SEQ ID NO: 1 que carece del péptido de tránsito.

[0042] Aunque estas variantes de genes se descubrieron usando un sistema de expresión de microalgas eucariotas, los genes son generalmente más útiles en formas que se conocen en la técnica, incluida su expresión en plantas superiores para producir aceites de triglicéridos alterados. Cuando se incorporan a una célula oleaginosa (por ejemplo, de una planta oleaginosa, algas (por ejemplo, microalgas)), las variantes de tioesterasas pueden alterar los perfiles de ácidos grasos de la célula para producir productos comerciales de alto valor nuevos o más económicos.

[0043] Los mutantes simples, dobles o triples se pueden usar para producir un aceite con una alta proporción de ácidos grasos de C8:0 a C10:0. Por ejemplo, la proporción C8/10 puede ser igual o mayor de 0,3; 0,5; 0,7; 0,9; 1,0; 1,2; 1,4; 1,6; 1,8; 2,0; 2,2; 2,4; 2,6; 2,8; o 3,0.

[0044] La divulgación también proporciona la biomasa residual de tales células después de la extracción de aceite, oleoquímicos, combustibles y productos alimenticios elaborados a partir de los aceites y métodos de cultivo de las células. En formas de realización variables, las células son células de microalgas, que incluyen células heterótrofas u heterótrofas obligadas, y células clasificadas como *Chlorophyta*, *Trebouxiphyceae*, *Chlorellales*, *Chlorellaceae*, o *Chlorophyceae*. Las células también pueden ser células vegetales o células de macroalgas. Se prefieren las células huésped que tienen una vía de síntesis de ácidos grasos tipo II. Aunque los ejemplos que se dan a continuación usan el *Trebouxiphyte Prototheca moriformis* como una célula huésped, los genes, las construcciones y los métodos descritos también pueden encontrar uso en cultivos de semillas oleaginosas. Los métodos para introducir estos genes en tales cultivos, tales como soja, maíz, colza, alazor, girasol y otros se conocen en la técnica; véanse, por ejemplo, las Patentes de EE.UU. N^{os}. 6331664, 5512482, 5455167, 5667997. Los ejemplos de oleoquímicos incluyen surfactantes y solventes hechos de ácidos grasos o aceites.

[0045] Por consiguiente, en un ejemplo de la divulgación, se proporciona una proteína no natural, un gen aislado que codifica la proteína no natural, un casete de expresión que expresa la proteína no natural, o una célula

huésped que comprende el casete de expresión, donde la proteína no natural tiene al menos el 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad con SEQ ID NO: 1 y comprende Y o F en la posición correspondiente a la posición 163 de SEQ ID NO: 1 y/o P, K, o A en la posición correspondiente a la posición 186 de SEQ ID NO: 1, y opcionalmente K en la posición correspondiente a la posición 228 de SEQ ID NO: 1.

[0046] En un ejemplo relacionado de la divulgación existe un método para producir un aceite de triglicéridos. El método incluye expresar, en una célula huésped, la proteína del mencionado inmediatamente antes, o una proteína que comprende al menos el 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de una de las SEQ ID NOs: 3-8 que tiene Y o F en la posición correspondiente a la posición 163 de SEQ ID NO: 1 y/o P, K, o A en la posición correspondiente a la posición 186 de SEQ ID NO: 1, y opcionalmente K en la posición correspondiente a la posición 228 de SEQ ID NO: 1. El método también incluye cultivar la célula huésped y aislar el aceite.

[0047] En otro ejemplo de la divulgación se proporciona un método para aumentar los ácidos grasos C8 y/o C10 en un perfil de ácidos grasos de un aceite producido por una célula huésped opcionalmente oleaginosa. El método incluye proporcionar un gen parental que codifique una enzima FATB, mutar el gen para tener Y o F en la posición correspondiente a la posición 163 de SEQ ID NO: 1 y/o P, K, o A en la posición correspondiente a la posición 186 de SEQ ID NO: 1, y opcionalmente K en la posición correspondiente a la posición 228 de SEQ ID NO: 1. El método también incluye expresar el gen mutado en la célula huésped y producir el aceite. El perfil de ácidos grasos del aceite aumenta, de este modo, en los ácidos grasos C8 y/o C10, relativo al gen parental. Opcionalmente, el gen que codifica la enzima FATB codifica una proteína con al menos el 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad con SEQ ID NO: 1, 13 o 14.

[0048] Como se detalla en el ejemplo 3, en comparación con el trabajo de la técnica anterior en *E. coli*, el descubrimiento de la ventaja de usar Ala o Thr en la posición 230 de Cpal FATB1 (SEQ ID NO: 13) en términos de la producción de C8+C10 y/o de la proporción C8/C10 aumentada, es nuevo e inesperado. Estas mutaciones nuevas son útiles solas, en combinación con una mutación en la posición 163, que incluye las mutaciones favorecedoras de C8 descritas aquí, en combinación con una mutación en la posición 186, que incluye las mutaciones favorecedoras de C8 descritas aquí, o en combinación con una mutación doble en las posiciones 163 y 186, que incluye las mutaciones favorecedoras de C8 descritas aquí. Por consiguiente, en un ejemplo de la divulgación, existe una proteína no natural, un gen aislado que codifica la proteína no natural, un casete de expresión que expresa la proteína no natural, o una célula huésped que comprende el casete de expresión, donde la proteína no natural tiene al menos el 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad con SEQ ID NO: 13 y A o K en la posición correspondiente a la posición 230 de SEQ ID NO: 13. Un método para producir un aceite incluye expresar, en una célula huésped, las proteínas no naturales descritas aquí, cultivar la célula y aislar el aceite.

Variantes de Acil-ACP Tioesterasas

[0049] Las variantes de TEs se pueden usar en construcciones genéticas y células oleaginosas genéticamente modificadas (por ejemplo, plantas, algas, microalgas) con uno o más genes exógenos para producir ácidos grasos, acilglicéridos, o derivados de los mismos. Por ejemplo, las microalgas o los cultivos de semillas oleaginosas que, naturalmente, o mediante modificación genética, producirían altos niveles de lípidos, se pueden diseñar (o diseñarse aun más) para expresar una variante exógena de acil graso-ACP tioesterasa, que puede facilitar la escisión de los ácidos grasos de la proteína portadora de acilos (ACP) durante la síntesis de ácidos grasos. Los ácidos grasos sintetizados se pueden incorporar a los glicéridos de acilo, incluidos los triacilglicéridos (TAGs, triglicéridos). Los TAGs pueden recuperarse o, mediante un procesamiento enzimático adicional dentro de la célula, o *in vitro*, producir otros compuestos útiles.

[0050] Las variantes de acil graso-ACP tioesterasas están diseñadas en base a la especificidad deseada para un grupo acilo graso en crecimiento (durante la síntesis de ácidos grasos) que tiene una longitud de cadena de carbono particular. Un dominio de especificidad se selecciona según su preferencia por un sustrato de acilo graso ACP particular y/o por su capacidad para influir, aumentar y/o promover la producción de ácidos grasos de una longitud de cadena de carbono deseada. Generalmente, las variantes de acil graso-ACP tioesterasas tienen una especificidad de sustrato preferencial por sustratos de acilo graso ACP de cadena media (por ejemplo, para liberar los ácidos grasos C8, C10, C12 y/o C14). En diferentes ejemplos de la divulgación, el dominio de especificidad en el extremo N-terminal de la acil-ACP tioesterasa es heterólogo (por ejemplo, debido a mutaciones puntuales y/o intercambio de dominios) al dominio catalítico del extremo C-terminal. En determinados ejemplos, la especificidad del sustrato de longitud de cadena de ácidos grasos y/o la preferencia del dominio de especificidad y el dominio catalítico es la misma o está dentro de 1-2 carbonos.

Optimización de codones para expresión

[0051] El ADN que codifica un polipéptido para expresarse en un microorganismo, por ejemplo, una variante de acil-ACP tioesterasa y un marcador seleccionable puede ser ADNc con optimización de codones. Los métodos para recodificar genes para la expresión en microalgas se describen en la patente de EE.UU. N.º 7,135,290. La información adicional para la optimización de codones está disponible, por ejemplo, en la base de datos Codon Usage Database en kazusa.o.jp/codon/. La tabla para el uso de codones preferidos de *Prototheca* también se proporciona en la Publicación de Patente de EE.UU. N.º 2012/0283460, Tabla 1.

Expresión y direccionamiento a plástidos

[0052] Las proteínas expresadas en el genoma nuclear de *Prototheca* se puede dirigir al plástido usando señales de direccionamiento de plástidos. Se conocen secuencias de direccionamiento de plástidos endógenas a *Chlorella*, tales como genes en el genoma nuclear de *Chlorella* que codifican proteínas que se dirigen al plástido; véanse, por ejemplo, los números de acceso de GenBank AY646197 y AF499684, y en una forma de realización, tales secuencias de control se usan en los vectores descritos aquí, por ejemplo, para dirigir la expresión de una proteína a un plástido de *Prototheca*.

[0053] Los ejemplos siguientes describen el uso de secuencias de direccionamiento de plástidos de algas para dirigir proteínas heterólogas al compartimento correcto en la célula huésped. Bibliotecas de ADNc fueron hechas usando células de *Prototheca moriformis* y *Chlorella protothecoides* y se describen en los ejemplos de Publicación de Patente de EE.UU. N.º 2012/0283460 y en la Solicitud de PCT N.º PCT/US2009/066142. Las secuencias de aminoácidos de las secuencias de direccionamiento de plástidos de algas identificadas a partir del direccionamiento de plástidos útil de las bibliotecas de ADNc de variantes de acil-ACP tioesterasas expresadas de forma recombinante se proporcionan en la Publicación de Patente de EE.UU. N.º 2012/0283460 y aquí. En diversas formas de realización, el péptido de tránsito de plástido comprende una secuencia de aminoácidos seleccionada de entre el grupo que consiste en

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRA, SGPRRPARPLPVR,
SGPRRPARPLPVRAAIASEVPVATTSPR, RPARPLPVRGRA,
RPARPLPVRAAIASEVPVATTSPR, RCGDLRRSAGSGPRRPARPLPVRGRA,
RCGDLRRSAGSGPRRPARPLPVRAAIASEVPVATTSPR, PARPLPVR,
PARPLPVRAAIASEVPVATTSPR, RPARPLPVR, and
RRPARPLPVRAAIASEVPVATTSPR.

[0054] Cuando las variantes de FATB nuevas se describan aquí, se entenderá que se puede usar una variedad de péptidos de tránsito de plástidos heterólogos. En otras palabras, el dominio de péptido no dirigido está más altamente conservado. Por consiguiente, las formas de realización descritas aquí presentan el dominio enzimático FATB nuevo con o sin una secuencia de direccionamiento de plástidos. Por ejemplo, cuando aquí se da una identidad de porcentaje a un nuevo gen FATB, la misma identidad se puede aplicar (cuando se especifique) a la misma secuencia en ausencia del péptido de direccionamiento. Un péptido de direccionamiento sustituto se puede usar opcionalmente en relación con tal secuencia.

Células huésped - Microorganismos productores de aceites o lípidos

[0055] Se puede usar cualquier especie de organismo que produzca lípidos y/o hidrocarburos adecuados, aunque se prefieren los microorganismos que producen naturalmente altos niveles de lípidos y/o hidrocarburos adecuados. La producción de hidrocarburos por microorganismos es revisada por Metzger et al. Appl Microbiol Biotechnol (2005) 66: 486-496 y A Look Back at the U.S. Department of Energy's Aquatic Species Program: Biodiesel from Algae, NREUTP-580-24190, John Sheehan, Terri Dunahay, John Benemann y Paul Roessler (1998).

[0056] Las consideraciones para la selección de microorganismos incluyen, además de la producción de lípidos o hidrocarburos adecuados para producir aceites, combustibles y oleoquímicos: (1) alto contenido de lípidos como un porcentaje de peso celular; (2) facilidad de crecimiento; (3) facilidad de ingeniería genética; y (4) facilidad de procesamiento de biomasa. En formas de realización particulares, el microorganismo de tipo salvaje o modificado genéticamente produce células que son al menos el 40 %, al menos el 45 %, al menos el 50 %, al menos el 55 %, al menos el 60 %, al menos el 65 %, o al menos el 70 % o más de lípidos. Los organismos preferidos crecen heterotróficamente (en azúcares en ausencia de luz) o se pueden diseñar para que lo hagan usando, por ejemplo, métodos descritos aquí. La facilidad de transformación y la disponibilidad de marcadores y promotores seleccionables, constitutivos o inducibles, que son funcionales en el microorganismo afectan a la facilidad de

ingeniería genética. Las consideraciones de procesamiento pueden incluir, por ejemplo, la disponibilidad de medios eficaces para lisar las células.

A. Algas

[0057] En una forma de realización, el microorganismo es una microalga. Los ejemplos no limitativos de microalgas que se pueden usar para la expresión de variantes de acil-ACP tioesterasas incluyen, por ejemplo, *Achnanthes orientalis*, *Agmenellum*, *Amphiprora hyaline*, *Amphora coffeiformis*, *Amphora coffeiformis lineae*, *Amphora coffeiformis punctata*, *Amphora coffeiformis taylori*, *Amphora coffeiformis tenuis*, *Amphora delicatissima*, *Amphora delicatissima capitata*, *Amphora* sp., *Anabaena*, *Ankistrodesmus*, *Ankistrodesmus falcatus*, *Boekelovia hooglandii*, *Borodinella* sp., *Botryococcus braunii*, *Botryococcus sudeticus*, *Bracteococcus minor*, *Bracteococcus medionucleatus*, *Carteria*, *Chaetoceros gracilis*, *Chaetoceros muelleri*, *Chaetoceros muelleri subsalsum*, *Chaetoceros* sp., *Chlorella anitrata*, *Chlorella Antarctica*, *Chlorella aureoviridis*, *Chlorella candida*, *Chlorella capsulate*, *Chlorella desiccata*, *Chlorella ellipsoidea*, *Chlorella emersonii*, *Chlorella fusca*, *Chlorella fusca* var. *vacuolata*, *Chlorella glucotropha*, *Chlorella infusionum*, *Chlorella infusionum* var. *actophila*, *Chlorella infusionum* var. *auxenophila*, *Chlorella kessleri*, *Chlorella lobophora* (cepa SAG 37.88), *Chlorella luteoviridis*, *Chlorella luteoviridis* var. *aureoviridis*, *Chlorella luteoviridis* var. *lutescens*, *Chlorella miniata*, *Chlorella minutissima*, *Chlorella mutabilis*, *Chlorella nocturna*, *Chlorella ovalis*, *Chlorella parva*, *Chlorella photophila*, *Chlorella pringsheimii*, *Chlorella protothecoides* (incluida cualquiera de las cepas UTEX 1806, 411, 264, 256, 255, 250, 249, 31, 29, 25), *Chlorella protothecoides* var. *acidicola*, *Chlorella regularis*, *Chlorella regularis* var. *minima*, *Chlorella regularis* var. *umbricata*, *Chlorella reisi*, *Chlorella saccharophila*, *Chlorella saccharophila* var. *ellipsoidea*, *Chlorella salina*, *Chlorella simplex*, *Chlorella sorokiniana*, *Chlorella* sp., *Chlorella sphaerica*, *Chlorella stigmatophora*, *Chlorella vanniellii*, *Chlorella vulgaris*, *Chlorella vulgaris* f. *tertia*, *Chlorella vulgaris* var. *autotrophica*, *Chlorella vulgaris* var. *viridis*, *Chlorella vulgaris* var. *vulgaris*, *Chlorella vulgaris* var. *vulgaris* f. *tertia*, *Chlorella vulgaris* var. *vulgaris* f. *viridis*, *Chlorella xanthella*, *Chlorella zofingiensis*, *Chlorella trebouxii*, *Chlorella vulgaris*, *Chlorococcum infusionum*, *Chlorococcum* sp., *Chlorogonium*, *Chroomonas* sp., *Chrysosphaera* sp., *Cricosphaera* sp., *Cryptocodinium cohnii*, *Cryptomonas* sp., *Cyclotella cryptica*, *Cyclotella meneghiniana*, *Cyclotella* sp., *Dunaliella* sp., *Dunaliella bardawil*, *Dunaliella bioculata*, *Dunaliella granulate*, *Dunaliella maritime*, *Dunaliella minuta*, *Dunaliella parva*, *Dunaliella peircei*, *Dunaliella primolecta*, *Dunaliella salina*, *Dunaliella terricola*, *Dunaliella tertiolecta*, *Dunaliella viridis*, *Dunaliella tertiolecta*, *Eremosphaera viridis*, *Eremosphaera* sp., *Ellipsoidon* sp., *Euglena*, *Franceia* sp., *Fragilaria crotonensis*, *Fragilaria* sp., *Gleocapsa* sp., *Gloeothamnion* sp., *Hymenomonas* sp., *Isochrysis* aff. *galbana*, *Isochrysis galbana*, *Lepocinclis*, *Micractinium*, *Micractinium* (UTEX LB 2614), *Monoraphidium minutum*, *Monoraphidium* sp., *Nannochloris* sp., *Nannochloropsis salina*, *Nannochloropsis* sp., *Navicula acceptata*, *Navicula biskanterae*, *Navicula pseudotenelloides*, *Navicula pelliculosa*, *Navicula saprophila*, *Navicula* sp., *Nephrochloris* sp., *Nephroselmis* sp., *Nitzschia communis*, *Nitzschia alexandrina*, *Nitzschia communis*, *Nitzschia dissipata*, *Nitzschia frustulum*, *Nitzschia hantzschiana*, *Nitzschia inconspicua*, *Nitzschia intermedia*, *Nitzschia microcephala*, *Nitzschia pusilla*, *Nitzschia pusilla elliptica*, *Nitzschia pusilla monoensis*, *Nitzschia quadrangular*, *Nitzschia* sp., *Ochromonas* sp., *Oocystis parva*, *Oocystis pusilla*, *Oocystis* sp., *Oscillatoria limnetica*, *Oscillatoria* sp., *Oscillatoria subbrevis*, *ParaChlorella kessleri*, *Pascheria acidophila*, *Pavlova* sp., *Phagus*, *Phormidium*, *Platymonas* sp., *Pleurochrysis carterae*, *Pleurochrysis dentate*, *Pleurochrysis* sp., *Prototheca wickerhamii*, *Prototheca stagnora*, *Prototheca portoricensis*, *Prototheca moriformis*, *Prototheca zopfii*, *PseudoChlorella aquatica*, *Pyramimonas* sp., *Pyrobotrys*, *Rhodococcus opacus*, *Sarcinoid chrysophyte*, *Scenedesmus armatus*, *Schizochytrium*, *Spirogyra*, *Spirulina platensis*, *Stichococcus* sp., *Synechococcus* sp., *Tetradron*, *Tetraselmis* sp., *Tetraselmis suecica*, *Thalassiosira weissflogii*, y *Viridiella fridericana*.

[0058] Las células huésped ilustrativas presentan células oleaginosas que producen perfiles de ácidos grasos alterados y/o una distribución regioespecífica alterada de ácidos grasos en glicerolípidos, y productos producidos de las células. Los ejemplos de células oleaginosas incluyen células microbianas que tienen una vía de biosíntesis de lípidos tipo II, incluidas las células oleaginosas plastídicas, tales como las de las algas oleaginosas. Los ejemplos específicos de células incluyen microalgas heterótrofas o eucariotas heterótrofas obligadas del filo *Chlorophyta*, la clase *Trebouxiophytae*, el orden *Chlorellales*, o la familia *Chlorellaceae*. Se proporcionan ejemplos de microalgas oleaginosas en las Solicitudes de Patente PCT publicadas WO2008/151149, WO2010/063032, WO2011/150410, y WO2011/150411, incluidas las especies de *Chlorella* y *Prototheca*, un género que comprende heterótrofos obligados. Las células oleaginosas pueden ser, por ejemplo, capaces de producir el 25 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 85 %, o aproximadamente el 90 % de aceite por peso celular, ± 5 %. Las publicaciones mencionadas anteriores también describen métodos para cultivar tales células y extraer aceite, especialmente de células de microalgas; tales métodos son aplicables a las células descritas aquí. En cualquiera de las formas de realización descritas aquí, las células pueden ser células heterótrofas que comprenden un gen de invertasa exógena para permitir que las células produzcan aceite a partir de una materia prima de sacarosa.

[0059] Las formas de realización ilustrativas de células huésped incluyen células oleaginosas recombinantes que expresan uno o más genes exógenos que codifican enzimas de biosíntesis de ácidos grasos, como se menciona en las reivindicaciones. Como resultado, algunas formas de realización presentan aceites celulares nunca antes obtenibles en un aceite celular. En algunos casos, los aceites celulares no eran obtenibles a partir de un aceite no vegetal o sin semillas, o no eran obtenibles en absoluto.

[0060] Las células oleaginosas producen un aceite de almacenamiento, que se puede almacenar en vesículas de almacenamiento de la célula. Se puede obtener un aceite celular crudo a partir de las células, al romper las células y aislar el aceite. Los aceites producidos se pueden refinar, blanquear y desodorizar (RBD), como se conoce en la técnica o como se describe en la WO2010/120939. Los aceites crudos o RBD se pueden usar en una variedad de productos o procesos alimenticios, químicos e industriales. Después de la recuperación del aceite permanece una biomasa residual valiosa. Los usos para la biomasa residual incluyen la producción de papel, plásticos, absorbentes, adsorbentes, como pienso para animales, para nutrición humana, o para fertilizante.

[0061] Cuando se proporciona un perfil de ácidos grasos de un aceite celular de triglicéridos, se entiende que este se refiere a una muestra no fraccionada del aceite de almacenamiento extraído de la célula analizada en condiciones en las que los fosfolípidos se han eliminado o con un método de análisis que es sustancialmente insensible a los ácidos grasos de los fosfolípidos (por ejemplo, usando cromatografía y espectrometría de masas). Dado que las células son oleaginosas, en algunos casos el aceite de almacenamiento constituirá la mayor parte de todos los TAGs en la célula.

[0062] En formas de realización variables, la célula huésped es una célula plastídica, por ejemplo, una microalga heterotrófica del filo *Chlorophyta*, la clase *Trebouxiophytae*, el orden *Chlorellales*, o la familia *Chlorellaceae*. En diversas formas de realización, la célula es oleaginosa y capaz de acumular al menos un 40 % de aceite por peso de células secas. La célula puede ser un heterótrofo obligado, tal como una especie de *Prototheca*, incluida *Prototheca moriformis* o *Prototheca zopfii*. El ácido nucleico que codifica las variantes de acil-ACP TEs descritas aquí también se puede expresar en algas o plantas autótrofas. Opcionalmente, la célula es capaz de usar sacarosa para producir aceite y se puede introducir un gen de invertasa recombinante para permitir el metabolismo de sacarosa, como se describe en las publicaciones PCT WO2008/151149, WO2010/063032, WO2011/150410, WO2011/150411, y la solicitud de patente internacional PCT/US12/23696. La invertasa puede someterse a optimización de codones e integrarse en un cromosoma de la célula, al igual que todos los genes mencionados aquí. El uso de codones para diferentes especies de algas y plantas de interés se conoce en la técnica y se puede encontrar, por ejemplo, en internet en la base de datos Codon Usage Database en kazusa.o.jp/codon/.

[0063] Los polinucleótidos que codifican las variantes de acil-ACP TEs descritas aquí también se pueden expresar en una amplia variedad de células huésped vegetales. De particular interés son las células vegetales de plantas implicadas en la producción de aceites vegetales para usos comestibles e industriales, que incluyen, por ejemplo, cultivos templados de semillas oleaginosas. Las plantas de interés incluyen, pero no de forma limitativa, semillas de uva (variedades de canola y alto contenido de ácido erúico), girasol, alazor, algodón, *Cuphea*, soja, cacahuete, coco y aceite de palma, y maíz. Véanse las Patentes de EE.UU. N.ºs: 5,850,022; 5,723,761; 5,639,790; 5,807,893; 5,455,167; 5,654,495; 5,512,482; 5,298,421; 5,667,997; y 5,344,771; 5,304,481.

Aceites con perfiles de ácidos grasos de origen no natural

[0064] Los aceites descritos aquí son diferentes de otros aceites de origen natural que tienen alto contenido de ácidos grasos de cadena media, tal como el aceite de palma, el aceite de semilla de palma, y el aceite de coco. Por ejemplo, los niveles de contaminantes, tales como los carotenoides, son mucho más altos en el aceite de palma y el aceite de semilla de palma que en los aceites descritos aquí. Los aceites de palma y semilla de palma contienen, en particular, alfa y beta carotenos y licopeno en cantidades mucho más altas que en los aceites descritos aquí. Además, más de 20 carotenoides diferentes se encuentran en el aceite de palma y de semilla de palma, mientras que los ejemplos demuestran que los aceites descritos aquí contienen muy pocas especies de carotenoides y niveles muy bajos. Asimismo, los niveles de compuestos de vitamina E, tales como los tocotrienoles, son mucho mayores en el aceite de palma, semilla de palma, y coco que en los aceites descritos aquí.

[0065] Generalmente, las cepas de *Prototheca* tienen muy pocos o ningún ácido graso con la longitud de cadena C8-C14. Por ejemplo, las cepas de *Prototheca moriformis* (UTEX 1435), *Prototheca krugani* (UTEX 329), *Prototheca stagnora* (UTEX 1442) y *Prototheca zopfii* (UTEX 1438) no producen (o cantidades indetectables) ácidos grasos C8, entre el 0-0,01 % de ácidos grasos C10, entre el 0,03-2,1 % de ácidos grasos C12 y entre el 1,0-1,7% de ácidos grasos C14.

[0066] En algunos casos, las células oleaginosas (por ejemplo, las cepas de *Prototheca*) que contienen un transgén que codifica una variante de acil graso-ACP tioesterasa tiene un perfil de ácidos grasos caracterizado por el 5-10 %, 10-20 %, 20-30 %, 30-40 %, 40-50 %, 50-60%, 60-70 %, 70-80 %, 80-90 %, o 90-99 % de ácidos grasos C8, C10, C12, o C14. En otros casos, las cepas de *Prototheca* que contienen un transgén que codifica una acil graso-ACP tioesterasa que tiene actividad hacia sustratos de acil graso-ACP de longitud de cadena C12 y C14 y produce ácidos grasos de longitud de cadena C12 y la longitud de cadena C14 en una proporción de 1:1 +/- 20 %.

[0067] En algunos casos, mantener las cepas de *Prototheca* transgénicas bajo presión constante y selectiva alta para retener genes exógenos es ventajoso debido al aumento en el ácido graso deseado de una longitud de cadena específica. También se pueden lograr altos niveles de retención de genes exógenos mediante la inserción de genes exógenos en los cromosomas nucleares de las células usando métodos y vectores de recombinación homóloga descritos aquí. También se contemplan células recombinantes que contienen genes exógenos integrados en cromosomas nucleares.

[0068] El aceite de microalgas también puede incluir otros constituyentes producidos por las microalgas, o incorporados en el aceite de microalgas del medio de cultivo. Estos otros constituyentes pueden estar presentes en cantidades variables, dependiendo de las condiciones de cultivo usadas para cultivar las microalgas, las especies de microalgas, el método de extracción usado para recuperar el aceite de microalgas de la biomasa y otros factores que pueden afectar a la composición de aceite de microalgas. Ejemplos no limitativos de tales constituyentes incluyen carotenoides, presentes de 0,1-0,4 microgramos/ml, clorofila presente de 0-0,02 miligramos/kilogramo de aceite, gamma-tocoferol presente de 0,4-0,6 miligramos/100 gramos de aceite, y tocotrienoles totales presentes de 0,2-0,5 miligramos/gramo de aceite.

[0069] Los otros constituyentes pueden incluir, sin limitación, fosfolípidos, tocoferoles, tocotrienoles, carotenoides (por ejemplo, alfa-caroteno, beta-caroteno, licopeno, etc.), xantófilas (por ejemplo, luteína, zeaxantina, alfa-criptoxantina y beta-criptoxantina), y varios compuestos orgánicos o inorgánicos.

