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

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
10 October 2013 (10.10.2013)   WIPO | PCT

(10) International Publication Number
WO 2013/152051 A2

(71) Applicant: LS9, INC. [US/US]; 600 Gateway Boulevard, South San Francisco, CA 94080 (US).

(72) Inventors; and
(71) Applicants : GREENFIELD, Derek, L. [US/US]; 600 Gateway Boulevard, South San Francisoc, CA 94080 (US). SCHIRMER, Andreas, W. [US/US]; 600 Gateway Boulevard, South San Francisco, CA 94080 (US). CLARKE, Elizabeth, J. [US/US]; 600 Gateway Boulevard, South San Francisco, CA 94080 (US). GROBAN, Eli, S. [US/US]; 600 Gateway Boulevard, South San Francisco, CA 94080 (US). DA COSTA, Bernardo, M. [US/US]; 600 Gateway Boulevard, South San Francisco, CA 94080 (US). HU, Zhihao [US/US]; 600 Gateway Boulevard, South San Francisco, CA 94080 (US).
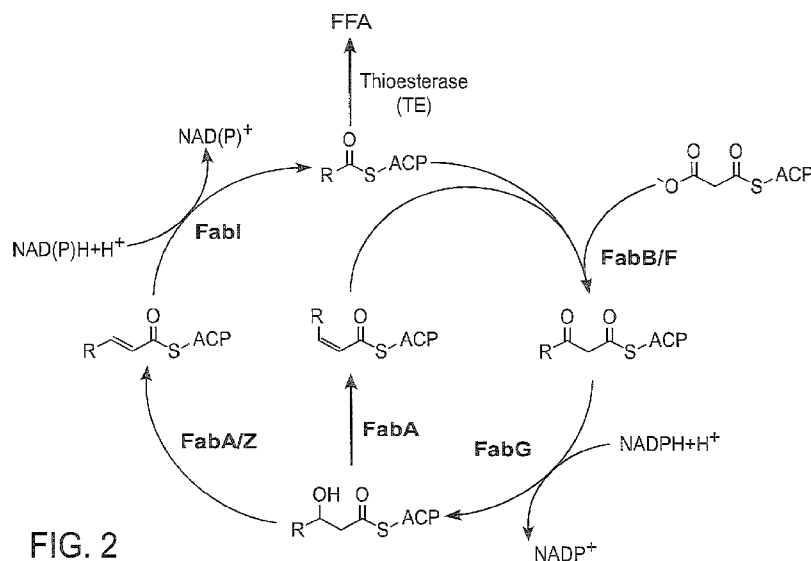
(74) Agent: HAJOS, Brigitte; LS9, Inc., 600 Gateway Boulevard, South San Francisco, CA 94080 (US).

[Continued on next page]

(54) Title: IMPROVED PRODUCTION OF FATTY ACID DERIVATIVES

FIG. 2

(57) Abstract: The disclosure relates to recombinant host cells including strain modifications effective to improve titer, yield and/or productivity of fatty acid derivatives. The disclosure further relates to cell cultures including the recombinant host cells for the fermentative production of fatty acid derivatives and compositions thereof.

WO 2013/152051 A2

WO 2013/152051 A2

IMPROVED PRODUCTION OF FATTY ACID DERIVATIVES

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001]   This application claims the benefit of U.S. Provisional Application No. 61/619,324 filed April 2, 2012, and is hereby incorporated by reference in its entirety.

SEQUENCE LISTING

[0002]   The instant application contains a Sequence Listing which has been submitted in ASCII format via EFS-Web and is hereby incorporated by reference in its entirety. Said ASCII copy, created on April 2, 2013, is named LS00042PCT_SL.txt and is 143,098 bytes in size.

FIELD

[0003]   The disclosure relates to recombinant host cells including strain modifications effective to improve titer, yield and/or productivity of fatty acid derivatives. The disclosure further relates to cell cultures including the recombinant host cells for the fermentative production of fatty acid derivatives and compositions thereof.

BACKGROUND

[0004]   Fatty acid derivatives including fatty aldehydes, fatty alcohols, hydrocarbons (alkanes and olefins), fatty esters (e.g., waxes, fatty acid esters, or fatty esters), and ketones denote important categories of industrial chemicals and fuels. These molecules and their derivatives have numerous applications including, but not limited to, use as surfactants, lubricants, plasticizers, solvents, emulsifiers, emollients, thickeners, flavors, fragrances, and fuels. Crude petroleum is currently a primary source of raw materials for producing petrochemicals and fuels. The two main classes of raw materials derived from petroleum are short chain olefins (e.g., ethylene and propylene) and aromatics (e.g., benzene and xylene isomers). These raw materials are derived from longer chain hydrocarbons in crude petroleum by cracking it at considerable expense using a variety of methods, such as catalytic cracking, steam cracking, or catalytic reforming. These raw materials can be used to make petrochemicals such as monomers, solvents, detergents, and adhesives, which otherwise cannot be directly refined from crude petroleum. Petrochemicals, in turn, can be used to make specialty chemicals, such as plastics, resins, fibers, elastomers, pharmaceuticals, lubricants, gels, and the like. Particular specialty

chemicals that can be produced from petrochemical raw materials include, but are not limited to, fatty acids, hydrocarbons, fatty aldehydes, fatty alcohols, esters, and ketones.

[0005]   Hydrocarbons, for example, have many commercial uses. As such, shorter chain alkanes and alkenes are used in transportation fuels. Longer chain alkenes are used in plastics, lubricants, and synthetic lubricants. In addition, alkenes are used as a feedstock to produce alcohols, esters, plasticizers, surfactants, tertiary amines, enhanced oil recovery agents, fatty acids, thiols, alkenylsuccinic anhydrides, epoxides, chlorinated alkanes, chlorinated alkenes, waxes, fuel additives, and drag flow reducers. Similarly, esters have many commercial uses. For example, biodiesel, an alternative fuel, is made of esters (*e.g.,* fatty acid methyl ester, fatty acid ethyl esters, *etc.*). Some low molecular weight esters are volatile with a pleasant odor which makes them useful as fragrances or flavoring agents. In addition, esters are used as solvents for lacquers, paints, and varnishes. Furthermore, some naturally occurring substances, such as waxes, fats, and oils are also made of esters. Esters are further used as softening agents in resins and plastics, plasticizers, flame retardants, and additives in gasoline and oil. In addition, esters can be used in the manufacture of polymers, films, textiles, dyes, and pharmaceuticals.

[0006]   Aldehydes are used to produce a large number of specialty chemicals. For example, aldehydes are used to produce polymers, resins (*e.g.,* Bakelite), dyes, flavorings, plasticizers, perfumes, pharmaceuticals, and other chemicals, some of which may be used as solvents, preservatives, or disinfectants. In addition, certain natural and synthetic compounds, such as vitamins and compounds used as hormones are aldehydes. Furthermore, many sugars contain aldehyde groups. Fatty aldehydes can be converted to fatty alcohols by chemical or enzymatic reduction. Similarly, fatty alcohols have many commercial uses as well. The shorter chain fatty alcohols are used in the cosmetic and food industries as emulsifiers, emollients, and thickeners. Due to their amphiphilic nature, fatty alcohols behave as nonionic surfactants, which are useful in personal care and household products, such as, for example, detergents. In addition, fatty alcohols are used in waxes, gums, resins, pharmaceutical salves and lotions, lubricating oil additives, textile antistatic and finishing agents, plasticizers, cosmetics, industrial solvents, and solvents for fats. Fatty alcohols such as aliphatic alcohols include a chain of 8 to 22 carbon atoms. Fatty alcohols usually have an even number of carbon atoms and a single alcohol group (-OH) attached to the terminal carbon. Some are unsaturated and some are branched. They are widely used in industrial chemistry. Most fatty alcohols in nature are found as waxes which are

esters with fatty acids and fatty alcohols. They are produced by bacteria, plants and animals. Currently, fatty alcohols are produced via catalytic hydrogenation of fatty acids produced from natural sources, such as coconut oil, palm oil, palm kernel oil, tallow and lard, or by chemical hydration of alpha-olefins produced from petrochemical feedstocks. Fatty alcohols derived from natural sources have varying chain lengths. The chain length of fatty alcohols is important and specific to particular applications. Dehydration of fatty alcohols to alpha-olefins can also be accomplished by chemical catalysis.

[0007]   Due to the inherent challenges posed by exploring, extracting, transporting and refining petroleum for use in chemical- and fuel products, there is a need in the art for a an alternate source which can be produced economically and efficiently for the use of chemical- and fuel production. Moreover, the burning of petroleum-based fuels has become a serious hazard to the environment, especially in light of the ever increasing population inhabiting the planet. Thus, there is a need for a petroleum replacement that does not cause the type of environmental damage created by exploring, extracting, transporting and refining petroleum.

[0008]   One option of producing renewable petroleum is by engineering host cells to produce renewable petroleum products. Biologically derived fuels and chemicals offer advantages over petroleum based fuels. Biologically derived chemicals such as hydrocarbons (*e.g.,* alkanes, alkenes, or alkynes), fatty alcohols, esters, fatty acids, fatty aldehydes, and ketones are directly converted from biomass to the desired chemical product. However, in order for the use of biologically-derived fatty acid derivatives from fermentable sugars or biomass to be commercially viable as a source for production of renewable chemicals and fuels, the process must be optimized for efficient conversion and recovery of product. The development of biologically derived fuels and chemicals has been one focus of research and development in recent years. Still, there remains a considerable need for improvements in the relevant processes and products in order for biologically-derived fuels and chemicals to become a commercially viable option. Areas that need improvement include the energy efficiency of the production process and the final product yield. The current disclosure addresses this need.

SUMMARY

[0009]   One aspect of the disclosure provides a recombinant host cell having a genetically engineered polynucleotide sequence, wherein the polynucleotide sequence codes for one or more polypeptides that have a specific enzymatic activity. The polynucleotide sequence is exogenous or endogenous to the host cell. As such, the disclosure provides a recombinant host cell having a genetically engineered polynucleotide sequence encoding one or more polypeptides, wherein the polypeptides have activity selected from the group including, but not limited to, 3-hydroxydecanoyl-[acp] dehydratase (E.C. 4.2.1.60) activity; ß-ketoacyl-ACP synthase I (E.C. 2.3.1.41) activity; ß -ketoacyl-ACP synthase II (E.C. 2.3.1.179) activity; [acp] S-malonyltransferase {malonyl-CoA-ACP transacylase} (E.C. 2.3.1.39) activity; 3-oxoacyl-{ß –ketoacyl}-ACP reductase (E.C. 1.1.1.100) activity; ß -ketoacyl-ACP synthase III (E.C. 2.3.1.180) activity; enoyl-ACP reductase (NADH) (E.C. 1.3.1.9) activity; enoyl-ACP reductase (NADPH) (E.C. 1.3.1.10) activity; 3-hydroxy-acyl-[acp] dehydratase (E.C. 4.2.1.59) activity; and trans-2, cis-3-decenoyl-ACP isomerase (E.C. 5.3.3.14) activity, wherein the recombinant host cell produce a fatty acid derivative composition at a higher titer, yield or productivity than a corresponding wild type host cell when cultured in a medium containing a carbon source under conditions effective to express the polynucleotide. In a related aspect, the recombinant host cell produces the fatty acid derivative composition at a higher titer, yield and/or productivity when the polypeptide is expressed in combination with at least one other polypeptide of the enzymatic activity. In another aspect, the recombinant host cell produces the fatty acid derivative composition at a higher titer, yield or productivity when the polypeptide is expressed in combination with at least five other polypeptides of the enzymatic activity. In yet another aspect, the recombinant host cell produces the fatty acid derivative composition at a higher titer, yield or productivity when expressed in combination with at least two or three or four or five or six or more polypeptides of the enzymatic activity. In another related aspect, the recombinant host cell includes one or more genetically engineered polynucleotide sequences that further code for a polypeptide that is an acyl carrier protein (ACP). ACP can be in expressed in combination with one or more of the polypeptides that code for any of the enzymatic activities, wherein the ACP further increases the titer, yield and/or productivity of the recombinant host cell when cultured under appropriate conditions. In yet another related aspect, a genetically engineered polynucleotide sequence further encodes a polypeptide that has accABCD activity (E.C. 6.4.1.2).

accABCD can be in expressed in combination with one or more of the polypeptides that code for any of the enzymatic activities, wherein the accABCD further increases the titer, yield and/or productivity of the recombinant host cell when cultured under appropriate conditions.

[0010]   Another aspect of the disclosure provides a recombinant host cell having a genetically engineered polynucleotide sequence encoding one or more polypeptides, wherein the polypeptides have enzymatic activity including, but not limited to, trans-2, cis-3-decenoyl-ACP isomerase activity (fabA or fabM); ß-ketoacyl-ACP synthase I (fabB); malonyl-CoA-ACP transacylase (fabD); ß -ketoacyl-ACP synthase I (fabF or fabB); ß -ketoacyl-ACP reductase (fabG); ß -ketoacyl-ACP synthase III (fabH); enoyl-ACP reductase (fabI or fabL or fabV or fabK); and 3-hydrox-acyl-[acp] dehydratase (fabA or fabZ); trans-2-enoyl-ACP reductase II (fabK). In a related aspect, the polypeptide is selected from fabA, fabB, fabD, fabF, fabG, fabH, fabI, fabL, fabV, fabZ, fabM, and fabK and or combinations thereof. In yet another related aspect, the polypeptide is selected from FabA from *Salmonella typhimurium* (NP_460041); FabB from *Escherichia coli* (NP_416826); FabD from *Salmonella typhimurium* (NP_460164); FabG from *Salmonella typhimurium* (NP_460165); FabH from *Salmonella typhimurium* (NP_460163); FabZ from *Salmonella typhimurium* (NP_459232); FabM from *Streptococcus mutans* (AAN59379); FabK from *Streptococcus pneumoniae* (AAF98273); FabV from *Vibrio cholera* (YP_001217283); FabF from *Clostridium acetobutylicum* (NP_350156); FabI *from Bacillus subtillis subsp. subtilis str. 168* (NP_389054); FabL from *Bacillus subtillis subsp. subtilis str. 168* (NP_388745); FabI from *Acinetobacter sp. ADP1* (YP_047630); FabI from *Marinobacter aquaeoli VT8* (YP_958813); FabI from *Rhodococcus opacus B4* (YP_002784194); FabH from *Acinetobacter sp. ADP1* (YP_046731); FabH from *Marinobacter aquaeoli VT8* (YP_958649); and FabH from *Rhodococcus opacus B4* (YP_00278448) or combinations thereof.

[0011]   The disclosure further contemplates a recombinant host cell having a genetically engineered polynucleotide sequence encoding an ACP polypeptide, wherein the recombinant host cell produces a fatty acid derivative composition at a higher titer, yield or productivity than a corresponding wild type host cell when cultured in a medium containing a carbon source under conditions effective to express the ACP polypeptide. In a related aspect, the genetically engineered polynucleotide sequence further encodes a polypeptide that has phosphopantetheinyl transferase (E.C. 2.7.8.7) activity. Herein, the genetically engineered polynucleotide sequence includes a *sfp* gene coding encoding a phosphopantetheinyl transferase (E.C. 2.7.8.7). In a

related aspect, a genetically engineered polynucleotide sequence further encodes a polypeptide that has accABCD activity (E.C. 6.4.1.2). ACP can be in expressed in combination with accABCD and/or a phosphopantetheinyl transferase, wherein the combination of any of the expressed polypeptides further leads to increases in the titer, yield and/or productivity of the recombinant host cell when cultured under appropriate conditions. In another related aspect, ACP is derived from the same organism as a terminal pathway enzyme expressed in the recombinant host cell, wherein the terminal enzyme cleaves any acyl-ACP species that is part of the fatty acid biosynthetic pathway. The ACP is exogenous or endogenous to the host cell.

[0012]   The disclosure further encompasses a recombinant host cell including a genetically engineered polynucleotide sequence including a transposon, wherein insertion of the transposon into a *yijP* gene affects a second gene flanking the *yijP* gene, wherein the second gene codes for a polynucleotide that is up- or down regulated, and wherein the up- or down regulated polynucleotide codes for a polypeptide that affects production of a fatty acid derivative composition when the host cell is cultured in a medium containing a carbon source under conditions effective to express the polypeptide. The *yijP* gene can be flanked by genes on either side. In a related aspect, the insertion of the transposon into the *yijP* gene results in inactivation of the *yijP* gene or a polynucleotide thereof, which affects one or more of the genes flanking the *yijP* gene, wherein the flanking gene or genes code for a polypeptide that affects production of a fatty acid derivative composition when the host cell is cultured in a medium containing a carbon source under conditions effective to express the polypeptide. In one related aspect, the flanking gene includes polynucleotides including, but not limited to, *ppc, yijO, frwD, pflC, pflD* or *argE*.

[0013]   Another aspect of the disclosure provides a recombinant host cell including a genetically engineered polynucleotide sequence encoding a phosphoenolpyruvate carboxylase (ppc) polypeptide, wherein the recombinant host cell produces a fatty acid derivative composition at a higher titer, yield or productivity than a corresponding wild type host cell when cultured in a medium containing a carbon source under conditions effective to express the ppc polypeptide.

[0014]   Still, another aspect of the disclosure provides a cell culture that includes any of the recombinant host cells presented herein (*supra*). The recombinant host cell is cultured in a medium such that the recombinant host cell produces fatty acid derivative compositions according to the genetic engineering methods presented herein (*supra*). In a related aspect, the

fatty acid derivative compositions produced by the recombinant host cells of the present disclosure include, but are not limited to, fatty acids, fatty esters, fatty alcohols, fatty aldehydes, alkanes, terminal olefins, internal olefins, and ketones. In another related aspect, the fatty acid derivative is a C6, C8, C10, C12, C13, C14, C15, C16, C17, or C18 fatty acid derivative. In yet another related aspect, the fatty acid derivative is a C10:1, C12:1, C14:1, C16:1, or C18:1 unsaturated fatty acid derivative. In a further related aspect, the fatty acid derivative composition comprises one or more of C8, C10, C12, C14, C16, and C18 fatty acid derivatives. The fatty acid derivative compositions produced by the cell cultures containing the recombinant host cells of the present disclosure include fatty acids, fatty aldehydes, fatty alcohols, fatty esters, alkanes, terminal olefins, internal olefins, and ketones. The disclosure further encompasses fatty acid derivative compositions that include fatty acid derivatives having a double bond at position 7 in the carbon chain between C7 and C8 from the reduced end of the fatty alcohol; fatty acid derivative compositions including unsaturated fatty acid derivatives; fatty acid derivative compositions including saturated fatty acid derivatives; and fatty acid derivative compositions including branched chain fatty acid derivatives.

[0015]   The disclosure further contemplates a cell culture containing any of the recombinant host cells presented herein, wherein the recombinant host cells have a titer that is at least about 5% greater than the titer of the corresponding wild type host cells when cultured under the same conditions as the recombinant host cells. Herein, the recombinant host cells have a titer of from about 1g/L to about 250g/L, and more specifically from about 90 g/L to about 120g/L. In a related aspect, the recombinant host cells have a yield that is at least about 10% to about 40%. In one aspect, the recombinant host cells have a yield of about 25%. Still encompassed herein is a cell culture containing any one of the recombinant host cells presented herein, wherein the productivity of the cell culture ranges from about 0.7mg/L/hr to about 3g/L/hr or higher.

[0016]   Another aspect of the disclosure provides methods of making a recombinant host cell, including genetically engineering the recombinant host cell such that the cell expresses a polypeptide sequence that is encoded by one or more polynucleotide sequences under specific culture conditions, wherein the polynucleotide sequence codes for one or more polypeptides that have a specific enzymatic activity.

BRIEF DESCRIPTION OF THE DRAWINGS

[0017]   The present disclosure is best understood when read in conjunction with the accompanying figures, which serve to illustrate the preferred embodiments. It is understood, however, that the disclosure is not limited to the specific embodiments disclosed in the figures.

[0018]   **Figure 1** presents an exemplary biosynthetic pathway for use in production of acyl CoA as a precursor to fatty acid derivatives in a recombinant microorganism. The cycle is initiated by condensation of malonyl-ACP and acetyl-CoA.

[0019]   **Figure 2** presents an exemplary fatty acid biosynthetic cycle, where malonyl-ACP is produced by the transacylation of malonyl-CoA to malonyl-ACP (catalyzed by malonyl-CoA:ACP transacylase (fabD)); then β-ketoacyl-ACP synthase III (fabH) initiates condensation of malonyl-ACP with acetyl-CoA. Elongation cycles begin with the condensation of malonyl-ACP and an acyl-ACP catalyzed by β-ketoacyl-ACP synthase I (fabB) and β-ketoacyl-ACP synthase II (fabF) to produce a β-keto-acyl-ACP, then the β-keto-acyl-ACP is reduced by β-ketoacyl-ACP reductase (fabG) to produce a β-hydroxy-acyl-ACP, which is dehydrated to a *trans*-2-enoyl-acyl-ACP by β-hydroxyacyl-ACP dehydratase (fabA or fabZ). FabA can also isomerize *trans*-2-enoyl-acyl-ACP to *cis*-3-enoyl-acyl-ACP, which can bypass fabI and can used by fabB (typically for up to an aliphatic chain length of C16) to produce β-keto-acyl-ACP. The final step in each cycle is catalyzed by enoyl-ACP reductase (fabI) that converts *trans*-2-enoyl-acyl-ACP to acyl-ACP. In the methods described herein, termination of fatty acid synthesis occurs by thioesterase removal of the acyl group from acyl-ACP to release free fatty acids (FFA). Thioesterases (*e.g., tesA*) hydrolyze thioester bonds, which occur between acyl chains and ACP through sulfydryl bonds.

[0020]   **Figure 3** illustrates the structure and function of the acetyl-CoA carboxylase (accABCD) enzyme complex. BirA biotinylates accB, the biotin carboxyl carrier protein, which is part of the acetyl-CoA carboxylase enzyme complex.

[0021]   **Figure 4** presents an overview of an exemplary biosynthetic pathway for production of fatty alcohol starting with acyl-ACP, where the production of fatty aldehyde is catalyzed by the enzymatic activity of acyl-ACP reductase (AAR) or thioesterase (TE) and carboxylic acid reductase (Car). The fatty aldehyde is converted to fatty alcohol by aldehyde reductase (also referred to as alcohol dehydrogenase).

8

[0022]   **Figure 5** presents an overview of two exemplary biosynthetic pathways for production of fatty esters starting with acyl-ACP, where the production of fatty esters is accomplished by a one-enzyme system or a three-enzyme-system.

[0023]   **Figure 6** presents an overview of exemplary biosynthetic pathways for production of hydrocarbons starting with acyl-ACP; the production of internal olefins is catalyzed by the enzymatic activity of OleABCD; the production of alkanes is catalyzed by the enzymatic conversion of fatty aldehydes to alkanes by way of aldehyde decarbonylase (ADC); and the production of terminal olefins is catalyzed by the enzymatic conversion of fatty acids to terminal olefins by a decarboxylase.

[0024]   **Figure 7** illustrates fatty acid derivative (Total Fatty Species) production by the MG1655 *E. coli* strain with the fadE gene attenuated (*i.e.,* deleted) compared to fatty acid derivative production by *E. coli* MG1655. The data presented in Fig. 7 shows that attenuation of the fadE gene did not affect fatty acid derivative production

[0025]   **Figure 8** shows malonyl-CoA levels in DAM1_i377 in log phase, expressing eight different C. *glutamicum* acetyl-CoA carboxylase (Acc) operon constructs.

[0026]   **Figure 9** shows intracellular short chain-CoA levels in *E. coli* DAM1_i377 in log phase expressing ptrc1/3_accDACB-birA±panK operon constructs. accDACB+birA is also referred to herein as accD+.

[0027]   **Figure 10** shows fatty acid methyl ester (FAME) production in *E. coli* strain DV2 expressing ester synthase 9 from *M. hydrocarbonoclasticus* and components of an acetyl-CoA carboxylase complex from *C. glutamicum*.

[0028]   **Figure 11** shows production of fatty alcohols by *E. coli* expressing the *Synechococcus elongatus* PCC7942 AAR together with the accD+ operon from *C.glutamicum* on a pCL plasmid. Triplicate samples are shown for the accD+ strains.

[0029]   **Figures 12A and 12B** show data for production of Total Fatty Species (mg/L) from duplicate plate screens when plasmid pCL_P$_{trc}$_tesA was transformed into each of the iFAB-containing strains shown in the figures and a fermentation was run in FA2 media with 20 hours from induction to harvest at both 32°C (Figure 12A) and 37°C (Figure 12B).

[0030]   **Figure 13** shows FAME production of *E. coli* DAM1with plasmid pDS57 and integrated fabHI operons. The fabH/I genes are from *Marinobacter aquaeoli* VT8 or from *Acinetobacter baylyi* ADP1. See Table 7 for a more details on the fabH/I operons in these strains.

[0031]   **Figure 14** shows FAME production of *E. coli* DAM1with plasmid pDS57 and different configurations of the *C. glutamicum* acc genes as well as integrated fabHI operons. The strains contain the fabH/I genes from *Rhodococcus opacus* or *Acinetobacter baylyi* ADP1. See Table 7 for more details on the fabH/I and acc operons.

[0032]   **Figure 15** shows FAME and FFA titers of two *E. coli* DAM1 pDS57 strains with integrated fabH/I genes strains selected from figure 13 compared to the control strain *E. coli* DAM1 pDS57.

