



- (51) International Patent Classification:
C12P 5/02 (2006.01) C12N 9/88 (2006.01)
- (21) International Application Number:
PCT/EP2016/065389
- (22) International Filing Date:
30 June 2016 (30.06.2016)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
15174554.4 30 June 2015 (30.06.2015) EP
- (71) Applicants: **TOTAL MARKETING SERVICES** [FR/FR]; 24 Cours Michelet La Défense 10, 92800 Puteaux (FR). **QINGDAO INSTITUTE OF BIOENERGY AND BIOPROCESS TECHNOLOGY (QIBEBT)** [CN/CN]; Chinese Academy of Sciences, 189 Songling Road, Qindao Shandong (CN).
- (72) Inventors: **FOURAGE, Laurent**; 8bis rue Michelet, 92150 Suresnes (FR). **LAEUFFER, Frédéric**; 36 rue du Caire, 75002 Paris (FR). **STRUB, Henri**; 56 rue de la chapelle Saint Jean, 60700 Pont Sainte Maxence (FR). **WANG, Yun**; 189 Songling Road, Qindao Shandong (CN). **XU, Jian**; 189 Songling Road, Qindao Shandong (CN). **XU, Huifang**; 189 Songling Road, Qindao Shandong (CN). **LI, Shengying**; 189 Songling Road, Qindao Shandong (CN).
- (74) Agents: **PAEMEN, Liesbet** et al.; E. Gevaertdreef 10a, 9830 Sint-Martens-Latem (BE).

- (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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).

Declarations under Rule 4.17:

— of inventorship (Rule 4.17(iv))

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))

— with sequence listing part of description (Rule 5.2(a))

(54) Title: PRODUCTION OF α -OLEFINS

(57) Abstract: The present invention relates to the biosynthesis of α -olefins. In particular, the invention provides methods for the production of medium-chain α -olefins, more particularly C₁₁ α -olefins, using a polypeptide with decarboxylase activity on free fatty acids with 8 to 14 carbons, in particular on C₈-C₁₂ free fatty acids, more particularly on C₁₂ free fatty acids, or a genetically engineered host cell expressing or overexpressing said polypeptide.



PRODUCTION OF α -OLEFINS

TECHNICAL FIELD

The application generally relates to biosynthetic processes. In particular, the
5 application relates to the biosynthesis of α -olefins.

BACKGROUND

Unsaturated hydrocarbons such as α -olefins are an industrially important group of
molecules which can serve as precursors for lubricants and surfactants in addition to
10 being used in fuels, or as feedstock for further chemical conversion to e.g. fuels,
polymers, plastics, textiles, solvents, adhesives, etc.

α -Olefins, more particularly linear α -olefins, have traditionally been produced from
petroleum sources through oligomerization of ethylene but it would be advantageous to
find alternative methods for producing such α -olefins, notably from renewable
15 resources. Moreover, the chemical synthesis results in a mixture of linear α -olefins
which exhibit different carbon chain length. Such mixture may not be desired as
feedstock for further chemical conversion such as the production of poly- α -olefins of
interest.

Microbial production of olefins has been reported. Examples include a three-gene
20 cluster responsible for generating alkenes with internal double bonds through the head-
to-head condensation of two fatty acyl-coenzyme A (acyl-CoA) molecules in
Micrococcus luteus (Beller et al. 2010 Appl Environ Microbiol 76:1212-1223); a unique
P450 decarboxylase OleT_{JE} from *Jeotgalicoccus* sp. ATCC 8456, which directly
decarboxylates C₁₂ to C₂₀ free fatty acids to form α -olefins in presence of H₂O₂ (WO
25 2009/085278, Rude et al. 2011 Appl Environ Microbiol 77:1718-1727, Liu et al. 2014
Biotechnology for Biofuels 7:28); another enzyme of the cyp152 P450 enzyme family,
namely P450_{BS β} from *Bacillus subtilis* subsp. *subtilis* str. 168, for which decarboxylase
activity on palmitic acid has been shown (WO 2009/085278, Rude et al. 2011); and a
type I polyketide synthase from *Synechococcus* sp. PCC 7002, which is capable of
30 transforming fatty acyl-ACPs into α -olefins via sequential polyketide synthase chain
elongation, keto reduction, sulfonation mediated by its sulfotransferase domain, and

the coupled hydrolysis and decarboxylation catalyzed by the thioesterase domain (Mendez-Perez et al. 2011 Appl Environ Microbiol 77:4264-4267).

In general, there is a need for alternative production methods of α -olefins, and more particularly improved processes that allow production of α -olefins at a lower cost and/or that are more environmental friendly. It is also desired to provide production methods which allow tailoring of the carbon chain length of the α -olefins, more particularly a production method for medium-chain α -olefins (i.e. α -olefins having a carbon chain length comprised between 7 and 13), more particularly for C₁₁ α -olefins.

10 SUMMARY OF THE INVENTION

The present invention solves one or more of the above described problems of the prior art. In particular, methods are provided for the production of α -olefins which allow controlling the carbon chain length of the α -olefins.

The present invention is based, at least in part, on the discovery that certain genomic sequences of *Alicyclobacillus acidocaldarius* and *Staphylococcus massiliensis* encode enzymes having free fatty acid decarboxylase activity. None of these organisms have previously been reported to produce terminal olefins. It has further been found that certain polypeptides can be identified which have specific decarboxylase activity on medium-chain free fatty acids, in particular C₈ to C₁₂ free fatty acids, more particularly on C₁₂ free fatty acids, thereby producing medium-chain α -olefins, in particular C₇ to C₁₁ α -olefins, more particularly C₁₁ α -olefins. The application provides examples thereof, more particularly an olefin-producing enzyme (Sm46) identified in *Staphylococcus massiliensis*, the P450_{BS β} fatty acid hydroxylase (Bs168) from *Bacillus subtilis* subsp. *subtilis* str. 168 and the decarboxylase enzyme identified in *Alicyclobacillus acidocaldarius*, have specific decarboxylase activity on C₁₂ free fatty acids, thereby producing C₁₁ α -olefins. The identification of these enzymes facilitates the further identification of similar enzymes from other organisms.

The present invention is in particular captured by any one or any combination of one or more of the below numbered aspects and embodiments (i) to (xvii) wherein:

(i) A method for the production of C₇-C₁₁ α-olefins comprising culturing a recombinant host cell comprising a recombinant nucleic acid encoding a decarboxylase enzyme under conditions suitable for the production of C₇ to C₁₁ α-olefins by said host cell, wherein the preferred substrate of said decarboxylase enzyme is an C₈-C₁₂ free fatty acid.

(ii) The method according to (i), wherein said nucleic acid encoding a decarboxylase enzyme comprises a nucleotide sequence having at least about 75%, preferably at least about 80%, more preferably at least about 95%, sequence identity to SEQ ID NO:10, SEQ ID NO:7, SEQ ID NO: 3 or SEQ ID NO: 5 and wherein said recombinant nucleic acid ensures expression or overexpression of said decarboxylase.

(iii) The method according to (i) or (ii), wherein said nucleic acid encodes a polypeptide comprising an amino acid sequence having at least 80%, preferably at least 90% identity, to SEQ ID NO:13, SEQ ID NO: 8, SEQ ID NO: 4 or SEQ ID NO: 6 or an active fragment of said polypeptide.

(iv) The method according to any one of (i) to (iii), wherein said nucleic acid encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:13.

(v) The method according to any one of (i) to (iv), wherein C₁₂ free fatty acids are the preferred substrate of said decarboxylase, and wherein said α-olefins are C₁₁ α-olefins.

(vi) The method according to any one of (i) to (v), wherein said host cell is cultivated in a medium comprising C₈-C₁₂ free fatty acids, preferably C₁₂ free fatty acids.

(vii) The method according to any one of (i) to (vi), wherein the host cell has further been genetically engineered to produce or overproduce C₈-C₁₂ free fatty acids, preferably C₁₂ free fatty acids.

(viii) The method according to (vii), wherein the host cell comprises a recombinant nucleic acid encoding an enzyme involved in the production of free fatty acids with a carbon chain length comprised between 8 and 12, more particularly a C₁₂ free fatty acid.

(ix) The method according to claim (viii), wherein the host cell comprises a recombinant nucleic acid encoding a thioesterase having activity on C₈ to C₁₂ acyl-ACP, preferably on C₁₂ acyl-ACP.

5

(x) The method according to any one of (i) to (ix), wherein the host cell is an oleaginous host cell.

10

(xi) The method according to any one of (i) to (x), wherein the host cell is selected from the group comprising bacteria, yeasts, fungi, plants and algae.

(xii) The method according to any one of (i) to (xi), further comprising the step of recovering the α -olefins from the host cell or the culture medium.

15

(xiii) A method for the production of poly- α -olefins comprising the following steps:

i) producing C₇-C₁₁ α -olefins according to a method according to any one of (i) to (xii); and

ii) performing an oligomerization reaction using the α -olefins obtained in step i) as monomer to produce an oligomer; and optionally,

20

iii) hydrogenating the oligomer produced in step ii).

(xiv) The method according to (xiii), wherein the poly- α -olefins are C₃₃ poly- α -olefins, wherein step i) comprises the production of C₁₁ α -olefins, and wherein the oligomerization reaction in step ii) is a trimerization reaction.

25

(xv) A composition comprising poly- α -olefins obtainable by the method according to (xiv), wherein at least 85%, preferably at least 90%, more preferably at least 95%, of the poly- α -olefins are C₃₃ poly- α -olefins.

30

(xvi) A polypeptide having decarboxylase activity, wherein said polypeptide comprises an amino acid sequence having at least 80%, preferably at least 90%, identity to SEQ ID NO:13.

35

(xvii) An isolated nucleic acid encoding an enzyme having decarboxylase activity, wherein said nucleic acid comprises a nucleotide sequence having at least about 80%,

more preferably at least about 95%, sequence identity to SEQ ID NO: 7 or SEQ ID NO: 5.

5 (xviii) The polypeptide according to (xvi) or the isolated nucleic acid according to (xvii), wherein C₈ to C₁₄ free fatty acids, preferably C₈ to C₁₂ free fatty acids, more preferably C₁₂ free fatty acids are the preferred substrate of said decarboxylase.

10 (xix) The isolated nucleic acid according to (xvii) or (xviii), wherein said nucleic acid comprises a nucleotide sequence having at least about 95% sequence identity to SEQ ID NO: 7.

(xx) A vector comprising at least one regulatory sequence operatively coupled to the nucleic acid sequence of any of (xvii) to (xix).

15 (xxi) A host cell comprising the nucleic acid according to any of claims (xvii) to (xix) integrated into its genome or the vector according to claim 20.

20 (xxii) The host cell of (xxi), which is an oleaginous eukaryotic microalgae or oleaginous yeast.

(xxiii) Use of a host cell according to (xxi) or (xxii) for the industrial production of lubricants.

25 (xxiv) Lubricant comprising poly- α -olefins obtainable by a method according to (xiii) or (xiv).

(xxv) Lubricant according to (xxiv), wherein said poly- α -olefins are biosourced poly- α -olefins.

30 (xxvi) Lubricant according to (xxiv) or (xxv), wherein at least 50%, preferably at least 85%, of said poly- α -olefins consist of C₃₃ poly- α -olefins.

35 (xxvii) Lubricant according to any one of (xxiv) to (xxvi), wherein said lubricant is an automotive lubricant.

BRIEF DESCRIPTION OF THE FIGURES

The teaching of the application is illustrated by the following Figures which are to be considered as illustrative only and do not in any way limit the scope of the claims.

Figure 1: Nucleotide (A,C,E,G) and amino acid (B,D,F,H) sequences of OleT_{JE} (A,B),
5 Bs168 (C,D), Aa162 (E,F) and Sm46 (G,H).

Figure 2: Plasmid map of pET28a.

Figure 3: Heterologous expression of genes encoding OleT_{JE}, Bs168, Aa162 and Sm46 in *E. coli*. SDS-PAGE analysis of lysates of *E. coli* cells carrying a recombinant plasmid comprising a gene encoding OleT_J, Bs168, Aa162 or Sm46, cultured under
10 conditions to induce expression of the transgene (0.4 mM IPTG) or not (uninduced). M: protein marker.

Figure 4: *In vitro* production of α -olefins by OleT_{JE}, Bs168, Aa162 and Sm46. Genes encoding OleT_J, Bs168, Aa162 and Sm46 were recombinantly expressed in *E. coli*. The supernatant of the *E. coli* lysates was used in an *in vitro* assay for α -olefin
15 biosynthesis in the presence of C₁₂, C₁₄ or C₁₆ fatty acid substrate (1 mM) and H₂O₂ (500 μ M). C₁₁, C₁₃ and C₁₅ α -olefin production by the different enzymes is indicated as the area under the gas chromatographic (GC) peak.

Figure 5: (A) Codon-optimized (for expression in *E. coli*) nucleotide sequences encoding Sm46 and a truncated variant Sm46-del29 that has the N-terminal 29 amino
20 acids deleted (SEQ ID NO:9 and SEQ ID NO:10, respectively) with flanking restriction sites. Codon-optimized (for expression in *E. coli*) nucleotide sequence (SEQ ID NO: 9) encoding Sm46 is shown; Nde I and Xho I restriction sites are indicated in bold, the deleted sequence to encode for the truncated variant Sm46-del29 is underlined. **(B)** Amino acid sequence of Sm46-del29 (SEQ ID NO:13).

Figure 6: *In vitro* conversion of free fatty acids into α -olefins by a truncated variant of Sm46: Sm46-del29. Conversion percentages of the free fatty acid substrates (A) and the corresponding α -alkene products (B) are shown.

Figure 7: *In vitro* production of α -olefins by Aa162, Bs168 and Sm46-del29. Recombinantly expressed and purified enzyme (0.2 μ M) was used in an *in vitro* assay
30 for α -olefin biosynthesis in the presence of C₁₄ fatty acid substrate. The conversion

percentages are shown for, from left to right, C₁₄ free fatty acid substrate, C₁₃ α -olefin, α -hydroxy C₁₄ fatty acid, β -hydroxy C₁₄ fatty acid.

DETAILED DESCRIPTION OF THE INVENTION

5 Unless otherwise defined, all terms used in disclosing the invention, including technical and scientific terms, have the meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. By means of further guidance, term definitions are included to better appreciate the teaching of the present invention.

As used herein, the singular forms "a", "an", and "the" include both singular and plural
10 referents unless the context clearly dictates otherwise.

The terms "comprising", "comprises" and "comprised of" as used herein are synonymous with "including", "includes" or "containing", "contains", and are inclusive or open-ended and do not exclude additional, non-recited members, elements or method steps. Where reference is made to embodiments as comprising certain elements or
15 steps, this encompasses also embodiments which consist essentially of the recited elements or steps.

The recitation of numerical ranges by endpoints includes all numbers and fractions subsumed within the respective ranges, as well as the recited endpoints.

The term "about" as used herein when referring to a measurable value such as a
20 parameter, an amount, a temporal duration, and the like, is meant to encompass variations of +/-10% or less, preferably +/-5% or less, more preferably +/-1% or less, and still more preferably +/-0.1% or less of and from the specified value, insofar such variations are appropriate to perform in the disclosed invention. It is to be understood that the value to which the modifier "about" refers is itself also specifically, and
25 preferably, disclosed.

All documents cited in the present specification are hereby incorporated by reference in their entirety.

Standard reference work setting forth the general principles of biochemistry includes Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, ed. Michal, G,
30 John Wiley and Sons, Inc., New York, US, 1999.

The term "olefin" or "alkene" refers herein to molecules composed of carbon and hydrogen, containing at least one carbon-carbon double bond. Olefins containing one carbon-carbon double bond are denoted herein as mono-unsaturated hydrocarbons and have the chemical formula C_nH_{2n} , where n equals at least two.

5 "Alpha-olefins", " α -olefins", "1-alkenes" or "terminal olefins" are used as synonyms herein and denote olefins or alkenes having a double bond at the primary or alpha (α) position. "Linear α -olefins" or "LAO" as used herein refer to α -olefins that have a linear hydrocarbon chain, whereas "branched α -olefins" have a branch on one or more carbon atoms of the hydrocarbon chain. The term "medium-chain α -olefins" is used
10 herein to denote α -olefins with 7 to 13 carbons and encompasses any one or more of C_7 α -olefins, C_8 α -olefins, C_9 α -olefins, C_{10} α -olefins, C_{11} α -olefins, C_{12} α -olefins, and C_{13} α -olefins. The term "uneven-numbered α -olefins" refers to α -olefins wherein the number of carbon atoms is not even in number. Thus, uneven-numbered medium-chain α -olefins encompass C_7 , C_9 , C_{11} and C_{13} α -olefins.

15 As used herein, the term "fatty acid" or "free fatty acid" means a carboxylic acid having the formula $RCOOH$, or a salt ($RCOO^-$) thereof. R represents an aliphatic group, preferably an alkyl group. Fatty acids can be saturated, mono-unsaturated, or poly-unsaturated. The term "medium-chain fatty acid" or "medium-chain free fatty acid" as used herein denotes a fatty acid or free fatty acid having 8 to 14 carbon atoms. The
20 term "even-numbered fatty acids" refers to fatty acids wherein the number of carbon atoms is even in number. Thus, even-numbered medium-chain fatty acids encompass C_8 , C_{10} , C_{12} and C_{14} fatty acids.

