(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 September 2020 (10.09.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/181091 A1

(51) International Patent Classification:
C12N 9/02 (2006.01) *C07K 14/195* (2006.01)
C12P 7/04 (2006.01)

(21) International Application Number:
PCT/US2020/021199

(22) International Filing Date:
05 March 2020 (05.03.2020)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/813,796 05 March 2019 (05.03.2019) US

(71) Applicant: **BP CORPORATION NORTH AMERICA INC.** [US/US]; 501 Westlake Park Boulevard, Houston, Texas 77079 (US).

(72) Inventors: **VALLE, Fernando**; 150 W. Warrenville Road, MC 200-1W, Naperville, Illinois 60563 (US). **KIM, Unmi**; 150 W. Warrenville Road, MC 200-1W, Naperville, Illinois 60563 (US).

(74) Agent: **GATTARI, Patrick G.**; Mcdonnell Boehnen Hulbert & Berghoff LLP, 300 South Wacker Drive, Suite 3100, Chicago, IL 60606 (US).

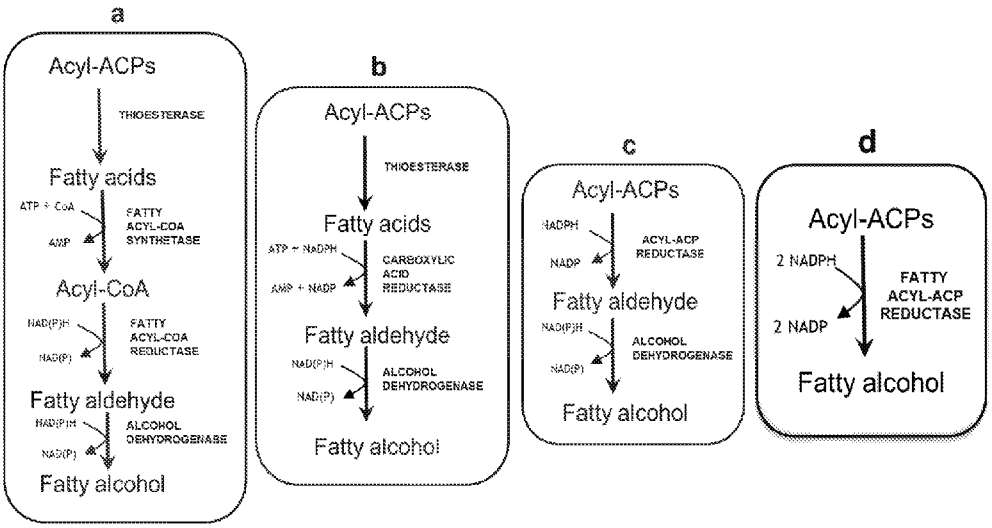
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

(54) Title: FATTY ACYL-ACP REDUCTASES AND THEIR USES

FIGURE 1



(57) Abstract: Fatty acyl-ACP reductase (FAR) enzymes are provided, along with their use in the production of fatty alcohol compositions in heterologous recombinant cells.

WO 2020/181091 A1

FATTY ACYL-ACP REDUCTASES AND THEIR USES

FIELD OF THE INVENTION

The invention relates to fatty acyl-ACP reductase (FAR) enzymes and their use in the production of fatty alcohol compositions.

BACKGROUND

Fatty alcohols (FOHs) are used for the commercial production of surfactants, soaps, detergents, cosmetics, lubricants, plastics, etc. More than 70% of commercial fatty alcohols are obtained from plant-derived oils, particularly palm oils. Between the years 2000 and 2010, the amount of fatty alcohols produced from palm and palm kernel oil grew by 100% (Biermann, Bornscheuer *et al.* 2011).

In 2010, approximately 50 million tons of palm and palm kernel oil were produced, representing more than 1/3 of the annual global production of all plant-derived oils (Biermann, Bornscheuer *et al.* 2011). Continuous growth of this industry and concerns over rainforest destruction have resulted in the implementation of more sustainable palm oil processes, as well as the development of more environmentally friendly approaches via production of fatty alcohols by fermentation of recombinant microorganisms. Several alternative biosynthetic routes have been developed (Reiser and Somerville 1997, Schirmer, Rude *et al.* 2010, Steen, Kang *et al.* 2010, Zheng, Li *et al.* 2012, Akhtar, Turner *et al.* 2013) and a summary of biosynthetic routes is presented in Figure 1. As can be seen, all four routes (corresponding to Figure 1, parts a-d) start with acyl-ACPs (also referred to herein as "fatty acyl-ACPs"). Acyl-ACPs are a common intermediary of the fatty acid biosynthetic pathway *e.g.* in bacteria. The differences between these four routes are the type, and number of enzymes required to produce fatty alcohols, as well as the energy (ATP), reducing equivalents (NAD(P)H) and the requirement for acyl-CoA. One disadvantage of routes a-c is that multiple enzymes are required, and so to optimize the production process, each pathway must be optimized for the recombinant cell used. Furthermore, the routes shown in Figure 1 parts a-c lead to the production of free fatty aldehydes as intermediates and the routes shown in Figure 1 parts a and b also lead to the production of fatty acids as intermediates. These intermediates are undesirable because aldehydes are known to be very reactive, while fatty acids have been shown to cause membrane stress (Lennen, Kruziki *et al.* 2011). A more attractive and efficient route (shown in Figure 1 part d) involves the direct conversion of fatty acyl-ACPs into fatty alcohols using a single FAR enzyme which catalyzes the complete reduction to alcohol, without the liberation of undesirable intermediates.

Certain plant-derived and bacterial FARs can utilize *E. coli* fatty acyl-ACPs, and directly produce fatty alcohols (Metz, Pollard *et al.* 2000, Doan, Carlsson *et al.* 2009, Hofvander, Doan *et al.* 2011). However, plant FARs preferentially use acyl-CoA intermediates that are not normally present in

bacteria and these enzymes utilize *E. coli* fatty acyl-ACPs with low efficiency. Plant plastid-located FAR enzymes that preferentially use fatty acyl-ACPs as substrates have now been reported (for a review see Rowland and Domergue 2012). *In vitro* characterization of bacterial FAR identified from *Marinobacter aquaeolei* VT8 suggests that bacterial FARs have a wider range of specificity and higher activity than plant FARs (Willis, Wahlen *et al.* 2011).

Bacterial FARs are more active than plant FARs and therefore represent ideal candidates for the production of fatty alcohols from fatty acyl-ACPs. However, only a limited number of functional bacterial FARs have been reported. For example, US 8,216,815 mentions fatty alcohol production by FARs identified in *Marinobacter algicola* and *Oceanobacter* sp. RED65 that directly produce fatty alcohols from fatty acyl-ACP in *E. coli*. Liu *et al.* also mentioned the use of *Marinobacter aquaeolei* FAR enzyme (Liu, Chen *et al.* 2016). Previously known FARs (such as FAR identified in *Marinobacter algicola*, and reported in US 8,216,815) are capable of producing fatty alcohols as short as C14, but these previously known FARs have not been demonstrated to produce highly desirable fatty alcohols having a shorter chain length. Considerable efforts have been dedicated to engineering known enzymes to produce shorter chain length fatty alcohols but, so far as the Inventors are aware, such attempts have not been successful enough to enable commercialization of a FAR-based process.

Based on the limited number of active FARs reported in the literature, there is an urgent and unmet need to identify new FAR enzymes that can be used for the production of fatty alcohols. There is likewise an urgent and unmet need to identify new FAR enzymes that can produce fatty alcohols with a chain length of less than C14.

SUMMARY OF THE INVENTION

The present invention provides new fatty acyl-ACP reductase (FAR) enzymes. The FAR enzymes of the invention possess highly desirable characteristics including *e.g.*: (i) efficient production of fatty alcohols from acyl-ACPs involving a single enzyme (i.e. the polypeptide of the invention); (ii) production of fatty alcohols without undesirable production of aldehyde and fatty acid intermediates; (iii) production of a high yield of fatty alcohols; (iv) production of a high proportion of saturated fatty alcohols; (v) production of a high proportion of highly desirable C14:0 fatty alcohol (also known as Myristyl alcohol); and (vi) production of highly desirable C12 fatty alcohol.

The invention provides a recombinant polypeptide comprising an amino acid sequence having at least 70% sequence identity to SEQ ID NO: 1.

The invention also provides an isolated nucleic acid which encodes a polypeptide of the invention. In some embodiments, the nucleic acid is codon optimized for expression in a prokaryotic cell. In some embodiments, the prokaryotic cell is selected from the group consisting of *Acinetobacter*,

Agrobacterium, *Escherichia*, *Cupriavidus*, *Clostridium*, *Rhodobacter*, *Marinobacter*, *Bacillus*, *Klebsiella*, *Tatumella*, *Pseudomonas*, *Ralstonia*, *Rhodococcus*, *Methylobacterium*, *Methylophilus*, *Methylococcus*, *Methylomicrobium*, *Methylomonas*, *Pantoea*, *Streptomyces*, *Parachlorella*, *Synechococcus*, *Synechocystis* and *Thermocynechococcus*.

In some embodiments, the prokaryotic cell is *E. coli*. In some embodiments, the nucleic acid sequence has at least 70% sequence identity to SEQ ID NO: 2.

In some embodiments, the nucleic acid sequence is codon optimized for expression in a eukaryotic cell. In some embodiments, the eukaryotic cell is selected from the group consisting of a yeast cell, a fungal cell, an algal cell and a plant cell.

In some embodiments, the yeast cell is selected from the group consisting of *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Schefferomyces*, *Rhodosporidium*, *Hansenula*, *Klockera*, *Schwanniomyces*, *Issatchenkia*, *Yarrowia* and *Rhodotorula*.

In some embodiments, the yeast cell is selected from the group consisting of *S. cerevisiae*, *C. lipolytica*, *R. glutinis*, *S. bulderi*, *S. barnetti*, *S. exiguus*, *S. uvarum*, *S. diastaticus*, *K. lactis*, *K. marxianus K. fragile*, *P. kudriavzevii*, *S. stipitis* and *I. orientalis*.

In some embodiments, the fungal cell is a filamentous fungal cell. In some embodiments, the filamentous fungal cell is selected from the group consisting of *Aspergillus*, *Penicillium*, *Rhizopus*, *Chrysosporium*, *Myceliophthora*, *Trichoderma*, *Humicola*, *Acremonium* and *Fusarium*. In some embodiments, the filamentous fungal cell is selected from the group consisting of *A. niger*, *A. oryzae*, *T. reesei*, *P. chrysogenum*, *M. thermophila*, and *R. oryzae*.

In some embodiments, the algal cell is selected from the group consisting of *Botryococcus*, *Nannochloropsis*, *Chlorella*, *Chlamydomonas*, *Dunaliella*, *Chaetoceros*, *Porphyridium*, *Scenedesmus* and *Pseudochlorococcum*. In some embodiments, the algal cell is selected from the group consisting of *B. braunii* and *N. gaditana*.

In some embodiments, the nucleic acid of the invention comprises a nucleic acid sequence having at least 70% sequence identity to SEQ ID NO: 3.

The invention also provides a vector comprising the nucleic acid of the invention. In some embodiments, the vector comprises: (a) an origin of replication; (b) a promoter sequence operably linked to said nucleic acid; and/or (c) a reporter gene. In some embodiments, the vector comprises a nucleic acid sequence having at least 70% sequence identity to SEQ ID NO: 4.

The invention also provides a recombinant cell engineered to express a polypeptide of the invention. In some embodiments, the recombinant cell is a prokaryotic cell. In some embodiments, the prokaryotic cell is selected from *Acinetobacter*, *Agrobacterium*, *Escherichia*, *Cupriavidus*,

Clostridium, Rhodobacter, Marinobacter, Bacillus, Klebsiella, Tatumella, Pseudomonas, Ralstonia, Rhodococcus, Methylobacterium, Methylophilus, Methylococcus, Methylomicrobium, Methylomonas, Pantoea, Streptomyces, Parachlorella, Synechococcus, Synechocystis and *Thermocynechococcus*.

In some embodiments, the prokaryotic cell is *E. coli*. In some embodiments, the recombinant cell is a eukaryotic cell. In some embodiments, the eukaryotic cell is selected from the group consisting of a yeast cell, a fungal cell, an algal cell and a plant cell.

In some embodiments, the yeast cell is selected from the group consisting of *Saccharomyces, Kluyveromyces, Candida, Pichia, Schizosaccharomyces, Schefferomyces, Rhodosporidium, Hansenula, Klockera, Schwanniomyces, Issatchenkia, Yarrowia* and *Rhodotorula*. In some embodiments, the yeast cell is selected from the group consisting of *S. cerevisiae, C. lipolytica, R. glutinis, S. bulderi, S. barnetti, S. exiguus, S. uvarum, S. diastaticus, K. lactis, K. marxianus, K. fragile, P. kudriavzevii, S. stipitis* and *I. orientalis*.

In some embodiments, the fungal cell is a filamentous fungal cell. In some embodiments, the filamentous fungal cell is selected from the group consisting of *Aspergillus, Penicillium, Rhizopus, Chrysosporium, Myceliophthora, Trichoderma, Humicola, Acremonium* and *Fusarium*. In some embodiments, the filamentous fungal cell is selected from the group consisting of *A. niger, A. oryzae, T. reesei, P. chrysogenum, M. thermophila,* and *R. oryzae*.

In some embodiments, the algal cell is selected from the group consisting of *Botryococcus, Nannochloropsis, Chlorella, Chlamydomonas, Dunaliella, Chaetoceros, Porphyridium, Scenedesmus* and *Pseudochlorococcum*. In some embodiments, the algal cell is selected from the group consisting of *B. braunii* and *N. gaditana*.

In some embodiments, the recombinant cell of the invention comprises one or more genetic modifications resulting in at least one, any two, any three, any four or all five of the following phenotypes: (a) an increase in fatty acyl-ACP synthesis; (b) a decrease in fatty acyl-ACP degradation; (c) an upregulation of fatty alcohol secretion; (d) an increase in flux through the fatty acid biosynthetic pathway; (e) an increase in tolerance to fatty alcohol.

In some embodiments, the recombinant cell comprises one or more genetic modifications resulting in a change in the degree of saturation of the fatty alcohols produced.