[0070] En algunos casos, el aceite extraído de las especies de *Prototheca* comprende no más de 0,02 mg/kg de clorofila. En algunos casos, el aceite extraído de las especies de *Prototheca* comprende no más de 0,4 mcg/ml de carotenoides totales. En algunos casos, el aceite de *Prototheca* comprende entre 0,40-0,60 miligramos de gamma-tocoferol por 100 gramos de aceite. En otros casos, el aceite de *Prototheca* comprende entre 0,2-0,5 miligramos de tocotrienoles totales por gramo de aceite.

[0071] Los aceites producidos de células huésped que expresan una variante de acil-ACP tioesterasa tendrá un perfil isotópico que lo distinga, por ejemplo, de aceites mezclados de otras fuentes. El valor de isótopo de carbono estable $\delta^{13}\text{C}$ es una expresión de la proporción de $^{13}\text{C}/^{12}\text{C}$ relativa a un estándar (por ejemplo PDB, carbonita del esqueleto fósil de Belemnita americana de la formación Peedee de Carolina del Sur). El valor de isótopo de carbono estable $\delta^{13}\text{C}$ (0/00) de los aceites se pueden relacionar con el valor $\delta^{13}\text{C}$ de la materia prima usada. En algunas formas de realización, los aceites se derivan de organismos oleaginosos cultivados heterotróficamente en azúcar derivado de una planta C4, tal como maíz o caña de azúcar. En algunas formas de realización, el $\delta^{13}\text{C}$ (0/00) del aceite es de 10 a -17 0/00 o de 13 a -16 0/00.

[0072] En diversos ejemplos de la divulgación, una célula huésped que expresa una variante de acil-ACP tioesterasa que comprende todos o los residuos determinantes de especificidad de un dominio de especificidad de una acil-ACP tioesterasa que prefiere C10 (por ejemplo, una acil-ACP tioesterasa de *Cuphea hookeriana*), y un dominio catalítico de una acil-ACP tioesterasa que prefiere C12 (por ejemplo, una acil-ACP tioesterasa de *Cuphea wrightii* o *Umbellularia californica*) produce un aceite que comprende al menos aproximadamente el 10 % de ácidos grasos C12:0, y al menos aproximadamente el 10 % de ácidos grasos C14:0.

[0073] En diferentes ejemplos de la divulgación, una célula huésped que expresa una variante de acil-ACP tioesterasa que comprende todos o los residuos determinantes de especificidad de un dominio de especificidad modificado de una primera acil-ACP tioesterasa que tiene una o ambas sustituciones His163→Tyr o Leu186→Pro (o en posiciones correspondientes a His163→Tyr o Leu186→Pro de SEQ ID NO: 61), y un dominio catalítico de una segunda acil-ACP tioesterasa produce un aceite que comprende al menos aproximadamente el 5 %, por ejemplo, al menos aproximadamente el 6 %, 7 %, 8 %, 9 %, 10 %, 12 %, 15 %, o más, de ácidos grasos C8:0 o al menos aproximadamente el 5 %, por ejemplo, al menos aproximadamente el 6 %, 7 %, 8 %, 9 %, 10 %, 12 %, 15 %, o más, de ácidos grasos C10:0 o una proporción de C8:0/C10:0 que es al menos aproximadamente el 5 %, por ejemplo, al menos aproximadamente el 6 %, 7 %, 8 %, 9 %, 10 %, 12 %, 15 %, o más. Según sea apropiado, el dominio de especificidad se puede derivar de una acil-ACP tioesterasa que prefiere C8:0, C10:0 o C12:0 e independientemente el dominio catalítico se puede derivar de una acil-ACP tioesterasa que prefiere C8:0, C10:0 o C12:0. El dominio de especificidad y el dominio catalítico puede ser de la misma o diferentes acil-ACP tioesterasas. En diversas formas de realización, una célula huésped que expresa una variante de acil-ACP tioesterasa que comprende todos o los residuos determinantes de especificidad de un dominio de especificidad modificado de una acil-ACP tioesterasa que prefiere C10 (por ejemplo, una acil-ACP tioesterasa de *Cuphea hookeriana* que tiene una o ambas sustituciones His163→Tyr o Leu186→Pro), y un dominio catalítico de una acil-ACP tioesterasa que prefiere C10 (por ejemplo, una acil-ACP tioesterasa de *Cuphea hookeriana*) produce un aceite que comprende al menos aproximadamente el 5 %, por ejemplo, al menos aproximadamente el 6 %, 7 %, 8 %, 9 %, 10 %, 12 %, 15 %, o más, de ácidos grasos C8:0 o al menos aproximadamente el 5 %, por ejemplo, al menos aproximadamente el 6 %, 7 %, 8 %, 9 %, 10 %, 12 %, 15 %, o más, de ácidos grasos C10:0 o una proporción de C8:0/C10:0 que es al menos aproximadamente el 5 %, por ejemplo, al menos aproximadamente el 6 %, 7 %, 8 %, 9 %, 10 %, 12 %, 15 %, o más.

[0074] En diversos ejemplos de la divulgación, una célula huésped que expresa una variante de acil-ACP tioesterasa que comprende todos o los residuos determinantes de especificidad de un dominio de especificidad de una acil-ACP tioesterasa que prefiere C14 (por ejemplo, una acil-ACP tioesterasa de *Cinnamomum camforum*), y un dominio catalítico de una acil-ACP tioesterasa que prefiere C12 (por ejemplo, una acil-ACP tioesterasa de *Cuphea wrightii* o *Umbellularia californica*) produce un aceite que comprende ácidos grasos C12:0 y ácido graso C14:0 en una proporción aproximada de 1:1; por ejemplo, una proporción de 1:1 +/- 20 %.

[0075] Además, las células huésped que expresan una variante de acil-ACP tioesterasa que comprende 5 o más residuos de aminoácidos que se extienden desde el extremo C-terminal de un dominio conector posicionado en el extremo N-terminal hasta el dominio hidrofóbico, producen un aceite que comprende ácidos grasos de longitud de cadena media relativamente elevados (por ejemplo, C8:0, C10:0, C12:0, C14:0) en comparación con las células huésped que expresan la misma acil-ACP tioesterasa sin un dominio conector. En diferentes ejemplos de la divulgación, las células huésped que expresan una variante de acil-ACP tioesterasa que comprende 5 o más residuos de aminoácidos que se extienden desde el extremo C-terminal de un dominio conector posicionado en el extremo N-terminal hasta el dominio hidrofóbico, producen un aceite que comprende ácidos grasos de longitud de cadena media aumentados al menos 1 vez, 2 veces, 3 veces, o más, en comparación con las células huésped que expresan la misma acil-ACP tioesterasa sin un dominio conector.

[0076] En un ejemplo específico de la divulgación, una célula recombinante comprende ácidos nucleicos operables para expresar un producto de un gen exógeno que codifica un gen exógeno de una variante de acil-ACP tioesterasa que codifica una acil-ACP tioesterasa activa que cataliza la escisión de ácidos grasos de cadena media de ACP. Como resultado, en un ejemplo, el aceite producido se puede caracterizar por un perfil de ácidos grasos elevado en los ácidos grasos C8, C10, C12 y/o C14 y reducido en los ácidos grasos C16, C18, y C18:1 como resultado de expresión de los ácidos nucleicos recombinantes. En varios ejemplos, el aumento en los ácidos grasos C8, C10, C12 y/o C14 es mayor del 1 %, 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 %, 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, del 75-85 %, del 70-90 %, del 90-200 %, del 200-300 %, del 300-400 %, del 400-500 %, o mayor del 500 %.

[0077] En algunas formas de realización, una modificación genética adicional para aumentar el nivel de ácidos grasos de cadena media en la célula o aceite de la célula incluye la expresión de un gen de aciltransferasa de ácido lisofosfatídico exógeno que codifica una aciltransferasa de ácido lisofosfatídico activa (LPAAT) que cataliza la transferencia de un grupo acilo graso de cadena media hasta la posición sn-2 de un acilgliceróester sustituido. En una forma de realización relacionada específica, tanto una acil-ACP tioesterasa exógena, como se reivindica, como una LPAAT se expresan de forma estable en la célula. Como resultado de la introducción de ácidos nucleicos recombinantes en una célula oleaginosa (y especialmente en una célula microbiana plasmática), una tioesterasa específica de cadena media exógena y una LPAAT exógena que cataliza la transferencia de un grupo acilo graso de cadena media hasta la posición sn-2 de un acilgliceróester sustituido, la célula se puede realizar para aumentar el porcentaje de un ácido graso de cadena media particular en los triacilglicéridos (TAGs) que produce 10, 20, 30, 40, 50, 60, 70, 80, 90 veces, o más. La introducción de la LPAAT exógena puede aumentar los ácidos grasos de cadena media en la posición sn-2 1, 2, 3, 4 veces o más, en comparación con la introducción de una acil-ACP tioesterasa que prefiere cadenas medias exógena sola. En una forma de realización, el ácido graso de cadena media es mayor del 30, 40, 50, 60, 70, 80, o 90 % de los ácidos grasos TAG producidos por la célula. En varias formas de realización, el ácido graso de cadena media es cáprico, caprílico, láurico, mirístico, y/o palmítico.

[0078] En diversas formas de realización, el gen que codifica una aciltransferasa de ácido lisofosfatídico (LPAAT) se selecciona del grupo que consiste en *Arabidopsis thaliana* 1-acil-sn-glicerol-3-fosfato aciltransferasa (n.º de acceso de GenBank AEE85783), *Brassica juncea* 1-acil-sn-glicerol-3-fosfato aciltransferasa (n.º de acceso de GenBank ABQ42862), *Brassica juncea* 1-acil-sn-glicerol-3-fosfato aciltransferasa (n.º de acceso de GenBank ABM92334), *Brassica napus* 1-acil-sn-glicerol-3-fosfato aciltransferasa (n.º de acceso de GenBank CAB09138), aciltransferasa de ácido lisofosfatídico reinhardtii de *Chlamydomonas* (n.º de acceso de GenBank EDP02300), aciltransferasa de ácido lisofosfatídico de nucifera de Cocos (n.º de acceso de GenBank AAC49119), aciltransferasa de ácido lisofosfatídico de *Limnanthes alba* (n.º de acceso de GenBank EDP02300), *Limnanthes douglasii* 1-acil-sn-glicerol-3-fosfato aciltransferasa (putativa) (n.º de acceso de GenBank CAA88620), *Limnanthes douglasii* acil-CoA: sn-1-acilglicerol-3-fosfato aciltransferasa (n.º de acceso de GenBank ABD62751), *Limnanthes douglasii* 1-acilglicerol-3-fosfato O-aciltransferasa (n.º de acceso de GenBank CAA58239), *Ricinus communis* 1-acil-sn-glicerol-3-fosfato aciltransferasa (n.º de acceso de GenBank EEF39377).

[0079] Alternativamente, o además de la expresión de una LPAAT exógena, la célula puede comprender ácidos nucleicos recombinantes que son operables para expresar una enzima KASI o KASIV exógena y opcionalmente para reducir o eliminar la actividad de una KASII, que es particularmente ventajoso cuando se expresa una acil-ACP tioesterasa que prefiere cadenas medias. La ingeniería de las células de *Prototheca* para sobreexpresar las enzimas KASI y/o KASIV conjuntamente con una acil-ACP tioesterasa que prefiere cadenas medias puede generar cepas en las que la producción de ácidos grasos C10-C12 es al menos aproximadamente el 40 % del total de ácidos grasos, por ejemplo, al menos aproximadamente el 45 %, 50 %, 55 %, 60 % o más, del total de

ácidos grasos. La producción de cadena media también se puede aumentar al suprimir la actividad de KASI y/o KASII (por ejemplo, utilizando un knockout o un knockdown). Un knockout cromosómico de diferentes alelos de KASI de *Prototheca moriformis* (UTEX 1435) junto con la sobreexpresión de una acil-ACP tioesterasa que prefiere cadenas medias puede conseguir perfiles de ácidos grasos que sean al menos aproximadamente el 60 % de ácidos grasos C10-C14, por ejemplo, al menos aproximadamente el 65 %, 70 %, 75 %, 80 %, 85 % o más de ácidos grasos C10-C14. Los ácidos grasos de cadena media elevados también se pueden obtener como resultado de la expresión de polinucleótidos en horquilla de ARN de KASI. Además de cualquiera de estas modificaciones, la producción de ácidos grasos insaturados o poliinsaturados se puede suprimir (por ejemplo, por knockout o knockdown) de una enzima SAD o FAD.

[0080] En una forma de realización, una de las células productoras de alto contenido de cadenas medias descritas anteriormente se diseña posteriormente para producir un aceite poliinsaturado bajo al hacer knockout o knockdown a una o más desaturasas de acilos grasos. Por consiguiente, el aceite producido tiene una alta estabilidad.

[0081] Los aceites o ácidos grasos de alto contenido de cadenas medias derivados de hidrólisis de estos aceites pueden ser particularmente útiles en aplicaciones alimentarias, de combustible y oleoquímicas, incluida la producción de lubricantes y surfactantes. Por ejemplo, los ácidos grasos derivados de las células se pueden esterificar, craquear, reducir a un aldehído o alcohol, aminor, sulfatar, sulfonar, o someter a otro proceso químico conocido en la técnica.

[0082] La invención, que se ha descrito en detalle anteriormente, se ejemplifica en los siguientes ejemplos, que se ofrecen para ilustrar, pero no limitar, la invención reivindicada.

EJEMPLOS

[0083] Los ejemplos siguientes se ofrecen para ilustrar, pero no limitar, la invención reivindicada.

Ejemplo de referencia 1: mutagénesis de *Cuphea hookeriana* FATB2

[0084] Modificamos la actividad y especificidad de una tiroesterasa FATB2 aislada originalmente de *Cuphea hookeriana* (Ch FATB2, acceso U39834), usando mutagénesis dirigida al sitio de H163 y L186 dentro del núcleo enzimático (H163 y L186 dentro de Ch FATB2).

[0085] Para los ejemplos anteriores, se utilizó una construcción de expresión que dirigió las variantes de FATB y los marcadores de selección al locus *Thi4* (biosíntesis de tiamina). Se usó un gen de resistencia a los antibióticos para seleccionar para la resistencia al antibiótico G418. El promotor de UAPA se usó para conducir las FATB. La construcción se ejemplifica en SEQ ID NO: 9.

[0086] Como se describe en la PCT/US2014/013676, descubrimos que el injerto del dominio de especificidad del extremo N-terminal de *Cuphea avigera* FATB1 (Ca FATB1) (figura 2B) sobre la *Cuphea hookeriana* FATB2 (figura 2A) mejora la actividad y la proporción C8-C10. Los transformantes de *Prototheca moriformis* que expresan los mutantes Ch FATB2 H163Y, L186P (D3130) mostraron un aumento de aproximadamente 2 veces en la suma promedio de C8-C10, así como un cambio en la especificidad de perfil de ácidos grasos relativo al tipo salvaje Ch FATB2 (D3042).

[0087] La His en la posición 163 dentro de la Ch FATB2 (figura 2A) está conservada altamente a través de las tioesterasas FATB. En contraste, la Leu en la posición 186 dentro de la Ch FATB es rara. En otras FATB, la posición 186 está ocupada típicamente por una Pro o Leu. Debido a estas observaciones y también al aumento de la actividad y el cambio en la especificidad del perfil de ácidos grasos de cepas de *Prototheca moriformis* que expresan el mutante Ch FATB2 H163Y, L186P (D3130), identificamos H163 y L186 como "puntos calientes" para mutagénesis y mutagénesis exhaustiva realizada en ambos H163 y L186 para explorar el efecto de combinaciones de aminoácido en la actividad de la Ch FATB2 cuando se expresa dentro del sistema modelo de *Prototheca moriformis*. Los detalles del sistema de clonación se dan en la PCT/US2014/013676.

[0088] Se generaron treinta y ocho variantes Ch FATB2 individuales y su efecto sobre la acumulación de ácidos grasos C8:0 y C10:0. Los transformantes con la suma de C8-C10 dentro de 3 desviaciones estándar sobre el control Ch FATB2 de tipo salvaje (D3598) se clasificaron como positivos y aquellos dentro de las 3 desviaciones estándar por debajo se puntuaron como negativos. Véase la tabla 1. Los transformantes restantes se clasificaron como neutros. Como se muestra en la tabla 1, *Prototheca moriformis*, transformada con seis de los mutantes Ch FATB2 (D3570, D3573, D3582, D3584, D3588, y D3599), acumuló los ácidos grasos C8:0-C10:0 dentro de 3 estándar sobre los transformantes que expresan el control Ch FATB2 de tipo salvaje (D3598).

TABLA 1 Análisis de variantes de FATB para la producción de ácidos grasos C8-C14 en <i>P. moriformis</i>.*						
Variante de <i>Ch</i> FAPT2		C8:0	C10:0	C12:0	C14:0	Suma de C8-C10
D3565-186V	promedio	1,82	6,86	0,20	1,64	8,68
	STDEV	0,28	0,76	0,02	0,05	1,04
D3566-186Y	promedio	1,94	7,17	0,22	1,72	9,11
	STDEV	0,45	1,11	0,03	0,08	1,56
D3567-186W (negativo)	promedio	0,78	3,98	0,16	1,66	4,76
	STDEV	0,06	0,20	0,05	0,02	0,26
D3568-186T	promedio	1,78	7,00	0,21	1,61	8,78
	STDEV	0,38	0,89	0,02	0,06	1,27
D3569-186S	promedio	1,85	7,05	0,20	1,61	8,9
	STDEV	0,32	0,75	0,02	0,06	1,07
D3570-186P (positivo)	promedio	2,86	8,68	0,25	1,65	11,54
	STDEV	0,25	0,61	0,02	0,05	0,86
D3571-186F	promedio	1,46	5,98	0,19	1,62	7,44
	STDEV	0,32	0,78	0,02	0,06	1,1
D3572-186M	promedio	1,48	6,13	0,18	1,59	7,61
	STDEV	0,23	0,46	0,01	0,03	0,69
D3573-186K(positivo)	promedio	2,62	8,74	0,23	1,54	11,36
	STDEV	0,66	1,33	0,03	0,05	1,99
D3574-186I	promedio	1,56	6,35	0,18	1,63	7,91
	STDEV	0,26	0,55	0,01	0,04	0,81
D3575-186H	promedio	1,97	7,38	0,19	1,60	9,35
	STDEV	0,29	0,58	0,01	0,03	0,87
D3576-186G	promedio	1,41	5,83	0,18	1,70	7,24
	STDEV	0,25	0,69	0,02	0,07	0,94
D3577-186E	promedio	2,14	7,91	0,21	1,68	10,05
	STDEV	0,54	1,29	0,03	0,08	1,83
D3578-186Q	promedio	2,07	7,69	0,20	1,61	9,76
	STDEV	0,46	0,96	0,02	0,06	1,42
D3579-186C	promedio	0,95	4,77	0,16	1,60	5,72
	STDEV	0,07	0,23	0,01	0,03	0,3
D3580-186N	promedio	2,25	7,93	0,21	1,55	10,18
	STDEV	0,33	0,79	0,02	0,04	1,12
D3581-186R	promedio	2,21	7,74	0,22	1,63	9,95
	STDEV	0,73	1,90	0,04	0,07	2,63
D3582-186A(positivo)	promedio	2,39	8,74	0,23	1,58	11,13
	STDEV	0,78	1,94	0,04	0,06	2,72
D3603-186D	promedio	1,91	7,02	0,19	1,54	8,93

ES 2 772 123 T3

	STDEV	0,41	1,14	0,02	0,05	1,55
D3583-163V(negativo)	promedio	0,00	0,12	0,03	1,89	0,12
	STDEV	0,00	0,07	0,04	0,21	0,07
D3584-163Y(positivo)	promedio	3,71	10,52	0,30	1,61	14,23
	STDEV	0,92	1,75	0,04	0,04	2,67
D3585-163W	promedio	1,11	4,88	0,18	1,67	5,99
	STDEV	0,12	0,28	0,01	0,04	0,4
D3586-163T(negativo)	promedio	0,00	0,01	0,01	1,78	0,01
	STDEV	0,00	0,03	0,02	0,13	0,03
D3587-163P(negativo)	promedio	0,00	0,01	0,01	1,84	0,01
	STDEV	0,00	0,03	0,03	0,14	0,03
D3588-163F(positivo)	promedio	3,79	10,82	0,31	1,59	14,61
	STDEV	0,54	0,77	0,01	0,03	1,31
D3589-163K(negativo)	promedio	0,00	0,01	0,06	1,79	0,01
	STDEV	0,00	0,02	0,01	0,07	0,02
D3590-163L	promedio	1,95	7,49	0,20	1,66	9,44
	STDEV	0,38	1,15	0,03	0,08	1,53
D3591-163I(negativo)	promedio	0,06	0,70	0,07	1,74	0,76
	STDEV	0,02	0,15	0,01	0,11	0,17
D3592-163G(negativo)	promedio	0,00	0,01	0,06	1,81	0,01
	STDEV	0,00	0,02	0,01	0,03	0,02
D3593-163E(negativo)	promedio	0,00	0,02	0,06	1,99	0,02
	STDEV	0,01	0,05	0,02	0,20	0,06
D3594-163Q(negativo)	promedio	0,06	0,69	0,07	1,74	0,75
	STDEV	0,05	0,36	0,01	0,06	0,41
D3595-163C(negativo)	promedio	0,00	0,02	0,01	1,80	0,02
	STDEV	0,00	0,05	0,02	0,17	0,05
D3596-163R(negativo)	promedio	0,00	0,01	0,01	1,92	0,01
	STDEV	0,00	0,04	0,02	0,36	0,04
D3597-163A(negativo)	promedio	0,00	0,00	0,01	1,72	0
	STDEV	0,00	0,00	0,03	0,14	0
D3600-163S12 (negativo)	promedio	0,00	0,00	0,02	1,74	0
	STDEV	0,00	0,00	0,03	0,12	0
D3601-163M(negativo)	promedio	0,02	0,76	0,02	1,75	0,78
	STDEV	0,05	0,16	0,04	0,15	0,21
D3602-163N(negativo)	promedio	0,00	0,00	0,01	1,74	0
	STDEV	0,00	0,00	0,02	0,07	0
D3609-163D(negativo)	promedio	0,00	0,00	0,01	1,80	0
	STDEV	0,00	0,00	0,02	0,15	0
D3598- Ch FATB2 de tipo salvaje	promedio	1,52	6,55	0,19	1,60	8,07
	STDEV	0,19	0,62	0,02	0,11	0,81

D3599- H163Y, L186P(positivo)	promedio	5,77	12,50	0,39	1,73	18,27
	STDEV	0,63	0,99	0,03	0,05	1,62
* 12 transformantes fueron cribados por mutante. La duración de la producción de lípidos en condiciones de bajo nivel de nitrógeno fue de 3 días.						

[0089] En resumen, hemos demostrado que es posible aumentar la actividad y cambiar la especificidad de perfil dentro de las tioesterasas FATB específicas de C8-C10 derivadas de *Cuphea hookeriana* utilizando mutagénesis de H163 y L186 dirigida al sitio dentro del dominio de especificidad del extremo N-terminal. Encontramos células que expresaban variantes que excedían la secuencia parental ChFATB2 en términos de la suma de la producción de C8:0 + C10:0, incluidas las cepas que producían aceites con perfiles de ácidos grasos donde la producción de C8 y C10 superaba el 9, 11, 14 del perfil.

Ejemplo de referencia 2: identificación de mutantes dobles en FATB

[0090] Según la capacidad demostrada para modificar la actividad y especificidad de una tioesterasa FATB2 aislada originalmente de *Cuphea hookeriana* (Ch FATB2, acceso U39834), usando mutagénesis de H163 y L186 dirigida al sitio se inició una segunda ronda de mutagénesis. Se generaron seis construcciones que combinan las mutaciones positivas de Rd1 (C8 + C10 dentro de 3 desviaciones estándar sobre el control Ch FATB2 de tipo salvaje (D3598)) (tabla 2).

TABLA 2

	Construcciones de mutaciones beneficiosas	
1)	163Tyr	186Lys
2)	163Tyr	186Ala
3)	163Phe	186Pro
4)	163Phe	186Lys
5)	163Phe	186Ala

[0091] Para los ejemplos anteriores, se usó una construcción de expresión que dirigía las variantes de FATB y los marcadores de selección al locus Thi4 (biosíntesis de tiamina). Un gen de resistencia a antibióticos se usó para seleccionar la resistencia al antibiótico G418. El promotor UAPA se usó para conducir FATB. La construcción se ejemplifica en SEQ ID NO: 9.

[0092] Se generaron cinco variantes individuales Ch FATB2 variante y se cuantificó su efecto sobre la acumulación de ácidos grasos C8:0 y C10:0. Los transformantes con la suma de C8-C10 dentro de las 3 desviaciones estándar sobre el control Ch FATB2 de tipo salvaje (D3598) se clasificaron como positivos (tabla 3) y aquellos dentro de las 3 desviaciones estándar por debajo se clasificaron como negativos (tabla 3). Los transformantes restantes se clasificaron como neutros. Como se muestra en la tabla 3, *Prototheca moriformis* transformada con tres de los mutantes Ch FATB2 (D3875, D3876 y D3885) acumuló los ácidos grasos C8:0-C10:0 dentro de 3 estándar sobre los transformantes que expresan el control Ch FATB2 de tipo salvaje (D3598).

TABLA 3

Variante de Ch FAPT2		C8:0	C10:0	C12:0	C14:0	Suma de C8-C10
D3875-163F, 186A (positivo)	promedio	4,66	12,40	0,34	1,61	19,02
	STDEV	1,27	2,39	0,05	0,19	3,73
D3876-163F, 186K (positivo)	promedio	5,25	13,12	0,36	1,55	20,28
	STDEV	1,19	2,05	0,05	0,03	3,31
D3877-163F, 186P (negativo)	promedio	0,00	0,00	0,00	1,90	1,90
	STDEV	0,00	0,00	0,00	0,28	0,28

D3884-163Y, 186A	promedio	4,29	11,69	0,32	1,52	17,81
	STDEV	0,58	1,06	0,03	0,04	1,69
D3885-163Y, 186K (positivo)	promedio	5,39	13,14	0,36	1,49	20,38
	STDEV	1,31	2,18	0,05	0,03	3,52
D3598- <i>Ch</i> FATB2 de tipo salvaje	promedio	1,14	5,72	0,15	1,56	8,57
	STDEV	0,27	0,77	0,06	0,05	1,12
D3599- H163Y, L186P	promedio	5,65	12,91	0,39	1,74	20,69
	STDEV	1,29	2,06	0,06	0,02	3,42

Ejemplo 3: mutaciones en la posición 230 de FATB

[0093] En el siguiente ejemplo demostramos la capacidad para modificar la actividad y especificidad de tres tioesterasas FATB aisladas originalmente de *Cuphea hookeriana* (Ch FATB2, acceso Uniprot U39834), *Cuphea palustris* (Cpal FATB1, acceso Uniprot Q39554, SEQ ID NO: 13) y *Cuphea avigera* FATB1 (Ca FATB1 acceso R4J2L6, SEQ ID NO: 14) usando mutagénesis dirigida al sitio de una Met conservada dentro del núcleo enzimático (M230 dentro de Cpal FATB1).