As used herein, the term "host cell" refers to a cell that can be used to produce an α -olefin as described herein. A host cell may be an isolated cell or a cell line grown in
25 culture, or a cell which resides in a living tissue or organism.

As used herein, the terms "microbial", "microbial organism" or "micro-organism" are intended to mean any organism that exists as a microscopic cell that is included within the domains of archaea, bacteria or eukaryotes. Therefore, the term is intended to encompass prokaryotic or eukaryotic cells or organisms having a microscopic size and
30 includes bacteria, archaea and eubacteria such as cyanobacteria of all species as well as eukaryotic micro-organisms such as fungi, including yeasts, and algae. The term

also includes cell cultures of any species that can be cultured for the production of a biochemical.

The term "oleaginous" as used herein with reference to a host cell denotes cells characterized by their lipid accumulation capability. Typically, their biomass contains
5 over 20% lipids in dry matter.

The "algae" group encompasses, without limitation, (i) several eukaryotic phyla, including the Rhodophyta (red algae), Chlorophyta (green algae), Dinoflagellata, Haptophyta, (ii) several classes from the eukaryotic phylum Heterokontophyta which include without limitation the classes Bacillariophyceae (diatoms), Eustigmatophyceae,
10 Phaeophyceae (brown algae), Xanthophyceae (yellow-green algae) and Chrysophyceae (golden algae), and (iii) the prokaryotic phylum Cyanobacteria (blue-green algae).

The term "algae" includes for example genera selected from : *Achnanthes*, *Amphora*,
Anabaena, *Anikstrodesmis*, *Arachnoidiscus*, *Aster*, *Botryococcus*, *Chaetoceros*,
15 *Chlamydomonas*, *Chlorella*, *Chlorococcum*, *Chorethron*, *Cocconeis*, *Coscinodiscus*,
Cyclotella, *Cylindrotheca*, *Dunaliella*, *Emiliana*, *Euglena*, *Fistulifera*, *Fragilariopsis*,
Gyrosigma, *Hematococcus*, *Isochrysis*, *Lampriscus*, *Monochrysis*, *Monoraphidium*,
Nannochloris, *Nannochloropsis*, *Navicula*, *Neochloris*, *Nephrochloris*, *Nephroselmis*,
Nitzschia, *Nodularia*, *Nostoc*, *Odontella*, *Oochromonas*, *Oocystis*, *Oscillatoria*,
20 *Pavlova*, *Phaeodactylum*, *Playtmonas*, *Pleurochrysis*, *Porphyra*, *Pseudoanabaena*,
Pyramimonas, *Stichococcus*, *Synechococcus*, *Synechocystis*, *Tetraselmis*,
Thalassiosira, and *Trichodesmium*.

The terms "genetically engineered" or "genetically modified" or "recombinant" as used herein with reference to a host cell denote a non-naturally occurring host cell, as well
25 as its recombinant progeny, that has at least one genetic alteration not found in a naturally occurring strain of the referenced species, including wild-type strains of the referenced species. Such genetic modification is typically achieved by technical means (i.e. non-naturally) through human intervention and may include, e.g., the introduction of an exogenous nucleic acid and/or the modification, over-expression, or deletion of
30 an endogenous nucleic acid.

The term "exogenous" or "foreign" as used herein is intended to mean that the referenced molecule, in particular nucleic acid, is not naturally present in the host cell.

The term “endogenous” or “native” as used herein denotes that the referenced molecule, in particular nucleic acid, is present in the host cell.

By “recombinant nucleic acid” when referring to a nucleic acid in a recombinant host cell, is meant that at least part of said nucleic acid is not naturally present in the host
5 cell in the same genomic location. For instance a recombinant nucleic acid can comprise a coding sequence naturally occurring in the host cell under control of an exogenous promoter, or it can be an additional copy of a gene naturally occurring in the host cell, or a recombinant nucleic acid can comprise an exogenous coding sequence under the control of an endogenous promoter.

10 By “nucleic acid” is meant oligomers and polymers of any length composed essentially of nucleotides, e.g., deoxyribonucleotides and/or ribonucleotides. Nucleic acids can comprise purine and/or pyrimidine bases and/or other natural (e.g., xanthine, inosine, hypoxanthine), chemically or biochemically modified (e.g., methylated), non-natural, or derivatised nucleotide bases. The backbone of nucleic acids can comprise sugars and
15 phosphate groups, as can typically be found in RNA or DNA, and/or one or more modified or substituted sugars and/or one or more modified or substituted phosphate groups. Modifications of phosphate groups or sugars may be introduced to improve stability, resistance to enzymatic degradation, or some other useful property. A “nucleic acid” can be for example double-stranded, partly double stranded, or single-stranded.

20 Where single-stranded, the nucleic acid can be the sense strand or the antisense strand. The “nucleic acid” can be circular or linear. The term “nucleic acid” as used herein preferably encompasses DNA and RNA, specifically including genomic, hnRNA, pre-mRNA, mRNA, cDNA, recombinant or synthetic nucleic acids, including vectors.

By “encoding” is meant that a nucleic acid sequence or part(s) thereof corresponds, by
25 virtue of the genetic code of an organism in question, to a particular amino acid sequence, e.g., the amino acid sequence of a desired polypeptide or protein. By means of example, nucleic acids “encoding” a particular polypeptide or protein, e.g. an enzyme, may encompass genomic, hnRNA, pre-mRNA, mRNA, cDNA, recombinant or synthetic nucleic acids.

30 Preferably, a nucleic acid encoding a particular polypeptide or protein may comprise an open reading frame (ORF) encoding said polypeptide or protein. An “open reading frame” or “ORF” refers to a succession of coding nucleotide triplets (codons) starting with a translation initiation codon and closing with a translation termination codon

known *per se*, and not containing any internal in-frame translation termination codon, and potentially capable of encoding a polypeptide or protein. Hence, the term may be synonymous with "coding sequence" as used in the art.

The nucleic acids taught herein may encode more than one polypeptide or protein.
5 Such nucleic acids are denoted as "polycistronic" nucleic acids and typically comprise several ORFs or coding sequences, each encoding a polypeptide or protein.

The terms "polypeptide" and "protein" are used interchangeably herein and generally refer to a polymer of amino acid residues linked by peptide bonds, and are not limited to a minimum length of the product. Thus, peptides, oligopeptides, polypeptides,
10 dimers (hetero- and homo-), multimers (hetero- and homo-), and the like, are included within the definition. Both full-length proteins and fragments thereof are encompassed by the definition. The terms also include post-expression modifications of the polypeptide, for example, glycosylation, acetylation, phosphorylation, etc. Furthermore, for purposes of the present invention, the terms also refer to such when including
15 modifications, such as deletions, additions and substitutions (e.g., conservative in nature), to the sequence of a native protein or polypeptide.

The term "enzyme" as used herein denotes a biological molecule that catalyzes a chemical reaction. The term encompasses single enzymes, i.e. single catalytic entities, as well as systems comprising more than one catalytic entity. The enzymes described
20 herein can naturally possess the recited activity or they can be engineered to exhibit said activity.

As used herein, "fatty acid enzyme" means any enzyme involved in fatty acid biosynthesis. As used herein, the term "fatty acid biosynthetic pathway" means a biosynthetic pathway that produces fatty acids. Fatty acid enzymes can be expressed
25 or over-expressed in a host cell to produce fatty acids.

As used herein, the terms "purify," "purified," or "purification" means the removal or isolation of a molecule from its environment by, for example, isolation or separation. As used herein, these terms also refer to the removal of contaminants from a sample. For example, when α -olefins are produced in a host cell, the olefins can be purified by the
30 removal of other cellular components (e.g., nucleic acids, polypeptides, lipids, carbohydrates, or other hydrocarbons). The terms "purify," "purified," and "purification" do not require absolute purity. They are relative terms.

As used herein, the terms “identity” and “identical” and the like refer to the sequence similarity between two polymeric molecules, e.g., between two nucleic acid molecules or polypeptides. Methods for comparing sequences and determining sequence identity are well known in the art. By means of example, percentage of sequence identity refers to a percentage of identical nucleic acids or amino acids between two sequences after alignment of these sequences. Alignments and percentages of identity can be performed and calculated with various different programs and algorithms known in the art. Preferred alignment algorithms include BLAST (Altschul, 1990; available for instance at the NCBI website) and Clustal (reviewed in Chenna, 2003; available for instance at the EBI website). Preferably, BLAST is used to calculate the percentage of identity between two sequences, such as the “Blast 2 sequences” algorithm described by Tatusova and Madden 1999 (FEMS Microbiol Lett 174: 247-250), for example using the published default settings or other suitable settings (such as, e.g., for the BLASTN algorithm: cost to open a gap = 5, cost to extend a gap = 2, penalty for a mismatch = -2, reward for a match = 1, gap x_dropoff = 50, expectation value = 10.0, word size = 28; or for the BLASTP algorithm: matrix = Blosum62, cost to open a gap = 11, cost to extend a gap = 1, expectation value = 10.0, word size = 3).

The term “renewable” is used herein to refer to a material (e.g. a molecule, a composition or a product) that can be produced or is derivable from a natural resource which is periodically (e.g., annually or perennially) replenished through the actions of plants of terrestrial, aquatic or oceanic ecosystems (e.g., agricultural crops, edible and non-edible grasses, forest products, seaweed, or algae), or microorganisms (e.g., bacteria, fungi, or yeast).

The term “renewable resource” refers to a natural resource that can be replenished within a 100 year time frame. The resource may be replenished naturally, or via agricultural techniques. Renewable resources include, for example but without limitation, plants, animals, fish, bacteria, fungi, yeasts, algae and forestry products. They may be naturally occurring, hybrids, or genetically engineered organisms. Natural resources such as crude oil, coal, and peat which take longer than 100 years to form are not considered to be renewable resources.

The term “bio-based content” refers herein to the amount of carbon from a renewable resource in a material as a percentage of the mass of the total organic carbon in the material, as determined by standard ASTM D6866.

The term “biosourced” with respect to a material (e.g. a molecule, a composition or a product) means that such material is derived from starting materials of renewable origin (i.e. from renewable resources). Accordingly, subject to typical measurement errors, a biosourced material has a bio-based content of at least 90%, preferably at least 95%,
5 more preferably at least about 96%, 97% or 98%, even more preferably at least about 99% such as about 100%.

The present application generally relates to the biosynthesis of olefins, in particular α -olefins, more particularly medium-chain α -olefins.

10 More particularly, the application provides nucleotide sequences encoding medium-chain free fatty acid decarboxylases and the polypeptides encoded thereby, recombinant organisms comprising said nucleotide sequences, methods of production of medium-chain α -olefins using said polypeptides or said recombinant organisms and products obtained by these methods.

15 Nucleotide sequences

The application provides nucleic acids encoding enzymes having decarboxylase activity on medium-chain free fatty acids, which are of interest for the production of the corresponding medium-chain α -olefins. Indeed, the inventors have identified a number of nucleic acid sequences which encode enzymes having decarboxylase activity on
20 medium-chain free fatty acids. In particular embodiments these nucleic acids encode decarboxylase enzymes for which medium-chain free fatty acids are preferred substrates.

Nucleotide sequences encoding enzymes having medium-chain free fatty acid decarboxylase activity, in particular C_8 - C_{12} free fatty acid decarboxylase activity, more
25 particularly C_{12} free fatty acid decarboxylase activity, and accordingly suitable for use in the methods envisaged herein include the sequence encoding the P450_{BS β} fatty acid hydroxylase (Bs168) from *Bacillus subtilis* subsp. *subtilis* str. 168 (SEQ ID NO:3); a genomic sequence of *Alicyclobacillus acidocaldarius* encoding an olefin-producing enzyme (Aa162) (SEQ ID NO:5); and a genomic sequence of *Staphylococcus*
30 *massiliensis* encoding an olefin-producing enzyme (Sm46) (SEQ ID NO:7), as well as variants of these sequences. Variant nucleotide sequences may for instance be codon-optimized sequences for recombinant expression in a host cell of choice. For instance,

nucleotide sequences having at least about 70% or 75%, preferably at least about 80%, more preferably at least about 85%, 90% or 95%, even more preferably at least about 96%, 97%, 98% or 99% sequence identity to SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:7 and encoding enzymes having medium-chain free fatty acid decarboxylase activity, in particular C₈-C₁₂ free fatty acid decarboxylase activity, more particularly C₁₂ free fatty acid decarboxylase activity, are envisaged for use in the methods disclosed herein. An exemplary variant nucleotide sequence of SEQ ID NO:7 is SEQ ID NO:9, which has been codon-optimized for recombinant expression in *E. coli*.

10 In particular embodiments, the decarboxylase enzymes encoded by the nucleic acid sequences envisaged herein have substrate preference for medium-chain free fatty acids, preferably C₈-C₁₂ free fatty acids, more preferably C₁₂ free fatty acids.

The inventors are the first to identify the nucleic acid sequence of *Alicyclobacillus acidocaldarius* (SEQ ID NO:5) and a genomic sequence of *Staphilococcus massiliensis* (SEQ ID NO:7) encoding olefin-producing enzymes (termed Aa162 and Sm46, respectively). More particularly, the nucleic acid sequence of *Alicyclobacillus acidocaldarius* (SEQ ID NO:5) was found to encode a decarboxylase (Aa162, SEQ ID NO:6) acting preferably on medium-chain free fatty acids, more particularly on C₁₂ free fatty acids, which can ensure an increased efficiency of the production of medium-chain α -olefins, more particularly C₁₁ α -olefins, and/or increased purity of said α -olefins. The genomic sequence of *Staphilococcus massiliensis* (SEQ ID NO:7) was found to encode an olefin-producing enzyme (Sm46, SEQ ID NO:8) that has decarboxylase activity on medium-chain free fatty acids. In particular, it was shown that Sm46 shows substrate preference for C₈-C₁₂ free fatty acids, more particularly C₁₂ free fatty acid was found to be the preferred substrate for Sm46. Also advantageously, the Sm46 polypeptide shows specific decarboxylase activity, i.e. the formation of co-products, in particular hydroxy fatty acids such as α - and β -hydroxy fatty acids, when converting free fatty acid substrate is minimal with said polypeptide.

These novel sequences allow the production of medium-chain α -olefins, in particular C₇-C₁₁ α -olefins, from the corresponding fatty acid substrates such as the production of C₁₁ α -olefins from C₁₂ free fatty acids. In addition, these sequences can be used to further identify sequences encoding enzymes having medium-chain free fatty acid

decarboxylase activity, including decarboxylases for which C₁₂ free fatty acids are the preferred substrate.

Variants (or mutants) of sequences identified herein can be naturally occurring or they can be man-made e.g. using genetic engineering techniques. Such techniques are well known in the art and include, for example but without limitation, site directed mutagenesis, random chemical mutagenesis, and standard cloning techniques. Other exemplary techniques for mutagenesis are recombination techniques such as DNA shuffling that use fragments of existing sequences and mix them in novel combinations.

Particularly envisaged herein are variant nucleotide sequences obtained by screening a shuffling library between the nucleotide sequences set forth in SEQ ID NO:1 and SEQ ID NO:7, and a shuffling library between the nucleotide sequences set forth in SEQ ID NO:3 and SEQ ID NO:5 for recombinant nucleotide sequences encoding recombinant polypeptides having a desired biological activity, in particular decarboxylase activity, more particularly (specific) medium-chain free fatty acid decarboxylase activity such as C₈-C₁₂ free fatty acid decarboxylase activity, even more preferably the preferred substrate of these decarboxylases is a C₈-C₁₂ free fatty acid. The technique of DNA shuffling is well known in the art. Reference can be made to Stemmer (1994. Nature 370:389-391) for an exemplary shuffling technique.

The nucleotide sequences envisaged herein encoding enzymes having medium-chain free fatty acid decarboxylase activity are of particular interest for the recombinant production of medium-chain α -olefins, e.g. through the (over)expression of said nucleic acid sequences in a recombinant host cell, as detailed below.