In some embodiments, the recombinant cell comprises one or more genetic modifications resulting in a change in the chain length of acyl-ACPs produced by the fatty acid biosynthetic pathway.

The invention also provides a recombinant cell transformed with a vector of the invention. In some embodiments, the recombinant cell is stably transformed with the nucleic acid of the invention. In some embodiments, the recombinant cell is transiently transformed with a nucleic acid of the invention.

The invention also provides a method of producing a fatty alcohol composition comprising: a) culturing a recombinant cell in a suitable culture medium, wherein the recombinant cell comprises a gene encoding a heterologous FAR enzyme having at least 70% sequence identity to SEQ ID NO: 1; and b) allowing expression of said gene, wherein said expression results in the production of a fatty alcohol composition.

In some embodiments, the method further comprises purifying the fatty alcohols from cell culture. In some embodiments, the method further comprises purifying the fatty alcohols from supernatant. In some embodiments, the method further comprises harvesting and lysing recombinant cell(s) to obtain fatty alcohols. In some embodiments, the method further comprises purifying the fatty alcohols from lysate.

In one embodiment, the invention provides a method of producing a fatty alcohol composition comprising: a) culturing a recombinant cell in a suitable culture medium, wherein the recombinant cell comprises a gene encoding a heterologous FAR enzyme having at least 70% sequence identity to SEQ ID NO: 1; b) allowing expression of said gene, wherein said expression results in the production of a fatty alcohol composition; and c) purifying the fatty alcohols from cell culture.

In one embodiment, the invention provides a method of producing a fatty alcohol composition comprising: a) culturing a recombinant cell in a suitable culture medium, wherein the recombinant cell comprises a gene encoding a heterologous FAR enzyme having at least 70% sequence identity to SEQ ID NO: 1; b) allowing expression of said gene, wherein said expression results in the production of a fatty alcohol composition; and c) purifying the fatty alcohols from supernatant.

In one embodiment, the invention provides a method of producing a fatty alcohol composition comprising: a) culturing a recombinant cell in a suitable culture medium, wherein the recombinant cell comprises a gene encoding a heterologous FAR enzyme having at least 70% sequence identity to SEQ ID NO: 1; b) allowing expression of said gene, wherein said expression results in the production of a fatty alcohol composition; and c) harvesting and lysing recombinant cell(s) to obtain fatty alcohols.

In one embodiment, the invention provides a method of producing a fatty alcohol composition comprising: a) culturing a recombinant cell in a suitable culture medium, wherein the recombinant cell comprises a gene encoding a heterologous FAR enzyme having at least 70% sequence identity to SEQ ID NO: 1; b) allowing expression of said gene, wherein said expression results in the production of a fatty alcohol composition; c) harvesting and lysing recombinant cell(s) to obtain fatty alcohols; and d) purifying the fatty alcohols from lysate.

In some embodiments, the recombinant cell is defined according to any one of claims 20 to 36.

The invention also provides a cell-free method for producing a fatty alcohol composition, the method comprising incubating a polypeptide according to the invention with fatty acyl-ACP and a reductant. In some embodiments, the reductant is NADPH. In some embodiments, the method further comprises purifying the fatty alcohol composition.

Typically, the fatty alcohol composition produced by a method of the invention comprises C12 fatty alcohol.

In some embodiments, the method further comprises reducing the fatty alcohol composition. In some embodiments, the method further comprises esterifying the fatty alcohol composition.

The invention also provides a fatty alcohol composition obtained by the method of the invention.

The invention also provides a composition comprising fatty alcohol composition or esterified fatty alcohol composition obtained by a method of the invention.

The invention also provides use of a polypeptide of the invention in a method of producing a fatty alcohol composition.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing aspects and many of the attendant advantages will become more readily appreciated as the same become better understood by reference to the following detailed description, when taken in conjunction with the accompanying drawings.

FIGURE 1 summarizes current technologies for the production of fatty alcohols from fatty acyl-ACPs in *E. coli*. Parts a) and b) depict routes based on the use of thioesterases to produce fatty acids, in which fatty acids are converted into fatty aldehyde intermediates by different enzymes. Part c) depicts a route that directly produces fatty aldehydes, which are then converted to fatty alcohols by alcohol dehydrogenases. Part d) depicts a route that directly converts fatty acyl-ACPs into fatty alcohols using a FAR enzyme, which catalyzes the 2-step reduction of fatty acyl-ACP to fatty alcohols without liberating fatty aldehyde or fatty alcohol intermediates. FARs of the invention produce fatty alcohols *via* the route depicted in part d).

FIGURE 2 shows low copy expression vector, pCL-BP2 with spectinomycin resistance cassette and multiple cloning sites for FAR expression under the Trc inducible promoter.

FIGURE 3 shows the relative amount of total fatty alcohols produced by *E. coli* expressing a FAR gene from *Aesturaiibacter salexigens* (AsFAR), *Massilia* sp. Root335 (Mas1FAR), *Marinobacter* (Mar1FAR), *Hahella ganghwensis* (HgFAR), *Marinobacter daepoensis* (MdFAR), *Duganella* (DugFAR), *Herbaspirillum* (HerFAR), *Spongiibacter marinus* (SmFAR; strain 16), *Zhongshania aliphaticivorans* (ZaFAR), *Marinobacter* (Mar2FAR), *Marinobacter lipolyticus* (MlFAR), and

Nanorana parkei (NpFAR). Data are relative to the total amount of fatty alcohols produced by SmFAR.

DETAILED DESCRIPTION OF THE INVENTION

Fatty Acyl-ACP Reductase Polypeptides

The present invention is based on the surprising discovery of a previously unknown dehydrogenase enzyme with fatty acyl-ACP reductase (FAR) activity derived from *Spongiibacter marinus* (herein "SmFAR") that catalyzes the direct reduction of fatty acyl-ACP into fatty alcohols, without the liberation of undesirable intermediates.

"Fatty acyl-ACP reductases" are sometimes referred to in the literature as "fatty acyl-CoA reductases" or "fatty acyl-CoA/ACP reductases" on the basis that they are frequently able to catalyze Acyl-CoA, as well as Acyl-ACP.

In comparison to previously reported FARs, the *Spongiibacter marinus* fatty acyl-ACP reductase (SmFAR) polypeptide of the invention produces a fatty alcohol composition having a distinct and advantageous fatty alcohol profile. The Inventors found that SmFAR can produce highly desirable shorter chain length fatty alcohols than previously reported FAR enzymes. The enzyme of the invention also produces a desirable and advantageously high proportion of saturated fatty alcohols, as compared to previously reported FAR enzymes. Advantageously, this can help simplify downstream processing when seeking to obtain saturated product.

SmFAR also comprises a unique polypeptide sequence that has low sequence identity to previously reported FARs. Sequence identity comparisons between SmFAR and publicly available protein sequences identified the nearest homolog as having ~65% sequence identity to SmFAR. Indeed, SmFAR has about ~50% sequence homology to FAR derived from *Marinobacter hydrocarbonoclasticus* or from *Marinobactoer algicola* ("MaFAR").

A FAR polypeptide of the invention comprises an amino acid sequence having at least 70% sequence identity to SEQ ID NO: 1. In some embodiments, the FAR polypeptide of the invention has at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to SEQ ID NO: 1. In some embodiments, the FAR polypeptide of the invention has at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to a functional fragment or catalytic domain of SEQ ID NO: 1.

The SmFAR polypeptide sequence is represented by SEQ ID NO: 1:

MRDLNRDKSSTIAALTGKQVLITGTTGFLGKVLLEKVLRTTPGVGGIYLLMRGNRHFRDATDRFREEVISSSIFDTLR
HELGQEGFEALIASRVHCVTGEITAPRFGLTRTVFTALADKLDVVINSAASVNFREELDRALEINTFSLENVAKLCELG
GNIPLVHVSTCYVNGYRSGDIHETSEGPVGLELPEGPGSFYETSALIERLQRLVEELRGQYEGRALKAKLIEHGAKEA
QAAGWNDTYTFTKWLGEQYLYKVMRGYSLTVVRPSIIESTLREPSPGWIEGVKVADAVLMAYARGKVAFFPGKRA
GVIDIIPADLVANGVLLSTAEQLLEPGQHRIYQCCSGGRNPLMLGDFIDHLVAEASSHHGDYDKLFKAPPRRRFTAV
DKRVFSSVAMCVQVALTIVGAILRRLGRRGKLRARRNIDTAVELSKIFSFYAQPNYCFHGDKLLSLSKSMGAVDQEL
FPVDAGVIDWQHYIRHVHMPGLNHYALQGQPSVRSQQRREEVKTAPKNLTATEA **(SEQ ID NO: 1)**

An example of an algorithm that is suitable for determining sequence similarity is the BLAST algorithm, which is described in Altschul *et al.*, 1990, J. Mol. Biol. 215:403-410. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence that either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. These initial neighborhood word hits act as starting points to find longer HSPs containing them. The word hits are expanded in both directions along each of the two sequences being compared for as far as the cumulative alignment score can be increased. Extension of the word hits is stopped when: the cumulative alignment score falls off by the quantity X from a maximum achieved value; the cumulative score goes to zero or below; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLAST program uses as defaults a word length (W) of 11, the BLOSUM62 scoring matrix (see Henikoff & Henikoff, 1992, Proc. Nat'l. Acad. Sci. USA 89:10915-10919) alignments (B) of 50, expectation (E) of 10, M'5, N'-4, and a comparison of both strands.

Any of the amino acid sequences described herein can be produced together or in conjunction with at least 1, *e.g.*, at least (or up to) 2, 3, 5, 10, or 20 heterologous amino acids flanking each of the C- and/or N-terminal ends of the specified amino acid sequence, and or deletions of at least 1, *e.g.*, at least (or up to) 2, 3, 5, 10, or 20 amino acids from the C- and/or N-terminal ends.

FAR polypeptides of the invention may have one or more (*e.g.*, up to 2, 3, 5, 10, 20, 30, 40, or 50) conservative amino acid substitutions relative to the polypeptide of SEQ ID NO: 1. Conservative substitutions can be chosen from among a group of amino acids having a similar side chain to the reference amino acid. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a

group of amino acids having sulphur-containing side chains is cysteine and methionine. Accordingly, exemplary conservative substitutions for each of the naturally occurring amino acids are as follows: Ala to Ser; Arg to Lys; Asn to Gln or His; Asp to Glu; Cys to Ser or Ala; Gln to Asn; Glu to Asp; Gly to Pro; His to Asn or Gln; Ile to Leu or Val; Leu to Ile or Val; Lys to Arg; Gln or Glu; Met to Leu or Ile; Phe to Met, Leu or Tyr; Ser to Thr; Thr to Ser; Trp to Tyr; Tyr to Trp or Phe; and, Val to Ile or Leu.

The invention also provides a fusion protein that includes at least a portion (*e.g.*, a fragment or domain) of a FAR polypeptide of the invention attached to one or more fusion segments, which are typically heterologous to the FAR polypeptide. Suitable fusion segments include, without limitation, segments that can provide other desirable biological activity or facilitate purification of the FAR polypeptide (*e.g.*, by affinity chromatography). Fusion segments can be joined to the amino or carboxy terminus of a FAR polypeptide. The fusion segments can be susceptible to cleavage.

Fatty Acyl-ACP Reductase Nucleic Acids

A FAR nucleic acid of the invention encodes a FAR polypeptide of the invention. The FAR nucleic acid may be isolated, synthetic or recombinant.

A FAR nucleic acid of the invention may be optimized for expression in a recombinant cell. "Optimized" nucleic acid sequences encode an amino acid sequence using codons that are preferred in a recombinant cell. The optimized nucleic acid sequence is typically engineered to retain completely or as much as possible of the amino acid sequence originally encoded by the starting nucleic acid sequence, which is also known as the "parental" sequence. Several methods for codon optimization are known in the art. Preferably, codon optimized sequences avoid nucleotide repeats and restriction sites that are utilized in cloning the FAR nucleic acids, by adjusting the settings in commercial software or by manually altering the sequences to substitute codons that introduce undesired sequences, for example with highly utilized codons in the heterologous organism of interest.

In some embodiments, the nucleic acid is codon optimized for expression in a prokaryotic cell. In some embodiments, the nucleic acid is codon optimized for expression in *Acinetobacter, Agrobacterium, Escherichia, Cupriavidus, Clostridium, Rhodobacter, Marinobacter, Bacillus, Klebsiella, Tatumella, Pseudomonas, Ralstonia, Rhodococcus, Methylobacterium, Methylophilus, Methylococcus, Methylomicrobium, Methylomonas, Pantoea, Streptomyces, Parachlorella, Synechococcus, Synechocystis* and *Thermocynechococcus*. In some embodiments, the nucleic acid is codon optimized for expression in *E. coli*. In some embodiments, the nucleic acid sequence has at least 70% sequence identity to SEQ ID NO: 2. In some embodiments, the nucleic acid sequence has at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least

98%, or at least 99%, or 100% sequence identity to SEQ ID NO: 2, over a region of at least about 10, *e.g.*, at least about 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, or 1557 nucleotides, or a catalytic domain thereof.