[0094] Recientemente se ha informado que la sustitución de la M230 conservada dentro de la Cpal FATB1 con Iso, Val, Phe o Leu aumentará la actividad enzimática de esta tioesterasa. Dado que estos resultados se obtuvieron usando *E. coli*, realizamos un cribado similar para ver si los resultados podrían ser reproducidos cuando se expresan en microalgas de *Prototheca moriformis*. Se generaron los mutantes de tipo salvaje y trece Cpal FATB1 M230 y se cuantificó su efecto sobre la acumulación de ácidos grasos C8:0. Como se muestra en la tabla 4, *Prototheca moriformis* transformada con seis de los mutantes Cpal FATB1 M230 (D3206, D3208, D3211, D3212, D3214, y D3215), mostró perfiles de ácidos grasos que eran similares a la cepa de alga huésped S6165 no transformada que probablemente se deba a que la mutación inactiva la enzima Cpal FATB1. En cambio, los transformantes de *Prototheca moriformis* que expresan uno de los restantes siete mutantes Cpal FATB1 M230 acumularon los ácidos grasos C8:0 en diferentes grados sobre el huésped S6165 no transformado. D3213 (M230P) fue menos eficaz que los transformantes de Cpal FATB1 de tipo salvaje (D3004), mientras que D3207 (M230L) mostró los mismos niveles de ácidos grasos C8:0 que el Cpal FATB1 de tipo salvaje. D3210 (M230A), D3216 (M230T), y D3217 (M230F) acumularon ~1-1,5 % más de C8:0 que el D3004 de tipo salvaje. Finalmente, D3132 (M230I) y D3209 (M230V) mostraron un aumento de 4 veces en niveles C8:0, en comparación con el D3004 de tipo salvaje. Aunque que estos resultados comparten alguna similitud con los datos publicados derivados de la expresión en *E. coli*, hay algunas excepciones notables. Por ejemplo, a diferencia de en *E. coli*, la sustitución de M230 con Leu no mejoró la acumulación de ácidos grasos C8:0 relativa al Cpal FATB1 de tipo salvaje. Además, reemplazar la M230 con una Ala o Thr aumentó la acumulación de C8:0 con respecto al Cpal FATB1 de tipo salvaje, que no se esperaba en base al cribado basado en *E. coli*.

TABLA 4.					
Impacto en los perfiles de ácidos grasos tras la expresión de Cpal FATB1 de tipo salvaje o el mutante indicado dentro de la cepa de alga <i>P. moriformis</i> S6165.					
Transformante		C8:0	C10:0	C12:0	C14:0
Cpal FATB1 de tipo salvaje - D3004	promedio	3,67	0,52	0,21	1,43
	mediana	2,98	0,44	0,19	1,45
M2301 - D3132	promedio	13,04	1,66	0,40	1,13
	mediana	12,13	1,53	0,37	1,15
M230K - D3206	promedio	0,01	0,01	0,06	1,46
	mediana	0,00	0,00	0,06	1,45
M230L - D3207	promedio	3,32	0,44	0,55	1,54

	mediana	3,45	0,44	0,56	1,53
M230G - D3208	promedio	0,05	0,01	0,07	1,58
	mediana	0,07	0,00	0,07	1,58
M230V - D3209	promedio	14,13	1,96	0,81	1,36
	mediana	14,31	1,97	0,80	1,36
M230A - D3210	promedio	4,06	0,35	0,47	1,82
	mediana	3,92	0,34	0,45	1,80
M230R - D3211	promedio	0,00	0,02	0,06	1,43
	mediana	0,00	0,00	0,06	1,44
M230H - D3212	promedio	0,00	0,05	0,10	1,49
	mediana	0,00	0,05	0,09	1,47
M230P - D3213	promedio	1,78	0,54	1,24	2,85
	mediana	1,65	0,52	1,20	2,77
M230D - D3214	promedio	0,00	0,03	0,05	1,50
	mediana	0,00	0,03	0,05	1,50
M230E - D3215	promedio	0,00	0,00	0,05	1,49
	mediana	0,00	0,00	0,05	1,46
M230T - D3216	promedio	5,83	0,57	0,39	1,48
	mediana	5,77	0,57	0,40	1,52
M230F - D3217	promedio	5,75	0,97	0,93	1,78
	mediana	5,24	0,91	0,89	1,77
parental S6165		0	0	0	1,50

Los datos mostrados es el promedio y la mediana de 12-14 transformantes individuales para cada construcción de expresión Cpal FATB1.

[0095] Debido a las discrepancias en el resultado entre la expresión de *E. coli* y *P. moriformis*, exploramos la consecuencia de generar mutantes en la posición paralela dentro de tioesterasas FATB específicas de C8-C10 derivadas de *C. hookeriana* (Ch FATB2) y *C. avigera* (Ca FATB1). La figura 2 muestra los resultados de sustituir la Met con Iso en los Ch FATB2 (D3455, M228I) y Lys con Met o Iso en los Ca FATB1 (D3458 y D3459, respectivamente). De manera interesante, los transformantes que expresan el mutante Ch FATB2 M228I (D3455) exhiben ~50 % menos del total de ácidos grasos C8:0-C14:0, en comparación con la expresión de Ch FATB2 de tipo salvaje (D3042). Los transformantes que expresan K228M Ca FATB1 (D3458) produjeron un nivel de ácidos grasos C8:0-C14:0 ~1,5 veces superior, en comparación con la Ca FATB1 de tipo salvaje (D3456), mientras que el K228I Ca FATB1 (D3459) fue ligeramente menos efectivo que la expresión de Ca FATB1 de tipo salvaje. De manera importante, los mutantes K228M y K228I Ca FATB1 mostraron una preferencia nueva de ácidos grasos. Ambos mutantes Ca FATB1 acumularon un porcentaje inferior de C8:0 relativo al total de C8:0-C14:0, en comparación con la Ca FATB1 de tipo salvaje. Además, los transformantes que expresan el mutante K228I Ca FATB1 (D3459) produjo ácidos grasos C12:0 y C14:0, que no se observaron con la Ca FATB1 de tipo salvaje o K228M.

[0096] En resumen, hemos demostrado que las conclusiones extraídas del cribado de la expresión de *E. coli* solo concuerdan parcialmente con nuestros datos derivados de expresar los mutantes Cpal FATB1 en nuestra plataforma de *P. moriformis*. Además, los fenotipos observados en la sustitución de la misma posición de aminoácidos dentro de los Ch FATB2 y Ca FATB1 no son lo que se esperaba en base al cribado de la expresión original de *E. coli*. Nuestros resultados demuestran que la plataforma de expresión y la tioesterasa influyen en los resultados de un estudio de mutagénesis.

Ejemplo de referencia 4

[0097] Además de los resultados mostrados en la tabla 3, descubrimos que los transformantes de *Prototheca moriformis* que expresan los mutantes Ch FATB2 H163Y, L186P, y 230K (D3599) mostraron un cambio en la especificidad de perfil de ácidos grasos relativo al mejor mutante Ch FATB2 (D3599). Por lo tanto, se generó un conjunto adicional de mutantes para alterar la actividad y especificidad de Ch FATB2, tabla 5. La mutación 228K corresponde a la posición 230 de *Cuphea palustris* FATB1 (SEQ ID NO: 13). El residuo 230 de *Cuphea palustris* FATB1 corresponde a M228 en la *Cuphea hookeriana* FATB2 y *Cuphea avigera* FATB1.

TABLA 5

	Construcciones de mutaciones beneficiosas		
1)	163Tyr	186Pro	228Lys
2)	163Tyr	186Lys	228Lys
3)	163Tyr	186Ala	228Lys
4)	163Phe	186Pro	228Lys
5)	163Phe	186Lys	228Lys
6)	163Phe	186Ala	228Lys

[0098] Se generaron cinco variantes individuales de Ch FATB2 y se cuantificó su efecto en la acumulación de ácidos grasos C8:0 y C10:0.

TABLA 6

Variante Ch FAPT2		C8:0	C10:0	C12:0	C14:0	Suma de C8-C10
D3886-163F, 186A, 228K	promedio	3,72	2,33	0,20	1,80	8,05
	STDEV	1,06	0,46	0,03	0,10	1,64
D3887-163F, 186K, 228K	promedio	5,16	2,97	0,25	1,88	10,25
	STDEV	1,34	0,61	0,04	0,12	2,11
D3888-163F, 186P, 228K	promedio	4,57	2,72	0,18	1,85	9,32
	STDEV	1,42	0,71	0,06	0,07	2,24
D3895-163Y, 186A, 228K	promedio	4,17	2,51	0,20	1,84	8,71
	STDEV	1,72	0,93	0,07	0,10	2,80
D3896-163Y, 186K, 228K	promedio	4,35	2,70	0,22	1,80	9,06
	STDEV	0,73	0,28	0,02	0,07	1,08
D3598- Ch FATB2 de tipo salvaje	promedio	1,14	5,72	0,15	1,56	8,57
	STDEV	0,27	0,77	0,06	0,05	1,12
D3519- H163Y, L186P, 228K	promedio	6,27	3,57	0,22	1,89	11,94
	STDEV	2,10	0,86	0,05	0,08	3,07

Ejemplo 5

[0099] En el siguiente ejemplo, demostramos la capacidad para modificar la actividad y especificidad de una tioesterasa FATA aislada originalmente de *Garcinia mangostana* (GmFATA, acceso O04792), usando mutagénesis dirigida al sitio que se dirige a seis posiciones de aminoácidos en la enzima. La razón para dirigirse a tres de los aminoácidos (G108, S111, V193) se basó en una investigación publicada por Facciotti, et al., *Nat Biotechnol.* (1999) 17(6):593-7. Los tres aminoácidos restantes (L91, G96, T156) se seleccionaron en base a la investigación realizada en Solazyme con otras tioesterasas.

[0100] Para probar el impacto de cada mutación en la actividad de la GmFATA, los genes de tipo salvaje y mutantes fueron clonados en un vector que permite la expresión dentro de la cepa de *P. moriformis* S3150. La tabla 7 resume los resultados a partir de un cribado de perfil de lípidos de tres días comparando la GmFATA de tipo salvaje con los 14 mutantes. Tres mutantes de GmFATA (D3998, D4000, D4003) aumentaron la cantidad de C18:0 al menos 1,5 veces, en comparación con la proteína de tipo salvaje (D3997). D3998 y D4003 fueron mutaciones que habían sido descritas por Facciotti et al (*NatBiotech* 1999) como sustituciones que aumentaban la actividad de GmFATA. En cambio, la mutación D4000 se basó en la investigación en Solazyme que demostró que esta posición influenció la actividad de las tioesterasas FATB.

Tabla 7

Cepa de alga	ADN #	SEQ ID NO:	GmFATA	C14:0	C16:0	C18:0	C18:1	C18:2
<i>P. moriformis</i> S3150		-	-	1,63	29,82	3,08	55,95	7,22
	D3997	15	<i>GmFATA de tipo salvaje</i>	1,79	29,28	7,32	52,88	6,21
	D3998	16	S111A, V193A	1,84	28,88	11,19	49,08	6,21
	D3999	17	S111V, V193A	1,73	29,92	3,23	56,48	6,46
	D4000	18	G96A	1,76	30,19	12,66	45,99	6,01
	D4001	19	G96T	1,82	30,60	3,58	55,50	6,28

	D4002	20	G96V	1,78	29,35	3,45	56,77	6,43
	D4003	21	G108A	1,77	29,06	12,31	47,86	6,08
	D4007	25	G108V	1,81	28,78	5,71	55,05	6,26
	D4004	22	L91F	1,76	29,60	6,97	53,04	6,13
	D4005	23	L91K	1,87	28,89	4,38	56,24	6,35
	D4006	24	L91S	1,85	28,06	4,81	56,45	6,47

	D4008	26	T156F	1,81	28,71	3,65	57,35	6,31
	D4009	27	T156A	1,72	29,66	5,44	54,54	6,26
	D4010	28	T156K	1,73	29,95	3,17	56,86	6,21
	D4011	29	T156V	1,80	29,17	4,97	55,44	6,27

Ejemplo 6

[0101] El lípido de almacenamiento de *P. moriformis* de tipo salvaje está compuesto de ~60 % de ácido oleico (C18:1), ~ 25-30 % de ácido palmítico (C16:0) y ~5-8 % de ácido linoleico (C18:2), con cantidades menores de ácido esteárico (C18:0), ácido mirístico (C14:0), ácido α -linolénico (C18:3 α) y ácido palmitoleico (C16:1). Este perfil de ácidos grasos surge de las actividades relativas y afinidades de sustrato de las enzimas de la vía biosintética endógena de ácidos grasos. La introducción del gen de tioesterasa FATA de *Garcinia mangostana* (GarmFATA1) en *P. moriformis* da como resultado aceites con niveles aumentados de estearato (C18:0). Además, demostramos que las mutaciones únicas G96A y G108A, y las dobles mutaciones (S111A, V193A) en GarmFATA1 aumentaron la acumulación de C18:0 relativa a la proteína nativa GarmFATA1.

[0102] En el presente ejemplo, evaluamos la actividad de tioesterasa de una serie de mutantes GarmFATA1 adicionales. Estos mutantes se generaron al combinar las mutaciones G96A, G108A, S111A y V193A descritas anteriormente en mutantes dobles, triples o cuádruples. Específicamente, evaluamos las combinaciones de mutantes GarmFATA1 (G96A, G108A), GarmFATA1 (G96A, S111A), GarmFATA1 (G96A, V193A), GarmFATA1 (G108A, S111A), GarmFATA1 (G108A, V193A), GarmFATA1 (G96A, G108A, S111A), GarmFATA1 (G96A, G108A, V193A), GarmFATA1 (G96A, S111A, V193A), GarmFATA1 (G108A, S111A, V193A), y GarmFATA1 (G96A, G108A, S111A, V193A). GarmFATA1 (G108A) se usó como un control, ya que, de todos los mutantes evaluados anteriormente, este dio el mejor aumento en los niveles de C18: 0 sobre la proteína nativa GARMFATA1. El cribado se efectuó en S5780, una cepa construida previamente en S5100 - una cepa de base oleica alta. S5780 se creó a través de la delección dirigida del alelo dominante SAD2-1, lo que reduce el índice de conversión de C18:0 a C18:1, y la sobreexpresión de PmKASII, lo que aumenta el alargamiento de C16:0 a C18:0. S5780 se transformó con construcciones que se dirigían al gen LPAT2 de *T. cacao* (TcLPAT2) y las combinaciones anteriormente mencionadas de los mutantes dirigidos al sitio GarmFATA1 al locus FATA-1. TcLPAT2 es altamente específico para la incorporación de ácidos grasos insaturados en la posición sn-2 de TAGs. La cepa S5780, que contiene una delección de un alelo estearoil ACP desaturasa (SAD), se realizó según las instrucciones en las aplicaciones de propiedad conjunta WO2010/063031, WO2011/150411, y/o WO2012/106560.

Construcción usada para la expresión de TcLPAT2 y GarmFATA1 (G96A, G108A) en el locus PmFATA1 - (pSZ5990)

[0103] En este ejemplo, se generó la cepa S5780, transformada con la construcción pSZ5990, que expresa el gen de *Saccharomyces carlbergensis* SUC2 (que permite su selección y crecimiento en medio que contiene sacarosa), una LPAT2 de *T. cacao* y una tioesterasa FATA1 de *G. mangostana* (G96A; G108A) dirigidos a la región genómica endógena de PmFATA1. La construcción pSZ5990 introducida para la expresión en S5780 se puede escribir como FATA-1 3' flanco::CrTub2-ScSUC2-PmPGH:separador1:PmG3PDH-1-TcLPAT2-PmATP:separador2:PmSAD2-2v2-CpSAD1tp_GarmFATA1(G96A, G108A)_FLAG-PmSAD2-1::FATA-1 5' flanco.

[0104] La secuencia del ADN transformante se proporciona en la figura 1. Los sitios de restricción relevantes en la construcción se indican en minúsculas, subrayados en negrita, y son de 5'-3' BspQI, KpnI, XbaI, MfeI, BamHI, AvrII, NdeI, NsiI, AflII EcoRI, SpeI, AscII, ClaI, SacI y BspQI, respectivamente. Los sitios BspQI delimitan los extremos 5' y 3' del ADN transformante. Las secuencias en minúsculas y negrita representan el ADN genómico de *P. moriformis*, que permite la integración dirigida en el locus FATA1 mediante recombinación homóloga. Siguiendo en la dirección 5' a 3', la tubulina 2 de *Chlorella reinhardtii*, que impulsa la expresión del gen de *S. cerevisiae* SUC2, se indica por un texto en minúsculas en recuadro. Las cursivas en mayúsculas indican el iniciador ATG y el terminador TGA para SUC2, mientras que la región codificante se indica con minúsculas en cursiva. La 3' UTR del gen de la fosfoglicerato deshidratasa (PGH) de *P. moriformis* se indica con un texto subrayado en minúsculas, seguido de tampón/secuencia de ADN separadora-1 indicada con un texto en cursiva, negrita y minúsculas. Inmediatamente después del nucleótido tampón hay un promotor endógeno de G3PDH-1 de *P. moriformis*, indicado con un texto en minúsculas en recuadro. Las cursivas en mayúsculas indican los codones del Iniciador ATG y el terminador TGA del gen LPAT2 de *T. cacao*, mientras que la cursiva en minúsculas indica el resto del gen. La 3' UTR del gen de trifosfato de adenosina de *P. moriformis* (ATP) se indica con un texto

subrayado en minúsculas, seguido de tampón/secuencia de nucleótidos espaciadora 2 indicada con un texto en cursiva, en negrita y minúsculas. Inmediatamente después de la secuencia espaciadora-2 está el promotor endógeno de PmSAD2-2 de *P. moriformis*, indicado por un texto en minúsculas en recuadro. Las cursivas en mayúsculas indican el iniciador ATG y el terminador TGA para el gen FATA1 de *G. mangostana*, mientras que la región codificante se indica con cursiva en minúsculas. El gen FATA1 se fusiona traduccionalmente con el péptido de tránsito de *C. protothecoides* estearoil ACP desaturasa-1 (CpSAD1) en el extremo N-terminal (indicado por un texto en cursiva, minúscula y subrayado) para el direccionamiento apropiado al cloroplasto y el marcador 3XFLAG en el extremo C-terminal (indicado por texto con doble subrayado, en cursiva, negrita y minúsculas). GarmFATA1 con el péptido de tránsito CpSAD y la secuencia 3XFLAG es seguido por la 3' UTR del gen endógeno de la estearoil ACP desaturasa-1 (SAD1) indicada con un texto subrayado en minúsculas. La secuencia genómica del gen endógeno FATA1 se indica con un texto en negrita y minúsculas. La construcción final fue secuenciada para asegurar marcos de lectura correctos y secuencias de direccionamiento, y se proporciona como SEQ ID NO:46.

[0105] Además de los genes LPAT2 de *T. cacao* y FATA1 de *G. mangostana* (G96A, G108A) dirigidos al locus PmFATA1 (pSZ5990), se designaron muchas otras construcciones que incorporan las diversas mutaciones anteriormente descritas para la transformación en S5780. Estas construcciones se resumen a continuación en la tabla 8:

TABLA 8		
Plásmido	SEQ ID NO:	Genotipo
pSZ5936	47	FATA-1::CrTUB2-ScSUC2-PmPGH:PmG3PDH-1-TcLPAT2-PmATP:PmSAD2-2-CpSAD1_GarmFATA1(G108A)_FLAG-PmSAD2-1::FATA-1_5'
pSZ5991	48	FATA-1_3'::CrTUB2-ScSUC2-PmPGH:PmG3PDH-1-TcLPAT2-PmATP:PmSAD2-2-CpSAD1_GarmFATA1(G96A,S111A)_FLAG-PmSAD2-1::FATA-1_5'
pSZ5986	49	FATA-1_3'::CrTUB2-ScSUC2-PmPGH:PmG3PDH-1-TcLPAT2-PmATP:PmSAD2-2-CpSAD1_GarmFATA1(G96A,V193A)_FLAG-PmSAD2-1::FATA-1_5'
pSZ5982	50	FATA-1_3'::CrTUB2-ScSUC2-PmPGH:PmG3PDH-1-TcLPAT2-PmATP:PmSAD2-2-CpSAD1_GarmFATA1(G108A,S111A)_FLAG-PmSAD2-1::FATA-1_5'
pSZ5983	51	FATA-1_3'::CrTUB2-ScSUC2-PmPGH:PmG3PDH-1-TcLPAT2-PmATP:PmSAD2-2-CpSAD1_GarmFATA1(G108A,V193A)_FLAG-PmSAD2-1::FATA-1_5'
pSZ6005	52	FATA-1_3'::CrTUB2-ScSUC2-PmPGH:PmG3PDH-1-TcLPAT2-PmATP:PmSAD2-2-CpSAD1_GarmFATA1(G96A, G108A, S111A)_FLAG-PmSAD2-1::FATA-1_5'
pSZ5984	53	FATA-1_3'::CrTUB2-ScSUC2-PmPGH:PmG3PDH-1-TcLPAT2-PmATP:PmSAD2-2-CpSAD1_GarmFATA1(G96A, G108A, V193A)_FLAG-PmSAD2-1::FATA-1_5'
pSZ6004	54	FATA-1_3'::CrTUB2-ScSUC2-PmPGH:PmG3PDH-1-TcLPAT2-PmATP:PmSAD2-2-CpSAD1_GarmFATA1(G96A, S111A, V193A)_FLAG-PmSAD2-1::FATA-1_5'
pSZ5985	55	FATA-1_3'::CrTUB2-ScSUC2-PmPGH:PmG3PDH-1-TcLPAT2-PmATP:PmSAD2-2-CpSAD1_GarmFATA1(G108A, S111A, V193A)_FLAG-PmSAD2-1::FATA-1_5'
pSZ5987	56	FATA-1_3'::CrTUB2-ScSUC2-PmPGH:PmG3PDH-1-TcLPAT2-PmATP:PmSAD2-2-CpSAD1_GarmFATA1(G96A, G108A, S111A, V193A)_FLAG-PmSAD2-1::FATA-1_5'
pSZ6018	47	FATA-1::CrTUB2-ScSUC2-PmPGH:PmG3PDH-1-TcLPAT2-PmATP:PmACPp1p-CpSAD1_GarmFATA1(G108A)_FLAG-PmSAD2-1::FATA-1_5'
pSZ6019	50	FATA-1::CrTub2-ScSUC2-PmPGH:PmG3PDH-1-TcLPAT2-PmATP:PmACP-P1p-CpSAD1tp_GarmFATA1(G108A, S111A)_FLAG-PmSAD2-1::FATA-1
pSZ6020	51	FATA-1::CrTub2-ScSUC2-PmPGH:PmG3PDH-1-TcLPAT2-PmATP:PmACP-P1p-CpSAD1tp_GarmFATA1(G108A, V193A)_FLAG-PmSAD2-1::FATA-1
pSZ6021	53	FATA-1::CrTub2-ScSUC2-PmPGH:PmG3PDH-1-TcLPAT2-PmATP:PmACP-P1p-CpSAD1tp_GarmFATA1(G96A, G108A, V193A)_FLAG-PmSAD2-1::FATA-1
pSZ6022	55	FATA-1::CrTub2-ScSUC2-PmPGH:PmG3PDH-1-TcLPAT2-PmATP:PmACP-P1p-CpSAD1tp_GarmFATA1(G108, S111A, V193A)_FLAG-PmSAD2-1::FATA-1
pSZ6023	49	FATA-1::CrTub2-ScSUC2-PmPGH:PmG3PDH-1-TcLPAT2-PmATP:PmACP-P1p-CpSAD1tp_GarmFATA1(G96A, V193A)_FLAG-PmSAD2-1::FATA-1

TABLA 8		
Plásmido	SEQ ID NO:	Genotipo
pSZ6026	48	FATA-1::CrTub2-ScSUC2-PmPGH:PmG3PDH-1-TcLPAT2-PmATP:PmACP-P1p-CpSAD1tp_GarmFATA1(G96A, S111A)_FLAG-PmSAD2-1:FATA-1
pSZ6028	52	FATA-1::CrTub2-ScSUC2-PmPGH:PmG3PDH-1-TcLPAT2-PmATP:PmACP-P1p-CpSAD1tp_GarmFATA1(G96A, G108A, S111A)_FLAG-PmSAD2-1::FATA-1

[0106] Todas estas construcciones tienen el mismo esqueleto de vector, marcador seleccionable, los mismos promotores, genes y 3' UTRs que pSZ5990 que difieren solo en las mutaciones en el GarmFATA1. Asimismo, las construcciones pSZ6019 t0 pSZ6023, pSZ6026 y pSZ6028 difieren en el promotor, que impulsa el mutante particular GarmFATA1. Mientras que en pSZ5990, el GarmFATA1 (G96A, G108A) es impulsado por el promotor PmSAD2-2v2, las diversas combinaciones mutantes GarmFATA1 en pSZ6019-pSZ6028 son impulsadas por el promotor PmACP-P1. Las secuencias de nucleótidos de varios mutantes GarmFATA1 usados en las construcciones anteriores se muestran en SEQ ID NOs: 47-56. La secuencia promotora de PmACP-P1, que es pSZ6019 - pSZ6028, se muestra en SEQ ID NO: 57. Los sitios de restricción relevantes, como texto en negrita, se muestran 5'-3' respectivamente.

[0107] Para determinar su impacto en perfiles de ácidos grasos, todas las construcciones anteriormente descritas se transformaron independientemente en cualquiera de los S5780. Los transformantes primarios se purificaron clonalmente y se cultivaron bajo condiciones estándar de producción de lípidos a pH5,0. Los perfiles resultantes a partir de un conjunto de clones representativos que surgen de la transformación de S5780 con pSZ5936 (D4933), pSZ5990 (D4950), pSZ5991 (D4951), pSZ5986 (D4948), pSZ5982 (4931), pSZ5983 (D4932), pSZ6005 (D4952), pSZ5984 (D4933), pSZ6004 (D4953), pSZ5985 (D4934), pSZ5987 (D4949), pSZ6018 (D4978), pSZ6019 (D4979), pSZ6020 (D4980), pSZ6021 (D4981), pSZ6022 (D4982), pSZ6023 (D4983), pSZ6026 (D4986), pSZ6028 (D4988) se muestran en las tablas 9-19 respectivamente.

[0108] La tabla 13 enumera los perfiles promedio de ácidos grasos y el consumo de glucosa (relativos a la cepa base S7485) para cada conjunto de transformantes. La interrupción de un alelo de FATA-1 reduce C16:0 al 1-2 %, mientras que la actividad de TcLPAT2 se manifiesta como un aumento del 1-1,5 % en C18:2 en estas cepas. D4993 Y D4978, que expresan el mutante GarmFATA1 (G108A), acumularon entre del 44,69 % al 45,33 % y del 34,26 al 50,94 % de C18:0, respectivamente. D4993 tiene el GarmFATA1 (G108A) impulsado por el promotor PmSAD2-2, mientras que, para D4978, el promotor PmACP-P1 impulsa el GarmFATA1 (G108A). Las cepas con las combinaciones (G96A, G108A), (G108A, S111A) y (G108A, V193A) acumularon consistentemente más C18:0 que el mutante único (G108A), con incrementos mínimos en C16:0. D4950 (G96A, G108A) produjeron más del 50 % de C18:0 en cepas múltiples. Los mutantes triples (G96A, G108A, S111A), (G96A, G108A, V193A) y (G96A, S111A, V193A) también produjeron generalmente un C18:0 mayor, pero a un coste de C16:0 aumentado. El mutante triple (G108A, S111A, V193A) y el mutante cuádruple (G96A, G108A, S111A, V193A) produjeron un C18:0 menor que el mutante único G108. El promotor PmACP-P1 generalmente dio como resultado más C18:0 que los impulsados por el promotor PmSAD2-2.

Tabla 9

Identificación de la muestra	Perfil de ácidos grasos						
	C14:0	C16:0	C18:0	C18:1	C18:2	C18:3 α	C20:0
S5780	0,75	7,07	30,32	51,61	5,96	0,79	2,16
S5780; T1402; D4993-7	0,65	4,66	45,33	38,86	7,42	0,64	1,50
S5780; T1402; D4993-4	0,66	4,62	45,22	38,64	7,70	0,67	1,51
S5780; T1402; D4993-2	0,63	4,54	44,94	39,11	7,59	0,66	1,54
S5780; T1402; D4993-8	0,65	4,52	44,92	39,22	7,62	0,65	1,50
S5780; T1402; D4993-9	0,64	4,60	44,69	39,45	7,52	0,64	1,48
S5780; T1395; D4978-1	0,72	5,22	50,94	32,58	7,49	0,67	1,43
S5780; T1395; D4978-6	0,68	5,15	49,26	34,74	7,17	0,65	1,45
S5780; T1395; D4978-2	0,78	6,21	43,12	39,62	7,01	0,72	1,57

Identificación de la muestra	Perfil de ácidos grasos						
	C14:0	C16:0	C18:0	C18:1	C18:2	C18:3 α	C20:0
S5780; T1395; D4978-3	0,79	6,90	34,26	48,01	6,41	0,79	1,91

[0109] La tabla 9 proporciona perfiles primarios de ácidos grasos de 3 días de cepas representativas S5780 transformadas con D4993 (pSZ5936) y D4978 (pSZ6018). Tanto pSZ5936 como pSZ6018 tienen el mutante GarmFATA1 (G108) impulsado por PmSAD2-2 o PmACP-P1, respectivamente.