Recombinant host cells

The application provides genetically engineered host cells capable of producing medium-chain α -olefins, wherein said host cells are characterized in that they comprise a recombinant nucleic acid encoding an enzyme having decarboxylase activity on C₈ to C₁₄ free fatty acids, preferably on C₁₂ free fatty acids as described hereinabove. For instance, these recombinant host cells may comprise an exogenous nucleic acid encoding a medium-chain free fatty acid decarboxylase, or they may over-express an endogenous nucleic acid encoding a medium-chain free fatty acid decarboxylase. In particular embodiments, the recombinant host cells comprise a recombinant nucleic acid comprising a nucleotide sequence having at least about 70% or 75%, preferably at

least about 80%, more preferably at least about 85%, 90% or 95%, even more preferably at least about 96%, 97%, 98% or 99% sequence identity to SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO:9 or SEQ ID NO:10 and encoding an enzyme having decarboxylase activity on free fatty acids with a carbon chain length
5 comprised between 8 and 14, preferably on C₈-C₁₂ free fatty acids, more preferably on C₁₂ free fatty acids. In particular embodiments, the recombinant host cells comprise a recombinant nucleic acid comprising a nucleotide sequence having at least about 70% or 75%, preferably at least about 80%, more preferably at least about 85%, 90% or 95%, even more preferably at least about 96%, 97%, 98% or 99% sequence identity to
10 SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO: 3, or SEQ ID NO: 5 and encoding a decarboxylase with preferred activity on C₈- C₁₄ free fatty acids, preferably on C₈- C₁₂ free fatty acids, more preferably on C₁₂ free fatty acids. In particular embodiments, the recombinant host cells comprise a recombinant nucleic acid comprising a nucleotide sequence that encodes a polypeptide comprising an amino
15 acid sequence of SEQ ID NO:8, SEQ ID NO:13, SEQ ID NO:4 or SEQ ID NO:6 or a functional variant, in particular a functional fragment, of said polypeptide.

The genetically engineered host cells disclosed herein may be further genetically modified through expression of one or more recombinant nucleic acids encoding an enzyme involved in a fatty acid biosynthetic pathway. For example, the recombinant
20 host cells may be further modified to overexpress a nucleic acid encoding an enzyme involved in an endogenous fatty acid biosynthetic pathway, or they may be further modified by introduction into the host cell of an exogenous nucleic acid encoding an enzyme involved in fatty acid synthesis. When expressed, the recombinant nucleic acid encoding the enzyme involved in fatty acid synthesis confers to the host cell the ability
25 to produce or overproduce a fatty acid.

Each step within a fatty acid biosynthetic pathway can be modified to produce or overproduce a fatty acid of interest. For example, known genes involved in the synthesis of fatty acids can be expressed or overexpressed in a host cell to produce a desired free fatty acid, or attenuated to inhibit production of a non-desired fatty acid.

30 For instance, production of hydrocarbons starting from acyl-ACP is ensured by thioesterases, whereafter the production of terminal olefins is catalyzed by a decarboxylase. Accordingly, exemplary genes involved in fatty acid synthesis include genes encoding thioesterase.

In a preferred embodiment, the host cells are further genetically engineered to express or overexpress a thioesterase to induce or increase free fatty acid production. The chain length of a fatty acid substrate is controlled by thioesterase, and hence, by (over)expressing a suitable thioesterase, a free fatty acid with desired carbon chain length can be obtained. Non-limiting examples of thioesterases are provided in Table 1. Preferably, medium-chain fatty acid (MCFA)-specific thioesterases are used for (over)expression in the recombinant host cells according to the invention.

Table 1: Thioesterases

GenBank accession number (UniProtKB/Swiss-Prot)	Source organism	Gene	Preferential product produced
AAC73596	<i>E. coli</i>	<i>tesA</i> without leader sequence	C _{18:1}
AAC73555	<i>E. coli</i>	<i>tesB</i>	
AAA34215 (Q41635); AAC49001	<i>Umbellularia Californica</i>	<i>fatB</i>	C _{12:0}
AAC49269 (Q39513)	<i>Cuphea hookeriana</i>	<i>fatB2</i>	C _{8:0} – C _{10:0}
AAC49269; AAC72881	<i>Cuphea hookeriana</i>	<i>fatB3</i>	C _{14:0} – C _{16:0}
AAC49151 (Q39473)	<i>Cinnamomum camphorum</i>	<i>fatB</i>	C _{14:0}
CAA85388	<i>Arabidopsis thaliana</i>	<i>fatB</i>	C _{16:1}
NP_189147; NP_193041	<i>Arabidopsis thaliana</i>	<i>fatA</i>	C _{18:1}
CAC39106	<i>Bradyrhizobium japonicum</i>	<i>fatA</i>	C _{18:1}

AAC72883	<i>Cuphea hookeriana</i>	<i>fatA</i>	C _{18:1}
AAL79361	<i>Helianthus annuus</i>	<i>fatA1</i>	
JF338905	<i>Cocos nucifera</i>		C _{12:0}

The thioesterase may be a thioesterase that is naturally present in higher plants. Two families of acyl-ACP thioesterases are present in higher plants: the “Class I” acyl-ACP thioesterases encoded by *FatA* genes, which are responsible for cleaving long-chain (for example, C₁₆ and C₁₈) unsaturated fatty acids from acyl-ACP, and the “Class II” acyl-ACP thioesterases encoded by *FatB* genes, which are active on saturated fatty acyl chains, and which can be specific for medium-chain (C₈-C₁₄) acyl-ACPs or which can be active on both medium- and long-chain fatty acyl-ACPs. Non-limiting examples of thioesterases which are MCFA-specific and naturally present in plants, and hence suitable for (over)expression in the recombinant host cells described herein, are thioesterases encoded by *FatB* genes, or the thioesterases described for instance in Voelker et al. (1992 Science 257:72-74) and Jing et al. (2011 Biochemistry 12:44). The thioesterase may also be an engineered thioesterase as described for instance in Voelker et al. (1994 Journal of Bacteriology 176:7320-7327).

Depending on the α -olefin of interest, the expression of thioesterase enzymes may be either induced (by introduction into the host cell of an exogenous nucleic acid encoding said enzyme), stimulated (by overexpression of an endogenous gene encoding said enzyme) or attenuated (by modification of an endogenous gene encoding said enzyme).

In some situations, C₁₂ free fatty acids can be produced by expressing or overexpressing thioesterases that use C₁₂-ACP (for example, accession numbers Q41635 and JF338905) and attenuating thioesterases that produce non-C₁₂ fatty acids. In other instances, C₁₄ free fatty acids can be produced by attenuating endogenous thioesterases that produce non-C₁₄ fatty acids and (over)expressing the thioesterases that use C₁₄- ACP (for example, accession number Q39473).

Acetyl-CoA, malonyl-CoA, and free fatty acid overproduction can be verified using methods known in the art, for example, by using radioactive precursors, HPLC, and GC-MS subsequent to cell lysis.

For the production of C₁₁ α -olefins, the host cells are preferably modified by the introduction of an exogenous nucleic acid encoding a thioesterase having preferential hydrolase activity towards C₁₂ acyl-ACP substrate such as Q41635 or JF338905 and/or upregulating endogenous genes encoding a thioesterase having preferential hydrolase activity towards C₁₂ acyl-ACP, and optionally downregulating endogenous genes encoding thioesterases that produce non-C₁₂ fatty acids.

The medium-chain free fatty acids produced by the fatty acid enzymes in the host cells envisaged herein are the substrates of the decarboxylase enzymes described herein.

Accordingly, particularly preferred host cells for the production of medium-chain α -olefins are recombinant host cells comprising:

- a recombinant nucleic acid encoding a thioesterase having preferential hydrolase activity towards medium-chain acyl-ACP substrate; and
- a recombinant nucleic acid encoding a medium-chain free fatty acid decarboxylase.

Particularly preferred host cells for the production of C₁₁ α -olefins are recombinant host cells comprising:

- a recombinant nucleic acid encoding a thioesterase having preferential hydrolase activity towards C₁₂ acyl-ACP substrate; and
- a recombinant nucleic acid encoding a C₁₂ free fatty acid decarboxylase.

In particular embodiments, the free fatty acid decarboxylase is an enzyme having at least about 70%, preferably at least about 80%, more preferably at least about 85%, 90% or 95%, even more preferably at least about 96%, 97%, 98% or 99% sequence identity to SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8 or SEQ ID NO:13 and having medium-chain free fatty acid decarboxylase activity, in particular C₈-C₁₂ free fatty acid decarboxylase activity. In particular embodiments, the fatty acid decarboxylase is Sm46 (SEQ ID NO:8), Sm46-del29 (SEQ ID NO:13), Bs168 (SEQ ID NO:4) or Aa162 (SEQ ID NO:6) or a functional variant thereof since specific production of medium-chain α -olefins, in particular C₇-C₁₁ α -olefins, more particularly C₁₁ α -olefins, was shown herein for these enzymes. Accordingly, in particular embodiments, the application provides a genetically engineered host cell comprising a recombinant nucleic acid comprising a nucleotide sequence having at least about 70% or 75%, preferably at least about 80%, more preferably at least about 85%, 90% or 95%, even

more preferably at least about 96%, 97%, 98% or 99% sequence identity to, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9 or SEQ ID NO:10 and encoding an enzyme having medium-chain free fatty acid decarboxylase activity, in particular C₈-C₁₂ free fatty acid decarboxylase activity. In further particular embodiments, the recombinant nucleic acid comprises a nucleotide sequence having at least about 70% or 75%, preferably at least about 80%, more preferably at least about 85%, 90% or 95%, even more preferably at least about 96%, 97%, 98% or 99% sequence identity to SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO: 3 or SEQ ID NO: 5 and encodes an enzyme having decarboxylase activity for which C₈-C₁₂ free fatty acids, more particularly C₁₁ free fatty acids, are the preferred substrate. In preferred embodiments, the host cell comprises a recombinant nucleic acid encoding a polypeptide having an amino acid sequence comprising SEQ ID NO:8, SEQ ID NO:13, SEQ ID NO:4 or SEQ ID NO:6 or a functional variant thereof. In further embodiments, the genetically engineered host cell further comprises a recombinant nucleic acid encoding a thioesterase having preferential hydrolase activity towards medium-chain acyl-ACP substrate, preferably towards C₈-C₁₂ acyl-ACP substrate, more preferably towards C₁₂ acyl-ACP substrate.

In particular embodiments, where the above nucleotide sequences are endogenously expressed by the host cell of the invention, it is envisaged that expression of these sequences can be increased specifically so as to ensure commercially relevant medium-chain α -olefin production.

More particularly, in particular embodiments, the host cell is selected to have a high endogenous thioesterase activity. Methods of selecting cells having particular properties are known in the art.

Any cell that can be suitably transformed with and/or genetically engineered to ensure (over)expression of one or more of the described recombinant nucleic acids can be used in the context of the present invention. The host cells disclosed herein can be any prokaryotic or eukaryotic organism or cell. Non-limiting examples of host cells include plant cells, bacterial cells, yeast cells, fungal cells, and algal cells. In embodiments, the host cells are genetically engineered bacteria, or genetically engineered fungi, in particular yeasts, genetically engineered algae, or genetically engineered plant cells.

Preferably, the host cells are oleaginous host cells. For example, the host cell may be an oleaginous bacterium, an oleaginous fungus, oleaginous yeast or an oleaginous

alga. Non-limiting examples of oleaginous yeasts include *Lipomyces starkeyi*, *Rhodospiridium toruloides*, *Rhodotorula glutinis*, and *Yarrowia lipolytica*. Non-limiting examples of oleaginous algae genera include *Botryococcus*, *Chaetoceros*, *Chlorella*, *Chlorococcum*, *Cylindrotheca*, *Dunaliella*, *Fistulifera*, *Isochrysis*, *Nannochloropsis*,
5 *Neochloris*, *Nitzschia*, *Pavlova*, *Scenedesmus*, *Skeletonema*, *Stichococcus* and *Tetraselmis*.

Thus, the genetically engineered host cells disclosed herein comprise a recombinant nucleic acid encoding a medium-chain free fatty acid decarboxylase disclosed herein, and optionally one or more recombinant nucleic acids encoding a fatty acid enzyme,
10 i.e. an enzyme involved in fatty acid synthesis. Additionally or alternatively the expression of one or more genes encoding enzymes involved in the production of fatty acids other than medium-chain free fatty acids may be suppressed, decreased or limited.

The methods for generating the genetically engineered host cells described herein
15 involve standard genetic modifications, for which well-established methods are available to the skilled person.

Genetic engineering of the host cells to contain a recombinant nucleic acid encoding a polypeptide or a fatty acid enzyme as described herein is accomplished in one or more steps via the design and construction of appropriate vectors and transformation of the
20 host cells with those vectors.

Electroporation and/or chemical (such as calcium chloride- or lithium acetate-based) transformation methods or *Agrobacterium tumefaciens*-mediated transformation methods as known in the art can be used.

The recombinant nucleic acid(s) encoding a decarboxylase enzyme and/or a fatty acid
25 enzyme that may be provided in the host cells in the context of the present invention typically comprises a coding sequence encoding the decarboxylase enzyme and/or the fatty acid enzyme placed under the transcriptional control of one or more promoters and one or more terminators, both of which are functional in the host cell.

Promoter and terminator sequences may be native to the host cell or exogenous to the
30 host cell. Useful promoter and terminator sequences include those that are highly identical (i.e. having an identities score of 90% or more, preferably 95% or more, most preferably 99% or more) in their functional portions compared to the functional portions

of promoter and terminator sequences, respectively, that are native to the host cell, particularly when the insertion of the recombinant nucleic acid is targeted at a specific site in the host genome. The use of native (to the host) promoters and terminators, together with their respective upstream and downstream flanking regions, can permit
5 the targeted integration of the recombinant nucleic acid into specific loci of the host genome.

Additionally or alternatively, the coding sequence may be native to the host cell or exogenous to the host cell.

Numerous vectors are known to practitioners skilled in the art, and selection of an
10 appropriate vector is a matter of choice. The vectors can either be cut with particular restriction enzymes or used as circular DNA.

The vectors taught herein preferably comprise (a combination of) a recombinant nucleic acid as described herein. In particular, a vector comprises (a combination of) the coding sequence of a polypeptide or a fatty acid enzyme as described herein and
15 associated promoter and terminator sequences. The vector may contain restriction sites of various types for linearization or fragmentation. Vectors may further contain a backbone portion (such as for propagation in *E. coli*) many of which are conveniently obtained from commercially available yeast or bacterial vectors. The vector preferably comprises one or more selection marker gene cassettes. A "selection marker gene" is
20 one that encodes a protein needed for the survival and/or growth of the transformed cell in a selective culture medium. Typical selection marker genes encode proteins that (a) confer resistance to antibiotics or other toxins such as chloramphenicol, zeocin (sh *ble* gene from *Streptoalloteichus hindustanus*), genetecin, melibiase (MEL5), hygromycin (aminoglycoside antibiotic resistance gene from *E. coli*), ampicillin,
25 tetracycline, or kanamycin (kanamycin resistance gene of Tn903), (b) complement auxotrophic deficiencies of the cell. Two prominent examples of auxotrophic deficiencies are the amino acid leucine deficiency (e.g. LEU2 gene) or uracil deficiency (e.g. URA3 gene). Cells that are orotidine-5'-phosphate decarboxylase negative (*ura3*-) cannot grow on media lacking uracil. Thus a functional URA3 gene can be used as a
30 marker on a cell having a uracil deficiency, and successful transformants can be selected on a medium lacking uracil. Only cells transformed with the functional URA3 gene are able to synthesize uracil and grow on such medium. If the wild-type strain does not have a uracil deficiency (as is the case with *I. orientalis*, for example), an

auxotrophic mutant having the deficiency must be made in order to use URA3 as a selection marker for the strain. Methods for accomplishing this are well known in the art. The selection marker cassette typically further includes a promoter and terminator sequence, operatively linked to the selection marker gene, and which are operable in the host.

Successful transformants can be selected for in known manner, by taking advantage of the attributes contributed by the marker gene, or by other characteristics (such as ability to produce α -olefins) contributed by the inserted recombinant nucleic acids. Screening can also be performed by PCR or Southern analysis to confirm that the desired insertions, and optionally deletions have taken place, to confirm copy number and to identify the point of integration of coding sequences into the host genome. Activity (such as α -olefin-producing activity) of the polypeptide encoded by the inserted coding sequence can be confirmed using known assay methods as described elsewhere herein.

Also disclosed herein are methods for obtaining a genetically engineered host cell capable of producing an α -olefin of interest as described herein, which method may comprise transforming a host cell with a recombinant nucleic acid encoding an enzyme having medium-chain free fatty acid decarboxylase activity as taught herein, more particularly a decarboxylase generating the α -olefin of interest and optionally one or more recombinant nucleic acids encoding a fatty acid enzyme as taught herein, more particularly a fatty acid enzyme involved in the synthesis of the substrate for the free fatty acid decarboxylase enzyme capable of generating the α -olefin of interest. In particular, the method may comprise the steps of:

- a) transforming a host cell with a recombinant nucleic acid encoding an enzyme having medium-chain free fatty acid decarboxylase activity as taught herein and optionally one or more recombinant nucleic acids encoding a fatty acid enzyme as taught herein; and
- b) selecting a host cell capable of producing an α -olefin of interest.

In particular embodiments, the method further comprises modifying said host cell so as to reduce the endogenous production of olefins other than the α -olefin of interest.

As detailed above, different genetic modifications are envisaged herein which induce medium-chain α -olefin production and/or increase the yield of (one or more particular)

medium-chain α -olefins in a host cell. Accordingly, the present invention also relates to the use of the genetically engineered host cells as described herein for the production of α -olefins, more particularly medium-chain α -olefins.