SEQ ID NO: 2 represents a nucleic acid sequence corresponding to the SmFAR polypeptide (SEQ ID NO: 1) codon optimized for expression in *E. coli*:

ATGCGTGATTTAAACCGCGATAAAAGTAGCACCATTGCCGCACTGACCGGCAAACAGGTGCTGATCACCGGC
ACCACCGGCTTTCTGGGTAAGGTGCTGCTGGAAAAAGTGCTGCGTACAACCCCGGGTGTTGGCGGCATTTATC
TGCTGATGCGCGGCAACCGTCACTTCCGCGACGCAACAGATCGTTTCCGCGAAGAAGTTATCAGTAGCAGCAT
CTTCGACACACTGCGCCATGAACTGGGTCAAGAAGGTTTTGAGGCCCTGATCGCAAGCCGCGTTCATTGTGTG
ACCGGCGAAATTACAGCCCCGCGCTTCGGTTTAACCCGCACCGTGTTTACCGCCCTGGCCGATAAACTGGACG
TGGTGATCAACAGCGCAGCCAGCGTGAACTTTCGTGAAGAACTGGATCGCGCCCTGGAGATCAATACCTTCA
GTCTGGAGAACGTGGCCAAACTGTGTGAGCTGGGCGGTAATATCCCTCTGGTGCACGTGAGCACCTGCTACG
TTAACGGCTACCGTAGCGGCGACATTCACGAAACTAGTGAAGGCCCGGTGGGTTTAGAATTACCGGAAGGCC
CGGGCAGCTTCTATGAAACTAGTGCCCTGATTGAACGCTTACAGCGTCTGGTGGAGGAACTGCGCGGCCAGT
ATGAGGGTCGTGCCCTGAAAGCCAAGCTGATTGAACATGGCGCCAAGGAAGCACAGGCAGCCGGCTGGAAC
GACACCTATACCTTCACAAAGTGGCTGGGTGAGCAGTACCTGTACAAGGTGATGCGCGGCTATAGTCTGACC
GTTGTGCGTCCGAGCATTATCGAGAGCACACTGCGTGAACCGAGTCCTGGTTGGATCGAAGGTGTGAAGGTT
GCAGACGCCGTGCTGATGGCATATGCCCGCGGCAAAGTGGCATTTTTTCCGGGCAAGCGCGCCGGTGTGATT
GATATCATCCCGGCAGATCTGGTGGCCAATGGTGTTCTGCTGAGCACAGCAGAACAGCTGCTGGAACCGGGC
CAACATCGCATCTATCAGTGTTGCAGTGGCGGTCGCAACCCGCTGATGCTGGGCGATTTTATTGACCACCTGG
TGGCCGAAGCCAGCAGCCATCACGGCGACTACGACAAACTGTTTAAAGCACCGCCGCGTCGCCGTTTTACCGC
CGTTGACAAGCGTGTGTTTAGCAGCGTGGCAATGTGCGTTCAGGTGGCCCTGACCATTGTTGGTGCAATCCTG
CGTCGTCTGGGCCGTCGTGGTAAGCTGCGTGCCCGTCGTAACATTGACACCGCAGTGGAGCTGAGCAAAATC
TTCAGCTTTTACGCCCAGCCGAACTACTGCTTTCACGGTGACAAACTGCTGAGCCTGAGCAAAAGCATGGGTG
CCGTGGATCAAGAACTGTTTCCGGTGGACGCAGGTGTGATCGACTGGCAGCACTATATTCGCCACGTTCACAT
GCCGGGCCTGAACCATTACGCACTGCAGGGTCAGCCGAGTGTGCGTAGCCAGCAACGCCGCGAGGAAGTGA
AAACCGCCCCTAAGAATCTGACCGCCACCGAGGCATGA **(SEQ ID NO: 2)**

In some embodiments, the nucleic acid is codon optimized for expression in a eukaryotic cell. In some embodiments, the nucleic acid is codon optimized for expression in a yeast cell, a fungal cell, an algal cell and a plant cell. In some embodiments, the nucleic acid is codon optimized for expression in *Saccharomyces, Kluyveromyces, Candida, Pichia, Schizosaccharomyces, Schefferomyces, Rhodosporidium, Hansenula, Klockera, Schwanniomyces, Issatchenkia, Yarrowia* or *Rhodotorula*. In some embodiments, the nucleic acid is codon optimized for expression in *S. cerevisiae, C. lipolytica, R. glutinis, S. bulderi, S. barnetti, S. exiguus, S. uvarum, S. diastaticus, K. lactis, K. marxianus K. fragile, P. kudriavzevii, S. stipitis* or *I. orientalis*.

In some embodiments, the nucleic acid is codon optimized for expression in a filamentous fungal cell. In some embodiments, the nucleic acid is codon optimized for expression in *Aspergillus, Penicillium,*

Rhizopus, Chrysosporium, Myceliophthora, Trichoderma, Humicola, Acremonium or *Fusarium*. In some embodiments, the nucleic acid is codon optimized for expression in *A. niger, A. oryzae, T. reesei, P. chrysogenum, M. thermophila*, or *R. oryzae*.

In some embodiments, the nucleic acid is codon optimized for expression in *Botryococcus, Nannochloropsis, Chlorella, Chlamydomonas, Dunaliella, Chaetoceros, Porphyridium, Scenedesmus* or *Pseudochlorococcum*. In some embodiments, the nucleic acid is codon optimized for expression in *B. braunii* or *N. gaditana*.

In some embodiments, the nucleic acid has at least 70% sequence identity to SEQ ID NO: 3. In some embodiments, the nucleic acid sequence has at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99%, or 100% sequence identity to SEQ ID NO: 3, over a region of at least about 10, *e.g.*, at least about 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, or 1557 nucleotides, or a catalytic domain thereof.

SEQ ID NO: 3 represents the native *S. marinus* FAR nucleic acid sequence corresponding to the SmFAR polypeptide sequence (SEQ ID NO: 1):

ATGCGGGATCTGAATCGAGATAAGTCGTCGACGATAGCGGCCCTGACGGGTAAGCAGGTGCTGATTACCGGT
ACAACGGGGTTCTTGGGCAAGGTGTTGTTAGAGAAGGTTCTGCGCACGACGCCGGGCGTGGGCGGTATTTAC
TTGCTCATGCGGGGCAATCGCCATTTTCGTGATGCCACTGACCGCTTTCGGGAAGAAGTGATCTCCTCCTCCAT
TTTTGACACCTTGCGGCATGAACTGGGGCAGGAGGGCTTTGAGGCGCTGATTGCAAGTCGCGTTCACTGTGT
GACGGGGGAAATTACCGCACCGCGGTTTGGTCTGACGCGCACAGTCTTTACGGCGTTGGCCGACAAGCTCGA
TGTGGTGATCAATTCCGCAGCGAGTGTAAATTTTCGTGAGGAGTTGGATCGAGCGCTGGAAATCAATACGTTC
TCCCTGGAAAACGTCGCGAAACTCTGTGAGTTGGGGGGGGAATATACCACTTGTTCACGTGTCCACCTGCTACG
TAAACGGCTACCGGTCGGGTGATATCCACGAGACGTCCGAGGGCCCTGTTGGGCTGGAGTTACCTGAGGGGC
CGGGTAGTTTCTACGAAACCTCGGCCTTGATTGAGCGCTTGCAGCGCCTTGTTGAGGAGTTGCGCGGGCAATA
CGAGGGACGGGCGCTGAAGGCCAAGCTCATAGAGCATGGCGCGAAGGAAGCGCAAGCTGCGGGCTGGAAC
GATACCTATACCTTCACCAAGTGGCTGGGTGAGCAATATTTGTACAAGGTGATGCGAGGCTACTCGCTGACGG
TTGTTCGTCCGTCGATCATCGAGAGCACACTGCGCGAGCCAAGCCCAGGGTGGATTGAGGGAGTCAAGGTCG
CCGACGCGGTGTTGATGGCATATGCGCGCGGTAAGGTTGCGTTCTTTCCCGGTAAGCGGGCCGGGGTGATCG
ATATTATTCCGGCCGATCTCGTGGCCAATGGGGTATTGCTCTCGACAGCCGAGCAGTTGCTCGAACCGGGCCA
GCATCGCATCTATCAATGTTGTAGCGGTGGCCGTAACCCACTAATGCTGGGAGACTTTATTGACCACTTGGTG
GCAGAAGCCAGCAGTCATCACGGTGACTACGACAAGCTTTTTAAGGCGCCGCCCCGCCGGCGATTCACGGCG
GTGGATAAGCGTGTATTCAGTAGCGTCGCAATGTGCGTTCAGGTGGCGCTGACGATTGTTGGCGCGATATTG
CGTCGTCTCGGACGGCGCGGCAAGCTTAGAGCGCGGCGAAATATCGACACGGCTGTCGAGCTCTCCAAAATA
TTCTCTTTCTATGCTCAGCCCAATTACTGCTTTCACGGCGACAAGCTGCTTTCATTATCGAAGTCGATGGGCGCT
GTCGATCAGGAGCTGTTCCCTGTGGATGCCGGCGTGATTGATTGGCAGCACTATATTCGTCACGTTCATATGC
CGGGATTAAACCACTATGCGCTGCAGGGGCAGCCCTCAGTAAGGTCACAGCAGCGGCGCGAGGAAGTGAAA
ACGGCGCCGAAAAACCTGACGGCGACCGAGGCTTGA **(SEQ ID NO: 3)**

Vectors

A vector of the invention comprises a FAR nucleic acid of the invention. A vector may comprise one or more of an origin of replication, a promoter sequence operably linked to a nucleic acid of the invention and a reporter gene or selectable marker.

Vectors according to the invention include, but are not limited to, pCL-BP2, pBR322, pBR327, pACYC184, pACYC177, pSC101, pCL1920, and pCL1921.

Vectors of the invention comprising an origin of replication include, but are not limited to, ColE1, pMB1, p15A, pSC101, R6K, RK2 and pRSF1010.

The promoter may be homologous or heterologous. The promoter may be constitutive or inducible.

In one embodiment, the promoter is inducible and is activated in the presence of an inducing agent. Inducing agents include, but are not limited to, sugars, metal salts, and antibiotics. In some cases, the promoter allows constitutive expression of the FAR polypeptide.

In one embodiment, promoters that are active at different stages of growth can be used.

The promoter sequence may be operable in a prokaryotic cell, for example an *E. coli* cell. In one embodiment the promoter is a Trc promoter.

The promoter sequence may be operable in a eukaryotic cell, for example a yeast cell, a fungal cell, an algal cell or a plant cell.

Where the recombinant cell is a fungal cell, the promoter can be a fungal promoter (including, but not limited to, a filamentous fungal promoter), a promoter operable in plant cells, or a promoter operable in mammalian cells. Mammalian, mammalian viral, plant and plant viral promoters can drive particularly high expression when the associated 5' UTR sequence (i.e., the sequence which begins at the transcription start site and ends one nucleotide before the start codon), normally associated with the mammalian or mammalian viral promoter is replaced by a fungal 5' UTR sequence. The source of the 5' UTR can vary provided it is operable in the filamentous fungal cell. In various embodiments, the 5' UTR can be derived from a yeast gene or a filamentous fungal gene. The 5' UTR can be from the same species as the recombinant cell or from a different species.

Promoters for recombinant expression in yeast are known in the art. Suitable promoters for *S. cerevisiae* include, but are not limited to, the MFα1 promoter, galactose inducible promoters such as GAL1, GAL7, and GAL10 promoters, glycolytic enzyme promoters including the TPI and PGK promoters, the TDH3 promoter, the TEF1 promoter, the TRP1 promoter, the CYC1 promoter, the CUP1 promoter, the PHO5 promoter, the ADH1 promoter, and the HDP promoter. A suitable promoter in the genus *Pichia* sp. is the AOX1 promoter.

Suitable reporter genes or selectable markers include, but are not limited to, a drug resistance gene, a metabolic enzyme, a factor required for survival of the recombinant cell, a fluorescent marker, or an enzyme that generates a detectable product. Cells transformed with the vector can be selected based on their ability to grow in the presence of inhibitors (*e.g.* antibiotics) or under conditions in which untransformed cells cannot grow.

The vector may be a high copy number vector, an intermediate copy number vector, or a low copy number vector. Preferably, the vector is a low copy number vector, *e.g.* vector pCL-BP2.

In some embodiments, the vector comprises a nucleic acid sequence having at least 70% sequence identity to SEQ ID NO: 4. In some embodiments, the vector comprises a nucleic acid sequence having at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to SEQ ID NO: 4. Vector corresponding to SEQ ID NO: 4 is also referred to herein as pCl-BP2.

CGCGCGCGAAGGCGAAGCGGCATGCATTTACGTTGACACCATCGAATGGTGCAAAACCTT
TCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATGTGAAACC
AGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTG
GTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGC
GGAGCTGAATTACATTCCCAACCGCGTGGCACAACAACTGGCGGGCAAACAGTCGTTGCT
GATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCCGTCGCAAATTGTCGCGGCGATT
AAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTCGATGGTAGAACGAAGCGGC
GTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATC
ATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTC
CGGCGTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAA
GACGGTACGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTG
TTAGCGGGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATC
TCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCG
GTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGC
CAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGG
TGCGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCC
GTCAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCT
GCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAA
AAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTC
ATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAA
TTAATGTGAGTTAGCGCGAATTGATCTGGTTTGACAGCTTATCATCGACTGCACGGTGCAC
CAATGCTTCTGGCGTCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAAAT
CACTGCATAATTCGTGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTGCGCCGA
CATCATAACGGTTCTGGCAAATATTCTGAAATGAGCTGTTGACAATTAATCATCCGGCTCG
TATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCGCCGCTGAGAA
AAAGCGAAGCGGCACTGCTCTTTAACAATTTATCAGACAATCTGTGTGGGCACTCGACCG
GAATTATCGATTAACTTTATTATTAAAAATTAAAGAGGTATATATTAATGTATCGATTAAA
TAAGGAGGAATAAACCATGGATCCGAGCTCGAGATCTGCAGCTGGTACCATATGGGAATT
CGAAGCTTACGTAGAACAAAAACTCATCTCAGAAGAGGATCTGAATAGCGCCGTCGACCA
TCATCATCATCATCATTGAGTTTAAACCAAATAAAACGAAAGGCTCAGTCGAAAGACTGG
GCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGG
GAGCGGATTTGAACGTTGCGAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGCCA
TAAACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTC
TACAAACTCTTTATGTGCTTAGTGCATCTAACGCTTGAGTTAAGCCGCGCCGCGAAGCGGC