[0033]   **Figure 16** is a diagrammatic depiction of the iFAB138 locus, including a diagram of cat-loxP-P_{T5} cassette integrated in front of iFAB138 (Figure 16A); and a diagram of the P_{T5}_iFAB138 region (Figure 16B).

[0034]   **Figure 17** shows that strain V668, which has the rph and ilvG genes repaired, produced a higher level of FFA than EG149, which has neither of the genes repaired.

[0035]   **Figure 18** is a diagrammatic depiction of a transposon cassette insertion in the yijP gene of strain LC535 (transposon hit 68F11). Promoters internal to the transposon cassette are shown, and may have effects on adjacent gene expression.

[0036]   **Figure 19** illustrates fatty alcohol production in *E. coli* DV2 expressing *Synechococcus elongatus* acyl-ACP reductase (AAR) and coexpressing various cyanobacterial acyl carrier proteins (ACPs). Details regarding the source of the ACPs are provided in Table 12.

[0037]   **Figure 20** illustrates fatty acid production in *E. coli* DV2 expressing leaderless *E. coli* *thioesterase* 'tesA and coexpressing a cyanobacterial acyl carrier protein (cACP) and *B. subtilis* sfp.

DETAILED DESCRIPTION

[0038]   **General Overview**

[0039]   The disclosure is based, at least in part, on the discovery that modification of various aspects of the fatty acid biosynthetic pathway in a recombinant host cell facilitates enhanced production of fatty acid derivatives by the host cell. The disclosure relates to compositions of fatty acid derivatives having desired characteristics and methods for producing the same. Further, the disclosure relates to recombinant host cells (*e.g.,* microorganisms), cultures of recombinant host cells, methods of making and using recombinant host cells, for example, use of cultured recombinant host cells in the fermentative production of fatty acid derivatives having desired characteristics.

[0040]   More specifically, the production of a desired fatty acid derivative composition (*e.g.,* acyl-CoA, fatty acids, fatty aldehydes, short and long chain alcohols, hydrocarbons, fatty alcohols, esters (*e.g.,* waxes, fatty acid esters, or fatty esters), terminal olefins, internal olefins, and ketones is enhanced by modifying the expression of one or more genes involved in a biosynthetic pathway for fatty acid, fatty ester, alkane, alkene, olefin, or fatty alcohol, production, degradation and/or secretion. The disclosure provides recombinant host cells which have been engineered to provide enhanced fatty acid biosynthesis relative to non-engineered or native host cells (*e.g.,* wild type host cells that function as control cells), which is accomplished, for example, through strain improvements. As such, the disclosure identifies polynucleotides useful in the recombinant host cells, methods, and compositions of the disclosure. It will be generally recognized that absolute sequence identity to such polynucleotides is not necessary. For example, changes in a particular polynucleotide sequence can be made and the encoded polypeptide screened for activity. Such changes typically comprise conservative mutations and silent mutations (*e.g.,* codon optimization). Genetically engineered or modified polynucleotides and encoded variant polypeptides can be screened for a desired function, including but not limited to, increased catalytic activity, increased stability, or decreased inhibition (*e.g.,* decreased feedback inhibition), using methods known in the art.

[0041]   The disclosure identifies enzymatic activities involved in various steps (*i.e.,* reactions) of the fatty acid biosynthetic pathways described herein according to Enzyme Classification (EC) number, and provides exemplary polypeptides (e.g.,enzymes) categorized by

11

such EC numbers, and exemplary polynucleotides encoding such polypeptides. Such exemplary polypeptides and polynucleotides, which are identified herein by Accession Numbers and/or Sequence Identifier Numbers (SEQ ID NOs), are useful for engineering fatty acid pathways in parental host cells to obtain the recombinant host cells described herein. The polypeptides and polynucleotides described herein are exemplary and non-limiting. The sequences of homologues of exemplary polypeptides described herein are available to those of skill in the art through various databases (e.g., rhw Entrez databases provided by the National Center for Biotechnology Information (NCBI) , the ExPasy databases provided by the Swiss Institute of Bioinformatics, the BRENDA database provided by the Technical University of Braunschweig, and the KEGG database provided by the Bioinformatics Center of Kyoto University and University of Tokyo, all which are available on the World Wide Web).

[0042]   **Definitions**

[0043]   As used in this specification and the appended claims, the singular forms "a," "an" and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a recombinant host cell" includes two or more such recombinant host cells, reference to "a fatty alcohol" includes one or more fatty alcohols, or mixtures of fatty alcohols, reference to "a nucleic acid coding sequence" includes one or more nucleic acid coding sequences, reference to "an enzyme" includes one or more enzymes, and the like.

[0044]   Accession Numbers: Sequence Accession numbers throughout this description were obtained from databases provided by the NCBI (National Center for Biotechnology Information) maintained by the National Institutes of Health, U.S.A. (which are identified herein as "NCBI Accession Numbers" or alternatively as "GenBank Accession Numbers") , and from the UniProt Knowledgebase (UniProtKB) and Swiss-Prot databases provided by the Swiss Institute of Bioinformatics (which are identified herein as "UniProtKB Accession Numbers").

[0045]   Enzyme Classification (EC) Numbers: EC numbers are established by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (IUBMB), description of which is available on the IUBMB Enzyme Nomenclature website on the World Wide Web. EC numbers classify enzymes according to the reaction they catalyze.

[0046]   As used herein, the term "nucleotide" refers to a monomeric unit of a polynucleotide that consists of a heterocyclic base, a sugar, and one or more phosphate groups. The naturally

occurring bases (guanine, (G), adenine, (A), cytosine, (C), thymine, (T), and uracil (U)) are typically derivatives of purine or pyrimidine, though it should be understood that naturally and non-naturally occurring base analogs are also included. The naturally occurring sugar is the pentose (five-carbon sugar) deoxyribose (which forms DNA) or ribose (which forms RNA), though it should be understood that naturally and non-naturally occurring sugar analogs are also included. Nucleic acids are typically linked via phosphate bonds to form nucleic acids or polynucleotides, though many other linkages are known in the art (e.g., phosphorothioates, boranophosphates, and the like).

[0047]   As used herein, the term "polynucleotide" refers to a polymer of ribonucleotides (RNA) or deoxyribonucleotides (DNA), which can be single-stranded or double-stranded and which can contain non-natural or altered nucleotides. The terms "polynucleotide," "nucleic acid sequence," and "nucleotide sequence" are used interchangeably herein to refer to a polymeric form of nucleotides of any length, either RNA or DNA. These terms refer to the primary structure of the molecule, and thus include double- and single-stranded DNA, and double- and single-stranded RNA. The terms include, as equivalents, analogs of either RNA or DNA made from nucleotide analogs and modified polynucleotides such as, though not limited to methylated and/or capped polynucleotides. The polynucleotide can be in any form, including but not limited to, plasmid, viral, chromosomal, EST, cDNA, mRNA, and rRNA.

[0048]   As used herein, the terms "polypeptide" and "protein" are used interchangeably to refer to a polymer of amino acid residues. The term "recombinant polypeptide" refers to a polypeptide that is produced by recombinant techniques, wherein generally DNA or RNA encoding the expressed protein is inserted into a suitable expression vector that is in turn used to transform a host cell to produce the polypeptide.

[0049]   As used herein, the terms "homolog," and "homologous" refer to a polynucleotide or a polypeptide comprising a sequence that is at least about 50% identical to the corresponding polynucleotide or polypeptide sequence. Preferably homologous polynucleotides or polypeptides have polynucleotide sequences or amino acid sequences that have at least about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or at least about 99% homology to the corresponding amino acid sequence or polynucleotide sequence. As used herein the terms sequence "homology" and sequence "identity" are used interchangeably.

[0050]   One of ordinary skill in the art is well aware of methods to determine homology between two or more sequences. Briefly, calculations of "homology" between two sequences can be performed as follows. The sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). In a preferred embodiment, the length of a first sequence that is aligned for comparison purposes is at least about 30%, preferably at least about 40%, more preferably at least about 50%, even more preferably at least about 60%, and even more preferably at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or about 100% of the length of a second sequence. The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions of the first and second sequences are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent homology between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps and the length of each gap, that need to be introduced for optimal alignment of the two sequences.

[0051]   The comparison of sequences and determination of percent homology between two sequences can be accomplished using a mathematical algorithm, such as BLAST (Altschul et al., J. Mol. Biol., 215(3): 403–410 (1990)). The percent homology between two amino acid sequences also can be determined using the Needleman and Wunsch algorithm that has been incorporated into the GAP program in the GCG software package, using either a Blossum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6 (Needleman and Wunsch, J. Mol. Biol., 48: 444–453 (1970)). The percent homology between two nucleotide sequences also can be determined using the GAP program in the GCG software package, using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. One of ordinary skill in the art can perform initial homology calculations and adjust the algorithm parameters accordingly. A preferred set of parameters (and the one that should be used if a practitioner is uncertain about which parameters should be applied to determine if a molecule is within a homology limitation of the claims) are a Blossum 62 scoring matrix with a gap penalty of 12, a gap extend penalty of 4, and a frameshift

gap penalty of 5. Additional methods of sequence alignment are known in the biotechnology arts (see, e.g., Rosenberg, BMC Bioinformatics, 6: 278 (2005); Altschul, et al., FEBS J., 272(20): 5101–5109 (2005)).

[0052]   As used herein, the term "hybridizes under low stringency, medium stringency, high stringency, or very high stringency conditions" describes conditions for hybridization and washing. Guidance for performing hybridization reactions can be found in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1 - 6.3.6. Aqueous and non-aqueous methods are described in that reference and either method can be used. Specific hybridization conditions referred to herein are as follows: 1) low stringency hybridization conditions -- 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by two washes in 0.2X SSC, 0.1% SDS at least at 50°C (the temperature of the washes can be increased to 55°C for low stringency conditions); 2) medium stringency hybridization conditions -- 6X SSC at about 45°C, followed by one or more washes in 0.2X SSC, 0.1% SDS at 60°C; 3) high stringency hybridization conditions -- 6X SSC at about 45°C, followed by one or more washes in 0.2.X SSC, 0.1% SDS at 65°C; and 4) very high stringency hybridization conditions -- 0.5M sodium phosphate, 7% SDS at 65°C, followed by one or more washes at 0.2X SSC, 1% SDS at 65°C. Very high stringency conditions (4) are the preferred conditions unless otherwise specified.

[0053]   An "endogenous" polypeptide refers to a polypeptide encoded by the genome of the host cell (e.g., parental microbial cell) from which the recombinant cell is engineered or derived.

[0054]   An "exogenous" polypeptide refers to a polypeptide which is not encoded by the genome of the parental microbial cell. A variant (i.e., mutant) polypeptide is an example of an exogenous polypeptide.

[0055]   The term "heterologous" generally means derived from a different species or derived from a different organism. As used herein it refers to a nucleotide sequence or a polypeptide sequence that is not naturally present in a particular organism. Heterologous expression means that a protein or polypeptide is experimentally added to a cell that does not normally express that protein. As such, heterologous refers to the fact that a transferred protein was initially derived from a different cell type or a different species then the recipient. For example, a polynucleotide sequence endogenous to a plant cell can be introduced into a bacterial host cell by recombinant

methods, and the plant polynucleotide is then a heterologous polynucleotide in a recombinant bacterial host cell.

[0056]   As used herein, the term "fragment" of a polypeptide refers to a shorter portion of a full-length polypeptide or protein ranging in size from four amino acid residues to the entire amino acid sequence minus one amino acid residue. In certain embodiments of the disclosure, a fragment refers to the entire amino acid sequence of a domain of a polypeptide or protein (*e.g.,* a substrate binding domain or a catalytic domain).

[0057]   As used herein, the term "mutagenesis" refers to a process by which the genetic information of an organism is changed in a stable manner. Mutagenesis of a protein coding nucleic acid sequence produces a mutant protein. Mutagenesis also refers to changes in non-coding nucleic acid sequences that result in modified protein activity.

[0058]   As used herein, the term "gene" refers to nucleic acid sequences encoding either an RNA product or a protein product, as well as operably-linked nucleic acid sequences affecting the expression of the RNA or protein (*e.g.,* such sequences include but are not limited to promoter or enhancer sequences) or operably-linked nucleic acid sequences encoding sequences that affect the expression of the RNA or protein (*e.g.,* such sequences include but are not limited to ribosome binding sites or translational control sequences).

[0059]   Expression control sequences are known in the art and include, for example, promoters, enhancers, polyadenylation signals, transcription terminators, internal ribosome entry sites (IRES), and the like, that provide for the expression of the polynucleotide sequence in a host cell. Expression control sequences interact specifically with cellular proteins involved in transcription (Maniatis et al., Science, 236: 1237-1245 (1987)). Exemplary expression control sequences are described in, for example, Goeddel, Gene Expression Technology: Methods in Enzymology, Vol. 185, Academic Press, San Diego, Calif. (1990).

[0060]   In the methods of the disclosure, an expression control sequence is operably linked to a polynucleotide sequence. By "operably linked" is meant that a polynucleotide sequence and an expression control sequence(s) are connected in such a way as to permit gene expression when the appropriate molecules (e.g., transcriptional activator proteins) are bound to the expression control sequence(s). Operably linked promoters are located upstream of the selected

polynucleotide sequence in terms of the direction of transcription and translation. Operably linked enhancers can be located upstream, within, or downstream of the selected polynucleotide.

[0061]   As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid, *i.e.,* a polynucleotide sequence, to which it has been linked. One type of useful vector is an episome (*i.e.,* a nucleic acid capable of extra-chromosomal replication). Useful vectors are those capable of autonomous replication and/or expression of nucleic acids to which they are linked. Vectors capable of directing the expression of genes to which they are operatively linked are referred to herein as "expression vectors." In general, expression vectors of utility in recombinant DNA techniques are often in the form of "plasmids," which refer generally to circular double stranded DNA loops that, in their vector form, are not bound to the chromosome. The terms "plasmid" and "vector" are used interchangeably herein, in as much as a plasmid is the most commonly used form of vector. However, also included are such other forms of expression vectors that serve equivalent functions and that become known in the art subsequently hereto. In some embodiments, a recombinant vector further comprises a promoter operably linked to the polynucleotide sequence. In some embodiments, the promoter is a developmentally-regulated, an organelle-specific, a tissue-specific, an inducible, a constitutive, or a cell-specific promoter. The recombinant vector typically comprises at least one sequence including (a) an expression control sequence operatively coupled to the polynucleotide sequence; (b) a selection marker operatively coupled to the polynucleotide sequence; (c) a marker sequence operatively coupled to the polynucleotide sequence; (d) a purification moiety operatively coupled to the polynucleotide sequence; (e) a secretion sequence operatively coupled to the polynucleotide sequence; and (f) a targeting sequence operatively coupled to the polynucleotide sequence. In certain embodiments, the nucleotide sequence is stably incorporated into the genomic DNA of the host cell, and the expression of the nucleotide sequence is under the control of a regulated promoter region. The expression vectors described herein include a polynucleotide sequence described herein in a form suitable for expression of the polynucleotide sequence in a host cell. It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of polypeptide desired, *etc.* The expression vectors described herein can be introduced into host cells to produce polypeptides, including fusion polypeptides, encoded by the polynucleotide sequences as described herein.

[0062]   Expression of genes encoding polypeptides in prokaryotes, for example, *E. coli,* is most often carried out with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion polypeptides. Fusion vectors add a number of amino acids to a polypeptide encoded therein, usually to the amino- or carboxy-terminus of the recombinant polypeptide. Such fusion vectors typically serve one or more of the following three purposes: (1) to increase expression of the recombinant polypeptide; (2) to increase the solubility of the recombinant polypeptide; and (3) to aid in the purification of the recombinant polypeptide by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant polypeptide. This enables separation of the recombinant polypeptide from the fusion moiety after purification of the fusion polypeptide. In certain embodiments, a polynucleotide sequence of the disclosure is operably linked to a promoter derived from bacteriophage T5.

[0063]   In certain embodiments, the host cell is a yeast cell, and the expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari et al., EMBO J., 6: 229-234 (1987)), pMFa (Kurjan et al., Cell, 30: 933-943 (1982)), pJRY88 (Schultz et al., Gene, 54: 113-123 (1987)), pYES2 (Invitrogen Corp., San Diego, CA), and picZ (Invitrogen Corp., San Diego, CA).

[0064]   In other embodiments, the host cell is an insect cell, and the expression vector is a baculovirus expression vector. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.,* Sf9 cells) include, for example, the pAc series (Smith et al., Mol. Cell Biol., 3: 2156-2165 (1983)) and the pVL series (Lucklow et al., Virology, 170: 31-39 (1989)).

[0065]   In yet another embodiment, the polynucleotide sequences described herein can be expressed in mammalian cells using a mammalian expression vector. Other suitable expression systems for both prokaryotic and eukaryotic cells are well known in the art; see, e.g., Sambrook et al., "Molecular Cloning: A Laboratory Manual," second edition, Cold Spring Harbor Laboratory, (1989).

[0066]   The term "corresponding wild type host cell" as referred to herein, means a cell that functions as a control cell. For example, if a polypeptide in a recombinant host cell is up-regulated, then the same polypeptide would exist at a lower level in the control cell. Conversely, if a polypeptide in a recombinant host cell is down-regulated, then the same polypeptide would

exist at a higher level in the control cell. Furthermore, the "recombinant or engineered host cell" is a microorganism used to produce one or more of fatty acid derivatives including, for example, acyl-CoA, fatty acids, fatty aldehydes, short and long chain alcohols, hydrocarbons, fatty alcohols, esters (*e.g.*, waxes, fatty acid esters, or fatty esters), terminal olefins, internal olefins, and ketones. In some embodiments, the recombinant host cell comprises one or more polynucleotides, each polynucleotide encoding a polypeptide having fatty acid biosynthetic enzyme activity.

[0067]   As used herein "acyl-CoA" refers to an acyl thioester formed between the carbonyl carbon of alkyl chain and the sulfhydryl group of the 4'-phosphopantethionyl moiety of coenzyme A (CoA), which has the formula R-C(O)S-CoA, where R is any alkyl group having at least 4 carbon atoms.

[0068]   As used herein "acyl-ACP" refers to an acyl thioester formed between the carbonyl carbon of alkyl chain and the sulfhydryl group of the phosphopantetheinyl moiety of an acyl carrier protein (ACP). The phosphopantetheinyl moiety is post-translationally attached to a conserved serine residue on the ACP by the action of holo-acyl carrier protein synthase (ACPS), a phosphopantetheinyl transferase. In some embodiments an acyl-ACP is an intermediate in the synthesis of fully saturated acyl-ACPs. In other embodiments an acyl-ACP is an intermediate in the synthesis of unsaturated acyl-ACPs. In some embodiments, the carbon chain will have about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, or 26 carbons. Each of these acyl-ACPs are substrates for enzymes that convert them to fatty acid derivatives.

[0069]   As used herein, the term "fatty acid derivative" means a "fatty acid" or a "fatty acid derivative", which may be referred to as a "fatty acid or derivative thereof". The term "fatty acid" means a carboxylic acid having the formula RCOOH. R represents an aliphatic group, preferably an alkyl group. R can comprise between about 4 and about 22 carbon atoms. Fatty acids can be saturated, monounsaturated, or polyunsaturated. A "fatty acid derivative" is a product made in part from the fatty acid biosynthetic pathway of the production host organism. "Fatty acid derivatives" includes products made in part from acyl-ACP or acyl-ACP derivatives. Exemplary fatty acid derivatives include, for example, acyl-CoA, fatty acids, fatty aldehydes, short and long chain alcohols, fatty alcohols, hydrocarbons, esters (*e.g.*, waxes, fatty acid esters, or fatty esters), terminal olefins, internal olefins, and ketones.

[0070]   A "fatty acid derivative composition" as referred to herein is produced by a recombinant host cell and typically comprises a mixture of fatty acid derivative. In some cases, the mixture includes more than one type of product (*e.g.*, fatty acids and fatty alcohols, fatty acids and fatty acid esters or alkanes and olefins). In other cases, the fatty acid derivative compositions may comprise, for example, a mixture of fatty alcohols (or another fatty acid derivative) with various chain lengths and saturation or branching characteristics. In still other cases, the fatty acid derivative composition comprises a mixture of both more than one type of product and products with various chain lengths and saturation or branching characteristics.

[0071]   As used herein, the term "fatty acid biosynthetic pathway" means a biosynthetic pathway that produces fatty acids and derivatives thereof. The fatty acid biosynthetic pathway may include additional enzymes or polypeptides with enzymatic activities besides the ones discussed herein to produce fatty acid derivatives having desired characteristics.

[0072]   As used herein, "fatty aldehyde" means an aldehyde having the formula RCHO characterized by a carbonyl group (C=O). In some embodiments, the fatty aldehyde is any aldehyde made from a fatty alcohol. In certain embodiments, the R group is at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, or at least 19, carbons in length. Alternatively, or in addition, the R group is 20 or less, 19 or less, 18 or less, 17 or less, 16 or less, 15 or less, 14 or less, 13 or less, 12 or less, 11 or less, 10 or less, 9 or less, 8 or less, 7 or less, or 6 or less carbons in length. Thus, the R group can have an R group bounded by any two of the above endpoints. For example, the R group can be 6-16 carbons in length, 10-14 carbons in length, or 12-18 carbons in length. In some embodiments, the fatty aldehyde is a C6, C7, C8, C9, C10, C11, C12, C13, C14, C15, C16, C17, C18, C19, C20, C21, C22, C23, C24, C25, or a C26 fatty aldehyde. In certain embodiments, the fatty aldehyde is a C6, C7, C8, C9, C10, C11, C12, C13, C14, C15, C16, C17, or C18 fatty aldehyde.

[0073]   As used herein, "fatty alcohol" means an alcohol having the formula ROH. In some embodiments, the R group is at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, or at least 19, carbons in length. Alternatively, or in addition, the R group is 20 or less, 19 or less, 18 or less, 17 or less, 16 or less, 15 or less, 14 or less, 13 or less, 12 or less, 11 or less, 10 or less, 9 or less, 8 or less, 7 or less, or 6 or less carbons in length. Thus, the R group can have an R group

bounded by any two of the above endpoints. For example, the R group can be 6-16 carbons in length, 10-14 carbons in length, or 12-18 carbons in length. In some embodiments, the fatty alcohol is a C6, C7, C8, C9, C10, C11, C12, C13, C14, C15, C16, C17, C18, C19, C20, C21, C22, C23, C24, C25, or a C26 fatty alcohol. In certain embodiments, the fatty alcohol is a C6, C7, C8, C9, C10, C11, C12, C13, C14, C15, C16, C17, or C18 fatty alcohol.

[0074]   The R group of a fatty acid derivative, for example a fatty alcohol, can be a straight chain or a branched chain. Branched chains may have more than one point of branching and may include cyclic branches. In some embodiments, the branched fatty acid, branched fatty aldehyde, or branched fatty alcohol is a C6, C7, C8, C9, C10, C11, C12, C13, C14, C15, C16, C17, C18, C19, C20, C21, C22, C23, C24, C25, or a C26 branched fatty acid, branched fatty aldehyde, or branched fatty alcohol. In particular embodiments, the branched fatty acid, branched fatty aldehyde, or branched fatty alcohol is a C6, C7, C8, C9, C10, C11, C12, C13, C14, C15, C16, C17, or C18 branched fatty acid, branched fatty aldehyde, or branched fatty alcohol. In certain embodiments, the hydroxyl group of the branched fatty acid, branched fatty aldehyde, or branched fatty alcohol is in the primary (C1) position.

[0075]   In certain embodiments, the branched fatty acid derivative is an iso-fatty acid derivative, for example an iso-fatty aldehyde, an iso-fatty alcohol, or an antesio-fatty acid derivative, an anteiso-fatty aldehyde, or an anteiso-fatty alcohol. In exemplary embodiments, the branched fatty acid derivative is selected from iso-C7:0, iso-C8:0, iso-C9:0, iso-C10:0, iso-C11:0, iso-C12:0, iso-C13:0, iso-C14:0, iso-C15:0, iso-C16:0, iso-C17:0, iso-C18:0, iso-C19:0, anteiso-C7:0, anteiso-C8:0, anteiso-C9:0, anteiso-C10:0, anteiso-C11:0,anteiso-C12:0, anteiso-C13:0, anteiso-C14:0, anteiso-C15:0, anteiso-C16:0, anteiso-C17:0, anteiso-C18:0, and an anteiso-C19:0 branched fatty alcohol.

[0076]   The R group of a branched or unbranched fatty acid derivative can be saturated or unsaturated. If unsaturated, the R group can have one or more than one point of unsaturation. In some embodiments, the unsaturated fatty acid derivative is a monounsaturated fatty acid derivative. In certain embodiments, the unsaturated fatty acid derivative is a C6:1, C7:1, C8:1, C9:1, C10:1, C11:1, C12:1, C13:1, C14:1, C15:1, C16:1, C17:1, C18:1, C19:1, C20:1, C21:1, C22:1, C23:1, C24:1, C25:1, or a C26:1 unsaturated fatty acid derivative. In certain embodiments, the unsaturated fatty acid derivative is a C10:1, C12:1, C14:1, C16:1, or C18:1 unsaturated fatty acid derivative. In other embodiments, the unsaturated fatty acid derivative is

unsaturated at the omega-7 position. In certain embodiments, the unsaturated fatty acid derivative comprises a cis double bond.