Producing α -olefins using recombinant host cells

5 In a further aspect, the invention provides methods for the production of medium-chain α -olefins, more particularly C_{11} α -olefins, which method comprises providing a genetically engineered host cell as described above and culturing said genetically engineered host cell in a culture medium so as to allow the production of medium-chain α -olefins. More particularly, the host cell is cultured under conditions suitable to ensure
10 expression or overexpression of the enzyme having medium-chain free fatty acid decarboxylase activity envisaged herein and optionally one or more fatty acid enzyme(s) involved in the synthesis of the substrate of the enzyme having medium-chain free fatty acid decarboxylase activity.

In particular embodiments, the host cells ensure a rate of α -olefin production, more
15 particularly medium-chain α -olefin production, which is sufficiently high to be industrially valuable. Indeed, the recombinant host cells disclosed herein may be capable of ensuring a high yield at limited production costs. Furthermore, they are capable of producing α -olefins of desired carbon chain length. Indeed, the decarboxylase enzymes envisaged herein preferably have substrate preference for a
20 medium-chain free fatty acid, in particular a C_8 - C_{12} free fatty acid, more particularly C_{12} free fatty acid. Also advantageously, the production of unwanted co-products such as hydroxyl fatty acids is minimal. Indeed, the polypeptides envisaged herein preferably have specific decarboxylase activity.

The recombinant host cells are cultured under conditions suitable for the production of
25 medium-chain α -olefins by the host cells. More particularly this implies "conditions sufficient to allow (over)expression" of the recombinant nucleic acid encoding a decarboxylase enzyme, which means any condition that allows a host cell to (over)produce a medium-chain free fatty acid decarboxylase or a fatty acid enzyme as described herein. Suitable conditions include, for example, fermentation conditions.
30 Fermentation conditions can comprise many parameters, such as temperature ranges, levels of aeration, and media composition. Each of these conditions, individually and in combination, allows the host cell to grow. To determine if conditions are sufficient to

allow (over)expression, a host cell can be cultured, for example, for about 4, 8, 12, 18, 24, 36, or 48 hours. During and/or after culturing, samples can be obtained and analyzed to determine if the conditions allow (over)expression. For example, the host cells in the sample or the culture medium in which the host cells were grown can be tested for the presence of a desired product. When testing for the presence of a desired product, assays, such as, but not limited to, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), TLC, HPLC, GC/FID, GC/MS, LC/MS, MS, can be used.

Exemplary culture media include broths or gels. Micro-organisms are typically grown in a culture medium comprising a carbon source to be used for growth of the micro-organism. Exemplary carbon sources include carbohydrates, such as glucose, fructose, cellulose, or the like, that can be directly metabolized by a micro-organism. In addition, enzymes can be added to the culture medium to facilitate the mobilization (e.g., the depolymerization of starch or cellulose to fermentable sugars) and subsequent metabolism of the carbon source. A culture medium may optionally contain further nutrients as required by the particular strain, including inorganic nitrogen sources such as ammonia or ammonium salts, and the like, and minerals and the like.

Other growth conditions, such as temperature, cell density, and the like are generally selected to provide an economical process. Temperatures during each of the growth phase and the production phase may range from above the freezing temperature of the medium to about 50°C.

The culturing step of the methods of the invention may be conducted aerobically, anaerobically, or substantially anaerobically. Briefly, anaerobic conditions refer to an environment devoid of oxygen. Substantially anaerobic conditions include, for example, a culture, batch fermentation or continuous fermentation such that the dissolved oxygen concentration in the medium remains between 0 and 10% of saturation. Substantially anaerobic conditions also includes growing or resting cells in liquid medium or on solid agar inside a sealed chamber maintained with an atmosphere of less than 1% oxygen. The percent of oxygen can be maintained by, for example, sparging the culture with an N₂/CO₂ mixture or other suitable non-oxygen gas or gasses.

The cultivation step of the methods described herein can be conducted continuously, batch-wise, or some combination thereof.

In particular embodiments, wherein phototrophic algae are used as host cells, the method for the production of medium-chain α -olefins may comprise providing algae genetically engineered to (over)produce medium-chain α -olefins as taught herein, and culturing said algae in photobioreactors or an open pond system using CO₂ and sunlight as feedstock.

In certain embodiments, the conditions suitable for the production of medium-chain α -olefins may further imply cultivating the host cells in a culture medium which comprises at least one fatty acid substrate, which is converted into an α -olefin by the decarboxylase encoded by the recombinant nucleic acid comprised in the host cell.

Preferably, the fatty acid substrate is a saturated free fatty acid substrate. Also preferably, the fatty acid substrate is a straight chain free fatty acid substrate. Also preferably, the fatty acid substrate is an even-numbered C₈-C₁₄ free fatty acid substrate (i.e. a free C₈, C₁₀, C₁₂, or C₁₄ free fatty acid substrate or any combination thereof), preferably an even-numbered C₈-C₁₂ free fatty acid substrate, more preferably a C₁₂ free fatty acid substrate.

Particularly intended herein is the production of medium-chain α -olefins, more particularly C₁₁ α -olefins. Medium-chain α -olefins, more particularly C₁₁ α -olefins can be obtained using a recombinant host cell described herein specifically modified for the production of medium-chain α -olefins, more particularly C₁₁ α -olefins.

In further embodiments, methods are provided for producing medium-chain α -olefins, more particularly C₁₁ α -olefins, which, in addition to the steps detailed above, further comprise the step of recovering the α -olefins from the host cell or the culture medium. Suitable purification can be carried out by methods known to the person skilled in the art such as by using lysis methods, extraction, ion exchange resins, electro dialysis, nanofiltration, etc.

Accordingly, methods are provided for the production of medium-chain α -olefins, more particularly C₁₁ α -olefins, which methods comprise the steps of:

- (i) providing a genetically engineered host cell as described herein;
- (ii) culturing the host cells in a culture medium under conditions suitable for the production of medium-chain α -olefins, and
- (iii) recovering the α -olefins from the host cell or the culture medium.

In particular embodiments, wherein oleaginous yeasts are used as host cells, the method for the production of medium-chain α -olefins may comprise the following steps:

(i) providing oleaginous yeasts genetically engineered to (over)produce medium-chain α -olefins as taught herein;

5 (ii) culturing said oleaginous yeasts in fermenters; and

(iii) recovering the α -olefins from the oleaginous yeast or the culture medium.

In particular embodiments, wherein phototrophic algae are used as host cells, the method for the production of medium-chain α -olefins may comprise the following steps:

10 (i) providing algae genetically engineered to (over)produce medium-chain α -olefins as taught herein;

(ii) culturing said algae in photobioreactors or an open pond system using CO₂ and sunlight as feedstock; and

(iii) recovering the α -olefins from the algae or the culture medium.

15 In particular embodiments the host cells are cultivated under conditions which allow secretion of α -olefins into the environment.

Typically, in the methods for the production of medium chain α -olefins envisaged herein, the decarboxylase expressed by the host cell is not secreted by said host cell and the α -olefin is produced inside the host cell. However, in particular embodiments or for particular applications, it is of interest to ensure secretion of the decarboxylase by the host cells provided herein. This can be of interest where the enzyme is envisaged to be active upon secretion into its environment. A secretion signal sequence can be operably linked to the nucleic acid encoding the free fatty acid decarboxylase to this end. In this connection, "operably linked" denotes that the sequence encoding the secretion signal peptide and the sequence encoding the polypeptide to be secreted are connected in frame or in phase, such that upon expression the signal peptide facilitates the secretion of the polypeptide so-linked thereto.

20
25

Polypeptides

Also disclosed herein are enzymes having medium-chain free fatty acid decarboxylase activity. In particular embodiments, these polypeptides are encoded by nucleic acids and nucleotide sequences described herein. Accordingly, disclosed herein are polypeptides encoded by a nucleic acid comprising a nucleotide sequence having at least about 70% or 75%, preferably at least about 80%, more preferably at least about

30

85%, 90% or 95%, even more preferably at least about 96%, 97%, 98% or 99% sequence identity to SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO:9 or SEQ ID NO:10 having medium-chain free fatty acid decarboxylase activity, in particular C₈-C₁₂ free fatty acid decarboxylase activity, preferably C₁₂ free fatty acid decarboxylase activity. Indeed, in particular embodiments, the enzymes having medium-chain free fatty acid decarboxylase activity, C₈-C₁₂ free fatty acid decarboxylase activity, preferably C₁₂ free fatty acid decarboxylase activity can be obtained by recombinant expression of said nucleic acids in a host cell. Methods for recombinant production of polypeptides are known in the art.

10 In particular embodiments, the polypeptides which are enzymes having medium-chain free fatty acid decarboxylase activity, in particular C₈-C₁₂ free fatty acid decarboxylase activity, more particularly C₁₂ free fatty acid decarboxylase activity, include polypeptides having an amino acid sequence comprising, consisting essentially of or consisting of, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8 or SEQ ID NO:13, and
15 variants, including active fragments, of these polypeptides having medium-chain free fatty acid decarboxylase activity, in particular C₈-C₁₂ free fatty acid decarboxylase activity, more particularly C₁₂ free fatty acid decarboxylase activity. Preferred polypeptides are those consisting of the amino acid sequence of SEQ ID NO:4 (i.e. Bs168), SEQ ID NO:6 (i.e. Aa162), SEQ ID NO:8 (i.e. Sm46), or SEQ ID NO:13 (i.e.
20 Sm46-del29) and functional variants of these polypeptides. Indeed, SEQ ID NO:8, and its active fragment of SEQ ID NO:13, are novel polypeptides provided herein having medium-chain free fatty acid decarboxylase activity, in particular C₈-C₁₂ free fatty acid decarboxylase activity, more particularly C₁₂ free fatty acid decarboxylase activity.

In particular embodiments, the polypeptides have decarboxylase activity with a preferred activity on medium-chain free fatty acids, in particular on a C₈-C₁₂ free fatty acid, more particularly on C₁₂ free fatty acids, i.e. the preferred substrate for these decarboxylase enzymes are medium-chain free fatty acids, in particular C₈-C₁₂ free fatty acids, more particularly C₁₂ free fatty acids. In particular embodiments, these enzymes have at least about 80%, 85%, 90% or 95%, even more preferably at least
25 about 96%, 97%, 98% or 99% sequence identity to SEQ ID NO:8, SEQ ID NO:13, SEQ ID NO:4 or to SEQ ID NO: 6. Preferred enzymes comprise or consist of SEQ ID NO:8, SEQ ID NO:13, SEQ ID NO:4 or SEQ ID NO: 6 and functional variants thereof such as but not limited to an active fragment or variant thereof which maintains the activity of
30

the enzyme comprising or consisting of SEQ ID NO:8, SEQ ID NO:13, SEQ ID NO:4 or SEQ ID NO: 6.

As used herein in connection to a decarboxylase enzyme, a "preferred substrate" refers to the free fatty acid for which the decarboxylase enzyme has the highest activity, i.e. when reacting with a number of free fatty acids, the decarboxylase enzyme has the highest activity when the substrate is the preferred free fatty acid substrate. The activity of a decarboxylase enzyme can be determined by measuring the concentration of the corresponding α -olefin product when the decarboxylase acid is reacted with a free fatty acid substrate. The α -olefin concentration in a reaction or culture medium can be measured by methods well known in the art, such as by GC/MS analysis. The substrate preference for a decarboxylase enzyme can hence be determined by calculating the conversion ratio of each free fatty acid substrate tested into corresponding α -olefin product, wherein the preferred substrate is the free fatty acid substrate with the highest conversion ratio. The substrate preference for a decarboxylase enzyme can also be determined by calculating the conversion ratio of each free fatty acid substrate tested, wherein the preferred substrate is the free fatty acid substrate with the highest conversion ratio. The "conversion ratio" or "substrate conversion ratio" means herein the ratio between consumed substrate (for instance fatty acid substrate concentration (e.g. in mM)) and produced reaction product(s) (for instance α -olefin concentration (e.g. in mM) and also potential co-product(s) is quantified (for instance hydroxyl fatty acid concentration (e.g. in mM)). The conversion ratio or conversion percentage for a specific reaction product can be calculated as the ratio between consumed substrate (for instance fatty acid substrate concentration (e.g. in mM)) and the specific reaction product produced (for instance α -olefin concentration (e.g. in mM)).

Also envisaged herein are variant polypeptides of the polypeptides described herein. It is understood that the variant polypeptides described herein may have conservative or non-essential amino acid substitutions, which do not have a substantial effect on the polypeptide functions. Whether or not a particular substitution will be tolerated (i.e., will not adversely affect desired biological properties) can be determined as described in Bowie et al. (1990) (Science 247:1306 1310). A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have

been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine), and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine).

Additional polypeptide variants are those in which additional amino acids are fused to the polypeptide, such as a secretion signal sequence, or a sequence which facilitates purification of the polypeptide.

Yet other polypeptide variants include functional or active fragments comprising at least about 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100 or 150 consecutive amino acids, and which retain the same biological function as Bs168, Aa162, or Sm46 (e.g. retain olefin-producing activity, more particularly improved decarboxylase activity towards medium-chain free fatty acids, more preferably on C₁₂ FFA substrate). Exemplary functional or active fragments include without limitation truncated forms of the polypeptides described herein, which retain the decarboxylase activity of Bs168, Aa162, or Sm46. These functional or active fragments hence retain at least the decarboxylase catalytic domain of the polypeptide, i.e. the part of the polypeptide that is involved in the decarboxylase reaction. A particular example of such truncated form is the polypeptide having an amino acid sequence of SEQ ID NO:13 (encoded by the nucleotide sequence of SEQ ID NO:10), which polypeptide is a functional fragment of Sm46 wherein the 29 N-terminal amino acids were deleted.

Functional variants of the polypeptides described herein retain the decarboxylase activity of the polypeptides. Accordingly, the functional variants may comprise or consist of an amino acid sequence having at least about 70%, preferably at least about 80%, more preferably at least about 85%, 90% or 95%, even more preferably at least about 96%, 97%, 98% or 99% sequence identity in the decarboxylase catalytic domain of SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:13. The polypeptide variants may have an amino acid sequence substantially identical to SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:13 or they may have an amino acid sequence having at least about 70%, preferably at least about 80%, more preferably at least about 85%, 90% or 95%, even more preferably at least about 96%, 97%, 98% or

99% sequence identity to SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:13. The polypeptide variants, in particular the active or functional fragments, may have an amino acid sequence wherein the 40 or less, preferably 35 or less, more preferably 30 or less such as 29 or less, N-terminal amino acids of SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8 or SEQ ID NO: 13 have been deleted.

The polypeptides disclosed herein have free fatty acid decarboxylase activity, in particular decarboxylase activity on C₈-C₁₄ free fatty acids, more particularly on even-numbered C₈-C₁₄ free fatty acids such as on C₈, C₁₀, C₁₂ and/or C₁₄ free fatty acids, even more particularly C₈-C₁₂ free fatty acid decarboxylase activity such as C₁₂ free fatty acid decarboxylase activity. Indeed, as shown in the experimental section, decarboxylase activity has been shown for OleT_{JE}, Bs168, Aa162, Sm46 or Sm46-del29, more particularly decarboxylase activity on C₁₂, C₁₄ and C₁₆ free fatty acids. The Sm46, Sm46-del29, Aa162 and Bs168 enzymes showed higher activity on C₁₂ free fatty acid, resulting in higher production of C₁₁ α-olefins.

Decarboxylase activity of a polypeptide can be assayed using routine methods. For example, the polypeptide can be contacted with a substrate, in particular a free fatty acid substrate, under conditions that allow the polypeptide to function. A decrease in the level of the substrate or an increase in the level of an α-olefin can be measured to determine decarboxylase or olefin-producing activity.

The polypeptides described herein may further catalyze the hydroxylation of fatty acids, in particular the α- and β-hydroxylation of fatty acids, as side reactions. Preferred polypeptides are those for which the decarboxylation activity is the dominant activity, i.e. which have specific decarboxylase activity. In preferred embodiments, the conversion of a free fatty substrate into hydroxy fatty acid, in particular α- and β-hydroxy fatty acid by the polypeptide is only marginal (such as less than 25%, preferably less than 20%, more preferably less than 10% of total products) as compared to the conversion of the fatty acid substrate into the corresponding α-olefin. In embodiments, the conversion percentage for hydroxy fatty acids is less than 25%, preferably less than 20%, more preferably less than 10%. The conversion percentage for hydroxy fatty acids can be calculated as the ratio between consumed substrate (for instance fatty acid substrate concentration (e.g. in mM)) and produced hydroxyl fatty acids (for instance hydroxyl fatty acid concentration (e.g. in mM)). Particular examples of polypeptides with specific decarboxylase activity are Sm46 or functional variants

thereof, including active fragments thereof, such as the truncated form of Sm46 wherein the 29 N-terminal amino acids were deleted (i.e. the polypeptide encoded by the nucleotide sequence of SEQ ID NO:10 or the polypeptide of SEQ ID NO:13).

5 The polypeptides envisaged herein can be produced by recombinant expression in a host cell. In particular embodiments, the polypeptide is secreted by the host cell.