Tabla 10

Identificación de la muestra	Perfil de ácidos grasos						
	C14:0	C16:0	C18:0	C18:1	C18:2	C18:3 α	C20:0
S5780	0,70	6,98	30,82	51,28	5,80	0,77	2,27
S5780; T1402; D4950-3	0,64	5,16	50,73	32,92	7,34	0,62	1,41
S5780; T1402; D4950-8	0,64	5,17	50,63	33,10	7,28	0,62	1,41
S5780; T1402; D4950-5	0,66	5,20	50,23	33,31	7,42	0,62	1,40
S5780; T1402; D4950-7	0,65	5,15	49,90	33,81	7,31	0,63	1,40
S5780; T1402; D4950-4	0,66	5,22	49,53	34,13	7,21	0,61	1,42

- 5 [0110] La tabla 10 proporciona perfiles primarios de ácidos grasos de 3 días de cepas representativas S5780 transformadas con D4950 (pSZ5990). PSZ5990 expresa el mutante GarmFATA1 (G96A, G108) impulsado por PmSAD2-2.

Tabla 11

Identificación de la muestra	Perfil de ácidos grasos						
	C14:0	C16:0	C18:0	C18:1	C18:2	C18:3 α	C20:0
S5780	0,81	7,56	31,15	50,19	6,12	0,82	2,18
S5780; T1402; D4951-18	0,73	5,11	46,22	36,56	7,92	0,72	1,54
S5780; T1402; D4951-8	0,70	4,80	42,65	40,59	7,77	0,72	1,58
S5780; T1402; D4951-3	0,70	4,82	42,42	40,74	7,76	0,71	1,58
S5780; T1402; D4951-4	0,69	4,82	42,28	40,88	7,76	0,73	1,60
S5780; T1402; D4951-15	0,72	4,95	42,07	40,72	8,00	0,73	1,58
S5780; T1395; D4986-21	0,79	5,78	48,77	33,99	7,54	0,69	1,48
S5780; T1395; D4986-18	0,77	5,77	48,43	34,61	7,32	0,65	1,46
S5780; T1395; D4986-23	0,78	5,66	47,64	35,30	7,44	0,69	1,49
S5780; T1395; D4986-15	0,75	5,52	47,60	35,80	7,21	0,67	1,50
S5780; T1395; D4986-1	0,84	6,38	46,95	34,55	8,29	0,64	1,33

- 10 [0111] La tabla 11 proporciona perfiles primarios de ácidos grasos de 3 días de cepas representativas S5780 transformadas con D4951 (pSZ5991) y D4986 (pSZ6026). PSZ5991 y pSZ6026 expresan el mutante GarmFATA1 (G96A, S111A) impulsado por PmSAD2-2 y PmACP-P1, respectivamente.

Tabla 12

Identificación de la muestra	Perfil de ácidos grasos						
	C14:0	C16:0	C18:0	C18:1	C18:2	C18:3 α	C20:0
S5780	0,82	7,51	30,86	50,34	6,27	0,86	2,16
S5780; T1388; D4948-6	0,70	4,93	46,65	36,92	7,66	0,68	1,50
S5780; T1388; D4948-10	0,66	4,79	46,23	37,72	7,51	0,68	1,43
S5780; T1388; D4948-11	0,72	5,05	46,18	36,89	8,04	0,67	1,47
S5780; T1388; D4948-4	0,72	5,11	46,11	36,97	8,00	0,66	1,45
S5780; T1388; D4948-9	0,72	5,06	46,09	36,96	8,05	0,67	1,45
S5780; T1395; D4983-25	0,73	5,85	49,47	32,96	7,77	0,62	1,49
S5780; T1395; D4983-14	0,68	5,25	48,53	35,02	7,32	0,63	1,52
S5780; T1395; D4983-27	0,70	5,66	48,35	34,56	7,56	0,62	1,50
S5780; T1395; D4983-18	0,67	5,30	48,26	35,35	7,29	0,62	1,51
S5780; T1395; D4983-13	0,68	5,31	48,09	35,59	7,27	0,63	1,48

[0112] La tabla 12 proporciona perfiles primarios de ácidos grasos de 3 días de las cepas representativas S5780 transformadas con D4948 (pSZ5986) y D4983 (pSZ6023). PSZ5986 y pSZ6023 expresan el mutante GarmFATA1 (G96A, V193A) impulsado por PmSAD2-2 y PmACP-P1, respectivamente.

Tabla 13

Identificación de la muestra	Perfil de ácidos grasos						
	C14:0	C16:0	C18:0	C18:1	C18:2	C18:3 α	C20:0
S5780	0,77	7,36	30,84	50,71	6,24	0,87	2,12
S5780; T1388; D4931-25	0,70	5,13	48,48	35,40	7,25	0,64	1,43
S5780; T1388; D4931-9	0,73	5,43	48,29	34,92	7,63	0,65	1,41
S5780; T1388; D4931-13	0,72	5,24	48,13	35,22	7,64	0,66	1,45
S5780; T1388; D4931-17	0,76	5,14	48,07	35,08	7,86	0,68	1,42
S5780; T1388; D4931-12	0,73	5,33	47,91	35,27	7,65	0,67	1,42
S5780; T1395; D4979-36	0,89	6,91	50,03	31,13	7,83	0,67	1,40
S5780; T1395; D4979-5	0,77	5,88	49,65	33,24	7,25	0,68	1,48
S5780; T1395; D4979-41	0,79	6,25	49,52	33,09	7,28	0,63	1,42
S5780; T1395; D4979-39	0,82	6,36	49,43	32,49	7,66	0,66	1,48
S5780; T1395; D4979-32	0,82	6,49	49,12	32,98	7,45	0,63	1,43

[0113] La tabla 13 proporciona perfiles primarios de ácidos grasos de 3 días de las cepas representativas S5780 transformadas con D4931 (pSZ5982) y D4979 (pSZ6019). PSZ5982 y pSZ6019 expresan el mutante GarmFATA1 (G108A, S111A) impulsado por PmSAD2-2 y PmACP-P1, respectivamente.

Tabla 14

Identificación de la muestra	Perfil de ácidos grasos						
	C14:0	C16:0	C18:0	C18:1	C18:2	C18:3α	C20:0
S5780	0,79	7,46	31,60	49,57	6,28	0,86	2,19
S5780; T1388; D4932-48	0,78	4,59	49,48	31,74	9,13	1,30	1,84
S5780; T1388; D4932-36	0,66	4,89	49,25	34,63	7,53	0,66	1,43
S5780; T1388; D4932-28	0,66	4,93	49,04	34,91	7,50	0,65	1,38
S5780; T1388; D4932-23	0,67	4,95	49,03	34,55	7,69	0,66	1,42
S5780; T1388; D4932-5	0,68	4,93	49,01	34,77	7,54	0,67	1,38
S5780; T1395; D4980-21	0,71	4,54	51,48	32,09	7,54	0,87	1,78
S5780; T1395; D4980-1	0,72	5,80	48,65	33,81	8,04	0,62	1,46
S5780; T1395; D4980-25	0,68	5,46	47,67	35,53	7,61	0,66	1,47
S5780; T1395; D4980-18	0,77	6,49	46,51	34,39	8,69	0,71	1,45
S5780; T1395; D4980-30	0,70	5,22	45,14	38,84	6,80	0,70	1,70

[0114] La tabla 14 proporciona perfiles primarios de ácidos grasos de 3 días de las cepas representativas S5780 transformadas con D4932 (pSZ5983) y D4980 (pSZ6020). PSZ5983 y pSZ6020 expresan el mutante GarmFATA1 (G108A, V193A) impulsado por PmSAD2-2 y PmACP-P1, respectivamente.

Tabla 15

Identificación de la muestra	Perfil de ácidos grasos						
	C14:0	C16:0	C18:0	C18:1	C18:2	C18:3α	C20:0
S5780	0,75	7,14	30,84	51,19	5,87	0,79	2,26
S5780; T1402; D4952-9	0,77	5,68	48,96	33,44	7,70	0,68	1,50
S5780; T1402; D4952-5	0,75	5,58	48,60	33,94	7,60	0,70	1,52
S5780; T1402; D4952-1	0,75	5,62	48,59	33,94	7,63	0,69	1,51
S5780; T1402; D4952-8	0,78	5,78	48,51	33,71	7,74	0,67	1,50
S5780; T1402; D4952-10	0,77	5,65	48,35	34,15	7,59	0,70	1,52
S5780; T1395; D4988-5	0,99	8,68	48,51	31,29	7,08	0,64	1,51
S5780; T1395; D4988-7	0,75	5,50	46,63	36,68	7,41	0,69	1,43
S5780; T1395; D4988-8	0,77	5,57	46,51	36,73	7,47	0,70	1,42
S5780; T1395; D4988-3	1,12	9,63	44,06	33,16	8,33	0,76	1,57
S5780; T1395; D4988-10	1,27	11,45	43,35	31,26	8,95	0,74	1,49

[0115] La tabla 15 proporciona perfiles primarios de ácidos grasos de 3 días de las cepas representativas S5780 transformadas con D4952 (pSZ6005) y D4988 (pSZ6028). PSZ6005 y pSZ6028 expresan el mutante GarmFATA1 (G96A, G108A, S111A) impulsado por PmSAD2-2 y PmACP-P1, respectivamente.

Tabla 16

Identificación de la muestra	Perfil de ácidos grasos						
	C14:0	C16:0	C18:0	C18:1	C18:2	C18:3 α	C20:0
S5780	0,79	7,46	31,60	49,57	6,28	0,86	2,19
S5780; T1388; D4933-12	0,67	5,48	50,40	32,85	7,31	0,63	1,36
S5780; T1388; D4933-9	0,69	5,68	50,20	32,58	7,55	0,65	1,41
S5780; T1388; D4933-8	0,66	5,46	50,07	33,20	7,35	0,63	1,39
S5780; T1388; D4933-2	0,70	5,66	49,99	32,81	7,61	0,63	1,38
S5780; T1388; D4933-5	0,85	5,84	41,97	39,70	7,94	0,97	1,36
S5780; T1395; D4981-1	0,63	5,07	37,33	46,45	6,75	0,76	2,00
S5780; T1395; D4981-3	0,71	5,70	34,96	47,88	7,02	0,86	1,88
S5780; T1395; D4981-7	0,70	5,87	34,44	48,58	6,52	0,78	2,04
S5780; T1395; D4981-4	0,75	6,18	33,78	48,83	6,61	0,83	1,98
S5780; T1395; D4981-8	0,71	6,42	33,38	49,33	6,05	0,78	2,21

[0116] La tabla 16 proporciona perfiles primarios de ácidos grasos de 3 días de las cepas representativas S5780 transformadas con D4933 (pSZ5984) y D4981 (pSZ6021). PSZ5984 y pSZ6021 expresan el mutante GarmFATA1 (G96A, G108A, V193A) impulsado por PmSAD2-2 y PmACP-P1, respectivamente.

Tabla 17

Identificación de la muestra	Perfil de ácidos grasos						
	C14:0	C16:0	C18:0	C18:1	C18:2	C18:3 α	C20:0
S5780	0,74	7,27	31,04	50,75	5,96	0,81	2,18
S5780; T1388; D4953-6	0,84	6,99	47,90	33,26	7,58	0,65	1,46
S5780; T1388; D4953-4	0,85	7,09	47,54	33,64	7,46	0,66	1,42
S5780; T1402; D4953-3	0,89	6,91	47,54	33,36	7,56	0,71	1,60
S5780; T1402; D4953-9	0,91	7,26	46,67	33,52	7,90	0,70	1,49
S5780; T1402; D4953-1	0,90	7,20	46,37	33,86	7,91	0,72	1,54

[0117] La tabla 17 proporciona perfiles primarios de ácidos grasos de 3 días de las cepas representativas S5780 transformadas con D4953 (pSZ6004). PSZ6004 expresa el mutante GarmFATA1 (G96A, S111A, V193A) impulsado por PmSAD2-2.

Tabla 18

Identificación de la muestra	Perfil de ácidos grasos						
	C14:0	C16:0	C18:0	C18:1	C18:2	C18:3 α	C20:0
S5780	0,78	8,24	30,24	50,34	6,00	0,79	2,23
S5780; T1402; D4934-20	0,84	7,10	46,71	34,60	7,34	0,65	1,47
S5780; T1402; D4934-15	0,78	6,76	44,01	38,09	6,88	0,65	1,59
S5780; T1402; D4934-24	1,03	10,69	39,82	33,95	11,12	0,71	1,36
S5780; T1402; D4934-14	0,77	6,83	38,68	43,31	6,51	0,71	1,88

Identificación de la muestra	Perfil de ácidos grasos						
	C14:0	C16:0	C18:0	C18:1	C18:2	C18:3 α	C20:0
S5780; T1402; D4934-16	0,75	6,91	35,57	46,50	6,20	0,71	1,92
S5780; T1395; D4982-1	0,00	6,19	39,51	41,35	8,23	0,78	1,92
S5780; T1395; D4982-2	0,03	7,02	35,52	46,24	6,59	0,81	1,89

[0118] La tabla 18 proporciona perfiles primarios de ácidos grasos de 3 días de las cepas representativas S5780 transformadas con D4934 (pSZ5985) y D4982 (pSZ6022). PSZ5985 y pSZ6022 expresan el mutante GarmFATA1 (G108A, S111A, V193A) impulsado por PmSAD2-2 y PmACP-P1, respectivamente.

Tabla 19

Identificación de la muestra	Perfil de ácidos grasos						
	C14:0	C16:0	C18:0	C18:1	C18:2	C18:3 α	C20:0
S5780	0,70	6,98	30,82	51,28	5,80	0,77	2,27
S5780; T1402; D4949-2	0,62	4,54	46,07	38,20	7,44	0,63	1,46
S5780; T1402; D4949-13	0,66	4,57	45,33	38,42	7,85	0,68	1,50
S5780; T1402; D4949-7	0,64	4,61	45,02	39,06	7,55	0,64	1,50
S5780; T1402; D4949-8	0,64	4,62	44,87	39,16	7,51	0,67	1,54
S5780; T1402; D4949-3	0,64	4,88	44,18	39,83	7,18	0,65	1,56

- 5 [0119] La tabla 19 proporciona perfiles primarios de ácidos grasos de 3 días de las cepas representativas S5780 transformadas con D4949 (pSZ5987). PSZ5985 expresa el mutante GarmFATA1 (G96A, G108A, S111A, V193A) impulsado por PmSAD2-2.

LISTADO DE SECUENCIAS INFORMAL

[0120]

SEQ ID NO: 1

Secuencia de aminoácidos de *Cuphea hookeriana* FATB2 ("ChFATB2") de tipo salvaje

MVAAAASSAFFPVPAPGASPKPGKFGNWPSSLSPSFKPKSIPNGGFQVKANDSAHPKANGSAVSLKSGSL
 NTQEDTSSSPPPRTFLHQLPDWSRLLTAITTVFVKSRRPDMHDRKSKRPDMLVDSFGLESTVQDGLVFRQ
 SFSIRSYEIGTDRTASIEITLMNHLQETSLNHCKSTGILLDGFGRTELMCKRDLIWVVIKMQIKVNRYPAW
 GDTVEINTRFSRLGKIGMRDWLISDCNTGEILVRATSAYAMNQKTRRLSKLPYEVHQEIVPLFVDSPV
 IEDSDLKVHKFKVKTGDSIQKGLTPGWNDLDVNQHVSNVKYIGWILESMPTVEVLETQELCSLALEYRREC
 GRDSVLESVTAMDPSKVGVRSQYQHLLRLEDGTAIVNGATEWRPKNAGANGAISTGKTSNGNSVS

SEQ ID NO: 2

CDS de FATB2 específica de C8/C10 de *Cuphea hookeriana*

CTGGATACCATTTTCCCTGCGAAAAACATGGTGGCTGCTGCAGCAAGTTCCGCATTCTTCCCTGTTCCA
GCCCCGGGAGCCTCCCCTAAACCCGGGAAGTTTCGGAATTGGCCCTCGAGCTTGAGCCCTTCCCTTCAAGC
CCAAGTCAATCCCCAATGGCGGATTTTCAGGTTAAGGCAAATGACAGCGCCCATCCAAAGGCTAACGGTTC
TGCAGTTAGTCTAAAGTCTGGCAGCCTCAACACTCAGGAGGACACTTCGTCGTCCCCCTCCTCCTCGGACT
TTCCTTCACCAGTTGCCTGATTGGAGTAGGCTTCTGACTGCAATCACGACCGTGTTCGTGAAATCTAAGA
GGCCTGACATGCATGATCGGAAATCCAAGAGGCCTGACATGCTGGTGGACTCGTTTGGGTGGAGAGTAC
TGTTTCAGGATGGGCTCGTGTTCGACAGAGTTTTTTCGATTAGGTCTTATGAAATAGGCACTGATCGAACG
GCCTCTATAGAGACACTTATGAACCACTTGCAGGAAACATCTCTCAATCATTGTAAGAGTACCGGTATTC
TCCTTGACGGCTTCGGTCGTACTCTTGAGATGTGTAAAAGGGACCTCATTTGGGTGGTAATAAAAAATGCA
GATCAAGGTGAATCGCTATCCAGCTTGGGGCGATACTGTGAGATCAATACCCGGTTCTCCCGGTGGGG
AAAATCGGTATGGGTGCGGATTGGCTAATAAGTGATTGCAACACAGGAGAAATCTTGTAAAGAGCTACGA
GCGCGTATGCCATGATGAATCAAAGACGAGAAGACTCTCAAAACTTCCATACGAGGTTACACAGGAGAT
AGTGCCTCTTTTTGTGCGACTCTCCTGTGATTGAAGACAGTGATCTGAAAGTGCATAAGTTTAAAGTGAAG
ACTGGTGATTCCATTCAAAGGGTCTAACTCCGGGGTGAATGACTTGGATGTCAATCAGCACGTAAGCA
ACGTGAAGTACATTGGGTGGATTCTCGAGAGTATGCCAACAGAAGTTTTGGAGACCCAGGAGCTATGCTC
TCTCGCCCTTGAATATAGCGGGAATGCGGAAGGGACAGTGTGCTGGAGTCCGTGACCGCTATGGATCCC
TCAAAGTTGGAGTCCGTTCTCAGTACCAGCACCTTCTGCGGCTTGAGGATGGGACTGCTATCGTGAACG
GTGCAACTGAGTGGCGGCCGAAGAATGCAGGAGCTAACGGGGCGATATCAACGGGAAAAGACTTCAAATGG
AAACTCGGTCTCTTAGAAGTGTCTCGGAACCTTCCGAGATGTGCATTTCTTTTCTCCTTTTTCATTTTGT
GGTGAGCTGAAAGAAGAGCATGTCGTTGCAATCAGTAAATTGTGTAGTTTCGTTTTTCGCTTTGCTTCGCT
CCTTTGTATAATAATATGGTCAGTCGTCTTTGTATCATTTTCATGTTTTTCAGTTTATTTACGCCATATAAT
TTTT

SEQ ID NO: 3

Secuencia de aminoácidos de Ch FATB2-D3570, el péptido de tránsito de algas de pSZ4689 está subrayado, la etiqueta de epítipo FLAG está en negrita y el residuo P186 está en negrita y en minúsculas

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRASSLSPSFKPKSIPNGGFQVKANDSAHPKA
NGSAVSLKSGSLNTQEDTSSSPPPRTFLHQLPDWSRLLTAITTVFVKSkrpdmhdkSKRPDMLVDS
FGLESTVQDGLVFRQSFIRSIEIGTDRTASIEITLMNHLQETSLNHCKSTGILLDGFGRTP**EMCKRD**
LIWVVIKMQIKVNRYPAGWDTVEINTRFSRLGKIGMRDWLISDCNTGEILVRATSAYAMMNQKTRR
LSKLPIYEVHQEIVPLFVDSPVIEDSDLKVHKFKVKTGDSIQKGLTPGWNDLDVNQHVSNNKYIGWIL
ESMPTEVLETQELCSLALEYRRECGRDSVLESVTAMDPSKVGVRSQYQHLLRLEDGTAIVNGATEWR
PKNAGANGAISTGKTSNGNSVSM**DKYKDHGDYKDHIDYKDDDDK***

SEQ ID NO: 4

Secuencia de aminoácidos de Ch FATB2-D3573, el péptido de tránsito de algas de pSZ4692 está subrayado, la etiqueta del epítipo FLAG está en negrita y el residuo K186 está en negrita y en minúsculas

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRASSLSPSFKPKSIPNGGFQVKANDSAHPKA
NGSAVSLKSGSLNTQEDTSSSPPPRTFLHQLPDWSRLLTAITTVFVKSkrpdmhdkSKRPDMLVDS
FGLESTVQDGLVFRQSFIRSIEIGTDRTASIEITLMNHLQETSLNHCKSTGILLDGFGRTP**kEMCKRD**
LIWVVIKMQIKVNRYPAGWDTVEINTRFSRLGKIGMRDWLISDCNTGEILVRATSAYAMMNQKTRR
LSKLPIYEVHQEIVPLFVDSPVIEDSDLKVHKFKVKTGDSIQKGLTPGWNDLDVNQHVSNNKYIGWIL
ESMPTEVLETQELCSLALEYRRECGRDSVLESVTAMDPSKVGVRSQYQHLLRLEDGTAIVNGATEWR
PKNAGANGAISTGKTSNGNSVSM**DKYKDHGDYKDHIDYKDDDDK***

SEQ ID NO: 5

Secuencia de aminoácidos de Ch FATB2-D3582, el péptido de tránsito de algas de pSZ4702 está subrayado, la etiqueta del epítipo FLAG está en negrita y el residuo A186 está en negrita y en minúsculas

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRASSLSPSFKPKSIPNGGFQVKANDSAHPKA
NGSAVSLKSGSLNTQEDTSSSPPPRTFLHQLPDWSRLLTAITTVFVKSKRPDMHDRKSKRPDMLVDS
FGLESTVQDGLVFRQSFSIRSYEIGTDRTASIEITLMNHLQETSLNHCKSTGILLDGFGRTaEMCKRD
LIWVVIKMQIKVNRYPAGDVTVEINTRFSRLGKIGMRDWLISDCNTGEILVRATSAYAMNQKTRR
LSKLPYEVHQEIVPLFVDSPVIEDSDLKVHKFKVKTGDSIQKGLTPGWNDLDVNQHVSINVKYIGWIL
ESMPTEVLETQELCSLALEYRRECGRDSVLESVTAMDPSKVGVR SQYQHLLRLEDGTAIVNGATEWR
PKNAGANGAISTGKTSNGNSVSMDYKDHDGDYKDHDIDYKDDDDK*

5 **SEQ ID NO: 6**

Secuencia de aminoácidos de Ch FATB2-D3584, el péptido de tránsito de algas de pSZ4704 está subrayado, la etiqueta del epítipo FLAG está en negrita y el residuo Y163 está en negrita y en minúsculas

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRASSLSPSFKPKSIPNGGFQVKANDSAHPKA
NGSAVSLKSGSLNTQEDTSSSPPPRTFLHQLPDWSRLLTAITTVFVKSKRPDMHDRKSKRPDMLVDS
FGLESTVQDGLVFRQSFSIRSYEIGTDRTASIEITLMN^yLQETSLNHCKSTGILLDGFGRTEMCKRD
LIWVVIKMQIKVNRYPAGDVTVEINTRFSRLGKIGMRDWLISDCNTGEILVRATSAYAMNQKTRR
LSKLPYEVHQEIVPLFVDSPVIEDSDLKVHKFKVKTGDSIQKGLTPGWNDLDVNQHVSINVKYIGWIL
ESMPTEVLETQELCSLALEYRRECGRDSVLESVTAMDPSKVGVR SQYQHLLRLEDGTAIVNGATEWR
PKNAGANGAISTGKTSNGNSVSMDYKDHDGDYKDHDIDYKDDDDK*

SEQ ID NO: 7

Secuencia de aminoácidos de Ch FATB2-D3588, el péptido de tránsito de algas de pSZ4709 está subrayado, la etiqueta del epítipo FLAG está en negrita y el residuo F163 está en negrita y en minúsculas

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRASSLSPSFKPKSIPNGGFQVKANDSAHP¹⁵KA
NGSAVSLKSGSLNTQEDTSSSPPPRTFLHQLPDWSRLLTAITTVFVKSKRPDMHDRKSKRPDMLVDS
FGLESTVQDGLVFRQSFSIRSYEIGTDRTASIEITLMN^fLQETSLNHCKSTGILLDGFGRTEMCKRD
LIWVVIKMQIKVNRYPAGDVTVEINTRFSRLGKIGMRDWLISDCNTGEILVRATSAYAMNQKTRR
LSKLPYEVHQEIVPLFVDSPVIEDSDLKVHKFKVKTGDSIQKGLTPGWNDLDVNQHVSINVKYIGWIL
ESMPTEVLETQELCSLALEYRRECGRDSVLESVTAMDPSKVGVR SQYQHLLRLEDGTAIVNGATEWR
PKNAGANGAISTGKTSNGNSVSMDYKDHDGDYKDHDIDYKDDDDK*

SEQ ID NO: 8

- 20 **Secuencia de aminoácidos de tipo salvaje Ch FATB2-D3598, el péptido de tránsito de algas de pSZ4243 está subrayado, la etiqueta del epítipo FLAG está en negrita y los residuos H163 y L186 están en negrita y en minúsculas**

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRASSLSPSFKPKSIPNGGFQVKANDSAHPKA
NGSAVSLKSGSLNTQEDTSSSPPPRTFLHQLPDWSRLLTAITTVFVKSKRPDMHDRKSKRPDMLVDS
FGLESTVQDGLVFRQSFSIRSYEIGTDRTASIEITLMN^hLQETSLNHCKSTGILLDGFGRTL^lEMCKRD
LIWVVIKMQIKVNRYPAGDVTVEINTRFSRLGKIGMRDWLISDCNTGEILVRATSAYAMNQKTRR
LSKLPYEVHQEIVPLFVDSPVIEDSDLKVHKFKVKTGDSIQKGLTPGWNDLDVNQHVSINVKYIGWIL
ESMPTEVLETQELCSLALEYRRECGRDSVLESVTAMDPSKVGVR SQYQHLLRLEDGTAIVNGATEWR
PKNAGANGAISTGKTSNGNSVSMDYKDHDGDYKDHDIDYKDDDDK*

SEQ ID NO: 9

- 25 **Secuencia de nucleótidos del vector de cribado de la expresión de ChFATB2, usando pSZ4243 como ejemplo. Los brazos de homología 5' y 3' que permiten la integración dirigida al locus Thi4 están indicados con minúsculas; el promotor CrTUB2 está indicado en mayúsculas y cursiva, que impulsa la expresión del**

gen de resistencia a neomicina indicado con minúsculas y cursiva, seguido por el terminador PmPGH 3'UTR destacado en mayúsculas. El promotor PmUAPA1 (indicado con texto en **negrita**) impulsa la expresión del ChFATB2 con optimización de codones (indicado con un texto en minúsculas y **negrita**) y se termina con la CvNR 3'UTR indicada con **negrita**, en minúsculas y subrayada. Los sitios de clonación de restricción y los fragmentos de ADN espaciadores están indicados con letras subrayadas y en mayúscula.