[0077]   As used herein, the term "clone" typically refers to a cell or group of cells descended from and essentially genetically identical to a single common ancestor, for example, the bacteria of a cloned bacterial colony arose from a single bacterial cell.

[0078]   As used herein, the term "culture" typical refers to a liquid media comprising viable cells. In one embodiment, a culture comprises cells reproducing in a predetermined culture media under controlled conditions, for example, a culture of recombinant host cells grown in liquid media comprising a selected carbon source and nitrogen. "Culturing" or "cultivation" refers to growing a population of recombinant host cells under suitable conditions in a liquid or solid medium. In particular embodiments, culturing refers to the fermentative bioconversion of a substrate to an end-product. Culturing media are well known and individual components of such culture media are available from commercial sources, *e.g.,* under the Difco™ and BBL™ trademarks. In one non-limiting example, the aqueous nutrient medium is a "rich medium" comprising complex sources of nitrogen, salts, and carbon, such as YP medium, comprising 10 g/L of peptone and 10 g/L yeast extract of such a medium. The host cell of a culture can be additionally engineered to assimilate carbon efficiently and use cellulosic materials as carbon sources according to methods described in U.S. Patents 5,000,000; 5,028,539; 5,424,202; 5,482,846; 5,602,030; WO 2010127318. In addition, in some embodiments the host cell is engineered to express an invertase so that sucrose can be used as a carbon source.

[0079]   As used herein, the term "under conditions effective to express a genetically engineered polynucleotide sequence" means any condition that allows a host cell to produce a desired fatty acid derivative. Suitable conditions include, for example, fermentation conditions.

[0080]   As used herein, "modified" or an "altered level of" activity of a protein, for example an enzyme, in a recombinant host cell refers to a difference in one or more characteristics in the activity determined relative to the parent or native host cell. Typically differences in activity are determined between a recombinant host cell, having modified activity, and the corresponding wild-type host cell (*e.g.,* comparison of a culture of a recombinant host cell relative to the corresponding wild-type host cell). Modified activities can be the result of, for example, modified amounts of protein expressed by a recombinant host cell (*e.g.,* as the result of increased

or decreased number of copies of DNA sequences encoding the protein, increased or decreased number of mRNA transcripts encoding the protein, and/or increased or decreased amounts of protein translation of the protein from mRNA); changes in the structure of the protein (*e.g.*, changes to the primary structure, such as, changes to the protein's coding sequence that result in changes in substrate specificity, changes in observed kinetic parameters); and changes in protein stability (e.g., increased or decreased degradation of the protein). In some embodiments, the polypeptide is a mutant or a variant of any of the polypeptides described herein. In certain instances, the coding sequences for the polypeptides described herein are codon optimized for expression in a particular host cell. For example, for expression in *E. coli*, one or more codons can be optimized as described in, *e.g.*, Grosjean et al., Gene 18:199-209 (1982).

[0081]   The term "regulatory sequences" as used herein typically refers to a sequence of bases in DNA, operably-linked to DNA sequences encoding a protein that ultimately controls the expression of the protein. Examples of regulatory sequences include, but are not limited to, RNA promoter sequences, transcription factor binding sequences, transcription termination sequences, modulators of transcription (such as enhancer elements), nucleotide sequences that affect RNA stability, and translational regulatory sequences (such as, ribosome binding sites (*e.g.*, Shine-Dalgarno sequences in prokaryotes or Kozak sequences in eukaryotes), initiation codons, termination codons).

[0082]   The terms "altered level of expression" and "modified level of expression" are used interchangeably and mean that a polynucleotide, polypeptide, or hydrocarbon is present in a different concentration in an engineered host cell as compared to its concentration in a corresponding wild-type cell under the same conditions.

[0083]   As used herein, the term "titer" refers to the quantity of fatty acid derivative produced per unit volume of host cell culture. In any aspect of the compositions and methods described herein, a fatty acid derivative is produced at a titer of about 25 mg/L, about 50 mg/L, about 75 mg/L, about 100 mg/L, about 125 mg/L, about 150 mg/L, about 175 mg/L, about 200 mg/L, about 225 mg/L, about 250 mg/L, about 275 mg/L, about 300 mg/L, about 325 mg/L, about 350 mg/L, about 375 mg/L, about 400 mg/L, about 425 mg/L, about 450 mg/L, about 475 mg/L, about 500 mg/L, about 525 mg/L, about 550 mg/L, about 575 mg/L, about 600 mg/L, about 625 mg/L, about 650 mg/L, about 675 mg/L, about 700 mg/L, about 725 mg/L, about 750 mg/L, about 775 mg/L, about 800 mg/L, about 825 mg/L, about 850 mg/L, about 875 mg/L, about 900

mg/L, about 925 mg/L, about 950 mg/L, about 975 mg/L, about 1000 mg/L, about 1050 mg/L, about 1075 mg/L, about 1100 mg/L, about 1125 mg/L, about 1150 mg/L, about 1175 mg/L, about 1200 mg/L, about 1225 mg/L, about 1250 mg/L, about 1275 mg/L, about 1300 mg/L, about 1325 mg/L, about 1350 mg/L, about 1375 mg/L, about 1400 mg/L, about 1425 mg/L, about 1450 mg/L, about 1475 mg/L, about 1500 mg/L, about 1525 mg/L, about 1550 mg/L, about 1575 mg/L, about 1600 mg/L, about 1625 mg/L, about 1650 mg/L, about 1675 mg/L, about 1700 mg/L, about 1725 mg/L, about 1750 mg/L, about 1775 mg/L, about 1800 mg/L, about 1825 mg/L, about 1850 mg/L, about 1875 mg/L, about 1900 mg/L, about 1925 mg/L, about 1950 mg/L, about 1975 mg/L, about 2000 mg/L (2g/L), 3g/L, 5g/L, 10g/L, 20g/L, 30g/L, 40g/L, 50g/L, 60g/L, 70g/L, 80g/L, 90g/L, 100g/L or a range bounded by any two of the foregoing values. In other embodiments, a fatty acid derivative is produced at a titer of more than 100g/L, more than 200g/L, more than 300g/L, or higher, such as 500 g/L, 700 g/L, 1000 g/L, 1200 g/L, 1500 g/L, or 2000 g/L. The preferred titer of fatty acid derivative produced by a recombinant host cell according to the methods of the disclosure is from 5g/L to 200g/L, 10g/L to 150g/L, 20g/L to 120g/L and 30g/L to 100g/L. In one embodiment, the titer of fatty acid derivative produced by a recombinant host cell according to the methods of the disclosure is about 1g/L to about 250g/L and more particularly, 90 g/L to about 120g/L. The titer may refer to a particular fatty acid derivative or a combination of fatty acid derivatives produced by a given recombinant host cell culture.

[0084]   As used herein, the "yield of fatty acid derivative produced by a host cell" refers to the efficiency by which an input carbon source is converted to product (*i.e.,* fatty alcohol or fatty aldehyde) in a host cell. Host cells engineered to produce fatty acid derivatives according to the methods of the disclosure have a yield of at least 3%, at least 4%, at least 5%, at least 6%, at least 7%, at least 8%, at least 9%, at least 10%, at least 11%, at least 12%, at least 13%, at least 14%, at least 15%, at least 16%, at least 17%, at least 18%, at least 19%, at least 20 %, at least 21%, at least 22%, at least 23%, at least 24%, at least 25%, at least 26%, at least 27%, at least 28%, at least 29%, or at least 30% or a range bounded by any two of the foregoing values. In other embodiments, a fatty acid derivative or derivatives is produced at a yield of more than 30%, 40%, 50%, 60%, 70%, 80%, 90% or more. Alternatively, or in addition, the yield is about 30% or less, about 27% or less, about 25% or less, or about 22% or less. Thus, the yield can be bounded by any two of the above endpoints. For example, the yield of a fatty acid derivative or

derivatives produced by the recombinant host cell according to the methods of the disclosure can be 5% to 15%, 10% to 25%, 10% to 22%, 15% to 27%, 18% to 22%, 20% to 28%, or 20% to 30%. In a particular embodiment, the yield of a fatty acid derivative or derivatives produced by the recombinant host cell is about 10% to about 40%. In another particular embodiment, the yield of a fatty acid derivative or derivatives produced by the recombinant host cell is about 25%. The yield may refer to a particular fatty acid derivative or a combination of fatty acid derivatives produced by a given recombinant host cell culture.

[0085]   As used herein, the term "productivity" refers to the quantity of a fatty acid derivative or derivatives produced per unit volume of host cell culture per unit time. In any aspect of the compositions and methods described herein, the productivity of a fatty acid derivative or derivatives produced by a recombinant host cell is at least 100 mg/L/hour, at least 200 mg/L/hour, at least 300 mg/L/hour, at least 400 mg/L/hour, at least 500 mg/L/hour, at least 600 mg/L/hour, at least 700 mg/L/hour, at least 800 mg/L/hour, at least 900 mg/L/hour, at least 1000 mg/L/hour, at least 1100 mg/L/hour, at least 1200 mg/L/hour, at least 1300 mg/L/hour, at least 1400 mg/L/hour, at least 1500 mg/L/hour, at least 1600 mg/L/hour, at least 1700 mg/L/hour, at least 1800 mg/L/hour, at least 1900 mg/L/hour, at least 2000 mg/L/hour, at least 2100 mg/L/hour, at least 2200 mg/L/hour, at least 2300 mg/L/hour, at least 2400 mg/L/hour, or at least 2500 mg/L/hour. For example, the productivity of a fatty acid derivative or derivatives produced by a recombinant host cell according to the methods of the may be from 500 mg/L/hour to 2500 mg/L/hour, or from 700 mg/L/hour to 2000 mg/L/hour. In one particular embodiment, the yield is about 0.7mg/L/h to about 3g/L/h. The productivity may refer to a particular fatty acid derivative or a combination of fatty acid derivatives produced by a given recombinant host cell culture.

[0086]   As used herein, the term "total fatty species" and "total fatty acid product" may be used interchangeably herein with reference to the amount of fatty alcohols, fatty aldehydes and fatty acids, as evaluated by GC-FID as described in International Patent Application Publication WO2008/119082. The same terms may be used to mean fatty esters and free fatty acids when referring to a fatty ester analysis.

[0087]   As used herein, the term "glucose utilization rate" means the amount of glucose used by the culture per unit time, reported as grams/liter/hour (g/L/hr). As used herein, the term "carbon source" refers to a substrate or compound suitable to be used as a source of carbon for

prokaryotic or simple eukaryotic cell growth. Carbon sources can be in various forms, including, but not limited to polymers, carbohydrates, acids, alcohols, aldehydes, ketones, amino acids, peptides, and gases (e.g., CO and CO_2). Exemplary carbon sources include, but are not limited to, monosaccharides, such as glucose, fructose, mannose, galactose, xylose, and arabinose; oligosaccharides, such as fructo-oligosaccharide and galacto-oligosaccharide; polysaccharides such as starch, cellulose, pectin, and xylan; disaccharides, such as sucrose, maltose, cellobiose, and turanose; cellulosic material and variants such as hemicelluloses, methyl cellulose and sodium carboxymethyl cellulose; saturated or unsaturated fatty acids, succinate, lactate, and acetate; alcohols, such as ethanol, methanol, and glycerol, or mixtures thereof. The carbon source can also be a product of photosynthesis, such as glucose. In certain preferred embodiments, the carbon source is biomass. In other preferred embodiments, the carbon source is glucose. In other preferred embodiments the carbon source is sucrose.

[0088]   As used herein, the term "biomass" refers to any biological material from which a carbon source is derived. In some embodiments, a biomass is processed into a carbon source, which is suitable for bioconversion. In other embodiments, the biomass does not require further processing into a carbon source. The carbon source can be converted into a biofuel. An exemplary source of biomass is plant matter or vegetation, such as corn, sugar cane, or switchgrass. Another exemplary source of biomass is metabolic waste products, such as animal matter (*e.g.,* cow manure). Further exemplary sources of biomass include algae and other marine plants. Biomass also includes waste products from industry, agriculture, forestry, and households, including, but not limited to, fermentation waste, ensilage, straw, lumber, sewage, garbage, cellulosic urban waste, and food leftovers. The term "biomass" also refers to sources of carbon, such as carbohydrates (*e.g.,* monosaccharides, disaccharides, or polysaccharides).

[0089]   As used herein, the term "isolated," with respect to products (such as fatty acids and derivatives thereof) refers to products that are separated from cellular components, cell culture media, or chemical or synthetic precursors. The fatty acids and derivatives thereof produced by the methods described herein can be relatively immiscible in the fermentation broth, as well as in the cytoplasm. Therefore, the fatty acids and derivatives thereof can collect in an organic phase either intracellularly or extracellularly.

[0090]   As used herein, the terms "purify," "purified," or "purification" mean the removal or isolation of a molecule from its environment by, for example, isolation or separation.

"Substantially purified" molecules are at least about 60% free (*e.g.,* at least about 70% free, at least about 75% free, at least about 85% free, at least about 90% free, at least about 95% free, at least about 97% free, at least about 99% free) from other components with which they are associated. As used herein, these terms also refer to the removal of contaminants from a sample. For example, the removal of contaminants can result in an increase in the percentage of fatty acid derivatives in a sample. For example, when a fatty acid derivative is produced in a recombinant host cell, the fatty acid derivative can be purified by the removal of host cell proteins. After purification, the percentage of fatty acid derivative in the sample is increased. The terms "purify," "purified," and "purification" are relative terms which do not require absolute purity. Thus, for example, when a fatty acid derivative is produced in recombinant host cells, a purified fatty acid derivative is a fatty acid derivative that is substantially separated from other cellular components (e.g., nucleic acids, polypeptides, lipids, carbohydrates, or other hydrocarbons).

[0091]   **Strain Improvements**

[0092]   In order generate a high titer, yield, and/or productivity of fatty acid derivatives, a number of modifications were made to the production host cells. FadR is a key regulatory factor involved in fatty acid degradation and fatty acid biosynthetic pathways (Cronan *et al., Mol. Microbiol.,* 29(4): 937-943 (1998)). The *E. coli* ACS enzyme FadD and the fatty acid transport protein FadL are components of a fatty acid uptake system. FadL mediates transport of fatty acids into the bacterial cell, and FadD mediates formation of acyl-CoA esters. When no other carbon source is available, exogenous fatty acids are taken up by bacteria and converted to acyl-CoA esters, which can bind to the transcription factor FadR and depress the expression of the fad genes that encode proteins responsible for fatty acid transport (FadL), activation (FadD), and β-oxidation (FadA, FadB, FadE, and FadH). When alternative sources of carbon are available, bacteria synthesize fatty acids as acyl-ACPs, which are used for phospholipid synthesis, but are not substrates for β-oxidation. Thus, acyl-CoA and acyl-ACP are both independent sources of fatty acids that can result in different end-products (Caviglia *et al., J. Biol. Chem.,* 279(12): 1163-1169 (2004)).

[0093]   There are conflicting speculations in the art as to the factors that can limit fatty acid biosynthesis in host cells, such as *E. coli.* One suggestion is that a limitation of the main precursors for fatty acid biosynthesis, for example, acetyl-CoA and malonyl-CoA can result in decreased synthesis of fatty acid derivatives. One approach to increasing the flux through fatty

acid biosynthesis is to manipulate various enzymes in the pathway (see Figures 1 and 2). Example 3 describes studies which show construction of fab operons that encode enzymes in the biosynthetic pathway for conversion of malonyl-CoA into acyl-ACPs and integration into the chromosome of an *E.coli* host cell. Without wanting to be bound by theory, this may increase the flux of fatty acid biosynthesis. The supply of acyl-ACPs from acetyl-CoA via the acetyl-CoA carboxylase (acc) complex and fatty acid biosynthetic (fab) pathway is another step that may limit the rate of fatty acid derivative production (see Figure 3). Example 2 shows the effect of overexpression of an optimized version of *E. coli Corynebacterium glutamicum* accABCD (±birA) demonstrated that such genetic modifications can lead to increased production of acetyl-coA and malonyl-CoA in *E. coli*.

[0094]   In another approach, mutations in the rph and ilvG genes in the *E. coli* host cell were shown to result in higher free fatty acid (FFA) production, which translated into higher production of fatty alcohol as shown in Example 4. In still another approach, transposon mutagenesis and high-throughput screening was carried out to find beneficial mutations that increase the titer or yield. As shown in Example 5, a transposon insertion in the yijP gene can improve the fatty alcohol yield in shake flask and fed-batch fermentations.

[0095]   **Generation of Fatty Acid Derivatives by Recombinant Host Cells**

[0096]   The present disclosure provides numerous examples of polypeptides (*i.e.*, enzymes) having activities suitable for use in the fatty acid biosynthetic pathways described herein. Such polypeptides are collectively referred to herein as "fatty acid biosynthetic polypeptides" or "fatty acid biosynthetic enzymes". Non-limiting examples of fatty acid pathway polypeptides suitable for use in recombinant host cells of the disclosure are provided herein. In some embodiments, the disclosure includes a recombinant host cell including a polynucleotide sequence which encodes a fatty acid biosynthetic polypeptide. The polynucleotide sequence, which includes an open reading frame encoding a fatty acid biosynthetic polypeptide and operably-linked regulatory sequences, can be integrated into a chromosome of the recombinant host cells, incorporated in one or more plasmid expression systems resident in the recombinant host cell, or both. In one embodiment, a fatty acid biosynthetic polynucleotide sequence encodes a polypeptide which is endogenous to the parental host cell (*i.e.*, the control cell) of the recombinant host cell that is being engineered. Some such endogenous polypeptides are overexpressed in the recombinant host cell. In another embodiment, the fatty acid biosynthetic

polynucleotide sequence encodes an exogenous or heterologous polypeptide. In other words, the polypeptide encoded by the polynucleotide is exogenous to the parental host cell. In yet another embodiment, the genetically modified host cell overexpresses a gene encoding a polypeptide (protein) that increases the rate at which the host cell produces the substrate of a fatty acid biosynthetic enzyme, *i.e.,* a fatty acyl-thioester substrate. In certain embodiments, the enzyme encoded by the expressed gene is directly involved in fatty acid biosynthesis. Such recombinant host cells may be further engineered to include a polynucleotide sequence encoding one or more fatty acid biosynthetic polypeptides (*i.e.,* enzymes involved in fatty acid biosynthesis). Examples of such polypeptides are polpeptides or proteins having thioesterase activity, wherein the recombinant host cell synthesizes fatty acids; or having thioesterase activity and carboxylic acid reductase (CAR) activity, wherein the recombinant host cell synthesizes fatty aldehydes and fatty alcohols; or having thioesterase activity, carboxylic acid reductase activity and alcohol dehydrogenase activity wherein the recombinant host cell synthesizes fatty alcohols; or having acyl-CoA reductase (AAR) activity wherein the recombinant host cell synthesizes fatty aldehydes and fatty alcohols; or having acyl-CoA reductase (AAR) activity and alcohol dehydrogenase activity wherein the recombinant host cell synthesizes fatty alcohols; or having fatty alcohol forming acyl-CoA reductase (FAR) activity, wherein the recombinant host cell synthesizes fatty alcohols; or having thioesterase activity, carboxylic acid reductase activity and aldehyde decarbonylase activity, wherein the recombinant host cell synthesizes alkanes; or having acyl-CoA reductase activity and aldehyde decarbonylase activity, wherein the recombinant host cell synthesizes alkanes; or having ester synthase activity wherein the recombinant host cell synthesizes fatty esters (*e.g.,* one enzyme system; see Figure 5); or having thioesterase activity, acyl-CoA synthase activity and ester synthase activity wherein the recombinant host cell synthesizes fatty esters (*e.g.,* three enzyme system; see Figure 5); or having OleA activity, wherein the recombinant host cell synthesizes aliphatic ketones; or having OleABCD activity, wherein the recombinant host cell synthesizes internal olefins; or having thioesterase activity and decarboxylase activity, wherein the recombinant host cell synthesizes terminal olefins; or combinations thereof. In some embodiments, at least one polypeptide encoded by a fatty acid biosynthetic polynucleotide is an exogenous (or heterologous) polypeptide (*e.g.,* a polypeptide originating from an organism other than the parental host cell, or a variant of a polypeptide native to the parental microbial cell) or an endogenous polypeptide

(*i.e.,* a polypeptide native to the parental host cell) wherein the endogenous polypeptide is overexpressed in the recombinant host cell.

[0097]   Table 1 below provides a listing of exemplary proteins which can be expressed in recombinant host cells to facilitate production of particular fatty acid derivatives.

[0098]   **Table 1: Gene Designations**

Gene Designation	Source Organism	Enzyme Name	Accession No.	EC Number	Exemplary Use
1. Fatty Acid Production Increase / Product Production Increase					
accA	*E. coli, Lactococci*	Acetyl-CoA carboxylase, subunit A (carboxyltransferase alpha)	AAC73296, NP_414727	6.4.1.2	increase Malonyl-CoA production
accB	*E. coli, Lactococci*	Acetyl-CoA carboxylase, subunit B (BCCP: biotin carboxyl carrier protein)	NP_417721	6.4.1.2	increase Malonyl-CoA production
accC	*E. coli, Lactococci*	Acetyl-CoA carboxylase, subunit C (biotin carboxylase)	NP_417722	6.4.1.2, 6.3.4.14	increase Malonyl-CoA production
accD	*E. coli, Lactococci*	Acetyl-CoA carboxylase, subunit D (carboxyltransferase beta)	NP_416819	6.4.1.2	increase Malonyl-CoA production
fadD	*E. coli W3110*	acyl-CoA synthase	AP_002424	2.3.1.86, 6.2.1.3	increase Fatty acid production
fabA	*E. coli K12*	β-hydroxydecanoyl thioester dehydratase/isomerase	NP_415474	4.2.1.60	increase fatty acyl-ACP/CoA production
fabB	*E. coli*	3-oxoacyl-[acyl-carrier-protein] synthase I	BAA16180	2.3.1.41	increase fatty acyl-ACP/CoA production
fabD	*E. coli K12*	[acyl-carrier-protein] S-malonyltransferase	AAC74176	2.3.1.39	increase fatty acyl-ACP/CoA production
fabF	*E. coli K12*	3-oxoacyl-[acyl-carrier-protein] synthase II	AAC74179	2.3.1.179	increase fatty acyl-ACP/CoA production
fabG	*E. coli K12*	3-oxoacyl-[acyl-carrier protein]	AAC74177	1.1.1.100	increase fatty acyl-ACP/CoA production

Gene Designation	Source Organism	Enzyme Name	Accession No.	EC Number	Exemplary Use
		reductase			
fabH	*E. coli K12*	3-oxoacyl-[acyl-carrier-protein] synthase III	AAC74175	2.3.1.180	increase fatty acyl-ACP/CoA production
fabI	*E. coli K12*	enoyl-[acyl-carrier-protein] reductase	NP_415804	1.3.1.9	increase fatty acyl-ACP/CoA production
fabR	*E. coli K12*	Transcriptional Repressor	NP_418398	none	modulate unsaturated fatty acid production
fabV	*Vibrio cholerae*	enoyl-[acyl-carrier-protein] reductase	YP_001217 283	1.3.1.9	increase fatty acyl-ACP/CoA production
fabZ	*E. coli K12*	(3R)-hydroxymyristol acyl carrier protein dehydratase	NP_414722	4.2.1.-	increase fatty acyl-ACP/CoA production
fadE	*E. coli K13*	acyl-CoA dehydrogenase	AAC73325	1.3.99.3, 1.3.99.-	reduce fatty acid degradation
fadR	*E. coli*	transcriptional regulatory protein	NP_415705	none	Block or reverse fatty acid degradation
2. Chain Length Control					
tesA (with or without leader sequence)	*E. coli*	thioesterase - leader sequence is amino acids 1-26	P0ADA1	3.1.2.-, 3.1.1.5	C18 Chain Length
tesA (without leader sequence)	*E. coli*	thioesterase	AAC73596, NP_415027	3.1.2.-, 3.1.1.5	C18:1 Chain Length
tesA (mutant of E. coli thioesterase I complexed with octanoic acid)	*E. coli*	thioesterase	L109P	3.1.2.-, 3.1.1.5	<C18 Chain Length
fatB1	*Umbellularia californica*	thioesterase	Q41635	3.1.2.14	C12:0 Chain Length
fatB2	*Cuphea hookeriana*	thioesterase	AAC49269	3.1.2.14	C8:0 - C10:0 Chain Length
fatB3	*Cuphea hookeriana*	thioesterase	AAC72881	3.1.2.14	C14:0 - C16:0 Chain Length
fatB	*Cinnamomum camphora*	thioesterase	Q39473	3.1.2.14	C14:0 Chain Length
fatB	*Arabidopsis*	thioesterase	CAA85388	3.1.2.14	C16:1 Chain Length

Gene Designation	Source Organism	Enzyme Name	Accession No.	EC Number	Exemplary Use
	thaliana				
fatA1	*Helianthus annuus*	thioesterase	AAL79361	3.1.2.14	C18:1 Chain Length
atfata	*Arabidopsis thaliana*	thioesterase	NP_189147 , NP_193041	3.1.2.14	C18:1 Chain Length
fatA	*Brassica juncea*	thioesterase	CAC39106	3.1.2.14	C18:1 Chain Length
fatA	*Cuphea hookeriana*	thioesterase	AAC72883	3.1.2.14	C18:1 Chain Length
tes	*Photbacterium profundum*	thioesterase	YP_130990	3.1.2.14	Chain Length
tesB	*E. coli*	thioesterase	NP_414986	3.1.2.14	Chain Length
fadM	*E. coli*	thioesterase	NP_414977	3.1.2.14	Chain Length
yciA	*E. coli*	thioesterase	NP_415769	3.1.2.14	Chain Length
ybgC	*E. coli*	thioesterase	NP_415264	3.1.2.14	Chain Length