Production of α -olefins

Also provided herein is the use of the medium-chain free fatty acid decarboxylase enzymes disclosed herein for the production of medium-chain α -olefins, more particularly C₁₁ α -olefins.

10 Some methods described herein relate to the production of medium-chain α -olefins, more particularly C₁₁ α -olefins, using a (purified) medium-chain free fatty acid decarboxylase enzyme disclosed herein and a free fatty acid substrate. Accordingly, disclosed herein is a method for the production of medium-chain α -olefins, more particularly C₁₁ α -olefins, comprising contacting a medium-chain free fatty acid
15 decarboxylase enzyme with a suitable free fatty acid substrate so as to produce medium-chain α -olefins, more particularly C₁₁ α -olefins.

In particular embodiments, the medium-chain free fatty acid decarboxylase is a polypeptide having an amino acid sequence of SEQ ID NO:4, SEQ ID NO:6 SEQ ID NO:8 or SEQ ID NO:13 or a functional variant, including a functional or active fragment,
20 of said polypeptide. In particular embodiments, the medium-chain free fatty acid decarboxylase is a polypeptide having an amino acid sequence having at least about 70%, preferably at least about 80%, more preferably at least about 85%, 90% or 95%, even more preferably at least about 96%, 97%, 98% or 99% sequence identity to SEQ ID NO:4, SEQ ID NO:6 SEQ ID NO:8 or SEQ ID NO:13. In particular embodiments, the
25 enzyme has a preferred activity on medium-chain free fatty acids, more preferably on C₈-C₁₂ free fatty acids. Most particularly, the enzyme is a decarboxylase having preferred activity on C₁₂ free fatty acids and having an amino acid sequence having at least about 85%, 90% or 95%, even more preferably at least about 96%, 97%, 98% or 99% sequence identity to SEQ ID NO:8, SEQ ID NO:13, SEQ ID NO: 4 or SEQ ID NO:
30 6. In particular embodiments, the enzyme has 95% amino acid sequence identity to SEQ ID NO:8, SEQ ID NO:13, SEQ ID NO: 4 or SEQ ID NO: 6.

For example, a host cell can be genetically engineered to overexpress a medium-chain free fatty acid decarboxylase as disclosed herein. The recombinant host cell can be cultured under conditions sufficient to allow (over)expression of the decarboxylase. Cell-free extracts can then be generated using known methods. For example, the host
5 cells can be lysed using detergents or by sonication. The overexpressed polypeptides can be purified using known methods, or the cell-free extracts can be used as such for the production of olefins. The host cells can also be genetically engineered to overexpress a medium-chain free fatty acid decarboxylase as disclosed herein and to secrete said polypeptide into the growth medium as described elsewhere. The secreted
10 polypeptides can then be separated from the growth medium and optionally purified using known methods without the need for obtaining cell-free extracts.

Next, free fatty acid substrates can be added to the cell-free extracts or (purified) enzymes and maintained under conditions to allow conversion of the free fatty acid substrates to α -olefins. The α -olefins can then be separated and purified using known
15 techniques.

Olefins having particular branching patterns, levels of saturation, and carbon chain length can be produced from free fatty acid substrates having those particular characteristics using the described methods. For example, the fatty acid substrate may be an unsaturated free fatty acid substrate (e.g. a monounsaturated free fatty acid
20 substrate), or a saturated free fatty acid substrate. The fatty acid substrate may be a straight chain free fatty acid substrate, a branched chain free fatty acid substrate, or a free fatty acid substrate that includes a cyclic moiety.

Preferably, the fatty acid substrate is a saturated free fatty acid substrate. Also preferably, the fatty acid substrate is a straight chain free fatty acid substrate. Also
25 preferably, the fatty acid substrate is an even-numbered C₈-C₁₄ free fatty acid substrate (i.e. a C₈, C₁₀, C₁₂ and/or C₁₄ free fatty acid substrate), more preferably an even-numbered C₈-C₁₂ free fatty acid substrate, most preferably a C₁₂ fatty acid substrate.

Particularly intended herein is the production of C₁₁ α -olefins. C₁₁ α -olefins can be obtained from a C₁₂ free fatty acid substrate using a C₁₂ free fatty acid decarboxylase
30 enzyme described herein. As shown in the experimental section, Sm46, Sm46-del29, Bs168 and Aa162 specifically show decarboxylase activity on C₁₂ free fatty acid substrates, i.e. C₁₂ free fatty acid substrate is the preferred substrate for α -olefin

production for Sm46, Sm46-del29, Bs168 and Aa162. Hence, in particular embodiments, the present invention relates to a method for the production of C₁₁ α -olefins, which method comprises contacting a C₁₂ free fatty acid decarboxylase encoded by a nucleic acid comprising a nucleotide sequence having at least about 5 70% or 75%, preferably at least about 80%, more preferably at least about 85%, 90% or 95%, even more preferably at least about 96%, 97%, 98% or 99% sequence identity to SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:3 or SEQ ID NO:5. In embodiments, the method comprises contacting a a polypeptide having an amino acid sequence comprising or consisting of SEQ ID NO:8, SEQ ID NO:13, SEQ ID NO:4 or 10 SEQ ID NO:6 or a functional variant of said polypeptide, preferably a polypeptide consisting of SEQ ID NO:8, SEQ ID NO:13, SEQ ID NO:4 or SEQ ID NO:6, with a C₁₂ free fatty acid substrate, preferably dodecanoic acid (or lauric acid).

Also provided herein are medium-chain α -olefins, more particularly the C₁₁ α -olefins, obtainable by the methods disclosed herein.

15 Medium-chain α -olefins

Also disclosed herein are the medium-chain α -olefins, and compositions comprising medium-chain α -olefins more particularly the C₁₁ α -olefins, produced by the methods as herein described. The methods described herein advantageously result in the production of homogenous α -olefins, wherein the α -olefins produced have a uniform 20 carbon chain length. These processes are hence more efficient than conventional processes which result in the production of mixture of α -olefins with different carbon chain length and which require separation of the different α -olefins for subsequent reactions.

Post-production processing

25 The produced α -olefins, more particularly medium-chain α -olefins such as C₁₁ α -olefins, can be used as or converted into a fuel, in particular a biofuel. These α -olefins, more particularly medium-chain α -olefins such as C₁₁ α -olefins, can also be used as starting material for the production of chemicals or personal care additives (e.g. polymers, surfactants, plastics, textiles, solvents, adhesives, etc.). They can also be 30 used as feedstock for subsequent reactions, such as hydrogenation and/or oligomerization reactions, to make other products.

A further aspect of the invention relates to a method for the production of poly- α -olefins (PAO), said method comprising:

- a) producing α -olefins, more particularly medium-chain α -olefins, according to a method disclosed herein;
- 5 b) oligomerizing the α -olefins produced in step a); and optionally
- c) hydrogenating the oligomer produced in step b).

In particular embodiments, a method is provided for the production of C_{33} PAOs, which comprises:

- a) producing C_{11} α -olefins according to a method disclosed herein;
- 10 b) trimerizing the C_{11} α -olefins produced in step a); and optionally
- c) hydrogenating the trimer produced in step b).

Oligomerization of medium-chain α -olefins in the presence of a catalyst is well known in the art. Catalysts that can be used for the oligomerization step are for example, but not limited to, $AlCl_3$, BF_3 , BF_3 complexes for cationic oligomerization, and metal based
15 catalysts like metallocenes.

Following the oligomerization step, residual unsaturation that is potentially present in the oligomers is saturated by catalytic hydrogenation resulting in saturated aliphatic hydrocarbons with one or more side branches.

The oligomers obtained by methods as described herein are known under the generic
20 name of poly- α -olefins (PAO). The PAO production methods described herein advantageously result in the homogenous production of PAOs of a well-defined carbon chain length. Accordingly, the application also provides a composition of PAO's obtainable by a PAO production method described herein, characterized in that at least 50%, preferably at least 85% or 90%, more preferably at least 95% such as 96%, 97%,
25 98% or even 99% of the PAOs have a well-defined carbon chain length such as C_{33} PAOs. Such a homogenous composition of poly- α -olefins has not previously been disclosed.

The methods provided herein allow obtaining a base oil with a well-defined viscosity. The PAOs, more particularly the C_{33} PAOs, obtainable by a method as described
30 herein can be used as base oils, which display very attractive viscosity indices, with the viscosity increasing with the number of carbons. These base oils can be used,

together with additives and optionally other base oils, to formulate lubricants. In particular, PAOs with a number of carbons of about 30, more particularly 33 carbons, are preferred for automotive lubricants.

5 Accordingly, in a further aspect, the present invention relates to the use of the medium-chain free fatty acid decarboxylase enzymes and the recombinant host cells described herein for the industrial production of lubricants.

Also provided herein are lubricants comprising poly- α -olefins, more particularly lubricants comprising poly- α -olefins which contain a more homogenous composition of poly- α -olefins, more particularly a high concentration of poly-olefins of a well-defined
10 length, such as those obtainable by a method as described herein. Indeed, the invention allows for the production of lubricants which are produced based on biosourced medium-chain alpha-olefins. More particularly, the invention allows for the provision of lubricants comprising poly- α -olefins, whereby at least 50%, preferably at least 85% or 90%, more preferably at least 95% such as 96%, 97%, 98% or even 99%
15 of said poly- α -olefins are poly- α -olefins of a well defined carbon chain length, such as C₃₃ poly- α -olefins. Lubricants comprising such a homogenous composition of poly- α -olefins have not previously been disclosed.

Further disclosed herein are methods for the production of alkanes, said methods comprising: (a) production of α -olefins more particularly medium-chain α -olefins,
20 according to a method disclosed herein; and (b) hydrogenation of the α -olefins obtained in step (a) to produce alkanes.

The present invention will now be further illustrated by means of the following non-limiting examples.

25

EXAMPLES

Example 1: Production of C₁₁ α -olefins by OleT_{JE}, Sm46, Bs168 and Aa162

Material and methods

Molecular cloning

The genes encoding OleT_{JE} (SEQ ID NO:1), Sm46 (SEQ ID NO:3), Bs168 (SEQ ID NO:5) and Aa162 (SEQ ID NO:7) were cloned into the *NdeI/BamHI* sites of the pET28a plasmid (Novagen, Fig. 2) by standard molecular biology techniques.

Protein overexpression and purification

5 *E. coli* BL21(DE3) cells (Novagen) carrying a recombinant plasmid or an empty PET28a plasmid (control) were cultured for several cycles to ensure best growth state at 37°C in 10 mL LB medium supplemented with 50 µg/ml of kanamycin, followed by inoculation (1:100 ratio) into 50 mL fresh Terrific Broth medium containing 50 µg/ml of kanamycin, 1 mM thiamine, 10% glycerol and a rare salt solution (6750 mg/l FeCl₃, 500
10 mg/l ZnCl₂, 500 mg/l CoCl₂, 500 mg/l Na₂MoO₄, 250 mg/l CaCl₂, 465 mg/l CuSO₄, and 125 mg/l H₃BO₃) at 37°C. Cells were grown at 37°C for 3 to 4 h until the optical density at 600 nm (OD₆₀₀) reached 0.6 to 0.8, at which 0.2 or 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and 0.5 mM δ-aminolevulinic acid were added, followed by 18 h of cultivation at 16°C.

15 Then, the cell cultures were recovered by centrifugation (at 4°C) and re-suspended in 1.5 mL 50 mM Tris buffer, pH 7.5, containing 0.1 M NaCl. The re-suspended cell cultures were disrupted by sonication and centrifuged at 13,000 x g for 30 min (4 °C). The supernatant was collected and the pellet was re-suspended in 1.5 mL 50 mM Tris buffer, pH 7.5, containing 0.1 M NaCl for further analysis. An SDS gel was run to
20 check the expression of the genes using 15 µL of supernatant or pellet solution mixed with 5 µL of 4x loading buffer (Fig. 3).

The amount of total proteins within the supernatant was quantified using Pierce BCA kit according to the manufacturer's instructions.

In vitro enzymatic assays

25 *In vitro* olefin biosynthesis was evaluated by adding to the supernatant a fatty acid substrate in DMSO to a final concentration of 0.2 mM (1:100 dilution) or 1 mM (1:10 dilution), and 500 µM H₂O₂. The reaction mixture was incubated at 28°C for 2-3 h, or stopped using 50 µL of 10 M HCl at t=0, 15, 30, 60 and 120 min for a dynamic test. Then, C₁₆ fatty acid was added at a final concentration of 0.2 mM (1:100 dilution) as an
30 internal standard. 300 µL reaction mixture of each reaction condition was extracted with equal volume of ethyl acetate, and the organic phase was analyzed for hydrocarbons by gas chromatography/mass spectrometry (GC/MS). In the meanwhile, undecene of

different concentrations (0, 0.05 mM, 0.1 mM, 0.2 mM and 0.5 mM) were tested to produce a standard curve.

Results

Figure 4 shows fatty acid decarboxylase activity for all tested enzymes OleT_{JE}, Sm46, 5 Bs168 and Aa162. Bs168 and Aa162 show preference for C₁₂ fatty acid substrate, thereby specifically producing C₁₁ α -olefins. In later experiments, Sm46 was also shown to have substrate preference for C₁₂ fatty acid substrate (data not shown).

10 **Example 2: Production of α -olefins by a truncated Sm46 from *Staphylococcus Massiliensis***

The construct pET28b-Sm46 of Example 1 did show poor expression in *E. coli* BL21(DE3), which did not improve upon subcloning into another expression vector pCWori or optimization of the expression conditions. In order to maximize expression 15 level and to assess potential of Sm46 as a decarboxylase enzyme with C₁₂ as preferred substrate, a truncated version of codon-optimized (for expression in *E. coli*) Sm46, Sm46-del29, was designed, which has the N-terminal 29 amino acids deleted, and the activity of the recombinant truncated Sm46-del29 was tested.

Material and methods

20 *Materials*

Fatty acid substrates and terminal alkene authentic standards were purchased from TCI (Shanghai, China). Antibiotics were obtained from Solar-Bio (Beijing, China). Other chemicals were purchased from Sigma Aldrich (St. Louis, MO, USA) or Ameresco (Solon, OH, USA). Molecular cloning kits, such as E.Z.N.A.TM Plasmid Miniprep Kit and Wizard SV Gel and PCR Clean-up System, were purchased from OMEGA Bio-Tek 25 (Jinan, China) and Promega (Madison, WI, USA) respectively. Oligonucleotides and codon-optimized gene were synthesized by Genewiz (Suzhou, China). The Pfu DNA polymerases and all restriction endonucleases were obtained from Takara (Dalian, China). Ni-NTA resin used for protein purification was from Qiagen (Valencia, CA, USA); Millipore Amicon Ultra centrifugal filters (Billerica, MA, USA) and PD-10 30 desalting columns were from GE Healthcare (Piscataway, NJ, USA).

Molecular cloning Sm46-del29 gene

The 1,275 bp fragment of the truncated gene, *Sm46-del29*, was PCR-amplified from the initial construct of pET28b-*Sm46* (see Example 1) using *Pfu* DNA polymerase and subsequently subcloned into pET28b vector at the Nde I/Xho I restriction sites. The sequences of primers used are as follows:
5 GTCCATATGGCAAAAAGCTGCCTAAAGTG (*Sm46-del29-F*, forward primer, SEQ ID NO:11) and GTACTIONGAGTTATTTGCGGGCAACACGCGG (*Sm46-del29-R*, reverse primer, SEQ ID NO:12). All recombinant plasmid constructs were confirmed by DNA sequencing (Sangon Biotech, Shanghai, China). Upon sequence verification, plasmids were used to transform *E. coli* strain BL21 (DE3) for protein expression.

10

Heterologous expression and purification of Sm46-del29

The *E. coli* BL21 (DE3) cells transformed with the recombinant pET28b-*Sm46-del29* plasmid were grown overnight at 37 °C with shaking at 220 rpm in LB medium containing 50 µg/ml kanamycin. The overnight culture was used as seed culture to inoculate (1:100 dilution) 1 to 3 litres of modified Terrific Broth containing 4 % glycerol, 1 mM thiamine, trace metal and the corresponding antibiotics. Cells were then grown at 15 37 °C for 3 to 4 h until the optical density at 600 nm (OD₆₀₀) reached ~ 0.6, at which point δ-aminolevulinic acid (0.5 mM final concentration) was supplemented and the expression of *Sm46-del29* was induced by the addition of 0.2 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Cells were then further cultured for 24 h at a reduced 20 temperature of 18°C before being harvested by centrifugation at 6000 rpm, 4 °C. The cell pellet was frozen at – 80 °C until required.