GTCGGCTTGAACGAATTGTTAGACATTATTTGCCGACTACCTTGGTGATCTCGCCTTTCAC
GTAGTGGACAAATTCTTCCAACTGATCTGCGCGCGAGGCCAAGCGATCTTCTTCTTGTCCA
AGATAAGCCTGTCTAGCTTCAAGTATGACGGGCTGATACTGGGCCGGCAGGCGCTCCATT
GCCCAGTCGGCAGCGACATCCTTCGGCGCGATTTTGCCGGTTACTGCGCTGTACCAAATGC
GGGACAACGTAAGCACTACATTTCGCTCATCGCCAGCCCAGTCGGGCGGCGAGTTCCATA
GCGTTAAGGTTTCATTTAGCGCCTCAAATAGATCCTGTTCAGGAACCGGATCAAAGAGTTC
CTCCGCCGCTGGACCTACCAAGGCAACGCTATGTTCTCTTGCTTTTGTCAGCAAGATAGCC
AGATCAATGTCGATCGTGGCTGGCTCGAAGATACCTGCAAGAATGTCATTGCGCTGCCATT
CTCCAAATTGCAGTTCGCGCTTAGCTGGATAACGCCACGGAATGATGTCGTCGTGCACAAC
AATGGTGACTTCTACAGCGCGGAGAATCTCGCTCTCTCCAGGGGAAGCCGAAGTTTCCAA
AAGGTCGTTGATCAAAGCTCGCCGCGTTGTTTCATCAAGCCTTACGGTCACCGTAACCAGC
AAATCAATATCACTGTGTGGCTTCAGGCCGCCATCCACTGCGGAGCCGTACAAATGTACG
GCCAGCAACGTCGGTTCGAGATGGCGCTCGATGACGCCAACTACCTCTGATAGTTGAGTC
GATACTTCGGCGATCACCGCTTCCCTCATGATGTTTAACTTTGTTTTAGGGCGACTGCCCTG
CTGCGTAACATCGTTGCTGCTCCATAACATCAAACATCGACCCACGGCGTAACGCGCTTGC
TGCTTGGATGCCCGAGGCATAGACTGTACCCCAAAAAAACAGTCATAACAAGCCATGAAA
ACCGCCACTGCGCCGTTACCACCGCTGCGTTCGGTCAAGGTTCTGGACCAGTTGCGTGAGC
GCATACGCTACTTGCATTACAGCTTACGAACCGAACAGGCTTATGTCCACTGGGTTCGTGC
CTTCATCCGTTTCCACGGTGTGCGTCACCCGGCAACCTTGGGCAGCAGCGAAGTCGAGGC
ATTTCTGTCCTGGCTGGCGAACGAGCGCAAGGTTTCGGTCTCCACGCATCGTCAGGCATTG
GCGGCCTTGCTGTTCTTCTACGGCAAGGTGCTGTGCACGGATCTGCCCTGGCTTCAGGAGA
TCGGAAGACCTCGGCCGTCGCGGCGCTTGCCGGTGGTGCTGACCCCGGATGAAGTGGTTC
GCATCCTCGGTTTTCTGGAAGGCGAGCATCGTTTGTTCGCCCAGCTTCTGTATGGAACGGG
CATGCGGATCAGTGAGGGTTTGCAACTGCGGGTCAAGGATCTGGATTTCGATCACGGCAC
GATCATCGTGCGGGAGGGCAAGGGCTCCAAGGATCGGGCCTTGATGTTACCCGAGAGCTT
GGCACCCAGCCTGCGCGAGCAGGGGAATTAATTCCCACGGGTTTTGCTGCCCGCAAACGG
GCTGTTCTGGTGTTGCTAGTTTGTTATCAGAATCGCAGATCCGGCTTCAGCCGGTTTGCCG
GCTGAAAGCGCTATTTCTTCCAGAATTGCCATGATTTTTTCCCCACGGGAGGCGTCACTGG
CTCCCGTGTTGTCGGCAGCTTTGATTCGATAAGCAGCATCGCCTGTTTCAGGCTGTCTATGT
GTGACTGTTGAGCTGTAACAAGTTGTCTCAGGTGTTCAATTTCATGTTCTAGTTGCTTTGTT
TTACTGGTTTCACCTGTTCTATTAGGTGTTACATGCTGTTCATCTGTTACATTGTCGATCTG
TTCATGGTGAACAGCTTTGAATGCACCAAAAACTCGTAAAAGCTCTGATGTATCTATCTTT
TTTACACCGTTTTCATCTGTGCATATGGACAGTTTTCCCTTTGATATGTAACGGTGAACAGT
TGTTCTACTTTTGTTTGTTAGTCTTGATGCTTCACTGATAGATACAAGAGCCATAAGAACCT
CAGATCCTTCCGTATTTAGCCAGTATGTTCTCTAGTGTGGTTCGTTGTTTTTGCGTGAGCCA
TGAGAACGAACCATTGAGATCATACTTACTTTGCATGTCACTCAAAAATTTTGCCTCAAAA
CTGGTGAGCTGAATTTTTGCAGTTAAAGCATCGTGTAGTGTTTTTCTTAGTCCGTTATGTAG
GTAGGAATCTGATGTAATGGTTGTTGGTATTTTGTCACCATTCATTTTTATCTGGTTGTTCT
CAAGTTCGGTTACGAGATCCATTTGTCTATCTAGTTCAACTTGGAAAATCAACGTATCAGT
CGGGCGGCCTCGCTTATCAACCACCAATTTCATATTGCTGTAAGTGTTTAAATCTTTACTTA
TTGGTTTCAAAACCCATTGGTTAAGCCTTTTAAACTCATGGTAGTTATTTTCAAGCATTAAC
ATGAACTTAAATTCATCAAGGCTAATCTCTATATTTGCCTTGTGAGTTTTCTTTTGTGTTAG
TTCTTTTAATAACCACTCATAAATCCTCATAGAGTATTTGTTTTCAAAAGACTTAACATGTT
CCAGATTATATTTTATGAATTTTTTTAACTGGAAAAGATAAGGCAATATCTCTTCACTAAA
AACTAATTCTAATTTTTCGCTTGAGAACTTGGCATAGTTTGTCCACTGGAAAATCTCAAAG
CCTTTAACCAAAGGATTCCTGATTTCCACAGTTCTCGTCATCAGCTCTCTGGTTGCTTTAGC
TAATACACCATAAGCATTTTCCCTACTGATGTTCATCATCTGAGCGTATTGGTTATAAGTG
AACGATACCGTCCGTTCTTTCCTTGTAGGGTTTTCAATCGTGGGGTTGAGTAGTGCCACAC
AGCATAAAATTAGCTTGGTTTCATGCTCCGTTAAGTCATAGCGACTAATCGCTAGTTCATT
TGCTTTGAAAACAACTAATTCAGACATACATCTCAATTGGTCTAGGTGATTTTAATCACTA
TACCAATTGAGATGGGCTAGTCAATGATAATTACTAGTCCTTTTCCTTTGAGTTGTGGGTA
TCTGTAAATTCTGCTAGACCTTTGCTGGAAAACTTGTAAATTCTGCTAGACCCTCTGTAAA
TTCCGCTAGACCTTTGTGTGTTTTTTTTGTTTATATTCAAGTGGTTATAATTTATAGAATAA
AGAAAGAATAAAAAAAGATAAAAAGAATAGATCCCAGCCCTGTGTATAACTCACTACTTT

AGTCAGTTCCGCAGTATTACAAAAGGATGTCGCAAACGCTGTTTGCTCCTCTACAAAACAG
ACCTTAAAACCCTAAAGGCTTAAGTAGCACCCTCGCAAGCTCGGGCAAATCGCTGAATAT
TCCTTTTGTCTCCGACCATCAGGCACCTGAGTCGCTGTCTTTTTCGTGACATTCAGTTCGCT
GCGCTCACGGCTCTGGCAGTGAATGGGGGTAAATGGCACTACAGGCGCCTTTTATGGATT
CATGCAAGGAAACTACCCATAATACAAGAAAAGCCCGTCACGGGCTTCTCAGGGCGTTTT
ATGGCGGGTCTGCTATGTGGTGCTATCTGACTTTTTGCTGTTCAGCAGTTCCTGCCCTCTGA
TTTTCCAGTCTGACCACTTCGGATTATCCCGTGACAGGTCATTCAGACTGGCTAATGCACC
CAGTAAGGCAGCGGTATCATCAACAGGCTTACCCGTCTTACTGT **(SEQ ID NO: 4)**

Recombinant cells

The invention provides recombinant cells engineered to express a polypeptide of the invention. The invention also provides recombinant cells transformed with a nucleic acid of the invention. The nucleic acid may be extrachromosomal, on a vector (typically a plasmid), which can be a low copy number vector, an intermediate copy number vector, or a high copy number vector. In some embodiments, the recombinant cell is transformed with a vector of the invention. In some embodiments, the recombinant cell is transiently transformed. The nucleic acid may be maintained episomally and thus comprise a sequence for autonomous replication, such as an autosomal replication sequence. Alternatively, the recombinant cell may be stably transformed wherein the nucleic acid is integrated in one or more copies into the genome of the cell. Integration into the cell's genome may occur at random by non-homologous recombination but preferably, the nucleic acid construct may be integrated into the cell's genome by homologous recombination, as is well known in the art.

In some embodiments, the recombinant cell is a prokaryotic cell. In some embodiments, the prokaryotic cell is selected from *Acinetobacter*, *Agrobacterium*, *Escherichia*, *Cupriavidus*, *Clostridium*, *Rhodobacter*, *Marinobacter*, *Bacillus*, *Klebsiella*, *Tatumella*, *Pseudomonas*, *Ralstonia*, *Rhodococcus*, *Methylobacterium*, *Methylophilus*, *Methylococcus*, *Methylomicrobium*, *Methylomonas*, *Pantoea*, *Streptomyces*, *Parachlorella*, *Synechococcus*, *Synechocystis* and *Thermocynechococcus*. In some embodiments, the prokaryotic cell is *E. coli*. Suitable cells of the bacterial genera include, but are not limited to, cells of *Lactobacillus*, *Pseudomonas*, and *Streptomyces*. Suitable cells of bacterial species include, but are not limited to, cells of *Bacillus subtilis*, *Bacillus licheniformis*, *Lactobacillus brevis*, *Pseudomonas aeruginosa*, and *Streptomyces lividans*. Typically, the recombinant cell is not *Spongibacter marinus*.

In some embodiments, the recombinant cell is a eukaryotic cell. In some embodiments, the eukaryotic cell is selected from the group consisting of a yeast cell, a fungal cell, an algal cell and a plant cell. In some embodiments, the yeast cell is selected from the group consisting of *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Schefferomyces*, *Rhodosporidium*, *Hansenula*, *Klockera*, *Schwanniomyces*, *Issatchenkia*, *Yarrowia* and *Rhodotorula*. In some embodiments, the yeast cell is selected from the group consisting of *S. cerevisiae*, *C. lipolytica*, *R.*

glutinis, *S. bulderi*, *S. barnetti*, *S. exiguus*, *S. uvarum*, *S. diastaticus*, *K. lactis*, *K. marxianus*, *K. fragile*, *P. kudriavzevii*, *S. stipitis* and *I. orientalis*. Suitable cells of yeast include, but are not limited to *C. albicans*, S. *pombe*, *H. polymorpha*, *P. pastoris*, *P. canadensis*, or *P. rhodozyma*.

In some embodiments, the fungal cell is a filamentous fungal cell. In some embodiments, the filamentous fungal cell is selected from the group consisting of *Aspergillus*, *Penicillium*, *Rhizopus*, *Chrysosporium*, *Myceliophthora*, *Trichoderma*, *Humicola*, *Acremonium* and *Fusarium*. In some embodiments, the filamentous fungal cell is selected from the group consisting of *A. niger*, *A. oryzae*, *T. reesei*, *P. chrysogenum*, *M. thermophila*, and *R. oryzae*. Suitable cells of filamentous fungal genera include, but are not limited to, cells of *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Coprinus*, *Coriolus*, *Corynascus*, *Chaetomium*, *Cryptococcus*, *Filobasidium*, *Gibberella*, *Hypocrea*, *Magnaporthe*, *Mucor*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Scytaldium*, *Schizophyllum*, *Sporotrichum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *tolypocladium* and *Trametes*. In certain aspects, the recombinant cell is a *Trichoderma sp.*, *Penicillium sp.*, *Humicola sp. (e.g., Humicola insolens)*; *Aspergillus sp.*, *Chrysosporium sp.*, *Fusarium sp.*, *or Hypocrea sp.*. Suitable cells can also include cells of various anamorph and teleomorph forms of these filamentous fungal genera.

Suitable cells of filamentous fungal species include, but are not limited to, cells of *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Chrysosporium lucknowense*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Bjerkandera adusta*, *Ceriporiopsis aneirina*, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis gilvescens*, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*, *Ceriporiopsis subrufa*, *Ceriporiopsis subvermispora*, *Coprinus cinereus*, *Coriolus hirsutus*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Neurospora intermedia*, *Penicillium purpurogenum*, *Penicillium canescens*, *Penicillium solitum*, *Penicillium funiculosum*, *Phanerochaete chrysosporium*, *Phlebia radiate*, *Pleurotus eryngii*, *Talaromyces flavus*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, and *Trichoderma viride*.

In some embodiments, the algal cell is selected from the group consisting of *Botryococcus*, *Nannochloropsis*, *Chlorella*, *Chlamydomonas*, *Dunaliella*, *Chaetoceros*, *Porphyridium*, *Scenedesmus* and *Pseudochlorococcum*. In some embodiments, the algal cell is selected from the group consisting of *B. braunii* and *N. gaditana*.

Recombinant cells may be cultured in conventional nutrient media modified as appropriate for activating promoters (if an inducible promoter is present), selecting transformants, or amplifying the nucleic acid sequence encoding the FAR polypeptide. Culture conditions, such as temperature, pH and the like, are those previously used with the recombinant cell selected for expression, and will be apparent to those skilled in the art. Many references are available for the culture and production of many cells, including cells of bacterial and fungal origin. Cell culture media in general are set forth in Atlas and Parks (eds.), 1993, The Handbook of Microbiological Media, CRC Press, Boca Raton, FL, which is incorporated herein by reference. Preferred culture conditions for a given recombinant cell may be found in the scientific literature and/or from the source of the recombinant cell such as the American Type Culture Collection (ATCC).

In cases where a FAR coding sequence is under the control of an inducible promoter, the inducing agent, *e.g.*, a sugar, metal salt or antibiotic, may be added to the medium at a concentration effective to induce expression of FAR polypeptide.