5

ccctcaactgcgacgctgggaaccttctccgggcaggcgatgtgcggtgggtttgcctccttggcacg
gctctacaccgctcgagtacgccatgagggcggtgatggctgtgtcggttgccacttcgtccagagacg
gcaagtcgtccatcctctgctgtgtggcgcgacgctgcagcagtcctctgcagcagatgagcgtg
actttggccatttcacgcactcgagtgtacacaatccatttttcttaagcaaatgactgctgattg
accagatactgtaacgctgatttcgctccagatcgcacagatagcgaccatgttgctgctctgaaa
atctggattccgaattcgaccctggcgctccatccatgcaacagatggcgacacttggtacaattcc
tgtcacccatcggcattggagcaggtccacttagattcccgatcacccacgcacatctcgctaatagt
cattcgttcgtgtcttcgatcaatctcaagtgagtggtgcatggatcttggttgacgatgcggtatgg
gtttgcgccgctggctgcagggctctgcccaaggcaagctaaccagctcctctccccgacaatactc
tcgcaggcaaaagccggtcacttgcttccagattgccataaactcaattatggcctctgtcatgcc
atccatgggtctgatgaatggtcacgctcgtgtcctgacggttccccagcctctggcggtccccctgcc
ccgcccaccagcccacgcgcgcggcagtcgctgccaaggctgtctcggaGGTACCCTTTCTTGCGC
TATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGA
TGATGCTTCGACCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGC
GCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAAC

ACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCC
TCTTCCTCTTCGTTTCAGTCACAACCCGCAAACCTCTAGAATATCAatgatcgagcaggacggcctcc
acgccggctccccgcgcctgggtggagcgctgttcggctacgactgggccagcagaccatcgg
ctgctccgacgcgcctgttcgcctgtccgcccagggccgcccgtgctgttcgtgaagaccgac
ctgtccggcgccctgaacgagctgcaggacgaggccgcccgcctgtcctggctggccaccaccggcg
tgccctgcgcgcctgtgctggacgtgggtgaccgaggccggccggaactggctgctgctggcgagg
gcccggccaggacctgctgtcctcccacctggccccgcgcgagaagggtgtccatcatggccgacgcc
atgcgcgcctgcacaccctggacccccgccacctgccccttcgaccaccaggccaagcaccgcacg
agcgcgcccgcaaccgcacatggaggccggcctgggtggaccaggacgacctggacgaggagcaccaggg
cctggccccgcgcgagctgttcgcccgcctgaaggcccgcatgcccgcgagcaggacctgggtgggtg
acccacggcgacgcctgcctgcccacatcatgggtggagaacggccgcttctccggcttcatcgact
gcggccgcctgggcgtggccgaccgctaccaggacatcgccctggccacccgcgacatcgccgagga
gtggggcgcgagtgggccgaccgcttctcctgggtgctgtacggcatcgccgcccccgactcccagcgc
atcgcccttctaccgcctgctggacgagttcttctgaCAATTGACGCCCGCGCGGCGCACCTGACCTG
TTCTCTCGAGGGCGCCTGTTCTGCCTTGCGAAACAAGCCCTGGAGCATGCGTGCATGATCGTCTCT
GGCGCCCCGCGCGCGGTTTGTGCCCCGCGCGGCGCGCGGCGCGGGGCGCATTGAAATTGTTG
CAAACCCACCTGACAGATTGAGGGCCAGGCAGGAAGGCGTTGAGATGGAGGTACAGGAGTCAAGT
AACTGAAAGTTTTTATGATAACTAACAACAAAGGGTCGTTTCTGGCCAGCGAATGACAAGAACAAGA
TTCCACATTTCCGTGTAGAGGCTTGCCATCGAATGTGAGCGGGCGGGCCGCGGACCCGACAAAACCC
TTACGACGTGGTAAGAAAAACGTGGCGGGCACTGTCCTGTAGCCTGAAGACCAGCAGGAGACGATC
GGAAGCATCACAGCACAGGATCCCGCGTCTCGAACAGAGCGCGCAGAGGAACGCTGAAGGTCTCGCC
TCTGTGCGACCTCAGCGCGGCATACACCACAATAACCACCTGACGAATGCGCTTGGTTCTTCGTCCA
TTAGCGAAGCGTCCGGTTCACACACGTGCCACGTGGCGAGGTGGCAGGTGACAATGATCGGTGGAG
CTGATGGTTCGAAACGTTTCACAGCCTAGGGATATCATAGCGACTGCTACCCCCGACCATGTGCCGAG
GCAGAAATTATATACAAGAAGCAGATCGCAATTAGGCACATCGCTTTGCATTATCCACACACTATTC
ATCGCTGCTGCGGCAAGGCTGCAGAGTGTATTTTTGTGGCCAGGAGCTGAGTCCGAAGTCGACGCG
ACGAGCGGCGCAGGATCCGACCCCTAGACGAGCACTGTCAATTTCCAAGCACGCAGCTAAATGCGCT
GAGACCGGGTCTAAATCATCCGAAAAGTGTCAAATGGCCGATTGGGTTTCGCTTAGGACAATGCGCT
GCGGATTCGCTCGAGTCCGCTGCCGGCCAAAAGGCGGTGGTACAGGAAGGCGCACGGGGCCAACCT
GCGAAGCCGGGGGCCCCGAACGCCGACCGCCGGCCTTCGATCTCGGGTGTCCCCCTCGTCAATTTCT
CTCTCGGGTGCAGCCACGAAAGTCGTGACGCAGGTCACGAAATCCGGTTACGAAAAACGCAGGTCTT
CGCAAAAACGTGAGGGTTTCGCGTCTCGCCCTAGCTATTCTGTATCGCCGGGTGACACCCACGTGCAG
AAAAGCCCTTGAATAACCCGGGACCGTGGTTACCGCGCCGCTGCACCAGGGGGCTTATATAAGCCC
ACACCACACCTGTCTCACCACGCATTTCTCCAACTCGCGACTTTTCGGAAGAAATTGTTATCCACCT
AGTATAGACTGCCACCTGCAGGACCTTGTGTCTTGCAGTTTGTATTGGTCCC GGCGCTCGAGCACGA
CAGATCTGGGCTAGGGTTGGCCTGGCCGCTCGGCACCTCCCTTTAGCCGCGCGCATCCGCGTTCCAG
AGGTGCGATTTCGGTGTGTGGAGCATTGTCTATGCGCTTGTGGGGTTCGTTCCGTGCGCGGGCGGGTCCG
CCATGGGCGCCGACCTGGGCCCTAGGGTTTGTTCGCGGCAAGCGAGCCCCCTCTACCTCGTTCGCC
CCCCGCGATTCCCTCTCTCTTGCAGCCACTAGTAACAatggccaccgcacccatttctcggcgcttc
aatgcccgcctgcggcgacctgcgtcgctcggcggggtccgggccccggcgcccagcgaggccccctcc
ccgtgcgcggggcgcgccctccagccctgagccctccttcaagcccaagtcacatcccccaacggcggtt
ccaggtgaaggccaacgacagcgcccaacccaaggccaacggctccgcctgagcctgaagagcggc
agcctgaacacccaggaggacacctcctccagccccccccccgcacettcctgcaccagctgcccg
actggagccgcctgctgaccgccatcaccacccgtgttcgtgaagtccaagcgccccgacatgcacga
ccgaagtccaagcgccccgacatgctgggtggacagcttcggcctggagtcacccgtgcaggacggc
ctgggtgttcggccagtccttctccatccgctcctacgagatcggcaccgacccgacccgcagcatcg
agaccctgatgaaccacctgcaggagacctccctgaaccactgcaagagcaccggcatcctgctgga
cggcttcggccgcaccctggagatgtgcaagcgcgacctgatctgggtgggtgatcaagatgcagatc
aagggtgaaccgcctacccccgcctggggcgacaccgtggagatcaacacccgccttcagccgcctgggca
agatcgggcatgggcccgcgactgggtgatctccgactgcaacacccggcgagatcctgggtgcgcgccac
cagcgccctacgcatgatgaaccagaagacccgcgcctgtccaagctgccctacgaggtgcaccag
gagatcggtgccccctgttcgtggacagccccgtgatcgaggactccgacctgaagggtgcacaagtca
aggtgaagaccggcgacagcatccagaagggcctgacccccggctggaacgacctggacgtgaacca
gcacgtgtccaacgtgaagtacatcggttggaatcctggagagcatgccaccgaggtgctggagacc
caggagctgtgtctccctggccctggagtaccgcccgcgagtgcgccgcgactccgtgctggagagcg
tgaccgccatggaccccagcaagggtgggcgtgcgctcccagtagcagcactgctgcgcctggagga

cggcaccgccatcgatgaacggcgccaccgagtgggcgccccagaacgcccggcgccaacggcgccatc
tccaccggcaagaccagcaacggcaactccgtgtccatggactacaaggaccacgacggcgactaca
aggaccacgacatcgactacaaggacgacgacgacaagtgaCTCGAGgcagcagcagctcgatagt
atcgacacactctggacgctgggtcggtgtgatggactgttgcgcccacacttgctgccttgacctgtg
aatatccctgccgctttttatcaaacagcctcagtggtgtttgatcttgtgtgtacgcgcttttgcgag
ttgctagctgcttgtgctatattgcgaataaccacccccagcatcccttccctcgtttcatatcgctt
gcatacccaaccgcaacttatttacgctgtcctgctatccctcagcgtgctcctgctcctgctcact
gccccctgcacagccttggtttgggctccgctgtattctcctggtagtgcaacctgtaaaccagca
ctgcaatgctgatgcacgggaagtagtggttgatgggaacacaaatgga**AAGCTGTATAGGGATAACAG**
GGTAATGAGCTCcagcgccatgccacgcccccttgatggcttcaagtacgattacgggtgttgattgt
gtgtttgttgcttagtggtcatggtttagaataatacacttgatttcttgctcacggcaatctcggc
ttgtccgcaggttcaacccccatttcggagctctcaggtcagccgcgcaatgaccagccgctacttcaa
ggacttgcacgacaacgccgaggtgagctatgtttaggacttgattggaaattgtcgtcgacgcata
ttcgcgctccgcgacagcacccaagcaaaatgtcaagtgcgttccgatttgctccgcaggtcgatg
ttgtgatcgctcggcgcggatccgcgggtctgtcctgcgcttacgagctgaccaagcaccttgacgt
ccgggtacgcgagctgagattcgattagacataaattgaagattaaacccgtagaaaaatttgatgg
tcgcgaaactgtgctcgattgcaagaaattgatcgtcctccactccgcaggtcgccatcatcgagca
ggcggttgctcccgccggcgccgctggctggggggacagctgttctcggccatgtgtgtacgtaga
aggatgaatttcagctggttttcgttgacagctgtttgtgcatgatttgtttcagactattgttga
atgttttttagatttcttaggatgcattgttctgcatgcgact

SEQ ID NO: 10

Secuencia de aminoácidos de la Cpal FATB1 de tipo salvaje (D3004, pSZ4241). El péptido de tránsito de algas está subrayado, la etiqueta del epitopo FLAG está en negrita y el residuo M230 está en negrita y en minúsculas.

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRASSSLSPSLKPKSIPNGGFQVKANASAHPK
ANGSAVTLKSGSLNTQEDTLSSSPPPRAFFNQLPDWSMLLTAITTVFVAPEKRWTMFDRKSKRPNML
MDSFGLERVVQDGLVFRQSFSIRSYEICADRTASIETVMNHVQETSLNQCKSIGLLDDGFGRSPEMC
KRDLIWVVTRMKIMVNRYPTWGDTIEVSTWLSQSGKIG**m**GRDWLISDCNTGEILVRATSVYAMMNQK
TRRFSKLPHEVRQEFAPHFLDSPPAIEDNDGKLQKFDVKTGDSIRKGLTPGWYDLVDVNQHVSNVKYI
GWILESMPTEVLETQELCSLTLEYRRECGRDSVLESVTSMDPSKVGDRFQYRHLRLEDGADIMKGR
TEWRPKNAGTNGAISTGKT**MDYKDHDGDYKDHDIDYKDDDDK**

SEQ ID NO: 11

Secuencia de aminoácidos de la Chook FATB2 (D3042, pSZ4243) de tipo salvaje. El péptido de tránsito de algas está subrayado, la etiqueta del epitopo FLAG está en negrita y el residuo M228 está en negrita y en minúsculas.

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRASSSLSPSFKPKSIPNGGFQVKANDSAHPKA
NGSAVSLKSGSLNTQEDTSSSPPPRTFLHQLPDWSRLLTAITTVFVKSKRPMDHDKSKRPDMLVDS
FGLESTVQDGLVFRQSFSIRSYEIGTDRTASIETLMNHLQETSLNHCKSTGILLDDGFGRTLEMCKRD
LIWVVIKMQIKVNRYPAWGDTVINTRFSRLGKIG**m**GRDWLISDCNTGEILVRATSAYAMMNQKTRR
LSKLPYEVHQEIVPLFVDSPVIEDSDLKVHKFKVKTGDSIQKGLTPGWNLDVDVNQHVSNVKYIGWIL
ESMPTEVLETQELCSLALEYRRECGRDSVLESVTAMDPSKVGRSQYQHLLRLEDGTAIVNGATEWR
PKNAGANGAISTGKTSNGNSVS**MDYKDHDGDYKDHDIDYKDDDDK**

SEQ ID NO: 12

Secuencia de aminoácidos de la Ca FATB1 de tipo salvaje (D3456, pSZ4532). El péptido de tránsito de algas está subrayado, la etiqueta del epítipo FLAG está en negrita y el residuo K228 está en negrita y en minúsculas.

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRAAINSRAHPKANGSAVSLKSGSLNTQEDTSSS
 PPPRTFLHQLPDWSRLLTAITTVFVKSKRPDMHDRKSKRPDMLMDSFGLESIVQEGLEFRQSFSIRS
 YEIGTDRTASIEITLMNYLQETSLNHCKSTGILLDGFGRTPEMCKRDLIWVVTMKMIKVNRYPAWGDT
 VEINTWFSRLGKIG**k**GRDWLISDCNTGEILIRATSAYATMNQKTRRLSKLPYEVHQEIAPLFVDSPP
 VIEDNDLKLHKFEVKTGDSIHKGLTPGWNDLDVNQHVSINVKYIGWILESMPTVEVLETQELCSLALEY
 RRECGRDSVLESVTAMDPTKVGGRSQYQHLLRLEDGTDIVKCRTEWRPKNPGANGAISTGKTSNGNS
 VS**MDYKDHDGDYKDHDIDYKDDDDK**

5 **SEQ ID NO: 13**

Secuencia de aminoácidos de *Cuphea palustris* FATB1

MVAAAASSACFPVPSPGASPKPGKLGWSSSLSPSLKPKSIPNGGFQVKANASAHPKANG
 SAVTLKSGSLNTQEDTLSSPPPRAFFNQLPDWSMLLTAITTVFVAPEKRWTMFDRKSKR
 PNMLMDSFGLERVVQDGLVFRQSFSIRSIEICADRTASIEVMNHVQETSLNQCKSIGLL
 DDGFGSPSEMCKRDLIWVVTMIMVNRYPTWGDITIEVSTWLSQSGKIGMRDWLISDCN
 TGEILVRATSVYAMNQKTRRFSKLPHEVRQEFAPHFLDSPPAIEDNDGKLQKFDVKTGD
 SIRKGLTPGWYDLVDVNQHVSINVKYIGWILESMPTVEVLETQELCSLTLEYRRECGRDSVLE
 SVTSMGPSKVGDRFQYRHLLRLEDGADIMKGRTEWRPKNAGTNGAISTGKT

SEQ ID NO: 14

Secuencia de aminoácidos de *Cuphea avigera* FATB1

MVAAAASSAFFSVVPVPGTSPKPGKFRIWPSSSLSPSFKPKPIPNGGLQVKANSRAHPKANG
 SAVSLKSGSLNTQEDTSSPPPRFTFLHQLPDWSRLLTAITTVFVKSKRPDMHDRKSKRPD
 MLMDSFGLESIVQEGLEFRQSFSIRSIEIGTDRTASIEITLMNYLQETSLNHCKSTGILLD
 GFGRTPEMCKRDLIWVVTMKMIKVNRYPAWGDTVEINTWFSRLGKIGKGRDWLISDCNTG
 EILIRATSAYATMNQKTRRLSKLPYEVHQEIAPLFVDSPPVIEDNDLKLHKFEVKTGDSI
 HKGLTPGWNDLDVNQHVSINVKYIGWILESMPTVEVLETQELCSLALEYRRECGRDSVLESV
 TAMDPTKVGGRSQYQHLLRLEDGTDIVKCRTEWRPKNPGANGAISTGKTSNGNSVS

SEQ ID NO: 15

Secuencia de aminoácidos de gen parental de tipo salvaje Gm FATA; D3997, pSZ5083. El péptido de tránsito de algas está subrayado y la etiqueta del epítipo FLAG está en negrita y mayúsculas.

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVGRGAIPPRIIVVSSSSSKVNPLKTEAVVSSGLA
 DRLRLGSLTEDGLSYKEKFIVRCYEVGINKTATVETIANLLQEVGCNHAQSVGYSTGGFSTTPTMRK
 LRLLIWVTARMHIEIYKPAWSDVVEIESWGQEGKIGTRDWILRDYATGQVIGRATSKWVMNQDT
 RRLQKVDVDVRDEYLVHCPRELRLAFPEENNSSLKKISKLEDPSQYSKLGLVPRRADLDMNQHVNNV
 TYIGWVLESMPQEIIDTHELQTITLDYRRECQHDDVVDLSLTSPEPSEDAEAVFNHNGTNGSANVSAN
 DHGCRNFLHLLRLSGNGLEINRGTEWRKKPTR**MDYKDHDGDYKDHDIDYKDDDDK**

SEQ ID NO: 16

Secuencia de aminoácidos del gen mutante Gm FATA S111A, V193A; D3998, pSZ5084. El péptido de tránsito de algas está subrayado, la etiqueta del epítipo FLAG está en negrita y mayúsculas y los residuos S111A, V193A están en negrita y minúsculas.

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRAIPPRIIVVSSSSSKVNPLKTEAVVSSGLA
DRLRLGSLTEDGLSYKEKFIVRCYEVGINKTATVETIANLLQEVGCNHAQSVGYSTGGF**a**TTPTMRK
 LRLIWVTARMHIEIYKYPAWSDVVEIESWGQEGKIGTRRDWILRDYATGQVIGRATSKWVMNQDT
 RRLQKVD**a**DVRDEYLVHCPRELRLAFPEENNSSLKKISKLEDPSQYSKLGLVPRRADLDMNQHVNNV
 TYIGWVLESMPQEIIDTHELQTITLDYRRECQHDDVVDLSLTSPEPSEDAEAVFNHNGTNGSANVSAN
 DHGCRNFLHLLRLSGNGLEINRGRTWRKKPTR**MDYKDHDGDYKDHDIDYKDDDDK**

5 **SEQ ID NO: 17**

Secuencia de aminoácidos del gen mutante Gm FATA S111V, V193A; D3999; pSZ5085. El péptido de tránsito de algas está subrayado, la etiqueta del epítipo FLAG está en negrita y mayúsculas y los residuos S111V, V193A están en negrita y minúsculas.

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRAIPPRIIVVSSSSSKVNPLKTEAVVSSGLA
DRLRLGSLTEDGLSYKEKFIVRCYEVGINKTATVETIANLLQEVGCNHAQSVGYSTGGF**v**TTPTMRK
 LRLIWVTARMHIEIYKYPAWSDVVEIESWGQEGKIGTRRDWILRDYATGQVIGRATSKWVMNQDT
 RRLQKVD**a**DVRDEYLVHCPRELRLAFPEENNSSLKKISKLEDPSQYSKLGLVPRRADLDMNQHVNNV
 TYIGWVLESMPQEIIDTHELQTITLDYRRECQHDDVVDLSLTSPEPSEDAEAVFNHNGTNGSANVSAN
 DHGCRNFLHLLRLSGNGLEINRGRTWRKKPTR**MDYKDHDGDYKDHDIDYKDDDDK**

10 **SEQ ID NO: 18. - secuencia de aminoácidos del gen mutante Gm FATA G96A; D4000, pSZ5086. El péptido de tránsito de algas está subrayado, la etiqueta del epítipo FLAG está en negrita y mayúsculas y el residuo G96A está en negrita y minúsculas.**

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRAIPPRIIVVSSSSSKVNPLKTEAVVSSGLA
DRLRLGSLTEDGLSYKEKFIVRCYEVGINKTATVETIANLLQEV**a**CNHAQSVGYSTGGFSTTPTMRK
 LRLIWVTARMHIEIYKYPAWSDVVEIESWGQEGKIGTRRDWILRDYATGQVIGRATSKWVMNQDT
 RRLQKVDVDVRDEYLVHCPRELRLAFPEENNSSLKKISKLEDPSQYSKLGLVPRRADLDMNQHVNNV
 TYIGWVLESMPQEIIDTHELQTITLDYRRECQHDDVVDLSLTSPEPSEDAEAVFNHNGTNGSANVSAN
 DHGCRNFLHLLRLSGNGLEINRGRTWRKKPTR**MDYKDHDGDYKDHDIDYKDDDDK**

15 **SEQ ID NO: 19 - secuencia de aminoácidos del gen mutante Gm FATA G96T; D4001, pSZ5087. El péptido de tránsito de algas está subrayado, la etiqueta del epítipo FLAG está en negrita y mayúsculas y el residuo G96T está en negrita y minúsculas.**

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRAIPPRIIVVSSSSSKVNPLKTEAVVSSGLA
DRLRLGSLTEDGLSYKEKFIVRCYEVGINKTATVETIANLLQEV**t**CNHAQSVGYSTGGFSTTPTMRK
 LRLIWVTARMHIEIYKYPAWSDVVEIESWGQEGKIGTRRDWILRDYATGQVIGRATSKWVMNQDT
 RRLQKVDVDVRDEYLVHCPRELRLAFPEENNSSLKKISKLEDPSQYSKLGLVPRRADLDMNQHVNNV
 TYIGWVLESMPQEIIDTHELQTITLDYRRECQHDDVVDLSLTSPEPSEDAEAVFNHNGTNGSANVSAN
 DHGCRNFLHLLRLSGNGLEINRGRTWRKKPTR**MDYKDHDGDYKDHDIDYKDDDDK**

20 **SEQ ID NO: 20 - secuencia de aminoácidos del gen mutante Gm FATA G96V; D4002, pSZ5088. El péptido de tránsito de algas está subrayado, la etiqueta del epítipo FLAG está en negrita y mayúsculas y el residuo G96V está en negrita y minúsculas.**

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRAIPPRIIVVSSSSSKVNPLKTEAVVSSGLA
DRLRLGSLTEDGLSYKEKFIVRCYEVGINKTATVETIANLLQEV**v**CNHAQSVGYSTGGFSTTPTMRK
 LRLIWVTARMHIEIYKYPAWSDVVEIESWGQEGKIGTRRDWILRDYATGQVIGRATSKWVMNQDT
 RRLQKVDVDVRDEYLVHCPRELRLAFPEENNSSLKKISKLEDPSQYSKLGLVPRRADLDMNQHVNNV
 TYIGWVLESMPQEIIDTHELQTITLDYRRECQHDDVVDLSLTSPEPSEDAEAVFNHNGTNGSANVSAN
 DHGCRNFLHLLRLSGNGLEINRGRTWRKKPTR**MDYKDHDGDYKDHDIDYKDDDDK**

SEQ ID NO: 21 - secuencia de aminoácidos del gen mutante Gm FATA G108A; D4003, pSZ5089. El péptido de tránsito de algas está subrayado, la etiqueta del epítipo FLAG está en negrita y mayúsculas y el residuo G108A está en negrita y minúsculas.

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRAIPPRIIVVSSSSSKVNPLKTEAVVSSGLA
DRLRLGSLTEDGLSYKEKFIVRCYEVGINKTATVETIANLLQEVGCNHAQSVGYST**a**GFSTTPTMRK
 LRLLIWVTARMHIEIYKYPAWSDVVEIESWGQEGKIGTRRDWILRDYATGQVIGRATSKWVMNQDT
 RRLQKVDVDVRDEYLVHCPRELRLAFPEENNSSLKKISKLEDPSQYSKLGVPRRADLDMNQHVNNV
 TYIGWVLESMPQEI IDTHELQTITLDYRRECQHDVVDVSLTSPEPSEDAEAVFNHNGTNGSANVSAN
 DHGCRNFLHLLRLSGNGLEINRGRTWRKKPTR**MDYKDHDGDYKDHDIDYKDDDDK**

- 5 **SEQ ID NO: 22 - secuencia de aminoácidos del gen mutante Gm FATA L91F; D4004, pSZ5090. El péptido de tránsito de algas está subrayado, la etiqueta del epítipo FLAG está en negrita y mayúsculas y el residuo L91F está en negrita y minúsculas.**

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRAIPPRIIVVSSSSSKVNPLKTEAVVSSGLA
DRLRLGSLTEDGLSYKEKFIVRCYEVGINKTATVETIAN**f**LQEVGCNHAQSVGYSTGGFSTTPTMRK
 LRLLIWVTARMHIEIYKYPAWSDVVEIESWGQEGKIGTRRDWILRDYATGQVIGRATSKWVMNQDT
 RRLQKVDVDVRDEYLVHCPRELRLAFPEENNSSLKKISKLEDPSQYSKLGVPRRADLDMNQHVNNV
 TYIGWVLESMPQEI IDTHELQTITLDYRRECQHDVVDVSLTSPEPSEDAEAVFNHNGTNGSANVSAN
 DHGCRNFLHLLRLSGNGLEINRGRTWRKKPTR**MDYKDHDGDYKDHDIDYKDDDDK**

- 10 **SEQ ID NO: 23 - secuencia de aminoácidos del gen mutante Gm FATA L91K; D4005, pSZ5091. El péptido de tránsito de algas está subrayado, la etiqueta del epítipo FLAG está en negrita y mayúsculas y el residuo L91K está en negrita y minúsculas.**

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRAIPPRIIVVSSSSSKVNPLKTEAVVSSGLA
DRLRLGSLTEDGLSYKEKFIVRCYEVGINKTATVETIAN**k**LQEVGCNHAQSVGYSTGGFSTTPTMRK
 LRLLIWVTARMHIEIYKYPAWSDVVEIESWGQEGKIGTRRDWILRDYATGQVIGRATSKWVMNQDT
 RRLQKVDVDVRDEYLVHCPRELRLAFPEENNSSLKKISKLEDPSQYSKLGVPRRADLDMNQHVNNV
 TYIGWVLESMPQEI IDTHELQTITLDYRRECQHDVVDVSLTSPEPSEDAEAVFNHNGTNGSANVSAN
 DHGCRNFLHLLRLSGNGLEINRGRTWRKKPTR**MDYKDHDGDYKDHDIDYKDDDDK**

- 15 **SEQ ID NO: 24 - secuencia de aminoácidos del gen mutante Gm FATA L91S; D4006, pSZ5092. El péptido de tránsito de algas está subrayado, la etiqueta del epítipo FLAG está en negrita y mayúsculas y el residuo L91S está en negrita y minúsculas.**

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRAIPPRIIVVSSSSSKVNPLKTEAVVSSGLA
DRLRLGSLTEDGLSYKEKFIVRCYEVGINKTATVETIAN**s**LQEVGCNHAQSVGYSTGGFSTTPTMRK
 LRLLIWVTARMHIEIYKYPAWSDVVEIESWGQEGKIGTRRDWILRDYATGQVIGRATSKWVMNQDT
 RRLQKVDVDVRDEYLVHCPRELRLAFPEENNSSLKKISKLEDPSQYSKLGVPRRADLDMNQHVNNV
 TYIGWVLESMPQEI IDTHELQTITLDYRRECQHDVVDVSLTSPEPSEDAEAVFNHNGTNGSANVSAN
 DHGCRNFLHLLRLSGNGLEINRGRTWRKKPTR**MDYKDHDGDYKDHDIDYKDDDDK**

SEQ ID NO: 25 - secuencia de aminoácidos del gen mutante Gm FATA G108V; D4007, pSZ5093. El péptido de tránsito de algas está subrayado, la etiqueta del epítipo FLAG está en negrita y mayúsculas y el residuo G108V está en negrita y minúsculas.