Purification of the His-tagged protein was carried out as described by Liu et al. (2014 25 Biotechnol. Biofuels 7:28) with minor modifications. All protein purification steps were performed at 4 °C. Specifically, the cell pellets were thawed and re-suspended in 40 ml lysis buffer (50 mM NaH₂PO₄, 300 mM NaCl, 10% glycerol and 10 mM imidazole, pH 8.0) through vortexing. After cell disruption by ultrasonication, the cell lysate was centrifuged at 12,000 × g for 30 min to remove cellular debris. To the clarified cell 30 lysate, 1 ml of Ni-NTA resin was added and gently mixed at 4°C for 1 h. The slurry was then loaded onto an empty column, and washed with approximately 100 ml of wash buffer (50 mM NaH₂PO₄, 300 mM NaCl, 10% glycerol and 20 mM imidazole, pH 8.0) until no proteins were detectable in flow-through. The bound target proteins were eluted with elution buffer (50 mM NaH₂PO₄, 300 mM NaCl, 10% glycerol and 250 mM

imidazole, pH 8.0). The eluates were pooled and concentrated with an Amicon Ultra centrifugal filter, and buffer exchanged on a PD-10 column into storage buffer (50 mM NaH₂PO₄, 300 mM NaCl 10% glycerol, pH 7.4). The final purified protein was flash-frozen in liquid nitrogen and stored at -80°C for later use.

5

UV-visible Spectroscopy—Spectroscopic characterization of Sm46-del29

Analysis of the UV-visible spectroscopic properties of *Sm46-del29* was performed on a Cary 60 UV-visible spectrophotometer (Varian, UK). The ferrous-CO complex of *Sm46-del29* was prepared by slow bubbling of carbon monoxide gas into a solution of ferric P450, and subsequent reduction of the protein with sodium dithionite. The CO-bound reduced difference spectrum of the enzyme was obtained through recording the respective spectrum before and after the addition of sodium dithionite (Omura and Sato 1964. J Biol Chem 239:2379-2385, Amaya et al. 2016 J Inorg Biochem. 158:11-16).

In vitro enzymatic assays

The fatty acid decarboxylation assays containing 1.0 μM *Sm46-del29*, 200 μM fatty acid substrate (from C₈ to C₂₀), 220 μM H₂O₂ in 200 μl of storage buffer were carried out at 28°C for 2 h. Reactions were quenched by the addition of 20 μl of 10 M HCl. Heptadecanoic acid was added as internal standard and the mixture was extracted by 200 μl ethyl acetate. Following extraction, the organic phase was collected and analyzed by gas chromatography as described below.

Gas chromatography (GC)

The GC analytical method for hydrocarbon and fatty acid samples was adapted from Guan et al. (2011 J Chromatogr A 1218:8289-8293). The analyses were performed on an Agilent 7890B gas chromatograph equipped with a capillary column HP-INNOWAX (Agilent Technologies, Santa Clara, CA, USA; cross-linked polyethylene glycerol, i.d. 0.25 μm film thickness, 30 m by 0.25 mm). The helium flow rate was set to 1 ml per minute. The oven temperature was controlled initially at 40°C for 4 min, then increased at the rate of 10°C per min to 250°C, and held for 5 min. The injecting temperature was set to 280°C with the injection volume of 1 μl under splitless injection conditions. The response factors between fatty acids and alkenes were determined by analyzing known authentic fatty acids (C₁₀–C₂₀), 1-alkenes (C₉–C₁₉) and 1-heptadecanoic acid standards as described in Liu et al. (2014 Biotechnol. Biofuels 7:28).

Results

The truncated variant of Sm46 (i.e. Sm46-del29) had decarboxylase activity and showed substrate preference for C₁₂ free fatty acid for α -olefin production among the tested fatty acid substrates (Fig. 6). Sm46-del29 was also able to decarboxylate C₁₀ free fatty acid, but technical problems (due to volatility) hindered detection of the formed C₉ α -olefin.

Example 3: Production of α -olefins and hydroxy fatty acids by Sm46-del29, Bs168 and Aa162

Recombinantly expressed and purified Bs168 and Aa162 from Example 1 and Sm46-del29 of Example 2 were reacted with C₁₄ fatty acid and the products of decarboxylation (C₁₃ α -olefin) and hydroxylation (α -OH-C₁₄ fatty acid and β -OH-C₁₄ fatty acid) reactions were analyzed.

All enzymes tested were able to decarboxylate myristic acid (C₁₄), but also catalyzed α - and β -hydroxylation of myristic acid as side reactions. Fatty acid decarboxylation was the dominant reaction for all tested enzymes, but Sm46-del29 formed less hydroxy fatty acids, indicating that these enzymes have specific decarboxylase activity.

CLAIMS

1. A method for the production of C₇-C₁₁ α-olefins comprising culturing a recombinant host cell comprising a recombinant nucleic acid encoding a decarboxylase enzyme under conditions suitable for the production of C₇ to C₁₁ α-olefins by said host cell, wherein the preferred substrate of said decarboxylase enzyme is an C₈-C₁₂ free fatty acid.
2. The method of claim 1, wherein said nucleic acid encoding a decarboxylase enzyme comprises a nucleotide sequence having at least about 75%, preferably at least about 80%, more preferably at least about 95%, sequence identity to SEQ ID NO:10, SEQ ID NO:7, SEQ ID NO: 3 or SEQ ID NO: 5 and wherein said recombinant nucleic acid ensures expression or overexpression of said decarboxylase.
3. The method according to claim 1 or 2, wherein said nucleic acid encodes a polypeptide comprising an amino acid sequence having at least 80%, preferably at least 90% identity, to SEQ ID NO:13, SEQ ID NO: 8, SEQ ID NO: 4 or SEQ ID NO: 6 or an active fragment of said polypeptide.
4. The method according to any one of claims 1 to 3, wherein said nucleic acid encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:13.
5. The method according to any one of claims 1 to 4, wherein C₁₂ free fatty acids are the preferred substrate of said decarboxylase, and wherein said α-olefins are C₁₁ α-olefins.
6. The method according to any one of claims 1 to 5, wherein said host cell is cultivated in a medium comprising C₈-C₁₂ free fatty acids, preferably C₁₂ free fatty acids.
7. The method according to any one of claims 1 to 6, wherein the host cell has further been genetically engineered to produce or overproduce C₈-C₁₂ free fatty acids, preferably C₁₂ free fatty acids.

8. The method according to claim 7, wherein the host cell comprises a recombinant nucleic acid encoding an enzyme involved in the production of free fatty acids with a carbon chain length comprised between 8 and 12, more particularly a C₁₂ free fatty acid.
- 5
9. The method according to claim 8, wherein the host cell comprises a recombinant nucleic acid encoding a thioesterase having activity on C₈ to C₁₂ acyl-ACP, preferably on C₁₂ acyl-ACP.
- 10
10. The method according to any one of claims 1 to 9, wherein the host cell is an oleaginous host cell.
11. The method according to any one of claims 1 to 10, wherein the host cell is
- 15 selected from the group comprising bacteria, yeasts, fungi, plants and algae.
12. The method according to any one of claims 1 to 11, further comprising the step of recovering the α -olefins from the host cell or the culture medium.
- 20
13. A method for the production of poly- α -olefins comprising the following steps:
- i) producing C₇-C₁₁ α -olefins according to a method according to any one of claims 1 to 12; and
- ii) performing an oligomerization reaction using the α -olefins obtained in step i) as monomer to produce an oligomer; and optionally,
- 25 iii) hydrogenating the oligomer produced in step ii).
14. The method according to claim 13, wherein the poly- α -olefins are C₃₃ poly- α -olefins, wherein step i) comprises the production of C₁₁ α -olefins, and wherein the oligomerization reaction in step ii) is a trimerization reaction.
- 30
15. A composition comprising poly- α -olefins obtainable by the method according to claim 14, wherein at least 85%, preferably at least 90%, more preferably at least 95%, of the poly- α -olefins are C₃₃ poly- α -olefins.

16. A polypeptide having decarboxylase activity, wherein said polypeptide comprises an amino acid sequence having at least 80%, preferably at least 90%, identity to SEQ ID NO:13.
- 5 17. An isolated nucleic acid encoding an enzyme having decarboxylase activity, wherein said nucleic acid comprises a nucleotide sequence having at least about 80%, more preferably at least about 95%, sequence identity to SEQ ID NO: 7 or SEQ ID NO: 5.
- 10 18. The polypeptide according to claim 16 or the isolated nucleic acid according to claim 17, wherein C₈ to C₁₄ free fatty acids, preferably C₈ to C₁₂ free fatty acids, more preferably C₁₂ free fatty acids are the preferred substrate of said decarboxylase.
- 15 19. The isolated nucleic acid according to claim 17 or 18, wherein said nucleic acid comprises a nucleotide sequence having at least about 95% sequence identity to SEQ ID NO: 7.
- 20 20. A vector comprising at least one regulatory sequence operatively coupled to the nucleic acid sequence of any of claims 17 to 19.
21. A host cell comprising the nucleic acid according to any of claims 17 to 19 integrated into its genome or the vector according to claim 20.
- 25 22. The host cell of claim 21, which is an oleaginous eukaryotic microalgae or oleaginous yeast.
23. Use of a host cell according to claim 21 or 22 for the industrial production of lubricants.
- 30 24. Lubricant comprising poly- α -olefins obtainable by a method according to claim 13 or 14.
25. Lubricant according to claim 24, wherein said poly- α -olefins are biosourced poly- α -olefins.
- 35

26. Lubricant according to claim 24 or 25, wherein at least 50%, preferably at least 85%, of said poly- α -olefins consist of C₃₃ poly- α -olefins.

5

27. Lubricant according to any one of claims 24 to 26, wherein said lubricant is an automotive lubricant.

1/10

A) SEQ ID NO:1: *Jeotgalicoccus* sp. ATCC 8456 terminal olefin-forming fatty acid decarboxylase gene, complete cds

ATGGCAACACTTAAGAGGGATAAGGGCTTAGATAATACTTTGAAAGTATTAAGCAAGGTTATC
TTTACACAACAAATCAGAGAAATCGTCTAAACACATCAGTTTTCCAAACTAAAGCACTCGGTGG
TAAACCATTTCGTAGTTGTGACTGGTAAGGAAGGCGCTGAAATGTTCTACAACAATGATGTTGTT
CAACGTGAAGGCATGTTACCAAACGTATCGTTAATACGCTTTTTGGTAAAGGTGCAATCCATA
CGGTAGATGGTAAAAACACGTAGACAGAAAAGCATTGTTTCATGAGCTTGATGACTGAAGGTAA
CTTGAATTATGTACGAGAATTAACGCGTACATTATGGCATGCGAACACACAACGTATGGAAAGT
ATGGATGAGGTAAATATTTACCGTGAATCTATCGTACTACTTACAAAAGTAGGAACACGTTGGG
CAGGCGTTCAAGCACCACCTGAAGATATCGAAAGAATCGCAACAGACATGGACATCATGATCGA
TTCATTTAGAGCACTTGGTGGTGCCTTTAAAGGTTACAAGGCATCAAAAAGAAGCACGTCGTCGT
GTTGAAGATTGGTTAGAAGAACAATTATTGAGACTCGTAAAGGGAATATTCATCCACCAGAAG
GTACAGCACTTTACGAATTTGCACATTGGGAAGACTACTTAGGTAACCCAATGGACTCAAGAAC
TTGTGCGATTGACTTAATGAACACATTCCGCCCATTAATCGCAATCAACAGATTCGTTTCATTC
GGTTTACACGCGATGAACGAAAACCCAATCACACGTGAAAAAATTAATCAGAACCTGACTATG
CATATAAATTCGCTCAAGAAGTTCGTTCGTTACTATCCATTCGTTCCATTCCTTCCAGGTAAAGC
GAAAGTAGACATCGACTTCCAAGGCGTTACAATTCCTGCAGGTGTAGGTCTTGCATTAGATGTT
TATGGTACAACGCATGATGAATCACTTTGGGACGATCCAAATGAATTCGCCCCAGAAAGATTCC
AACTTGGGACGGATCACCATTTGACCTTATTCACAAGGTGGTGGAGATTACTGGACAAATCA
CCGTTGTGCAGGTGAATGGATCACAGTAATCATCATGGAAGAAACAATGAAATACTTTGCAGAA
AAAATAACTTATGATGTTCCAGAACAAGATTTAGAAGTGGACTTAAACAGTATCCCAGGATACG
TTAAGAGTGGCTTTGTAATCAAAAATGTTTCGCCAAGTTGTAGACAGAACATAA

B) SEQ ID NO:2

MATLKRDKGLDNTLKVLKQGYLYTTNQRNRLNTSVFQTKALGGKPFVVVTGKEGAEMFYNNDDVV
QREGMLPKRIVNTLFGKGAIH TVDGKKHVDRKALFMSLMTEGNLNYVRELRTLWHANTQRMES
MDEVNIYRESIVLLTKVGTRWAGVQAPPEDIERIAATDMDIMIDSFALGGAFKGYKASKEARRR
VEDWLEEQIIETRKGNIHPPGEGTALYEFAHWEDYLGPNPMSRSTCAIDLMNTFRPLIAINRFVSF
GLHAMNENPITREKIKSEPDYAYKFAQEVRRYYPFVFPFLPGKAKVDIDFQGV TIPAGVGLALDV
YGTTHDESLWDDPNFRPERFETWDGSPFDLIPQGGGDYWTNHRCAGEWITVIIMEETMKYFAE
KITVDVPEQDLEVDLNSIPGYVKS GFVIKNVREVVDR

FIG. 1

2/10

C) SEQ ID NO:3: *Bacillus subtilis* subsp. *subtilis* str. 168 chromosome, complete genome

ATGAATGAGCAGATTCCACATGACAAAAGTCTCGATAACAGTCTGACACTGCTGAAGGAAGGGT
ATTTATTTATTAAAAACAGAACAGAGCGCTACAATTCAGATCTGTTTCAGGCCCGTTTGTGGG
AAAAAACTTTATTTGCATGACTGGCGCTGAGGCGGCGAAGGTGTTTTATGATACGGATCGATTC
CAGCGGCAGAACGCTTTGCCAAGCGGGTGCAGAAATCGCTGTTTGGTGTTAATGCGATTCAGG
GAATGGATGGCAGCGCGCATATCCATCGGAAGATGCTTTTTCTGTCAATTGATGACACCGCCGCA
TCAAAAACGTTTGGCTGAGTTGATGACAGAGGAGTGGAAAGCAGCAGTCACAAGATGGGAGAAG
GCAGATGAGGTTGTGTTATTTGAAGAAGCAAAGAAATCCTGTGCCGGGTAGCGTGCTATTGGG
CAGGTGTTCCGTTGAAGGAAACGGAAGTCAAAGAGAGAGCGGATGACTTCATTGACATGGTCGA
CGGTTCCGGTGCTGTGGGACCGCGGCATTGGAAAGGAAGAAGAGCAAGGCCGCGTGCGGAAGAG
TGGATTGAAGTCATGATTGAAGATGCTCGTGCCGGCTTGCTGAAAACGACTTCCGGAACAGCGC
TGCATGAAATGGCTTTTACACACAAGAAGATGGAAGCCAGCTGGATTCCCGCATGGCAGCCAT
TGAGCTGATTAATGTACTGCGGCCTATTGTGCGCCATTTCTTACTTTCTGGTGTTTTAGCTTTG
GCGCTTCATGAGCATCCGAAGTATAAGGAATGGCTGCGGTCTGGAAACAGCCGGGAAAGAGAAA
TGTTTGTGCAGGAGGTCCGCAGATATTATCCGTTCCGCCCCGTTTTTAGGGGCGCTTGTCAAAA
AGATTTTGTATGGAATAACTGTGAGTTTAAGAAGGGCACATCGGTGCTGCTTGATTTATATGGA
ACGAACCACGACCCTCGTCTATGGGATCATCCCGATGAATCCGGCCGGAACGATTTGCGGAGC
GGGAAGAAAATCTGTTTGATATGATTCCCTCAAGGCGGGGGGCACGCCGAGAAAGGCCACCGCTG
TCCAGGGGAAGGCATTACAATTGAAGTCATGAAAGCGAGCCTGGATTCCTCGTCCATCAGATT
GAATACGATGTTCCGGAACAATCACTGCATTACAGTCTCGCCAGAATGCCATCATTGCCTGAAA
CGGCTTCGTAATGAGCGGAATCAGACGAAAAAGTTAA

D) SEQ ID NO:4

MNEQIPHDKSLDNSLTLKKEGYLF IKNRTERYNSDLFQARLLGKNFICMTGAEAAKVFYDTRF
QRQNALPKRVQKSLFGVNAIQGMDGSAHIRKMLFLSLMTPPHQKRLAELMTEEWKAAVTRWEK
ADEVVLFEAKEILCRVACYWAGVPLKETEVKERADDFIDMVDAFGAVGPRHWKRRARPRAE
WIEVMIEDARAGLLKTTSGTALHEMAFHTQEDGSQLD SRMAAIELINVL RP IVAI SYFLVFSAL
ALHEHPKYKEWLRSNSREREMFVQEVRRYYPFGPFLGALVKKDFVWNNCEF KKGTSVLLDLYG
TNHDPRLWDHPDEF RPERFAEREENLFD MIPQGGGHA EKGHRCPEGITIEVMKASLDFLVHQI
EYDVPEQSLHYS LARMP SLPESGFVMSGIRRS

FIG. 1 (continued)