3. Saturation Level Control*

Sfa	*E. coli*	Suppressor of fabA	AAN79592, AAC44390	none	increase monounsaturated fatty acids
fabA	*E. coli K12*	β-hydroxydecanoyl thioester dehydratase/isomerase	NP_415474	4.2.1.60	produce unsaturated fatty acids
GnsA	*E. coli*	suppressors of the secG null mutation	ABD18647.1	none	increase unsaturated fatty acid esters
GnsB	*E. coli*	suppressors of the secG null mutation	AAC74076.1	none	increase unsaturated fatty acid esters
fabB	*E. coli*	3-oxoacyl-[acyl-carrier-protein] synthase I	BAA16180	2.3.1.41	modulate unsaturated fatty acid production
des	*Bacillus subtilis*	D5 fatty acyl desaturase	O34653	1.14.19	modulate unsaturated fatty acid production

4. Product Output: Wax Production

AT3G51970	*Arabidopsis thaliana*	long-chain-alcohol O-fatty-acyltransferase	NP_190765	2.3.1.26	wax production
ELO1	*Pichia angusta*	Fatty acid elongase	BAD98251	2.3.1.-	produce very long chain length fatty acids
plsC	*Saccharomyces cerevisiae*	acyltransferase	AAA16514	2.3.1.51	wax production
DAGAT/DGAT	*Arabidopsis thaliana*	diacylglycerol acyltransferase	AAF19262	2.3.1.20	wax production
hWS	*Homo sapiens*	acyl-CoA wax	AAX48018	2.3.1.20	wax production

Gene Designation	Source Organism	Enzyme Name	Accession No.	EC Number	Exemplary Use
		alcohol acyltransferase			
aft1	*Acinetobacter sp. ADP1*	bifunctional wax ester synthase/acyl-CoA:diacylglycerol acyltransferase	AAO17391	2.3.1.20	wax production
ES9	*Marinobacter hydrocarbonoclasticus*	wax ester synthase	ABO21021	2.3.1.20	wax production
mWS	*Simmondsia chinensis*	wax ester synthase	AAD38041	2.3.1.-	wax production
5. Product Output: Fatty Alcohol Output					
		thioesterases (see above)			increase fatty acid/fatty alcohol production
BmFAR	*Bombyxmori*	FAR (fatty alcohol forming acyl-CoA reductase)	BAC79425	1.1.1.-	convert acyl-CoA to fatty alcohol
acr1	*Acinetobacter sp. ADP1*	acyl-CoA reductase	YP_047869	1.2.1.42	reduce fatty acyl-CoA to fatty aldehydes
yqhD	*E. coli W3110*	alcohol dehydrogenase	AP_003562	1.1.-.-	reduce fatty aldehydes to fatty alcohols; increase fatty alcohol production
alrA	*Acinetobacter sp. ADP1*	alcohol dehydrogenase	CAG70252	1.1.-.-	reduce fatty aldehydes to fatty alcohols
BmFAR	*Bombyxmori*	FAR (fatty alcohol forming acyl-CoA reductase)	BAC79425	1.1.1.-	reduce fatty acyl-CoA to fatty alcohol
GTNG_1865	*Geobacillusthermodenitrificans NG80-2*	Long-chain aldehyde dehydrogenase	YP_001125970	1.2.1.3	reduce fatty aldehydes to fatty alcohols
AAR	*Synechococcus elongatus*	Acyl-ACP reductase	YP_400611	1.2.1.42	reduce fatty acyl-ACP/CoA to fatty aldehydes
carB	*Mycobacterium smegmatis*	carboxylic acid reductase protein	YP_889972	6.2.1.3, 1.2.1.42	reduce fatty acids to fatty aldehyde
FadD	*E. coli K12*	acyl-CoA synthetase	NP_416319	6.2.1.3	activates fatty acids to fatty acyl-CoAs
atoB	*Erwiniacarotovora*	acetyl-CoA acetyltransferase	YP_049388	2.3.1.9	production of butanol
hbd	*Butyrivibriofibrisolvens*	Beta-hydroxybutyryl-CoA dehydrogenase	BAD51424	1.1.1.157	production of butanol

Gene Designation	Source Organism	Enzyme Name	Accession No.	EC Number	Exemplary Use
CPE0095	*Clostridium perfringens*	crotonasebutyryl-CoA dehydryogenase	BAB79801	4.2.1.55	production of butanol
bcd	*Clostridium beijerinckii*	butyryl-CoA dehydryogenase	AAM14583	1.3.99.2	production of butanol
ALDH	*Clostridium beijerinckii*	coenzyme A-acylating aldehyde dehydrogenase	AAT66436	1.2.1.3	production of butanol
AdhE	*E. coli CFT073*	aldehyde-alcohol dehydrogenase	AAN80172	1.1.1.1 1.2.1.10	production of butanol
6. Fatty Alcohol Acetyl Ester Output					
		thioesterases (see above)			modify output
acr1	*Acinetobacter sp. ADP1*	acyl-CoA reductase	YP_047869	1.2.1.42	modify output
yqhD	*E. Coli K12*	alcohol dehydrogenase	AP_003562	1.1.-.-	modify output
AAT	*Fragaria x ananassa*	alcohol O-acetyltransferase	AAG13130	2.3.1.84	modify output
7. Product Export					
AtMRP5	*Arabidopsis thaliana*	Arabidopsis thaliana multidrug resistance-associated	NP_171908	none	modify product export amount
AmiS2	*Rhodococcus sp.*	ABC transporter AmiS2	JC5491	none	modify product export amount
AtPGP1	*Arabidopsis thaliana*	Arabidopsis thaliana p glycoprotein 1	NP_181228	none	modify product export amount
AcrA	*CandidatusPr otochlamydiaa moebophila UWE25*	putative multidrug-efflux transport protein acrA	CAF23274	none	modify product export amount
AcrB	*CandidatusPr otochlamydiaa moebophila UWE25*	probable multidrug-efflux transport protein, acrB	CAF23275	none	modify product export amount
TolC	*Francisellatul arensis subsp. novicida*	Outer membrane protein [Cell envelope biogenesis,	ABD59001	none	modify product export amount
AcrE	*Shigellasonnei Ss046*	transmembrane protein affects septum formation and cell membrane permeability	YP_312213	none	modify product export amount

Gene Designation	Source Organism	Enzyme Name	Accession No.	EC Number	Exemplary Use
AcrF	*E. coli*	Acriflavine resistance protein F	P24181	none	modify product export amount
tll1619	*Thermosynech ococcus elongatus [BP-1]*	multidrug efflux transporter	NP_682409 .1	none	modify product export amount
tll0139	*Thermosynech ococcus elongatus [BP-1]*	multidrug efflux transporter	NP_680930 .1	none	modify product export amount
8. Fermentation					
replication checkpoint genes					increase output efficiency
umuD	*Shigellasonnei Ss046*	DNA polymerase V, subunit	YP_310132	3.4.21.-	increase output efficiency
umuC	*E. coli*	DNA polymerase V, subunit	ABC42261	2.7.7.7	increase output efficiency
pntA, pntB	*Shigellaflexne ri*	NADH:NADPH transhydrogenase (alpha and beta subunits)	P07001, P0AB70	1.6.1.2	increase output efficiency
9. Other					
fabK	*Streptococcus pneumoniae*	trans-2-enoyl-ACP reductase II	AAF98273	1.3.1.9	Contributes to fatty acid biosynthesis
fabL	*Bacillus licheniformis DSM 13*	enoyl-(acyl carrier protein) reductase	AAU39821	1.3.1.9	Contributes to fatty acid biosynthesis
fabM	*Streptococcus mutans*	trans-2, cis-3-decenoyl-ACP isomerase	DAA05501	4.2.1.17	Contributes to fatty acid biosynthesis

[0099]   **Production of Fatty Acids**

[00100]   The recombinant host cells may include one or more polynucleotide sequences that comprise an open reading frame encoding a thioesterase, *e.g.,* having an Enzyme Commission number of EC 3.1.1.5 or EC 3.1.2.– (for example, EC 3.1.2.14), together with operably-linked regulatory sequences that facilitate expression of the protein in the recombinant host cells. In the recombinant host cells, the open reading frame coding sequences and/or the regulatory sequences are modified relative to the corresponding wild-type gene encoding the thioesterase. The activity

of the thioesterase in the recombinant host cell is modified relative to the activity of the thioesterase expressed from the corresponding wild-type gene in a corresponding host cell. In some embodiments, a fatty acid derivative composition including fatty acids is produced by culturing a recombinant cell in the presence of a carbon source under conditions effective to express the thioesterase. In related embodiments, the recombinant host cell comprises a polynucleotide encoding a polypeptide having thioesterase activity, and one or more additional polynucleotides encoding polypeptides having other fatty acid biosynthetic enzyme activities. In some such instances, the fatty acid produced by the action of the thioesterase is converted by one or more enzymes having a different fatty acid biosynthetic enzyme activity to another fatty acid derivative, such as, for example, a fatty ester, fatty aldehyde, fatty alcohol, or a hydrocarbon.

[00101]   The chain length of a fatty acid, or a fatty acid derivative made therefrom, can be selected for by modifying the expression of particular thioesterases. The thioesterase will influence the chain length of fatty acid derivatives produced. The chain length of a fatty acid derivative substrate can be selected for by modifying the expression of selected thioesterases (EC 3.1. 2.14 or EC 3.1.1.5). Hence, host cells can be engineered to express, overexpress, have attenuated expression, or not express one or more selected thioesterases to increase the production of a preferred fatty acid derivative substrate. For example, C_{10} fatty acids can be produced by expressing a thioesterase that has a preference for producing C_{10} fatty acids and attenuating thioesterases that have a preference for producing fatty acids other than C_{10} fatty acids (e.g., a thioesterase which prefers to produce C_{14} fatty acids). This would result in a relatively homogeneous population of fatty acids that have a carbon chain length of 10. In other instances, C_{14} fatty acids can be produced by attenuating endogenous thioesterases that produce non-C_{14} fatty acids and expressing the thioesterases that use C_{14}-ACP. In some situations, C_{12} fatty acids can be produced by expressing thioesterases that use C_{12}-ACP and attenuating thioesterases that produce non-C_{12} fatty acids. For example, C12 fatty acids can be produced by expressing a thioesterase that has a preference for producing C12 fatty acids and attenuating thioesterases that have a preference for producing fatty acids other than C12 fatty acids. This would result in a relatively homogeneous population of fatty acids that have a carbon chain length of 12. The fatty acid derivatives are recovered from the culture medium with substantially all of the fatty acid derivatives produced extracellularly. The fatty acid derivative composition produced by a recombinant host cell can be analyzed using methods known in the art, for

example, GC-FID, in order to determine the distribution of particular fatty acid derivatives as well as chain lengths and degree of saturation of the components of the fatty acid derivative composition. Acetyl-CoA, malonyl-CoA, and fatty acid overproduction can be verified using methods known in the art, for example, by using radioactive precursors, HPLC, or GC-MS subsequent to cell lysis. Additional non-limiting examples of thioesterases and polynucleotides encoding them for use in the fatty acid pathway are provided in PCT Publication Application No. WO2010/075483, expressly incorporated by reference herein.

[00102]   **Production of Fatty Aldehydes**

[00103]   In one embodiment, the recombinant host cell produces a fatty aldehyde. In some embodiments, a fatty acid produced by the recombinant host cell is converted into a fatty aldehyde. In some embodiments, the fatty aldehyde produced by the recombinant host cell is then converted into a fatty alcohol or a hydrocarbon. In some embodiments, native (endogenous) fatty aldehyde biosynthetic polypeptides, such as aldehyde reductases, are present in the host cell (*e.g., E. coli*) and are effective to convert fatty aldehydes to fatty alcohols. In other embodiments, a native (endogenous) fatty aldehyde biosynthetic polypeptide is overexpressed. In still other embodiments, an exogenous fatty aldehyde biosynthetic polypeptide is introduced into a recombinant host cell and expressed or overexpressed. A native or recombinant host cell may comprise a polynucleotide encoding an enzyme having fatty aldehyde biosynthesis activity (e.g., a fatty aldehyde biosynthetic polypeptide or a fatty aldehyde biosynthetic polypeptide or enzyme). A fatty aldehyde is produced when the fatty aldehyde biosynthetic enzyme is expressed or overexpressed in the host cell. A recombinant host cell engineered to produce a fatty aldehyde will typically convert some of the fatty aldehyde to a fatty alcohol. In some embodiments, a fatty aldehyde is produced by expressing or overexpressing in the recombinant host cell a polynucleotide encoding a polypeptide having fatty aldehyde biosynthetic activity such as carboxylic acid reductase (CAR) activity. CarB, is an exemplary carboxylic acid reductase. In practicing the disclosure, a gene encoding a carboxylic acid reductase polypeptide may be expressed or overexpressed in the host cell. In some embodiments, the CarB polypeptide has the amino acid sequence of SEQ ID NO: 7. In other embodiments, the CarB polypeptide is a variant or mutant of SEQ ID NO: 7. Examples of carboxylic acid reductase (CAR) polypeptides and polynucleotides encoding them include, but are not limited to FadD9 (EC 6.2.1.-, UniProtKB Q50631, GenBank NP_217106, SEQ ID NO: 34), CarA (GenBank

ABK75684), CarB (GenBank YP889972; SEQ ID NO: 33) and related polypeptides described in PCT Publication No. WO2010/042664 and US Patent No. 8,097,439, each of which is expressly incorporated by reference herein. In some embodiments the recombinant host cell further comprises a polynucleotide encoding a thioesterase. In some embodiments, the fatty aldehyde is produced by expressing or overexpressing in the recombinant host cell a polynucleotide encoding a fatty aldehyde biosynthetic polypeptide, such as a polypeptide having acyl-ACP reductase (AAR) activity. Expression of acyl-ACP reductase in a recombinant host cell results in the production of fatty aldehydes and fatty alcohols (see Figure 4). Native (endogenous) aldehyde reductases present in a recombinant host cell (*e.g., E. coli*), can convert fatty aldehydes into fatty alcohols. Exemplary acyl-ACP reductase polypeptides are described in PCT Publication Nos. WO2009/140695 and WO/2009/140696, both of which are expressly incorporated by reference herein. A composition comprising fatty aldehydes (a fatty aldehyde composition) is produced by culturing a host cell in the presence of a carbon source under conditions effective to express the fatty aldehyde biosynthetic enzyme. In some embodiments, the fatty aldehyde composition comprises fatty aldehydes and fatty alcohols. Typically, the fatty aldehyde composition is recovered from the extracellular environment of the recombinant host cell, *i.e.,* the cell culture medium.

[00104]   Production of Fatty Alcohols

[00105]   In some embodiments, the recombinant host cell includes a polynucleotide encoding a polypeptide (an enzyme) having fatty alcohol biosynthetic activity (a fatty alcohol biosynthetic polypeptide or a fatty alcohol biosynthetic enzyme), and a fatty alcohol is produced by the recombinant host cell. A composition comprising fatty alcohols (a fatty alcohol composition) may be produced by culturing the recombinant host cell in the presence of a carbon source under conditions effective to express a fatty alcohol biosynthetic enzyme. In some embodiments, the fatty alcohol composition comprises fatty alcohols, however, a fatty alcohol composition may comprise other fatty acid derivatives. Typically, the fatty alcohol composition is recovered from the extracellular environment of the recombinant host cell, *i.e.,* the cell culture medium. In one approach, recombinant host cells have been engineered to produce fatty alcohols by expressing a thioesterase, which catalyzes the conversion of acyl-ACPs into free fatty acids (FFAs) and a carboxylic acid reductase (CAR), which converts free fatty acids into fatty aldehydes. Native (endogenous) aldehyde reductases present in the host cell (*e.g., E. coli*) can convert the fatty

aldehydes into fatty alcohols. In some embodiments, native (endogenous) fatty aldehyde biosynthetic polypeptides, such as aldehyde reductases present in the host cell, may be sufficient to convert fatty aldehydes to fatty alcohols. However, in other embodiments, a native (endogenous) fatty aldehyde biosynthetic polypeptide is overexpressed and in still other embodiments, an exogenous fatty aldehyde biosynthetic polypeptide is introduced into a recombinant host cell and expressed or overexpressed. In some embodiments, the fatty alcohol is produced by expressing or overexpressing in the recombinant host cell a polynucleotide encoding a polypeptide having fatty alcohol biosynthetic activity which converts a fatty aldehyde to a fatty alcohol. For example, an alcohol dehydrogenase (aldehyde reductase, e.g., EC 1.1.1.1), may be used in practicing the disclosure. As used herein, an alcohol dehydrogenase refers to a polypeptide capable of catalyzing the conversion of a fatty aldehyde to an alcohol (e.g., a fatty alcohol). One of ordinary skill in the art will appreciate that certain alcohol dehydrogenases are capable of catalyzing other reactions as well, and these non-specific alcohol dehydrogenases also are encompassed by the alcohol dehydrogenase. Examples of alcohol dehydrogenase polypeptides useful in accordance with the disclosure include, but are not limited to AlrA of *Acinetobacter sp. M-1* (SEQ ID NO: 3) or AlrA homologs such as AlrAadp1 (SEQ ID NO: 4) and endogenous *E. coli* alcohol dehydrogenases such as YjgB, (AAC77226) (SEQ ID NO: 5), DkgA (NP_417485), DkgB (NP_414743), YdjL (AAC74846), YdjJ (NP_416288), AdhP (NP_415995), YhdH (NP_417719), YahK (NP_414859), YphC (AAC75598), YqhD (446856) and YbbO [AAC73595.1]. Additional examples are described in International Patent Application Publication Nos. WO 2007/136762, WO2008/119082 and WO 2010/062480, each of which is expressly incorporated by reference herein. In certain embodiments, the fatty alcohol biosynthetic polypeptide has aldehyde reductase or alcohol dehydrogenase activity (EC 1.1.1.1).

[00106] In another approach, recombinant host cells have been engineered to produce fatty alcohols by expressing fatty alcohol forming acyl-CoA reductases or fatty acyl reductases (FARs) which convert fatty acyl-thioester substrates (e.g., fatty acyl-CoA or fatty acyl-ACP) to fatty alcohols. In some embodiments, the fatty alcohol is produced by expressing or overexpressing a polynucleotide encoding a polypeptide having fatty alcohol forming acyl-CoA reductase (FAR) activity in a recombinant host cell. Examples of FAR polypeptides useful in accordance with this embodiment are described in PCT Publication No. WO2010/062480 which is expressly incorporated by reference herein. Fatty alcohol may be produced via an acyl-CoA

dependent pathway utilizing fatty acyl-ACP and fatty acyl-CoA intermediates and an acyl-CoA independent pathway utilizing fatty acyl-ACP intermediates but not a fatty acyl-CoA intermediate. In particular embodiments, the enzyme encoded by the over expressed gene is selected from a fatty acid synthase, an acyl-ACP thioesterase, a fatty acyl-CoA synthase and an acetyl-CoA carboxylase. In some embodiments, the protein encoded by the over expressed gene is endogenous to the host cell. In other embodiments, the protein encoded by the overexpressed gene is heterologous to the host cell. Fatty alcohols are also made in nature by enzymes that are able to reduce various acyl-ACP or acyl-CoA molecules to the corresponding primary alcohols. See also, U.S. Patent Publication Nos. 20100105963, and 20110206630 and US Patent No. 8097439, expressly incorporated by reference herein. Strategies to increase production of fatty alcohols by recombinant host cells include increased flux through the fatty acid biosynthetic pathway by overexpression of native fatty acid biosynthetic genes and/or expression of exogenous fatty acid biosynthetic genes from different organisms in the production host such that fatty alcohol biosynthesis is increased.

[00107]   **Production of Esters**

[00108]   As used herein, the term "fatty ester" may be used with reference to an ester. A fatty ester as referred to herein can be any ester made from a fatty acid, for example a fatty acid ester. In some embodiments, a fatty ester contains an A side and a B side. As used herein, an "A side" of an ester refers to the carbon chain attached to the carboxylate oxygen of the ester. As used herein, a "B side" of an ester refers to the carbon chain comprising the parent carboxylate of the ester. In embodiments where the fatty ester is derived from the fatty acid biosynthetic pathway, the A side is contributed by an alcohol, and the B side is contributed by a fatty acid. Any alcohol can be used to form the A side of the fatty esters. For example, the alcohol can be derived from the fatty acid biosynthetic pathway. Alternatively, the alcohol can be produced through non-fatty acid biosynthetic pathways. Moreover, the alcohol can be provided exogenously. For example, the alcohol can be supplied in the fermentation broth in instances where the fatty ester is produced by an organism. Alternatively, a carboxylic acid, such as a fatty acid or acetic acid, can be supplied exogenously in instances where the fatty ester is produced by an organism that can also produce alcohol. The carbon chains comprising the A side or B side can be of any length. In one embodiment, the A side of the ester is at least about 1, 2, 3, 4, 5, 6, 7, 8, 10, 12, 14, 16, or 18 carbons in length. When the fatty ester is a fatty acid methyl ester, the A side of the

ester is 1 carbon in length. When the fatty ester is a fatty acid ethyl ester, the A side of the ester is 2 carbons in length. The B side of the ester can be at least about 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, or 26 carbons in length. The A side and/or the B side can be straight or branched chain. The branched chains can have one or more points of branching. In addition, the branched chains can include cyclic branches. Furthermore, the A side and/or B side can be saturated or unsaturated. If unsaturated, the A side and/or B side can have one or more points of unsaturation. In one embodiment, the fatty ester is produced biosynthetically. In this embodiment, first the fatty acid is activated. Non-limiting examples of "activated" fatty acids are acyl-CoA, acyl ACP, and acyl phosphate. Acyl-CoA can be a direct product of fatty acid biosynthesis or degradation. In addition, acyl-CoA can be synthesized from a free fatty acid, a CoA, and an adenosine nucleotide triphosphate (ATP). An example of an enzyme which produces acyl-CoA is acyl-CoA synthase. In some embodiments, the recombinant host cell comprises a polynucleotide encoding a polypeptide, *e.g.,* an enzyme having ester synthase activity, (ester synthase polypeptide or an ester synthase).

[00109] A fatty ester is produced by a reaction catalyzed by the ester synthase polypeptide expressed or overexpressed in the recombinant host cell. In some embodiments, a composition comprising fatty esters fatty ester is produced by culturing the recombinant cell in the presence of a carbon source under conditions effective to express an ester synthase. In some embodiments, the fatty ester composition is recovered from the cell culture. Ester synthase polypeptides include, for example, an ester synthase polypeptide classified as EC 2.3.1.75, or any other polypeptide which catalyzes the conversion of an acyl-thioester to a fatty ester, including, without limitation, a thioesterase, an ester synthase, an acyl-CoA:alcohol transacylase, an acyltransferase, or a fatty acyl-CoA:fatty alcohol acyltransferase. For example, the polynucleotide may encode wax/dgat, a bifunctional ester synthase/acyl-CoA:diacylglycerol acyltransferase from Simmondsia chinensis, Acinetobacter sp. Strain ADP, Alcanivorax borkumensis, Pseudomonas aeruginosa, Fundibacter jadensis, Arabidopsis thaliana, or Alkaligenes eutrophus. In a particular embodiment, the ester synthase polypeptide is an *Acinetobacter sp.* diacylglycerol O-acyltransferase (wax-dgaT; UniProtKB Q8GGG1, GenBank AAO17391) or *Simmondsia chinensis* wax synthase (UniProtKB Q9XGY6 , GenBank AAD38041. In another embodiment, the ester synthase polypeptide is for example ES9 (a wax ester synthase from *Marinobacter hydrocarbonoclasticus* DSM 8798, UniProtKB A3RE51(SEQ

41

ID NO: 6); ES8 of *Marinobacter hydrocarbonoclasticus* DSM8789 (GenBank Accession No. ABO21021; SEQ ID NO:7); GenBank ABO21021, encoded by the ws2 gene; or ES376 (another wax ester synthase derived from *Marinobacter hydrocarbonoclasticus* DSM 8798, UniProtKB A3RE50, GenBank ABO21020, encoded by the ws1 gene. In a particular embodiment, the polynucleotide encoding the ester synthase polypeptide is overexpressed in the recombinant host cell. In some embodiments, a fatty acid ester is produced by a recombinant host cell engineered to express three fatty acid biosynthetic enzymes: a thioesterase enzyme, an acyl-CoA synthetase (fadD) enzyme and an ester synthase enzyme (*e.g.,* three enzyme system; see Figure 5). In other embodiments, a fatty acid ester is produced by a recombinant host cell engineered to express one fatty acid biosynthetic enzyme, an ester synthase enzyme (*e.g.,* one enzyme system; see Figure 5). Non-limiting examples of ester synthase polypeptides and polynucleotides encoding them suitable for use in these embodiments include those described in PCT Publication Nos. WO2007/136762 and WO2008/119082, and WO/2011/038134 (three enzyme system) and WO/2011/038132 (one enzyme system), each of which is expressly incorporated by reference herein. The recombinant host cell may produce a fatty ester, such as a fatty acid methyl ester, a fatty acid ethyl ester or a wax ester in the extracellular environment of the host cells.