A recombinant cell of the invention may comprise one or more genetic modifications to increase the yield and/or recovery of fatty alcohol. Such genetic modifications may result in at least one, any two, any three, any four or all five of the following phenotypes: (a) an increase in fatty acyl-ACP production; (b) a decrease in fatty acyl-ACP degradation; (c) an upregulation of fatty alcohol secretion; (d) an increase in flux through the fatty acid biosynthetic pathway; and (e) an increase in tolerance to fatty alcohol.

(a) Increase in fatty acyl-ACP production

The yield of fatty alcohols may be improved by increasing the production of the SmFAR substrate, fatty acyl-ACP. Fatty acyl-ACP production may be increased by various modifications of the fatty acid biosynthetic (FAB) pathway in which fatty acyl-ACP is an intermediate. This may include overexpression of endogenous genes and/or expression of exogenous genes from different organisms. Methods for increasing fatty acyl-ACP production are known in the art (see for example, US 8,110,670, incorporated herein by reference). For example, increased fatty acyl-ACP production may be achieved in bacteria by overexpression of malonyl-CoA-ACP transacylase (fabD), β-ketoacyl-ACP synthase III (fabH), β-ketoacyl-ACP synthases (fabB and fabF), β-ketoacyl-ACP reductase (fabG), 3-hydrox-acyl-ACP dehydratase (fabA) and/or enoyl-ACP reductase (fabI).

(b) Decrease in acyl-ACP degradation

The yield of fatty alcohols may be improved by decreasing the degradation of the SmFAR substrate, fatty acyl-ACP. Methods to increase fatty alcohol yield by reducing acyl-ACP degradation are known in the art (see for example, Liu, Chen *et al.* 2016, incorporated herein by reference). In the FAB pathway, fatty acyl-ACP is converted to fatty acids by thioesterases. Reduced degradation of fatty

acyl-ACP may therefore be achieved in bacteria by, for example, downregulation or deletion of thioesterases.

(c) Upregulation of fatty alcohol secretion

Increasing the recovery of fatty alcohols may be achieved by increasing the rate of fatty alcohol secretion from the recombinant cell. Methods for increasing the secretion of fatty alcohols by microorganisms are known in the art, *e.g.* by overexpression of efflux pumps (see for example, Zhang *et al.* Biochem. Eng. J. **133**: 149-156 (2018), incorporated herein by reference). For example, prokaryotic multidrug resistance (MDR) transporters can export a wide variety of substrates, including fatty alcohols, and so recombinant upregulation of these transporters may lead to increased fatty alcohol secretion. Examples of MDR transporters in *E. coli* include MdtJ, Bcr and MdtH.

(d) Increase in flux through the fatty acid biosynthetic pathway

The main precursors for the fatty acid biosynthetic pathway are malonyl-CoA and acetyl-CoA. Increasing the availability of these precursors may increase the rate of fatty acyl-ACP generation and therefore increase flux through the FAB pathway. Methods of increasing flux through the FAB pathway are known in the art (see for example, WO2013/152051, incorporated herein by reference). In bacteria, overexpression of acetyl-CoA carboxylase (accA, accB, accC, and accD), which catalyzes the conversion of acetyl-CoA to malonyl-CoA, may increase the yield of fatty acyl-ACPs.

(e) Increase in tolerance to fatty alcohol

High levels of intracellular fatty alcohols could potentially lead to toxic effects in recombinant cells. If required, increased tolerance to fatty alcohols may be achieved by upregulating secretion of fatty alcohols from the recombinant cell (see point (c) above).

The foregoing phenotypes may be introduced singly or in combinations of two, three, four, five or more modifications. One or more of the foregoing phenotypes may be obtained by increasing or decreasing expression of an endogenous protein (*e.g.*, by at least a factor of about 1.1, about 1.2, about 1.5, about 2, about 5, about 10 or about 20) or a result of introducing expression of a heterologous polypeptide. For avoidance of doubt, "decreasing" or "reducing" gene expression encompasses eliminating expression. Decreasing (or reducing) the expression of an endogenous protein may be accomplished by inactivating one or more (or all) endogenous copies of a gene in a cell. A gene may be inactivated by deletion of at least part of the gene or by disruption of the gene *e.g.* by deleting some or all of a gene coding sequence or regulatory sequence whose deletion results in a reduction of gene expression in the cell.

Fatty alcohol production

The invention provides a method of producing a fatty alcohol composition comprising: a) culturing a recombinant cell in a suitable culture medium, wherein the recombinant cell comprises a gene encoding a heterologous FAR enzyme having at least 70% sequence identity to SEQ ID NO: 1; and b) allowing expression of said gene, wherein said expression results in the production of a fatty alcohol composition.

Recombinant cells of the invention may be cultured under suitable conditions in a liquid or solid medium. In some embodiments, recombinant cells of the invention undergo fermentation. Fermentation conditions include batch, fed-batch and continuous fermentation. Classical batch fermentation is a closed system, wherein the composition of the medium is not subject to artificial alterations during fermentation. In fed-batch fermentation, the substrate is added in increments as fermentation progresses. In both classical batch fermentation and batch-fed fermentation, the product(s) remain in the bioreactor until the end of the process. Batch and fed-batch fermentation are common and well-known in the art. In continuous fermentation, a defined medium is added continuously to the bioreactor and an equal volume of product containing medium is removed simultaneously. Continuous fermentation aims to maintain steady state growth conditions. Methods for modulating nutrients and growth factors for continuous fermentation processes as well as techniques for maximizing the rate of product formation are well known in the art of industrial microbiology. The fermentation process may be an aerobic or an anaerobic fermentation process.

The fermentation process is typically run at a temperature that is optimal for growth of the recombinant FAR-expressing cell. Fermentation is typically carried out at a temperature within the range of from about 10°C to about 70°C, from about 15°C to about 60°C, from about 20°C to about 50°C, or from about 25°C to about 40°C. Fermentation is typically carried out for a period of time within the range of from about 8 to 240 hours, from about 12 hours to about 168 hours, from about 16 hours to about 144 hours, from about 20 hours to about 120 hours, or from about 24 hours to about 72 hours. Fermentation is typically carried out at a pH in the range of 4 to 8, in the range of 5 to 7, or the range of 5.5 to 6.5.

The method of the invention may comprise purifying fatty alcohols from cell culture to provide a purified fatty alcohol composition.

The method of the invention may comprise purifying fatty alcohols from culture supernatant to provide a purified fatty alcohol composition.

The method of the invention may comprise harvesting and lysing recombinant cells to release intracellular fatty alcohols from the recombinant cell. The method of the invention may further

comprise purifying the fatty alcohol composition from the lysate to provide a purified fatty alcohol composition.

The invention also provides a cell-free method for producing a fatty alcohol composition comprising incubating a FAR polypeptide of the invention with acyl-ACP and a reductant. In some embodiments, the reductant is NADPH. Typically a cell-free method of the invention involves incubating the FAR polypeptide of the invention under suitable conditions of temperature, pH, and ionic strength with a suitable substrate, *e.g.* acyl-ACP, and a suitable reductant, *e.g.* NAD(P)H. A cell-free method of the invention may comprise use of an immobilized FAR polypeptide of the invention. In one embodiment, an immobilized FAR polypeptide of the invention is immobilized on an inert, insoluble substrate. An immobilized FAR polypeptide may have greater stability compared with a soluble form and may be more readily recovered from reaction mixtures. A cell-free method of the invention may further comprise purifying fatty alcohols from the cell-free reaction mixture to provide a purified fatty alcohol composition.

Wherein the fatty alcohol composition was produced in a recombinant cell of the invention, purifying the fatty alcohol composition typically comprises separating fatty alcohols from components of the culture medium, fermentation process, and/or cellular material. Wherein the fatty alcohol composition was produced in a cell-free method of the invention, purifying the fatty alcohol composition typically comprises separating fatty alcohols from the cell-free reaction mixture.

Fatty alcohol compositions may be purified using conventional methods known in the art. For example, fatty alcohols may be purified by solvent extraction with a water immiscible solvent. In some embodiments, purification of fatty alcohols occurs contemporaneously with biosynthesis of fatty alcohols.

Fatty alcohol composition

Fatty alcohol compositions may be characterized using various analytical methods known in the art. Fatty alcohol characteristics described herein, such as % fatty alcohol distribution and % saturation, are typically determined using gas chromatography–mass spectrometry (GCMS). An exemplary GCMS method is provided herein in Example 3.

Fatty alcohol compositions produced by a method of the invention typically comprise C12 fatty alcohol. Fatty alcohol compositions produced by a method of the invention typically comprise saturated C12 fatty alcohol (i.e. "C12:0" fatty alcohol). In one embodiment, the fatty alcohol composition comprises at least 1%, at least 2%, at least 3%, at least 4%, at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, or at least 90% of C12 fatty alcohol.

In one embodiment, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, or at least 90% of the fatty alcohols in the fatty alcohol composition are saturated fatty alcohols. In one embodiment, at least 45% of the fatty alcohols in the fatty alcohol composition are saturated fatty alcohols.

FARs of the invention provide fatty alcohol compositions with highly desirable characteristics (discussed herein). In some embodiments, modification(s) are introduced to the recombinant cell that alter the characteristics of fatty alcohol compositions produced by FARs of the invention. For example, the recombinant cell may be modified to change the proportion of saturation of the fatty alcohol and/or modified to change the chain-length of fatty acyl-ACPs produced by the FAB pathway.

The FAB pathway produces a range of saturated and unsaturated fatty acyl-ACPs which can be used as substrates by a FAR enzyme. Modifying the proportion of saturated and unsaturated fatty acyl-ACPs produced by the recombinant cell may allow the degree of saturation in the fatty alcohol composition to be altered. Methods to increase the proportion of saturated fatty acyl-ACPs produced by the FAB pathway are known in the art (see for example Heath and Rock, JBC, **271**:27795-27801 (1996), incorporated herein by reference). As disclosed in WO 2013/096092 and US 9260727, in *E. coli*, overexpression of β-hydroxyacyl-ACP dehydratase enzymes FabA or FabZ, and/or enoyl-ACP reductase enzyme FabI, increased the level of intermediates in the saturated branch of the FAB pathway and thereby increase the degree of saturated fatty acyl-ACPs produced.

The FAB pathway produces fatty acyl-ACPs of varying chain lengths, with C14-C18 fatty acyl-ACPs typically being the most abundant. The FAB pathway may be engineered to increase the proportion of shorter chain fatty acyl-ACPs produced thereby increasing the production of shorter chain fatty alcohols. Methods to alter the chain-length of fatty acyl-ACPs produced by the FAB pathway are known in the art (see for example, Torella *et al.* PNAS, **110**(28):11290-11295 (2013), incorporated herein by reference). In *E. coli*, an increase in shorter chain fatty acyl-ACP production may be achieved by altering the chain length specificity of FAB enzymes involved in chain length elongation (*e.g.* FabG, FabB or FabF).

In one embodiment, a method of the invention further comprises reducing the fatty alcohol composition to produce alkanes. Methods of reducing fatty alcohols are known in the art. In one embodiment, reduction of fatty alcohols is chemically-mediated. In one embodiment, reduction of fatty alcohols is biologically-mediated *e.g.* by a microorganism. In one embodiment, alkanes of the invention are purified from the fatty alcohol composition from which they are derived.

In one embodiment, a method of the invention further comprises esterifying the fatty alcohol composition to produce esters. Methods of esterifying fatty alcohols are known in the art. In one embodiment, the fatty alcohols are reacted with a carboxylic acid to form esters. In one embodiment, esterification is carried out in the presence of a strong catalyst, *e.g.* sodium hydroxide. In one

embodiment, esterification is catalyzed by an enzyme that catalyzes the conversion of fatty alcohols to esters. In one embodiment, esters are purified from the fatty alcohol composition from which they are derived. In one embodiment, purified esters of the invention are used as biodiesel fuel.

In one embodiment, the method of the invention further comprises processing fatty alcohol compositions, alkane compositions and/or ester compositions to provide a fuel composition.

In one embodiment, the method of the invention further comprises processing fatty alcohol compositions, alkane compositions and/or ester compositions to provide a cleaning composition (*e.g.* detergent).

In one embodiment, the method of the invention further comprises processing fatty alcohol compositions, alkane compositions and/or ester compositions to provide an industrial composition (*e.g.* lubricant).

In one embodiment, the method of the invention further comprises processing fatty alcohol compositions, alkane compositions and/or ester compositions to provide a cosmetic composition *e.g.* shampoo or soap.

The invention also provides fuel compositions, cleaning compositions (*e.g.* detergents), industrial compositions (*e.g.* lubricants), and cosmetic compositions (*e.g.* shampoos, soaps *etc.*) comprising or produced from fatty alcohol compositions, alkane compositions and/or ester compositions produced by the methods disclosed herein.

Publications cited herein and the material for which they are cited are hereby specifically incorporated by reference in their entireties.

The following examples are provided for the purpose of illustrating, not limiting, the material disclosed herein.

EXAMPLE 1

In the hope of identifying novel FAR enzymes, phylogenetic trees were constructed based on BLAST searches utilizing the known active FARs from *M. aquaeolei, M. algicola* and *Oceanobacter* sp. RED65 protein sequences. Based on these trees, a set of 24 protein sequences, distantly related to these 3 FARs were selected (see Table 1). FAR homologs were synthesized with codon optimized sequences for recombinant expression in *E. coli* by Genewiz (South Plainfield, NJ). The synthesized genes were cloned into vector pTrcHis2A (Invitrogen Inc. Carlsbad, CA) by Genewiz with the addition of restriction sites NcoI and HindIII.

Table 1. Identity of FARs selected for analysis.