- 20 MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRAIPPRIIVVSSSSSKVNPLKTEAVVSSGLA
DRLRLGSLTEDGLSYKEKFIVRCYEVGINKTATVETIANLLQEVGCNHAQSVGYST**v**GFSTTPTMRK
 LRLLIWVTARMHIEIYKYPAWSDVVEIESWGQEGKIGTRRDWILRDYATGQVIGRATSKWVMNQDT
 RRLQKVDVDVRDEYLVHCPRELRLAFPEENNSSLKKISKLEDPSQYSKLGVPRRADLDMNQHVNNV
 TYIGWVLESMPQEI IDTHELQTITLDYRRECQHDVVDVSLTSPEPSEDAEAVFNHNGTNGSANVSAN
 DHGCRNFLHLLRLSGNGLEINRGRTWRKKPTR**MDYKDHDGDYKDHDIDYKDDDDK**

SEQ ID NO: 26 - secuencia de aminoácidos del gen mutante Gm FATA T156F; D4008, pSZ5094. El péptido de tránsito de algas está subrayado, la etiqueta del epítipo FLAG está en negrita y mayúsculas y el residuo T156F está en negrita y minúsculas.

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRAIPPRIIVVSSSSSKVNPLKTEAVVSSGLA
DRLRLGSLTEDGLSYKEKFIVRCYEVGINKTATVETIANLLQEVGCNHAQSVGYSTGGFSTTPTMRK
 LRLLIWTARMHIEIYKYPAWSVDVVEIESWGQEGGKIG**f**RRDWILRDYATGQVIGRATSKWVMNNDT
 RRLQKVDVDVRDEYLVHCPRELRLAFPEENNSSLKKISKLEDPQSQYSLGLVPRRADLDMNQHVNNV
 TYIGWVLESMPQEIIDTHELQTITLDYRRECQHDVVDVSLTSPEPSEDAEAVFNHNGTNGSANVSAN
 DHGCRNFLHLLRLSGNGLEINRGRTWRKKPTR**MDYKDHDGDYKDHDIDYKDDDDK**

- 5 **SEQ ID NO: 27 - secuencia de aminoácidos del gen mutante Gm FATA T156A; D4009, pSZ5095. El péptido de tránsito de algas está subrayado, la etiqueta del epítipo FLAG está en negrita y mayúsculas y el residuo T156A está en negrita y minúsculas.**

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRAIPPRIIVVSSSSSKVNPLKTEAVVSSGLA
DRLRLGSLTEDGLSYKEKFIVRCYEVGINKTATVETIANLLQEVGCNHAQSVGYSTGGFSTTPTMRK
 LRLLIWTARMHIEIYKYPAWSVDVVEIESWGQEGGKIG**a**RRDWILRDYATGQVIGRATSKWVMNNDT
 RRLQKVDVDVRDEYLVHCPRELRLAFPEENNSSLKKISKLEDPQSQYSLGLVPRRADLDMNQHVNNV
 TYIGWVLESMPQEIIDTHELQTITLDYRRECQHDVVDVSLTSPEPSEDAEAVFNHNGTNGSANVSAN
 DHGCRNFLHLLRLSGNGLEINRGRTWRKKPTR**MDYKDHDGDYKDHDIDYKDDDDK**

- 10 **SEQ ID NO: 28 - secuencia de aminoácidos del gen mutante Gm FATA T156K; D4010, pSZ5096. El péptido de tránsito de algas está subrayado, la etiqueta del epítipo FLAG está en negrita y mayúsculas y el residuo T156K está en negrita y minúsculas.**

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRAIPPRIIVVSSSSSKVNPLKTEAVVSSGLA
DRLRLGSLTEDGLSYKEKFIVRCYEVGINKTATVETIANLLQEVGCNHAQSVGYSTGGFSTTPTMRK
 LRLLIWTARMHIEIYKYPAWSVDVVEIESWGQEGGKIG**k**RRDWILRDYATGQVIGRATSKWVMNNDT
 RRLQKVDVDVRDEYLVHCPRELRLAFPEENNSSLKKISKLEDPQSQYSLGLVPRRADLDMNQHVNNV
 TYIGWVLESMPQEIIDTHELQTITLDYRRECQHDVVDVSLTSPEPSEDAEAVFNHNGTNGSANVSAN
 DHGCRNFLHLLRLSGNGLEINRGRTWRKKPTR**MDYKDHDGDYKDHDIDYKDDDDK**

- 15 **SEQ ID NO: 29 - secuencia de aminoácidos del gen mutante Gm FATA T156V; D4011, pSZ5097. El péptido de tránsito de algas está subrayado, la etiqueta del epítipo FLAG está en negrita y mayúsculas y el residuo T156V está en negrita y minúsculas.**

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRAIPPRIIVVSSSSSKVNPLKTEAVVSSGLA
DRLRLGSLTEDGLSYKEKFIVRCYEVGINKTATVETIANLLQEVGCNHAQSVGYSTGGFSTTPTMRK
 LRLLIWTARMHIEIYKYPAWSVDVVEIESWGQEGGKIG**v**RRDWILRDYATGQVIGRATSKWVMNNDT
 RRLQKVDVDVRDEYLVHCPRELRLAFPEENNSSLKKISKLEDPQSQYSLGLVPRRADLDMNQHVNNV
 TYIGWVLESMPQEIIDTHELQTITLDYRRECQHDVVDVSLTSPEPSEDAEAVFNHNGTNGSANVSAN
 DHGCRNFLHLLRLSGNGLEINRGRTWRKKPTR**MDYKDHDGDYKDHDIDYKDDDDK**

- 20 **SEQ ID NO: 30 - secuencia de nucleótidos del vector de expresión del gen parental de tipo salvaje GmFATA (D3997, pSZ5083). Los brazos de homología 5' y 3' que permiten la integración dirigida al locus Thi4 están indicados con minúsculas; el promotor CrTUB2 está indicado en mayúsculas y cursiva, que impulsa la expresión del marcador de selección de neomicina indicado con minúsculas y cursiva, seguido por el terminador PmPGH 3'UTR subrayado en mayúsculas. El promotor PmUAPA2-1 (indicado con texto en negrita) impulsa la expresión del gen GmFATA (indicado con un texto en minúsculas y negrita) y se termina con la CvNR 3'UTR indicada con negrita, en minúsculas y subrayada. Los sitios de clonación de restricción y los fragmentos de ADN espaciadores están indicados con letra subrayada y en mayúsculas.**

ES 2 772 123 T3

ccctcaactgcgacgctgggaaccttctccgggcaggcgatgtgctggtgggtttgcctccttggcacg
gctctacaccgctcgagtagccatgagggcggatggctgtgtcggttgccacttcgtccagagacg
gcaagtgcgtccatcctctgctgtgtggcgcgacgctgcagcagtcctctgcagcagatgagcgtg
actttggccatttcacgcactcgagtgtacacaatccatttttcttaaagcaaagtactgctgattg
accagatactgtaacgctgatttcgctccagatcgcacagatagcgaccatgtttgctgctgtgaaa
atctggattccgaattcgacctggcgtccatccatgcaacagatggcgacacttggtacaattcc
tgtcacccatcggtatggagcaggtccacttagattcccgatcacccacgcacatctcgctaatagt
cattcgttcgtgtcttcgatcaatctcaagtgagtgtgcatggatcttggttgacgatgcggtatgg
gtttgcgcgctggctgcagggctctgcccaggcaagctaaccagctcctctccccgacaatactc
tcgcaggcaaagccggtcacttgccctccagattgccaataaactcaattatggcctctgtcatgcc
atccatgggtctgatgaatggtcacgctcgtgtcctgaccttccccagcctctggcgtccccctgcc
ccgcccaccagcccacgcccgcgcggcagtcgctgccaaggctgtctcggaGGTACCCTTTCTTGCGC
TATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGA
TGATGCTTCGACCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGC
GCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAAC
ACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCC
TCTTCTTCTTCGTTTCAGTCACAACCCGCAAACTCTAGAATATCAatgatcgagcaggacggcctcc
acgcccggctccccgcgcctgggtggagcgcctgttcggctacgactgggcccagcagaccatcgg
ctgctccgacgccgcctgttccgcctgtccgcccagggccgccccgtgctgttcgtgaagaccgac
ctgtccggcgccctgaacgagctgcaggacgaggccgcccgcctgtcctggctggccaccaccggcg

[illegible]

ccgccggtctgtcctgcgcttacgagctgaccaagcaccctgacgtccgggtacgcgagctgagatt
cgattagacataaattgaagattaaacccgtagaaaaatttgatggtcgcgaactgtgctcgattg
caagaaattgatcgctcctccactccgcaggtcgccatcatcgagcagggcggttgctcccgccggcg
cgctggctggggggacagctgttctcgccatgtgtgtacgtagaaggatgaatttcagctggttt
tcgttgacagctgtttgtgcatgattgtttcagactattgttgaaatgttttagatttcttagga
tgcatgatttgtctgcatgcgact

SEQ ID NO: 31 - secuencia de nucleótidos del gen mutante GmFATA S111A, V193A (D3998, pSZ5084). El promotor, 3'UTR, el marcador de selección y los brazos de direccionamiento son los mismos que los que se describen en SEQ ID NO: 30.

atggccaccgcatccactttctcggcgttcaatgcccgcgtgcggcgacctgcgtcgctcgggcggt
ccgggccccggcgcccagcgaggccccctccccgtgcgcggggcgcccacccccccgcatcatcg
ggtgtcctcctcctcctccaaggtgaacccccctgaagaccgaggccgtggtgtcctccggcctggcc
gaccgcctgcgcctgggtccttgaccgaggacggcctgtcctacaaggagaagttcatcgctgcgt
gctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcaggaggtggg
ctgcaaccacgcccagtcctgtgggtactccaccggcggttcgccaccacccccaccatgcgcaag
ctgcgcctgatctgggtgaccgcccgcgtgcacatcgagatctacaagtaccccgctggtccgacg
tggtggagatcgagtcctggggccaggcgagggaagatcggcaccccgccgactggatcctgcg
cgactacgccaccggccaggtgatcggcgcgccacctccaagtgggtgatgatgaaccaggacacc
cgccgcctgcagaaggtggacgcggacgtgcgcgacgagtagctggtgactgcccccgcgagctgc
gcctggccttccccgaggagaacaactcctccctgaagaagatctccaagctggaggacccccccca
gtactccaagctgggcctggtgccccgcgcgcgcgacctggacatgaaccagcacgtgaacaacgtg
acctacatcggttgggtgctggagtccatgccccaggagatcatcgacaccacgagctgcagacca
tcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactcctgacctccccgagcc
ctccgaggacgcccagggcgtgttcaaccacaacggcaccaacggctccgccaacgtgtccgccaac
gaccacggctgcccgaacttcttgacctgctgcgcctgtccggcaacggcctggagatcaaccgcg
gccgcaccgagtggcgcaagaagcccaccgcgtggactacaaggaccacgacggcgactacaagga
ccacgacatcgactacaaggacgacgacgacaagtga

SEQ ID NO: 32 - secuencia de nucleótidos del gen mutante GmFATA S111V, V193A (D3999, pSZ5085). El promotor, 3'UTR, el marcador de selección y los brazos de direccionamiento son los mismos que los que se describen en SEQ ID NO: 30.

atggccaccgcatccactttctcggcgttcaatgcccgcgtgcggcgacctgcgtcgctcgggcggt
ccgggccccggcgcccagcgaggccccctccccgtgcgcggggcgcccacccccccgcatcatcg
ggtgtcctcctcctcctccaaggtgaacccccctgaagaccgaggccgtggtgtcctccggcctggcc
gaccgcctgcgcctgggtccttgaccgaggacggcctgtcctacaaggagaagttcatcgctgcgt
gctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcaggaggtggg
ctgcaaccacgcccagtcctgtgggtactccaccggcggttcgtcaccacccccaccatgcgcaag
ctgcgcctgatctgggtgaccgcccgcgtgcacatcgagatctacaagtaccccgctggtccgacg
tggtggagatcgagtcctggggccaggcgagggaagatcggcaccccgccgactggatcctgcg
cgactacgccaccggccaggtgatcggcgcgccacctccaagtgggtgatgatgaaccaggacacc
cgccgcctgcagaaggtggacgcggacgtgcgcgacgagtagctggtgactgcccccgcgagctgc
gcctggccttccccgaggagaacaactcctccctgaagaagatctccaagctggaggacccccccca
gtactccaagctgggcctggtgccccgcgcgcgcgacctggacatgaaccagcacgtgaacaacgtg
acctacatcggttgggtgctggagtccatgccccaggagatcatcgacaccacgagctgcagacca
tcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactcctgacctccccgagcc
ctccgaggacgcccagggcgtgttcaaccacaacggcaccaacggctccgccaacgtgtccgccaac
gaccacggctgcccgaacttcttgacctgctgcgcctgtccggcaacggcctggagatcaaccgcg
gccgcaccgagtggcgcaagaagcccaccgcgtggactacaaggaccacgacggcgactacaagga
ccacgacatcgactacaaggacgacgacgacaagtga

SEQ ID NO: 33 - secuencia de nucleótidos del gen mutante GmFATA G96A (D4000, pSZ5086). El promotor, 3'UTR, el marcador de selección y los brazos de direccionamiento son los mismos que los que se describen en SEQ ID NO: 30.

atggccaccgcatccactttctcggcgttcaatgcccgcgtgcgggcgacctgcgtcgctcgggcgggct
ccggggccccggcgcccagcgaggccccctccccgtgcgcggggcgcgccatccccccccgcatcatcg
ggtgtcctcctcctcctccaaggtgaacccccctgaagaccgaggccgtggtgtcctccggcctggcc
gaccgcctgcgcctgggctccctgaccgaggacggcctgtcctacaaggagaagttcatcgtgcgt
gctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcaggaggtggc
gtgcaaccacgcccagtcctgtgggctactccaccggcggttctccaccacccccaccatgcgcaag
ctgcgcctgatctgggtgaccgcccgcgtgcacatcgagatctacaagtaccccgctggtccgacg
tggtggagatcgagtcctggggccaggcgagggaagatcggcacccgcgcgactggatcctgcg
cgactacgccaccggccagggtgatcgcccgccacctccaagtgggtgatgaaccaggacacc
cgccgcctgcagaaggtggacgtggacgtgcgcgacgagtagctggtgcactgccccgcgagctgc
gcctggccttccccgaggagaacaactcctccctgaagaagatctccaagtggaggacccccccca
gtactccaagctgggcctggtgccccgcgcgcccacctggacatgaaccagcacgtgaacaacgtg
acctacatcggttgggtgctggagtccatgccccaggagatcatcgacacccacgagctgcagacca
tcacctggactaccgcccgcgagtgccagcacgacgagctggtggactcctgacctccccgagcc
ctccgaggacgcccagggcgtgttcaaccacaacggcaccaacggctccgccaacgtgtccgccaac
gaccacggctgcccgaacttcttgacacgtgctgcgcctgtccggcaacggcctggagatcaaccgcg
gccgcaccgagtgggcgcaagaagcccaccgcgtggactacaaggaccacgacggcgactacaagga
ccacgacatcgactacaaggacgacgacgacaagtga

SEQ ID NO: 34 - secuencia de nucleótidos del gen mutante GmFATA G96T (D4001, pSZ5087). El promotor, 3'UTR, el marcador de selección y los brazos de direccionamiento son los mismos que los que se describen en SEQ ID NO: 30.

atggccaccgcatccactttctcggcgttcaatgcccgcgtgcgggcgacctgcgtcgctcgggcgggct
ccggggccccggcgcccagcgaggccccctccccgtgcgcggggcgcgccatccccccccgcatcatcg
ggtgtcctcctcctcctccaaggtgaacccccctgaagaccgaggccgtggtgtcctccggcctggcc
gaccgcctgcgcctgggctccctgaccgaggacggcctgtcctacaaggagaagttcatcgtgcgt
gctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcaggaggtgac
gtgcaaccacgcccagtcctgtgggctactccaccggcggttctccaccacccccaccatgcgcaag
ctgcgcctgatctgggtgaccgcccgcgtgcacatcgagatctacaagtaccccgctggtccgacg
tggtggagatcgagtcctggggccaggcgagggaagatcggcacccgcgcgactggatcctgcg
cgactacgccaccggccagggtgatcgcccgccacctccaagtgggtgatgaaccaggacacc
cgccgcctgcagaaggtggacgtggacgtgcgcgacgagtagctggtgcactgccccgcgagctgc
gcctggccttccccgaggagaacaactcctccctgaagaagatctccaagtggaggacccccccca
gtactccaagctgggcctggtgccccgcgcgcccacctggacatgaaccagcacgtgaacaacgtg
acctacatcggttgggtgctggagtccatgccccaggagatcatcgacacccacgagctgcagacca
tcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactcctgacctccccgagcc
ctccgaggacgcccagggcgtgttcaaccacaacggcaccaacggctccgccaacgtgtccgccaac
gaccacggctgcccgaacttcttgacacgtgctgcgcctgtccggcaacggcctggagatcaaccgcg
gccgcaccgagtgggcgcaagaagcccaccgcgtggactacaaggaccacgacggcgactacaagga
ccacgacatcgactacaaggacgacgacgacaagtga

SEQ ID NO: 35 - secuencia de nucleótidos del gen mutante GmFATA G96V (D4002, pSZ5088). El promotor, 3'UTR, el marcador de selección y los brazos de direccionamiento son los mismos que los que se describen en SEQ ID NO: 30.

atggccaccgcatccactttctcggcgttcaatgcccgcgtgcgggcgacctgcgtcgctcggcgggct
 ccgggccccggcgcccagcgaggccctccccgtgcgcggggcgcgccatccccccccgcgcatcatcgt
 ggtgtcctcctcctcctccaaggtgaacccctgaagaccgaggccgtggtgtcctccggcctggcc
 gaccgcctgcgcctgggtccttgaccgaggacggcctgtcctacaaggagaagttcatcgtgcgt
 gctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcaggaggtggg
 gtgcaaccacgcccagtcctgtgggtactccaccggcggttctccaccacccccaccatgcgcaag
 ctgcgcctgatctgggtgaccgcccgcgcatgcacatcgagatctacaagtaccccgctgggtccgacg
 tgggtggagatcgagtcctggggccaggcgagggaagatcggcacccgcgcgactggatcctgcg
 cgactacgccaccggccagggtgatcgccgcgcccacctccaagtgggtgatgatgaaccaggacacc
 cgccgcctgcagaaggtggacgtggacgtgcgcgacgagtagctgggtgactgccccgcgagctgc
 gcctggccttccccgaggagaacaactcctccctgaagaagatctccaagtggaggacccctccca
 gtactccaagctgggcctgggtgccccgcgcgcccacctggacatgaaccagcacgtgaacaacgtg
 acctacatcggttgggtgctggagtccatgccccaggagatcatcgacacccacgagctgcagacca
 tcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactcctgacctccccgagcc
 ctccgaggacgcccaggccgtgttcaaccacaacggcaccaacggctccgccaacgtgtccgccaac
 gaccacggctgcccgaacttcttgacacgtgctgcgcctgtccggcaacggcctggagatcaaccgcg
 gccgcaccgagtgccgcaagaagcccaccgcgcatggactacaaggaccacgacggcgactacaagga
 ccacgacatcgactacaaggacgacgacgacacaagtga

SEQ ID NO: 36 - secuencia de nucleótidos del gen mutante GmFATA G108A (D4003, pSZ5089). El promotor, 3'UTR, el marcador de selección y los brazos de direccionamiento son los mismos que los que se describen en SEQ ID NO: 30.

atggccaccgcatccactttctcggcgttcaatgcccgcgtgcgggcgacctgcgtcgctcggcgggct
 ccgggccccggcgcccagcgaggccctccccgtgcgcggggcgcgccatccccccccgcgcatcatcgt
 ggtgtcctcctcctcctccaaggtgaacccctgaagaccgaggccgtggtgtcctccggcctggcc
 gaccgcctgcgcctgggtccttgaccgaggacggcctgtcctacaaggagaagttcatcgtgcgt
 gctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcaggaggtggg
 ctgcaaccacgcccagtcctgtgggtactccaccggcggttctccaccacccccaccatgcgcaag
 ctgcgcctgatctgggtgaccgcccgcgcatgcacatcgagatctacaagtaccccgctgggtccgacg
 tgggtggagatcgagtcctggggccaggcgagggaagatcggcacccgcgcgactggatcctgcg
 cgactacgccaccggccagggtgatcgccgcgcccacctccaagtgggtgatgatgaaccaggacacc
 cgccgcctgcagaaggtggacgtggacgtgcgcgacgagtagctgggtgactgccccgcgagctgc
 gcctggccttccccgaggagaacaactcctccctgaagaagatctccaagtggaggacccctccca
 gtactccaagctgggcctgggtgccccgcgcgcccacctggacatgaaccagcacgtgaacaacgtg
 acctacatcggttgggtgctggagtccatgccccaggagatcatcgacacccacgagctgcagacca
 tcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactcctgacctccccgagcc
 ctccgaggacgcccaggccgtgttcaaccacaacggcaccaacggctccgccaacgtgtccgccaac
 gaccacggctgcccgaacttcttgacacgtgctgcgcctgtccggcaacggcctggagatcaaccgcg
 gccgcaccgagtgccgcaagaagcccaccgcgcatggactacaaggaccacgacggcgactacaagga
 ccacgacatcgactacaaggacgacgacgacacaagtga

SEQ ID NO: 37 - secuencia de nucleótidos del gen mutante GmFATA L91F (D4004, pSZ5090). El promotor, 3'UTR, el marcador de selección y los brazos de direccionamiento son los mismos que los que se describen en SEQ ID NO: 30.

atggccaccgcatccactttctcggcgttcaatgcccgcgtgcgggcgacctgcgtcgctcggcgggct
 ccgggccccggcgcccagcgaggccctccccgtgcgcggggcgcgccatccccccccgcgcatcatcgt
 ggtgtcctcctcctcctccaaggtgaacccctgaagaccgaggccgtggtgtcctccggcctggcc
 gaccgcctgcgcctgggtccttgaccgaggacggcctgtcctacaaggagaagttcatcgtgcgt
 gctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacttcttgacgaggtggg
 ctgcaaccacgcccagtcctgtgggtactccaccggcggttctccaccacccccaccatgcgcaag

ctgcgcctgatctgggtgaccgcccgcacatcgagatctacaagtaccccgctggtccgacg
tggtggagatcgagtcctggggccaggcgagggcaagatcggcaccgcccgcgactggatcctgcg
cgactacgccaccggccaggtgatcgccgcgccacctccaagtgggtgatgatgaaccaggacacc
cgccgctgcagaaggtggacgtggacgtgcgcgacgagtagctggtgactgccccgcgagctgc
gcctggccttccccgaggagaacaactcctccctgaagaagatctccaagtggaggacccctccca
gtactccaagtgggcctggtgccccgcgcgcccacctggacatgaaccagcacgtgaacaacgtg
acctacatcggttgggtgctggagtccatgccccaggagatcatcgacaccacgagctgcagacca
tcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactcctgacctccccgagcc
ctccgaggacgccgagggcgtgttcaaccacaacggcaccaacggctccgccaacgtgtccgccaac
gaccacggctgcccgaacttctctgcacctgctgcgcctgtccggcaacggcctggagatcaaccgcg
gccgcaccgagtgccgcaagaagcccaccgcacatggactacaaggaccacgacggcgactacaagga
ccacgacatcgactacaaggacgacgacgacaagtga

SEQ ID NO: 38 - secuencia de nucleótidos del gen mutante GmFATA L91K (D4005, pSZ5091). El promotor, 3'UTR, el marcador de selección y los brazos de direccionamiento son los mismos que los que se describen en SEQ ID NO: 30.

atggccaccgcatccactttctcggcgttcaatgcccgcctgcggcgacctgcgtcgctcgggcggt
ccgggccccggcgcccagcgaggccctccccgtgcgcggggcgccca tcccccccgcatcatcgt
ggtgtcctcctcctcctccaaggtgaacccctgaagaccgagggcgtggtgtcctccggcctggcc
gaccgctgcgcctgggtccttgaccgaggacggcctgtcctacaaggagaagttcatcgtgcgt
gctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacaagtgcaggaggtggg
ctgcaaccacgcccagtcctggtggtactccaccggcggttctccaccacccccaccatgcgcaag
ctgcgcctgatctgggtgaccgcccgcacatcgagatctacaagtaccccgctggtccgacg
tggtggagatcgagtcctggggccaggcgagggcaagatcggcaccgcccgcgactggatcctgcg
cgactacgccaccggccaggtgatcgccgcgccacctccaagtgggtgatgatgaaccaggacacc
cgccgctgcagaaggtggacgtggacgtgcgcgacgagtagctggtgactgccccgcgagctgc
gcctggccttccccgaggagaacaactcctccctgaagaagatctccaagtggaggacccctccca
gtactccaagtgggcctggtgccccgcgcgcccacctggacatgaaccagcacgtgaacaacgtg
acctacatcggttgggtgctggagtccatgccccaggagatcatcgacaccacgagctgcagacca
tcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactcctgacctccccgagcc
ctccgaggacgccgagggcgtgttcaaccacaacggcaccaacggctccgccaacgtgtccgccaac
gaccacggctgcccgaacttctctgcacctgctgcgcctgtccggcaacggcctggagatcaaccgcg
gccgcaccgagtgccgcaagaagcccaccgcacatggactacaaggaccacgacggcgactacaagga
ccacgacatcgactacaaggacgacgacgacaagtga

SEQ ID NO: 39 - secuencia de nucleótidos del gen mutante GmFATA L91S (D4006, pSZ5092). El promotor, 3'UTR, el marcador de selección y los brazos de direccionamiento son los mismos que los que se describen en SEQ ID NO: 30.

atggccaccgcatccactttctcggcgttcaatgcccgcctgcggcgacctgcgtcgctcgggcggt
ccgggccccggcgcccagcgaggccctccccgtgcgcggggcgccca tcccccccgcatcatcgt
ggtgtcctcctcctcctccaaggtgaacccctgaagaccgagggcgtggtgtcctccggcctggcc
gaccgctgcgcctgggtccttgaccgaggacggcctgtcctacaaggagaagttcatcgtgcgt
gctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaactcgctgcaggaggtggg
ctgcaaccacgcccagtcctggtggtactccaccggcggttctccaccacccccaccatgcgcaag
ctgcgcctgatctgggtgaccgcccgcacatcgagatctacaagtaccccgctggtccgacg
tggtggagatcgagtcctggggccaggcgagggcaagatcggcaccgcccgcgactggatcctgcg
cgactacgccaccggccaggtgatcgccgcgccacctccaagtgggtgatgatgaaccaggacacc
cgccgctgcagaaggtggacgtggacgtgcgcgacgagtagctggtgactgccccgcgagctgc
gcctggccttccccgaggagaacaactcctccctgaagaagatctccaagtggaggacccctccca
gtactccaagtgggcctggtgccccgcgcgcccacctggacatgaaccagcacgtgaacaacgtg
acctacatcggttgggtgctggagtccatgccccaggagatcatcgacaccacgagctgcagacca
tcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactcctgacctccccgagcc
ctccgaggacgccgagggcgtgttcaaccacaacggcaccaacggctccgccaacgtgtccgccaac
gaccacggctgcccgaacttctctgcacctgctgcgcctgtccggcaacggcctggagatcaaccgcg

gccgcaccgagtgggcgcaagaagcccacccgcatggactacaaggaccacgacggcgactacaagga
ccacgacatcgactacaaggacgacgacgacaagtga

SEQ ID NO: 40 - secuencia de nucleótidos del gen mutante GmFATA G108V (D4007, pSZ5093). El promotor, 3'UTR, el marcador de selección y los brazos de direccionamiento son los mismos que los que se describen en SEQ ID NO: 30.