3/10

E) SEQ ID NO:5: ALICYCLOBACILLUS ACIDOCALDARIUS LA1 CTG162, WHOLE GENOME SHOTGUN SEQUENCE

ATGAATCAGTGCATTCCGCGCGATCGAACGTTTGACAGCAGCCTCGCCTTGATAAAGGAAGGGT
ATTTGTTTCATCAAAAATCGAGTTGATCAATACCAATCCGACATCTTCGAAGCGCGTCTCCTCCT
GGAAAATGTGGTATGCATGCACGGAGCAGAGGCGGCAAACTCTTCTACAATACGGAAGTGT
CAACGCCAAGGTGCTCTTCCGAAGCGGGTTCAAAGACGCTTTTCCGAGAAAACGCCATCCAAA
CCCTTGATGGTACAGCGCATCTTACCCGTAAGCAGCTGTTTCTGTCTGTTGACGCCGGATCA
AGAAAAATCCCTTGCAGCGCTCGCGACAACGCAGTGGAGGGAGTGCGCGAAGGTATGGGAGAAC
GCGGATAGGGTTGTGCTATTTGAAGAGGCCAAGCGGATGTTATGTCTGGATCGCATGTCAGTGG
CCGGGGTTCCGCTGGATGAATCGGAGGTGTCAAAGCGGGCCGACGATTTTGGGGCGATGGTGG
CGGTTTTGGAGCGGTTGGTCCGCGACATTGGAAAGGCCGGAGAGCTCGGGCCAGAGCAGAAGCA
TGGCTCCGGCAGATGATTGACGAGATACGAATCGGATTGCGTAGTGTAGATGAACATACGCCGC
TCCATGTGGTGGCCTTTTGGCGTGACGTGAATGGAAACCTCTTGGATGCTCAGATGGTTGCAAT
CGAGTTAATCAATCTGCTACGACCCATCGTAGCTATTTCTACTTTTCATCACGTTTTTCAGCCCTG
GCCCTGCACGAACACCCGACATGGCGAGACCGATTGAAGGCGCGCAATGAAGCGGATATCGAGA
TGTTTGTGCAAGAGGTTTCGTCGCTACTATCCGTTTCGCGCCATTTCTCGGTGCCAGAGTGAAAA
GGATTTTGTGTGGAGGGGATACGAATTTAAAAGAGGGACCCTTGTGTTGCTGGATGTGTATGGA
ACCCATCATGATGCCCGCCTCTGGGATTCCCAAATGAGTTTCGACCCGAACGATTCATGAGAA
AAACAGTTGGGCGTTTTGATTTGATTCCTCAAGGTGGAGGGGACTCTCACACCGGTCATCGTTG
CCCTGGTGAAGGCGCCACCATCGAGATTATGAAGGCGAGCGTGGATTTTCTGGTTAACCAAATT
GACTTCGAAGTGCCCGCTCAGGACCTCAGTTACAGATTGGATGTTATGCCGACGTTGCCAAAGA
GCGGATTTGTGCTGACCCATGTTTCATCGGAAGTTCATAGCTTCTCCGACCATTGCTACACCTAA
TGTTTCTGAAGCTCTTCCCTCAGAAGTCTAA

F) SEQ ID NO:6

MNQCI PRDRTFDSSLALIKEGYLF IKNRVDQYQSD IFEARLLLENV VCMHGAEAAKLFYNTELF
QRQ GALPKRVQKTLFGENAIQTLDGTAHLHRKQLFLSLLTPDQEKSLATLATTQWRECAKVWEN
ADR VVLFEEAKRMLCRIACQWTGVPLDESEVSKRADDFGAMVDAFGAVGPRHWKGRRARARAEA
WLRQMIDEIRIGLRVDEHTPLHVVAFWRDVNGNLLDAQMVAIELINLLRP IVAISTFITFSAL
ALHEHPTWRDR LKARNEAD IEMFVQEVRRYYPFAPFLGARVKKDFVWRGYEFKRGTLVLLDVY
G THHDARLWDS PNEFRPERFMRKTVGPFDLIPQGGDSHTGHRCPGEGATIEIMKASVDFLVNQI
DFEVPAQDLSYRLDVMPTLPKSGFVLTHVHRKFIASPTIATPNGSEALPSEV

FIG. 1 (continued)

4/10

G) SEQ ID NO:7: STAPHYLOCOCCUS MASSILIENSIS S46 CONTIG01, WHOLE GENOME SHOTGUN SEQUENCE

ATGTTTGTAGATTCGATACTTGTGTTAAGATTAATTTATTA AAAACGGGTATACAATTAGAAA
 TGAAAAATGGGGGAATCAAAGTGGCAAAGAACTACCTAAGGTTAAAGGCCTAGATAACACAGT
 AGACATTATTAAAGGCGGGTATACATACGTACCTGGCAAATTAGAAGAATTTGATTCTAAAGCA
 TTTGAAGTACGCGCATTAGGCGGTAAGAAAATTGCTGTTATGAGCGGTAAGAAGCGGCAGAAA
 TTTTCTATGATAATGAAAAAATGGAAAGACAAGGTACTTTACCAAACGTATCGTAAACACTTT
 ATTTGGTAAAGGTGCAATTCATACAACCTGCTGGTAAGAAGCACGTTGACCGTAAAGCTTTATTT
 ATGTCACCTTATGACAGATGAAAATCTTAACTACTTACGTGAATTAACACGTAATTATTGGTTCA
 TGAATACTGAACGTATGCAAAGCATGGATAAAGTTAACGTATATAACGAATCAATTTATATGTT
 AACTAAAATCGGCTTCCGTTGGGCTGGTATCATCCAAACGCCTGAAGAAGCAGAACAAAATGCG
 AAAGACATGGATACTATGATTAACTCATTCGTATCTTTAGGTTTACGCTTACAAAGGTTATAAGA
 AAGCTAAAAAGCACGTAAACGTGTTGAAGATTTCTTAGAAAAACAAATTATCGATGTGCGTAA
 AGGTAAATTACACCCTGAAGAAGGTACTGCGTTATACGAATTCGCGCATTGGGAAGATTTAAAC
 GATAACCCAATGGATTCTCACTTATGTGCAGTAGACTTAATGAACGTTGTGCGCCATTAGCTG
 CAATCAACCGTTTCATCAGCTATGGTGTAAAGTATTAATCGAATTCGATCAAGAAAAAGAAAA
 ATTACGTCTTGAAAATAATGAAGACTATGCGTATAAATTCGCTCAAGAAGTACGTGCTATCTTC
 CCATTCGTACCATACTTACCAGGTAGAGCAGCTGTTGATTTAGAATATGACGGCTACAAAATCC
 CTGCAGGTATGATGACAGCATTAGATGTTTATGGTACGACACATGATGAAGATTTATGGGAAAA
 CCCAGACCAATTCAATCCTAACCCTTTGATAACTGGGACGGTAGCCATTGACTTAATTCCA
 CAAGGTGGCGGTGACTTCTATACGAACCACAGATGTGCTGGTGGTGGATCACAGTTATCATT
 TGGAAGAAACAATGAAATATTTGCGGAATAAGATTGAATTTGATGTACCGTCTCAAGATTTATC
 AGTTAAGCTTGATAAATTACCAGGTAACGTAACAAGCGGTACAATCATTAGTAATGTACGTCCA
 CGTGTGCGCGTAAATAA

H) SEQ ID NO:8

MFVDSILVLRNLNLLKTGIQLEMKNNGGIKVAKKLPKVKGLDNTVDI IKGGYTYVPGKLEEFDSKA
 FEVRALGGKKI AVMSGKEAAE IFYDNEKMERQGTLPKRI VNTLFGKGAIHTTAGKKHVDRKALF
 MSLMTDENLNYLRELTRNYWFMNTERMQSMDKVN VYNESIYMLTKIGFRWAGIIQTPEEAEQNA
 KDMDTMINSFVSLGSAYKGYKKAKKARKRVEDFLEKQIIDVRKGLHPEEGTALYEF AHWEDLN
 DNPMDSHLCAVDLMNVVRPLAAINRFISYGVKVLIEFDQEKEKLRLENNEDYAYKFAQEVRRIF
 PFVPLYLPGRAAVDLEYDGYKIPAGMMTALDVYGTTHDEDLWENPDQFNPNRFDNWDGSPFDLIP
 QGGDFYTNHRCAGEWITVIIMEETMKYFANKIEFDVPSQDLSVKLDKLPGNVTSGTII SNVRP
 RVARK

FIG. 1 (continued)