FAR	Name	Species Name	GenBank Protein ID
1	AsFAR	*Aestuariibacter salexigens*	WP_051275700.1
2	Ole1FAR	*Oleiphilus*	WP_068554578.1
3	Mas1FAR	*Massilia sp. Root335*	WP_056451153.1
4	BbFAR	*Burkholderiales bacterium*	OGB29195.1
5	MmFAR	*Marinobacter mobilis*	SDX22907.1
6	HcFAR	*Hahella chejuensis*	WP_011398823.1
7	MhFAR	*Marinobacter hydrocarbonoclasticus*	WP_011785966.1
8	Mar1FAR	*Marinobacter*	WP_014870673.1
9	HgFAR	*Hahella ganghwensis*	WP_020407662.1
10	MdFAR	*Marinobacter daepoensis*	WP_029655611.1
11	AbFAR	*Acidobacteria bacterium*	OFV84852.1
12	RrFAR	*Rubrobacter radiotolerans*	WP_038682271.1
13	Mas2FAR	*Massilia*	WP_047825758.1
14	DugFAR	*Duganella*	WP_070357154.1
15	HerFAR	*Herbaspirillum*	WP_052452145.1
16	SmFAR	*Spongiibacter marinus*	WP_027874719.1
17	ZaFAR	*Zhongshania aliphaticivorans*	WP_008249708.1
18	Mar2FAR	*Marinobacter*	WP_053114882.1
19	Ole2FAR	*Oleiphilus*	WP_068482091.1
20	Ole3FAR	*Oleiphilus*	WP_068760013.1
21	Mar3FAR	*Marinobacter*	WP_070055810.1
22	MlFAR	*Marinobacter lipolyticus*	WP_012139620.1
23	NpFAR	*Nanorana parkeri*	XP_018415248.1
24	B1FAR	*Brachybacterium*	WP_062947317.1

EXAMPLE 2

Cloning of FAR into expression vector

To study FAR expression *in vivo*, the 24 FAR synthetic genes were cloned into the low-copy expression vector pCL-BP2 (see Figure 2). This vector was assembled from 4 synthetic DNA fragments (IDT, Caralville, IA), and is based on vector pCL1920 (Lerner and Inouye 1990), with some modifications, including the Trc promoter and optimized translation and termination sequences. The pCL-BP2 DNA sequence is shown in SEQ ID NO: 4. FAR genes were amplified by PCR using primers shown in Table 2. PCR products were cloned into vector pCL-BP2 using the In-Fusion cloning kit (Takara Bio) per manufacturer's protocol using linearized pCL-BP2 by restriction enzymes NcoI and HindIII. All DNA constructs were transformed into TOP10 *E. coli* strain and sequence verified.

Table 2. PCR primer sequences used to amplify FAR for cloning into pCL-BP2

FAR	Forward primer	Reverse primer
AsFAR	SEQ ID NO: 4 GAGGAATAAACCATGGAAAGCAGTACCACCTT	SEQ ID NO: 5 TGTTCTACGTAAGCTTCATGCGGCCTGATTCACGC
Ole1FAR	SEQ ID NO: 6 GAGGAATAAACCATGACCCAGGAAAGCAGCAT	SEQ ID NO: 7 TGTTCTACGTAAGCTTCAGTCGGTCACACTCAGCT
Mas1FAR	SEQ ID NO: 8 GAGGAATAAACCATGACCACCAGTATCAACGA	SEQ ID NO: 9 TGTTCTACGTAAGCTTCAGCGCACTGCCTGCGGTG
BbFAR	SEQ ID NO: 10 GAGGAATAAACCATGAATACCGCCCCTGATCT	SEQ ID NO: 11 TGTTCTACGTAAGCTTCAACGAACATTAACGGTGG
MmFAR	SEQ ID NO: 12 GAGGAATAAACCATGACCAAACACCAGATCGA	SEQ ID NO: 13 TGTTCTACGTAAGCTTCAGGCGGCTTTACGGGCGC
MhFAR	SEQ ID NO: 14 GAGGAATAAACCATGAACTATTTTTTTAACCGG	SEQ ID NO: 15 TGTTCTACGTAAGCTTCACCAGTAAATGCCGCGCA
Mar1FAR	SEQ ID NO: 16 GAGGAATAAACCATGAATTATTTCGTTACAGG	SEQ ID NO: 17 TGTTCTACGTAAGCTTCACCAATGAATGCCACGCA
HgFAR	SEQ ID NO: 18 GAGGAATAAACCATGAACTATTTCGTGACAGG	SEQ ID NO: 19 TGTTCTACGTAAGCTTCACCAGTGGATGCCACGCA
MdFAR	SEQ ID NO: 20 GAGGAATAAACCATGAATTACTTTTTAACCGG	SEQ ID NO: 21 TGTTCTACGTAAGCTTCACCAATAAATACCACGCA
AbFAR	SEQ ID NO: 22 GAGGAATAAACCATGAATCACTTAGATCCGTG	SEQ ID NO: 23 TGTTCTACGTAAGCTTCAAAACGGCACGGTTTCGG
RrFAR	SEQ ID NO: 24 GAGGAATAAACCATGACAGATGTGGACCGCAG	SEQ ID NO: 25 TGTTCTACGTAAGCTTCACAGGGCCAGACTTTTGA
Mas2FAR	SEQ ID NO: 26 GAGGAATAAACCATGACCATCACCCCGCTGAG	SEQ ID NO: 27 TGTTCTACGTAAGCTTCATGCGGCGGCTTCAGGTT
DugFAR	SEQ ID NO: 28 GAGGAATAAACCATGAAACCTGCACTGAACGT	SEQ ID NO: 29 TGTTCTACGTAAGCTTCAAATTTTCACCGGGCGGG
HerFAR	SEQ ID NO: 30 GAGGAATAAACCATGACCCAACCGGTGCAGAA	SEQ ID NO: 31 TGTTCTACGTAAGCTTCAACGGGCAACCTGCGGGG
SmFAR	SEQ ID NO: 32 GAGGAATAAACCATGCGTGATTTAAACCGCGA	SEQ ID NO: 33 TGTTCTACGTAAGCTTCATGCCTCGGTGGCGGTCA
ZaFAR	SEQ ID NO: 34 GAGGAATAAACCATGGATACCGCAAGCAGCAA	SEQ ID NO: 35 TGTTCTACGTAAGCTTCAGGCGGCCTGTTTCTGGC
Mar2FAR	SEQ ID NO: 36 GAGGAATAAACCATGAACTATTTTTTTAACCGG	SEQ ID NO: 37 TGTTCTACGTAAGCTTCACCAATAAATGCCGCGCA

Ole2FAR	SEQ ID NO: 38 GAGGAATAAACCATGAGCTACTTCATTACCGG	SEQ ID NO: 39 TGTTCTACGTAAGCTTCACCAATGAATACCGCGCA
Ole3FAR	SEQ ID NO: 40 GAGGAATAAACCATGAGTTATTTCCTGACCGG	SEQ ID NO: 41 TGTTCTACGTAAGCTTCAGGCGAAGGCCACTTGTT
Mar3FAR	SEQ ID NO: 42 GAGGAATAAACCATGAATTATTTCCTGACCGG	SEQ ID NO: 43 TGTTCTACGTAAGCTTCACCAATGAATGCCGCGCA
MlFAR	SEQ ID NO: 44 GAGGAATAAACCATGAACTATTTTCTGACCGG	SEQ ID NO: 45 TGTTCTACGTAAGCTTCACCAGTAGATACCGCGCA
NpFAR	SEQ ID NO: 46 GAGGAATAAACCATGACCAGCATCCCGGAGTA	SEQ ID NO: 47 TGTTCTACGTAAGCTTCAATAACGCATTGTGCTAC
B1FAR	SEQ ID NO: 48 GAGGAATAAACCATGAGCCATACCAGTCCTGC	SEQ ID NO: 49 TGTTCTACGTAAGCTTCACCATTCCAGGATCTGCC

EXAMPLE 3

Evaluation of fatty alcohol production

Recombinant TOP10 *E. coli* strains expressing FAR were grown overnight at 30°C in SOB medium containing 100 ug/mL spectinomycin. Overnight cultures were used to inoculate 50 mL SOB medium with 100 ug/mL spectinomycin to a starting OD600 of 0.1. Cultures were incubated at 30°C shaking at 200 RPM until OD600 reached between 0.6 – 0.8, FAR expression was induced with isopropyl-β-D-thiogalactoside (IPTG) to a final concentration of 1 mM. Induced cultures were incubated under the same condition for about ~20 hours (overnight). 2 x 10 mL cultures were centrifuged at 4000 RPM, 24°C for 10 min. Supernatant and pellet were stored at -20°C or -80°C, respectively.

Fatty alcohols were extracted from both supernatant and pellets separately. Cell pellets were allowed to thaw at room temperature, followed by re-suspension in 0.5 mL heptane, vortexed for 30 seconds, incubated at 35°C for 1 min and vortexed again for 30 seconds. Extracts were centrifuged at 4000 RPM, 10°C for 5 min and the organic layers were analyzed by GC-MS. Supernatants were thawed at room temperature and vortexed to mix well. 1 mL supernatant was transferred to a different extraction vial, 2 mL of heptane were added and vortexed for 1 min at room temperature. The vials caps were carefully opened to allow any gas to vent out, before vortexing again for 1 min. Vials were centrifuged at 4000 RPM, 10°C for 1 min, and 1 mL of the organic heptane layers were transferred into clean vials. Samples were blow-dried with nitrogen and re-dissolved in 0.5 mL of heptane for GC-MS analysis.

2 µl samples were analyzed by GC-MS (Inlet: splitless or split ratio 1:2) on an HP-Innowax column (30 m, 0.25 mm, 0.25 µm, Agilent P/N:1909N-133). GC method started at 60°C (hold for 2 min), increased the temperature with a rate of 10°C/min to 250°C and held for 15 min with total flow, 41 mL/min with 0.5 min delay. Quantitative analysis was performed using a calibration curve made from reference fatty alcohol standards: C8:0-OH, C10:0-OH, C12:0-OH, C14:0-OH, C15:0-OH, C16:0-OH, C16:1-OH, C18:0-OH, C18:1-OH, C18:2-OH, C18:3-OH. Calibration curve ranges were 0.5 mg/L to 100 mg/L with LOD: 0.1 mg/L and LOQ: 0.2 mg/L. Mass spec parameters set-up consisted

of solvent delay at 3.5 min; mass scan range between 35 to 800 mu; collision energy at 70v. Compounds were identified by comparing observed spectral data with the NIST library. Agilent MassHunter software was used for quantitative analysis by calculating concentrations off of an established standard curve.

Evaluations of 24 FAR genes expressed in *E. coli* indicated that not all FAR enzymes were functional in *E. coli*. Surprisingly, of the 24 FAR genes expressed in *E. coli*, only 12 produced fatty alcohols (Figure 3). Analysis of the data showed that under the conditions used, FAR homologs that were more closely related to *M. aquaeolei, M. algicola* and *Oceanobacter* sp. RED65 protein sequences did not always produce fatty alcohols. Unexpectedly, this indicates that sequence homology to previously reported active FARs is not a good indicator when seeking to identify novel active FARs.

Based on level of fatty alcohols produced, some FAR enzymes were more active than others. Of the 12 tested, only the FARs *Spongiibacter marinus* (strain 16), *Aestuariibacter salexigens* (strain 1) and *Zhongshania aliphaticivorans* (strain 17), produced fatty alcohols to a level high enough to make them attractive for applied purposes. As shown in Figure 3, SmFAR (strain 16) markedly outperformed other FARs in the total level of fatty alcohols produced.

EXAMPLE 4

Production of fatty alcohols in *E. coli* strain W3110K

Recombinant W3110K strains expressing a gene encoding a FAR enzyme from either *Spongiibacter marinus*, *Aestuariibacter salexigens* or *Zhongshania aliphaticivorans* were grown overnight at 30°C in SOB medium containing 100 ug/mL spectinomycin. Overnight cultures were used to inoculate 50 mL LB medium with 1% glucose and 100 ug/mL spectinomycin to a starting OD_{600} of 0.1. Cultures were incubated at 30°C shaking at 200 RPM until OD_{600} reached between 0.6 – 0.8, FAR expression was induced with isopropyl-β-D-thiogalactoside (IPTG) to a final concentration of 1 mM. Induced cultures were incubated under the same conditions for about ~20 hours (overnight). 10 mL cultures were collected for fatty alcohol analysis. Extraction and quantification of fatty alcohols were performed as described in Example 3, except 5 ml of heptane was used for extraction.

Comparison of fatty alcohol production in SmFAR with MaFAR

US 8,216,815 describes characteristics of fatty alcohol compositions produced by FAR obtained from *Marinobactoer algicola* (denoted "MaFAR"). Table 3 provides a comparison between the fatty alcohol composition described in US 8,216,815 and the fatty alcohol composition produced using a FAR of the present invention.

Table 3. Comparison between fatty alcohol compositions produced by SmFAR and MaFAR

% distribution	C12:0	C14:0	C14:1	C16:0	C16:1	C18:0	C18:1	Saturated	Unsaturated
SmFAR	1	31	ND	13	48	0	7	45	55
MaFAR	ND	8	ND	30	30	ND	32	38	62

ND = not detected

As set out in Table 3, under the conditions described in US 8,216,815, the fatty alcohol composition produced by MaFAR contained 8% C14 fatty alcohol, 60% C16 fatty alcohol (30% C16:0, 30% C16:1), and there was no detection of <C14 fatty alcohol. On the contrary, SmFAR yielded a fatty alcohol composition containing 31% C14 fatty alcohol, 61% C16 fatty alcohol (13% C16:0, 48% C16:1), and, desirably, a detectable amount of C12 fatty alcohol. Furthermore, the distribution of saturated vs. unsaturated fatty alcohol was different between MaFAR and SmFAR. Advantageously, total saturated fatty alcohol produced using SmFAR was ~45%, whereas MaFAR achieved a total saturation level of ~38%.

Distribution of fatty alcohols produced in SmFAR, ZaFAR and AsFAR

As shown in Table 4, the recombinant *E. coli* strains produced between 228.5 mg/L to 930.3 mg/L of fatty alcohols. The fatty alcohols produced by the recombinant cells had varying levels of distribution between chain lengths. Analysis of the type of fatty alcohols produced by each of the FARs indicate that C14:0 and C16:0 fatty alcohols are produced at levels greater than 45% of total fatty alcohols produced.

Table 4. distribution of fatty alcohols produced by *E. coli* W3110K overexpressing a FAR gene from *S. marinus* (SmFAR), *Z. aliphaticivorans* (ZaFAR), or *A. salexigens* (AsFAR).