[00110]   Production of Hydrocarbons

[00111]   This aspect of the disclosure is based, at least in part, on the discovery that altering the level of expression of a fatty aldehyde biosynthetic polypeptide, for example, an acyl-ACP reductase polypeptide (EC 6.4.1.2) and a hydrocarbon biosynthetic polypeptide, *e.g.,* a decarbonylase in a recombinant host cell facilitates enhanced production of hydrocarbons by the recombinant host cell. In one embodiment, the recombinant host cell produces a hydrocarbon, such as an alkane or an alkene (*e.g.,* a terminal olefin or an internal olefin) or a ketone. In some embodiments, a fatty aldehyde produced by a recombinant host cell is converted by decarbonylation, removing a carbon atom to form a hydrocarbon. In other embodiments, a fatty acid produced by a recombinant host cell is converted by decarboxylation, removing a carbon atom to form a terminal olefin. In some embodiments, an acyl-ACP intermediate is converted by decarboxylation, removing a carbon atom to form an internal olefin or a ketone (see Figure 6). In some embodiments, the recombinant host cell comprises a polynucleotide encoding a polypeptide (an enzyme) having hydrocarbon biosynthetic activity (a hydrocarbon biosynthetic polypeptide or a hydrocarbon biosynthetic enzyme), and the hydrocarbon is produced by

expression or overexpression of the hydrocarbon biosynthetic enzyme in a recombinant host cell. An alkane biosynthetic pathway from cyanobacteria consisting of an acyl–acyl carrier protein reductase (AAR) and an aldehyde decarbonylase (ADC), which together convert intermediates of fatty acid metabolism to alkanes and alkenes has been used to engineer recombinant host cells for the production of hydrocarbons (Figure 6). The second of two reactions in the pathway through which saturated acyl-ACPs are converted to alkanes in cyanobacteria entails scission of the C1-C2 bond of a fatty aldehyde intermediate by the enzyme aldehyde decarbonylase (ADC), a ferritin-like protein with a binuclear metal cofactor of unknown composition. In some embodiments, the hydrocarbon is produced by expressing or overexpressing in the recombinant host cell a polynucleotide encoding a polypeptide having hydrocarbon biosynthetic activity such as an aldehyde decarbonylase (ADC) activity (*e.g.*, EC 4.1.99.5). Exemplary polynucleotides encoding an aldehyde decarbonylase useful in accordance with this embodiment include, but are not limited to, those described in PCT Publication Nos. WO2008/119082 and WO2009/140695 which are expressly incorporated by reference herein and those sequences presented in Table 2 below. In some embodiments the recombinant host cell further comprises a polynucleotide encoding a fatty aldehyde biosynthesis polypeptide. In some embodiments the recombinant host cell further comprises a polynucleotide encoding an acyl-ACP reductase. See, for example, Table 2 below.

[00112] Table 2: Exemplary Hydrocarbon Biosynthetic Polynucleotides and Polypeptides

Protein name	Polypeptide sequence	Nucleotide sequence	Sequence
Decarbonylase (ADC)	SEQ ID NO: 35	SEQ ID NO: 36	Synechococcus elongatus PCC7942 YP.sub.--400610 (Synpcc7942.sub.--1593)
Acyl-ACP Reducatase (AAR)	SEQ ID NO: 37	SEQ ID NO: 38	*Synechococcus elongatus PCC7942 YP_400611 (Synpcc7942_1594)*
Decarbonylase (ADC)	SEQ ID NO: 39	SEQ ID NO: 40	Prochlorococcus mariunus CCMP1986 PMM0532
Acyl-ACP Reducatase (AAR)	SEQ ID NO: 41	SEQ ID NO: 42	Prochlorococcus marinus CCMP1986 PMM0533 (NP_892651)

[00113] In some embodiments, a composition comprising is produced by culturing the recombinant cell in the presence of a carbon source under conditions effective to express the Acyl-CoA reductase and decarbonylase polynucleotides. In some embodiments, the

hydrocarbon composition comprises saturated and unsaturated hydrocarbons. However, a hydrocarbon composition may comprise other fatty acid derivatives. Typically, the hydrocarbon composition is recovered from the extracellular environment of the recombinant host cell, *i.e.,* the cell culture medium. As used herein, an alkane refers to saturated hydrocarbons or compounds that consist only of carbon (C) and hydrogen (H), wherein these atoms are linked together by single bonds (*i.e.,* they are saturated compounds). Olefins and alkenes refer to hydrocarbons containing at least one carbon-to-carbon double bond (*i.e.,* they are unsaturated compounds). Terminal olefins, α-olefins, terminal alkenes, and 1-alkenes refer to the same compounds with reference to α-olefins or alkenes with a chemical formula $CxH2x$, distinguished from other olefins with a similar molecular formula by linearity of the hydrocarbon chain and the position of the double bond at the primary or alpha position. In some embodiments, a terminal olefin is produced by expressing or overexpressing in the recombinant host cell a polynucleotide encoding a hydrocarbon biosynthetic polypeptide, such as a polypeptide having decarboxylase activity as described, for example, in PCT Publication No. WO2009/085278 which is expressly incorporated by reference herein. In some embodiments the recombinant host cell further comprises a polynucleotide encoding a thioesterase. In other embodiments, a ketone is produced by expressing or overexpressing in the recombinant host cell a polynucleotide encoding a hydrocarbon biosynthetic polypeptide, such as a polypeptide having OleA activity as described, for example, in PCT Publication No. WO2008/147781, which is expressly incorporated by reference herein. In related embodiments, an internal olefin is produced by expressing or overexpressing in the recombinant host cell a polynucleotide encoding a hydrocarbon biosynthetic polypeptide, such as a polypeptide having OleCD or OleBCD activity together with a polypeptide having OleA activity as described, for example, in PCT Publication No. WO2008/147781, expressly incorporated by reference herein.

[00114]   **Recombinant Host Cells and Cell Cultures**

[00115]   Strategies to increase production of fatty acid derivatives by recombinant host cells include increased flux through the fatty acid biosynthetic pathway by overexpression of native fatty acid biosynthetic genes and expression of exogenous fatty acid biosynthetic genes from different organisms in the production host. As used herein, a recombinant host cell or engineered host cell refers to a host cell whose genetic makeup has been altered relative to the

corresponding wild-type host cell, for example, by deliberate introduction of new genetic elements and/or deliberate modification of genetic elements naturally present in the host cell. The offspring of such recombinant host cells also contain these new and/or modified genetic elements. In any of the aspects of the disclosure described herein, the host cell can be selected from the group consisting of a plant cell, insect cell, fungus cell (*e.g.,* a filamentous fungus, such as *Candida sp.,* or a budding yeast, such as *Saccharomyces sp.*), an algal cell and a bacterial cell. In one preferred embodiment, recombinant host cells are recombinant microorganisms. Examples of host cells that are microorganisms, include but are not limited to cells from the genus *Escherichia, Bacillus, Lactobacillus, Zymomonas, Rhodococcus, Pseudomonas, Aspergillus, Trichoderma, Neurospora, Fusarium, Humicola, Rhizomucor, Kluyveromyces, Pichia, Mucor, Myceliophtora, Penicillium, Phanerochaete, Pleurotus, Trametes, Chrysosporium, Saccharomyces, Stenotrophamonas, Schizosaccharomyces, Yarrowia,* or *Streptomyces.* In some embodiments, the host cell is a Gram-positive bacterial cell. In other embodiments, the host cell is a Gram-negative bacterial cell. In some embodiments, the host cell is an *E. coli* cell. In other embodiments, the host cell is a *Bacillus lentus* cell, a *Bacillus brevis* cell, a *Bacillus stearothermophilus* cell, a *Bacillus lichenoformis* cell, a *Bacillus alkalophilus* cell, a *Bacillus coagulans* cell, a *Bacillus circulans* cell, a *Bacillus pumilis* cell, a *Bacillus thuringiensis* cell, a *Bacillus clausii* cell, a *Bacillus megaterium* cell, a *Bacillus subtilis* cell, or a *Bacillus amyloliquefaciens* cell. In other embodiments, the host cell is a *Trichoderma koningii* cell, a *Trichoderma viride* cell, a *Trichoderma reesei* cell, a *Trichoderma longibrachiatum cell,* an *Aspergillus awamori* cell, an *Aspergillus fumigates* cell, an *Aspergillus foetidus* cell, an *Aspergillus nidulans* cell, an *Aspergillus niger* cell, an *Aspergillus oryzae* cell, a *Humicola insolens* cell, a *Humicola lanuginose* cell, a *Rhodococcus opacus* cell, a *Rhizomucor miehei* cell, or a *Mucor michei* cell. In yet other embodiments, the host cell is a *Streptomyces lividans* cell or a *Streptomyces murinus* cell. In yet other embodiments, the host cell is an Actinomycetes cell. In some embodiments, the host cell is a *Saccharomyces cerevisiae* cell. In other embodiments, the host cell is a cell from a eukaryotic plant, algae, cyanobacterium, green-sulfur bacterium, green non-sulfur bacterium, purple sulfur bacterium, purple non-sulfur bacterium, extremophile, yeast, fungus, an engineered organism thereof, or a synthetic organism. In some embodiments, the host cell is light-dependent or fixes carbon. In some embodiments, the host cell has autotrophic activity. In some embodiments, the host cell has photoautotrophic activity, such as in the presence of light. In some embodiments, the host cell is heterotrophic or mixotrophic in the

45

absence of light. In certain embodiments, the host cell is a cell from *Arabidopsis thaliana,*
Panicum virgatum, Miscanthus giganteus, Zea mays, Botryococcuse braunii, Chlamydomonas
reinhardtii, Dunaliela salina, Synechococcus Sp. PCC 7002, Synechococcus Sp. PCC 7942,
Synechocystis Sp. PCC 6803, Thermosynechococcus elongates BP-1, Chlorobium tepidum,
Chlorojlexus auranticus, Chromatiumm vinosum, Rhodospirillum rubrum, Rhodobacter
capsulatus, Rhodopseudomonas palusris, Clostridium ljungdahlii, Clostridiuthermocellum,
Penicillium chrysogenum, Pichia pastoris, Saccharomyces cerevisiae, Schizosaccharomyces
pombe, Pseudomonas fluorescens, or *Zymomonas mobilis.*

[00116]   **Production of Fatty Acid Derivative Compositions by Recombinant Host cells**

[00117]   A large variety of fatty acid derivatives can be produced by recombinant host cells
comprising strain improvements as described herein, including, but not limited to, fatty acids,
acyl-CoA, fatty aldehydes, short and long chain alcohols, hydrocarbons (*e.g.,* alkanes, alkenes or
olefins, such as terminal or internal olefins), fatty alcohols, esters (*e.g.,* wax esters, fatty acid
esters (e.g., methyl or ethyl esters)), and ketones. In some embodiments of the present
disclosure, the higher titer of fatty acid derivatives in a particular composition is a higher titer of
a particular type of fatty acid derivative (*e.g.,* fatty alcohols, fatty acid esters, or hydrocarbons)
produced by a recombinant host cell culture relative to the titer of the same fatty acid derivatives
produced by a control culture of a corresponding wild-type host cell. In such cases, the fatty acid
derivative compositions may comprise, for example, a mixture of the fatty alcohols with a
variety of chain lengths and saturation or branching characteristics. In other embodiments of the
present disclosure, the higher titer of fatty acid derivatives in a particular compositions is a
higher titer of a combination of different fatty acid derivatives (for example, fatty aldehydes and
alcohols, or fatty acids and esters) relative to the titer of the same fatty acid derivative produced
by a control culture of a corresponding wild-type host cell.

[00118]   **Engineering Host cells**

[00119]   In some embodiments, a polynucleotide (or gene) sequence is provided to the host
cell by way of a recombinant vector, which comprises a promoter operably linked to the
polynucleotide sequence. In certain embodiments, the promoter is a developmentally-regulated,
an organelle-specific, a tissue-specific, an inducible, a constitutive, or a cell-specific promoter.
In some embodiments, the recombinant vector includes at least one sequence including, but not

limited to, (a) an expression control sequence operatively coupled to the polynucleotide sequence; (b) a selection marker operatively coupled to the polynucleotide sequence; (c) a marker sequence operatively coupled to the polynucleotide sequence; (d) a purification moiety operatively coupled to the polynucleotide sequence; (e) a secretion sequence operatively coupled to the polynucleotide sequence; and (f) a targeting sequence operatively coupled to the polynucleotide sequence. The expression vectors described herein include a polynucleotide sequence described herein in a form suitable for expression of the polynucleotide sequence in a host cell. It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of polypeptide desired, etc. The expression vectors described herein can be introduced into host cells to produce polypeptides, including fusion polypeptides, encoded by the polynucleotide sequences as described herein. Expression of genes encoding polypeptides in prokaryotes, for example, *E. coli*, is most often carried out with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion polypeptides. Fusion vectors add a number of amino acids to a polypeptide encoded therein, usually to the amino- or carboxy- terminus of the recombinant polypeptide. Such fusion vectors typically serve one or more of the following three purposes (1) to increase expression of the recombinant polypeptide; (2) to increase the solubility of the recombinant polypeptide; and (3) to aid in the purification of the recombinant polypeptide by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant polypeptide. This enables separation of the recombinant polypeptide from the fusion moiety after purification of the fusion polypeptide. Examples of such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin, and enterokinase. Exemplary fusion expression vectors include pGEX (Pharmacia Biotech, Inc., Piscataway, NJ; Smith et al., *Gene, 67*: 31-40 (1988)), pMAL (New England Biolabs, Beverly, MA), and pRITS (Pharmacia Biotech, Inc., Piscataway, N.J.), which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant polypeptide. Examples of inducible, non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, *Gene* (1988) 69:301-315) and pET 11d (Studier *et al.*, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, Calif. (1990) 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac

fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident λ prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter. Suitable expression systems for both prokaryotic and eukaryotic cells are well known in the art; see, e.g., Sambrook et al., "Molecular Cloning: A Laboratory Manual," second edition, Cold Spring Harbor Laboratory, (1989). Examples of inducible, non-fusion *E. coli* expression vectors include pTrc (Amann et al., *Gene, 69*: 301-315 (1988)) and PET 11d (Studier et al., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA, pp. 60-89 (1990)). In certain embodiments, a polynucleotide sequence of the disclosure is operably linked to a promoter derived from bacteriophage T5. In one embodiment, the host cell is a yeast cell. In this embodiment, the expression vector is a yeast expression vector. Vectors can be introduced into prokaryotic or eukaryotic cells via a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.,* DNA) into a host cell. Suitable methods for transforming or transfecting host cells can be found in, for example, Sambrook et al. (*supra*). For stable transformation of bacterial cells, it is known that, depending upon the expression vector and transformation technique used, only a small fraction of cells will take-up and replicate the expression vector. In order to identify and select these transformants, a gene that encodes a selectable marker (*e.g.,* resistance to an antibiotic) can be introduced into the host cells along with the gene of interest. Selectable markers include those that confer resistance to drugs such as, but not limited to, ampicillin, kanamycin, chloramphenicol, or tetracycline. Nucleic acids encoding a selectable marker can be introduced into a host cell on the same vector as that encoding a polypeptide described herein or can be introduced on a separate vector. Cells stably transformed with the introduced nucleic acid can be identified by growth in the presence of an appropriate selection drug.

[00120]   **Host Cells**

[00121]   As used herein, an engineered or recombinant host cell is a cell used to produce a fatty acid derivative composition as further described herein. A host cell is referred to as an engineered host cell or a recombinant host cell if the expression of one or more polynucleotides or polypeptides in the host cell are altered or modified as compared to their expression in a corresponding wild-type host cell (*e.g.,* control cell) under the same conditions. In any of the aspects of the disclosure described herein, the host cell can be selected from the group consisting

48

of a eukaryotic plant, algae, cyanobacterium, green-sulfur bacterium, green non-sulfur bacterium, purple sulfur bacterium, purple non-sulfur bacterium, extremophile, yeast, fungus, engineered organisms thereof, or a synthetic organism. In some embodiments, the host cell is light dependent or fixes carbon. In some embodiments, the host cell has autotrophic activity. Various host cells can be used to produce fatty acid derivatives, as described herein.

[00122]   **Mutants or Variants**

[00123]   In some embodiments, the polypeptide is a mutant or a variant of any of the polypeptides described herein. The terms mutant and variant as used herein refer to a polypeptide having an amino acid sequence that differs from a wild-type polypeptide by at least one amino acid. For example, the mutant can comprise one or more of the following conservative amino acid substitutions: replacement of an aliphatic amino acid, such as alanine, valine, leucine, and isoleucine, with another aliphatic amino acid; replacement of a serine with a threonine; replacement of a threonine with a serine; replacement of an acidic residue, such as aspartic acid and glutamic acid, with another acidic residue; replacement of a residue bearing an amide group, such as asparagine and glutamine, with another residue bearing an amide group; exchange of a basic residue, such as lysine and arginine, with another basic residue; and replacement of an aromatic residue, such as phenylalanine and tyrosine, with another aromatic residue. In some embodiments, the mutant polypeptide has about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, or more amino acid substitutions, additions, insertions, or deletions. Preferred fragments or mutants of a polypeptide retain some or all of the biological function (*e.g.,* enzymatic activity) of the corresponding wild-type polypeptide. In some embodiments, the fragment or mutant retains at least 75%, at least 80%, at least 90%, at least 95%, or at least 98% or more of the biological function of the corresponding wild-type polypeptide. In other embodiments, the fragment or mutant retains about 100% of the biological function of the corresponding wild-type polypeptide. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without affecting biological activity may be found using computer programs well known in the art, for example, LASERGENETM software (DNASTAR, Inc., Madison, WI). In yet other embodiments, a fragment or mutant exhibits increased biological function as compared to a corresponding wild-type polypeptide. For example, a fragment or mutant may display at least a 10%, at least a 25%, at least a 50%, at least a 75%, or at least a 90% improvement in enzymatic activity as compared to the corresponding

wild-type polypeptide. In other embodiments, the fragment or mutant displays at least 100% (*e.g.*, at least 200%, or at least 500%) improvement in enzymatic activity as compared to the corresponding wild-type polypeptide. It is understood that the polypeptides described herein may have additional conservative or non-essential amino acid substitutions, which do not have a substantial effect on the polypeptide function. Whether or not a particular substitution will be tolerated (*i.e.*, will not adversely affect desired biological function, such as carboxylic acid reductase activity) can be determined as described in Bowie et al. (Science, 247: 1306-1310 (1990)). 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).

[00124] Variants can be naturally occurring or created in vitro. In particular, such variants can be created using genetic engineering techniques, such as site directed mutagenesis, random chemical mutagenesis, Exonuclease III deletion procedures, or standard cloning techniques. Alternatively, such variants, fragments, analogs, or derivatives can be created using chemical synthesis or modification procedures. Methods of making variants are well known in the art. These include procedures in which nucleic acid sequences obtained from natural isolates are modified to generate nucleic acids that encode polypeptides having characteristics that enhance their value in industrial or laboratory applications. In such procedures, a large number of variant sequences having one or more nucleotide differences with respect to the sequence obtained from the natural isolate are generated and characterized. Typically, these nucleotide differences result in amino acid changes with respect to the polypeptides encoded by the nucleic acids from the natural isolates. For example, variants can be prepared by using random and site-directed mutagenesis. Random and site-directed mutagenesis are described in, for example, Arnold, Curr. Opin. Biotech., 4: 450-455 (1993). Random mutagenesis can be achieved using error prone PCR (see, e.g., Leung et al., Technique, 1: 11-15 (1989); and Caldwell et al., PCR Methods Applic., 2: 28-33 (1992)). In error prone PCR, PCR is performed under conditions where the copying

fidelity of the DNA polymerase is low, such that a high rate of point mutations is obtained along the entire length of the PCR product. Briefly, in such procedures, nucleic acids to be mutagenized (e.g., a polynucleotide sequence encoding a carboxylic reductase enzyme) are mixed with PCR primers, reaction buffer, $MgCl_2$, $MnCl_2$, Taq polymerase, and an appropriate concentration of dNTPs for achieving a high rate of point mutation along the entire length of the PCR product. For example, the reaction can be performed using 20 fmoles of nucleic acid to be mutagenized, 30 pmole of each PCR primer, a reaction buffer comprising 50 mM KCl, 10 mM Tris HCl (pH 8.3), 0.01% gelatin, 7 mM $MgCl_2$, 0.5 mM $MnCl_2$, 5 units of Taq polymerase, 0.2 mM dGTP, 0.2 mM dATP, 1 mM dCTP, and 1 mM dTTP. PCR can be performed for 30 cycles of 94 °C for 1 min, 45 °C for 1 min, and 72 °C for 1 min. However, it will be appreciated that these parameters can be varied as appropriate. The mutagenized nucleic acids are then cloned into an appropriate vector, and the activities of the polypeptides encoded by the mutagenized nucleic acids are evaluated. Site-directed mutagenesis can be achieved using oligonucleotide-directed mutagenesis to generate site-specific mutations in any cloned DNA of interest. Oligonucleotide mutagenesis is described in, for example, Reidhaar-Olson et al., Science, 241: 53-57 (1988). Briefly, in such procedures a plurality of double stranded oligonucleotides bearing one or more mutations to be introduced into the cloned DNA are synthesized and inserted into the cloned DNA to be mutagenized (e.g., a polynucleotide sequence encoding a CAR polypeptide). Clones containing the mutagenized DNA are recovered, and the activities of the polypeptides they encode are assessed. Another method for generating variants is assembly PCR. Assembly PCR involves the assembly of a PCR product from a mixture of small DNA fragments. A large number of different PCR reactions occur in parallel in the same vial, with the products of one reaction priming the products of another reaction. Assembly PCR is described in, for example, U.S. Patent 5,965,408. Still another method of generating variants is sexual PCR mutagenesis. In sexual PCR mutagenesis, forced homologous recombination occurs between DNA molecules of different, but highly related, DNA sequences in vitro as a result of random fragmentation of the DNA molecule based on sequence homology. This is followed by fixation of the crossover by primer extension in a PCR reaction. Sexual PCR mutagenesis is described in, for example, Stemmer, Proc. Natl. Acad. Sci., U.S.A., 91: 10747-10751 (1994). Variants can also be created by in vivo mutagenesis. In some embodiments, random mutations in a nucleic acid sequence are generated by propagating the sequence in a bacterial strain, such as an E. coli strain, which carries mutations in one or more of the DNA repair pathways. Such

"mutator" strains have a higher random mutation rate than that of a wild-type strain. Propagating a DNA sequence (*e.g.,* a polynucleotide sequence encoding a CAR polypeptide) in one of these strains will eventually generate random mutations within the DNA. Mutator strains suitable for use for in vivo mutagenesis are described in, for example, International Patent Application Publication No. WO1991/016427. Variants can also be generated using cassette mutagenesis. In cassette mutagenesis, a small region of a double-stranded DNA molecule is replaced with a synthetic oligonucleotide "cassette" that differs from the native sequence. The oligonucleotide often contains a completely and/or partially randomized native sequence. Recursive ensemble mutagenesis can also be used to generate variants. Recursive ensemble mutagenesis is an algorithm for protein engineering (*i.e.,* protein mutagenesis) developed to produce diverse populations of phenotypically related mutants whose members differ in amino acid sequence. This method uses a feedback mechanism to control successive rounds of combinatorial cassette mutagenesis. Recursive ensemble mutagenesis is described in, for example, Arkin et al., Proc. Natl. Acad. Sci., U.S.A., 89: 7811-7815 (1992). In some embodiments, variants are created using exponential ensemble mutagenesis. Exponential ensemble mutagenesis is a process for generating combinatorial libraries with a high percentage of unique and functional mutants, wherein small groups of residues are randomized in parallel to identify, at each altered position, amino acids which lead to functional proteins. Exponential ensemble mutagenesis is described in, for example, Delegrave et al., Biotech. Res, 11: 1548-1552 (1993). In some embodiments, variants are created using shuffling procedures wherein portions of a plurality of nucleic acids that encode distinct polypeptides are fused together to create chimeric nucleic acid sequences that encode chimeric polypeptides as described in, for example, U.S. Patents 5,965,408 and 5,939,250. Insertional mutagenesis is mutagenesis of DNA by the insertion of one or more bases. Insertional mutations can occur naturally, mediated by virus or transposon, or can be artificially created for research purposes in the lab, e.g., by transposon mutagenesis. When exogenous DNA is integrated into that of the host, the severity of any ensuing mutation depends entirely on the location within the host's genome wherein the DNA is inserted. For example, significant effects may be evident if a transposon inserts in the middle of an essential gene, in a promoter region, or into a repressor or an enhancer region. Transposon mutagenesis and high-throughput screening was done to find beneficial mutations that increase the titer or yield of a fatty acid derivative or derivatives.