atggccaccgcatccactttctcggcgttcaatgcccgcgtgcgggcgacctgcgtcgctcgggcgggct
ccgggccccggcgccccagcgaggccctccccgtgcgcggggcgcgccatccccccccgcatcatcgt
ggtgtcctcctcctcctccaaggtgaacccctgaagaccgaggccgtggtgtcctccggcctggcc
gaccgcctgcgccctgggctccctgaccgaggacggcctgtcctacaaggagaagttcatcgtgcgt
gctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcaggaggtggg
ctgcaaccacgcccagtcctgtgggctactccaccgtcggttctccaccacccccaccatgcgcaag
ctgcgcctgatctgggtgaccgcccgcgtgcacatcgagatctacaagtaccccgctggtccgacg
tggtggagatcgagtcctggggccaggcgagggaagatcggcaccccgcgcgactggatcctgcg
cgactacgccaccggccaggtgatcgggcgcgccacctccaagtgggtgatgatgaaccaggacacc
cgccgcctgcagaaggtggacgtggacgtgcgcgacgagtagctggtgactgcccccgcgagctgc
gcttgcccttccccgaggagaacaactcctccctgaagaagatctccaagtggaggacccctccca
gtactccaagctgggcctggtgccccgcccgcgcgacctggacatgaaccagcacgtgaacaacgtg
acctacatcggtgggtgctggagtccatgccccaggagatcatcgacaccacgagctgcagacca
tcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactcctgacctccccgagcc
ctccgaggacgcccagggccgtgttcaaccacaacggcaccaacggctccgccaacgtgtccgccaac
gaccacggctgcccgaacttcttgacctgctgcgcctgtccggcaacggcctggagatcaaccgcg
gccgcaccgagtgggcgcaagaagcccacccgcatggactacaaggaccacgacggcgactacaagga
ccacgacatcgactacaaggacgacgacgacaagtga

SEQ ID NO: 41 - secuencia de nucleótidos del gen mutante GmFATA T156F (D4008, pSZ5094). El promotor, 3'UTR, el marcador de selección y los brazos de direccionamiento son los mismos que los que se describen en SEQ ID NO: 30.

atggccaccgcatccactttctcggcgttcaatgcccgcgtgcgggcgacctgcgtcgctcgggcgggct
ccgggccccggcgccccagcgaggccctccccgtgcgcggggcgcgccatccccccccgcatcatcgt
ggtgtcctcctcctcctccaaggtgaacccctgaagaccgaggccgtggtgtcctccggcctggcc
gaccgcctgcgccctgggctccctgaccgaggacggcctgtcctacaaggagaagttcatcgtgcgt
gctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcaggaggtggg
ctgcaaccacgcccagtcctgtgggctactccaccggcggttctccaccacccccaccatgcgcaag
ctgcgcctgatctgggtgaccgcccgcgtgcacatcgagatctacaagtaccccgctggtccgacg
tggtggagatcgagtcctggggccaggcgagggaagatcggttccgcccgcgactggatcctgcg
cgactacgccaccggccaggtgatcgggcgcgccacctccaagtgggtgatgatgaaccaggacacc
cgccgcctgcagaaggtggacgtggacgtgcgcgacgagtagctggtgactgcccccgcgagctgc
gcttgcccttccccgaggagaacaactcctccctgaagaagatctccaagtggaggacccctccca
gtactccaagctgggcctggtgccccgcccgcgcgacctggacatgaaccagcacgtgaacaacgtg
acctacatcggtgggtgctggagtccatgccccaggagatcatcgacaccacgagctgcagacca
tcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactcctgacctccccgagcc
ctccgaggacgcccagggccgtgttcaaccacaacggcaccaacggctccgccaacgtgtccgccaac
gaccacggctgcccgaacttcttgacctgctgcgcctgtccggcaacggcctggagatcaaccgcg
gccgcaccgagtgggcgcaagaagcccacccgcatggactacaaggaccacgacggcgactacaagga
ccacgacatcgactacaaggacgacgacgacaagtga

SEQ ID NO: 42 - secuencia de nucleótidos del gen mutante GmFATA T156A (D4009, pSZ5095). El promotor, 3'UTR, el marcador de selección y los brazos de direccionamiento son los mismos que los que se describen en SEQ ID NO: 30.

atggccaccgcatccactttctcggcggttcaatgcccgcgtgcgggcgacctgcgtcgctcggcgggct
ccgggccccggcgcccagcgaggccctccccgtgcgcggggcgcccatcccccccgcatcatcg
ggtgtcctcctcctcctccaaggtgaacccctgaagaccgaggccgtggtgtcctccggcctggcc
gaccgcctgcgcctgggctccctgaccgaggacggcctgtcctacaaggagaagttcatcgtgcgt
gctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcaggaggtggg
ctgcaaccacgcccagtcctgtgggctactccaccggcggttctccaccacccccaccatgcgcaag
ctgcgcctgatctgggtgaccgcccgcgtgcacatcgagatctacaagtaccccgctggtccgacg
tggtggagatcgagtcctggggccaggcgagggaagatcggcgcgcgcgcgactggatcctgcg
cgactacgccaccggccagggtgatcgccgcgcgcacctccaagtgggtgatgatgaaccaggacacc
cgccgcctgcagaaggtggacgtggacgtgcgcgacgagtagctggtgcactgccccgcgagctgc
gcctggccttccccgaggagaacaactcctccctgaagaagatctccaagtggaggacccctccca
gtactccaagctgggcctggtgccccgcgcgcgcgacctggacatgaaccagcacgtgaacaacgtg
acctacatcggttgggtgctggagtccatgccccaggagatcatcgacaccacgagctgcagacca
tcacctggactaccgcccgcgagtgccagcacgacgagctggtggactcctgacctccccgagcc
ctccgaggacgcccagggcgtgttcaaccacaacggcaccaacggctccgccaacgtgtccgccaac
gaccacggctgcccgaacttctctgcacctgctgcgcctgtccggcaacggcctggagatcaaccgcg
gccgcaccgagtgccgcaagaagcccaccgcgtggactacaaggaccacgacggcgactacaagga
ccacgacatcgactacaaggacgacgacgacaagtga

SEQ ID NO: 43 - secuencia de nucleótidos del gen mutante GmFATA T156K (D4010, pSZ5096). El promotor, 3'UTR, el marcador de selección y los brazos de direccionamiento son los mismos que los que se describen en SEQ ID NO: 30.

atggccaccgcatccactttctcggcggttcaatgcccgcgtgcgggcgacctgcgtcgctcggcgggct
ccgggccccggcgcccagcgaggccctccccgtgcgcggggcgcccatcccccccgcatcatcg
ggtgtcctcctcctcctccaaggtgaacccctgaagaccgaggccgtggtgtcctccggcctggcc
gaccgcctgcgcctgggctccctgaccgaggacggcctgtcctacaaggagaagttcatcgtgcgt
gctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcaggaggtggg
ctgcaaccacgcccagtcctgtgggctactccaccggcggttctccaccacccccaccatgcgcaag
ctgcgcctgatctgggtgaccgcccgcgtgcacatcgagatctacaagtaccccgctggtccgacg
tggtggagatcgagtcctggggccaggcgagggaagatcggaagcgccgcgactggatcctgcg
cgactacgccaccggccagggtgatcgccgcgcgcacctccaagtgggtgatgatgaaccaggacacc
cgccgcctgcagaaggtggacgtggacgtgcgcgacgagtagctggtgcactgccccgcgagctgc
gcctggccttccccgaggagaacaactcctccctgaagaagatctccaagtggaggacccctccca
gtactccaagctgggcctggtgccccgcgcgcgcgacctggacatgaaccagcacgtgaacaacgtg
acctacatcggttgggtgctggagtccatgccccaggagatcatcgacaccacgagctgcagacca
tcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactcctgacctccccgagcc
ctccgaggacgcccagggcgtgttcaaccacaacggcaccaacggctccgccaacgtgtccgccaac
gaccacggctgcccgaacttctctgcacctgctgcgcctgtccggcaacggcctggagatcaaccgcg
gccgcaccgagtgccgcaagaagcccaccgcgtggactacaaggaccacgacggcgactacaagga
ccacgacatcgactacaaggacgacgacgacaagtga

SEQ ID NO: 44 - secuencia de nucleótidos del gen mutante GmFATA T156V (D4011, pSZ5097). El promotor, 3'UTR, el marcador de selección y los brazos de direccionamiento son los mismos que los que se describen en SEQ ID NO: 30.

atggccaccgcatccactttctcggcggttcaatgcccgcgtgcggcgacctgcgtcgctcggcgggct
ccggggccccggcgcccagcgaggccccctccccgtgcgcggggcgccatccccccccgcattatcg
ggtgtcctcctcctcctccaaggtgaacccccgaagaccgagggcggtggtgtcctcggcctggcc
gaccgcctgcgcctgggctccctgaccgaggacggcctgtcctacaaggagaagttcatcgtgcgt
gctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgcaacctgctgcaggaggtggg
ctgcaaccacgcccagtcctgtgggctactccaccggggcttctccaccacccccaccatgcgaag
ctgcgcctgatctgggtgaccgcccgcattgcacatcgagatctacaagtaccccgctggtccgacg
tggtggagatcgagtcctggggccaggggcgagggcaagatcggcgtgcgcgcgactggatcctgcg
cgactacgccaccggccagggtgatcgcccgccaccctccaagtgggtgatgaaccaggacacc
cgccgcctgcagaaggtggacgtggacgtgcgcgacgagtagctggtgcactgccccgcgagctgc
gcctggccttccccgaggagaacaactcctccctgaagaagatctccaagtggaggacccccccca
gtactccaagctgggctggtgccccgcgcgcgacctggacatgaaccagcacgtgaacaacgtg
acctacatcggtggtgctggagtccatgccccaggagatcatcgacaccacgagctgcagacca
tcacctggactaccgcccgcgagtgccagcacgacgagctggtggactcctgacctccccgagcc
ctccgaggacgcccagggcgtgttcaaccacaacggcaccaacggctccgccaacgtgtccgccaac
gaccacggctgcccgaacttctctgcacctgctgcgcctgtccggcaacggcctggagatcaaccgcg
gccgcaccgagtgccgcaagaagcccaccgcgcatggactacaaggaccacgacggcgactacaagga
ccacgacatcgactacaaggacgacgacgacaagtga

SEQ ID NO: 45

Secuencia de aminoácidos del gen parental de tipo salvaje Gm FATA; el péptido señal está eliminado

IPRIIVVSSSSSKVNPLKTEAVVSSGLADRLRLGSLTEDGLSYKEKFIVRCYEVGINKTATVETIA
NLLQEVGCNHAQSVGYSTGGFSTTPTMRKLRLIWVTARMHIEIYKYPAWSVVEIESWGQEGKIGT
RRDWILRDYATGQVIGRATSKWMMNQDTRRLQKVDVDVRDEYLVHCPRELRLAFPEENNSSLKKIS
KLEDPSQYSKLGIVPRRADLDMNQHVNNVTYIGWVLESMPQEIIDTHELQTITLDYRRECQHDDVDV
SLTSPSEDAEAVFNHNGTNGSANVSANDHGCNRFLLHLLRLSGNGLEINRGRTWRKKPTR

SEQ ID NO: 46

Secuencia de nucleótidos del ADN transformante contenido en el plásmido pSZ5990 transformado en S5780

gctcttcccaactcagataataccaatacccctccttctcctcctcatccattcagtacccccccc
ttctcttcccaaagcagcaagcgctggcttacagaagaacaatcggttccgcaaagtcgccgag
cactgcccgcgaggcgccgcccagcagcccgcttgccacacaggcaacgaatacattcaataggg
ggcctcgcaaatggaaggagcggtaaagggtagaggagcactgcgcacaaggggctgtgcaggag
tgactgactgggcgggcagacggcgcccgggcgaggcaagcaggggaagattgaagcggcaggg
aggaggatgctgattgaggggggcacgcagtcctctcttgaccgggataaggaagcaaatattcg
gccggttggtgtgtgtgtgcacgttttcttctcagagtcgtgggtgtgcttcaggaggagatat
aagcagcaggatcgaaatcccgcgaccagcggttcccatccagccaaccacctgtcggtacccttt
cttgcgctatgacacttccagcaaaaggtaggcggggctgcgagacggcttcccggcgtgcatgca
acaccgatgatgcttcgacccccgaagctccttcggggctgcatggcgctccgatgccgctccag
ggcgagcgctgtttaaatagccaggcccccgattgcaaagacattatagcgagctaccaagccata
ttcaaacacctagatcactaccacttctacacaggccactcgagcttgtgatcgcaactccgctaagg
ggcgccctcttctccttctcgtttcagtcacaacccgcaaacctctagatatcaATGctgctgcaggcc
ttcctgttctgctggccggcttcgccgccaagatcagcgccctccatgacgaacgagacgtccgacc
ccccctggtgacttcacccccacaagggctggatgaacgacccccaacggcctgtggtacgacga
gaaggacgccaagtggcacctgtacttccagtacaaccogaacgacaccgtctgggggacgcccttg

ttctggggccacgccacgtccgacgacctgaccaactgggaggaccagcccatcgccatcgccccga
 agcgcaacgactccggcgcttctccggctccatgggtgggactacaacaacacotccggcttctt
 caacgacacccatcgacccgcgcagcgctgcgtggccatctggacctacaacaccccgagtcgag
 gagcagtacatctcctacagcctggacggcggtacaccttcaccgagtagcagaagaaccccggtgc
 tggccgccaactccacccagttccgcgacccgaaggtcttcttggtacgagccctcccagaagtggat
 catgaccgcggccaagtcccaggactacaagatcgagatctactcctccgacgacctgaagtccgtg
 aagctggagtcgcggttcgccaacgagggcttccctcggtaccagtagcagtgccccggcctgatcg
 aggtccccaccgagcaggaccccagcaagtcctactgggtgatgttcatctccatcaacccccggcgc
 cccggccggcggtccttcaaccagtacttcgtcggcagcttcaacggcaccacttcgaggccttc
 gacaaccagtcccgctgggtggacttcggcaaggactactacgcccctgcagaccttcttcaacaccg
 acccgacctacgggagcgccctgggcatcgctgggcctccaactgggagtactccgccttcgtgcc
 caccaacccctggcgctcctccatgtccctcgtgcgcaagttctccctcaacaccgagtagcaggcc
 aaccgggagacgggagctgatcaacctgaaggccgagccgatcctgaacatcagcaacgcgggccccct
 ggagccggttcgccaccaacaccacgttgacgaaggccaacagctacaacgtcgacctgtccaacag
 caccggcacccctggagttcgagctgggtgacgccgtcaacaccacccagacgatctccaagtccgtg
 ttcgcggacctctccctctggttcaagggcctggaggaccccgaggagtacctccgcatgggcttcg
 aggtgtccgcgtcctccttcttctggtgacccgcgggaacagcaaggtgaagttcgtgaaggagaacc
 ctacttcaccaaccgcatgagcgtgaacaaccagcccttcaagagcgagaacgacctgtcctactac
 aaggtgtacggcttgctggaccagaacatcctggagctgtacttcaacgacggcgacgtcgtgtcca
 ccaacacctacttcatgaccacccgggaacgccttgggtccgtgaacatgacgacgggggtggacaa
 cctgttctacatcgacaagtccaggtgcgcgaggtcaagTGAcaattgacgccccgcgcggcgacc
tgacctgttctctcgaggcgccctgttctgccttgcgaacaagccctggagcatgcgtgcatgat
cgtctctggcgccccgcgcgggttctgcctcgcggggcgccgcggccgcggggcgcattgaa
attgttgcaaaccacacctgacagattgagggccaggcaggaaggcgttgagatggaggtacagga
gtcaagtaactgaaagttttatgataactaacaacaaagggctggttctggccagcgaatgacaag
aacaagattccacatttccgtgtagaggttgcctcgaatgtgagcggggcgggccgcggacccgac
aaaacccttacgacgtggtgaagaaaaacgtggcgggcactgtccctgtagcctgaagaccagcagga
gacgatcggaagcatcacagcacaggatcccgctctcgaacagagcgcgagagggaacgctgaagg
tctcgccctctgtcgacctcagcgcgccatacaccacaataaccacctgacgaatgcgcttggttct
tcgtccattagcgaagcgtccggttcacacacgtgccacgttggcgaggtggcaggtgacaatgatc
ggtggagctgatggtcgaaacgttcacagcctaggtacgcgcgtcagcctacacgtcttctccgata
cctttccctcattgcattttatgccagactgggtccagcctgggtgggtgctcccgctcgattgct
cgtgtcggaggcggggacccccgctctctctatttatcactgcctctccccgaccaaccctgacga
ctgtaaccctgccagaaacaattcagcctcatcaaaccgagttgtgcacaaggcgactaatttttt
agtcgggaaacaaccgccttcagaagcatccggacgggggtagcgaggctgtgtcgagcgccgtgg
ggatctggccggtgaggtgcccgaaatccgtgtacagctcagcggtgggatcatcgacccccggga
tcacgacccccgtggggccggggccccggaccctataactaaaagccgacgccagtgcaaaaccacaa
acatttactccttaatcctccctcctccttcatacacaccacaaagtaataactcacccatATGgc
 catcgccgcgcgcgcgtgatcgtgcccttgggcctgctgttcttcatctccggcctgggtggtgaac
 ctgatccaggccctgtgcttcgtgctgatccgccccctgtccaagaacacctaccgcaagatcaacc
 gcgtggtggccgagctgctgtggctggagctgatctggctggtggactggtgggccccgctgaagat
 caaggtgttcatggaccccgagtccttcaacctgatgggcaaggagcacgccctgggtggtggccaac
 caccgctccgacatcgactggctggtgggtggctgctggcccagcgctccggctgcctgggctccg
 ccctggccgtgatgaagaagtccccaagttcctgcccgatcggtggtccatgtggttctccga
 gtacctgttccctggagcgctcctgggccaaggacgagaacacctgaaggccggcctgcagcgccctg
 aaggacttccccgcgcccttctggctggccttcttcgtggagggcaccgccttcacccaggccaagt
 tcctggccgcccaggagtagcgcgcctcccagggcctgcccatcccccgcaacgtgctgatccccg
 caccaagggcttcgtgtccgcctgtcccacatgcgctccttcgtgcccgcctctacgacatgacc

gtggccatcccccaagtccctccccctccccaccatgctgcgccctgttcaagggccagccctccgtgg
 tgacgtgcacatcaagcgctgcctgatgaaggagctgcccagaccgacgaggccgtggccagtg
 gtgcaaggacatgttcgtggagaaggacaagctgctggacaagcacatcgccgaggacaccttctcc
 gaccagcccatgcaggacctgggcccggcccatcaagtccttgcctggtgggtggcctcctgggctgcc
 tgatggcctacggcgccctgaagttcctgcagtgtcctccctgctgtcctcctggaagggcatcgc
 cttcttccctgggtgggcctggccatcgtgaccatcctgatgcacatcctgatcctgttctcccagtec
 gagcgctccacccccgccaaggtggccccggcaagcccaagaacgacggcgagacctccgaggccc
 gccgcgacaagcagcagTGA**atgcat**atgtggagatgtaggggtggctgactcgttggaggtgggtgt
 ttttttttatcgagtgcgcggcgccgcaaacgggtcccttttttatcgaggtgttcccaacgccgcac
 cgccctcttaaaacaacccccaccaccacttgtcgaccttctcgtttgttatccgccacggcgcccc
 ggagggggcgtcgtctggccgcgcgggcagctgtatcgccgcgctcgtccaatggtgtgtaatcttg
 gaaagataataatcgatggatgaggaggagagcgctgggagatcagagcaaggaatatagttggca
 cgaagcagcagcgtactaagctgtagcgtgttaagaaagaaaaactcgtgttaggctgtattaatc
 aaggagcgtatcaataattaccgacctatacctttatctccaacccaatcgccgg**cttaaggatcta**
agtaagattcgaagcgctcgacctgcccgaaggactgcagcccatgtcgtagtgaccgccaatgt
aagtgggctggcggttccctgtacgtgagtcacagtcactgcacgcgcaccacctctcgaccggca
ggaccaggcatcgcgagatacagcgcgagccagacacggagtgccgagctatgcgcacgctccaact
agatatcatgtggatgatgagcatgaattcctggctcgggcctcgtgctggcactccctcccatgcc
gacaaccttctgctgtcaccacgaccacgatgcaacgcgacacgacctggggactgatcggtt
actgcacctgcatgcaattgtcacaagcgcatactccaatcgatatccgtttgatttctgtgaaaac
tcgctcgaccgcccgcgtcccgaggcagcgatgacgtgtgcgtgacctgggtgttctcgtcgaaagg
ccagcaaccccaaatcgaggcgatccggagattgggatctgatccgagcctggaccagatcccca
cgatgcggcacgggaactgcacgactcggcgcggaaccagcttctgtaaatgccagattggtgtc
cgataccttgatttgccatcagcgaaacaagacttcagcagcgagcgtatttgccggggcggtgctacc
agggttgcatacattgcccatttctgtctggaccgctttaccggcgagaggggtgagttgatgggt
tggcaggcatcgaaacgcgcgtgcatggtgtgtgtgtctgttttcggctgcacaatttcaatagtcg
gatgggcgacggtagaattgggtgttgcgctcgctgcatgcctcgccccgtcgggtgtcatgaccg
ggactggaatccccctcgcgacctcctgctaacgctcccgactctcccgcccgcgcgaggaatag
actctagttcaaccaatcgaca**actagt**ATGgccaccgcatccactttctcggcgttcaatgccgcg
tgccggcgacctgcgtcgtcggcgggctccgggccccggcgcccagcgaggccctccccgtgcgcg
ggcgcgccatccccccccgcatcatcgtggtgtcctcctcctccaaggtgaacccccctgaagac
 cgaggccgtggtgtcctccggcctggccgaccgctgcgcctgggtccttgaccgaggacggcctg
 tcctacaaggagaagttcatcgtgcgtgctacgaggtgggcatcaacaagaccgccaccgtggaga
 ccatcgccaacctgctgcaggaggtggcgtgcaaccacgcccagtcctggtgggctactccaccgcccg
 cttctccaccacccccaccatgcgcaagctgcgcctgatctgggtgaccgcccgcacatgcacatcgag
 atctacaagtaccccgctggtccgacgtggtggagatcgagtccctggggccaggggcgagggaaga
 tcggcaccgcgcgcgactggatcctgcgcgactacgccaccggccaggtgatcgccgcgcgcacctc
 caagtgggtgatgatgaaccaggacacccgcgcctgcagaaggtggacgtggacgtgcgcgacgag
 tacctggtgactgccccgcgagctgcgcctggccttccccgaggagaacaactcctcctgaaga
 agatctccaagctggaggaccctcccagtaactccaagctgggctggtgccccgcgcgcgcgacct
 ggacatgaaccagcacgtgaacaacgtgacctacatcggtggtgctggagtccatgcccaggag
 atcatcgacacccacgagctgcagaccatcacctggactaccgcccgcgagtgccagcacgacgag
 tgggtggactccctgacctcccccgagccctccgaggacgcccaggccgtgttcaaccacaacggcac
 caacgggtccgccaacgtgtccgccaacgaccacggctgccgcaacttctgcacctgctgcgcctg
 tccggcaacggcctggagatcaaccgcggccgcacccaggtggcgcaagaagcccaccgcg**atggact**
acaaggaccacgacggcgactacaaggaccacgacatcgactacaaggacgacgacgaagTGAat
cgatggagcgacgagtgctgcgtgcggggctggcgaggagtgaggacgccctcctcgtcctctctgttc
 tgaacggaacaatcgccacccccgcgctacgcgccacgcacatcgagcaacgaagaaacccccgatg

ataggttgcggtggctgccgggatatagatccggccgcacatcaaagggccctccgccagagaaga
agctcctttcccagcagactccttctgctgccaaaacattctctgtccacagcaacaccaaaggat
gaacagatcaacttgcttccgcgtagcttcctcggctagcgtgcttgcaacaggtccctgcacta
ttatcttctgctttcctctgaattatgcggcaggcgagcgctcgtcttgccgagcgctccttcgcg
ccgccctcgctgatcgagtgtacagtcaatgaatggtgagctcttgttttccagaaggagttgctcc
ttgagcctttcattctcagcctcgataacctccaaagccgctctaattgtggagggggttcgaa**ccg**
aatgctgctgaacgggaaggaggaggagaaagagtgagcaggaggaggttcagaaatgagaaatga
gaggtgaaggaacgcataccctatgcccttgcaatggacagtgttctggccaccgccaccaagactt
cgtgtcctctgatacatcatgcgattgattacgttgaatgcgacggccgggtcagccccggacctccac
gcaccggtgctcctccaggaagatgcgcttgtcctccgccatcttgcaagggtcaagctgctccaa
aactcttggggcggttccggacggacggctaccgcggtgcgccctgaccgccactgttcggaagc
agcgcgctgcatggggcagcgccgctgcggtgcgccacggaccgcatgatccaccggaaaagcgca
cgcgctggagcgcgcagaggaccacagagaagcggaagagacgccagtactggcaagcaggctgggtc
ggtgccatggcgcgctactaccctcgctatgactcgggtcctcggccggctggcggtgctgacaatt
cgtttagtgagcagcgactccattcagctaccagtcgaactcagtgggcacagtgactcccgctctt
c

SEQ ID NO: 47 - secuencia de nucleótidos del péptido de tránsito CpSAD fusionado con el gen GarmFATA1 (G108A) y la etiqueta 3X FLAG en pSZ5936 y pSZ6018

actagtATGgccaccgcatccactttctcggcggttcaatgcccgctgcggcgacctgcgtcgctcgg
cgggctccgggccccggcgccccagcgaggccccctccccgtgcgc**gggcgcgcc**atccccccccgc
catcgtggtgtcctcctcctcctccaaggtgaacccccctgaagaccgagggcgtggtgtcctccggc
ctggccgaccgctgcgcttgggtccttgaccgaggacggcctgtcctacaaggagaagttcatcg
tgcgctgctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcagga
ggtgggctgcaaccacgcccagtcctggtggtactccaccgcccgttctccaccacccccaccatg
cgcaagctgcgctgatctgggtgaccgcccgcgcatgcacatcgagatctacaagtaaccccgctggt
ccgacgtggtggagatcgagtcctggggccaggggcgagggcaagatcggcacccgcccgcgactggat
cctgcgcgactacgccaccggccaggtgatecgccgcgcgccacctccaagtgggtgatgatgaaccag
gacacccgcccgcctgcagaaggtggacgtggacgtgcgcgacgagtaacctggtgactgccccgcg
agctgcgcctggccttccccgaggagaacaactcctcctgaagaagatctccaagctggaggaccc
ctcccagtaactccaagctgggctggtgccccgcgcgcgcgacctggacatgaaccagcacgtgaac
aacgtgacctacatcggtgggtgctggagtccatgccccaggagatcatcgacacccacgagctgc
agaccatcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactcctgacctcccc
cgagccctccgaggacgcccaggccgtgttcaaccacaacggcaccaacggctccgccaacgtgtcc
gccaacgaccaggtgcccgaacttctgcacctgctgcgcctgtccggcaacggcctggagatca
accgcgccgcaccgagtgggcgcaagaagcccaccgc**atggactacaaggaccacgacggcgacta**
caaggaccacgacatcgactacaaggacgacgacgacaagTGA**atcgat**

SEQ ID NO:48 - secuencia de nucleótidos del péptido de tránsito CpSAD fusionado con el gen GarmFATA1 (G96A, S111A) y la etiqueta 3X FLAG en pSZ5991 y pSZ6026

actagtATGgccaccgcatccactttctcggcggttcaatgcccgctgcggcgacctgcgtcgctcgg
cgggctccgggccccggcgccccagcgaggccccctccccgtgcgc**gggcgcgcc**atccccccccgc
catcgtggtgtcctcctcctcctccaaggtgaacccccctgaagaccgagggcgtggtgtcctccggc
ctggccgaccgctgcgcttgggtccttgaccgaggacggcctgtcctacaaggagaagttcatcg
tgcgctgctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcagga
ggtggcgtgcaaccacgcccagtcctggtggtactccaccggcggttctgccaccacccccaccatg
cgcaagctgcgctgatctgggtgaccgcccgcgcatgcacatcgagatctacaagtaaccccgctggt
ccgacgtggtggagatcgagtcctggggccaggggcgagggcaagatcggcacccgcccgcgactggat
cctgcgcgactacgccaccggccaggtgatecgccgcgcgccacctccaagtgggtgatgatgaaccag
gacacccgcccgcctgcagaaggtggacgtggacgtgcgcgacgagtaacctggtgactgccccgcg

agctgcgcttggccttccccgaggagaacaactcctccctgaagaagatctccaagctggaggaccc
ctcccagtaactccaagctgggctggtgccccgcgcgcgcgacctggacatgaaccagcacgtgaac
aacgtgacctacatcggtgggtgctggagtccatgccccaggagatcatcgacacccacgagctgc
agaccatcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactccctgacctccc
cgagccctccgaggacgcgcgaggccgtgttcaaccacaacggccaccaacggctccgccaacgtgtcc
gccaaacgaccacggctgcccgaacttctgcaactgctgcgctgtccggcaacggcctggagatca
accgcgccgcgacccgagtgggcgcaagaagccccaccgcgatggactacaaggaccacgacggcgacta
caaggaccacgacatcgactacaaggacgacgacgacaagTGAatcgat