FAR	Total (mg/L)	% distribution								
		C10:0	C12:0	C14:0	C16:0	C16:1	C18:0	C18:1	saturated	unsaturated
16 SmFAR	930.3	0	1	31	13	48	0	7	45	55
17 ZaFAR	425.9	1	1	29	30	30	0	10	60	40
1 AsFAR	228.5	3	1	25	33	30	0	8	62	38

In conclusion, the FAR discovered in this work advantageously makes possible the direct production of fatty alcohols, without producing undesirable intermediates. In particular the *S. marinus* FAR produced a high proportion of C14:0 fatty alcohol, also known as Myristyl alcohol. This alcohol has numerous commercial applications, and is currently obtained from coconut and palm kernel oils. The FAR of the invention also produced highly desirable C12 fatty alcohol which, to the best of the Inventors'

knowledge, has not been achieved by FARs in the literature. Moreover, the FAR of the invention provided a high yield of fatty alcohols, and which contain a high proportion of saturated fatty alcohols.

The FAR polypeptide of the invention is of considerable value for the production of fatty alcohols. Moreover, as the FAR polypeptide is already capable of producing shorter chain fatty alcohols (namely C12) than previously reported FARs, the Inventors believe that less strain engineering might be required (in comparison to strains expressing other FARs), when seeking to obtain even shorter chain fatty alcohols.

While illustrative embodiments have been illustrated and described herein, it will be appreciated that various changes can be made therein without departing from the spirit and scope of the contents described herein.

REFERENCES

Akhtar, M. K., N. J. Turner and P. R. Jones (2013). "Carboxylic acid reductase is a versatile enzyme for the conversion of fatty acids into fuels and chemical commodities." Proceedings of the National Academy of Sciences **110**(1): 87-92.

Biermann, U., U. Bornscheuer, M. A. R. Meier, J. O. Metzger and H. J. Schäfer (2011). "Oils and Fats as Renewable Raw Materials in Chemistry." Angewandte Chemie International Edition **50**(17): 3854-3871.

Doan, T. T. P., A. S. Carlsson, M. Hamberg, L. Bülow, S. Stymne and P. Olsson (2009). "Functional expression of five Arabidopsis fatty acyl-CoA reductase genes in Escherichia coli." Journal of Plant Physiology **166**(8): 787-796.

Hofvander, P., T. T. P. Doan and M. Hamberg (2011). "A prokaryotic acyl-CoA reductase performing reduction of fatty acyl-CoA to fatty alcohol." FEBS Letters **585**(22): 3538-3543.

Lennen, R. M., M. A. Kruziki, K. Kumar, R. A. Zinkel, K. E. Burnum, M. S. Lipton, S. W. Hoover, D. R. Ranatunga, T. M. Wittkopp, W. D. Marner and B. F. Pfleger (2011). "Membrane Stresses Induced by Overproduction of Free Fatty Acids in Escherichia coli." Applied and Environmental Microbiology 77(22): 8114-8128.

Lerner, C. G. and M. Inouye (1990). "Low copy number plasmids for regulated low-level expression of cloned genes in Escherichia coli with blue/white insert screening capability." Nucleic acids research **18**(15): 4631-4631.

Liu, Y., S. Chen, J. Chen, J. Zhou, Y. Wang, M. Yang, X. Qi, J. Xing, Q. Wang and Y. Ma (2016). "High production of fatty alcohols in Escherichia coli with fatty acid starvation." Microbial Cell Factories **15**(1): 129.

Metz, J. G., M. R. Pollard, L. Anderson, T. R. Hayes and M. W. Lassner (2000). "Purification of a Jojoba Embryo Fatty Acyl-Coenzyme A Reductase and Expression of Its cDNA in High Erucic Acid Rapeseed." Plant Physiology **122**(3): 635-644.

Reiser, S. and C. Somerville (1997). "Isolation of mutants of Acinetobacter calcoaceticus deficient in wax ester synthesis and complementation of one mutation with a gene encoding a fatty acyl coenzyme A reductase." Journal of Bacteriology **179**(9): 2969-2975.

Rowland, O. and F. Domergue (2012). "Plant fatty acyl reductases: Enzymes generating fatty alcohols for protective layers with potential for industrial applications." Plant Science **193-194**: 28-38.

Schirmer, A., M. A. Rude, X. Li, E. Popova and S. B. del Cardayre (2010). "Microbial Biosynthesis of Alkanes." Science **329**(5991): 559-562.

Steen, E. J., Y. Kang, G. Bokinsky, Z. Hu, A. Schirmer, A. McClure, S. B. del Cardayre and J. D. Keasling (2010). "Microbial production of fatty-acid-derived fuels and chemicals from plant biomass." Nature **463**: 559.

Willis, R. M., B. D. Wahlen, L. C. Seefeldt and B. M. Barney (2011). "Characterization of a Fatty Acyl-CoA Reductase from Marinobacter aquaeolei VT8: A Bacterial Enzyme Catalyzing the Reduction of Fatty Acyl-CoA to Fatty Alcohol." Biochemistry **50**(48): 10550-10558.

Zheng, Y.-N., L.-L. Li, Q. Liu, J.-M. Yang, X.-W. Wang, W. Liu, X. Xu, H. Liu, G. Zhao and M. Xian (2012). "Optimization of fatty alcohol biosynthesis pathway for selectively enhanced production of C12/14 and C16/18 fatty alcohols in engineered Escherichia coli." Microbial Cell Factories **11**(1): 65.

CLAIMS

1. A recombinant polypeptide comprising an amino acid sequence having at least 70% sequence identity to SEQ ID NO: 1.

2. An isolated nucleic acid which encodes the polypeptide according to claim 1.

3. The nucleic acid according to claim 2, wherein the nucleic acid is codon optimized for expression in a prokaryotic cell.

4. The nucleic acid according to claim 3, wherein the prokaryotic cell is selected from the group consisting of *Acinetobacter, Agrobacterium, Escherichia , Cupriavidus, Clostridium, Rhodobacter, Marinobacter, Bacillus, Klebsiella, Tatumella, Pseudomonas, Ralstonia, Rhodococcus, Methylobacterium, Methylophilus, Methylococcus, Methylomicrobium, Methylomonas, Pantoea, Streptomyces, Parachlorella, Synechococcus, Synechocystis* and *Thermocynechococcus*.

5. The nucleic acid according to claim 4, wherein the prokaryotic cell is *E. coli*.

6. The nucleic acid according to claim 5 comprising a nucleic acid sequence having at least 70% sequence identity to SEQ ID NO: 2.

7. The nucleic acid according to claim 2, wherein the nucleic acid sequence is codon optimized for expression in a eukaryotic cell.

8. The nucleic acid according to claim 7, wherein the eukaryotic cell is selected from the group consisting of a yeast cell, a fungal cell, an algal cell and a plant cell.

9. The nucleic acid according to claim 8, wherein the yeast cell is selected from the group consisting of *Saccharomyces, Kluyveromyces, Candida, Pichia, Schizosaccharomyces, Hansenula, Klockera, Schwanniomyces, Schefferomyces, Rhodosporidium, Issatchenkia, Yarrowia* and *Rhodotorula*.

10. The nucleic acid according to claim 9, wherein the yeast cell is selected from the group consisting of *S. cerevisiae, C. lipolytica, R. glutinis, S. bulderi, S. barnetti, S. exiguus, S. uvarum, S. diastaticus, K. lactis, K. marxianus K. fragile, P. kudriavzevii, S. stipitis,* and *I. orientalis*.

11. The nucleic acid according to claim 8, wherein the fungal cell is a filamentous fungal cell.

12. The nucleic acid according to claim 11, wherein the filamentous fungal cell is selected from the group consisting of *Aspergillus*, *Penicillium*, *Rhizopus*, *Chrysosporium*, *Myceliophthora*, *Trichoderma*, *Humicola*, *Acremonium* and *Fusarium*.

13. The nucleic acid according to claim 12, wherein the filamentous fungal cell is selected from the group consisting of *A. niger*, *A. oryzae*, *T. reesei*, *P. chrysogenum*, *M. thermophila*, and *R. oryzae*.

14. The nucleic acid according to claim 8, wherein the algal cell is selected from the group comprising *Botryococcus*, *Nannochloropsis*, *Chlorella*, *Chlamydomonas*, *Dunaliella*, *Chaetoceros*, *Porphyridium*, *Scenedesmus* and *Pseudochlorococcum*.

15. The nucleic acid according to claim 14, wherein the algal cell is selected from the group comprising *B. braunii* and *N. gaditana*.

16. The nucleic acid according to claim 2 comprising a nucleic acid sequence having at least 70% sequence identity to SEQ ID NO: 3.

17. A vector comprising the nucleic acid according to any one of claims 2 to 16.

18. The vector according to claim 17 comprising:

(a) an origin of replication;

(b) a promoter sequence operably linked to said nucleic acid; and/or

(c) a reporter gene.

19. The vector according to claim 17 or claim 18 comprising a nucleic acid sequence having at least 70% sequence identity to SEQ ID NO: 4.

20. A recombinant cell engineered to express the polypeptide according to claim 1.

21. The recombinant cell according to claim 20, wherein the recombinant cell is a prokaryotic cell.

22. The recombinant cell according to claim 21, wherein the prokaryotic cell is selected from the group consisting of *Acinetobacter*, *Agrobacterium*, *Escherichia* , *Cupriavidus*, *Clostridium*,

Rhodobacter, Marinobacter, Bacillus, Klebsiella, Tatumella, Pseudomonas, Ralstonia, Rhodococcus, Methylobacterium, Methylophilus, Methylococcus, Methylomicrobium, Methylomonas, Pantoea, Streptomyces, Parachlorella, Synechococcus, Synechocystis, and *Thermocynechococcus.*

23. The recombinant cell according to claim 22, wherein the prokaryotic cell is *E. coli.*

24. The recombinant cell according to claim 20, wherein the recombinant cell is a eukaryotic cell.

25. The recombinant cell according to claim 24, wherein the eukaryotic cell is selected from the group consisting of a yeast cell, a fungal cell, an algal cell and a plant cell.

26. The recombinant cell according to claim 25, wherein the yeast cell is selected from the group consisting of *Saccharomyces, Kluyveromyces, Candida, Pichia, Schizosaccharomyces, Hansenula, Klockera, Schwanniomyces, Schefferomyces, Rhodosporidium, Issatchenkia, Yarrowia* and *Rhodotorula.*

27. The recombinant cell according to claim 26, wherein the yeast cell is selected from the group consisting of *S. cerevisiae, C. lipolytica, R. glutinis, S. bulderi, S. barnetti, S. exiguus, S. uvarum, S. diastaticus, K. lactis, K. marxianus, K. fragile, P. kudriavzevii, S. stipites,* and *I. orientalis.*

28. The recombinant cell according to claim 25, wherein the fungal cell is a filamentous fungal cell.

29. The recombinant cell according to claim 28, wherein the filamentous fungal cell is selected from the group consisting of *Aspergillus, Penicillium, Rhizopus, Chrysosporium, Myceliophthora, Trichoderma, Humicola, Acremonium* and *Fusarium.*

30. The recombinant cell according to claim 29, wherein the filamentous fungal cell is selected from the group consisting of *A. niger, A. oryzae, T. reesei, P. chrysogenum, M. thermophila,* and *R. oryzae.*

31. The recombinant cell according to claim 25, wherein the algal cell is selected from the group consisting of *Botryococcus, Nannochloropsis, Chlorella, Chlamydomonas, Dunaliella, Chaetoceros, Porphyridium, Scenedesmus* and *Pseudochlorococcum.*

32. The recombinant cell according to claim 31, wherein the algal cell is selected from the group consisting of *B.* and *N. gaditana.*

33. The recombinant cell according to any one of claims 20 to 32 comprising one or more genetic modifications resulting in at least one, any two, any three, any four or all five of the following phenotypes:

(a) an increase in acyl-ACP synthesis;

(b) a decrease in acyl-ACP degradation;

(c) an upregulation of fatty alcohol secretion;

(d) an increase in flux through the fatty acid biosynthetic pathway; and

(e) an increase in tolerance to fatty alcohol.

34. The recombinant cell according to any one of claims 20 to 33 comprising one or more genetic modifications resulting in a change in the degree of saturation of the fatty alcohols produced.

35. The recombinant cell according to any one of claims 20 to 34 comprising one or more genetic modifications resulting in a change in the chain-length of the acyl-ACPs produced by the fatty acid biosynthetic pathway.

36. A recombinant cell transformed with the vector according to any one of claims 17 to 19.

37. The recombinant cell according to claim 36, wherein the recombinant cell is stably transformed with the nucleic acid according to any one of claims 2 to 16.

38. The recombinant cell according to claim 36, wherein the recombinant cell is transiently transformed with the nucleic acid according to any one of claims 2 to 16.

39. A method of producing a fatty alcohol composition comprising:

a) culturing a recombinant cell in a suitable culture medium, wherein the recombinant cell comprises a gene encoding a heterologous fatty acyl-ACP reductase enzyme having at least 70% sequence identity to SEQ ID NO: 1; and

b) allowing expression of said gene, wherein said expression results in the production of a fatty alcohol composition.

40. The method according to claim 39, wherein the recombinant cell is defined according to any one of claims 20 to 38.

41. The method according to claim 39 or claim 40, the method further comprising purifying the fatty alcohol composition from cell culture.

42. The method according to claims 39 or claim 41, the method further comprising purifying the fatty alcohol composition from supernatant.

43. The method according to any one of claims 39 to 42, the method further comprising harvesting and lysing recombinant cell(s) to obtain a fatty alcohol composition.

44. The method according to any one of claims 39 to 43, the method further comprising purifying the fatty alcohol composition from lysate.

45. The method according to any one of claims 39 to 44, wherein the fatty alcohol composition comprises C12 fatty alcohol.

46. The method according to any one of claims 39 to 45, wherein at least 45% of the fatty alcohols in the fatty alcohol composition are saturated fatty alcohols.

47. The method according to any one of the claims 39 to 46, the method further comprising reducing the fatty alcohol composition.

48. The method according to any one of claims 39 to 47, the method further comprising esterifying the fatty alcohol composition.

49. A fatty alcohol composition obtained by the method according to any one of claims 39 to 48.

50. A composition comprising the fatty alcohol composition or esterified fatty alcohol composition obtained by the method of any one of claims 39 to 49.