[00125]   **Culture Recombinant Host Cells and Cell Cultures/Fermentation**

[00126]   As used herein, the term "fermentation" broadly refers to the conversion of organic materials into target substances by host cells, for example, the conversion of a carbon source by recombinant host cells into fatty acids or derivatives thereof by propagating a culture of the recombinant host cells in a media comprising the carbon source. As used herein, the term "conditions permissive for the production" means any conditions that allow a host cell to produce a desired product, such as a fatty acid or a fatty acid derivative. Similarly, the term "conditions in which the polynucleotide sequence of a vector is expressed" means any conditions that allow a host cell to synthesize a polypeptide. Suitable conditions include, for example, fermentation conditions. Fermentation conditions can comprise many parameters, including but not limited to temperature ranges, levels of aeration, feed rates and media composition. Each of these conditions, individually and in combination, allows the host cell to grow. Fermentation can be aerobic, anaerobic, or variations thereof (such as micro-aerobic). Exemplary culture media include broths or gels. Generally, the medium includes a carbon source that can be metabolized by a host cell directly. In addition, enzymes can be used in the medium to facilitate the mobilization (e.g., the depolymerization of starch or cellulose to fermentable sugars) and subsequent metabolism of the carbon source. For small scale production, the engineered host cells can be grown in batches of, for example, about 100 mL, 500 mL, 1 L, 2 L, 5 L, or 10 L; fermented; and induced to express a desired polynucleotide sequence, such as a polynucleotide sequence encoding a CAR polypeptide. For large scale production, the engineered host cells can be grown in batches of about 10 L, 100 L, 1000 L, 10,000 L, 100,000 L, and 1,000,000 L or larger; fermented; and induced to express a desired polynucleotide sequence. Alternatively, large scale fed-batch fermentation may be carried out. The fatty acid derivative compositions described herein are found in the extracellular environment of the recombinant host cell culture and can be readily isolated from the culture medium. A fatty acid derivative may be secreted by the recombinant host cell, transported into the extracellular environment or passively transferred into the extracellular environment of the recombinant host cell culture. The fatty acid derivative is isolated from a recombinant host cell culture using routine methods known in the art.

[00127]   **Products Derived From Recombinant Host Cells**

[00128]   As used herein, "fraction of modem carbon" or fM has the same meaning as defined by National Institute of Standards and Technology (NIST) Standard Reference Materials

(SRMs4990B and 4990C, known as oxalic acids standards HOxI and HOxII, respectively. The fundamental definition relates to 0.95 times the 14C /12C isotope ratio HOxI (referenced to AD 1950). This is roughly equivalent to decay-corrected pre-Industrial Revolution wood. For the current living biosphere (plant material), fM is approximately 1.1. Bioproducts (e.g., the fatty acid derivatives produced in accordance with the present disclosure) comprising biologically produced organic compounds, and in particular, the fatty acid derivatives produced using the fatty acid biosynthetic pathway herein, have not been produced from renewable sources and, as such, are new compositions of matter. These new bioproducts can be distinguished from organic compounds derived from petrochemical carbon on the basis of dual carbon-isotopic fingerprinting or ^{14}C dating. Additionally, the specific source of biosourced carbon (e.g., glucose vs. glycerol) can be determined by dual carbon-isotopic fingerprinting (see, e.g., U.S. Patent No. 7,169,588, which is herein incorporated by reference). The ability to distinguish bioproducts from petroleum based organic compounds is beneficial in tracking these materials in commerce. For example, organic compounds or chemicals comprising both biologically based and petroleum based carbon isotope profiles may be distinguished from organic compounds and chemicals made only of petroleum based materials. Hence, the bioproducts herein can be followed or tracked in commerce on the basis of their unique carbon isotope profile. Bioproducts can be distinguished from petroleum based organic compounds by comparing the stable carbon isotope ratio (^{13}C/^{12}C) in each sample. The ^{13}C/^{12}C ratio in a given bioproduct is a consequence of the ^{13}C/^{12}C ratio in atmospheric carbon dioxide at the time the carbon dioxide is fixed. It also reflects the precise metabolic pathway. Regional variations also occur. Petroleum, C3 plants (the broadleaf), C4 plants (the grasses), and marine carbonates all show significant differences in ^{13}C/^{12}C and the corresponding δ^{13}C values. Furthermore, lipid matter of C3 and C4 plants analyze differently than materials derived from the carbohydrate components of the same plants as a consequence of the metabolic pathway. Within the precision of measurement, ^{13}C shows large variations due to isotopic fractionation effects, the most significant of which for bioproducts is the photosynthetic mechanism. The major cause of differences in the carbon isotope ratio in plants is closely associated with differences in the pathway of photosynthetic carbon metabolism in the plants, particularly the reaction occurring during the primary carboxylation (i.e., the initial fixation of atmospheric CO_2). Two large classes of vegetation are those that incorporate the "C3" (or Calvin-Benson) photosynthetic cycle and those that incorporate the "C4" (or Hatch-Slack) photosynthetic cycle. In C3 plants, the primary CO_2

54

fixation or carboxylation reaction involves the enzyme ribulose-1,5-diphosphate carboxylase, and the first stable product is a 3-carbon compound. C3 plants, such as hardwoods and conifers, are dominant in the temperate climate zones. In C4 plants, an additional carboxylation reaction involving another enzyme, phosphoenol-pyruvate carboxylase, is the primary carboxylation reaction. The first stable carbon compound is a 4-carbon acid that is subsequently decarboxylated. The CO_2 thus released is refixed by the C3 cycle. Examples of C4 plants are tropical grasses, corn, and sugar cane. Both C4 and C3 plants exhibit a range of $^{13}C/^{12}C$ isotopic ratios, but typical values are about -7 to about -13 per mil for C4 plants and about -19 to about -27 per mil for C3 plants (see, e.g., Stuiver et al., Radiocarbon 19:355 (1977)). Coal and petroleum fall generally in this latter range. The 13C measurement scale was originally defined by a zero set by Pee Dee Belemnite (PDB) limestone, where values are given in parts per thousand deviations from this material. The "δ13C" values are expressed in parts per thousand (per mil), abbreviated, ‰, and are calculated as follows:

[00129] $\delta^1 3C\ (‰) = [(^{13}C/^{12}C)\ \text{sample-}\ (^{13}C/^{12}C)\ \text{standard}]/\ (^{13}C/^{12}C)\ \text{standard} \times 1000$

[00130]   Since the PDB reference material (RM) has been exhausted, a series of alternative RMs have been developed in cooperation with the IAEA, USGS, NIST, and other selected international isotope laboratories. Notations for the per mil deviations from PDB is $\delta^{13}C$. Measurements are made on CO_2 by high precision stable ratio mass spectrometry (IRMS) on molecular ions of masses 44, 45, and 46. The compositions described herein include bioproducts produced by any of the methods described herein, including, for example, fatty aldehyde and alcohol products. Specifically, the bioproduct can have a $\delta^{13}C$ of about -28 or greater, about -27 or greater, -20 or greater, -18 or greater, -15 or greater, -13 or greater, -10 or greater, or -8 or greater. For example, the bioproduct can have a $\delta^{13}C$ of about -30 to about -15, about -27 to about -19, about -25 to about -21, about -15 to about -5, about -13 to about -7, or about -13 to about -10. In other instances, the bioproduct can have a $\delta^{13}C$ of about -10, -11, -12, or -12.3. Bioproducts produced in accordance with the disclosure herein, can also be distinguished from petroleum based organic compounds by comparing the amount of ^{14}C in each compound. Because ^{14}C has a nuclear half-life of 5730 years, petroleum based fuels containing "older" carbon can be distinguished from bioproducts which contain "newer" carbon (see, e.g., Currie, "Source Apportionment of Atmospheric Particles", Characterization of Environmental Particles, J. Buffle and H. P. van Leeuwen, Eds., 1 of Vol. I of the IUPAC Environmental Analytical

Chemistry Series (Lewis Publishers, Inc.) 3-74, (1992)). The basic assumption in radiocarbon dating is that the constancy of ^{14}C concentration in the atmosphere leads to the constancy of ^{14}C in living organisms. However, because of atmospheric nuclear testing since 1950 and the burning of fossil fuel since 1850, ^{14}C has acquired a second, geochemical time characteristic. Its concentration in atmospheric CO_2, and hence in the living biosphere, approximately doubled at the peak of nuclear testing, in the mid-1960s. It has since been gradually returning to the steady-state cosmogenic (atmospheric) baseline isotope rate (^{14}C /^{12}C) of about 1.2 x 10-12, with an approximate relaxation "half-life" of 7-10 years. (This latter half-life must not be taken literally; rather, one must use the detailed atmospheric nuclear input/decay function to trace the variation of atmospheric and biospheric ^{14}C since the onset of the nuclear age.) It is this latter biospheric ^{14}C time characteristic that holds out the promise of annual dating of recent biospheric carbon. ^{14}C can be measured by accelerator mass spectrometry (AMS), with results given in units of "fraction of modern carbon" (fM). fM is defined by National Institute of Standards and Technology (NIST) Standard Reference Materials (SRMs) 4990B and 4990C. As used herein, "fraction of modern carbon" or "fM" has the same meaning as defined by National Institute of Standards and Technology (NIST) Standard Reference Materials (SRMs) 4990B and 4990C, known as oxalic acids standards HOxI and HOxII, respectively. The fundamental definition relates to 0.95 times the ^{14}C /^{12}C isotope ratio HOxI (referenced to AD 1950). This is roughly equivalent to decay-corrected pre-Industrial Revolution wood. For the current living biosphere (plant material), fM is approximately 1.1. The compositions described herein include bioproducts that can have an fM ^{14}C of at least about 1. For example, the bioproduct of the disclosure can have an fM ^{14}C of at least about 1.01, an fM ^{14}C of about 1 to about 1.5, an fM ^{14}C of about 1.04 to about 1.18, or an fM ^{14}C of about 1.111 to about 1.124.

[00131]   Another measurement of ^{14}C is known as the percent of modern carbon (pMC). For an archaeologist or geologist using ^{14}C dates, AD 1950 equals "zero years old". This also represents 100 pMC. "Bomb carbon" in the atmosphere reached almost twice the normal level in 1963 at the peak of thermo-nuclear weapons. Its distribution within the atmosphere has been approximated since its appearance, showing values that are greater than 100 pMC for plants and animals living since AD 1950. It has gradually decreased over time with today's value being near 107.5 pMC. This means that a fresh biomass material, such as corn, would give a ^{14}C signature near 107.5 pMC. Petroleum based compounds will have a pMC value of zero.

Combining fossil carbon with present day carbon will result in a dilution of the present day pMC content. By presuming 107.5 pMC represents the ^{14}C content of present day biomass materials and 0 pMC represents the ^{14}C content of petroleum based products, the measured pMC value for that material will reflect the proportions of the two component types. For example, a material derived 100% from present day soybeans would give a radiocarbon signature near 107.5 pMC. If that material was diluted 50% with petroleum based products, it would give a radiocarbon signature of approximately 54 pMC. A biologically based carbon content is derived by assigning "100%" equal to 107.5 pMC and "0%" equal to 0 pMC. For example, a sample measuring 99 pMC will give an equivalent biologically based carbon content of 93%. This value is referred to as the mean biologically based carbon result and assumes all the components within the analyzed material originated either from present day biological material or petroleum based material. A bioproduct comprising one or more fatty acid derivatives as described herein can have a pMC of at least about 50, 60, 70, 75, 80, 85, 90, 95, 96, 97, 98, 99, or 100. In other instances, a fatty acid derivative described herein can have a pMC of between about 50 and about 100; about 60 and about 100; about 70 and about 100; about 80 and about 100; about 85 and about 100; about 87 and about 98; or about 90 and about 95. In yet other instances, a fatty acid derivative described herein can have a pMC of about 90, 91, 92, 93, 94, or 94.2.

[00132]   **Screening Fatty Acid Derivative Compositions Produced by Recombinant Host Cells**

[00133]   To determine if conditions are sufficient to allow expression, a host cell can be cultured, for example, for about 4, 8, 12, 24, 36, or 48 hours. During and/or after culturing, samples can be obtained and analyzed to determine if the conditions allow expression. For example, the host cells in the sample or the 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 product, assays, such as, but not limited to, TLC, HPLC, GC/FID, GC/MS, LC/MS, MS, can be used. Recombinant host cell cultures are screened at the 96 well plate level, 1 liter and 5 liter tank level and in a 1000L pilot plant using a GC/FID assay for "total fatty species".

[00134]   **Utility of Fatty Acid Derivative Compositions**

[00135]   A fatty acid is a carboxylic acid with a long aliphatic tail (chain), which is either saturated or unsaturated. Most naturally occurring fatty acids have a chain of an even number of

carbon atoms, from 4 to 28. Fatty acids are usually derived from triglycerides. When they are not attached to other molecules, they are known as "free" fatty acids. Fatty acids are usually produced industrially by the hydrolysis of triglycerides, with the removal of glycerol. Palm, soybean, rapeseed, coconut oil and sunflower oil are currently the most common sources of fatty acids. The majority of fatty acids derived from such sources are used in human food products. Coconut oil and palm kernel oil (consist mainly of 12 and 14 carbon fatty acids). These are particularly suitable for further processing to surfactants for washing and cleansing agents as well as cosmetics. Palm, soybean, rapeseed, and sunflower oil, as well as animal fats such as tallow, contain mainly long-chain fatty acids (e.g., C18, saturated and unsaturated) which are used as raw materials for polymer applications and lubricants. Ecological and toxicological studies suggest that fatty acid-derived products based on renewable resources have more favorable properties than petrochemical-based substances. Fatty aldehydes are used to produce many specialty chemicals. For example, aldehydes are used to produce polymers, resins (e.g., Bakelite), dyes, flavorings, plasticizers, perfumes, pharmaceuticals, and other chemicals, some of which may be used as solvents, preservatives, or disinfectants. In addition, certain natural and synthetic compounds, such as vitamins and hormones, are aldehydes, and many sugars contain aldehyde groups. Fatty aldehydes can be converted to fatty alcohols by chemical or enzymatic reduction. Fatty alcohols have many commercial uses. Worldwide annual sales of fatty alcohols and their derivatives are in excess of U.S. $1 billion. The shorter chain fatty alcohols are used in the cosmetic and food industries as emulsifiers, emollients, and thickeners. Due to their amphiphilic nature, fatty alcohols behave as nonionic surfactants, which are useful in personal care and household products, such as, for example, detergents. In addition, fatty alcohols are used in waxes, gums, resins, pharmaceutical salves and lotions, lubricating oil additives, textile antistatic and finishing agents, plasticizers, cosmetics, industrial solvents, and solvents for fats. The disclosure also provides a surfactant composition or a detergent composition comprising a fatty alcohol produced by any of the methods described herein. One of ordinary skill in the art will appreciate that, depending upon the intended purpose of the surfactant or detergent composition, different fatty alcohols can be produced and used. For example, when the fatty alcohols described herein are used as a feedstock for surfactant or detergent production, one of ordinary skill in the art will appreciate that the characteristics of the fatty alcohol feedstock will affect the characteristics of the surfactant or detergent composition produced. Hence, the characteristics of the surfactant or detergent composition can be selected for by producing

particular fatty alcohols for use as a feedstock. A fatty alcohol-based surfactant and/or detergent composition described herein can be mixed with other surfactants and/or detergents well known in the art. In some embodiments, the mixture can include at least about 10%, at least about 15%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, or a range bounded by any two of the foregoing values, by weight of the fatty alcohol. In other examples, a surfactant or detergent composition can be made that includes at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or a range bounded by any two of the foregoing values, by weight of a fatty alcohol that includes a carbon chain that is 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, or 22 carbons in length. Such surfactant or detergent compositions also can include at least one additive, such as a microemulsion or a surfactant or detergent from non-microbial sources such as plant oils or petroleum, which can be present in the amount of at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or a range bounded by any two of the foregoing values, by weight of the fatty alcohol. Esters have many commercial uses. For example, biodiesel, an alternative fuel, is comprised of esters (e.g., fatty acid methyl esters, fatty acid ethyl esters, etc.). Some low molecular weight esters are volatile with a pleasant odor, which makes them useful as fragrances or flavoring agents. In addition, esters are used as solvents for lacquers, paints, and varnishes. Furthermore, some naturally occurring substances, such as waxes, fats, and oils are comprised of esters. Esters are also used as softening agents in resins and plasticizers, flame retardants, and additives in gasoline and oil. In addition, esters can be used in the manufacture of polymers, films, textiles, dyes, and pharmaceuticals. Hydrocarbons have many commercial uses. For example, shorter chain alkanes are used as fuels. Longer chain alkanes (e.g., from five to sixteen carbons) are used as transportation fuels (e.g., gasoline, diesel, or aviation fuel). Alkanes having more than sixteen carbon atoms are important components of fuel oils and lubricating oils. Even longer alkanes, which are solid at room temperature, can be used, for example, as a paraffin wax. In addition, longer chain alkanes can be cracked to produce commercially valuable shorter chain hydrocarbons. Like short chain alkanes, short chain alkenes are used in transportation fuels. Longer chain alkenes are used in plastics, lubricants, and synthetic lubricants. In addition, alkenes are used as a feedstock to

produce alcohols, esters, plasticizers, surfactants, tertiary amines, enhanced oil recovery agents, fatty acids, thiols, alkenylsuccinic anhydrides, epoxides, chlorinated alkanes, chlorinated alkenes, waxes, fuel additives, and drag flow reducers. Ketones are used commercially as solvents. For example, acetone is frequently used as a solvent, but it is also a raw material for making polymers. Ketones are also used in lacquers, paints, explosives, perfumes, and textile processing. In addition, ketones are used to produce alcohols, alkenes, alkanes, imines, and enamines. Lubricants are typically composed of olefins, particularly polyolefins and alpha-olefins. Lubricants can either be refined from crude petroleum or manufactured using raw materials refined from crude petroleum. Obtaining these specialty chemicals from crude petroleum requires a significant financial investment as well as a great deal of energy. It is also an inefficient process because frequently the long chain hydrocarbons in crude petroleum are cracked to produce smaller monomers. These monomers are then used as the raw material to manufacture the more complex specialty chemicals. The disclosure is further illustrated by the following examples. The examples are provided for illustrative purposes only. They are not to be construed as limiting the scope or content of the disclosure in any way.

EXAMPLES

[00136]   **EXAMPLE 1**

[00137]   **Production Host Modifications – Attenuation of Acyl-CoA Dehydrogenase**

[00138]   This example describes the construction of a genetically engineered host cell wherein the expression of a fatty acid degradation enzyme is attenuated.

[00139]   The fadE gene of *Escherichia coli* MG1655 (an *E. coli* K strain) was deleted using the Lambda Red (also known as the Red-Driven Integration) system described by Datsenko et al., Proc. Natl. Acad. Sci. USA 97: 6640-6645 (2000), with the following modifications:

[00140]   The following two primers were used to create the deletion of fadE:

Del-fadE-F

5'-AAAAACAGCAACAATGTGAGCTTTGTTGTAATTATATTGTAAACATATT GATTCCGGGGATCCGTCGACC (SEQ ID NO: 9); and

Del-fadE-R

5'-AAACGGAGCCTTTCGGCTCCGTTATTCATTTACGCGGCTTCAACTTTCC TGTAGGCTGGAGCTGCTTC (SEQ ID NO: 10)

[00141]   The Del-fadE-F and Del-fadE-R primers were used to amplify the kanamycin resistance (KmR) cassette from plasmid pKD13 (described by Datsenko et al., supra) by PCR. The PCR product was then used to transform electrocompetent *E. coli* MG1655 cells containing pKD46 (described in Datsenko et al., supra) that had been previously induced with arabinose for 3-4 hours. Following a 3-hour outgrowth in a super optimal broth with catabolite repression (SOC) medium at 37°C, the cells were plated on Luria agar plates containing 50 μg/mL of Kanamycin. Resistant colonies were identified and isolated after an overnight incubation at 37°C. Disruption of the fadE gene was confirmed by PCR amplification using primers fadE-L2 and fadE-R1, which were designed to flank the *E. coli* fadE gene.

[00142] The fadE deletion confirmation primers were:

fadE-L2 5'-CGGGCAGGTGCTATGACCAGGAC (SEQ ID NO: 11); and

fadE-R1 5'-CGCGGCGTTGACCGGCAGCCTGG (SEQ ID NO: 12)

[00143]   After the fadE deletion was confirmed, a single colony was used to remove the KmR marker using the pCP20 plasmid as described by Datsenko et al., supra. The resulting MG1655 *E. coli* strain with the fadE gene deleted and the KmR marker removed was named *E. coli* MG1655 ΔfadE, or *E. coli* MG 1655 D1. Fatty acid derivative (total fatty species) production by the MG1655 *E. coli* strain with the fadE gene deleted was compared to fatty acid derivative production by *E. coli* MG1655. The deletion of the fadE gene did not affect fatty acid derivative production (Figure 7). A number of exemplary host cell strains are described herein, examples of which are described below in Table 3.

[00144]   **Table 3**: Genetic Characterization of *E. coli* Strains

Strain	Genetic Characterization
DV2	MG1655 F-, λ-, ilvG-, rfb-50, rph-1, ΔfhuA::FRT, ΔfadE::FRT
DV2.1	DV2 fabB::fabB[A329V]
D178	DV2.1 entD::FRT_P_{T5}_entD
EG149	D178 ΔinsH-11::P_{LACUV5}-iFAB138
V642	EG149 rph+
SL313	V642 lacIZ::P_{A1}_'tesA/pDG109
V668	V642 ilvG$^+$
LC397	V668 lacIZ::P_{TRC}_'tesA(var)_kan
SL571	V668 lacIZ:: P_{TRC}_'tesA(var)_FRT
LC942	SL571 attTn7::P_{TRC}_'tesA(var)

DG16	LC942/pLC56
V940	LC397/pV171.1
D851	SL571 yijP::Tn5-cat/pV171.1
BD64	DV2 ΔinsH-11::P$_{LACUV5}$-iFAB138 loxP_P$_{T5}$_fadR
DAM1	DV2 attTn7::P$_{TRC}$_tesA_fadD
Shu.002	DV2 ΔinsH-11::P$_{T5}$-iFAB138 loxP_P$_{T5}$_fadR

Plasmids: pDG109, pLC56 and pV171.1 both are pCL_P$_{trc}$_carB_tesA_alrA_fabB_fadR operon with variable expression of carB and tesA. iFAB138 is SEQ ID NO: 19.

[00145]   **EXAMPLE 2**

[00146]   **Increased Flux Through The Fatty Acid Synthesis Pathway – Acetyl CoA Carboxylase Mediated**

[00147]   **Fatty Ester Production:**

[00148]   The main precursors for fatty acid biosynthesis are malonyl-CoA and acetyl-CoA (Figure 1). It has been suggested that these precursors limit the rate of fatty acid biosynthesis in *E. coli*. In this example, synthetic acc operons [*Corynebacterium glutamicum* accABCD (±birA)] were overexpressed and the genetic modifications led to increased acetyl-coA and malonyl-CoA production in *E. coli*. In one approach, in order to increase malonyl-CoA levels, an acetyl-CoA carboxylase enzyme complex from *Corynebacterium glutamicum* (*C. glutamicum*) was overexpressed in *E. coli*. Acetyl-CoA carboxylase (acc) consists of four discrete subunits, accA, accB, accC and accD (Figure 3). The advantage of *C. glutamicum* acc is that two subunits are expressed as fusion proteins, accCB and accDA, respectively, which facilitates its balanced expression. Additionally, *C. glutamicum* birA, which biotinylates the accB subunit (Figure 3) was overexpressed. Exemplary *C. glutamicum* birA DNA sequences are presented as SEQ ID NO: 55 and SEQ ID NO: 56. A *C. glutamicum* birA protein sequence is presented as SEQ ID NO: 57.

[00149]   The synthetic operons of the *C. glutamicum* acc genes were cloned in the following way in OP80 (see WO2008/119082 as incorporated-by-reference herein) *P*trc1-accDACB, *P*trc3-accDACB, *P*trc1-accCBDA and *P*trc3-CBDA. *P*trc1 and *P*trc3 are derivatives of the commonly used *P*trc promoter, which allow attenuated transcription of target genes. Note that the native sequences were amplified from the chromosomal DNA as they showed favorable codon usage (only the codon for Arg6 in accCB was changed). The *C. glutamicum* birA gene was codon optimized and obtained by gene synthesis. It was cloned downstream of the acc genes in all four operon constructs. Below we refer to the operon configuration accDACB as accD- and the

operon configuration accDACB+birA as accD+. The resulting plasmids were transformed into
E. coli DAM1_i377, which contains integrated copies (i) of leaderless thioesterease *'tesA* and
acyl-CoA synthetase *fadD* from *E.coli* and Ester synthase 9 (ES9) from *Marinobacter*
hydrocarbonoclasticus (SEQ ID NO: 6). All genes are controlled by *P*trc promoters. The strains
were grown in 5NBT media (described below) in shake flasks and were analyzed for malonyl-
CoA using short chain-CoA assay described below. Figure 8 shows that six of the eight *C.*
glutamicum acc±birA constructs showed elevated levels of malonyl-CoA in logarithmic phase
demonstrating their functionality in *E. coli*. It was noted that coexpression of birA further
increased malonyl-CoA levels in the ptrc1/3_accDACB strains, in particular with the plasmid
containing the Ptrc3-accDACB-birA operon configuration (plasmid pAS119.50D; SEQ ID NO:
62).