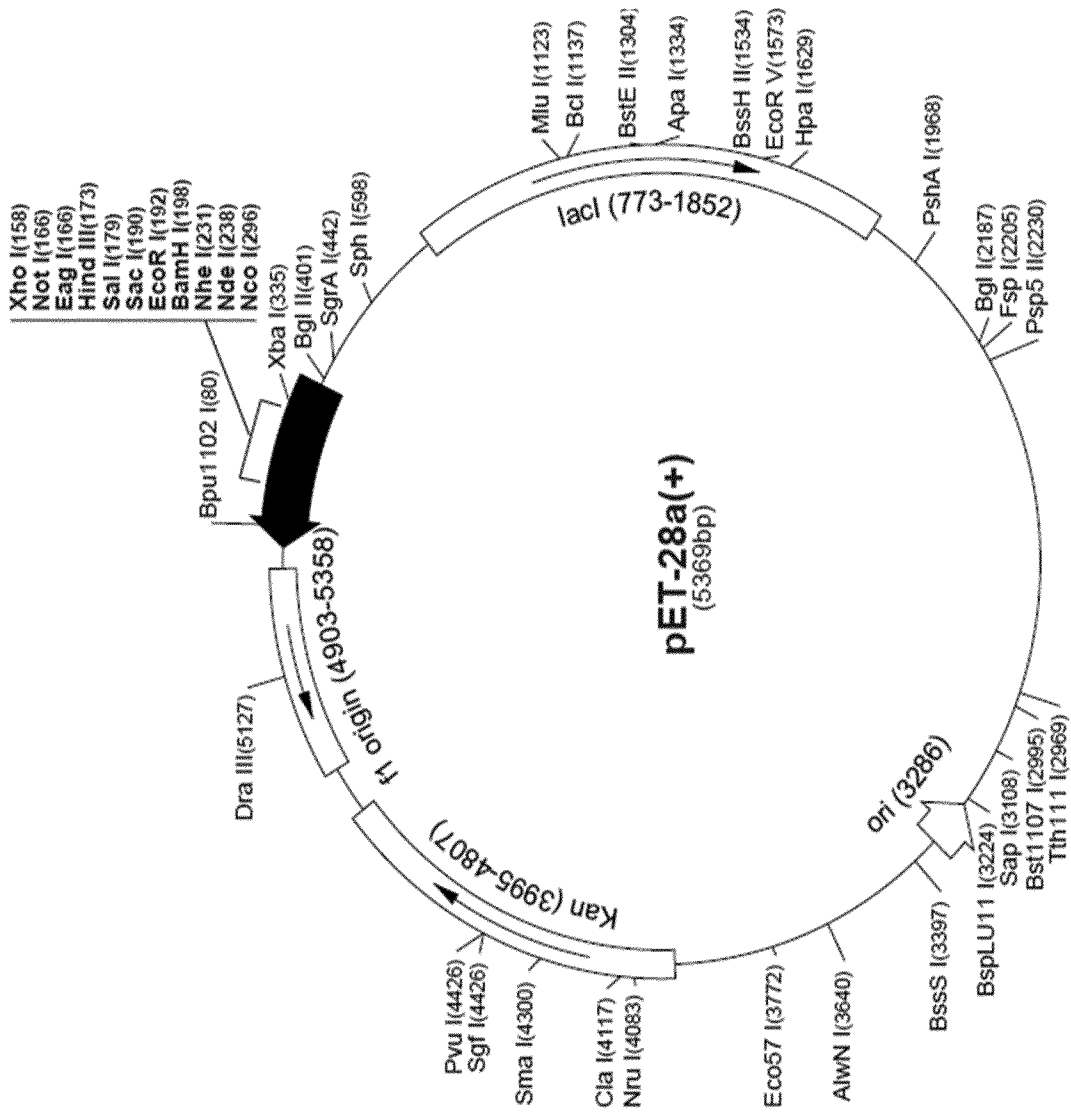


FIG. 2

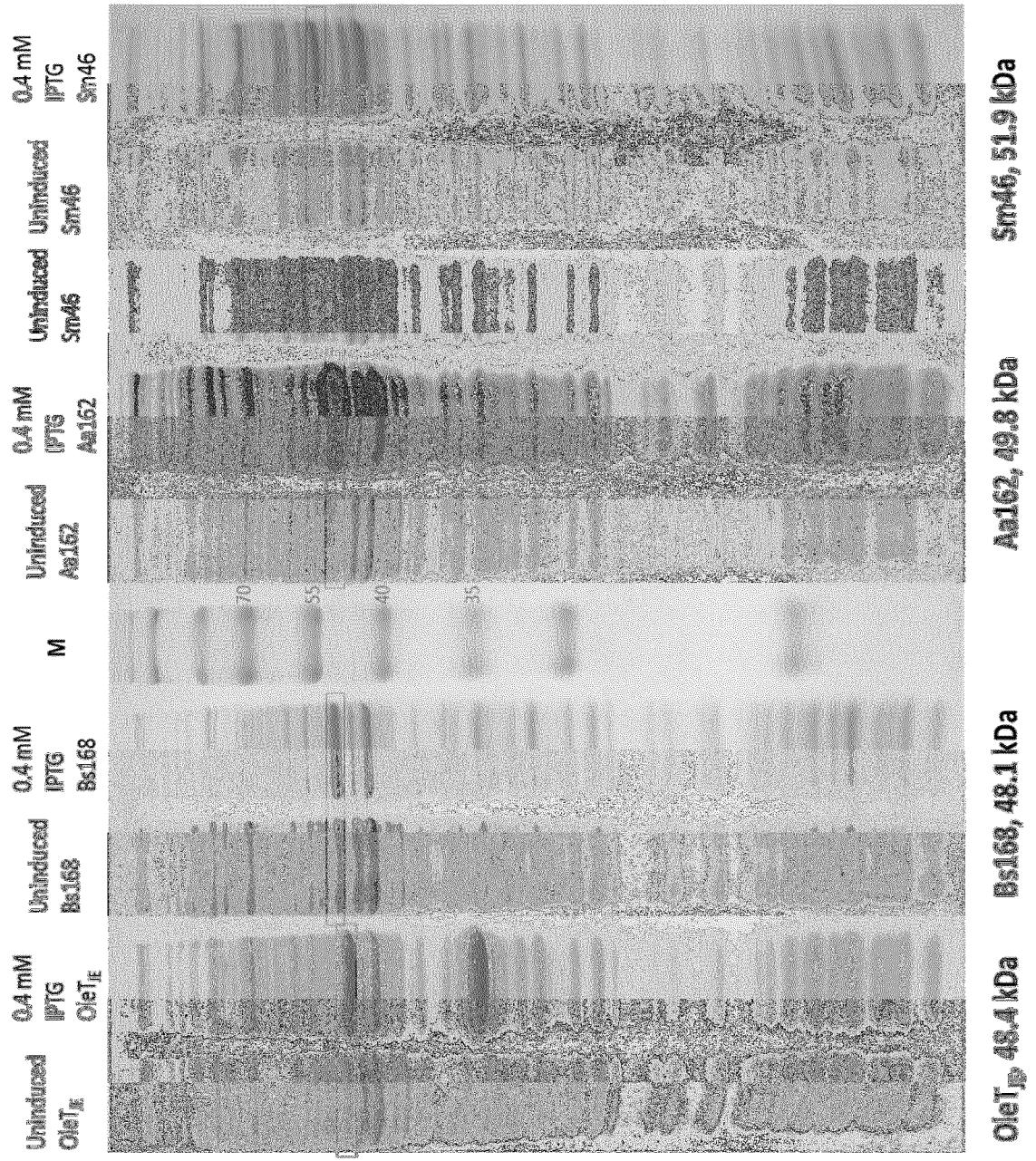


FIG. 3

7/10

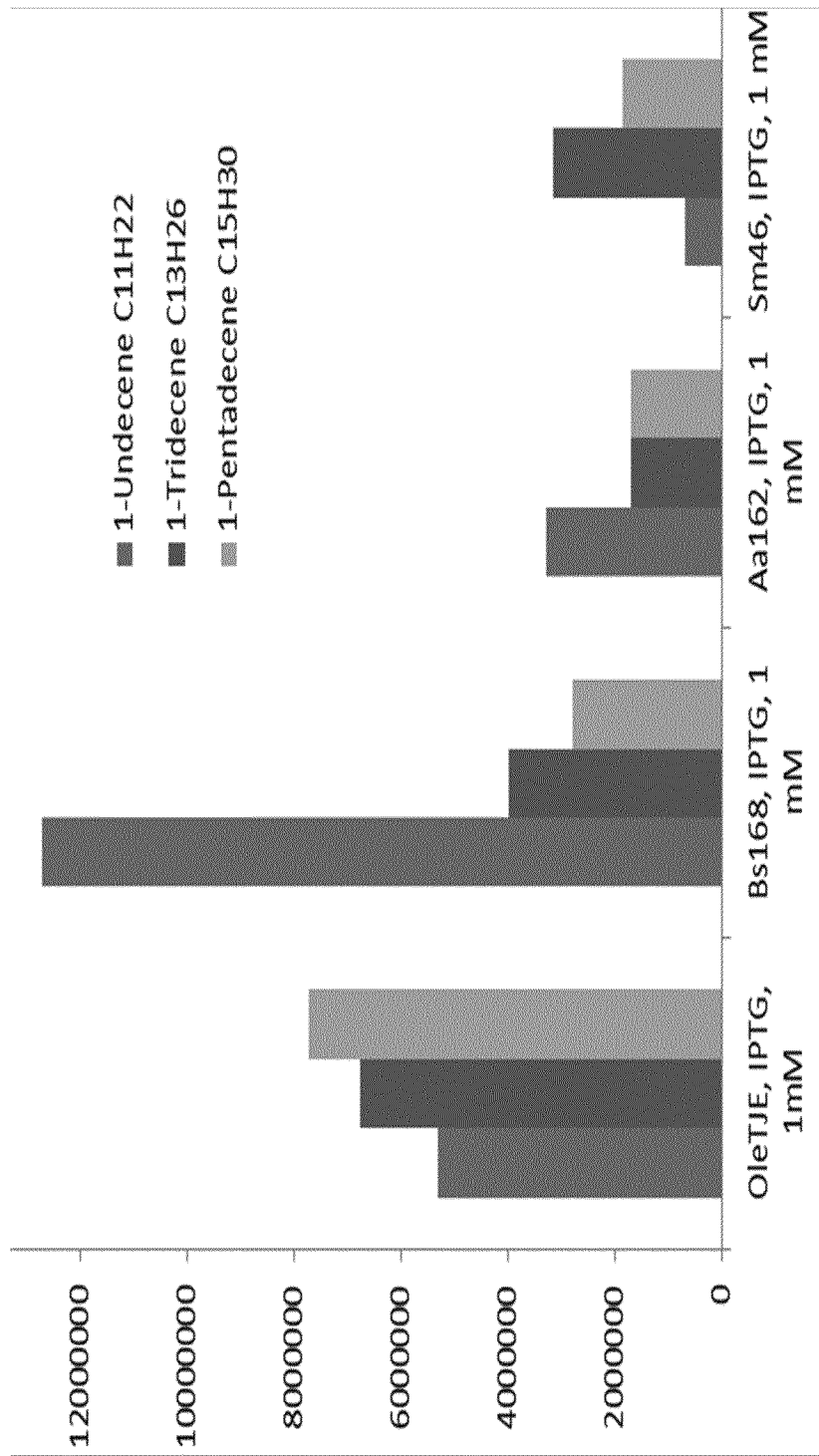


FIG. 4

8/10

(A) Codon-optimized (for expression in *E. coli*) nucleotide sequence encoding Sm46 from *Staphylococcus Massiliensis* strain S46 (SEQ ID NO:9) and flanking restriction sites

CATATGTTCGTGGATAGCATTCTGGTTCTGCGCCTGAACCTGCTGAAGACAGGCATCCAGCTGG
 AGATGAAGAACGGTGGCATCAAAGTGGCAAAAAAGCTGCCTAAAGTGAAAGGTCTGGACAACAC
 CGTGGACATCATCAAGGGTGGCTATACCTACGTGCCTGGCAAACCTGGAGGAGTTCGACAGCAAA
 GCATTTCGAAGTGC GCGCCCTGGGTGGCAAGAAGATCGCAGTGATGAGCGGCAAGGAAGCCGCCG
 AGATTTTTTATGATAACGAAAAAATGGAGCGTCAGGGTACCCTGCCGAAGCGCATCGTGAACAC
 ACTGTTTCGGTAAAGGCGCCATTTCATACCACCGCCGGCAAGAAACATGTGGATCGCAAGGCACTG
 TTCATGAGTCTGATGACCGATGAAAATTTAAATTATCTGCGCGAACTGACACGCAACTATTGGT
 TTATGAATACAGAACGCATGCAGAGCATGGATAAAGTGAATGTGTACAATGAAAGCATTATAT
 GCTGACCAAAAATTGGCTTCCGCTGGGCCGGTATCATTTCAGACCCCTGAAGAGGCCGAGCAGAAT
 GCCAAAGACATGGACACCATGATCAACAGCTTTGTGAGCCTGGGCAGCGCCTACAAGGGTTACA
 AAAAAGCCAAGAAAGCCC GCAAGCGCTGGAAGATTTTCTGGAGAAAACAATTATCGACGTTCC
 TAAAGGCAAACCTGCATCCGGAGGAAGGTACCGCCCTGTACGAATTCGCCATTGGGAAGACCTG
 AACGATAACCCGATGGACAGCCATCTGTGCGCCGTTGATCTGATGAACGTTGTTCCGCCGCTGG
 CAGCAATTAACCGCTTCATTAGCTACGGCGTTAAAGTGCTGATCGAATTCGACCAGGAAAAAGA
 AAAGCTGCGCCTGGAGAACAACGAGGACTACGCCTACAAGTTCGCACAGGAAGTGCGCCGTATC
 TTTCCGTTTCGTGCCTTACTTACCGGGTTCGCGCCGCCGTTGGATCTGGAGTATGATGGCTATAAGA
 TCCCGGCCGGTATGATGACCGCCCTGGATGTTTACGGTACCACACACGATGAGGATCTGTGGGA
 GAATCCGGATCAGTTCAACCCGAATCGTTTTGATAACTGGGACGGCAGTCCGTTTGATCTGATT
 CCGCAGGGCGGTGGCGATTTCTACACCAATCATCGTTGCGCCGGCGAGTGGATCACCGTGATTA
 TTATGGAAGAAACAATGAAATACTTTGCCAACAAAATTGAATTCGATGTGCCGAGTCAGGACCT
 GAGCGTTAAACTGGACAAACTGCCTGGCAACGTGACCAGCGGTACCATCATTAGCAACGTGCGT
 CCGCGTGTGGCCGCAAATAA**CTCGAG**

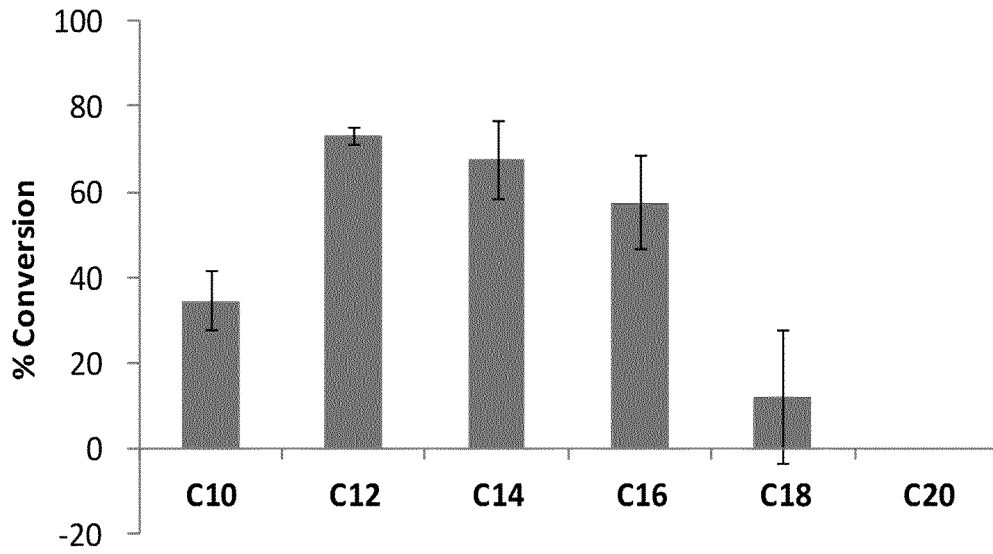
(B) SEQ ID NO: 13

AKKLPKVKGLDNTVDIIKGGYTYVPGKLEEFDSKAFEVRLGGKKIAVMSGKEAAEIFYDNEKM
 ERQGTLPKRIVNTLFGKGAIHHTAGKKHVDRKALFMSLMTDENLNYLRELTRNYWFMNTERMQS
 MDKVNYNESIYMLTKIGFRWAGIIQTPEEAEQNAKDMDTMINSFVSLGSAYKGYKKAKKARKR
 VEDFLEKQIIDVRKGLHPEEGTALYEFAHWEDLNDNPMDSHLCAVDLMNVVRPLAAINRFISY
 GVKVLIIEFDQEKEKLRLNEDYAYKFAQEVRRIFPFVYLPGRAAVDLEYDGYKIPAGMMTAL
 DVYGTTHDEDLWENPDQFNPNRFDNWDGSPFDLIPQGGDFYTNHRCAGEWITVIIMEETMKYF
 ANKIEFDVPSQDLSVKLDKLPGNVTSGTIISNVRPRVARK

FIG. 5

9/10

A



B

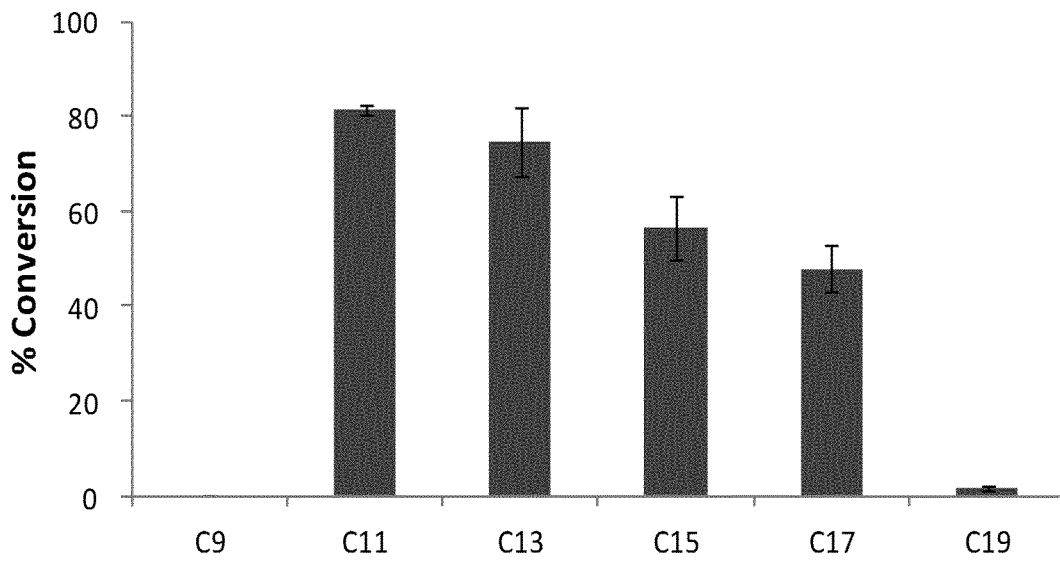


FIG. 6

10/10

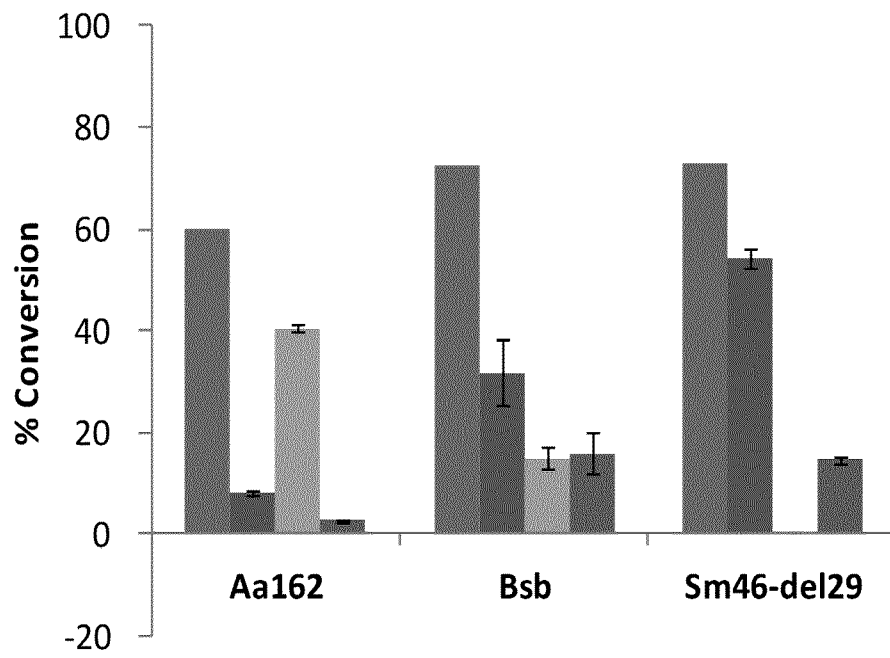


FIG. 7

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2016/065389

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12P5/02 C12N9/88
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)
C12P C12N C10M
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, Sequence Search, EMBASE, WPI Data

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] 27 May 2015 (2015-05-27), "SubName: Full=Cytochrome P450 family protein {ECO:0000313 EMBL:EKU50422.1}";", XP002761949, retrieved from EBI accession no. UNIPROT:K9B6J8	16-22
A	Database accession no. K9B6J8 the whole document ----- -/--	1-12,23

Further documents are listed in the continuation of Box C.

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 21 November 2016	Date of mailing of the international search report 29/11/2016
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Authorized officer Schneider, Patrick

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2016/065389

C(Continuation). 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] 10 February 2009 (2009-02-10), "SubName: Full=Cytochrome P450 {ECO:0000313 EMBL:EED06866.1}";", XP002764068, retrieved from EBI accession no. UNIPROT:B7DSS6 Database accession no. B7DSS6	17,18, 20-22
A	the whole document	1-12,23
X	US 2012/253088 A1 (ALIBHAI MURTAZA F [US] ET AL) 4 October 2012 (2012-10-04) paragraphs [0066], [0069]; examples 16,17; sequences 13,14	1-3,5-12
X	US 2013/130952 A1 (LUO SHUJI [US] ET AL) 23 May 2013 (2013-05-23) the whole document	13-15, 24-27
A	YI LIU ET AL: "Hydrogen peroxide-independent production of ?-alkenes by OleTJE P450 fatty acid decarboxylase", BIOTECHNOLOGY FOR BIOFUELS, BIOMED CENTRAL LTD, GB, vol. 7, no. 1, 24 February 2014 (2014-02-24), page 28, XP021179059, ISSN: 1754-6834, DOI: 10.1186/1754-6834-7-28 the whole document	1-27
A	M. A. RUDE ET AL: "Terminal Olefin (1-Alkene) Biosynthesis by a Novel P450 Fatty Acid Decarboxylase from Jeotgalicoccus Species", APPLIED AND ENVIRONMENTAL MICROBIOLOGY, vol. 77, no. 5, 7 January 2011 (2011-01-07), pages 1718-1727, XP055231228, US ISSN: 0099-2240, DOI: 10.1128/AEM.02580-10 the whole document	1-27
	----- -/--	

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2016/065389

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>MENDEZ-PEREZ D ET AL: "Modular synthase-encoding gene involved in [alpha]-olefin biosynthesis in Synechococcus sp. strain PCC 7002", APPLIED AND ENVIRONMENTAL MICROBIOLOGY, AMERICAN SOCIETY FOR MICROBIOLOGY, US, vol. 77, no. 12, 29 April 2011 (2011-04-29), pages 4264-4267, XP002715053, ISSN: 1098-5336, DOI: 10.1128/AEM.00467-11 [retrieved on 2011-04-29] the whole document</p> <p style="text-align: center;">-----</p>	1-27

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2016/065389

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 4, 16, 19(completely); 1-3, 5-12, 17, 18, 20-23(partially)

Method to produce C7-C11 alpha-olefins using a recombinant host cell expressing the decarboxylase SEQ ID NO:8 from *Staphylococcus massiliensis* encoded by SEQ ID NO:7 and its N-terminally truncated version SEQ ID NO:13 (29 amino acids missing) encoded by the codon-optimized SEQ ID NO:10, using C8-C12 free fatty acids as substrate.

- 1.1. claims: 1-3, 5-12(all partially)

Same as invention 1 for the use of the decarboxylase SEQ ID NO:4 from *Bacillus subtilis* encoded by SEQ ID NO:3.

2. claims: 1-3, 5-12, 17, 18, 20-23(all partially)

Same as invention 1 for SEQ ID NOs: 5 and 6, and an isolated nucleic acid having SEQ ID NO:5, a corresponding host cell and its use to produce lubricants.

3. claims: 13-15, 24-27

A method for producing poly-alpha-olefins comprising producing olefins according to inventions 1 to 3 and oligomerizing said olefin and a composition or lubricant comprising said poly-alpha-olefin.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2016/065389

Patent document cited in search report	Publication date	Patent family member(s)	Publication date	
US 2012253088	A1	04-10-2012	US 2011196180 A1	11-08-2011
			US 2012253088 A1	04-10-2012
			WO 2009085278 A1	09-07-2009

US 2013130952	A1	23-05-2013	NONE	