51. Use of a polypeptide according to claim 1 in a method of producing a fatty alcohol composition.

FIGURE 1

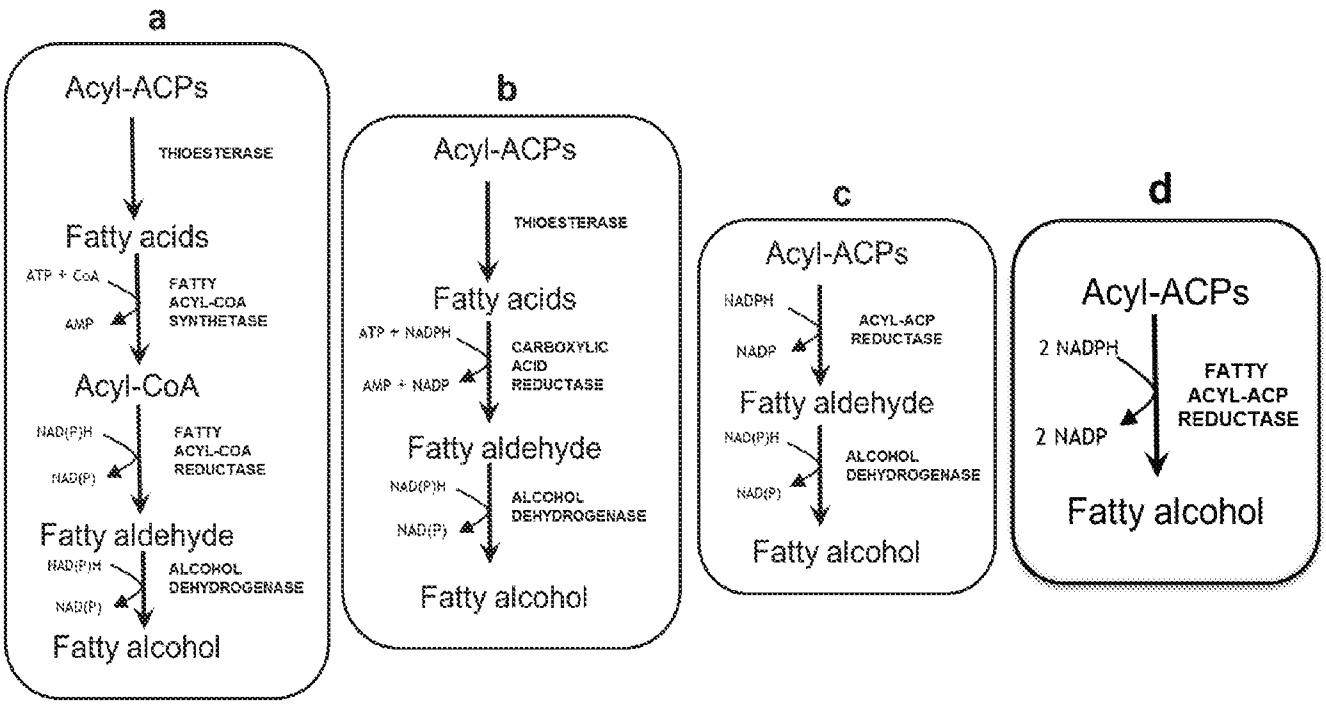

a
Acyl-ACPs
THIOESTERASE
Fatty acids
ATP + CoA
AMP
FATTY ACYL-COA SYNTHETASE
Acyl-CoA
NAD(P)H
NAD(P)
FATTY ACYL-COA REDUCTASE
Fatty aldehyde
NAD(P)H
NAD(P)
ALCOHOL DEHYDROGENASE
Fatty alcohol
b
Acyl-ACPs
THIOESTERASE
Fatty acids
ATP + NADPH
AMP + NADP
CARBOXYLIC ACID REDUCTASE
Fatty aldehyde
NAD(P)H
NAD(P)
ALCOHOL DEHYDROGENASE
Fatty alcohol
c
Acyl-ACPs
NADPH
NADP
ACYL-ACP REDUCTASE
Fatty aldehyde
NAD(P)H
NAD(P)
ALCOHOL DEHYDROGENASE
Fatty alcohol
d
Acyl-ACPs
2 NADPH
2 NADP
FATTY ACYL-ACP REDUCTASE
Fatty alcohol

FIGURE 2

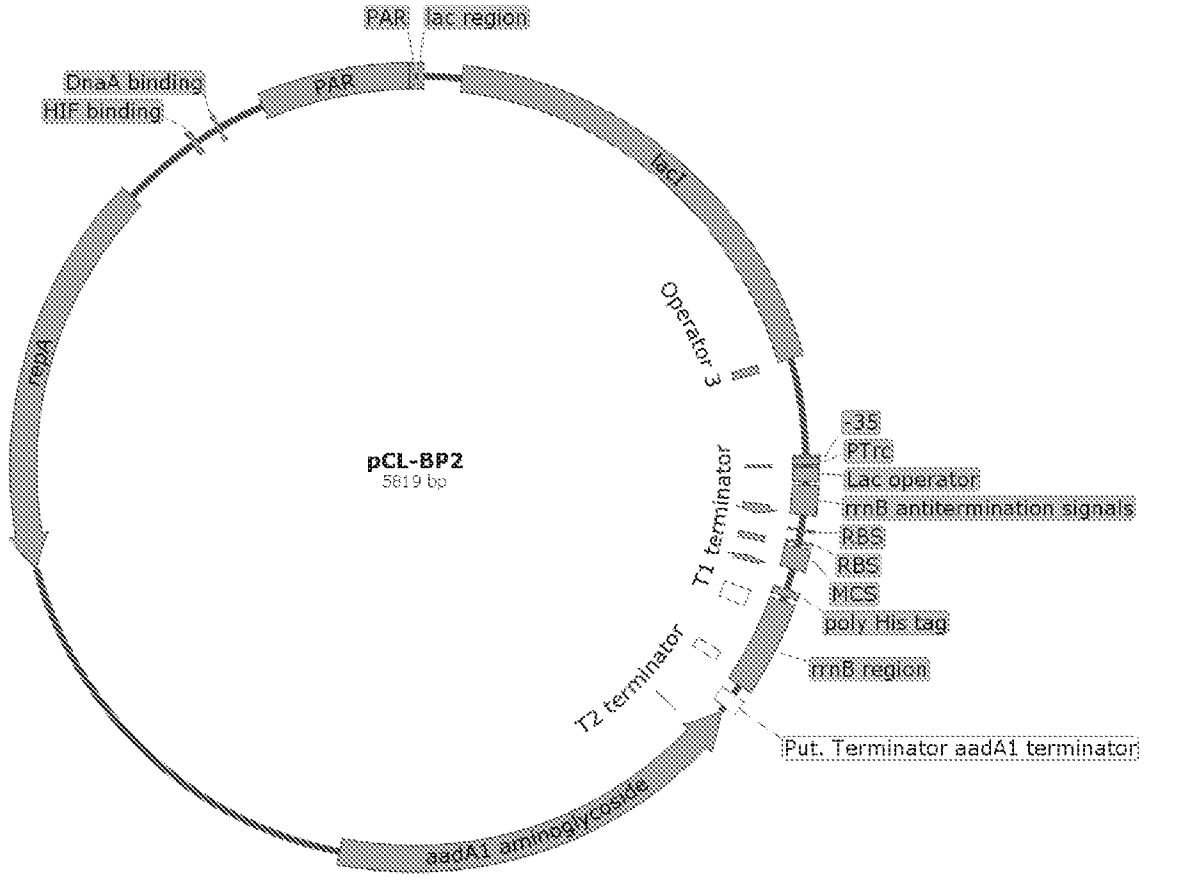
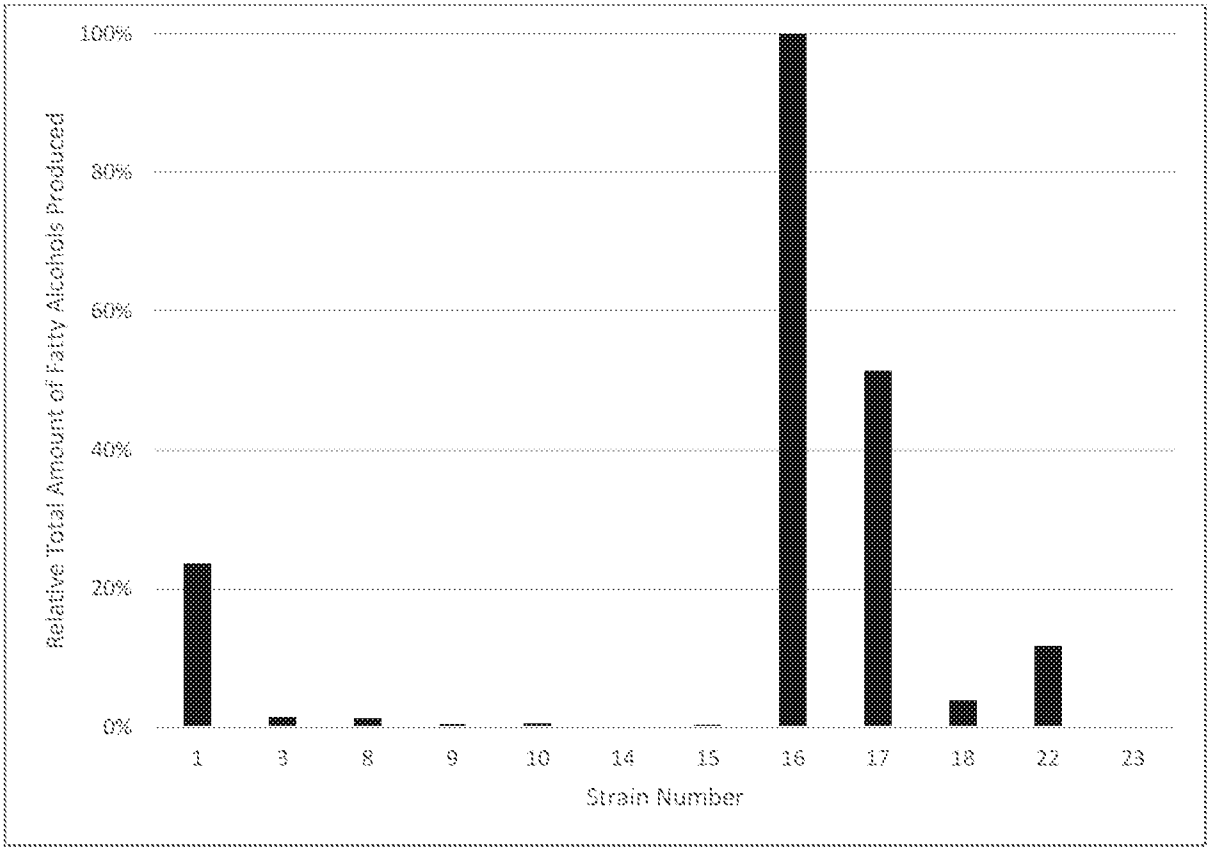

FAR
lac region
DnaA binding
HIF binding
FAR
lacI
Operator 3
pCL-BP2
5819 bp
T1 terminator
T2 terminator
aadA1 aminoglycoside
35
PTrc
Lac operator
rrnB antitermination signals
RBS
RBS
MCS
poly His tag
rrnB region
Put. Terminator aadA1 terminator

FIGURE 3


100%
80%
60%
40%
20%
0%
Relative Total Amount of Fatty Alcohols Produced
1
3
8
9
10
14
15
16
17
18
22
23
Strain Number

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2020/021199

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N9/02 C12P7/04 C07K14/195
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12N C12P C12R C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, BIOSIS, CHEM ABS Data, EMBASE, WPI Data, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE UniProt [Online] 31 January 2018 (2018-01-31), Tully B.J., Graham E.D., Heidelberg J.F.: "Full=Dehydrogenase; Spongiibacter sp.", XP055706350, retrieved from EBI accession no. UNIPROT:A0A2E9SL90 Database accession no. A0A2E9SL90	1,2,17, 18,20,21
Y	sequence -----	3-16,19, 22-48
X	WO 2013/159055 A1 (CODEXIS INC [US]) 24 October 2013 (2013-10-24)	49-51
Y	paragraphs [0006] - [0008] paragraphs [0090] - [0094]; claims ----- -/--	3-16,19, 22-48

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
22 June 2020	30/06/2020

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Burkhardt, Peter

1

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2020/021199

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2016/159869 A1 (BIOPETROLIA AB [SE]) 6 October 2016 (2016-10-06)	49-51
Y	page 20, lines 5-9 page 23, line 26 - page 24, line 9; claims; sequence 8	3-16,19, 22-48

X	WO 2012/006114 A2 (CODEXIS INC [US]; BEHROUZIAN BEHNAZ [US] ET AL.) 12 January 2012 (2012-01-12)	49-51
Y	paragraph [0150] - paragraph [0160]; claims; sequence 13	3-16,19, 22-48

X	WO 2013/096082 A1 (CODEXIS INC [US]) 27 June 2013 (2013-06-27)	49-51
Y	paragraph [0077] - paragraph [0088]; claims	3-16,19, 22-48

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2020/021199

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2013159055 A1	24-10-2013	US 2015133698 A1	14-05-2015
		US 2016244787 A1	25-08-2016
		WO 2013159055 A1	24-10-2013
WO 2016159869 A1	06-10-2016	CN 108271383 A	10-07-2018
		EP 3277827 A1	07-02-2018
		US 2018112240 A1	26-04-2018
		WO 2016159869 A1	06-10-2016
WO 2012006114 A2	12-01-2012	CA 2804019 A1	12-01-2012
		EP 2585587 A2	01-05-2013
		US 2012009640 A1	12-01-2012
		US 2014093929 A1	03-04-2014
		WO 2012006114 A2	12-01-2012
WO 2013096082 A1	27-06-2013	EP 2794884 A1	29-10-2014
		EP 2794893 A1	29-10-2014
		US 2014336423 A1	13-11-2014
		US 2014357727 A1	04-12-2014
		WO 2013096082 A1	27-06-2013
		WO 2013096092 A1	27-06-2013