[00150]   In order to test the effect of combining panK and acc-birA overexpression, the
optimized panK gene was cloned downstream of birA in ptrc1/3_accDACB-birA. Pantothenate
kinase panK (or CoaA) catalyzes the first step in the biosynthesis of coenzyme A, an essential
cofactor that is involved in many reactions, *e.g.,* the formation of acetyl-CoA, the substrate for
acetyl-CoA carboxylase. The resulting plasmids were transformed into DAM1_i377, grown in
5NBT (+TVS1) media in shake flasks, and the strains were analyzed for short-chain-CoAs using
the method described below. As shown in Figure 9, in log phase panK coexpression further
increased malonyl-CoA levels and also increased acetyl-CoA levels demonstrating that panK can
further increase the malonyl-CoA levels. The impact of coexpressing an acetyl-CoA carboxylase
enzyme complex on fatty ester production was evaluated by expressing ester synthase 9 (SEQ ID
NO: 6) with and without acc genes in another *E. coli* production host. More specifically,
plasmids OP80 (vector control), pDS57 (with ES9), pDS57-accD- (with ES9 and accDACB) or
pDS57-accD+ (with ES9 and accDACB-birA; SEQ ID NO: 63) were transformed into *E. coli*
strain DV2 and the corresponding transformants were selected on LB plates supplemented with
100 mg/L of spectinomycin.

[00151]   Two transformants of each plasmid were independently inoculated into LB medium
supplemented with 100 mg/L of spectinomycin and grown for 5-8 hours at 32°C. The cultures
were diluted 30-fold into a minimal medium with the following composition: 0.5 g/L NaCl, 1
mM MgSO$_4$ x 7 H$_2$O, 0.1 mM CaCl$_2$, 2 g/L NH$_4$Cl, 3 g/L KH$_2$PO$_4$, 6 g/L Na$_2$HPO$_4$, 1 mg/L
thiamine, 1x trace metal solution, 10 mg/L ferric citrate, 100 mM Bis-Tris (pH7.0), 30 g/L
glucose and 100 mg/L spectinomycin. After over-night growth at 32°C, the cultures were diluted

10-fold in quadruplicate into minimal medium of the same composition except that the media contained 1 g/L instead of 2 g/L NH4Cl and was supplemented with 1 mM IPTG and 2% (v/v) methanol. The resulting cultures were then grown at 32°C in a shaker. The production of fatty acid methyl esters (FAMEs) was analyzed by gas chromatography with flame ionization detector (GC-FID). The samples were extracted with butyl acetate in a ratio of 1:1 vol/vol. After vortexing, the samples were centrifuged, and the organic phase was analyzed by gas chromatography (GC). The analysis conditions were as follows: instrument: Trace GC Ultra, Thermo Electron Corporation with Flame ionization detector (FID) detector; column: DB-1 (1% diphenyl siloxane; 99% dimethyl siloxane) CO1 UFM 1/0.1/5 01 DET from Thermo Electron Corporation, phase pH 5, FT: 0.4 μm, length 5m, id: 0.1mm; inlet conditions: 250°C splitless, 3.8 m 1/25 split method used depending upon sample concentration with split flow of 75 mL/m; carrier gas, flow rate: Helium, 3.0mL/m; block temperature: 330°C; oven temperature: 0.5 m hold at 50°C, 100°C/m to 330°C, 0.5 m hold at 330°C; detector temperature: 300°C; injection volume: 2μL; run time/flow rate: 6.3 m/3.0 mL/m (splitless method), 3.8 m/1.5 mL/m (split 1/25 method), 3.04 m/1.2mL/m (split 1/50 method). FAMEs produced are shown in Figure 10. The expression of ES9 by itself in *E. coli* DV2 led to FAME production above the control DV2 OP80. Coexpression of the *C. glutamicum* acetyl-CoA carboxylase complex led to an approx. 1.5-fold increase in FAMEs and the additional expression of the *C. glutamicum* biotin protein ligase led to an approx. 5-fold increase in FAMEs. These results suggest that the increased supply of malonyl-CoA improves the ability of ES9 to convert intermediates of the fatty acid biosynthetic machinery to fatty acid methyl esters in *E. coli*.

[00152]   *Short-chain-CoA assay:* 15 ml falcon tubes were prepared with 0.467 ml 10% TCA with crotonyl-CoA as internal standard and overlayed with 2 ml of silicone oil. The tubes were chilled on ice and fermentation broth equivalent to 1 ml OD600 = 31.2 was carefully layered on top of the silicone oil. The samples were centrifuged at 11,400 g at 4°C for four 4 min cycles. For each sample, a 400 ml aliquots of the TCA/cellular extract was removed and placed in a fresh Eppendorf tube for neutralization with 1 ml Octylamine (in CHCl3). After vortexing, the samples were centrifuged for 30 sec at 13,000 g. 200 ml of the top layer was filtered using a 0.2 um PTFE syringe filter and then subjected to LC-MS/MS analysis.

[00153]   *Description of media used in experiments:*

Media ID						
4N-BT	5N-BT	FA2	FA2.1	FA2.3	Concentration	Ingredient
0.5	0.5	0.5	0.5	0.5	g/L	NaCl
2	2	2	2	2	g/L	NH_4Cl
3	3	3	3	3	g/L	KH_2PO_4
6	6	6	6	6	g/L	Na_2PO_4
1	1	1	1	1	mM	$MgSO_4$
0.1	0.1	0.1	0.1	0.1	mM	$CaCl_2$
1	1	1	1	1	mg/L	thiamine
0.2	0.2	0.1	0.1	0.1	M	Bis-Tris pH7
0.1	0.1	0.05	0.1	0.1	%	Triton X-100
1	1	1	1	1	x	Trace Minerals
27	27	10	10	10	mg/L	$FeCl_2.6H_2O$
40	50	30	30	35	g/L	glucose

1000 fold concentrated Trace Vitamins Solution

0.06 g/L Riboflavin
6 g/L Niacin
5.4 g/L Pantothenic Acid
1.4 g/L Pyridoxine
0.06 g/L Biotin
0.01 g/L Folic Acid

1000 fold concentrated Trace Metal Solution

2 mL/L Concentrated hydrochloric acid
0.5 g/L boric acid
1.9 g/L cupric sulfate, pentahydrate, USP
1 g/L zinc chloride anhydrous
2 g/L sodium molybdenate dehydrate
2 g/L calcium chloride dehydrate

[00154]   Fatty Alcohol Production:

[00155]   The impact of coexpressing an acetyl-CoA carboxylase enzyme complex on Fatty alcohol production was evaluated by expressing the Acyl-ACP reductase (AAR) from *Synechococcus elongatus* (SEQ ID NO: 38) with and without acc genes in *E. coli* DV2. The accD+ operon configuration was selected as it gave the best results when coexpressed with ester synthase (see previous example). The accDABC-birA operon was cloned downstream from the *aar* gene in pLS9-185 (a pCL1920 derivative) using Infusion technology (Clontech Laboratories, Inc., Mountain View, CA). The resulting plasmid was transformed into *E. coli* DV2 and the

corresponding transformants were selected on LB plates supplemented with 100 mg/L of spectinomycin. Fatty alcohols produced are shown in Figure 11. The coexpression of AAR and accD+ led to a *ca.* 1.5-fold increase in fatty alcohol titers as compared to the AAR only control (pLS9-185). The data were reproducible (triplicate samples were shown). These results demonstrate that increasing malonyl-CoA levels lead to improved fatty acid production when this acyl-ACP reductase is used. In addition, Example 3 describes co-expression of acc genes together with entire fab operons.

[00156] EXAMPLE 3

[00157] Increased Flux Through The Fatty Acid Synthesis Pathway – iFABs

[00158] Fatty Acid Derivative Production:

[00159] Strategies to increase the flux through the fatty acid synthesis pathway in recombinant host cells include both overexpression of native *E. coli* fatty acid biosynthesis genes and expression of exogenous fatty acid biosynthesis genes from different organisms in *E. coli*. In this study, fatty acid biosynthesis genes from different organisms were combined in the genome of *E. coli* DV2 (Table 3) under the control of the lacUV5 promoter and integrated into the IS5-11 site. Sixteen strains containing iFABs 130 – 145 were evaluated. The detailed structure of iFABs 130 – 145 is presented in Tables 4 and 5.

[00160] Table 4: Components from Different Species used in iFABs 130-145

Abbreviation	Full Description
St_fabD	Salmonella typhimurium fabD gene
nSt_fabH	Salmonella typhimurium fabH gene with the native RBS
sSt_fabH	Salmonella typhimurium fabH gene with a synthetic RBS
Cac_fabF	Clostridium acetobutylicum (ATCC824) fabF gene
St_fabG	Salmonella typhimurium fabG gene
St_fabA	Salmonella typhimurium fabA gene
St_fabZ	Salmonella typhimurium fabZ gene
BS_fabI	Bacillus subtilis fabI gene
BS_FabL	Bacillus subtilis fabL gene
Vc_FabV	Vibrio chorlerae fabV gene
Ec_FabI	Escherichia coli fabI gene

[00161] Each "iFAB" included various fab genes in the following order: 1) an enoyl-ACP reductase (BS_fabI, BS_FabL, Vc_FabV, or Ec_FabI); 2) a b -ketoacyl-ACP synthetase III

(St_fabH); 3) a malonyl-CoA-ACP transacylase (St_fabD); 4) a b-ketoacyl-ACP reductase
(St_fabG); 5) a 3-hydroxy-acyl-ACP dehydratase (St_fabA or St_fabZ); 6) a b -ketoacyl-ACP
synthetase II (Cac_fabF). Note that St_fabA also has trans-2, cis-3-decenoyl-ACP isomerase
activity and that Cac_fabF has b -ketoacyl-ACP synthetase II and b -ketoacyl-ACP synthetase I
activities (Zhu et al., BMC Microbiology 9:119 (2009)). See Table 5, below for the specific
composition of iFABs 130 – 145.

[00162]   Table 5: Composition of iFABs 130 – 145

ifab	BS_fabI	BS_fabL	Vc_fabV	Ec_fabI	nSt_fabH	sSt_fabH	St_fabD	St_fabG	St_fabA	St_fabZ	Cac_fabF
ifab130	1	0	0	0	1	0	1	1	1	0	1
ifab131	1	0	0	0	1	0	1	1	0	1	1
ifab132	1	0	0	0	0	1	1	1	1	0	1
ifab133	1	0	0	0	0	1	1	1	0	1	1
ifab134	0	1	0	0	1	0	1	1	1	0	1
ifab135	0	1	0	0	1	0	1	1	0	1	1
ifab136	0	1	0	0	0	1	1	1	1	0	1
ifab137	0	1	0	0	0	1	1	1	0	1	1
ifab138	0	0	1	0	1	0	1	1	1	0	1
ifab139	0	0	1	0	1	0	1	1	0	1	1
ifab140	0	0	1	0	0	1	1	1	1	0	1
ifab141	0	0	1	0	0	1	1	1	0	1	1
ifab142	0	0	0	1	1	0	1	1	1	0	1
ifab143	0	0	0	1	1	0	1	1	0	1	1
ifab144	0	0	0	1	0	1	1	1	1	0	1
ifab145	0	0	0	1	0	1	1	1	0	1	1

[00163]   The plasmid pCL_P$_{trc}$_tesA was transformed into each of the strains and a
fermentation was run in FA2 media with 20 hours from induction to harvest at both 32°C and
37°C. Data for production of *Total Fatty Species* from duplicate plate screens is shown in
Figures 12A and 12B. From this screen the best construct was determined to be DV2 with
iFAB138. The sequence of iFAB138 in the genome of EG149 is presented as SEQ ID NO: 19.

[00164]   **Fatty Ester Production:**

[00165]   A full synthetic fab operon was integrated into the *E. coli* chromosome and evaluated
for increased FAME production by expression in *E. coli* DAM1 pDS57. In addition, four
synthetic acc operons from *Corynebaterium glutamicum* were coexpressed and evaluated for
improved FAME productivity. Several strains were obtained that produced FAMEs at a faster
rate and higher titers. The sixteen different iFAB operons (Table 5) were put under the control
of the lacUV5 promoter and integrated into the IS5-11 site of *E. coli* DAM1. These strains were
named DAM1 ifab130 to 145. They were transformed either with pDS57 (containing ester

synthase 377) or pDS57 co-expressing different versions of acc operons, see above) for evaluation of FAME production. Exemplary plasmids are described in Table 6.

Table 6: Plasmids containing Ester Synthase ES9 (from *Marinobacter hydrocarbonclasticus*) and Synthetic acc Operons (from *Corynebactrium glutamicum*)

Plasmid	Genes
pTB.071	pDS57-accCBDA
pTB.072	pDS57- accCBDA-birA
pTB.073	pDS57- accDACB
pTB.074	pDS57- accDACB-birA

pDS57 = pCL_ptrc- ES9

[00166]   The DAM1 ifab strains were analyzed in 96-well plates (4NBT medium), shake flasks (5NBT medium) (see above for medium description) and in fermenters at 32°C. The best results were obtained in 96-well plates and in shake flasks, where several DAM1 ifab strains with pDS57-acc-birA plasmids showed higher FAME titers. In particular, DAM1 ifab131, ifab135, ifab137, ifab138 and ifab143 with pDS57-accDACB-birA showed 20-40% improved titers indicating that in these strains a higher flux through the fatty acid pathway was achieved, which resulted in a better product formation rate (these results were reproducible in several independent experiments).

[00167]   Effect of overexpressing fabH and fabI on Fatty Acid Methyl Ester (FAME) Production:

[00168]   Strategies to increase the flux through the fatty acid synthesis pathway in recombinant host cells include both overexpression of native fatty acid biosynthesis genes and expression of heterologous fatty acid biosynthesis genes. FabH and fabI are two fatty acid biosynthetic enzymes that have been shown to be feedback inhibited (Heath and Rock, JBC 271: 1833-1836 (1996)). A study was conducted to determine if FabH and FabI might be limiting the rate of FAME production. FabH and fabI homologues (from *E. coli, B. subtilis, Acinetobacter baylyi* ADP1, *Marinobacter aquaeoli* VT8, and *Rhodococcus opacus*) were overexpressed as a synthetic operon and evaluated in *E. coli* DAM1 pDS57 (a strain observed to be a good FAME producer). In one approach, fabHfabI operons were constructed from organisms that accumulate

waxes (*A. baylyi, M. aquaeoli*) or triacylglycerides (*R. opacus*) and integrated into the chromosome of *E. coli* DAM1 pDS57. In a related approach, a synthetic acc operons from *C. glutamicum* were co-expressed (as described in Example 2, above). Eleven different fabHI operons were constructed (assembled in vitro) as summarized in Table 7. The fabHI operons were put under the control of IPTG inducible lacUV5 promoter and integrated into the IS5-11 site of E. coli DAM1. These strains were named as shown in the table below. They were transformed either with pDS57 (containing ester synthase 377) or pDS57 coexpressing different versions of acc operons for evaluation of FAME production.

[00169] Table 7: Genotype of Integrated fabHI Operons

Strain	Genotype of additional fab operon	Plasmid
stEP117	DAM1 ΔinsH::P$_{LACUV5}$ (snyRBS) EcfabH (synRBS) Bsfabl::kan	pDS57
stEP118	DAM1 ΔinsH::P$_{LACUV5}$ (snyRBS) EcfabH (synRBS) BsfabL::kan	pDS57
stEP127	DAM1 ΔinsH::P$_{LACUV5}$ (EcRBS) EcfabH (EcRBS) Bsfabl::kan	pDS57
stEP128	DAM1 ΔinsH::P$_{LACUV5}$ (EcRBS) EcfabH (EcRBS) BsfabL::kan	pDS57
stEP129	DAM1 ΔinsH::P$_{LACUV5}$ (EcRBS) ADP1fabH (EcRBS) ADP1fabl::kan	pDS57
stEP130	DAM1 ΔinsH::P$_{LACUV5}$ (snyRBS) ADP1fabH (synRBS) ADP1fabl::kan	pDS57
stEP131	DAM1 ΔinsH::P$_{LACUV5}$ (snyRBS) VT8fabH1 (synRBS) VT8fabl::kan	pDS57
stEP132	DAM1 ΔinsH::P$_{LACUV5}$ (snyRBS) VT8fabH2 (synRBS) VT8fabl::kan	pDS57
stEP133	DAM1 ΔinsH::P$_{LACUV5}$ (EcRBS) VT8fabH1 (synRBS) VT8fabl::kan	pDS57
stEP134	DAM1 ΔinsH::P$_{LACUV5}$ (EcRBS) VT8fabH2 (synRBS) VT8fabl::kan	pDS57
stEP151	DAM1 ΔinsH::P$_{LACUV5}$ (snyRBS) Rofabl (synRBS) RofabH::kan	pDS57
stEP153	DAM1 ΔinsH::P$_{LACUV5}$ (EcRBS) ADP1fabH (EcRBS) ADP1fabl::kan	pDS57-accCBDA
stEP154	DAM1 ΔinsH::P$_{LACUV5}$ (EcRBS) ADP1fabH (EcRBS) ADP1fabl::kan	pDS57-accDACB
stEP155	DAM1 ΔinsH::P$_{LACUV5}$ (EcRBS) ADP1fabH (EcRBS) ADP1fabl::kan	pDS57-accCBDA-birA
stEP156	DAM1 ΔinsH::P$_{LACUV5}$ (EcRBS) ADP1fabH (EcRBS) ADP1fabl::kan	pDS57-accDACB-birA
stEP157	DAM1 ΔinsH::P$_{LACUV5}$ (snyRBS) EcfabH (synRBS) Bsfabl::kan	pDS57-accCBDA
stEP158	DAM1 ΔinsH::P$_{LACUV5}$ (snyRBS) EcfabH (synRBS) Bsfabl::kan	pDS57-accCBDA-birA
stEP159	DAM1 ΔinsH::P$_{LACUV5}$ (EcRBS) EcfabH (synRBS) Bsfabl::kan	pDS57-accCBDA
stEP160	DAM1 ΔinsH::P$_{LACUV5}$ (EcRBS) EcfabH (synRBS) Bsfabl::kan	pDS57-accCBDA-birA
stEP161	DAM1 ΔinsH::P$_{LACUV5}$ (EcRBS) VT8fabH1 (synRBS) VT8fabl::kan	pDS57-accCBDA
stEP162	DAM1 ΔinsH::P$_{LACUV5}$ (EcRBS) VT8fabH1 (synRBS) VT8fabl::kan	pDS57-accCBDA-birA
stEP163	DAM1 ΔinsH::P$_{LACUV5}$ (EcRBS) VT8fabH2 (synRBS) VT8fabl::kan	pDS57-accCBDA
stEP164	DAM1 ΔinsH::P$_{LACUV5}$ (EcRBS) VT8fabH2 (synRBS) VT8fabl::kan	pDS57-accCBDA-birA

Bs: *Bacillus subtilis*; Ec: *Escherichia coli*; ADP1: *Acinetobacter* sp. ADP1; VT8:

Marinobacter aquaeolei VT8; Ro: *Rhodococcus opacus* B4

[00170] The DAM1 ifabHI strains were analyzed in 96-well plates (4NBT medium), shake flasks (5NBT medium) and in fermenters at 32°C. In a shake flask, a number of the ifabHI strains carrying pDS57 plasmid performed better than the control DAM1 pDS57strain, reaching 10 to 15% higher FAME titers (Figure 13). Additional increase in FAME titers was obtained when ifabHI strains were transformed with pDS57-acc-birA plasmids, in particular an increase of 50% in FAME titers was observed in strain StEP156 (DAM1 IS5-11::lacUV5(ecRBS)ADP1fabH (ecRBS)ADP1fabI pDS57-accDACB-birA) (Figure 14).

[00171] Some of the strains with ifabHI were run in fermenters, where an increase in FAME titers, specific productivity and yield was also observed (Figure 15), indicating that in these strains a higher flux through the fatty acid pathway was achieved, which resulted in a better product formation rate. In particular stEP129 (DAM1 5-11::UV5(ecRBS)ADP1fabH (ecRBS)ADP1fabI pDS57) showed higher FAME titers and yield in several independent fermentation runs. Other combinations of fabH and fabI may be used to achieve similar effects. Although FAME is exemplified here, this approach to alter fatty acid biosynthetic genes is a useful approach to increase production of any fatty acid derivative.

[00172] **Effect of inserting a strong promoter in front of operon FAB138 on Fatty Acid Methyl Ester (FAME) production:**

[00173] The lacUV5 promoter of iFAB138 was replaced by a T5 promoter (SEQ ID NO: 2) leading to higher levels of expression of iFAB138, as confirmed by mRNA analysis. The expression of iFAB138 from the T5 promoter resulted in a higher titer, yield and productivity of fatty esters. Strain shu.002 (Table 3) is isogenic to strain BD64 (Table 3) except that it contains the T5 promoter controlling expression of the iFAB138 operon (SEQ ID NO: 19).

[00174] **Table 8: Primers used to Generate iT5_138 Cassette and Verify its Insertion in New Strains**

Primer Name	SEQ ID NO	Sequence
DG405	20	TTGTCCATCTTTATATAATTTGGGGGTAGGGTGTTCTTTATGTAAAAAAAC gtttTAGGATGCATATGGCGGCC

DG406	21	GATAAATCCACGAATTTTAGGTTTGATGATCATTGGTCTCCTCCTGCAGGTG CGTGTTCGTCGTCATCGCAATTG
DG422	22	ACTCACCGCATTGGTGTAGTAAGGCGCACC
DG423	23	TGAATGTCATCACGCAGTTCCCAGTCATCC
EG744	24	CCATCTTCTTTGTACAGACGTTGACTGAACATG
EG749	24	GCACCATAGCCGTAATCCCACAGGTTATAG
oTREE047	26	TGTCATTAATGGTTAATAATGTTGA

[00175]   Primers DG405 and DG406 (Table 8) were used to amplify a cat-loxP and T5 promoter cassette adding 50bp homology to each end of the PCR product, such that it could be integrated into any strain replacing the lacUV5 promoter regulating expression of the iFAB138 operon. The cat-loxP-T5 promoter was transformed into BD64/pKD46 strain. Transformants were recovered on LB + chloramphenicol plates at 37°C overnight, patched to a fresh LB + chloramphenicol plate, and verified by colony PCR using primers DG422 and DG423. Plasmid pJW168 (Palmeros *et al., Gene* 247: 255-264 (2000)) was transformed into strain BD64 i-cat-loxP-T5_138 and selected on LB + carbenicillin plates at 32°C. In order to remove the cat marker, expression of the cre-recombinase was induced by IPTG. The plasmid pJW168 was removed by growing cultures at 42°C. Colonies were patched on LB + chloramphenicol and LB + carbenicillin to verify loss of pJW168 and removal of cat marker, respectively. The colony was also patched into LB as a positive control, all patched plates were incubated at 32°C. The removal of the cat marker was confirmed by colony PCR using primers DG422 and DG423. The resulting PCR product was verified by sequencing with primers EG744, EG749 and oTREE047, the strain was called shu.002. Figure 16 shows the iFAB138 locus: a diagram of the cat-loxP-P_{T5} cassette integrated in front of FAB138 (Figure 16A) and a diagram of the P_{T5}_iFAB138 region (Figure 16B). The sequence of the cat-loxP-T5 promoter integrated in front of iFAB138 with homology to integration site is presented as SEQ ID NO: 1 and the sequence of the iT5_FAB138 promoter region with homology to integration site is presented as SEQ ID NO: 2. There are a number of conditions that can lead to increased fatty acid flux. In this example increased fatty acid flux was achieved by altering the promoter strength of operon iFAB138. The expression of iFAB138 from the T5 promoter was beneficial, nonetheless, when this promoter change was combined with the insertion of yijP::Tn5 cassette further improvements were observed in titer, yield and productivity of fatty acid esters and other fatty acid derivatives (data not shown).

[00176]   **EXAMPLE 4**

[00177]   **Increasing the Amount of Free Fatty Acid (FFA) Product By Repairing the rph and ilvG Mutations**

[00178]   The ilvG and rph mutations were corrected in this strain resulting in higher production of FFA. Strains EG149 and V668 (Table 3) were transformed with pCL_P$_{trc}$_tesA. Fermentation was run at 32°C in FA2 media for 40 hours to compare the FFA production of strains EG149 and V668 with pCL_P$_{trc}$_tesA. Correcting the rph and ilvG mutations resulted in a 116% increase in the FFA production of the base strain with pCL_P$_{trc}$_tesA. As seen in Figure 17, V668/ pCL_P$_{trc}$_tesA produced more FFA than the EG149/ pCL_P$_{trc}$_tesA control. Since FFA is a precursor to the LS9 products, higher FFA production is a good indicator that the new strain can produce higher levels of LS9 products.