SEQ ID NO: 49 - secuencia de nucleótidos del péptido de tránsito CpSAD fusionado con el gen GarmFATA1 (G96A, V193A) y la etiqueta 3X FLAG en pSZ5986 y pSZ6023

actagtATGgccaccgcatccactttctcggcgttcaatgcccgcgtgcggcgacctgcgctcgctcgg
cggtccgggccccgcgcgcgcagcagggccctccccgtgcgcgggcgcgccatcccccccgcat
catcggtggtgtctctcctcctcctccaaggtgaacccccctgaagaccgagggcgtggtgtctccggc
ctggccgaccgctgcgctgggtcctgaccgaggacggcctgtcctacaaggagaagtcatcg
tgcgctgctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcagga
ggtggcggtgcaaccacgcccagtcctggtgggtactccaccggcggttctccaccacccccaccatg
cgcaagctgcgctgatctgggtgaccgcccgcgcatgcacatcgagatctacaagtaccccgctggt
ccgacgtggtggagatcgagtcctggggccagggcgagggcaagatcggcacccgcgcgactggat
cctgcgcgactacgccaccggccaggtgatcggcgcgcgcacccctccaagtgggtgatgatgaaccag
gacaccgcgcgcctgcagaaggtggacgcggacgtgcgcgacgagtagctggtgactgccccgcg
agctgcgcttggccttccccgaggagaacaactcctccctgaagaagatctccaagctggaggaccc
ctcccagtaactccaagctgggctggtgccccgcgcgcgcgacctggacatgaaccagcacgtgaac
aacgtgacctacatcggtgggtgctggagtccatgccccaggagatcatcgacacccacgagctgc
agaccatcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactccctgacctccc
cgagccctccgaggacgcgcgaggccgtgttcaaccacaacggccaccaacggctccgccaacgtgtcc
gccaaacgaccacggctgcccgaacttctgcaactgctgcgctgtccggcaacggcctggagatca
accgcgccgcgacccgagtgggcgcaagaagccccaccgcgatggactacaaggaccacgacggcgacta
caaggaccacgacatcgactacaaggacgacgacgacaagTGAatcgat

SEQ ID NO: 50 - secuencia de nucleótidos del péptido de tránsito CpSAD fusionado con el gen GarmFATA1 (G108A, S111A) y la etiqueta 3X FLAG en pSZ5982 y pSZ6019

actagtATGgccaccgcatccactttctcggcgttcaatgcccgcgtgcggcgacctgcgctcgctcgg
cggtccgggccccgcgcgcgcagcagggccctccccgtgcgcgggcgcgccatcccccccgcat
catcggtggtgtctctcctcctcctccaaggtgaacccccctgaagaccgagggcgtggtgtctccggc
ctggccgaccgctgcgctgggtcctgaccgaggacggcctgtcctacaaggagaagtcatcg
tgcgctgctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcagga
ggtggcggtgcaaccacgcccagtcctggtgggtactccaccggcggttctcgccaccacccccaccatg
cgcaagctgcgctgatctgggtgaccgcccgcgcatgcacatcgagatctacaagtaccccgctggt
ccgacgtggtggagatcgagtcctggggccagggcgagggcaagatcggcacccgcgcgactggat
cctgcgcgactacgccaccggccaggtgatcggcgcgcgcacccctccaagtgggtgatgatgaaccag
gacaccgcgcgcctgcagaaggtggacgtggacgtgcgcgacgagtagctggtgactgccccgcg
agctgcgcttggccttccccgaggagaacaactcctccctgaagaagatctccaagctggaggaccc
ctcccagtaactccaagctgggctggtgccccgcgcgcgcgacctggacatgaaccagcacgtgaac
aacgtgacctacatcggtgggtgctggagtccatgccccaggagatcatcgacacccacgagctgc
agaccatcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactccctgacctccc
cgagccctccgaggacgcgcgaggccgtgttcaaccacaacggccaccaacggctccgccaacgtgtcc
gccaaacgaccacggctgcccgaacttctgcaactgctgcgctgtccggcaacggcctggagatca
accgcgccgcgacccgagtgggcgcaagaagccccaccgcgatggactacaaggaccacgacggcgacta
caaggaccacgacatcgactacaaggacgacgacgacaagTGAatcgat

SEQ ID NO: 51 - secuencia de nucleótidos del péptido de tránsito CpSAD fusionado con el gen GarmFATA1 (G108A, V193A) y la etiqueta 3X FLAG en pSZ5983 y pSZ6020

actagtATGgccaccgcatccactttctcggcggttcaatgcccgctgcggcgacctgcgtcgctcggcgggctccgggccccggcgcccagcgaggccccctccccgtgcgcg**gggcgcgcc**atcccccccgcatcatcgtggtgtcctcctcctcctccaaggtgaacccccctgaagaccgagggcgtggtgtcctccggcctggccgaccgcctgcgcctgggctccctgaccgaggacggcctgtcctacaaggagaagtcatcgtgcgtgctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcaggaagtgggtgcaaccacgcccagtcctggtgggtactccaccgcccgttctccaccacccccaccatgcgaagctgcgcctgatctgggtgaccgcccgcacatcgagatctacaagtaccccgccctggtccgacgtggtggagatcgagtcctggggccaggcgagggcaagatcggcacccgcccgcgactggatcctgcgcgactacgccaccggccaggtgatcggccgcgcacacctccaagtgggtgatgatgaaccagacacccgcccgcctgcagaaggtggacgcggacgtgcgcgacgagtagctggtgactgccccgcgagctgcgcctggccttccccgaggagaacaactcctcctgaagaagatctccaagctggaggacccctcccagtagtccaagctgggctggtgccccgcgcgcgcgacctggacatgaaccagcacgtgaacaacgtgacctacatcggtgggtgctggagtccatgccccaggagatcatcgacacccacgagctgcagaccatcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactccctgacctcccccgagccctccgaggacgcccaggccgtgttcaaccacaacggcaccaacggctccgccaacgtgtccgccaacgaccacggctgccgcaacttctgcacctgctgcgcctgtccggcaacggcctggagatcaaccgcggccgcacccgagtggtggcgcaagaagcccacccgc**atggactacaaggaccacgacggcgactacaaggaccacgacatcgactacaaggacgacgacgacaag**^{TGA}atcgat

SEQ ID NO: 52 - secuencia de nucleótidos del péptido de tránsito CpSAD fusionado con el gen GarmFATA1 (G96A, G108A, S111A) y la etiqueta 3X FLAG en pSZ6005 y pSZ6028

actagtATGgccaccgcatccactttctcggcggttcaatgcccgctgcggcgacctgcgtcgctcggcgggctccgggccccggcgcccagcgaggccccctccccgtgcgcg**gggcgcgcc**atcccccccgcatcatcgtggtgtcctcctcctcctccaaggtgaacccccctgaagaccgagggcgtggtgtcctccggcctggccgaccgcctgcgcctgggctccctgaccgaggacggcctgtcctacaaggagaagtcatcgtgcgtgctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcaggaagtgggtgcaaccacgcccagtcctggtgggtactccaccgcccgttctccaccacccccaccatgcgaagctgcgcctgatctgggtgaccgcccgcacatcgagatctacaagtaccccgccctggtccgacgtggtggagatcgagtcctggggccaggcgagggcaagatcggcacccgcccgcgactggatcctgcgcgactacgccaccggccaggtgatcggccgcgcacacctccaagtgggtgatgatgaaccagacacccgcccgcctgcagaaggtggacgtggacgtgcgcgacgagtagctggtgactgccccgcgagctgcgcctggccttccccgaggagaacaactcctcctgaagaagatctccaagctggaggacccctcccagtagtccaagctgggctggtgccccgcgcgcgcgacctggacatgaaccagcacgtgaacaacgtgacctacatcggtgggtgctggagtccatgccccaggagatcatcgacacccacgagctgcagaccatcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactccctgacctcccccgagccctccgaggacgcccaggccgtgttcaaccacaacggcaccaacggctccgccaacgtgtccgccaacgaccacggctgccgcaacttctgcacctgctgcgcctgtccggcaacggcctggagatcaaccgcggccgcacccgagtggtggcgcaagaagcccacccgc**atggactacaaggaccacgacggcgactacaaggaccacgacatcgactacaaggacgacgacgacaag**^{TGA}atcgat

SEQ ID NO: 53 - secuencia de nucleótidos del péptido de tránsito CpSAD fusionado con el gen GarmFATA1 (G96A, G108A, V193A) y la etiqueta 3X FLAG en pSZ5984 y pSZ6021

actagtATGgccaccgcatccactttctcggcggttcaatgcccgctgcggcgacctgcgtcgctcggcgggctccgggccccggcgcccagcgaggccccctccccgtgcgcg**gggcgcgcc**atcccccccgcatcatcgtggtgtcctcctcctcctccaaggtgaacccccctgaagaccgagggcgtggtgtcctccggcctggccgaccgcctgcgcctgggctccctgaccgaggacggcctgtcctacaaggagaagtcatcgtgcgtgctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcaggaagtgggtgcaaccacgcccagtcctggtgggtactccaccgcccgttctccaccacccccaccatg

cgcaagctgcgccctgatctgggtgaccgcccgcacatcgagatctacaagtaacccgcctggt
ccgacgtggtggagatcgagtcctggggccaggcgagggcaagatcggcacccgcgcgactggat
cctgcgcgactacgccaccggccaggtgatcgccgcgcacacctccaagtgggtgatgatgaaccag
gacacccgcgcctgcagaaggtggacgcggacgtgcgcgacgagtaacctggtgactgccccgcg
agctgcgccctggccttccccgaggagaacaactcctcctgaagaagatctccaagtggaggacc
ctcccagtaactccaagtgggctggtgccccgcgcgcgcgacctggacatgaaccagcacgtgaac
aacgtgacctacatcggtggtggtgagtgccatgccccaggagatcatcgacacccacgagctgc
agaccatcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactcctgacctcccc
cgagccctccgaggacgccgaggccgtgttcaaccacaacggcaccaacggctccgccaacgtgtcc
gccaacgaccacggctgccgcaacttctgacctgtgcgccctgtccggcaacggcctggagatca
accgcgccgcacccgagtggtgcaagaagcccacccgcgatggactacaaggaccacgacggcgacta
caaggaccacgacatcgactacaaggacgacgacgacaag^{TGA}atcgat

SEQ ID NO: 54 - secuencia de nucleótidos del péptido de tránsito CpSAD fusionado con el gen GarmFATA1 (G96A, S111A, V193A) y la etiqueta 3X FLAG en pSZ6004

actagt^{ATG}gccaccgcatccactttctcggcgttcaatgcccgcctgcgggcgacctgcgctcgctcgg
gggctccggggcccccggcgcacagcgaggccccctccccgtgcgcggggcgcgcccatccccccccgcac
catcggtggtgtcctcctcctcctccaaggtgaacccccctgaagaccgagggcgtggtgtcctccggc
ctggccgacccgctgcgccctgggctccctgaccgaggacggcctgtcctacaaggagaagtcatcg
tgcgctgctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcagga
ggtggcggtgcaaccacgcccagtcctggtgggtactccaccggcggttcgccaccacccccaccatg
cgcaagctgcgccctgatctgggtgaccgcccgcacatcgagatctacaagtaacccgcctggt
ccgacgtggtggagatcgagtcctggggccaggcgagggcaagatcggcacccgcgcgactggat
cctgcgcgactacgccaccggccaggtgatcgccgcgcacacctccaagtgggtgatgatgaaccag
gacacccgcgcctgcagaaggtggacgcggacgtgcgcgacgagtaacctggtgactgccccgcg
agctgcgccctggccttccccgaggagaacaactcctcctgaagaagatctccaagtggaggacc
ctcccagtaactccaagtgggctggtgccccgcgcgcgcgacctggacatgaaccagcacgtgaac
aacgtgacctacatcggtggtggtgagtgccatgccccaggagatcatcgacacccacgagctgc
agaccatcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactcctgacctcccc
cgagccctccgaggacgccgaggccgtgttcaaccacaacggcaccaacggctccgccaacgtgtcc
gccaacgaccacggctgccgcaacttctgacctgtgcgccctgtccggcaacggcctggagatca
accgcgccgcacccgagtggtgcaagaagcccacccgcgatggactacaaggaccacgacggcgacta
caaggaccacgacatcgactacaaggacgacgacgacaag^{TGA}atcgat

SEQ ID NO: 55 - secuencia de nucleótidos del péptido de tránsito CpSAD fusionado con el gen GarmFATA1 (G108A, S111A, V193A) y la etiqueta 3X FLAG en pSZ5985 y pSZ6022

actagt^{ATG}gccaccgcatccactttctcggcgttcaatgcccgcctgcgggcgacctgcgctcgctcgg
gggctccggggcccccggcgcacagcgaggccccctccccgtgcgcggggcgcgcccatccccccccgcac
catcggtggtgtcctcctcctcctccaaggtgaacccccctgaagaccgagggcgtggtgtcctccggc
ctggccgacccgctgcgccctgggctccctgaccgaggacggcctgtcctacaaggagaagtcatcg
tgcgctgctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcagga
ggtgggctgcaaccacgcccagtcctggtgggtactccaccggcggttcgccaccacccccaccatg
cgcaagctgcgccctgatctgggtgaccgcccgcacatcgagatctacaagtaacccgcctggt
ccgacgtggtggagatcgagtcctggggccaggcgagggcaagatcggcacccgcgcgactggat
cctgcgcgactacgccaccggccaggtgatcgccgcgcacacctccaagtgggtgatgatgaaccag
gacacccgcgcctgcagaaggtggacgcggacgtgcgcgacgagtaacctggtgactgccccgcg
agctgcgccctggccttccccgaggagaacaactcctcctgaagaagatctccaagtggaggacc
ctcccagtaactccaagtgggctggtgccccgcgcgcgcgacctggacatgaaccagcacgtgaac
aacgtgacctacatcggtggtggtgagtgccatgccccaggagatcatcgacacccacgagctgc
agaccatcacctggactaccgcccgcgagtgccagcacgacgacgtggtggactcctgacctcccc
cgagccctccgaggacgccgaggccgtgttcaaccacaacggcaccaacggctccgccaacgtgtcc
gccaacgaccacggctgccgcaacttctgacctgtgcgccctgtccggcaacggcctggagatca

accgcggccgcaccgagtgggcgcaagaagcccacccgcatggactacaaggaccacgacggcgacta
caaggaccacgacatcgactacaaggacgacgacgacaagTGAatcgat

SEQ ID NO: 56 - secuencia de nucleótidos del péptido de tránsito CpSAD fusionado con el gen GarmFATA1 (G96A, G108A, S111A, V193A) y la etiqueta 3X FLAG en pSZ5987

actagtATGgccaccgcatccactttctcggcggttcaatgcccgctgcgggcgacctgcgtcgetcgg
cgggctccgggccccggcgcccagcgaggccccctccccgtgcgcgggcgcgccatccccccccgc
catcgtggtgtcctcctcctcctccaaggtgaacccccctgaagaccgagggcgtggtgtcctccggc
ctggccgaccgcctgcgcctgggctccctgaccgaggacggcctgtcctacaaggagaagtcatcg
tgcgctgctacgaggtgggcatcaacaagaccgccaccgtggagaccatcgccaacctgctgcagga
ggtggcggtgcaaccacgcccagtcctgtgggctactccaccgcccgttcgccaccacccccaccatg
cgcaagctgcgcctgatctgggtgaccgcccgcgatgcacatcgagatctacaagtaccccgctggt
ccgacgtggtggagatcgagtcctggggccaggggcgagggcaagatcggcaccgcgcgcgactggat
cctgcgcgactacgccaccggccaggtgatcgccgcgcgcacacctccaagtgggtgatgatgaaccag
gacaccgcgcgcctgcagaaggtggacgcggacgtgcgcgcgacgagtaacctggtgactgccccgcg
agctgcgcctggccttccccgaggagaacaactcctccctgaagaagatctccaagctggaggaccc
ctcccagtaactccaagctgggctggtgccccgcgcgcgcgacactggacatgaaccagcacgtgaac
aacgtgacctacatcggctgggtgctggagtccatgccccaggagatcatcgacaccacgagctgc
agaccatcacctggactaccgcgcgagtgccagcacgacgacgtggtggactcctgacctcccc
cgagccctccgaggacgcgcgaggccgtgttcaaccacaacggcaccaacggctccgccaacgtgtcc
gccaacgaccacggctgcgcgaacttctgcacctgctgcgcctgtccggcaacggcctggagatca
accgcggccgcaccgagtgggcgcaagaagcccacccgcatggactacaaggaccacgacggcgacta
caaggaccacgacatcgactacaaggacgacgacgacaagTGAatcgat

SEQ ID NO: 57 - secuencia de nucleótidos del promotor PmACP-P1 en pSZ6019-pSZ6023, pSZ6026 y pSZ6028

Gaattc	gcctgctcaagcgggcgctcaacatgcagagcgtcagcgagacgggctgtggcgatcgcg
	gacggacgaggccgcctctgcctggttgaactgagcgtcagcgctggctaaggggagggagactca
	tccccaggctcgcgccagggtctgatcccgctctcgggcggtgatcggcgcgcatgactacgaccca
	acgacgtacgagactgatgtcgggtcccgacgaggagcgccgcgaggcactcccgggccaccgacat
	gtttacaccgaccgaaagcactcgctcgtatccattccgtgcgcccgcacatgcacatcttttgggt
	accgacttcggtcttgttttaaccctacgacctgccttccaaggtgtgagcaactcgcccgacatg
	accgaggggtgatcatccgatccccaggccccagcagccccctgccagaatggctcgcgctttccagc
	ctgcaggcccgtctccaggtcgacgcaacctacatgaccacccaatctgtcccagaccccaaca
	ccctccttcctgcttctctgtgatcgctgatcagcaaca
	actagt

REIVINDICACIONES

1. Acil graso-ACP tioesterasa no natural que tiene al menos el 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad con SEQ ID NO: 45 y comprende una alanina (A) en la posición correspondiente a la posición 74 de SEQ ID NO: 45 (G96).
- 5 2. Tioesterasa no natural según la reivindicación 1, que también comprende una alanina (A) en la posición correspondiente a la posición 89 de SEQ ID NO: 45 (S111) y/o una alanina (A) en la posición correspondiente a la posición 171 de SEQ ID NO: 45 (V193) y/o una alanina (A) en la posición correspondiente a la posición 86 de SEQ ID NO: 45 (G108).
- 10 3. Gen aislado que codifica una tioesterasa no natural, donde la tioesterasa no natural tiene al menos el 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad con SEQ ID NO: 45 y comprende una alanina (A) en la posición correspondiente a la posición 74 de SEQ ID NO: 45 (G96).
- 15 4. Gen aislado según la reivindicación 3, que codifica una tioesterasa no natural, donde la tioesterasa no natural también comprende una alanina (A) en la posición correspondiente a la posición 89 de SEQ ID NO: 45 (S111) y/o una alanina (A) en la posición correspondiente a la posición 171 de SEQ ID NO: 45 (V193) y/o una alanina (A) en la posición correspondiente a la posición 86 de SEQ ID NO: 45 (G108).
5. Casete de expresión que expresa una tioesterasa no natural, donde la tioesterasa no natural tiene al menos el 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad con SEQ ID NO: 45 y comprende una alanina (A) en la posición correspondiente a la posición 74 de SEQ ID NO: 45 (G96).
- 20 6. Casete de expresión según la reivindicación 5, que expresa una tioesterasa no natural, donde la tioesterasa no natural también comprende una alanina (A) en la posición correspondiente a la posición 89 de SEQ ID NO: 45 (S111) y/o una alanina (A) en la posición correspondiente a la posición 171 de SEQ ID NO: 45 (V193) y/o una alanina (A) en la posición correspondiente a la posición 86 de SEQ ID NO: 45 (G108).
7. Célula huésped que comprende el casete de expresión según la reivindicación 5 o la reivindicación 6.
8. Célula huésped según la reivindicación 7, donde la célula es una célula de microalga oleaginosa.
- 25 9. Célula huésped según la reivindicación 8, donde la célula es una célula de *Prototheca*.
10. Célula huésped según la reivindicación 9, donde la célula es una célula de *Prototheca moriformis*.
11. Método para producir un aceite de triglicéridos, donde el método comprende expresar, en una célula huésped, la tioesterasa no natural según la reivindicación 1 o la reivindicación 2; cultivar la célula huésped; y aislar el aceite.
- 30 12. Método para producir un aceite de triglicéridos, donde el método comprende expresar, en una célula huésped, el casete de expresión según la reivindicación 5 o la reivindicación 6; cultivar la célula huésped; y aislar el aceite.
- 35 13. Método para aumentar los ácidos grasos C18:0 en un perfil de ácidos grasos de un aceite producido por una célula huésped opcionalmente oleaginosa, donde el método comprende proporcionar un gen parental que codifique una enzima acil graso-ACP tioesterasa A (FATA), mutar el gen para codificar una tioesterasa no natural que tiene al menos el 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad con SEQ ID NO: 45 y comprende una alanina (A) en la posición correspondiente a la posición 74 de SEQ ID NO: 45 (G96); expresar el gen mutado en la célula huésped; y producir el aceite, por el cual los ácidos grasos C18:0 son incrementados en el perfil de ácidos grasos del aceite.
- 40 14. Método según la reivindicación 13, que también comprende mutar el gen para codificar una tioesterasa no natural que tiene una alanina (A) en la posición correspondiente a la posición 89 de SEQ ID NO: 45 (S111) y/o una alanina (A) en la posición correspondiente a la posición 171 de SEQ ID NO: 45 (V193) y/o una alanina (A) en la posición correspondiente a la posición 86 de SEQ ID NO: 45 (G108).

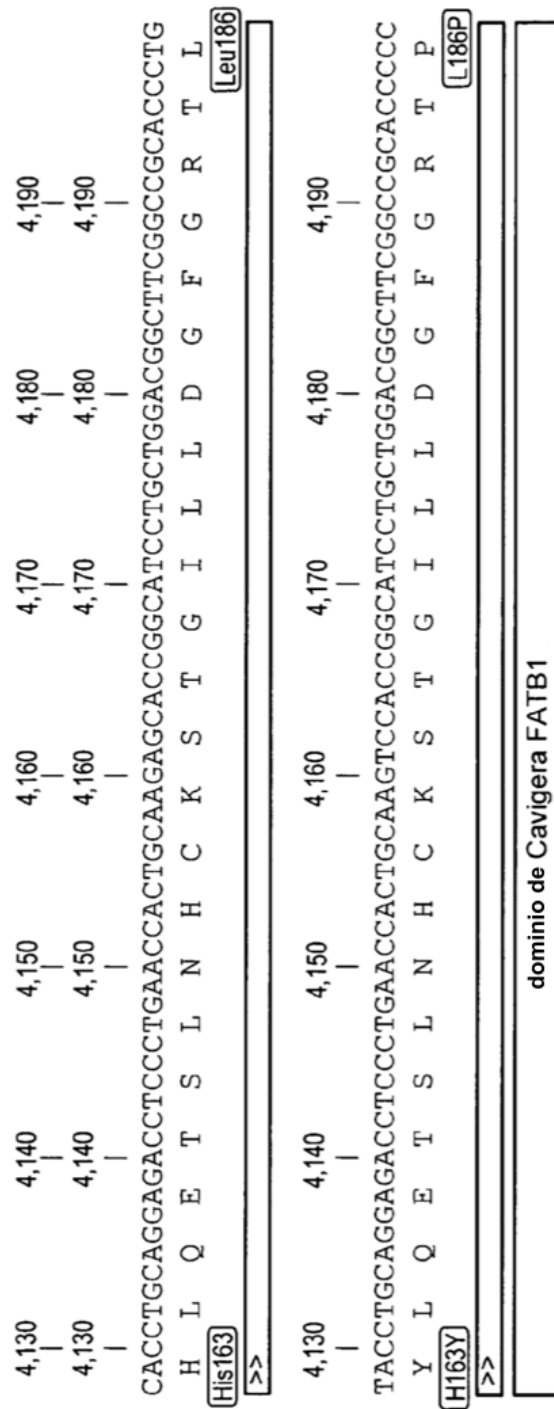


Figura 1

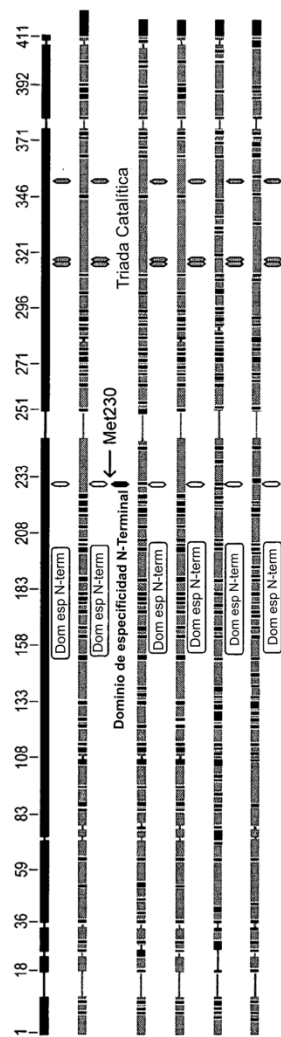


Figura 2A



Figura 2B

Figura 3A

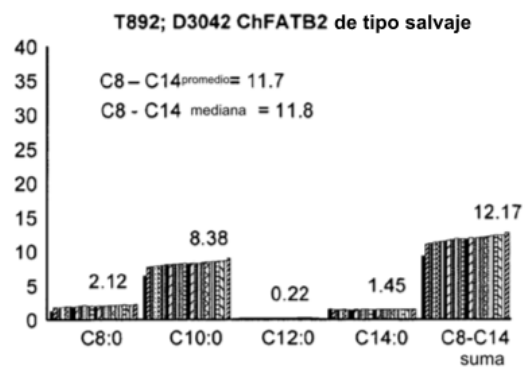


Figura 3B

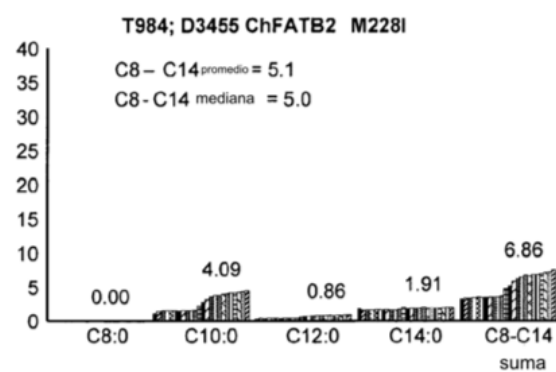


Figura 3C

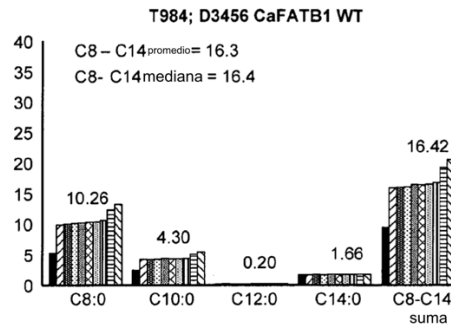


Figura 3D

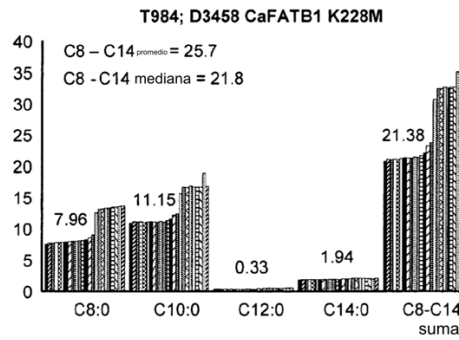


Figura 3E