[00179]   **EXAMPLE 5: Increased Production of Fatty Acid Derivatives by Transposon Mutagenesis – yijP**

[00180]   **Fatty Alcohol Production:**

[00181]   To improve the titer, yield, productivity of fatty alcohol production by *E. coli*, transposon mutagenesis and high-throughput screening was carried out and beneficial mutations were sequenced. A transposon insertion in the yijP strain was shown to improve the strain's fatty alcohol yield in both shake flask and fed-batch fermentations. The SL313 strain produces fatty alcohols. The genotype of this strain is provided in Table 3. Transposon clones were then subjected to high-throughput screening to measure production of fatty alcohols. Briefly, colonies were picked into deep-well plates containing LB, grown overnight, inoculated into fresh LB and grown for 3 hours, inoculated into fresh FA2.1 media, grown for 16 hours, then extracted using butyl acetate. The crude extract was derivatized with BSTFA (N,O-bis[Trimethylsilyl]trifluoroacetamide) and analyzed using GC/FID. Spectinomycin (100mg/L) was included in all media to maintain selection of the pDG109 plasmid. Hits were selected by choosing clones that produced a similar total fatty species as the control strain SL313, but that had a higher percent of fatty alcohol species and a lower percent of free fatty acids than the control. Strain 68F11 was identified as a hit and was validated in a shake flask fermentation using FA2.1 media. A comparison of transposon hit 68F11 to control strain SL313 indicated that 68F11 produces a higher percentage of fatty alcohol species than the control, while both strains produce similar titers of total fatty species. A single colony of hit 68F11, named LC535, was

72

sequenced to identify the location of the transposon insertion. Briefly, genomic DNA was
purified from a 10 mL overnight LB culture using the kit ZR Fungal/Bacterial DNA MiniPrep™
(Zymo Research Corporation, Irvine, CA) according to the manufacturer's instructions. The
purified genomic DNA was sequenced outward from the transposon using primers internal to the
transposon:

DG150 5'-GCAGTTATTGGTGCCCTTAAACGCCTGGTTGCTACGCCTG-3'(SEQ ID
NO: 27)

DG131 5'- GAGCCAATATGCGAGAACACCCGAGAA-3'(SEQ ID NO: 28)

[00182]   Strain LC535 was determined to have a transposon insertion in the yijP gene
(Figure18). yijP encodes a conserved inner membrane protein whose function is unclear. The
yijP gene is in an operon and co-transcribed with the ppc gene, encoding phosphoenolpyruvate
carboxylase, and the yijO gene, encoding a predicted DNA-binding transcriptional regulator of
unknown function. Promoters internal to the transposon likely have effects on the level and
timing of transcription of yijP, ppc and yijO, and may also have effects on adjacent genes frwD,
pflC, pfld, and argE. Promoters internal to the transposon cassette are shown in Figure 18, and
may have effects on adjacent gene expression. Strain LC535 was evaluated in a fed-batch
fermentation on two different dates. Both fermentations demonstrated that LC535 produced fatty
alcohols with a higher yield than control SL313, and the improvement was 1.3-1.9% absolute
yield based on carbon input. The yijP transposon cassette was further evaluated in a different
strain V940, which produces fatty alcohol at a higher yield than strain SL313. The yijP::Tn5-cat
cassette was amplified from strain LC535 using primers:

LC277 5'-
CGCTGAACGTATTGCAGGCCGAGTTGCTGCACCGCTCCCGCCAGGCAG-3'
(SEQ ID NO: 29)

LC278 5'-
GGAATTGCCACGGTGCGGCAGGCTCCATACGCGAGGCCAGGTTATCCAACG-
3'
(SEQ ID NO: 30)

[00183]   This linear DNA was electroporated into strain SL571 and integrated into the
chromosome using the lambda red recombination system. Colonies were screened using primers
outside the transposon region:

DG407 5'-AATCACCAGCACTAAAGTGCGCGGTTCGTTACCCG-3'(SEQ ID NO:
31)
DG408 5'-ATCTGCCGTGGATTGCAGAGTCTATTCAGCTACG-3'(SEQ ID NO: 32)

[00184]   A colony with the correct yijP transposon cassette was transformed with the
production plasmid pV171.1 to produce strain D851. D851 was tested in a shake-flask
fermentation against strain V940 that does not contain the yijP transposon cassette. The result of
this fermentation showed that the yijP transposon cassette confers production of a higher percent
of fatty alcohol by the D851 strain relative to the V940 strain and produces similar titers of total
fatty species as the V940 control strain. Strain D851 was evaluated in a fed-batch fermentation
on two different dates. Data from these fermentations is shown in Table 9 which illustrates that
in 5-liter fed-batch fermentations, strains with the yijP::Tn5-cat transposon insertion had an
increased total fatty species ("FAS") yield and an increase in percent fatty alcohol ("FALC ").
The terms "total fatty species" and "total fatty acid product" may be used interchangeably
herein with reference to the amount of fatty alcohols, fatty aldehydes and free fatty acids, as
evaluated by GC-FID as described in International Patent Application Publication WO
2008/119082. The same terms may be used to mean fatty esters and free fatty acids when
referring to a fatty ester analysis. As used herein, the term "fatty esters" includes beta
hydroxy esters.

[00185]   Table 9: Effect of yijP transposon insertion on titer and yield of FAS and FALC

Strain	FAS Titer	FAS Yield	Percent FALC	FALC Yield
V940	68 g/L	18.7%	95.0%	17.8%
D851	70 g/L	19.4%	96.1%	18.6%
V940	64 g/L	18.4%	91.9%	16.9%
D851	67 g/L	19.0%	94.0%	17.8%

[00186]   **Tank Fermentation Method:**

[00187]   To assess production of fatty acid and fatty acid derivatives in tank a glycerol vial of
desired strain was used to inoculate 20mL LB + spectinomycin in shake flask and incubated at

32°C for approximately six hours. 4mL of LB culture was used to inoculate 125mL Low PFA Seed Media (below), which was then incubated at 32°C shaker overnight. 50mL of the overnight culture was used to inoculate 1L of Tank Media. Tanks were run at pH 7.2 and 30.5°C under pH stat conditions with a maximum feed rate of 16g/L/hr glucose.

[00188]   Table 10: Low P FA Seed Media:

Component	Concentration
NH_4Cl	2 g/L
NaCl	0.5 g/L
KH_2PO_4	1 g/L
$MgSO_4$-$7H_2O$	0.25 g/L
$CaCl_2$-$2H_2O$	0.015 g/L
Glucose	20 g/L
TM2 Trace Minerals solution	1 mL/L
Ferric citrate	10 mg/L
Bis Tris buffer (pH 7.0)	100 mM
Spectinomycin	115 mg/L

[00189]   Table 11: Tank Media

Component	Concentration
$(NH_4)_2SO_4$	0.5 g/L
KH_2PO_4	3.0 g/L
Ferric Citrate	0.034 g/L
TM2 Trace Minerals Solution	10 mL/L
Casamino acids	5 g/L
Post sterile additions	
$MgSO_4$-$7H_2O$	2.2 g/L
Trace Vitamins Solution	1.25 mL/L
Glucose	5 g/L
Inoculum	50 mL/L

[00190]   Further studies suggest that the improved titer and yield of FAS and FALC in strains with the yijP transposon insertion is due to reduction in the activity of phosphoenolpyruvate carboxylase (ppc). A ppc enzyme assay was carried out in-vitro in the following strains to evaluate this hypothesis.

1) Δppc = DG14 (LC942 Δppc::cat-sacB/pLC56)

2) wt-ppc = DG16 (LC942/pLC56)

3) yijP::Tn5 = DG18 (LC942 yijP::Tn5-cat/pLC56)

[00191] Ppc activity was measured in cells grown in a shake flask fermentation using a standard shake flask protocol in FA2.3 media (described above) and harvested 12-16 hours after induction. Approximately 5 mL of cells were centrifuged and the cell paste was suspended in BugBuster Protein Extraction Reagent (Novagen) with a protease inhibitor cocktail solution. The cell suspension was incubated with gentle shaking on a shaker for 20 min. Insoluble cell debris was removed by centrifugation at 16,000 ×g for 20 min at 4 °C followed by transferring the supernatant to a new tube. Ppc activity in the cell lysate was determined by a coupling reaction with citrate synthase using following reaction mixture: 0.4 mM acetyl-CoA, 10 mM phosphoenolpyruvate, 0.5 mM monobromobimane, 5 mM $MgCl_2$, 10 mM $NaHCO_3$, and 10 units citrate synthase from porcine heart in 100 mM Tris-HCl (pH 8.0). The formation of CoA in the reaction with citrate synthase using oxaloacetate and acetyl-CoA was monitored photometrically using fluorescent derivatization of CoA with monobromobimane. The Ppc assay results showed that the yijP::Tn5-cat transposon cassette decreased the Ppc activity in the cell by 2.7 fold compared to wild type cells. The cells with deletion of ppc did not grow well and the activity was about 10 times lower than wild type cells. The results also indicate that the highest yield of fatty alcohol production requires a level of Ppc expression lower than the wild-type level. Proteomics data was also collected to assess the abundance of the Ppc protein in two strains with and without the yijP::Tn5-cat transposon cassette. Protein samples were collected from strains V940 and D851 grown in bioreactors under standard fatty alcohol production conditions (described above). Samples were taken at two different time points: 32 and 48hours and prepared for analysis.

[00192] Sample collection and protein isolation was carried out as follows:

[00193] 20 ml of fermentation broth were collected from each bioreactor at each time point. Samples were quenched with ice-cold PBS and harvested by centrifugation (4500 rpm/10min) at 4°C. Cell pellet was washed with ice-cold PBS and centrifuged one more time and stored at -80°C for further processing.

[00194] Total protein extraction was performed using a French press protocol. Briefly, cell pellets were resuspended in 7 ml of ice-cold PBS and French pressed at 2000 psi twice to ensure complete lysing of the bacteria. Samples were centrifuged for 20 min at 10000 rpm at 4°C to separate non-lysed cells and cell debris from the protein fraction. Total protein concentration of

clear lysate was determined using BCA Protein Assay Reagent. Samples were diluted to 2 mg

proteins/ml concentration and frozen at -80°C.

[00195]   Samples were resuspended in the appropriate buffer and trypsinized overnight at 37°C

and lyophilized. Fragmented protein samples were labeled with isotopically enriched

methylpiperazine acetic acid at room temperature for 30 min. Labeled samples were separated

using cation exchange liquid chromatography and subjected to mass spectroscopy analysis using

an ion trap mass spectrometer. Raw data was normalized using background subtraction and bias

correction.

[00196]   Proteomics data showed a significant reduction in the relative abundance of Ppc

protein in D851 strain when compared to V940 at 32 hours and 48 hours. D851 had about 15%

of the Ppc levels of V940 at 32 hours and about 35% of the Ppc levels of V940 at 48 hours.

These data show that the yijP::Tn5-cat transposon cassette results in a significant reduction in

Ppc abundance in the cell. This suggests that the observed benefits to fatty alcohol production by

strains harboring the yijP::Tn5-cat transposon hit is due to reducing the amount of Ppc protein.

[00197] These results suggest that altering ppc activity can improve the yield of fatty acid

derivatives. There are a number of ways to alter the expression of the ppc gene, and the yijP

transposon insertion is one way to accomplish this. Without wanting to be bound by theory,

if the effect of reducing phosphoenolpyruvate carboxylase activity is to limit the flow of

carbon through the TCA cycle, one could achieve similar results by decreasing the activity of

citrate synthase (gltA) or slowing the TCA cycle by decreasing the activity of any of the

enzymes involved in the TCA cycle.

[00198]   **EXAMPLE 6**

[00199]   **Increased Flux Through The Fatty Acid Synthesis Pathway –Acyl Carrier**

Protein (ACP) Mediated Fatty Alcohol Production

[00200]   When terminal pathway enzymes from sources other than *E. coli* are expressed in *E.*

coli as the heterologous host to convert fatty acyl-ACPs to products, limitations may exist in the

recognition, affinity and/or turnover of the recombinant pathway enzyme towards the *E. coli*

fatty acyl-ACPs. Note that although ACP proteins are conserved to some extent in all organisms,

their primary sequence can differ significantly. To test this hypothesis the *acp* genes from

several cyanobacteria were cloned downstream from the *Synechococcus elongatus* PCC7942

acyl-ACP reductase (AAR) present in pLS9-185, which is a pCL1920 derivative. In addition, the

sfp gene (Accession no. X63158; SEQ ID NO: 53)_from *Bacillus subtilis*, encoding a

phosphopantetheinyl transferase with broad substrate specificity, was cloned downstream of the

respective *acp* genes. This enzyme is involved in conversion of the inactive apo-ACP to the

active holo-ACP. The plasmids constructed are described in Table 12.

[00201]   **Table 12: Plasmids Coexpressing Cyanobacterial ACP with and without *B.***
***subtilis* sfp Downstream from *S. elongatus* PCC7942 AAR**

Base plasmid	ACP Source	ACP SEQ ID NO. (DNA/Polypeptide)	Without sfp	With sfp
pLS9-185	*Synechococcus elongatus* 7942	49/50	pDS168	pDS168S
pLS9-185	*Synechocystis* sp. 6803	45/46	pDS169	not available
pLS9-185	*Prochlorococcus marinus* MED4	47/48	pDS170	pDS170S
pLS9-185	*Nostoc punctiforme* 73102	43/44	pDS171	pDS171S
pLS9-185	*Nostoc sp.* 7120	51/52	pDS172	pDS172S

[00202]   All the *acp* genes were cloned with a synthetic RBS into the *Eco*RI site immediately

downstream of the *aar* gene in pLS9-185 using InFusion technology (Clontech Laboratories,

Inc., Mountain View, CA). The *Eco*RI site was reconstructed downstream of the *acp* gene.

Similarly, the *B. subtilis sfp* gene was InFusion cloned into this *Eco*RI site along with a synthetic

RBS. All plasmids were transformed into *E. coli* MG1655 DV2 (Table 3). The control for these

experiments was the expression of AAR alone (pLS9-185). The results from standard shake

flask fermentation experiments are shown in Figure 19. Significant improvement in fatty alcohol

titers were observed in strains containing the plasmids pDS171S, pDS172S, pDS168 and

pDS169 demonstrating that ACP overexpression can be beneficial for fatty alcohol production,

in this case presumably by aiding in the recognition, affinity and/or turnover of acyl-ACPs by the

heterologous terminal pathway enzyme. (See Table 12 for the source of the ACPs and presence

or absence of *sfp*).

[00203]   **Fatty Acid Production:**

[00204] In order to evaluate if the overexpression of an ACP can also increase free fatty acid

production, one cyanobacterial ACP gene with sfp was amplified from pDS171s (Table 12) and

cloned downstream from 'tesA into a pCL vector. The resulting operon was under the control of

the *P*trc3 promoter, which provides slightly lower transcription levels than the *P*trc wildtype promoter. The construct was cloned into *E. coli* DV2 and evaluated for fatty acid production. The control strain contained the identical plasmid but without cyanobacterial ACP and *B. subtilis* sfp. The results from a standard microtiter plate fermentation experiment are shown in Figure 20. Significant improvement in fatty acid titer was observed in the strain coexpressing the heterologous ACP demonstrating that ACP overexpression can be beneficial for fatty acid production, in this case presumably by increasing the flux through the fatty acid biosynthetic pathway.

We claim:

1. A recombinant host cell comprising a genetically engineered polynucleotide sequence encoding one or more polypeptides, wherein said polypeptides have activity selected from the group consisting of 3-hydroxydecanoyl-[acp] dehydratase (E.C. 4.2.1.60) activity; ß-ketoacyl-ACP synthase I (E.C. 2.3.1.41) activity; ß -ketoacyl-ACP synthase II (E.C. 2.3.1.179) activity; [acp] S-malonyltransferase {malonyl-CoA-ACP transacylase} (E.C. 2.3.1.39) activity; 3-oxoacyl-{ß –ketoacyl}-ACP reductase (E.C. 1.1.1.100) activity; ß -ketoacyl-ACP synthase III (E.C. 2.3.1.180) activity; enoyl-ACP reductase (NADH) (E.C. 1.3.1.9) activity; enoyl-ACP reductase (NADPH) (E.C. 1.3.1.10) activity; 3-hydroxy-acyl-[acp] dehydratase (E.C. 4.2.1.59) activity; and trans-2, cis-3-decenoyl-ACP isomerase (E.C. 5.3.3.14) activity, and wherein said recombinant host cell produces a fatty acid derivative composition at a higher titer, yield or productivity than a corresponding wild type host cell when cultured in a medium containing a carbon source under conditions effective to express said polynucleotide.

2. The recombinant host cell of claim 1, wherein said recombinant host cell produces said fatty acid derivative composition at a higher titer, yield or productivity when said polypeptide is expressed in combination with at least one other polypeptide of said activity.

3. The recombinant host cell of claim 1, wherein said recombinant host cell produces said fatty acid derivative composition at a higher titer, yield or productivity when said polypeptide is expressed in combination with at least five other polypeptides of said activity.

4. The recombinant host cell of claim 1, wherein said trans-2, cis-3-decenoyl-ACP isomerase activity is fabA or fabM.

5. The recombinant host cell of claim 1, wherein said ß-ketoacyl-ACP synthase II is fabF.

6. The recombinant host cell of claim 1, wherein said malonyl-CoA-ACP transacylase is fabD.

7. The recombinant host cell of claim 1, wherein said ß -ketoacyl-ACP synthase I is fabF or fabB.

8. The recombinant host cell of claim 1, wherein ß -ketoacyl-ACP reductase is fabG.

9. The recombinant host cell of claim 1, wherein ß -ketoacyl-ACP synthase III is fabH.

10. The recombinant host cell of claim 1, wherein enoyl-ACP reductase is fabI or fabL or fabV or fabK.

11. The recombinant host cell of claim 1, wherein 3-hydrox-acyl-[acp] dehydratase is fabA or fabZ.

12. The recombinant host cell of claim 1, wherein trans-2-enoyl-ACP reductase II is fabK.

13. The recombinant host cell of claim 1, wherein said polypeptide is selected from the group consisting of fabA, fabB, fabD, fabF, fabG, fabH, fabI, fabL, fabV, fabZ, fabM, and fabK.

14. The recombinant host cell of claim 1, wherein said polypeptide is selected from the group consisting of FabA from *Salmonella typhimurium* (NP_460041); FabB from *Escherichia coli* (NP_416826); FabD from *Salmonella typhimurium* (NP_460164); FabG from *Salmonella typhimurium* (NP_460165); FabH from *Salmonella typhimurium* (NP_460163); FabZ from *Salmonella typhimurium* (NP_459232); FabM from *Streptococcus mutans* (AAN59379); FabK from *Streptococcus pneumoniae* (AAF98273); FabV from *Vibrio cholera* (YP_001217283); FabF from *Clostridium acetobutylicum* (NP_350156); FabI from *Bacillus subtillis subsp. subtilis str.* 168 (NP_389054); FabL from *Bacillus subtillis subsp. subtilis str.* 168 (NP_388745); FabI from *Acinetobacter sp. ADP1* (YP_047630); FabI from *Marinobacter aquaeoli* VT8 (YP_958813); FabI from

Rhodococcus opacus B4 (YP_002784194); FabH from *Acinetobacter sp. ADP1*
(YP_046731); FabH from *Marinobacter aquaeoli* VT8 (YP_958649); and FabH from
Rhodococcus opacus B4 (YP_00278448).

15. The recombinant host cell of claim 1, wherein said polypeptide encoded by said
    genetically engineered polynucleotide sequence is exogenous.

16. The recombinant host cell of claim 1, further comprising a genetically engineered
    polynucleotide sequence encoding an acyl carrier protein (ACP).

17. The recombinant host cell of claim 16, wherein said ACP further increases said titer,
    yield or productivity.

18. A recombinant host cell comprising a genetically engineered polynucleotide sequence
    encoding an ACP polypeptide, wherein said recombinant host cells produces a fatty acid
    derivative composition at a higher titer, yield or productivity than a corresponding wild
    type host cell when cultured in a medium containing a carbon source under conditions
    effective to express said ACP polypeptide.

19. The recombinant host cell of claim 18, wherein said recombinant host cell further
    comprises a recombinant polynucleotide encoding a polypeptide that has
    phosphopantetheinyl transferase (E.C. 2.7.8.7) activity.

20. The recombinant host cell of claim 19, wherein said polynucleotide encodes a *sfp* gene.

21. The recombinant host cell of claims 1or 18, wherein said recombinant host cell further
    comprises a recombinant polynucleotide encoding a polypeptide that has accABCD
    activity (E.C. 6.4.1.2).

22. The recombinant host cell of claim 1or 18, wherein said ACP is derived from the same
    organism as a terminal pathway enzyme expressed in said recombinant host cell.

23. The recombinant host cell of claim 18, wherein said ACP is exogenous.

24. A recombinant host cell comprising a genetically engineered polynucleotide sequence comprising a transposon, wherein insertion of said transposon into a *yijP* gene affects a second gene flanking said *yijP* gene, wherein said second gene codes for a polynucleotide that is up- or down regulated, and wherein said up- or down regulated polynucleotide codes for a polypeptide that affects production of a fatty acid derivative composition when said host cell is cultured in a medium containing a carbon source under conditions effective to express said polypeptide.

25. The recombinant host cell of claim 24, wherein insertion of said transposon into said *yijP* gene results in inactivation of said yijP gene or a polynucleotide thereof, which affects said second gene flanking said *yijP* gene, wherein said second gene codes for a polypeptide that affects production of a fatty acid derivative composition when said host cell is cultured in a medium containing a carbon source under conditions effective to express said polypeptide.

26. The recombinant host cell of claim 24, wherein said second gene or polynucleotide thereof is selected from the group consisting of *ppc, yijO, frwD, pflC, pflD* or *argE.*

27. A recombinant host cell comprising a genetically engineered polynucleotide sequence encoding a phosphoenolpyruvate carboxylase (ppc) polypeptide, wherein said recombinant host cell produces a fatty acid derivative composition at a higher titer, yield or productivity than a corresponding wild type host cell when cultured in a medium containing a carbon source under conditions effective to express said ppc polypeptide.

28. A cell culture comprising the recombinant host cell according to any of claims 1 through 27.

29. The cell culture of claim 28, wherein said medium comprises a fatty acid derivative composition.

30. The cell culture of claim 29, wherein the fatty acid derivative composition comprises at least one fatty acid derivative selected from the group consisting of fatty acid, a fatty ester, a fatty alcohol, a fatty aldehyde, an alkane, a terminal olefin, an internal olefin, and a ketone.

31. The cell culture of claim 28, wherein the fatty acid derivative is a C6, C8, C10, C12, C13, C14, C15, C16, C17, or C18 fatty acid derivative.

32. The cell culture of claim 28, wherein the fatty acid derivative is a C10:1, C12:1, C14:1, C16:1, or C18:1 unsaturated fatty acid derivative.

33. The cell culture of claim 29, wherein the fatty acid derivative composition comprises one or more of C8, C10, C12, C14, C16, and C18 fatty acid derivatives.

34. The cell culture of claim 29, wherein the fatty acid derivative composition comprises fatty acids.

35. The cell culture of claim 29, wherein the fatty acid derivative composition comprises fatty aldehydes.

36. The cell culture of claim 29, wherein the fatty acid derivative composition comprises fatty alcohols.

37. The cell culture of claim 29, wherein the fatty acid derivative composition comprises fatty esters.

38. The cell culture of claim 29, wherein the fatty acid derivative composition comprises alkanes.

39. The cell culture of claim 29, wherein the fatty acid derivative composition comprises terminal olefins.

40. The cell culture of claim 29, wherein the fatty acid derivative composition comprises internal olefins.

41. The cell culture of claim 29, wherein the fatty acid derivative composition comprises ketones.

42. The cell culture of claim 29, wherein the fatty acid derivative composition comprises fatty acid derivatives having a double bond at position 7 in the carbon chain between C7 and C8 from the reduced end of the fatty alcohol.

43. The cell culture of claim 29, wherein the fatty acid derivative composition comprises unsaturated fatty acid derivatives.

44. The cell culture of claim 29, wherein the fatty acid derivative composition comprises saturated fatty acid derivatives.

45. The cell culture of claim 29, wherein the fatty acid derivative composition comprises branched chain fatty acid derivatives.

46. The cell culture of claim 28, wherein said recombinant host cell has a titer that is at least about 5% greater than a titer of said corresponding wild type host cell when cultured under the same conditions as the recombinant host cell.

47. The cell culture of claim 46, wherein said recombinant host cell has a titer of from about 1g/L to about 250g/L.

48. The cell culture of claim 47, wherein said recombinant host cell has a titer of from about 90 g/L to about 120g/L.

49. The cell culture of claim 28, wherein said recombinant host cell has a yield that is at least about 10% to about 40%.

50. The recombinant host cell of claim 49, wherein said recombinant host cell has a yield of about 25%.

51. The cell culture of claim 28, wherein said productivity ranges from about 0.7mg/L/hr to about 3g/L/hr.

FIG. 1

FIG. 2

ATP + HCO$_3^-$ ADP + Pi

accC

Biotin Carboxylase

BCCP   accB   BCCP

birA

Carboxyltransferase

accAD

Malonyl-CoA   Acetyl-CoA

FIG. 3

Acyl-ACP

Thioesterase
(TE)

Acyl-ACP
Reductase (AAR)

Fatty Acid   **Fatty Aldehyde**

Carboxylic Acid
Reductase (Car)

Aldehyde reductase or
alcohol dehydrogenase

Fatty Alcohol

FIG. 4

FIG. 5

Three Enzyme System

Acyl-ACP
↓ Thioesterase (TE)
Fatty Acid
↓ Acyl-CoA synthetase (fadD)
Acyl-CoA

Ester synthase

+/- alcohol produced by production host or added to the culture medium

→ Fatty ester
Wax ester

One Enzyme System

Acyl-ACP

Ester synthase

FIG. 6

FIG. 7

FIG. 8

DAM1_377 with *C. glutamicum* Acc and BirA expression:
Mal-CoA in log phase

FIG. 9

DAM1-i377: Short-chain CoA levels in log phase, acc-panK coexpression

FIG. 10

FIG. 11

FIG. 12A

Duplicate plate screen at 32°C

FIG. 12B

Duplicate plate screen at 37°C

Total fatty species (mg/L)

FIG. 13

FIG. 14

FIG. 15

FIG. 16A

FIG. 16B

FIG. 17

FIG. 18

FIG. 19

FIG. 20